



**POLITECNICO
DI TORINO**

POLITECNICO DI TORINO

Master Degree course in Environmental and Land Engineering

Master Degree Thesis

Increasing Rainfall Observation Density in Urban Areas through the Integration of Crowdsourced Personal Weather Stations

Supervisors

Dr. Paola MAZZOGLIO

Dr. Benedetta MOCCIA

Candidate

Elaheh GHAFFARIPOUR

ACADEMIC YEAR 2024-2025

Acknowledgements

I would like to express my sincere gratitude to **Dr. Paola Mazzoglio** for her constant guidance, insightful feedback, and encouragement throughout the development of this work. I am also grateful to **Dr. Benedetta Moccia** from the **Hydrology Group at La Sapienza University of Rome** for her collaboration and constructive discussions.

I wish to express my deepest appreciation to my parents, whose unwavering support, love, and belief in me have made this achievement possible. They have always encouraged me to pursue my dreams and overcome every challenge with determination. My heartfelt thanks go to my partner for being by my side through every step of this journey. His constant support and faith in me have been a true source of strength and light, even during the most difficult times. Finally, I am grateful to my friends for their kindness, patience, and companionship throughout this period.

Abstract

Accurate rainfall monitoring is fundamental for hydrological modeling and flood-risk management, yet the spatial density of official rain gauge networks remains insufficient to capture short-range variability, especially in urban areas. The growing availability of crowdsourced Personal Weather Stations (PWSs) provides an opportunity to increase the density of rainfall observations at minimal cost. This study evaluates the reliability of citizen-owned PWSs across eleven Italian cities and investigates their integration with Official Weather Stations (OWSs) to enhance rainfall observation density in Rome and Turin, validated against radar-derived accumulations.

Rainfall data from Netatmo PWSs (5-minute resolution) and regional OWS networks were collected and aggregated to hourly totals to ensure temporal consistency between datasets. After applying a 10% missing-data threshold, PWS–OWS pairs were compared in terms of annual rainfall totals, zero-rainfall frequencies, and short-duration annual maxima. The analysis revealed that while PWSs reproduce rainfall occurrence with high temporal consistency, they tend to underestimate rainfall volumes, particularly during high-intensity events.

Two flash-flood events—Rome (8 June 2021) and Turin (22 June 2021)—were analysed to assess whether integrating selected PWSs (within $\pm 10\%$ and $\pm 20\%$ of the OWS reference) improves spatial rainfall representation. Interpolation with Inverse Distance Weighting (IDW) and Ordinary Kriging (OK) was compared against radar-derived accumulations. The integration of raw PWSs led to varying outcomes: in Turin, the denser network improved the spatial agreement with radar, while in Rome, where OWS coverage was already high, the added PWSs resulted in a marginal reduction in agreement. These contrasting results highlight that the effectiveness of PWS integration is case-dependent, influenced by network density, rainfall characteristics, and local station geometry.

Finally, the open-source quality control algorithm (PWS-pyQC) was applied to the PWS data, and the same interpolation–validation workflow was repeated. The quality control effectively removed inconsistent records and enhanced internal data reliability. Overall, this work demonstrates that crowdsourced PWSs, when properly filtered, can complement official networks to increase rainfall observation density and support high-resolution urban rainfall monitoring.

Contents

1	Introduction	5
1.1	Established rainfall measurement systems	6
1.2	Crowdsourced rainfall observations through personal weather stations . .	6
1.3	Research objectives	8
1.4	Research questions	8
2	Data and study area	11
2.1	Rain gauges	11
2.2	Personal weather stations	12
2.3	Weather radar	13
2.4	Study areas	13
3	Evaluation of raw PWS data	15
3.1	Methodology	15
3.2	Comparison between PWSs and OWSs	16
3.2.1	National comparison of annual rainfall totals	16
3.2.2	National comparison of zero-rainfall frequencies	18
3.2.3	National comparison of 1-hour annual maximum rainfall	19
3.2.4	Case-study analysis: Rome and Turin	20
3.3	Discussion and interpretation	23
4	Rainfall interpolation and radar validation with raw PWS data	27
4.1	Methodology	27
4.2	Spatial comparison with radar data	29
4.2.1	Turin — 22 June 2021	29
4.2.2	Rome — 8 June 2021	32
4.3	Discussion and interpretation	35
5	Spatial interpolation and radar validation with quality-controlled PWS data	37
5.1	Methodology	38
5.2	Spatial comparison with radar data	40
5.2.1	Turin — 22 June 2021	41
5.2.2	Rome — 8 June 2021	43
5.3	Discussion and interpretation	45

6	Discussion, conclusions, and future perspectives	47
A	Additional figures and sensitivity tests	49
A.1	Example of rejected OK interpolation from OWS-only configuration . . .	49
	Bibliography	51

Chapter 1

Introduction

Precipitation is a key component of the global energy and water cycles and exerts a strong influence on the Earth’s climate system [1]. Among its forms, rainfall plays a fundamental role in both terrestrial and aquatic ecosystems, providing essential goods and services that support human well-being [2].

Rising global temperatures are altering the hydrological cycle [3], with one of the most pressing concerns being the intensification of extreme precipitation events [4]. These climatic changes, compounded by population growth, urban expansion into flood-prone regions, and insufficient progress in flood-control strategies, are expected to intensify flood risks and their social, environmental, and economic impacts [5]. In this context, enhancing our ability to understand, model, and predict rainfall, particularly extreme events, remains a critical priority for both scientific research and policymaking [6].

Rainfall exhibits high spatial and temporal variability, which necessitates dense observational networks to adequately capture. While rain gauges remain the cornerstone of precipitation measurement and are widely regarded as the de facto standard for hydrometeorological observations, their deployment is often constrained by financial and logistical limitations, leading to sparse spatial coverage and leaving large regions under-observed [1, 7]. Furthermore, in some countries, data accessibility is further hampered by decentralization and by the reluctance of certain agencies to make climate records publicly available [8].

Despite advances in remote sensing, the current accuracy of ground-based weather radars, the spatial coverage of rain gauges, and the spatiotemporal resolution of satellite products remain insufficient for critical applications such as short-term weather forecasting, flood prediction, and water balance monitoring. To address these limitations, so-called “opportunistic sensors” have emerged as complementary sources of precipitation information. These include technologies such as commercial microwave links (CMLs) [9, 10] and crowdsourced personal weather stations (PWSs) [11]. When combined with standardized metadata in line with WIGOS (WMO Integrated Global Observing System) guidelines, these novel data sources offer the potential to fill observational gaps and provide fit-for-purpose information for hydrometeorological applications.

1.1 Established rainfall measurement systems

Rain gauges (from simple accumulation vessels to tipping-bucket, weighing and optical gauges) give physically direct, ground-truth measurements of precipitation intensity and depth, and form the basis of most climatological and hydrological records [12]. However, gauge networks are unevenly distributed: while many land areas host networks, only a minority of regions reach gauge densities sufficient to capture the strong spatial variability of rainfall. While improvements in temporal resolution have enhanced the ability of the rain gauges to capture short-duration rainfall events, gauge networks face significant challenges. Long-term continuous records are relatively scarce, and sub-daily datasets are even less common [13].

Ground-based weather radars complement gauges by providing spatially continuous fields of precipitation reflectivity from which rainfall rates can be derived. Radars excel at representing the structure and dynamics of precipitation systems and resolving mesoscale and convective features that gauges miss. Nonetheless, radar retrievals are sensitive to the drop-size distribution (DSD), beam geometry (beam elevation and “lifting” with range), attenuation and ground clutter; these factors can create range-dependent biases and blind zones [14, 15]. Dual-polarization radars improved hydrometeor classification and reduced some uncertainties, but radar networks remain spatially limited and often coincide with the same well-observed regions that host dense gauge networks, leaving many areas under-sampled [12].

Satellites cannot provide continuous sampling of the Earth’s surface, yet compared to in situ networks their coverage is far more extensive and spatially consistent. At a large scale, satellite-based precipitation products deliver near-global observations and synoptic perspectives that are particularly valuable in regions with sparse or absent ground-based networks [12]. These products rely on a variety of meteorological sensors and retrieval algorithms, producing datasets that can be tailored to user needs across different spatial and temporal scales. While satellites help to overcome many of the spatial limitations of gauges and radars, they are also prone to biases because they measure precipitation processes indirectly from space. Such biases may arise from diurnal sampling gaps, instrument calibration, retrieval algorithm design, or atypical surface and atmospheric conditions that the algorithms fail to properly account for. Recent advances in instrumentation and retrieval methodologies have, however, reduced some of these systematic errors, offering improved accuracy compared to earlier generations of satellite products [16].

1.2 Crowdsourced rainfall observations through personal weather stations

Crowdsourcing has traditionally been understood as obtaining data by engaging large groups of people, but recent innovations extend this definition to include information collected from public sensors connected via the Internet [17]. This evolution means that citizens are no longer only consumers of information but also active producers, contributing environmental observations through devices such as smartphone sensors, amateur weather stations, citizen science initiatives, and even social media platforms [18].

Personal weather stations have recently emerged as a promising crowdsourced solution to enhance rainfall and meteorological observations, particularly in urban areas. These are low-cost, citizen-installed sensors, often part of home automation systems, that automatically upload local weather data in real time to online platforms, creating dense decentralized networks of observations [19, 20]. The affordability and ease of deployment of PWSs have led to tens of thousands of stations being distributed worldwide, vastly increasing the spatial density of measurements available at minimal cost [17, 20]. In addition, the near-real-time transmission of observations through cloud-based platforms supports timely monitoring of intense rainfall events and enhances their potential for nowcasting and early-warning systems [21]. Urban and hydrological studies are beginning to capitalize on this high-density coverage. For example, crowdsourced PWS data have been used to map fine-scale rainfall patterns and fill observational gaps in cities, improving flood monitoring and modeling in fast-responding catchments [22, 23]. There is also growing integration of PWS observations into operational systems: recent studies demonstrate that assimilating or merging PWS data with conventional networks, weather radar, and numerical models can improve high-resolution rainfall estimation and forecasting [24–26]. The increasing number of case studies reporting successful applications of PWS data, from real-time urban flood warning to model validation, underscores the considerable potential of these crowdsourced networks for hydrometeorology and climatology. Indeed, national meteorological services have begun to explore PWS datasets as a valuable supplement to official observations, given their ability to provide hyperlocal information in near real-time [7, 20].

Despite their promise, PWS networks face a number of technical and observational limitations that must be addressed before they can fully transition into reliable scientific data sources. A key challenge is data quality. Unlike official weather stations, PWSs are installed by amateurs without strict siting guidelines, which often leads to suboptimal placement (e.g. too close to buildings, on balconies, or under partial shelter) and inconsistent maintenance [27]. These issues can introduce biases or errors. For instance, a rain gauge mounted on an uneven surface or sheltered from wind will record less precipitation than an optimally placed gauge. Hardware and design limitations of the low-cost sensors further affect accuracy. Many PWSs use tipping-bucket rain gauges with small collectors, which tend to underestimate rainfall during intense downpours due to overflow or mechanical limits [28, 29]. Empirical evaluations have shown that while PWS rainfall totals agree reasonably well with reference gauges for moderate rains, they markedly underestimate high-intensity events; beyond roughly 30 mm/h rainfall rates, PWS measurements diverge and fail to capture peak storm accumulations [30, 31]. In one study, clusters of citizen stations under-recorded heavy rainfall by 20–40 % when no corrections were applied [30]. Moreover, the lack of uniform sensor models, different PWS brands and designs in the network, leads to heterogeneous performance and bias characteristics from one station to another [28]. Social and spatial biases in where PWSs are deployed also present a concern: recent analyses in U.S. and Canadian cities found PWSs tend to be over-represented in higher-income or flood-prone neighborhoods, meaning crowdsourced data coverage is not truly uniform across the urban landscape [22, 32]. All these factors contribute to larger uncertainty and potential bias in PWS datasets compared to official

measurements, reinforcing the need for careful data validation before use [20].

Efforts are underway to improve the reliability of crowdsourced PWS data through automated quality control (QC) methods. These algorithms perform statistical and spatio-temporal consistency checks. For example, comparing each station’s readings with nearby stations or with physical plausibility ranges to detect outliers and correct errors [33, 34]. More advanced approaches calibrate crowdsourced observations against authoritative sources (such as rain gauges or radar) to adjust biases and improve accuracy [25, 30]. Comparative studies have found that different QC techniques each have strengths and weaknesses depending on station density and environmental factors, underscoring the need for robust, standardized QC procedures and common benchmark datasets to evaluate new methods [35]. In parallel, meteorological agencies have begun integrating vetted PWS observations into their operations – for instance, assimilating filtered PWS data (surface pressure, temperature, humidity, rainfall) into high-resolution weather models – which has shown improved local forecasting after excluding low-quality stations [26, 36]. Community initiatives (e.g. promoting proper sensor siting and maintenance) are also helping to enhance data quality and increase trust in citizen-collected observations [7, 28]. Thanks to these advancements, personal weather stations are gradually being recognized as valuable components of the environmental monitoring infrastructure [20].

In summary, personal weather stations stand at the forefront of the crowdsourced data revolution, offering a cost-effective, spatially dense, and real-time source of meteorological information, while also motivating parallel exploration of new IoT-based observation systems. With continued improvements in data quality assurance and integration, PWS networks and their opportunistic counterparts are poised to substantially enrich urban hydrometeorological monitoring and resilience efforts in the years to come.

1.3 Research objectives

The main objective of this research is to evaluate the reliability and potential of Personal weather stations for enhancing rainfall monitoring and hydrological applications in urban areas. By integrating citizen-owned sensors with official weather networks, this study aims to assess the quality of PWS rainfall data and explore their contribution to increasing spatial resolution of observations. The research further investigates the effectiveness of established quality control techniques for filtering PWS data, compares rainfall estimates obtained from official weather station (OWS) datasets and from the combined OWS–PWS datasets with radar observations, and examines whether the integration of PWS networks provides added value in capturing localized precipitation dynamics that are often missed by conventional monitoring systems.

1.4 Research questions

This thesis seeks to answer the following key questions:

- To what extent do rainfall measurements from personal weather stations differ from those of nearby official weather stations, and how consistent are these differences across diverse climatic settings?

- How effective is the open-source quality control algorithm (PWS-pyQC) in improving the reliability of PWS rainfall data?
- Can the integration of quality-controlled PWS data with traditional rainfall networks enhance the spatial representation of precipitation relative to unadjusted radar fields?
- Under which network and event conditions can PWS data provide significant added value for capturing extreme rainfall patterns?

Chapter 2

Data and study area

In this chapter, the data sources and study areas considered in this research are described. Section 2.1 introduces the first source of rainfall data, namely the records obtained from official weather stations. Section 2.2 presents the second source of data, consisting of measurements from personal weather stations. Section 2.3 outlines the third source of data, derived from the national weather radar network operated by the Italian Civil Protection Department. Finally, Section 2.4 provides an overview of the study areas, including the climatic context of the selected cities.

