ABSTRACT: Although large-scale atmospheric reanalyses are now providing physical, realistic fields for many variables, precipitation remains problematic. As shown in recent studies, using a regional model to downscale the global reanalysis only marginally alleviates the precipitation simulation problems. For this reason, later-generation analyses, including the recent National Centers for Environmental Prediction regional reanalysis, are using precipitation assimilation as a methodology to provide superior precipitation fields. This methodology can also be applied to regional model simulations to substantially improve the precipitation fields downscaled from global reanalysis. As an illustration of the regional model precipitation assimilation impact, the authors describe here an extended-range simulation comparison over South America. The numerical experiments cover the beginning of the Large-Scale Biosphere–Atmosphere wet season campaign of January 1999. Evaluations using radiosonde datasets obtained during this campaign are provided as well. As will be shown, rain-rate assimilation not only increases the regional model precipitation simulation skill but also provides improvements in other fields influenced by the precipitation. Because of the potential impact on land surface features, the authors believe they will ultimately be able to show improvements in monthly to seasonal
forecasts when precipitation assimilation is used to generate more accurate land
surface initial conditions.

**KEYWORDS:** Precipitation assimilation; Regional model; South America

1. Introduction

During the past several years, reanalysis products have provided many large-scale meteorological fields that are not only useful for validating coarse-scale climate models, but can also be used as forcing for various regional application models (e.g., hydrologic models). These reanalysis products have been subject to much evaluation, and, unfortunately, precipitation and associated hydrologic products remain problematic.

Reanalysis precipitation is classified as a \(C\) variable, which means that it is not assimilated, but entirely determined by the model 6-h forecasts (Kalnay et al. 1996). Janowiak et al. (Janowiak et al. 1998), Roads and Betts (Roads and Betts 2000), and Roads and Chen (Roads and Chen 2000) showed that although large-scale patterns of precipitation fields from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) global reanalysis project [R-1 (Kalnay et al. 1996); R-2 (Kanamitsu et al. 2002)] agree somewhat with observations, smaller-scale regions, particularly monsoon regions, had notable problems.

As shown previously by Roads et al. (Roads et al. 2003), downscaling these reanalysis products with regional models did not produce any notable improvement in South American precipitation. In fact, in many cases regional simulations produce degraded results compared to global analyses. This was quite surprising since it was thought that regional models would improve a number of regional features such as the South American low-level jet (SALLJ), which is characterized by an increase of the horizontal wind speed at lower-troposphere levels along the eastern part of the Andes. Transport of moisture from the Amazon region toward the southern part of South America, and moisture carried out from the Amazon basin, often condenses and precipitates in the region of the SALLJ convergence, producing explosive mesoscale convective complexes downstream of the low-level jet core, with a maximum of precipitation during the night (Nogués-Paegle and Berbery 2000).

This lack of improvement was especially discouraging since regional models provide a means to investigate the influence of regional features on the atmospheric circulation by means of a better characterization of topography, coastal shape, land–sea ice distribution, contrasts in soil use, etc. While there are a number of research avenues that need to be explored to determine why current regional simulations are not able to substantially improve the reanalysis, we believe that incorporation of the precipitation data may be a useful interim tool to provide better regional simulations than simply using the driving reanalysis alone.

Several techniques have been proposed to improve the analysis of the moisture and divergence, which is essential to reproduce the tropical circulation mainly maintained by gravitational modes due to diabatic forcings (Puri and Bourke 1982), such as Donner (Donner 1988), Heckley et al. (Heckley et al. 1990), Puri
and Miller (Puri and Miller 1990), Puri and Davidson (Puri and Davidson 1992), Aonashi (Aonashi 1993), and Kasahara et al. (Kasahara et al. 1994). Precipitation assimilation as part of a physical initialization (PI) procedure has also been proposed to reduce the analysis imbalance due mostly to misrepresentation of the moisture fields by the analyses.

The PI procedure was first described by Krishnamurti et al. (Krishnamurti et al. 1984), and subsequently in Krishnamurti et al. (Krishnamurti et al. 1988, 1991), where the specific humidity vertical profiles are adjusted by reversing a modified Kuo cumulus parameterization. Manobianco et al. (Manobianco et al. 1994) used the satellite-derived rain rates to scale the model’s profiles of total latent heating generated by the Kuo cumulus convection parameterization scheme as well. Basically the former PI schemes utilize the ratio of “observed” and model-predicted rain rates to adjust their vertical profiles in a global model.

