Transportation of Large Wind Components: A Review of Existing Geospatial Data

Meghan Mooney and Galen Maclaurin
National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-6A20-67014
September 2016

Contract No. DE-AC36-08GO28308

Transportation of Large Wind Components: A Review of Existing Geospatial Data

Meghan Mooney and Galen Maclaurin
National Renewable Energy Laboratory

Prepared under Task No. EPSA.Z330

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-6A20-67014
September 2016

Contract No. DE-AC36-08GO28308

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

This report is available at no cost from the National Renewable Energy
Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect http:/www.osti.gov/scitech

Available for a processing fee to U.S. Department of Energy
and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI http://www.osti.gov
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS http://www.ntis.gov
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

Acknowledgments

The U.S. Department of Energy's Office of Energy Policy and Systems Analysis funded this project. We thank David Schoeberlein and Zachary Clement for their careful review and Diana Bauer, Christopher Mone, Jason Cotrell, Donna Heimiller, and Mark Jacobson for input and assistance on this project. Opinions represented in this report are the authors' own and do not reflect the view of the U.S. Department of Energy or the U.S. government.

List of Acronyms

AWEA	American Wind Energy Association
FAF	Freight Analysis Framework
GIS	Geographic Information System(s)
HSIP	Homeland Security Infrastructure Program
LCV	long combination vehicles
MW	megawatt
NBI	National Bridge Inventory
NREL	National Renewable Energy Laboratory
NTAD	National Transportation Atlas Database
ORNL	Oakridge National Laboratory
OSOW	oversized and overweight

Executive Summary

This report features the geospatial data component of a larger project evaluating logistical and infrastructure requirements for transporting oversized and overweight (OSOW) wind components. The goal of the larger project was to assess the status and opportunities for improving the infrastructure and regulatory practices necessary to transport wind turbine towers, blades, and nacelles from current and potential manufacturing facilities to end-use markets. The tandem report in this project, titled *Transportation of Large Wind Components: A Permitting and Regulatory Review*, focuses on the regulatory aspects of OSOW transportation, whereas the purpose of this report is to summarize existing geospatial data on wind component transportation infrastructure and to provide a data gap analysis, identifying areas for further analysis and data collection.

A literature review highlighted the unique challenges of transporting large wind components due to physical infrastructure limitations caused by the size of wind turbines and the multi-modal nature of its transport. In particular, weight for nacelles, turning radius for blades and weight, turning radius, and vertical and horizontal clearance for tower sections are the major infrastructure-related barriers impeding component transport. Additionally, transportation of wind energy components is often multimodal, making use of road, rail, and waterway options, and each mode entails different infrastructure-specific challenges. Physical limitations of infrastructure are often further complicated by regulatory challenges, which vary significantly between local and/or state jurisdictions.

The collection and assessment of existing geospatial infrastructure data sets raised concerns and challenges for data integration and subsequent routing analyses. Issues identified included differences in the scale at which data were created, topological errors (i.e., logical inconsistencies in the data), differences in naming conventions, and inconsistent attribute or spatial coverage. In addition to these issues, data sets specific to wind energy deployment could not be identified—such as locations of specialized crane services contracted to install turbines. Data describing certain overhead obstructions were also unavailable for most of the country, such as clearances for streetlights, electrical and telecommunication wires, and encroaching trees. We conducted three routing scenarios to highlight these limitations of existing geospatial data for routing OSOW wind component transport. Our findings indicate that overall the existing state of transportation infrastructure data provides the necessary base layers (e.g., road and rail networks and bridge inventories); however, the level of detail, attribute and spatial completeness, and logical consistency of the data impose significant challenges for routing OSOW wind components from the manufacturer to the remote wind site.

Table of Contents

List of Figures

List of Tables

1 Introduction

Assessments of wind energy site suitability rely primarily on resource potential and distance to transmission infrastructure, while the challenges of transporting wind energy components from the manufacturer to a potential site receives less attention (Cotrell et al. 2014a; Cotrell et al. 2014b). The dimensions and weight of wind components often exceed the limits of U.S. infrastructure, making them difficult to transport. Because of the difficulties associated with transportation logistics, sites that are extremely remote or that are on complex terrain are more costly sites to develop. As wind components continue to get larger, the challenges of transporting these components are likely to intensify.

Currently, transportation makes up about 3%–8% of the total land-based wind capital costs in the United States, and with projected increases in turbine sizes, these percentages are expected to increase significantly (Cotrell et al. 2014a; Cotrell et al. 2014b; Zayas et al. 2015). Despite this projected increase, little is known about the state of U.S. transportation infrastructure for transporting these large wind components. Yet, understanding transportation infrastructures' influence on wind development is important to eliminate hurdles that are preventing wind development. As part of this effort to better understand infrastructure requirements for wind component transportation across the United States, this project aims to (1) review the literature on transportation and logistical barriers for wind energy, (2) collect data based on these known infrastructure requirements, and (3) perform a data gap analysis on the existing data in order to identify existing problems, missing data, and logical consistencies. This report is organized around these three objectives. This report also showcases potential application of the collected data and some of the data-specific problems we found in three different wind component routing scenarios in the United States.

Coupled with the physical issues of transporting large wind components are the hurdles involved with the required permitting process. For most state and local regulatory authorities, permits are required for truck transport of oversized and overweight (OSOW) wind components through a given jurisdiction's boundaries. Often each regulatory jurisdiction requires unique transporting requirements and restrictions as part of the permitting process, and such permits must be negotiated individually with each regulatory unit (Cotrell et al. 2014a; Cotrell et al. 2014b). This report does not focus on the regulatory aspects of OSOW wind transport, which are addressed in concurrent research published in a separate technical report, titled *Transportation of Large Wind Components: A Permitting and Regulatory Review* (Levine and Cook, forthcoming). Although this report does not specifically detail these permitting hurdles, it is still very much connected to the infrastructure requirements for wind component transportation, and therefore, the state and local permitting requirements are incorporated in our investigation of data needs.

2 Literature Review

This literature review is organized into two main categories: (1) general summary of the literature, and (2) transportation breakpoints. This review was not meant to look broadly at infrastructure and logistics for OSOW transport. Instead, it was focused on the scope of our subproject goals of collecting spatial data related to U.S. wind component transportation.

2.1 General Summary of the Literature

Our review found a general lack of literature regarding wind component transportation and a limited number of peer-reviewed sources. Consequently, most of the substantive sources found in our review are from non-peer-reviewed sources—mostly white papers, industry-focused news articles, and presentation slides. This review attempts to unpack some of the emerging themes and gaps within the literature.

2.1.1 Emerging Themes

The majority of existing literature takes a conceptual approach, relying heavily on qualitative information and conversations with industry stakeholders (Cotrell et al. 2014a; Cotrell et al. 2014b; Roy 2016). These conversations reveal a rich knowledge base from industry experiences with transportation logistics; however, quantitative data appear to be mostly proprietary. Discussed in this sub-section are a few of the notable themes that reemerge throughout the literature—namely, the challenges associated with turbine size and the multi-modal nature of transporting wind components. Both of these themes are particularly important when collecting and assessing spatial dimensions of wind component transportation infrastructure in the United States.

The most common theme across the literature is related to the challenges associated with the size and weight of wind components. Simply put, wind turbines are large, heavy, and extremely difficult to transport. The dimensions and weight of wind components place limits on the types of feasible routes, due to the larger turning radius, tall clearance requirements, and road weight restrictions of the OSOW loads. Indeed, due to the limits placed on the size and weight of turbine components, those that are too heavy for a particular mode or too big to fit under older bridges must traverse an alternate, and potentially more costly, route. At the same time, certain states, counties, and municipalities require different permitting for OSOW transport that may be more expensive, limited to certain hours of the day, or may completely restrict the passage of OSOW transport all together. In traditional route planning, the fastest, most cost-effective, and direct route is generally chosen. However, in OSOW freight transportation, the “best route” is dependent on not only time and distance, but it is also adjusted for limitations and barriers, including both physical and regulatory barriers, that limit mobility. To make matters more complex, these physical and regulatory barriers vary significantly across mode and geographic space. Specific limitations and barriers addressed throughout the literature are summarized in Table 1.

Table 1. Limitations and Barriers to Transporting Wind Components

Limitation	*Description*
Lack of uniformity for permits	There is a high degree of variability in permitting cost and specific restrictions across states and within states (e.g., counties and municipalities). Trucking wind components is often cheapest for long hauls; however, this mode requires permitting for OSOW cargo. These permits are often negotiated and purchased for each specific "pass-through" jurisdiction. Permits vary based on processing time, cost, expected bonding, and specific restrictions including allowed sizes, weights, trailer overhangs, time of day, the number of required OSOW escorts and signage. Part of the permitting process involves road bonding, whereby jurisdictions require entire routes, or even just specific subsections, to be bonded and assessed for current damage and verification that transport will not cause any further damage.
Lack of uniformity for the physical makeup of wind components	Wind component dimensions and weight lack uniformity. This high degree of variability makes it difficult to assign size and weight breakpoints because these breakpoints will vary based on the turbine design. In addition, turbines are getting larger and the variability that already exists will likely only intensify as new technologies, such as longer segmented blades, begin to proliferate in the market. With difficulty assigning breakpoints, making transport decisions becomes highly conditional and difficult to model.
Size and weight of turbines	Turbines are large and heavy, and therefore difficult to transport by nature. Size barriers exist across all modes, although they are more prolific for trucking. Barriers are related to curvature and turning radiuses, road-weight limits, street overhangs, vertical and horizontal clearances of tunnels and bridges, and even permitting barriers through certain jurisdictions.
Burden of proof shifted to the wind industry	The permitting process is a negotiated process, and during these negotiations, the industry must be able to justify why traveling through a particular jurisdiction is safe and economically beneficial. The industry must be able to justify how transporting OSOW components is safe for traffic and road integrity, and at the same time, they must be able to prove that transportation will stimulate jobs and the economy, whether through direct or indirect stimulus.
High variability in mode cost	The cost of transport on a given mode varies tremendously within and across modes depending on the size/weight configurations of the component being transported, the company used, the distance traveled, and the relative location of manufacturing and project sites. This high variability in costs introduces another layer of complexity and nuance to modeling wind transportation networks and determining the best path.

