

Responses to Editor's and Reviewers' Comments
HESS-2017-752

Technical note: Space-time analysis of rainfall extremes in Italy: clues from a reconciled dataset

Libertino, Andrea*; Ganora, Daniele; Claps, Pierluigi
*andrea.libertino@polito.it

We greatly appreciate the insightful comments from the editors and reviewers. The comments from the editors and reviewers have been reproduced below, interspersed with our indented responses (R#1, R#2 and R#4 refers to comments from the anonymous reviewers 1,2 and 4 respectively, R#AM are comment from the reviewer Alberto Montanari and E refers to Editor's comments). The comments have been grouped when dealing with similar topics. The line numbering refers to the numbering of the revised manuscript. To help identifying the updates made to the manuscript, a copy with the track changes is attached too.

R#1_C0 *The presentation of the manuscript is clear and concise. The authors credit the sources of the data (even for those 12 Institutions out of 21 which have not yet provided data, see Table 1, "Under request" under the column "Digitized data availability")*

R#AM_C2 *Table 1: I do not understand the meaning of "under request". Does this mean that data are not yet available? Were the data already requested? Figure 2 shows that information for some of the regions labeled as "under request" is already available and therefore the whole picture is not completely clear to me.*

R#4_C4 *Table 1: It is unclear what "under request" means here. Does it mean "available upon request"?*

We were referring to those cases in which the agencies do not provide data directly on their websites but they provide it "upon request". In the revised manuscript the "Under request" wording is substituted in Table 1 by the more explicit "available upon request".

R#1_C1 *Data policy 1. The I-RED data policy is not clearly specified while it must be clearly presented to the reader both in the main text and briefly mentioned in the abstract too. The authors' work is valuable for the research community and the society whether I-RED could be made totally public or not. However, I believe most readers would be interested in knowing if I-RED is accessible in total or only in part. For instance, could you add a column to Table 1 to make it clear if you are allowed to re-distribute the data through I-RED? Is there any website where the reader could access the public part of I-RED?*

R#1_C2 *Data policy 2. You made clear in the manuscript that I-RED is related to CUBIST, which can be accessed on the polito website and as you write it is "first important attempt of making the large Italian hydrological heritage freely available in computer-readable format" (page 2, line 19). My question is: why CUBIST data is freely available and I-RED data has restricted access?*

R#1_C3 *Data policy 3. How could you reconcile these two statement: (1) page 2 line 27 "...the Italian law adopted an Open Source policy for the public data..." and (2) your statement on "data availability" where you write that you have signed agreements with the data provider (which are all public institutions) that restrict the use of the data to the aims of your project? In my view, if you get the dataset only after signing an agreement that limit the use of a dataset then the dataset is not openly available (by definition).*

R#AM_C6 *Data availability is a potential issue. Data-bases are useful when they are readily available. It would be interesting to discuss data availability in the body of the paper.*

E *Please, explain into detail the availability of data (such as upon request), as this seems to be the major plus of the paper for a wide audience that there is a set of data available to many researchers not familiar with the data set in Italy and how to get the data.*

We thank the reviewers and the editor for the suggestions. Data availability is one of our main concerns, that is why we have reported in the caption of Table 1 all the websites where the reader could find the original regional datasets. In the revised version of the manuscript have described the procedure to access the data that we have merged and harmonized in the I-RED in the "data availability" box. Considering the relevance of the topic, some lines about the accessibility of the I-RED have been also added in the abstract (P1, LL10-11) and in the manuscript (P7, LL4-5). Substantially, as the owners of the data want to supervise the use of their data and have restricted the access to the aim of our project, we can provide the full database only to those who will join our project and collaborate in its development. We plan also to provide the access to the regional databases to those who will provide us a clearance from the regional authorities in which they are interested.

We understand that it is not easy to understand why the data of the CUBIST dataset is effectively open to everybody while the equivalent data from the local agencies have a restricted access, but this is mainly a drawback

of the change in the owner of the network that we have described in the manuscript. The data included in the CUBIST database are the digitized version of the data included in the Hydrological Yearbooks of the *SIN-SIMN* that have been made freely available in a PDF format under the Hydrological Yearbooks Project of the Italian National Institute for Environmental Protection and Research (*ISPRA*) (<http://www.isprambiente.gov.it/it/progetti/acque-interne-e-marino-costiere-1/progetto-annali>). As digital version of public data, they are public too. When the *SIMN* have been dismantled every local authority adopt a different policy for data providing and distribution. Generally, even if they are still available for research and non-profit purposes, they are monitored from the competent authorities. From this emerges the difficulties in making the *I-RED* dataset fully open access. The Italian law concerning the open data mainly refers to the D.Lgs.82/2005, D.M.10/11/2011, D.Lgs. 36/2006, L. 221/2012, D.Lgs. 179/2012, L. 114/2014 as clarified at P2, L33. Substantially it requires that all the data of public interest collected from public authorities should be made freely available on the internet for non-commercial use. The complex framework of the hydrological authorities answered differentially to the law. Some of the regional authorities started immediately to provide the data openly on the internet as they become available. Other regions, to make sure of the non-commercial use of the data, provide them only upon request, after a certification of the purposes for which the data would be used. Other regions are still developing the platforms for the digitalization and the distribution of the data, and their local archives are still not officially ready for the publication. Moreover, there is a gap in some regions between the dismissal of the *SIMN* and the full operability of the local authorities: in this period data could have been collected in a non-systematic way from authorities different from the definitive one. This involves further complexity in the definition of who and how should work on patching and publishing the data.

R#1_C4 *Conclusions and Future plans. Page 10, line 14. "The final aim is to make the update of the database systematic and unsupervised." A few words on how the authors plan to achieve the goal would do here. In particular, if the author could present any link with national and/or international project/activities this would strengthen their statement. For example, I'm aware of initiatives aimed at collecting Italian datasets such as ArcCIS (<https://www.arcis.it/wp/en/home-2/>) or ISPRA-SCIA (http://www.scia.isprambiente.it/home_new.asp). Do the authors have any contact with them? In particular, it seems to me that ISPRA-SCIA is doing a work on the establishment of a national database that is very similar to I-RED, though mostly for aggregated values. Do you plan to join your efforts with other institutions? Do you have any contact with international institutions? There are several ongoing projects at an international level to collect and organize in-situ observations (See for example COPERNICUS: <https://climate.copernicus.eu/global-land-and-marine-observations-database>)*

We have clarified the part of the manuscript related to the future plans. As clarified at P11, LL4-14 the automatic and unsupervised update of the database can be obtained by strengthening the collaboration with the data providers, in the framework of joined projects, as the one that led to the development of the *ArcCIS* dataset, collecting updated rainfall and temperature data from a group of regional authorities in Northern Italy. The consistence with other projects working at other spatio-temporal scales will be also carried out in order to automatically analyse the consistence of the developed dataset and integrating it with the existing ones, e.g. the *SCIA* dataset, referring to the 24 hours and daily scale. Joined projects with international institutions will be evaluated and endorsed in order to make available the *I-RED* database in larger frameworks for trans-boundary exchange of precipitation data.

R#4_C3 *Page 3 Lines 14-15: How frequently were data types (b) and (c) inconsistent with one another, and what would be the source of this inconsistency if it's the same station?*

As clarified at P5, L1-8 of the revised manuscript, the sources of the inconsistencies could be various, according to the evolution of the monitoring systems of the different regions and often is due to the joint management of interregional basins. The different regional authorities often have adopted different codes/names for the same station, the first step has been thus to identify the presence of duplicate stations with same/similar name covering different time intervals. Sometimes, even for the same station, neighboring regions can provide different data for the same years. This can be, e.g., due to the fact that sometimes regions share rainfall data before their validation and official publication. That is why we decided, in case of inconsistencies, to preserve the data from the competent authority.

