Overcoming the challenge of initial parameter estimation for event-based hydrological models

Luciana K. Cunha, Project Manager, WEST Consultants, Folsom, CA, leunha@westconsultants.com
David C. Curtis, Sr. Vice President, WEST Consultants, Folsom, CA, dcurtis@westconsultants.com
Travis Stanford, U.S. Army Corps of Engineers, Fort Worth District, Fort Worth, TX, Travis.W.Stanford@usace.army.mil
Jerry Cotter, U.S. Army Corps of Engineers, Fort Worth District, Fort Worth, TX, Jerry.L.Cotter@usace.army.mil

Abstract

An innovative approach that bridges the gap between event-based hydrological modeling and continuous soil moisture accounting approaches is discussed. Hydrologic Engineering Center (HEC) Hydrological Modelling System (HMS), developed by the US Army Corps of Engineers (USACE), is one of the most widely used hydrologic modeling programs in the world. HMS was used in the project, but the developed methodology can be extended to any event-based hydrological model.

HMS is applied to forecast reservoirs operations both by the USACE and outside agencies. While HMS can be ran continuously, it is mainly applied operationally in its event-based configuration to avoid the need of continuous maintenance. Event-based hydrological models do not explicitly address hydrologic processes that take place between storms such as evapotranspiration and percolation of water through the soil column. In HMS, hydrologic losses from incident rainfall are modeled simplistically using an initial loss and a continuing loss parameter. In an operational environment, these parameters are adjusted once catchment response is observed and compared to the HMS response. Time spent waiting for the initial catchment response needed to adjust the HMS loss parameters is time lost for flood mitigation or planning reservoir operations.

An innovative approach was developed that applies simulated soil moisture obtained from the Noah North American Land Data Assimilation System version 2 (NLDAS-v2) to update HMS parameters before an event is initiated. The application of NLDAS-v2 allows the estimation of initial losses after long dry periods when flow and rainfall indices are likely to be zero. NLDAS-v2 has the advantage of being continuously updated and publicly available online.

The methodology was applied for 24 reservoirs in the Ft. Worth District of the USACE. For each reservoir, statistical analyses were performed to determine the best soil moisture indicators and to define the relationships between initial moisture condition, rainfall, and runoff. A spreadsheet tool was developed that incorporates the NASA soil moisture data and information on historical storms to provide improved estimates of initial and continuous loss parameters for HMS in a real-time river forecast environment.
Introduction

The US Army Corps of Engineers (USACE) Fort Worth District (SWF) routinely produces inflow forecasts for each of the District’s dams that support decisions related to reservoir operations. Historically, one of the challenges to accurately forecast inflows to District reservoirs is the estimation of hydrologic losses. Current methods in use by the Fort Worth District to estimate hydrologic losses involve simple nomographs and HEC-HMS hydrologic modeling software which provides information regarding timing, rate of rise, and peak pool elevation. Both methodologies are supported, in large measure, by professional judgement. The general consensus from the District’s hydrologic forecasters is that the current methods are inadequate, especially for rainfall events occurring after prolonged dry periods or for multiple storm events occurring in series.

The goals of the project was to develop and implement an easy-to-use method for predicting losses associated with storm events that contribute runoff to 24 dams of interest (Figure 1). A methodology was developed to bridge the gap between event-base hydrological models commonly used by the USACE and the need for timely and accurate runoff forecasts even after long dry periods. The method takes advantage of existing systems continuously operated by NASA or NOAA. For this project, the North America Land Data Assimilation System (NLDAS) was applied. The National Water Model (NWM) operated and maintained by the National Oceanic and Atmospheric Administration (NOAA) can also be used, but at the time of the project the NWM reanalysis period was not available.

![Figure 1. Location of the dams of interest](image)
Datasets

Daily runoff time series were provided by the SWF for all sites of interest. The period of record available varies for each location, but all time series cover the period of 1981 to 2016. Basin average rainfall values were calculated based on the Parameter-elevation Regression on Independent Slopes Model (PRISM) AN81d (Daly and Bryant, 2013). PRISM accounts for topographic influences on precipitation patterns and amounts.

