

Evaluating Multi-Scale Flow Predictions for the Connecticut River Basin

Muhammet Omer Dis¹, Emmanouil Anagnostou^{1,*}, Flamig Zac², Humberto Vergara² and Yang Hong²

¹Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut, USA ²School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma, USA

Abstract

This case study evaluates a computationally efficient distributed hydrological model, named Coupled Routing and Excess Storage (CREST), for flood modeling of basins in the Connecticut River Basin (CRB). Simulation of discharges is performed by forcing CREST with a long record (eight years) of high resolution radar-rainfall data and potential evapotranspiration maps derived from the North American Regional Reanalysis. The model performance is evaluated against observed streamflows obtained from United States Geological Survey gauging stations at outlet and interior points of various CRB sub-basins. CREST parameters were calibrated based on a three year record (2005-2007) and validated for the remaining data period (2003-2004 and 2008-2009). The model performance evaluation is based on different metrics, including the Nash-Sutchliffe Coefficient of Efficiency (NSCE), Mean Relative Error (MRE), Root Mean Square Error (RMSE), and Pearson Correlation Coefficient (PCC). The analysis shows that CREST slightly underestimated the peak flows, but exhibited a generally good capability in simulating the stream flow variability for the CRB basins. Specifically, NSCE, MRE, RMSE, and PCC values of hourly flow simulations varied from 0.31 to 0.58, -0.06 to 0.13, 61 to 121 (mm) and 0.60 to 0.83, respectively. At daily time scale the performance metrics exhibited improved values indicating that CREST has sufficient accuracy for long term multi-scale hydrologic simulations.

Keywords: Distributed hydrological model implementation; Calibration and validation; High temporal resolution; CREST; Connecticut river basin

Introduction

Rainfall-runoff modeling has a long history; the first hydrologist that used rainfall-runoff model was an Irish engineer Thomas James Mulvaney (1822-1892) who published his work in 1851. During the last few decades, a number of conceptual and physically-based models have been developed and used for simulation of floods [1-6]. The analysis of hydrological model simulations and their spatiotemporal fluctuations can be used as vital tools to support management activities such as flood risk management, water supply and improving water quality. According to Brakendridge et al. [7], their study shows that an increasing population density has caused a greater risk regarding the natural processes, such as flooding, which makes it crucial to identify such risky areas. Dehotin and Braud [8] mentioned the importance of distributed hydrologic models, indicating that they are valuable mechanisms to spatially and realistically study the prediction of water balance components. Hydrologic modeling can be employed to evaluate flood mitigation alternatives to flood and drought risks. They can be used to evaluate and study the impact of land use on water resources. Additionally, Moriasi et al. [9] states that hydrological modeling is an essential tool for managing water quantity and quality. The spatial structure of distributed hydrological models can inspect and evaluate the heterogeneities of watershed characteristics and its parameters [8,10].

Even though distributed models capture sufficient details and realistic catchment characteristics, over parameterization can be a problem in calibration. Dehotin and Braud [8] stated the growing of concerns about optimum parameterization. They indicated that the major concern is the contrasting aspects of model complexity versus data availability. Rozalis et al. [11] stated that simplicity versus complexity of hydrological models is still a controversial subject for better representation of a catchment. In their study, a simpler model was used with minimum number of parameters. The smaller number of parameters does not require calibration to decrease uncertainty over ungauged areas. Moreover, Bergstrom [12] claimed an excessive increase in model complexity does not always improve the quality of the results. Brandt [13] supported this statement and further elaborated in his study that going from complex to simpler model does not affect model performance. Thus, using a model with optimal complexity relative to data availability and resolution is the key to improving hydrologic predictability. Also, it is worth mentioning that one of the important parameterizations in the modeling are the initial conditions and associated parameters sets. The initial condition of a model mostly depends on catchment area, catchment topography, antecedent moisture condition, ground water table position and land use. Hence, for distributed hydrological modeling, parameterizations of these initial conditions and its spatial variability are the fundamental factors for runoff simulations, especially for extreme rainfall events [14]. The influence of initial conditions on the model output reduces with increasing the simulation time period [15]. Additionally, a spinup or run-up period can be used to diminish the sensitivity to initial conditions. In this study a one-year spin up period was used in the model simulation. Therefore, the number of initialization parameters was efficiently reduced.

Two other significant issues for hydrological modeling are related to calibration and validation methodology, specifically the metrics used to analyze and evaluate simulation results [9,10,12,16,17]. In this study, we evaluated the accuracy of a grid-based distributed hydrologic model,

*Corresponding author: Emmanouil Anagnostou, Civil and Environmental Engineering, UCONN, USA, Tel: 860-486-6806; E-mail: manos@engr.uconn.edu

Received June 09, 2015; Accepted June 29, 2015; Published July 03, 2015

Citation: Omer Dis M, Anagnostou E, Zac F, Vergara H, Hong Y (2015) Evaluating Multi-Scale Flow Predictions for the Connecticut River Basin. Hydrol Current Res 6: 208. doi:10.4172/2157-7587.1000208

Copyright: © 2015 Omer Dis M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

which was calibrated and validated in the Connecticut River Basin. The performance of the model related to calibration and verification was evaluated using the Nash-Sutchliffe Coefficients of Efficiency (NSCE), Mean Relative Error (MRE), Root Mean Square Error (RMSE), and the Pearson Correlation Coefficient (PCC), which compared supplemental normalized flow simulations with obtained hourly observations obtained for the basin. For this analysis, a continuous simulation was applied for the hydrological analysis in the Connecticut River Basin using the Coupled Routing and Excess Storage (CREST) Hydrological Model. CREST model was selected for this study because the model currently represents both a national [18] and global flood simulation system [1]. Similar to other distributed models, catchments are represented as grid cells to simulate spatio-temporal variation of water, energy fluxes, and storage.

