

Physical controls on the scale-dependence of ensemble streamflow forecast dispersion

G. Mascaro¹, E. R. Vivoni², and R. Deidda¹

¹Dipartimento di Ingegneria del Territorio, Università di Cagliari, Cagliari, Italy

²School of Earth and Space Exploration & School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, Arizona, USA

Received: 16 March 2010 – Revised: 15 June 2010 – Accepted: 8 July 2010 – Published: 23 July 2010

Abstract. The accuracy of ensemble streamflow forecasts (ESFs) is impacted by the propagation of uncertainty associated with quantitative precipitation forecasts (QPFs) through the physical processes occurring in the basin. In this study, we consider consistent ESFs (i.e., observations and ensemble members are equally likely) and we study the effect of basin area (A) and antecedent rainfall (AR) on the ESF dispersion, a metric of flood forecast skill. Results from a set of numerical experiments indicate that: (i) for small basins ($\lesssim 180 \text{ km}^2$), ESF dispersion is mainly dominated by the runoff generation process and does not depend on the basin size A ; (ii) for larger areas, ESF dispersion decreases with A according to a log-linear relation due to the decreasing variability of ensemble QPFs and, possibly, to the channel routing process. In addition, we found that, regardless the basin size, the ESF dispersion decreases as AR increases, and that the influence of AR is larger for basins with fast response times. Physical controls (land cover, soil texture and morphometric features) on the analyzed basin response confirm these interpretations.

1 Introduction

The increasing availability of computational resources and recent modeling tools have allowed the development of flood forecasting systems combining meteorological and hydrological models with downscaling and data assimilation techniques (e.g., Westrick et al., 2001; Bacchi et al., 2002; Siccardi et al., 2005). In addition, ensemble techniques have been recently adopted to account for different sources of uncertainty in input data and model parameterization and struc-

ture (Schaake et al., 2007; Verbunt et al., 2007). In parallel with these efforts, new verification tools have been designed to rigorously test the quality of hydrometeorological forecasts, to evaluate the sensitivity of each source of uncertainty, and to identify improvements aimed at increasing forecast skill (see, for example, the report developed by the US National Weather Service available on-line at http://www.nws.noaa.gov/oh/rfcdev/docs/Final_Verification_Report.pdf).

In a recent study of ensemble hindcasts from a hydrometeorological forecasting system, Mascaro et al. (2010) showed how two verification metrics, the Verification Rank Histogram (Anderson, 1996; Hamill and Colucci, 1997; Talagrand et al., 1997) and the Continuous Ranked Probability Score (Wilks, 2006; Hersbach, 2000), can be adapted to evaluate ensemble flood forecasts. In particular, the authors quantified several attributes of forecast quality (consistency, reliability, resolution and uncertainty) associated with ensemble streamflow forecasts (ESFs). They found that Quantitative Precipitation Forecasts (QPFs) characterized by ensemble consistency (or reliability) led to higher flood forecast accuracy, as compared to ESFs obtained from QPFs with ensemble deficiencies (overdispersion and underdispersion). An ensemble forecast is consistent when observations are statistically indistinguishable from ensemble members (Anderson, 1997). If this condition occurs, the spread or dispersion of the ESF is directly related to the skill of the forecast system (Wilks, 2006).

The ensemble flood forecasting verification experiments conducted by Mascaro et al. (2010) also revealed two important challenges for assessing the uncertainty of hydrometeorological forecasts. The first is related to the spatiotemporal attenuation of a QPF as it is transformed into streamflow by the physical processes in a basin (e.g., Vivoni et al., 2007a): analyses by Mascaro et al. (2010) indicated that the basin smoothing effect is higher for larger areas, implying a greater need for accurate ensemble QPFs for smaller basins.



Correspondence to: G. Mascaro
(gmascaro@unica.it)

A second challenge is the strong influence of the rainfall prior to the forecasted event on the flood forecast skill: the higher the antecedent rainfall (AR), the lower the propagation of QPF errors into ESFs quality, while, as AR decreases, the ensemble QPF accuracy becomes more critical.

This paper contributes to the better understanding of these two issues by quantifying the influence of basin size and antecedent rainfall on the forecast skill of a hydrometeorological system, measured here by the ensemble streamflow dispersion. In this exercise, we only account for the uncertainty associated with the rainfall forecasts, whereas all the other sources of uncertainty, such as model parameters and basin initial conditions, are currently neglected.

2 Data and experimental setup

The main objective of the experimental design is to create synthetic “observations” under controlled conditions in order to allow the generation of ensemble QPFs and ESFs with known characteristics. Thus, a database of “observed” rainfall events at high spatiotemporal resolution and corresponding “observed” streamflow was created over the Baron Fork basin in Oklahoma (Fig. 1) during summer periods (June–August) of 1997–2005. This was achieved by utilizing the Space Time RAINfall (STRAIN) multifractal downscaling model (Deidda, 2000) and the TIN-based Real-time Integrated Basin Simulator (tRIBS) (Ivanov et al., 2004a) hydrological model, according to the following steps (Mascaro et al., 2010):

1. Rainfall data estimated by the Next-Generation Radar (NEXRAD) network, available on a regular 4-km grid with a resolution of 1 h, were aggregated within a square domain of side $L=256$ km centered on the Baron Fork (Fig. 1a) over an accumulation time $T=16$ h. This resulted in 138 consecutive 16-h-long rainfall events for each summer, with average precipitation R . Figure 2a shows an example for summer 2000.
2. The STRAIN model was used to downscale the mean rainfall R from the coarse scale $L \times L \times T$ to the fine scale $\lambda \times \lambda \times \tau$, where $\lambda=4$ km and $\tau=15$ min (i.e., 5 levels of downscaling). In this way, the “observed” high-resolution rainfall events were created. To further explain this step, Fig. 2b reports the time series of the mean areal precipitation (MAP) computed over the Baron Fork basin from the high-resolution data.
3. The tRIBS model, calibrated using real precipitation and streamflow data of summer 2000 (Mascaro et al., 2010), was forced with the “observed” fine-scale rainfall, to obtain the “observed” streamflow response for each summer (Fig. 2c). Simulations were conducted for 15 interior locations in order to study the dependence on basin area and other characteristics (Fig. 1b, Table 1).

