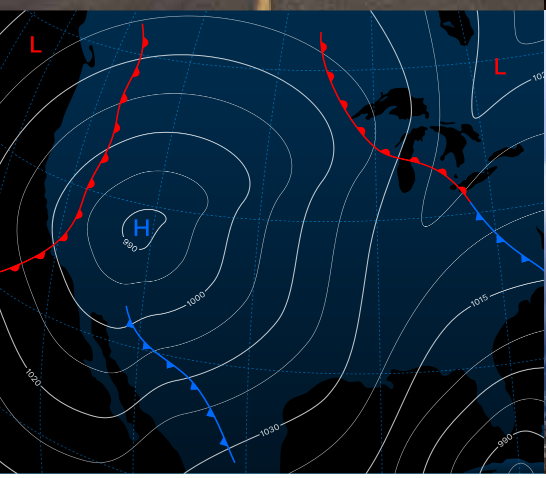
FULL REPORT

# Weather Dataset Needs for Planning and Analyzing Modern Power Systems




A Report of the Energy Systems Integration Group's Weather Datasets Project Team

**October 2023**

ESIG
ENERGY SYSTEMS INTEGRATION GROUP

## About ESIG

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation. More information is available at https://www.esig.energy.

## ESIG's Publications

This full report (and a high-resolution version for printing), a shorter summary report, a stand-alone version of Section 2 (meteorology fundamentals), and fact sheets are available at https://www.esig.energy/weather-data-for-power-system-planning. All ESIG publications can be found a https://www.esig.energy/reports-briefs.

## Get in Touch

To learn more about the topics discussed in this report or for more information about the Energy Systems Integration Group, please send an email to info@esig.energy.

Design: David Gerratt/NonprofitDesign.com
Production management and editing: Karin Matchett/tomorrowsfootprint.com

# Weather Dataset Needs for Planning and Analyzing Modern Power Systems

**A Report of the Energy Systems Integration Group's Weather Datasets Project Team**


## Project Lead

**Justin Sharp**, Sharply Focused

## Writing Team

**Justin Sharp**, Sharply Focused

**Michael Milligan**, Milligan Grid Solutions

**Hannah Bloomfield**, Newcastle University, UK

## Main Project Team Contributors

**Priya Sreedharan**, GridLab

**Julia Matevosyan**, Energy Systems Integration Group

**Jared Lee**, National Center for Atmospheric Research

**Erik Smith**, EPRI

**Andrea Staid**, EPRI

**Michael Craig**, University of Michigan

**James Wilczak**, National Oceanic and Atmospheric Administration

**Derek Stenclik**, Telos Energy

**Ana Dyreson**, Michigan Technological University

**David Brayshaw**, University of Reading

**Jeff Freedman**, University at Albany, State University of New York

**Laurent Dubus**, RTE

**Carlo Brancucci**, encoord

**John Zack**, Meso, Inc.

**David McQueen**, Australian Bureau of Meteorology



## Acknowledgments

**Grant Buster** (National Renewable Energy Laboratory) and **Irene Schicker** (GeoSphere Austria) provided important contributions in Section 2 with respect to machine learning techniques.


This report was produced by a project team made up of diverse members with diverse viewpoints and levels of participation. Specific statements may not necessarily represent a consensus among all participants.


## Suggested Citation

Energy Systems Integration Group. 2023. *Weather Dataset Needs for Planning and Analyzing Modern Power Systems* (Full Report). A Report of the Weather Datasets Project Team. Reston, VA. https://www.esig.energy/weather-data-for-power-system-planning.

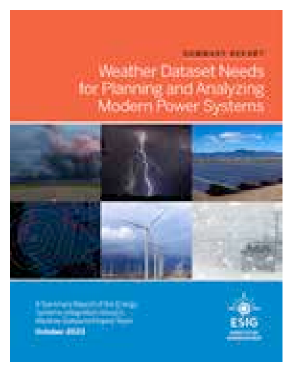
ESIG
ENERGY SYSTEMS
INTEGRATION GROUP

## Summary Versions of the Report and Meteorology Overview

This full report (and a high-resolution version for printing) is accompanied by a shorter summary report, an executive summary, and fact sheets, as well as a stand-alone version of Section 2 titled "Meteorology 101: Meteorological Data Fundamentals for Power System Planning," an overview of meteorology, data, and modeling for readers of the summary report who would like a deeper dive into those areas.

*Weather Dataset Needs for Planning and Analyzing Modern Power Systems* was produced by a project team convened by the Energy Systems Integration Group to assess the gaps in existing weather data used in power system planning, and outline a process for producing ideal weather datasets for planning studies for increasingly weather-dependent electric power systems. The report provides details on what is needed and why, outlines the status of and gaps in existing data and methods, and describes an approach to building a solid, long-term planning solution.

These documents are available a https://www.esig.energy/weather-data-for-power-system-planning.


# Contents

# Abbreviations Used

| | |
|---|---|
| **BTM** | Behind the meter |
| **DHI** | Diffuse horizontal irradiance |
| **DNI** | Direct normal irradiance |
| **ECMWF** | European Center for Medium-Range Weather Forecasting |
| **ERA5** | Fifth-Generation ECMWF Atmospheric Re-Analysis of the Global Climate |
| **GAN** | Generative adversarial network |
| **GCM** | Global climate model |
| **GHI** | Global horizontal irradiance |
| **HRRR** | High-Resolution Rapid Refresh Model |
| **IRP** | Integrated resource plan |
| **MCP** | Measure, correlate, and predict |
| **MERRA** | Modern-Era Retrospective Analysis for Research and Applications |
| **NCAR** | National Center for Atmospheric Research |
| **NCEP** | National Centers for Environmental Prediction |
| **NREL** | National Renewable Energy Laboratory |
| **NSRDB** | National Solar Radiation Database |
| **NWP** | Numerical weather prediction |
| **WIND** | Wind Integration National Dataset |
| **WTK-LED** | WIND Toolkit Long-term Ensemble Dataset |

# Executive Summary

The electricity system is rapidly transitioning from a system mostly powered by fossil fuels to one in which wind, solar, hydro, and nuclear generators provide most of the generating capacity and energy. At the same time, energy-limited resources such as battery storage are rapidly becoming more prevalent, and behind-the-meter generation is blurring the lines between generation and load and between transmission and distribution. Concurrently, load is fundamentally changing as transportation and heating electrify. These widespread changes lead to the increasing weather-dependence of supply and demand, making power system planning dramatically more complex and requiring much more comprehensive weather data for robust system planning.

**To robustly quantify possible supply/demand combinations in future planning scenarios requires long time series of temporally coincident weather variables that accurately describe the weather impacts concurrently affecting the electricity system.**

## Increasing Weather Dependence and Weather Complexity

The electric power system has always been affected by the weather. Demand has long been modulated by weather conditions, temperature in particular, which also impacts thermal generator capacity (especially gas turbines), cooling water availability, and transmission capacity. Hydro power is obviously impacted by the environmental water cycle. All types of extreme weather impact generators, fuel supplies, and transmission and distribution infrastructure. Going forward, available generation will increasingly be defined by the weather occurring at the location of every wind or solar plant. The behavior in time and space of multiple weather variables—in particular, temperature, wind, and solar irradiance—increasingly affects the amount of generation possible.

The result is a much greater range of possible outcomes for supply and demand, as they will be driven by the behavior of multiple variables in time and space, as opposed to largely temperature impacts on large load centers. Often, between 10 and 40 years or even more of weather data are needed to capture this range. To robustly quantify the range and probability of possible supply/demand combinations in future planning scenarios requires long time series of temporally coincident weather variables that accurately describe the frequency distribution and evolution of all the weather impacts concurrently affecting the electricity system.

## Accounting for the Increased Weather Dependence and Complexity in Power Systems Models

Power system planners' efforts to develop resource portfolios that are cost-effective, reliable, and resilient require accurate estimates of increasingly weather-dependent generation and load. Although the use of weather observations would be ideal, this is not practical, and a modeling methodology must be used to synthesize the time series evolution of the necessary variables as accurately as possible. Planners study the power system with a variety of models that can simulate the operation of existing and hypothetical power systems, determine optimal capacity build-outs, and assess resource adequacy, which measures the probability or likelihood of a power system having insufficient resources to meet load. To

**The need for appropriate weather data to plan for the rapidly increasing weather dependence of both generation and load is becoming acute.**

evaluate a power system's dependence on weather, power system models require highly detailed weather data including the key variables influencing solar, wind, traditional generation, transmission, and load.

While in the past, the impact of temperature on demand was by far the biggest weather driver, these relationships are changing. The electrification of transportation and heating are making the end-use loads more susceptible to weather extremes, especially in winter. In addition, as the energy transition proceeds, data for wind speed and solar irradiance are critical for defining wind and solar generation patterns. Increasing weather dependence and complexity also mean that weather associated with outages and derating for transmission and all generation types can cause common mode failures that dramatically amplify weather impacts.

The need for appropriate weather data to study and plan for these impacts is becoming acute.

## A Need for Much Higher Fidelity Weather Data That Meet Several Important Criteria

The most pressing need is to be able to estimate the supply of wind and solar generation in current and future resource portfolios. This requires that the weather driving these generators is accurately quantified at every plausible location where such generators exist or may be built, including customer-sited generation behind the meter. In addition, the data must represent the chronological evolution of weather variables in order to model and optimize the charge and discharge of energy storage.

While weather modeling of the power system has improved considerably in the last several years, there are still major gaps and inaccuracies in the data available to power system planners. Planners lack the necessary information to properly quantify and mitigate reliability risks for power systems transitioning to a fundamentally new resource mix.

No currently available datasets meet all the above criteria for power systems studies in U.S. geography. The National Renewable Energy Laboratory (NREL) Wind Integration National Database (WIND) Toolkit meets some of the criteria for wind generation estimates, and the NREL National Solar Radiation Database (NSRDB) meets some of them for solar generation estimates. Together these provide the rudimentary datasets that power system modelers are typically using today, but this approach is not a tenable solution looking ahead.

**The work required to achieve a long-term solution to weather data needs is not trivial, but it is manageable and is much less costly than blindly building trillions of dollars of infrastructure without the basic tools to cost-effectively optimize it and assess its reliability.**

The Energy Systems Integration Group convened a project team to assess the gaps in existing weather data used in power system planning and outline a process for producing ideal weather datasets for planning studies for increasingly weather-dependent electric power systems of the future. This report provides details on what is needed and why, outlines the status of and gaps in existing data and methods, and describes an approach to building a solid, long-term solution (Table ES-1, p. x). The work required is not trivial, but it is manageable and is much less costly than blindly building trillions of dollars of infrastructure without the basic tools to cost-effectively optimize it and assess its reliability. Continental-scale multi-decadal assessments are becoming more commonplace, particularly over Europe (e.g., through academic institutions or interactive climate services), and some of these tools have global modeling capabilities for renewable system components. However, none of these assessments meet all of the above requirements, either (primarily due to inadequate resolution and insufficient validation).

## Disconnect Between the Power System Modeling and Meteorology Communities

The processes that drive weather involve complex interactions of many variables, especially phenomena

**TABLE ES-1**

**The Main Attributes of Time Series Data Necessary to Meet General Power System Modeling Needs**

| | |
|---|---|
| **Including the necessary variables** | Include the necessary variables at sufficient spatio-temporal resolution and accuracy to reflect actual conditions that define the generation potential at current and future wind/solar sites and temperature at load centers |
| **Covering multiple decades with ongoing extension** | Cover multiple decades with consistent methodology and be extended on an ongoing basis to capture the most recent conditions and allow climate trends to be identified |
| **Coincident and physically consistent** | Are coincident and physically consistent, in space and time, across weather variables |
| **Validated** | Are validated against real conditions with uncertainty quantified |
| **Documented** | Are documented transparently and in detail, including limitations and a guide for usage |
| **Periodically refreshed** | Are periodically refreshed to account for scientific and technological advancements |
| **Available and accessible** | Publicly available, expertly curated, and easily accessible |

Source: Energy Systems Integration Group.

such as local circulations that impact wind generation and cloud and aerosol (including smoke) processes that impact solar generation. The number of observations that would be needed to accurately describe the factors defining the amount of generation in different renewable energy portfolios is orders of magnitude higher than what is currently available or realistically possible. Fortunately, the atmosphere follows physical laws, and, using the available observations and sophisticated computer programs that model the laws governing atmospheric processes, it is possible to fill in many of the data gaps.

Weather model output, however, has limitations. All too often, synthetic weather data produced by these models are either used in power system modeling as if they are equivalent to high-quality observations, or, on the other end of the scale, model output is rejected in favor of simpler, easier-to-understand observational records that are then extrapolated using statistical methods with dubious scientific basis. Both outcomes lead to study results that have greater uncertainty than is typically advertised and may result in poor downstream decisions when model-synthesized data that "seem reasonable" are assumed to accurately reflect actual present or future conditions.

**When significant amounts of weather-driven renewable resources are being evaluated, power system analysis needs to involve meteorologists with an understanding of power systems, who can advise on the best meteorological data sources to use for a task and shine light on the possible biases and uncertainties that these choices will produce.**

For this reason, when significant amounts of weather-driven renewable resources are being evaluated, power system analysis needs to involve meteorologists with an understanding of power systems—who can advise on the best meteorological data sources to use for a task, shine light on the possible biases and uncertainties that these choices will produce, and work together with power system modelers to refine both meteorological modeling and power system modeling codes and processes to provide the most effective coupling between the two. Although cross-sector engagement is increasing, and many meteorologists are now working on energy transition research problems, the level of engagement is still too low. There is a need for more direct engagement by the meteorology community in the energy sector, with the meteorology and power systems communities working together and each learning about the needs, constraints, and capabilities of the other field.

## Poor Data Validation and a Lack of Sector Cooperation

Power system models' reliance on synthesized weather data makes it critical to have confidence in these underlying data—and to understand their limitations and uncertainties—so that the resultant power system analyses can be trusted and their reliability quantified. Confidence is developed by validating and quantifying the uncertainty of the synthetic data, a process that compares the model data to as many ground truth observations as possible, across as much time and space as possible. However, few observations exist for the validation process.

To assess the quality of synthesized data, it is critical that all available observations are utilized. This may mean installing some new observation stations in places where wind and solar are likely to be deployed, but the most obvious and cheapest solution is to make the thousands of observations now available at existing (and future) wind and solar plants generally available. Keeping data proprietary is counterproductive to promoting a transition to wind and solar generation and must change to enable more accurate system modeling, including a better understanding of the accuracy and uncertainty of synthetic time series of weather data. As counterintuitive as it may seem, it is highly likely that time series generation estimates for periods in the past that are being used for power system planning models are of considerably *lower* quality than the time series forecasts of generation produced for operations a day or two in advance of the time they are estimating. This is a direct consequence of the lack of validation and data sharing.

## Climate Change Adds More Degrees of Freedom

In the near term—certainly the next five years, but likely the next one to two decades—the overall impact of climate change on the accuracy of power system studies

**In the near term, the overall impact of climate change on the accuracy of power system studies will be small relative to the enormous effect of increasingly weather-driven capacity, energy-limited resources, and electrification.**

will be small relative to the enormous effect of increasingly weather-driven capacity, energy-limited resources, and electrification.

The impact of climate change cannot be ignored, however, as it is likely to have increasing impacts on the power grid, including temperature-driven record loads and impacts on transmission and generation driven by more extreme and/or frequent weather. If climate extremes combine with common-mode failures because, for example, temperatures fall outside of those typically experienced by the grid, the impacts could be profound. However, in this report the focus is on delivering tangible information about modeling the distribution of coincident weather outcomes on grids with increasing levels of wind and solar. These data are based upon real observations of the atmosphere in which the gaps between observations have been filled by meteorological models that produce results that are physically consistent with the initial observations and internally consistent across weather variables.

Long time series produced in this way (especially if they are extended in an ongoing fashion) will reveal emerging climate trends. However, predicting the future climate requires a different approach with much more uncertainty, not just in the modeling, but (among other things) in assumptions about emissions pathways. The report includes a short section that discusses the key caveats of climate change on the report's central concerns and, as appropriate throughout, notes how climate change could impact its conclusions. It also introduces some of the work and techniques being undertaken in the area of climate change impacts on power systems. A deeper treatment of this topic is recommended for a future task force.

## A Roadmap for Meeting Weather Input Needs in Power System Modeling

There is an urgent need to develop one or more datasets that can become the standard for the power/electricity sector to use now, and moving forward for the foreseeable future, as weather inputs to planning studies including renewable energy integration studies, resource adequacy assessments, capacity expansion planning, and integrated resource planning. Thoughtfully produced, archived, and curated, such data would also be valuable in other important tasks associated with the power system, including renewable energy resource assessments and renewable energy performance analyses, as well as being extremely useful for foundational research work to examine the relationships between supply and demand and weather patterns/climate signals, and for establishing possible climate trends.

There can be no reliable energy transition without broadly available, consistent, weather datasets for power system studies that meet the criteria outlined above. Given public policies that promote or require increases in renewable energy, the necessary data can be considered a public good—one that is government funded, publicly available, and routinely maintained.

There are two stages in the development of an ideal weather dataset.

### STAGE 1: Validate and Refine Requirements and Confirm Need and Fitness

The initial stage of building an ideal weather dataset would convene a technical review committee composed of expert power system stakeholders, experienced energy meteorologists familiar with how power system modeling

**There is an urgent need to develop one or more datasets that can become the standard for the electricity sector to use as weather inputs to planning studies including renewable energy integration studies, resource adequacy assessments, capacity expansion planning, and integrated resource planning.**

is performed for both hypothetical studies and actual utility or system planning, experienced numerical weather prediction (NWP) modelers whose experience covers high-resolution modeling and data assimilation, and experts in NWP post-processing methodologies including bias correction and downscaling techniques employing machine learning techniques.

The technical review committee would vet and refine the dataset requirements; determine possible methods to create the sample datasets; using three to seven candidate methods, produce sample datasets; and determine whether the candidate datasets add value over the controls. It would select the method with the best combination of cost and accuracy and move to Stage 2.

### STAGE 2: Produce Historical Archive and Ongoing Process

Once the value of a dedicated process to produce a high-fidelity archive is established, the next step is to build the archive and operationalize the process of ongoing extension using the method selected in Stage 1. The main decisions would be how far the archive will go back and when operational extension will be performed; the rest of the process of developing the data should be relatively straightforward and automated.

At this stage, curation of the data will be key to its usability and to understanding its limitations and uncertainty. The following issues would need to be thought through:

- How to ensure that users can efficiently access the data they need
- Building out of a broad observation network to be used in properly validating high-resolution output, in data assimilation where NWP-based solutions are deployed, and in post-processing to reduce systematic errors
- Ongoing validation
- User education
- Documentation of alternative data sources

---

With rising levels of wind, solar, and storage and increased electrification, power system planning is becoming more complex and more weather-dependent—with a greater need to accurately model the impacts of weather variables on resource adequacy and system reliability. Accurate power system analysis requires time series data for key weather variables that are temporally coincident, have sufficiently high spatial and temporal resolution, and are robustly validated. The availability of such an ideal weather dataset, together with education and coordination between the meteorology and power system communities, will equip system planners to guide future resource siting and build-out for a reliable, high-renewables grid.

# Introduction

The impacts of weather in the electricity sector have always been important, as weather modulates demand and impacts much of the infrastructure traditionally used to generate and deliver electricity. This relationship is growing stronger with the increase in weather-driven renewable generation and the growing electrification of sectors such as buildings and transportation. To plan and operate the electricity system reliably and cost-effectively, it is imperative to gain more complete knowledge of potential weather impacts to the system. Future reliability is at risk unless we understand the range of different supply and demand balance possibilities that are driven by physically plausible weather combinations—especially combinations that drive demand, generator availability, fuel availability (both renewable resources and traditional supply), and transmission capacity in ways that stress system reliability and resilience. This requires more accurate, more detailed, longer, chronological weather datasets than are available today to describe the range and likelihood of different concurrent weather impacts on electricity system components—especially renewable generation—and load.

**Future reliability is at risk unless we understand the range of different supply and demand balance possibilities that are driven by physically plausible weather combinations—especially combinations that drive demand, generator availability, fuel availability, and transmission capacity in ways that stress system reliability and resilience.**

The Energy Systems Integration Group convened a project team to explore the linkages created by the rising weather dependence of the electricity system; describe the need for, and the nature of, data to represent these linkages; and outline a process for producing ideal weather datasets for power system planning studies for the increasingly weather-dependent systems of the future. This report explores the linkages created by the rising weather dependence of the electricity system and describes the need for, and the nature of, data to represent these linkages. While the same weather relationships impact both electricity system operations and planning, as well as the development and operation of individual renewable power plants, the nature of the data needed is different for each application. The emphasis here is on the weather data inputs needed for power system planning studies for power systems with high levels of renewables. This requires correlated, time-synchronized data for wind, solar, and temperature observations. The report does not include discussion of datasets necessary for planning hydropower generation, as hydro's time granularity of months and seasons does not require the same level of correlation as needed for wind, solar, temperature, and load. The objective is to provide information that is useful both to power system experts who use these data and to the meteorologists involved in creating and providing data and advising on their use.

Although there is a growing number of (mostly model-based) meteorological datasets available for power system planning, there are major challenges and gaps to fill. These include:

- The lack of sufficiently complete, accurate meteorological data to correctly encapsulate possible conditions on the electricity system
- A lack of detailed, objective validation of model-produced weather data against ground truth, especially at times when inaccuracy could yield erroneous conclusions about system adequacy
- Misconceptions among data users about the complex nature, limitations, and applicability of the data that are available
- An overly simplistic application of meteorological inputs in estimating the output of renewable generators
- Misconceptions among meteorological data providers about how the data they make available are applied

**The correlation among different weather variables means that the common practice of increasing the range of possible scenarios by making separate random draws for different variables in correlated datasets is not appropriate. To cover the range of possible outcomes, much longer datasets need to be produced.**

Currently, there are gaps not just in the data needed, but in understanding of the needs and capabilities of the power system engineering and atmospheric sciences communities. Bridging these knowledge gaps is one of the primary goals of this report. One consequence of this goal is that, for any individual person, certain content may seem somewhat basic and other content may seem advanced. Where this happens, readers are encouraged to reflect on how they can help others with different expertise or can seek support across the increasingly large overlap of meteorology and power systems engineering professions. Readers may also refer to the glossary at the end of the report.

This report shows how methods and data used in power system analysis need to become much more sophisticated and must evolve away from a focus on load- and temperature-driven reliability events. The data and models being used are insufficient to accurately determine the important relationships between load and renewables across sufficiently long periods for probabilistic analysis. This is especially important given that time series data for different weather variables are not independent. Their correlation means that the common practice of increasing the range of possible scenarios by making separate random draws for different variables in correlated datasets is not appropriate. To cover the range of possible outcomes, much longer datasets need to be produced.

## Recent Studies

Several recent studies have recognized these needs. For example, the changing nature and increased importance of weather dependence, and why data to quantify it are essential, are illustrated in Novacheck et al. (2021) and

Bloomfield et al. (2021). Recent studies that incrementally improve methods for incorporating weather in power system models include Voisin et al. (2018) and Nahmacher et al. (2016), with some incorporating stochastic methods to represent a wider range of scenarios and extremes, including Wang et al. (2016) and Dyreson et al. (2022), or stochastic methods that capture hydrometeorological variability, as in Su et al. (2020). Work is being done in Europe to produce a weather inputs database that is future-proofed—able to serve the needs of power system modelers and planners as the transition to a high-renewables grid proceeds—and captures some aspects of expected climate change pathways, as discussed in Dubus et al. (2022). The Energy Systems Integration Group (ESIG) Redefining Resource Adequacy Task Force's report discusses how system modeling techniques must change, including consideration of weather dependencies, to better quantify resource adequacy in the evolving system and points to the insufficiencies of currently available data (ESIG, 2021).

**No datasets exist that meet the requirements with sufficient accuracy, spatial and temporal resolution, or record length to capture all of the possible drivers of supply and demand balance in the new paradigm.**

## Lack of Data of Sufficient Quality

While there have been great strides in the availability of high-resolution meteorological data for power system modeling studies over the past decade (especially fields for determining wind and solar output), no datasets exist that meet the requirements with sufficient accuracy, spatial and temporal resolution, or record length to capture all of the possible drivers of supply and demand balance in the new paradigm. This is specifically true for the United States and generally true globally. In addition, the data that are available have not been sufficiently validated to assess the uncertainty of their representations of truth, and thus their appropriateness for use in power system planning. Where validation has been performed, biases and limitations have been discovered even in the current best-in-class data available (see, for example, Sharp (2022)). Data limitations have led to gross simplifications in weather inputs even where practitioners are earnestly attempting to robustly address the added complexity. To fill data gaps, scientifically questionable "bootstrapping" methods are being used to synthesize long data records from whatever limited data are available from operational projects, leading to combinations of weather variables that are either not physically plausible or occur in the bootstrapped data at frequencies that are not representative of observed frequencies.[1] In some cases, modeling limitations prevent the incorporation of multiple years of weather data even where they do exist.

Another important element affecting weather-driven supply and demand in power system planning is, of course, climate change. In the past it was assumed that a sufficiently long record of past conditions (typically at least 30 years) was representative and predictive of the range of future weather conditions. However, the assumption of a stationary climate is no longer valid. This report is focused on the types of weather data that are crucial to collect or simulate using models for application in rigorous power system planning studies, with the complex topic of modeling climate change being mostly beyond its scope. However, while it is reasonable to assume that the overall distribution of weather impacting wind and solar generators will not change radically in the coming one to two decades, especially compared to the impact of adding large amounts of renewable generation to power systems, the influence of climate change is important to keep in mind, especially regarding how changes in temperature may concurrently drive demand and the availability of generation to meet it. Therefore, we include caveats that climate change places on the discussions below where appropriate and summarize these in a brief section at the end of the report. The project team views the present work as a necessary precursor to a broader discussion of how to incorporate climate change into power system modeling weather inputs, which is worthy of follow up work.[2]

## Key Needs Highlighted by This Report

There are many expert meteorologists and power systems engineers, but few people have more than a basic grasp of *both* meteorology and electricity systems, or fully understand just how much more complex a weather-driven system is than one in which weather mainly modulates demand. In addition to resulting at times in overly simple methods to synthesize longer datasets, the lack of holistic understanding creates additional issues that must be addressed now even as more complete data become available. For example, model data are often used as if they have the accuracy and degree of uncertainty of observations, yet their representativeness in time and space is a function of the model configuration used and model inputs. In addition, weather models can have higher accuracy for some weather conditions and lower accuracy for others; therefore, the quality of modeled weather is also a function of the weather that is occurring at any given time and place. At the other end of the spectrum, mistrust of weather model data sometimes results in useful model data being passed over in favor of inputs that are simpler and more familiar but less complete. For instance, overly simple models are often used to extrapolate data from one location to another or to estimate one weather variable using another variable. There is an urgent need for education, coordination, and cooperation between power system experts, meteorologists, and climatologists (Coughlin and Goldman, 2008; Craig et al., 2022; Bloomfield et al., 2022).

**Model data are often used as if they have the accuracy and degree of uncertainty of observations, yet their representativeness in time and space is a function of the model configuration used and model inputs.**

These data and modeling challenges are leading to the inappropriate "black box" application of meteorological inputs, the use of methods with many limitations to fill data gaps, and a lack of appropriate consideration of the uncertainty in modeling results that may lead to poor decisions in power system planning. With hundreds of billions of dollars of new infrastructure being built for the energy transition, it behooves the sector to address

1 Bootstrapping is a statistical procedure that resamples a single dataset to create many simulated samples.

2 Discussions of the challenges that climate change adds to this subject can be found in Craig et al. (2018), Craig et al. (2019), and Bloomfield et al. (2021), and examples of studies that incorporate climate change in power system modeling in at least one meteorological variable include McFarland et al. (2015), Schlott et al. (2018), Craig et al. (2019), Miara et al. (2019), Peter (2019), Turner et al. (2019), Steinberg et al. (2020), Voisin et al. (2020), Hill et al. (2021), Ralston Fonseca et al. (2021), and Cohen et al. (2022). In addition, the Electric Power Research Institute recently launched its Climate Resilience and Adaptation initiative (Climate READi), which takes a broad approach to climate change impacts on the electricity system, aiming to help people across the industry identify optimal investments for power systems' climate resilience.

these challenges. While these challenges are not trivial, the cost of overcoming them is trivial compared to the risks presented by the current data inadequacies.

This report documents the weather-related challenges in power system planning and describes a transdisciplinary approach to overcoming them. First, the project team gives an overview of the different sources of meteorological data for power systems engineers and provides context from which meteorologists can better understand the data needs of power systems engineers. We describe the datasets that are currently available and show how they are not sufficient for the tasks at hand. We conclude by offering recommendations for producing robust, future-proof datasets that better serve the power systems community and examples of possible approaches for putting existing data to use.

The report's major sections are as follows:

- **Section 1**, "The Challenges of the Evolving Weather/Energy Nexus," looks holistically at the impacts of the weather on electricity systems, describes systems' increased weather dependence, and articulates the consequences and resulting meteorological needs of power system planning processes.
- **Section 2**, "Meteorological Data Fundamentals for Power System Planning," presents an overview of meteorological data for the power systems audience, covering the different sources of data available to the community and their benefits and limitations.
- **Section 3**, "Weather Inputs Needed for System Planning," describes the weather-sensitive inputs to system planning models and discusses the weather data needed to produce the inputs. For instance, power system models' ability to estimate wind generation requires knowledge of wind speed and direction as well as several other secondary weather variables.
- **Section 4**, "An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data," provides a comprehensive list of weather variables needed for system planning studies and articulates in detail the necessary attributes of such data (resolution, length, validation, etc.). The section goes on to describe the data currently available and compare them to the desired data and attributes.
- **Section 5**, "Project Description for Producing Robust Weather Inputs Data," reiterates the urgent need for comprehensive datasets that meet the requirements outlined in Section 4, and outlines a plan for producing the required data.
- **Section 6**, "Guidance for Using Existing Weather Inputs," presents examples of promising approaches for effectively using existing weather data: what the gaps and consequences of these gaps are, how to use the existing data most effectively, and what to be aware of as improved data become available.
- **Section 7**, "The Impact of a Changing Climate," briefly discusses the consequences of climate change for power system modeling and the weather inputs used.
- **Section 8**, "Summary and Next Steps," summarizes the key takeaways of this report including the current status, need for, and benefits of taking action; the requirements for data that will address the current gaps and limitations; and how to produce such data.

SECTION 1

# The Challenges of the Evolving Weather/Energy Nexus

Until recently, the biggest weather impact on electricity systems was temperature modulating load, with the coldest winter days and warmest summer days setting the amplitude and timing of peak demand, and extreme weather events driving outages of generation, transmission, and distribution. Today, as wind and solar resources comprise an increasingly large portion of the supply, on some systems the effect of temperature on demand has now been surpassed by the influence of weather impacting supply via wind and solar generation. At the same time, the weather dependence of load is increasing due to electrification of heating, cooling, and transportation (Figure 1, p. 7).[3]

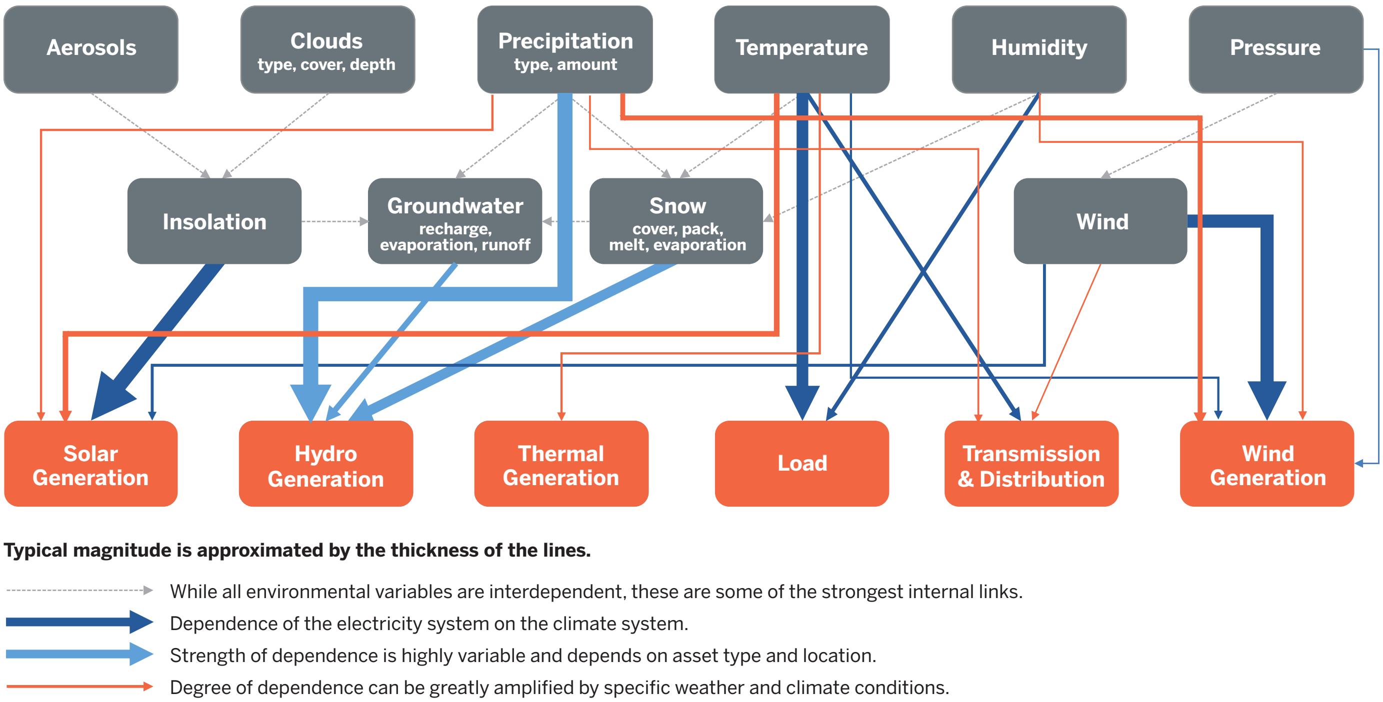
These variables are interrelated in complex ways that vary according to the daily, seasonal, and interannual variability of weather. In addition, the distribution of different combinations of each may be evolving in time due to changes in the Earth-system energy balance that is being driven by rising greenhouse gas concentrations. This linkage is particularly complex, as it is not yet well understood how a changing climate will affect weather that impacts wind and solar resources, and there is little certainty around how quickly greenhouse gas emissions will be abated.

The remainder of this section will discuss the nexus of weather and energy and its evolution, with a focus on the

3 For definitions of terms that some readers may be unfamiliar with, please see the glossary at the end of the report.

FIGURE 1

## Electricity System Weather-Dependence




**The primary linkages between variables in the weather and climate system (gray) and the electricity system (orange). There are many feedbacks between the environmental variables; the strongest links are shown in dashed gray lines. Dark blue lines indicate direct dependencies that are most important in everyday operation of the electricity system, while orange lines indicate dependencies that do not typically have a large impact on a daily basis but can have a profound impact in particular circumstances or combinations. For instance, freezing temperatures, high humidity and/or freezing rain can cause wind generation to become unavailable due to icing, and extreme winds can damage transmission and distribution infrastructure. Light blue lines denote where the strength of dependence is highly variable and depends on asset type and location.**

Source: Energy Systems Integration Group.

implications for power system planning and modeling. Appendix A provides additional insight into the intersection of weather and energy including a broader look at how it impacts other power system functions such as real-time operations and renewable generation resource assessment and performance assessment.

## Increasingly Weather-Driven Supply

Outside of hydro-dominated electricity systems, temperature was, until recently, the primary weather variable impacting electricity supply. While the impact of weather on supply is still smaller than on demand today, the effects are significant and will continue to grow. The efficiency of thermal plants is reduced with warmer ambient air and higher cooling-water temperatures. Temperature also influences the reliability of thermal and renewable generators, with periods outside of typical ranges more likely to see forced outages (Murphy, Sowell, and Apt, 2019). Additionally, cold temperatures can affect the availability of the natural gas supply to gas-fired generation through conflicts with residential heating, decreased pipeline pressure, and increased failures of gas transportation infrastructure. Transmission and distribution systems are also impacted by weather. Temperature and wind can affect transmission line ratings, precipitation patterns on scales of days to years impact hydro output, and drought increases the likelihood of wildfires that can impact the transmission system. And weather extremes including icing, snow, high winds, and lightning all impact transmission and distribution systems.

Today, supply-side weather dependence is increasing rapidly due to the rising levels of wind and solar generation, with wind speed and direction and solar irradiance becoming chief drivers. Wind power output is especially sensitive to small changes in wind conditions, because the power density in moving air (wind) is proportional to the third power of wind speed (if the wind speed doubles, the power density increases by a factor of $2^3=8$). Both wind and solar generation are impacted by temperature (which on calm, hot days can significantly reduce solar panel efficiency), humidity, and precipitation (of all types), as will be detailed in Section 3, "Weather Inputs Needed for System Planning." Solar and (to a lesser extent) wind generation are also impacted by smoke and other atmospheric aerosol loads, both of which are strongly influenced by weather patterns.

Weather impacts on renewables today and going forward are higher in amplitude than the chiefly temperature-driven weather impacts on the electricity system historically. Both wind and solar resources can go to zero for periods of time. In addition, unlike for hydro generation in which the impacts of weather events affecting water inflow exhibit significant buffering in time and space, changes in weather at the location of wind and solar generators impact their output immediately. The weather impacts on variable renewables are also more complex. Wind and solar generators can be located in a diverse range of locations, and resource conditions can vary considerably across short distances and change rapidly in time. The conditions in one part of a wind plant, for example, can vary considerably from those in another part of the same plant, especially in complex topography.

Increases in average ambient temperatures and the frequency of extremely warm days will lead to more weather-related transmission derates and more thermal plant derates and outages related to ambient temperatures and cooling water temperatures. It is also speculated that rain, snow, and extreme cold may become more prevalent as the climate changes (though the atmospheric science community has not reached a consensus on this matter), which would further increase weather-related outages for all generator types.

This significantly greater supply-side weather dependence and the lack of data to fully quantify the impact at any given location and time (past or future) is a primary motivator for this report.

## Increasingly Weather-Influenced Demand and Increasing Share of Weather-Dependent Distributed Energy Resources

The electrification and decarbonization of transportation and the built environment will dramatically increase loads. Increased use of electric heat will increase winter loads, and increased air conditioning use, driven by overall warmer temperatures and general growth in adoption, will increase summer loads. Efficient heat pump technology reduces the overall load growth of the transition to electric heat, but in colder environments, on the coldest days loads will spike when less efficient supplemental resistive heating is necessary. The use of electric transportation will of course greatly increase overall loads and has a weather-dependent component associated with energy used to condition the vehicle cabin.

The impact of weather on demand shape and amplitude is being changed by increasing levels of behind-the-meter (BTM) generation (mostly rooftop solar). The impact of BTM generation is essentially another incarnation of the increasingly weather-driven supply discussed above, but it has the additional complexity of tending to be highly distributed and embedded in demand profiles because detailed resource data are typically not available.

**Sensitivity of load to weather is increasing and becoming more complex. Models used to predict load growth and models used to forecast load are increasing in sophistication and require more and higher fidelity data.**

Thus, the sensitivity of load to weather is increasing and becoming more complex. Temperature dependence is increasing, and variables impacting solar generation

are becoming important. As a result, models used to predict load growth and models used to forecast load are increasing in sophistication and require more and higher fidelity data.

## Increasingly Vulnerable Transmission and Distribution Infrastructure

The dependence of transmission and distribution systems on weather and weather extremes is well known. However, the vulnerability of transmission and distribution systems is rising due to changes in extreme events and their frequency, especially events like region-wide temperature excursions. The transmission of power will play an important role in the adoption and integration of variable renewable resources, and some weather events will concurrently stress every part of the power system; therefore, the planning of mitigation and response strategies requires a better understanding of the interplay of weather drivers of supply and demand and concurrent weather drivers of transmission and distribution impacts.

## Increasing Interdependence of Weather Influence on System Components and Increasing Complexity

While the ongoing increase in weather dependence across supply, demand, and transmission and distribution is clear, there is a concurrent increase in the complexity of the weather impacts on these areas as well as increased interdependence among the weather impacts. The relative uniformity of temperature across a region meant that in the past, relatively simple relationships could be developed between the temperature at a small number of weather observation sites within population centers and the load expected within a balancing area, and between the temperature measurements near thermal generating facilities and the outage probability at the facility.

However, simple relationships between individual observational data points and wind and solar resource within a region are much less common, and the resource can vary dramatically across short distances. For example, weather that brings cold temperatures may result in above-normal wind speeds in one part of the region and below-normal wind speeds tens of miles away, or possibly even closer. The same is true for cloud cover. These relationships between weather variables and solar or wind power output are defined by interactions between the big picture (synoptic and mesoscale) weather conditions (which primarily drive temperature) and local surface features; the relationships are highly complex because they are defined by multiple weather parameters. At the same time, wind and solar generation facilities are widely distributed and located away from populous areas where weather observations are prevalent. They also cover large areas that can have significant weather variation within them. Therefore, observations that are representative of conditions at existing or future renewable facilities are often not available, and where they are, they usually do not cover a representative and sufficiently long period.

To recap, multiple weather variables are driving both supply and demand in ways that range from strongly synergistic to strongly antagonistic, and these variables vary significantly across an electricity system footprint and across time. Data defining the distribution of these

**Multiple weather variables are driving both supply and demand in ways that range from strongly synergistic to strongly antagonistic, and these variables vary significantly across an electricity system footprint and across time. Data defining the distribution of these variables in time and space are not currently collected (or modeled) at anything close to the required fidelity.**

variables in time and space are not currently collected (or modeled) at anything close to the required fidelity.

As noted above, thermal generation derates and outages also have weather dependencies, as do transmission and distribution. When combined with the rapidly growing share of variable energy resources in a power system's generation mix, this results in very complicated relationships linking weather and the supply of, demand for, and transportation of electricity, especially when considering that the line between generation and consumption, and transmission and distribution, is now blurred by BTM generation. Periods of risk and their causes are radically shifting, and the nature and diversity of system stress are becoming much more complex, because they are driven by coincident weather impacts on renewable supply (which will vary across the footprint), load, generator availability, transmission, distribution, and hydro generation. As a result, each part of the system can no longer be considered independently. It is crucial to have data that can allow the envelope of possible supply and demand balance combinations to be quantified for existing and hypothetical future electricity systems to assess system reliability across the range of expected weather conditions.

## Climate Change: The Wild Card in the Deck

Climate change poses an additional layer of complexity and uncertainty in power systems' weather dependence. In the past, climatology that was measured over a sufficiently long period was assumed to be stationary, so historical temperature and (for hydro) water inflow observations were mostly sufficient to model the primary impact of weather on future loads and generation. However, the assumption of stationarity is no longer valid. Not only is the power system becoming much more influenced by the weather, but the weather itself is increasingly deviating from historical norms.

The effects of a changing climate on electricity systems are potentially large, especially as these systems become increasingly weather-dependent, and some studies have started to modify weather input data to assess electricity system sensitivity to temperature and precipitation changes. There is some degree of scientific consensus regarding expectations around average temperature and to some extent changes in its extremes. There is also

some consensus regarding precipitation changes, although there is still considerable uncertainty. However, the impacts of climate change on wind and solar resources are only just beginning to be examined at scales necessary to model their impact on supply and assess how these impacts correlate to temperature and precipitation changes. Global climate models (GCMs) generally cannot predict changes in wind and solar resources at sufficient spatio-temporal resolution for use in system planning models. When output from different GCMs is downscaled (using methods discussed in the next section), large disagreements are seen between the different models for wind and solar irradiance, and these will change in the future. Further, there are limited options to determine which resultant dataset, if any, is likely to be most representative of future conditions. In short, the uncertainties associated with the impacts of climate change on the electricity system are potentially large and not yet well understood.

That said, large changes are generally not expected in the overall spatio-temporal wind and solar resource distributions over the coming decade. For these reasons, and because, at least in the short term, it is more urgent to quantify the impact of weather variables on rapidly expanding renewables and rapidly changing demand, this report does not attempt to provide definitive recommendations around producing and using weather inputs that incorporate climate change projections into resource adequacy and other system planning studies. Section 7, "The Impact of a Changing Climate," summarizes central questions about the impacts of climate change on supply and demand in future power systems and the areas where research and development are needed.

## The Need to Understand and Quantify Electricity System Weather Dependence

Because the impact of weather on the electricity system has broadened, it is no longer enough to simply ensure there is enough generation to meet the peak loads for the climatologically hottest and/or coldest days. While in the future the overall electricity consumption will still peak on the hottest and coldest days, demand will no longer be primarily described as a function of time of year, day of week, time of day, and temperature, and utility-scale generation will no longer be simply a function of available capacity and outage rates. Both demand and supply have large, rapidly growing components that are influenced in numerous ways by different weather variables. The weather variables all vary in time and space and in ways that are interrelated. This increase in the number of weather variables, and the number of locations at which these variables have a significant impact, means that much more weather data are needed to estimate the weather impact on the electricity system at any given moment. Further, weather data spanning many years are needed, because it is necessary to determine the range of possible outcomes of these variables and their likelihood of occurring. The weather variables in these data must coincide in time and represent realistic chronology of weather patterns, because the variables are interrelated by atmospheric physics and this physics also defines their evolution in time. Thus, the data must come from observations or be synthesized by models that correctly represent atmospheric physics.

The key weather data required to analyze weather impacts on power systems have several critical gaps: either they are not observed at the required locations, the observation record is too short to meet all the needs of the sector (especially for modeling used in system planning), or (typically for operational purposes) a prediction of future conditions is required. To synthesize these data, physics-based models are commonly used, particularly numerical weather prediction (NWP) models.

It is important to note that the term "prediction" is something of a misnomer in this context. While it is true that these models are used to predict future states of the atmosphere, they are also often used to synthesize past data in a way that is consistent with atmospheric physics. Therefore, the terms "forecast" and "prediction" are often used throughout this report to mean making estimates about periods in the past, which can be confusing to non-meteorologists. See Box 1.

**While in the future the overall electricity consumption will still peak on the hottest and coldest days, demand will no longer be primarily described as a function of time of year, day of week, time of day, and temperature, and utility-scale generation will no longer be simply a function of available capacity and outage rates.**

BOX 1

### Three Definitions of "Forecast"

The term "forecast" is used here in three distinct ways:

- To predict what is expected in **the operational time frame** (e.g., day-ahead or hour-ahead) to conduct reliable and efficient market and power system operations.
- To predict how a time series of a parameter such as load may change in **future years** based upon historical relationships of weather and past outcomes, and the expected overall change in the parameter's magnitude.
- To estimate a time series for a weather variable **for a period in the past** at locations for which no observational data are available, by using available weather data and a model. Typically, numerical weather prediction (NWP) models are used, and the main variants of this process are reforecasting and reanalysis, which are covered in detail in Section 3, "Weather Inputs Needed for System Planning." While the data being predicted are for a period in the past, the models are the same as those used for weather forecasting, hence the term "forecast" is often used by meteorologists to refer to this modeling of periods in the past. However, because the term "forecast" has a strong connotation of "future" for most people, in this report we strive to use terms other than "forecast" when referring to the past—such as reforecast, reanalysis, modeled, and simulated—to minimize confusion.

## Operations

To operate the electricity system efficiently, an operator needs to use the lowest-cost generators available whenever possible, which are usually wind, solar, and hydro, as these have no fuel costs. Wind and solar generation are driven by weather patterns occurring in real time, the same weather patterns driving demand. Thus, operators need to have not just quality forecasts of temperature at load centers to predict demand, but also quality forecasts of the variables defining wind and solar output at every wind and solar generation facility. Wind and solar generation forecasts have advanced considerably in the past decade and can provide a reasonably good view of what renewable generation will be available for immediate dispatch, in the next few hours, and a day or two ahead. Operators can use these along with appropriate operating margins to ensure that low-marginal-cost wind and solar generation can meet as much load as possible, while ensuring that other resources are available to meet the balance of load in a least-cost and secure manner, as well as make up any shortfall due to imperfect load, wind, and/or solar forecasts and unplanned generator outages.

## System Planning

Forecasts in the operational time horizon, however, do not provide the information needed for the many ways power system planning tools and methods incorporate weather data. For example, ensuring that sufficient generation capacity is available in periods of scarcity requires knowledge of the conditions that will drive scarcity, which are increasingly weather-dependent and increasingly complex. This is the realm of system planning, where the objective is to ensure that resource adequacy—supply meeting demand—is maintained in a manner that meets cost and policy goals. If a policy goal is to reduce the use of fossil fuel generation, non-fossil resources must be deployed in a way that is cost-effective and, as more fossil plants are retired, still ensures resource adequacy. If the non-fossil resources include wind and solar, this means having a complete picture of how the portfolio of those resources will perform across all reasonably expected demand situations while accounting for the availability of hydro and thermal generation and any transmission and distribution constraints. This requires a database of weather information that has sufficient geographical and temporal resolution to

estimate the concurrent effects of weather on every weather-driven system component.

In power system planning, probabilistic resource adequacy analysis evaluates whether a power system has sufficient resources to serve demand, across a wide range of uncertainty arising from fluctuations in load, fluctuations in the availability of renewable resources, and unplanned generator outages. Each of these uncertainties is fundamentally driven by weather variability.

To quantify this uncertainty, resource adequacy analysis is conducted in a probabilistic manner across hundreds or thousands of samples, each of which varies load, renewable resource availability, and forced outages, to determine the likelihood of different scenarios occurring. It is critical that these samples capture the appropriate range of potential weather outcomes, which requires evaluating many years of weather data. Often, between 10 and 40 years or even more of weather data are needed to capture this range, with the necessary time series increasing in length as the number of interdependent weather variables increases the possible range of supply and demand outcomes.

The introduction of large shares of energy storage and flexible loads also requires that data be evaluated in a *chronological* fashion, as the availability of these energy-limited resources during a scarcity event is dependent on the system conditions in hours both preceding and following the event. Weather data need to be available chronologically on at least hourly timescales. In addition, the probabilistic resource adequacy analysis should evaluate *correlation* across uncertainties. While traditional resource adequacy assessments treated load, availability of renewables, and forced outage rates as uncorrelated, actual practice has shown that each of these can be highly correlated due to underlying weather conditions. This is especially true during extreme weather events when the system is most stressed.

**If renewable generators, especially wind plants, are strategically placed not only where average production is good but where high production occurs when demand is high, this resource diversity leads to more steady supply and fewer risky periods.**

The ability to estimate the performance of wind and solar resources using a full/accurate understanding of their weather dependence will also give system planners the tools to refine their approach to siting these resources. Resource diversity can play a powerful role in smoothing out the amount of renewable generation available across a system's footprint. However, if renewable plants are sited only where the resource is best on average on an annual basis, for example, this lack of resource diversity means that times of high demand and/or low resource become periods with risk of scarcity. However, if renewable generators, especially wind plants, are strategically placed not only where average production is good but *where high production occurs when demand is high*, this resource diversity leads to more steady supply and fewer risky periods. The current ad hoc build-out of renewable resources is unlikely to see the true benefits of that diversity. Although even intentionally planned resource diversity will not eliminate all mismatches between demand and the timing of high renewable energy supply, weather patterns are governed by dynamical rules and tend to cluster around certain outcomes (for instance, the wind does always blow behind a cold front). Thus, careful planning of renewable resource diversity, and transmission, can greatly mitigate variability and supply/ demand mismatch.

