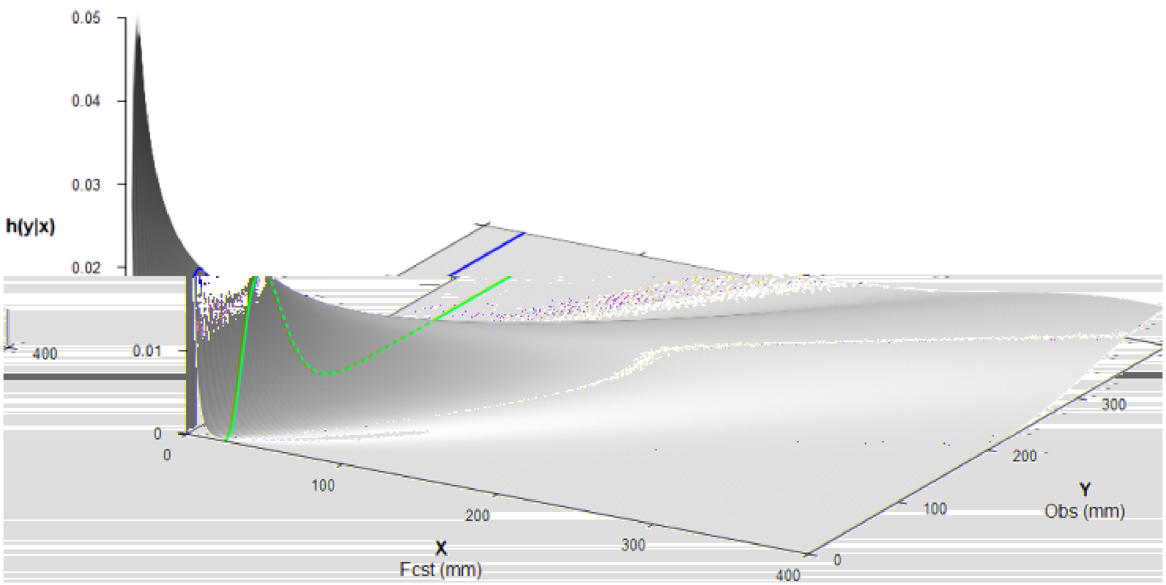
Ensemble Forecasting with the Hydrologic Ensemble Forecast Service (HEFS)

As Implemented at the
California-Nevada River Forecasting Center
(CNRFC)



Operational configuration described: 05/01/2022
This document last updated: 10/05/2022

Please email questions or comments about this document to cnrfc.webmaster@noaa.gov, and include "HEFS at CNRFC" in the subject line.

Table of Contents

1 Introduction

The purpose of this document is to help clarify how National Weather Service (NWS) Hydrologic Ensemble Forecasts are generated at the CNRFC. It is hoped that an increased understanding of the fundamentals, process, and limitations will lead toward (1) more informed and appropriate applications by users and (2) ideas for improvements and refinements by researchers and collaborators.

Hydrologic forecasts provide value to a variety of sectors including flood management, reservoir management, water resources management, hydropower, navigation, and recreation. Historically, hydrologic forecasts have been single-value (deterministic) and of short duration (a day or two) given the uncertainty in the weather forecast. Improvements in weather forecast skill has led to longer forecast durations (e.g. 5 days) in some locations. Probabilistic forecasts (usually regression based) have been restricted to seasonal volume forecasts associated with snowmelt (e.g. April-July volume).

Over the past two decades it has become clear that water resource and emergency managers need more than a single-value forecast. They are managing the risks of their actions (or inactions) and the associated costs. Risk is the product of probability and consequence. They understand the consequences. They need the probability. They need probabilistic hydrologic forecasts for short, medium, and long-range decision making.

Figure 1 - Forecast Use by Temporal Range

Probabilistic forecasts can be generated in multiple ways. The most common are through error propagation and through ensemble techniques. For reasons associated with feasibility and application, ensemble techniques have been the focus of the National Weather Service's development efforts for some time. Progress has been attributable to a growing acceptance that uncertainty is something that can be leveraged to make more informed decisions (National Research Council of the National Academies 2006) and substantial community support as evidenced through the success of the Hydrological Ensemble Prediction Experiment (Schaake et al. (2007), www.hepex.org).

Sources of uncertainty which contribute to uncertainty in the streamflow forecast include: meteorology, hydrology, and flow regulation. Assuming that the (1) hydrologic model is well-conceived and well-calibrated, (2) observations used to drive the model are representative and quality controlled, (3) the individual running the model is well-trained and experienced, and (4) observed and near-term regulated flows are well-defined, then the majority of uncertainty in CNRFC streamflow forecasts typically arises from the uncertainty in future weather (precipitation and air temperature forecasts).

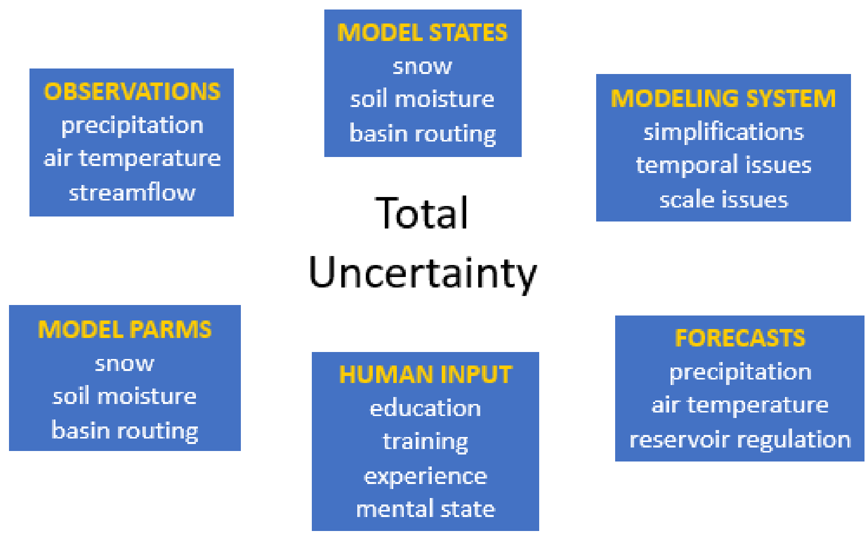


Figure 2 - Sources of Streamflow Forecast Uncertainty

The NWS effort to develop a methodology and toolset capable of generating reliable short, medium, and long-range hydrologic ensembles began in about 2001. Prototype efforts took nearly 10 years to make their way into operations. Today, the CNRFC uses the Hydrologic Ensemble Forecast Service (HEFS) to issue forecasts daily at 285 locations (**Figure 3**). HEFS forecasts are operationally relied upon by water, emergency, environmental, hydropower, and recreation managers to manage risk and improve outcomes.

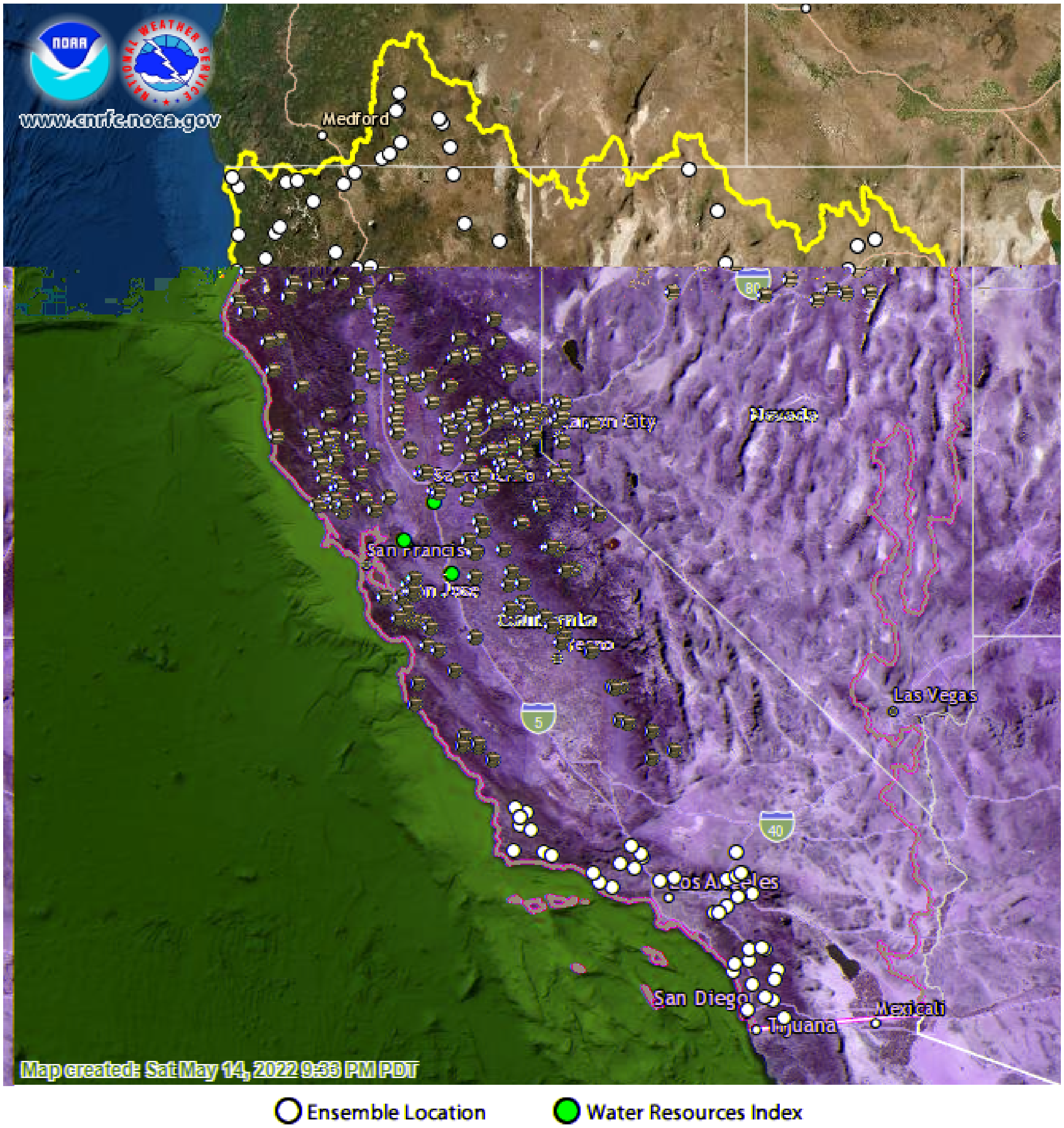


Figure 3 - CNRFC HEFS Forecast Locations

2 Summary Points

The CNRFC uses the HEFS to construct meteorological ensembles from single-valued forecasts. This process, known as statistical post-processing, leverages information from the record of past forecasts and observations. Important aspects of this process as implemented at the CNRFC, and characteristics of the resulting streamflow ensemble forecast, are listed below.

- The CNRFC hydrologic models require three meteorological inputs (hydrologic forcings) for simulation of runoff. These are 1) precipitation, 2) surface (2 meters above the ground) air temperature, 3) and freezing level. The CNRFC uses the HEFS to generate ensemble forcings for precipitation and air surface temperature. Work is underway at the CNRFC to configure the HEFS to also generate an ensemble forecast of freezing level.

- The precipitation and surface air temperature ensembles generated by the HEFS are bias-reduced and exhibit consistent variability with respect to the record of forecasts. Furthermore, the HEFS employs a method known as the Schaake Shuffle (Clark et al., 2004) to ensure that the ensembles are temporally and spatially consistent across locations.

- The CNRFC ensemble streamflow forecasts, which result from applying the ensemble forecasts of precipitation and surface air temperature to the hydrologic models, reflect meteorological uncertainty only. Hydrologic uncertainty is not accounted for.

- The CNRFC uses the HEFS to develop ensemble forcings on a 6-hour time step. The forcings are applied to the hydrologic models at the elevation zone scale (one to three zones per modeled subbasin). The CNRFC configures the HEFS to generate ensembles spanning the first 28 days of the forecast. For morning (T0 = 12z) forecasts, the CNRFC appends ensemble forcingss for days 29 to 365 generated outside of the HEFS using raw climatology to support water supply forecasts.

- The HEFS consists of two primary components: the Meteorologic Ensemble Forecast Processor (MEFP), and 2) the Ensemble Postprocessor (EnsPost). The CNRFC uses the MEFP, but does not run the EnsPost. The EnsPost is designed to modify streamflow forecast ensemble members to minimize bias and incorporate uncertainty attributed to hydrologic modeling. The CNRFC plans to test EnsPost pending completion of planned improvements. A timeline for testing and implementation has not been established.

3 Important Concepts

Before diving into details of how HEFS works and how it is applied at the CNRFC, several prerequisite concepts are introduced in the following paragraphs.

Time of Forecast (T0)

The time at which the forecast portion of the hydrologic forcings begins is referred to as T0 (pronounced “T zero”). CNRFC issues regular forecasts once per day reflecting T0 = 12z during summer, and twice per day (reflecting T0 = 12z, 18z) the rest of the year. Note that “z” indicates time zone, which is UTC or “zulu” time. During active weather CNRFC will issue forecasts as often as four times per day (reflecting T0 = 12z, 18z, 00z, 06z). The times at which forecasts are actually issued are typically 2 to 3 hours after the T0 reflected by the forecast. This is because time is required to develop the CNRFC official single-valued forecast for the CNRFC region, quality check hydrologic data, run hydrologic models and inspect results, and review bulletins. Even though there is a 2 to 3 hour difference between T0 and when the forecast is issued, the forecast is generally referred to by the T0 value it reflects. The CNRFC strives to issue all forecasts (T0 = 12z, 18z, 00z, 06z) by 9am, 3pm, 9pm, and 3am local time respectively.

HAS Forcings and CNRFC Official Single-valued Forecasts

The CNRFC creates single-valued forcings, referred to as the “HAS forcings”, which are applied to hydrologic models to produce the CNRFC official single-valued streamflow forecasts. The forcings span 20 days, from T0 - 10 days to T0 + 10 days. A group of acronyms used to reference the six basic components of HAS forcings are listed in **Table 1**. In these acronyms, Q = quantitative, P = Precipitation. T = Temperature, Z = Freezing Level, E = Estimate, and F = Forecast.

Table 1 - Acronyms for Portions of the HAS Forcings

T0

Forcing Type	10 days before T0	10 days after T0
Precipitation	HAS QPE	HAS QPF
Air Temperature	HAS QTE	HAS QTF
Freezing Level	HAS QZE	HAS QZF

An important point with respect to developing forcings for ensemble forecasts is that *the only portions of the HAS forcings used to create ensemble forcings are:*

- Forecast portion: days 1 - 3 of the HAS QPF,
- Observed portion: all 10 days of the HAS QPE, HAS QTE, and HAZ QZE.

Each of the six parts of the HAS forcings have unique data sources, which vary with forcing type and day of forcing. These are summarized in **Table 2**.

Table 2 - Sources of HAS Forcings

T0

Forcing Type	Days before T0		Days after T0		
	10 - 2	1	1 - 3	4 - 6	7 - 10
Precipitation	prev. Hydro QC	Hydro QC	WPC / NBM	WPC	NBM
Air Temperature	prev. URMA/RTMA	URMA/RTMA	NBM		
Freezing Level	prev. HRRR	HRRR	GFS / ECMWF blend		GFS

WPC -	Weather Prediction Center
NBM -	National Blend of Models
GFS -	Global Forecast System
ECMWF -	European Center for Medium-Range Weather Forecasts
Hydro QC -	CNRFC Hydrologic QC process (“prev.” = from previous forecast)
URMA -	Unrestricted Real-time Mesoscale Analysis
RTMA -	Real-time Mesoscale Analysis
HRRR -	High-resolution Rapid Refresh

In order to provide a more complete picture of the CNRFC forecast process, details on how the CNRFC single-valued forcings are developed is provided in the following paragraphs.

The 10-days of observed forcings are created as follows. On each day of forecast, the most recent 6-hour periods (since the previous forecast), of gauge precipitation are quality checked by the Hydro forecaster. These point values are normalized to the monthly Parameter-elevation Regressions on Independent Slopes Model (PRISM) normals. The normalized values are then distance-weighted to each precipitation grid cell, and then restored at each grid cell. The result is a set of 6-hour precipitation grids reflecting observed precipitation at gage locations and monthly PRISM patterns at grids in between. For surface temperature, grids from the URMA are used for hours 24, 18, and 12, prior to T0. For hour 6 prior to T0, the temperature grid from the RTMA is adopted. For freezing level, 6-hour grids from the HRRR are adopted. For the 9 days prior, observed 6-hour grids from previous forecasts are used.

The 10-day CNRFC single-valued forcings (or HAS forecast) are created as follows. For forecast days 1 - 3, during periods of quiet weather, precipitation grids from the NBM or WPC guidance are typically adopted without changes. For forecast days 1 - 3, during periods of more active weather, precipitation forecasts from the NBM or WPC guidance and current observations are still considered, but with the HAS forecaster providing additional review and input. During active weather, the HAS forecaster will typically review additional models, latest observations from gages on the ground, radar, and satellite. The HAS forecaster uses the Graphical Forecast Editor (GFE) software tool to display and edit NWP grids to create the final adjusted 6-hr precipitation grids. GFE grid editing capabilities include:

- Use grids from two different atmospheric models for two different time steps.
- Blend grids from two different models for the same time step.
- Manually edit grids to be more consistent with latest observations.
- Manually edit grids based on experience with model performance during past events.

.
For forecast days 4 - 6, the HAS forecaster usually adopts the precipitation grids from WPC guidance, but has the option to do otherwise if the situation warrants. For forecast days 7 - 10, precipitation grids from the NBM are adopted. Temperature forecasts (days 1 - 10) are based solely on NBM guidance. For freezing level, the HAS forecaster creates a blend of GFS and ECMWF grids for days 1 - 6. For days 7-10, GFS freezing level grids are adopted. With 6-hour forcing grids for precipitation, surface temperature and freezing level defined for the full 20-day simulation period, zone-average values are then extracted to obtain the HAS forcings required to execute the hydrologic models.

During operations, the HAS forcings are applied to the hydrologic models. The Hydro forecaster inspects the results of the simulation, primarily by comparing recent simulated flows against observed flows. Occasionally the Hydro forecast will apply a “modifier” to the model to improve model performance. Modifiers can be made to model states, parameters, or time series (including forcings). Modifiers can be applied during the HAS forecast or in simulations outside of operations, such as update states simulations. As long as a modifier has not expired, it will be applied in the next operational forecast. With the HAS forcings applied, and the hydrologic models run and results reviewed, the CNRFC official single-valued streamflow forecast is issued.

