Improving the spatial monitoring of rainfalls for flood assessment

Franck Schoefs
Professor, Nantes Université, École Centrale Nantes, CNRS, GeM, UMR 6183, IUML, FR 3473, F-44000 Nantes, France

Chengjun Cao
Graduate Student, Nantes Université, École Centrale Nantes, CNRS, GeM, UMR 6183, F-44000 Nantes, France

Nicholas Gooding
PhD, Nantes Université, École Centrale Nantes, CNRS, GeM, UMR 6183, IUML, FR 3473, F-44000 Nantes, France

ABSTRACT: Flood assessment for early warning systems rely on the confidence on data. Rainfalls are usually estimated by radar with a low accuracy but a fine spatial mesh and rain gauges with a high accuracy but a low density. Adding rain gauges could be of added value for assessing the flood. First, a method for analyzing the role of added rain gauges on the reduction of uncertainties of spatial average rainfall is proposed. Kriging method with external drift was applied for interpolation estimation. The uncertainty of the interpolation is also quantified. Also, the distribution of the weights of the rain gauges was analyzed simultaneously using Thiessen polygon. When comparing rainfall from rain gauges with different amounts of additions, the impact of change is evaluated using the indicator of mean area rainfall. In this study, we analyze the distribution of rain gauges in the Landro basin and the corresponding changes in precipitation. The spatial distribution of rain gauges within the basin was optimized and improved. Finally, the rainfall data was applied to the Iber software and the changes in river flow were obtained by means of the corresponding hydrological model. The evolution of the probability of false alarm and the probability of good assessment are also estimated. This work is carried out in the AAFLOOD Atlantic Area Interreg project.

1. INTRODUCTION

Floods are the climate change phenomenon that affects most people in the Atlantic area, causing run-off in urban areas and overflowing rivers through increasingly intense rainfall. In the last 20 years, more than 1,500 lives and 52,000,000 euros have been lost in the EU. This is a problem for national and regional governments, but also for local authorities and communities, water companies, scientific and non-governmental organizations. AA- FLOODS aims to unite all these actors to create and test new tools, plans and regulations to enhance flood risk management and response.

As one of the most important factors influencing flooding, the regional rainfall will have a direct impact on the frequency of flooding in the area. Therefore, in order to monitor floods, the effective detection of the amount of rainfall in the respective area will be one of the key methods.

In practice, rain gauges have always been one of the main sources of precipitation data. In more developed and technologically advanced countries, satellites and radar are also playing an increasingly important role in acquiring precipitation data, especially in providing real-time data.

In the 1990s, statistical methods of spatial information were introduced into the algorithms for the joint estimation of precipitation by radar and rain gauges, which are still being developed and improved. Compared to other methods, which focus on the random errors present in the measured data, the estimation process takes better account of the temporal and spatial correlation between the data and has a higher estimation accuracy.

The Kriging method in spatial information statistics is particularly well suited to dealing with data that include both sparse, high-precision measurements (rain gauges) and dense, error-rich measurements (radar, satellites). The Kriging method was first used mainly for climate analysis, such as annual precipitation distribution (Dingman et al. 1988) (Hevesi et al. 1992). Later, the Kriging method was used to merge rain gauges with radar data for optimal unbiased estimation (Sun et al. 2000). In addition to this, with the widespread use of meteorological satellites, rainfall estimation algorithms based on satellite imagery are used and have proven to be highly effective. As a result, for countries or regions where the number of rain gauges is insufficient and the distribution is sparse, (Grimes et al. 1999) use kriging method with external drift (KED) to combine satellite data with rain gauge data to obtain regional rainfall estimates. (Delrieu et al. 2014) also used the KED method to merge rainfall gauge data with radar data to quantify the accuracy of rainfall estimates. In turn, (Pardo-Igúzquiza 1998) provided a comprehensive comparison of the ordinary kriging, cokriging, kriging with external drift and the classical Thiessen method in terms of estimating rainfall. In the end they obtained that the kriging method with external drift gives the most coherent results according to cross-validation.

