MULTI-SCALE CALIBRATION OF KINEROS-DWEPP, A COMBINED PHYSICALLY-BASED HYDROLOGIC MODEL AND PROCESS-BASED SOIL EROSION MODEL

Shea Burns, Research Specialist, USDA-ARS, Tucson, AZ, shea.burns@ars.usda.gov; Phil Guertin, University of Arizona (UA), Tucson, AZ; David C. Goodrich, USDA-ARS, Tucson, AZ.

INTRODUCTION

This paper compares a step-wise, multi-scale calibration of a physically-based hydrologic model, DWEPP, to a more traditional calibration performed at a single scale. The Dynamic Water Erosion Prediction Project model (DWEPP; Bulygina et al., 2007) combines the KINematic Runoff and EROSion model (KINEROS2; Semmens et al., 2008), an established physically-based hydrologic model, and the Water Erosion Prediction Project model (WEPP; Flanagan et al., 2001), a proven process-based soil erosion model. KINEROS2 has been shown to perform well when representing runoff (Miller et al., 2007) but the representation of sediment transport and erosion processes are poor (Canfield and Goodrich, 2006). Using DWEPP, Bulygina (2007) calibrated at plot and hillslope scales and found satisfactory performance of sediment yield dynamics; however, the model was not tested at the small watershed scale and performance across scales, a substantial challenge in both hydrology and erosion models, was not evaluated. The step-wise, multi-scale approach used in this exercise attempts to improve upon traditional calibrations where uncertainty and poor performance are common when moving across spatial scales.

The calibrations use data collected at the plot, hillslope, and small watershed scale from the USDA-ARS Southwest Watershed Research Center’s Walnut Gulch Experimental Watershed (WGEW). Rainfall and runoff data were collected using a rainfall simulator (Paige et al., 2003) for the plot scale, and via a network of rain gages and flumes for the hillslope and watershed scales (Renard et al., 1993, 2008). Calibrations start at the plot scale where several key parameters for plot-scale runoff and erosion processes are calibrated and fixed. At the hillslope-scale, plot-scale parameters remain fixed from the previous calibration and another set of key parameters for the hillslope-scale are calibrated. Finally, at the small watershed scale, plot- and hillslope-scale parameters remain fixed from the prior calibrations, and channel parameters are calibrated. A traditional calibration was also performed at the small watershed scale for comparison to the performance of the step-wise, multi-scale calibrations.

The semi-arid WGEW, located in southeastern Arizona (31° 43’N, 110° 41’W), is 149 km² and surrounds the town of Tombstone, Arizona (Figure 1). It has a mean annual temperature of 17.7°C and a mean annual precipitation of 350mm, the majority of which is a result of high-intensity convective thunderstorms in the summer monsoon season. The watershed is ephemeral and nearly all channel runoff is generated by rainfall from the summer monsoon. The watershed has 27 distinct soil units mapped by the USDA-NRCS in the Soil Survey Geographic (SSURGO) Database. Vegetation is characterized by shrubland in the lower two-thirds of the watershed and grassland in the upper third (Renard et al., 1993, 2008).
Lucky Hills, a set of subwatersheds located in the shrubland portion of the WGEW, was chosen as the study area (Figure 1). Lucky Hills is instrumented with two raingages and five flumes. Within the Lucky Hills study area, the Lucky Hills 104 (LH104) watershed contains the nested watersheds Lucky Hills 102 (LH102) and Lucky Hills 106 (LH106), all with digital runoff recording flumes. The LH104, LH106, and LH102 watershed areas are 4.41, 0.367, and 1.653 hectares, respectively. LH104 and its subwatersheds present a good location for this study because of the availability of the rainfall, runoff, and sediment time-series data required for calibration at each scale of interest.

MODELLING FRAMEWORK

**Dynamic Water Erosion Prediction Project (DWEPP) model** The Dynamic Water Erosion Prediction Project (DWEPP) model is the product of combining the process-based WEPP model with the physically-based KINEROS2 model to create a dynamic version of WEPP. DWEPP can estimate runoff, sediment deposition, channel erosion, and sediment yield where it occurs within a hillslope or watershed like the original WEPP model (Flanagan et al., 2007), but can also estimate within-storm hydrology and sediment dynamics (Bulygina et al., 2007).