One of the purposes of this work is to examine the performance of the CREST model with calibrated parameters over interior sub-basin stations. Calibrating a model at a larger scale basin and using these parameters to model un gauged locations is one of the popular ways in situations where no observations are available. According to Bingeman et al. [19], a calibrated hydrological model based on stream flow data should be applicable on other sub-catchments without recalibration of the parameters set. State variables of basin characteristics related to land cover and soil texture for a given set of watersheds can be applied to other watersheds that are hydrologically similar without extensive recalibration [20]. Kouwen et al. [21] study also confirmed the fact that a parameter set can effectively transferred between watersheds to simulate peak flow without further calibration in southern Ontario. Additionally, Xie et al. [22] transferred calibrated parameters from data-rich areas to data-sparse areas and the results showed promising in estimating daily runoff with transferred variables. Hence, CREST was calibrated against stream flow observations at sub-basin outlets and was used to simulate stream flows based on the same parameter set at its interior nested sub-basins without further calibration.

This case study aims to evaluate the multi-basin scale predictability of stream flow in the Connecticut River Basin using the CREST model following calibration procedures suggested in the literature. Potential impacts of these fluctuations are important for water supplies, during the peak seasons of water demand, and mitigating flooding risk associated with peak flows. This study presents an opportunity to enhance and refine the estimation of hydrological processes via the CREST model and understand its calibration, parameterization and validation procedures over a mid-latitude basin. The main objectives of this work are: (1) use available observations to assess the accuracy of simulations, (2) determine the spatial variability of calibration parameters, and (3) understand the applicability of calibrated parameter values applied to interior catchments in order to predict flows at un gauged basin locations. Based on the topography and USGS observation gauges, the Connecticut River Basin was divided into nine sub-basins for all CREST simulations. Knowledge of the fluctuations and the accuracy of the predictions can assist in performing simulations over un gauged locations. For this purpose, three of the sub-basins, which have enough interior locations for un gauged analysis, were further subdivided into sub-watersheds in order to perform an analysis of parent basin calibration parameter runs.

The paper is organized as follows. Section 2 describes the study area, data, and the CREST hydrological model. Section 3 provides information regarding the model calibration, parameterization and validation with analysis of evaluated statistics. Section 4 provides a discussion of the results of the hydrological simulation using the Page 2 of 13

CREST model for the Connecticut River Basin. The conclusions and future work are discussed in the last section.

Study Region and Data

The Connecticut River Basin (CRB) is the study basin for this work and the CRB is a major river basin in New England. Runoff from the CRB discharges to the Connecticut River. The Connecticut River starts in Quebec Canada and runs through Connecticut (CT), Massachusetts (MA), New Hampshire (NH), and Vermont (VT) and empties into Long Island Sound in Old Lyme Connecticut. The total watershed area contributing to the Connecticut River is approximately 28,500 km². There are approximately 390 towns and cities located within the watershed with a total population of approximately 2.3 million people. The land uses that are within the watershed consist of forest, agriculture, residential and water. Approximately, 79% is covered by forested, 11% by agriculture and the remaining area is covered by residential and water. The CT River flows for about 660 km and provides hydroelectric power, is navigable up to Windsor Locks, used for irrigation, and is used for recreation [23].

Instantaneous records of river discharges obtained from nine United States Geological Survey (USGS) gauging stations within the Connecticut River Basin were analyzed in this study. The location of the gauging stations and their contributing sub-basins are illustrated in Figure 1. The black dots represent the streamflow gauges, which are labeled with USGS station numbers. The locations of USGS monitoring stations were used to determine the number of sub-basins based on the location and topography and their watersheds area. Given a particular gauging station along with the use of DEM data, we are able



Figure 1: Map of Connecticut River Basin showing the drainage areas of the subbasins used in this study. The black dots represent streamflow gauges labeled with USGS station numbers.

to determine the area contributing runoff at each station. The data from the nine stream gauges were used to calibrate and verify continuous simulations from the CREST model across watersheds with drainage areas ranging from 200 to 25,000 km².

Zanon et al. [24] has argued that the flood response could be reasonably well reproduced when using high resolution rainfall observations. Even though continuous rain gauge data are available for the area, there are several gaps in these data records. As Hirpa et al. [25] reported unequal time period data between observation gauges might affect the certainty of parameter estimations. Additionally, Du et al. [26] have stated that there are many stations in the USA that do not have extensive flow records or have limited flow data. The gauge data measurements are available in 5-min, 15-min, and 30-min intervals, with 15-min intervals being the most common. Thus, the finer time scale gauge measurements were converted into hourly stream flow averages in units of m3/s from January 2002 to December 2009. The input data sets consist of gridded radar-rainfall (mm/h) and Potential Evapotranspiration (PET) (mm/3h) data. The precipitation data was extracted from the WSR-88D Stage IV product obtained from the North American Regional Reanalysis. The Stage IV radar rainfall fields were used to force the model at the hourly time step and 4 km spatial grid resolution. PET data based on the North American Regional Reanalysis (NARR) available at 32 km spatial and 3-hourly temporal resolution [27] were used as forcing variable for the hydrological model.

In addition to the radar rainfall and the PET data, High Resolution Digital Elevation Model (DEM) data was used to delineate the various watersheds and to perform river routing of the stream flow simulations.

