score

D4.6- Validation of citizen science data

DATE OF DELIVERY - DD/MM/YYYY

AUTHORS Saul Crowley (UCD), José Gómez Barrón (UCD), Morgane Miss (UCD), Domitille Deschamps (UCD), Luca Baldini (CNR), Tasneem Ahmed (ATU), Arghadyuti Banerjee (ATU), Giovanni Serafino (MBI), Andrea Rucci (MBI)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003534

DOCUMENT TRACKS DETAILS

Project acronym	SCORE
Project title	Smart Control of the Climate Resilience in European Coastal Cities
Starting date	01.07.2021
Duration	48 months
Call identifier	H2020-LC-CLA-2020-2
Grant Agreement No	101003534

Deliverable Information	
Deliverable number	D4.6
Work package number	WP4
Deliverable title	Validation of citizen science data
Lead beneficiary	University College Dublin
Author(s)	Saul Crowley (UCD), José Gómez Barrón (UCD), Morgane Miss (UCD), Domitille Deschamps (UCD), Luca Baldini (CNR), Tasneem Ahmed (ATU), Arghadyuti Banerjee (ATU), Giovanni Serafino (MBI), Andrea Rucci (MBI)
Due date	30/08/2024
Actual submission date	30/08/2024
Type of deliverable	Report
Dissemination level	Confidential

VERSION MANAGEMENT

Revision table			
Version	Name	Date	Description
V 0.1	Saul Crowley, UCD	09/08/2024	First draft
V 0.2	Cécil Meulenberg, ZRS Koper	12/08/2024	Updated draft internally reviewed
V 0.3	Filippo Giannetti, UNIPI	22/08/2024	Updated draft after contribution from partners
V 0.4	Saul Crowley, UCD	26/08/2024	Final updates after all contributions
V1.0	Iulia Anton, Salem Gharbia, ATU	27/08/2024	Final version

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym / Abbreviation	Meaning / Full text
BTE	Built-In Test Equipment
CCLL	Coastal City Living Lab
DT-EWS	Data-Twin Early Warning System
EWSS	Early Warning Support System
GUI	Graphical User Interface
MAD	Median Deviation Absolute
RMSE	Root Mean Square Error
STAPI	Sensor Things Application Programming Interface
SIP	SCORE ICT Platform
WMS	World Meteorological Organization

BACKGROUND: ABOUT THE SCORE PROJECT

SCORE is a four-year EU-funded project aiming to increase climate resilience in European coastal cities.

The intensification of extreme weather events, coastal erosion and sea-level rise are major challenges to be urgently addressed by European coastal cities. The science behind these disruptive phenomena is complex, and advancing climate resilience requires progress in data acquisition, forecasting, and understanding of the potential risks and impacts for real-scenario interventions. The Ecosystem-Based Approach (EBA) supported by smart technologies has potential to increase climate resilience of European coastal cities; however, it is not yet adequately understood and coordinated at European level.

SCORE outlines a co-creation strategy, developed via a network of 10 coastal city 'living labs' (CCLLs), to rapidly, equitably and sustainably enhance coastal city climate resilience through EBAs and sophisticated digital technologies.

The 10 coastal city living labs involved in the project are: Sligo and Dublin, Ireland; Barcelona/Vilanova i la Geltrú, Benidorm and Basque Country, Spain; Oeiras, Portugal; Massa, Italy; Piran, Slovenia; Gdansk, Poland; Samsun, Turkey.

SCORE will establish an integrated coastal zone management framework for strengthening EBA and smart coastal city policies, creating European leadership in coastal city climate change adaptation in line with The Paris Agreement. It will provide innovative platforms to empower stakeholders' deployment of EBAs to increase climate resilience, business opportunities and financial sustainability of coastal cities.

The SCORE interdisciplinary team consists of 28 world-leading organisations from academia, local authorities, RPOs, and SMEs encompassing a wide range of skills including environmental science and policy, climate modelling, citizen and social science, data management, coastal management and engineering, security and technological aspects of smart sensing research.

EXECUTIVE SUMMARY

This document is a deliverable of the SCORE project, funded under the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003534.

The aim of this document is to outline the dual objectives essential to the success of the SCORE project: first, providing an overview of the methods and practices for validating citizen science data, and second, integrating these validated data with institutional monitoring systems. Validating citizen science data is crucial to ensuring the accuracy and reliability of information collected by non-professional volunteers. Once validated, this data can be integrated into institutional systems to enhance environmental monitoring, particularly in urban coastal areas where detailed and localised data can significantly improve monitoring capabilities and decision-making.

The integration of low-cost, widely distributed citizen science sensors offers significant benefits by providing detailed and localised environmental data, complementing existing institutional networks. However, the report also highlights the need for rigorous validation processes to ensure the reliability and accuracy of this data. In detail, it outlines various validation techniques, including calibration, cross-validation with local reference instruments, and the use of consistency checks to address different types of errors.

At the core of the validation process is the SCORE ICT Platform (SIP), which acts as a centralised hub for data management and sharing, real-time interactions, and stakeholder communication. The report details how the SIP will employ consistency checks —temporal, spatial, and climatological— as well as collocation methods like dual and triple collocation to maintain data integrity and filter out anomalies.

In addition, the Early Warning Support system (EWSS) is discussed, emphasising how it benefits from the inclusion of validated citizen science data to improve flood risk assessments and other environmental predictions. The report stresses the importance of regular calibration, a robust validation framework, and the autonomous management of data integrity by the EWSS to ensure that only reliable information is used for critical environmental assessments.

Finally, the report recommends investing in capacity-building initiatives for citizen scientists, including training on data collection and sensor maintenance, and establishing a clear framework for integrating citizen science data with institutional datasets. This approach aims to maximise the potential of citizen science data while ensuring high standards of data quality and reliability.

LINKS WITH OTHER PROJECT ACTIVITIES

The data generated from citizen science activities, as described in WP4, plays a crucial role in the work outlined in WP5 and WP8. WP5 "Pre/post-EBA Interventions Evidence Collection and Knowledge Marketplace," focuses on creating and implementing the SCORE ICT Platform. This platform is designed to centralise and manage various data sources, including those from citizen science efforts. It supports the integration, analysis, and sharing of this data to enhance environmental knowledge and decision-making.