R#AM_C1 *Abstract: I would avoid the term "explosive rainfall". I understand the reason why the authors introduced it in the first sentence of the abstract, but I still believe that it would be advisable to use terms with a well-defined technical meaning.*

We have used the term “explosive rainfall”, recently become popular in the media and in the common language, to more easily reach a wider audience. The term is not commonly adopted by the average scientific community to which this journal is addressed, while meteorologists and climatologists are more inclined to relate heavy thunderstorms to explosive cyclones (see e.g. Kouroutzoglou et al, 2011) and this motivates the adoption of the term. However, we accept the suggestion to refer to a more technically sound term, as in P1, L1 of the revised manuscript.

R#AM_C4 Page 9, line 14: “A record-breaking event is defined as the annual value that exceeds all the previous ones.” Such definition implies a greater frequency of events at the beginning of the record. Did the authors consider identifying record-breaking events by fixing a threshold for rainfall intensity, basing on information that may be extracted by the whole record of observations (without introducing any assumption on the underlying probability distribution).

R#AM_C4: Page 9, line 14: it is stated that “At this stage, only nationwide record-breakings are considered, pulling up all the data together year by year.” I do not understand how data were pulled together. Did the authors pool together sites with different climatic behaviours? Does this mean that record-breaking events at the local level may have been discarded?

We are aware that to carry on analysis that involves pooling up of the data requires homogeneity controls and the wide domain under analysis does not warrant for climatic homogeneity. However, we propose a countrywide record-breaking analysis to provide a general assessment on the large rainfall amounts we are referring to when speaking about the “extremes of the extremes”. Location of the record breaking also point out their distribution in space. This analysis does not allow to consider record breaking events at the local level. We will clarify this in the revised version of the manuscript. On the other hand, our research on the topic is ongoing, as we are trying to identify a meaningful threshold for selection of homogeneous extraordinary events at the country scale. This additional analysis is still under development and will require a significant amount of efforts. We have then decided to devote a different paper to this analysis, that can have implication in terms of significance of increasing trends of short duration rainstorms.

R#1_C5 Figure 4. Excellent Figure. Page 8, lines 7-8 (“When short...appears.”). I agree with your statement, nonetheless this could also be an artifact due to the color scale chosen. Is there any particular reason for your choices of min/max values for color scale of the different durations? Does the min/max values correspond to any percentile of the distribution of values, for example?

The reviewer is right, the apparent distribution of the median values could be also an artifact of the chosen color scale (the min/max values correspond to the min/max values of the distributions). Nonetheless, the growing coefficient of variation of the national distribution of the median values when larger durations are considered seems to confirm a different distribution of the median values across the durations. These considerations are included in the revised manuscript at P8, LL12-14.

R#AM_C3 Page 3, line 11: it is stated that “Considering that most of the provided data have been validated from the related authorities, they are considered reliable and, at first, included directly in the I-RED” It would be interesting to discuss the validation tests that have been considered by the authorities. Extreme rainfall data may be affected by relevant uncertainty, it would be useful to mention the gauging methods, what kind of checks have been considered by the authorities and so on.

R#4_C1 Some discussion is needed on how the rainfall data at these stations is collected, as well as how the practice of recording rainfall data over time may have changed. For example, there may be some stations that reported 24 h rainfall totals for a longer period of time than 1 h rainfall recordings. This additional discussion could be aided by a plot similar to Figure 2d for the shortest (1 h) and longest (24 h) durations. Information on the types of gauges at these stations could be critical to understanding how capable the dataset is at capturing the most extreme rainfall amounts.

The reviewers are right. Fully detailed information on the measurement methods and on the validation of the data would be an interesting and useful information to be associated with the data. Unfortunately, despite we can certainly state that they respect the WMO standards we can not provide this kind of information as it is not generally published by the surveying authorities. The data managed from the SIMN are usually from manual gauges, like the ones described in Acquavita et al., 2016. The regional agencies then, with different timing, started to substitute the manual tipping-bucket rain gauges with automatic ones. Despite the information of the rain gauge model adopted is lacking for almost all the regions (except some rare cases, like the Veneto Region that provides a full report on the data: http://www.arpa.veneto.it/temi-ambientali/agrometeo/file-e-allegati/atlan-te-precipitazioni/20_Strumenti%20e%20criteri%20di%20osservazione%20e%20di%20gestione%20dei%20dati%20-

%20La%20serie%20pluviometrica%201984-2010%20dell2019ARPAV.pdf) we can suppose that the rain gauges adopted from ARPA Piemonte, described in Acquavota et al., 2016, are representative of the ones adopted from the different regional authorities. All the data are thus recorded with a time resolution that allows to capture the extreme rainfall amounts for all the considered durations (despite sometimes for some durations some data are lacking due to different motivations, e.g., missing reading of the record from the operator, etc.). These information are reported at P2, LL20-31 of the revised manuscript.

It is also not easy to find detailed information on the validation procedures adopted, despite many of the regional authorities have undersigned the document from the Italian Institute for Environmental Protection and Research about the validation procedure that should be adopted for the validation of meteorological databases. As the document is available only in Italian (Barbero et al, 2016) a summary is translated and reported in Appendix A and cited at P4, L3.

R#1_C4 Figure 2. Panels b and c shows exactly the same histogram. Please, check it.

R#4_C2 Figure 2 (b) and (c): These histograms look identical. Is it true that the same number of stations that reported 24 h rainfall also reported 1 h rainfall?

The distribution of the length classes is similar across the different duration. As it is not possible nor useful to show such similar behavior twice, figure 2 has been edited in the revised manuscript. Now panel (b) shows the number of series longer than fixed length values and, for underlying the differences across the durations, panel (c) report the number of null values per duration. Some lines on the source of the null values are added at P7, LL25-27.

R#2_C1 Figure 2: There is not any official abbreviation of years as "Y" so probably it would be more clear instead of t(Y) to write just "Years". Panels b and c: please decrease the size of fonts in the X-axis so it can be read more easily and also change the label to "Length (years)" or something similar as no.years is confusing (also check the panel d: do you mean years or data?). This suggestion is for any other Figure, e.g., for Figure 5 where "t(Y)" is used.

Figure 2 and 5 of the revised manuscript have been amended according to the reviewer's suggestions.

R#2_C2 Analysis presented in Figure 4. There is a dense network, more than sufficient to provide kriging estimates for the whole Italy. It would much more useful in my opinion not only because dots may overlap but because you will provide estimates also in places where there is no information. So, I would suggest to construct kriging maps of the statistics analyzed.
E Some supplementary material on the spatial variability of rainfall in Italy (maps) will also be more than welcome.

Considering the exploratory nature of our analysis we did prefer to provide point statistics, which –at this stage– indicate the main features of rain variability at the country scale but do not address the goal of providing robust estimates in ungauged areas. Considering the many problems we have faced in the regional analysis in the North-West of Italy (Libertino et al, 2018) the uneven distribution of the rain gauges in space and the high variability of record length make the whole kriging analysis a challenge deserving a separate work. In partially following the reviewer's suggestion, however, interpolated maps will be added as supplementary material to more clearly represent rainfall indices variability over space.

R#2_C3 Mean is quite robust in general, yet here you prefer only the median. Of course, it can be affected by outliers yet common methods of fitting distributions, e.g., product moments or L-moments are using mean values. So, in my opinion you should provide also the maps of the mean value.

R#2_C4 If I understood well you have estimated the mean values of L-CV, L-skew and L-kurt of all duration. Of course you are dealing with maxima and we are expecting the shape characteristics to be close yet this is not necessarily true. If indeed these summary statistics are close among the different duration, please report it or else provide different maps for each duration.

