A Simplified Diagnostic Model of Orographic Rainfall for Enhancing Satellite-Based Rainfall Estimates in Data-Poor Regions

CHRIS FUNK AND JOEL MICHAELSEN

Geography Department, University of California, Santa Barbara, Santa Barbara, California

(Manuscript received 26 July 2003, in final form 28 April 2004)

ABSTRACT

An extension of Sinclair's diagnostic model of orographic precipitation ("VDEL") is developed for use in data-poor regions to enhance rainfall estimates. This extension (VDELB) combines a 2D linearized internal gravity wave calculation with the dot product of the terrain gradient and surface wind to approximate terrain-induced vertical velocity profiles. Slope, wind speed, and stability determine the velocity profile, with either sinusoidal or vertically decaying (evanescent) solutions possible. These velocity profiles replace the parameterized functions in the original VDEL, creating VDELB, a diagnostic accounting for buoyancy effects. A further extension (VDELB*) uses an on/off constraint derived from reanalysis precipitation fields. A validation study over 365 days in the Pacific Northwest suggests that VDELB* can best capture seasonal and geographic variations. A new statistical data-fusion technique is presented and is used to combine VDELB*, reanalysis, and satellite rainfall estimates in southern Africa. The technique, matched filter regression (MFR), sets the variance of the predictors equal to their squared correlation with observed gauge data and predicts rainfall based on the first principal component of the combined data. In the test presented here, mean absolute errors from the MFR technique were 35% lower than the satellite estimates alone. VDELB assumes a linear solution to the wave equations and a Boussinesq atmosphere, and it may give unrealistic responses under extreme conditions. Nevertheless, the results presented here suggest that diagnostic models, driven by reanalysis data, can be used to improve satellite rainfall estimates in data-sparse regions.

1. Introduction

This research addresses the need for a computationally light model of orographic precipitation to support rainfall estimation in food-insecure regions. Many drought-prone regions have complicated terrain and limited rainfall data. Afghanistan, Ethiopia, Central America, and southern Africa have seen recent famines that affected millions of people. Thus, we often know the least about precipitation in many of the regions where it has the most direct impact on livelihoods. Because precipitation and yields are tightly linked in these drought-prone regions (Fig. 1), improvements in rainfall estimation will extend our capacity to monitor, anticipate, and mitigate famine. In this paper we describe and validate a diagnostic rainfall estimation procedure, derived from Sinclair's (1994, hereinafter S94) "VDEL" formulation, extended with a vertical-velocity calculation based on linearized 2D internal gravity waves (IGW).

The extended diagnostic model (VDELB) has been used to represent orographic influences in the 36-yr (1961–96) Collaborative Historical African Rainfall Model (CHARM) precipitation climatological description (Funk et al. 2003). CHARM precipitation has mean absolute errors of about 0.5 standard deviation. Seasonal totals correspond fairly well to independent agroclimatic indicators such as national-level maize production statistics (Fig. 1). This paper documents more fully the VDELB estimation process and demonstrates its application to satellite rainfall estimation (RFE) in southern Africa. Enhanced RFE fields can provide earlier identification of potentially food-insecure regions.

The problem may be framed by comparing average precipitation from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996) with high-quality interpolated gauge data from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) dataset (Daly et al. 1994) (Fig. 2). In the ideal case, a model could translate coarse-scale atmospheric data into precipitation fields that more closely resemble the true spatial–temporal distribution.

Given sufficient resources, an embedded regional-scale or orographic model, such as the Clark–Hall (Clark and Hall 1991; Clark et al. 1996), the Center for Analysis and Prediction of Storms Advanced Regional Prediction System (Xue et al. 2000, 2001), the Regional
Fig. 1. Scatterplot of relative national Zimbabwe production (1990/91–1995/96 production years) and seasonal (Sep–Mar) CHARM precipitation in the North and South Matabeleland, Masvingo, and Mashonaland crop-growing regions in Zimbabwe (shaded in map). The maize production index was created by 1) screening out agricultural units that reported abnormally low values in any season between 1980/81 and 1999/2000, 2) totaling the remaining units, and 3) dividing each season’s total by the average of all the seasons to create an anomaly index. Production statistics were kindly provided by the Regional Remote Sensing Unit in Harare.

