

**A methodology for merging multisensor precipitation estimates based
on expectation-maximization and scale recursive estimation**

Rohit Gupta, V. Venugopal, and Efi Foufoula-Georgiou

St. Anthony Falls Laboratory, Department of Civil Engineering, University of Minnesota
2 Third Avenue S.E, Minneapolis, MN 55414 USA

Submitted to *Journal of Geophysical Research-Atmospheres*

November, 2004

1st Revision: April 2005

2nd Revision: August 2005

Corresponding author: efi@umn.edu

Abstract

Scale recursive estimation (SRE) is a Kalman filter-based methodology, which can be used to produce optimal (in terms of bias and minimum variance) estimates of a field at any desired scale given uncertain and sparse observations at different scales. SRE requires the specification of the state equation, which describes the variability of the precipitation process across scales, and the observation equation, which relates the observations to the state. Typical models for describing the multiscale rainfall variability are the multiplicative cascade models. However, in order to convert them into the additive form needed by SRE, one needs to work in the log space, creating thus a problem in handling zero-intermittency in a satisfactory way. In this paper, we propose an alternative approach, based on a data-driven identification methodology, which operates directly on the data and does not require a pre-specified multiscale model structure. Rather, system identification and estimation is performed simultaneously via a likelihood-based Expectation-Maximization (EM) procedure. The merits of the proposed approach versus approaches based on multiplicative cascade models are explored via several examples of synthetic and real precipitation fields. For practical application, the proposed approach will need to be extended to include the temporal evolution of storms. This extension presents theoretical challenges and until these are addressed, a simple alternative is explored of coupling the EM-SRE approach with a spatial downscaling methodology to merge precipitation observations available at different spatial and temporal scales. An example application is presented motivated by its relevance to the Global Precipitation Measuring (GPM) mission.

Index Terms – Precipitation, Kalman filtering, Scale Recursive Estimation, EM algorithm, Maximum likelihood estimation

1. Introduction

Precipitation is one of the most inhomogeneous and fast evolving hydrometeorological processes in space and time. The multiscale variability observed in precipitation is due to the nesting of small, transient storm elements within larger long-lived elements. In order to increase the accuracy of atmospheric and hydrologic predictions, accurate precipitation estimates are required for model initialization, data assimilation, and also model verification. A variety of sensors, for example, raingauges, radars, and satellites are used to obtain precipitation related measurements. Each measurement technique has some advantages and limitations. Raingauges and radars, for example, provide relatively the most accurate precipitation measurements but with limited coverage. On the other hand, infrared sensors on geostationary satellites provide a broad and continuous coverage but with limited accuracy, and microwave sensors on polar orbiting satellites stand somewhere in between. In order to produce accurate precipitation estimates, an obvious solution is to merge these disparate sources of measurements and exploit the advantages that each measurement technique has to offer.

Scale Recursive Estimation (SRE) [see *Chou et al.*, 1994a, b for original references] has recently been proposed as a methodology for merging multi-sensor, multi-scale precipitation measurements in order to obtain estimates of precipitation and their error statistics at desired spatial scales, for the purpose of model verification [*Tustison et al.*, 2003] or data assimilation [*Kumar, 1999; Primus et al.*, 2001]. The SRE methodology, which has its roots in Kalman filtering, explicitly takes into account the disparate (in scale) measurement sources and their sensor-dependent uncertainty. This methodology requires a multiscale stochastic model to describe the scale-to-scale variability of spatial precipitation. Several such models have been explored in the past

[e.g., see *Gupta and Waymire, 1993; Lovejoy and Schertzer, 1991; Kumar and Foufoula-Georgiou, 1993a, b; Harris et al., 1997; among others*] but a class of models that naturally fits into the SRE framework (because they can be brought into the recursive additive form required by SRE) is that of multiplicative cascade models [e. g., *Gupta and Waymire, 1993; Over and Gupta, 1994*]. These models have been used for rainfall applications within the SRE framework by *Primus et al. [2001]* and [2003].

There are two (interrelated) problems that arise in using multiplicative cascade models for spatial rainfall within the SRE methodology. First, in order to bring these models into the additive state-scale recursive equation required by SRE, one has to work in the logarithmic space. Since spatial rainfall fields contain zero values (due to intermittency) a small threshold is usually used to replace the zeros with nonzero values during the SRE procedure. Sensitivity of the fitted model parameters and SRE estimates to the chosen threshold value was reported in *Tustison et al. [2003]* although this issue was not pursued further. Second, by construction, multiplicative cascade models produce fields, which are non-zero everywhere within the modeling domain. If an imposed small threshold value were to be used to define “zeros” (as values below the chosen threshold) the statistics of these zeros would be completely predetermined by the cascade model parameters and would follow a power law distribution (i.e., zero areas of all sizes would be expected to be present). This might be a restriction, if the statistics of the zero areas do not follow power law distributions. Thus, both of the above issues pose limitations in considering multiplicative cascade models for rainfall within the SRE framework.

Motivated by these limitations, this paper proposes an alternative approach to scale recursive estimation based on a data-driven system identification methodology, which operates directly on the data (and not their logs) and does not require a

prespecified multiscale model structure (in that sense, the proposed approach is referred to as “nonparametric”). This is accomplished by a likelihood-based Expectation Maximization (EM) on scale-recursive dynamics on trees [e.g., *Kannan et al.*, 2000] which identifies and estimates the model recursively (and dynamically) from the available multiscale/ multisensor observations with no fixed structure of the process dynamics. As such, it provides a valuable alternative in many practical situations. The merits of the proposed nonparametric EM-SRE approach compared to parametric approaches are documented based on a suite of numerical experiments.

In practical applications, merging of multiscale observations has to be performed continuously over time giving rise to the need to have the temporal structure of precipitation also taken into account. Extending the SRE methodology to dynamic (spatio-temporally varying) fields is not a simple task. The challenge of multiscale estimation of dynamic systems lies in the prediction step, which requires untangling the spatial mixing due to temporal dynamics. This step can be involved even in simple dynamics such as diffusion processes. Research on SRE of dynamic fields includes that of *Ho et al.* [1996]. The idea behind their approach is that the multiscale models for the updated and predicted estimation errors are propagated through time in the same way that Kalman filter propagates the error covariances, but in a more computationally efficient manner, i.e., without computing or storing the full error covariance matrix. They introduced a reduced-order spatially interpolated multiscale model and its efficiency was demonstrated in several applications. Before these methodologies are explored towards the problem of merging multiscale spatio-temporal precipitation observations, it is worth considering simpler methodologies, which can provide insight into the problem. Such a simple methodology which relies on combining the EM-SRE methodology with a

downscaling (spatial or spatio-temporal) scheme to produce space-time merged precipitation products is explored in this paper via an example motivated by the sampling specifications of the Global Precipitation Measuring (GPM) mission.

This paper is structured as follows. In the next section, a brief overview of the SRE framework is presented while leaving the mathematical details for the appendix. Section 3 focuses on log-normal and bounded log-normal multiplicative cascades and numerical experiments are carried out to determine their merits and limitations for precipitation representation within SRE. In section 4, the non-parametric EM-SRE methodology is presented. Section 5 demonstrates, through numerical experiments, some advantages of the non-parametric over the multiplicative cascade parametric models. Section 6 presents a case study that is of potential relevance to the GPM mission. Namely, the EM-SRE framework is combined with spatial downscaling to accommodate the merging of observations available at different scales and different times. Finally, conclusions and open problems for future research are presented in section 7.

2. Scale Recursive Estimation Framework

A multiscale process can be represented on an inverted tree, as shown in Figure 1. The tree can essentially be seen as a way of connecting the information about the process at different scales. Each node on the tree corresponds to a unique combination of scale and spatial location and is given the location index λ , and the spatial scale index $m(\lambda)$, which is assumed in our case to be the same for all nodes at the same spatial scale. The multiscale stochastic models of interest are specified in terms of scale recursive dynamic equations defined on the tree. Specifically, if $X(\lambda)$ denotes the value of the process state

at node λ , the evolution of the multiscale process from coarse ($\gamma\lambda$) to fine (λ) scale is of the form:

$$X(\lambda) = A(\lambda)X(\gamma\lambda) + B(\lambda)W(\lambda) \quad (1)$$

where $X(\lambda)$ is the zero-mean state of the system, $A(\lambda)$ and $B(\lambda)$ are parameters that control the scale-to-scale variability of the process, and $W(\lambda) \sim N(0,1)$ is a noise component, which is independent of the state. The term $A(\lambda)X(\gamma\lambda)$ represents a coarse-to-fine scale prediction or interpolation, $B(\lambda)W(\lambda)$ represents the higher resolution detail added in going from one scale to the next finer scale. It is noted that the state $X(\lambda)$ can be a multi-dimensional vector with different variables but in this work, the state (spatial precipitation) is considered to be a scalar. It is also worthwhile to note that although the system parameters, such as $A(\lambda)$, $B(\lambda)$ and $W(\lambda)$, can vary with both scale and location, in this work, a special case is assumed in which the parameters are constant at each scale, i.e., the parameters are independent of location. Along with the estimate of the state, computing the error statistics or uncertainty of the estimates is also of interest. Defining the variance of the state as $P_x(\lambda) = E[X^2(\lambda)]$, using eqn. (1) and the fact that the state and the noise terms are independent, propagation of the variance $P_x(\lambda)$ can be shown to evolve from coarse to fine scale according to a Lyapunov equation as:

$$P_x(\lambda) = A^2(\lambda)P_x(\gamma\lambda) + B^2(\lambda) \quad (2)$$

This equation shows how the variance of the state at one location relates to that of its parent. The coarse-to-fine scale model of eqn. (1) can be inverted to give a model evolving from fine to coarse scale, which can be written in the form:

$$X(\gamma\lambda) = F(\lambda)X(\lambda) + W^*(\lambda) \quad (3)$$

where $W^*(\lambda) \sim N(0, Q(\lambda))$, and $F(\lambda)$ can be obtained from the parameters of the coarse-to-fine scale model, as the ratio of the variances of the states [see *Chou et al.*, 1994a]:

$$F(\lambda) = \frac{P_x(\gamma\lambda)}{P_x(\lambda)} A(\lambda) \quad (4)$$

Also, by taking $E[X^2(\gamma\lambda)]$, the variance of the state can simply be shown to evolve from fine to coarse scale as:

$$P_x(\gamma\lambda) = F^2(\lambda)P_x(\lambda) + Q(\lambda) \quad (5)$$

In order to incorporate the measurements of the process at different scales, a measurement model, which relates the measurements to the state of the system at a given location, is necessary. This model takes the form:

$$Y(\lambda) = C(\lambda)X(\lambda) + V(\lambda) \quad (6)$$

where $Y(\lambda)$ represents the measured quantity, $C(\lambda)$ relates the state to the measurement, and $V(\lambda) \sim N(0, R(\lambda))$, is the measurement error. In this work, $C(\lambda)$ is assumed to be equal to 1 since the measurement and the state represent the same quantity i.e. precipitation. This may not always be the case and, in general, $C(\lambda)$ can be a complex often non linear relationship between the measured quantity and the state of the system. The measurement model given by eqn. (6) takes into account the measurement uncertainty $V(\lambda)$, which in all practical cases differs from one scale to another scale because typically different instruments or sensors are employed to observe the process at different scales.

