Atmospheric Research 282 (2023) 106548

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Nowcasting extreme rain and extreme wind speed with machine learning techniques applied to different input datasets

Sandy Chkeir, Aikaterini Anesiadou, Alessandra Mascitelli, Riccardo Biondi *

Università degli Studi di Padova, Dipartimento di Geoscienze, Padova, Italy

ARTICLE INFO

Keywords:
Nowcasting
Machine learning
Weather
Extreme rain
Extreme wind speed

ABSTRACT

Predicting extreme weather events in a short time period and their developing in localized areas is a challenge. The nowcasting of severe and extreme weather events is an issue for air traffic management and control because it affects aviation safety, and determines delays and diversions. This work is part of a larger study devoted to nowcasting rain and wind speed in the area of Malpensa airport by merging different datasets. We use as reference the weather station of Novara to develop a nowcasting machine learning model which could be reusable in other locations. In this location we have the availability of ground-based weather sensors, a Global Navigation Satellite System (GNSS) receiver, a C-band radar and lightning detectors. Our analysis shows that the Long Short-Term Memory Encoder Decoder (LSTM E/D) approach is well suited for the nowcasting of meteorological variables. The predictions are based on 4 different datasets configurations providing rain and wind speed nowcast for 1 h with a time step of 10 min. The results are very promising with the extreme wind speed probability of detection higher than 90%, the false alarms lower than 2%, and a good performance in extreme rain detection for the first 30 min. The configuration using just weather stations and GNSS data in input provides excellent performances and should be preferred to the other ones, since it refers to the pre-convective environment, and thus can be adaptable to any weather conditions.

1. Introduction

1.1. Severe weather patterns

Monitoring and predicting extreme atmospheric events is very challenging, especially when they develop locally in a short time range. Despite the enhancement in remote sensing due to the use of satellite measurements and improvements in model parameterizations, there are still large uncertainties on the knowledge of the dynamical processes of deep convective systems and severe weather events.

Severe and extreme weather events cause many deaths, injuries and damages every year, accounting for the major economic damages in several countries (Pielke Jr et al., 2003; Emanuel, 2005), and they are one of the major risks for aviation safety. The number and the intensity of such phenomena increased in the last decades in some areas of the globe including Europe (Hov et al., 2013; Rädler et al., 2018; Rädler et al., 2019).

Related to the increase in the frequency of such events, in recent years, there has been a growing interest in the study of natural disasters. The loss related to natural catastrophes worldwide increased by a factor of about three within the last 4 decades (Hoeppe, 2016). These phenomena are mainly associated with weather-related events, e.g., storms and floods. There is some justification in assuming that changes in the atmosphere, which include a widespread increase of water vapor content in the lowest layers of the troposphere (Rädler et al., 2018), are playing a relevant role (Hoeppe, 2016). Several studies have been performed in this field, leading to interesting considerations regarding the spatial and temporal evolution of intense meteorological phenomena and assessing possible future scenarios by the end of the 21th century (Rädler et al., 2019). In this context, our study focuses on the development of an efficient nowcasting model for extreme rainfall and extreme wind speed in the area of Milano Malpensa airport with the future objective of creating a supporting tool for aviation managers.

Heavy rainfalls show significant trends over Europe (Van den Besselaar et al., 2013; Maraun, 2013), and usually result in impactful floods. Daily and sub-daily heavy precipitation is indeed one of the most significant hazards (Scoccimarro et al., 2015), and rainfall events are projected to intensify with climate change (Ban et al., 2015; Stott, 2016).

* Corresponding author.
E-mail address: riccardo@biondiriccardo.it (R. Biondi).

https://doi.org/10.1016/j.atmosres.2022.106548
Received 16 June 2022; Received in revised form 28 November 2022; Accepted 28 November 2022
Available online 2 December 2022

Heavy precipitations are often connected with the lightning formation. According to the literature, the simulation of the number of lightning events is higher in central and southern Europe, specifically across high altitude, mountainous regions, and this is in accordance with what has been observed by the lightning detection networks (Rädler et al., 2018). Another interesting aspect in the projections is that they show an increase in some parts of Central and Eastern Europe of the wind shear in unstable situations, probably related to specific conditions of air mass displacement (van Delden, 2001). It is now clear that the determination of descriptive and predictive algorithms for the atmospheric behavior (Bonafoni et al., 2019; Laviola et al., 2020; Mascitelli et al., 2020) is crucial in an area such as Europe, which is strongly affected by severe weather phenomena.

Europe has a large and dense coverage of meteorological sensors including weather stations, lightning detectors, ground-based Global Navigation Satellite System (GNSS) receivers and radars. However, the management of all the networks is scattered within a number of different agencies and it is extremely difficult to collect all the needed datasets at European level to perform a climatological study aimed at identifying sensitive areas.

We select northern Italy as a hotspot for our analyses for three different reasons:

- it is an area with a high frequency of extreme weather events with positive trends in the last decades (Rädler et al., 2019);
- there is a relatively easy access to meteorological data;
- it is of high interest for aviation purposes for the presence of two large airports (Milano Malpensa and Bergamo Orio al Serio).

1.2. Severe weather forecast

The atmospheric Water Vapor (WV) plays a key role in the storm development; Trenberth (2012) shows that a small increase of WV is amplified in weather systems, promoting more intense extreme events. Eighty-five percent of strong convection (overshooting the tropopause) comes from small systems with a short lifetime (Liu and Zipser, 2005) and due to this reason, it is hard to be detected and monitored. Several studies have shown the capabilities of ground-based GNSS for studying and predicting severe weather events (Adams et al., 2017; Sapucci et al., 2019; Mascitelli et al., 2020). The GNSS measurements allow the estimation of the atmospheric Integrated WV (IWV), which is the engine of the convection (Bonafoni et al., 2019). Moreover, there is evidence that an abrupt increase in the total lightning discharge rate (Williams et al., 1999) often precedes severe weather occurrences on the ground (Williams et al., 1999; Darden et al., 2010) with a lag time of a few minutes to nearly one hour (Wu et al., 2018).

The idea of this work is to combine the capabilities of a dense network of weather stations and ground-based GNSS receivers with the radar and lightning detectors (D'Adderio et al., 2020) for nowcasting extreme rainfall and extreme wind speed with high temporal and spatial resolution to support the aviation management (Solazzo et al., 2020). Airports are bottlenecks in the Air Traffic Management (ATM) network and are especially impacted by thunderstorms in their vicinity (Gultepe et al., 2019). The importance of severe weather nowcasting for aviation and the performances and limitations of an operational Numerical Weather Prediction (NWP) model have already been shown by the German Weather Service (James et al., 2018). Bonelli and Marcacci (2008) developed a thunderstorm tracking algorithm in northern Italy based on radar and lightning data; Kohn et al. (2011) developed an algorithm based only on lightning data to nowcast thunderstorm in the Mediterranean with promising results but with the outlook to include additional dataset types and increase the period of availability. Recently it has been demonstrated that the machine learning Long Short-Term Memory (LSTM) techniques provides significant results for forecasting (Tekin et al., 2021) and nowcasting purposes (Shi et al., 2017; Ayzel et al., 2020) providing better results than the well-known and complex Weather and Research Forecasting (WRF) model, with potential for efficient and accurate weather forecasting up to 12 h (Hewage et al., 2021). A tentative development of LSTM model and deep generative model to nowcast precipitation has been recently done by using only radar reflectivity as input (Klocek et al., 2021; Ravuri et al., 2021) and combining radar and weather stations data together (Zhang et al., 2021). Machine learning also provided excellent performances in predicting wind speed (Khosravi et al., 2018). Cramer et al. (2017) evaluated the performances of different machine learning algorithms to predict the rainfall in 42 cities with different climatological characteristics; Benevides et al. (2019) were the first ones to develop a machine algorithm including weather stations data together with GNSS WV.

With this work we pose the basis to develop a rainfall and wind speed nowcasting model including different input data in a localized area. The model is run for several stations around the airport with the objective to create a polygon highlighting the highest risk for the aviation. The results presented in this paper are used to define the highest temporal resolution we can achieve and to choose the better model and configuration which can be applied to the other stations.

2. Data

2.1. Data collection and pre-processing

For this study, 4 different types of data are gathered: weather station, GNSS, radar and lightning data. Each type of data is crucial for the development of the method. The weather station data define the environmental conditions useful to study the convection, the GNSS data give more detailed insights on the WV in the atmosphere, the radar data provide information on the existence of different types of precipitation processes, and the lightning data denote the existence of extreme precipitation events.

This study is part of the ALARM project (Soler, 2021) targeting Milano Malpensa airport from an aviation safety perspective. In this paper we focus on a specific location, Novara, close to Malpensa airport, because it is the only location of the hotspot area for which we have the availability of all the aforementioned datasets. The data acquisition was not possible for all variables in the same period so we report the details of data availability in Table 1, and the map of the different locations involved in Fig. 1.

2.1.1. Weather station data

The weather data of interest are pressure, temperature, relative humidity, rain, wind speed and wind direction. They are provided, on demand, by the regional agency for environmental protection of Piemonte (http://www.arpa.piemonte.it/) with a temporal resolution of 10 min (Table 1). All the parameters (except rain) are provided as averaged values, while rain is the accumulated value over the 10-min time period. Not all the sensors are exactly collocated and some data were corrupted, so we needed to merge the data of different stations, considering similar environmental/climatological conditions and the same height above mean sea level (Fig. 1). For these reasons, we set a maximum distance between the stations not larger than 20 km and we converted the pressure and temperature according to the height. The closest wind sensor to Novara is in Cameri (5 km from Novara) and the closest pressure sensor is in Arconate (20 km from Novara and 15 km from Cameri). The pressure data from Arconate (elevation 182 m) is adjusted to the elevation of Novara station (151 m) by using the barometric formula as in Berberan-Santos et al. (1997). All the outliers of each parameter were removed.

2.1.2. GNSS data

We use the Zenith Total Delay (ZTD) time series provided, just for the purpose of the project by the Geomatics Research & Development s.r.l. (GReD) (Lagasio et al., 2019), with a temporal resolution of 30 s. We average the data with 10-min sampling to make them consistent with the

Table 1
List of variables used in this study. The availability of radar reflectivity only refers to severe events with reflectivity higher than 36 dBz.

Variable	Availability	Unit	Station type	Location	Lat (°)	Lon (°)	Elevation (m)	Period
Temperature	100%	°C	weather	Novara	45.44	8.63	151	2010–2020
Pressure	98.85%	mb	weather	Arconate	45.54	8.84	182	2010–2019
Rain	100%	mm	weather	Novara	45.44	8.63	151	2010–2020
Wind direction	94.49%	o	weather	Cameri	45.54	8.69	173	2010–2020
Wind speed	99.38%	m/s	weather	Cameri	45.54	8.69	173	2010–2020
Humidity	99.86%	%	weather	Novara	45.44	8.63	151	2010–2020
ZTD	99.07%	m	GNSS	Novara	45.44	8.61	219	2011–2020
Average Reflectivity	100%	dBZ	radar	–	–	–	–	2012–2019
Maximum Reflectivity	100%	dBZ	radar	–	–	–	–	2012–2019
# of Lightning	100%		lightning					2015–2019
Type of Lightning	100%		lightning					2015–2019
Lightning height	100%	m	lightning					2015–2019

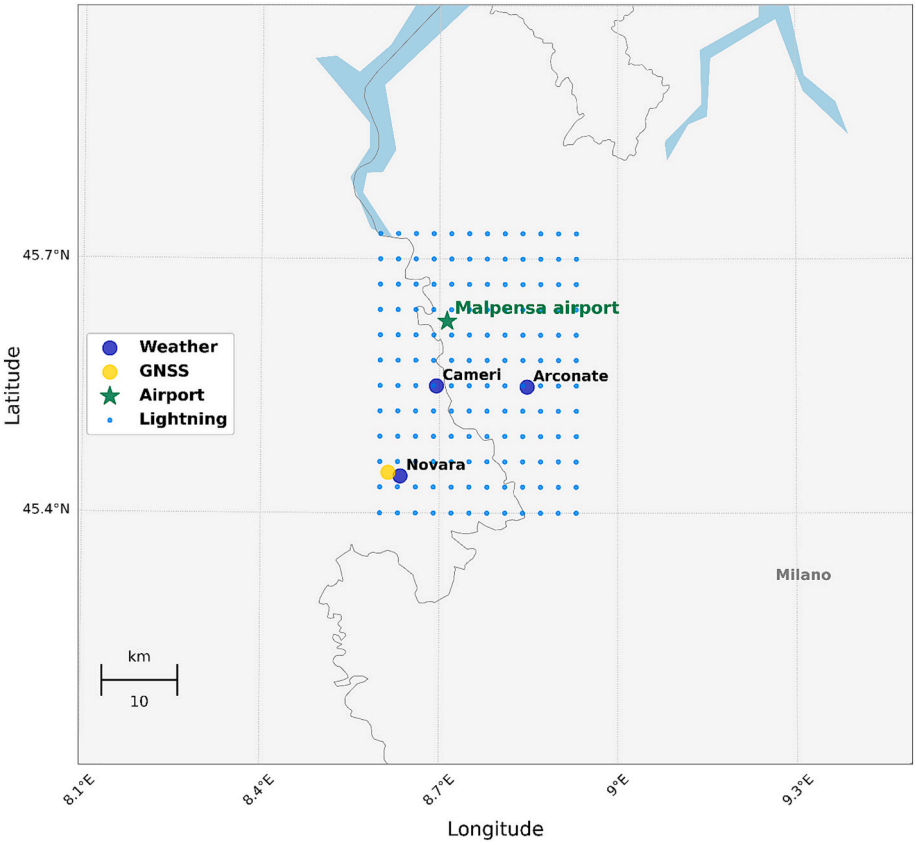

Fig. 1. Map of the locations of Novara, Malpensa airport, weather and GNSS stations and the wide area of lightning stations' coverage.

meteorological data temporal resolution and we remove all the outliers.