Fig. 2. Shading shows annual PRISM precipitation (Daly et al. 1994). Contours delineate isohyets from the NCEP-NCAR reanalysis.
\[ w_z = u_z \cdot \nabla z. \]  
\[ (1) \]

The orographic vertical velocity at pressure level \( p \) \([w_z(p)]\) is then estimated by a function based on the pressure at the surface and top of the precipitating layer and a user-defined parameter \( \gamma \):

\[ w_z(p) = w \left( \frac{p - p_s}{p_t - p_s} \right)^{\tan(\gamma/4)}. \]  
\[ (2) \]

We have used 300 hPa as the top of the precipitating layer. Tests in the Pacific Northwest over 365 days (1983) suggested that the VDEL formulation had difficulty replicating rainfall across the full geographic/seasonal domain for any given choice of \( \gamma \). While it might be possible to vary \( \gamma \) in space and time, improving model performance, we instead replace Eq. (2) with a function derived from linearized gravity waves (Bradley 1985; Cotton and Anthes 1989; Durran 1990; Bruintjes et al. 1994). This step removes the need to set \( \gamma \) explicitly.

3. Method

a. Linearized internal gravity waves—Background

Queney (1948), Scorer (1955), and Eliassen and Palm (1960) described linear gravity wave responses in stratified flows over orography. Early research by Smith (1979, 1982) looked at orographic effects in a quasigeostrophic framework. Carruthers and Choularton (1982) developed a model based on the linearized vorticity equation. Hunt et al. (1988) and Carruthers and Hunt (1990) expanded this approach, which generates three unique solutions for a lower, middle, and upper layer of the atmosphere. Clark and Peltier (1977, 1984), Peltier and Clark (1979), and Clark and Farley (1984) noted the importance of wave-breaking effects. Durran (1986) used numerical simulations to examine amplification mechanisms in large-amplitude mountain waves, demonstrating that when the parameter \((N \times \text{mountain height/wind speed})\) is less than 0.3 and the static stability has a two-layer structure, the nonlinear response can differ significantly from the linear solution. When nonlinear waves are much larger than the linear solutions, the highest stability is found in the lowest layer, the flow resembles a hydraulic jump (Long 1935), and wave breaking may occur (Durran 1986, 1990, 1992).

Although nonlinear effects are important, a recent study by Dornbrack and Nappo (1997) compared linear and nonlinear models over a range of realistic stabilities and found essentially similar results. In the real atmosphere waves interact with each other, turbulence, and the mean flow in complicated ways, yet, despite these difficulties, linear theory is still useful as a first-order estimation procedure that is easy to understand (Nappo 2002). This property makes it a reasonable foundation for diagnostic applications.

b. Linearized internal gravity waves—Description

Consider a two-dimensional Boussinesq atmosphere in which the perturbations of the atmospheric state variables are independent of and small in comparison with their background state. Under these conditions, the linearized momentum, mass, and thermodynamic conservation equations may be combined to provide a single expression for the terrain-induced vertical-velocity perturbations. If air with a mean wind speed \( \mathbf{u} \) flows across a set of sinusoidal ridges, with heights determined by the horizontal wavenumber \( k \), such that \( h(x) = h_m \cos(kx) \), then the terrain-induced vertical-velocity perturbations \((w')\) will depend only on \( x \) and \( z \), and the Taylor–Goldstein wave equation (Taylor 1931; Goldstein 1931; Nappo 2002) may be written as (Durran 1990, p. 62; Holton 1992, p. 204)

\[
\left( \frac{\partial^2 w'}{\partial x^2} + \frac{\partial^2 w'}{\partial z^2} \right) + \frac{N^2}{\mathbf{u}^2} w' = 0. \]  
\[ (3) \]

The solutions to these equations (Durran 1990) may be written as

\[
w = \Re\{Ae^{ikx+mi} + Be^{ikx-mi}\},\]  
\[ (4) \]

where \( A \) and \( B \) are complex coefficients and \( \Re \) denotes the real part of the solution. The vertical structure of these waves will be determined by the dispersion relationship

\[
m^2 = \left( \frac{N^2}{\mathbf{u}^2} \right) - k^2, \]  
\[ (5) \]

where \( m \) and \( k \) are the vertical and horizontal wave-numbers. Because the flow at the boundary must be parallel to the surface, the following solutions may be derived (Durran 1990):

if \( (\pi k > N) \),

\[
w_x(x, z) = -\pi h_m k \exp(-\mu z) \sin(kx) \quad \text{and} \]

if \( (\pi k < N) \),

\[
w_x(x, z) = -\pi h_m k \sin(kx + mz), \]  
\[ (6) \]

where \( m \) is calculated according to Eq. (5), and \( \mu \) is calculated as

\[
\mu = (k^2 - N^2/\mathbf{u}^2)^{1/2}. \]  
\[ (7) \]

These correspond to the evanescent and sinusoidal forms of the vertically trapped waves.