Soil moisture conditions for each watershed was characterized by the North American Land Data Assimilation System version 2 dataset (NLDAS-v2). NLDAS-v2 datasets are available at the hourly scale for the period of 1979 to present and are updated daily with a latency of 3 or 4 days. The spatial resolution of the NLDAS dataset is 13 km in space (Xia et al 2016).

NLDAS-v2 provides the results of four land surface models forced by the best available observations. For a review of these models refer to Cai et al (2014). The results from Noah-2.8 and VIC-4.0.5 are applied in this report. Noah-2.8 was selected since this model is used as the land surface component in multiple weather forecast systems. The Noah-2.8 model was developed as the land component of the NOAA NCEP mesoscale Eta model (Xia et al 2012). VIC was selected since it is the model currently used to generate CMIP5 hydrology projections over the contiguous U.S. by the U.S. Bureau of Reclamation and collaborators (U.S. Bureau of Reclamation, 2014). This dataset can be potentially used to better understand the effects of climate change on runoff generation. VIC was also included in the project to verify if the methods applied in this project are sensitive to the land surface model applied.

Data for the period 1981 to 2016, which correspond to the period of available daily runoff data, were applied. The Noah-2.8 soil moisture variables used in this project include: soil moisture content (SOILM) for soil depths of 0 to 40 in, 0 to 80 in, 4 to 16 in, 16 to 40 and 40 to 80 in; liquid soil moisture content (LSOIL) for soil depths of 0 to 4 in, 4 to 16 in, 16 to 40 in and 40 to 80 in; and moisture availability (MSTAV) for soil depths of 0 to 40 in and 0 to 80 in.

The VIC-4.0.5 model is a macroscale, semi-distributed, grid-based, hydrologic model developed at the University of Washington and Princeton University (Wood et al., 1997). NLDAS-2 applies the full water and energy balance version of VIC (version 4.0.5). The model includes three soil layers, with a 10 cm top layer (SOILM 1) and four deeper layers of spatially varying thicknesses (SOILM 2 and 3). Soil moisture (SOILM) and moisture availability (MSTAV) for the 0-40 inches layer (SOILM 0-40) and for the total column (SOILM Total) are also provided. The VIC model utilizes sub-grid vegetation tiles, and includes a two-layer energy balance snow model (Cherkauer et al., 2003) that uses sub-grid elevation bands to represent the impact of elevation on temperature, precipitation, and snow.

The time series of basin average precipitation and Noah 16-40 inch layer liquid soil moisture content (LSOIL) for the Benbrook Lake Reservoir for a dry (2011) and a wet (1991) year show the influence of initial soil moisture on runoff generation (Figure 2). LSOIL was similar in the beginning of both years. However, LSOIL decreased significantly in 2011 due to low precipitation and high temperature. By October 2011, LSOIL was significantly below average, while in October 1991 LSOIL was near average. For both years, a significant rainfall event (more than 4 inches/day) was registered in October. The resultant maximum peak inflow for this event was only 0.04 inch/day in 2011, while in 1991 a maximum peak inflow of 2.23 inch/day was observed.

1 Weather Research and Forecasting (WRF) regional atmospheric model, the NOAA NCEP coupled Climate Forecast System (CFS) and the Global Forecast System (GFS)
Methodology

The HMS deficit and constant loss model (DCLM) applies an initial value of loss at the beginning of the rainfall event, prior to the beginning of surface runoff, which is a function of the current moisture state (MS) of the watershed. Interception loss (IL) accounts for rainfall intercepted by vegetation, retention on the surface (depression storage) and initial infiltration into the soil. As the watershed becomes saturated, the initial loss approaches zero. The drier the MS of the watershed, the higher IL will be.

