

# LINKING MESO-SCALE AND MICROSCALE MODELS: USING BLUP FOR DOWNSCALING

Katherine Campbell, Los Alamos National Laboratory  
Los Alamos, NM 87545

**KEY WORDS:** disaggregation, kriging, power-law variograms, precipitation

## Abstract

This paper describes the down-scaling link between the regional atmospheric model and the surface hydrology model currently implemented as part of the Los Alamos National Laboratory's Rio Grande Coupled Environmental Modeling Project. Downscaling has been extensively explored in recent years, usually in the context of Global Circulation Models, and approaches proposed in the literature are reviewed for usability in the context of coupled environmental modeling. The LANL project requires a method which can be conditioned on the gridded output of the meso-scale atmospheric model, as well as covariates available at smaller scales (at least topographic variables). Best linear unbiased prediction, or BLUP (essentially kriging) statistical models provide one well-developed method for making such spatially conditioned estimates. The adaptation, utility and drawbacks of this method for predicting across spatial scales are discussed.

## 1. Coupled Environmental Modeling

Linking numerical environmental models together to simulate the complete hydrogeologic cycle is a natural approach to studying the potential effects of global climate change on regional ecosystems. Well-developed models exist for global-scale and regional atmospheric processes, for surface hydrological processes including both overland and river flows, and for the subsurface component of the hydrologic cycle. When run independently, boundary conditions must be provided for each model. These conditions can take the form of arbitrary specifications (e.g., no-flow boundaries or sinks), input data (e.g., observed daily temperature ranges and precipitation), or a physical parameterization of a boundary process (e.g., re-radiation of energy from the surface of the earth.) The linked models, by contrast, might be run as a closed, or more nearly closed, system, in which at least some of the boundary conditions required by one model are provided by other models.

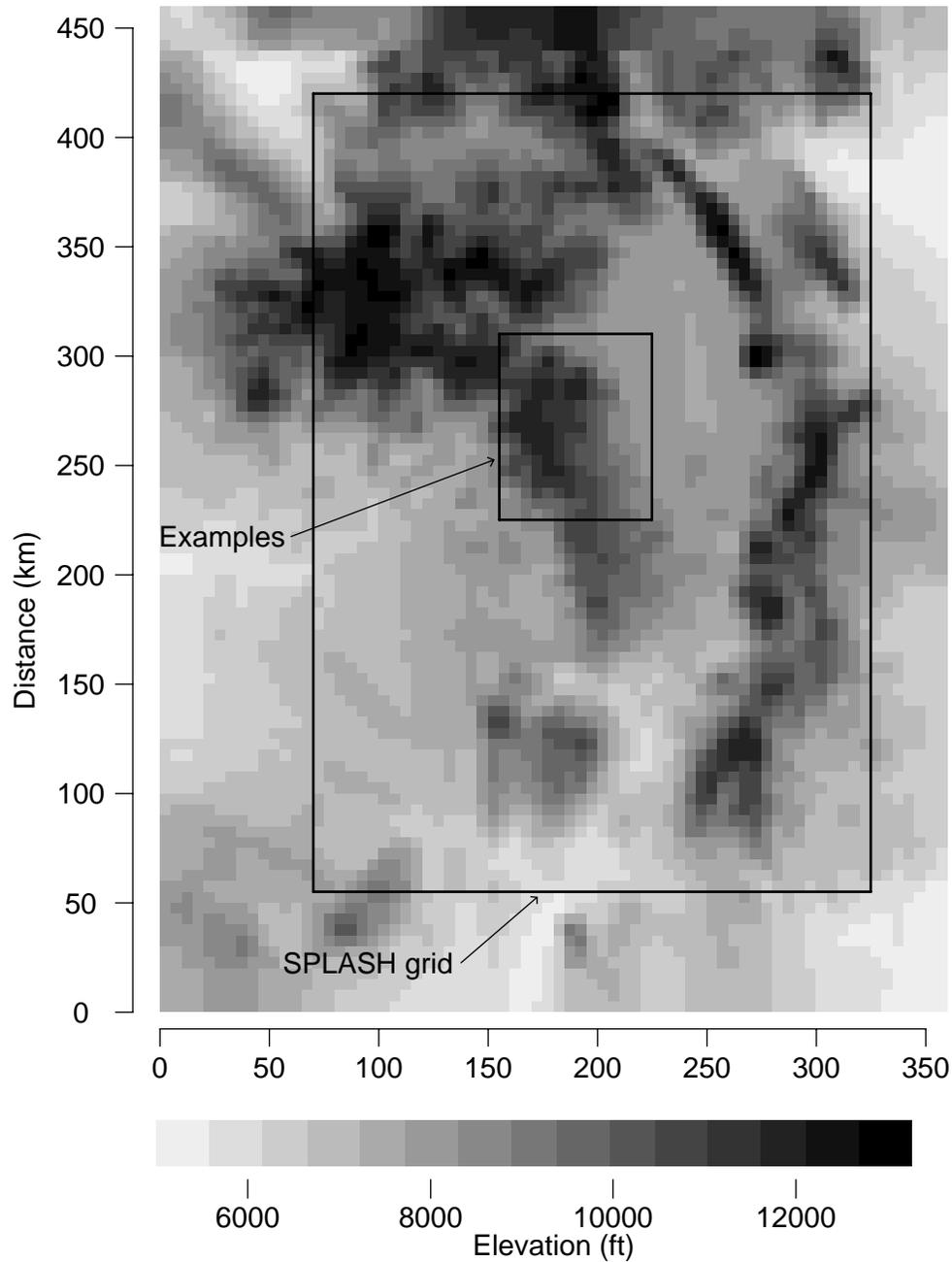
One problem which arises in implementing this program is the mismatch among both the time and space scales that are appropriate to each type of model. Global circulation models (GCMs) are run at spatial scales of 50 to 1000 km, and even the finest of these grids (which

already requires extensive computational resources) is too coarse to resolve the effects of topographic relief, land-use, and other significant surface features. A meso-scale atmospheric model is typically run with a spatial resolution of 5 to 100 kilometers. Again, the lower end of this range represents the current state of the art with the best computational resources available. Moreover, moving to smaller grid sizes entails not only corresponding increase in computational resources but also the development of new microphysics parameterizations for the atmospheric processes involved. Atmospheric processes are generally modeled with time steps on the order of one minute or less.