Infiltration is modeled as a function of infiltrated depth as represented by the three parameter Smith-Parlange infiltration model (Parlange et al., 1982). Runoff is routed dynamically using the 1D kinematic wave model and the continuity equation as presented in Equation (1):

\[
\frac{\partial h}{\partial t} + \frac{\partial Q}{\partial x} = q(x, t)
\]

(1)

where \( Q \) is discharge per unit width, \( h \) is the depth of flow, \( t \) is time, \( x \) is distance along the flow path, and \( q \) is the lateral inflow rate. The 1D dynamic routing of sediment in DWEPP is also represented by a continuity equation as in Equation (2):

\[
\frac{\partial CA}{\partial t} + \frac{\partial (CQ)}{\partial x} = S
\]

(2)

where \( C \) is sediment concentration, \( A \) is cross-sectional area of flow, \( Q \) is flow discharge, \( t \) is time, \( x \) is distance down the flow path, and \( S \) is the source/sink term for sediment.

DWEPP maintains the same representation of a watershed, a cascade of hillslopes and channels, as KINEROS2 (Figure 2). Hillslopes are abstracted into one-dimensional rectangular overland flow planes that contribute laterally to channels. Channel segments have two lateral contributors and up to ten upstream elements, channel or hillslope.

**Automated Geospatial Watershed Assessment (AGWA) Tool** AGWA is a geographic information system (GIS) interface and framework used to organize data and parameterize, integrate, and visualize two watershed hydrologic models, KINEROS2 and the Soil & Water Assessment Tool (SWAT) (Miller et al. 2007; [http://www.tucson.ars.ag.gov/agwa](http://www.tucson.ars.ag.gov/agwa)). AGWA is available as AGWA 1.5 for users with Environmental Systems Research Institute (ESRI) ArcView 3.x and as AGWA 2.0 for users with ESRI ArcGIS 9.x. AGWA requires basic digital data to parameterize KINEROS2 including a digital elevation model (DEM), a polygon soils map and associated soils database and look-up tables, and a classified land cover grid and associated look-up table. AGWA 2.0 is used in this research to delineate and discretize the different watershed configurations and then derive a priori input parameter values for hillslope elements for KINEROS2, which then are used for DWEPP.
METHODOLOGY

Data collection For this study, rainfall simulator events on four replicate 2m by 6m plots in the Lucky Hills Study area were used from the summer of 2007. The Walnut Gulch Rainfall Simulator, an oscillating boom variable intensity simulator (Paige et al., 2003) was used to simulate rainfall events. The events selected were recommended by Stone (personal communication, 2008) as the most current and appropriate for this application. The four events used in the study were wet runs beginning at an intensity of 61.7 mm/hr and moving up to 103.6, 126.2, 152.9, and 179.6 mm/hr, sequentially. Each intensity was held constant until a condition of steady state discharge was reached. The wet run was preceded 45 minutes earlier by a dry run of constant intensity at 61.7 mm/hr for 45 minutes to achieve a consistent soil moisture across the plot.

Rainfall is measured at two locations within the Lucky Hills Study area (Figure 1). A detailed description of the rain gage history, design, measurement, and network in the WGEW is described by Goodrich et al. (2008). Runoff is measured at three different locations within the Lucky Hills Study area (Figure 1). LH104 and LH102 watersheds have Santa Rita or Smith flumes and LH106 has an H flume installed (Stone et al., 2008). The data quality for both flumes types is considered “good to excellent” (Stone et al., 2008). Sediment is measured at three different locations within the Lucky Hills Study area (Figure 1). LH104 and LH102 watersheds have traversing slot sediment samplers and LH106 has a pump sampler (Nichols et al., 2008). Associated discharge measurements are integral to calculating sedigraphs and sediment yields for runoff events. The pump sampler is restricted to measuring suspended sediment smaller than 6.35 mm. The traversing slot sampler is restricted to measuring suspended sediment smaller than 13 mm. Neither sampler measures bed load. Sample collection times are stored on and synchronized with the runoff data logger for the flume.

