TRMM rainfall data estimation over the Peruvian Amazon-Andes basin and its assimilation into a monthly water balance model

WALDO SVEN LAVADO CASIMIRO[1,2], DAVID LABAT[2], JEAN LOUP GUYOT[3], JOSYANE RONCHAIL[4] & JUAN JULIO ORDOÑEZ[1]

1 *Servicio Nacional de Meteorología e Hidrología, SENAMHI, Casilla 11 1308, Lima 11, Perú*
wlavado@senamhi.gob.pe

2 *LMTG-Université de Toulouse-CNRS-IRD-OMP, 14 Avenue Edouard Belin, F-31400 Toulouse, France*

3 *LMTG-Université de Toulouse-CNRS-IRD-OMP, IRD, CP 7091 Lago Sul, 71619-970 Brasília DF, Brazil*

4 *Université Paris 7 (IRD)* and *LOCEAN, Boite 100, 4 Place Jussieu, F-75252 Paris Cedex 05, France*

Abstract The Peruvian Amazon-Andes basin corresponds to about 10% of the total Amazonian basin and is characterized by sparse rainfall data, particularly over the lowland zone (rainforest). We compare the 3B43 product of the Tropical Rainfall Measuring Mission (TRMM) with raingauge data over two sub-basins (Urubamba and Tambo) in the Ucayali basin located in the Peruvian Amazon-Andes basin. The spatial distribution of the 3B43 product is 0.25° × 0.25° (approx. 27.8 × 27.8 km) and data are at a monthly scale. The period of comparison between on-site rainfall and 3B43 TRMM data is from January 1998 to December 2007. Comparison between on-site rainfall observations and 3B43 product is carried out using correlation coefficient and relative error. Improvement of the TRMM rainfall data is then proposed based on on-site rainfall data. After analysis of the 3B43 product, three sets of distributed rainfall data (*in situ*, TRMM and on-site TRMM improved) were used as input in a GR2M monthly water balance model to simulate discharge at Urubamba and Tambo. Classical statistical overcalibration and validation procedure show a better accuracy in flow simulations, using only original TRMM data over Urubamba basin and improved data over Tambo basin.

Key words Amazon basin; TRMM; Andes; Peru; monthly water balance model

INTRODUCTION

The Amazon Basin (AB) is the largest basin on the planet, with a drainage area of 6 200 000 km^2 and mean annual discharge rate of 209 000 m^3 s^{-1} (~5% of all the above-water lands) (Molinier *et al.*, 1996; Marengo, 2006). AB is one of the regions with the highest rainfall in the world and a major water vapour source (Espinoza *et al.*, 2009). Also, over the AB it is possible to have extreme events such as the dramatic drought observed in 2005 (Marengo *et al.*, 2008; Zeng *et al.*, 2008).

Rainfall stations over the Peruvian Amazon basin (~10% of total Amazonian basin) are poorly distributed. Mean rainfall at basin scale over this region is mostly underestimated (Guyot, pers. comm.). The Rainfall 3B43 product of the Tropical Rainfall Measuring Mission has been used with good results over distinct parts of the world by e.g. Sorooshian *et al.* (2000), Adeyewa & Nakamura (2003), Chiu *et al.* (2006), Dinku *et al.* (2007) and Islam & Uyeda (2007), who compare or improve on-site rainfall land data. Introduction of TRMM rainfall products in hydrological and water balance models has already been documented (see e.g. Wilk *et al.*, 2006; Collischonn *et al.*, 2008; Su *et al.*, 2008).

In Peru measurements of hydro-climatological variables over the Amazon basin are administered by the National Meteorology and Hydrology Service SENAMHI (http://www.senamhi.gob.pe). Most of the hydrometeorological data are related to climate variables (rainfall, temperature, etc.). Since 2001, the HYBAM project (http://www.ore-hybam.org) has carried out flow measurement missions over the largest Amazon-Andes rivers in the Peruvian territory (Yerren *et al.*, 2004; Lavado *et al.*, 2009c). Results of flow measurements along with meteorological observation from SENAMHI have, for the first time, allowed the development of a complete database over the Peruvian Amazon-Andes region (PAB) which has been used in this study. This complete database has enabled monthly hydrological models to be developed for the first time in the Peruvian Amazon zone (Lavado *et al.*, 2009a,d). This paper compares on-site rainfall data with 3B43

TRMM rainfall data. Improvements of the original 3B43 TRMM data are also compared. The assimilation of these three data sets of rainfall in the GR2M monthly water balance models was evaluated.

AVAILABLE DATA AND METHODS

Two sub-basins (Urubamba and Tambo) in Ucayali basin were used in this study (Fig. 1 and Table 1). Watershed boundaries were plotted at base from Maldonadillo and Tambo hydrological gauging stations. Watershed subdivisions were made using the Digital Elevation Model (DEM) provided by the US National Aeronautics and Space Administration (NASA) through the Shuttle Radar Topography Mission, SRTM (http://www2.jpl.nasa.gov/srtm).