Coupling a meso-scale atmospheric model to a GCM has the potential to reduce the mismatch in scales between the atmospheric and hydrologic components of the modeling system. However, surface and near-surface hydrology, at least in a region of high topographic relief or highly variable vegetation cover, must be modeled at a spatial scale on the order of 50 to 2000 m, although significantly larger time steps (15 minutes to several hours) are adequate to describe the evolution of the important processes. Subsurface processes can generally be modeled at larger spatial scales than surface processes (although vertical discrimination can be important to capture thin but hydrologically significant stratigraphic layers), and at much slower time-scales.

In summary, even if well-established physical process models are selected, non-trivial rescaling interfaces must be constructed in order to link them into a truly coupled model. The Los Alamos National Laboratory's Rio Grande Coupled Environmental Modeling Project is attempting to link a meso-scale atmosphere model (the Regional Atmospheric Modeling System, or RAMS, Pielke et al. 1992) and a surface hydrology model (Simulator for Processes of Landscapes: Surface/subsurface Hydrology, or SPLASH, Martens et al., in preparation, and Xiao et al., 1996). The completed system will also integrate a subsurface model and an explicit river model. RAMS simulations require the use of two-way interactive, nested grids, of which the largest (80 km spacing) covers most of the western United States and parts of Canada and Mexico), while the smallest (5 km spacing) covers the upper Rio Grande in southern Colorado and northern New Mexico. Figure 1 shows the elevations in the

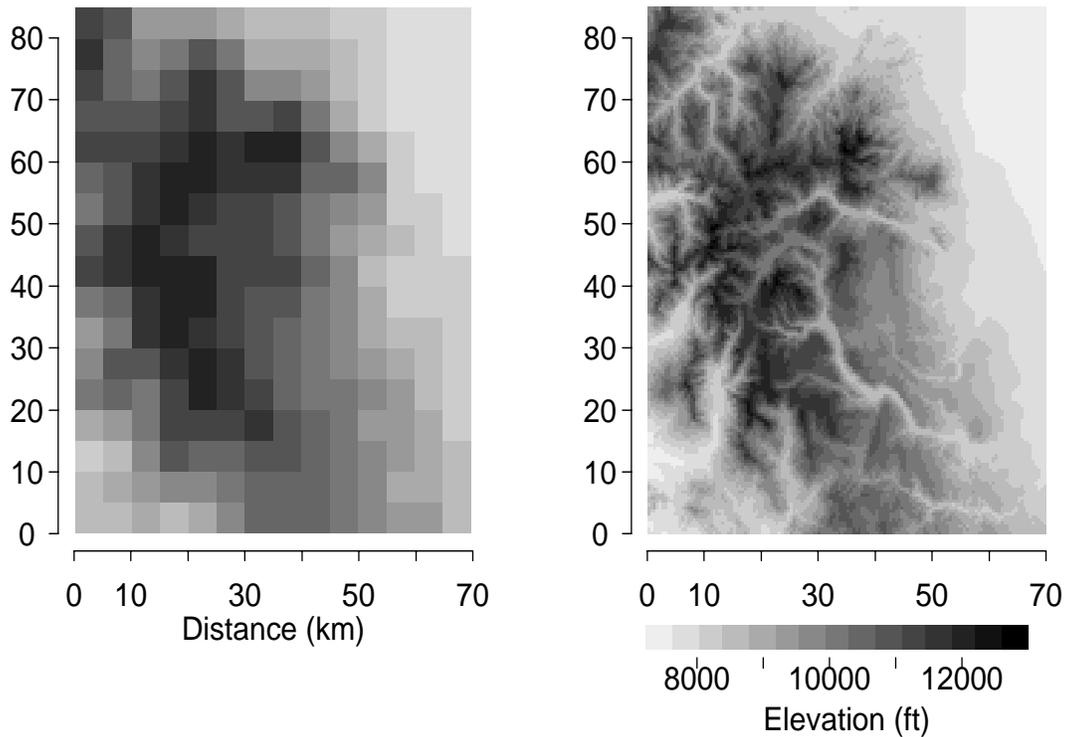
**Figure 1: Topography of the northern Rio Grande basin at 5-kilometer resolution**



bottom layer of the RAMS grid for the smallest grid. The Rio Grande heads in the San Juan mountains of Colorado in the northwest part of this scene (between 300 and 350 km north in the coordinate system shown) and flows east and then south into New Mexico, staying to the west of the north-south Sangre de Cristo range. The SPLASH subgrid encloses all of the area that drains

into the Rio Grande above Cochiti Dam, approximately 100 km north in this coordinate system. The examples later in this paper all come from the “Examples” region, which is expanded in Figure 2. Figure 2 also shows every fifth point of the SPLASH grid in this region, revealing significant detail in this mountainous terrain that is masked at the 5 km scale.