Methods of Analysis

Model

A raster-based distributed hydrologic model, Coupled Routing and Excess Storage (CREST) was implemented over Connecticut River Basin. The CREST model is a hybrid modeling strategy, developed by the University of Oklahoma and NASA SERVIR-an acronym meaning

Parameter Name	Description	Suggested range
coem	The slope flow speed multiplier	3.33, 100.0
river	The channel flow speed multiplier	3.33,100.0
under	The horizontal velocity under the ground (Generally hydraulic conductivity used for this velocity)	0.01,51.0
leako	The overland reservoir discharge multiplier	0.01,1.0
leaki	The interflow reservoir discharge multiplier	0.01,1.0
th	The flow accumulation needed for a cell to be marked as a channel cell. If a cells flow accumulation is greater than th then the cells slope flow speed is multiplied by river	0.01,30.0
pwm	The maximum soil water capacity (depth integrated pore space) of the soil layer	0.01,250.0
pb	The exponent of the variable infiltration curve.	0.01,4.0
pim	The impervious area ratio	0.00,100.0
pke	The multiplier to convert between input PET and local actual ET	0.01,1.0
pfc	The soil saturated hydraulic conductivity	0.01,51.0
iwu	The initial value of soil water. (This is a percentage of the pwm)	0.0,100.0
iso	The initial value of overland reservoir	0.01,100.0
isu	The initial value of interflow reservoir	0.01,100.0

Table 1: CREST model parameters with their value ranges.

to "to serve" in Spanish- Project Team [1]. The model is spatially distributed rainfall-runoff model dedicated to simulate flow discharges at regional and global scales aimed to represent hydrological processes associated with floods and droughts. Distributed CREST model can be applicable to almost any kind of hydrological problems and provides important advantage over existing models under different land cover and soil type scenarios with user defined spatiotemporal resolutions. The CREST model provides water managers with information about streamflow amounts that can enable better decision-making regarding water resources, floods, and agriculture.

CREST model uses a combination of DEM, PET, and precipitation input data with different user defined spatiotemporal resolutions. The processes modeled include rainfall-runoff generation and capacity for cell-to-cell routing, canopy interception, infiltration, evaporation, recharge baseflow, sub-grid cell variability of soil moisture storage capacity and routing processes at the sub-grid scale. The model controls the maximum storage of the infiltrating water and yield surface runoff generation with connected layers within the soil profile. Cell-to-cell flow routing of direct surface runoff is applied using a kinematic wave assumption. Besides, coupling between the flow simulation and routing component via feedback mechanisms provides realistic applications of the hydrologic variables (i.e. soil moisture). The CREST model has been discussed extensively in previous papers [1,28] and details of the model can be found therein.

CREST model contains several default parameters. These parameters have value ranges initially specified based on land cover and soil type data. The definitions of parameters with their value ranges for CRB are listed in Table 1. The model contains 14 parameters, and three of them are related to initial soil conditions, (the initial value of soil water "iwu", initial value of overland reservoir "iso", and initial value of interflow reservoir "isu") that can be adjusted using warmup period. The majority of the parameter ranges come from physical considerations. For example, the slope flow speed multiplier (coem) can be regarded as the inverse of manning's roughness, and "river" is similar to "coem" but for river channels [29]. Parameter "under" represents the horizontal velocity of subsurface flow, and hydraulic conductivity is used for this velocity [1,30]. "leako" (leaki) is overland (interflow) reservoir multipliers, which varies from 0.01 to 1 [1]. "pwm" is the capacity of the soil to hold water [31]. "pb" is a parameter related to infiltration and described as the exponent of the variable infiltration curve while 'pim' is the percentage of impervious area, which is derived from land cover data [1]. "pke" is a multiplier factor to convert PET to local actual evapotranspiration [29]. "pfc" is the soil saturated hydraulic conductivity which is derived based on soil type data [1]. The initial values of these parameters are adjusted through a calibration procedure utilizing using measured stream flows at gauging stations [6]. This is discussed at the next section.

Model Calibration and Validation

CREST model calibration and validation were carried out over Connecticut River Basin at various selected observation stream gauges (Figure 2). Bergstrom [12] defined model calibration as a process that model parameters are arranged to make model results to meet the measurements. However, if the number of parameters used in the calibration is large, automatic calibration is a better option to reduce labor-intensive. Additionally, automatic approach for model calibration abbreviates the time with the advantage of speed and power of high performance computers as well as the approach eliminates the kinds of subjective human judgments [9,16]. The auto-calibration routine based on Differential Evolution Adaptive Metropolis (DREAM) method

Page 4 of 13



was devised to calibrate CREST over a number of watershed areas in Connecticut River Basin based on hourly stream flow data for the period 2005 to 2007. The algorithm runs to mitigate sum of squared residuals to estimate posterior parameters using multiple Markov Chain Monte Carlo chains from prior iterations [15].

The numbers of parameters which were subject to calibration were limited to 14 as defined by Wang et al. [1]. Calibrated values are listed in Table 2 along with the acceptable intervals (Table 1) used to constrain the parameters search. Each basin was calibrated and validated separately to capture the spatial variability of the parameters over the region. Even though all parameters are in the acceptable range, the parameter values vary across basins. For example, the slope flow speed multiplier "coem" varied between 10.39 and 74.33 with a mean of 34.93 and 0.61 coefficient of variation. The maximum soil water capacity "pwm" parameter has the maximum coefficient of variation, i.e 1.15.

Parameter Name	01192500	01127500	01193500	01208500	01188090	01122500	01127000	01205500	01184000
coem	74.33	43.52	53.54	50.06	14.55	12.3	18.07	10.39	37.63
river	99.19	80.89	99.06	86.67	79.61	98.53	80.02	90.45	7.16
under	33.69	7.25	38.96	0.9	50.43	4.16	35.3	13.64	17.79
leako	0.02	0.36	0.74	0.4	0.08	0.02	0.1	0.23	0.1
leaki	0.94	0.72	0.84	0.6	0.73	0.53	0.52	0.42	0.23
th	24.76	16.81	15.29	26.82	21.85	13.32	11.62	16.72	13.69
pwm	32.89	3.03	7.59	1.51	2.52	1.91	3.02	7.56	14.3
pb	2.06	0.49	2.45	2.46	0.79	1.31	2.31	1.79	0.38
pim	1.42	0.93	0.99	0.04	0.85	1.08	1.5	1.05	0.69
pke	0.44	0.95	0.99	0.84	0.79	0.17	0.96	0.84	0.72
pfc	44.88	3.48	38.37	50.49	45.68	37.24	35.36	39.72	46.45
iwu	71.33	0.59	91.16	75.92	7.86	91.28	70.48	42.6	89.78
iso	9.17	15.14	2.11	4.77	92.52	40.6	59.05	65.2	80.21
isu	1.45	51.77	12.9	81.45	13.06	48.05	98.21	26.71	85.11

 Table 2: Parameters values determined from calibration for each gauged sub basin.

