

Quantile Regression for Power Grid Damage Prediction

Arnis Staško^{1*}, Andrejs Bondarenko¹, Eduards Otomers¹, Dmitrijs Bļizņuks¹,
Ilze Birzniece¹, Rihards Sirmāis², Kārlis Kristofers Vēliņš², and
Jānis Grundspenķis¹

¹Institute of Applied Computer Systems, Riga Technical University, 6A Kipsalas Street,
1048, Riga, Latvia

²JSC “Sadales tīkls”, 1 Šmerļa Street, 1160, Riga, Latvia

arnis.stasko@rtu.lv, andrejs.bondarenko@rtu.lv, eduards.otomers@rtu.lv,
dmitrijs.bliznuks@rtu.lv, ilze.birzniece@rtu.lv, rihards.sirmais@sadalestikls.lv,
karlis-kristofers.velins@sadalestikls.lv, janis.grundspenkis@rtu.lv

Abstract. This article addresses the challenge of predicting electrical grid damage in the Latvian regions of Zemgale and Kurzeme using machine learning and quantile regression techniques. The study leverages a comprehensive dataset spanning 2009–2025, incorporating environmental, infrastructure, seasonal, and weather-related features. To handle the significant data imbalance, various spatial and temporal aggregation strategies are tested. Feature selection and aggregation are carefully tailored to preserve correlations between predictors and damage counts. The XGBoost algorithm, configured for quantile regression, is employed with extensive hyperparameter tuning and cross-validation. Results indicate that models trained at the maintenance-region scale (6-cell) achieve the best overall balance between accuracy and reliability, with a mean absolute error of 4.09 and an R2 of 0.53 for daily damage prediction. While the state region (2-cell) model excels in predicting high damage count events (30 or more), it also provides more precise minimum and maximum bounds, offering actionable uncertainty bounds for operational planning. The findings highlight the importance of spatial aggregation and feature engineering in improving predictive performance for rare-event forecasting in power grids. Several promising directions for further development have been identified, and the obtained results are sufficiently promising to support the application of the proposed methodology to broader geographical regions.

Keywords: Electrical Grid, Damage Prediction, Machine Learning, Quantile Regression, Climate Factors.

* Corresponding author

© 2026 Arnis Staško, Andrejs Bondarenko, Eduards Otomers, Dmitrijs Bļizņuks, Ilze Birzniece, Rihards Sirmāis, Kārlis Kristofers Vēliņš, and Jānis Grundspenķis. This is an open access article licensed under the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).

Reference: A. Staško, A. Bondarenko, E. Otomers, D. Bļizņuks, I. Birzniece, R. Sirmāis, K.K. Vēliņš, J. Grundspenķis, “Quantile Regression for Power Grid Damage Prediction”, Complex Systems Informatics and Modeling Quarterly, CSIMQ, no. 48, pp. 76–93, 2026. Available: <https://doi.org/10.7250/csimq.2026-48.04>

Additional information. Author ORCID iD: A. Staško – <https://orcid.org/0000-0003-2876-5497>, A. Bondarenko – <https://orcid.org/0000-0002-4103-6814>, E. Otomers – <https://orcid.org/0009-0002-7092-6389>, D. Bļizņuks – <https://orcid.org/0000-0003-4252-9220>, I. Birzniece – <https://orcid.org/0000-0002-5775-6138>, J. Grundspenķis – <https://orcid.org/0000-0003-2526-4662>, PII S225599222600270X. Received: 17 July 2026. Accepted: 28 October 2026. Available online: 31 October 2026.

1 Introduction

Electrical power grid networks are prone to seasonal and weather-related outages. These outages can be caused by extreme wind, increased humidity, precipitation, the arrival of birds, or other factors. When more damage occurs that the on-call repair crews can handle simultaneously, the time customers spend without electricity increases significantly. On the other hand, keeping too many repair crews on active duty increases the costs of maintaining the network. In either case, the end consumer suffers. It is essential to predict damage counts to set the plans of repair crews on duty and promptly address the damage that has occurred.

Consequently, this challenge is inherently structured as a socio-physical-cybernetic system, in which three interconnected domains dynamically interact. In the physical domain, grid assets (e.g., overhead distribution lines, aging wooden poles) interact with environmental and meteorological stressors (wind gusts, temperature, precipitation). In the socio domain, human and operational dimensions predominate, including repair crew scheduling, labor cost constraints, and customer quality of life. In the cybernetic domain, data-driven feedback and control mechanisms driven by predictive machine learning models process physical weather inputs to optimize social decision-making and crew dispatch strategies. Viewing damage prediction through this framework ensures that model predictions are not evaluated in statistical isolation, but as decision-support signals that directly balance physical grid safety against social and economic operational costs.

To operationalize this socio-physical-cybernetic perspective within a real-world distribution grid, the authors aim to investigate how accurately it is possible to predict the number of faults in Latvian electrical networks, focusing on the regions of Zemgale and Kurzeme as representative case studies. Kurzeme is highly exposed to coastal maritime fronts with severe wind gust dynamics, whereas Zemgale features flat inland agricultural terrain subject to soil saturation and river basin flooding. Together, they encompass a diverse mix of overhead electricity line infrastructure.

Meteorological and climate-related environmental conditions are among the primary external drivers of power grid reliability. Rapidly changing meteorological variables, such as wind speed, precipitation, temperature, humidity, and atmospheric pressure, directly influence the occurrence of weather-induced failures on hourly and daily timescales. At the same time, more persistent climate and environmental characteristics, including proximity to the sea, long-term soil moisture conditions, and seasonal patterns, shape the baseline vulnerability of electrical infrastructure by modifying local weather exposure and long-term asset deterioration. Together, these short-term meteorological conditions and longer-term climate factors essentially affect the resilience of power distribution networks.

This study contributes to the growing body of research on weather-induced failures by investigating how both dynamic meteorological conditions and more stable climate-related environmental factors can be incorporated into predictive models of electrical network failures. We evaluate the predictive value of rapidly evolving weather observations together with geographically persistent variables across different spatial aggregation levels. Although the objective is not to attribute failures to climate change, the proposed methodology provides insights into how weather variability and regional climate characteristics jointly influence power system reliability. The findings, therefore, contribute to climate adaptation research by supporting improved operational preparedness and more resilient management of electrical infrastructure under varying environmental conditions.

Various studies have explored the prediction of power grid failures, but each operates in a specific region, at a different scale, using different datasets and prediction models, or even with different forecasting goals. A study of Orange and Rockland Utilities, Inc.'s electrical network in New York State, US [1], explores LSTM capabilities for predicting the number of outages. A study in the Guangdong region of China [2] explores the use of various convolutional networks to predict whether an electricity line will experience more than four outages per day. Another study in 29 counties across eight states of the US [3] proposes using the NGBoost model to predict the

percentage of customers without power. A study in the Albacete and Valencia regions of Spain [4] focuses on deep learning ensemble methods and predicting either incident count or extreme incident count events. Lastly, in China Southern Power Grid Co.'s electrical network [5], the focus is on classifying fault types and fault durations rather than predicting the exact number of faults. A comprehensive related literature review on outage prediction research is published in our previous work [6]. Existing studies generally combine meteorological data with power grid outage data to predict damage either for extreme events, such as cyclones or hurricanes, or for longer time spans. The review highlights that prior research has focused on predicting power grid failures in extreme weather conditions, and that models developed for extended periods suffer from significant data imbalance. The review also noted that no studies were found concerning the Baltic states, with Finland representing the closest geographically comparable context to Latvia.