The PI scheme was also applied to regional models. Yap (Yap 1995) applied the PI scheme to the Florida State University (FSU) regional gridpoint model, and Nunes and Cocke (Nunes and Cocke 2004) subsequently developed this procedure for the FSU regional spectral model, again with a Kuo-type cumulus convection scheme. Nunes and Cocke (Nunes and Cocke 2004) also studied the impact of the boundary condition on the regional model assimilation by using the PI scheme and found out that the global base fields have little influence on the regional model fields in the inner domain, despite the base fields being part of the regional solution.

All the above PI procedures were developed to improve initial condition and, consequently, to increase the short-term rainfall forecasting mainly over the Tropics. We have now redesigned PI for climate simulations in order to improve regional downscaling by means of continuous precipitation assimilation. In particular, we have now included the new PI into the Experimental Climate Prediction Center Regional Spectral Model (ECPC-RSM) code, which was previously described by Juang and Kanamitsu (Juang and Kanamitsu 1994), Anderson et al. (Anderson et al. 2001), and Roads (Roads 2004). In this study the modified PI was applied to the simplified Arakawa–Schubert cumulus convection parameterization scheme (SAS) for long-term simulations over South America, where satellite rain rates were continuously assimilated.

In sections 2 and 3 we describe the model, the assimilation procedure, and data, followed by the experiment results for the extended simulation. Some concluding remarks and future goals are presented in section 4.

2. Methodology

2.1. Model

The ECPC-RSM is a hydrostatic, primitive equation model, with similar physics as the driving R-2 Global Spectral Model (Kanamitsu et al. 2002). However, the regional simulations include an updated Oregon State University land surface model (Pan and Mahrt 1987) with an increased number of vegetation types. The convection parameterization is the same as R-2, that is, SAS, originally designed by Pan and Wu (Pan and Wu 1994). The ECPC-RSM in this study had a horizontal
resolution of about 60 km, which was selected to be compatible with the resolution of the satellite rain-rate dataset, and 28 vertical layers. A Mercator projection was used for the regional grid. Figure 1 shows the ECPC-RSM 60-km orography for the model domain. The small box represents the area where the radiosondes used as validation data are launched.

2.2. Precipitation assimilation

A “perfect observation” assumption is being used by the precipitation assimilation scheme described below, where “observed” or estimated rain rates are a strong constraint. That is to say, these data are assimilated without knowledge of their error characteristics. This is not the ideal hypothesis due to the uncertainty associated with precipitation estimates, but for now, this will be our starting point to determine the impact of the assimilation on the analysis, especially the removal of the model’s systematic precipitation errors by reducing the inconsistencies between the model and the coarser forcing analysis parameterizations.

Figure 1. ECPC-RSM 60-km orography over South America. The box represents the radiosonde dataset area used in the comparison. Rebio Jarú, ABRACOS, and Rolim Moura are WETAMC/LBA stations, and Alta Floresta is a WMO station.
The precipitation assimilation scheme proposed is based on a modified PI procedure where a complete numerical model with full physics is integrated in time in order to bring the model’s precipitation closer to the “observed” or prescribed rain rates. We then applied the modified PI (only rainfall nudging) to ECPC-RSM using SAS, although preliminary experiments have shown that the procedure also works for the relaxed Arakawa–Schubert convection parameterization (Nunes et al. 2004). The modified PI basically adjusts the specific humidity vertical profile, taking into account the difference between “observed” and predicted rain rates as given by

\[ q_k^a = q_k + \frac{g}{p_s} \eta \Delta_k (R_{obs} - R_{mod}). \]  

(1)

Here \( q_k^a \) and \( q_k \) are the specific humidity obtained from rainfall nudging and predicted by the model, respectively; \( \Delta_k (R_{obs} - R_{mod}) \) is a partition of the total rain-rate difference, where \( R_{obs} \) is the “observed” rain rate, \( R_{mod} \) is the model’s rain rate, and \( \eta \equiv \eta(k) \) is a parameter of nudging. The subscript \( k \) represents a particular layer. In this study, the simplest assumption to represent the \( \eta \) vertical distribution was made constrained by

\[ \frac{p_s}{g} \int_{\sigma_t}^{\sigma_b} (q^a - q) \ d\sigma = R_{obs} - R_{mod}, \]  

