Journal of
Water & Climate Change

Journal of Water and Climate Change Vol 13 No 8, 3004 doi: 10.2166/wcc.2022.302

A systematic quantitative review on the performance of some of the recent short-term rainfall forecasting techniques

Shejule Priya Ashok and Sreeja Pekkat*

Department of Civil Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
*Corresponding author. E-mail: sreeja@iitg.ac.in

SPA, 0000-0001-5895-2002; SP, 0000-0001-9166-5590

ABSTRACT

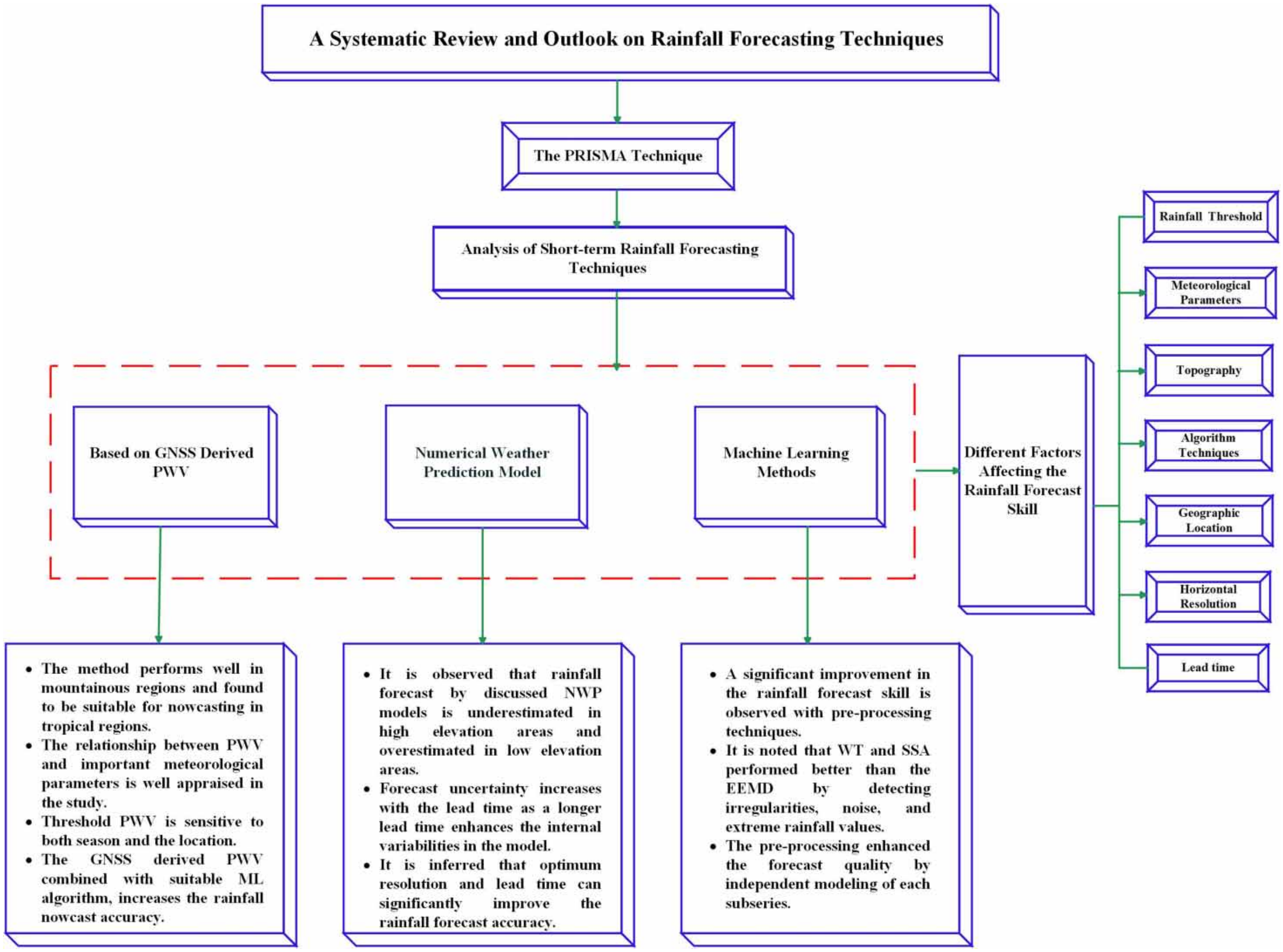
Rainfall forecasting is a high-priority research problem due to the complex interplay of multiple factors. Despite extensive studies, a systematic quantitative review of recent developments in rainfall forecasting is lacking in the literature. This study conducted a systematic quantitative review of statistical, numerical weather prediction (NWP) and machine learning (ML) techniques for rainfall forecasting. The review adopted the preferred reporting items for systematic reviews and meta-analyses (PRISMA) technique for screening keywords and abstracts, leading to 110 qualified papers from multiple databases. The impact of rainfall threshold, meteorological parameters, topography, algorithm techniques, geographic location, the horizontal resolution of the model, and lead time on rainfall forecast was examined. The review shows the importance of precipitable water vapor (PWV) along with other meteorological parameters for accurate nowcasting in coastal and mountainous regions. An increase in rainfall forecast uncertainty with an increase in the lead time makes the NWP model less popular for the short-term forecast. The pre-processing techniques increased the accuracy of ML techniques by considering extreme values and detecting the irregularly distributed multi-scale features of rainfall in space and time. Future research can focus on hybrid models with improved accuracy for nowcasting. The output from the hybrid model serves as input for the decision support system required for urban flood risk management.

Key words: forecast skill, lead time, pre-processing techniques, rainfall forecast, resolution, systematic review

HIGHLIGHTS

- The PRISMA method is applied for a systematic quantitative review.
- The Global Navigation Satellite System-derived precipitable water vapor (PWV) is found to be capable of analyzing the real-time profile of water vapor content.
- Forecast can be improved by considering additional meteorological parameters along with the PWV.
- A longer lead time in the NWP model enhances the forecast uncertainty.
- There is a significant improvement in the forecast by machine learning models after pre-processing.

GRAPHICAL ABSTRACT



1. INTRODUCTION

According to the IPCC (Intergovernmental Panel on Climate Change) reports, climate change is responsible for extreme weather events (Pachauri *et al.* 2014; IPCC 2021), which lead to floods and droughts around the world. On a global scale, flooding causes significant damage to human lives and properties (Balica *et al.* 2012). The recent Sixth Assessment Report (AR6) comprises three groups which are for the physical science basis, impacts, adaptation, vulnerability, and mitigation, with a greater focus on regional information, which can be utilized for climate risk assessments (IPCC 2021). With changing hydro-climatology, the frequency and severity of extreme rainfall events are likely to increase, leading to frequent flooding (Imhoff *et al.* 2020). While these natural disasters cannot be prevented, the resilience of society toward these events can be improved by adopting state-of-the-art management techniques. One of the methods is the early and timely forecast of extreme events providing a sufficient lead time for preparedness. A reliable mitigation and management measure can be adopted through a decision support system if flooding can be predicted well in advance by integrating flood modeling and rainfall forecasting. Recent technical advances have significantly enhanced the rainfall forecasting skill (Bauer *et al.* 2015; Bhomia *et al.* 2019), which has steadily evolved over the years. In this context, nowcasting is a useful tool for short-term rainfall forecasting in specific urban settings.

The development in technology and knowledge about the atmospheric processes is required for the advances in rainfall forecast techniques (Benjamin *et al.* 2019; Randall *et al.* 2019). Table 1 summarizes the evolution of rainfall forecasting techniques from the 1900 to date and the predominant methods adopted during different periods. There is a gradual development from statistical methods (blind of processes) in the pre-2000 age to more process-based revolutionary big data and internet of things (IoT) for refined rainfall forecasting in the present. This was possible only due to the advancement in the real-time

Table 1 | Evolution of rainfall forecasting techniques

Time period		Methods	Observational data
Pre-2000	1900–1940	• Predominance of statistical methods • Method of correlation coefficients • Frontal analysis • Multiple regression equations	• Surface observations • Graphs, maps for determining pressure contours
	1941–2000	• Correlated upper-level air/waves with the surface pressure pattern • During 1981–1990, focus was on long-range forecast (LRF) methods, i.e., stochastic methods • Power regression models • NWP models	• Aircraft used to collect real-time weather data • Satellite and radar observations became popular • Use of sea surface temperature, mean sea level pressure • Zonal and meridional wind data
Post-2000	2000–2017	• Progress in meteorology • Use of high-speed computational facility • Radar rainfall forecast • Advent of empirical/ML methods	• Ground-based/weather satellites • Deployment of extensive radar networks
Current research	2018–2021	• Focus on speedy and effective algorithms • Era of big data, IoT for rainfall forecast • Considered the dependence of rainfall on a number of atmospheric variables • Efforts to enhance forecast reliability	• Focus on collecting data of high temporal and spatial resolution • Advanced radar data • Wireless and fully automated system for data collection

measurements of weather data and high computational developments including supercomputers (Golding 2000; Kumar *et al.* 2017; Wiston & Mphale 2018). The role of weather satellites deployed by various nations and the advanced weather radar system for microclimate monitoring has gone a long way in the improvement of rainfall forecasting (Randall *et al.* 2019). The present-day research is mainly focused on very short-term nowcasting and very long-term decadal forecasting, an overlap of interannual variability and long-term climate change (Krishnan & Sugi 2003; Vareed *et al.* 2013; Sun *et al.* 2014; Liu *et al.* 2015; Foresti *et al.* 2016; Salvi *et al.* 2017; Choudhury *et al.* 2019; Smith *et al.* 2019; Imhoff *et al.* 2020). While nowcasting is mandatory for the timely prediction of flooding at the local scale, decadal forecasting is important for long-term planning considering the impact of climate variabilities.

The numerical weather prediction (NWP) models are becoming increasingly popular for short-term rainfall forecasting. The different short-term rainfall forecasting NWP models adopted by different countries are summarized in Table 2. The NWP models are classified into global models and mesoscale models. The global models are used for the medium-range forecast as they cannot run at high resolution. In contrast, mesoscale models are used for the short-range forecast. The mesoscale models require weather forecasts obtained from global models for initialization and adjusting the boundary conditions (Diagne *et al.* 2013; Cogan 2016; Ramírez & Vindel 2017). Global and mesoscale models can provide surface weather details and be efficiently used for climatic simulations at a given region of interest (Pu & Kalnay 2019). The global model has global coverage but less ability to resolve explicitly convective systems, while the regional (mesoscale) model with fine-grid spacing (a few kilometers) performed better in accurately analyzing convective-scale features (WMO 2017).

STEPS is a probabilistic precipitation forecast approach that combines an extrapolation nowcast and a NWP forecast (Bowler *et al.* 2006). The Flash Flood Guidance System (FFGS) in the USA is the hydro-meteorological modeling system that combines remote sensing of rainfall (radar/satellite) and NWP output to provide early measures on flash floods for the next 6 h (WMO 2017).

The quantitative rainfall nowcasting techniques used by different countries are based on radar echo-tracking and extrapolation to produce real-time forecasts. The more sophisticated methods involve blending multiple observation systems with the NWP output for an accurate nowcast (WMO 2017). Based on the extrapolation method, they are classified as ‘area trackers’ and ‘cell trackers’. The IMD uses a Short-range Warning of Intense Rainstorms in Localized System (SWIRLS-2) and a Warning Decision Support System (WDSS) for nowcasting (Roy *et al.* 2019). The SWIRLS uses an area-tracking extrapolation method for radar echoes, while the WDSS uses centroid tracking for nowcasting in the Indian region (Li & Lai 2004; Lakshmanan

Table 2 | Summary of NWP models adopted by meteorological centers in different countries

Center	Country	Model	Application
BOM	Australia	STEPS	Short-term rainfall forecast
KMA	Korea	MAPLE/KLAPS	Short-term rainfall forecast
ECMWF	Europe	SWIRLS	Short-term/nowcasting
IMD	India	WRF/GFS	Short-term rainfall forecast
HKO	Hong Kong	SWIRLS	Nowcasting
NOAA	USA	GFS	Short-term rainfall forecast
JMA	Japan	MSM	Very short-term forecast
UKMO	UK	NIMROD	Very short-term forecast
Meteo France	France	AROME	Short-term rainfall forecast
CPTEC	Brazil	BAM	Short-term rainfall forecast
SCMO	China	TITAN/TREC	Nowcasting
CMC	Canada	GEPS	Quantitative precipitation forecast

BOM, Bureau of Meteorology; KMA, Korean Meteorological Administration; ECMWF, European Centre for Medium-range Weather Forecasts; IMD, Indian Meteorological Department; HKO, Hong Kong Observatory; NOAA, National Oceanic and Atmospheric Administration; JMA, Japan Meteorological Administration; UKMO, United Kingdom's Meteorological Office; Meteo France, France Meteorological Service; CPTEC, Centro de Previsão do Tempo e Estudos Climáticos (Portuguese for Center for Weather Forecast and Climatic Studies); SCMO, Shanghai Central Meteorological Observatory; CMC, Canadian Meteorological Centre; STEPS, Short-Term Ensemble Prediction System; MAPLE, McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation; KLAPS, Korea Local Analysis and Prediction System; SWIRLS, short-range warning of intense rainstorms in localized systems; WRF, Weather Research and Forecasting; GFS, Global Forecasting System; MSM, Mesoscale Model; NIMROD, Nowcasting and Initialisation for Modelling Using Regional Observation Data System; AROME, Applications of Research to Operations at Mesoscale; BAM, Brazilian Global Atmospheric Model; TITAN, Thunderstorm Identification, Tracking, Analysis, and Nowcasting; TREC, tracking radar echoes by correlation vectors; GEPS, Global Ensemble Prediction System.

et al. 2007). The Korean Meteorological Administration (KMA) employed the McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation (MAPLE) and the Korea Local Analysis and Prediction System (KLAPS) for short-term rainfall forecast (Germann & Zawadzki 2002, 2004). This technique uses variational echo-tracking with semi-Lagrangian advection of radar reflectivity and correlation of the forecast with the observation. The Shanghai Central Meteorological Observatory (SCMO) uses Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) for storm forecasting and rainfall nowcasting (Dixon & Wiener 1993). The TITAN uses a cell-tracking method for echo movement tracking and extrapolation. The Hong Kong Observatory (HKO) has developed the SWIRLS model for rainfall nowcast operations, a robust method for tracking the existing rain system (Li & Lai 2004). The United Kingdom's Meteorological Office (UKMO) developed an automated NIMROD (Nowcasting and Initialisation for Modelling Using Regional Observation Data), which uses pixel-based linear extrapolation of radar echo for quantitative precipitation forecast.