2.1 Rain gauges

Rainfall measurements were obtained from rain gauges installed and maintained by the regional environmental agencies. For each city, data were collected from the official meteorological station networks by identifying all automatic gauges located within the municipal boundaries through the respective agency websites. The datasets retrieved from the sources listed in Table 2.1 were downloaded for the year 2021 and 2022. Not all stations, however, provided continuous records for the full interval, and temporal coverage therefore varied depending on operational history and data availability.

In addition, the temporal resolution differed among stations, ranging from one-minute to hourly accumulations depending on the instrumentation. As these stations belong to official monitoring networks, the raw datasets were assumed to have already undergone standard quality-control procedures. Only preliminary checks were carried out, such as the identification of missing values or evident outliers, while the validated structure of the data was preserved. Station coordinates were extracted in geographic format (EPSG:4326) and subsequently transformed into the projected system UTM Zone 33N (EPSG:32633) for use in spatial analyses, including distance calculations between official and personal weather stations.

Table 2.1: Official weather station web portals for the study cities.

City	Web Portal
Ancona	http://app.protezionecivile.marche.it/sol/indexjs.sol?lang=it
Bologna	https://simc.arpae.it/dext3r/
Catania	http://www.sias.regione.sicilia.it/frameset_download.htm
Catanzaro	https://www.cfd.calabria.it/
Firenze	https://www.sir.toscana.it/
Genova	https://www.arpal.liguria.it/
Milano	https://www.arpalombardia.it/
Padova	https://www.arpa.veneto.it/
Potenza	https://protezionecivile.regione.basilicata.it/protcivibas/home.jsp
Roma	https://protezionecivile.regione.lazio.it/
Torino	https://www.arpa.piemonte.it/

2.2 Personal weather stations

Personal weather station platforms such as Netatmo and Weather Underground (WU) make high-frequency observations of variables including rainfall, temperature, humidity, and wind available online, both in real time and as historical records. In practice, several commercial and volunteer PWS networks are currently in operation.

The Netatmo Weathermap is a global network of smart home weather stations. Each Netatmo station typically consists of an indoor module measuring temperature, humidity, and pressure, with optional outdoor modules for rainfall and wind. The rain gauge operates as a tipping-bucket with a 13 cm funnel, capable of measuring precipitation in increments of 0.101 mm across a range of 0.2 - 150 mm/h. Observations are automatically uploaded to the Netatmo cloud service at approximately five-minute intervals. The aggregated data are displayed through the Netatmo Weathermap platform, which also provides an Application Programming Interface (API) for data retrieval.

The Weather Underground Wundermap is one of the largest PWS networks, comprising more than 250,000 active stations worldwide. It integrates observations from multiple station brands, including rainfall, temperature, humidity, and wind, and provides access through its own API. AeriWeather's PWSweather is another commercial platform that allows users to manage their stations via an online dashboard and to access data through APIs offered under contributor agreements. In addition, Meteonetwork (Italy) maintains one of the largest regional PWS networks in Europe. Other examples include commercial services such as Davis Instruments' WeatherLink and AmbientWeather, as well as volunteer initiatives such as the Citizen Weather Observer Program (CWOP, United States) and the Met Office Weather Observations Website (WOW, United Kingdom), which allow privately owned stations to report data into national systems.

In this research, the Netatmo platform was adopted as the primary PWS data source. Netatmo was selected because of its widespread deployment, its relatively high spatial

density in Europe, and the accessibility of its data through an API. Previous studies have emphasized that Netatmo is particularly advantageous among PWS platforms because it combines a high network density with standardized instrumentation and open data access. The Netatmo API enables programmatic retrieval of station observations together with metadata, including station coordinates. By leveraging this dense network and open interface, a large, geo-referenced dataset of rainfall measurements was compiled. In contrast, some regional PWS platforms lack open APIs, and others provide less extensive coverage.

2.3 Weather radar

The RADAR-DPC platform is a national-scale weather monitoring system operated by the Italian Civil Protection Department (Dipartimento della Protezione Civile). It provides near real-time visualization of current and recent meteorological phenomena by integrating raw observations from the national radar network, rain-gauge and temperature station networks, meteorological satellites, and lightning detection systems. All products are distributed as open-access web services.

The system is supported by the WIDE (Weather Ingestion Data Engine), which automates data ingestion, quality control, and the fusion of multiple sources before publishing the processed products. This infrastructure also enables time-series queries, allowing past observations and model outputs to be retrieved for analysis. A suite of radar-derived products is generated, including Vertical Maximum Intensity (VMI) and Surface Rainfall Total (SRT). For this thesis, the Surface Rainfall Intensity (SRI) product was used. SRI is developed through operational chains at the Centro Funzionale Centrale (CFC) by combining data from the national radar and rain-gauge networks. The product provides an estimate of ground-level rainfall intensity in millimetres per hour, updated every five minutes.

The SRI rainfall grids are supplied as georeferenced rasters in the WGS84/UTM Zone 33N projection with a spatial resolution of $1 \text{ km} \times 1 \text{ km}$. Data values are stored as 32-bit floats, and missing or invalid values are indicated by a nodata flag (-9999), as specified in the metadata files. The high temporal resolution ensures that rapid changes in convective rainfall can be effectively monitored.

The RADAR-DPC system can be accessed via an interactive online dashboard, which provides real-time visualisation and options for downloading data. Through this platform, a comprehensive and user-friendly set of meteorological products is made available under open license, enabling researchers, agencies, and practitioners to freely access and use the information. In this study, the SRI product was specifically employed for the case studies of Rome and Turin, where it served as a reference for evaluating the interpolated rainfall fields obtained from the combined PWS and OWS datasets.

2.4 Study areas

For the first part of this research, eleven Italian cities were selected in order to ensure coverage of diverse geographical contexts and climatic regimes. The selection included urban

areas distributed across the north, center, and south of the country, thus representing a wide range of rainfall conditions and hydrological settings. In chapter 3 and afterwards, particular attention will be given to Rome and Turin, which serve as the principal case studies.

Ancona represents the central Adriatic coast, with a Mediterranean climate of mild winters, hot summers, and rainfall maxima in autumn and winter. Bologna, in the Po Valley, exemplifies the continental regime with year-round precipitation and peaks in spring and autumn. Catania, on the Ionian coast of Sicily, typifies the warm southern Mediterranean climate, with wet winters and dry summers. Catanzaro, in Calabria at higher elevation, reflects a sub-Mediterranean highland regime with relatively high winter rainfall. Florence, in the Arno Valley, has a transitional Mediterranean climate, with precipitation concentrated in autumn and spring. Genoa, on the Ligurian coast, has a fully Mediterranean climate with abundant rainfall, particularly in autumn. Milan, in Lombardy, represents the populous urban climate of the western Po Valley, with hot, humid summers, cool winters, and frequent fog. Padua, in the Veneto plain, reflects the eastern Po Valley regime, with rainfall distributed throughout the year. Potenza, in the Apennines, exemplifies a high-elevation climate with snowy winters and relatively high annual rainfall.

Rome, the capital of Italy, has a classic Mediterranean climate with hot, dry summers and mild, wet winters. Turin, located at the foot of the Alps, has a temperate continental climate with cold winters, warm summers, and rainfall peaks in spring and autumn.

Chapter 3

Evaluation of raw PWS data

This chapter presents the evaluation of raw rainfall data from personal weather stations against nearby official weather stations across eleven Italian cities during the period 2021–2022. The objective is to assess the reliability of raw PWS records in reproducing long-term rainfall totals and extreme events before applying any advanced quality control procedures.

The chapter is organized into three main sections. Section 3.1 describes the methodology adopted, including preliminary control, PWS–OWS pairing, and the statistical indicators used for comparison. Section 3.2 presents the agreement analysis between PWSs and OWSs through a set of figures and tables, with particular attention to Roma and Torino as key case studies. Finally, Section 3.3 discusses the main results, highlighting patterns of agreement and limitations in the raw PWS datasets.

3.1 Methodology

In order to evaluate the reliability of raw PWS data, all available time series for eleven Italian cities (Ancona, Bologna, Catania, Catanzaro, Florence, Genoa, Milan, Padua, Potenza, Rome, and Turin) were considered for the period 2021–2022.

A preliminary control was applied to ensure data completeness: stations with more than 10% missing data in a given year were excluded from further analysis. The remaining PWSs were then paired with their nearest OWS within the same city. Distances between stations were computed using the UTM Zone 33N projection (EPSG:32633) to minimize distortion across Italy.

For each PWS–OWS pair and year, the following statistics were computed:

- Annual precipitation totals (mm).
- Percentage of zero-rainfall records (%).
- Annual maximum rainfall depths (mm) for durations of 1, 3, 6, 12, and 24 hours.

3.2 Comparison between PWSs and OWSs

This section presents the agreement analysis between PWSs and OWSs to evaluate the reliability of raw PWS rainfall data before quality control. The comparison is organised to progressively examine different aspects of measurement consistency across the Italian network. First, three national-scale analyses are performed to assess the agreement between PWSs and OWSs in terms of annual rainfall totals (Table 3.1; Fig. 3.1), the frequency of zero-rainfall observations (Fig. 3.2), and the dependence of bias on rainfall intensity using 1-hour annual maxima (Fig. 3.3). These complementary metrics provide an overall view of how raw PWS data reproduce rainfall magnitude, occurrence, and extremes across the country. Finally, two case studies (Rome and Turin) are analysed in greater detail (Figs. 3.4–3.7) to investigate local-scale variability, inter-station consistency, and the capability of PWSs to capture extreme rainfall events. All results refer to raw PWS datasets prior to any quality-control or bias-correction procedures.

3.2.1 National comparison of annual rainfall totals

This subsection evaluates the overall agreement between PWSs and OWSs in terms of annual rainfall totals across eleven Italian cities during 2021–2022. The comparison results, summarised in Table 3.1, report the total number of available PWSs, the subset retained after applying the 10% missing-data criterion, the mean PWS–OWS distance, the mean relative bias of annual rainfall totals, and the proportion of pairs falling within $\pm 10\%$ and $\pm 20\%$ of the OWS reference.

Across most cities and years, the mean relative bias is negative, indicating a general underestimation of annual rainfall by raw PWS data. Typical values range between -10% and -30% , with Catania exhibiting the largest deficits (up to -72%) and Padua 2021 being nearly unbiased (-1%). Rome 2022 represents the opposite behaviour, showing a marked overestimation of $+37.0\%$. The fraction of pairs within $\pm 10\%$ of OWS totals seldom exceeds 30–35%, increasing to roughly 40–50% when the tolerance is relaxed to $\pm 20\%$. These results demonstrate that unfiltered PWS datasets display significant spatial variability and systematic biases.

Such heterogeneity is consistent with previous studies indicating that the reliability of PWS data is strongly affected by installation quality and infrequent maintenance. Mechanical obstructions (e.g., insects or twigs) and improper levelling of tipping-bucket gauges can lead to systematic underestimation, while occasional handling or cleaning by PWS owners may cause false tips and short-term overestimation [29]. Design limitations and the lack of calibration in consumer-grade gauges can also introduce structural bias, particularly under high-intensity rainfall [28]. These physical and operational factors help explain the inter-city variability observed in the present analysis.

Table 3.1: Summary of annual rainfall agreement between PWSs and OWSs across Italian cities (2021–2022). Metrics include the total number of available PWSs, the number of valid PWS–OWS pairs after applying the 10% missing-data criterion, the mean PWS–OWS distance, the mean relative bias in annual rainfall totals, and the percentage of pairs within $\pm 10\%$ and $\pm 20\%$ of the OWS reference.

City	Year	Total PWSs	Valid Pairs	Mean Distance (km)	Mean Relative Bias (%)	Within $\pm 10\%$ (%)	Within $\pm 20\%$ (%)
Ancona	2 021	5	1	2.12	2.0	100.0	100.0
Ancona	2 022	5	1	2.12	−5.8	100.0	100.0
Bologna	2 021	25	7	3.40	−14.5	42.9	57.1
Bologna	2 022	25	9	3.31	−11.8	33.3	55.6
Catania	2 021	8	2	7.01	−42.3	0.0	0.0
Catania	2 022	8	1	9.91	−72.0	0.0	0.0
Catanzaro	2 021	5	1	0.73	−17.4	0.0	100.0
Catanzaro	2 022	5	1	2.35	−36.7	0.0	0.0
Florence	2 021	26	9	4.23	−20.8	11.1	33.3
Florence	2 022	26	10	2.87	−11.4	20.0	50.0
Genoa	2 021	64	24	1.38	−11.1	33.3	45.8
Genoa	2 022	64	27	1.28	−17.2	25.9	37.0
Milan	2 021	91	34	2.58	−14.0	26.5	50.0
Milan	2 022	91	40	2.33	−16.8	25.0	40.0
Padua	2 021	14	8	2.65	−1.0	37.5	87.5
Padua	2 022	14	7	2.81	10.1	28.6	57.1
Potenza	2 021	2	0	1.95	–	–	–
Potenza	2 022	2	0	1.95	–	–	–
Rome	2 021	148	60	2.31	−23.1	18.3	41.7
Rome	2 022	148	86	3.52	37.0	16.3	38.4
Turin	2 021	45	18	2.09	−25.2	16.7	27.8
Turin	2 022	45	18	2.25	−27.2	27.8	50.0

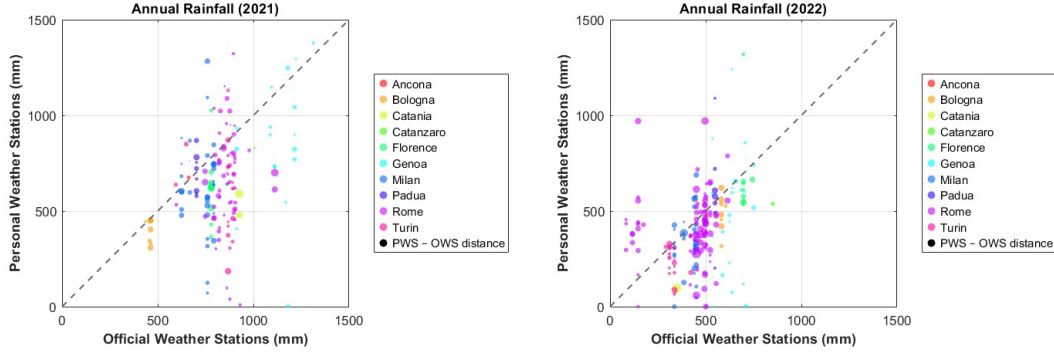


Figure 3.1: Annual rainfall comparison between PWSs and OWSs across all cities (2021–2022). Marker size indicates PWS–OWS distance; colours represent cities; the dashed line denotes perfect agreement (1:1).