In order to compute the estimates of the process and their error statistics at every scale, the fine-to-coarse and coarse-to-fine state evolution equations, and the

measurement model, are all integrated together to form a single estimation framework. The multiscale estimates are computed from an upward sweep (a filtering step) in which the information is passed from one scale to the next coarsest scale, and a downward sweep (a smoothing step), which proceeds from coarse to fine scales. The upward and downward sweeps together represent a generalization of the Rauch-Tung-Striebel (RTS) smoothing algorithm [see *Chou et al.*, 1994a]. The upward sweep is a filtering step, which computes $E[X(\lambda) | Y_\lambda]$ for all nodes where Y_λ is all the data in the subtree below node λ . This is done recursively from the nodes at the bottom of the tree (finest scale) to the node at the top of the tree (coarsest scale) using an extension of Kalman filter to trees. The upward sweep consists of an initialization step, which is followed by the measurement update, scale propagation, and merging step. In the initialization step, the state is initialized to the global mean of the process, which is zero by definition, and the error variance at the smallest scale is initialized to the variance as predicted by the multiscale model. In the measurement update step, the state and the error variance are updated via the Kalman filter if the measurements are available at the scale of analysis. The updated state and error variance are then propagated to the next coarser scale using the prescribed multiscale model. Finally, because of the discrepancy in the number of pixels or nodes between various scales, the predicted state and error variance from the last step are combined through a weighted average (this is called the merging step). At this point, the upward sweep (which started at the finest scale, and went up to the coarsest scale) is complete, and the downward sweep begins. The downward sweep can be seen as a smoothing step, which computes $E[X(\lambda) | Y]$ for all the nodes where Y is the data in the entire tree. This step allows for information exchange between adjacent nodes, as those nodes have contributed to the same upward sweep estimates of the state and its

error variance. This step runs recursively from the top of the tree to the bottom. It uses the final solution of the previous filtering step as the initial point of the recursion. The mathematical details of this algorithm are given in Appendix A. The reader is also referred to *Chou et al.* [1994a, b] (see also *Kumar, 1999; Primus et al., 2001* and *Tustison et al., 2003*), for further details on the algorithm.

3. Multiplicative Cascades within Scale Recursive Estimation

The scale-recursive estimation framework requires the specification of a model which describes the multiscale variability of the process under study. A class of popular multiscale precipitation models is that of multiplicative cascade models (e.g., *Gupta and Waymire, 1993; Over and Gupta, 1994* and *Lovejoy and Schertzer, 1995*). A multiplicative cascade begins with the mean value at the root scale and operates in a (usually dyadic) tree to distribute the “mass” via successive application of a distributive operation at many scales. To evolve from one scale to the next finest scale on the tree, the process values are determined by multiplication of the values at the parent scale with “weights” drawn from a distribution. As such, the multiplicative cascade can be put into the recursive form given by:

$$\chi_c(\lambda) = \chi_c(\gamma\lambda)\omega(\lambda) \quad (7)$$

where $\chi_c(\lambda)$ is the value of the process at scale λ , $\chi_c(\gamma\lambda)$ is the value of the process at the parent node, $\gamma\lambda$ and $\omega(\lambda)$ are the multiplicative cascade weights. In order to incorporate this cascade model into the SRE framework, it is necessary to express it in an additive form, which can be achieved by taking the logs of eqn. (7):

$$\ln \chi_c(\lambda) = \ln \chi_c(\gamma\lambda) + \ln \omega_c(\lambda) \quad (8)$$

Details of two commonly used multiplicative cascade models (the lognormal cascade (LN) and bounded lognormal cascade (BLN)) and how these can be incorporated into the SRE framework can be found in *Tustison et al.* [2003] (see also Appendix B for a brief account). Spatial precipitation exhibits zero intermittency, and thus working with the logs of the data is problematic because $\log(0)$ is undefined. A simple way to handle the zero values is by replacing them with a small positive value (e.g., sensor detection threshold) at all scales or by replacing them at each scale differently based on the minimum value at that scale. For example, as suggested by *Tustison et al.* [2003], zeros can be replaced by:

$$\ln(0) \equiv \ln(\min Y(\lambda)) - c \quad (9)$$

where c is a chosen parameter, and $\min Y(\lambda)$ is the minimum observation at scale λ .

To quantify the sensitivity of the threshold value chosen to replace the zeros in the precipitation field on the estimated cascade parameters, a set of numerical experiments was performed. The log-normal and bounded log-normal cascades (see Appendix B) were fitted to two precipitation storms: a summer convective storm over Kansas City, Missouri (July 4th, 1995) as observed by the KEAX radar, and a tropical storm over Darwin, Australia (January 27th, 1998). The spatial resolution of both precipitation fields was $2 \times 2 \text{ km}^2$. A weighted least squares fitting was done by minimizing the difference between the theoretical variance (see Eqns. (B3) and (B7)) and the empirical variance (computed from the logarithms of the observed data at all available scales) [see also *Tustison et al.*, 2003 for further details]. It is worth mentioning here that we have plotted the variance of the log of the fields versus scale because it is this variance that is propagated from small to large scales and vice-versa in the SRE methodology.

Figures 2 and 3 show the theoretical (for the fitted LN and BLN cascades) and the empirical multiscale variance curves for the KEAX and Darwin fields, respectively, and for several thresholds, used to define the zeros: 10^{-4} mm/h, 10^{-3} mm/h, and a scale-dependent threshold as in eqn. (9) with parameter c set to 1 [see *Tustison et al.*, 2003 for this selection, which corresponds to approximately 10^{-3} mm/h at the smallest scale]. These figures also show the values of the estimated parameters for the LN and BLN cascades. Several observations can be made from these figures. First, it is noted that depending on the chosen threshold, the empirical variance (i.e., variance of the log of the observed field) changes significantly. This, naturally, affects the parameters of the fitted multiplicative cascade models (see Figures 2 and 3). As the cascade model parameters are directly related to the parameters of the multiscale state-space equation given by (1), and the propagation of variance given by (2), poorly estimated model parameters will affect the merging and the SRE estimates at any scale of interest [see also *Tustison et al.*, 2003].

In the absence of parametric multiscale models which can explicitly handle the zero-intermittency of rainfall, it is worth exploring non-parametric models, which can be incorporated into the SRE framework for merging multisensor observations. The simplest such model is based on utilizing a data-defined variance reduction curve (VRC), i.e., how the variance of the process changes with scale, via a look-up table or graph, and without approximating it with an apriori model, such as a multiplicative cascade. This approach bypasses the problem of zeros, but requires that such a curve can be reliably computed from available observations at multiple scales. Notice that if only a few scales of observation are available, some form of interpolation would have to be performed such that a VRC is defined over all scales of interest. Besides trivial linear interpolation,

another form of interpolation could be achieved through the process of aggregation of the high resolution observations, if available, and computation of the variance of the aggregated fields at the scales of interest. As shown in *Tustison* [2001], when an accurate VRC is available, the VRC-SRE approach works well. However, the estimation of an accurate VRC might not always be feasible owing to sparsity of data. It makes sense in these cases then, to follow an approach by which the multiscale structure is not explicitly prescribed (neither in form such as in multiplicative cascades, nor in how the variance changes with scale such as in the VRC) but it is left to be recursively estimated from all the available observations at all available scales. In the next section, such an approach for the simultaneous system identification and scale recursive estimation is proposed.

4. Expectation-Maximization System Identification Approach

An Expectation Maximization (EM) algorithm for estimation of the parameters of a multiscale stochastic process based on scale recursive dynamics on trees was introduced in *Kannan et al.* [2000] building on results of *Chou et al.* [1994a, b]. This approach was used to provide maximum likelihood (ML) estimates of the parameters for the general class of nonhomogeneous trees (i.e., nonuniform or irregular branching) with no fixed structure for the process dynamics. This approach altogether eliminates the need to prescribe a priori the type of a multiscale model; rather it uses the measurements available at multiple scales and dynamically evolves the multiscale state-space equation given by (1). Figure 4 shows a simple illustration of the EM algorithm. In general, the parameter set of the multiscale recursive framework for which we intend to find maximum likelihood estimates is $\{A(\lambda), B(\lambda), C(\lambda), R(\lambda) \mid \lambda \in T\}$, where T is the set

of all nodes in the tree. Let this parameter set be collectively denoted by θ . For our problem, we will fix the parameters $A(\lambda)$ and $C(\lambda)$ based on our understanding of the process and parameter $R(\lambda)$ based on the information on sensor uncertainties. So, the only parameter for which we intend to find the maximum likelihood estimate is $B(\lambda)$. Hence the parameter set θ only contains $B(\lambda)$.

In Maximum Likelihood identification of the multiscale state-space model using the EM algorithm, the E-Step of the algorithm involves the computation of the expected log likelihood of the observed data and missing data (all states and missing observations). For a single run i.e. for a single sequence of observations, the expected log likelihood is

$$E_{\theta} \{L(X, Y, \theta) | Y\}$$

where $L(X, Y, \theta)$ is the log-likelihood function that is defined as the joint log probability of the states and measurements, and is given by [Kannan et al., 2000, Digalakis et al., 1993]

$$\begin{aligned} L(X, Y, \theta) = & - \sum_{\lambda \in (T-0)} \{ \log |B(\lambda)B^T(\lambda)| + [X(\lambda) - A(\lambda)X(\gamma\lambda)]^T [B(\lambda)B^T(\lambda)]^{-1} [X(\lambda) - A(\lambda)X(\gamma\lambda)] \} \\ & - \sum_{\lambda \in T} \{ \log |R(\lambda)| + [Y(\lambda) - C(\lambda)X(\lambda)]^T R^{-1}(\lambda) [Y(\lambda) - C(\lambda)X(\lambda)] \} + \text{constant} \end{aligned} \quad (10)$$

The Expectation or E-Step computes the conditional expectations of complete-data sufficient statistics whereas the maximization or M-Step uses these statistics to re-estimate the model parameters. The computation of expected log likelihood depends on three expectations i.e. $E[X(\lambda) | Y]$, $E[X(\lambda)X^T(\lambda) | Y]$, $E[X(\lambda)X^T(\gamma\lambda) | Y]$, where Y is the data on the entire tree. The M-Step of the algorithm is described first before showing how the above expectations are computed in the E-Step.