Globally, high spatial and temporal resolution radar networks and satellite data are in use for weather forecasts (Liguori & Rico-Ramirez 2014; Heuvelink *et al.* 2020). Satellite-based information is used in mountainous regions and other areas where a limited number of rain gauge measurements are available (Duan & Bastiaanssen 2013; Zhu *et al.* 2020). The studies reported that rainfall forecast for the next 0–6 h can be obtained by integrating the NWP models and radar extrapolation techniques (Bowler *et al.* 2006; He *et al.* 2013; Liguori & Rico-Ramirez 2014; Wang *et al.* 2016; Chu *et al.* 2018; Shehu & Haberlandt 2021). It is explicit that several models have been developed in the recent past for short-term rainfall forecasting. Since rainfall depends on several factors and has complex nonlinear associations, the same model may yield different results in different regions. The main objective of this systematic quantitative review is to appraise such variabilities associated with some of the recent models used for short-term rainfall forecasting. The review summarizes the findings and factors affecting the forecasting accuracy by considering rainfall forecasting methods at multiple time scales. The different methods considered in this study are (1) statistical methods, (2) the NWP model, and (3) machine learning (ML).

2. METHODOLOGY

This study aims for a systematic review (SR) for mapping the findings related to short-term rainfall forecasting reported by previous researchers to arrive at a generalized conclusion. This SR is mainly performed to address the following research questions: (1) what are the different short-term rainfall forecasting techniques, (2) how different factors affect the forecast accuracy, and (3) how forecasting accuracy is improved. The methodology of SR performed in this study is explained by a flowchart as presented in Figure 1. The preferred reporting items for SRs and meta-analyses (PRISMA) method (Debray *et al.* 2017) is applied to screen the relevant studies for SR. The SR includes the following steps: (1) identification/formulation of research questions, (2) identification of keywords using Boolean operators such as AND and OR for search strategy, (3) extracting and screening relevant publications satisfying the inclusion criteria and excluding the others. The inclusion criteria considered in this SR are (a) case studies related to rainfall forecasting, (b) factors affecting the nowcast, short, and medium-range forecast, (c) forecast accuracy, (d) impact factor of journals divided into three categories (high impact factor >30, moderate impact factor 10–3, and low impact factor <3), and (e) relevant conference proceedings.

According to the PRISMA flowchart shown in Figure 1, 2,100 papers were identified from the electronic databases such as Elsevier, IEEE, Nature, Springer, Scopus, Taylor and Francis, Web of Science, and Wiley. The manuscripts written in the English language and available in the database are considered. The keywords selected for SR include short-term rainfall forecast, Global Navigation Satellite System (GNSS)-derived precipitable water vapor (PWV) techniques, pre-processing techniques, ML, and NWP. The review identifies efficient forecasting techniques based on region, model type/resolution/lead time, and pre-processing methods. During the initial article reading, duplicated papers were removed. After removing the duplicates, articles reduced to 1,200 numbers. Further screening was done based on the title review. After reading the abstract, 200 articles were selected. From these 200 papers, articles satisfying the given setting were selected after full manuscript reading. Initially, the articles were shortlisted mainly based on the keywords. Those articles satisfying the criteria presented by PRISMA were further studied. The articles were then categorized into short-, medium-, and long-range and nowcasting category on the basis of the lead time of prediction. There are a number of statistical techniques available, but this review focused on recent advancements such as GNSS-derived PWV. Hence, other conventional statistical techniques were excluded, and only selected papers were considered. After screening ML techniques with pre-processing, a further filter was applied, and those methods that were highly efficient in noise removal and completely interpreting the time series were selected. In the case of NWP models, studies showing the influence of location, lead time, and horizontal resolution were included, and the remaining were rejected. Finally, this process yielded 110 eligible articles for further analysis. The flowchart of article coding is summarized in Figure 2. The papers selected were categorized based on the methods used for forecast, which include (a) statistical, (b) NWP, and (c) ML techniques. To examine the variability among them,

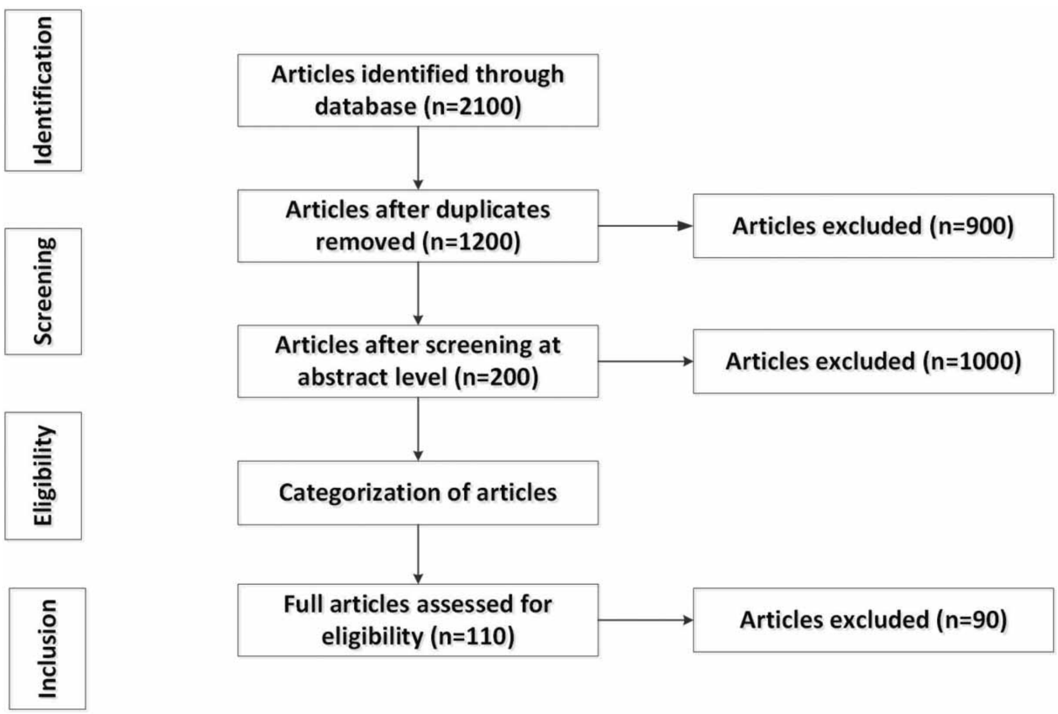


Figure 1 | Flowchart of PRISMA (preferred reporting items for SRs and meta-analyses) framework adopted for the SR.

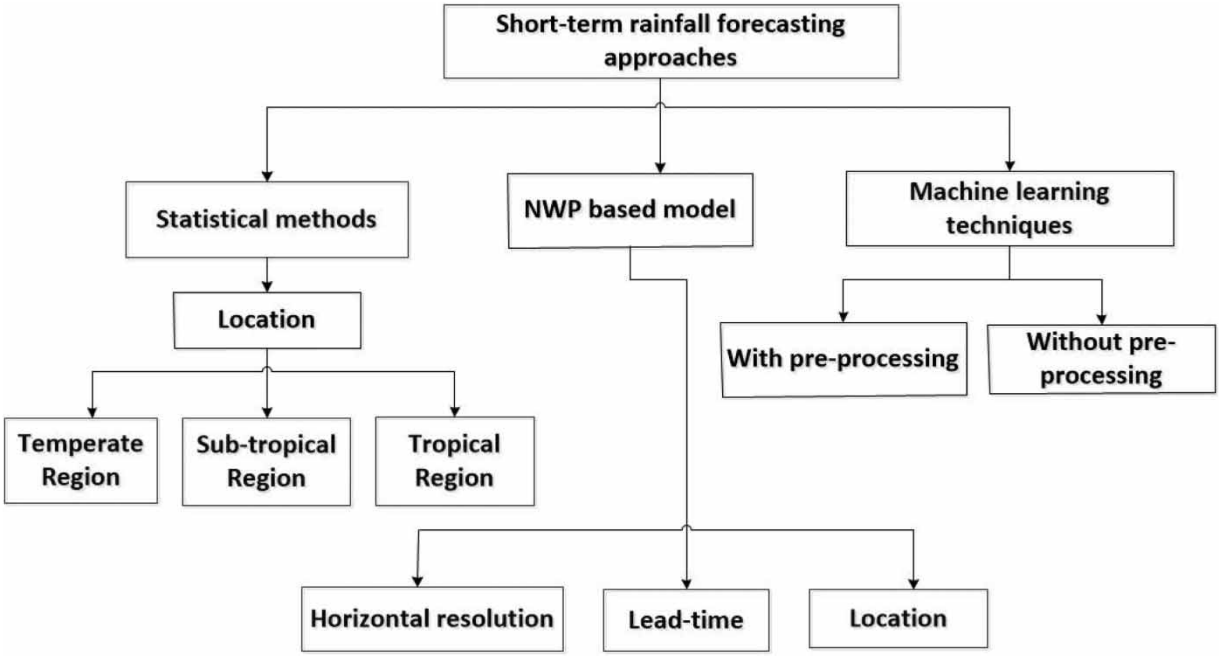


Figure 2 | Details of article coding for manuscripts selected for short-term rainfall forecasting.

(1) the statistical method was further classified based on location, i.e., temperate, subtropical, and tropical regions; (2) the NWP method was classified based on resolution, lead time, and location, and (3) ML methods were studied based on pre-processing techniques.

3. DESCRIPTIVE ANALYSIS OF SHORT-TERM RAINFALL FORECASTING TECHNIQUES

3.1. Short-term rainfall forecast based on GNSS-derived PWV

The IMD and other weather forecasting agencies have been using statistical methods for rainfall forecast for more than 100 years (Thapliyal 1982; Gowariker *et al.* 1989; Parthasarathy *et al.* 1993; Singh & Pai 1996; Thapliyal 1997; Rajeevan *et al.* 2000; Gadgil *et al.* 2005; Munot & Kumar 2007; Rajeevan *et al.* 2007; Kumar *et al.* 2012). The traditional statistical models are based on the long-term measurement of rainfall and its dependence on various meteorological parameters (Abbot & Marohasy 2018). Some of the statistical models like auto-regressive moving average (ARMA) and auto-regressive integrated moving average (ARIMA) methods were applied in the nonlinear hydrological process. However, its accuracy mainly depends on user knowledge and experience (Anh *et al.* 2019). These models mainly fall under the paradigm of stationarity. However, natural processes like rainfall are chaotic in nature. Despite so many efforts in the forecast by statistical methods, there is still enough scope for improving the forecast efficiency. This review focused on the more advanced statistical technique based on GNSS-derived PWV.

The GNSS-derived PWV is a convincing approach for rainfall analysis and forecasting and has wide applications in rainfall forecasting, global climate analysis (Yao *et al.* 2017), improving NWP (Gutman *et al.* 2003; Gendt *et al.* 2004; Guo *et al.* 2021). This forecasting technique comes under statistical methods for short-term rainfall forecast, which is further classified based on the location. The GNSS-derived PWV reflects the water vapor stored in a vertical air column above a certain area (Manandhar *et al.* 2018). The GNSS comprises a satellite constellation that transmits a radio signal through the atmosphere, which is received by the ground-based GNSS receiver. The atmospheric water vapor interferes with the propagated signal along the path, causing a delay referred to as 'tropospheric delay'. This tropospheric delay mainly accounts for zenith total delay (ZTD), from which the PWV can be obtained by a conversion factor (Shi *et al.* 2015; He *et al.* 2019; Li *et al.* 2020; Zhao *et al.* 2020). The PWV is the atmospheric water vapor expressed as the height of an equivalent column of liquid water (Manandhar *et al.* 2018). It is observed that in coastal and mountainous regions, the convection process arises due to rapid spatial variation in water vapor content. The GNSS is capable of analyzing the real-time distribution of water vapor content that can be used for precipitation nowcasting (Yuan *et al.* 2014). Therefore, PWV is considered to be a more reliable parameter for rainfall forecasts in mountainous areas (Kawase *et al.* 2006).

This review summarizes GNSS-derived PWV in the temperate, subtropical, and tropical regions for rainfall nowcast as shown in Tables 3–5, respectively. These tables represent how the true forecast rate (TFR) and the false forecast rate (FFR) change with the change in the precipitation threshold, lead time of the forecast, length of data, and algorithm applied for the rainfall forecast/nowcast. The TFR is defined as the ratio of the number of correctly forecasted rainfall events to the

Table 3 | GNSS-derived PWV for the rainfall nowcast in the temperate region

Author	Study region	Threshold	Lead time	Algorithm	Data	TFR (%)	FFR (%)
Benevides *et al.* (2015)	Temperate region of Lisbon Portugal	1.5 mm/h	1–6 h	Least-square fitting analysis	2010–2012 hourly data	75	65
Benevides *et al.* (2019)	Temperate region of Lisbon Portugal	0.5 mm/h	1 h	Nonlinear auto-regressive exogenous neural network model	2011–2015 hourly data	71.9	23.3
Łoś *et al.* (2020)	Central and northern Poland	–	0–2 h	Random forest classifier	2017–2019 hourly data	87	–

TFR, true forecast rate; FFR, false forecast rate.