This solution is limited in several important ways. The two-dimensional representation ignores important effects such as flow splitting and blocking. We assume that the perturbations in the state fields are small and that the linearized equations are adequate. This assumption may be violated, especially when dealing with large topographic obstacles, and is related to several significant weather phenomena, such as clear-air turbulence and trapped lee waves.
Three-dimensional linear perturbation theory (Smith 1980) suggests that unrealistic results are obtained when $hN\tau^{-1} > 1$, where $h$ is the height of the mountain; the energy produced by gravity waves under these conditions will exceed the kinetic energy incident on the mountain, causing the flow to avoid the obstacle. If $N = 0.01$ s$^{-1}$ and $\overline{\tau} = 10$ m s$^{-1}$, then this criteria is exceeded when $h > 1000$ m. Our model implementation will tend to overestimate in these cases. Smith’s (1980) three-dimensional formulation also predicts the tendency of the low-level flow to split around mountains, with the velocity perturbations spreading laterally to the windward and leeward sides. These aspects of flow over and around complex terrain are not captured in the two-dimensional VDELB formulation.

c. Combining 2D linearized IGWs with the VDELB diagnostic model

The $\bar{u}_y \cdot \nabla z$ term in Eq. (1) and the $-\overline{\tau}h_{k}k$ term in Eq. (6) determine the relative magnitude of the vertical-velocity response. If we assume that $\overline{\tau} = u_s$, then we can rewrite the original diagnostic equation for the orographically induced vertical motion [Eq. (1)]:

$$u_s(x, z) = w_s \exp(-\mu z) \sin(kz)$$

if $(u, k > N)$,

$$w_s(x, z) = w_s \sin(kx + mz).$$

Although this approximation is not exact, it does allow us to incorporate some of the attributes of IGWs into VDELB. The atmospheric response will depend on whether $Nu_s$ is less than $k$. We have used NCEP–NCAR reanalysis wind fields at the local terrain height to represent $u_s$.

Because $k$ varies inversely with the characteristic ridge width $L_z$, narrow ridges will tend to be associated with vertically decaying displacement fields. Positive values of $m$ [Eq. (5)], associated with broad terrain features and moderate wind speeds, propagate energy vertically in the atmosphere and can produce clouds upstream or downstream from a terrain feature. Typical values for $u_s$ and $N$ in the lower troposphere are about $10^3$ m s$^{-1}$ and $10^{-2}$ s$^{-1}$, respectively. Thus, for typical ridge widths with $L_z > 6000$ m, or values of $k$ less than $10^{-3}$ s$^{-1}$, vertically propagating internal gravity waves will dominate.

For the wide mountain limit, Queney (1948) found that for a simple 2D idealized mountain the terrain profile is reproduced at every integral of $2\pi(\overline{\tau}/N)$. The profiles of vertical motion associated with this wide mountain case will vary as a function of $N$ and $\overline{\tau}$. High winds and low stability will create sustained vertical motion at distances approaching the scale height of the troposphere, 7–8 km (Fig. 3). Low winds under stable conditions will create rapid transitions between ascending and descending motions, with wavelengths of under 1 km. Vertically decaying evanescent solutions (not shown) have also been included in the model implementation.