The M-Step:

The parameter set of the multiscale state space model for which we intend to find out maximum likelihood estimates is $\{B(\lambda) \mid \lambda \in \mathbb{T}\}$. Maximizing the expected likelihood using multivariate regression to obtain new estimates of the parameters [Kannan *et al.*, 2000] gives:

$$\hat{B}(\lambda) = [E\{X(\lambda)X^T(\lambda)\} - \hat{A}(\lambda)E\{X(\gamma\lambda)X^T(\lambda)\}]^{1/2} \quad (11)$$

The M-step of the EM algorithm can also be used to update other parameters of the multiscale state space model. The reader is referred to Appendix C for details.

The E-Step:

The E-Step of the algorithm computes the expected quantities required in the right hand side of the above equations. For a complete set of observations Y_0 , these quantities can be written in terms of the observations and the smoothed estimates of the state and their associated error covariance [Kannan *et al.*, 2000]:

$$E\{X(\lambda) \mid Y_0\} = \hat{X}_s(\lambda) \quad (12)$$

$$E\{X(\lambda)X^T(\lambda) \mid Y_0\} = P_s(\lambda) + \hat{X}_s(\lambda)\hat{X}_s^T(\lambda) \quad (13)$$

$$E\{X(\lambda)X^T(\gamma\lambda) \mid Y_0\} = P_s(\lambda, \gamma\lambda) + \hat{X}_s(\lambda)\hat{X}_s^T(\gamma\lambda) \quad (14)$$

where $P_s(\lambda, \gamma\lambda) \equiv E\{[X(\lambda) - \hat{X}_s(\lambda)][X(\gamma\lambda) - \hat{X}_s(\gamma\lambda)]^T \mid Y_0\}$. The terms $\hat{X}_s(\lambda)$ and $P_s(\lambda)$ are computed by the downward sweep of the Rauch-Tung-Striebel (RTS) algorithm discussed previously (i.e., the extension of the Kalman filter to dyadic trees). The remaining term required is $P_s(\lambda, \gamma\lambda)$, which can easily be shown to be computed using terms from the RTS downward sweep as explained in detail in Appendix C.

When observations at some node are missing, these missing observations and the unseen state information are jointly treated as missing data by the EM algorithm. So the expectations of the observation terms used in the M-Step are given by

$$E\{Y(\lambda) | Y_0\} = \begin{cases} Y(\lambda) & \text{if observed} \\ C(\lambda)E\{X(\lambda) | Y_0\} & \text{if missing} \end{cases} \quad (15)$$

$$E\{Y(\lambda)Y^T(\lambda) | Y_0\} = \begin{cases} Y(\lambda)Y^T(\lambda) & \text{if observed} \\ R(\lambda) + C(\lambda)E\{X(\lambda)X^T(\lambda) | Y_0\}C^T(\lambda) & \text{if missing} \end{cases} \quad (16)$$

$$E\{Y(\lambda)X^T(\lambda) | Y_0\} = \begin{cases} Y(\lambda)E\{X^T(\lambda) | Y_0\} & \text{if observed} \\ C(\lambda)E\{X(\lambda)X^T(\lambda) | Y_0\} & \text{if missing} \end{cases} \quad (17)$$

Thus the EM-based estimation method can also be used to estimate the parameters from sequences of incomplete observations, which is very convenient for practical applications.

5. Application of the EM-SRE Algorithm

5.1. Testing Convergence and Estimation Accuracy

To investigate the convergence of the EM-SRE algorithm and the accuracy of the estimated parameters, numerical experiments were conducted using synthetic data of known multiscale structure. Spatial fields were generated using the state-space recursive equation, (eqn. (1)) and the measurement equation (eqn. (6)), with node- and scale-invariant parameters, $A(\lambda) = A$, $B(\lambda) = B$, $C(\lambda) = C$, $R(\lambda) = R$. As discussed before, for our application relating to precipitation, the parameters A and C are equal to 1. Several measurement uncertainty levels were investigated, but here we report results of almost perfect observations, with $R(\lambda) = 0.001$. The spatial fields were generated by varying the parameter $B(\lambda)$, which controls the information that is added when one

moves from coarse to fine scale, and the square of its value gives the difference in the process (in our case, log-rainfall) variance between those scales (since $A(\lambda) = 1$ in eqn. (2)). In this example, the parameter $B(\lambda)$ for all scales was set to 4 mm/h. Table 1 shows how the value of $B(\lambda)$ changes with several iterations of the EM algorithm. The results are reported in Table 1 for the case when the initial value of $B(\lambda)$ was chosen to be 1.0, although different initial values were tried. In every case the algorithm converged (after only a few iterations) to approximately the same value of $B(\lambda)$. The convergence criterion used was that the relative absolute difference between the values of the log-likelihood function from successive iterations is less than 10^{-2} . A more stringent condition on the log-likelihood function convergence required more iterations but did not significantly improve the estimate of $B(\lambda)$.

5.2. Example Performance of the EM-SRE Algorithm

Having tested the accuracy and convergence of the EM algorithm, we proceed with test applications on real precipitation fields. The radar-observed hourly precipitation field at a spatial resolution of $2 \times 2 \text{ km}^2$ over Darwin, Australia (January 27, 1998), which was discussed in section 3 was used for this analysis. The field contained 33% of zero values as computed from the available highest resolution ($2 \times 2 \text{ km}^2$) field over the region of interest (see Figure 5). As has been seen in Figure 3, depending on how the zero values were treated, different parameters of the LN and BLN cascade models were obtained. These parameters are expected to result in different SRE-merged estimates, along with their uncertainty. For example, using a lognormal cascade as the underlying multiscale model with the threshold-dependent parameters shown in Figure 3, the statistics of the estimated field at $4 \times 4 \text{ km}^2$ were found to exhibit significant dependence

on the threshold used, e.g., the standard deviation of the estimated field changed from 7.22 to 5.56 to 3.93 mm/h for the three thresholds of 10^{-4} , 10^{-3} , and scale-dependent threshold, respectively. Given this sensitivity of the cascade models to zero intermittency, the EM-SRE algorithm, which does not require any apriori model specification, appears to be a suitable alternative.

It is noted that the Kalman filtering recursive estimation has optimal performance for Gaussian distributions. In practical applications, approximate Gaussianity is typically achieved by applying transformations to the original data and working in the transformed space. When multiplicative cascade models are used within the SRE framework, they require working in the log space in order to transform the multiplicative structure of the models to the additive form required by the state-space equation. Although this log-transformation introduces difficulty in handling the zeros, it achieves an approximate Gaussianity of the PDF of rainfall apart from the possible mass at (or close to) 0 coming from the spatial intermittency (or from the replacement of zeros with arbitrarily small values). The proposed EM-SRE approach works in the real space and thus avoids the shortcoming of having to handle the zeros in the log-space. Approximate Gaussianity (apart again from the possible mass at zero) is achieved by applying a power transformation to the data. For the Darwin observations (33% zeros at a scale of $2 \times 2 \text{ km}^2$), a power of 0.17 was found to achieve the closest approximation to a Gaussian distribution of the non-zero values at all scales. It is noted that although the proposed SRE-EM approach is preferable to multiplicative cascades for not having to deal with zeros in the log-space, it does not truly overcome the intermittency problem, when it comes to achieving approximate Gaussianity. A large mass at zero will always prevent a good approximation to Gaussianity. This issue is further discussed in the conclusions. For

the Darwin rainfall example, the field at $2 \times 2 \text{ km}^2$ and $16 \times 16 \text{ km}^2$ were considered known (with zero observational uncertainty) and the EM-SRE algorithm was used to estimate the field at two intermediate scales: $4 \times 4 \text{ km}^2$ and $8 \times 8 \text{ km}^2$. The “true” fields at any scale were obtained by aggregation of the highest resolution $2 \times 2 \text{ km}^2$ fields. The results of this application are summarized in Table 2, and also displayed in Figure 5 for the $4 \times 4 \text{ km}^2$ estimated field. Similar results were obtained when observational uncertainty was introduced. It is noticed from Table 2 that, for the estimated field at $4 \times 4 \text{ km}^2$, the EM-SRE scheme is able to reproduce about 93% of the variability of the precipitation field while no bias is observed in the estimated field. Similarly, for the estimated field at $8 \times 8 \text{ km}^2$, the percentage reproduction of variability is about 97% again with no bias in the estimated field. Table 2 also reports the mean uncertainty values for the estimated fields at both estimation scales, as well as the values of the root mean square error (RMSE). It is noted that in comparing the estimated (merged) and true precipitation fields, the simple measures of performance reported in Table 2 are adequate and there is no need to use more sophisticated measures (e.g., multiscale or combined amplitude-distance measures presented in *Zepeda-Arcé et al.*, [2000] and *Venugopal et al.*, [2005] for forecast verification applications). This is because in multiscale merging applications, it is unlikely that the observed field at one scale (by one sensor) will exhibit drastically different features, e.g., significantly misplaced high rainfall areas, than the observed field at another scale (by another sensor) and thus magnitude-based measures would mostly be adequate to capture the differences in the compared fields. If this is not deemed to be the case in some applications, the above mentioned more sophisticated comparison measures might be used in addition to simple measures.

5.3. Effect of Missing Values on EM-SRE

In practice, there might be many missing values in the observed precipitation fields especially at the finer resolution due to errors in the recording instrument or due to the fact that a finer resolution sensor might not completely cover a large area over all times of interest. For that reason, various numerical experiments were conducted to study the effect of missing values in the precipitation field on the merged product produced by the proposed EM-SRE approach.