Sources: Cotrell et al. 2014a; Cotrell et al. 2014b

The issue of size for turbine component transportation is not likely to subside in the near future. In fact, wind turbines have grown significantly since the 1980s and continue to rise today (AWEA, n.d.; Cotrell et al. 2014a; Cotrell 2014b; Zayas et al. 2015). As new technologies develop for generating wind energy in areas with low resource potential, wind turbines are expected to almost double in height in the near future, from the common 80 m to as large as 150 m and 160 m at hub height (U.S. FAA restrictions permitting) (AWEA, n.d.; Cotrell et al. 2014a; Cotrell 2014b). The expected increase in size will ultimately lead to higher transportation costs (Zayas et al. 2015), and the estimated proportion of total land-based wind capital costs for wind component transportation (about 3%–8%) is expected to increase significantly (Cotrell et al. 2014a; Cotrell 2014b). This cost increase is because increased rotor diameter and hub heights drive higher expenses related to transportation and installation, requiring specialized transportation trailers and support vehicles that can only be transported on certain U.S. highways, as well as an increase in special permitting, rare special-purpose cranes needed for turbine assembly, and the number of loads needed for turbine delivery (Zayas et al. 2015). A number of solutions have been proposed to offset the increased costs associated with larger turbine component transportation, including investments in advanced trailer technologies, switching to temporary onsite manufacturing to avoid transportation hurdles, or encouraging more collaboration between permitting authorities to make uniform permits and wind corridors (see Appendix A for a detailed list of proposed solutions).

Another reoccurring theme throughout the literature is the multi-modal nature of turbine component transportation. Currently, wind components are transported using a variety of different modes—including barge/ship, rail, and truck—in order to transport components from the manufacturing site to the remote wind site (AWEA, n.d.; Del Franco 2015; Cotrell 2014a; Cotrell 2014b; Zayas et al. 2015). For instance, one 150-MW project can require as many as 689 truckloads, 140 railcars, and 8 ships to complete the transportation process (AWEA, n.d.). Yet, given the multi-modal nature, there is no universal formula used to determine which modes are used for which type of scenario. Rather, the choice of any given mode is highly conditional, depending largely on the variable costs associated with the manufacturer used, permitting required, distance traveled, and the component(s) being transported. Table 2 lists some of the main variables that influence transportation costs, and inevitably, route and mode choice.

Table 2. Factors Influencing Route Choice and Transportation Costs

Mode	***Influencing Factors***
All modes	• Component being transported—size and weight configurations • Fuel costs • Distance traveled • Proximity of manufacturing facilities and wind sites • Labor costs • Storage facilities (*if applicable*) • Number of needed mode changes • Transportation company used
Rail	• Tracking rights and rail ownership • Rail-line partnerships or agreements • Age and dimensions of railways, tunnels, and bridges • Track radius/ track curvature
Truck	• Permitting uniformity • Regional permitting associations, multi-state or multi-jurisdiction permits • Number of required permits—number of pass-through regulatory jurisdictions • Requirements for each permit (e.g., overhang limitations, daylight hour restrictions, road restrictions, bonding requirements, escort requirements) • Available information and permitting process • Road type • Road and bridge clearances and weight limits • Number of exits/ turns and road curvature • Weather, time of day, and season of travel • Number and roadway condition of pass-through places (e.g., urban areas, cities, towns)
Ship/barge	• Port fees • Port clearances • Weather and season of travel • Channel depth

Regardless of the mode choice, trucking will likely be used at the very least to connect other modes to the remote project site. This is important because trucking is subject to special permitting avoided by other modes and it often requires more logistical planning. In addition, trucking takes longer and can require up to eight oversized loads for a single turbine—one for the nacelle, three for the blades, and four for the tower sections (Coleman 2009; Anderson 2012). The number of loads required to deliver a turbine by truck is perhaps unavoidable; however, the permitting restrictions can be reduced to more manageable feats.

2.2 Summary of Transportation Breakpoints

Transportation infrastructure in the United States places limits and barriers (e.g., breakpoints) on the size, shape, and weight of OSOW loads, which specifically impacts transport of blades, the nacelle, and tower sections. Listed in Table 3 are the size, shape, and weight breakpoints for transporting each of these three OSOW wind components. It is important to note that—through conversations with industry experts and a detailed examination of the literature—to date, wind components have been designed specifically to meet these transportation breakpoints. These transportation breakpoints are particularly important in identifying which barriers exist for different modes and along which routes in order to ensure transport from the manufacturing facility to the remote site.

Table 3. Transportation Breakpoints[a]

	Component Breakpoint	Hub Height/ MW Affected	Notes
Tower Components	Width: 4.3 or 4.6 m in diameter [A,C] Length: 52 or 63m [D] Weight: ~ < 80,000 lbs. (truck) [B]	Width: 80-m–160-m turbines and any turbine larger than ~1.9 MW [C]	[A] Rolled steel can be used to overcome this breakpoint, though it greatly increases capital costs. [B] Weight is likely not going to play a large role in transportation of a tower component in the future due to rolled steel and tower segmentation. [C] Most turbines today are influenced by this transportation breakpoint. [D] The length breakpoint accommodates for turning radiuses. Segmentation of tower sections is how this breakpoint is achieved.
Blades	Length: 52 m–63 m [A, C] Width: (aka blade root) 4.3 m–4.6 m [A] Weight: ~ < 80,000 lbs. (truck) [D]	Length: Potentially affects 2.2-MW–3.8-MW turbines [C] Width: 4.3–7.3 MW. [B]	[A] New technologies in blade design could allow for longer blades despite the transport breakpoint. [B] This is based on future turbines with longer blades. The length breakpoint potentially affects future turbine installations with lengths up to 80 m and 4.3 MW–7.3 MW. [C] This potentially affects 2.2-MW–3.8-MW turbines; however, there appears to be no absolute limit on length, yet longer lengths increase turning radius and this seems to be the accepted breakpoint in order to accommodate turning radius (~120 ft). At the same time, longer lengths mean longer trailer overhangs, and most states limit these overhangs. [D] The weight breakpoint does not appear to affect blade transportation because blades are the lightest component.
Nacelle	Length: 11.7 m [A] Height: 4.3 m–4.6 m [B] Weight: ~< 80,000 lbs. (truck); ~< 102 tonnes, ~225,000 lbs (rail) [C,D]	Weight: 3–5 MW [E]	[A] Limited to the 19-axle trailer load deck. [B] Designed to fit under overpasses and traffic controls. [C] 225,000 lbs. is approximately the weight limit of rail transport (~260,000). [D] Trucks and nacelles (total load) need to be less than 80,000 lbs. in order to comply with U.S. interstate restrictions. In some states, special permits can be purchased to increase the weight by a fixed amount. In order to keep nacelles under this weight, the nacelle components are separated as much as possible. [E] Unless future technology can significantly reduce nacelle weight, any future turbines of 3 MW–5 MW will be affected by this breakpoint, and nacelles will be too heavy to transport on the road even with internal components shipped separately. This will lead to a stronger reliance on rail for the majority of the transportation, and any trucking to the site will need special and costly permits.

[a] Transportation breakpoints are based on size and shape thresholds commonly agreed upon by industry members. They represent the agreed maximum size of a given wind component in order to be able to transport it on a

particular mode given a particular barrier. They are based on both physical barriers in mode capacity as well as common permitting requirements.

Sources: AWEA, n.d.; Cotrell et al. 2014a; Cotrell et al. 2014b; Zayas et al. 2015

2.2.1 Key Breakpoint Considerations for Geospatial Data

Important considerations for geospatial data include:

- Nacelles are the most difficult component to transport due to weight. Because rail is capable of transporting heavier components, rail is ideal for transporting nacelles.
- With height restrictions, transporting tower segments on bridges and through tunnels is the principal concern.
- The dimensions of blades and the turning radius involved with specific turns and road curvature are the main concerns for the length breakpoint.
- These breakpoints are rough estimates. Most are estimated based on conversations with industry members. A lack of hard data makes it difficult to determine more accurate breakpoints.

2.3 Existing Gaps Within the Literature

A few notable gaps exist within the literature. First, the specific size breakpoints (as detailed in Section 2.2) for transporting turbine components are wide-ranged estimates stemming from interviews with industry stakeholders, and therefore, validation of breakpoints have not been assessed. This is a significant gap because it affects the limits placed on the design of turbines, potentially leading to deployment of smaller turbines than many transportation systems might otherwise allow, and at the same time, it will lead to extremely conservative transport modeling and influence route choice decisions. Second, there is little discussion or analysis on the costs and benefits of implementing novel transportation solutions. For instance, the use of specialized trailers and segmented blade design (both of which are implemented in Europe) could work within the confines of U.S. transportation restrictions while allowing larger blade deployment. Alternatively, the implementation of “pop-up” on-site manufacturing facilities will greatly alter the role that transportation plays in wind site development by potentially alleviating some challenges altogether. Third, there is a significant lack of data regarding variability in permitting, as well as transportation costs and route choice decisions as they relate to the larger supply chain. From a data science perspective, this is particularly significant because without understanding the dynamics between the supply chain and mode and route choices—and the role in which permitting complicates the decision-making process—it is increasingly more difficult to model such characteristics at the national level. Perhaps such variation is too nuanced for typology creation, but such conclusions are hard to make without examining many scenarios across both time and space.