These VDELB vertical-velocity profiles are used in conjunction with the VDELB parameterizations of relative humidity effects and a theoretical conversion factor, derived from the Clausius–Clapeyron equation and the first law of thermodynamics (Haltiner 1971). The resulting diagnostic relationship is integrated from the surface to the top of the precipitating layer, taken as 300 hPa. Please see S94 for more details.

d. The VDELB* model

The terrain-induced vertical-velocity estimates obtained from Eq. (8) may be combined with large-scale vertical velocities: $w_i = w_o + w_s$. In our studies we have used vertical-velocity profiles from the 4-times-per-day NCEP–NCAR reanalysis to represent $w_s$. Initial tests with the VDELB model (described in section 4) showed that the IGW parameterization better represented the spatial and seasonal rainfall variability. Both VDELB and VDELB*, however, overestimated precipitation during all seasons. We corrected for this by assuming that orographic effects can only enhance a preexisting synoptic system. This can be achieved through modifying $w_s$;
if \( r > 0 \), \( w_i = w_i + w_i \) and
if \( r = 0 \), \( w_i = 0 \),

\[
\text{where } r \text{ is the reanalysis precipitation rate in millimeters per day. This effectively limits orographic rainfall to days and locations with large-scale precipitation. We refer to this constrained model as VDELB*.}
\]

It is important to note that the reanalysis vertical velocity \( w_i \) does not incorporate mesoscale effects. In the reanalysis \( w_i \) is derived from solutions to the primitive equations for vorticity, divergence, virtual temperature, logarithm of surface pressure, and specific humidity on a T62 spectral grid with 28 unequally spaced sigma levels (Kanamitsu 1989; Kanamitsu et al. 1991; Sela 1980). Area-averaged elevations are used on the T62 resolution as well, with an approximate resolution of 2°. The reanalysis does not incorporate mesoscale orographic features. The vertical velocity at the surface is assumed to be zero, and the change in vertical velocity in sigma coordinates is calculated by the Arakawa (1972) vertical finite-differencing scheme. Equation (9) thus makes the valid assumption that the large-scale divergence is independent of the local terrain-induced vertical-velocity fields.

e. Matched filter regression

Section 5 presents an application in southern Africa that fuses reanalysis, VDELB*, and satellite rainfall estimates. We briefly describe the fusion technique here.

Matched filter regression (MFR) is a statistical technique that merges multiple fields to produce an estimate of a single target. It borrows from remote sensing the use of matched filters to isolate a desired signal (Stocker et al. 1990; Tu et al. 1997; Stocker and Schum 1997; Villeneuve et al. 1999; Funk et al. 2001). MFR has been used by Famine Early Warning Systems Network (FEWSNET) scientists to generate crop water satisfaction forecasts for southern Africa for the 2002/03 season (Funk et al. 2002).

MFR takes into account the following two factors: 1) the correlation between each predictor and the target and 2) the convergence of evidence within the predictors themselves. In the first step, predictors are scaled to have a mean of 0 and a variance of \( r^2 \), where \( r \) is each predictor’s correlation with the observations. This weighting gives better-correlated fields more weight in the second step, in which the first principal component (PC1) of the predictors is calculated. Because the values of PC1 will be high where the inputs covary, PC1 indicates a convergence of evidence, moderated by the strength of each input’s relation to the target (correlation). Bi-variate regression is then used to translate PC1 into estimates of rainfall.

If we let \( \mathbf{r} \), \( \mathbf{v} \), and \( \mathbf{s} \) be vectors representing \( n \) \( (n = n_{\text{col}} \times n_{\text{var}}) \), gridded estimates of precipitation based on the NCAR–NCEP reanalysis (\( \mathbf{r} \)), VDELB* model (\( \mathbf{v} \)), and satellite rainfall fields (\( \mathbf{s} \)), then we can describe the MFR process as

\[
\mathbf{r}_i = c_{\mathbf{v}_i} (\mathbf{v}_i - \mathbf{v}) \sigma^{-1}, \quad \forall i \in [1, \ldots, n],
\]

\[
\mathbf{v}_i = c_{\mathbf{v}_i} (\mathbf{v}_i - \mathbf{v}) \sigma^{-1}, \quad \forall i \in [1, \ldots, n],
\]

\[
\mathbf{s}_i = c_{\mathbf{s}_i} (\mathbf{s}_i - \mathbf{s}) \sigma^{-1}, \quad \forall i \in [1, \ldots, n],
\]

\[
\mathbf{m} = \text{PC1}([\mathbf{r}^{T} \mathbf{v}^{T} \mathbf{s}]), \quad \text{and}
\]

\[
\text{mfrfe}_i = b_0 + b_i \mathbf{m}_i, \quad \forall i \in [1, \ldots, n].
\]

The correlations \((c_{\mathbf{r}_i}, c_{\mathbf{v}_i}, c_{\mathbf{s}_i})\), means \((\mathbf{r}, \mathbf{v}, \mathbf{s})\), and standard deviations \((\sigma_{\mathbf{r}_i}, \sigma_{\mathbf{v}_i}, \sigma_{\mathbf{s}_i})\) are calculated based on the subset of locations coincident with the stations. These values are used to create standardized fields in which the expected variances will be equivalent to the squared correlations with the gauge data [Eq. (10)]. The first principal component of these combined datasets [Eq. (11)] provides a weighted combination of \( \mathbf{r}, \mathbf{v}, \) and \( \mathbf{s} \), which can be linearly rescaled to estimate the observed data [Eq. (12)].

