

Assessing the potential of complex artificial neural networks for modelling small-scale soil erosion by water

Nils Barthel¹, Simone Ott¹, Benjamin Burkhard¹, and Bastian Steinhoff-Knopp²

¹Leibniz University Hannover, Institute of Earth System Sciences, Physical Geography and Landscape Ecology Section, Schneiderberg 50, 30167, Hannover, Germany

²Thünen-Institute, Coordination Unit Climate Soil Biodiversity, Bundesallee 49, 38116, Braunschweig, Germany

Correspondence: Nils Barthel (barthel@phygeo.uni-hannover.de)

Abstract. Modelling soil erosion by water is essential for developing effective mitigation strategies and preventing on- and off-site damages in agricultural areas. So far, complex artificial neural networks have rarely been applied in small-scale soil erosion modelling, and their potential still remains unclear. This study compares the performance of different neural network architectures for modelling soil erosion by water at a small spatial scale in agricultural cropland. The analysis is based on erosion rate data at a 5 m × 5 m resolution, derived from a 20-year monitoring programme, and covers 458 hectares of cropland across seven investigation areas in northern Germany. Nineteen predictor variables related to topography, climate, management, and soil properties were selected as inputs to assess their interrelationships with observed erosion patterns. A single-layer neural network (SNN), a deep neural network (DNN), and a convolutional neural network (CNN) were applied and evaluated against a random forest (RF) model used as a benchmark. A leave-one-area-out validation is applied to evaluate how well the models generalize to areas withheld entirely during training. While all models tended to underestimate high erosion rates, they often successfully captured the underlying spatial patterns. The CNN and DNN achieved the highest predictive performance, with the CNN producing the lowest root mean squared error (RMSE: 2.187) and the highest F1 score (0.373), and the DNN achieving the lowest mean absolute error (MAE: 0.856) and performing largely comparably to the CNN. This study demonstrates the potential of complex neural networks to capture erosion patterns at the field-to-landscape scale and provides insights into the relevance of the chosen predictor variables, as well as key modelling limitations, such as the underestimation of very high erosion rates in unseen areas. It also highlights the need for more comprehensive datasets to improve generalization capabilities of the models.

1 Introduction

Soil erosion by water causes the loss of topsoil, along with the displacement of organic carbon and nutrients, degrading agricultural land and contributing to off-site damages such as water eutrophication and road blockages (Issaka and Ashraf, 2017). These processes have significant and long-lasting environmental and economic consequences in affected areas worldwide. A combination of anthropogenic, climatic, and topographic factors unique to each region defines the extent and distribution of soil erosion by water. Modelling the interactions among these factors at fine spatial scales can help reveal key drivers and the

resulting erosion patterns. The resulting insights can support the development of targeted strategies to mitigate further soil degradation (Borrelli et al., 2018; Igwe et al., 2017).

Various approaches have been applied to estimate soil erosion by water across different spatial scales, ranging from global (Guerra et al., 2020), continental (Panagos et al., 2021), national (Plambeck, 2020), to small field plots (Anache et al., 2018). The respective modelling approaches vary in data requirements, input variables, underlying principles, and complexity. Model outputs are typically classified into erosion severity, susceptibility categories or expressed as continuous predictions of soil erosion rates over space and time. Depending on the intended output and available data, a variety of model types have been developed, including process-based, empirical, and, more recently, machine learning models (Borrelli et al., 2021).

Examples of physical process-based models are WEPP (Nearing et al., 1989; Pieri et al., 2007), Erosion 3D (Schmidt et al., 1999), and EuroSEM (Morgan et al., 1998). Empirical models include the Universal Soil Loss Equation (USLE; Wischmeier and Smith 1978), its revised version (RUSLE; Renard 1997), and further adaptations (Borrelli et al., 2021). One reason for the popularity of the USLE and its variations is their relatively low complexity, making these models time- and cost-efficient (Alewell et al., 2019; Avand et al., 2023; Kumar et al., 2022). The apparent simplicity of the USLE contrasts with the complexity of calculating its individual factors, which results in a wide range of individual USLE-based applications and estimated loss rates due to variations in input data and methods (Fiener et al., 2020).

