The use of personal weather station observation for improving precipitation estimation and interpolation

András Bárdossy[1], Jochen Seidel[1], and Abbas El Hachem[1]

[1]Institute for Modelling Hydraulic and Environmental Systems, University of Stuttgart, D-70569 Stuttgart, Germany

Correspondence: Jochen Seidel (jochen.seidel@iws.uni-stuttgart.de)

Abstract. The number of personal weather stations (PWS) with data available online through the internet is increasing gradually in many parts of the world. The purpose of this study is to investigate the applicability of these data for the spatial interpolation of precipitation for high intensity events of different durations. Due to unknown errors and biases of the observations rainfall amounts of the PWS network are not considered directly. Instead, only their temporal order is assumed to be
5 correct. The crucial step is to find the stations with informative measurements. This is done in two steps, first by selecting the locations using time series of indicators of high precipitation amounts. The remaining stations are checked whether they fit into the spatial pattern of the other stations. Thus, it is assumed that the percentiles of the PWS network are accurate. These percentiles are then *translated* to precipitation amounts using the distribution functions which were interpolated using the information from German National Weather Service (DWD) data only. The suggested procedure was tested for the State of
10 Baden-Württemberg in Germany. A detailed cross validation of the interpolation was carried out for aggregated precipitation amounts of 1, 3, 6, 12 and 24 hours. For each aggregation nearly 200 intense events were evaluated. The results show that filtering the secondary observations is necessary as the interpolation error after filtering and data transformation decreases significantly. The biggest improvement is achieved for the shortest time aggregations.

1 Introduction

15 Comprehensive reviews on the current state of citizen science in the field of hydrology and atmospheric sciences were published by Buytaert et al. (2014) and Muller et al. (2015). Both of these reviews give a detailed overview of the different forms of citizen science data and highlight the potential to improve knowledge and data in the fields of hydrology and hydro-climatology. One type of information which is of particular interest for hydrology are data from in situ sensors. In recent years, the amount of low cost personal weather stations (PWS) has increased with an incredible speed. Data from PWS are published online
20 on internet portals such as Netamto (www.netatmo.com) or Weather Underground (www.wunderground.com). These stations provide weather observations which are available in real time as well as for the past. This is potentially very useful to complement systematic weather observations of national weather services, especially with respect to precipitation, which is highly variable in space and time. Traditionally rainfall is interpolated using point observations. The shorter the time aggregation the higher the variability of rainfall becomes, and the more the quality of interpolation deteriorates (Bárdossy and Pegram, 2013;
25 Berndt and Haberlandt, 2018). In consequence, the number of interpolated precipitation products with sub-daily resolution is

low, but such data would be required for many hydrological applications (Lewis et al., 2018). Additional information such as radar measurements can improve interpolation (Haberlandt, 2007), however, radar rainfall is still highly prone to different kinds of errors (Villarini and Krajewski, 2010). Against the backdrop of low precipitation station densities, the additional data from PWS has a high potential to improve the information of spatial and temporal precipitation characteristics. However, one

30 of the major drawbacks from PWS precipitation data is their trustworthiness. There is little systematic control on the placing and correct installation and maintenance of the PWS, so it is usually not known whether a PWS is set up according to the international standards published by the WMO (World Meteorological Organization, 2008). The measured data itself may have unknown errors which can be biased and contain independent measurement errors, too. Therefore, the data from PWS networks cannot be regarded to be as reliable as those of professional networks operated by national weather services or environmental

35 agencies. Hence, the use of PWS data requires specific efforts to account for these errors. For air temperature measurements, Napoly et al. (2018) developed a quality control (QC) procedure to filter out suspicious measurements from PWS stations that are caused e.g. by solar exposition or incorrect placement. For precipitation, de Vos et al. (2017) investigated the applicability of personal stations for urban hydrology in Amsterdam, Netherlands. They reported results of a systematic comparison of an official observation of the Royal Netherlands Meteorological Institute (KNMI) and a PWS Netatmo rain gauge. This provides

40 information on the quality of measurements in case of correct installation of the devices. As many of the PWS may be placed without consideration of the WMO standards, the results of these comparisons cannot be transferred to the other PWS observations. In a more recent study, de Vos et al. (2019) developed a QC methodology of PWS precipitation measurements based on a combined official rain gauge and radar product over the Netherlands. This however can be problematic as radar data has errors as well (e.g. attenuation, clutter, beam blockage) and thus the quantitative precipitation estimation (QPE) is

45 often uncertain Villarini and Krajewski (2010). Furthermore, on the shorter time scales effects such as attenuation or wind drift lead to a disagreement between radar data and rain gauge data (Yan and Bárdossy, 2019). In addition, the study by de Vos et al. (2019) does not provide a guideline on how to use the measurements of the PWS if no radar observations are available. Overall, the data from PWS rain gauges may provide useful information for many precipitation events and may also be useful for real time flood forecasting, but data quality issues have to be overcome. In this paper we focus on the use of PWS data for

50 the interpolation of intense precipitation events. We propose a two fold approach based on indicator correlations and spatial patterns to filter out suspicious measurements and to use the information from PWS indirectly. The basic assumption hereby is that many of the stations may be biased but are correct in the temporal order. For the spatial pattern, information from a reliable precipitation network, e.g. from a national weather service is required. These measurements are considered to be more trustworthy than the PWS data, however, the number of such stations is usually much lower. This paper is organized as fol-

55 lows: After the introduction, the methodology to find useful information and the subsequent interpolation steps are described. The described procedure was used for precipitation events of the last four years in the federal state of Baden-Württemberg in South-West Germany. The results of the interpolation and the corresponding quality of the method are discussed in section 4. The paper ends with a discussion and conclusions.

2 Study Area and Data

60 The federal state of Baden-Württemberg is located in South-West Germany and has an area of approximately 36,000 km^2. The annual precipitation varies between 600 and 2100 mm (Deutscher Wetterdienst, 2020), and the highest amounts are recorded in the higher elevations of the mountain ranges of the Black Forest. The rain gauge network of the German Weather Service (DWD) in Baden-Württemberg (referred to as primary network from here on) currently comprises 111 stations for the study period with high temporal resolution data (Fig. 1). The gauges used in this network are typically weighing gauges.
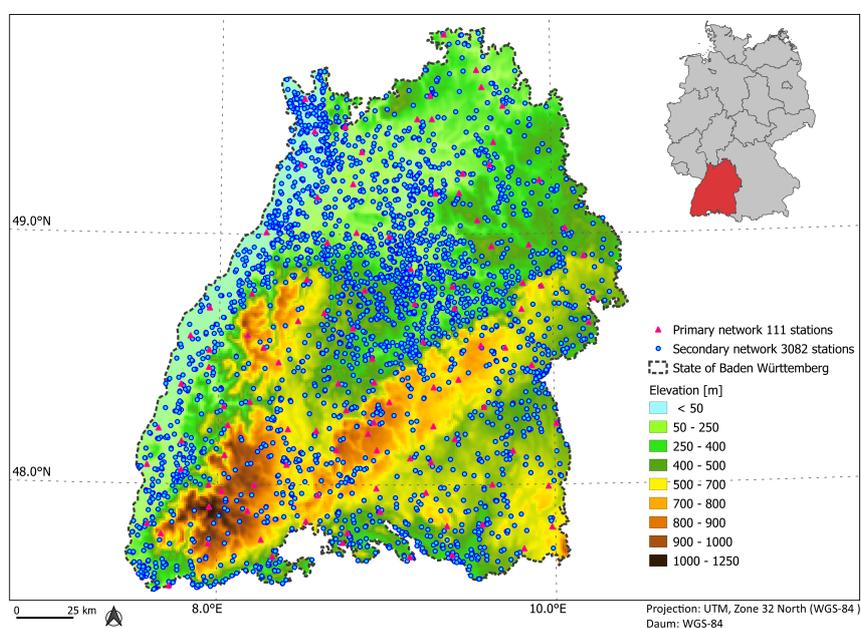

Figure 1. Map Of the federal state of Baden-Württemberg showing the topography and the location of the DWD (primary) and Netatmo (secondary) gauges.