2.1.3. Radar data

The radar is located in Monte Lema (Canton of Ticino, Switzerland) at an altitude of 1630 m, about 70 km north of Novara and it covers the entire Lombardy and Piedmont regions. It is a C-band radar with a fixed elevation angle (tilt) sequence of 20 tilts per 5-min period (Joss et al., 1998). We use the coordinates, area, velocity, average and maximum reflectivity of the convective cell with a temporal resolution of 5 min coming from the Thunderstorms Radar Tracking (TRT) algorithm (Nisi et al., 2018). The TRT uses a threshold of 36 dBZ to select the convection and separates each single storm and cell from nearby storms using the Max Echo product. We therefore have, for a given time period, the reflectivity information only above this threshold. To make this data consistent with the other datasets, we resample the radar data with a 10-min sampling. The measurement corresponding to a convective cell is assigned to a station if the convective cell is collocated with the station itself. To achieve the 10-min sampling from the initial 5-min one, the maximum value of the respective variable of interest is picked. The TRT products were provided just for the purpose of the project by Meteowiss.

2.1.4. Lightning data

The lightning data are provided just for the purpose of the project by the Earth Networks Total Lightning Network (ENTLN) (Liu and Heckman, 2012). There are 2 LF-VHF stations in the proximity of Milano Malpensa. Fig. 1 shows the coverage of those data in relation to Novara location. The area covered is of about 20 × 30 km^2. The available information is the time, location, number of lightnings related to the covered area and time period, intensity, type of lightning and the average height of intra cloud lightning. The type of lightning, based on the available information, can either be intra-cloud (lightning that happens completely inside the cloud) or cloud-to-ground (lightning that occurs between the cloud and the ground).

2.1.5. Overview of the input dataset

The meteorological data consist of a continuous time series of 12 parameters (Table 1), Our model needs continuous time series so all the possible gaps must be filled and cannot be removed to preserve the time order of the samples. We filled the gaps using the random regression imputation method as explained in Section 3.3.1, double checking that the standard deviation of the distribution did not significantly change compared to the original dataset distribution (Kang, 2013).

As an example, Fig. 2 shows the time series from 2012 to 2020 of the weather station data, GNSS and radar data (lightning are not reported since provided as a binary information). To give an idea of how these data look like during an extreme event, we also report (Table 2) the absolute values for the extreme rain event on 15 May 2015 at 14:40.

2.2. Definition of extremes

The occurrence of an extreme weather event can be defined if its intensity exceeds a specified threshold. Such a threshold can be either connected to an absolute value or to a relative one such as a percentile value of a given set of data over a period of time. In the literature, there are different indices used to characterize the extremes (Myhre et al., 2019; Grazzini et al., 2020; Laurila et al., 2021). In this study extreme rainfall or wind speeds are defined as events in which the accumulated rain or mean wind speed exceeds the 95th percentile with respect to the reference. We define as reference the dataset created by using 14 stations (Bergamo, Biella, Busto Arsizio, Cavaria, Como, Gozzano, Luino, Milano, Molteno, Novara, Pavia, Lonate Pozzolo, Varese and Verbania) located within a radius of about 50 km from Malpensa airport for a period of 11 years (2010−2020). Only events with the accumulated rain or average wind speed value larger than 0 were accounted by extracting the percentile values of rainfall or wind speed. According to these rules, the threshold for extreme rainfall is computed as 1.6 mm of accumulated rain in 10 min (corresponding to a rain rate of 9.6 mm/h), and the threshold for the average wind speed is set to 5.7 m/s. The extremes from radar are already selected by the TRT algorithms according to the aforementioned threshold of 36 dBZ.

The threshold used to define extreme rainfall is also confirmed by the International Civil Aviation Organization (ICAO) draft Document 9837 - Manual on Automatic Observing Systems at Aerodromes (Chapter 6, 6.2.1.2) and the World Meteorological Organization (WMO) No. 8 (Annex to Chapter 14) reporting a value of 10 mm/h, which is very similar to our threshold of 9.6 mm/h. However, the extreme wind speed value is not available in the above-mentioned documents but the Canadian Avalanche Association defines strong wind speed higher than 11.11 m/s (Roeger et al., 2003) and WMO Volume 1 reports as reference the Beaufort scale setting the threshold of strong winds at 10.8 m/s. After a discussion with the ALARM project ATM stakeholders, we agreed that 10 m/s could be the best reference for our analyses.

3. Theory and methods

3.1. Methods

Weather is a continuous, data-intensive, multi-dimensional, complex dynamic, and non-linear process of the Earth's atmospheric system (Maqsood et al., 2003), and these properties make weather forecasting a challenging task in atmospheric science. Meteorologists and atmospheric scientists developed several methodologies and approaches

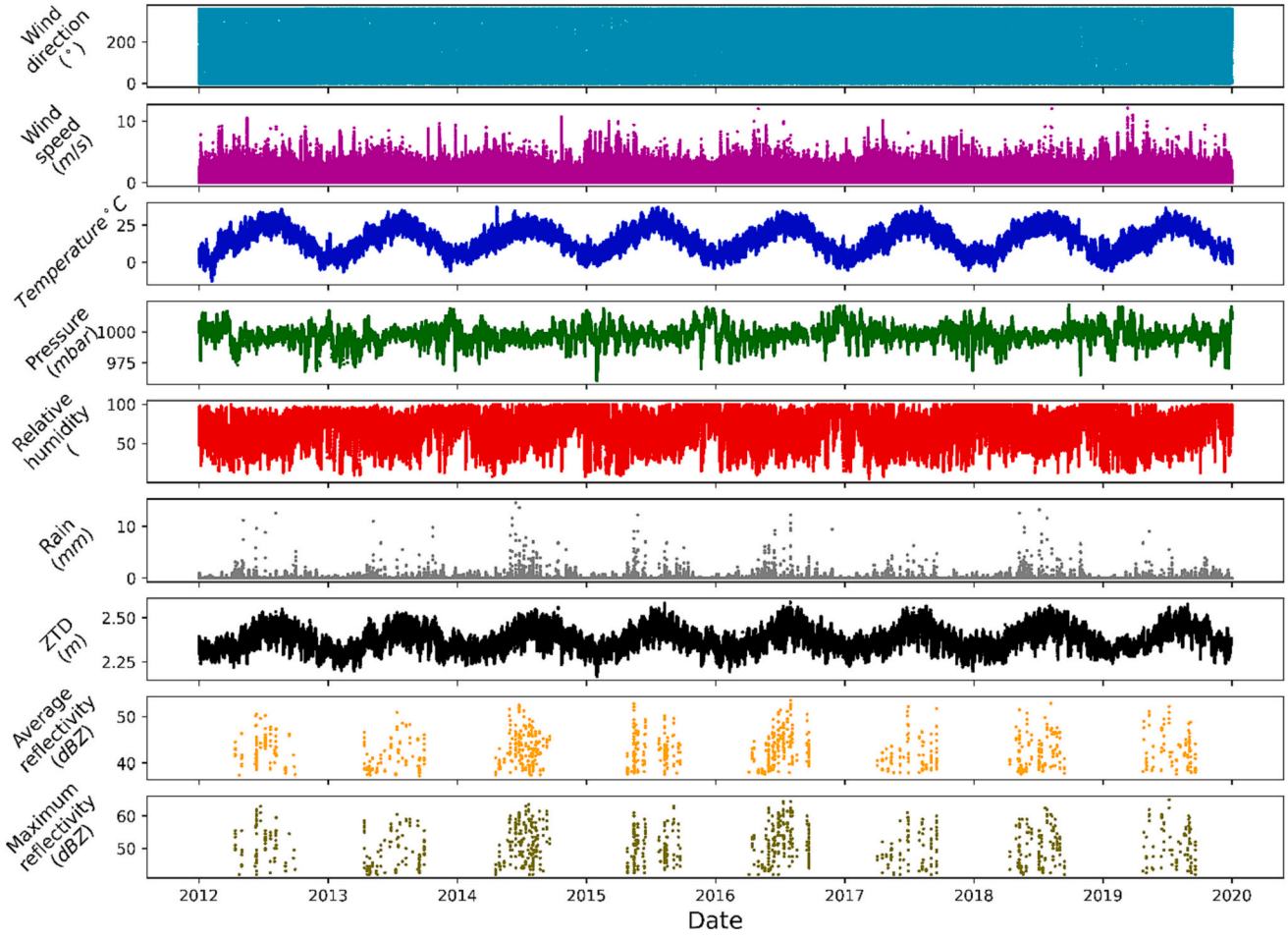

Fig. 2. Time series inputs of the nowcasting model from 2012 to 2020.

Table 2
Example of nowcasting model input dataset for an extreme rainfall event.

Date Time	Wind Direction [°]	Wind Speed [m/s]	Temp. [°C]	Press. [mbar]	Rel. Hum. [%]	Rain [mm/10 m]	ZTD [mm]	Avg dBz	Max dBz
15/05/2015 14:40	349.00	6.40	13.07	986.85	91.67	7.60	2.35	47.87	57.00
15/05/2015 14:50	351.00	6.40	12.63	987.05	92.33	5.20	2.35	47.72	57.00
15/05/2015 15:00	343.00	7.60	12.20	987.25	93.00	3.60	2.35	0.00	0.00
15/05/2015 15:10	347.00	5.80	12.07	987.45	93.67	1.80	2.36	44.27	56.00
15/05/2015 15:20	349.00	5.60	11.93	987.55	94.33	0.60	2.37	42.05	54.50
15/05/2015 15:30	358.00	6.20	11.80	987.75	95.00	2.40	2.36	42.38	53.50

based on physical and numerical simulations to predict the weather (Powers et al., 2017). Nevertheless, the developed models have a number of limitations, mainly, the incomplete understanding of the climate system, the imperfect ability to transform the acquired knowledge into accurate mathematical equations and the limited power of computers (Legates, 2002). Due to the learning ability from historical data of Artificial Neural Networks (ANNs) and machine learning approaches (Bochenek and Ustrnul, 2022), the field gained attention in the artificial intelligence community and data-driven approaches in weather forecasting started to be used. To understand the complex dynamics of convective cells in the atmosphere, temporal relationships must be considered to detect when we have heavy rainfalls, which is one of the meteorological conditions characterizing the convection. Forecasting a feature variable for multiple time steps in the future using typical neural networks is crucial to capture temporal dependencies. Thus, the so-called Recurrent Neural Networks (RNNs) are usually used for this purpose due to their adaptive nature with time-series data. Predictive models have recently been replaced by RNN-based approaches (Hong, 2008). The advent of deep learning techniques on feature representations and prediction tasks (Liu et al., 2014) increased the model performance. The ability to consider not only the initial atmospheric state prior to the prediction (single past value), but also the sequence of those events leading to the actual state, should provide more predictive power. In our framework, we are interested to nowcast rain and wind speed as target features related to convection development. For a target feature like the rain that is hard to be predicted, a well built and tested forecasting algorithm must be provided with careful attention to the pre-processing and post-processing steps corresponding to the prediction task. To accomplish this goal, we investigate, evaluate and compare three neural network-based models: a shallow and deep ANN as a baseline forecasting model, a Long Short-Term Memory (LSTM) model, and an Encoder-Decoder LSTM-based model (E/D LSTM).

3.1.1. ANNs

The concept of ANNs was first introduced by Mcculloch and Pitts (1943) as a computational model for biological neural networks. Such architectures provide a methodology for solving many types of non-linear problems that are difficult to be solved by traditional techniques. Most meteorological processes often exhibit temporal and spatial variability, with non-linear relationship between the different parameters. The ANNs have the capability to extract the existing relationship between the inputs and outputs of a process, without the physics being explicitly provided and that is why these algorithms are called data-driven models and are well suited for nowcasting tasks.

A single-layered ANN, with a single output, is known as the perceptron. Multi-Layer Perceptrons (MLPs) are the most commonly used architecture for ANNs. The composition of MLPs contains layers of neurons with an input layer, an output layer, and a hidden layer (at least one). The layers of the ANN are interlinked with each other by developing a multi-layered architecture, and this makes the model essentially complex for the ANN processing. Details on the ANN structure and architecture are reported in Appendix A.