Table 4 | GNSS-derived PWV for the rainfall nowcast in the subtropical region

Author	Study region	Threshold	Lead time	Algorithm	Data	TFR (%)	FFR (%)
Yao *et al.* (2017)	Zhejiang Province, China	0.6–0.8 mm/h	5.15 h	Precise Point Positioning (PPP) data-processing software	2014–2015 hourly data	82	66
Zhao *et al.* (2018)	Zhejiang Province, China	–	2–6 h	Least-square fitting time- series analysis	2014–2015 hourly data	90	60–66
Zhao *et al.* (2020)	Zhejiang Province, China	–	2–6 h	Precise Point Positioning (PPP) data-processing software	2014–2015 hourly data	>95	<30
Li *et al.* (2020)	Subtropical region of Hong Kong	1.1–1.7 mm/h	5.15 h	Pre-processing based on WMO, U.S. National Weather Service criteria	2010–2019 hourly data	95.5	28.9
Li *et al.* (2022)	Subtropical region of Hong Kong	Anomaly-based percentile thresholds	4.13 h	GNSS data acquisition and pre-processing	2010–2019 hourly data	97.6	13.4

Table 5 | GNSS-derived PWV for the rainfall nowcast in the tropical region

Author	Study region	Threshold	Lead time	Algorithm	Data	TFR (%)	FFR (%)
Manandhar *et al.* (2018)	Tropical region of Singapore	0.3–0.4 mm/h	5 min	–	2010–2013	87.7	38.6
Manandhar *et al.* (2019a)	Tropical region of Singapore	–	5 min	Data-driven ML technique (SVM)	2012–2015	80.4	20.3
Manandhar *et al.* (2019b)	Tropical region of Singapore	0.2–0.3 mm/h	45–60 min	–	2010–2016	79.62	50.38
Liu *et al.* (2019)	Tropical region of Singapore	0.1 mm/h	10–60 min	Improved BP-NN	2010–2012	>96	<40
Biswas *et al.* (2021)	Tropical region of Singapore and Brazil	0.7–0.9 mm	6 h	GPS-derived atmospheric gradient and residual	2010–2013 and 2016	87	36.6

SVM, support vector machine; BP-NN, back-propagation neural network.

actual number of rainfall events. The FFR represents the false alarm situations and is defined as the ratio of the number of falsely forecasted rainfall events (no rainfall occurred) to the actual number of rainfall events.

Researchers (Benevides *et al.* 2015; Yao *et al.* 2017; Zhao *et al.* 2020) compared longer duration series of PWV values with the rainfall series over the same period of time. They found that when the factors such as PWV, PWV variation, and the rate of change of PWV reached a particular value, the rainfall probability increased significantly. This particular value is known as the rainfall threshold. From Tables 3–5, it is observed that different threshold gives different forecast results. The TFR values obtained from the studies ranged between 71.9 and 96%, while the FFR values varied from 20.3 to 66.6%. This indicates that overall PWV values are a good indicator of rainfall occurrence. In temperate regions, rainfall forecast was obtained with a lead time of 1–6 h, in subtropical regions 1–5.15 h, and in tropical regions 5–60 min. For meteorological nowcasting suggested, the lead time is up to 30 min. From Tables 3–5, it is observed that GNSS-derived PWV can be successfully applied for nowcasting in tropical regions. Also, the forecast accuracy is found to be more in tropical and subtropical regions than in temperate regions (Li *et al.* 2020).

In all these studies, hourly data are used for analysis with the data length of 1 year, 2, 3, 6, and 9 years. It is observed that with an increase in data length, rainfall forecast probability beyond 2 days can be obtained (Seco *et al.* 2012; Benevides *et al.* 2015), thereby allowing time for preparedness during extreme events. In tropical regions, mostly convective rainfall occurs and has a duration of less than 30 min (Manandhar *et al.* 2019b). Therefore, hourly sample data are not good in this region. Different algorithms yield different forecast results. Rainfall depends on a number of parameters such as PWV, temperature, pressure, and relative humidity. The method, which can encompass most of these factors, gives an accurate TFR. A detailed explanation of the influence of PWV and different algorithms on rainfall forecast is explained below.

3.1.1. Relationship between PWV and rainfall

A positive relationship was observed between PWV and rainfall (Champollion *et al.* 2004; Bastin *et al.* 2007; Yan *et al.* 2009; Brenot *et al.* 2014; Zhao *et al.* 2018), with PWV increasing just before the rainfall event and decreasing after the rainfall (Yao *et al.* 2017; Barindelli *et al.* 2018). However, the rise in PWV will not always cause rainfall if other factors are not favorable (Sharifi *et al.* 2015). Certain external factors, such as thermodynamic variations, need to be considered as it affects the rainfall (Shoji 2013). Thermodynamic properties are mainly associated with temperature profile and moisture content determining the convection process (Pall *et al.* 2007; Lepore *et al.* 2015). Thermodynamic indices are important to understand the atmospheric instability, which ultimately triggers the rainfall (Ajilesh *et al.* 2020). Manandhar *et al.* (2019b) stated that for temperate and subtropical regions, an apparent change in PWV values is observed certain hours before the rain. In contrast, PWV values are very high for tropical regions and change marginally before the rain.

3.1.2. Relationship between threshold PWV and rainfall

Researchers (Jin & Luo 2009; Manandhar *et al.* 2018; Zhao *et al.* 2018) have studied the influence of the seasonal variation in PWV on the rainfall forecast. The seasonal variation in the PWV value was found to be more on a rainy day as compared to a non-rainy day (Manandhar *et al.* 2018). Therefore, it is understood that there is a threshold PWV beyond which rainfall can occur. According to Jin & Luo (2009), the threshold PWV value for rainfall forecast is the location and season-specific. Seasonal variations in PWV were observed over many GNSS stations. The authors noted that, during the intermonsoon season, PWV values are high as compared to the monsoon season. As the temperature in the intermonsoon is high, it can hold more water vapor, causing an increase in the PWV values. Zhao *et al.* (2018) also proved the same results (the PWV value is more in summer than in winter) based on the rainfall forecast experiment. Researchers concluded that the threshold PWV value is sensitive to the location (Jin & Luo 2009; Benevides *et al.* 2015; Yao *et al.* 2017; Manandhar *et al.* 2018, 2019b). Manandhar *et al.* (2018) stated that the threshold value also changes with the location. The PWV range in temperate, subtropical, and tropical regions is about 0–45, 0–80, and 30–70 mm, respectively. The PWV values in tropical regions are higher because of high temperature and relative humidity. Yao *et al.* (2017) noted that for the same threshold, at different stations, rainfall forecast results are found to be different. This shows that threshold values vary with the location. It is observed that if the calculated PWV value (evaluating criteria) exceeds the threshold value, then the probability of rainfall occurrence is high (Benevides *et al.* 2015). Researchers stated that the threshold of maximum PWV contributes to rainfall forecast in the tropics, whereas, in temperate and subtropical regions, the threshold of the maximum rate of the increment of PWV is the deciding factor (Benevides *et al.* 2015; Yao *et al.* 2017; Manandhar *et al.* 2019b).

3.1.3. Influence of different algorithms

Previous studies revealed that rainfall depends on a myriad of atmospheric parameters. Researchers implemented different ML algorithms by considering different parameters to yield an accurate forecast (Rahimi *et al.* 2018; Sangiorgio *et al.* 2018; Khaniani *et al.* 2021; Zhao *et al.* 2022). Figures 3 and 4 show that different algorithms applied at the same geographical station give different forecast rates. Manandhar *et al.* (2019a) improved their previous study by applying the support vector machine (SVM), an ML technique to classify rainy and non-rainy events. From Table 3, it is noted that forecast accuracy reached 87% when ML-based random forest (RF) was applied for storm nowcasting (Łoś *et al.* 2020). Employing SVM, it was noted that the FFR was reduced by 18.3%, as shown in Table 5. Liu *et al.* (2019) have applied an improved back-propagation neural network (BP-NN) algorithm for short-term rainfall forecast by incorporating more meteorological parameters such as temperature, relative humidity, dew point, and pressure. Interestingly, the TFR obtained was more than 96% when more factors were considered. Biswas *et al.* (2021) applied a new method considering significant weather features, horizontal tropospheric gradient, and atmospheric residual. The TFR was found to be 87% and the FAR reduced to 36.6% for the next 6 h prediction, as shown in Table 5. Therefore, it shows that it is necessary to include the relevant meteorological parameters and suitable algorithms for improving the TFR and reducing the FFR and evaluate the same for specific regions and seasons.

3.1.4. Effect of using multiple meteorological parameters on the rainfall forecast

Several studies have been carried out throughout the world considering the effect of meteorological parameters on rainfall forecast skills (Holloway & Neelin 2010; Seco *et al.* 2012; Chen & Li 2013; Sharifi *et al.* 2015; Suparta & Alhasa 2015; Priego *et al.* 2017; Yeh *et al.* 2018; Mawandha *et al.* 2019; Guo *et al.* 2021). A summary of meteorological parameters considered in the literature is listed in Table 6. Figure 5 shows the TFR and FFR values for the rainfall forecast in the temperate and subtropical regions. From Table 6, it is seen that Benevides *et al.* (2015) has considered PWV as a primary factor for the

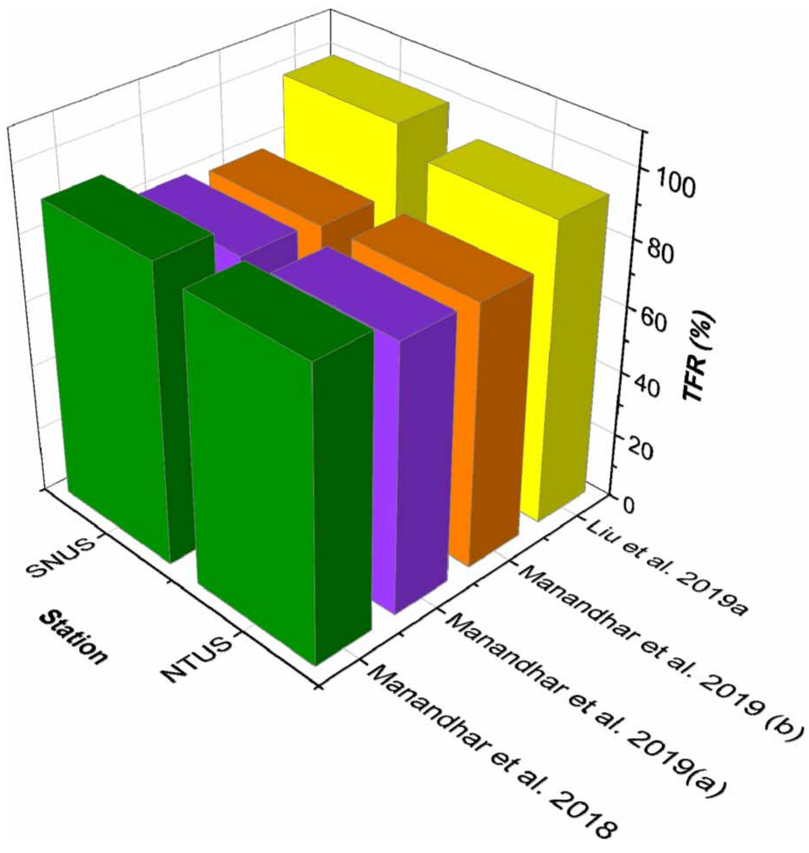


Figure 3 | Comparison of the true forecast rate affected by different algorithms for the same geographical location (the tropical region of Singapore).

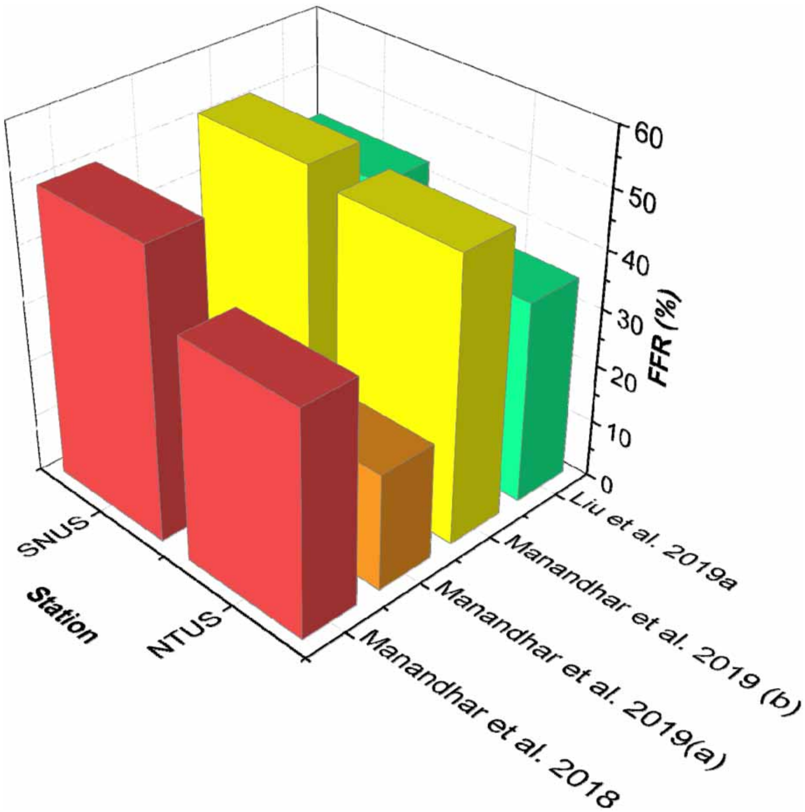


Figure 4 | Comparison of the false forecast rate affected by different algorithms for the same geographical location (the tropical region of Singapore).

rainfall forecast. Benevides *et al.* (2019) considered other meteorological parameters such as temperature, pressure, and relative humidity along with the PWV, which resulted in the reduction of the FFR by 5% as compared to the previous study.

A few studies carried out in the temperate region of Austria, Europe, and Italy include Karabatić *et al.* (2011), Guerova *et al.* (2016), and Barindelli *et al.* (2018). Karabatić *et al.* (2011) observed that prediction errors mainly occurred due to uneven topography in the alpine region of Austria. Therefore, the author has considered the effect of station height, latitude, and temperature gradient while calculating/extrapolating the pressure at the station. Guerova *et al.* (2016) studied the GNSS meteorology in Europe and presented state-of-the-art weather prediction, climate monitoring, and the assimilation of GNSS products into the NWP models. Barindelli *et al.* (2018) evaluated the relationship between PWV time variations and rainfall events. The author observed the peak in PWV when the rain clouds approached the station, followed by a decrease in the value by 5–10 mm when they moved past the station.

Yao *et al.* (2017) considered PWV variation, monthly PWV, and the rate of change of PWV values. In this case, the obtained TFR was 81% and the FFR was 66%. In the subtropical region, Li *et al.* (2020) considered two new predictors, PWV decrement and the rate of the PWV decrement, for the first time, along with the PWV values for short-term rainfall forecast/nowcast. With this addition, it was noted that TFR improved around 20% from 75 to 95.5% as compared to the studies in the temperate region, and the FFR significantly reduced to 28.9%. Zhao *et al.* (2018) considered the PWV value ZTD as the main parameter. The TFR was quite good with a value of 90%, but the FFR was around 65%, which was not an improvement. Therefore, to improve results, Zhao *et al.* (2020) considered additional parameters such as PWV first derivative, ZTD variation, and ZTD first derivative, and noted that FFR was significantly reduced to 30% with the same TFR of 90%. Li *et al.* (2022) used seven predictors, including hourly PWV and its six types of derivatives, showing the overall picture of PWV variation prior to the rainfall event. The TFR was found to be as high as 97.6%, and the FFR reduced to 13.4% for a prediction window of 4.13 h. This comparison plot is shown in Figure 5.

