

# Precipitation downscaling using random cascades: a case study in Italy

B. Groppelli, D. Bocchiola, and R. Rosso

Politecnico di Milano, Milano, Italy

Received: 12 February 2010 – Revised: 18 March 2010 – Accepted: 7 June 2010 – Published: 2 July 2010

**Abstract.** We present a Stochastic Space Random Cascade (SSRC) approach to downscale precipitation from a Global Climate Model (hereon, *GCMs*) for an Italian Alpine watershed, the Oglio river (1440 km<sup>2</sup>). The SSRC model is locally tuned upon Oglio river for spatial downscaling (approx. 2 km) of daily precipitation from the NCAR Parallel Climate Model. We use a 10 years (1990–1999) series of observed daily precipitation data from 25 rain gages. Scale Recursive Estimation coupled with Expectation Maximization algorithm is used for model estimation. Seasonal parameters of the multiplicative cascade are accommodated by statistical distributions conditioned upon climatic forcing, based on regression analysis. The main advantage of the SSRC is to reproduce spatial clustering, intermittency, self-similarity of precipitation fields and their spatial correlation structure, with low computational burden.

## 1 Introduction

Global Climate Models *GCMs* deliver meteorological variables with a fine resolution in time, but usually for a coarse spatial grid (50–500 km), and they usually reproduce poorly the statistics of historical records at the spatial scales of interest in hydrology. So, a proper tailoring is required for local use, before any accurate guess about hydrologic cycle can be ventured (e.g. Lammering and Dwyer, 2000; Burlando and Rosso, 1991, 2002a). Downscaling in space of outputs of back calculations from climatic models requires appropriate data assimilation schemes (e.g. Bocchiola, 2007; Kang and Ramirez, 2007). A class of methods that accounts for intermittency and self-similarity properties of precipitation is

statistical downscaling via Stochastic Space Random Cascades (SSRC) (Tessier et al., 1993; Over and Gupta, 1994, 1996; Veneziano and Langousis, 2005, 2009; Veneziano et al., 2006). A considerable effort has been recently devoted towards multi scale data assimilation schemes using scale recursive estimation SRE based upon the SSRC theory (Primus et al., 2001; Tustison et al., 2003; Bocchiola and Rosso, 2006; Gupta et al., 2006; Bocchiola, 2007).

This paper addresses downscaling of precipitation, at resolution of approximately  $2 \times 2$  km<sup>2</sup>, for the 1440 km<sup>2</sup> Oglio river basin (closed at Sarnico) in the Retiche Italian Alps (Fig. 1). The study is aimed to evaluate scenarios of future water resources within the Oglio river watershed, in a window of 50 years or so. We previously identified (see Groppelli and Pengo, *master thesis*) the *GCM* that best reproduces precipitation in the study area, i.e. the NCAR Parallel Climate Model (henceforth *NCAR\_PCM*) (Washington et al., 2000; Meehl et al., 2000). This provides the most accurate estimates of mean precipitation in the area, and it is the best candidate for downscaling in the region. For model calibration, we use a 10-year (1990–1999) series of observed daily precipitation data from 25 rain gages within the watershed.

