



Rationalization of automatic weather stations network over a coastal urban catchment: A multivariate approach

Mousumi Ghosh^a, Jitendra Singh^{b,c}, Sheeba Sekharan^a, Subimal Ghosh^{a,d,e}, P.E. Zope^e, Subhankar Karmakar^{a,c,d,*}

^a Interdisciplinary Program in Climate Studies, Indian Institute of Technology Bombay, Mumbai 400076, India

^b Environmental Science and Engineering Department, Indian Institute of Technology Bombay, Mumbai 400076, India

^c School of the Environment, Washington State University, Vancouver 98686, WA, USA

^d Centre for Urban Science and Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

^e Department of Civil engineering, Indian Institute of Technology Bombay, Mumbai 400076, India

ARTICLE INFO

Keywords:

AWS
Rationalization
PCA
TOPSIS
Flood mapping

ABSTRACT

The establishment and maintenance of an exhaustive hydrometeorological network are challenging tasks in densely populated coastal cities having erratic rainfall patterns. The current study proposes a robust statistical framework to rationalize an existing Automatic Weather Station (AWS) network which monitors multiple hydrometeorological observations to obtain maximum information at an optimal cost. This framework combines multivariate statistical approaches and multi attribute decision making techniques. We have demonstrated the proposed framework utilizing rainfall and relative humidity information from the existing AWS network across Mumbai city. Principal Component Analysis (PCA) is performed on daily rainfall and relative humidity datasets, followed by Technique of Order Preference by Similarity to Ideal Solution (TOPSIS) to rank the stations based on their capability to capture spatiotemporal variability. Out of the initial 60 AWSs established by the Municipal Corporation of Greater Mumbai (MCGM) with the primary objective of addressing Mumbai city's flood-susceptibility problem, 35 AWS data are used for this analysis on grounds of data completeness and reliability. Our analysis reveals that the spatiotemporal information of relevant hydrometeorological observations can be proficiently collected by a rationalized network of 22 AWS and that this may bring down network maintenance costs. The flood inundation and hazard maps for the Mithi catchment, one of the major flood hotspots of the city, are derived from the existing network and the rationalized network. The comparison of maps highlights the high accuracy of the rationalized network to reproduce the spatial characteristics of floods across the catchment. Our study presents a first-of-its-kind attempt to evaluate the performance of the rationalized network at flood inundation and hazard level derived from comprehensive 1-D and 2-D hydrodynamic approaches. The proposed generic framework can be utilized to reassess the precision and efficiency of various existing hydrometeorological monitoring networks at a regional-to-national scale and to achieve maximum spatiotemporal information from multiple hydrometeorological observations but at optimum maintenance cost.

1. Introduction

An incessant rise is being observed in frequency and intensity of hydroclimatic extremes (such as flood, rainstorm, drought, heatwave) over the last few decades locally, regionally as well as globally, attributable to increasing greenhouse emissions (Meehl and Washington 1993; IPCC 2013), anthropogenic activities (Kishatwal et al. 2010; Vittal et al. 2013; Shastri et al. 2015), and natural climate variability (Trenberth 2005; IPCC 2007). These hydroclimatic extremes which have

significantly increased over the last few decades are accompanied by huge livelihood and socioeconomic losses almost every year (Kron 2005). Real-time monitoring and analysis of hydrometeorological observations that quantify the intensity of such extremes will facilitate water resource planners and policymakers to strategize and adopt different mitigative measures. Studying such extremes though is found to be particularly challenging in developing countries due to the unavailability of sufficient long-term hydrometeorological data both at a finer spatial and temporal scale. High-density weather monitoring

* Corresponding author at: Interdisciplinary Program in Climate Studies, Indian Institute of Technology Bombay, Mumbai 400076, India.

E-mail address: skarmakar@iitb.ac.in (S. Karmakar).

networks that record data at sub-daily and sub-hourly scales are thus an indispensable requirement in urban areas to improve the quality and accuracy of statistical analysis for research purposes (Sherly et al. 2015). Hydrometeorological observations such as rainfall, temperature, relative humidity, wind speed, direction, and other relevant data measured along with the calculated ones such as dew point and wind chill are automatically transmitted or recorded by AWS network established over a region (Ahmad et al., 2017). Rainfall is the most dynamic hydrometeorological observation and shows greater spatiotemporal variability than other observations such as relative humidity and temperature, particularly over the tropical region. High variability in the amount of rainfall at a sub-hourly scale over a small area often results in flash floods, especially in urban catchments.

Design, establishment, and evaluation of adequate AWS networks are areas with huge research potential to obtain high-quality hydrometeorological estimates which are accurate and representative of the spatiotemporal variability in a region (Mishra and Coulibaly 2009). Ideally, the number of stations in an AWS network and their spatial distribution and design should be thoughtfully set up such that they provide maximum quantitative information about the data required by end-users at appropriate time intervals. The planning and evaluation of the hydrometeorological network are influenced by the user requirements and time step of the meteorological observations. The networks should be capable of providing relevant information for various purposes like the establishment of water distribution, quality and irrigation networks for water resources management, flood forecasting in flood-prone areas, the impact of climate change on water resources (Mishra and Coulibaly 2009; Varekar et al. 2016; Varekar et al. 2021). Therefore, the adequate design and evaluation of the network will vary as per the scope of the network to cater to the requirements of hydrologists, climatologists, hydrogeologists, and other stakeholders. Furthermore, it also varies according to the requirement of timestep or frequency of monitoring (sub-daily, daily, monthly) for different hydrometeorological observations. Such quantification will, however, be accompanied by higher costs of installation, operation, and maintenance (Agarwal et al. 2018), and hindrances through geographic and environmental constraints. Additionally, some stations may include redundant information without necessarily adding any new value in information content or decreasing uncertainty (Stosic et al., 2017). In developing countries, administrative authorities are usually apprehensive about the chances of equipment vandalism in station premises that disrupt their smooth operation and maintenance. Therefore, strategic rationalization of station networks is not a luxury but a necessity to extract optimal information from the monitored observations, and reassess and redesign the AWS networks to meet all basic requirements as cost-effectively as possible (Mishra and Coulibaly 2009). Rationalization is used to assess, reevaluate, design, and re-design new or existing hydrometeorological and water quality networks (Mishra and Coulibaly, 2009; Mustafa et al. 2014; Varekar et al. 2015; Varekar et al. 2016; Varekar et al. 2021) for the selection of significant/optimal stations, parameters, and frequency of measurement across the desired location.

Various methods have been discussed in the literature for the design and assessment of hydrometeorological networks. The extensively used approaches for network assessment and design can be summarized as (1) statistical, (2) information theory, (3) spatial interpolation, (4) optimization, and (5) hybrid approaches, which have been discussed comprehensively by Mishra and Coulibaly 2009. Statistical approaches are commonly applied to hydrologically homogenous regions to quantify uncertainty related to statistical estimates viz., reduction in variance (Matalas and Gilroy 1968; Putthividhya and Tanaka 2012), dimensionality reduction through principal component analysis (Morin et al. 1979; Dai et al., 2017) and cluster analysis (Kar et al. 2015; Tiwari et al. 2020), maximizing regional information using generalized least squares (Moss and Tasker 1991), all with the basic objective of minimizing deviation. Spatial interpolation techniques utilize point- and areal-based

values to determine data at any other point within the area on the basis of spatial dependence between neighboring observations (Camera et al. 2014; Manz et al. 2016). Kriging-based geostatistical approach, which can also be combined with other rationalization techniques stands out as the most extensively executed spatial interpolation statistical technique (Jewell and Gaussiat, 2015; Cecinati et al. 2018). The information theory approach utilizes the principle of maximum entropy, i.e., gauging stations should have minimum mutual information but maximum information content (Shannon 1948; Cover and Thomas 1991; Mogheir et al. 2006). The hybrid approaches combine two or more methods or utilize the output of one method as an input to another to improve the accuracy of measurements (Chen et al. 2008; Adhikary et al. 2015). Some studies apply expert knowledge and opinion utilizing their experience and take decisions based on user needs, regional topographical characteristics, and surveys to rationalize a network (Skok, 2006). Very few studies have however considered multiple hydrological and hydrometeorological observations while designing a network. Keum and Coulibaly 2017 proposed a decision support framework to design a multivariable hydrometeorological network by simultaneously utilizing discharge and streamflow data. Lately, remotely sensed data are also being increasingly implemented for hydrometeorological monitoring purposes owing to their superior ability in capturing better spatial and temporal resolution data (Li et al. 2016). Application of this approach is however constrained by uncertainty, low-resolution data (Dai et al., 2017), lack of continuous long-term data (Huang et al. 2020), and is often combined with ground monitored in-situ data which are relatively more accurate and can be utilized for validation of remotely sensed data for network rationalization (Wang et al., 2017). The above-mentioned real-world practical problems often tend to keep the administrative authorities fixated on ground-based monitoring stations for hydrological and hydrometeorological data, while progressively working towards the inclusion of dense AWS in existing manual monitoring networks.

Several countries have set up hydrometeorological stations following the guidelines of the World Meteorological Organization (WMO) and Commission for Instruments and Methods of Observation (CIMO). For instance, Canada has the highest density of AWS network with 384 weighting type gauging stations, whereas the USA, Germany, Sweden, Slovakia, and Norway have established networks of 331, 134, 111, 83, and 70 hydrometeorological stations respectively (Nitu and Wong 2010). Designing station networks and deciding their optimum number is hindered by region-specific practical and physical challenges such as hydro-climatic constraints, space limitations, accessibility and security issues, equipment vandalism risks, etc., particularly in densely populated countries like India. In such cases, hydrometeorological station networks that are already set up conventionally can be reassessed by employing various rationalization techniques. Although several studies exist that focus on designing rainfall and discharge monitoring networks (Al-Zahrani and Husain, 1998; St-Hilaire et al. 2003; Xu et al. 2013; Adhikary et al. 2015), research on rationalization of AWS for simultaneous monitoring of multiple hydrometeorological observations (such as precipitation, relative humidity, temperature, etc.) can hardly be found in the existing literature. Given the rise in frequency of individual extreme events and compound extreme events, i.e., the occurrence of multiple events concurrently which are driven by multiple hydrometeorological observations that are inter-dependent as well as interconnected (Hao et al., 2018), there has been an escalation in research related to the analysis of dependence and influence of multiple hydrometeorological observations over each other, since it is their combination which makes an extreme event unprecedented and amplifies the adverse impacts (Zscheischler et al. 2018). This further re-emphasizes the necessity and importance of consistent and accurate monitoring of multiple hydrometeorological datasets required to carry out these studies. It should also be noted that the existing rationalization approaches used to design the optimal network of single observation may not be applicable for AWS networks that monitor multiple observations. Additionally, the validation of a rationalized hydrometeorological

station network is important so that it fulfills the purpose of installation, which has been less commonly observed in the past literature.

Hence, our study proposes a comprehensive framework for the rationalization of AWS networks that addresses the aforementioned gaps in existing rationalization approaches. The notion of the developed rationalization framework in this study is synonymous with the notion of optimization. Optimization techniques are typically based either on conventional mathematical programming techniques (MPT) such as linear programming, dynamic programming, multiple criteria analysis, and fuzzy-based approaches or evolutionary techniques like genetic

algorithm, particle swarm optimization, etc. (Varekar and Karmakar, 2017). Our proposed framework consists of multivariate statistical approaches to rationalize the existing AWS networks based on multiple hydrometeorological observations. The proposed rationalization framework is not a planning tool for the placement of stations rather a tool for performance evaluation of an existing network to capture the temporal variability of hydrometeorological observations. More importantly, our framework attempts to validate whether the rationalized networks are economically sustainable and as effective as original networks in capturing the quality and reliability of the information. For

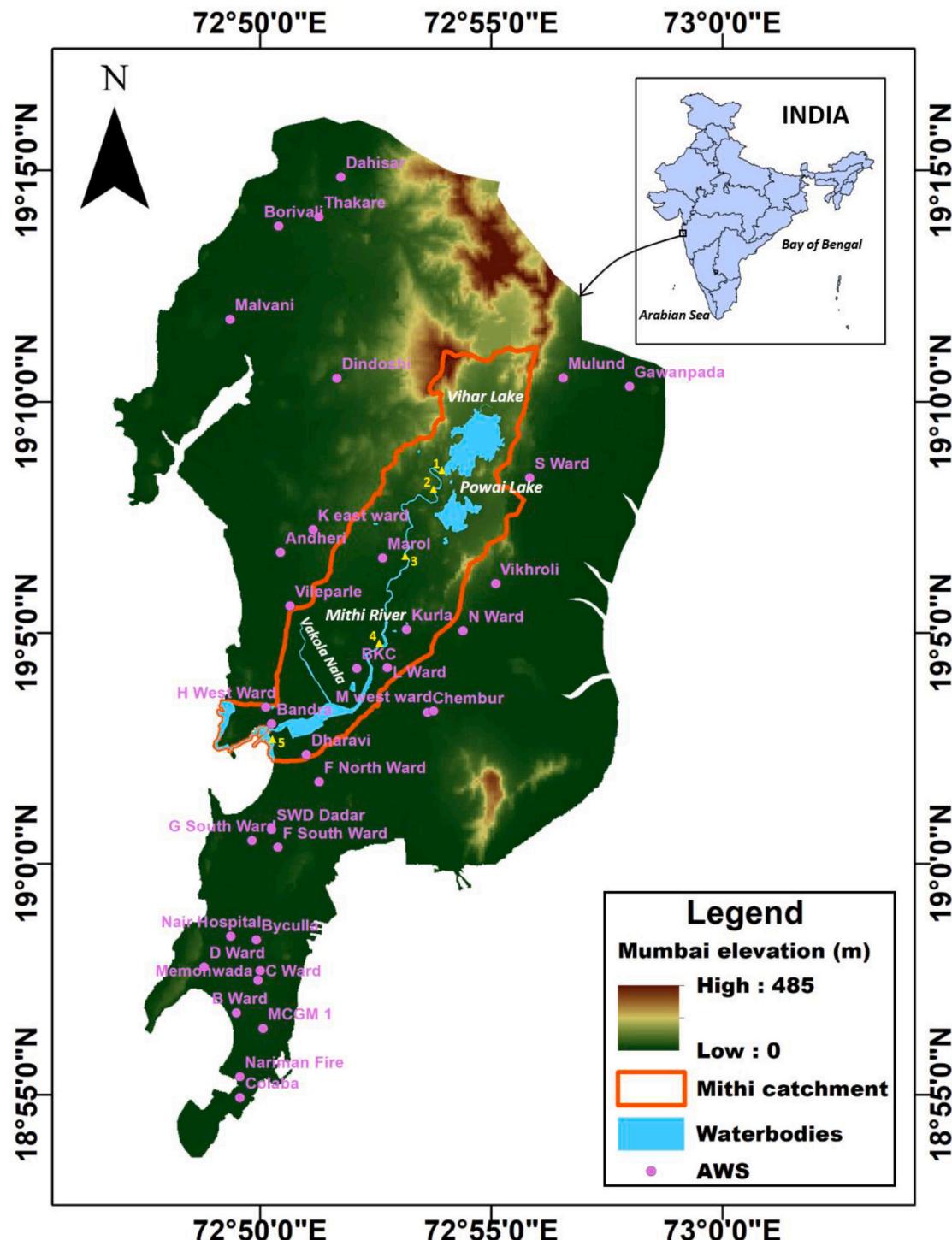


Fig. 1. Location of 35 MCGM Automatic Weather Stations (AWS), major water bodies like Mithi river, Vihar, and Powai lakes, and Mithi catchment in Mumbai used for the current study. 1–5 triangular symbols mark the change in gradients along the catchment.

demonstrating the validation of the rationalized network at flood inundation and hazard level, we employed our framework over the highly flood-prone Mumbai city (Hallegeatte et al., 2013) that exhibits high spatiotemporal variability of rainfall (Singh et al. 2017). Nonetheless, it will prove most useful for thickly populated tropical urban catchments, especially the ones located along coastlines wherein high spatiotemporal variability is observed in hydrometeorological observations and installation plus maintenance of dense AWS networks are challenging tasks. The existing networks can be evaluated with this framework by addressing the above-mentioned challenges and being cost-effective.