When studying rainfall, it is common to use mean area rainfall (MAR) as a description. The most frequently used method for studying the calculation of mean area rainfall is the Thiessen polygon method. Based on the distribution of the polygons an estimate can be made of the weight of each component.

In this paper, the area chosen for study is the Landro Basin in Galicia. Five historical storm events are focused and their rainfall data are obtained from rain gauges, satellite and radar records from the area. Rainfall data from rain gauges and satellites are merged by using the kriging with external drift method and spatially interpolated to obtain a continuous rainfall data surface from discrete rainfall data points. Simulations are also carried out by adding different numbers of new rain gauges to the study area to obtain a comparison of the simulated rainfall data under different numbers of rain gauges with the real rainfall data to obtain the best distribution of rain gauges.

The simulated rainfall data will then be transferred to the hydrological model of the Landro basin through the Iber software to calculate the river flow in the basin for each storm event.

2. CASE STUDY

2.1. Study area and rainfall data

In the present study, the subject is the Landro Basin in Galicia. The Landro basin location and sub-basin discretization are the richly colored parts of Figure 1. The outlet is on the northern border of the basin.

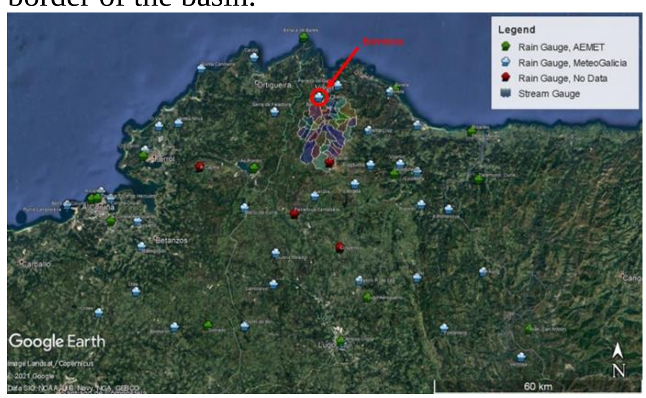

Figure 1: Map of Landro Basin and around area.

There are three sources of rainfall data used in this paper, rain gauges, satellites and radar. The distribution of rain gauges in the Landro Basin is shown in the Figure 1. The rain gauge data used in this study is from MeteoGalicia, which is labelled in white. The rain gauges labelled in red, shown as 'no data' in the figure are those that have been added in recent years and have not been switched on to collect data and therefore have no available data. As there is only one available rain gauge within the basin, the name of the station is Borreiros.

Expect the Borreiros station, there are many other rain gauges distributed around the basin. Therefore, in order to obtain more rain gauge data, we extended our data collection to the vicinity of the basin. However, given that rain gauges far from the basin may contribute less to rainfall estimates within the basin, a range of study areas were considered to assess when rain gauges may not be useful for flood modelling. In total, five study 'areas' were considered. These areas are shown as Figure 2. The rain gauges data is recorded as 10-minute intervals. The sizes and number of stations in each zone are presented in Table 1.

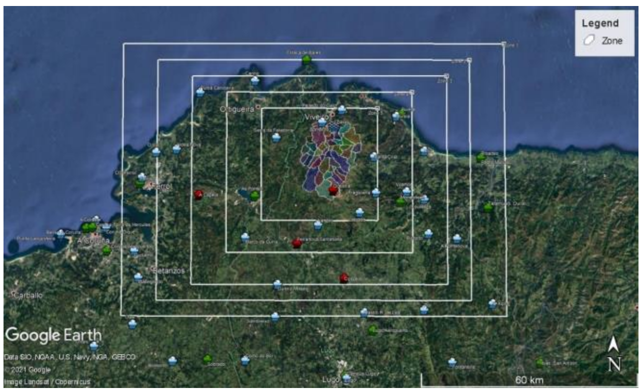

Figure 2: Different zones for Landro Basin. The innermost circle is zone 5, and the outermost circle is zone 1.

