The South American Land Data Assimilation System (SALDAS) 5-Year Retrospective Atmospheric Forcing Datasets

Luis Gustavo G. de Goncalves
Earth System Science Interdisciplinary Center (ESSIC) - University of Maryland also at Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD, USA

William J. Shuttleworth
Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, USA

Daniel Vila
Cooperative Institute of Climate Studies (CICS), ESSIC/UMD, College Park, MD, USA

Eliane Larroza, Marcus J. Bottino, Dirceu L. Herdies, Jose A. Aravequia, Joao G. Z. De Mattos
Centro de Previsao do Tempo e Estudos Climaticos/ Instituto Nacional de Pesquisas Espaciais, Cachoeira Paulista, SP, Brazil

David L. Toll, Matthew Rodell
Hydrological Sciences Branch, NASA – Goddard Space Flight Center, Greenbelt, MD, USA

Paul Houser
Center for Research on Environment and Water, Institute of Global Environment and Society,
Calverton, Maryland, USA

Corresponding author address: Luis Gustavo, de Goncalves, ESSIC/UMD and NASA/GSFC, NASA Goddard Space Flight Center, Code 614.3, Greenbelt Rd
Greenbelt, MD 20771.
E-mail: luis.g.degoncalves@nasa.gov

ABSTRACT

The definition and derivation of a 5-year, 0.125º, 3-hourly atmospheric forcing dataset for the South America continent is described which is appropriate for use in a Land Data Assimilation System and which, because of the limited surface observational networks available in this region, uses remotely sensed data merged with surface observations as the basis for the precipitation and downward shortwave radiation fields. The quality of this data set is evaluated against available surface observations. There are regional difference in the biases for all variables in the dataset, with biases in precipitation of the order 0-1 mm/day and RMSE of 5-15 mm/day, biases in surface solar radiation of the order 10 W/m^2 and RMSE of 20 W/m^2, positive biases in temperature typically between 0 and 4 K, depending on region, and positive biases in specific humidity around 2-3 g/Kg in tropical regions and negative biases around 1-2 g/Kg further south.

1. Introduction

Land Surface Models (LSMs) are an important component of Numerical Weather Prediction (NWP) and Global Climate Models and also used in surface hydrology assessments. They provide description of the soil-vegetation system which is the lower boundary condition to the atmosphere and provides the feedback to the atmosphere from the underlying land surface. Several studies have shown that surface storage of water and energy is important in land-atmosphere systems at regional and global scale (e.g. Betts et al., 1996; Koster and Suarez, 1999; Fennessey and Shukla, 1999; Koster et al., 2004; de Goncalves et al., 2006a) and that surface states, such as soil moisture and temperature, can impact atmospheric numerical model predictions.

There are continuing efforts to increase the accuracy (and, as a result, complexity) of the representation of the processes involved in the soil-vegetation-atmosphere system. However, realistic results will only ensue is these models are provided with realistic forcing data. Such forcing data typically comprises air temperature and humidity, wind speed, surface pressure, radiation and precipitation, but the number and nature of the forcing variables vary with the purposes of the LSM. These atmospheric forcing data may be provided from surface and remotely sensed observations, may be model derived, or may be a combination of both modeled and observation information if this is advantageous. Land Data Assimilation Schemes (LDAS: Mitchell et al., 2004) have been successfully employed to provide improved initial surface fields of soil moisture in near real time, for use in predictive meteorological models and to address land-surface management issues. LDAS comprise two-dimensional arrays of LSMs set up to match the grid squares used in the predictive model, which are then forced by model-derived

near-surface fields supplemented, to the maximum extent possible, with surface observations of meteorological variables.

An important challenge when implementing LDAS is the scarcity of comprehensive land-surface data at the spatial and temporal resolutions at which they operate (Maurer et al., 2002). Providing adequate observations of precipitation is particularly problematic because precipitation is so spatially variable and often only point sample data from well-separated rain gauges are available. Some regions of the globe (e.g., North America, Europe, and Japan) have reasonably dense observational coverage. However, others do not, including significantly in the context of the present paper, South America, which has very sparse surface data coverage that is biased toward populated centers near the edge of the continent or along the main river courses where important cities are located (de Goncalves et al., 2006b). Currently, LDAS modelers must rely heavily on atmospheric analyses and remote sensing products for forcing in these regions (Rodell et al., 2004).

This paper describes the creation of forcing data appropriate for use in the South American Land Data Assimilation System (SALDAS) where surface data is limited, using South America as an example, and the validation of the resulting dataset against the observations that are available in this region. These same data have also been adopted as regional forcing data for the model comparisons that are being made in the Large-scale Biosphere Atmosphere experiment Model Intercomparison Study (LBA-MIP; see protocol at http://www.climatemodeling.org/lba-mip/)

2. SALDAS Forcing Data

The SALDAS forcing data are a combination of atmospheric fields necessary for land surface modeling for South America which are derived by combining modeled and observation based sources.

a. Model-Calculated Data

The forcing data cover the entire continent of South America and are build around the model-calculated values of air temperature, wind speed and specific humidity at 2m, surface pressure, downward shortwave and longwave surface radiation, and precipitation from South American Regional Reanalysis (SARR). These SARR data (Aravequia et al. 2007), which were released in 2006 by CPTEC/INPE (Centro de Previsão do Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais), are a medium-term, dynamically consistent, high-resolution, high-frequency, atmospheric dataset covering South American. Currently they are available for a 5-year period from 2000 to 2004. The SARR data are derived using the modified version of the Eta model (Chou and Herdies, 1996) and the Regional Physical-space Statistical System (RPSAS) data assimilation scheme applied at 40Km horizontal resolution and 38 vertical levels. This system integrates upper air and surface observations from several sources over South America, including vertical soundings from the RACCI/LBA and SALLJEX field campaigns over the Amazon and the low-level jet regions along the Andes, respectively. The quality of the reanalysis is assumed to be superior to the operational Eta model analyses because the model and data assimilation systems remained frozen during the

analysis, a larger number of observation were used, and more output fields were saved therefore allowing more comprehensive evaluation.

The topography used in the Eta model when calculating the SARR differs substantially from the SALDAS topography which is derived from USGS GTOPO30 global 30 second elevation map (Row, Hastings, and Dunbar, 1995), and adjustments in the air temperature, humidity, surface pressure and downward longwave radiation are required to allow for these differences in altitude. The air temperature and surface pressure at 2 m are adjusted using the standard vertical atmospheric lapse rate, specific humidity is adjusted by assuming a constant relative humidity between the two elevations, and the longwave radiation is corrected based on the ratio between vapor pressure at the two levels and temperature between the two levels applied to the Stefan-Boltzmann equation. For a more detailed description of the elevation correction procedures, see Cosgrove et al. (2003).