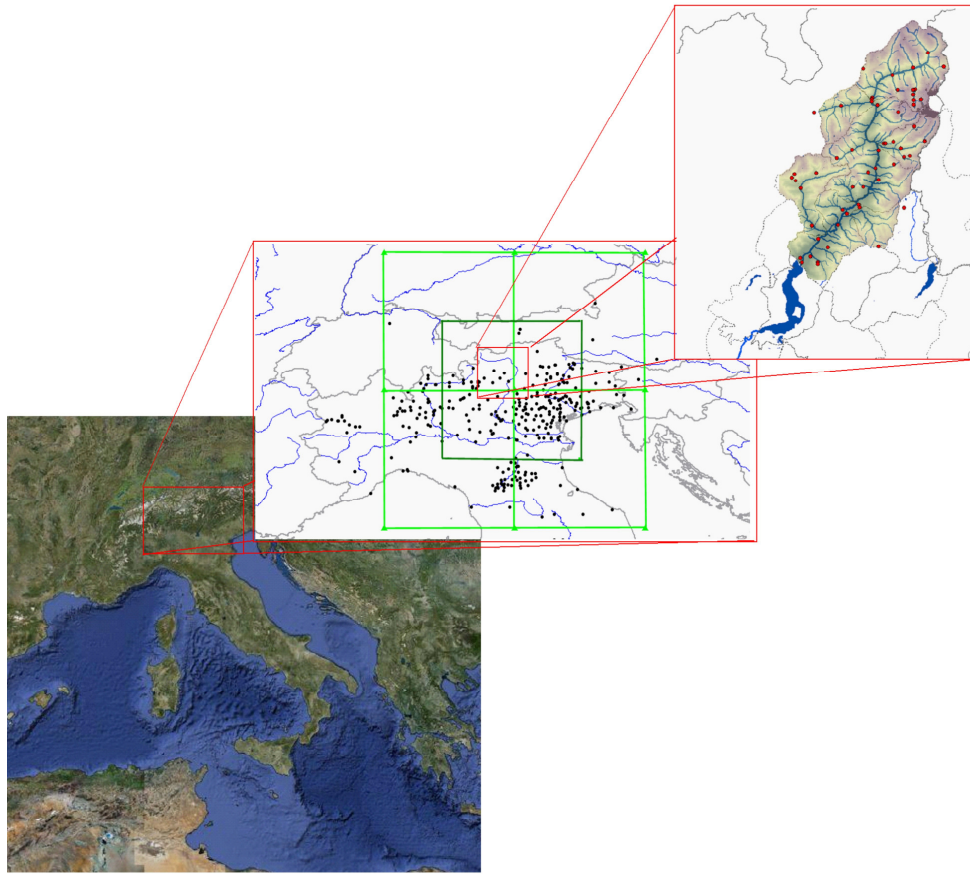
## 2 Downscaling technique

### 2.1 Daily precipitation bias

Visual analysis of daily precipitation series from the *NCAR\_PCM* ( $R_{GCM}$ ) against those observed at ground level (average areal precipitation using the 25 rain gauges within the Oglio watershed, namely  $R_{GAO}$ ) indicates considerably different patterns, concerning both rain rate and intermittency (i.e. the sequence of dry and wet spells). Therefore, one needs to correct  $R_{GCM}$  to obtain daily precipitation consistent



Correspondence to: B. Groppelli  
(bibiana.groppelli@polimi.it)



**Fig. 1.** The study area: Oglio river basin in the Retiche Italian Alps. Clear cells denote the NCAR\_PCM grid and the light dots are the raingauge network in the Oglio river basin.

with  $R_{GAO}$ . Here we do so using a random multiplicative process that accounts for intermittence, termed  $Bias_{GAO}$

$$\begin{aligned}
 Bias_{GAO} &= R_{GAO} / R_{GCM} = B_{GAO} B_0 W_0 \\
 \left. \begin{aligned} P(B_0 = p_0^{-1}) &= p_0 \\ P(B_0 = 0) &= 1 - p_0 \end{aligned} \right\} E[B_0] &= p_0^{-1} p_0 + 0 (1 - p_0) = 1 \\
 W_0 = e^{(w_0 - \sigma_{w_0}^2/2)} &\left. \right\} E[W_0] &= 1 \\
 w_0 &= N(0, \sigma_{w_0}^2)
 \end{aligned} \tag{1}$$

where  $B_{GAO}$ ,  $p_0$ , and  $\sigma_{w_0}^2$  are model parameters to be estimated from data.  $B_{GAO}$  is a constant term that forces the average daily value of  $R_{GAO}$  to match its sample value, because of GCM underestimation of rainfall during wet spells. The term  $B_0$  is a  $\beta$  model generator (see Over and Gupta, 1994). It is the probability that the rain rate  $R_{GAO}$  for a given day is zero, conditioned upon  $R_{GCM}$  being positive, and it is modelled here by a binomial distribution. The term  $W_0$  is a "strictly positive" generator. It is used to add a proper amount of variability to precipitation during wet spells. This approach is used for consistency with the SSRC approach, as described further on.