The electrical network damage prediction task faces severe dataset imbalance, with cause and effect at a detailed level but low probability. When dividing geographical space into grid cells and analyzing hourly data, only a few cell/hour pairs contain damage events. There are many more samples without damage than with damage under the same weather conditions. When aggregating cells and/or hours, weather observations show similar extreme values or relatively normal mean values per grid cell and/or time interval. Under these conditions, conventional mean-based regression techniques fail, as they optimize for expected average outcomes and, consequently, underestimate rare damage cases. Therefore, quantile regression [7] is specifically chosen to overcome these limitations because it models the conditional quantiles of the damage response rather than the conditional mean.

The article is organized as follows. Section 2 outlines the research methodology, Section 3 introduces domain characteristics, Section 4 describes the feature selection, Section 5 continues with the feature aggregation approach, Section 6 describes model training, validation, and hyperparameter selection, Section 7 discusses the results, and Section 8 concludes the article and outlines directions for future research.

2 Research Methodology

This study uses the design science research methodology (DSRM) [8] as its goal is to design, develop, and validate a machine learning model for power grid damage prediction. Consequently, the research is organized according to the following DSRM activities: problem identification and motivation, definition of objectives, design and development, demonstration, evaluation, and communication [9].

2.1 Problem Identification and Motivation

In extreme weather conditions, the number of damages in electricity grid networks increases significantly. Especially considering the recent changes in weather conditions in Latvia, with increasingly frequent heavy rains and strong gusts of wind. Technical repair crew resources planned for normal conditions become insufficient. If all available repair crew's resources are occupied fixing damage, then those reporting subsequent damages must wait longer to receive assistance. Households left without electricity both affect people's quality of life and reduce the grid operator's income due to the services not being provided. On the other hand, increasing the size of the repair crew disproportionately increases maintenance costs if the employees are not busy. In both cases, costs increase for the service recipients.

While existing literature addresses electrical damage prediction, current work regarding Eastern Europe is limited. Earlier works present damage prediction in Finland [10] and Norway [11]. Research presented in [10] focuses on convective storms and the damage caused during these storms. In contrast, [11] deals with transmission line damage prediction, which have different characteristics compared to those of distribution power lines. Therefore, additional research is

needed on damage prediction in electricity distribution networks, taking into account various weather conditions in Eastern Europe.

2.2 Objectives for a Solution

The goal of the solution is to predict the daily number of electrical network damages so that it is possible to warn service recipients promptly and plan the amount of repair crew resources. Considering the probabilistic nature of damage, where the number of historical damages experienced under similar conditions varies, and the limited results of our previous research [6], the damage forecast should be expressed as a probabilistic forecast bounded by lower and upper prediction intervals.

2.3 Design and Development

The design and development of machine learning models is based on the machine learning workflow adopted from [12] as presented in Figure 1.

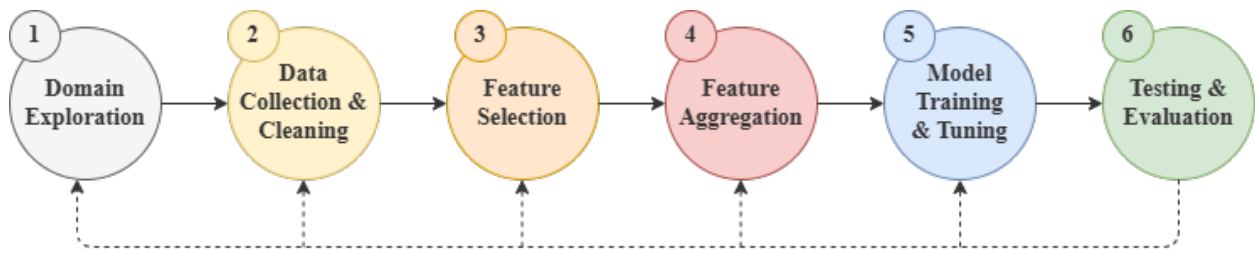


Figure 1. Machine learning workflow

During the domain exploration phase, interviews were conducted with JSC “Sadales tīkls”[†] network experts to discuss grid infrastructure parameters, identify historical damage causes, and evaluate data availability inside the organization, from partners, and open public sources. The identified domain characteristics are presented in Section 3. In the data gathering phase, primary infrastructure data and weather observations were collected. Additionally, public environmental open data were integrated, custom attributes engineered, and data cleaning and preprocessing performed.

Following preprocessing, features were evaluated and selected (Section 4) based on predictive relevance and data quality. Next, feature aggregation was performed, experimenting with different aggregation strategies and approaches (Section 5).

Model training and hyperparameter optimization were carried out using a cross-validation approach to prevent overfitting (Section 6). The resulting models were assessed on an independent testing dataset with metrics suitable for regression tasks. To understand and validate the underlying model logic, SHapley Additive exPlanations (SHAP) [13] were utilized to interpret feature impact on the model output. The days with increased damage counts were individually reviewed, and model predictions were evaluated.

Intermediate damage prediction results were continuously reviewed together with domain experts. Guided by expert feedback, the machine learning pipeline was iteratively refined, allowing for progressive improvement of the feature space, hyperparameters, and overall prediction relevance and usability.

2.4 Demonstration

To demonstrate the performance of the developed models, historical damage data from 2025 (which were not used for model training) is used, paying special attention to days when the number of damages reaches 30 or more per day (Section 7).

[†] <https://sadalestikls.lv/en>

2.5 Evaluation

To evaluate the model's performance, the testing data set as described in Section 2.4 is used. Results and evaluation notes are presented in Section 7. For each sample in the testing dataset, the prediction together with the lower and upper bounds were calculated. The predicted values were aggregated into a total daily damage forecast to measure all models at the same granularity level. This daily data aggregation aligns precisely with the operator's planning horizon, as repair crew shifts and logistics are managed based on a daily operational cycle.

To reflect both statistical validity and the specific nature of the tasks, model performance is evaluated using a combination of classical regression metrics and probabilistic prediction interval metrics. The mean absolute error (MAE) [14] value is used to measure the average error between the predicted and actual values. In the context of damage prediction, the results of this metric do not give full feedback on the accuracy of the model, since the absolute value treats overestimated and underestimated errors as the same, and particularly large errors are balanced out by small errors. In electricity grid emergency management, large miscalculations during extreme storms carry far worse consequences than small errors during routine conditions. Therefore, root mean squared error (RMSE) [14] is used to penalize large errors heavily, ensuring the model is evaluated on its robustness during high-impact events. While both MAE and RMSE are used to measure the error values of predictions, the coefficient's of determination (R^2) value [14] is used to measure the accuracy of the overall model and how well the model captures trends in the overall data. For regression tasks, the best possible R^2 value is 1, while lower MAE and RMSE values are preferred.

Specifically, for quantile regression, prediction interval coverage probability (PICP) was used [15]. It is possible to measure how often the actual value falls within the predicted interval. By using the mean prediction interval width (MPIW) [16], the average width of the predicted intervals was measured. Both of these metrics are used together to evaluate the accuracy and fit of the predicted intervals.