Hydrologic Models

Hydrologic models are used to transform the 6-hour forcings of mean areal precipitation, air temperature, and rain/snow elevation into subbasin runoff. Air temperature is the areally-averaged value at 2 meters above the elevation of the centroid of the modeled area. Rain/snow elevation is at the elevation of the centroid. The first 10-days of forcings, prior to T0, reflect observations. For all subbasins, the forecast portion of the forcings, after T0, are the HAS forecast when producing the CNRFC official single-valued streamflow forecasts, and multi-valued (ensemble) for when producing ensemble streamflow forecasts.

Each modeled area is a subbasin elevation zone, with snow accumulation/melt processes and soil moisture/runoff processes modeled independently in each. Each CNRFC subbasin is composed of up to 3 such zones, depending on the elevation range of the subbasin. Typically, each elevation zone consists of the subbasin area at: 0 to 5,000 ft, 5,000 to 8,000 ft, or greater than 8,000 ft. As such, the zones are separated by the 5000 ft and 8,000 ft elevation contours. An example of elevation zones is shown in **Figure 5**.

Within each subbasin, each elevation zone is represented by a SNOW-17 (Anderson, 1976) model and a SAC-SMA (Burnash, 1973) model. SNOW-17 and SAC-SMA models are continuous simulation models. Model states of both models are saved with each simulation (forecast), and provide initial states for the next forecast. SNOW-17 is a one-dimensional conceptual temperature-index model, but includes a simplified energy budget solution for rain on snow conditions. The inputs to SNOW-17 are the three forcings for the elevation zone, and the output is a depth time series of rain on the relative portion of non-snow covered area plus melt from the relative portion of snow covered area. This areal average "rain + melt" depth series for the elevation zone is input to the SAC-SMA model. SAC-SMA is a one-dimensional lumped-parameter conceptual soil moisture model, in which subsurface storage in upper and lower soil layers is represented by a system of connected storage tanks spanning both layers. The resulting area-average runoff depth time series for each zone is then computed.

The runoff depth series for all elevation zones are then weighted by zone area to obtain the areal average runoff depth series for the subbasin. This series is then transformed, by a unit hydrograph representing the subbasin area, to obtain the simulated runoff hydrograph for the subbasin. At many locations in the CNRFC system, the unit hydrograph is used to change 6-hour SAC-SMA runoff depth series to a 1-hour streamflow series. This is helpful for calibrating to and forecasting peak flows in fast-responding watersheds.

If quality-checked streamgage data are available at the subbasin outlet, the observed hydrograph is merged with the simulated runoff hydrograph before routing downstream. Rather than an abrupt step from observed to simulated, a transition time, unique for each location, is used to blend the flows. The hydrologic forecaster compares simulated and observed and runoff hydrographs to assess hydrologic model performance, and when warranted, will make adjustments to model parameters, states, or related time series.

Components of the HEFS

The HEFS has two primary components: 1) the Meteorologic Ensemble Forecast Processor (MEFP), and 2) the Ensemble Postprocessor (EnsPost). Each of these also has a parameter estimator (PE) component: the MEFPPE and EnsPostPE, respectively. These components and their position in the forecast sequence are shown in **Figure 4**.

The MEFP is used to generate meteorologic ensemble forecasts of precipitation and temperature which display historically consistent bias correction and spread. Outside of the HEFS, these ensembles are used to force the hydrologic models, which results in streamflow forecast ensembles. At this point in the process, the streamflow ensembles will reflect any biases in the hydrologic models, and will display no uncertainty due to the hydrologic models.

Though not presently used at the CNRFC, the EnsPost is designed to be executed once the streamflow ensemble forecast has been created by the MEFP. The EnsPost adjusts the streamflow ensembles to reduce bias and incorporate uncertainty (spread) attributed to hydrologic initial states, parameters, and modeled processes. The resulting streamflow ensembles would then reflect minimal bias (both meteorological and hydrologic) and exhibit historically consistent spread reflecting meteorologic and hydrologic uncertainty. Note that the CNRFC does plan to test EnsPostPE and EnsPost components in the future, once further

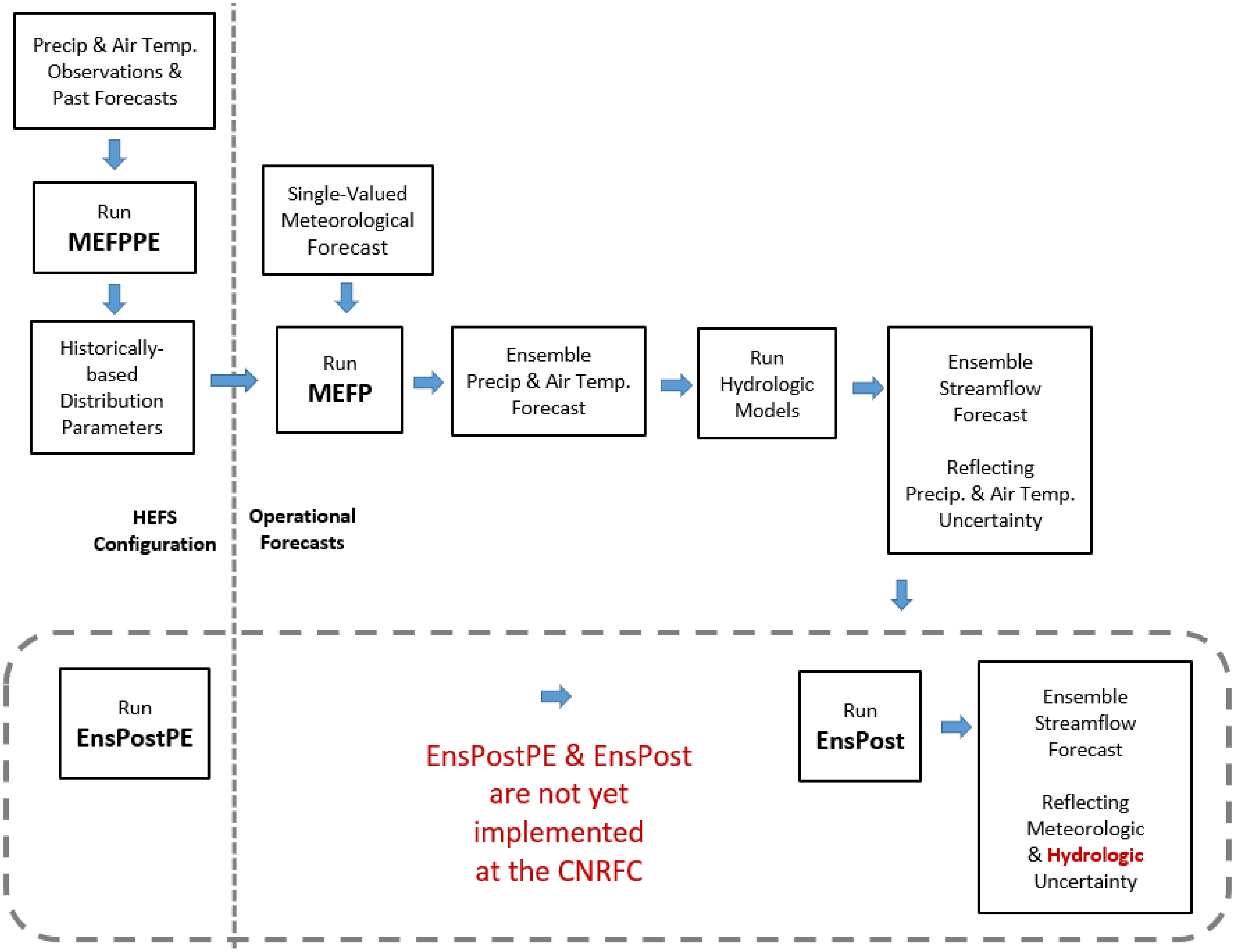


Figure 4 - Components of the HEFS as implemented at the CNRFC (May 2022)

refinements have been incorporated to the components.

MEFP Single-valued Inputs and Spatial Resolution

The CNRFC has configured the MEFP to use two single-valued current forecast series, one for precipitation and the other for air temperature, spanning forecast days 1 through 28. Each 6-hour single-valued series is generated from 6-hour forecast grids. Days 1 - 3 of the precipitation forecast are defined by the HAS QPF. All other portions of the MEFP single-valued inputs are defined by the mean of the GEFSv12 ensemble. **Table 3** lists the sources of these grids and spatial resolution.

Table 3 - Spatial Resolution of Forecast Grids

<table>
<tr><th rowspan="2">Forcing Type</th><th colspan="3">Forecast Day</th></tr>
<tr><th>1 - 3</th><th>4 - 10</th><th>11 - 28</th></tr>
<tr><td>Precipitation</td><td>HAS QPF
4.9 km x 4.9 km
(9.2 sq mi)</td><td>GEFSv12 mean
0.25 deg x 0.25 deg
(228 sq mi)</td><td>GEFSv12 mean
0.5 deg x 0.5 deg
(914 sq mi)</td></tr>
<tr><td>Air Temperature</td><td colspan="2">GEFSv12 mean
0.25 deg x 0.25 deg
(228 sq mi)</td><td>GEFSv12 mean
0.5 deg x 0.5 deg
(914 sq mi)</td></tr>
</table>

notes: 1 - length and area dimensions are nominal for Northern California region.
2 - HAS QPF is provided on the HRAP grid.

Six-hour values, areally-averaged over each subbasin elevation zone, are computed by weighting grid cell values contained by, or overlapping, each elevation zone. The result is a 6-hour single-valued meteorological forecast time series for each elevation zone. These are the MEFP single-valued inputs. **Figure 5** shows grid cell outlines overlaying CNRFC basins and basin zones for a region in the Northern Sierra. In this region, higher elevation subbasins are composed of two elevation zones, while lower elevation subbasins are defined by a single elevation zone. In this region, the 5,000 ft elevation contour separates lower and upper elevation zones.

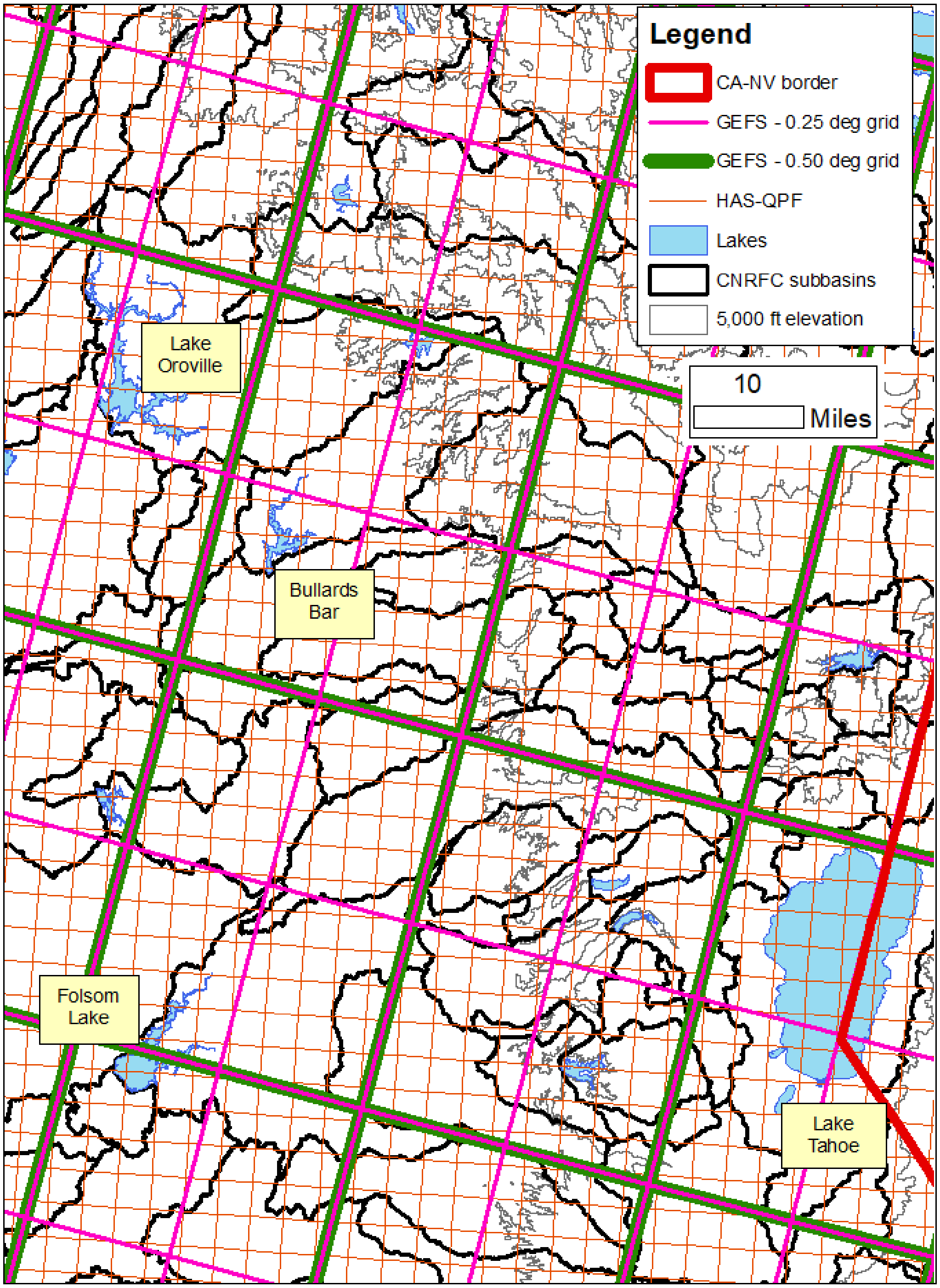


Figure 5 - Forecast Grids Overlaying CNRFC Basin Zones

4 Steps for Creating the CNRFC Ensemble Forecast

Overview of Method

The HEFS employs statistical postprocessing to develop ensembles of 6-hour time series of precipitation and air temperature (2 meters above ground surface). These ensembles, often referred to as the “hydrologic forcing ensembles”. Outside of HEFS, the ensembles are applied to the hydrologic models. A third forcing series, the rain/snow elevation, is also applied to the hydrologic models, but only as a single-valued series. The hydrologic models are then simulated, n times, with n being the number of members in each ensemble. Each simulation computes subbasin runoff, and also routes and combines streamflow hydrographs. This results in an ensemble of simulated streamflow hydrographs at each location in the hydrologic modeling system. At locations which are deemed by CNRFC as official ensemble forecast locations, the streamflow ensembles are post-processed to generate probabilistic forecast information in various formats.

Developing the ensemble forecasts begins with the MEFPPE, which computes statistical distribution parameters relating past forecast values to observations, for a collection of forecast time windows (canonical events). This is done only once as part of configuration of the HEFS. During operations, the MEFP requires as input the two 6-hour single-valued current forecast series, one of precipitation and one of air temperature, extracted from the 6-hour grids in **Table 3**. These two series must exist for every elevation zone in the system. For each canonical event, the MEFP extracts from the single-valued forecast series the value of the canonical event, and uses it as the *condition* for defining the distribution of observations associated with that forecast value. The conditional distribution is quickly generated on-the-fly, because the parameters which define the distribution have been previously stored by the MEFPPE. For each canonical event, the MEFP draws n samples from the corresponding conditional distributions, where n is the desired number of ensemble members to create. It is at this step in the process that bias in the single-valued forecast series is minimized, because the sampled distributions represent observations. The samples drawn from the distributions also reflect historically-consistent spread, because they are drawn using stratified sampling, and again representing observations. For both forcing types, the MEFP inputs the samples to a process known as the Schaake Shuffle, which is configured by the CNRFC to generate the meteorologic ensemble forecasts of elevation-zone precipitation and air temperature for forecast days 1 - 28. The Schaake Shuffle also leverages historical event patterns to generate ensembles that are spatially and temporally consistent across subbasins. The CNRFC uses these 28-day ensemble forecasts to generate ensemble streamflow forecasts at forecast times of T=12z, 18z, 00z, and 06z. For 12z forecasts, which are issued every day, the CNRFC appends ensemble forecast forcings for days 29 - 365, resulting in a 365-day ensemble forecast. The appended portion of the T=12z forecast is not generated by the HEFS, but is generated by the CNRFC using raw climatology.

The following sections describe how the MEFP component of the HEFS creates ensemble forecasts of air temperature and precipitation at the CNRFC. The presentation here is intended to provide the reader with a basic understanding of the process. For those seeking greater detail on the process, references are provided.

Preparatory steps performed once, as part of the HEFS configuration, include:

A. Collect Supporting Data
B. Define Canonical Events
C. Compute Statistical Parameters

Operational steps performed for each ensemble forecast include:

D. Extract Current Forecast Inputs
E. Generate Conditional Probability Surfaces
F. Generate and Sample Conditional Probability Distributions
G. Generate Forcings Ensembles (Schaake Shuffle)
H. Extend Ensembles to 365 days using Climatology
I. Apply Forcings Ensembles to Hydrologic Models

Steps A, B, & C - Preparation

A - Collect observed and past forecast data

Historical data used to calibrate the hydrologic models are listed in **Table 4**. The first three rows provide sources of the three observed forcings. As indicated in the third column, more than one source was needed to complete the period of record for precipitation and freezing level. CNRFC creates precipitation grids from GHCN gage data for the early portion of the precipitation record. For the later portion CNRFC uses the archived operational HAS QPE grids. Observed temperature grids from NOAA's Analysis of Record for Calibration (AORC) are used for the full calibration record. Freezing level is shown as "computed" because the majority of the record is based on NWP models. In 2019, CNRFC began using archived operational HAS QZE grids to extend the freezing level record. Observed streamflow data are not used by HEFS, but are listed for completeness as a needed data component for hydrologic models calibration.