Table 1: Distribution information for rain gauges in each zone.

Zone	*Station Numbers*	*Area(km^2)*	*Density(km2/STA)*
1	*25*	*8,050*	*252*
2	*18*	*5,811*	*265*
3	*15*	*4,103*	*228*
4	*10*	*2,295*	*177*
5	*6*	*1,004*	*168*

Several sources of satellite precipitation data are available in the study area, with different historical records and temporal and spatial resolutions. Satellite precipitation data from the PERSIAN-CSS system are applied in this study, recording rainfall data from 2003 to the present, from 43.80°N to 43.16°N and 8.32°W to 7.00°W, with a resolution of 0.04° per hour.

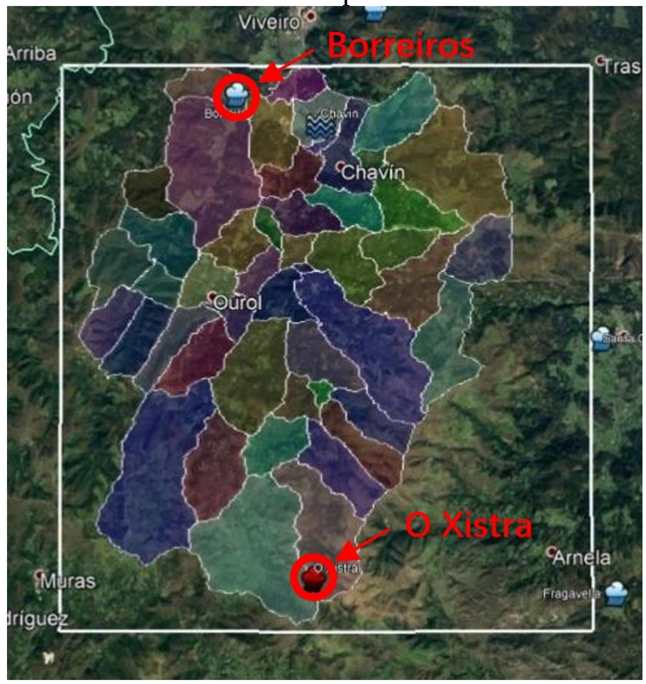

Figure 3: The boundaries of MeteoMapa.

Table 2: Selected storm events.

Storm	*Start*	*End*	*Hourly Timesteps*
1	*12-Nov-2019 00:00:00*	*13-Nov-2019 10:00:00*	*35*
2	*13-Nov-2019 16:00:00*	*16-Nov-2019 07:00:00*	*64*
3	*21-Nov-2019 23:00:00*	*23-Nov-2019 16:00:00*	*42*
4	*16-Jan-2020 08:00:00*	*17-Jan-2020 10:00:00*	*27*
5	*11-Aug-2020 12:00:00*	*13-Aug-2020 02:00:00*	*39*

The radar data come from a WRM200 model radar from Vaisala. In this study, the rainfall data from the radar and the rain gauge data will be conditionally combined, called 'MeteoMapa'. This data will be considered as the standard value of rainfall data for the current area, i.e. defined as the true rainfall. It is used to compare the uncertainty in the estimated rainfall after interpolation and as a reference.

The boundaries of the MeteoMapas grid around the Landro basin are shown in Figure 3.

2.2. Storm events

The larger 5 historical storm events are selected for this study which are identified by first looking at the daily trends in the MeteoMapas data and then identifying the start and end points of the storms by having zero or near zero rainfall across the basin. The duration of each storm event is shown in table 2.