(2)

where \( \sigma_t \) is the \( \sigma \) surface at the upper limit level, and \( \sigma_b \) represents the top of the surface layer. However, the scheme can incorporate a prescribed moistening profile, which could eventually be obtained from the model or even data profiles. The limits for the vertical integration can be changed to where \( R_{mod} > 0 \). This methodology is based on a supposition that in the limit the total precipitation (evaporation) difference was used to modify the moisture profile. The adjustment takes place when either “observed” or predicted rain rates are greater than 5 mm day\(^{-1}\). This particular threshold was chosen based on the performance of a few extended simulations. A lower threshold resulted in a too-dry bias. This scheme is totally dependent on the cumulus and large-scale parameterizations, since they determine the latent heat release after the humidity adjustment takes place. The advantage of using this scheme, in comparison to standard PI procedures, is that the moisture adjustment can occur even in the absence of the model’s precipitation. As proposed in the former PI schemes, we assume that the uncertainties in the humidity field are much larger than in the temperature, as was also recently suggested by Hou et al. (Hou et al. 2000).

The original moisture adjustment is based on the vertical structure function described in Krishnamurti et al. (Krishnamurti et al. 1991) for a modified Kuo convection scheme, where, to minimize the drift from the observed data, a Newtonian relaxation toward the analyses was applied to the spectral coefficients of vorticity, divergence, and surface pressure. The specific humidity and temperature were left out in order to preserve the moistening and heating rates from the precipitation assimilation scheme. In contrast to the original PI scheme, Newtonian relaxation was not used here, because the regional model’s boundary conditions (i.e., the base field, which is part of the regional model total field) were updated.
each 6 h using reanalysis. This reduces any substantial drift due to moisture adjustment and avoids the strong damping introduced by nudging the spectral coefficients toward the coarser-scale analysis, and makes the regional spectral model suitable for use with this type of scheme. In addition to this, Newtonian relaxation toward the reanalysis fields was not employed in this particular study because we wanted to assure that the rainfall assimilation itself was able to bring the rest of the prognostic variables closer to the boundary fields.

2.3. Datasets

2.3.1. Reanalysis sets

2.3.1.1. National Centers for Environmental Prediction reanalysis. Initial and boundary conditions for all experiments were provided by the NCEP–Department of Energy (DOE) Atmospheric Model Intercomparison Project (AMIP-II) reanalysis (R-2) described previously by Kanamitsu et al. (Kanamitsu et al. 2002). R-2 has a triangular spectral truncation of 62 waves, or T62, corresponding to a horizontal resolution of about 200 km at the equator, and 28 vertical layers. In R-2, the observed 5-day mean precipitation based on rain gauge and satellite observations was adopted to adjust the soil moisture of the top soil layer (Kanamitsu et al. 2003).

2.3.1.2. European Centre for Medium-Range Weather Forecasts reanalysis. The 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis project (ERA-40) (Simmons and Gibson 2000) was used as an independent dataset to evaluate the ECPC-RSM results. The ERA-40 analysis methodology allows the direct use of nonconventional data such as those provided through satellite measurements. ERA-40 uses a spectral global forecast model with triangular truncation of 159 waves and 60 levels in the vertical, including a well-resolved boundary layer and stratosphere (Betts et al. 2003).

2.3.2. Station datasets

2.3.2.1. LBA and World Meteorological Organization radiosondes. During the Wet Season Atmospheric Mesoscale Campaign/Large-Scale Biosphere–Atmosphere Experiment (WETAMC/LBA) radiosonde sites in Rondonia, Brazil, were subject to a quality control process based on visual inspection, plausibility, and spatial and physical consistency (Longo et al. 2002). We selected three of these radiosonde sites based on their frequency, namely, ABRACOS (Ouro Preto d’Oeste), Reserva Biologica do Jaru (Rebio Jaru), and Rolim de Moura, respectively, 10.75°S, 62.37°W; 10.14°S, 61.91°W; and 11.70°S, 61.78°W (Figure 1). We also included the radiosonde site of Alta Floresta in Mato Grosso, Brazil, located at 9.87°S, 56.10°W (Figure 1), which is one of the World Meteorological Organization (WMO) stations, among the validation radiosonde data.

2.3.2.2. Surface synoptic observations. Daily values of precipitation (millimeters) and maximum and minimum temperatures (degrees Celsius) were used as validation data. They were obtained from the Centro de Previsão do Tempo e Estudos Climáticos (CPTEC), São Paulo, Brazil, from surface synoptic observa-
tion (SYNOP) reports transmitted internationally through the Global Telecommuni-
cation System (GTS).