These corrections can be significant and their calculation is an essential step in the calculation of the SALDAS forcing data. The Eta model coordinate system represents topography as steps (Bryan 1969) in order to preserve conserved properties in its finite difference schemes (Mesinger et al. 1988). Consequently, the elevation corrections are greatest in the Andes where rapid changes in altitude induce large step changes in the Eta coordinates. Longwave radiation corrections are up to 20 W/m^2 in regions where there are no abrupt changes in topography (Figure 1a, non-shaded areas), but can be up to 100 W/m^2 in, for example, the Andes (Figure 1a, shaded areas). Similarly, corrections in surface pressure are 30 hPa in fairly flat regions (Figure 1b non-shaded areas) but up to 200 hPa in mountainous regions (Figure 1b non-shaded areas); in specific humidity, 2

g/Kg in fairly flat regions (Figure 1c, non-shaded areas) but up to 10 g/Kg in mountainous regions (Figure 1c, shaded areas), and in temperature, 5 K in fairly flat regions (Figure 1d, non-shaded areas) but up to 20 K for high mountains (Figure 1c, shaded areas) .

Since the main goals of SALDAS is to provide more realistic and accurate datasets over South America than already available from existing global reanalyzes, downward shortwave radiation and precipitation are observation based and derived from GOES satellite measurements and real time TRMM Multisatellite Precipitation Analysis (TMPA-RT) retrievals (Huffman et al, 2007), respectively. In order to reduce the bias of satellite rainfall retrievals, TMPA-RT is combined with surface rain gauges (when available), using additive and multiplicative methods (Vila et al, 2008). The datasets are linearly interpolated in space and time to 1/8 º resolution and 3-hourly frequency, respectively. However, the downward shortwave radiation which is adjusted following changes in the zenithal angle which is expressed as a function of hour of the day and latitude.

b. Downward Shortwave Radiation

The surface solar radiation fluxes used in the SALDAS forcing fields were derived from satellite radiance measurements from the Geostationary Operational Environmental Satellite (GOES)-8 visible imagery using a simplified physical model, GL1.2, developed at the Divisao de Satelites e Sistemas Ambientais (DSA – Division for Satellites and Environmental Systems) in CPTEC (Ceballos et al., 2004). The GL1.2 model considers two broadband spectral intervals for tropospheric radiative transfer: it is

assumed the ultraviolet and visible intervals are essentially non-absorbing and can be processed as a single interval, and that near-infrared intervals have negligible atmospheric scattering and very low cloud transmittance. The current version of the GL1.2 model does not include the effect of aerosols.

In order to make use of GL1.2 data, values are first spatially transposed from their standard 4 km output grid to the 0.125^0 SALDAS grid and then aggregated from their 30-min native temporal resolution to the 3-hour interval used in SALDAS. When GL1.2 data is not available, SARR estimates of downward shortwave radiation is substituted. The monthly average percentage availability of satellite-derived solar radiation data is shown in Figure 2 for the period 2000-2004. The colored bars are the monthly average for 3-hour periods within the day and the black line the all day average. At the time of writing, DSA is reprocessing the GL1.2 data and it is anticipated the percentage availability of satellite-derived downward shortwave radiation data used in the SALDAS forcing dataset is expected to increase once these new data become available.

c. Precipitation

The data sources for the daily surface precipitation observations used in the SALDAS forcing dataset include those provided by the World Meteorological Organization (WMO) supplemented by an INPE compilation of data available from the following agencies:

(a) Agência Nacional de Energia Eléctrica (ANEEL; National Agency for Electrical Energy);

(b) Agência Nacional de Águas (ANA; National Water Agency);

(c) Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME; Meteorology and Hydrologic Resources Foundation of Ceará);

(d) Superintendência do Desenvolvimento do Nordeste (SUDENE; Superintendence for Development of the Northeast);

(e) Departamento de Águas e Energia Elétrica do Estado de São Paulo (DAEE; Department of Water and Electrical Energy for the State of São Paulo), in collaboration with the Centro de Previsão de Tempo e Estudos Climáticos (CPTEC; Brazilian Weather Forecast and Climate Studies Center); and

(f) Technological Institute of Paraná (SIMEPAR).

These surface observations are then combined with the Experimental Real-Time TRMM Multi-Satellite Precipitation Analysis (TMPA-RT), which precipitation retrieval algorithm has been shown to approximately reproducing the histogram of precipitation based on surface observations and also to be reasonably effective in detecting large events at the daily timescale (Huffman et al, 2007). The merging technique used to generate the SALDAS precipitation production includes additive and multiplicative correction schemes to remove the bias of satellite retrievals (Vila et al., 2008). The TMPA-RT data were selected in preference to alternative data sets because, among other characteristics, because it doesn't include rain gauge data. This product (also called 3B42RT), for example, is produced at 3-hourly, 0.25 x 0.25 º resolution, but Rozante et al. (2008) showed that, although 70% of TRMM rainfall estimates correlate with observations with a correlation coefficient in the range 0.5-0.75, TRMM tends to underestimate precipitation by up to 50% and have low correlation (< 0.3) over regions where warm clouds are present, particularly during the austral winter. A Version 6 of the

TRMM 3B42 product is also available. In this case, the biases in satellite estimates are corrected at a monthly time scale using GPCC surface station datasets. However, as shown by Vila et al. (2008), two reasons favor the use of an alternative product for land surface modeling in South America, specifically:

(a) the addition of observations from local South American agencies and Brazilian automatic weather stations means the number of surface observations in the CPTEC/INPE database is about four times larger than GPCC datasets (Figure 3); and

(b) the technique used involves making a daily correction for precipitation bias in contrast to monthly correction in version-6 3B42 data, with consequent better agreement at the daily and sub-daily timescale, resulting in better characterization of diurnal cycles when applied to land surface models which typically run at sub-hourly time-steps.

3. Validation Studies for the SALDAS Atmospheric Forcing Data

Because of the large extent of the modeled domain and limited number of observations available, the validation strategy adopted was to divide continental South America into the sub-regions characterized by different climate regimes shown in Figure 4 (Chou et al., 2002; de Goncalves et al., 2006a). Hereafter these regions are referred to as: NO (North - 47W/83W; 17S/11N); NE (Northeast - 33W/47W; 17S/11N); and CS (Central-South - 33W/83W; 47S/17N). The CPTEC/INPE database has a collection of different surface observation networks over South America from regional and national centers and agencies with private and federal jurisdiction as described in previous

sections. Relevant ground measurements were selected from this data base for each region to evaluate the SALDAS atmospheric forcing data, recognizing that there are likely to be scale differences between the single point observations and model derived gridded datasets.

a. Precipitation

To achieve the best quality precipitation product in SALDAS, all the available surface observations are merged with TRMM data. However, for the purpose of validation, a cross-validation approach adapted from Chen et al (2002) was used in which the gauge reports from 10% randomly selected groups of climate stations were withdrawn from the merging process, with the remaining 90% of the stations then used in the bias-removal process. This procedure was repeated 10 times so that each gauge was withdrawn once. The bias-corrected TRMM precipitation estimate was then compared with the corresponding observation to examine the performance of the merging and bias correction technique (Vila et al., 2008). For the year 2004, the evaluation of the merged SALDAS precipitation was also compared against a similar evaluation of TRMM 3B42RT (real time) precipitation product the same 10% subset of observations. The mean monthly bias expressed in mm/day for TRMM (purple bars) and SALDAS (blue bars) is show as a function of the time of year in Figure 5, for the entire South American continent and each of the regions NO, NE and CS. The solid line in this figure is the observed mean monthly precipitation for the whole continent and the separate regions, also expressed in mm/day.