3.2.2 National comparison of zero-rainfall frequencies

Across all cities and years (2021–2022), the average difference between PWS and OWS zero frequencies was extremely small, with global means of $+0.09\%$ (2021), and -0.34% (2022), and standard deviations of about $1.6\text{--}1.8\%$. These values demonstrate that PWSs reproduced the dry–wet occurrence patterns of the official network with high consistency, with no systematic tendency to over- or under-detect rainfall events. The tight clustering of points around the 1:1 line further confirms that discrepancies observed in annual or sub-hourly rainfall totals originate mainly from volume-measurement bias rather than from event-detection failures.

These results confirm that PWSs closely reproduce the temporal pattern of wet and dry periods observed by the official network. The near-perfect alignment of most pairs along the 1:1 line indicates that the sensors are reliable in detecting rainfall occurrence even when their measured volumes differ. This finding supports the interpretation that the discrepancies discussed in Section 3.2 originate primarily from measurement bias and mechanical sensitivity rather than missed or false rainfall events. Consequently, the zero-rainfall comparison provides additional evidence that unfiltered PWS data are temporally consistent and suitable for use in subsequent quality-control and interpolation analyses.

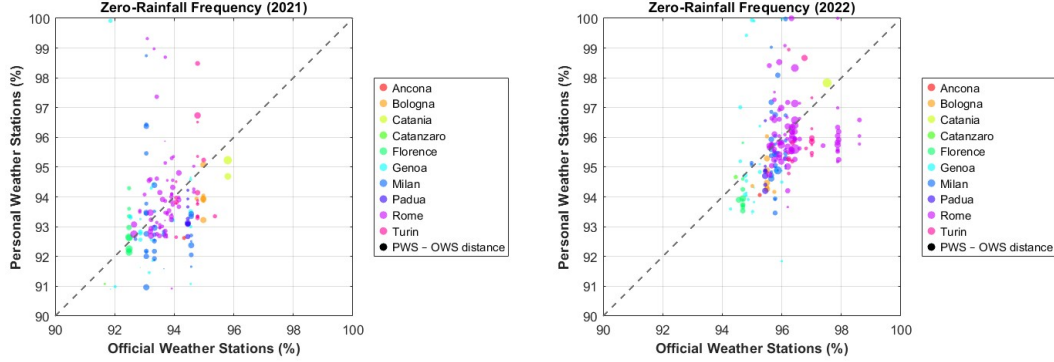


Figure 3.2: Percentage of zero-rainfall records in PWSs versus OWSs across all cities (2021–2022). Each marker represents a PWS–OWS pair, coloured by city; the dashed line denotes perfect agreement (1:1).

3.2.3 National comparison of 1-hour annual maximum rainfall

To investigate whether bias magnitude depends on rainfall intensity, a correlation analysis was performed between the OWS 1-hour annual maximum rainfall and the corresponding PWS relative deviation. A negative correlation indicates that underestimation increases with rainfall magnitude, while a positive correlation suggests growing overestimation at higher intensities. The resulting statistics were computed for each city and for the aggregated dataset over the years 2021–2022.

Although correlations were estimated for individual cities, the number of valid OWS–PWS pairs was often too limited for meaningful statistical interpretation. In most cities, the number of valid OWS–PWS pairs did not exceed ten to twenty, which is insufficient for reliable correlation analysis. Statistical studies demonstrate that correlations derived from small samples are unstable and may deviate markedly from the true population value. The minimum sample size required for stable correlation estimates depends on the expected effect size and study context; however, in most research settings, samples smaller than approximately $n = 150$ yield unreliable results, and reasonable trade-offs between accuracy and confidence are achieved only when sample sizes approach about 250 observations [37].

Consequently, this study interprets only the results from the aggregated datasets, which combine all station pairs from the eleven cities. The 2021 and 2022 datasets, comprising 160 and 197 valid pairs respectively, meet the minimum requirements for statistical robustness.

The aggregated results reveal a consistent and statistically significant negative relationship between OWS 1-hour annual maxima and the corresponding PWS relative deviations. In 2021, the correlation was $r = -0.23$ ($p = 0.003$; Spearman $\rho = -0.44$, $p < 0.001$), indicating that PWS underestimation intensifies with increasing rainfall magnitude. This pattern strengthened in 2022 ($r = -0.44$, $p < 0.001$; $\rho = -0.46$, $p < 0.001$), confirming that bias between PWS and OWS measurements becomes more pronounced

during high-intensity events. The observed negative correlation between rainfall intensity and PWS bias confirms that low-cost sensors systematically underestimate extreme precipitation, consistent with the mechanical constraints of tipping-bucket gauges [28,29].

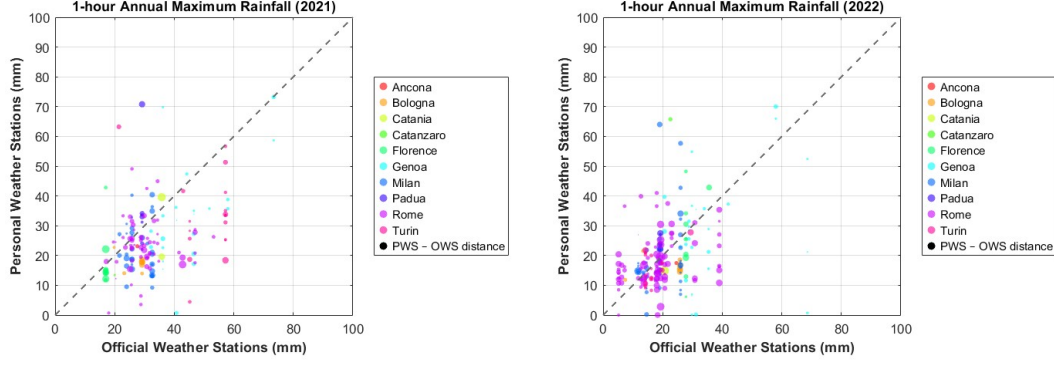


Figure 3.3: Comparison of 1-hour annual maximum rainfall between PWSs and OWSs across all cities (2021-2022). Marker size indicates PWS–OWS distance; colours represent cities; the dashed line denotes perfect agreement (1:1).

3.2.4 Case-study analysis: Rome and Turin

Among the investigated cities, Rome and Turin were selected for detailed evaluation owing to their contrasting climatic and geographical settings and the availability of continuous multi-year PWS–OWS datasets. Rome, located in central Italy under a Mediterranean regime, frequently experiences convective storms and short-duration flash floods, whereas Turin, situated in the northwestern Po Valley at the base of the Alps, is characterised by mixed convective and orographic rainfall with marked seasonal variability. Both cities also have coinciding radar observations for significant flood events (8 June 2021 in Rome and 22 June 2021 in Turin), which allows, in the following chapters, a comparison between interpolated rainfall maps derived from (OWS + PWS) and radar-derived fields. These contrasting meteorological conditions make them ideal test sites to investigate whether integrating PWSs—particularly those within the $\pm 10\%$ – $\pm 20\%$ bias buffer relative to OWSs—can enhance the spatial representation of rainfall and improve agreement with radar estimates.

Figures 3.4 and 3.5 show the PWS–OWS pairwise comparison of annual rainfall totals for Rome and Turin, with shaded areas corresponding to $\pm 10\%$ and $\pm 20\%$ deviation from the OWS reference.

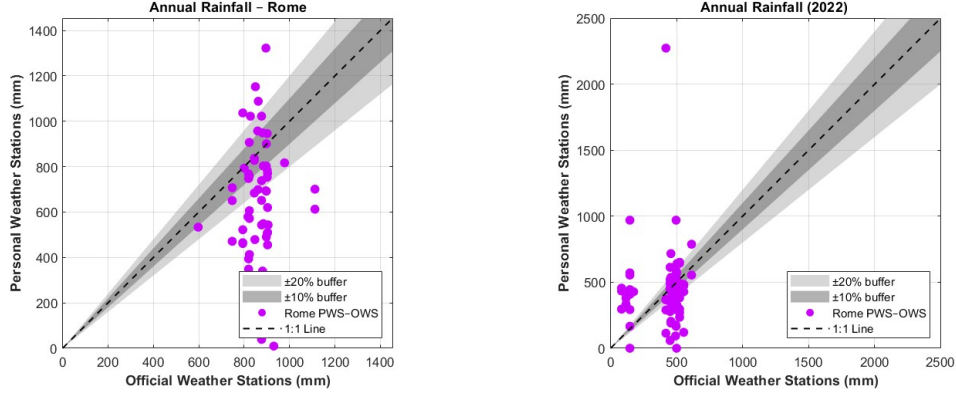


Figure 3.4: Annual rainfall agreement between PWSs and OWSs in Rome for 2021 and 2022. Shaded bands represent $\pm 10\%$ and $\pm 20\%$ deviation from the 1:1 line.

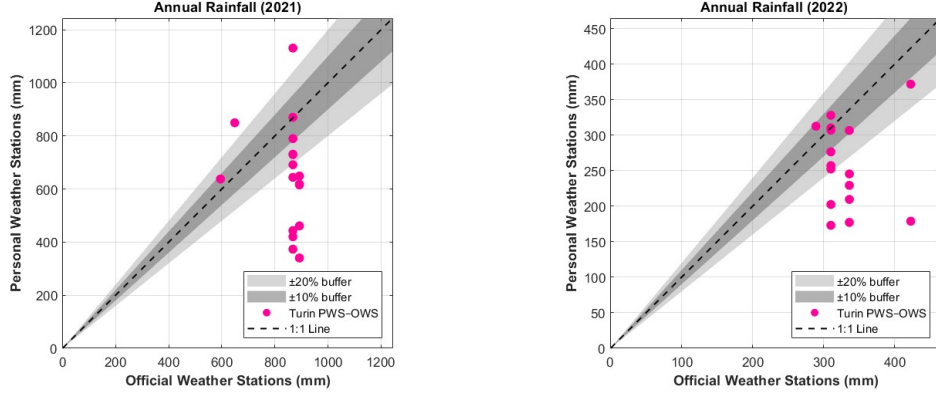


Figure 3.5: Annual rainfall agreement between PWSs and OWSs in Turin for 2021 and 2022. Shaded bands represent $\pm 10\%$ and $\pm 20\%$ deviation from the 1:1 line.

In Rome, the mean relative bias shifted from an underestimation of -23.1% in 2021 to an overestimation of $+37.0\%$ in 2022, with roughly one fifth of pairs within $\pm 10\%$ of OWS values. In Turin, underestimation persists in both years (-25.2% and -27.2%), although the share of pairs within $\pm 20\%$ increased from 28% to 50%. Given the coexistence of over- and underestimation across stations, a uniform city-level multiplier cannot meaningfully reduce error without worsening it for a substantial subset of PWSs. Consequently, accuracy improvements depend on a rigorous quality-control and bias-correction workflow that accounts for station-specific behaviour and event conditions.

To further characterise the performance of the two station networks, the variability of annual precipitation totals was assessed using boxplots (Fig. 3.6). These plots summarise the statistical distribution of annual rainfall recorded by PWSs and OWSs in Rome and Turin, providing a visual comparison of their internal consistency. In both cities, PWS datasets display substantially wider interquartile ranges and a larger overall spread than the corresponding OWS records, indicating higher variability among individual sensors.

Such dispersion reflects the heterogeneous behaviour of citizen-owned gauges, which are affected by differing installation environments, maintenance practices, and calibration states. In contrast, the narrower and more uniform distributions observed for OWSs confirm the reliability expected from professionally maintained instruments.

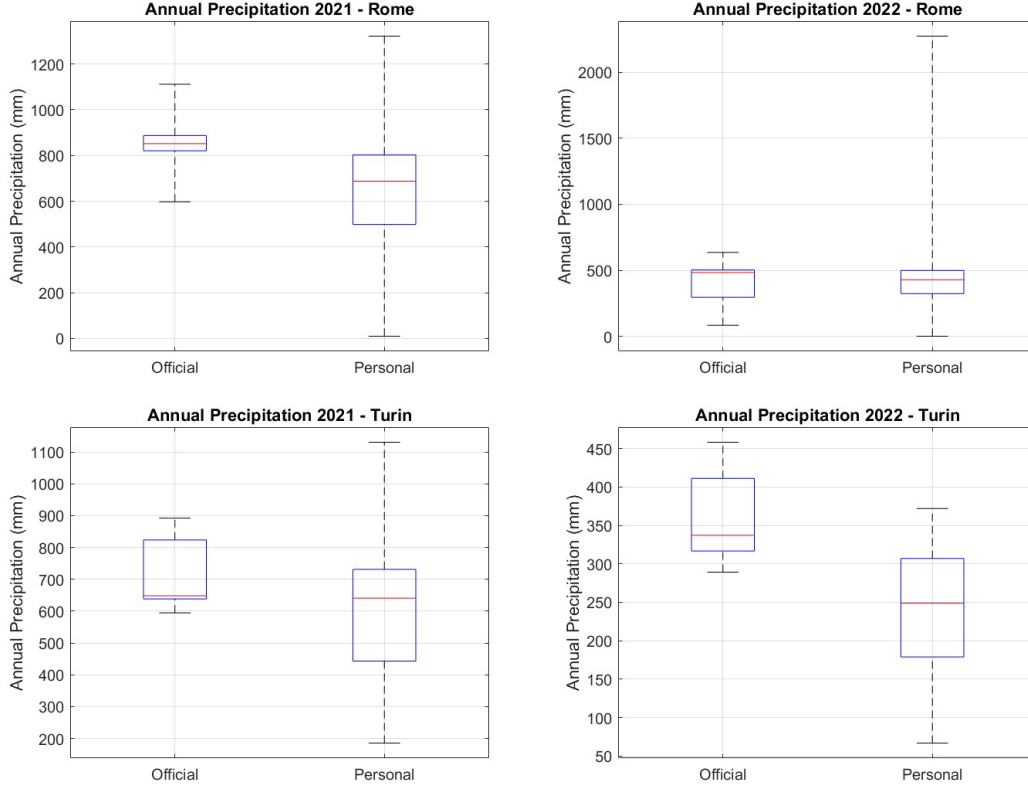


Figure 3.6: Boxplots of annual precipitation from OWSs and PWSs for Rome and Turin (2021–2022).

To evaluate the representativeness of PWSs during intense rainfall, an “extreme-capture” index was computed for each duration (1–24 h), defined as the percentage of PWSs recording at least 80% of the maximum rainfall measured by any OWS in the same city and year (Fig. 3.7). In Rome, capture efficiency increases with duration, reaching approximately 60–65% at 6–12 h in 2021 and remaining above 55% across all durations in 2022. This pattern indicates that PWSs better represent multi-hour storm totals than short-duration peaks, consistent with the mechanical limits of tipping-bucket sensors, which tend to underestimate during high-intensity bursts. The relatively uniform behaviour observed in 2022 suggests improved internal consistency of the PWS network, likely due to the exclusion of incomplete records and a denser, more stable set of active stations.

