Advancing the use of satellite rainfall datasets for flood prediction in ungauged basins: The role of scale, hydrologic process controls and the Global Precipitation Measurement Mission.

Faisal Hossain and Nitin Katiyar
Department of Civil and Environmental Engineering
Tennessee Technological University, Cookeville, USA

1.0 INTRODUCTION

Floods account for about 15 % of the total death toll related to natural disasters, wherein typically more than 10 million lives are either displaced or lost each year internationally (Hossain, 2006). Rainfall is the primary determinant of floods and its intimate interaction with the landform (i.e., topography, vegetation and channel network) magnified by highly wet antecedent conditions leads to catastrophic flooding in medium (i.e., 1000 ~ 5000 km^2) and large (i.e., > 5000 km^2) river basins. Furthermore, floods are more destructive over tropical river basins that lack adequate surface stations necessary for real-time rainfall monitoring – i.e., the ungauged river basins (Hossain and Katiyar, 2006a) (see Figure 1, left panel).

However, flood prediction is becoming ever more challenging in these medium-to-large river basins due to the systematic decline of in situ rainfall networks world-wide. The gradual erosion of these conventional rainfall data sources has lately been recognized as a major concern for advancing hydrologic monitoring, especially in basins that are ungauged or already sparsely instrumented (Stokstad, 1999; Shikhlomanov et al., 2002). As a collective response, the hydrologic community have recently established partnerships for the development of space-borne missions for cost-effective, yet global, hydrologic measurements. The most pertinent example in the context of flood prediction is the Global Precipitation Measurement (GPM) mission for global monitoring of rainfall (Smith et al., 2007). Hence, there is no doubt that the hydrologic community as a whole will gradually become dependent on GPM for a substantial part of its rainfall data needs for hydrologic research and operational monitoring.

GPM now beckons hydrologists as an opportunity to improve flood prediction capability in ungauged basins. However, before the potential of GPM can be realized, there are a number of hydrologic issues that must be addressed. Our success in leveraging the GPM to improve flood prediction will depend largely on the recognition of these issues and the feedback provided by hydrologists on the assessment of satellite rainfall data to the satellite data producing community (Hossain and Lettenmaier, 2006). The purpose of this chapter is to articulate these hydrologic issues that require further research and highlight the recent progress made in understanding them in the hope that satellite rainfall data can be used in hydrologic models more effectively.

2.0 OVERVIEW OF SATELLITE RAINFALL REMOTE SENSING AND GPM

The heritage of GPM originated two decades ago when Infrared (IR) radiometers on geostationary satellites were launched to provide high resolution measurement (Griffith et al., 1978). While geostationary IR sensors have substantial advantages in that they provide essentially continuous observations, a major limitation is that the quantity being sensed, cloud top temperature, is not directly related to precipitation (Huffman et al., 2001). Subsequently, space-borne passive microwave (PMW) radiometers evolved as a more dependable alternative (in terms of accuracy) a decade later. PMW sensors work on the principle that naturally emitted radiation in the microwave frequencies greater than 20 GHz is dictated by the composition of atmospheric hydrometeors. PMW sensors are considered more accurate under most conditions for precipitation estimation over land than their IR counterparts.

In 1997, the Tropical Rainfall Measuring Mission (TRMM), the first space-borne active microwave (AMW) precipitation radar (TRMM-PR), was launched. Although radar generally is the most accurate remote sensing technique for precipitation estimation, radar technology is costly, and TRMM-PR has limited spatial coverage (at latitudes between about 35^{o} S and 35^{o} N) with a sampling frequency about once per day. Therefore, the constellation of PMW sensors continue to represent a compromise between IR sensors and TRMM-PR in terms of sampling frequency, accuracy, and global coverage. GPM is therefore being planned now as a global constellation of low earth orbiting satellites (some of them existing) carrying various PMW sensors (Smith et al., 2007). It will essentially be an extension of the TRMM mission in space and time, which would provide near-global coverage of land areas, and would formally incorporate a means of combining precipitation radar with PMW sensors to maximize sampling and retrieval accuracy. The GPM Core satellite will be similar in concept to the TRMM satellite, and will house precipitation radar of improved accuracy as well as a PMW sensor (Figure 1, right panel). Through this configuration, GPM aims to provide coherent global precipitation products with temporal resolution ranging from 3 to 6 hours and spatial resolution in the range 25-100 km^2 (Smith et al., 2007; see also http://gpm.gsfc.nasa.gov).

A major benefit offered by the GPM program would be the increased availability of microwave rainfall data that will be cooperatively provided from multiple platforms by several independent programs at a high temporal resolution (~3 hours). It must however be noted by the hydrologist that, the microwave overpasses yield only instantaneous rainfall estimates rather than accumulated rainfall totals that are typically used as input in hydrologic models. Caution and thoughtful preprocessing are needed before investigating the usefulness of satellite rainfall data for flood prediction (discussed in detail later). Furthermore, hydrologists need to be cognizant of the current availability of a large number of 'combined' satellite algorithms that function on the basis of both geostationary IR and PMW data. It is currently not known what role IR-based algorithms, if any, will continue to play for flood prediction during the GPM-era as the frequency of the more accurate PMW data increases many folds. Promising newer algorithms that combine the IR data more intelligently and yet manage to retain the strength of PMW algorithms should be kept in the hydrologist's shortlist of potential input data sources over ungauged basins (Joyce et al., 2004).

3.0 CURRENT KNOWLEDGE GAPS ON SATELLITE-BASED FLOOD PREDICTION

3.1 THE PROCESS-BASED KNOWLEDGE GAP

Understanding the current knowledge gaps on satellite based flood prediction is critical to successful application of satellite rainfall data over regions lacking access to a conventional rainfall data source. The central theme on the current knowledge gap deals with the hydrologic implications of uncertainty of the satellite rainfall estimates. This satellite estimation uncertainty manifests as a result of the mismatches in the identification of rainy and non-rainy areas by the satellite algorithm, while considerable hydrologic implication exists for this uncertainty due to the spatial scaling properties of the river basin (Wood et al., 1990). The focus of this chapter is however, mostly on the former issue (i.e., satellite rainfall uncertainty) in an independent manner, even though we recognize that the basin scaling properties would play a definitive role in dictating the optimal use of satellite rainfall data for flood prediction.

Although there are several sources of uncertainty that complicate our understanding of flood prediction accuracy (see for example, Georgakakos et al., 2004), the principal source of uncertainty is, undoubtedly, rainfall (Kavetski et al., 2006a, 2006b; Hossain et al., 2004a, 2004b; Krzyzstofowicz, 1999, 2001). In a recent study, Syed et al. (2004) corroborated this further by demonstrating that 70%-80% of the variability observed in the terrestrial hydrologic cycle is, in fact, attributed to rainfall. For the case of satellite rainfall estimation, this uncertainty can lead to unacceptably large uncertainties in runoff simulation (Nijssen and Lettenmaier, 2004). Thus, if satellite rainfall data are to be critically assessed of the opportunities they possess for river flood prediction over large ungauged basins, it is important that we first understand the error propagation that is associated with satellite-estimated rainfall.