## 2.2 Stochastic space random cascade

The spatial variability of rainrates is here modelled using a homogeneous SSRC (Schertzer and Lovejoy, 1987; Gupta and Waymire, 1993; Over and Gupta, 1994, 1996; Deidda, 2000; Veneziano and Langousis, 2005; Veneziano et al., 2006).

Spatial rainfall distribution is modelled as a branching tree structure (e.g. Bocchiola, 2007; Groppelli et al., 2010). Each layer in the tree represents a lattice, where the size of the cells (or nodes) is coincident with the resolution (or scale) associated to the samples of the rainfall field obtained with some measurement device(s), or somehow estimated.

The node at the coarsest resolution is called "root" node, while the nodes at the finest resolution are called "leaves". Named  $R_0$  the average rainfall rate at the synoptic scale, the dimensionless rainfall rate in any cell, indexed by  $i$ , at a generic scale,  $s$ , namely  $X_s^i = R_s^i / R_0$  is  $X_s^i = X_0 \prod_{j=1}^s Y_j^i$ , where  $X_0$  is  $X$  at a root node and the operator  $Y$  is a "generator" of the cascade at a given scale, with statistics  $E[Y_s^i] = 1$  and  $Var[Y_s^i] = \sigma_{Y_s}^2$ . Rainfall process displays spatial

intermittence, this meaning that the process has a finite probability mass at zero [Kedem and Chiu, 1987; Kumar and Foufoula-Georgiou, 1994]. Based upon Over and Gupta (1996) we modelled the cascade generator here as the product of two independent generators. This is given by:

$$Y_s^i = B_s^i W_s^i \quad (2)$$

where  $W_s$  is a “strictly positive” generator, modelling the rainfall process for the rainy areas, and  $B_s$  is a  $\beta$  model generator, i.e. the probability that rain rate in a cell at scale  $s$  is zero, conditioned on its parent being positive

$$P(B_s^i = 0) = 1 - b^{-\beta} = 1 - p \quad \text{and} \quad P(B_s^i = b^\beta) = b^{-\beta} = p \quad (3)$$

$$E[B_s^i] = 0(1 - b^{-\beta}) + b^\beta b^{-\beta} = 1$$

where  $b$  is the branching number and  $\beta$  is a parameter estimated from the data set (as in Over and Gupta, 1994). In the rainy areas, one also requires to model the pdf of  $W$ . This is accommodated here using a scaled lognormal variable (see e.g. Marsan et al., 1996; Over and Gupta, 1994) as  $W_s^i = e^{(w_s^i - \sigma_{ws}^2/2)}$  and  $E[W_s^i] = 1$  with  $w_s^i - \sigma_{ws}^2/2 = \text{Log } W_s^i$  and  $w_s^i \approx N(0; \sigma_{ws}^2)$ .

### 3 Model estimation

To estimate  $B_{\text{GAO}}$  and SSRC model parameters we used an analogous approach.  $B_{\text{GAO}}$  in Eq. (1) was estimated as the mean value of the ratio of  $R_{\text{GAO}}$  and  $R_{\text{GCM}}$  for the period (1990–1999). The estimation of  $p_0$  is directly performed using the observed probability of non null rainfall upon gauges conditioned upon GCM’s non null rainfall, and one can use this estimated value for simulation. We did this seasonally, because a preliminary analysis showed that the average duration of dry spells depends upon season. The variance of  $W_0$ , namely  $\sigma_{w_0}^2$ , was also estimated seasonally, using a maximum likelihood approach, performed using a modified version of the SRE-EM algorithm in time [e.g. Bocchiola and Rosso, 2006]. The estimated parameter values are reported in Table 1.

The SSRC model is usually tested and tuned against spatially distributed remotely sensed precipitation estimates, e.g. from ground radar and/or satellites (e.g. Over and Gupta, 1994; Bocchiola, 2007). However, long and accurate series of observed precipitations are seldom available from remote sensing devices, and we are not aware of similar data bases for the catchment of interest. Here, the SRE-EM approach is used for SSRC model estimation from sparse rain gage data, i.e. to evaluate process noise  $\sigma_{ws}^2$  (Groppelli et al., 2010).