For South America (SA) as a whole (Figure 5a, upper left), the bias in SALDAS precipitation data are substantially better relative to observations than the TRMM 3B42RT precipitation data during the southern hemisphere spring and summer (i.e. in the wet season), while in the winter when the observed rainfall is also smaller, the bias is small in both cases. This general behavior is observed in all regions (NO, NE and SC) but each region differs to some extent, with SC region (Figure 5d) most similar to SA as a whole because this is the largest region and most of gauge stations are located in this region. The precipitation products northeast region (Figure 5c) both show a systematic negative bias consistent with the results of Rozante et al. (2008) and Vila et al (2008), but there is a relative improvement during the first part of the year when gauge data is included in the SALDAS analysis. The biases in the merged data is worst in northern region (Figure 5b), presumably because gauge stations are very sparse and unevenly distributed in this region and the impact gauge correction is therefore inhibited.

Figure 6 shows the monthly mean Root Mean Square Error (RMSE), expressed in mm/day for the same regions and same time period of time using a similar analysis to that made for bias. There is an improvement when gauge data is merged with the 3BR2RT product in all regions throughout the year, with most improvement in the southern region (Figure 6d) and limited benefit when gauge data is included in the northern region (Figure 6b). Consistent with the results for bias and RSME, the correlation coefficient is systematically improved by merging with gauge data in all regions (Figure 7).

b. Downward Shortwave Radiation

Ceballos et al. (2004) evaluated the GL 1.2 model retrievals against surface data during the year 2002. Three precision pyranometers were used to provide measured daily irradiation representative of rural, urban industrial, and urban coastal areas, while 90 automatic stations using Li-Cor pyranometers in the CPTEC network (http://www.cptec.inpe.br) provided irradiation data for monthly comparisons. The daily means biases were of the order of 5 W/m^2, with standard deviation of ~15 W/m^2, while the monthly means showed a bias of approximately 10 W/m^2 with standard deviation of less than 20 W/m^2. The larger errors were found in highly industrialized or heavily agricultural areas with a high concentration of aerosols. Bearing in mind that the SALDAS downward solar radiation data is a derivative of the original GL1.2 dataset (aggregated from 30 minute and 0.04º to 3 hour and 0.125 º resolution, see section 2) and also involves inclusion of SARR radiation fields when GOES data is missing, a re-evaluation was performed to establish the acceptability of the SALDAS data.

During year 2004, the SALDAS solar radiation data were compared against daily average values from the automatic station network described by Ceballos et al. (2004) with the results evaluated separately over each of the CS, NE and NO regions in the top, middle and bottom panels of Figure 8, respectively. The bars in the left hand side panels in Figure 8 express the mean monthly bias over the region in W/m^2 with the error bars representing standard deviation in this bias, while the dashed lines in these figures is the RMSE, also in W/m^2. There is an annual variation but, on average, the SALDAS forcing data tend to overestimate observed radiation in the CS and NO regions but to underestimate observations in the NE region. Nonetheless, the correlation between SALDAS daily averages and observations shown coefficient given in the panels on the

right of Figure 8 is quite good, with a correlation coefficient greater than 0.7 and 0.82 in the CS region. The variation in the bias through the year in CS maintains has the same pattern as that reported by Ceballos et al. (2004), with overestimation during the winter months associated with increased pollution. In the NO region, radiation is overestimated, which Ceballos et al. (2004) suggests is due partly to the high concentration of aerosols during the biomass burning season (not included in the model) and partly due to errors in the retrieval algorithm when atmospheric precipitable water is high. Previous studies of satellite estimates of solar radiation (Whitlock et al., 1995; Pinker et al., 2001; Stackhouse et al., 2001) have reported a mean deviation of ±10 W/m^2 with a standard deviation less than 20 W/m^2 for grid squares hundreds of kilometers on the side. However, for SALDAS grid cells (of the order of 12 km x 12 km) the monthly mean errors exceed these values in a few instances. In addition, there is additional error when SARR estimates are used to replace missing GOES data because the Eta model parameterization is known to underestimation cloud cover and consequently also to overestimate the downward shortwave radiation reaching the surface.

c. Temperature and Specific Humidity

Figures 9 and Figure 10 show comparisons between SALDAS data and observations for 2 m temperature and specific humidity expressed in terms of the bias and standard deviation calculated for South America as a whole and for separate regions, as in previous sections. The monthly mean daily temperature shows significantly different results in different regions, with a relatively smaller bias and RMSE in the NO region but larger values in the NE region (Figure 9). In the NO region, the bias in temperature is

small and it has little seasonality because there is little variation in annual temperature in this region. In the semiarid NE and sub-tropical CS regions, there is a clear seasonality in the bias and in the CS region in particular, there is a negative bias during the austral winter. The overall bias in temperature integrated over the entire continent reflects this seasonal dependency, with mean monthly values reaching 2K.

SALDAS near surface specific humidity data consistently overestimates observations by about 2-3 g/Kg in NE and NO regions throughout the year but consistently underestimates by about 1-2 g/Kg in the CS region. This pattern may be related to shortcomings in the SARR atmospheric water distribution or meridional advection between tropical and sub-tropical South America.

4. Summary and Discussion

A 5-year, 0.125º, 3-hourly atmospheric forcing dataset was derived for the South America continent in support the South American Land Data Assimilation System (SALDAS) initiative which can be used for a variety of applications including weather and climate simulations and water management. The backbone of the resulting product is the South American Regional Reanalysis (SARR) data, but this supplemented by remotely sensed data merged with surface observations as the basis for the precipitation and downward shortwave radiation fields. The quality of the forcing data sets was evaluated against available surface observations, recognizing the limited observing network in South America. There are regional difference in the biases for all variables, with biases in precipitation typically of the order 0-1 mm/day and RMSE of 5-15 mm/day, biases in surface solar radiation typically of the order 10 W/m^2 and RMSE of

20 W/m^2, positive biases in temperature typically between 0 and 4 K depending on region, and positive biases in specific humidity around 2-3 g/Kg in tropical regions and negative biases of around 1-2 g/Kg further south.

Acknowledgements

This study was supported by the NASA–LBA Ecology (Group CD36) Project under grant NNX06AG91G and NASA Postdoctoral Program under the Oak Ridge Associated Universities (ORAU).

References

Aravéquia, J. A.; Herdies, D. L.; Sapucci, L. F.; Andreoli, R. V.; Ferreira, S. F.S.; Gonçalves, L.G.G. (2007), Reanálise Regional 2000-2004 sobre a América do Sul com o Modelo RPSAS/ETA: Descrição do Experimento e dos Produtos Derivados. *Boletim da SBMET.* V.31 n.02.