Weather stations used in this work are part of the Peruvian National Meteorology and Hydrology Service, SENAMHI (http://www.senamhi.gob.pe) network: 24 rainfall stations were used (Fig. 1). Rainfall and temperature data (mean, maximum and minimum) are available at the monthly scale and include the period from January 1998 to December 2007.

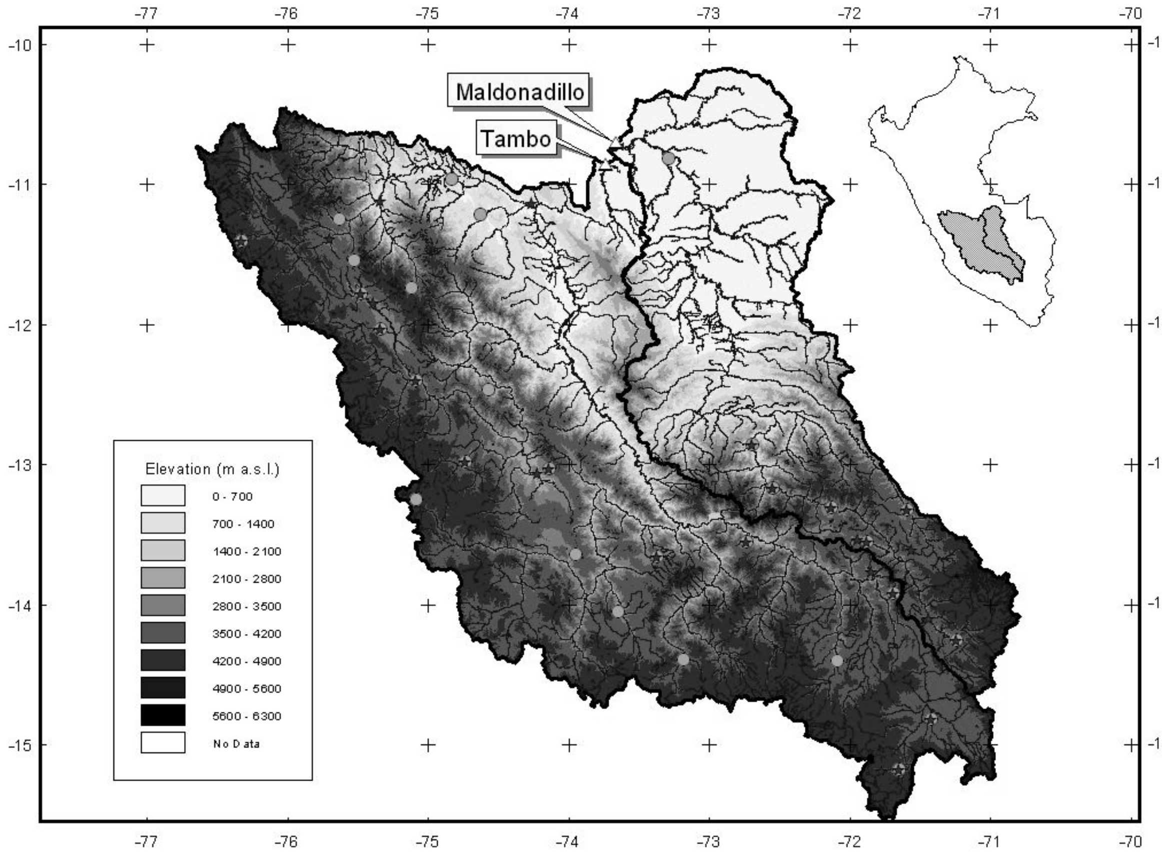

Fig. 1 Location of the study region. Basins boundary takes as a base the Tambo and Maldonadillo gauging station. Circles are rainfall stations, stars are temperature stations.

Table 1 Flow stations, locations, drainage area (Dr. Ar.) and flow record used in this study.

Flow Station	River	Latitude (°S)	Longitude (°W)	Altitude (m a.s.l.)	Dr. Ar. (km^2)	R (mm $year^{-1}$)	E (mm $year^{-1}$)	F ($m^3 s^{-1}$)
Maldonadillo	Urubamba	10.74	73.71	287	60301	985	996	2289
Tambo	Tambo	10.85	73.71	290	135252	754	1020	1824

R: Rainfall; E: Evapotranspiration; F: Flow.