Standard metrics treat overestimation and underestimation symmetrically. In utility grid management, however, costs are inherently asymmetric: underestimating damage results in unserved demand, severe regulatory penalties, and public safety risks, whereas overestimating damage leads to idle crew labor costs. Quantile loss [17] (also known as Pinball loss [18]) is selected to evaluate the model's performance across risk thresholds by penalizing overestimation and underestimation asymmetrically at specific quantiles (lower bound and upper bound). The cumulative quantile loss across the interval serves as the primary metric for overall probabilistic accuracy.

2.6 Communication

The developed models are being integrated into the JSC "Sadales tīkls" damage prediction solution for surface power grid damage prediction before a storm (ST-ODA), so that daily predictions are communicated within the organization and, in case of long term success in production, also shared with industry experts later. Additionally, the results of this study are made available to the scientific community through this publication.

3 Domain Characteristics

A historical dataset on power grid damage statistics over the Kurzeme and Zemgale regions of Latvia is used in this study. The dataset covers the period from 2009 to 2025, identifying local damage events on a map. Given that damage rarely affects the same power pole, it does not make sense to predict damage to each pole separately. For this research, several grid resolution options are considered (Figure 2): state regions (2 cells), maintenance regions (6 cells), county regions (11 cells), and weather observation regions (467 cells).



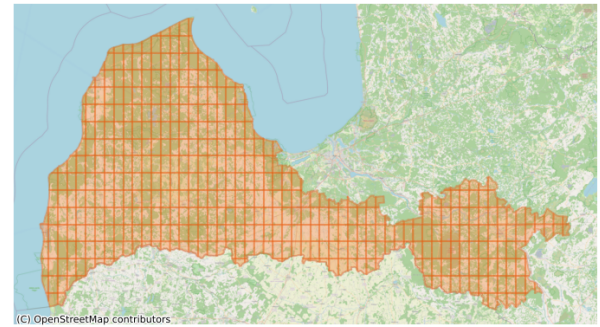
(a) State regions



(b) Maintenance regions



(c) County regions



(d) Weather observation regions

Figure 2. Grid resolution options

Available data contains observations from 6,209 days (149,016 hours). Combining these with grid resolution options, eight potential datasets are obtained: for state regions 12,088 and 290,112 records; for maintenance regions 36,264 and 870,336 records; for county regions 66,484 and 1,595,616 records; and for weather observation regions 2,822,548 and 67,741,152 records. Moving from higher resolution to lower resolution minimizes dataset imbalance (Figure 3). Considering that the imbalance in the hourly data is extremely high and the first experiments with regression models showed poor results (especially when evaluating the confusion matrices), either by missing almost all damage events or exceedingly overpredicting damage occurrences depending on the model's sensitivity, hourly datasets will not be used further in this study.

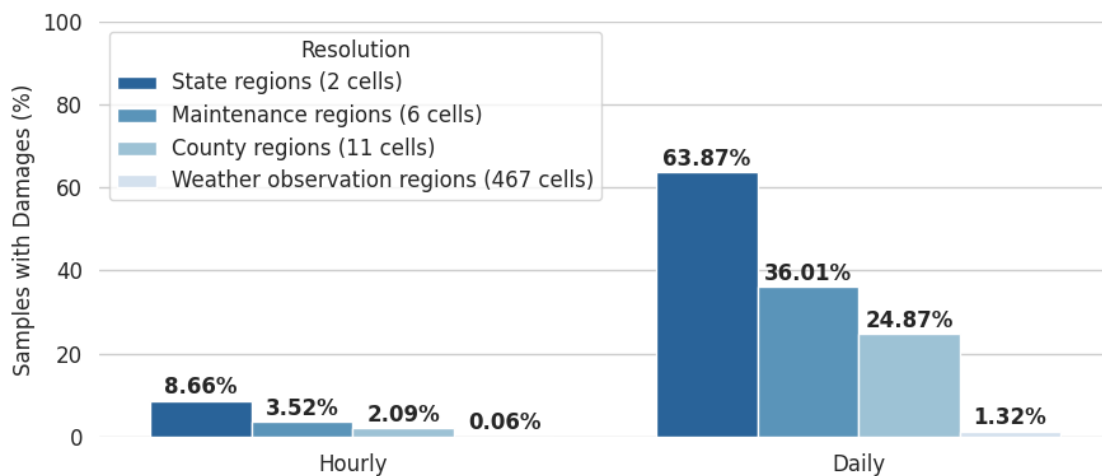


Figure 3. Dataset balance across space-time scales

While the dataset balance on a daily time scale shows many records with damages, the imbalance is transferred to the damage count distribution (Figure 4), containing fewer samples with high damage count per cell per day. In the next section, feature selection is discussed.

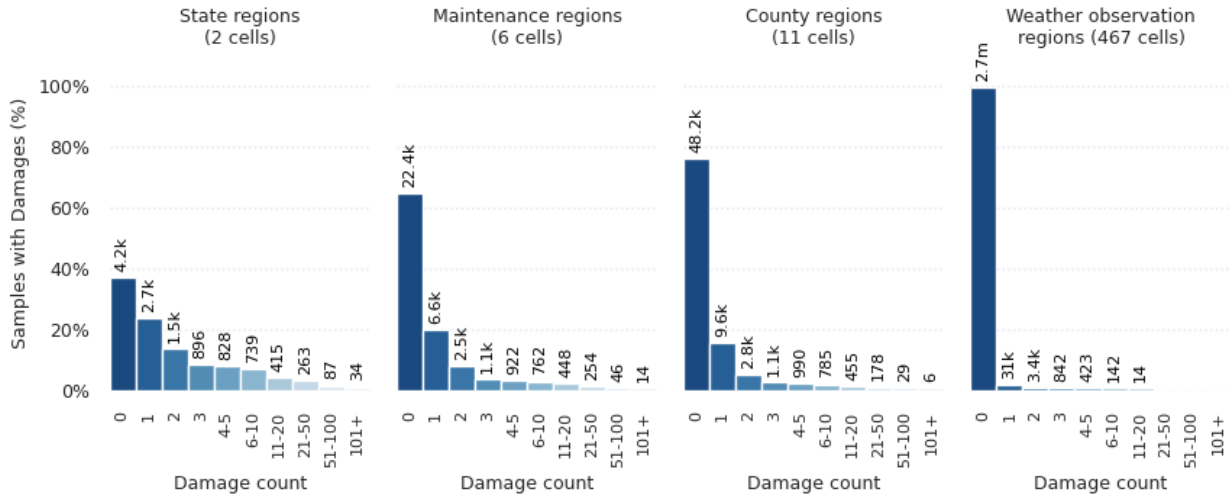


Figure 4. Dataset balance after daily aggregation across space scales

4 Feature Selection

Four groups of features were collected: environment characteristics, infrastructure characteristics, seasonal specifics, and weather observations.

Environment characteristics. To capture environmental characteristics for each grid cell, elevation above sea level and the distance to the Baltic Sea were calculated. Higher elevated regions could have a stronger wind impact than lower regions. Trees in regions closer to the coastline could have adapted to stronger winds than those in inland regions.