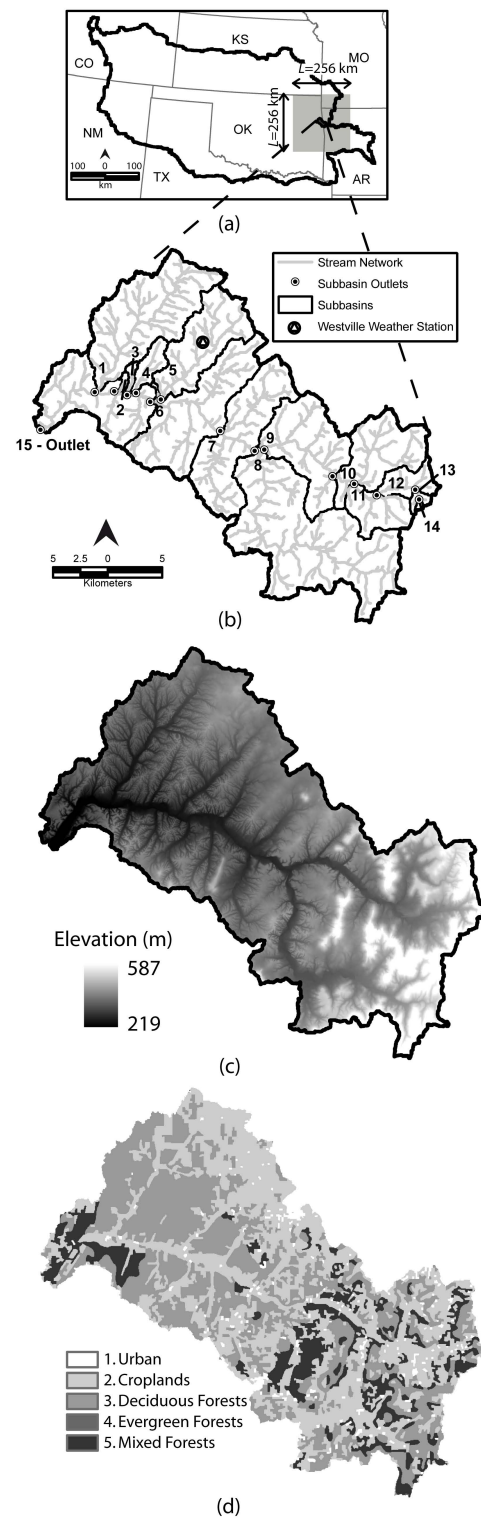


Fig. 1. (a) Baron Fork basin location within the Arkansas Red River basin. The coarse-scale domain $L \times L$ used for rainfall downscaling is plotted with a gray square. (b) Baron Fork basin boundaries, channel network, boundaries and outlets of the 15 sub-basins (Table 1), and the Westville meteorological station. (c) US Geological Survey (USGS) Digital Elevation Model (30-m); (d) Land Cover map derived from the USGS Land Use and Land Cover (LULC) dataset.

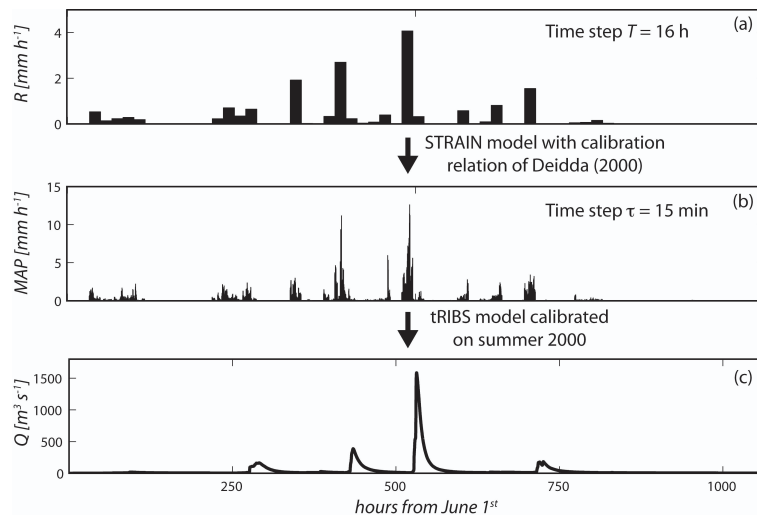


Fig. 2. Synthetic “observed” database generation. (a) Time series of NEXRAD rainfall estimates R during summer 2000, aggregated in the coarse spatial domain $L \times L$ (with $L=256$ km) centered on the Baron Fork over an accumulation time $T=16$ h. (b) Time series of Mean Areal Precipitation (MAP) over Baron Fork, calculated by averaging the high-resolution (4-km, 15-min) “observed” rainfall fields, obtained by downscaling the coarse rainfall R with the STRAIN model. (c) Time series of the “observed” hydrograph at the Baron Fork outlet, produced by the tRIBS model with the “observed” fine-scale rainfall forcing.

Table 1. Baron Fork sub-basin characteristics (Vivoni et al., 2006): area (A), maximum distance to the sub-basin outlet (L), relief ratio (S), drainage density (D_d); time of concentration (T_c) from Kirpich (1940): $T_c=0.000325 \times L^{0.77} \times S^{-0.385}$, where units are L [m] and S [m km^{-1}]; minimum, mean and maximum elevation (min_{EL} , μ_{EL} , max_{EL}); mean slope (μ_{SL}); percentages of basin area covered by land cover classes 1 and 2 (LC_{1+2}) and 3, 4 and 5 (LC_{3+4+5}). For land cover classes, see Fig. 1d.

Sub-basin	A (km^2)	L (km)	S (m km^{-1})	D_d (km^{-1})	T_c (h)	min_{EL} (m)	μ_{EL} (m)	max_{EL} (m)	μ_{SL} (%)	LC_{1+2} (%)	LC_{3+4+5} (%)
1	108.23	25.73	6.06	0.99	5.78	230.42	316.54	386.47	12.52	39.23	60.77
2	1.41	2.59	34.01	0.83	0.51	235.29	282.36	323.38	14.42	13.14	86.86
3	2.67	4.52	21.44	0.77	0.93	237.73	290.04	334.65	15.07	14.08	85.92
4	12.14	8.06	14.94	0.81	1.67	237.73	302.01	358.12	18.69	5.26	94.74
5	65.06	19.90	9.26	0.83	4.03	248.10	327.61	432.49	9.09	58.43	41.57
6	610.06	50.33	6.81	0.84	9.26	243.83	360.08	586.71	10.00	51.63	48.37
7	450.26	40.01	8.11	0.84	7.25	261.51	374.52	586.71	10.01	51.04	48.96
8	365.25	35.03	9.09	0.82	6.27	268.21	385.55	586.71	10.55	48.23	51.77
9	182.91	29.78	9.49	0.82	5.44	272.48	386.59	554.97	10.54	52.52	47.48
10	106.91	18.64	13.41	0.84	3.32	304.98	407.74	554.97	10.50	55.12	44.88
11	49.07	12.72	19.10	0.87	2.16	311.98	413.04	554.97	11.11	57.52	42.48
12	21.18	9.03	24.92	0.87	1.50	329.98	427.24	554.97	12.08	54.44	45.56
13	4.29	3.53	51.27	0.77	0.55	373.98	467.99	554.97	17.40	21.94	78.06
14	0.78	1.33	112.77	0.30	0.19	396.98	479.37	546.97	19.22	18.21	81.79
15 (Outlet)	808.39	67.26	5.47	0.86	12.59	218.84	346.54	586.71	11.02	47.30	52.70

The “observed” series of rainfall and streamflow were treated as synthetic truth without accounting for any kind of observational error. Once the “observed” database was generated, we randomly extracted $N_{\text{EV}}=100$ coarse-scale precipitation events R , whose values range from 0.16 to 4.07 mm h^{-1} , with an average of 0.86 mm h^{-1} . Average values of the corresponding MAP from the “observed” rainfall at 4-km resolution over the 15 sub-basins range from 0.73

to 0.97 mm h^{-1} . For each event, we performed hydrometeorological hindcasts according to the event-based approach shown in Fig. 3. At time t^* , we assume that a Numerical Weather Prediction (NWP) model issues a rainfall forecast R for the next T hours in the domain L by L . Then, the STRAIN downscaling model is used to create an ensemble of $N_{\text{ENS}}=50$ consistent (fine-scale) QPFs and, subsequently, tRIBS simulates the ESFs at 15 nested locations.

