



The Burstinator: A Random Forest for Downburst Detection Using Radar and Lightning Data

Florian Ackermann¹, Martin Aregger¹, Daniele Nerini², Simone Balmelli², Alessandro Hering², Josué Gehring², and Monika Feldmann^{1,3}

¹Institute of Geography and Oeschger Centre for Climate Change Research, University of Bern, Switzerland

²Federal Office of Meteorology and Climatology – MeteoSwiss, Switzerland

³Now at Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland

Correspondence: Florian Ackermann (florianvalentinackermann@gmail.com), Monika Feldmann (monika.feldmann@env.ethz.ch)

Abstract. Operational weather forecasting in Switzerland currently lacks targeted diagnostic tools for severe convective gusts, leaving their occurrence and detailed physical properties insufficiently quantified. In this study, we perform case studies of wind gust-producing convective cells to link 3D C-band polarimetric radar signatures and lightning with gust occurrence. We find that reflectivity cores, K_{DP} cores, and mid-altitude radial convergence are linked to surface wind gusts. These associations informed the development of the Burstinator, a Random Forest machine-learning model designed to distinguish between thunderstorms with and without severe gusts. The Burstinator yields a 0.66 detection probability paired with a false alarm rate of 0.08, demonstrating superior skill over the baseline WDRAFT. Predictor importance analysis indicates that radar-based features are the top predictors. Finally, we apply the Burstinator to a two-year dataset of 5-min resolution radar data to characterize the spatial distribution of the convective gust frequency. Our findings highlight how integrating machine learning with dual-polarization radar and lightning observations can advance convective wind detection and severe weather monitoring across Switzerland.

1 Introduction

Downbursts caused by severe thunderstorms can produce straight-line wind gusts on the surface, which pose risks to buildings, infrastructure, and human safety in Europe. Of all weather-related hazards, wind gusts cause the most injuries and the second most fatalities in Europe (Pilorz et al., 2023). While extensive investigation into convective wind hazards has historically centered on North America, European scientific literature on downburst dynamics remains comparatively sparse, with few operational classification tools deployed across European weather services (Pistotnik et al., 2011; MétéoSuisse, 2024; Parodi et al., 2019; Fujita, 1981; Pilorz et al., 2023).

Wet downbursts are intense downdrafts driven by strong negative buoyancy resulting from evaporative cooling and hydrometeor loading (Böhm, 1992; Markowski and Richardson, 2010). These events are subclassified into microbursts when smaller in horizontal diameter than 4 km, and macrobursts when they are larger (Fujita, 1990). Researchers in the United States have studied the wind fields and environmental characteristics linked to downbursts (Fujita, 1981; Fujita and Wakimoto, 1981;



Atkins and Wakimoto, 1991; Kuster et al., 2016; Mahale et al., 2016), and identified precursors, such as localized peaks in reflectivity, in specific differential phase shift (K_{DP}), and in mid-altitude radial convergence (MARC), along with differential reflectivity (Z_{DR}) holes, and lightning activity (Kuster et al., 2016, 2021; Mahale et al., 2016; Amiot et al., 2019; Farnell et al., 2017; Williams et al., 1989). Furthermore, increased lightning activity is linked to the occurrence of severe convective events such as downbursts (Baker et al., 1999; Williams et al., 1989; Farnell et al., 2017).

Nevertheless, only a limited number of algorithms explicitly target the detection or prediction of convective surface wind gusts. Most operational algorithms aim to capture the generalized severity of a convective system, hail, or intense precipitation events. Some efforts have been made focusing on convective surface wind gusts: Smith et al. (2004) designed a downburst prediction and detection algorithm by applying a linear discriminant analysis on reflectivity, Doppler velocity, and environmental predictors; and recently, Xiao et al. (2023) developed a deep-learning model on radar reflectivity for minute-level convective gust nowcasting in Eastern China, using interpolations of measured surface wind speeds as the ground truth.

A key reason for the limited number of such algorithms lies in observational constraints. The main factor limiting studies of downburst occurrences is the lack of data, as there is no gridded ground reference for maximum gusts available. Until recently, only measurements by stations and European Severe Weather Database (ESWD) reports could be used as reliable markers for convective wind gusts. In Switzerland, users can report severe wind among other weather events using the MeteoSwiss application since fall 2021 (MeteoSwiss, 2026). The addition of crowd-sourced reports increases the usable data considerably. We hereby make an attempt using these crowd-sourced wind data to derive additional data points for convective winds.

In addition to the scarcity of ground observations, radar-based detection is challenged by inherent measurement limitations. A common limitation of weather radars in complex terrain is beam blockage by the orography, which motivated the placement of Swiss radars at higher altitudes (Germann et al., 2022). Positioning weather radars at higher elevations minimizes beam blockage at the expense of losing near-surface scanning coverage (Germann et al., 2022). Further limitations include beam broadening and increasing elevation with distance, attenuation by heavy precipitation, and folded Doppler velocity values due to Nyquist velocity overshoot (Fabry, 2015; Feldmann et al., 2020). These factors complicate the linkage between radar-observed signatures and surface-level wind gusts.

Given these challenges, machine learning has recently attracted attention as a promising approach to optimizing conventional extrapolation methods and potentially improving warning accuracy for convection-related hazards (Prudden et al., 2020). Machine learning has been successfully applied to various severe weather problems (e.g., Hill et al. (2020); Hill and Schumacher (2021); Gensini et al. (2021); Loken et al. (2022); Gagne et al. (2017); McGovern et al. (2017)), as it can integrate heterogeneous data sources and capture non-linear relationships. For example, Leinonen et al. (2023) developed a neural network architecture for multi-hazard 2D forecasting in the Swiss radar domain, and Wolfensberger et al. (2021) applied a polarimetric-radar-based Random Forest regressor to refine precipitation estimations.

Most existing radar studies on convective wind gusts focus on flat orographic settings in the United States, leaving a research gap regarding convective gust detection across complex terrain, such as in Switzerland. Linking radar patterns to wind gusts is a relatively unexplored subject. Consequently, evaluating the extent to which observations from high-altitude radars can yield



meaningful indicators of damaging ground-level downburst outflow remains a critical research objective. The Swiss radar network, characterized by its placement in complex orography, provides a unique opportunity to address this question.

We aim to tackle the prevalent absence of a convective wind product in Switzerland by proposing a diagnostic algorithm and by pursuing the following research questions:

- A) Which radar signatures are indicative of convective surface wind gusts in Switzerland?
- B) In terms of storm properties, how are thunderstorms that we labeled 'gust-producing' in our dataset different from thunderstorms that we labeled 'non-gust-producing'?
- C) Can a Random Forest model trained on radar and lightning data reliably distinguish between thunderstorms that produce wind gusts and those that do not, and does it achieve satisfactory predictive performance?
- D) Where in the Swiss radar domain were convective gusts most active between 2024 and 2025?

2 Data

The dataset covers five consecutive warm seasons (May through October, 2019-2023) extending up to a range of 140 km across the Swiss radar domain (see Fig. 12).

2.1 Radar data

Operational weather monitoring by MeteoSwiss relies on a network of five dual-polarization C-band Doppler radars, which scan up to a 246 km range across 20 elevation angles (-0.5° to 40°) at 5-min intervals (Germann et al., 2022). Polar radar measurements are collected at a gate resolution of 500 m in range by 1° in azimuth (Germann et al., 2016). We utilize reflectivity (Z), radial velocity (V_D), and the uncorrected differential phase (Φ_{DP}). Convective storm boundaries were identified using Thunderstorms radar tracking (TRT), the MeteoSwiss product for cell tracking (Hering et al., 2004; Feldmann et al., 2023). TRT relies on dynamic reflectivity thresholds that adjust according to the evolutionary stage of the thunderstorm. We calculated K_{DP} by applying the Vulpiani method (Vulpiani et al., 2012) on the uncorrected differential phase (Φ_{DP}), and we derived radial shear (RS) using the linear least squares derivative on radial velocity (V_D) (Feldmann, 2023; Mahalik et al., 2019; Smith et al., 2021).

2.2 Surface measurements

We use the maximum 1-second gust at a temporal resolution of 10 min from official weather service stations in Switzerland and in Germany, Italy, and Austria, located within 20 km of Switzerland. To match the 5-min update cycle of the radar, we mapped each 10-min maximum station gust measurement across three consecutive radar scans. Each radar timestep is assigned the maximum gust value from the temporally closest overlapping 10-min intervals. Data processing involves rejecting all non-severe wind values ($<25 \text{ m s}^{-1}$) (Markowski and Richardson, 2010), and all non-plausible values ($>61 \text{ m s}^{-1}$). The validity range of the Swiss anemometers ends at 60 m s^{-1} . We set the upper boundary to 61 m s^{-1} , because a station measured 60.3 m s^{-1} during the devastating La Chaux-de-Fonds thunderstorm in 2023 - a value declared valid later (MétéoSuisse, 2024).



Stations >1500 m a.s.l., and all stations in the Alpine range are excluded, because they might be influenced by lower free tropospheric winds (Brönnimann, 2020).

90 2.3 Crowd-sourced data

Category	Description
Gentle breeze	Small twigs constantly moving
Strong breeze	Strong branches moving
Stormy wind	Breaking of thin branches, garden furniture blown over
Storm	Breaking of strong branches, damage to houses
Hurricane	Large trees uprooted, roofs damaged

Table 1. Crowd-sources wind report categories and descriptions.

Table 1 shows the crowd-sourced wind report categories from the MeteoSwiss application, given by a qualitative expression of the wind damage. We used the categories 'Strong breeze', 'Stormy wind', 'Storm', and 'Hurricane'. To belong to the gust-positive category, at least 'Storm' must be achieved. Because crowd-sourced wind reports launched in October 2021, we also incorporated 40 ESWD reports from 2019 to October 2021 (Dotzek et al., 2009). To align non-station observational data with the 5-min update interval of the radar network, both ESWD entries and crowd-sourced reports were mapped to the closest 5-min UTC timestep. We included crowd-sourced wind reports from the adjacent countries to Switzerland, because we excluded all stations, crowd-sourced reports, and ESWD reports from the Alpine range.

2.4 Lightning data

The lightning data stem from the company Météorage, which operates a network of lightning detection sensors detecting Cloud-to-Ground strokes and Intra-Cloud pulses, estimating their location, polarization, and current, and grouping them into flashes (Pédeboy et al., 2018; Bonnet, 2019). The localization technique combines a triangulation of the time measurements of the arriving electromagnetic waves with a measurement of the incoming wave direction (Erdmann et al., 2020). Lightning location metrics exhibit a median spatial error margin of 100–500 m and timestamping accuracy within 50 ns (Schulz et al., 2016).

105 2.5 Dataset properties

We refer to cells as the areas of thunderstorms that are outlined by the TRT algorithm, and we refer to a cell at one 5-min timestep as an event. Events can be either positive, meaning the cell caused a severe gust; negative, meaning the cell did not



cause a severe gust; or uncertain. Our approach to defining thunderstorm cells as either gust-positive or gust-negative for a given timestep is outlined in the methods section.

