https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

Improving object-oriented radar based nowcast by a nearest neighbour approach

Bora SHEHU[1], Uwe HABERLANDT[1]
[1]Institute for Hydrology and Water Resources Management, Leibniz University Hannover, Germany

Correspondence to: Bora Shehu (shehu@iww.uni-hannover.de)

Abstract.

The nowcast of rainfall storms at fine temporal and spatial resolutions is quite challenging due to the erratic nature of rainfall at such scales. Typically, rainfall storms are recognized by radar data, and extrapolated in the future by the Lagrangian persistence. However, storm evolution is much more dynamic and complex than the Lagrangian persistence, leading to short forecast horizons especially for convective events. Thus, the aim of this paper is to investigate the improvement that past similar storms can introduce to the object-oriented radar based nowcast. Here we propose a nearest neighbour approach that measures first the similarity between the "to-be-nowcasted" storm and past observed storms, and later uses the behaviour of the past most similar storms to issue either a single nowcast (by averaging the 4 most similar storm-responses) or an ensemble nowcast (by considering 30 most similar storm-responses). Three questions are tackled here: i) what features should be used to describe storms in order to check for similarity? ii) how to measure similarity between past storms? and iii) is this similarity useful for storm oriented nowcast? For this purpose, individual storms from 110 events in the period 2000-2018 recognized within the Hannover Radar Range (R~115km^2), Germany, were used as a basis for investigation. A "leave-one-event-out" cross-validation is employed to train and validate the nearest neighbour approach for the prediction of the area, mean intensity, the x and y velocity components of the "to-be-nowcasted" storm for lead times up to + 3 hours. Prior to the training, two importance analyses methods (Pearson correlation and partial information correlation) are employed to identify the most important predictors. The results indicate that most of storms behave similarly, and the knowledge obtained from such similar past storms can improve considerably the storm nowcast compared to the Lagrangian persistence especially for convective events (storms shorter than 3 hours) and longer lead times (from 1 to 3 hours). The nearest neighbour approach seems promising, nevertheless there is still room for improvement by either increasing the sample size or employing more suitable methods for the predictor identification.

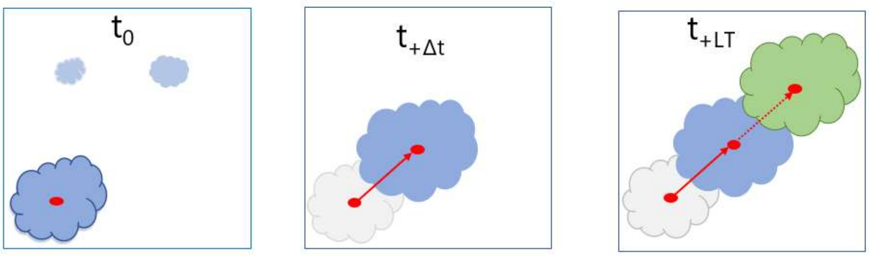
Keywords:

Rainfall nowcast, Lagrangian persistence, probabilistic nowcast, similar storms, nearest neighbour

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

1. Introduction

Typically, radar based nowcasts are used for short-term rainfall nowcast. The rainfall is either considered as an object (a set of radar grid cells with the intensity above a threshold that moves together as a unit and is regarded as a storm (Dixon & Wiener, 1993; Johnson et al., 1998)) or as an intermittent field (intensity is moving from one pixel of the radar image to the other (Ruzanski et al., 2011; Zahraei et al., 2012)). Whilst the field-based approach of rainfall nowcasting has gained popularity recently, here the focus is only on the object-oriented forecast, thus on the nowcasting of storms. In such forecast three mains steps are performed (illustrated in **Figure 1**): a) first the storm is identified – so a group of grid cells with intensity higher than a threshold is recognized in the radar image at time t_0, b) the storm identified is then tracked for the time $t_0+\Delta t$ (where Δt is the temporal resolution of the radar data) and velocities are assigned to the movement of the storm, and finally c) the storm as lastly observed at time t (when the forecast is issued) is extrapolated at a specific lead time (the time in the future when the forecast is needed) t_{+TL}, with the last observed velocity vector. This is a linear extrapolation of the storm structure in the future, considering the spatial intensity distribution within the storm and the movement of the storm as constant in time - also referred to as Lagrangian Persistence (Germann et al., 2006). Applications of such storm-based nowcast are common in literature like TITAN, HyRaTrac, Konrad etc. (Han et al., 2009; Hand, 1996; Krämer, 2008; Lang, 2001; C. E. Pierce et al., 2004).



a) Storm Identification *b) Storm Tracking* *c) Storm Extrapolation*

Figure 1 *The main steps of an object-based radar nowcast. Blue indicates the current state of the storm at any time t, grey indicates the past states of the storm (at t-Δt), and green indicates the future states of the storm (t_{+LT}) (Shehu, 2020)*

One of the main drawbacks of radar-based forecast, is that a storm has to be first identified in order to be extrapolated in the future. In other words, the storm cannot be predicted before it has started anywhere in the region, only the movement can be predicted. As already discussed in Bowler et al., (2006) and Jensen et al. (2015), these birth errors cause the radar forecast to be used only for short lead time forecast (up to 3 hours), and for longer lead times a blending between a Numerical Weather Prediction Model (NWP) and radar based nowcast should be used instead (Codo & Rico-Ramirez, 2018; Foresti et al., 2016; Jasper-Tönnies et al., 2018). Nevertheless, for short lead times (1-2 hours) the radar based nowcast is still preferred as it outperforms the NWP nowcasts (Berenguer et al., 2012; Jensen et al., 2015; Lin et al., 2005; Zahraei et al., 2012). Apart from the birth errors, other sources of the errors in the object-oriented nowcast can be attributed to storm identification, storm tracking and Lagrangian extrapolation (L. Foresti & Seed, 2015; C. Pierce et al., 2012; Rossi et al., 2015).

Many works have been already conducted to investigate the role that different intensity thresholds for the storm identification, or that different storm tracking algorithms have on the nowcasting results (Goudenhoofdt & Delobbe, 2013; Han et al., 2009; Hou & Wang, 2017; Jung & Lee, 2015; Kober & Tafferner, 2009). Very high intensity thresholds may be suitable for convective storms, however can cause false splitting of the storms and which can affect negatively the tracking algorithm. Thus, one has to be careful in adjusting the intensity threshold dynamically over the radar field and type of storm. Storm tracking algorithm can be improved if certain relationships are learned from past observed dataset

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

(like a Fuzzy approach in Jung & Lee (2015) or a tree-based structure in Hou & Wang (2017)), but there is still a limit that the tracking improvement cannot surpass due to the implementation of the Lagrangian persistence (Hou & Wang, 2017).

The errors due to the Lagrangian persistence are particularly high for convective events at longer lead times (past 1 hour) as the majority of convective storms last no longer than 60 minutes (Goudenhoofdt & Delobbe, 2013; Wilson et al., 1998). At these lead times, the persistence fails to predict mainly the death of these storm cells, while for shorter lead times it fails to represent the growing/decaying rate and the changing movement of a storm cell (Germann et al., 2006). For stratiform events, since they are more persistent in nature, Lagrangian persistence can potentially give reliable results up to 2 or 3 hours lead time (Krämer, 2008). Nevertheless studies have found that for fine spatial (1km^2) and temporal (5min) scales, the Lagrangian Persistence can yield reliable results up to 20-30 min lead time, which is also known in the literature as the predictability limit at such scales (Grecu & Krajewski, 2000; Kato et al., 2017; Ruzanski et al., 2011). For object-oriented radar based nowcast, this predictability limit can be extended up to 1 hour for stratiform events and up to 30-45min for convective events if a better radar product (merged with rain gauge data) is fed into the nowcast model (Shehu & Haberlandt, 2021). Past these lead times, the errors due to the growth/decay and death of the storm govern.

Nevertheless, one can estimate roughly these processes by acknowledging previously observed storm cells. For instance, if it is known that a storm is of convective nature, most probably it will die from 20 min to two hours of the storm birth, the peak intensities happen mainly in the afternoon or evening, and that they dissipate fast after the peak intensity has been reached. Such characteristics of storm behaviour can be analysed from the past observation (Goudenhoofdt & Delobbe, 2013; Zawadzki, 1973). As stated by (Gallus et al., 2008), the rainfall characteristics are related to the morphological characteristics of the storm itself. Thus, it is to be expected that an extensive observation of past storm behaviours can be very useful in creating and establishing new nowcasting rules (Wilson et al., 2010).

An implementation of such learning from previous observed storms (with focus only on the object-based nowcast and not the field-based one) show for instance Hou & Wang (2017) where a Fuzzy classification scheme was implemented to improve the tracking and matching of storms which resulted in an improved nowcast, and Zahraei et al. (2013) where a SOM algorithm was used to predict the birth and decay of storms on coarse scales extending the predictability of storms by 20 %. These studies suggest that past observed storms may be useful in extending the predictability limit of the storms at hand. Thus, the aim of this study is to investigate if non-linear relationships learned from past observed storms can surpass the Lagrangian persistence and extend the predictability limit of different storms. For this purpose, a nearest neighbour method is developed at the storm scale, which is used to first recognize similar storms in the past, and then assign their behaviours to the "to-be-nowcasted" storm.

The nearest neighbour method has been used in the field of hydrology mainly for classification , regression or resampling purposes (e.g. Lall & Sharma (1996)) but there are some examples of prediction as well (Galeati, 1990). The assumption of this method is that similar events are described by similar predictors, and thus if one identifies the predictors successfully, similar events that behave similarly can be identified. For a new event, the respective response is then obtained by averaging the responses of past k – most similar storms. The k-value can be optimized by minimizing a given cost function. Because of the averaging, the response obtained, will be a new one, satisfying thus the condition that nature doesn't repeat itself, but nevertheless it is confined within the limits of the observed events. Consequently, a k-NN nowcast is unable to predict extreme behaviours outside of the observed range.

The application of the k-NN seems reasonable as an extension of the object-oriented radar based nowcast. It can be used instead of the Lagrangian persistence in step 3, for the extrapolation of rainfall storms into the future. The benefit of the k-NN application is that one can either give a single or an ensemble nowcast; since k-neighbours can be selected as similar to a storm at hand, a probability based on the similarity rank, can be issued at each of the past storm, providing

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

so an ensemble of responses. Ensemble nowcasts are more preferred for rainfall nowcasts due to the high uncertainty associated with rainfall predictions at such fine scales (Germann & Zawadzki, 2004).

Before applying a k-NN for the storm nowcast, question that arise are I) what features are more important when describing a storm, II) how to evaluate similarity between storms and III) how to use their information for the nowcasting of the storm at hand. To answer these questions and to achieve the aim of this study, the paper is organized as follows. First in Section 2 the study area is described, following with the structure of the k-NN method in Section 3.1 where: the generation of the storm database is discussed in Section 3.1.1, the predictors selected and target variables are given in in Section 3.1.2, the methods used for predictor identification in Section 3.1.3, and different application of the k-NN in Section 3.1.4. The training and the performance criteria are shown in Section 3.2 followed by the results in Section 4 separated into predictors influence (Section 4.1), single k-NN (Section 4.2) and ensemble k-NN performance (Section 4.3). Finally, the study is closed by derived conclusions and outlook in Section 5.