The next section presents a validation of the VDELB model in the Pacific Northwest. Section 5 then applies the VDELB and MFR to improve rainfall estimates in southern Africa.

4. Model validation—Northwestern United States

This section describes the implementation of the VDELB model in a data-rich area—the Pacific Northwest (Fig. 4, left panel). Three hundred sixty-five days (the year 1983) of daily gauge observations (Fig. 4, right panel) were compared with results from the VDEL, VDELB, and VDELB* estimators. Station data were generously provided by P. Thornton and S. Running, who have developed statistical interpolation models of this same region (Thornton et al. 1997). Data from the National Weather Service Cooperative Observatory Network and the Natural Resources Conservation Services Snowpack Telemetry System (SNOWTNET—U.S. Department of Agriculture Soil Conservation Service) were combined to provide over 1100 stations. Data quality was verified, and stations missing more than 73 observations for 1983 (20% of 365 days) were dropped. This filter left 813 spatially well-distributed stations in the test region.

The National Geophysical Data Center’s 5-min Digital Elevation Model was used to represent the surface topography. Four-times-per-day temperature, specific humidity, wind velocity, geopotential height, and precipitation-rate reanalysis data (Kalnay et al. 1996) were used to generate precipitation estimates. The atmospheric fields are on a 2.5° geographic grid, and the surface fluxes (precipitation) are arrayed on an approximately 1.875° lattice. Note that NCEP–NCAR do not incorporate precipitation data into their reanalysis process, so that the precipitation rate is an estimated variable.

Five different model runs were made for the 365 days
in 1983 at an 8-km², 6-h resolution. VDEL was tested at three different parameter settings: \( \gamma = (1.5, 1.0, 0.5) \). Because \( \gamma \) controls the rate at which the terrain-induced vertical velocity decays, decreasing \( \gamma \) leads to increasing amounts of rain. The VDELB and VDELB* formulations, based respectively on Eqs. (8) and (9), were also evaluated.

a. Model accuracy assessment

Time series (Fig. 5) of the 365 days of observed gauge and the IGW-based, reanalysis-constrained VDELB* reveal strong levels of correspondence at the regional scale, with a correlation of 0.88 and a mean bias error of 0.2 mm day\(^{-1}\) (7% of annual mean of 3 mm day\(^{-1}\)). The overall skill \([1 - (\text{standard error/standard deviation of observations})]\) of the VDELB* in comparison with at-station daily values is substantially lower (0.47). Accuracy at local/daily scales is thus very low.

Figure 5 suggests that the model tends to overestimate extreme events. This overestimation is likely due to a combination of the linear IGW assumption; the lack of mass, moisture, or thermodynamic constraints; and the limitations inherent in the 2D analysis framework. Unlike 3D representations (Smith 1980), the VDEL and VDELB models cannot replicate flow splitting or blocking and are, thus, likely to overestimate vertical velocities.

Although the magnitude of heavy precipitation events is overestimated, their timing is surprisingly well represented. This result does not hold during the summer, when mesoscale convection will tend to produce small amounts of rainfall not captured by either the reanalysis or VDELB* models.
Plots of the five models averaged over the winter, spring, summer, and autumn seasons (Fig. 6) of 1983 show signs of the 2.5° footprint of the reanalysis atmosphere. The IGW-based VDELB does a better job of recreating the strong longitudinal precipitation gradients found in this region and performed substantially better than the best VDEL model run, with an annual mean absolute error (MAE) of 2.8 versus 4.4 mm day$^{-1}$. Both diagnostic procedures have difficulty recreating the large-scale (synoptic/seasonal) patterns of variability. The constrained VDELB* estimates are more accurate (MAE values of 1.8 mm day$^{-1}$) and realistic (see for comparison the PRISM field in Fig. 2). By combining synoptic process information from the reanalysis fields with the local terrain-induced precipitation patterns, the VDELB* model recreates reasonably well the spatial and temporal patterns of rainfall in this area.