2. Study area and data description

The proposed framework of rationalization has been demonstrated over Mumbai city, an urban Indian catchment that extends over an area of 438 km² between 18.00°–19.20° N and 72.00°–73.00° E (Fig. 1). Mumbai, being the financial capital of India, is home to over 20 million people. It faces incessant precipitation events almost every year owing to its orography, erratic monsoon rains under the influence of southwest monsoons during June–September, and urban heat island effect. Hence this city, being a low-lying watershed and due to its geographic location along the coastline, is often subjected to substantial flooding. The city which was originally a cluster of 7 islands, is primarily a low-lying area with an average elevation of 14 m except in certain areas of Northern Mumbai where elevation rises to around 400 m at the hill of Sanjay Gandhi National Park. The rapidly increasing population and unprecedented expansion and urbanization have further increased its vulnerability to floods. The Mithi river plays a crucial role in the drainage network of Mumbai since multiple outfalls including Vakola, a major Nala, drain the excess water due to rains directly into the river. It originates from Vihar Lake and gathers water from Vihar, Tulsi, and Powai lakes, and travels 18.4 km accompanied by rapid change in a gradient from upper to lower reaches. The variation in elevation (points 1–5 in Fig. 1) along the catchment comprises of 4 reaches (Zope et al. 2015):

- 1–2: Very steep bed gradient (1:200) between the origin of Mithi River to Jogeshwari Vikhroli Link Road (JVLR).
- 2–3: Steep bed gradient (1:450) between JVLR road to Mathuradas Vasani (MV) road.
- 3–4: Moderate bed gradient (1:850) between MV road to Chatrapati Sivaji Terminal (CST) bridge.
- 4–5: Flat bed gradient (1:4000) between CST bridge to Mahim causeway in Bandra Kurla Complex (BKC) area.

Hence, during a heavy precipitation event, the water carrying capacity of the river channel gets exceeded which results in the overtopping of stormwater across the river banks. The insufficient water carrying capacity of the river, which further confluences with the Arabian Sea at Mahim creek - a low-lying area with a high density of multi-storied office buildings - has resulted in the Mithi river catchment being identified as one of the most flood-prone regions in Mumbai. Situated along the western coastline of the Indian subcontinent, the city experiences high spatiotemporal variability of rainfall (Lokanadham et al. 2009; Zope et al. 2015). Several studies in the past have investigated rainfall pattern and their associated dynamics (Lei et al. 2008), predicted extreme rainfall using statistical techniques (Nayak and Ghosh 2013; Shastri et al. 2017), or used physics-based numerical weather prediction models to forecast rainfall over Mumbai (Bohra et al. 2006; Paul et al. 2018; Patel et al. 2019; Patel et al. 2020). Most of these studies have highlighted the limitation of unavailability and reliability of sub-hourly rainfall data, which is highly essential for providing decent forecasts especially for urban areas where few hours of rainfall may result in flash floods. Furthermore, during the 2005 floods, the two rainfall-recording weather stations set up by IMD (at Colaba and Santacruz) proved inadequate to be representative of the entire rainfall of Mumbai due to high variability of daily rainfall ranging from 7 cm in

Colaba to 94 cm in Santa Cruz. After this infamous heavy rainfall event, where Mumbai witnessed around 944 mm rainfall over 24 h, resulting in huge loss of life and property, 26 AWS were installed by MCGM in June 2006 (MCGM, 2007), which later increased to 60 AWS over a period of time for continuous and consistent measurement. These AWS can record other significant hydrometeorological observations along with rainfall such as relative humidity, minimum and maximum temperature, wind speed and direction, pressure.

Mumbai city being the commercial capital of India faces high economic diversity, space constraints, and a rapid increase in unplanned settlements. These ground problems intensify the possibility of sabotage of measuring instruments thus impeding the implementation of any conventional rationalization technique in the city. The governing authority MCGM has thus installed the AWS at existing local fire station premises as they being government owned are justifiably safer and secure from local vandalism. However, continued monitoring and inventorization of various hydrometeorological observations are quite cost-intensive. The proposed framework aims to identify the set of AWS which shows high temporal variability and the remaining stations can be curtailed down, which may substantially reduce the maintenance and operation costs.

Singh et al. 2017 investigated the sub-hourly rainfall data recorded from 26 AWS stations for June, July, August, and September (JJAS) months of years 2013 and 2014 and found pattern-less high spatiotemporal variation in the rainfall. Before this, the available rainfall data observations for 2006–2014 were thoroughly analyzed and the sub-hourly rainfall data for 26 out of 60 AWS for the years 2013 and 2014 were found to be complete. The current study further considers the hydrometeorological observations for the JJAS months of 2015–2018. The rainfall and relative humidity data of 35 AWS over Mumbai for these 4 years with a temporal resolution of 15-min has been utilized in this analysis. These stations have been selected based on the consistency, reliability, and completeness of these two hydrometeorological observations. Multiple hydrometeorological observations for a greater number of stations may be considered to demonstrate this framework when a sufficient database is procured over the subsequent years. The sub-hourly rainfall and relative humidity data were used to inspect their spatiotemporal variability over Mumbai. However, for further analysis, since the relative humidity does not show much variation throughout the day and the presence of a large number of zero values in sub-hourly rainfall datasets did not provide any significant and clear results, the current study utilized the rainfall and relative humidity on a daily time scale for demonstrating the proposed rationalization framework. Details of the rain gauging station network have been shown in Fig. 1 and Table 1.

2.1. Spatio-temporal characteristics of rainfall

The sub-hourly rainfall data of five representative extreme rainfall days of each year between 2015 and 2018 were analyzed to understand the spatial behavior of the rainfall across the city. The nonparametric Kendall τ correlation coefficients for rainfall between every possible pair of the 35 AWS were plotted against the Euclidean distance between the corresponding stations (Fig. 2). The nonparametric Kendall τ metric can consider both continuous as well as ordinal values and does not have to suffice to normal distribution only and therefore overcomes the limitations linked with parametric correlation coefficients, such as Pearson's correlation coefficient (Vittal et al. 2015). Hence for x number of AWS (in the present study $x = 35$), a set of $(x(x-1)/2)$ values, i.e., 595 correlation coefficients were computed for each of the 5 extreme rainfall days during 2015–2018. It is observed that closer stations show higher correlations while farther stations show lower correlations. The significant correlations between pairs of AWS are diminishing very fast with an increase in distance between the stations.

Fig. 3 represents the diurnal distribution of the sub-hourly rainfall data. The rainfall data used has been obtained by spatially averaging the

Table 1

Details of AWS operated by MCGM used in the current study.

Sl. No.	Name of Station	Latitude (N)	Longitude (E)	Elevation (m)
1	Andheri Fire Station	19.11227	72.84067	14
2	B Ward Fire Station	18.9461	72.82481	9
3	Bandra Fire Station	19.05039	72.83748	6
4	BKC Fire Station	19.07038	72.86812	3
5	Borivali Fire Station	19.22991	72.83994	10
6	Byculla Fire Station	18.97242	72.832	11
7	C Ward Fire Station	18.9614	72.83339	9
8	Chembur Fire Station	19.05456	72.89361	13
9	Colaba Fire Station	18.91547	72.82608	10
10	D Ward Fire Station	18.96262	72.81316	9
11	Dahisar Fire Station	19.24764	72.86245	17
12	Dharavi Fire Station	19.03929	72.84995	6
13	Dindoshi Fire Station	19.17503	72.86099	27
14	F North Ward Fire Station	19.02942	72.85461	6
15	F South Ward Fire Station	19.00593	72.83973	9
16	G South Ward Fire Station	19.00832	72.83038	8
17	Gawanpada Fire Station	19.17205	72.96637	4
18	H West Ward Fire Station	19.05639	72.8354	12
19	K east ward Fire Station	19.12036	72.85237	22
20	Kurla Fire Station	19.08449	72.88607	11
21	L Ward Fire Station	19.07065	72.8792	10
22	M westward Fire Station	19.05497	72.89578	20
23	Malvani Fire Station	19.19625	72.82244	7
24	Marol Fire Station	19.1101	72.87751	25
25	MCGM 1 Fire Station	18.94044	72.83442	15
26	Memonwada Fire Station	18.95782	72.83264	8
27	Mulund Fire Station	19.17516	72.94255	28
28	N Ward Fire Station	19.08393	72.90644	14
29	Nair Hospital Fire Station	18.9737	72.8227	10
30	Nariman Fire Fire Station	18.92307	72.8261	15
31	S Ward Fire Station	19.13903	72.93045	14
32	SWD Dadar Fire Station	19.01231	72.83742	7
33	Thakare Fire Station	19.23324	72.85449	20
34	Vikhroli Fire Station	19.10094	72.91817	17
35	Vileparle Fire Station	19.09292	72.8441	8

rainfall measured using the arithmetic mean technique at all the considered AWS over the study area. The probability density functions (PDFs) of the sub-hourly (15-min) rainfall have been derived for four different duration spells in a day: 0–6 h (red color, night), 6–12 h (blue color, morning), 12–18 h (green color, afternoon), 18–0 h (grey color, evening). The pie charts represent the quantity of rainfall during corresponding spells. The standard deviations corresponding to each PDF have also been presented in parenthesis inside the pie chart to describe the spatial distribution of rainfall during different spells. The PDFs and standard deviation values indicate that the spatial variation of the rainfall is highest during evening spells and lowest during morning spells in 2015 and 2016. However, in 2017, the spatial variation of rainfall is high during all spells. The year 2017 received a substantially higher amount of rainfall than the other three years. This year also witnessed the worst flooding due to heavy rainfall between 28th and 29th August 2017, since the historically infamous heavy precipitation event of July 26th, 2005. In 2018, the highest and lowest spatial variation in rainfall is observed during night and evening spell respectively. The pie charts show that the majority of rainfall during 2015 and 2018 occurred during the morning spells. Similarly, the highest rainfall in 2016 and 2017 occurred during night and afternoon periods respectively. The evening spells for all four years show comparatively less quantity of rainfall than other spells. Fig. S1 illustrates the monthly distribution of rainfall for all four years. The June and July months contribute maximum to the total rainfall for all four years as compared

to the other two months. This random variation in the diurnal distribution of rainfall reveals its patternless behavior for the study area.

2.2. Spatio-temporal characteristics of relative humidity

The sub-hourly relative humidity data of five representative extreme rainfall days of each year between 2015 and 2018 were analyzed to understand the spatial behavior of the relative humidity across the city. The nonparametric Kendall τ correlation coefficients for relative humidity between every possible pair of the 35 AWS were plotted against the Euclidean distance between the corresponding stations (Fig. S2). Unlike rainfall, a high correlation is found between almost all AWS within the study area for relative humidity observations for the 5 extreme days of rainfall between 2015 and 2018. The number of significant correlations between a pair of AWS is greater than the non-significant correlations out of the total 595 correlation coefficients computed for all possible pairs of AWS for all 4 years. This can be attributed to the lower variation in relative humidity observations as compared to rainfall which is a more dynamic observation. The diurnal distribution of the spatially distributed sub-hourly relative humidity data is analyzed in Fig. S3. The PDFs of the sub-hourly (15-min) relative humidity have been derived for the four different duration spells in a day: 0–6 h (red color, night), 6–12 h (blue color, morning), 12–18 h (green color, afternoon), 18–0 h (grey color, evening). Unlike rainfall, the relative humidity shows clear diurnal patterns for the study area considered. The relative humidity is highest during night spells and lowest during afternoon spells for years 2015–2018. It is similar during the morning and evening for the concerned period. Interestingly, spatial variation of relative humidity is also similar during these spells as indicated by the PDFs and standard deviation values. The standard deviation values are more or less similar in 2015 and 2016. However, in the year 2017, the standard deviation values are higher which denotes a higher degree of spatial variation.