2.3.3. Precipitation datasets

2.3.3.1. Special Sensor Microwave Imager. Special Sensor Microwave Im-
ager (SSM/I) radiometer measurements from satellites launched by the Defense
Meteorological Satellite Program (DMSP) were used to provide daily rain-rate
estimates for all experiments. The National Oceanic and Atmospheric Admin-
istration (NOAA)/National Environmental Satellite, Data, and Information Service
(NESDIS) SSM/I algorithm (Ferraro and Marks 1995), based on scattering and
emission methods, was used to provide the rain-rate estimates. Scattering-based
retrieval algorithms are usually used over land, and they do not perform well on
systems with little or no ice particles, that is, warm-rain systems (Ferraro and
Marks 1995). For this reason they may not be appropriate to detect rain in Nor-
deste, where most of the rain is produced by warm clouds. In the absence of SSM/I
data, NESDIS outgoing longwave radiation (OLR) data were blended into the
rainfall dataset (Gairola and Krishnamurti 1992; Arkin 1979). These daily esti-
mates were then linearly interpolated in time to attain rain rates for the regional
model time step of 3 min. The SSM/I–OLR precipitation estimates were provided
on a Gaussian grid of about 70 km with defined values from 60°N to 60°S, and
then bilinearly interpolated to the ECPC-RSM model grid.

2.3.3.2. One-degree-daily Global Precipitation Climatology Project. The
one-degree-daily precipitation estimates included into the Global Precipitation
Climatology Project (1DD-GPCP; Huffman et al. 2001) evaluate the model pe-
rformance for simulating rainfall patterns for PI and Control experiments.
The algorithm is divided into two parts: the threshold-matched precipitatio
n index (TMPI) from 40°N to 40°S, which is based on a merged infrared (IR) radiometer
on geostationary satellite (geo-IR) as in Arkin and Meisner (Arkin and Mei
sner 1987) with low-Earth-orbit IR, and rescaled daily Television and Infrared Obser-
vation Satellite (TIROS) Operational Vertical Sounder (TOVS) for regions outside
the latitude belt covered by the merged geo-IR (Susskind et al. 1997).

3. Extended simulation experiment

To illustrate the advantages of using precipitation assimilation to improve the
climate regional downsampling, the following experiments were performed over
South America during part of the WETAMC/LBA campaign in January 1999. Our
experiment consisted of continuous assimilation of the daily SSM/I–OLR precipi-
tation estimates using modified PI as described in section 2.1. Long-term integra-
tions of ECPC-RSM are presented below. SAS was chosen for the RSM continu-
os simulations. Control simulations, in which no initialization was used, were
also performed.

3.1. Precipitation patterns

The ECPC-RSM executed continuous precipitation assimilation for January 1999.
Figure 2 shows the monthly mean precipitation for 60-km ECPC-RSM resolution
Figure 2. Total precipitation (mm day$^{-1}$) for Jan 1999: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, (d) ERA-40, (e) SSM/I–OLR, and (f) 1DD-GPCP.
Figure 2. (Continued)
Figure 2. (Continued)
with modified PI, Control (without precipitation assimilation), R-2, ERA-40, SSM/I–OLR, and 1DD-GPCP estimates, respectively. The PI simulation agrees quite well with the SSM/I–OLR estimates. Comparing Figures 2a and 2e, we note that spurious rain around the Andes Cordillera present in Figures 2b–d was removed by the precipitation assimilation. Figures 2b–d also show excessive precipitation over the intertropical convergence zone. The 1DD-GPCP (Figure 2f) displays a similar precipitation pattern of SSM/I estimates, but the precipitation is more spread due to its resolution, and the rainfall cores over Paraguay and the center of Argentina are not shown.

Table 1 presents spatial correlation coefficients and root-mean-square errors (rm ses) between PI, Control experiments, and SSM/I–OLR rainfall estimates for the mean precipitation fields. Comparisons with R-2, ERA-40, and SSM/I–OLR estimates were also included in Table 1. From Table 1, we can see significant improvement in the monthly mean precipitation for January 1999, when precipitation assimilation (PI) was present in the regional simulation.

Figure 3 shows two more scores, namely, equitable threat score (ETS) (Schaefer 1990) and Bias for the daily accumulated values during the month against the SSM/I values. ETS and Bias show the positive impact of using PI compared to the Control, especially for rain rates between 5 and 10 mm day$^{-1}$. In particular, note that higher precipitation values are also well represented by the PI simulations (Figures 3a,b). These scores are used to show that the model is not able to reproduce the SSM/I pattern without the inclusion of the precipitation data.