Infrastructure characteristics. Statistics about overhead electrical line length and line types (bare wire lines, lines with insulated wires, suspended cable lines) were collected. Likewise, information about electrical line locations (lines in populated areas, forests, orchards, scrublands, agricultural lands, swamps, meadows, production areas, near rivers, ditches, roads or railroads) was obtained. Regions with a higher number of overhead lines of a specific type and location may experience more power grid damage due to weather and natural conditions. As older wooden poles are more vulnerable, pole types and pole ages are also considered. Because historical asset data were unavailable for this research, past pole ages were reconstructed retrospectively. We decremented the known pole age annually; if the calculated age fell below zero, we simulated a replacement event by assuming the prior pole had been replaced at the age of 30 years.

Seasonal specifics. Sun elevation impacts heating of the air and consequently the movement of air masses affecting winds, storms, and precipitation. Wind impact on trees is higher with leaves than without them. According to [19], in Latvia, leaf growth begins in April and ends in the middle of May, while leaf discoloration and falling begin in the middle of September and end by late October. It is accepted that wind impact on trees without leaves is 40–50%, while 100% with full leaves [20]. Therefore, a leaf factor was calculated according to the rules presented in Table 1.

Weather observations. Historical hourly weather observation data from OpenMeteo.com [21] were used, including temperature, relative humidity, dew point, precipitation, rain, snowfall, surface pressure, boundary layer height, wind direction, wind speed, wind speed in gusts, and weather code. Except for soil moisture and soil temperature, daily historical data were used due to missing values in hourly data.

Table 1. Leaf factor calculation

Period	Leaf status	Leaf factor
Jan 1–Mar 31	No leaves	0.45
Apr 1–May 14	Growing leaves	gradually increasing from 0.45 to 1.00
May 15–Sep 14	Full leaves	1.00
Sep 14–Oct 31	Falling leaves	gradually decreasing from 1.00 to 0.45
Nov 1–Dec 31	No leaves	0.45

5 Feature Aggregation

The available data shows that at high detailing (grid size and time scale) exists signs of correlation between attributes and damage count. For instance, exploration of wind speed in gusts and actual damages locates damages in regions and hours with higher wind speeds in gusts. At the same time, in many grid cells and hours with the same wind speed in gusts, there is no damage at all.

When minimum (min), maximum (max), mean, and standard deviation (std) values are aggregated across multiple grid cells and/or into daily values, the overall picture becomes even less clear. Taking the maximum across cells or hours increases the share of records with high gust speeds, while taking the mean smooths the values into a daily average that carries little information. As a result, peaks in wind speed, whether in one grid cell or one hour, are no longer visible in the aggregated data. The key question is therefore how to aggregate attributes so that correlation is preserved and predictions improve. To address this challenge, we propose and apply the aggregation technique best suited to each attribute.

For the environmental attributes mean elevation and the mean and minimum distances to Baltic Sea for aggregated cells are calculated. For electric line type statistics and environment statistics, total length in meters is counted. For pole type statistics, a count of poles of a certain type for aggregated cells is calculated. For pole age statistics, bins of 0–5, 6–10, ..., 56–60 are calculated. For seasonally specific attributes, the leaf factor is the same at any scale, but for sun elevation the maximum value per day is used. For weather observation attributes, temperature, relative humidity, surface pressure, boundary layer height, wind speed, and wind speed in gusts, min, max, mean, and std values are calculated. For precipitation, rain, and snowfall, only the sum across all hours and/or grids is calculated. For snow depth, min, max, and mean are calculated. For soil moisture and soil temperature, only the mean value is used.

Special attention is paid to weather code, wind direction, wind speed, and wind speed in gusts. For the weather code, time integration is performed, keeping a value of how many hours across the day and/or joint grids the weather code was present. Hourly wind direction is normalized to the eight sides of the sky – “N”, “NE”, “E”, “SE”, “S”, “SW”, “W”, “NW”, and the integrated time by wind direction is used for aggregation. For wind speed and wind speed in gusts, the hour count with a certain speed threshold is calculated (fifteen attributes each hour; count of speed = 0, >0, >=5, >=10, >=15, >=16, >=17, >=18, >=19, >=20, >=21, >=22, >=23, >=24, >=25).

Additionally, after consultation with the Latvian Environment, Geology and Meteorology Centre[‡], a simplified joint attribute warning level was calculated (Table 2) for each hour, and then time integration was applied for aggregation, receiving four attributes: warn_level_0 – no warning, warn_level_1 – yellow warning, warn_level_2 – orange warning, warn_level_3 – red warning.

The full aggregated attribute list used for training is given in Appendix A. In the next section, the model training is presented.

[‡] SLLC “Latvian Environment, Geology and Meteorology Centre”, Available: <https://videscentrs.lv/gmc.lv/>

Table 2. Warning level thresholds

Warning level	Wind speed in gusts	Operand	Precipitation	Warning level code
Normal	less than 15 m/s	AND	less than 10 mm/3h	0
Yellow warning	from 15 m/s	OR	from 10 mm/3h	1
Orange warning	from 20 m/s	OR	from 20 mm/3h	2
Red warning	from 25 m/s	OR	from 30 mm/3h	3

6 Training & Hyperparameter Tuning

According to The State of Machine Learning Competitions 2025 Edition[§], gradient-boosted decision trees remain the primary tabular competition winner’s modeling tool, sometimes as part of an ensemble alongside neural nets. Favoring XGBoost [22] training speed, GPU support, and built-in ability to perform multiple quantile regression thresholds in a single model, it was selected for the experiments. For quantile regression, the following quantile thresholds are used: 0.05 for minimum bound prediction, 0.50 for base prediction, and 0.95 for maximum bound prediction. To stabilize the variance and handle the right-skewness of the target variable, log-transformation is used. To create a machine learning model with each of the datasets, hyperparameter search is performed using a cross-validation approach, followed by training and testing of the final model.

During hyperparameter search for each step, five models are built and tested using a cross-validation approach. For cross-validation, the training dataset is split as presented in Table 3.

Table 3. Cross-validation settings

Fold	Training data	Validation data
#1	2009–2019	2020
#2	2009–2020	2021
#3	2009–2021	2022
#4	2009–2022	2023
#5	2009–2023	2024

Hyperparameter search is managed by Optuna [23] with a limit of 500 trials according to the selected parameter space (Listing 1.1). For each step, mean cross-validation results of the quantile loss of five models are used to assess average model performance.

Listing 1.1. XGBoost hyperparameter search space

```

1 params = {
2     "objective": "reg:quantileerror",
3     "quantile_alpha": [0.05, 0.50, 0.95],
4     "device": "cuda",
5     "tree_method": "hist",
6     "early_stopping_rounds": 25,
7     "n_estimators": trial.suggest_int("n_estimators", 100, 1000, step=100),
8     "max_depth": trial.suggest_int("max_depth", 3, 10),
9     "learning_rate": trial.suggest_float("learning_rate", 0.01, 0.2, log=True),
10    "subsample": trial.suggest_float("subsample", 0.6, 1.0),
11    "colsample_bytree": trial.suggest_float("colsample_bytree", 0.6, 1.0),
12    "random_state": 42
13 }
```

[§] <https://mlcontests.com/state-of-machine-learning-competitions-2025/>

Optuna search results and the selected hyperparameters for each model/scale are presented in Table 4.