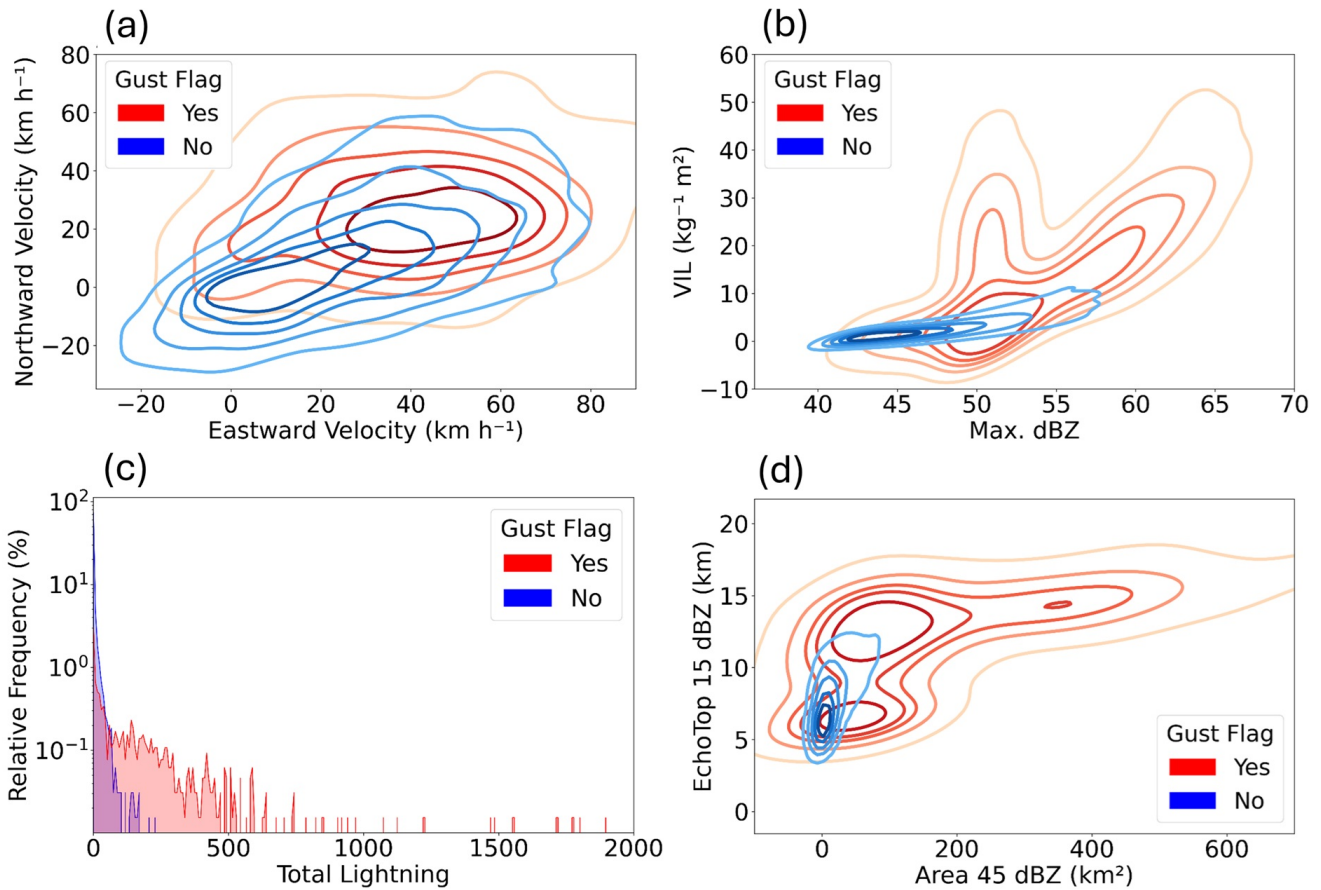


Figure 1. Characteristics of gust-positive and gust-negative thunderstorm cells. Subplots (a), (b), and (d) display Kernel Density Estimates (KDE), where contour lines delimit smoothed regions of uniform probability density. Five contour lines divide the plots into six distinct zones of ascending probability density. As an illustration, the center-most zone contains 16.66 % of the dataset, while the next surrounding zone accounts for 33.33 %. Subplot (c) is a histogram featuring a logarithmic scale on the x-axis. A small number of extreme positive events exceed 2000 total lightning strikes; nevertheless, the horizontal axis is truncated at 2000 to preserve visual legibility.

110 Figure 1 illustrates the storm characteristics associated with both gust-producing and non-gust-producing convective events within our dataset. Positive events have a stronger mean north-eastward motion than negative events (Fig. 1a). Since this is the predominant storm motion direction in Switzerland, this emphasizes faster storm motion. Fast-moving storms are more likely to produce severe surface gusts due to the downward transport of momentum. While non-gust events generally maintain VIL levels below 10 kg m⁻², gust-producing convective cells accumulate higher VIL masses alongside higher maximum reflectivity profiles. (Fig. 1b). Figure 1c shows a bulk of positive events that have 100-500 lightning strikes, while negative events have



fewer lightning strikes. In contrast to gust-negative thunderstorms that usually attain maximum 15 dBZ echo top heights of 5 to 10 km, gust-positive cells are taller and can ascend to 15 km. (Fig. 1d). The >45 dBZ reflectivity area is also larger for positive events than for negative events.

3 Methods

120 3.1 Ground truth dataset

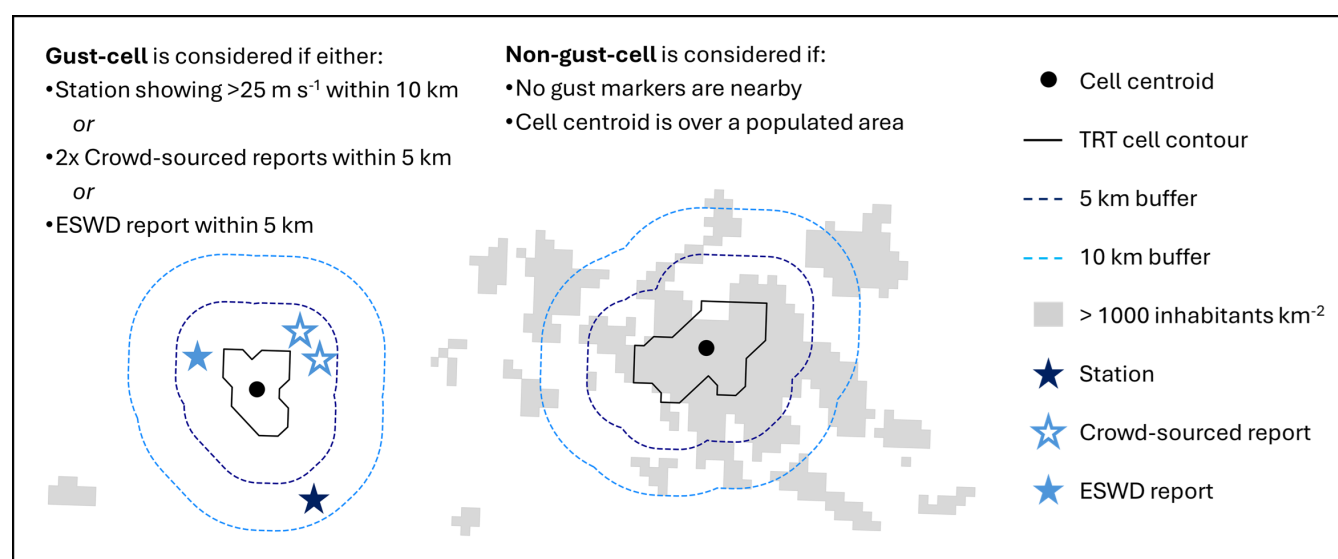


Figure 2. Ground truth labeling schematic. The non-gust-cell condition was only applied from 2022 to 2023 and between 06:00 and 22:00 local time, as crowd-sourced wind reports were launched in October 2021. If none of the shown conditions apply, the event is labeled uncertain and is excluded.

A gust marker is either a station showing $>25 \text{ m s}^{-1}$ wind, a crowd-sourced severe wind report, or an ESWD severe wind report. Whenever the required number of source-specific gust markers falls within the designated spatial buffer surrounding a TRT cell boundary, we attribute the gust to convective storm dynamics and classify the cell as gust-positive for that timestep (Fig. 2). A cell is labeled a positive event if a station showing $>25 \text{ m s}^{-1}$ is within 10 km distance, or if two crowd-sourced reports (one of which is stronger or equal to 'Storm' and the other is stronger or equal to 'Strong breeze') are within 5 km distance, or if one ESWD report is within 5 km distance. We chose an increased spatial buffer for the stations because they are more reliable than reports. The trustworthiness of the gust marker source influences the buffer distance because we want to generate a sufficiently large dataset. In cases where a gust marker lies near two or more candidate cells, attribution is prioritized to the most mature cell.

130 Figure 3 displays the counts of gust markers that were attributed to positive events for each year and month. While station data span the entire 5-year period, crowd-sourced reports started in late October 2021. The information gain due to crowd-

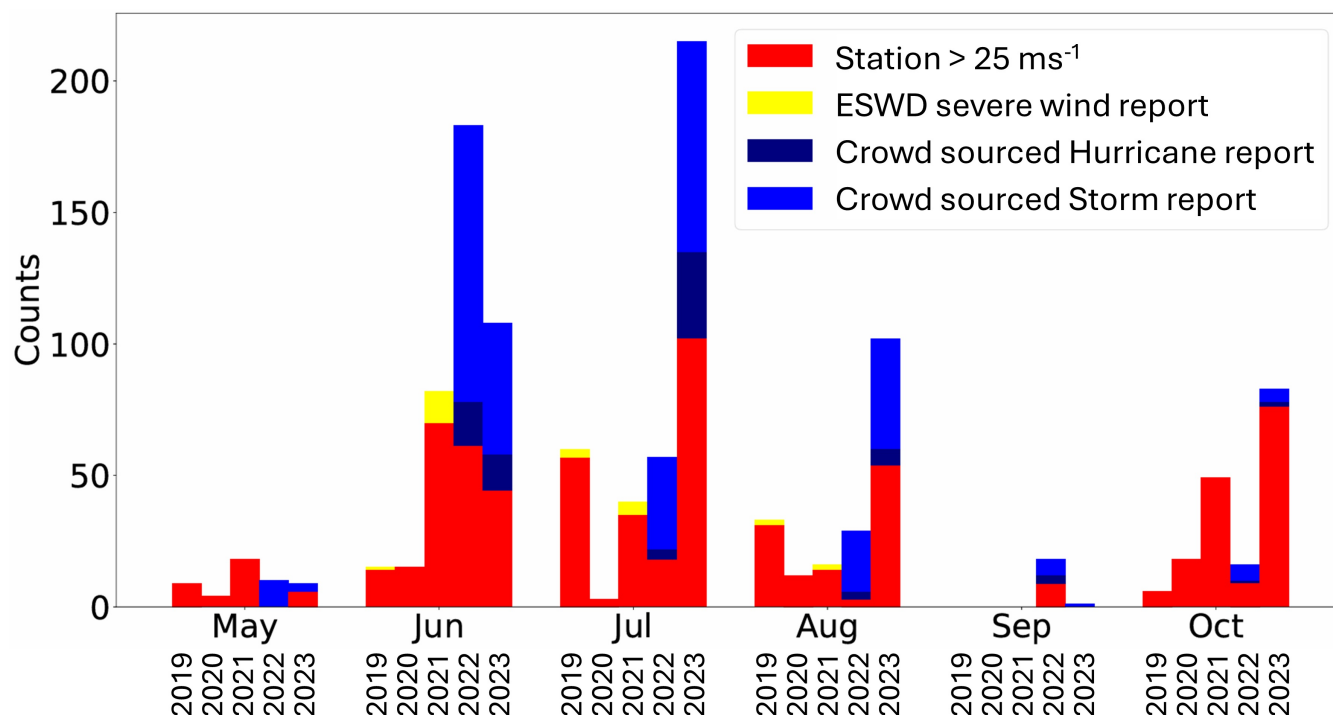


Figure 3. Cumulative bulk of stations, ESWD, and crowd-sourced reports attributed to cells, which we defined as gust-positive, per month and year. A cell can bind multiple stations, ESWD, and crowd-sourced reports at the same time. All cell-associated gust markers are stacked in this plot. Crowd-sourced reports of 'strong breeze' and 'stormy wind' are not shown.

sourcing is especially pronounced in June, July, and August of 2022 and 2023. After crowd-sourcing was introduced, the number of gust markers roughly doubled.

For a negative label, the cell centroid must overlay a region of >1000 inhabitants per km^2 while maintaining an absence of gust markers in the designated buffer zones (Fig. 2). A population density threshold is chosen to increase certainty that the thunderstorm did not cause convective gusts, rather than the gust remaining unnoticed by the population. Negative labeling was applied exclusively during daytime hours (06:00 to 22:00 UTC), when an active population is expected to submit reports during severe wind events.

Figure 4a depicts the gust-positive events, which are densely spread in the Ticino region south of the Alps. The gust-negative events are shown in Figure 4b, and highlight the most populated areas in Switzerland, since they are defined only for these areas (see Fig. 2). Most negative events are situated on the Swiss plateau to the North of the Alps, whereas the North-South distribution is more balanced in the positives.

