

Hydrologic Ensemble Forecasting Service (HEFS)

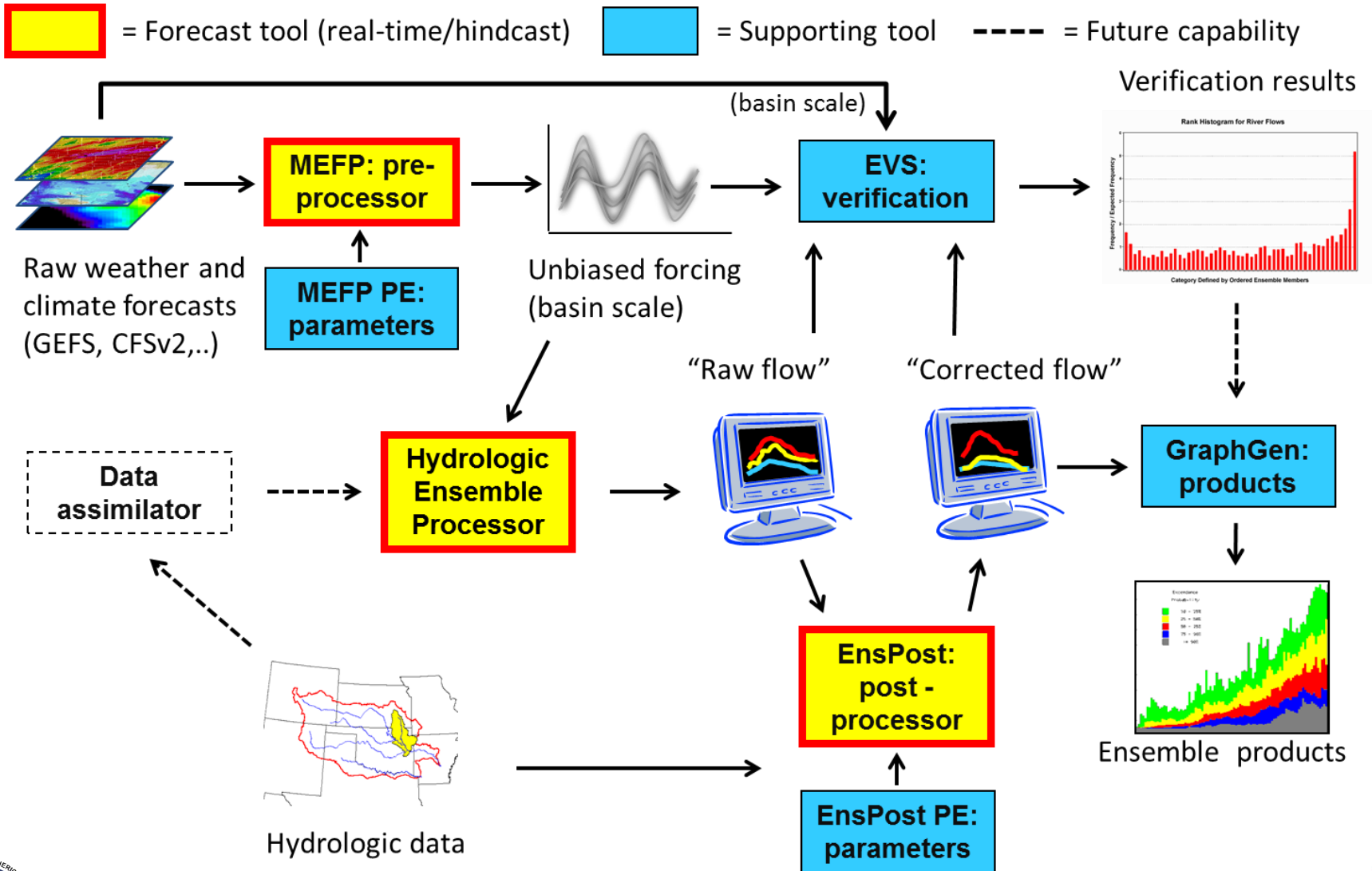
Seminar B What's New for HEFSv1

Hank Herr

HEFS Training Workshop (4), Silver Spring, MD
September 19, 2013



HEFS Components



Overview of HEFSv1 Changes

❑ Diagnostics

- o Tools have been added to allow for better QC of historical and RFC archived forecast/observed time series used as input to MEFPPE
 - Use MEFPPE to identify bad data
 - Use external tools to fix bad data
- o MEFPPE parameter diagnostic tools have been added
 - Based on earlier XEFS EPP3 work with John Schaake and Rob Hartman
 - Working on guidance for how to use diagnostic tools in decision making with MEFPPE