An error propagation of satellite rainfall estimates for flood prediction applications requires the derivation of the probability distribution of simulated stream flow involving the following three components: (1) a probabilistically formulated satellite rainfall model that can simulate realistic and equi-probable random traces of satellite-like rainfall estimates; (2) a deterministic or probabilistic hydrologic model for the rainfall-runoff transformation to floods; and (3) a Monte Carlo (MC) framework linking (1) and

(2). The fully random MC sampling can be currently considered the most preferred method for such uncertainty analysis due to ever-increasing computing power (Hossain et al., 2004c). Other reasons for the widespread preference of MC techniques are their lack of restrictive assumptions and completeness in sampling the error structure of the random variables (Beven and Freer, 2001; Kremer, 1983).

However, the traditional MC approach to modeling stream-flow error propagation exhibits limited physical insight of the role played by each hydrologic process control comprising the flood phenomenon. This is because the hydrologic model in a typical MC uncertainty assessment would be applied as a black-box unit for transforming the rainfall to runoff. While the more sophisticated physically-based and fully distributed hydrologic models are capable of simulating the individual water cycle components in the continuum of space and time, the problem of identifying the make up of streamflow as a function of various runoff components nevertheless persists. The derived runoff error distribution is thus consequently marginal, regardless of the type of hydrologic model use (conceptual-lumped or physically-based and distributed) because of its functional integration over the major hydrologic processes on the land surface. This distribution can not be isolated into components that can be linked directly to the individual hydrologic process controls or the nature of its physical representation (such as, infiltration, base flow, evaporation, etc.). To the best of our knowledge, there has not been any successful attempt to relate the marginal distribution of streamflow simulation error to these individual process controls. Existing literature provides little indication of a coherent agenda to explore the role of hydrologic process controls in the context of advancing satellite-based flood prediction over un-gauged river basins.

But how exactly does the study of flood prediction uncertainty as a function of hydrologic process controls and satellite rainfall estimation error help in serving the greater scientific agenda of protecting mankind from the flooding hazard? For any given river basin, flood prediction needs are unique (e.g., one may be interested in stream-flow prediction at the basin outlet or distributed simulation of water levels for the entire river network). A hydrologist is faced with a wide variety of geology, soils, initial wetness, vegetation, land use and topographic characteristics that affect the relationship between rainfall and runoff in the most unique ways. This relationship consequently affects the relationship between rainfall estimation error and runoff simulation error. While detailed information on the land surface may not always be available, especially for ungauged basins, approximate characteristics such as dominant overland flow mechanism (saturation excess vs. infiltration excess), extent of evapotranspiration (low vs. high vegetation), flow regime in channels (low-mild vs. steep channel slopes) are reasonable to be known a priori. Thus, if the role played by each hydrologic process control in transforming the rainfall estimation error to stream-flow error could be understood, then, ideally, one would be better poised to wisely select a hydrologic model with "commensurate" process representation that yields "acceptable" error propagation. These would allow the hydrologist to make an informed decision on his choice for a hydrologic model for flood mitigation purposes in an ungauged basin on the basis of the quality of satellite rainfall data available to him.

3.2 THE SCALE-BASED KNOWLEDGE GAP

Another knowledge gap that we must recognize in the context of flood prediction is the complexity of the error structure of satellite rainfall data. Unlike radar rainfall estimation, where after careful quality control and error adjustments, the residual error is associated primarily with a random component that usually has modest space-time correlation, satellite precipitation retrieval uncertainty is associated with correlated rain/no-rain detection and false alarm error characteristics as well as systematic and random rain rate error components with non-negligible spatio-temporal correlation lengths. Hossain and Anagnostou (2006a) have recently demonstrated the complex nature of this satellite error structure using a ground validation site over the Oklahoma Mesonet region. Furthermore, different satellite rainfall algorithms would have different error characteristics, while the combined multi-sensor algorithms may be expected to have a more complex error structure depending on the type of calibration data used in the making. Nonetheless, most attempts to characterize errors in satellite precipitation retrievals to date portray the error structure using metrics that can be argued as overly simplistic, and ultimately misleading relative to the hydrologic potential of GPM. For instance, error metrics limited to 'bias' and 'random error' parameters have been used to define the minimum success criteria of GPM and other community efforts like the Pilot Evaluation of High Resolution Precipitation Products (PEHRP, Turk et al., 2006). For flood prediction, these metrics are probably not adequate, even though they may serve a very useful purpose at meteorological scales. The desire for progression to finer (spatial) scales in satellite precipitation estimation is in fact counter-balanced by increasing dimensionality of the retrieval error, which has a consequently complex effect on the propagation through land surface-atmosphere interaction simulations. This in turn has tremendous implications for the spatial and temporal scales at which hydrologic models can reasonably be implemented, or rather, the scale at which optimal data use is feasible.

As an example, consider the dynamic process of vertical soil moisture transport. The water flux in soil is governed by the cumulative effect of infiltration, runoff, gradient diffusion, gravity, and soil water extraction through roots for canopy transpiration. All these processes exhibit dynamic variability in the ranges of minutes to hours over scales of cm^2 to km^2. However, satellite precipitation algorithms of today cannot hope to resolve these resolutions. It is even doubtful if the future space-borne precipitation remote sensing can independently deliver the rainfall data at the resolution where surface hydrology is dominant, which is at considerably smaller space-time scales than the typically coarser meteorologic scale at which satellite data is produced. As a minimum, there is a need to understand the spatial resolution to which satellite products can realistically be disaggregated (see Margulis and Entekhabi, 2001 and Venugopal et al., 1999 for example) and to estimate the resulting error structure, and its interaction with hydrologic models which produce flood forecasts. The scale incongruity between meteorological process data and its hydrologic application represents a competing trade off between lowering the satellite retrieval error versus modeling land-vegetation-atmosphere processes at the finest scale possible. While much remains to be done to define these trade-offs towards optimal use of satellite rainfall data in hydrologic models, it may well not be possible to implement GPM products at scales as fine as those cited above (e.g., 5 km), and that the 25-100 km resolutions suggested by Smith et al (2004) may perhaps be more realistic and reliable for the hydrologist for flood prediction.

A major problem encountered in application of satellite rainfall data is that the frequency of complex mis-matches (with the ground-truth) increases as the satellite rainfall data is progressively reduced in scale (as alluded earlier in section 3.1). We can demonstrate this phenomenon through an example on the detection performance of two types of satellite rainfall types. Figure 2 demonstrates typical detection capabilities for rain and no-rain for two different sensors (PMW – left panel; IR-right panel) using the most accurate space-borne rainfall data derived from the TRMM Precipitation Radar (PR) as 'ground-truth' (data product name 2A25). The presence of definitive spatial structures of detection of rain and no-rain as a function of sensor-type is clearly evident. This detection capability is also known to be strongly influenced by scale and rainfall rates (Hossain and Anagnostou, 2006).