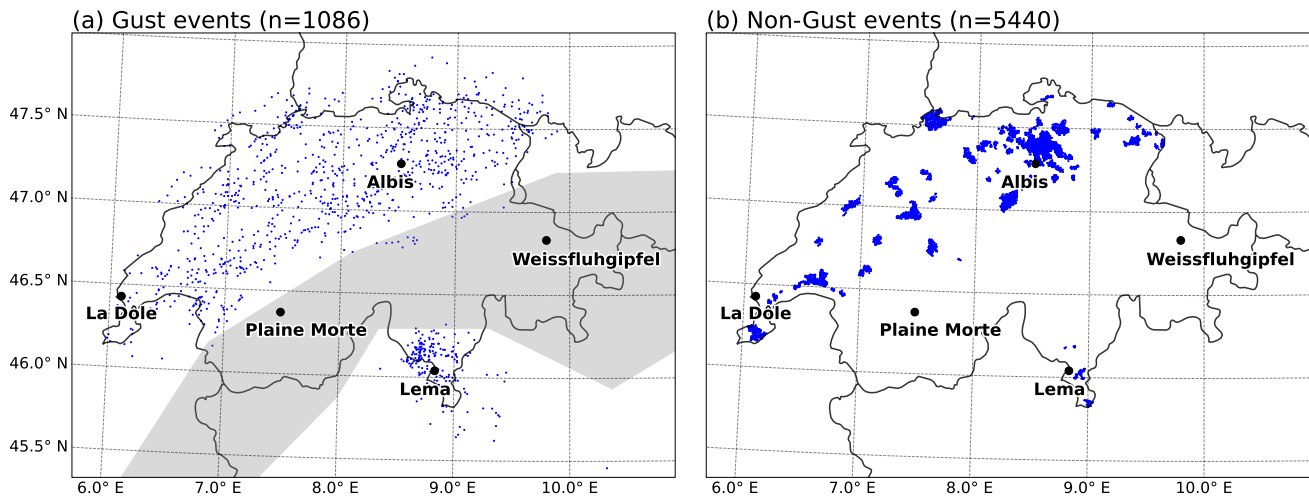


Figure 4. The black dots represent the Swiss radars. Subplot (a) shows the centroids of the TRT cells labeled as gust-positive for the period 2019 to 2023. Subplot (b) shows the centroids of the TRT cells labeled as gust-negative for the same period.

3.2 Random forest classifier

Random Forests combine multiple decision trees constructed from random subsamples of the dataset, utilizing a randomized subset of predictors to evaluate every split (Breiman, 2001). It has been shown that Random Forests are especially successful with large numbers of variables and are broad enough to adapt to various classification problems, since the element of randomness makes Random Forests especially resistant to overfitting (Biau and Scornet, 2016). We aimed to optimize the Random Forest on the Critical Success Index (CSI) (see Appendix B). A high CSI indicates good detection probability and few false alarms. Predictor selection was informed by the case study analyses demonstrating that enhanced reflectivity, K_{DP} , and mid-altitude radial shear act as radar indicators of downbursts. We remapped polarimetric reflectivity, K_{DP} , and radial shear data from the original single-radar polar space onto a gridded Cartesian 3D field for the extraction of radar predictors. The 3D grid has a horizontal resolution of 1000 m onto 1000 m, and a vertical resolution of 200 m. A detailed description of the radar predictor extraction process is contained in Aregger et al. (2026). Each gate is transferred to its respective location in the gridded 3D field. In instances where multiple radar gates coincide within the same Cartesian grid box, the maximum observed value is retained. The final volume-integrated predictors were then extracted for each cell and timestep. All predictors are described in Appendix (Table B1). In addition, we added lagged predictors for Intra-cloud lightning and total lightning. Lags are values from previous timesteps, so, for example, lag 1 is taken from the state of the cell 5 min before the valid time.

Figure 5 depicts the workflow. We iteratively replaced missing predictor values using a Bayesian ridge. To maintain dataset integrity, we excluded cells exhibiting opposing gust labels within a 15-min window, as well as cells younger than 15 min, which lacked defined values for the 15-min lagged predictors. Because events in the training set should ideally be uncorrelated

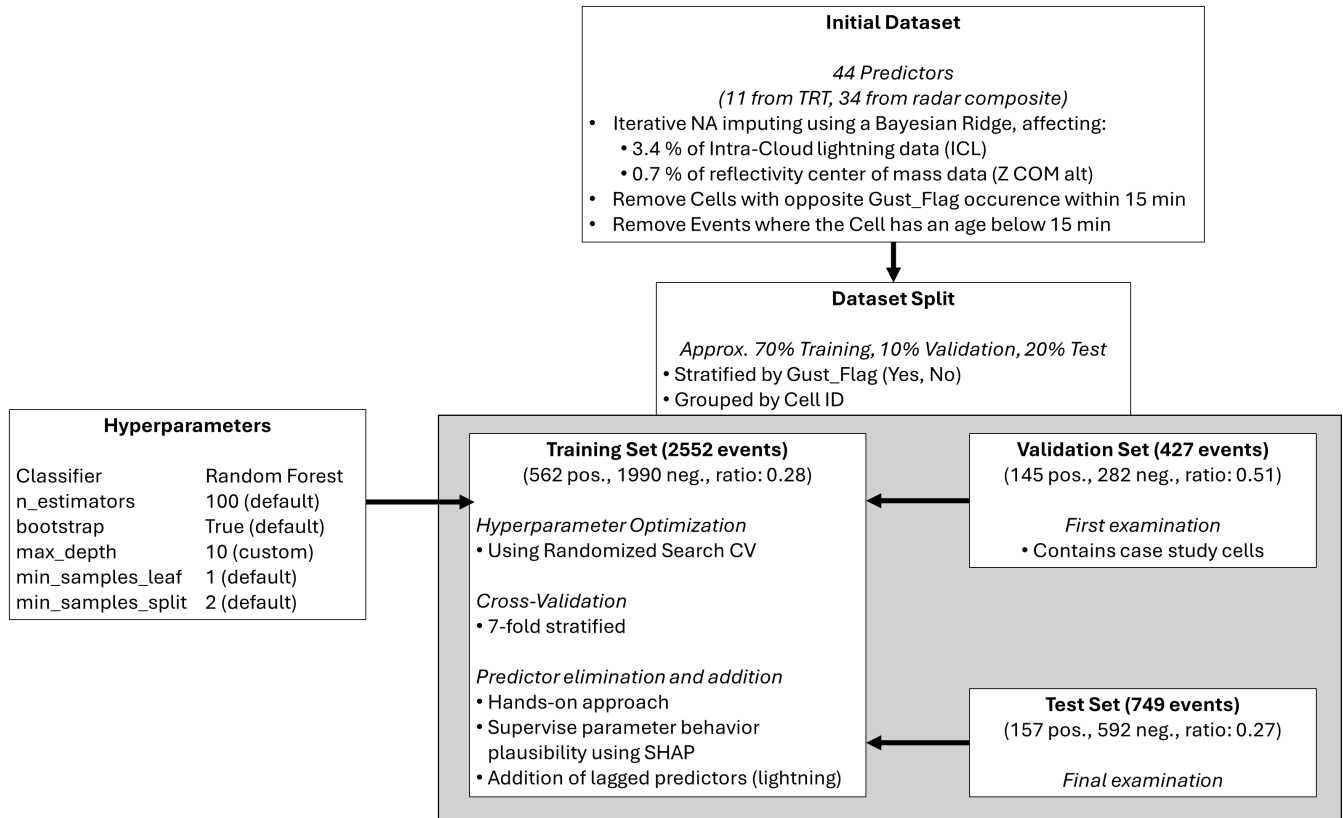


Figure 5. Workflow architecture

to events in the test set, we assigned cells for their entire lifespan to a single dataset split (Chase et al., 2022). All case study cells were assigned to the validation set. The dataset was partitioned into approximately 70 % training, 10 % validation, and 20 % testing subsets through stratification by gust label to preserve proportional positive-to-negative event ratios. Then, all case study cells were allocated to the validation set, thereby increasing the positive-to-negative ratio to 0.51. Cross-validation is a data resampling method and can be used to tune hyperparameters and to reduce overfitting (Berrar, 2019). We performed a stratified 7-fold random subsampling cross-validation on the training set to assess the generalization ability of our model (Fig. 5).

3.3 Predictor importance

We assessed predictor importance using permutation-based and SHAP-based ranking methods. Permutation importance quantifies the drop in model performance after randomly permuting a predictor, thereby breaking its association with the response (Gregorutti et al., 2017). We use the single-pass approach, which permutes each predictor independently, and the multi-pass approach, which successively retains already identified important predictors as they are permuted, to better account for correlated



features. In addition, we computed SHAP importance, which is based on Shapley-value theory and summarizes each predictor's average absolute contribution to the model predictions, providing a complementary attribution-based ranking (Lundberg and Lee, 2017).

3.4 Baseline WDRAFT

To provide a quantitative reference point for evaluating our model, we required a baseline. Formulated by Steward (1991), the wind gust estimator w (ms^{-1}) functions as a gridded, column-integrated product intended to quantify downdraft velocity inside a convective precipitation shaft. It is based on the "Cloud Top Penetrative Downdraft Mechanism" and indicates a strong downward velocity when VIL (Vertically Integrated Liquid) is high and EchoTop is low. The equation for w is as follows:

$$w = ((20.63 \text{ms}^{-2})(VIL) - (3.125 \cdot 10^{-6} \text{s}^{-2})(EchoTop)^2)^{0.5}$$

MeteoSwiss has been using WDRAFT (ms^{-1}), which is an adaptation of w using EchoTop20 and an additional cell displacement velocity (in ms^{-1}) term since 2012 (Trefalt, 2017) (Appendix A). Although w as an approximation of the "Cloud Top Penetrative Downdraft Mechanism" was successfully implemented in the United States, WDRAFT did not perform satisfactorily in Switzerland. Due to EchoTop frequently overshooting the velocity term, the following extraction of the square root of the resulting negative number generates an error. Thus, we removed the squaring of EchoTop, which improved WDRAFT's performance in Switzerland and provided a sufficient baseline for our studies:

$$WDRAFT = (20.63 \text{ms}^{-2} \cdot VIL - 3.125 \times 10^{-6} \text{s}^{-2} \cdot EchoTop20)^{0.5} + \frac{velocity}{10.8}$$

We decided to use WDRAFT to investigate whether a machine learning approach combined with a variety of predictors performs better than a more "simple" approach. We define WDRAFT as gust-indicating when it exceeds 25ms^{-1} within a TRT cell, constituting severe wind per definition (Markowski and Richardson, 2010).

4 Results

4.1 Case study

The following case study shows the predictive potential of the selected variables. Figure 6 shows a cell over East Switzerland on 2023-08-24 at 18:55 UTC. Stations measured up to 31.6ms^{-1} winds, and crowd-sourced reports were issued around the cell. The 1° Weissfluhgipfel scan shows strong gradients of opposing radar-relative velocities (Fig. 6a). Such areas are highlighted by positive radial shear, indicating strong convergence within the thunderstorm (Fig. 6b). These convergent shear zones are associated with the pressure deficit induced by negatively buoyant air masses, which drag in surrounding air from the upper levels. Furthermore, the Albis radar 1° scan captured $>55 \text{dBZ}$ reflectivity, and K_{DP} up to 5°km^{-1} , indicating heavy precipitation loading (Figs. 6c and 6d). The other case study cells showed similar radar properties concerning reflectivity, K_{DP} , and radial shear, as outlined in Ackermann (2025). These findings align with previous radar-based downburst observations