65   In order to asses the spatial variability within a dense network of primary gauges, the precipitation data from the municipality of Reutlingen (located about 30 km south of the state capital Stuttgart) was additionally used. This city operates a dense network of 12 weighing rain gauges (OTT Pluvio^2) since 2014 in an area of 87 km^2 (not shown in Fig 1). Furthermore, three Netatmo rain gauges were installed at the Institute's own weather station on the Campus of the University of Stuttgart, where a Pluvio^2 weighing rain gauge is installed as well. This allows a direct comparison between the gauges from the primary network and the

70 secondary network in the case the latter are installed and maintained correctly. For PWS, the Netatmo network was selected (https://weathermap.netatmo.com). Data from this PWS network (referred to as secondary network from here onwards) can be downloaded with the Netatmo API in different temporal resolutions down to 5 minutes. The Netatmo rain gauges are plastic tipping buckets which have an opening orifice of 125 cm^2 (compared to 200 cm^2 of the primary network). Since these devices are not heated, their usage is limited to liquid precipitation. To take this into account, data from secondary stations were only

75 used in case the average daily air temperature at the nearest DWD station was above 5 °C. The number of gauges from the

secondary network varies over time. The time from 2015 to 2019 was considered for this study. At the end of the time period over 3000 stations from the secondary network were available. One can see that many stations have less than one year of observations. The proposed methodology cannot accommodate these stations, but in the future it is likely that a large portion of them can be considered.
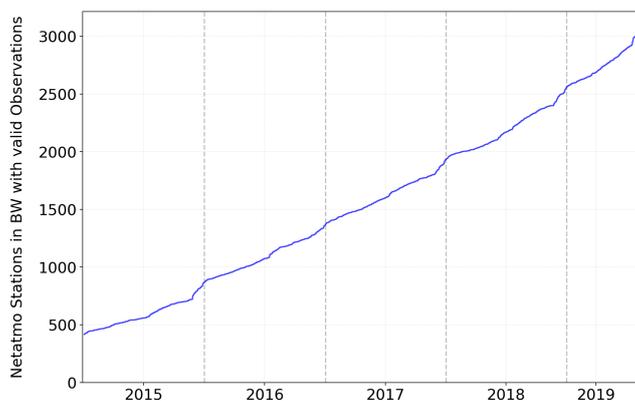

Figure 2. Development of the number of online available Netatmo rain gauges.

80   Figure 2 shows the number of secondary stations as a function of time. The stations from the secondary network show an uneven distribution in space, which mainly reflects the population density and topography of the study area. The number of secondary stations is higher in densely populated areas are such as in the Stuttgart metropolitan area and the Rhine-Neckar Metropolitan Region. Furthermore, there are no secondary network stations above 1000 m a.s.l., however the primary network only has one station above 1000 m (at the Feldberg summit at 1496 m) as well.

3   Methodology

85

It is assumed that the secondary stations may have individual measurement problems, (e.g. incorrect placement, lack of and/or wrong maintenance, data transmission problems) and due to their large number there is no possibility to check their proper placing and functioning directly. Furthermore, at many locations (especially in urban areas) there is no possibility to set up the rain gauges so that they fulfil the WMO standards. Therefore, the first goal is to filter out stations which deliver data
90   contradicting the observations of the primary network which meet the WMO standards. Two filters are applied - the first one compares the secondary time series with the closest primary series with the focus on intense precipitation. The second filter is designed to remove individual contradicting observations using a spatial comparison.

3.1   High intensity indicator based filtering

This relationship is independent of a possible station bias and is only important for high intensities, since for most hydrological
95   applications low precipitation values play a minor role. A secondary station is useful if this relationship holds. Unfortunately

the assumption can only be checked for selected test locations. Since is not intended to use the data from the secondary stations directly, their temporal ranks which are considered as indicator series of intense precipitation are used for this purpose instead.

Observations from the primary and secondary network are available at short time steps and can be be aggregated to different Δt durations. The usefulness of the secondary data is investigated for different time aggregations. $Z_{\Delta t}(x,t)$ is the (partly
100 unknown) precipitation at location x and time t integrated over the time interval Δt. It is assumed that this precipitation is measured by primary network at locations $\{x_1,\ldots,x_N\}$. The measurements of the secondary network are indicated as $Y_{\Delta t}(y_j,t)$ at locations $\{y_1,\ldots,y_M\}$. Note that Y is considered to be a random field, and thus methods like Co-Kriging or Kriging with an external drift are not applicable.

In order to identify stations which are likely to deliver reasonable data for high intensities, indicator correlations are used.
105 For a selected variable $U = Z$ or $U = Y$ and probability α the indicator series

$$I_{\alpha,,U}(x,t) = \begin{cases} 1 & \text{if } F_{u,\Delta t}\left(U_{\Delta t}(x,t)\right) > \alpha \\ 0 & \text{else} \end{cases} \tag{1}$$

For any two locations corresponding to the primary network x_i and x_j and any α and Δt the correlation (in time) of the indicator series is $\rho_{Z,\alpha,\Delta t}(x_i,x_j)$ and provides an information on how precipitation series vary in space. This indicator correlation usually decreases with increasing separation distance. This decrease is not at the same rate everywhere and not
110 the same for different thresholds and aggregations. For the secondary network indicator correlations $\rho_{Z,Y,\alpha,\Delta t}(x_i,y_j)$ with the series in the primary network can be calculated. This can then be compared to the indicator correlations of the primary network.

The sample size has a big influence on the variance of the indicator correlations. Therefore, to take into account the limited interval of availability of the secondary observations, indicator correlations of the primary network corresponding to the same periods for which the secondary variable is available are used for the comparison. This is done individually for each secondary
115 site. A secondary station is flagged as suspicious if its indicator correlations with the nearest primary network points are below the lowest indicator correlation corresponding to the primary network for the same time steps and at the same separation distance. This means if:

$$\rho_{Z,Y,\alpha,\Delta t}(x_i,y_j) < \min\left\{\rho_{Z,\alpha,\Delta t}(x_k,x_m)\,;\, \| x_k - x_m \| \approx \| x_i - y_j \|\right\} \tag{2}$$

then the secondary station shows weaker association to the primary than what one would expect from primary observations.
120 In this case it is reasonable to discard the measured time series corresponding to the secondary network at location y_i. This procedure can be repeated for a set of selected α values. High α-s (dependent on the aggregation interval Δt are preferred as the goal is to improve precipitation estimation for strong precipitation events.

3.2 Precipitation amount estimation for secondary observations

After the selection of the potentially useful secondary stations the next step is to correct their observations. The distribution
125 function of the measured precipitation values at locations x_i of the primary and at locations y_j of the secondary network are denoted as $F_{x_i,\Delta t}(z)$ and $G_{y_j,\Delta t}(z)$ respectively. The basic assumption for the suggested approach is that the measured

precipitation data from the secondary network may be biased in their values but are good in their order (at least for high intensities). This means that if at times t_1 and t_2:

$$Y_{\Delta t}(y_i, t_1) < Y_{\Delta t}(y_i, t_2) \Rightarrow Z_{\Delta t}(y_i, t_1) < Z_{\Delta t}(y_i, t_2) \tag{3}$$

130   This means that the measured precipitation amount from the secondary network is likely to have an unknown bias, but the order of values at a location is preserved. This assumption is likely to be reasonable for high precipitation intensities. Thus, the percentile of the precipitation observed at a given time at a secondary location can be used for the estimation of precipitation amounts. Since this is a percentile and not a precipitation amount it has to be converted to a precipitation amount for further use. This can be done using the distribution function of precipitation amounts corresponding to the location y_j and the aggregation

135   Δt. As the secondary observation could be biased their distribution $G_{y_j, \Delta t}$ cannot be used for this purpose. Thus, one needs an unbiased estimation of the local distribution functions.

Distribution functions based on long observation series are available for the locations of the primary network. For locations of the secondary network they have to be estimated via interpolation. This can be done by using different geostatistical methods. A method for interpolating distribution functions for short aggregation times is presented in Mosthaf and Bardossy (2017).

140   Another possibility is to interpolate the quantiles corresponding to selected non-percentiles or interpolating percentiles for selected precipitation amounts. Another alternative to estimate distribution functions corresponding to arbitrary locations is to use functional kriging (Giraldo et al., 2011) to interpolate the distribution functions directly. The advantage of interpolating distribution functions is that they are strongly related to geographical locations of the selected location and to topography. These variables are available in high spatial resolution for the whole investigation domain. Additionally, observations from

145   different time periods and time aggregations can also be taken into account as co-variates.

In this paper Ordinary Kriging (OK) is used for the interpolation of the quantiles and for the percentiles to construct the distribution functions both for the locations of the secondary observations and for the whole interpolation grid. For a given aggregation Δt, time t and target secondary location y_j the observed percentile of precipitation is:

$$P_{\Delta t}(y_j, t) = G_{y_j, \Delta t}\left(Y_{\Delta t}(y_j, t)\right) \tag{4}$$

150   For the observations of the primary network the quantiles of the precipitation distribution at the primary stations are selected. The distributions at the primary stations are based on the same time steps as those which have valid observations at the target secondary station. This way a possible bias due to the short observation period at the secondary location can be avoided. The quantiles are:

$$Q_{\Delta t}(x_i) = F_{\Delta t, x_i}^{-1}\left(P_{\Delta t}(y_j, t)\right) \tag{5}$$

155   These quantiles are interpolated using OK to obtain an estimate of the precipitation at the target location.

$$Z_{\Delta t}^{o}(y_j, t) = \sum_{i=1}^{n} \lambda_i Q_{\Delta t}(x_i) \tag{6}$$

Here the λ_i-s are the weights calculated using the Kriging equations. Note that the precipitation amount at the target location is obtained via interpolation, but the interpolation is not using the primary observations corresponding to the same time, but instead is using the quantiles corresponding to the percentile of the target secondary station observation. Thus these values

160 may exceed all values observed at the primary stations at time t. Note that this correction of the secondary observations is non-linear. This procedure is used for all locations which were accepted after application of the temporal filter.