Betts, A. K., J. H. Ball, A. C. M. Beljaars, M. J. Miller, and P. A. Viterbo (1996), The land surface-atmosphere interaction: A review based on observational and global modeling perspectives, *J. Geophys. Res.*, **101(D3)**, 7209–7225

Bryan, K. (1969), A numerical method for the study of the circulation of the World Ocean, *J. Comput. Phys.*, **4***,* 347–376

Chou, S. C.; Herdies, D. L., 1996: Test runs using Eta model over South America. In: 15th Conference on weather Analysis and Forecasting, 1996, Virginia. *15th conference on weather analysis and forecasting*, 1996. v. 1.

Ceballos, J. C., M. J. Bottino, and Souza, J.M. (2004), A simplified physical model for assessing solar radiation over Brazil using GOES 8 visible imagery, pp. 1–14*, J. Geoph. Resear.*, vol. **109**, D02211

Cosgrove, B.A., Dag Lohmann, Kenneth E. Mitchell, Paul R. Houser, Eric F. Wood, John C. Schaake, Alan Robock, Curtis Marshall, Justin Sheffield, Qingyun Duan, Lifeng Luo, R. Wayne Higgins, Rachel T. Pinker, J. Dan Tarpley, and Jesse Meng, (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research*, Vol. **108**, No. D22, 8842, doi:10.1029/2002JD003118

de Goncalves, L. G. G., W. J. Shuttleworth, S. C. Chou, Y. Xue, P. R. Houser, D. L. Toll, J. Marengo, and M. Rodell (2006a), Impact of different initial soil moisture fields on Eta model weather forecasts for South America, *J. Geophys. Res.*, ***111***, D17102, doi:10.1029/2005JD006309

de Goncalves, L. G. G., W. J. Shuttleworth, B. Nijssen, E. J. Burke, J. A. Marengo, S. C. Chou, P. Houser, and D. L. Toll (2006b), Evaluation of model-derived and remotely sensed precipitation products for continental South America, *J. Geophys. Res.*, ***111***, D16113, doi:10.1029/2005JD006276.

Fennessey, M. J., and J. Shukla. 1999. Impact of initial soil wetness on seasonal atmospheric prediction*, Journal of Climate* **12**: 3167-3180.

Huffman G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker and D. B. Wolff (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales*, J. Hydrometeor.*,**8**, 38-55.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, M. Kistler, A. Berg, R. Reichle, M. Rodell, and J. Famiglietti (2004), Realistic initialization of land surface states: impacts on subseasonal forecast skill, *J. Hydromet*., **5 (6),** 1049-1063

Koster, R.D., and M.J. Suarez, (1999), A simple framework for examining the interannual variability of land surface moisture fluxes. *J. Clim*., **12 (7)** 1911-1917.

Liou, K.-N. (1980), An Introduction to Atmospheric Radiation. Academic Press. Londres, 392pp.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *J. Clim*., **15**, 3237–3251.

Mesinger, F., T. L. Black, and Z. I. Janjic (1988), A summary of the NMC step mountain (Eta) coordinate model, paper presented at Workshop on Limited-Area Modeling Intercomparison, Natl. Cent. for Atmos. Res., Boulder, Colo.

Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *J. Geophys. Res*., **109**, D07S90,doi:10.1029/2003JD003823.

Pinker, R. T., I. Laszlo, and B. Zhang (2001), Pathfinder large scale radiative fluxes: Data availability and their use in climate research, in IRS 2000: Current Problems in Atmospheric Radiation, pp. 481 – 484, A. Deepak, Hampton, Va.

Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.-J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, The Global Land Data Assimilation System, Bull. *Amer. Meteor.* Soc., **85 (3),** 381-394, 2004.

Rozante, J.R., S.S. da Costa, L.G.G. de Goncalves and D. Vila, Comparison between satellite and ground-based rainfall measurements: validation of the TRMM monthly and annual daily mean precipitation over South America. Submitted to Journal of Hydrology.

Row, L.W., Hastings, D.A., and Dunbar, P.K., 1995. TerrainBase Worldwide Digital Terrain Data - Documentation Manual, CD-ROM Release 1.0. National Geophysical Data Center, Boulder, Colorado.

Stackhouse, P. W., S. K. Gupta, S. J. Cox, M. N. Chiacchio, and J. C. Mikovitz (2001), The WCRP/GEWEX Surface Radiation Budget Project release 2: An assessment of surface flux at 1 degree resolution, in IRS2000: Current Problems in Atmospheric Radiation, pp. 485 – 488, A. Deepak, Hampton, Va.

Vila, D., L.G.G. de Goncalves, J.R. Rozante and D.L. Toll, Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimations over Continental South America. *Submitted to Journal of Hydrometeorology*.

Whitlock, C. H., T. P. Charlock, W. F. Staylor, R. T. Pinker, I. Laszlo, A. Ohmura, H. Gilgen, T. Konzelman, R. C. Di Pasquale, C. D. Moats, S. R. LeCroy, and N. A. Ritchey (1995), First global WCRP shortwave surface radiation budget dataset, Bull. *Am. Meteorol. Soc*., **76**, 905– 922.