Table 4. Hyperparameter search results

Model Parameter	State regions (2 cells)	Maintenance regions (6 cells)	County regions (11 cells)	Weather observation regions (467 cells)
Duration [‡]	23 minutes	25 minutes	25 minutes	107 minutes
Best trial	432	448	371	173
Quantile loss	1.058	0.474	0.298	0.016
XGBoost hyperparameters				
n_estimators	600	500	300	900
max_depth	5	7	6	4
learning_rate	0.041	0.022	0.018	0.172
subsample	0.69	0.82	0.85	0.98
colsample_bytree	0.64	0.65	0.60	0.99

Using these settings, final models with the full training set 2009–2024 is built and tested on the dataset with data from 2025. The results are presented and analyzed in the next section.

7 Results

Model-specific prediction results on the testing set with data from 2025 are calculated. Due to different number of records and dataset imbalance, the generic model results are barely comparable. Therefore, in Table 5, aggregated daily predictions for joint Kurzeme and Zemgale were evaluated. They show how well each model, trained at different spatial scales, predicts total damage count per day.

Table 5. Total daily damage prediction results

Model Metric	State regions (2 cells)	Maintenance regions (6 cells)	County regions (11 cells)	Weather observation regions (467 cells)
MAE	4.47	4.09	4.14	11.46
RMSE	5.72	5.36	5.47	13.80
R2	0.47	0.53	0.51	-2.09
PICP	0.95	1.00	1.00	1.00
MPIW	25.32	35.38	44.09	90.10
Quantile Loss (Min)	0.42	0.55	0.57	0.57
Quantile Loss (Prediction)	2.24	2.05	2.07	5.73
Quantile Loss (Max)	0.96	1.21	1.63	3.93

The best-performing model is the 6-cell model with respective MAE 4.09, RMSE 5.36, R2 0.53, and quantile loss for prediction 2.05 values. The 6-cell, 11-cell, and 467-cell models predict lower and upper bounds of damage per day such that the actual damage count falls within these limits, as shown by the PICP. However, this was performed at the expense of a wider corridor (min and max predictions). The 2-cell model shows better results for the lower and upper bounds, respectively,

[‡] During the Optuna search and training, all models were built on NVIDIA RTX PRO 6000 Blackwell Server Edition

having the best MPIW value of 25.32 and better quantile loss results of 0.42 for minimum and 0.96 for maximum damage count prediction per day.

Each model's aggregated predictions together with the minimum and maximum bounds and actual damage value as shown in Figure 5. Figure 5 shows that the 11-cell and 467-cell models always predict 0 for the minimum bound. The 467-cell model also almost always predicts 0 for base prediction. This can be explained by the imbalance of the dataset, namely, the high number of records of cases without damages. The 2-cell, 6-cell, and 11-cell models overpredict days with small damage counts, but pivotal is that they have prediction and upper bound peaks for increased damage count days.

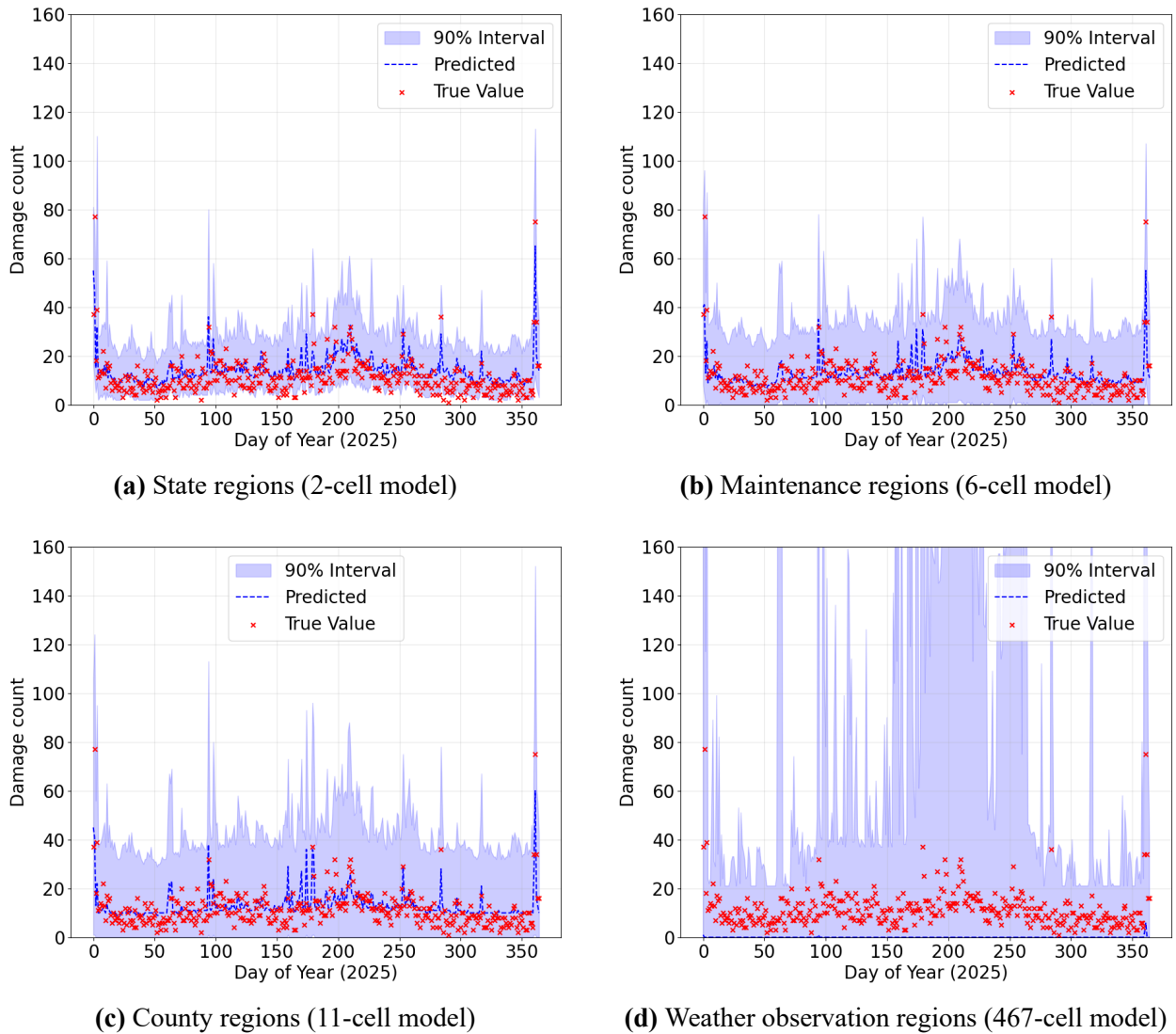


Figure 5. Quantile regression results on aggregated daily predictions

For a detailed analysis, days with either predicted value (with any of the four models) or the actual value (30 or more damages per day) are presented in Table 6. Evaluating individual predictions and comparing model performance, it is visible that 2-cell and 6-cell models strongly outperform 11-cell and 467-cell models. These results also confirm that the 2-cell model more accurately determines the minimum and maximum prediction bounds. Somewhat surprisingly, the 2-cell model shows even better baseline prediction results than the 6-cell model. This is explained by the fact that the 6-cell model results are more accurate at a lower number of damages per day.