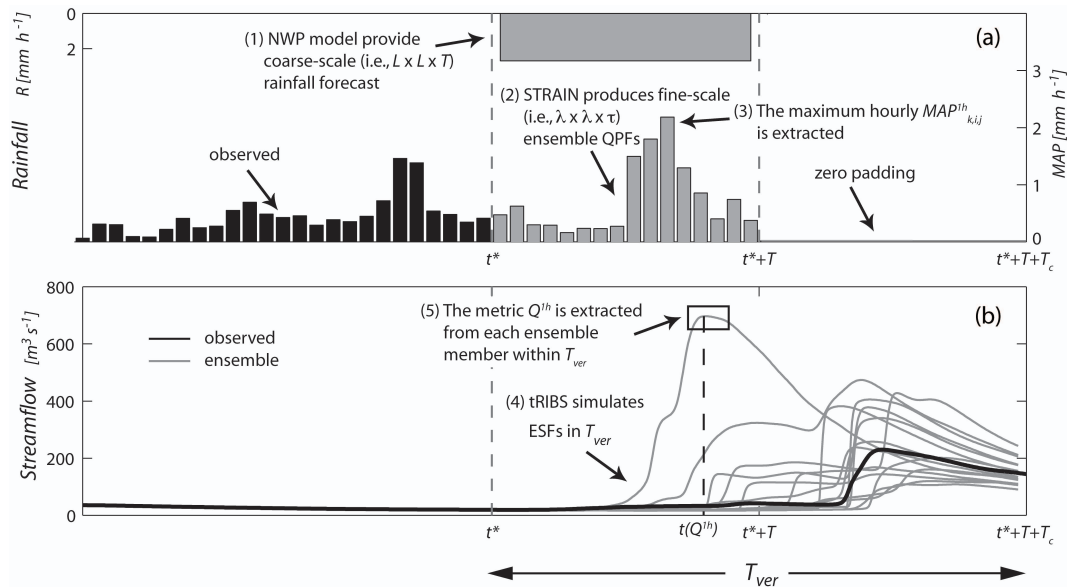


Fig. 3. Event-based approach for the hydrometeorological hindcasts. In (a): (1) precipitation is observed up to t^* , when a NWP model issues a coarse rainfall forecast R for the next T hours over the $L \times L$ domain, (2) the STRAIN model produces fine-scale ($\lambda \times \lambda \times \tau$) consistent ensemble QPFs between t^* and (t^*+T) , and (3) the metric $MAP_{k,i,j}^{1h}$ used to quantify the dispersion of ensemble QPF is extracted from each rainfall member. In (b): (4) the tRIBS model, forced by the ensemble QPFs from t^* to (t^*+T) and zero padding rainfall from (t^*+T) up to (t^*+T+T_c) , simulates the ESFs, and (5) the metric $Q_{k,i,j}^{1h}$ used to measure the ESF dispersion is extracted from each streamflow member within the time window T_{ver} at time $t(Q^{1h})$. Note that the zero-padding interval in the rainfall input is added to include the delay in the streamflow response due to the basin response time.

We underline that the basin state at t^* is determined by a continuous simulation with “observed” rainfall from the beginning of each summer. In this exercise, we did not actually use NWP outputs to determine R , but instead relied on the aggregated NEXRAD rainfall data as a measure of the coarse-scale prediction. This limitation could be readily addressed in future studies.

3 Analysis methods

3.1 Metrics of ESF and QPF dispersion

We introduce some notation to describe the analyses conducted in this study. As previously mentioned, a total of $N_{EV}=100$ events (i.e., 100 total events) were selected and, for each event $k=1, \dots, N_{EV}$, we (i) created an ensemble of $N_{ENS}=50$ consistent QPFs (i.e., 50 ensemble members) over the domain $L \times L \times T$ at resolution $\lambda \times \lambda \times \tau$ and (ii) simulated the corresponding ESFs for each sub-basin $j=1, \dots, 15$. The generic i -th member ($i=1, \dots, N_{ENS}$) of the ensemble QPF generated to predict the k -th event is expressed as $QPF_{k,i}$, while the corresponding member of the ESF simulated in sub-basin j is $ESF_{k,i,j}$.

The first analysis was aimed to characterize the variability of the ensemble rainfall forecasts with the basin area (or scale). For this purpose, for each forecast event k and sub-basin j , we computed the MAP from each $QPF_{k,i}$ and

extracted the hourly maximum, labeled as $MAP_{k,i,j}^{1h}$ (see Fig. 3). To quantify the ensemble rainfall dispersion for each sub-basin, we adopted a metric of the coefficient of variation (CV) of the $MAP_{k,i,j}^{1h}$, computed as:

$$CV_{k,j}^{MAP^{1h}} = \frac{\sigma_{k,j}^{MAP^{1h}}}{\mu_{k,j}^{MAP^{1h}}}, \quad (1)$$

where $\sigma_{k,j}^{MAP^{1h}}$ and $\mu_{k,j}^{MAP^{1h}}$ are the standard deviation and the average of the $N_{ENS}=50$ ensemble $MAP_{k,i,j}^{1h}$. This led to $N_{EV}=100$ values of $CV_{k,j}^{MAP^{1h}}$ for each j sub-catchment.

Next, we performed a similar analysis for the corresponding ESFs. From each $ESF_{k,i,j}$, we extracted a metric $Q_{k,i,j}^{1h}$, defined as the maximum discharge volume accumulated over a duration of 1 h in the verification time window $T_{ver} = T + T_{c,j}$ (Fig. 3), where $T_{c,j}$ is the time of concentration of the sub-basin j (see Table 1). Subsequently, we calculated the CV of the ensemble $Q_{k,i,j}^{1h}$ as:

$$CV_{k,j}^{Q^{1h}} = \frac{\sigma_{k,j}^{Q^{1h}}}{\mu_{k,j}^{Q^{1h}}}, \quad (2)$$

where $\sigma_{k,j}^{Q^{1h}}$ and $\mu_{k,j}^{Q^{1h}}$ are the standard deviation and the average of the $N_{ENS}=50$ ensemble $Q_{k,i,j}^{1h}$. As in the other case, $N_{EV}=100$ values of $CV_{k,j}^{Q^{1h}}$ were calculated for each j sub-catchment.