3.1.2. Long Short-Term Memory (LSTM) Cell

The LSTM is a special kind of an RNN that is capable of learning the long-term temporal dependencies in sequential data more accurately than conventional RNNs and is powerful for modeling long-range dependencies (Ibrahim et al., 2020). A basic LSTM network of two-layered LSTM with fully connected hidden layers of 100 neurons already showed good performance in a Sequence to Sequence Weather forecasting (Zaytar and El Amrani, 2016). The LSTM networks are usually used as the baseline for weather long and short-term forecasting applications (Al Sadeque and Bui, 2020) getting incremental results such as the stacked LSTMs for temporal forecasting (Zaytar and El Amrani, 2016) and the spatio-temporal stacked LSTMs (Karevan and Suykens, 2018).

The main criterion of the LSTM network is the memory unit h_t, also called an internal unit, which can memorize the temporal state. Given a sequential input $x = \{x_1, x_2, \ldots, xt, xt+1, \ldots, x_n\}$, where n is the total number of samples, a LSTM cell can process x by using its memory unit, and in contrast to a conventional RNN, the LSTM cell is shaped by the addition or removal of information through three controlling gates: input gate i_t, output gate o_t, and forget gate f_t. Thanks to these gates, the LSTM overcomes the vanishing gradient problem unlike RNNs (Ibrahim et al., 2020). Moreover, it allows it to better handle long input sequences by updating and controlling the flow of the information in the block using these gates according to the equations (Ibrahim et al., 2020):

$$f_t = \sigma_g\left(w_f x_t + u_f h_{t-1} + b_f\right) \quad (1)$$

$$i_t = \sigma_g(w_i x_t + u_i h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma_g(w_o x_t + u_o h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \sigma_h(w_c x_t + u_c h_{t-1} + b_c) \quad (4)$$

$$h_t = o_t \otimes \sigma_h(c_t) \quad (5)$$

where x_t is the current passed input, h_t is the current hidden state, w, u and b stand for the weight matrices and biases, the nonlinear functions $\sigma_g(.)$ $\left(\sigma_g(x) = \frac{1}{1+e^{-x}}\right)$ and $\sigma_h(.)$ (tan_h) are the sigmoid and hyperbolic tangent functions and $\otimes$ denotes element-wise multiplication. Eqs. (1), (2) and (3) establish gate activations, Eq. (4) determines the new cell state c_t, where the 'memories' are stored or deleted, and Eq. (5) is the final output. The memory unit h_t defined in Eq. (5) refers to the hidden state vectors of the LSTM layer shown in Fig. 4. Each LSTM unit has a hidden state, where h_{t-1} represents the hidden state of the previous timestep and h_t is the hidden state of the current timestep. With the mechanism of the three gates, each LSTM unit computes the cell states represented by c_{t-1} and c_t for the previous and current time steps, respectively. The hidden state is known as short term memory and the cell state is known as long term memory of each LSTM unit in the LSTM layer.

3.1.3. LSTM encoder-decoder

A simple baseline LSTM network is sufficient to make single predictions based on the state of the atmosphere leading up to that time (i.e. 1 lead time step); however, our goal is to nowcast multiple lead time steps into the future and to provide predictions of the target features 1 h ahead with a time sampling of 10 min. We can accomplish this objective using dense LSTM networks at the expense of reduced model performance. We want a neural network architecture that can readily provide

accurate predictions arbitrarily far into the future. We can relate such a technique to a vector-based sequence to sequence learning problem, and implement E/D structure with more than one RNN/LSTM cell (Fig. 3). The E/D RNN/LSTM-based architecture (Sutskever et al., 2014) has become popular after its success in displacing classical phrase-based statistical machine translation systems for state-of-the-art results. In machine translation, the key idea is to encode the source sentence as a fixed-length vector and then use the decoder to generate a translation. In our work, we use the E/D LSTM approach for time series prediction instead of word translation. Deep stacked LSTMs can be applied to time series data as a sequence to sequence based encoder-decoder model. This takes a sequence of features as input (input sequence of n time steps), and outputs a target sequence (output sequence of multiple m time steps) as a continuation to the input target sequence. The dimensions n and m are not always equal; in our case, we have previous information about the past 2 h ($n = 12$) and we predict 1 h in the future (m = 6) as shown in Fig. 3.

The model for every recurrent node follows a standard LSTM structure with the gates to control the states of the memory cell. The additional tasks which compose the E/D structure are: the Repeat Vector layer, which repeats the final output vector from the encoding layer as a constant input to each timestep of the decoder, and the time Distributed Dense layer, which applies the same Dense (fully-connected) operation to every timestep of a 3D tensor.

3.2. Experiments

3.2.1. Feature representation and preparation

In this work, we have a multivariate driving series (13 weather features) from different sources: 6 features for weather stations (rain, relative humidity, pressure, temperature, wind speed and wind direction), 1 feature for GNSS (ZTD), 4 features for lightning (number of lightnings, cloud-to-cloud, cloud-to-ground, intra-average-height), and 2 features for radar (average and maximum dBZ). As explained previously, all the features are sampled at a 10-min rate. Since we aim to study the impact of each type of data source, as well as the significance of the availability of datasets (Table 1), on the prediction of extreme rains and extreme winds, the priority is given to the temporal information of the data with two changing factors: input configuration (i.e. what input features are fed to the model) and the time slot of available data. Google Colab and Jupyter Notebooks were selected as platforms for building and managing the datasets before and during the training.

A number of pre-processing steps are done on the raw data:

- Filling the missing feature values per data source. The time series of each feature is relatively long (availability between 4 years and 3 months, and 9 years) with some short temporal gaps. For our purposes, we decided to fill the lightning feature by adding zeros at the time steps in which the lightning is not present. We used the same filling method for the radar feature in all the instances where we did not have any reflectivity information (where reflectivity was lower than 36 dBZ). Finally, we use straight-forward interpolation or machine learning regression methods to fill the gaps of weather parameters, radar and GNSS. Among the tested methods, the Stochastic Random Regression (SRR) (Enders, 2010) has shown to provide more consistent values than the linear/average interpolation. The SRR relies on the use of the function 'random imputation' which replaces missing values with some random observed values of the variable and it is repeated iteratively for all the variables containing missing values. In this way, they can serve as parameters in the linear regression model in order to estimate other variable values.

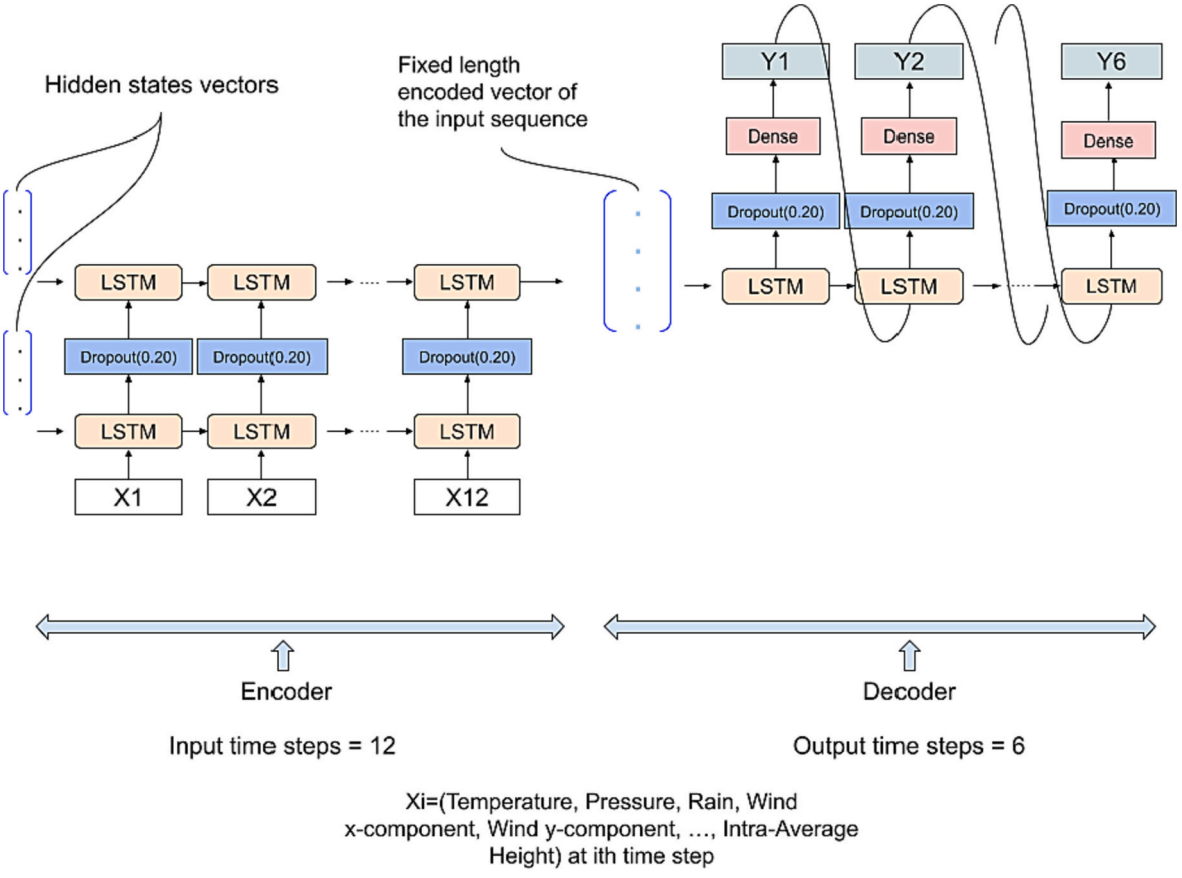

Fig. 3. Encoder-decoder LSTM-based model diagram.

- Wind speed and direction (available in degrees) features were reformulated into a wind vector of x and y components because the distribution of wind vectors is simpler for the model to be correctly interpreted. With the wind speed and wind direction, we can obtain the component vector winds, u and v, as follows:

$$u = ws^{*} \, cos(\theta) \tag{6}$$

$$v = ws^{*} \, sin(\theta) \tag{7}$$

where ws represents the wind speed information and θ is the wind direction in radians.

- The time is also an additional feature (besides other weather parameters) to improve the performance of the model because it provides information on the periodicity of the parameter, so we generate time signals (day sin, day cos, year sin, year cos), deriving a sine transform and cosine transform of the date-time, adding 4 new features to the dataset.

Our models perform best when the inputs and outputs are stationary. By definition, a time series is non-stationary when its observations depend on time, because, in this case, the statistics (mean and variance) also change with the time. This is a challenging problem for machine learning techniques since they cannot guarantee the development of reliable prediction models with not stationary features (Livieris et al., 2021). Thus, we devote some time to study the stationarity for each feature set using the commonly used Dickey-Fuller and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) statistical tests (Gagniuc, 2017). Based on these tests, the stationary features are rain, day sine and day cosine, while relative humidity, temperature, pressure, wind vector components, ZTD, lightning, and radar are non-stationary or weakly stationary. It is possible to transform the non-stationary features into stationary features by computing the differences between the original features and the same features lagged by a specified period (Livieris et al., 2021):

$$\text{difference}(t) = \text{observation}(t) - \text{observation}(t - T) \tag{8}$$

where $T = 1$ lag period. In our specific case, T is the seasonality of the features which is the frequency repeating the pattern.

The stationary features are given in input to the model unchanged, the non-stationary features are first transformed in differences (stationary) and then given in input to the model in this form. When the model provides the outputs, the rain is provided in the right format (because it is stationary) while the wind is provided in difference of wind vector components format, so we must get the original value inverting the Eq. (8) and then recompute the wind speed from Eq. (6) and Eq. (7).

3.2.2. Model preparation

When the time series of features are pre-processed and prepared, we need to transform them into a format correctly interpreted by the model. The transformation process consists of 3 stages:

1) *Data split*. The chosen weather data is divided into three selected datasets corresponding to the train, validation and test sets (70%, 20%, 10%, respectively).
2) *Scaling*. All the data are rescaled to values between 0 and 1 using Min-Max normalization defined as:

$$x_{norm} = \frac{x - min(x)}{max(x) - min(x)} \tag{9}$$

where x is the feature variable to rescale. The normalization is performed on each feature separately.

3) *Supervised problem*. The time series dataset is composed by all the meteorological features mentioned in the dataset section. For the forecasting machine learning model we transform the time series data into two subsets: the input observations of the previous 2 h (X) and the target observations of the next 1 h (Y). This is achieved using a window generator capable of:
   - Generating windows with input and target features from the time series data, specifying the targets we want to predict;
   - Splitting windows into pairs, input feature windows and target feature windows;
   - Creating batches of window pairs for each of test, training, and validation datasets;
   - Creating the input feature vector with input width of 12 time-steps (lag observations);
   - Defining the targets to be predicted as the rain, the x and y wind components of the generated wind vector with target width of 6 time-steps in the future;
   - Shifting the parameter by 6 time-steps.

The windows generated are validated by using a split function and we compute X, and Y matrices for each of the training, test, and validation dataset as shown in Fig. 4.