Figure 4 displays ETS and Bias for PI, Control, and SSM/I compared to the 1DD-GPCP estimates. The PI curve in Figure 4a shows higher ETS than Control, which is not surprising because it mostly follows the SSM/I pattern. The SSM/I lower ETS and higher Bias (Figure 4b) for threats lower than 1 mm day$^{-1}$ might be due to the SSM/I interpolation to the model grid. PI still maintains a better performance than Control. Comparisons with rain gauge will be shown further in this section.

In Figure 5, the temporal evolution of daily precipitation values (Figure 5a) and correlation coefficients for precipitation versus the 1DD-GPCP precipitation estimates (Figure 5b) are displayed. PI is not perfectly adjusted to the SSM/I curve, because the procedure does not totally adjust the rain at one single time step. Control has the lower correlation values during the entire simulation. The mean values are (in millimeters per day): 3.70, 4.69, 3.98, 3.86 for PI, Control, SSM/I, and GPCP, respectively. The average precipitation difference between the Control and the PI simulations is around 1 mm day$^{-1}$, which is about 26% of the mean “observed” value. PI has the lowest mean value of precipitation and is closer to the GPCP mean value in comparison to the control experiment.

### Table 1. Spectral models vs SSM/I–OLR estimates for Jan 1999.

<table>
<thead>
<tr>
<th>Monthly mean precipitation (mm day$^{-1}$)</th>
<th>Correlation coefficient/rmse (mm day$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td>0.98/0.97</td>
</tr>
<tr>
<td>Control</td>
<td>0.42/6.24</td>
</tr>
<tr>
<td>R-2</td>
<td>0.56/4.64</td>
</tr>
<tr>
<td>ERA-40</td>
<td>0.58/4.43</td>
</tr>
</tbody>
</table>
Figure 3. Scores for daily precipitation (mm day$^{-1}$) over South America during Jan 1999: (a) ETS and (b) Bias. The blue curve indicates PI experiment, and the red curve indicates Control. Comparisons were made against SSM/I–OLR estimates.
Figure 4. Scores for daily precipitation (mm day$^{-1}$) over South America during Jan 1999: (a) ETS and (b) Bias. The blue curve indicates PI experiment, red curve represents Control, and the green curve shows SSM/I–OLR estimates. Comparisons were made against GPCP estimates.
Figure 5. Temporal evolution during Jan 1999 over South America of (a) daily precipitation values (mm day$^{-1}$) and (b) spatial correlation coefficients (PI, Control, and SSM/I vs GPCP). The blue curve indicates PI experiment, the red curve represents Control, the green curve shows the SSM/I–OLR estimates, and the yellow curve corresponds to 1DD-GPCP daily values.
3.2. Prognostic variable patterns

We performed some additional evaluation of the scheme, using the reanalysis fields. The R-2 serves as the base field of the regional solution. For this reason we compare the PI and the Control with the R-2 as a measure of how the precipitation assimilation leads the model to deviate from the base field. Besides the R-2 field comparisons, we also employ an independent reanalysis set, the ERA-40.

Table 2 shows the spatial correlation coefficient and the rmse of the temperature, the relative humidity, and the horizontal wind mean values at lower- and upper-tropospheric levels, respectively, 850 and 300 hPa, against the regional model base field, that is, R-2. The most significant difference for both correlation and rmse is for the relative humidity fields, which is related to the prognostic variables: temperature and specific humidity, which is directly adjusted by the scheme. The increased PI rmse at 300 hPa suggests that PI dramatically changes the specific humidity at upper levels, although the temperature changes are not significant.

Table 3 displays the same scores of Table 2 but against the ERA-40. The Control rmse for relative humidity are now increased compared to the PI, which indicates that the R-2 and the Control underestimate the humidity values from the lower to the upper levels, at least compared to the ERA-40 and the PI.