Lavado *et al*. (2009b) presented a satisfactory relationship between the FAO Penman-Monteith and Hargreaves-Samani (HE) methods (Hargreaves & Samani, 1985); the modified Hargreaves-Samani method (HEm) is used for our region of study, given by:

$$\mathrm{HE} = \frac{00023(T + 17.8)(T\max - T\min)^{0.5}\,\mathrm{Ra}}{\lambda} \tag{1}$$

$$\mathrm{HEm} = 0.6484\mathrm{HE} + 1.146 \tag{2}$$

where HEm is evapotranspiration (mm d^{-1}); T is mean temperature (°C); Tmax is maximum temperature (°C); Tmin is minimum temperature (°C); and Ra is extraterrestrial radiation (MJ m^{-2} d^{-1}). The formula to compute Ra is given by Allen *et al*. (1998) and λ is latent heat of vaporization (MJ kg^{-1}); λ in this study is assumed as 2.45 MJ Kg^{-1} (Allen *et al*., 1998).

In order to work with soil data, we have used the methodology described by Dieulin *et al*. (2006), based on the data from the Digital Soil Map of the World (FAO/UNESCO, 1981). Soil water holding capacity is named WHC in this study.

Spatial distribution of the 3B43 TRMM product is at 0.25° × 0.25° (~ 27.8 × 27.8 km) and is at monthly scale since January 1998. The period of comparison between measured point rainfall and 3B43 TRMM product is from January 1998 to December 2007.

Improvement of TRMM rainfall data was made from on-site rainfall stations, using the following equations (Collischonn, 2006):

$$P'(x_i, y_i) = P(x_i, y_i) - P_m(x_i, y_i) \tag{3}$$

where $P'(x_i, y_i)$ is the fluctuation of TRMM with on-site rainfall; $P(x_i, y_i)$ is on-site rainfall estimate located over the cell (x_i, y_i); $P_m(x_i, y_i)$ is the rainfall estimate from TRMM over the cell (x_i, y_i). Conceptually, the methodology interpolates values of P', instead of on-site rainfall $P(x_i, y_i)$, because we consider that the variability of the first is lower than the second. Thus:

$$\hat{P}'(x, y) = \sum_{i=1}^{n} \lambda_i \cdot P'(x_i, y_i) \tag{4}$$

where $\hat{P}'(x, y)$ is the fluctuation (or residual) interpolated in the cell (x, y). For any (x, y) location, TRMM rainfall is improved using the following equation:

$$\hat{P}(x, y) = P_m(x, y) + \hat{P}'(x, y) \tag{5}$$

In summary, three sets of rainfall data were used for our analysis: on-site rainfall (Obs.R); TRMM rainfall data (TRMM); and improved rainfall (Obs.R TRMM).

The GR2M model version used in this study is described by Niel *et al*. (2003), and is based on GR2M (Edijatno & Michel, 1989; Kabouya, 1990), and reviewed by Makhlouf & Michel (1994); this model is a two-reservoir conceptual model.

Rainfall and evapotranspiration are spatially interpolated over cells from 0.25° × 0.25° (~ 27.8 × 27.8 km), based on the kriging method. The soil water holding capacities from the Digital Soil Map of the World (FAO/UNESCO, 1981) were interpolated over our region of study to obtain values at 0.25° × 0.25° grids, following the methodology given by Dieulin *et al*. (2006).

In this study, an automatic calibration is carried out using the Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) that is a modified version of the original SCE-UA global optimization algorithm (Duan *et al*., 1992) using the Metropolis Hastings strategy (Metropolis *et al*., 1953; Hastings, 1970), for detailed description of this method see Vrugt *et al*. (2003).

In hydrological modelling two tests are commonly used: Nash-Sutcliffe efficiency (Nash; Nash & Sutcliffe, 1970) and relative volume error expressed in percent (%V), where $Q_{o,j}$ is the observed value; $\overline{Q_o}$ is the mean of the observed values; $Q_{s,j}$ the model simulated values; and n the number of data points:

$$\text{Nash} = 100 \times \left[1 - \frac{\sum_{j=1}^{n} \left(Q_{o,j} - Q_{o,j} \right)^2}{\sum_{j=1}^{n} \left(Q_{o,j} - \overline{Q_o} \right)^2} \right] \quad (6)$$

$$\%V = 100 \times \left[\frac{\sum_{j=1}^{n} Q_{s,j} - \sum_{j=1}^{n} Q_{o,j}}{\sum_{j=1}^{n} Q_{o,j}} \right] \quad (7)$$

RESULTS

In order to compare TRMM rainfall (3B43 product) and on-site rainfall data from 1998 to 2007, the statistics: Pearson correlation coefficient and relative error ((TRMM – Obs.R)/Obs.R) between both sets of data were compared (see Fig. 2).