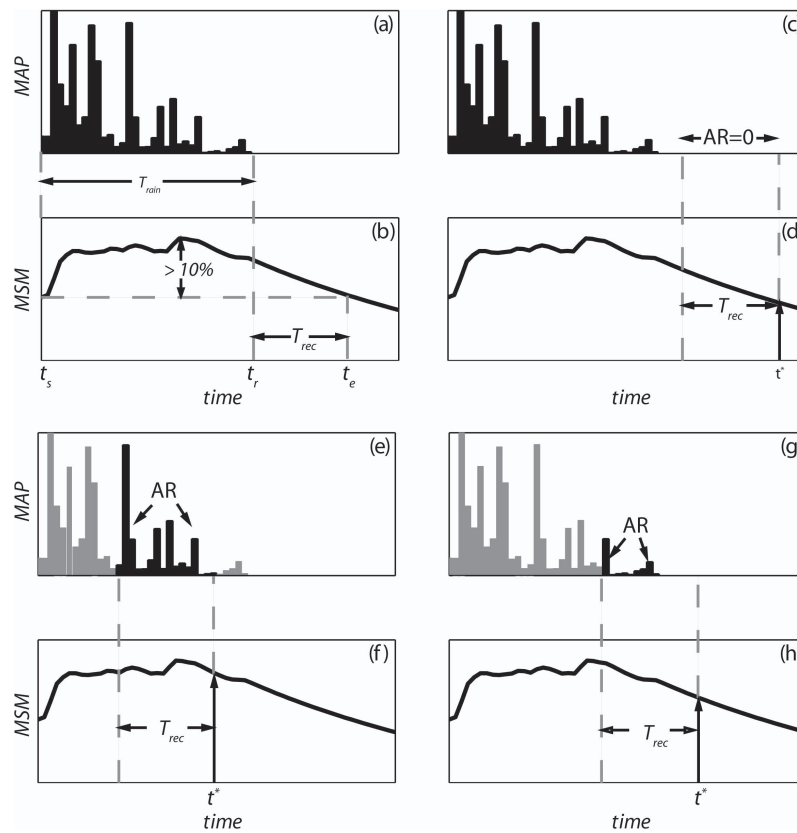


Fig. 4. Schematic of the approach used to estimate the recession time T_{rec} . **(a)–(b)** Time series of MAP and mean soil moisture (MSM) in the top 10 cm. Rainfall starts at time t_s and ends at time t_{er} , with duration $T_{\text{rain}} = (t_{\text{er}} - t_s)$. MSM decreases after t_{er} and, at time t_e , it reaches the initial value at t_s . The recession time is $T_{\text{rec}} = (t_e - t_{\text{er}})$. T_{rec} was estimated considering rainfall events causing an increase of MSM larger than 10% of its initial value. **(c)–(h)** Computation of AR in three cases: the forecast event starts well after a storm (c, d); the forecast event starts during a storm (e, f); and the forecast event starts few hours after a storm (g, h).

3.2 Quantification of basin initial condition

To quantify the basin condition at the beginning of each forecast event k , we computed the antecedent rainfall $\text{AR}_{k,j}$ (a variable readily available in operational forecasting), as the MAP over the sub-basin j averaged over T_{rec} hours prior to the beginning of the event. T_{rec} captures the period during which the antecedent basin conditions impact the response to a rainfall event. Here, T_{rec} was estimated through the soil moisture recession behavior averaged over the basin (MSM), as shown in Fig. 4a–b. At time t_s , a rainfall event increases MSM until the rainfall end at time t_{er} when the MSM recession starts. At time t_e , MSM returns to the value prior to the event at time t_s . In this way, we identify the rainfall duration, $T_{\text{rain}} = (t_{\text{er}} - t_s)$, and soil moisture recession time, $T_{\text{rec}} = (t_e - t_{\text{er}})$. We used a single value of T_{rec} computed as follows: (i) tRIBS model run with “observed” rainfall provided MAP and MSM for the nine summers, (ii) we identified rainfall events that caused an increase in MSM larger than 10%, and (iii) for each selected event, we computed T_{rec} and calculated the mean value over all the events. In the case of Baron Fork, we obtained an average $T_{\text{rec}} = 86$ h.

To better understand how the antecedent rainfall computed over T_{rec} hours prior to t^* (i.e., the beginning of the forecast event) can provide information on the basin initial condition, let us refer to Fig. 4c–h where we presented three possible cases with varying t^* for the same event. When t^* is far from an antecedent event (Fig. 4c–d), the previous rainfall does not affect the basin state and soil moisture redistribution has already occurred. In this case, $\text{AR} = 0$. When a forecast event starts in the middle of or immediately after the storm (Fig. 4e–f), the basin is likely wet and its conditions at t^* is influenced by the previous event. In this case, AR is high. Finally, when t^* is placed some time after the end of a rainfall event (Fig. 4g–h), the soil moisture has started to recess, but the basin conditions still have memory of the past rainfall event. In this situation, AR is low.

We acknowledge that AR is not a perfect proxy of the basin initial condition since it is not a direct reflection of the absolute wetness of a basin. However, since our experiments are focused on summer events with rapid soil moisture recessions, AR can be reasonably used as a proxy of initial conditions. In the following, we present interesting evidence

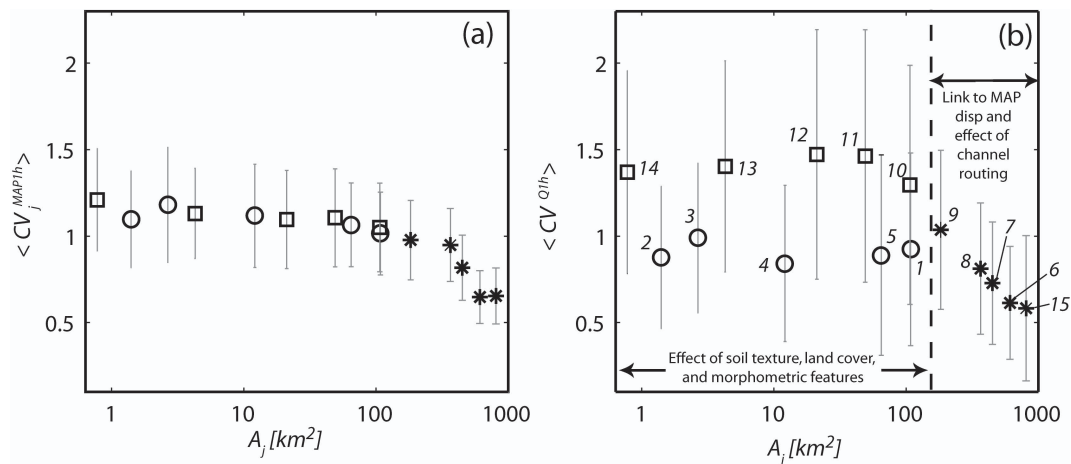


Fig. 5. (a) Variability of ensemble QPFs, quantified by $\langle CV_j^{MAP1h} \rangle$ (index $j = 1, \dots, 15$ refers to the basins listed in Table 1), as a function of the basin size, A_j . (b) Variability of ESFs, quantified by $\langle CV_j^{Q1h} \rangle$, as a function of A_j . Vertical bars indicate the standard deviations of the 100 CVs. The three different markers (circle, square and asterisk) are used to indicate the three groups of sub-basins as explained in the text.

of linkages between AR and ensemble dispersion. In future studies, we plan to evaluate the effect of basin initial state on ensemble forecast quality by: (i) investigating how the soil moisture patterns at the beginning of each forecast event interacts with ensemble QPFs characterized by different statistical properties, and (ii) evaluating how this interaction affects the ESF dispersion.