Similarly, WP8, "Development of Integrated Early Warning Support and Spatial Digital Twin Solution Prototypes," is dedicated to developing a GIS-based Early Warning Support system (EWSS) and Digital Twin platform. This system leverages the data collected from various sources, including citizen science, to provide real-time environmental monitoring and predictive capabilities.

TABLE OF CONTENT

INDEX OF TABLES

1. INTRODUCTION

1.1. Background and Objectives

This report "D4.6: Validation of Citizen Science Data. Report on Validation Algorithms for Data Collected During Citizen Science Activities'' addresses the essential task of ensuring the quality and reliability of data collected through citizen science initiatives using low-cost sensing technology. These sensors are deployed in various Coastal City Living Labs (CCLLs) as part of a broader effort to enhance climate resilience in European coastal cities. The process for selecting these low-cost sensors has been presented in 'D4.2-Report on low-cost sensors viable for citizen science activities.' [1] The SCORE project framework for citizen science has been detailed previously in 'D4.3-Citizen science DIY framework' [2], while specific citizen science activities in each CCLL will be showcased in 'D4.5-Citizen science activities in CCLLs' (forthcoming).

The emergence of citizen science as a valuable tool for environmental monitoring creates the opportunity for innovative approaches to data collection. In this project, low-cost sensors are employed by citizen scientists to gather data that is complementary to the high-quality data collected by institutional sensors. These low-cost sensors, while accessible and easy to deploy, often require additional validation to ensure their data is accurate, reliable, and useful for scientific and policy-making purposes.

Task 4.5 focuses on developing and implementing algorithms to validate the data from these low-cost sensors. The goal is to evaluate their potential ability to supplement institutional data used in coastal city early-warning systems. This validation effort aims to ensure that the data collected by citizen scientists meets the necessary standards for use in critical decision-making processes related to climate resilience.

1.2. Definitions

In environmental applications, sensor validation is the process of confirming the overall accuracy and reliability of measurements, as well as their ability to support specific applications. This validation process includes calibration and the verification (or monitoring) of this calibration.

The terms calibration, verification, and validation are defined by the International Organization for Standardization (ISO/TS 19101-2:2008, ISO/TS 19159-1:2014) and the Committee on Earth Observation Satellites (CEOS) for satellite remote sensing of environmental variables.

Calibration refers to the process of quantitatively defining the system's responses to known, controlled signal inputs. Proper calibration of sensors, and providing detailed instructions on sensor operation and maintenance is essential to provide high-quality data collections and ensure streamlined validation.

In contrast, validation requires independent means to assess the quality of the data products derived from the sensors and system output. While calibration may involve laboratory tests, validation requires measurements in the field, with a focus on the application and purpose of data usage (fitness to be used). Overall, validation assesses the adequacy of sensors to fulfil specific objectives.

The verification process checks current data to ensure their accuracy and consistency. Verification often involves real-time monitoring to promptly identify any behaviours that can be attributed to the malfunctioning of some sensors.

1.3. Importance of Data Validation in Citizen Science

The validation of data from low-cost sensors is pivotal for a number of reasons. Firstly, it ensures that the data is accurate and reliable, thereby making it useful for integration with institutional data. Proper data validation enables the creation of a database that can enhance early-warning systems, providing timely and precise supportive information about climate hazards. This, in turn, helps in the development of effective mitigation and adaptation strategies.

Validation procedures further help to identify and correct any discrepancies in the data collected by low-cost sensors. By comparing data from these sensors with data from other sensor networks or high-quality reference instruments, the validation process can highlight any inconsistencies and guide necessary adjustments. This ensures that the data contribute positively to the overall understanding of local climate conditions and hazards.

Validation also plays a critical role in maintaining the credibility of citizen science activities. When the data collected by citizen scientists is validated and deemed reliable, it builds trust among stakeholders, including scientists, policymakers, and the general public. This trust is essential for the continued support and expansion of citizen science initiatives, which are increasingly recognised as valuable contributors to environmental monitoring and climate resilience efforts.

2. OVERVIEW OF CITIZEN SCIENCE SENSORS IN CCLLS

2.1. CCLL Sensor Deployment and Management

Citizen Science plays a pivotal role in the SCORE project, where local communities actively participate in environmental monitoring to enhance climate resilience. The CCLLs are integral to this initiative, identifying major hazards that require vigilant monitoring to mitigate their impact on coastal cities. This collaborative approach ensures that the monitoring activities are aligned with the specific needs and vulnerabilities of each community. SCORE aims to mitigate vulnerability to coastal hazards by deploying a dense network of low-cost sensors and empowering citizens to participate in hazard co-monitoring. These affordable sensors have to provide consistent data for trend analysis and support early warning systems, as well as that they might monitor the effectiveness of various ecosystem-based adaptations (EBAs) implemented within SCORE.

Initially, the CCLLs engaged in discussions and assessments to pinpoint the critical environmental hazards, such as coastal and riverine flooding, air quality deterioration, and shoreline erosion. These identified hazards served as the foundation for the project WP4 team to develop a curated catalogue of low-cost sensors (https://sensors.score-eu-project.eu/). This catalogue was designed to offer a range of sensors capable of monitoring various environmental parameters relevant to the identified hazards. The CCLLs then selected sensors from this catalogue, tailoring their choices to their respective local context and monitoring needs.