2. Study Area and Data

The study area is located in northern Germany, and lies within the Hannover Radar Range as illustrated in **Figure 2**. The radar station is situated at the Hannover Airport, and it covers an area with a radius of 115 km. The Hannover radar data are C-band data provided by German Weather Service (DWD), and measure the reflectivity at an azimuth angle of 1° and at 5 min scans (Winterrath et al., 2012). The reflectivity measurements are converted to intensity according to Marshall-Palmer relationship with the coefficients a=256 and b=1.42 (Bartels et al., 2004). The radar data were corrected from the static clutters and erroneous beams and then converted to Cartesian Coordinate system (1 km^2 and 5 min) as described in Berndt et al. (2014). Additionally, following the results from Shehu & Haberlandt (2021), a conditional merging between the radar data and 100 gauge recording with the radar range at 5 min time steps was performed. The period from 2000 to 2018 was used as a basis for this investigation, from which 110 events with different characteristics were extracted (see Shehu & Haberlandt (2021) or Shehu (2020)). These events were selected for urban flood purposes, thus contain mainly convective events and few stratiform events.

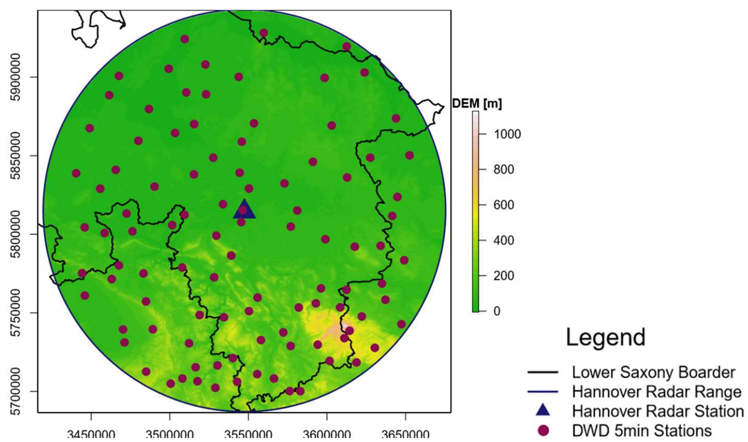


***Figure 2** The available recording rain gauges (red) and radar (blue) inside the study area.*

3. Methods

3.1 Developing the k-NN model

3.1.1 Generating the storm database

Each of the selected events contains many storms, whose identification and tracking was performed on the basis of the HyRaTrac algorithm in the hindcast mode (Krämer, 2008; Schellart et al., 2014). A storm is initialized if a group of radar grid cells (> 64) has a reflectivity higher than Z=20dBz, while storms are recognized as convective – if a group

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

bigger than 16 radar grid cells has an intensity higher than 25 dBz, and as stratiform – if a group bigger than 128 radar grid cells has an intensity higher than 20 dBz. The tracking of individual storms in consecutive images is done by the optimization of the cross-correlation between the last 2 images (t=0 and t-5 min), and local displacement vectors for each storm are calculated. In case a storm is just recognized, then global displacement vectors based on cross-correlation of the entire radar image are assigned to them.

Thus, a dataset with several types of storms is built and saved. The storms are saved with an ID based on the starting time and location, and for each time step of the storm evolution the spatial information is saved. Here the spatial rainfall intensities of a storm at a particular time step (in 5min) of the storms' life, is referred to as the "state" of the storm. A storm that has been observed for 15 minutes, consists of three "states" each occurring at a 5 min time step. For each of the storm states an ellipsoid is fitted to the intensities in order to calculate the major and minor axis and the orientation angle of the major axis. This storm database is the basis for developing the k-NN method and for investigating the similarity between storms. Some characteristics of the identified storms like duration, mean area, maximum intensity, number of splits/merges, local velocity components, and ellipsoidal features, are shown in the **Figure 3**.

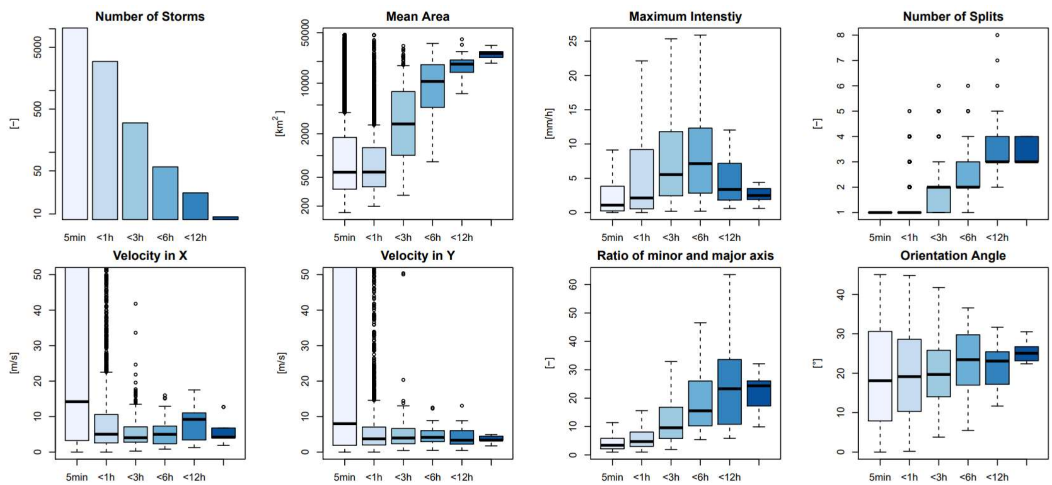


***Figure 3** Different properties of the storms recognized from 110 events separated into 6 groups according to their duration (shown in different shades of blue).*

As seen from the number of storms for each duration in **Figure 3**, the unmatched storm cells make the majority of the storms recognized. These are storms that last just 5 min (one-time step) as the algorithm fails to track them at consecutive time steps. These "storms" can either be dynamic clutters from the radar measurement, as they are characterized by small area, circular shapes (small ratio of minor and major axis) and by very high velocities, or artefacts created by low intensity thresholds used for the storm identification, or finally produced by the unrepresentativeness of the volume captured by the radar station. Apart from the unmatched storms, the majority of the remaining storms are of convective nature: storms with short duration (shorter than 6 hours), high intensity and low areal coverage.

Here two types of convective storms are distinguished: local convective with very low coverage and low intensity, and mesoscale convective which are responsible for floods (very high intensity) and have a larger coverage. The stratiform storms characterized by large area, long duration and low intensities, as well as meso- γ scale convective events with duration up to 6 hours, are not very well represented by the dataset as only a few of them are present in the selected events (respectively circa 20 and 50 storms). Therefore, it is to be expected that the k-NN approach may not yield very good results for such storms due to the low representativeness. From the characteristics of the storms illustrated in **Figure 3**, it

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

can be seen that for stratiform storms that live longer than twelve hours the variance of the characteristics is quite low (when compared to the rest of the storms) which can either be attributed to the persistence of such storms or to the low representativeness in the database. Thus, even though the data size for stratiform is quite small, the k-NN may still deliver good results as characteristics of such storms are more similar.

3.1.2 Selecting features for similarity and target variables

At first storms are treated like objects that manifest certain features (predictors) like area, intensity, lifetime etc., at each state of the storms' life until the storm dies (and the predictors are all set to zero). The features of the objects are categorized into present and past features, as illustrated in **Figure 4** (shown respectively in blue and grey). The present features describe the current state of the storm at the time of nowcast (denoted with t_0 in **Figure 4**), and are calculated from one state of the storm. To compute certain features, an ellipsoid is fitted to each state of the storm. The past features, on the other hand, describe the predictors of the past storm states (denoted with t_{-1}, t_{-2} in **Figure 4)** and their change over the past life of the storm. For example, the average area from time t_{-2} to t_{-1} is a past feature. A pre-analysis of important predictors showed that the average features over the last 30 minutes are more suitable as past predictors than the averages over last 15 or 60 min or than the calculation of past changing rates. Therefore, averages over past 30 minutes are computed here:

$$P_{30} = \sum_{i=t_0}^{t-30min} P_i/6 \,, \quad (1)$$

where P_i is the predictors value at time i, and P_{30} the average value of the predictor over last 30min. The selected features (both present and past) that are used here to describe storms as objects and hence tested as predictors are shown in **Table 1**. The present features help to recognize storms that are similar at the given state when the nowcast is issued (blue storm in **Figure 4**) and the past ones give additional information about the past evolution of the storm (average of grey storms in **Figure 4**).

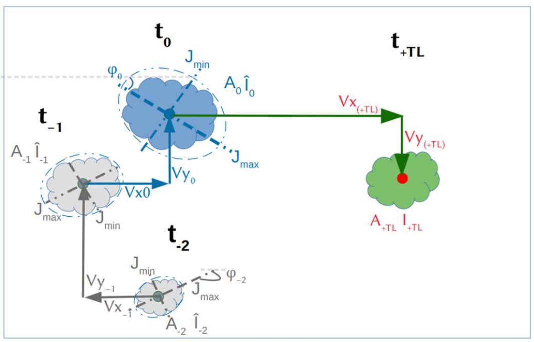


***Figure 4** The features describing the past (grey) and present (blue) states of the storm used as predictors to nowcast the future states of the storm (green) at a specific lead time (T_{+LT}) that is described by 4 target variables (in red). The nowcast is issued time t_0.*

The aim of these features is to recognize the states of previously observed storms that are most similar to the current one (shown in blue in **Figure 4**) of the "to-be-nowcasted" storm. Once the most similar past storm states are recognized, their respective future states at different lead times can be assigned as the future behaviour (shown in green in **Figure 4**) of the current state of the "to-be-nowcasted" storms. Since the storms are regarded as objects with specific features, future behaviours at different lead times are determined by four target variables: area (A_{+LT}), mean intensity (I_{+LT}) and velocity in X (Vx_{+LT}) and Y (Vy_{+LT}) direction. Additionally, the total lifetime of the storm is considered as a fifth target (L_{tot}). Theoretically, the total lifetime is predicted indirectly when any of the first four targets is set to zero, however here it is considered as an independent variable in order to investigate if similar storms have similar lifetime durations.

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

For each state of each observed storm in the database, the past and present features of that state with its' respective future states of the five target variables from +5min to +180min (every 5 min) lead times are saved together and form the predictor-target database that is used for the development of the k-NN nowcast model. A summary of the predictors and target variables calculated per state is given in **Table 1**. Before training and validating the k-NN method, an importance analysis is performed for each of the target variables in order to recognize the most important predictors. As the predictors have different ranges, prior to the importance analysis and the k-NN application, they are normalized according to their median and range between the 0.05 and 0.95 quantiles:

$$normP_i = \frac{P_i - Q_{Pi}^{0.5}}{Q_{Pi}^{0.95} - Q_{Pi}^{0.05}}, \qquad (2)$$

where P is the actual value, $normP$ the normalized value, and $Q_{Pi}^{0.5}$, $Q_{Pi}^{0.95}$, $Q_{Pi}^{0.05}$ the quantiles 0.5, 0.05 and 0.95 of the i^{th} predictors' vector. The reason why these quantiles were used for the normalization instead of the typical mean and maximum to minimum range, is that some outliers are present in the data. For instance, very high and unrealistic velocities are present in some convective storms where the tracking algorithm fails to capture adequate velocities (Han et al., 2009). Thus, to avoid the influence of these outliers, the given range is employed.