4 Results and discussion

4.1 Effect of basin area and physical controls

The relation between the variability of ensemble rainfall input and basin area is illustrated in Fig. 5a, where the average of the $CV_{k,j}^{MAP1h}$ over the 100 events for each sub-basin j , indicated with $\langle CV_j^{MAP1h} \rangle$, is plotted versus the corresponding area A_j in logarithmic scale, together with the standard deviations of the CVs (vertical bars). The meaning of the three kinds of markers (circles, squares, and asterisks) is defined later on. A preliminary analysis was conducted to assess the statistical significance of results based on averages over 100 events, finding that, in each sub-basin j , $\langle CV_j^{MAP1h} \rangle$ converges to an approximately constant value when more than 50 randomly selected events are considered. In Fig. 5a, we can see that $\langle CV_j^{MAP1h} \rangle$ (i.e., the variability of the ensemble MAP in each sub-basin j) tends to decrease as the basin area increases since averaging precipitation values over larger domains introduces an increasing smoothing effect. Note that the MAP variability is constant for areas smaller than 16 km^2 since this is the highest spatial resolution of the rainfall fields. Note also that, since the ensemble QPFs were generated by STRAIN using the same parameter set, the high-resolution rainfall fields are characterized by the same multifractal properties in all sub-basins, and, thus, the pattern shown in Fig. 5a is only related to the basin size.

The analogous relation between the dispersion of the ESFs and the basin area is reported in Fig. 5b, where the mean of the $N_{EV}=100$ values of $CV_{k,j}^{Q1h}$, indicated with $\langle CV_j^{Q1h} \rangle$, is plotted vs. A_j , along with the standard deviation of the CVs as vertical bars. In this case, the convergence of $\langle CV_j^{Q1h} \rangle$ as a function of number of events is achieved when more than 75 randomly selected events are used. Interestingly, Fig. 5b reveals a different behavior as compared to Fig. 5a. For areas smaller than $\sim 180 \text{ km}^2$, it is possible to identify two groups of sub-basins where ESF dispersion is not sensitive to A : (a) sub-basins 2, 3, 4, 5, and 1 (order by increasing size; circles), located in the downstream parts of the basin, exhibit a $\langle CV_j^{Q1h} \rangle$ near 0.93 and a smaller variability of the CVs, and (b) sub-basins 14, 13, 12, 11, and 10 (squares), located in the upper parts of the basin, have a $\langle CV_j^{Q1h} \rangle$ of about 1.40 and a higher standard deviation of the CVs. For areas larger than $\sim 180 \text{ km}^2$, the $\langle CV_j^{Q1h} \rangle$ decreases with A_j according to a log-linear relation and reaches the value of 0.58 for the entire basin (asterisks). Clearly, the scale-dependence of the ensemble QPF dispersion is transformed significantly by the watershed, leading to a different scale-dependence of the ensemble streamflow dispersion.

These results also suggest that other features than the catchment area are responsible for the differing behavior between the two groups of small- and medium-sized basins outlined above. To investigate this issue, we analyzed here two factors that condition the basin response time (Vivoni et al., 2007b). A first factor is the basin morphometric (terrain) characteristics that influence runoff travel time over hillslope paths and within stream channels. In particular, high terrain slopes can reduce the basin response time, leading to a lower capacity of the basin to smoothen the rainfall variability. A second factor is the runoff generation influenced by soil and land cover characteristics. In tRIBS, four runoff generation

mechanisms are simulated including saturation excess, infiltration excess, perched subsurface stormflow, and groundwater exfiltration (Ivanov et al., 2004a). The occurrence of a specific runoff type is the result of the interaction of static characteristics of the catchment such as topography, land cover and soil texture, and dynamics factors such as antecedent wetness and rainfall. Here, we are comparing the intrinsic basin characteristics; therefore, in the following, we only focus on the static features affecting runoff generation.

The land cover and soil texture distributions within the sub-basins affect runoff generation. For example, rainfall intercepted by the trees in forested areas and water infiltrated in permeable forest soils lead to a delayed runoff response (e.g., Ivanov et al., 2004b; Norbiato et al., 2009). As a result, we expect that basins with slower runoff generation will attenuate rainfall variability to a greater extent, and this will lead to a lower ESF dispersion. Another factor influencing runoff generation is the presence of regions of flow convergence. We evaluated the possible effect of flow convergence differences by calculating the topographic index (Beven and Kirkby, 1979) defined as $\lambda = \ln(A/\tan\beta)$, where A is the contributing area and β is the land-surface slope angle, as in Forman et al. (2008). We did not find any marked differences across sub-basins, implying that this factor should not affect the observed sensitivity of the ESF dispersion. Therefore, we did not include the effect of flow convergence and only considered land cover and soil texture distributions.

To address the first physical control (morphometric properties), we focused on the relief ratio S , computed for each sub-basin as the elevation range divided by the length of the main channel (Table 1). Sub-basins with high relief ratio are likely subjected to quicker runoff response time, as compared to those with lower S . To address the second physical controls (land cover and soil texture), we refer the reader to Fig. 1d depicting the spatial distribution of urban, croplands, deciduous, evergreen and mixed forests. Following Ivanov et al. (2004b), this map was also used as a surrogate for the spatial variability of soil properties. The calibrated parameters used in our study indicate more permeable soils in forested areas with greater rainfall interception, as compared to croplands and urbanized areas. Table 1 reports the aggregated percentage LC_{1+2} of the basin covered by urban and cropland classes and the aggregated percentage LC_{3+4+5} of forest classes 3, 4, and 5.

To interpret the differences between the two groups of sub-basins, let us consider Fig. 6 reporting the scatterplot of LC_{3+4+5} vs. S for the 15 sub-catchments. Sub-basins 1–4 of group (a) (plotted with circles) are placed in the upper-left part of the plot, with high LC_{3+4+5} (mostly forested) and low-moderate S . In these sub-catchments, runoff generation is likely slower due to a larger amount of intercepted and infiltrated water. The other sub-basin of group (a) (sub-basin 5) has a lower value of LC_{3+4+5} (41.57%), but this is compensated by a low S (9.26 m km^{-1}). As a result, the ESF dispersion of group (a) is relatively low (Fig. 5b), as

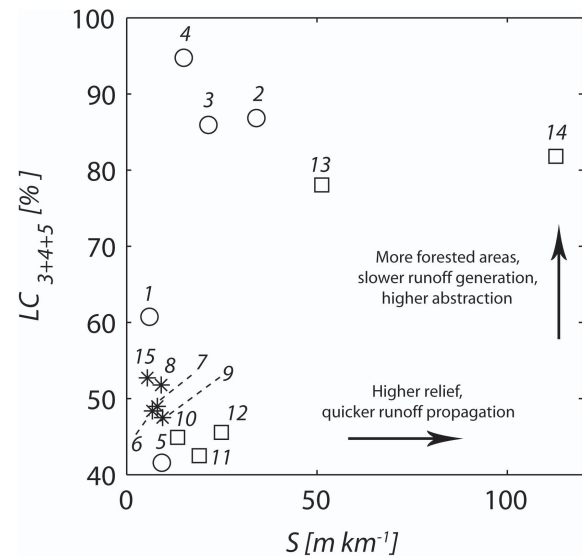


Fig. 6. Scatterplot of LC_{3+4+5} vs. S (see Table 1). Sub-basins of group (a) are indicated with circles, those of group (b) with squares and the remaining sub-catchments with asterisks.