In Figure 6, the temporal evolution of the spatial correlation coefficients at intervals of 6 h are displayed for the prognostic variables: temperature (Figure 6a) and horizontal wind, which was separated into zonal (Figure 6c) and meridional (Figure 6d) components, and for relative humidity (Figure 6b) at 850 hPa. The solid lines correspond to correlation values obtained from comparison with R-2, and the dashed lines represent the correlation coefficients against ERA-40. The PI simulation (blue line) cannot be easily distinguished from the Control, except for the relative humidity panel, where the correlation values are greater than the Control during the simulation for the R-2 and the ERA-40 comparisons.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relative humidity (%)</th>
<th>Temperature (K)</th>
<th>Horizontal wind (m s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level (hPa)</td>
<td>850</td>
<td>300</td>
<td>850</td>
</tr>
<tr>
<td>PI</td>
<td>0.92/5.60</td>
<td>0.81/12.32</td>
<td>0.97/0.81</td>
</tr>
<tr>
<td>Control</td>
<td>0.87/7.70</td>
<td>0.78/8.83</td>
<td>0.96/0.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Relative humidity (%)</th>
<th>Temperature (K)</th>
<th>Horizontal wind (m s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level (hPa)</td>
<td>850</td>
<td>300</td>
<td>850</td>
</tr>
<tr>
<td>PI</td>
<td>0.91/7.16</td>
<td>0.73/12.68</td>
<td>0.94/1.06</td>
</tr>
<tr>
<td>Control</td>
<td>0.81/10.44</td>
<td>0.38/21.55</td>
<td>0.94/1.20</td>
</tr>
</tbody>
</table>
Figure 6. Spatial correlation coefficient temporal evolution during Jan 1999 over South America at 850 hPa for (a) temperature, (b) relative humidity, (c) zonal wind component, and (d) meridional wind component. The solid blue and red curves represent the PI and Control, respectively, vs R-2, and the dashed blue and red curves correspond to PI and Control, respectively, vs ERA-40.
Figure 7 shows the spatial correlation coefficients for 300 hPa, as in Figure 6. As shown in Table 3, the low performance of the Control simulation compared to the ERA-40 is enhanced by the red dashed line in Figure 7b, which corresponds to the relative humidity field upper-level field. Temperature, zonal, and meridional wind component correlations (Figures 7a, 7c, and 7d, respectively) are not sig-
Figure 7. Spatial correlation coefficient temporal evolution during Jan 1999 over South America at 300 hPa for (a) temperature, (b) relative humidity, (c) zonal wind component, and (d) meridional wind component. The solid blue and red curves represent the PI and Control, respectively, vs R-2, and the dashed blue and red curves correspond to PI and Control, respectively, vs ERA-40.
significantly improved by the PI procedure, and are sometimes outperformed by the Control simulation.

The PI, Control, R-2, and ERA-40 mean values for relative humidity, temperature, and horizontal wind are displayed below for January 1999. The mean pattern of these variables will enhance the areas where either the PI, or the Control, or even both simulations systematically fail. At least two well-defined episodes of the
South Atlantic convergence zone (SACZ), which is usually related to strong convective activity over the Amazon region, occurred in January 1999 [the first and the third week of this month; the third week is described in Herdies et al. (Herdies et al. 2002)].

From Figure 8a, we can see that PI brings the relative humidity fields at 850 hPa closer to the R-2 (Figure 8c) and the ERA-40 (Figure 8d) than the Control (Figure 8b) (see Tables 2 and 3 where the correlation coefficients and rmse are displayed). For example, compare the PI, and the Control, to the R-2 and the ERA-40, especially over the SACZ and the Amazon region.

Figure 9 displays the relative humidity fields at 300 hPa for the PI, the Control, the R-2, and the ERA-40. The ERA-40 (Figure 9d) shows a band of higher relative humidity values extending from the Amazon region to the South Atlantic Ocean. A similar pattern is found in Figure 9a, the PI simulation. This pattern does not exist in Figure 9b, the Control, and it is only insinuated by the R-2 (Figure 9c). We speculate that the ERA-40 analysis methodology, which includes more satellite data, might be responsible for that well-defined band.

Figures 10 and 11 show the mean temperature patterns over South America at 850 and 300 hPa, respectively. At 850 hPa, the PI (Figure 10a) agrees with the R-2 (Figure 10b); the Control (Figure 10b) shows increased temperature values over the central-west region of Brazil. Table 2 suggests some advantage to the Control (Figure 11b) for 300 hPa compared to the PI (Figure 11a), although the R-2 (Figure 11c) and the ERA-40 (Figure 11d) patterns are similar.