The weather data needed for power system modeling for planning purposes are central to the purpose of this report and are described in more detail in Section 3, "Weather Inputs Needed for System Planning." Subsequent sections then discuss how to provide the needed data and offer guidance around how to proceed in the meantime until data specific to the need become available.

## Analysis of Renewable Resource Operation, Planning, and Performance

In addition to the broad categories of system operations and system planning, weather data are also needed for a range of renewable resource development activities, including identification of prospective sites, evaluation of projects' generation expectations and variability, generator placement (especially for wind turbines), and optimal sizing and siting of battery storage. These activities require the same data and a similar temporal length as those needed for power system planning, but they focus on small geographical areas and the specific variables associated with the resource type. At the same time, they often require more detailed spatial resolution. Weather data are also vital for renewable resource project operations and maintenance, and power scheduling and participation in market processes handled by renewable facility operators. Like for system operations, these data are usually current observations or short-range forecasts of future conditions but are focused on the specific project geography. Lastly, weather data are used in performance analysis of renewable resource projects to determine the fuel (renewable resource) availability and other environmental conditions and compare them to ' generator output.

In summary, the weather and electricity sector nexus is strong and growing, and there are many needs for weather data in power system planning and operations. This report is focused primarily on the data needed for system planning, specifically, for probabilistic resource adequacy analysis and capacity expansion modeling. The weather data needs in these areas span large geographical areas across dozens of historical and future weather years. The report describes the needs in detail and proposes an approach to robustly fulfilling them.

SECTION 2

# Meteorological Data Fundamentals for Power System Planning

This section gives an overview of the nature of meteorological data available for use in power system planning. It is important for users to be aware of the challenges when applying weather data to power system planning. Today's available observations are generally too sparse to be used for renewable generation estimates, and data from weather models often used as a proxy for observations have limitations that need to be understood to estimate their impact on the results of power system modeling. Further, simple models sometimes used to synthesize the wind and solar profiles for a given day based on observed predictors like temperature may appear to produce a long time series that looks as though it reflects reality quite well; however, careful validation will usually reveal a poor match with reality, especially when one looks at coincident combinations of different variables across a region. (For definitions of terms that some readers may be unfamiliar with, please see the glossary at the end of the report.)

## Need for Accurate, Long-Duration, Chronological Weather Datasets for Power System Studies

Power systems span continents, with weather events in one corner of the grid having an impact on operations hundreds of miles away. Therefore, analyzing how weather will impact the electricity system means knowing, with a reasonable degree of certainty, the evolution of weather in time and space that impacts electricity system

supply (generation), demand (load), transmission, and distribution.

Accurately representing the state of a modern electricity system where wind and solar generation are distributed over wide areas and often far from load centers, requires:

- Knowledge of the weather variables driving the generation potential at the location of every weather-driven generator, as well as every potential generation location if portfolio expansion modeling is being conducted
- Knowledge of the weather variables driving demand at load centers
- Details of weather affecting other system assets, for example, weather that may cause thermal generator derates or outages or changes to the transmission or distribution system. In addition, the hydrological state needs to be known if there is significant hydro generation.

**Long records are crucial to capture the range of atypical weather combinations that produce weather-related risk, and because the number of variables increases, the range of atypical combinations that produce risk also grows and requires longer records to capture.**

Planners use historical time series of weather records to project likely future scenarios of supply and demand, adjusting for known or predicted changes in both. Power system studies, especially resource adequacy analysis, require many years (ideally several decades) of chronological weather data that capture the range of potential weather variables affecting load, resource availability, and forced outages. Long records are crucial to capture the range of atypical weather combinations that produce weather-related risk, and because the number of variables increases, the range of atypical combinations that produce risk also grows and requires longer records to capture. In addition, energy-limited resources (such as storage and flexible demand) create the requirement that the weather data not only be physically consistent in space, but also accurately represent the correct chronological evolution of the weather, as this will impact how they are managed (for instance, how storage is charged and discharged). Because many power systems analysis tasks attempt to evaluate future portfolios of weather-driven generation, including determining where those generators should be built, the weather data need to be known not only at the location of current generators, but all other plausible generator locations.

## Synthesizing Weather Datasets with Models: Why Do It and Why Is It Difficult?

When available, direct observations are the most accurate way to characterize atmospheric variables. However, such an archive is not available, and it would be impractical to build, as it would require a much denser network of atmospheric measurements than currently exists, with instruments every two or three kilometers in some locations. This would be prohibitively expensive to build and maintain. In any event, it would take at least a decade of gathering observations before anything close to a representative archive would be available.

As a result, models are used to fill in the temporal and spatial gaps. These range from simple models, often developed by power systems engineers with little or no meteorological training, to highly sophisticated physics-based weather models involving millions of lines of code and running on the world's most powerful supercomputers. Some of the latest artificial intelligence methods are also starting to be deployed in conjunction with physics-based models, to reduce the enormous computational requirements of running the physical models at high spatial resolution.

Simple models are easy to understand but usually inaccurate. On the other hand, physics-based models tend to produce data that are much more accurate, but it is important to understand that synthetic data produced this way can still contain large errors even when they look realistic. In addition, expert knowledge is required to understand the inherent uncertainties in the modeling process, because the same weather model can produce vastly different output depending on how it is configured. The addition of artificial intelligence can further obfuscate how data are derived.

The atmosphere has many variables, including wind speed and direction; temperature; pressure; water vapor concentration (humidity); hydrometeor concentration, phase, and size (cloud droplets, rain drops, cloud ice, hail, snow); aerosol type, concentration, and size; incoming solar radiation; and outgoing infrared radiation. Each variable interacts with the others and responds to characteristics of the Earth's surface: altitude, slope, reflectivity, roughness, temperature, moistness, etc. The relationships among all of these factors are highly non-linear: multiple different variables influence one another, and changes may be muted or amplified in different circumstances. However, these relationships follow well-defined physical laws and are not random. This creates a dynamic, constantly changing environmental system with an almost infinite number of possible states with some variables changing rapidly over small distances.

Like the atmospheric system, the power system is also interconnected in time and space. Events occurring in one part of the system impact others and evolve in time. This is also true of the interactions between the two systems (e.g., a change in wind speed at the location of a wind generator affects the evolution of the weather elsewhere *and* changes the electricity system state). Therefore, each state has a specific impact on supply, demand, and other weather-influenced components of the electricity system.

Therefore, models that synthesize data for use in power system analysis ideally should capture the physical and dynamical relationships between weather variables and produce weather states that are physically plausible, evolve realistically in time and space, and produce distributions of conditions like those that are observed. A primary motivation of this section is to help users of synthetic data better understand how difficult this is and communicate the limitations these challenges often confer onto synthetic data.

**Models that synthesize data for use in power systems analysis ideally should capture the physical and dynamical relationships between weather variables and produce weather states that are physically plausible, evolve realistically in time and space, and produce distributions of conditions like those that are observed.**

### Importance of Understanding the Types and Sources of Data Uncertainties

The difference between an observation and reality ("truth") is mainly a function of the measurement uncertainty of the instrument used to take the observation. However, the difference between synthetic weather data and truth, in addition to being subject to uncertainties in all the observations used in the modeling process, is mostly a function of the modeling method. Therefore, synthetic weather data have much more inherent uncertainty than weather observations. This is intuitive to most users when simple models are used, but it is also true of data that are synthesized by complex numerical weather prediction (NWP) methods, including reanalysis and reforecast datasets (discussed in detail shortly), which are widely used. While these methods use observations as inputs and produce detailed outputs with realistic weather patterns that reflect the input observations, the uncertainty of model output data is *not* similar to that of direct meteorological observations. In addition, the uncertainty in synthetic data produced using physics-based models is not uniform in time and space, between different weather regimes and geographies, or for different configurations of the same model. At the time of writing, this is not fully understood even by many savvy power system modelers, and it is almost never acknowledged in reports communicating the analysis of power system modeling that utilizes these inputs.

**The uncertainty in synthetic data produced using physics-based models is not uniform in time and space, between different weather regimes and geographies, or for different configurations of the same model.**

Furthermore, it must be remembered that few synthetic model data have been robustly validated against observations, in large part because in many cases such validation is not possible because the modeling was performed specifically to fill gaps where observations were unavailable. It is not correct to assume that if model output is similar to an available observation in one part of the model domain, that output in other parts of the domain will also be accurate.

**Regardless of the source of synthetic weather data, validation and uncertainty quantification are essential steps to ensure that invalid conclusions are not drawn from studies that utilize synthetic weather inputs.**

This is not to say that synthetic weather data are not useful. When model configurations are thoughtfully designed to produce output for use in subsequent power system modeling, it is possible to produce valuable data. However, it must be understood that these data have much more inherent uncertainty than those coming from weather observations. As a result, regardless of the source of synthetic weather data, validation and uncertainty quantification are essential steps to ensure that invalid conclusions are not drawn from studies that utilize synthetic weather inputs.

The following discussion gives a basic description of the different sources of weather data for power system modeling. It is designed to help non-meteorologists make the best use of the guidance given in the rest of this report and be able to intelligently use weather inputs in power system modeling. This discussion covers different types of weather observations, ways in which weather data can be synthesized using models, and the pros and cons of different approaches. It describes how model implementation and configuration can impact the output data, explains why this might matter for different applications, and discusses the importance of validating model data. More detailed information on these subjects can be found in Appendix B.

## Weather Observations

Weather observations are usually the most precise way to quantify atmospheric conditions and should be used wherever practical; however, observations at the necessary locations or across the required time period often are not available for assessing weather impacts on the electricity system. Therefore, observations are usually used to validate and determine the uncertainty of model data and/or to bias-correct the data by identifying systematic relationships between model output and truth.

Weather observations are recorded by instruments that measure quantities such as temperature, pressure, humidity, wind, precipitation type; cloud type, level, and coverage; and visibility. For in situ observations, the measurement device is physically located where the observation is taken, and for remotely sensed observations, the instrument is physically removed from the locations being observed, such as on orbiting satellites.

In situ measurements are appealing, as their uncertainty and quality is usually easy to quantify and the instruments are cheap relative to remote-sensing devices; however, their spatial coverage is limited and tends to be clustered around population centers. Examples of in situ instruments include thermometers, anemometers, precipitation gauges, and barometers. In situ temperature measurements, typically taken at airports, have historically been the primary dataset used to assess the impact of temperature on electricity demand. Such records typically span several decades, and sometimes more than a century. In situ measurements are also taken at wind and solar facilities, but these records are much shorter, and the data typically are not available for input into power system modeling applications (see Box 2, p. 18).

BOX 2

### Observations Made at Existing Renewable Resource Facilities

Most existing renewable resource facilities in the U.S. are equipped with instruments to collect meteorological data, and observation archives for these facilities would be very valuable for validating and bias-correcting model data. These facilities' data collection is done in part because the Federal Energy Regulatory Commission (FERC) Order 764 requires that transmission operators be provided with temperature, wind speed and direction, and atmospheric pressure from each wind generation facility on their systems and be provided with temperature, atmospheric pressure, and irradiance from each solar generation facility, to aid in power generation predictions used in system operations. Many other countries have similar requirements. However, these data are usually not made public and so cannot be used for power system modeling studies. One result of this is the paradoxical situation where reconstruction of past generation estimates for planning is less exact than forecasting of future generation for operations.

**The only current way to produce the required data is to use models. However, broad access to weather archives for existing weather-driven power plants is necessary to validate and bias-correct model data.**

For analyses of electricity systems for planning studies, multi-decadal records are needed covering all possible current and future generation sites. Because most renewable generation facilities have been operational for under a decade, even if they were available, observational records are not long enough to be able to fully capture the distribution of weather-driven generation outcomes. In addition, data at operational plants are not always a good proxy for generation at future plants more than a few miles away. Thus, the only current way to produce the required data is to use models. However, broad access to weather archives for existing weather-driven power plants is necessary to validate and bias-correct model data. Further, access to power and availability archives would allow much better generation estimates to be produced from model-synthesized weather data. For these reasons, the project team strongly recommends policy changes to improve overall access to observation archives for existing weather-driven power plants.

Remotely sensed data are a crucial input to models that are commonly used to produce datasets for power system analysis today and going forward. Remote-sensing instruments either observe atmospheric data from somewhere distant from the measurement location (known as passive sensing) or send out a signal and observe the interaction of the signal with the atmosphere (known as active sensing). Remote-sensing instruments can gather data from large areas or volumes either by having a wide field of view or by scanning. Examples are cameras (passive sensors) and weather radars (active sensors). Modern remote sensing has revolutionized NWP, which requires the best possible estimate of the state of the atmosphere to forecast future states and/or synthesize a more detailed picture of the weather than is available from observations alone. More details about in situ and remote-sensing observations can be found in Appendix B.

While observations are typically the most reliable measure of atmospheric conditions, they have major drawbacks.

- Observations are typically spatially sparse and often located in places that are not representative of the important meteorological properties driving supply and demand across a region. For example, many surface observing stations are located at airports, and those in populated areas tend to be the best maintained. This means the temperature data may be useful for developing relationships with load, but wind and solar data are unlikely to be representative of remote regional wind and solar plants.
- The instruments used by different observing networks are of vastly different quality and are maintained to different standards; quality control can be very tedious. One should not assume that one temperature, wind, or other measurement is as accurate as the next.
- Remotely sensed data are often voluminous and complex. They may require expert processing and interpretation, and measurements

are often not uniformly organized in time and space. The sensing devices are typically expensive.

- Data discontinuities and biases can result from instrument updates, updated instrument calibrations, station relocation, and even environmental changes around the observation (e.g., new buildings or increased shading by trees).

## Model Data

As noted above, the network of observations is insufficient to provide a representative view of generation potential for current and future renewables, so the observations we have need to be augmented with model-synthesized data. This section explores the limitations of simple models and of the sophisticated NWP and machine learning methods used to produce more comprehensive datasets. More detailed information about NWP can be found in Appendix B.

### Modeling the Atmosphere's Complex Behavior

Because of the complex nature of the atmosphere, simple statistical models using variable(s) observed at one site (for example, temperature) are rarely able to estimate other variables at the same site, let alone at other locations. Any suggestion that such modeling is possible should be viewed with deep skepticism in all but the simplest cases. However, while the atmosphere is complex, its evolution in time and space does follow well-defined physical laws related to conservation of energy, momentum, and mass, and these laws can be described with mathematical equations. Solving these equations is the basis of physics-based modeling, which is widely used to produce synthesized weather data for a range of uses, including power system analysis.

In some cases, such as in the production of irradiance data for the National Solar Radiation Database (NSRDB), models are diagnostic and use observational data to infer (diagnose) an estimate for the value of a related quantity. An everyday example of a diagnostic model is seen in a mercury thermometer. The thermometer measures the expansion of mercury, and the diagnostic model converts this to temperature. But most physics-based models used to synthesize atmospheric data are prognostic: if the state of the atmosphere is known at many locations, such models can estimate the state of the atmosphere at other locations and other times. This is the realm of NWP models, commonly known as weather models or weather forecast models. While predicting the future state of the atmosphere is the most familiar use of NWP to the public, NWP models can also be used together with observations to estimate a denser array of historical meteorological data than is available from observations alone. (See Box 1, p. 11, for the three definitions of "forecast" used in this report.) NWP enables the production of datasets that are representative of the distribution of past weather conditions concurrently impacting wind, solar, and load and that capture the chronological evolution of these conditions in a realistic way.

Data produced by NWP models adhere to the physical laws governing atmospheric motions and processes and are produced on convenient regular geographic grids, with even temporal spacing. The distribution of each variable in time and space and its relationship to every other variable is consistent with these laws, which is important for producing chronological time series data of variables that represent the evolution of plausible weather scenarios. This means the data meet many of the requirements for use in modern power system modeling where wind and solar generation is broadly dispersed. However, as discussed below, while NWP models can provide reasonable estimates, even they are far from perfect, and their output should not be viewed as a near-perfect representation of truth. Many factors associated with the input data and model configuration affect these models' output model, which can deviate significantly from reality.

**While NWP models can provide reasonable estimates, even they are far from perfect, and their output should not be viewed as a near-perfect representation of truth. Many factors associated with the input data and model configuration affect these models' output model, which can deviate significantly from reality.**

## Simple Statistical Models

An estimate of the meteorological conditions at a particular location is often needed because there are gaps in an observing record or because the required variable is not measured. Simple statistical models are often employed to fill such gaps in the observed record by using correlations of the available observations at the site of interest and observations at a nearby location to predict the missing data. These models can be useful for minimal data filling or for extrapolating a record using data from a nearby site with a longer time series, but it is critical that their validity and uncertainty is evaluated carefully, because such models rarely capture the range of possible outcomes and can produce false and misleading data that will impact downstream analysis. One easy way to check the validity of simple statistical models is to withhold some of the data from the dataset used to establish the relationship and check how well the model predicts the withheld data.

**It is critical to carefully evaluate the validity and uncertainty of simple statistical models, because these rarely capture the range of possible outcomes and can produce false and misleading data that will impact downstream analysis.**

An example of a simple statistical model is "measure, correlate, and predict" (MCP), which is frequently used in wind resource assessment. Here, a (usually linear) correlation is developed between observations at an airport or other nearby observing location that has a long, good-quality meteorological record, and observations measured at a prospective renewable resource site. The relationship is used to put the data from the resource assessment measurement campaign into the context of the longer climate record to allow production estimates to be corrected up or down. If a good correlation can be established between the two observations, this method can work reasonably well to normalize average annual, monthly, and (with a strong correlation) daily output of a short measurement campaign (for example, two years) to the longer-term average. MCP is often applied with a long-term reference of about a decade. Because the climatological norm is considered as requiring 30 years to capture, most, but not all, of the average monthly variability can be captured in this way.

MCP-like methods are also sometimes used to synthesize load time series. Here, measurements from two sites are correlated, and a simple transfer function is developed that allows periods without observations to be estimated. Because load and temperature are typically strongly correlated, this application is typically useful if applied with careful validation.

Other models apply simple empirical rules, for example, the assumption that a constant wind shear in the lower atmosphere can be used to extrapolate the wind speed at one height using data at another height. Similar rules can estimate temperature using constant lapse rates (the change in temperature with height). Such empirical rules can be useful in some applications, but do not produce the required level of accuracy in others. Therefore, it is important for a data user to know when empirical rules have been applied and understand the nature and impact of the uncertainty introduced.

Another example of the use of statistical models is to predict the daily profile of wind and solar generation based on the temperature regime influencing load for days that occur around the same time of year; this use is problematic. For instance, the assertion may be that if a warm January day has a particular solar shape, solar generation on other warm January days for which no solar data exist will have the same daily profile. This seems intuitively compelling; however, the reality is much more complex. Cool summer days can be sunny, hot summer days can be cloudy, and, as anyone living in the U.S. Midwest knows very well, the coldest winter days are often blazingly sunny. When one includes additional coincident variables like wind speed, the situation quickly becomes complex, especially if correlations are being attempted between the conditions of two or more variables at different sites.

Statistical and empirical models like MCP typically relate one or two predictors (e.g., input variables like wind and temperature at location A) to the output variable (e.g., wind at location B) in a way that the output variable being predicted varies in a simple linear fashion with the input variable (first order, as opposed to quadratic, cubic, or higher order). These models are rough empirical approximations not representative of all the physical laws at play. They can produce apparently reasonable distributions with average errors but lead to very large errors in any given hour. This is problematic if the large error correlates with a weather condition that causes electricity system stress. Another problem with statistical and empirical models is overfitting, where a complex relationship between multiple variables is found within a sample, but validation outside of the sample shows that the apparent prediction capability is not present.

**Models that attempt to reproduce the wind and solar profiles for a given day based on predictors like temperature may appear to produce a reasonable long time series where the range of output variables looks as though it reflects reality quite well. But careful validation will usually reveal a poor match with reality, especially when one looks at coincident combinations of variables impacting load and wind and solar generation across a region.**

Models that attempt to reproduce the wind and solar profiles for a given day based on predictors like temperature may appear to produce a reasonable long time series where the range of output variables looks as though it reflects reality quite well. But careful validation will usually reveal a poor match with reality, especially when one looks at coincident combinations of variables impacting load and wind and solar generation across a region.

## Numerical Weather Prediction Models

While NWP models are best known as the basis of modern-day weather forecasts, NWP methods are also: (1) the core component in datasets utilized as weather inputs for power system modeling, and (2) used for global climate modeling designed to understand the potential consequences of anthropogenic climate change. NWP models mathematically represent the physical laws governing the weather and can be used together with observations to estimate a denser array of historical meteorological data than is available from observations alone. (See Box 1, p. 11, for three definitions of "forecast.")

However, there are many sources of uncertainty and approximation related to the data used as inputs to the NWP process and the specific model used.

NWP methods can be used to produce other types of datasets that are commonly used in power system planning, including reanalysis data[4] (defined in the footnote and described in more detail below), as well as in the high-resolution downscaling of both reanalysis data and global climate model (GCM) output. In what follows we introduce the basic principles of NWP modeling, some of the model configuration choices that have the most influence on the applicability of NWP data to power system modeling tasks, and gridded weather analysis data, reanalysis data, and downscaling. Warner (2011) provides an excellent summary of best practices for NWP modeling. Other recent works that provide useful summaries of NWP modeling for renewable resource applications include Haupt et al. (2017, 2019) and Jiménez et al. (2019).

**Power system analyses where weather risk is high and important for decision-making should probably engage a meteorologist who is well versed in NWP to explore potential pitfalls and ensure that erroneous conclusions are not drawn from downstream power system modeling results because of imperfect weather inputs.**

Power system modelers using NWP output must have a basic knowledge of things that impact the accuracy of the data they are using and the situations where larger errors might show up—the devil is in the details. Power system analyses where weather risk is high and important for decision-making should probably engage a meteorologist who is well versed in NWP to explore potential pitfalls and ensure that erroneous conclusions are not drawn from downstream power system modeling results because of imperfect weather inputs. Above all, one should always remember that garbage in will result in garbage out. Having enough knowledge to know when to question the quality of weather inputs is essential.

### Basic NWP Principles

Atmospheric processes adhere to physical laws that can be described mathematically as a system of regular and partial differential equations. If we perfectly describe these laws mathematically and we perfectly know the state of the entire atmospheric system at a given time, then we can, in theory, determine the entire atmospheric state at all future times. This situation is known as an initial value problem. NWP is the branch of atmospheric science dedicated to determining the initial value as accurately as possible and solving the initial value problem for subsequent times by representing as closely as possible the physical laws governing the motions and processes that are occurring, using mathematical equations that can be solved using numerical methods. NWP models are physics-based models (sometimes referred to as physical models) that perform this modeling on computers. They require extensive and accurate data inputs (the initial value) and apply these inputs to the physics-based equations to model the atmosphere, including the development and decay of weather systems and their movement across a geographical area. NWP models can either be run over the entire globe or over a particular region of interest.

By discretizing the three-dimensional model domain into grid cells (i.e., grid volumes),[5] NWP models represent and predict values for numerous variables (including temperature, wind speed, and solar irradiance) at every grid cell in the domain, regardless of whether or not an observation exists for that grid cell. Because the modeling is physically based, where interpolation/extrapolation of observations leads to initial conditions that are not consistent with the physics of the system in some locations (typically due to a lack of

4 A weather analysis is a process that takes available weather data and uses it together with knowledge of the laws of physics to estimate the state of the atmosphere and is the first step in the forecasting process. It can be done manually, but today it is typically done using computer codes. Reanalysis is the term used for a similar process that occurs after the fact when all of the possible data are available, including what would have been the future state of the atmosphere. Through the use of sophisticated computer codes, reanalysis reconciles all the data from observations and past, current, and future model estimates in an effort to produce the most accurate weather analysis possible.

5 NWP models track the state of the atmosphere at a finite number of grid points. The closer together these grid points are in the horizontal and vertical, the higher the resolution of the model. It takes at least three grid points to represent a simple feature on the Earth's surface or in the atmosphere. An intuitive example of the representation of terrain is that a V-shaped valley requires three grid points to resolve, and a U-shaped valley requires four, so if the grid points are 1 km apart, the smallest valley that can be represented is 2 km wide. All features below this scale are not explicitly resolved.

data), the model will tend to evolve the fields to remove the physical imbalance; this adjustment process will usually produce a more accurate representation of the atmosphere than was available by simple interpolation of available observations. This is a powerful feature of NWP that is particularly useful in regions of complex topography where fields may vary rapidly with distance and observations are sparse.

NWP models can be run at different grid spacing in both the horizontal and the vertical, which determines the granularity (or resolution) of the geography and attendant physical processes that the model can simulate. A high-resolution grid is critical for power system studies so that the weather impacting existing and potential future wind, solar, and other plants can be accurately determined, along with concurrent weather impacting load. These weather data can then be used in power system models to evaluate how weather will affect the concurrent performance of these resources and loads on the power grid so that studies can identify potential points of weather-driven reliability risk.

It is reasonably intuitive that we cannot measure the state of the atmosphere perfectly even at one location (due to measurement uncertainty), let alone everywhere. In addition, we do not have the computer power necessary to represent every turbulent eddy or cloud droplet explicitly even if these details could be measured. Moreover, numerical methods are inherently approximations because they deal with finite differences versus the infinitesimal differences of pure calculus. Therefore, perfectly predicting

weather variables at any given time or place is not possible. In addition to these limitations in our ability to model the atmosphere, the laws governing the atmosphere's behavior involve non-linear interactions among many variables. Systems like this are highly sensitive to small changes in the initial conditions, and their behavior is inherently chaotic. Small perturbations in the initial state ultimately result in large differences in the future state. The metaphor that a butterfly flapping its wings in Africa can affect the development and path of a hurricane in North America is apt.[6]

The amount of time a modeled system remains predictable depends on how accurately the initial state is measured, the dynamics in the system, and the length scales of interest. Therefore, since measurements can never be performed everywhere or with perfect accuracy, and since those observations cannot be perfectly represented by analytical functions, even with infinite computer resources, there are fundamental limits to the accuracy of the predictions that NWP models can make. That is, while the data are useful, they are imperfect, and these imperfections must be quantified and considered when the data are used as an input to power system modeling. Predictability depends on the scale of the weather features of interest, on the order of minutes for small-scale phenomena a few meters across (such as dust devils), to a few weeks for the planetary

**In addition to the limitations in our ability to model the atmosphere, the laws governing the atmosphere's behavior involve non-linear interactions among many variables. Such systems are highly sensitive to small changes in the initial conditions, and their behavior is inherently chaotic. Small perturbations in the initial state result in large differences in the future state.**

6 The atmospheric system was where chaotic systems were first explored. Edward Lorenz showed in his famous 1972 talk, "Predictability: Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?," that for such a system, while the exact present determines the future, the approximate present does not approximately determine the future (Lorenz, 1972).

waves[7] encircling the Earth that are thousands of kilometers across and drive large-scale weather systems (Judt, 2018, 2020).

Figure 2 provides a simplified representation of how atmospheric data are represented in an NWP model and the process of iteratively running such a model. All NWP modeling starts with an initial condition that is a three-dimensional representation of the atmosphere. The initial condition is produced by taking a first guess of the atmospheric state (also known as the background field) from a prior model run (usually a short-range prediction of one, three, six, or twelve hours) and adjusting it using as many sources of observational weather data as possible, including surface observations, balloon soundings, radar data, ground- and space-based remote-sensed information, and aircraft data. This is a complex process that incorporates the observations into the model in a way that considers both model and observational uncertainty and produces an initial condition that is physically consistent with the model topography. For regional NWP models, lateral boundary conditions must also be specified at regular intervals (typically every one to six hours) for the entire duration of the simulation, from either a global model or a larger regional NWP model. These lateral boundary conditions are another source of model error; eventually this "boundary creep" can contaminate results throughout the domain. See Warner, Peterson, and Treadon (1997) for more in-depth treatment of lateral boundary conditions for NWP modeling.

After weather observations are assimilated (data assimilation to be explained further below), the atmospheric state at the next time step is determined by numerically solving

FIGURE 2

The NWP Cycle and Representation of Atmospheric Data on a Model Grid

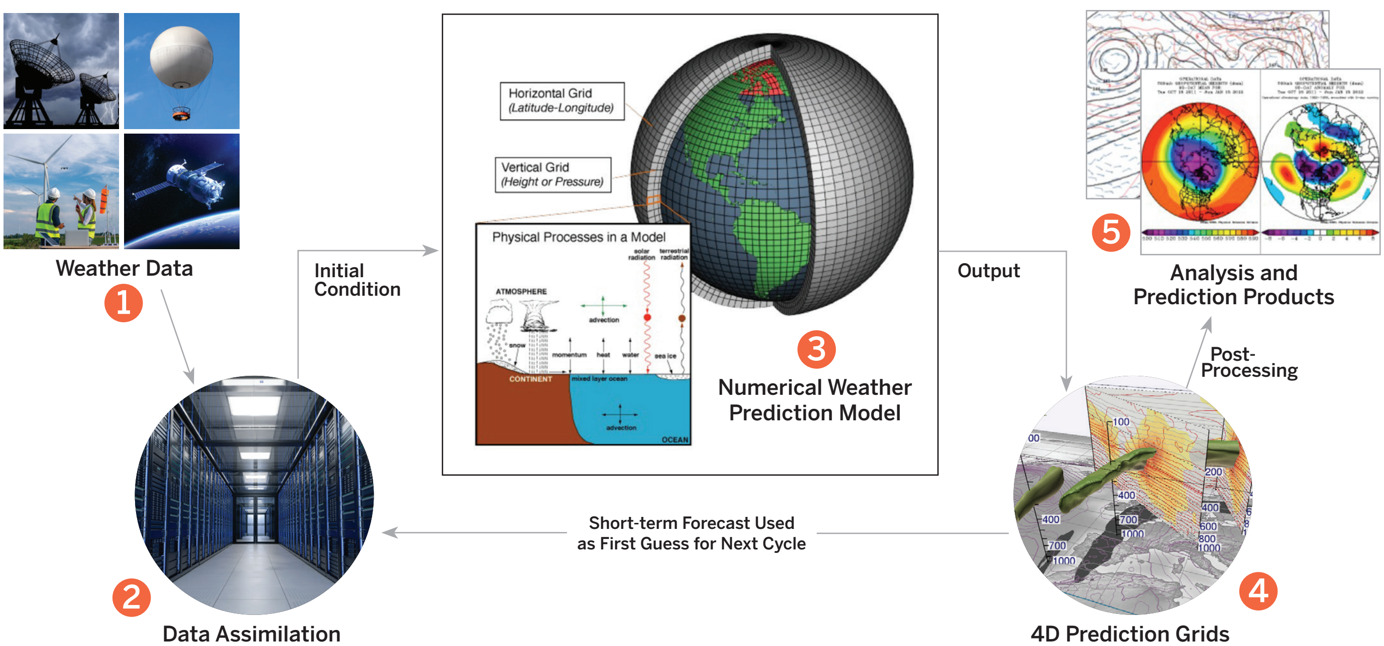



**Illustration of the cyclical NWP process. Gridded weather data output from a prior NWP iteration becomes the background field (or first guess) to the next iteration. This first guess is then nudged toward observations, while keeping it consistent with differences between how the model configuration represents the physical world. The NWP calculations are then performed and the result post-processed according to the use case, while a short-range forecast feeds the next cycle.**

Source: Justin Sharp.

7 Planetary waves (also known as Rossby waves) can be thought of as broad undulations in the jet stream, and they drive the large-scale weather patterns (periods of storminess and quiescence). There are typically four to eight waves (ridges and troughs) spanning the globe. They result from the rotation of the Earth and are modified by temperature gradients as well as interactions with surface features and other processes that move energy around.


the governing equations at each grid point. This process is repeated until the user-configured model end time is reached. Gridded model output is written to a file at regular intervals.[8] The prediction accuracy depends on the accuracy of the initial condition, how many time step iterations are performed, and how accurately the atmospheric conditions can be represented in the model. The latter is a function of the model resolution (for example, a fair weather cumulus cloud that is 500 m across cannot be represented in a model with 10 km grid spacing) and of how well the physical processes can be represented and solved in computer codes, which itself is a function of the accuracy of the numerical methods used to solve the governing equations and whether those equations can even be represented at the scales being modeled.

## The Impact of Model Resolution

Understanding the importance of model resolution is crucial, as small-scale features can have a strong impact on the weather that drives wind generation, solar generation, and load. Static features in the real world—such as steep valleys or sharp transitions from forest to grassland or ocean—that occur at scales smaller than the grid spacing will not be accurately represented in the model. Consequently, the effects of these features will be represented inaccurately or not at all. Similarly, fine-scale weather phenomena like sharp warm or cold fronts or small thunderstorm cells will be different in model space than in reality. Differences between model data and reality are particularly important to consider in regions with complex (i.e., hilly or mountainous) topography. This is because the smaller-scale weather phenomena are a projection of the larger-scale weather pattern onto the topography and associated surface characteristics (like gaps, passes, slopes, and roughness). Therefore, where model topography is considerably different from actual topography, even if the large-scale weather pattern is correctly modeled, the projection of it onto the smaller scale will be consistently incorrect, and modeled values may be very different from those of reality.

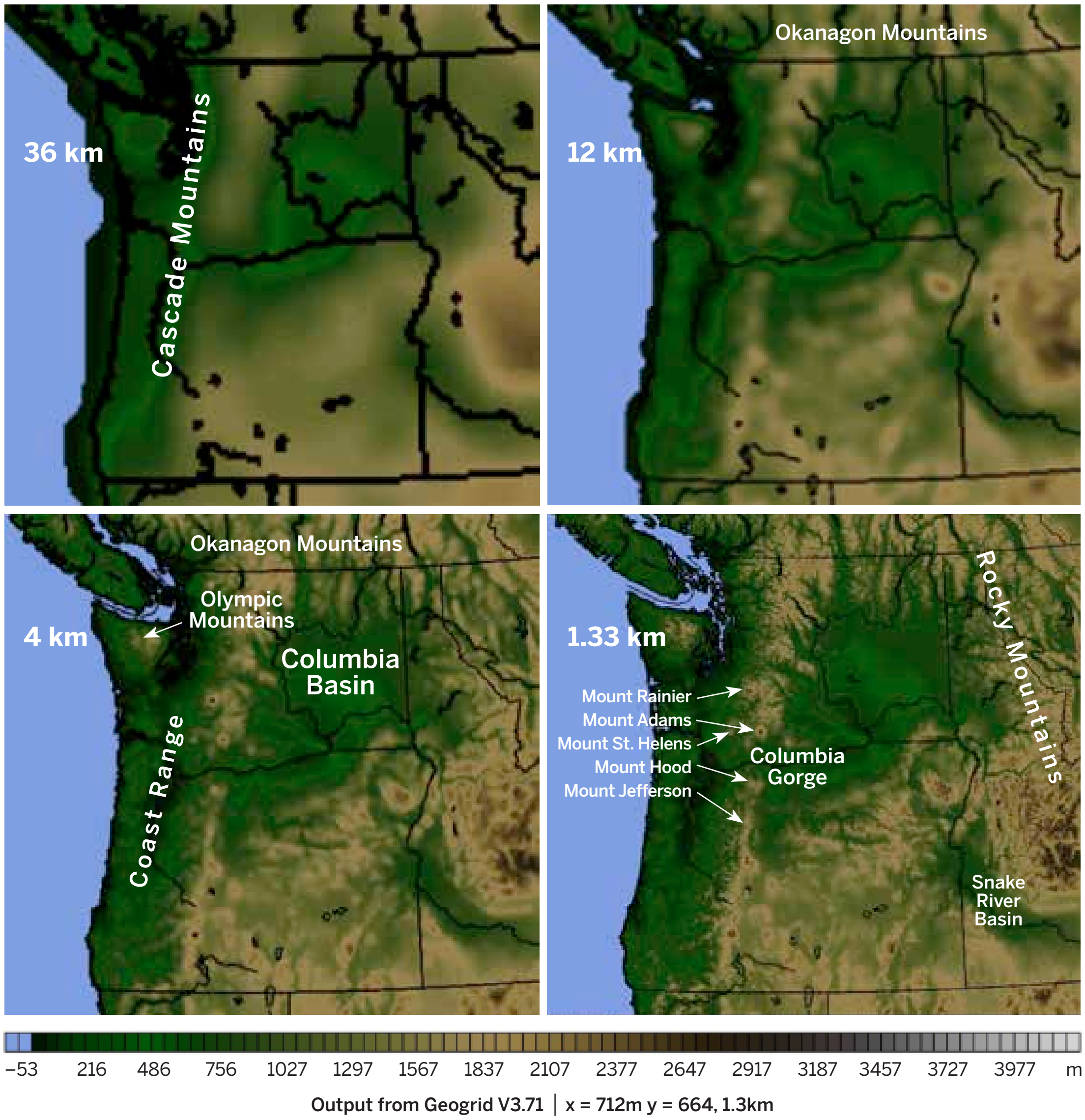
> **Where model topography is considerably different from actual topography, even if the large-scale weather pattern is correctly modeled, the projection of it onto the smaller scale will be consistently incorrect, and modeled values may be very different from those of reality.**

**Horizontal resolution.** Figure 3 (p. 26) provides a vivid example of the impact of model resolution on model topography. This poorly represented terrain in turn profoundly affects how local-scale weather features such as the flow through mountain gaps (known as gap flows), sea breezes, and mountain-valley circulation evolve in the NWP model in response to larger-scale weather systems. In the western U.S., these phenomena drive the power output of many gigawatts of wind energy facilities, and areas of clouds and clearing associated with mountain ridges could impact vast swaths of solar generation, especially in the future.

8 A common misconception is that the integration time step is the same as the output time interval. This is rarely true, although the integration time step represents the minimum output interval. The integration time step is a function of the model spatial resolution (with higher resolution requiring a short time step) and is usually a few seconds to a few minutes. The output interval just defines how frequently the atmospheric state is archived.

FIGURE 3

## Model Representation of U.S. Pacific Northwest Topography at Different Grid Space Resolutions




**Topography represented in the four progressively finer-scale domains used for the University of Washington's Department of Atmospheric Sciences's operational NWP model. The four domains have a grid spacing of 36 km (top left), 12 km (top right), 4 km (bottom left), and 1.33 km (bottom right).**

Source: University of Washington. Available at the web page Pacific Northwest Mesoscale Model Weather Forecasts: Information (https://a.atmos.washington.edu/wrfrt/info.html).

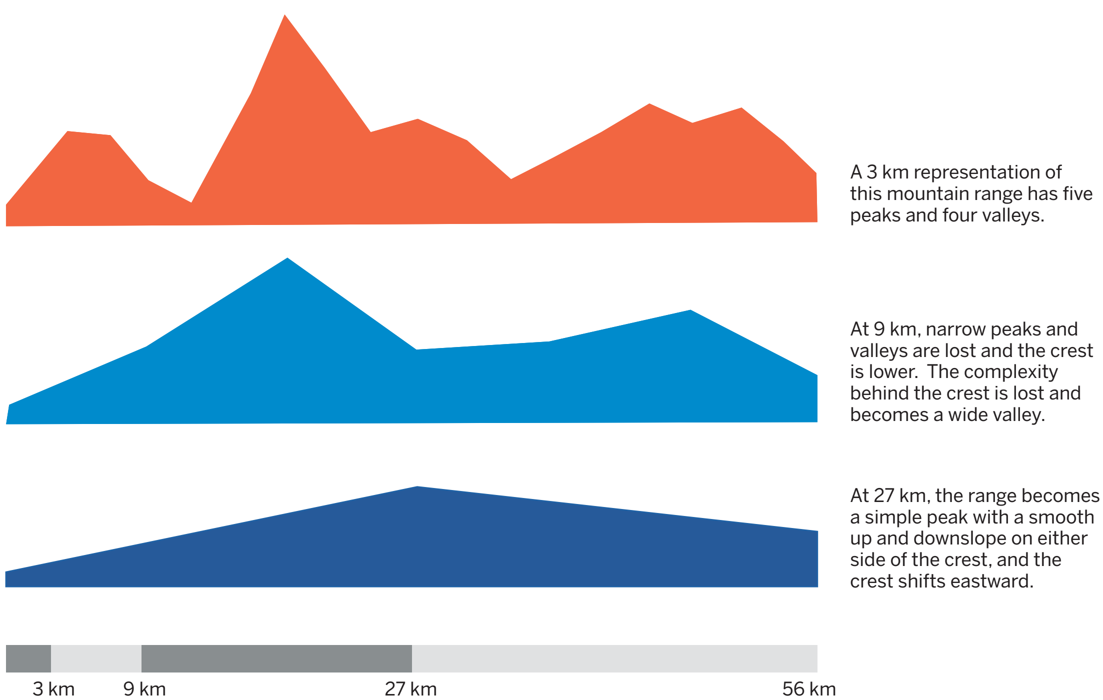
Figure 3 shows the model representation of topography in the Pacific Northwest at horizontal grid spacing of 36 km, 12 km, 4 km, and 1.33 km. This includes the Columbia Gorge, where the actual elevation is less than 100 m at river level, with steep sidewalls rising rapidly to the crest of the Cascade Mountains at a height of over 1000 m. Mount Hood (3429 m) and Mount Adams (3743 m) lie to the south and north of the gap. Several other large volcanoes are located in this region, as are several mountain ranges and a large inland basin. The key message here is that at low resolutions, many of the topographic features like tall mountains, steep canyons, and river drainages are not properly represented; thus, the weather they drive in reality will diverge from the weather that develops in the model. For comparison, a configuration with a grid spacing of 1.33 km has 784 grid points and 729 grid cells within the same geographic area as a single 36 km grid cell represented by four corner points.

At a grid spacing of 36 km (which is close to the resolution of the frequently used ERA5 reanalysis data-set, discussed below), the gross features of the terrain are present, including smoothed versions of the mountains and rivers; however, tall volcanoes and mountain ranges are barely captured. At 12 km grid spacing, the largest peaks can be seen as only smooth areas of high terrain, and, similarly, low passes appear as smooth valleys. Details of the Columbia Gorge and the coastal range can be seen. At 4 km grid spacing, most of the major mountain gaps, tall mountains, and valleys in the main mountain ranges can be seen, and the important lowlands are represented as being near sea level as in reality. It is not until a 1.33 km grid spacing is used that the Columbia Gorge is resolved accurately. Resolving the Gorge is crucial to correctly predicting the wind generation from the large number of wind farms at its eastern terminus.

Figure 4 shows hypothetical cross-sections through terrain similar to that in Figure 3. Using 3 km, 9 km, and 27 km grid spacing, it illustrates the profound differences in surface elevation and terrain features at different resolutions. The divergence between each model resolution and reality affects elevation-dependent values such as surface temperature and precipitation phase, but more importantly, model resolution affects how meteorological phenomena like cold pools, downslope winds, and upslope precipitation evolve in the model.

FIGURE 4

### Hypothetical Cross Sections Showing Model Representations of a Complex Topography at Different Grid Spacing




**The top plot shows a cross-section of hypothetical complex topography represented at 3 km grid spacing. The middle plot uses the average of sets of three 3 km points for each 9 km point. In the bottom plot, three 9 km points were averaged to get to each 27 km point.**

Source: Justin Sharp.

Model resolution has similar impacts on land surface characteristics with the placement of urban, forest, farm, and desert areas all becoming progressively more accurate as resolution increases. Each type of land surface has different values defining properties like surface roughness, albedo, emissivity, and heat capacity, properties that have a profound influence on how the real and modeled atmosphere respond to the surface.[9] For example, surface roughness dramatically impacts the low-level wind speed, wind shear, and turbulence. Albedo, emissivity, and heat capacity change the rates of surface heating and cooling, affecting low-level temperature and therefore mixing of the low-level air, which in turn impacts the vertical distribution of near-surface wind speed, temperature, and humidity.

**While it is typically understood that lower-resolution models will not properly predict the details of air flow in complex topography, it is often mistakenly believed that these models will predict the broad features of the flow and that this output can then be statistically corrected. However, if the model topography cannot properly support conditions that cause a phenomenon, the phenomenon may be absent altogether from model output.**

While it is typically understood that lower-resolution models will not properly predict the details of air flow in complex topography, it is often mistakenly believed that these models will predict the broad features of the flow and that this output can then be statistically corrected. However, if the model topography cannot properly support conditions that cause a phenomenon, the phenomenon may be absent altogether from model output. For example, in Figure 4 there are no valleys whatsoever in the 27 km resolution cross-section; therefore, it is impossible for the model to create the valley inversion (and the calm conditions that come with it)

that will sometimes be present in reality.[10] In other cases, for example, where a deep valley exists in reality but the model resolution is only sufficient to represent a shallow mountain pass, the model may produce the gap winds that occur in reality but their magnitude is very muted compared to reality. Both examples directly impact the accuracy of wind generation estimates. The errors in the way the weather evolves will also propagate downstream and grow as the simulation progresses, potentially impacting other regions.

Figure 5 (p. 29) provides an illustrative example, again from the Pacific Northwest, of the impact of resolution on the wind field. Just as a large boulder in a river creates a wake behind it where the flow is slower or may even reverse, Mount Hood creates a significant wake in the atmosphere, and in the right conditions this wake can persist for tens of kilometers and impact many wind plants downstream. The top of Figure 5 shows the wind field simulated with an NWP model running at 1 km grid spacing so that it has sufficient resolution to resolve both the wake and the atmospheric waves (manifested in the figure as periodic lines of stronger wind) created by narrow ridges. Areas of stronger winds are also seen behind some slopes and associated with width changes

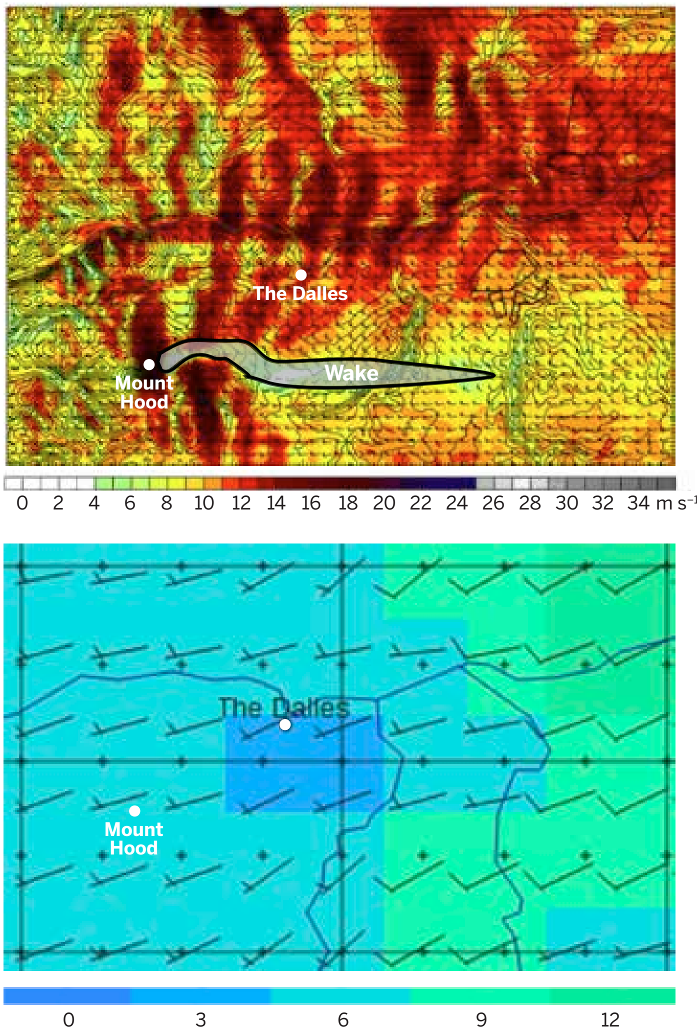
9 Albedo is the diffuse reflectivity of a surface. Surfaces with an albedo of 1 reflect all the sunlight that hits them, while those with an albedo of 0 absorb it all. Emissivity is the effectiveness of a material for emitting thermal (visible light and infrared) radiation. Heat capacity is the amount of heat supplied to a unit mass of material to achieve a unit temperature rise.

10 An inversion is an atmospheric layer where temperature increases with height. Inversions often occur in winter in basins and valleys because surface cooling on the valley sides causes cold air to drain down the valley floor. Without significant daytime heating it is difficult to remove the cold layer, as the air above it is warmer and does not mix down into it.

in the Columbia Gorge. These structures were first indicated in high-resolution NWP simulations like this one and confirmed to exist through analysis of turbine winds.[11] Their presence was subsequently evaluated in detail in the Wind Forecast Improvement Project Part 2 (WFIP2) field campaign (Draxl et al., 2021).

FIGURE 5

### Wakes and Waves Observable in a 1 km, But Not 30 km, Simulation of the Columbia Gorge in the U.S. Pacific Northwest




**The output from a 1 km Weather Research and Forecasting (WRF) simulation (top) clearly shows mountain wake and wave activity to the east of Mount Hood, whereas the output from the 30 km ERA5 dataset (bottom) for the same hour in April 2010 does not show this activity.**

Sources: Iberdrola Renewables (top), and Sharply Focused with data from the European Center for Medium-Range Weather Forecasting (bottom).

However, the bottom of Figure 5 shows the output resulting from a much lower grid spacing of about 30 km. The difference is dramatic, because simulations at this resolution cannot capture the detailed structures seen in the 1 km simulation. The topography that causes the phenomena does not exist at this resolution, and the grid spacing is insufficient to represent the rapidly varying wind field. The waves observed in reality and in 1 km simulations create significant mixing of the lower atmosphere and impact the evolution of the airflow. Therefore, a lower-resolution model, in addition to being unable to resolve the features of the terrain, will not capture the impact these features have as the model proceeds, causing the model output to diverge significantly from reality.