The spatiotemporal analysis of the sub-hourly rainfall and relative humidity data is an attempt to verify the earlier conclusions established by Singh et al., 2017 with refurbished data sets from the years 2015 to 2018. The study further ascertained the claims and previous findings of high spatial as well as temporal variability of rainfall across the city. However, the spatial and temporal variability of relative humidity is comparatively less than that of rainfall. This substantiates the requirement of a dense AWS network for reliable and consistent monitoring of all the observations, which will enable the creation of a high quality and long-term data set for future research and analyses. Nevertheless, a dense AWS network is accompanied by high maintenance and operation costs. The present work is a novel and well-thought-out attempt to explore the hydrometeorological dataset (created from the AWS installed by MCGM) and establish a framework to rationalize the AWS network based on the ability of the stations to capture the variability of the multiple observations. Rationalization of AWS over this area is further challenging due to pattern-less observations as concluded from the spatiotemporal analysis and hence this framework can be implemented over other coastal urban catchments.

3. Methodology

3.1. Description of the proposed framework

The framework proposed for the current study is a two-stage technique represented in Fig. 4 with detailed elucidation in the subsequent sections. The notion of the first stage is to quantify the variability of multiple hydro-meteorological observations by utilizing a multivariate statistical approach. The second stage prioritizes the AWS based on their ability to capture the spatiotemporal variability of these observations through a multi-attribute decision-making technique. The variation of these observations is evaluated through Principal Component Analysis (PCA) following which the stations are ranked through Technique of

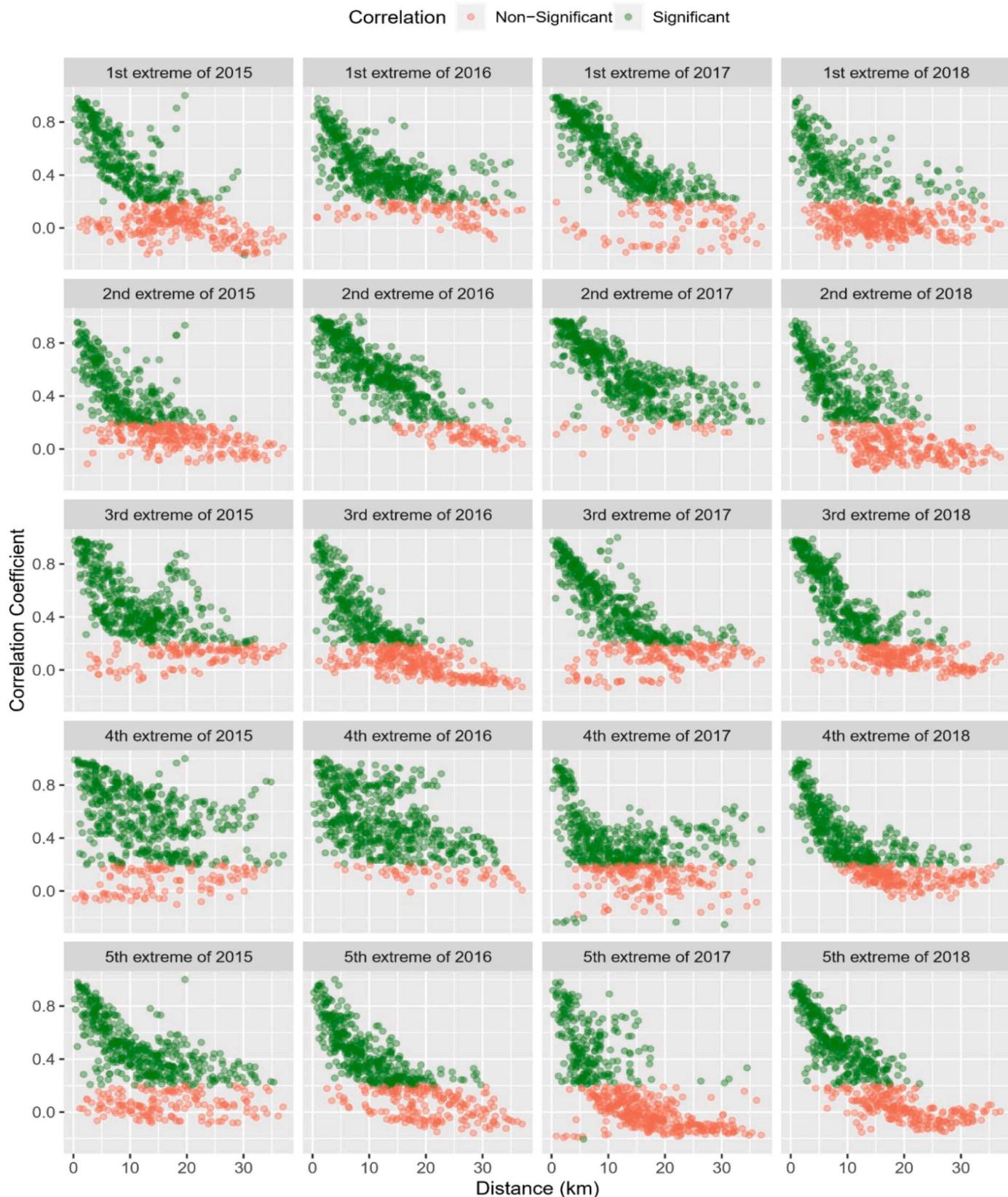


Fig. 2. Nonparametric Kendall's τ correlation coefficients between rainfall at every possible pair of the rain gauge stations against the Euclidean distance between corresponding stations, for each of the 5 extreme days for 2015–2018. Green circles represent significant correlations that lie above the 5% level of significance while orange circles that lie below this level represent non-significant correlations.

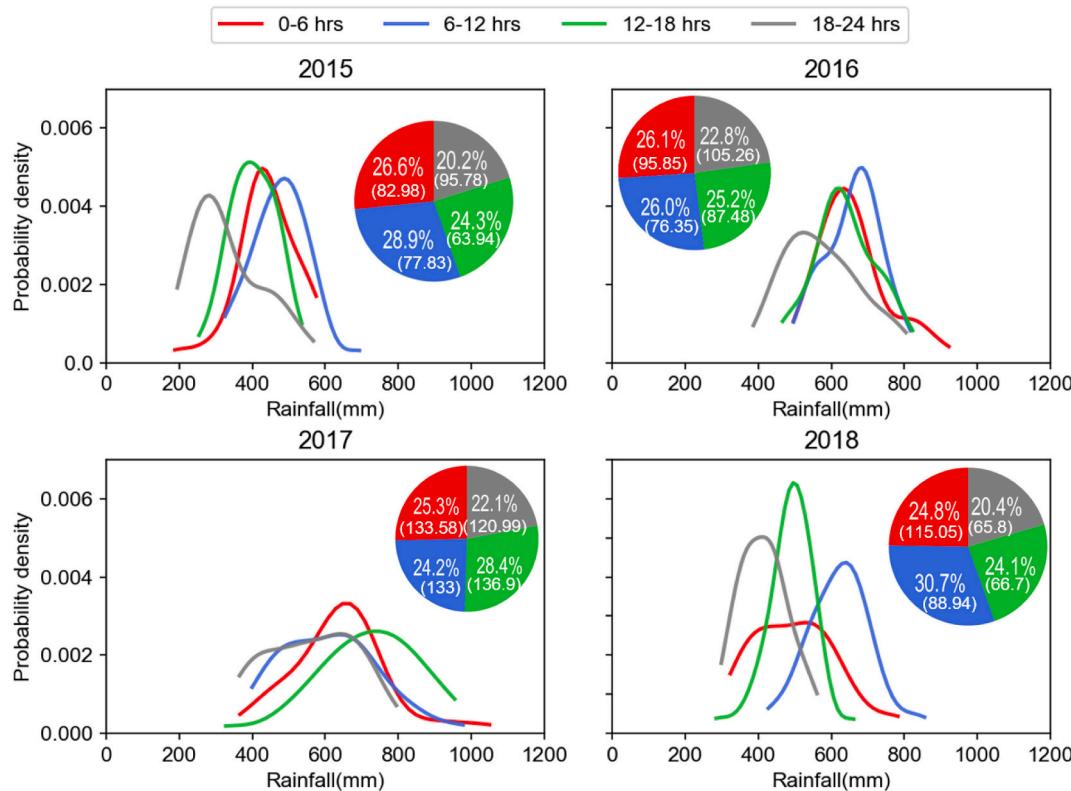


Fig. 3. Diurnal distribution of rainfall for years 2015–2018 in terms of the probability density functions of sub-hourly (15 min) rainfall and pie-chart reflecting the total amount of rainfall (in %) during all four duration spells: 0–6 h (red color, night), 6–12 h (blue color, morning), 12–18 h (green color, afternoon), 18–0 h (grey color, evening); values in parenthesis in piechart represent the standard deviation of sub-hourly rainfall data

Order Preference by Similarity to Ideal Solution (TOPSIS), a multiple attribute decision-making method.

3.1.1. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) aims to decorrelate and reduce the dimensionality of a dataset consisting of a large number of interrelated variables while conserving the variation to the largest extent possible (Jackson 1991). This technique is widely implemented in a selection of hydrological, hydrometeorological as well as water quality monitoring stations. The stations depicting similar characteristics i.e., areal homogeneity over a period of time can be grouped for rationalization of the network (Morin et al. 1979). In this method, the Eigen decomposition is performed on the covariance matrix of the concerned variables. The estimated Eigenvalues and corresponding Eigenvectors are arranged in order of variability explained. The original concerned variables are multiplied to the Eigenvectors to obtain principal components. A few components are selected which account for most of the variability in the concerned variables. Hence, the information about the most important variables is retained by PCs, which can describe the whole dataset with the least loss of information (Helena et al. 2000; Singh et al. 2004). Following this, FA further reduces the contribution of less significant variables, by extracting varifactors (VFs), a new set of variables obtained through rotation of axis defined by PCA. PCs comprise a weighted linear combination of original observable variables while VFs comprise of the unobservable, latent, and hypothetical variables (Wunderlin et al. 2001).

In the present study, the first stage of the proposed framework comprises of PCA in which the AWS are considered as variables for the two hydrometeorological observations, rainfall (P) and relative humidity (RH). The significant PCs are extracted for each of the observations individually and subsequently subjected to varimax rotation to obtain VFs. In the beginning, the hydrometeorological observations

concerning all variables (AWS stations) i.e., $\{P\}_{35 \times 1}$ and $\{RH\}_{35 \times 1}$ normalized to assign equal weightage during the analysis. Following this, the eigenvalues $\{\lambda\}_{35 \times 1}$ and eigenvectors $\{V\}_{35 \times 1}$ for each parameter are computed using the corresponding covariance matrices. The principal components accounting for a major proportion of the variation in the dataset are utilized further. In the present analysis, the first principal components, i.e., the ones which explain 83% and 87% of the variance for rainfall and relative humidity respectively are considered for varimax rotation as it captured most of the variability. Finally, the factor loading matrix is established and varimax rotation is conducted on it to identify the significant variables (AWS in the present context). A higher factor loading value corresponds to greater temporal variability in hydrometeorological data since each variable represents individual AWS in this study and PCA is performed over each variable. The variability of the parameters is thus obtained over each station and the temporal variability is quantified. The stations with a factor loading value greater than 0.65, the design threshold value considered in this study, are considered significant (Ouyang 2005; Musthafa et al. 2014). Since the observations are highly dynamic and non-linear, this conservative threshold is selected. This analysis completes the first stage of the framework, which estimates the factor loading of all AWS for all the observations and are considered for the second stage of the framework.

3.1.2. Multiple-attribute decision-making (MADM)

The second stage of the framework implements the Multi-attribute decision making (MADM) method, a class of discrete multi-objective optimization methods, in which a set of alternatives/options are assessed based on certain criteria to make an optimal selection and ranking the alternatives with the greatest level of satisfaction (Dyer et al. 1992). Amongst the various MADM methods, the Technique of Order Preference by Similarity to Ideal Solution (TOPSIS) and Analytic Hierarchy Process (AHP) are the most widely executed methods. In the

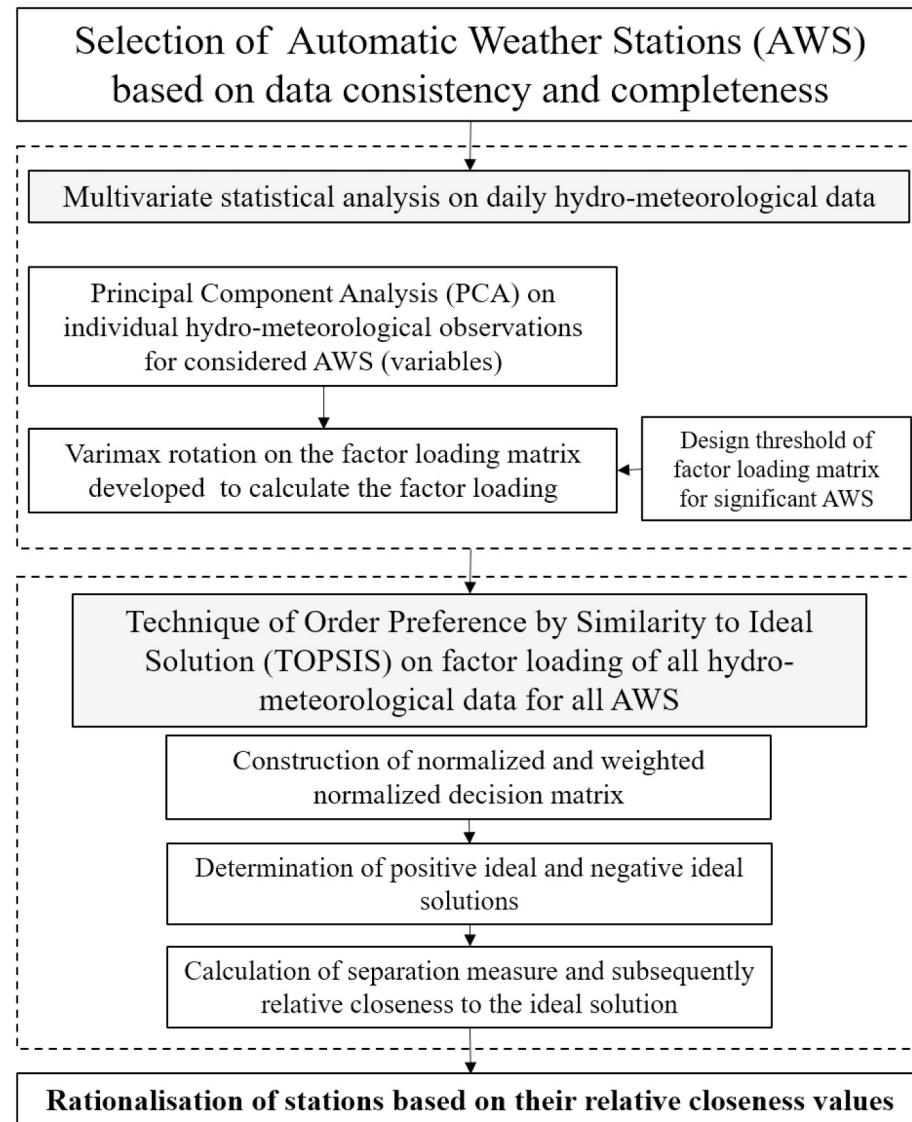


Fig. 4. Framework for rationalization of AWS stations based on their ability to capture spatiotemporal variability.

