Norwegian
Meteorological
Institute

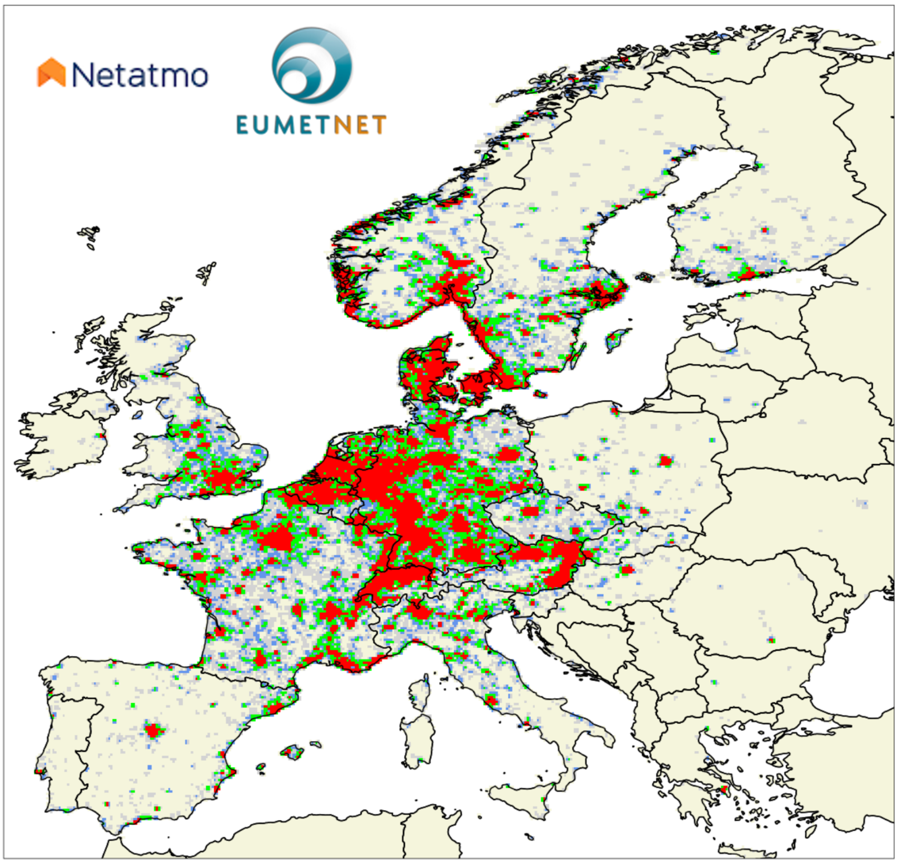
METreport

No. 15/2022
ISSN 2387-4201
Restricted

Quality control for data from personal weather stations - EUMETNET R&D Study A1.05

Cristian Lussana, Frederik De Roo and Thomas N. Nipen

The Norwegian Meteorological Institute, Oslo, Norway



Norwegian
Meteorological
Institute

METreport

Title	**Date**
Quality control for data from personal weather stations - EUMETNET R&D Study A1.05	November 25, 2022
Section	**Report no.**
Division for Climate Services	15/2022
Author(s)	**Classification**
Cristian Lussana, Frederik De Roo and Thomas N. Nipen	◯ Free ● Restricted

Abstract

In the context of EUMETNET tender "R&D Study A1.05, R&D on Quality Control Tools for Observations from Personal Weather Stations", MET Norway has applied several automatic quality tests to hourly precipitation data over Europe. The observational dataset, made available by EUMETNET, consists of measurements from NETATMO (`netatmo.com`) rain gauges over Europe for the year 2020. The research questions addressed in our study are: i) investigate accuracy and precision of crowdsourced data with respect to data provided by traditional weather stations; ii) exploratory use of crowdsourced data for the reconstruction of precipitation fields over Europe, with focus on the impact of quality control procedures. MET Norway has applied spatial consistency tests (SCTs) from the *Titanlib* library of automatic quality control routines (`https://github.com/metno/titanlib`). The results show that SCTs can be effectively applied to the continental-scale observational network and that the reconstructed fields reproduce realistic and spatially coherent patterns of precipitation.

Keywords

hourly precipitation, crowdsourced data, Europe, quality control, spatial statistics

Disciplinary signature	Responsible signature
Hans Olav Hygen	Cecilie Stenersen

Contents

1 Introduction

In the context of EUMETNET R&D study A1.05 (EUMETNET tender "R&D Study A1.05, R&D on Quality Control Tools for Observations from Personal Weather Stations"), the Norwegian meteorological institute (MET Norway) has applied several automatic quality control tests to hourly precipitation data over Europe. The observational dataset, made available by EUMETNET, consists of measurements from personal weather stations, which in this document we will refer to as crowdsourced data. Specifically, we have used measurements from NETATMO (`netatmo.com`) weather stations observing precipitation.

The research questions addressed in our study are: i) investigate accuracy and precision of crowdsourced data with respect to data provided by traditional weather stations; ii) exploratory use of crowdsourced data for the reconstruction of precipitation fields over Europe, with special focus on automatic quality control.

We have applied quality control procedures developed by MET Norway during the Titan (*Båserud et al.*, 2020) project for the quality control of Scandinavian data. The automatic quality control checks have been further developed in a library named Titanlib (C++ / R / Python) (*Abraham et al.*, 2022).

The document is organized as follows: Section 2 describes the data used and presents a comparison of NETATMO observations against traditional observations. Section 3 describes the quality control procedures. In addition, we present a general strategy for the optimization of quality control routines given a specific application. Section 4 presents the results of the optimization of the quality control routines together with the reconstruction of precipitation fields, for four case studies, over Europe. All figures are reported after the Conclusions. Besides, the document includes three Appendices. In the first two, we present the configurations used to run the quality control routines. In the last appendix, we list the websites where the Titanlib code can be downloaded.

2 Data

We have used crowdsourced observations of hourly precipitation measured by NETATMO private weather stations (`netatmo.com`). The dataset is the "eumetnet-netatmo" (EUMETNET Sandbox Netatmo Network Data) provided by EUMETNET to the partners of "EUMETNET Obs CA Study A1.05". The dataset covers the year 2020. In addition to that, we have completed the EUMETNET sandbox with NETATMO observations that were available at MET Norway, this allowed us to slightly increase the number of observations and to consider data from September 2019 to October 2022 (instead of the year 2020 only). The data flow is the following: the original data received from EUMETNET or NETATMO are converted into 10-min precipitation amounts and stored into NetCDF files; then, we have accumulated the 10-min precipitation into hourly totals.

The dataset covers a large part of Europe. We have focused our attention on four case studies, which will be described in Sec. 4.0.3. The spatial distribution of the observational network is shown in Fig. 1. The station density is shown as the average number of stations close to each of the points of a regular grid covering the domain. The average is computed considering all samples (i.e. hours) in 2020. Panel *a* shows the average number of crowdsourced observations within a circular region with radius of 1 km from each point. The other panels are similar to panel *a*, but they display averages when considering larger circular regions (i.e. 3 km, 5 km and 10 km). Figure 1 shows that the network is denser in Central Europe than in other regions. However, in the main cities, all over Europe, the number of stations rises rapidly.

We have studied the precision and accuracy of crowdsourced hourly precipitation data and the results are shown in Figs. 2- 4. The crowdsourced data has been evaluated against reference data, such as the hourly precipitation measured by traditional weather stations managed by national weather services: MET Norway; the Swedish Meteorological and Hydrological Institute (SMHI); the Finnish Meteorological Institute (FMI). The hourly observations of precipitation used (both NETATMO and traditional data) have been quality controlled with automatic procedures. Furthermore, we have considered only precipitation measurements when temperatures observed at reference stations were greater than 2°C because crowdsourced data are measured by non-heated rain gauges.

The closer the crowdsourced stations are to a reference station, the better the agreement of their observations should be. On the other hand, the bigger the region considered around each reference observation, the larger the number of crowdsourced stations we can use to obtain more robust and resistant statistics. Figs. 2, Figs. 3 and Figs. 4 show the results when we consider radii of 1km, 3km and 5 km, respectively.

Given a specific observation time, the comparison has been done in two different ways: a) we compare each reference observation with its crowdsourced nearest neighbour (green lines in panels *b*); b) alternatively, for each reference observation, we extract several percentiles (i.e. the 10-th, the 25th, the median, the 75-th and the 90-th) from the empirical distribution of crowdsourced data belonging to the circular region chosen and we compare them with reference observation (black lines and grey regions in panels *b*). The data are collected for the hours in the period of time from 2019-09-01 to 2022-10-01. Then, the collected samples (i.e. the pairs (reference observation, nearest neighbour crowdsourced observation); the n-tuple (reference observation, 10-th percentile,...,90-th percentile)) are divided, on the basis of the reference precipitation amount, into several precipitation classes (e.g. hourly precipitation between 0 mm and 0.1 mm, 0.1 and 0.2, and so on). Finally, for each parameter (i.e. crowdsourced nearest neighbour and the percentiles), the statistics considered -and shown in the figures- are the medians of the collected samples in the precipitation classes. In the figures, we display also the statistics we obtain when the roles of reference and crowdsourcing data are inverted (blue lines).

The *a* panels in Figs. 2- 4 show the spatial distributions of the stations used and the insets on the top left show the number of observations belonging to each of the precipitation classes considered (note that the y-axis scale is logarithmic). The grey dots in the insects show the number of samples in each precipitation class, when the subdivision is based on the reference precipitation. The blue dots show the number of samples belonging to each precipitation class but on the basis of the crowdsourced precipitation. The *b* panels show the statistics of the crowdsourced data conditional to the reference precipitation classes. As stated above, the blue line shows the opposite, that is the statistics of the reference precipitation conditional to the crowdsourced precipitation classes.

The results show a rather good agreement between reference observations and crowdsourced data, especially for precipitation values between 0 mm/h and 4-6 mm/h. As expected, the agreement is better when we consider smaller circular regions of 1 km around the reference stations (Fig. 2). However, even when larger regions of 5 km are considered (Fig. 4), we still have a reasonable agreement. For instance, for the 5 km radius, the linear regression between the (median of) crowdsourced and reference observations is such that the reference precipitation equals 0.71 times the reference precipitation.

The deviations between the two conditional probabilities (i.e. the black and the blue lines) confirms that the crowdsourced data underestimates the reference precipitation, particularly for intense precipitation.

3 Methods

Atmospheric observations are affected by gross measurement errors (GEs) when their values are independent of the atmospheric conditions (e.g. a rain gauge placed under a roof). In the following, for the sake of brevity, observations affected by GEs are referred to as "bad" observations, while the others are "good" observations.