Error grids were created by subtracting each day’s closest VDELB estimate from corresponding gauge observations. These daily error values were interpolated to the study grid and summed to provide seasonal composites. Averaged over the seasons, these plots provide estimates of the mean bias error (Fig. 7). It has been convincingly argued that the areal average of mean absolute error is a good performance statistic for model evaluation (Willmott et al. 1985); seasonal and annual mean bias and absolute error values are listed in Table 1. Three-dimensional theory suggests that splitting will occur when $hN/\pi > 1$, and it appears that the largest errors in the VDELB* model may be associated with relatively weak winds under relatively stable but moist conditions in the spring.

b. VDELB model found useful

Even the lowest MAE values in Table 1 are still very large when compared with the mean (e.g., the annual VDELB* daily MAE of 1.8 mm day$^{-1}$ is 60% of the annual mean of 3 mm day$^{-1}$), but at-station estimates of uncertainty are unforgiving, because they do not ac-
FIG. 7. VDEL, VDELB, and VDELB* mean bias errors for four seasons of 1983 (mm day$^{-1}$). Error fields based on 365 daily model runs are compared with 880+ daily gauge observations. At-station errors were calculated and then interpolated to the study grids to produce error maps.

Table 1. Mean daily precipitation rates (mm day$^{-1}$) calculated at observed locations, together with mean bias errors (MBE) and mean absolute errors (MAE). Error statistics were calculated by comparison with the model grid cell containing the rain gauge. Statistics are based on 365 daily fields for 1983.

<table>
<thead>
<tr>
<th></th>
<th>Winter (Dec–Feb)</th>
<th>Spring (Mar–May)</th>
<th>Summer (Jun–Aug)</th>
<th>Autumn (Sep–Nov)</th>
<th>Annual mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gauge mean</td>
<td>4.6</td>
<td>2.7</td>
<td>1.8</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>VDELB* Mean</td>
<td>5.8</td>
<td>3.2</td>
<td>0.9</td>
<td>3.1</td>
<td>3.2</td>
</tr>
<tr>
<td>VDELB* MBE</td>
<td>1.3</td>
<td>0.5</td>
<td>-1.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>VDELB MBE</td>
<td>2.6</td>
<td>1.5</td>
<td>1.5</td>
<td>1.8</td>
<td>1.9</td>
</tr>
<tr>
<td>VDELB MBE</td>
<td>3.0</td>
<td>1.8</td>
<td>4.0</td>
<td>3.7</td>
<td>3.1</td>
</tr>
<tr>
<td>VDELB MAE</td>
<td>3.4</td>
<td>1.9</td>
<td>3.1</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>VDELB (γ = 0.5) MBE</td>
<td>8.1</td>
<td>6.1</td>
<td>9.3</td>
<td>8.5</td>
<td>8.0</td>
</tr>
<tr>
<td>VDELB (γ = 0.5) MAE</td>
<td>9.3</td>
<td>6.6</td>
<td>9.7</td>
<td>8.9</td>
<td>8.6</td>
</tr>
<tr>
<td>VDELB (γ = 1) MBE</td>
<td>5.6</td>
<td>4.6</td>
<td>7.4</td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td>VDELB (γ = 1) MAE</td>
<td>7.0</td>
<td>5.2</td>
<td>7.8</td>
<td>6.9</td>
<td>6.7</td>
</tr>
<tr>
<td>VDELB (γ = 1.5) MBE</td>
<td>2.5</td>
<td>2.5</td>
<td>4.8</td>
<td>3.6</td>
<td>3.3</td>
</tr>
<tr>
<td>VDELB (γ = 1.5) MAE</td>
<td>4.5</td>
<td>3.3</td>
<td>5.3</td>
<td>4.4</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Count for the cancellation of errors arising from spatial smoothing. The magnitude of the daily VDELB MAE, for example, compares favorably with the monthly MAE values of climatologically aided interpolation (CAI; Willmott and Robeson 1995) fields in data-sparse regions of southern Africa (Fig. 8). The CAI precipitation fields are created by interpolating anomalies from long-term climatological means and are a commonly refer
enced dataset produced by the University of Delaware Geography Department. Because typical monthly MAE values over southern Africa for interpolated gauge data in January are about 30 mm month$^{-1}$ (about 1 mm day$^{-1}$) it is very plausible that the VDELB model can be used to improve rainfall estimates—a topic explored more fully in the next section.