The simulation starts with a warm-up period to reduce the initial condition effect. Specifically, the three parameters related to initial conditions (iwu, iso, and isu) are model states that were adjusted automatically by running CREST with a warm-up period. Marshall and Randhir [23] observed in their study using a 40-year study period over Connecticut River Basin that the maximum snow accumulation occurs in the month of January and decreases exponential through April. To avoid snowmelt effects in parameter calibration the period of January through the middle of March was excluded from the error analysis.

Finally, after parameter calibration, CREST was validated for a four year period. Refsgard [10] defined model validation as a process of illustrating that the calibrated parameter set without further adjustments is capable to reproduce flows in different times other than the calibration period. This process was done again without any interruption during snow accumulation and melting span, but we excluded the snow process period (January 1-March 15) from the analysis.

Evaluation Indexes

Different error metrics have been used in hydrologic model validation exercises of past studies [9,16,32]. Multi criteria error functions aid to explain various aspects of hydrographs. In this study, the model efficiency to predict stream flow at basin outlets was demonstrated qualitatively by plotting time-series of observed and simulated stream flows and determining error statistics based on hydrographs normalized by the corresponding catchment area. The error metrics are listed below:

Nash-Sutcliffe Coefficients of Efficiency (NSCE):

$$NSCE = 1 - \frac{\sum_{i=1}^{n} (Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^{n} (Q_{o,i} - \overline{Q}_o)^2} \dots \dots (1)$$

where $Q_{o,i}$ is measured stream flow of the time ith; $Q_{s,i}$ is simulated stream flow of the time ith; $\overline{Q_o}$ is the average of entire observed streamflow values; and n is the total number of time steps. The value of NSCE varies from -∞ to 1. On the one hand, a value of NSCE \leq 0 represents that the model does not have capability to use the observed mean as a predictor; on the other hand, NSCE=1 indicates that simulation results are capturing the measurements perfectly.

Mean Relative Error (MRE):

$$MRE = \frac{\sum_{i=1}^{n} (Q_{s,i} - Q_{o,i})}{\sum_{i=1}^{n} Q_{o,i}}$$
(2)

MRE gives an indication of how close predictions are relative to the observations. A value of MRE=0 shows that the simulated total amount of discharges is unbiased to observations.

Root Mean Square Error (RMSE):

$$RMSE(\%) = \frac{100}{\overline{Q}_{n}} * \sqrt{\frac{\sum_{i=1}^{n} (Q_{o,i} - Q_{s,i})^{2}}{n}} \dots$$
(3)

RMSE measures the magnitude of the differences between simulated and observed discharges relative to mean observed discharge value. Therefore, a low RMSE indicates better fit and the value of zero signifies the perfect fit.

Pearson Correlation Coefficient (PCC):

$$PCC = \frac{\sum_{i=1}^{n} (Q_{o,i} - \overline{Q}_{o})(Q_{s,i} - \overline{Q}_{s})}{\sqrt{\sum_{i=1}^{n} (Q_{o,i} - \overline{Q}_{o})^{2} * \sum_{i=1}^{n} (Q_{s,i} - \overline{Q}_{s})^{2}}}$$
(4)

where $\overline{Q_s}$ is the average of entire simulated streamflow values. PCC represents how well the linear relationship between measurements and predictions. This value ranges from -1 to 1.If PCC=0, then there is no relation between the two variables. The closer either -1 or 1 indicates stronger correlation between them.

Boyle et al. [16] have shown that RMSE is sensitive to the peak flows and can strongly bias recession error characteristics. PCC, on the other hand, reflects the collinearity between simulations and observations and correlation-based measures have excessive sensitivity to peak flows [32]. Moreover, MRE measures the average tendency of simulated against observed data and explains this tendency with overestimations or underestimations [9]. Servat and Dezetter [33] pointed out that even though NSCE error metric shows some weakness with low flows, it is the best objective function to provide extensive information on hydrograph prediction accuracy.

Results

CREST model was applied for the period between 2002 and 2009 to represent the hydrologic response of the Connecticut River Basin and several of its sub-basins. The calibration of the model was carried out over nine sub-basins from 2005 to 2007. Then, CREST stream flow simulations were compared to the measured flows at the outlets of the nine sub-basins and interior locations of three of the catchments over the seven-year time period using 2002 as the model spin-up. Simulations for the various basin sizes are summarized below showing an overall good agreed with observations for different event magnitudes.

In Figure 3, observed and simulated normalized hourly discharges were plotted along with measured rainfall for the period 2004 and 2005 at multiple stations. For brevity, three of the spatially various CRB sub-basins associated with different sizes are illustrated. Results, in the Figure 3, from first and second columns belong to validation and calibration periods, respectively. Three different size basins (small, middle, and large scale catchments), namely B0128500 (673 km²), B01122500 (1,046 km²), and B01184000 (25,019 km²) catchments, are visualized from top to bottom. As it can be seen from the hydrographs, overall, model performed well over CRB and the variability depicted in the flow simulation is close agreement with observed flows. Even though the timing of the peaks is estimated well, CREST simulations tend to slightly underestimate some of the peak flows. As it can be seen from the hydrographs in April 2005 simulations underestimated observations associated with small amounts of rainfall. However, it would be expected that light rainfall produce smaller runoff and this mismatch in terms of peak discharges can be explained by uncertainties in either rainfall or stream flow observational data. However, although slight underestimation is captured in qualitative stream flow comparisons, the majority of the simulated discharges were attained close to the measurements.