3 Existing Data

The literature review (Section 2) identified two major themes important for the project goals of collecting geospatial data of wind component transportation in the United States. First, the size of wind components leads to a number of transportation barriers (see Table 1), and second, turbines are transported using a variety of modes. In addition, the literature review summarized the key breakpoints associated with transporting each wind component for each type of transportation mode. These findings informed our data collection to best reflect the current landscape of U.S. wind transportation infrastructure requirements. In this section, we detail the landscape of existing data related to turbine transportation—including the specific data that would be required to create a national database and the actual data that we collected.

3.1 Assessment of Data Needs

We have identified two main data requirements—(1) data must be comprehensive, in terms of geographic coverage and attribute detail, and (2) data must be publically available. To meet the first requirement, data sets must cover all potential geographic regions, modes of transport, supply chain scenarios, and barriers. Second, in order to comply with the U.S. government and U.S. Department of Energy standards of data transparency and accessibility, we have placed a higher precedent on obtaining publically available data over data from proprietary sources. To our knowledge, no single database currently exists that meets these two requirements.[1] Because no single data set exists that meets the project needs, we compiled data sets from a variety of sources amassing all necessary data sets for full coverage of desired layers, attributes, and geographic regions.

In order to compile a comprehensive database, we collected a large number of disparate data products. Figure 1 illustrates our data vision for a comprehensive database on wind component transportation in the United States, which describes our data needs from the organizational standpoint of a transportation network. This is a common approach taken in the field of transportation geography, in which transportation systems are most often understood as a network made up of links, nodes, and barriers, whereby two sites are connected if they share a link (or series of links) and do not have any barriers along the connecting links that limit the transportation flow. As such, while compiling data, we categorized our data needs along these three network components—transportation routes (or modes), critical sites (e.g., origins, destinations, and connector sites), and barriers and limitations associated with particular modes' ability to transport wind components. As illustrated, the modes included in our needs assessment include truck, rail, and waterway (ship/barge). The network sites include origins (e.g., manufacturing facilities), destinations (e.g., existing wind site locations or potential areas for development), and transition nodes (e.g., rail yards, storage facilities, and ports). Finally, the barriers we incorporated include both physical and regulatory barriers that may impede turbine mobility along a particular mode. Examples of physical barriers include vertical clearances along tunnels and bridges, turning radiuses involved in road exits or curvature, and road or bridge weight limits. Examples of regulatory barriers include truck pass-through jurisdictions that may have their own laws and regulations regarding OSOW permitting and clusters of traffic accidents

---

[1] *Proprietary datasets, such as TomTom, ProMiles, and NavTeQ exist; however, these data sets are often designed for general navigation, not OSOW transportation, and they are often burdened with the same spatial and attribute errors as publically available data sets.*

that indicate potential areas of high traffic congestion or dangerous intersections for transporting OSOW vehicles. Some of the data sets listed in Figure 1 are data layers (or objects) that have their own spatial geometries, while others are just attribute data (i.e., tables) that are associated with particular layers. In addition, some layers are obtained through spatial querying and network calculations. Furthermore, some of the data sets listed in Figure 1 are currently missing from our data inventory, and the reasons for their absence are discussed in Section 4.

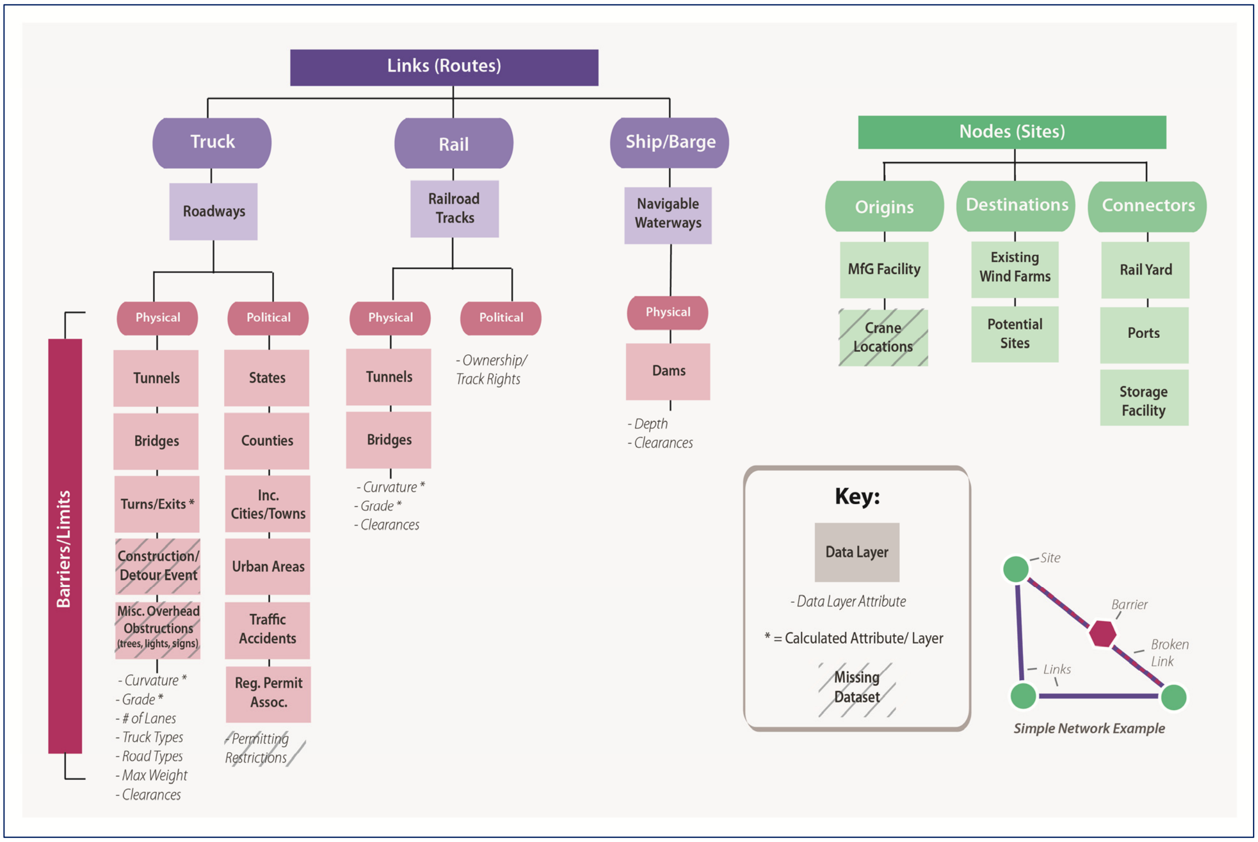


Figure 1. Network perspective of data needs

3.2 Data Acquisition

As previously mentioned, data sets were collected from a variety of sources to meet our data requirements. These sources include the U.S. census, the National Transportation Atlas Database (NTAD), and, when public sources could not be identified, data were obtained from proprietary sources like Homeland Security Infrastructure Program (HSIP), among others. Because these data sets came from different sources, they were all collected at a variety of different spatial scales and with different standards for precision and accuracy. Tables 5–7 in Appendix B list the data sets we collected by their data type, source, scale, and other explanatory fields such as important attributes they contain, attributes that should be used for relational joining, and a description of the data set. In total, 23 individual data layers were collected that make up 34 out of 39 of the desired non-calculated data sets or attributes listed in Figure 1.

4 Data Gap

In addition to collecting data and assessing different routing scenarios, this project performed a data gap analysis on the state of the existing data. Although we were able to collect the majority of the necessary data, problems of completeness and consistency still existed within these data sets. This section aims to provide a high-level overview of some of the main data gaps we found. In particular, this section outlines the three main data gaps this research has found—namely, missing data sets, data discrepancies in the existing data collected, and tool-based data gaps. The overarching purpose of this data gap analysis is to shed some light on the areas needed for further analysis and data collection. Specific gaps we found while performing scenario-based routing analyses are detailed in the following section (Section 5), and more detailed information on data discrepancies is provided in Appendix C.

4.1 Missing Data Sets

Out of 39 desired data sets or attributes listed in Figure 1, 5 are currently missing. There are two general reasons why we have missing data sets. First, some of these missing data sets—such as crane locations and multi-level permitting restrictions—require timely data creation. The second reason we have missing data sets is because some data sets—such as construction and detour events and overhead obstructions like streetlights, signs, or encroaching trees—are simply difficult to represent in a national geographic information system (GIS) database. Indeed, these fine-resolution data sets may never be fully available at the national scale due to the high degree of local variability and the dynamic nature of these transportation barriers. For instance, potential barriers such as encroaching trees or detour events are so specific to a particular location and they are constantly changing, so they are difficult to capture in a national spatial data set. The fact that some transportation barriers will never fully be satisfied in a GIS database underscores the key reason why many regulatory officials stress the importance of getting boots-on-the-ground to confirm that a particular route is safe for travel. Nonetheless, using GIS to understand wind component infrastructure capacity does provide a higher-level assessment of the barriers and limitations along particular modes and can help companies and local officials sift through potential routing options.