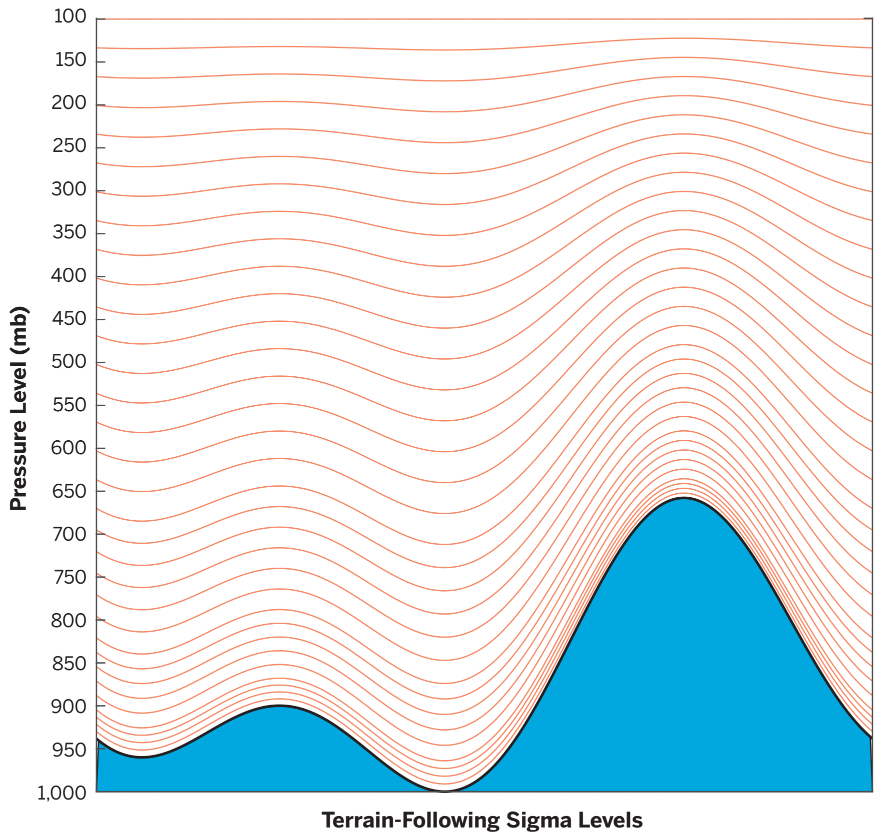
**Vertical resolution.** Vertical resolution is also important. Vertical gradients of atmospheric properties like wind speed, temperature, and humidity tend to be largest near the surface, as this is where most of the sun's energy is transferred to the atmosphere during the day and where most cooling occurs at night. The surface is also where most evaporation of water occurs and where topography and land surface characteristics have the largest impacts on weather. Therefore, higher vertical resolution is needed near the Earth's surface, but, in the interest of model efficiency, lower vertical resolution can be used higher in the atmosphere. A hybrid coordinate system is therefore used in NWP models that follows the terrain near the surface and gradually migrates toward a non-terrain-following coordinate away from the surface, as illustrated in Figure 6 (p. 30). This allows the strong surface gradients to be resolved regardless of the elevation of the terrain while reducing the resolution needed farther above the surface.

Figures 7 and 8 (p. 31) provide a schematics of a three-dimensional grid illustrating terrain-following coordinates in 3D and the high-level aspects of performing NWP (solving the forecast equations with the available computer resources). Many details in the topography, surface properties (water, grass, woodland), and weather features (like clouds) cannot be resolved at the grid spacing used (where data only exist at the intersections of the grid lines), illustrating the importance of resolution. Figure 8 also shows how the weather stations (white and red icons) do not coincide with the grid points. Subgrid-scale

11 Observed by Justin Sharp and meteorologists at Iberdrola Renewables.

**FIGURE 6**

## Illustration of a Hybrid Coordinate System Used in NWP Models




**Vertical gradients of atmospheric properties are largest near the surface, necessitating higher resolution there. However, the surface does not have constant elevation, and there is also no need to perform calculations below ground. Therefore, a hybrid coordinate system is used which follows the surface elevation at ground level and gradually relaxes with height above ground to a constant-pressure level coordinate away from terrain. The blue area is a cross-section profile of the terrain, and the bold black line references sigma = 1, which represents the model surface level. Each orange line above represents a sigma level at which properties of the atmosphere are calculated. The levels follow the terrain most closely near the ground regardless of pressure (a proxy for elevation above sea level) and in this example are closest together near the ground, which is how they are configured in actual NWP models.**

Source: Justin Sharp.

parameterization schemes are used for processes that cannot be explicitly modeled, introduced below.

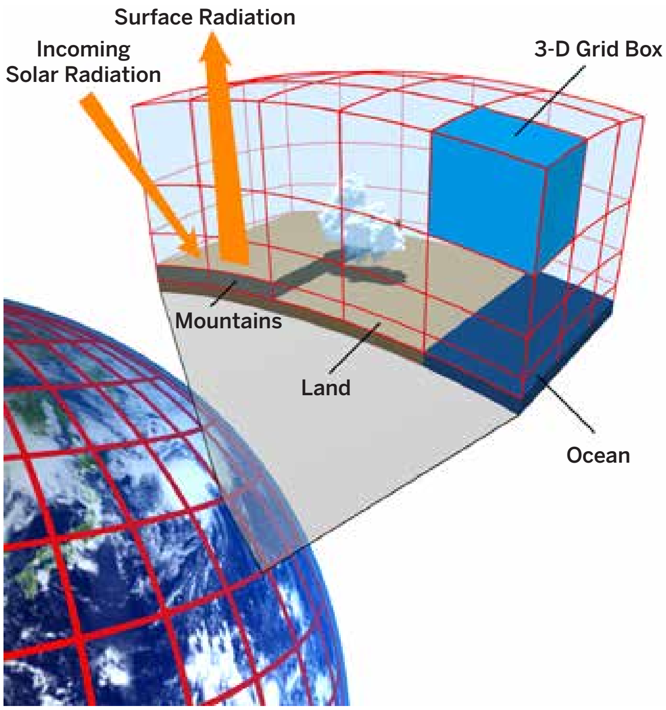
### The Impact of Parameterizations

Even as computer resources have allowed for a dramatic increase in the resolution at which NWP models can be run, there are still physical processes relevant to power system planning that cannot be modeled, as they occur at scales smaller than the grid scale of even the highest-resolution model configurations, are too poorly understood or too complex to model explicitly, or occur too rapidly. These processes that cannot be explicitly modeled must be parameterized.

Figure 9 (p. 32) shows physical processes and features that need to be parameterized by NWP. The average (or bulk) effects of these processes can be determined using reasonable statistical relationships that are based on well-validated empirical observations or using sub-model processes that, while physics-based, determine the bulk average properties within the grid cell. For instance, the physics defining how raindrops form and fall to earth is well understood and can be modeled explicitly, but modeling the condensation, growth, and coalescence of every cloud droplet and raindrop is impractical for NWP purposes. Instead, a parameterization—also known as a scheme—is used which is essentially a

## FIGURES 7 AND 8

### Illustrations of How Features are Discretized in an NWP Model Domain



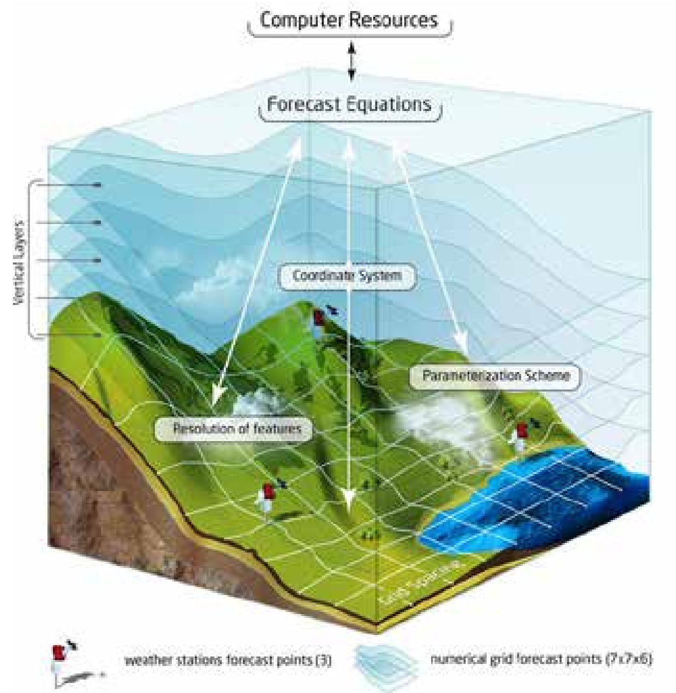




**Figure 7 shows how some or all of the planetary domain is broken down into grid cells in an NWP model, while Figure 8 zooms into a small sub-domain. This shows how model grid cells follow the terrain near the surface, how a single grid cell does not perfectly represent everything within it, and how weather observations do not typically coincide with model grid points.**

Sources: Figure 7: COMET® website at http://meted.ucar.edu/ of the University Corporation for Atmospheric Research, sponsored in part through cooperative agreement(s) with the National Oceanic and Atmospheric Administration.  Figure 8: meteoblue (https://content.meteoblue.com/en/research-education/educational-resources/weather-model-theory/model-domain)

sub-model that simulates a particular meteorological process. In the case of the formation of cloud droplets, it is known as the cloud microphysics parameterization. The scheme provides a physically sound approximation of the bulk effect of the physical processes occurring in the formation of clouds and precipitation.

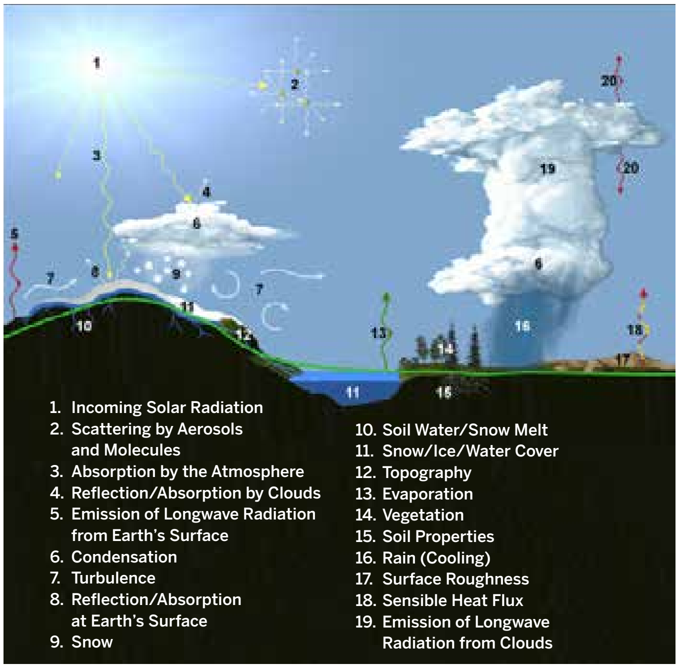
Because parameterizations are approximations, there are often several different versions that perform the same task, and each version may contain adjustable coefficients, settings, or parameters that can be tuned to make the approximation more accurate in different circumstances. For example, different schemes, and different parameter settings within a scheme, might work better in different seasons of the year, in different regions, or at different model resolutions. Sometimes schemes performing some of the different tasks in Figure 9 (p. 32) may be designed to work well together, while others should not be used concurrently. Others sacrifice accuracy in favor of lower computational overhead; this is common for operational forecasting applications where timeliness is vital. The choice of parameterizations and corresponding parameter settings within a scheme is usually based on informed experimentation and validation, and the consequences of the choices can be profound. It is important for data users to at least be aware that the choice of parameterizations can impact output biases.

Figure 10 (p. 33) shows the sensitivity of hub-height wind speeds to changes in parameter settings related to turbulence and surface roughness.[12] The model configuration is identical in both cases, including the

12 The parameter settings tune the Mellor-Yamada-Nakanishi-Niino (MYNN) planetary boundary-layer parameterization (Nakanishi and Niino, 2006) and MM5 surface-layer parameterization (Jiménez et al., 2012). The surface layer and planetary boundary-layer (PBL) parameterizations are codes that handle the complex atmospheric physics associated with exchanges between the surface and the atmosphere, including things like exchange of heat and moisture with the surface and interactions between the free atmosphere and the layer impacted by the surface.

FIGURE 9

## Commonly Parameterized Components of NWP




**A summary of the various parts of NWP modeling that are parameterized.**

Source: COMET® website at http://meted.ucar.edu/ of the University Corporation for Atmospheric Research, sponsored in part through cooperative agreement(s) with the National Oceanic and Atmospheric Administration. 

choice of parameterization schemes. The only modification is in the choice of settings for parameters related to turbulence and surface roughness. The upper plot shows line plots of wind speed for many different combinations, while the lower plot translates these wind speeds to estimates of wind generation. Note that the spread in wind speed solutions is significant, but is greatly amplified by the cubic relationship between wind speed and power output. If completely different schemes were used, versus just fine-tuning the parameters, the impacts could be even larger.

### Data Assimilation

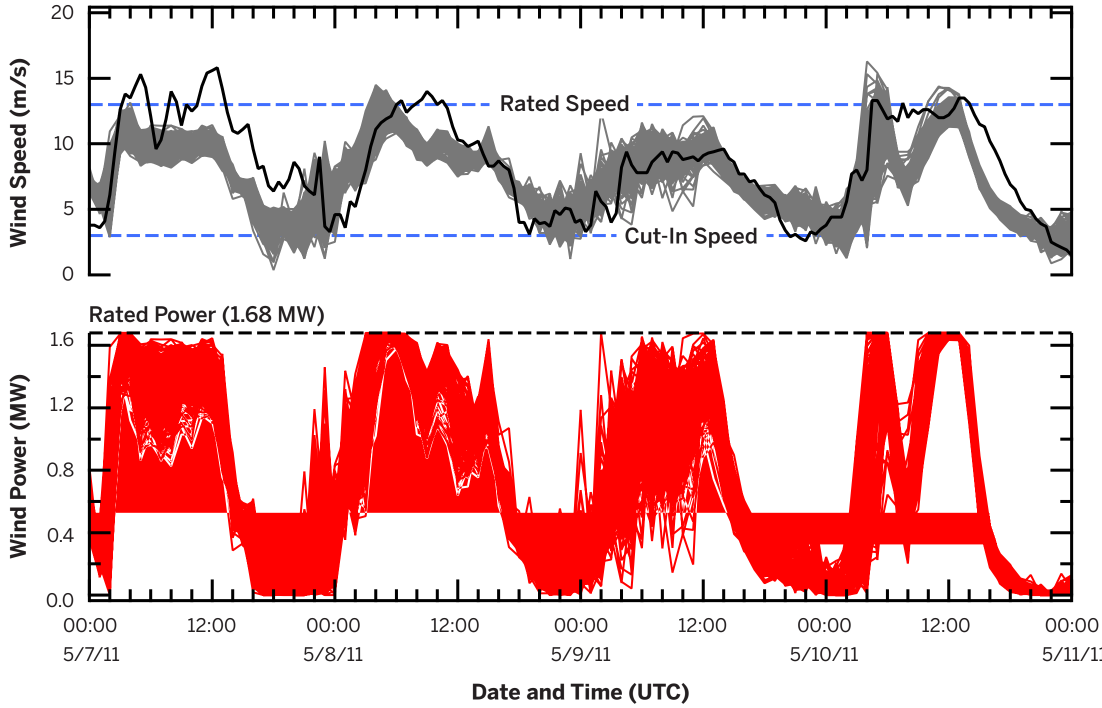
In the data assimilation component of the NWP modeling cycle (see Figure 2, p. 24), the model first-guess field—the best initial guess at the state of the atmosphere usually obtained from a prior short-range NWP forecast—is adjusted to produce an initial condition that is as close to reality as possible. This process uses observations that are collected throughout the atmosphere, including at the surface. As illustrated by the three weather stations shown in Figure 8 (p. 31), these observations typically are not collocated with the model grid points. The data assimilation process melds the observations with the first-guess field in a way that nudges the first guess toward observational truth and takes care of spreading the influence of the observation to nearby grid points, while at the same time maintaining the mathematical balance of the model's representation of the Earth's surface and atmosphere, which has less detail than reality. This is one of the most difficult-to-grasp aspects of NWP.

In short, we want the model initial condition to represent real-world conditions seen in observations as closely as possible, but at the same time it is important for the new initial condition to be as close to mathematical balance as possible in model space, and not lose important details about the state of the atmosphere that prior model runs have inferred. Models contain a less detailed representation than reality of things like terrain slope, elevation, and surface attributes like roughness, albedo, and heat capacity. These differences between model space and real space are largest close to the surface, especially where the real surface details are complex. Observations are generally more accurate than the model first guess, but the first guess contains details that have developed in the prior NWP cycle as the dynamic fields (temperature, pressure, wind speed and direction, etc.) have adjusted to the static model fields (topography, slope, land surface characteristics, etc.) in order to balance the physical equations. These details are also most important near the surface and where surface details are complex. It may be the case that, for instance, a temperature observation in a valley might be more accurate than the model first guess; however, that observation should not carry much weight because it represents a phenomenon that is at a scale the model cannot represent.

Thus, data assimilation is about much more than creating a new initial condition by interpolating available observations onto a grid: the assimilation process seeks to strike a balance between pushing initial condition features that drive weather at scales the model can represent toward observed truth, while maintaining the details that have been inferred by the model in regions where observations are sparse. In addition, assimilation accounts for differences that are due to the different level of detail the model

**FIGURE 10**

## The Different Outcomes When Using Different Parameter Settings with the Same Model Configuration




**Traces of wind speed and wind power for many different iterations of a model run with everything held constant except parameters related to turbulence and surface roughness. The upper plot shows the range of wind speeds generated by numerous model runs. The lower plot translates these wind speed differences to estimates of wind generation. The wider spread seen in the lower plot shows the profound impact of parameter choice when the cubic relationship between wind speed and power output amplifies these differences.**

Source: Yang et al. (2016).

resolution can represent relative to reality. For example, if the model surface elevation is higher than the real elevation where an observation was taken, the temperature expected in the model will be different from that observed in reality. If these differences between the model first guess and observations are naïvely pushed toward the observations, then the model initial condition may be moved far from physical balance (in model space), and, just like the real Earth system, a physics-based model will respond to remove imbalance when model integration starts. If the imbalances are large, then phenomena that are physically unrealistic (like strong winds that would not occur in reality) will develop and the model may even become unstable and cause the simulation to fail (i.e., crash).

### NWP Principles Takeaways

Power system planners often need data of a higher spatial resolution than are available from observations and need these data to be representative of the real conditions that are occurring in time and space, including how different weather variables coincide. NWP provides a way to synthesize such data. Because many of the meteorological features driving weather variables that impact supply and demand, especially those determining renewable resource generation, are driven by topography or small-scale weather features, the NWP modeling must either be conducted at sufficiently high resolution or use a post-processing method (described later in this section).

Running at higher resolution is usually the more accurate approach. However, it is not a panacea. First, even if vertical resolution is held constant, the computational resources needed to increase horizontal resolution scale by at least the third power because the number of required time steps increases by the same factor as the resolution change to keep the model computationally stable. Hence, a 1 km simulation takes at least 27 times the resources of a 3 km simulation, and takes 27,000 times the resources of a 30 km simulation. The volume of output data also expands by the power of two, as does time to output them. And while the structures in high-resolution models can look very compelling, they are difficult to validate due to the small number of observations that are available relative to the number of grid points. This is especially true in complex terrain, where the meteorological fields are most in need of validation but observations tend to be sparse. Ultimately, a compromise must be made between the benefits of higher resolution and the computational and data storage resources that are available.

**While the structures in high-resolution models can look very compelling, they are difficult to validate due to the small number of observations that are available relative to the number of grid points.**

In addition, regardless of resolution used, NWP models depend on an accurate initial condition. How good this starting point is depends on past model runs and on the amount and quality of available observations. Data assimilation takes a representation of the atmosphere produced by a prior model run and applies observations to it to produce the initial condition. This process is very complicated and is a source of significant uncertainty that varies in time and space.

Even when using high-resolution configurations, some processes that need to be modeled still occur at scales finer than the model grid scale. These are represented by subgrid-scale parameterizations. Many different choices of parameterizations and associated settings exist, and their choice in the model configuration can greatly impact the model accuracy. Further, some work better in some locations, seasons, and/or weather regimes than others.

Some grasp of these factors is necessary when utilizing data produced by NWP processes to ensure that the data are applied appropriately versus being considered as a simple proxy for observations.

### NWP Applications Relevant to Synthesizing Power System Weather Inputs

It is widely recognized that the basis for modern day weather forecasting is the regular collection and assimilation of data into NWP models and then running those models to produce a forecast of the weather expected in the coming days, and this use case is deployed to produce source data for operational load and generation forecasting. NWP models can also be used to produce estimates of weather conditions for many power system analysis tasks. This section describes the key applications of NWP modeling that are relevant to power system applications.

#### Producing Operational Forecast Data

NWP models are the foundation of all operational weather forecasting products including forecasts produced for the power sector. While the operational application is not the main focus of this report, a short description is given so that the process can be compared to how NWP is used to produce other datasets that are this report's central concern. Additionally, there are some instances in which archived operational NWP has been utilized for power systems analysis, and a few words need to be said about this.

When producing NWP output for operational forecasts of future weather, the first few forecast hours need to be produced as fast as possible, as they are only valuable if they represent a forecast of the future; if they are not produced quickly, they become an estimate of conditions in the past. This means making compromises regarding when to cut off ingestion for the data assimilation cycle so that a good enough initial condition can be produced, and the process of integrating the NWP model to produce estimates of future atmospheric conditions can begin. Choices also need to be made about model resolution and parameterizations that prioritize model speed as well as accuracy. Lastly, choices of output variables and output frequency need to be made that provide the best overall value for all end users, not just those in the power sector.

Operational forecast models are regularly updated to incorporate the latest enhancements in NWP techniques and increasing computer power. Thus, the model configuration is not static in time, meaning that output resolution, skill, and biases are not constant.

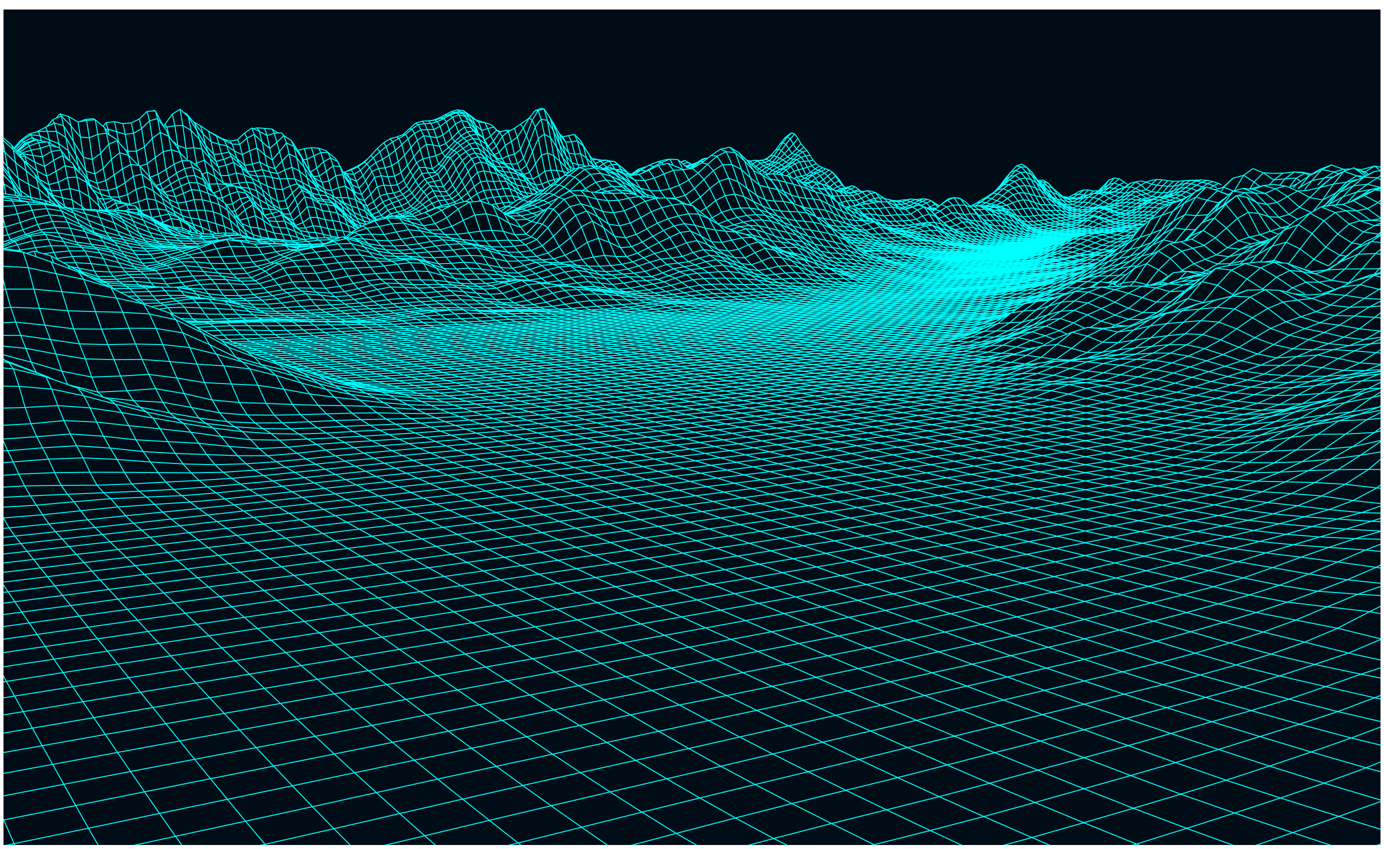
Factors like early data cut-off, configurations set up for speed, and dissemination designed to be for all generic users mean that NWP data produced for operational weather forecasting by national forecast centers (e.g., the National Oceanic and Atmospheric Administration's National Centers for Environmental Prediction (NOAA/NCEP) in the U.S.) is not ideal for use in power system modeling. For power system uses we would like weather inputs to be the best possible representation of the state of the atmosphere, with variables and output level selected to match sector needs and model configuration being held constant to prevent unexpected changes in model biases.

**Reanalysis datasets have many strengths. However, they are often misunderstood as being able to serve as a proxy for observations, and thus are often misused.**

### Producing Reanalysis Data

One of the most widely used types of atmospheric data, including for power system analysis, is reanalysis data. Reanalysis datasets have many strengths. However, they are often misunderstood as being able to serve as a proxy for observations, and thus are often misused.

Reanalysis datasets are produced by NWP modeling systems configured specifically to produce as accurate a representation of the atmosphere as possible for a given model resolution and the available input data across long periods. Unlike using NWP to produce a prediction of the future, reanalysis seeks to produce a spatially and temporally complete representation of *conditions in the past* by incorporating all useful observations and using the model physics to approximate atmospheric states



where and when no observations are available. Reanalysis data provide an easy-to-use, four-dimensional representation of the state of the atmosphere, often across the entire globe, using a single consistent method. The data are provided on regularly spaced grids, across long time periods (years, and often decades) at moderately high horizontal grid spacing (typically tens of kilometers, sometimes better) and moderately high temporal resolution (usually one- to three-hour intervals). Moreover, all the variables output at all locations are time-coincident and physically consistent: the same physical phenomena simultaneously impact every variable, and the cross-correlations between the variables are captured in the model. Another feature of reanalysis datasets is that, unlike for operational forecasts, the same configuration of the model is used to produce the entire archive, which means modeling system skill is static. All of these features are important for power system work.

However, reanalysis datasets are often misused by end users. Many believe that reanalysis data can be used as a proxy for actual observations and that they have a similar accuracy level. But it is crucial to understand that reanalysis data are an *estimate* of the state of the atmosphere across an area, not a finite point. Reanalysis data are not observations. The quality of the estimate depends on several factors including the quality of the model used to produce the reanalysis, the configuration of parameters within the models, the horizontal and vertical resolution used in the modeling, and the quality, quantity, and distribution of observations assimilated into the model. In addition, the representativeness of the reanalysis output can differ under different atmospheric conditions and in different regions. As was described in the previous section, when weather conditions are heavily influenced by phenomena at scales smaller than an NWP model configuration can resolve (phenomena too small for the model to "see"), the results can deviate substantially from the reality of the finite point where observations are measured. While the variables are all physically consistent according to the mathematical relationships governing the atmospheric system, this consistency exists at the resolution of the model and is only as good as the background field and observations going into the reanalysis *and* the ability of the model to resolve the phenomena present at the resolution used.

To begin the reanalysis, the first background field used (the first guess of the atmospheric state) comes from archived output from a high-quality operational forecast model's initial condition. For example, a reanalysis dataset beginning in 1990 would utilize an initial condition from 00 UTC January 1, 1990. All available observational data are assimilated into this analysis. Reanalysis employs the most sophisticated assimilation methods available to produce the reanalysis field. This is usually a method called 4D-Var, which considers not just how observations vary in space, but how they vary in time and space relative to the model background field and relative to short-range model predictions forward and backward in time.[13] Observations from about six hours either side of analysis time are analyzed for this purpose. The output from this process is the first interval of the reanalysis. NWP integration then moves the reanalysis state forward to the next output time, for example, 01 UTC January 1, 1990. The short-range forecast from this step then becomes the first guess for the next assimilation cycle.[14] The data assimilation cycle is then repeated with appropriate observation archives, followed by integration to the next reanalysis time. The process repeats until the entire dataset has been created. This usually takes months or years and millions of CPU (central processing unit) hours utilizing a supercomputer.

Raw model output from the reanalysis process is archived, but the data provided to users are usually processed into datasets that provide a standard set of atmospheric variables on a regular grid that is typically mapped to a sphere on constant-pressure levels. The transformation process can lead to the loss of useful resolution in the vertical and interpolation artifacts in the horizontal grid, and expert users may want to use raw grids where available. See Appendix B for details.

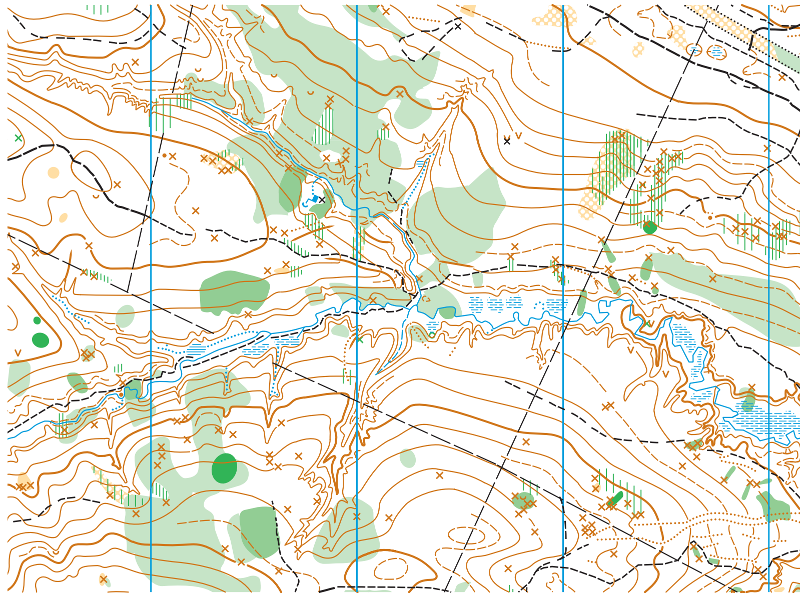
13 An accessible introduction to 4D-Var, which is used in producing ERA-5 reanalysis data—probably the most utilized reanalysis datasets for renewable energy applications—can be found at: https://www.ecmwf.int/en/about/media-centre/news/2017/20-years-4d-var-better-forecasts-through-better-use-observations.

14 The short NWP integration creates fields like accumulated precipitation that are also archived.

## Deriving Downscaled Regional Datasets

NWP and GCM models only resolve atmospheric phenomena at a scale equal to about six to eight times the grid resolution (Skamarock, 2004). For instance, a 30 km model will resolve weather features that have a length scale of 180 km or more, which is much broader than many regional-scale weather impacts. However, downscaling can be used to produce higher-resolution datasets from lower-resolution ones, although it must be used with care. Output produced by the low-resolution models can be used as input to higher-resolution NWP models in order to reproduce the atmospheric conditions present in phenomena occurring at smaller scales that are driven by larger-scale weather patterns. For example, a low-resolution GCM can produce the strong winds associated with a deep low-pressure system (a large-scale phenomenon), but it cannot translate these winds to the heavy precipitation that will result from these winds along a steep mountain range (a small-scale phenomenon), because the mountains cannot be properly represented in the GCM. Examples of other small-scale phenomena driven by the large scale include circulations like sea breezes, gap winds, mountain-valley circulations, thunderstorm cells, cloudiness on the windward side of hills and mountains, and clouds clearing on the leeward side. Some of these smaller-scale phenomena are known to affect wind energy generation, especially in regions of more complex terrain, and other small-scale phenomena impact solar generation. These phenomena typically occur at scales below those of most national operational NWP forecast models (although this is changing as computer power increases) and well below the scales resolved by best-in-class reanalyses like the European Center for Medium-Range Weather Forecasting's (ECMWF's) ERA5 and the U.S. National Aeronautics and Space Administration's (NASA's) MERRA-2 or any current GCMs. Downscaled NWP output is also produced for certain operational forecasting needs, for instance, fire weather and air quality, which require very high-resolution modeling.

Importantly, the process of downscaling can be applied to historical output, like reanalysis output for use in power system modeling, or to the output of GCMs. The best-in-class Wind Integration National Dataset (WIND) Toolkit dataset from the National Renewable Energy Laboratory (NREL) is produced this way.



When performing downscaling, the lower-resolution initial condition is first interpolated onto the higher-resolution model grid in a process that also adjusts the meteorological fields to account for the different elevations present in the higher-resolution domain. Once the NWP modeling begins, the effects on the initial field from the higher-resolution terrain will cause the meteorological fields to realign and include the impact of the finer-scale topography that causes such phenomena as channeling of the wind, forced lifting over terrain, damming of cold stable air behind narrow gaps, and differences in heating across slopes. This adjustment process is known as spin-up, and once the model is spun up, the output will represent the phenomena present at the finer scales.

Because the area being downscaled is regional versus global, the weather entering and exiting the edges of the domain needs to be provided to the model as it runs forward in time. These boundary conditions from the larger-scale analysis or forecast feed the edges of the finer-scale domain with accurate data about the larger-scale weather pattern. This keeps the fine-scale domain anchored to the larger scales that are well represented in the lower-resolution data, while at the same time allowing the model to fill in the smaller-scale effects in a physically consistent way. In some cases (where the larger-scale features are trusted), scale-selective nudging can also be used to ensure that the larger-scale features within the domain do not drift during the finer-scale forecast run. This means the model can run for longer without needing to be reinitialized. This creates fewer seams in the model output and minimizes computationally expensive spin-up time, the output from which is not useful (generally the first few hours).

An example of the power of downscaling to yield more accurate representations of the weather fields is modeling in complex topography, such as the western U.S., the Appalachian Mountains, or the European Alps. Better-resolved mountain barriers will better block cold, stable air in the models, and better-resolved steeper mountain slopes can accelerate winds more in line with reality, which can be to speeds several times larger than seen in lower-resolution models. As with all NWP output, once the model is spun up, the resultant downscaled data are physically consistent between weather variables. For instance, a sharp mountain barrier will be much taller at high resolution and thus reduce the air flow at lower levels (an impact on wind speed) across a barrier, compared to air flow modeled by a lower-resolution model. This in turn can change the temperature on the downstream side of the mountain because the air is coming from a different elevation with different atmospheric stability. At the same time, a gap or pass in the mountain barrier shown in downscaled data will be better defined and lower in elevation, also reflecting reality more closely. This will create stronger winds in its lee, and the air immediately downstream of the gap will be colder than in the original low-resolution output; it may also be drier and remove fog present in nearby locations not impacted by the gap. These more accurate representations of the weather fields will result in more accurate estimates of the wind and solar resources in the region, as well as temperature at load centers and weather-related outages at traditional generators. The more accurate representations also greatly improve estimates of precipitation that occurs in steep terrain that may feed a hydro system.

NWP downscaling is a powerful tool for providing consistent information about local effects and developing long time series at a level of detail not possible with available observations. However, it must be used with care for precisely this reason. The lack of observations means that only a small fraction of the NWP data points can be validated against ground truth,[15] so it is especially important to make sure that the model output is validated where it can be to understand how well the model is performing. It is also important to remember that since model performance will vary with weather regime, validation should be more than just calculating average errors.

**NWP downscaling is a powerful tool for providing consistent information about local effects and developing long time series at a level of detail not possible with available observations. However, it must be used with care for precisely this reason. The lack of observations means that only a small fraction of the NWP data points can be validated against ground truth.**

### Producing Global Climate Models

GCMs can produce datasets that represent weather conditions for decades into the future. Therefore, GCM output is potentially useful if one wants to simulate conditions affecting the electricity system in a future affected by climate change, although it must be understood that there are considerable uncertainties in climate predictions, and expert climatologists should be engaged to understand how large the signal is relative to the model uncertainty. GCMs are, at their core, a type of NWP model, and like other forms of NWP output, the data from these models are dynamically consistent across output fields. While the core atmospheric modeling functions of GCMs are basically the same as those of other NWP models, GCMs have tighter coupling to modeling of other aspects of the Earth system such as the cryosphere, oceans, and atmospheric chemistry (including greenhouse gas concentrations), because over long time frames, feedbacks between these systems become increasingly important. GCMs also typically use much lower resolution to make long simulations computationally tractable, although, like regular NWP models, GCM resolution is constantly improving. Using

15 Ground truth is actual wind and solar realization measured with instrumentation, as opposed to data from a model that is estimating the quantity.

a GCM, it is possible to create accurate representations of the distribution of weather over longer periods. We know this because GCMs can accurately recreate historical distributions of, for example, temperature and rainfall across broad regions. The premise of climate modeling is that if statistical descriptors of the past climate can be simulated accurately, then simulations of the future will provide insight into how those distributions change as greenhouse gas concentrations change, and the climate warms.[16]

It was noted above that the atmosphere is inherently chaotic and thus completely unpredictable at time scales beyond two to three weeks. Therefore, just like a standard NWP model, when a GCM is given a reasonable initial condition, it can accurately predict the evolution of the weather systems in this initial condition with some skill for a week or two, and, just like a standard NWP model, its prediction skill will fade beyond this horizon. How then is it possible to make predictions about Earth's future climate with GCMs? This paradox is explained by the fact that chaos theory states that within the apparent randomness of a chaotic system there are underlying patterns, feedback loops, repetition, and self-organization. The objective when running a GCM is not to predict the weather at any given time in the future, but rather to predict the distribution of future weather events that can evolve at the scales the model simulates, for different Earth system conditions (like the amount of $CO_2$ in the atmosphere).

While GCMs can potentially simulate conditions affecting the electricity system in a future affected by climate change, this matter is considerably more complicated than it first appears, and there are several important caveats to understand before considering using GCM output for this purpose. These caveats, briefly laid out below, form the basis for why this report does not focus on power system weather inputs under climate change.

Because there is no observational method to validate the predictions of a GCM in the future, the standard validation process is to use GCMs to simulate conditions over the last century or so, using the changes that are known to have occurred in the atmosphere (like increasing $CO_2$ and the oscillation of solar output through the 11-year solar sunspot cycle)[17] as a boundary condition. These simulations have been found to produce consistent and reasonably accurate results using many different GCMs (Hausfather, 2017). Once a GCM configuration is validated by showing it can produce a reasonable estimate of past climate, it is assumed that it can be used to model many future decades for different scenarios (such as changing $CO_2$ concentrations or changes in atmospheric aerosols). The results from these simulations are compared between different GCM models, and where they are similar for the same changing boundary conditions (e.g., $CO_2$ concentration), a higher degree of confidence is ascribed to the predicted distribution changes.

Almost all GCMs indicate significant future warming, and many produce patterns of temperature and precipitation changes that are similar to one another. However, there is much more uncertainty around how wind and irradiance patterns might change. Further, GCMs do not run at sufficient resolution to be able to diagnose how large-scale changes even in fields like temperature and precipitation may translate to changes at smaller scales in regions of more complex topography, which are necessary to model for power system planning. One approach to examining these smaller-scale changes is to use the GCM output as input to higher-resolution NWP models in order to downscale it as described in the previous sub-section. When this is done, there is again some consistency in results for temperature and precipitation. However, the results of downscaling exercises are mostly inconclusive when examining phenomena like local wind circulations and cloud cover.

## Post-Processing of NWP Data

NWP output and climate projections are not free of errors. Sources of error include the initial conditions used in NWP models and how they are constructed

16 GCMs are not only used for studies of anthropogenic climate change and can provide insight into changes due to any manner of slow changes in the Earth system.

17 The energy output of the sun oscillates over time in a quasi-regular and predictable way, with an average cycle from the solar minimum through solar maximum and back to the minimum of approximately 11 years.

(observations, assimilation), accompanied by boundary condition errors and model physics errors. For power system uses, even high-resolution model output can display significant deficiencies, resulting in systematic biases and less weather variability than expected. For example, even comparatively fine resolutions of NWP simulations provide an average temperature for each grid box of, say, 2 km x 2 km for local scale and 20 km x 20 km for a global scale, which can fail to reflect variability that has important impacts for both supply and demand in future power systems.

Post-processing can address some of the above deficiencies, by enhancing NWP or GCM output using simple methods such as determining and removing bias errors or performing more complex tasks for applications such as wind plant production estimates. The conversion from grid box to point estimates (point-based post-processing) or from coarse grid box to very fine grid box (grid-based post-processing) is called calibrated post-processing. Promising new machine learning methods offer an advanced form of grid-based post-processing, with the possibility of downscaling NWP output without the large computational expense of running very high-resolution NWP simulations.

#### Statistical Post-Processing

A wide range of techniques are used for post-processing. Regardless of the method used, post-processing should produce an estimate as close as possible to the truth, while respecting the climatological probabilities and producing results that are physically consistent between the different meteorological parameters. Given enough training data (i.e., observations that can be compared with NWP output), these methods can improve both spatial and temporal representation of NWP estimates, but care should be exercised because the techniques tend to smooth the data and produce outputs that underrepresent the upper and lower tails of variables like temperature, wind speed, and irradiance. In addition, the large amount of observational data needed to train them is often not available. Thus, while statistical post-processing can improve NWP output accuracy by some measures, it can also adversely impact important aspects of the original data distribution that could affect results when the data are used for tasks like resource adequacy analysis.

A simple post-processing example is bias correction in combination with a distribution correction, where one corrects the current estimate with the model's bias and distribution of errors from past estimates. For ensembles,[18] other methods based on the idea of a weather generator can be used to search for past simulations that are very close to the current forecast and use the past corresponding observation as new forecast, such as the analog ensemble (AnEn) approach (e.g., Delle Monache et al., 2013; Alessandrini et al., 2015a, 2015b; Alessandrini and McCandless, 2020). Statistical methods are relatively easy to implement and apply, once the data are available and prepared.

Machine learning and other artificial intelligence methods can also be used to improve NWP output. The link between model and observations contains non-linear relationships, which are difficult to capture with traditional statistical methods; however, using non-linear machine learning methods such as support vector machines, decision trees, and artificial neural networks, these relationships can be detected between observational data and NWP output. Once trained, machine learning methods can correct other NWP output. However,

18 An ensemble in the context of NWP is a set of NWP simulations utilizing different NWP models or configurations and/or slightly different initial conditions. The resulting sets of output can be statistically analyzed and the dispersion between them utilized to assess simulation uncertainty.

these methods can be challenging to design, need a lot of tuning and computing power, and require a significant amount of data to train on (ideally at least a year to capture all four seasons, and preferably multiple years to account for inter-annual variability).

Recent advances in machine learning are indicating that in the near future there is the possibility that some of these methods may not only be able to correct and/or downscale NWP output, but may, given enough existing NWP training data and observations, actually be able to produce better estimates by operating on low-resolution NWP output and observations than can be produced using high-resolution models. While detailing these developments is beyond the scope of this report, they are likely to become very important within the lifecycle of this document, and interested readers are referred to McGovern et al. (2019) and Lam et al. (2022) for more details.

### Generative Machine Learning for Weather and Climate Data

Recent advances in machine learning techniques for computer vision and generative models have inspired a new class of methods for the post-hoc downscaling of NWP outputs. Generative models can learn and sample virtually any conditional joint probability distribution such that they can produce realistic multivariate spatio-temporal fields given some conditional input. For example, a generative model can be trained to produce continuous gridded multivariate (e.g., wind, temperature, etc.) data-sets that are physically realistic across both space and time given a lower dimensional input such as a set of point observations or a low-resolution climate model dataset. These methods promise to reduce the burden-some computational requirements of high-resolution NWP simulations while maintaining high-quality data outputs. If these methodologies can be proven to work well, they will enable the production of higher-resolution and longer time series of weather input data suitable for power system modeling applications, as well as ensembles of these datasets that capture the uncertainty of the weather inputs and therefore allow electricity system studies to model sensitivity to this uncertainty.

Deep convolutional neural networks (CNNs) have been recently shown to excel at a wide range of computer vision tasks, including meteorological applications (Alzubaidi et al., 2021; McGovern et al., 2019). These networks are designed to the dimensionality and structure of image, video, and NWP simulation data. This results in powerful non-linear parametric models that can learn to emulate physical phenomena such as the momentum balance for wind flows on a spatio-temporal grid, much in the same way that finite-difference or finite-volume methods execute physical equations from cell to cell in NWP models. Note that this comparison between trained convolutional operators and physics-based finite-difference/finite-volume methods cannot be directly proven for large dimensional relationships such as multidimensional weather fields but can be demonstrated in simple 1D examples (Rackauckas et al., 2021), which supports the utility of these trained models in physical domains. The result is a learned model that can emulate a physical simulation similar to an NWP but at a fraction of the computational cost.

In practice, a major problem is that a basic convolutional network can exhibit regression to the mean in the form of blurring or smoothing when producing forecasts or enhancing the resolution of data. This can result in an underestimation of extremes such as heavy rainfall intensities at small spatial scales (Ayzel, Scheffer, and Heistermann, 2020). One solution to this problem is adversarial training with generative adversarial networks (GANs) (Stengel et al., 2020; Hess et al., 2022; Wang et al., 2021; Rosencrans et al., 2023; Gagne et al., 2018), where a generative model must produce data that are not only accurate but also sufficiently realistic to fool a discriminative network. That is, the generative model produces outputs that are mathematically and statistically indistinguishable from NWP outputs from the perspective of a sophisticated classification model. For downscaling data with GANs (often called "super-resolving"), the generative network is trained to produce an enhancement of the low-resolution input data that the discriminator believes is similar to real data, while simultaneously minimizing the numerical deviation from a corresponding true high-resolution dataset. This method has been shown to be effective in creating highly realistic enhancements for many types of data.

GANs with deep convolutional networks have only recently been applied to the task of downscaling NWP data, but have already shown considerable promise with high-quality physics-based validation of the outputs (Stengel et al., 2020). To the knowledge of the authors, only a small handful of public datasets have been published at the time of this writing that leverage GANs to downscale historical reanalysis data or future climate data (Buster et al., 2023; Rosencrans et al., 2023; Hess et al., 2022). However, several additional wind datasets are known to be in development that leverage GANs to do a final spatio-temporal enhancement on coarse NWP data instead of running the NWP down to the final desired resolution. The benefit of this hybrid NWP+GAN approach is a significant reduction in computational costs compared to what would be required by a full high-resolution NWP simulation (estimated at one to two orders of magnitude in compute time savings).

The main drawbacks of using GANs for downscaling are that this requires significant investment in machine learning expertise, machine learning–specific computing infrastructure, and high-quality training data, and can result in a loss of methodological interpretability including the possibility for data outputs that do not respect physical constraints. This last problem is clearly the most concerning, as low-quality data with poor physical constraints could compromise power system planners' ability to accurately predict and plan for future system needs. The methods described above have the potential to greatly benefit the renewable energy and meteorological communities, but rigorous validation needs to be of the utmost priority. Statistical benchmarking, validation against ground-truth observations, and careful examination of physical data characteristics like turbulence should all be regular practice when implementing these methods.

## Crucial Takeaways for Power Systems Modelers Using NWP and GCM Data

In summary, NWP is a complex subject with many nuances. It requires expert knowledge to determine what model resolution, parameterizations, and parameter settings are best for the problem being solved and/or the best compromise between accuracy and computational burden. When performing long simulations across broad regions, configurations that work well in one region or season may perform poorly in others. Understanding the limitations and possible pitfalls of the models' output requires deep knowledge of NWP systems. Some meteorologists without deep NWP backgrounds are not fully aware of these limitations and may recommend inappropriate usage of these models in power system planning. Even meteorologists with NWP backgrounds are sometimes unaware of how the data are being used and might recommend different approaches if they were. It is essential to have a feedback loop between power systems modelers and NWP experts when NWP data

**Even meteorologists with NWP backgrounds are sometimes unaware of how the data are being used and might recommend different approaches if they were. It is essential to have a feedback loop between power systems modelers and NWP experts when NWP data are being used for weather inputs into power system analysis.**

are being used for weather inputs into power system analysis.

Using data derived from NWP seems compelling because their regular format and general geographical and temporal completeness make them easy to use. But it is essential to understand that NWP data are not the same as observations—even data coming from reanalysis datasets that are often touted as suitable substitutions for observations. In addition, the performance of one NWP model or configuration is not a predictor of the performance of another model or even the same model used in a different region, with a different configuration, or with different input data. Even with well-chosen selections of resolution, parameterizations, and other configurable options, NWP models sometimes perform poorly. This poor performance does not occur randomly and is often related to specific atmospheric conditions and/or regions. When these factors align with weather situations that result in stress on the electricity system, the weather inputs going into power system models may be poor and compromise the results. Garbage in, garbage out.

Therefore, it is crucial that for any study using NWP data as a proxy for observations, NWP data not be utilized as a black box dataset that is equivalent to quality-controlled observations. Users need to have at least a basic understanding of how the data were produced or engage with a meteorologist who has an NWP background—and ideally an understanding of how weather data are used in power systems models—who can guide them in whether the data are appropriate for the application at hand. As part of this process, to ensure the appropriateness and accuracy of a modeled dataset for power system planning, users should review

**It is crucial that for any study using NWP data as a proxy for observations, NWP data not be utilized as a black box dataset that is equivalent to quality-controlled observations.**

a comprehensive validation report for NWP data being used that has been performed within the context of the power system modeling use case. If a comprehensive validation report is not available, such a validation should be performed. NREL recommends such validation be performed before using the WIND Toolkit data.[19] Unfortunately, such validation is uncommon, and those validations that have been performed, such as for the overall NREL NSRDB and WIND Toolkit datasets, have looked mostly at bulk average statistics for a handful of sites and have not evaluated the dataset accuracy in the context of electricity system risk periods.[20] It is important to note that these limited validations indicate significant differences between the NWP data and the ground truth; however, the results are not widely publicized. For example, a narrowly targeted simple evaluation of the WIND Toolkit data during a period of system stress in the western U.S. indicated substantial over-predictions of wind energy potential in the U.S. Pacific Northwest.[21] But such validations are not standard industry practice. This lack of validations is due in part to the limited data availability to perform thorough evaluations and in part to a lack of understanding of the need. The project team recommends the development of a best practices guide for validating weather inputs prior to use, and suggests that this be an integral part of any project that is developed to address the need for better weather input datasets.

19 https://www.osti.gov/biblio/1166659/.

20 For validations for the overall NSRDB, see https://doi.org/10.1016/j.rser.2018.03.003 and https://www.nrel.gov/docs/fy22osti/83015.pdf. For the WIND Toolkit, see https://www.nrel.gov/docs/fy15osti/61740.pdf and https://www.nrel.gov/docs/fy14osti/61714.pdf.

21 https://gridlab.org/wp-content/uploads/2022/05/GridLab_California-2030-Meteorological-Deep-Dive.pdf.

SECTION 3

# Weather Inputs Needed for System Planning

The preceding sections outlined the weather-energy nexus, provided a broad overview of the use of weather data in the electricity sector, and introduced the types of weather data that are available. This context highlights the trade-offs and pitfalls to consider when determining how to analyze weather impacts on increasingly weather-dependent power systems and what datasets to apply. The report now turns to the incorporation of meteorological impacts into system planning studies. Here, the project team describes in detail the different uses of weather data in these studies, how the data have typically been sourced and employed, how the data needs are changing in more weather-dependent systems, and the ramifications of this for the applicability of currently available data.