Model Setup AGWA 2.0 was used to create watershed configurations in the study area, parameterize them, and write KINEROS2 parameter files, which were then customized for use in DWEPP. A 1m resolution Light Detection And Ranging (LiDAR) digital elevation model (DEM) projected to UTM Zone 12 with NAD83 datum was used to delineate and discretize the watershed into model elements and calculate their slopes. Multiple contributing source areas (CSA) were used in the discretization of LH104 to represent the different hillslope areas. The CSA defines the threshold at which first order (headwater) streams are defined. The discretized watershed was parameterized using the AZ671 – Cochise County, Arizona, Douglas-Tombstone Part soil survey (Soil Survey Staff, 2008), the National Land Cover 2001 Dataset (Homer et al., 2004), and associated AGWA look-up tables. Initial DWEPP specific parameters values not derived by AGWA were obtained from Alberts et al. (1995) and Bulygina (2007).

Plot Scale The rainfall simulator plot was represented in the model as a single overland flow element, extracted from a KINEROS2 parameter file written by AGWA 2.0. It was then customized with the geometric and topographic properties of the rainfall simulator plots. The element length was set to 6 meters (the plot length); the element width was set to 2.01 meters (the plot width); the element slope was set to 0.16, the measured slope used by Buono (2009) in rainfall simulator studies at Lucky Hills; and canopy cover was set to 0.17, the measured cover value for the rainfall simulator studies in Buono (2009). Interception depth was set to 0.612 mm
(calculated from a 3.6 mm interception depth for the vegetation type multiplied by 17% cover on the plot).

**Hillslope Scale** Three hillslope configurations were used, representing LH106, LH102, and LH104 as single overland flow elements. Each configuration was evaluated post-calibration to determine which best represented the observed data. LH104 was discretized using AGWA 2.0 with CSAs of 0.3663 and 1.6528 ha to create KINEROS2 parameter files from which the LH106 and LH102 hillslope elements were extracted. The LH104 hillslope element for the third configuration was extracted from KINEROS2 parameter file created from a discretization with a CSA of 4.4077 ha and a watershed outlet further downstream.

**Watershed Scale** Two watershed configurations were used, representing CSAs where LH106 and LH102 are discrete hillslopes. Each configuration was evaluated post-calibration to determine which best represented the observed data. The configurations used the KINEROS2 parameter files previously created for the extraction of the LH106 and LH102 hillslope elements at the hillslope scale. The LH106 configuration contained 13 hillslope and 5 channel elements; the LH102 configuration contained 8 hillslope and 3 channel elements.

**Model Calibration** Model calibration was performed using PEST, a model-independent parameter estimation and optimization software package (Doherty, 2004; [http://www.pesthomepage.org/](http://www.pesthomepage.org/)). PEST uses a nonlinear estimation technique, the Gauss-Marquardt-Levenberg method. It works by minimizing an objective function of the form:

\[
\phi = \sum_{i=1}^{m} (w_i r_i)^2
\]

where \( m \) is the number of observations, \( w_i \) is the weight associated with the \( i \)’th observation, and \( r_i \) is the residual (observed minus predicted) of the \( i \)’th observation.

PEST may exhibit poor performance or have trouble lowering the objective function under a variety of circumstances including but not limited to situations where: derivatives are not calculated with sufficient precision; parameters are highly correlated; the model has a highly nonlinear response to parameter change; initial parameter values are poor; observations are insensitive to initial parameter values; parameter values reach their upper or lower bounds; or PEST calibration parameters are set up poorly. In the case of highly correlated or highly insensitive parameters, PEST may also encounter non-uniqueness with those parameters. PEST is also susceptible to finding a local minimum in the objective function instead of the global objective function minimum. PEST and DWEPP communicate interactively through batch files (Figure 3).
**Calibration Scheme** Two sets of calibrations were performed to compare, contrast, and distinguish possible approaches to the calibration problem (Figure 4). The calibrations are grouped into two scenarios described in more detail below. Group 1 was designed to incorporate the calibration of different sets of parameters with the plot, hillslope, and watershed scales over multiple events in a step-wise fashion, beginning at the plot scale and moving through the hillslope to the watershed scale. Group 2 was designed to incorporate the calibration of different parameters at the watershed scale over multiple events in a lumped fashion and then scales down to evaluate calibrated hillslope parameters at the hillslope scale.