**Figure 2: Topography of the example area at resolutions of 5 km (RAMS) and 100 m (SPLASH)**



## 2. Approaches

The general downscaling problem, particularly the problem of predicting regional weather based on the output of GCMs, has received a great deal of attention in recent years. For a recent review, see Wilby and Wigley (1997). Deterministic methods include embedding limited area climate models within the larger global circulation models in order to move from GCM scales to scales at which surface effects such as topography and vegetation can usefully be incorporated, which is critical if the atmospheric model is to receive feedback from the surface model. Another deterministic approach uses simplified parameterization of the subgrid dynamics (e.g., Leung and Ghan 1998).

The much larger class of stochastic methods includes many variations, of which we review only a few. Regression methods develop linear or non-linear relationships between subgrid-scale parameters (generally single-site observations) and coarser-scale predictor variables. The latter are typically EOFs (empirical orthogonal functions) for the large-scale pressure fields, information which could also be generated by a GCM. For the disaggregation problem, another step would be necessary to distribute the output of the regression estimator from the estimated locations

across the gridded landscape. One method for this is the PRISM algorithm (Daly et al., 1994), which makes use of the local topography. A recent special issue of the *Journal of Geographic Information and Decision Analysis* (Dubois et al., 1998) contains a number of articles applying kriging, splines, adaptive kernel estimators and other spatial interpolation methods to rainfall and other meteorological variables.

Weather pattern approaches develop probability distributions for the subgrid-scale parameters conditioned on a classification of synoptic atmospheric information or “weather states”. Such classification schemes can be subjective or objective. For downscaling applications the classes must be based on or related to indices of atmospheric circulation such as pressure fields that can be produced by GCMs. These methods provide spatial simulations; the temporal component must be provided by the forcing sequence of weather states. In the interesting modification of Hughes et al. (1999), the forcing relationship is mediated by a hidden Markov model, which improves the ability of the model to reproduce wet- and dry-spell frequencies. Depending on what is described by the conditional probability distributions, further distribution

of simulated point values across a grid may be needed for the disaggregation problem.

The scaling models of Perica and Fofoula-Georgiou (1996) and Waymire, Over and Gupta (e.g., Gupta and Waymire 1993, Over and Gupta 1994) explicitly start with gridded data and develop conditional probability distributions from which to simulate on a smaller grid. The parameters of these models are scaling exponents (Perica and Fofoula-Georgiou) or multiscaling functions (Waymire et al.) and variances, which can be estimated from the gridded information. Perica and Fofoula-Georgiou show that they may also be parameterized as functions of the convective available potential energy, a measure of convective instability in the pre-storm environment, while Over and Gupta (1996) have examined the evolution of the parameters over time within the course of a storm.

Because they do not require external parameter estimation such as is needed for regression and weather pattern methods, these scaling methods seem more promising for the present application. One problem with them at present is that while simulations using these methods provide realistic images, the average over all simulations is simply the original large-scale information. In order to reproduce the effects of topography, which are pronounced in the northern part of the Rio Grande basin, it is essential to condition the output not only on the large-scale gridded data, but also on some spatially-distributed covariates. Topographic covariates including elevation, slope and aspect are available on much finer grids than the output of the meso-scale atmospheric model.

By far the best-developed algorithm for conditioning predictions simultaneously on observations of the field of interest and also on spatially-distributed covarying fields is kriging, which we here consider as a special form of best linear unbiased prediction or BLUP (cf. Campbell 1991). The remainder of this paper explores the possibilities of adapting this algorithm to the problem of statistical disaggregation in the present context.

### 3. BLUP (Kriging) for Downscaling

The probabilistic model for observations which underlies Best Linear Unbiased Prediction, or BLUP, is the mixed effects linear model

$$y = X\beta + Zu + e \quad (3-1)$$

where  $y$  is a vector of  $m$  observable random variables,  $X$  and  $Z$  are known matrices,  $\beta$  is a vector of  $p$  unknown parameters having fixed values ("fixed effects"),  $u$  and

$e$  are vectors of  $q$  and  $m$ , respectively, random variables ("random effects") such that  $E(u) = 0$ ,  $E(e) = 0$ , and

$$\text{Var} \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \sigma^2 \quad (3-2)$$

where  $G$  and  $R$  are known positive definite matrices and  $\sigma^2$  is a positive constant (which may or may not be known.) Given observations  $y$ , the BLUP procedure provides estimates  $\hat{\beta}$  of  $\beta$  and  $\hat{u}$  of  $u$  which are linear functions of the data  $y$ , unbiased in the sense that the average value of the estimate (with respect to the distribution of  $y$ ) is equal to the expected value of the quantities being estimated, and best in the sense of having minimum mean square error within the class of linear unbiased estimates. Explicit expressions for  $\hat{\beta}$  and  $\hat{u}$  are given by

$$\hat{\beta} = (X^T Q^{-1} X)^{-1} X^T Q^{-1} y \quad (3-3)$$

and

$$\begin{aligned} \hat{u} &= GZ^T Q^{-1} [I - X(X^T Q^{-1} X)^{-1} X^T Q^{-1}] y \\ &= GZ^T Q^{-1} (y - X\hat{\beta}) \end{aligned} \quad (3-4)$$

where

$$Q = R + ZGZ^T = E[(Zu + e)(Zu + e)^T]. \quad (3-5)$$

(Equation 3.5 shows that  $R$  and  $G$  need not be of full rank as long as  $Q$  is invertible. In particular, either may be zero. If  $G$  is zero, Eq. 3-3 reduces to the generalized least squares estimate for  $\beta$ .)