these sub-basins have slower runoff responses, either due to high abstractions or long travel times. In contrast, sub-basins of group (b) (10–14, squares) are located in two different areas of the plot: sub-basins 10–12 are in the lower-left part, with S similar to those of sub-basins 1–4, but a much lower LC_{3+4+5} , resulting in a faster runoff generation due to a larger percentage of urban areas and croplands; sub-basins 13 and 14 are placed in the upper-right part of the plot, with a high percentage of forests ($LC_{3+4+5} \sim 80\%$), but also a very high relief ratio ($S = 51.27$ and 112.77 m km^{-1}). Overall, these characteristics lead to a faster response in sub-basins of group (b), due to lower abstractions or quicker runoff propagation. Faster sub-basin responses lead to greater sensitivity to rainfall variability and a higher ESF dispersion. For completeness, Fig. 6 also reports the other sub-catchments (6–9 and 15, asterisks), which are characterized by an intermediate behavior between the two groups.

For areas larger than 180 km^2 , the ESF dispersion decreases with the basin area. This is generated by the decreasing relation between dispersion of ensemble QPFs and A (Fig. 5a), which is reflected on the same relation for the ESF dispersion. In addition, we argue that over this range of basin areas, runoff generation may become secondary to channel routing processes (also see Vivoni et al., 2006). This is supported by the ensemble streamflow simulations of Carpenter and Georgakakos (2006) who showed a log-linear decrease in the ensemble flow range for large basin areas. Further, as the channel network length was increased (due to intrinsic characteristics of the basins), these authors obtained a steeper decrease in the ESF dispersion. Thus, in light of analyses presented here and results of Carpenter

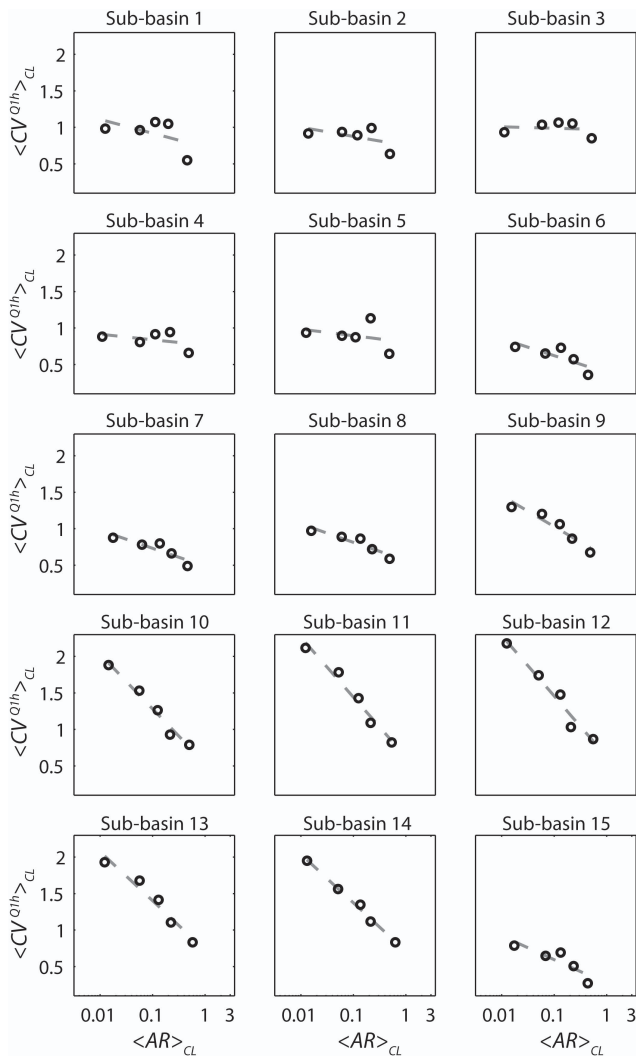


Fig. 7. Relation between ESF dispersion and AR (mm h^{-1}) in each sub-basin.

and Georgakakos (2006), we conclude that channel routing can be a physical factor explaining the decrease of ensemble streamflow dispersion for large areas. Presently, it is unclear which basin characteristics control the transition in ESF dispersion behavior between scale ranges dominated by runoff generation at small A and channel routing at large A .

4.2 Effect of antecedent rainfall and physical controls

The relation between the ESF dispersion and the antecedent rainfall is illustrated in Fig. 7. Each panel refers to a sub-basin j and was constructed as follows. First, we sorted the $\text{AR}_{k,j}$ values in increasing order and divided them into 5 AR classes with equal number of events (20). For each class, we computed the average of: (i) the 20 $\text{AR}_{k,j}$ values of events falling in that class, and (ii) the corresponding $\text{CV}_{k,j}^{\text{Q}^{1\text{h}}}$. To simplify the notation, these mean quantities

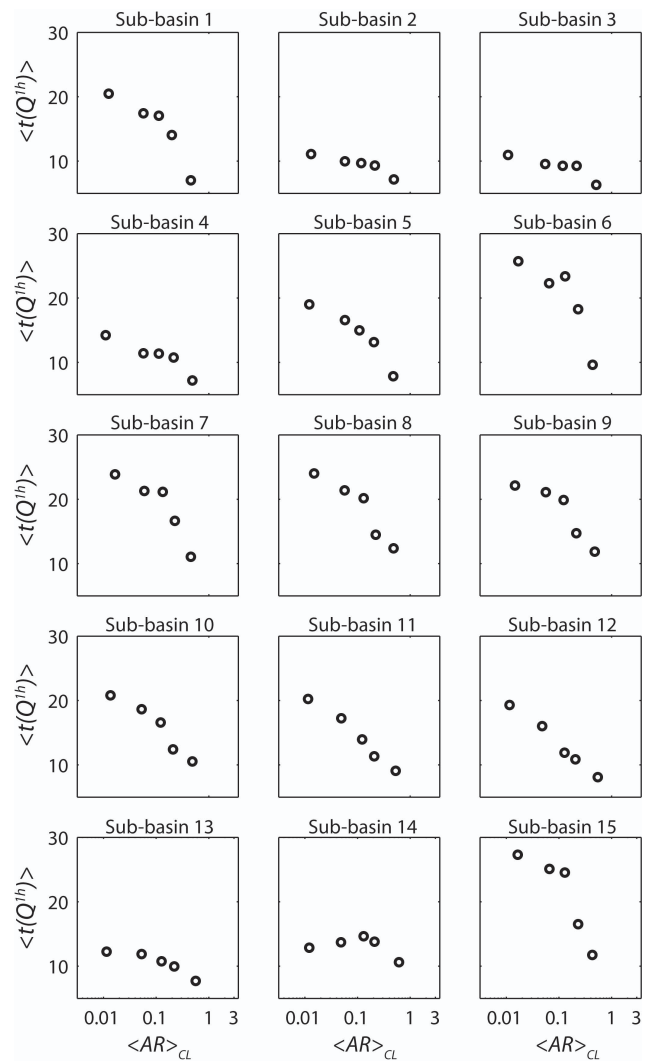


Fig. 8. Relation between $\langle t(Q^{1\text{h}}) \rangle$ (measured in hours from t^*) and AR (mm h^{-1}) in each sub-basin.