X and Y are represented in the form of 3D tensors for the input feature and target/output feature, respectively:

- $X \in R^{m \times l \times n}$, each sample observation $Xj = \{X1, X2, \ldots, X12\}$ and each $Xi = \{x1, x2, \ldots, xn\}$ *with* $j = (1, 2, \ldots, m)$ *and* $i = (1, 2, \ldots, l = 12)$, where m is the total number of samples, l is the past time steps, and n is the total number of input features and its variable depending on the input configurations.
- $Y \in R^{m \times \acute{l} \times 3}$, each sample observation $Yj = \{Y1, Y2, \ldots, Y6\}$ and each $Yi = \{y1, y2, y3\}$ *with* $j = (1, 2, \ldots, m)$ *and* $i = (1, 2, \ldots, \acute{l} = 6)$, where m is the total number of samples, $\acute{l}$ is the future time steps for 3 target features of the rain and wind vector components.

3.2.3. Model post-processing

When the nowcast is obtained, the output is post-processed following the steps below:

- The nowcasting is performed by using the differenced test data (Eq. (8));
- The nowcasts are denormalized to return them into their original scale;
- The difference values are inverted to the original test data.

After these three steps, the wind speed values are consistent with the observations; however, the rain values are largely underestimated because most of the rain values are zero or close to zero. This issue is not new and there are several approaches towards eliminating it, including artificial neural networks, machine learning and statistical methods (Hemri et al., 2014; Jeon et al., 2016; Li et al., 2019). To improve the nowcasting of the rain, we therefore apply a post-processing procedure, which is described in Section 4.2.

3.2.4. Model training and evaluation

We trained the three models, MLP network as a forecasting baseline, baseline LSTM, and E/D LSTM architectures, by using the settings provided in Table 3. As a loss function, we use the Mean-Square Error (MSE), and as evaluation metrics, we use MSE, Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE), to compare the accuracy level and the performance of each model architecture trained:

$$MSE = \frac{1}{N} \Sigma_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{10}$$

$$MAPE = \frac{100\%}{N} \Sigma_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{11}$$

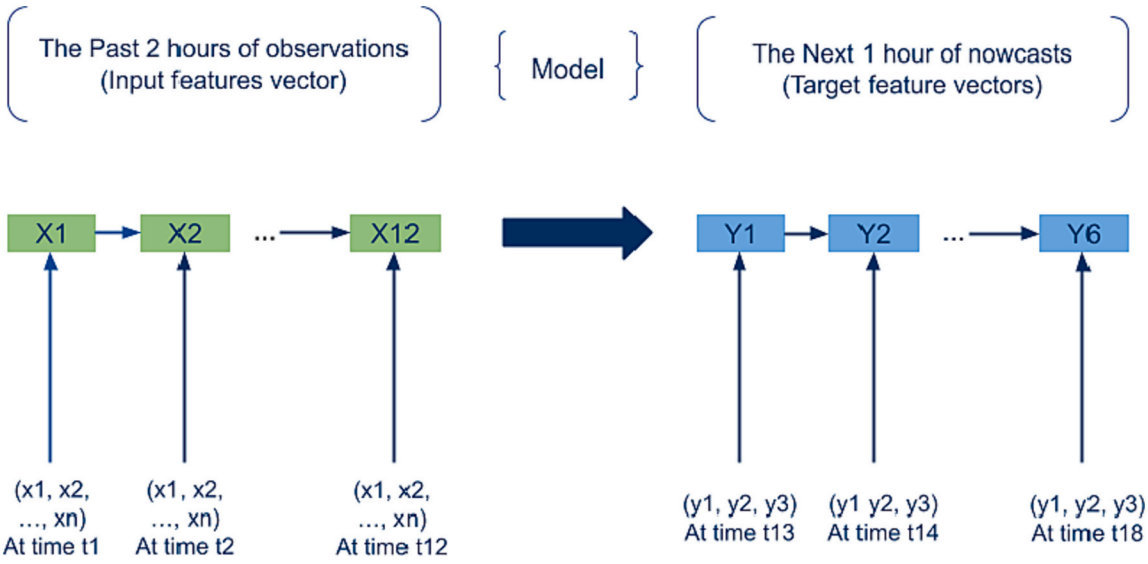

Fig. 4. Input/Output data flow of each observation.

Table 3
Model Setting for the training, and 5 repeated experiments on each network architecture.

Epochs	150 training iterations
Batch size	256
Optimization technique	Adaptive moment estimation (Adam) optimizer with a learning rate = 0.01
Network weight initialization	"glorot" or "xavier" uniform initialization
Dense layers	Activations = "relu" number of neuron = 32
LSTM layers	Activations: hyperbolic tangent (tanh) function Shuffle = False (order is respected for time series forecasting purpose) lstm units = 32
Regularization type	Dropout = 0.15 (15%)

$$\mathrm{MAE} = \frac{1}{N}\Sigma_{i=1}^{N}|y_i - \widehat{y}_i| \tag{12}$$

where N is the number of samples, y_i is the i^{th} true observation, and $\widehat{y}_i$ is the i^{th} predicted observation.

In order to choose the best parameter setting with the lowest validation error, we use the early stopping technique, which allows us to select the model associated with the parameter setting with minimum validation error (Fathi and Maleki Shoja, 2018). The evaluation is performed in both the training and validation datasets. We conduct three experiments for each of the 3 model architectures independently, using the diagnostic approach defined by Brownlee (2014), and each experimental scenario is run 5 times because the random initial conditions for an LSTM network can be very different any time a configuration is trained. The comparison of the models is based on the training and validation dataset of all the features (weather station parameters, GNSS ZTD, radar dBZ and lightning parameters). To compare the experiments, we use the Root-Mean-Square Error (RMSE) as a metric, which is defined as: $RMSE = \sqrt{MSE}$. Each model is then evaluated at the end of each epoch, and the RMSE scores are saved. We ran three different models (Table 4) with the same input/output configuration to evaluate their performances and choose the best model to be kept for the future analysis. We used all the parameters (weather stations, GNSS, radar, and lightning) as input, and rain and wind speed as output. All the metrics (MSE, MAPE, MAE) show that the E/D LSTM method converges more rapidly and to lower values than to the other models. In Table 4 we report the RMSE for the training and validation datasets, where we can see that the best performances are given by E/D LSTM model except for the rain validation dataset; however, the RMSE for the E/D LSTM (0.182 mm) is really close to the lowest value of the LSTM (0.179 mm). For these reasons, we will show all the results obtained with the E/D LSTM model.

Table 4
Average RMSE for rain and wind speed nowcasting with three different tested models after optimization: MLP, LSTM and E/D LSTM on Novara station data (all features).

Model architecture	Rain Train	Rain Val	Wind Train	Wind Val
MLP	0.150	0.193	0.134	0.140
LSTM	0.135	**0.179**	0.092	0.096
E/D LSTM	**0.129**	0.182	**0.078**	**0.081**

The optimization of the E/D LSTM model is done on the training and validation datasets, where 10% of training data serves as a validation set, and the validation dataset is used for each hyperparameter in order to choose the optimal value. The test dataset is used as new data to evaluate the optimal configuration and generalize the model configuration (Tibshirani and Friedman, 2017). A diagnostic approach is used to investigate the running performances by computing the RMSE over time (epochs) and assessing possible improvements. The other hyperparameters are investigated using a box and whisker plot (Appendix B) of the computed error scores of the validation set on each target (Brownlee, 2017). The tuning, using box plots, is a tradeoff of average performance and variability of that performance, with an ideal result having a low mean error with low variability, meaning that it is good and reproducible. In Table 5 we report in bold the hyperparameters used in this analysis.

Table 5
Optimized Hyperparameters of the E/D LSTM model.

Optimized Hyperparameters of the E/D LSTM model hyperparameters using the same diagnostic approach	
Number of training epochs	{100, **110**, 150, 200, 500}
Dropout	{10%, 15%, **20%**}
Batch size	{64, 128, **256**}
Number of neurons for Dense layers	{32, **64**, 128, 256}
Number of LSTM units for LSTM layers	{18, 32, **64**, 128}

3.2.5. Nowcasting performance assessment

To assess the performance of the nowcasting model, we use 3 different metrics commonly used in the relevant literature (Zhang et al., 2021). These metrics are defined based on the True Positive (TP), False Positive (FP) True Negative (TN), and False Negative (FN), which denote the successfulness of a nowcasting model in relation to an observed event. An event can either refer to a rain/wind event (an instance of non-zero rain/wind value) or to a rain/wind extreme event (an instance of rain/wind value above the extreme threshold). Table 6 shows the relationship between observed and nowcasted events. We define:

- TP as the number of observed events that were successfully nowcasted by the model;
- FP as the number of non-observed events that were nowcasted by the model;
- FN as the number of observed events that were not predicted;
- TN the number of non-observed events that were correctly not predicted.

Based on Table 6, we can define the assessment metrics, which are categorical statistics with values in the range 0–1, but we prefer to report them in percentage (therefore multiplied by a factor of 100):

- The Probability of Detection (POD) is the percentage of observed events that were successfully nowcasted. Therefore, the closer the value is to 100, the better it is.

$$POD = \frac{TP}{TP + FN} \bullet 100$$

- The False Alarm Rate (FAR) is the percentage of nowcasted events which did not happen in reality. The closer this value is to 0, the better it is.

$$FAR = \frac{FP}{TP + FP} \bullet 100$$

- The Critical Success Index (CSI) is the percentage of correctly predicted events to the total number of positive events nowcasted or needed. The best score for this metric is 100 and the worst is 0.

$$CSI = \frac{TP}{TP + FP + FN} \bullet 100$$

4. Results

We run the E/D LSTM model with 4 different configurations:

- Version 1 (v1) configuration includes the weather station data (WS) and the GNSS data (GNSS) with an availability of 9 years (2012−2020);
- Version 2 (v2) configuration includes the WS, the GNSS and the lightning data (LIGH) with an availability of 4 years and 3 months (2015–2019);
- Version 3 (v3) configuration includes the WS, the GNSS and the radar data (RADAR) with an availability of 8 years (2012–2019);
- Version 4 (v4) configuration includes the WS, the GNSS, the RADAR data and the LIGH with an availability of 4 years and 3 months (2015–2019).

The performances of the model depend on the input features used and on the length of the training dataset, so a real comparison between the configurations can be obtained only considering a common period of data availability. Table 7 reports the RMSE obtained with the E/D LSTM with the 4 configurations (within the respective period available) and Table 8 reports the RMSE for the same configurations, but training the model in the common period of 4 years and 3 months. The comparison of Table 7 with Table 8 shows that the shortest data availability does not impact much on the model performances, in fact the RMSE for the rain training in v1 (v3) is even lower when we decrease the period from 9 (8) to 4.25 years, the rain test RMSE increases by 11% (7%), the wind training RMSE increases by 6% (improves) and the wind test RMSE increases by 8% (improves). In general, the performances to nowcast the rain are steady when changing the input configuration but we lower the RMSE by 12% when we add the RADAR and LIGH to the WS and GNSS.

Ideally, a larger dataset should improve a machine learning model performance, but just in case the dataset well represents all the possible combinations and variability of the inputs. In our case, the 4.25-year dataset is more homogeneous than the 9-year one with a smaller variability of the input parameters, thus the model provides better performances with a smaller dataset available.

4.1. Wind speed nowcasting

The nowcasting of wind speed performs very well providing wind speeds with an RMSE of 0.081 m/s (Table 7-8), with 100% POD and just 4.3% FAR in all configurations at all time-steps (Table 9). When we focus on the extreme wind speeds (exceeding 5.7 m/s) (Table 10), the lowest CSI is achieved with v1 at 10-min nowcasting with 83.2% POD and just 1.3% FAR, then the CSI increases with the lead time steps of the nowcasting and with the addition of features to the model input. It is interesting to notice that the improvement of the CSI is due to higher values of the POD, while the FAR are always lower than 1.8%. All the CSI for nowcasting wind speeds between 30 min to 1 h are higher than 90% reaching the best performance of 95.8%. When we focus on the extreme wind speeds according to ATM stakeholders' requirements (exceeding 10 m/s) (Table 11), the lowest CSI is achieved with v1 at 10-min nowcasting with 83.3% POD and 0% FAR, then the CSI increases with the lead time steps apart from the 60 min lead time, where under all configurations there is a drop of CSI due to the fact that FAR increases. It is interesting to notice that the improvement of the CSI from 10 to 50 min lead time is due to higher values of the POD, while the FAR are always 0%. However, we should note that by setting the extreme wind speed to 10 m/s, we are statistically closer to the 99th percentile of the non-zero

Table 6
Performance definition according to the correct/false observation/nowcast.

Event Nowcast	Event observed		
	Yes	No	
Yes	TP	FP	Nowcast Yes (TP + FP)
No	FN	TN	Nowcast No (FN + TN)
	Observation Yes (TP + FN)	Observation No (FP + TN)	

Table 7
Average RMSE for rain and wind speed nowcasting with 4 input configurations of the total features available within the time slots provided by the WS, GNSS, RADAR and LIGH.