3.3 Event based spatial filtering

While some stations may work properly in general, due to unforeseen events (such as battery failure or transmission errors) at certain times they may deliver individual false values. In order to filter out these errors a simple geostatistical outlier detection

165 method is used as described in Bárdossy and Kundzewicz (1990). For a given aggregation Δt, time t and target secondary location y_j the precipitation amount is estimated via OK using the observations of aggregation Δt at time t of primary stations. This value is denoted as $Z^*_{\Delta t}(y_j, t)$. If the precipitation amount at the secondary station estimated using (6) differs very much from $Z^*_{\Delta t}(y_j, t)$, the secondary location is discarded for the interpolation. As limit for the difference 3 times the Kriging standard deviation was selected. Formally:

170
$$\left| \frac{Z^*_{\Delta t}(y_j, t) - Z^o_{\Delta t}(y_j, t)}{\sigma_{\Delta t}(y_j, t)} \right| > 3 \tag{7}$$

This means that if the estimated precipitation at the secondary location does not fit into the pattern of the primary observations then it is discarded. Note that this filter is not discarding secondary observations which differ from the primary - it only removes those where there is a strong local disagreement. This procedure is most frequently removing false zeros at secondary observations which are due to temporary loss of connection between the rain gauge module and the Netatmo base station.

175 Note that this method could also be applied using the percentiles.

This and the previous procedure allow the selection of secondary data which can be used for precipitation interpolation.

3.4 Interpolation of the precipitation amounts

Once the percentiles of the secondary locations are converted to precipitation amounts one can use different Kriging procedures for the interpolation over a grid in the target region. The simplest solution is to use OK. For aggregations of one day or longer the

180 orographic influence should be taken into account. This can be done by using External Drift Kriging (Ahmed and de Marsily, 1987).

The problem with these procedures is that the precipitation amounts of the secondary network are more uncertain than those of the primary network. To reflect this difference a modified version of Kriging as described in Delhomme (1978) is applied. This allows a reduction of the weights for the secondary stations.

185 ### 3.4.1 Kriging using uncertain data

Suppose that for each point y_i time t and time aggregation Δt there is an unknown error of the percentiles $\varepsilon(y_i, t)$ which has the following properties:

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

1. Unbiased :

$$E[\varepsilon(y_i, t)] = 0 \tag{8}$$

2. Uncorrelated :

$$E[\varepsilon(y_i, t)\varepsilon(y_j, t)] = 0 \text{ if } i \neq j \tag{9}$$

3. Uncorrelated with the parameter value:

$$E[\varepsilon(y_i, t)Z(y_i, t)] = 0 \tag{10}$$

For the primary network we assume that $\varepsilon(x_i, t) = 0$.

The interpolation is based on the observations

$$\{u_1, \ldots, u_N\} = \{x_1, \ldots, x_N\} \cup \{y_1, \ldots, y_M\} \tag{11}$$

For any location x

$$Z^*_{\Delta t}(x, t) = \sum_{i=1}^{n} \lambda_i \left(Z(u_i, t) + \varepsilon(u_i, t) \right) \tag{12}$$

To minimize the estimation variance an equation system similar to the OK system has to be solved, namely:

$$\sum_{j=1}^{n} \lambda_j \gamma(u_i - u_j) + \lambda_i E[\varepsilon(u_i, t)^2] + \mu = \gamma(u_i - x) \ i = 1, \ldots, n$$

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{13}$$

Note that OK is a special case of this procedure with the additional assumption $\varepsilon(y_j, t) = 0$. This system leads to an increase of the weights for the primary and a decrease of the weights for the secondary network. For each time step and percentile the variances of the random error terms $\varepsilon(y_i, t)$ is estimated from the interpolation error of the distribution functions. This interpolation method is referred to as Kriging using uncertain data (KU).

4 Application and Results

The section describing the application of the methodology is divided into three parts. First the rationale of the assumptions is investigated. As a second step the methodology is applied on a large number of intense precipitation events on different time aggregations using a cross validation approach. This allows an objective judgment of the applicability of the results. Finally the results of the interpolation on a regular grid are shown and compared.

Table 1. Statistics of three Netatmo stations compared to a Pluvio weighing gauge for April to October 2019 at the IWS Meteorological Station

	1h				6h				24h			
	Pluvio	N07	N10	N11	Pluvio	N07	N10	N11	Pluvio	N07	N10	N11
p_0 [-]	0.92	0.84	0.84	0.91	0.82	0.75	0.84	0.82	0.59	0.56	0.65	0.59
mean [mm]	1.24	1.46	1.80	1.41	3.46	4.04	4.24	3.89	5.78	7.28	7.51	7.02
std [mm]	2.15	2.52	4.49	2.52	4.86	5.77	7.55	5.71	8.46	10.49	11.52	10.33
25% [mm]	0.18	0.20	0.10	0.20	0.39	0.33	0.30	0.40	0.48	0.63	0.58	0.58
50% [mm]	0.51	0.71	0.50	0.61	1.49	1.41	0.91	1.21	2.36	2.78	1.62	2.58
75% [mm]	1.34	1.72	1.41	1.52	4.60	5.33	4.14	4.95	7.82	9.87	11.26	9.95
max [mm]	19.84	22.62	44.74	22.22	23.28	28.58	44.74	27.98	45.62	55.55	56.16	55.55

All statistics except for the p_0 values are based on non-0 values

4.1 Justification of the methods

For a direct comparison between the secondary rain gauges and devices from the primary network, three Netatmo rain gauges we installed next to a Pluvio[2] weighing rain gauge (the same type as regularly used by the DWD) at the Institute for Modelling Hydraulic and Environmental Systems' (IWS) own weather station on the Campus of the University of Stuttgart. With this data
215 from 15 May to 15 October 2019 a direct comparison between the different devices used in the primary and secondary network was possible.

Table 1 shows statistics of the three devices compared to those of the reference station. The table shows that the secondary stations overestimated precipitation amounts by about 20 %. Furthermore, on can observe that the deviation between the reference and the Netatmo gauge are not linear, hence a data correction of the secondary gauges using a linear scaling factor
220 is not sufficient. Figure 3 shows scatter plots of hourly rainfall data and the corresponding percentiles from the three Netatmo gauges and a reference station.

Figure 3 shows that for high percentiles their occurrence is the same for the primary and the secondary devices. Although this is only one example with a relatively short time period it does support our assumption that the quantiles between primary and secondary stations are similar for higher precipitation intensities. However, one secondary device (N10) delivered data
225 which deviates substantially from the other measurements. This was caused by an interrupted connection between the rain sensor and the base station. In this case, the total sum of precipitation over a longer time period was transferred at once (i.e. in one single measurement interval) when the connection was established again. This leads to an extreme outlier which falsifies the results. The first filtering procedure can identify such problems effectively.