were indicated with $\langle \text{AR} \rangle_{\text{CL}}$ and $\langle \text{CV}^{\text{Q}^{1\text{h}}} \rangle_{\text{CL}}$, without reference to the sub-basin j . Overall, Fig. 7 shows that the ESF dispersion decreases as AR increases. This implies that ensemble flood forecasts following high AR have less dispersion as compared to forecasts after a low AR in the same sub-basin. A possible physical interpretation for this finding is as follows. As previously described, rainfall prior to a forecasted event influences the basin state (soil moisture content, streamflow) at t^* , the beginning of the forecast (Figs. 3 and 4) (e.g., Fedora and Beschta, 1989; Vivoni et al., 2006; Merz and Blöschl, 2009). If the amount of rainfall prior to t^* is large (high AR), most of the basin area has begun to respond, possibly leading to the streamflow rising limb. Under these conditions, the runoff production in each ensemble hydrological simulation will be mainly dominated by a deterministic pattern due to the prior storm, and the ESF dispersion will be reduced. In contrast, if no or little rainfall

Table 2. Values of parameters K_0 and K_1 of Eq. (3) and the linear correlation coefficients r between $\langle CV^{Q^{1h}} \rangle_{CL}$ and $\log(\langle AR \rangle_{CL})$ in each sub-basin.

Sub-basin	K_0	K_1	r
1	0.76	0.05	-0.88
2	0.68	0.05	-0.72
3	0.83	0.04	-0.77
4	0.66	0.05	-0.89
5	0.77	0.04	-0.76
6	0.45	0.06	-0.79
7	0.55	0.06	-0.83
8	0.60	0.07	-0.92
9	0.78	0.08	-0.73
10	0.90	0.12	-0.74
11	1.08	0.10	-0.75
12	1.06	0.11	-0.98
13	1.02	0.09	-0.94
14	1.00	0.08	-0.98
15	0.37	0.09	-0.93

prior to t^* has been observed (AR=0 or low AR), the ESF dispersion will be mainly affected by the ensemble QPFs variability and, to a less extent, by a deterministic component due to a prior storm (when $AR \neq 0$). In this case, the ESF dispersion will be likely larger.

The validity of this hypothesis would imply that, for high AR, the time $t(Q^{1h})$ where the maximum $Q_{k,i,j}^{1h}$ of each ensemble streamflow member is found within T_{ver} (Fig. 3) is shifted towards t^* , meaning that we expect a decreasing relation between $t(Q^{1h})$ and AR. This is confirmed in Fig. 8, where the average of $t(Q^{1h})$ over the same classes used in Fig. 7 ($\langle t(Q^{1h}) \rangle$) are plotted versus $\langle AR \rangle_{CL}$. As an additional consequence of these results, forecasts with high AR are less affected by the presence of deficiencies (e.g., bias, underdispersion or overdispersion) in the ensemble QPFs, as compared to ESFs obtained with lower AR.

As a further analysis, we investigated the differences in the relation between ESF dispersion and AR by fitting the following regression in the 15 sub-basins (Fig. 7):

$$\langle CV^{Q^{1h}} \rangle_{CL} = K_0 - K_1 \log(\langle AR \rangle_{CL}), \quad (3)$$

where K_0 and K_1 are parameters reported in Table 2, together with the linear correlation coefficients of the regressions (see Fig. 7, dashed gray lines). Parameter K_0 accounts for the magnitude of ESF dispersion, while the slope K_1 controls the rate of decreasing of ESF dispersion with the antecedent rainfall. In the extreme situation where $K_1=0$, AR does not play a role in ESF dispersion, whereas as K_1 increases, the effect of AR becomes larger. The relation between the parameters K_1 and the sub-basin area is illustrated in Fig. 9. This clearly reveals that it is possible to detect different patterns for the same two groups of sub-basins identi-

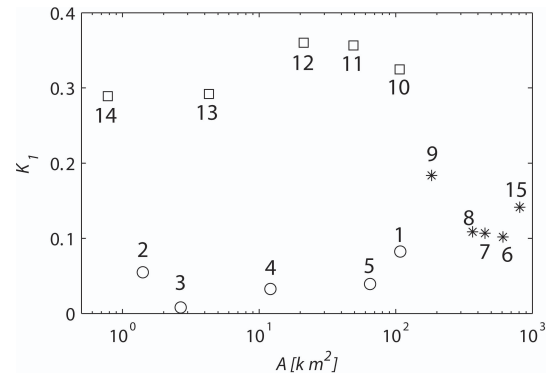


Fig. 9. Relation between parameter K_1 for each sub-basin, and the corresponding sub-basin area.

fied in Fig. 5b. In fact, sub-basins in group (a) (plotted with circles) are characterized by a lower K_1 as compared to sub-basins in group (b) (squares). The remaining sub-basins (asterisks), at large basins areas, have values of K_1 intermediate between those of the other two groups. This is consistent with previous results and interpretations. The physical factors that increase or decrease the runoff response times (soil texture, land cover, terrain slope) also play an important role in determining if the antecedent rainfall has a strong or weak control on the ESF dispersion. Those sub-basins in group (a) with slow response times (permeable, forested soils with high abstractions and low slopes) are less sensitive to AR (low K_1). In other words, the physical controls in these sub-basins lead to an ESF dispersion for high AR which is not considerably smaller than the one obtained for low AR, due to the greater spatiotemporal attenuation of the ensemble QPFs. In sub-basins of group (b) with faster response times (less permeable croplands with lower abstractions and higher slopes), the ESF dispersion is attenuated to a greater extent with higher AR (high K_1) due to the more limited basin smoothing that occurs in the rainfall-runoff transformation.

5 Summary and conclusions

Ensemble techniques have been recently adopted in hydrometeorological systems to account for different sources of uncertainty. However, a limited number of studies have been devoted to investigate the effect of the basin characteristics and initial state on the skill of ensemble forecasts. In this study, we quantified the influence of two important factors on the accuracy of the ensemble streamflow forecasts (ESFs): (i) the attenuation effect on the spatiotemporal variability of ensemble rainfall forecasts due to the physical processes in a basin, and (ii) the amount of rainfall prior to the forecast event, which is one component affecting the initial forecast state. For this purpose, we conducted a synthetic experiment where we generated consistent (or reliable) ESFs for

a large set of hydrometeorological events over the Baron Fork basin (808 km²) and 14 sub-basins (minimum area of 0.78 km²). To quantify flood forecast skill, we introduced a metric measuring the ESF dispersion. In this exercise, the ensemble rainfall forecasts are characterized by the same statistical properties in all sub-basins. Results of the study can be summarized in the following two points:

1. The variability of the ensemble rainfall forecasts has a clear decreasing pattern with basin area, while the physical processes governing the rainfall-runoff transformation lead to a relation between ESF dispersion and basin area which requires a careful interpretation. For small areas, the ESF dispersion is constant and its value depends on the basin features controlling runoff generation (land cover and soil texture) and travel time (morphometric features). For larger areas, ESF dispersion decreases with A according to a log-linear relation, due to the decreasing variability of the QPFs as the area increases, and, possibly, to the response time in the channel network.
2. ESF dispersion decreases as the total amount of rainfall prior to the forecasted event (AR) increases. This is explained considering that, for high AR, the runoff production in each ensemble hydrological simulation is mainly affected by a deterministic component due to the previous storm runoff. As a consequence, the ESF dispersion will be reduced.