The January 27, 1998 hourly precipitation field at $2 \times 2 \text{ km}^2$ resolution over Darwin, Australia was used again as the base field, and two cases were analyzed: in the first case, it was assumed that there were no missing values in any of the input fields ($2 \times 2 \text{ km}^2$ and $32 \times 32 \text{ km}^2$ field) while in the second case, approximately 54% values (randomly sampled) in the fine scale precipitation field were considered missing, but the coarse resolution field at $32 \times 32 \text{ km}^2$ was assumed to be complete. Notice that in this case, identifying and fitting a multiplicative cascade model to the fine resolution field would be problematic; in fact, even the computation of the Variance Reduction Curve (VRC) at fine resolution would present a problem, making the proposed EM-SRE approach an attractive alternative.

Table 3 summarizes the results of this experiment. It is seen that for the case of no missing values, the EM-SRE scheme is able to reproduce about 93% of the variability of the precipitation field while no bias is observed in the estimated field. On the other hand, when there are missing values in the fine scale field (54% missing values), the proposed approach slightly over-estimates the mean and the standard deviation, although overall the statistics of the estimated $8 \times 8 \text{ km}^2$ field compare well to the case of no missing values. The importance of large scale information can also be inferred from the

results of this experiment. Specifically, as there are no missing values in the $32 \times 32 \text{ km}^2$ field, the information available at the large scale is utilized to fill the missing gaps in the estimated field at $8 \times 8 \text{ km}^2$ during the smoothing step. It can be seen from Table 3 that the mean uncertainty of the estimated field increases for the case of missing values. For the first case of no missing values, all the estimated $8 \times 8 \text{ km}^2$ pixel values have the same uncertainty. In the second case, however, pixels that have missing values (or are in the vicinity of missing values) have higher uncertainty. This is because the SRE scheme in the downward sweep propagates the uncertainty in a way that considers the neighborhood dependence of the spatially close nodes (pixels). This is clearly demonstrated in Figure 6 which displays a histogram of estimation uncertainty (error) in the $8 \times 8 \text{ km}^2$ field for the case of no missing (Figure 6 top) and 54% missing data (Figure 6 bottom), respectively.

6. Merging Infrequent High-resolution and Frequent Low-resolution Observations: A Case Study of Relevance to GPM

The SRE merging methodology has been applied so far to static spatial precipitation fields, i.e., at one instant of time or accumulations over a period of time. As discussed in the introduction, most applications would require merging observations sampled at different spatial and temporal scales, and thus would require an extended SRE methodology that can explicitly incorporate the temporal evolution of the precipitation field. Development of such a methodology is a research issue in itself, and is beyond the scope of the present paper. In this section, we explore an alternative simple methodology which couples the EM-SRE scheme with a spatial downscaling scheme to merge rainfall observations at different spatio-temporal scales, as for example those anticipated to result

from GPM (*Smith et al.*, 2002). Specifically we demonstrate via a numerical example that one can take advantage of the more frequent low-resolution observations and spatial downscaling to complement the less frequent high-resolution observations for a consistent, in both space and time, merged product.

The numerical experiment consists of 2 cases (see Figure 7). Case 1 is the ideal (best) case where observations are available at all times (only 3 hours are considered here for demonstration purposes; $t = 1, 2,$ and 3 hours) at the coarse (16 km) and fine (2 km) resolutions. Case 2 is a scenario in which the high resolution observations (2 km) are available only at $t = 1$ hr, but the coarse resolution (16 km) observations are available at all times ($t = 1, 2,$ and 3 hrs). In Case 2, the merging is done by two methods: a simple method (called Case 2a) in which the 2 km spatial structure at $t=1$ hr is assumed to hold true for $t = 2$ and 3 hrs also (i.e., no temporal evolution of the field over 3 hours), and only a simple renormalization of the total water depth is performed to preserve the observed 16km fields at $t = 2$ and 3 hrs; and a more sophisticated case (called Case 2b) at which the 16 km fields at $t = 2$ and 3 hrs are spatially downscaled [using the statistical downscaling parameters estimated from the 2 km field at $t = 1$ hr and the method of *Perica and Foufoula-Georgiou*, 1996a, b] and then SRE merging is performed on the original 16 km and the downscaled 2 km fields. For comparison purposes, the 3-hour aggregated fields are computed and are shown in Figure 8 for Cases 1, 2a and 2b. As expected, Case 1 gives a merged product which is the closest to the “true” field as for this case complete fine and coarse scale observations were available at all times. In Case 2b, it is observed that spatial downscaling (which considered the dynamic evolution of the field at the large scale to infer via downscaling its dynamic evolution at small scale)

significantly improved the merged product as compared to Case 2a where the fine-scale field was assumed static over a period of 3 hours.

It is noted that when coarser-scale observations are not available frequently over time, the spatial downscaling scheme used here would not be adequate to capture the small scale dynamics and one would need to implement the dynamic (spatio-temporal) downscaling model of *Venugopal et al.* [1999] where the fine-scale precipitation field could be propagated over a period of time preserving both the spatial multiscale structure of rainfall, and its temporal persistence. That of course, requires some model parameters (specifically the spatial scaling parameter H and the dynamic scaling exponent z) to be known a-priori for this type of storm. Although evidence exists that the spatial downscaling parameter H can be related to the Convective Available Potential Energy (CAPE) in the pre-storm environment [e.g., see *Perica and Foufoula-Georgiou*, 1996a] and thus it can be dynamically updated based on observable meteorological quantities, it is still not clear how the dynamic scaling parameter z could be related to physical observables of the storm.

7. Conclusions

The work presented in this paper proposes a framework for merging multi-scale multisensor precipitation observations via an Expectation-Maximization Scale Recursive Estimation (EM-SRE) algorithm. The framework explicitly takes into account the measurement disparity (in scale) and the measurement noise, and it can easily handle missing observations at any scale. The EM algorithm is used in conjunction with SRE to iteratively perform, in parallel, system identification and estimation of the multiscale state-space model. The proposed approach is a data-driven approach and does not

assume any model a priori, such as multiplicative cascades; rather it identifies and estimates the model recursively based on the available measurements at multiple scales. The proposed EM-SRE approach appears to be a promising technique to merge precipitation estimates available at different scales especially when lack of high resolution observations and/or presence of high zero-intermittency preclude the identification and reliable estimation of parametric multiscale models. The presence of zero-intermittency in spatial rainfall is partially treated by the proposed approach by not having to work in the log space, as required when using the multiplicative cascades. However, the presence of zeros (both in the multiplicative cascades and in the proposed approach) precludes close approximation to Gaussianity required for optimality of the Kalman filter methodology. In the example cases considered here, the % of zeros was of the order of 30% at the highest resolution, and the results were satisfactory. For higher % of zeros, it is recommended that a preprocessing of the observations be performed to identify the zeros that come from the “inside” of the storm versus the ones that are “outside” of the storm (these will show up as zeros at all scales – even the larger ones). In that way, one can define the area over which merging is to be done and exclude the background non-rainy areas. Practical applications will not only have to deal with zero intermittency, but also with the time evolution of the storm. Theoretical extension of the EM-SRE framework to include time was not considered in this paper. However, motivated by practical applications related to the upcoming Global Precipitation Measuring (GPM) mission, a simple alternative methodology which couples the spatial EM-SRE approach with a spatial downscaling scheme was explored and was found promising based on a limited number of case studies. Future research should undertake a

more extensive testing and also address the extension of the proposed framework to space-time nonparametric multiscale estimation and merging.

Acknowledgements. This research was funded by NASA (Grants NAG5-12909 and NAG5-13639) and NSF (ATM-0130394), and NSF (EAR-0120914) as part of the National Center for Earth-Surface Dynamics (NCED) at the University of Minnesota). Computer resources were provided by the Minnesota Supercomputer Institute, Digital Technology Center at the University of Minnesota. All this support is gratefully acknowledged.

Appendix A. Details of the SRE framework

Before giving details of the SRE algorithm, it is necessary to define some terms:

- $Y_\lambda = \{Y(s) \mid s = \lambda \text{ or } s \text{ is a descendant of } \lambda\}$ is the set of measurements at all nodes below λ including the measurement at node λ .
- $Y_\lambda^+ = \{Y(s) \mid s \text{ is a descendant of } \lambda\}$ is the set of measurements at all nodes below λ excluding the measurement at node λ .
- $X(\lambda \mid \lambda)$ is used in place of $X(\lambda \mid Y_\lambda)$, which is the best estimate of $X(\lambda)$ given measurements at λ and all the nodes below λ .
- $X(\lambda \mid \lambda+)$ is used in place of $X(\lambda \mid Y_\lambda^+)$, which is the best estimate of $X(\lambda)$ given measurements at all nodes below λ .
- Similar notations are used for P .

A.1 Initialization at the Finest Scale

For each node λ at the finest scale, the following prior values are assigned:

$$X(\lambda \mid \lambda+) = 0, \quad \forall \lambda \ni m(\lambda) = m_{\max} \quad (\text{A.1})$$

$$P(\lambda \mid \lambda+) = P_x(\lambda), \quad \forall \lambda \ni m(\lambda) = m_{\max} \quad (\text{A.2})$$

A.2 Upward Sweep

The upward sweep computes the best estimates of the state $X(\lambda)$ at node λ given measurements at or below node λ . It consists of three steps at each scale:

A.2.1 Measurement Update Step:

$$X(\lambda \mid \lambda) = X(\lambda \mid \lambda+) + K(\lambda)[Y(\lambda) - C(\lambda)X(\lambda \mid \lambda+)] \quad (\text{A.3})$$

$$P(\lambda \mid \lambda) = P(\lambda \mid \lambda+)[1 - K(\lambda)C(\lambda)] \quad (\text{A.4})$$

where $K(\lambda)$ is the Kalman gain, a weight which is optimally chosen such that it minimizes the expected error variance of the state. The Kalman gain is given by

$$K(\lambda) = \frac{P(\lambda | \lambda+)C(\lambda)}{P(\lambda | \lambda+)C^2(\lambda) + R(\lambda)} \quad (\text{A.5})$$

A.2.2 Scale Propagation Step:

$$X(\gamma\lambda | \lambda_i) = F(\lambda_i)X(\lambda_i | \lambda_i) \quad (\text{A.6})$$

$$P(\gamma\lambda | \lambda_i) = F^2(\lambda_i)P(\lambda_i | \lambda_i) + Q(\lambda) \quad (\text{A.7})$$

where $Q(\lambda)$ is given by

$$Q(\lambda) = P_x(\gamma\lambda) - \frac{A^2(\lambda)P_x^2(\gamma\lambda)}{P_x(\lambda)} \quad (\text{A.8})$$