3.1.3 Selection of most relevant predictors

The application of the k-NN method can be relevant if there is a clear connection between the target variable and the features describing this target variable. For instance, in the case of Galeati (1990), a physical background backed up the connection between target variable (discharge) and the features (daily rainfall volume and mean temperature). In the case of the storms at such fine temporal and spatial scales, due to the erratic nature of the rainfall itself, there are no physical related information that can be extracted from radar data. Different features of the storm itself can be investigated for their importance to the target variable. Nevertheless, the identification of such features (referred here as predictors) is difficult because it is bounded to the set of the available data and the relationships considered. Commonly a strong correlation between the predictors selected and the target variable is used as an indicator of a strong relationship between them. However, the relationship between predictors and target variables may still be of non-linear nature, thus another predictor important analysis should be advised when selecting the predictors. Sharma & Mehrotra (2014) proposed a new methodology, designed specifically for the k-NN approach, where no prior assumption about the system type is required. The method is based on a metric called the Partial Information Correlation and is computed from the Partial Information as:

$$PIC = \sqrt{(1 - \exp(-2PI))}, \qquad (3)$$

where PIC is the Partial Information Correlation and the PI is the Partial Information. The Partial Information itself is a modification of the Mutual Information in order to measure partial dependency between the predictors and the target variable, by adding predictors one at a time (step-wise procedure). The evaluation of PIC needs a pre-existing identified predictor from which the computation can start. If the pre-defined predictor is correctly selected, then through the Equation (3), the method is able to recognize and leave out the new predictors which are not related to the response and which don't bring additional value to the existing relationship between the current predictors and target variable. Relative weights for the k-NN regression application can be derived for each predictor, as a relationship between the PIC metric and the associated partial correlation:

$$\alpha_j = PIC_{X,Xj|X(-j)} \frac{s_{Y|X(-j)}}{s_{X|X(-j)}}, \qquad (4)$$

where X is the predictor, Y the target response, $S_{Y|X(-j)}$ the scaled conditional standard deviations between the first predictor and the target, $S_{Xj|X(-j)}$ the scaled conditional standard deviations between the additional predictor and the first one, and the

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

α_j the predictors weight. The R package NPRED was used for the investigation of the PIC derived importance weights (Sharma et al., 2016).

***Table 1** List of all the past and present features of the storm object that are investigated for their importance as predictors, and the respective target variables calculated for different lead times.*

	Features	Symbol
Present Features	number of storm cells within the storm region	Cells [-]
	current storm lifetime at time of nowcast	L_{now} [min]
	area of the storm	A [km^2]
	mean spatial intensity	I_{ave} [mm/h]
	maximum spatial intensity	I_{max} [mm/h]
	standard deviation of the spatial intensities	I_{sd1} [-]
	standard deviation of intensities groups inside the storm	I_{sd2} [-]
	global velocity of the entire radar image	V_g [m/s]
	x and y component of the local velocity of the storm region	V_x, V_y [m/s]
	major and minor axis of the ellipsoid and their ratio	J_{max}, J_{min}[km] J_r [-]
	orientation angle of the major axis of the ellipsoid	Φ [°]
Past Features	average area over the last 30 min of storm existence	A_{30} [km^2]
	average mean intensity over the last 30 min of storm existence	$Iave_{30}$ [mm/h]
	average maximum intensity over the last 30 min of storm existence	$Imax_{30}$ [mm/h]
	average standard deviation of intensity over the last 30 min of storm existence	$Isd1_{30}$ [-]
	average standard deviation of intensity groups over the last 30 min of storm existence	$Isd2_{30}$ [-]
	average global velocity over the last 30 min of storm existence	Vg_{30} [m/s]
	average x and y component of the local velocity over the last 30 min of storm existence	Vx_{30}, Vy_{30} [m/s]
	average value of the major and minor axis of the ellipsoid and their ratio over the last 30 min of storm existence	$Jmax_{30}$, $Jmin_{30}$ [km] Jr_{30} [-]
	average major axis orientation of the ellipsoid over the last 30 min of storm existence	Φ_{30} [°]
Target Variables	Total lifetime of the storm	L_{tot} [min]
	Predicted Area and Intensity at LT from +5min to +180min	A_{+LT} [km^2], $Iave_{+LT}$ [mm/h],
	Predicted Velocity X and Y at LT from +5min to +180min	Vx_{+LT}, Vy_{+LT} [m/s]

Here in this study, these two importance analyses are used to determine the most important predictors and their respective weights in the k-NN similarity calculation. For each target variable the most important predictor identified from Pearson Correlation, is given to the PIC metric as the first predictor. The analysis is complex due to the presence of several predictors, 38 states of future behaviour for each target variable (for each 5min between +5min to +180 min lead times), and different times of nowcast; the weights were calculated first for three lead times +15min, +60min and +180 min, and for three storm groups separated according to their duration <60min, 60min-180min, and > 3 hours. Here the averages weights over these groups and lead times are calculated and used as a reference for each importance analysis. The k-NN errors with these average weights are compared in Section 4.1.

3.1.4 Developing the k-NN structure

The structure of the proposed k-NN approach at the storm scale is illustrated at **Figure 5** - left) the current "to-be-nowcasted" storm is shown, while at – right) the past observed storms. First in Step 1, the Euclidean distance between the most important predictors (either present or past predictors), of past storm states and the current one is calculated to identify the most-similar states of the past storms (distance between the blue shapes at left and right side of **Figure 5**):

$$E_d = \sqrt{\sum_{i=1}^{N} w_i \cdot (X_i - Y_i)^2}, \quad (5)$$

where w is the weight of the respective i^{th} predictor, X the predictor of the "to-be-nowcasted" storm, Y the predictor of a past observed storm, N the total number of predictors used and E_d the Euclidian distance between the "to-be-nowcasted" and a past observed storm. The assumption made here is that the smaller the distance, the higher the similarity of future behaviour between the selected storms and the "to-be-nowcasted" storm. Therefore, in ***Step 2*** these distances are ranked in an ascending order and 30 past storm states with the smallest distance are selected (***Step 3***). Once the similar past storm states have been recognized (the blue-shape in **Figure 5** - right), the future states of these storms (the green-shapes in **Figure 5** - right, each for a specific lead time from the occurrence of the selected similar blue-state), are treated as future states (the green-shape in **Figure 5** - left) of the "to-be-nowcasted" storm. In Step 4, either a single or an ensemble nowcast is issued. If a single nowcast is selected, then the green-instances of the k-neighbours are averaged with weights for each lead time:

$$R_{new} = \sum_{i=1}^{k} Pr_i \cdot R_i, \quad (6)$$

where k is the number of neighbours obtained from optimization, R and Pr are respectively the response and weight of the i^{th} neighbour and the R_{new} the response of the "to-be-nowcasted" storm as averaged from k neighbours. Contrary, if a probabilistic nowcast is selected, 30-ensembles are issued independently; to each neighbour a probability is assigned according to their rank with the "to-be-nowcasted" storm:

$$Pr_i = \frac{(1/R_i)}{\sum_{i=1}^{k}(1/R_i)}, \quad (7)$$

where k is the selected number of neighbours and R and Pr are respectively the rank and the probability weights of the i^{th} neighbour. The probability weights calculated here are as well used for computation of the single nowcast in Equation (6). Only neighbours that display a distance lower than 0.5 are selected for both single and ensemble nowcast in order to minimize the influence of non-similar storms.

Since the performance of the single k-NN nowcast is highly dependent on the number of k – neighbours used for the averaging, a prior training is to be done in order to select the right k-neighbours that yield the best performance. The application of the k-NN (and consequently its training) can either be done per each target variable independently, or for all target variables grouped together. In the first approach, the dependency of the target variables between one another is not assured, they are predicted independently from one another. This is referred here as the target-based k-NN and is denoted in the results as VS1. The main advantage of this application is that, since the relationship between the target variables are not kept, new storms can be generated. Theoretically, the predicted variables should have a lower error since the training is done specifically per each variable, nevertheless this approach doesn't say much if similar storms behave similarly. Thus, it is used here as a benchmark for best possible training that can be reached by the k-NN with the current selected predictor set. In the second approach, the relationships between target variables as exhibited by previous storms are kept. The storm structure and the relationship between features are maintained as observed and thus the question if similar storms behave similarly can be answered. This is referred here as the storm-based k-NN and is denoted in the results as VS2. In this study the two approaches are used (respectively called VS1 and VS2) to understand the potential and the actual improvement that the k-NN can bring to the storm nowcast.

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

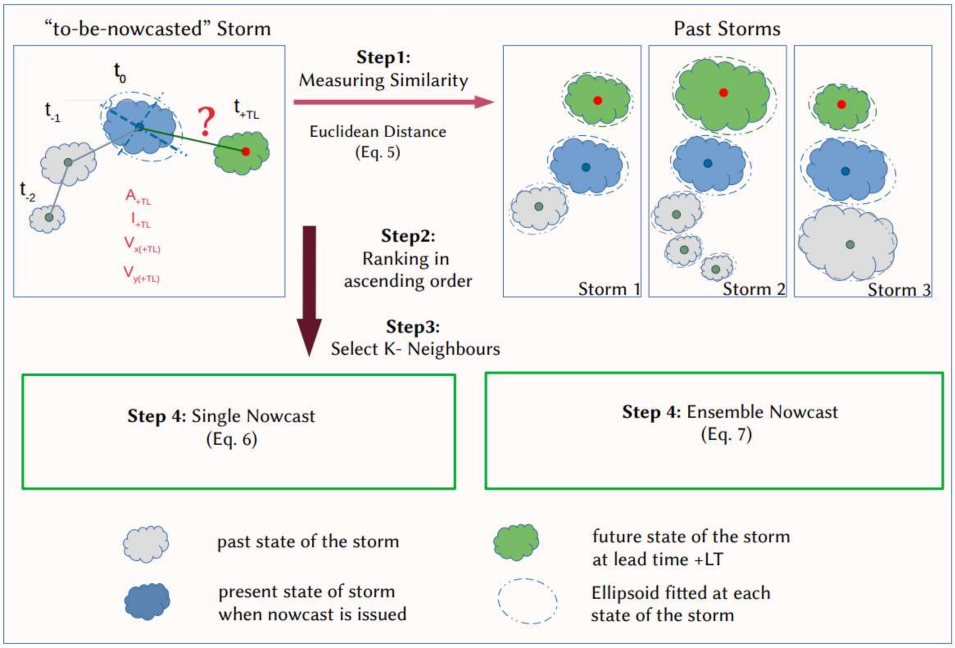


Figure 5 *The main steps involved in the k-NN based nowcast with the estimation of similar storms (Step 1 to 3) and assigning the future responses of past storm as the new response of the "to-be-nowcasted" storm either in a single nowcast (Step4-left) or in an ensemble nowcast (Step4-right).*