Titanlib is a library of functions implementing data quality control checks aiming at identifying bad observations within datasets of near-surface atmospheric variables. In this study, we focus on hourly precipitation as measured by a network of crowdsourced observations over Europe. The main idea behind the Titanlib checks is to exploit the expected spatial consistency of the atmospheric variables within a single observation time (i.e. we do not consider time series, only "time slices", which are assumed to be independent of each other). In fact, sudden and abrupt changes measured by a single observation only (i.e. outliers), and not by its neighbouring observations, are usually associated with GEs.

The type of quality checks considered are the so-called spatial consistency tests (SCTs *Lussana et al.*, 2010). SCTs can mistake very local weather conditions with GEs. In the case of precipitation, a rain gauge placed too close to large obstacles, such as buildings or trees, would measure values strongly influenced by local turbulence. Let us assume that its neighbouring rain gauges have not all been placed too close to obstacles, then when raining we would notice a certain variability in their measurements. In particular, instruments that have been placed in sites that have a significantly different exposure compared to their neighbours, will often report outlier observations. In this case, we will say that those rain gauges are affected by large representativeness errors (REs), which for most practical applications can be identified with GEs and adequately reported. Therefore, a GE can also originate from an observation that is measuring a very local atmospheric phenomenon, which is characterized by spatial scales not adequately resolved (locally) by the observational network.

The descriptions of the SCTs applied are presented in Sec. 3.1. The subsequent application of the SCTs to the selected case studies is straightforward, because the information required to run the procedures are available in the Eumetnet sandbox. However, we need to specify the values of the SCT parameters. An important part of this study is the description, in Sec. 3.2, of a strategy that can be used to optimize the SCT parameters.

3.1 Spatial consistency tests (SCTs)

Two different types of spatial tests have been used: SCT dual and SCT resistant.

SCT dual can be used to identify observations that are measuring no-precipitation while most of the neighbouring observations report precipitation or, vice-versa, when observations measuring precipitation are surrounded by a majority of no-precipitation measurements. The Titanlib routine used is *sct_dual()* and its description can be found in the Titanlib-wiki `https://github.com/metno/titanlib/blob/master/src/sct_dual.cpp`.

SCT resistant is used to identify GEs when precipitation observations deviate significantly from the expected values estimated from neighbouring observations. Deviations can be either an underestimate or an overestimate of the expected values. The Titanlib routine used is *sct_resistant()* and its description and algorithm can be found in the wiki: `https://github.com/metno/titanlib/wiki/Spatial-consistency-test-resistant` . Furthermore, SCT resistant is also described in detail in the MET report by *Lussana and Båserud* (2021a).

In this study, we compare the two modes of operation of SCT resistant: 1) basic; 2) advanced. Both methods are described at the same web page reported above. The basic mode implements SCT as described by *Lussana et al.* (2010). The advanced mode is designed to recognize spatial patterns in the observation representativeness error (*Lussana and Båserud*, 2021a). In fact, when the SCT identifies clear spatial patterns of bad observations, it might be that REs have been mistaken as GEs. Therefore, the advanced mode should flag less bad observations then the basic mode, because it should better distinguish between GEs and REs. The advanced method has been developed -and tested- mostly for temperature and for that variable it shows some benefits (*Lussana and Båserud*, 2021b,c). In this study, we are testing it on precipitation.

3.2 Strategy for the optimization of SCT resistant (Titan tuner)

The aim is to choose the best configuration for SCT resistant and apply it to a selection of case studies.

The best configuration of a generic check is identified through the minimization of a cost function, which is dependent on: a) the type and magnitude of the GEs under consideration; b) the prior knowledge we have on both the observational network and the specific application considered; c) empirical data collected through tailored experiments. The procedure applied has been described by *Alerskans et al.* (2022) for temperature over Denmark. In this study, it is applied to hourly precipitation over Europe.

The first step is the definition of a specific type of GE (e.g. rain gauges overestimate precipitation of a fixed amount). Then, several experiments are conducted, each of them designed such that slightly different SCT resistant configurations are used. For each experiment, a contingency table is obtained as explained below. For each observation, the outcome of the check is a binary event of the type: "yes, the observation is affected by GE" (bad observation) or "no, the observation is not affected by GE" (good observation). All observations are tested simultaneously, however the contingency table is based only on the outcomes obtained on a selected subset of observations, which are chosen among all the others. This subset of observations constitutes the reference dataset (not to be confused with the reference dataset defined in Sec. 2). A single consistency table is the result of two independent runs of the check: an "unperturbed" run, where the reference data are assumed to be all good observations, is used to set the false alarms (i.e. reference good is flagged as bad) and the correct negatives (i.e. reference good is flagged as good); and a "perturbed" run, where the reference data are assumed to be all bad observations, because their values has been modified to simulate the effect of the GE under study. The perturbed run is used to set the hits (i.e. reference bad is flagged as bad) and misses (i.e. reference bad is flagged as good).

The choice of the best configuration relies on the definition of a cost function and its subsequent minimization. For a single experiment, the cost is defined as:

$$\mathrm{Cost} = 1 - \mathrm{H} + \mathrm{FAR} \cdot m \tag{1}$$

Where: H is the hit rate; FAR is the false alarm rate; m is a penalty factor (*Alerskans et al.*, 2022).

The selection of the reference dataset is an important step in the optimization procedure and it depends on the specific type of GE under study. First, we need to decide whether an observation is suitable as a candidate reference observation. It is worth remarking that: i) we assume that all the reference observations are good observations; ii) the reference observations must be far apart from each other because we do not want the perturbation over one observation to influence the checks on the others. The first assumption is that the reference observations must be good observations. For the selection of the SCT resistant reference dataset, we consider observations greater than 2 mm and specifically, those that have: more than 4 nearby neighbours within a circular region with radius of 10 km and they must have measured a value that is confirmed by the neighbours (i.e. within plus or minus one standard deviation from the mean of the neighbours). The 2 mm threshold is introduced because we are more interested in detecting GEs when some rain has been registered, such that the spatial variability of precipitation is significant enough to make it challenging for SCT resistant. The second assumption is satisfied by imposing that the selected observations must be more than 50 km apart from each other. At this point, the selected observations constitute the reference dataset.

Note that it is possible to select more than one random reference subset of observations that satisfy the distance-based requirement. As a consequence, for a single observation time, it is possible to build up the contingency table based on more than one reference dataset, for instance by implementing a bootstrapping procedure, such that the results are less dependent on the random choice of the reference dataset.

One last important recommendation. The NETATMO-based observational network of crowd-sourced data is extremely unevenly distributed in space (see Sec. 2). As a consequence, the optimization of the parameters will better represent urban areas, where the network is denser, than rural areas. For this reason, it is always a good idea to refine the tuning of the parameters with e.g. Titantuner (see Sec. 3.1).

4 Results

4.0.1 SCT dual parameter optimization

The procedure used to optimize the SCT dual parameters is one of “trial-and-errors". A useful tool that we have used to evaluate our choices on the parameters and their impacts on the outcomes is TitanTuner `https://github.com/metno/titantuner`.

The parameters chosen are reported in Appendix A.

4.0.2 SCT resistant parameter optimization

The GE types considered are:

- observation measures 50% less than what expected
- observation measures 90% less than what expected
- observation measures 100% more than what expected
- observation measures 500% more than what expected

The results of the optimization experiments (see Sec. 3.2) are shown in Figs. 5- 8. The figures show the Receiver Operating Characteristic (ROC) curves for the different GE types. Each figure includes two ROC curves, one for SCT resistant basic (blue) and one for advanced (red). In all cases, SCT resistant basic performs better than the advanced test. This implies that work still needs to be done to adapt SCT resistant advanced for precipitation. In Sec. 4.0.3, the results shown have been obtained using SCT resistant basic.

In general, better results (i.e. higher hit-rates and lower false-alarm-rates) are achieved in detecting GEs of the type “observation measures more than what expected”.

On the basis of the interpretation of Figs. 5- 8, we have tried Titantuner with the best values of the threshold, which have been reported in the captions of the figures. The final configuration chosen is reported in Appendix B, where the threshold is set to 1 (same as in Fig. 8). Note that for this particular study we did not have a specific application in mind, other than the figures presented in Sec. 4.0.3 may have to look realistic. Other choices are possible for different applications.

In Figures 9- 10, two examples of the outcomes of SCT resistant (configuration as in Sec. B) are presented for the case study Storm Alex (see Sec. 4.0.3). Fig. 9 shows the bad observations detected over the Alps, while Fig. 10 shows the bad observations over South-East England. In both cases, the SCT detects several stations and a typical type of observation error is those of "rain gauges stuck to 0 mm" while the neighbours measure precipitation.

4.0.3 Case Studies

The four case studies considered are:

- Storms Ciara and Denis, from 2020-02-09 to 2020-02-10. Figs. 11- 12
- Convective precipitation over Central Europe, from 2020-06-13 to 2020-06-15. Figs. 13- 14
- Convective precipitation over the United Kingdom and western Europe. from 2020-08-11 to 2020-08-12. Figs. 15- 16
- Storm Alex. from 2020-10-02 to 2020-10-04. Figs. 17- 18

For each case study, we present: i) a figure with the reconstruction of the precipitation field after the application of SCT resistant and ii) a figure on the differences in reconstructing precipitation with and without SCTs. The precipitation fields are reconstructed onto a regular grid with resolution of 0.1 x 0.1 degrees. For each grid cell including at least one crowdsourced observation, the reconstructed value is set to the average of the crowdsourced observations.

The first case study reconstructs the passage over Europe of Storms Ciara and Denis during February 2020. The precipitation field is shown in Fig. 11. Precipitation has been observed on the British isles, Central and Northern Europe. The most intense precipitation has been registered in Western Norway. The effects of the SCTs are shown in Fig. 12 and mostly it reduces precipitation amounts.

The second case study is a case of convective summer precipitation over Central Europe. The most intense precipitation has been measured over Austria and Germany, as shown in Fig. 13. The convective pattern shows scattered thunderstorms all over Germany and they have been well observed by the dense network of observational data. The impacts of SCT dual and resistant are shown in Fig. 14.

The third case study is a case of convective precipitation over the United Kingdom and western Europe that happened in August 2020. The observed precipitation mostly covers Great Britain, Western and Central Europe. For instance, the precipitation maxima over Madrid is clearly visible in the map of Fig. 15. The impacts of the SCTs are shown in Fig. 16. Among all figures showing the SCT impacts, this is the one that shows the least impact, possibly because there is a large portion of the domain with no precipitation.

The fourth case study reconstructs the passage over Europe of Storm Alex (Fig. 17). The higher precipitation amounts have been measured over the Alps, the southern part of Great Britain and over France. As shown in Fig. 18, in this case the impact of the SCTs is also limited. However, SCT resistant is having a bigger impact than SCT dual.

In Table 1, the average frequency of bad observation flagged by the two SCTs are reported. SCT dual is flagging more observations than SCT resistant, however from Figs. 12- 14- 16- 18 it is clear that SCT resistant is having more impact on the final reconstructed fields. It is also worth remarking that the frequency of rejection is limited to a few percent of the total set of observations.