In general, similar interpolated maps were estimated using both data sets considering the mean multi-annual for the 1998-2007 period over the two basins analysed. However, under-estimation of rainfall are observed using TRMM data mostly in the lower zones of the Urubamba basins (region with sparse distribution of rainfall stations, see Fig. 1). Correlation coefficients are mostly significant (α = 0.05) with better values over the boundary region of the two basins analysed.

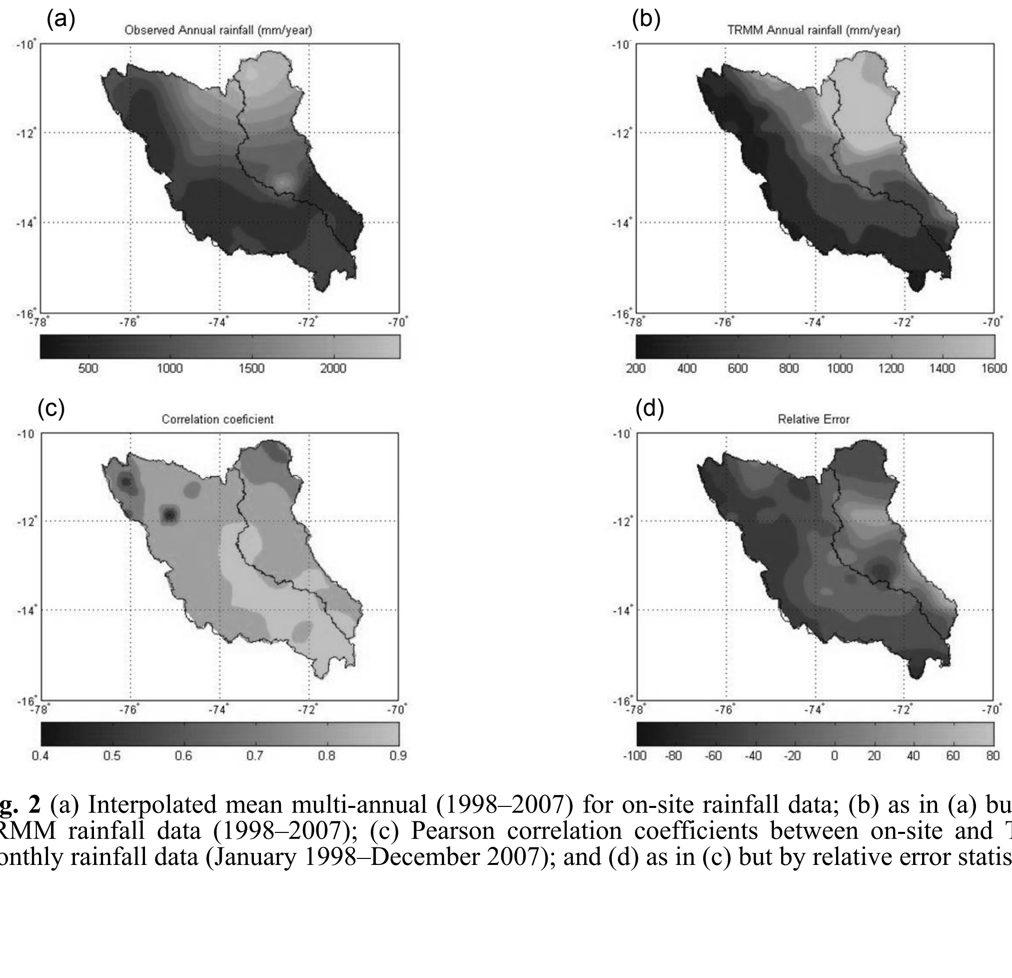

Fig. 2 (a) Interpolated mean multi-annual (1998–2007) for on-site rainfall data; (b) as in (a) but using TRMM rainfall data (1998–2007); (c) Pearson correlation coefficients between on-site and TRMM monthly rainfall data (January 1998–December 2007); and (d) as in (c) but by relative error statistic.