The success in resolving the scale incongruity to a level practically feasible for flood prediction will therefore rest on the feedback between hydrologists and meteorologists (the typical algorithm and data producers). Even though the efforts at addressing hydrologic prediction uncertainty (Beven and Binley, 1992) are probably as mature as the efforts to characterize uncertainty of remote sensing of rainfall (North and Nakamoto, 1989), both efforts have evolved independently. This lack of feedback can be attributed to the absence of proper metrics and frameworks that are interpretable by both end-user hydrologists and producer meteorologists. Two satellite rainfall algorithms with similar bias and root mean squared error (RMSE) can have much different error propagation properties when used in hydrologic models (Lee et al., 2004). Thus hydrologists today are therefore left with inadequate metrics to identify optimal data use and thereby communicate to the data producers on the desired minimum criteria for a satellite mission to be effective in flood prediction at pertinent scales and geographic location.

4.0 MODELING SATELLITE RAINFALL ERROR COMPLEXITY

Current satellite error models target mostly PMW sensor retrievals focusing primarily on the sampling uncertainty due to the low frequency of satellite overpasses (for a detailed review see Astin, 1997; Steiner et al. 1996; Bell, 1987; Bell et al., 1990; Steiner et al., 2003; Gebremichael and Krajewski, 2004). Recently, Hossain and Anagnostou (2004a,b) have provided evidence that a detailed decomposition of the satellite rainfall error structure with explicit formalization of the uncertainty in rainy/non-rainy area delineation can contribute to improving our understanding of the implications of satellite estimation error on land surface simulation parameters for fine-scale hydrological processes (such as floods and soil moisture dynamics). There are also other notable models formulated recently to characterize the error structure of satellite rainfall data that may be of interest to the hydrologist to advance satellite based flood prediction (Bellerby and Sun, 2005; Teo, 2006).

4.1 A TWO DIMENSIONAL SATELLITE RAINFALL ERROR MODEL (SREM2D)

Motivated by the current state of the art in error modeling and the challenges faced by the need for high-resolution satellite rainfall data in hydrology, a mathematical formalization of a space-time error model, named *SREM2D*, was recently developed by Hossain and Anagnostou (2006a). *SREM2D* had the following design objectives in mind during its conceptualization: (1) It should function as a filter wherein the hydrological implications of fine-scale components of the satellite precipitation error structure can be explicitly determined by coupling it with a hydrological/land surface model; Thus, *SREM2D*-based experiments should provide the much needed focus to meteorologists for the development of next-generation of satellite rainfall products with enhanced societal applications; These experiments should also help hydrologist identify the optimality criterion for using a given satellite rainfall dataset in a hydrologic model; (2) It should be modular in design with the capability to allow uncertainty assessment of *any* satellite rainfall algorithm; (3) It should be conceptualized in an algorithmic fashion so that it is easy to code numerically by a user wishing to make use of the model for his/her own scientific agenda. *SREM2D* uses as input “reference” rain fields of higher accuracy and resolution representing the “true” surface rainfall process, and stochastic space-time formulations to characterize the multi-dimensional error structure of satellite retrieval. The algorithmic approach of *SREM2D* is aimed at generating realistic ensembles of satellite rain fields from the most definitive “reference” rain fields that would preserve the estimation error characteristics at various scales of aggregation. By propagating the simulated ensembles in a hydrologic model, *SREM2D* would therefore allow the understanding of the implications of satellite rainfall error structure and scale complexity on streamflow simulation.

The major dimensions of error structure in satellite estimation modeled by *SREM2D* are (1) the joint probability of successful delineation of rainy and non-rainy areas accounting for a spatial structure; (2) the temporal dynamics of the conditional rainfall estimation bias (rain > 0 unit); and (3) the spatial structure of the conditional (rain > 0 unit) random deviation. The spatial structure in *SREM2D* is modeled as spatially correlated Gaussian random fields while the temporal pattern of the systematic deviation is modeled using a lag-one autoregressive process. The spatial structures for rain and no-rain joint detection probabilities are modeled using Bernoulli trials of the uniform distribution with a correlated structure. This correlation structure is generated from

Gaussian random fields transformed to the uniform distribution random variables via an error function transformation. There are nine error parameters in total. Complete details on *SREM2D* can be found in Hossain and Anagnostou (2006a).

5.0 CURRENT PROGRESS ON CLOSING THE KNOWLEDGE GAP

5.1 ON SCALE BASED KNOWLEDGE GAP

Comparison of *SREM2D*-simulated satellite rainfall with actual satellite rainfall data produced by NASA (IR-3B41RT; Huffman et al., 2003) has shown that a complex and multi-dimensional error modeling technique (such as *SREM2D*) can preserve the estimation error characteristics across scales with marginal deviations. Upon comparison with less complex and commonly applied error modeling strategies, it was also shown that these (simpler) approaches typically underestimated sensor retrieval error standard deviation by more than 100% upon aggregation, which, for *SREM2D*, was found to be below 30% (Hossain and Anagnostou, 2006a). More recent studies have further demonstrated that understanding of the hydrologic implications of satellite-rainfall data overland can be significantly improved through the use of *SREM2D* in a hydrologic error propagation framework. This is a promising finding as it would allow a more reliable investigation of the optimality criterion for using satellite rainfall data in hydrologic models. Two aspects were examined in detail: (1) soil moisture dynamics and (2) ensemble rainfall data generation. For understanding the impact of satellite rainfall uncertainty on soil moisture dynamics, the Common Land Model (CLM; Dai et al., 2003) was coupled with *SREM2D* to propagate ensembles of simulated satellite rain fields for the prediction of soil moisture at 5 cm depth region. It was observed that *SREM2D* captured the spatiotemporal characteristics of soil moisture uncertainty with higher consistency than a simpler bi-dimensional error modeling strategy (Figure 3, upper panels; Hossain and Anagnostou, 2005b). In a subsequent follow-up study, further insights were revealed from the pursuit of the scientific query: *Can a multidimensional satellite rainfall error model perform realistic ensemble generation of satellite rainfall data of improved accuracy for a satellite retrieval technique?* Using as reference, ground radar (WSR-88D) rainfall fields, the scale-dependent multidimensional error structure for satellite rainfall algorithms was determined. Next, by reversing the definition of reference and corrupted rain fields produced by *SREM2D,* the inverse multidimensional error structure of WSR-88D rainfall fields with respect to the satellite rainfall data was identified. *SREM2D* was then run in the inverse mode to generate reference-like realizations of rainfall. The accuracy of inverse-SREM2D rainfall ensemble was observed to be consistently higher than the simpler inverse error-modeling scheme for the IR-3B41RT product (Figure 3, lower panels).

Because most attempts to characterize errors in satellite precipitation retrievals to date portray the error structure of satellite rainfall estimates using metrics that are overly simplistic, and ultimately misleading relative to the true hydrologic potential of satellite rainfall data, a complex and multi-dimensional error modeling strategy that is compatible with the dynamic nature of land surface hydrologic processes is needed to advance optimal data use in hydrologic models.