Table 4 - Observed Data Sources

calibration por = wy 1980 - 2021

Forcing	Type	Source & POR
precipitation	observed	GHCN 10/1979 - 09/2003 QPE 10/2003 - 09/2021
temperature	observed	AORC 02/1979 - 09/2021
freezing level	computed	ERA5 01/1979 - 12/2018 GFS 01/2019 - 09/2019 QZE 10/2019 - 09/2021
streamflow	observed	USGS daily (primary source)

GHCN - Global Historical Climate Network
QPE - CNRFC Quantitative Precipitation Estimate
AORC - NOAA Analysis of Record for Calibration
ERA5 - ECMWF Reanalysis version 5
GFS - Global Forecast System
QZE - CNRFC Quantitative Freezing Level Estimate
USGS - United States Geological Survey

Table 5 lists sources of past forecast data for calibrating the MEFP. Past forecasts can be historical forecasts or hindcasts depending on data source. Hindcasts of short-term (days 1 - 3) precipitation are challenging to develop because the forecast process can be complex for active weather situations and involve expert judgment that cannot be replicated by a computer program. In order to address this challenge, CNRFC uses historical HAS QPF forecasts from water years 2010 - 2021 for the first 3 days of precipitation forecasts.

Table 5 - Sources of Past Forecast Data

Forecast days 1 - 3, por = wy 2010 - 2021

Forcing	Type	Source
precipitation	historical forecasts	HAS QPF

Forecast days 1 - 15, por = wy 1980 - 2020

Forcing	Type	Source
precipitation	hindcast	GEFS mean
temperature	hindcast	GEFS Tmin, Tmax

B - Define Canonical Events

Canonical events are a set of time windows within forecast days 1 - 28. The motivation for using canonical events is to capture the skill in the forecast precipitation and temperature at different temporal scales. The MEFPPE uses canonical events to compute statistical parameters relating past forecasts to observations. The MEFP uses canonical events to extract canonical event values from the current single-valued forecasts, and ultimately to generate the ensemble forecast series.

A unique set of canonical events is defined for precipitation, and another for air temperature. Each canonical event is defined by a duration (aggregation period) and a lead time relative to T0. The value assigned to a precipitation canonical event is simply the total precipitation amount during the event. For temperature canonical events, two values are of interest: a representative maximum value (Tmax) and a representative minimum value (Tmin). Tmax and Tmin are described further in E1.

There are two types of canonical events: base events and modulation events. Base events do not overlap and have no gaps in between. Modulation events span multiple base events and can overlap one another. At the CNRFC, precipitation is represented by 35 base and 8 modulation events. Temperature is represented by 14 base events and 0 modulation events. The total number of canonical events used to represent precipitation and temperature for the first 28 days of the ensemble forecast is therefore 57. The sequence and duration of these events are shown in **Table 6** for precipitation and in **Table 7** for temperature.

Table 6 - Canonical Events for Precipitation

Table 7 - Canonical Events for Temperature

Day of forecast	1	2	3	4	5	6	7	8	9	10	11	12	13	14	...
Base Events (numbered)	1	2	3	4	5	6	7	8	9	10	11		12		...

Day of forecast (continued)	... 15	16	17	18	19	20	21	22	23	24	25	26	27	28
Base Events (numbered)	... 13							14						

C - Compute Statistical Parameters

The HEFS uses statistical parameters to describe the relationships between past forecasts and observations. For temperature, the five parameters listed below are computed for each canonical event by the MEFPPE and saved for operational use by the MEFP.

μ_x = mean of observations
σ_x = standard deviation of observations
μ_y = mean of past forecasts
σ_y = standard deviation of past forecasts
γ = Pearson's product-moment correlation coefficient

Together, the 5 parameters can be used to define a joint probability surface, with the first pair of parameters defining the marginal distribution of observations and the second pair defining the marginal distribution of past forecasts. The fifth parameter defines how well observations are predicted by forecasts.

In order to compute the 5 parameters, the MEFPPE extracts data pairs of forecasts and observations (of precipitation and air temperature) for each available past forecast. For each past forecast, one data pair will be extracted for each canonical event. Next, the MEFPPE pools the extracted data pairs. The CNRFC has configured the MEFPPE to create a pool of data pairs for every 5th calendar day, using 61-day windows. Suppose the first day for which parameters are to be computed is January 1. Then for each canonical event, the MEFPPE pools all data pairs having forecast calendar days within plus or minus 30 days (a 61-day window) of January 1. If a past forecast is available for each day, then the number of data pairs for each canonical event will be equal to 61 times the number of years of past forecasts (11 for HAS QPF, 41 for GEFS). For each canonical event, the 5 parameters are then computed and saved. The MEFPPE then advances the 61-day time window by 5 days, and repeats the process to compute and store the 5 parameters associated with January 6. The process is repeated until the 5 parameters have been computed and stored for every 5th day of the calendar year. Operationally, the MEFP adopts the parameter set corresponding to the "5th day" nearest to the current forecast day.

Steps D, E, F, G, & H - Operational Forecasts with HEFS

D - Extract Current Forecast Inputs

For each basin elevation zone, three 6-hour single-valued series representing the current forecast are required by the MEFP: precipitation, maximum air temperature, and minimum air temperature. These series are created by extracting 6-hour values from forecast grids. Sources for the current single-valued forecast series are listed in **Table 5**. Below, **Table 8** shows for precipitation the temporal position of base (grey shading) and modulation (blue or yellow shading) canonical events. **Table 9** shows for maximum and minimum temperature the temporal position of base events (grey shading). For temperature there are no modulation events. For each canonical event, the MEFP extracts the values for all canonical events from the 6-hour single-valued input series. Canonical event values for precipitation are simply the total precipitation during the event window. Computation of canonical event values of temperature is described in E1. Note in **Tables 8** and **9** that each breaks the 28-day MEFP forecast period into days 1 - 14 in the upper half and days 15 - 28 in the lower half.

Table 8 - Current Forecast Canonical Inputs - Precipitation

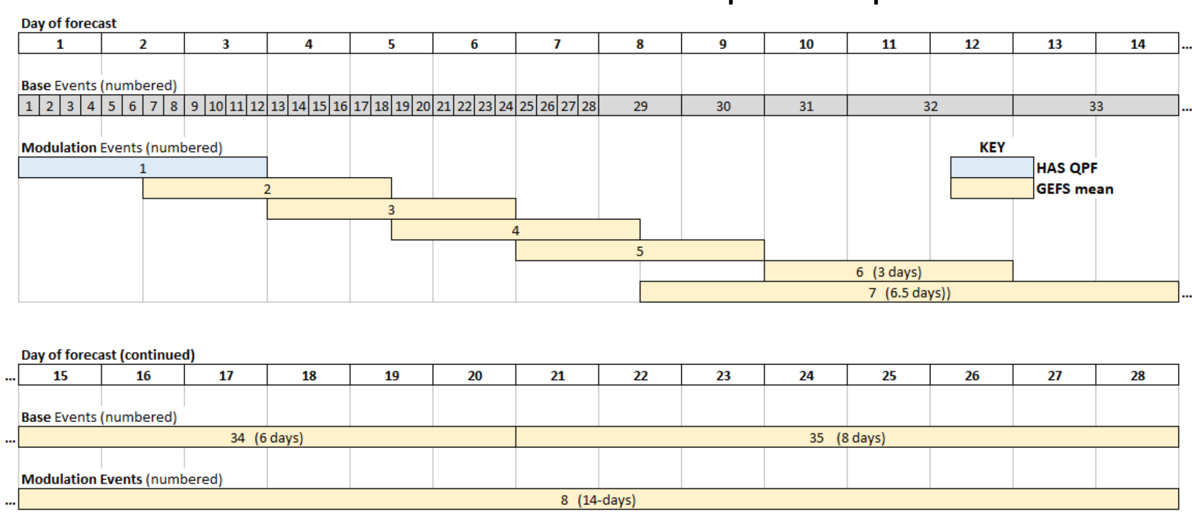
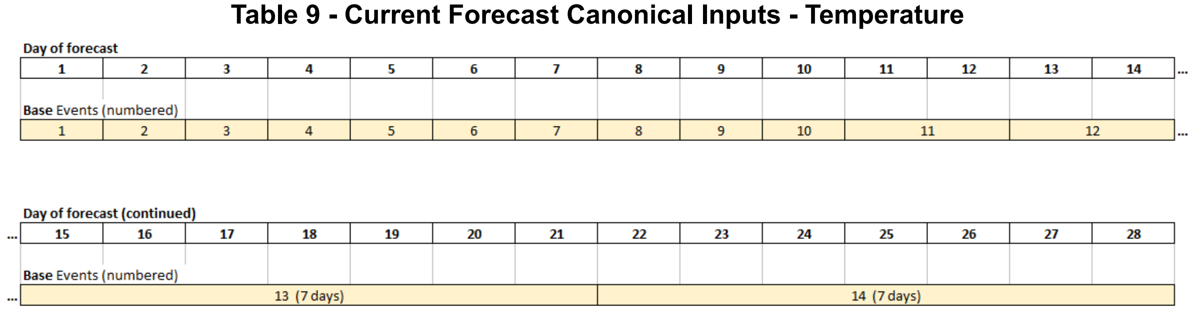


Table 9 - Current Forecast Canonical Inputs - Temperature



It is critical that past forecast sources used by the MEFPPE to compute statistical parameters are consistent with current forecast sources. This is because the MEFP will use the statistical parameters to describe the bias and uncertainty associated with the current forecast canonical values. **Table 10** shows that past forecast sources (**Table 5**) and current forecast sources (**Tables 8** and **9**) are consistent.

Table 10 - Consistency of Forecast Sources

		Source	
Forcing	**Forecast Days**	**Past Forecasts**	**Current Forecasts**
Precipitation	1 - 3	HAS QPF	HAS QPF
Precipitation	4 - 28	GEFS Mean	GEFS Mean
Temperature	1 - 28	GEFS Tmin, Tmax	GEFS Tmin, Tmax

E - Generate Joint Probability Distributions and Conditional Probability Surfaces

In this document, the term "conditional probability surface" refers to the mathematical surface that is obtained when a joint probability surface (observations vs. forecasts) is divided by the marginal distribution of forecasts. The resulting surface, when sliced at a forecast value, always yields a conditional probability distribution with area of 1.0. The shape of the surface reveals the effect that the magnitude of the forecast has on the conditional distribution of observations.

A detailed mathematical description of the procedures implemented by MEFP is provided by Herr and Krzysztofowicz (2005) and Lu et al (2011). This section applies parts of that procedure to example data sets of temperature and precipitation, and intentionally tries to avoid the more advanced topics in the reference paper. The description here is intended to visually illustrate how conditional probability surfaces for temperature and precipitation are created. In the next section, the process by which these surfaces are used by MEFP during forecast operations is described.

E1 - Temperature

As configured by the CNRFC, temperature canonical events can consist of 24-hour (12z to 12z), or n x 24-hour time windows. For each canonical event, representative maximum and minimum temperature values (Tmax and Tmin) are of interest. For forecasts, Tmax and Tmin are instantaneous values. For observations, Tmax and Tmin are 6-hour average values for periods 18z-00z and 06z-12z. This is because, as configured at CNRFC, instantaneous forecast values are used from forecasts (past and current), while hydrologic models are forced with 6-hour average temperature time series.

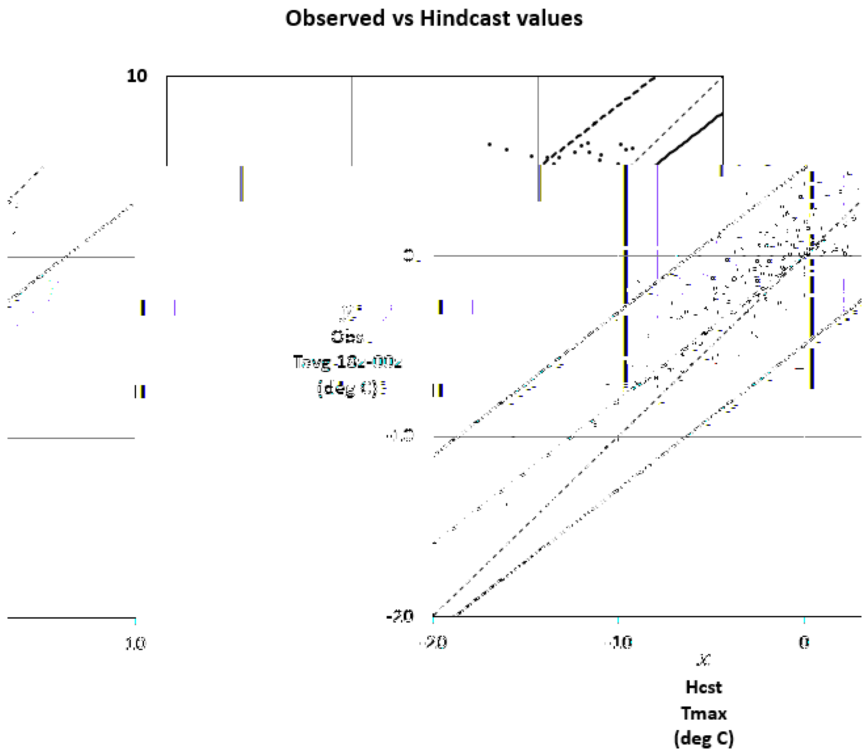


Figure 6 - Data Pairs and Confidence Intervals (Tmax)

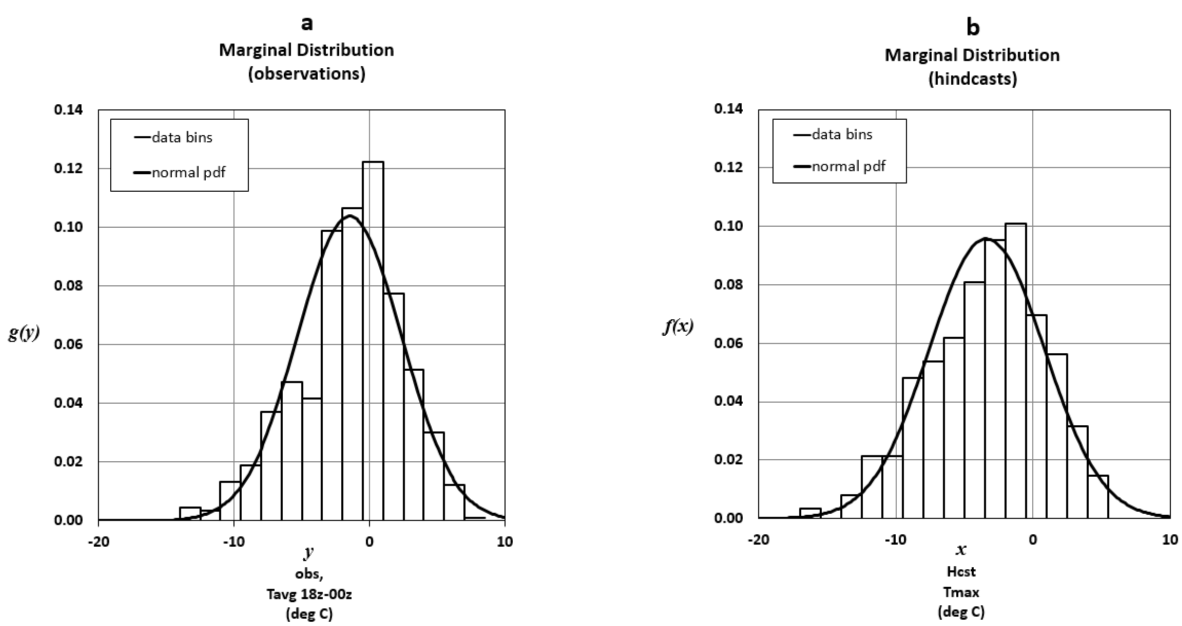


Figure 7 - Marginal Distributions and Data Bins (Tmax)

For 24-hour canonical events, all 6-hour observed values (18z-00z and 06z-12z) and instantaneous forecast values are computed from the 24-hour period. For n x 24-hour canonical events, all 6-hour observed values (18z-00z and 06z-12z) and instantaneous forecast values within the canonical event are set to the averages of values computed for each 24-hour period within the canonical event.

As a result, two joint probability distributions are needed:

1. One relating observed 18z-00z Tavg to forecast Tmax, and
2. One relating observed 06z-12z Tavg to forecast Tmin

Marginal Distributions

In this example, a 1-day canonical event temperature hindcast data set of 594 data pairs for the upper basin of North Fork Dam of the American River is plotted in **Figure 6**. The extracted data are for December 27 (12z to 12z). Each data pair consists of an x and y value in which x is the past forecast instantaneous maximum value Tmax, and y is the observed 6-hour average value Tavg for 18z - 00z. A normal distribution is fit to each sample group to obtain the two resulting marginal distributions, $f(x)$ and $g(y)$. These are shown in **Figures 7a and 7b**. Data bins are also shown for comparison.

Joint Probability Distribution

For one canonical event, the bivariate normal density function (**Equation 1**) is used to estimate the joint probability distribution $h(x,y)$ of observations (y) and past forecast canonical event inputs (x). $h(x,y)$ is completely defined by the four marginal parameters: μ_x , σ_x , μ_y , σ_y , and correlation coefficient γ. The five parameters are computed directly from the samples.

Equation 1

$$h(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\gamma^2}} exp\left[\frac{1}{2\sqrt{1-\gamma^2}}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 + 2\gamma\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right]$$

in which:

$h(x,y)$ = joint probability distribution (in this case bivariate normal)
μ_x = sample mean of past forecasts (x)
σ_x = sample standard deviation of past forecasts (x)
μ_y = sample mean of observations (y)
σ_y = sample standard deviation of observations (y)
γ = Pearson's product-moment correlation coefficient

Computed values of the 5 parameters for this example are μ_x = -3.37, σ_x = 4.17, μ_y = -1.48, σ_y = 3.84, and γ = 0.80. The resulting joint probability surface, $h(x,y)$, is shown in **Figure 8**:

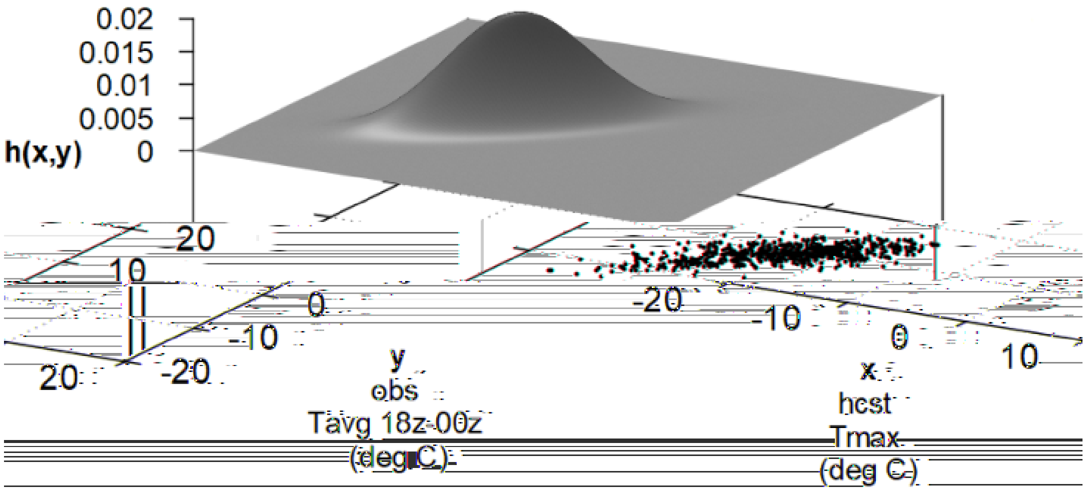


Figure 8 - Joint Probability Surface (Tmax)

A useful property of the bivariate normal distribution is that the dependence structure between x and y is linear and requires computation of only one additional parameter, γ, from the data. As shown by **Equations 2a** and **2b**, the conditional mean value of y is a straight line, and the conditional standard deviation of y is a constant. These properties represent the change in distribution of observations with respect to forecast value for one canonical event.