Figure 4 Step-wise (left) and lumped (right) calibrations on all precipitation events.
Calibration Group 1  Calibration group 1 (Figure 4) begins with P1 at the plot scale with the calibration of parameters hydraulic conductivity (KS) and initial soil saturation (SAT) against the hydrographs from four replicate wet rainfall simulator plots. Despite the plot scale calibration simulating a wet run of the rainfall simulator, SAT was not set to 1 (100% saturation) and instead was calibrated simultaneously with KS to maximize the performance of the model. P2 followed P1 and utilized the optimized parameters from P1 in the calibration of Manning’s roughness coefficient (MAN) against the same hydrographs as P1. P3 followed P2 and used the optimized parameters from the previous calibrations in the calibration of rill Manning’s roughness coefficient (RMAN), interrill erodibility (KI), and rill erodibility (KR) against the sediment yields from the same four rainfall simulator plots.

For the hillslope scale in group 1, separate calibrations were performed at LH106, LH102, and LH104 to compare the performance of different representative hillslope areas, KS was calibrated in H1\textsubscript{106}, H1\textsubscript{102}, and H1\textsubscript{104} against the hydrographs from 17 events at LH106, 18 events at LH102, and 18 events at LH104, respectively. With the exception of the first calibration P1, optimized parameters from previous calibrations were utilized in H1\textsubscript{106}, H1\textsubscript{102}, and H1\textsubscript{104}. Optimized parameter SAT from P1 was not utilized at the hillslope scale because the events used at the hillslope scale had a minimum 24 hour period without precipitation, as opposed to the saturated condition of the rainfall simulator events; SAT was defaulted to 0.2 instead. Optimized parameter KS from P1 was not used at the hillslope scale either, because of the relationship between rainfall intensity and infiltration (Hawkins and Cundy, 1987; Paige and Stone, 2003; Paige et al., 2002; Stone and Paige, 2003) and the high rainfall rates used in the rainfall simulator events. Calibrations H2\textsubscript{106}, H2\textsubscript{102}, and H2\textsubscript{104} followed H1\textsubscript{106}, H1\textsubscript{102}, and H1\textsubscript{104} with the calibration of sediment transport capacity (TC), interrill covered fraction (IC), and shear stress adjustment factor (TA) against the sediment yield of the events. H2\textsubscript{106}, H2\textsubscript{102}, and H2\textsubscript{104} used optimized parameters from previous calibrations along the same calibration path (with the exception of P1, for stated reasons).

Moving to the watershed scale in group 1 and continuing from the H2\textsubscript{106} and H2\textsubscript{102} calibrations (H2\textsubscript{104} was not used as it represents the entire modeling area), separate calibrations were performed using LH106 and LH102 as the representative hillslopes. Prior to the watershed scale, all calibration parameters were plane element parameters because they represented processes more appropriate to the calibration scale and because a channel network was not yet represented; at the watershed scale, all calibration parameters were channel element parameters because channel processes begin to dominate at larger scales (plane element parameters used calibration values from the hillslope calibrations). In W1\textsubscript{106} and W1\textsubscript{102}, channel parameters KS and coefficient of variation of KS (CV) were calibrated against the hydrographs from 18 events. Optimized parameters from previous calibrations along the same calibration path were used (with the exception of P1). W2\textsubscript{106} and W2\textsubscript{102} followed W1\textsubscript{106} and W1\textsubscript{102} with the calibration of channel parameter MAN against the hydrographs from the same events; optimized parameters from previous calibrations along the same calibration path were used (with the exception of P1). W3\textsubscript{106} and W3\textsubscript{102} followed W2\textsubscript{106} and W2\textsubscript{102} with the calibration of channel parameters rain splash erosion coefficient (SPLASH) and soil cohesion coefficient (COH) against the sediment yields from the same events; optimized parameters from previous calibrations along the same calibration path were used (with the exception of P1).
Calibration Group 2 Calibration group 2 (Figure 4) forgoes calibration at the plot and hillslope scales and instead performs a lumped calibration at the watershed scale beginning with WL1\textsubscript{106} and WL1\textsubscript{102}. Plane and channel parameters were calibrated simultaneously against the hydrographs from 18 events for WL1\textsubscript{106} and WL1\textsubscript{102}. WL2\textsubscript{106} and WL2\textsubscript{102} followed WL1\textsubscript{106} and WL1\textsubscript{102}, respectively, with the simultaneous calibration of plane and channel parameter MAN against the hydrographs from the same events; optimized parameters from the previous calibration along the same calibration path were used. WL3\textsubscript{106} and WL3\textsubscript{102} followed WL2\textsubscript{106} and WL2\textsubscript{102}, respectively, with the simultaneous calibration of plane parameters RMAN, KI, KR, TC, IC, and TA and channel parameters SPLASH and COH against the sediment yields from the same events; optimized parameters from the previous calibration along the same calibration path were used. Finally, moving back to the hillslope scale with the same events, the optimized plane parameters from the three previous calibrations were used in HL\textsubscript{106} and HL\textsubscript{102} to evaluate and compare the performance of the step-wise calibration to the lumped calibration at the hillslope scale.