A.2.3 Merging Step:

$$P(\gamma\lambda | \gamma\lambda+) = \left[\frac{\{1 - q(\gamma\lambda)\}}{P_x(\gamma\lambda)} + \sum_{i=1}^{q(\gamma\lambda)} \frac{1}{P(\gamma\lambda | \lambda_i)} \right]^{-1} \quad (\text{A.9})$$

$$X(\gamma\lambda | \gamma\lambda+) = P(\gamma\lambda | \gamma\lambda+) \sum_{i=1}^{q(\gamma\lambda)} \frac{X(\gamma\lambda | \lambda_i)}{P(\gamma\lambda | \lambda_i)} \quad (\text{A.10})$$

A.3 Downward Sweep

The filtered estimates at the root node are the smoothed estimates and are used as the starting point in the downward sweep. The smoothed estimates at the remaining nodes are found by distributing the information back down the tree.

$$X_s(\lambda) = X(\lambda | \lambda) + J(\lambda)[X_s(\gamma\lambda) - X(\gamma\lambda | \lambda)] \quad (\text{A.11})$$

$$P_s(\lambda) = P(\lambda | \lambda) + J^2(\lambda)[P_s(\gamma\lambda) - P(\gamma\lambda | \lambda)] \quad (\text{A.12})$$

where, $J(\lambda)$ is a weighting coefficient, which is given by

$$J(\lambda) = F(\lambda) \frac{P(\lambda | \lambda)}{P(\gamma\lambda | \lambda)} \quad (\text{A.13})$$

Appendix B: Multiplicative cascades

Log-normal Cascades

For the log-normal cascade, the cascade weights at all scales come from the same lognormal distribution:

$$\omega(\lambda) = e^{(\sigma Z - \sigma^2/2)} \quad (\text{B.1})$$

where $Z \sim N(0,1)$, and σ is the model parameter which relates to the variability of the process. It can be shown [e.g., *Tustison et al.*, 2003] that the log-normal cascade model parameter σ can be related to the parameter of the SRE state equation (eqn. (1)) in the following way:

$$B(\lambda) = \sigma \quad (\text{B.2})$$

Also, the variance of the log-process can be shown to be:

$$P_X(\lambda) = P_X(\lambda_0) + m(\lambda)\sigma^2 \quad (\text{B.3})$$

where λ_0 is the root-scale, and $m(\lambda)$ is the index of scale, as shown in Figure 1.

Bounded Log-Normal Cascades

For the bounded log-normal cascade, the cascade weights have the same form as those of the lognormal cascade, except that now σ is a function of scale and thus is given by $\sigma(\lambda)$. Using the subscript “bc” to refer to the bounded log-normal cascade, these weights may be written as [e.g., *Menabde*, 1998; *Menabde and Sivapalan*, 2000]:

$$\omega_{bc}(\lambda) = e^{(\sigma_{bc}(\lambda)Z - \sigma_{bc}(\lambda)^2/2)} \quad (\text{B.4})$$

The model parameter $\sigma_{bc}(\lambda)$ is chosen so that the cascade weights follow a specified change with scale:

$$\sigma_{bc}(\lambda) = \sigma_1 2^{-(m(\lambda)-1)H} \quad (\text{B.5})$$

where $m(\lambda)$ is the scale index that has a value of zero for the coarsest scale. This formulation requires $\sigma_{bc}(\lambda)$ to be initialized at scale $m(\lambda) = 1$ with the value $\sigma_{bc}(\lambda) = \sigma_1$, which must be specified. The parameter H controls how fast the variance of the weights decays with the increasing scale index $m(\lambda)$. It can be shown [e.g., see *Tustison et al.*, 2003] that the bounded log-normal cascade model parameter $\sigma_{bc}(\lambda)$ can be related to the parameter of the state equation given by (1) in the following way:

$$B(\lambda) = \sigma_{bc}(\lambda) \tag{B.6}$$

The variance of the log-process can be shown to be:

$$P_X(\lambda) = P_X(\lambda_0) + \sigma_1^2 \left[\frac{(1 - 4^{-m(\lambda)H})}{1 - 4^{-H}} \right] \tag{B.7}$$

Appendix C: EM Algorithm

Equations of the M-Step for updating all the parameters of the multiscale state space model in order to maximize the expected log-likelihood function are given as follows:

$$\hat{A}(\lambda) = [E\{X(\lambda)X^T(\gamma\lambda)\}][E\{X(\gamma\lambda)X^T(\gamma\lambda)\}]^{-1} \quad (\text{C.1})$$

$$\hat{B}(\lambda) = [E\{X(\lambda)X^T(\lambda)\} - \hat{A}(\lambda)E\{X(\gamma\lambda)X^T(\lambda)\}]^{1/2} \quad (\text{C.2})$$

$$\hat{C}(\lambda) = [E\{Y(\lambda)X^T(\lambda)\}][E\{X(\lambda)X^T(\lambda)\}]^{-1} \quad (\text{C.3})$$

$$\hat{R}(\lambda) = [E\{Y(\lambda)Y^T(\lambda)\} - \hat{C}(\lambda)E\{X(\lambda)Y^T(\lambda)\}] \quad (\text{C.4})$$

Computation of $P_s(\gamma, \gamma\lambda)$ (required for E-Step):

$P_s(\lambda, \gamma\lambda)$ is computed directly in the downward sweep using the result that the smoothed error is a Gauss-Markov process. The smoothed error $\tilde{x}_s(\lambda) = x(\lambda) - \hat{x}_s(\lambda)$ has been shown to be modeled as a multiscale process [Luetzgen and Willsky, 1995a, b; Luetzgen et al., 1993].

$$\tilde{x}_s(\lambda) = J(\lambda)\tilde{x}_s(\gamma\lambda) + \tilde{w}(\lambda) \quad (\text{C.5})$$

where, $\tilde{w}(\lambda)$ is white noise and has zero mean with covariance given by

$$P(\lambda | \lambda) - P(\lambda | \lambda)F^T(\lambda)P^{-1}(\gamma\lambda | \lambda)F(\lambda)P(\lambda | \lambda) \quad (\text{C.6})$$

Thus,

$$\begin{aligned} P_s(\lambda, \gamma\lambda) &= E\{\tilde{x}_s(\lambda)\tilde{x}_s^T(\gamma\lambda) | Y_0\} \\ &= E\{[J(\lambda)\tilde{x}_s(\gamma\lambda) + \tilde{w}(\lambda)]\tilde{x}_s^T(\gamma\lambda) | Y_0\} \end{aligned}$$

$$\text{Or, } P_s(\lambda, \gamma\lambda) = J(\lambda)P_s(\gamma\lambda) \quad (\text{C.7})$$

References

- Chou, K. C., A. S. Willsky, and A. Benveniste, Multiscale recursive estimation, data fusion, and regularization, *IEEE Trans. Automat. Control*, 39, 464–478, 1994a.
- Chou, K. C., A. S. Willsky, and R. Nikoukhah, Multiscale systems, Kalman filters, and Riccati equations, *IEEE Trans. Automat. Control*, 39, 479–492, 1994b.
- Digalakis, V. V., J. Rohlicek, and M. Ostendorf, ML estimation of a stochastic linear system with the EM algorithm and its application to speech recognition, *IEEE Trans. Speech Audio Proc.*, 1(4), 431–442, 1993.
- Gupta, V. K., and E. C. Waymire, A statistical analysis of mesoscale rainfall as a random cascade, *J. Appl. Meteorol.*, 32, 251–267, 1993.
- Harris, D., A. W. Seed, M. Menabde, and G. L. Austin, Factors affecting multiscaling analysis of rainfall time series, *Nonlinear Processes Geophys.*, 4(3), 137–156, 1997.
- Ho, T., P. Fieguth, and A. Willsky, Multiresolution stochastic models for efficient solution of large-scale space-time estimation problems, *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, (VI), 3098-4001, 1996.
- Kannan, A., M. Ostendorf, W. C. Karl, D. A. Castañon, and R. K. Fish, ML parameter estimation of a multiscale stochastic process using the EM algorithm, *IEEE Trans. Signal Proc.*, 48(6), 1836-1840, 2000.
- Kumar, P., A multiple scale state-space model for characterizing subgrid scale variability of near-surface soil moisture, *IEEE Trans. Geosci. Remote Sens.*, 37, 182–197, 1999.
- Kumar, P., and E. Foufoula-Georgiou, A multicomponent decomposition of spatial rainfall fields, 1, Segregation of large and small-scale features using wavelet transforms, *Water Resour. Res.*, 29(8), 2515–2532, 1993a.

- Kumar, P., and E. Foufoula-Georgiou, A multicomponent decomposition of spatial rainfall fields, 2, Self-similarity in fluctuations, *Water Resour. Res.*, 29(8), 2533–2544, 1993b.
- Lovejoy, S., and D. Schertzer, Multifractal analysis techniques and the rain and cloud fields from 10^{-3} to 10^{-6} m, in *Non-linear Variability in Geophysics: Scaling and Fractals*, edited by D. Schertzer and S. Lovejoy, *Kluwer Acad.*, Norwell, Mass., 1991.
- Luetzgen, M. R., and A. S. Willsky, Multiscale smoothing error models, *IEEE Trans. Automat. Cont.*, 40, 173–175, 1995a.
- Luetzgen, M. R., and A. S. Willsky, Likelihood calculation for a class of multiscale stochastic models, with application to texture discrimination, *IEEE Trans. Image Proc.*, (4), 194–207, 1995b.
- Luetzgen, M. R., W. C. Karl, A. S. Willsky, and R. R. Tenney, Multiscale representations of Markov random fields, *IEEE Trans. Signal Proc.*, 41, 3377–3396, 1993.
- Menabde, M., Bounded lognormal cascades as quasi-multiaffine random processes, *Nonlinear Processes Geophys.*, 5(2), 63–68, 1998.
- Menabde, M., and M. Sivapalan, Modeling of rainfall time series and extremes using bounded random cascades and Lévy-stable distributions, *Water Resour. Res.*, 36(11), 3293–3300, 2000.
- Over, T. M., and V. K. Gupta, Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large scale forcing. *J. Appl. Meteorol.*, 33(12), 1526 – 1542, 1994.
- Perica, S. and E. Foufoula-Georgiou, Linkage of scaling and thermodynamic parameters of rainfall: results from midlatitude mesoscale convective systems, *J. Geophys. Res.*, 101(D3), 7431-7448, 1996a.