The secondary measurement devices can lead to very different biases depending on where and how they are installed. This
230 can be seen comparing the distribution functions of hourly precipitation accumulations corresponding to a set of very close primary stations with those of the secondary stations in the same area. Figure 4 shows the distribution functions of three primary and four secondary stations in the city of Reutlingen. While the distribution functions of the primary network are nearly
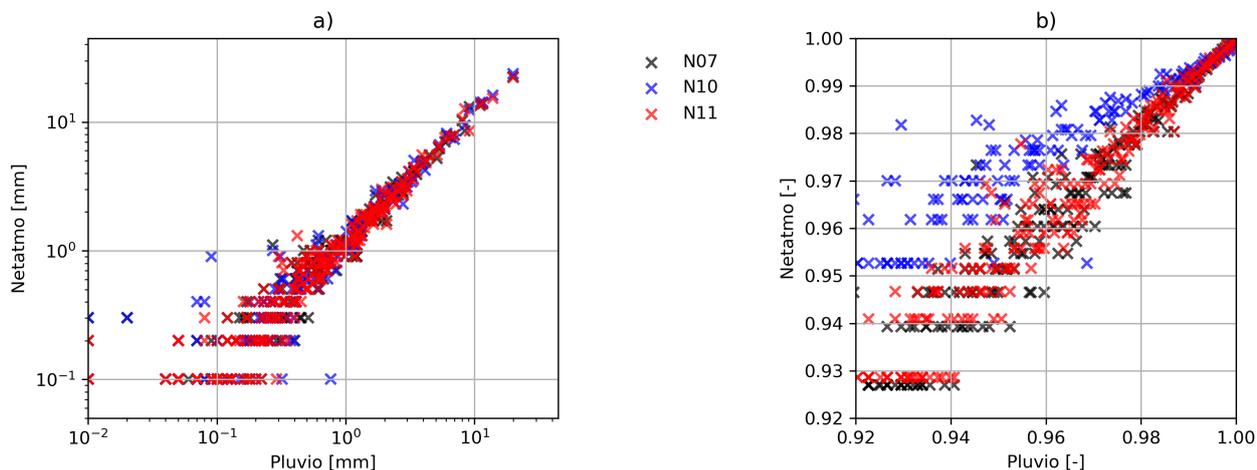
Figure 3. Scatter plot showing a) the hourly rainfall values (axes log-scaled) and b) the corresponding upper percentiles > 0.92 (right) between the Pluvio[2] weighing gauge and three Netatmo gauges (N07, N10, N11) at the IWS Meteorological Station.
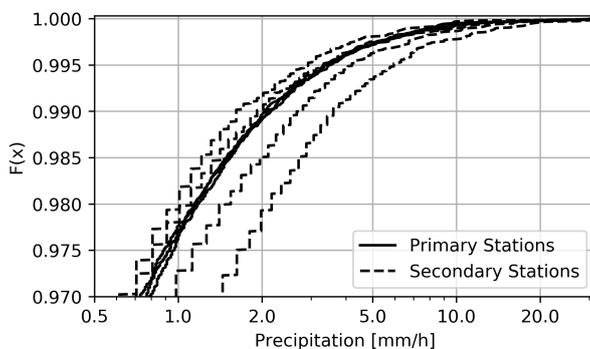

Figure 4. The upper part of empirical distribution functions of three primary stations (solid lines) and four secondary stations (dashed lines) from a small area in the city of Reutlingen based on a sample size of 15990 data pairs (hourly precipitation).

identical, those of the nearest secondary stations vary significantly. Some over and others underestimate the amounts significantly. This example supports the concept of the paper, namely that secondary data require filtering and data transformations before use. While the distributions differ, the probability of precipitation (p_0) ranges from 0.90 to 0.91 and is thus very similar for both types of stations indicating that the occurrence of precipitation can be well detected by the secondary network.

4.1.1 Application of the filters

Indicator correlations were calculated for different temporal aggregations and for a large number of different α values in the range between 95 and 99 %. Figure 5 shows the indicator correlations for one hour aggregation and the 99 % quantile using

240   pairs of observations of the primary-primary and the primary and secondary network as a function of station distance. The
indicator correlations of the pairs of the primary network show relatively high values and a slow decrease with increasing
distance. In contrast if the indicator correlations are calculated using pairs with one location corresponding to the primary and
one to the secondary network the scatter increased substantially. Secondary stations for which the indicator correlations are
very small in the sense of equations (2) are considered as unreliable and are removed from further treatment. A relatively large
245   distance tolerance was used as the density of the primary stations is much lower than the density of the secondary stations. On
the right panel the indicator correlations corresponding to the remaining secondary stations shows a similar spatial behaviour
as the primary network. In our case 881 stations remained after the application of this filter. This number is small compared
to the total number of available secondary stations, but note that the shortest records were removed and low correlations may
occur as a consequence of short observation periods, and in the future with increasing number of measurements some of these
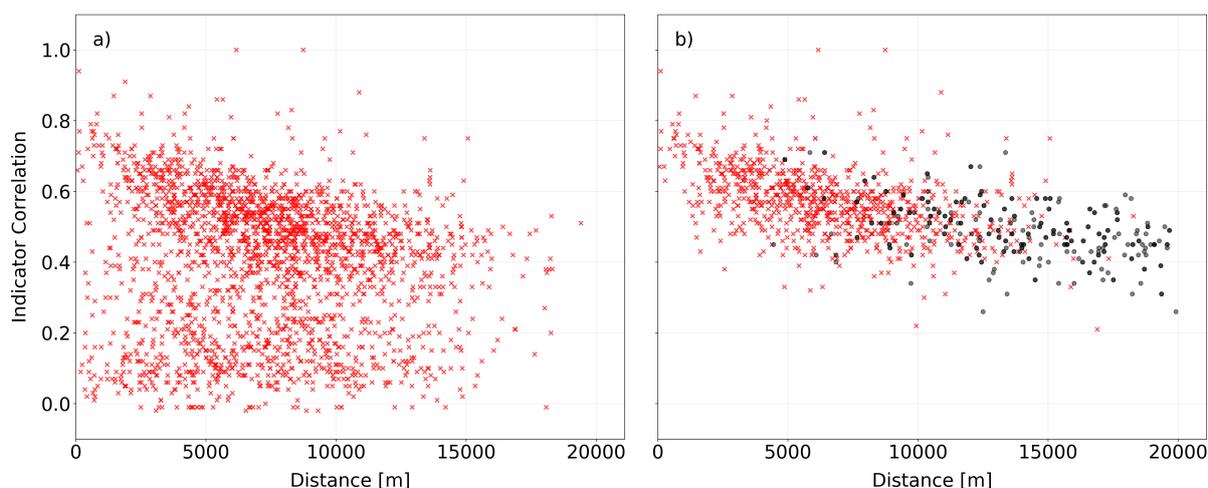250   stations may be reconsidered.

Figure 5. Indicator correlations for 1h temporal resolution and α = 0.99 between the secondary network and the nearest primary network
stations before (left) and after (right) applying the filter (red Xes). The black dots refer to the indicator correlation between the primary
network stations.

The second filter was applied for each event individually. The number of removed measurements was below 5 %. The
secondary filter did not play an important role in the procedure.

4.2   Cross validation results

As there is no ground truth available the quality of the procedure had to be tested by comparing omitted observations and their
255   estimates obtained after the application of the method.

The cross validation was carried out for a set of different time aggregations Δt and a set of selected events. Only times with
intense precipitation were selected, as for low intensity cases the interpolation based on the primary network is sufficiently

Hydrology and
Earth System
Sciences
Discussions

Table 2. Statistics of the selected intense precipitation events based on the primary network.

Temporal resolution	1 Hour	3 Hours	6 Hours	12 Hours	24 Hours
Number of intense events	185	190	190	195	195
Events between October-March	1	16	29	48	57
Events between April-September	184	174	161	147	138
Minimum of the maxima [mm]	28,01	31,2	33,35	34,9	35,5
Maximum of the maxima [mm]	122,3	158,2	158,4	160	210,3
p_0 (mean of all stations and events)	0,9	0,84	0,77	0,68	0,55

p_0 is defined here as precipitation <0.1mm

accurate. Table 2 shows some characteristics of the selected events. For short time periods nearly all events were from the summer season, while for longer aggregation the number of winter season events increased, but their portion remained below 30 %. Note the high portion of zeros for all aggregations.

The improvement obtained through the use of secondary data is demonstrated using a cross validation procedure. The primary network is randomly split into 10 subsets of 10 or 11 stations each. The data of each of these subsets was removed and subsequently interpolated using two different configurations of the data used, namely a) only other primary network stations (Reference 1) and b) using the other primary and the secondary network stations (Reference 2). For the latter case, the interpolations were carried out using the primary station data and the following configurations:

- C1: All secondary stations

- C2: Secondary stations remaining after the application of the temporal filter

- C3: Secondary stations remaining after application of the temporal and the event based spatial filter

- C4: Secondary stations remaining after application of the temporal and the event based spatial filter and considering uncertainty (KU)

The results were compared to the observations of the removed stations. The comparison was done for each location using all time steps and at each time step using all locations. Different measures including those introduced in Bárdossy and Pegram (2013) were used to compare the different interpolations. The results were evaluated for each time aggregation.

First, the measured and interpolated values were compared for each individual station and the Pearson (r) and Spearman correlations (ρ) of the observed and interpolated series were calculated. Table 3 shows the results for the different configurations used for the interpolation.

There is no improvement if no filter is applied - except a very slight improvement for 1 hour durations. This is mainly due to the better identification of the wet and dry areas. The use of the filters (and the subsequent transformation of the precipitation values) leads to an improvement of the estimation - the temporal filter being the most important. The spatial filter further improves the correlation while the additional consideration of the uncertainty of the corrected values at the secondary network

Table 3. Percentage of the stations with improved temporal correlation (compared to interpolation using primary stations only) for the configurations C1-C4.