In Section 4 that follows, "An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data," the project team defines the characteristics of meteorological datasets that are needed for use as weather inputs for various power systems studies—the spatial and temporal resolution, the required variables, the time series length, and how data are produced, documented, managed, and made available —and gives specific guidance regarding the required and desired criteria for these attributes. We review and evaluate currently available data sources against these criteria to reveal important gaps, and, in Section 5, "Project Description for Producing Robust Weather Inputs Data," we describe a clear way forward for robustly filling these gaps. Section 6, "Guidance for Using Existing Weather Inputs," then looks at how weather inputs are currently used in power system modeling, discusses how existing datasets are applied and highlights their limitations, and offers examples of methods that may be able to mitigate the deficiencies to some degree. It summarizes do's and don'ts for accounting for the uncertainties inherent in all weather datasets. (For definitions of terms that some readers may be unfamiliar with, please see the glossary at the end of the report.)

## A Critical Need for Comprehensive Weather Datasets Targeted to Power System Modeling

Power system modeling applications are quite broad, as are various study objectives, falling generally into the categories of renewable integration studies, integrated resource plans (IRPs) or similar planning studies, and resource adequacy studies (see Box 3). All of these evaluate load, resource mix, and transmission scenario(s). While these scenarios typically represent future configurations, the data used, including the weather inputs, are usually based on conditions in the past for which measurements are available, either for direct use

**BOX 3**

### Power System Modeling Categories and Their Respective Data Needs

Weather data inputs are used extensively in power system planning, modeling, and operations. Although there are many different types of power system models, the most relevant for our purposes here are those used in the following three planning activities. At the heart of all planning, operational, and resource adequacy modeling is the requirement for the various simulations to proceed chronologically through one or more years. Planning models typically require hourly data, and operational models often use a five-minute time scale. Therefore, all renewable energy datasets need to faithfully preserve the chronology throughout the entire time period on either an hourly or five-minute time scale.

**Renewable integration studies** typically use models that simulate power system operations with various levels of renewable resources. As more renewable resources have been added to the power system and more regions in the U.S. have adopted ambitious renewable energy targets, these studies are evolving to incorporate very high levels of renewable resources and focus on how the power system could be operated—specifically to balance short-term fluctuations and uncertainty in wind and solar production—under these scenarios.

**IRPs or similar planning studies** are used in many state jurisdictions. The models used for this type of study can vary to some degree, but they most often include some type of planning/optimization model that can evaluate long-term costs and benefits of alternative resource mixes. These planning studies are sometimes augmented by more detailed operational models that require higher time resolution and more accurately simulate power system operations. Planning studies, and some operational modeling, often also include a resource adequacy assessment.

Planning models require a very large input dataset that can be used to choose the most effective combination and location of wind, solar, and other resources that are consistent with the planning objectives (optimized capacity expansion). Planning models used in renewable energy studies evaluate many alternative renewable resource build-out scenarios, performing what can be thought of as a "search" function to find the best combination of resources. This means that data for many renewable resources will be evaluated as candidate sites; hence, data for a very large number of renewable resource locations must be available for the planning models. The time resolution needed for renewable resource data for these models is hourly, and for as many years as possible (ideally three or more decades, though this is not always feasible), so as to guide the selection of the best long-term locations for renewable resource development.

**Resource adequacy studies** can be part of an IRP or carried out separately. Resource adequacy analysis typically requires hourly data and is an investigation of the ability of the power supply to reliably meet demand across a range of uncertainties. Resource availability, the probability of generators being out of service, and other factors are used to calculate one or more reliability metrics, which may include loss-of-load expectation, expected unserved energy, or heat maps that show times of expected supply risk. These studies are used to determine the total amount of resources that are needed to ensure reliability. The results of resource adequacy studies are being increasingly driven by the changing resource mix that includes more renewable resources and fewer traditional resources.

or as inputs for data synthesis. This difficulty is associated with all aspects of the power system, not only renewable generation. Sometimes the historical data are modified to model changes expected in the future, for example, to account for load growth, or changes in weather as a result of climate change.

There are many sources of uncertainty in power system modeling, especially when addressing future conditions. We do not have a good understanding of how load will evolve as transportation and heating electrify; where and in what quantity wind, solar, storage, and transmission will be built; or how much dispatchable generation will remain on the system and how flexible it will be. In addition, as already discussed, the increase in weather-driven influences on the power system makes it important to capture the inter-annual variability of weather impacts on both demand and the availability of thermal and non-thermal resources. But despite the uncertainty of future system conditions, we do have the capability to estimate, using capacity expansion models, the mix of wind, solar, transmission, dispatchable generation, and storage that meet reliability needs at least cost and to determine, using production cost models, how current and future generation mixes will perform under the full range of potential weather conditions.

Since every portion of the system is becoming increasingly interwoven—with weather conditions being the consistent linkage—the modeling efforts require quality weather data in order to obtain quality results. Modeling approaches that assume that the behavior of different system elements is independent will not properly describe the envelope of possible operating conditions, and long time series of concurrent weather data are needed. Databases are needed that include the concurrent weather variables that will impact load, wind, solar, hydro, and thermal generation, in current and future system configurations. These need to be long enough to capture weather variability and infrequent severe weather events, of high enough resolution to get a reasonable assessment of generation at any current or future renewable generation facility, and physically consistent so that the estimates for all resource types are based on the same underlying weather conditions.

**Databases need to be long enough to capture weather variability and infrequent severe weather events, of high enough resolution to get a reasonable assessment of generation at any current or future renewable generation facility, and physically consistent so that the estimates for all resource types are based on the same underlying weather conditions.**

**High-risk events do not have to be "extreme" in the classical sense to pose risks. As more wind and solar generation is added to the resource mix, combinations of moderate events, such as low winds and moderately cold temperatures, will strain the grid.**

### Ability to Capture High-Stress Weather Periods

The power system is often most at risk during periods of high-stress weather, leading to increasing interest in obtaining a better appreciation and understanding of what constitutes high-risk weather from a grid reliability perspective, as well as the impact of such weather. Analyzing these risks requires knowledge of all the coincident weather impacts. It is important to note, however, that high-risk events do not have to be "extreme" in the classical sense to pose risks. As more wind and solar generation is added to the resource mix, combinations of moderate events, such as low winds and moderately cold temperatures, will strain the grid (Novacheck et al., 2021).

There is also mounting evidence that the frequency and intensity of certain types of extreme weather have already increased (e.g., extreme heat, droughts, and heavy precipitation and environmental conditions they affect, such as wildfires and flooding). These are projected to continue to rise, and there is an expectation that climate change will affect all weather variables impacting the electricity system. As noted, the confluence of the above factors with climate change makes this a very difficult problem indeed, with very large inherent uncertainty; therefore, this report is focused on evaluating the changing electricity system under the current climate. However, better planning and operational strategies are needed in anticipation of extreme events, hence the discussion in Section 7, "The Impact of a Changing Climate."

### Physically Consistent Weather Inputs

Overarching in all discussions of weather inputs for power system studies is that not only has the weather dependence of the electricity system increased, but this increase has led to much more interaction of weather impacts across different elements of the electricity system, bringing increased complexity. Because impacts are related, the data inputs used must be coincident and physically consistent.

For example, the temperature at a load center is being driven by the same weather pattern driving wind and solar generation at the same time in other locations. This means that planning studies for such a system should not use approaches similar to those traditionally employed by power system modeling communities for evaluating hydro risk, where different hydro years are randomly drawn and then evaluated against loads for different weather years and random draws of generation outage. This approach assumes independence between hydro resource availability, hourly load, and other thermal generation outages, which, while not completely true, will not impact traditional study results in a profound way. But the interdependence of different weather-influenced variables means that using simple Monte Carlo methods to evaluate each variable independently is not valid in a system with significant shares of wind and solar generation.

Not capturing this dependence has a detrimental impact on the results because, while enough independent random draws will cover most possible combinations of wind generation, solar generation, and load, the sample will include many combinations that are not physically realistic and will not correctly represent the probabilities of different combinations that can occur. The interactive effects and resource diversity actually improve system reliability in some circumstances. For example, several consecutive days with well-below-normal temperature across a region do happen quite frequently, and this drives multi-day periods of high load. Periods of several days with little or no wind generation across a balancing area also occur quite frequently, as do multi-day periods of well-below-normal solar generation. If the probability of such days occurring concurrently across the same footprint was as high as the frequency their independent occurrence implies, an electricity system that is predominantly renewable resource–based would be impractical. The same is true for the independent combination of hot days with low-wind and low-solar days. However, the atmosphere follows physical rules, causing certain weather variables to correlate with others. For example, cold air in a location does not just happen. It is a result of atmospheric processes like clear skies in a location (which are correlated with high solar generation) or the movement of cold air from one place to another (which is correlated with high wind generation). In a high-renewables energy system, these atmospheric rules define the distribution of the supply and demand balance in a region, including the tails of that distribution, which are crucial to understand when assessing system risk as levels of renewables rise.

Data representing weather that impacts supply, demand, and generator outages that are used to analyze the electricity system should reflect these physical rules; we refer to such data as being physically consistent (Box 4, p. 48). Further, they must have chronological consistency,

**The atmosphere follows physical rules, causing certain weather variables to correlate with others. For example, cold air in a location does not just happen. It is a result of atmospheric processes like clear skies in a location (which is correlated with high solar generation) or the movement of cold air from one place to another (which is correlated with high wind generation). In a high-renewables energy system, these atmospheric rules define the distribution of the supply and demand balance in a region, including the tails of that distribution, which are crucial to understand when assessing system risk as levels of renewables rise.**

evolving in space and time in a way that also obeys the same physical rules that connect the weather patterns at any point in time with the weather patterns before and after that time. Energy storage and demand response increase the importance of chronological consistency, as batteries must be charged prior to providing energy to the grid and demand response is only available for limited periods of time.

Observations are always the most reliable representation of weather inputs at a point, as they obviously obey atmospheric rules (to within measurement error) and do not have any modeling uncertainty. However, because weather-driven generation is widely distributed, and solar irradiance and wind speed and direction can change dramatically over small distances, much more geographically granular weather data are now needed than can be gathered with available observations. Data to fill the gaps must be synthesized by models able to reproduce as closely as possible the observed patterns of coincident variables that define wind and solar generation and that modulate load. These synthesized weather data can then be used to determine the hourly generation potential at all current and possible future wind and solar facilities, and to estimate demand, over a long enough period to account for the range of portfolio supply and demand possibilities.

If the weather inputs are synthesized by models, the differences between the real atmosphere and the estimates from the modeling process must be well quantified, for weather fields at a given instant as well as the distribution of overall outcomes. Additionally, it is usually not valid to combine temperature output from one modeling method with wind from another and irradiance from yet another, because this will yield physical inconsistencies between the variables. Similarly, even if three instantiations of the same model were used in the above example, and the resolution of two of them was different and the third used a longer simulation to produce the results, the resulting data will have some physical inconsistency. When the modeling methods used are fundamentally similar and operate on the same input data, these inconsistencies might be small enough to not impact downstream analysis, but this should be verified and not assumed.

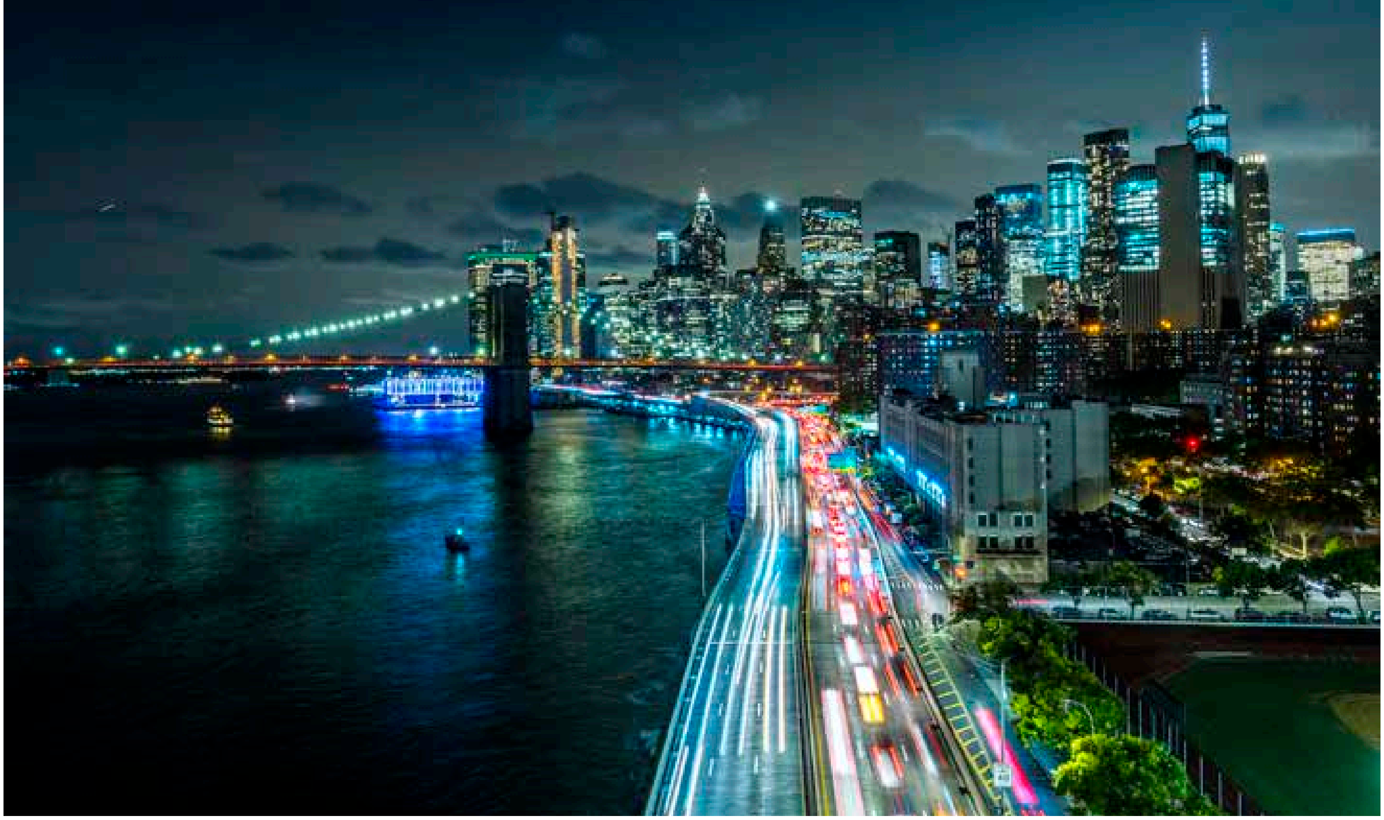
Model-synthesized meteorological datasets meeting these criteria do not currently exist. Some datasets exist that partially meet the requirements, but it is both possible and essential to create datasets with the required spatio-temporal scale and length that contain the coincident weather inputs driving wind, solar, hydro,

**BOX 4**

### The Importance of Physically Consistent Data for Different Weather Variables

This report urges the use of physically consistent data for different weather variables. Generally, this means that if the data are being synthesized, then a physics-based modeling method should be used, and the same instance of that model used to simultaneously produce all the variables. This ensures that they adhere to the physical laws and are all subject to the same inaccuracies from a single modeling process. However, the availability of synthetic weather data currently meeting these criteria is limited. The predominant datasets that are available that do meet these criteria come from reanalysis techniques or operational NWP models and have other drawbacks. For example, reanalysis data have a lower spatial resolution than is available from other datasets.

This issue was discussed by the ESIG project team, and, ultimately, the group decided on a compromise, opting not to be overly prescriptive, and suggests that, until datasets meeting all requirements become available, in some instances it is reasonable to source weather variables from more than one model as long as the following apply: all of the models are physics-based; the modeling uses similar spatial resolution; there is significant overlap in the source observations used in the modeling process; and the level of coherence between the models is evaluated. An example of where a combination of two datasets is often used is the use of high-resolution wind data from the NREL WIND Toolkit and high-resolution solar data from the NSRDB, instead of using a lower-resolution single source like ERA5. The rationale here is that any physical inconsistency is less problematic than the issues resulting from lower spatial resolution. However, the impact of this compromise needs to be more thoroughly investigated than it has been to date.



All of the relevant interdependent weather variables need to be derived in a physically consistent manner for coincident points in time and space. This is not a simple task. However, the cost of creating such datasets is trivial relative to the peril of flying blind.

and load, as well as concurrent information pertinent to events that can cause weather-driven outages and derates of system components. All of the relevant interdependent weather variables need to be derived in a physically consistent manner for coincident points in time and space. This is not a simple task. However, the cost of creating such datasets is trivial relative to the peril of flying blind.

The remainder of this section examines the components of power system modeling that are impacted by weather inputs. Although these components are described separately, it should be kept in mind that in the evolving electricity system, the same weather is simultaneously impacting all of them. The subsequent section, "An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data," describes the attributes of the meteorological datasets that are urgently needed, discusses currently available datasets and their limitations within this context, and proposes a path to bridging the gaps.

## Load Data Synthesis and/or Normalization

Overall demand for electricity follows regular, predictable patterns that are a function of the time of day, day of week, and time of year. These patterns are associated with the rhythm of human activity, changes in daylight hours, and seasonal changes in weather. Day-to-day weather also modulates these fluctuations in a predictable way, because it impacts the amount of energy used to condition indoor spaces. Temperature is by far the most influential variable on the built environment's consumption, while humidity, solar irradiance, and wind have a secondary impact. However, the relatively simple linkage of temperature to load is changing dramatically due to the increased prevalence of behind-the-meter distributed energy resources (DERs), especially solar photovoltaic (PV) generation and storage.

Another recent change is the addition of new loads due to electrification of transport and space heating, which also makes loads more sensitive to weather, though this is largely an amplification of the existing temperature dependence. The increasing use of demand response complicates the relationship between load and temperature.

Here we consider the weather-driven components of load synthesis. The process of estimating future loads, especially beyond two or three years, from past usage by accounting for changes in usage from factors like electrification, energy efficiency, and demand response is a broader area of emerging research. The discussion also does not consider demand changes due to climate change. Rather, we describe how weather drivers are used to extrapolate short load time series records to longer periods by building relationships with long records of weather drivers. That is, given a starting point consisting of a load record (that may have already been modified to account for expected load growth) and coincident weather record, the following discussion looks at how to produce climatologically representative long-term records of load using long enough weather records and how to weather-normalize existing load records to extract non-weather-driven trends.

## Establishing a Temperature/Load Relationship

When conducting power system modeling, the need is to have a long load time series that represents the expected typical load of the period being analyzed while encapsulating the impact of the range of weather conditions that have occurred historically. This can be done by establishing a temperature/load relationship. For example, if we have an hourly load record in 2022 and we know the weather in 1991 through 2022, we can estimate what the load time series would have been in 2022 if everything was held constant except for the weather, which instead was the same as it was in 1991, 1992 . . . or 2021. Similarly, we may adjust the 2022 load to estimate the load in 2024 in a way that accounts for non-weather-related load changes (for example, from increased use of electric vehicles) and then estimate the load time series for the weather in other years. This is useful because, while load records from the past years may be available to utility planners, past electricity usage is not usually a good predictor of future demand over the long term. The overall average load, the load shape, and the peak load evolve over time due to changes in population, technology, and usage. In addition, available historical load records may not be long enough to cover the range of climatologically possible weather conditions. Fortunately, robust relationships—either known to be relatively stable or that can be adjusted to accommodate known load growth—can be developed between temperature records at major load centers and the coincident load for recent years.

Once the temperature/load relationship has been established, it can be used in two ways:

- To normalize longer records of load to determine the trend in load without the impact of temperature variability from year to year. That is, a multi-year load record can be adjusted to some standard temperature year to determine how load is changing over time. These changes can then be applied to actual or synthetic loads to extrapolate them into the future.
- To build a long synthetic load record that represents how recent load conditions would vary across many weather years. This is done by using the load/temperature relationship and a long temperature record (from the same observing sites used to build the load/temperature relationship) to project recent load conditions onto a climatologically representative set of past weather years in order to better encapsulate possible load outcomes.

The above approaches allow a climatologically representative series of hourly load data and peak load to be developed for current and future systems for use in IRP or similar planning studies and resource adequacy studies, using surface temperature observations from weather stations within load centers. This is attractive because it is easy to understand, and high-quality surface data, usually from commercial airports, are nearly always available from several locations within major load centers. The sites are usually well documented, and the data are typically quality controlled, have low uncertainty, and contain few gaps.

## Accounting for Behind-the-Meter Distributed Energy Resources

Rising levels of behind-the-meter (BTM) DERs—variable generation and storage—have made the process of representing weather impacts on load much more complex. The weather dependence of load is now a function of temperature driving demand and the weather variables driving BTM generation. Since most BTM generation is solar PV, solar irradiance is the primary variable, but temperature and wind speed also affect solar output. Further, snow or ice on the panels of BTM generation may be a significant risk especially where rooftop solar is, or may become, a non-negligible winter-time generation source, as the impact will be largest during periods when load is also high. Similar impacts can be experienced from smoke during wildfire season (Juliano et al., 2022). BTM storage greatly further complicates determining load since the timing of charging and discharging is an unknown variable. All of these factors degrade the relationship between temperature and load; therefore, the impact of BTM DERs on load must be removed before the methodologies discussed above can be used to synthesize gross load time series for multiple weather years.

Fortunately, digital metering provides a way, in theory, to extract BTM generation from the net load to recover gross load, so this process is becoming easier. However, the BTM generation data that coincide with the synthesized load day must then be reapplied to obtain the actual load net of BTM DERs. For historical periods prior to collection of BTM DER generation data, this will need to be estimated from weather data. And all BTM generation data will need to be scaled to the expected BTM DER capacity for the future period being studied. This means that quality data are needed for coincidental irradiance and, ideally, also wind and temperature, all at higher granularity than was needed for the simple relationship between load and temperature. As mentioned above, while the temperature throughout a load center is not uniform, it tends to follow consistent patterns so that a small number of temperature observations can explain most of the variance in gross load. But this is not true of irradiance, as regional cloudiness follows more complex patterns. The wind field contains even more small-scale impacts, but—unlike BTM solar—assuming no sudden technological breakthroughs in small-scale wind generators, BTM wind does not look likely to become a significant contributor to power generation. Further details around estimating wind and solar generation are discussed in more detail in the renewable energy sub-section below.

### Accounting for Climate Change and the Urban Heat Island Effect

When determining the inputs needed to synthesize load, it is also important to consider the impact of climate change and the urban heat island effect of increasing

urbanization. For the purpose of creating temperature-to-load relationships over the last few years, these impacts are likely not particularly profound. However, if one wants to then use a temperature dataset going back 30 or more years to examine the envelope of normalized load, the temperature from the earlier part of the record may not be representative of the current climate. If one is conducting a planning process for a decade or more into the future, the distribution of past temperatures may be even less representative. While these impacts are not the main subject of this report, these concerns, combined with the strong linkage of load to temperature, should not be ignored. There are several ways to deal with this, from simple trend analysis to the use of data from global climate models. These are described in Section 7, "The Impact of a Changing Climate."

## Weather Data for Developing Time Series Data of Wind and Solar Output

Since wind and solar generation are completely weather-dependent, the rapid increase in their deployment has led to an urgent need for data that can be used to derive wind and solar generation.

### Essential Weather Variables

In addition to the obvious variables of hub-height wind speed (for wind generation) and irradiance (for solar generation), other weather variables also play a role—

and in many cases their role becomes larger during periods of weather-related system risk. For wind generators, changes in wind speed and direction across the rotor diameter affect output. Temperature affects air density and atmospheric turbulence, and turbines have an operating temperature range outside of which they are forced into outage to protect the equipment (e.g., Al-Rasheedi et al., 2021; Al-Khayat et al., 2021). Humidity, temperature, and precipitation all impact blade icing conditions; light icing greatly reduces machine efficiency, and moderate build-up soon leads to outage conditions. Precipitation (or the lack of it) can change blade soiling conditions and lightning causes blade pitting, both of which affect machine efficiency. For solar generation, panel efficiency is significantly affected by back panel temperature, which is a function of the ambient air temperature and wind speed. Panel soiling reduces output and is a function of wind, surface type and dryness, and precipitation occurrence (e.g., Al-Rasheedi et al., 2020). Solar output can be curtailed by snow or ice on the panels and is vastly reduced by smoke and other heavy atmospheric aerosol loads, both of which are strongly influenced by weather patterns.

### The High Degree of Complexity in Developing a Historically Complete Time Series Estimate of Possible Generating Potential

To perform accurate studies on existing generation portfolios means estimating as accurately as possible a time series of the expected generating history of the variable resources across many weather years, just as was described above for load. These time series must be long enough to determine the range of possible wind or solar generation, and they must be concurrent with load and weather-driven outage and derating risks for other non-renewable generation and transmission.

In an ideal world, power system models would use actual generation from each facility, aggregating if necessary to the level needed by the analysis being performed. However, the current fleet of wind and solar generators has a limited operating history and lacks representativeness of the expected future build-out. Existing generators are clustered in regions where there is a union of high expected output and existing transmission; they do not reflect the number and geographical diversity that is expected in the future. Further, to perform capacity expansion analysis and/or analysis that includes future wind and solar build-out requires data that encompass all possible future locations of wind and solar generation as well. At many locations where wind and solar might be built in the future, there is no observational record at all.

As a result, developing a historically complete time series estimate of possible generating potential is many orders of magnitude more difficult than synthesizing a load dataset that covers likely historical demand fluctuations due only to temperature. The story is further complicated for wind, because generation is proportional to the cube of wind speed, and wind distributions vary dramatically over short distances, at different hub heights, and for different rotor diameters—especially in complex terrain, which is where much of the best wind resource is located.

For this reason, the current representation of coincident wind and solar generation data used alongside load data in power system modeling is limited and relies on combinations of methods and data that are widely accepted as insufficient. The primary methods used are outlined in the next sub-section, along with data needed for each.

Where a limited study does allow the use of observational data, care must be exercised to ensure the proper understanding of whether the generation data include periods of curtailment due to operational constraints, and, if they do, an understanding of whether the curtailed energy is added back into the actual production and the method used to do this.

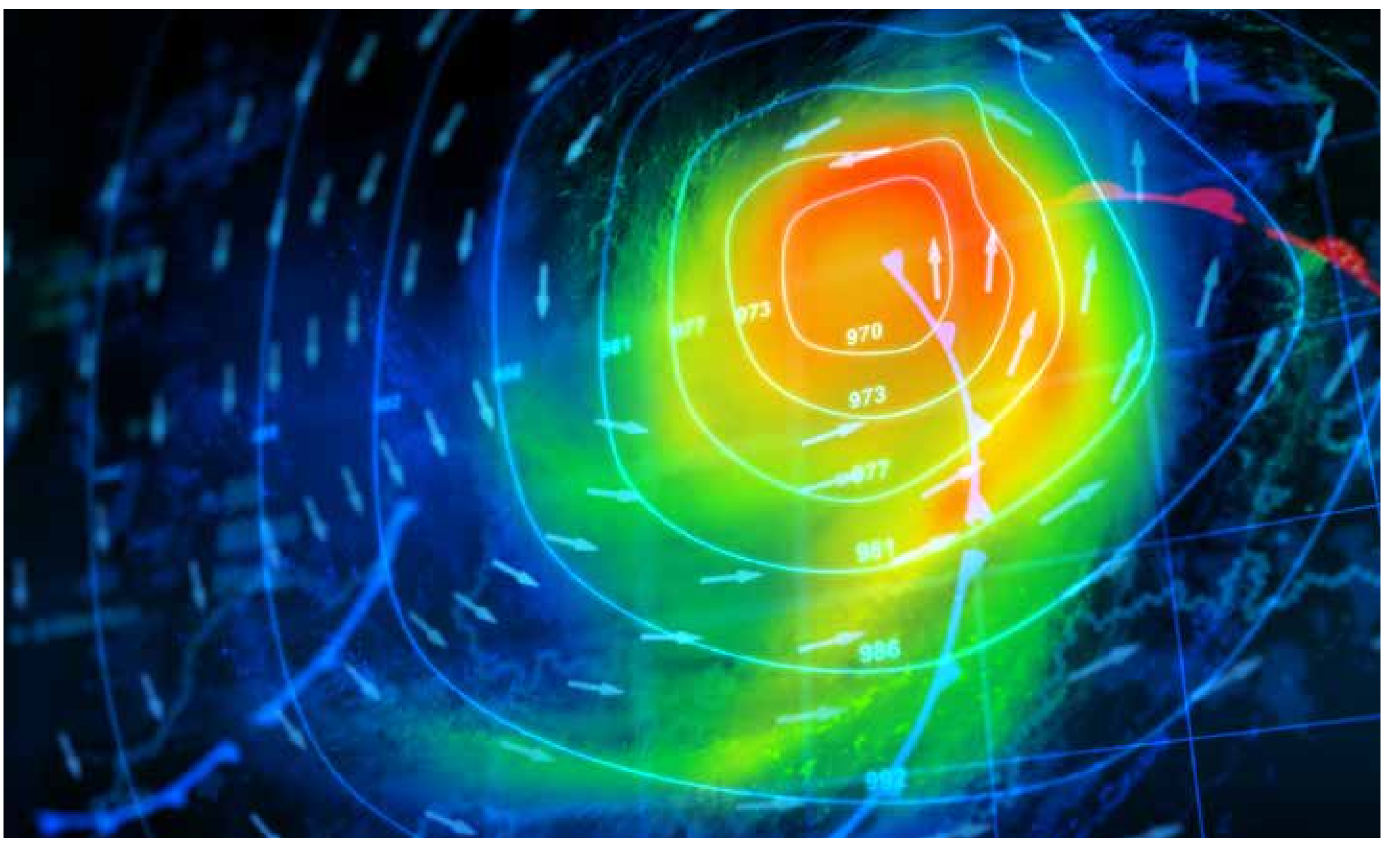
**The current representation of coincident wind and solar generation data used alongside load data in power system modeling is limited and relies on combinations of methods and data that are widely accepted as insufficient.**

Modelers also need to be careful not to represent future wind plants with legacy wind turbine technologies. Wind turbine towers are getting taller, rotor diameters larger, and overall machine design more efficient, thus increasing energy yield relative to older technologies. In some cases, a larger rotor may be used in slightly less windy areas, effectively lowering the cut-in wind speed and increasing the turbine capacity factor. Solar panels and tracking systems are also getting more efficient, and falling costs are pushing up inverter loading ratios (the number of panels installed per MWac of power) to increase capacity factors.

## Extrapolation and Synthesis: Two Methods to Produce Datasets Long Enough to Capture Climatology

When the period of operational data is not long enough to meet a study's needs, the options are to extrapolate to a longer record or produce a completely synthetic estimate of output. Both methods have advantages and disadvantages, but it is worth noting at the outset that while the use of empirical correlations to extrapolate longer generation time series for renewable resources is intuitively easy to grasp for non-meteorologists, it should be regarded with skepticism even if other options seem limited.

Extrapolation methods relate the observed generation to meteorological variables to create an empirical power curve in a similar fashion to creating the relationship between temperature and load. If done at the level of an individual wind or solar facility, this will implicitly account for loss factors like wind plant wakes, solar inverter losses and clipping,[22] collector system losses, and substation losses. However, unless the data used to create the empirical power are carefully prepared, the function will also implicitly account for the average effects of other loss factors like output curtailment, equipment availability, icing, snow on panels, or high wind cut-out. These are factors that it would be best not to include in the extrapolation. In addition, site-level wind and solar data are difficult to procure and often poor in quality, especially the detailed data defining how overall generation is impacted by other operating conditions such as generator availability, and output curtailment due to transmission constraints or market conditions.




22 Because of the variable shape of solar output through the day and year, and because inverter capacity is very expensive, solar facility inverters are often sized smaller than the installed capacity of solar panels because the energy lost during the relatively limited times that output exceeds the inverter sizing is worth less than the cost of a larger inverter. The resultant effect is called clipping.

Synthesizing generation estimates without reference to actual generation data uses power curves specific to the installed equipment (for example, model specification for the wind turbine, PV panel, and inverter). The advantage of this is that it is more generic and can be used for hypothetical future plants, but the disadvantage is that loss factors are not based on actual power plant configurations in the field.

Whether extrapolation or synthesis methods are used, it is worth noting that if aggregated estimates of output for an area are needed, the aggregation usually needs to be done by estimating the output at each facility and summing the results. This is because wind and solar resources can change over short times and distances in ways that do not follow simple relationships to "average" conditions. In addition, different sites employ different technologies that do not directly scale. However, as is discussed below, some current practices attempt to extrapolate production at a regional level. But at best, this produces a coarse relationship between regional meteorological conditions and renewable output and is generally not recommended as the amount and geographical diversity of renewable resources increases.

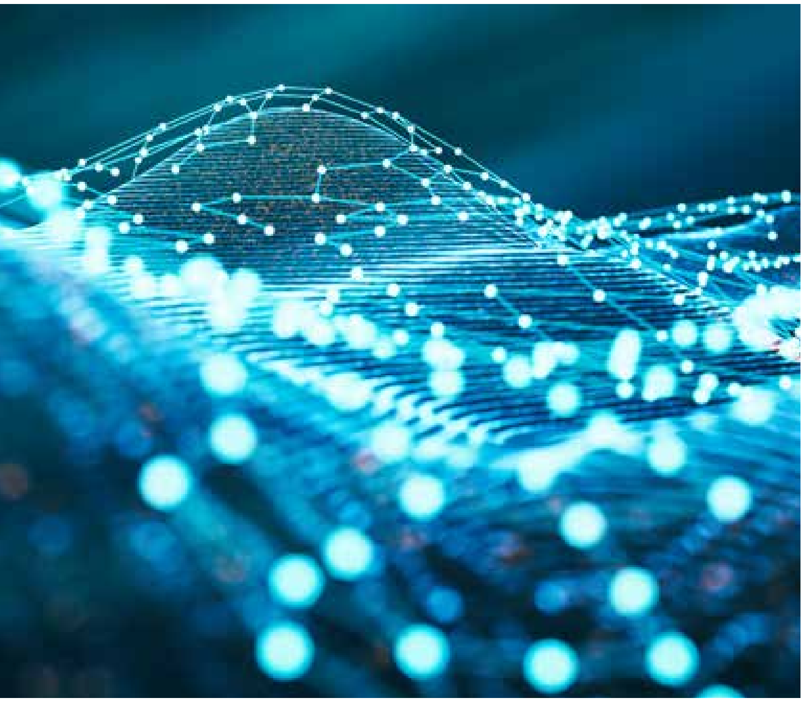
### Estimating the Generation of Renewable Resources Using NWP or Other Physical Models

Renewable generation can be estimated using time series data of weather produced by numerical weather prediction (NWP) models and/or by processing data from satellite observations using other physics-based models. Both sources can provide high-resolution gridded data of some, or all, of the variables needed to estimate generation potential. From time series data of weather, it is possible to estimate the generation of any existing or hypothetical renewable resource site. The weather data from grid points near an existing or hypothetical renewable generation facility can be passed through power curves (usually based on equipment parameters and standardized loss factors, but they may be derived from actual plant data) to develop an estimate of expected generation, and can then be aggregated as needed for the power system analysis being performed.

The methodology of using gridded data to produce generation estimates at current and expected renewable generation locations and then aggregating is appealing because, as noted in Section 2, "Meteorological Data Fundamentals for Power System Planning," gridded datasets are easy to process. However, the quality of the generation estimates is governed by the quality of the model used to create the weather data. For example, the considerations discussed in Section 2 including resolution, parameterizations, and any weather patterns that impact systematic errors must be understood when using NWP models. It is essential to validate model output in the context of a given study. As discussed above, even if a broad validation study has shown low average bias and error rates across a model domain, large deviations from reality may exist at certain times and/or locations.

**It is essential to validate NWP model output in the context of a given study because, even if a broad validation study has shown low average bias and error rates across a model domain, large deviations from reality may exist at certain times and/or locations.**

Aside from the uncertainties of using model-based inputs, there are not currently any datasets available that meet the requirements of (a) providing all the necessary weather data for a long enough time period, and (b) having sufficient resolution to properly estimate the necessary variables, to estimate generation across a long enough time period, especially in locations other than flat plains. These gaps, and possible ways to fill them, are discussed in more detail in the "Guidance for Using Existing Weather Inputs" section below. For now, it is worth noting that the National Renewable Energy Laboratory's (NREL's) WIND Toolkit provides enough resolution to capture most features driving wind resources but currently does not cover a long enough period, while the European Center for Medium-Range Weather Forecasting's (ECMWF's) ERA5 dataset covers a long enough temporal period and is regularly extended but does not resolve many regional or local weather features driving renewable resources (Molina, Gutierrez, and Sanchez, 2021). The National Solar Radiation Database (NSRDB) provides good overall estimates of solar irradiance, but validation suggests it may not capture some of the short-term variability sufficiently for power



systems studies (Habte, Sengupta, and Lopez, 2017). It is regularly extended, but its 23-year length is not quite long enough to use concurrently with wind and load data to capture the full envelope of concurrent variability.

### Estimating the Generation of Renewable Resources Using Historical Generation and Empirical Correlations

The second category of methods often used in IRPs and similar planning studies is to use actual historical renewable generation data for a region and attempt to develop empirical relationships between the generation data and a longer time series of weather observations from one or more nearby sites. These methods benefit from being easy to understand and simple to implement. They also use standard meteorological observations, which are relatively easy to acquire for long time periods. However, because they are not physics-based, these methods usually suffer from the same issues as using a Monte Carlo simulation that does not correctly reflect the dependence of each weather input on the others, and thus on different components of the electricity system. Where models generating data in this way are used, it is important that the data are validated, not just to verify that the overall distribution of outcomes for wind or solar generation looks realistic, but to confirm that the data the models produce meet the concurrency and chronology requirements—otherwise, the data will not represent the overall balance of supply and demand situations that actually occur. Inaccuracies inherent in the methodologies described here will grow larger as the installed capacity of wind and solar increases.

One empirical method often used is a flavor of the "measure, correlate, and predict" method discussed in Section 2, "Meteorological Data Fundamentals for Power System Planning," that is used to extrapolate load records based on temperature records. The performance of this method declines as the time interval being predicted gets smaller (e.g., annual adjustment is more accurate than monthly, which is more accurate than daily) because non-linear physical effects are more prevalent at smaller time scales. Moreover, hourly estimates using relationships like these are usually a poor representation of actual conditions, especially for wind energy and especially if the long time series record is not a hub-height wind speed near the site of interest. The method will yield a statistical distribution that seems reasonable, but the data will not be coincident or physically consistent between different variables or locations. For example, if MCP is applied to predict the wind field at two different locations, or to predict the wind and temperature at the same location, the output fields will not represent the concurrent physical state of those variables for the specific time and location.

Another method sometimes used to extend time series of regional renewable output is to take a long record of a weather observation together with the shorter overlapping regional renewable resource generation record. The generation record is first normalized to current capacity values to account for the ongoing increase in renewable resource installations. Ideally, this process would take into account the impact of technology changes, since newer capacity usually generates more energy for similar wind speeds, but typically, if these adjustments are made, they are rudimentary in nature. Then, for periods with no output data, days that are most "similar" are found in the overlapping record and applied to the day without data. For example, if a 10-year time series of wind generation is available and a 30-year record is desired, while a nearby long observing record contains 30 or more years of temperature data that overlap, then the average temperature for a day in the past for which renewable output is unknown (say, March 17, 1991) may be compared to surrounding days in the past for which there is renewable output (e.g., all days from March 15 through 19 from 2011 to 2020). The renewable output for the day with the closest temperature match (e.g., March 16, 2015) is then used as a proxy for wind generation.

The idea in this example is to maintain some of the relationship between load (via the temperature variable and time of year) and renewable output. However, the actual combination of possible effects driving temperature at the observing site has limited overlap with the effects driving renewable generation at the many distributed locations. Hence, even if the distributions of possible renewable generation outcomes obtained with this method approximate those seen in the observational record (which they often will, since the resulting data are composed of duplicates of the original observations), much of the concurrency in time series variables is lost. In addition, because regional operational data are usually used, geographical diversity effects are lost. The data may also contain artifacts like curtailment, outage effects, and operational issues that will bias the results and are difficult to remove without analyzing the facility-level data, which are usually not available.

## Water Inflow Data for Hydro Generation

While not the focus of this report, there is significant variability in hydro generation capacity from month to month and year to year, and this variability needs to be accounted for in planning studies. Hydro variability is typically evaluated by examining different combinations of hydro years, load years, and outage possibilities. There are several ways to do this. One of the simplest options is to iterate through the combinations of load years, hydro years, and different generation outage conditions. For example, if one has a 20-year load time series and 10 years of water inflow data, there are 200 different combinations of load and hydro. Each of these 200 combinations can be simulated multiple times with different generation unit outage combinations selected for each simulation hour using Monte Carlo methods that can vary in complexity from completely random to weighted according to the unit history, season, and/or weather. In this example, if 100 iterations were performed for each hydro and load combination, the total number of simulations would be 20,000. However, there are more sophisticated ways to increase the sample space, including creating many additional load years by combining load days from different years in ways that are plausible based on analysis of past load trends. For instance, one can use a Markov chain approach that randomly walks between days in different weather bins according to historically observed transitions. Examples of some of these approaches can be found in Hart and Mileva (2022).

The methodologies above assume that available hydro power capacity is independent of day-to-day weather and outage probabilities of any generation on the system. Because of the inertia inherent in the water cycle processes and the management of water through hydro systems, these assumptions are mostly valid.

While the assumption of the independence of water inflow data from short-term weather variables allows the methods described above to be used to increase sample size, water inflow data are still needed for assets present in a region being modeled, and the longer the period available the more thoroughly hydro variability can be accounted for, and the more uncertainty will be reduced.

## Weather Data to Estimate Outages and Derating Likelihood for All Asset Types

Outage data for bulk system assets are collected by the North American Electric Reliability Corporation (NERC) as part of the Generating Availability Data System (GADS) and Transmission Availability Data System (TADS) programs. For generator assets, multiple variables are calculated, and the most relevant for resource adequacy studies is the forced outage rate. There are many variations on how this rate is represented, but they are the cornerstone of the resource adequacy studies, and the forced outage rates of thermal resources and the time-varying production levels of variable renewable resources generally drive the study results.

Traditional modeling methods for resource adequacy use these forced outage rates to calculate the probability of a resource shortage in every hour of the study (one or more years). Currently, most modeling frameworks assume that these outages are uncorrelated and the forced outage rates are generally not linked to weather. This means that current resource adequacy models cannot generally

account for correlated outages of hydro-thermal resources or correlations between thermal outages and wind or solar availability, including those outages that are linked together by severe weather. Yet studies show that weather that creates peak loads across a region also increases the likelihood of generation and transmission outages (Murphy, Sowell, and Apt, 2019; Allen-Dumas, Binita, and Colin, 2019).[23] Recent events such as winter storms Elliott (2022) and Uri (2021) and summer heat waves, together with an increased focus on climate risks, have brought more attention to this area.

Extreme weather can result in multiple parts of the system failing simultaneously (known as common mode failures), while very high or very low temperatures simultaneously drive high loads. For example, during extreme winter weather, coal piles can freeze, gas turbine air inlets can become iced, and gas supply to generators can be limited due to conflicts with demand from commercial and residential heating, frozen wellheads, or loss of pressure at the gas compression stations. Wind turbines' blades can become iced, or turbines can shut down due to low temperatures, and solar panels can be covered in snow and ice. In the summer, equipment can overheat, and a lack of cooling water (or cooling water that is too warm) may result in thermal plant derating or outages.

Incorporating these relationships in power system simulations requires detailed records of outage events at each generation plant and accurate local weather data to correlate to the events. Such data are easy to record, and the problem is more one of reporting and sharing than of technical feasibility. Generation outage or derate codes that are used in an event are not always reflective of the original cause of outage or derate. Generation outage codes are also often not granular enough to be able to distinguish various causes of outage or derating related to different weather phenomena. In some cases, the weather data are regarded as proprietary and are not shared in ways that would help in system analysis. For generating resources that come online in the future, weather information at potential future sites will help

23 See also Dison, Dombrowsky, and Carden (2022).

analyze risk. Several of the datasets that are already available, to be discussed in Section 4, "An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data," can in many cases be used as a proxy for this. For cases where more detailed data are needed, the proposal in Section 5 for producing more complete wind and solar datasets will also yield much of the required meteorological data, although additional work may need to be done to model water intake temperatures for thermal plants.

## Weather Data to Assess Transmission and Distribution Risk

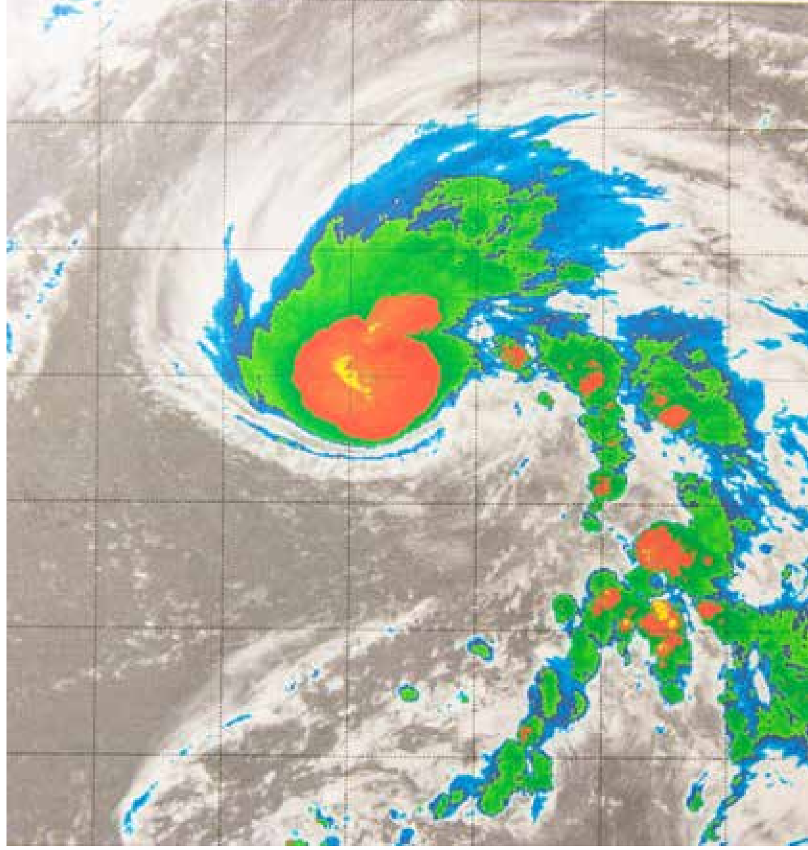
The actual energy a transmission or distribution line can carry is a function of environmental conditions. Electricity passing through any conductor will result in resistive heating, and as the conductor warms, it will expand and sag. This heating is a function of the conductor type and the current being carried. Meanwhile, the heat dissipation rate is a function of ambient temperature, wind speed, and solar insolation. This effect of environmental conditions means that the true limits of a line—the dynamic line rating—can be different from the static rating, and in many cases additional capacity is available. In operating electricity systems, temperature sensors on transmission lines can allow additional current to be carried when conditions allow. This is particularly useful as wind generation increases, because lines may be able to carry wind energy that otherwise would require expensive line upgrades to accommodate. On the other hand, extreme temperatures can result in line deratings. Analyzing these effects in power system models requires concurrent information about wind, temperature, and irradiance along the line path.

Transmission and distribution can also be forced out by wildfire (or wildfire risk), lightning, extreme winds, and extreme icing. Wildfire risks usually result in temporary de-energization to prevent ignitions, active fires cause de-energization to protect equipment, and smoke from wildfires can cause short-circuit faults on the power lines leading to line outages if attempts to clear the fault (by disconnecting and reconnecting the line) are unsuccessful. Lightning damage is typically recovered from quickly. However, extreme winds and icing can be catastrophic, as the loss of a major line can take weeks or even months to repair, during which time the system is much more vulnerable to other outages or derates. Many of the same risks apply to both transmission and distribution, but distribution outages usually result in the loss of local load and thus are not a major risk to broader system reliability (although the increase in BTM generation makes this separation less clear).

While the impacts of weather on the transmission and distribution systems are mostly limited to extreme events, such events often occur as part of the common mode failures described above for generator outages and derates. Therefore, as weather datasets are developed to address the other system components discussed in this section, it is worth keeping transmission and distribution impacts in mind, and, where possible, ensuring that new meteorological datasets provide synergistic coverage of these impacts.

## Incorporating Supply and Demand Forecast Errors into Power System Modeling

During power system operations, forecasts of load, wind, and solar for the operating horizon are used to facilitate market function, security-constrained unit commitment, and real-time dispatch. Here, the term "forecast" is being used according to the first usage in Box 1 (p. 11), an estimate of future conditions in an operational time frame. To reduce confusion we will call such forecasts "operational forecasts." During actual operations, the load and generation will have been forecasted multiple times at different lead times, most commonly day ahead, hour ahead, and a few minutes ahead. Power system analyses of market operations, unit commitment, and/or dispatch, such as those performed for integration studies with production cost models, can make an assumption of perfect foresight, but it is more realistic to incorporate imperfect forecasts that have a similar level of accuracy as that achieved by operational forecasts. Therefore, for this use case, time series that provide estimates of load and wind and solar generation for use in the modeling also need corresponding operational forecasts for each time interval, and each time interval typically requires an operational forecast for several different lead times. Thus, if the weather data being used to estimate reality are being synthesized, we also need to do another level of synthesis to produce the corresponding operational forecasts, and "forecast" data need to have the same "accuracy" characteristics relative to the synthetic data as the real operational forecasts have relative to the real data.



The simplest way to incorporate a level of uncertainty or error into power system modeling is to determine the typical forecast error level in operations by comparing to actual loads and generation at various operating gate closure times (e.g., day ahead, hour ahead, and 10 minutes ahead). This can be done at varying levels of sophistication that account for different times of day, different seasons, and different weather regimes. It is especially important to assess the accuracy of operational forecasts during periods of weather transition, as this tends to be when the largest errors in wind, solar, and load forecasts occur, and the errors often exhibit synchronicity. Once typical error levels are known, they can be used to develop synthetic operational forecasts for use in power system models by perturbing the load, wind, and solar data according to the error levels. These forecasts can be scaled to perform sensitivity tests at differing levels of accuracy. Ideally, forecasts used in operations are probabilistic and contain information about the distribution of uncertainty. Probabilistic forecasts provide more information that can be used, for example, to create scenarios for stochastic modeling or characterize system risk.

This approach has some shortcomings. One is that operational forecast data are often proprietary. Additionally, operational forecast data do not provide information about the forecastability of future generators; that is, the skill of a power output forecast for a wind farm cannot be determined until the wind farm is built and the forecast for any given time can be compared to actual output at that time. And, as noted, forecastability is a function of weather regime. Therefore, to provide an accurate representation of forecast accuracy for any given time in a power system modeling exercise requires forecast data that are representative of the forecast skill for the weather occurring for the time and location being modeled. For this reason, the NREL WIND Toolkit contains a wind forecast component in addition to the main dataset intended to represent ground truth.[24] The Toolkit contains forecast datasets at several lead times representing typical decision points so that simulated functions like unit commitment and dispatch can be modeled based on data that mimic the accuracy and autocorrelation of then-state-of-the-art wind forecasts. However, forecasting quality has improved significantly, and the WIND Toolkit forecasts no longer represent state-of-the-art accuracy. Since the WIND Toolkit was produced in 2014, the deployment of renewables has increased dramatically, and since existing renewable assets are generally located within the same microclimates as most future assets, the accuracy of future forecasts can be better determined now than in the past, using statistics from forecasts for existing assets.