Machine learning approaches have been increasingly used to identify areas at risk of soil erosion by water and to capture the complex relationships between environmental predictors and observed erosion patterns. This includes the application of methods such as random forests (RF; Garosi et al. 2019; Ghosh and Maiti 2021; Jaafari et al. 2022), with several studies reporting that RF can outperform methods such as support vector machines (SVMs) and generalized additive models (GAMs). While these studies highlight the potential of RF models, their main focus has been on predicting and classifying gully-erosion susceptibility.

Another promising machine learning approach for soil erosion modelling is the application of artificial neural networks (ANNs). However, existing studies in this field have often employed simple neural network architectures, typically with only a single hidden layer. These single-layer networks have not only been used to model soil erosion susceptibility (De la Rosa et al., 1999) but also to predict quantitative soil erosion rates at the plot scale (Licznar and Nearing, 2003). These early applications demonstrate the potential of ANNs, but single-layer networks may not fully utilize the capacity of more complex architectures to capture non-linear relationships between influencing variables and the spatial distribution and magnitude of soil erosion (Avand et al., 2023).

More complex, multi-layered ANNs have rarely been applied to predict soil erosion rates quantitatively. Instead, they have been mostly used to model a classified erosion susceptibility (Golkarian et al., 2023; Khosravi et al., 2023; Sarkar and Mishra, 2018) or have been restricted to gully erosion (Ghorbanzadeh et al., 2020; Saha et al., 2021). In cases where ANNs have been applied to quantify continuous soil erosion rates, studies often rely on limited datasets, such as those collected using erosion pins over just one year (Gholami et al., 2021; Sahour et al., 2021).

While machine learning methods have shown promise in modelling soil erosion, their application to predict continuous erosion rates at fine spatial scales remains limited, particularly when using neural networks. Currently, most studies rely on

input data with spatial resolutions between 5 and 300 metres, with higher-resolution datasets often restricted to small areas, such as individual plots (Borrelli et al., 2021; Parsons, 2019). This lack of research combining neural networks with high-resolution data is partly due to the limited availability of long-term monitoring schemes at the landscape scale, which provides spatially explicit erosion data as ground truth for model training and validation (Batista et al., 2025).

To address this gap, this study utilized long-term soil-erosion monitoring data spanning more than two decades and collected across seven study areas in northern Germany (Steinhoff-Knopp and Burkhard, 2018). This monitoring dataset, together with high-spatial-resolution predictor variables, served as input for different machine learning models. Specifically, three neural networks of increasing complexity and a random forest, used as a benchmark model, were systematically compared to evaluate their ability to predict continuous soil-erosion rates. To assess model robustness and generalizability, we applied a leave-one-area-out validation strategy, in which each study area was withheld in turn for testing while the models were trained on the remaining areas.

The goal of this study was to contribute to a more comprehensive understanding of the strengths and limitations of different machine learning approaches for soil erosion modelling by aiming to answer the following research questions:

1. Which machine learning approach achieves the highest predictive performance in estimating spatial patterns of continuous soil erosion rates at the field-to-landscape scale?
2. What are the challenges and limitations when reproducing erosion patterns within the same areas and when extrapolating to previously unseen areas?
3. Which predictor variables are most important for predicting soil erosion by water at the field-to-landscape scale across different machine learning models?

2 Methods

2.1 Study area

This study uses data from seven distinct study areas in Lower Saxony, Germany, located between 53.1° N and 51.9° N latitude and 8.3° E and 10.5° E longitude. Five of these regions are clustered in the southeastern part of Lower Saxony, and named *Leine-Innerstebergland*, while the other two are located in the Northeast and Southwest (Fig. 1). The cultivated cropland within all the areas covers 458 hectares, is prone to erosion, and represents different soil types, relief characteristics, and management conditions (Capelle and Lüders, 1985; Capelle, 1990).

A large part of the soils – predominantly Luvisols and Cambisols – has a high silt content due to loess or sand-loess deposits, increasing the erodibility of the topsoil (Steinhoff-Knopp and Burkhard, 2018). The mean slope across all study areas is 4.01°, with the southern regions having a steeper relief (4.81°) compared to the northern (2.35°) and the western region (3.68°). The primary crop is winter wheat, followed by winter barley, rapeseed, sugar beet, maize, and potatoes. All farmers within the monitored areas practice conventional farming and implement various soil conservation practices, including reduced tillage, contour-parallel tillage, cover crops, grassed tramlines, and drainage systems.