The reviewer is right: most of the common methods of fitting distributions consider the mean as best descriptor of the central tendency. However, in this descriptive phase we wanted to represent the variability of the central tendency of the series as unbiased as much as possible with respect to the outliers. On the other hand, we admit information of the mean can be of interest for the readers: therefore, we have added a map of the mean of the extremes in the supplementary material as reported at P8, L11 of the revised manuscript. Brief considerations on the behaviour of the L-moment statistics across the different durations have been added in the revised version

of the manuscript at lines P8, LL18-19 and P9, L1, and the maps of the ratios for the different durations added in the supplementary materials.

R#2_C5 *It is not clear to me if the maxima values have emerged from a sliding-window process or from a fixed-block (non-overlapping). In the latter case the user of this database should know this fact in order to correct the data by the Hershfield factor. Please comment on that and clarify.*

Annual maxima have been collected with a sliding-windows procedure. The information is added in the revised version of the manuscript at P2, L20.

R#2_C6 *You can use plain text for L-CV. In Figure 4 is plain while in the text you are using Italics. Please re-check the text for minor typos, e.g., line 26, p6 replace "an unique" with "a unique".*

R#4_C5 *There are a few typos in the paper, including the following: Page 1 Line 24: "dismantledment" should be "dismantlement", Page 1 Line 25: "stucked" should be "stuck", Page 2 Line 2: "metodologies" should be "methodologies".*

The typos reported by the reviewer have been amended and the manuscript double-checked for more typos.

References

- Acquaotta, F., Fratianni, S., & Venema, V. (2016). Assessment of parallel precipitation measurements networks in Piedmont, Italy. *International Journal of Climatology*, 36(12), 3963-3974.
- Barbero S. et al. (2017): Linee guida per il controllo di validità dei dati idro-meteorologici. ISPRA – Manuali e Linee Guida 156/2017. Roma, 2017.
- Kouroutzoglou, J., Flocas, H. A., Keay, K., Simmonds, I., Hatzaki, M. (2011). Climatological aspects of explosive cyclones in the Mediterranean. *International Journal of Climatology*, 31(12), 1785-1802.
- Libertino A., P. Allamano, F. Laio, and P. Claps. Regional-scale analysis of extreme precipitation from short and fragmented records. *Advances in Water Resources* 2018, 112, 147-159.

Technical note: Space-time analysis of rainfall extremes in Italy, clues from a reconciled dataset

Andrea Libertino¹, Daniele Ganora¹, and Pierluigi Claps¹

¹Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Torino, ITALY.

Correspondence: Andrea Libertino (andrea.libertino@polito.it)

Abstract. Like other Mediterranean areas, Italy is prone to the development of events with ~~explosive~~-significant rainfall intensity, lasting for several hours. The main triggering mechanisms of these events are quite well known but the aim of developing rainstorm hazard maps compatible with their actual probability of occurrence is still far from being reached. A systematic frequency analysis of these occasional highly intense events would require a complete countrywide dataset of sub-daily rainfall records, but this kind of information was still lacking for the Italian territory. In this work several sources of data are gathered, for assembling the first comprehensive and updated dataset of extreme rainfall of short duration in Italy. The resulting dataset, referred to as Italian Rainfall Extreme Dataset (*I-RED*), includes the annual maximum rainfalls recorded in 1 to 24 consecutive hours from more than 4500 stations across the country, spanning the period between 1916 and 2014. A detailed description of the spatial and temporal coverage of the *I-RED* is presented, together with an exploratory statistical analysis aimed at providing preliminary information on the climatology of extreme rainfall at the national scale. Due to some legal restrictions, the database can be provided only under certain conditions. Taking into account the potentialities emerging from the analysis, a description of the ongoing and planned future work activities on the database is provided.

1 Introduction

Italy can boast of a role at the highest level in the development of meteorological observations (Brunetti et al., 2006), with 6 meteorological stations operating since the eighteenth century (Bologna, Milano, Roma, Padova, Palermo and Torino), and 15 stations with observation starting in the first half of the nineteenth century. First attempts of performing a systematic collection of monthly rainfall data go back to as early as 1880 when the National Office for Meteorology and Climate was funded. The National Hydrographic Service (*SIN*) and the National Hydrographic and Mareographic Service (*SIMN*) collected annual maxima values for 1-3-6-12 and 24 hours durations in the Hydrological Yearbooks from 1917 to early 2000s (the final publication year depends on the local agencies of the *SIMN*). The Legislative Decree 112/1998 dismantled the *SIMN*, transferring its tasks to the 19 administrative regions and the 2 autonomous provinces of Trento and Bolzano. These authorities were designated as local Operational Centres and Regional Environmental Agencies to deal with hydro-meteorological monitoring and civil protection issues.

~~In spite~~-In spite of the huge heritage of data, only a small fraction of the Italian rainfall data is available in a computer-readable format. Moreover, the ~~dismantledment~~-dismantlement of the National Service led to a lack of update of the national database

of extreme rainfall that is still ~~stueked~~stuck, for some regions, at the beginning of the '90s. This has led to a very fragmented framework: updated rainstorm hazard assessments are actually only available for some regions and only at the regional scale (see, e.g., Ubaldi et al. (2014); Libertino et al. (2017)). Various regional studies present different ~~metodologies~~methodologies and are sometimes based on very different data densities and record lengths (e.g., Claps et al., 2016), but only few updated
5 analyses on short-duration rainfall in a over-regional framework are available (e.g., Rudari et al., 2005; Allamano et al., 2009).

In view of the assembling of the first comprehensive dataset of extreme rainfall of short duration in Italy several major sources of data have been analysed. The resulting dataset, referred to as Italian Rainfall Extremes Database (*I-RED*), includes data from more than 4500 stations across the country, spanning the period between 1916 and 2014, and refers to annual maximum rainfall recorded in 1 to 24 consecutive hours (exact durations available are 1-3-6-12 and 24 hours).

10 The following sections describes the sources of the data, the work carried out for the merging of the database and the operations that are still required for making it suitable for a nationwide robust rainfall frequency analyses. A preliminary analysis of the extreme rainfall regime at the national scale is also presented.

2 Merging the *I-RED* Dataset

2.1 Data sources

15 As a follow-up of the activities of the Italian National Group for the Prevention of the Hydrogeological Disasters (*GNDCI*) a comprehensive nationwide hydrological information system has been set up, within the “*CUBIST* project”, funded by the Italian Ministry of Education and Research (*PRIN* 2005). The database includes about 6000 pluviographs and pluviometers, 700 temperature stations and about 400 river basins (Claps et al., 2008) and is available at: <http://www.cubist.polito.it>; accessed: 2017-10-26. In the detail, the database includes rainfall data from 1900 to 2001 (depending on the region) and constitute the
20 first important attempt of making the large Italian hydrological heritage freely available in a computer-readable format.

~~The~~The annual maxima data for different durations included in the *CUBIST* dataset are extracted with a sliding-windows process from manual tipping-bucket rain gauge data, equipped with a recording system that writes on diagram paper. Further information on the characteristics of the stations can be found in Acquotta et al. (2016). The number of data per year is not constant across the analysed period, being increasing in time as more stations have been installed in the recent years. Data
25 availability decreases in the period of the Second World War, as many records have been missed in that period. After 1980, with the progressive dismissal of the *SIMN* and the development of the local hydrographic authorities, data availability decreases rapidly until 2001, when the rain gauges still under the *SIMN* were taken over by the local Operational Centres.