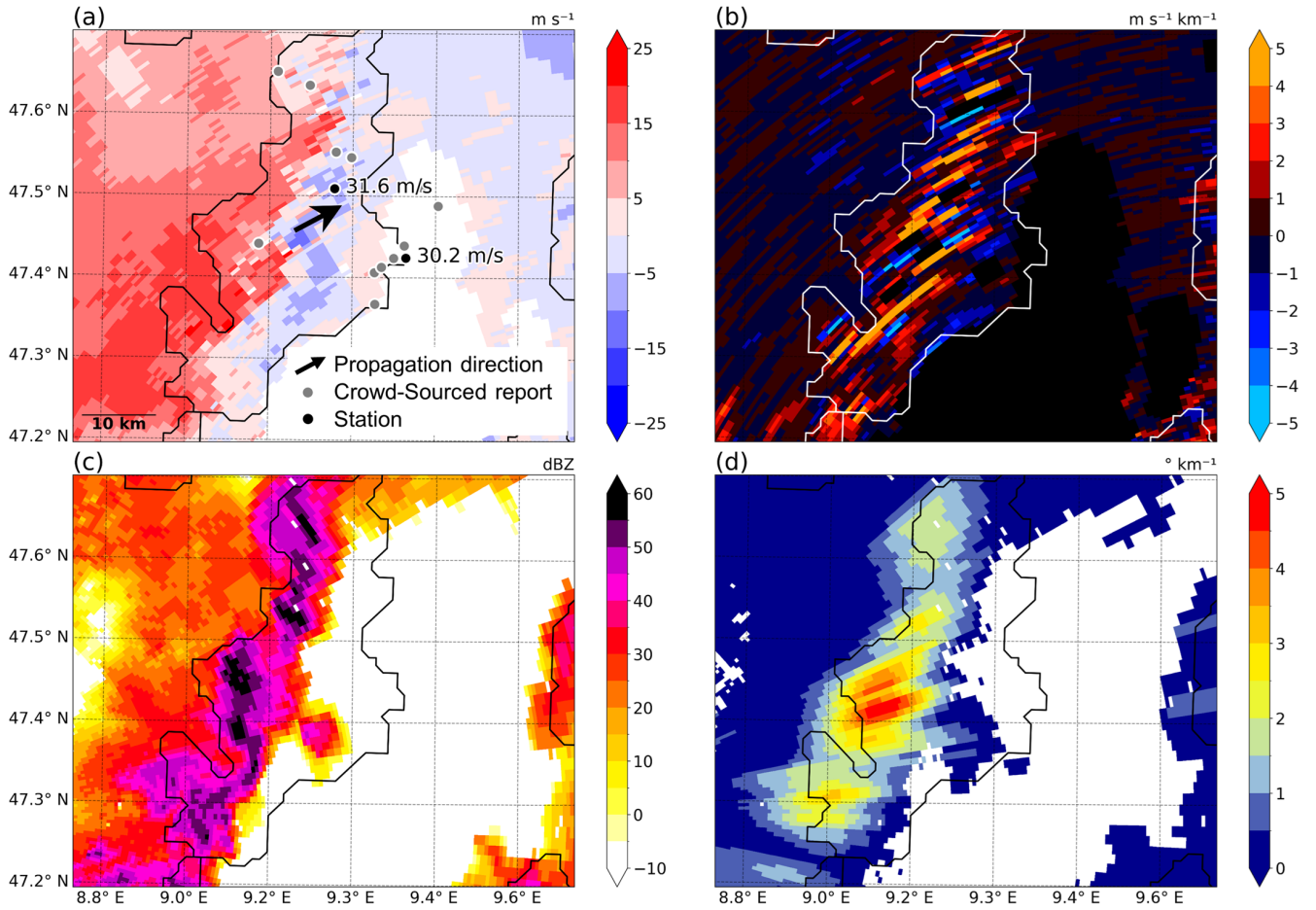


Figure 6. Situation on 2023-08-24 at 18:55. (a) is the radial velocity measured by the Weissfluhgipfel 1° elevation scan. Red depicts movement toward the radar, and blue movement away from the radar. The arrow indicates the cell displacement direction. The grey points are crowd-sourced wind reports, stronger or equal to 'strong breeze'. The black dots are weather stations, and the values are the maximum per-second wind gust speeds between 18:50 UTC and 19:00 UTC. Subplot (b) displays radial shear derived via the linear least squares derivative operator from subplot (a), showing convergent radial shear in red and divergent radial shear in blue. (c) is reflectivity measured by the Albis 1° elevation scan, and (d) is K_{DP} measured by the same scan.

(Kuster et al., 2021; Mahale et al., 2016; Amiot et al., 2019). We also observed Z_{DR} cores, indicating strong updrafts. However, we rarely observed a clear descending motion of reflectivity cores, K_{DP} cores, or Z_{DR} holes, as observed previously before strong surface wind gusts (Isaminger, 1988; Kuster et al., 2021; Mahale et al., 2016).

Figure 7 shows how the Burstinator performs on the case study cell from Fig. 6. The severity of the cell intensified between 17:45 and 17:50 UTC. The Burstinator produced the first positive at 17:50 UTC, matching the true Gust_Flag. The SHAP waterfall plot at 17:50 UTC (Fig. 7a) indicates that $Z > 40$ had the strongest impact on the positive prediction. At 18:55 UTC,

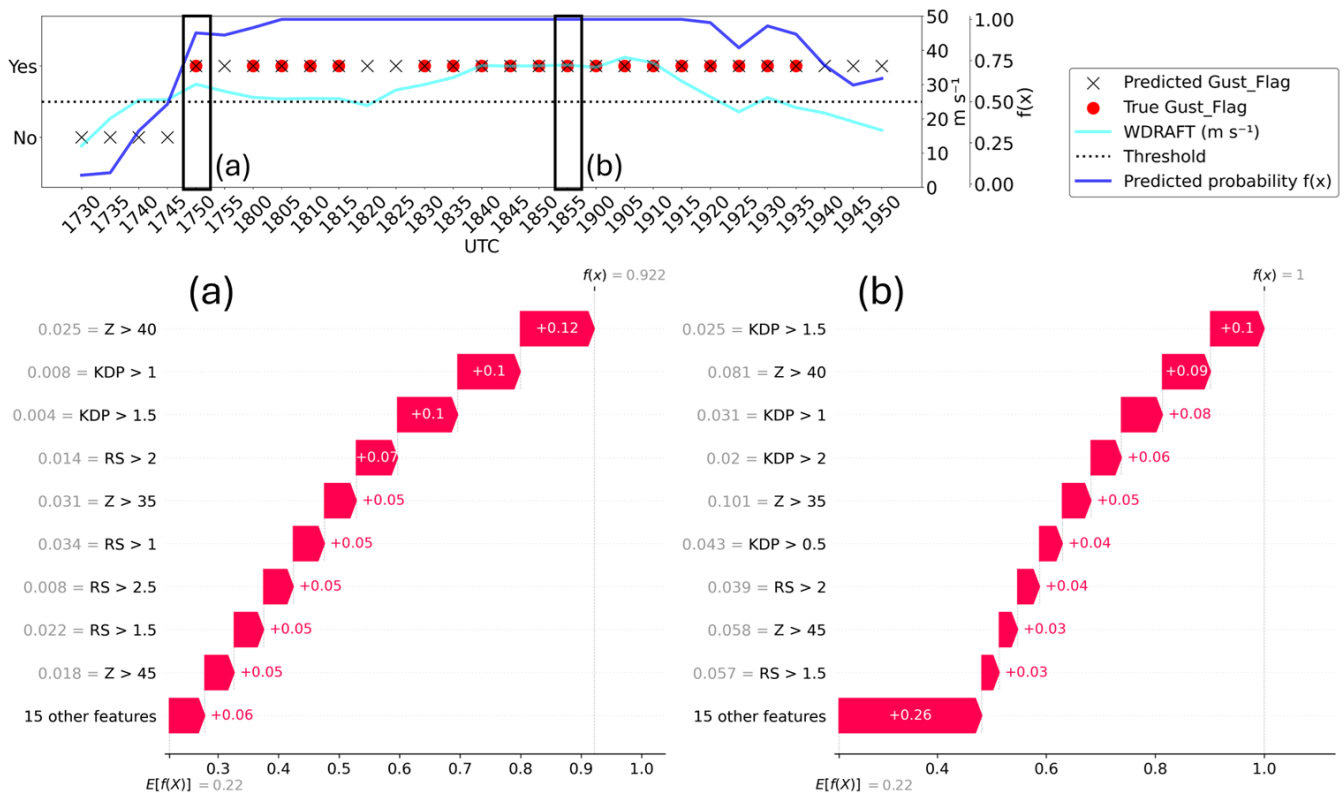


Figure 7. Temporal continuation of the case study cell from Fig. 6, assigned to the validation set. The plot shows the true Gust_Flag, the Burstinator’s prediction, predicted probability, and the WDRAFT values. The waterfall plots are valid for 17:45 UTC (a) and for 18:55 UTC (b). Date: 2023-08-24.

$KDP > 1.5$ had the strongest impact, and the Burstinator produced a true positive with a predicted probability of 1. The high predicted probability also matched our understanding of the case, as the cell clearly induced damaging gusts at that time, indicated by several crowd-sourced reports, and wind speeds of up to 31.6 m s^{-1} (Fig. 6a).



4.2 Model performance and predictor importance

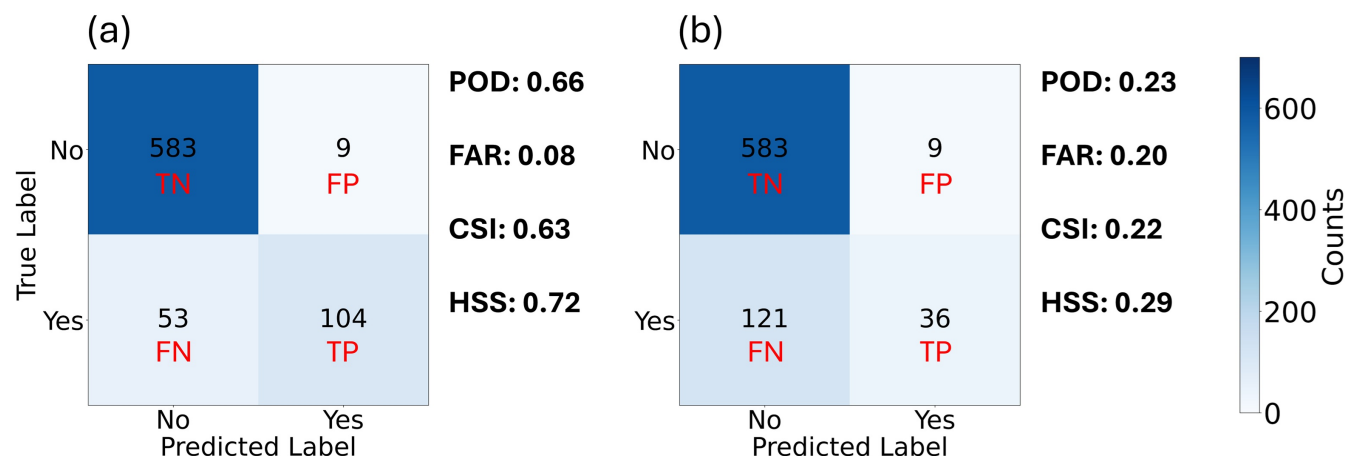


Figure 8. Confusion matrices and scores for the Burstinator (a) and WDRAFT (b) performing on the test set. The scores are defined in Appendix B.

	Precision	POD/Recall	F1-Score	Support
Negative	0.92	0.98	0.95	592
Positive	0.92	0.66	0.77	157
Accuracy			0.92	749
Macro Average	0.92	0.82	0.86	749
Weighted Average	0.92	0.92	0.91	749

Table 2. Classification report for the Burstinator performing on the test set.

210 In figure 8, the confusion matrices for test set predictions by the Burstinator (Fig. 8a) and WDRAFT (Fig. 8b) reveal that while both models correctly identified 583 negative cases and yielded 9 false positives, the Burstinator demonstrated superior capacity in identifying positive events, with 104 true positives compared to 36 for WDRAFT. WDRAFT produced considerably more false negatives (121) than true positives, whereas the Burstinator reduced the number of false negatives to 53. Compared to the baseline WDRAFT (POD: 0.23, FAR: 0.20), the scores indicate that the Burstinator yields a greater POD of 0.66 along with a lower FAR of 0.08. Consequently, the Burstinator outscores WDRAFT in Critical Success Index with a CSI of 0.63 versus 0.22. Moreover, a higher HSS of 0.72 for the Burstinator compared to 0.29 for WDRAFT demonstrates the machine-learning framework's refined skill in separating gust-positive from gust-negative events.