Equations 2a and 2b

$$(x - \mu_x)$$

$$\cdot \sqrt{1-\gamma^2} \qquad \sigma_{y|x=x_o} = \sigma_y$$

in which:

$\mu_{y|x=x_o}$ = mean of observed values y given that forecast $x = x_o$,

$\sigma_{y|x=x_o}$ = standard deviation of observed values y given that forecast $x = x_o$.

Conditional Probability Surface

While the joint probability surface is informative, it is the conditional probability surface, given by **Equation 3**, that we are most interested in.

Equation 3

$$h(y|x) = \frac{h(x,y)}{f(x)},$$

In **Equation 3**, $f(x)$ is the marginal distribution of past forecast samples. We are in effect normalizing the joint probability surface to the marginal distribution of past forecasts. The resulting conditional probability surface for this example is shown in **Figure 9**.

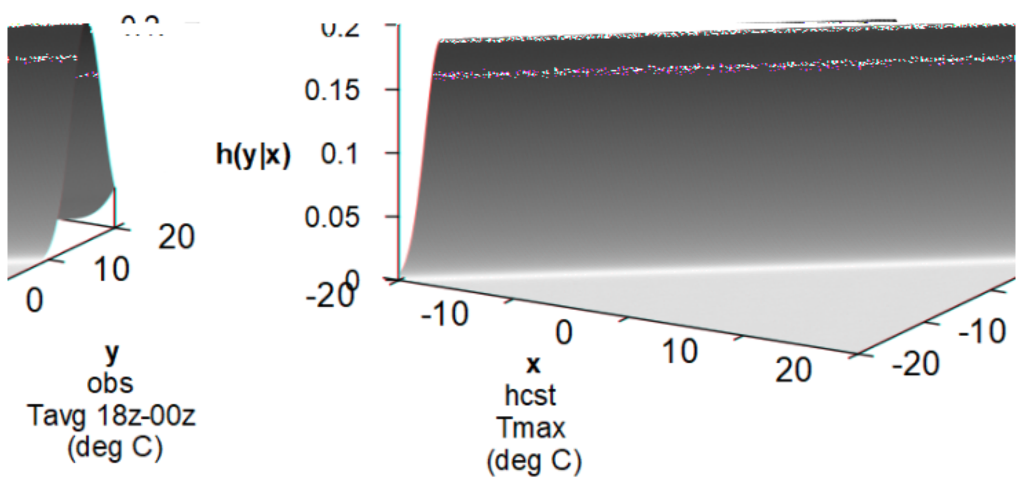


Figure 9 - Conditional Probability Surface (Tmax)

This surface clearly reflects the dependence structure defined by **Equations 2a and 2b** (see confidence intervals in **Figure 6**). When this surface is "sliced" with a y-z plane at a selected x value, the resulting intersection is a normal distribution. If a different value of x is chosen for the slice, the same distribution will result, but will be shifted in the y direction. Note that when the surface was developed, the x axis represented the *past* forecast canonical input value, but when used operationally, the x axis represents the *current* forecast canonical input value. For this reason, consistency between current and past forecast sources is essential, as indicated in **Table 10**.

For the canonical event represented by **Figure 9**, if the current forecast canonical input value were a maximum temperature (Tmax) of 10 deg C, then the conditional normal distribution associated with that value defines the corresponding range of uncertainty of the 18z-00z average temperature (Tavg) forecast. Following the above steps, another surface for the Tmin case would also be developed from data pairs of Tmin and 06z-12z Tavg.

E2 - Precipitation

Unlike temperature, precipitation is intermittent. Properly accounting for the probability of both rain and non-rain conditions is a mixed distribution problem. The MEFP provides two options for computing mixed distribution statistics: 1) Explicit Precipitation Intermittency Treatment (EPT) described in Herr and Krzysztofowicz (2005) and Wu et al. (2011), and 2) Implicit Treatment (IPT) described in Wu et al. (2011). For short lead times the differences between IPT and EPT methods are small, while EPT is much more skillful for longer lead times. For this reason, the EPT method is used at the CNRFC.

For simplicity, this document does not attempt to describe the details of mixed distribution computations, but instead focuses on the "wet-wet" case. In this example, the CNRFC configured a value of 0.254 mm as the minimum 3-day precipitation value considered non-zero.

Marginal Distributions

An example data set of 952 wet-wet data pairs for the upper basin of North Fork Dam of the American River is plotted in **Figure 10**. The extracted data are for December 7 (center of 61-day extraction window). Each data pair reflects precipitation over a 3-day time period starting at T0.

The CNRFC configures MEFP to use the gamma distribution to describe each of the two marginal distributions: $f(x)$ for forecast values (x), and $g(y)$ for observations (y). Each gamma distribution is defined by two parameters: the shape factor (α) and the scale factor (β). The density function form of the gamma distribution $f(x)$ is shown in **Equation 4**.

Equation 4

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \cdot x^{\alpha-1} e^{-\frac{x}{\beta}}$$

in which:
x = forecast value
α = shape factor
β = scale factor
$\Gamma(\alpha)$ = gamma function

Estimates of α and β are computed from the sample mean (μ) and sample standard deviation (σ), as shown in **Equations 5a** and **5b**. Corresponding values of α_y and β_y are also computed for $g(y)$.

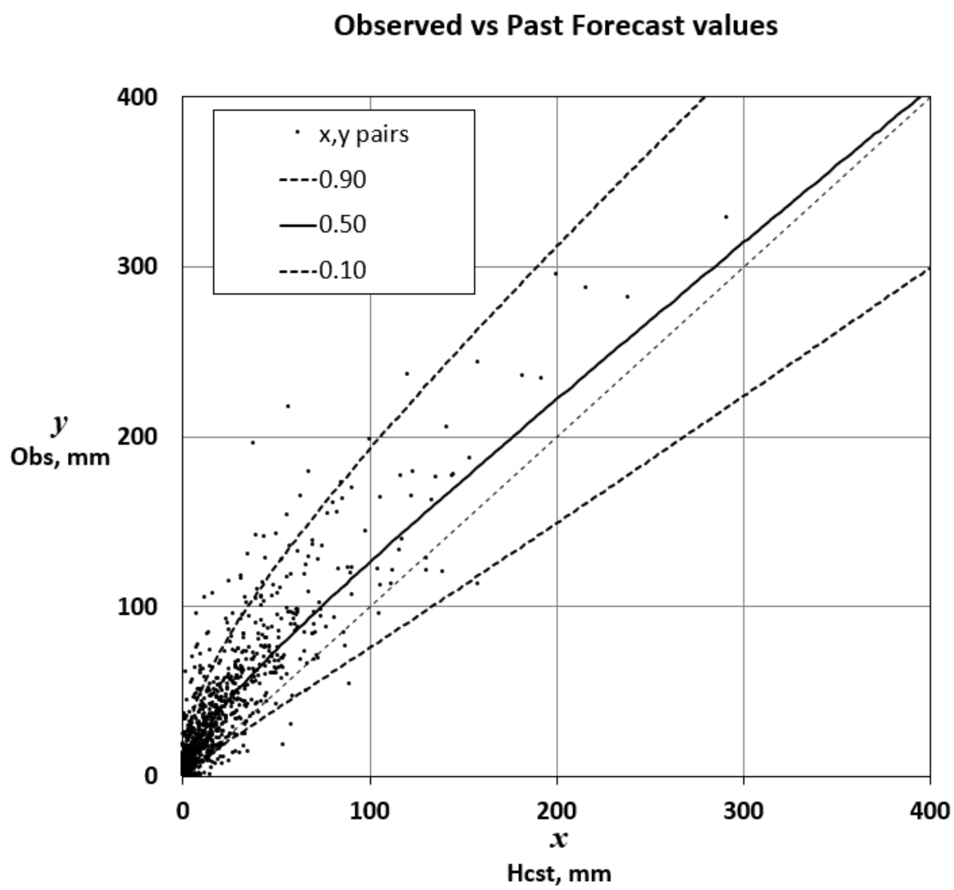


Figure 10 - Data Pairs and Confidence Intervals of "wet-wet" Precipitation)

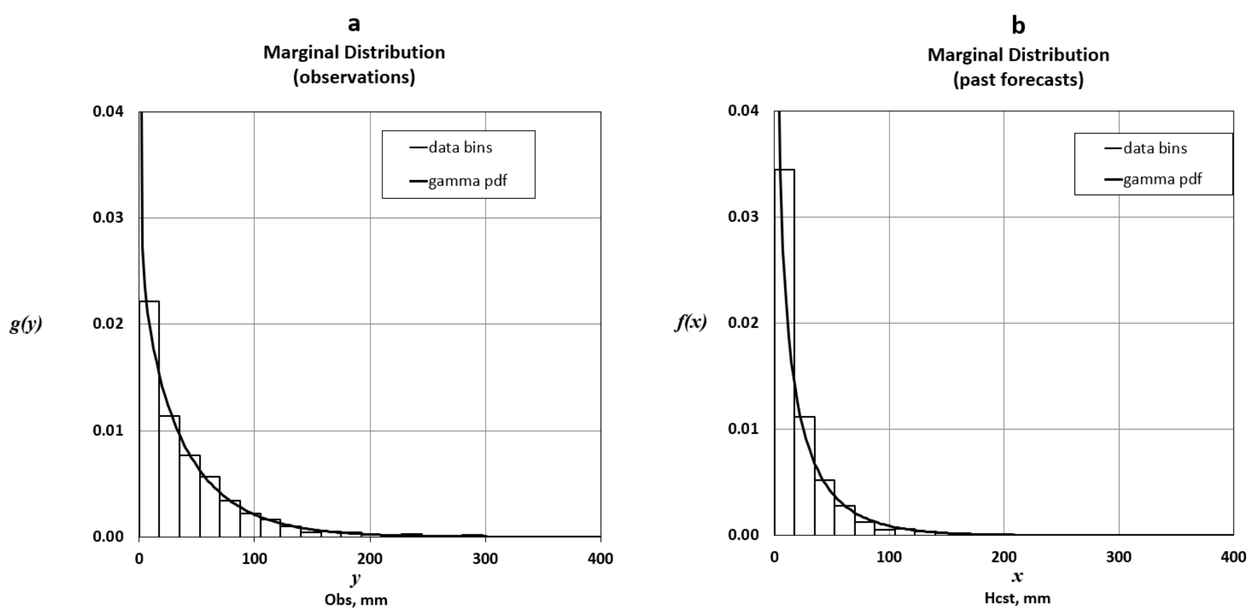


Figure 11 - Marginal Distributions of "wet-wet" Precipitation

Equations 5a and 5b

$$\alpha_x = \frac{\mu_x}{\beta_x} \quad \text{= shape factor for } f(x)$$

$$\beta_x = \frac{{\sigma_x}^2}{\mu_x} \quad \text{= scale factor for } f(x)$$

The resulting two marginal distributions, $f(x)$ and $g(y)$, are shown in **Figures 11a** and **11b**. Data bins are also shown for comparison. The computed Pearson's product-moment coefficient (γ) is 0.882 (in x-y space).

Joint Probability Distribution

If precipitation data were fairly normally distributed, then the same approach as for temperature could be taken for generating the joint probability surface $h(x, y)$ and conditional probability surface $h(y|x)$. Instead, the HEFS uses the normal quantile transform (NQT) to transform the forecast (x) and observed (y) values into more normally distributed data sets of u and v. The NQT iis defined by **Equations 6a** and **6b**.

Equations 6a and 6b

$$u = NQT(x) = Q^{-1}(F(x)) \quad \text{and} \quad v = NQT(y) = Q^{-1}(G(y))$$

in which

$NQT()$ = normal quantile transform
$Q^{-1}()$ = inverse of cumulative standard normal distribution
$F(x)$ = cumulative marginal distribution of x
$G(y)$ = cumulative marginal distribution of y
x = forecast value
y = observed value

u-v data pairs are plotted in **Figure 12**, and marginal distributions of u and v are plotted in **Figures 13a** and **13b**.

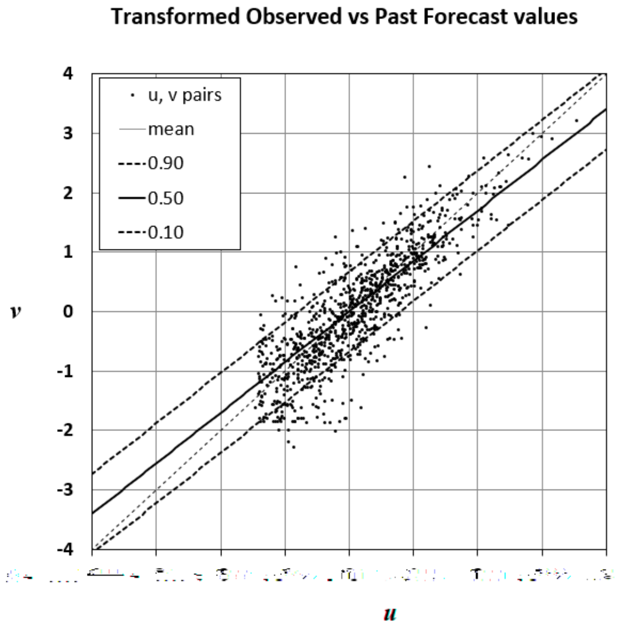


Figure 12 - Data Pairs and Confidence Intervals of Transformed "wet-wet" Precipitation

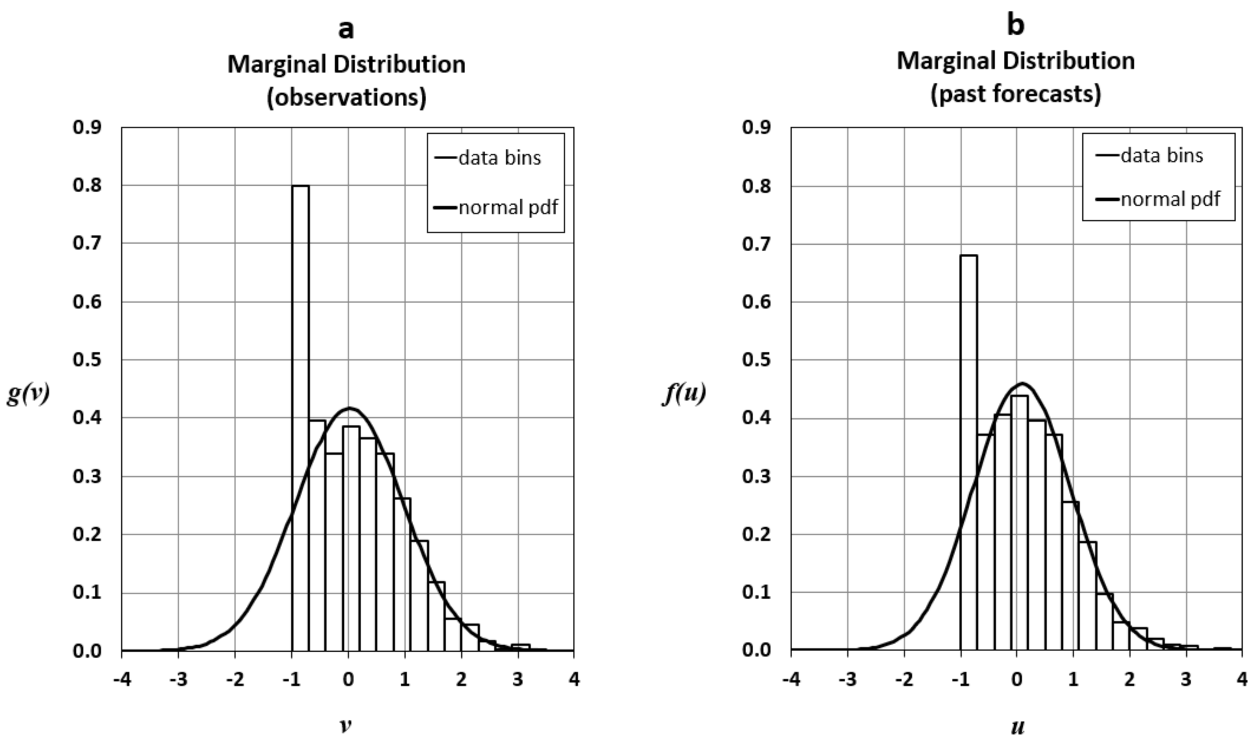


Figure 13 - Marginal Distributions of Transformed "wet-wet" Precipitation

The bivariate standard normal distribution (**Equation 7**) is then fit to the collection of u-v data pairs. The distribution has only one parameter, Pearson's product-moment correlation coefficient, γ, computed from transformed values. For this example γ = 0.851.

Equation 7

$$h(u,v) = \frac{1}{2\pi\sqrt{1-\gamma^2}} \cdot exp\left[-\frac{u^2+v^2-2\gamma uv}{2(1-\gamma^2)}\right]$$

The resulting bivariate standard normal distribution for this example is shown in **Figure 12**.