Case study	f SCT dual (%)	f SCT resistant (%)
Storms Ciara and Denis	7.3	3
Convective prec. over Central Eu.	2.9	1.5
Convective prec. over the U.K. and western Eu.	1.6	1.2
Storm Alex	4.7	2.3

Table 1: Average frequency of observations flagged as “bad” (f, units %) in the four case studies for SCT dual and SCT resistant.

5 Conclusions

Hourly precipitation is characterized by large spatial variability and the instrumentation used by NETATMO is rather different from those of the traditional weather stations, often used as reference stations. Despite all of this, the crowdsourced data constitutes an accurate and precise source of precipitation data. However, it must be considered that accuracy and precision of the analyzed data decrease as precipitation increases.

The observational network of crowdsourced data allows for the reconstruction of precipitation fields that are spatially coherent and seems to represent spatial patterns of precipitation from the mesoscale down to city neighbourhoods, where enough observations are available. Quality control procedures, such as the SCTs presented here, are needed in order to remove outliers and observations affected by gross-measurements errors. It is worth remarking that the Titanlib SCT functions used were able to quality control the whole observational dataset in a few minutes, running on a common desktop computer.

In our experiments we have removed only small percentages of the available data. The quality control procedures can be optimized with respect to specific applications, and in this document an optimization strategy has been described and proposed as a general guide.

6 Figures

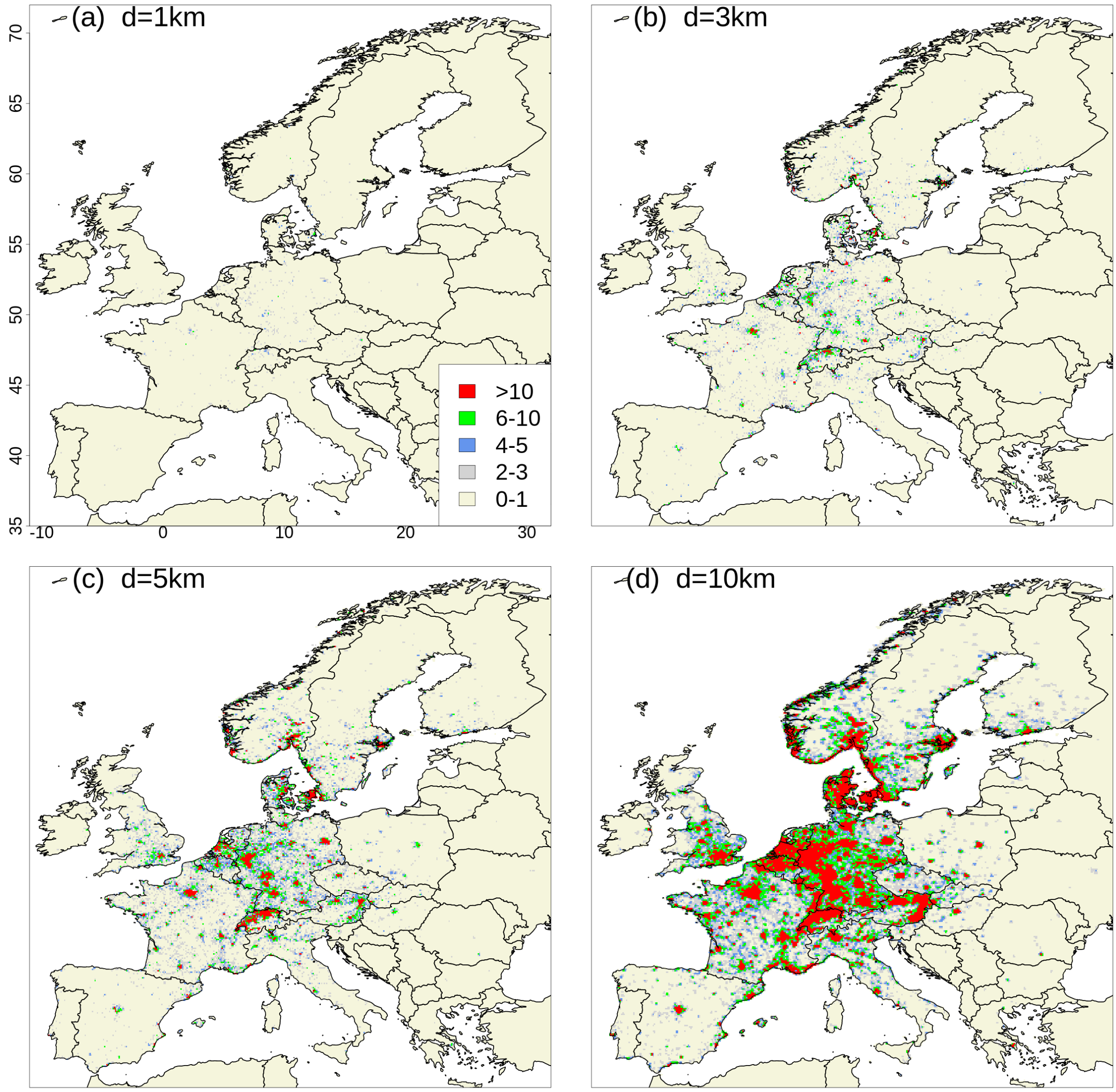


Figure 1: Average spatial density of crowdsourced observations over Europe in the year 2020. The panels show the average number of observations in a circular region of radius d around each point.

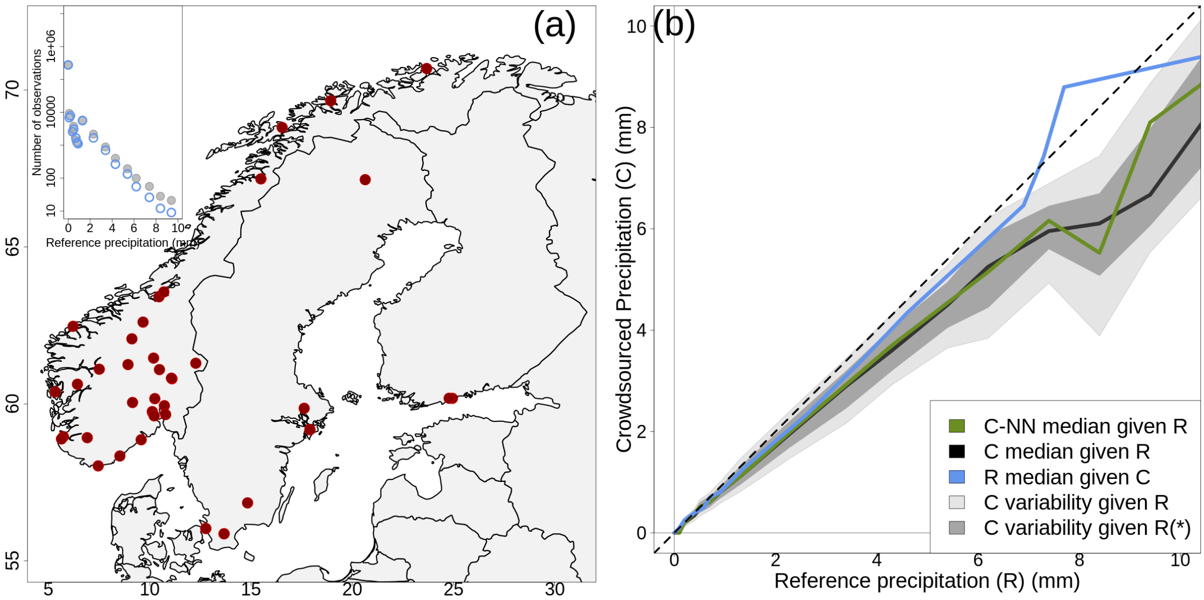


Figure 2: Accuracy and precision of crowdsourced hourly precipitation totals with respect to reference observations. The time period considered is from 2019-09-01 to 2022-10-01. The crowdsourced observations used lie within a circular region of radius equal to 1 km from the reference stations. We have used only those cases when at least 3 crowdsourced observations were simultaneously available. Panel *a* shows the location of the 42 reference stations and the inset on the top right shows the number of observations as a function of the reference precipitation class (R). Panel *b* shows the crowdsourced precipitation (C) as a function of R. For each R class, we have considered several empirical distribution of crowdsourced values and we have shown: the median of the nearest observation (green); the median of all observations within the circular region (black); the expected variability within the circular region is shown as the range of values between the (median of the) 10-th and the (median of the) 90-th percentiles (light grey region); a second estimate of the variability is shown considering the interquartile range of values (i.e. between the 25-th and the 75-th percentile, grey region). The blue line shows the median of the distribution of the reference values as a function of the classes of C.

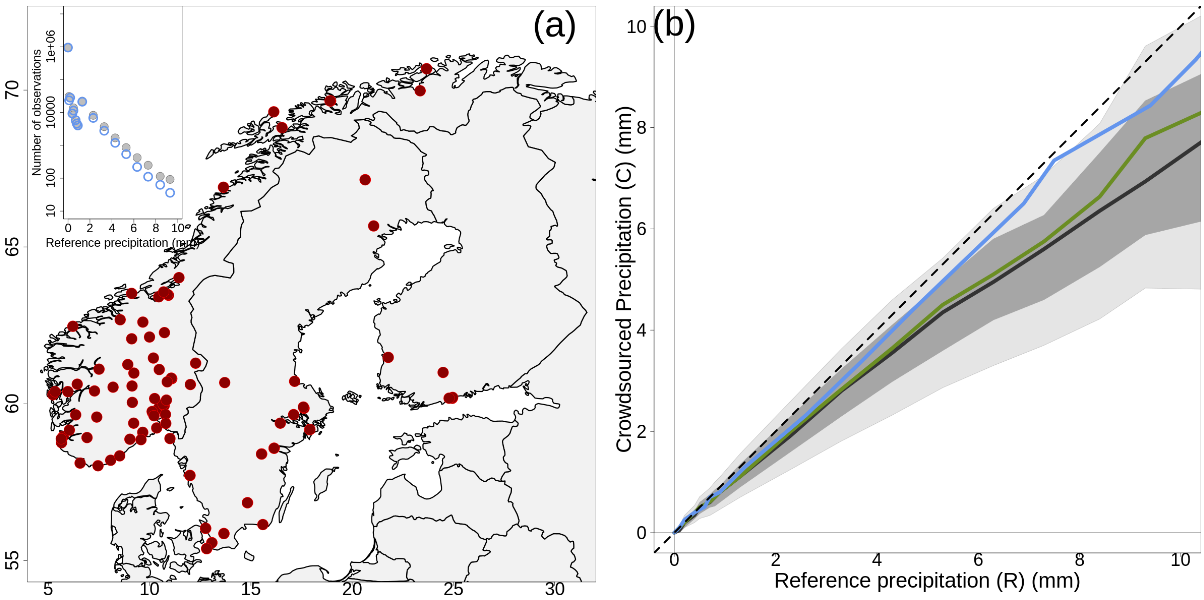


Figure 3: Accuracy and precision of crowdsourced hourly precipitation totals with respect to reference observations (see the caption of Fig. 2) when the radius of the circular region is 3 km and the required minimum number of stations is set to 5. The number of stations is 98.