current study, TOPSIS, a utility-based compensatory technique, which was developed by [Hwang and Yoon \(1981\)](#) has been implemented to rank the alternatives (AWS in the current problem) based on their ability to capture the spatiotemporal variability of various hydrometeorological observations (rainfall and relative humidity for this study). This method has an edge over other methods as (1) a huge number of alternatives and criteria can be evaluated in this method; (2) it does not require a large number of subjective inputs; (3) it has a programmable and easy to understand algorithm; and (4) it has a comparatively consistent ranking scheme since the chances of rank reversal are reduced ([Kalbar et al. 2012](#)). Moreover, TOPSIS has no restriction on the number of attributes unlike AHP, which may not work efficiently as the number of attributes (both criteria and alternatives) increases. A distance-based method is utilized in this approach for quantification and thereafter comparison of the various alternatives over a set of criteria, i.e., rainfall and relative humidity observations in this case study. TOPSIS utilizes a computationally efficient mathematical algorithm to rank the alternatives and has found wide applicability as decision support means in many relevant fields such as environmental decision-making, water resources management, marketing management related problems. However, this robust technique has not been applied in the rationalization of hydrometeorological monitoring networks to our knowledge. The steps

involved in the second stage of the framework have been enumerated below.

- The factor loading values obtained from FA performed for each criterion, i.e., rainfall and relative humidity are normalized for all the alternatives, i.e., AWS and a vector normalized score matrix is constructed as per eq. 1:

$$n_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad \forall j \quad (1)$$

where x_{ij} is the value of i^{th} alternative for j^{th} criteria of the original score matrix; n_{ij} comprises the values of the normalized matrix. In this case, we have 35 alternatives and 2 criteria.

- In the next step, a weighted normalized score matrix is constructed by assigning weights to the criteria j as per eq. 2,

$$w n_{ij} = n_{ij} * w_j \quad \forall i, j \quad (2)$$

where w_j is the weight assigned to the criteria and wn_{ij} refers to the values of the normalized matrix.

The positive ideal solutions (I^+) and negative ideal solutions (I^-) are determined as per eqs. 3 and 4 respectively

$$I^+ = \left\{ \left(\max_i w_{nj} | j \in J \right), \left(\min_i w_{nj} | j \in J' \right); \forall i \right\} = \{ w_{n_1}^+, w_{n_2}^+, \dots, w_{n_m}^+ \} \quad (3)$$

$$I^- = \left\{ \left(\min_i w_{nj} | j \in J \right), \left(\max_i w_{nj} | j \in J' \right); \forall i \right\} = \{ w_{n_1}^-, w_{n_2}^-, \dots, w_{n_m}^- \} \quad (4)$$

where J and J' correspond to benefit and cost criteria respectively and $J + J' =$ the Total number of criteria. In our study, all the criteria considered are benefit types.

- iii. The separation measures from both positive ideal solution (S^+) as well as negative ideal solution (S^-) are calculated from the n -dimensional Euclidean distance between criteria as per eqs. 5 and 6 respectively.

$$S_i^+ = \sqrt{\sum_{i=1}^n (w_{nj} - w_{n_j}^+)^2} \quad \forall i \quad (5)$$

$$S_i^- = \sqrt{\sum_{i=1}^n (w_{nj} - w_{n_j}^-)^2} \quad \forall i \quad (6)$$

- iv. Following this, the relative closeness (R_i^+) to the positive ideal solution is calculated for all the alternatives as per eq. 7.

$$R_i^+ = \frac{S_i^+}{S_i^+ + S_i^-} \quad (7)$$

- v. Finally, the alternatives are ranked based on the descending order of relative closeness values. The TOPSIS renders an optimal solution that is closest to the positive ideal solution and farthest from the negative ideal solution.

3.2. Assessment of the proposed framework

The credibility of the proposed multivariate statistical framework developed for the rationalization of AWS is assessed by evaluating whether the rationalized stations are representative of the variation of hydrometeorological observations over the entire study area. In the current study, PDFs for the two spatially averaged hydrometeorological observations considered i.e., daily rainfall and relative humidity are derived for all the AWS and the rationalized AWS. The PDFs for daily rainfall and relative humidity, which are spatially averaged values obtained using the arithmetic averaging technique, have been derived for all the original AWS network and the rationalized AWS network. PDFs have been estimated using kernel density estimation, which is a non-parametric approach. During the initial stage of analysis, the PDFs were derived for all stations individually for higher spatial resolution like that of flood inundation and hazard maps. However, representing multiple PDFs for all 35 stations in a single plot did not disclose any prominent pattern and thus no conclusion could be deduced from it. Therefore, two sets of data were used for comparison, the first set consisting of arithmetic average values of rainfall and relative humidity for all AWS in the city and the second set consisting of arithmetic average values of the rationalized AWS obtained from the analysis. The PDFs are compared by computing the Kullback-Leibler (KL) Divergence values for each parameter individually (Kullback 1959). The KL divergence value also known as information divergence or relative entropy, is widely

implemented in data mining literature, which interprets the non-symmetric measure of the difference between two PDFs for a particular variable. It can be defined by eq. 8:

$$D_{KL}(X(a) \parallel Y(a)) = \int_{-\infty}^{\infty} X(a) \ln \frac{X(a)}{Y(a)} da \quad (8)$$

where $X(a)$ and $Y(a)$ are the probability distributions for considered observations for all the AWS and rationalized AWS respectively and $D_{KL}(X(a) \parallel Y(a))$ is the divergence of $Y(a)$ from $X(a)$ and is a non-negative measure. This implies that when $D_{KL}(X(a) \parallel Y(a))$ is zero, both the distributions are identical. The correlation coefficients for the hydrometeorological observations rainfall and relative humidity between spatially averaged data from all stations of the original AWS network and 22 individual stations of the rationalized AWS network are also determined using eq. 9.

$$r = \frac{n \sum_{i=0}^n P_i Q_i - \sum_{i=0}^n P_i \sum_{i=0}^n Q_i}{\sqrt{\left[n \sum_{i=0}^n P_i^2 - \left(\sum_{i=0}^n P_i \right)^2 \right] \left[n \sum_{i=0}^n Q_i^2 - \left(\sum_{i=0}^n Q_i \right)^2 \right]}} \quad (9)$$

where P_i and Q_i represent the value of hydrometeorological observation on the i^{th} day for spatial averaged AWS (the original network of 35 stations) and each station of the rationalized AWS network, respectively. The term 'n' represents the total number of observations.

Further, the present study for the first time attempts to evaluate the efficiency of this proposed rationalization framework at the inundation and hazard level of floods, since the foremost objective of the study was to evaluate the usefulness of the existing AWS network. The dense network of Mumbai city was set up by the administrative authorities to study and analyze the spatiotemporal pattern of the hydrometeorological observations and address the frequent flooding scenario of the area after the infamous flooding event of 26th July 2005. The city witnessed extremely high rainfall over a 24-h period which started at 2 PM and continued till the next day 2 PM when an average 94 cm rainfall was recorded accompanied by a high tide of 4.48 m. The event was heavily localized and had high variability over different areas of Mumbai where Bhandup, Colaba, Dharavi, Malabar hill, Vihar lake, and Santacruz recorded 81 cm, 7 cm, 49 cm, 7 cm, 104 cm, and 94 cm rainfall respectively (Bhohra et al. 2006). This resulted in flooding over most of the parts of the city, with low lying areas like Dharavi and Bandra-Kurla Complex being the worst affected ones. The transportation network of the city also came to a halt due to excessive waterlogging over roads, railway tracks, and airports. This event resulted in economic losses worth 2 billion US\$ and around 500 human deaths. In the current study, a hydrodynamic flood model was developed for the highly flood-prone Mithi river catchment and also due to the availability of fine-resolution DEM (an essential requirement for any flood model), over this area. A 3-way coupled flood model has been developed in MIKE FLOOD platform (Vojinovic and Tutulic 2009), over Mithi river catchment in Mumbai (Fig. 1), considering all flood influencers i.e., flows through the channel, overland, storm-water drains, and tidal influences to generate flood inundation and hazard maps. Rainfall is a significant hydrometeorological parameter that is required in any flood model and hence we have used the sub-hourly rainfall data obtained from AWS in this study. The MIKE FLOOD modelling framework developed for the area which integrates MIKE 11 for river channel flow and MIKE 21 HD FM for flood plain flow is represented in Fig. 5. Table 2 enlists the data used for the development of the hydrodynamic flood modelling framework. The MIKE 11 model utilizes an implicit finite difference scheme by Abbott and Ionescu (1967) to solve the continuity and momentum equations and simulate the river flow using suitable boundary conditions i.e., upstream discharge and water levels (DHI 2017). The river shape file along with its cross-section details which were surveyed across different chainages of the river and Vakola Nala and channel roughness

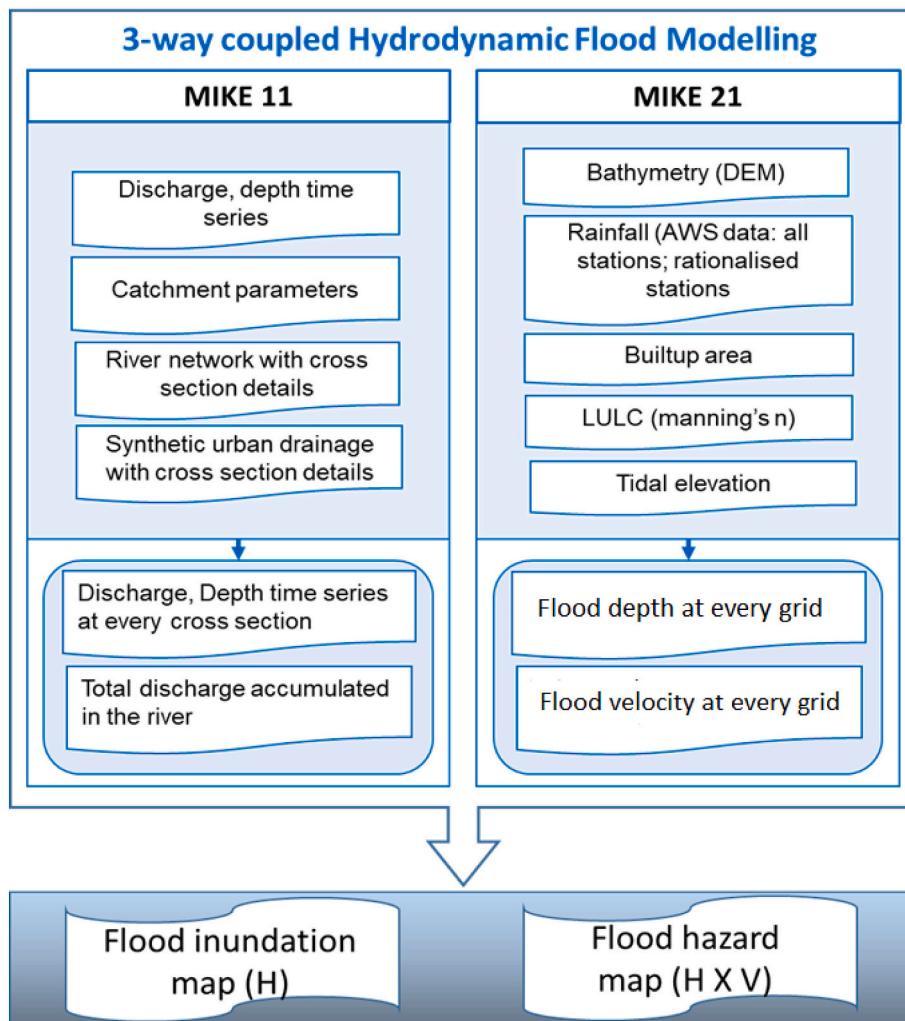


Fig. 5. 3-way coupled MIKE FLOOD hydrodynamic flood modelling framework proposed for flood inundation and hazard mapping.

is provided as input to the MIKE 11 model. The cross-sections of stormwater networks located in the catchment are extracted from DEM with the MIKE HYDRO tool and the water depth and discharge can be simulated at various stretches of the river. The MIKE 21 HD FM model utilizes an unstructured non-orthogonal triangular mesh based on two-dimensional incompressible Reynolds-averaged Navier-Stokes equations to generate numerical solutions under hydrostatic pressure and Boussinesq assumptions (Papaioannou et al. 2016; Mohanty et al. 2020). This mesh created by DEM data, land open and lake boundary along with building layers, to solve shallow water equations (Kadam and Sen 2012) provides an advantage of the flexibility of resolution over the modelling area over the classical rectangular grid but with an optimal design. Hence, the significant elevation, complex structures, and heterogenous areas within the same modelling area can be given greater importance by increasing the resolution over those areas as compared to the homogenous areas. This simplifies the computational complexity in terms of lesser time requirements without affecting the model outputs. Manning's resistance values for various land use classes are provided to account for flood plain roughness (Mohanty et al. 2018). The MIKE hydrodynamic flood model can also perform simulations over a rapid change in elevations and steep slope regions which gives it an edge over flood models. Finally, the 1-D model with stormwater drainage network is coupled with the 2-D model, and 3-way coupled model developed is run for two sets of rainfall data of a heavy rainfall event which occurred in 18th–19th June 2015: 1st set of run considers for all AWS present within the Mithi river catchment, i.e., 8 stations; 2nd set of run considers

the rationalized AWS present within the catchment, i.e., 5 stations. The Thiessen polygon method has been used to areally distribute the point rainfall values of AWS across the catchment which is provided as an input to the flood model. The flood maps derived from these two sets of runs are compared to evaluate the accuracy of the rationalized network to that of the entire AWS network in terms of the spatial extent of the flood. Conventionally, the flooding extent is expressed in terms of flood depth in the majority of the flood-related studies. However, another important component of the flood, i.e., deep water flows which may alleviate the extent of damage during an extreme flood scenario should also be incorporated during the analysis of these events. To address this issue, in our present analysis, flooding is represented as flood hazard maps (considering the product of depth and velocity which signifies momentum) along with the classical flood inundation maps (considering depth only) which have been utilized in previous studies (Tingsanchali and Karim 2005; Ghosh et al. 2020).