215

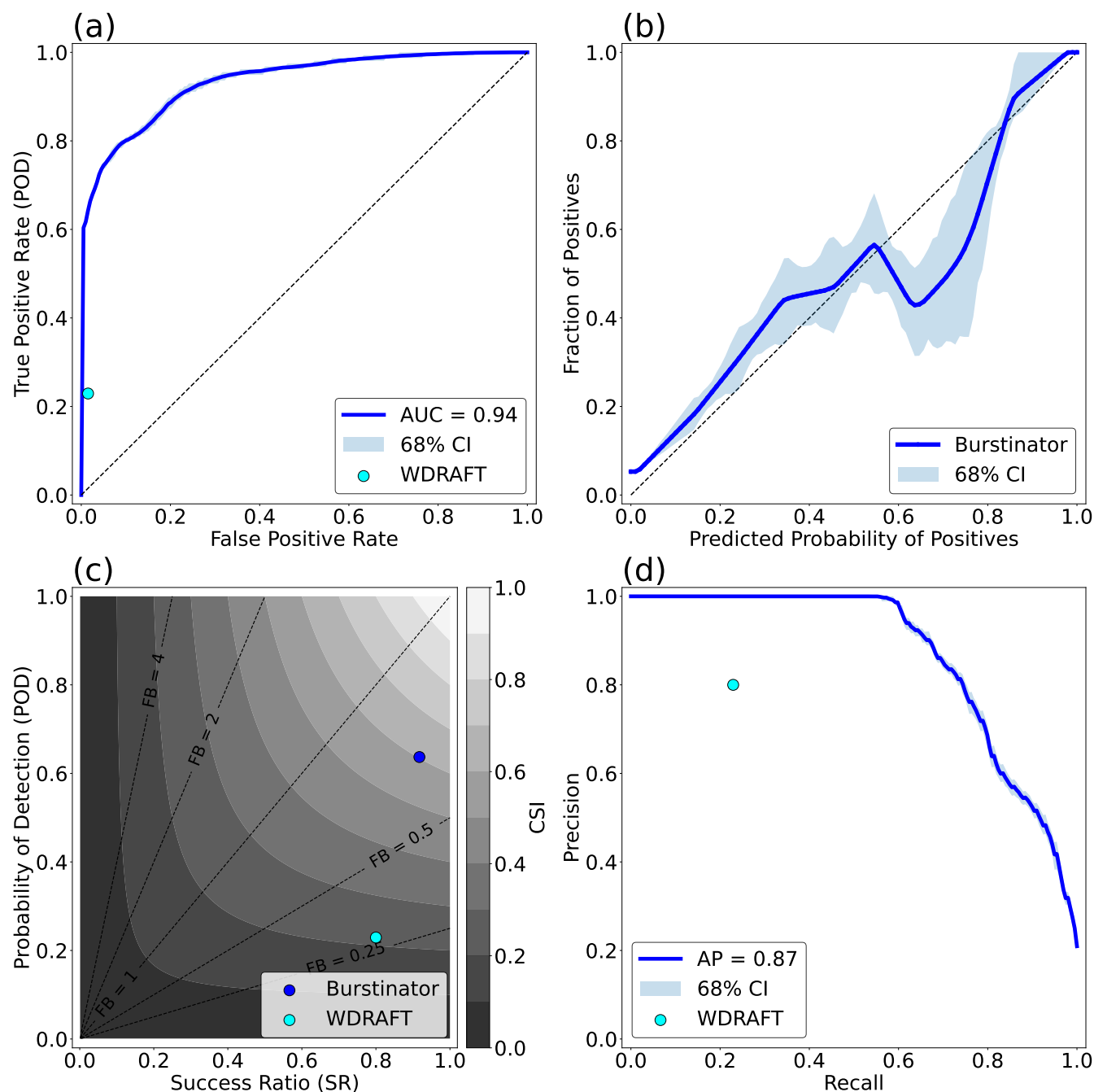


Figure 9. Performance diagrams from the evaluation of the Burstinator against the test set, averaged over 100 runs. Subplot (a) illustrates the Receiver Operating Characteristic (ROC) curve alongside a 68 % confidence interval, while subplot (b) presents the 10-bin calibration curve, also showing the 68 % confidence intervals. (c) are the performance diagrams with corresponding CSI and frequency biases (FB) from the Burstinator and WDRAFT. (d) is the precision-recall curve with the 68 % confidence intervals.



Figure 9a highlights the discriminatory performance through the ROC curve, where a sharp upper-left corner and an AUC of 0.94 confirm a strong true positive rate relative to false positives across all decision thresholds. Figure 9b illustrates model calibration, indicating that while the Burstinator underestimates event likelihood in the 30–40 % predicted probability bin, it tends toward overconfidence when assigning probabilities of 60–80 %. Figure 9c is the performance diagram displaying the Burstinator, having a frequency bias of around 0.75. WDRAFT has an even lower frequency bias of 0.25. Figure 9c indicates that both models tend to underpredict positive events; however, the Burstinator is substantially closer to unbiased performance than WDRAFT, which shows a much stronger underprediction. Figure 9d depicts the trade-off between precision and recall across different thresholds, where the Burstinator has an average precision (AP) of 0.87. The Burstinator achieves a strong precision-recall balance overall, with an average precision of 0.87, whereas WDRAFT operates at high precision but very low recall, indicating that it captures only a small fraction of the observed events.

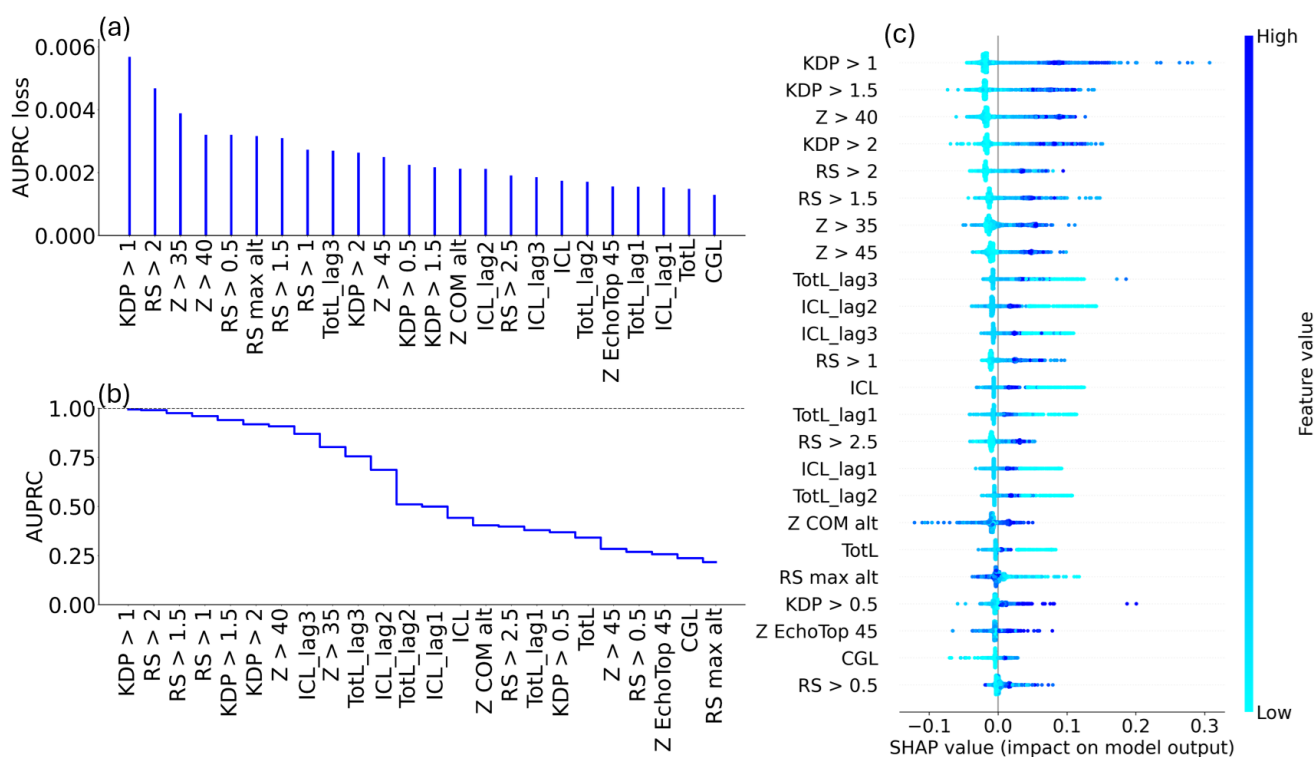


Figure 10. Predictor importance charts. (a) and (b) display skexplain Permutation Importances computed from a subsample of 50 % of the training data averaged over 50 permutation iterations. The evaluation metric is the area under the precision-recall curve (AUPRC). (a) is the backward single-pass permutation importance, and (b) is the backward multipass permutation importance. (c) is the SHAP importance. Each point is a single observation. The position on the x-axis denotes the SHAP value. Positive SHAP values correspond to a stronger push towards a positive prediction. The colors denote the predictor values. The predictors are ranked by their mean absolute SHAP value.



Figure 10a shows the backward single-pass permutation importance for the Burstinator, displaying the performance loss when a single predictor is shuffled randomly. The most important predictors in Figure 10a are K_{DP} ($KDP > 1$), radial shear (RS > 2), and reflectivity ($Z > 35$). Figure 10b is the backward multi-pass permutation importance, where the predictors remain shuffled, starting with the most important single-pass permuted predictor. Among the most important predictors in Figure 10b are K_{DP} and radial shear predictors. Figure 10c is the SHAP predictor importance ranking, and shows K_{DP} ($KDP > 1$ and $KDP > 1.5$), and reflectivity ($Z > 40$) among the most important predictors. We deem the model physically coherent, as higher reflectivity-, K_{DP} - and radial shear values generally correspond to stronger positive prediction output (Fig. 10c). Noticeably, all charts in Figure 10 display a mixture of K_{DP} , reflectivity, and radial shear predictors among the most important features. Furthermore, the greater lags of lightning (e.g., TotL_lag3, or ICL_lag3) generally rank higher than unlagged lightning (TotL, ICL, CGL), which is indicative of lightning jumps preceding storm hazards.

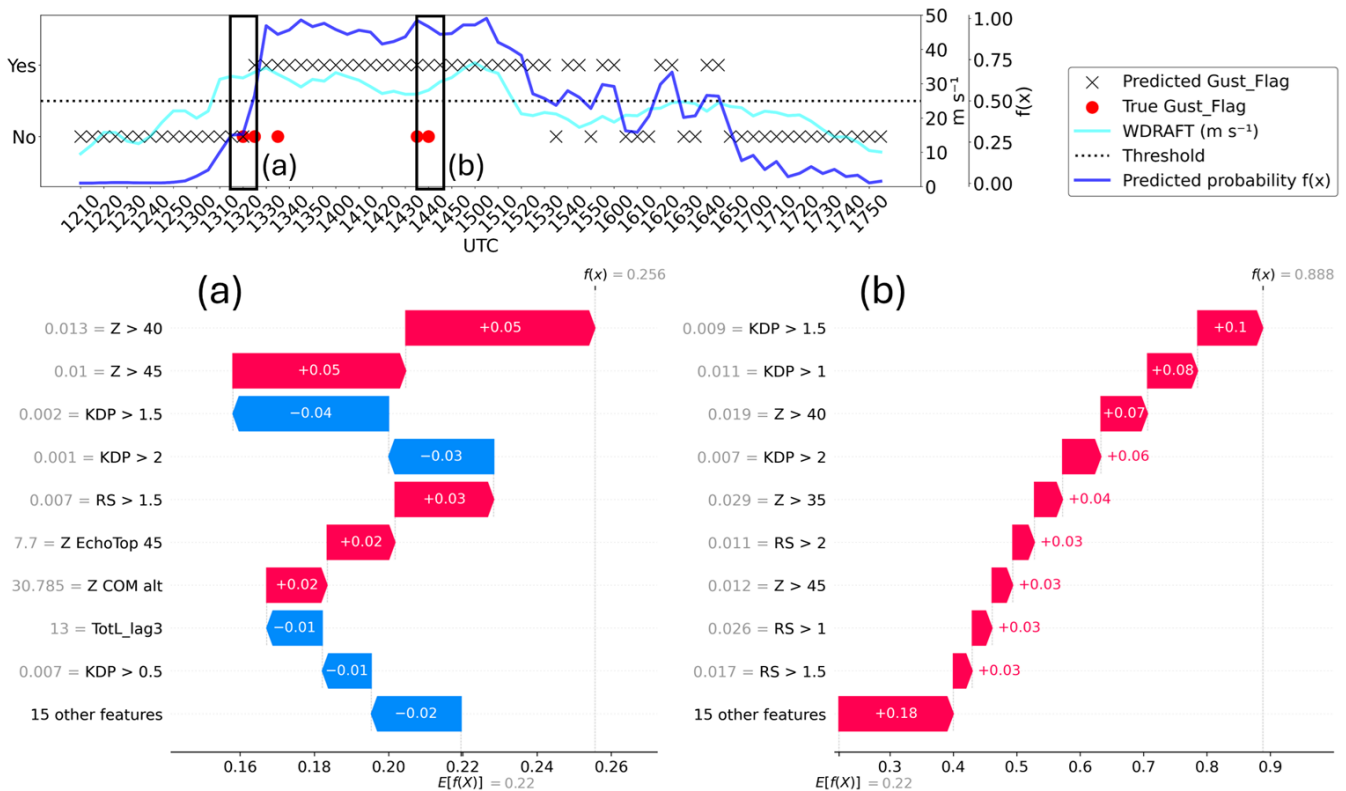


Figure 11. Temporal continuation of a TRT cell on 2023-07-12 in the validation set showing the true Gust_Flag, the Burstinator’s prediction, predicted probability, and the WDRAFT values. The Burstinator produced a true negative at 13:20 (a) and a false positive at 14:40 (b). (a) and (b) are SHAP waterfall plots showing the predictor values on the y-axis, and their impact on the model prediction.