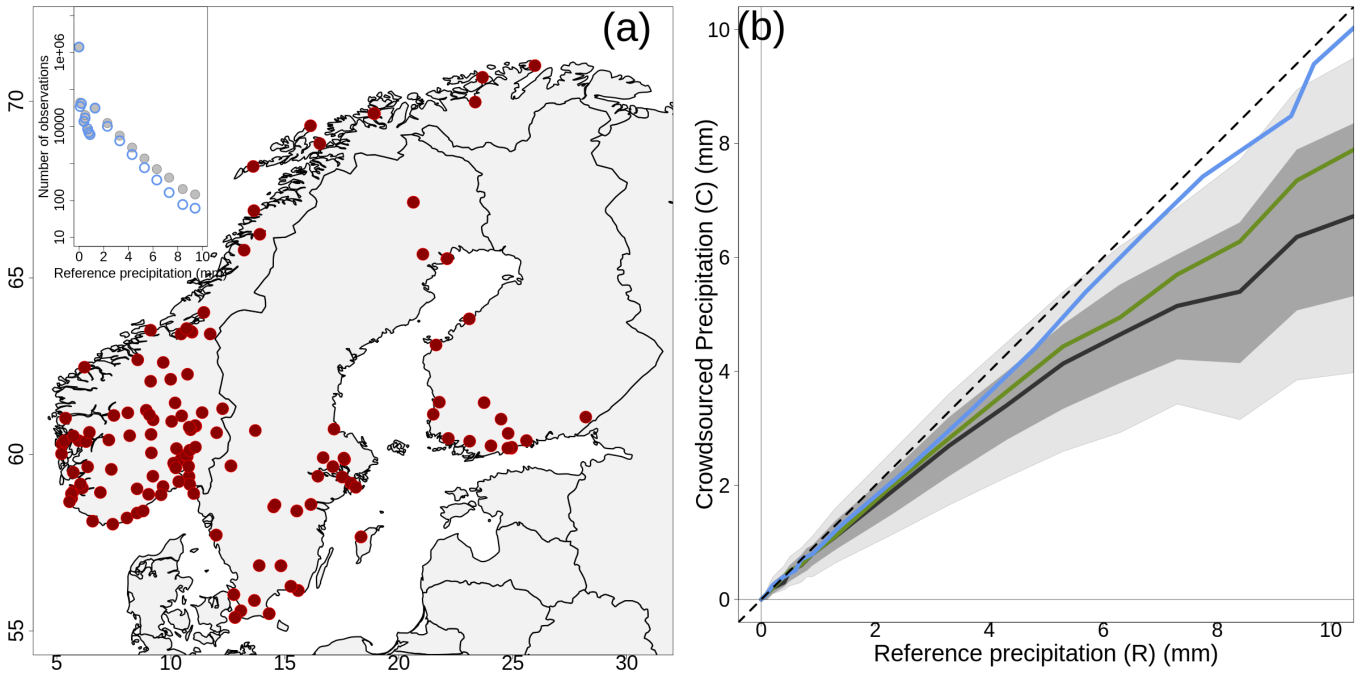


Figure 4: Accuracy and precision of crowdsourced hourly precipitation totals with respect to reference observations (see the caption of Fig. 2) when the radius of the circular region is 5 km and the required minimum number of stations is set to 5. The number of stations is 146.

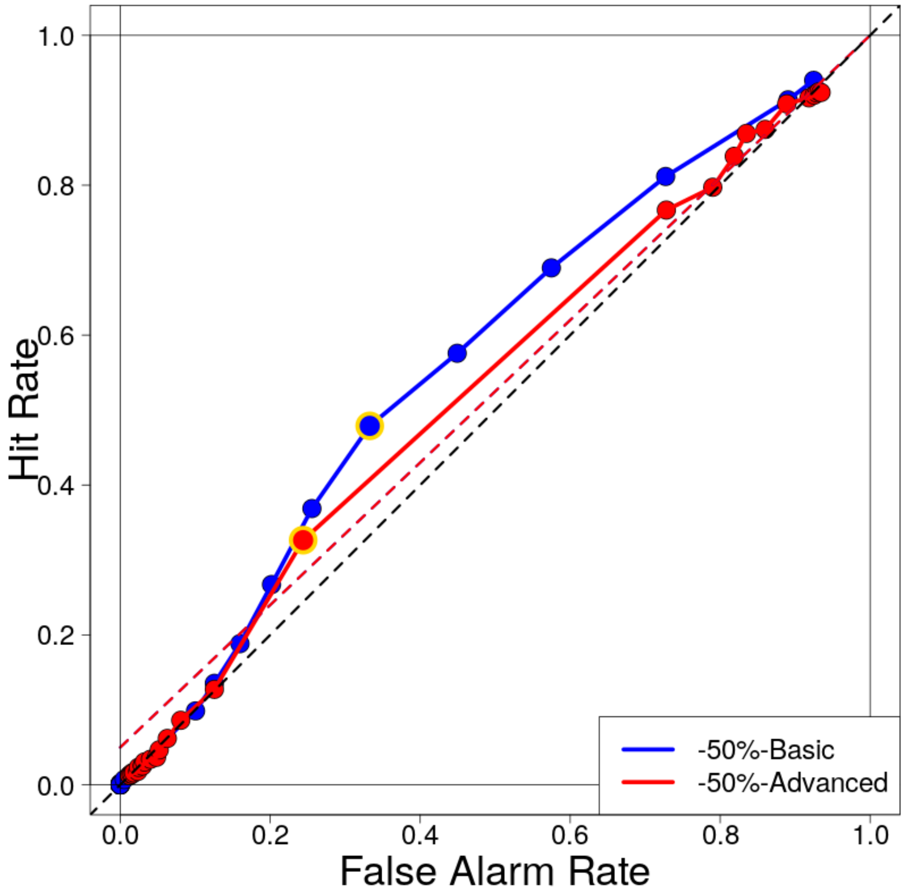


Figure 5: ROC curves for SCT resistant basic (blue) and advanced (red) for different thresholds. The type of GE considered is "observation measures 50% less than what expected". The best results (i.e. minimum of the cost function, points marked with a gold circle) have been obtained when: i) basic, threshold is 0.5; ii) advanced, threshold is 2.

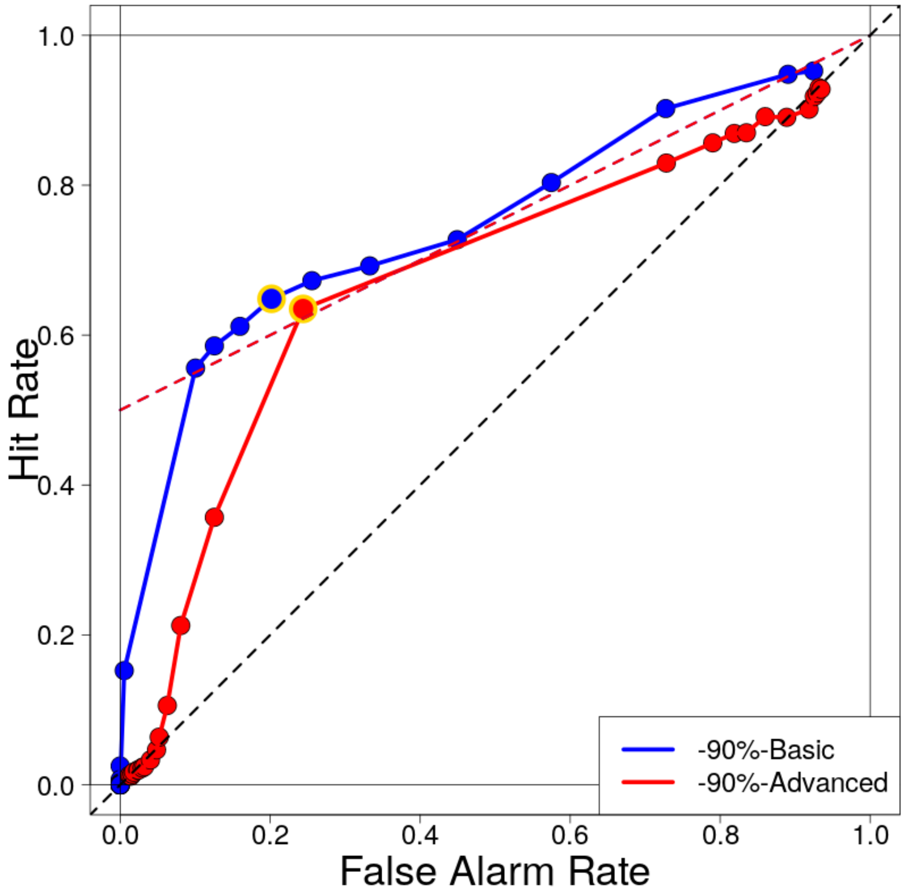


Figure 6: ROC curves, same layout as Fig. 5. In this case, the GE type is “observation measures 90% less than what expected”. The best results are obtained for: i) basic, threshold is 0.7; ii) advanced, threshold is 2.

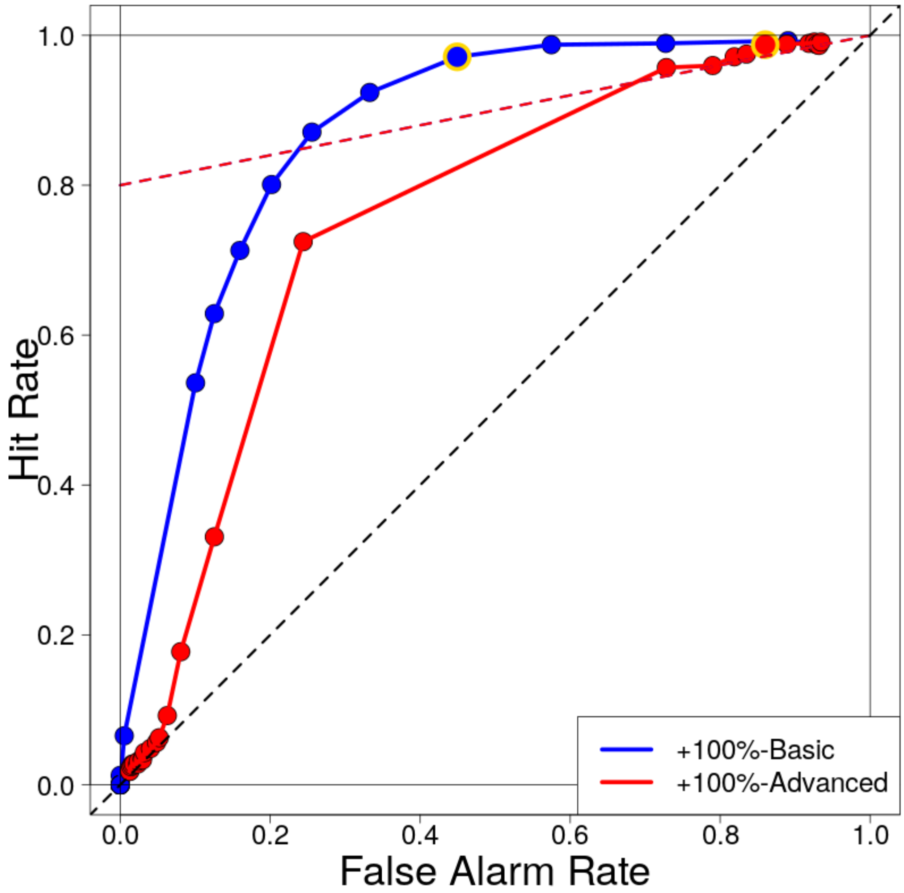


Figure 7: ROC curves, same layout as Fig. 5. In this case, the GE type is "observation measures 100% more than what expected". The best results are obtained for: i) basic, threshold is 0.4; ii) advanced, threshold is 0.6.