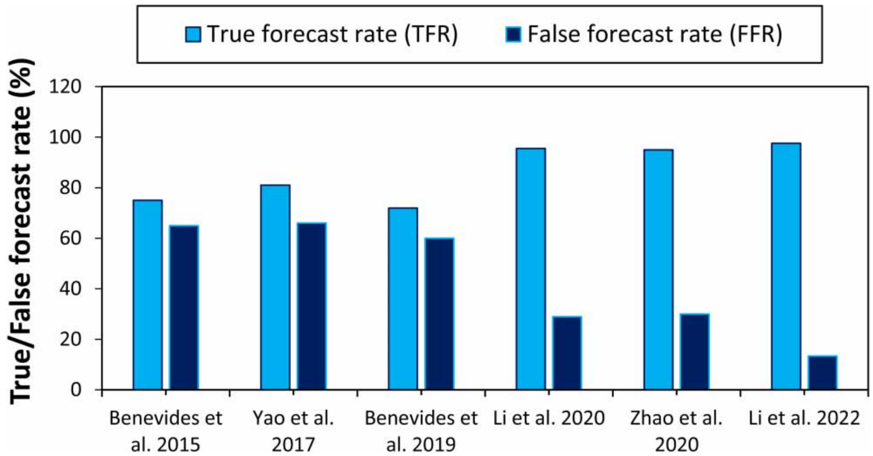


Figure 5 | Statistics (true/false forecast rate) related to rainfall forecasting for different studies.

From Table 6, it is seen that only one study has considered the impact of solar radiation on rainfall forecast. Very few researchers have considered the decrement of PWV and first-order and second-order derivatives of PWV values. However, the performance is marginally high for the three-factor model (three-meteorological parameters) and five-factor models (five meteorological parameters). Li *et al.* (2020) obtained a better forecast by using a combination of PWV increment, PWV increment rate, PWV decrement, PWV decrement rate, and PWV values. This is one of the reasons for the better performance of the method in the tropical and subtropical regions compared to the temperate region. It is evident from the above discussion that the PWV, when reinforced with other meteorological parameters, improved the TFR of short-term rainfall forecasts. Further studies are needed for optimizing and identifying the meteorological parameters required for accurate nowcasting.

3.2. NWP models

The NWP model represents the numerical simulation of atmospheric conditions to forecast the evolution of weather. A high-resolution model is required for the detailed representation and understanding of the atmospheric condition. The chaotic nature of the weather significantly affects the NWP forecast. A summary of rainfall forecast by popular NWP models is listed in Table 7. The different NWP models presented in Table 7 include the fifth-generation mesoscale model (MM5), Australian Community Climate Earth-System Simulator (ACCESS), Global/Regional Integrated Model system (GRIMs), WRF, Advanced Research WRF (ARW), GFS, the UKMO, Advanced Regional Prediction System (ARPS), and the Unified Model (UM). Table 7 describes various studies based on different factors such as the location of the study, different models used for short-term rainfall forecast, resolution of the model, and the forecast lead time. The following section explains the factors that influence the rainfall forecast by NWP models.

3.2.1. Influence of horizontal resolution on the rainfall forecast accuracy of NWP models

The NWPs are mostly grid models, in which the horizontal resolution is defined as the spacing between the grid cells. For other models with a global domain, such as spectral models, it is related to the number of waves that can be resolved by the model (Giunta *et al.* 2019). It was observed that the accuracy of the forecast can be enhanced with finer grid cells (Subramanian & Gopalakrishnan 2020). The short-term rainfall forecast skill depends on the location, season, and model resolution (Mass *et al.* 2002; Das *et al.* 2008). Researchers have performed sensitivity studies to understand the impact of horizontal resolution on NWP model forecast (Martin 1998; Gallus 1999; Goswami *et al.* 2012; Jang & Hong 2014; Li *et al.* 2016; Wang *et al.* 2016). To assess the performance of different methods, the statistical measures, including the probability of detection (POD), bias score (BS), false alarm ratio (FAR), critical success index (CSI), and equitable threat score (ETS), were used and are listed in Table 8.

Table 9 shows the effect of horizontal resolution on the rainfall forecast skill. Mass *et al.* (2002) applied the MM5 model and found that rainfall is underestimated at 12 km resolution as the coarser-resolution model does not identify small-scale

Table 6 | List of meteorological parameters considered in the literature

	Study									
Predictor	Benevides *et al.* (2015)	Yao *et al.* (2017)	Manandhar *et al.* (2018)	Zhao *et al.* (2018)	Benevides *et al.* (2019)	Liu *et al.* (2019)	Manandhar *et al.* (2019a)	Manandhar *et al.* (2019b)	Li *et al.* (2020)	Zhao *et al.* (2020)
PWV value	×	×	√	√	√	√	√	×	√	√
PWV variation/ PWV increment	√	√	×	×	×	×	×	√	√	√
PWV increment rate/first derivative	×	√	×	×	×	×	×	√	√	√
ZTD variation	×	×	×	×	×	×	×	×	×	√
ZTD first derivative	×	×	×	√	×	×	×	×	×	√
PWV second derivative	×	×	√	×	×	×	×	×	×	×
Monthly PWV	×	√	×	×	×	×	×	√	×	×
Seasonal PWV	×	×	√	×	×	×	×	×	×	×
Pressure	×	×	×	×	√	√	×	×	×	×
Temperature	×	×	×	×	√	√	√	×	×	×
Relative humidity	×	×	×	×	√	√	√	×	×	×
Dew temperature	×	×	×	×	×	√	√	×	×	×
PWV decrement	×	×	×	×	×	×	×	×	√	×
PWV decrement rate	×	×	×	×	×	×	×	×	√	×
Solar radiation	×	×	×	×	×	×	√	×	×	×

PWV, precipitable water vapor; ZTD, zenith total delay.

features. However, increasing the resolution to 4 km over predicts the rainfall, thereby reducing the overall forecast skill. Jang & Hong (2014) applied the GRIMs with different horizontal resolutions of 25, 50, and 100 km for the quantitative forecast of heavy rainfall events over the Korean peninsula. It was noted that with enhanced resolution, complex topography can be well represented with an improved ETS.

According to Kumar *et al.* (2016), increasing resolution improves rainfall forecasting skills. The POD increases by 40% when the resolution decreases from 45 to 5 km, and the BS increases by 24.3%. This indicated that better topographic representation positively impacts rainfall forecasting in the mountainous region. Sridevi *et al.* (2018) found that a high-resolution model has less overestimation with a BS of 1–1.25, while a low-resolution model has a significant overestimation with a BS of 1.5–2, implying that higher resolution improves forecast performance. The same results were observed by other studies reported in the literature (Li *et al.* 2016; Sharma *et al.* 2017, 2021). Li *et al.* (2016) observed that at high resolution (5 km), the BS was proportional to the rainfall threshold.

It is concluded that the standard verification metrics (POD, ETS, CSI, and BS) increase with an increase in the horizontal resolution with a sufficient reduction in the FAR. The accuracy improvement was mainly due to the accurate prediction of moisture and temperature with increased horizontal resolution. The above discussion noted that coarser horizontal resolution may not accurately represent land surface characteristics and topography, ultimately influencing the rainfall forecast accuracy. The finer resolution provided better orographic and mesoscale features and can be considered as an essential step toward accurate short-term rainfall forecasts. However, the model skill does not necessarily relate to increased horizontal resolution (Wang *et al.* 2004). The computational cost increases with the resolution (Mass *et al.* 2002) and also the increased resolution comes at the cost of substantial computational effort.

Table 7 | Summary of literature with different NWP models reviewed in the present study

Reference	Coverage	Model	Resolution	Lead time
Mass *et al.* (2002)	Washington	MM5	4/12/36 km	24 h
Goswami *et al.* (2012)	India	MM5	10/30/90/60 km	24 h
Shrestha *et al.* (2012)	Southeast Australia	ACCESS	80 km/37.5 km/12 km/5 km	1–24 h
Jang & Hong (2014)	Korea	GRIM	25/50/100 km	24 h
Kumar *et al.* (2016)	India	WRF	5/15/45 km	24 h
Li *et al.* (2016)	China	WRF	5/10/15/20/30/45 km	24/48/72/96 h
Prakash *et al.* (2016)	India	GFS UKMO	22/17 km	Days 1–5
Shahrban *et al.* (2016)	Southeast Australia	ACCESS	12 km	13–24 h
Wang *et al.* (2016)	China	ARPS	3 km	1–6 h
Jee & Kim (2017)	Korea	WRF	5 km – outer 1 km – inner	18/12/6 h
Moya-Álvarez *et al.* (2018)	Peru	WRF	0.75/3/6/18	10 days
Chu *et al.* (2018)	China	WRF-ARW	3 km	1–6 h
Jabbari *et al.* (2020)	Korea	WRF	1/2/4/8/12/16/20 km	12/24/36/48/60/72 h
Sridevi *et al.* (2018)	India	GFS	25 km	Days 1–5
Zhou *et al.* (2018)	China	WRF	10 km	36 h
Bhomia *et al.* (2019)	India	WRF	$0.25° \times 0.25°$	24/48 h
Sharma *et al.* (2021)	India	UM	10/40 km	24/48/72 h

MM5, fifth-generation mesoscale model; ACCESS, Australian Community Climate Earth-System Simulator; GRIMs, Global/Regional Integrated Model system; WRF, Weather Research and Forecasting; ARW, Advanced Research WRF; GFS, Global Forecast System; UKMO, the UK Met Office Unified Model; ARPS, Advanced Regional Prediction System; UM, Unified Model.

Table 8 | Description of statistical measures appearing in this review for assessing model performance

Acronym	Full form	Description
POD	Probability of Detection	Fraction of events correctly forecasted
ETS	Equitable Threat Score	Account for the hits that would occur purely due to random chance
CSI	Critical Success Index	Fraction of all correctly diagnosed observed and forecast events (excluding false and missed alarms)
FAR	False Alarm Ratio	Fraction of events that were actually non-events
BS	Bias Score	The ratio of predicted to observed rain
RMSE	Root Mean Square Error	Average error magnitude
TFR	True Forecast Rate	The ratio of the number of correctly forecasted rainfall events to the actual number of rainfall events
FFR	False Forecast Rate	The ratio of the number of forecasted rainfall events but no rainfall actually occurred to the actual number of rainfall events
r	Correlation Coefficient	Measures the degree of a linear relationship between observed and forecasted data
R^2	Coefficient of Determination	How well the model represents the data
MAE	Mean Absolute Error	Used for error characterization of a model
NSE	Nash–Sutcliffe Efficiency	To measure the model performance
CE	Coefficient of Efficiency	To check how well the observed and forecasted value fits

Table 9 | Effect of horizontal resolution on rainfall forecast skill

Reference	Horizontal resolution change (km)		Improvement			
			POD/ETS (%)	CSI (%)	BS (%)	FAR reduction (%)
Mass *et al.* (2002)	12	4	–	–	35	–
Jang & Hong (2014)	100	50	13	–	–	–
Kumar *et al.* (2016)	45	5	40	–	24.3	–
Li *et al.* (2016)	45	20	12	9.0	3.0	13.6
Sharma *et al.* (2017)	40	17	29	–	–	24
Sridevi *et al.* (2018)	25	12	–	–	20	–
Sharma *et al.* (2021)	45	10	10	9.0	–	10

POD, probability of detection; ETS, equitable threat score; CSI, critical success index; FAR, false alarm ratio; BS, bias score.

3.2.2. Influence of lead time on rainfall forecast accuracy

The lead time is defined as the length of time between the forecast issuance of the event and the occurrence of the forecasted event (Subramanian & Gopalakrishnan 2020). For many flood warning/forecasting studies, a lead time of 6–48 h is considered optimal (Herath *et al.* 2016). However, the forecast uncertainty increases with the lead time due to limited knowledge about the complex atmospheric processes. Therefore, the research efforts are directed toward increasing the lead time along with forecast accuracy, so that real-time measures can be taken against extreme events. The effect of lead time on rainfall forecasting skills was studied based on theoretical and modeling studies (Das *et al.* 2008; Shrestha *et al.* 2012; Jang & Hong 2014; Wang *et al.* 2016; Jee & Kim 2017; Chu *et al.* 2018; Sridevi *et al.* 2018; Bhomia *et al.* 2019; Jabbari *et al.* 2020; Sharma *et al.* 2021). Das *et al.* (2008) applied different mesoscale models, namely the MM5, Regional Spectral Model (RSM), the Eta model, and the WRF model over India, to check the forecast skill. It is noted that the mesoscale models performed better for the 1-day forecast, and the model's skill decreases with an increase in the lead time. The results of the performance score are summarized in Table 10. Shrestha *et al.* (2012) found that POD, which suggests the ability of the model to correctly diagnose the event, decreases with the lead time (POD reduces from 70 to 30%). The FAR increases, and the statistical measurement of the mean difference given by the BS fluctuates as the lead time increases. According to Jang & Hong (2014), the root mean square error (RMSE) increases, and bias decreases as the lead time increases by 24 h. The decrease in forecasting skills with an increase in the lead time was also endorsed by other literature (Chu *et al.* 2018; Sridevi *et al.* 2018; Jabbari *et al.* 2020).

Results reported in the literature (Shrestha *et al.* 2012; Jee & Kim 2017; Bhomia *et al.* 2019; Sharma *et al.* 2021) clearly show that FAR increased with the lead time, while CSI and POD decreased with the lead time. A higher lead time causes large internal variabilities (i.e., chaotic variabilities of the climate caused by the model itself), resulting in high uncertainty in the forecast skill (Lafaysse *et al.* 2014). Therefore, it is crucial to determine the optimum lead time for NWP models before employing it for rainfall forecasting.