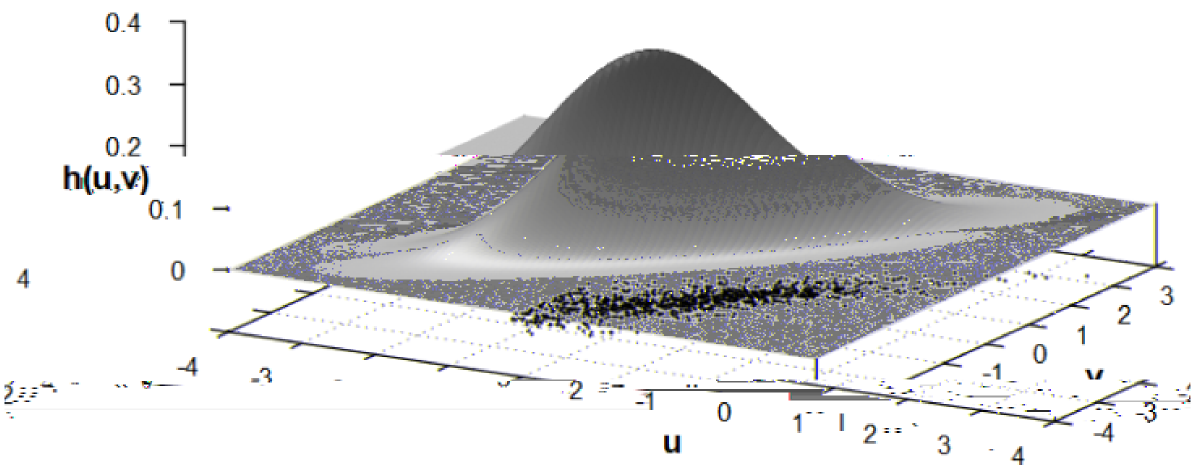


Figure 14 - Bivariate Standard Normal Surface

The linear dependence structure between v as a function of u is given by **Equations 8a** and **8b**, which are analogs of **Equations 2a** and **2b**. The only parameter required is γ, computed from transformed data values.

Equations 8a and 8b

$$\mu_{v|u=u_o} = \gamma \cdot u$$

$$\sigma_{v|u=u_o} = \sqrt{1-\gamma^2}$$

in which:

$\mu_{v|u=u_o}$ = mean of observed values v given that forecast $u = u_o$

$\sigma_{v|u=u_o}$ = standard deviation of observed values v given that forecast $u = u_o$

The surface $h(u, v)$ is then back-transformed from u-v space to x-y space to obtain the bivariate meta-Gaussian distribution $h(x, y)$. This is done using **Equation 9** below. which is equation 19 in Herr and Krzysztofowicz (2005). In so doing, the original marginal distributions are enforced, and the dependence structure, linear in the u-v domain, becomes non-linear in the x-y domain.

Equation 9

$$h(x, y) = \frac{f(x)g(y)}{\sqrt{1-\gamma^2}q\left(Q^{-1}\big(G(y)\big)\right)} \times q\left(\frac{Q^{-1}\big(G(y)\big) - \gamma Q^{-1}\big(F(x)\big)}{\sqrt{1-\gamma^2}}\right)$$

in which:

$h(x, y)$ = joint probability distribution
$mg(x, y)$ = bivariate meta-Gaussian distribution
$q(\,)$ = standard normal density function
$Q^{-1}(\,)$ = inverse of cumulative standard normal distribution function
$F(x)$ = cumulative marginal distribution of forecast value (x)
$G(y)$ = cumulative marginal distribution of observed value (y)
$f(x)$ = marginal probability density function of forecast value (x)
$g(y)$ = marginal probability density function of observed value(y)

In **Equation 9**, as the gamma functions $f(x)$ and $g(y)$ are each defined by shape and scale parameters, the total number of parameters required to define $h(x, y)$ is five: α_x, β_x, α_y, β_y, and γ. The computed values of these parameters for this example are: α_x = 0.54, β_x = 41.6, α_y = 0.86, β_y = 47.3, and Υ = 0.882 (0.851 in u-v space). The resulting surface, $h(x, y)$, is shown in **Figure 15**.

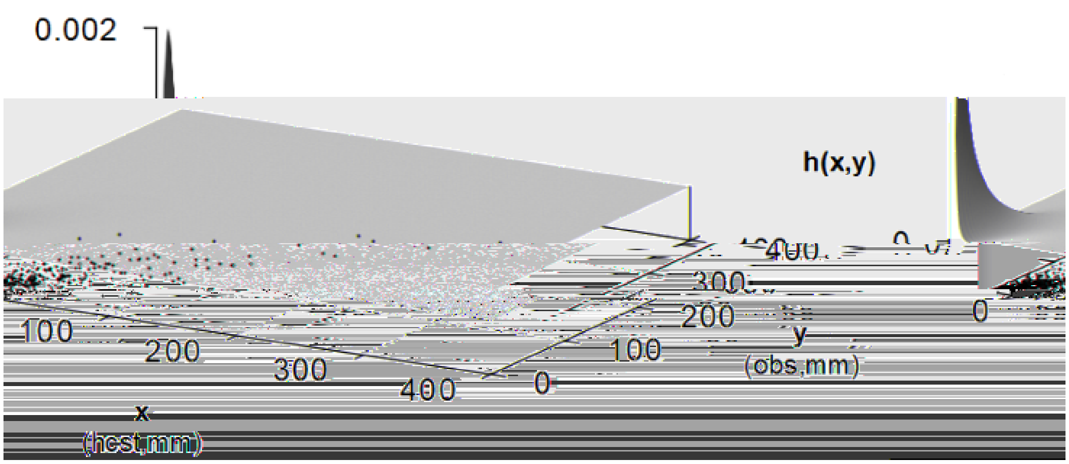


Figure 15 - Joint Probability Surface (Precipitation)

The conditional probability surface $h(y|x)$ for the canonical event is obtained from **Equation 3**.

Equation 3 (repeated)

$$h(y|x) = \frac{h(x,y)}{f(x)}$$

Equation **3** for this example is shown in **Figure 16.**

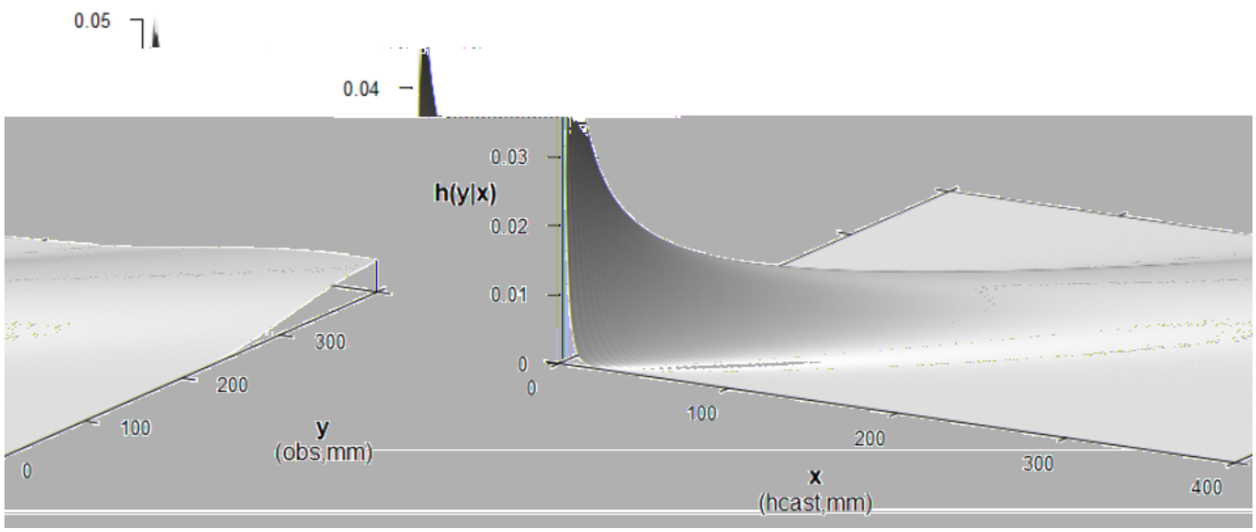


Figure 16 - Conditional Probability Surface (Precipitation)

This surface reflects the back-transformed dependence structure defined by **Equations 8a and 8b** (confidence intervals are shown **Figure 10**). Note that when the surface was developed, the x axis represented the *past* forecast value, but when used operationally, the x axis represents the *current* forecast value. For this reason, consistency between current and past forecast sources is essential, as indicated in **Table 10**.

F - Generate and Sample Conditional Probability Distributions

With conditional probability surfaces defined (**Figure 9**, **Figure 16**), MEFP "slices" each surface at the corresponding current forecast canonical input value. The slice reveals the conditional distribution of observations associated with the current forecast value. The conditional distribution is inherently bias-corrected and reflects historically consistent spread, because it is a distribution of observations (not past forecasts).

When sampling the conditional probability distributions (slices), the cumulative form of the distribution, $H_{Y|X}(y|x)$ is used. $H_{Y|X}(y|x)$ is given by **Equation 10**, which is Equation 20 in Herr and Krzysztofowicz (2005).

Equation 10

$$H_{Y|X}(y|x) \;=\; Q\left(\frac{Q^{-1}(G(y)) - \gamma Q^{-1}(F(x))}{\sqrt{1-\gamma^2}}\right)$$

in which:

$H_{Y|X}(y|x)$= cumulative conditional probability distribution
$Q(\ )$ = cumulative standard normal distribution function
$Q^{-1}(\ )$ = inverse of cumulative standard normal distribution function
$F(x)$ = cumulative (gamma) distribution of forecast value (x)
$G(y)$ = cumulative (gamma) distribution of observed value (y)

Sampling is done using equally spaced increments of cumulative probability along the vertical H(y|x) axis. The values of non-exceedance probability (NEP) used for sampling are given by Weibull plotting plotting positions, which are defined by :

Equation 11

$$NEP = r\,/\,(n + 1)$$

where: NEP = non-exceedance probability (Weibull plotting position)
r = rank of sample (from largest to smallest)
n = sample size

Note that the maximum value of NEP computed by NEP is less than 1.0 and the minimum value of NEP is greater than 0.

F1 - Precipitation

Using the conditional probability surface developed for precipitation in the previous section, **Figure 17a** shows the conditional probability surface for precipitation from **Figure 16**, with two colored curves indicating slicing the surface at two hypothetical current forecast canonical input values of 25 and 200 mm on the x axis. The resulting magenta curve is the conditional probability distribution given a current forecast value of 25 mm, and the resulting blue curve is the conditional probability distribution given a current forecast value of 200 mm. The resulting difference in mean and spread of the conditional distributions is seen in **Figure 17b**. **Figure 17c** shows the cumulative form, H(y|x), of the same distributions. Note that during operations, only one current forecast canonical input value, and therefore only one conditional probability distribution, is produced for each canonical event.

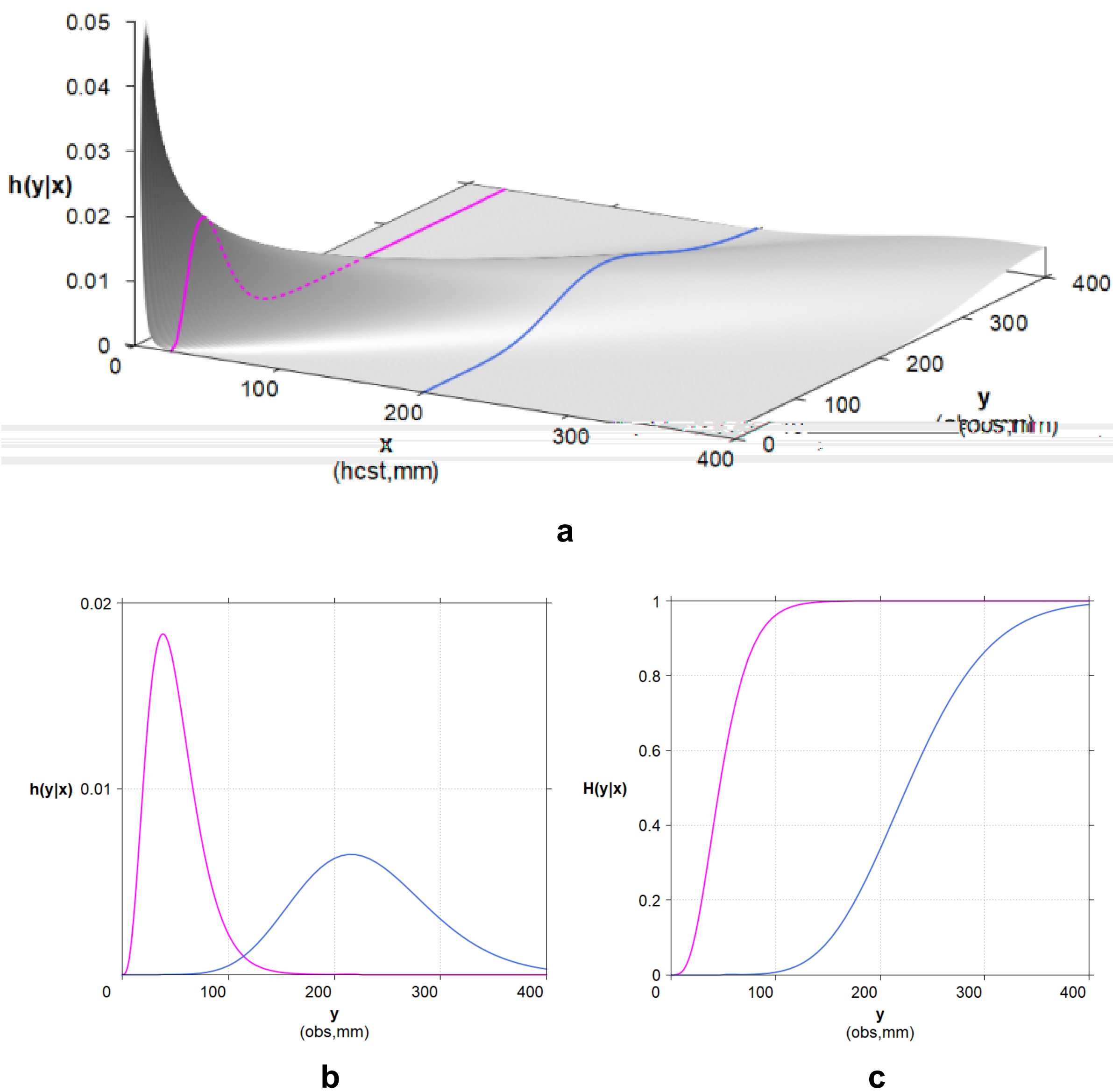
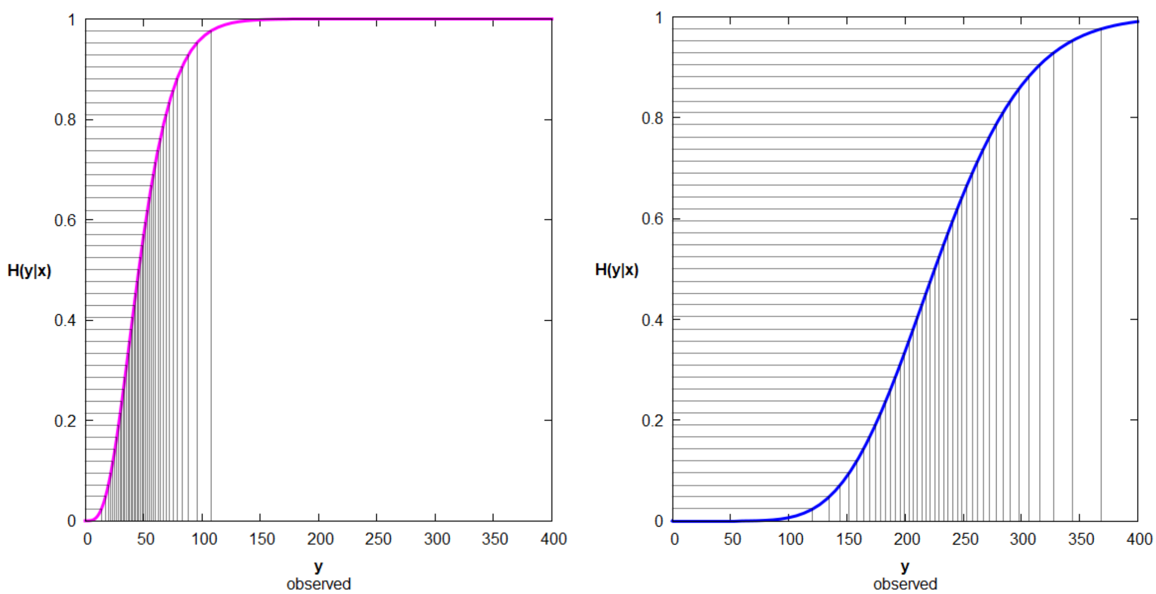


Figure 17 - Conditional Probability Distributions for One Canonical Event (precipitation)

At the CNRFC, the MEFP generates ensemble forcings spanning forecast days 1 - 28, for all forecast times. For the morning (T0=12z) forecast only, the CNRFC also computes ensemble members from raw climatology for forecast days 29 - 365). In creating the raw climatology portion of the ensemble forecast, the number of ensemble members is 41. In order to have the MEFP generate the same number ensemble members for forecast days 1 - 28, 41 samples are drawn from the cumulative conditional probability distribution for each canonical event.

The samples must be drawn so as to be unbiased, in order that the resulting distribution of samples is reflective of the original continuous distribution. It is also desirable that the sampling method is repeatable. To satisfy these requirements, the MEFP draws samples from the cumulative form of the conditional probability distribution. **Figures 18a and 18b** illustrate drawing 41 samples from cumulative probability distributions corresponding to hypothetical current forecast canonical input values of 25 mm and 200 mm. The resulting sample values are indicated on the horizontal axes.



Forecast = 25 mm | Forecast = 200 mm

a | **b**

Figure 18 - Sampling of Conditional Probability Distributions (precipitation)

F2 - Temperature

The process for generating conditional probability distributions for temperature is fundamentally the same as for precipitation, but for each canonical event two surfaces are required:

- Tmax forecast vs 18z-00z Tavg observed (example shown in **Figure 9**).
- Tmin forecast vs 06z-12z Tavg observed.

Operationally, MEFP would "slice" each surface at their respective forecast values, and two samples drawn by stratified sampling: one sample of 41 Tmin values and one sample of 41 Tmax values.

