Overview of the Hydrologic Ensemble Forecast Service (HEFS)

Release 0.1.2

June 1, 2012

National Weather Service
Office of Hydrologic Development

I. Document history

Action	Person	Date
First draft	james.brown@hydrosolved.com	16th April 2012
Updated draft	james.brown@hydrosolved.com	22nd April 2012
Added Table 5	kevin.he@noaa.gov	25th April 2012
Input from Mark Fresch	mark.a.fresch@noaa.gov	26th April 2012
Added Table 3	satish.regonda@noaa.gov	26th April 2012
Added Table 2	limin.wu@noaa.gov	26th April 2012
Added Table 4	xuning.tan@noaa.gov	26th April 2012
Input from Bob Mills	robert.mills@noaa.gov	28th April 2012
Final edits and additions	james.brown@hydrosolved.com	30th April 2012
Created 0.1.2	james.brown@hydrosolved.com	21st May 2012
Document review	nickolas.heen@noaa.gov	30th May 2012
Final edits for 0.1.2	james.brown@hydrosolved.com	1st June 2012

CONTENTS

1. Purpose and Scope of this Document

This document aims to:

1) provide a non-technical overview of the Hydrologic Ensemble Forecast Service (HEFS) and;
2) summarize the key steps involved in preparing the HEFS for operational forecasting, hindcasting, and verification.

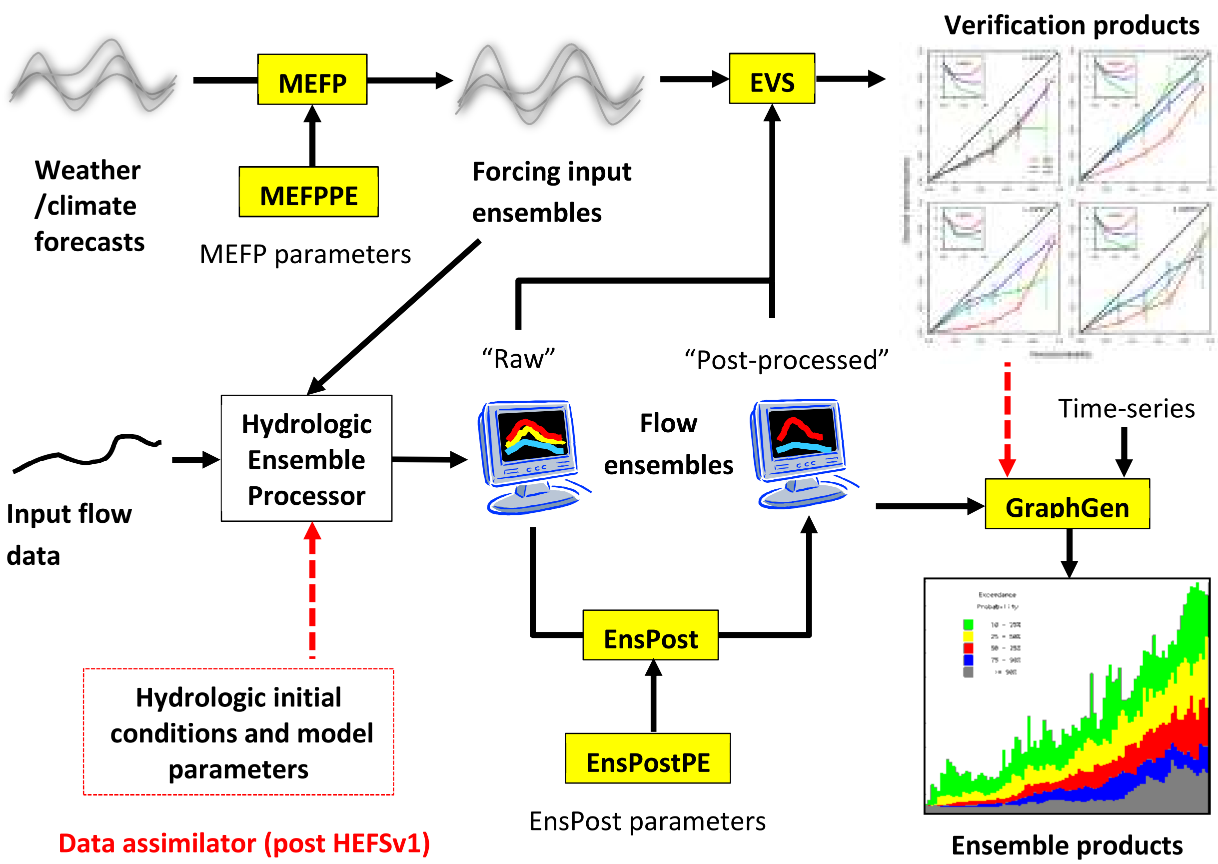
Detailed descriptions of the individual components of the HEFS, including the science behind the software, the detailed functionality and options, and how to apply the HEFS, are provided in the various users' manuals for the HEFS components. This document provides an overview, rather than detailed guidance for specific applications. However, future revisions to this document will accommodate illustrative applications or case studies of the HEFS as an "end-to-end" system.

2. Introduction to the HEFS

The HEFS was developed by the Office of Hydrologic Development (OHD) of the U.S. National Weather Service (NWS). The HEFS issues hydrologic forecasts that are "uncertainty aware", i.e. they provide information about forecast uncertainty. This is achieved by issuing an *ensemble* of possible values of the forecast variables. Unlike single-valued or "deterministic" forecasts, which comprise a *single* estimate of the forecast variable at each time and location, an ensemble forecast provides a set of possible values. An ensemble forecasting system, such as the HEFS, translates or "propagates" an ensemble of inputs (e.g. precipitation and temperature) through a hydrologic model to provide an ensemble of outputs (streamflow). Ensemble forecasting provides a convenient way to quantify and trace the movement of uncertainty through hydrologic models, which otherwise require fixed values of the inputs (i.e. fixed values of temperature and precipitation). In ensemble forecasting, the hydrologic models, such as SAC-SMA and SNOW-17, are executed repeatedly. Each execution uses a different value for the inputs and, by implication, provides one possible value for the outputs (streamflow). By collecting together the ensemble of outputs, the forecasts can be used to develop probability statements, such as the probability of flooding (fraction of members that exceed flood stage), or to use in other modeling tools or decision support systems.