3.2.3. Influence of location

The location and its topography play an important role in the short-term rainfall forecast efficiency. Shrestha *et al.* (2012) showed that the NWP models, such as Australian Community Climate Earth-System Simulator-VICTAS (ACCESS-VT) and Australian Community Climate Earth-System Simulator-Australia (ACCESS-A), overestimated rainfall up to 60% in low elevation and underestimated rainfall up to 30% in high elevation. Shahrban *et al.* (2016) also proved similar results with the ACCESS-A model showing the overestimation of rainfall in low precipitation areas and underestimation in high rainfall areas. Kumar *et al.* (2016) noted that with an increase in resolution, better topographic representations over hilly areas were obtained, improving the forecasting efficiency over the mountainous regions compared to plain areas. Bhomia *et al.* (2019) found that a high correlation between IMD observed (ground observation) and WRF forecasted results over central India with less bias and standard deviation (SD), indicating a better forecast efficiency. The Western Ghats (WG) and the north-east (NE) region of India receive the highest amount of monsoon rainfall during the month of June, July, August, and September (JJAS). Over the WG region (72 °E–76 °E, 13 °N–21 °N), the correlation between the IMD observed rainfall and WRF forecasted rainfall was good, but the increased BS indicated overestimation. High SD was found over the NE region

Table 10 | Studies highlighting the effect of lead time on NWP model performance

			Improvement (%)				
Study	**Lead time**		**POD/RS**	**CSI/ETS**	**Bias**	**FAR/RMSE**	**POD/RS**
Das *et al.* (2008)	Day 1	Day 3	MM5-Western India	–	ETS = 44.5	–	–
			ETA-Western India	–	25.5	–	–
			MM5-Eastern India	–	39	–	–
			ETA-Eastern India	–	10	–	–
Shrestha *et al.* (2012)	Short	Long	30	33	–	31	
Jang & Hong (2014)	24 h	48 h	–	–	–	RMSE = 13.67	
Wang *et al.* (2016)	1 h	6 h	–	11	11.55	–	
Jee & Kim (2017)	6 h	18 h	4	ETS-2	–	3	
Chu *et al.* (2018)	1 h	6 h	–	9	–	–	
Jabbari *et al.* (2020)	0–12 h	61–72 h	Event-2001	–	–	28	–
			Event-2007	–	–	11	–
			Event-2011	–	–	24	–
Sridevi *et al.* (2018)	Day 1	Day 5	12.5 km	RS = 2	–	–	–
			25 km	RS = 1	–	–	–
Bhomia *et al.* (2019)	24 h	48 h		–	7	4	5
Sharma *et al.* (2021)	Day 1	Day 3	11	7.5	–	11	

POD, probability of detection; ETS, equitable threat score; CSI, critical success index; FAR, false alarm ratio; RMSE, root mean squared error; RS, ratio score.

(90 °E–98 °E, 22 °N–30 °N), showing low forecast efficiency. This suggests the lower accuracy of WRF over high terrain areas. The BS was found to be less over the NW and SE regions of India, suggesting the underestimation of rainfall. Sridevi *et al.* (2018) found that the rainfall is overestimated by 1–3 mm by the GFS model compared to observed rainfall in different parts of India. It is apparent that more studies are needed to clearly draw the bounds within which NWP models work efficiently with respect to resolution, lead time, and location.

3.3. ML methods

The above discussion clearly highlights the need for alternate methods of short-term rainfall forecasting that are versatile and, at the same time, provide high efficiency. In the recent past, researchers have attempted to resolve some of the drawbacks of statistical and physically based models by using the capabilities of ML approaches or hybrid models (Hong 2008; Sumi *et al.* 2012; Cramer *et al.* 2017; Balamurugan & Manojkumar 2021; Ridwan *et al.* 2021). Artificial neural network (ANN), k-nearest neighbor (KNN), SVM, decision tree (DT), and RF are some of the popular ML models that were employed to handle complex nonlinear association (Hong & Pai 2007; Hong 2008; Sumi *et al.* 2012; Akrami *et al.* 2014; Cramer *et al.* 2017; Abbot & Marohasy 2018). These models are entirely based on the relationship among input and output variables and do not require knowledge of the underlying physical process (Solomatine & Ostfeld 2008). The ML models were applied for short-term as well as long-term rainfall forecasts.

3.3.1. Pre-processing of ML techniques

It was noted that the peak values of rainfall and lag effect are not efficiently captured by ML models (Dawson & Wilby 2001; Jain & Srinivasulu 2004; De Vos & Rientjes 2005). Therefore, researchers proposed pre-processing the rainfall data before model application to improve the forecast accuracy (Chau & Wu 2010; Kalteh 2017; Ouyang & Lu 2018; Li & Zhang 2019).

In this review, three pre-processing algorithms were studied: (1) singular spectrum analysis (SSA); (2) wavelet analysis (WA); and (3) ensemble empirical mode decomposition (EEMD). These three pre-processing techniques are powerful mathematical tools that analyze the internal structure of non-stationary time series (Chau & Wu 2010; Feng *et al.* 2015; Ouyang & Lu 2018). Furthermore, these techniques are reported to be capable of removing the noise, providing time–frequency domains of the analyzed signal, and giving insights into the physical aspect of the data series, which helps improve the forecast quality. Table 11 shows the performance comparison of various ML methods with and without pre-processing. It is seen that ML methods were used for daily rainfall forecast (Partal & Kisi 2007; Chau & Wu 2010; Kisi & Cimen 2012; Sumi *et al.*

2012; Unnikrishnan & Jothiprakash 2020) and monthly rainfall forecast (Nourani *et al.* 2009; Ramana *et al.* 2013; Feng *et al.* 2015; Kalteh 2017; Ouyang & Lu 2018; Li & Zhang 2019) in different parts of the world. Conjunction/hybrid models performed better than a single model in both cases. The ML models were used to forecast rainfall a few days to a few months ahead; however, the forecast accuracy reduces with an increase in the lead time. The three data pre-processing techniques stated above were mainly used by studies presented in Table 11. The performance of each of these techniques for rainfall forecasting is explained below.

3.3.1.1. SSA as a pre-processing technique. The effectiveness of SSA as a data pre-processing tool for nonlinear time-series analysis was demonstrated by earlier studies (Sivapragasam *et al.* 2001; Marques *et al.* 2006; Chau & Wu 2010; Zhang *et al.* 2011; Latifoğlu *et al.* 2015; Wang *et al.* 2015; Kalteh 2017; Li & Zhang 2019; Unnikrishnan & Jothiprakash 2020). It is a well-developed, data-adaptive, non-parametric methodology and successfully analysed the internal structure of the time series. Some studies have employed SSA for monthly rainfall forecasts (Latifoğlu *et al.* 2015; Kalteh 2017; Li & Zhang 2019). Before applying the model, SSA decomposes the time series into different components such as trend, periodicity, cyclic components, and noise. The model is then applied to each subcomponent for forecasting the data series. Figure 6 indicates that SSA pre-processing helps to reduce the rainfall forecasting error and improves ML model performance. According to Chau & Wu (2010), SSA applied to a hybrid ANN–SVR (support vector regression) model significantly reduced RMSE (Figure 6(d)) as compared to only the ANN model. The author concluded that three local models could better approximate the rainfall characteristics than a single global model. Unnikrishnan & Jothiprakash (2020) applied the SSA technique to separate stationary and non-stationary components. The author modeled the stationary part by the ARIMA model and the non-stationary part by the ANN model. The forecasted components from both models were added to get the daily rainfall forecast for the whole year, improving previous research limitations. Figure 6 a) to e) indicates that SSA pre-processing helps to reduce the rainfall forecasting error and improves ML model performance.

3.3.1.2. WA and EEMD as a pre-processing technique. Researchers (Partal & Kisi 2007; Wang *et al.* 2009; Adamowski & Chan 2011; Kisi & Cimen 2012; Ramana *et al.* 2013; Feng *et al.* 2015; Ouyang & Lu 2018) demonstrated the feasibility of hybrid WT models in different hydrological operations. The study performed by Ramana *et al.* (2013) stated that the performance of WA-ANN was better than ANN (Figure 7(a)), which was attributed to the ability of wavelet-based models to capture variation in nonlinear dynamics of the temperature. It is established that temperature plays a crucial role in the rainfall process for hilly areas. The efficiency index increased by 30% after applying the wavelet transform. Feng *et al.* (2015) applied the WA-SVM model for monthly rainfall forecast in arid regions. The WA-SVM model performed better than the ANN-SVM model, providing a multi-time ahead rainfall forecast. It was noted that the WA-SVM model does not require much information related to physical processes and can be applied when limited input information is available. The performance at three stations before and after pre-processing is given in Figure 7(b). The authors successfully predicted 1-, 3-, and 6-month-ahead rainfall forecasts using the WA-SVM model, which was superior to ANN and SVM models.

Ouyang & Lu (2018) applied multi-gene genetic programming (MGGP) and echo state networks (ESN) methods for monthly rainfall forecasts. The author used SSA, wavelet transform (WT), and ensemble empirical mode decomposition (EEMD) as data pre-processing techniques. WT and SSA performed better, while the performance of EEMD at all three stations was inferior. Among all, the WT technique was recommended for short-term rainfall forecast as it can capture the exact locality of any variation in data series (Ramana *et al.* 2013; Ouyang & Lu 2018). The study conducted by Partal & Kisi (2007) applied a wavelet and neuro-fuzzy conjunction model for a 1-day ahead daily precipitation forecast (Figure 7(c)). The determination coefficient by the neuro-fuzzy method was around 0.1, while the conjunction model increased 8–9 times, significantly improving the results. This may be due to the efficient forecast of extreme values by the conjunction model. In another study, wavelet transform was combined with support vector regression (WSVR) for daily precipitation forecast (Kisi & Cimen 2012). The mean absolute error (MAE), RMSE, Nash–Sutcliffe Efficiency (NSE), R^2 value for the single SVR model (without pre-processing), and the hybrid model WSVR is presented in Table 11. Figure 7(d) represents a reduction in MAE and RMSE values after pre-processing at two different stations in Turkey. Previous studies conclude that WT is a superior tool for detecting irregularly distributed multi-scale rainfall features in space and time (Partal & Kisi 2007; Kisi & Cimen 2012; Ouyang & Lu 2018).

Table 11 | Performance comparison of ML models before and after pre-processing for rainfall forecast

Author	Study region	Data	Model	Performance index (with pre-processing) MAE (mm)	RMSE (mm)	CE/ NSE	r/R^2	Performance index (without pre-processing) MAE (mm)	RMSE (mm)	CE/ NSE	r/R^2
Partal & Kisi (2007)	Turkey	Daily precipitation data 1987–2001 Afyon station	WT + Neuro-fuzzy	–	1.06	–	0.913	–	3.53	–	0.037
		Izmir station		–	2.17	–	0.913	–	6.83	–	0.124
		Mugla station		–	3.07	–	0.881	–	8.14	–	0.146
Nourani *et al.* (2009)	Iran	Monthly precipitation 1973–1999	WT + ANN	–	–	–	0.784	–	–	–	0.31
Chau & Wu (2010)	China	Daily rainfall Wuxi station 1988–2007	ANN + SVR + SSA	–	4.18	0.87	–	–	10.59	0.17	–
		Zhenwan station 1989–1998	ANN + SVR + SSA	–	3.18	0.92	–	–	10.68	0.09	–
Kisi & Cimen (2012)	Turkey	Daily Precipitation 1987–2001 Afyon station	WT + SVM	9.0	21.4	0.647	0.815	14.2	38.7	0.154	0.103
		Izmir station		13.6	46.5	0.593	0.782	19.6	71.6	0.037	0.276
Sumi *et al.* (2012)	Japan	Daily rainfall 1975–2009	ANN + Multi-model + PCA	–	7.555	0.9973	–	–	14.633	0.9880	–
Ramana *et al.* (2013)	India	Monthly rainfall 1901–1975	WA + ANN	–	63.01	94.78	0.974	–	163.79	64.73	0.807
Feng *et al.* (2015)	China	Monthly rainfall 1960–2012 Yeniugou station	WA + SVM	10.424	12.642	0.863	0.929	12.018	12.568	0.806	0.905
		Qilian station	WA + SVM	7.828	12.689	0.892	0.945	11.57	18.777	0.762	0.875
		Tuole station	WA + SVM	7.345	11.574	0.888	0.943	11.66	18.92	0.7	0.888
Kalteh (2017)	Iran	Monthly precipitation data 1986–2005	ANN + SSA	–	52.257	0.731	0.858	–	91.096	0.183	0.444
Ouyang & Lu (2018)	China	Monthly rainfall 1964–2013 Bamiansha-zhao, Chaganhua, and Chatai station	MGGP + ESN + EEMD + WT + SSA	1.7703	2.6787	0.9953	–	3.8910	5.15570	0.9850	–
Li & Zhang (2019)	China	1983–2013 Monthly average precipitation data	SSA + DA + SVR	5.6120	7.4430	0.9751	0.9782	23.8732	30.4194	0.3430	0.5383
Unnikrishnan & Jothiprakash (2020)	India	1961–2013 Daily rainfall data	SSA–ARIMA–ANN	8.29	15.58	0.69	0.68	9.06	21.83	0.39	0.41

WA, wavelet analysis; EEMD, ensemble empirical mode decomposition; SSA, singular spectrum analysis; ARIMA, auto-regressive integrated moving average; ANN, artificial neural network; WA, wavelet analysis; WT, wavelet transform; ESN, echo state networks; SVR, support vector regression; ANFIS, adaptive neuro-fuzzy inference system; DA, dragonfly algorithm; MGGP, multi-gene genetic programming.