We allow estimation of  $\sigma_{ws}^2$  to vary without any regular structure in scales, so making no assumptions concerning the scale structure of precipitation (e.g. by using regular scale invariance, either bounded or unbounded, as e.g. in Tustison et

al., 2003). Gupta et al. (2006), and Bocchiola (2007) demonstrated that so doing better process noise estimates (i.e.  $\sigma_{ws}^2$ ), are obtained. Consistently, downscaling was carried out using unconstrained weights.

The estimation of beta generator  $B$  is generally based upon calculation of scaling of sample moments of order zero at different resolution in space (e.g. Over and Gupta, 1994). Using of sparse gauge networks does not allow direct evaluation of  $B$ , because of scarce spatial representativeness of rain gauges. To overcome this issue we proceeded as follows. For every wet day, we used SRE-EM to produce an interpolated (honouring observed data) rainfall fields at  $2 \times 2 \text{ km}^2$  from the rain gauge observations. Then, we applied a threshold of precipitation, used to set to zero low estimated precipitation (see e.g. Perica and Foufoula-Georgiou, 1996). In this way, areas where low precipitation is attained are considered as dry, obtaining an intermittent rainfall field. The threshold is dynamically evaluated for each day and depends upon the least observed rainfall intensity for that day  $R_{\text{Gmin}}$ , as  $R_{\text{Tresh}} = e^{(\ln[R_{\text{Gmin}}] - c)}$ . The critical value of  $c$  is selected so as to fulfil the criteria of average spatial intermittency as observed, and changes with season according to our findings. In Table 1 we report the obtained values of  $c$  and the corresponding spatial intermittence, expressed as the fraction of area actually covered with rainfall, or Fractional Wetted Area,  $\text{FWA}_G$  and  $\text{FWA}_S$  (from Gauges and Simulated). It is important to notice that spatial intermittence, or FWA changes from storm to storm, depending on the spatial distribution of rainfall. We used the so obtained zero field to estimate the beta generator parameter (i.e.  $p$  in Eq. 3) for each storm.

#### 3.1 Dependence of cascade parameters upon climate forcing

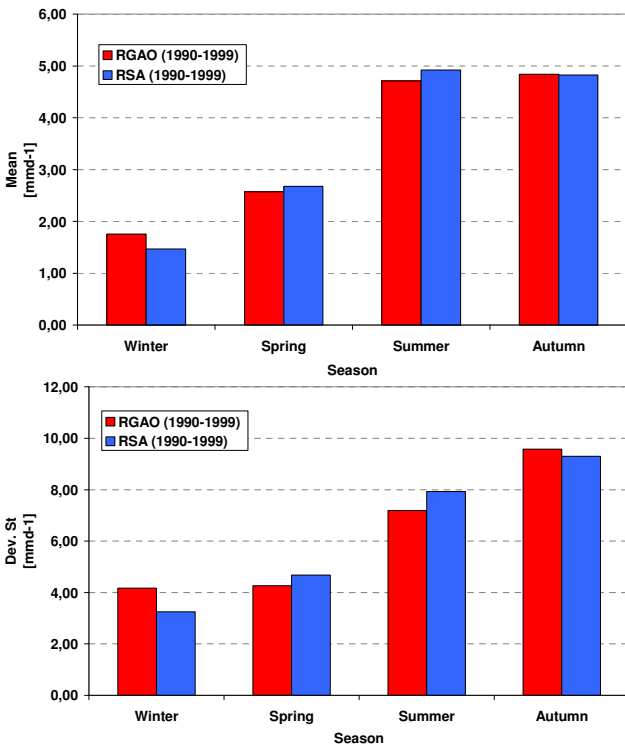
Following Over and Gupta (1994), we tentatively linked FWA, given by parameter  $p$ , to external scale forcing, as expressed by average precipitation in the area, approximated here by  $R_{\text{GA}}$ . According to Over and Gupta (1994), we use a functional dependence,  $1 - p / (1 - b^2) = (R_{\text{GAO}} / R_{\text{max}})^k$ , where  $R_{\text{max}}$  and  $k$  are empirically estimated parameters. The so estimated seasonal parameters are reported in Table 1. The  $\sigma_{ws}^2$  parameters displayed a limited storm to storm variability, particularly at the finest scale, and their scatter plot at each scale could be accommodated using seasonally valid LN distributions (not shown for shortness).