❑ Other changes include configuration changes, bug fixes, and enhancements related to the science and software of HEFS

- o Behavior of and output from MEFPPE, EnsPostPE, MEFP, and EnsPost all change as a result of this release
 - Example: By default, modulation events are no longer included when MEFP generates ensembles

Required Actions for HEFSv1

❑ Perform configuration changes

- o May have already been done
- o Read the release notes
- o Summarizing list provided in Appendix of these slides

❑ Re-estimate all parameter files

- o Use your existing parameter estimation stand-alone (with updated jars)
- o Perform QC for historical and RFC archived forecast/observed data
 - Described later
- o MEFP temperature parameters
 - GEFS reforecast data read incorrectly in previous release (FogBugz 1166)
- o MEFP precipitation parameters
 - CFSv2 used fixed 'Data Window' in previous release (FogBugz 1213)
- o EnsPost parameters
 - Bug fix related to how CDFs are computed in the EnsPostPE

Required Actions for HEFSv1

- ❑ **Configuration changes when using raw climatology**
 - CHPS transformations should be used to append raw climatology to the end of MEFP generated ensembles
 - Instructions in *MEFP Configuration Guide: Forecast Components Section 6.1.3*
 - Resampled climatology will still be generated using MEFP
 - Applies to MARFC and NERFC for NYCDEP
- ❑ **MEFP adapter run file property change if ensemble members are constructed based on calendar year instead of water year**
 - Set the memberIndexingYear property to be “calendarYear”
 - Hydrologic water year (Oct 1 – Sep 30) is the default setting
 - Matches standard ESP climatology forecasting
 - Applies to MARFC and NERFC for NYCDEP
 - To be consistent with the HEFS hindcasts generated at OHD

Things to Watch Out For

- ❑ **The default settings for a new adapter run file property behaviorIfEventMissing changes the behavior of MEFP**
 - If a canonical event cannot be computed for a source using the input time series provided due to missing data, an error is thrown
 - More sensitive to missing data
 - Old data more likely to yield errors in adapter run due to missing values at end of the time series
 - The option 'fillClimatology' will make it run as before, but then it is easier for bad/old data to be processed by MEFP and go unnoticed
- ❑ **Preprocessing workflows (import and spatial interpolation) may not error out if data to import is not provided**
 - Ex: GEFS workflow will use old data in spatial interpolation if import module fails to import any data
 - Error will not occur until MEFP uses the data as input (see above)

Time Series Diagnostics Why?