Continuing loss (CL) is the average rate of loss throughout the rest of the storm and accounts for infiltration, percolation and evaporation. CL reflects the physical properties of the watershed (e.g. soil, land cover, topography) and the storm event (e.g. rainfall intensity, duration, spatial and temporal distribution). For the same watershed, CL varies only based on the properties of the rainfall event (RE).

From the start of the rainfall event, the water balance equation can be expressed as:

\[ P = IL(MS) + CL(RE) \cdot t + QF \]

where P is precipitation, IL(MS) is initial loss as a function of the moisture state of the watershed, CL is continuous loss as a function of the rainfall event properties (RE), t is the time for which rainfall is larger than zero and QF is the runoff resulting from the rainfall event. QF is expressed as the total inflow to the reservoir minus baseflow (BF). BF is the portion of the inflow that comes from the sum of deep subsurface flow and delayed shallow subsurface flow. P, IL(MS), and QF are measured in inches, while CL(RE) is measured in inches per day.
Initial and continuous losses were estimated using daily basin average precipitation estimated from PRISM AN81D and runoff data provided by the SWF. The methods applied, as well as the results, are presented in the following sections.

**Initial Loss Estimation:** Initial loss is a function of the current state of saturation of the watershed. In the DCLM, the current moisture state of the watershed is specified as water deficit. The model requires the definition of two parameters: maximum soil deficit and soil deficit at the beginning of the simulation. In this section, a methodology to estimate maximum soil deficit and soil deficit as a function of the current moisture state of the watershed is presented.

From the available discharge and precipitation time series, it is possible to retrieve several indicators of hydrologic conditions prior to the flood events. To determine the optimal indicator for the hydro-meteorological conditions, multiple indicators were initially calculated and tested. These included: base flow at the beginning of each event, accumulated precipitation and discharge for 5, 15, and 30 days prior to the flood, the antecedent precipitation index (API) and antecedent flow index (AFI). Continuous time series of all indicators were generated.

Both rainfall and runoff present advantages and disadvantages when used to estimate soil moisture. Rainfall has the advantage of accounting for storm events even when no runoff is generated. Runoff has the advantage of being the result of the integration of multiple hydrologic processes at multiple spatial and temporal scales (rainfall, infiltration, evapotranspiration, flow transport), and therefore represents the overall state of the watershed. However, runoff indices are affected by season and do not necessarily reflect changes caused by rains during the previous week.

While rainfall and runoff indicators are suitable in many regions, that is not the case for Texas. All regions in Texas experience long periods of droughts, and during these periods rainfall and runoff indicators tend toward zero and are not able to capture the severity of the drought. To overcome this limitation, in this project soil moisture time series generated based on land surface models are applied. While measured soil moisture information is not readily available, simulated soil moisture data is available as output from land surface models. Since the soil moisture data is the output of a model, it is subject to model and data uncertainties. However, these datasets can be very useful since land surface models take into consideration other processes (e.g. evapotranspiration, soil percolation) that are not directly accounted for when applying observed precipitation and runoff datasets to estimate moisture indices.

To improve the real-time estimation of initial loss, relationships that relate initial losses to the current state of the watershed were developed. Various relationships were developed based on multiple moisture indicators with the goal of selecting the relationship that provides the best estimate of initial losses for each watershed.

Initial losses were estimated by evaluating historical rainfall-runoff events. Storm events that generated no or insignificant runoff were selected. To quantify insignificant runoff, a threshold was defined. The initial losses for these events were plotted against soil moisture indicators and linear or exponential relationships were fit to the data. Two examples are shown for the Benbrook Lake watershed based on (A) AFI and (B) Noah liquid soil moisture content (0 to 40-inch) (Figure 3). In this plot, different seasons are represented by markers of different shape and color. In general, the initial losses decrease as moisture conditions increases.

A methodology to extract semi-envelope curves was applied. An envelope curve of a series of points is represented by the smooth curve that outline all extreme points. The semi-envelope curves were derived by relaxing the requirement that all extreme points should be below the
curve. The semi-envelope curve focuses on the decay rate of the data, instead of encompassing all extreme points.