Table 1: Environmental hazards and monitored parameters

Hazard	Monitored Parameters
Coastal erosion	Shoreline and topography; water level; surface waves
Coastal flooding	Water level and surface waves; precipitation; atmospheric pressure; air temperature
Drought	Precipitation; atmospheric pressure; air temperature
Heat waves, heath domes	Air quality and UV levels; precipitation; atmospheric pressure; air temperature
Lands and river flooding	Precipitation; atmospheric pressure; air temperature, water level; surface waves
Landslide	Precipitation; atmospheric pressure; air temperature
Sea level rise	Shoreline and topography; water level; surface waves
Storm surge	Water level; surface waves; precipitation; atmospheric pressure; air temperature

A full summary of the low-cost sensors utilised in the project can be found in the Appendix section. These low-cost sensors are crucial for enabling local communities in the CCLLs to actively engage through citizen science initiatives. The data collected from these citizen science projects is directly uploaded to the SCORE ICT platform, where it complements institutional data and aids in validating models for the SCORE Early-Warning Support system. This citizen science approach is particularly valuable because coastal monitoring is typically complex (due to the multitude of hazards simultaneously occurring) and therefore expensive, often relying solely on high-cost standard instruments that offer limited spatial and temporal resolution.

Given the reliance on these citizen-collected datasets, ensuring their accuracy and reliability is paramount. The most common planned validation method is cross-validation with reference instruments. Other methods include statistical methods and comparison with reliable data sources, Table 2 summarises these methods by CCLL.

Table 2: CCLL validation methods

CCLL	Monitored Parameter	Method
Barcelona/Vilanova	Intensified atmospheric conditions	Weather stations
Sligo	Sea level rise	Sea water levels
Massa	Coastal erosion	Shoreline, topography
Gdansk	Statistical analysis	Statistical methods
Benidorm	Statistical analysis	Statistical methods
Barcelona/Vilanova	Comparison with reliable data	Rain sensors, water level
Benidorm	Comparison with reliable data	Photography

3. VALIDATION ALGORITHMS AND METHODOLOGIES

3.1. Introduction to Validation Algorithms

Validation algorithms play a crucial role in assessing model performance by comparing predictions to real-world data or a separate validation dataset. These methods help determine whether a model can generalise effectively to unseen data. In the context of low-cost sensors, data validation becomes even more critical due to several reasons:

- Accuracy and reliability: Validation confirms that a model accurately represents the real-world system or the process(es) it aims to simulate. This is particularly vital in applications where incorrect predictions could have serious consequences, for example healthcare applications to monitor personal health status or autonomous driving.
- Prevention of overfitting: Overfitting occurs when a model learns not only the underlying patterns in the training data, but also to recognise noise and outliers. Unlike traditional data training, which involves using all available data, low-cost sensors may not have extensive training data. Validation ensures that the model generalises well to new data, reducing the risk of overfitting.
- Enhancing model quality: Validation identifies and rectifies errors, ensuring consistent and reliable model behaviour. By catching issues early in the development process, it contributes to scalability and flexibility while minimising costs.
- Detecting errors and anomalies: Low-cost sensors are susceptible to manufacturing inconsistencies and environmental factors. Data validation helps identify and correct these errors, ensuring accurate analysis.
- Ensuring consistency across devices: Variations in sensor performance make consistency challenging. Validation standardises data, making it comparable across different sensors and ensuring uniformity across the collected data.
- Improving model performance: Validating large datasets from low-cost sensors enhances model inputs, leading to better accuracy in predictions for applications like environmental monitoring and citizen science.
- Maintaining data quality: Over time, low-cost sensors may degrade, causing measurement drift. Regular data validation detects these drifts early, prompting recalibration or sensor replacement to maintain high data quality.
- Supporting decision-making: Reliable, validated data from low-cost sensors is essential for informed decision-making in fields such as smart cities, environmental management and public health.

To ensure the accuracy and reliability of data from low-cost sensors, the SCORE project leverages a multi-faceted approach to guarantee the accuracy and reliability of data collected from low-cost sensors.

Several validation criteria must therefore be applied. Initially, calibration and comparison with high-precision reference instruments are essential for the majority of sensors being used by the CCLLs. Regular calibration against these references can correct biases in sensor readings and verify their accuracy. Statistical techniques such as regression analysis or error detection algorithms can then be employed to compare sensor outputs with reference values, thereby quantifying and addressing discrepancies caused by environmental variations or sensor malfunctions.

Maintaining data integrity and reliability involves continuous monitoring and validation processes. Using outlier detection algorithms are recommended to handle outliers indicating potential sensor issues or external interferences and to ensure reproducibility through testing to confirm that repeated measurements under the same conditions yield consistent results. Establishing validation rules and thresholds for acceptable data ranges automatically flags and excludes invalid data, complemented by automated quality control processes that continuously validate incoming data.

3.2. Criteria for Validating Low-Cost Sensor Data

Calibration and verification of calibration are essential for correcting biases in sensor readings. Common methods include data-driven techniques to adjust outputs based on reference measurements. This often involves adjustments for environmental factors like temperature and humidity, which can significantly impact sensor performance. Low-cost sensors typically undergo validation by comparing their data with that from high-accuracy instruments, aiding in identifying and correcting discrepancies.

Ensuring data consistency over time requires regular stability checks. This process includes monitoring for sensor drift and performing recalibrations as necessary. Employing statistical measures, such as the *Coefficient of Determination* (R^2), *Root Mean Square Error* (RMSE), and *Mean Absolute Error* (MAE), evaluates sensor accuracy. These metrics quantify alignment between sensor readings and expected values, highlighting any deviations. See Appendix 2 for definitions of all the statistical measures referenced in this report.

Sensors must be tested and adjusted for their specific operational environments. For coastal areas, considerations include factors like saltwater corrosion and varying atmospheric conditions. Utilising multiple sensors for redundancy helps in cross-verifying data, enhancing reliability. This redundancy enables the identification and correction of anomalies or errors, ensuring the robustness of the collected data.

3.3. Approaches to Algorithm-Based Data Validation

Machine learning models can be trained on labelled data from reference instruments to predict and correct low-cost sensor outputs. These models are capable of performing anomaly detection, and can identify readings that deviate from expected patterns. Statistical techniques, such as regression analysis, assess relationships between sensor data and reference measurements to address biases. Time-series analysis methods like an autoregressive integrated moving average can also detect trends, seasonality, and anomalies.