The universal kriging equations can be derived from the model in Eq. 3-1. Specifically, consider the problem of predicting the values of  $y^{(2)} = \{y(t_i)\}$  at  $n$  locations  $t_{m+1}, \dots, t_{m+n}$  given observations  $y^{(1)} = \{y(t_i)\}$  at  $m$  locations  $t_1, \dots, t_m$ . If  $N = m + n$ , replace  $X$  by  $SX$  and  $Z$  by  $SZ$  in the above equations, where  $S$  is the  $m \times N$  "sampling matrix" of the form  $\begin{bmatrix} I_m & 0_{m \times n} \end{bmatrix}$  ( $I_m$  is the  $m \times m$  identity matrix and  $0_{m \times n}$  denotes the  $m \times n$  matrix of zeros.) The columns of the  $N \times p$  matrix  $X$  may include a column of ones, one or more columns of coordinates (e.g., an x-coordinate, a y-coordinate, or elevation), or columns of other covariates. For normal kriging,  $Z = SI_N$ , and the  $N \times N$  matrix  $G$  is estimated as usual from the variogram.  $e$  is identified with the measurement errors in the  $m$  observations to complete the general kriging model, but frequently  $R=0$ . Specifically,

$$\hat{\beta} = (\mathbf{X}^{(1)\text{T}} \mathbf{Q}^{-1} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)\text{T}} \mathbf{Q}^{-1} \mathbf{y}^{(1)} \quad (3-6)$$

and

$$\hat{u} = \begin{bmatrix} \mathbf{G}_{11} \\ \mathbf{G}_{21} \end{bmatrix} \mathbf{Q}^{-1} (\mathbf{y}^{(1)} - \mathbf{X}^{(1)} \hat{\beta}) \quad (3-7)$$

for  $\mathbf{Q} = \mathbf{R} + \mathbf{G}_{11}$ , where  $\mathbf{X} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{S}\mathbf{X} \\ \mathbf{P}\mathbf{X} \end{bmatrix}$  ( $\mathbf{P}$  is

the  $n \times N$  matrix  $\begin{bmatrix} \mathbf{0}_{n \times m} & \mathbf{I}_n \end{bmatrix}$ ) partitions  $\mathbf{X}$  into the first  $m$  rows corresponding to the observations and the last  $n$  rows corresponding to the prediction points, and

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{S}\mathbf{G}\mathbf{S}^{\text{T}} & \mathbf{S}\mathbf{G}\mathbf{P}^{\text{T}} \\ \mathbf{P}\mathbf{G}\mathbf{S}^{\text{T}} & \mathbf{P}\mathbf{G}\mathbf{P}^{\text{T}} \end{bmatrix} \quad (3-8)$$

is the similar partitioning of  $\mathbf{G}$ . Finally, the kriged predictions are  $\hat{y}^{(2)} = \mathbf{X}^{(2)} \hat{\beta} + \mathbf{P} \hat{u} = \mathbf{A} \mathbf{y}^{(1)}$ . (Equations for the entries in the matrix  $\mathbf{A}$  can be obtained from Eqs. 3-6 and 3-7, and these are what are usually referred to as the “kriging equations”.)

For downscaling, we will consider taking  $\mathbf{Z}$  to be something more general than  $\mathbf{S}\mathbf{I}_N$ . For example, for a phenomenon  $y$  which is strongly elevation-dependent such as temperature, but possibly with spatially varying dependence on elevation, we may incorporate elevation both as a fixed effect and as a random effect. Specifically, to use elevation as a random effect, we take

$\mathbf{Z} = \mathbf{S} \begin{bmatrix} \mathbf{I}_N & \mathbf{A} \end{bmatrix}$ , where  $\mathbf{A}$  is a diagonal matrix of elevations, and partition the  $2N$ -vector of random effects as  $\begin{bmatrix} \mathbf{u} \\ \mathbf{w} \end{bmatrix}$ , where the random effect  $w$  corresponds

to the spatially correlated coefficients of elevation. If  $\mathbf{u}$  and  $\mathbf{w}$  are uncorrelated and the covariance matrix of

$\begin{bmatrix} \mathbf{u} \\ \mathbf{w} \end{bmatrix}$  is partitioned as  $\begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}$ , then the matrix  $\mathbf{Q}$  defined

in Eq. (3.5) becomes  $\mathbf{Q} = \mathbf{R} + \mathbf{S}\mathbf{G}\mathbf{S}^{\text{T}} + \mathbf{S}\mathbf{A}\mathbf{F}\mathbf{A}^{\text{T}}\mathbf{S}^{\text{T}}$ . Then the  $n$  predicted values  $y(t_i)$  are computed as  $\hat{y}^{(2)} = \mathbf{X}^{(2)} \hat{\beta} + \hat{u}^{(2)} + \mathbf{A}^{(2)} \hat{w}^{(2)}$ , where again (cf. Eqs. 3-6 and 3-7)

$$\hat{\beta} = (\mathbf{X}^{(1)\text{T}} \mathbf{Q}^{-1} \mathbf{X}^{(1)})^{-1} \mathbf{X}^{(1)\text{T}} \mathbf{Q}^{-1} \mathbf{y}^{(1)}, \quad (3-9)$$

$$\hat{u}^{(2)} = \mathbf{G}_{21} \mathbf{Q}^{-1} (\mathbf{y}^{(1)} - \mathbf{X}^{(1)} \hat{\beta}), \quad (3-10)$$

and now

$$\hat{w}^{(2)} = \mathbf{F}_{21} \mathbf{A}^{(1)} \mathbf{Q}^{-1} (\mathbf{y}^{(1)} - \mathbf{X}^{(1)} \hat{\beta}), \quad (3-11)$$

$\mathbf{A} = \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{(2)} \end{bmatrix}$  being the obvious partition of  $\mathbf{A}$ .