In Turin, performance is substantially lower in 2021, with only 20–25% of PWSs capturing at least 80% of the OWS maximum. This widespread underestimation points to reduced catch efficiency caused by intense convective events and installation-related

issues such as partial sheltering or poor levelling. However, in 2022 the performance improves markedly, with capture efficiency rising to nearly 80% for 3 h and about 70% for 6 h events, before gradually declining for longer durations. This trend confirms that PWSs reproduce multi-hour accumulations more reliably than short, high-intensity rainfall peaks, while also demonstrating that network consistency and rainfall regime significantly influence measurement accuracy. The overall behaviour observed for both cities aligns with previous studies reporting systematic underestimation of short-duration extremes by uncorrected PWS datasets [30], and a generally better agreement with official records at longer aggregation scales [38]. These findings highlight that while raw PWS data can adequately represent intense rainfall over several hours, short-term peaks remain prone to mechanical and exposure-induced undercatch, reinforcing the need for subsequent quality control and bias correction.

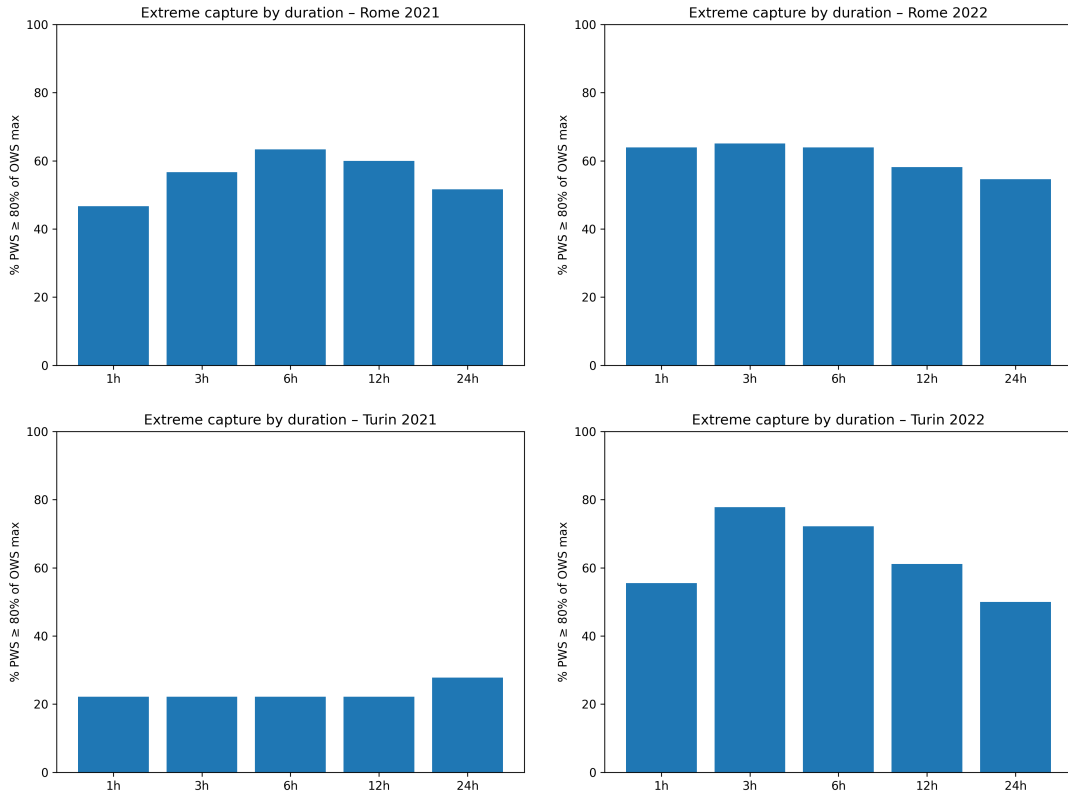


Figure 3.7: Percentage of PWSs capturing at least 80% of the maximum OWS rainfall for each duration (1–24 h) in Rome and Turin (2021–2022).

3.3 Discussion and interpretation

At the national scale, raw PWS data show a prevailing tendency to underestimate rainfall relative to OWS observations, with bias magnitude increasing under high-intensity

conditions. This systematic undercatch reflects the mechanical limits of low-cost tipping-bucket gauges, which often lose efficiency during intense rainfall. The pattern, however, is not uniform across the network: isolated episodes of strong overestimation were identified, particularly in Rome 2022, where mean PWS totals exceeded OWS values by more than 50%. Such anomalies are consistent with false-tipping artefacts during cleaning or maintenance, as reported in the literature, and highlight the heterogeneous behaviour of citizen-operated instruments. The “consistent underestimation” observed at the national scale should therefore be interpreted as a dominant trend rather than a universal rule.

As shown in Section 3.2, PWSs nonetheless reproduce rainfall occurrence with remarkable fidelity. The mean difference in zero-rainfall frequency relative to OWSs remained below 1% across all cities and years, confirming that discrepancies mainly arise from volume bias rather than event-detection failures. Because this analysis was performed at hourly resolution, minor undercatch during very short or light events may still occur, yet at the temporal scales relevant for hydrological analysis PWSs demonstrate reliable detection capability.

Spatially, PWS performance depends strongly on climatic and operational context. Northern and central cities such as Milan, Genoa, and Padua exhibited smaller deviations from OWS references, while southern locations like Catania and Catanzaro showed pronounced underestimation, likely linked to more convective rainfall and less optimal station siting. Rome and Turin exemplify these contrasts: Rome alternated between under- and overestimation depending on rainfall regime, whereas Turin persistently underestimated but displayed improved internal consistency over time. These results confirm that both climate and maintenance practices play a decisive role in determining network reliability.

The case studies further highlight how PWS networks behave during intense rainfall events. As detailed in Section 3.2, Turin initially displayed low extreme-capture efficiency but improved substantially in 2022, while Rome maintained relatively high values across both years. These trends indicate that PWSs reproduce multi-hour rainfall accumulations more effectively than short, high-intensity peaks, and that network consistency and rainfall regime exert a strong influence on their performance. The overall behaviour observed for both cities aligns with previous evaluations of PWS datasets [30, 38], which reported notable underestimation for sub-hourly extremes but progressively closer agreement with official gauges at longer durations.

Although PWSs capture rainfall timing and longer-duration totals reasonably well, large interquartile ranges in the annual boxplots reveal substantial variability among individual sensors. This dispersion shows that, while the network is temporally coherent, it remains spatially heterogeneous and affected by station-specific biases. Because over- and underestimation coexist within the same city, applying a single correction factor would not improve overall accuracy. Instead, these findings emphasise the need for a station-level quality-control procedure capable of identifying and correcting individual anomalies.

In summary, the analyses in this chapter establish the baseline reliability of raw PWS datasets and clarify the limits of their direct hydrological applicability. Unfiltered PWS data provide temporally consistent yet quantitatively biased rainfall information, with accuracy shaped by both environmental and operational conditions. These findings

lay the foundation for Chapter 4, which investigates whether integrating selected PWSs can enhance the spatial representation of rainfall through interpolation techniques and improve agreement with radar-derived estimates during two flash-flood events.

Chapter 4

Rainfall interpolation and radar validation with raw PWS data

This chapter investigates whether the inclusion of uncorrected observations from personal weather stations enhances the spatial representation of rainfall during short-duration, high-intensity events. Building upon the reliability assessment of raw PWS data presented in Chapter 3, the analyses focus on two flash-flood episodes that affected Rome on 8 June 2021 and Turin on 22 June 2021. In Rome, a yellow-level alert was issued by the Civil Protection the previous day, warning of scattered but locally severe thunderstorms. The storm produced heavy convective rainfall, particularly in the northern districts such as Ponte Milvio, where localized flooding occurred within a few hours. In Turin, the event was triggered by a violent convective cell that delivered up to 95.8 mm of rain in a day and nearly 70 mm in one hour, causing widespread street flooding and numerous emergency interventions across northern districts reported by ARPA Piemonte.

4.1 Methodology

Rainfall analyses were carried out using three complementary datasets: official weather stations, personal weather stations, and radar-derived accumulations from the national Radar-DPC network.

Radar accumulations were computed from five-minute Surface Rainfall Intensity provided by the Dipartimento della Protezione Civile. For each event, all available SRI GeoTIFFs were aggregated over a 24-hour window consistent with the gauge integration period (00:00–24:00 UTC). A conservative missing-data criterion was applied: if a pixel contained a NoData value in any time step, the entire pixel was flagged as missing in the final daily grid. The daily composites were then cropped to the bounding boxes of Rome and Turin and stored in a common spatial reference system (EPSG:32633 – UTM 33N). The interpolation grid resolution was fixed at 1 km² to maintain consistency between radar and gauge-based products. This workflow ensured that radar and station accumulations were evaluated over identical spatial and temporal domains for both case studies.

Spatial rainfall fields derived from OWSs were generated using two interpolation techniques: Inverse Distance Weighting (IDW) and Ordinary Kriging (OK). IDW served as the deterministic baseline, while OK represented the geostatistical alternative. Three spherical, exponential and Gaussian variogram models were tested for OK. Model parameters for both techniques were optimised through leave-one-out cross-validation at station locations, selecting the configuration that minimised the root-mean-square error (RMSE) between observed and predicted rainfall values.

The same interpolation procedures were subsequently applied to combined OWS-PWS datasets. Two PWS subsets were considered: (i) stations whose annual rainfall totals for 2021 were within $\pm 10\%$ of the corresponding OWS reference, and (ii) those within $\pm 20\%$, representing a more inclusive threshold. This selection ensured that only stations with broadly consistent long-term behaviour contributed to the spatial analysis.

This city-dependent performance is consistent with evidence from other studies, mentioned the choice of interpolation technique depends strongly on local topography, data density, and rainfall variability. In topographically complex or data-sparse regions, geostatistical and regression-based methods such as kriging can produce unstable spatial patterns at daily scales, while simpler deterministic approaches like IDW often yield more robust and physically realistic results, even if some smoothing occurs [39]. These considerations support the methodological decision to retain, for each city, the interpolator that achieved the best overall agreement with radar accumulations.

The interpolated rainfall grids were evaluated against radar-derived accumulations over the common valid-data mask, defined by overlapping non-missing pixels. The comparison employed a set of complementary statistical metrics, including the Pearson correlation coefficient (r) and Spearman correlation coefficient (ρ), RMSE, mean absolute error (MAE), bias, Nash–Sutcliffe efficiency (NSE), and the Structural Similarity Index (SSIM). Visual examination of spatial patterns and difference maps supported the quantitative assessment.

Since the RADAR–DPC fields used in this study are not gauge-adjusted, they cannot be regarded as accurate in terms of absolute rainfall magnitude. Consequently, error-based indicators (MAE, RMSE, Bias, NSE) are interpreted qualitatively, whereas pattern-oriented measures— ρ and SSIM—are considered the most reliable descriptors of spatial agreement between interpolated and radar-derived rainfall fields.

The Spearman rank metric is particularly useful for this type of comparison, as it captures whether regions of higher and lower rainfall coincide, regardless of systematic bias. Previous studies have shown that Spearman correlation often mirrors Pearson correlation in precipitation datasets, but remains more robust when the distributions are heavy-tailed or affected by measurement inconsistencies [40]. The SSIM, originally developed for image analysis, quantifies the similarity between two spatial fields by jointly comparing local luminance, contrast, and structural information. When applied to gridded rainfall maps, SSIM provides an integrated score (ranging from -1 to 1) that measures how well the spatial organization of precipitation is preserved beyond pointwise error metrics. As noted by different studies, SSIM can be extended to quantify structural differences between two geophysical datasets, with a value of 1 representing a perfect structural match. Thus, higher SSIM values indicate closer alignment of rainfall patterns, providing a meaningful

measure of spatial fidelity complementary to correlation analysis [41].

The following section presents the outcomes of these interpolations and their comparison with radar accumulations for the two flash-flood events, focusing on how the integration of PWSs influenced the spatial accuracy of rainfall fields in Rome and Turin.

4.2 Spatial comparison with radar data

This section presents the spatial analysis of rainfall fields derived from gauge-based interpolation and their comparison with unadjusted RADAR–DPC accumulations for the two case studies: Turin (22 June 2021) and Rome (8 June 2021). The objective is to evaluate whether integrating raw PWSs into the OWSs improves the spatial representation of rainfall patterns. The analysis focuses on spatial agreement rather than absolute accuracy. Pattern-based metrics, namely the Spearman correlation coefficient and the Structural Similarity Index, are therefore considered the most meaningful indicators, whereas absolute error measures (RMSE, MAE, Bias, NSE) are reported only for qualitative comparison.

4.2.1 Turin — 22 June 2021

The network configuration for Turin (EPSG:32632) is shown in Figure 4.1. The OWSs includes only a few gauges, mostly concentrated in the northern and central sectors of the city, with limited coverage in the southern and western areas. The integration of PWSs markedly increases the number of available stations, particularly across the central and southern districts, where few official gauges are present. This denser configuration helps fill several spatial gaps in the monitoring network, with the $\pm 10\%$ and $\pm 20\%$ thresholds representing progressively broader PWS subsets.

Several OK configurations were initially tested using spherical, exponential, and Gaussian variogram models optimised through leave-one-out cross-validation. Although OK reproduced the overall rainfall gradients, it tended to oversmooth local maxima and produced lower spatial correspondence with the radar field. Consequently, IDW was adopted as the preferred interpolator for the Turin event. A representative OK output obtained with the exponential variogram model is included in Appendix to illustrate the over-smoothing effect.