The model outputs from the hydrological simulations are summarized in Table 3A and 3B for calibration and validation periods, respectively. Basin areas, observed and simulated mean discharges, error metrics results (NSCE, MRE, RMSE, and PCC), and percentage of unavailable observations during the analysis are illustrated. During calibration period, as it can be seen from Table 3, NSCE values vary between 0.31 and 0.68 for the nine-basins while MRE results range from -6% to 13%. Additionally, 0.60 is the smallest value with hourly time resolution for PCC metric. The demonstrated NSCE, MRE, and PCC values show that the model simulations have good agreement with measurements in the calibration data period. In terms of RMSE error metric, however, at some stations, it is difficult to appreciate the quality of the results.

In the validation period, on the other hand, Nash is lower (ranging between 0.12 and 0.58) for hourly resolution, while PCC values ranged between 0.42 and 0.77 for the nine-basins. It is interesting enough, during validation, overall MRE exhibited underestimation in the range of 4% to 26%. However RMSE values dropped between 2 and 20% in the basins, which indicate improved random error.

Moriasi et al. [9] have shown that hydrologic model performance is better for longer time steps (i.e. annual versus monthly) and claimed that simulation statistics improve as a function of time resolution. Fernandez et al. [34] supported this statement and reported that NSCE values during calibration in their study in the range of 0.36 and 0.66 for daily and monthly time resolutions, respectively. Hence, daily error metrics were calculated and illustrated in Table 3C.

The study illustrates a better agreement between estimated and measured flows at daily scale and model results become quite realistic. For instance, in B01205500, NSCE values increase from 0.31 to 0.47 during calibration and from 0.16 to 0.30 in validation period for hourly and daily resolutions, respectively. While PCC values improve from 0.60 to 0.70 (calibration period) and 0.46 to 0.59 (validation period) as a function of time resolution in the same catchment, no change is observed in terms of the MRE values, which is expected given that resolution primarily affects the random component of error. Similar improvements are also captured for the other basins (Table 3C). Moreover, RMSE values drop significantly and become less than 100% during both validation and calibration in the basins with the exception of B01193500 (103.91%).

Analysis of the error metrics with their annual fluctuations were, additionally, examined in Figure 4. The figure provides an illustration of simulations versus observation streamflow variability. Results reported in Figure 4A-D show the hourly calculated NSCE, MRE, RMSE, and

Page 6 of 13



PCC statistics, respectively for each year in the 7-year time period. Multiple transparent dots represent the basin statistic and darkness of the dots is increased as a function of watershed size. Analysis of the results acquired for 7-year gives more reasonable amount for smaller basins. The majority of NSCE, MRE, PCC values ranged in acceptable levels during both calibration and validation. As expected, NSCE and PCC values for calibration are better than the validation periods and increase with longer time steps. Moreover, model outputs are quite realistic in between 2004 to 2008, while results diverged in the year 2003 and 2009. Thus, that can yield a bias into the optimization process with low NSCE, RMSE, and PCC.

Three nested subbasins were considered as ungauged interior catchments namely B01122500 (Figure 5), B0112700 (Figure 6), and B01205500 (Figure 7). After model was calibrated for the outlet of the parent basin, the parameters were fixed, and then CREST was used to simulate the interior catchments. Figures 5-7 show the comparison between model outputs and measurements via statistical error functions with annual variability. The catchments were ordered with increasing area. Differences in model responses increase with increasing size of the catchment. Overall, results show that there is good agreement between the predictions and the measurements.

In Figure 5, statistical metrics of NSCE, MRE, RMSE, and PCC

presented for different years and different interior subbasins of parent basin B01122500. The results show that the highest NSCE values are obtained in the larger basin (B01122500) while the small basin (B0119500) exhibits the poorest agreement in terms of the error metric. Basin size dependency is also observed in RMSE metric values with the highest value reported for the smaller basins. MRE results, on the other hand, are within +/- 0.5, and the overall performance of the model in terms of PCC is above 0.5.

To test calibration, validation periods, and resolution effect on the error metrics, Table 4 statistics were calculated and showed the results as a summary. In Table 4, basin statistics are visualized at hourly and daily resolution during calibration and validation periods for ungauged interior subbasins of parent B01122500 basin. As it can be seen from table, number of time steps in the observed data is decreased as a function of catchment area. In addition, it is important to note that discharge point of small basin is at the furthest location from the calibrated downstream outlet. It is expected that the simulation performance is decreased as a function of catchment area. Additionally, slight improvements are observed changing time resolution from hourly to daily.

Table 5 shows calculated basin statistics for interior subbasin of parent B01127000 at hourly time resolution during calibration

Page 7 of 13

Gauge Name	Basin Area	Basin Area (A)Calibration										
Gauge Name		Q _o [n	n³/s]	Q [m³/s]		NaN Data		MDE	DM05 (9/1	500		
	[km²]	μ	σ	μ	σ	[%]	NSCE	MRE	RMSE [%]	PCC		
01192500	190.11	4.38	4.74	4.85	3.89	0.11	0.41	0.11	83.28	0.67		
01127500	231.29	5.31	8.55	5.42	7.64	5.29	0.68	0.02	90.59	0.83		
01193500	259.00	6.70	10.62	7.11	8.58	4.90	0.42	0.06	121.25	0.66		
01208500	673.40	14.03	18.98	13.79	14.20	17.13	0.35	-0.02	109.36	0.61		
01188090	979.02	22.32	29.14	25.28	28.85	3.92	0.40	0.13	101.32	0.70		
01122500	1,046.36	23.09	26.85	24.75	18.98	1.77	0.64	0.07	70.03	0.81		
01127000	1,846.66	41.74	44.58	45.85	34.66	0.10	0.54	0.10	72.34	0.74		
01205500	3,998.94	90.83	114.55	89.73	94.74	1.08	0.31	-0.01	104.98	0.60		
01184000	25,019.28	690.49	531.64	648.82	452.39	1.22	0.37	-0.06	61.18	0.65		