Ensemble forecasting relies on a combination of physically-based and statistical modeling. The HEFS comprises both physically-based hydrologic models (e.g. SAC-SMA, SNOW-17) and statistical modeling of the forecast uncertainties. The components of the HEFS are implemented within a modular software framework (*Figure 1*). The HEFS modules aim to quantify and, where possible, reduce the uncertainties at various stages in the hydrologic modeling process, as well as generate outputs for operational forecasting.

Figure 1: overview of the HEFS



Statistical models rely on historical observations to determine historical forecast errors. This requires statistical modeling of the relationship between the past forecasts and observations. If this relationship is relatively constant or "stationary" in time, past forecasting errors provide a statistical guide to future forecasting errors. These statistical models provide the basis to generate the ensemble of inputs required by the hydrologic models (the Meteorological Ensemble Forecast Processor or MEFP) and to correct for consistent errors (biases) in the streamflow predictions (the Ensemble Post-processor or EnsPost). Of course, the ensembles generated by these models must be physically, as well as statistically, reasonable. In particular, they must reproduce observed patterns of forcing and streamflow in *space and time*. For example, adjacent basins could have similar precipitation amounts at any given time. They must also reproduce the observed relationships *between* variables. For example, precipitation will not fall as snow at high air temperatures. Moreover, they aim to reproduce these patterns at several different accumulation volumes (e.g. daily, monthly, etc.).

The statistical modeling in the HEFS is conducted in two parts. First, a Parameter Estimator (PE) is used to estimate the parameters of each statistical model. The parameters must be estimated from a long and consistent record of paired predictions and observations. This is necessary to minimize sampling uncertainty; that is, to provide parameter estimates that are reasonable and not too "noisy." Secondly, the estimated parameters are applied in real time to the "raw" operational forecasts, whether from the forcing models (e.g. GFS, CFSv2) or streamflow models (SAC-SMA and SNOW-17). For example, the MEFP Parameter Estimator (MEFPPE) estimates the parameters of the relationship between the historical observations and the "raw" temperature and precipitation forecasts from weather and climate models, such as the Global Forecast System (GFS). In real time, the MEFP uses the parameters from the MEFPPE together with the raw operational forecasts from the GFS. Using this information, the MEFP produces an ensemble of temperature and precipitation forecasts for input to the hydrologic models. The MEFP accounts for any biases in the raw forecasts via the model parameters estimated with the MEFPPE. The MEFP accommodates several sources of raw forecasts, including the single-valued operational forecasts from the

RFCs, the GFS, the Climate Forecast System (v.2), historical observations ("climatological forcing") and, shortly, the Global Ensemble Forecast System (GEFS). These raw forecasts are integrated seamlessly into the MEFP, in order to provide bias-corrected forcing from less than 1 day out to almost one year (to be extended in future).

For streamflow, the EnsPost Parameter Estimator (EnsPostPE) models the statistical relationship between the streamflow predictions (hydrologic simulations) and observations. Using the parameters estimated by the EnsPostPE, the EnsPost makes an adjustment to the raw streamflow forecasts in real time. This adjustment accounts for any biases identified in the historical raw forecasts used to calibrate the EnsPostPE. Both conceptually and practically, the MEFP and EnsPost are similar; they both aim to produce a forecast ensemble that is statistically consistent with the past observations under similar conditions to the "raw" inputs. However, they use different sources of information (forcing versus streamflow) and hence account for different sources of uncertainty. In this context, the total uncertainty in hydrologic forecasting may be factored into two main sources of uncertainty and bias, namely the input or "forcing uncertainties" and the "hydrologic uncertainties." The latter comprise all sources of uncertainty and bias in the hydrologic modeling. Since the MEFP only accounts for the forcing uncertainty, the EnsPost must account for the hydrologic uncertainty. In order to model the hydrologic uncertainties separately from the forcing uncertainties, the EnsPost uses hydrologic simulations. Hydrologic simulations comprise observed forcing and hence the forcing uncertainties are effectively eliminated (or at least minimized, if the observations are relatively error-free).

As indicated above, the MEFPPE and the EnsPostPE are similar in practice, as well as conceptually. Thus, both PEs use a simplified form of statistical modeling that invokes rather strict assumptions about the processes that generated the sample data. Since both PEs model the "scatter" between two variables (observed versus forecast precipitation for MEFPPE and observed versus simulated streamflow for EnsPostPE), they model a "bivariate" relationship. Both PEs assume that this bivariate relationship follows a normal or "Gaussian" probability distribution, and model the sample data accordingly. The main advantage of this assumption is that the normal distribution has very few parameters to estimate. To assist with this, the raw data are transformed using a "Normal Quantile Transform." In other words, the statistical modeling is conducted in a space that is consistent with the normal distribution. Once the model parameters have been estimated by the PEs in transformed space, they are applied in real time (by the MEFP or the EnsPost) to generate ensemble members. Finally, the ensemble members are back-transformed into original space (e.g. to streamflow in cubic feet per second). For the MEFPPE, precipitation creates an additional complexity of being a "mixed" variable (i.e. *if* precipitation occurs, then it occurs with a given amount), but that is beyond the scope of this overview.