Lastly, geographical diversity and the increasing deployment of storage will also begin to erode the impact of short-range (minutes ahead) forecast errors, but forecasts in the hours-ahead to day-ahead range increase in importance due to the need to optimize storage charge and discharge. With all these factors in mind, for some applications, it may be necessary to generate a forecast dataset in a similar manner to that produced for the WIND Toolkit, in order to provide the best possible reflection of expected skill of predictions used for unit commitment and dispatch tasks in power system models.

24 To the knowledge of the authors, the WIND Toolkit is the only dataset designed specifically for wind integration studies that contains a companion "forecast" dataset.

SECTION 4

# An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data

This section gives a list of the data and data attributes needed for a weather inputs database used for power system modeling. These comprehensive data meet the requirements to account for and study the increasingly weather-driven aspects of the electricity system. This section describes the required variables, spatial and temporal resolution, and time series length. We examine the extent to which the currently available data can be used on their own or collectively to meet these needs against these criteria to reveal the gaps. Then, having concluded that no combinations of data or datasets are available that fully meet the needs of power system modelers and planners, a proposal is offered in Section 5 for how these gaps can be robustly filled.[25]

**The most fundamental and urgent weather data need for conducting system planning studies in most regions is accurate wind speed, solar irradiance, and temperature data for every plausible location in the region of interest, and, if imports and exports are to be correctly modeled, for all interconnected regions. However, to develop power production and load profiles, several additional variables are required to describe the complete range of weather impacts.**

## The Data and Attributes of Ideal System Planning Weather Inputs

Currently, the most fundamental and urgent weather data need for conducting system planning studies in most regions is accurate wind speed, solar irradiance, and temperature data for every plausible location in the region of interest, and, if imports and exports are to be correctly modeled, for all interconnected regions. However, to develop power production and load profiles, several additional variables are required to describe the complete range of weather impacts discussed in the previous section. Most power system modeling requires hourly or better granularity of these data, with 5-minute intervals the preferred granularity for some production cost modeling applications to assess ramping capability and reserve needs. To properly capture overall weather variability, several decades of data are needed. Datasets meeting these criteria would make it possible to carry out power system studies that evaluate the impact of potential renewable resource expansions as well as enable assessments of extreme weather and its impact on grid reliability. Such studies are a prerequisite for planning and building a reliable low-carbon power grid with large amounts of weather-driven generation capacity.

Table 1 (p. 61) lists the main attributes of time series data necessary to meet general power system modeling needs.

In addition, a specific organization or entity will need to assume responsibility for curating the data, including providing quality control and validation, flagging issues, advising on best practices for its use, and evaluating the need for and scope of periodic updates.

We want meteorological data produced for power system models to reflect actual conditions, but sufficient observations are not available for any study beyond a very

25 For definitions of terms that some readers may be unfamiliar with, please see the glossary at the end of the report.

**TABLE 1**

## The Main Attributes of Time Series Data Necessary to Meet General Power System Modeling Needs

| | |
|---|---|
| **Including the necessary variables** | Include the necessary variables at sufficient spatio-temporal resolution and accuracy to reflect actual conditions that define the generation potential at current and future wind/solar sites and temperature at load centers |
| **Covering multiple decades with ongoing extension** | Cover multiple decades with consistent methodology and be extended on an ongoing basis to capture the most recent conditions and allow climate trends to be identified |
| **Coincident and physically consistent** | Are coincident and physically consistent, in space and time, across weather variables |
| **Validated** | Are validated against real conditions with uncertainty quantified |
| **Documented** | Are documented transparently and in detail, including limitations and a guide for usage |
| **Periodically refreshed** | Are periodically refreshed to account for scientific and technological advancements |
| **Available and accessible** | Publicly available, expertly curated, and easily accessible |

Source: Energy Systems Integration Group.

small spatial scale, and a sufficiently long history to capture climatology will likely never be available. This, and the other attributes in the list above strongly suggest that the production of such a dataset will need to use numerical weather prediction (NWP) modeling approaches and that high-resolution reanalysis and/or downscaling NWP methodologies are the best fit. Generative adversarial network (GAN) machine learning methods also show promise for producing sufficient spatio-temporal resolution at lower overall computational cost than using only high-resolution NWP modeling, and other statistical post-processing methods could be applied to correct known NWP model biases. Each of these methods are detailed in Section 3 above, "Weather Inputs Needed for System Planning."

Their use allows the resultant dataset to be anchored on as many observations as possible, while at the same time the full dynamics and physics of the NWP system can produce dynamically consistent and realistic fields where observations are not available, especially in complex topography. In the discussions below there is a general assumption that use of NWP methods will be necessary to satisfy the requirements for power system modeling weather inputs. As described in Section 2, "Meteorological Data Fundamentals for Power System Planning," several factors impact the quality of a dataset that has NWP at its foundation, and these are reiterated where appropriate below.

**The use of NWP modeling allows the resultant dataset to be anchored on as many observations as possible, while at the same time the full dynamics and physics of the NWP system can produce dynamically consistent and realistic fields where observations are not available, especially in complex topography.**

### **ATTRIBUTE 1:** Includes the Necessary Variables Across Required Regions with Sufficient Spatial and Temporal Resolution to Meet Power System Modeling Study Needs

#### Sufficient Spatial Resolution

Datasets must be produced by processes that provide sufficient spatial resolution to accurately resolve the phenomena impacting supply and demand. This means:

- Knowing temperature in enough detail to accurately predict its impact on load
- Specifying variables driving wind and solar in enough detail to quantify the generating potential at every plausible generation site
- Having information about weather phenomena at a scale that can be used to estimate their impact on thermal generation derates and outages, transmission, and distribution

Wind resource is the limiting factor in determining how to use models to fill in gaps in observations, because the wind field is heavily influenced by topography and near land and water interfaces. To be able to estimate the output from wind plants driven by (or even just modulated by) phenomena like sea breezes, gap flows,[26] and mountain-valley circulations—at the granularity needed for system planning—will require that the data points are no more than 2 km from the point of interest, and ideally much less. Even in the Great Plains of the U.S. the wind field varies significantly across rolling hills. In this case, the sources of this variability are much simpler than in areas of complex terrain and can be corrected by statistical post-processing in places with an observational record (as discussed in Section 3); however, observations to drive these corrections are not available in many locations.

Figure 11 (p. 63) illustrates why spatial resolution is important, using examples from two wind plants in the Pacific Northwest. The Big Horn wind plant is built in steep terrain, and the turbines at the north end are at elevations several hundred meters higher than those at the south end. The top left panel shows how, in the summer, the wind resource is primarily driven by gap flow through the Columbia Gorge, which can be quite shallow and often does not reach the higher-elevation turbines in the northern part of the plant. The bottom left panel shows how, in the winter, the higher-elevation part of the wind plant often experiences strong winds from the southwest as storm systems pass by, but the lower elevations remain entrenched in a layer of cold, stable air that prevents the momentum from mixing down. On the top right, it can be seen that higher ridges to the west sometimes excite semi-stationary mountain waves that will migrate slowly across the plant, leading to extreme differences in wind speed across short distances. The bottom right shows how the Klondike (KL1, KL2, KL3, KL3A), Hay Canyon (HC), and Star Point wind plants, which sit a short distance to the southwest, are also impacted by mountain waves. In addition, this area can come under the influence of the wake generated by Mount Hood about 100 km to its west. The impact of the wake can be seen in the low wind speeds cutting through the Klondike 3A wind farm in this graphic.

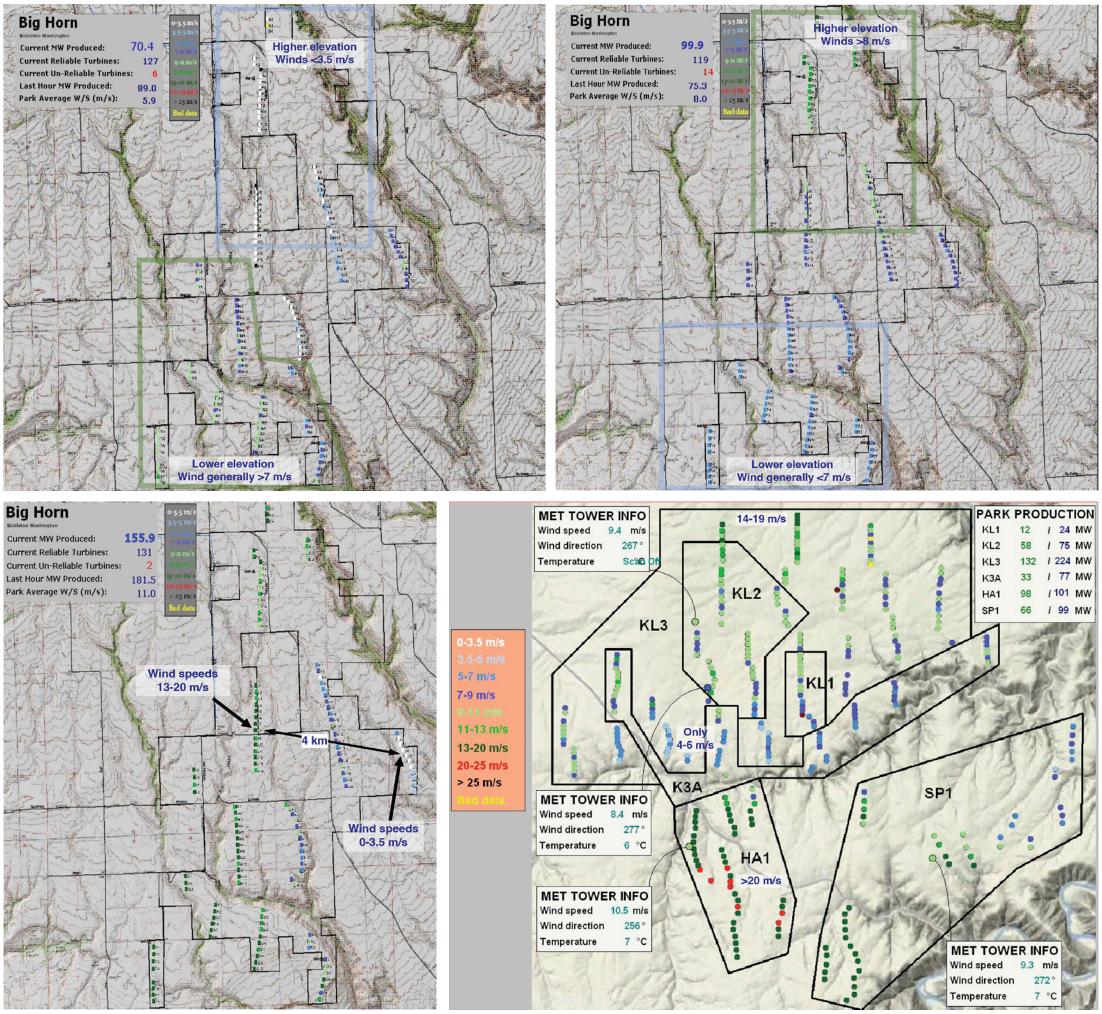
**While it is not necessary to resolve the wind speed at every turbine in a power systems dataset, it is important that the dataset estimates the impacts of features like those described in the example at least to a level where the regional wind generation can be accurately modeled.**

Section 3 describes the importance of model resolution when using NWP to produce a representation of atmospheric structures that are influenced by terrain and surface characteristics. While it is not necessary to resolve the wind speed at every turbine in a power system dataset, it is important that the dataset estimates the impacts of features like those described in this example at least to a level where the regional wind generation can be accurately modeled. For instance, in the Columbia Gorge there are many wind plants like Big Horn that span considerable elevation or, like Klondike and Big Horn, are impacted by waves and wakes. If an NWP system is used at a resolution that does not resolve the existence

26 Gap flows are a phenomenon driven by the interaction of atmospheric pressure fields with topographical features like mountain gaps, passes, gorges, canyons, and channels.

**FIGURE 11**

## Turbine Wind Speed Maps for Two Wind Plants in the Columbia Gorge of the U.S. Pacific Northwest, Showing Horizontal and Vertical Variabilityof NWP




**Top left: A shallow summertime gap flow event causes high output at lower-elevation southern turbines while the northern turbines are below cut-in. Top right: Strong southwest winds during a winter storm power turbines at higher elevations, while stable air prevents this momentum reaching lower-elevation machines. Bottom left: Mountain waves lead to dramatic east-west output differences. Bottom right: The wake of Mount Hood at a nearby plant on a plateau on the south side of the Gorge.**

Source: Iberdrola Renewables.

of features like those in Figure 11—if it cannot "see" them—it will not be able to produce wind fields that correctly estimate the hourly output from these facilities. In this case grid spacing of at most 1.33 km is necessary to resolve the features (Sharp and Mass, 2002). Output from NWP modeling performed at a lower resolution will require refinement by downscaling with higher-resolution NWP or methods like the GAN machine learning technique described in Section 3, or will need to be statistically corrected. Statistical correction will be difficult in remote locations (where future wind plants may be sited) because observations are not available or are of poor quality. Note that GAN downscaling requires some high-resolution output to train the downscaling method and has yet to be fully proven, so data produced this way need to be especially well validated. In any event, including when only high-resolution NWP is used, it is crucial that the model data are validated (see

Achieving the type of resolution where wind shear across the turbine rotor can be resolved is difficult in NWP modeling, and validation studies should be performed to determine the model's skill in describing reality rather than assuming that vertical structure is accurate in all weather regimes.

below) as extensively as observations allow in order to understand the limitations of the modeling method.

Sufficient vertical resolution is also needed in models used to derive wind data. Near the surface, the desire is to realistically capture some of the vertical structure of the atmosphere across the rotor layer as it evolves through the diurnal cycle. Enough vertical levels are needed so that model output at multiple heights near the surface is not merely an extrapolation across three or fewer model height levels. This will generally improve the accuracy of different hub-height predictions. In addition, the strength of the wind resource is often heavily influenced by regions of strong atmospheric stability near the surface, and a lack of vertical resolution may result in the sharp vertical gradients in fields like wind and temperature not being sufficiently resolved. Lastly, if feasible, model levels should be configured so that the output in the rotor plane requires minimal interpolation. It should be recognized, though, that achieving the type of resolution where wind shear across the turbine rotor can be resolved is difficult in NWP modeling, and validation studies should be performed to determine the model's skill in describing reality rather than assuming that vertical structure is accurate in all weather regimes.

Resolution is also important in some regions for defining the complexity of fields impacting solar generation potential, especially if NWP output is being used instead of model-processed satellite data to predict clouds and aerosol components. NWP output may also be valuable for producing temperature fields for more advanced treatment of load forecasting. Here again, sufficient resolution is crucial if small-scale features have significant effects, such as in coastal cities like San Francisco, where large temperature gradients can exist across short distances due to marine effects.

### Data for Wind Generation Estimation

To make an approximation of wind generation, wind speed near hub-height is needed. Wind speed accuracy is a major factor in reducing uncertainty in the generation estimate, primarily because generation scales with the third power of wind speed in the part of a wind turbine power curve between cut-in and the knee of the power curve (Figure 12, p. 65). The sensitivity of the generation estimate to changes in wind speed is further complicated by the region of zero generation below cut-in speed, the flattening of the cubic curve toward rated speeds (as the blades begin to pitch), and the impacts of high wind speed cut-out and hysteresis at high wind speeds.

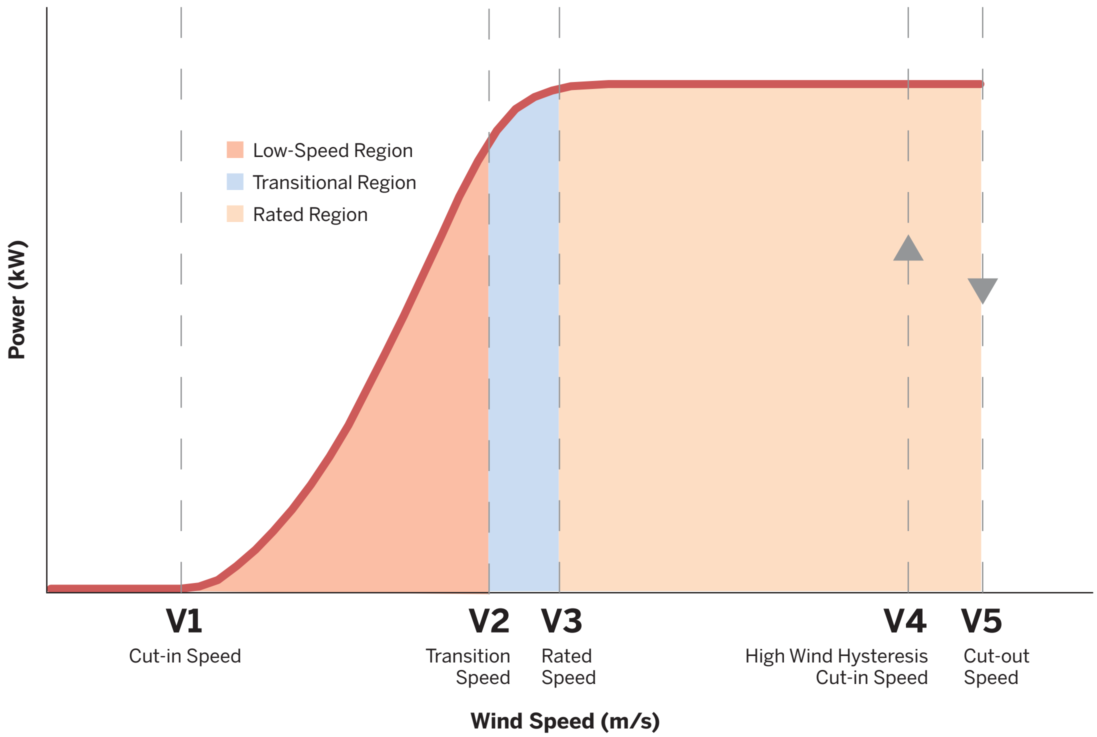
Since wind observations are not available at the spatial and temporal density needed for power system modeling, the most accurate alternative is output from NWP models. Wind is calculated at each point in an NWP model at every time step in the modeling process, and since the internal model time step is usually considerably shorter than the shortest required interval for power system modeling (about 5 minutes), NWP models can produce the required temporal resolution. No model can consistently predict wind speeds to within the 1–2 m/s range that is needed, so model data must be validated against wind observations wherever possible to develop insight into the skill and uncertainty of the model and the resultant impacts on generation estimates. If possible, some form of bias correction should be applied if model distributions are found to deviate considerably from observations.

Datasets for power system modeling need to include wind data at several levels from the surface through to 300 meters to provide wind speed throughout the rotor layer for many different possible hub heights, including those of the tallest offshore turbines. It is crucial *not* to extrapolate near-surface (10 m) winds to hub height, because at night the surface decouples from the free

It is crucial not to extrapolate near-surface (10 m) winds to hub height, because at night the surface decouples from the free atmosphere above so that hub-height winds increase while surface winds decrease.

FIGURE 12

## Wind Turbine Power Curve




**Wind turbines are sensitive to small changes in wind conditions. V1 is the cut-in wind speed, the speed above which a turbine begins generating power. V3 is the rated wind speed, the speed at which the turbine reaches its rated power output; at speeds higher than this, no additional power available in the wind is captured, as the generator cannot further increase its output. V5 is the cut-out wind speed, the speed at which the pitch of the turbine blades reduces the output to 0 to protect the turbine. Operation of the turbine is suspended until the wind speed has slowed to V4 before it goes back up again, cycling between V4 and V5.**

Source: Energy Systems Integration Group.

atmosphere above so that hub-height winds increase while surface winds decrease. Wind direction is useful for refining generation estimates, as wind plant power curves can be developed that account for the different levels of waking from different directions.

Temperature is useful, ideally at regular intervals from the surface to the top of the blade-swept area, although a single value somewhere between 50 m and 100 m will suffice. It is used to determine air density, which is a secondary variable in calculating power generation and can be important when building time series for prospective sites that do not have any generation data that can be utilized to create a plant power curve (which bakes in density effects).

Temperature at multiple levels through to the top of the blade-swept area is useful to determine other features of the operating environment. First, it gives an indication of the presence of strong surface stability. The presence of strong surface stability is an indicator that the model winds might be less accurate than usual because NWP models are notoriously poor at handling stable boundary layers and mix them out too quickly. Second, hub-height temperature can be used to determine whether cold or hot weather shutdown is likely. Lastly, relative humidity is also a useful variable, as the combination of temperature and relative humidity is an indicator that icing might be present. Icing dramatically impacts turbine performance and often shuts turbines down completely.

Wind speed and direction, temperature, and relative humidity are all variables that are part of core NWP calculations and thus will be available at the temporal frequency required for power system modeling.

### Data for Solar Generation Estimation

Producing a first-order estimate of solar photovoltaic (PV) production requires global horizontal irradiance (GHI), while estimation of concentrated solar power production requires direct normal irradiance (DNI). Surface measurement density of these variables is quite sparse; therefore, they must be modeled. The two ways to do this with the coverage and accuracy required are through using NWP models or using models that derive these variables directly from satellite observations.[27] Both methods produce model data that are anchored to observations, but it should be understood that they are not direct observations; the nature and magnitude of uncertainties in the estimates vary and have implications when using them to determine expected output from solar generators. More details can be found in Appendix B.

Solar generation is also significantly impacted by panel backplane temperature, which is largely a function of ambient temperature and wind speed. Of course, measurements of temperature on the panels are by far the best estimate of panel temperature since it is heavily impacted by local microscale effects—a panel on a black roof will get hotter than one on a rack in a grassy field. But measurements typically are not available across the required time frames, so having somewhat accurate near-surface wind and temperature data in a convenient dataset is helpful in adjusting the output expectations. For existing plants, long NWP-based records can be tuned against actual observations.

Weather data indicating the presence of weather impacts that can dramatically impact generation estimates are also valuable. Indications of frozen precipitation (snow and ice) allow assessment of the risk that output estimates may be dramatically incorrect due to panels being covered by snow. And while it is not possible for an NWP model to predict wildfire ignitions, it is possible to estimate wildfire risk based on temperature, relative humidity, and

27 The data assimilation process in NWP performs a similar task of deriving irradiance data from observational satellite measurement for the NWP initial condition.

wind speed. Including diagnostics based on these variables would allow users of power system modeling to determine periods where modified inputs could be used to stress-test high-risk periods.

### Data to Estimate Gross Demand

As noted numerous times, temperature is the driving weather variable for demand. However, humidity, cloud cover, solar angle, and wind speed also contribute. These variables are measured at many surface observing sites and tend to be of higher quality and density in high population areas which also define load. The data are also always available in datasets derived from NWP as required variables in atmospheric modeling. It is important to note that raw NWP temperature data, even from reanalysis datasets, may differ from actual measured surface temperature observations, because they represent the average of the grid cell and so may deviate due to differences in model elevation and land surface characteristics relative to actual observations, especially if the grid cell is quite large. However, if it is necessary to augment observations with model data to extend the time series length, sufficient time series of measured temperature are nearly always available to bias-correct the model output.

### Data to Estimate Hydro Potential

The weather dependence for hydro power is ultimately manifested mostly in streamflow variability, which is a function of rainfall, snowfall and snowmelt, and ground-water charge. Therefore, precipitation amount and type and the temperature of snowpack all apply. Unlike weather impacting wind and solar generation, there is considerable delay between weather occurring and the effect on hydro capacity, and human-directed water management also plays a large role. Given that the data challenges for hydro are different from those of solar, wind, and load—being multi-sectoral and significantly dependent on hydrological modeling to capture the effect of precipitation, water inflow, and surface run-off, and because historical water inflow and independent stochastic selection methods are likely sufficient for handling weather impacts on hydro power in many power system applications—weather data for hydro will not be considered in this section.

### Data to Estimate Outage and Derate Probabilities and Other Weather Influences

Section 3, "Weather Inputs Needed for System Planning," outlined the weather affecting outages and derates of all electricity system assets. Near-surface variables for fields such as temperature, wind speed, and frozen precipitation are all needed. Weather observations probably provide enough coverage of the extent and duration of frozen precipitation to allow the impacts to be handled in power system models that are sophisticated enough to include it now or in the future, but NWP methods will produce an estimate of frozen precipitation in a convenient, easy-to-process gridded format, so it is recommended that these data be archived.

### Other Meteorological Data

#### General Meteorological Data Defining Atmospheric State

Output from NWP models contains data necessary to restart the NWP process with another model (or the same one with a different configuration) or to perform advanced post-processing like the GAN methodology. Thus, saving it could be useful for refining the output. The data are also useful for research and data-mining tasks that could inform power system modelers of trends and uncertainties in a model dataset. Therefore, if considerable investment is made to produce such datasets, it does not make sense to throw away all of the data not immediately of value to power system modeling. However, high-resolution models spanning continental-sized domains can contain tens of millions (possibly hundreds of millions) of grid points, each with a suite of variables, for every output hour. Thus, for long time series data, compromises must be made. Where modeling approaches are used, it is recommended that as much near-surface information as possible be archived, as well as data from levels typically used to analyze and characterize meteorological regimes. However, it does not make sense to be overly prescriptive here, and the exact definition is best left to a technical review committee (see Section 5).

## Forecast Data[28]

Any dataset representing forecasted estimates for use in unit commitment and dispatch functions of power system planning should contain all the information needed to estimate hourly load at any node and wind and solar at any generator as if it was being forecast for real-time operations. The forecast horizon needs to satisfy unit commitment and dispatch gate closure lead times (and ideally market function lead times as well), accounting for the time needed in an operational setting to produce the forecast. The forecast data need to have the same correlation with the "truth" dataset (in terms of accuracy) as current state-of-the-art for load and generation forecasting systems. Because load and generation forecast skill is evolving and system gate closures get updated, and because planning studies that model forecast uncertainty are a less common subset of power system modeling applications, it is recommended that a separate effort determine where it is valuable to produce wind, solar, and load forecasts as companions to ground truth datasets for system planning use, and whether it is best to produce them separately from an effort to develop a ground truth dataset.

## Recommendations

To support accurate modeling of wind speed and direction, datasets should have horizontal resolution sufficient to resolve the wind field and vertical resolution sufficient to resolve surface inversions, several levels within the rotor plane, and sharp inversions capping flows driven by phenomena such as sea breezes and topography. Meeting these requirements for wind will usually provide for improved representativeness of solar and temperature data as well when the same source of model data is used. As noted, vertical resolution is variable in NWP models; the selection of levels is beyond the scope here and needs careful consideration by experts based on the application. Horizontal resolution of 2 km or better is required if complex topography is present. The topography of the continental U.S. is complex enough from the Rocky Mountains west and in the Appalachian Mountains that a grid spacing of 1 km is recommended, although 2 km might provide passable data. While the Midwest likely does not need this level of resolution, splitting the country into different domains introduces its own problems (such as dataset seams) and is not recommended. In addition, since shallow cumulus clouds and deep convection are both common in the Midwest (and profoundly and rapidly impact wind and solar resource, as well as temperature), this also makes resolutions of 2 km or better desirable since modern NWP models can explicitly resolve these phenomena at these resolutions.

For some applications, power system modelers would like data that have a 5-minute time resolution to align with typical dispatch intervals. This would allow the intra-hour variability of load and renewable resources to be assessed in production cost models. However, many power system modeling efforts utilize hourly data to reduce computation time and make problems more tractable.

Assuming that NWP is used as part of the process to produce the necessary data, it is technically feasible to produce data at 5-minute intervals. Even at a relatively low resolution like 10 km grid spacing, most NWP models would be integrated at a time step of one minute or less. The higher the resolution, the shorter the time step needed to maintain numerical stability. However, there

28 To clarify, "forecast data" in this context refers to weather data used to prepare forecasts of load and wind and solar generation that are then used in power system models to reflect operational foresight relative to the estimates of the load, wind, and solar generation that represent the ground truth of what occurred for the same period.

are some caveats for outputting this frequently. First, shortwave and longwave radiation parameterizations in NWP models are performed less frequently than the dynamical time step, as they are computationally expensive, although newer schemes like the Fast All-Sky Radiation Model for Solar Applications (FARMS) developed by NREL allow for fast radiation calculations every model time step with minimal degradation in accuracy (Xie, Sengupta, and Dudhia, 2016). It should be noted that the reference configuration of the Mesoscale and Microscale Meteorology Laboratory's WRF-Solar® model (Jiménez et al., 2016) employs both FARMS as well as a traditional two-stream shortwave and longwave radiation parameterization scheme (a rapid radiative transfer model (RRTMG)) to capture aerosol-cloud-radiation feedbacks that are critical for accurate irradiance predictions, as well as improving the accuracy of radiation calculations between model time steps.

Second, data users should realize that NWP models represent average changes over grid cells and will not capture all the variability that exists. Third, if a reanalysis method is used for data synthesis, 5-minute output is unlikely to properly capture the evolution, as the frequency is higher than many of the observations being assimilated. Lastly, outputting high-resolution gridded data at 5-minute intervals can create input-output bottlenecks as data are written to storage, and it also dramatically increases the volume of data created.

Reconciling all of these trade-offs, below is a summary of the recommended specifications of dataset variables and spatial and temporal requirements.

Required data at a time interval of no less than 15 minutes, and horizontal grid spacing of 2 km or better:

- Wind speed and direction at 10 m, 25 m, 75 m, 100 m, 125 m, 150 m, 200 m, 300 m
- Temperature at 2 m, 10 m, 25 m, 75 m, 100 m, 125 m, 150 m, 200 m, 300 m
- Relative humidity at 2 m, 100 m, 300 m, or alternatively a post-processed icing risk field
- GHI, DNI, and diffuse horizontal irradiance (DHI)

Recommended data at an interval of no less than hourly, with a grid spacing of 2 km or better:

- Accumulated rainfall and snowfall, and precipitation type (hourly)
- All other model surface data and 2D fields
- All data from native model levels below 1 km above ground level. This will be useful for academic and applied research.
- Primary prognostic data (air temperature, pressure, water vapor mixing ratio, horizontal and vertical wind components) interpolated to standard meteorological pressure levels from the surface to 300 hPa (1000, 925, 850, 700, 500, and 300 hPa)[29]

### ATTRIBUTE 2: Covers Multiple Decades with Consistent Methodology and Is Continuously Extended

Weather input datasets need to cover a climatologically valid time span if they are to capture the inherent variability in the atmosphere. Typically, atmospheric scientists have considered a 30-year period as sufficient to capture most of the variability that is expected. However, even longer periods are required to capture the tail—extreme weather events that are critical when assessing power system reliability. Ideally the longest datasets possible are desired to capture as much variability as possible and derive information about events in the tails of the distribution.

The other side of the coin is the impact of climate change. While datasets going back 60 years or more are probably representative of variability and conditions for wind and solar resources, there is no question that overall temperature distributions have changed, and this is likely beginning to impact other weather fields. Longer datasets are more likely to reveal climate change signals, and datasets that are continuously extended in the future are the best way to ensure that trends can be detected and evaluated as they develop.

Datasets covering large areas for long durations are essential to capture the full range of possible conditions and long-term trends. NWP methods are core to

29 Prognostic variables provide information about atmospheric state that can be used to both describe the state and predict the future state. These are the most useful variables that meteorologists can use to understand how the atmosphere is evolving and are the basis for performing forecasting tasks.

producing these. Importantly, the data availability to produce high-quality initial conditions for NWP modeling has been enabled by weather satellites, which is important to recognize when deciding how far into the past to develop power system weather inputs using NWP modeling techniques. While we want the longest dataset possible, we do not want to utilize poor-quality data—which generally means going back no further than 1990. The year 1978 is generally considered the beginning of the satellite era for weather prediction purposes, and our ability to monitor the detailed atmospheric state has improved dramatically since then. (More details about the impact of weather satellites can be found in Appendix B.) As new remote-sensing instruments are deployed, initial condition quality continues to improve, albeit much more slowly now. When using datasets created with NWP methodologies, it should be recognized that the quality of the data is a function of the observations going into them, as well as the model resolution. Even when using modern models to assimilate data, earlier periods, especially before the satellite era, contain higher biases and deficiencies because of this. These may be difficult to detect because the observing network contained much less detailed information than the grid data.

**Extreme caution should be exercised if using NWP data archived from operational forecasting as weather inputs to energy system modeling. The model configuration used to generate the operational forecast data is unlikely to be consistent throughout the entire period of interest. In addition, operational models prioritize timeliness over producing the best possible initial condition and predictions.**

Another possible source of inconsistency in a multi-decadal dataset is the use of non-standardized model set-ups. NWP data that are archived from operational forecasting are sometimes used as weather input to energy system modeling. However, this should be done with extreme caution, for two reasons. First, the model used to generate the operational forecast data is unlikely to be consistent throughout the period of interest. Operational models are regularly updated to incorporate new developments from the research community (for example, improved parameterizations) or to increase

resolution or size of the region being modeled as computational power increases. Each time an update is made, model biases and error levels may change. This might seem trivial, but consider that above cut-in, within the low-speed region of the wind turbine power curve (see Figure 12, p. 65), a 2 m/s change in wind speed can triple the power output. Thus, a small systematic change in wind speed bias when the model configuration changes can appear as a significant change in average generation. Aside from potentially changing the supply and demand balance in a power system model, a power system modeler with no visibility to model configuration changes may attribute a shift in average wind output like this to climate change.

The second issue with using operational model output is, as noted in Section 3, data assimilation in operational models is optimized to starting model integration at the earliest possible moment so that the forecast is timely, rather than to producing the best possible initial condition. Thus, operational model output is generally inferior to output from the same NWP configuration performed retrospectively utilizing all available data to produce the best initial condition (as, for example, in reanalysis).

Another consideration when determining dataset length is that the longer the dataset, the more computational resources are needed to create it, and subsequently refresh it (see Attribute 6, future-proofed).

### Recommendation

Datasets produced as weather inputs to system planning models should extend back to at least 1990 and should use a consistent methodology throughout. Ideally, datasets should go back as far as possible, but documentation should be clear about the increased uncertainty in earlier years, especially prior to 1978. Longer datasets also require more computational resources, and if trade-offs need to be made between producing data prior to 1990 and other attributes like resolution, future-proofing, and continuous extension, then limiting the historical duration is preferred.

Datasets aimed at power system modeling users should be extended in an ongoing fashion using the same consistent methodology. Continuous extension is essential and is far more important than extending the record back many decades. This will ensure that the latest gridded data are always available for power system modeling and to compare against new observations (particularly at renewable resource sites) that can be used to validate the model performance. Continuous extension of the dataset will also allow any trends in climate to be observed and will provide accurate, easily accessible weather information to analyze outages and future extreme events.

## ATTRIBUTE 3: Coincident and Physically Consistent Across Weather Variables

Given the increasingly weather-correlated behavior of supply and demand, time series variables must be coincident in time to maintain correlations between related phenomena that impact supply, demand, and infrastructure risks. Assuming that the instruments used are reasonably accurate, observational data achieve this; an observation taken with one observing platform will be consistent with another taken at the same time. However, as has been noted, there is an insufficient density of observational data to meet power system modeling needs, so data must be synthesized with models and the output variables must be physically consistent. It is vital to realize that if different weather variables used as inputs for power system modeling are sourced from different meteorological models, or if the models are not physics-based, it is unlikely that the time-coincident data will be physically consistent *even if* the inputs to the meteorological models are the same consistent set of weather observations. The inconsistency can lead to combinations of weather variables that are not physically reasonable and combinations that have a different likelihood of occurrence in the synthesized time series than in reality.

**Given the increasingly weather-correlated behavior of supply and demand, time series variables must be coincident in time to maintain correlations between related phenomena that impact supply, demand, and infrastructure risks.**

This data incongruency must be minimized because it will produce incorrect distributions of net load and may result in non-plausible outcomes. This in turn can lead, for example, to sub-optimal portfolio optimizations in capacity expansion models or, in the case of tails in the distribution events, can produce resource adequacy findings that are inconsistent with reality. For instance, if irradiance at a location is over-estimated because it is estimated using an imperfect relationship with irradiance at another location, while another method over-estimates wind and temperature on cold cloudy days, the compounding of the errors will lead to an overly optimistic supply and demand balance that is considerably different from reality, even though the data are all coincident.

When the same physics-based model configuration is used for all required variables, these incongruencies will not occur, though of course the model data themselves can differ from reality as discussed at length in preceding sections. However, if the meteorological models are statistical in nature, or even if two different physics-based models are used to source different variables, inconsistencies will occur. For statistical models the inconsistencies are likely to be profound. If two different physics-based models are used, the incongruency will be smaller but may still be significant, especially if output, while time-coincident, comes from models that have simulated different lengths of time from their starting point or have significantly different resolution.

For instance, such an inconsistency can result from the common combination of data from the National Renewable Energy Laboratory (NREL) Wind Integration National Dataset (WIND) Toolkit with the NREL National Solar Radiation Database (NSRDB). The WIND Toolkit data come from a model that runs multi-day simulations. That is, they begin with an initial condition and then the model predicts several days' worth of data. Meanwhile, the NSRDB data are created using a different physical model (the Physical Solar Model) that processes new satellite data every hour with prediction of future times. Although the large-scale weather pattern present in the simulations used to produce the WIND Toolkit is nudged back toward observations throughout the prediction period,[30] it is possible for local-scale wind (and cloud) structures that the model develops in response to topography to become inconsistent with the cloud fields in NSRDB. This incongruency is probably nowhere near as serious as that which could arise from using different statistical models; however, no literature could be found by the ESIG project team that explores its magnitude. Since these two sources of data are frequently used together by power system modelers, the impact on the balance of supply and demand should be investigated if the datasets are to continue to be used together. This serves as a cautionary example of how even apparently coincident datasets that are commonly used together might not be physically consistent.

#### Recommendation

Observations or physics-based models, as opposed to statistical models, should be used wherever possible. When combinations of observations and one or more physics-based models are used, even though the times are coincident, some validation must be performed to ensure that the resultant combinations of variables produce physically reasonable outcomes and that the differences between these outcomes and those seen in reality are understood and quantified.

**Observations or physics-based models, as opposed to statistical models, should be used wherever possible. When combinations of observations and one or more physics-based models are used, even though the times are coincident, some validation must be performed.**

### ATTRIBUTE 4: Validated with Uncertainty Quantified

Output data produced by any type of model, even if the model inputs are well-quality-controlled observations, must be robustly validated and the uncertainty must be quantified. The data should not be expected to perfectly match actual observations, but the degree to which they do not needs to be known for each variable of interest—in the case of power system modeling, primarily wind, irradiance, and temperature—and as a function of

30 See "Deriving Downscaled Regional Datasets" in Section 2.

location, elevation, time of day, and time of year. In addition, it is important to pay particular attention to errors and biases that occur in weather regimes where the combination of moderate to high load and low wind and/or solar resource produces high net loads. If these scenarios coincide in a systematic fashion with significant errors or biases in the modeled variables contributing to supply and demand, then this must be considered when stating confidence in the study results that use the weather inputs. The issue should then become part of a feedback process to improve the weather inputs either through post-processing or improvement of the underlying model.

Aside from validating the data and quantifying the possible magnitude of errors, if using NWP, it is worthwhile to create ensemble datasets. Here, the same initial data are passed through different models or the same model with different configurations, or slightly perturbed versions of the initial data are run through the same model (or some combination of the two). Doing this produces multiple potential realizations of the atmospheric state, forming an envelope of "truth." Even where ground truth observations are not available for validation, the spread of the data within the different ensembles provides a measure of the uncertainty of the model data and can also be utilized downstream to run several instances of a power system analysis and examine the spread of outcomes.

#### Recommendation

Datasets produced for the purposes of power system analysis should include validation as a core part of the project to create them. This validation should pay particular attention to high-risk scenarios, for example, weather regimes yielding resource adequacy concerns, where biases and errors could lead to incorrect conclusions. For example, a resource adequacy study is less concerned about the accuracy of average annual capacity factor of wind and solar resources, and more concerned about the accuracy—and associated probabilities—of sustained

low-wind and low-solar periods. While a dataset that accurately predicts annual capacity factors but not outlier events may be appropriate for a solar or wind plant developer, it is insufficient for power system reliability analysis. In addition, new model datasets should use ensemble techniques to produce more than one estimate of weather inputs so that sensitivity of the power system models to weather inputs can be evaluated.

**Dataset validation should pay particular attention to high-risk scenarios, such as weather regimes yielding resource adequacy concerns. While a dataset that accurately predicts annual capacity factors but not outlier events may be appropriate for a solar or wind plant developer, it is insufficient for power system reliability analysis.**

### ATTRIBUTE 5: Documented in Detail and Transparently

Documentation of weather datasets used as inputs for power system studies is critical. It must be detailed and cover the items identified below. It should also transparently highlight the strengths and weaknesses of the methodology employed and provide guidance regarding how weaknesses may impact power system modeling efforts.

#### Recommendations

Documentation should include:

- Everything needed for an independent entity to recreate the data, including model configuration and input data sources. This also allows outside entities to test and critique the methodologies used.

- Validation results and measures of uncertainty, including ongoing validation as the dataset is extended.
- An accessible tutorial that educates non-meteorologist users in how the data were produced. The tutorial should help users understand the differences they should expect between the dataset and the actual truth that could theoretically be measured if a microscale observing network was possible. It should make it clear that any gridded dataset will be imperfect and describe the dataset's limitations and possible flaws.
- A clear description of the format of the dataset so that the necessary information can easily be parsed by end users.
- Descriptions of each variable provided along with advice about the known issues regarding the modeling of each variable that might be relevant to power system modeling. For example, for data that have been produced by regional downscaling using an NWP model, during cold periods with strong surface inversions the inversions tend to be eroded faster than in reality, resulting in time series data of temperature, wind speed, and low-lying cloud/fog (and thus irradiance) that progressively drift away from reality until the model initial condition is refreshed. This will obviously impact estimates of wind and solar generation and load that are derived from the data.

**Plans should be made—and budget assigned—to update the entire database at the point where improved science and methods can produce a materially more useful dataset.**

### ATTRIBUTE 6: Future-Proofed

Plans should be made at the beginning of any project producing power system planning weather inputs to make sure that, within reason, it continues to represent the state of the art.

#### Recommendation

Aside from continuously extending any dataset produced for power system modeling, plans should be made—and budget assigned—to update the entire database at the point where improved science and methods can produce a materially more useful dataset. For example, the output from the existing method could be compared to that of the very latest methods each year for a sample of the dataset. When output from a test run shows 10 percent improvement of core metrics, the entire dataset would be recreated using the latest methods.

### ATTRIBUTE 7: Publicly Available, Easily Accessible, and Standardized

To move toward the next generation of power system modeling techniques, quality weather inputs are essential. As we have seen, producing such datasets, while possible, is no small task. The data volume will also be very large. Therefore, those datasets that are produced should be broadly available, easy to access, and provided in a standardized manner.

#### Recommendations

- Create a data standard for weather inputs. The standard would define the format of the data and indicate which data are mandatory and which are optional. It will also put the data into three categories:
  - Data that will be routinely used by power system modelers. This should include everything that is needed, but no more, in order to minimize complexity and data volume. This will largely be fields like wind, temperature, relative humidity on geometric height levels above ground level, and two-dimensional fields like precipitation (amount and phase), and surface solar irradiance.
  - Data that may be needed for more in-depth analysis of the power system as a function of meteorological conditions likely to be of interest to those doing a deeper dive. This will be a more complete set of meteorological variables available on pressure levels and/or the native NWP model vertical coordinate. The recommendations for Attribute 1 provide a possible list for a technical review committee to consider as a starting point.
  - All other output deemed worth keeping relative to the cost of archiving it.

## Need for a Dedicated Team to Provide Curation and Advice

In the transition to a weather-driven electricity system, meteorological data are crucial, and the production and validation of datasets for use in power system modeling is complex and expensive—though many orders of magnitude less costly than the components of the power system being analyzed or the value that quality data will provide. Further, while the goal is to make weather input data as easy to use as possible, the data will sometimes contain significant differences relative to actual conditions, and understanding the nature of these differences is a specialized task.

For these reasons, and because there will be a need for technical and logistical management of processes to continuously update and periodically improve datasets, it is recommended that a dedicated team curate the weather dataset or datasets produced for use in the public domain. The team would manage data distribution, educational outreach, and general advice on data use; provide limited advice to specific users (while users requiring extensive assistance would be put in touch with a network of experts); collect user feedback and evaluate requests for changes and additions; perform validation and quality control; and flag issues associated with use of the data that could affect the conclusions drawn from power system modeling efforts that use the data. The team would also monitor the availability of other non-public data in the sphere and provide information to users as available.

The curation team could provide a certification service that evaluates how public weather input data are utilized for important decision-making and provide a seal of approval.

## Comparison of Requirements and Currently Available Data

Here we introduce the most pertinent data currently available for power system analysis and score them against the seven required attributes outlined above. Table 2 (p. 76), summarizes some of the most useful available datasets, including some that have recently been introduced or are currently in development, and indicates where they do and do not meet the required attributes.

Using the United States as the area of interest, if we apply the criteria from earlier in this section to the available datasets listed in Table 2, then most of the datasets are eliminated, because the complex topography from the Rocky Mountains westward and the Appalachian Mountains eastward requires geographical spacing of 4 km or less to represent many of the phenomena driving renewable resources. As is detailed below, the longer, more frequently updated reanalysis datasets like ERA5 (with 30 km grid spacing) are nowhere close to providing the required spatial resolution, and observations are far too sparse.

**If we apply the seven criteria to the available datasets listed in Table 2, most of the datasets are eliminated, because the complex topography from the Rocky Mountains westward and the Appalachian Mountains eastward requires geographical spacing of 4 km or less to represent many of the phenomena driving renewable resources.**

The datasets that remain are the NREL WIND Toolkit, the NREL NSRDB, and the operational forecast archive of the National Oceanic and Atmospheric Administration's (NOAA's) High-Resolution Rapid Refresh (HRRR) model. Aside from there being limited options, Table 2 shows that the useful datasets do not completely meet the other required attributes for power system modeling weather inputs. Only NSRDB is longer than a decade and regularly updated. It is unclear how well any of the datasets represent the renewable generation output in different locations or how well they capture the variability, because detailed, systematic validations against actual wind and solar generation output have not been performed. A critical need is to produce updates at regular intervals of one year or less to continuously extend the dataset using the exact same methodology, and only NSRDB meets this.

Regular updating, in addition to providing data to model the most recent (and thus, often the most relevant) periods, meets another need: to provide time series data that can capture the evolving signature of climate change. This may not be a critical need today, but the research

TABLE 2

## Summary of Current Power System Modeling Weather Input Data Sources

| | Spatial Resolution | Temporal Resolution | Length | Continuously Extended | Correct Variables/ Levels | Coincident and Coherent | Validated/Uncertainty Quantified for Power System Use | Detailed Documentation | Future-Proofed | Availability/ Ease of Access | Curation and Advice | Region Covered |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **MERRA-2**[a] | ~60 km | 60 min | 1980–present | Yes | Yes/No | Yes | No | | Probably | | Basic | Global |
| **ERA5**[b] | ~30 km | 60 min | 1940–present | Yes | Yes/No | Yes | Some | | Yes | | Good | Global |
| **HRRR**[c] | 3 km | 15 min | 2014–present | Yes | Yes/No | Yes/No | No | | Unideal | | Basic | U.S. |
| **WIND Toolkit**[d] | 2 km | 5 min | 2007–2014 | No | Yes/Yes | Yes | Yes | | No | | Basic | Various |
| **WTK-LED**[e] | 2 km/4 km | 5 min | 3 year/ 20 year | No | Yes/Yes | Yes | Not yet | Not yet | No | Unknown, dataset not yet available | | Various |
| **NSRDB**[f] | 4 km/ 60 km | 30 min | 1998–present | Yes | Yes/No | Solar only | Yes | | Yes | | Basic | Most of globe |
| **CERRA**[g] | 11 km/5.5 km | 60 min | 1980–present | | No/Yes | No solar | Yes | | Possibly | | Basic | Europe |
| **CONUS404**[h] | 4 km | 60 min/ 15 min (precip) | 1980–2020 | No | Unknown/ Probably | Yes | Not the intended use | | | | | Continental U.S. |
| **BARRA**[i] | 12 km/ 1.5 km | 60 min | 1990–2019 | No | Yes/ Probably | Yes | | | | Fee-based | | Australia/ New Zealand |
| **Public Observing Networks**[j] | Non-uniform, variable density | 1 hr or less | Variable | Yes | Yes/No | Mostly | Varies. Not for power systems | Varies | Usually | Usually easy | Varies | Global |
| **Renewable Energy Project Data**[k] | Non-uniform, variable density | Usually minutes | Variable but rarely more than 10 years | Varies | Yes/ Usually | Yes | Usually | Varies, but usually poor | Varies | Usually poor | Usually none | Very limited |
| **Proprietary Statistically Derived VRE Shapes**[l] | Non-uniform, variable density | Usually hourly | Variable. Rarely reliable long records. | Varies | Usually incomplete | No | Partial | See note | No | | None | Very limited |

■ Fully Met ■ Close to Being Met ■ Partially Met ■ Met in a Very Limited Way ■ Not Met at All ■ Not Enough Info. for Determination

**Summary of the most applicable datasets globally that are (or can be) used to provide weather inputs for power system analysis tasks, especially for providing estimate of site-level generation, and concurrent weather-driven load and generation outage risks. The degree to which the needs of each column heading are met is estimated with color coding. See documentation for each dataset for all details.**

Source: Energy Systems Integration Group.