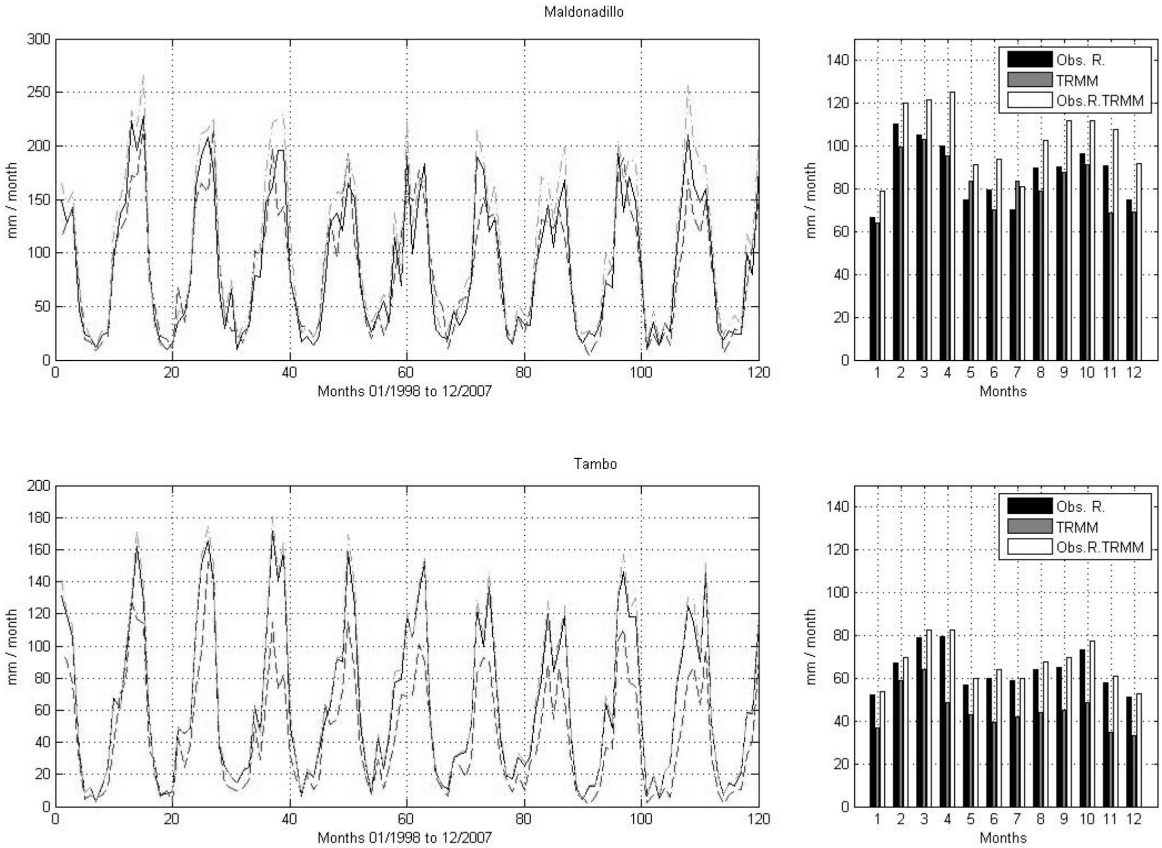

Fig. 3 Top: Comparison of mean monthly rainfall data sets interpolated over Maldonadillo basin. Bottom: Comparison of mean monthly rainfall data sets interpolated over Tambo basin. Black line is on-site rainfall data, dashed line is TRMM rainfall data and dotted dashed line is improved observed rainfall TRMM data.

Underestimation from using TRMM data shows in the interpolated relative error as well; negative values are located mostly over the western zone of the two basins analysed. Figure 3 shows comparison of the mean areal value of the three data sets of rainfall used over the Maldonadillo and Tambo basins; on average, good estimations were exhibited for Maldonadillo basin and underestimations were evident over Tambo basin. Nevertheless, improved rainfall data (Obs.R TRMM) show good relationships with on-site rainfall data.

Table 2 shows statistical Nash and %*V* for the two basins analysed as well as for the three sets of rainfall data. Figures 4 and 5 show parameter samples and hydrographs during the

Table 2 Comparison of Nash and relative volume error (%*V*) for the three sets of rainfall data used. Better values are shown in bold.

Rainfall set	Statistic	Maldonadillo:		Tambo:	
		Cal.	Val.	Cal.	Val.
Obs.R	Nash	0.80	0.87	**0.92**	**0.92**
	%*V*	0.57	3.90	**0.01**	**8.83**
TRMM	Nash	0.84	0.81	0.57	0.40
	%*V*	2.22	7.38	12.23	32.35
Obs.R TRMM	Nash	**0.92**	**0.92**	0.86	0.82
	%*V*	**0.66**	**8.04**	1.35	5.31