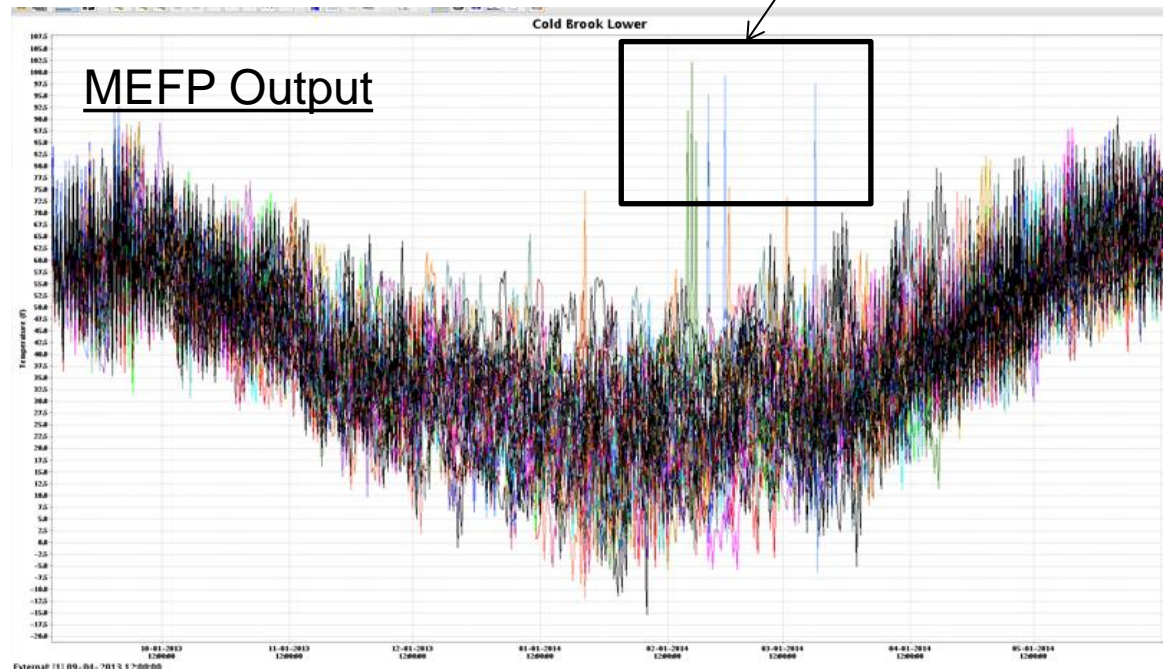
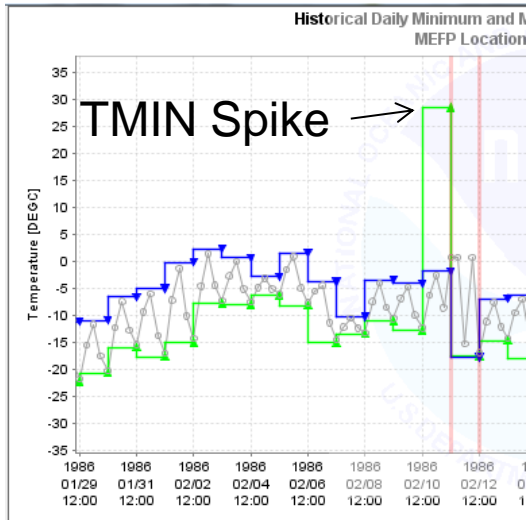
MTRN6LWR Datacard Snippet

1828146	27.135	13.485	3.495	33.280	33.280	33.280
1828147	39.500	33.280	33.280	33.280	33.280	33.280
2828148	33.585	33.280	33.280	33.280	33.280	33.280
2828149	33.280	33.280	36.708	33.280	33.280	33.280
2828150	33.280	33.280	33.280	33.280	33.280	33.280
2828151	33.280	33.280	33.280	33.280	4.869	33.280
2828152	33.280	33.280	-2.398	33.280	33.280	33.280
2828153	7.537	33.280	33.280	33.280	33.280	33.280
2828154	33.280	33.280	33.280	33.280	33.280	33.280

Many repeated 33.280 values
Does not fit diurnal pattern

Unreasonable
Spikes

Historical TMIN/TMAX



Time Series Diagnostics

Why?