After the late '80s, indeed, the local Environmental Agencies started to support the *SIMN* in its work. Gradually, the 21 regional hydrological services took over the networks and the tasks of the national one. In this period most of the old manual tipping-bucket rain gauges have been substituted with automatic stations, similar to the one described in Acquotta et al. (2016) for the Piemonte region. Each hydrological service adopted its own rules for the publication of the collected data and, even if
30 the Italian law adopted ~~an Open Source policy~~an open source policy for non-commercial uses (under D.Lgs.82/2005, D.Lgs.36/2006, D.M.10/11/2011, L.221/2012, D.Lgs.179/2012 and L.114/2014), an updated database of the

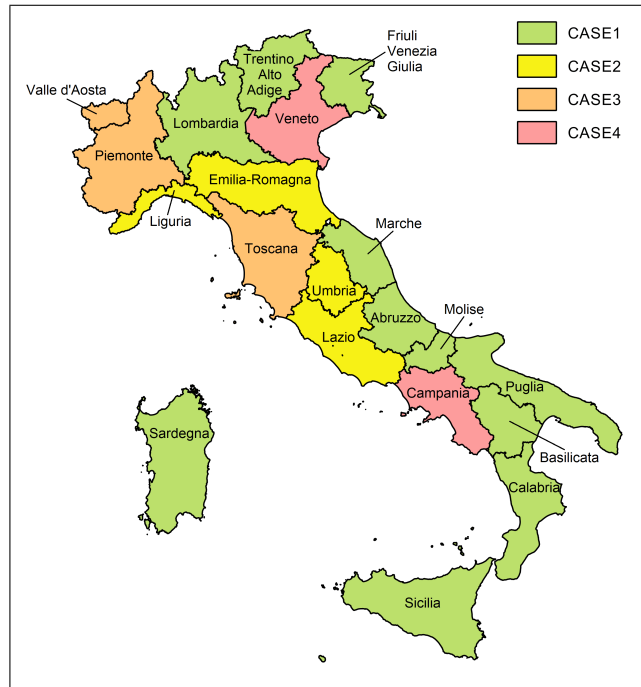


Figure 1. Names of the Italian regions and type of datasets provided by the regional authorities. The cases refer to the ~~bulleted-bullet~~ list of section 2.2.

annual rainfall maxima for sub-daily duration at the national scale is still lacking. For the scopes of this research, the different agencies have been contacted, requesting the regional annual maxima datasets for sub-daily durations. The regions of Italy are shown in ~~Figure-figure~~ 1 together with the type of data provided, that will be described in the following section. Table 1 lists the name of the local authorities and the regional codes, aimed at identifying them in the database. The public availability of the original dataset is also reported. ~~The different providers adopt various data policies, raising significant problems with the publication of the merged dataset.~~

2.2 Cleaning and merging operations

Merging and harmonizing the different datasets is a quite long and difficult operation, that is still ongoing. The different Operational Centres provided different types of datasets, with different temporal coverages and spatial reference systems.

10 Duplicate stations are often present in the databases of ~~neighboring-neighbouring~~ regions.

The first steps of this work have been carried out at the regional scale. For each region all the data falling inside the regional boundaries have been considered. These data, according to the setting of the databases of the local Operational Centers, could belong to one of these 3 categories:

- a) Data from the *CUBIST* database for the 1900-2001 period already available from the former national service

Table 1. Regions of Italy with the assigned code and the related local Operational Center with references to the availability of digitized data.

CD	Region	Operational Center	Digitized data availability
01	Abruzzo	Ufficio Idrografico e Mareografico Regione Abruzzo	Under-Available upon request
02	Basilicata	Dipartimento Protezione Civile Regione Basilicata	Available in ¹
03	Calabria	Centro Funzionale Multirischi - ARPACAL	Available at ²
04	Campania	Centro Funzionale Regione Campania	Under-Available upon request
05	Emilia-Romagna	ARPA Emilia-Romagna	Under-Available upon request
06	Friuli Venezia Giulia	Ufficio Idrografico Regione Autonoma Friuli Venezia Giulia	Under-Available upon request
07	Lazio	Centro Funzionale Regione Lazio	Under-Available upon request
08	Liguria	ARPAL-CFMI-PC	Partially available at ³
09	Lombardia	ARPA Lombardia	Available at ⁴
10	Marche	Dipartimento di Protezione Civile Regione Marche	Available at ⁵
11	Molise	Centro Funzionale Regione Molise	Under-Available upon request
12	Piemonte	ARPA Piemonte	Partially available at ⁶
13	Puglia	Dipartimento di Protezione Civile Regione Puglia	Available at ⁷
14	Sardegna	ARPAS	Under-Available upon request
15	Sicilia	Osservatorio delle Acque Regione Siciliana	Under-Available upon request
16	Toscana	Servizio Idrografico Regionale Toscana	Available at ⁸
17	Trento *	Centro Funzionale Provincia Autonoma di Trento	Available at ⁹
18	Bolzano - Alto Adige *	Ufficio Idrografico Provincia Autonoma di Bolzano - Alto Adige	Under-Available upon request
19	Umbria	Regione Umbria	Under-Available upon request
20	Valle d'Aosta	Centro Funzionale Regione Autonoma Valle d'Aosta	Under-Available upon request
21	Veneto	ARPAV	Under-Available upon request

* the Autonomous Provinces of Trento and Bolzano - Alto Adige, together, constitute the region Trentino Alto Adige (CD: 22)

¹ Manfreda, S., Sole, A. and De Costanzo, G.: Le precipitazioni estreme in Basilicata, Editrice Universo Sud, 2015.

² ARPACAL: Centro Funzionale Multirischi, <http://www.cfd.calabria.it/>, accessed: 2016-08-01

³ ARPAL: Consultazione Dati Meteorologici, <http://www.cartografiar.regione.liguria.it/SiraQualMeteo/Fruizione.asp>, accessed: 2016-08-01

⁴ ARPA Lombardia: Progetto Strada, <http://idro.arpalombardia.it/pmapper-4.0/map.phtml>, accessed: 2016-08-01

⁵ Protezione Civile Regione Marche: Annali Idrologici Regione Marche, <http://console.protezionecivile.marche.it>, accessed: 2016-08-01

⁶ ARPA Piemonte: Banca dati meteorologica, <http://www.regione.piemonte.it/ambiente/aria/rilev/ariaday/annali/meteorologici>, accessed: 2016-08-01

⁷ Protezione Civile Puglia: Annali Idrologici - Parte I, <http://www.protezionecivile.puglia.it/centro-funzionale/analisielaborazione-dati>, accessed: 2016-08-01

⁸ SIR Toscana: Settore Idrologico Regionale, <http://www.sir.toscana.it/>, accessed: 2016-08-01

⁹ Centro Funzionale di Protezione Civile Provincia Autonoma di Trento: Meteotrentino, <http://www.meteotrentino.it/>, accessed: 2016-08-01

b) Data provided by the regional authority

c) Data provided by the regional authorities of the ~~neighboring~~-neighbouring regions, falling out of their regional borders

Observations dating before 1916 have been discarded, as considered not significant and too unevenly distributed. Considering that most of the provided data have been validated from the related authorities, they are considered reliable and, at

first, included directly in the *I-RED*. ~~In~~ For information on the validation procedures, please refer to the Appendix A and to Barbero S. et al. (2017). In the presence of inconsistencies between the type b) and type c) data, preliminary manual merging was carried out. The sources of the inconsistencies could be various, according to the evolution of the monitoring systems of the different regions and often is due to the joint management of interregional basins. The different regional authorities often have adopted different codes/names for the same station, the first step has been thus to identify the presence of duplicate stations with same/similar name covering different time intervals. Sometimes, even for the same station, neighbouring regions can provide different data for the same years. This can be, e.g., due to the fact that sometimes regions share rainfall data before their validation and official publication. If the same station was found in the database of more ~~neighboring~~ neighbouring regions a first attempt of merging the series together was carried out, by analysing the data recorded year by year. If the merging was not feasible, higher priority was given to the data provided by the authority of the considered region (that is usually also the owner of the network). This allowed to avoid the presence of duplicate series in the *I-RED*.