3.2 Training of the k-NN and performance assessment

3.2.3 Training the single k-NN nowcast

The training of the k-NN is done based on the 5189 storms extracted from 110 events on a "leave-one-out" cross-validation. Since the "not" matched storms can either be dynamic clutter or artefacts of the tracking algorithm, they are left outside of the k-NN training and validation. The assumption is here that an improvement of the radar data or tracking algorithm would eliminate the "not" matched storms, hence we focus only on the improvement that the k-NN can introduce to the matched storms. "Leave-one-event-out" cross-validation means here that the storms of each event have to be nowcasted by considering as a past database the storms from the remaining 109 events. The objective function is the minimization of the absolute error between predicted and observed target variables at lead times from +5min to +180 min:

$$Error_{target} = \sum_{i=1}^{N} (Pred_{i,+L} \ - Obs_{i,+LT})/N, \qquad (8)$$

where the *Pred* is the predicted response and *Obs* the observed response for the i^{th} storm and $+LT$ the lead time. The results of the storms' nowcast are also dependent on the time of nowcast in respect to the storms' life (time step of the storm existence when the nowcast is issued). If the time of nowcast is 5min, only the present predictors are used for the calculation of storm similarity, and as higher the time of nowcast as more predictors are available for the similarity calculation. It is expected for the nowcast to perform worse at the first 5min of the storm existence, as the velocities are not assigned properly to the storm region and the past predictors are not yet calculated. Therefore, the training is done separately for three different groups of nowcast times, in order to achieve a proper training of the k-NN model: Group 1 – Nowcast issued at 1st timestep of storm recognition, Group 2 – Nowcast issued between 30min to 1 hour of storm evolution, and Group 3 – Nowcast issued between 2 and 3 hours of storm evolution. The k-number with the lowest absolute error averaged over all the events for most of the lead times (as per Equation (8)) is selected as a representative for the

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.
Hydrology and Earth System Sciences
Discussions

single nowcast. For the training, the mean instead of the median is computed over each group in order to account as well for the influence of outliers.

3.2.3 Validating the k-NN single and ensemble nowcast

Once the important predictors are identified and the k-NN has been trained, the performance of both single and ensemble k-NN is assessed also in a "leave-one-event-out" cross-validation mode. Two performance criteria are used to assess the performance: i) absolute error per lead time and target variable (as in the training of the k-NN in Equation (8), and ii) the improvement (%) per each lead time and target variable that the k-NN approach introduces to the nowcast when compared to the Lagrangian persistence;

$$Error_{impr}\ [\%] = 100 \cdot \frac{(|Error_{ref}| - |Error_{new}|)}{|Error_{ref}|}, \tag{9}$$

where the $Error_{new}$ is the error manifested by the k-NN, the $Error_{ref}$ the error manifested by the Lagrangian persistence and the $Error_{impr}$ the improvement in reducing the error per each lead time. For the ensemble application of the k-NN method additional criteria were employed. For each storm the number of time steps where the observed target variable was within the range of the ensemble nowcast was computed in order to give an idea how effective the ensemble range is depending on the lead time. Moreover, the number of ensembles that have a smaller error than the Lagrangian persistence were computed for each lead time and target variable.

As stated earlier the results depend on the time of nowcast and also storm duration (in regard to available storms). Therefore, the performance criteria for both single and ensemble nowcast were computed separately for different storm durations and time of nowcasts as illustrated in **Table 2**. It is important to mention as well, that since one event may contain many storms of similar nature, when leaving one event out for the cross-validation, the number of available storms is actually lower than the numbers given in **Table 2**. This is particularly affecting the performance of the storms longer than 6 hours, as the "leave-one-event-out" cross-validation causes fewer available storms for the similarity computation.

Table 2 *The selected storm durations and time of nowcast for the performance calculation of the single and ensemble nowcast and the respective number of storms for each case.*

Storm living shorter than 30 min		Storms living within 0.5 - 3 hours		Storms living longer than 3 hours	
Time of Nowcast	No. Storms	Time of Nowcast	No. Storms	Time of Nowcast	No. Storms
5 min	4106	5 min	994	5min	89
15 min	2265	1h	370	2h	89
30 min	271	3h	6	6h	33

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.
Hydrology and Earth System Sciences
Open Access
EGU
Discussions

4. Results:

4.1 Predictors Importance Analysis

Figure 6 illustrates the results of the two important analysis methods (Pearson correlation and partial information correlations - PIC) for each of the target variable and their average over the 5 variables. The stronger the shade of the green colour, the more important is the predictor for the target variable. The weights given here are averaged from the weights calculated at three different lead times and storm durations. First the Pearson Correlation weights were advised for the identification of the most important predictors. From the results it is clear that the autocorrelation has a higher influence, as the target variables are mostly correlated with their respective past and present values. This influence logically is higher for the shorter lead times and smaller for the longer lead times. For longer lead times the importance of other predictors, that are not related directly with the target variable, increases. Similar patterns can be observed among the Area, Intensity and Total Lifetime target variables, indicating that these three variables may be dependent on each other, and on similar predictors like: current lifetime, area, standard deviation of intensity, the major and minor ellipsoidal axis and the global velocity. On the other hand, are the velocity components, which seem to be highly dependent on the autocorrelation and slightly correlated to area and ellipsoidal axes. It has to be mentioned that apart for the standard deviation intensities also the mean, median, and maximum spatial intensities were investigated. Nevertheless, it was found that the I_{sd1} and I_{sd2} had the higher correlation weights, and since there is a high collinearity between these intensity predictors, they were left out of the predictor's importance analysis.

Method	Target	Present Predictors												Past Predictors - averaged from last 30 min									
		Cells	L_{now}	A	PI_{sd1}	PI_{sd2}	V_g	V_x	V_y	J_{max}	J_{min}	J_r	Φ	A	PI_{sd1}	PI_{sd2}	V_g	V_x	V_y	J_{max}	J_{min}	J_r	Φ
Pearson Correlation	A	0.09	0.18	0.67	0.15	0.48	0.05	0.00	0.00	0.50	0.49	0.09	0.00	0.65	0.17	0.00	0.07	0.00	0.06	0.51	0.49	0.12	0.00
	I	0.00	0.07	0.11	0.36	0.14	0.04	0.00	0.00	0.12	0.12	0.00	0.04	0.10	0.33	0.13	0.00	0.00	0.05	0.12	0.11	0.05	0.04
	Vx	0.00	0.00	0.10	0.02	0.04	0.16	0.21	0.00	0.08	0.00	0.00	0.03	0.09	0.00	0.00	0.18	0.28	0.00	0.09	0.00	0.00	0.00
	Vy	0.00	0.05	0.00	0.00	0.05	0.00	0.00	0.15	0.04	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.04	0.22	0.05	0.04	0.00	0.00
	L_{tot}	0.00	0.11	0.36	0.10	0.22	0.09	0.00	0.00	0.22	0.20	0.05	0.05	0.34	0.00	0.21	0.10	0.00	0.00	0.22	0.20	0.08	0.07
	Average	0.00	0.08	0.25	0.13	0.18	0.07	0.10	0.10	0.19	0.16	0.05	0.04	0.24	0.10	0.08	0.07	0.10	0.10	0.19	0.17	0.05	0.02
Partial Information Correlation	A	0.00	0.08	0.15	0.00	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.33	0.00	0.07	0.00	0.00	0.33	0.00
	I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Vx	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	Vy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	L_{tot}	0.00	0.15	0.13	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.00	0.00	0.00	0.11	0.33	0.00
	Average	0.00	0.05	0.06	0.00	0.00	0.09	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.20	0.01	0.20	0.02	0.13	0.00

Figure 6 *Strength of relationship between the selected predictors and the target variables based on two predictors identification methods: upper –correlation, and lower –PIC weights. The green shade indicates the strength of the relationship: with 0 for no relationship at all, and 1 for highest dependency.*

The application of the PIC analyses requires that the most important predictors should be introduced to the analysis first. Hence based on the Pearson correlation values the following most important predictors were selected: Area –A, Intensity –PI_{sd1}, - Velocity X – Vx_{30}, Velocity Y –Vy_{30}, Total Lifetime – A. The results of the PIC analysis are shown in the lower row of **Figure 6**. For storm duration lower than 3 hours, where a lot of zeros are present, the PIC methods seems to be unable to converge to stable results or to identify important predictors. For the intensity and velocity components, the PIC identifies only 1 important predictor which, in the case of the Intensity and Velocity in the Y direction, does not correspond with the most important predictor fed first in the analysis. Contrary for Total Lifetime and Area, only for storms that last longer than 3 hours, the method is able to converge and give the most important predictors; for Area - A, Vg, past Vy_{30} and the L_{now}, while for Total Lifetime - A, Vel_g, L_{now} and $Jmin_{30}$. At the moment it is unclear why the PIC method is unable to perform well for all of the target variables and storm groups. One reason might be that only the Area and Total Lifetime are dependent on the chosen target variables. Another reason might be that for the other target variables the heavy-tail of the probability distribution and the high zero sample size may influence the calculation of the joint and mutual probability distribution. The reason why the method is performing poorly for the application at hand, even though developed specifically for the k-NN application, is not completely understood and is not investigated further on for the time being since it is outside the scope of this paper.

Overall, the results from the Pearson correlation seem more robust and stable (throughout the lead times and storm groups) than the PIC method; the importance weights increase with the lifetime of the storm and decrease with higher lead time. These behaviours are expected as with increasing lead time the uncertainty becomes bigger and with increasing lifetime the storm dynamic becomes more persistent (due to the large scales and the stratiform movement involved). Moreover, the important predictors do not change drastically from one lead time or storm group to the other, as seen in the PIC. Therefore, the predictors estimated from the correlation with the given weights in **Figure 6** are used as input to the k-NN application. In order to make sure that the predictor set from the Pearson correlation was the right one, the improvement in the single k-NN training error of using these predictors instead of the ones from PIC are shown in **Figure 7**. The results shown in this figure are computed according to the Equation (9) (where "new" is k-NN with correlation weights, and "ref" is the k-NN with PIC weights) for the target-based k-NN approach (solid lines) and storm-based k-NN approach (dashed lines) and are averaged for three groups of nowcast times as indicated in the training of k-NN (Section 3.2.3) and as well in the legend of **Figure 7**.