Calibration algorithms play a crucial role in maintaining sensor accuracy. Dynamic calibration continuously updates parameters based on real-time comparisons with reference instruments. Multi-point calibration involves using multiple reference points to achieve greater accuracy, particularly useful for diverse sensor types like those in the SCORE project. Data fusion techniques combine inputs from multiple sensors to enhance overall data quality and reliability.

Various validation methods ensure data integrity and reliability. Statistical validation involves outlier detection and regression analysis to identify trends and patterns. Contextual validation checks for consistency with expected environmental conditions and known data patterns. Record-level validation includes completeness checks and identifying gaps in data collection. Automated validation employs scripts and rules within data processing pipelines for real-time data integrity checks.

The project team is currently working on implementing an algorithm to collect and compare data from relevant platforms with locally collected, continuously updated reference values. This algorithm is being developed as part of the ongoing improvements to the SIP and is expected to be completed before the project's conclusion. While specific algorithms will vary depending on the sensor type, the underlying principle of consistent validation remains the same.

3.4. Data stream quality automated evaluation and pre-processing by the DT-EWS

The EWSS collects and processes the data from the data streams acquired by the sensors and operates some checks on their integrity and consistency, also taking actions to prepare them to be used for the current scenarios evaluation. Each data stream is read and passed through a chain of data-preparation processes:

- *Data normalization and check*: Time series are made compliant with a common system format, i.e., with the system time parameters and measurement units. Missing or negative values are detected, the latter replaced by empty entries, and if the percentage of missing data in the scenario time interval exceeds a specific threshold, data are discarded. In all these cases, the system raises a warning through the graphical user interface (GUI) that can be read by the user, with the details about the problem of the specific sensor.

- *Outliers detection*: A procedure aimed at detecting possible wrong or suspect values is applied. These kinds of techniques represent a common step in the pre-processing of data for analysis or model-building purposes. More specifically, the EWSS implements algorithms based on the Median Absolute Deviation (MAD) and Isolation Forest to identify outliers in the time series (see the documentation in D8.5 for details). At the end of the procedure, all the points identified as possible outliers are removed. Like in the previous step, in case of detection, the information is reported to the user as a warning related to the specific sensor through the GUI.

- *Data imputation*: Once the time series are cleaned up, if they are considered valid, i.e., the number of removed/missing elements does not exceed the fixed threshold in the scenario time interval, empty entries are imputed in order to get a complete time series. This task is achieved in two possible ways: By adopting either interpolation, or median imputation, both selectable in the configuration stage. In the first case, the time series values are linearly interpolated, keeping into account the actual frequency. In the second case, the median of the values in the series interval is used.

At the end of these steps, collected data from different data streams belonging to a specific time series and location are processed together, in order to obtain a single time series associated to a relevant scenario. At this stage, for the specific element, it is necessary to have at least one valid time series that has been previously collected. If this is not the case, the system notifies through the GUI that the element itself cannot be used for the creation of the scenario.

The result of these steps is therefore a collection of valid time series acquired from the sensors that can be fed into the process of creation of the current scenario. Indeed, these values are then channelled into the feature aggregation procedure, which represent the last preprocessing step needed to define the current scenario evaluated by the EWSS.

4. LOCAL LEVEL VALIDATION IN CCLLS

4.1. Specific Procedures and Protocols for Routine Measurements

This section presents the specific procedures and protocols for routine measurements, focusing on the calibration of sensors used in the CCLLs. The tables below detail the calibration methods employed for some of the most popular and established sensors currently in use by the 'frontrunner' CCLLs. These frontrunner CCLLs were selected to pilot these processes, establishing best practices for sensor deployment and calibration. Their role as pilots was intended to provide a framework for the 'follower' CCLLs, which subsequently can now utilise these examples when deploying their own sensors in collaboration with citizen scientists.

While this section concentrates on the sensors that have been successfully implemented by the frontrunners, it is important to note that many other sensors such as the DIY Wave Gauge (https://sensors.score-eu-project.eu/sensor/diy-wave-gauge/) and the Smart Pebbles (https://sensors.score-eu-project.eu/sensor/smart-pebbles/) are still in the pilot phase. As these pilots continue, the corresponding calibration protocols and documentation are being developed based on the collective experiences of all CCLLs. These experiences will be compiled into an overall catalogue of instructional manuals which will be included on the sensor catalogue website. This evolving catalogue will serve as a resource for future sensor deployments, offering guidance and support to other CCLLs and similar initiatives beyond the scope of this project.

Table 3: MINKE water quality sensor calibration itinerary

Sensor	Typical criteria for data validation	Guidelines to follow
MINKE Water quality sensors	*Calibration Process*	Order and Accuracy: Depending on the type of probe that is calibrated, calibration must follow a specific order: it can either be a three-point calibration—midpoint first, then low point, and finally high point, a 2-point calibration—dry point and 0 mg/L (for dissolved oxygen e.g.), or single-point calibration. Inaccuracies in this order or values can lead to erroneous data. Cleaning: The sensor probes need to be cleaned with distilled water between each calibration step to prevent contamination. Stable Readings: Ensuring the sensor readings are stable before recording calibration values is crucial. Unstable readings can lead to incorrect calibration and thus poor data quality. Storage Solutions: The pH and ORP probes need to be kept in a storage solution when not in use to prevent them from drying out, which can affect their accuracy.
	Installation and Setup	Environment: Sensors should be installed in environments avoiding moisture accumulation, temperature and humidity transients, and direct airflow, all of which can impact sensor accuracy. Protection: Outdoor installations require protective enclosures to shield the sensors from moisture and other environmental factors. Filtering foam or other protective measures can be used to protect the sensor pad. Circuit Configuration: Properly configuring circuits, such as switching from UART to I2C mode for communication, is necessary for correct sensor operation.

	Maintenance	**Regular Cleaning**: Sensors, especially those deployed in the field, need frequent cleaning to prevent biofouling and other contaminants from affecting measurements. The maintenance can vary from weekly to less frequent depending on the environment. **Electrolyte Solutions**: Some probes, like those for dissolved oxygen, require periodic replacement of internal electrolyte solutions and cleaning of residues.