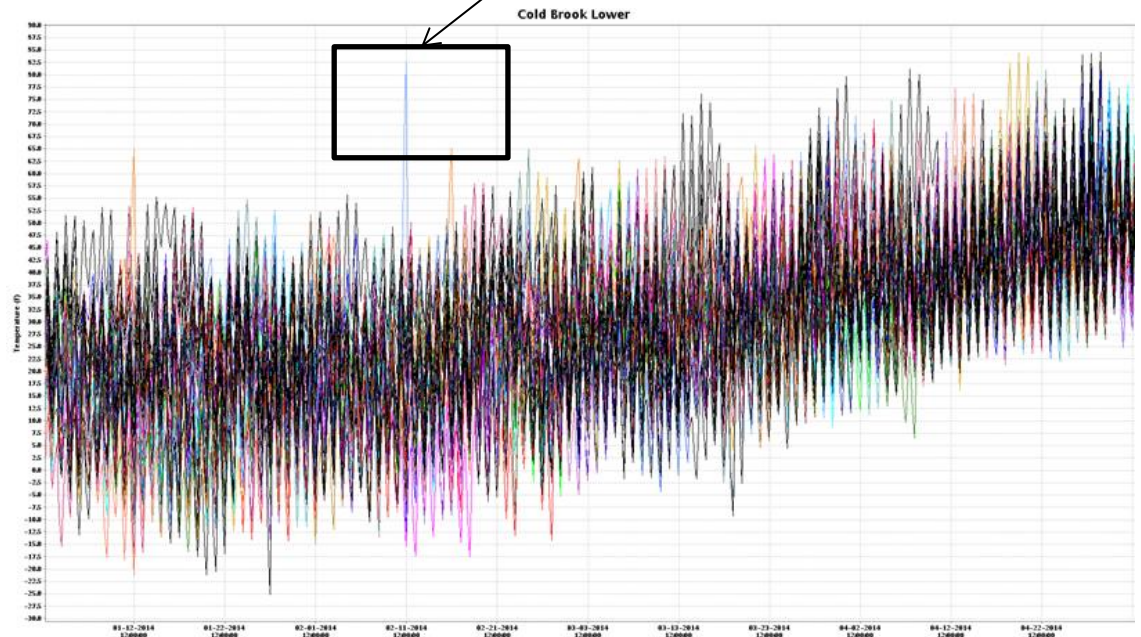
MTRN6LWR Datacard Snippet

1828146	27.135	13.485	3.495	19.469	26.166	12.036
1828847	5.417	16.394	23.322	15.514		
2828148	10.707	21.242	27.728	21.349	18.255	22.532
2828149	24.357	9.755	3.070	11.340	15.978	0.550
2828150	-6.595	10.380	20.414	4.162	-3.521	10.214
2828151	19.079	12.559	9.330	16.076	19.620	7.120
2828152	1.337	10.990	16.736	5.610	0.384	8.048
2828153	12.795	5.378	1.866	11.441	17.117	8.858
2828154	4.936	18.668	26.480	12.085	5.312	22.893

Copied from another year to fit the diurnal pattern

MEFP Output

Only one spike left



Time Series Diagnostics (Demo)

❑ Quality control historic/archive data

- o Applicable to **Historical Data Panel** and **RFC Forecasts Panel** within the MEFPPE
- o Checks are made when the steps are performed (or data imported):
 - Missing data
 - Gross range (0 – 100 mm for precip, -100 – 100 degC for temperature)
 - Minimum temperature exceeds maximum temperature
 - Check results are stored in a file within the MEFPPE run area
- o **Location Summary Panels** are enhanced
 - Icon has been added to display for which locations historical or RFC archived forecast data is questionable
- o **Diagnostics Panel** is enhanced
 - Questionable data is highlighted by light red marks/zones within graphic
- o GraphGen chart viewing panel enhanced to allow for more easily navigating and examining questionable data values

Time Series Diagnostics (Demo)

□ Quality control steps

- o Upon installing and configuring HEFSv1 in the parameter estimation stand-alone, perform the steps associated with the **Historical Data Panel** and **RFC Forecasts Panel** again so that MEFPPE can detect questionable data
- o Examine the **Location Summary Panels** to identify locations that have questionable data
- o Use the **Diagnostics Panel** to examine historical data and archived RFC forecast data to determine if questionable data is bad data
- o Replace bad data with reasonable data values at the source
 - Historical time series
 - Modify datacard files and re-import the data into CHPS –or– use the FEWS Data Editor:
<https://publicwiki.deltares.nl/display/FEWSDOC/04+Data+Display+and+Data+Editor#04DataDisplayandDataEditor-DataEditor>
 - RFC archived forecast time series with observations
 - Modify data in archive database
 - Copy RFC forecast files, modify by hand, and import the files
- o Perform step again to incorporate fixed data (or import RFC data)

Estimated Parameters Diagnostics (Demo)

- ❑ **Diagnostics added to allow for a cursory examination of estimated parameters**
 - Based on diagnostics employed by John Schaake and Rob Hartman
 - Block plot displays parameters in three dimensions (day of year, event, and parameter value)
 - Parameter value displayed based on color
 - Table view also available with same color coding for background of cells
 - In addition to viewing raw parameter values, the following are allowed:
 - Displaying difference between observed and forecast event means
 - Displaying difference between correlations across forecast sources
 - Clicking on blocks or table cells opens up a scatter plot of the raw canonical event values used in parameter estimation
 - Tools provided for navigating the chart and table and for selecting events for which to display values

Questions?