### 4 Model validation and results

We validate the model by downscaling of GCM’s control run for the period 1990–1999. We first validate  $\text{Bias}_{\text{GAO}}$  by *i*) comparing sample fraction of wet days in the simulated series  $R_{\text{SA}}$ ,  $p_{0S}$ , against its sample value from  $R_{\text{GA}}$  series,  $p_0$  and *ii*) comparing the second order statistics of

**Table 1.** Bias<sub>G<sub>AO</sub></sub> parameters. Estimated parameter  $c$  and related average spatial intermittence (FWA) and Regression Parameters  $p/R_{G_{AO}}$ . Symbols are explained in Sects. 2 and 3. All significant results ( $\alpha = 0.05$ ).

Season	$B_{G_{AO}}$ [.]	$\sigma_{W0}^2$ [.]	$FWA_G$ [%]	$FWA_S$ [%]	$c$ [.]	$k$ [.]	$R_{max}$ [mmd <sup>-1</sup> ]	$p$ -value [.]
Winter	1.458	0.27	60%	59%	<b>0.5</b>	0.150	0.384	$< 10^{-5}$
Spring	1.329	0.20	48%	47%	<b>1.4</b>	0.140	0.484	$< 10^{-5}$
Summer	1.669	0.45	48%	48%	<b>1.0</b>	0.137	0.461	$< 10^{-5}$
Autumn	2.005	0.38	47%	47%	<b>1.3</b>	0.111	0.502	$< 10^{-5}$



**Fig. 2.** Validation Statistics back-casting. Symbols are explained in Sects. 3 and 4.

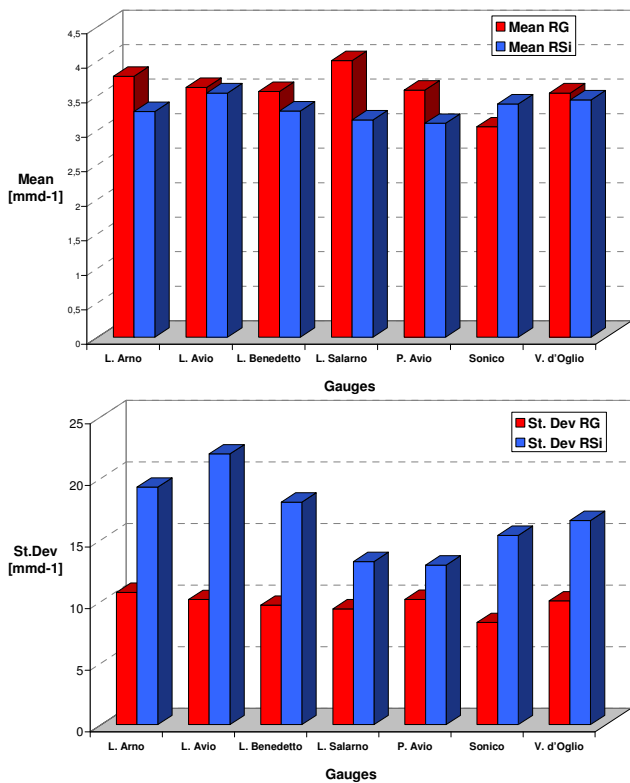
estimated daily area rainfall  $R_{SA} = R_{GCM} \cdot \text{Bias}_{G_{AO}}$  against those of  $R_{G_{AO}}$ . The fraction of wet days  $p_0(i)$  is well described, and its simulated yearly average value is  $E[p_{0S}] = 38.4\%$ , against  $E[p_0] = 36.7\%$ , with  $p$ -value = 0.45. Similarly good agreement was found seasonally, as illustrated in Fig. 2. This indicates suitability of the model to correctly depict daily wet-dry spells sequence. Then, (ii) we investigate the relation between (year round) average of the daily simulated rainfall  $R_{SA}$ ,  $E[R_{SA}] = 3.52 \text{ mmd}^{-1}$  (with  $\text{Dev.St}[R_{SA}] = 7.02 \text{ mm d}^{-1}$ ) and the observed value  $R_{G_{AO}}$ ,  $E[R_{G_{AO}}] = 3.47 \text{ mm d}^{-1}$  ( $\text{Dev.St}[R_{G_{AO}}] = 6.81 \text{ mm d}^{-1}$ ). Good agreement is therefore seen ( $p$ -value = 0.61 and  $p$ -value = 0.03, for mean and standard deviation respectively) and similar