RESULTS AND DISCUSSION

The step-wise calibrations that use LH102 as the representative hillslope have the best overall performance in the calibrations as judged by the R-Index statistic and the slope and R^2 values of the 1:1 plots (Table 1). The R-Index (Equation 4) is a simple calculation that can quickly compare performance between calibrations and show if the calibration is under- or over-predicting the observed data.

$$R-Index = \frac{\text{predicted}}{\text{observed}}$$  \hspace{1cm} (4)

The increased performance of the LH102 based calibrations over the others could be a result of several factors. First, the size of the LH102 hillslope may simply be the best representative hillslope size for this study; LH106 as a hillslope is too small and LH104 as a hillslope is too large. Parameters, such as KS, cover, and MAN, are spatially variable parameters but each element can only have one value. There is a size threshold above or below which the single value may not represent the processes occurring at that particular scale well. LH102 appears to be closer to that threshold than the other watersheds. For the most part, the optimized parameter values for the LH102 based step-wise calibrations are not problematic. The problematic parameter for the LH102 based step-wise calibrations is channel KS (1000 mm/hr); channel KS is problematic for the other calibrations as well. Because channel KS is the dominant active calibration parameter affecting runoff rates at the watershed scale, it is forced to unrealistically large values to try and minimize the residuals of the high runoff rates (for large events, the magnitude of the residuals for high runoff rates is larger than the residuals of low runoff rates). The impact of this is observed in the R-Index and slopes of the runoff volumes and peak runoff rates because the runoff volumes show a larger trend of underestimating than do the peak runoff rates (Table 1). This is because as channel KS is increased to reduce the residuals of the high runoff rates in the hydrographs, the runoff volumes decrease. Channel MAN is also larger than hillslope MAN, which is the opposite of what is observed in the real world. Both MAN values are within realistic ranges however, and may be explained by the improved performance of the hillslope over the watershed scale calibrations.
As expected when calibrating against a multiple events with a range of rainfall inputs and runoff and sediment yields, some will perform better than others as the optimized parameter values will be a compromise between the large and small events. This also results in the intermediate events outperforming both the large and small events. This is observed at both the hillslope and watershed scale in Figure 5, which shows a small event, intermediate event, and large event. The small and large events are under- and over-estimated, but the intermediate event matches the observed data well.

The poor performance of some events may also be attributed to using an incorrect initial soil moisture, despite taking steps to minimize this. This is highlighted by some small events performing well while others perform poorly, though observed hydrographs and hyetographs are comparable in magnitude and shape (Figure 6). Measured soil moisture data, if available, and the calibration of the matric potential parameter could improve the fitment of these events.