As indicated above, the HEFS aims to produce ensembles that are physically, as well as statistically, reasonable. Thus, for all temporal and spatial scales of interest, the ensemble forecasts should capture similar patterns in space and time as those *observed* under the same conditions. Within the MEFPPE, the spatial and temporal patterns in temperature and precipitation are preserved by "shuffling" the ensemble members using historical observations. This ensures the relative ordering (ranking) of the members at adjacent times or locations is consistent with those of the observations *on the same dates* in the historical record. The re-ordering technique is known as the "Schaake Shuffle." In the EnsPostPE, the temporal correlations between streamflow amounts at adjacent times are modeled with an "autoregressive" model; this exploits the persistence in streamflow over time.

Alongside the components of the HEFS that produce operational ensemble forecasts, there are tools for hindcasting and verification. Hindcasting is necessary to produce the long and consistent record of historical forecasts needed to properly evaluate the HEFS at particular locations. Without a long record of forecasts from a frozen version of the HEFS, it would be difficult to evaluate forecast quality with reasonably small sampling uncertainty, i.e. with reasonable confidence in the verification results. The Ensemble Verification System (EVS) allows for the verification of the HEFS hindcasts from which guidance can be developed for operational use of the HEFS (e.g. about the conditions under which performance might be impaired and what to look for).

3. Limitations of the HEFS Version 1.0

The HEFS is an operational system and is subject to regular enhancements. These include phased enhancements and bug fixes, which are based on scientific evaluation and software testing. The phased enhancements are implemented in "Development Releases" (DRs). Scientific evaluation requires hindcasting and verification, which are time-consuming and resource intensive. Also, the research-to-operations transition of the HEFS will lead to several novel applications that may require further testing and evaluation. The HEFS Version 1.0 has several known limitations, of which some will be addressed in the planned DRs. The main limitations are summarized in *Table 1*.

Table 1: *key limitations of the HEFS Version 1.0*

Limitation	Potential impact	Plans to address
Limited functionality for quality controlling the HEFS components, including the calibration of the MEFP and EnsPost (i.e. the PEs) and real-time application, i.e. how to identify problematic forecasts	Difficult to tune parameters of the MEFP and EnsPost for particular applications. Relies more on hindcasting and verification, which is time consuming. In real-time application, "problem" forecasts may be difficult to screen	The need for improved diagnostic information will be addressed in DR2 and beyond by adding functionality to the MEFPPE and EnsPostPE (statistics/plots). Real-time screening of HEFS forecasts is a potential future enhancement after HEFSv1

Inability to **explicitly** account for some sources of uncertainty, notably in hydrologic model states and parameters, and in observations. Instead, their effects are accounted for indirectly (by EnsPost)	More difficult to isolate particular problems or sources of uncertainty to be addressed by enhancements. Lumping of different sources of uncertainty runs risk that some aren't properly accounted for and relies heavily on (quality of) observed flow and EnsPost. Reliance on manual MODs rather than data assimilation (DA) could lead to inaccurate model states and improper accounting for uncertainty	No plans to address inaccurate model states until at least HEFSv2 through DA.
Limitations of the simple assumptions made by the HEFS components, notably MEFP and EnsPost, when addressing complex hydrometeorological / hydrologic conditions	Many specific instances, but key examples include river regulations, extreme events and cases where the residuals of the fitted models (MEFP and EnsPost) are not normal. Also, the space-time modeling adopted by the MEFP and EnsPost is quite simplistic	This will be addressed in guidance for the specific components, beginning with DR1
Limited sources of raw forcing forecasts. Currently limited to RFC/HPC, GFS, and CFSv2	Failure to accommodate potentially valuable forcing information, such as forecasts from the SREF. However, this ideally requires a suitable archive of hindcasts	Plan to include GEFS in DR3. Additional sources of forcing information will not be addressed in HEFSv1, The GEFS should significantly improve the medium-range forcing, and hence streamflow, forecasts when compared to GFS
Limited flexibility of the science algorithms. Inability to choose an algorithm for a particular situation, based on guidance	Some scope for "tuning" the MEFP and EnsPost, but limited scope for changing the underlying modeling approach to suit the application. When the assumptions of the MEFP and the EnsPost are not fully met, there are no alternatives to apply	No plans to increase the flexibility of the science algorithms. The provision of a tool box of techniques for the forcing and streamflow was originally planned for HEFSv2

Limited pre-defined products or templates for communicating the outputs from the HEFS. While the GraphGen and the EVS are both flexible, templates are also needed for HEFS products	Potential confusion about how best to communicate the outputs from the HEFS or lack of consistency between RFCs (some of which may be justified)	This is an ongoing effort and will be improved by knowledge of how the HEFS is being applied in practice. This is not, primarily, a software issue, but related to the development of templates and guidance for applying the GraphGen (operationally) and the EVS (for hindcasting)
Limitations of the underlying hydrometeorological and hydrologic models used in the CHPS	This is broad problem. Examples include limitations of the lag/K routing approach, inability of the raw forcing models to capture convection, difficulties in calibrating Snow-17 etc.	No specific plans to address these limitations. In terms of routing, the three-parameter Muskingum model was being investigated for HEFSv2
Limited hindcasting and verification of the HEFS components, as well as “end-to-end” applications	Limited insight into the quality and skill of the HEFS ensembles under varied conditions, including situations where the HEFS performs less well. Limited guidance on how to apply the HEFS in practice	This is currently being addressed through four phases of hindcasting and verification, mainly focused on the different sources of raw forcing information (via MEFP) and the application of EnsPost. The hindcasting and verification will be used to develop improved guidance and build trust in the HEFS
Limited ability to plot large datasets in GraphGen. For example, inability to plot hourly data for ~40 ensemble members for more than ~240 days	Reduced scope for visualization of long-range predictions at an hourly timestep	In practice, it should be possible to visualize the long-range forecasts at reduced frequency (e.g. 6-hourly or daily)

4. Key Steps in Applying the HEFS

There are five key stages in applying the HEFS, namely:

1) configuring the RFC forecast locations to accommodate the HEFS;

2) collecting the data needed to calibrate the HEFS components (including the generation of simulated streamflows);
3) calibrating the HEFS components, notably the MEFP and EnsPost (using the MEFPPE and EnsPostPE, respectively), including quality-control of the estimated parameter values;
4) applying the "end-to-end" HEFS in real time, including the generation of forecast products; and
5) evaluating the HEFS offline through long-term hindcasting and verification (in order to provide guidance for real-time application).