3. METHODOLOGY

3.1. Spatial Interpolation

3.1.1. Interpolation by Kriging method with external drift

In this study, the Kriging with external drift method is used to merge rainfall data from rain gauges and satellites for Storm Events 1 to 5 and Zones 1 to 5, respectively, to obtain interpolated rainfall conditions. Figure 4 shows one of these cases.

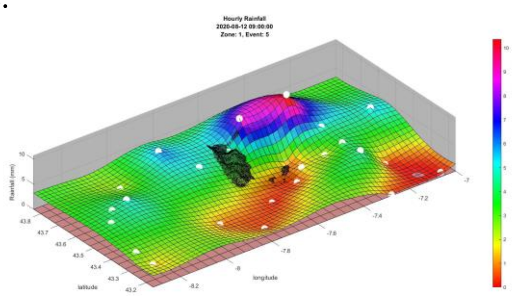

Figure 4: Interpolated rainfall estimates for Storm Event 5 within Zone 1 at 12-Aug-2020 09:00:00.

In figure 4, the z-axis is the rainfall in mm at each location at 19 h. The white dot markers in the figure represent the rainfall data from the rain gauge. And the highly colorful grid surface represents the estimated rainfall after interpolation. This surface takes into account the rain gauge data and is shown as having passed through each rain gauge point. Besides, the highly dense black grid in the figure represents the MeteoMapa data. It can be observed that there is a clear difference between the estimated and real values but also a similar trend.

3.1.2. Quantifying estimation uncertainty

After the interpolation has been performed, the uncertainty in the spatial interpolation needs to be quantified by cross-validation. Since in spatial interpolation the rainfall data obtained is estimated, the main purpose of this step is to explore the uncertainty in the estimated rainfall results. The images for the cross-validation are again plotted through MATLAB.

For the uncertainty, figures 5 and 6 show two cases of the results, that as the study area increases and more rain gauges become available, the standard deviation of the rainfall gradually decreases, and the uncertainty decreases to some extent.

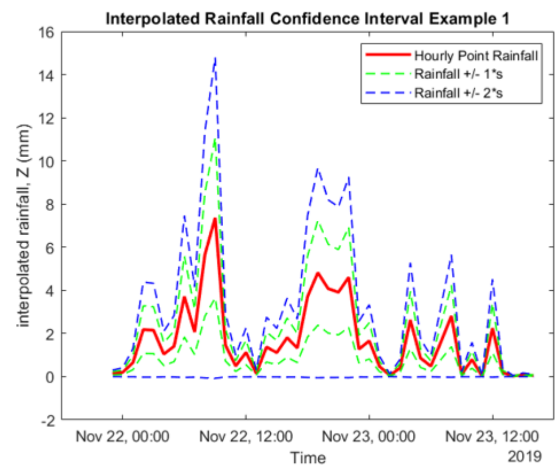

Figure 5: Estimation uncertainty for storm 3 zone 1.

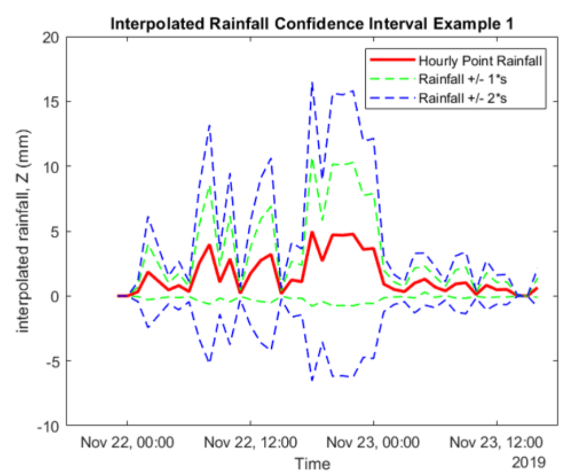

Figure 6: Estimation uncertainty for storm 3 zone 4.