5.2 ON PROCESS BASED KNOWLEDGE GAP

5.2.1 Use of a Modular Hydrologic Modeling Platform

In order to understand the implication of satellite rainfall error on hydrologic processes, we recently developed an open-book watershed model (Figure 4). The design was modular wherein specific hydrologic processes could be conveniently altered or added to make a process-based understanding of satellite rainfall error propagation (as discussed in Section 3.1). A square-grid volume domain was used where the individual processes of overland flow and infiltration to the subsurface were linked to simulate the response of the unsaturated zone to rainfall (*Figure* 1). In the open-book model, the generated surface and subsurface runoff were calculated at each time-step from knowledge of the time-varying infiltration (or recharge to the soil) and by keeping track of the soil water storage for each grid volume at every timestep. The overland flow was then routed along the direction of steepest gradient for each grid surface until it laterally drains into the main channel. The streamflow was modeled as a 1-D kinematic flow using Manning's equation. Evapo-transpiration and 2-D saturated zone flow were assumed insignificant in our conceptualizations as our goal was to focus primarily on streamflow simulation at the timescales of flooding where the surface and the unsaturated zones are considered hydrologically the most dynamic regions. A point to note herein is that the 'Depth to bedrock' shown in Figure 4 is basically the depth of the effective soil column. Complete details on the open-book watershed model can be found in Katiyar and Hossain (2006).

5.2.2 The Hydrologic Process Conceptualizations

To understand the role played by specific hydrologic process control conceptualization, three types of rainfall-runoff conceptualizations were considered for computing excess rainfall over a grid volume (see Figure 5). These conceptualizations were: (1) A simple statistical parameterization to compute excess rainfall; (2) A linear storage-discharge conceptualization for surface and subsurface runoff generation; (3) A non-linear storage-discharge conceptualization for surface and subsurface runoff generation. The overland and river flow components of the model remained the same. These process conceptualizations employed basically a mass-balance approach and are presented briefly as follows.

For the statistical model, the precipitation *p(t)* was partitioned into infiltration to the soil water store as $\boldsymbol{a}p(t)$, and surface runoff (quickflow/overland flow) as $(1-\boldsymbol{a})p(t)$. The subsurface flow draining from the grid volume's soil water store is assumed insignificant (at flooding timescales) and the soil water storage is updated at each time-step on the basis of the recharge only. When it equaled the maximum storage capacity of S_b (computed as $D\varphi$; D= depth of effective soil column and φ is porosity), all precipitation was consequently transformed as surface runoff with no recharge.

For the linear storage-discharge conceptualization, the following water balance equation was used for each grid volume,

$$\frac{ds(t)}{dt} = p(t) - q_{se}(t) - q_{ss}(t) \qquad (1)$$

where, $s(t)$ is the soil water storage, $p(t)$ is the precipitation, $q_{se}(t)$ is the overland saturation-excess flow and $q_{ss}(t)$ is the sub-surface flow at time t. The $q_{ss}(t)$ and $q_{se}(t)$ were computed as follows,

$$q_{ss} = \frac{s(t) - S_f}{t_c} \qquad \text{if } s(t) > S_f \qquad (2a)$$

$$q_{ss} = 0 \qquad \text{if } s(t) < S_f \qquad (2b)$$

where, S_f is the soil moisture storage at field capacity (defined by the soil type) and t_c is the grid response time to subsurface flow. t_c is approximated from Darcy's law assuming a triangular groundwater aquifer and hydraulic gradient approximated by ground slope.

$$t_c = \frac{L\phi}{2K_s \tan\beta} \qquad \text{(2c)}$$

Herein, L is the grid size, K_s the saturated hydraulic conductivity and β is the grid slope. The sub-surface flow draining out from each grid volume is not routed within the soil medium for the same reason that it would comprise an insignificant component of the total flood volume. However, this model conceptualization represented a complexity level higher than the previous statistical parameterization because of the use of mass balance equation and physically-based watershed parameters to identify the saturation excess runoff. The overland saturation excess flow q_{se}(t) was computed as follows,

$$q_{se} = \frac{s(t) - S_b}{\Delta t} \qquad \text{if } s(t) > S_b \qquad \text{(3a)}$$

$$q_{se} = 0 \qquad \text{if } s(t) < S_b \qquad \text{(3b)}$$

where S_b is the soil's storage capacity computed as Dφ (D= effective soil column depth and φ is porosity).

Finally, for the non-linear storage-discharge conceptualization, the subsurface runoff in the linear model was reformulated as follows,

$$q_{ss} = \left[\frac{s - s_f}{a}\right]^{1/b} \qquad \text{if } s > s_f \qquad \text{(4a)}$$

$$q_{ss} = 0 \qquad \text{if } s < s_f \qquad \text{(4b)}$$

Here, parameters 'b' defines the degree of non-linearity in the storage-discharge relationship, while 'a' replaces t_c in Equation 2a. Figure 6 summarizes all three process conceptualizations to showcase the gradual increase in model complexity.

5.2.3 The Error Propagation Framework

SREM2D was coupled to the open-book models to understand the response of satellite rainfall error to spatial scaling on river flow prediction uncertainty. We used, as our reference, quality-controlled ground radar (WSR-88D) rainfall data over the Oklahoma Mesonet region. Satellite rainfall error parameters were derived for satellite rainfall algorithm (3B41RT) that is produced by NASA (Huffman et al., 2003) on a real-time basis. *SREM2D* was then used to corrupt radar rainfall fields in a space-time framework to simulate satellite-like rain estimates. The satellite rainfall error propagation in streamflow prediction was assessed in a MC framework for the three model types across two spatial scales of aggregations - 0.25 degree and 0.50 degree. The 15 MC realizations of *SREM2D* generated rainfields that were propagated through each open-book model conceptualization yielded corresponding uncertainty limits in streamflow simulation.

Two contrasting issues were considered in the error propagation. If either the uncertainty limits were predicted too narrowly or the whole ensemble envelope is biased (i.e., the reference streamflow is consistently outside the prediction error bounds), then a comparison with in-situ/reference measurements would suggest that the combined rainfall-model complexity structure was invalid for the satellite rainfall error. If, on the other hand they were predicted too widely, then it could be concluded that the hydrologic modeling structure had little predictive capability. The dichotomous nature of '*structural validity*' and '*predictive ability*' was quantified by the *Exceedance Probability (EP* Equation 5*)* and *Uncertainty Ratio (UR* Equation 6*),* respectively, as follows:

$$EP = \frac{\text{Number of times reference streamflow exceeds the uncertainty limits}}{\text{Total number of timesteps}} \qquad (7)$$

$$UR = \frac{\text{Uncertainty in runoff volume simulation (beween uncertainty limits)}}{\text{Observed Runoff Volume}} \quad (8)$$

Table 1 summarizes the findings of the error propagation experiment as a function of scale and hydrologic process conceptualizations. The global picture that emerges from this table can summarized as follows:

(1) statistical parameterization for excess rainfall results in increased sensitivity of satellite rainfall error to streamflow prediction uncertainty; this sensitivity, however, responds favorably to scaling towards improving the model's structural validity at larger scales of aggregation;

(2) inclusion of a linear/non-linear reservoir for subsurface flow representation visibly smoothens the hydrologic simulations and reduces the runoff simulation uncertainty;

(3) insignificant beneficial impact is observed through the inclusion of non-linearity in the storage-discharge relationship and it may so be that the scale of application is already so large that satellite rainfall error is insensitive to any further increase in process complexity;

(4) there is strong indication that hydrologic process complexity plays a definitive role in accurately capturing streamflow variability on the basis of model driven by downscaled satellite input data .