Hindcasting and verification are *not* part of the operational forecasting process and will not be conducted on operational hardware. However, they are central to building trust in the HEFS components, as well as developing verification products and guidance for real-time application. In future, therefore, hindcasting and verification should become routine steps when adding new forecast locations, re-calibrating hydrologic models or otherwise making "significant" changes to the CHPS or HEFS configurations. However, given the limited time and resources available to implement the HEFSv1 at the RFCs, preliminary hindcasting and verification results will be supplied to the RFCs by the HEFS Development Team for selected forecast locations.

The key steps in applying the HEFS are summarized for each software component in *Tables 2-6*. The aim is to provide an overview of the major steps involved in applying the HEFS, rather than detailed guidance or functionality. Detailed guidance can be found in the user's manuals for each component. Blank entries in some columns imply that there are no "assumptions" or "things to watch for."

Table 2: *key stages in applying the MEFPPE*

Steps	Assumptions	Things to watch for
1. Acquire historical data Acquire historical MAP/MAT data from the CHPS database and create historical data files in PI-XML for the MEFPPE to use. The files can be exported manually by the user prior to running MEFPPE or can be exported via a panel within the MEFPPE.	MEFPPE has been installed in a CHPS stand-alone and all needed historical MAP and MAT time series have been imported into the CHPS database. The time series should be the same MAP and MAT time series that are used to drive the standard ESP forecasts at an RFC.	The locations available for parameter estimation in the MEFPPE are based on the historical data made available via the exported PI-XML files. The historical time series can be viewed via the MEFPPE.
2. Process historical data Create faster-access binary files containing historical MAP/MAT data, to be stored with the estimated parameters for access during operational ensemble generation. Also, some data processing is performed for the historical temperature data, computing the historical 24h minimum and maximum temperature values. For precipitation, the historical data is used without change.		The processed historical data can be viewed via the MEFPPE. For MAP data, no processing is performed. Therefore, the processed historical data should be identical to the exported historical data.
3. Create HPC/RFC archive The archive of past QPF/QTF and corresponding observed values is provided in ASCII text files. It can be created using MEFPPE if the data is in the "vfypairs" table of the archive database or manually, external to the MEFPPE, and then imported.	Archives of past QPF/QTF along with corresponding observed values of several years are available.	The archived QPF/QTF should have been created in the past using the same process that is used for current operational forecasts. The time series can be viewed via the MEFPPE. These archives are necessary to estimate the MEFP parameters for the RFC forecast data source.

4. Acquire GFS data Acquire the reforecast data files for the GFS forecast source via the MEFPPE interface.	MEFPPE acquires the data as needed via SFTP.	The time series can be viewed via the MEFPPE. These archives are necessary to estimate MEFP parameters for the GFS forecast data source.
5. Acquire CFSv2 data Acquire the reforecast data files for the CFSv2 forecast source via the MEFPPE interface.	MEFPPE acquires the data as needed via SFTP.	The time series can be viewed via the MEFPPE. These archives are necessary to estimate the MEFP parameters for the CFSv2 forecast data source.
6. Estimate parameters Specify estimation options and estimate the parameters of the MEFP for whichever forecast sources will be used to generate the ensembles operationally. Examine the quality of the estimated parameters to determine their acceptability		Only basic estimation options should be modified by users, initially. Those options are described in the MEFP manual. Default values for advanced options should be used in most cases until experience and understanding has been gained with the science of the MEFP.
7. Accept (zip) parameters Create zip files of parameters to be exported during operational ensemble generation.		MEFP parameters are zipped and stored in FEWS module data set files. Since the MEFP is typically executed for a group of locations at one time, and only one exportDataSetActivity can be specified in a module configuration file, all parameters for a group must be zipped together.

Table 3: *key stages in applying the EnsPostPE*

Steps	Assumptions	Things to watch for
1. Collect and prepare data Collect the historical simulated streamflows (SQIN) and corresponding observed streamflows (QME); for downstream locations, the SQIN should comprise the local and routed contributions, i.e. the total flow. A longer period of record is preferred (e.g. 20+ years). When hindcasting, this should be coordinated with the hindcast period.	QME and SQIN are available at 6-hr or daily time step. Missing values are represented as -999.00 Data files are in in the piXML format	The historical simulations should originate from the same configuration of CHPS that is used operationally It is critical that the time system in the piXML files is correct. Typically, the QME is stored in Data Card format (in local time) and converted to piXML. In converting the files, the data must be shifted correctly and/or the time system correctly identified in the piXML. For SQIN, this is handled in the CHPS configuration, but the “import” configurations related to the time system should be set properly The data in the piXML files should correspond to the period of the record for which EnsPost calibration is desired The “location id” needs to be same in both observed and simulated data files The “parameter id” should be QME for observations, i.e., 24-hr average value, and, for simulations, either SQIN or QINE The “time step unit” and “flow units” should be checked for correctness

2. Evaluate flows Develop annual hydrographs from the QME, and identify dominant seasonality such as wet- or dry- seasons Calculate the error (simulated flow – observed flow) and plot for the different months/seasons Calculate basic verification metrics between simulated flows and associated observed flows to inform calibration choices.	Initially, default values may be chosen for the parameters in the EnsPostPE. Over time, experience should lead to more informed choices Seasons must be formed of consecutive months and the number of months in each season should be similar (for sample size reasons) Verification metrics and errors should be calculated on the daily time scale (i.e. using QME)	Fewer months per season results in fewer samples in that season. This may increase sampling uncertainty/noise
3. Set PE options This includes specifying the seasons and advanced parameter options in the EnsPostPE	While some parameters, such as choice of seasons, are relatively intuitive, other parameters require an understanding of the technical details and calculation of verification metrics of the EnsPostPE ensembles generated for different options. Without that understanding, the default options are preferred	The parameters are pertinent to the skill in the simulated streamflow and the basin hydrology. Parameter choices can, therefore, vary with basin
4. Run the EnsPostPE This estimates the parameter values for the EnsPost	That the correct data for both variables, i.e. QME and SQIN, are uploaded where required	Observed and simulated streamflows must be available for each location Display the QME and SQIN data within the EnsPostPE first. Data in the display should correspond to the period of the record for which the calibration of EnsPost is desired