5. Merging VDELB* with satellite rainfall estimates in southern Africa

In this test application, VDELB* estimates are merged with satellite and reanalysis rainfall fields using matched filter regression. Satellite RFE for Africa are primarily based on the length of time that satellite-observed cloud-top temperatures remain above a temperature threshold. These estimates are routinely produced by the Climate Prediction Center (CPC; Xie and Arkin 1997, 1998; Herman et al. 1997) and the Tropical Applications of Meteorology Using Satellite and Other Data (TAMSAT) group at the University of Reading (Grimes et al. 1999). We combine CPC RFE fields for a single month (January 1997) with VDELB and reanalysis rainfall estimates and use a comparison with Global Historical Climate Network (GHCN) data (Peterson and Vose 1997) to demonstrate that the merged matched-filter (MF) RFE are more accurate than the RFE alone.

Figure 9 shows the satellite RFE, gauge locations, and associated gauge–RFE scatterplot. Although the correlation between the observed and estimated precipitation is reasonable ($r = 0.74$), the mean absolute error (74 mm month$^{-1}$) is very high: 49% of the mean of the gauge values.

Comparison with NCEP–NCAR reanalysis precipitation fields and VDELB* estimates revealed similar levels of correlation (0.79 and 0.76, respectively). Even though the data constraining the reanalysis are sparse in sub-Saharan Africa, it does a reasonable job of representing precipitation variability. A similar finding was presented in Funk et al. (2003), which showed that reanalysis data were capable of accurately representing submonthly temporal rainfall variability in Mali and Kenya. It does appear, however, that nonstationarity within the reanalysis fields may create spurious trends, which must be corrected by appropriate use of station data.

The correlations between the station and the RFE, reanalysis, and VDELB fields were used in conjunction with the MFR calculation to produce an improved rainfall estimate (Fig. 10). Fusion increased the correlation (0.84) and reduced the MAE by 35% to 48 mm day$^{-1}$. The outlier to the west of Lake Kivu is more realistic, and the very heavy rainfall on the western shore of Madagascar is better reproduced. The lower rainfall rates over southern Mozambique appear to be more realistic, although no station data were located in this region.

Figure 11 shows a more detailed comparison of the errors (estimated − observed). In some cases (A: eastern Cape, B: central Madagascar, and D: central Malawi), the MF RFE overestimates the precipitation by more than 100 mm. These regions are associated with steep gradients in topography (Fig. 12) and low elevation. The steep gradients make it likely that the linear assumptions of the IGW model are not met, producing unrealistically strong atmospheric responses. Because the elevation of these locations is low, the effect of overestimating terrain-induced vertical velocities may be increased, because the bias can affect precipitation at all pressure levels. Dealing more effectively with low elevation, high-terrain gradient locations could improve model performance.

At higher elevations (C: central Madagascar in Fig. 11), inclusion of orographic effects can improve rainfall estimates. Of interest is that both the satellite and matched-filter rainfall fields underestimated precipitation in Zambia and Zimbabwe (E in Fig. 11). The cause of this bias is not clear, but it might be possible to use a more spatially explicit correction scheme. The satellite RFE tends to dramatically overestimate rainfall in the Congo and Zaire (cases F and G)—a problem that is mitigated by the MFR approach.

6. Summary and discussion

Although nonhydrostatic high-resolution numerical weather prediction models are the best way of representing atmospheric processes in complex terrain, their computational burden will continue to make their application over large domains of space and time difficult. The diagnostic VDELB model described and tested here offers a means of approximating some of the effects of local terrain quickly over large areas at mesoscale resolutions and daily-or-better time steps. This model, or
refinements of it, could be used to represent subgrid orographic processes within global numerical weather prediction models, to enhance satellite-based rainfall estimates, or to improve interpolated gauge surfaces. We presented a brief proof of concept that illustrates this point for a single month in southern Africa.