While the findings represented a useful first step, our study has limitations like any other investigation. Hence, it is important to extend the investigation involving a wider range of research objectives and more complex hydrologic process representation. Examples of extension could be: (1) repeat the investigation using real-world watersheds in a range of climatic conditions, distributed hydrologic models and elevation data and thereby understand the utility of the open-book approach as a physically consistent proxy for investigating optimality criteria; (2) increase the complexity of the hydrologic processes through more physically-based process equations (i.e., Richards equation or Green and Ampt for infiltration; energy-balance method for representation of evapo-transpiration etc.); and finally, (3) explore scaling behavior at finer space-time resolutions (< 0.25 degree and < 1 hourly). Currently the Land Information System of NASA (LIS) provides the hydrologic state of the land at 0.5 hour 1 X 1 km^2 resolution using satellite datasets that are subsequently ingested for numerous societal applications such as weather prediction, agricultural planning, army operations etc (Kumar et al., 2006). We hope that extension of our work along these directions can consequently help us achieve a firmer understanding of the optimality criteria for use of remotely-sensed rainfall data from space-borne platforms in hydrologic models.

6.0 CONCLUDING REMARKS

For advancing the use of satellite rainfall data for flood prediction, there are basically two major issues related to rainfall data uncertainty that hydrologists need to recognize in the context of flood prediction– i) the role played by hydrologic process controls; and ii) the role played by scale. We have highlighted the progress made by us on the understanding of these two issues. Much work remains to be done towards a more complete understanding on optimal use of satellite rainfall data in hydrologic models. It is however, equally important to initiate the work in anticipation of a successful leveraging of GPM. We have argued in this chapter, as has bee argued previously by others, that unless there is a shift in paradigm, the conventional assessment frameworks and metrics for estimation of rainfall from satellite sensors will probably remain inadequate for hydrologic purposes such as flood prediction. We also argued that greater emphasis must be placed on development of hydrologically relevant precipitation estimation algorithms, and that this will require involvement of a broader cross-section of the hydrologic community. We therefore hope that identification of these key issues, as discussed in this chapter, will usher a new era for hydrologists working on optimal use of satellite rainfall data in anticipation of GPM.

We would like to close this chapter with a candid discussion of the limitations and disclaimers associated with our study that readers should be aware of. For example, while we have predominantly focused on floods, the choice of appropriate error metrics would most probably be dictated by the flood type (high/extreme floods versus low/frequent floods). Furthermore, the hydrologic implication of satellite rainfall error would be strongly influenced by the hydrologic variable (or predictand) in question. Again, this chapter focused on floods, while soil moisture, which plays a critical role in partitioning of rainfall into runoff, would not have the same implication as streamflow. The reader is referred to the work of Hossain and Anagnostou (2005a, 2005b) where a detailed investigation has been carried out for soil moisture.

7.0 REFERENCES

Anagnostou, E.N. (2004). Overview of Overland Satellite Rainfall Estimation for Hydro-Meteorological Applications. *Surveys in Geophysics,* 25(5-6), pp. 511-537.

Astin, I. 1997. A survey of studies into errors in large scale space-time averages of rainfall, cloud cover, sea surface processes and the earth's radiation budget as derived from low orbit satellite instruments because of their incomplete temporal and spatial coverage. *Surveys in Geophysics* 18: 385–403.

Bell, T.L. 1987. A space-time stochastic model of rainfall for satellite remote-sensing studies. *Journal of Geophysical Research* 92D:9631–9643.

Bell, T.L., Abdullah, A., Martin, R.L., and North, G.R. 1990. Sampling errors for satellite-derived tropical rainfall—Monte Carlo study using a space-time stochastic model. *Journal of Geophysical Research* 95D: 2195–2205.

Bellerby, T., and Sun, J. 2005. Probabilistic and ensemble representations of the uncertainty in IR/Microwave precipitation product. *Journal of Hydrometeorology*, 6: 1032-1044.

Beven, K.J., and Freer, J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249:11–29.

Beven, K.J., Binley, AM. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* 6:279–298.

Gebremichael, M., and Krajewski, W.F. 2004. Characterization of the temporal sampling error in space-time-averaged rainfall estimates from satellites. *Journal of Geophysical Research* 109(D11), (doi: D11110 0.1029/2004JD004509).

Georgakakos, K.P., Seo, D-J., Gupta, H., Schaake, J.C., and Butts, M. 2004. Towards the characterization of stream-flow simulation uncertainty through multimodel ensembles. *Journal of Hydrology* 298:223–241.

Griffith, C.G., Woodley, W.L., and Grube, P.G. 1978. Rain estimation from geosynchronous satellite imagery-visible and infrared studies. *Monthly Weather Review* 106:1153–1171.

Hossain, F. 2006. Towards formulation of a fully space-borne system for early warning of floods: Can cost-effectiveness outweigh flood prediction uncertainty? *Natural Hazards* 37(3):263–276 (doi :10.1007/s11069-005-4645-0).

Hossain, F., and Anagnostou, E.N. 2006a. A two-dimensional satellite rainfall error model. *IEEE Transactions Geosciences and Remote Sensing* 44 (6): 1511-1522 (doi: 10.1109/TGRS.2005.863866).

Hossain, F., and Anagnostou, E.N. 2006b. Assessment of a multi-dimensional satellite rainfall error model for ensemble generation of satellite rainfall data. *IEEE Geosciences and Remote Sensing Letters,* 3(3): 419-423.

Hossain, F., and D.P. Lettenmaier. 2006. Flood Prediction in the Future: Recognizing Hydrologic Issues in Anticipation of the Global Precipitation Measurement Mission. *Water Resources Research* Opinion Paper. (In press; available online: http://iweb.tntech.edu/fhossain/papers/WRRHossainLettenmaier.pdf).

Hossain, F., and Katiyar, N. 2006a. Improving flood forecasting in international river basins. *EOS Transactions (AGU)* 87(5):49–50.

Hossain, F. and Katiyar, N. 2006b. Sensitivity Analyses of Satellite Rainfall Estimation Error to Open-Book Hydrologic Models of Varying Levels of Conceptualisation and Spatial Aggregation, *Hydrological Sciences Journal.* (Revised and in review).

Hossain, F., and Anagnostou, E.N. 2005a. Numerical investigation of the impact of uncertainties in satellite rainfall estimation and land surface parameters for simulation of soil moisture. *Advances in Water Resources* 28(12):1336–1350 (doi:10.1016/j.advwatres.2005.03.013).

Hossain, F. and Anagnostou, E.N. 2005b. Using a multi-dimensional satellite rainfall error model to characterize uncertainty in soil moisture fields simulated by an offline land surface model *Geophysical Research Letters* vol. 32(L15402) (doi:10.1029/2005GL023122).