(a) OWSs only



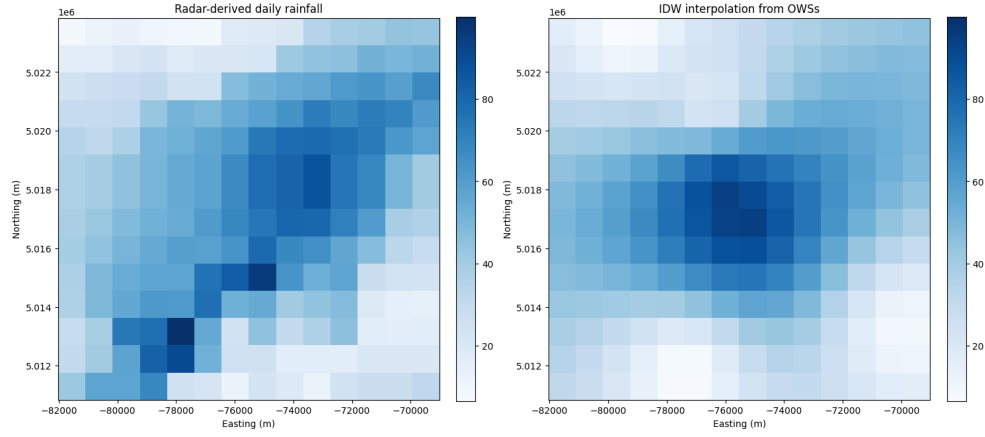
(b) OWSs + PWSs ($\pm 10\%$)



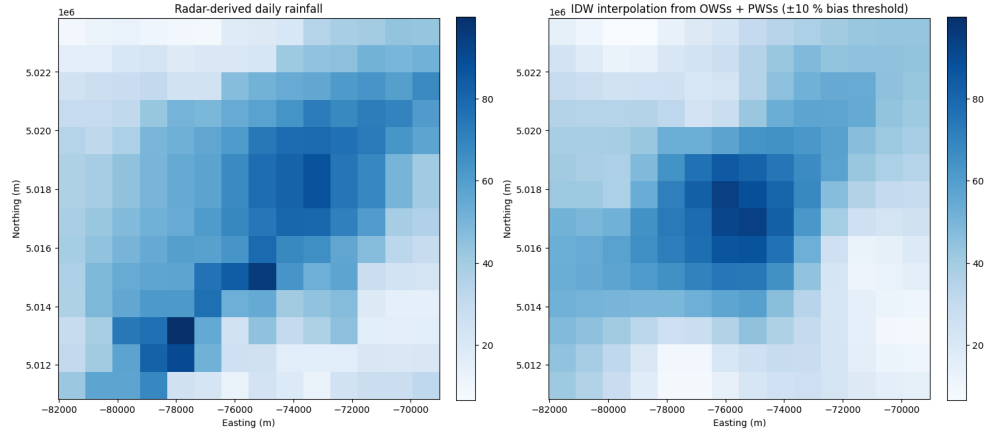
(c) OWSs + PWSs ($\pm 20\%$)

Figure 4.1: Location of (a) OWSs, (b) OWSs with PWSs within $\pm 10\%$ (shown in dark blue), and (c) OWSs with PWSs within $\pm 20\%$ (shown in light blue) within the Turin administrative boundary (EPSG:32632).

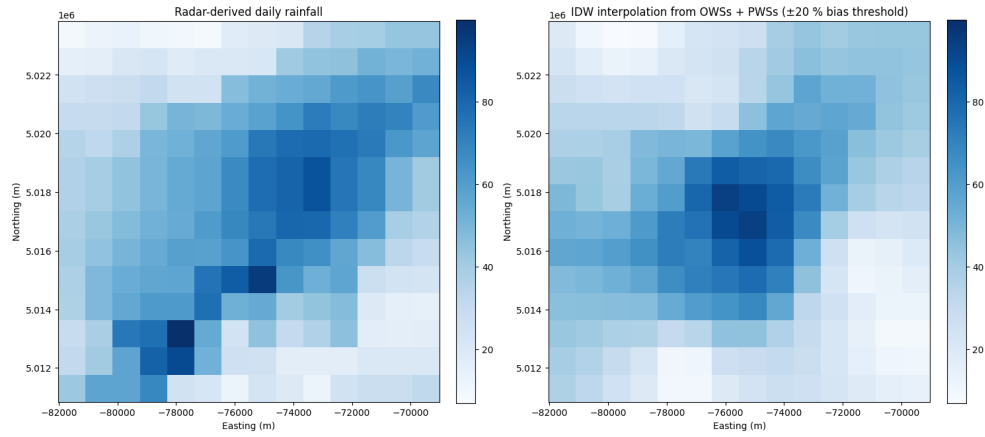
Figure 4.2 compares radar and interpolated rainfall maps. The OWS-only interpolation reproduces the location of the main rainfall maximum. When integrating PWSs within $\pm 10\%$, the field becomes smoother, with small changes in ρ and SSIM. Further expansion to the $\pm 20\%$ subset leads to a slightly more homogeneous field and a modest improvement in correlation and pattern similarity. All panels share identical colour limits to allow consistent visual comparison.



(a) Radar vs. OWS-only



(b) Radar vs. OWS + PWSs ($\pm 10\%$)



(c) Radar vs. OWS + PWSs ($\pm 20\%$)

Figure 4.2: Comparison between radar-derived and interpolated daily rainfall for Turin (22 June 2021): (a) Radar vs. OWSs-only, (b) Radar vs. OWSs with PWSs ($\pm 10\%$), (c) Radar vs. OWSs with PWSs ($\pm 20\%$) (EPSG:32633).

Table 4.1: Statistical comparison between radar and interpolated rainfall fields for Turin (22 June 2021).

Configuration	ρ	SSIM	r	MAE [mm]	RMSE [mm]	Bias [mm]	NSE
OWS only	0.710	0.616	0.683	11.873	17.195	−3.533	0.325
OWS + PWSs ($\pm 10\%$)	0.706	0.622	0.686	12.673	17.528	−5.696	0.298
OWS + PWSs ($\pm 20\%$)	0.714	0.638	0.701	12.364	16.977	−5.330	0.342

Overall, the inclusion of raw PWS observations produced small but consistent improvements in the spatial structure of rainfall fields over Turin. These improvements were mainly reflected in pattern-based indicators (Spearman’s ρ and SSIM), which quantify spatial coherence.

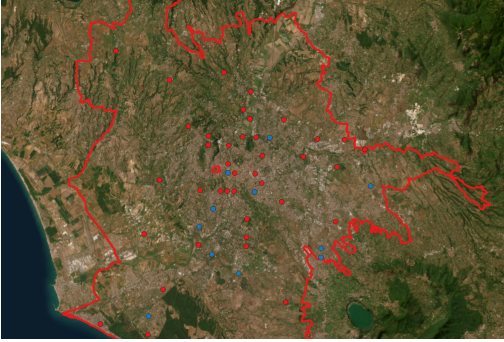
4.2.2 Rome — 8 June 2021

The configuration of official and personal weather stations across Rome (EPSG:32633) is shown in Figure 4.3. The OWS network is fairly dense and evenly distributed across most of the municipal area, with the highest concentration of gauges in the central and northern sectors and slightly fewer stations toward the western and coastal zones. Adding PWSs under the $\pm 10\%$ and $\pm 20\%$ bias thresholds increases the total station density, particularly across the central and southern parts of the city, reducing several of the remaining gaps in the official network.

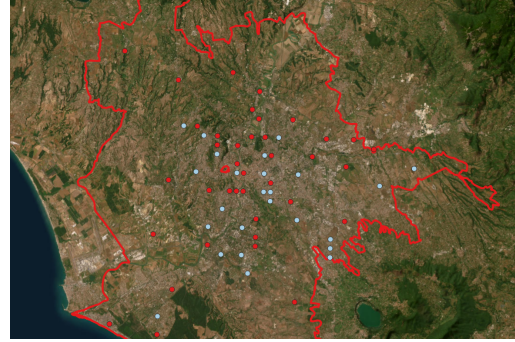
For Rome, interpolation was performed using the OK method, applying the same grid and valid-data mask as for the radar fields. Figure 4.4 illustrates the radar and interpolated rainfall distributions. The radar shows a strongly convective event with local maxima exceeding 120 mm day^{-1} , mainly in the north-western sector. The OWS-only interpolation captures the general orientation of the rainfall core but smooths the gradients. Integrating PWSs within $\pm 10\%$ leaves the overall pattern largely unchanged, whereas the $\pm 20\%$ configuration disperses the rainfall core and reduces the spatial agreement with radar. To prevent visual saturation caused by radar overestimation, interpolation maps use the same colour limits but interpretations rely on relative rather than absolute values.



(a) OWSs only

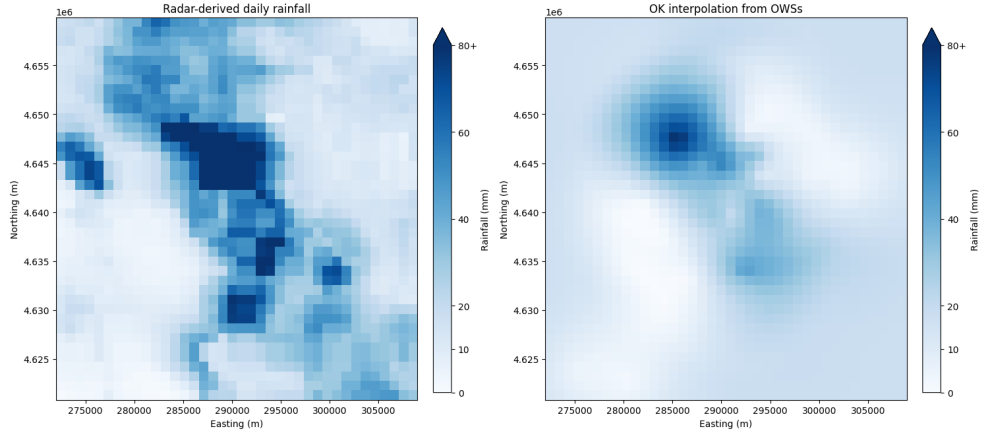


(b) OWSs + PWSs (± 10 %)

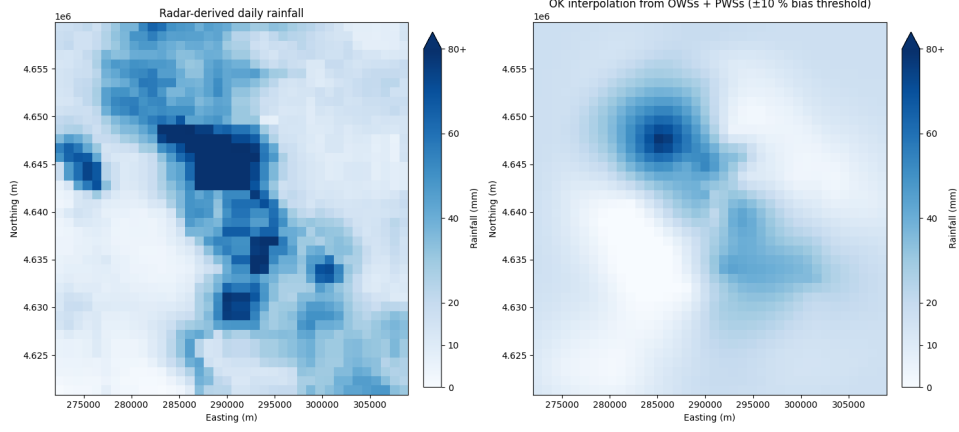


(c) OWSs + PWSs (± 20 %)

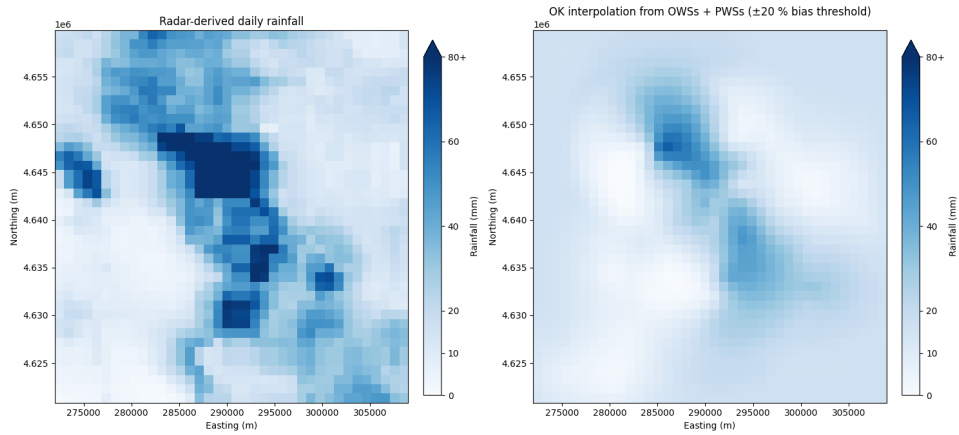
Figure 4.3: Location of (a) OWSs, (b) OWSs with PWSs within ± 10 % (shown in dark blue), and (c) OWSs with PWSs within ± 20 % (shown in light blue) within the Rome administrative boundary (EPSG:32633).



(a) Radar vs. OWS-only



(b) Radar vs. OWS + PWSs ($\pm 10\%$)



(c) Radar vs. OWS + PWSs ($\pm 20\%$)

Figure 4.4: Comparison between radar-derived and interpolated daily rainfall for Rome (8 June 2021): (a) Radar vs. OWSs-only, (b) Radar vs. OWSs with PWSs ($\pm 10\%$), and (c) Radar vs. OWSs with PWSs ($\pm 20\%$) (EPSG:32633).

Table 4.2: Statistical comparison between radar and interpolated rainfall fields for Rome (8 June 2021).

Configuration	ρ	SSIM	r	MAE [mm]	RMSE [mm]	Bias [mm]	NSE
OWS only	0.743	0.412	0.743	12.171	17.776	−10.007	0.290
OWS + PWSs ($\pm 10\%$)	0.720	0.418	0.740	12.261	17.683	−10.029	0.298
OWS + PWSs ($\pm 20\%$)	0.683	0.379	0.711	13.266	19.295	−11.377	0.164

Overall, the results for Rome indicate that the addition of raw PWS observations did not enhance the spatial consistency with radar and, at higher inclusion thresholds, slightly degraded the agreement. This behaviour likely reflects the uneven spatial distribution and variable reliability of raw PWSs within the urban area.

4.3 Discussion and interpretation

The comparative analysis between Turin and Rome demonstrates that the influence of integrating raw PWSs into rainfall interpolation is highly case-dependent. The two metropolitan areas differ both in the structure of their observing networks and in the interpolation approach that best reproduces radar-derived spatial patterns.

In Turin, the official network comprises only a few stations, mostly concentrated in the northern part of the city and along the Po Valley, leaving the southern and western sectors poorly covered. Personal Weather Stations are more numerous in the central and southern districts, thus effectively filling observational gaps and extending spatial representativeness in areas without official gauges. Given this sparse and uneven distribution, the deterministic IDW method provided the most robust and physically consistent rainfall fields, whereas OK tended to oversmooth rainfall pattern (see Appendix). The inclusion of PWSs led to slight improvements in spatial continuity and structural agreement with radar, as reflected by modest increases in Spearman’s ρ and SSIM. Although these changes were not large, they confirm that even raw PWSs can meaningfully enhance spatial representation when official networks are sparse.

In Rome, the situation is markedly different. The official network consists of about 39 stations, providing substantially denser and more homogeneous coverage across most of the municipal area. Many PWSs are located within already well-instrumented zones, especially in the urban core and northern districts, offering limited additional information. Under these conditions, OK produced slightly better agreement with radar and cross-validation results than IDW, suggesting that a variogram-based geostatistical weighting was more effective when a larger number of gauges was available. However, adding raw PWS data did not improve—and in some cases slightly degraded—the spatial correspondence, as indicated by small decreases in ρ and SSIM, particularly for the $\pm 20\%$ inclusion threshold, where measurement uncertainty increases as the selection criterion becomes more permissive. This deterioration likely reflects both the redundancy of PWSs located near OWSs and the variable reliability of individual PWS measurements.

Overall, these results show that the benefit of incorporating raw PWS data depends strongly on the relative density and reliability of the existing network. When PWSs

complement a sparse configuration, as in Turin, they can improve the depiction of spatial variability even without filtering; when the baseline network is already dense and the additional sensors are redundant or noisy, as in Rome, the improvement is negligible or negative. These findings emphasise the need for systematic quality control and station selection prior to their operational use, an aspect addressed in Chapter 5, where the same analyses are repeated after applying the PWS-pyQC algorithm.