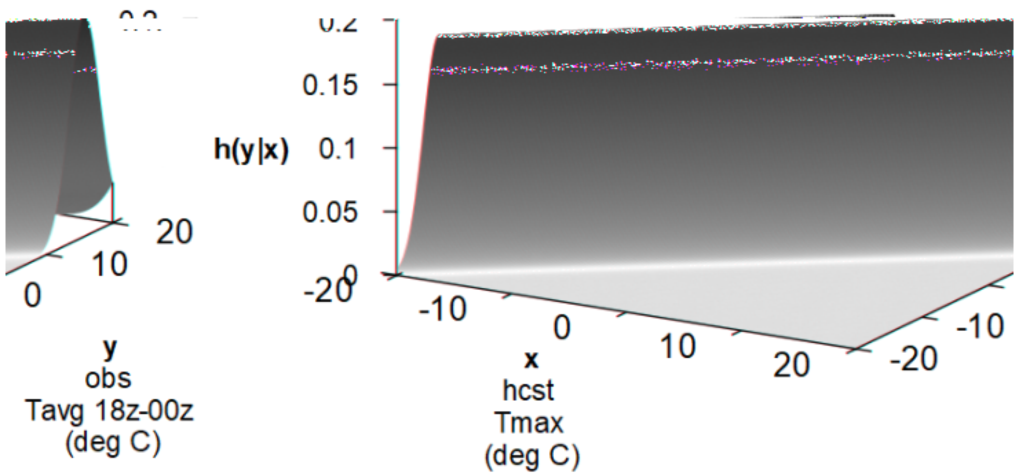


Figure 9 (repeated) - Conditional Probability Surface (Tmax)

G - Create Hydrologic Ensemble Forcings (Schaake Shuffle)

Samples drawn from the conditional probability distributions for all canonical events are input to a procedure known as the Schaake Shuffle. The procedure is a simple and efficient method used to preserve the space-time statistical properties of climatology among multiple hydro-meteorological variables across multiple forecast locations for ensemble forecasting. For this application at CNRFC, once MEFP has drawn the 41 samples from each conditional probability distribution for each canonical event for a forcing type, the Schaake Shuffle transforms the samples into a 41-member ensemble spanning 28 days.

G1 - Precipitation

Beginning with precipitation, it is useful to organize the sample as shown in **Table 11**. Note that column width does not reflect canonical event duration. The top row provides an arbitrary id number for each canonical event. Precipitation is represented by 35 base events and 8 modulation events. The first column contains an arbitrary number id for each sample. Each column represents the individual sample values drawn from the conditional probability distribution for the canonical event represented by the column. Data values, shown as “b” and “m”, represent base event and modulation event sample values. The last row contains the correlation coefficients (between observed and past forecast samples) for each canonical event. These values are indicated by “c”. The canonical event values represent accumulated precipitation during each canonical event.

Table 11 - Precipitation Samples for Schaake Shuffle (Forecast Days 1 - 28)

	Precipitation Base Events																																			Precipitation Modulation Events							
Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	1	2	3	4	5	6	7	8
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	m	m	m	m	m	m	m	m
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	m	m	m	m	m	m	m	m
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	m	m	m	m	m	m	m	m
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	m	m	m	m	m	m	m	m
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
41	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b	m	m	m	m	m	m	m	m
Corr. Coeff.	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c

The Schaake Shuffle ranks base events extracted from the 41-year historical record, and assigns year labels to the base events of corresponding rank. A detailed example of the Schaake Shuffle applied to precipitation is provided in the attachment “Schaake Shuffle Step-by-Step Example”. If there were no modulation events, each set of base events having matching year labels would be merged to form the set of ensemble members (this is how it is done for temperature). However, precipitation has modulation events, which overlap base events and sometimes other modulation events. To handle this, the Schaake Shuffle considers all (base and modulation) events, in order from lowest correlation to highest. Typically, modulation events will have higher correlation, and will apply after base events, thus “modulating” the base events. When a modulation event is considered, base events contained by the modulation event are adjusted to be consistent with the modulation event.

G2 - Temperature

As described in E1 - Temperature, conditional probability distributions that relate Tmax forecasts and Tmin forecasts to Tavg (18z to 00z) and Tavg (06z to 12z) observations respectively, for each day of the forecast, are used to construct the temperature ensemble. For the two 6-hour periods not addressed (00z to 06z, 12z to 18z), the MEFP interpolates to fill gaps. This is shown in Table **12**.

Table 12 - HEFS Processing of Temperature (Tmin and Tmax)

Time	MEFPPE		MEFP	
Time Period	Observed	Past Forecast	Current Forecast	Source of Ensemble Member 6-hour Tavg value
12z to 18z	--	--	--	interpolation
18z to 00z	Tavg	Tmax	Tmax	sample from distribution
00z to 06z	--	--	--	interpolation
06z to 12z	Tavg	Tmin	Tmin	sample from distribution

Temperature is represented by 14 base events, and no modulation events. Each base event is a 1-day period from which the Tmin and Tmax values are extracted from past forecasts and the current forecast. Samples of 6-hour Tavg are drawn from the corresponding conditional probability distributions. **Tables 13a** and **13b** show how the sampled values can be organized, with "b" and "c" indicating base event sample and correlation values respectively.

Tables 13a and 13b - Temperature Samples for Schaake Shuffle (Forecast Days 1 - 28)

13a

Tavg (06z - 12z)
Base Events

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
41	b	b	b	b	b	b	b	b	b	b	b	b	b	b
Corr. Coeff.	c	c	c	c	c	c	c	c	c	c	c	c	c	c

13b

Tavg (18z - 00z)
Base Events

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	b	b	b	b	b	b	b	b	b	b	b	b	b	b
2	b	b	b	b	b	b	b	b	b	b	b	b	b	b
3	b	b	b	b	b	b	b	b	b	b	b	b	b	b
4	b	b	b	b	b	b	b	b	b	b	b	b	b	b
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
41	b	b	b	b	b	b	b	b	b	b	b	b	b	b
Corr. Coeff.	c	c	c	c	c	c	c	c	c	c	c	c	c	c

The Schaake Shuffle treats the values in **Tables 13a** and **13b** completely separately. With **Table 13a**, base events are extracted from the 41-year historical record, and year labels are assigned to the base events of corresponding rank. There are no modulation events, so no further adjusting of values is necessary. The process is repeated for the values in **Table 13b**. The temperature ensembles are created by merging (alternating 06z-12z and 18z-00z) Tavg values having the same year label, and then interpolating in time between those values to obtain the missing 00z-06z and 12z-18z Tavg values. Each resulting set of values having matching year labels is an ensemble member.

H - Extend Ensembles to 365 days using Climatology

For the CNRFC morning forecast (T0 = 12z), the ensemble is extended to span days 29 - 365 of the forecast. This part of the ensemble is defined *outside* of the HEFS using raw climatology, in which each member corresponds to one year in the 41-year historical record. Raw climatology is also used to define ensemble members for freezing level for days 11 through 365. The resulting 365-day *forecasts* are then merged with the 10-day single-valued *observed* forcings (QPE, QTE, and QZE) to create the 375-day (from T0 - 10 days to T0 + 365 days) forcing series required to execute the hydrologic models. A summary of the resulting forcings is provided in **Table 14**. Note that work is underway at CNRFC to configure MEFP to generate a freezing level ensemble for days 1 through 10.

Table 14 - Overview of Ensemble Forcings

T0

	Days before T-0	**Days after T-0**				
Forcing	**10 - 1**	**1 - 3**	**4 - 6**	**7 - 10**	**11 -28**	**29 - 365**
precipitation	HAS QPE	ensemble (MEFP)				ens. (climatology)
temperature	HAS QTE	ensemble (MEFP)				ens. (climatology)
freezing level	HAS QZE	HAS QZF			ens. (climatology)	

(white)	ensemble
(grey)	single-valued series

I - Apply Forcings Ensembles to Hydrologic Models

The hydrologic models are configured to and executed within the framework of the Community Hydrologic Prediction System (CHPS). The ensemble forcings of **Table 14** are applied to the hydrologic models. Note that for the T0 = 12z forecast, the ensemble forecast extends 365 days. For all other forecast times (T0 = 0z, 6z, or 18z) the ensemble forecast extends 28 days.

The ensemble forcings are applied across all subbasins (3 forcings for each subbasin elevation zone) one member at a time. This ensures that historically-based spatial and temporal patterns embedded by the Schaake Shuffle are preserved. The results of applying the ensemble forcings to the hydrologic models are ensemble streamflow forecasts, reflecting only meteorologic uncertainty. At each ensemble forecast location, the ensemble streamflow forecast consists of 41 members. These streamflow forecasts are then further processed (outside of the HEFS) to create various probabilistic displays, including plots of short-term discharge and seasonal runoff volume.

Within CHPS, any non-expired modifiers made in previous simulations, such as HAS forecasts or update states simulations, are included in the ensemble simulations. The only exception is the time series change (TSCHNG) modifier when applied to a discharge series will not be included in the ensemble simulations.

5 MEFP Limitations

This section is intended to list the more significant limitations of MEFP. Some limitations, particularly those relating to issues of consistency, are not limitations of the HEFS methodology, but are limitations on data availability. Other limitations, particularly those contributing to underprediction of rare events, will be addressed to some extent with the release of the HEFS version 2.

Only Meteorological Uncertainty is Considered

The MEFP system is a very robust and stable system that can be implemented easily into an operational environment. Verification studies done by the CNRFC have shown that the MEFP program provides statistically reliable spread across seasons, lead times, and event size. However, the uncertainty is limited to that associated with forecasted precipitation and temperature. The HEFS does have hydrologic uncertainty components (EnsPostPE and EnsPost), but these have not yet been implemented at the CNRFC. This is because previous testing on an early version of HEFS indicated additional refinement of these components would be appropriate before testing further.

The MEFP does not Generate Freezing Level Ensembles

Another MEFP limitation is the quality of temperature estimates during winter storms. Since MEFP can only be parameterized for temperature and precipitation, uncertainty in rain-snow elevation estimates is derived within SNOW-17 when the lapsed temperature forecast is used to estimate the rain-snow elevation. MEFP creates 6-hour temperature ensembles derived from daily maximum and minimum forecasts. This method works well when describing the daily diurnal pattern during clear sky situations. But it does not work well for precipitation events when variations between daily maximum and minimum temperature are compressed or even non-existent. This occurs when storm attributes, such as frontal passage, overwhelm the normal diurnal pattern. In these situations, the diurnal pattern forecast can be overstated and result in incorrect precipitation typing (rain or snow). This can be problematic for basins where watershed area changes dramatically with just a slight change in elevation. In these cases, a very large area of the watershed could be modeled as snow falling due to unreasonable low temperature estimates from the diurnal temperature estimates. This issue could be improved by adding a third parameter to the MEFP - freezing level. Also, parameterizing temperatures based on 6-hour records rather than daily maximums and minimums would also be an improvement. Because of this, CNRFC has configured HEFS to use the single-valued freezing level estimate (HAS-QZF) for all ensemble members for the first 10 days of the forecast. This change eliminates any uncertainty in precipitation typing, but does provide a more realistic estimate of where it is raining and snowing in a watershed. So MEFP temperature uncertainty impacts are limited to snowmelt processes modeled by SNOW-17 during the first 10 days of the forecast run.

CNRFC notes that other RFCs are using the two temperature slots in the MEFPPE to forecast two of the synoptic times and then interpolating the other two. So there is a mitigation/workaround for situations where the diurnal-cycle modeling is inappropriate.

Conditionality of Meteorological Uncertainty

In addition, as described above, the MEFP is calibrated using samples across a moving 61-day window. The samples likely represent a diverse set of atmospheric conditions that do not have the same predictability. As such, it is possible that MEFP provides over-dispersed ensembles when atmospheric conditions are more predictable (strongly forced frontal system) and under-dispersed ensembles when atmospheric conditions are less predictable (a cut-off low or convective). However, deriving conditional distributions could run into issues related to inadequate sample sizes.

Limited Ensemble Spread in Late Season Snowmelt Forecasts

While HEFS forecasts generated by MEFP reflect uncertainty in meteorology, uncertainty in the current state of the hydrologic models is not reflected. With respect to snowmelt forecasts, it is important to recognize that uncertainty in the modeled snowpack is not reflected. While hydrologic models are periodically updated to reflect latest available snow course measurements, the resulting values of basin zone snow-water equivalent (SWE) are single-valued best estimates. Uncertainty about these estimates is not modeled.

Consistency in Forecast Models and Methods

Current and past forecast data sets should be as consistent as possible to avoid introducing errors in bias or spread into the ensemble forecast. Ideally, the current operational forecast model, and associated forecast methods, would be exactly consistent with the model and methods reflected in past forecasts. However, the operational forecast model and methods are adjusted with time in order to provide a best forecast. Past forecasts in the form of reforecasts will typically reflect a single “frozen” version of the model, and past forecasts in the form of archived forecasts will reflect any changes in models or methods during the record. Efforts are made to build data sets that are as consistent as possible, but they are not perfectly consistent.

An example of a known inconsistency at the CNRFC is described here. At the CNRFC, the HAS QPF for days 1 - 3 is used as the single-valued precipitation input to the MEFP. Archived forecasts for the period of record wy 2010 - 2021 (**Table 5**) supplied to the MEFPPE for computation of statistical parameters. However, the National Blend of Models (NBM), which is a component of the current HAS QPF, is only reflected in the HAS QPF forecast archive for the last few years of the record. Through testing, the CNRFC determined that ensembles computed using the HAS-QPF for days 1 - 3, still out performed GEFSv12 even with the inconsistent representation of the NBM.

Consistency in Period of Record

The HEFS computations can also be affected by inconsistencies in period of record. There is a period of record of historical data that the Schaake shuffle draws upon to rank historical canonical events. This period of record should equal that from which canonical events are drawn for the MEFPPE to compute statistical parameters. Computations to create ensembles beyond the MEFP period (28 days at the CNRFC), which can be based on raw or sampled climatology, should also reflect the same period of record to prevent sudden changes at the transition.

Only One Parametric Distribution Option

The MEFP is limited to the gamma distribution for fitting the marginal distributions of observations and forecasts. This can be an issue when trying to get a good fit at the tails of a distribution. The quantile-quantile plots in **Figures 20 and 21** show how observations and forecasts for an example location, French Meadows (FMDC1), do not have a good fit in the upper tail of the distribution. Verification studies at the CNRFC have shown that this issue can adversely affect the reliability of the ensemble forecast associated with a large event (greater than about 200 mm over 3 days at FMDC1), resulting in a low bias.

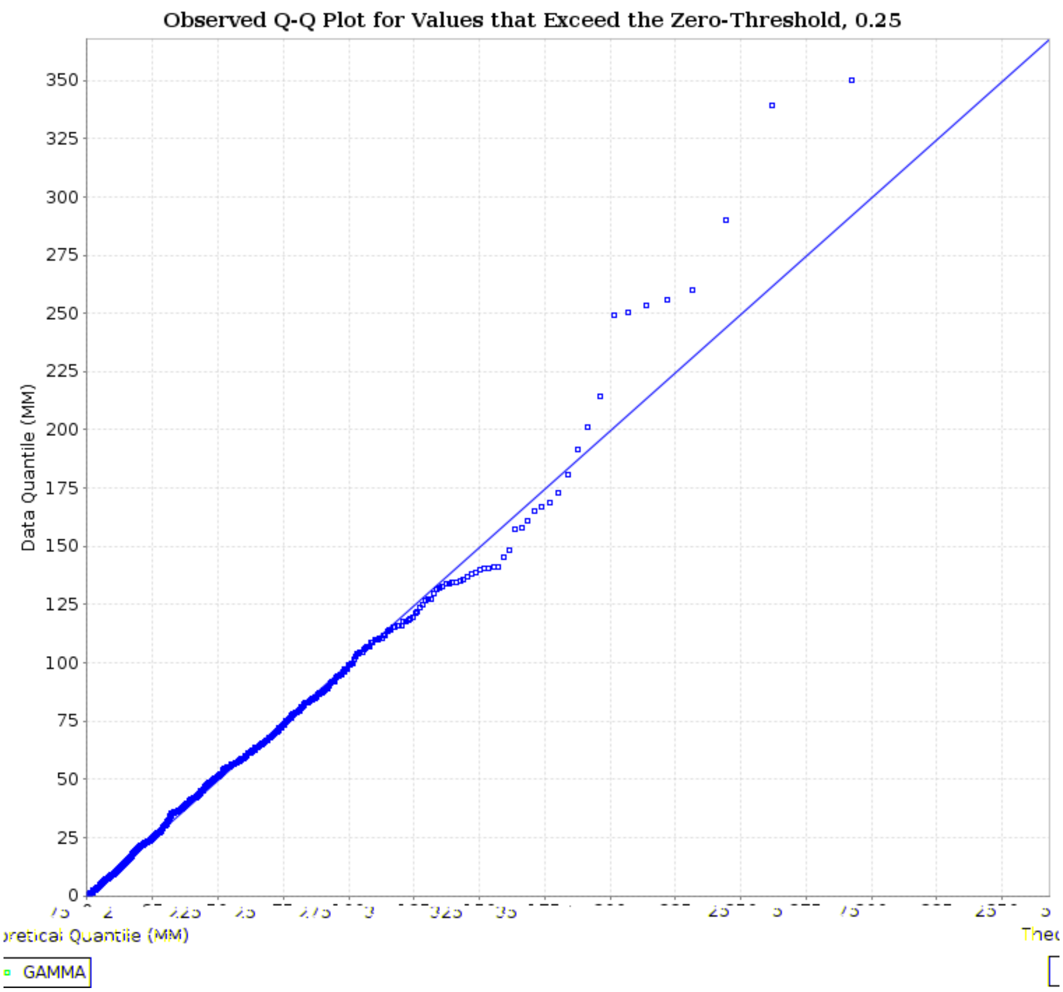


Figure 20 - French Meadows *Observed* Data & Theoretical Quantiles for 3-day Total Precipitation for January 26th 60-day Window Sample Size