Semi-envelope curves are used instead of envelope curves since initial losses that result in saturated or almost saturated soils were probably not observed for the full range of initial states of the watershed. Moreover, uncertainties arise from using daily instead of hourly precipitation and runoff data as well as from including multiple day events without a more robust representation of evapotranspiration.

![Figure 3.](image-url) (A) AFI and (B) liquid soil moisture content at the 0 to 40-inch soil layer versus initial loss data (inches) for BNBT2. Values for different seasons are identified by color and marker shape. JFM: January, February, and March; AMJ: April, May and June; JAS: July, August and September, and OMD: October, November and December.)
**Continuous and total loss estimation:** Storm events generating significant runoff were selected from the available time series of precipitation and runoff. For each event, initial loss was estimated using the soil moisture indicator for the beginning of the event and the exponential or linear curves described in the previous section. Once the initial loss is known, continuous loss can be calculated as a function of known properties of the event, including total rainfall, number of days with rain, average daily rainfall and the moisture condition in the beginning of the event. An optimization procedure was implemented to determine which moisture indicator best describes initial losses for the basins of interest, and which characteristics of the rainfall event best predict continuous loss.

The procedure to estimate initial and continuous losses is schematically represented in Figure 4. The goodness of fit of the relationships are evaluated based on the coefficient of determination ($\rho^2$) and the normalized RMSE (CVRMSE). The coefficient of determination ($\rho^2$) is a number that indicates the proportion of the variance in the dependent variable that is explained by the independent variable(s). The root mean squared error (RMSE) provides a measure of the differences between predicted and observed values. The normalized RMSE (CVRMSE) is adopted since allows for the direct comparison of different datasets or models.

![Figure 4. Optimization procedure to select the best moisture indicator and relationships to predict the characteristics of flood events (dependent variables)](image-url)
Runoff event selection: Figure 5 shows a schematic representation of the procedure used to select main runoff events. The figure contains all the steps required to define the events.

![Runoff event selection process: schematic representation](image)

1) Event identification: identify runoff that is larger than the 98th runoff percentile;
2) Beginning of event (Tb): look backwards from Tr minus concentration time and identify two days with rainfall equal to zero or the end of the previous event;
3) End of event (Te): look forward from Tr and identify the time for which the three criteria are valid:
   1. Runoff less than half of the 98th runoff percentile
   2. No significant rainfall for the last Te days
   3. No significant decay in runoff for 2 consecutive time steps

Figure 5. Runoff event selection process: schematic representation

**Event-based hydrological analysis: results**

**Initial and Continuous Loss Estimation:** The best soil moisture indicator varied for each basin and variable being predicted. To be consistent and to allow the direct comparison among different basins of interest a unique soil moisture indicator was selected for all sites. The Noah SOILM 0-40 inches layer was selected since its performance for all basins and variables does not differ significantly from the performance obtained with the best NLDAS soil moisture indicator.

Statistically significant relationships between hydrological losses and the following variables were identified: a) Soil moisture state at the beginning of rainfall; b) Storm total rainfall volume; c) Duration of rainfall; and d) Peak day rainfall intensity.

The r-squared and CV(RMSE) for relationships to estimate total loss, total continuous loss, and peak runoff based on Noah SOILM 0-40-inch data are shown in Figure 6. For all watersheds, the NLDAS relationships explained from 73% (LEWT2) to 96% (WTYT2) of the variation in total loss, 61% (PCCT2) to 95% (GGLT2 and WTYT2) of the variation in total continuous loss and 32% (JFNT2) to 75% (BNBT2) of the variation in peak flow. CVRMSE is shown in Figure 7. Total loss and total continuous loss present similar errors, while errors for peak flow are considerably higher. This is expected since the estimation of peak flow using daily data is limited especially for small watershed with lag time of 1 day or less.
Figure 6. R-squared for relationships to estimate total loss, total continuous loss, and peak runoff based on soil moisture data (Noah SOILM 0-40) for watersheds of interest (Figure 1). Sub-basins are identified in Figure 1. Watersheds are identified by B: Brazos; G: Guadalupe; N: Neches; R: Red and T: Trinity watershed.