**See the figure footnotes on the following page.**

### Table 2 Footnotes

a MERRA-2. The resolution of MERRA-2 (Modern-Era Retrospective Analysis for Research and Applications) is typically insufficient for weather input use in power system analysis.

b ECMWF (European Center for Medium-Range Weather Forecasting) Reanalysis v5. ERA5 has insufficient resolution to diagnose regional or local weather, yet it is widely used for power system analysis.

c High-Resolution Rapid Refresh (HRRR). The HRRR is an operational model and therefore configured to balance accuracy with speed. It undergoes regular configuration updates, so model skill is changing in time. Occasionally, major updates may occur that can create step changes in model biases.

d Wind Integration National Dataset Toolkit. The years 2007 through 2013 cover the U.S., and 2014 uses a different configuration that includes Mexico and Canada.

e WTK-LED (WIND (Wind Integration National Dataset) Toolkit Long-term Ensemble Dataset) is still in production, and there is little current documentation. There are three years at 2 km, and 20 years at 4 km that are downscaled to 2 km with the machine learning GAN (generative adversarial network) approach. In addition, one year of ensemble data is being produced to aid in quantifying uncertainty.

f NSRDB (National Solar Radiation Database). Irradiance resolution is 4 km. Other variables are interpolated from MERRA-2 data using an unvalidated method. These data are generally not appropriate as weather inputs to power system analysis, forcing NSRDB to be used in combination with other datasets, which creates consistency issues.

g CERRA (Copernicus Regional Reanalysis for Europe). Ensembles at 11 km. Does not include all weather variables.

h CONUS404. A 4 km, long-term regional hydroclimate reanalysis over the conterminous United States (CONUS), 1979–2020. Developed by the U.S. Geological Survey to assess hydrological climatology, but may be useful to repurpose for power system analysis.

i Bureau's Atmospheric High-Resolution Regional Reanalysis for Australia. A 12 km reanalysis with 1.5 km domains over four cities in Australia.

j Many public observing networks exist with variable density, quality, and applicability.

k Observed data from renewable energy facilities is of course applicable to variable renewable energy, but quality varies from site to site and is typically proprietary. Data across the upper portion of the rotor sweep is often not measured.

l Often used proprietary data. The same shape is often assumed across broad areas. Validations are not rigorous, and methodologies are usually not fully documented in a transparent way. Output usually includes only a single weather variable.

required to extract climate signals will be complex and challenging and will require many years of consistent data created for the purpose. Lastly, elements of the usable datasets must be combined to provide data to represent both wind and solar generation reasonably well. Hence, there is some risk of physical inconsistency between weather variable fields. While this effect is probably reasonably small, it should be investigated if the sector plans to continue to utilize data from different sources.

### ERA5 and Other Global Reanalysis Datasets

There are several global reanalysis datasets. (See Section 3, "Weather Inputs Needed for System Planning," for a detailed description of this type of dataset.) The most well known are the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) Reanalysis,[31] the National Aeronautics and Space Administration's Modern-Era Retrospective Analysis for Research and Applications (MERRA), MERRA-2 (an update of MERRA),[32] the Interim ECMWF Atmospheric Re-Analysis of the Global Climate (ERA-Interim),[33] and the fifth-generation ECMWF Atmospheric Re-Analysis of the Global Climate (ERA5).[34] These datasets provide an estimate of all the main variables that define the state of the atmosphere, as well as the state of the interface with the land and ocean surface on easy-to-use three-dimensional grids for every time interval in the dataset. Data include latitudinal and longitudinal wind components, temperature, humidity, liquid and frozen water content, and geopotential height in three dimensions; two-dimensional fields like irradiance, accumulated precipitation and snowfall, soil and water temperature, and model topography and land use; and often many other derived fields. The datasets span multiple decades and have a temporal resolution between one and six hours. Each is or was regularly extended with the latest weather data until deprecated by a subsequent improved dataset designed to take its place.

The focus here is on ERA5 (which was preceded by ERA-Interim), since for the purpose of weather inputs to the energy sector, it is far superior to the others. The NCEP/NCAR Reanalysis was one of the earliest available global reanalysis datasets and is still being regularly extended, but its resolution is far too coarse for the needs of power system modeling. MERRA-2, which replaced MERRA, has a finer resolution than either MERRA or the NCAR/NCEP reanalysis, but is still much too coarse to use in any capacity for power system modeling without downscaling first. (Note that MERRA-2 provides the meteorological companion dataset to NSRDB, discussed below.)

ERA5 is a global reanalysis dataset on which important meteorological fields defining the state of the atmosphere are represented on a 0.25°x0.25° grid, with 137 terrain-following vertical levels. (Section 3 includes an explanation of terrain-following coordinates.) The data that are typically served to users are interpolated onto a regular Cartesian grid with regular 30 km spacing. The reanalysis is performed using the ECMWF Integrated Forecasting System (IFS) model and its 4D-Var data assimilation system, which are widely considered best in class. The output has been broadly validated and is found to produce meteorological fields that are representative of observations, especially in simple terrain. The archive extends back several decades and is regularly updated. ECMWF commits significant resources to quality-controlling the output. The modeling system is clearly documented, and the data are easy to access for any region of interest on the planet. For these reasons, ERA5 is an attractive dataset that is widely used, including for power system planning studies.

However, while ERA5 is unquestionably the best global reanalysis dataset currently available, it is not a panacea. Average validation statistics are very good, but the horizontal grid spacing of 30 km is insufficient to produce detailed meteorological fields present in complex

31 https://www.ncei.noaa.gov/products/weather-climate-models/reanalysis-1-2.

32 https://gmao.gsfc.nasa.gov/reanalysis/MERRA/ and https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/.

33 https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-interim.

34 The landing page for ERA5 information is https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5. From here there are links to detailed documentation. Documentation for the other reanalysis datasets can also be found online.

While ERA5 is unquestionably the best global reanalysis dataset currently available, and average validation statistics are very good, the horizontal grid spacing of 30 km is insufficient to produce detailed meteorological fields present in complex topography.

topography, fields that are crucial to resolve for estimating renewable generation in these areas. As an example, a validation study including complex topography in southern Europe showed that variables like wind speed can exhibit average correlation coefficients in the range of 0.5 when compared to observations (Molina, Gutiérrez, and Sánchez, 2021). Poor correlations between the reanalysis data and observations are also found for other important variables such as temperature and precipitation when the combination of weather and terrain produces phenomena like valley cold pools, and large deviations from reality have been observed west of the Rocky Mountains in the United States.[35]

The main way in which ERA5 fails to meet the criteria for a long-term historical dataset for use in power system planning is its horizontal grid spacing (see Table 2, p. 76). Other, less serious limitations are that (a) the output is only available at hourly intervals, and (b) easily accessible data for near-surface levels are only available at 10 m and 100 m. Because of ERA5's low resolution, using it to estimate renewable generation for power system modeling can produce large errors, especially in regions of complex terrain, which are often good locations for renewables development. (See "The Impact of Model Resolution" on p. 25.) However, because ERA5 is such a good dataset overall, it is possible that it might be used as the input to downscaling methodologies, and it is valuable in regions with simple topography.

## The High-Resolution Rapid Refresh Model (HRRR)

The HRRR is an operational limited-area model that runs on a rapid update cycle and covers the continental United States. New observational data are assimilated every hour, followed by a short forecast run (currently either 18 or 48 hours ahead, depending on the time of day), meaning that a new analysis is available every hour. Because the model is high resolution (currently 3 km grid spacing) and tethered to reality with frequent data assimilation, it offers many of the benefits of reanalysis but with high resolution. However, the fact that it is an operational model is a major drawback. To get the model refreshed with new observations and update the short-term forecast, strict data cut-off times need to be enforced. (See the discussion of data assimilation in Section 2, "Meteorological Data Fundamentals for Power System Planning.") Thus, many fewer observations will make it into the analysis than in the case of, for example, ERA5.

The HRRR is high resolution and tethered to reality with frequency data assimilation; however, the fact that it is an operational model is a major drawback, and the data are not future-proofed.

In addition, the HRRR model configuration and code are updated quite frequently, which might seem like a good thing, but it introduces changing biases into the time series data. At some point a major model change is likely to happen, such as an increase in horizontal resolution or a change in dynamical core, and this will create a data discontinuity. Lastly, the model has only been running since 2014, so the time series is too short for use in power system modeling.

Despite these flaws, the HRRR may be a good choice to provide weather inputs for some modeling exercises in which it is not essential to have a long time history to cover all possible conditions. Possible examples are renewable integration studies within the continental U.S. that aim to study periods since 2014, capacity expansion studies, and perhaps production cost modeling studies focused on reserve and flexibility needs. However, resource adequacy studies will require longer and more consistent time series data than are provided by this dataset.

35 As seen in unpublished client work performed by Justin Sharp of Sharply Focused that contrasts reanalysis datasets with observations.

## The WIND Toolkit

The NREL WIND Toolkit dataset was produced specifically to provide weather inputs to wind integration studies.[36] The team that created it went to significant lengths to tune the model configuration so that wind speed autocorrelation and spatial covariance accurately represented the scales being examined. They also chose a 2 km grid spacing to ensure that most weather features important to wind generation were resolved. Data are output at 5-minute intervals to provide the granularity needed to resolve wind ramping events. The WIND Toolkit is unique in that a companion dataset containing "forecasts" was also created. For each hour in the WIND Toolkit output, there is an accompanying set of values that represents what forecasts of the weather at that hour would be for different lead times that correspond to power systems' operational gate closure times. The weather forecasts were then used to produce power forecasts at thousands of possible wind generation sites. The forecasts were tuned to have a similar skill to state-of-the-art forecasts. Of course, forecasting has improved since 2014, so the skill of these forecasts is lower than is possible today.

**The skill of the WIND Toolkit's model is good relative to what can be expected from NWP, but the differences are large enough to matter in power systems applications. The most significant issue is an overall high bias in the wind speed.**

In one validation study the dataset was compared to wind observations located on tall meteorological towers at 13 sites around the U.S. (Draxl et al., 2015). The comparisons were reasonably good but by no means perfect. The daily shapes of the wind averages showed some differences between the model and observational data, as did the wind speed distributions. The skill of the model is good relative to that expected from NWP, but the differences are large enough to matter in power system applications. The most significant issue is an overall high bias in the wind speed. In another validation (King, Clifton, and Hodge, 2014), wind speeds from the WIND Toolkit at the locations around 284 real and hypothetical wind plants were used to calculate wind power at each plant. The "plants" were designed to represent either existing wind generation facilities or places throughout the country that were reasonable possibilities for future wind plants. Among other comparisons, the aggregate capacity factor for existing plants in the Midcontinent Independent System Operator (MISO) and Electric Reliability Council of Texas (ERCOT) territories was compared to power data derived from the WIND Toolkit. The simulated aggregate output was found to be reasonably consistent with reality, but again, there were significant differences in daily output shape and in energy volume. Of most concern was a tendency to over-predict the wind speed, yielding capacity factors that were 5 to 10 percentage points too high when aggregated across broad U.S. regions like ERCOT and MISO.

This finding points to how critical it is not only to produce an easy-to-use dataset, but also to ensure that it is validated in detail. While it would be desirable for the WIND Toolkit data not to exhibit this bias, at least the bias is documented, which is not the case for many other datasets. It is also important to make sure that inaccuracies are communicated. While the over-prediction has been noted by several users of the dataset, it is perhaps not as widely broadcast as it should be. The WIND Toolkit data contain significant errors during some critical weather regimes. For instance, Sharp (2022) found that during periods of low wind resource across the entire western U.S., the WIND Toolkit often greatly over-predicted the wind speed in the Pacific Northwest, yielding generation estimates for the large amount of wind in the Bonneville Power Administration balancing area that were much too high. These errors are due to NWP models often struggling to represent stable boundary layers and mixing momentum to the surface from aloft too quickly during cold stable weather.

The output of the original WIND Toolkit covers the period from 2007 through 2013. An additional year, 2014, was added later but used a different model set up, and the 2014 data have different biases. Thus, the dataset clearly does not meet the multi-decadal requirement, the requirement for regular extension, or the requirement for consistent model configuration.

36 https://www.nrel.gov/grid/wind-toolkit.html.

The objective when the WIND Toolkit was produced was to provide the best wind inputs to integration models, and the NWP model used was tuned to do this. At the time it was produced, solar predictions from the model used were not particularly good. Therefore, when using the WIND Toolkit data, a companion needs to be found for solar data. Usually the NSRDB is used, but this brings up some of the issues mentioned above for Attribute 3 (coincident and physically consistent across weather variables).

Despite its flaws, the WIND Toolkit is still one of the best available datasets for providing wind inputs to power systems models and is widely used. However, the flaws highlight the importance of validating data before use, and of taking the findings into account so as not to draw erroneous conclusions. In addition, its limited length means that users will often seek to extend the dataset using statistical methods. This needs to be done with great care (see Section 6, "Guidance for Using Existing Weather Inputs," for details).

The WIND Toolkit is now rather antiquated, and NWP modeling has advanced considerably since the Toolkit was produced because of general advancements and targeted programs like the Wind Forecast Improvement Projects and the Solar Forecast Improvement Projects funded by the U.S. Department of Energy. Subsequent projects have extended the geographical scope of the WIND Toolkit data to Canada and Mexico as well as several Asian locations using a similar methodology. New projects are now underway to create the WIND Toolkit Long-term Ensemble Dataset (WTK-LED) using updated models. This will feature three years of 2 km grid spacing simulations over the continental U.S. and Alaska with 5-minute output, and 20 years over North America at an hourly output. The longer time series will then be downscaled using the GAN machine learning approach described in Section 3 to ultimately provide 20 years of 2 km output with 5-minute temporal resolution. In addition, an ensemble of model runs was generated for 2018, and this is used to provide uncertainty quantification. A limited validation has been performed that compares the WTK-LED wind speeds to lidar observations taken at a wind plant in flat terrain in Oklahoma and two lidars offshore from the East Coast (Pronk et al., 2022). The validation also compares these observations to the ERA5 dataset in order to assess the value of the WTK-LED relative to existing data. The validation indicates that WTK-LED-predicted wind speed profiles show a limited negative bias offshore (~ -0.5 m/s) and a slight positive bias at the land-based site (~ +0.5 m/s). ERA5 shows a significant negative bias at both locations (~ -1 m/s), with a larger bias at the land-based site, but ERA5 outperformed the WTK-LED in terms of the centered root-mean-square error (cRMSE) and correlation coefficient, for both the land-based and offshore cases, in all atmospheric stability conditions.

**Despite its flaws, the WIND Toolkit is still one of the best available datasets for providing wind inputs to power systems models and is widely used. However, the flaws highlight the importance of validating data before use, and of taking the findings into account so as not to draw erroneous conclusions.**

Work on the WTK-LED is ongoing, so there are few published results at this time. It will be particularly interesting to see how well data from the GAN downscaling approach compare to corresponding raw NWP output and how both compare to actual field observations. If the project is successful and the validation shows accurate results, the new dataset would meet most, though not all, of the criteria for power system weather inputs. The main issues would be a lack of ongoing extension, lack of future-proofing, and lack of dedicated curation. In addition, the project is designed explicitly to produce wind data, so it is unclear whether it will produce useful concurrent solar irradiance data or whether a different dataset will need to be used for this, potentially yielding physical consistency issues.

### The National Solar Radiation Database (NSRDB)

NREL's NSRDB is a database of solar irradiance that covers the period 1998–2021 (as of July 2023).[37] It is extended annually to cover the previous year. The data currently use the Physical Satellite Model (PSM) to derive historical global horizontal, direct normal, and diffuse horizontal irradiance.[38] At the time of writing, the data for the U.S. are available for 4 km x 4 km grid cells for the period 1998–2021 and for 2 km x 2 km grid cells for the period 2019–2021. Output is available at 30-minute intervals throughout the period of record and at 5-minute intervals from 2019 onward. In addition to the United States, the NSRDB has been extended for several other countries.[39] The geographical and temporal resolution of these extensions varies depending on the available satellite data in each area.

**In addition to being regularly extended, the NSRDB data are periodically refreshed throughout the entire period of record to incorporate new methodologies and improved inputs. This is the type of future-proofing that is needed for weather inputs to be most useful to power system models.**

Research and development to further improve the data accuracy and usefulness is ongoing. Satellite observations are also improving as more advanced instruments are deployed. Thus, in addition to being regularly extended, the NSRDB data are periodically refreshed throughout the entire period of record to incorporate new methodologies and improved inputs. This is the type of future-proofing that is needed for weather inputs to be most useful to power system models. The one drawback of updates to NSRDB is that new instruments have only limited value for periods prior to their deployment.[40] However, to ensure a consistent record, when the entire record cannot be refactored as a result of an update, the old version is still provided. For instance, 2 km data have been available since 2019, but 4 km data are still provided as well so that they are consistent with the rest of the dataset.

The NSRDB data have been validated against surface observations (Buster et al., 2022; Habte, Sengupta, and Lopez, 2017; Sengupta et al., 2015b), but there is a lack of publicly available, high-quality surface solar radiation measurements in the U.S. Seven, nine, and 20 sites were compared in the 2015, 2017, and 2022 studies, respectively. In addition, comparisons of point measurements at surface stations to the pixels in the NSRDB is not really an apples-to-apples comparison. Long-term biases at the seven stations compared in the validation study were relatively small, so there is reasonable confidence that the overall values derived with the method are relatively accurate for monthly and annual averages. However, correlation between hourly and sub-hourly observations and NSRDB data is poor. All three reports show significant biases with overestimates on clear days and underestimates on cloudy days, as well as MAE/RMSE metrics that can be higher than 40% overall, and higher still on cloudy days. This level of error might have a significant impact in power system modeling. As with the discussion of the WIND Toolkit above, it is very positive that these types of validations have been carried out, but it is unclear whether users of the data are aware how large the errors might nonetheless be.

The NSRDB also contains time series data of wind and temperature and some other commonly used meteorological fields on the same 4 km (2 km since 2018) grid. These data are provided to aid in calculations

37 https://nsrdb.nrel.gov/data-sets/us-data.

38 See Sengupta et al. (2015a) for a description of the Physical Satellite Model.

39 https://nsrdb.nrel.gov/data-sets/international-data.

40 There is some improvement that observations collected from more advanced instruments can bring to data collected before the new instrument deployment. The new data help to define overall temporal and spatial variability as well as local-scale features that are anchored to terrain. These may be used by machine learning algorithms to "enhance" the prior data.

of temperature and wind impacts on solar generation. These fields are interpolated from data in the MERRA-2 reanalysis, which has a grid spacing of about 60 km; the same issues raised in the discussion of MERRA-2 data apply to them. The interpolation can add confusion because the resampled data appear to have higher resolution than they really do. An inverse-distance weighting is used to cast the 60 km wind data to a 4 km grid, while temperature is linearly interpolated and then adjusted to the altitude of the high-resolution grid using a lapse rate correction that typically will not represent real atmospheric conditions. Because of the coarse source resolution and non-physics-based interpolation, it is recommended that non-solar NSRDB data fields not be used as inputs for calculations of weather impacts on loads or wind generation in power system modeling.

NSRDB appears to meet many of the criteria described above for use in power system modeling. There are decades of observations, the resolution is acceptable, it is continuously extended, and it has been validated and is documented. However, only the irradiance components have an appropriate resolution for power system modeling, and validation reports raise questions as to the applicability of the irradiance data, too. Careful validation of power estimates against observed output is required.

### Public Weather Observations

There is a huge number of public weather stations located throughout the world, and when they provide the right data, with the right attributes for use in power system modeling, they should always be preferred, as observations are always better than model data. Observed temperature data are often available for long enough periods, at high enough density for use in determining the weather impact on load. However, weather observations are typically much denser and higher quality in urban areas; in less densely populated areas where wind and solar generators tend to be located, observations are sparse. In addition, publicly available weather data are not designed to capture the information needed to estimate variable renewable generation. For instance, public stations *very* rarely measure solar irradiance, and the wind is measured at a height of 10 m and not within the rotor plane. This is a significant issue because 10 m wind and hub-height wind follow opposite diurnal profiles, with 10 m wind peaking during the afternoon and hub-height wind peaking at night. There are some quasi-public high-density observation networks, such as the New York State Mesonet (Brotzge et al., 2020), which do have pyranometers that measure GHI at all 126 standard sites statewide and higher-quality radiation flux sensors at 17 enhanced sites, but gradual degradation, and occasional recalibration or outright replacement, of some of the radiation sensors can lead to changing and/or inconsistent observation error characteristics across a network. Particularly for solar radiation observations, users must be aware of instrument quality and calibration issues that can affect measurement uncertainty.

**Public weather stations should always be preferred when they provide the right data, with the right attributes for use in power system modeling, as observations are always better than model data. Temperature data are often available for long enough periods, at high enough density for use in determining the weather impact on load. However, observations are sparse in less densely populated areas where wind and solar generators tend to be located.**

### Proprietary Time Series

As wind and solar capacity has increased, private consultancies that perform tasks such as resource adequacy studies using power system models have had to begin to consider the impact of wind and solar generation. As should be abundantly clear by now, doing so is no simple task, especially since the data needed are not readily available. To their credit, these companies have tried to make do with what data they have and have developed some innovative approaches to estimate renewables generation. However, because filling wind and solar data voids is complex, it behooves downstream consumers of these data or of products derived from them to ask questions. Consumers of the data need to ensure that the methodology is scientifically defensible and that any limitations and their impacts on power system study outcomes are well understood. See Table 2 (p. 76). Ideally, methods should be peer reviewed.

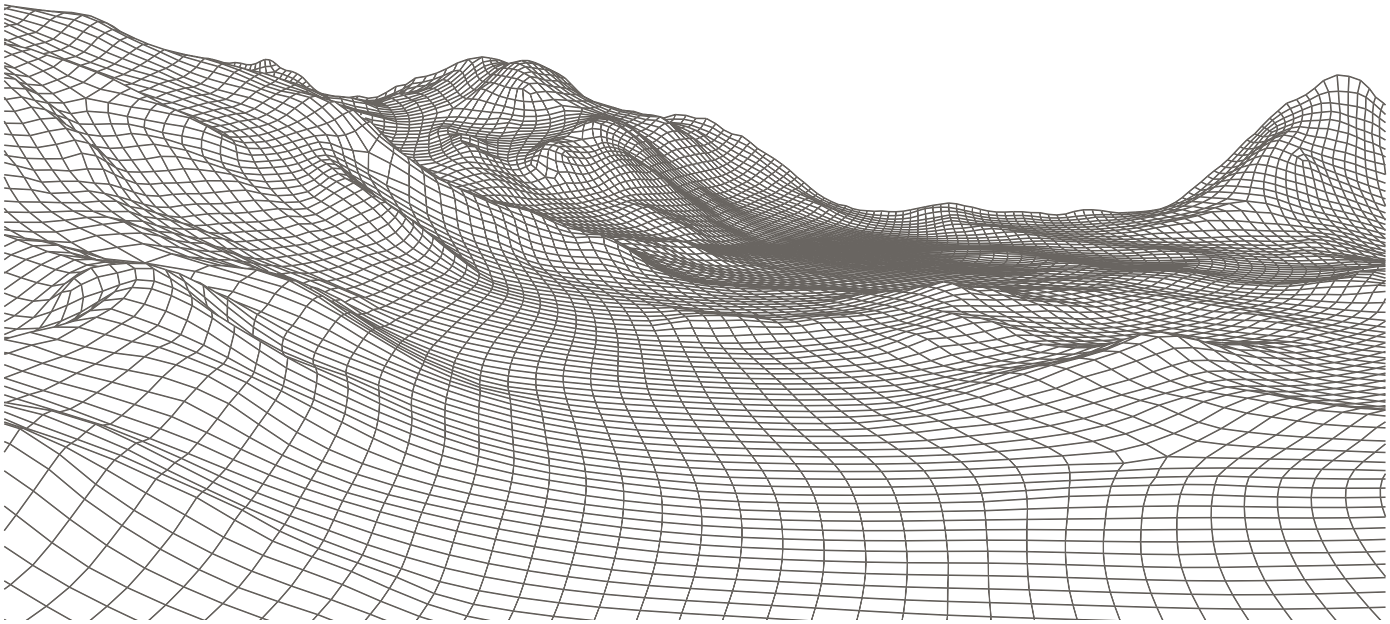
SECTION 5

# Project Description for Producing Robust Weather Inputs Data

The discussion in Section 4 indicates that there are several datasets available that meet some of the requirements of the power system modeling sector for weather inputs, but all of them fall short in one area or another, largely because they either are too low in resolution, do not have a long enough time history, are antiquated, or do not capture all necessary weather variables in a physically consistent fashion.

**All existing datasets fall short of what is needed for power system modeling, as they either are too low in resolution, do not have a long enough time history, are antiquated, or do not capture all necessary weather variables in a physically consistent fashion.**

There is an urgent need to develop one or more datasets that can become the standard for the power/electricity sector to use now and moving forward for the foreseeable future. In addition to activities like renewable energy integration studies, resource adequacy assessments, capacity expansion modeling, and integrated resource planning that use power system models that must represent the increasing weather dependence of the electric power system, other areas would also greatly benefit from such a dataset, including renewable resource assessments and renewable resource performance analyses. In addition, if properly designed and archived, a comprehensive high-resolution dataset would be extremely



useful for foundational research work to examine the relationships between load and renewable resources, and broader weather patterns and climate signals, as well as for establishing possible climate trends. A dataset with the attributes described in Section 4 would also be a leap forward in the state-of-the-science for describing the condition of the atmosphere at high resolution and thus would be of great interest to many other sectors that are weather data stakeholders.

There can be no reliable energy transition without broadly available, consistent, weather datasets for power system studies that meet the seven criteria outlined in this report (see Table 1, p. 61). Given public policies that promote or require increases in renewable resources, this dataset should be considered a public good—one that is government funded, publicly available, and routinely maintained.

While what is being proposed is not trivial, the computer power needed is considerably less than that currently used by the National Oceanic and Atmospheric Administration (NOAA) for its operational forecasting efforts, and it is inexpensive compared to its value: providing accurate information guiding the deployment of trillions of dollars of renewable assets, specifically, where to locate and how to interconnect them in order to minimize cost and maximize reliability.

The current gaps and limitations in weather inputs for power system modeling could be addressed through the creation of a comprehensive, public dataset meeting all the requirements discussed in Section 4. The objective would be to produce time series data that can be used to realistically assess weather impacts on supply and demand in a high-renewables system.

The dataset would have the following attributes: (a) sufficient spatial resolution; (b) sufficient temporal resolution; (c) including the necessary variables in space and time; (d) covering multiple decades with consistent methodology and being extended on an ongoing basis; (e) coincident and coherent across all weather variables; (f) validated with uncertainty quantified; (g) documented transparently and in detail, including limitations and a guide for usage; (h) future-proofed; and (i) publicly available and easily accessible. It would be ideal for an entity with sufficient resources to have responsibility for curating the data, performing ongoing validation, flagging issues, and advising on the dataset's use.

The project would likely use either a high-resolution reanalysis or reforecast method, or a hybrid of high- and moderate-resolution solutions with one of these methods plus downscaling using machine learning methods. It would proceed in two stages: first, a technical review committee would refine the dataset requirements, assess methods for creating a sample dataset, and preside over a request for proposals to create one or more sample datasets that are thoroughly evaluated to assess accuracy expectations for the second phase; second, a high-fidelity archive would be built using the selected methodology, and the process of ongoing extension would be implemented.

## **STAGE 1:** Validate and Refine Requirements and Confirm Need and Fitness

The initial stage of building an ideal weather dataset would convene a technical review committee composed of:

- Expert power system stakeholders
- One or two experienced energy meteorologists who are familiar with the big picture of how power system modeling is performed for both hypothetical studies and actual utility or system planning
- Experienced NWP modelers whose experience covers high-resolution modeling and data assimilation
- Experts in NWP post-processing methodologies including bias correction and downscaling techniques employing machine learning techniques

The technical review committee would perform the following steps:

1. **Vet and refine the dataset requirements described in Section 4**.
2. **Determine possible methods to create the sample datasets**:
   a. Select a period for which data will be produced. This may be a period of a year or a selection of dates intentionally chosen to cover different regimes that are important to system modeling. A recent year and/or a period that overlaps with that of past observational campaigns like the jointly sponsored NOAA/U.S. Department of Energy Wind Forecast Improvement Projects should be used in order to aid validation.[41] Whatever period is chosen should be one where as many quality observations as possible can be obtained to validate the fields that impact wind and solar generation across a broad range of geographies and weather regimes.

   b. Select candidate methods for dataset production. Ideally, candidate methods would be selected in an open and transparent competitive process. For example, a request for proposals could be broadcast, allowing interested parties to submit proposals describing the methodology they believe will best fulfill the requirements. Submissions would then be reviewed, and the most promising ones invited to produce sample datasets. This would ensure the maximum likelihood that candidate methods would include the latest innovations to maximize accuracy and provide a range of options and price points. Another avenue could be a cooperative agreement with NOAA to produce a high-resolution reanalysis dataset based on the current High-Resolution Rapid Refresh Model (HRRR) configuration. This would have the advantage of largely mimicking the current operational set-up and would be highly synergistic and useful to other sectors. Incorporating both approaches would provide the optimal information with which to determine the source(s) that provide the most effective and efficient pathway to producing the full historical and ongoing dataset(s).

   c. Using three to seven candidate methods, produce sample datasets.
3. **Compare the candidate methods and determine their value relative to using continuously extended datasets that exist today.** Datasets that could be compared to the sample dataset include the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) and ERA5 reanalysis datasets and the NOAA/High-Resolution Rapid Refresh (HRRR) operational forecast archive. National Solar Radiation Database (NSRDB) solar irradiance data could also be compared.

41 See https://psl.noaa.gov/renewable_energy/wfip/ and https://psl.noaa.gov/renewable_energy/wfip2/.

a. Obtain as much observational validation data as possible, with a focus on meteorological observations relevant to power system modeling. If possible, use industry outreach to obtain the most relevant observation data.

b. Rigorously validate the sample datasets and the control datasets against observations.

c. If quality power-output data can be obtained from a representative set of renewable resource facilities, create a post-processing model and train it to predict power based on the candidate and control model output to determine whether the new datasets better predict the overall characteristics of generation than the controls. This is important, because while it is certain that low-resolution datasets like MERRA-2 and ERA5 will not accurately predict the wind features present in complex terrain, it is still important to determine how accurately low-resolution output might predict power by statistical means (e.g., building a relationship between site generation and model wind speed/irradiance). It is unlikely that they will more accurately predict power output, but the experience of operational wind power forecasters has been that statistical models relating NWP output to project power can be as valuable as improved wind speed predictions. If such statistical processing of low-resolution model data can yield power predictions on par with those from high-resolution models, it may be worth spending effort looking at ways to utilize existing data.

   Of course, using a method like this would not be possible where no generation history exists, and it is likely that we need the large modeling effort that is being proposed. However, since most new renewable generation is now built near currently operational plants, if such a method works for existing sites using existing low-resolution, easy-to-obtain data from datasets like ERA5, and if operational output data can be obtained from sufficient numbers of existing projects, then developing methods of estimating generation at locations with no history using the existing reanalysis and generation history from existing nearby plants would be a much simpler, cheaper, and quicker solution.

4. Determine whether the candidate datasets add value over the controls. Assuming they do, select the method with the best combination of cost and accuracy and move to Stage 2.

## STAGE 2: Produce Historical Archive and Ongoing Process

Once the value of a dedicated process to produce a high-fidelity archive is established, the next step is to build the archive and operationalize the process of ongoing extension using the method selected in Stage 1. The main decisions at this point would be how far the archive will go back and when operational extension will be performed (for instance, are data for January 5, 2023, produced on January 10, 2023; are data for January 2023 produced in March 2023; or are data for all of 2023 produced sometime in 2024). The rest of the process of developing the data should be relatively straightforward and automated.

At this stage, curation of the data will be key to its usability and to understanding its limitations and uncertainty. The following issues would need to be thought through:

- **Data access:** Data volumes will be very large (many petabytes), and users will need a way to efficiently access the data they need.
- **Observation network:** A broad observation network will need to be built out, both through building new observational assets and by obtaining meteorological data from existing renewable plants. To properly validate high-resolution output, more observations will be needed. These observations will also be valuable in data assimilation where numerical weather prediction NWP-based solutions are deployed, and in post-processing to reduce systematic errors.
  - Regions where wind and solar plants exist or may be built should be targeted, as these are often regions with no currently available public measurements. Where there are public data, these rarely measure wind at the elevations required, and almost never record solar radiation. To obtain the required density of observations will require educating renewable resource project owners on the value of (confidentially) sharing observations to improve ground truth data and getting the renewable energy

sector to understand that improved meteorological datasets are in the interest of the entire sector.

  – Interaction with system operators and regulators may also be needed to help secure meteorological data.

  – In a limited number of cases, new observing networks may need to be deployed, either temporarily or permanently, to assess the quality of the data being produced in important data voids.

- **Ongoing validation:**

  – The data are only valuable if there is confidence in their accuracy. While no dataset will ever be perfect, understanding and communicating the flaws can prevent incorrect downstream conclusions from being drawn, as well as lead to methods to improve it.

  – Low-frequency, high-impact events should be identified and differences between available observations and the model data for these events analyzed in detail to determine how well tail events are captured. Sufficient human resources should be deployed so that high-impact events can be documented in detail to produce a library of such events for future stress-testing of the power system.

- **User education:** Providing insight into how and why the data might differ from ground truth will help to ensure that they are applied correctly. This will also reduce the misuse of existing weather datasets, because users will become more informed about the nuances and limitations of physical model–based datasets and learn best practices for their use. There should also be outreach efforts to promote the use of the data and report back on findings when they are used in important research.

- **Documentation of alternative data sources:** It may be helpful for the project to produce a central knowledge repository describing other energy meteorology datasets and their uses and limitations. This would be valuable for users and would provide insight into ways that data in any dataset can be improved by being refreshed. It is possible that the project could be further expanded to become a repository for the actual data from other efforts as well, allowing it to become a one stop shop and promoting ongoing innovation.

SECTION 6

# Guidance for Using Existing Weather Inputs

Datasets available today for power system modeling have significant shortcomings. At the time of writing, the National Renewable Energy Laboratory's (NREL) Wind Integration National Dataset (WIND) Toolkit best fits the overall needs for public weather datasets in the United States, especially for any studies within complex terrain, and these data are often augmented with data from the National Solar Radiation Database (NSRDB) for estimating solar generation.

For studies in simple, relatively flat terrain, terrain, ERA5 (the fifth-generation European Center for Medium-Range Weather Forecasting (ECMWF) Atmospheric Re-Analysis of the Global Climate) can be considered because its longer history and regular updates may outweigh the issue of low resolution in these regions. The ESIG project team creating this report also notes that several new datasets are in the works, including extensions to the WIND Toolkit, but none meet all of the required criteria for power system modeling. Table 3 (p. 90) provides a summary of how well the combination of the WIND Toolkit and NSRDB datasets, and the stand-alone ERA5 dataset, currently meet power system modeling needs and highlights the gaps and weaknesses of each. Table 2, on p. 76 in Section 4, "An Ideal Weather Inputs Database for Power System Planning, and Comparison to Currently Available Data," provides a more comprehensive table covering many other datasets.

**Despite the shortcomings in available data, the work of power system planners and modelers must move forward. This section describes the most important gaps, proposes ways to fill or work around them, and highlights the resulting uncertainties that system planners need to be aware of.**

Despite the significant shortcomings in available data, the work of power system planners and modelers must move forward. Therefore, in this section we describe the most important gaps, propose ways to fill or work around them, and highlight the resulting uncertainties that system planners need to be aware of.[42]

Unless it is known that meteorological data come from observations, a dataset may have significant deviations from reality—and these deviations can be complex in nature. It is best to always consult a meteorologist when applying weather inputs to power system modeling efforts if the consequences of the analysis have any gravity. The importance of consulting a meteorologist becomes even greater when datasets are extended, to ensure that the extension methodology used allows the objectives of the study to be accurately met. Maintaining the consistency through time and space of the weather variables is vital but difficult, especially when using statistical bootstrapping methods.

Conversely, it is important for meteorologists who are extending the existing weather data inputs for power system modeling to consult with power system experts in order to fully comprehend how the data will be used prior to executing intensive computer simulations to produce these datasets. This will ensure that methodologies will produce results that are consistent with the sector's needs.

42 For definitions of terms that some readers may be unfamiliar with, please see the glossary at the end of the report.

TABLE 3

## Summary of Best Available Public Datasets to Estimate Site-Level Generation at All Current and Future Wind and Solar Assets in All Regions of the United States

| Attribute | WIND Toolkit/NSRDB Combination | | ERA5 |
|---|---|---|---|
| | For Wind/Load | For Solar | Wind/Solar/Load |
| **Has required temporal resolution**[a] | 5-min produced | 5-min since 2019 | Hourly |
| **Has required spatial resolution** | 2 km | 4 km; 2 km since 2019 | 30 km |
| **Includes multiple heights above the surface** | | N/A | |
| **Available for several decades** | 8 years[b] | Since 1998 | Yes |
| **Has regular updates** | Nothing formal | Annual | Daily (7-day lag) |
| **Is future-proofed** | Ad hoc | Yes | Yes |
| **Is long enough to detect climate signals** | Unlikely | Possibly | Yes |
| **Models are adequately validated** | | | |
| **Accuracy assessed, including for risk periods** | Against tall meteorology towers | Limited | Limited |
| **Variability assessed, against reality** | Limited | Limited | Several studies |
| **Assessed power system modeling applicability?** | Designed for this | No studies found | No studies found |
| **Provides companion "forecasts"**[c] | Produced | No, but possible | No |
| **Is based on consistent input observations and/or models** | Yes, except 2014 | Yes | Yes (single modeling system) |
| **Physical consistency between wind/solar** | No; impact should be investigated | | |
| **Well documented and easy to use** | | | |
| **Limitations are clearly specified** | | | |

Fully Met Close to Being Met Partially Met Met in a Very Limited Way Not Met at All

**This summarizes key attributes of the three best available public datasets that can provide a reasonable estimate of site-level generation at all current and future wind and solar assets in all regions of the United States. The WIND Toolkit and NSRDB are typically used in tandem, with the WIND Toolkit providing data for estimating wind generation (and possibly loads) and NSRDB being used for estimating solar generation, because neither provides acceptable accuracy for both variables. This introduces physical consistency issues, as described in Box 4 on p. 48).**

a All datasets have hourly data. Five-minute data were produced for the WIND Toolkit, but NREL reports that they are no longer available.

b Data from the years 2007 through 2013 use a different configuration compared to 2014. An extension is being produced for the whole dataset.

c For definitions of "forecast" used in this report, see Box 1, p. 11. NREL reports that the forecast dataset is no longer available but other sources may have an archive.

Notes: WIND = Wind Integration National Dataset; NSRDB = National Solar Radiation Database; ERA5 = Fifth-Generation ECMWF Atmospheric Re-Analysis of the Global Climate. Climate; N/A = not applicable.

Source: Energy Systems Integration Group.

Various approaches can be used to supplement existing data and to perform stress tests of the power system for extreme weather. A few examples are:

- Using data sources such as surface measurements in addition to the synthetic datasets
- Extending existing data sources to represent higher levels of renewable resources
- Extrapolating existing datasets to include more years of consistent data
- Extending existing datasets to evaluate the impact of extreme weather

Below, the project team urges caution if using new or ad hoc methods and describes why. Then it discusses promising practices for applying any meteorological datasets to power system models and provides one example for each from recent work. These examples help light the way for additional evolution of approaches to filling the existing data gaps.

## The Need for Caution If Considering the Use of New or Ad Hoc Methods

Some recently developed methods may be unproven or invalid. Currently, many power system planners are developing heuristics or bootstrapping approaches to fill the gaps and limitations in current weather datasets. Typically, these bootstrapping methods will take a short period (several years) of available, modeled wind and solar data and extrapolate that to long multi-decadal datasets using correlations in daily temperature. While these methods are well intentioned (and often necessary, given the unavailability of data), they have limitations. Power system planners will want to use caution if utilizing new or ad hoc methods, as some recently developed methods may lead to invalid modeling results.

The emergence of unproven or invalid methods is not new. As early as 2012, perhaps earlier, it was recognized that new methods for bootstrapping or otherwise using statistical approaches to develop wind/solar profiles was, in some cases, becoming increasingly untethered from the way that weather actually works. For example:

> For integration studies, it is critical to use time-synchronized wind and load data to ensure that underlying weather drivers are properly accounted for in the statistical analyses and operational simulations. Unfortunately, the existing datasets are becoming somewhat stale. This is leading to the development of questionable ad hoc methods that attempt to create high-quality, consistent wind datasets to be used for integration studies. Unfortunately, such methods will likely compromise integration analysis and interconnection planning until more current datasets can be developed and kept up to date (Milligan et al., 2012).

**Even when using a complete dataset from a reputable source, it is bad practice to assume that if a dataset covers the geographical region of interest at a spatial and temporal fidelity necessary for power system modeling, the data provide appropriate accuracy for the task at hand.**

Even when using a complete dataset from a reputable source, it is bad practice to assume that if a dataset covers the geographical region of interest at a spatial and temporal fidelity necessary for power system modeling, the data provide appropriate accuracy for the task at hand. The temporal length and spatial breadth of many gridded datasets, as well as their ease of use, can seem compelling, and users often treat these grids as black boxes representative of reality. However, while the data are usually anchored to some atmospheric observations, most of them are modeled; in addition, many meteorological phenomena, especially those driving wind resource, are scale-dependent and may not be resolved by the modeling method employed, or for all weather regimes. It is therefore crucial to understand how datasets used as meteorological inputs to power system modeling might deviate from reality.

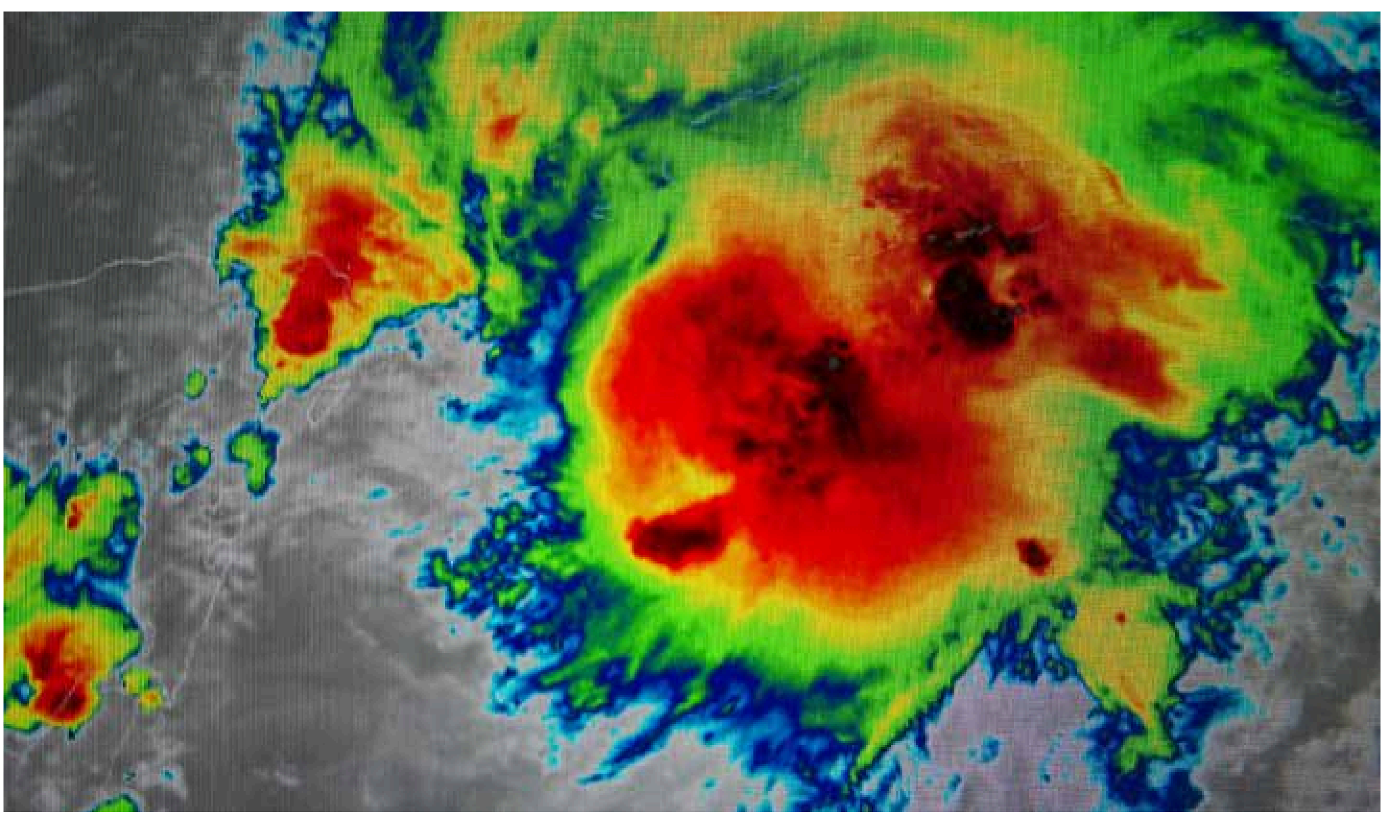
Typically, even when appropriate meteorological inputs are available, they do not cover the right time window or a long enough period. This leads to efforts to extrapolate and extend the data that are often questionable. Examples of ad hoc approaches include, but are not limited to:

- Scaling wind or solar input linearly to represent more renewable capacity or, more generally, making linearity assumptions about weather data

- Mixing and matching data from different time periods or different datasets
- Creating synthetic bootstrap datasets without considering the concurrency and consistency comments above
- Assuming that locations near to one another experience the same meteorology
- Assuming linear (or even polynomial) linkages between different weather variables and using these to extrapolate the time series of one weather variable from another. Weather variables are highly non-linear and rarely, if ever, interact in linear ways. For instance, solar and wind generation does have a relationship with temperature, but temperature alone cannot be used to predict the wind or solar energy that will be generated.
  - Assuming linear linkages is problematic because relationships between temperature and other weather variables are non-linear, complex, and change with time and weather regime. For example, for a single location, a period with very cold temperatures could be accompanied by high wind on some occasions and low wind on others.
  - For wind and solar resources coincident with temperatures driving load, shorter datasets with known coincidence and consistency between variables are better than long datasets with some or all the data being synthetically created with untested relationships.
- Using statistical and probabilistic simulation methods for producing renewable resource time series that are untested or unproven to be valid.

## Examples of Practices for Consideration When Applying or Extending Existing Datasets to Evaluate Weather Risk

There are, however, approaches to applying and extending existing datasets that can yield reasonable/usable results, if done carefully and with the necessary expertise and communication among parties. Meteorological datasets that provide high spatial fidelity, while anchored by some atmospheric observations, are almost always derived in part by models. If a model is simple and easy to understand by a non-meteorologist, it likely is not very accurate. Conversely, better models are very complicated, but this can create an assumption that



the complexity implies accuracy. If the complexity created a near-perfect representation of weather conditions influencing the power system, then the data could be used with confidence as a proxy for observations. However, this is not the case, because, while atmospheric modeling is tremendously effective and valuable, it is simply not possible for it to produce near-perfect results in all circumstances due to limitations in observations, computer power, and approximations in model representation of the atmosphere (see Section 2, "Meteorological Data Fundamentals for Power System Planning," for details).

When using any weather data for power system modeling (even if the dataset is already considered complete enough for use), it should be standard practice to consult a meteorologist who is well versed in the attributes of those data. The teams that produce these datasets can be excellent sources of information, but typically they do not have the budget or staff to provide case-by-case help. It is even more important to get advice from a qualified meteorologist if one wishes to extend an existing dataset. Carrying this out requires experts with knowledge of the observations and model(s) used to produce the dataset and of the phenomena that predominantly affect the renewable resources in the system being evaluated.

For example, if one is using a dataset produced by a numerical weather prediction (NWP) model to estimate wind generation, then a meteorologist with experience interpreting NWP wind output (ideally for the purpose of wind energy prediction) in the region of the study should be consulted. Similarly, for the National Solar Radiation Database (NSRDB), extending the dataset requires a meteorologist with some knowledge of the method used to derive surface irradiances from satellite data and of the nuances of the data validation program. If the meteorology expert is educated by the power system modeler about the conditions that cause stress to the system being evaluated, they can provide specific insights into the uncertainties in the data being used that can ensure that risk is being correctly quantified. They will also be able to suggest methods to extend datasets (and the caveats that result from doing so) or propose alternative methods to more completely evaluate the potential areas of weather-driven risk.

**Meteorological experts should be consulted who are knowledgeable with the observations and model(s) used to produce the dataset and with the phenomena that predominantly affect the renewable resources in the system. If the meteorology expert is educated by the power systems modeler about the conditions that cause stress to the system being evaluated, they can provide specific insights into the uncertainties in the data being used that can ensure that risk is being correctly quantified.**

### Representing Higher Levels of Renewable Resources with Existing Data Sources That Do Not Completely Overlap

A recent report from GridLab examines alternative pathways that California could take to reach its clean energy targets by 2030 (Stenclik, Welch, and Sreedharan, 2022). To represent a future build-out of more wind and solar generation in the power system model representation of the generation portfolio, additional wind and solar capacity was added to supplement the existing plants' capacities. This optimization approach that determined where to put the capacity used the NREL WIND Toolkit and NSRDB data as the weather input, and modeled as many years as possible in the PLEXOS model to gain a better understanding of the potential future resource mixes through time.

The study team recognized that the datasets did not fully overlap and consulted a meteorologist to evaluate how the data limitations impacted the study and to sanity-check methodologies that attempted to mitigate the limitations. Because there are 22 years of solar data and only 8 years of wind data, the study team applied the full 22-year set of solar data to further examine its impact over a longer term. Wind data were used in a couple of different ways, recognizing the absence of the common weather driver for the years during which there were solar data but no wind data. One approach was to use the annual wind profile with the lowest wind output (2012) from the entire wind dataset. The study recognized this limitation, but determined that because California's current and projected solar capacity is significantly higher than the wind capacity, having more years of solar data was more important than a shorter dataset with correlated wind and solar. While this approach provides some insights, a better dataset would have allowed for greater insight and certainty on how the California grid would perform over a longer period of time.

The study also used the locations for all existing wind and solar plants from the U.S. Energy Information Administration (EIA) Form 860 database to develop generation profiles consistent for incumbent plants.[43] Because one of the study's objectives was to assess the ability of California's grid to tap into potentially available resource across the West, wind and solar generation was estimated for the entire United States portion of the Western Interconnection. Both the wind and solar data from the NSRDB and WIND Toolkit were input to the NREL System Advisor Model (SAM) to create the power profiles at each desired location.[44] This resulted in multiple years of hourly, time-synchronized wind and solar power production data which the study authors say "is critical to understanding the multi-year variability of the wind and solar resources, the likelihood of multi-day sustained low renewable resource production, and the characteristics of outlier events" (Stenclik, Welch, and Sreedharan, 2022, 18-19).

Once the process of determining the capacity expansion to a future renewable resources portfolio had been completed, a time series of estimated renewable generation for each plant in the expanded portfolio could be calculated for the eight years that had overlapping wind and solar data, and outlying days with low renewable generation identified at different levels of geographical aggregation.

## Assessing the Quality of Power System Model Weather Inputs During Periods of System Stress

As a separate part of the study in the above example, days where renewable resource was particularly low relative to demand were investigated by the team meteorologist. It was found that some issues in the NWP representation of the renewable resource during these periods of system stress resulted in a substantial

over-prediction of the available wind resource (Sharp, 2022). In this case, the portfolio had enough dispatchable capacity to cover the difference between actual resource and model estimates, but this may not be the case for studies involving higher levels of renewables or more retirement of gas capacity. This illustrates the importance of having a subject matter expert evaluate power system modeling weather inputs.

## Extrapolating Existing Data Sources to Longer Time Series

There is a range of current practices that attempt to fill the data void using bootstrapping methods, and some are better than others. Hart and Mileva (2022) provide an example of one of the better methodologies, which they used in a study that describes a weather-synchronized simulation approach to resource adequacy analysis. A combination of the Federal Energy Regulatory Commission (FERC) Form 714, NREL WIND Toolkit, NSRDB, EIA Forms, and Western Electricity Coordinating Council (WECC) datasets were used to provide hourly time series data for load, wind, solar, thermal outage/derate, as well as lower-resolution hydro data. Concurrent weather data for several sites were obtained from the National Oceanic and Atmospheric Administration's National Centers for Environmental Information

43 See https://www.eia.gov/electricity/data/eia860/.

44 The System Advisor Model is a free techno-economic software model that facilitates decision-making for people in the renewable energy industry. See https://sam.nrel.gov/.