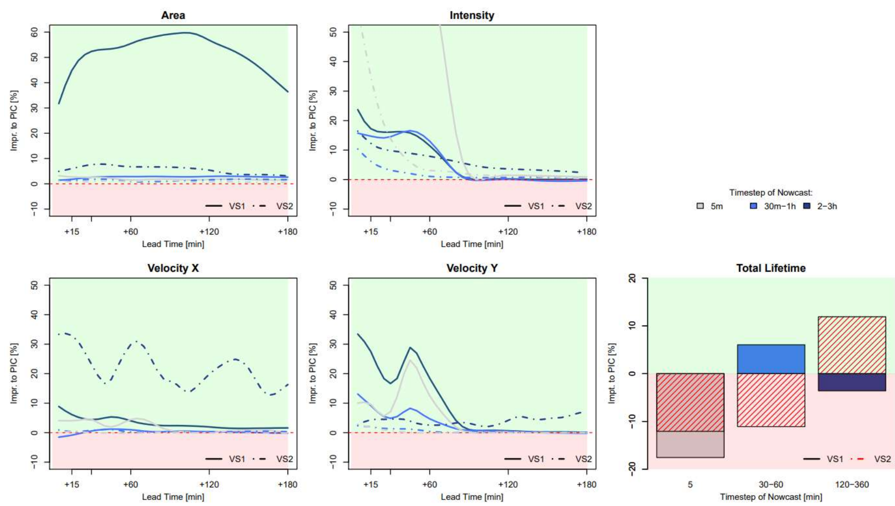


***Figure** 7 The average error improvement per lead time and target variable from applying the k-NN (VS1 target-based, VS2 storm-based) with the predictors and weights derived by the Pearson correlation instead of PIC. The improvements are averaged for different times of nowcast. The green plot region indicates a positive improvement of the correlation predictors in comparison to the PIC, and the red region indicates a deterioration.*

The results from **Figure 7** indicate that for the Area, Intensity, and Velocity components, the Pearson correlation weights improve the performance of target-based k-NN up to 30% compared to the PIC weights. This happens mainly for the short lead times throughout the three groups of nowcast times. For longer lead times there seems to be no significant difference between the predictors sets. Nevertheless, here the mean over the grouped storms is shown to illustrate the influence of the outliers. In the case of the median, the Pearson correlation has the clear superiority compared to the PIC predictors set. The same cannot be said for the Total Lifetime as a target variable, here not always the Pearson correlation weights give the best results for all the nowcast times. In fact, here the k-NNs based on the PIC weights seem to be more appropriate and yielded better results. However, as the other 4 target variables are better for the Pearson correlation, this

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

predictor set was selected for all applications of the k-NN (with different weights according to **Figure 6**) to keep the results consistent with one another. A further analysis was done that proved that the application of the correlation weights produces lower errors than the non-weighted k-NN application (all weights are assigned to 1 to the most important predictors from Pearson correlation).

4.2 Training of the single k-NN nowcast

Once the most important predictors and their weights are determined, the training of the single k-NN nowcast for the two k-NN applications (storm-based and target-based) was performed. The optimal k-value obtained from minimizing the absolute error produced by k-NN are shown in **Figure 8**. The results are averaged for the given nowcast times (see legend), each lead time and target variables for both k-NN applications (VS1 target-based and VS2 storm-based).

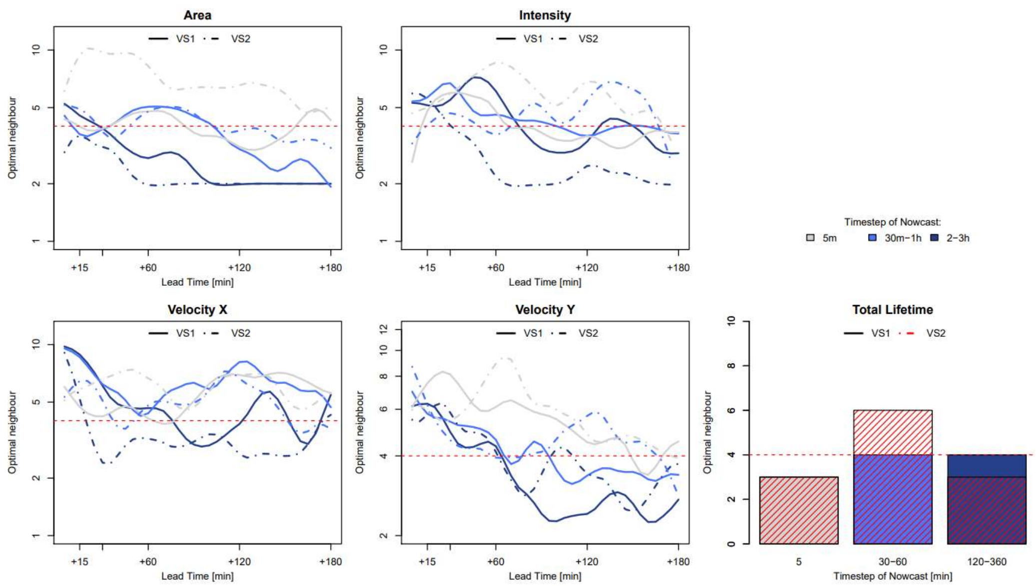


Figure 8 *The training of the k-NN per target variable based on predictors and weights derived from Pearson correlation analysis: the optimal selected "k" neighbours yielding the lowest absolute errors. Two k-NN applications are shown here – VS1 in solid line and VS2 in dashed line. The results of the training are averaged for given group of nowcast times. The red dashed horizontal line in the second row indicates the k=4 that is chosen in this study for the k-NN application.*

Overall, it seems that averaging between 2 to 10 neighbours give the best results depending on the lead time, and there is a clear decreasing trend of neighbours with increasing lead time. The best achieved k-numbers from the two k-NN applications are different from one another at some lead times, nevertheless they seem to converge around k=3 or k=4 neighbours. Here for both application the k=4 was selected (indicated with red dashed horizontal line in **Figure 8**) as a better compromise between different lead times, nowcast times and target variables.

4.3 Results of the single 4-NN nowcast

The absolute errors of the 4-NN determinist nowcast run for both target- and storm-based approaches are shown in **Figure 9** for each lead time and target variable. The results are grouped according to the storm duration; i) upper row – for storms that lived 30min, ii) middle row – for storms that lived up to 3 hours and iii) lower row – for storms that lived longer than 3 hours, and are averaged per nowcast times given in **Table 2**. To getter a better overview of the majority

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.
Hydrology and Earth System Sciences
Discussions

of storms, the median results were shown here instead of the mean ones. As shown as well in the training of the 4-NN, the target-based k-NN exhibits lower Area, Intensity and Velocity errors than the event-based 4-NN.

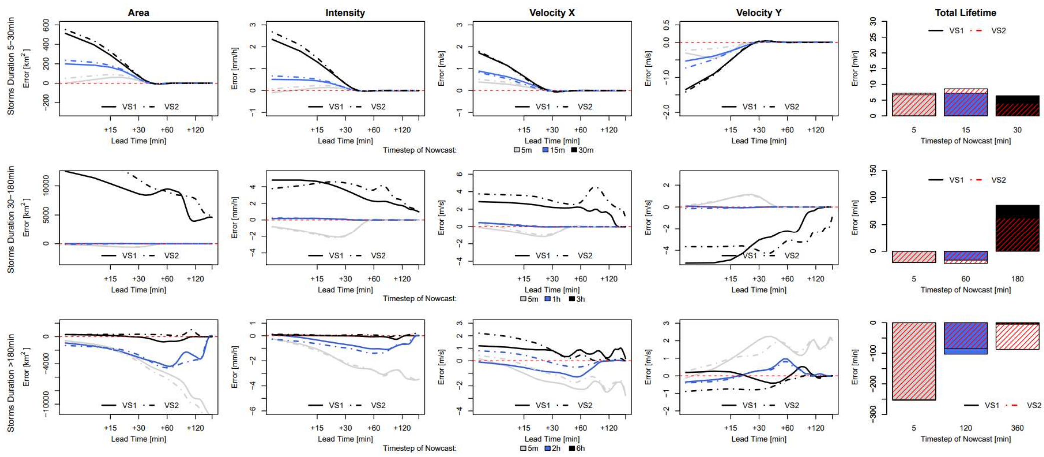


Figure 9 *The median error for each target variable (Area, Intensity, Velocity in X and Y direction and Total Lifetime) based on two 4-NN applications: -VS1 in solid and VS2 in dashed lines. The median errors are computed over storms that are: shorter than 30 min (upper row), than 3 hours (middle row), and longer than 3 hours (lower row), and over given times of nowcast.*

For storm living less than 30 minutes, the error is decreasing with the lead time and past LT+30 min is zero, as the deaths of the storms have been captured successfully. The Total Lifetime of the majority of the storms can be captured with only 5 min overestimation regardless of the nowcast time. The errors for the 4 target variables (except Total Lifetime) are lower for the earlier nowcast times than for the later ones. This is explained by the sample size, as with increasing nowcast time, the sample size becomes smaller and thus 4-NN may not find the suitable neighbours. For very short lead times in these storms (up to LT+15min), the errors of event-based are between 10% (for Area, Intensity and Total Lifetime) to 20% (for Velocity components) higher than the target-based 4-NN. For Area and Intensity, the errors are consistently higher than the target-based, however for the Total Lifetime and Velocity components there are certain nowcast times and lead times where the errors from the storm-based are up to 50% lower. Past 30 min lead times there is no difference between the two 4-NN approaches as both of them predict the death of storms correctly.

For the storms living up to three hours, the same behaviour is, more or less, observed. The only difference is for nowcasts issued at the 36th timestep of storm existence. Here it is clear that the 4-NN fails to capture the death of storm that live exactly three hours, however this is attributed to the number of available storms with duration of 3 hours (median over 6 storms available). Since the Area, Intensity and Total Lifetime are overestimated and not converging to zero for high lead times, it is clear that the nearest neighbours are being selected from the longer storms that do not dissipate within the next 3 hours. The differences between the two 4-NN approaches are visible for lead times up to 2 hours (except the nowcast at 36th time step of the storms life), afterwards the errors are converging to zero for the 4 target variables. The storm-based 4-NN produces circa 10% lower errors than the target-based one for the nowcast times lower than 30min, while for later nowcast times the errors are clearly higher than the target based one (up to 100% higher). As the sample size is the same for both approaches, it seems like storm-based may be more appropriate at the beginning of the storm's life and that these storms behave more similarly at the first 30 minutes of their evolution rather than in their later life.

For the storms that live longer than 3 hours (under 100 storms available) the same problem, as in the nowcast issued at the 36th time step of the previous storms, is present. The Total Lifetime is clearly underestimated (up to 100min)

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.
Hydrology and Earth System Sciences
Open Access
Discussions

as due to database the information is taken from shorter storms. It is important to notice here, that although 70 storms are present, because of the "leave-one-event-out" validation, the storm database is actually smaller. Nevertheless, the error is manifested here differently: as the long storms are more persistent in their features: The Area, Intensity and Velocity components are captured better for the short lead times with the error increasing at higher lead times. Here as well the nowcast issued at the earlier stages of the storm's life exhibit higher errors than in the later stages. Especially for the nowcast at the 6th hour of storms existence, the errors are quite low for all 5 target variables due to the persistence of the stratiform storms. For this group of long storms, the storm-based nowcast yields errors from 0 up to more than 100% higher than the target-based one, with only few exceptions depending on the time of nowcast and variable. It is clear that the storm-based 4-NN is more influenced by the number of available storms than the target-based approach.

Figure 10 shows the improvement that the 4-NN introduces to the nowcast when compared to the Lagrangian persistence (either target- or storm-based) and are averaged per lead time for each of the three group of storms and the respective times of nowcast. Since the Lagrangian Persistence doesn't issue a Total Lifetime nowcast, only the four target variables (Area, Intensity and Velocity components) are considered. The green area indicates the percent of improvement from the application of the 4-NN approach, and the red area indicates the percent of deterioration from the 4-NN application (Lagrangian persistence is better).