Once merged, for each region, type b) and type c) datasets, the resulting dataset has to be merged with the type a) dataset. This operation has been quite complex, as the overlapping period between the different dataset was different for each region and because most of the authorities did not tracked the change in the name/code of the stations. The different procedures performed, according to the type of the dataset that the region has provided (as reported in figure 1) can be summarized as follows:

1. *Regions that digitized the whole SIMN database for their areal domain and provide a complete merged database.* The provided data were inserted in the *I-RED* without editing and without considering the *CUBIST* series. Only for the Abruzzo and Molise regions some preliminary refinement was needed, as the two regions were divided in 1963, and the databases of the two regions partially overlap. The stations were then divided according to the actual regional boundaries and the duplicate series removed.
2. *Regions that provided datasets including data from their actual regional network partially merged with subsets of digitized data from the SIMN Hydrological Yearbooks.* As not all the *SIMN* datasets were digitized from the local authorities, the dataset lacked part of the stations included in the *CUBIST* database. To maximize the available information, data from the regional databases and the *CUBIST* one were manually analyzed and merged, in order to avoid duplicate values. Stations with same name and similar coordinates were merged together in the presence of a two-year consistent overlapping period. In the presence of inconsistencies between the values recorded by the two stations two stations in the overlapping period, they were treated as different stations and renamed. If that was not possible to unravel any doubt, the stations were considered as separate entities. For the Liguria region, the information in ARPAL (2013) was used to overcome the lack of information on the continuity of the series.
3. *Regions that provided two different datasets: one containing the whole digitized data from the SIMN stations and another containing the digitized data from their actual networks.* The data of the two databases were merged together, the overlapping period manually ~~analyzed~~ analysed to avoid overlapping, and the *CUBIST* database ignored. The operation

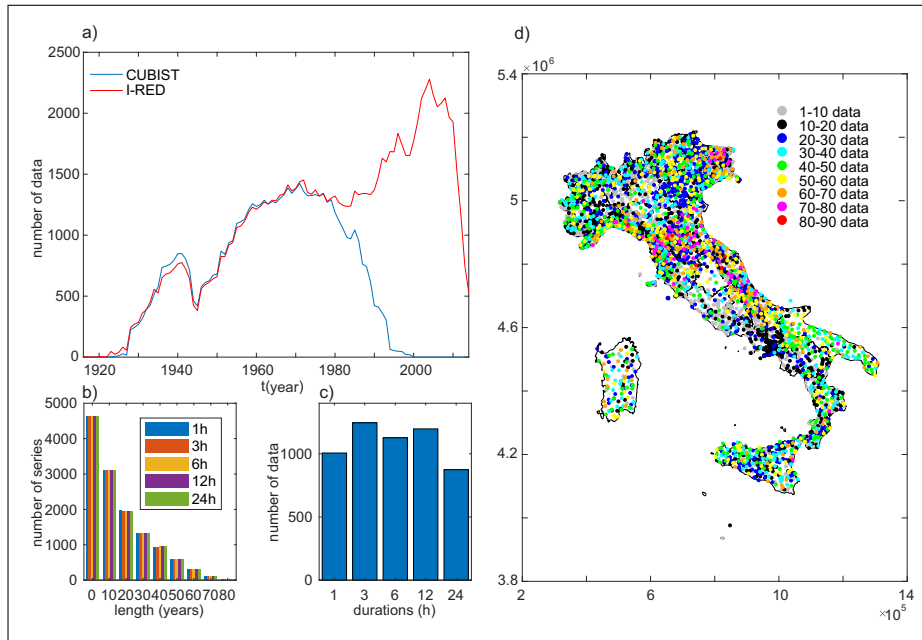


Figure 2. (a) Data availability per year in the *I-RED* and *CUBIST* databases (the smallest value across the 5 considered duration is reported per each year). (b) Number of series per length class longer than fixed threshold values in the *I-RED* (red) and *CUBIST* (blue) databases for durations 1 per duration (null values are ignored) and 24. (c) hours Number of null values per duration. (d) Length of the series in the *I-RED* database represented in space: the color refers to the minimum length among the 5 available durations. If more stations overlap due to the resolution of the picture, the one with the longer series appears on the top.

was made possible by the collaboration of *ARPA* Piemonte, for Piemonte and Valle d’Aosta, and of the Università degli Studi di Firenze, for Toscana.

4. *Regions that provided only the data recorded from the network they actually manage.* All the information concerning the *SIMN* stations was lacking. The provided dataset was therefore merged with the whole *CUBIST* database for the considered regions. Duplicate values were excluded analyzing manually the overlapping period, if present.

With the application of the above described rules, 20 complete regional datasets have been obtained. The regional datasets were finally merged together to generate the *I-RED*. After the merging phase some reliability check has been performed, in order to detect any problematic or incorrect information. They includes the identification and removal of the duplicate data/stations and reliability checks on the larger values of the dataset, comparing them to the absolute record-breaking events for all the durations (see *Libertino (2017)*), aimed at detecting inconsistencies in rainfall series. If any suspect value was found, its year of occurrence was compared, when referring to recent years, with the data from event reports or newspapers. If the data refers to a *SIMN* station, the *Hydrological Yearbooks* were consulted. If no evidence was found, the related authority was

contacted. Most of the operations need human supervision, and a thorough verification work. If it is not possible to unravel any doubt the suspect value is discarded.

Due to the complexity of the check operations, further efforts and collaborations with the regional authorities are still ongoing to increase the consistency of the database. Nevertheless, to date (October 2017) the *I-RED* includes more than 4500 stations nationwide and constitute the largest updated dataset of annual maxima for Italy.

Considering that most of the regional authorities supervise the use and widespread of their datasets for preventing improper uses, a detailed description on how to access the *I-RED* is reported in the “Data availability” box.

In the following, the spatio-temporal distribution of the assembled data will be described.

3 Main features of the *I-RED* database

The number of data available per year in the *I-RED* is reported in figure 2a, as compared with that of the *CUBIST* database. As every station is related to ~~an~~ a unique value of annual maxima for a given duration, the presence of a measurement implies the presence of a station. The number of available stations increases with time, and drastically grows after the dismissal of the *SIMN* and the development of the local agencies. The decrease after 2010 can be attributed to the fact that not all the regions have published the data for the most recent years.

The smaller size of the *I-RED* compared the *CUBIST* database in some years can be due to:

- The presence, before 1945, in the *CUBIST* database of data from territories lost by Italy after World War II (e.g. , Istria) or from ~~neighboring~~ neighbouring countries, not included in the *I-RED*;
- The fact that regional agencies could have decided for different reasons not to include data or stations from the *SIMN* dataset in their database. Part of these data could therefore be lost not considering the *CUBIST* database for these regions.

Considering the limited significance of the information loss, further efforts for including these data will be planned only in a future stage of the development of the database.

For a descriptive analysis of the rainfall data, all the assembled time series are classified according to their length. Results are shown in ~~Figure 2band 2c for the 1 and 24 hours durations, respectively. The other durations (not shown) report a similar behaviour. figure 2b.~~ Considering the short life of the rain gauges installed by the regional Operational Centers, ~~most of the new series are a large percentage of the series is~~ shorter than 20 years but the contribution of the *CUBIST* database allows for a significant amount of longer series(~~see, .~~ The series with more than 80 years of data are for the 1, 3, 6, 12, 24 hours durations are respectively 16,14,15,14 and 17. In general, all the durations report a similar behaviour, despite some differences in the distribution of the null values as shown in figure 2c. The reasons that lead to missing data only for certain durations can be various and related to either the measuring, the recording or the storage of the data (e.g., the 60-70 data classmissed reading of the record from the operator, data classified as not-valid in the validation phase, etc.).