3.2. Quantifying the impact of changing the rain gauges number

With spatial interpolation it is possible to obtain rainfall estimates for a given storm event at a specific time within a defined region. However, such results only lead to haphazard visual data of the rainfall and do not allow for direct comparisons between different situations. To be able to understand the variation in rainfall estimates for different rainfall gauge amounts, then, an indicator is needed to characterize the rainfall situation. (Xu et al. 2013) and (Zeng et al. 2018) analyzed the effect of the rain gauge density and distribution on the uncertainty of block hydrological models with different catchment sizes. They apply the mean areal rainfall (MAR) as an indicator to describe rainfall conditions, which is used to draw conclusions.

Therefore, mean area rainfall will also be applied here as an impact indicator. In this study, rainfall z(x,t) in the study area is a variable with respect to location and time. MAR is defined as the average value of estimated rainfall at all points in the area at each time step. The formula is shown below:

$$MAR = \frac{1}{T}\frac{1}{A}\int_{T_0}^{T_e}\int_{0}^{Z_e} z(x,T)dxdT \qquad (1)$$

where T is storm event time duration, A is study region area, T_0 is storm events start time, T_e is storm events end time, Z_e is rainfall in the last position, x is location in the study area.

3.3. Location of the added rain gauges

Once the indicators to evaluate rainfall are available, the following step is to optimize the distribution of rain gauges. In the process of optimization, new rain gauges will have to be added at locations where no rain gauge data is available. For simulating the additional rain gauges, it is necessary to extract the data for the corresponding locations from the MeteoMapa data, which represents the real rainfall. As the existing MeteoMapa data range only covers the Landro basin, even less than area 5, and does not include the surrounding areas. Therefore the scope for improving the spatial distribution of rain gauges is limited to the MeteoMapa coverage. This is shown in the figure 3.

As shown in the map, two rain gauge stations are marked in the study area, Borreiros is the original rain gauge station and provides rainfall data for all the storm events. The other rain gauge station, O Xistra, is a recent addition. It is still in the data collection stage and does not have rainfall data for the events. Therefore, O Xistra will be treated as one of the new rain gauges whose location has already been determined.

In addition to O Xistra, three additional rain gauges will be added separately. During this process, it will be necessary to control that each rain gauge has approximately equal weights to ensure that each gauge has the same range of influence to guarantee maximum efficiency. In calculating the weight of the area affected by each rain gauge, the Thiessen polygon is used to calculate the area that each gauge can affect and then compared to the whole area to obtain the weight.

O Xistra already is the first rain gauge to be added in the region. The affected area for each gauges is shown in the top left of figure 7.

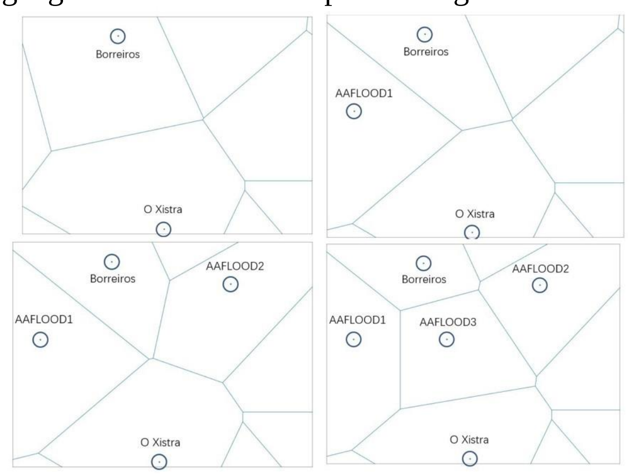

Figure 7: Weight distribution for adding rain gauge. Top left shows the addition of the first rain gauge. Top right shows the addition of the second rain gauge. Bottom left shows the third rain gauge being added. Bottom right shows the fourth rain gauge being added.

In this situation, the overall boundaries represent the area under study. The dot at the top represents the already existing rain gauge Borreiros and the dot at the bottom represents the additional rain gauge O Xistra. The area of the polygon encircling each gauge indicates the range that it affects, while the other polygons represent the impact of the other surrounding rain gauges that are not within the study area.