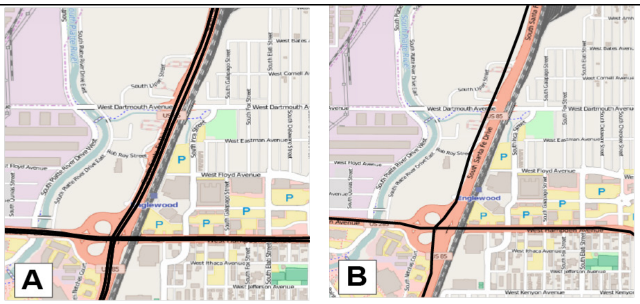
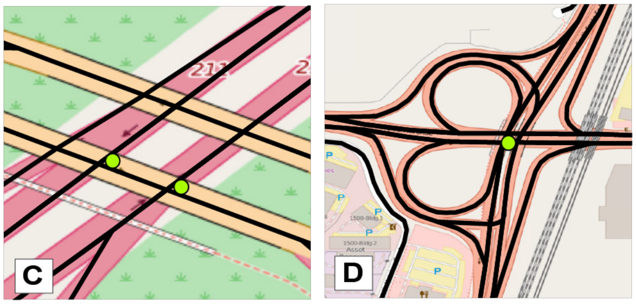
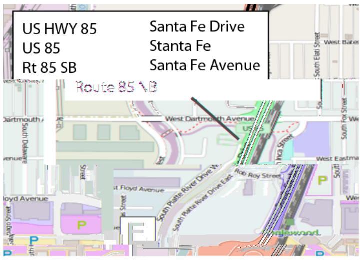
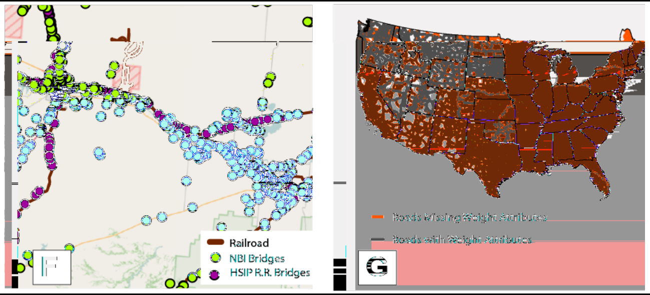
Other data sets that are missing are not actually listed in Figure 1. These data sets could potentially help make future analysis more robust, and they include data related to wind component supply chains, heavily utilized corridors for transportation, and cost information associated with using particular modes. The reason these data sets were not included in our initial vision of data requirements (Figure 1) is because we see them as the next-level type data that we would need to perform analyses that are more robust. In other words, these data sets must be informed in part by the phase-one-type analysis of the data we have set out to collect in this project. These data sets could potentially make up a new network category branch off of the barriers and limitations—such as costs and incentives—which are costs assigned with using a particular mode, traveling along a particular corridor, or utilizing a particular storage site, for example, that might make it more favorable than any other alternative. Though these data sets are next-level-type data, they are worth mentioning for future research, and they would help to fill in the literature gap previously discussed regarding the fact that little is known about the dynamics between costs of mode and route choice, permitting requirements, and the supply chain.

4.2 Data Discrepancies

We have identified four types of data discrepancies common throughout the collected data—(1) differences in scale, (2) topology errors, (3) differences in attribute-labeling conventions, and (4) spatial and attribute coverage (Table 4). Differences in scale are attributed to the fact that data sets were collected from a variety of sources, which were created for different purposes, and were often created at different cartographic scales (e.g., scale of 1:250,000 compared to 1:5,000). These differences in scale have profound impacts on the precision and accuracy of the data, and they often lead to the other three data discrepancy problems. Partly caused by differences in scale, topology errors are inaccuracies in the spatial relationships between two or more geographic objects. For example, the spatial relationship between a bridge and road is a topological intersection; however, we might have a topology error where a bridge does not intersect the road. Topological accuracy is a prerequisite for many GIS operations and spatial analysis in general. Without proper topology, we cannot perform spatial queries or high-level analysis. Differences in attribute-labeling conventions are not explicitly a spatial problem; however, it can complicate spatial analysis. For example, a single road can be identified as "Main Street" in one data set, but in another data set, the same road is labeled as "Route 11," or "Main St." These discrepancies in naming conventions can complicate or even inhibit relational joins between data sets. Finally, spatial and attribute coverage is an issue when data sets are missing critical attribute information for some spatial objects or have incomplete coverage of spatial objects. See Appendix C for more information on these four data discrepancies.

Most of these data discrepancies are directly related to the fact that data sets come from different sources. Each of these errors exists in one form or another in every data set we collected; however, some data sets seem to be particularly susceptible to such problems. Combined, these four data discrepancies introduce significant analytical challenges—such as failed joins, failed spatial overlays or identification of spatial relationships, and erroneous calculations. Though many of these problems are common throughout the field of GIS, in excess, their role in impeding meaningful analysis should not be underestimated—they must be addressed before any future analysis can take place.

Table 4. Commonly Found Data Discrepancies

Data Issue	Examples
Differences in Scale *Each data set was created at a different geographic scale*	
Topology Errors *Problems with the spatial relationships among spatial objects*	
Attribute-Labeling Conventions *Differences in attribute-naming leads to join issues*	
Coverage *Differences in attribute coverage or spatial coverage*	

A. Road layer from U.S. census, collected at the county level; scale ~1:5,000.
B. National Transportation Atlas Database Freight Analysis Framework (NTAD FAF) national freight road network layer; scale 1:100,000.
C. National Bridge Inventory (NBI) bridge topology issue. Bridges do not intersect with road intersections. Two points are used to represent a divided highway interchange.
D. NBI bridge topology issue. Bridges do not intersect with road intersections. Only one point is used to represent a divided highway interchange.
E. Various attribute-naming conventions observed across different data sets referencing the same road feature.
F. Example of spatial coverage issues across two different bridge data sources—(1) NBI railroad bridges in green, and (2) Homeland Security Infrastructure Program (HSIP) railroad bridges in purple. If NBI railroad bridges were the same as the HSIP railroad bridges, then no purple points should be visible. Instead, some of the HSIP bridges are not present in the NBI data set.
G. Example of attribute coverage issues within a single data set (NTAD FAF roads). The attribute represented is a long combination vehicle (LCV) type, which refers to the type of trucking trailer allowed on a given route and is a proxy for road weight.

4.3 Lack of Analysis Tools

The final data-related gap worth mentioning is the availability of automated tools to assist industry stakeholders in the routing and permitting process. According to stakeholders, there are a few states that utilize some form of automated GIS tool to help in their routing and permitting process. Many of these tools are purchased from a provider such as TomTom, ProMiles, or AutoTurn, and they provide point-to-point turn directions for the transportation company based on similar infrastructure data that we collected and listed in Appendix B. Though many of these states use some form of automated system, there is no national-level tool or set of standard data sets used in state tools to help alleviate the process. Each state does something different—some do not use these automated systems at all, and for those that do, their tools can only help with routing within the state, not across state-lines. At the same time, these tools all utilize different underlying data, and therefore, the data used to inform the routing could be significantly different from state to state. Stakeholders also indicated the need for an application that records dynamic or small-scale barriers, such as overarching trees, low hanging signs and lights, or construction and other detour events. Having some sort of user-generated tool for data creation could easily be integrated into the current road bonding process and be updated regularly by regulatory officials with news of such events.

5 Routing Scenarios

We examined the functionality of the data collected through transportation scenarios in three regions—(1) Wisconsin and Minnesota, (2) Texas, and (3) the Northeast. These three scenarios were chosen based on conversations with industry stakeholders that revealed that these regions not only vary significantly in terms of their OSOW infrastructure, but they also vary greatly within the permitting realm.

For each region, we performed a rail and trucking routing analysis to first examine whether the data were complete enough to represent the transportation network connecting major wind manufacturing facilities (e.g., Tier I wind manufacturing facilities) to known wind site locations, and second, to better understand some of the limitations within the data. See Figure 2 for a map of U.S. wind turbine and wind component manufacturing facilities. Due to a lack of detailed data on waterways, we only considered assessing rail and truck routes for these three routing scenarios. Unfortunately, due to all of the topological issues discussed in Section 4.2, we were unable to automate the routing scenarios. Instead, the routing and decision making was performed manually using the same set of guiding principles—such as mode choice and time and distance forces adjusted for barriers and limitations—that an automated network algorithm would consider. Furthermore, because the analysis was performed manually, it was highly time intensive and consequently, we were unable to look at the entire supply-transport chain of all three major wind components for each region. Instead, we split the transportation of the three major wind components across the three regions, whereby we tailored the routing analysis to the specific transportation breakpoints of the wind component being transported in each region.