Computationally, this is reasonably efficient. If the downscaled values in a small grid are conditioned on a  $3 \times 3$  or  $5 \times 5$  neighborhood of large-scale gridded values, then the dimension of the matrix  $\mathbf{Q}$  is 9 or 25. In the case where the only random effect is  $u$ , this inversion only needs to be done once, but in general it will need to be done for each large-scale grid. If storage is less important than computational time, these inversions can be performed once and stored, provided that the matrices  $\mathbf{G}$ ,  $\mathbf{F}$  and  $\mathbf{A}$  are not functions of time. (In the current implementation,  $\mathbf{A}$  is a function of time only for precipitation, and  $\mathbf{S}$  and  $\mathbf{F}$  are fixed for all variables.)

#### 4. Variogram models for downscaling

Turning next to the problem of estimating  $\mathbf{S}$  and  $\mathbf{F}$ , note that the  $m$  observations in the present context consist of observations on the large-scale grid; their “supports” are five-kilometer squares. The  $n$  values to be predicted have as supports much smaller squares, 100 m on a side. Thus we are dealing here not with a point process  $y(t)$  but with averages over supports of different dimensions, i.e. processes  $y_V(t)$  and  $y_v(t)$  which are averages of a point process  $y(t)$  over 5-km and 100m supports  $V$  and  $v$ , respectively. If we knew the autocovariance function for the underlying point process, we could compute the variograms of the averaged processes. However, the RAMS output provides only “observations” of the larger averaged process  $y_V(t)$ . In this case it is useful, if not essential, to consider variogram models that reflect some built-in assumptions about scaling, so that we can get from the variogram of the  $y_V$  process to the cross-variogram of the  $y_V$  and  $y_v$  processes, which is needed to calculate  $\mathbf{G}_{21}$  and  $\mathbf{F}_{21}$ .

The easiest variogram to scale with changing support are those associated with a simple scaling point process. A point process  $y(t)$  is “simple scaling” if for each  $\lambda > 0$  there is a multiplier  $C_\lambda$  such that the joint

distribution of  $\{y(t_1), \dots, y(t_k)\}$  is the same as that of

$$\left\{ C_\lambda^{-1} y(\lambda t_1), \dots, C_\lambda^{-1} y(\lambda t_k) \right\}.$$

It is straightforward to show that, as a function of  $\lambda$ ,  $C_\lambda$  must be of the form  $\lambda^H$ . Averages of such processes scale similarly;  $y_{(\lambda)}(\lambda t)$  is distributed like  $\lambda^H y_{(1)}(t)$ , where (in  $d$  dimensions)

$$y_{(\lambda)}(t) = \frac{1}{\lambda^d} \int f\left(\frac{t-s}{\lambda}\right) y(s) ds \quad (4-12)$$

for some kernel function  $f$  (in this paper,  $f$  is a simple “boxcar” function.) Finally, the variograms of such averages must be related by

$$\frac{\gamma_{(\lambda)}(\lambda h)}{\gamma_{(\kappa)}(\kappa h)} = \left(\frac{\lambda}{\kappa}\right)^{2H}. \quad (4-13)$$

The variogram  $\gamma(h)$  of simple scaling point process must be proportional to  $s^{2H}$ , which is of course unbounded, so that the covariance of the process is not defined. One modification to this simple model that has been proposed in the literature takes explicit account of the fact that both the support of the measurements and the domain in which the process is realized are finite to derive bounded variogram models together with explicit rules for scaling these models as the support and domain are changed. Specifically, Di Federico and Neuman (1997) observed that a power law variogram of the form

$$\gamma(s) = \frac{1}{2} [Y(x+s) - Y(x)]^2 = C_0 s^{2H} \quad (4-14)$$

can be written as a weighted sum

$$\gamma(s) = \int_0^\infty \gamma(s, n) \frac{dn}{n} \quad (4-15)$$

in two different ways. For  $0 < H < 1/2$ ,  $\gamma(s, n)$  can be chosen to be an exponential variogram,

$$\gamma(s, n) = \sigma^2(n) [1 - \exp(-sn)] \quad (4-16)$$

where  $n^{2H} \sigma^2(n) = 2HC_0/\Gamma(1-2H)$ . Alternatively, for  $0 < H < 1$ ,  $\gamma(s, n)$  can also be a Gaussian variogram,

$$\gamma(s, n) = \sigma^2(n) \left[ 1 - \exp\left(-\frac{\pi n^2 s^2}{4}\right) \right], \quad (4-17)$$

where now  $n^{2H} \sigma^2(n) = \left(\frac{4}{\pi}\right)^H 2HC_0/\Gamma(1-H)$ . Truncated forms of Eq. (4-13), where the limits of inte-

gration are taken to be  $n_l$  and  $n_u$  with  $0 \leq n_l < n_u \leq \infty$ , provide bounded variogram models. Di Federico and Neuman propose that the finite cutoffs  $n_l$  and  $n_u$  be related to the linear dimensions of the domain,  $L$ , and the sample support,  $l$ , respectively, by means of simple proportionality:  $n_u = 1/(\mu l)$  and a lower cutoff

$n_l = 1/(\mu L)$ , resulting in three-parameter variogram models ( $\mu$ ,  $H$  and  $C_0$ , assuming that  $l$  and  $L$  are known) for the exponential and Gaussian cases. In particular, this choice for  $n_u$  results in a variogram which scales in  $n_u$ , i.e., in the support, as required by Eq. (4-12). Analytical expressions for these models, which involve the incomplete gamma function, are provided in their paper. In application we have introduced a fourth parameter  $b$  relating the unknown domain size  $L$  to the sample support  $l$ :  $L = bl$ .