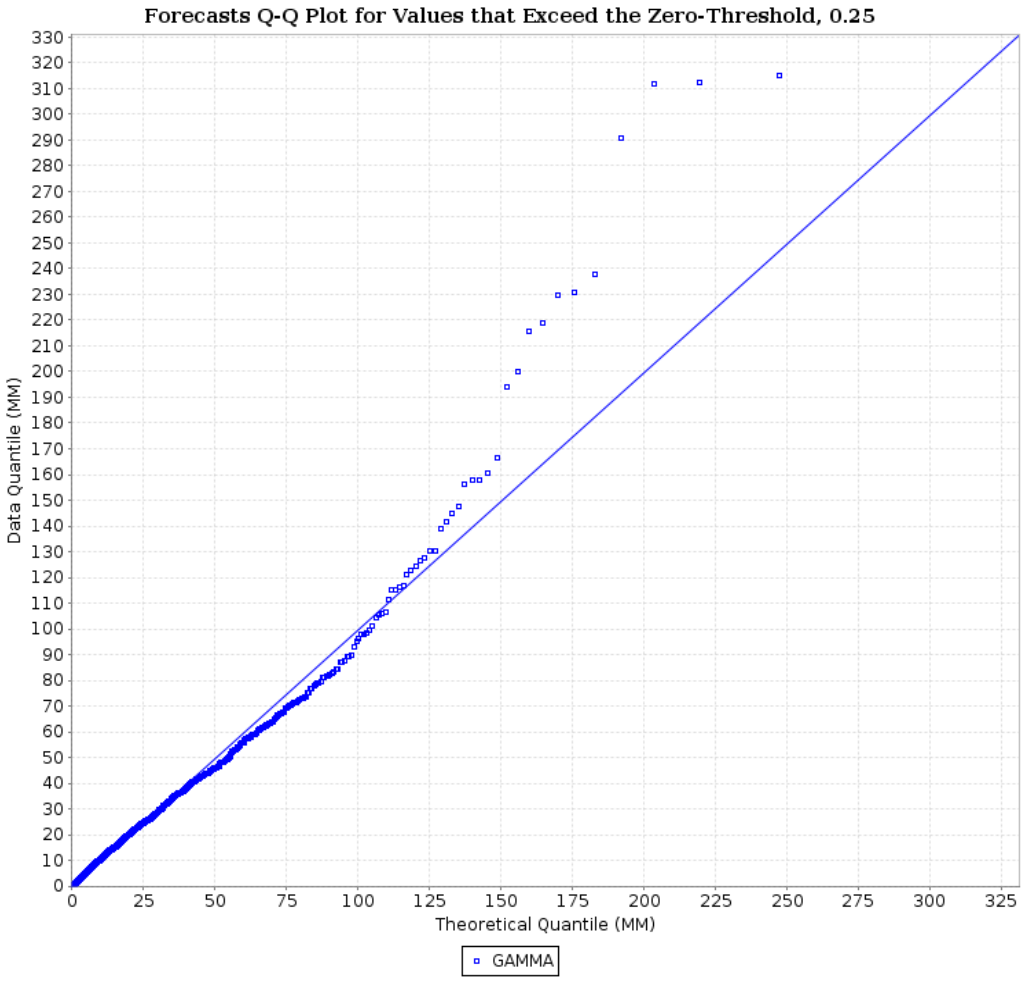


Figure 21 - French Meadows *Forecast* Data & Theoretical Quantiles for 3-day Total Precipitation for January 26th 60-day Window Sample Size

Type-II Conditional Bias

Type-II conditional bias (T2CB) in this case refers to the tendency of MEFP to systematically underestimate the most extreme observed precipitation amounts. In contrast, smaller forecasts are reasonably unbiased, conditional upon the forecast amount (aka small Type-I conditional bias or good "reliability"). The main reason for this is a "regression dilution" or "attenuation effect" (see Wikipedia for description), which is common with regression-type statistical post-processors, such as the MEFP. Methods for reducing the effect of T2CB are under consideration for implementation in the HEFS v2.

Lack of Smoothness between Canonical Event Boundaries

Each canonical event is a separate statistical model. When these models are brought together in a forecast horizon, without any kind of smoothing (as is the case with the MEFP), then any differences in the statistical behavior between these events (e.g., merely due to sampling uncertainty) will translate into discontinuities in the forecast horizon. They are generally most prominent for temperature because it is a smoothly varying time-series. At the CNRFC, this lack of smoothness can also occur when transitioning from the last day of the MEFP-generated ensemble forecast (day 28) to the first day of the ensemble forecast developed outside of MEFP using raw climatology. This occurs when there is an inconsistency between the “no-skill” baseline adopted by the MEFP for periods of forecast forcing, which is known as “resampled climatology”, and the raw climatology used after the period of forecast forcing. This typically occurs when the period of record for the forecast forcing (and hence resampled climatology) is different from the period of record used for raw climatology.

The Schaake Shuffle is not Flow Dependent

The MEFP uses observed time-series that begin on the same historical month/day/hour in each of N historical years. It is purely conditional upon the month, day and hour at which the forecast is issued, nothing else. For example, if there is an extreme atmospheric river on 21 January 2024 at location XYZ, but there are no similar cases on or near that calendar day in the historical record, then the Schaake shuffle will provide a poor representation of the space time patterns because it will use largely dry conditions to shuffle an extremely wet forecast. In that case the Schaake shuffle will effectively randomize the inputs (since dry values all have tied ranks). Alternatives to the Schaake shuffle exist, each with unique strengths and weaknesses. One such alternative under consideration is adopting a “flow dependent” approach, in which shuffling is conditioned on the current forecast (GEFS for example) state of the atmosphere. This has the potential benefit of being more likely to capture extreme conditions if the forecasting model is more skillful than climatology (which is, effectively, what the Schaake shuffle relies on), but also has limitations which are beyond the scope of this document.

6 HEFS Products

The phrase “HEFS products”, as used in this section, refers to the streamflow forecast ensembles and the variety of probabilistic products obtained derived from them. A variety of HEFS products are disseminated through the CNRFC website. The simplest are the actual streamflow forecast ensemble time series, which can be downloaded in csv format. Hourly ensemble csv data include regulation effects and are provided out to 30 days. Daily ensemble csv data which do not include regulation effects are provided out to 365 days.

The hourly 30-day HEFS forecasts can be downloaded for an entire forecast group in csv format at: 30-day HEFS (**Figure 22**). Through this website, a user can also obtain older ensemble forecast csv files through the “Forecast Groups Archive” at the bottom of the page. There is also an option to obtain the current ensemble forecast in csv format for a single location by entering in the five character ID in the “Individual Points” section.

Ensemble traces can also be viewed and downloaded in csv format for a given location in the interactive short range peak exceedance plot for every location where HEFS results are available. **Figure 23** shows an example for the West Walker River. All of the traces can be displayed on the plot by clicking the “View Model Traces” button to the right of the graph. The hourly csv data can be obtained by clicking on the 5 letter ID above the graphic where it says “CSV Ensemble File Download”. There are other short range graphics that can be viewed on the CNRFC website, such as probabilistic accumulated volumes, and daily box plots and histograms.

There are also a number of long range volume plots for many HEFS locations as well. There are graphical displays for forecast monthly, seasonal (April-July), water year, and multi-year volumes. Above the water year accumulation plots (**Figure 24**) the daily 365-day ensemble time series can be accessed by clicking on the five letter ID next to the “CSV Ensemble File Download” text. There is also a 365-day HEFS csv download site similar to the hourly one where a user can download an HEFS forecast csv file for all locations in a given forecast group, obtain older HEFS forecasts for a given date, and also get a current 365-day HEFS for a specified location. This site can be accessed by clicking on the “Forecast Group” text to the right of the “CSV Ensemble File Download”.

CALIFORNIA NEVADA RIVER FORECAST CENTER

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

HOME HYDROLOGY WEATHER CLIMATE RESEARCH / OUTREACH LINKS SEARCH ABOUT US

Short Range Hourly Ensemble CSV File Download

Forecast Groups

Forecast Group		Filename	Date/Time Last Modified	Size	ID Help Doc
Klamath		2022081612_klamath_hefs_csv_hourly.zip	16-Aug-2022 08:26 AM PDT	852K	
North Coast		2022081612_NorthCoast_hefs_csv_hourly.zip	16-Aug-2022 08:25 AM PDT	260K	
Russian/Napa		2022081612_RussianNapa_hefs_csv_hourly.zip	16-Aug-2022 08:38 AM PDT	143K	
Upper Sacramento		2022081612_UpperSacramento_hefs_csv_hourly.zip	16-Aug-2022 07:48 AM PDT	968K	
Feather/Yuba		2022081612_FeatherYuba_hefs_csv_hourly.zip	16-Aug-2022 07:49 AM PDT	899K	
Cache/Putah		2022081612_cacheputah_hefs_csv_hourly.zip	16-Aug-2022 08:12 AM PDT	45K	
American		2022081612_american_hefs_csv_hourly.zip	16-Aug-2022 08:13 AM PDT	1054K	
Lower Sacramento		2022081612_LowerSacramento_hefs_csv_hourly.zip	16-Aug-2022 09:03 AM PDT	59K	
Central Coast		2022081612_CentralCoast_hefs_csv_hourly.zip	16-Aug-2022 08:26 AM PDT	90K	
Southern California		2022081612_SouthernCalifornia_hefs_csv_hourly.zip	16-Aug-2022 08:10 AM PDT	328K	
Tulare		2022081612_Tulare_hefs_csv_hourly.zip	16-Aug-2022 08:13 AM PDT	276K	
San Joaquin		2022081612_SanJoaquin_hefs_csv_hourly.zip	16-Aug-2022 08:13 AM PDT	646K	
North San Joaquin		2022081612_n_sanjoaquin_hefs_csv_hourly.zip	16-Aug-2022 08:21 AM PDT	148K	
East Sierra		2022081612_EastSierra_hefs_csv_hourly.zip	16-Aug-2022 08:22 AM PDT	1031K	
Humboldt		2022081612_Humboldt_hefs_csv_hourly.zip	16-Aug-2022 08:38 AM PDT	134K	

Note 1: Data represented in the files are in kcfs (thousands of cubic feet per second).

Note 2: Each location includes 41 ensemble members with the first column starting with year 1980 and the last column ending with year 2020. Each historical year is more meaningful beyond the first couple weeks when climatology drives the spread in the ensemble.

Individual Points

Enter desired location(s) below (some older browsers may allow only one file download at a time):

Download

Forecast Groups Archive

Forecast Group	Start Date	Start Hour (Forecast Cycle)	End Date	End Hour (Forecast Cycle)
Klamath	08/16/2022	12 UTC	08/16/2022	12 UTC

Download

Figure 22 - Hourly 30-day Ensemble Streamflow Forecasts

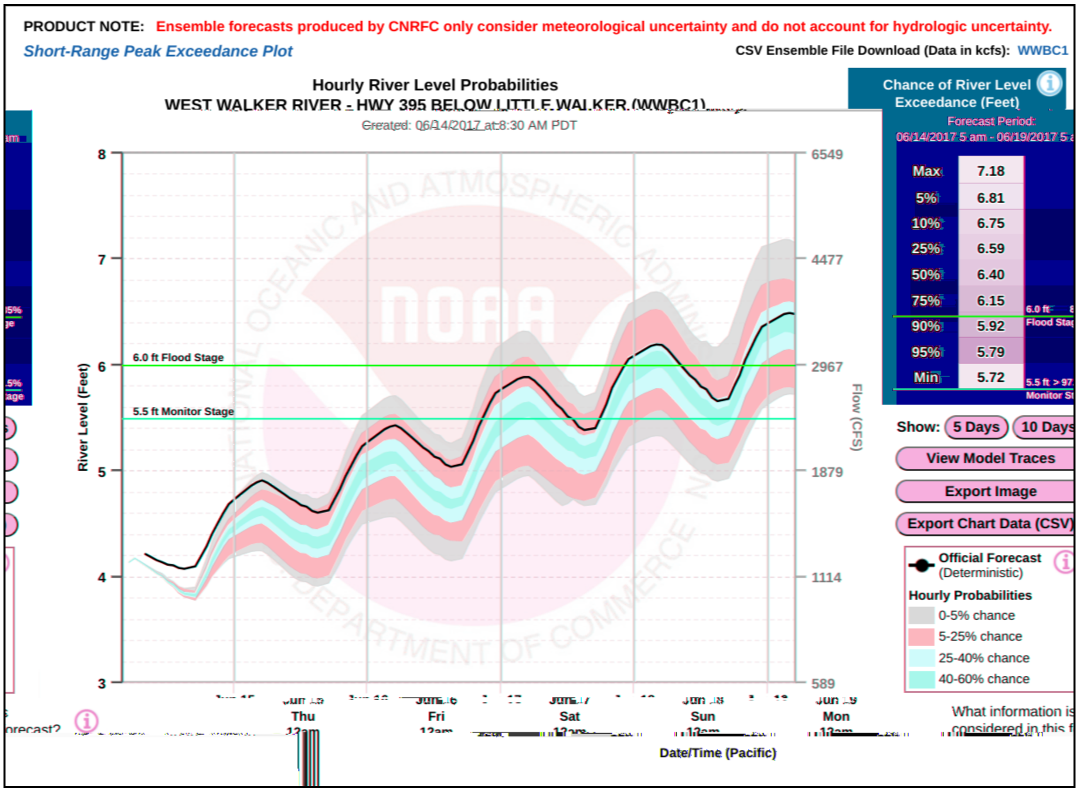


Figure 23 - Short Range Ensemble Graphic

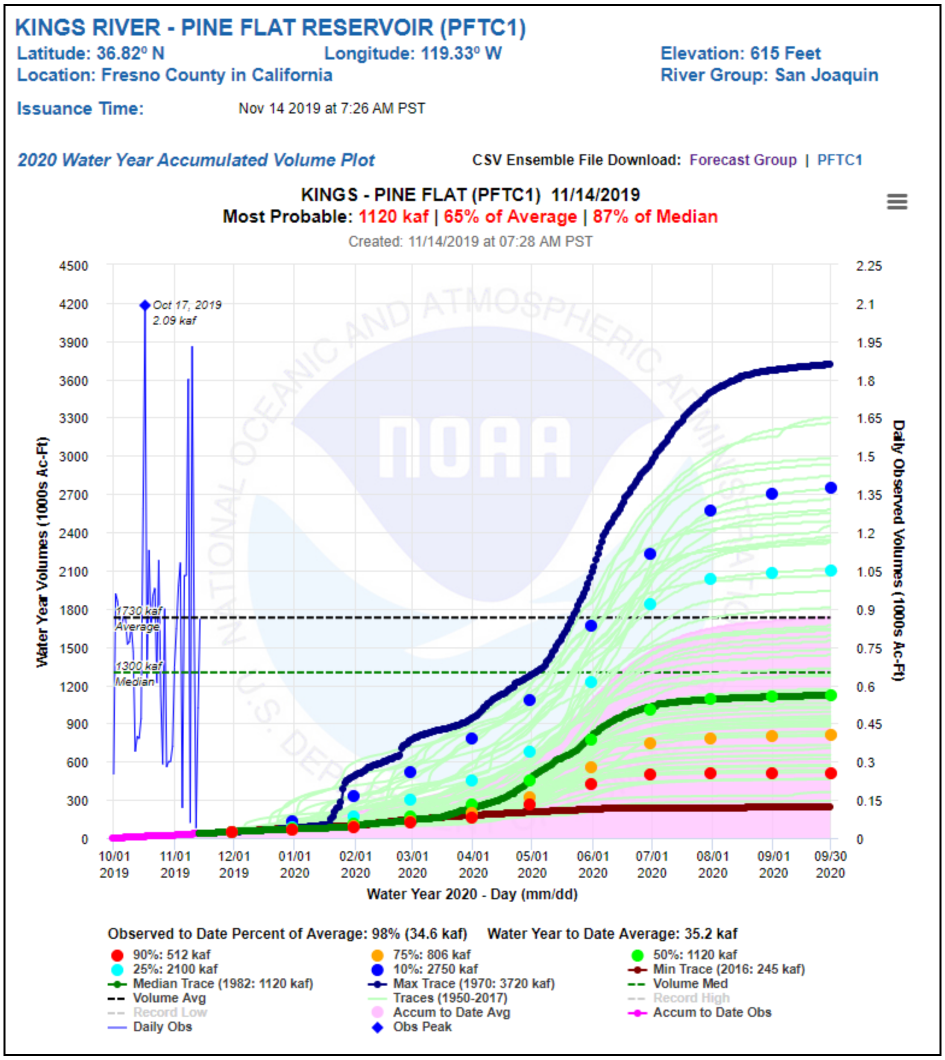


Figure 24 - Long Range Water Year Accumulation Plot

7 References

Anderson, E.A., 'A Point Energy and Mass Balance Model of a Snow Cover', NOAA Technical Report NWS 19, 150 pp, February 1976.

Burnash, R. J. C., Ferral, R. L., and McGuire, R. A. 'A Generalized Streamflow Simulation System, Conceptual Modeling for Digital Computers', National Weather Service, California Department of Water Resources, March 1973.

Clarke, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., Wilby, R., 2004. The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields. Journal of Hydrometeorology, 5 (1), 243-262.

Demargne, j., Wu, L., Regonda, S. K., Brown, J. D., Lee, H., He, M., Seo, D., Hartman, R., Herr, H., Fresch, Schaake, J., and Zhu, Y., The science of NOAA's operational hydrologic ensemble forecast service., Bull. Amer. Meteor. Soc., 2014, Jan., 79-98.

Herr, H.D., Krzysztofowicz, R., 2005. Generic probability distribution of rainfall in space: the bivariate model. J. Hydrol., 306, 234–263.

Wu, L., Seo, D., Demargne, J., Brown, J., Cong, S., Schakke, J., 2011. Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. J. Hydrol., 399, 281-298.

8 Attachment "Schaake Shuffle Step-by-Step Example"

THIS PAGE INTENTIONALLY LEFT BLANK

Schaake Shuffle Step-by-Step Example

- Assume we have a 24 hour precipitation forecast. We have 4 6-hour base events and one 24 hour total modulation event – 5 canonical events.
- Assume we have 10 ensemble members that matches the number of years we have in our climatology.
- The Schaake Shuffle will compute the mapping of the ensemble members one canonical event at a time starting from lowest correlation, and ending with the highest correlated canonical event.
- Base and modulation event precipitation amounts will be shuffled based on the climatological ordering technique.