(NOAA/NCEI). The available data only overlapped across the period 2007–2014 and were extended to 2007–2020, with extreme care taken to preserve the underlying weather drivers. The authors used weather binning methods in which the weather variables explaining the hourly net load at key locations driving supply and demand were determined using principal component analysis of the hourly net load and used the principal components to divide days into different bins of weather modes. A Markov chain method was then used to randomly walk between weather bins in a fashion that was consistent with real daily weather transitions, and on each day Monte Carlo methods randomly selected hourly load, wind, solar, and thermal shapes from options within the same regime bin. This allowed them to greatly expand the potential system conditions with combinations that are plausible.[45]

Hart and Mileva also performed weather-synchronized simulations where weather data from available model time series were used to simulate coincident conditions for the overlapping period of wind, solar, load, thermal, and hydro data. They obtained results that were similar to the Monte Carlo analysis, though the loss-of-load expectation in the synchronized study was slightly higher. They found that the weather-synchronized simulation approach was superior to Monte Carlo because it "provides confidence that the findings reflect actual physical weather phenomena and all relevant spatial and temporal correlations" (Hart and Mileva, 2022, 14). Methods such as this are therefore promising, and they offer much more plausible results than "blind" Monte Carlo sampling (using only daily temperature correlations, for example) that may destroy the underlying weather drivers for wind, solar, demand, and hydro. However, this study also found that the principal component driving loss-of-load expectation on the Western Interconnection was still temperature. These results suggest that at current levels of wind and solar in the United States, it is possible to use the currently available data and some promising approaches to bootstrapping datasets that allow for reasonable optimizations in planning tasks like capacity expansion and reasonable estimates of resource adequacy; however, as renewable capacity increases, higher-quality, longer datasets will become essential.

## Extending Datasets to Examine Extreme Weather, High-Impact Low-Probability Impacts

There is increasing interest and urgency in capturing grid performance during extreme storms, one category of high-impact, low-probability events. The 2014 NERC Integration of Variable Generation Task Force (IVGTF) Task 1.6 report discusses the use of probabilistic modeling in power system applications, including to model high-impact low-probability events (NERC, 2014).[46] Because these events are relatively rare, we do not have sufficient data to really capture the probabilities involved. There is an increasing need for scenario analysis and stress testing to see how the grid can operate during these unusual events. This is discussed further in NERC (2014).

Two recent reports develop stress tests based on plausible potential events that may make it challenging for the grid to operate reliably. The first study, the California Pathways report, developed several such stress tests (Stenclik, Welch, and Sreedharan, 2022). These included early gas retirements in California, low hydro power availability, coal retirements in the Western Interconnection, limiting imports to the California grid, multi-year demand variability, combined stressors, and demand flexibility. Each of these stress cases was modeled with an eight-year weather dataset, and with three alternative resource portfolios, resulting in 192 years of simulations, plus additional simulations across the 20-year demand data that yielded a total of 264 years of simulations.

Notably, each of these stress tests did not attempt to alter the wind/solar datasets to represent some type of extreme wind/solar performance, as doing so would have likely destroyed the common weather driver that links demand, wind, and solar generation. However, the resulting stress tests account for the common weather driver across the many modeling inputs, providing confidence that no artificial or implausible weather patterns were assumed in the analysis. Additionally, the eight-year dataset captured actual storms across the interconnection. Of course, a longer dataset would further increase confidence in the results of the stress tests and could reveal other potential events of concern.

45 See Hart and Mileva (2022) Section 2.1.1 and Appendix B for more details.

46 See NERC (2014) pages 14-15 for why it is difficult to assemble a proper dataset to analyze high-impact, low-probability events.

A second study, carried out by GE Energy on behalf of the Natural Resources Defense Council, developed and modeled two stress tests for the Eastern Interconnection to evaluate the benefit of eliminating key transmission bottlenecks during extreme weather (GE Energy Consulting, 2022). The first stressor modeled was a heat wave in 2035, the basis of which was the three-day summer heat wave of August 2018. The second stressor was a winter polar vortex, also in 2035. This event was based upon the 2014 polar vortex that saw demand about 40 percent higher than normal at the same time that generation outages increased because of the cold weather. The demand profile was increased to simulate the higher load during the storm, but the overall pattern of demand remained the same. The report showed that interregional transmission would have significant value by reducing or preventing electricity shortages during the extreme weather that was evaluated. Similar to the California Pathways report, the Natural Resources Defense Council report would have benefited from a more complete and robust dataset that contains more actual extreme weather events.

As discussed above, it is essential that data inputs used for power system modeling that include renewable resources ensure underlying consistency in the weather drivers. Additionally, a more comprehensive validation program should be carried out, especially now that much more wind and solar are expected to connect to the grid in coming years. With an annually updated wind/solar database, such validation could be extremely informative and more comprehensive because more renewable plants will be online.

Power system modelers should resist the urge to tamper with weather/wind/solar data unless they are working closely with a qualified meteorologist.

Some ad hoc methods could potentially be valid, but are unproven, while others are not plausible. Although actual wind/solar production data may be useful in some types of studies, without complementing these datasets with data that can reasonably represent future renewable resources, the usefulness of these actual data for resource adequacy studies may be somewhat limited until a more robust, annually updated, dataset can be developed. Mixing renewable resources and demand data from different years will destroy the underlying weather linkage and will result in invalid modeling outputs. Even though the results may "seem reasonable," there is no reason to assume that they are reasonable. Above all, remember the aphorism: garbage in = garbage out.

## Do's and Don'ts

The following list of do's and don'ts when using existing weather datasets in power system modeling studies are adapted from Stenclik (2022).

### Do consult a meteorologist; don't go it alone.

Power system planning, particularly resource adequacy analysis, sits at the intersection of engineering, economics, and meteorology. Too often, power system engineers and planners develop assumptions about weather without consulting with a meteorologist or atmospheric scientist familiar with the use of weather data in power system modeling. The truth is that cross-disciplinary analysis is required, and power system engineers need to exercise caution when bootstrapping datasets, especially for outlier events like Winter Storm Elliott in North America in late December 2022.

### Do model stressors to all resource types. Don't assume that extreme weather impacts only renewables.

Winter Storm Uri in February 2021 and Winter Storm Elliott in December 2022 showed that all resource types are affected by the weather. Natural gas is susceptible to fuel scarcity, wind and solar are dependent on atmospheric conditions, coal piles freeze, and all equipment sees increased outages during extreme conditions. All too often, however, system planners assume that only variable renewable resources are affected by the weather. This is emphatically not the case. The impact of weather on all resources needs to be modeled across a wide range of conditions to ensure that sufficient resources are available to meet load even during conditions likely to cause common mode failures.

### Do stress-test systems against as many future weather realizations as possible. Don't make investment decisions using single weather years.

Power systems need to be stress-tested against potential high-impact, low-probability events, for both the current and future power grids. But planners need to proceed with caution. These events may not fit neatly in the conventional planning reserve margin and the one-day-in-ten-years loss-of-load expectation framework that our grids are planned to accommodate. These events should not just drive investment in *more* resources, but rather investment in a more resilient grid generally. Stress-testing, in conjunction with typical probabilistic analysis, is needed.

### Don't just evaluate a doomsday planning scenario. Do use data reflecting likely correlations among stressors.

When trying to prepare for worst-case scenarios, it can be tempting to develop an infeasible "what if" situation. Rather than utilize robust atmospheric and meteorological analysis (see the first item in this list), power system planners often develop a doomsday scenario where everything goes wrong simultaneously. Load spikes, all wind and solar generation drops to zero, and transmission interconnections with neighbors are unavailable. But while there are certainly correlations across these stressors, it is important to base analysis on likely or potential weather conditions rather than synthetic stress events.

### Do consider weather in neighboring grids. Don't assume each power system is an island.

While the North American power grid is made of a smorgasbord of independent system operators, regional transmission organizations, utilities, and various balancing authorities, transmission links most of them across the Eastern and Western Interconnections. During extreme weather events, these transmission links can offer significant reliability benefits; however, many independent system operators, regional transmission organizations, and utilities plan for islanded conditions without support from neighboring balancing authorities. This leaves a lot of value on the table and makes interregional transmission development difficult. Ongoing discussions at FERC and elsewhere are rightfully considering ways to improve these interconnections to help support reliability and resource adequacy.

SECTION 7

# The Impact of a Changing Climate

Over the coming decades, power systems will continue to evolve as levels of storage, renewables, and other new technologies increase, conventional generators retire, and transmission and distribution systems are modified. At a regional scale, these changes will likely have a larger impact on bulk resource adequacy than climate change will (e.g., Bloomfield et al., 2021). As this report and others (e.g., Craig et al., 2022) have shown, it is critical to develop appropriate datasets to model this weather-induced variability and to enhance the capability of power system models to be able to ingest the multi-decadal datasets needed to model demand and renewable generation variability accurately.

While climate science is unable to quantify the impacts of climate change on wind and solar resources with certainty at this point, it is advanced enough to begin to predict trends in temperature and, to some degree, precipitation. There is evidence that urbanization and climate change are already changing temperature distributions, and it is expected that future conditions will include increases in the average number of cooling degree days (thus increasing the use of air conditioning), increased extreme temperature excursions, changes in humidity during the cooling season, and changes in the length of hot spells. Changes in the nature of the cool season are also expected, with an overall decrease in heating degree days (leading to less need for space heating) combined with the possibility that extreme cold waves may become more likely and/or severe in some locations. Hotter summers and longer dry periods are also resulting in a rapid increase in wildfire frequency, intensity, and size, and these fires impact the transmission and distribution system.

But while it is unquestionable that the impacts of weather and climate on demand are changing, predominantly through increasing temperatures, the question of how wind and solar resources will change is more complex and requires significant engagement between the power sector and the climate science community. Because of this, and because the task of gathering appropriate data to model wind and solar generation is already immense, this report has not delved into the challenge of integrating climate projections into power system planning.

However, climate change cannot be ignored. At smaller scales, climate change could have a more magnified impact, as demonstrated by public safety power shutoffs in the U.S. West. And we are already seeing an increase in the number and intensity of extreme weather events occurring due to climate change around the world[47]—extreme heat, heavy precipitation, and prolonged drought—and this will increase in the future. These extreme weather events impact the reliability and resilience of energy systems and change the envelope of uncertainty surrounding different possible weather-driven outcomes, and with it the likelihood of events that stress the system. Extreme events do not have to be "extreme" in the classical sense to pose risks to system operation. As more wind and solar are added to the resource mix, *combinations* of events, such as low winds and moderately cold temperatures, may strain the grid. In the future, weather that appears unremarkable to an individual (a low-wind, cloudy, moderately cold day) could be a high-impact stressed grid condition.

Therefore, when using weather inputs for power system modeling that are developed based on past conditions, there is a need to contextualize the results with some consideration that the climate is changing. Some of this change will already be captured in long time series data (e.g., a comparison of the results from the 1980s with 2010s), but the changes will not be representative of the future, especially for studies that extend over a longer period (say, 10 to 20 years in the future). The changing climate will impact all facets of the power system, motivating the need for better planning and operational strategies in anticipation of extreme events.

**Extreme events do not have to be "extreme" in the classical sense to pose risks to system operation. In the future, weather that is considered unremarkable—such as a low-wind, cloudy, moderately cold day—may result in a high-impact stressed grid condition.**

Multiple dimensions of resilience come into play when considering the impacts of climate change on power system operation. Consequences of a potentially disruptive event can be minimized by the system's ability to:

- Withstand the impact of a disruptive event (through advanced hardening, protection measures, or proactive decisions on revision of future technical standards)
- Respond to the conditions (through real-time operations such as generator redispatch and transmission switching)
- Recover quickly (through targeted restoration, getting repair crews in place, and replacement hardware)

Each of these dimensions requires an understanding of the impacts of expected weather and the ability to forecast it at an appropriate lead time. Characterizing the future landscape of extreme weather events is often done by looking at the past, for example, using multiple decades of reanalysis data to look at the most extreme event seen in a region. However, because of climate change, the past is now an insufficient representation of the range of future conditions. This is due to the non-stationarity of many of the meteorological drivers of power system behavior, such as near-surface temperature, precipitation, solar radiation, and wind speeds.

To understand these impacts, there is a need to use climate projections, which can inform our understanding of changes to the mean climate and vulnerabilities of power systems to extreme weather. This section provides an overview of ways that climate change will need to be considered when interpreting output from power system

47 See https://www.forbes.com/sites/mariannelehnis/2022/12/29/2022-was-a-year-of-record-breaking-extreme-weather-events/?sh=5b13c338736b for examples from 2022.

models that are evaluating future conditions, and includes a brief discussion of how the current state of the science influences confidence in the importance and magnitude of different effects.

## Ways in Which Climate Change Will Affect the Applicability of Historical Data and Modeling Outcomes

Here we focus on the response of climate models to increases in greenhouse gas emissions and how this may theoretically impact the electricity sector, and discuss the limitations of these climate models for power system modeling.

The uncertainty in future climate projections arises from three distribution sources:

- The choice of climate data years used for the simulation (internal variability)
- The choice of climate model (model uncertainty)
- The choice of climate projection scenario (scenario uncertainty)

Climate scientists have spent significant time and resources to quantify the relative magnitude of these sources of uncertainty at different temporal and spatial scales (e.g., Hawkins and Sutton, 2008), and this can be useful for a discussion of the applicability of historical data. Despite the significant work done to date, the large uncertainty remaining for many atmospheric variables makes power system planners' work complex, as it remains difficult to prepare electricity infrastructure to survive such events at an acceptable level of reliability.

### Temperature

The climate science community has high confidence when it comes to the impact of climate change on variables like near-surface temperature. Global mean temperatures are increasing at a rate proportional to anthropogenic greenhouse gas emissions. Multiple decades ahead, the main uncertainty comes from the choice of climate projection scenarios (i.e., the extent to which policies and technological developments drive greenhouse gas reductions). There is an increase in the frequency of types of events most closely related to near-surface temperatures, such as multi-day heat waves, hot


days, and tropical nights (when temperatures do not drop below 20°C (68°F) during the night). These increasing temperatures will lead to changes in electricity demand, through reductions in demand for heating and increased demand for cooling, with changes in cooling outweighing changes in heating (Deroubaix et al., 2021; Bloomfield et al., 2021). Increased temperatures can also lead to a reduction in solar photovoltaic (PV) performance due to reduced panel efficiency (Feron et al., 2021). Other impacts of increasing temperatures include changing load profile shapes, reduced cooling water for traditional power plants, and reduced transmission capacity (Panteli and Mancarella, 2015).

### Precipitation

Known thermodynamic responses of the atmosphere to climate change suggest that storms will become more intense due to increases in available precipitable water, although there is very large uncertainty in climate models and internal variability in the number and intensity of extreme storms. As air warms, its capacity to hold water increases (known as the Clausius-Clapeyron rate), and heavy rainfall events are expected to become more frequent. The warmer and moister atmosphere and oceans also suggests that the strongest hurricanes will become more intense, with more rainfall and possibly increasing size, which would affect new areas (Gensini, Ramseyer, and Mote, 2014). There is, however, very large model uncertainty and internal variability in the number and intensity of extreme storms and hurricanes. The prevalence of severe weather events like extreme precipitation events and extreme winds (straight-line, hurricanes, and tornados) may all impact grid infrastructure. Snowpack and melt timing are impacted by climate change in a regionally specific manner, resulting in changes to the timing of hydro power availability and system constraints that can lead to oversupply and undersupply issues (Craig et al., 2018).

Climate change has also altered the historical pattern of droughts. Warmer, drier conditions result in the drying up of water sources such as lakes and rivers. Droughts are projected to become more frequent and longer and have more severe consequences, such as what is seen in the western United States. These droughts directly impact hydro power production and the availability of cooling water for thermal generation plants.

### Wind Speed and Solar Irradiance

It is much more difficult to project changes in wind speed and solar irradiance distributions in time and space, as these are impacted very strongly by internal variability (see, for example, Bloomfield et al. (2021) for an example over Europe). Over the historical period there is some evidence in the scientific literature for global stilling—a reduction in the magnitude of global-mean near-surface wind speeds—from the 1960s to 2010s, which is thought to be due to dynamical changes in reductions in the pole-to-equator temperature gradient (another result of climate change, as the poles warm much more rapidly than the equator (Zeng et al., 2019)). However, the global-mean wind speeds have increased again since 2010, and this behavior can be tied to atmospheric variability.

The general response of surface solar irradiance is that there is an increase in cloudy days, which results in reduced solar PV output (Haupt et al., 2016). However, for both wind and solar power the day-to-day variability is much larger than the climate change signal. For now, the important thing is to include the day-to-day variability in power system studies for accurate results (Yin, Molini, and Porporato, 2020). Ultimately, risk characterization of extreme events will need to include a regional study of expected hazards and robust sensitivity analyses across the spectrum of uncertainties present in a future climate.

> **There is an urgent need for interaction between the climate science communities and power system modeling communities to bridge the gap between known methods for processing historical data and the need to incorporate the impacts of climate change.**

## Need for Interaction Between the Climate Science and Power System Modeling Communities

Many of the changes outlined in this section highlight that historical time series such as those described in previous sections of this report are not appropriate for thinking about a future climate. A simple way that this is often dealt with is to de-trend the historical temperature record, bringing temperature levels up to those experienced in the present day (or projecting them forward to potential future levels such as was done in Bloomfield, et al. (2022)). However, for variables other than near-surface temperature, removing these trends can be complex due to the difficulty in separating climate change–induced trends from internal variability. This highlights an urgent need for interaction between the climate science communities and power system modeling communities to bridge this gap between known methods for processing historical data and the need to incorporate the impacts of climate change.

## Evaluating How Climate Change Affects Specific Weather Events That Adversely Affect Power Systems

Severe weather events like extreme temperature, precipitation, and wind all impact grid infrastructure. Evidence points toward climate change increasing the probability of severe weather events, due to more energy availability in the atmosphere (mostly because of increased moisture content), and dynamical changes as the equatorial-to-polar temperature gradients change. However, with the exception of extreme temperature events, this work is far from being settled science. With this in mind, power system modeling studies, especially those looking beyond the next five years, need to consider how high-impact, low-probability event risks such as those below might be taken into account with at least moderate certainty:

- Extreme temperatures leading to more frequent outages and derates of transmission and generation assets that coincide with high loads
- Very low resource availability for renewables that may happen but for which there is not yet a compelling signal relative to the noise of internal variability
- Drought and shifting timing of precipitation that increase the power grid's exposure to wildfire-related risks (from wildfires damaging equipment or requiring de-energization to prevent electrical equipment from igniting fires)

How can power system modelers incorporate possible climate scenarios and articulate their uncertainty? How can energy meteorologists monitor these impacts and trends going forward? These questions are beyond the scope of this report, but power system modelers and atmospheric scientists must be thinking about them and working together to make sure that the right problems are being addressed.

## Current Capabilities of Climate Projection Datasets

Although climate scientists have some understanding of the response of meteorological variables to climate change, the exact way that these changing meteorological variables transfer into power system impacts is complex. Issues with using future climate model simulations include:

- Available spatial resolution of datasets
- Available temporal resolution of datasets
- The need to fully incorporate future climate uncertainty

Throughout the climate modeling community there are multiple types of future climate simulations available for climate impact modeling. These include both global and regional climate models. All types of climate model data will require some calibration before use in power system modeling, and this is strongly encouraged so that these modeling results are not biased.

Global climate models are generally of lower spatial and temporal resolution (to account for having to model the whole globe), but these large-scale modeling simulations allow for an understanding of how local-scale behavior can link to large-scale forcings, such as impacts of different teleconnection patterns (climate anomalies that are related to each other at large distances) like El Niño–Southern Oscillation or the Pacific–North American Pattern. They can also help with understanding of the

drivers of local changes through analysis of phenomena like jet stream position or pole-to-equator temperature gradients. Global climate models are generally run at horizontal grid spacing of 100–200 km with output at daily intervals.[48] Examples of commonly used global climate models are the CMIP5 and CMIP6 ensembles, which contain data from multiple modeling centers and form part of the Intergovernmental Panel on Climate Change reports.[49] Data are normally available for a historical period (sometimes as far back as 1850) through to 2100 to be used for integrated assessment modeling to aid in future policy decisions. Some state-of-the-art High Resolution Model Intercomparison Project experiments are now providing global climate model data at hourly resolution and approximately 50 km grid spacing, but these are very limited.[50]

Regional climate models are generally of much higher spatial and temporal resolution (~10–50 km and 1- to 3-hourly resolution) than global climate models and can be thought of as "downscaled" global climate models. This higher resolution allows for much more information to be given at a particular location and for the analysis of diurnal cycles of relevant variables, which is useful for power system modeling. However, unpacking the underlying meteorological drivers of a system's behavior may be more difficult, and models may have some issues around the boundaries of the domain. Generally, regional models are available in large ensembles with different combinations of global climate models providing boundary conditions and regional climate models performing the downscaling, so these can provide a good assessment of climate model uncertainty. An example of this is the European EURO-CORDEX experiment and a similar project for North America called NA-CORDEX.[51] Although some regional climate models are created by statistically downscaling global climate data (e.g., using machine learning), and this can be done in-house at an institution, users must use caution, as the downscaled models may not accurately represent all the key climate processes.

Some "ultra-high-resolution" regional climate models are also available, often run using the Weather Research and Forecasting model (WRF). Pryor, Barthelmie, and Shepherd (2020) include an example run at 4 km. These models are useful as they run at a scale referred to as "convection-permitting," so they can represent clouds more realistically, which is important for modeling renewable generation. These simulations are also operating at a scale consistent with the recommendations for realistic operation of a wind turbine parameterization (Fitch et al., 2012). However, although these high-resolution simulations may seem the obvious choice for wind power modeling, they are often only from one climate model for a single emissions pathway, and therefore focus primarily on accurately modeling internal variability rather than scenario and model uncertainty. Thus, the choice of climate modeling tool above will be very dependent on the science question being considered.

The use of multiple climate change scenarios is particularly useful for power system modeling that is highly dependent on changing temperatures, as there is relatively little internal variability and model uncertainty at the end of the century. In contrast, for wind and solar generation modeling, a range of models and many simulation years are needed to capture the high uncertainty in internal and model variability.

A key point to note regarding the use of climate projections is that for the model to be useful it needs to have a good representation of the historical climate variables. This is particularly important when the weather data are being used as an input to power system models for which the impact is highly non-linear. For example, a small bias in wind speed leads to a large bias in wind power generation because of the cubic relationship between wind speed and wind power between the cut-in and rated wind speeds (see Figure 12, p. 65). Also, while changes in demand with temperature are somewhat linear for typical winter and summer temperatures, the

48 The actual model time step (or internal temporal resolution) is much less than this. The reason for the widely spaced output interval is to manage the large volumes of data and because for global climate models it isn't expected that output at high resolution temporal scale is representative of reality. The objective is to study trends, not create hourly data that can be used for downstream applications.

49 See the World Climate Research Programme's Coupled Model Intercomparison Project at https://www.wcrp-climate.org/wgcm-cmip.

50 See, for example, https://www.primavera-h2020.eu/.

51 See https://www.euro-cordex.net/ and https://cordex.org/domains/region1-north-america/.

overall relationship of demand sensitivity to temperature is U-shaped, and the tails (the warmest and coldest days) can deviate significantly from linear relationships. Bias correction or calibration techniques are commonly used in climate science to take a multi-decadal historical period of climate model data and compare the parameters of its distribution to historical observations. The mean, variance, or skewness of the distributions can then be corrected (with similar corrections applied to the future climate period, if we assume the bias is stationary) to give more representative data. Bias correction of climate model data is strongly encouraged to make the climate model output as useful as possible when modeling impacts.

Global climate projections are *generally* only available on a daily temporal resolution (with limited high-resolution modeling efforts available at sub-daily resolution). This report has noted the importance of high temporal resolution data for modeling wind and solar accurately. Where high temporal resolution data are not available, there is potential for the daily cycles from historical data to be used to model sub-daily needs, but the potential for climate change to impact historical daily cycles should be considered if relying on historical sub-daily behavior.

Given the coarse temporal resolution of climate data, care should be taken when thinking about impacts at the site of a particular wind plant, or at a county level. The climate model grid boxes provide average conditions over the whole grid cell, and if this grid cell is complex (e.g., containing mountains, coastline, or cities), the results might not be representative of the site.

The applicability of climate change projections for energy system modeling is an emerging area of research. Multiple papers are now available on the impacts of climate change on individual power system aspects (e.g., demand, wind power, and solar power), and historical data from extreme events can be "adjusted" to show how they might have behaved under climate change (e.g., as Bloomfield et al. (2022) do over Europe for extreme temperature scenarios), but results from full power system modeling simulations are more limited. A great deal of research is developing in this field, and future reports will be able to expand on this.

SECTION 8

# Summary and Next Steps

The electric power system is increasingly weather-dependent, with demand as well as supply being influenced by common weather patterns. Consequently, power system planning is increasingly complex and requires longer, higher-resolution datasets with coincident variables (temperature, wind speed/direction, irradiance, etc.). These changes present a number of challenges to the power system modeling community, which needs to simulate power system operations for a variety of reasons including capacity expansion and resource adequacy planning. These studies need to evaluate system behavior across a wide range of potential weather conditions. More weather dependence and complexity require more accurate and comprehensive weather data.

This report has outlined how weather data are used in power system modeling, discussed what power system modelers' and planners' weather data needs are, presented the seven attributes of an ideal dataset, and described a path to creating the ideal weather datasets. While the cost of creating such datasets is not trivial, the cost is low compared to the risks posed by the current data inadequacies—relative to the peril of flying blind.

## Increasing Weather Dependence and Weather Complexity

Going forward, available generation will increasingly be defined by the weather occurring at the location of every wind or solar plant; multiple weather variables in particular, temperature, wind, and solar irradiance, now affect the amount of generation possible. Demand has also long been modulated by weather conditions, especially temperature, and the electrification of

transportation and building conditioning is further increasing its weather dependence. Thus, the possible range of electricity system outcomes is becoming more diverse and far more complex.

In the future, demand will no longer be primarily described as a function of time of year, day of week, time of day, and temperature, and utility-scale generation will no longer be simply a function of available capacity and outage rates. Rather, supply as well as demand will be heavily affected by weather patterns, causing high-risk periods in which weather variables lead to decreased supply and increased load simultaneously.

High-risk events do not have to be "extreme" in the classical sense to pose risks. As more wind and solar generation is added to the resource mix, combinations of events, such as low winds and moderately cold temperatures, will strain the grid.

## The Need for Better Weather Data

When evaluating possible future scenarios of power system build-out in power system planning studies, power system models use historical weather data to determine the possible operating conditions that can occur, and their likelihood. Ideally, the weather data used to determine weather impacts on supply and demand would come from a reliable observational record of past conditions. However, because of the diversity of renewable resource generating locations, such a record does not exist and is not practical to produce. Therefore, models are used to synthesize historical weather datasets.

To be useful, models capture a similar range of possible weather scenarios impacting elements of the electricity system as is observed. Models must be able to represent the level of temporal and spatial granularity that defines the weather impacting load and generation and capture a range and distribution of possible weather scenarios impacting elements of the electricity system similar to that which is observed. Several weather datasets exist and are used by power system planners to estimate load and renewable resource production for use in planning studies. However, no datasets exist that meet the requirements with sufficient accuracy, spatial and temporal resolution, or record length to capture all the possible drivers of supply and demand balance in the new paradigm.

Currently available datasets either:

- Are too low in resolution
- Do not have a long enough time history, and therefore cannot capture the full range of lower-probability events, which are often high-risk periods that must be modeled accurately in system planning analyses
- Are antiquated
- Do not capture all necessary weather variables in a physically consistent fashion

**Models that attempt to reproduce the wind and solar profiles based on predictors like temperature may appear to produce a reasonable long time series, but careful validation will usually reveal a poor match with reality. The limitations are not well documented, the level of uncertainty is not currently well quantified, and the power system sector's understanding of the data is poor.**

Yet model data are often used as if they have the accuracy and degree of uncertainty of observations, when in fact their representativeness in time and space is a function of the model configuration and inputs used. Models that attempt to reproduce the wind and solar profiles for a given day based on predictors like temperature may appear to produce a reasonable long time series where the range of output variables looks as though it reflects reality quite well, but careful validation will usually reveal a poor match with reality, especially when one looks at coincident combinations of different variables across a region. These limitations are not well documented, the level of uncertainty is not currently well quantified, and the power system sector's understanding of the data is poor.

In some cases, synthetic data are being used as if they are direct observations of weather conditions or validated model results, with little or no consideration for how imperfections might impact results and conclusions in the power system studies in which they are being deployed. In other cases, synthetic datasets are rejected, and alternative simpler solutions are deployed that are typically even more problematic.

Observations of key weather variables used in power system modeling inputs are critical to validating the quality of weather model output. These need to be taken at a density that provides some confidence in the weather model output in locations where they impact load and generation. While temperature observations are widely available for validating synthetic temperature data within load centers, there are very few observations available to validate synthetic data produced to estimate wind and solar generation. One solution to this dilemma is to use the many observations made at wind and solar generation facilities; however, these are rarely made available. Broad access to observation archives for existing weather-driven power plants would be very valuable for validating and bias-correcting model data, and it is recommended that facility owners be incented or required to share these data.

## The Attributes of Better Weather Datasets

The most pressing need is to be able to estimate the supply of wind and solar generation in current and future power system portfolios. This means accurately quantifying the weather driving these generators at every plausible location where they exist or may be built, including behind the meter. In addition, the data must represent the chronological evolution of weather to model and optimize charge and discharge of battery storage and demand response.

These data need to:

- Include the necessary variables at sufficient temporal resolution (at least hourly, with 5-minute data needed for some purposes) with sufficient accuracy and spatial resolution[52] to produce meteorological fields that are representative of actual conditions that define the generation potential at current and future wind and solar sites (including those behind the meter) and temperature at load centers
- Cover multiple decades with a consistent methodology so that the range of expected conditions can be quantified, and be extended on an ongoing basis to capture the most recent conditions
- Be coincident, chronological, and physically consistent across weather variables
- Be validated against real conditions with uncertainty quantified
- Be documented transparently and in detail, including a description of their limitations and a guide for usage
- Be periodically refreshed to account for scientific and technological advancements so the data remain relevant
- Be publicly available, expertly curated, and easily accessible

**It is crucial to have comprehensive, standardized, public domain datasets designed specifically for power system modeling activities if the energy transition and renewable energy build-out are to proceed in a cost-effective way while ensuring system reliability and resilience.**

## The Benefits of Better Data

It is crucial to have comprehensive, standardized, public domain datasets designed specifically for power system modeling activities if the energy transition and renewable resource build-out are to proceed in a cost-effective way while ensuring system reliability and resilience. Such datasets will enable more thorough and accurate representation of weather impacts on the supply of renewable resources for any combination of resources, as well as the concurrent weather impacts on demand, traditional generation sources, and transmission. This is essential for renewable integration studies, resource adequacy assessments, capacity expansion modeling, and integrated resource planning.

The data will also benefit other parts of the electricity sector including renewable resource assessments and renewable resource performance analyses. If properly designed and archived, a high-resolution dataset would

52 Grid spacing 2 km or better in complex topography due to rapid variation in the wind field, and 10 km grid spacing or better over flat land or open ocean; and a vertical resolution of 10–50 m in the lowest 300 m of the atmosphere or as high as can be demonstrated to be valuable relative to the computational cost.

be extremely useful for foundational research work to examine the relationships between load and renewable resources, between broader weather patterns and climate signals, and for establishing possible climate trends. Datasets like the one proposed in this report would be a leap forward in the state of the science for describing the condition of the atmosphere at high resolution and thus would be of tremendous interest to many other sectors that are weather data stakeholders.

The work required is not trivial. But it is manageable and, importantly, much less costly than blindly building trillions of dollars of infrastructure without the basic tools to cost-effectively optimize it and assess its reliability.

## Producing the Data We Need

Weather patterns affecting wind and solar production need to be properly captured, which will involve using numerical weather prediction (NWP) models run at resolutions of 2 km grid spacing or less. It is possible that applying generative adversarial network (GAN) machine learning techniques might allow the number of high-resolution simulations to be reduced and thus reduce computational expense.

The first step in producing the weather data needed for power system planning now and going forward is for a technical review committee to vet and refine the dataset requirements. The technical review committee would then determine how to produce the best possible dataset that fits budget constraints by identifying candidate methods and produce a short time series of data (e.g., one year) to compare with widely used datasets like ERA5, the High-Resolution Rapid Refresh (HRRR) model, and the National Solar Radiation Database (NSRDB). Great strides can be made for weather inputs used to estimate wind generation relative to these existing datasets, but improving on the solar data from NSRDB might be more challenging.

Once the methodology for creating the needed dataset has been established, the next step will be to operationalize its production and archive. It will be essential to ensure that the system is well documented, that the data produced are validated against actual conditions, and that data users are provided with the transparent information they need to understand how the data differ from reality and how this might impact power system modeling results.

**The weather is complex, as is the electricity system, and few people have more than a basic grasp of both fields. The lack of holistic understanding is leading to the misapplication of data that can result in invalid power system modeling results and poor decision-making. There is an urgent need for coordination, cooperation, and education between power system experts, meteorologists, and climatologists.**

## The Importance of Cross-Disciplinary Cooperation

The weather is complex, as is the electricity system. Few people have more than a basic grasp of both fields. The lack of holistic understanding is leading to the misapplication of data that can result in invalid power system modeling results and poor decision-making. There is an urgent need for coordination, cooperation, and education between power system experts, meteorologists, and climatologists. It is crucial that power system modelers clearly articulate their data needs, and just as important that the providers of weather and climate data understand how the data are being applied in power system modeling and engage with power system planners to ensure they understand the limitations of the data that are being provided.

With rising levels of wind, solar, and storage and increased electrification, power system planning is becoming more complex and more weather-dependent—with a greater need to accurately model the impacts of weather variables on resource adequacy and system reliability. Accurate modeling requires a validated, high-resolution dataset with a long time series for key weather variables. The availability of such an ideal weather dataset, together with education and coordination between the meteorology and power system communities, will equip system planners to guide future resource siting and build-out for a reliable, high-renewables grid.

# References


Al-Rasheedi, M., C. A. Gueymard, M. Al-Khayat, A. Ismail, J. A. Lee, and H. Al-Duaj. 2020. "Performance Evaluation of a Utility-Scale Dual-Technology Photovoltaic Power Plant at the Shagaya Renewable Energy Park in Kuwait." *Renewable and Sustainable Energy Reviews* 133: 110139. https://doi.org/10.1016/j.rser.2020.110139.

Al-Rasheedi, M., M. Al-Khayat, C. A. Gueymard, S. E. Haupt, B. Kosović, A. Al-Qattan, and J. A. Lee. 2021. "Performance Analysis of a 10 MW Wind Farm in a Hot and Dusty Desert Environment. Part 1: Wind Resource and Power Generation Evaluation." *Sustainable Energy Technologies and Assessments* 47: 101487. https://doi.org/10.1016/j.seta.2021.101487.

Al-Khayat, M., M. Al-Rasheedi, C. A. Gueymard, S. E. Haupt, B. Kosović, A. Al-Qattan, and J. A. Lee. 2021. "Performance Analysis of a 10 MW Wind Farm in a Hot and Dusty Desert Environment. Part 2: Combined Dust and High-Temperature Effects on the Operation of Wind Turbines." *Sustainable Energy Technologies and Assessments* 47: 101461. https://doi.org/10.1016/j.seta.2021.101461.

Alessandrini, S., L. Delle Monache, S. Sperati, and G. Cervone. 2015a. "An Analog Ensemble for Short-Term Probabilistic Solar Power Forecast." *Applied Energy* 157: 95-100. https://doi.org/10.1016/j.apenergy.2015.08.011.

Alessandrini, S., L. Delle Monache, S. Sperati, and J. N. Nissen. 2015b. "A Novel Application of an Analog Ensemble for Short-Term Wind Power Forecasting." *Renewable Energy* 76: 768–781. https://doi.org/10.1016/j.renene.2014.11.061.

Alessandrini, S., and T. McCandless. 2020. "The Schaake Shuffle Technique to Combine Solar and Wind Power Probabilistic Forecasting." *Energies* 13(10): 2503. https://doi.org/10.3390/en13102503.

Allen-Dumas, M. R., K. C. Binita, and C. Colin. 2019. *Extreme Weather and Climate Vulnerabilities of the Electric Grid: A Summary of Environmental Sensitivity Quantification Methods.* Oak Ridge, TN: Oak Ridge National Laboratory. https://basc.pnnl.gov/library/extreme-weather-and-climate-vulnerabilities-electric-grid-summary-environmental-sensitivity.

Alzubaidi, L., J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan. 2021. "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions." *Journal of Big Data* 8: 53. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00444-8#Sec21.

Ayzel, G., T. Scheffer, and M. Heistermann. 2020. "RainNet v1.0: A Convolutional Neural Network for Radar-Based Precipitation Nowcasting." *Geoscientific Model Development* 13(6): 2631–2644. https://gmd.copernicus.org/articles/13/2631/2020/.

Bloomfield, H. C., D. J. Brayshaw, A. Troccoli, C. M. Goodess, M. De Felice, L. Dubus, P. E. Bett, and Y.-M. Saint-Drenan. 2021. "Quantifying the Sensitivity of European Power Systems to Energy Scenarios and Climate Change Projections." *Renewable Energy* 164: 1062–1075. https://doi.org/10.1016/j.renene.2020.09.125.

Bloomfield, H. C., D. J. Brayshaw, M. Deakin, and D. Greenwood. 2022. "Hourly Historical and Near-Future Weather and Climate Variables for Energy System Modelling." *Earth System Science Data* 14(6): 2749–2766. https://doi.org/10.5194/essd-14-2749-2022.

Brotzge, J. A., J. Wang, C. D. Thorncroft, E. Joseph, N. Bain, N. Bassill, N. Farruggio, et al. 2020. "A Technical Overview of the New York State Mesonet Standard Network." *Journal of Atmospheric and Oceanic Technology* 37(10): 1827–1845. https://doi.org/10.1175/JTECH-D-19-0220.1.

Buster, G., M. Bannister, A. Habte, D. Hettinger, G. Maclaurin, M. Rossol, M. Sengupta, Y. Xie. 2022. "Physics-Guided Machine Learning for Improved Accuracy of the National Solar Radiation Database." *Solar Energy* 232: 483-492. https://doi.org/10.1016/j.solener.2022.01.004.

Buster, G., B. Benton, A. Glaws, and R. King. 2023. "Super-Resolution for Renewable Energy Resource Data with Climate Change Impacts (Sup3rCC)." Golden, CO: National Renewable Energy Laboratory. https://doi.org/10.25984/1970814.

Cohen, S. M., A. Dyreson, S. Turner, V. Tidwell, N. Voisin, and A. Miara. 2022. "A Multi-Model Framework for Assessing Long- and Short-Term Climate Influences on the Electric Grid." *Applied Energy* 317: 119193. https://doi.org/10.1016/j.apenergy.2022.119193.

Coughlin, K., and C. Goldman. 2008. *Physical Impacts of Climate Change on the Western U.S. Electricity System: A Scoping Study.* LBNL-1249E. Berkeley, CA: Lawrence Berkeley National Laboratory. https://doi.org/10.2172/944431.

Craig, M., S. Cohen, J. Macknick, C. Draxl, O. J. Guerra, M. Sengupta, S. E. Haupt, B.-M. Hodge, and C. Brancucci. 2018. "A Review of the Potential Impacts of Climate Change on Bulk Power System Planning and Operations in the United States." *Renewable and Sustainable Energy Reviews* 98: 255–267. https://doi.org/10.1016/j.rser.2018.09.022.

Craig, M. T., I. L. Carreño, M. Rossol, B.-M. Hodge, and C. Brancucci. 2019. "Effects on Power System Operations of Potential Changes in Wind and Solar Generation Potential under Climate Change." *Environmental Research Letters* 14(3): 034014. https://doi.org/10.1088/1748-9326/aaf93b.

Craig, M. T., J. Wohland, L. P. Stoop, A. Kies, B. Pickering, H. C. Bloomfield, J. Browell, et al. 2022. "Overcoming the Disconnect Between Energy System and Climate Modeling." *Joule* 6(7): 1405–1417. https://doi.org/10.1016/j.joule.2022.05.010.

Delle Monache, L., F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight. 2013. "Probabilistic Weather Prediction with an Analog Ensemble." *Monthly Weather Review* 141(10): 3498–3516. https://doi.org/10.1175/MWR-D-12-00281.1.

Deroubaix, A., I. Labuhn, M. Camredon, B. Gaubert, P.-A. Monerie, M. Popp, J. Ramarohetra, Y. Ruprich-Robert, L. G. Silvers, and G. Siour. 2021. "Large Uncertainties in Trends of Energy Demand for Heating and Cooling Under Climate Change." *Nature Communications* 12: 5197. https://doi.org/10.1038/s41467-021-25504-8.

Dison, J., A. Dombrowsky, and K. Carden. 2022. *Accrediting Resource Adequacy Value to Thermal Generation.* Report prepared by Astrapé Consulting for Advanced Energy Economy. https://info.aee.net/hubfs/Accrediting%20Resource%20Adequacy%20Value%20to%20Thermal%20Generation-1.pdf.

Draxl, C., A. Clifton , B.-M. Hodge, and J. McCaa. 2015. "The Wind Integration National Dataset (WIND) Toolkit." *Applied Energy* 151: 355–366. https://doi.org/10.1016/j.apenergy.2015.03.121.

Draxl, C., R. P. Worsnop, G. Xia, Y. Pichugina, D. Chand, J. K. Lundquist, J. Sharp, et al. 2021. "Mountain Waves Can Impact Wind Power Generation." *Wind Energy Science* 6(1): 45–60. https://doi.org/10.5194/wes-6-45-2021.

Dubus, L., D. J. Brayshaw, D. Huertas-Hernando, D. Radu, J. Sharp, W. Zappa, and L. P. Stoop. 2022. "Towards a Future-Proof Climate Database for European Energy System Studies." *Environmental Research Letters* 17(12): 121001. https://doi.org/10.1088/1748-9326/aca1d3.

Dyreson, A., N. Devineni, S. W. D. Turner, T. De Silva M., A. Miara, N. Voisin, S. Cohen, and J. Macknick. 2022. "The Role of Regional Connections in Planning for Future Power System Operations Under Climate Extremes." *Earth's Future* 10(6): e2021EF002554. https://doi.org/10.1029/2021EF002554.

ESIG (Energy Systems Integration Group). 2021. *Redefining Resource Adequacy for Modern Power Systems.* A Report of the Redefining Resource Adequacy Task Force. Reston, VA: Energy Systems Integration Group. https://www.esig.energy/wp-content/uploads/2021/08/ESIG-Redefining-Resource-Adequacy-2021.pdf.

Feron, S., R. R. Cordero, A. Damiani, and R. B. Jackson. 2021. "Climate Change Extremes and Photovoltaic Power Output." *Nature Sustainability* 4(3): 270–276. https://doi.org/10.1038/s41893-020-00643-w.

Fitch, A., J. B. Olson, J. K. Lundquist, J. Dudhia, A. K. Gupta, J. Michalakes, and I. Barstad. 2012. "Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model." *Monthly Weather Review* 140(9): 3017–3038. https://doi.org/10.1175/MWR-D-11-00352.1.

Gagne, D. J., S. E. Haupt, D. Nychka, H. Christensen, A. Subramanian, and A. Monahan. 2018. "Generation of Spatial Weather Fields with Generative Adversarial Networks." Presentation at the 4th Conference on Stochastic Weather Generators (SWGEN 2018). Boulder, CO: University Corporation for Atmospheric Research (UCAR). http://n2t.net/ark:/85065/d7cz3b2c.

GE Energy Consulting. 2022. *Economic, Reliability,and Resiliency Benefits of Interregional Transmission Capacity: Case Study Focusing on the Eastern United States in 2035.* Schenectady, NY. https://www.nrdc.org/sites/default/files/ge-nrdc-interregional-transmission-study-report-20221017.pdf.

Gensini, V. A., C. Ramseyer, and T. L. Mote. 2014. "Future Convective Environments Using NARCCAP." *International Journal of Climatology* 34(5): 1699–1705. https://doi.org/10.1002/joc.3769.

Habte, A., M. Sengupta, and A. Lopez. 2017. "Evaluation of the National Solar Radiation Database (NSRDB): 1998–2015." Golden, CO: National Renewable Energy Laboratory. https://www.nrel.gov/docs/fy17osti/67722.pdf.

Hart, E., and A. Mileva. 2022. *Advancing Resource Adequacy Analysis with the GridPath RA Toolkit: A Case Study of the Western U.S*. Berkeley, CA: GridLab. https://gridlab.org/gridpathratoolkit/.

Haupt, S. E., J. Copeland, W. Y. Y. Cheng, Y. Zhang, C. Ammann, and P. Sullivan. 2016. "A Method to Assess the Wind and Solar Resource and to Quantify Interannual Variability over the United States Under Current and Projected Future Climate." *Journal of Applied Meteorology and Climatology* 55(2): 345–363. https://doi.org/10.1175/JAMC-D-15-0011.1.

Haupt, S. E., P. A. Jiménez, J. A. Lee, and B. Kosović. 2017. "Principles of Meteorology and Numerical Weather Prediction." In G. Kariniotakis (ed.), *Renewable Energy Forecasting: From Models to Applications.* Woodhead Publishing, Cambridge, MA. https://doi.org/10.1016/B978-0-08-100504-0.00001-9.

Haupt, S. E., B. Kosović, J. A. Lee, and P. A. Jiménez. 2019. "Mesoscale Modeling of the Atmosphere." In P. Veers (ed.), *Wind Power Modeling and Simulation. Volume 1: Atmosphere and Plant.* Stevenage, UK: IET Publishing. https://doi.org/10.1049/PBPO125F_ch3.

Hausfather, Z. 2017. "Analysis: How Well Have Climate Models Projected Global Warming?" *Carbon Brief*, October 5, 2017. https://www.carbonbrief.org/analysis-how-well-have-climate-models-projected-global-warming/.

Hawkins, E., and R. Sutton. 2008. "The Potential to Narrow Uncertainty in Regional Climate Predictions." *Bulletin of the American Meteorological Society* 90(8): 1095–1108. https://doi.org/10.1175/2009BAMS2607.1.

Hess, P., M.Drüke, S. Petri, F. M. Strnad, and N. Boers. 2022. "Physically Constrained Generative Adversarial Networks for Improving Precipitation Fields from Earth System Models." *Nature Machine Intelligence* 4: 828–839. https://doi.org/10.1038/s42256-022-00540-1.

Hill, J., D. Rupp, N.Voisin, and G. W. Characklis. 2021. "The Effects of Climate Change on Interregional Electricity Market Dynamics on the U.S. West Coast." Preprint, Earth and Space Science Open Archive. http://dx.doi.org/10.1002/essoar.10506741.1.

Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante. 2012. "A Revised Scheme for the WRF Surface Layer Formulation." *Monthly Weather Review* 140(3): 898–918. https://doi.org/10.1175/MWR-D-11-00056.1.

Jiménez, P. A., S. Alessandrini, S. E. Haupt, A. Deng, B. Kosović, J. A. Lee, and L. Delle Monache. 2016. "The Role of Unresolved Clouds on Short-Range Global Horizontal Irradiance Predictability." *Monthly Weather Review* 144(9): 3099–3107. https://doi.org/10.1175/MWR-D-16-0104.1.

Jiménez, P. A., J. A. Lee, B. Kosović, and S. E. Haupt. 2019. "Solar Resource Evaluation with Numerical Weather Prediction Models." In J. Polo, L. Martín-Pomares, and A. Sanfilippo (eds.), *Solar Resources Mapping: Fundamentals and Applications*, Springer Cham. https://doi.org/10.1007/978-3-319-97484-2_7.

Judt, F. 2018. "Insights into Atmospheric Predictability Through Global Convection-Permitting Model Simulations." *Journal of the Atmospheric Sciences* 75(5): 1477–1497. https://doi.org/10.1175/JAS-D-17-0343.1.

Judt, F. 2020. "Atmospheric Predictability of the Tropics, Middle Latitudes, and Polar Regions Explored Through Global Storm-Resolving Simulations." *Journal of the Atmospheric Sciences* 77(1): 257–276. https://doi.org/10.1175/JAS-D-19-0116.1.

Juliano, T. W., P. A. Jiménez, B. Kosović, T. Eidhammer, G. Thompson, L. K. Berg, J. Fast, A. Motley, and A. Polidori. 2022. "Smoke from 2020 United States Wildfires Responsible for Substantial Solar Energy Forecast Errors." *Environmental Research Letters* 17(3): 034010. https://doi.org/10.1088/1748-9326/ac5143.

King, J., A. Clifton, and B. M. Hodge. 2014. *Validation of Power Output for the WIND Toolkit.* NREL/TP-5D00-61714. Golden, CO: National Renewable Energy Laboratory. https://doi.org/10.2172/1159354.

Lam, R., A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, et al. 2022. "GraphCast: Learning Skillful Medium-Range Global Weather Forecasting." Preprint. arXiv:2212.12794v2 [cs.LG]. https://doi.org/10.48550/arXiv.2212.12794.

Lorenz, E. 1972. "Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?" Presentation to the Global Atmospheric Research Program, American Association for the Advancement of Science, Washington, DC, December 29, 1972, printed as an appendix in E. Lorenz, *The Essence of Chaos*, University of Washington Press, 1995. https://climate.envsci.rutgers.edu/climdyn2017/LorenzButterfly.pdf.

McFarland, J., Y. Zhou, L. Clarke, P. Sullivan, J. Colman, W. S. Jaglom, M. Colley, et al. 2015. "Impacts of Rising Air Temperatures and Emissions Mitigation on Electricity Demand and Supply in the United States: A Multi-Model Comparison." *Climatic Change* 131(1): 111–125. https://doi.org/10.1007/s10584-015-1380-8.

McGovern, A., R. Lagerquist, D. J. Gagne II, G. E. Jergensen, K. L. Elmore, C. R. Homeyer, and T. Smith. 2019. "Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning." *Bulletin of the American Meteorological Society* 100(11): 2175–2199. https://doi.org/10.1175/BAMS-D-18-0195.1.

Miara, A., S. M. Cohen, J. Macknick, C. J. Vörösmarty, F. Corsi, Y. Sun, V. Tidwell, R. Newmark, and B. M. Fekete. 2019. "Climate-Water Adaptation for Future U.S. Electricity Infrastructure." *Environmental Science and Technology* 53(23): 14029–14040. https://doi.org/10.1021/acs.est.9b03037.

Milligan, M., E. Ela, D. Lew, D. Corbus, Y. Wan, B. Hodge, and B. Kirby. 2012. "Operational Analysis and Methods for Wind Integration Studies." *IEEE Journal on Sustainability* 3(4): 612–619. http://dx.doi.org/10.1109/TSTE.2011.2160881.