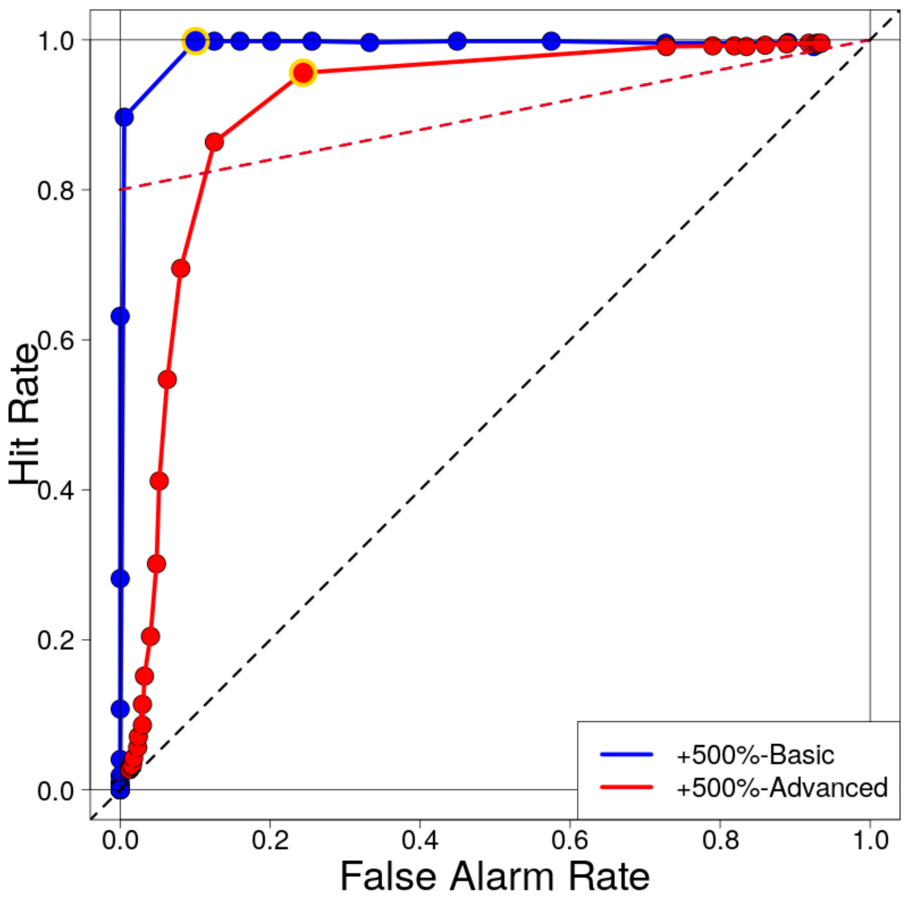


Figure 8: ROC curves, same layout as Fig. 5. In this case, the GE type is “observation measures 500% more than what expected”. The best results are obtained for: i) basic, threshold is 1; ii) advanced, threshold is 2.

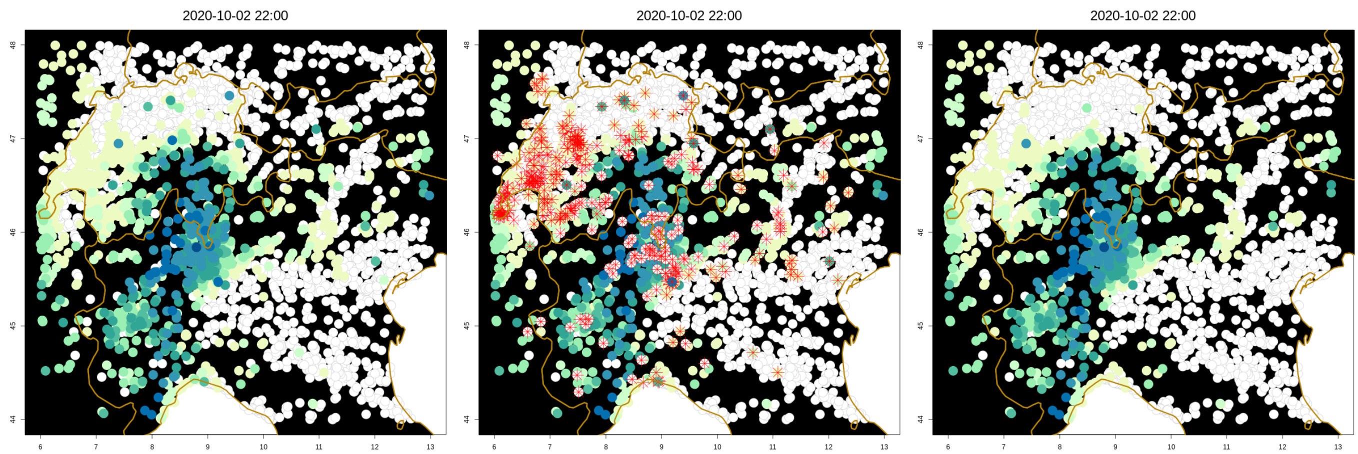


Figure 9: Example of SCT application over the Alps for the Storm Alex case study. On the left, the whole dataset of crowdsourced observations is shown (no quality control). In the middle, observations rejected by the SCTs are marked with red symbols. On the right, quality controlled observations only.

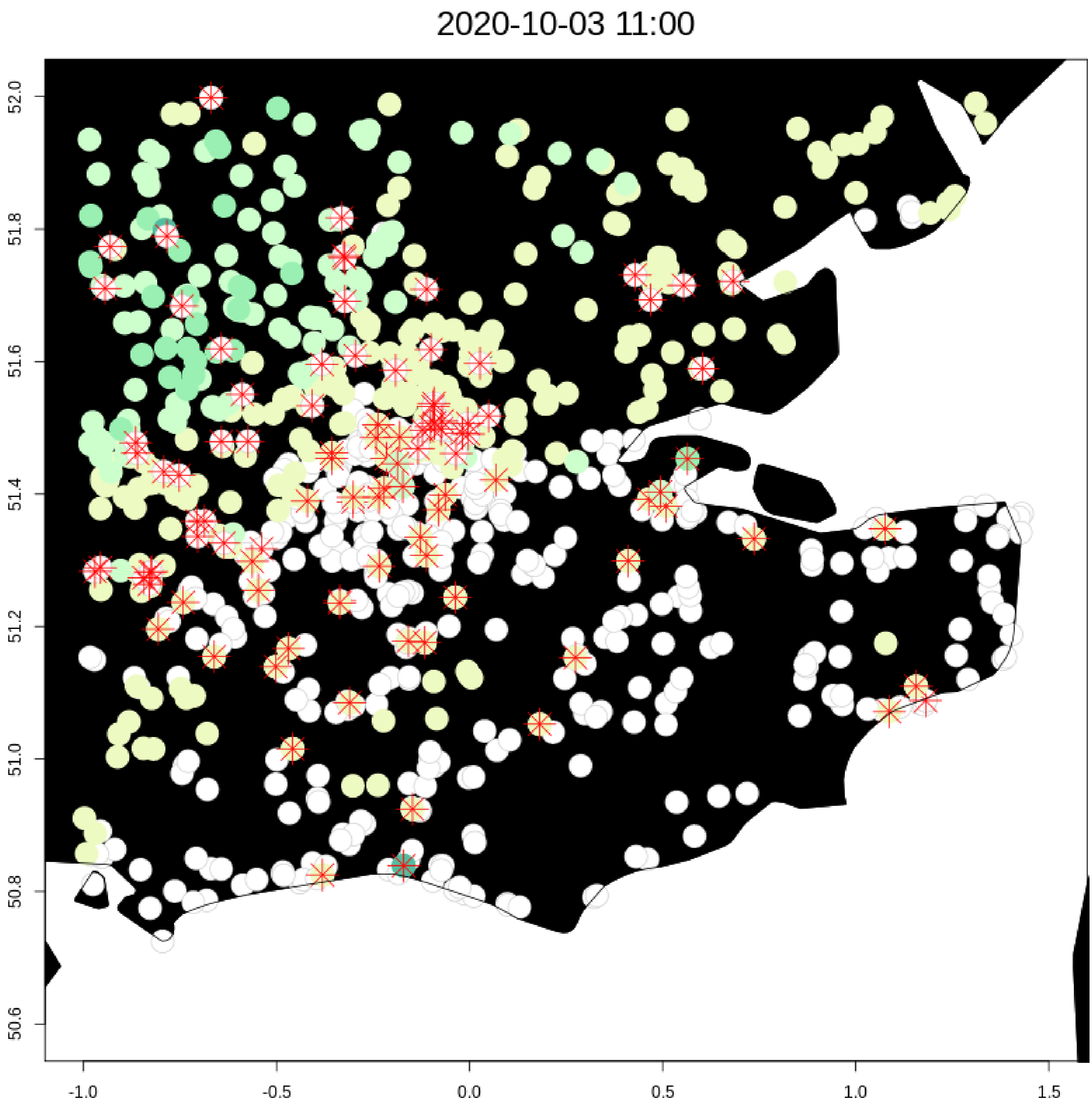


Figure 10: Example of SCT application over the London area and South-East England for the Storm Alex case study. The observations rejected by the SCTs are marked with red symbols.