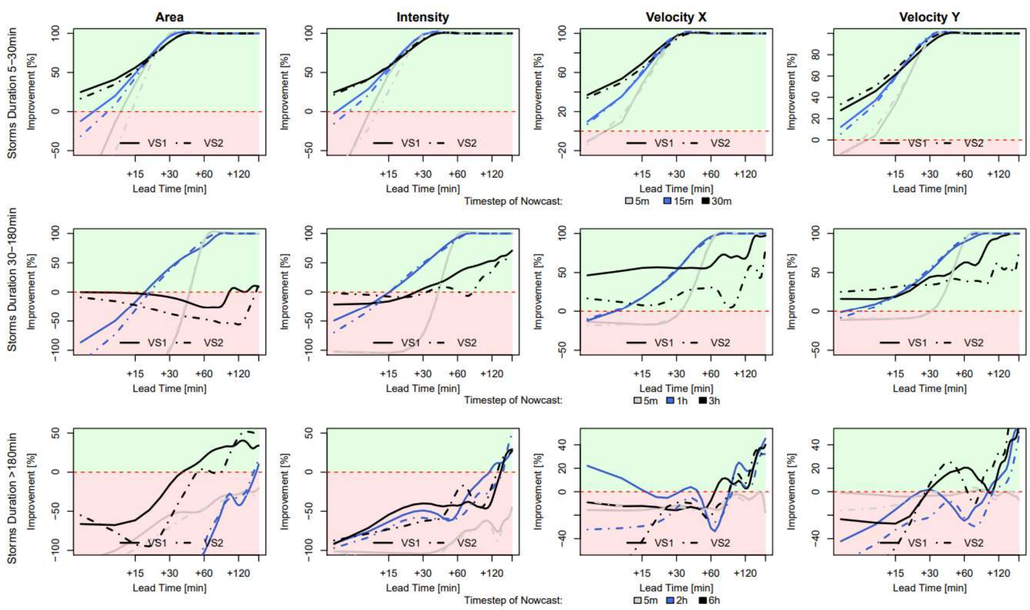


***Figure 10** The median improvements that the single 4-NN nowcast can introduce in the nowcast of the target variables (Area, Intensity, Velocity in X and Y direction) in comparison to the Lagrangian persistence. The results are shown for each 4-NN application: VS1 in solid and VS2 in dashed lines and are calculated separately for storms that live shorter than 30 min (upper row), shorter than 3 hours (middle row) and longer than 3 hours (lower row), and for the respective times of nowcast. The green region of the plot indicates a positive improvement (better nowcast by the 4-NN application) and the red region indicates a deterioration (better nowcast by the Lagrangian persistence).*

For the 30min storms, the 4-NN approach (both target- and storm-based) are considerably better than the Lagrangian persistence: improvement is higher than 50% from the LT+15min and up to 100% from LT+30min. The improvement is greater for nowcast at 3rd timestep of storm existence (when the persistence predictors are considered). It

Hydrology and Earth System Sciences
Open Access
Discussions

is clear than due to the autocorrelation, the Lagrangian persistence is more reliable for the short lead times and for earlier times of nowcast. However, after 15 min lead times and for times of nowcast near to the dissipation of the storms, non-linear relationships govern, and hence the improvement from the nearest neighbour are more significant. The target-based 4-NN results in higher improvements than the storm-based one only for lead time up to 30min, with storm-based improvements being 10-40% less than the target-based. Past these lead time the improvements are for both 100%.

For the storms that live between 30 min to 3 hours, the improvements are introduced first after LT+15 or +30 min depending on the time of nowcast: with increasing time of nowcast increases the improvement as well. The only exception is for the nowcast of Area on the 36th timestep of the storm existence, where no clear improvement of the 4-NN approaches could be seen before LT+2h. However, this low improvement for the nowcast issued at the 36th timestep of storms life was expected following the poor performance of the 4-NN shown in **Figure 9**. Regarding the difference of the two 4-NN approaches, with few exceptions, the storm-based nowcast introduces 5-40% less improvements than the target-based. Another exception is the nowcast at the 36th time step, where the storm-based improvements are clearly lower, especially for the higher lead times, than the target-base (up to 100% or more).

For storms living longer than 3 hours, the improvements are present for lead times higher than 2 hours. Since the features of the long storms (mostly of stratiform nature) are persistence in time, is understandable for the Lagrangian Persistence to deliver better nowcast up to LT+2h. Past this lead time non-linear transformations should be considered. Here, even though the storm database is small, the non-linear predictions based on the 4-NN capture better these transformations than the persistence. The improvement introduces by the storm-based are generally from 20-100% lower than the improvements introduced from the target based.

To conclude, the 4-NN single nowcast brings up to 100% improvements for lead times higher than the predictability limit of the Lagrangian persistence and are dependent mainly on the storm type and the size of database. Overall for all of the storms the improvement is mainly at the high lead times and later times of nowcast, as the k-NN is capturing particularly well the death of the storms. The results from the long events are suffering the most from the small size of the database. This was anticipated, as the events were mainly selected from convective and mesoscale convective events that have the potential to cause urban floods. A bigger database, with more stratiform events included, will introduce a higher improvement to the Lagrangian persistence. These improvements are expected to be higher for lead times longer than 2 hours, but is yet to be seen if a larger database can as well behave better than the persistence even for lead times shorter than the predictability limit. Regarding the two different 4-NN approaches, the storm-based performs around 20% worse than the target-based nowcast, introducing generally 40% lower improvements to the Lagrangian persistence. These values are valid mostly for the first 4 target variables and not for the Total Lifetime. Regarding the Total Lifetime, both of these approaches deliver more or less same results, indicating that similar storms have similar life times. The main differences between these two approaches lie between the growth/decay processes, which the target-based 4-NN can capture better. Nevertheless, it has to be mentioned that the target-based approach is profiting more from the selected predictors and their respective weights. A more suitable predictor set and weights, may actually improve the results of the storm-based 4-NN considerably.

4.4 Results of the ensemble 30-NN nowcast

Figure 11 illustrates the minimum error achieved from the best ensemble member of 30 most similar storms for the "to-be-nowcasted" storm. The results are shown as in the previous Figures per each lead time and target variable, for storms divided into 3 groups according to their duration and averaged depending on the time of nowcast. For the 30min long storms, the errors of the best ensemble are typically lower than the single 4-NN nowcast for all the variables, lead times and time of nowcasts. Here only the nowcasts issued at the 1st timestep of storm existence have errors slightly higher than zero for short lead times (up to LT+15min), apart from that, regardless of the 30-NN approach, all errors are zero. This suggests once again that storms in this duration behave similarly and their response can be predicted adequately by similar neighbours.

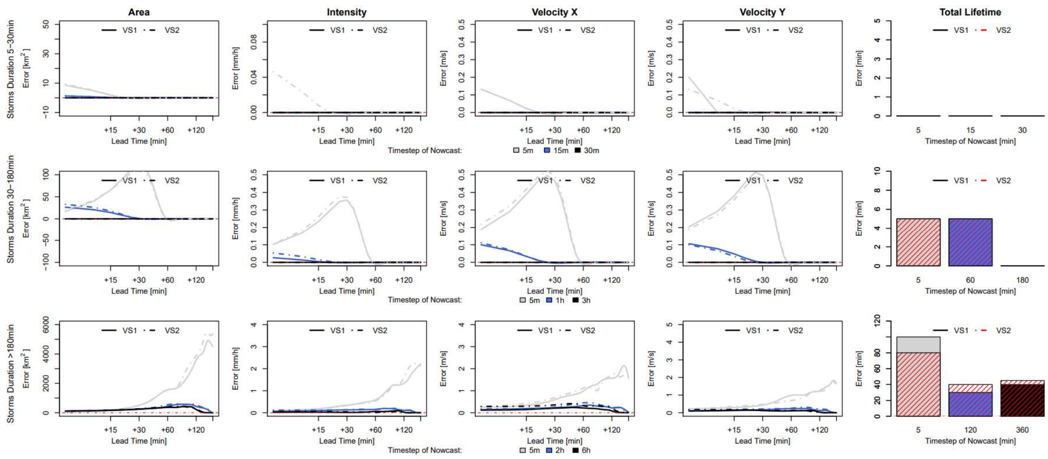


Figure 11 The median of absolute errors of the best issued ensemble member (best possible nowcast) for each of the target variables (Area, Intensity, Velocity in X and Y direction and Total Lifetime) based on two 30-NN application: VS1 shown with solid and VS2 with dashed lines. The median errors are calculated over storms that lived shorter than 30 min (upper row), shorter than 3 hours (middle row), longer than 3 hours (lower row), and for the respective times of nowcast.

For storms that live shorter than 3 hours, the error of the best ensemble member is decreasing with increasing lead time and timestep of nowcast. The difference between the target- and storm-based nowcasts is within the range of the single 4-NN nowcast for the first 4 target variables, with storm-based ensemble having 10%-30% higher errors in the first 30 min of the nowcast than the target-based. For the Total Lifetime, both of the ensembles exhibit very similar errors for most of the nowcast times. It is worth mentioning here, that for the nowcast at the 36th time step of storms' existence the errors are much lower than the single 4-NN nowcast. This proves that the most similar storm is within the 30 members, but not within the first 4 neighbours selected in the case of the single 4-NN nowcast.

Due to the unrepresentativeness in the database, the errors of the longer storms are considerably higher than the other storm groups, and the errors of the first 4 target variables are increasing with the lead time and decreasing with the nowcast time. These results correspond to the ones from the single 4-NN nowcast. However here unlike the other storm groups, the differences between the storm-based and target-based approach are visible past 30 min lead time, with the storm-based errors being 15-35% higher than the target-based. As the best ensemble is between the 30 most similar storms (with zero errors for shorter lead times), then the given predictors set is failing to capture the most similar storms within the 4 closest storms (or the rank average of the 4 closest storms is not the best solution possible). This is understandable as the predictor's weights and the training of the k-NN was focused mainly on shorter storms.

Overall the ensemble results are better than the single 4-NN nowcast, suggesting that the best responses are obtained by singular neighbours (either the closest one or within the 30 neighbours) and not by averaging. Thus, there is still room for improving the single 4-NN nowcast by selecting better the important predictors and their weights or averaging differently the nearest neighbours. However, the results from **Figure 11** emphasize that similar storms do behave similarly, as the error is almost zero, and that the developed k-NN on the given database with 30 ensembles gives satisfactory results.

To investigate the use of the ensemble spread in the nowcast, in **Figure 12,** the percent of time steps, where the observed values fell within the ensemble range, is calculated for each storm duration group, nowcast time and lead time.

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

For the short storms (duration < 30 min), almost at 100% of the time steps, the observed target values fall within the ranges of the ensemble 30-NN. This value decreases slightly for storms with duration up to 3 hours, but still is higher than 80%. However, for the long storms (longer than 3 hours), the range of the ensembles captures the observed value better for shorter lead times and for longer times of nowcast. For longer lead times and early times of nowcast, more than 50% of the time steps are representing adequately the observed target variables. While the ensemble range is satisfactory for the short storms, improvements should be done, either by increasing the database of stratiform events or selecting different predictors, in order for the ensemble range to represent the observed target variables adequately. There is hardly any difference between the storm-based and the target-based nowcast: for short storms (duration shorter than 3 hours) independent of the time of nowcast and variables, at the storm-based the number of time steps are less than 1% fewer than the ones from the target-based, and for longer storms less than 5%. This suggests that the suitability of the ensemble range does not depend on the k-NN approach but mostly on the past storms available.