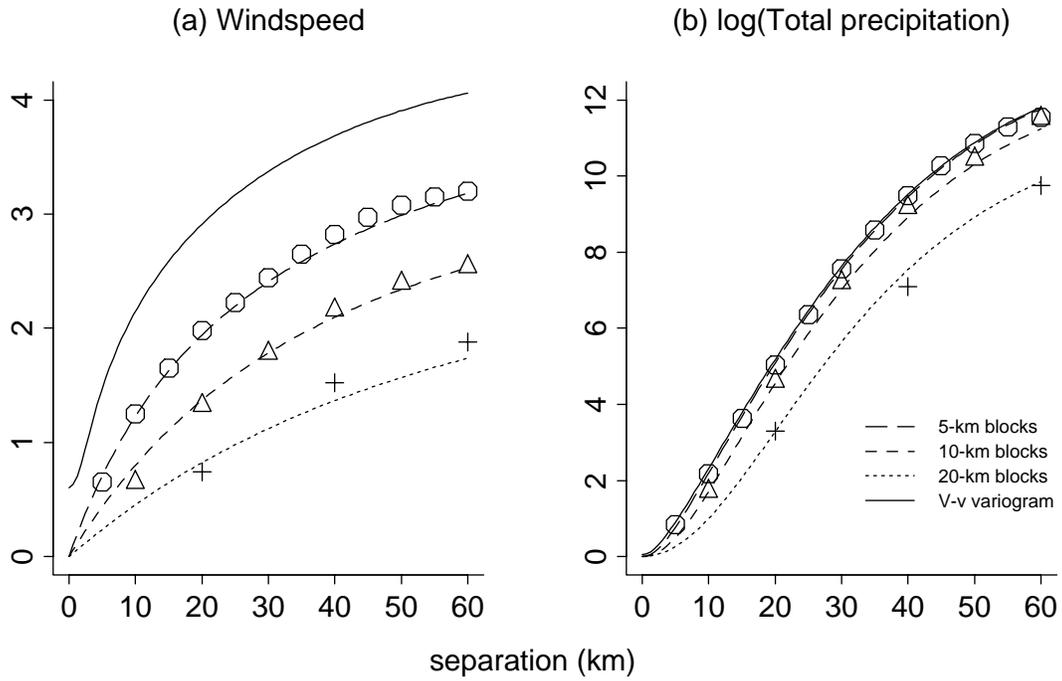
Typically we find  $\mu$  is about 1 and  $b$  is on the order of 30. Given empirical variograms computed not only on the 5-km gridded output from RAMS but also 2 x 2 and 4 x 4 averages of this output (i.e., 10- and 20-km grids),  $h$  can be estimated independently of the other parameters using Eq. (4-12), and given  $H$ ,  $\mu$  and  $b$ ,  $C_0$

is a function of the “sill”  $\sigma^2(n_l) - \sigma^2(n_u)$ . Thus fitting the four parameters of the covariance model by a standard minimization procedure is in theory a reasonably stable procedure and might be implemented or at least updated at each time step. In practice, fitting this model seems to be as much of an art as fitting other popular variogram models. Some examples are shown in Figure 3. To date, therefore, we have fixed these parameters for an entire run, but investigation of their dependence on large-scale parameters (such as noted by Perica and Foufoula-Georgiou) is a possibility for future investigation.

## 5. Results

The equations of Sections 3 and 4 have been implemented in a FORTRAN program to downscale eight variables, as shown in Table 1. Six of these fields are provided as 20-minute averages by the RAMS calculations, and are downscaled and passed on to SPLASH with the same frequency. Precipitation (as liquid and solid precipitation) is provided at 2-minute intervals but downscaled only when total precipitation “tips a bucket” at 1 mm accumulation in some RAMS cell (but not less frequently than once every 20 minutes.) Liquid and solid precipitation are summed to a total mm equivalent and converted to the log scale for downscaling; zeros in the RAMS grid are set equal to a small positive value for this purpose and the downscaled