Temporal aggregation	1 hour		3 hours		6 hours		12 hours		24 hours	
Number of events	185		190		190		195		195	
Correlation measure	r	ρ	r	ρ	r	ρ	r	ρ	r	ρ
C1: Primary and all secondary without filter OK	60	68	40	57	31	49	22	34	17	32
C2: Primary and secondary using temporal filter OK	81	91	75	90	73	90	64	84	52	81
C3: Primary and secondary using temporal and spatial filter OK	81	92	75	93	73	92	69	92	56	87
C4: Primary and secondary using temporal and spatial filter KU	81	92	75	92	74	91	70	91	56	86

r Pearson correlation, ρ Spearman correlation.

resulted in a marginal improvement. As the secondary stations are not uniformly distributed over the investigated domain the gain of using them is also not uniform. Highest improvements were achieved in and near urban areas with a high density of secondary stations, less improvement was achieved in forested areas with few secondary stations.

The measured and interpolated results were also compared for each event in space and (r) and (ρ) the observed and the
285 interpolated spatial patterns were calculated as well. Table 4 shows the results for the different configurations C1 to C4 used for the interpolation.

Table 4. Percentage of the stations with improved spatial correlation (compared to interpolation using primary stations only) for the configurations C1-C4.

Temporal aggregation	1 hour		3 hours		6 hours		12 hours		24 hours	
Number of events	185		190		190		195		195	
Correlation measure	r	ρ	r	ρ	r	ρ	r	ρ	r	ρ
C1: Primary and all secondary without filter OK	83	68	72	52	63	49	53	49	49	46
C2: Primary and secondary using temporal filter OK	96	97	90	93	90	93	84	89	80	85
C3: Primary and secondary using temporal and spatial filter OK	96	97	92	94	93	94	89	92	84	89
C4: Primary and secondary using temporal and spatial filter KU	93	94	90	92	90	93	84	89	80	87

r Pearson correlation, ρ Spearman correlation.

The use of secondary stations leads to a frequent improvement of the spatial interpolation even in the unfiltered case. The reason for this is that the spatial pattern is reasonably well captured by the secondary network. With increasing time aggregation the improvement disappears as the role of the bias increases. As in the case of the temporal evaluation the first filter (and the
290 subsequent transformation of the precipitation values) leads to the highest improvement. The spatial filter plays a marginal role, and the consideration of the uncertainty leads to a slight reduction of the quality of the spatial pattern. The improvement is smaller for higher temporal aggregations.

Hydrology and
Earth System
Sciences
Discussions
Open Access

EGU

Finally all results were compared in both space and time. Here the root mean squared error (RMSE) was calculated for all events and control stations. Table 5 shows the results for the different configurations used for the interpolation.

Table 5. RMSE (mm) for all stations and events.

Temporal aggregation	1 hour	3 hours	6 hours	12 hours	24 hours
Number of events	185	190	190	195	195
Reference 1: Primary stations only OK	5.97	6.97	7.34	7.71	8.35
C1: Primary and all secondary without filter OK	6.21	44.79	18.43	10.01	24.16
C2: Primary and secondary using temporal filter OK	4.83	6.05	6.61	7.33	8.29
C3: Primary and secondary using temporal and spatial filter OK	4.84	6.07	6.58	7.19	8.12
C4: Primary and secondary using temporal and spatial filter KU	4.82	6.02	6.53	7.15	8.08

295 The improvement is high for each aggregation. The temporal filter is important to improve interpolation quality. The spatial filter and the consideration of the uncertainty of the secondary stations are of minor importance. The improvement is the largest for the shortest aggregation (1 hour) where the RMSE decreased by 20 % and the smallest for the 24 hours aggregation with an improvement of 4 %. Decreasing spatial variability and increasing regularity with increasing time aggregation is the reason for these differences.

300 **4.3 Selected Case Studies**

As the cross validation results were showing improvements, the data transformations and subsequent interpolations were carried out for all selected events. As an illustration four case studies are shown and discussed here.

The first example (Fig. 6) shows the results of the interpolation of a 1 hour aggregated precipitation amount for the time period from 15:00 to 16:00 on June 11, 2018. The top panels of this figure show three different precipitation interpolations for
305 this event:

a) using the combination of the two station networks after application of the filters and transformation of the secondary data

b) using the primary network only

c) using raw data from the secondary network only

310 The panels in the bottom row of Figure 6 show d) the difference between b) and a), and e) the difference between b) and c). The three images a) to c) are similar in their rough structure, but there are important differences in the details. The interpolation using the primary network leads to a relatively smooth surface. The unfiltered secondary station based interpolation is highly variable and shows distinct patterns such as small dry and wet areas. The combination after filtering and transformation is more detailed than the primary interpolation, and in some regions these differences are high. The map of the difference between the
315 primary and the secondary station based interpolation (Fig. 6 e)) shows large regions of underestimation and overestimation by
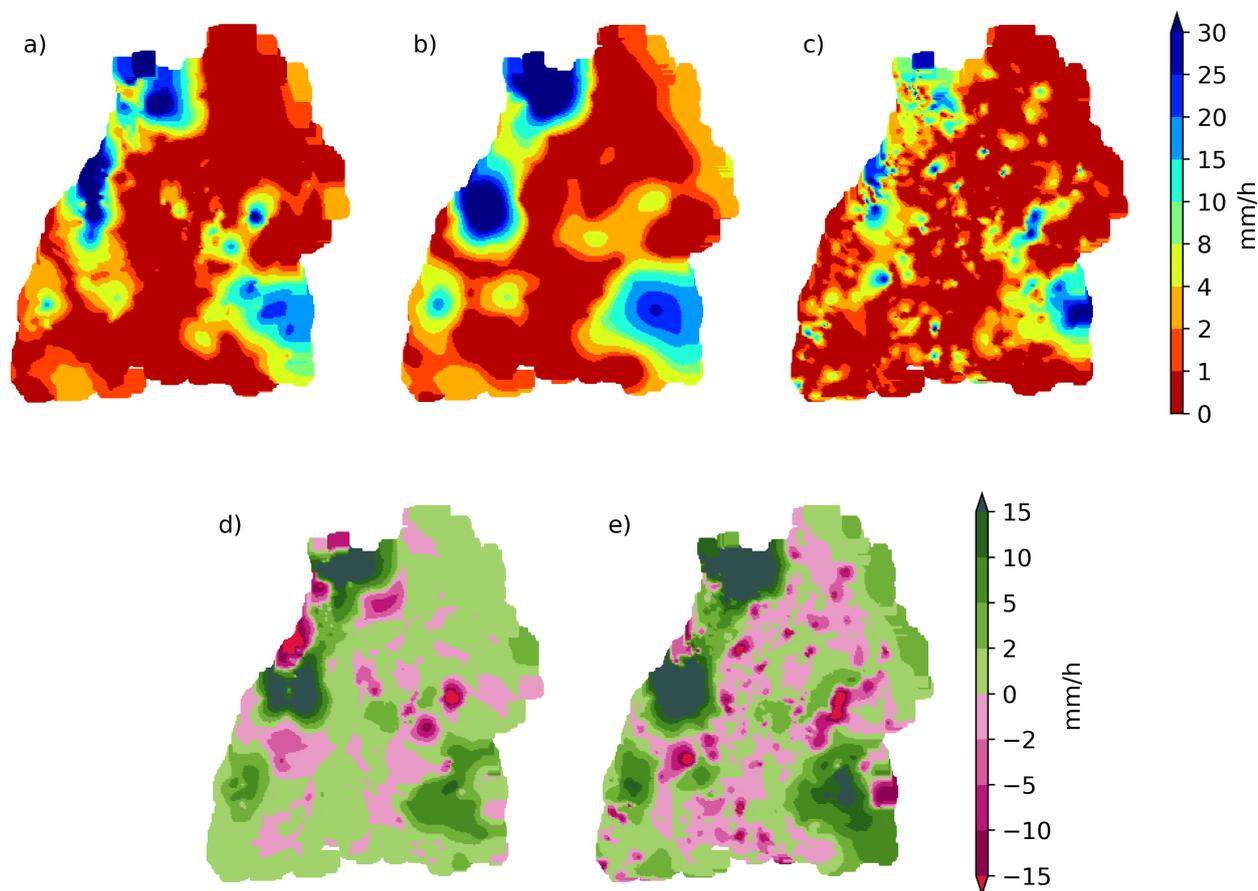
the secondary network. The differences between the primary and the filtered interpolations using transformed secondary data in panel d) is much smaller but in some regions the differences are still quite large, e.g. in the north-eastern part of the study area. In both cases, negative and positive differences occur. Note that for this data the cross validation based on the primary observations showed an improvement of r from 0.36 to 0.77, of ρ from 0.55 to 0.76 and a reduction of the RMSE from 12.5 to 320 8.2.