Figure 11 shows examples of a true negative (Fig. 11a) and of a false positive (Fig. 11b) during the lifecycle of a thunderstorm, which was in the validation set. Some reflectivity predictors had a positive impact on the Burstinator’s prediction



240 at 13:20 UTC, and WDRAFT exhibited $>30 \text{ ms}^{-1}$ (Fig. 11a). The event at 13:20 UTC had large precipitation content, indicated by high WDRAFT and high reflectivity in the Burstinator. At 14:40 UTC, the Burstinator predicted a positive by 89 % probability, with all predictors indicating a gust-positive event by their SHAP values (Fig. 11b). In addition, WDRAFT also exhibited $>25 \text{ ms}^{-1}$. Yet, this is labeled a negative event in the ground truth dataset because the cell had its center point in a populated area, and no reports were issued. Thus, the Burstinator and WDRAFT hit a false alarm. However, regarding the
245 predicted probability of the Burstinator and the WDRAFT value, we assume a wrong label in the ground truth here.



4.3 Prediction of unseen data

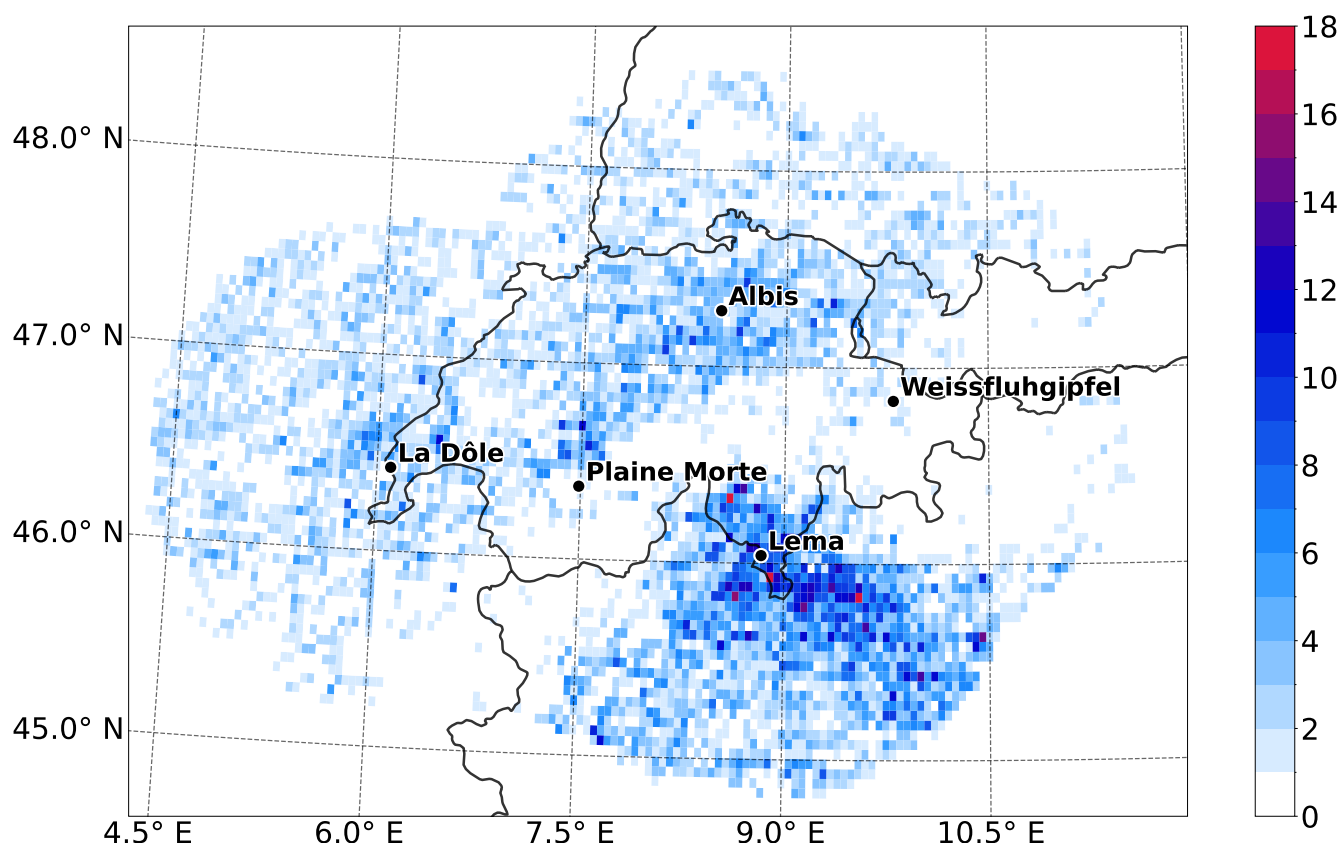


Figure 12. Two-year convective downburst occurrence map. The colors denote the cumulative positions of the cell centroid during positive predictions made by the Burstinator for 5-min TRT cells in the Swiss radar domain for 2024 and 2025, between May and October each year.

Figure 12 is the convective gust frequency map from 2024 to 2025, which we derived by predicting the binary gust class for every 5-min TRT cell using the Burstinator. The Ticino and Northern Italy are the most active convective gust regions in the Swiss radar domain. This is unsurprising, given that these regions are thunderstorm hotspots (Feldmann et al., 2023). North of the Alps, local downburst frequency peaks are present close to the radars. Increased frequencies of downbursts around La Dôle and Albis match the spatial distribution of ordinary thunderstorms (Feldmann et al. (2023) fig. 3). In contrast, the frequencies in the Northern prealps between Plaine Morte and Albis also match the spatial distribution of stronger thunderstorm types, such as hail- or rainstorms. The Burstinator also predicts gusts in the Alpine range, although we removed all Alpine events from the ground truth. Despite the Alpine range being less prone to severe thunderstorms, we nonetheless expect some strong gust events.



	Total	Positive	Negative	Positive %
Events	655'237	10'080	645'157	1.5
Cell IDs	155'474	1'588	153'886	1.0

Table 3. Total positive and negative classifications generated by the Burstinator for all 5-min events and individual cell IDs evaluated on unseen data spanning May through October across 2024 and 2025.

Table 3 shows the numbers of positive and negative predictions between May and October 2024 and 2025. The Burstinator classifies 1.5 % of events as positive. Within this period, there were a total of 155'474 TRT cells in the Swiss radar domain. Only a small fraction of 1 % of these TRT cells were classified gust-positive by the Burstinator during their lifespan. This is an indication of not "over-warning".

260 **5 Discussion**

We have outlined through previous case studies that reflectivity cores, K_{DP} cores, and convergence are co-occurring with strong surface wind gusts in Switzerland (Ackermann, 2025). While strong K_{DP} signatures indicate high concentrations of oblate raindrops (Frugis, 2020; Kuster et al., 2021) and high reflectivity denotes hydrometeor mass loading comprising hail or graupel (Donaldson, 1961; Fabry, 2015), convergence marked by positive radial shear reflects air advection triggered by dry-air entrainment or the low-pressure anomaly left by a descending cold pool (Funk et al., 1998; Hong et al., 2021; Lemon and Parker, 1996; Eilts et al., 1996).

Whereas reflectivity and K_{DP} cores indicate heavy precipitation loading, large radial shear hints at the negative buoyancy. These findings align with preexisting observations of wet downbursts (Kuster et al., 2016, 2021; Mahale et al., 2016; Markowski and Richardson, 2010). We extracted reflectivity- K_{DP} -, and radial shear predictors for model training because the case studies showed that these radar parameters are indicative of downbursts in Switzerland.

Training a model on downburst detection requires a reliable and sufficiently large ground truth of convective wind gusts. Switzerland - in comparison to almost any other country - has an exceptionally dense measurement network. Nonetheless, constructing a ground-truth wind gust dataset is inherently challenging due to the absence of continuous gridded surface wind data in Switzerland; consequently, positive labeling criteria represent a balance between minimizing false positive attributions and preserving an adequately sized dataset for machine learning. Aside from stations and ESWD reports, the model is trained on extensive crowd-sourced data. Furthermore, without crowd-sourcing, negative labelling would have been impossible. Crowd-sourced reports offer broad coverage and great spatiotemporal detail, but they are inherently uncertain and require quality control (Aregger et al., 2026; Kopp et al., 2024). Models trained on such data should therefore be validated on independent observations first. Crowd-sourced observations introduce potential uncertainties through subjective intensity overestimation, a lack of direct mapping between qualitative damage classes and the 25 ms^{-1} severe wind threshold, and reporting time lags.



Temporal precision is further constrained by station measurement intervals (10-min maximum 1-second gusts), which complicate precursor-to-downburst lag estimation. Moreover, spatial buffers for positive events and unverified reporting assumptions for negative events in populated zones contribute to labeling noise, necessitating future dataset revisions — such as raising population density cutoffs and tightening spatial margins — as longer observational archives become available. Figure 11 shows the difficulty of negative labelling: Despite the Burstinator and WDRAFT both strongly suggesting a positive event at, for example, 14:40 UTC (Fig. 11b), the event is labelled negative in the ground truth due to no nearby crowd-sourced wind reports at that time.

Utilizing the ground-truth dataset, we trained a Random Forest model to identify gust-positive events; test set evaluations demonstrate that the Burstinator achieves a higher POD and CSI alongside a lower FAR than the baseline WDRAFT — a promising outcome given Switzerland’s lack of existing operational convective wind detectors. Although confusion matrices confirm that false negatives remain the primary limitation of the Burstinator, its frequency bias is closer to unity than that of WDRAFT, which underpredicts positive cases far more frequently. The calibration curve indicates that the Burstinator is overconfident in events with a 70-80 % probability of being positive and underconfident in events with a 30-40 % predicted probability. An AUC of 0.94 demonstrates the Burstinator’s strong class discrimination, reflecting a 94 % probability that a randomly chosen gust-positive event will receive a higher predicted probability score than a randomly chosen negative event. On one hand, we argue that the Burstinator performs well because we use a variety of sources to extract predictors, namely reflectivity, radial velocity, K_{DP} , and lightning. The feature importances reveal that K_{DP} , radial shear, and reflectivity features are the most important feature classes, whereas lightning is less important. Both permutation and SHAP ranking agree that K_{DP} ($KDP > 1$) is the most important feature. This result is expected, as K_{DP} is generally regarded as a more robust radar parameter for detecting heavy precipitation than reflectivity. At C-band wavelengths, heavy precipitation leads to significant radar attenuation, whereas K_{DP} is largely unaffected by this attenuation.

The downburst frequency map aligns with the literature, revealing that the Southern Prealps and, to a lesser extent, the Northern Prealps, show frequent downburst activity (Feldmann et al., 2023). Furthermore, our work complements the mesocyclone detection algorithm of Feldmann (2023) by extending detection capabilities from mid-level rotational signatures to surface-impacting convective wind gusts. In summary, the machine learning framework demonstrates superior capability for radar- and lightning-based severe weather detection compared to traditional, single-formula estimators like WDRAFT.

From a climate-change perspective, convective wind gusts are thermodynamically expected to intensify in a warmer atmosphere (Prein, 2023). However, projections for Europe remain uncertain due to intermodel and regional variability (Púčik et al., 2017; Kotarba, 2026). This uncertainty underscores the importance of robust warning systems, both for future research and for effective adaptation and preparedness. In addition, severe convective storms are often multi-hazard, as wind, precipitation, hail, and lightning commonly co-occur, and this is likely to remain relevant for future severe-weather risk.