5. Create zip (parameter) file Create groups and name them, and then zip the parameter files for each group; each group consists of one or more locations. The zip file name will match the group name with “.zip” added as a file extension		The zip file name needs to be consistent with names in the operational configuration
6. Check parameter values Once the parameter values have been estimated, they should be checked for plausibility	The EnsPostPE user documentation provides some guidance, but diagnostic information will be improved with future DRs	Visualize the parameters using the EnsPostPE GUI. Typically the regression coefficient, b, is high for high flow values Parameters for all months within a single season should have the same values
7. Verify EnsPostPE output This is an optional step, but recommended. The EnsPostPE provides corrected ensembles that can be verified directly. By comparing these ensembles for different options of the PE, both against the observed flows and the simulated flows, the options used in the PE may be better informed	This step is time consuming and requires knowledge of how to evaluate the performance of the EnsPostPE ensembles (e.g. using the EVS). However, it might be considered for critical locations or where time permits	

Table 4: *key stages in configuring and running the HEFS operationally*

Steps	Assumptions	Things to watch for
1. Obtain forcing grids Set up a cron to download NCEP's GFS and CFSv2 grids and populate the data files to appropriate CHPS import subdirectories The cron should run once a day to pull the 12Z CFSv2 grids (prate, tmax and tmin) and the 00Z GFS grids	The GFS import directory and import CFSv2 subdirectories have been created	Because of the lag at NCEP, the CFSv2 grid is actually for 12Z the previous day. The GFS grids comprise one forecast per day
2. Import forcing grids Create/modify configuration files to run ImportHEFSGrids workflow in CHPS The new ImportHEFSGrids workflow will import "raw" CFSv2 and GFS grids data to CHPS database and interpolate the grids to cover the extended CONUS domain only	The variables defining the GFS and CFSv2 grid data import directories have been properly set in the CHPS global.properties.file The directories to hold GFS and CFSv2 mapstack files have been created	It may be helpful to separate the GFS/CFSv2 import workflow from general importGrid workflow, so both the GFS and CFSv2 data can be imported successfully into CHPS database
3. Convert forcing grids Create/modify configuration files to run Grid Data Model Adapter in CHPS The Grid Data Model Adapter will convert the GFS and CFSv2 data grids to equivalent ASCII files for use by MEFP	GFS and CFSv2 data for precipitation and temperature has been successfully imported into CHPS	New module instance and new workflow need to be created and registered properly in CHPS configurations

4. Configure/run MEFPPE See steps in Table 2 to run MEFPPE to estimate MEFP parameters for the required locations		
5. Configure/run MEFP Create new module configuration files for MAP and MAT at the required locations; create new workflow for MEFP adapter; register the newly created Module instances and workflows. MEFP will generate the ensemble forecast time series used to force the hydrologic models		
6. Set-up for hydro. run Modify existing streamflow preprocessing configuration files to add MEFP forcing data to the hydro forecast run; create new workflow to run MEFP-hydro to generate operational flow forecast	MEFP forcing data for the selected locations have been generated successfully in the CHPS database	
7. Run EnsPostPE See steps in Table 3 to run the EnsPostPE to generate the EnsPost parameters for the required locations		

8. Configure/run EnsPost Create new module configuration files for the EnsPost Adapter; modify the workflow created in Step 6 to add EnsPost to the MEFP-ESP run so that the flow forecasts will be post-processed	Step 5 and Step 6 have been finished successfully	
9. Visualize results Create/modify configurations to set up GraphGen to visualize the forecast results		

Table 5: *key stages in configuring and running the HEFS in hindcast mode*

Steps	Assumptions	Things to watch for
1. Collect data Collect historical streamflow (QME), precipitation (MAP), and temperature (MAT) observations, and simulated streamflow (SQIN) data for each test location. A record of 20+ years is ideal.	MAP and MAT data are available on 6-hrly time step; QME and SQIN data are available on daily or 6-hrly time step. Ideally: a) there are no missing data; and b) all data files are in PI-XML format.	When importing these data into CHPS, make sure the time step in the "Import" configuration files are consistent with the time step of the data.
2. Create Forcing Hindcasts Run MEFP to produce required temperature and precipitation forcing for targeted HEFS hindcasting period (e.g. climatology, GFS, GEFS, and CFS). Collect forcing from other sources (e.g. RFC/HPC) if required.	The latitude and longitude information for the centroid of the test location is available (which is required by MEFP). MEFP is calibrated (refer to Table 2 for detailed description).	If there are missing climatology, GFS, GEFS, or CFS data, MEFP will not produce the corresponding forcing hindcasts during those missing periods. For the RFC/HPC forcing, the data must faithfully represent the operational forcing conditions if using the hindcasts to guide operations.
3. Create files to import MEFP hindcasts The HEFS hindcasting run is conducted in batch mode for multiple T0s. It is necessary to import the MEFP hindcasts at the start of hindcasting for each T0. The following script/configuration files need to be created for this purpose: a) a shell script to copy MEFP hindcasts into CHPS import directory; b) a module file to call this script and import the data into CHPS; c) an import workflow to call this module.	MEFP hindcasts are produced and stored in the designated directory.	Make sure the ensemble ID index in the MEFP hindcasts are consistent with its counterpart in the forecast configuration files (refer to Step 5); the new workflow and module file should be registered.