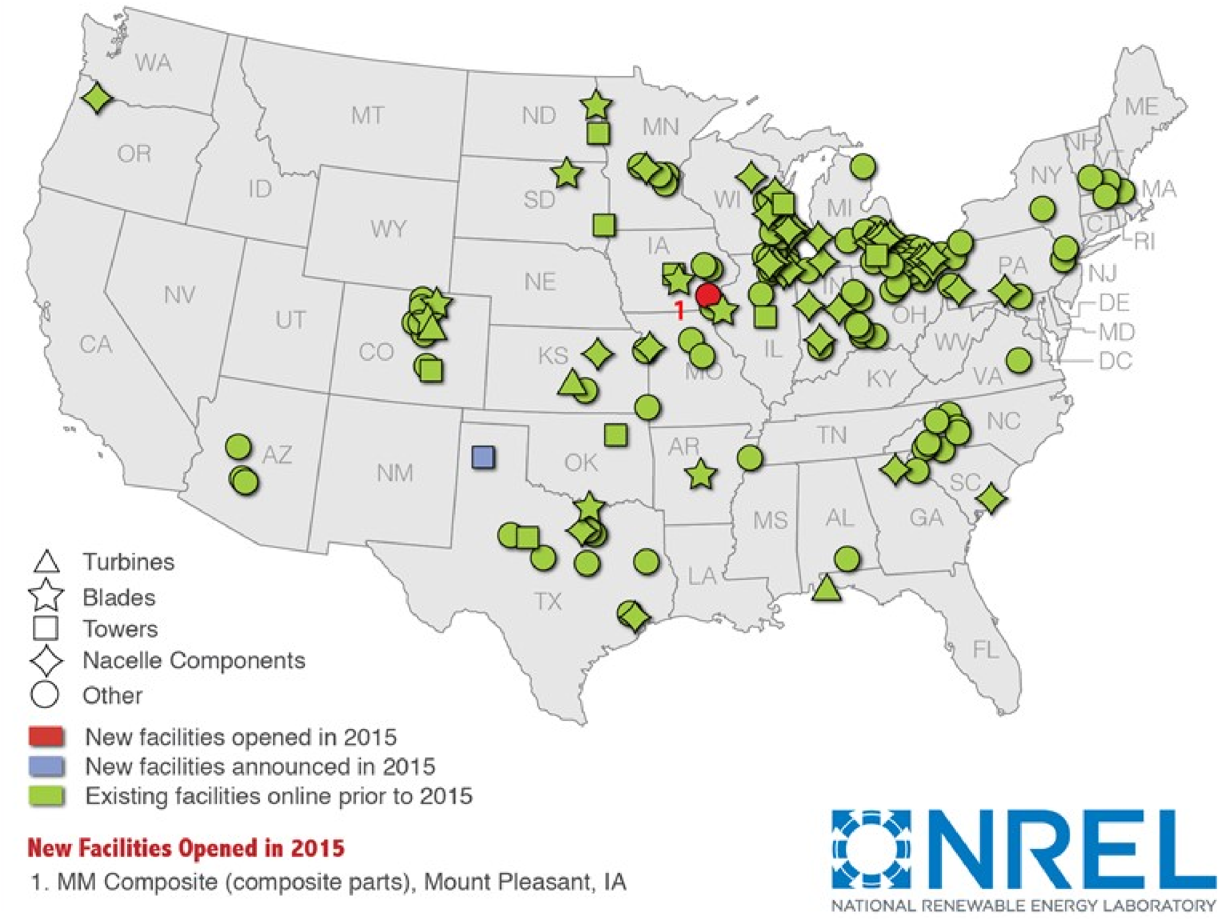


Figure 2. Wind turbine and component manufacturing facilities

5.1 Minnesota-Wisconsin Routing for Blades

We examined Minnesota and Wisconsin because there are a number of wind facilities, and these states have invested significantly in reconditioning their main corridors to facilitate OSOW transportation. We conducted a routing scenario for transporting blade components from the closest major blade manufacturer, located close to the Minnesota border in Grand Forks, North Dakota, to a wind site located in the center of Wisconsin just 40 miles north of Madison (Figure 3). In the routing analysis, we not only identified the most direct routes, but we also factored in blade-related transportation breakpoints such as the turning radiuses of road exits and road curvature. We identified nine specific route changes, or exits, that required a turn from one road to another and 35 potentially sharp turns in the road's curvature. It is important to note that these barriers are only identified as being a potential barrier. Due to the underlying road data, any calculations of turning radiuses will be inherently flawed with wide margins of error; therefore, it is best to use GIS-routing analysis to identify potential barriers that then must be followed up either through detailed aerial imagery or through a boots-on-the-ground inspection.

When compared to the other scenarios, the Minnesota-Wisconsin scenario stands out because we only examined one mode (trucking) and did not consider the potential route for rail. The reason for excluding rail in this scenario is largely related to the missing attribute information of the rail data set in the Minnesota-Wisconsin region. In order for a rail route to be chosen, we set a parameter that looked for rail lines that are either owned by the same company or lines where a single company has tracking rights. In this region, however, roughly half of the ownership and tracking attributes were missing, and therefore, due to incomplete data, rail was excluded from the routing analysis as a potential mode.

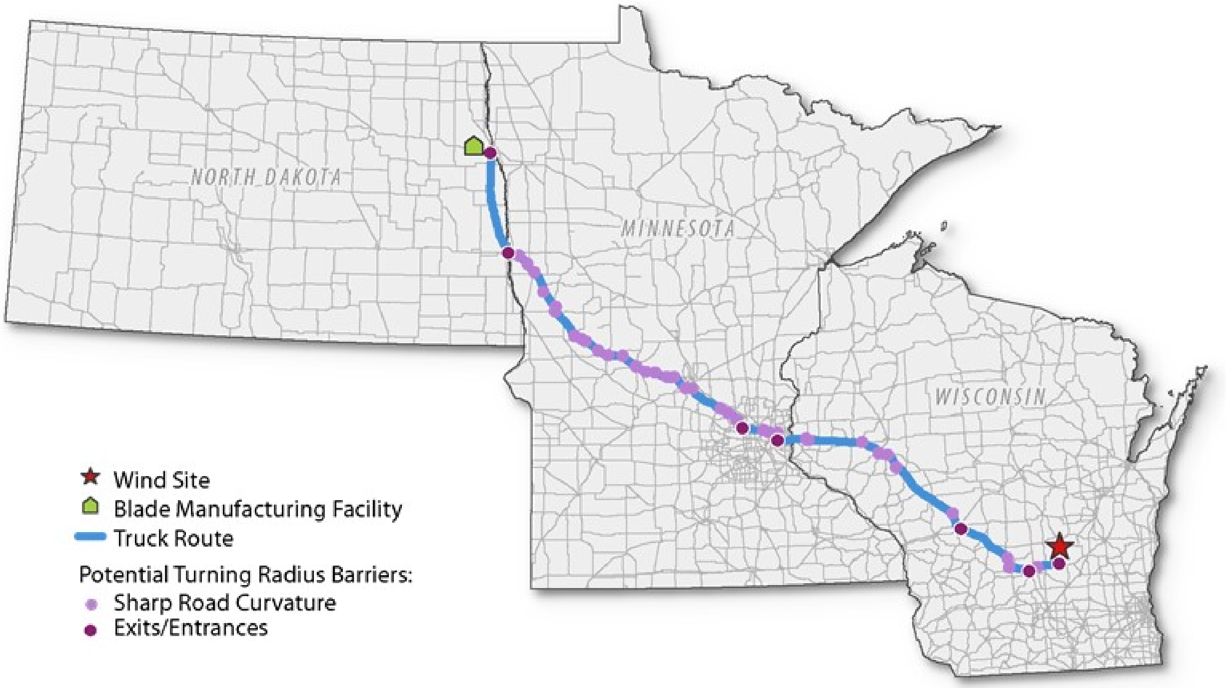


Figure 3. Minnesota-Wisconsin scenario routes for transporting blades

Data sources: Census 2015 county-level road layer; NREL 2015 manufacturing locations for manufacturing site; Ventyx 2013 wind farms for wind site.

5.2 Texas Routing for Nacelle Components

The Texas scenario examined routing the turbine nacelle from the hypothetical closest major manufacturer—located in Amarillo, Texas—to the wind site located in central Texas, just 50 miles northwest of Killeen. Because nacelles are so heavy, their mass plays the largest role in truck route determination. Mapped in Figure 4 are the two best routes that connect the manufacturing site to the remote wind farm—one rail and one truck route—along with their known weight restrictions. As shown, the best possible route is actually the railway route, which in this instance, happens to connect door-to-door from the manufacturing site to the wind site. Because there is a direct connection, choosing this route does not require any truck transportation on public roads, and therefore, would not require dealing with the permitting process.

It is worth noting that, out of the 165 *potential*[2] bridges identified along the rail route, none have bridge load attributes, and therefore, we were unable to identify whether these bridges are capable of carrying the heavy nacelles. Nonetheless, rail has a standard weight limit of 250,000 lbs per railcar, which is better suited for the masses associated with a nacelle when compared to trucking. U.S. interstates and state priority freight highways are mostly limited to 80,000 lbs (for the truck and the trailer combined). Though these bridge weight limits are unknown, we do have data related to weight limits of *potential* road bridges, and as seen in Figure 4, there are a considerable number of *potential* bridges with weight restriction barriers along the chosen route—namely, 67 out of 314. For these reasons, our analysis has determined that the rail route is the best option here.