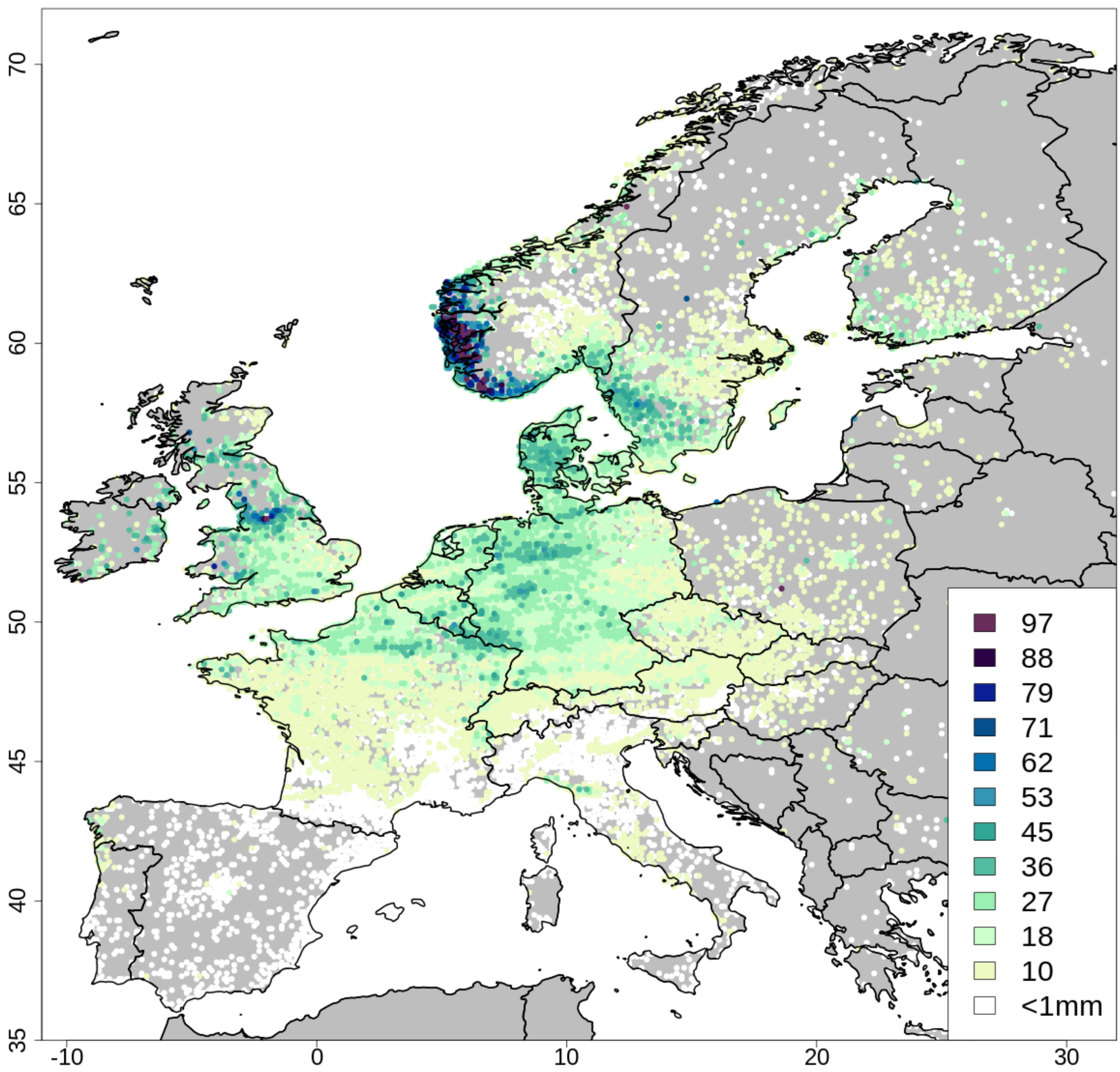


Figure 11: Storms Ciara and Dennis. Precipitation totals from 2020-02-09 00 UTC to 2020-02-11 00 UTC. The precipitation is accumulated using observations after SCT resistant has been applied.

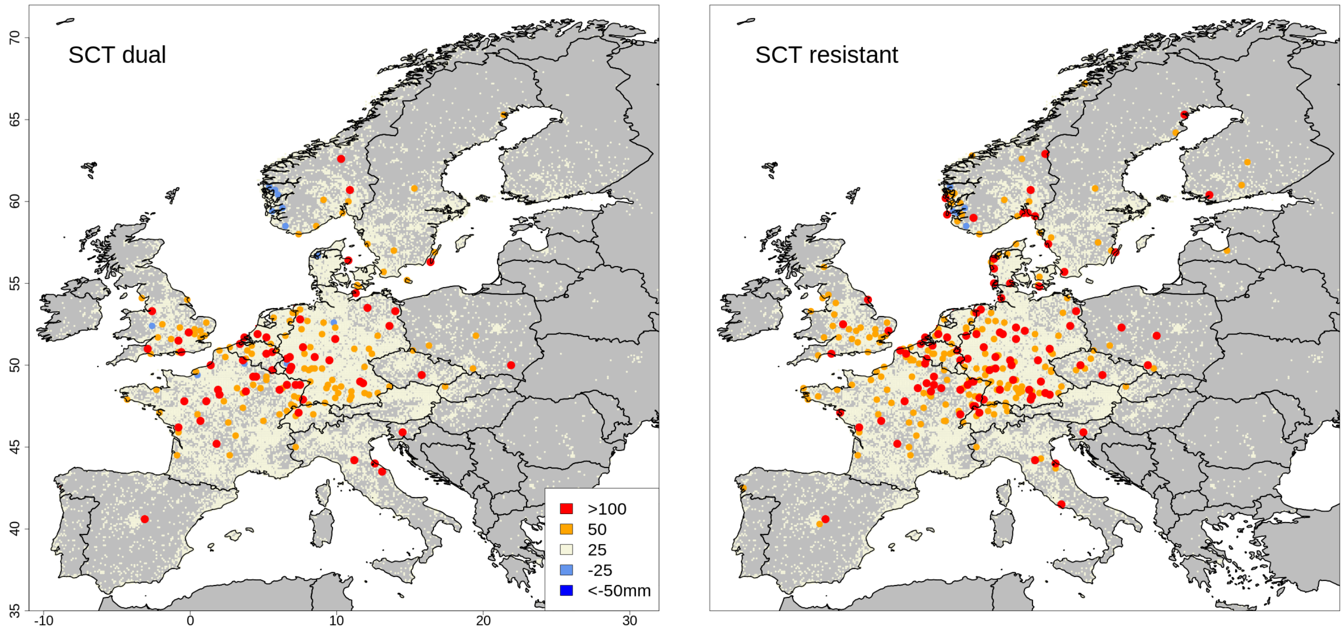


Figure 12: Storms Ciara and Dennis. Difference between total observed precipitation without and with the quality checks (i.e. raw totals minus quality controlled totals). The time period considered is the same as in Fig.11. In the left panel, SCT dual has been applied. In the right panel, SCT resistant has been applied.

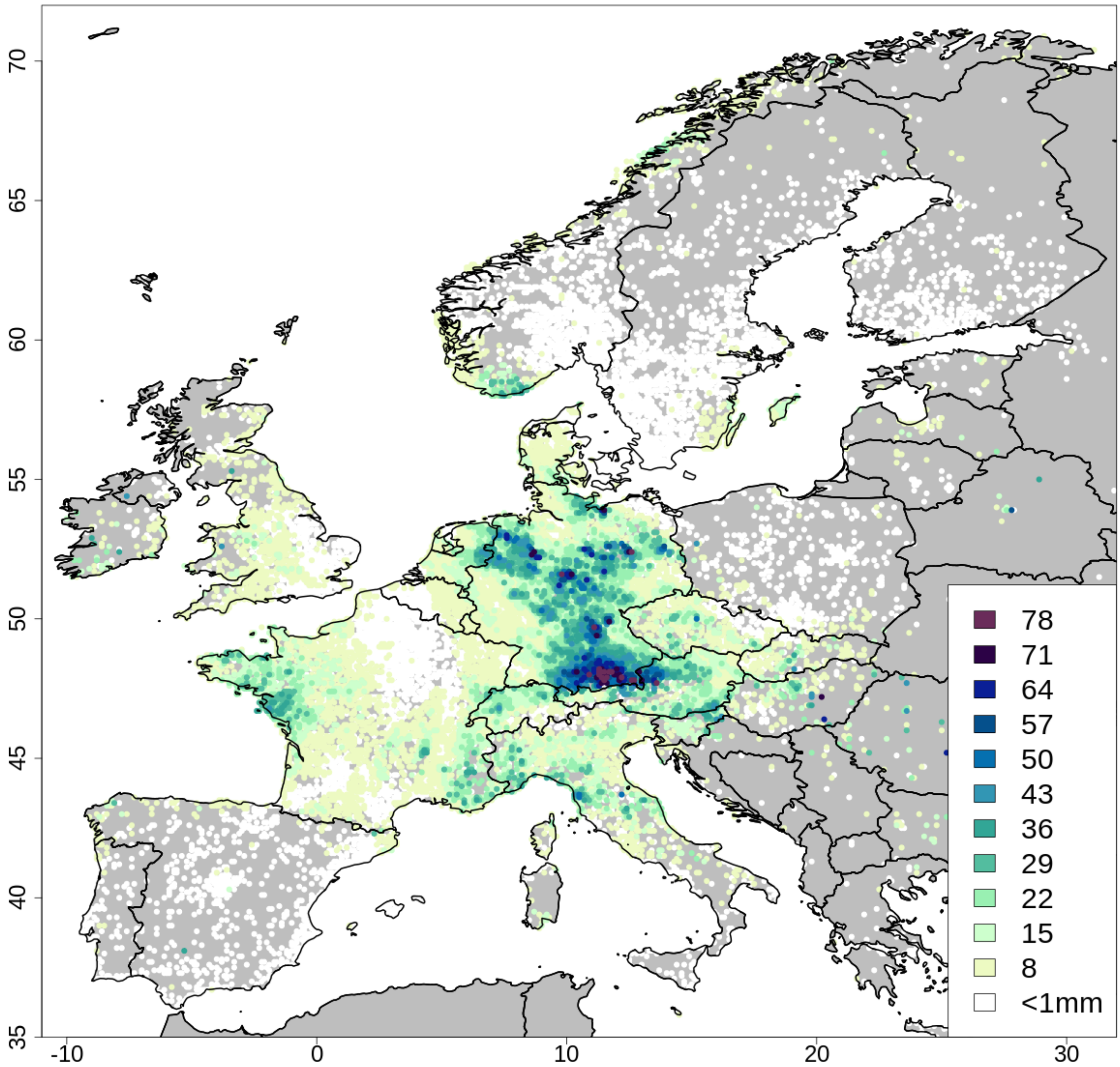


Figure 13: Convective precipitation over Central Europe. Precipitation totals from 2020-06-13 00 UTC to 2020-06-16 00 UTC. The precipitation is accumulated using observations after SCT resistant has been applied.