Molina, M. O., C. Gutiérrez, and E. Sánchez. 2021. "Comparison of ERA5 Surface Wind Speed Climatologies over Europe with Observations from the HadISD Dataset." *International Journal of Climatology* 41(10): 4864–4878. https://doi.org/10.1002/joc.7103.

Murphy, S., F. Sowell, and J. Apt. 2019. "A Time-Dependent Model of Generator Failures and Recoveries Captures Correlated Events and Quantifies Temperature Dependence." *Applied Energy* 253: 113513. https://doi.org/10.1016/j.apenergy.2019.113513.

Nahmmacher, P., E. Schmid, M. Pahle, and B. Knopf. 2016. "Strategies Against Shocks in Power Systems: An Analysis for the Case of Europe." *Energy Economics* 59: 455–465. https://doi.org/10.1016/j.eneco.2016.09.002.

Nakanishi, M., and H. Niino. 2006. "An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog." *Boundary-Layer Meteorology* 119: 397–407. https://doi.org/10.1007/s10546-005-9030-8.

NERC (North American Electric Reliability Corporation). 2014. *Task 1.6: Probabilistic Methods.* https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/IVGTF%20Task%201-6_09182014.pdf.

Novacheck, J., J. Sharp, M. Schwarz, P. Donohoo-Vallett, Z. Tzavelis, G. Buster, and M. Rossol. 2021. "The Evolving Role of Extreme Weather Events in the U.S. Power System with High Levels of Variable Renewable Energy." NREL/TP-6A20-78394. Golden, CO: National Renewable Energy Laboratory. https://www.nrel.gov/docs/fy22osti/78394.pdf.

O'Connell, M., N. Voisin, J. Macknick, and T. Fu. 2019. "Sensitivity of Western U.S. Power System Dynamics to Droughts Compounded with Fuel Price Variability." *Applied Energy* 247: 745–754. https://doi.org/10.1016/j.apenergy.2019.01.156.

Panteli, M., and P. Mancarella. 2015. "Influence of Extreme Weather and Climate Change on the Resilience of Power Systems: Impacts and Possible Mitigation Strategies." *Electric Power Systems Research* 127: 259–270. https://doi.org/10.1016/j.epsr.2015.06.012.

Peter, J. 2019. "How Does Climate Change Affect Electricity System Planning and Optimal Allocation of Variable Renewable Energy?" *Applied Energy* 252: 113397. https://doi.org/10.1016/j.apenergy.2019.113397.

Pronk, V., N. Bodini, M. Optis, J. K. Lundquist, P. Moriarty, C. Draxl, A. Purkayastha, and E. Young. 2022. "Can Reanalysis Products Outperform Mesoscale Numerical Weather Prediction Models in Modeling the Wind Resource in Simple Terrain?" *Wind Energy Science* 7(2): 487–504. https://doi.org/10.5194/wes-7-487-2022.

Pryor, S. C., R. J. Barthelmie, and T. J. Shepherd. 2020. "20% of U.S. Electricity from Wind Will Have Limited Impacts on System Efficiency and Regional Climate." *Scientific Reports* 10: 541. https://www.nature.com/articles/s41598-019-57371-1.

Rackauckas, C., Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, and A. Edelman. 2021. "Universal Differential Equations for Scientific Machine Learning." Digital preprint. https://arxiv.org/abs/2001.04385.

Ralston Fonseca, F., M. Craig, P. Jaramillo, M. Bergés, E. Severnini, A. Loew, H. Zhai, et al. 2021. "Effects of Climate Change on Capacity Expansion Decisions of an Electricity Generation Fleet in the Southeast U.S." *Environmental Science & Technology* 55(4): 2522–2531. https://doi.org/10.1021/acs.est.0c06547.

Rosencrans, D., B. Benton, G. Buster, A. Glaws, R. King, J. K. Lundquist, J. Gu, and G. Maclaurin. 2023. "Wind Resource Data for Southeast Asia Using a Hybrid Numerical Weather Prediction with Machine Learning Super Resolution Approach." Golden, CO: National Renewable Energy Laboratory; Washington, DC: U.S. Agency for International Development. https://pdf.usaid.gov/pdf_docs/PA021299.pdf.

Schlott, M., A. Kies, T. Brown, S. Schramm, and M. Greiner. 2018. "The Impact of Climate Change on a Cost-Optimal Highly Renewable European Electricity Network." *Applied Energy* 230: 1645–1659. https://doi.org/10.1016/j.apenergy.2018.09.084.

Sengupta, M., A. Habte, A. Lopez, and Y. Xi. 2015a. "Delivering Data for the PV User from Physics-Based Satellite Models." Presentation at the PV Solar Resource Workshop, February 27, 2015. Golden, CO: National Renewable Energy Laboratory. https://www.nrel.gov/pv/assets/pdfs/2015_pvmrw_161b_sengupta.pdf.

Sengupta, M., A. Weekley, A. Habte, A. Lopez, and C. Molling. 2015b. "Validation of the National Solar Radiation Database (2005-2012)." Preprint. Presented at the European PV Solar Energy Conference and Exhibition, Hamburg, Germany, September 14-18, 2015. Golden, CO: National Renewable Energy Laboratory. https://doi.org/10.13140/RG.2.1.4103.3367.

Sharp, J. 2022. *Meteorological Deep Dive of Low Renewable Energy Periods in Accelerated 2030 California Clean Electricity Portfolios*. Berkeley, CA: GridLab. https://gridlab.org/california-2030-study/.

Sharp, J., and C. F. Mass. 2002. "Columbia Gorge Flow: Insights from Observational Analysis and Ultra-High Resolution Model Simulation." *Bulletin of the American Meteorological Society* 18: 75–79.

Skamarock, W. 2004. "Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra." *Monthly Weather Review* 132(12): 3019–3032. https://doi.org/10.1175/MWR2830.1.

Steinberg, D. C., B. K. Mignone, J. Macknick, Y. Sun, K. Eurek, A. Badger, B. Livneh, and K. Averyt. 2020. "Decomposing Supply-Side and Demand-Side Impacts of Climate Change on the U.S. Electricity System Through 2050." *Climatic Change* 158: 125–139. https://doi.org/10.1007/s10584-019-02506-6.

Stenclik, D. 2022. "Best Practices for Modeling Extreme Weather in Power Systems." *Telos Energy*, blog. December 26, 2022. https://www.telos.energy/post/best-practices-for-modeling-extreme-weather-in-power-systems.

Stenclik, D., M. Welch, and P. Sreedharan. 2022. *Reliably Reaching California's Clean Electricity Targets: Stress Testing Accelerated 2030 Clean Portfolios.* Berkeley, CA: GridLab. https://gridlab.org/california-2030-study/.

Stengel, K., A. Glaws, D. Hettinger, and R. N. King. 2020. "Adversarial Super-Resolution of Climatological Wind and Solar Data." *Proceedings of the National Academy of Sciences* 117: 16805–16815. https://doi.org/10.1073/pnas.1918964117.

Su, Y., J. D. Kern, P. M. Reed, and G. W. Characklis. 2020. "Compound Hydrometeorological Extremes Across Multiple Timescales Drive Volatility in California Electricity Market Prices and Emissions." *Applied Energy* 276: 115541. https://doi.org/10.1016/j.apenergy.2020.115541.

Turner, S. W. D., N. Voisin, J. Fazio, D. Hua, and M. Jourabchi. 2019. "Compound Climate Events Transform Electrical Power Shortfall Risk in the Pacific Northwest." *Nature Communications* 10(1). https://doi.org/10.1038/s41467-018-07894-4.

Voisin, N., A. Dyreson, T. Fu, M. O'Connell, S. Turner, T. Zhou, and J. Macknick. 2020. "Impact of Climate Change on Water Availability and its Propagation Through the Western U.S. Power Grid." *Applied Energy* 276: 115467. https://doi.org/10.1016/j.apenergy.2020.115467.

Voisin, N., M. Kintner-Meyer, D. Wu, R. Skaggs, T. Fu, T. Zhou, T. Nguyen, and I. Kraucunas. 2018. "Opportunities for Joint Water–Energy Management: Sensitivity of the 2010 Western U.S. Electricity Grid Operations to Climate Oscillations." *Bulletin of the American Meteorological Society* 99(2): 299–312. https://doi.org/10.1175/BAMS-D-16-0253.1.

Wang, B., Y. Zhou, P. Mancarella, and M. Panteli. 2016. "Assessing the Impacts of Extreme Temperatures and Water Availability on the Resilience of the GB Power System." 2016 IEEE International Conference on Power System Technology (POWERCON), pp. 1–6, Wollongong, NSW, Australia. https://doi.org/10.1109/POWERCON.2016.7754047.

Wang, J., Z. Liu, I. Foster, W. Chang, R. Kettimuthu, and V. Rao Kotamarthi. 2021. "Fast and Accurate Learned Multiresolution Dynamical Downscaling for Precipitation." *Geoscientific Model Development* 14(10): 6355–6372. https://doi.org/10.5194/gmd-14-6355-2021.

Warner, T. T., R. A. Peterson, and R. E. Treadon. 1997. "A Tutorial on Lateral Boundary Conditions as a Basic and Potentially Serious Limitation to Regional Numerical Weather Prediction." *Bulletin of the American Meteorological Society* 78(11): 2599–2618. https://doi.org/10.1175/1520-0477(1997)078<2599:ATOLBC>2.0.CO;2.

Warner, T. T. 2011. "Quality Assurance in Atmospheric Modeling." *Bulletin of the American Meteorological Society* 92(12): 1601–1610. https://doi.org/10.1175/BAMS-D-11-00054.1.

Xie, Y., M. Sengupta, and J. Dudhia. 2016. "A Fast All-Sky Radiation Model for Solar Applications (FARMS): Algorithm and Performance Evaluation." *Solar Energy* 135 (October): 435–445. https://doi.org/10.1016/j.solener.2016.06.003.

Yang, B., Y. Qian, L. K. Berg, P.-L. Ma, S. Wharton, V. Bulaevskaya, H. Yan, Z. Hou, and W. J. Shaw. 2016. "Sensitivity of Turbine-Height Wind Speeds to Parameters in Planetary Boundary-Layer and Surface-Layer Schemes in the Weather Research and Forecasting Model." *Boundary-Layer Meteorology* 162: 117–142. https://doi.org/10.1007/s10546-016-0185-2.

Yin, J., A. Molini, and A. Porporato. 2020. "Impacts of Solar Intermittency on Future Photovoltaic Reliability." *Nature Communications* 11: 4781. https://www.nature.com/articles/s41467-020-18602-6.

Zeng, Z., A. D. Ziegler, T. Searchinger, L. Yang, A. Chen, K. Ju, S. Piao, et al. 2019. "A Reversal in Global Terrestrial Stilling and Its Implications for Wind Energy Production." *Nature Climate Change* 9: 979–985. https://www.nature.com/articles/s41558-019-0622-6.

# Glossary

### Background field

**A first guess of the atmospheric state used in numerical weather prediction, usually obtained from the output of a prior model run.** Available observations are applied to the background field during the data assimilation process to arrive at the model initial condition.

### Bootstrapping

**A statistical procedure that resamples a dataset with limited data points to synthetically create a broader sample size with the objective of more completely capturing the range of possible values.** For example, one can bootstrap three years of temperature data at a site into 50 years using relationships with daily temperature at nearby sites.

### Boundary conditions

**In numerical weather prediction, the conditions at the edge of the region being modeled that define how the state of the system outside of the area being modeled propagates into the simulation at the edges**

### Capacity factor

**A normalized measure of output of a generator (whether conventional coal or gas, solar, wind, etc.) expressed as a percentage of total possible generation during the same time period.** For example, a generator with a maximum output of 100 MW producing 50 MW is running with a 50% capacity factor.

### Downscaling

**In NWP, the process of taking model output at a given fidelity (grid spacing) and applying it to a model that uses a higher spatial (and possibly temporal) resolution.** Because the higher-resolution model represents both static features (like model topography) and dynamic features (like temperature and wind speed) with more fidelity, the downscaled results will reflect the features that are produced as a result of the interaction of the large-scale flow with smaller-scale environmental features.

### Dynamically consistent

**Adhering to relationships determined by physical laws that bind together different atmospheric variables in time and space according to well-defined mathematical relationships (a synonym of physically consistent).** Observations of weather variables (for example, temperature and wind speed) that are coincident in time and space are always dynamically consistent with each other, as are the output fields from physics-based models.

Though output fields from a physics-based model are dynamically consistent with each other, this does not mean that physics-based model fields from two different models or from a model and observations will be dynamically consistent with each other. Data output from a statistical model, or multiple statistical models, will not be dynamically consistent unless the statistical model implicitly captures the dynamics of the system (this is starting to happen with advanced machine learning techniques).

### Energy-limited resource

**A resource that can only deliver a limited amount of energy before becoming unavailable.** For example, batteries and pumped storage can only provide generation until the energy stored in them is depleted, at which point they must be recharged. Likewise, there is a limit to the amount of time that a demand response program can ask consumers to cut back on their usage.

### Ensemble dataset

**A dataset containing two or more versions of the same data created in different ways.** For example, a time series estimate of temperature, wind speed, and irradiance for the exact same times and locations, but generated by three different configurations of model physics, is a three-member ensemble, often called a physics ensemble.

### Firm capacity

**A colloquial term usually used either to indicate generators that are almost always available to generate at rated output when they are online unless they are forced into an outage by unforeseen circumstances like equipment failure (traditionally applied to coal, natural gas, and nuclear plants) or to indicate the amount of capacity a generator is likely to provide during times of tight supply conditions.** All generators can be considered as firm under certain conditions. For example, a 100 MW solar plant can be considered as being able to provide 80 MW of firm capacity in the middle of the day on a cloud free day when no clouds formation is expected.

### Future-proof

**Able to serve the needs of power system modelers and planners as the transition to a high-renewables grid proceeds, and as new methods emerge that may make prior iterations of the data defunct**

### Gate closure time

**In wholesale electricity markets, a lead time before actual operations when decisions are made.** For example, all of the bids for a day-ahead market must be received before a certain time the previous day so that they can be used in the unit commitment process. Real-time markets and the sending of dispatch instructions also have closure times. Wind and solar generation forecasts that are used to inform these processes need to bereceived before the closure time to be useful. Another example is the time when natural gas has to be purchased ahead of when it is needed on the power system.

### Global climate model

**A model designed to simulate the physical processes occurring in the Earth system to produce output representing the state of the system for years, decades, or even centuries, into the future.** Global climate models couple together all of the components affecting the climate system—the atmosphere, oceans, cryosphere (ice), and land surface—including modeling and tracking components that may change in time, such as greenhouse gas concentrations, ice coverage, ocean temperatures, aerosols, solar irradiance, orbital changes, vegetation coverage, etc., all of which feedback on one another.

### Grid point

**A discrete point in space where the properties of a system being modeled are represented and tracked.** The distance between a model's grid points is one of the primary determinants of how well the model can represent the real world.

### Gridded data

**A dataset that contains the value of different fields (for example, temperature and wind) across many grid points.** Gridded data can be 1D (a line of points), 2D (an array of points representing a horizontal or vertical slice), or 3D (an array of points representing a volume) in space, and can also have a time component (i.e., each grid is repeated for different times).

### Ground truth

**The actual observed value for a quantity, as opposed to a model-synthesized or -forecasted value**

### Initial condition

**The state of a time-dependent dynamical system.** For example, the initial condition in a numerical weather prediction model represents, as closely as possible, the state of the atmosphere at the starting point of a simulation that is aimed at predicting the future atmospheric state.

### Lead time

**Distance into the future of a forecast.** For NWP models, lead time is usually expressed as how far into the future the forecast is, relative to the time that the initial condition represents, not the time difference between when the forecast is issued and when the forecast is for (valid time). For example, if the state of the atmosphere is observed at 4:00 am, assimilated into an initial condition by 6:00 am, and used to produce a forecast at 7:50 am that provides a forecast of condition at 10:00 pm, the lead time is stated as 18 hours (from 4:00 am to 10:00 pm). It is stated this way because the further into the future being predicted, the lower the accuracy is likely to be.

In contrast, in the power sector lead times are considered as referencing the time into the future from *the time the forecast was issued*. So, if the NWP output from the example above was fed into a generation prediction model as soon as it was produced at 7:50 am, and produces a forecast by 8:00 am of the expected wind generation at 10:00 pm, the lead time will usually be stated as 14 hours (8:00 am to 10:00 pm). This can cause considerable confusion.

### Monte Carlo methods

**A class of modeling that uses repeated random sampling from a probability distribution and uses the sampled value as the input to a deterministic equation or algorithm.** By repeated random sampling within the input distribution, the range of output possibilities can be determined provided the input probabilities are known and process inputs are independent.

### Normalize

**To scale related quantities so that they can be compared.** For example, take wind farms A, B, and C in three different locations. The wind farms have maximum output capacities of 50 MW, 100 MW, and 200 MW. A's average annual output is 25 MW, B's is 40 MW, and C's is 80 MW. To answer the question, "Which produces the most power relative to its size?" the data need to be normalized to a standard scale, for example, output as a percentage of maximum capacity (the definition of capacity factor, given above). We then find that A has a capacity factor of 50% while B and C have capacity factors of 40%.

### Numerical method

**The process of using an algorithm to solve or approximate the solution to a problem, usually using a computer that iterates through many calculations to converge on a result.** Numerical methods enable the solution of mathematical problems, such as those involved in weather prediction, that are difficult or impossible to solve analytically.

### Numerical weather prediction

**Solving the equations that govern the state and motion of the atmosphere using numerical methods, so that if the state of the atmosphere at one time is known, the state at a nearby time can be estimated**

### Overfitting

**When a complex relationship between multiple variables is found within a data sample of model output, but, when the relationship is applied to data outside of the sample, validation shows it is not robust**

### Parameterization (or scheme)

**A sub-model that empirically simulates a particular meteorological process that either cannot be modeled explicitly or is computationally too difficult to model explicitly**

### Physics-based (or physical) model

**A model of atmospheric processes that adheres to physical laws and can be described mathematically as a system of regular and partial differential equations** (see also "numerical method")

### Physically consistent

**A synonym of dynamically consistent**

### Resource adequacy

**An analytical process that evaluates whether there are enough resources on the power system to serve load across a wide range of potential conditions, including load, weather, and unexpected outages.** The results help determine the amount of resources needed to be able to serve load at a specified level of reliability.

### Reanalysis

For context, weather analysis is a process that takes available weather data and uses them together with knowledge of the laws of physics to estimate the state of the atmosphere and is the first step in the forecasting process. **Reanalysis is similar to weather analysis but occurs after the fact when all of the possible data are available, including what would have been the future state of the atmosphere.** Through the use of sophisticated computer codes, reanalysis reconciles all the data from observations and past, current, and future model estimates in an effort to produce the most accurate weather analysis possible.

### Thermal generator

**A generator using combustion, fission, or geothermal heat to drive a turbine to produce electricity—i.e., coal, natural gas, and nuclear plants.** Turbines may be turned directly by hot exhaust gases (e.g., a gas combustion turbine), by steam produced from the heat source (e.g., coal and nuclear plants), or by a combination of both (gas combined-cycle plants).

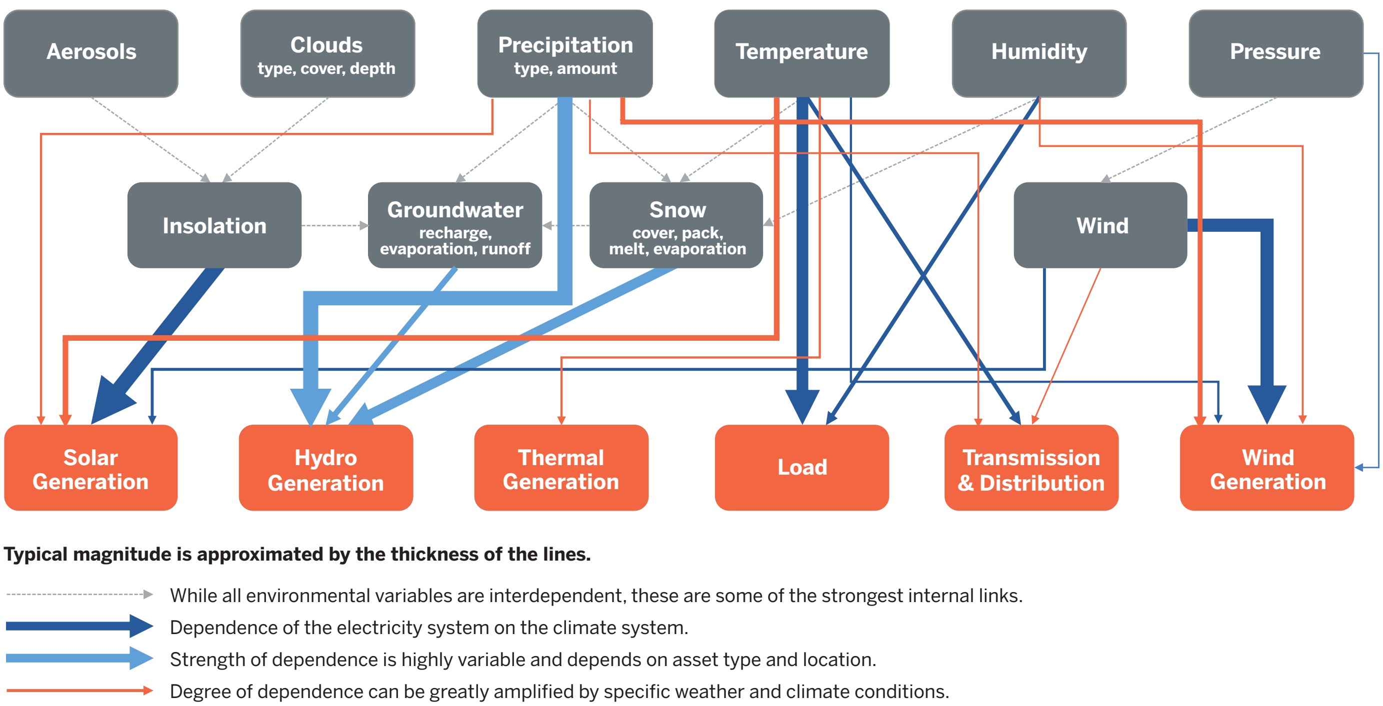
### Transfer function

**A mathematical function that relates one (or several) input variables to an output variable.** For example, a wind turbine power curve is a transfer function that relates wind speed and air density to power output.

# Appendix A: Additional Insights into Linkages Between Electricity Systems and Meteorology

As documented in Section 1, weather and climate interact with power systems in numerous different interconnected ways. The most important variables are temperature, insolation, and wind speed. Load, wind generation, and solar generation are all influenced to some degree by these fields. Their relative level of importance depends on an electric system's load types and the relative amount of wind and solar generation compared to hydro and thermal generators. Precipitation is very important to systems with significant hydro generation, and in times of drought can also affect cooling water for thermal generators. Figure A.1 shows the interactions between the key variables, depicts the other environmental factors affecting each

**FIGURE A.1**
Electricity System Weather-Dependence




**The primary linkages between variables in the weather and climate system (gray) and the electricity system (orange). There are many feedbacks between all of the environmental variables; the strongest links are shown in dashed gray lines. Dark blue lines indicate direct dependencies that are most important in everyday operation of the electricity system, while orange lines indicate dependencies that do not typically have a large impact on a daily basis but can have a profound impact in particular circumstances or combinations. For instance, freezing temperatures, high humidity and/or freezing rain can cause wind generation to become unavailable due to icing, and extreme winds can damage transmission and distribution infrastructure.**

Source: Energy Systems Integration Group.

component of a power systems, and illustrates the complexity of the coupled Earth system and electricity system. While the strongest internal linkages between the different environmental factors are shown, every weather variable impacts all others in some way. Some combinations of weather variables, such as severe winter storms with low temperatures and frozen precipitation, or hot, dry days with strong gusty winds, can pose profoundly amplified risks to the power system.

The electric power sector uses meteorological data for three general purposes: (a) development of weather-dependent generation and assessment of the performance of that generation, (b) system planning and adequacy, and (c) operations. Each has its own spatial and temporal requirements for assessment of weather and climate, and thus uses different types of weather and climate data. The body of this report focuses on system planning and adequacy studies. Here we provide some background on how data are used for the other two categories to allow interested readers to compare and contrast them and explore synergies that exist in applicability of existing and proposed datasets.

Figure A.2 (p. 124) illustrates a typical flow of weather data into the three main power system use cases (orange boxes), showing the relationships of the data sources (light blue boxes). All data, except for projections of future climate, are ultimately derived from the environmental observations concerning the state of the atmosphere (light orange box), but there is a modeling layer (dark blue boxes) where a great deal of computation occurs to take the weather observation data that are available and extrapolate them to a regular grid that provides estimates of the key variables seen in Figure A.1. This modeling process is essential because the available near-surface observations do not begin to provide the necessary level of detail, especially to estimate wind and solar generation at all of the locations where estimates are needed. However, the modeling layer also introduces uncertainty that cannot be ignored, but, unfortunately, often is.

The weather and climate data are then used both as an input to produce estimates of past, present, or future states of the electricity system (medium blue boxes) and directly in analysis done for each use case. The dotted box denotes processes where it is sometimes desirable

to utilize climate predictions in order to assess the impacts of climate change.

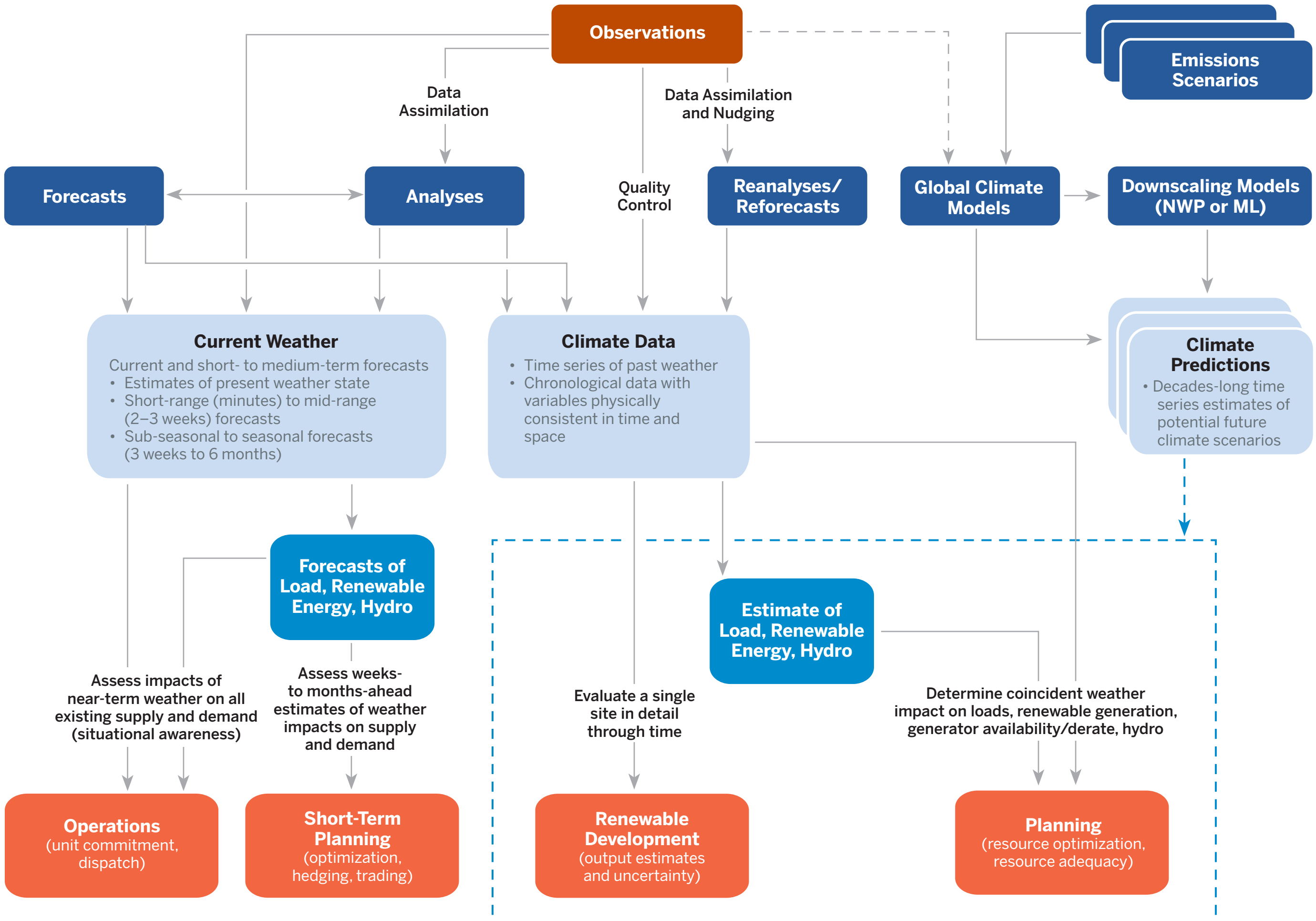
When using meteorological data in power system planning, it is important to understand the source of the data and how well they capture the spatial and/or temporal variability of weather phenomena that affects the electricity system components one is trying to assess. The quality and applicability of model outputs depend on the quality of the model used, its resolution (the fidelity with which the model is run), and the quality of the original observations. For example, if performing resource adequacy studies on a winter-peaking system, output from a model that commonly over-predicts wind speed on cold days in regions with large amounts of wind generation is unlikely to properly capture the correlated weather risk of low temperatures driving both loads and low wind generation.

## Weather Data for Resource Assessment and Performance Assessment of Renewable Energy Projects

Meteorological data are the main input to the resource assessment process that determines the expected output of proposed renewable energy projects. The objective of these assessments is to obtain the best overall energy estimate possible within a reasonable budget and time span. It is crucial to maximize accuracy and minimize uncertainty in the estimate of plant output to determine

FIGURE A.2

## Flow of Weather Data into the Primary Power Systems Processes That Require Meteorological Data for Analysis Tasks




**Weather data (light blue and dark blue boxes) flow into the primary power systems processes (orange boxes) that require meteorological data for analysis tasks. In some cases, this includes tasks that estimate states of the electricity system (medium blue boxes) that are impacted by the state of the atmosphere. Solid lines indicate the flow of data that are output from one process and become an input to another. The dashed line from observations to global climate models indicates that the data are used in validation of average global climate model outputs versus as an input to their production.**

Notes: ML = machine learning; NWP = numerical weather prediction.

Source: Justin Sharp, adapted for the Energy Systems Integration Group.

the viability of a renewable energy project and obtain financing. The estimate is usually broken down by average output for the month and hour of day and also presented in the form of 8,760 time series data, showing estimates for the average expected output for each hour of the year. Variability and uncertainty are also estimated. Gross output for proposed wind and solar projects needs to be adjusted for losses, many of which are weather-related items such as blade icing; blade pitting due to lightning; high wind cut-out; high and low temperature cut-out for wind projects; and panel or blade soiling, smoke, and snow cover on solar panels.

For wind energy, the wind at multiple heights is needed to determine the shear across the rotor, and one needs to determine the effective wind across the rotor sweep and the suitability of the turbine locations from the perspective of shear forces, turbulence, inflow angle, and peak expected wind. In addition, the resource can vary considerably across a site due to microscale meteorological effects, usually driven by topographical effects. Models are usually used to estimate these effects across the multiple turbine locations, with a handful of short observing campaigns being performed to evaluate the validity, bias, and uncertainty of the model output. Wake effects are also estimated using models once the turbine layout has been determined.

The primary variables to predict wind generation are of course wind speed and direction. Wind generation is also a function of air density, which can be derived from temperature and pressure. In addition, relative humidity is important for assessing icing risk. Solar irradiance is most important for solar generation, and photovoltaic generation is also significantly impacted by panel temperature, which is a function of ambient temperature and wind speed. The sources of all these data points need to be concurrent because they are related. For example, generation lost to icing depends not just on the amount of icing, but how strong the wind is when it occurs.

To meet the above goals, resource assessment usually involves conducting measurement campaigns at locations of interest that last one or more years. Usually a combination of met masts, sodars, and lidars are used so that the wind can be measured at multiple height levels at multiple sites. Where masts are used, temperature is usually measured at two or more levels so that atmospheric stability can be assessed. The observational data need to have few gaps, be taken frequently enough to assess variability and (in the case of wind energy) turbulence, and provide for hourly averaging. Thus, observations are typically taken every five minutes using platforms deployed specifically for resource assessment that are closely monitored and maintained. Because measurement campaigns are not usually long enough to determine a full climatology, the observed data then need to be put into the context of the broader climate. This is usually done using measure, correlate, and predict (MCP) methodologies that use overlapping periods between the measurement campaign and longer time series from either numerical weather prediction output, nearby in-situ observations, or remote-sensed data, to

normalize the campaign data relative to typical climate. This dataset should be at least a decade long, and ideally span multiple decades to fully capture the interannual variability. However, it does not have to be of equivalent quality so long as the measurement campaign and spatial modeling properly capture the characteristics of the resource.

After a wind or solar project is built, the resource still needs to be measured for several reasons, including providing a way to measure the power generation performance of the project relative to the wind and solar resource to ensure that it is operating as expected, to compare with the original resource assessment and provide feedback into the accuracy of the process, and to provide site data for situational awareness and forecasting in operations. Usually, the resource assessment measurements used prior to construction are designed to be temporary and are removed when the project is constructed, with new permanent observations replacing them. In the case of wind facilities, wake effects of the new project need to be considered in the placement and use of these observations and when comparing them to the pre-build data.

Proposed and operational plant meteorology data need to be highly representative of the actual conditions at the plant, have quantifiable uncertainty, and be long enough to capture climatological variability and extremes in the fields driving output and possible outages/derates. For wind, the data must provide insights into how the resource varies across the site footprint. However, they do not need to cover areas outside the project. This contrasts with data for grid operations, which need to cover resource information for all generators in the balancing area, and for planning/adequacy studies, where data need to cover all possible places where renewable projects are proposed.

Ideally, some way to estimate the possible range of resource changes in an evolving climate should also be available, but research into this is in its infancy. The best currently available methodology is to continue to keep consistent, high-quality records at project sites and begin to look for trends.

## Real-Time Operation of the Electricity System

Weather inputs to electricity system operation focus mostly on situational awareness and short-range forecasting—what is happening now and what is predicted to happen in the next few days. Some focus is placed on medium range (one to two weeks) and seasonal projections (the next few months), but most of the emphasis is on the next two days. Weather inputs are used to:

- Forecast expected demand, net of any behind-the-meter variable generation
- Forecast wind, solar, and hydro generation
- Determine the likelihood of derates and outages of all generation types and transmission
- Assess the quality of forecasts relative to current information
- Assess risks to transmission and distribution infrastructure due to wind, icing, or fire

This information relies on observational and forecast data that then feed into power markets, unit commitment, and dispatch processes to ensure that supply and demand are balanced in the most effective way given generation and demand forecasts.

The core attributes of these weather inputs are that they:

- Are timely: current conditions must be available in near-real time, and forecasted conditions need to be available in time for market and operational gate closures.
- Are reasonably accurate: greater accuracy would be helpful, but the current accuracy is acceptable. (One priority for improvement is to improve forecasting for high-impact, low-probability events.)
- Provide a general estimate of net load and bulk forecast: the geographical area of consideration is the generation and load footprint, with the priority being estimating the net load accounting for behind-the-meter renewables and the bulk forecast of output of wind and solar plants (as opposed to intra-site detail).

# Appendix B: Weather 201

This appendix augments Section 2, "Meteorological Data Fundamentals for Power System Planning," for the benefit of readers who would like more details.

## In-Situ Observations

In-situ observations provide measurements specific to their location. In-situ measurements are appealing, as their uncertainty and quality are usually easy to quantify, the instrumentation is relatively cheap, and they often have long records. However, their spatial coverage is typically limited.

In-situ measurements have been taken around the world for centuries, and long records are available at some sites. Examples include thermometers, precipitation gauges, and barometers. Uncertainty and accuracy depend on the instrument specifications, placement, and maintenance. Most in-situ observations are fixed in space and are typically surface-based, with towers used to gather measurements from multiple near-surface levels. Another form of in-situ measurement uses radiosondes, an instrument package carried aloft by a weather balloon. These instruments report the instrument location as part of the data collected.

## Remotely Sensed Observations

Remote-sensing instruments either observe atmospheric data from somewhere remote from the measurement location (passive sensing) or send out a signal and observe the interaction of the signal with the atmosphere (active sensing). This means that remote-sensing devices can gather data from large areas or volumes by scanning across them. Examples include cameras (a passive sensor) flying on orbiting satellites and weather radars (an active sensor that sends out a pulse of radio waves and measures the reflected signal).

Remotely sensed data from a vast array of instruments located both on satellites and on the ground are now recorded in large quantities, often at high spatial and temporal resolution. Examples are weather radars, atmospheric sounders, and atmospheric imagers. These instruments usually measure at multiple locations along a line or within a volume. Often, the instruments are space-based, in which case they may either be in geostationary orbits, which always have the same field of view of the Earth and thus provide frequent observations within their view, or be in an orbit that transits different parts of the planet, thus covering a broader field of view but with less frequent observations at any given location. Remotely sensed data have revolutionized our ability to diagnose the four-dimensional state of the atmosphere and are a critical input to models that produce widely used gridded datasets derived from numerical weather prediction (NWP) and other types of modeling.

Some major complexities are associated with remotely sensed data that need to be understood if one is using the data directly without expert guidance. The quantities measured sometimes have complex relationships to the atmospheric variables that are derived from them and require significant processing to arrive at the atmospheric data. Further, atmospheric conditions can affect sensitivity, accuracy, and range. For example, weather radar measures atmospheric reflectivity, and this is a function of precipitation type among other factors, and heavy precipitation will limit range.[1] The instrument response may be quite nuanced; therefore, care is needed in interpretation of data. For example, lidar and radar, both of which can be used to remotely sense wind, can "see" farther in clear conditions; however, if the air is exceptionally clean,

1 Rain has a much higher radar reflectivity than snow, except melting snow produces more reflection. Large hail produces even larger returns than rain or snow.

these instruments will not be able to sense the wind conditions. For scanning instruments, the volume being sensed increases with distance from the radar and the average resolution decreases, because the scan produces ever larger concentric circles. Similarly, visible satellite imagers can detect the tops of clouds, but the same clouds prevent the imager from seeing clouds at other levels.

## The Impact of the Era of Satellite Remote Sensing on Weather Observations and Modeling

The year 1978 is generally considered the beginning of the satellite era for weather prediction purposes. Continuous monitoring by weather satellites began in 1974, and the first polar-orbiting environmental satellite (POES) was launched in 1978. The POES program greatly improved the data available for assimilation, as polar-orbiting satellites orbit at a much lower altitude (about 850 km above the surface, versus 35,780 km for the Geostationary Operational Environmental Satellite (GOES)), allowing much higher-resolution sampling. These satellites also use active sounding sensors that in many cases can penetrate clouds and provide more information about the environment, including ocean temperature and surface winds on the ocean, and can estimate temperature and humidity profiles. Subsequent satellites have been equipped with increasingly sophisticated and high-resolution instrumentation, leading to a dramatic increase in the quality of atmospheric analyses as observations from large volumes of the atmosphere became available.

## Numerical Weather Prediction

All weather inputs for operational load, wind, and solar forecasts in the electricity sector are based on foundational data coming from government-operated NWP programs. This is because the process of collecting and assimilating data is costly and requires cooperation across nations, and the models themselves require vast quantities of computer resources. In some cases, additional NWP tasks are performed by users or providers in the energy sectors in the process of producing sector-specific products, but for the most part, at this time, the NWP output of the major national centers—the European Center for Medium-Range Weather Forecasting, the UK Meteorological Office, the U.S. National Oceanographic and Atmospheric Administration's National Centers for Environmental Prediction, and the Canadian Meteorological Center—is difficult to improve upon in a timely and cost-effective manner. Most providers in the energy sector focus on statistical post-processing of the raw NWP data, usually using machine learning techniques.

### Reanalysis Output Refactoring

Raw model output from the reanalysis process is archived, but the data provided to users are usually refactored into datasets that provide a standard set of atmospheric variables on a regular grid that is typically mapped to a sphere with multiple vertical levels. For spectral models, the raw model archive consists of spectral coefficients or gridded data on a reduced Gaussian grid,[2] so it is usually interpolated to a fixed latitude and longitude spacing when provided to end users. This means that grid spacing in the north-south direction is constant but west-east spacing varies with latitude. For example, a 0.25° latitude x 0.25° longitude grid has north-south spacing of 27.8 km everywhere,[3] while the west-east spacing is 27.8*cos(latitude), which is 27.8 km at the equator, 24.1 km at 30 degrees, 19.7 km at 45 degrees, and 13.9 km at 60 degrees. It is important to note that in this case the apparent increased horizontal west-east resolution at higher latitudes is an artifact of this interpolation and is not an indication of increased resolution at high latitudes.

Reanalysis data are usually provided to end users on familiar vertical coordinates like height or pressure levels. For example, ERA5 (Fifth-Generation ECMWF Atmospheric Re-Analysis of the Global Climate) data are provided at 25 hPa intervals starting from 1000 hPa. However, the native model output represented on a terrain-following vertical coordinate has far better vertical resolution near the surface. This can be useful

2 A discussion of spatial referencing, reduced Gaussian grids, and spectral coefficients can be found at https://confluence.ecmwf.int/display/CKB/ERA5%3A+What+is+the+spatial+reference.

3 The polar circumference of Earth is 40,008.8 km, or 111.13 km per degree.

for wind energy purposes as it provides wind speed estimates at several levels across the rotor diameter, for those willing to deal with transforming from the native format.

### Deep Convolutional Neural Networks for Downscaling NWP Output

Recent advances in machine learning techniques for computer vision have inspired a new class of methods for the post-hoc downscaling of NWP outputs. These methods promise to reduce the burdensome computational requirements of high-resolution NWP simulations while maintaining high-quality data outputs. If these methodologies can be proven to work well, they will enable the production of higher resolution and longer time series of weather input data suitable for power system modeling applications, as well as ensembles of these datasets that capture the uncertainty of the weather inputs and therefore allow electricity system studies to model sensitivity to this uncertainty.

Deep convolutional neural networks (CNNs) have been recently shown to excel at a wide range of computer vision tasks, including generative models. Convolutional kernels are designed to match the dimensionality and structure of image, video, and NWP simulation data, and the convolutions are repeatedly layered to extract and process data features at both large and small scales. The result is a powerful nonlinear parametric model that can learn physical phenomena such as the momentum balance for wind flows on a gridded hypercube in much the same way that finite-difference or finite-volume methods operate in NWP models.

In practice, a major problem is that a naïve convolutional network can exhibit regression to the mean in the form of blurring when producing forecasts or enhancing the resolution of data. Statistically, this may be a reliable output for the convolutional network that will minimize its objective function, but it greatly reduces the practical value of the data. One solution to this problem is adversarial training with generative adversarial networks (GANs), where a generative model must produce data that are not only conditionally accurate but also sufficiently realistic to fool a discriminative network. For downscaling data with GANs (often called super-resolving), the generative network is trained to produce an enhancement of the low-resolution input data that the discriminator believes is similar to real data, while simultaneously minimizing the numerical deviation from a corresponding true high-resolution dataset. This method has been shown to be effective in creating highly realistic enhancements for many types of data.

GANs with deep convolutional networks have only recently been applied to the task of downscaling NWP data but have already shown considerable promise with high-quality physics-based validation of the outputs (Stengel et al., 2020). To the knowledge of this project team, only a single public dataset has been published at the time of this writing that leverages GANs to downscale climate data, in this case a precipitation dataset from CMIP6 (Hess et al., 2022).[4] However, several wind datasets are known to be in development that leverage GANs to do a final spatio-temporal enhancement on coarse NWP data instead of running the NWP down to the final desired resolution. The benefit of this hybrid NWP+GAN approach is a significant reduction in computational costs compared to what would be required by a full high-resolution NWP simulation (estimated at more than two orders of magnitude in compute time savings).

The main drawbacks of using GANs for downscaling are that this requires significant investment in machine learning expertise, machine learning–specific computing infrastructure, and high-quality training data, and can result in a loss of methodological interpretability including the possibility for data outputs that do not respect physical constraints. This last problem is clearly the most concerning, as low-quality data with poor physical constraints could compromise power system planners' ability to accurately predict and plan for future system needs. The methods described above have the potential to greatly benefit the renewable energy and meteorological communities, but rigorous validation needs to be of the upmost priority. Statistical benchmarking, validation against ground-truth observations,

4 See the World Climate Research Programme's Coupled Model Intercomparison Project at https://www.wcrp-climate.org/wgcm-cmip.

and careful examination of physical data characteristics like turbulence should all be regular practice when implementing these methods.

## Data Produced for Solar Generation Calculations

Satellites contain instruments that measure the reflection of solar radiation within the atmosphere and the emission of infrared radiation by it. Several factors including the presence of water vapor, clouds, and the temperature profile all impact these measurements, and through complex model algorithms these measurements can be used to make inferences about the properties of the atmosphere that result in the measurements and/or about irradiance at the surface. The methods can be used specifically to produce irradiance measures, as is the case for the National Solar Radiation Database (NSRDB), or in conjunction with NWP models where the assimilation process uses the measurements to improve the initial condition and then the model determines the evolution of the surface irradiance properties.

Irradiance data produced by NWP models are subject to many of the same caveats regarding model resolution that have been highlighted for wind data. In addition, radiation calculations are computationally expensive because they model all the reflection, absorption, emission, and scatter of both longwave and shortwave radiation throughout the atmosphere and by the ground. Because of this computational intensity, they are usually performed at longer time step intervals than other model calculations. For example, the calculation may be performed as infrequently as every 30 to 60 minutes, although every 5 to 15 minutes is more common. This is important, because short-interval irradiance data in some NWP datasets may be static or interpolated between radiation calculation periods even if other fields are updated more frequently. Also, most NWP models only need to calculate global horizontal irradiance (GHI) as part of the modeling radiative processes and may not calculate direct normal irradiance (DNI) and/or diffuse horizontal irradiance (DHI). However, many modern models (for example, WRF-Solar) have options that allow GHI at the ground to be calculated as frequently as the regular model time step. They also have options that allow direct irradiance at the surface to be calculated. From this, DNI can be calculated, and together with GHI, DHI can be deduced.

The NSRDB has 4 km grid spacing, which is reasonably good, but a finer grid is better, especially when dealing with smaller clouds. The National Center for Atmospheric Research has developed the MAD-WRF model (Multi-sensor Advection Diffusion Weather Research and Forecasting) for intra-day forecasting applications, which uses satellite observations (and surface-based ceilometer observations, where available) to correct the cloud and other model fields at initialization (Jiménez et al., 2022).[5] It inserts clouds where the model has none (and estimates the level(s) at which to add cloud and modify other model fields accordingly) and eliminates clouds where the satellite shows that none exist. Application of newer techniques like this will further improve irradiance data in future datasets.

5 See https://ral.ucar.edu/solutions/products/mad-wrf.

# Appendix References

Hess, P., M. Drüke, S. Petri, F. M. Strnad, and N. Boers. 2022. "Physically Constrained Generative Adversarial Networks for Improving Precipitation Fields from Earth System Models." *Nature Machine Intelligence* 4: 828–839. https://doi.org/10.1038/s42256-022-00540-1.

Jiménez, P. A., J. Dudhia, G. Thompson, J. A. Lee, and T. Brummet. 2022. "Improving the Cloud Initialization in WRF-Solar with Enhanced Short-Range Forecasting Functionality: The MAD-WRF Model." *Solar Energy* 239: 221–233. https://doi.org/10.1016/j.solener.2022.04.055.

Stengel, K., A. Glaws, D. Hettinger, and R. N. King. 2020. "Adversarial Super-Resolution of Climatological Wind and Solar Data." *Proceedings of the National Academy of Sciences* 117(20): 16805–16815. https://doi.org/10.1073/pnas.1918964117.

## PHOTOS

Cover:
(Top left) © iStockphoto/sharply-done
(Top middle) © iStockphoto/marcrossmann
(Top right) © Thinkstock/MonaMakela
(Lower left) © iStockphoto/seamartini
(Lower middle) © iStockphoto/DustyPixel
(Lower right) © iStockphoto/shaunl

p. xi: © iStockphoto/Shaiith

p. xiii: © iStockphoto/zhongguo

p. 1: © iStockphoto/lovelyday12

p. 3: © iStockphoto/janiecbros

p. 6: © iStockphoto/WichienTep

p. 9: © iStockphoto/Michael Edwards

p. 10: © iStockphoto/Thossaphol

p. 12: © iStockphoto/gorodenkoff

p. 14: © iStockphoto/Kinwun

p. 17: © iStockphoto/kruwt

p. 20: © iStockphoto/blackdovfx

p. 23: © iStockphoto/AndreyPopov

p. 24: Figure 2
(#1 from upper left clockwise)
© TebNad; © SergeyKlopotov; © DSCimage; © 3DSculptor
(#2) © SweetBunFactory
(#3) © National Oceanic and Atmospheric Administration
(#4) © Creative Commons/Rautenhaus et al.
(#5) © National Oceanic and Atmospheric Administration

p. 25: © iStockphoto/monsitj

p. 28: © iStockphoto/FrankRamspott

p. 35: © MakcnmNbackiok

p. 37: © iStockphoto/Stramyk

p. 40: © iStockphoto/gorodenkoff

p. 41: © iStockphoto/monsitj

p. 44: © iStockphoto/2ndLookGraphics

p. 49: © iStockphoto/pidjoe

p. 51: © iStockphoto/Adam Calaitzis

p. 53: © iStockphoto/Petrovich9

p. 55: © iStockphoto/koto_feja

p. 57: © iStockphoto/Jurkos

p. 59: © iStockphoto/TrongNguyen

p. 66: © iStockphoto/Bilanol

p. 68: © iStockphoto/Marc Bruxelle

p. 70: © iStockphoto/aapsky

p. 73: © iStockphoto/Yunpeng Li

p. 77: © iStockphoto/ysbrandcosijn

p. 81: © iStockphoto/daverhead

p. 84: © iStockphoto/yangphoto

p. 85: © iStockphoto/DamienGeso

p. 86: © iStockphoto/Laurence Dutton

p. 88: © iStockphoto/Marcus Millo

p. 92: © iStockphoto/Peter Blottman Photography

p. 94: © iStockphoto/Perytskyy

p. 96: © iStockphoto/MikeMareen

p. 97: © iStockphoto/skynesher

p. 98: © iStockphoto/da-kuk

p. 100: © iStockphoto/filo

p. 102: © iStockphoto/Elen11

p. 104: © iStockphoto/su tim

p. 105: © iStockphoto/Songpon Ruengsamut

p. 123: © iStockphoto/FrankRamspott

p. 125: © iStockphoto/ph2212

# Weather Dataset Needs for Planning and Analyzing Modern Power Systems

**A Report of the Energy Systems Integration Group's Weather Datasets Project Team**

---

The report is available at https://www.esig.energy/weather-data-for-power-system-planning.

To learn more about ESIG's work on this topic, please send an email to info@esig.energy.

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation, particularly with respect to clean energy. More information is available at https://www.esig.energy.

ESIG
ENERGY SYSTEMS INTEGRATION GROUP