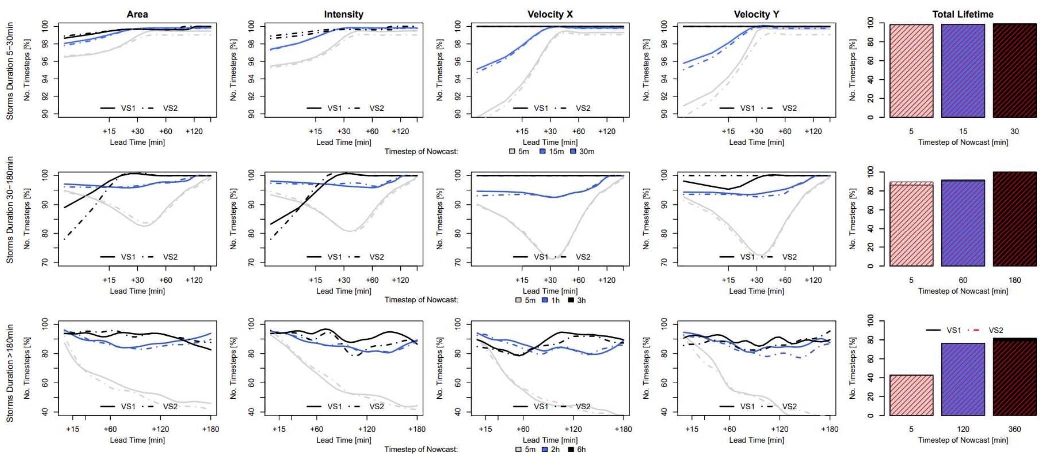


__Figure 12__ The percent of time steps when the observed target variable (Area, Intensity, Velocity in X and Y direction and Total Lifetime) is within the range of the ensembles for two 30-NN applications – VS1 in solid and VS2 in dashed lines. The percent of time steps is calculated for storms that lived shorter than 30 min (upper row), shorter than 3 hours (middle row) and longer than 3 hours (lower row), and for the respective nowcast times.

Figure 13 demonstrates the number of ensembles that yielded a better nowcast than the Lagrangian persistence (better ensembles). This percent of ensembles is computed and shown for each storm duration group, time of nowcast and lead time. For all the three groups it is visible that the number of better ensembles increases considerably with the lead time – suggesting that the ensemble predictions are particularly useful for the longer lead times where the single nowcast is not able to capture the storm evolution. For short storms (duration shorter than 30min) the number of ensembles is low for lead times up to 30 min and in this range the ensembles are worse for the early times of nowcasts. However, past this lead time, the number of better ensembles is more than 80 % (24 ensembles) with no clear difference between different times of nowcast. This coincides with the predictability limit of the Lagrangian persistence at such scales. Thus, it makes sense that the ensemble nowcasts behave better after the predictability limit of the persistence is reached. Moreover, for these storms the difference between the two types of 30-NN is insignificant (less than 1% for all target variables and times of nowcasts).

For storms that live shorter than 3 hours, the results are slightly worse than the very short storms. Here as well the number of better ensembles increases drastically for all the target variables between LT+15min to LT+30min. Interesting in this storm group are the results from the nowcast at the 3 hours of storm existence that exhibit different behaviours than the other nowcast times. However, this is expected as the Lagrangian persistence performs particularly poorly because it cannot model the storms deaths. Here as well the difference between the two types of 30-NN is insignificant, although a bit higher than for the very short storms (~2.5% difference).

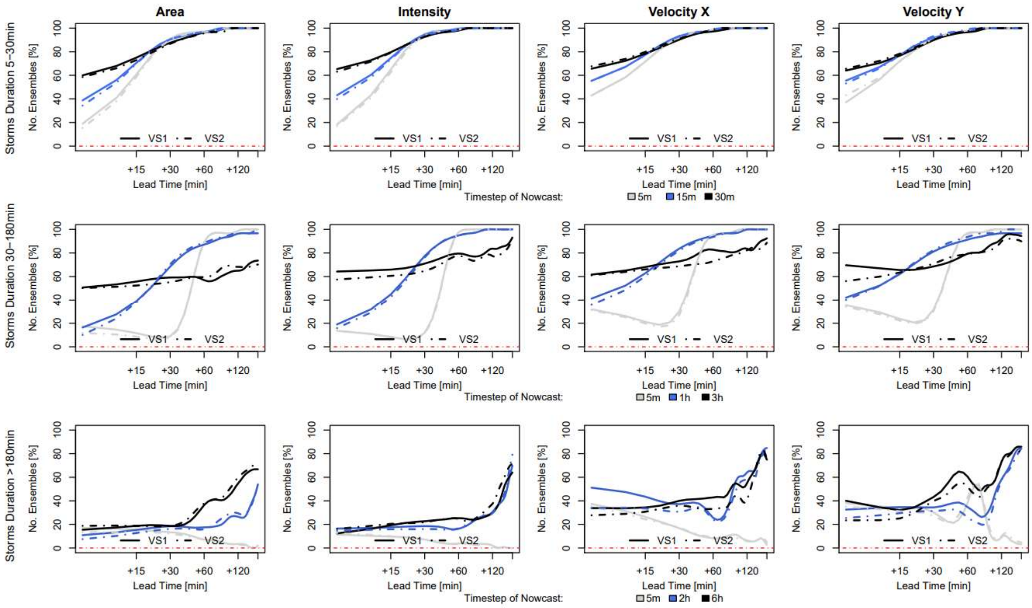


Figure 13 *The percent of ensembles for each of the two 30-NN applications (VS1 in solid and VS2 in dashed lines) that yielded better nowcasts than the Lagrangian persistence based on each target variable (Area, Intensity, Velocity in X and Y direction). The percent of ensembles is calculated for storms shorter than 30 min (upper row), shorter than 3 hours (middle row), longer than 3 hours (lower row), and for the respective times of nowcast.*

For the longer storms the percent of better ensembles is increasing with the time of nowcast and are increasing mainly for LT+45min to LT+60min, but still not as high as in the other storm groups. The worse performance is at nowcasts at the 1st time step of the storm where the percent of better ensembles is quite low (between 1 and 0 ensembles) for the LT+180min for all of the target variables. What is interesting from these storms, is that the percent of better ensembles is higher at the Velocity components than in the Area and Intensity predictions. This suggest the velocity components are more persistent (see **Figure 3**) and easier to be predicted from similar storms. Still it is worth mentioning that the percent of better ensembles is almost never zero. Even with a small database for the long storms, the 30-NN can recognize 1-5 similar past storms that can give useful information in improving the nowcast when compared to the Lagrangian persistence.

5. Conclusions

Accurate predictions of rainfall storms at fine temporal and spatial scales based on radar data are quite challenging to achieve. The errors associated with the radar measurements, identification and tracking of the storms, and more importantly the extrapolation of the storms in the future based on the Lagrangian persistence, are limiting the forecast horizons of such radar based nowcasts to 30-45 min for convective storms and to 1 hour for stratiform events. The focus of this paper was the improvement of the storm-oriented radar based nowcasts by considering other non-linear behaviour for future extrapolation instead of the Lagrangian persistence. For this purpose, a nearest neighbour approach was proposed that predicts future behaviours based on past observed behaviours of similar storms. The method was

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

developed and validated for the Hannover Radar Range where storms from 110 events were pooled together and used in a "leave-one-event-out" cross-validation. From 110 events a total of around 5200 storms with different morphology were identified and tracked in order to build up the database for the k-NN implementation. The storms were treated as ellipses and for each state of the storms' evolution different features (describing both present and past states) were computed. The k-NN approach was developed on these features to predict the behaviour of the storms in the future (for lead times up to 3 hours) through 5 target variables (Area, Intensity, Velocity in X and Y direction and Total Lifetime).

First an importance analysis was performed in order to recognize the most important predictors for each of the target variable. Two different approaches were employed for this purpose: Pearson correlation, and Partial Information Correlation (PIC). A comparison of these two methods revealed that for the application at hand the Pearson Correlation is more reliable at determining important predictors, and delivers 5%-30% better results than the PIC method. However, the PIC seems promising mainly for determining the most important predictors of the Area and Total Lifetime for storms longer than 3 hours, and is still recommended to investigate for further works. The Area, Intensity and Total Lifetime of the storms seem to be co-dependent on one another and on the features that describe their evolution. In particularly the variance of the spatial intensity is an important predictor for the three of them. On the other hand, the velocity components are dependent as well more on features that describe their evolution. Nevertheless, there is still a dependency of the area and velocity components, and should be included when predicting each other.

The weights derived from the Person correlation were used for the similarity estimation of different storms based on the Euclidian distance. Two k-NN approaches were developed on two measurement of similarity: a) target-based approach – similarity was computed for each target independently and indicates the best performance possible by the given predictors and weights, and b) storm-based approach – similarity was computed for each storm keeping the relationship between the target variables. For the two approaches a single (averaging the 4 closest neighbours) and an ensemble (with 30 nearest neighbours) nowcast were issued for all of the storms in "leave-one-event-out" cross-validation mode. In the single nowcast the difference between the two lied mainly at short lead times (up to 30 min) with the event-based results yielding 10-30% higher errors than the target-based ones. Exception was the Total Lifetime where the storm-based prediction was almost the same as the target-based approach. However, at higher lead times the difference between the two became insignificant, as the death processes was captured well for the majority of the storms. The same behaviours were observed as well in the ensemble nowcast, with target-based ensembles being slightly better than the storm-based nowcast.

To investigate what value each of the two k-NN approaches introduces to the nowcast, their errors (for both single and ensemble nowcast) were compared to the errors produced by the Lagrangian persistence. For both of the approaches the improvement was up to 100% for convective storms for lead times higher than 15 min, and up to 50% for mesoscale storms for lead times higher than 2 hours. The results were particularly good for the small convective storms due to the high number of storms available in the database. For the mesoscale storms (with duration longer than 3 hours) the improvements were not satisfactory due to the small sample size of such long storms. An increment in the sample size is expected to improve the performance of the k-NN for these storms as well. However, when consulting the ensemble k-NN application it seems that, even for these storms and the given database, there are at least 5-10 ensemble members that are better than the Lagrangian persistence. This emphasizes not only the importance of the ensemble nowcast in comparison to the single one, but also the importance of nearest neighbour method in its potential to improve the nowcast.

Overall the results suggest that if the database is big enough, storms that behave similarly can be recognized by their features, and their responses are useful in improving the nowcast up to 3 hours lead times. Further improvements can be achieved if the predictors importance is estimated better (i.e. Monte Carlo approach, or neural networks) or if additional predictors are included from other data sources like satellite data, Numerical Weather Prediction Models etc.

© Author(s) 2021. CC BY 4.0 License.

A different averaging of the neighbours, either different weights or k-neighbours, may as well improve performance and match the results of the single nowcast with the ensemble one at least for the short lead times. In conclusion, the results seem promising at the storms scale, nevertheless is still to be seen if the methodology applied here can introduce improvements as well at the local scale, i.e. validation with the measurements from the rain gauge observations.