The spatial distribution of the stations is shown in figure 2d. The color scale refers to the number of the available data per each series. The minimum number across the 5 duration is considered. One can clearly distinguish that, even if the whole

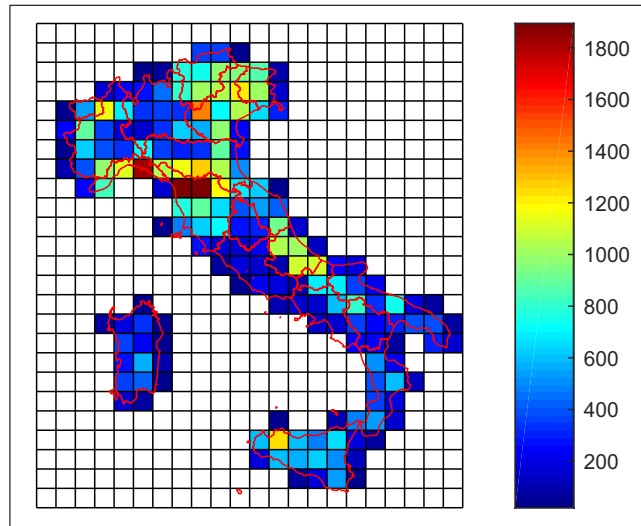


Figure 3. Number of station-year per cell over a 50 km grid.

national territory is represented, the density of the stations widely changes across the nation. To show the relevance of the non-uniformity, a gridded domain with a mesh size of 50 km is introduced. Figure 3 shows the number of station-year, i.e., the total number of data per cell, showing all the available data of the stations located within the cell. If data consistency changes for the different durations, the shortest one is considered. The non uniformity of the network density clearly emerges at first sight, with some cells presenting almost 10 times the number of data of other cells. The most densely gauged cells can be found in the North-West of the country, in particular in Liguria region, in the northern Toscana and, in the North-East.

4 Descriptive statistical analysis of rainstorms in Italy

A preliminary descriptive analysis of the characteristics of extreme rainfalls at the national scale has been carried out on the newly developed *I-RED* database. Series with a minimum length of 20 years of data have been considered in this analysis. This length constraint leads to a subset of 1974 series available for the analysis, out of the original 4686. For each duration, the median of the series is depicted in figure 4. The median is used as a robust estimator of the central tendency of a series, less sensitive than the mean to the presence of outliers. As common methods of fitting distributions, e.g., product moments or L-moments are using mean values for representing the central tendency, maps of the mean for the different durations are attached in the supplementary material.

~~When short durations are considered, stations with larger median values seem to be evenly distributed across the country. Moving to longer durations, a marked spatial clusterizations of the series with larger median value appears~~ Some geographical areas are characterized by clusters of large median values and these clusters appear consistent across the different durations. Furthermore, at the country-wide scale we observe that the coefficient of variation of the medians increases for increasing durations, suggesting a wider range of variability of the corresponding median values.

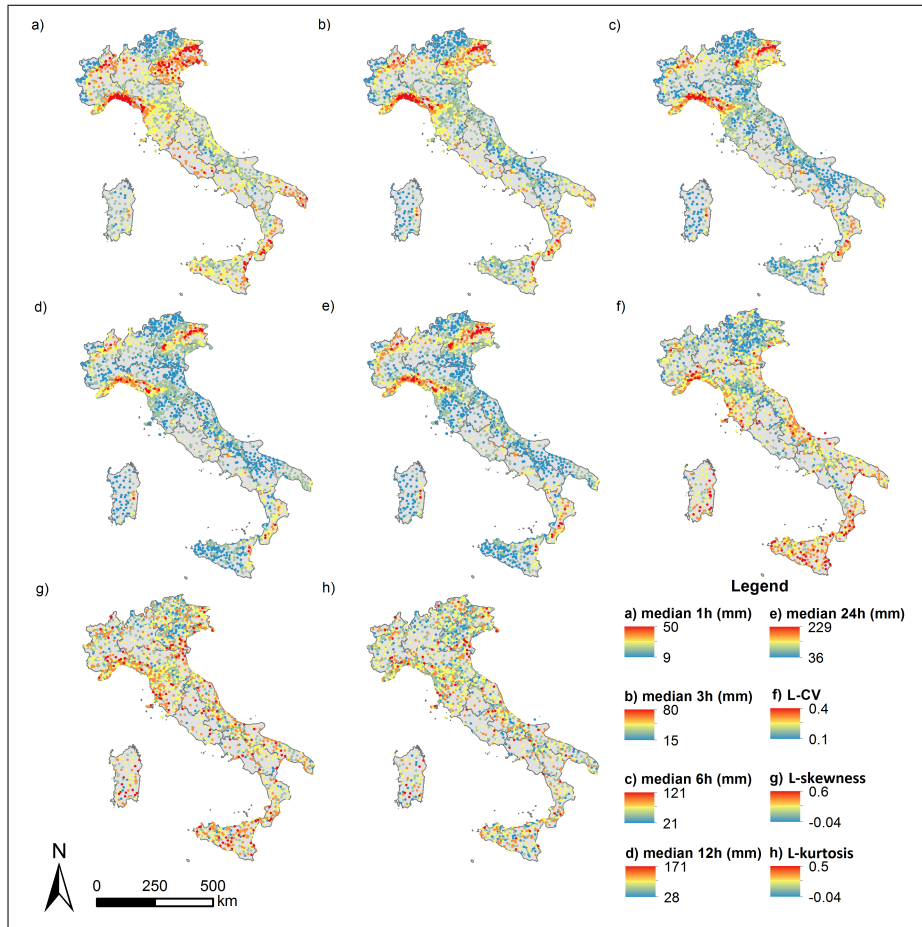


Figure 4. Median values of the *I-RED* series from 1 (a) to 24 (e) hours. Average statistics for the five durations considered: (f) $L-CV$, (g) L-skewness and (h) L-kurtosis. Series with more than 20 data are considered.

For each series, the sample L-moments (Hosking and Wallis, 1997) have then been computed to describe the shape of the empirical distribution of the records. The mean L-moments ratios among the different durations give information respectively on the dispersion ($L-CV$), skewness (L-skewness) and “peakedness” (L-kurtosis) of the empirical distributions. All the above statistics are mapped in figure 4. Considering that the L-moments ratios show similar behaviour for the considered durations, we decided for simplicity to include in the paper only the average ones. The maps for the different durations are reported in the supplementary material. Panel (f) shows that the coastal areas and the islands are generally characterized by a higher variability in the annual maxima series, presenting larger $L-CV$ values. The northern part of the peninsula, even if characterized by large median values, shows lower $L-CV$, which is typical of areas with large average rainfall values. It is harder to identify a precise spatial pattern in the distribution of the skewness and kurtosis values (panels (g) and (h)). Coastal and island areas seem to generally show larger skewness values, confirming the influence of the Mediterranean sea on