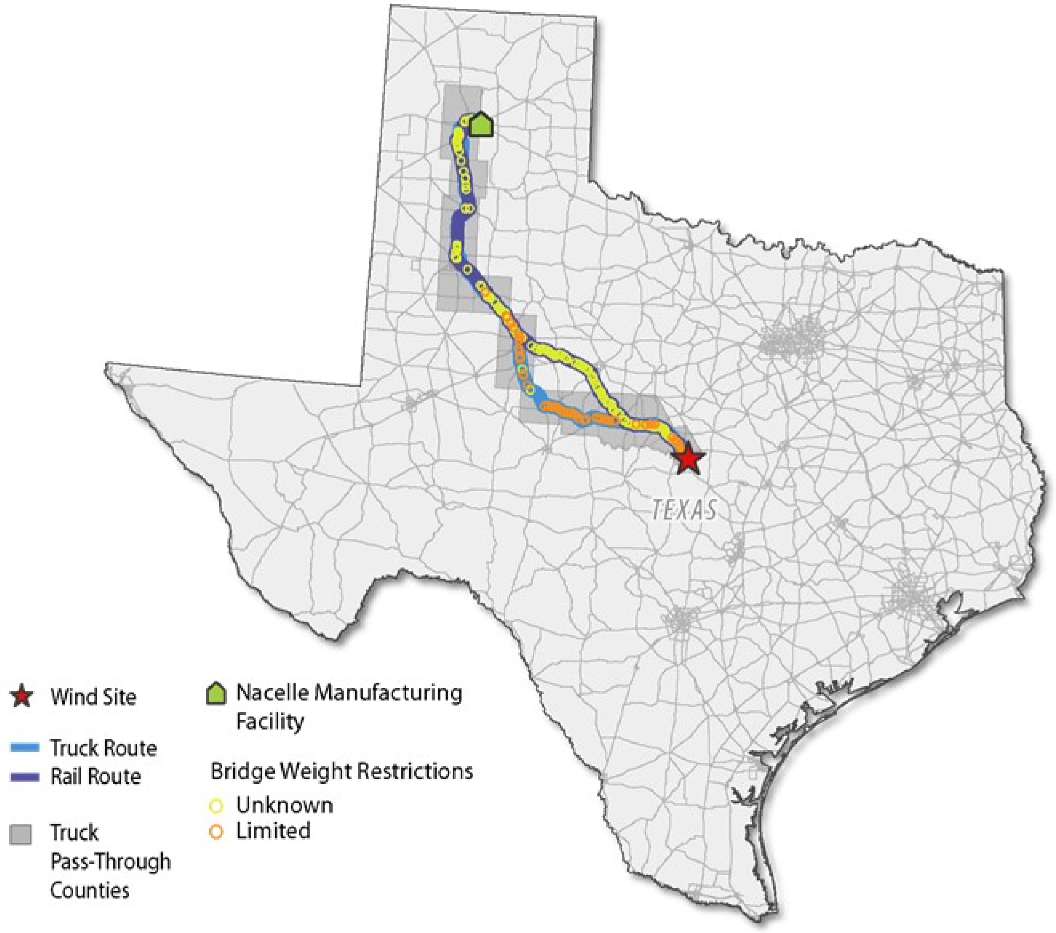


Figure 4. Texas scenario routes for transporting nacelles

Data Sources: Census 2015 county-level road layers; NTAD 2015 railroad layer; HSIP Gold 2013 railroad bridges for bridges along the rail route only; HSIP Gold 2013 tunnels; NBI bridges for bridges along roads only; NREL 2015 manufacturing locations for manufacturing site; Ventyx 2013 wind farms for wind site; U.S. census 2015 counties for pass-through counties; NTAD FAF national road layer for road weights.

---

[2] *Due to topology issues, all bridges within 100 m of the route were grouped together as a* potential *bridge along our mapped route. There is a very good chance that up to half of these bridges might not actually fall along our route. There is also a good chance that a single bridge is being represented by two or more bridges depending on the data for the specific bridge that was collected and tabulated (some bridges are represented by one point while others are represented by multiple points for each division of a divided highway).*

5.3 Northeast Routing for Towers

Figure 5 illustrates the routing analysis for transporting tower components in the Northeast. Because towers are influenced by all major breakpoints, we factored height (clearances), weight limits, and turning radiuses into our routing analysis. Figure 5 shows two potential routes for two different modes, rail and truck, that connect the tower manufacturing facility (Ventower) to the wind farm location on Deer Island, located just outside of Boston, Massachusetts. To determine the best route for each mode, each potential route was first filtered for low vertical clearances. The northern truck route, which takes a similar path to the rail-route shown in Figure 5, was deemed an unviable route and excluded. Next, weight restrictions of roads, rail lines, and bridges were examined to determine whether or not these routes would be able to handle a minimum of at least 80,000 lbs, and thus a 13-axle or 19-axle trailer carrying a tower. Unfortunately, the only data we could obtain regarding weight were from the bridge data sets,[3] and as the yellow circles in Figure 5 indicate, even this data set provides an incomplete picture of weight barriers along these routes. Out of the 376 *potential* bridges along the truck route and 810 *potential* railroad bridges, 125 road bridges and 808 rail bridges were found to be missing data for bridge design load, and therefore, were categorized as unknown. Also, for the trucking route, we found 26 potential bridges with known weight limits that would prohibit tower transport. Finally, we factored in turning radiuses by looking at the number of road changes required and found that there were 15 turns and exits associated with the trucking route mapped in Figure 6. Due to the potential turning radius issues, as well as the 26 known *potential* bridge weight limits, our analysis found the rail route to be the best route and mode for transporting the tower component from Ventower to Deer Island. However, the persisting issue is that none of the rail bridges have weight limits or vertical clearance limits that might make them impassable.

---

[3] *Data on railway weight restrictions were not available; however, we can assume that all freight rail lines are capable of transporting the nacelle. Additionally, the data regarding road weight in the national freight route data set is missing LCV-type attributes for nearly all of the freight roads located in the eastern half of the United States. The LCV-type could have been used as a proxy for weight on roads.*

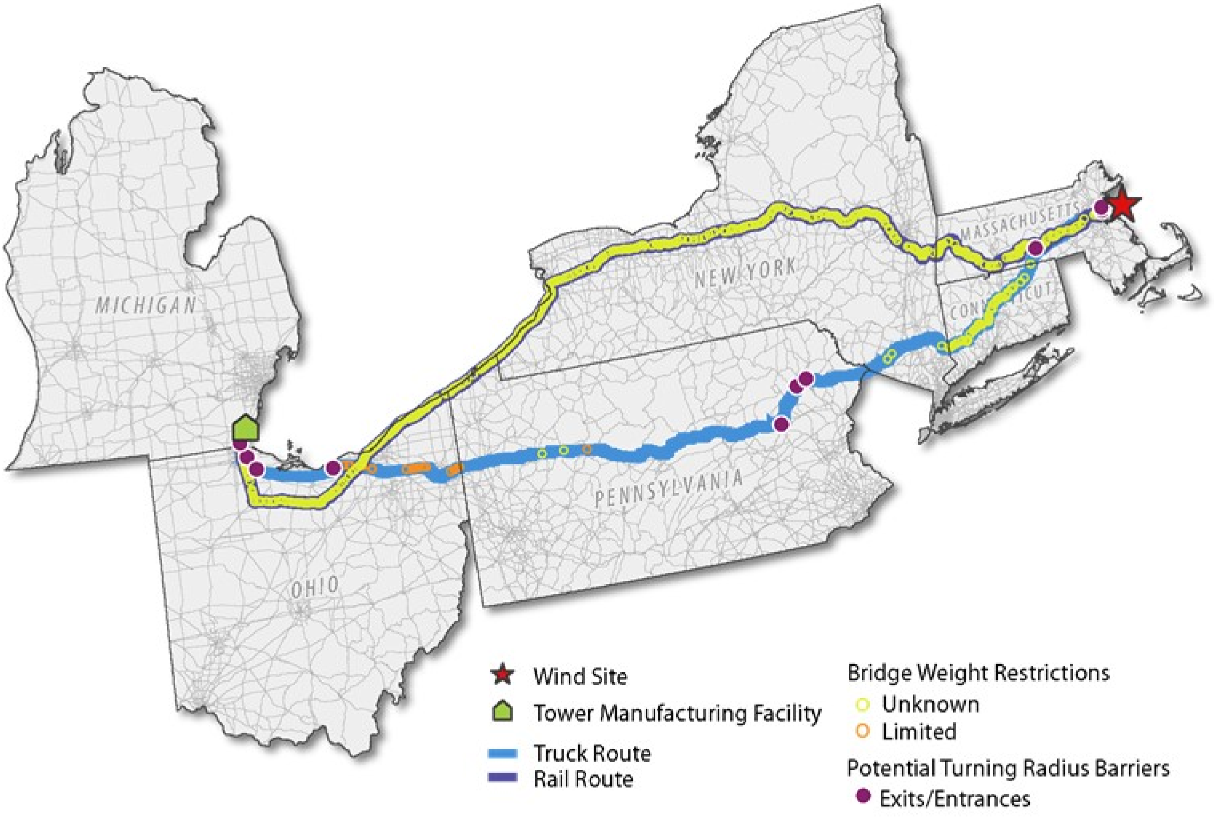


Figure 5. Northeast scenario routes for transporting towers

Data Sources: Census 2015 county-level road layers; NTAD 2015 railroad layer; HSIP Gold 2013 railroad bridges for bridges along the rail route only; HSIP Gold 2013 tunnels; NBI bridges for bridges along roads only; NREL 2015 manufacturing locations for manufacturing site; Ventyx 2013 wind farms for wind site; NTAD FAF national road layer for road weights.