Input Configuration	Period	Rain Train	Rain Test	Wind Train	Wind Test
(v1) WS + GNSS	9 years	0.157	0.105	0.084	0.085
(v2) WS + GNSS+LIGH	4.25 years	0.133	0.117	0.086	0.089
(v3) WS + GNSS+RADAR	8 years	0.146	0.110	0.086	0.088
(v4) WS + GNSS+RADAR+LIGH	4.25 years	0.129	0.118	0.078	0.081

Table 8
Average RMSE for rain and wind speed nowcasting of the 4 input configurations within the common time slot available when all features are available (4 years and 3 months).

Input Configuration	Rain Train	Rain Test	Wind Train	Wind Test
(v1) WS + GNSS	0.136	0.118	0.089	0.092
(v2) WS + GNSS+LIGH	0.133	0.117	0.086	0.089
(v3) WS + GNSS+RADAR	0.134	0.118	0.083	0.085
(v4) WS + GNSS+RADAR+LIGH	0.129	0.118	0.078	0.081

wind speed data instead of the 95th percentile. In the observation dataset, 190 events overpass the wind speed threshold of 5.7 m/s while just 12 events overpass the threshold of 10 m/s. In the nowcasted data with threshold 10 m/s, the number of extreme events varies between 10 and 14 with the lead time.

Fig. 5 resumes the CSI of the wind speed nowcasting model for the 4 different configurations and with different data availability or selection (extremes). In Fig. 6, the left panel shows the comparison between the observed extreme winds and the nowcasted winds in v4 at 10 and 60 min; the right panel shows the corresponding scatter plot, where it is clear that the correlation is higher for the 60 min (1) than for the 10 min (0.97).

4.2. Rain nowcasting

Fig. 7 shows the observed rain values (gray) and the corresponding nowcasting (green). It is clear that the nowcasting has a trend similar to the observations, but there is a bias between the two, which depends on the intensity of the rain. The similarity on the trends decreases with the lead time step, higher at 10 min (correlation 0.95) and lower at 60 min (correlation 0.63) as shown in Table 12. For this reason, we decide to apply a post-processing algorithm to improve the results (Fig. 7, orange line). It is important to note that we are not interested only in correcting the nowcasting outcomes, but also in using the correction technique as a tool to correct the future predictions, therefore, we use the observed data only to make our algorithm learn and extract corrective parameters. The technique is applied to all the instances of common non-zero observed and nowcasted. The nowcasted data are split into groups based on their percentile's magnitude (Wang et al., 2016; Li et al., 2019) with a separation step of 5% from 0 to 95 and 2% from 95 to 100 in order to catch the biggest possible level of detail for the rarest events. As a next step, a linear regression is implemented between the nowcasted data and the difference between the observed and nowcasted data (error). The slope and intercept are used to retrieve the corrected value in each group by using the equation:

$$nc_i = n_i + n_i \bullet s_i + ic_i \tag{13}$$

where nc is the nowcasted corrected value, n is the original nowcasted value, s is the slope, ic is the intercept and i refers to the respective percentile group.

The performance of the rain nowcasting is reduced compared to the wind speed as we could expect from the literature (Khosravi et al., 2018; Zhang et al., 2021). From the nowcasts, we find that the model underestimates the rainfall values (Fig. 7) because the number of training samples is small for higher rain values and so the neural network is more biased towards the prediction of lower values. This is addressed as a general problem with the unbalanced dataset in deep learning-based techniques (Wang et al., 2016). All of the three developed models, however, estimate the wind speed accurately from past rain data used in the input configurations. Therefore, the RMSE for the rain nowcasting is 0.118 mm (Table 7-8). The performance improves when we decrease the period of v1 availability from 9 years (Table 13) to 4.25 (Table 14), while it is the opposite for v3. The best performances are provided by the v4 with 77.1% CSI at 10 min corresponding to 89.5% POD and 15.2% FAR, but also the configuration v1 is surprisingly providing high

Table 9
Model performances for nowcasting wind speed using all the data available from each source.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
All data	100	4.9	95.1	100	4.3	95.7	100	4.7	95.3	100	4.3	95.7
common data	100	4.3	95.7	100	4.3	95.7	100	4.3	95.7	100	4.3	95.7

Table 10
Model performances for nowcasting extreme wind speed using the common period of availability and 5.7 m/s as threshold.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
10 m	83.2	1.3	82.3	84.4	1.8	83.1	85.8	0.6	85.3	86.3	1.8	85.0
20 m	92.1	1.7	90.7	89.6	1.1	88.7	90.5	1.7	89.1	91.1	1.7	89.6
30 m	94.7	1.6	93.3	93.2	0.6	92.7	94.7	0.6	94.2	94.2	0.6	93.7
40 m	95.3	1.1	94.3	93.8	0.6	93.3	95.3	0.5	94.8	95.3	0.5	94.8
50 m	96.8	1.1	95.8	92.7	0.6	92.2	96.3	1.1	95.3	95.8	1.1	94.8
60 m	94.7	1.6	93.3	93.2	1.1	92.3	95.8	2.2	93.8	94.2	0.6	93.7

Table 11
Model performances for nowcasting extreme wind speed using the common period of availability and 10 m/s as threshold.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
10 m	83.3	0.0	83.3	83.3	0.0	83.3	100.0	0.0	100.0	100.0	0.0	100.0
20 m	83.3	0.0	83.3	91.7	0.0	91.7	91.7	0.0	91.7	91.7	0.0	91.7
30 m	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0
40 m	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0
50 m	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0
60 m	100.0	14.3	85.7	100.0	7.7	92.3	100.0	7.7	92.3	100.0	7.7	92.3

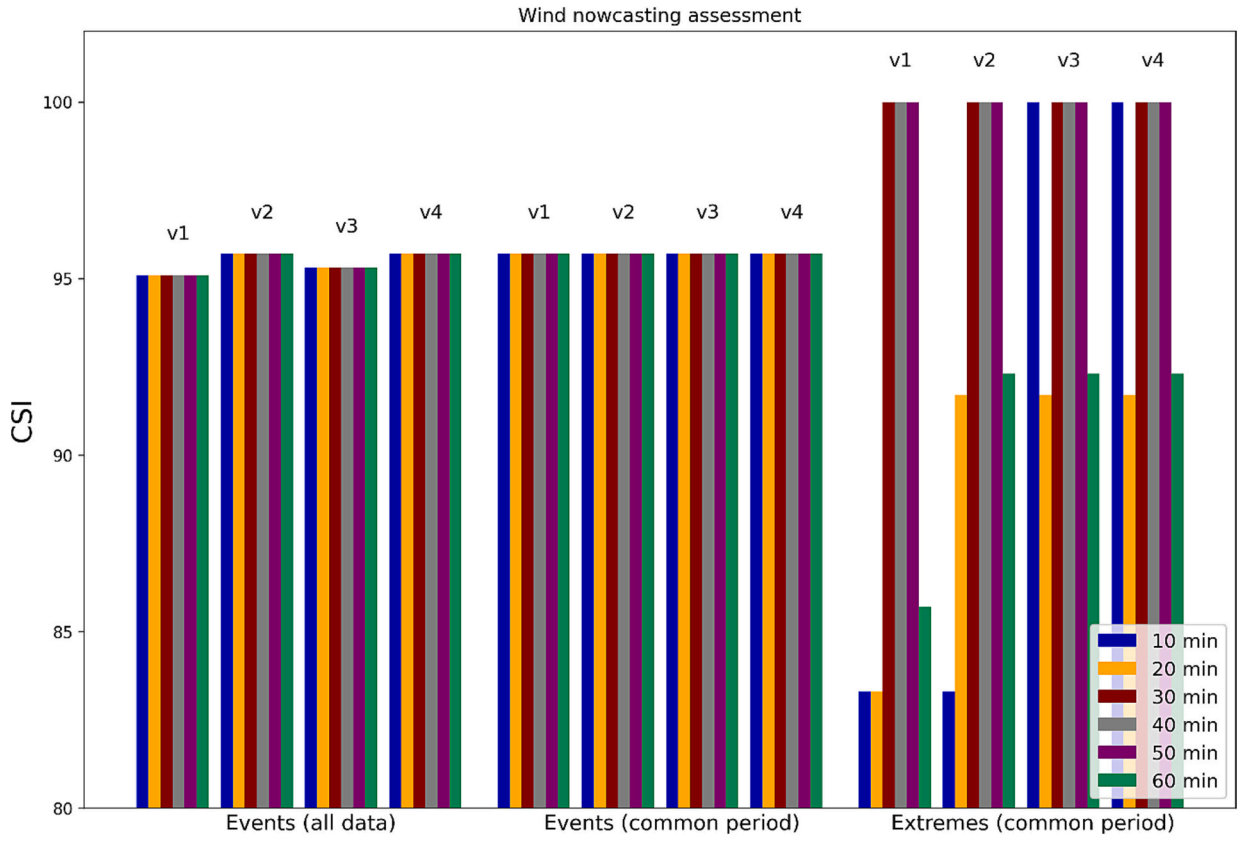

Fig. 5. Percentage Critical Success Index (CSI) of the wind speed nowcasting model for different configurations and datasets availability.

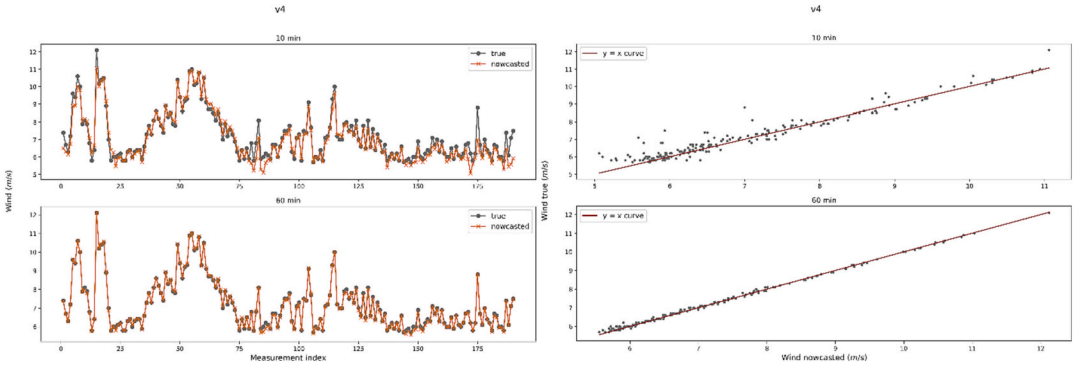

Fig. 6. Observed extreme winds (blue) and the nowcasted winds (orange) in v4 at 10 (top) and 60 (bottom) minutes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

performance at all the lead time steps with CSI ranging from 72.5% at a 10-min time step to 75.5% at a 50-min time step. However, the v1 provides much less FAR than the v4 and similar values of CSI (Table 14).

The nowcasting of extremes (Table 15) provides high values of POD at 10- and 20-min time steps but also the FAR have high values bringing to a non-performant CSI for lead times larger than 30 min. The best performance for the detection of the rain extremes is achieved by using the simplest configuration v1 including just the weather stations and the GNSS ZTD. In general, we can say that the radar and lightning data contribute to improve the POD in rain nowcasting but they also largely increase the FAR especially when we focus on extreme events.

Fig. 8 resumes the CSI of the rain nowcasting model for the 4 different configurations and with different data availability or selection (extremes).