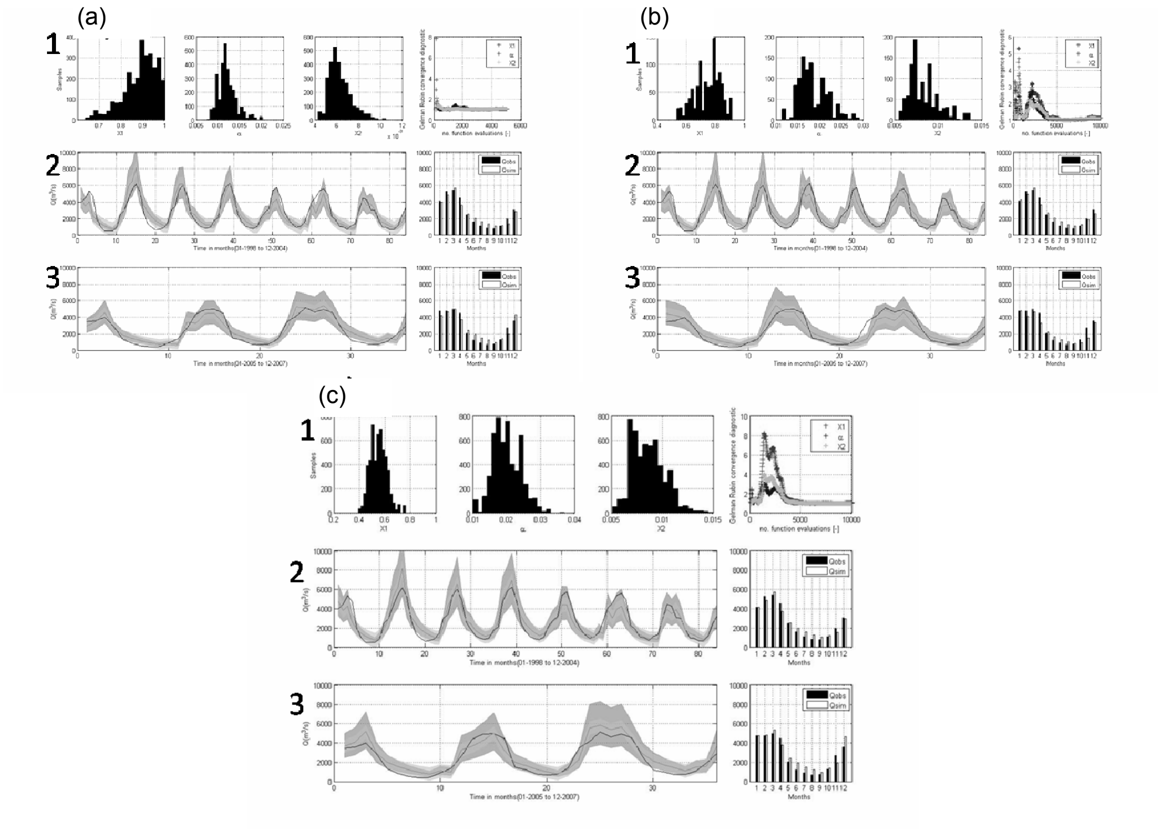

Fig. 4 Calibration and validation procedure for GR2M model over the Maldonadillo basin: (a) Using on-site rainfall data set: 1: histograms of parameters in the model and evolution of the Gelman and Rubin scale reduction. 2: hydrograph prediction uncertainty for the calibration period (01/1998 to 12/2004) and monthly observed (Qobs) and simulated flows (Qsim). 3: hydrograph prediction uncertainty for the validation period (01/2005 to 12/2007) and monthly observed (Qobs) and simulated flows (Qsim). (b) as in (a) but by TRMM rainfall data. (c) as in (a) but for improved observed rainfall TRMM data. Observed flows are represented by the black line. The grey line denotes simulated flow. The dark shaded area denotes the prediction uncertainty that results from parameter uncertainty. The light shaded area denotes the additional prediction uncertainty that results from model and measurement uncertainty.

calibration and validation procedures. In summary, introduction of TRMM rainfall data without improvements into the GR2M model over Maldonadillo basin gave good results in simulating monthly flow (Nash = 0.84 and %V = 2.22 during the calibration procedure), but un-acceptable results over Tambo basin (Nash = 0.57 and %V = 12.23 during the calibration procedure). Using improved rainfall data (Obs.R TRMM) led to good results for both basins.

CONCLUSION

TRMM data and its improved data (Obs.R TRMM) appear to describe well the hydrological regimes over sparse regions in the Peruvian Amazon-Andes, but on-site rainfall data are indispensable for validating and improving rainfall data sets. Future work in this region will focus on improving the TRMM daily-data (3B42 product) using observed information as well as hydrological models at the daily time scale.

Acknowledgements This study was supported by the Institute of Research for Development (IRD), France and the Peruvian National Meteorology and Hydrology Service, SENAMHI. The authors wish to thank all those involved in the TRMM mission as well as those who worked