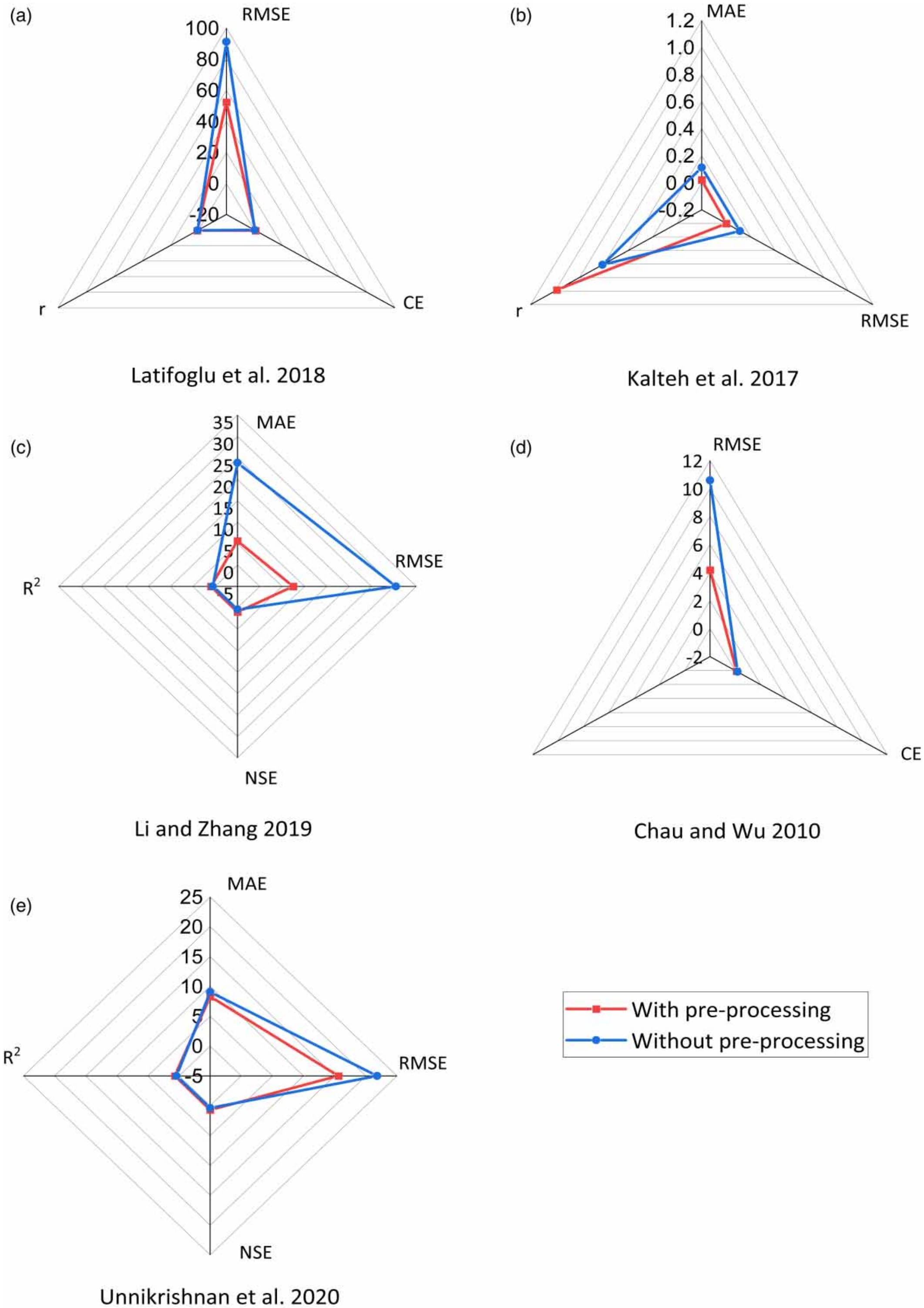


Figure 6 | a) to e) Radar chart showing the performance skill of ML-based model with and without SSA pre-processing for rainfall forecasting.

From the above discussion, the pre-processing technique considerably improved the model performance mainly by detecting the irregular components and forecasting the extreme values. Traditional ML methods failed to capture the peak values efficiently, which can be significantly improved with hybrid models. From Table 11, it is seen that error indices (i.e., RMSE

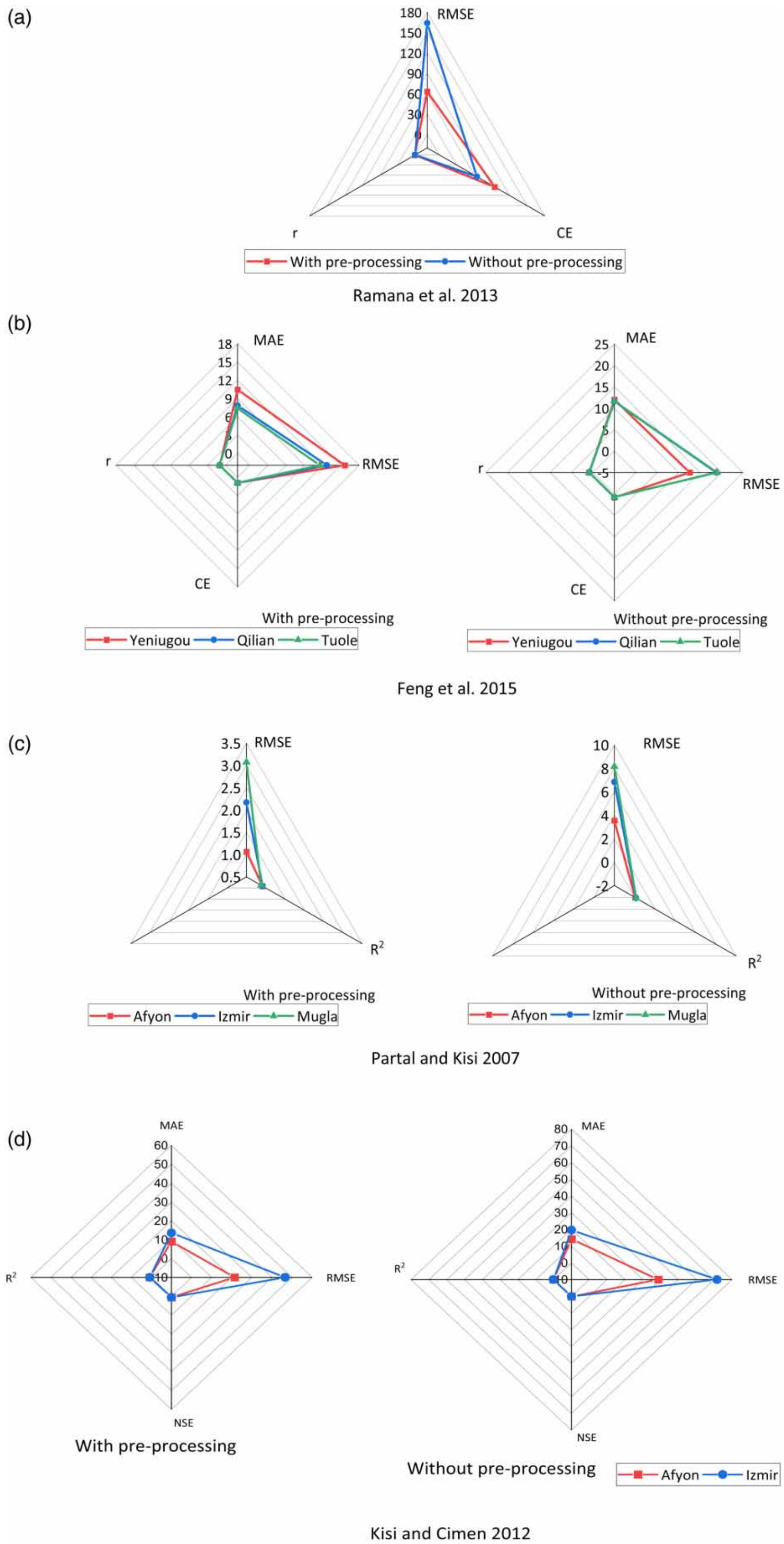


Figure 7 | Radar chart showing the performance skill of ML-based model with and without pre-processing (wavelet analysis and ensemble empirical mode decomposition) for rainfall forecasting.

and MAE) reduce with the application of pre-processing techniques, which may be attributed to the removal of fluctuations in rainfall series.

4. SUMMARY OF THE REVIEW AND THE WAY FORWARD

4.1. Important observations and recommendations from this review

This review paper deals with the PRISMA method of an SR and critical analysis of 110 selected papers from various databases, constituting 1,200 papers without duplication for short-term rainfall forecasting techniques. The study evaluated statistical procedures, physically based numerical weather forecasting models, and ML techniques for the said purpose. The GNSS-derived PWV was found to be capable of analyzing the real-time profile of water vapor content. The method performs well in mountainous regions as it is less affected by altitude and is found to be suitable for nowcasting in tropical regions. The relationship between rainfall, PWV, derivatives of PWV, and other meteorological parameters such as temperature, pressure, relative humidity, and solar radiation is well appraised in the study. It is noted that the threshold PWV value is sensitive to both the season and the location. It is concluded that when GNSS-derived PWV is combined with a suitable ML algorithm, the rainfall nowcast accuracy increases.

The discussed NWP models, ACCESS-VT, ACCESS-A, WRF, and GFS, show that rainfall is underestimated in high elevation areas and overestimated in low elevation areas. It is noted that uncertainty in the forecast increases with the lead time as a longer lead time enhances the internal variabilities in the model. It is concluded that rainfall forecast by the NWP model depends on location, model resolution, season, topography, and forecast lead time. It is also inferred that optimum resolution and lead time can significantly improve forecast accuracy.

The importance of ML techniques in short-term rainfall forecast with and without pre-processing is studied in the present study. A significant improvement in rainfall forecasting is observed with pre-processing techniques. Out of the three pre-processing techniques, the WT and SSA performed better than the EEMD by detecting irregularities, noise, and extreme rainfall values. The pre-processing enhanced the forecast quality by independent modeling of each subseries.

4.2. Future recommendations and development needed

The threshold PWV value is season and location-specific. Additional studies are required for the tropical, subtropical, and temperate regions to strengthen the observed results. Efforts need to be made to generalize threshold PWV by drawing its relationship with seasonal variation, location, and easily observed meteorological parameters. The review highlights the need to explore the optimum data length for rainfall forecast in the temperate, tropical, and subtropical regions. Different ML algorithms can be assimilated with GNSS-derived PWV techniques to improve the nowcasting skill. This review highlights the need to perform more sensitivity studies to understand the influence of different parameters like topography, resolution, and lead time on the forecast skill of NWP models. This review focused only on a few specific NWP models, which needs to be extended further to evaluate their efficiency for short-term rainfall forecast. While considering the ML techniques, only three pre-processing techniques were reviewed. Based on the desired output, other pre-processing techniques must be explored. This review did not explicitly consider the influence of ground information, radar, and satellite data affecting rainfall forecasting. Further research is needed to reduce the static noise in the radar data, which affects the forecast efficiency.

The past few decades have witnessed the frequent occurrence of high-intensity, short duration extreme rainfall events including cloud bursts, which cannot be forecasted with sufficient accuracy. This calls for continuous review and updating of existing short-term rainfall forecasting/nowcasting techniques. Apart from this, the development of new techniques needs to focus on efficiently considering the non-stationarity in rainfall time series. A location-based high-resolution nowcasting model is recommended specifically for urban catchments. The densely populated urban areas will have high temperatures, creating local heat differences resulting in localized rainfall events. Therefore, it is important to quantify the impact of urbanization on rainfall nowcast skill. Due to the complex atmospheric process, there is great difficulty in real-time forecasting. Recently, Ravuri *et al.* (2021) developed an observation-driven approach using deep generative models (DGMs) for skillful nowcast. The potential of DGMs for accurate rainfall nowcasting should be explored in detail. Its performance in terms of lead time, incorporating uncertainty at multiple spatio-temporal scales, and forecasting high-intensity rainfall events should be studied in detail. Similarly, efforts are needed to improve the capability of NWP models by considering the nonlinearities and randomness in rainfall events. For this purpose, the possibility of coupling NWP with the stochastic model (De Luca & Capparelli 2022) can be further evaluated and demonstrated. The rainfall forecast/nowcast with an adequate lead time serves as a mandatory input to the hydrologic models used for urban flood forecasting. Such an integrated module, along with a suitable decision support system, is the need of the hour for effective urban flood management.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abbot, J. & Marohasy, J. 2018 Forecasting of medium-term rainfall using artificial neural networks: case studies from Eastern Australia. In: *Engineering and Mathematical Topics in Rainfall* (T. V. Hromadka & P. Rao, eds.). IntechOpen.

Adamowski, J. & Chan, H. F. 2011 A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology* **407** (1–4), 28–40.

Ajilesh, P., Rakesh, V., Sahoo, S. K. & Himesh, S. 2020 Observed and model-simulated thermodynamic processes associated with urban heavy rainfall events over Bangalore, India. *Meteorological Applications* **27** (1), e1854.

Akrami, S. A., Nourani, V. & Hakim, S. J. S. 2014 Development of non-linear model based on wavelet-ANFIS for rainfall forecasting at Klang Gates Dam. *Water Resources Management* **28** (10), 2999–3018.

Anh, D., Duc Dang, T. & Pham Van, S. 2019 Improved rainfall prediction using combined pre-processing methods and feed-forward neural networks. *J*^{2} **2** (1), 65–83.

Balamurugan, M. S. & Manojkumar, R. 2021 Study of short term rain forecasting using machine learning based approach. *Wireless Networks* **27** (8), 5429–5434.

Balica, S. F., Wright, N. G. & Van der Meulen, F. 2012 A flood vulnerability index for coastal cities and its use in assessing climate change impacts. *Natural Hazards* **64** (1), 73–105.

Barindelli, S., Realini, E., Venuti, G., Fermi, A. & Gatti, A. 2018 Detection of water vapor time variations associated with heavy rain in northern Italy by geodetic and low-cost GNSS receivers. *Earth, Planets and Space* **70** (1), 1–18.

Bastin, S., Champollion, C., Bock, O., Drobinski, P. & Masson, F. 2007 Diurnal cycle of water vapor as documented by a dense GPS network in a coastal area during ESCOMPTE IOP2. *Journal of Applied Meteorology and Climatology* **46** (2), 167–182.

Bauer, P., Thorpe, A. & Brunet, G. 2015 The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55. https://doi.org/10.1038/nature14956.

Benevides, P., Catalao, J. & Miranda, P. M. A. 2015 On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall. *Natural Hazards and Earth System Sciences* **15** (12), 2605–2616.

Benevides, P., Catalao, J. & Nico, G. 2019 Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors. *Remote Sensing* **11** (8), 966.

Benjamin, S. G., Brown, J. M., Brunet, G., Lynch, P., Saito, K. & Schlatter, T. W. 2019 100 years of progress in forecasting and NWP applications. *Meteorological Monographs* **59**, 13–11.

Bhomia, S., Kumar, P. & Kishtawal, C. M. 2019 Evaluation of the weather research and forecasting model forecasts for Indian summer monsoon rainfall of 2014 using ground based observations. *Asia-Pacific Journal of Atmospheric Sciences* **55** (4), 617–628.

Biswas, A. N., Lee, Y. H. & Manandhar, S. 2021 Rainfall forecasting using GPS derived atmospheric gradient and residual for tropical region. *IEEE Transactions on Geoscience and Remote Sensing* **60**, 1–10.

Bowler, N. E., Pierce, C. E. & Seed, A. W. 2006 STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society* **132** (620), 2127–2155.

Brenot, H., Wautelet, G., Warnant, R., Neméghaire, J. & Van Roozendael, M. 2014 *GNSS Meteorology and Impact on NRT Position*. Rotterdam, The Netherlands.