good fitting is observed in each season for the mean values, illustrated in Fig. 2. More critical behaviour is found when dealing with seasonal standard deviation. Albeit numerical values of simulated and observed standard deviations are somewhat close,  $p$ -values for Fisher's F test are low, thus indicating improper variances. However, this may be due to somewhat high number of seasonal daily data used here (circa 900 values/season), making goodness of fit test very severe.

Then, we validated the SSRC by *i*) verifying the agreement of the simulated  $FWA_S$ , with its sample value from the observed series  $FWA_G$ , both upon the catchment as a whole, and upon each single rain gauge, and *ii*) comparing the second order statistics of simulated yearly cumulated precipitation and daily precipitation upon single rain gauges  $i$ ,  $R_{Scumi}$ ,  $R_{Si}$  with the observed series of  $R_{Gcum}$  and  $R_{Gi}$ . The model generally does well in preserving spatial intermittency, and the yearly average simulated wetted area at the finer scale is  $E[FWA_S] = 0.51$  against  $E[FWA_G] = 0.50$  ( $p$ -value = 0.65). Generally acceptable agreement is observed, also seasonally (not reported for the sake of shortness).

Second order (year round) statistics of the simulated daily precipitation upon single rain gauges  $R_{Si}$ , are illustrated in Fig. 3, together their observed counter part,  $R_{Gi}$ . We report here only stations featuring a complete data base for the ten years in study, for robustness. Concerning daily values, the mean values of simulated daily precipitation,  $R_{Si}$  are reasonably correct. Again here, comparison of  $R_{Si}$  against  $R_{Gi}$  shows some issues in term of standard deviation, with simulated values that are somewhat overestimated. However, the high number of considered daily values (i.e. 3652 values) may result again here into considerably low  $p$ -values. Further, at station standard deviation was seen to be considerably affected by few high precipitation events, as obtained by SSRC. Concerning yearly cumulated precipitation  $R_{Scumi}$ , comparison against  $R_{Gcum}$ , displays satisfactory results for both mean and standard deviation (not shown here for shortness).

In Fig. 4 we show a sample snapshots of simulated historic precipitation (namely 14 March 1996), illustrating the concept of FWA and precipitation clustering as reproduced by use of our approach. The event has low spatial intermittency.