5. Discussion and conclusions

The machine learning techniques for meteorological purposes have recently been applied with good results. The assimilation of weather radar data to 4-DVAR in WRF has shown an improvement of 30% in short-lived precipitation nowcasting (Thiruvengadam et al., 2020), and, in general, the assimilation of new and different types of datasets contributes to better severe weather nowcasting (Kohn et al., 2011; Benevides et al., 2019; Zhang et al., 2021). The use of machine learning techniques has also improved the nowcasting performances. There have been several attempts to provide inputs to machine learning models over a long time series of different datasets, which led to increased performances in rainfall prediction (Benevides et al., 2019: Ayzel et al., 2020; Hewage et al., 2021). In this work, we use the standard weather stations data, radar and lightning data already widely presented in the literature

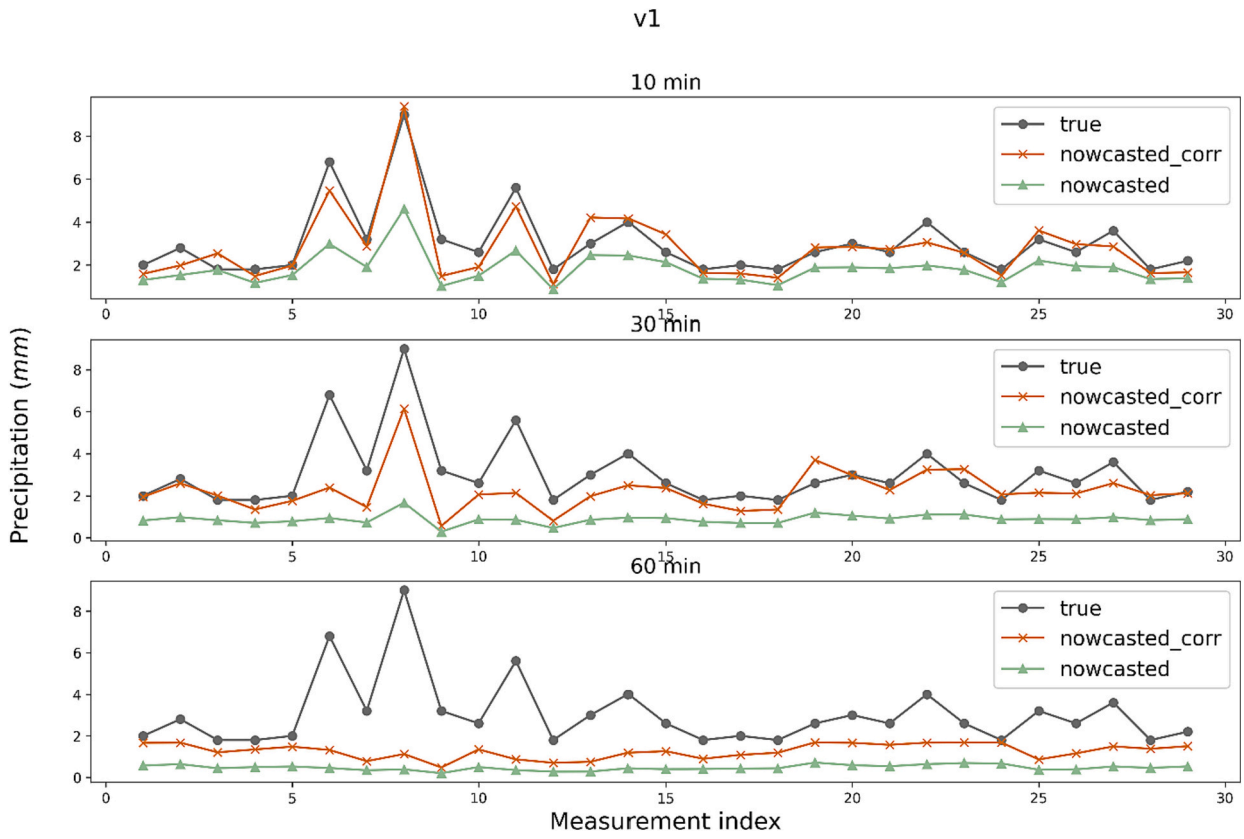

Fig. 7. Observed extreme rain (gray), the nowcast rain (green) and corrected nowcast (orange) in v1 at 10 (top), 30 (middle) and 60 (bottom) minutes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 12
RMSE and correlation (Corr) between the observations and the rain post-processed nowcasting.

	WS + GNSS		WS + GNSS+LIGH		WS + GNSS+RADAR		WS + GNSS+RADAR+LIGH	
	RMSE	Corr	RMSE	Corr	RMSE	Corr	RMSE	Corr
10 m	0.31	0.95	0.57	0.95	0.32	0.93	0.42	0.91
20 m	0.40	0.90	0.67	0.88	0.41	0.87	0.58	0.79
30 m	0.46	0.81	0.77	0.69	0.47	0.80	0.67	0.64
40 m	0.50	0.72	0.83	0.44	0.51	0.73	0.71	0.51
50 m	0.52	0.66	0.69	0.45	0.53	0.69	0.73	0.39
60 m	0.54	0.63	0.61	0.44	0.55	0.68	0.73	0.36

Table 13
Model performances for nowcasting rain using all the data available from each source.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
10 m	62.5	3.3	61.2	60.0	15.4	54.0	72.3	7.0	68.6	89.5	15.2	77.1
20 m	68.0	6.6	64.9	61.0	17.3	54.1	74.9	6.4	71.3	87.5	14.8	76.0
30 m	69.8	8.5	65.6	63.4	16.9	56.1	74.4	6.0	71.0	85.4	14.8	74.3
40 m	70.3	9.6	65.4	64.5	16.1	57.3	72.9	5.2	70.1	84.4	13.5	74.6
50 m	70.5	10.3	65.3	63.6	16.1	56.6	72.1	4.8	69.6	83.7	12.8	74.6
60 m	70.4	10.8	64.9	62.8	16.3	56.0	71.3	4.9	68.7	82.5	13.1	73.3

(Franch et al., 2020; Klocek et al., 2021; Zhang et al., 2021) to feed the LSTM model in order to optimize it. We also want to highlight the relevance that GNSS data have recently shown for nowcasting applications (Benevides et al., 2019; Łoś et al., 2020; Guo et al., 2021). The final objective of this study is to create a reliable end-to-end nowcasting model using a LSTM-based model able to nowcast extreme rain and extreme wind speed of weather phenomena quickly developing at small spatial scale near airports. For this reason, we focus on the Malpensa airport area, chosen as a hotspot of the project ALARM, around which we have the availability of 14 weather stations, 8 GNSS stations and the radar reflectivity. The model has been first tested on a single location, Novara, because additionally to the weather station, GNSS and radar, we also have the availability of lightning detectors. The model tested and optimized in Novara will be applied to the entire area of Malpensa and, later, to another hotspot of the project, Brussels Zaventem airport. The number of extreme wind speed events, with extreme threshold defined

Table 14
Model performances for nowcasting rain using the common period of availability.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
10 m	77.2	7.8	72.5	60.0	15.4	54.0	51.0	10.7	48.0	89.5	15.2	77.1
20 m	79.0	9.1	73.2	61.0	17.3	54.1	53.0	10.4	49.9	87.5	14.8	76.0
30 m	79.0	8.6	73.6	63.4	16.9	56.1	52.4	11.6	49.1	85.4	14.8	74.3
40 m	78.7	6.9	74.4	64.5	16.1	57.3	52.6	9.3	49.9	84.4	13.5	74.6
50 m	78.5	4.8	75.5	63.6	16.1	56.6	52.6	8.2	50.3	83.7	12.8	74.6
60 m	77.8	4.7	74.9	62.8	16.3	56.0	52.8	7.9	50.5	82.5	13.1	73.3

Table 15
Model performances for nowcasting extreme rain using the common period of availability.

	(v1)			(v2)			(v3)			(v4)		
	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI	POD	FAR	CSI
10 m	87.5	12.5	77.8	87.5	6.7	82.4	87.5	6.7	82.4	87.5	26.3	66.7
20 m	81.3	7.1	76.5	75.0	25.0	60.0	93.8	21.1	75.0	81.3	35.0	56.5
30 m	87.5	12.5	77.8	68.8	21.4	57.9	75.0	7.7	70.6	68.8	35.3	50.0
40 m	87.5	50.0	46.7	50.0	20.0	44.4	81.3	7.1	76.5	62.5	37.5	45.5
50 m	75.0	55.6	38.7	37.5	33.3	31.6	75.0	60.0	35.3	37.5	40.0	30.0
60 m	75.0	36.8	52.2	6.3	50.0	5.9	62.5	63.0	30.3	25.0	50.0	20.0

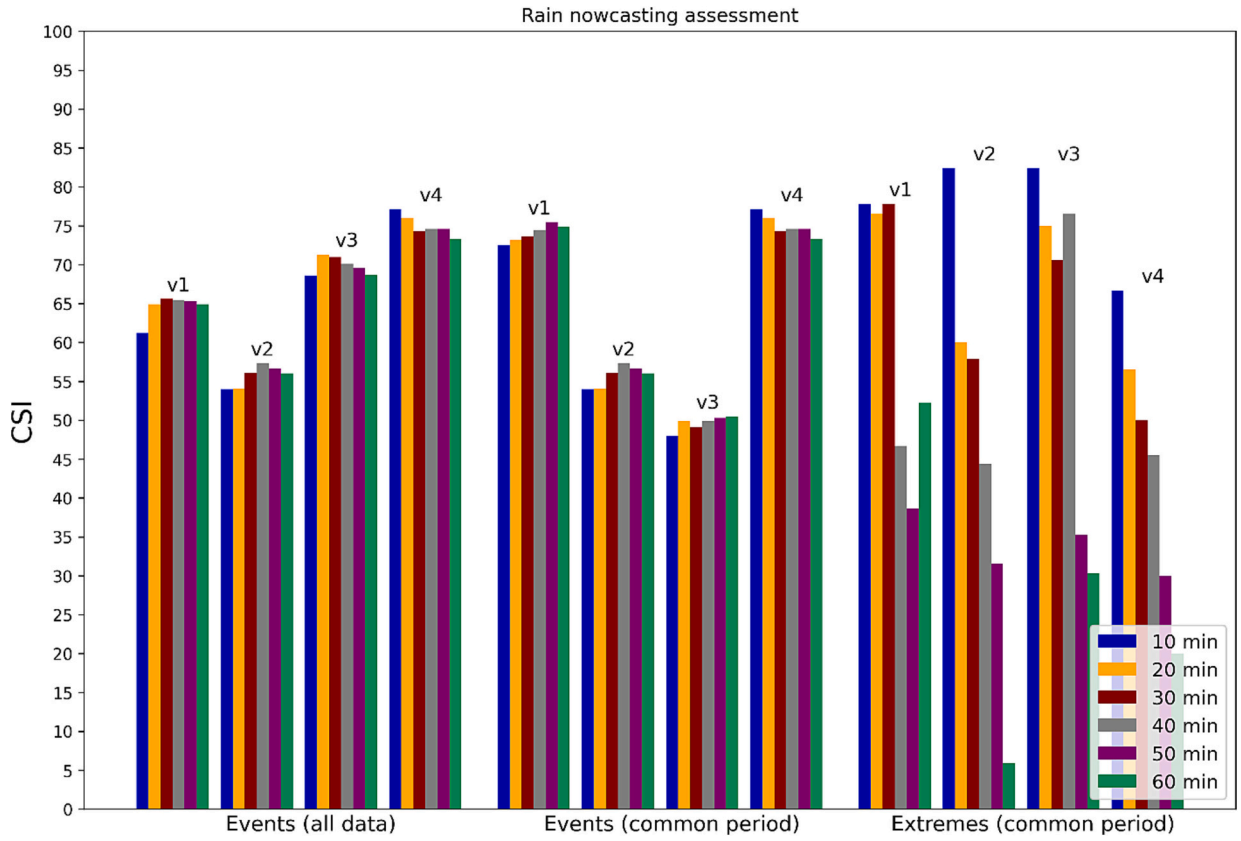

Fig. 8. Percentage Critical Success Index (CSI) of the rain nowcasting model for different configurations and datasets availability.

by the stakeholders (10 m/s), is not statistically significant (just 10–14 cases depending on the lead time); however, it is an important achievement for the scope of the project because it shows that the model is able to nowcast rare events strongly impacting the aviation and providing a real contribution to the ATM. This work shows that the algorithm has excellent performances, a relevant statistic will be obtained in the near future studies applying the model to the 14 stations around Milano Malpensa.

The E/D LSTM model has been chosen within the different tested ones since it provided the best performances in our case. The performances of the model are very good with a RMSE of 0.118 mm for the rain nowcasting and a RMSE between 0.081 and 0.092 m/s for the wind speed. However, while the detection for the extreme wind absolute value is highly stable, providing CSI higher than 80% with the 10-min lead time and higher than 90% with the lead time between 20 and 60 min, the model to nowcast the extreme rain absolute value has still some issues which must be solved applying post-processing corrections.

We want to highlight that radar reflectivity and lightnings are the

"signature" of severe weather already well developed, and they can be used to understand if the phenomena are going to intensify or weaken. The weather stations data and the GNSS ZTD can be used for the same purpose, but also and specially to understand if severe weather is going to form or not. In this regard, the capabilities of the v1 model configuration can be considered more relevant. The configuration WS + GNSS "captures" the quickly developing events since it is sensitive to the pre-convection environment, while the radar and lightning data provide information just when the severe event is already developed, as explained in Sections 2.1.3 and 2.1.4. This is clear from Table 10 (the common period for extreme events), where the performance of v3 remarkably increases. The combined use of different datasets does not always lead to better performances. In our specific case, we must consider that the addition of lightning and radar means a shorter time period of availability but also a more homogeneous dataset.

Our results in some cases show performances improving with the lead time. This is true for the wind speed in general (even though the FAR are present just for the 60-min lead time), but not for the rain where the performances increase with the lead time only using the whole dataset available with configurations v1 and v2. In case of extreme rain with the common data period availability, all the configurations show decreasing performances with lead time.

The reasons for this behavior can mainly be three:

- The model is not optimized according to the POD/FAR/CSI performances, but according to the RMSE, which always decreases with the lead time;
- We nowcast two parameters at the same time (rain and wind speed), and we must find a balance in the optimization of the model, which is not the best if they are considered individually;
- There is no connection between the thresholds which we decided to use for determining the POD and FAR and the nowcasting model.

With this study, we pose the basis to create a future operational product for the ATM. The model developed and optimized for Novara will be run in other 13 stations around Milano Malpensa to create a polygon showing to the air traffic controllers what area is at risk in the next one hour.

CRediT authorship contribution statement

Sandy Chkeir: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Aikaterini Anesiadou:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Alessandra Mascitelli:** Data curation, Formal analysis. **Riccardo Biondi:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work takes place in the frame of the ALARM project funded by the SESAR Joint Undertaking (JU) Programme under grant agreement No 891467 and the CARGO project funded by the SESAR JU Programme under grant agreement No 783287.