- Perica, S. and E. Foufoula-Georgiou, A model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *J. Geophys. Res.*, 101(D21), 26,347-26,361, 1996b.
- Primus, I., D. McLaughlin, and D. Entekhabi, Scale-recursive assimilation of precipitation data, *Adv. Water Resour.*, 24, 941-953, 2001.
- Smith, E.A., A. Mehta, and J.M. Shepherd, Description of Global Precipitation Measurement (GPM) Mission. *GPM Report Series No. 6* (E.A. Smith, J.W. Adams, eds.), NASA/Goddard Space Flight Center, Tech Memo., Greenbelt, MD, 2002.
- Tustison, B., Multiscale techniques for the verification of quantitative precipitation forecasts, M.S. Thesis, *University of Minnesota*, June 2001.
- Tustison, B., D. Harris, and E. Foufoula-Georgiou, Scale-recursive estimation for multi-sensor QPF verification: A preliminary assessment, *J. Geophys. Res.*, 108(D8), 8377-8390, 2003.
- Venugopal, V., E. Foufoula-Georgiou, and V. Sapozhnikov, A space-time downscaling model for rainfall, *J. Geophys. Res.*, 104(D16), 19,705-721, 1999.
- Venugopal, V., S. Basu, and E. Foufoula-Georgiou, A new metric for comparing precipitation patterns with an application to ensemble forecasts, *J. Geophys. Res.*, 110, doi:10.1029/2004JD005395, 2005.
- Zepeda-Arce, J., E. Foufoula-Georgiou, and K. Droegemeier, Space-time rainfall organization and its role in validating quantitative precipitation forecasts, *J. Geophys. Res.*, 105(D8), 10,129-10,146, 2000.

Figure Captions:

- Figure 1:** Representation of a multiscale process on a grid and its associated quadtree.
- Figure 2:** Variance of the natural log of the observed field and the fitted lognormal and bounded log-normal cascade model variance versus scale for the hourly precipitation field over Kansas observed by the KEAX radar. The fitting was done for various thresholds to show the sensitivity to zero-intermittency. Thresholds were set to (a) 10^{-3} mm/h; (b) 10^{-4} mm/h; and (c) $\ln[\min Y(\lambda)] - c$, (where $c = 1$, and $\min Y(\lambda)$ is the minimum observed value at scale λ).
- Figure 3:** Same as Figure 2, but for the radar-observed precipitation field over Darwin, Australia.
- Figure 4:** A simple illustration of the EM algorithm.
- Figure 5:** Illustration of the results of merging the 2×2 km² and 16×16 km² precipitation fields (over Darwin, Australia) via the EM-SRE methodology to produce a merged product at 4×4 km². Comparison of the spatial autocorrelation structure of the observed and estimated fields is also displayed.
- Figure 6:** Probability distribution of the uncertainty of the estimated field at 8×8 km² when (a) there are no missing values in the 2×2 km² input field; and (b) when there are about 54% missing values in the 2×2 km² input field.
- Figure 7:** Schematic of the two scenarios considered to explore the merging of more frequent low resolution observations with less frequent high resolution observations.

Figure 8: Illustration of the utility of using a spatial downscaling scheme in conjunction with the EM-SRE methodology for the purpose of merging infrequent high-resolution observations with more frequent low-resolution observations.

Table Captions:

- Table 1:** Estimation of parameter $B(\lambda)$ as a function of the number of the EM algorithm iterations. The convergence criterion is on the relative change of the log-likelihood function in successive iterations (see text).
- Table 2:** Statistics of the actual and the estimated field at a spatial resolution of $4 \times 4 \text{ km}^2$ and $8 \times 8 \text{ km}^2$ for the Darwin storm. The estimated field (via the EM-SRE algorithm) is compared to the actual field at two scales in terms of RMSE and Bias. The mean uncertainty of the estimated field is also reported.
- Table 3:** Statistics of the actual and the estimated field at $8 \times 8 \text{ km}^2$. The estimated field (via the EM-SRE algorithm) for both cases (no missing values and missing values in the input field) is compared to the actual field at $8 \times 8 \text{ km}^2$ in terms of RMSE and Bias. The mean uncertainty of the estimated field is also reported for both the cases.

Table 1: Estimation of parameter $B(\lambda)$ as a function of the number of the EM algorithm iterations. The convergence criterion is on the relative change of the log-likelihood function in successive iterations (see text).

EM Iterations	Parameter $B(\lambda)$
Initial Value	1.00
Iter. 1	3.74
Iter. 2	3.90
Iter. 3	4.00
Iter. 4	4.12
Iter. 5	4.12
Iter. 6	4.12
Actual Value	4.00

Table 2: Statistics of the actual and the estimated field at a spatial resolution of $4 \times 4 \text{ km}^2$ and $8 \times 8 \text{ km}^2$ for the Darwin storm. The estimated field (via the EM-SRE algorithm) is compared to the actual field at two scales in terms of RMSE and Bias. The mean uncertainty of the estimated field is also reported.

Observation Scale	Estimation Scale	Actual or Estimated	Statistical Parameter				
			Mean (mm/hr)	Std. Dev. (mm/hr)	RMSE (mm/hr)	Bias (mm/hr)	Mean Uncertainty (mm/hr)
$2 \times 2 \text{ km}^2$ and $16 \times 16 \text{ km}^2$	$4 \times 4 \text{ km}^2$	Actual	1.74	4.14	-	-	-
		Estimated	1.74	3.87	0.29	0.0	0.90
$2 \times 2 \text{ km}^2$ and $16 \times 16 \text{ km}^2$	$8 \times 8 \text{ km}^2$	Actual	1.74	3.69	-	-	-
		Estimated	1.74	3.59	0.10	0.0	0.35

Table 3: Statistics of the actual and the estimated field at $8 \times 8 \text{ km}^2$. The estimated field (via the EM-SRE algorithm) for both cases (no missing values and missing values in the input field) is compared to the actual field at $8 \times 8 \text{ km}^2$ in terms of RMSE and Bias. The mean uncertainty of the estimated field is also reported for both the cases.

Input Fields	Est. Scale	Actual or Estimated	Statistical Parameter				
			Mean (mm/hr)	Std. Dev. (mm/hr)	RMSE (mm/hr)	Bias (mm/hr)	Mean Uncertainty (mm/hr)
$2 \times 2 \text{ km}^2$ (no missing values) and $32 \times 32 \text{ km}^2$	$8 \times 8 \text{ km}^2$	Actual	1.96	4.92	-	-	-
		Estimated	1.96	4.61	0.35	0.0	1.30
$2 \times 2 \text{ km}^2$ (54% missing values) and $32 \times 32 \text{ km}^2$	$8 \times 8 \text{ km}^2$	Actual	1.96	4.92	-	-	-
		Estimated	2.04	5.08	1.29	-0.08	1.54

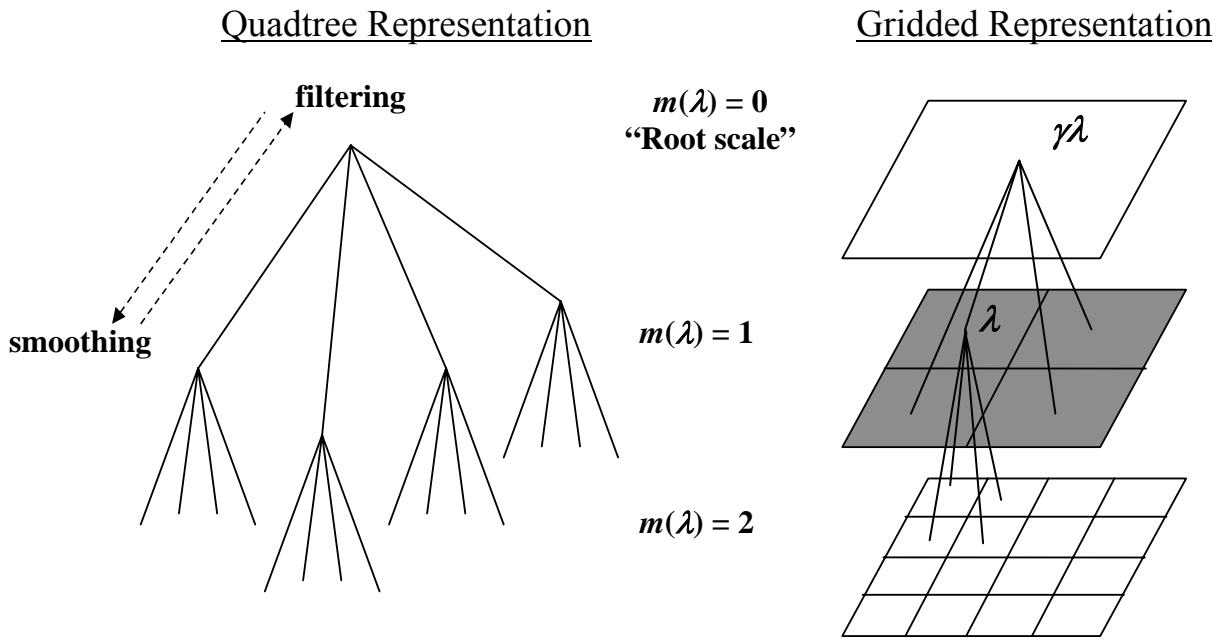


Figure 1: Representation of a multiscale process on a grid and its associated quadtree.