Table 4: HOBO water level sensor calibration itinerary

Sensor	Typical criteria for data validation	Guidelines to follow
HOBO water level sensor	*Calibration Process*	Calibration Process • Order and Accuracy: Each pressure sensor in the sensor end is individually calibrated at multiple pressures and temperatures over the calibrated range. Ensure the logger is within the specified operation range (0 to 207 kPa or 0 to 850 kPa depending on the model) for accurate measurements. Avoid exceeding the burst pressure of the sensor. • Stable Readings: Ensure the logger has reached full temperature equilibrium (approximately 20 minutes) before starting measurements to ensure accuracy. • Reference Water Level: Set a precise reference water level at the beginning of each deployment using the HOBOconnect app. This step is crucial for accurate water level data. • Water Density: Configure the correct water density in the app to ensure accurate water level readings. Cleaning Biofouling: Regularly inspect the sensor end for biological growth, which can affect pressure sensor accuracy. Clean any biofouling to maintain sensor performance. Storage Solutions Battery Maintenance: Replace the AA batteries as needed, especially before long deployments. Properly store the logger in a dry environment to prevent any damage to the electronic components.
HOBO water level sensor	*Installation and setup*	Environment • Avoid Moisture Accumulation: The logger's top-end unit is weatherproof but not waterproof. Avoid continuous, highly saturated environments to prevent condensation inside the logger. • Temperature Equilibrium: Ensure the logger reaches temperature equilibrium before starting measurements. Avoid sudden temperature changes that can affect accuracy. Protection Protective Enclosures: Use stilling wells for deployments in lakes, rivers, or streams to protect the sensor from currents, wave action, and debris. Ensure proper ventilation and shading for well deployments to minimize condensation. Circuit Configuration Logger Configuration: Properly configure the logger using the HOBOconnect app. Ensure Bluetooth connectivity for data download and logger setup.

	Maintenance	Regular Cleaning • Sensor Inspection: Periodically check the sensor end for any contaminants or biofouling. Clean the pressure sensor's face to maintain accuracy. • Battery Replacement: Replace the batteries regularly, especially for deployments longer than a year or in extreme temperature conditions. Use fresh dielectric grease (comes with the sensor) on contacts to maintain a good electrical connection. Compensating for Drift Regular Reference Checks: Regularly update the reference water level and water density during deployment to compensate for any drift. This ensures the long-term accuracy of the logger data.

Table 5: BRESSER WIFI weather station calibration itinerary

Sensor	Typical criteria for data validation	Guidelines to follow
BRESSER WIFI ClearView 7in1	*Calibration Process*	No calibration process is recommended by the manufacturer, although calibration can be undertaken via the device's captive portal if required. However, the manufacturer advises against this.
	Installation and setup	The user must place the sensor above a non-heat radiating surface such as gravel, bare earth, greenery…(etc) at least 5 metres away from buildings and high vegetation. The station is to be mounted around 5 metres above the ground, pointing north. At this point, the user is directed to connect the sensor to the base and connect the system to the Wi-Fi and Weather Underground account.
	Maintenance	No frequent maintenance is required for this unit. The user is provided with an email if their device is disconnected. In this case, they are instructed to check the batteries in the sensor or perform a connection refresh.

5. VALIDATION AT THE SIP LEVEL

The SCORE ICT Platform (SIP) is a centralised data repository and knowledge-sharing hub designed to enhance collaboration among project partners and stakeholders. By providing access to datasets, maps, and geostories, the aim of the SIP is to support the development and implementation of climate resilience strategies in coastal cities. Serving as a backbone for effective communication and data exchange, it enables greater coordination among stakeholders, which is vital to achieving the project's goals. The platform's capability to enable real-time interactions and its integration with various technologies make it indispensable in contemporary environmental monitoring and conservation. Errors can arise from various sources, including gross errors due to device malfunctions, systematic biases, and random errors inherent to measurement processes. Additionally, representativeness errors, stemming from spatial and temporal discrepancies, must be addressed to ensure data accuracy. By incorporating error modelling and validation processes, the SIP can maintain the integrity of integrated data, enhancing its reliability for environmental monitoring and decision-making.

5.1. Approaches on Data Validation Algorithms on the SIP

Consistency checks

According to WMO, measuring an environmental variable implies a set of operations for determining the value of such a variable in a given measurement unit. The difference between the value of the measurement and the true value (unknown) of the measured quantity is the error.

Measurements of the field of a physical variable (e.g. temperature, precipitation) are obtained via remote sensing or via a network of pointwise sensors that are combined through spatial interpolation methods to reconstruct a 2-D field of a given variable. However, the quality of the reconstructed field is linked to the error of each sensor of the network and of course, to the representativeness error. Both institutional networks and citizen science networks should apply some automatic check to assess the validity of a datum collected by a sensor

A first level of check is performed on elementary raw data (or gross data), and consists in the application of a basic procedure for verifying the absence of low level anomalies, such as malfunctions, instabilities or interferences. Some sensors are equipped with Built-In Test Equipment (BITE) allowing remote control of basic sensor parameters. Dealing with the citizen science network, using sensors with a basic BITE is recommended, in order to properly flag meaningless measurements or anomalous functioning. However, whether BITE is present or not, *inner consistency* tests should be performed, which at simplest level, consist in a check detection of unrealistic measurements (i.e., physically unacceptable or outside of range of measurements of a sensor). Other checks are based on detection of inconsistencies between different measurements (a trivial example could be detection of rain with measured temperature well below the zero leading to non-physical measurements).

Other checks are the following:

- *Temporal consistency checks*: These checks are based on checking the maximum and minimum degree of variability of the data over time and are intended to identify any anomalies between temporally contiguous data or with respect to the values that have historically occurred in a given site. With regard to the minimum variability allowed, the temporal consistency check procedures are aimed at ascertaining the presence of persistence in the series of measured values, i.e. the persistence over time of a value that is the same in time or, conversely, characterised to excessive fluctuations. Such a check depends on the properties of a measured physical quantity and could differ, for example for precipitation or for temperature. Some threshold criteria can be adopted, but the design of thresholds should rely on an analysis of available time series, but a fine tuning is not strictly necessary.