6 Conclusions

This study trained and validated the Burstinator Random Forest model to detect TRT cells causing wind gusts. We trained the Burstinator on a five-year ground truth, which we built using crowd-sourced wind reports, station data, and ESWD reports. We
 315 aimed to train a skillful model that is physically coherent and feasible to operate in real-time.

First, we demonstrated that radar signatures linked to heavy precipitation loading and mid-altitude convergence are indicative of downbursts in Switzerland. We then extracted these signature metrics for each gust-positive and gust-negative cell timestep, utilizing them as predictors to train the Random Forest machine learning model. The radar products reflectivity and K_{DP} capture the heavy precipitation loading, and radial shear captures the descent of the cold pool. The combination of these radar predictors
 320 with lightning predictors yielded satisfying results. The Burstinator shows superior skill in distinguishing gust-positive from gust-negative thunderstorms compared to the WDRAFT baseline (CSI of 0.63 vs 0.22). The impact of predictor values on the prediction is physically coherent, and K_{DP} ($KDP > 1$) emerged as the top predictor across both permutation and SHAP importance.

To sum up, our work answers the research questions as follows:

325 A) Enhanced reflectivity, K_{DP} , and convergent radial shear are indicative of wet downbursts or strong downdrafts in Switzerland.

B) Gust-producing thunderstorms have higher precipitation loading, stronger north-eastward propagation, and show increased lightning activity compared to gust-negative thunderstorms.

C) The random forest provides an effective underlying machine-learning framework. Combining radar with lightning predic-
 330 tors generates a well-functioning algorithm. Our final product is more skilled in distinguishing gust-positive from gust-negative thunderstorms than the baseline.

D) The primary downburst hotspot is the Southern Prealpine range, and the secondary hotspot is the Northern Prealpine range.

Our work attempts to use crowd-sourced wind reports and examines downbursts in Switzerland. We show that machine learn-
 335 ing offers a robust framework for classifying convective storm cells based on their potential to produce damaging winds. The Burstinator Random Forest algorithm developed in this study can diagnose thunderstorms in real-time, is physically grounded and interpretable, and can thus be applied as an operational product. However, uncertainties related to the ground truth must be considered: The ground truth was constructed by linking gust markers to TRT storm cells; however, the dataset retains uncertainties regarding report reliability, temporal accuracy, and the passive assumption that an absence of reports in populated
 340 daytime areas implies no severe winds occurred.

Despite this, we strongly encourage the use of crowd-sourced reports for model training, as they provide crucial data, making machine learning more effective. Furthermore, using these untouched data, we developed a skillful algorithm in the field of downburst detection - a field where operational products are scarce to non-existent.



Code and data availability. The code is available at <https://github.com/florianvalentinackermann-oss/Burstinator>, the data is available at
345 10.5281/zenodo.22031097.



Appendix A: MeteoSwiss WDRAFT formula

$$WDRAFT = \left(20.63 ms^{-2} \cdot VIL - 3.125 \times 10^{-6} s^{-2} \cdot EchoTop20^2 \right)^{0.5} + \frac{velocity}{10.8} \quad (A1)$$



Appendix B: Full Predictor List

Predictor	Description	Source
$Z > 35$	Volume-integrated percentage of pixels where Z exceeds 35 dBZ	Composite
$Z > 40$	Volume-integrated percentage of pixels where Z exceeds 40 dBZ	Composite
$Z > 45$	Volume-integrated percentage of pixels where Z exceeds 45 dBZ	Composite
$Z \text{ EchoTop } 45$	Maximum height where Z equals 45 dBZ	TRT Algorithm
$Z \text{ COM alt}$	Volume-integrated height of the center of mass of Z	Composite
$KDP > 0.5$	Volume-integrated percentage of pixels where K_{DP} exceeds $0.5^\circ \text{ km}^{-1}$	Composite
$KDP > 1$	Volume-integrated percentage of pixels where K_{DP} exceeds 1° km^{-1}	Composite
$KDP > 1.5$	Volume-integrated percentage of pixels where K_{DP} exceeds $1.5^\circ \text{ km}^{-1}$	Composite
$KDP > 2$	Volume-integrated percentage of pixels where K_{DP} exceeds 2° km^{-1}	Composite
$RS > 0.5$	Volume-integrated percentage of pixels where radial shear exceeds $0.5 \text{ m s}^{-1} \text{ km}^{-1}$	Composite
$RS > 1$	Volume-integrated percentage of pixels where radial shear exceeds $1 \text{ m s}^{-1} \text{ km}^{-1}$	Composite
$RS > 1.5$	Volume-integrated percentage of pixels where radial shear exceeds $1.5 \text{ m s}^{-1} \text{ km}^{-1}$	Composite
$RS > 2$	Volume-integrated percentage of pixels where radial shear exceeds $2 \text{ m s}^{-1} \text{ km}^{-1}$	Composite
$RS > 2.5$	Volume-integrated percentage of pixels where radial shear exceeds $2.5 \text{ m s}^{-1} \text{ km}^{-1}$	Composite
$RS \text{ max alt}$	Volume-integrated height of the maximum radial shear value	Composite
CGL	Cloud-to-Ground lightning	TRT Algorithm
ICL*	Intra-Cloud lightning	TRT Algorithm
TotL*	Total lightning (CGL + ICL)	TRT Algorithm

Table B1. Table overview of predictors, with asterisked predictors incorporating lagged values from 5, 10, and 15 min before the event time step.



Appendix C: Performance Metrics

350 We use the following scores to evaluate the Burstinator. Events associated with gusts are positives, whereas events unassociated with gusts are negatives.

Actual / Predicted	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table C1. Contingency table showing the terms and relationships of actual and predicted classes in a binary classification model.

$$\text{POD/Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (\text{C1})$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (\text{C2})$$

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{POD}}{\text{Precision} + \text{POD}} \quad (\text{C3})$$

355
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (\text{C4})$$

$$\text{Macro Average} = \frac{1}{N} \sum_{i=1}^N s_i \quad \text{where } s_i \text{ is the score for class } i \quad (\text{C5})$$

$$\text{Weighted Average} = \frac{\sum_{i=1}^N w_i s_i}{\sum_{i=1}^N w_i} \quad \text{where } w_i \text{ is the support of class } i \quad (\text{C6})$$

$$\text{FAR} = \frac{\text{FP}}{\text{TP} + \text{FP}} \quad (\text{C7})$$

$$\text{CSI} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (\text{C8})$$



$$360 \quad \text{HSS} = \frac{2(\text{TP} \times \text{TN} - \text{FP} \times \text{FN})}{(\text{TP} + \text{FN})(\text{FN} + \text{TN}) + (\text{TP} + \text{FN})(\text{FP} + \text{TN})} \quad (\text{C9})$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (\text{C10})$$

$$\text{SR} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (\text{C11})$$



Author contributions. This work was conducted as F.A.'s MSc thesis. D.N., S. B., J.G., M.A., and M.F. supervised the thesis and designed the project. M.A. contributed to code development and data processing. A.H. provided input on the use of operational thunderstorm products
365 at MeteoSwiss. F.A. wrote the manuscript and created the figures. All authors provided editorial input and feedback on the final manuscript.

Competing interests. The position of M.F. at the University of Bern was funded by the Mobiliar Insurance Group. This had no influence any part of this study.

Acknowledgements. We sincerely thank Olivia Martius for her guidance during the thesis. We further thank Daniel Wolfensberger and Ulrich Hamann for valuable discussions. The position of M.A. was funded through the SNSF Synergia Project scClim (CRSII5_201792).



370 References

- Ackermann, F.: Radar Diagnostics for Convective Wind Gusts, https://www.geography.unibe.ch/unibe/portal/fak_naturwis/e_geowiss/c_igeogr/content/e39592/e1608235/e1611002/pane1611003/e1745421/Ackermann.Florian_2025-compressed_ger.pdf, 2025.
- Amiot, C. G., Carey, L. D., Roeder, W. P., McNamara, T. M., and Blakeslee, R. J.: C-Band Dual-Polarization Radar Signatures of Wet Downbursts around Cape Canaveral, Florida, *Weather and Forecasting*, 34, 103–131, <https://doi.org/10.1175/WAF-D-18-0081.1>, 2019.
- 375 Aregger, M., Martius, O., Germann, U., and Hering, A.: Hail detection and sizing for storm objects: A random forest framework using Swiss C-band radar and crowdsourced reports, *Artificial Intelligence for the Earth Systems*, -1, <https://doi.org/10.1175/AIES-D-26-0001.1>, 2026.
- Atkins, N. T. and Wakimoto, R. M.: Wet Microburst Activity over the Southeastern United States: Implications for Forecasting, *Weather and Forecasting*, 6, 470–482, [https://doi.org/10.1175/1520-0434\(1991\)006<0470:WMAOTS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1991)006<0470:WMAOTS>2.0.CO;2), 1991.
- Baker, M., Blyth, A., Christian, H., Latham, J., Miller, K., and Gadian, A.: Relationships between Lightning Activity and Various Thunder-
 380 cloud Parameters: Satellite and Modelling Studies, *Atmospheric Research*, 51, 221–236, [https://doi.org/10.1016/S0169-8095\(99\)00009-5](https://doi.org/10.1016/S0169-8095(99)00009-5), 1999.
- Berrar, D.: Cross-Validation, in: *Encyclopedia of Bioinformatics and Computational Biology*, pp. 542–545, Elsevier, <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>, 2019.
- Biau, G. and Scornet, E.: A Random Forest Guided Tour, *TEST*, 25, 197–227, <https://doi.org/10.1007/s11749-016-0481-7>, 2016.
- 385 Böhm, J. P.: A General Hydrodynamic Theory for Mixed-Phase Microphysics. Part I: Drag and Fall Speed of Hydrometeors, *Atmospheric Research*, 27, 253–274, [https://doi.org/10.1016/0169-8095\(92\)90035-9](https://doi.org/10.1016/0169-8095(92)90035-9), 1992.
- Bonnet, M.: Status of lightning detection performances and limitations of existing systems, in: *International Colloquium on Lightning and Power systems*, 2019.
- Breiman, L.: Random Forests, *Machine Learning*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- 390 Brönnimann, S.: Climate of the Free Troposphere and Mountain Peaks, pp. 1–20, Oxford University Press, <https://doi.org/10.1093/acrefore/9780190228620.013.755>, 2020.
- Chase, R. J., Harrison, D. R., Burke, A., Lackmann, G. M., and McGovern, A.: A Machine Learning Tutorial for Operational Meteorology. Part I: Traditional Machine Learning, *Weather and Forecasting*, 37, 1509–1529, <https://doi.org/10.1175/WAF-D-22-0070.1>, 2022.
- Dotzek, N., Groenemeijer, P., Feuerstein, B., and Holzer, A. M.: Overview of ESSL’s Severe Convective Storms Research Using the European
 395 Severe Weather Database ESWD, *Atmospheric Research*, 93, 575–586, <https://doi.org/10.1016/j.atmosres.2008.10.020>, 2009.
- Erdmann, F., Defer, E., Caumont, O., Blakeslee, R. J., Pédeboy, S., and Coquillat, S.: Concurrent Satellite and Ground-Based Lightning Observations from the Optical Lightning Imaging Sensor (ISS-LIS), the Low-Frequency Network Meteorage and the SAETTA Lightning Mapping Array (LMA) in the Northwestern Mediterranean Region, *Atmospheric Measurement Techniques*, 13, 853–875, <https://doi.org/10.5194/amt-13-853-2020>, 2020.
- 400 Fabry, F.: *Radar Meteorology: Principles and Practice*, Cambridge University Press, 1 edn., <https://doi.org/10.1017/CBO9781107707405>, 2015.
- Farnell, C., Rigo, T., and Pineda, N.: Lightning Jump as a Nowcast Predictor: Application to Severe Weather Events in Catalonia, *Atmospheric Research*, 183, 130–141, <https://doi.org/10.1016/j.atmosres.2016.08.021>, 2017.
- Feldmann, M.: Supercell Thunderstorms in the Alpine Region - From Weather Radar Observations to Idealized Modeling, Ph.D. thesis, Lausanne, EPFL, <https://doi.org/10.5075/EPFL-THESIS-10232>, 2023.
- 405