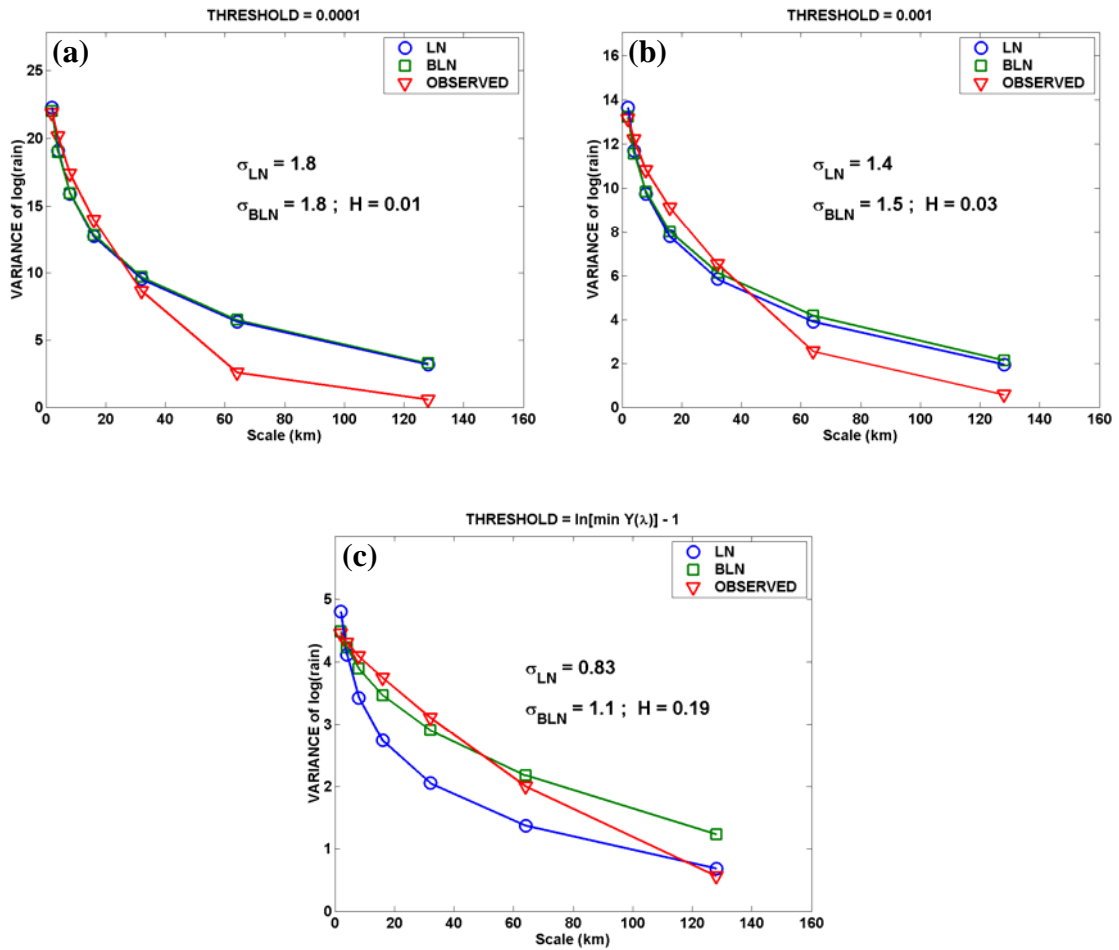


Figure 2: Variance of the natural log of the observed field and the fitted lognormal and bounded log-normal cascade model variance versus scale for the hourly precipitation field over Kansas observed by the KEAX radar. The fitting was done for various thresholds to show the sensitivity to zero-intermittency. Thresholds were set to (a) 10^{-3} mm/h; (b) 10^{-4} mm/h; and (c) $\ln[\min Y(\lambda)] - c$, (where $c = 1$, and $\min Y(\lambda)$ is the minimum observed value at scale λ).

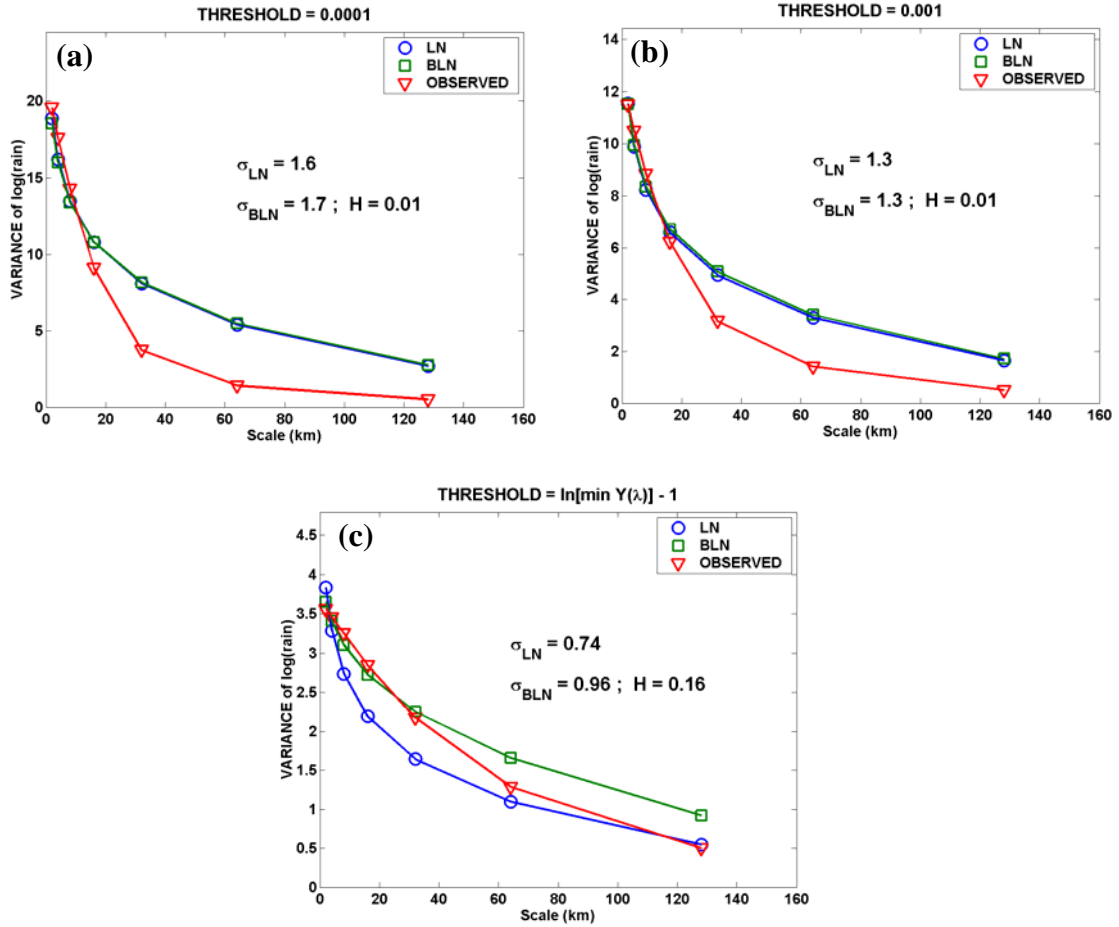


Figure 3: Same as Figure 2, but for the radar-observed precipitation field over Darwin, Australia.

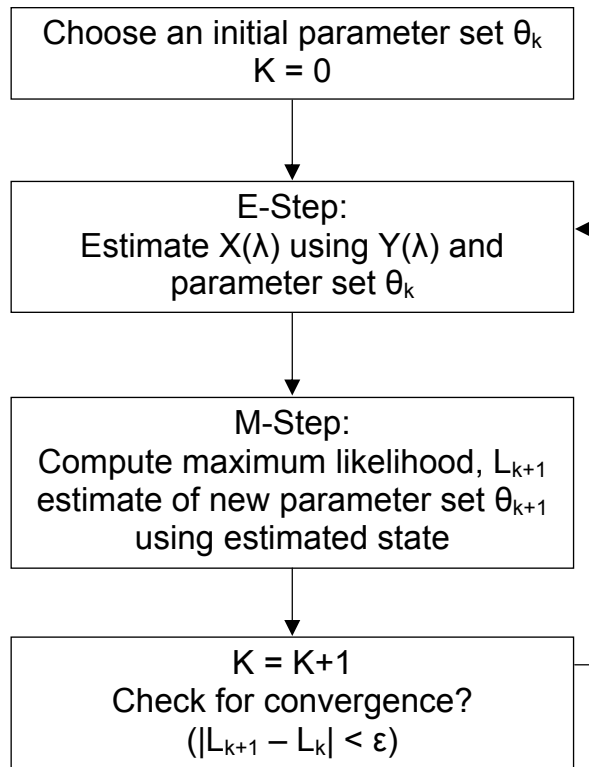


Figure 4: A simple illustration of the EM algorithm.

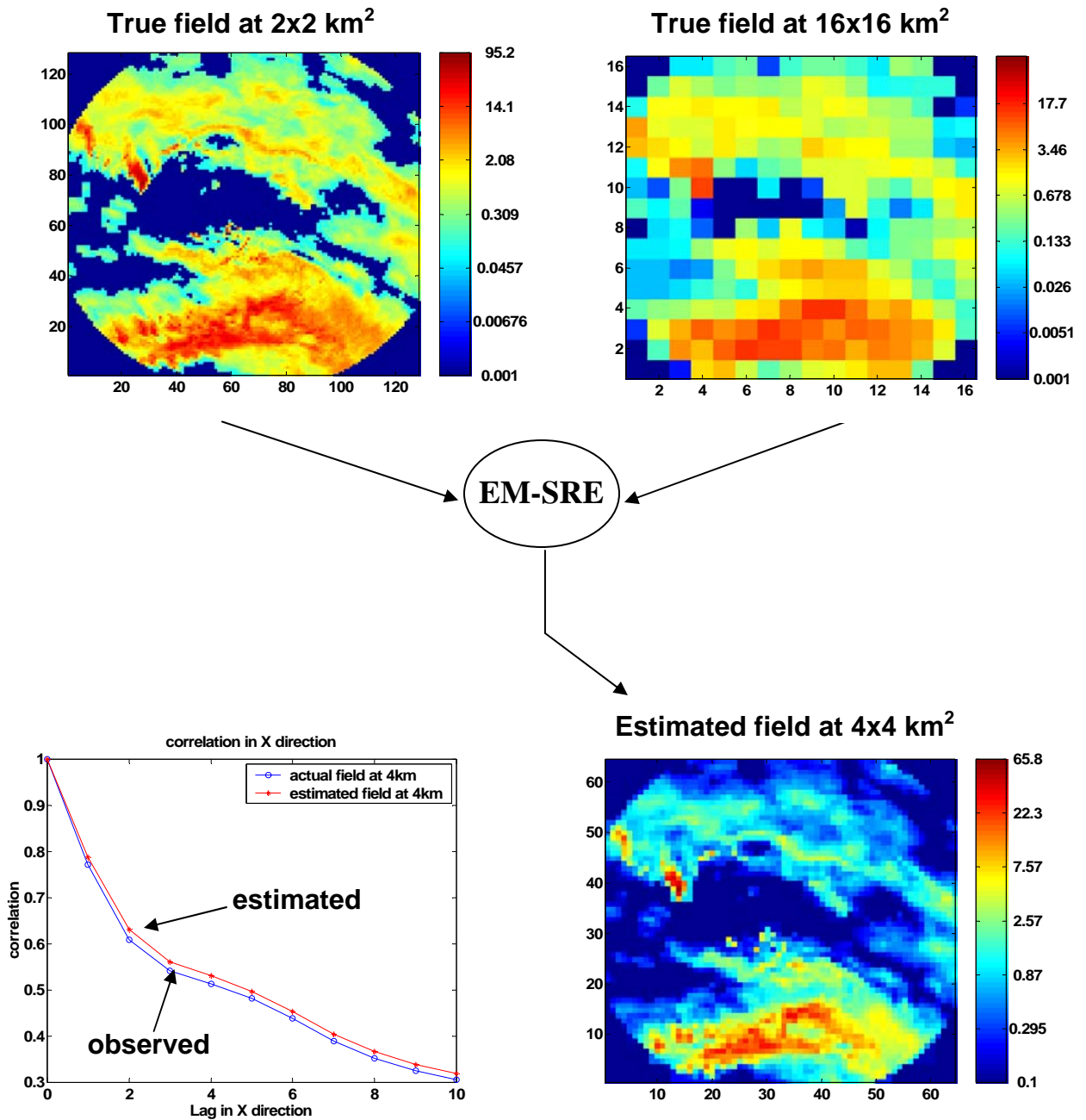


Figure 5: Illustration of the results of merging the 2x2 km² and 16x16 km² precipitation fields (over Darwin, Australia) via the EM-SRE methodology to produce a merged product at 4x4 km². Comparison of the spatial autocorrelation structure of the observed and estimated fields is also displayed.