Geographically speaking, the Northeast stands out from the West: States are closer together, population densities are higher, and transportation infrastructure is often older. These characteristics make it incredibly difficult to find a route with no physical barriers, like low vertical clearances or weight restrictions, a route that avoids major urban areas, and a route with the least number of state and local permitting requirements. Unlike the other two regions examined here, there are no major Tier I wind manufacturing facilities located in the northeastern region of the United States—the closest is in Detroit. Therefore, any trucking-based transportation of wind turbines from a Tier I manufacturing facility to the Northeast will likely have to deal with the regulatory hurdles of crossing many states, counties, and municipalities.

6 Discussion

This project has summarized existing geospatial data on transportation infrastructure for wind components, highlighting the common data-related challenges encountered. The data gap analysis identified areas for further analysis, data collection, and limitations of existing tools used in the industry. The routing scenarios illustrated some of the complexities in routing wind turbines and reinforced many of our findings presented in the data gap. From the data assessment, we found that much of the data needed to create a comprehensive database for OSOW wind transportation in the United States already exists. However, there are a number of data discrepancies that must be addressed before any further spatial analysis on infrastructure requirements could take place. Because the data were collected from a variety of sources and for entirely different purposes, they contain significant errors such as those related to spatial scale, topology, attribute-naming conventions, and spatial and attribute coverage inconsistencies; therefore, compiling them into the same database will require a considerable amount of cleaning and preprocessing in order to deal with many of the data discrepancies reported. Indeed, given the current state of the data, we have found that it will be challenging to perform meaningful analysis until these fundamental data discrepancies have been addressed. With that said, this report has identified that there is a need for a national-based OSOW wind transportation infrastructure database that could be utilized in existing automated systems or in other tools that would alleviate many of the hurdles in the permitting process.

6.1 Future Work

Future work on understanding the capacity of infrastructure systems for wind component transportation should further examine network and spatial analysis constructs and the connectivity of infrastructure systems across the United States. This would involve creating detailed spatial data on regulatory and permitting requirements (e.g., escort requirements, allowable overhangs, maximum weight limits) at various cartographic scales in order to spatially assess permitting uniformity or areas of contention (bottlenecks). Doing so would help identify wind component transportation corridors, both within and across states, and based on both physical and regulatory bottlenecks, that could be targeted for state-to-state (or jurisdiction-to-jurisdiction) permitting uniformity or infrastructure-based development. This would promote the transport of wind components along these routes and break down many of the physical and regulatory barriers that increase the costs of wind deployment. Before any of this could be done, however, it is clear that substantial work would be needed in data cleaning, repairing topology, and data integration.

References

Anderson, J. 2012. "Experience and the Evolution of Wind Power Project Costs in the United States." Stanford, CA: Stanford Institute for Economic Policy Research. Accessed November 10, 2015. ftp://ftp.repec.org/opt/ReDIF/RePEc/sip/12-021.pdf.

AWEA Transportation Logistics Working Group. n.d. "Meeting the Wind Power Challenge – Industry Perspective." Accessed September 9, 2016. http://highwaytransport.transportation.org/Documents/Follidori.pptx.

Cotrell, J., T. Stehly, J. Johnson, J.O. Roberts, Z. Parker, G. Scott, and D. Heimiller. 2014a. "Analysis of Transportation and Logistics Challenges Affecting the Deployment of Larger Wind Turbines: Summary of Results." Golden, CO: National Renewable Energy Laboratory. Accessed November 10, 2015. http://www.nrel.gov/docs/fy14osti/61063.pdf.

Cotrell, J., T. Stehly, J. Johnson, J.O. Roberts, Z. Parker, G. Scott, and D. Heimiller. 2014b. "Land-Based Wind Turbine Transportation and Logistics Barriers and Their Effects on U.S. Wind Markets." Golden, CO: National Renewable Energy Laboratory. Accessed November 10, 2015. http://www.nrel.gov/docs/fy14osti/61780.pdf.

Del Franco, Mark. 2015. "How 'Wind Turbine Logistics' Became A Four-Letter Word." North American Windpower. Accessed November 10, 2015. http://www.nawindpower.com/issues/NAW1412/FEAT_02_How-Wind-Turbine-Logistics-%20Became-A-Four-Letter-Word.html.

Levine, Aaron and Jeff Cook. Forthcoming. Transportation of Large Wind Components: A Permitting and Regulatory Review. Golden, CO: The National Renewable Energy Laboratory.

Roy, M. 2016. "Finding and Using Data to Identify and Evaluate Corridors for Transporting Multi-State, Multi-Modal Oversize/Overweight Freight." Accessed November 11, 2015. http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=3639.

Zayas, Jose, Michael Derby, Patrick Gilman, and Shreyas Ananthan. 2015. "Enabling Wind Power Nationwide." Washington, D.C.: U.S. Department of Energy. Accessed September 9, 2016. http://www.osti.gov/scitech/biblio/1220457.

Appendix A: Proposed Solutions

Table 5. Potential Solutions for Transporting Taller Towers

Solution	*Description*
On-site manufacturing	Onsite manufacturing has been proposed as one potential solution to avoid most transportation limitations involved with turbine component transportation. Whether or not such an approach will be economically viable is still yet to be determined. The creation of the manufacturing site will require more land needed for the wind project that will likely go vacant after the turbines have been installed. At the same time, the costs associated with creating temporary manufacturing facilities may actually be more expensive than dealing with transportation logistics.
Design innovations	A number of different innovations for turbine design have been proposed that could help alleviate some transportation breakpoints and allow for bigger turbine development. For instance, segmented blades could allow for longer blades and easier transport of blades while nacelle weight technologies could reduce the weight of high-megawatt nacelles, and thus, increase its transportability. Additionally, tower material manufacturing methods and design could lead to cheaper and lighter components.
Regional permitting/multi-state permits	Regional and multistate permitting could be centralized to facilitate the permitting process between regions. There are a few notable examples, including WASHTO (multistate), TxPROS (Texas), Wisconsin OS/OW Freight Network, SASHTO (various southern states), MAASTO (Mid-Atlantic), and NASTO (North-Atlantic) (AWEA).
Specialized trailers	Specialized trailers that can carry heavier loads and reduce turning radiuses allow for larger turbines around difficult turns.
Design and permitting uniformity	Dimension design uniformity of turbines will allow for ease of transport given the breakpoint conditions. Permitting uniformity would unify the permitting process and specific restrictions placed by jurisdictions.
Collaboration with state and local governments	Collaboration with state and local governments would mostly justify the need for permit uniformity across states and regions.
Creation of wind corridors	The creation of wind corridors, or route networks, that connect manufacturing sites to potential wind sites could centralize the permitting process required on these routes.

Sources: Cotrell et al. 2014a; Cotrell et al. 2014b

Appendix B: Data Catalogue

Table 6. Data Sets Collected for Network Routes

Network Category	Category Type	Feature Type	Description	Last Update	Licensing	Source	Join Attributes	Important Attributes	Scale	Priority
Routes	Rail	Rail	Railroad lines	2015	Public	NTAD 2015; FRA 2015	Rail ID/name	Track owner, tracking right companies	1:100k	1
Routes	Rail	Rail	Railroad nodes	2015	Public	NTAD 2015; FRA 2015	Rail ID/name	--	1:100k	2
Routes	Road	Road	NHS's Freight Road Network	2010	Public	NTAD 2015; FAF v3.4	Name, type	LCV (trailer) type permitted, NHS class, traffic flow, projected traffic, VMT, etc.	1:100k	1
Routes	Road	Road	All roads, by state and county	2015	Public	U.S. census	Name, type	Route class and type	NA (~1:5k)	1
Routes	Water	Water	Navigable waterways	2011	Public	ONL 1993; NTAD 2015	River name	Type/class	--	1

Table 7. Data Sets Collected for Network Sites

Network Category	Category Type	Feature Type	Description	Last Update	Licensing	Source	Join Attributes	Important Attributes	Scale	Priority
Sites	Destination sites	Potential	Individual turbine locations	2013	Private	Ventyx 2013	--	Site name, height	--	3
Sites	Destination sites	Potential	Boundaries of wind farm sites	N/A	Private	Ventyx 2013	--	Site name, capacity	--	3
Sites	Origin sites	Manufacturing	Wind-related manufacturing facilities	2014	Public	NREL 2015	--	Type of facility, component class	--	1
Sites	Transit nodes	Multimodal node	Multimodal facilities	2015	Public	DOT BTS; NTAD 2015	--	Mode, mode types serviced, name	1: 24k	3
Sites	Transit nodes	Rail yard	Rail yard locations	2011	Private	ORNL 2011; HSIP Gold 2013	Rail names, IDs	Rail names using facility, length, width, type of facility, number of tracks, storage capacity	--	2
Sites	Transit nodes	Ports	World Port Index port locations	2016	Public	WPI 2016	--	Channel depth, port type, shelter type, overhead limit	--	1