Hossain, F., Anagnostou, E.N., and Dinku, T. 2004a. Sensitivity analyses of satellite rainfall retrieval and sampling error on flood prediction uncertainty. *IEEE Transactions on Geosciences and Remote Sensing* 42(1).

Hossain, F., Anagnostou, E.N., Dinku, T., and Borga, M. 2004b. Hydrological model sensitivity to parameter and radar rainfall estimation uncertainty. *Hydrological Processes* 18 (doi:10.1002/hyp.5659).

Hossain, F., Anagnostou, E.N., and Lee Khil-Ha. 2004c. A non-linear and stochastic response surface method for Bayesian estimation of uncertainty in soil moisture simulation from a land surface model. *INVITED PAPER—Special Issue in Non-Linear Processes in Geophysics* 11:1–15 (SREF-ID: 1607-7946/npg/2004-11-1).

Huffman, G.J., Adler, R.F. Morrissey, M.M. and others. 2001. Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology,* 2, pp 36-50.

Huffman, G.J., R.F. Adler, E.F. Stocker, D.T. Bolvin, and E.J. Nelkin. 2003. Analysis of TRMM 3-hourly multi-satellite precipitation estimates computed in both real and post-real time. *12th Conference on Satellite Meteorology and Oceanography*, Long Beach, California, Feb. 9-13, 2003.

Joyce, R.J. and Janowiak, J.E., Arkin, P.E., and Xie, P. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3): 487-503.

Katiyar, N., and Hossain, F. 2006. An open-book watershed model for prototyping space-borne flood monitoring systems in international river basins. *Environmental Modeling and Software* (In review: available online http://iweb.tntech.edu/fhossain/papers/EnvSoft_GPM.pdf).

Kavetski, D., Kuczera, G., and Franks, S.W. 2006a. Bayesian analysis of input uncertainty in hydrological modeling: 2 Application. *Water Resources Research,* 42(W03408) (doi:10.1029/2005WR004376).

Kavetski, D., Kuczera, G., and Franks, S.W. 2006b. Bayesian analysis of input uncertainty in hydrological modeling: 1 Theory. *Water Resources Research* 42 (W03407) (doi:10.1029/2005WR004368).

Kremer, J.N. 1983. Ecological implications of parameter uncertainty in stochastic simulation. *Ecological Modeling* 18:187–207.

Krzyzstofowicz, R. 1999. Bayesian theory of probabilistic forecasting via deterministic hydrological model. *Water Resources Research* 35(9):2739–2750.

Krzyzstofowicz, R. 2001. The case for probabilistic forecasting in hydrology. Journal of Hydrology 249: 2–9.

Kumar, S. V., C.D. Peters-Lidard, Y. Tian, J. Geiger, P.R. Houser, S. Olden, L. Lighty, J.L. Eastman, P. Dirmeyer, P. Doty, P., J. Adams, E.F. Wood, and J. Sheffield. 2006. LIS - An Interoperable Framework for High Resolution Land Surface Modeling.

Environmental Modeling and Software (In press; copy available at http://lis.gsfc.nasa.gov) (doi:10.1016/j.envsoft.2005.07.004) .

Lee, K.H., and Anagnostou, E.N. (2004). Investigation of the nonlinear hydrologic response to precipitation forcing in physically based land surface modeling. *Canadian Journal of Remote Sensing* 30(5), 706-716.

Margulis, S., and Entekhabi, D. (2001). Temporal disaggregation of satellite-derived monthly precipitation estimates and resulting propagation of error in partitioning of water at the land surface. *Hydrology and Earth System Sciences,* 5(1): 27-38.

Nijssen, B., and Lettenmaier, D.P. 2004. Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites. *Journal of Geophysical Research* 109(D02103).

North, G. R., and S. Nakamoto, 1989: Formalism for comparing rain estimation designs. *Journal of Atmospheric and Oceanic Technology,* 6:985–992.

Shiklomanov, A.I., Lammers, R.B., and Vörösmarty, C.J. 2002. Widespread decline in hydrological monitoring threatens pan-arctic research. *EOS Transactions* 83(2):16–17.

Smith E., Asrar, G., Furuhama, Y., Ginati, A., Kummerow, C., Levizzani, V., Mugnai, A., Nakamura, K., Adler, R. Casse, V., Cleave, M., Debois, M., Durning, J., Entin, J., Houser, P., Iguchi, T., Kakar, R., Kaye, J., Kojima, M., Lettenmaier, D.P., Luther, M., Mehta, A., Morel, P., Nakazawa, T., Neeck, S., Okamoto, K., Oki, R., Raju, G., Shepherd, M., Stocker, E., Testud, J., and Wood, E.F. 2004. The international global precipitation measurement (GPM) program and mission: An overview. In *Measuring Precipitation from Space: EURAINSAT and the Future,* (Eds.) V. Levizzani and F.J.

Turk, Kluwer Academic Publishers (In press; copy available at http://gpm.gsfc.nasa.gov).

Steiner, M. 1996. Uncertainty of estimates of monthly areal rainfall for temporally sparse remote observations. *Water Resources Research,* 32:373–388.

Steiner, M., Bell, T.L., Zhang, Y., and Wood, E.F. 2003. Comparison of two methods for estimating the sampling-related uncertainty of satellite rainfall averages based on a large radar dataset. *Journal of Climate,* 16:3759–3778.

Stokstad, E. 1999. Scarcity of rain, stream gages threatens forecasts. *Science* 285:1199.

Syed, T.H., Lakshmi, V., Paleologos, E., Lohmann, D., Mitchell, K., and Famiglietti, J. 2004. Analysis of process controls in land surface hydrological cycle over the continental United States. *Journal of Geophysical Research* 109(D22105), doi: 10.1029/2004JD004640.

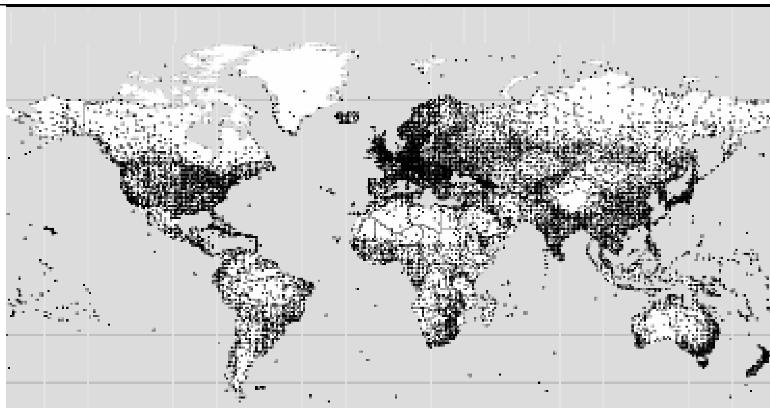
Teo, C.K. 2006. Satellite-based rainfall estimation and its application to crop yield modeling. *PhD Thesis (unpublished)*, Department of Meteorology, University of Reading.

Turk, J., Bauer, P., and Ebert, E.E., and Arkin, P.A. 2005. Satellite-Derived Precipitation Verification Activities within the International Precipitation Working Group (IPWG), *14th Conference on Satellite Meteorology and Oceanography*, P2.15.