List of Figures

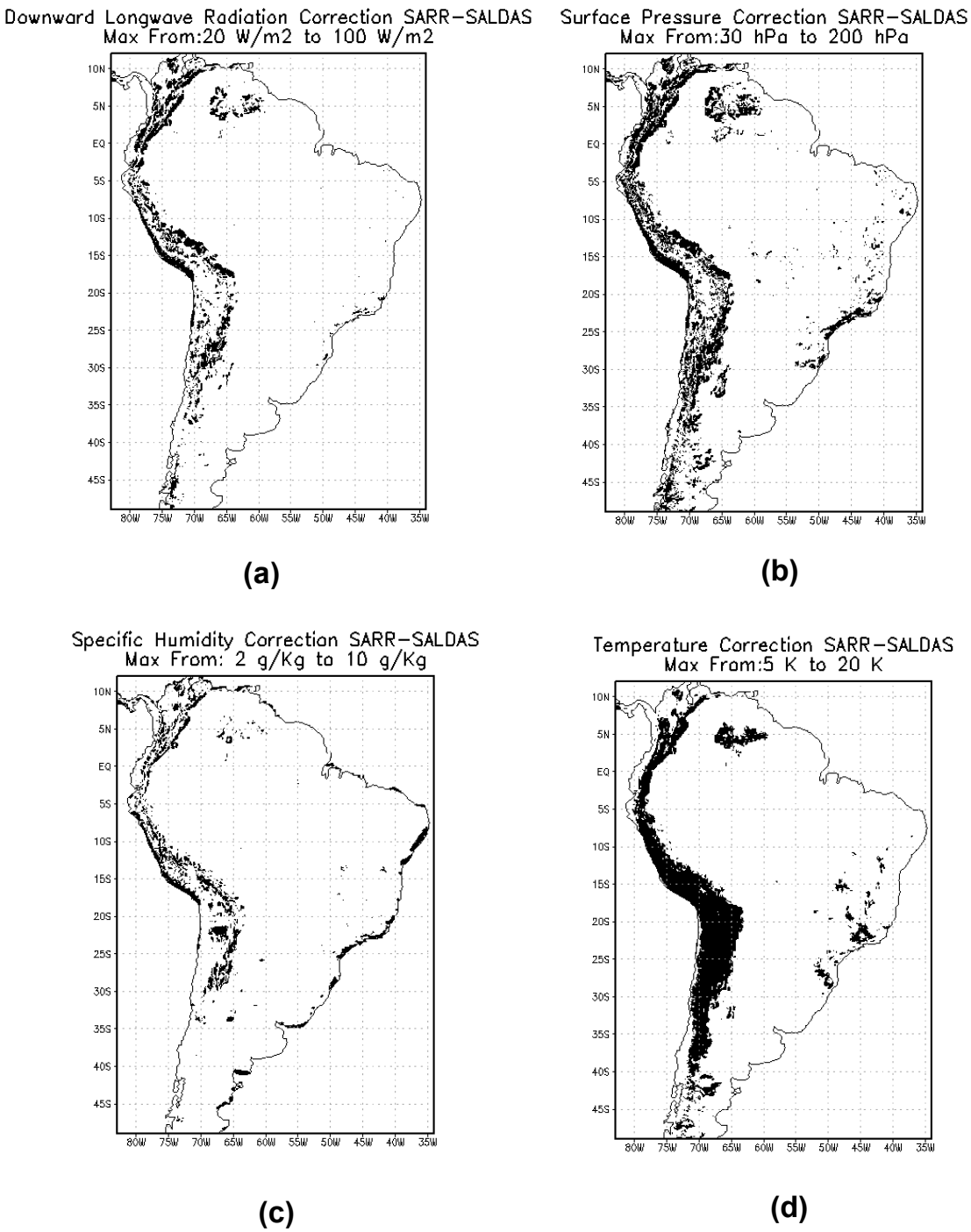

FIG 1. Elevation correction between SARR Eta vertical coordinate and SALDAS topography for downward longwave radiation at surface (a), surface pressure (b), surface specific humidity (c) and surface air temperature (d).

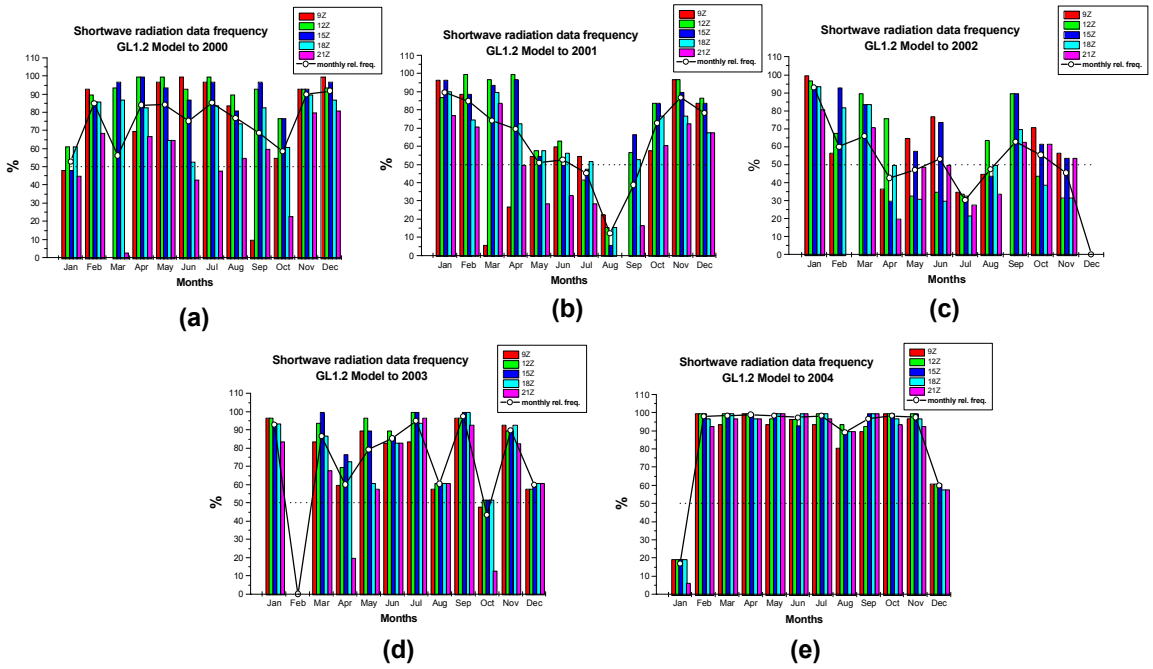

FIG 2. Monthly percentage of GOES/GL 1.2 satellite retrievals contributed to SALDAS forcing downward shortwave radiation from 2000-2004 [(a)-(e) respectively]

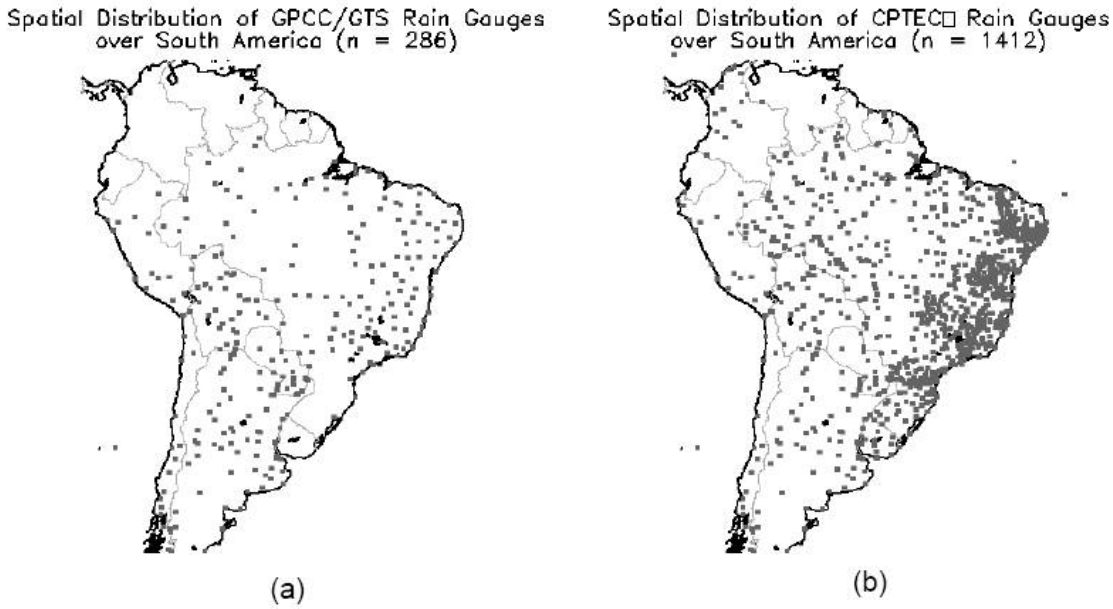

FIG 3. Typical GPCC/GTS (a) and CPTEC database rain gauges distribution over continental South America during the period of study (2000-2004).