**Figure 3: Some variogram models (di Frederico and Neuman)**



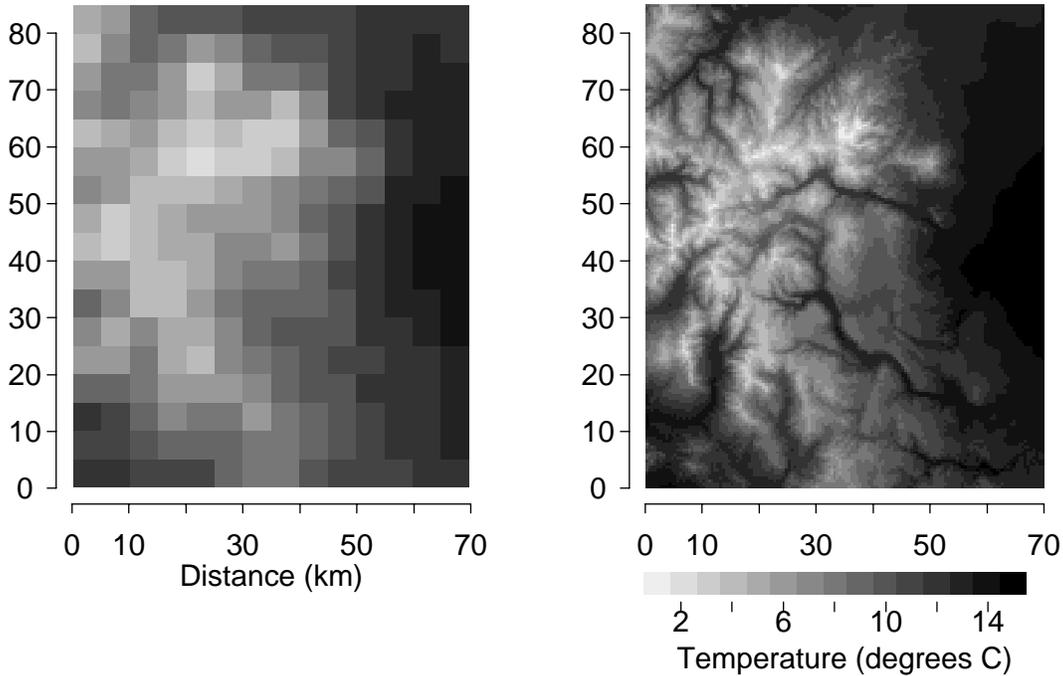
**Table 1: Downscaled Variables**

Variable	Transformation	X	Z
Pressure (mb)	--	$[1, h]$ ~ ~	$[I, [h_{ii}]]$
Temperature (°C)	--	$[1, h]$ ~ ~	$[I, [h_{ii}]]$
Mixing ratio (g/g)	--	1 ~	I
Shortwave radiation (Watts/m <sup>2</sup> )	--	1 ~	I
Longwave radiation (Watts/m <sup>2</sup> )	--	1 ~	I
Windspeed (m/sec)	--	1 ~	I
Total precipitation (mm)	log	--	I
Fraction snow	logit	$[1, h]$ ~ ~	I

values are later truncated back to zero. The fraction that is snow is computed, and if there is a mixture of rain and snow in the scene, then this fraction is converted to logits and downscaled. SPLASH accepts either rain or snow, but not both, in each of its cells, so the downscaled logits are simply transformed back to ones or zeros depending on whether they are positive or negative.

As Table 1 shows, all variables are downscaled with at least one fixed effect (a constant) except for total precipitation, which is discussed below. Elevation is a second fixed effect for pressure, temperature, longwave radiation and the fraction snow. Elevation is also used to augment the Z matrix for pressure and temperature; the bottom two layers of RAMS data provide information to estimate the covariance matrix F for a second random effect  $w$ . The ratio of the difference between these bottom two layers and the corresponding differences in elevation is also used to bias the corresponding coefficient in the  $\beta$  vector; in particular, the sign of the coefficient is constrained to match the sign suggested by the RAMS data. The BLUP calculations for  $\beta$ ,  $\hat{u}$  and (if present)  $\hat{w}$  are performed on a subgrid of the 50 x 50 SPLASH cells within each 5-km RAMS cells, based on RAMS data in a 25 x 25 km neighborhood of that cell, and then further distributed to

**Figure 4: Downscaling temperature (highly elevation dependent)**



the remainder of the SPLASH cells by simple interpolation, using detailed SPLASH elevation data to multiple  $\hat{\beta}$  and  $\hat{w}$ , as appropriate. Where elevation is an important covariate, the result closely follows elevation, as in the case of temperature (see Figure 4.) It may also be noted in Figure 4 that the overall SPLASH scene is slightly warmer than the RAMS scene; this is because the elevations at which temperature is reported in the lowest layer of the RAMS model are on the average about 150 m above the actual land surface.

For a variable like shortwave radiation with no non-constant covariates, the estimate is essentially a smooth of the RAMS data. However, even in this case the fact that the covariance model for the off-diagonal blocks of the G matrix (corresponding to the cross-covariance between the two scales) is not the same as the covariance model for the first on-diagonal block (the autocovariance at the larger scale) allows predictions on the small-scale grid that are slightly outside the range of the large-scale data.

One thing which BLUP does not do naturally is to conserve total precipitation within the area covered by a RAMS cell, which will be desirable especially when the feedback loop to RAMS is implemented. Conservation can be approximated however by (a) forcing the mean in the log scale to be equal to the RAMS mean, and further

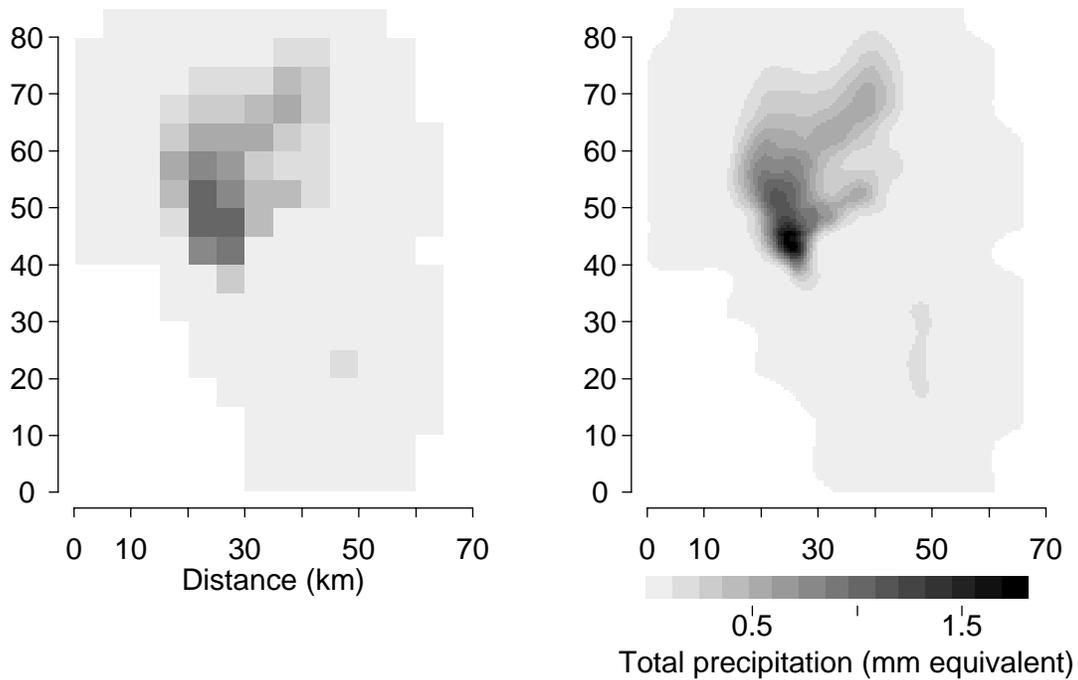
(b) accounting for the well-known bias, a function of the logarithmic variance, which arises when transforming back to the linear scale from the log scale. Thus we do not allow the algorithm to estimate even a constant term in this case, but set  $X=0$  and modify the results in each RAMS grid as suggested above. In the scene used for the examples in this paper, where the maximum total precipitation is just above 1 mm, the resulting absolute error in conservation do not exceed 0.015 mm, and the relative errors are less than 10% except for very small values of total precipitation, and less than 2% where the RAMS precipitation exceeds 0.2 mm. This is in spite of the fact that the maximum value on the SPLASH grid is almost 70% larger than the maximum value on the RAMS grid (Figure 5).