Figure 7. CV(RMSE) for relationships to estimate total loss, total continuous loss, and peak runoff based on soil moisture data (Noah SOILM 0-40). Sub-basins are identified in Figure 1. Watersheds are identified by B: Brazos; G: Guadalupe; N: Neches; R: Red and T: Trinity watershed.

Operational Tools: The results for all watersheds are provided in two main products developed to support event-based forecast:

1. Summary sheet: an individual summary sheet was created for each basin containing all the relevant information generated in this project. The summary sheet was designed taking into consideration forecasters need in the advance of an event. The document can be used to evaluate the current condition of the watershed in comparison to past conditions, and to make predictions if a computer is not available. The document can also be used to compare the response of different watersheds.
2. **Total Runoff Tool**: The results of the regression analyses were compiled in Excel files designed to support event-based runoff estimation at the daily scale. The Excel files are referred as the “Total Runoff Tools” (TRT). The TRTs contain basically the same information shown in the Summary Sheets in a way to support automated calculations for real time operations.

For brevity, the summary sheet for only one watershed will be discussed and included in this paper. Figure 8 and Figure 9 contain the first and second pages of the Benbrook Lake Watershed summary sheet. The following sections are included in the summary sheet:

1. **General Statistics** (Figure 8): This section contains general characteristics of the watershed, including the SHED code, name of the dam it drains to, USGS Site ID number, drainage area, mean annual precipitation, mean annual flow, mean annual loss, lag time, curve number, main land use, the 1, 5, 25, 50, 75, 95, and 99 quantiles for daily rainfall, daily runoff, initial condition, and seasonal initial condition. It also shows a map of the basin location.

2. **Daily changes** (Figure 8): This section contains two plots that show how initial soil moisture affects daily changes in runoff and soil moisture:

3. The plot on the left shows daily change in runoff as a function of initial soil moisture condition (x-axis) and accumulated rainfall for the lag time in days plus one (color-coded). This plot can be applied to understand the limits of daily runoff change as a function of the current soil moisture condition of the basin.

4. It is important to note that significant daily change in runoff does not occur when the soil is dry, but dry soil conditions vary from basin to basin. For example, for BNBT2 the maximum observed daily change in runoff is 2-inch/day and it occurred when soil moisture condition (SOILM 0-40-inch layer) was approximately 13.7 inches. For LVNT2 the maximum observed daily change in runoff is 3.8-inch/day and it occurs when soil moisture condition was approximately 14.4. For both basins the maximum daily runoff change occurred when the 2-day accumulated precipitation was higher than 4-inch. For TBLT2, the maximum observed daily change in runoff is approximately 0.23 inches/day, even when the soil is close to saturation.

5. The plot on the right shows the 4-day change in soil moisture as a function of initial soil moisture condition (x-axis) and the 4-day accumulated rainfall (color-coded). This plot allows the operator to estimate minimum and maximum changes in soil moisture that might have occurred during the NLDAS latency period. This plot can be used to obtain an initial estimation of the updated soil moisture. This plot also provides an estimation of the physical limits of continuous losses for each initial soil moisture condition. This information should be used with caution since the results shown in the plot are for simulated soil moisture properties over the entire watershed on the period of 4 days as simulated by a land surface model. Therefore, this information should not be used to determine the maximum infiltration over a short period of time and at a specific location in the basin.

6. **Soil moisture and daily runoff** (Figure 8): This section contains two plots that show how initial soil moisture affects daily runoff and how soil moisture distribution changes with the day of the year:

   a. The plot in the left shows daily runoff as a function of initial soil moisture condition (x-axis) and accumulated rainfall for the lag time in days plus one (color-coded). This plot can be applied to understand the limits of daily runoff as a function of the current soil moisture condition of the basin. Daily runoff
increases as soil moisture increases. Maximum daily runoff also varies significantly from basin to basin.

b. The plot in the right presents the distribution of soil moisture for each day of the year based on the period of 1981 to 2016. Shaded areas define the minimum and maximum, the 5th and 95th percentile, and the 25th and 75th percentile of soil moisture for each day of the year. The black line indicates the mean soil moisture for each day. This figure allows the operator to contextualize the current state of the soil in relation to what has been observed in previous years. This plot is also useful during drought situations to better understand current drought intensity.