Figure 6. Interpolated precipitation for the time period 15:00 to 16:00 on June 11, 2018 (upper panel), and the differences between primary and combination, and primary and secondary data based interpolations. Panel a) shows the result after applying the filtering, b) the interpolation from the primary network and c) the one from the secondary network. Panels d) and e) depict the differences between b) and a) and b) and c) respectively.

Another interpolated 1 hour accumulation corresponding to 17:00 to 18:00 on September 6, 2018 is shown on Figure 7. These pictures show a similar behaviour to those obtained for June 11 (Fig. 6). Here, a high local rainfall in the southern central part of the study area was obviously not captured by the secondary network, leading to a large local underestimation in panel e). Furthermore, a larger area with precipitation in the primary network in the northern central in panel b) is significantly

325   reduced in size by the rainfall/no-rainfall information from the secondary network in panel c). For this case, the cross validation

based on the primary observations showed an improvement of r from 0.61 to 0.86, of ρ from 0.59 to 0.72 and a reduction of
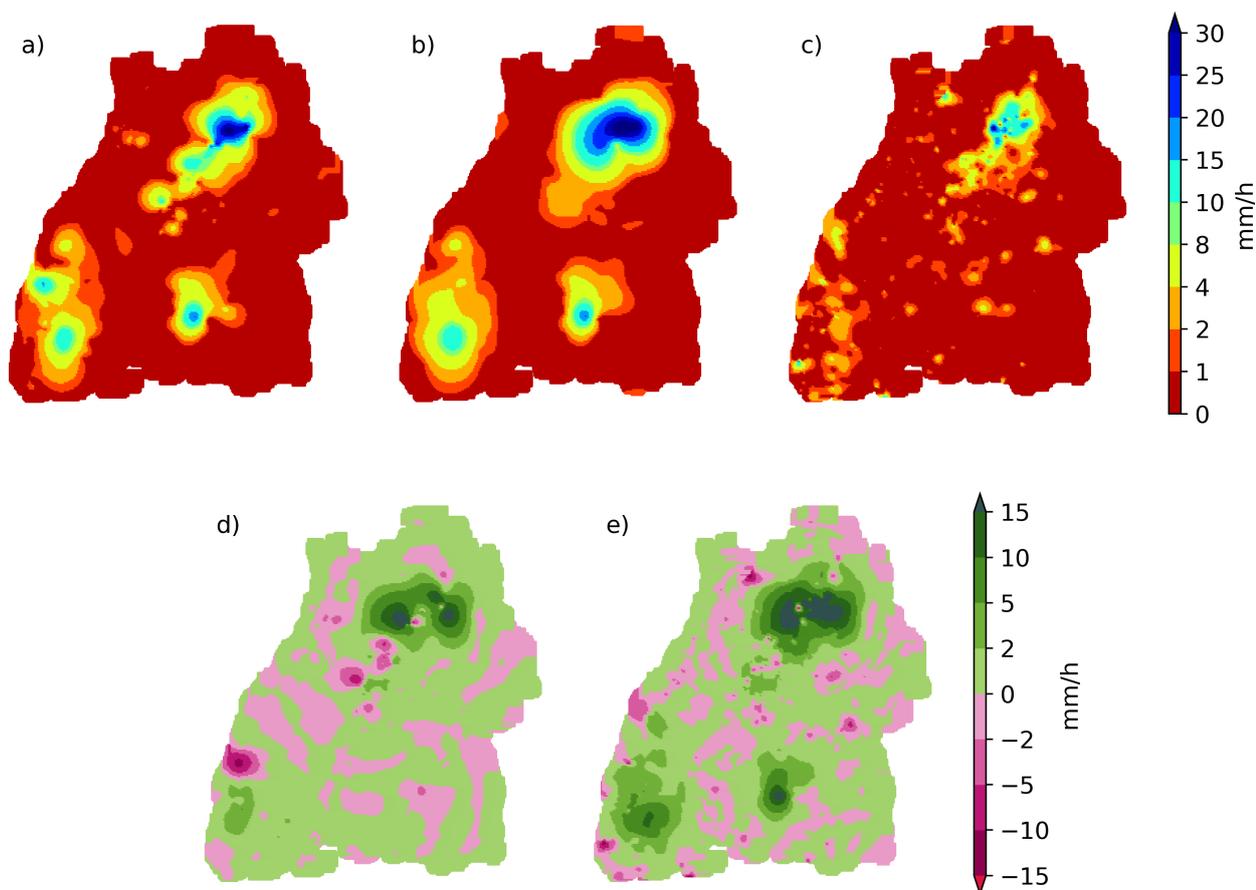
the RMSE from 5.65 to 3.75.

Figure 7. Interpolated precipitation for the time period 17:00 to 18:00 on September 6, 2018 (upper panel) and the differences between
primary and combination and primary and secondary data based interpolations. Panel a) shows the result after applying the filtering, b) the
interpolation from the primary network and c) the one from the secondary network. Panels d) and e) depict the differences between b) and a)
and b) and c) respectively.

The following two case studies show two interpolation examples for 24 hours which was the longest time aggregation in this study. Figure 8 shows the maps corresponding to the precipitation of 0:00 to 24:00 on May 14, 2018. The behaviour of the

330 interpolations is similar to the 1 hour cases shown above, the unfiltered and untransformed secondary interpolation is irregular and shows a systematic underestimation. Due to the longer aggregation, the local differences are less contrasting as in the case of hourly maps. The combination contains more details and the transition between high and low intensity precipitation is more complex. The difference between the primary (panel b)) and the combination based interpolation in panel a) is relatively smaller than for the 1 hour aggregations. This is caused by the reduction of the variability with increasing number of observations. Note