The VDELB model assumes a Boussinesq atmosphere and a linear solution to the internal gravity wave equations. The linear assumption is associated with errors, especially in conjunction with large-amplitude mountains. Although the VDEL model does not make this assumption, validation in the Pacific Northwest (section 4) suggests that the IGW extension improves performance while removing the need to set the param-

Fig. 9. Validation of CPC Herman RFE for Jan 1997 based on GHCN station data. Station locations are marked by triangles. The error statistics were calculated at station (and are thus very high). Several outliers are denoted by arrows. The “S” location may be associated with unrealistically high rainfall, linked to stratus over southern Mozambique.

Fig. 10. Validation of MF RFE for Jan 1997 based on GHCN station data. Station locations are marked by triangles. The error statistics compare point values (stations) with areal averages and, thus, are very high. Several outliers are denoted by arrows. The S location may be associated with unrealistically high rainfall, linked to stratus over southern Mozambique.
FIG. 11. This plot compares satellite RFE and matched filter regression (MFRFE) errors for Jan 1997; A–G represent especially large errors discussed in the text and noted on Figs. 10 and 11. The numbers within parentheses denote the longitude and latitude of the stations.

FIG. 12. Absolute value of terrain gradient linearly stretched from white to black over the range of observed values. The elevation data used were a resampled (averaged) version of the USGS Global 30 Arc-Second Elevation (GTOPO30) dataset.

Parameter controlling the VDELB model’s vertical-velocity profile. The VDELB appears promising, but the linear assumption appears to be associated with a tendency to systematically overestimate heavy-precipitation events (Fig. 5), especially in regions with large mountains (Fig. 7; case B in Figs. 10 and 11). Steep slopes at low elevations (cases A and D in Figs. 10 and 11) may also be overestimated.

Comparison with GHCN data suggests that the model performance is reasonable in subtropical areas. Tropical regions are more problematic, and considerably higher errors were found north of about 14°S. Both the RFE and MF RFE fields tended to underestimate rainfall in the Zambia–Zimbabwe region and to underestimate rainfall in the central and eastern Democratic Republic of the Congo.

In tropical regions, mesoscale convective processes will play a greater role. Smith and Lin (1982) found that the amplitude of thermally generated gravity waves would typically equal or exceed the amplitude of terrain-induced waves (Nappo 2002). The current model does not represent convection, and thus it underpredicts in the Congo–Zaire basin. An explicit representation of convectively forced IGWs (Orville 1965, 1968; Wyszogrodski 2000) might improve model performance in tropical regimes. The current model might be applicable in simpler tropical areas, such as Central America. Areas with complex mesoscale forcing, such as portions of equatorial eastern Africa, will likely require approaches that are more sophisticated. Preliminary modeling applications in central Asia appear to be promising and potentially useful, given the extreme scarcity of data in this region.

Several improvements to the VDELB model could likely enhance its performance. One missing component is any treatment of moisture conservation or rain-shadow effects. These enhancements could reduce the positive bias found in rain-shadow regions. The time-delay effects associated with hydrometeor effects (Smith 2003) should also be added. Difficulties associated with
the VDELB model’s linear assumption might be mitigated by running the model at a coarser resolution (say, ~25 km × 25 km), similar to global satellite precipitation products. The plausibility of using a full 3D mountain wave model should also be explored. An improved version of the model, with better representations of tropical orographic effects, might be effectively used to enhance these global rainfall estimates.

Acknowledgments. This work has been motivated and funded by the U.S. Agency for International Development Famine Early Warning System, through the U.S. Geological Survey (USGS) International Program (Cooperative Agreement DO1 99CRAGG0029). Chris Funk received support from a STAR fellowship provided by the U.S. Environmental Protection Agency (U-915612-010). The astute comments of our anonymous reviewers improved this manuscript considerably, as did suggestions provided by scientists from the USGS (Jim Verdin and Jim Rowland) and UCSB (Dar Roberts and Keith Clark). Previous work by and discussions with the Climate Prediction Center (Jim Miller, Tim Love, Ping Xie, and Kevin Laws) and the Regional Remote Sensing Unit in Harare (Kennedy Masamvu, Tinashe Pinx) and the Regional Remote Sensing Unit in Harare (Kennedy Masamvu, Tinashe Magadzire) helped to inspire this research effort. Jim Verdin originated the idea of using this orographic model as a means of enhancing rainfall estimates for food-security applications.

REFERENCES


