

Memorandum

To: Science Advisory Committee
From: Dan Opdyke, TPWD
Re: Hydrology-Based Environmental Flow Regime (HEFR) Potential Enhancements and Modifications
Date: January 21, 2010

This memorandum contains potential enhancements and modifications to HEFR that were gleaned from the recent HEFR survey, as well as discussions with HEFR users over the past several months. Enhancements are loosely grouped by category. The order of proposed enhancements within each category is random.

Functionality

1. Enhanced Functionality for Intermittent and Ephemeral Streams

Issue

As the SB 3 process moves west, more and more locations will be characterized by intermittent and ephemeral flows. Despite not always having flowing water, many of these locations will be judged to have a sound ecological environment. It is unclear if the standard outputs from HEFR will be satisfactory for these locations.

Proposed Solution

It is not immediately obvious what functionality should be included into HEFR. Ideally, a group of interested parties from the drier portions of the new basins (Nueces, Guadalupe-San Antonio, and Colorado-Lavaca) would be convened to discuss possible proposals. This could only succeed if there is enough interest among the parties involved. Ideas are welcome.

At a minimum, HEFR can be coded to output the historical frequency of zero flows and statistics describing the duration of zero flow events and intervals between events.

Anticipated Effort

Unknown. Simple statistics on zero flow events would require approximately 32 hours of coding and 6 hours of writing for the HEFR report.

2. Link Wet/Average/Dry Statistics Explicitly to Wet/Average/Dry Periods

Issue

Currently, HEFR statistics for wet, average, and dry hydrologic conditions are all calculated from the entire period of record. For implementation scenarios where the wet,

average, and dry flow recommendations only apply during defined wet, average, and dry conditions (as in the Sabine-Neches BBEST report), it is tempting to want to recode the statistics so that the wet, average, and dry recommendations are explicitly based on statistics drawn from historical wet, average, and dry periods, respectively. HEFR development staff has had numerous conversations on this topic with multiple parties in support of this idea, yet remain unconvinced that this is a clear step forward. This is an important conceptual and coding decision that would benefit by explicit discussion of the various advantages and disadvantages with input from the SAC and interested parties.

Note that explicit calculation of wet, average, and dry statistics based on historical wet, average, and dry periods would be antithetical to the two instream flow approaches proposed by the Trinity-San Jacinto BBEST, as both approaches recommended simple flow magnitudes and long-term attainment frequencies. For these two approaches, the concepts of wet, average, and dry are used as convenient nomenclature only and have no practical effect on the recommendations.

Proposed Solution

I am agreeable to preparing a short memo describing perceived disadvantages of this proposal. If someone is willing to develop a companion memo of advantages, a useful discussion could be held at a SAC meeting to establish a path forward.

Anticipated Effort

Unknown. Recoding HEFR to allow for optional binning of years or seasons into wet, average, and dry categories prior to statistical computations could be complicated, as this would impact all subsequent calculations.

3. Additional Calculation Options for Subsistence Flows

Issue

HEFR currently calculates subsistence flow recommendations based on a summary statistic (the default is the median) of the subsistence flows identified in the hydrographic separation step. HEFR also allows the user to input a “floor” for subsistence flows, based on outside information, e.g., related to maintaining water quality. This floor can be the 7Q2 or other metric. The subsistence flow calculations had a relatively low average score of 3.5 out of 5 in the HEFR survey. Additionally, several members of the Sabine-Neches BBEST advocated the use of the 5th percentile flow for subsistence flows; this statistic has been used for subsistence flows in the literature, including in the Colorado River LSWP study.

Proposed Solution

The 5th percentile (sometimes called Q95, in reference to its equivalent definition as the 95th percent exceedence level) and other simple low-flow statistics (7Q2, 7Q5, and 7Q10 were suggested by a survey respondent) could be added as options in HEFR.

Anticipated Effort

Eight hours to poll users for recommendations on relevant statistics. Twenty-four hours of coding.

4. Allow for Varying Percentiles of Volume, Duration, and Flow in the Percentile-Based Approach for Episodic Events.

Issue

A survey respondent suggested that HEFR should be recoded to allow user-input percentiles not just for wet, average, and dry (which already is allowed in HEFR), but also for the three episodic event characteristics of volume, duration, and peak flow. For example, the user might want to specify a wet-level high flow pulse using the 50th percentile for volume, the 40th percentile for duration, and the 65th percentile for peak flow.

Proposed Solution

This could be coded into HEFR. It is unknown how often users would have sufficient information to specify these variables, yet still need HEFR to generate environmental flows.

Anticipated Effort

Sixteen hours of coding. Limited edits to the HEFR report.

5. Calculation of Stream Power

Issue

A survey respondent suggested that HEFR calculate stream power and other engineering calculations.