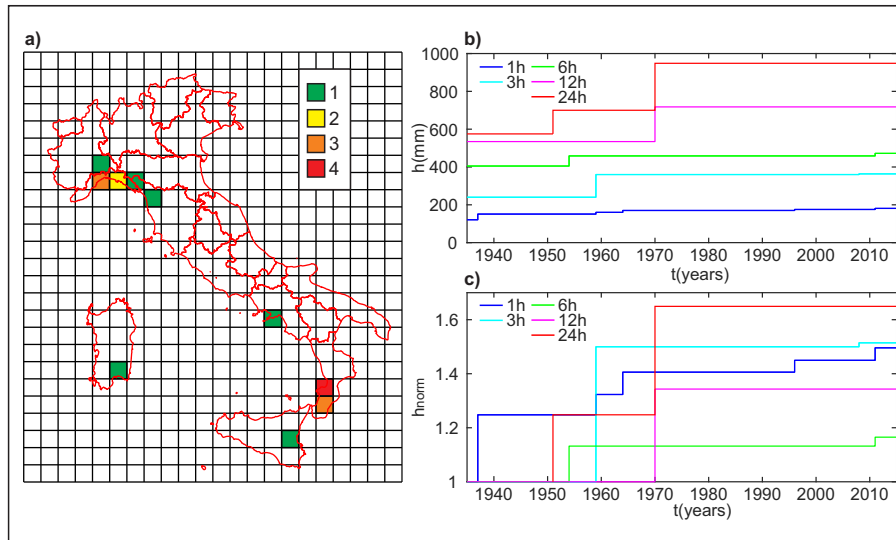


Figure 5. (a) Number of record-breaking events per cell over a 50 km grid. Record breaking rainfall depths for the 5 considered durations from 1935 to 2015 (b) in absolute values and (c) normalized on the 1935 values.

the climate of these areas. [All the aforementioned maps have been also interpolated for visualization purposes with ordinary kriging; detailed results are reported in the supplementary material.](#)

The significance of the developed dataset allows also to preliminary explore the rainfall events sometimes referred to as “black swans” (Blöschl et al., 2015), showing extraordinary intensities even when compared with the population of annual maxima. In Italy, many of these events have been studied as individual extraordinary events (e.g., Rebora et al., 2013; Fiori et al., 2014), due to the large recorded intensities and to their severe consequences, but the fragmented configurations of the national database have prevented a systematic treatment of this population of “extremes of the extremes” (Snorrason et al., 2002). A preliminary investigation on the occurrence of very-extreme events at the national scale has been performed and summarized in figure 5a that shows the spatial distribution of the record-breaking rainfall events for the considered durations from 1935 to 2015. A record-breaking event is defined as the annual value that exceeds all the previous ones. At this stage, only nationwide ~~record-breakings~~ [record-breaking](#) are considered, pulling up all the data together year by year. Record breaking rainfall amounts can provide a picture of the spatio-temporal distribution of the major weather anomalies in the country. Analysing record-breaking events has some advantages from both an operational and a statistical point of view. Due to the significant amounts recorded, these events can be easily verified combining different sources of information, and, moreover, this kind of analysis does not require any assumption on the underlying probability distribution (Coumou et al., 2013). The spatial distribution of the events seems to suggest a clusterization of these phenomena in some areas of the country: the eastern part of Liguria and northern part of Toscana and the extreme south of Calabria. Localized events also occurred in Campania, Sicilia and Sardegna. All of these areas are generally characterized by complex orography in the proximity of the coastline: a framework that can promote the development of particularly intense phenomena (Furcolo et al., 2015). At-site systematic

analysis of the record breaking events, which are expected to provide useful information for characterizing the extreme rainfall regime in the country (Lehmann et al., 2015), is now possible thanks to the consistency of the new *I-RED* database. Figure 5b shows the record-breaking evolution in time (for each duration); the occurrence of a new record-breaking is represented by an increasing step in the line. Panel (c) report the same ~~record-breakings~~record-breaking, whose values are normalized by the 1935 values.

5 Conclusions

The first comprehensive dataset of extreme rainfall in Italy, called *I-RED*, has been presented here. It is a significant source of information, able to provide unprecedented knowledge on the characteristics of heavy precipitation in Italy and on the possible rainfall regime changes in the last century. Further efforts will be addressed to increase the spatial data homogeneity and coverage in time, by including the data of the most recent years and, eventually, by contacting the local authorities for requesting assistance in the merging of the series. The final aim is to make the update of the database systematic and unsupervised. This can be obtained by strengthening the collaboration with the data providers, in the framework of joined projects, as the one that led to the development of the *ArCIS* (Pavan et al., 2013) dataset, collecting updated rainfall and temperature data from a group of regional authorities in Northern Italy. Collaborations with other projects, focused on different spatial or temporal scales, will be also explored in order to automatically and efficiently analyse the consistency of the *I-RED* dataset and to integrate it with the existing ones. A possible target is the *SCIA* dataset (Desiato et al., 2007) referring to the 24-hour and daily scale. Joined projects with international institutions will be evaluated and endorsed in order to make available the *I-RED* database in larger frameworks for trans-boundary exchange of precipitation data. In the meanwhile the *I-RED* will be used for exploring the different outcomes provided by this preliminary analysis, e.g., assessing the influence of the spatial distribution of the stations on the observation of record-breaking extreme events, evaluating the presence of trends in the temporal distribution of the “black swans” and analysing the statistical predictability of these kind of events on such a wide and complex domain.

Data availability. The original data can be requested to the authorities reported in table 1. Some of the agreements signed with the data providers, aimed at monitoring the correct use of the data, restrict their use to the aims of the authors’ project. Due to these legal restrictions, the full or partial access to the *I-RED* can be provided:

- to research individuals or groups in the framework of the authors’ project;
- to research individuals or groups not collaborating with the authors’ project, upon evidence of permission received by the involved regional agencies, reported in table 1.

For further details and queries, please refer to the corresponding author.

Appendix A: Guidelines for the quality check of hydro-meteorological data

Extracted and translated in English from Barbero S. et al. (2017)

A1 Quality control

[...] The attribution of a certain level of quality to the measured data passes through the process of validation of the data themselves, which consists in analysing all the data collected in terms of completeness, reasonableness and in eliminating erroneous values. Data validation (validity check) is only one of the Quality Control operational procedures (QC) consisting of a set of procedures and rules to ensure that a measurement system achieves and maintains a specific quality level initially established. The periodic calibration of the instruments, the periodic inspection of the sites and the preventive maintenance also belong to the QC process. The QC can be applied both in real time (real time quality control) that in delayed time, according to the needs of sharing, using and storing data nationally and internationally (e.g., World Meteorological Organization, 1993, 2010a, b). Moreover, the QC must be integrated into an effective and well-coordinated Quality Management (QM). The QM is, in fact, is expressed through the joint application of Quality Assurance (QA) and of the QC (World Meteorological Organization, 2014). The QA is the set of planned and systematic activities applied within a quality management system to provide the level of confidence with which the quality requirements are met. Basically, the QC is a system of activities to provide a quality product while the QA is a system of activity designed to verify that the quality control system is functioning properly. The main objectives of the QM are the identification, quantification and reduction of errors. Errors can be made for both technical reasons (i.e., due to the used methods and technologies) and procedural reasons (i.e., linked to an unclear or ineffective management or to the lack of adequate preparation of the operators). Furthermore, errors can be made during detection (e.g., the sensor does not read correctly), when the observation is transcribed (transposition of digits, shifting of dates, etc.) and during data transmission and storage (computer errors, errors of digitization, etc.). Many of these errors can be prevented by an appropriate QA, others must be identified and corrected through the QC procedures.