3.3. Rank reversibility analysis

Most of the decision-making methods primarily rely on the subjective judgments of decision-makers based on their expertise and experience. Similarly, all the attributes considered in TOPSIS are also associated with weights assigned as subjective information, which leads to uncertainty in the ranks of alternatives (Ishizaka and Nemery, 2013). Therefore, the alternatives should be analyzed by assessing the uncertainty apportioned with them due to weights by subjecting them to

Table 2

Details of data utilized in the development of 3-way coupled MIKE FLOOD framework.

Model	Data	Source	Remarks
MIKE 11	Network editor	Municipal Corporation of Greater Mumbai (MCGM)	Shapefiles of Mithi river catchment with the river and other water bodies
	Cross-section C/S editor	MCGM	C/S details at various chainages for Mithi River and Vakola Nala
	Hydrodynamic editor	Monitoring at site	Initial water depth and discharge
	Boundary editor	Monitoring at site	Time series (TS) of the discharge/ water depth at start and end of Mithi River
MIKE 21 HD FM	Digital elevation model (DEM)	World DEM	DEM for Mithi catchment with a built-up area
	Rainfall time series	MCGM automatic weather station data, IMD and Sherly et al. (2015)	Hourly rainfall data for the heavy rainfall event of 18th -19th June 2015
	Tidal time series	Indian National Centre for Ocean Information Services (INCOIS)	Tidal time series at the Mahim creek for the heavy rainfall event of 18th -19th June 2015
	Land use land cover (LULC)	National Remote Sensing Centre (NRSC), Hyderabad	LULC map of the Mithi catchment (to define the Manning's resistance of the flood plain)
	Building layer data	MCGM	Building layer data of Mithi catchment

numerous simulations with uniformly distributed values around the given weights (Yadav et al. 2019). In the current study, weights assigned to the loading values of the two hydrometeorological observations

considered are varied starting from 0%, 25%, 40%, and 50% in both directions. The change in the mean and standard deviation of relative closeness values due to variation in weights are evaluated to quantify the uncertainty associated.

4. Results and discussion

The two important hydrometeorological observations, rainfall, and relative humidity, recorded at the 35 Automatic Weather Stations are used in the analysis for rationalization of the network by using a multivariate statistical approach. In the first stage of the framework, FA/PCA is performed to identify principal stations for each of the observations. Fig. 6 represents the loading values of rainfall and relative humidity measured at a daily scale for all the stations from June to September of 2015–2018. The stations with loading values greater than 0.65 represented by blue color are considered significant and the stations with values less than 0.65 are represented by red color. Around 20 AWS have high loading values for rainfall, while 26 stations have high loading values for relative humidity. Hence, it can be deduced that about 55% of the stations are capable of capturing the variability reasonably well for each parameter. While some stations are found to have low loading values for rainfall, the loading values of relative humidity for most stations are high.

In the second stage of the framework, TOPSIS is performed on the loading values of rainfall and relative humidity obtained by FA/PCA by giving each parameter equal weight and ranking the AWS stations in terms of their ability to capture their variability. The stations are ranked in decreasing order of their relative closeness values deduced from eq. 7 (Fig. 7). When the relative closeness to ideal solution values for 35 AWS obtained from TOPSIS is plotted in decreasing order, it is observed that there is a steep change in slope after 22nd rank. It can be seen that 15 significant stations as per rainfall variability and all the stations as per relative humidity (as obtained from the PCA/FA) are included in the first

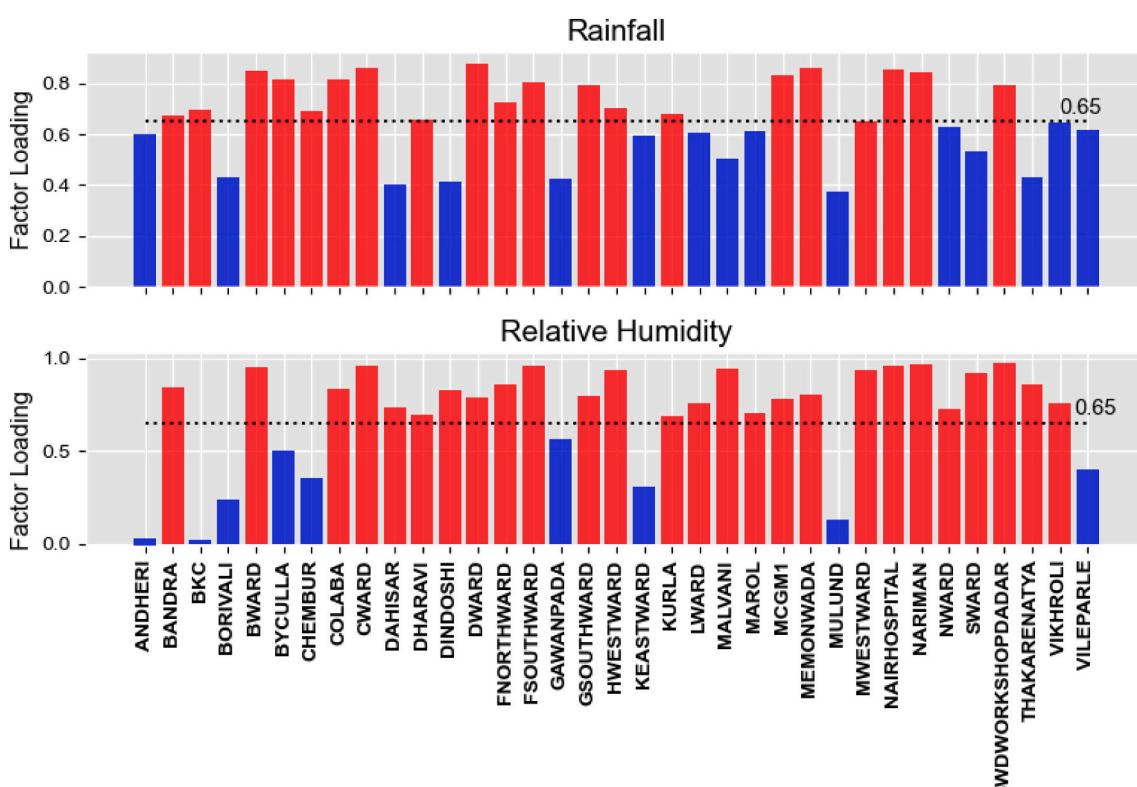


Fig. 6. Bar plots of factor loading values of Rainfall and Relative humidity (attributes selected for TOPSIS) for the stations considered in the study. Blue colored bars represent the stations that show significant (high) temporal variation and red-colored bars represent the stations with a non-significant (low) temporal variation.

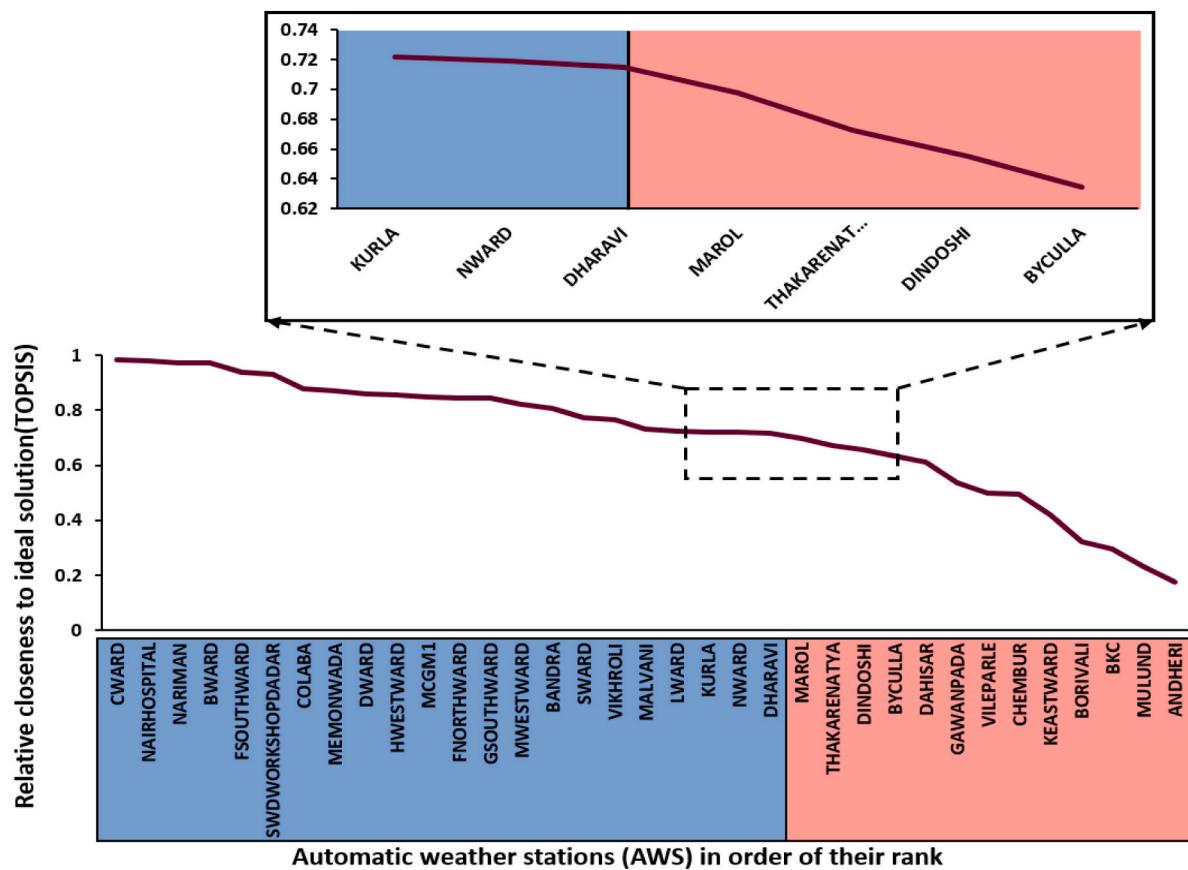


Fig. 7. Automatic weather stations identified to be significant and non-significant in terms of their ability to capture the spatiotemporal variation of rainfall and relative humidity.

22 ranked significant stations as found in the analysis by TOPSIS. This can be attributed to the fact that rainfall is a more dynamic observation than relative humidity and has greater variability during the monsoon season. Therefore, rainfall influences the final ranking of AWS more than relative humidity. The western ghats which run through the states of Maharashtra, Karnataka, and Kerala influence the monsoons in Mumbai to a great extent. The moisture-laden strong monsoon winds from the Arabian sea are orographically lifted upon collision against the Western Ghats (Jenamani et al. 2006). Along with this, phenomena such as the north-westward movement of low-pressure systems along monsoon trough from the Bay of Bengal, differential heating, northward moving mesoscale vortex over the north-east Arabian sea also plays a major role in the Monsoon over Mumbai overall resulting in high spatiotemporal variations of rainfall in the city (Lei et al. 2008).

Based on Fig. 7, the first 22 stations are considered significant and represented in Fig. 8. The observations monitored at the non-significant stations as obtained by this approach are suggested to be evaluated further. If no measurement errors and discrepancies are found during monitoring of the hydrometeorological observations at a station, and yet the loading values for the station is less, it thus implies that the measurements are satisfactory. Hence, it is evident that the temporal variability of these stations is insignificant and thus can be curtailed down to a number that is representative of measurements recorded at all the non-significant stations. The significant stations (represented in blue color) in Fig. 8 which show high temporal variability than non-significant stations (represented in red color) should be given greater importance for consistent and precise monitoring since these stations are crucial during the occurrence of any extreme event. This can help in the overall reduction of maintenance and operations costs by curtailing or

skeletonizing the existing network. On the other hand, if any discrepancies are found in the measurements of these observations for the non-significant stations, the cluster of these AWS may be periodically re-evaluated to re-design the AWS network. The pre- and post-evaluation of the AWS network through this novel framework can help in improved monitoring of the hydrometeorological observations at an optimal cost.

The proposed rationalized framework is further evaluated to assess whether the rationalized stations can be representative of all the stations in the area. The PDFs derived for spatially averaged (using the arithmetic averaging technique) daily rainfall and relative humidity data for all 35 stations and the top 22 stations according to TOPSIS are illustrated in Fig. 9. The KL-divergence values calculated by eq. 8 for rainfall and relative humidity plots are found to be 0.009 and 0.08 respectively, which reflects the fact that the PDFs for each parameter may be considered identical for the two cases to a greater extent. It can be concluded that the 22 stations are capable of capturing the pattern of both rainfall and relative humidity and are representative of the rainfall and relative humidity of the entire Mumbai according to measurements recorded at the stations. The correlation coefficients between spatially averaged data of the original AWS network and data observed at each of the 22 stations of the rationalized network are found to be more than 0.8 and 0.6 for rainfall and relative humidity respectively. Hence, the high correlation observed between the spatial average of both observations with that of the individual rationalized stations validates our proposed framework.