The weight of this additional rainfall gauge is shown in the Table 3.

Table 3: Weights of additional rainfall gauges.

Name of added gauge	Weight
O Xistra	31%

The second to fourth rain gauges are named AAFLOOD1-AAFLOOD3. Figure 7 shows the distribution of weights for each case separately.

The weights of the added rain gauges are shown in Table 4-6.

Table 4: Weights of additional rainfall gauges.

Name of added gauge	Weight
O Xistra	23%
AAFLOOD1	22%

Table 5: Weights of additional rainfall gauges.

Name of added gauge	Weight
O Xistra	22%
AAFLOOD1	22%
AAFLOOD2	22%

Table 6: Weights of additional rainfall gauges.

Name of added gauge	Weight
O Xistra	17%
AAFLOOD1	18%
AAFLOOD2	17%
AAFLOOD3	18%

4. RESULTS

After the location of the additional rain gauges is determined, the mean areal rainfall is calculated by spatial interpolation for different numbers of rain gauges. For reducing the non-essential effects, some treatment of each storm event is required. A small part of the beginning and end of the storm event is removed to a reasonable extent.

To better visualize the change in mean areal rainfall each time a new rain gauge is added, the mean areal rainfall of two adjacent cases at same time point is subtracted to obtain the amount of change between cases.

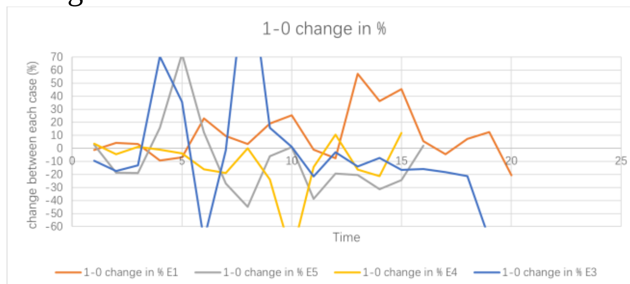

Figure 8: Mean areal rainfall change between with added one gauge and without added gauge.

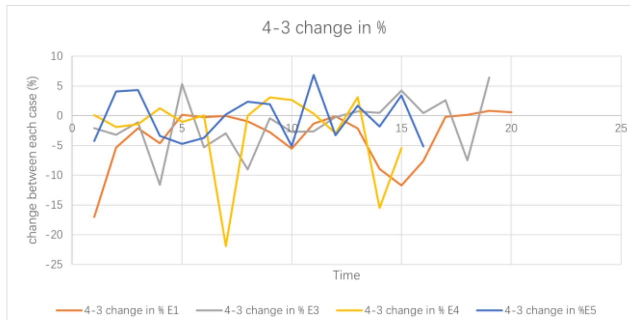

Figure 9: Mean areal rainfall change between with added four gauges and added three gauges.

And based on the characteristics of the storm events themselves, the five events can be divided into two categories. The first category contains events 1, 3, 4 and 5, which are characterized by a single peak of rainfall amount in the middle of the rainfall period. The other category is Event 2, which is characterized by multiple rainfall peaks during the rainfall process and a more even distribution of rainfall throughout the rainfall process. Event 2 has a much longer effective rainfall duration than the storm event in the first category. Therefore, event 2 cannot be placed in the image at the same time as the other events. The corresponding images that can be obtained are as follows: figures 8 and 9.

For each storm event, the magnitude of the change decreases as the number of rain gauges increases. Adding the first rain gauge results in a range of changes from +70% to -60%. As for the second added rain gauge case, there is a change in the range of +40% to -60%. Adding the third rain gauge causes a change in the range of +15% to -25%. Lastly, adding the fourth rain gauge causes a change in the range of +7% to -20%. The same change trend also occurred to storm event 2.