Champollion, C., Masson, F., Van Baelen, J., Walpersdorf, A., Chéry, J. & Doerflinger, E. 2004 GPS monitoring of the tropospheric water vapor distribution and variation during the 9 September 2002 torrential precipitation episode in the Cévennes (southern France). *Journal of Geophysical Research: Atmospheres* **109**, D24102.

Chau, K. W. & Wu, C. L. 2010 A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. *Journal of Hydroinformatics* **12** (4), 458–473.

Chen, J. & Li, G. 2013 Diurnal variations of ground-based GPS-PWV under different solar radiation intensity in the Chengdu Plain. *Journal of Geodynamics* **72**, 81–85.

Choudhury, D., Mehrotra, R., Sharma, A., Sen Gupta, A. & Sivakumar, B. 2019 Effectiveness of CMIP5 decadal experiments for interannual rainfall prediction over Australia. *Water Resources Research* **55** (8), 7400–7418.

Chu, H., Liu, M., Sun, M. & Chen, L. 2018 Rainfall nowcasting by blending of radar data and numerical weather prediction. In: *Understanding of Atmospheric Systems with Efficient Numerical Methods for Observation and Prediction*. IntechOpen, London, pp. 53–70.

Cogan, J. L. 2016 *Change in Weather Research and Forecasting (WRF) Model Accuracy with Age of Input Data from the Global Forecast System (GFS)*. Computational and Information Sciences Directorate, Army Research Laboratory Adelphi United States..

Cramer, S., Kampouridis, M., Freitas, A. A. & Alexandridis, A. K. 2017 An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. *Expert Systems with Applications* **85**, 169–181.

Das, S., Ashrit, R., Iyengar, G. R., Mohandas, S., Das Gupta, M., George, J. P. & Dutta, S. K. 2008 Skills of different mesoscale models over Indian region during monsoon season: forecast errors. *Journal of Earth System Science* **117** (5), 603–620.

Dawson, C. W. & Wilby, R. L. 2001 Hydrological modelling using artificial neural networks. *Progress in Physical Geography* **25** (1), 80–108.

Debray, T. P., Damen, J. A., Snell, K. I., Ensor, J., Hooft, L., Reitsma, J. B. & Moons, K. G. 2017 A guide to systematic review and meta-analysis of prediction model performance. *BMJ* **356**, i6460.

De Luca, D. L. & Capparelli, G. 2022 Rainfall nowcasting model for early warning systems applied to a case over Central Italy. *Natural Hazards* **112** (1), 501–520.

De Vos, N. J. & Rientjes, T. H. M. 2005 Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. *Hydrology and Earth System Sciences* **9**, 111–126.

Diagne, M., David, M., Lauret, P., Boland, J. & Schmutz, N. 2013 Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews* **27**, 65–76.

Dixon, M. & Wiener, G. 1993 TITAN: Thunderstorm identification, tracking, analysis, and nowcasting A radar-based methodology. *Journal of Atmospheric and Oceanic Technology* **10** (6), 785–797.

Duan, Z. & Bastiaanssen, W. G. M. 2013 First results from Version 7 TRMM 3b43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sensing of Environment* **131**, 1–13.

Feng, Q., Wen, X. & Li, J. 2015 Wavelet analysis-support vector machine coupled models for monthly rainfall forecasting in arid regions. *Water Resources Management* **29** (4), 1049–1065.

Foresti, L., Reyniers, M., Seed, A. & Delobbe, L. 2016 Development and verification of a real-time stochastic precipitation nowcasting system for urban hydrology in Belgium. *Hydrology and Earth System Sciences* **20** (1), 505–527.

Gadgil, S., Rajeevan, M. & Nanjundiah, R. 2005 Monsoon prediction – why yet another failure? *Current Science* **88** (9), 1389–1400.

Gallus Jr, W. A.Jr. 1999 Eta simulations of three extreme precipitation events: sensitivity to resolution and convective parameterization. *Weather and Forecasting* **14** (3), 405–426.

Gendt, G., Dick, G., Reigber, C., Tomassini, M., Liu, Y. & Ramatschi, M. 2004 Near real time GPS water vapor monitoring for numerical weather prediction in Germany. *Journal of the Meteorological Society of Japan. Ser. II* **82** (1B), 361–370.

Germann, U. & Zawadzki, I. 2002 Scale-dependence of the predictability of precipitation from continental radar images. Part I: description of the methodology. *Monthly Weather Review* **130** (12), 2859–2873.

Germann, U. & Zawadzki, I. 2004 Scale dependence of the predictability of precipitation from continental radar images. Part II: probability forecasts. *Journal of Applied Meteorology* **43** (1), 74–89.

Giunta, G., Salerno, R., Ceppi, A., Ercolani, G. & Mancini, M. 2019 Effects of model horizontal grid resolution on short- and medium-term daily temperature forecasts for energy consumption application in European cities. In: *Advances in Meteorology* **2019**, 1561697.

Golding, B. W. 2000 Quantitative precipitation forecasting in the UK. *Journal of Hydrology* **239** (1–4), 286–305.

Goswami, P., Shivappa, H. & Goud, S. 2012 Comparative analysis of the role of domain size, horizontal resolution and initial conditions in the simulation of tropical heavy rainfall events. *Meteorological Applications* **19** (2), 170–178.

Gowariker, V. A. S. A. N. T., Thapliyal, V., Sarker, R. P., Mandal, G. S. & Sikka, D. R. 1989 Parametric and power regression models: new approach to long range forecasting of monsoon rainfall in India. *Mausam* **40** (2), 115–122.

Guerova, G., Jones, J., Douša, J., Dick, G., de Haan, S., Pottiaux, E., Bock, O., Pacione, R., Elgered, G., Vedel, H. & Bender, M. 2016 Review of the state of the art and future prospects of the ground-based GNSS meteorology in Europe. *Atmospheric Measurement Techniques* **9** (11), 5385–5406.

Guo, M., Zhang, H. & Xia, P. 2021 Exploration and analysis of the factors influencing GNSS PWV for nowcasting applications. *Advances in Space Research* **67** (12), 3960–3978.

Gutman, S. I., Sahm, S., Jebb Stewart, S. B., Smith, T. & Schwartz, B. 2003 *A new Composite Observing System Strategy for Ground-Based Meteorology*. *12th Symposium on Meteorological Observations and Instrumentation*, Paper No.5.2, Long Beach, CA, Feb. 9-13.

He, S., Raghavan, S. V., Nguyen, N. S. & Liong, S. Y. 2013 Ensemble rainfall forecasting with numerical weather prediction and radar-based nowcasting models. *Hydrological Processes* **27** (11), 1560–1571.

He, Q., Zhang, K., Wu, S., Zhao, Q., Wang, X., Shen, Z. & Liu, X. 2019 Real-time GNSS-derived PWV for typhoon characterizations: a case study for super typhoon Mangkhut in Hong Kong. *Remote Sensing* **12** (1), 104.

Herath, S., Jha, R. & Dutta, D. 2016 Assessing lead time in flood forecasting for better emergency response. In: *2nd Workshop on Comparative Study on Urban Earthquake Disaster Management*, October. Kobe International Exhibition Hall, Japan.

Heuvelink, D., Berenguer, M., Brauer, C. C. & Uijlenhoet, R. 2020 Hydrological application of radar rainfall nowcasting in the Netherlands. *Environment International* **136**, 105431.

Holloway, C. E. & Neelin, J. D. 2010 Temporal relations of column water vapor and tropical precipitation. *Journal of the Atmospheric Sciences* **67** (4), 1091–1105.

Hong, W. 2008 Rainfall forecasting by technological machine learning models. *Applied Mathematics and Computation* **200** (1), 41–57.

Hong, W. C. & Pai, P. F. 2007 Potential assessment of the support vector regression technique in rainfall forecasting. *Water Resources Management* **21** (2), 495–513.
Imhoff, R. O., Brauer, C. C., Overeem, A., Weerts, A. H. & Uijlenhoet, R. 2020 Spatial and temporal evaluation of radar rainfall nowcasting techniques on 1,533 events. *Water Resources Research* **56** (8), e2019WR026723.
IPCC 2021 Chapter three: human influence on the climate system. In: *Climate Change 2021: The Physical Science Basis. Contribution of Working Group14 I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu & B. Zhou, eds.). IPCC, Geneva, Switzerland.
Jabbari, A., So, J. M. & Bae, D. H. 2020 Precipitation forecast contribution assessment in the coupled meteo-hydrological models. *Atmosphere* **11** (1), 34.
Jain, A. & Srinivasulu, S. 2004 Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resources Research* **40** (4), W04302.
Jang, J. & Hong, S. Y. 2014 Quantitative forecast experiment of a heavy rainfall event over Korea in a global model: horizontal resolution versus lead time issues. *Meteorology and Atmospheric Physics* **124** (3), 113–127.
Jee, J. B. & Kim, S. 2017 Sensitivity study on high-resolution WRF precipitation forecast for a heavy rainfall event. *Atmosphere* **8** (6), 96.
Jin, S. & Luo, O. F. 2009 Variability and climatology of PWV from global 13-year GPS observations. *IEEE Transactions on Geoscience and Remote Sensing* **47** (7), 1918–1924.
Kalteh, A. M. 2017 Enhanced monthly precipitation forecasting using artificial neural network and singular spectrum analysis conjunction models. *INAE Letters* **2** (3), 73–81.
Karabatić, A., Weber, R. & Haiden, T. 2011 Near real-time estimation of tropospheric water vapour content from ground based GNSS data and its potential contribution to weather now-casting in Austria. *Advances in Space Research* **47** (10), 1691–1703.
Kawase, H., Takeuchi, Y., Sato, T. & Kimura, F. 2006 Precipitable water vapor around orographically induced convergence line. *Sola* **2**, 25–28.
Khaniani, A. S., Motieyan, H. & Mohammadi, A. 2021 Rainfall forecast based on GPS PWV together with meteorological parameters using neural network models. *Journal of Atmospheric and Solar-Terrestrial Physics* **214**, 105533.
Kisi, O. & Cimen, M. 2012 Precipitation forecasting by using wavelet-support vector machine conjunction model. *Engineering Applications of Artificial Intelligence* **25** (4), 783–792.
Krishnan, R. & Sugi, M. 2003 Pacific decadal oscillation and variability of the Indian summer monsoon rainfall. *Climate Dynamics* **21** (3), 233–242.
Kumar, A., Pai, D. S., Singh, J. V., Singh, R. & Sikka, D. R. 2012 *Statistical Models for Long-Range Forecasting of Southwest Monsoon Rainfall Over India Using Step Wise Regression and Neural Network*.
Kumar, P., Ojha, S. P., Singh, R., Kishtawal, C. M. & Pal, P. K. 2016 Performance of weather research and forecasting model with variable horizontal resolution. *Theoretical and Applied Climatology* **126**, 705–713. https://doi.org/10.1007/s00704-015-1607-7.
Kumar, P., Kishtawal, C. M. & Pal, P. K. 2017 Impact of ECMWF, NCEP, and NCMRWF global model analysis on the WRF model forecast over Indian region. *Theoretical and Applied Climatology* **127**, 143–151.
Lafaysse, M., Hingray, B., Mezghani, A., Gailhard, J. & Terray, L. 2014 Internal variability and model uncertainty components in future hydro-meteorological projections: the Alpine Durance basin. *Water Resources Research* **50** (4), 3317–3341.
Lakshmanan, V., Smith, T., Stumpf, G. & Hondl, K. 2007 The warning decision support system–integrated information. *Weather and Forecasting* **22** (3), 596–612.
Latifoğlu, L., Kişi, Ö. & Latifoğlu, F. 2015 Importance of hybrid models for forecasting of hydrological variable. *Neural Computing and Applications* **26** (7), 1669–1680.
Lepore, C., Veneziano, D. & Molini, A. 2015 Temperature and CAPE dependence of rainfall extremes in the eastern United States. *Geophysical Research Letters* **42** (1), 74–83.
Li, P. W. & Lai, E. S. 2004 Short-range quantitative precipitation forecasting in Hong Kong. *Journal of Hydrology* **288** (1–2), 189–209.
Li, W. & Zhang, J. 2019 An innovated integrated model using singular spectrum analysis and support vector regression optimized by intelligent algorithm for rainfall forecasting. *Journal of Autonomous Intelligence* **2**, 46. https://doi.org/10.32629/jai.v2i1.37.
Li, Y., Lu, G., Wu, Z., He, H., Shi, J., Ma, Y. & Weng, S. 2016 Evaluation of optimized WRF precipitation forecast over a complex topography region during flood season. *Atmosphere* **7** (11), 145.
Li, H., Wang, X., Wu, S., Zhang, K., Chen, X., Qiu, C. & Li, L. 2020 Development of an improved model for prediction of short-term heavy precipitation based on GNSS-derived PWV. *Remote Sensing* **12** (24), 4101.
Li, H., Wang, X., Choy, S., Jiang, C., Wu, S., Zhang, J. & Zhang, K. 2022 Detecting heavy rainfall using anomaly-based percentile thresholds of predictors derived from GNSS-PWV. *Atmospheric Research* **265**, 105912.
Liguori, S. & Rico-Ramirez, M. A. 2014 A review of current approaches to radar-based quantitative precipitation forecasts. *International Journal of River Basin Management* **12** (4), 391–402.
Liu, J., Wang, J., Pan, S., Tang, K., Li, C. & Han, D. 2015 A real-time flood forecasting system with dual updating of the NWP rainfall and the river flow. *Natural Hazards* **77** (2), 1161–1182.
Liu, Y., Zhao, Q., Yao, W., Ma, X., Yao, Y. & Liu, L. 2019 Short-term rainfall forecast model based on the improved BP–NN algorithm. *Scientific Reports* **9** (1), 1–12.