Appendix A. Appendix

In an ANN, the processing part is performed in the hidden layer which executes two operational functions: summation function and transfer function, also known as an activation function. This process is summed together on each hidden neuron and it is denoted as (Ibrahim et al., 2020):

$$h_j = f_h\left(\sum_{i=1}^{p} w_{ij}x_i + b_j\right) \tag{14}$$

where h_j, b_j are respectively the output and bias vectors of the jth hidden neuron/node in the hidden layer, w_{ij} is the connection weight from the ith input node x_i to the jth hidden node, $f_h(.)$ is the activation function in the hidden layer, and p is the number of input nodes. Therefore, the future target feature can be predicted in the output layer as (Ibrahim et al., 2020):

$$\hat{y}_k = f_o\left(\sum_{j=1}^{m} v_{jk}h_j + b_k\right) \tag{15}$$

where $\hat{y}_k, b_k$ are respectively the output and bias vectors of the kth output neuron in the output layer, w_j is the connection weight from the jth hidden node to the output node, $f_o(.)$ is the activation function for the output layer, and m is the total number of samples. The bias parameter is used to regulate the output of the neuron in association with the weighted sum of the inputs.

Fig. A1 shows the block diagram of the simple ANN-based nowcasting model chosen as a forecasting baseline in our study. Each sample of our dataset is a 2D matrix composed of the past observations for each input feature. As an output, each sample is a 2D matrix composed of the future time steps for each target feature.

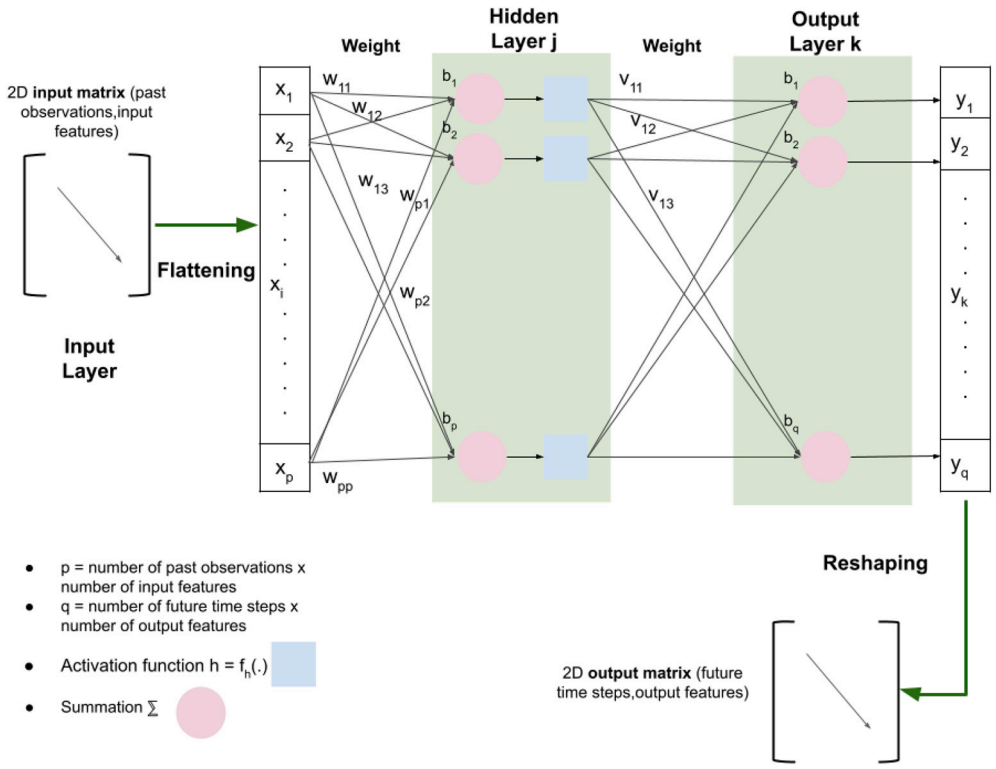

Fig. A1. ANN processing scheme.

Appendix B. Appendix

We have several hyperparameters to tune for both meteorological targets: rain and wind speed. Any experiment is summarized by the minimum, first quartile, median, third quartile, and maximum. In Figs. B1 and B2, the box plot rectangle encloses the median value, with an end at each quartile. Together with the box, the whiskers show how big is the range between the minimum and the maximum and the circles show the outliers. Larger ranges indicate wider distribution and more scattered score errors among the experimental runs. Larger boxes mean more variable data. When the medians are similar, a smaller box shows better performances.

The number of neurons affects the learning capacity of the network. Generally, more neurons are able to learn more from the problem at the cost of longer training time.

The rain validation dataset (Fig. B1) shows the lowest variability with 32 neurons and the lowest median with 256 neurons. The presence of outliers is clear with 64 and 128 neurons.

The wind speed validation dataset (Fig. B2) shows the lowest variability and lowest extremes with 32 neurons; however, the lowest median is achieved with 128 neurons but with the highest variability.

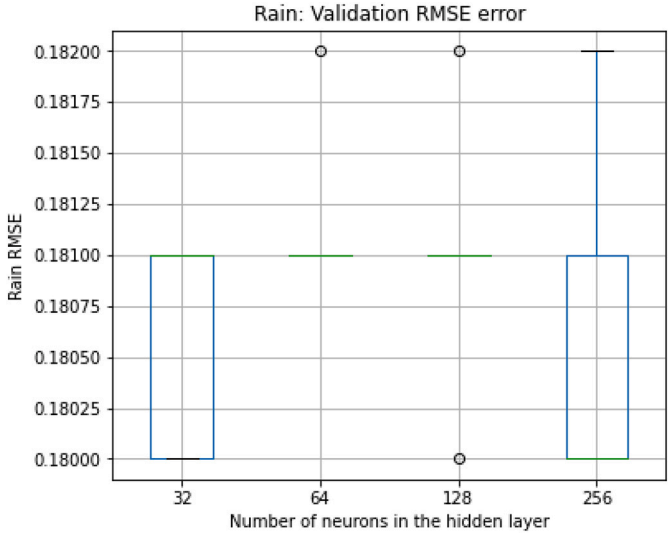

Fig. B1. Box plot for RMSE of the rain validation dataset, reporting the median (green), the first and third quartiles (extremes of the blue rectangle), minimum and maximum (black), and outliers (circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

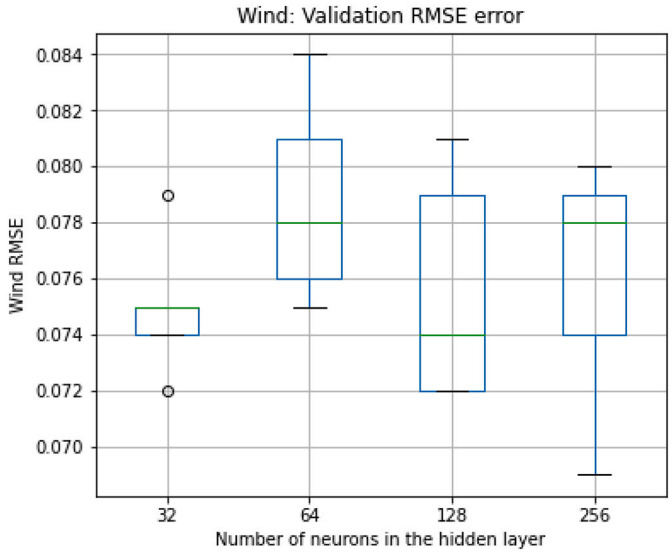

Fig. B2. Box plot for RMSE of the wind speed validation dataset, reporting the median (green), the first and third quartiles (extremes of the blue rectangle), minimum and maximum (black), and outliers (circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

References

Adams, D.K., Barbosa, H.M., De Los, Gaitán, Ríos, K.P., 2017. A spatiotemporal water vapor–deep convection correlation metric derived from the Amazon dense GNSS meteorological network. Mon. Weather Rev. 145 (1), 279–288.

Al Sadeque, Z., Bui, F.M., 2020. A deep learning approach to predict weather data using cascaded LSTM network. In: 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE). IEEE, pp. 1–5. https://doi.org/10.1109/CCECE47787.2020.9255716.

Ayzel, G., Scheffer, T., Heistermann, M., 2020. RainNet v1. 0: a convolutional neural network for radar-based precipitation nowcasting. Geosci. Model Dev. 13 (6), 2631–2644. https://doi.org/10.5194/gmd-2020-30.

Ban, N., Schmidli, J., Schär, C., 2015. Heavy precipitation in a changing climate: does short-term summer precipitation increase faster? Geophys. Res. Lett. 42, 1165–1172. https://doi.org/10.1002/2014GL062588.

Benevides, P., Catalao, J., Nico, G., 2019. Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors. Remote Sens. 11 (8), 966. https://doi.org/10.3390/rs11080966.

Berberan-Santos, M.N., Bodunov, E.N., Pogliani, L., 1997. On the barometric formula. Am. J. Phys. 65 (5), 404–412. https://doi.org/10.1119/1.18555.

Bochenek, B., Ustrnul, Z., 2022. Machine learning in weather prediction and climate analyses—applications and perspectives. Atmosphere 13 (2), 180. https://doi.org/10.3390/atmos13020180.

Bonafoni, S., Biondi, R., Brenot, H., Anthes, R., 2019. Radio occultation and ground-based GNSS products for observing, understanding and predicting extreme events: a review. Atmos. Res. 230, 104624 https://doi.org/10.1016/j.atmosres.2019.104624.

Bonelli, P., Marcacci, P., 2008. Thunderstorm nowcasting by means of lightning and radar data: algorithms and applications in northern Italy. Nat. Hazards Earth Syst. Sci. 8 (5), 1187–1198. https://doi.org/10.5194/nhess-8-1187-2008.

Brownlee, J., 2014. Discover feature engineering, how to engineer features and how to get good at it. In: Machine Learning Process.

Brownlee, J., 2017. Long short-term memory networks with python: develop sequence prediction models with deep learning. Machine Learning Mastery.

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 Syst. Appl. 85, 169–181. https://doi.org/10.1016/j.eswa.2017.05.029.

D'Adderio, L.P., Pazienza, L., Mascitelli, A., Tiberia, A., Dietrich, S., 2020. A combined IR-GPS satellite analysis for potential applications in detecting and predicting lightning activity. Remote Sens. 12, 1031. https://doi.org/10.3390/rs12061031.

Darden, C.B., Nadler, D.J., Carcione, B.C., Blakeslee, R.J., Stano, G.T., Buechler, D.E., 2010. Utilizing total lightning information to diagnose convective trends. Bull. Am. Meteorol. Soc. 91 (2), 167–176. http://www.jstor.org/stable/26232859.

Emanuel, K.A., 2005. Increasing destructiveness of tropical cyclones over the past 30 years. Nature 436, 686–688. https://doi.org/10.1038/nature03906.

Enders, C.K., 2010. Applied Missing Data Analysis. Guilford press.

Fathi, E., Maleki Shoja, B., 2018. Chapter 9—Deep neural networks for natural language processing. In: Gudivada, V.N., Rao, C.R. (Eds.), Handbook of Statistics, 38, pp. 229–316. https://doi.org/10.1016/bs.host.2018.07.006.

Franch, G., Maggio, V., Coviello, L., Pendesini, M., Jurman, G., Furlanello, C., 2020. TAASRAD19, a high-resolution weather radar reflectivity dataset for precipitation nowcasting. Sci. Data 7 (1), 1–13. https://doi.org/10.1038/s41597-020-0574-8.

Gagniuc, P.A., 2017. Markov Chains: From Theory to Implementation and Experimentation. John Wiley & Sons.

Grazzini, F., Craig, G.C., Keil, C., Antolini, G., Pavan, V., 2020. Extreme precipitation events over northern Italy. Part I: a systematic classification with machine-learning techniques. Q. J. R. Meteorol. Soc. 146 (726), 69–85. https://doi.org/10.1002/qj.3635.

Gultepe, I., Sharman, R., Williams, P.D., Zhou, B., Ellrod, G., Minnis, P., Trier, S., Griffin, S., Yum, S., Gharabaghi, B., Feltz, W., 2019. A review of high impact weather for aviation meteorology. Pure Appl. Geophys. 176 (5), 1869–1921. https://doi.org/10.1007/s00024-019-02168-6.

Guo, M., Zhang, H., Xia, P., 2021. Exploration and analysis of the factors influencing GNSS PWV for nowcasting applications. Adv. Space Res. 67 (12), 3960–3978. https://doi.org/10.1016/j.asr.2021.02.018.

Hemri, S., Scheuerer, M., Pappenberger, F., Bogner, K., Haiden, T., 2014. Trends in the predictive performance of raw ensemble weather forecasts. Geophys. Res. Lett. 41, 9197–9205. https://doi.org/10.1002/2014GL062472.