Here are the raw ensemble member precipitation values generated by MEFP prior to being shuffled by the Schaake Shuffle method

6-hour base event periods

24 aggregation modulation event

	0-6	6-12	12-18	18-24
ens #	FMAP (in)	FMAP (in)	FMAP (in)	FMAP (in)
1	0.60	1.00	0.85	0.20
2	0.44	0.55	0.24	0.40
3	0.35	0.40	0.17	0.13
4	0.47	0.65	0.32	0.17
5	0.30	0.32	0.21	0.09
6	0.38	0.53	0.73	0.10
7	0.33	0.35	0.25	0.05
8	0.49	0.75	0.38	0.15
9	0.41	0.64	0.21	0.16
10	0.40	0.44	0.30	0.11
	0.78	0.72	0.68	0.63

0-24
FMAP (in)
3.08
1.83
1.23
1.80
1.09
1.81
1.13
2.13
1.67
1.48
0.82

correlation values

	0-6	6-12	12-18	18-24
ens #	FMAP (in)	FMAP (in)	FMAP (in)	FMAP (in)
1	0.60	1.00	0.85	0.20
2	0.44	0.55	0.24	0.40
3	0.35	0.40	0.17	0.13
4	0.47	0.65	0.32	0.17
5	0.30	0.32	0.21	0.09
6	0.38	0.53	0.73	0.10
7	0.33	0.35	0.25	0.05
8	0.49	0.75	0.38	0.15
9	0.41	0.64	0.21	0.16
10	0.40	0.44	0.30	0.11

0-24
FMAP (in)
3.08
1.83
1.23
1.80
1.09
1.81
1.13
2.13
1.67
1.48

0-6	6-12	12-18	18-24	0-24
0.78	0.72	0.68	0.63	0.82
4	3	2	1	5

Based on the correlation values, the 18-24 hour period will get shuffled first because it has the lowest correlation. The modulation event will be applied last because it has the highest correlation.

- Let's start with the Schaake Shuffle being applied to the lowest correlated period: 18-24 hr.
- We want to map these 10 members to historical years as part of the Schaake Shuffle.
- We apply the Schaake Shuffle method to each base event, one at a time.
- We will step through this example step by step for base event 18-24hr.
- The shuffling for the 18-24 hour period is associated with the corresponding historical precipitation amount ordering.

MEFP unshuffled precipitation 6-hr precipitation values

	0-6	6-12	12-18	18-24
ens #	FMAP (in)	FMAP (in)	FMAP (in)	FMAP (in)
1	0.60	1.00	0.85	0.20
2	0.44	0.55	0.24	0.40
3	0.35	0.40	0.17	0.13
4	0.47	0.65	0.32	0.17
5	0.30	0.32	0.21	0.09
6	0.38	0.53	0.73	0.10
7	0.33	0.35	0.25	0.05
8	0.49	0.75	0.38	0.15
9	0.41	0.64	0.21	0.16
10	0.40	0.44	0.30	0.11

Historical precipitation values for corresponding forecast periods

	0-6	6-12	12-18	18-24
year	MAP (in)	MAP (in)	MAP (in)	MAP (in)
1990	0.05	0.52	0.26	0.09
1991	0.37	0.33	0.43	0.62
1992	0.27	0.18	0.30	0.42
1993	0.07	0.49	0.11	0.04
1994	0.48	0.70	0.52	0.24
1995	0.54	1.30	0.83	0.32
1996	0.35	0.25	0.09	0.03
1997	0.11	0.27	0.13	0.06
1998	0.02	0.61	0.45	0.20
1999	0.51	0.90	0.72	0.66

Ensemble Member (18-24 hour period)	
ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked **Ensemble** (18-24 hour period)	
ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

First, the 10 ensemble values are ranked by forecast value for the base event of interest.

Ensemble Member
(18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked **Ensemble**
(18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values
(18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

The 10 historical precipitation amounts are determined for the given forecast period.

Ensemble Member
(18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked **Ensemble**
(18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values
(18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked **Historical**

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

The 10 historical precipitation amounts are then ranked.

The highest ranked precipitation ensemble value is assigned the historical year with the largest precipitation amount.

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	**0.40**

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	**1**	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	**1**	2	**0.4**

The second highest ranked precipitation ensemble value is assigned the historical year with the second largest precipitation amount.

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	**0.20**
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	**2**	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	**2**	1	**0.2**
1999	1	2	0.4

This process is repeated for all ensemble values.

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	**0.17**
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	**3**	4	**0.17**
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	**0.16**
4	0.17
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	**0.15**
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member
(18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble
(18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	**0.15**
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values
(18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	**0.32**
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	**5**	**0.24**
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	**5**	**8**	**0.16**
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	**0.13**
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	**0.13**
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member
(18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble
(18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	**0.11**
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values
(18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	**0.11**
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member
(18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble
(18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	**0.10**
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values
(18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	**0.1**
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	**0.09**
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	**0.09**
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked Ensemble (18-24 hour period)

ens #	FMAP (in)
7	**0.05**
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked Historical

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	**0.05**
1993	9	5	0.09
1997	8	6	0.1
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.2
1999	1	2	0.4

- Now let's look at the ensemble members for the second lowest correlated base event: 12-18 hour forecast period.

- Ensemble members are ranked just like for the 18-24 hour base event

Ensemble Member (12-18 hour period)

ens #	FMAP (in)
1	0.85
2	0.24
3	0.17
4	0.32
5	0.21
6	0.73
7	0.25
8	0.38
9	0.21
10	0.30

Ranked **Ensemble** (12-18 hour period)

ens #	FMAP (in)
3	0.17
5	0.21
9	0.21
2	0.24
7	0.25
10	0.30
4	0.32
8	0.38
6	0.73
1	0.85

Historical values for the base event are selected and ranked just like for the 18-24 base event

Ensemble Member (12-18 hour period)

ens #	FMAP (in)
1	0.85
2	0.24
3	0.17
4	0.32
5	0.21
6	0.73
7	0.25
8	0.38
9	0.21
10	0.30

Ranked **Ensemble** (12-18 hour period)

ens #	FMAP (in)
3	0.17
5	0.21
9	0.21
2	0.24
7	0.25
10	0.30
4	0.32
8	0.38
6	0.73
1	0.85

Historical Values (12-18 hour period)

year	MAP (in)
1990	0.26
1991	0.43
1992	0.30
1993	0.11
1994	0.52
1995	0.83
1996	0.09
1997	0.13
1998	0.45
1999	0.72

Ranked **Historical**

year	Hist Rank	MAP (in)
1996	10	0.09
1993	9	0.11
1997	8	0.13
1990	7	0.26
1992	6	0.30
1991	5	0.43
1998	4	0.45
1994	3	0.52
1999	2	0.72
1995	1	0.83

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	3	0.17
1993	9	5	0.21
1997	8	9	0.21
1990	7	2	0.24
1992	6	7	0.25
1991	5	10	0.30
1998	4	4	0.32
1994	3	8	0.38
1999	2	6	0.73
1995	1	1	0.85

Results from Base Event 18-24 hr

Ensemble Member (18-24 hour period)

ens #	FMAP (in)
1	0.20
2	0.40
3	0.13
4	0.17
5	0.09
6	0.10
7	0.05
8	0.15
9	0.16
10	0.11

Ranked **Ensemble** (18-24 hour period)

ens #	FMAP (in)
7	0.05
5	0.09
6	0.10
10	0.11
3	0.13
8	0.15
9	0.16
4	0.17
1	0.20
2	0.40

Historical Values (18-24 hour period)

year	MAP (in)
1990	0.09
1991	0.62
1992	0.42
1993	0.04
1994	0.24
1995	0.32
1996	0.03
1997	0.06
1998	0.20
1999	0.66

Ranked **Historical**

year	Hist Rank	MAP (in)
1996	10	0.03
1993	9	0.04
1997	8	0.06
1990	7	0.09
1998	6	0.20
1994	5	0.24
1995	4	0.32
1992	3	0.42
1991	2	0.62
1999	1	0.66

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	7	0.05
1993	9	5	0.09
1997	8	6	0.10
1990	7	10	0.11
1998	6	3	0.13
1994	5	8	0.16
1995	4	9	0.15
1992	3	4	0.17
1991	2	1	0.20
1999	1	2	0.40

Results from Base Event 12-18 hr

Ensemble Member (12-18 hour period)

ens #	FMAP (in)
1	0.85
2	0.24
3	0.17
4	0.32
5	0.21
6	0.73
7	0.25
8	0.38
9	0.21
10	0.30

Ranked **Ensemble** (12-18 hour period)

ens #	FMAP (in)
3	0.17
5	0.21
9	0.21
2	0.24
7	0.25
10	0.30
4	0.32
8	0.38
6	0.73
1	0.85

Historical Values (12-18 hour period)

year	MAP (in)
1990	0.26
1991	0.43
1992	0.30
1993	0.11
1994	0.52
1995	0.83
1996	0.09
1997	0.13
1998	0.45
1999	0.72

Ranked **Historical**

year	Hist Rank	MAP (in)
1996	10	0.09
1993	9	0.11
1997	8	0.13
1990	7	0.26
1992	6	0.30
1991	5	0.43
1998	4	0.45
1994	3	0.52
1999	2	0.72
1995	1	0.83

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1996	10	3	0.17
1993	9	5	0.21
1997	8	9	0.21
1990	7	2	0.24
1992	6	7	0.25
1991	5	10	0.30
1998	4	4	0.32
1994	3	8	0.38
1999	2	6	0.73
1995	1	1	0.85

Results from Base Event 6-12 hr

Ensemble Member (6-12 hour period)

ens #	FMAP (in)
1	1.00
2	0.55
3	0.40
4	0.65
5	0.32
6	0.53
7	0.35
8	0.75
9	0.64
10	0.44

Ranked **Ensemble** (6-12 hour period)

ens #	FMAP (in)
5	0.32
7	0.35
3	0.40
10	0.44
6	0.53
2	0.55
9	0.64
4	0.65
8	0.75
1	1.00

Historical Values (6-12 hour period)

year	MAP (in)
1990	0.52
1991	0.33
1992	0.18
1993	0.49
1994	0.70
1995	1.30
1996	0.25
1997	0.27
1998	0.61
1999	0.90

Ranked **Historical**

year	Hist Rank	MAP (in)
1992	10	0.18
1996	9	0.25
1997	8	0.27
1991	7	0.33
1993	6	0.49
1990	5	0.52
1998	4	0.61
1994	3	0.70
1999	2	0.90
1995	1	1.30

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1992	10	5	0.32
1996	9	7	0.35
1997	8	3	0.40
1991	7	10	0.44
1993	6	6	0.53
1990	5	2	0.55
1998	4	9	0.64
1994	3	4	0.65
1999	2	8	0.75
1995	1	1	1.00

Results from Base Event 0-6 hr

Ensemble Member (0-6 hour period)

ens #	FMAP (in)
1	0.60
2	0.44
3	0.35
4	0.47
5	0.30
6	0.38
7	0.33
8	0.49
9	0.41
10	0.40

Ranked **Ensemble** (0-6 hour period)

ens #	FMAP (in)
5	0.30
7	0.33
3	0.35
6	0.38
10	0.40
9	0.41
2	0.44
4	0.47
8	0.49
1	0.60

Historical Values (0-6 hour period)

year	MAP (in)
1990	0.05
1991	0.37
1992	0.27
1993	0.07
1994	0.48
1995	0.54
1996	0.35
1997	0.11
1998	0.02
1999	0.51

Ranked **Historical**

year	Hist Rank	MAP (in)
1998	10	0.02
1990	9	0.05
1993	8	0.07
1997	7	0.11
1992	6	0.27
1996	5	0.35
1991	4	0.37
1994	3	0.48
1999	2	0.51
1995	1	0.54

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1998	10	5	0.30
1990	9	7	0.33
1993	8	3	0.35
1997	7	6	0.38
1992	6	10	0.40
1996	5	9	0.41
1991	4	2	0.44
1994	3	4	0.47
1999	2	8	0.49
1995	1	1	0.60

Results from Modulation Event 0-24hr Total

Ensemble Member
(24 hour period)

ens #	FMAP (in)
1	3.73
2	3.03
3	1.57
4	2.18
5	1.39
6	1.24
7	1.46
8	1.80
9	1.39
10	1.32

Ranked **Ensemble**
(24 hour period)

ens #	FMAP (in)
6	1.24
10	1.32
5	1.39
9	1.39
7	1.46
3	1.57
8	1.80
4	2.18
2	3.03
1	3.73

Historical Values
(24 hour period)

year	MAP (in)
1990	0.92
1991	1.75
1992	1.17
1993	0.71
1994	1.94
1995	2.99
1996	0.72
1997	0.57
1998	1.28
1999	2.79

Ranked **Historical**

year	Hist Rank	MAP (in)
1997	10	0.57
1993	9	0.71
1996	8	0.72
1990	7	0.92
1992	6	1.17
1998	5	1.28
1991	4	1.75
1994	3	1.94
1999	2	2.79
1995	1	2.99

Historical Year Label Mapping

year label	Ens# rank	Orig Ens #	FMAP (in)
1997	10	6	1.24
1993	9	10	1.32
1996	8	5	1.39
1990	7	9	1.39
1992	6	7	1.46
1998	5	3	1.57
1991	4	8	1.80
1994	3	4	2.18
1999	2	2	3.03
1995	1	1	3.73

Sorted Base Events are Combined into shuffled ensemble time series

0-6 hr

year label	FMAP (in)
1990	0.33
1991	0.44
1992	0.40
1993	0.35
1994	0.47
1995	0.60
1996	0.41
1997	0.38
1998	0.30
1999	0.49

6-12 hr

year label	FMAP (in)
1990	0.55
1991	0.44
1992	0.32
1993	0.53
1994	0.65
1995	1.00
1996	0.35
1997	0.40
1998	0.64
1999	0.75

12-18 hr

year label	FMAP (in)
1990	0.24
1991	0.30
1992	0.25
1993	0.21
1994	0.38
1995	0.85
1996	0.17
1997	0.21
1998	0.32
1999	0.73

18-24 hr

year label	FMAP (in)
1990	0.11
1991	0.20
1992	0.17
1993	0.09
1994	0.15
1995	0.16
1996	0.05
1997	0.10
1998	0.13
1999	0.40

Combined Results for the 4 base events

	FMAP (in)				
year label	0-6 hr	6-12 hr	12-18 hr	18-24 hr	total
1990	0.33	0.55	0.24	0.11	1.23
1991	0.44	0.44	0.30	0.20	1.38
1992	0.40	0.32	0.25	0.17	1.14
1993	0.35	0.53	0.21	0.09	1.18
1994	0.47	0.65	0.38	0.15	1.65
1995	0.60	1.00	0.85	0.16	2.61
1996	0.41	0.35	0.17	0.05	0.98
1997	0.38	0.40	0.21	0.10	1.09
1998	0.30	0.64	0.32	0.13	1.39
1999	0.49	0.75	0.73	0.40	2.37

- **Now we apply the modulation event last since it has the highest correlation**
- **The 24-hour modulation event is shuffled like the base events (see previous graphic)**
- **6-hour values are summed up over 24 hours**

6-hour values aggregated over 24 hour period

	FMAP (in)				
year label	0-6 hr	6-12 hr	12-18 hr	18-24 hr	total
1990	0.33	0.55	0.24	0.11	1.23
1991	0.44	0.44	0.30	0.20	1.38
1992	0.40	0.32	0.25	0.17	1.14
1993	0.35	0.53	0.21	0.09	1.18
1994	0.47	0.65	0.38	0.15	1.65
1995	0.60	1.00	0.85	0.16	2.61
1996	0.41	0.35	0.17	0.05	0.98
1997	0.38	0.40	0.21	0.10	1.09
1998	0.30	0.64	0.32	0.13	1.39
1999	0.49	0.75	0.73	0.40	2.37

Shuffled 24 hour Modulation Event Values

year label	FMAP (in)
1990	1.39
1991	1.80
1992	1.46
1993	1.32
1994	2.18
1995	3.73
1996	1.39
1997	1.24
1998	1.57
1999	3.03

A factor is calculated that will be applied uniformly to the individual 6-hour precipitation values so that the 6-hr summations equal the modulation event totals

24-hour totals from 6-hour values

year label	total
1990	1.23
1991	1.38
1992	1.14
1993	1.18
1994	1.65
1995	2.61
1996	0.98
1997	1.09
1998	1.39
1999	2.37

Modulation event shuffled totals

FMAP (in)
1.39
1.80
1.46
1.32
2.18
3.73
1.39
1.24
1.57
3.03

Ratio of modulation event values to 6-hr aggregations over 24-hr period

Scale Factor
1.13
1.30
1.28
1.12
1.32
1.43
1.42
1.14
1.13
1.28

A factor is calculated that will be applied uniformly to the individual 6-hour precipitation values. In this case, since the modulation event is applied last, the scaled 6-hour precipitation 24 hour totals equals the shuffled modulation event totals.

Ratio of modulation event values to 6-hr aggregations over 24-hr period

New 6-hr totals equal modulation event totals

	FMAP (in)				
year label	0-6 hr	6-12 hr	12-18 hr	18-24 hr	total
1990	0.33	0.55	0.24	0.11	1.23
1991	0.44	0.44	0.30	0.20	1.38
1992	0.40	0.32	0.25	0.17	1.14
1993	0.35	0.53	0.21	0.09	1.18
1994	0.47	0.65	0.38	0.15	1.65
1995	0.60	1.00	0.85	0.16	2.61
1996	0.41	0.35	0.17	0.05	0.98
1997	0.38	0.40	0.21	0.10	1.09
1998	0.30	0.64	0.32	0.13	1.39
1999	0.49	0.75	0.73	0.40	2.37

X

Scale Factor
1.13
1.30
1.28
1.12
1.32
1.43
1.42
1.14
1.13
1.28

=

	FMAP (in)				
year label	0-6 hr	6-12 hr	12-18 hr	18-24 hr	total
1990	0.37	0.62	0.27	0.12	1.39
1991	0.57	0.57	0.39	0.26	1.80
1992	0.51	0.41	0.32	0.22	1.46
1993	0.39	0.59	0.24	0.10	1.32
1994	0.62	0.86	0.50	0.20	2.18
1995	0.86	1.43	1.21	0.23	3.73
1996	0.58	0.50	0.24	0.07	1.40
1997	0.43	0.45	0.24	0.11	1.24
1998	0.34	0.73	0.36	0.15	1.58
1999	0.63	0.96	0.93	0.51	3.03

- In this example, the modulation event had the highest correlation, so it was ordered and applied last after all base events.
- If the modulation event had a lower correlation than one of the 6-hour base events, the modulation event would be applied before the 6-hour base event with the higher correlation.
- So the modulation event would be applied prior to all of the 6-hour base events being shuffled, and assigned historical year labels.
- In this case, climatological values are used for the base events that have not gone through the Schaake Shuffle when computing the modulation scale factor.
- The climatological values for the 6-hour base event with the higher correlation are replaced with shuffled MEFP values after the modulation event has been applied.