The performance of the rationalized AWS network is further evaluated at flood inundation and hazard levels. The 3-way coupled hydrodynamic MIKE FLOOD model which integrates both MIKE11 (1-D)

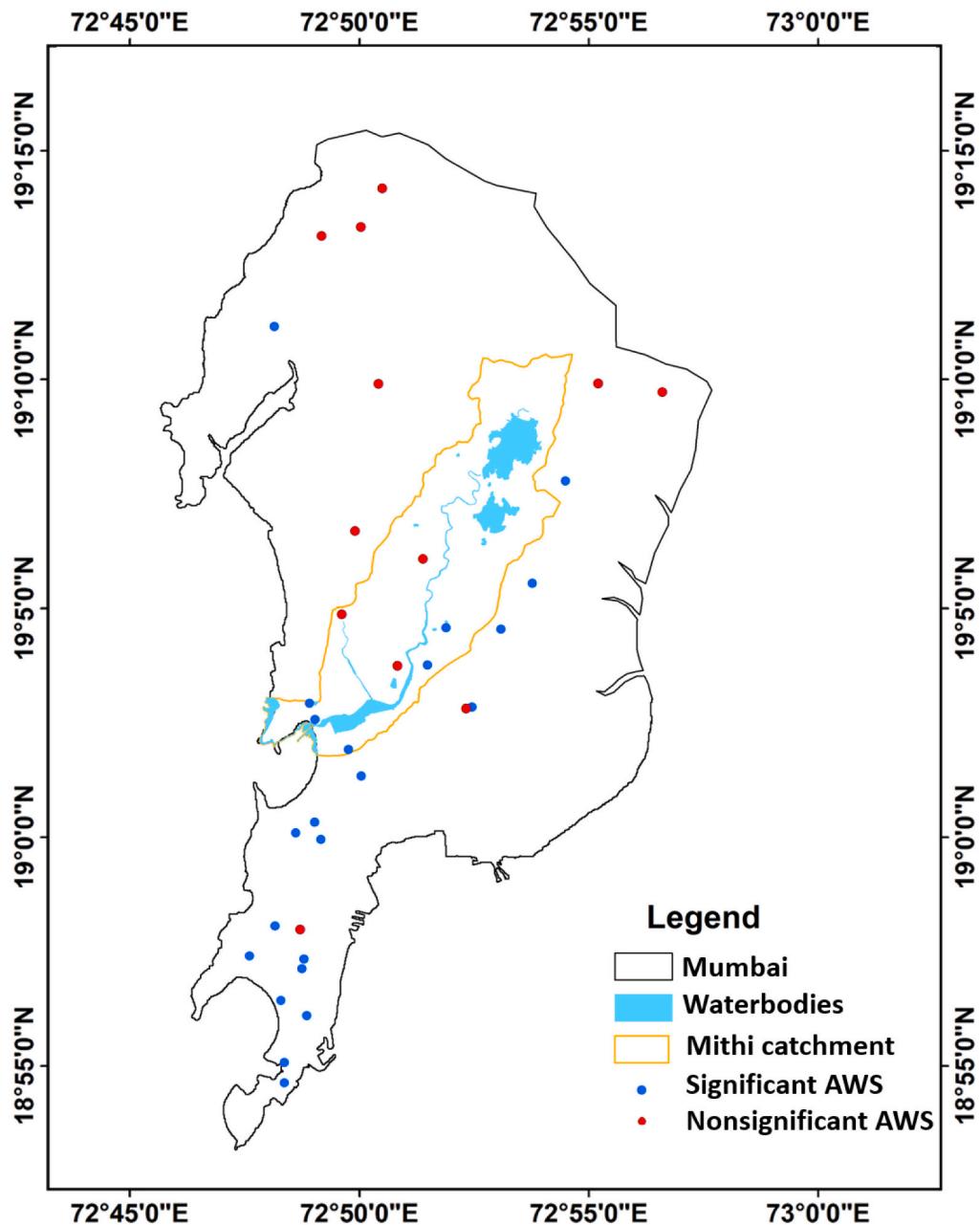


Fig. 8. Location of AWS stations ranked as per TOPSIS. Blue dots represent significant stations and Red dots represent non-significant stations in terms of capturing spatiotemporal variation of rainfall and relative humidity.

model with stormwater drainage network and MIKE 21 HD FM (2D) model is used in the present case study to derive flood inundation and hazard maps for a very heavy rainfall event of 18th–19th June 2015. It is observed that 8 out of 35 AWS considered in the study are located inside the Mithi river catchment which has been considered for flood modelling due to the availability of fine-resolution DEM for the area. The Mithi river catchment is delineated to perform experiments at inundation and hazard levels and two sets of runs are conducted: 1st set of run considers all 8 AWS present within the catchment; 2nd set of run considers 5 AWS (comes within the rationalized AWS network) present within the catchment. The flood inundation is represented in terms of flood depth and the hazard is quantified by taking into consideration both depth and velocity as per Australian Flood Hazard Classification as per eqs. 10 and 11 respectively (AIDR, 2017).

$$D_i = \begin{cases} I, 0 \leq d \leq 0.2 \\ II, 0.2 < d \leq 0.5 \\ III, 0.5 < d \leq 1 \\ IV, 1 < d \leq 1.2 \\ V, 1.2 < d \leq 1.5 \\ VI, 1.5 < d \leq 3 \\ VII, d > 3 \end{cases} \quad (10)$$

Where D_i represents the index of flood inundation for i^{th} class and d represents the value of depth in terms of 'm'.

$$H_i = \begin{cases} I, 0 \leq (d, v) \leq 0.3 \\ II, 0.3 < (d, v) \leq 0.6 \\ III, 0.6 < (d, v) \leq 0.8 \\ IV, 0.8 < (d, v) \leq 1.0 \\ V, 1.0 < (d, v) \leq 4.0 \\ VII, (d, v) > 4.0 \end{cases} \quad (11)$$

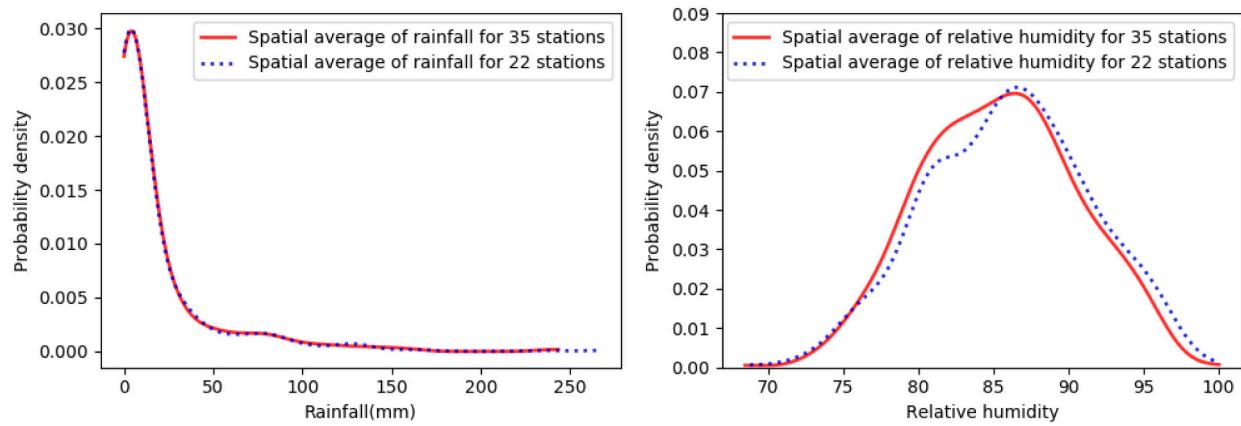


Fig. 9. Comparison of PDFs for spatially averaged (a) rainfall and (b) relative humidity for the entire AWS network and rationalized AWS network.

Where H_i represents the index of flood hazard for i^{th} class and (d, v) represents the product of depth and velocity in terms of ' m^2/s '. Figs. 10 and 11 represent the flood inundation and hazard maps derived for heavy rainfall event of 18-19th June 2015 respectively for all AWS and rationalized AWS obtained as per TOPSIS. Inundation (D) is categorized into low (I, II), medium (III, IV), and high (V, VI, VII) categories based on depth (d) only (eq. 10) and Hazard (H) is categorized into low (I, II), medium (III, IV), and high (V, VI) categories in terms of the product of both depth (d) and velocity (v) (eq. 11) based on the impact over human and economic livelihood. The number of cells and their percentage and spatial similarity under each category for two sets of rainfall data are illustrated in Tables 3 and 4 for flood inundation and hazard maps respectively. The spatial similarity is a statistical measure that shows the cell-to-cell concurrence in inundation and hazard maps for rainfall measured at stations within the rationalized network with respect to maps for rainfall measured at all the stations of the original AWS network within the catchment. It is observed that the information

captured by the skeletonized/ rationalized AWS network is in consensus with the original/existing network, which gets reflected at the inundation level as well as hazard level, i.e., number and percentage of cells under *low*, *medium*, and *high* categories for the existing AWS network are simulated as effectively as the rationalized network by the hydrodynamic flood modelling framework developed both in terms of inundation and hazard. The spatial similarity between the maps obtained using the rainfall measured with the rationalized network with respect to maps obtained with the original network is also very high. The spatial similarities for high and low categories are relatively greater than that of medium categories for both flood inundation and hazard maps.

This ensures effective performance by the rationalized AWS network for flood modelling. During flood inundation and hazard mapping, the *medium* and *high* flood inundation and hazard areas are discretely introspected and analyzed since they are more harmful from the viewpoint of human casualties and damages to livelihood, properties, economy, and cause greater inconvenience. The rainfall considered for

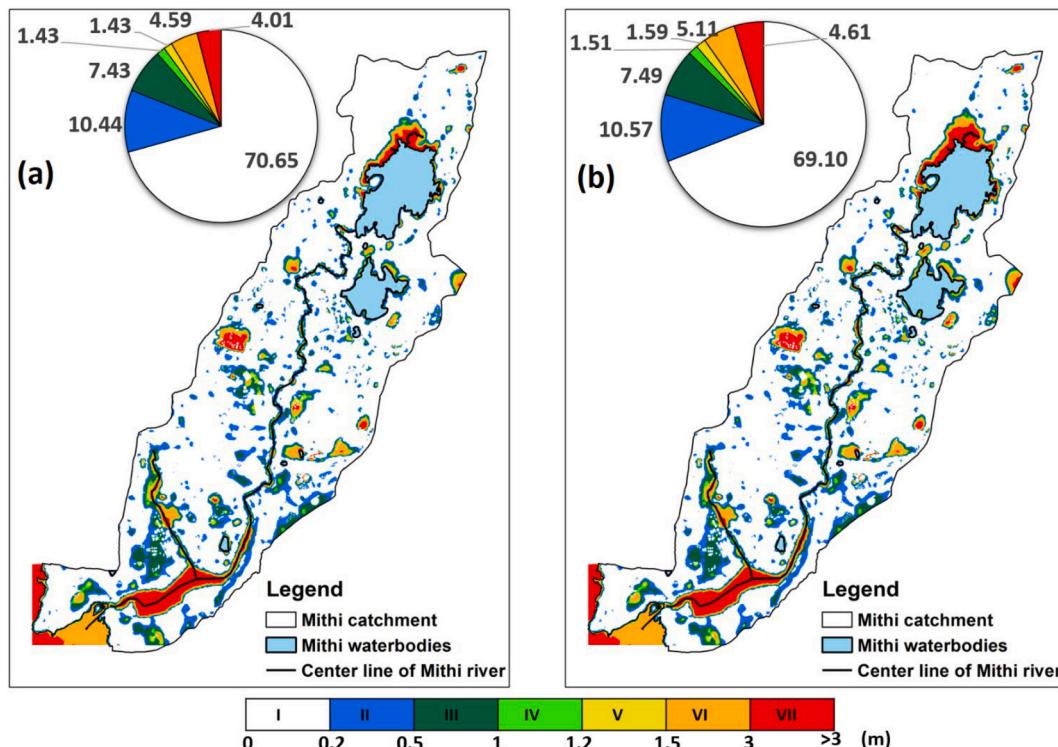


Fig. 10. Flood inundation maps for rainfall of (a) Total AWS network over the Mithi Catchment and (b) Rationalized AWS network over the Mithi Catchment in Mumbai; pie chart represents the % of area under each inundation class.

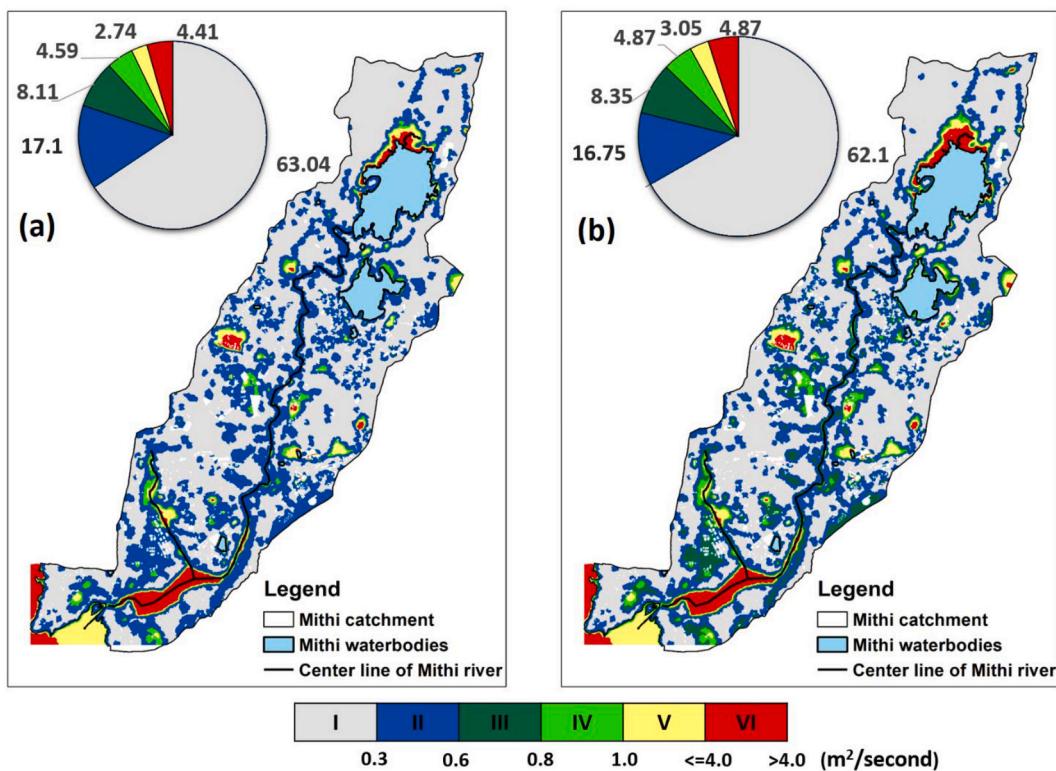


Fig. 11. Comparison between flood hazard maps for rainfall of (a) Total AWS network over the Mithi Catchment and (b) Rationalized AWS network over the Mithi Catchment in Mumbai; pie chart represents the % of area under each hazard class.