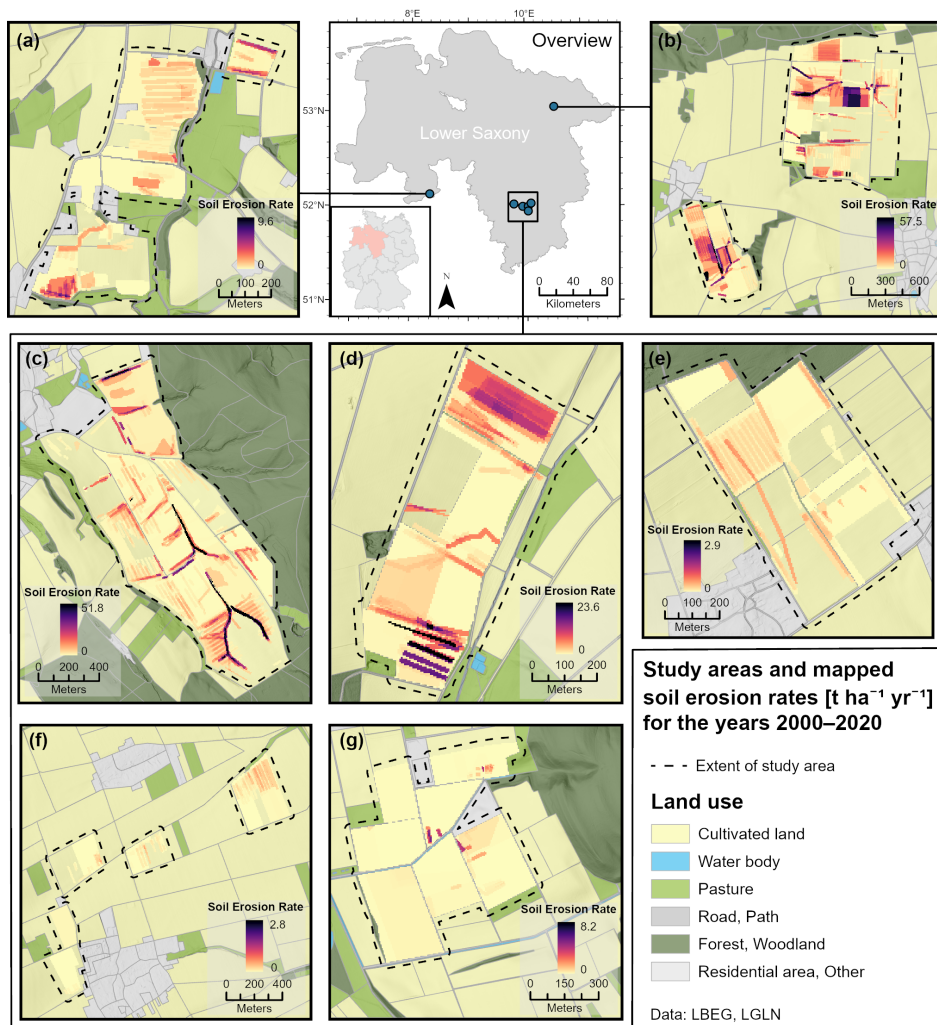


Figure 1. Overview of study areas in Lower Saxony and mapped soil erosion rates in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brüggen.

2.2 Data collection

The soil erosion dataset used in this study was derived from a long-term monitoring programme funded by the Lower Saxony State Authority for Mining, Energy and Geology (LBEG), covering the years 2000 to 2020. Exceptions include the investigation areas Adenstedt, where monitoring began in 2002, and Klein Ilde, where monitoring stopped in 2015. The methodology of the field surveys is based on the recommendations by Rohr et al. (1990) and the instructions by DVWK (1996) and Botschek et al. (2021). To implement a more efficient workflow, the mobile mapping application *EroPad* was used since 2010 (see Steinhoff et al. 2013). The field surveys were carried out each year after the snowmelt (usually in February or March) and usually within

one week after each erosive rainfall event ($> 12.7 \text{ mm h}^{-1}$; see DIN 19708:2022-08) throughout the summer months. Surveys were also conducted when farmers reported erosion features.