Appendix

List of Changes for HEFSv1

(see the release notes for a complete list with full details)



Changes for HEFSv1

MEFP/MEFPPE Configuration Changes

- ❑ **Grids.xml**
 - o <y> changed to 50 from 56 (FogBugz 1086)
- ❑ **MEFP_MAP_to_GMT.xml**
 - o Set <ignoreMissing> flag to false on sample transformation
- ❑ **MEFP_MAT_to_TAMN_TAMX.xml**
 - o Set <ignoreMissing> flag to false on sample transformation
- ❑ **MEFPPE.xml (PI-service configuration)**
 - o Includes GMT translated MAT time series query
- ❑ **MEFP_GEFS_TFMN_6to24.xml, MEFP_CFSv2_TFMN_6to24.xml**
 - o Fixed so that minimum is being computed (FogBugz 1195)
- ❑ **FGroup_MEFP_CFSv2*_LaggedEnsemble.xml**
 - o Removed some configuration causing warnings (FogBugz 1202)

Changes for HEFSv1

MEFP/MEFPPE Configuration Changes

- ❑ **MEFP_Forecast.xml, MEFP_FMAT_Forecast.xml (workflows)**
 - Changes made to allow for running diurnal pattern module in ensemble mode using multiple processors to improve performance

- ❑ **IMPORTANT: Raw climatology option in MEFP will be removed in the near future**
 - For HEFSv1, resampled climatology will be the default
 - Reason:
 - MEFP does not output true raw climatology for temperature (MAT)
 - Capability is already available through CHPS transformations
 - It is recommended that CHPS transformation be used to append raw climatology to MEFP output if raw must be used
 - See Section 5.1.3 of the *MEFP Configuration Guide: Forecast Components*

Changes for HEFSv1 Bug Fixes

☐ MEFP/MEFPPE

- o Fixed problem involving not removing tmp files (FogBugz 1091)
- o Fix to lessen chance of run-time information file being partially written when CHPS/MEFPPE crashes for any reason (FogBugz 1093)
- o Included missing jar in release (FogBugz 1119)
- o Logic problem due to the first day of the year with successfully computed parameters being the last day (FogBugz 1124)
- o RFC forecast source fixed to include correct unit in files constructed from archive database data (FogBugz 1133)
- o TMAX GEFS data being assigned to TMIN and vice versa (FogBugz 1166)

☐ EnsPost/EnsPostPE

- o Bug fix related to how CDFs are computed in the EnsPostPE
- o Bug fix related to how the CDFs and parameters are applied in EnsPost

Changes for HEFSv1 Small Enhancements

□ MEFP/MEFPPE

- o Proper default start/end years assigned to the RFC, GFS, GEFS, and CFSv2 forecast source estimation options (FogBugz 1094)
- o Improved messaging related to failures in estimating parameters (for example, insufficient data) (FogBugz 1106)
- o Added Default Button for each estimation option; clicking the button recovers the delivered default setting (FogBugz 1186)
- o Run file property memberIndexingYear added (FogBugz 1200)
 - Other run file properties also added (FogBugz 1213)
- o Various other changes tracked in FogBugz 1213
 - Modulation events removed from default canonical event list and are not included by default when MEFP generates ensembles
 - behaviorIfEventMissing flag added to adapter properties (defaults to erroring out if an event cannot be computed)
 - CFSv2 pairing window uses standard estimation option instead of fixed window

Changes for HEFSv1 Small Enhancements

□ EnsPost/EnsPostPE

- o Added an option to allow for outputting the model time-scale 24-hour results from EnsPost, instead of the 6-hourly disaggregated values (FogBugz 1159)
- o Various changes tracked in FogBugz 1191