- *Cross-checks with other quantities collected close to the sensors*: These checks are based on checking the measurements of the variable under examination with other correlated quantities measured in the same site (e.g. comparison of temperature with solar radiation) or at a certain distance for which the representativeness error can be negligible.
- *Spatial consistency checks:* These checks assume the existence of a sort of spatial correlation between the contemporary measurements of a quantity taken at (more or less) neighbouring stations. This category of check takes advantage of the availability of more sensors in the network. The implementation of such consistency check depends on the nature of the physical variable observed. An a priori knowledge of spatial property of a field is necessary, or otherwise specific statistical analysis of available network measurements are requested.
- *Climatological checks*: These checks are based on the comparison of the quantity under examination with some parameters obtained from long term historical series.

All these checks can be easily implemented to achieve automatic procedures to detect suspect data. In general, the detection of suspect data should trigger an inspection of the sensor and if suspects are confirmed, maintenance operations carried out by skilled technicians should start. In the context of a citizen science network, the adoption of detection of suspect data through automatic procedures is essential. Citizen scientists should be promptly made aware of issues with the sensor they manage, and in principle try to fix them. Controls are currently being established to flag issues and automatically notify citizens that their sensor needs maintenance.

Dual co-location

Whereas consistency checks target the detection of anomalies, the error of sensors still needs to be assessed. The comparison with a co-located reference sensor (it can be called Dual Collocation, DC) is the more common approach used to compare the relative uncertainties of measurement systems. The method assumes that two sensors measure the same things. In this case, one is assumed as the "true" and errors of a pair of sensor measurements are independent, the standard deviation of the difference $d_{i,j} = X_i - T_j$ coincides with the instrumental errors. In case T is replaced by another measurement Y characterised by a specific error, it is not possible to disentangle the contribution to such standard deviation due to the natural variability of the observed phenomenon from the error of sensors. This is the most common approach to verify calibration. However, needed is a co-located reference sensor for example, a sensor belonging to an institutional sensor network, for which a calibration protocol exists and is applied. However, this is not always possible. Nevertheless, in a context of consistency checks, monitoring of intercalibration of sensors can be helpful.

Triple co-location

In contrast, the triple co-location is a mathematical method that does not require a reference sensor to evaluate product error statistics. The method was introduced in 1998 for a satellite application [1], but has gained popularity recently, being increasingly applied to a wide variety of sensors and environmental parameters. However, one problem seems to be emerging and that is the existence of three sensors in a single point. Here, the three measurements could be obtained by the monitoring sensor, while the other two measurements can be obtained by interpolation of fields obtained from interpolation of institutional networks, or remote sensing. The important thing is to monitor with time, and compare the results of triple co-locations to detect anomalous values.

The method is based on an error model that relates estimates to true values. The most common error model assumed to apply the method [2] is described by

$R_i=\alpha_i+\beta_i T+\varepsilon_i$, i = 1, 2, 3 (1)

being R the estimate/measurement of a variable, T is the "true" (but unknown) value, while the error terms are α (additive error), β (multiplicative error), and ε (residual error). An alternative method is the multiplicative error model

$Ri=a_i T^{\beta i} e^{\varepsilon i}$, i = 1, 2, 3 (2)

That can be linearized as

$r_i = \alpha_i + \beta_i\, t + \varepsilon^i$, $i=1, 2, 3$ (3)

by taking the natural logarithm of both members and using $r_i = \ln(R_i)$ $\alpha_i = \ln(a_i)$, $r_i = \ln(R_i)$, and $t = \ln(T)$.

There are different procedures to solve the system of equations by optimising RMSE (root-mean-square error). Some implementations are available in public software repositories (see [3] https://github.com/HamedAlemo/MTC, or [4] https://github.com/kaighin/ETC) and also to derive important metrics to evaluate the uncertainties.

The triple co-location method requires certain assumptions beyond those related to the error models discussed in the previous equations. These include the independence of the "true" values from the "noise" in the sensor measurements and the independence of the components of the error terms within those equations. Additionally, the method assumes stationarity, meaning that the "true" values and noise maintain a constant mean and standard deviation over time, as opposed to being mobile. Another key assumption is representativeness, where sensors must observe a phenomenon that shares the same properties across measurements. Although some of these assumptions may only be loosely verified, it is generally accepted that the method can still be practically applied, yielding results that are less reliable but nonetheless valuable.

Outputs are the estimated error standard deviation of the three measurements, and the correlation coefficient between the "true" T and each measurement time series X_i.

In the case of citizen science networks, the method can be applied in a test bed in which three sensors are co-located, but also to implement the consistency check defined above, or assuming stationarity and representativeness at some distances (they depend on the correlation of the observed variables), or considering, for example, a triplet made of the value of the sensors, the value obtained through interpolation from institutional sensor networks, and values obtained by some models. In any case the application of the methods should be tuned to the environmental variable to be estimated.

6. INTEGRATION WITH INSTITUTIONAL DATA ON THE SCORE SIP/DT

6.1. Complementing Institutional Data with Citizen Science Data

Citizen science sensors, in general, can have a relevant impact in the EWSS operations, whose results inform users about the risk of flood in the CCLLs urban area. The EWSS employs data from institutional sensors, but it can also be fed with data from citizen sensors, that can be more finely distributed over the study area, providing a more detailed picture of the ongoing scenario the EWSS must evaluate. Hence, provided that sensors are correctly calibrated and their data quality confirmed, citizen science can decisively impact on the current scenario assessment, also helping in the re-analysis of past events.

Once the citizen science sensors will be installed and calibrated, they can be added as data sources to the sensor things application programming interface (STAPI) platform. First of all, a careful cataloguing on the STAPI of the related data streams is needed, making the measured parameters easily findable and usable by the system. The acquired data must be catalogued in such a way that the system can differentiate them, e.g., distinguishing rain rate (expressed in mm/h) and cumulated precipitation (in mm, over a specified time span) and correctly interpret and use their values. In case of multiple data produced by the same sensor, it is important to distinguish among different streams and ingest to the EWSS of the DT only the ones that are really relevant for its monitoring operations (e.g., rain rate, sea level, river discharge, sewage network pipes level).