Venugopal, V., Foufoula-Georgio, E., and Sapozhinikhov, V. 1999. A space-time downscaling model for rainfall. *Journal of Geophysical Research- Atmospheres,* 104(D16), pp. 19705-19721.

Table 3. *UR* and *EP* values as a function of spatial scale and process conceptualization.

0.25 degree		**0.50 degree**	
UR	**EP**	**UR**	**EP**
Statistical Model			
0.588	0.607	0.670	0.3064
% Change upon aggregation to 0.5 degree		+14.0 %	- 40.0%
Linear Model			
0.492	0.514	0.594	0.450
% Change upon aggregation to 0.5 degree		+20%	-12.5%
Non-Linear Model			
0.561	0.557	0.668	0.476
% Change upon aggregation to 0.5 degree		+19%	-14.5%



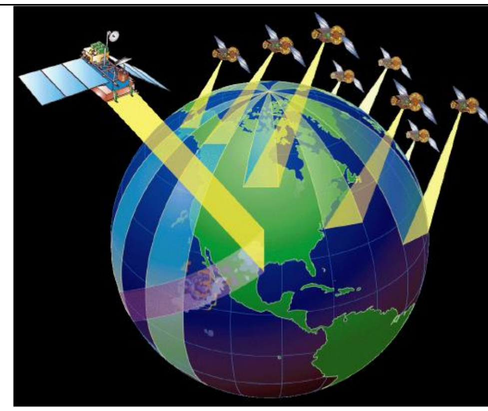


Figure 1. Left panel – global distribution of in-situ rainfall gages showing the sparse and unevenness in the underdeveloped world (source: http://www.cpc.noaa.gov). Right Panel - Constellation of GPM satellites. The larger satellite on the left represents the core with a radar on board, while the rest carry polar orbiting PMW sensors (source: http://gpm.gsfc.nasa.gov).

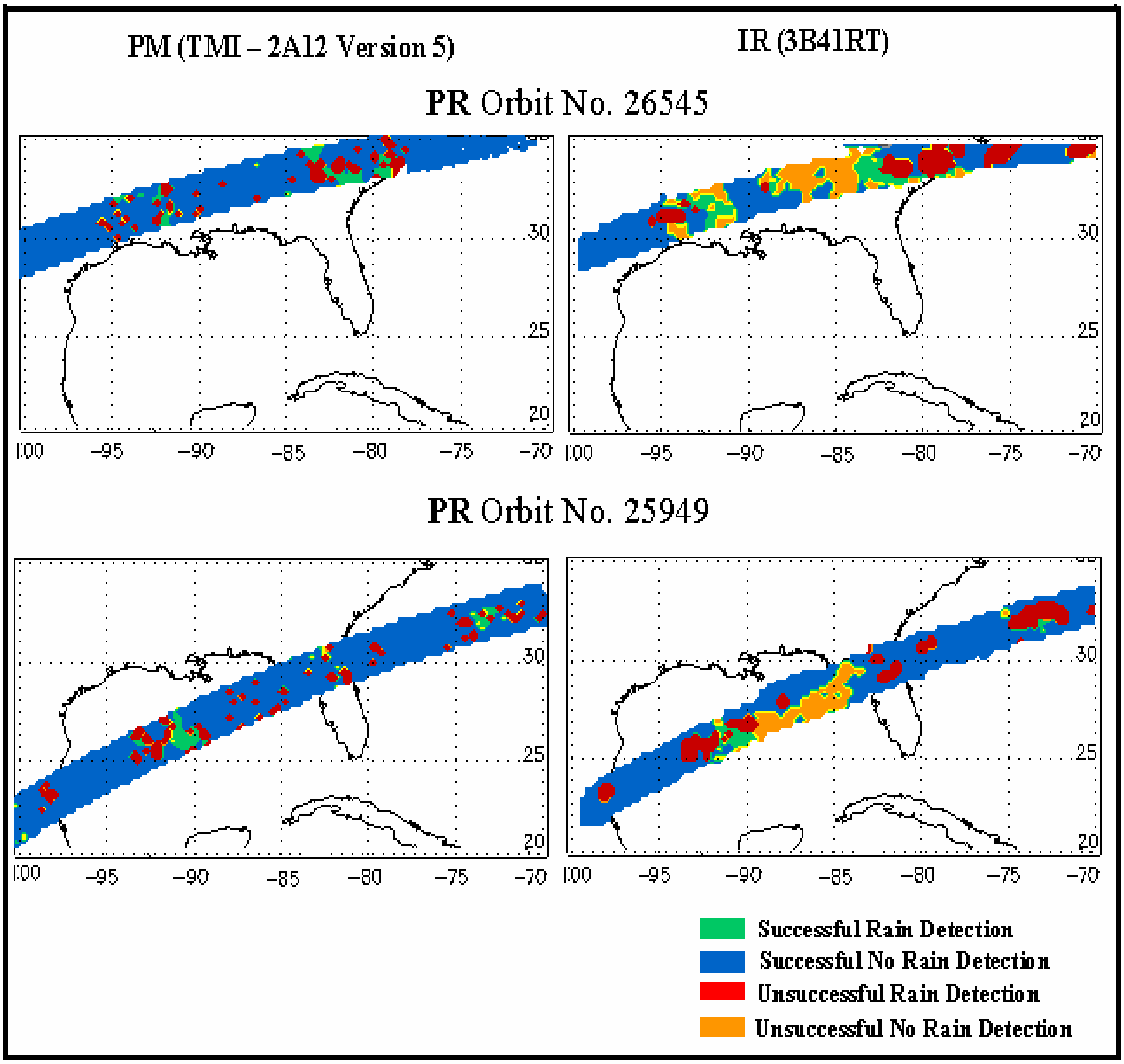


Figure 2. Successful and unsuccessful rain and no-rain detection by MW and IR sensors referenced with TRMM-PR observations. [Taken from Hossain and Anagnostou – 2006a; Reprinted with kind permission from Institute of Electrical and Electronics Engineering Transactions in Geosciences and Remote Sensing].