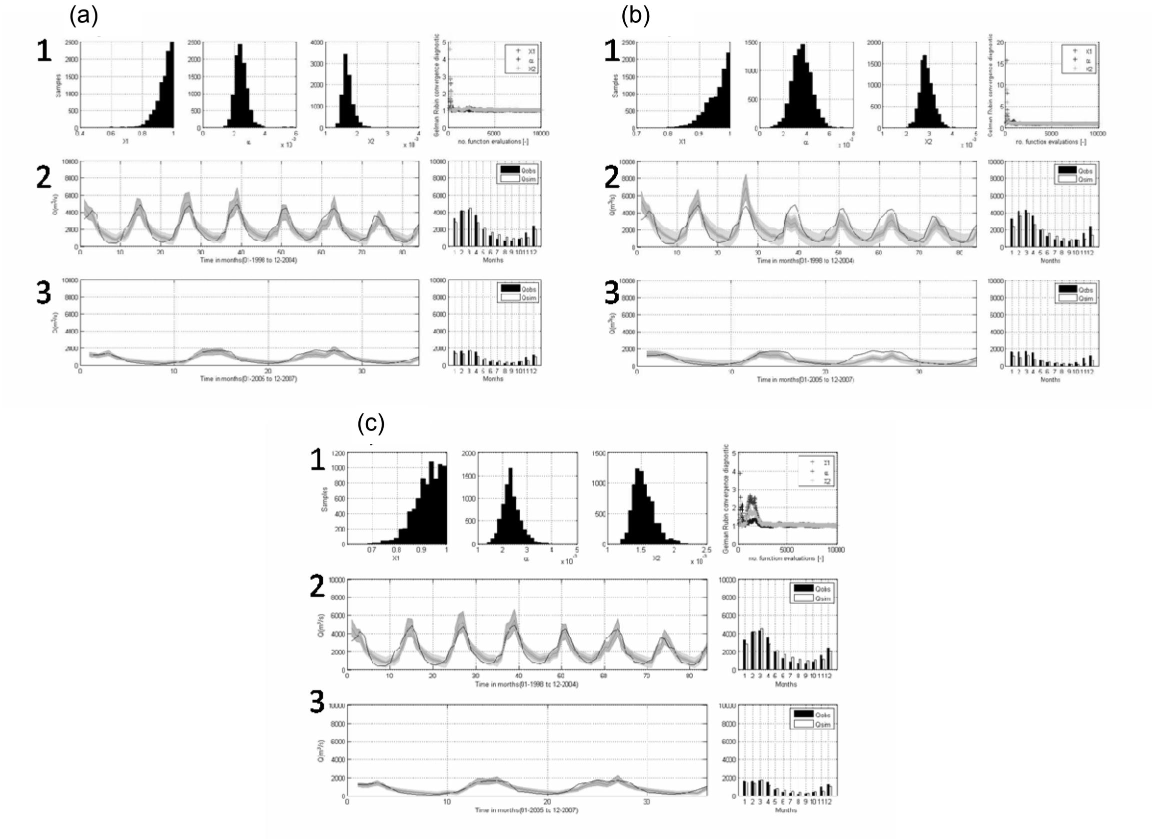

Fig. 5 Calibration and validation procedure for GR2M model over the Tambo basin; details as Fig. 4.

hard to collect data in the field. Thanks to Bruno Collischonn for provide the code to read TRMM data. Dedicated to Flavia.

REFERENCES

Adeyewa, Z. D. & Nakamura, K. (2003) Validation of TRMM radar rainfall data over major climatic regions in Africa. *J. Appl. Met.* **42**(2), 331–347.

Allen, G. R., Periera, L. S., Raes, D. & Smith, M. (1998) Crop evapotranspiration: guidelines for computing crop water requirements. *FAO Irrig. Drain. Paper 56*. Food and Agricultural Orgainzation of the United Nations, Rome, Italy.

Collischonn, B. (2006) Uso de precipitação estimada pelo satélite TRMM em modelo hidrológico distribuído, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.

Collischonn, B., Collischonn, W. & Tucci, C. E. M. (2008) Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *J. Hydrol.* **360**(1-4), 207–216.

Chiu, L. S., Liu, Z., Vongsaard, J., Morain, S., Budge, A., Neville, P. & Bales, C. (2006) Comparison of TRMM and water district rain rates over New Mexico. *Adv. Atmos. Sci*. **23**(1), 1–13.

Dieulin, C., Boyer, J. F., Ardoin-Bardin, S. & Dezetter, A. (2006) The contribution of GIS to hydrological modelling. In: *Climate Variability and Change—Hydrological Impacts* (Proc. Conf. La Habana, Cuba, November 2006), 68–74. IAHS Publ. 308. IAHS Press, Wallingford, UK.

Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J. & Ropelewski, C. F. (2007) Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* **28**(7-8), 1503–1526.

Duan, Q., Sorooshian, S. & Gupta, V. (1992) Effective and efficient global optimization for conceptual rainfall–runoff models. *Water Resour. Res.* **28**(4), 1015–1031.

Edijatno & Michel, C. (1989) Un modèle pluie–débit journalier à trois paramètres. *La Houille Blanche* **2**, 113–121.