100 The field surveys aimed to collect data on the spatial distribution and extent of three main types of soil erosion by water: linear (i.e. rill) erosion, sheet erosion, and sheet-to-linear erosion (erosion systems showing both features of sheet and linear erosion). The losses by linear erosion features were surveyed by measuring their volume and extrapolated to an 8 m buffer around their occurrence to account for the dispersion of soil loss due to tillage practices by farmers. Deposition sites were also recorded, but only qualitatively, and were therefore not included in the modelling. A detailed description and evaluation of the
 105 survey dataset is provided in Steinhoff-Knopp and Burkhard (2018). For the modelling process, soil erosion rates ($\text{t ha}^{-1}\text{yr}^{-1}$) were calculated by overlaying all surveyed erosion features in a grid resolution of $5 \text{ m} \times 5 \text{ m}$ (Fig. 1).

Table 1. Overview of the variables used in the soil erosion modelling process. A detailed description, including processing steps, can be found in Appendix A.

Variable	Description	Unit
Altitude	Based on digital elevation model	m
Slope	Slope steepness in the direction of maximum elevation change	degree (°)
Slope length	Accumulated downslope flowpath length	m
Aspect 360	Direction of maximum elevation change [0 - 360°]	degree (°)
Aspect 180	Normalized aspect direction [0 - 180°]	degree (°)
Plan curvature	Curvature perpendicular to the direction of maximum slope	-
Profile curvature	Curvature in the direction of maximum slope	-
Flow line curvature	Curvature along flow lines	-
Topographic position index	Relative altitude compared to surrounding terrain	-
Flow accumulation	Upslope contributing area	-
Wetness index	SAGA wetness index	-
Divergence-convergence index (DCI)	Index of water-flow divergence or convergence (3×3 grid-cell area)	-
DCI 10	Divergence–convergence index based on a 10×10 grid-cell area	-
Machining direction (MD)	Orientation of field tramlines [0 - 180°]	degree (°)
MD vs. aspect	Angle between MD and Aspect 180 [0 - 90°]	degree (°)
R factor	USLE rainfall erosivity factor	$\text{MJ mm (h ha yr)}^{-1}$
K factor	USLE soil erodibility factor	t h (MJ mm)^{-1}
LS factor	USLE topographic factor (slope length and steepness)	-
C factor	USLE crop and management factor	-

Nineteen predictor variables (i.e. features or covariates) potentially influencing soil erosion were selected for the modelling process (Table 1). These variables include factors related to topography, climate, soil properties, and agricultural land management. The topographic variables were derived from a digital elevation model (DEM) with a native resolution of $1 \text{ m} \times 1 \text{ m}$,

110 resampled to $5 \text{ m} \times 5 \text{ m}$, which serves as the basis for all derived topographic variables. The USLE R-factor (erosivity of rainfall), originally calculated at a grid resolution of $1 \text{ km} \times 1 \text{ km}$, was resampled to match the DEM-based variables. The USLE K-factor (erodibility of topsoil), derived from a soil map at a scale of 1:50 000 (LBEG, 2017), and the USLE C-factor (crop cover and management factor) were rasterized to the $5 \text{ m} \times 5 \text{ m}$ grid. The C-factor was determined for each field using agricultural land-management data collected during the field surveys and through interviews with the respective farmers.
115 Further information on the methods to determine each predictor variable is provided in Appendix A.

Several of the variables, especially the topographic variables, were derived from each other or capture related aspects, which could lead to a strong correlation between them (Avand et al., 2023; Jaafari et al., 2022). Therefore, the degree of the linear relationship between variables was assessed by using the Pearson correlation coefficient and the variance inflation factor (VIF). The Pearson correlation coefficient is used to gain information on the pairwise linear correlation between the variables, with
120 ± 0.7 being considered a threshold for strong correlation (Schober et al., 2018). The VIF is used to assess multicollinearity across all variables, evaluating whether variables with high pairwise correlations can still provide valuable information for the final predictions (Ebrahimi-Khusfi et al., 2021; O'Brien, 2007). Variables with a VIF below 5 are considered to have a low level of multicollinearity and are used for the modelling (Daoud, 2017).

2.3 Machine Learning Models

125 This section provides a brief introduction to the four distinct machine learning models employed in this study. The model implementations and hyperparameter selection process are described in Sect. 2.5.