- Feldmann, M., James, C. N., Boscacci, M., Leuenberger, D., Gabella, M., Germann, U., Wolfensberger, D., and Berne, A.: R2D2: A Region-Based Recursive Doppler Dealiasing Algorithm for Operational Weather Radar, *Journal of Atmospheric and Oceanic Technology*, 37, 2341–2356, <https://doi.org/10.1175/JTECH-D-20-0054.1>, 2020.
- Feldmann, M., Hering, A., Gabella, M., and Berne, A.: Hailstorms and Rainstorms versus Supercells—a Regional Analysis of Convective Storm Types in the Alpine Region, *npj Climate and Atmospheric Science*, 6, 19, <https://doi.org/10.1038/s41612-023-00352-z>, 2023.
- 410 Fujita, T. T.: Tornadoes and Downbursts in the Context of Generalized Planetary Scales, *Journal of the Atmospheric Sciences*, 38, 1511–1534, [https://doi.org/10.1175/1520-0469\(1981\)038<1511:TADITC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<1511:TADITC>2.0.CO;2), 1981.
- Fujita, T. T.: Downbursts: Meteorological Features and Wind Field Characteristics, *Journal of Wind Engineering and Industrial Aerodynamics*, 36, 75–86, [https://doi.org/10.1016/0167-6105\(90\)90294-M](https://doi.org/10.1016/0167-6105(90)90294-M), 1990.
- 415 Fujita, T. T. and Wakimoto, R. M.: Five Scales of Airflow Associated with a Series of Downbursts on 16 July 1980, *Monthly Weather Review*, 109, 1438–1456, [https://doi.org/10.1175/1520-0493\(1981\)109<1438:FSOAAW>2.0.CO;2](https://doi.org/10.1175/1520-0493(1981)109<1438:FSOAAW>2.0.CO;2), 1981.
- Gagne, D. J., McGovern, A., Haupt, S. E., Sobash, R. A., Williams, J. K., and Xue, M.: Storm-Based Probabilistic Hail Forecasting with Machine Learning Applied to Convection-Allowing Ensembles, *Weather and Forecasting*, 32, 1819–1840, <https://doi.org/10.1175/WAF-D-17-0010.1>, 2017.
- 420 Gensini, V. A., Converse, C., Ashley, W. S., and Taszarek, M.: Machine Learning Classification of Significant Tornadoes and Hail in the U.S. Using ERA5 Proximity Soundings, *Weather and Forecasting*, <https://doi.org/10.1175/WAF-D-21-0056.1>, 2021.
- Germann, U., Boscacci, M., Gabella, M., and Schneebeli, M.: Weather Radar in Switzerland, in: *From Weather Observations to Atmospheric and Climate Sciences in Switzerland*, edited by Willemse, S. and Furger, M., pp. 165–188, vdf Hochschulverlag AG ETH Zürich, Zürich, 2016.
- 425 Germann, U., Boscacci, M., Clementi, L., Gabella, M., Hering, A., Sartori, M., Sideris, I. V., and Calpini, B.: Weather Radar in Complex Orography, *Remote Sensing*, 14, 503, <https://doi.org/10.3390/rs14030503>, 2022.
- Gregorutti, B., Michel, B., and Saint-Pierre, P.: Correlation and variable importance in random forests, *Statistics and Computing*, 27, 659–678, <https://doi.org/10.1007/s11222-016-9646-1>, 2017.
- Hering, A., Morel, C., Sényi, S., Galli, G., Ambrosetti, P., and Boscacci, M.: Nowcasting Thunderstorms in the Alpine Region Using a Radar Based Adaptive Thresholding Scheme, *Proceedings of ERAD*, pp. 206–211, 2004.
- 430 Hill, A. J. and Schumacher, R. S.: Forecasting Excessive Rainfall with Random Forests and a Deterministic Convection-Allowing Model, *Weather and Forecasting*, <https://doi.org/10.1175/WAF-D-21-0026.1>, 2021.
- Hill, A. J., Herman, G. R., and Schumacher, R. S.: Forecasting Severe Weather with Random Forests, *Monthly Weather Review*, 148, 2135–2161, <https://doi.org/10.1175/MWR-D-19-0344.1>, 2020.
- 435 Isaminger, M. A.: A Preliminary Study of Precursors to Huntsville Microbursts, 1988.
- Kopp, J., Hering, A., Germann, U., and Martius, O.: Verification of weather-radar-based hail metrics with crowdsourced observations from Switzerland, *Atmospheric Measurement Techniques*, 17, 4529–4552, <https://doi.org/10.5194/amt-17-4529-2024>, 2024.
- Kotarba, A.: Trends in Severe Convective Storm Activity over Europe (1983–2024), <https://doi.org/10.5194/egusphere-egu26-4918>, 2026.
- Kuster, C. M., Heinselman, P. L., and Schuur, T. J.: Rapid-Update Radar Observations of Downbursts Occurring within an Intense Multicell Thunderstorm on 14 June 2011, *Weather and Forecasting*, 31, 827–851, <https://doi.org/10.1175/WAF-D-15-0081.1>, 2016.
- 440 Kuster, C. M., Bowers, B. R., Carlin, J. T., Schuur, T. J., Brogden, J. W., Toomey, R., and Dean, A.: Using KDP Cores as a Downburst Precursor Signature, *Weather and Forecasting*, 36, 1183–1198, <https://doi.org/10.1175/WAF-D-21-0005.1>, 2021.



- Leinonen, J., Hamann, U., Sideris, I. V., and Germann, U.: Thunderstorm Nowcasting With Deep Learning: A Multi-Hazard Data Fusion Model, *Geophysical Research Letters*, 50, e2022GL101 626, <https://doi.org/10.1029/2022GL101626>, 2023.
- 445 Loken, E. D., Clark, A. J., and McGovern, A.: Comparing and Interpreting Differently Designed Random Forests for Next-Day Severe Weather Hazard Prediction, *Weather and Forecasting*, 37, 871–899, <https://doi.org/10.1175/WAF-D-21-0138.1>, 2022.
- Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, in: *Advances in Neural Information Processing Systems*, edited by Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., vol. 30, Curran Associates, Inc., https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf, 2017.
- 450 Mahale, V. N., Zhang, G., and Xue, M.: Characterization of the 14 June 2011 Norman, Oklahoma, Downburst through Dual-Polarization Radar Observations and Hydrometeor Classification, *Journal of Applied Meteorology and Climatology*, 55, 2635–2655, <https://doi.org/10.1175/JAMC-D-16-0062.1>, 2016.
- Mahalik, M. C., Smith, B. R., Elmore, K. L., Kingfield, D. M., Ortega, K. L., and Smith, T. M.: Estimates of Gradients in Radar Moments Using a Linear Least Squares Derivative Technique, *Weather and Forecasting*, 34, 415–434, <https://doi.org/10.1175/WAF-D-18-0095.1>,
 455 2019.
- Markowski, P. and Richardson, Y.: *Mesoscale Meteorology in Midlatitudes*, John Wiley & Sons, Chichester, UK, 2010.
- McGovern, A., Elmore, K. L., Gagne, D. J., Haupt, S. E., Karstens, C. D., Lagerquist, R., Smith, T., and Williams, J. K.: Using Artificial Intelligence to Improve Real-Time Decision-Making for High-Impact Weather, *Bulletin of the American Meteorological Society*, 98, 2073–2090, <https://doi.org/10.1175/BAMS-D-16-0123.1>, 2017.
- 460 MétéoSuisse: Orage dévastateur du 24 juillet 2023 à La Chaux-de-Fonds, https://www.meteosuisse.admin.ch/dam/jcr:419af3d7-60c2-4207-8ed5-f3c555e8689a/Orage-La-Chaux-de-Fonds_Arbeitsbericht.pdf, 2024.
- MeteoSwiss: Meteo reports, <https://www.meteoswiss.admin.ch/services-and-publications/applications/meteo-reports.html#tab=meteo-reports-map>, 2026.
- Parodi, A., Lagasio, M., Maugeri, M., Turato, B., and Gallus, W.: Observational and Modelling Study of a Major Downburst Event in Liguria:
 465 The 14 October 2016 Case, *Atmosphere*, 10, 788, <https://doi.org/10.3390/atmos10120788>, 2019.
- Pilorz, W., Laskowski, I., Surowiecki, A., and Łupikasza, E.: Fatalities Related to Sudden Meteorological Events across Central Europe from 2010 to 2020, *International Journal of Disaster Risk Reduction*, 88, 103 622, <https://doi.org/10.1016/j.ijdr.2023.103622>, 2023.
- Pistotnik, G., Holzer, A. M., Kaltenböck, R., and Tschannett, S.: An F3 Downburst in Austria—A Case Study with Special Focus on the Importance of Real-Time Site Surveys, *Atmospheric Research*, 100, 565–579, <https://doi.org/10.1016/j.atmosres.2010.10.011>, 2011.
- 470 Prein, A. F.: Thunderstorm Straight Line Winds Intensify with Climate Change, *Nature Climate Change*, 13, 1353–1359, <https://doi.org/10.1038/s41558-023-01852-9>, 2023.
- Prudden, R., Adams, S., Kangin, D., Robinson, N., Ravuri, S., Mohamed, S., and Arribas, A.: A Review of Radar-Based Nowcasting of Precipitation and Applicable Machine Learning Techniques, <https://doi.org/10.48550/ARXIV.2005.04988>, 2020.
- Pédeboy, S., Barnéoud, P., Defer, E., and Coquillat, S.: Analysis of the Intra-Cloud lightning activity detected with Low Frequency Lightning
 475 Locating Systems, in: *25th Lightning Detection Conference & 7th International Lightning Meteorology Conference*, Fort Lauderdale, Florida, USA, 2018.
- Púčik, T., Groenemeijer, P., Rädler, A. T., Tijssen, L., Nikulin, G., Prein, A. F., Van Meijgaard, E., Fealy, R., Jacob, D., and Teichmann, C.: Future Changes in European Severe Convection Environments in a Regional Climate Model Ensemble, *Journal of Climate*, 30, 6771–6794, <https://doi.org/10.1175/JCLI-D-16-0777.1>, 2017.



- 480 Schulz, W., Diendorfer, G., Pedebay, S., and Poelman, D. R.: The European Lightning Location System EUCLID – Part 1: Performance
 Analysis and Validation, *Natural Hazards and Earth System Sciences*, 16, 595–605, <https://doi.org/10.5194/nhess-16-595-2016>, 2016.
- Smith, B. R., Sandmæl, T., Mahalik, M. C., Elmore, K. L., Kingfield, D. M., Ortega, K. L., and Smith, T. M.: Corrigendum, *Weather and
 Forecasting*, 36, 1131–1133, <https://doi.org/10.1175/WAF-D-20-0125.1>, 2021.
- Smith, T. M., Elmore, K. L., and Dulin, S. A.: A Damaging Downburst Prediction and Detection Algorithm for the WSR-88D, *Weather and
 Forecasting*, 19, 240–250, [https://doi.org/10.1175/1520-0434\(2004\)019<0240:ADDPAD>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0240:ADDPAD>2.0.CO;2), 2004.
- 485 Trefalt, S.: Hail and Severe Wind Gusts in the Convective Season in Switzerland, Ph.D. thesis, University of Bern, Bern, 2017.
- Vulpiani, G., Montopoli, M., Passeri, L. D., Gioia, A. G., Giordano, P., and Marzano, F. S.: On the Use of Dual-Polarized C-Band
 Radar for Operational Rainfall Retrieval in Mountainous Areas, *Journal of Applied Meteorology and Climatology*, 51, 405–425,
<https://doi.org/10.1175/JAMC-D-10-05024.1>, 2012.
- 490 Williams, E. R., Weber, M. E., and Orville, R. E.: The Relationship between Lightning Type and Convective State of Thunderclouds, *Journal
 of Geophysical Research: Atmospheres*, 94, 13 213–13 220, <https://doi.org/10.1029/JD094iD11p13213>, 1989.
- Wolfensberger, D., Gabella, M., Boscacci, M., Germann, U., and Berne, A.: RainForest: A Random Forest Algorithm for Quantitative
 Precipitation Estimation over Switzerland, *Atmospheric Measurement Techniques*, 14, 3169–3193, [https://doi.org/10.5194/amt-14-3169-
 2021](https://doi.org/10.5194/amt-14-3169-2021), 2021.
- 495 Xiao, H., Wang, Y., Zheng, Y., Zheng, Y., Zhuang, X., Wang, H., and Gao, M.: Convective-Gust Nowcasting Based on Radar Reflectivity
 and a Deep Learning Algorithm, *Geoscientific Model Development*, 16, 3611–3628, <https://doi.org/10.5194/gmd-16-3611-2023>, 2023.