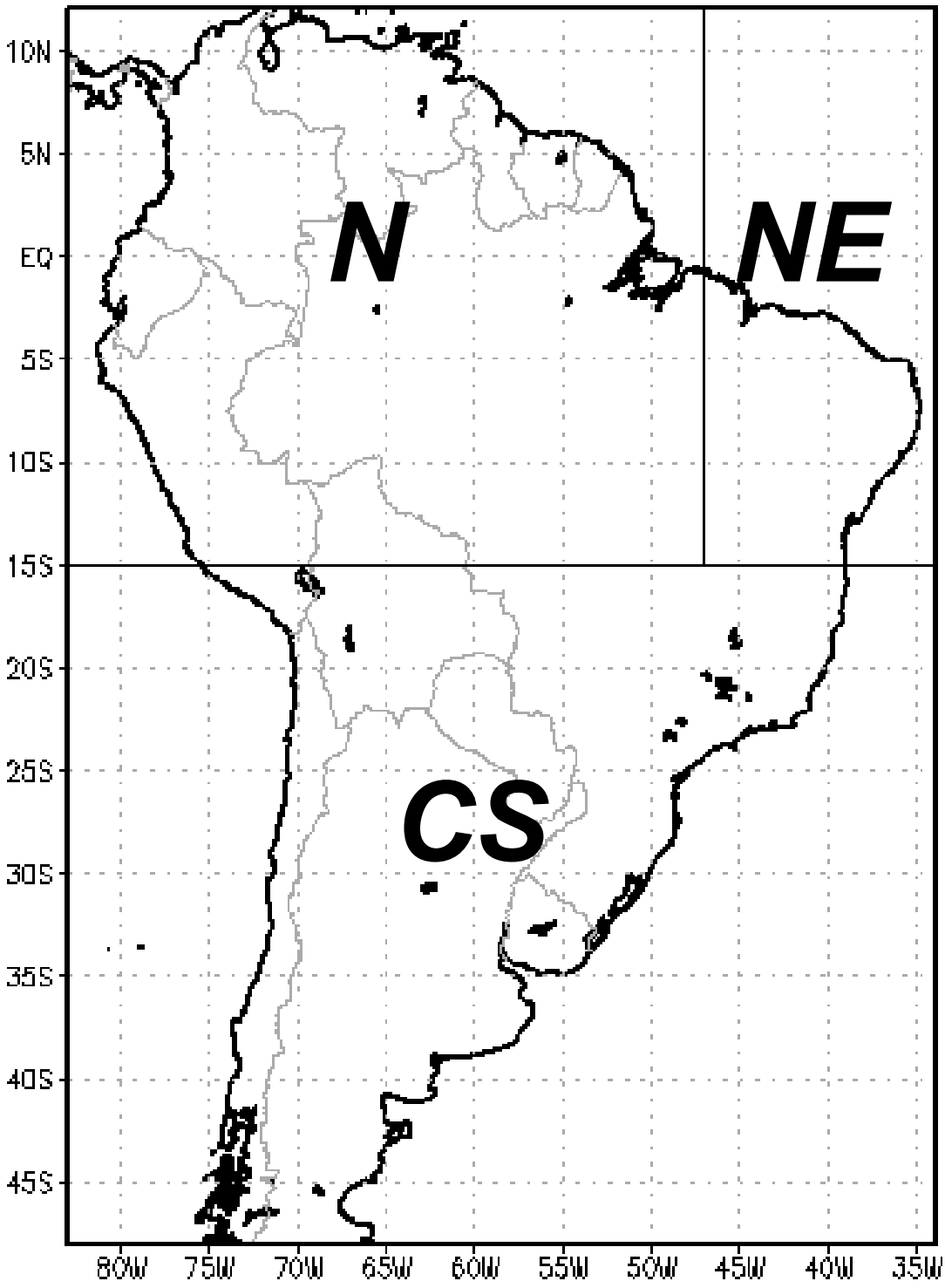

FIG 4. Division of the continent in three distinct climatic regions (N, NE and CS) based on annual precipitation regimes.

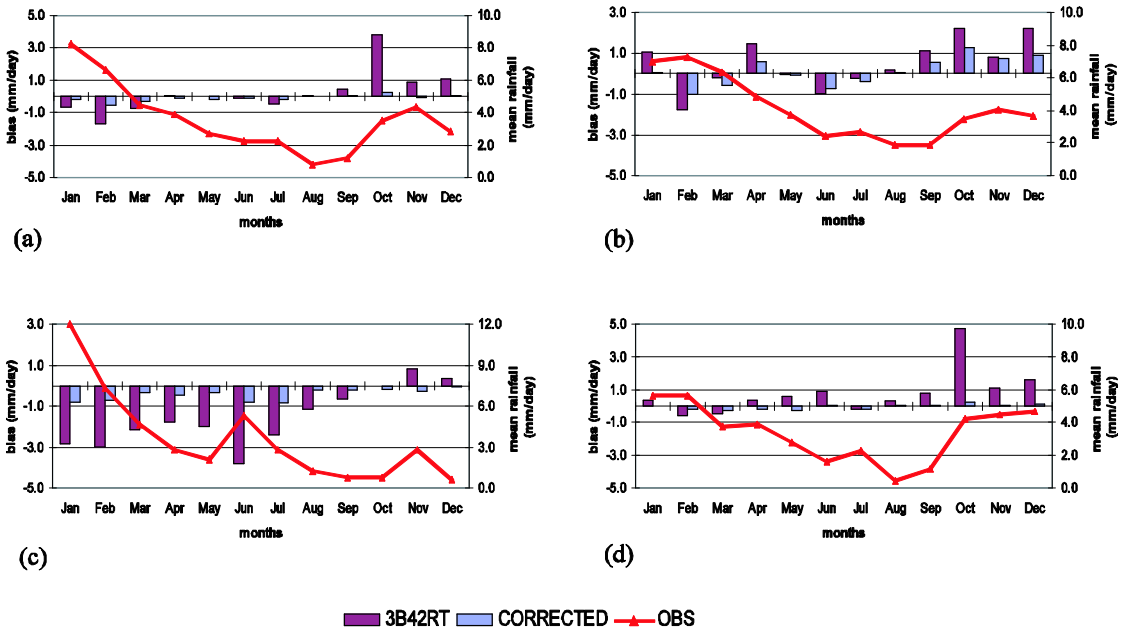

FIG 5. Mean monthly SALDAS (blue bars) and TRMM (purple bars) precipitation in mm/day as a function of the time of the year for South America (a), North (b), Northeast (c) and Central-South (d) regions. The solid line is the observed mean monthly precipitation also in mm/day (scales on the right y-axis on each panel).