Chapter 5

Spatial interpolation and radar validation with quality-controlled PWS data

Unlike official weather stations, which are installed and maintained according to standardized protocols, personal weather stations are highly heterogeneous and often lack consistent metadata or calibration records. This heterogeneity makes it challenging to enforce uniform data-quality standards. Traditional quality-control frameworks typically rely on extensive historical observations and validated reference instruments—resources that are rarely available for citizen-operated networks [42]. Therefore, to ensure the reliability and usability of rainfall data collected from PWSs, the application of a dedicated QC procedure is required.

In recent years, several QC approaches have been developed to address these challenges, each differing in methodology and scope. For instance, some algorithms can operate solely on PWS datasets without requiring primary reference stations such as official rain gauges. Others, rely on reliable official gauges, often referred to as primary stations, to characterize the expected spatial correlation of rainfall and to identify anomalous values. In addition, certain methods identify entire stations as unreliable when systematic inconsistencies or poor agreement with neighboring sensors are detected. More advanced approaches, on the other hand, operate at finer temporal resolutions, flagging only specific time intervals affected by measurement errors while retaining valid observations [35].

The performance of each method depends strongly on local conditions, data completeness, and station density. Consequently, no single QC algorithm can be universally recommended, as their effectiveness is inherently case-dependent and varies with the availability and characteristics of both PWS and official data [35]. In this work, the open-source Python package PWS-pyQC [43] was applied, a widely used QC framework specifically designed for rainfall observations, to the PWS datasets of Turin and Rome. The quality-controlled stations were then integrated with the official weather stations, and the same spatial interpolation and radar cross-validation workflow described in Chapter 4 was repeated to assess whether filtering PWS data improves agreement with radar-derived rainfall fields.

5.1 Methodology

The PWS-pyQC algorithm was applied to the PWS rainfall datasets as provided by its developers. The framework combines statistical and geostatistical analyses to identify unreliable sensors, correct systematic biases, and flag erroneous observations based on their consistency with a trusted reference network [34, 35]. In this study, the official weather station network served as the reference dataset, providing both the spatial dependence structure and the climatological rainfall distribution required for calibration. The workflow consists of three sequential modules: indicator-based filtering, bias correction through quantile mapping, and event-based filtering.

The method assumes that, while PWS magnitudes may be biased, the temporal rank ordering of rainfall at a given PWS is largely reliable for higher intensities. Formally, for two time steps t_1, t_2 at station y_i ,

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

where Y denotes PWS observations and Z the (partly unknown) precipitation [34].

High-intensity indicator-based filtering (IBF)

The IBF identifies secondary stations whose time series show inconsistent rainfall behaviour compared to the reference network. Each station's precipitation record is transformed into a binary indicator series marking exceedances above a selected high-intensity quantile, typically between the 98th and 99th percentiles for hourly data. For the primary and secondary networks, the indicator series are defined as follows:

$$I_{\alpha, \Delta t, Z}(x, t) = \begin{cases} 1 & F_{x, \Delta t}(U_{\Delta t}(x, t)) > \alpha, \\ 0 & \text{otherwise,} \end{cases} \quad I_{\alpha, \Delta t, Y}(y_j, t) = \begin{cases} 1 & G_{y_j, \Delta t}(Y_{\Delta t}(y_j, t)) > \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

where $U_{\Delta t}(x, t)$ and $Y_{\Delta t}(y_j, t)$ denote precipitation at time t aggregated over Δt at reference and PWS locations, respectively, and $F_{x, \Delta t}$, $G_{y_j, \Delta t}$ are their empirical distribution functions [34].

For any two locations x_i and x_j in the primary network, the correlation of their indicator series, $\rho_{Z, \alpha, \Delta t}(x_i, x_j)$, describes the spatial dependence of high-intensity rainfall. For a PWS y_j , the indicator correlation with its nearest reference station, $\rho_{Z, Y, \alpha, \Delta t}(x_i, y_j)$, is compared with the distribution of correlations among primary stations at comparable distances. If the following condition holds:

$$\rho_{Z, Y, \alpha, \Delta t}(x_i, y_j) < \min\{\rho_{Z, \alpha, \Delta t}(x_k, x_m) ; \|(x_k - x_m) - (x_i - y_j)\| < \Delta d\},$$

the secondary station exhibits a weaker association with the reference network than expected and is discarded. This filtering is based solely on the co-occurrence of intense rainfall events and is therefore insensitive to systematic over- or under-estimation biases [34, 35].

Bias correction by quantile mapping

After the IBF, the remaining PWSs are corrected for local bias in rainfall magnitude while preserving their temporal rank order. For each secondary station y_j , the observed rainfall percentile is computed as

$$P_{\Delta t}(y_j, t) = G_{y_j, \Delta t}(Y_{\Delta t}(y_j, t)),$$

and the corresponding precipitation quantiles of the reference network are obtained from the cumulative distributions of the primary stations:

$$Q_{\Delta t}(x_i) = F_{x_i, \Delta t}^{-1}(P_{\Delta t}(y_j, t)).$$

These quantiles are interpolated to the PWS location using ordinary Kriging to yield the corrected precipitation amount

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

where λ_i are Kriging weights. This transformation replaces the raw PWS rainfall value with an adjusted amount consistent with the climatological distribution of the reference network while maintaining the relative ranking of events [34, 35].

Event-based spatial filtering (EBF)

Even after bias correction, individual PWS readings can deviate substantially from neighbouring observations because of transmission errors or false zeros. To identify such outliers, a geostatistical consistency check is performed at each time step. For every corrected PWS value $Z_{\Delta t}^o(y_j, t)$, the expected rainfall is estimated by ordinary Kriging from all reference stations (excluding the PWS itself):

$$Z_{\Delta t}^*(y_j, t) = \text{OK}(Z_{\Delta t}(x_i, t)),$$

with associated Kriging standard deviation $\sigma_{\Delta t}(y_j, t)$. An observation is flagged as inconsistent when the standardized residual exceeds three standard deviations:

$$\left| \frac{Z_{\Delta t}^*(y_j, t) - Z_{\Delta t}^o(y_j, t)}{\sigma_{\Delta t}(y_j, t)} \right| > 3.$$

Measurements satisfying equation above are removed for that time step, while the station itself remains in the dataset. This procedure effectively eliminates transmission spikes or false zero values and retains the overall spatial pattern of reliable measurements. Only a very small fraction of PWS records—typically about 1 % or less—are rejected by this geostatistical outlier filtering, indicating that this step fine-tunes data quality by discarding only the most evident errors [34, 35, 44].

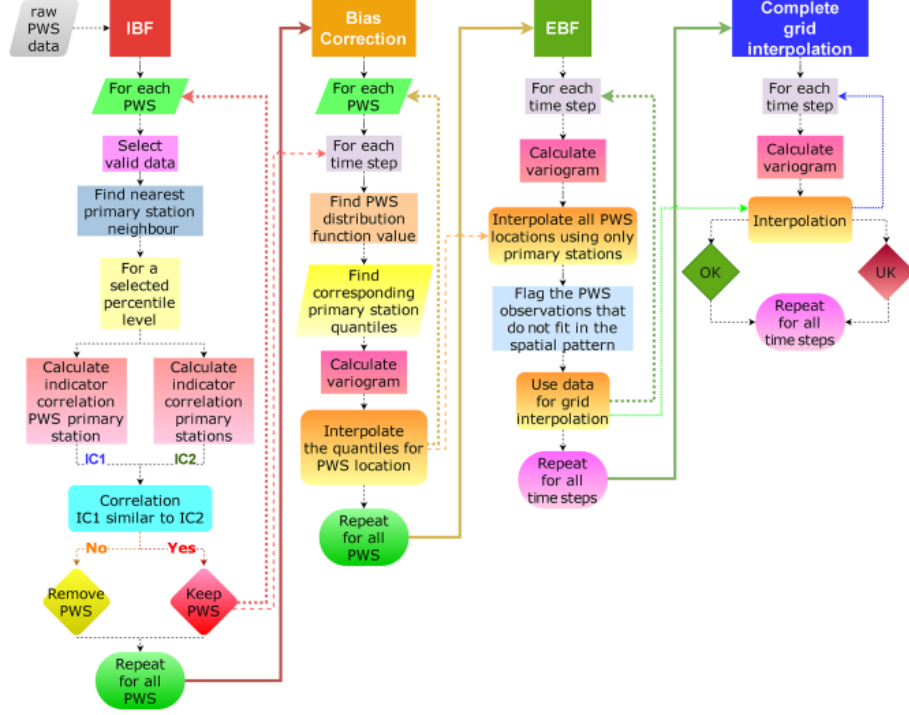


Figure 5.1: Workflow of the PWS-pyQC quality-control framework showing its three main modules: Indicator-Based Filtering, Bias Correction, and Event-Based Filtering. The diagram illustrates the sequential processing from raw PWS data to the final interpolated rainfall grid. Adapted from Bárdossy *et al.* (2020).

After the QC pipeline, the filtered and bias-corrected PWS datasets were merged with the official gauge observations. Spatial interpolation and radar cross-validation were then repeated using the same configuration and performance metrics adopted in Chapter 4. This allowed the assessment of whether the quality-controlled PWS data improved the spatial representation of rainfall and its agreement with radar-derived estimates.

5.2 Spatial comparison with radar data

This section presents the spatial validation of rainfall fields derived from gauge-based interpolation after applying the PWS-pyQC pipeline. The analysis follows the same structure and methodology of Chapter 4, focusing on the flash-flood events of 22 June 2021 (Turin) and 8 June 2021 (Rome). For both case studies, the interpolation grid (1 km²) and the valid-data mask were kept identical to Chapter 4 to enable one-to-one comparison. The statistical evaluation employs the same metrics— ρ , SSIM, r , MAE, RMSE, Bias, and NSE—with emphasis on pattern-based indicators, since the RADAR-DPC fields are not gauge-adjusted.

5.2.1 Turin — 22 June 2021

Figure 5.2 shows the spatial configuration of OWSs and the PWSs that passed all three stages of the PWS-pyQC workflow (indicator-based filtering, quantile-mapping bias correction, and event-based filtering). The retained QC-PWSs complement the sparser official network, particularly across the central and southern districts that were under-sampled with OWSs alone.

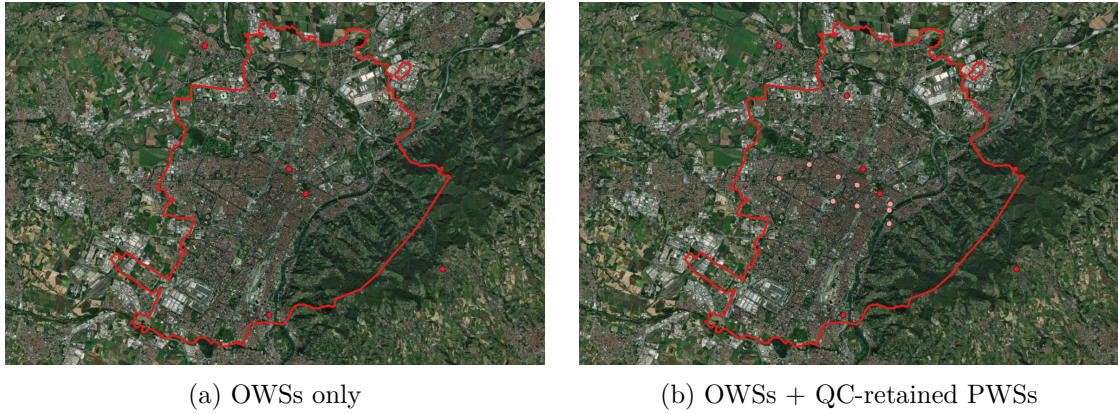


Figure 5.2: Location of the stations used for interpolation within the Turin administrative boundary (EPSG:32632): (a) OWSs only and (b) OWSs with QC-retained PWSs (shown in pink).

Rainfall fields were interpolated using IDW, as in Chapter 4. Figure 5.3 compares the radar daily accumulation with the IDW interpolations obtained using only OWSs and using the combined dataset of OWSs and PWSs that passed the quality control. Identical colour limits ($0\text{--}100\text{ mm day}^{-1}$) was applied to ensure a direct visual comparison.

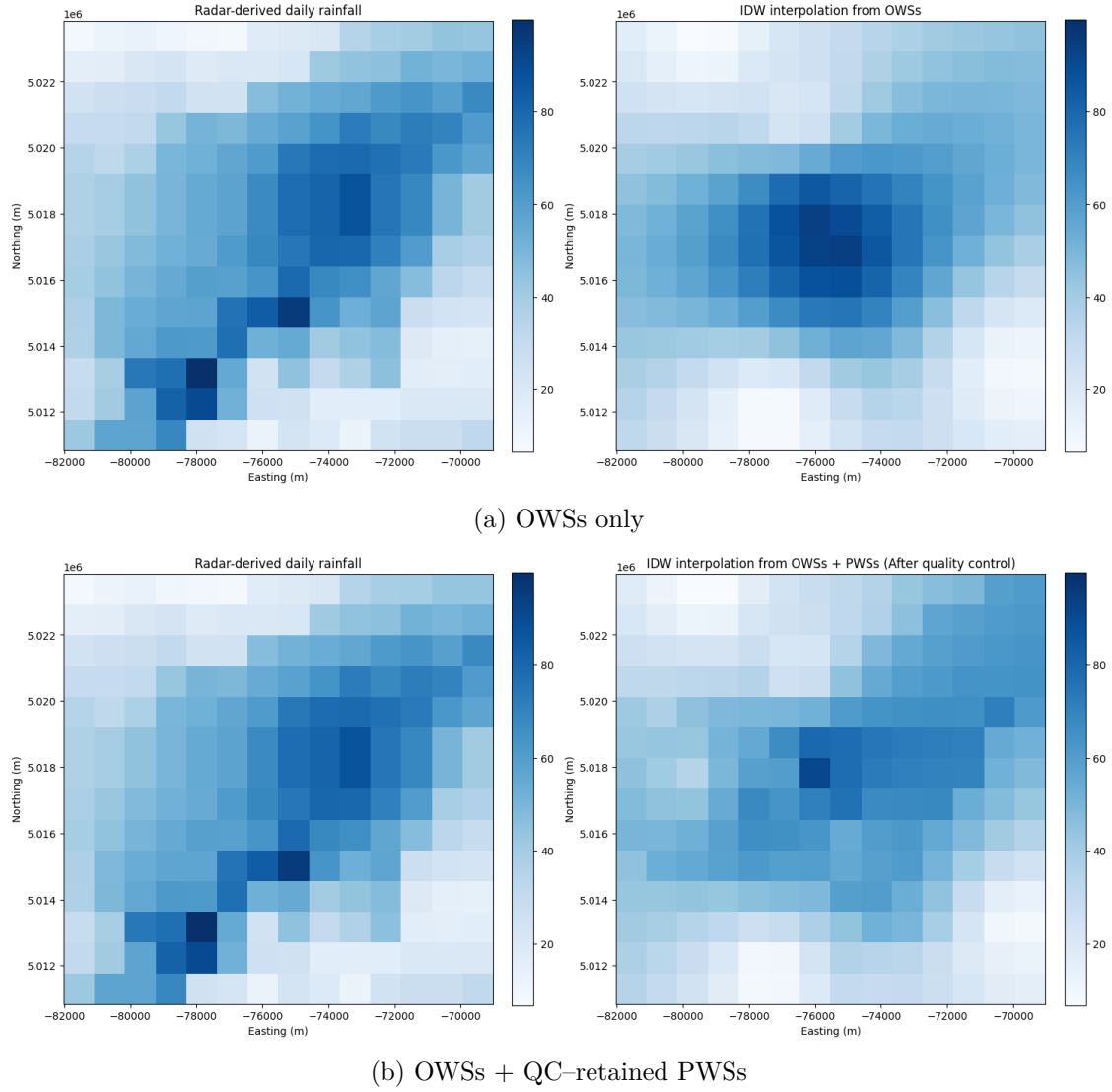


Figure 5.3: Comparison between radar-derived and interpolated daily rainfall for Turin (22 June 2021). Radar-derived daily accumulation (left panels) vs. IDW interpolation (right panels): (a) OWSs-only and (b) OWSs + QC-retained PWSs.