335 that for this data the cross validation based on the primary observations showed an improvement of r from 0.57 to 0.8, of ρ from 0.57 to 0.82 and a reduction of the RMSE from 15.99 to 13.61.
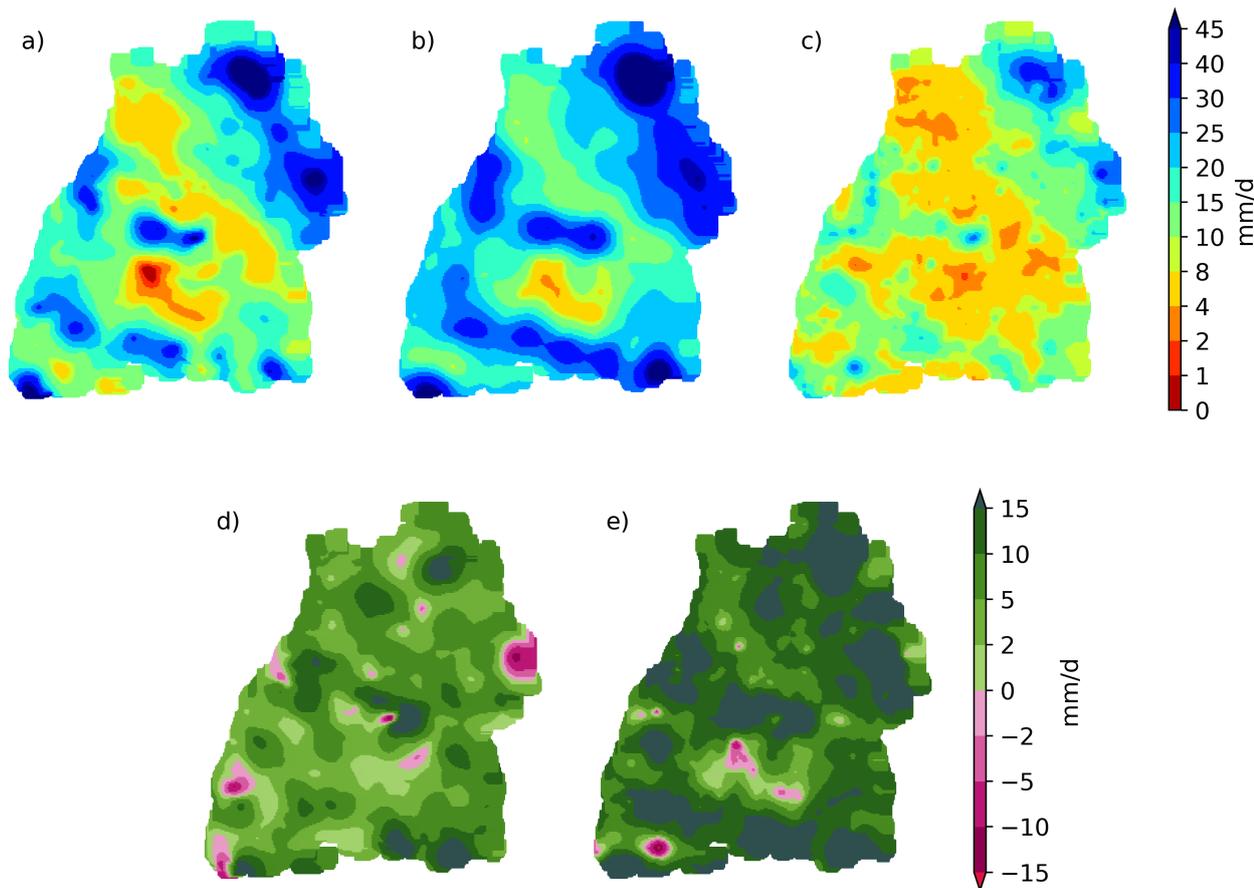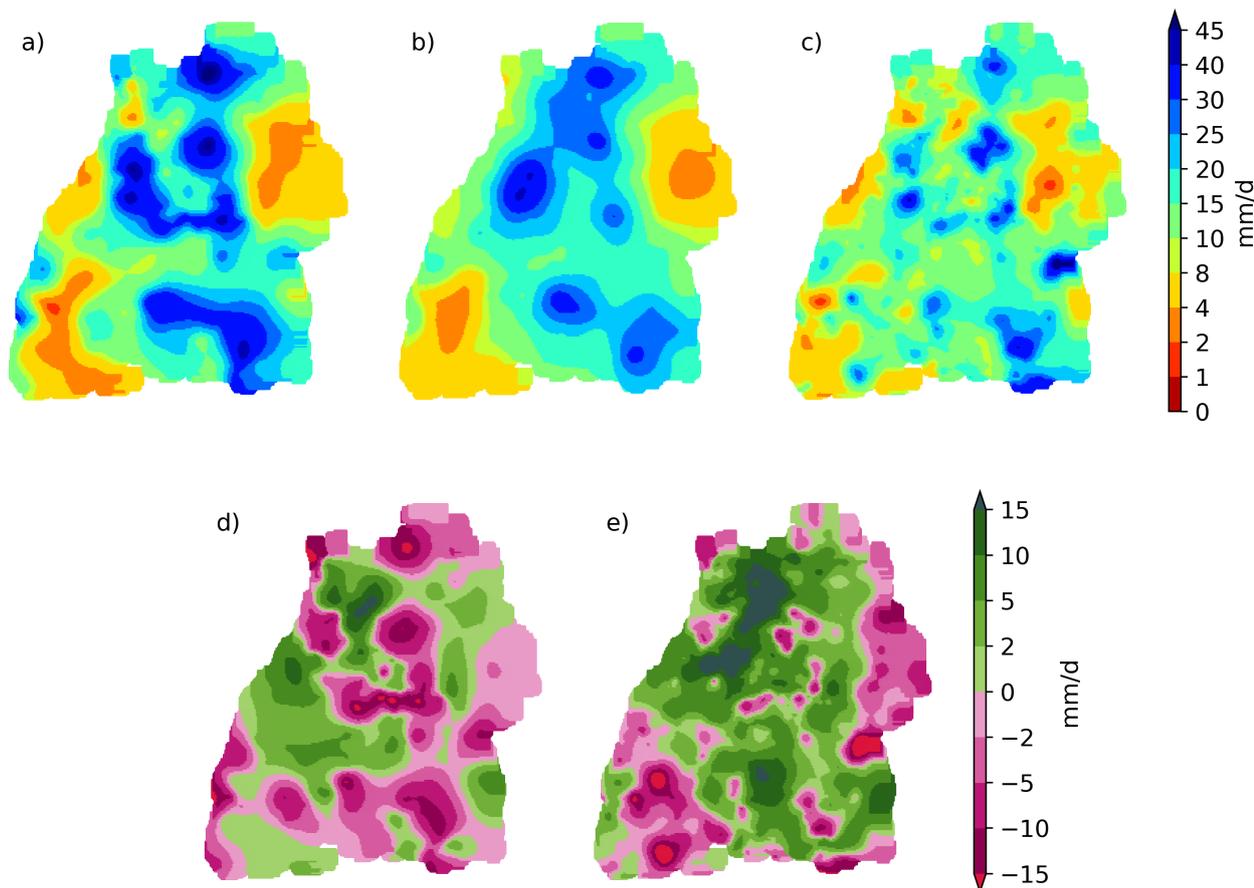

Figure 8. Interpolated precipitation for the time period for a 24h event from 0:00 to 24:00 on May 14, 2018 (upper panel) and the differences between primary and combination and primary and secondary data based interpolations. Panel a) shows the result after applying the filtering, b) the interpolation from the primary network and c) the one from the secondary network. Panels d) and e) depict the differences between b) and a) and b) and c) respectively.

Another interesting 24 hour event which was recorded on July 28, 2019 is shown in figure 9. The map based on the raw secondary data in panel c) shows very scattered intense rainfall. The combination of the primary and secondary observations changes the structure and the connectivity of these area with intense precipitation. The cross validation for this event showed

340 an improvement of r from 0.32 to 0.75, of ρ from 0.42 to 0.77 and a reduction of the RMSE from 14.77 to 10.21.

Figure 9. Interpolated precipitation for the time period for a 24h event from 0:00 to 24:00 on July 28, 2019 (upper panel) and the differences between primary and combination and primary and secondary data based interpolations. Panel a) shows the result after applying the filtering, b) the interpolation from the primary network and c) the one from the secondary network. Panels d) and e) depict the differences between b) and a) and b) and c) respectively.

The results of the filtering algorithm for the other events show a similar behaviour. The differences between primary and combined interpolation can be both positive and negative for all temporal aggregations. In general, the secondary network provides more spatial details, which could be very important for hydrological modelling of meso-scale catchments.

5 Discussion and conclusion

345 In this study an approach was presented on how data from PWS can be used to improve precipitation interpolation. The fundamental problem hereby is that the precipitation data from PWS are prone to various errors. An individual QC is either time consuming if a larger number of PWS is used or relies on other data sources as reference, such as precipitation estimates from weather radars which have an appropriate spatial and temporal resolution (de Vos et al., 2019). The approach presented in this study based on a combination of a reliable but spatially sparse primary network and a secondary network with numerous

350 but also potentially biased and/or faulty observations. For all temporal resolutions, using the unfiltered secondary network data substantially increased the RMSE values. Hence, a direct application of the raw secondary data leads to a deterioration of the interpolation quality. Therefore, a filtering of data from the secondary network is essential. The applied filters in this study may be conservative by rejecting more stations than absolutely needed, but this is important in order to obtain robust results. The length of times series from the current secondary network will increase and subsequently more observations which

355 were currently discarded due to the uncertainty caused by the short time series may also become useful. Furthermore, it can also be expected that the number of secondary stations will continue to increase, thus one can expect further improvements of the quality of precipitation maps on all temporal aggregations. A comparison of the spatial characteristics of the time series of primary and secondary stations can be used to filter out stations with unreliable data. Observed precipitation values at the remaining secondary stations can be transformed to become unbiased using the observed percentiles and the distributions

360 at the primary stations as shown in Appendix A. This transformation does not require an independent ground truth of best estimation of precipitation at the secondary locations. A second spatial filter can be applied to find occasional faulty values at the used secondary stations. The cross validation results of a large number of different intense precipitation events show that with the presently available secondary stations after application of the two filters and the data transformation one can improve interpolation quality significantly. The improvement is the biggest for hourly time aggregations with a reduction of

365 the RMSE by 20 % , while for daily values the improvement is around 4 %. The spatial precipitation patterns are improved after corrections with the help of secondary network observations, especially for the short time scales. In particular, the spatial extent of precipitation fields are modified by the rainfall/no-rainfall information from the dense secondary network data. Finally, we want to highlight the differences of the approach used in this study compared to precipitation estimation using weather radar, since this type is often used when rainfall fields with a high temporal and spatial resolution are required.

370 – Secondary stations measure precipitation on the ground whereas radar measures reflectivity in higher elevations. Therefore, rain measured by radar may be advected by wind.

– Secondary stations measure point precipitation, radar measures spatial aggregations over large volumes.

– Radar measurements have problems with attenuation, secondary stations do not.

– Radar resolution is relatively uniform, secondary stations form an irregular network.

375 These differences are not listed here to compete between the two forms of additional information, but to point out that their different behaviour may be used for an effective combination. The method presented here requires a relatively dense primary

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

network. The use of secondary stations in regions with sparse reliable networks seems to be also possible but will require further research.