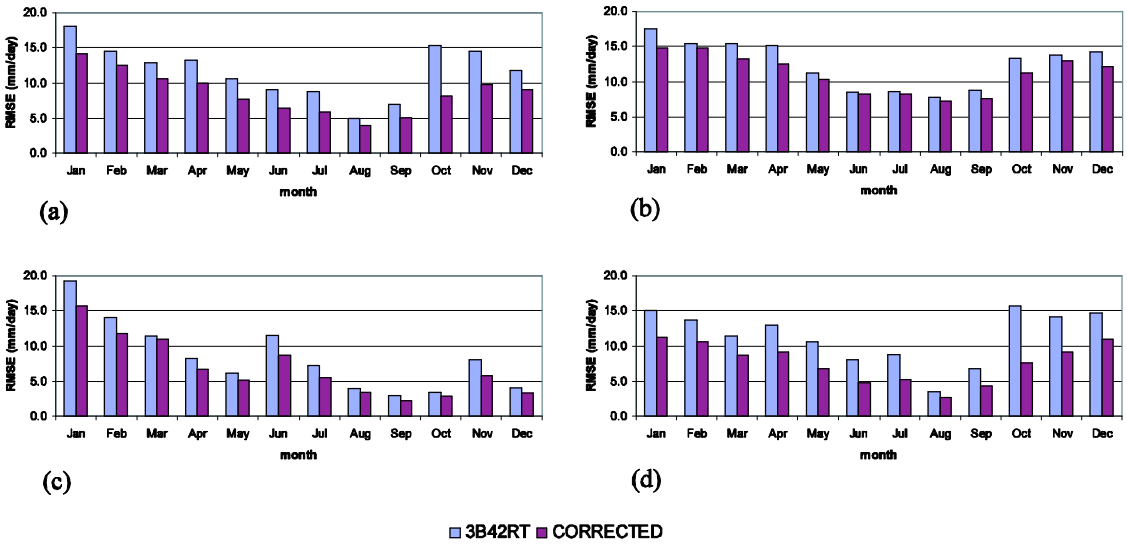

FIG 6. Monthly mean Root Mean Square Error (RMSE) for SALDAS (blue bars) and TRMM (purple bars) precipitation, expressed in mm/day for the same regions and period shown on Figure 5.

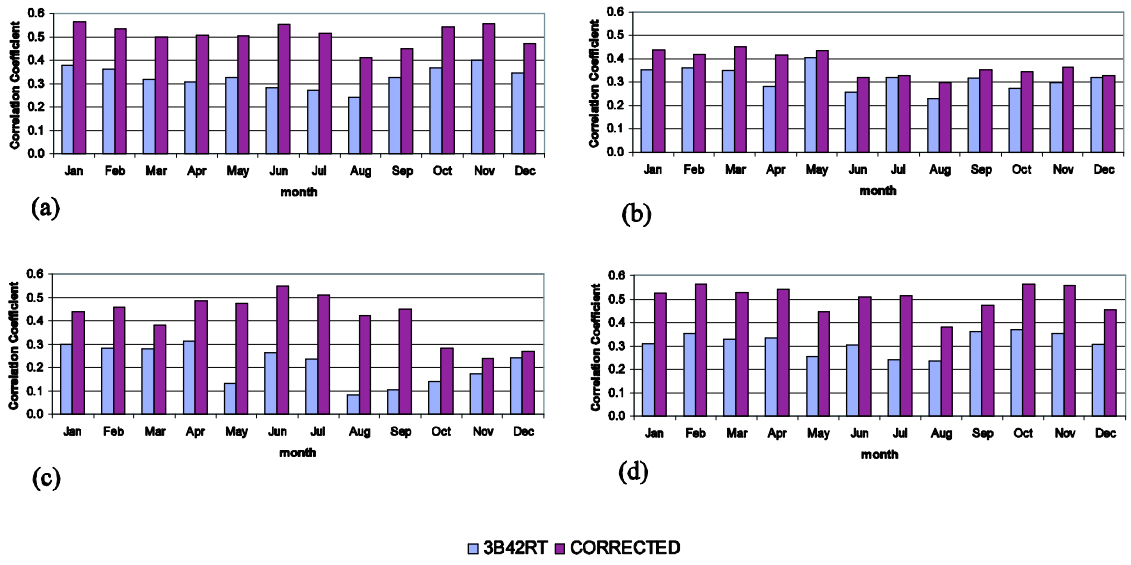

FIG 7. Same as Figure 6 except that values are for correlation coefficient.

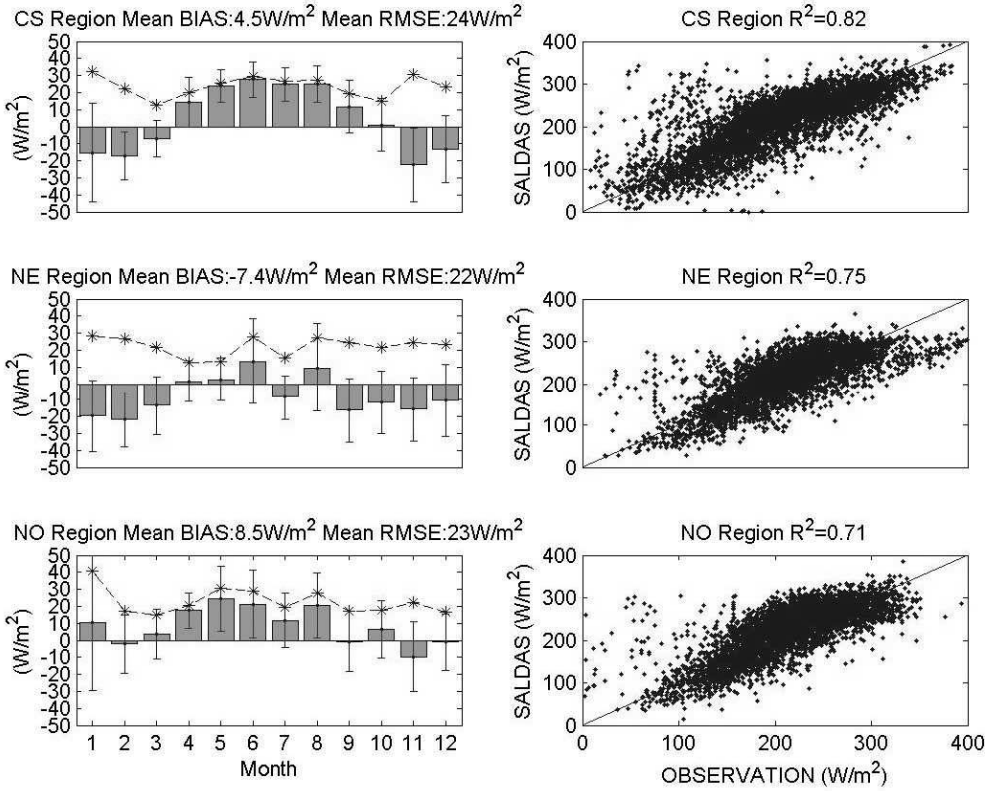

FIG 8. On the left column: year 2004 downward shortwave radiation mean monthly bias (shaded bars), bias standard deviation (error bars) and RMSE (dashed line) in W/m2 for CS (top), NE (middle) and NO (bottom) regions. On the right column: daily correlation coefficient between observed and SALDAS downward shortwave radiation in W/m2 for CS (top), NE (middle) and NO (bottom) regions.