	Basin Area					(-,				
Gauge Name		Q ₀ [I	n³/s]	Q _s [m³/s]		NaN Data	NSCE	MDE		PCC
	[km²]	μ	σ	μ	σ	[%]	NOCE	MIKE		FCC
01192500	190.11	4.20	3.35	4.00	2.40	1.79	0.21	-0.05	71.12	0.50
01127500	231.29	5.84	7.87	5.03	6.54	12.00	0.58	-0.14	87.64	0.77
01193500	259.00	6.63	9.08	6.86	7.10	13.21	0.29	0.04	115.00	0.58
01208500	673.40	17.88	16.33	13.15	9.76	4.24	0.24	-0.26	79.68	0.57
01188090	979.02	24.06	16.83	22.85	16.02	3.64	0.24	-0.05	61.18	0.60
01122500	1,046.36	26.05	23.91	22.31	15.62	5.35	0.44	-0.14	68.54	0.68
01127000	1,846.66	43.90	37.72	43.43	29.58	4.06	0.37	-0.01	68.19	0.63
01205500	3,998.94	97.80	95.21	82.89	57.22	0.02	0.16	-0.15	89.12	0.46
01184000	25,019.28	578.45	455.76	529.10	295.32	3.12	0.12	-0.09	74.09	0.42

(B)VALIDATION

Gauge Name	Basin Area		Cali	bration	(C) Daily Resolution				
	[km²]	NSCE	MRE	RMSE	PCC	NSCE	MRE	RMSE	PCC
01192500	190.11	0.44	0.11	76.75	0.69	0.21	-0.05	64.82	0.53
01127500	231.29	0.72	0.02	79.38	0.85	0.60	-0.14	79.17	0.78
01193500	259.00	0.50	0.06	103.91	0.72	0.35	0.04	97.85	0.62
01208500	673.40	0.47	-0.01	91.83	0.69	0.34	-0.26	69.81	0.67
01188090	979.02	0.44	0.13	93.16	0.73	0.25	-0.05	57.58	0.61
01122500	1,046.36	0.65	0.07	68.51	0.81	0.44	-0.14	67.07	0.69
01127000	1,846.66	0.57	0.10	68.68	0.76	0.41	-0.01	64.84	0.65
01205500	3,998.94	0.47	-0.01	83.77	0.70	0.30	-0.15	65.96	0.59
01184000	25,019.28	0.38	-0.06	59.89	0.65	0.12	-0.09	73.27	0.43

Table 3: Gauged basin statistics determined at hourly time resolution during calibration period in (A) and during validation period in (B). Additionally, their daily statistics are reported in (C).

period in section (A) and during validation period in section (B). Additionally, their daily statistics are reported in part (C). It is observed that, overall NSCE results, in B01124000, are close to each other during calibration and validation periods. MRE value reduces from 0.41 to 0.23 during calibration and validation, respectively. While RMSE drops 20% from calibration to validation period, slight change is obtained in PCC with the value being around 0.7. Slight changes are also obtained as function of the time resolution for all the metrics in calibration an validation periods.

On the other hand, in Figure 6, variability is noted with respect to year for the same parent basin (B01127000) and nested catchment. The same trend is seen in the error metrics results of two basins spanning a 7-yr period with higher accuracy for the large basin (B01127000). This suggests that, overall the model provides a reasonably good description of the fluctuation.

Finally, calibrated largest parent basin (B01205500) is examined

with interior nested basins in Figure 7. The smallest basin (01199000) tends to overestimate in terms of MRE results in addition to negative bias for the other two basins. NSCE results are below 0.5 while PCC values are above 0.3 in the 7-yr period. RMSE values improve as function of basin size.

The model efficiency is summarized in Table 6 during validation (section A) and calibration (section B) periods (hourly) as well as daily resolution (section C) for B01199000, B01200500 (considered as internal ungauged basins) and B01205500 (parent basin). MRE results show that the values are in the range of -0.02 and 0.21 for calibration and -0.15 and 0.07 for validation. Correlation is around 0.6 during calibration while it drops to 0.46 for the parent basin and improves to 0.65 for the medium basin (B01200500) in validation period. Specifically for the daily simulations, all RMSE values are below 100%. Examination of these statistics reported in the table shows that observations of the interior basins are well represented with CREST simulations.

Page 8 of 13









Page 9 of 13

Page 10 of 13

Gauge Name Basin Are	Basin Area					(A)Calibra	ation			
	-	QO[m ³ /s]		QS	6 [m³/s]	NaN Data	NECE	MDE	DMSE 19/1	PCC
	[km ²]	μ	σ	μ	σ	[%]	NOCE	WIRE		
01119500	313.39	7.21	12.08	5.60	3.25	10.88	0.26	-0.22	144.00	0.66
01122000	450.66	10.43	12.61	11.18	13.31	2.16	0.39	0.07	94.36	0.71
01122500	1,046.36	23.09	26.85	24.75	18.98	1.77	0.64	0.07	70.03	0.81

- ··	Basin Area					(B)Validati	on			
Gauge Name		QO[ı	m³/s]	QS [m³/s]		NaN Data	NECE	MDE	DM6E 19/1	DCC
	[km²]	μ	σ	μ	σ	[%]	NSCE	WIRE	RIVISE [%]	FCC
01119500	313.39	7.69	7.71	5.03	2.11	16.02	0.09	-0.35	95.40	0.53
01122000	450.66	10.47	10.10	9.80	11.18	2.68	0.15	-0.06	89.14	0.62
01122500	1,046.36	26.05	23.91	22.31	15.62	5.35	0.44	-0.14	68.54	0.68

Gauge Name	Basin area		Calib	ration		(C) Daily Resolution Validation				
	[km ²]	NSCE	MRE	RMSE	PCC	NSCE	MRE	RMSE	PCC	
01119500	313.39	0.28	-0.22	135.32	0.67	0.09	-0.35	91.47	0.53	
01122000	450.66	0.44	0.07	88.54	0.74	0.18	-0.06	84.68	0.64	
01122500	1,046.36	0.65	0.07	68.51	0.81	0.44	-0.14	67.07	0.69	

 Table 4:
 Same as in Table 3, but for the ungauged interior subbasins of parent 01122500 basin.