To take this a step further and visualize the changes in mean area rainfall, a box plot can be used to express the situation as figure 10.

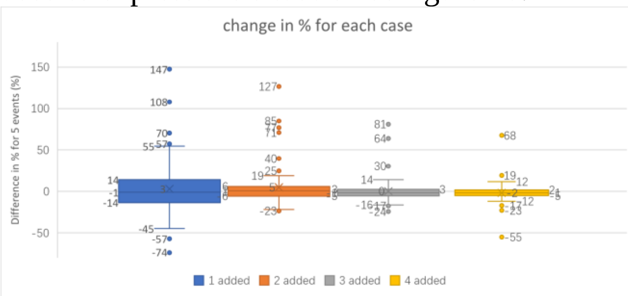

Figure 10: Box plot for mean areal rainfall change at all events with all gauges added case.

Figure 10 contains all of the storm events and clearly illustrate that for all the events adding gauges reduce the mean value of the changes. The scatter means that overestimate at a given time can be significantly (147% for highest point of the box plot by adding 1 gauges). And the underestimation is lower for each case.

For further analysis, the mean value of MAR for each storm event is calculated.

$$\overline{MAR} = \frac{1}{T}\int_{T_0}^{T_e}\int_{0}^{Z_e} MARdT \qquad (2)$$

where $\overline{MAR}$ is average of MAR.

This indicator gives the trend in MAR with different numbers of rain gauges. The results are shown in Figure 11. MAR tends to decrease as the number of rain gauges increases in events 3, 4 and 5, while the average of all MARs tends to converge as the number of rain gauges increases. It is reasonable to infer that if the number of rain gauges continues to increase, there will be a monotonic trend in the average values for all storm events.

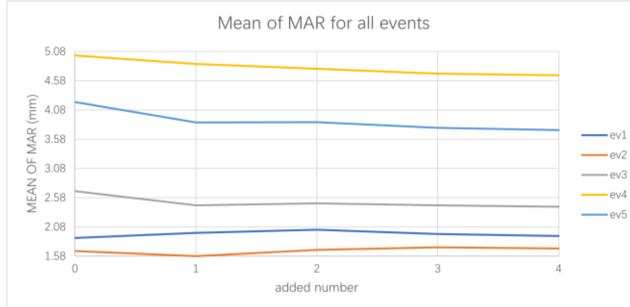

Figure 11: Mean value of MAR for all events in all cases.

One of the aims of the AA-FLOOD project is the prevention and monitoring of floods. Iber is a two-dimensional hydrodynamical model for the simulation of free surface flow, morpho dynamics, transport processes and habitat in rivers and estuaries. Therefore, after finishing the rainfall analysis, the rainfall data is imported into the Iber software to obtain the river runoff for the corresponding rainfall scenario.

The basin is discretized in the Iber and the rivers and tributaries in the basin are prominently marked. By importing the rainfall data into the model, a graphical representation of the river flow at the outlet can be obtained, using storm event 1 as an example as shown in figure 12. It could be observed that the addition of different numbers of rain gauges also has an effect on the estimation of river flows. By taking the average of the flows for each event, the graph could be obtained as figure 13. The trend in average river flows is broadly consistent with the trend in the average of MAR.

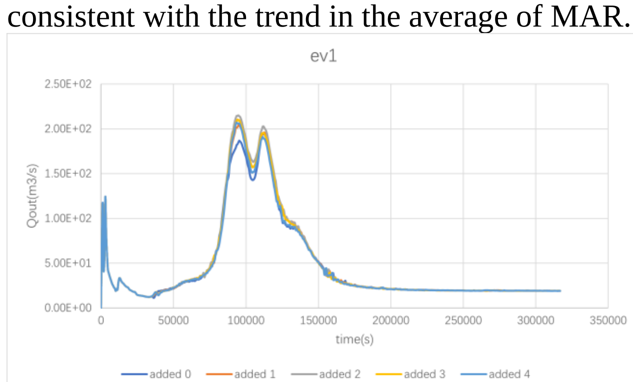

Figure 12: River flows for different numbers of rain gauge additions in event 1.