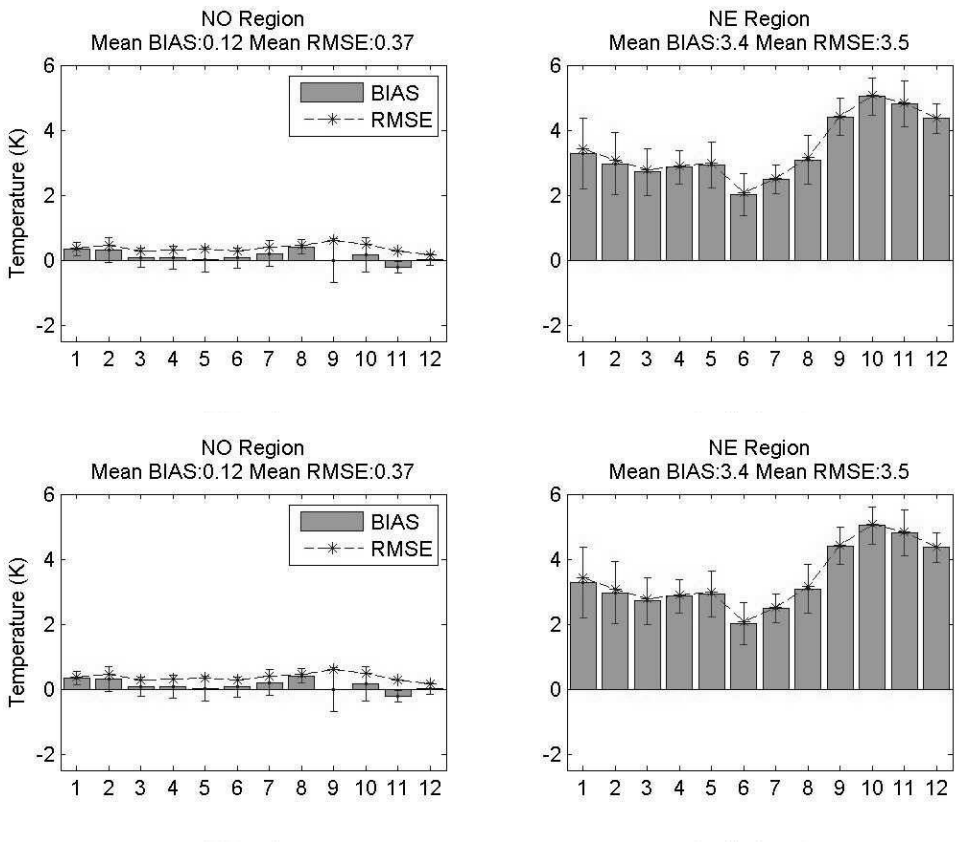

FIG 9. Year 2004 mean monthly temperature bias (shaded bars), bias standard deviation (error bars) and RMSE (dashed line) in K for NO, NE, CS and South American regions.

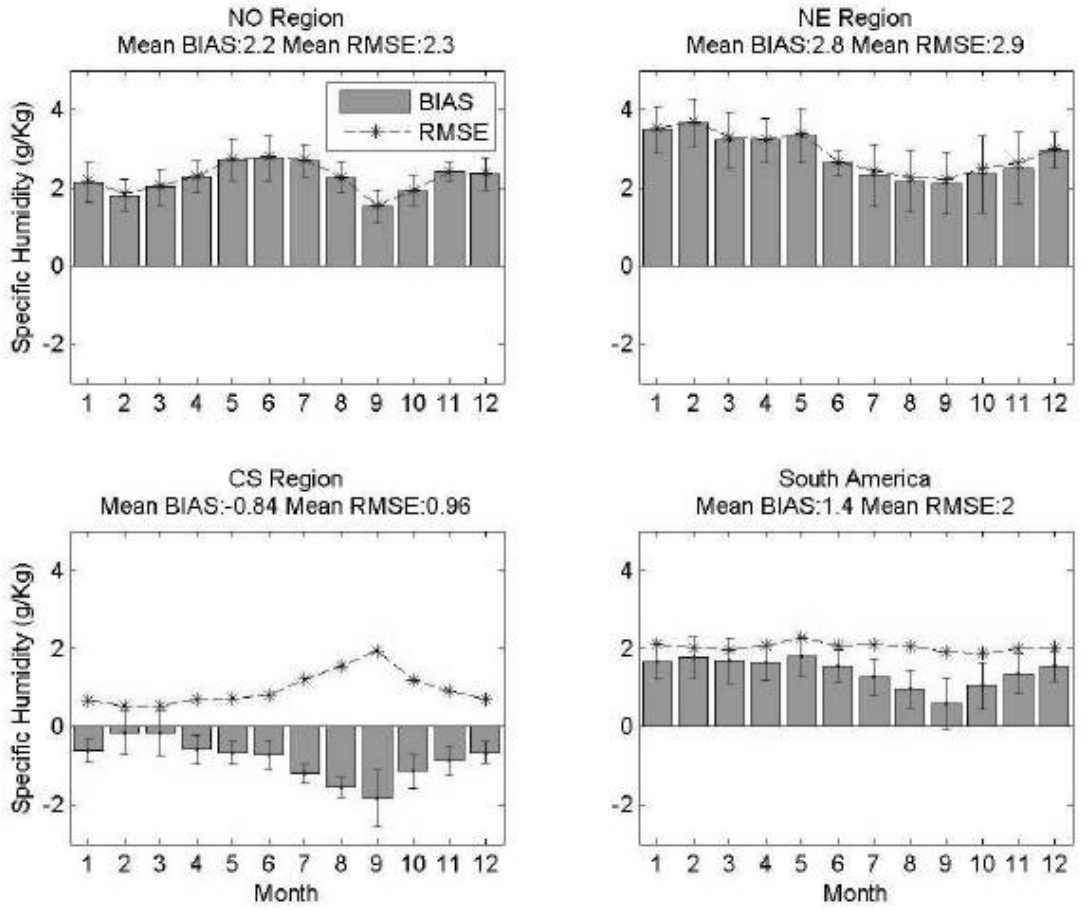

FIG 10. Year 2004 mean monthly specific humidity bias (shaded bars), bias standard deviation (error bars) and RMSE (dashed line) in g/Kg for the NO, NE, CS and South American regions.