Analysis of SHAP explanations shows (Figure 6) that models predict more damages in regions with older poles, higher soil temperatures, and more overhead lines in populated areas. Similarly, damage predictions increase during periods with more hours under Level 1 weather warnings, as

well as when sun elevation, temperature, boundary layer height, precipitation, and gusty wind speeds (especially above 15 meters per second) are higher. These observations coincide with hypotheses put forward by experts: older poles tend to break more frequently, the amount of damage increases during warmer weather conditions, and clearing solitary trees near power lines is often prohibited in populated areas. On the other hand, strong winds act as the primary cause of falling trees and their branches. Combining SHAP explanations with expert insights increases the reliability of the models.

Table 6. Analysis of days with ≥ 30 actual or predicted damages

		2-cell			6-cell			11-cell			467-cell		
Date	Actual	Min	Pred	Max	Min	Pred	Max	Min	Pred	Max	Min	Pred	Max
2025-01-01	37	16	55	81	8	40	82	1	45	106	0	1	363
2025-01-02	77	11	46	74	4	41	96	1	41	124	0	0	432
2025-01-04	39	9	32	110	2	26	87	0	22	95	0	0	268
2025-04-05	32	14	36	80	6	35	78	1	38	113	0	1	399
2025-06-24	14	15	29	49	5	31	68	1	36	93	0	0	375
2025-06-29	37	12	22	64	4	31	77	1	36	96	0	0	366
2025-06-30	25	10	21	54	3	26	65	1	34	86	0	0	358
2025-07-17	32	9	25	44	1	19	52	0	18	66	0	0	206
2025-07-30	32	12	33	54	2	27	60	0	26	77	0	0	215
2025-09-11	29	11	31	49	3	28	56	0	29	75	0	0	351
2025-10-12	36	11	29	49	3	27	60	0	28	78	0	0	289
2025-12-27	34	9	32	84	5	28	62	1	29	83	0	2	309
2025-12-28	75	12	65	113	11	55	107	1	60	152	0	6	442
2025-12-29	34	7	20	47	3	18	51	0	16	61	0	0	164

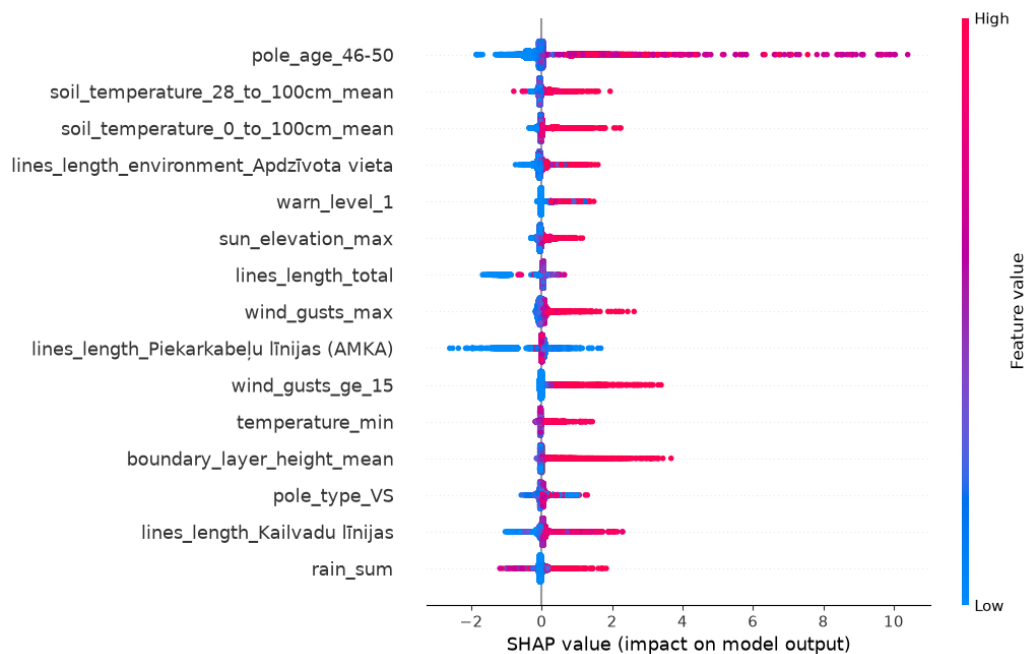


Figure 6. SHAP explanations for 6-cell model predictions

8 Conclusion

When the damage count is relatively small, the predicted value may vary without serious consequences. Days with high damage counts are of utmost interest for grid maintenance operation planning. The modeling results and prediction analysis showed satisfactory forecasting results if the final forecast is evaluated together with the minimum and maximum prediction bounds. The 2-cell model, while having lower overall prediction results, showed the best results in predicting high damage count events (30 or more) along with more precise minimum and maximum bounds. Therefore, the 2-cell model is favored over other models for service crew resource planning.

A limitation of this study is the use of actual historical weather data, as the weather forecasts issued a day before the damage occurred were not available. It should be noted that weather forecasting models are constantly evolving, and historical forecasts produced by models used ten years ago would therefore be of limited value. Instead, forecasts for previous years generated with the currently used weather prediction models would be needed, as these are the models applied in operational planning today and in the near future.

In future work, additional infrastructure data, forest vegetation models, and wetland maps could be tested for model training, especially if accurate historical information about the conditions on the day before the damage was available. Other machine learning models should also be tested, experimenting with different grid sizes, quantile thresholds, and broader geographical regions within Latvia and neighboring countries.

Methodologically, while our cross-validation strategy was chosen to preserve temporal integrity, preventing data leakage by ensuring models are never trained on future data to predict the past, it assumes a degree of data stability. Future work should investigate a rolling-horizon approach to better capture finer-grained seasonality and address potential distribution shifts over extended timelines. Regarding model scope and benchmarking, our hyperparameter tuning relied primarily on quantile error to align directly with our quantile interval goals. Although comparative evaluations against alternative forecasting frameworks (such as SARIMAX) and uncertainty quantification techniques (such as conformal prediction) offer valuable complementary perspectives, they were beyond the scope of this study and its available technical resources. While quantile-based metrics effectively guide model optimization, supplementing model results with the continuous ranked probability score (CRPS) would offer an even more comprehensive summary of probabilistic forecast performance.

Finally, interesting work could focus on the analysis and comparison of SHAP explanations between minimum, predicted, and maximum value explanations. In the field of user interaction and information presentation, it would also be useful to experiment with simultaneously displaying the damage predictions in grid cells on a map, along with the minimum and maximum limits, in a user-friendly way.

Funding

This work was funded by the European Union's NextGenerationEU instrument through the Latvia's Recovery and Resilience Plan, project "Development of a new big data-driven surface power grid damage prediction system (ST ODA) for more efficient prevention of power supply disruptions."