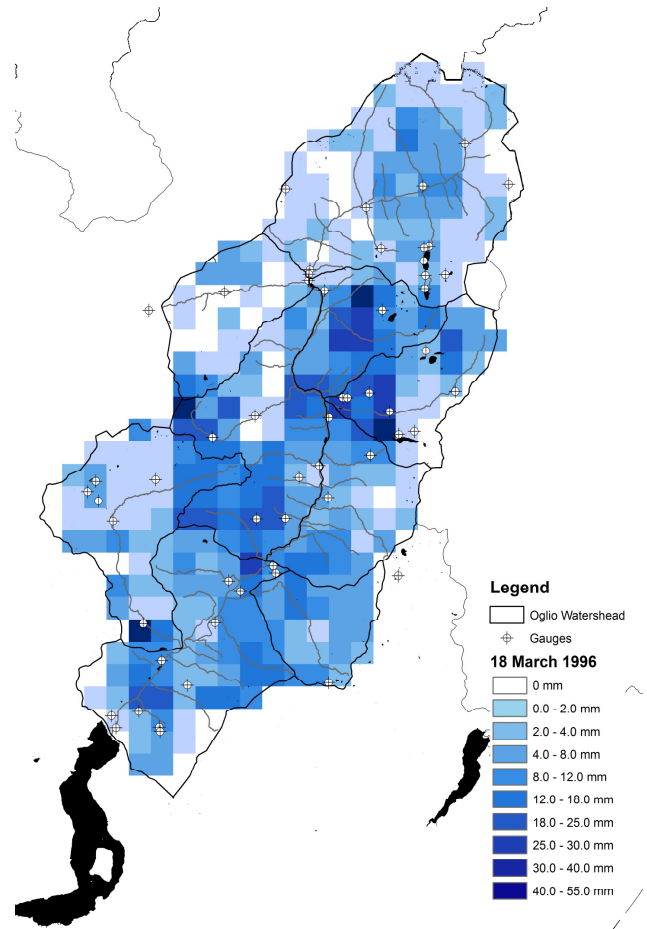


**Fig. 3.** Back-casting station wise comparison of second order moments of daily precipitation (Mean and St.Dev). All stations feature 10 years of data. Symbols are explained in Sections 3 and 4.

Notice the capability of the model to reproduce considerably clustered precipitation fields.

### 5 Conclusions and future developments

The proposed methodology allows downscaling of precipitation from a GCM, under the hypothesis that it can be well depicted using a random cascade. The approach given by  $BIAS_{GAO}$  plus SSRC seems capable of reproducing reasonably well the single site precipitation field in the area. The daily simulated values display a somewhat enhanced variability than that observed in historic precipitation, but yet it shows a good interpretation of spatial intermittency and average daily rainfall. Notice further that such increased variance, albeit theoretically remarkable may be in practice of little importance. In fact, it is related to few randomly (i.e. according to a lognormal distribution) high values of precipitation that do occur within few cells during few events, being in practice very unlikely to change either the main characteristics of the daily rainfall events, or their hydrological consequences. In the future, we will carry out further studies to highlight the cause and significance of such increased variability. The methodology requires the estimation of the



**Fig. 4.** Spatial ( $2 \times 2 \text{ km}^2$ ) daily precipitation for one event from the back-casting rainfall series using SSRC. 14 March 1996 – Event with low spatial intermittency (i.e. high value of  $p$ ).

cascade process parameters, which is here carried out using a maximum likelihood estimation method based upon on a multiscale support. Next developments will include use of the downscaled precipitation to provide future hydrological scenarios for the Oglio river.

*Acknowledgements.* The present paper reports work carried out under the umbrella of the CARIPANDA project, funded by the CARIPLO foundation of Italy under the direction of the ADAMELLO park authority, which is here acknowledged also for supporting with logistic aid. ENEL Produzione Milano is acknowledged for providing precipitation data from their stations. We acknowledge the international modeling groups (IPCC and NCAR) for providing their data for analysis, for collecting and archiving the model data and for organizing the model data analysis activity.