Gauge Name	Basin Area		(A)Calibration											
- augo rumo		QO	m³/s]	QS [m³/s]	NaN Data	NECE	MDE	DM8E 19/1	BCC				
	[km ²]	μ	σ	μ	σ	[%]	NOCE	WIKE	RIVISE [%]	FCC				
01124000	401.45	9.32	11.04	13.18	8.97	4.70	0.35	0.41	95.20	0.70				
01127000	1,846.66	41.74	44.58	45.85	34.66	0.10	0.54	0.10	72.34	0.74				

Gauge Name	Basin area	(B)V	/alidatio	n						
eauge		QO[m³/s]	QS	[m³/s]	NaN Data	NECE	MDE	DMCE 10	
	[km²]	μ	σ	μ	σ	[%]	NSCE	WIRE	RIVISE [7	
01124000	401.45	9.75	8.72	11.99	7.57	9.25	0.34	0.23	72.72	0.67
01127000	1,846.66	43.90	37.72	43.43	29.58	4.06	0.37	-0.01	68.19	0.63
Gauge Name	Basin Area	Calibration				Validation	n (C) Daily I	Resolutior	1	
	[km²]	NSCE	MRE	RMSE	PCC	NSCE	MRE	RM	SE	PCC
01124000	401.45	0.38	0.42	92.52	0.71	0.36	0.23	69.	86	0.68
01127000	1,846.66	0.57	0.10	68.68	0.76	0.41	-0.01	64.	84	0.65

Table 5: Same as in Table 3, but for the ungauged interior subbasin of parent 01127000 basin.

Figures 8 and 9 illustrate the comparison of the simulations in terms of quantiles at the basins outlets and interior ungauged locations, respectively. MRE and RMSE error metrics were calculated at 0.2, 05, 07, 0.9 and 0.95 quantiles and shown with each spatially distributed station with transparency dots. Reduction of the RMSE and increment of the MRE tendency are observed from low to high flows. The poorer performances of the model in terms of MRE and RMSE are obtained at the higher quantiles and considered to be related to inadequate representation of the peak events. Figures confirm the tendency of underestimations at peak discharges and this consistency supports the argument of using frequency analysis with large flow data to improve model results.

Conclusions

This case study has demonstrated the implementation of the CREST

model, and its calibration and validation processes at high temporal resolution with spatial distributed gauged observations as well as ungauged/poorly gauged catchments over Connecticut River Basin. The primary objective of this study is calibration and validation processes of distributed hydrological model based on stream flow observations and provide explicitly statistical analysis of the simulations via different objective functions. Another objective of this study is producing flow data for insufficient historical records at gauging stations using calibrated CREST model and predicting flows over un gauged locations of the Connecticut River Basin. In this manner, the basin was partition into 9 sub basins to represent distributed parameters sets via CREST model based on the USGS monitoring stations and their locations, and the model was evaluated with several important statistics with supplemental hydrographs.

Page 11 of 13

	Basin Area					(A)Ca	libratic	n		
Gauge Name	Buomyarou	QO[m ³ /s]		QS [m³/s]		NaN Data	NOOF	MDE		DCC
	[km ²]	μ	σ	μ	σ	[%]	NSCE	MKE	RMSE [%]	PUU
01199000	1,642.05	35.87	39.35	43.37	44.76	6.37	0.00	0.21	109.91	0.58
01200500	2,579.63	57.99	65.09	56.90	42.27	2.36	0.36	-0.02	89.78	0.60
01205500	3,998.94	90.83	114.55	89.73	94.74	1.08	0.31	-0.01	104.98	0.60

Gauge Name	Rasin Area	(B)Validation											
	BusinAlcu	QO[m³/s]		QS [m³/s]		NaN Data	NECE	MDE	DM6E 19/1	DCC			
	[km ²]	μ	σ	μ	σ	[%]	NSCE			PCC			
01199000	1,642.05	40.75	35.37	43.46	35.54	2.81	0.30	0.07	72.67	0.65			
01200500	2,579.63	65.17	54.80	58.52	41.43	8.77	0.32	-0.10	69.13	0.60			
01205500	3,998.94	97.80	95.21	82.89	57.22	0.02	0.16	-0.15	89.12	0.46			

Gauge Name	Basin Area [km ²]	Calibration				Validation (C) Daily Resolution			
		NSCE	MRE	RMSE	PCC	NSCE	MRE	RMSE	PCC
01199000	1,642.05	0.17	0.21	99.55	0.62	0.36	0.07	68.88	0.67
01200500	2,579.63	0.39	-0.02	86.88	0.63	0.37	-0.10	66.33	0.63
01205500	3,998.94	0.47	-0.01	83.77	0.70	0.30	-0.15	65.96	0.59

Table 6: Same as in Table 3, but for the ungauged interior subbasins of parent 01205500 basin.



CREST model was calibrated for three years and validated in four years during the period 2002-2009. Calibration of the model was done over the period 2005 to 2007 and validations were carried out over two distinct periods: 2003 to 2004 and 2008 to 2009. The model was calibrated against 14 parameters within specified limited boundaries while a few of these parameters were adjusted using 2002 as spin up period for plausible initial conditions so that model predicts rainfallrunoff responses closely to the flow measurements and reflects the behavior of watershed system in a various manner. The figures and tables demonstrated the model's accuracy for the simulation of the rainfall-runoff transformation at 9 stations and un gauged interior locations. In this study, several statistical indicators and supplemental graphical illustrations were applied for evaluation of the CREST model performance. Using multiple statistics helps to cover a different aspect of the hydrographs. Based on the error metrics and graphical comparisons, CREST is considered to provide accurate rainfall-runoff simulations during calibration and validation period. The results also indicate satisfactory in model performance over Connecticut River Basin.