We highlight two new contributions of this study. In a past study, Carpenter and Georgakakos (2006) found that channel routing is a dominant process responsible for the variability of the ESF dispersion for relatively large basin areas. In our study, we confirmed this result and also found that, in small basins, runoff generation becomes an important factor controlling the ESF dispersion. This was demonstrated by the differences in the basin properties (soil texture, land cover, slope) affecting these physical processes. A second novel finding is that the same basin physical properties also control the impact of the antecedent rainfall on the flood forecast skill. In basins with slower response times that are able to attenuate the rainfall variability, the ESF dispersion decreases with a lower rate as AR increases. In contrast, in basins with a faster response, the effect of the antecedent rainfall on the streamflow response is larger. In conclusion, this study advances the understanding of the role of the key physical factors affecting the scale-dependence of the ensemble flood forecast dispersion.

Acknowledgements. We thank the NSF EPSCOR program for access to the Linux cluster at New Mexico Tech. We also thank funding from NASA Terrestrial Hydrology Program (Grant NNX08AD49G) and the Assegno di Ricerca Ministeriale dell'Università di Cagliari.

We thank reviewers for their comments that helped to improve the quality of the manuscript.

Edited by: G. Boni

Reviewed by: L. Garrote, R. Rudari, and another anonymous referee

References

- Anderson, J. L.: A method for producing and evaluating probabilistic forecasts from ensemble model integrations, *J. Climate*, 9, 1518–1530, 1996.
- Anderson, J. L.: The impact of dynamical constraints on the selection of initial conditions on ensemble predictions: Low-order perfect model results, *Mon. Weather Rev.*, 125, 2969–2983, 1997.
- Bacchi, B., Buzzi, A., Grossi, G., and Ranzi, R.: Flood forecasting in a midsize catchment in the southern Alps: Recent experiences on the use of coupled meteorological and hydrological models, in: *Proceedings of 3rd Plinius Conference on Mediterranean Storms*, Baja Sardinia, edited by: Deidda, R., Mugnai, A., and Siccardi, F., Consiglio Nazionale delle Ricerche, Roma, Italy, 201–208, 2002.
- Beven, K. J. and Kirkby, M. J.: A physically-based variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24(1), 43–69, 1979.
- Carpenter, T. M. and Georgakakos, K. P.: Discretization scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modelling, *J. Hydrol.*, 328, 242–257, 2006.
- Deidda, R.: Rainfall downscaling in a space-time multifractal framework, *Water Resour. Res.*, 36, 1779–1794, doi:10.1029/2000WR900038, 2000.
- Fedora, M. A. and Beschta, R. L.: Storm runoff simulation using an antecedent precipitation index (API) model, *J. Hydrol.*, 112(1–2), 121–133, doi:10.1016/0022-1694(89)90184-4, 1989.
- Forman, B. A., Vivoni, E. R., and Margulis, S. A.: Evaluation of ensemble-based distributed hydrologic model response with disaggregated precipitation products, *Water Resour. Res.*, 44, W12409, doi:10.1029/2008WR006827, 2008.
- Hamill, T. M. and Colucci, S. J.: Verification of Eta-RSM short-range ensemble forecasts, *Mon. Weather Rev.*, 125, 1312–1327, 1997.
- Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, *Weather Forecast.*, 15, 559–570, 2000.
- Ivanov, V. Y., Vivoni, E. R., Bras, R. L., and Entekhabi, D.: Catchment hydrologic response with a fully distributed triangulated irregular network model, *Water Resour. Res.*, 40, W11102, doi:10.1029/2004WR003218, 2004a.
- Ivanov, V. Y., Vivoni, E. R., Bras, R. L. and Entekhabi, D.: Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach, *J. Hydrol.* 298, 80–111, 2004b.
- Kirpich, Z. P.: Time of concentration of small agricultural watersheds, *Civil Eng.*, 10(6), 362 pp., 1940.

- Mascaro, G., Vivoni, E. R., and Deidda, R.: Implications of Ensemble Quantitative Precipitation Forecast Errors on Distributed Streamflow Forecasting, *J. Hydrometeorol.*, 11, 69–86, doi:10.1175/2009JHM1144.1, 2010.
- Merz, M. and Blöschl, G.: A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, *Water Resour. Res.*, 45, W01405, doi:10.1029/2008WR007163, 2009.
- Norbiato, D., Borga, M., Merz, R., Blöschl, G., and Carton, A.: Controls on event runoff coefficients in the eastern Italian Alps, *J. Hydrol.*, 375(3–4), 312–325, doi:10.1016/j.jhydrol.2009.06.044, 2009.
- Schaake, J. C., Hamill, T. M., Buizza, R., and Clark, M.: HEPEX, the Hydrological Ensemble Prediction Experiment, *B. Am. Meteorol. Soc.*, 88, 1541–1547, 2007.
- Siccardi, F., Boni, G., Ferraris, L., and Rudari, R.: A hydrometeorological approach for probabilistic flood forecast, *J. Geophys. Res.*, 110, 1–9, doi:10.1029/2004JD005314, 2005.
- Talagrand, O., Vautard, R., and Strauss, B.: Evaluation of probabilistic systems, in: *Proceedings, ECMWF Workshop on Predictability*, ECMWF, Reading, Berkshire, United Kingdom, 125, 1–25, 1997.
- Verbunt, M., Walser, A., Gurtz, J., Montani, A., and Schr, C.: Probabilistic flood forecasting with a Limited-Area Ensemble Prediction System: selected case studies, *J. Hydrometeorol.*, 8, 807–909, 2007.
- Vivoni, E. R., Entekhabi, D., Bras, R. L., Ivanov, V. Y., Van Horn, M. P., Grassotti, C., and Hoffman, R. N.: Extending the predictability of hydrometeorological flood events using radar rainfall nowcasting, *J. Hydrometeorol.*, 7, 660–677, 2006.
- Vivoni, E. R., Entekhabi, D., and Hoffman, R. N.: Error propagation of radar rainfall nowcasting fields through a fully-distributed flood forecasting model, *J. Appl. Meteorol. Clim.*, 46(6), 932–940, 2007a.
- Vivoni, E. R., Entekhabi, D., Bras, R. L., and Ivanov, V. Y.: Controls on runoff generation and scale-dependence in a distributed hydrologic model, *Hydrol. Earth Syst. Sci.*, 11, 1683–1701, doi:10.5194/hess-11-1683-2007, 2007b.
- Westrick, K. J. and Mass, C. F.: An evaluation of a high-resolution hydrometeorological modeling system for prediction of a cool-season flood event in a coastal mountainous watershed, *J. Hydrometeorol.*, 2, 161–180, 2001.
- Wilks, D. S.: *Statistical Methods in the Atmospheric Sciences*, 2nd edn., Academic Press, 627 pp., 2006.