Proposed Solution

Stream power and other potential engineering calculations could be explored and added to HEFR. Such equations would likely require additional input from the user, e.g., stream slope.

Anticipated Effort

Unknown, but probably modest – if/when appropriate equations can be identified.

6. Duration of Subsistence Flows

Issue

In BBEST deliberations, the duration of historical and future periods at and below the subsistence flow level has been discussed. HEFR currently provides no information on the historical duration of continuous periods of flows at and below subsistence flows.

Proposed Solution

HEFR could be modified to record the duration and timing of each period less than or equal to subsistence flows. HEFR would then output summary tables and/or figures of such periods (e.g., minimum, median, and maximum duration, both annual and binned by season; statistics describing typical intervals between such periods could also be generated). One or two of these statistics could be incorporated into the main HEFR output table of flow recommendations.

Anticipated Effort

Thirty-two hours of coding. Four hours of writing in HEFR report. Suggestions on appropriate statistics (if any) for the final flow regime are appreciated.

7. Fix Multipeaks_Multiplier Anomaly in Overbank Flow Calculations

Issue

Currently, when the multipeaks_multiplier is used to disaggregate long overbank events, the separate pieces of the disaggregated event are all assigned to the overbank category. However, most of these “pieces” probably do not contain a flow in excess of the 1.5 year return flow. Thus, these pieces should probably be reclassified and tracked as high flow pulses.

This issue has lingered since HEFR was created and was proposed for correction in March 2009 but received a low priority and was not addressed. It was not an issue for the existing BBESTs, as both groups turned off the multipeaks_multiplier functionality.

Proposed Solution

Recode HEFR to process overbanks before high flow pulses and reclassify overbank days as high flow pulse days if/when they are identified as separate events but do not exceed the small flood threshold (default 1.5 year return flow). Another “solution” would be to conclude that disaggregating long overbank events is not necessary and to turn this function off for overbank events.

Anticipated Effort

Twenty-four hours of coding. Limited edits to the HEFR report.

Outputs

1. Regression Diagnostics

Issue

Currently, the “alternative, frequency-based approach” uses log-linear and quadratic regression fits between duration and volume against peak flow. These regressions are often poorly behaved, particularly in the vicinity of flow recommendations. The quadratic method often predicts negative values for duration or volume at low peak flows.

Proposed Solution

While not a solution, per se, this proposal would automatically generate regression diagnostic plots in conjunction with the regression fits that are already generated. Such plots would likely include: (1) regression residuals versus peak flow, (2) quantile-quantile plots to graphically assess normality, and (3) regression residuals versus fitted values. It is not proposed that such outputs would be used to statistically accept or reject regression fits at some significance level. Such statistical rigor is misplaced in the context of these regressions. However, these graphical outputs would allow the user to more easily select the best regression line, and also highlight if both regressions are fundamentally unacceptable.

Anticipated Effort

Twelve hours of coding by TWDB coder. Four hours of writing to add description to HEFR report.

2. Improved Predictive Relationships between Duration, Volume, and Peak Flow

Issue

Currently, the “alternative, frequency-based approach” uses log-linear and quadratic regression fits between duration and volume against peak flow. These regressions are often poorly behaved, particularly in the vicinity of flow recommendations. The quadratic method often predicts negative values for duration or volume at low peak flows.

Proposed Solution

Quadratic and log-linear regressions may not be the best choice for this task. Alternative methods, such as running 25th, 50th, and 75th percentiles could provide superior fits in the vicinity of flow recommendations.

Anticipated Effort

The advice of an experienced statistician would be invaluable in identifying appropriate methods. If a relatively straightforward method can be identified, it could probably be coded in 20 hours or so. A few additional hours for describing the method in the HEFR report would also be warranted.

3. Multipeaks_Multiplier Highlighting

Issue

High flow pulses and overbank events that are split up by the multipeaks_multiplier algorithm are not highlighted as such.

Proposed Solution

These could be highlighted in the output sheets so that the user understands where the multipeaks_multiplier algorithm made a difference.

Anticipated Effort

Three hours of coding.

4. Bankfull Highlighting

Issue

In the frequency-based approach, high flow pulses and overbank events are pooled into one population to improve statistical calculations. One side effect is that the user can inadvertently select high flow pulses that exceed bankfull, or overbank events that are below bankfull. Such selections are inconsistent with the conceptual premise of these flow components.

Proposed Solution

HEFR could be modified to allow the user to input the estimated bankfull flow. Then, any high flow pulses in excess of bankfull, or overbank events below bankfull, would generate a warning during computations and in the output tables.

Anticipated Effort

Six hours of coding.

5. Enhance Graphical and Tabular Displays of Outputs

Issue

More charts and tables describing HEFR outputs could be helpful to users.

Proposed Solution

Some brainstorming and review of BBEST reports would identify helpful charts and tables.