7. Initial condition update (Figure 9): This section presents the equations to update the soil moisture condition as a function of the current soil moisture and the accumulated rainfall for latency period of 3 or 4 days.

8. Runoff relationships (Figure 9): This section presents the equations to estimate initial loss, continuous loss, total loss and peak runoff for a storm event. The section also contains a plot that allows the forecaster to visually estimate an approximate value of total runoff as a function of total rainfall, initial soil moisture (color-coded) and peak rainfall intensity (marker size). Automated calculations based on these equations can be performed in the TRT.

9. Spatial variability (Figure 9): This section contains maps that allow the forecaster to evaluate the spatial variability of hydrological losses in the basin of interest. The map on the left shows spatial variability of hydrological losses for each HUC 12. The map on the right shows the loss spatial factor k, which represents the percentage of hydrological losses that occurs in each sub-area. This maps can be used to estimate uncertainties in the quantification of losses depending on the location of the storm in the basin.

Conclusions

A novel methodology to estimate hydrologic losses and total runoff at the event scale incorporating daily data was developed and encoded in Excel spreadsheets for operational use. The methodology relies on the analytical results derived from daily runoff, precipitation, and soil moisture indicator data. Rainfall events generating little or no runoff were used to establish relationships to estimate initial losses. Significant runoff events were used to establish relationships to estimate continuous and total losses based on the properties of the rainfall and initial soil moisture conditions. The results provides two main products to support event-based forecast, a summary sheet, and a Total Runoff Tool. The proposed methodology offers a significant improvement over current methods used to estimate hydrologic losses.
Figure 8. Runoff event selection process: schematic representation
**BNBT2:** Benbrook Lake, USGS: 08046500
Drainage area: 430 mi²

### Initial condition update

<table>
<thead>
<tr>
<th>Latency</th>
<th>Rain=0</th>
<th>Rain&gt;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-day</td>
<td>(IC = 0.96 \times IC_{(t-3)} + 0.15)</td>
<td>(IC = 0.94 \times \sum_{t=1}^{3} R_t + 0.95 \times IC_{(t-3)} + 0.36)</td>
</tr>
<tr>
<td>4-day</td>
<td>(IC = 0.95 \times IC_{(t-4)} + 0.19)</td>
<td>(IC = 0.94 \times \sum_{t=1}^{4} R_t + 0.93 \times IC_{(t-4)} + 0.45)</td>
</tr>
</tbody>
</table>

### Runoff relationships

- **Number of events:** 93
- **Average event duration:** 5 days
- \(IL = 5.75 - 0.54 \times IC\)
- \(CL = 1.58 + 0.54 \times TR - 0.11 \times IC\)
- \(TL = IL + CL\) or
- \(TL = 5.93 + 0.55 \times TR - 0.50 \times IC\)
- \(TR = TR - TL\)

\(PR = -2.16 + 0.31 \times PR + 0.18 \times IC\)

**IC:** Initial condition based on NLDAS-v2 NOAH SOILM 0-40 inches layer (in); \(IL\) - Initial loss (in); \(TR\) - Total rainfall (in); \(CL\) - Continuous loss (in); \(NL\) - Number of days rainfall; \(TL\) - Total loss (in); \(TR\) - Total runoff (in); \(PR\) - Peak rainfall (in/day); \(PR = \) peak runoff (in/day); \(AR\) - Average daily rainfall (in/day)

### Spatial variability

**HUC 12 - Losses**

- **Loss spatial factor (k):**

\[\sum_{k=1}^{1} TL_1 = TL \times k_1\]

**Figure 9.** Runoff event selection process: schematic representation
References