Table 5.1: Turin (22 June 2021): radar vs. interpolated fields (IDW). Metrics computed on the common valid-data mask.

Configuration	ρ	SSIM	r	MAE [mm]	RMSE [mm]	Bias [mm]	NSE
OWS only	0.710	0.616	0.683	11.873	17.195	-3.533	0.325
OWS + QC-PWS	0.742	0.648	0.711	10.159	15.401	-2.773	0.458

The inclusion of QC-PWSs improved all principal indicators. Pattern agreement

strengthened (ρ : 0.710→0.742; SSIM: 0.616→0.648), while error measures decreased (MAE: 11.9→10.2 mm; RMSE: 17.2→15.4 mm). NSE increased from 0.325 to 0.458, and the absolute bias reduced in magnitude. These improvements indicate a more consistent correspondence between the interpolated rainfall field and the spatial distribution observed by radar. Visually (Fig. 5.3), the QC-enhanced interpolation better delineates the rainfall core over northern Turin and suppresses peripheral artefacts present in the OWS-only field.

5.2.2 Rome — 8 June 2021

Figure 5.4 presents the distribution of OWSs and QC-retained PWSs within the Rome boundary. The preserved PWSs are mainly concentrated in the central and southern sectors, adding observation points to an already dense official network.

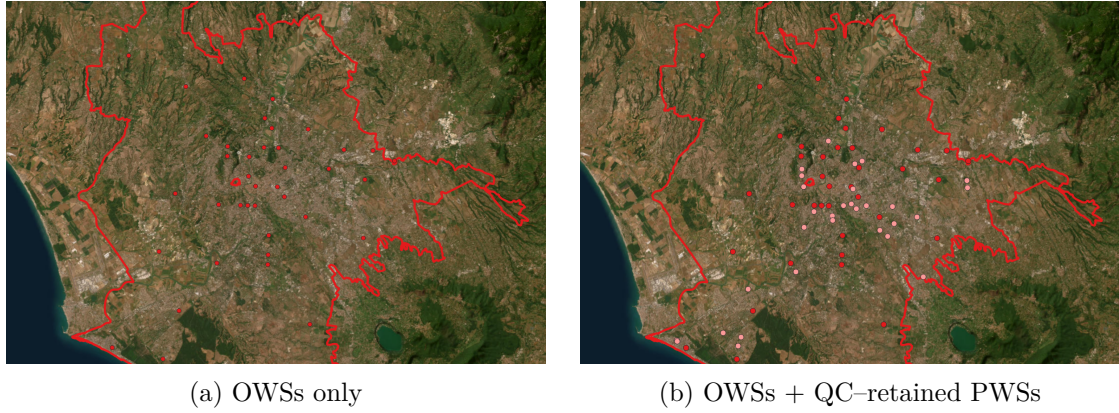


Figure 5.4: Location of the stations used for interpolation within the Rome administrative boundary (EPSG:32633): (a) OWSs only and (b) OWSs + QC-retained PWSs (shown in pink).

Interpolation was performed using OK, consistent with Chapter 4. Figure 5.5 compares the radar accumulation with the OK interpolations from using only OWSs and using the combined dataset of OWSs and PWSs that passed the quality control.

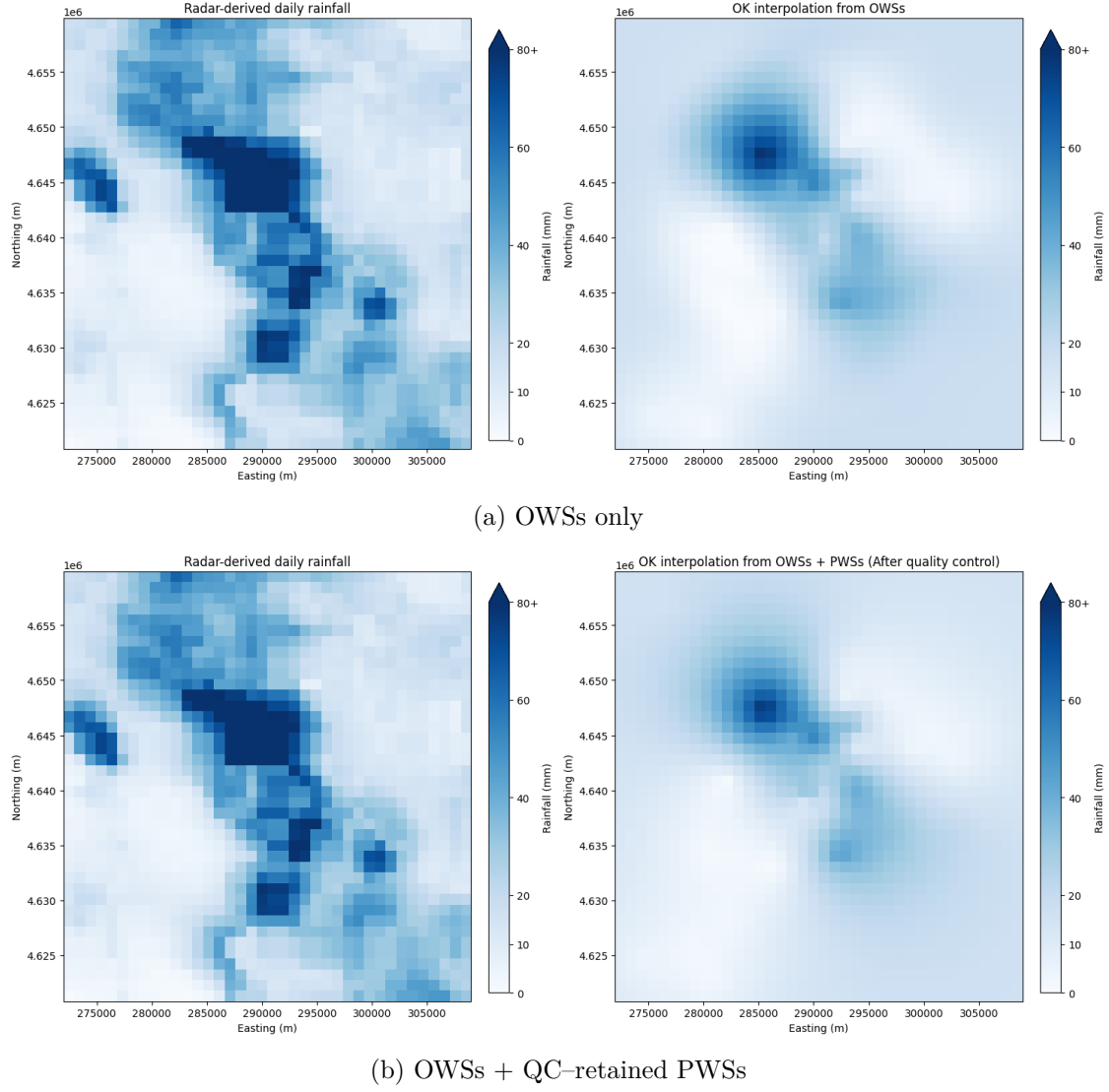


Figure 5.5: Comparison between radar-derived and interpolated daily rainfall for Rome (8 June 2021). Radar-derived daily accumulation (left panels) vs. OK interpolation (right panels): (a) OWSs-only and (b) OWSs + QC-retained PWSs.

Table 5.2: Rome (8 June 2021): radar vs. interpolated fields (OK). Metrics computed on the common valid-data mask.

Configuration	ρ	SSIM	r	MAE [mm]	RMSE [mm]	Bias [mm]	NSE
OWS only	0.743	0.412	0.743	12.171	17.776	-10.007	0.290
OWS + QC-PWS	0.812	0.439	0.787	12.142	17.813	-10.779	0.287

In Rome, the integration of QC-filtered PWSs produced a moderate improvement in

pattern-based metrics (increase of ρ from 0.743 to 0.812 and SSIM from 0.412 to 0.439), while absolute error metrics and NSE changed minimally. Visual inspection suggests a slightly sharper representation of the north-western rainfall core after QC, but the overall spatial structure remains close to the OWSs-only field, reflecting the already high density and quality of the official network.

5.3 Discussion and interpretation

The results obtained after applying the PWS-pyQC procedure confirm that systematic filtering and bias correction significantly enhance the spatial usability of crowdsourced rainfall observations. In Turin, where the official station network is sparse, integrating QC-retained PWSs led to consistent gains across all metrics. Notably, Spearman’s rank correlation and the Structural Similarity Index increased, indicating improved spatial coherence relative to the radar reference. RMSE and bias (interpreted qualitatively) also decreased, and the NSE improved, suggesting better pattern alignment with the radar field. The rainfall maxima were more clearly delineated in the interpolated maps, reflecting the added spatial detail from PWS integration.

In Rome, improvements were more modest and primarily confined to pattern-based metrics. Given the already dense and evenly distributed OWSs, the additional PWSs contributed limited new spatial information. Nonetheless, the QC step effectively filtered out anomalous PWS records, helping to preserve the consistency of the interpolated field.

Relative to the results in Chapter 4, the quality-controlled datasets consistently yielded higher ρ and SSIM values in both cities, reinforcing the importance of pre-processing. These gains demonstrate that PWS integration can enhance spatial rainfall mapping—particularly in undersampled areas—provided that robust QC procedures are applied. It is important to note, however, that the radar accumulations used here are not gauge-adjusted and are subject to their own uncertainties. As such, improvements in absolute error metrics (e.g., RMSE, bias, NSE) are interpreted cautiously and primarily reflect increased spatial agreement with the radar pattern, rather than definitive accuracy gains.

Overall, the PWS-pyQC workflow successfully retained the informative portion of the PWS network while suppressing inconsistent outliers. This finding supports the broader conclusion that crowdsourced sensors—when rigorously filtered—can be effectively integrated into operational urban rainfall monitoring, particularly for enhancing the spatial structure of rainfall estimates during high-impact events.

Chapter 6

Discussion, conclusions, and future perspectives

Crowdsourced rainfall data acquired from thousands of personal weather stations have the potential to fill observational gaps in official rain gauge networks, particularly where network density has declined over recent decades. When properly quality-controlled, PWS observations can complement existing networks, improving spatial interpolation and hydrological model calibration [23, 30, 33, 34, 45]. Their high spatial density enables a detailed representation of rainfall variability, which is especially relevant in complex urban environments [20]. Maintaining dense professional networks in cities is costly and logistically challenging, whereas PWSs are low-cost and easily deployed by individuals or communities, allowing large-scale participation at minimal expense [29].

This thesis evaluated the potential of crowdsourced PWSs to enhance rainfall monitoring in urban areas through a multi-stage analysis. Reliability was first assessed across eleven Italian cities (2021–2022), revealing a consistent underestimation at high rainfall intensities—also reported in previous studies [29, 31, 38]. Extending this assessment to cities with different climatic regimes confirmed that the underestimation trend is widespread, though its magnitude varies by context. Two representative case studies, Turin and Rome, were then analysed to examine how integrating PWS data with official networks influences spatial rainfall estimation. Interpolated rainfall maps derived from raw and quality-controlled datasets were compared with unadjusted RADAR–DPC accumulations—used here as a spatial benchmark for assessing pattern agreement—by applying inverse distance weighting and ordinary kriging. The application of the open-source PWS-pyQC algorithm improved spatial pattern correspondence, particularly in data-sparse conditions, and restored the consistency of interpolation in Rome, where raw PWS data had previously degraded agreement with radar. These findings highlight both the opportunities and the limitations of integrating citizen-collected observations into hydrological analyses, demonstrating their value for strengthening rainfall monitoring in data-scarce urban environments.

Together, these analyses address the research questions defined in Chapter 1. The nationwide comparison quantified the magnitude and consistency of PWS measurement bias across diverse climatic contexts. Applying the PWS-pyQC algorithm demonstrated

its effectiveness in improving the reliability of crowdsourced data through rigorous station-level filtering. The Turin and Rome case studies further showed that integrating quality-controlled PWSs enhances the spatial representation of rainfall relative to unadjusted radar fields, with the greatest benefits occurring under sparse official network conditions.

The quality control applied in this study was intentionally conservative, excluding entire stations that failed spatial-dependence consistency checks rather than flagging individual erroneous measurements. This approach improved the reliability of retained data but reduced spatial coverage in some areas, affecting local representativeness. The performance of any quality-control algorithm also depends on network density and data availability; future studies could therefore compare alternative open-source frameworks such as PWSQC or GSDR-QC to evaluate methodological sensitivity [35]. The comparison with radar fields relied on unadjusted RADAR–DPC accumulations, so error metrics were interpreted in terms of spatial correspondence rather than absolute rainfall accuracy. Furthermore, only two single-day flash-flood events were analysed, and interpolation parameters were fixed for each city, without multi-event cross-validation.

Future research should extend similar analyses across a wider range of events and climatic contexts to assess robustness. Comparative benchmarking of different quality-control algorithms and coupling quality-controlled PWS data with radar–gauge adjustment or machine-learning-based data fusion could further enhance rainfall field accuracy. Beyond methodological advances, addressing the social and spatial dimensions of citizen science remains crucial. Although the growing number of PWSs increases the potential for fine-scale rainfall monitoring, their distribution often reflects social and geographic biases—densities tend to be higher in wealthier or flood-prone neighbourhoods [22]. In addition, network sustainability is not guaranteed: a “peak Netatmo” effect has been observed, where participation may plateau or decline [46], and key institutional actors, such as emergency services and local authorities, are still insufficiently involved [47]. Strengthening institutional engagement and integrating PWS networks into public monitoring frameworks will therefore be essential for long-term continuity and reliability.