A2 Levels of the validity check

The first level of data validation is performed on the raw data (or gross data), i.e., the data at the original temporal resolution with which they are transmitted or detected at the measuring station and consists in the application of basic procedures for verifying the validity of the data. These checks aim at indicating malfunctions, instability, or interference. In the case of data coming from automatic measuring stations the validity checks are applied to the “meteorological message” coming from the station in the transcoding phase of the message that for the transmission must comply with certain rules. The checks carried out will therefore be related to the expected formats within a given message, to the date and time stamps, to the location of measuring station, to the codes of stations and sensors and to the presence of duplicate elements. This category of checks includes: syntax controls (e.g., alphabetic characters appearing in a text that should be numeric) which, if incorrect, can mine the transcoding process; logical controls that refer to both the intrinsic characteristics of the magnitude (e.g., World Meteorological Organization, 1993) and to the limits imposed by technical characteristics of the instrument used,

in terms of measuring range (e.g., for a rain gauge: 0-300 mm/h), resolution (e.g., for a rain gauge: 0.1 or 0.2 mm), and limits in the operating temperatures (e.g., 0-70°C for an unheated or -30-70°C for a heated rain gauge). The first level of controls can have three types of outcomes: control passed, suspected data or control not passed. In the first two cases the data is then subjected to subsequent checks, in the third the data is considered incorrect and discarded. The first level controls are performed on the elementary data, i.e., the data to the temporal aggregation derived from the measurement station. At this stage it is appropriate also carry out internal consistency checks. These are checks that are based on the comparison between synchronous values of different variables somehow related (e.g., by physical laws), so as to highlight any inconsistencies between the data. The second level of data validation consists of a series of “consistency” checks:

- Time consistency checks: they are based on the verification of the respect of a maximum and minimum level of variability of data over time and they have the purpose of identifying any anomalies between data temporally contiguous or with respect to the values that have historically occurred in a given site. Concerning the allowed minimum variability, consistency verification procedures are aimed at ascertaining the presence of persistence of measured values in the series, consisting in the lasting over time of a same or similar value;
- Cross checks with other quantities recorded at the same station: they are based on the control of the considered data with reference to other related quantities measured at the same site. E.g., temperature comparison with solar radiation;
- Spatial consistency checks: they are based on the hypothesis of gradual variability of the observed quantity in space and therefore on the existence of a sort of spatial correlation between the contemporaneous measures carried out in neighbouring stations. However, when dealing with rainfall, the hypothesis of gradual variability is lesser acceptable when smaller temporal aggregations are considered;
- Climatological checks: they are based on comparing the quantity under examination with some parameters derived from the whole historical series (e.g., tests based on the comparison with percentiles calculated on specific time intervals). The data are validated, at first, using automatic procedures. However, for the evaluation of the so-called “suspicious” data, a manual revision by qualified personnel is required to decide for every case if validate the suspect data, reject it as not valid, or fix it if possible.

25 *Competing interests.* No competing interests are present.

Acknowledgements. The authors thank Prof. Enrica Caporali and Dr. Valentina Chiarello for their assistance in preparing and screening the Toscana regional dataset and Dr. Stefano Macchia for his contribution in collecting and cleaning the data. Data providers reported in table 1 are acknowledged.

References

- Acquaotta, F., Fratianni, S., and Venema, V.: Assessment of parallel precipitation measurements networks in Piedmont, Italy, *International Journal of Climatology*, 36, 3963–3974, 2016.
- Allamano, P., Claps, P., Laio, F., and Thea, C.: A data-based assessment of the dependence of short-duration precipitation on elevation, *Physics and Chemistry of the Earth, Parts A/B/C*, 34, 635–641, 2009.
- 5 ARPAL: Atlante Climatico della Liguria, Final report, Regione Liguria, Genova, IT, 2013.
- Barbero S. et al.: Linee guida per il controllo di validità dei dati idro-meteorologici, 156, ISPRA – Manuali e Linee Guida, Roma 2017, 2017.
- Blöschl, G., Gaál, L., Hall, J., Kiss, A., Komma, J., Nester, T., Parajka, J., Perdigão, R. A. P., Plavcová, L., Rogger, M., Salinas, J. L., and Viglione, A.: Increasing river floods: fiction or reality?, *Wiley Interdisciplinary Reviews: Water*, 2, 329–344,
- 10 <https://doi.org/10.1002/wat2.1079>, <http://dx.doi.org/10.1002/wat2.1079>, 2015.
- Brunetti, M., Maugeri, M., Monti, F., and Nanni, T.: Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, *International journal of climatology*, 26, 345–381, 2006.
- Claps, P., Barberis, C., Agostino, M. D., Gallo, E., Laguardia, G., Laio, F., Miotto, F., Plebani, F., Vezzù, G., Viglione, A., and Zanetta, M.: Development of an Information System of the Italian basins for the CUBIST project, in: *EGU General Assembly 2008*, 2008.
- 15 Claps, P., Caporali, E., Chiarello, V., R., D., De Luca, D., Giuzio, L., Libertino, A., Lo Conti, F., Manfreda, S., Noto, V., and Versace, P.: Stima operativa delle piogge estreme sul territorio nazionale: nuovi metodi e possibili sinergie, in: *Atti del XXXV Convegno Nazionale di Idraulica e Costruzioni Idrauliche*, 2016.
- Coumou, D., Robinson, A., and Rahmstorf, S.: Global increase in record-breaking monthly-mean temperatures, *Climatic Change*, 118, 771–782, 2013.
- 20 Desiato, F., Lena, F., and Toreti, A.: SCIA: a system for a better knowledge of the Italian climate, *Bollettino di Geofisica Teorica ed Applicata*, 48, 351–358, 2007.
- Fiori, E., Comellas, A., Molini, L., Rebora, N., Siccardi, F., Gochis, D., Tanelli, S., and Parodi, A.: Analysis and hindcast simulations of an extreme rainfall event in the Mediterranean area: The Genoa 2011 case, *Atmospheric Research*, 138, 13–29, <https://doi.org/10.1016/j.atmosres.2013.10.007>, 2014.
- 25 Furcolo, P., Pelosi, A., and Rossi, F.: Statistical identification of orographic effects in the regional analysis of extreme rainfall, *Hydrological Processes*, pp. 1342–1353, 2015.
- Hosking, J. R. M. and Wallis, J. R.: *Regional frequency analysis: an approach based on L-moments*, Cambridge University Press, 1997.
- Lehmann, J., Coumou, D., and Frieler, K.: Increased record-breaking precipitation events under global warming, *Climatic Change*, 132, 501–515, 2015.
- 30 Libertino, A.: *Advances in the space-time analysis of rainfall extremes*, Ph.D. thesis, Politecnico di Torino, 2017.
- Libertino, A., Allamano, P., Laio, F., and Claps, P.: Regional-scale analysis of extreme precipitation from short and fragmented records, *Advances in Water Resources*, pp. –, <https://doi.org/https://doi.org/10.1016/j.advwatres.2017.12.015>, <https://www.sciencedirect.com/science/article/pii/S0309170817307558>, 2017.
- Pavan, V., Antolini, G., Agrillo, G., Auteri, L., Barbiero, R., Bonati, V., Brunier, F., Cacciamani, C., Cazzuli, O., Cicogna, A., et al.: The
- 35 ARCIS project, *Italian Journal of Agrometeorology*, 2, 2013.
- Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., and Parodi, A.: Extreme rainfall in the mediterranean: what can we learn from observations?, *Journal of Hydrometeorology*, 14, 906–922, 2013.

- Rudari, R., Entekhabi, D., and Roth, G.: Large-scale atmospheric patterns associated with mesoscale features leading to extreme precipitation events in Northwestern Italy, *Advances in Water Resources*, 28, 601–614, 2005.
- Snorrason, A., Finnsdóttir, H. P., and Moss, M. E.: The extremes of the extremes: extraordinary floods, 271, *International Assn of Hydrological Sciences*, 2002.
- 5 Uboldi, F., Sulis, A., Lussana, C., Cislighi, M., and Russo, M.: A spatial bootstrap technique for parameter estimation of rainfall annual maxima distribution, *Hydrology and Earth System Sciences*, 18, 981–995, 2014.
- World Meteorological Organization: *Guide on the Global Data-processing System*, 305, WMO, 1993.
- World Meteorological Organization: *Manual on the Global Data-processing System*, 485,1, WMO, 2010a.
- World Meteorological Organization: *Manual on Global Observing System*, 544,1, WMO, 2010b.
- 10 World Meteorological Organization: *Guide to Meteorological Instruments and Methods of Observation*, 8, WMO, 2014.