6. Funding

This research was funded by the German Federal Ministry of Education and Research (BMBF) grant number Förderkennzeichen 03G0846B.

7. Acknowledgements

The results presented in this study are part of the research project "Real-time prediction of pluvial floods and induced water contamination in urban areas (EVUS)", funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung BMBF) who are gratefully acknowledged. We are also thankful for the provision and right to use the data from the German National Weather Service (Deutscher Wetterdienst DWD).

8. References

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, a., Kohler, O., Gerlach, N., & MeteoSolutions GmbH. (2004). *Projekt RADOLAN - Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer)*. 111.

Berenguer, M., Surcel, M., Zawadzki, I., Xue, M., & Kong, F. (2012). The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: Intercomparison among numerical models and with Nowcasting. *Monthly Weather Review*, *140*(8), 2689–2705. https://doi.org/10.1175/MWR-D-11-00181.1

Berndt, C., Rabiei, E., & Haberlandt, U. (2014). Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *Journal of Hydrology*, *508*, 88–101. https://doi.org/10.1016/j.jhydrol.2013.10.028

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*, 2127–2155. https://doi.org/10.1256/qj.04.100

Codo, M., & Rico-Ramirez, M. A. (2018). Ensemble radar-based rainfall forecasts for urban hydrological applications. *Geosciences (Switzerland)*, *8*(8), 297. https://doi.org/10.3390/geosciences8080297

Dixon, M., & Wiener, G. (1993). TITAN: thunderstorm identification, tracking, analysis, and nowcasting - a radar-based methodology. *Journal of Atmospheric & Oceanic Technology*, *10*(6), 785–797. https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2

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. https://doi.org/10.5194/hess-20-505-2016

Galeati, G. (1990). A comparison of parametric and non-parametric methods for runoff forecasting. *Hydrological Sciences Journal*, *35*(1), 79–94. https://doi.org/10.1080/02626669009492406

Gallus, W. A., Snook, N. A., & Johnson, E. V. (2008). Spring and summer severe weather reports over the midwest as a function of convective mode: A preliminary study. *Weather and Forecasting*, *23*(1), 101–113. https://doi.org/10.1175/2007WAF2006120.1

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.

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

https://doi.org/10.1175/1520-0450(2004)043<0074:SDOTPO>2.0.CO;2

Germann, U., Zawadzki, I., & Turner, B. (2006). Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *Journal of the Atmospheric Sciences*, *63*(8), 2092–2108. https://doi.org/10.1175/JAS3735.1

Goudenhoofdt, E., & Delobbe, L. (2013). Statistical characteristics of convective storms in belgium derived from volumetric weather radar observations. *Journal of Applied Meteorology and Climatology*. https://doi.org/10.1175/JAMC-D-12-079.1

Grecu, M., & Krajewski, W. F. (2000). A large-sample investigation of statistical procedures for radar based short-term quantitative precipitation forecasting. *Journal of Hydrology*, *239*(1–4), 69–84. https://doi.org/10.1016/S0022-1694(00)00360-7

Han, L., Fu, S., Zhao, L., Zheng, Y., Wang, H., & Lin, Y. (2009). 3D convective storm identification, tracking, and forecasting - An enhanced TITAN algorithm. *Journal of Atmospheric and Oceanic Technology*, *26*(4), 719-732. https://doi.org/10.1175/2008JTECHA1084.1

Hand, W. H. (1996). An object-oriented technique for nowcasting heavy showers and thunderstorms. *Meteorological Applications*, *3*, 31–41. https://doi.org/10.1002/met.5060030104

Hou, J., & Wang, P. (2017). Storm tracking via tree structure representation of radar data. *Journal of Atmospheric and Oceanic Technology*, *34*, 729–747. https://doi.org/10.1175/JTECH-D-15-0119.1

Jasper-Tönnies, A., Hellmers, S., Einfalt, T., Strehz, A., & Fröhle, P. (2018). Ensembles of radar nowcasts and COSMO-DE-EPS for urban flood management. *Water Science and Technology*, *2017*(1), 27–35. https://doi.org/10.2166/wst.2018.079

Jensen, D. G., Petersen, C., & Rasmussen, M. R. (2015). Assimilation of radar-based nowcast into a HIRLAM NWP model. *Meteorological Applications*, *22*(3), 485–494. https://doi.org/10.1002/met.1479

Johnson, J. T., Mackeen, P. L., Witt, A., Mitchell, E. D., Stumpf, G. J., Eilts, M. D., & Thomas, K. W. (1998). The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm. *Weather and Forecasting*, *13*(2), 263–276. https://doi.org/10.1175/1520-0434(1998)013<0263:TSCIAT>2.0.CO;2

Jung, S. H., & Lee, G. (2015). Radar-based cell tracking with fuzzy logic approach. *Meteorological Applications*, *22*, 716–730. https://doi.org/10.1002/met.1509

Kato, R., Shimizu, S., Shimose, K. I., Maesaka, T., Iwanami, K., & Nakagaki, H. (2017). Predictability of meso-γ-scale, localized, extreme heavy rainfall during the warm season in Japan using high-resolution precipitation nowcasts. *Quarterly Journal of the Royal Meteorological Society*, *153*(704), 1406–1420. https://doi.org/10.1002/qj.3013

Kober, K., & Tafferner, A. (2009). Tracking and nowcasting of convective cells using remote sensing data from radar and satellite. *Meteorologische Zeitschrift*, *18*(1), 75–84. https://doi.org/10.1127/0941-2948/2009/359

Krämer, S. (2008). *Quantitative Radardatenaufbereitung für die Niederschlagsvorhersage und die Siedlungsentwässerung*. Leibniz Universität Hannover. Phd thesis.

L. Foresti, & Seed, A. (2015). On the spatial distribution of rainfall nowcasting errors due to orographic forcing. *Meteorological Applications*, *22*(1), 60–74.

Lall, U., & Sharma, A. (1996). A Nearest Neighbor Bootstrap For Resampling Hydrologic Time Series. *Water Resources Research*, *32*, 679–693. https://doi.org/10.1029/95WR02966

Lang, P. (2001). Cell tracking and warning indicators derived from operational radar products. *Proceedings of the 30th International Conference on Radar Meteorology, Munich, Germany*, *i*, 245–247. http://link.springer.com/10.1007/s00376-014-0003-z

Lin, C., Vasić, S., Kilambi, A., Turner, B., & Zawadzki, I. (2005). Precipitation forecast skill of numerical weather prediction models and radar nowcasts. *Geophysical Research Letters*, *32*, L14801.

https://doi.org/10.1029/2005GL023451

Pierce, C. E., Ebert, E., Seed, A. W., Sleigh, M., Collier, C. G., Fox, N. I., Donaldson, N., Wilson, J. W., Roberts, R., & Mueller, C. K. (2004). The nowcasting of precipitation during Sydney 2000: An appraisal of the QPF algorithms. *Weather and Forecasting*, *19*(1), 7–21. https://doi.org/10.1175/1520-0434(2004)019<0007:TNOPDS>2.0.CO;2

Pierce, C., Seed, A., Ballard, S., Simonin, D., & Li, Z. (2012). *Nowcasting. Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications* (J. Bech and J. L. Chau (ed.); pp. 97–142). https://doi.org/10.5772/39054

Rossi, P. J., Chandrasekar, V., Hasu, V., & Moisseev, D. (2015). Kalman filtering-based probabilistic nowcasting of object-oriented tracked convective storms. *Journal of Atmospheric and Oceanic Technology*, *32*(3), 461–477. https://doi.org/10.1175/JTECH-D-14-00184.1

Ruzanski, E., Chandrasekar, V., & Wang, Y. (2011). The CASA nowcasting system. *Journal of Atmospheric and Oceanic Technology*, *28*(5), 640–655. https://doi.org/10.1175/2011JTECHA1496.1

Schellart, A., Liguori, S., Krämer, S., Saul, A., & Rico-Ramirez, M. A. (2014). Comparing quantitative precipitation forecast methods for prediction of sewer flows in a small urban area. *Hydrological Sciences Journal*, *59*(7), 1418–1436. https://doi.org/10.1080/02626667.2014.920505

Sharma, A., & Mehrotra, R. (2014). An information theoretic alternative to model a natural system using observational information alone. *Water Resources Research*, *50*(1), 650–660. https://doi.org/https://doi.org/10.1002/2013WR013845

Sharma, A., Mehrotra, R., Li, J., & Jha, S. (2016). A programming tool for nonparametric system prediction using Partial Informational Correlation and Partial Weights. *Environmental Modelling & Software*, *83*, 271–275. https://doi.org/https://doi.org/10.1016/j.envsoft.2016.05.021

Shehu, B. (2020). *Improving the rainfall nowcasting for fine temporal and spatial scales suitable for urban hydrology*. Leibniz Universität Hannover. PhD thesis.

Shehu, B., & Haberlandt, U. (2021). Relevance of merging radar and rainfall gauge data for rainfall nowcasting in urban hydrology. *Journal of Hydrology*, *594*, 125931. https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125931

Tonidandel, S., & LeBreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business and Psychology*, *26*(1), 1–9. https://doi.org/10.1007/s10869-010-9204-3

Tonidandel, S., LeBreton, J. M., & Johnson, J. W. (2009). Determining the statistical significance of relative weights. In *Psychological Methods* (Vol. 14, Issue 4, pp. 387–399). American Psychological Association. https://doi.org/10.1037/a0017735

Wilson, J. W., Crook, N. A., Mueller, C. K., Sun, J., & Dixon, M. (1998). Nowcasting Thunderstorms: A Status Report. *Bulletin of the American Meteorological Society*, *79*(10), 2079–2099. https://doi.org/10.1175/1520-0477(1998)079<2079:NTASR>2.0.CO;2

Wilson, J. W., Feng, Y., Chen, M., & Roberts, R. D. (2010). Nowcasting challenges during the Beijing olympics: Successes, failures, and implications for future nowcasting systems. *Weather and Forecasting*, *25*(6), 1691–1714. https://doi.org/10.1175/2010WAF2222417.1

Winterrath, T., Rosenow, W., & Weigl, E. (2012). On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management. *IAHS-AISH Publication*, *351*, 323–329.

Zahraei, A., Hsu, K. lin, Sorooshian, S., Gourley, J. J., Hong, Y., & Behrangi, A. (2013). Short-term quantitative precipitation forecasting using an object-based approach. *Journal of Hydrology*, *483*, 1–15. https://doi.org/10.1016/j.jhydrol.2012.09.052

Zahraei, A., Hsu, K. lin, Sorooshian, S., Gourley, J. J., Lakshmanan, V., Hong, Y., & Bellerby, T. (2012). Quantitative

https://doi.org/10.5194/hess-2021-248
Preprint. Discussion started: 10 May 2021
© Author(s) 2021. CC BY 4.0 License.

Precipitation Nowcasting: A Lagrangian Pixel-Based Approach. *Atmospheric Research*, *118*, 418–434. https://doi.org/10.1016/j.atmosres.2012.07.001

Zawadzki, I. I. (1973). Statistical Properties of Precipitation Patterns. *Journal of Applied Meteorology*, *12*(3), 459–472. https://doi.org/10.1175/1520-0450(1973)012<0459:spopp>2.0.co;2