The CREST hydrological model estimates the hourly flow at spatially various catchments well, but with relatively large errors at peak quantiles. Overall, when we take into account statistical metrics and normalized hydrographs of the model results, it can be concluded that



the CREST is capable of reproducing continuous hourly stream flows in Connecticut River Basin, which allows model use for flood management applications in both gauged and un gauged basins. Results with CREST illustrated that model performance was satisfactory in representing amplitude and timing of the flow peaks, and that the models performs better in predicting the entire hydrograph. The model application and its calibrated parameters can be used for future work such as comparing to satellites-driven flow, flash flood predictions, flood frequency, and future climatic analysis on local and national scale as well.

Acknowledgements

This work was supported by the Northeast Utilities Foundation Endowment in Environmental Engineering. The authors would like to thank the anonymous reviewers for their valuable comments and feedback.

References

- Wang J (2011) The coupled routing and excess storage (CREST) distributed hydrological model. Hydrological Science Journal 56: 84-98.
- Odoni N, Lane S (2010) Knowledge-theoretic models in hydrology. Progress in Physical Geography 34: 151-171.
- 3. Dingman S (2002) Physical hydrology, Upper Saddle River, NJ: Prentice-Hall Inc.
- 4. Beven KJ (2000) Rainfall-Runoff Modeling. Chichester: John Wiley and sons Itd.
- Chow V, Maidment D, Mays L (1988) Applied Hydrology. International Editions. Singapore: McGraw-Hill.
- Crawford N, Linsley R (1966) Digital simulation in hydrology: Stanford Watershed Model IV. Technical Report 39, Palo Alto, CA: Stanford University.
- Brakenridge G (2003) Flood warnings, flood disaster assessments, and flood hazard reduction: the roles of orbital remote sensing. Pasadena, CA, Jet Propulsion Laboratory, National Aeronautics and Space Administration.
- Dehotin J, Braud I (2008) Which spatial discretization for distributed hydrological models? Proposition of a methodology and illustration for medium to largescale catchments.Hydrology and Earth System Sciences 12: 769-796.
- 9. Moriasi D (2007) Model evaluation guidelines for systematic quantification

of accuracy in watershed simulations. American Society of Agricultural and Biological Engineers 50: 885-900.

- Refsgaard J (1997) Parameterization, calibration and validation of distributed hydrological models. Journal of Hydrology 198: 69-97.
- Rozalis S, Morin E, Yair Y, Price C (2010) Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions. Journal of Hydrology 394: 245-255.
- 12. Bergstrom S (1991) Principles and confidence in hydrological modeling.Nordic Hydrology 22: 123-136.
- Brandt M (1990) Simulation of runoff and nitrate transport from mixed basins in Sweden.Nordic Hydrology 21: 13-34.
- Noto L, Ivanov V, Bras R, Vivoni E (2008) Effects of initialization on response of a fully distributed hydrological model. Journal of Hydrology 352: 107-125.
- Vrugt J (2009) Accelerating Markov Chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. International Journal of Nonlinear Sciences and Numerical Simulation 10: 273-290.
- 16. Boyle D, Gupta H, Sorooshian S (2000) Toward improved calibration of hydrological models: combining the strengths of manual and automatic methods.Water Resources Research 36: 3663-3674.
- Beven K (1989) Changing ideas in hydrology-The case of physically-based models.Journal of Hydrology 105: 157-172.
- Hoedjes J (2014) A Conceptual Flash Flood Early Warning System for Africa, Based on Terrestrial Microwave Links and Flash Flood Guidance.ISPRS International Journal of Geo-Information 3: 584-598.
- Bingeman A, Kouwen N, Soulis E (2006) Validation of the hydrological processes in a hydrological model Journal of Hydrologic Engineering 11: 451-463.
- Motovilov Y, Gottschalk L, Engeland K, Rodhe A (1999) Validation of a distributed hydrological model against spatial observations. Agricultural and Forest Meteorology 98–99: 257-277.
- Kouwen N (1993) Grouped response units for distributed hydrologic modelling. Journal of Water Resources Planning and Management 119: 289-305.
- 22. Xie Z (2007) Regional Parameter Estimation of the VIC Land Surface Model: Methodology and Application to River Basins in China. Journal of Hydrometeorology 8: 447-468.

Page 13 of 13

- 23. Marshall E, Randhir T (2008) Effect of climate change on watershed system: a regional analysis Climatic Change 89: 263-280.
- Zanon F (2010) Hydrological analysis of a flash flood across a climatic and geologic gradient: The September 18, 2007 event in Western Slovenia. Journal of Hydrology 394: 182-197.
- Hirpa F, Gebremichael M, Over T (2010) River fluctuation analysis: effect of watershed area.Water Resources Research 46: 12529.
- 26. Du B, Ji X, Harmel R, Hauck L (2009) Evaluation of a watershed model for estimating daily flow using limited flow measurements. Journal of the American Water Resources Association 45: 475-484.
- 27. Mesinger F (2005) North American Regional Reanalysis.Bulletin of the American Meteorological Society 343-360.
- Khan S (2011) Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria Basin: implications for hydrologic prediction in ungauged basins. IEEE Transactions on Geoscience and Remote Sensing 49: 85-95.

- USDA-NRCS (1994) State Soil Geographic Database (STATSGO) data users guide, Washington, DC: Data Use Information.National Soil Survey Center Misc. Publ. No. 1492.
- 30. Nash J (1957) The form of the instantaneous unit hydrograph.IAHS Publ 45: 114-121.
- Cosby B, Hornberger G, Clapp R, Ginn T (1984) A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. Water Resources Research 20: 682-690.
- Legates D, McCabe G (1999) Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation.Water Resources Research 35: 233-241.
- Servat E, Dezetter A (1991) Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese Savannah area. Hydrological Sciences Journal 36: 307-330.
- 34. Fernandez G, Chescheir G, Skaggs R, Amatya D (2005) Development and testing of watershed-scale models for poorly drained soils. Transactions of the American Society of Agricultural Engineers 48: 639-652.