At the moment, the EWSS is almost exclusively exploiting data from official sensors. It is worth noticing that it repeatedly performs an analysis of the received data, to check their completeness and consistency, and to point out the presence of eventual outliers, as already shortly described in Section 3. Of course, citizen science sensors must undergo the same following pre-processing stage. In case it is not possible to obtain the data series for the simulation because the original data streams lack of integrity or consistency, the system behaves differently with official and citizen science sensors. For the former, it raises warnings and informs the users through the graphical interface, so that users can take actions to check the consistency of the received data and of the sensors functioning. If the data are not consistent and the sensor is not properly working, the user can decide to “deactivate” it, in the sense that the data stream it generates are not included in the EWSS runs, to avoid affecting the EWSS projections with wrong information. In the case of citizen science sensors, on the other hand, the system will autonomously take the action of sensor “deactivation”. This is due to the fact that official sensors are considered more reliable, since they have been installed by expert operators, correctly calibrated, and managed by institutional entities that continuously check their functioning. They can produce inconsistent data streams in case of accidental malfunction and, in general, every warning raised by the system needs a careful evaluation and a decision by users/operators with relevant expertise. Conversely, citizen science sensors, being cheaper, managed by citizens, and potentially more prone to failure or malfunction, can easily have a negative impact on the run results of the EWSS. Therefore, to guarantee the robustness of the system in making projections of flood risk, citizen science sensors exhibiting a suspect behaviour will be automatically excluded, until they are not fixed and manually reintroduced by a user in the EWSS. This allows a more reliable impact of the data from citizen science sensors on the EWSS results.

7. CONCLUSION

The SCORE project's approach to integrating citizen science sensors and data with institutional monitoring systems offers significant potential to enhance environmental monitoring, particularly in the context of urban coastal areas. This report highlights insights and practical strategies for implementing this integration effectively. The deployment of low-cost sensors across various CCLLs has proven valuable in capturing a more detailed and localised understanding of environmental conditions. These sensors enable broader data coverage and can complement institutional networks, especially in areas that are otherwise hard to monitor. Despite their potential, low-cost and citizen science sensors pose challenges in data reliability. The report underscores the importance of rigorous validation processes, including calibration, cross-validation with local reference instruments, and consistency checks to ensure data accuracy. The SIP has emerged as a crucial tool for centralising data, facilitating real-time interactions, and ensuring effective communication among stakeholders. The platform's integration with various data validation algorithms will further support the accuracy and reliability of the data collected.

The implementation of consistency checks (temporal, spatial, and climatological) and co-location methods (dual and triple co-location) has been identified as essential for maintaining the integrity of data from both institutional and citizen science sources. These methods will help to filter out anomalies and improve overall data quality. Furthermore, the EWSS will benefit from the inclusion of citizen science data, provided that these data streams are validated and deemed reliable. The ability of the EWSS to process and utilise this data has the potential to improve the accuracy of flood risk assessments and other critical environmental predictions.

Citizen science data, when effectively integrated with institutional data, can significantly enhance the resolution and scope of environmental monitoring efforts. However, this integration demands data management practices to avoid potential issues of unreliable data. Therefore, it is essential to implement regular calibration of all citizen science sensors against high-precision instruments. This should be coupled with a validation framework which includes both automatic and manual data checks to ensure the highest possible data quality. By continuing the close cooperation with the development team of WP5, the SIP will continue to evolve in this respect as a central hub for data management, validation, and dissemination. Its role in ensuring real-time data consistency and facilitating stakeholder collaboration. Further SIP development will be particularly carried out in the context of integrating new data sources. The EWSS should also be further developed to autonomously manage data integrity, particularly with respect to citizen science sensors. Enhancing the system's ability to filter and flag inconsistent data will ensure that only reliable information informs critical environmental assessments.

To maximise the potential of citizen science data, it is crucial to invest in capacity-building initiatives among stakeholders as citizen scientists. This includes training in data collection and sensor maintenance, thereby increasing the reliability of the data they contribute. With this in mind, a clear framework for integrating citizen science data with institutional datasets is being established through the SCORE project, to be finalised by the end of the project for others to follow.

REFERENCES

[1] Cocco, C., Pilla, F., Oduori, G., Gómez-Barrón, J. P., Crowley, S., Calleo, Y., Banerjee, A., Anton, I., Gharbia, S., Nalakurthi, S.-R., et al. (2024). D4.2-Report on low-cost sensors viable for citizen science activities. Zenodo. https://doi.org/10.5281/zenodo.11242082

[2] Crowley, S., Cocco, C., & Pilla, F. (2024). D4.3-Citizen science DIY framework [Data set]. Zenodo. https://doi.org/10.5281/zenodo.11242108

[3] Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modelling and calibration using triple collocation. Journal of Geophysical Research: Oceans, 103(C4), 7755-7766. https://doi.org/10.1029/97JC03180

[4] Zwieback, S., Scipal, K., Dorigo, W., & Wagner, W. (2012). Structural and statistical properties of the collocation technique for error characterization. Nonlinear Processes in Geophysics, 19(1), 69-80. https://doi.org/10.5194/npg-19-69-2012

[5] Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., & Stoffelen, A. (2015). Characterization of precipitation product errors across the United States using multiplicative triple collocation. Hydrology and Earth System Sciences, 19(1). https://doi.org/10.5194/hess-19-1-2015

[6] McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. Geophysical Research Letters, 41(16), 6229-6236. https://doi.org/10.1002/2014GL061322

APPENDIX 1

Low-cost Sensors for Citizen Science

The SCORE project uses a catalogue of sensors selected to monitor a variety of parameters and environmental hazards. Below is a summary of these sensors and details on their respective use cases. The catalogue is available online at https://sensors.score-eu-project.eu/

Air Quality and Meteorological Parameters

This category encompasses sensors that measure atmospheric conditions, including temperature, humidity, pressure, wind speed, wind direction, and solar radiation. These parameters are crucial for understanding weather patterns, climate change, and air quality.