References

- [1] T. Zhao, E. Satoshi, M. Yue, M. Jensen, A. Marschilok, B. Nugent, B. Cerruti, and C. Spanos, "Machine Learning-based Dynamic Granular Electric Outage Forecasting," in *2023 Resilience Week (RWS)*, 2023, pp. 1–5. Available: <https://doi.org/10.1109/RWS58133.2023.10284644>
- [2] J. Li, Y. Gao, A. Lei, Y. Mei, W. Tao, L. Wang, T. Li, Y. Wang, and J. Liu, "Prediction of Power Failure Caused by Extreme Weather Based on Spatiotemporal Data," in *2024 4th International*

- Conference on New Energy and Power Engineering (ICNEPE)*, 2024, pp. 777–780. Available: <https://doi.org/10.1109/ICNEPE64067.2024.10860621>
- [3] L. Peter, K. Knezovic, M. Jawad, N. Katiyar, M. Subasic, and A. Shirsat, “Outage Prediction for Extreme Weather Events Based on Machine Learning,” in *CIREC Chicago Workshop 2024: Resilience of Electric Distribution Systems*, vol. 2024, 2025, pp. 305–309. Available: <https://doi.org/10.1049/icp.2024.2621>
 - [4] L. Prieto-Godino, C. Peláez-Rodríguez, J. Pérez-Aracil, J. Pastor-Soriano, and S. Salcedo-Sanz, “Predicting Weather-Related Power Outages in Large Scale Distribution Grids with Deep Learning Ensembles,” *International Journal of Electrical Power & Energy Systems*, vol. 170, article 110811, 2025. Available: <https://doi.org/10.1016/j.ijepes.2025.110811>
 - [5] S. Zhong, X. Qin, J. Li, J. Yao, J. Li, and Z. Lin, “Research on Intelligent Fault Identification and Power Outage Prediction Algorithm of Power Grid under Extreme Weather Based on Deep Learning,” in *2025 International Conference on Smart & Sustainable Technology (INCSST)*, 2025, pp. 1–7. Available: <https://doi.org/10.1109/INCSST64791.2025.11210386>
 - [6] A. Stasko, E. Otomers, A. Bondarenko, R. Sirmais, K. K. Vēliņš, D. Bļizņuks, and I. Birzniece, “Baseline Study of Machine Learning Applications for Power Grid Damage Prediction in Latvia,” *Automatic Control and Computer Sciences*, vol. 60, no. 3, pp. 425–438, 2026. Available: <https://doi.org/10.3103/S0146411626700549>
 - [7] R. Koenker, *Quantile Regression*, ser. Econometric Society Monographs. Cambridge University Press, 2005. Available: <https://doi.org/10.1017/CBO9780511754098>
 - [8] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design Science in Information Systems Research,” *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004. Available: <https://doi.org/10.2307/25148625>
 - [9] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems research,” *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007.
 - [10] R. Tervo, J. Karjalainen, and A. Jung, “Short-term prediction of electricity outages caused by convective storms,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8618–8626, 2019. Available: <https://doi.org/10.1109/TGRS.2019.2921809>
 - [11] E. S. Kiel and G. H. Kjølle, “A Bayesian Network Approach to Predicting Transmission Line Down Times,” in *Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*, 2020, pp. 3725–3732. Available: https://doi.org/10.3850/978-981-14-8593-0_3776-cd
 - [12] C. Shearer, “The CRISP-DM Model: the New Blueprint for Data Mining,” *Journal of Data Warehousing*, vol. 5, no. 4, pp. 13–22, 2000.
 - [13] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 4768–4777.
 - [14] V. Plevris, G. Solorzano, N. P. Bakas, M. A. El, and A. B. Seghier, “Investigation of Performance Metrics in Regression Analysis and Machine Learning-Based Prediction Models,” in *8th European Congress on Computational Methods in Applied Sciences and Engineering*, 2022. Available: <https://doi.org/10.23967/eccomas.2022.155>
 - [15] D. L. Shrestha and D. P. Solomatine, “Machine Learning Approaches for Estimation of Prediction Interval for the Model Output,” *Neural Networks*, vol. 19, no. 2, pp. 225–235, 2006. Available: <https://doi.org/10.1016/j.neunet.2006.01.012>
 - [16] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, “Lower Upper Bound Estimation Method for Construction of Neural Network-Based Prediction Intervals,” *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 337–346, 2011. Available: <https://doi.org/10.1109/TNN.2010.2096824>
 - [17] R. Koenker and G. Bassett, “Regression Quantiles,” *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978. Available: <http://www.jstor.org/stable/1913643>
 - [18] I. Takeuchi, Q. V. Le, T. D. Sears, and A. J. Smola, “Nonparametric Quantile Estimation,” *Journal of Machine Learning Research*, vol. 7, no. 45, pp. 1231–1264, 2006.
 - [19] G. Kalvane and A. Kalvans, “Seasonality in the Landscape of Latvia: a Phenological Calendar,” *Folia Geographica*, vol. 18, pp. 95–103, 2020. Available: <https://doi.org/10.22364/fg.18.11>
 - [20] M. R. Raupach, “Simplified Expressions for Vegetation Roughness Length and Zero-Plane Displacement as Functions of Canopy Height and Area Index,” *Boundary-Layer Meteorology*, vol. 71, pp. 211–216, 1994. Available: <https://doi.org/10.1007/BF00709229>

- [21] P. Zippenfenig, “Open-Meteo.com Weather API,” 2023, software license: CC-BY-4.0. Available: <https://github.com/open-meteo/open-meteo>
- [22] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794. Available: <https://doi.org/10.1145/2939672.2939785>
- [23] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 2623–2631. Available: <https://doi.org/10.1145/3292500.3330701>

Appendix A Aggregated Attributes

Table A1. Aggregated attribute list

Aggregated attribute (original label)	Source attribute	Aggregation
elevation_mean	elevation	mean()
distance_to_sea_min	distance_to_sea	min()
distance_to_sea_mean	distance_to_sea	mean()
leaf_factor	leaf_factor	value()
sun_elevation_max	sun_elevation	max()
lines_length_Posma garums, m	lines	sum(length)
lines_length_Kailvadu līnijas	lines[type=Kailvadu līnijas]	sum(length)
lines_length_Līnijas ar izolētiem vadiem	lines[type=Līnijas ar izolētiem vadiem]	sum(length)
lines_length_Piekarkabeļu līnijas	lines[type=Piekarkabeļu līnijas]	sum(length)
lines_length_populated_Y	lines[populated_area=Yes]	sum(length)
lines_length_populated_N	lines[populated_area=No]	sum(length)
lines_length_forest_Jā	lines[forest=Yes]	sum(length)
lines_length_forest_Nē	lines[forest=No]	sum(length)
lines_length_environment_Apdzīvota vieta	lines[environment=Apdzīvota vieta]	sum(length)
lines_length_environment_Augļu dārzs	lines[environment=Augļu dārzs]	sum(length)
lines_length_environment_Krūmājs	lines[environment=Krūmājs]	sum(length)
lines_length_environment_Lauksaimniecībā izmantojama zeme	lines[environment=Lauksaimniecībā izmantojama zeme]	sum(length)
lines_length_environment_Meža zeme	lines[environment=Meža zeme]	sum(length)
lines_length_environment_Purvs	lines[environment=Purvs]	sum(length)
lines_length_environment_Pļava	lines[environment=Pļava]	sum(length)
lines_length_environment_Ražošanas teritorija	lines[environment=Ražošanas teritorija]	sum(length)
lines_length_environment_Upe/Grāvis	lines[environment=Upe/Grāvis]	sum(length)
lines_length_environment_Ceļš	lines[environment=Ceļš]	sum(length)
lines_length_environment_Dzelzceļš	lines[environment=Dzelzceļš]	sum(length)
pole_types_Pi	poles[type=Pi]	count()
pole_types_St	poles[type=St]	count()
pole_types_ZS	poles[type=ZS]	count()
pole_types_VS	poles[type=VS]	count()
pole_types_Koka	poles[material=Koka]	count()
pole_types_Metāla	poles[material=Metāla]	count()
pole_types_Dzelzsbetona	poles[material=Dzelzsbetona]	count()
pole_types_Pārējie	poles[material=Pārējie]	count()
pole_age_1-5	poles[age=1..5]	count()
pole_age_6-10	poles[age=6..10]	count()
pole_age_11-15	poles[age=11..15]	count()
pole_age_16-20	poles[age=16..20]	count()
pole_age_21-25	poles[age=21..25]	count()
pole_age_26-30	poles[age=26..30]	count()
pole_age_31-35	poles[age=31..35]	count()
pole_age_36-40	poles[age=36..40]	count()
pole_age_41-45	poles[age=41..45]	count()
pole_age_46-50	poles[age=46..50]	count()
pole_age_51-55	poles[age=51..55]	count()
pole_age_56-60	poles[age=56..60]	count()
temperature_min	temperature	min()