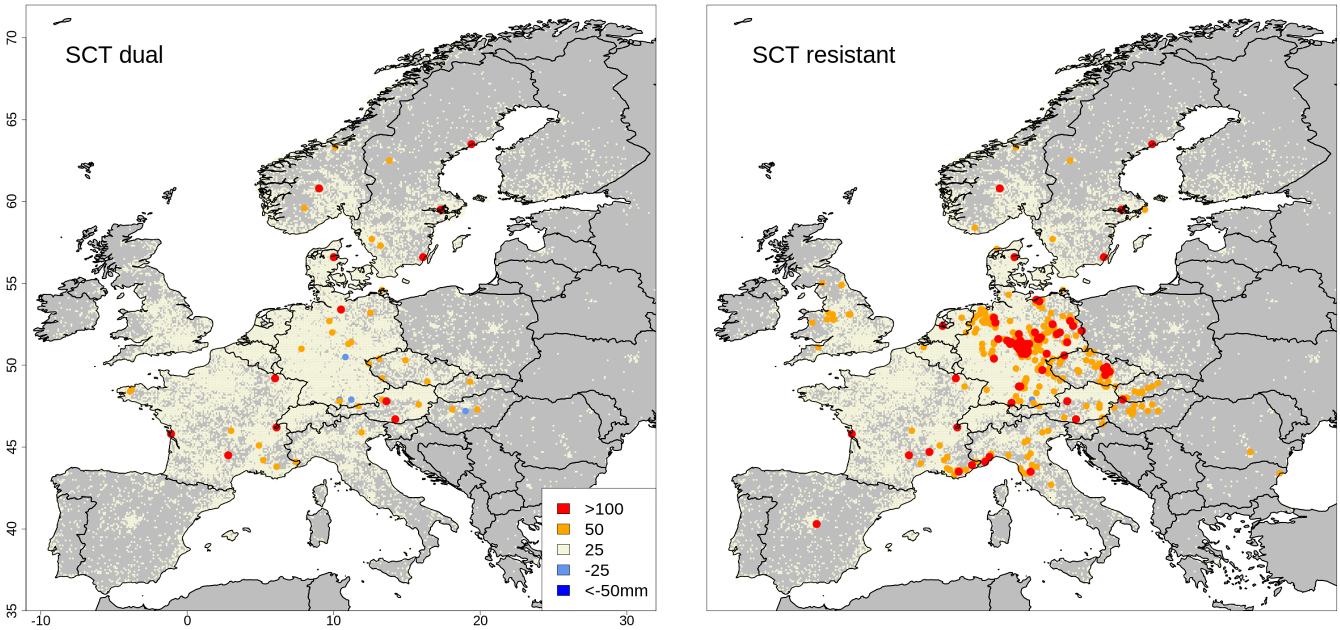


Figure 14: Convective precipitation over central Europe. Difference between total observed precipitation without and with the quality checks (i.e. raw totals minus quality controlled totals). The time period considered is the same as in Fig.13. In the left panel, SCT dual has been applied. In the right panel, SCT resistant has been applied.

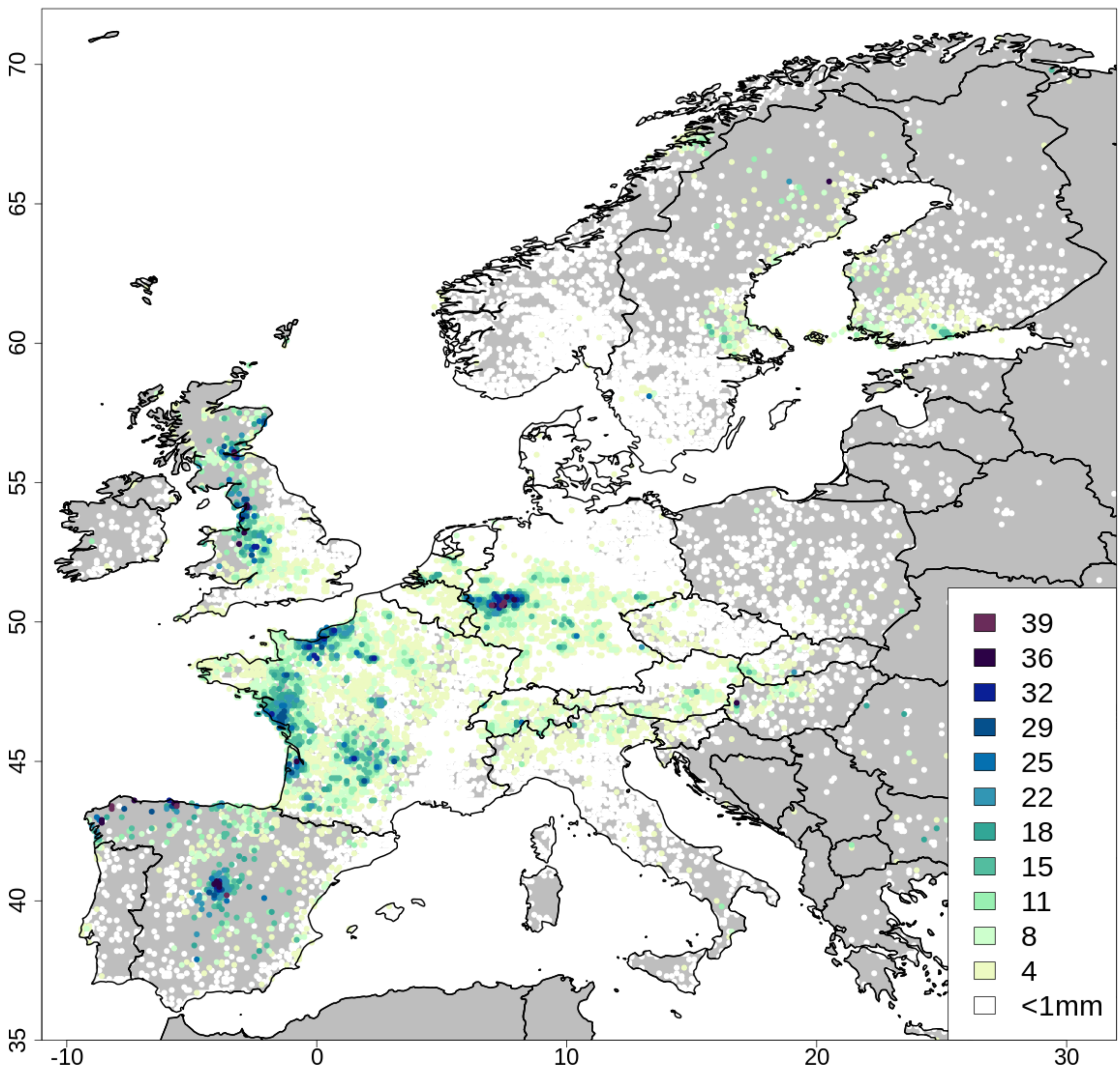


Figure 15: Convective precipitation over the United Kingdom and western Europe. Precipitation totals from 2020-08-11 00 UTC to 2020-08-13 00 UTC. The precipitation is accumulated using observations after SCT resistant has been applied.

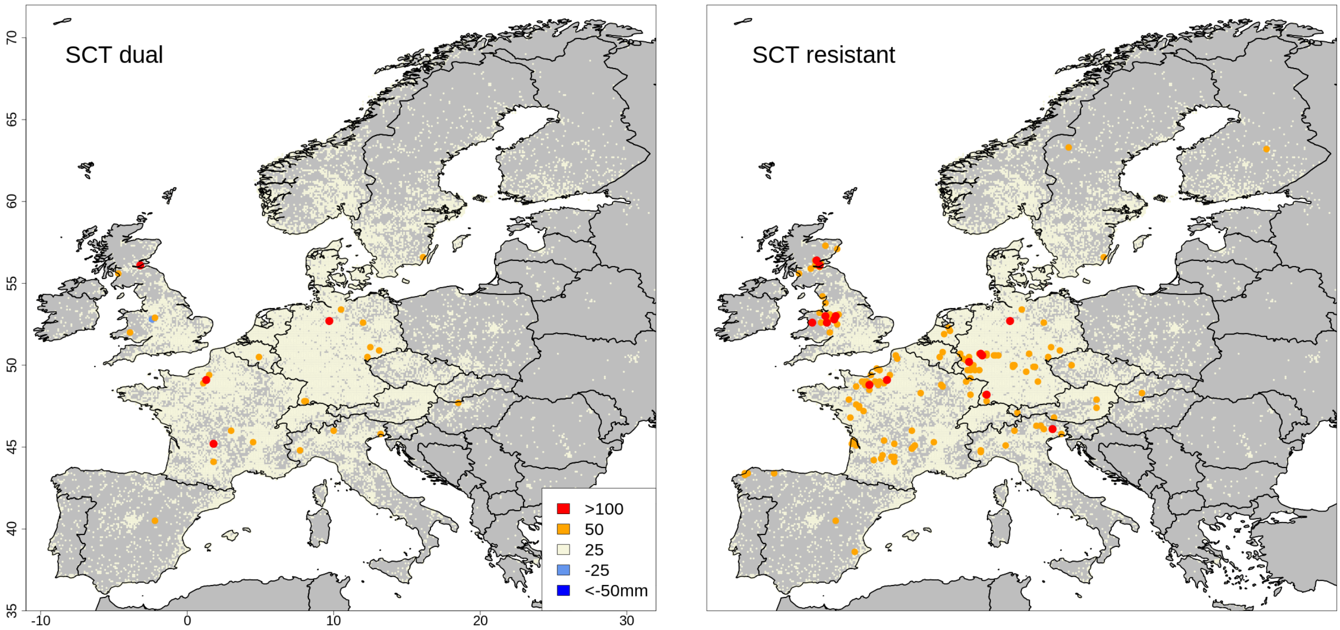


Figure 16: Convective precipitation over the United Kingdom and western Europe. Difference between total observed precipitation without and with the quality checks (i.e. raw totals minus quality controlled totals). The time period considered is the same as in Fig.15. In the left panel, SCT dual has been applied. In the right panel, SCT resistant has been applied.