- BRESSER WIFI ClearView Weather Station with 7-in-1 Sensor: Measures a comprehensive set of meteorological parameters including temperature, humidity, atmospheric pressure, precipitation, wind speed, wind direction, and UV levels.
- WSG Sonic Anemometer: Specifically measures wind speed and direction.
- BP260 Pressure Sensor: Measures atmospheric pressure.

Water Level and Flow

These sensors are designed to monitor the level and movement of water bodies, such as rivers, lakes, oceans, and underground aquifers. They are essential for flood prediction, water resource management, and coastal erosion studies.

- DIY wave gauge: Indirectly measures wave characteristics and water levels using pressure sensors.
- Onset HOBO water level loggers: Directly measures water levels.
- MaxBotix Ultrasonic sensors: Measures water levels using ultrasonic technology.
- Lidar-lite V3HP: Measures water levels using laser rangefinding.
- LoRaWAN LIDAR ToF Distance sensor LLDS12: Measures distance, which can be used to determine water levels in specific applications.
- Radar-Operated Tide Level Sensor: Specifically measures tide levels.
- Float-Operated Tide Level Sensor: Measures tide levels using a float and encoder system.
- PLS Level Sensor: Measures water levels based on pressure.
- RLS Level Sensor: Measures water levels using radar.
- ULS Level Sensor: Measures water levels using ultrasonic technology.
- Art Sewer: Specifically measures wastewater levels using radar.

Precipitation

These sensors measure the amount of rainfall, which is vital for hydrological modelling, agriculture, and disaster management.

- SmartLNB: Measures rainfall.
- Optical rain sensor: Measures rainfall based on optical reflection.
- R102 Pluviometre: Measures rainfall.

Soil and Water Quality

These sensors monitor the chemical and physical properties of soil and water, crucial for agriculture, environmental monitoring, and water resource management.

- Smart Citizen Water/Soil Station: Measures various water quality parameters like pH, dissolved oxygen, conductivity, ORP, temperature, and soil moisture.
- iMoisture - Soil Moisture Sensor: Specifically measures soil moisture.

Coastal and Shoreline Monitoring

These tools are used to observe and measure changes in coastal environments, including erosion, sedimentation, and sea level rise.

- Kite Aerial Photography (KAP): Captures images for analysis of coastal erosion and shoreline changes.
- Smart Pebbles: Used for studying beach morphology and erosion.
- Fixed Camera and PTZ Camera: Capture visual data for coastal monitoring.

Remote Sensing and Data Transmission

These technologies enable the collection and transfer of data from remote locations, facilitating real-time monitoring and analysis.

- Remote Telemetry Units (RTUs) + sensors: A platform for connecting various sensors and transmitting data.
- LoRaWAN LIDAR ToF Distance sensor LLDS12: Uses LoRaWAN for data transmission.
- SenseCAP S2120 8-in-1 LoRaWAN Weather Sensor: Uses LoRaWAN for data transmission.

Note: Some sensors, like the BRESSER weather station, provide data on multiple parameters. The categorization here is based on the primary parameter or the most significant use case.

APPENDIX 2

Definitions of statistical measures:

- **Anomaly Detection:** Anomaly detection identifies data points or patterns that significantly deviate from the norm or expected behaviour, often indicating errors or significant changes in the data.
- **Autoregressive Integrated Moving Average (ARIMA):** ARIMA is a time-series analysis method that models and forecasts data based on its own past values and past forecast errors, used to detect trends, seasonality, and anomalies.
- **Calibration Algorithms:** Calibration algorithms are techniques used to adjust sensor outputs to improve accuracy, typically involving comparison with reference instruments.
- **Coefficient of Determination (R^2):** R^2 is a statistical measure that indicates how well the observed outcomes are replicated by the model, based on the proportion of the total variation in the dependent variable that is explained by the independent variables. A value of 1 indicates perfect correlation, while 0 indicates no correlation.
- **Contextual Validation:** Contextual validation checks the consistency of data with expected environmental conditions and known data patterns to ensure it aligns with theoretical or historical expectations.
- **Data Fusion:** Data fusion combines data from multiple sensors or sources to enhance the overall quality and reliability of the information, providing a more comprehensive and accurate picture of the observed phenomena.
- **Dynamic Calibration:** Dynamic calibration is an approach where sensor parameters are continuously updated based on real-time comparisons with reference instruments, helping to maintain sensor accuracy.
- **Mean Absolute Error (MAE):** MAE measures the average magnitude of errors in a set of predictions, without considering their direction. It is the average of the absolute differences between predicted and observed values, with lower values indicating a more accurate model.
- **Multi-Point Calibration:** Multi-point calibration involves using multiple reference points to adjust and improve the accuracy of sensor measurements, particularly useful for sensors with varying characteristics or operational conditions.
- **Record-Level Validation:** Record-level validation involves completeness checks and identifying gaps in data collection to ensure the dataset is accurate and comprehensive, focusing on the integrity of individual data records.
- **Regression Analysis:** Regression analysis is a statistical technique used to assess the relationships between a dependent variable and one or more independent variables, helping to understand how changes in independent variables affect the dependent variable.
- **Root Mean Square Error (RMSE):** RMSE is a measure of the differences between the predicted and observed values in a model. It calculates the square root of the average squared differences between predicted and observed values, providing an overall measure of prediction accuracy. Lower RMSE values indicate better model performance.
- **Statistical Validation:** Statistical validation involves using statistical methods such as outlier detection and regression analysis to ensure the integrity and reliability of data, identifying trends, patterns, and anomalies in the dataset.
- **Time-Series Analysis:** Time-series analysis involves methods like ARIMA to detect trends, seasonality, and anomalies in data over time.
- **Validation Methods:** Validation methods are processes used to ensure data integrity and reliability, including statistical, contextual, and record-level validation, as well as automated scripts and rules in data processing pipelines.