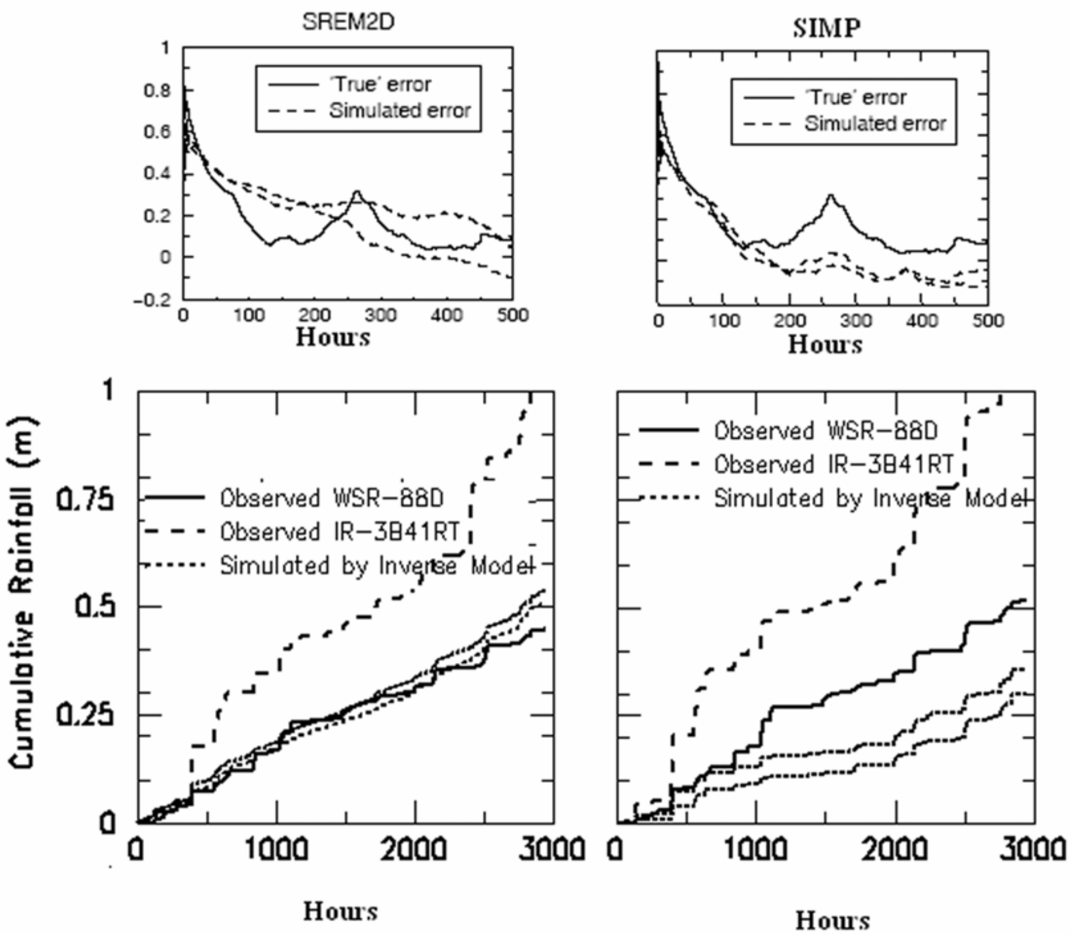


Figure 3. Hydrologic implications of using a multidimensional satellite rainfall error modeling strategy such as *SREM2D*.

Upper panels: Temporal correlogram of error field for near surface (5 cm) soil moisture simulated by CLM driven by simulated satellite rainfall data based on two error modeling schemes—*SREM2D* (left panel) and SIMP (right panel). The solid line represents the true soil moisture error dynamics on the basis of actual IR-3B41RT data. The dashed line represents the range of soil moisture error dynamics based on simulation by error model. SIMP represents the commonly used error modeling strategy in literature based on simple error statistics (From Hossain and Anagnostou, 2005b; *Reprinted with kind permission from American Geophysical Union*)

Lower panels: Ensemble envelopes (i.e. uncertainty range) of satellite-retrieved cumulative hyetographs (dotted lines) for two error-modeling schemes—*SREM2D* (left panel) and SIMP (right panel). Solid line represents "true" rainfall cumulative hyetograph from WSR-88D estimates, while the dashed line is the rainfall cumulative hyetograph from actual IR-3B41RT data. (From Hossain and Anagnostou, 2006b; *Reprinted with kind permission from the Institution of Electrical and Electronic Engineers*)

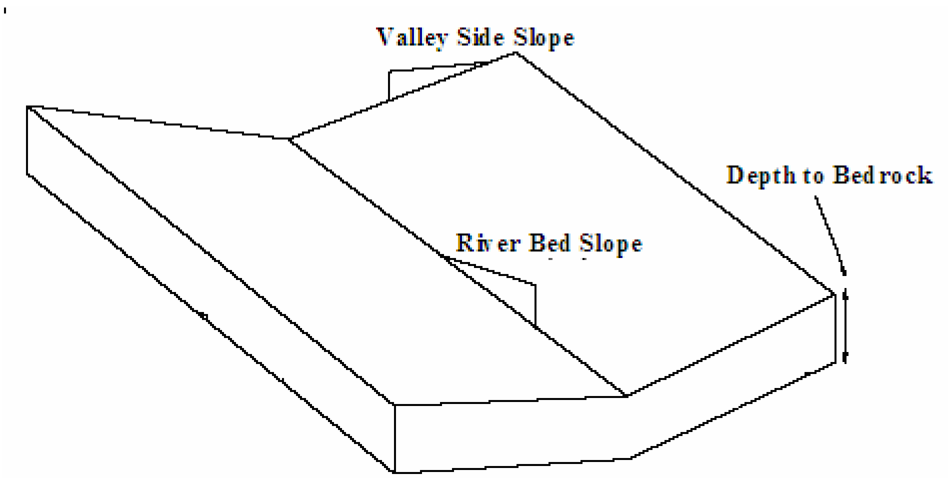


Figure 4. Geometric representation of the open-book watershed topography. Here, the depth to bedrock basically refers to the assumed depth of the effective soil column.

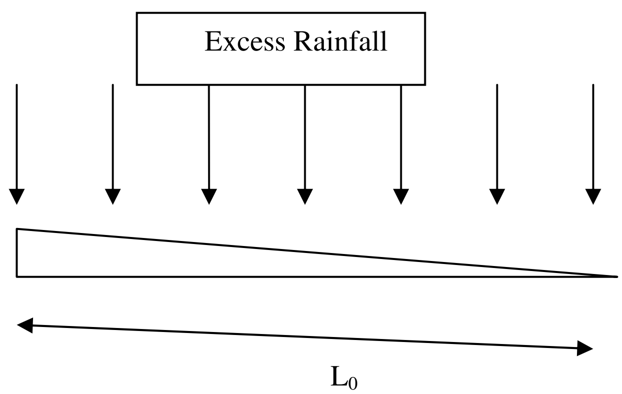
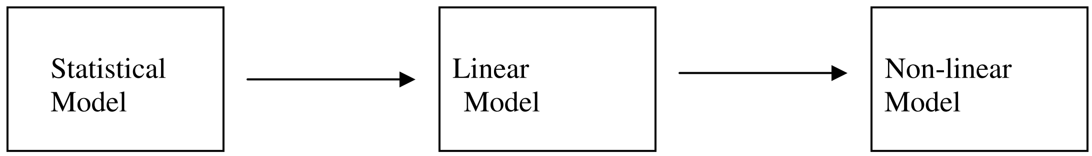


Figure 5. Overland flow routing from excess rainfall over pixels/zones.



$\text{Recharge} = \boldsymbol{a}\ p(t)$ $\qquad q_{ss} = \frac{s(t) - S_f}{t_c}$ $\qquad q_{ss} = \left[\frac{s - s_f}{a}\right]^{1/b}$

Increasing level of process complexity in rainfall-runoff transformation

Figure 6. The three rainfall-runoff process conceptualizations in the order of increasing complexity from left to right. The process difference is shown along with the model name.