Table 3

Comparison of flood inundation levels for *high*, *medium*, and *low* categories in terms of number of cells for (a) Total AWS network and (b) Rationalized AWS network over the Mithi Catchment in Mumbai.

Description of Inundation	Number of cells (%)		Spatial similarity (%)
	All AWS stations within Mithi catchment (8 stations)	Rationalized stations within Mithi catchment (5 stations)	
High (VI, VII)	59,999 (8.61%)	60,713 (8.72%)	99.8
Medium (IV, V)	19,945 (2.86%)	21,608 (3.10%)	93.4
Low (I, II, III)	616,571 (88.52%)	614,194 (88.18%)	99.3

Table 4

Comparison of flood hazard levels for *high*, *medium*, and *low* categories in terms of number of cells for (a) Total AWS network and (b) Rationalized AWS network over the Mithi Catchment in Mumbai.

Description of Hazard	Number of cells (%)		Spatial similarity (%)
	All AWS stations within Mithi catchment (8 stations)	Rationalized stations within Mithi catchment (5 stations)	
High (V, VI)	49,845 (7.16%)	55,187 (7.92%)	99.7
Medium (III, IV)	88,453 (12.7%)	92,159 (13.23%)	92.1
Low (I, II)	558,217 (80.14%)	549,169 (78.85%)	98.6

rationalized stations gives conservative results, i.e., the number of *high* and *medium* cells are relatively greater than rainfall utilized for all stations, which is beneficial for demarcating the flood zones, and are of particular interest to the governing authorities and policymakers for better management and preparedness. Hence this analysis enables us to validate our proposed rationalization framework at rainfall and relative

humidity level (Fig. 9), flood inundation level (Fig. 10), and hazard level (Fig. 11).

The second stage of the study utilizes a multi-attribute decision-making technique to rank the stations attributing equal weightage to the criteria (hydrometeorological observations) considered, which necessitates a sensitivity analysis of this technique. Therefore, a rank reversibility analysis is conducted for the stations obtained from this approach by varying the weights assigned to the loading values from 0%, 25%, 40%, and 50% in both directions. For instance, the attribute rainfall has 0.5 as the initially assigned weight, TOPSIS is performed with 1000 values in each of [0.5 ± 25%], [0.5 ± 40%], and [0.5 ± 50%] intervals to evaluate the mean and standard deviation of the relative closeness value for all the stations with respect to the deterministic value, i.e., 0.5. Fig. 12 represents the mean and standard deviation values of relative closeness values due to variation in weights for the rationalized network consisting of 22 AWS. Subsequently, Fig. S4 represents the mean and standard deviation values of relative closeness due to variation in weights for the original AWS network. It is observed that the mean of relative closeness values for each of the top-ranked stations (rationalized AWS network) as well as of all the stations (original AWS network) remains nearly constant due to variation in weights. The standard deviation is also very less for the top-ranked stations and increases gradually with a decrease in relative closeness values. It is also observed that the standard deviation is less for the bottom-ranked stations which can be attributed to the lower temporal variability of these stations. This validates the robustness of the multi-attribute decision-making method, TOPSIS, i.e., the rank of the stations remains unaffected due to variation in weights. The proposed framework is generic and applicable in other climatic conditions and geographical locations, especially in developing and underdeveloped countries with similar social and economic setups where extensive hydro-meteorological monitoring is hindered by space and budgetary constraints and security issues. The existing networks can be evaluated with this framework while addressing the above-mentioned challenges and being cost-effective.

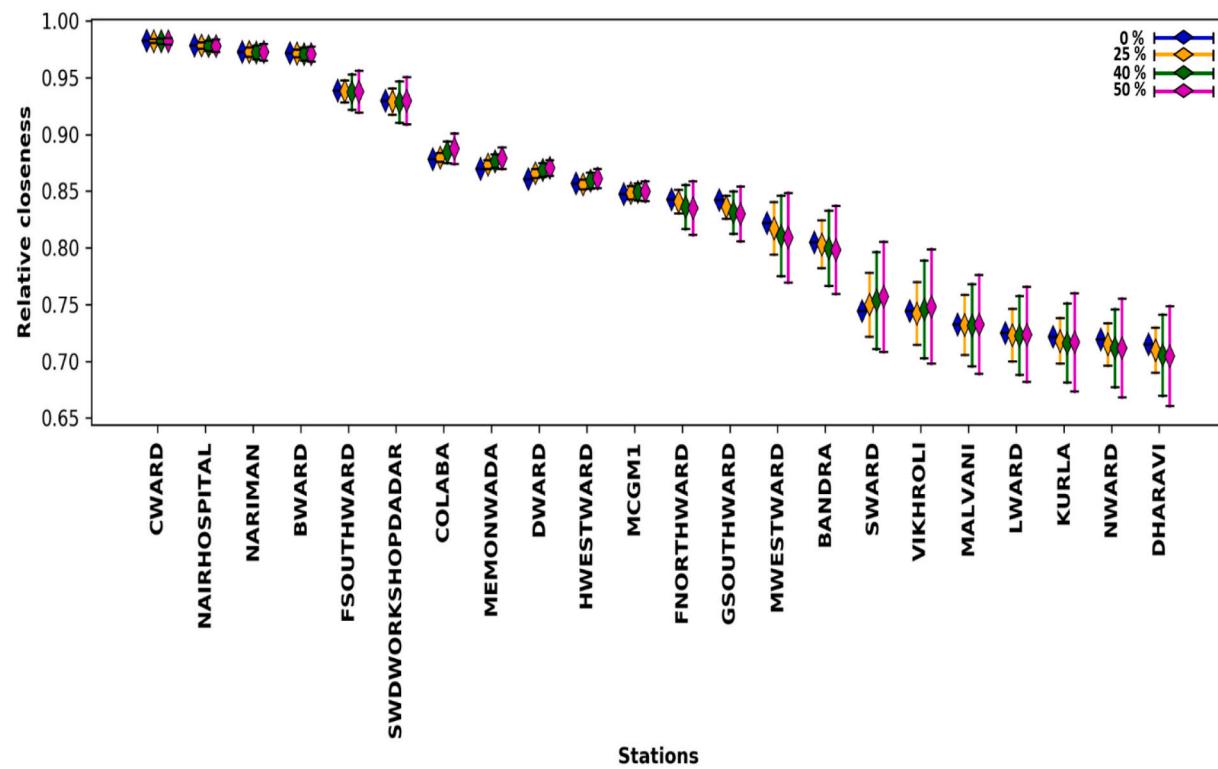


Fig. 12. Mean and standard deviation of closeness to the ideal solution for the top ranks of stations due to variation in weights.

5. Conclusion

Continuous and consistent availability of spatial data, best achieved through well-established dense weather monitoring networks or AWS, are essential for capturing the seemingly increasing spatiotemporal variability in hydrometeorological observations. Also, inappropriate placement of an existing monitoring station network may erroneously capture spatial characteristics of hydrometeorological observations, whereas extensively dense networks would raise their maintenance cost by several folds. Hence, to strike a balance between such limitations, an optimal rationalization of AWS networks needs to be accomplished. The current study presents a robust statistical framework to rationalize an AWS network. The proposed framework consists of a two-step approach followed by its comprehensive validation at a flood inundation level. In the first step, PCA is performed for each hydrometeorological observation individually at a particular AWS to determine the loading values for all stations. Following this, TOPSIS, a widely implemented MADM technique (Hwang and Yoon 1981), is employed on the loading values of each variable with equal weight by considering each station and observations as alternative and attribute, respectively. Lastly, the AWSs are ranked based on the decreasing order of the relative closeness values to the ideal solution. The proposed framework is demonstrated using sub-hourly precipitation and relative humidity information from the existing AWS network in Mumbai, India. Mumbai being a flood-prone city with a random and difficult-to-predict precipitation pattern (Singh et al. 2017), it was imperative for the city authorities to establish an exhaustive AWS network to capture the spatiotemporal characteristics of its precipitation. Post 2005 Mumbai flash floods (Bohra et al. 2006), MCGM took the initiative to tackle its flood-susceptibility problem and installed a dense network of about 60 AWS intending to collect consistent spatiotemporal information of several hydrometeorological observations. 35 out of the 60 stations with consistent and continuously available data points are considered for our analysis (Fig. 1). In our study, only 22 of 35 stations are found competent to represent the spatiotemporal pattern of precipitation and relative humidity of the entire city (Fig. 9). A comparison of

flood inundation and hazard maps derived from our network of rationalized stations and all stations present over the Mithi river catchment considered for flood modelling shows that our rationalized network can satisfactorily capture the catchment's flooding pattern (Fig. 10 and Fig. 11). Rank reversibility analysis of all AWSs further validates the robustness of the proposed framework (Fig. 12 and S2). Our study reveals the redundancy observed in AWS datasets and can help in the evaluation of the existing framework pre- and post-monsoon and also demonstrates that adopting a rationalized AWS network may help in the reduction of the maintenance cost. Although the proposed framework is elucidated using two observations as per data availability and consistency, it can be extended to rationalize a hydrometeorological network with multiple hydrometeorological observations and a greater number of stations. The framework is efficient at capturing the temporal variability for each of the stations considered irrespective of its elevation. This versatile framework can therefore be applied even for a set of stations with a wide range of elevation. Besides, the rank reversibility analysis shows that the rank of stations (in the context of capturing the temporal variability) remains unaltered due to variation in weights of different hydrometeorological observations. This further ascertains the credibility of the framework to be applied over areas with large differences in elevation.

Research studies on the evaluation and assessment of a hydrometeorological network involve multidisciplinary science, including hydrology, geology, morphology, topology, statistics, optimization, hydrometeorology, and geospatial technology. Also, the existing rationalization techniques are rigorous and mathematically intricate. Thus, realistically speaking, getting acquainted with such interdisciplinary perspectives behind the application of rationalization algorithms may probably be way outside the comfort zones of stakeholders who may thus resist implementing them. In addition, disaster authorities in developing/ underdeveloped countries may feel compelled to install AWS networks at potentially safe and accessible locations owing to reasons such as space constraints for stations' setup, security measures to prevent equipment vandalism or other unforeseen issues. It is to be

noted that the current study considers rationalization of an existing AWS network, rather than the design of a network over an ungauged catchment. The proposed statistical framework for AWS network rationalization can be deemed as an effective evaluation tool that can cater to the above-mentioned adversities and can be utilized on a timely basis to ensure the competence of any existing AWS network in capturing spatiotemporal characteristics of hydrometeorological observations over a study area.

Author contribution statement

SK and MG designed the problem. MG and JS performed the analysis with inputs from SK and SG. MG and JS prepared the figures with input from SK and SG. MG wrote the manuscript with inputs from SK, SG, JS and SS. SK and PEZ supplied data resources. All authors reviewed the manuscript.

Declaration of Competing Interest

None.

Acknowledgments

The work is supported by the Ministry of Earth Sciences (MoES), Govt. of India (Project No. 14MES003) and Department of Science & Technology (SPLICE-Climate Change Programme), Government of India (Project reference number DST/CCP/CoE/140/2018, Grant Number: 00000000000010013072 (UC ID: 18192442)). The authors would like to thank the MCGM for providing the shapefile and network cross-section details for the study area. The authors are also grateful to NRSC for providing the LULC maps, IMD Pune for rainfall data, and Indian National Centre for Ocean Information Services, Hyderabad for tidal elevation data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.atmosres.2021.105511>.