4. Create/adapt UpdateStates configuration These files are required to generate initial conditions for the hindcasting run. The following changes are necessary: a) add test locations to a new locationSet; b) make a copy the UpdateStates workflow corresponding to the forecast group (which contains the test locations), rename it appropriately; c) in the renamed UpdateStates workflow, change the locationSet containing the forecast group to the locationSet created in 4a.	The historical MAP, MAT data should have been imported into CHPS before running the UpdateStates workflow, as these data serve as input to the UpdateStates workflow.	The new workflow and module files should be registered appropriately.
5. Create/adapt hindcast configuration files a) make a copy of the forecast workflow corresponding to the forecast group, rename it appropriately, update the locationSet accordingly; b) make a copy of the MergeMAP(/MAT) module file corresponding to the forecast group, rename it appropriately, change the locationID to be the locationSet defined in 4a; c) create a module file for each test location to export hindcast data produced for the location; d) create an ESP export workflow to call the export module(s) from 5c.	Initial states are generated for the hindcast run.	In the new forecast workflow, the ensemble ID index should be consistent with that of the original MEFP hindcasts. All new workflow and module files should be registered appropriately.

6. Create EnsPost configuration files a) create a module file for each test location to run EnsPost; b) create a module file for each test location to export EnsPost-processed data; c) create an EnsPost export workflow to call the export module(s) from 6b.	EnsPost is calibrated (refer to Table 3 for details). The calibrated parameter files are stored in a specific directory.	All new workflow and module files should be registered appropriately.
7. Configure the EVS If running the EVS within CHPS (rather than standalone): a) create a project file for each test location (refer to Table 6 for details); b) zip the project file and the observed streamflow data at this location into a new file, move the zipped file to the designated directory; c) create an EVS module file to execute EVS for each test location (run EVS and export verification results); d) create an EVS workflow to call the EVS module files		The location identifier and variable identifier in the EVS project file should match those in the PI-XML files exported by CHPS. All new workflow and module files should be registered appropriately.
8. Augment forecast workflow for hindcast run Add the MEFP import workflow (3c), ESP export workflow (5d), EnsPost module files (6a), and EnsPost export workflow (6c) to the forecast workflow (5a)	All relevant files to be added are produced earlier	Import workflow, forecast module for upstream and downstream locations, ESP export workflow, EnsPost module files, and EnsPost export workflow should all be allocated in sequence in the forecast workflow

9. Generate initial conditions for the hindcast a) execute an import run (import MAP, MAT, QME, and SQIN) and a pre-process run (to produce MAPE) to prepare data for the update states run; b) run the UpdateStates workflow (4b) in batch mode during the hindcast period		Warm up the model (using at least a 2-year period) before running the UpdateStates workflow in batch mode
10. Generate streamflow hindcasts and EnsPost-processed data Execute a forecast run (run workflow created in Step 8 in batch mode) to produce streamflow hindcasts and run EnsPost to post-process the hindcasts, and export both datasets (before and after applying EnsPost)	Initial states are produced in Step 9	Define the forecast length appropriately (consistent with those of the MEFP hindcasts)
11. Visualize streamflow hindcasts and EnsPost-processed data Visualization may be conducted in the CHPS database viewer or GraphGen		
12. Run EVS workflow If running EVS workflow, execute a verification run (run workflow created in Step 7d) to produce verification products for streamflow hindcasts at test locations		Define the parameter ID appropriately (corresponding to the variables to be verified) in EVS module files (7c)

Table 6: *key stages in configuring the EVS to verify the HEFS hindcasts*

Steps	**Assumptions**	**Things to watch for**
1. Collect data Collect data for each variable (e.g. precipitation), forecast location and scenario (e.g. streamflow with EnsPost) to be verified. Are "sufficient" data available? Generally hindcasting will be used to provide a long record (5+ years), ideally 20+ years	Data are available. Precise data needed depends on what is being verified. At least forecasts and observations for pairing. Other variables may be used to "condition on" (e.g. investigate quality of precipitation forecasts when temperature is below freezing). If evaluating skill, the forecasts for the baseline are also needed (e.g. "ESP")	If forecasts and observations are measured at different times or cover different control volumes (e.g. 6-hours versus daily), a strategy is needed to pair the forecasts and observations
2. Create EVS project file Use a template or existing EVS project file or start from scratch (create a new project)	When running the EVS in standalone mode, no CHPS configuration is required. When running in CHPS mode, the project file is zipped and placed in an appropriate location for CHPS to access (see Table 5)	Generally, it is simpler to start with an existing EVS project file or "template" because many verification studies are similar (e.g. just applied to different locations). The HEFS Development Team can provide templates.
3. Add Verification Units A Verification Unit (VU) is required for each variable, location and scenario to verify. The VU identifies the location of the data, the time-scales to be verified, and the metrics to compute, among other things. An EVS project file generally contains several VUs		Verification with the EVS can be very time-consuming, depending on the size of the dataset being verified. Thus, careful thought about the aims of the verification study can save considerable time overall