Anticipated Effort

Unknown, but probably modest.

Ease of Use

1. Batch Mode

Issue

To generate HEFR outputs, the user must manually load and run each input data set individually. This can be tedious and inhibit exploratory evaluations of multiple sites.

Proposed Solution

A batch mode option could be added to HEFR. This would likely entail VBA code in Excel to use a list of cells containing the names of text files that hold USGS data and a separate file containing one set of HEFR input parameters. The VBA code would then run HEFR through each of the USGS data files using the one input parameter set.

Another possible solution that would not allow for batch mode, but would facilitate multiple runs, would be to recode HEFR to remember input settings from the last run. That way user inputs, especially complex seasonality, would not need to be manually re-entered each time HEFR is run (this was suggested by a survey respondent).

Anticipated Effort

A text file format suitable for containing all possible HEFR inputs (including highly flexible seasons, regressions, etc.) would need to be developed. Total coding effort for batch mode is estimated at 50 hours. Total coding effort for “memory” between runs is estimated at 40 hours. These tasks are largely coincident and 50 hours would probably suffice for both.

2. Overhaul User Interface, but not Functionality, of MBFIT

Issue

MBFIT is currently an Excel spreadsheet with little/no VBA code and no protection of cells. Various users have recreated and modified MBFIT, leading to possible versioning and QA/QC issues. Additionally, the IHA/MBFIT spreadsheet scored a relatively low average of 3 out of 5 in the HEFR survey.

Proposed Solution

Two related solutions present themselves. First, MBFIT could be wholly incorporated into the HEFR Excel Add-in. Alternatively, MBFIT could be recast as a stand-alone Excel Add-in. Either way, recoding MBFIT using VBA would reduce potential for corruption and allow for continued confidence in the calculations.

Anticipated Effort

This task is conceptually straightforward, as the MBFIT equations already exist. Improved visualization of outputs is probably warranted to facilitate parameter selection. Forty hours of coding are expected.

3. Change HEFR Defaults

Issue

HEFR default parameters facilitate model runs and provide guidance to the user. Lessons learned in the past year might be used to change HEFR defaults to improve the ease of use of the model.

Proposed Solution

Review HEFR applications by the Sabine-Neches and Trinity-San Jacinto BBESTs to determine what parameters were used and if replacements to the existing defaults are suggested.

Anticipated Effort

Ten hours of review. Four hours of coding. Two hours of editing HEFR report.

4. Recode HEFR outside of Excel and/or Allow Web Services

Issue

A survey respondent suggested that incorporation of HEFR into Excel limits functionality. This respondent also suggested allowing inputs from web services.

From my perspective, HEFR was specifically developed within Excel to facilitate data analysis and explorations by HEFR users, including contractors, many of whom are familiar with Excel. Recoding HEFR outside of Excel could inhibit, but certainly not prohibit, such analyses. I am not sure what potential functionality is being limited by Excel, as computations in VBA are highly flexible.

Web services do not appear to be mutually exclusive with Excel, as UT Center for Research in Water Resources (CRWR, led by Dr. David Maidment) researchers are working on an Excel spreadsheet that queries USGS flow data.

Proposed Solution

If desired, a different computing language and/or software platform could be identified and existing HEFR computations and tools migrated to the new platform.

Web services (presumably USGS flow data, at least to start) could be explored.

Anticipated Effort

Significant effort is anticipated for recoding HEFR. Depending on the computational language and platform used, a new coder would potentially be needed. Modest effort is anticipated for web services, assuming assistance is available from CRWR staff.

Documentation

1. Rename Episodic Events Calculation Methods

Issue

Currently, the two episodic events calculation methods are specifically referred to as the “original, percentile-based approach” and the “alternative, frequency-based approach.” These names were selected to reduce confusion at the time of the development of the second approach, however the names are cumbersome and themselves confusing.

Proposed Solution

Rename to the “Percentile-based” and “Frequency-based” approaches – unless anyone has a better idea.

Anticipated Effort

Very little coding, some editing of HEFR document if/when a future release is warranted.

2. Enhance Documentation

Issue

The HEFR report is nearly one year old. Some modifications to the text may be warranted (over and above those related to enhancements described herein). One survey respondent suggested expansion of Section 4 (“Decision Points...”) as well as interpretation of the outputs. Apparently decision points are a hot topic these days, as evidenced by the forthcoming memoirs of a certain Texan and former president. However, one presumes that such memoirs will contain little guidance with respect to HEFR decision points.

An adaptation of the Sabine-Neches Decision Matrix (page 8 of their report) could be very useful to the reader. Another survey respondent suggested increased documentation related to MBFIT.

Proposed Solution

Review, edit, and expand HEFR report.

Anticipated Effort

Forty hours of writing and editing. Additional review by SAC members.