Table 8. Data Sets Collected for Network Barriers

Network Category	Category Type	Feature Type	Description	Last Update	Licensing	Source	Join Attributes	Important Attributes	Scale	Priority
Barriers	Physical	Dams	Dams	2014	Public	National Atlas; NTAD 2015	River name	Dam purpose	1:2mil	3
Barriers	Physical	Bridges	Railroad Bridges	2009	Private	ORNL; HSIP Gold 2013	Rail line ID/name	Minimum vertical clearance, design load, length, number of tracks, age	--	1
Barriers	Physical	Tunnels	Railroad Tunnels	2011	Private	ORNL; HSIP Gold 2013	Rail line ID/name	Through route, origin/destination route, routes using tunnel, length, terrain, number of tracks, vulnerability, azimuth, age	--	1
Barriers	Physical	Tunnels	Road and Rail Tunnels	N/A	Private	HSIP Gold 2013	Rail/road ID/name;	Designed level of service, intersection feature, minimum vertical clearance, functional classification, year built (age), number of lanes on/under, design load, live load, width, median	--	1
Barriers	Physical	Bridges	Road Bridges	2011	Public	NBI; NTAD 2015	Road on/under bridge	Road on/under bridge, vertical clearance, horizontal clearance, design load, live loads, number of lanes on/under structure, age, width, median	--	1
Barriers	Regulatory	Urban areas	Urbanized Areas Boundaries	2013	Public	US Census 2010; NTAD 2015	State	--	1:500k	2
Barriers	Regulatory	States	State Boundaries (Coarse)	2015	Public	U.S. Census 2010	State	--	1:500k	1
Barriers	Regulatory	Cities/towns	Cities and Town Boundaries.	2015	Public	U.S. Census 2010	State, county	Name, type of place (legal description)	1:500k	1
Barriers	Regulatory	Counties	County Boundaries	2015	Public	U.S. Census 2010	State, name	Name	1:500k	1
Barriers	Regulatory	Consolidated city	Consolidated City Boundaries	2015	Public	U.S. Census 2010	State, county	Name	1:500k	1
Barriers	Regulatory	Accidents	Vehicle Accident Location Points	2014	Public	FARS v.2; NTAD 2015	State/county/city name	Number of accidents, type of accidents	--	2
Barriers	Regulatory	Permitting association	State Boundaries of Permitting Associations	2016	Public	NREL 2016	State	Name of association	1:500k	1

Appendix C: More on Data Discrepancies

Detailed in this section are the common data discrepancies present in the array of existing data we collected—namely, issues of scale, topological errors, differences across attribute-naming conventions, and issues related to spatial and attribute coverage.

Differences in Scale

Differences in scale exist across every data set collected. These differences are caused, in part, by the fact that each data set has a different origin and were created for different purposes; therefore, each data set was created at different geographic scales. For example, two different road layers were collected in Table 4-A and 4-B, illustrating these issues of scale. Table 4-A on the left is a road layer data set made available by the U.S. census at the county level, while Table 4-B is the same road feature that comes from a national-level data set of freight roads, the Freight Analysis Framework (FAF). Though both data sets in 4-A and 4-B represent the same geographic features (i.e., roads), they were collected for different purposes and at different scales. These differences in scale have profound impacts on the precision and accuracy of the data, whereby the larger-scale data set in Table 4-A is both a more precise and more accurate representation of the roadway. In addition, these differences in scale lead to a host of other data discrepancy issues, as discussed in the following sections. Unfortunately, the data set shown in Table 4-B has more detailed attributes related to the roadway, such as what type of trailer the road allows (e.g., dethatched 3+ axle), which is critical to our project needs while Table 4-A has very limited, if any at all, useful attribute information tied to the features. In order to merge and join the best parts of each data set—the spatial geometries of Table 4-B and the attribute information of 4-A—complex querying was required. Differences in scale can result in related data discrepancies, such as topological errors, attribute labeling, and data coverage.

Topology Errors

Topology errors are errors in the spatial relationships between two spatial objects. The most common topological error among our data sets was feature intersections, particularly amongst bridges and road layers. For instance, of the 650,000 bridges in our NBI data set, none intersected with the most spatially accurate road layer (the same phenomenon holds true for the rail bridges and railroad layers, both of which come from the same source—HSIP). Again, part of this is related to the scales in which these data sets were acquired—different spatial scales lead to data offsets in precision and accuracy, and therefore, topological errors. Another problem, however, is related to choices in feature representation (e.g., choosing to represent a feature as a point instead of a polygon or as a series of points). This problem is perhaps best understood by looking at representations in Table 4-C and 4-D. In both images, the bridge points do not intersect the road intersection nodes, which is likely because of the spatial scales of each data set. Nevertheless, even if these two data sets were at the same spatial scale, this topology error would still exist because the real-life three-dimensional bridge is abstracted as a single coordinate point rather than a polygon. To make matters more complex, the bridge data set is not consistent with its point feature representation between the two images—Table 4-C shows a single bridge represented as two separate points for each division of the interstate road that is under the bridge, and Table 4-D shows a single bridge point for all road divisions. If we take this issue and explore it further, we see that the bridge is not represented across lane divisions for the road-on-the-bridge or the road-under-the-bridge. Figure 2 further illustrates this feature representation issue from Table 4-C. In Figure 6, panel C-1 represents the original bridge representation as two points for each road division. Figure 6-C-2 shows an areal image of the bridge in question, and C-3 is a topologically corrected representation of how C-1 should be represented if all divided highway lane intersections were represented. At the interchange of two divided highways, there must be at least four bridge points representing a single bridge in order to account for these topology errors—and if there are any separated exit or entrance ramps intersected by the bridge, additional points must be included.

From left to right: C-1 is a screenshot of the original bridge point representation; C-2 is an areal image of the bridge; and C-3 is a topologically corrected version of the bridge layer.

Figure 6. Topology representation issue explained, from Table 4-C

Bridge topology is important not only for identifying potential barriers in clearance or weight, it is also important for creating an accurate network topology for road intersections. Topology errors arise when road interchanges—or road junctions that are vertically separated by one or more ramps—are incorrectly treated as road intersection nodes, which connect two routes together. Fixing these bridge topology issues is no simple feat. As mentioned earlier, there is no clear pattern in how bridges are topologically represented (e.g., represented as one-point feature versus multiple-point features), and this introduces a layer of complexity when trying to topologically correct each bridge point to a road intersection node. Not only do we need to identify the closest bridge to each road intersection with the most similar name, we also need to make sure we duplicate the bridge for each road lane division. Further complicating the issue, bridges must also be separated into different types of bridges, and these different typologies must be treated differently when correcting for these topology errors. For instance, bridges that allow a road to cross a water feature must have a topological rule whereby the bridge intersects only one road feature, whereas an interchange bridge must intersect at least two roads and should have multiple points that represent the same bridge in order to account for all road lane divisions (e.g., Figure 6-C-3. This research has identified at least eight different typology classes for bridges that should be considered when correcting for topological errors—and with over 600,000 bridges, there are many corrections that must be made before any analysis can take place. In this sub-section, we only discussed topological issues related to bridges and road features; however, similar errors exist across all the data sets. Therefore, in order to standardize the data set for even the most basic spatial analysis, a considerable amount of effort must be put into correcting for these topology errors.

Differences in Attribute Naming/Labeling Conventions

Another common data discrepancy issue is related to differences between attribute-naming and labeling conventions, both across data sets and within. This issue is so common that no single data set we collected was immune to this discrepancy; however, some have higher prevalence of such issues. In particular, the attribute issues within the bridge and road layers seem to be especially widespread. Table 4-E illustrates how a single geographic feature might be labeled in a variety of ways. Some attributes are labeled similarly and are off by just a few characters (e.g., “Santa Fe Dr.” and “Santa Fe Drive”), while others are labeled using different nomenclature conventions altogether (e.g., “US HWY 85” and “Santa Fe Dr.”). Such differences in attribute labeling lead to problems when performing database relational joins. Simply put, if two text strings are not the same, the join will fail. One potential solution for differences in text strings is called fuzzy joining, which looks for similarly structured strings and joins the most similar together; however, fuzzy joining is only a solution for similarly labeled attributes. If two layers use different labeling conventions (e.g., “US HWY 85” and “Santa Fe Dr.”), we cannot use fuzzy joining and must develop a more complex solution to join the data sets together (e.g., generating a lookup table for all such combinations). The inability to perform joining due to this attribute discrepancy is yet another layer of complexity that hinders our ability to preprocess and clean the data for a national-level infrastructure analysis in an automated manner.

Coverage Issues

The final data discrepancy issue is related to coverage, or problems regarding the completeness of the data. Not only do we have missing data, we also have the problem of incomplete data, whereby data sets are missing critical attribute information for many spatial objects or are missing spatial coverage all together. Table 4-F and 4-G illustrate these coverage issues for spatial and attribute coverage, respectively. First, panel F exemplifies the issue of spatial coverage, whereby we can see that two bridge data sets from different sources do not align. Given the nature of the data sets, we would expect there to be more green bridge data points than purple bridge data points. This is because these data come from the NBI and represent all bridges, including railroad bridges, whereas the purple points are from the HSIP data set for railroad bridges only. However, some bridge data records in the HSIP data set are not represented in the NBI data set and vice versa. The fact that these two data sets differ illustrates this issue of spatial coverage, whereby we are missing data in one, or potentially both, of the bridge data layers. At the same time, due to other issues of topology, scale, and attribute-naming conventions, any attempt to deal with these spatial coverage issues is further complicated by the difficulties associated with differentiating missing features from the redundancy between the two data sets.

The second type of coverage issue present among these data is related to missing attribute information. Missing attribute information can be seemingly random, in the case of bridges missing vertical clearance attributes, or they can have a clear spatial pattern. The latter is the case for road long combination vehicles (LCV) type attributes, which is a field that can be used as a proxy for road weight. As depicted in Table 4-G, LCV-type attributes for road layers are missing for most of the United States except for the Northwest. The fact that attributes are missing for particular layers is incredibly important because it limits the amount of information we have to make informed decisions about routing or infrastructure connectivity. Surely, incomplete data are a common problem that exists in much of the data world; however, the level of incompleteness that exists across the data sets collected is significantly high and must be addressed before any real analysis can be performed using these data.