4. Subset data if needed For each VU, it is possible to apply conditions to extract subsets from the overall dataset. For example, the data may be broken into particular seasons and verification conducted separately for each season	That sufficient data are available to compute the verification metrics for the subsets of data identified, otherwise the sampling uncertainties may be large	Large sampling uncertainties (i.e. "noisy" or misleading verification results) when using small sample sizes
5. Configure metrics In general, it is necessary to compute several verification metrics. One metric cannot provide a complete picture of forecast quality. However, some metrics are more suitable to particular problems. The EVS metrics are grouped into single-valued, ensemble and skill metrics	Requires some preparatory thought about the verification metrics that should be computed (i.e. what is the aim of the verification study?) and any thresholds that should be used to compute them (e.g. flood thresholds). Is it necessary to quantify the sampling uncertainty (e.g. small sample size)?	Consider the computational time when choosing a large number of metrics and computing them at a large number of thresholds
6. Run the verification Two steps are conducted by the EVS when running a VU: 1) the forecasts and observations are paired; and 2) the verification metrics are computed with the paired data.	That forecasts and observations are available at the same times and for the same accumulation volumes, otherwise the pairs will not be computed. That "sufficient" data are available for verification (generally several years)	Verification may be time consuming and CPU/memory intensive. When applying the EVS to large datasets, it may be necessary to increase the maximum memory allocated to the EVS **before** start-up (see start-up options in the user's manual)
7. Check verification pairs All verification results from the EVS reflect the verification pairs that were computed by the EVS. It is, therefore, critical to check some of these pairs against the raw data before relying on the results from the EVS	That the pairing has been conducted correctly. Without correct pairing of the forecasts and observations, the verification results will be meaningless.	The verification pairs are stored in an XML format with times in UTC. The pairs reflect any aggregation requested when defining the VU needed to do the pairing (e.g. aggregation of 6-hourly forecasts to pair with 24-hourly observations)

8. Display/interpret results The verification results may be viewed outside of the EVS (using the written outputs) or with the interactive viewer inside the EVS GUI Interpreting the verification results requires time and practice. The EVS user's manual provides some guidance on the meaning of the different metrics. Some metrics are more intuitive than others. For example "skill scores" show the relative quality of one forecasting system (e.g. MEFP) given a baseline (e.g. climatology)	Some familiarity with the verification metrics available in the EVS and with ensemble verification more generally. An awareness of the application and audience for the verification results.	Some metrics provide relative measures of quality and others are expressed in forecast units (e.g. CFS for streamflow). With the latter, some care is needed, because the "meaning" of these units will vary substantially between locations. Take care with interpreting results that are based on small sample sizes. The numerical (XML) outputs from the EVS provide sample sizes. Sampling uncertainties can also be evaluated explicitly, but this is time consuming

5. Glossary of Terms, Acronyms and Abbreviations

Weather and Hydrology

1. **Climatology** – The science that deals with the phenomena of climates or climatic conditions. Climatology also refers the historical record of observations (e.g. mean areal averages of actual temperature and precipitation) used to drive a model
2. **Canonical Event** – a partitioning of time scales in order to account for the varying skill of the different forcing inputs to MEFP (e.g., RFC QPF/QTF, GFS, and CFSv2)
3. **MAP** – Mean Areal Precipitation over a "basin"/watershed
4. **MAT** – Mean Areal Temperature over a "basin"/watershed

Weather Data

1. **.grib** file - A binary file format designed to store large amounts of gridded data (a large time series of grids received from NCEP)
2. **pedtsep** – A sequence of letters that identifies a type of data; In the Standard Hydrologic Exchange Format (SHEF), different types of data are keyed by a seven-character parameter code represented by the character string "PEDTSEP". This string is broken down as follows:

   PE = Physical Element (precipitation, gage height, temperature, etc.)
   D = Duration Code (instantaneous, hourly, daily, etc.). The duration code character (D) combined with the physical element (PE), describe the vast majority of observed hydrometeorological data. The duration code describes the period to which an observed or computed increment applies
   T = Type Code (observed data, forecast data, etc.)
   S = Source Code (further refines the type code which may indicate how data was created or transmitted
   E = Extremum Code (maximum value, minimum value, etc.)
   P = Probability Code (Chance value is at/below the specified value, e.g., 90%, 10%)

Example: 6-hour precipitation would be encoded PPQ, where **PP**Q represents **incremental precipitation** and the PP**Q** represents a **6-hour** duration

For more, see: https://ocwws.weather.gov/intranet/whfs/SHEF/Explain_duration2.shtml

3. **SHEF** – Standard Hydrologic Exchange Format. A standard ASCII format for exchanging data at the National Weather Service (NWS)

Models, Algorithms and Techniques

1. **Aggregation and Disaggregation** – forming larger or smaller control volumes, respectively
2. **Bias** – A systematic difference between an estimate of some quantity and its "true" value (generally meaning observed)
3. **Calibration** – A process of estimating model parameters based on observations and corresponding (raw) predictions. In pre- and post-processing, calibration has a second meaning, namely to correct probabilistic biases in ensemble forecasts by increasing their reliability
4. **CFS/v2** – Climate Forecast System. A fully coupled model representing the interaction between the Earth's oceans, land and atmosphere that can generate a forecast for 45 days, a full season, or 9 months. See also: http://cfs.ncep.noaa.gov/
5. **Disaggregation** – (see aggregation/disaggregation)
6. **Forcings** – The model inputs (e.g., precipitation and temperature) that drive/"force" a hydrologic model
7. **Ensemble Forecast** – A collection of equally likely predictions of the future states of a hydrologic system, based on sampling of the different sources of uncertainty and propagating them through a hydrologic modeling system (such as CHPS). An "ensemble trace" comprises two or more forecast lead times

8. **EnsPost** – Ensemble Post-Processor. A statistical technique that accounts for hydrologic uncertainties and biases separately from the forcing uncertainties and biases
9. **ESP** – Ensemble Streamflow Prediction. In NWS operations, this has the specific meaning of forcing the NWS River Forecast System with a sample of observations from the same dates in the past, i.e. climatological forcing. Some RFCs have augmented the original ESP algorithms to account for additional information
10. **Forecast Issue Time** – The date/time at which a forecast is initiated, also known as "T0." This differs from the Forecast Valid Time
11. **Forecast Valid Time** – The time at which a forecast is valid
12. **Forecast Lead time** – The difference between the Forecast Valid Time and the Forecast Issue Time
13. **GFS** - Global Forecast System. One of the operational forecast models run at NCEP. The operational GFS is run four times daily, with forecasts out to 384 hours. The GFS was also "frozen" in 1997 (the "frozen GFS") and used to generate hindcasts (i.e. retrospective forecasts) beginning in 1979, which are used to calibrate the MEFP
14. **Global Ensemble Forecast (GEFS) system** – an enhanced version of the GFS that produces ensemble forecasts
15. **iyr –** initial year
16. **lyr –** last year
17. **Hindcast** - a retrospective forecast or reforecast. A forecast where the issue time (T0) is in the past, based upon the conditions at the chosen T0, but using a current model (which may not have been available on the original forecast date). Reforecast is a term frequently used for weather models
18. **MOS** – Model Output Statistics
19. **NAM** - The operational North American Meso (NAM, formerly Eta) is run four times per day (00,06,12,18Z), all cycles run to 84-h