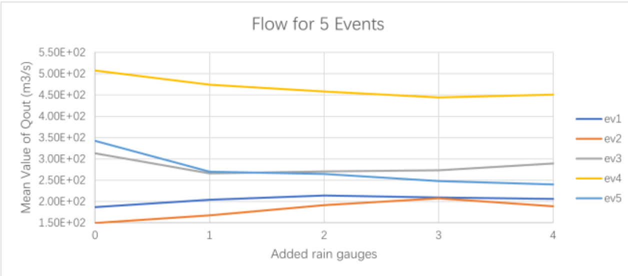

Figure 13: Mean value of river flows for different number of rain gauge additions over all events.

5. CONCLUSIONS

This paper examines in some detail the optimization of rain gauge distributions in the Landro basin of Galicia. In terms of spatial interpolation, kriging with external drift is flexibly applied to combine rainfall data measured by rain gauges with satellite data to obtain estimated rainfall amounts. The uncertainty of the interpolation method is also quantified. Afterwards, the location of additional rain gauges within the study area is determined. In this process, the Thiessen polygon method is applied to calculate the weights of the additional rain gauges.

Once the location of the additional rain gauges is determined, the rainfall for each situation needs to be analyzed. At this stage, the mean area rainfall is chosen as an indicator to quantify the impact of the additional rain gauges. Based on the analysis of mean area rainfall, the change in mean area rainfall is compared for different numbers of additional rain gauges. It could conclude that the additional rain gauges effectively improve the accuracy of rainfall measurements in the Landro Basin area. As the number of additional rain gauges increased, the accuracy of the measurements gradually improves and the over- or underestimation of rainfall decreases.

Finally it is concluded that when the number of additional rain gauges is sufficiently high, the trend of rainfall in the average will tend to be monotonic.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the AA-FLLOD Interreg AAproject (2018-2023).

6. REFERENCES

Delrieu, G., A. Wijbrans, B. Boudevillain, D. Faure, L. Bonnifait, and P.-E. Kirstetter. 2014. “Geostatistical radar–raingauge merging: A novel method for the quantification of rain estimation accuracy.” *Adv. Water Resour.*, 71: 110–124.

Dingman, S. Lawrence., D. M. Seely-Reynolds, and R. C. Reynolds III. 1988. “Application of Kriging to Estimating Mean Annual Precipitation in a Region of Orographic Influence1.” *JAWRA J. Am. Water Resour. Assoc.*, 24 (2): 329–339.

Grimes, D. I. F., E. Pardo-Igúzquiza, and R. Bonifacio. 1999. “Optimal areal rainfall estimation using raingauges and satellite data.” *J. Hydrol.*, 222 (1): 93–108.

Hevesi, J. A., A. L. Flint, and J. D. Istok. 1992. “Precipitation Estimation in Mountainous Terrain Using Multivariate Geostatistics. Part II: Isohyetal Maps.” *J. Appl. Meteorol. Climatol.*, 31 (7): 677–688. American Meteorological Society.

Pardo-Igúzquiza, E. 1998. “Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography.” *Int. J. Climatol.*, 18 (9): 1031–1047.

Sun, X., R. G. Mein, T. D. Keenan, and J. F. Elliott. 2000. “Flood estimation using radar and raingauge data.” *J. Hydrol.*, 239 (1): 4–18.

Xu, H., C.-Y. Xu, H. Chen, Z. Zhang, and L. Li. 2013. “Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China.” *J. Hydrol.*, 505: 1–12.

Zeng, Q., H. Chen, C.-Y. Xu, M.-X. Jie, J. Chen, S.-L. Guo, and J. Liu. 2018. “The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach.” *J. Hydrol.*, 563: 106–122.