Figures 12 and 13 display the streamlines associated with the horizontal wind fields at 850 and 300 hPa, respectively, where the shaded values correspond to the horizontal wind magnitudes. Although the spatial correlation coefficient and rmse (Tables 2 and 3) do not show any substantial improvement for the PI low-level horizontal wind field (Figure 12a) compared to the Control (Figure 12b), we do see that the PI wind field at 850 hPa now presents a northwesterly flow crossing eastern Bolivia, western and the south of Brazil, better indicating the SALLJ (Figure 12a), which is usually associated with an increase of the mean precipitation in this location (see Figures 2a, 2e). This precipitation core is not represented in Control (Figure 2b), and ERA-40 (Figure 2d), or properly by R-2 (Figure 2c). At least one episode of SALLJ was documented during the third week of January 1999 by the WETAMC/LBA (Marengo et al. 2002). A cyclonic circulation is also shown over Paraguay, Argentina, and southern Brazil in Figure 12a, to the right of the precipitation core, where moisture is being supplied from the South Atlantic Ocean. A cyclonic pattern can also be seen over the northeastern coast of Uruguay in Figure 12c (R-2), in general agreement with ERA-40 (Figure 12d). There is not a similar pattern in Figure 12b (Control).

In Figure 13, the Bolivian high summer circulation in the upper-troposphere levels is clearly represented in the PI and the Control experiments (Figures 13a,b). In Figures 13c and 13d, the R-2 and the ERA-40 wind fields divide the upper-level wind pattern into two cores. This might be due to their coarser resolution. The maximum of cyclonic vortex over the Nordeste depicted for both reanalyses (Figures 13c,d) is also represented by the PI field (Figure 13a), although it is displaced to the northwest.

In any event, all comparisons with the R-2 and the ERA-40 fields have shown that precipitation assimilation does not noticeably degrade the rest of the reanalysis.
Figure 8. Relative humidity (%) for Jan 1999 at 850 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 8. (Continued)
Figure 9. Relative humidity (%) for Jan 1999 at 300 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 10. Temperature (K) for Jan 1999 at 850 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 10. (Continued)
Figure 11. Temperature (K) for Jan 1999 at 300 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 11. (Continued)
Figure 12. Horizontal wind (m s\(^{-1}\)) for Jan 1999 at 850 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 13. Horizontal wind (m s$^{-1}$) for Jan 1999 at 300 hPa: (a) ECPC-RSM with PI, (b) ECPC-RSM Control, (c) R-2, and (d) ERA-40.
Figure 13. (Continued)
fields during the simulation. Thus, extended simulation using continuous precipitation assimilation demonstrates the consistency of the present methodology. That is, the drift from the forcing base field after one-month simulation is likely insignificant in comparison to the observational data error on tropical regions.

### 3.3. Comparison with WETAMC/LBA observations

Figures 14–16 display vertical profiles for the PI, the Control, three of the WETAMC/LBA radiosondes, and one WMO radiosonde for temperature, dewpoint temperature, and wind speed. The PI and the Control profiles were obtained from a bicubic interpolation of the model values to the exact location of the radiosonde sites. They were averaged only when and where (vertical levels) the observations were available. The blue line represents PI, the red line displays Control, and the green line corresponds to the observation for all vertical profiles. The small plots on the right side represent the differences between the PI (blue lines), the Control (red line) (S) profiles, and the observation (O).

The average temperature profiles (Figures 14a,d) show that for Rebio Jaru (Figure 14a), Rolim de Moura (Figure 14b), ABRACOS (Figure 14c), and Alta Floresta (Figure 14d) the mean differences are irrelevant. Figure 15 shows significant differences between the dewpoint temperatures for all selected sites (Figures 15a–d) from the lower levels to the upper limit level for the adjustment (around 300 hPa). The Control profile is systematically drier than the PI.

The wind speed profiles are displayed in Figure 16. The control experiment profiles tend to show a maximum of speed around 700 to 600 hPa. Except for Rebio Jaru (Figure 16a), this tendency is not followed by the observation. However, the PI profile is not better than the Control. Actually, the PI profiles overestimate the wind speed above the 500-hPa level. On the other hand, the Control seems to underestimate the wind speed, except for Alta Floresta (Figure 16d).

### 3.4. Comparisons with SYNOP dataset

The SYNOP dataset density is represented in Figure 17, for the 24-h accumulated precipitation (Figure 17a), the maximum daily temperature (Figure 17b), and the minimum daily temperature (Figure 17c). Orange to red points indicate that over 20 station data are available during January 1999 at that location. One concern with using observed data is the quality and frequency of observation, especially in the Tropics. Mass et al. (Mass et al. 2002) pointed out that a sparse observational network can result in a mistaken model validation.