Overall, the outcomes of this research demonstrate that crowdsourced personal weather stations, when properly quality-controlled, can meaningfully contribute to improving the spatial representation of rainfall in urban areas. By combining a systematic evaluation of PWS reliability across multiple cities with event-based validation against radar data, this work provides new evidence on both the potential and the constraints of citizen-collected observations. The findings confirm that the benefit of integrating PWS data depends strongly on network density and data quality, emphasizing the importance of rigorous quality control prior to their hydrological use. Ultimately, this study contributes to the growing field of citizen hydrometeorology by providing a practical framework for assessing, filtering, and integrating PWS rainfall data to support improved rainfall monitoring and, in future studies, applications such as urban flood-risk assessment.

Appendix A

Additional figures and sensitivity tests

A.1 Example of rejected OK interpolation from OWS-only configuration

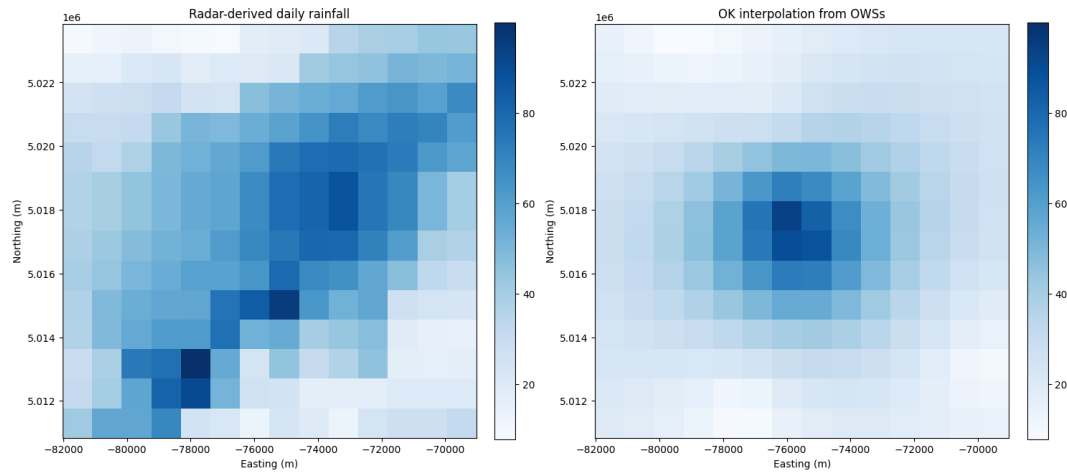


Figure A.1: Radar-derived daily rainfall (left) and example of ordinary kriging interpolation using only official weather stations (right) for the Turin case study. The OK realisation shown here was discarded due to its weak agreement with the radar pattern and its artificially smoothed spatial structure.

Bibliography

- [1] Chris Kidd, Andreas Becker, George J Huffman, Catherine L Muller, Paul Joe, Gail Skofronick-Jackson, and Dalia B Kirschbaum. So, how much of the Earth’s surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98(1):69–78, 2017.
- [2] Jennie Barron and Jessica Calfoforo Salas. *Rainwater harvesting: a lifeline for human well-being*. UNEP/Earthprint, 2009.
- [3] Yves Trambly, Maria Carmen Llasat, Christophe Randin, and Erika Coppola. Climate change impacts on water resources in the Mediterranean. *Regional Environmental Change*, 20(3):83, 2020.
- [4] Marco Luppichini, Monica Bini, Roberto Giannecchini, and Giovanni Zanchetta. High-resolution spatial analysis of temperature influence on the rainfall regime and extreme precipitation events in north-central Italy. *Science of The Total Environment*, 880:163368, 2023.
- [5] Sebastiaan N Jonkman. Global perspectives on loss of human life caused by floods. *Natural hazards*, 34(2):151–175, 2005.
- [6] Benedetta Moccia, Elena Ridolfi, Claudio Mineo, Fabio Russo, and Francesco Napolitano. On the occurrence of extreme rainfall events across Italy: Should we update the probability of failure of existing hydraulic works? *Water Resources Management*, 38(11):4069–4082, 2024.
- [7] Irene Garcia-Marti, Aart Overeem, Jan Willem Noteboom, Lotte de Vos, Marijn de Haij, and Kirien Whan. From proof-of-concept to proof-of-value: Approaching third-party data to operational workflows of national meteorological services. *International Journal of Climatology*, 43(1):275–292, 2023.
- [8] Paola Mazzoglio, Ilaria Butera, and Pierluigi Claps. I2-RED: a massive update and quality control of the Italian annual extreme rainfall dataset. *Water*, 12(12):3308, 2020.
- [9] Aart Overeem, Hidde Leijnse, and Remko Uijlenhoet. Country-wide rainfall maps from cellular communication networks. *Proceedings of the National Academy of Sciences*, 110(8):2741–2745, 2013.
- [10] Aart Overeem, JC R. Robinson, Hidde Leijnse, Gert-Jan Steeneveld, BK P. Horn, and Remko Uijlenhoet. Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters*, 40(15):4081–4085, 2013.
- [11] Lotte De Vos, Hidde Leijnse, Aart Overeem, and Remko Uijlenhoet. The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrology and Earth System Sciences*, 21(2):765–777, 2017.

- [12] Chris Kidd and George Huffman. Global precipitation measurement. *Meteorological Applications*, 18(3):334–353, 2011.
- [13] Philip D Jones, Keith R Briffa, Timothy J Osborn, Anders Moberg, and Hans Bergström. Relationships between circulation strength and the variability of growing-season and cold-season climate in northern and central Europe. *The Holocene*, 12(6):643–656, 2002.
- [14] M Kitchen and PM Jackson. Weather radar performance at long range simulated and observed. *Journal of Applied Meteorology and Climatology*, 32(5):975–985, 1993.
- [15] Witold F Krajewski, Martha C Anderson, William E Eichinger, Dara Entekhabi, Brian K Hornbuckle, Paul R Houser, Gabriel G Katul, William P Kustas, John M Norman, Christa Peters-Lidard, et al. A remote sensing observatory for hydrologic sciences: A genesis for scaling to continental hydrology. *Water Resources Research*, 42(7), 2006.
- [16] Thomas M Smith, Phillip A Arkin, John J Bates, and George J Huffman. Estimating bias of satellite-based precipitation estimates. *Journal of Hydrometeorology*, 7(5):841–856, 2006.
- [17] CL Muller, Lee Chapman, Samuel Johnston, Chris Kidd, Samuel Illingworth, Giles Foody, Aart Overeem, and RR Leigh. Crowdsourcing for climate and atmospheric sciences: current status and future potential. *International Journal of Climatology*, 35(11):3185–3203, 2015.
- [18] Andrew T Campbell, Shane B Eisenman, Nicholas D Lane, Emiliano Miluzzo, and Ronald A Peterson. People-centric urban sensing. In *Proceedings of the 2nd annual international workshop on Wireless internet*, pages 18–es, 2006.
- [19] Matthew F McCabe, Matthew Rodell, Douglas E Alsdorf, Diego G Miralles, Remko Uijlenhoet, Wolfgang Wagner, Arko Lucieer, Rasmus Houborg, Niko EC Verhoest, Trenton E Franz, et al. The future of Earth observation in hydrology. *Hydrology and earth system sciences*, 21(7):3879–3914, 2017.
- [20] Claudia Hahn, Irene Garcia-Marti, Jacqueline Sugier, Fiona Emsley, Anne-Lise Beaulant, Louise Oram, Eva Strandberg, Elisa Lindgren, Martyn Sunter, and Franziska Ziska. Observations from personal weather stations—EUMETNET interests and experience. *Climate*, 10(12):192, 2022.
- [21] Radosław Drożdżoń, Damian Absalon, and Ewa Łupikasza. The possibility of using personal weather station networks to verify and evaluate local extreme phenomena. In *AIP Conference Proceedings*, volume 2186, page 120003. AIP Publishing LLC, 2019.
- [22] Alexander B Chen, Jonathan L Goodall, T Donna Chen, and Zihao Zhang. Flood resilience through crowdsourced rainfall data collection: Growing engagement faces non-uniform spatial adoption. *Journal of Hydrology*, 609:127724, 2022.
- [23] Kay Khaing Kyaw, Emma Baietti, Cristian Lussana, Valerio Luzzi, Paolo Mazzoli, Stefano Bagli, and Attilio Castellarin. Private sensors and crowdsourced rainfall data: Accuracy and potential for modelling pluvial flooding in urban areas of Oslo, Norway. *Journal of Hydrology X*, 25:100191, 2024.
- [24] JM Nielsen, CZR van de Beek, S Thorndahl, Jonas Olsson, CB Andersen, JCM

- Andersson, MR Rasmussen, and JE Nielsen. Merging weather radar data and opportunistic rainfall sensor data to enhance rainfall estimates. *Atmospheric Research*, 300:107228, 2024.
- [25] Aart Overeem, Hidde Leijnse, Gerard van der Schrier, Else van den Besselaar, Irene Garcia-Marti, and Lotte Wilhelmina de Vos. Merging with crowdsourced rain gauge data improves pan-European radar precipitation estimates. *Hydrology and Earth System Sciences*, 28(3):649–668, 2024.
- [26] Christine Sgoff, Walter Acevedo, Zoi Paschalidi, Sven Ulbrich, Elisabeth Bauernschubert, Thomas Kratzsch, and Roland Potthast. Assimilation of crowd-sourced surface observations over Germany in a regional weather prediction system. *Quarterly Journal of the Royal Meteorological Society*, 148(745):1752–1767, 2022.
- [27] Ming Li, Quanxi Shao, Joel Janek Dabrowski, Ashfaqur Rahman, Andrea Powell, Brent Henderson, Zachary Hussain, and Peter Steinle. An automatic quality evaluation procedure for third-party daily rainfall observations and its application over Australia. *Stochastic Environmental Research and Risk Assessment*, 37(7):2473–2493, 2023.
- [28] Simon Bell, Dan Cornford, and Lucy Bastin. How good are citizen weather stations? Addressing a biased opinion. *Weather*, 70(3):75–84, 2015.
- [29] LW De Vos, AM Droste, MJ Zander, Aart Overeem, H Leijnse, BG Heusinkveld, GJ Steeneveld, and R Uijlenhoet. Hydrometeorological monitoring using opportunistic sensing networks in the Amsterdam metropolitan area. *Bulletin of the American Meteorological Society*, 101(2):E167–E185, 2020.
- [30] Nathalie Rombeek, Markus Hrachowitz, Arjan Droste, and Remko Uijlenhoet. Evaluation of high-intensity rainfall observations from personal weather stations in the Netherlands. *EGUsphere*, 2024:1–25, 2024.
- [31] Cristian Lussana, Emma Baietti, Line Båserud, Thomas Nils Nipen, and Ivar Ambjørn Seierstad. Exploratory analysis of citizen observations of hourly precipitation over Scandinavia. *Advances in Science and Research*, 20:35–48, 2023.
- [32] Priyadharshini Sakthivel and Raja Sengupta. Spatial bias in placement of citizen and conventional weather stations and their impact on urban climate research: A case study of the Urban Heat Island effect in Canada. *Urban Climate*, 59:102280, 2025.
- [33] Lotte Wilhelmina de Vos, Hidde Leijnse, Aart Overeem, and Remko Uijlenhoet. Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. *Geophysical Research Letters*, 46(15):8820–8829, 2019.
- [34] András Bárdossy, Jochen Seidel, and Abbas El Hachem. The use of personal weather station observation for improving precipitation estimation and interpolation. *Hydrology and Earth System Sciences Discussions*, 2020:1–23, 2020.
- [35] Abbas El Hachem, Jochen Seidel, Tess O’hara, Roberto Villalobos Herrera, Aart Overeem, Remko Uijlenhoet, András Bárdossy, and Lotte De Vos. A guide to using three open-source quality control algorithms for rainfall data from personal weather stations. *Hydrology and Earth System Sciences*, 28(20):4715–4731, 2024.
- [36] Alan Demortier, Marc Mandement, Vivien Pourret, and Olivier Caumont. Assimilation of temperature and relative humidity observations from personal weather

- stations in AROME-France. *Natural Hazards and Earth System Sciences*, 25(1):429–449, 2025.
- [37] Felix D Schönbrodt and Marco Perugini. At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5):609–612, 2013.
- [38] Robert J Lascano, Timothy S Goebel, Dennis C Gitz III, and John E Stout. Evaluation of a Wireless Solar Powered Personal Weather Station. *Agricultural Sciences*, 15(1):36–53, 2024.
- [39] Alice Crespi, Michael Matiu, Giacomo Bertoldi, Marcello Petitta, and Marc Zebisch. A high-resolution gridded dataset of daily temperature and precipitation records (1980–2018) for trentino–South Tyrol (North-Eastern Italian Alps). *Earth System Science Data Discussions*, 2021:1–27, 2021.
- [40] Dariela A Vázquez-Rodríguez, Víctor H Guerra-Cobián, José L Bruster-Flores, Carlos R Fonseca, and Fabiola D Yépez-Rincón. Evaluating the performance and applicability of satellite precipitation products over the rio grande–san juan basin in northeast Mexico. *Atmosphere*, 15(7):749, 2024.
- [41] V Houssou and J Carreau. Spatial pattern regression for gridded meteorological data: A precipitation and temperature case study. *Les Cahiers du GERAD ISSN*, 711:2440, 2025.
- [42] Alexander B Chen, Madhur Behl, and Jonathan L Goodall. Trust me, my neighbors say it’s raining outside: Ensuring data trustworthiness for crowdsourced weather stations. In *Proceedings of the 5th Conference on Systems for Built Environments*, pages 25–28, 2018.
- [43] Abbas El Hachem. pws-pyqc: Personal weather station (pws) quality control using python, 2020.
- [44] Maximilian Graf, Abbas El Hachem, Micha Eisele, Jochen Seidel, Christian Chwala, Harald Kunstmann, and András Bárdossy. Rainfall estimates from opportunistic sensors in Germany across spatio-temporal scales. *Journal of Hydrology: Regional Studies*, 37:100883, 2021.
- [45] Tess O’Hara, Fergus McClean, Roberto Villalobos Herrera, Elizabeth Lewis, and Hayley J Fowler. Filling observational gaps with crowdsourced citizen science rainfall data from the Met Office Weather Observation Website. *Hydrology Research*, 54(4):547–556, 2023.
- [46] Lee Chapman, Simon Bell, and Sophie Randall. Can crowdsourcing increase the durability of an urban meteorological network? *Urban Climate*, 49:101542, 2023.
- [47] Mohammad Gharesifard, Uta Wehn, and Pieter van der Zaag. Towards benchmarking citizen observatories: Features and functioning of online amateur weather networks. *Journal of Environmental Management*, 193:381–393, 2017.