Łoś, M., Smolak, K., Guerova, G. & Rohm, W. 2020 GNSS-based machine learning storm nowcasting. *Remote Sensing* **12** (16), 2536.
Manandhar, S., Lee, Y. H., Meng, Y. S., Yuan, F. & Ong, J. T. 2018 GPS-derived PWV for rainfall nowcasting in tropical region. *IEEE Transactions on Geoscience and Remote Sensing* **56** (8), 4835–4844.
Manandhar, S., Dev, S., Lee, Y. H., Meng, Y. S. & Winkler, S. 2019a A data-driven approach for accurate rainfall prediction. *IEEE Transactions on Geoscience and Remote Sensing* **57** (11), 9323–9331.
Manandhar, S., Lee, Y. H. & Meng, Y. S. 2019b GPS-PWV based improved long-term rainfall prediction algorithm for tropical regions. *Remote Sensing* **11** (22), 2643.
Marques, C. A. F., Ferreira, J. A., Rocha, A., Castanheira, J. M., Melo-Gonçalves, P., Vaz, N. & Dias, J. M. 2006 Singular spectrum analysis and forecasting of hydrological time series. *Physics and Chemistry of the Earth, Parts A/B/C* **31** (18), 1172–1179.
Martin, G. 1998 *An Outstanding Performance by the Eta-10 for the Southern California Storm of 23 February 1998*. Western Region Technical Attachment.
Mass, C. F., Ovens, D., Westrick, K. & Colle, B. A. 2002 Does increasing horizontal resolution produce more skillful forecasts? The results of two years of real-time numerical weather prediction over the Pacific Northwest. *Bulletin of the American Meteorological Society* **83** (3), 407–430.
Mawandha, H. G., Kishimoto, M. & Oishi, S. 2019 GNSS-based PWV application for short term rainfall prediction in mountainous region. In: *IOP Conference Series: Earth and Environmental Science*, November, Vol. 355, No. 1, p. 012070. IOP Publishing.
Moya-Álvarez, A. S., Gálvez, J., Holguín, A., Estevan, R., Kumar, S., Villalobos, E. & Silva, Y. 2018 Extreme rainfall forecast with the WRF-ARW model in the Central Andes of Peru. *Atmosphere* **9** (9), 362.
Munot, A. A. & Kumar, K. K. 2007 Long range prediction of Indian summer monsoon rainfall. *Journal of Earth System Science* **116** (1), 73–79.
Nourani, V., Alami, M. T. & Aminfar, M. H. 2009 A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence* **22** (3), 466–472.
Ouyang, Q. & Lu, W. 2018 Monthly rainfall forecasting using echo state networks coupled with data pre-processing methods. *Water Resources Management* **32** (2), 659–674.
Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R. & van Ypserle, J. P. 2014 *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland, p. 151.
Pall, P., Allen, M. R. & Stone, D. A. 2007 Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO_2 warming. *Climate Dynamics* **28** (4), 351–363.
Partal, T. & Kişi, Ö. 2007 Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology* **342** (1–2), 199–212.
Parthasarathy, B., Kumar, K. R. & Munot, A. A. 1993 Homogeneous Indian monsoon rainfall: variability and prediction. *Proceedings of the Indian Academy of Sciences, Earth and Planetary Sciences* **102** (1), 121–155.
Prakash, S., Mitra, A. K., Momin, I. M., Rajagopal, E. N., Milton, S. F. & Martin, G. M. 2016 Skill of short-to medium-range monsoon rainfall forecasts from two global models over India for hydro-meteorological applications. *Meteorological Applications* **23** (4), 574–586.
Priego, E., Jones, J., Porres, M. J. & Seco, A. 2017 Monitoring water vapour with GNSS during a heavy rainfall event in the Spanish Mediterranean area. *Geomatics, Natural Hazards and Risk* **8** (2), 282–294.
Pu, Z. & Kalnay, E. 2019 Numerical weather prediction basics: models, numerical methods, and data assimilation. In: *Handbook of Hydro-Meteorological Ensemble Forecasting* (Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake, eds.). Springer, Berlin, Heidelberg, pp. 67–97.
Rahimi, Z., Shafri, H. Z. M. & Norman, M. 2018 A GNSS-based weather forecasting approach using nonlinear auto regressive approach with exogenous input (NARX). *Journal of Atmospheric and Solar-Terrestrial Physics* **178**, 74–84.
Rajeevan, M., Guhathakurta, P. & Thapliyal, V. 2000 New models for long range forecasts of summer monsoon rainfall over North West and Peninsular India. *Meteorology and Atmospheric Physics* **73** (3), 211–225.
Rajeevan, M., Pai, D. S., Anil Kumar, R. & Lal, B. 2007 New statistical models for long-range forecasting of southwest monsoon rainfall over India. *Climate Dynamics* **28** (7), 813–828.
Ramana, R. V., Krishna, B., Kumar, S. R. & Pandey, N. G. 2013 Monthly rainfall prediction using wavelet neural network analysis. *Water Resources Management* **27** (10), 3697–3711.
Ramírez, L. & Vindel, J. M. 2017 Forecasting and nowcasting of DNI for concentrating solar thermal systems. In: *Advances in Concentrating Solar Thermal Research and Technology* (M. J. Blanco & L. Ramirez Santigosa, eds.). Woodhead Publishing, Sawston, UK, pp. 293–310.
Randall, D. A., Bitz, C. M., Danabasoglu, G., Denning, A. S., Gent, P. R., Gettelman, A. & Thuburn, J. 2019 100 years of earth system model development. *Meteorological Monographs* **59**, 12.1–12.66.
Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P. & Mohamed, S. 2021 Skillful precipitation nowcasting using deep generative models of radar. *Nature* **597** (7878), 672–677.
Ridwan, W. M., Sapitang, M., Aziz, A., Kushiar, K. F., Ahmed, A. N. & El-Shafie, A. 2021 Rainfall forecasting model using machine learning methods: case study Terengganu, Malaysia. *Ain Shams Engineering Journal* **12** (2), 1651–1663.
Roy, S. S., Mohapatra, M., Tyagi, A. & Bhowmik, S. R. 2019 A review of nowcasting of convective weather over the Indian region. *Mausam* **70** (3), 465–484.

Salvi, K., Villarini, G. & Vecchi, G. A. 2017 High resolution decadal precipitation predictions over the continental United States for impacts assessment. *Journal of Hydrology* **553**, 559–573.

Sangiorgio, M., Barindelli, S., Guglieri, V., Biondi, R., Solazzo, E., Realini, E. & Guariso, G. 2018 A comparative study on machine learning techniques for intense convective rainfall events forecasting. In: *International Conference on Time Series and Forecasting* (I. Rojas, ed.). Springer, Cham, pp. 305–317.

Seco, A., Ramírez, F., Serna, E., Prieto, E., García, R., Moreno, A. & Priego, J. E. 2012 Rain pattern analysis and forecast model based on GPS estimated atmospheric water vapor content. *Atmospheric Environment* **49**, 85–93.

Shahrban, M., Walker, J. P., Wang, Q. J., Seed, A. & Steinle, P. 2016 An evaluation of numerical weather prediction based rainfall forecasts. *Hydrological Sciences Journal* **61** (15), 2704–2717.

Sharifi, M. A., Sam Khaniani, A. & Joghataei, M. 2015 Comparison of GPS precipitable water vapor and meteorological parameters during rainfalls in Tehran. *Meteorology and Atmospheric Physics* **127** (6), 701–710.

Sharma, K., Ashrit, R., Bhatla, R., Mitra, A. K., Iyengar, G. R. & Rajagopal, E. N. 2017 Skill of predicting heavy rainfall over India: improvement in recent years using UKMO global model. *Pure and Applied Geophysics* **174** (11), 4241–4250.

Sharma, K., Ashrit, R., Kumar, S., Milton, S., Rajagopal, E. N. & Mitra, A. K. 2021 Unified model rainfall forecasts over India during 2007–2018: evaluating extreme rains over hilly regions. *Journal of Earth System Science* **130** (2), 1–13.

Shehu, B. & Haberlandt, U. 2021 Relevance of merging radar and rainfall gauge data for rainfall nowcasting in urban hydrology. *Journal of Hydrology* **594**, 125931.

Shi, J., Xu, C., Guo, J. & Gao, Y. 2015 Real-time GPS precise point positioning-based precipitable water vapor estimation for rainfall monitoring and forecasting. *IEEE Transactions on Geoscience and Remote Sensing* **53** (6), 3452–3459.

Shoji, Y. 2013 Retrieval of water vapor inhomogeneity using the Japanese nationwide GPS array and its potential for prediction of convective precipitation. *Journal of the Meteorological Society of Japan. Ser. II* **91** (1), 43–62.

Shrestha, D. L., Robertson, D. E., Wang, Q. J., Pagano, T. C. & Hapuarachchi, P. 2012 Evaluation of numerical weather prediction model precipitation forecasts for use in short-term streamflow forecasting. *Hydrology & Earth System Sciences Discussions* **9** (11), 12563–12611.

Singh, J. P. & Pai, D. S. 1996 An oceanic model for the prediction of southwest monsoon rainfall over India. *Mausam* **47** (1), 91–98.

Sivapragasam, C., Liong, S. Y. & Pasha, M. F. K. 2001 Rainfall and runoff forecasting with SSA–SVM approach. *Journal of Hydroinformatics* **3** (3), 141–152.

Smith, D. M., Eade, R., Scaife, A. A., Caron, L. P., Danabasoglu, G., DelSole, T. M. & Yang, X. 2019 Robust skill of decadal climate predictions. *Npj Climate and Atmospheric Science* **2** (1), 1–10.

Solomatine, D. P. & Ostfeld, A. 2008 Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics* **10** (1), 3–22.

Sridevi, C., Kumar Singh, K., Suneetha, P., Reval Durai, V. & Kumar, A. 2018 Rainfall forecast skill of global forecasting system (GFS) model over India during summer monsoon 2015. *Geofizika* **35** (1), 40–52.

Subramanian, K. & Gopalakrishnan, T. 2020 *Agrometeorological Advisory Services India: An Assessment*. Centre for Science and Environment, New Delhi.

Sumi, S. M., Zaman, M. & Hirose, H. 2012 A rainfall forecasting method using machine learning models and its application to the Fukuoka city case. *International Journal of Applied Mathematics and Computer Science* **22** (4), 841–854.

Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J. & Pinto, J. 2014 Use of NWP for nowcasting convective precipitation: recent progress and challenges. *Bulletin of the American Meteorological Society* **95** (3), 409–426.

Suparta, W. & Alhasa, K. M. 2015 Modeling of zenith path delay over Antarctica using an adaptive neuro fuzzy inference system technique. *Expert Systems with Applications* **42** (3), 1050–1064.

Thapliyal, V. 1982 Stochastic dynamic model for long range prediction of monsoon rainfall in Peninsular India. *Mausam* **33** (4), 399–404.

Thapliyal, V. 1997 Preliminary and final long range forecast for seasonal monsoon rainfall over India. *Journal of Arid Environments* **36** (3), 385–403.

Unnikrishnan, P. & Jothiprakash, V. 2020 Hybrid SSA-ARIMA-ANN model for forecasting daily rainfall. *Water Resources Management* **34** (11), 3609–3623.

Vareed Joseph, P., Gokulapalan, B., Nair, A. & Sheela Wilson, S. 2013 Variability of summer monsoon rainfall in India on inter-annual and decadal time scales. *Atmospheric and Oceanic Science Letters* **6** (5), 398–403.

Wang, Y., Leung, L. R., McGregor, J. L., Lee, D. K., Wang, W. C., Ding, Y. & Kimura, F. 2004 Regional climate modeling: progress, challenges, and prospects. *Journal of the Meteorological Society of Japan. Ser. II* **82** (6), 1599–1628.

Wang, W., Jin, J. & Li, Y. 2009 Prediction of inflow at three gorges dam in Yangtze River with wavelet network model. *Water Resources Management* **23** (13), 2791–2803.

Wang, Y., Guo, S., Xiong, L., Liu, P. & Liu, D. 2015 Daily runoff forecasting model based on ANN and data pre-processing techniques. *Water* **7** (8), 4144–4160.

Wang, G., Yang, J., Wang, D. & Liu, L. 2016 A quantitative comparison of precipitation forecasts between the storm-scale numerical weather prediction model and auto-nowcast system in Jiangsu, China. *Atmospheric Research* **181**, 1–11.

Wiston, M. & Mphale, K. M. 2018 Weather forecasting: from the early weather wizards to modern-day weather predictions. *Journal of Climatology & Weather Forecasting* **6** (2), 1–9.

WMO 2017 *Guidelines for Nowcasting Techniques*. WMO. Available at: https://library. wmo. int/opac/doc_num. php

Yan, X., Ducrocq, V., Poli, P., Hakam, M., Jaubert, G. & Walpersdorf, A. 2009 Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall. *Journal of Geophysical Research: Atmospheres* **114**, D03104.

Yao, Y., Shan, L. & Zhao, Q. 2017 Establishing a method of short-term rainfall forecasting based on GNSS-derived PWV and its application. *Scientific Reports* **7** (1), 1–11.

Yeh, T. K., Shih, H. C., Wang, C. S., Choy, S., Chen, C. H. & Hong, J. S. 2018 Determining the precipitable water vapor thresholds under different rainfall strengths in Taiwan. *Advances in Space Research* **61** (3), 941–950.

Yuan, Y., Zhang, K., Rohm, W., Choy, S., Norman, R. & Wang, C. S. 2014 Real-time retrieval of precipitable water vapor from GPS precise point positioning. *Journal of Geophysical Research: Atmospheres* **119** (16), 10044–10057.

Zhang, Q., Wang, B. D., He, B., Peng, Y. & Ren, M. L. 2011 Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. *Water Resources Management* **25** (11), 2683–2703.

Zhao, Q., Yao, Y. & Yao, W. 2018 GPS-based PWV for precipitation forecasting and its application to a typhoon event. *Journal of Atmospheric and Solar-Terrestrial Physics* **167**, 124–133.

Zhao, Q., Liu, Y., Ma, X., Yao, W., Yao, Y. & Li, X. 2020 An improved rainfall forecasting model based on GNSS observations. *IEEE Transactions on Geoscience and Remote Sensing* **58** (7), 4891–4900.

Zhao, Q., Liu, Y., Yao, W. & Yao, Y. 2022 Hourly rainfall forecast model using supervised learning algorithm. *IEEE Transactions on Geoscience and Remote Sensing* **60**, 1–9.

Zhou, J., Zhang, H., Zhang, J., Zeng, X., Ye, L., Liu, Y. & Chen, Y. 2018 WRF model for precipitation simulation and its application in real-time flood forecasting in the Jinshajiang River Basin, China. *Meteorology and Atmospheric Physics* **130** (6), 635–647.

Zhu, D., Wang, G., Ren, Q. & Ilyas, A. M. 2020 Hydrological evaluation of hourly merged satellite–station precipitation product in the mountainous basin of China using a distributed hydrological model. *Meteorological Applications* **27** (2), e1909.

First received 27 February 2022; accepted in revised form 11 July 2022. Available online 20 July 2022