As precipitation uncertainty is possibly the most important factor for the uncertainty in rainfall/runoff modelling the im-
380 provement of precipitation interpolation achieved by this paper may contribute to a reduction of the uncertainty of hydrological modelling. Furthermore, the real time availability of the data of secondary networks may help to improve the quality of flood forecasts. Moreover, this study can help to improve gridded precipitation products for shorter time scales. Other procedures for the efficient use of secondary data may also be considered. Specifically, the interpolation of precipitation amounts using Quantile Kriging (Lebrenz and Bárdossy, 2019) may lead to better results. Due to the large number of zeros occurring for short
385 aggregation intervals however, this procedure has to be modified, for example by combining it with the approach developed by Bárdossy (2011). Traditional geostatistical interpolation methods use values measured at the same time interval only. However, for shorter temporal aggregations where advection can play a role, values measured at previous time steps may also be relevant. The shorter the time aggregation Δt, the more important the temporal aspect becomes. This aspect is not treated here in detail and requires additional research.

390 *Data availability.* The precipitation data was obtained from the Climate Data Center of the German Weather Service (https://opendata. dwd.de/climate_environment/CDC). The data from the Netamo stations was downloaded using the Netatmo API (https://dev.netatmo.com/ apidocumentation).

Appendix A: Transformation of Precipitation Amounts at Secondary Stations

This appendix illustrates the calculation for the transformation of precipitation amounts at secondary stations as described in
395 section 3.2. For simplicity consider 4 primary stations at the corners of a square and the secondary station being in the center of the square. This configuration ensures that the Ordinary Kriging weights of the primary station with respect to the secondary station are all equal to $1/4$ independently of the variogram.

The observed precipitation amounts at the stations are 3.1, 1.8, 3.0 and 2.1 mm for a selected event. The secondary station reported 1.7 mm rainfall. This corresponds to the 0.99 non-exceedence probability of precipitation for the specific secondary
400 station. The precipitation quantiles at the primary stations corresponding to the 0.99 probability are 3.2, 3.5, 3.1 and 3.0 mm. Interpolation of these values gives 3.2 mm which is the value assigned to the secondary station instead of the value of 1.7 mm. This value is greater than all the four primary observations. The reason for this is that the primary observations all correspond to lower percentiles. Note that the interpolation of the primary values corresponding to the event for the secondary observation location would be 2.5 mm. Figure A1 illustrates this example.

405 *Author contributions.* AB designed the study, AEH implemented the filtering algorithm for the sduy area. JS conducted the case studies in the chapter for the justification of the methods. All authors contributed to the writing, reviewing and editing of the manuscript.
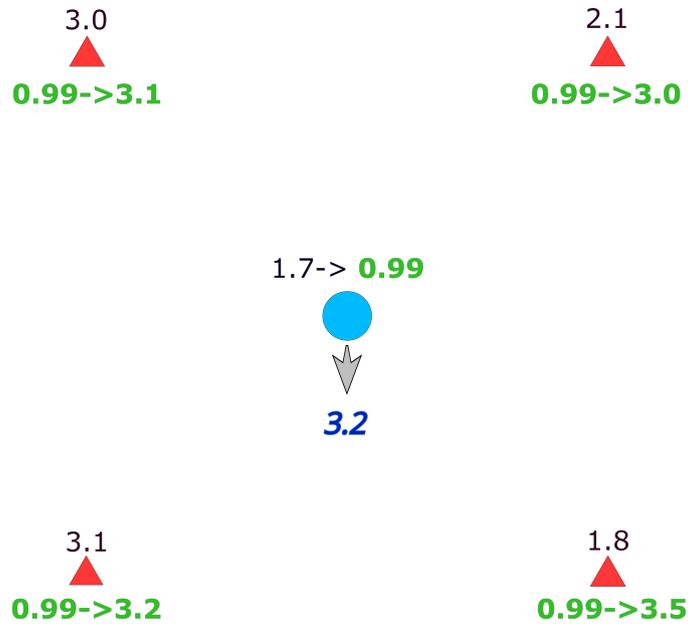
3.0
▲
0.99->3.1

2.1
▲
0.99->3.0

1.7-> **0.99**
●
⬇
3.2

3.1
▲
0.99->3.2

1.8
▲
0.99->3.5

Figure A1. Example for Transformation of precipitation amounts at a secondary station.

References

410   Ahmed, S. and de Marsily, G.: Comparison of geostatistical methods for estimating transmissivity using data transmissivity and specific capacity., Water Resources Research, 23, 1717–1737, 1987.

Bárdossy, A.: Interpolation of groundwater quality parameters with some values below the detection limit., Hydrology and Earth System Sciences, 15, 2763 – 2775, 2011.

Bárdossy, A. and Kundzewicz, Z.: Geostatistical methods for detection of outliers in groundwater quality spatial fields, Journal of Hydrology,
415   115, 343–359, https://doi.org/10.1016/0022-1694(90)90213-H, 1990.

Bárdossy, A. and Pegram, G.: Interpolation of precipitation under topographic influence at different time scales., Water Resources Research, 49, 4545–4565, 2013.

Berndt, C. and Haberlandt, U.: Spatial interpolation of climate variables in Northern Germany - Influence of temporal resolution and network density, Journal of Hydrology: Regional Studies, 15, 184 – 202, https://doi.org/https://doi.org/10.1016/j.ejrh.2018.02.002, http://www.
420   sciencedirect.com/science/article/pii/S2214581817303361, 2018.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T., Bastiaensen, J., De Bièvre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G., and Zhumanova, M.: Citizen science in hydrology and water resources: Opportunities for knowledge generation, ecosystem service management, and sustainable development, Frontiers in Earth Science, 2, https://doi.org/10.3389/feart.2014.00026, 2014.

425   de Vos, L., Leijnse, H., Overeem, A., and Uijlenhoet, R.: The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam, Hydrology and Earth System Sciences, 21, 765–777, https://doi.org/10.5194/hess-21-765-2017, 2017.

de Vos, L., Leijnse, H., Overeem, A., and Uijlenhoet, R.: Quality Control for Crowdsourced Personal Weather Stations to Enable Operational Rainfall Monitoring, Geophysical Research Letters, 46, 8820–8829, https://doi.org/10.1029/2019GL083731, 2019.

Delhomme, J.: Kriging in the hydrosciences, Advances in Water Resources, 1, 251–266, https://doi.org/10.1016/0309-1708(78)90039-8,
430   1978.

Deutscher Wetterdienst: Deutscher Klimaatlas, https://www.dwd.de/klimaatlas, 2020.

Giraldo, R., Delicado, P., and Mateu, J.: Ordinary kriging for function-valued spatial data, Environmental and Ecological Statistics, 18, 411–426, https://doi.org/10.1007/s10651-010-0143-y, 2011.

Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, Journal
435   of Hydrology, 332, 144–157, 2007.

Lebrenz, H. and Bárdossy, A.: Geostatistical interpolation by quantile kriging, Hydrology and Earth System Sciences, 23, 1633–1648, https://doi.org/10.5194/hess-23-1633-2019, 2019.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H. J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P., and Woods, R.: A rule based quality control method for hourly rainfall data and a 1-km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr, Journal
440   of Hydrology, 564, 930 – 943, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.07.034, http://www.sciencedirect.com/science/article/pii/S0022169418305390, 2018.

Mosthaf, T. and Bardossy, A.: Regionalizing nonparametric models of precipitation amounts on different temporal scales, Hydrology and Earth System Sciences, 21, 2463–2481, https://doi.org/10.5194/hess-21-2463-2017, 2017.

Muller, C., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., and Leigh, R.: Crowdsourcing for
445   climate and atmospheric sciences: current status and future potential, International Journal of Climatology, 35, 3185–3203,
       https://doi.org/10.1002/joc.4210, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.4210, 2015.

Napoly, A., Grassmann, T., Meier, F., and Fenner, D.: Development and Application of a Statistically-Based Quality Control for Crowd-
       sourced Air Temperature Data, Frontiers in Earth Science, 6, 118, https://doi.org/10.3389/feart.2018.00118, https://www.frontiersin.org/
       article/10.3389/feart.2018.00118, 2018.

450   Villarini, G. and Krajewski, W. F.: Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall,
       Surveys in Geophysics, 31, 107–129, https://doi.org/10.1007/s10712-009-9079-x, https://doi.org/10.1007/s10712-009-9079-x, 2010.

World Meteorological Organization: Guide to meteorological instruments and methods of observation., World Meteorological Organization,
       Geneva, Switzerland, oCLC: 288915903, 2008.

Yan, J. and Bárdossy, A.: Short time precipitation estimation using weather radar and surface observations: With rainfall displacement
455   information integrated in a stochastic manner, Journal of Hydrology, 574, 672–682, https://doi.org/10.1016/j.jhydrol.2019.04.061, 2019.