Continued on next page

Table A1. continued

Aggregated attribute (original label)	Source attribute	Rules
temperature_max	temperature	max()
temperature_mean	temperature	mean()
temperature_std	temperature	std()
relative_humidity_min	relative_humidity	min()
relative_humidity_max	relative_humidity	max()
relative_humidity_mean	relative_humidity	mean()
relative_humidity_std	relative_humidity	std()
dew_point_min	dew_point	min()
dew_point_max	dew_point	max()
dew_point_mean	dew_point	mean()
dew_point_std	dew_point	std()
precipitation_sum	precipitation	sum()
rain_sum	rain	sum()
snowfall_sum	snowfall	sum()
snow_depth_min	snow_depth	min()
snow_depth_max	snow_depth	max()
snow_depth_mean	snow_depth	mean()
surface_pressure_min	surface_pressure	min()
surface_pressure_max	surface_pressure	max()
surface_pressure_mean	surface_pressure	mean()
surface_pressure_std	surface_pressure	std()
boundary_layer_height_min	boundary_layer_height	min()
boundary_layer_height_max	boundary_layer_height	max()
boundary_layer_height_mean	boundary_layer_height	mean()
boundary_layer_height_std	boundary_layer_height	std()
soil_moisture_0_to_100cm_mean	soil_moisture_0_to_100cm	mean()
soil_moisture_0_to_7cm_mean	soil_moisture_0_to_7cm	mean()
soil_moisture_28_to_100cm_mean	soil_moisture_28_to_100cm	mean()
soil_moisture_7_to_28cm_mean	soil_moisture_7_to_28cm	mean()
soil_temperature_0_to_100cm_mean	soil_temperature_0_to_100cm	mean()
soil_temperature_0_to_7cm_mean	soil_temperature_0_to_7cm	mean()
soil_temperature_28_to_100cm_mean	soil_temperature_28_to_100cm	mean()
soil_temperature_7_to_28cm_mean	soil_temperature_7_to_28cm	mean()
wind_direction_E	wind_direction[direction=E]	count()
wind_direction_N	wind_direction[direction=N]	count()
wind_direction_NE	wind_direction[direction=NE]	count()
wind_direction_NW	wind_direction[direction=NW]	count()
wind_direction_S	wind_direction[direction=S]	count()
wind_direction_SE	wind_direction[direction=SE]	count()
wind_direction_SW	wind_direction[direction=SW]	count()
wind_direction_W	wind_direction[direction=W]	count()
weather_code_0	weather_code[code=0]	count()
weather_code_1	weather_code[code=1]	count()
weather_code_2	weather_code[code=2]	count()
weather_code_3	weather_code[code=3]	count()
weather_code_51	weather_code[code=51]	count()
weather_code_53	weather_code[code=53]	count()
weather_code_55	weather_code[code=55]	count()
weather_code_61	weather_code[code=61]	count()
weather_code_63	weather_code[code=63]	count()

Continued on next page

Table A1. continued

Aggregated attribute (original label)	Source attribute	Rules
weather_code_65	weather_code[code=65]	count()
weather_code_71	weather_code[code=71]	count()
weather_code_73	weather_code[code=73]	count()
weather_code_75	weather_code[code=75]	count()
wind_speed_min	wind_speed	min()
wind_speed_max	wind_speed	max()
wind_speed_mean	wind_speed	mean()
wind_speed_std	wind_speed	std()
wind_speed_e_0	wind_speed[speed = 0]	count()
wind_speed_g_0	wind_speed[speed > 0]	count()
wind_speed_ge_5	wind_speed[speed >= 5]	count()
wind_speed_ge_10	wind_speed[speed >= 10]	count()
wind_speed_ge_15	wind_speed[speed >= 15]	count()
wind_speed_ge_16	wind_speed[speed >= 16]	count()
wind_speed_ge_17	wind_speed[speed >= 17]	count()
wind_speed_ge_18	wind_speed[speed >= 18]	count()
wind_speed_ge_19	wind_speed[speed >= 19]	count()
wind_speed_ge_20	wind_speed[speed >= 20]	count()
wind_speed_ge_21	wind_speed[speed >= 21]	count()
wind_speed_ge_22	wind_speed[speed >= 22]	count()
wind_speed_ge_23	wind_speed[speed >= 23]	count()
wind_speed_ge_24	wind_speed[speed >= 24]	count()
wind_speed_ge_25	wind_speed[speed >= 25]	count()
wind_gusts_min	wind_gusts	min()
wind_gusts_max	wind_gusts	max()
wind_gusts_mean	wind_gusts	mean()
wind_gusts_std	wind_gusts	std()
wind_gusts_e_0	wind_gusts[speed = 0]	count()
wind_gusts_g_0	wind_gusts[speed > 0]	count()
wind_gusts_ge_5	wind_gusts[speed >= 5]	count()
wind_gusts_ge_10	wind_gusts[speed >= 10]	count()
wind_gusts_ge_15	wind_gusts[speed >= 15]	count()
wind_gusts_ge_16	wind_gusts[speed >= 16]	count()
wind_gusts_ge_17	wind_gusts[speed >= 17]	count()
wind_gusts_ge_18	wind_gusts[speed >= 18]	count()
wind_gusts_ge_19	wind_gusts[speed >= 19]	count()
wind_gusts_ge_20	wind_gusts[speed >= 20]	count()
wind_gusts_ge_21	wind_gusts[speed >= 21]	count()
wind_gusts_ge_22	wind_gusts[speed >= 22]	count()
wind_gusts_ge_23	wind_gusts[speed >= 23]	count()
wind_gusts_ge_24	wind_gusts[speed >= 24]	count()
wind_gusts_ge_25	wind_gusts[speed >= 25]	count()
warn_level_0	warn_level[level=normal](wind_gusts, precipitation)	count()
warn_level_1	warn_level[level=yellow](wind_gusts, precipitation)	count()
warn_level_2	warn_level[level=orange](wind_gusts, precipitation)	count()
warn_level_3	warn_level[level=red](wind_gusts, precipitation)	count()