<table>
<thead>
<tr>
<th>Calibration</th>
<th>Volume Mean</th>
<th>Volume Slope</th>
<th>Volume R²</th>
<th>Peak Flow Mean</th>
<th>Peak Flow Slope</th>
<th>Peak Flow R²</th>
<th>Sediment Yield Mean</th>
<th>Sediment Yield Slope</th>
<th>Sediment Yield R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>0.976</td>
<td>0.979</td>
<td>0.994</td>
<td>0.827</td>
<td>1.129</td>
<td>0.931</td>
<td>1.823</td>
<td>1.277</td>
<td>0.893</td>
</tr>
<tr>
<td>H2_106</td>
<td>0.628</td>
<td>0.914</td>
<td>0.847</td>
<td>0.829</td>
<td>1.012</td>
<td>0.889</td>
<td>0.206</td>
<td>0.688</td>
<td>0.751</td>
</tr>
<tr>
<td>H2_102</td>
<td>0.754</td>
<td>1.055</td>
<td>0.903</td>
<td>0.829</td>
<td>1.012</td>
<td>0.889</td>
<td>0.682</td>
<td>0.857</td>
<td>0.82</td>
</tr>
<tr>
<td>H2_104</td>
<td>0.676</td>
<td>0.995</td>
<td>0.822</td>
<td>0.694</td>
<td>0.973</td>
<td>0.787</td>
<td>0.291</td>
<td>0.802</td>
<td>0.811</td>
</tr>
<tr>
<td>W3_106</td>
<td>0.704</td>
<td>0.884</td>
<td>0.885</td>
<td>0.754</td>
<td>1.038</td>
<td>0.853</td>
<td>0.519</td>
<td>0.968</td>
<td>0.884</td>
</tr>
<tr>
<td>W3_102</td>
<td>0.698</td>
<td>0.972</td>
<td>0.865</td>
<td>0.756</td>
<td>1.038</td>
<td>0.823</td>
<td>0.935</td>
<td>1.223</td>
<td>0.875</td>
</tr>
<tr>
<td>WL3_106</td>
<td>0.706</td>
<td>0.915</td>
<td>0.874</td>
<td>0.775</td>
<td>1.024</td>
<td>0.853</td>
<td>0.676</td>
<td>0.854</td>
<td>0.858</td>
</tr>
<tr>
<td>WL3_102</td>
<td>0.653</td>
<td>0.986</td>
<td>0.852</td>
<td>0.709</td>
<td>1.053</td>
<td>0.816</td>
<td>1.303</td>
<td>1.866</td>
<td>0.871</td>
</tr>
<tr>
<td>HL_106</td>
<td>0.491</td>
<td>0.788</td>
<td>0.788</td>
<td>0.563</td>
<td>0.872</td>
<td>0.852</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HL_102</td>
<td>0.603</td>
<td>0.973</td>
<td>0.879</td>
<td>0.692</td>
<td>0.948</td>
<td>0.825</td>
<td>0.602</td>
<td>0.805</td>
<td>0.793</td>
</tr>
</tbody>
</table>

Table 1: Summary table showing relative performance (green is better, red is worse).
Figure 5 Small, intermediate, and large events at the hillslope (left) and watershed (right) scales depicting both good and bad hydrograph simulation.

Figure 6 Similar events showing a poor and strong performance, possibly due to inaccurate initial soil moisture conditions for the poor performing event.
CONCLUSIONS

The step-wise calibrations were able to outperform the lumped calibrations for both hydrology and sediment. Hillslope parameters from the lumped calibration at the watershed scale were used to compare the performance of a single hillslope element using those parameters to a single hillslope element using parameters from the step-wise calibration; although the lumped calibration at the watershed scale performs somewhat comparably to the final calibration of the step-wise calibrations, it clearly is not representing hillslope processes well and is relying solely on channel processes to increase calibration performance.

Because the step-wise calibrations were moderately successful in simulating observed hydrographs and sediment yields, and they clearly outperformed the lumped calibrations at the hillslope scale, it shows there is value in collecting data at multiple scales. In this study KS could not be scaled from rainfall simulator plots to the hillslope or watershed due to the difference in rainfall intensities and relationship between rainfall intensity and infiltration rate; however, the other parameters could be scaled from the plot to the hillslope and from the hillslope to the watershed effectively. Constraining the parameter space through the step-wise, multi-scale calibration process reduces uncertainty associated with applying parameters derived in the calibration to other scales in comparison to the lumped watershed scale calibration. This is encouraging for those who have invested time and effort in plot scale studies and data collection, in contrast with the discouraging discovery in the past that KS values from plot scale studies do not scale up.

REFERENCES