Espinoza, J. C., Ronchail, J., Guyot, J. L., Filizola, N., Noriega, L., Lavado, W., Pombosa, R. & Romero, R. (2009) Spatio–temporal rainfall variability in the Amazon Basin countries (Brazil, Peru, Bolivia, Colombia and Ecuador). *Int. J. Climatol.* (accepted).

FAO/UNESCO (1981) CD-ROM: *Soil Map of the World*, Ten volumes. UN Food and Agriculture Organization, Rome, Italy.

Hargreaves, G. H. & Samani, Z. A. (1985) Reference crop evapotranspiration from temperature. *Appl. Engng Agric.* **1**(2), 96–99.

Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1), 97–109.

Islam, M. N. & Uyeda, H. (2007) Use of TRMM in determining the climatic characteristics of rainfall over Bangladesh. *Remote Sens. Environ.* **108**(3), 264–276.

Kabouya, M. (1990) Modélisation pluie–débit au pas de temps mensuel et annuel en Algérie septentrionale. These de doctorat Université Paris-Sud, Paris, France.

Lavado, C. W. S., Labat, D., Guyot, J. L., Ardoin-Bardin, S. & Ordoñez, J. J. (2009a) Monthly water balance models in the Amazon drainage basin of Peru: Ucayali River basin. *Hydrol. Sci. J.* (submitted).

Lavado, C. W. S., Labat, D., Guyot, J. L., Boulet, G., Ordoñez, J. J. & Solis, O. (2009b) Comparison of reference evapotranspiration models with the standard FAO Penman-Monteith model in the Peruvian Amazon-Andes basin. *J. Hydrol*. (submitted).

Lavado, C. W. S., Labat, D., Guyot, J. L., Ronchail, J., Espinoza, J. & Ordoñez, J. (2009c) Recent trends in rainfall, temperature and evapotranspiration in the Peruvian Amazon-Andes basin: Huallaga and Ucayali basins. *Int. J. Climatol.* (submitted).

Lavado, C. W. S., Labat, D., Guyot, J. L., Vrugt, J. & Ordoñez, J. (2009d) Monthly water balance models to predict possible climatic change impacts using three IPCC AR-4 models in the Peruvian Amazon-Andes basin. *J. Hydrol*. (submitted).

Makhlouf, Z. & Michel, C. (1994) A two-parameter monthly water balance model for French watersheds. *J. Hydrol.* **162**(3-4), 299–318.

Marengo, J. (2006) On the hydrological cycle of the Amazon basin: a historical review and current state-of-the-art. *Revista Brasileira de Meteorologia* **21**(3), 1–19.

Marengo, J. A., Nobre, C. A., Tomasella, J., Oyama, M. D., De Oliveira, G. S., De Oliveira, R., Camargo, H., Alves, L. M. & Brown, I. F. (2008) The drought of Amazonia in 2005. *J. Climate* **21**, 495–516.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953) Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**(6), 1087–1092.

Molinier, M., Guyot, J.-L., Oliveira, E. D. & Guimarães, V. (1996) Les régimes hydrologiques de l'Amazone et de ses affluents. In: *L'hydrologie tropicale: géoscience et outil pour le développement* (ed. by P Chevalier & B. Pouyaud), 209–222. IAHS Publ. 238. IAHS Press, Wallingford, UK. Available at: http://www.iahs.info/redbooks/238.htm.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models, Part I. A discussion of principles. *J. Hydrol.* **10**(3), 282–290.

Niel, H., Paturel, J.-E. & Servat, E. (2003) Study of parameter stability of a lumped hydrologic model in a context of climatic variability. *J. Hydrol.* **278**(1-4), 213–230.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B. & Braithwaite, D. (2000) Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Met. Soc.* **81**(9), 2035–2046.

Su, F. G., Hong, Y. & Lettenmaier, D. P. (2008) Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *J. Hydromet.* **9**(4), 622–640.

Vrugt, J. A., Gupta, H. V., Bouten, W. & Sorooshian, S. (2003) A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **39**(8), 1201.

Wilk, J., Kniveton, D., Andersson, L., Layberry, R., Todd, M. C., Hughes, D., Ringrose, S. & Vanderpost, C. (2006) Estimating rainfall and water balance over the Okavango River Basin for hydrological applications. *J. Hydrol.* **331**(1-2), 18–29.

Yerren, J., Lavado, W., Fraizy, P. & Guyot, J. (2004) Les régimes hydrologiques dans le bassin amazonien du Pérou. (Los regimenes hidrológicos en la Cuenca Amazónica del Perú). III LBA Scientific Conference: Surface Hydrology and Water Chemistry, LBA, Brasilia, Brazil.

Zeng, N., Yoon, J. H., Marengo, J. A., Subramaniam, A., Nobre, C. A., Mariotti, A. & Neelin, J. D. (2008) Causes and impacts of the 2005 Amazon drought. *Environ. Res. Lett.* **3**, 014002.