20. **NGM** - The Nested Grid Model, a 48-hour numerical model of the atmosphere run twice daily by NCEP.
21. **NWP** - Numerical Weather Prediction
22. **PoP** – Probability of precipitation
23. **pobs –** Precipitation Observed
24. **pfsct –** Precipitation Forecast
25. **Reforecast** – See Hindcast
26. **Simulation** – A hydrologic model run that uses observed forcings
27. **SREF** – Short-Range Ensemble Forecast (SREF) system. An NCEP model that issues short-range ensemble forecasts
28. **System Time, a.k.a., T0/forecast time** – The date/time selected to begin a forecast
29. **Skill** – A measure of relative quality of a forecast system with respect to a baseline. The measure used for skill could vary (e.g. the mean error of one system relative to another)
30. **T0 – issue time, also known as the System/Forecast/Basis Time** - The date/time at which a forecast is initiated
31. **ts** – time series. A list (array/vector) of times with temperature/precipitation forecasts/measurements provided at each point in time

Mathematics/Statistics/Probability

1. **Bivariate Correlation** – Correlation between two variables
2. **Brier Score (BS)** – the average squared deviation between the predicted probabilities that a discrete event occurs (such as flooding) and the observed outcome (0 or 1)
3. **CRPS** – Continuous ranked probability score. The integral square difference between a forecast probability distribution and the observed outcome. It is typically averaged over many such cases (known as the "mean CRPS")

4. **CDF** – Cumulative distribution function

5. **Normal/Gaussian Distribution** – A simple, theoretical, probability distribution with two parameters (mean and standard deviation). The multivariate normal distribution, which describes several forecast times, locations or variables, is completely defined by a vector of means (one for each variable) and a covariance matrix

6. **NQT** – Normal Quantile Transform. A transformation made to a data sample so that it follows a normal probability distribution (i.e. so that the histogram of values would appear normal)

7. **Probability Distribution** – a function that describes the probability of each possible event associated with a random variable. A discrete random variable, such as the possibility of flooding, is described with a discrete probability distribution. A continuous random variable, such as temperature, is described with a continuous probability distribution or probability density function (see entry for PDF). A mixed random variable, such as precipitation, is described with a mixed probability distribution (i.e. precipitation cannot be negative)

8. **PDF** – Probability Density Function. A probability distribution for a continuous random variable, such as temperature. Probability density provides a *relative* measure of probability (a density). The actual probability of falling between two values is determined by integrating the PDF between those values

9. **RPS** – Ranked Probability Score. An extension of the Brier Score to several discrete probability categories (such as low, medium and high flows). Extension to all possible categories of a continuous variable is equivalent to the CRPS

OHD and NWS Systems

1. **AWIPS** – Advanced Weather Interactive Processing System

2. **CHPS** (pronounced “chips”) – Community Hydrologic Prediction System.

3. **EnsPost** – Ensemble Post-Processor

4. **EPP3** – Ensemble Preprocessor – A (Fortran) pre-cursor to MEFP (Java)

5. **EVS** – Ensemble Verification System
6. **ESP** – Ensemble streamflow prediction
7. **FEWS** – Flood Early Warning System. Developed by a company in the Netherlands, Deltares (formerly Delft) and written in Java. See also CHPS
8. **IFD** – Interactive Forecast Display (FEWS GUI)
9. **IVP** – Interactive Verification Program
10. **HEFS** – Hydrologic Ensemble Forecast Service
11. **MEFP** – Meteorological Ensemble Forecast Processor. A (Java based) rewrite of EPP3 (Fortran)
12. **NWSRFS** – National Weather Service River Forecast System. Replaced by CHPS
13. **PI** – Published Interface
14. **RAX** – RFC Archive Database – An archive of RFC forecasts and observed data stored in a Postgres database
15. **SREF** – Short-Range Ensemble Forecast system
16. **WHFS** – WFO Hydrologic Forecast System
17. **XEFS** – Experimental Ensemble Forecast System. The experimental precursor to HEFS

Organizations/Companies

1. **CPC** – Climate Prediction Center
2. **Deltares** (formerly Delft) – Netherlands company that developed FEWS which is “wrapped by” CHPS
3. **HPC** – Hydrometeorological Prediction Center
4. **HSD** – Hydrologic Services Division
5. **HDSB** – Hydrologic Data Systems Branch

6. **HSEB** – Hydrologic Software Engineering Branch. Part of the Office of Hydrologic Development (OHD)
7. **HSMB** – Hydrologic Science and Modeling Branch. Part of the Office of Hydrologic Development (OHD)
8. **NCEP** – National Centers for Environmental Prediction
9. **OCWWS** – Office of Climate, Water, and Weather Services

Software/Systems Development

1. **ASCII** – American Standard Code for Information Interchange.
2. **DR** – Development Release
3. **Fortran** – A general-purpose, procedural, imperative programming language
4. **GUI** – Graphical User Interface
5. **IR** – Interim Release
6. **SOA** – Service Oriented Architecture. An approach to developing software that emphasizes developing software in the form of interoperable services.
7. **XML** – eXtensible Markup Language. XML is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.

   See also:

   http://www.weather.gov/glossary/

   http://www.crh.noaa.gov/dtx/glossary.php

   http://www.nws.noaa.gov/mdl/synop/acronyms.php

   Also, consider using the search feature on any of the above web sites.