The precipitation, maximum, and minimum temperature model outputs were interpolated to the station locations and then compared to the observed values only for the dates and places where they are available. Blue and red represent negative and positive biases, respectively. Figures 18a and 18b show that PI has a precipitation positive bias over northeastern and central Brazil (Figure 18a) induced by the SSM/I–OLR increased precipitation values over these regions, and the opposite is true for the Control (Figure 18b). Despite this, the maximum temperature (Figure 18c) shows that PI is not significantly affected and presents lower bias values, especially over the Nordeste and Amazon regions compared to the Control (Figure 18d). Both have a negative bias over Chile and Argentina. Figures 18d and 18e show that the minimum temperature bias values are not significantly different from PI to Control.
Figure 14. Composite of the average vertical profiles of air temperature (°C) at (a) Rebio Jaru, (b) Rolim Moura, (c) ABRACOS, and (d) Alta Floresta. The small plots on the right are the simulation (S) differences from the observation (O). The green curve represents the station values, the blue curve indicates the PI experiment, and the red corresponds to the control experiment.
Figure 15. Composite of the average vertical profiles of dewpoint temperature (°C) at (a) Rebio Jaru, (b) Rolim Moura, (c) ABRACOS, and (d) Alta Floresta. The small plots on the right are the simulation (S) difference from the observation (O). The green curve represents the station values, the blue curve indicates the PI experiment, and the red corresponds to the control experiment.
Figure 16. Composite of the average vertical profiles of wind speed (m s$^{-1}$) at (a) Rebio Jaru, (b) Rolim Moura, (c) ABRACOS, and (d) Alta Floresta. The small plots on the right are the simulation (S) differences from the observation (O). The green curve represents the station values, the blue curve indicates the PI experiment, and the red corresponds to the control experiment.
Figure 17. Station data density for Jan 1999 for (a) 24-h accumulated precipitation, (b) maximum daily temperature, and (c) minimum daily temperature.
4. Concluding remarks

Including precipitation assimilation in a regional simulation is useful for downscaling precipitation fields provided by a different coarser model. As shown here, a long-term regional simulation could be performed without decreasing the model’s skill for other prognostic variables. In fact, for humidity-related variables, the procedure did improve the model skill for long-term simulations. From the extended simulations we also noticed that the model systematic errors such as the spurious rainfall over the Andes Cordillera were removed. The advantage of using our PI scheme was that the humidity profile was adjusted even though there was no model-produced precipitation.

In this study, we were particularly interested in investigating the impact of a continuous insertion of precipitation for simulation of monthly mean values over South America. Additional improvements are possible in these regional downscaling simulations with modern precipitation datasets that not only have higher spatial resolution but also higher temporal resolution. This may be especially important in regions with low-level jet activity, which tends to strongly peak during the nighttime.

Our ultimate goal in assimilating precipitation was not only to improve the precipitation fields but also coupled land–atmosphere interactions with land surface processes that are heavily dependent upon accurate precipitation. This coupled methodology could provide a potentially improved alternative to current
Figure 18. Model simulation minus station data for Jan 1999: for 24-h accumulated precipitation (mm) (a) PI and (b) Control; with maximum daily temperature (°C) (c) PI and (d) Control; and minimum daily temperature (°C) (e) PI and (f) Control.
Figure 18. (Continued)
Figure 18. (Continued)
offline data assimilations. Research to investigate other possible improvements is currently underway for the United States and other domains and will be reported in subsequent studies.

Acknowledgments. This research was funded by NOAA Cooperative Agreement NA17RJ1231, USDA FS 01-CA-11272169-149, 02-CA-112272166-056, and NASA NAG8-1875. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA, USDA, or NASA. This study used ECMWF ERA-40 data obtained from the ECMWF data server and NCEP–DOE AMIP-II reanalysis obtained from the NCEP (NOMADS2) data server. WETAMC/LBA radiosondes were obtained from the Department of Atmospheric Science, University of São Paulo (USP), Brazil. The SSM/I–OLR estimates were provided by Florida State University. We thank the anonymous reviewers and the editor for their helpful comments.

References


——, and A. Betts, 2000: NCEP–NCAR and ECMWF reanalysis surface water and energy budgets for the Mississippi River basin. *J. Hydrometeor.*, 1, 88–94.