References

- Danish Hydraulic Institute (DHI), 2017. Mike 21Hydrodynamic Module, User Guide, MIKE by DHI, 2017.
- Australian Institute for Disaster Resilience (AIDR), 2017 Guideline 7–3, Supporting Document for Implementation of Australian Disaster Resilience Handbook 7 Managing the Floodplain: A Guide to Best Practice in Flood Risk Management in Australia.
- Abbott, M.B., Ionescu, F., 1967. On the numerical computation of nearly horizontal flows. *J. Hydraul. Res.* 5 (2), 97–117.
- Adhikary, S.K., Yilmaz, A.G., Muttill, N., 2015. Optimal design of rain gauge network in the Middle Yarra River catchment, Australia. *Hydrol. Process.* 29 (11), 2582–2599.
- Agarwal, A., Marwan, N., Rathinasamy, M., Ozturk, U., Merz, B., Kurths, J., 2018. Optimal design of hydrometric station networks based on complex network analysis. *Hydrol. Earth Syst. Sci. Discuss.* 1–21.
- Ahmad, L., Habib, Kanth R., Parvaze, S., Sheraz, Mahdi S., 2017. Automatic Weather Station: Experimental Agrometeorology: A Practical Manual. Springer (ISBN: 978-3-319-69184-8).
- Al-Zahrani, M., Husain, T., 1998. An algorithm for designing a precipitation network in the south-western region of Saudi Arabia. *J. Hydrol.* 205 (3), 205–216.
- Bohra, A.K., Basu, S., Rajagopal, E.N., Iyengar, G.R., Gupta, M.D., Ashrit, R., Athiyaman, B., 2006. Heavy rainfall episode over Mumbai on 26 July 2005: Assessment of NWP guidance; *Curr. Sci.* 90 (9), 1188–1194.
- Camera, C., Bruggeman, A., Hadjinicolaou, P., Pashiardis, S., Lange, M.A., 2014. Evaluation of interpolation techniques for the creation of gridded daily precipitation ($1 \times 1 \text{ km}^2$): Cyprus, 1980–2010. *J. Geophys. Res. Atmos.* 119 (2).
- Cecinati, F., Moreno-Rodenas, A.M., Rico-Ramirez, M.A., ten Veldhuis, M.-C., Langeveld, J.G., 2018. Considering rain gauge uncertainty using kriging for uncertain data. *Atmosphere* 9, 446.
- Chen, M., Shi, W., Xie, P., Silva, V.B.S., Kousky, V.E., Higgins, R.W., Janowiak, J.E., 2008. Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.-Atmos.* 113, D04110.
- Cover, T.M., Thomas, J.A., 1991. Elements of Information Theory. Wiley, New York.
- Dai, Q., Bray, M., Zhuo, L., Islam, T., Han, D., 2017. A scheme for rain gauge network design based on remotely sensed rainfall measurements. *J. Hydrometeorol.* 18, 363–379.
- Dyer, J.S., Fishburn, P.C., Steuer, R.E., Wallenius, J., Zonts, S., 1992. Multiple criteria decision making, multiattribute utility theory: the next ten years. *Manag. Sci.* 38 (5), 645–654.
- Ghosh, M., Mohanty, M.P., Kishore, P., Karmakar, S., 2020. Performance evaluation of potential inland flood management options through a three-way linked hydrodynamic modelling framework for a coastal urban watershed. *Hydrol. Res.* <https://doi.org/10.2166/nh.2020.123>.
- Hallegratte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. *Nat. Clim. Chang.* 3 (9), 802–806.
- Hao, Z., Singh, V.P., Hao, F., 2018. Compound extremes in Hydroclimatology: a review. *Water.* 10, 718.
- Helena, B., Pardo, R., Vega, M., Barrado, E., Fernández, J.M., Fernández, L., 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. *Water Res.* 34, 807–816.
- Huang, Y., Zhao, H., Jiang, Y., Lu, X., 2020. A Method for the Optimized Design of a Rain Gauge Network Combined with Satellite Remote Sensing Data. *Remote Sens.* 12, 194.
- Hwang, C.L., Yoon, K., 1981. Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag, New York. <https://doi.org/10.1007/978-3-642-48318-9>.
- IPCC, 2007. In: Parry, M.L., Canziani, O.F., Palutik, J.P., van der Linden, P.J., Hanson, C. E. (Eds.), Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ Press, Cambridge, UK.
- IPCC, 2013. Climate change 2013: the physical basis. In: Contribution of Working Group 1 to the Fifth Assessment Report of the IPCC. Cambridge University Press, New York.
- Ishizaka, A., Nemery, P., 2013. Multi-Criteria Decision Analysis: Methods and Software. John Wiley & Sons.
- Jackson, J.E., 1991. A User's Guide to Principal Components. Wiley, New York.
- Jenamani, R.K., Bhan, S.C., Kalsi, S.R., 2006. Observational/forecasting aspects of the meteorological event that caused a record highest rainfall in Mumbai. *Curr. Sci.* 90 (10), 1344–1362.
- Jewell, S.A., Gaussiat, N., 2015. An assessment of kriging-based rain-gauge–radar merging techniques. *Q. J. R. Meteorol. Soc.* 141 (2015), 2300–2313.
- Kadam, P., Sen, D., 2012. Flood inundation simulation in Ajoy River using MIKE-FLOOD. *ISH J. Hydraulic Eng.* 18, 129–141.
- Kalbar, P.P., Karmakar, S., Asolekar, S.R., 2012. Selection of an appropriate wastewater treatment technology: a scenario-based multiple-attribute decision-making approach. *J. Environ. Manag.* 113, 156–169.
- Kar, A.K., Lohani, A.K., Goel, N.K., Roy, G.P., 2015. Rain gauge network design for flood forecasting using multi-criteria decision analysis and clustering techniques in lower Mahanadi river basin, India. *J. Hydrol. Regional Stud.* 4, 313–332.
- Keum, J., Coulibaly, P., 2017. Information theory-based decision support system for integrated design of multivariable hydrometric networks. *Water Resour. Res.* 53 (7), 6239–6259.
- Kishtawal, C.M., Niyogi, D., Tewari, M., Pielke, R.A., Shepherd, J.M., 2010. Urbanization signature in the observed heavy rainfall climatology over India. *Int. J. Climatol.* 30 (13), 1908–1916.
- Kron, W., 2005. Flood Risk = Hazard • Values • Vulnerability. *Water Int.* 30, 58–68.
- Kullback, S., 1959. Information Theory and Statistics. John Wiley & Sons. (Republished by Dover Publications in 1968; reprinted in 1978).
- Lei, M., Niyogi, D., Kishtawal, C., Pielke Sr., R.A., Beltran Przekurat, A., Nobis, T.E., Vaidya, S.S., 2008. Effect of explicit urban land surface representation on the simulation of the 26 July 2005 heavy rain event over Mumbai, India. *Atmos. Chem. Phys.* 8 (20), 5975–5995.
- Li, Y., Grimaldi, S., Walker, J.P., Pauwels, V., 2016. Application of remote sensing data to constrain operational rainfall-driven flood forecasting: a review. *Remote Sens.* 8, 456.
- Lokanadham, L., Gupta, K., Nikam, V., 2009. The indian society for hydraulics journal of hydraulic engineering characterization of spatial and temporal distribution of monsoon rainfall over mumbai, 15(2).
- Manz, B., Buytaert, W., Zulkafli, Z., Lavado, W., Willems, B., Alberto Robles, L., Rodriguez-Sanchez, J.-P., 2016. High-resolution satellite-gauge merged precipitation climatologies of the Tropical Andes. *J. Geophys. Res. Atmos.* 121, 1190–1207.
- Matalas, N.C., Gilroy, E.J., 1968. Some comments on regionalization in hydrologic studies. *Water Resour. Res.* 4 (6), 1361–1369.
- MCGM, 2007. Greater Mumbai Disaster Management Action Plan 1.
- Meehl, G.A., Washington, W.M., 1993. South Asian summer monsoon variability in a model with doubled atmospheric carbon dioxide concentration. *Science* 260, 1101–1104.
- Mishra, A.K., Coulibaly, P., 2009. Developments in hydrometric network design: a review. *Rev. Geophys.* 47, 1–24.
- Mogheir, Y., Singh, V.P., de Lima, J.L.M.P., 2006. Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza strip, Palestine. *Hydrogeol. J.* 14, 700–712.
- Mohanty, M.P., Sherly, M.A., Karmakar, S., Ghosh, S., 2018. Regionalized design rainfall estimation: an appraisal of inundation mapping for flood management under data-scarce situations. *Water Resour. Manag.* 32 (14), 4725–4746.
- Mohanty, M.P., Vittal, H., Yadav, V., Ghosh, S., Rao, G.S., Karmakar, S., 2020. A new bivariate risk classifier for flood management considering hazard and socio-economic dimensions. *J. Environ. Manag.* 255, 109733.

- Morin, G., Fortin, J.P., Sochanska, W., Lardeau, J.P., Charbonneau, R., 1979. Use of principal component analysis to identify homogenous precipitation stations for optimal interpolation. *Water Resources Research*. 15 (6), 1841–1850.
- Moss, M.E., Tasker, G.D., 1991. An intercomparison of hydrological network design technologies. *Hydrol. Sci. J.* 36, 209–221.
- Musthafa, O.M., Karmakar, S., Harikumar, P.S., 2014. Assessment and rationalization of water quality monitoring network: a multivariate statistical approach to the Kabbini River (India). *Environ. Sci. Pollut. Res.* 21, 10045–10066.
- Nayak, M.A., Ghosh, S., 2013. Prediction of extreme rainfall event using weather pattern recognition and support vector machine classifier. *Theor. Appl. Climatol.* 114, 583–603.
- Nitu, R., Wong, K., 2010. CIMO survey on national summaries of methods and instruments for solid precipitation measurement at automatic weather stations. In: *WMO Instruments and Observing Methods Rep.* 102, WMO/TD-1544 (57 pp).
- Ouyang, Y., 2005. Evaluation of river water quality monitoring stations by principal component analysis. *Water Res.* 39, 2621–2635.
- Papaioannou, G., Loukas, A., Vasilides, L., Aronica, G.T., 2016. Flood inundation mapping sensitivity to riverine spatial resolution and modelling approach. *Nat. Hazards* 83 (1), 117–132.
- Patel, P., Ghosh, S., Kaginekar, A., Islam, S., Karmakar, S., 2019. Performance evaluation of WRF for extreme flood forecasts in a coastal urban environment. *Atmos. Res.* 223, 39–48.
- Patel, P., Karmakar, S., Ghosh, S., Niyogi, D., 2020. Improved simulation of very heavy rainfall events by incorporating WUDAPT urban land use/land cover in WRF. *Urban Clim.* 32, 100616.
- Paul, S., Ghosh, S., Mathew, M., Devanand, A., Karmakar, S., Niyogi, D., 2018. Increased spatial variability and intensification of extreme monsoon rainfall due to urbanization. *Sci. Rep.* 8, 3918.
- Putthividhya, A., Tanaka, K., 2012. Optimal rain gauge network design and spatial precipitation mapping based on geostatistical analysis from co-located elevation and humidity data. *Int. J. Environ. Sci. Dev.* 40 (2), 124–129.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell. System Tech. J.* 27 (379–423), 623–656.
- Shastri, H., Paul, S., Ghosh, S., Karmakar, S., 2015. Impacts of urbanization on Indian summer monsoon rainfall extremes. *J. Geophys. Res.-Atmos.* 120 (2), 496–516.
- Shastri, H., Ghosh, S., Karmakar, S., 2017. Improving global forecast system of extreme precipitation events with regional statistical model: Application of quantile-based probabilistic forecasts. *J. Geophys. Res. Atmos.* 122, 1617–1634.
- Sherly, M.A., Karmakar, S., Chan, T., Rau, C., 2015. Design rainfall framework using multivariate parametric nonparametric approach. *J. Hydrol. Eng.* 21, 4015049.
- Singh, K.P., Malik, A., Mohana, D., Sinha, S., 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti river (India)—a case study. *Water Res.* 38, 3980–3992.
- Singh, J., Sekharan, S., Karmakar, S., Ghosh, S., Zope, P.E., Eldho, T.I., 2017. Spatio-temporal analysis of sub-hourly rainfall over Mumbai, India: is statistical forecasting futile? *J. Earth Syst. Sci.* 126, 38.
- Skok, G., 2006. Analytical and practical examples of estimating the average nearest-neighbor distance in a rain gauge network. *Meteorol. Z.* 15, 565–573.
- St-Hilaire, A., Ouarda, T.B.M.J., Lachance, M., Bobee, B., Gaudet, J., Gignac, C., 2003. Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. *Hydrol. Process.* 17 (18), 3561–3580.
- Stosic, T., Stosic, B., Singh, V.P., 2017. Optimizing streamflow monitoring networks using joint permutation entropy. *J. Hydrol.* 552, 306–312.
- Tingsanchali, T., Karim, M.F., 2005. Flood hazard and risk analysis in the southwest region of Bangladesh. *Hydrol. Process.* 19, 2055–2069.
- Tiwari, S., Jha, S.K., Singh, A., 2020. Quantification of node importance in rain gauge network: influence of temporal resolution and rain gauge density. *Sci. Rep.* 10, 9761.
- Trenberth, K.E., 2005. The impact of climate change and variability on heavy precipitation, floods, and droughts. In: Anderson, M.G. (Ed.), *Encyclopedia of Hydrological Sciences*. Wiley, Hoboken, NJ, USA.
- Varekar, V., Karmakar, S., 2017. Rationalization of water quality monitoring network. In: *The Water-Food-Energy Nexus: Processes, Technologies, and Challenges*, 94. CRC Press, Taylor & Francis Group.
- Varekar, V., Karmakar, S., Jha, R., 2015. Design of sampling locations for river water quality monitoring considering seasonal variation of point and diffuse pollution loads. *Environ. Monit. Assess.* 187, 376.
- Varekar, V., Karmakar, S., Jha, R., 2016. Seasonal rationalization of river water quality sampling locations: a comparative study of the modified Sanders and multivariate statistical approaches. *Environ. Sci. Pollut. Res.* 23, 2308–2328.
- Varekar, V., Yadav, V., Karmakar, S., 2021. Rationalization of water quality monitoring locations under spatiotemporal heterogeneity of diffuse pollution using seasonal export coefficient. *J. Environ. Manag.* 277, 111342.
- Vittal, H., Karmakar, S., Ghosh, S., 2013. Diametric changes in trends and patterns of extreme rainfall over India from pre-1950 to post-1950. *Geophys. Res. Lett.* 40 (12), 3253–3258.
- Vittal, H., Singh, J., Kumar, P., Karmakar, S., 2015. A framework for multivariate data-based at-site flood frequency analysis: Essentiality of the conjugal application of parametric and nonparametric approaches. *J. Hydrol.* 525, 658–675.
- Vojinovic, Z., Tutulic, D., 2009. On the use of 1D and coupled 1D-2D modelling approaches for assessment of flood damage in urban areas. *Urban Water J.* 6 (3), 183–199.
- Wang, K., Guan, Q., Chen, N., Tong, D., Hu, C., Peng, Y., Dong, X., Yang, C., 2017. Optimizing the configuration of precipitation stations in a space-ground integrated sensor network based on spatial-temporal coverage maximization. *J. Hydrol.* 548, 625–640.
- Wunderlin, D.A., Diaz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C., Bistoni, M.A., 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba Argentina). *Water Res.* 35, 2881–2894.
- Xu, H., Xu, C.Y., Chen, H., Zhang, Z., Li, L., 2013. Assessing the influence of rain gauge network density and distribution on hydrological model performance in a humid region of China. *J. Hydrol.* 505, 1–12.
- Yadav, V., Karmakar, S., Kalbar, P.P., Dikshit, A.K., 2019. PyTOPS: a Python based tool for TOPSIS. *SoftwareX*. 9, 217–222.
- Zope, P.E., Eldho, T.I., Jothiprakash, V., 2015. Impacts of urbanization on flooding of a coastal urban catchment: a case study of Mumbai City, India. *Nat. Hazards* 75, 887–908.
- Zscheischler, J., Westra, S., van den Hurk, B.J.J.M., et al., 2018. Future climate risk from compound events. *Nat. Clim. Chang.* 8, 469–477.