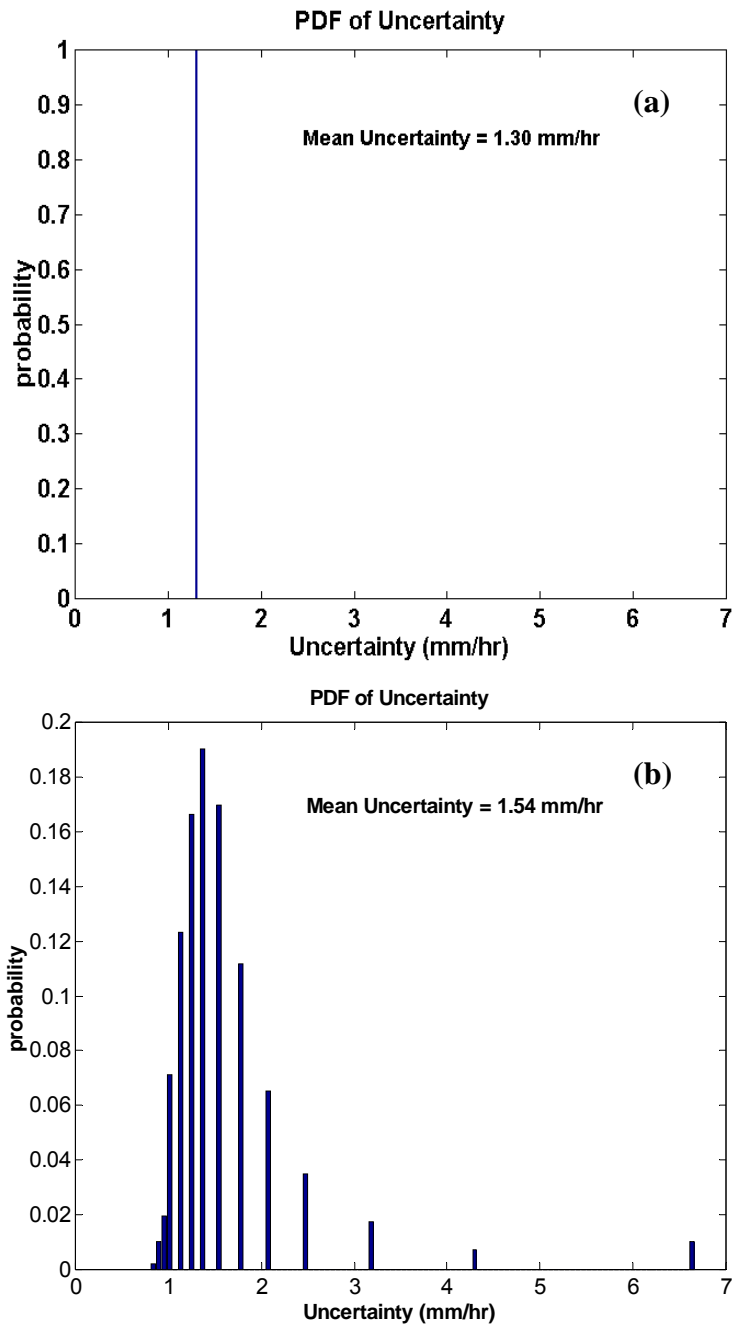


Figure 6: Probability distribution of the uncertainty of the estimated field at $8 \times 8 \text{ km}^2$ when (a) there are no missing values in the $2 \times 2 \text{ km}^2$ input field; and (b) when there are about 54% missing values in the $2 \times 2 \text{ km}^2$ input field.

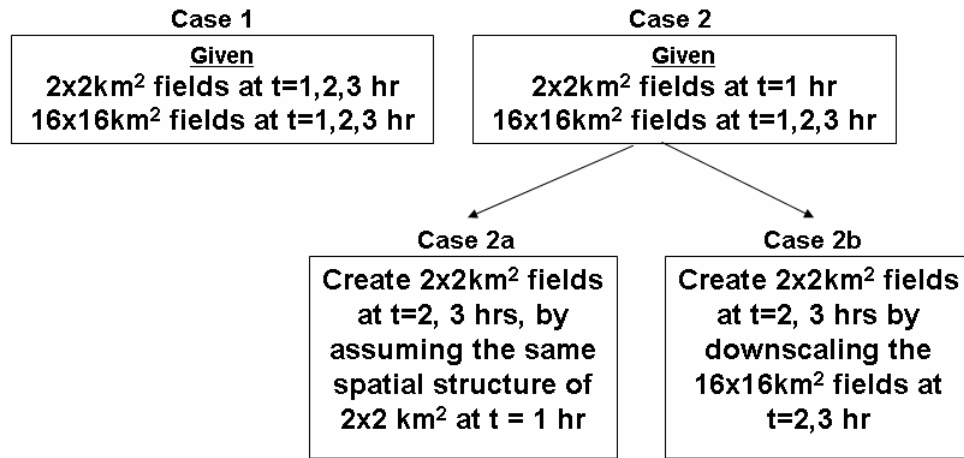


Figure 7: Schematic of the two scenarios considered to explore the merging of more frequent low resolution observations with less frequent high resolution observations.

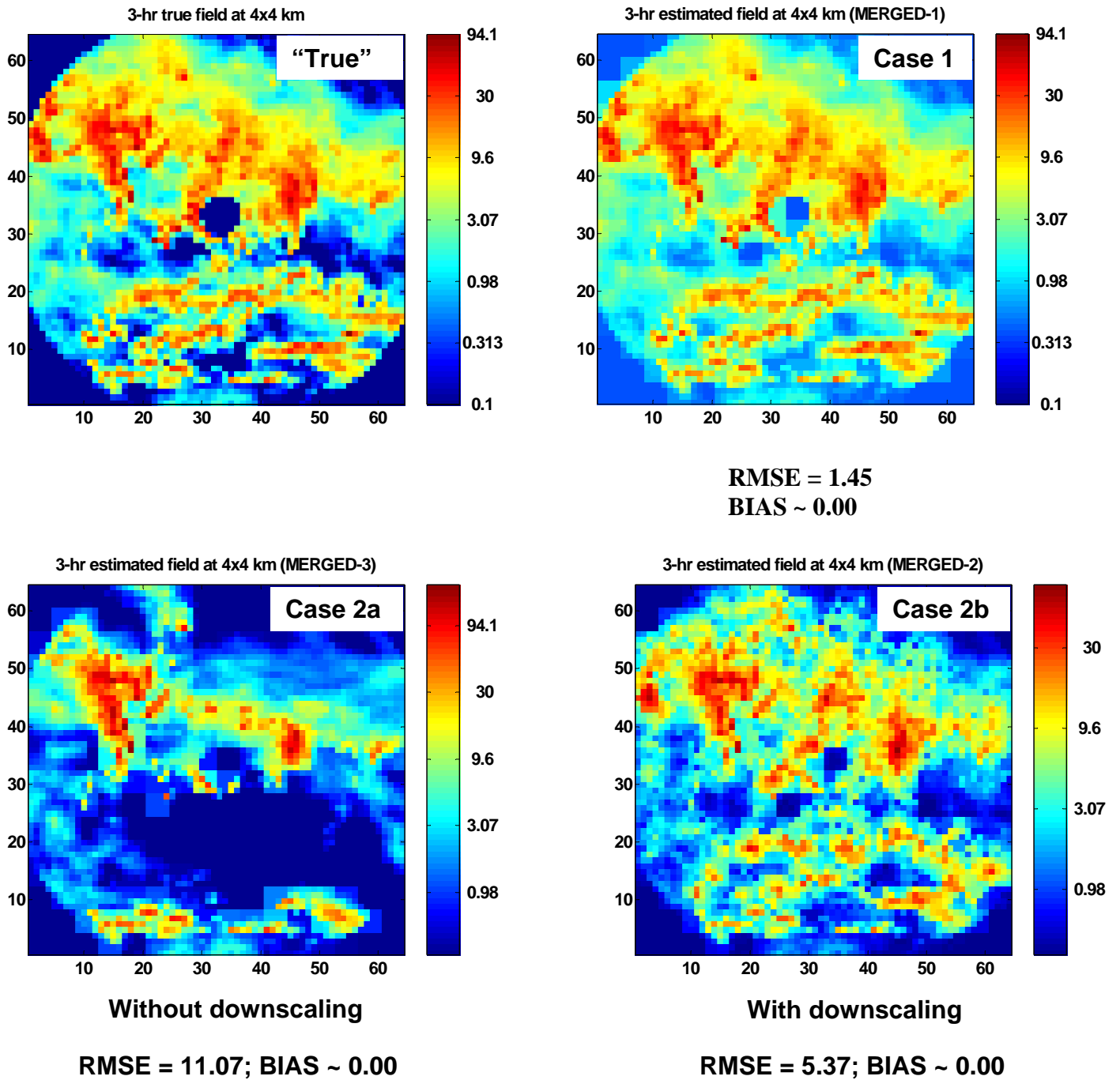


Figure 8: Illustration of the utility of using a spatial downscaling scheme in conjunction with the EM-SRE methodology for the purpose of merging infrequent high-resolution observations with more frequent low-resolution observations.