Edited by: R. Deidda  
 Reviewed by: one anonymous referee

## References

- Bocchiola, D. and Rosso, R.: Real time flood forecasting at dams: a case study in Italy, *Int. J. Hydropower Dams*, 13 (1), 92–99, 2006
- Bocchiola, D.: Use of Scale Recursive Estimation for multisensor rainfall assimilation: a case study using data from TRMM (PR and TMI) and NEXRAD, *Adv. Water Res.*, 30, 2354–2372, 2007.
- Burlando, P. and Rosso, R.: Effects of transient climate change on basin hydrology. 1. Precipitation scenarios for the Arno River, central Italy, *Hydrol. Proc.* 16, 1151–1175, 2002a.
- Deidda, R.: Rainfall downscaling in a space-time multifractal framework, *Water Resour. Res.*, 36(7), 1779–94. doi:10.1029/2000WR90003, 2000.
- Groppelli, B. and Pengo, A.: (*Master Thesis*) Sensitività climatica dei deflussi fluviali nei corsi d'acqua alpini. Politecnico di Milano, Italy available upon request (in Italian), 2005.
- Groppelli B., Bocchiola, D., and Rosso, R.: Spatial downscaling of precipitation using random cascades: a case study in Italy, *Water Resour. Res.*, submitted, 2010.
- Gupta, V. K. and Waymire, E.: A statistical analysis of meso-scale rainfall as a random cascade *J. Appl. Meteorol.*, 32, 251–267, 1993.
- Gupta, R., Venugopal, V. and Foufoula-Georgiou, E.: A methodology for merging multisensor precipitation estimates based on expectation-maximization and scale-recursive estimation, *J. Geophys. Res.-Atmos.*, 111(D2), D02102, doi:10.1029/2004JD005568, 2006.
- Kang, B. and Ramírez, J.: Response of Streamflow to Weather Variability under Climate Change in the Colorado Rockies, *J. Hydrol. Eng.*, doi:10.1061/ASCE1084-06992007, 12, 63 pp., 2007.
- Kedem, B. and Chiu, L. S.: Are rain rate process self-similar?, *Water Resour. Res.*, 23(10), 1816–1818, 1987.
- Kumar, P. and Foufoula Georgiou, E.: Characterizing multiscale variability of zero intermittency in spatial rainfall, *J. Appl. Meteorol.*, 33, 1516–1525, 1994.
- Lammering, B. and Dwyer, I.: Improvement of water balance in land surface schemes by random cascade disaggregation of rainfall, *Int. J. Climatol.*, 20, 681–695, 2000.
- Marsan, D., Schertzer, D., and Lovejoy, S.: Causal space-time multifractal processes: predictability and forecasting of rain fields. *J. Geophys. Res.*, 101, 26333–26346, 1996.
- Meehl, G. A., Collins, W. D., Boville, B. A., Kiehl, J. T., Wigley, T. M. L., and Arblaster, J. M.: Response of the NCAR Climate System Model to Increased CO<sub>2</sub> and the Role of Physical Processes, *J. Climate*, 13, 1879–1898, 2000.
- Over, T. M. and Gupta, V. K.: Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large scale forcing, *J. Appl. Meteorol.*, 33, 1526–1542, 1994
- Over, T. M. and Gupta, V. K.: A space-time theory of mesoscale rainfall using random cascade, *J. Geophys. Res.*, 101, 26319–26331, 1996.
- Perica, S. and Foufoula-Georgiou, E.: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *J. Geophys. Res.*, 101(21), 263647–263661, 1996.
- Primus, I., McLaughlin, D., and Entekhabi, D.: Scale-recursive Assimilation of Precipitation Data, *Adv. Water Resour.*, 24, 941–953, 2001.
- Tessier, Y., Lovejoy, S. and Schertzer, D.: Universal multifractals: theory and observations for rain and clouds, *J. Appl. Meteorol.*, 223–250, doi:10.1175/1520-0450, 1993.
- Tustison, B., Foufoula-Georgiou, E., and Harris, D.: Scale-recursive estimation for multisensor quantitative precipitation forecast verification: a preliminary assessment, *J. Geophys. Res.*, 10(D8), 8377, doi:10.1029/2001JD001073, 2003.
- Veneziano, D. and Langousis A.: The Areal Reduction Factor: a Multifractal Analysis, *Water Resour. Res.*, 41, W07008, doi:10.1029/2004WR003765, 2005.
- Veneziano, D., Langousis, A., and Furcolo, P.: Multifractality and Rainfall Extremes: A Review, *Water Resour. Res.*, 42, W06D15, doi:10.1029/2005WR004716, 2006.
- Veneziano, D., Langousis, A.: Scaling and Fractals in Hydrology, In: *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*, Edited by: B. Sivakumar, World Scientific, 145 pp. 2009.
- Washington, W. M., Weatherly, J. W., Meehl, G. A., Semtner Jr., A. J., Bettge, T.W., Craig, A. P., Strand Jr., W. G., Arblaster, J., Wayland, V. B., James, R. and Zhang, Y.: Parallel climate model (PCM) control and transient simulations, *Clim. Dynam.*, 16, 755–774, 2000.