Hewage, P., Trovati, M., Pereira, E., Behera, A., 2021. Deep learning-based effective fine-grained weather forecasting model. Pattern. Anal. Applic. 24 (1), 343–366. https://doi.org/10.1007/s10044-020-00898-1.

Hoeppe, P., 2016. Trends in weather related disasters–Consequences for insurers and society. Weather Clim. Extrem. 11, 70–79. https://doi.org/10.1016/j.wace.2015.10.002.

Hong, W.-C., 2008. Rainfall forecasting by technological machine learning models. Appl. Math. Comput. 200, 41–57. https://doi.org/10.1016/j.amc.2007.10.046.

Hov, Ø., Cubasch, U., Fischer, E., Höppe, P., Iversen, T., Gunnar Kvamstø, N., Kundzewicz, W.Z., Rezacova, D., Rios, D., Duarte Santos, F., Schädler, B., 2013. Extreme Weather Events in Europe: Preparing for Climate Change Adaptation. Norwegian Meteorological Institute.

Ibrahim, M., Alsheikh, A., Al-Hindawi, Q., Al-Dahidi, S., ElMoaqet, H., 2020. Short-Time Wind Speed Forecast Using Artificial Learning-Based Algorithms, Computational Intelligence and Neuroscience, 2020. Article ID 8439719. https://doi.org/10.1155/2020/8439719.

James, P.M., Reichert, B.K., Heizenreder, D., 2018. NowCastMIX: automatic integrated warnings for severe convection on nowcasting time scales at the German weather service. Weather Forecast. 33 (5), 1413–1433. https://doi.org/10.1175/WAF-D-18-0038.1.

Jeon, S., Paciorek, C.J., Wehner, M.F., 2016. Quantile-based bias correction and uncertainty quantification of extreme event attribution statements. Weather Clim. Extrem. 12, 24–32. https://doi.org/10.1016/j.wace.2016.02.001.

Joss, J., Schädler, B., Galli, G., Cavalli, R., Boscacci, M., Held, E., Della Bruna, G., Kappenberger, G., Nespor, V., Spiess, R., 1998. Operational Use of Radar for Precipitation Measurements in Switzerland. Vdf Hochschulverlag, ETH Zürich. ISBN 3–7281–2501-6.

Kang, H., 2013. The prevention and handling of the missing data. Korean J. Anesthesiol. 64 (5), 402–406.

Karevan, Z., Suykens, J.A., 2018. Spatio-Temporal Stacked LSTM for Temperature Prediction in Weather Forecasting arXiv preprint arXiv:1811.06341.

Khosravi, A., Machado, L., Nunes, R.O., 2018. Time-series prediction of wind speed using machine learning algorithms: a case study Osorio wind farm, Brazil. Appl. Energy 224, 550–566. https://doi.org/10.1016/j.apenergy.2018.05.043.
Klocek, S., Dong, H., Dixon, M., Kanengoni, P., Kazmi, N., Luferenko, P., Lv, Z., Sharma, S., Weyn, J., Xiang, S., 2021. MS-nowcasting: Operational Precipitation Nowcasting with Convolutional LSTMs at Microsoft Weather arXiv preprint arXiv: 2111.09954.
Kohn, M., Galanti, E., Price, C., Lagouvardos, K., Kotroni, V., 2011. Nowcasting thunderstorms in the Mediterranean region using lightning data. Atmos. Res. 100 (4), 489–502. https://doi.org/10.1016/j.atmosres.2010.08.010.
Lagasio, M., Pulvirenti, L., Parodi, A., Boni, G., Pierdicca, N., Venuti, G., Rommen, B., 2019. Effect of the ingestion in the WRF model of different Sentinel-derived and GNSS-derived products: analysis of the forecasts of a high impact weather event. Eur. J. Remote Sens. 52, 16–33. https://doi.org/10.1080/22797254.2019.1642799.
Laurila, T.K., Sinclair, V.A., Gregow, H., 2021. Climatology, variability, and trends in near-surface wind speeds over the North Atlantic and Europe during 1979–2018 based on ERA5. Int. J. Climatol. 41 (4), 2253–2278. https://doi.org/10.1002/joc.6957.
Laviola, S., Levizzani, V., Ferraro, R., Beauchamp, J., 2020. Hailstorm detection by satellite microwave radiometers. Remote Sens. 12, 621. https://doi.org/10.3390/rs12040621.
Legates, D.R., 2002. Limitations of climate models as predictors of climate change. Brief Anal. 396.
Li, W., Chen, J., Li, L., Chen, H., Liu, B., Xu, C.-Y., Li, X., 2019. Evaluation and bias correction of S2S precipitation for hydrological extremes. J. Hydrometeorol. 20, 1887–1906. https://doi.org/10.1175/JHM-D-19-0042.1.
Liu, C., Heckman, S., 2012. Total lightning data and real-time severe storm prediction. In: TECO-2012: WMO Tech. Conf. on Meteorological and Environmental Instruments and Methods of Observation.
Liu, C., Zipser, E.J., 2005. Global distribution of convection penetrating the tropical tropopause. J. Geophys. Res.-Atmos. 110 (D23) https://doi.org/10.1029/2005JD006063.
Liu, J.N., Hu, Y., You, J.J., Chan, P.W., 2014. Deep neural network based feature representation for weather forecasting. In: Proceedings on the International Conference on Artificial Intelligence (ICAI). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
Livieris, I.E., Stavroyiannis, S., Iliadis, L., et al., 2021. Smoothing and stationarity enforcement framework for deep learning time-series forecasting. Neural Comput. & Applic. 33, 14021–14035. https://doi.org/10.1007/s00521-021-06043-1.
Łoś, M., Smolak, K., Guerova, G., Rohm, W., 2020. GNSS-based machine learning storm nowcasting. Remote Sens. 12 (16), 2536. https://doi.org/10.3390/rs12162536.
Maqsood, I., Khan, M.R., Abraham, A., 2003. Weather forecasting models using ensembles of neural networks. In: Intelligent Systems Design and Applications. Springer, Berlin, Heidelberg, pp. 33–42. https://doi.org/10.1007/978-3-540-44999-7_4.
Maraun, D., 2013. When will trends in European mean and heavy daily precipitation emerge? Environ. Res. Lett. 8, 014004 https://doi.org/10.1088/1748-9326/8/1/014004.
Mascitelli, A., Federico, S., Torcasio, R., Dietrich, S., 2020. Assimilation of GPS zenith total delay estimates in RAMS NWP model: impact studies over Central Italy. Adv. Space Res. https://doi.org/10.1016/j.asr.2020.08.031.
Mcculloch, W., Pitts, W., 1943. A logical calculus of ideas immanent in nervous activity. Bull. Math. Biophys. 5, 127–147.
Myhre, G., Alterskjær, K., Stjern, C.W., Hodnebrog, Ø., Marelle, L., Samset, B.H., Stohl, A., 2019. Frequency of extreme precipitation increases extensively with event rareness under global warming. Sci. Rep. 9 (1), 1–10. https://doi.org/10.1038/s41598-019-52277-4.
Nisi, L., Hering, A., Germann, U., Martius, O., 2018. A 15-year hail streak climatology for the Alpine region. Q. J. R. Meteorol. Soc. 144 (714), 1429–1449. https://doi.org/10.1002/qj.3286.
Pielke Jr., R.A., Rubiera, J., Landsea, C., Fernández, M.L., Klein, R., 2003. Hurricane vulnerability in Latin America and the Caribbean: normalized damage and loss potentials. Nat. Hazards Rev. 4 (3), 101–114. https://doi.org/10.1061/(ASCE)1527-6988(2003)4:3(101).
Powers, J.G., Klemp, J.B., Skamarock, W.C., Davis, C.A., Dudhia, J., Gill, D.O., Coen, J.L., Gochis, D.J., Ahmadov, R., Peckham, S.E., Grell, G.A., 2017. The weather research and forecasting model: overview, system efforts, and future directions. Bull. Am. Meteorol. Soc. 98 (8), 1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1.
Rädler, A.T., Groenemeijer, P., Faust, E., Sausen, R., 2018. Detecting severe weather trends using an additive regressive convective hazard model (AR-CHaMo). J. Appl. Meteorol. Climatol. 57 (3), 569–587. https://doi.org/10.1175/JAMC-D-17-0132.1.
Rädler, A.T., Groenemeijer, P.H., Faust, E., Sausen, R., Púčik, T., 2019. Frequency of severe thunderstorms across Europe expected to increase in the 21st century due to rising instability. npj Clim. Atmos. Sci. 2 (1), 1–5. https://doi.org/10.1038/s41612-019-0083-7.
Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., 2021. Skilful precipitation nowcasting using deep generative models of radar. Nature 597 (7878), 672–677. https://doi.org/10.1038/s41586-021-03854-z.
Roeger, C., Stull, R., McClung, D., Hacker, J., Deng, X., Modzelewski, H., 2003. Verification of mesoscale numerical weather forecasts in mountainous terrain for application to avalanche prediction. Weather Forecast. 18 (6), 1140–1160. https://doi.org/10.1175/1520-0434(2003)018<1140:VOMNWF>2.0.CO;2.
Sapucci, L.F., Machado, L.A., de Souza, E.M., Campos, T.B., 2019. Global P ositioning S ystem precipitable water vapour (GPS-PWV) jumps before intense rain events: a potential application to nowcasting. Meteorol. Appl. 26 (1), 49–63. https://doi.org/10.1002/met.1735.
Scoccimarro, E., Villarini, G., Vichi, M., Zampieri, M., Fogli, P.G., Bellucci, A., Gualdi, S., 2015. Projected changes in intense precipitation over Europe at the daily and subdaily time scales*. J. Clim. 28, 6193–6203. https://doi.org/10.1175/JCLI-D-14-00779.1.
Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D.Y., Wong, W.K., Woo, W.C., 2017. Deep learning for precipitation nowcasting: A benchmark and a new model. In: Advances in Neural Information Processing Systems, 30, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.
Solazzo, E., Tournigand, P.Y., Barindelli, S., Guglieri, V., Realini, E., Nisi, L., Biondi, R., 2020. Understanding Severe Weather Events at Airport Spatial Scale. IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, pp. 5372–5375. https://doi.org/10.1109/IGARSS39084.2020.9323598.
Soler, M., 2021. D1.1 - Project Management Plan. ALARM Project. Funded by the SESAR Joint Undertaking (JU) under grant agreement No 891467.
Stott, P., 2016. How climate change affects extreme weather events. Science 352, 1517–1518. https://doi.org/10.1126/science.aaf7271.
Sutskever, I., Vinyals, O., Le, Q.V., 2014. Sequence to sequence learning with neural networks. Adv. Neural Inf. Proces. Syst. 27.
Tekin, S.F., Karaahmetoglu, O., Ilhan, F., Balaban, I., Kozat, S.S., 2021. Spatio-Temporal Weather Forecasting and Attention Mechanism on Convolutional lstms arXiv preprint arXiv:2102.00696.
Thiruvengadam, P., Indu, J., Ghosh, S., 2020. Significance of 4DVAR radar data assimilation in weather research and forecast model-based nowcasting system. J. Geophys. Res.-Atmos. 125 https://doi.org/10.1029/2019JD031369t e2019JD031369.
Tibshirani, S., Friedman, H., 2017. The Elements of Statistical Learning, 2nd edition.
Trenberth, K.E., 2012. Framing the way to relate climate extremes to climate change. Clim. Chang. 115 (2), 283–290. https://doi.org/10.1007/s10584-012-0441-5.
van Delden, A., 2001. The synoptic setting of thunderstorms in western Europe. Atmos. Res. 56, 89–110. https://doi.org/10.1016/S0169-8095(00)00092-2.
Van den Besselaar, E.J.M., Klein Tank, A.M.G., Buishand, T.A., 2013. Trends in European precipitation extremes over 1951–2010. Int. J. Climatol. 33 (12), 2682–2689. https://doi.org/10.1002/joc.3619.
Wang, G., Wang, D., Yang, J., Liu, L., 2016. Evaluation and correction of quantitative precipitation forecast by storm-scale NWP model in Jiangsu, China. Adv. Meteorol. 2016, 1–13. https://doi.org/10.1155/2016/8476720.
Williams, E., Boldi, B., Matlin, A., Weber, M., Hodanish, S., Sharp, D., Goodman, S., Raghavan, R., Buechler, D., 1999. The behavior of total lightning activity in severe Florida thunderstorms. Atmos. Res. 51 (3–4), 245–265. https://doi.org/10.1016/S0169-8095(99)00011-3.
Wu, F., Cui, X., Zhang, D.L., 2018. A lightning-based nowcast-warning approach for short-duration rainfall events: development and testing over Beijing during the warm seasons of 2006–2007. Atmos. Res. 205, 2–17. https://doi.org/10.1016/j.atmosres.2018.02.003.
Zaytar, M.A., El Amrani, C., 2016. Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. Int. J. Comput. Appl. 143 (11), 7–11.
Zhang, F., Wang, X., Guan, J., Wu, M., Guo, L., 2021. RN-Net: a deep learning approach to 0–2 h rainfall nowcasting based on radar and automatic weather station data. Sensors 21 (6), 1981. https://doi.org/10.3390/s21061981.