Other comparisons between the RAMS and downscaled SPLASH grids for this example are shown in Table 2. In particular, the downscaled scene has more rain and less snow than the RAMS scene (which includes both solid and liquid precipitation in several cells), because the algorithm is putting more rain into some narrow but deep valleys that are not resolved by the 5-km RAMS grid.

## 6. Conclusions

Although not an intrinsically scaling procedure, BLUP can be adapted to the downscaling problem by

**Figure 5: Downscaling precipitation**



**Table 2: Comparison of RAMS and Downscaled Precipitation**

Indicator	RAMS	SPLASH
Fraction of scene with precipitation	73.5%	75.1%
Maximum precipitation in one cell	1.07 mm	1.78 mm
Fraction of precipitation that is snow	68.3%	64.0%

making some assumptions (similar to those used in explicitly scaling methods) about the behavior of the covariance structure as scales change. Unlike cascade or wavelet methods, estimation of the variogram is the only place that cross-scale information can be used, however. The main advantage of BLUP is its ability to incorporate spatially distributed covariates such as elevation, about which there may be information at finer scales than provided by the atmospheric model. To the extent that such covariates are important determinants of a meteorological variable (as for pressure and temperature) BLUP is probably the best approach. We have had moderate success as well applying it to precipitation, although here it is probably not the best

possible model. Methods which make more explicit use of cross-scale information are intuitively more appealing for simulation of precipitation, and might be even more useful if some way to condition them on topographic variables such as slope and aspect could be developed.

## References

- K. Campbell (1991). Discussion of "That BLUP is a Good Thing --The Estimation of Random Effects", by G. K. Robinson. *Statistical Science* **6**, pp. 32-34, February 1991.
- C. Daly, R. P. Neilson and D. L. Phillips (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Applied Meteorology* **33**, pp. 140-158.
- V. Di Federico and S. P. Neuman (1997). Scaling of random fields by means of truncated power variograms and associated spectra. *Water Resources Research* **33**, pp. 1075-1088.
- G. Dubois, J. Malczewski and M. de Cort (editors, 1998). *Spatial Interpolation Comparison 97*. Special Issue of the *Journal of Geographic Information and Decision Analysis* **2**(2). On the Web at [http://publish.uwo.ca/~jmalczew/gida\\_4.htm](http://publish.uwo.ca/~jmalczew/gida_4.htm).

- V. K. Gupta and E. Waymire (1993). A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. *J. Applied Meteorology* **32**, pp. 251-267.
- J. P. Hughes, P. Guttorp and S. P. Charles (1999). A non-homogeneous hidden Markov model for precipitation occurrence. *J. Royal Statistical Society C* **48**, pp. 15-30.
- L. R. Leung and S. J. Ghan (1998). Parameterizing Sub-grid Orographic Precipitation and Surface Cover in Climate Models. *Monthly Weather Review* **126**, pp. 3271-3291.
- S. N. Martens, S. L. Ustin, and Wallender, W.W. (in preparation.) SPLASH: Simulator for Processes of Landscapes: Surface/Subsurface Hydrology.
- T. M. Over and V. K. Gupta (1994). Statistical Analysis of Mesoscale Rainfall: Dependence Random Cascade Generator on Large-Scale Forcing. *J. Applied Meteorology* **33**, pp. 1526-1542.
- T. M. Over and V. K. Gupta (1996). A space-time theory of mesoscale rainfall using random cascades. *J. Geophysical Research* **101**, pp. 26,319-26,331.
- S. Perica and E. Foufoula-Georgiou (1996). Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *J. Geophysical Research* **101**, pp. 26,347-26,361.
- R. A. Pielke, W. R. Cotton, R. L. Walko, C. J. Tremback, M. E. Nicholls, M. D. Moran, D. A. Wesley, T. J. Lee and J. H. Copeland (1992). A comprehensive meteorological modeling system – RAMS. *Meteorology and Atmospheric Physics* **49**, pp. 69-91.
- R. L. Wilby and T. M. L. Wigley (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography* **21**, pp. 530-548.
- Q. F. Xiao, S. L. Ustin, and W. W. Wallender (1996). A spatial and temporal continuous surface-subsurface hydrologic model. *J. Geophysical Research* **101**, pp. 29,565-29,584.