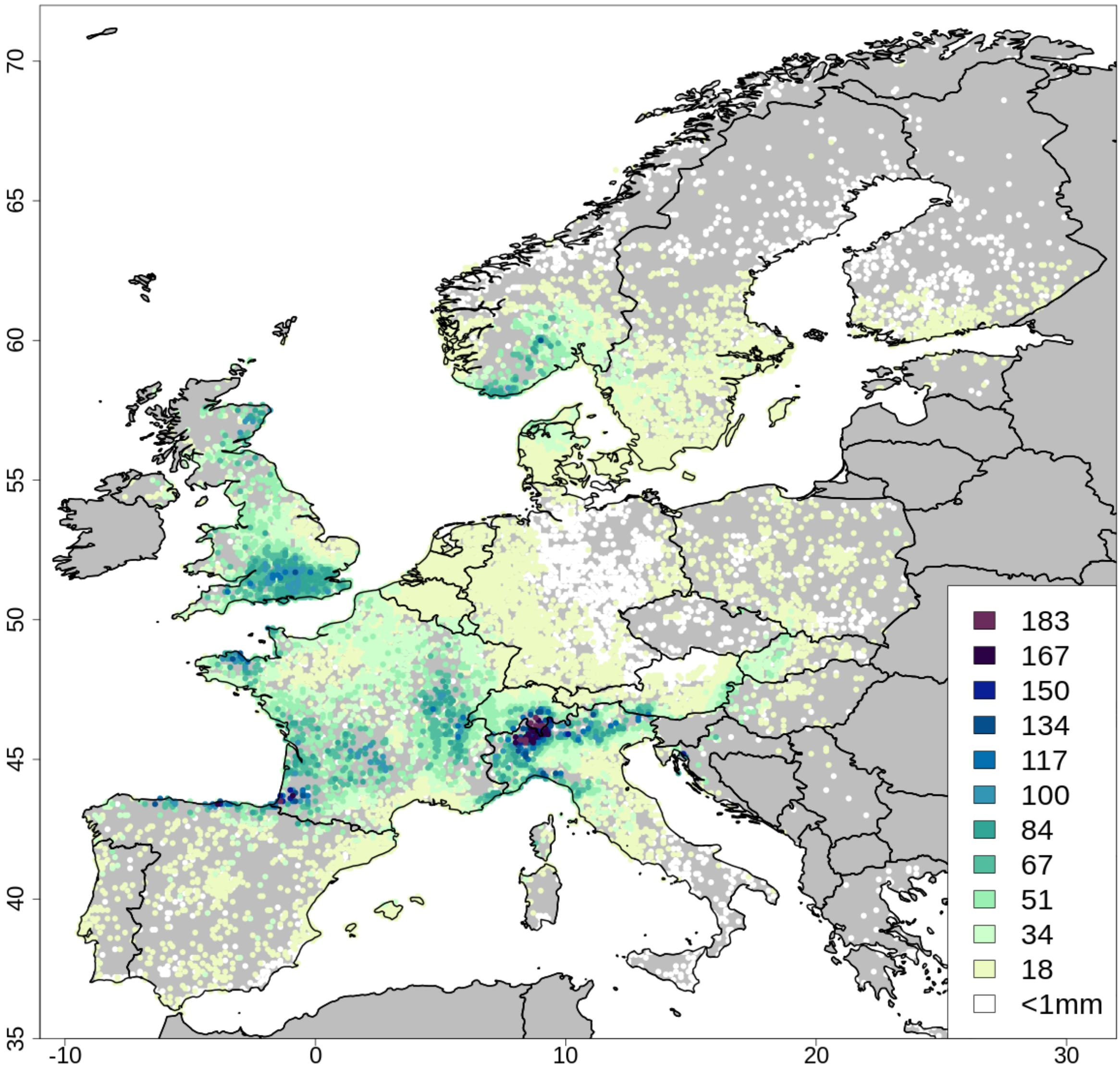


Figure 17: Storm Alex. Precipitation totals from 2020-10-02 00 UTC to 2020-10-05 00 UTC. The precipitation is accumulated using observations after SCT resistant has been applied.

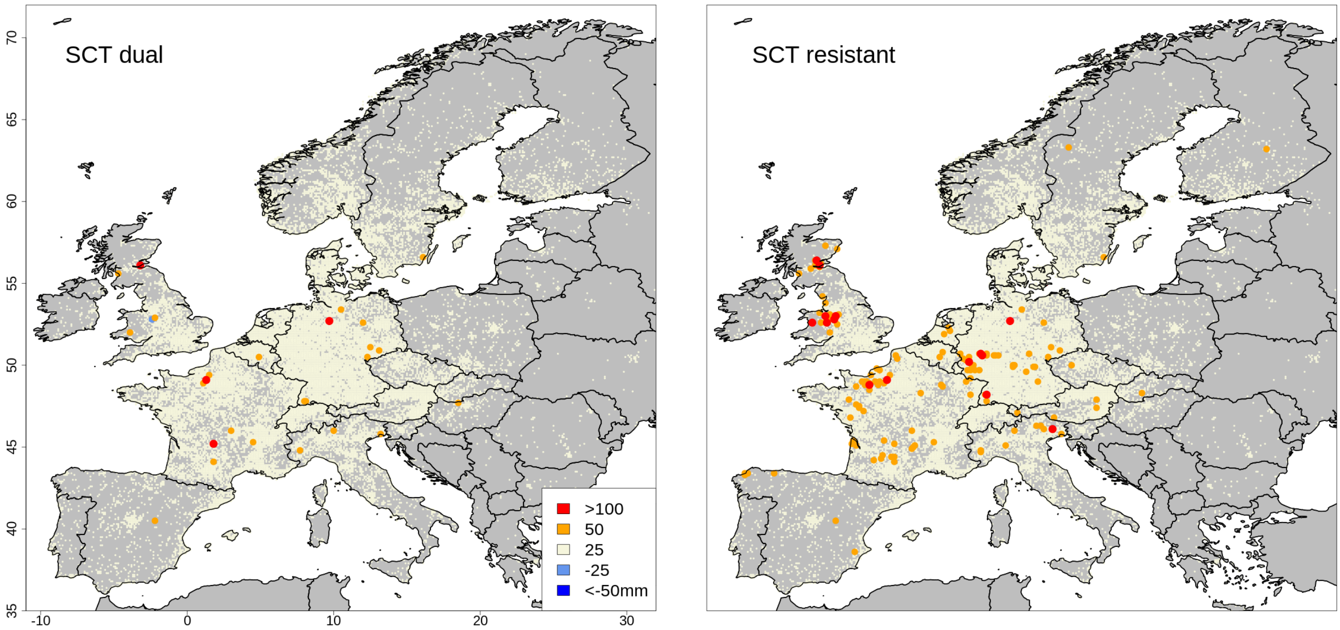


Figure 18: Storm Alex. Difference between total observed precipitation without and with the quality checks (i.e. raw totals minus quality controlled totals). The time period considered is the same as in Fig.17. In the left panel, SCT dual has been applied. In the right panel, SCT resistant has been applied.

A SCT dual settings

The SCT dual function signature is (`https://github.com/metno/titanlib/blob/master/src/sct_dual.cpp`):

```
ivec titanlib::sct_dual( const Points& points,
                         const vec& values,
                         const ivec& obs_to_check,
                         const vec& event_thresholds,
                         ConditionType condition,
                         int num_min_outer,
                         int num_max_outer,
                         float inner_radius,
                         float outer_radius,
                         int num_iterations,
                         float min_horizontal_scale,
                         float max_horizontal_scale,
                         int kth_closest_obs_horizontal_scale,
                         float vertical_scale,
                         const vec& test_thresholds,
                         bool debug)
```

The parameter values we have used to obtain the results presented in Sec. 4 are:

```
event_thresholds= 0.1 mm
condition= "greater than"
num_min_outer= 5
num_max_outer= 20
inner_radius= 10 km
outer_radius= 25 km
num_iterations= 10
min_horizontal_scale= 1 km
max_horizontal_scale= 10 km
kth_closest_obs_horizontal_scale= 3
vertical_scale= 200000 m (i.e. not used)
test_thresholds= 0.01
```

B SCT resistant settings

The SCT resistant function signature is (https://github.com/metno/titanlib/blob/master/src/sct_resistant.cpp):

```
ivec titanlib::sct_resistant( const Points& points,
                              const vec& values,
                              const ivec& obs_to_check,
                              const vec& background_values,
                              BackgroundType background_elab_type,
                              int num_min_outer,
                              int num_max_outer,
                              float inner_radius,
                              float outer_radius,
                              int num_iterations,
                              int num_min_prof,
                              float min_elev_diff,
                              float min_horizontal_scale,
                              float max_horizontal_scale,
                              int kth_closest_obs_horizontal_scale,
                              float vertical_scale,
                              const vec& values_mina,
                              const vec& values_maxa,
                              const vec& values_minv,
                              const vec& values_maxv,
                              const vec& eps2,
                              const vec& tpos,
                              const vec& tneg,
                              bool debug,
                              bool basic,
                              vec& scores)
```

The precipitation values have been transformed prior to executing the SCT. Specifically, we have applied a Box-Cox transformation with a transformation parameter equal to 0.5.

The parameter values we have used to obtain the results presented in Sec. 4 are:

```
background_elab_type= "MedianOuterCircle"
num_min_prof= 1 (not used with "MedianOuterCircle")
min_elev_diff= 100 (not used with "MedianOuterCircle")
num_min_outer= 10
num_max_outer= 30
```

```
inner_radius= 20 km
outer_radius= 50 km
num_iterations= 10
min_horizontal_scale= 1 km
max_horizontal_scale= 10 km
kth_closest_obs_horizontal_scale= 3
vertical_scale= 200000 m (i.e. not used)
values_mina= values - 500 mm
values_maxa= values + 500 mm
values_minv= values - 0.0001 mm
values_maxv= values + 0.0001 mm
basic= TRUE
eps2= 0.1
tpos= 1
tneg= 1
```

Note that: values_mina; values_maxa; values_minv and values_maxv are used to implement "short-cuts" to optimize the execution time of the routine. In this case, we have chosen values to avoid the routine taking the short-cuts.

C Code availability

All code is available in the GitHub repositories:

- Titanlib, `https://github.com/metno/titanlib`
- TitanTuner, `https://github.com/metno/titantuner`

References

Abraham, I. R., E. Alerskans, C. Lussana, T. N. Nipen, L. Oram, and I. A. Seierstad (2022), A strategy for the optimization of quality control checks available in the titanlib open library, *Tech. rep.*, EMS Annual Meeting 2022, doi:https://doi.org/10.5194/ems2022-178.

Alerskans, E., C. Lussana, T. N. Nipen, and I. A. Seierstad (2022), Optimizing spatial quality control for a dense network of meteorological stations, *Journal of Atmospheric and Oceanic Technology*, *39*(7), 973 – 984, doi:10.1175/JTECH-D-21-0184.1.

Båserud, L., C. Lussana, T. N. Nipen, I. A. Seierstad, L. Oram, and T. Aspelien (2020), Titan automatic spatial quality control of meteorological in-situ observations, *Advances in Science and Research*, *17*, 153–163, doi:10.5194/asr-17-153-2020.

Lussana, C., and L. Båserud (2021a), A spatial consistency test for the quality control of meteorological observations. part i: Methodology, *Tech. Rep. 11*, Norwegian Meteorological Institute.

Lussana, C., and L. Båserud (2021b), A spatial consistency test for the quality control of meteorological observations. part ii: Idealized experiments, *Tech. Rep. 12*, Norwegian Meteorological Institute.

Lussana, C., and L. Båserud (2021c), A spatial consistency test for the quality control of meteorological observations. part iii: Experiments on real-world data, *Tech. Rep. 13*, Norwegian Meteorological Institute.

Lussana, C., F. Uboldi, and M. R. Salvati (2010), A spatial consistency test for surface observations from mesoscale meteorological networks, *Quarterly Journal of the Royal Meteorological Society*, *136*(649), 1075–1088.