2.3.1 Random forest

A random forest (RF) model combines multiple decision trees to create an ensemble model to improve prediction accuracy (Breiman, 2001). A decision tree splits data based on feature values across multiple levels of nodes, with each branch representing different decision paths leading to predictions (Kingsford and Salzberg, 2008). Each tree is trained on a random subset
130 of the variables, and, in the case of regression tasks, as applied in this study, the final prediction is obtained by averaging the outputs of all trees.

Although mostly limited to classification outputs, RF has shown promise in previous studies on soil erosion modelling (Garosi et al., 2019; Ghosh and Maiti, 2021; Jaafari et al., 2022). Therefore, in this study, RF serves as a benchmark for
135 evaluating the performance of neural networks against other machine learning models. The RF model used in this study consisted of 200 decision trees, with a maximum tree depth of 20, a minimum of five samples required to split an internal node, and at least two samples per leaf.

2.3.2 Single-hidden layer neural network

Artificial neural networks consist of an input layer, an output layer, and one or more so-called hidden layers of neurons (or
140 nodes) in between (Rumelhart et al., 1986). Each neuron in these layers is interconnected through weights. During training,

these weights are updated through an optimization process involving multiple iterations, a loss function, and, most commonly, backpropagation. The weighted sum of the inputs is passed through an activation function, which applies a non-linear transformation to determine the neuron's output, which is then passed on to the subsequent layer. In a fully connected (or dense) layer, each neuron is connected to every neuron in the subsequent layer through individual weights.

145 Three different types of neural networks were compared in this study, each with a distinct and increasingly complex architecture. The first type is a neural network with a single dense layer (SNN) of 128 neurons between the input and the output layer (see Wythoff 1993). The Rectified Linear Unit (ReLU) activation function was applied in both the hidden and output layers in this and the following two variations of neural network models, to introduce non-linear relationships and constrain predictions to positive values (see Krizhevsky et al. 2017; Nair and Hinton 2010). Additionally, L_2 regularization, also known as Ridge
150 regularization, was employed across all neural networks to reduce overfitting (see Hoerl and Kennard 1970; Ng 2004). The Adam optimizer was used to determine the learning rate during training (see Kingma and Ba 2014).

2.3.3 Deep neural network

The interconnected neurons and the corresponding weights enable neural networks to capture complex non-linear relationships. Increasing the number of hidden layers enhances this capability while increasing training time and the likelihood of overfitting
155 the model to the training data. Such a neural network with multiple hidden layers is commonly known as a deep neural network (DNN) or a deep learning model (LeCun et al., 2015). In this case, the DNN was modified to have three hidden layers, each with 256 neurons.

2.3.4 Convolutional neural network

Convolutional neural networks (CNNs) are specifically designed to capture spatial relationships in grid-like data structures,
160 such as images (LeCun et al., 2015). This is achieved through convolutional layers that receive input patches containing the neighbouring grid cells around each target pixel. These patches provide local spatial context, enabling the network to learn spatial patterns and relationships relevant for predicting soil erosion (Krizhevsky et al., 2017; LeCun et al., 2015). In this study, each input consisted of a 7×7 grid-cell patch. Within these patches, the convolutional layers extracted spatial features using 256 filters per layer. Two dense layers with 256 neurons each were added following the convolutional layers. The dense layers
165 utilize the learned spatial patterns from the convolutional layers to produce the final predicted output.

2.4 Validation metrics

The results were compared using different validation metrics. To determine the overall differences between the modelled and the mapped soil erosion rate, the root mean squared error (RMSE) and the mean absolute error (MAE) were used. The F1 score, which is the harmonic mean of precision and recall, was used to evaluate the ability of the models to predict within
170 different ranges of soil erosion severity (Chinchor, 1992). For this, the continuous predicted output was categorized into six discrete classes: no erosion ($0 \text{ t ha}^{-1} \text{ yr}^{-1}$), very low erosion (> 0 to $< 0.25 \text{ t ha}^{-1} \text{ yr}^{-1}$), low erosion (0.25 to $< 1 \text{ t ha}^{-1} \text{ yr}^{-1}$),

medium erosion (1 to $< 2 \text{ t ha}^{-1} \text{ yr}^{-1}$), high erosion (2 to $< 5 \text{ t ha}^{-1} \text{ yr}^{-1}$), and very high erosion ($\geq 5 \text{ t ha}^{-1} \text{ yr}^{-1}$). A weighted average F1 score was calculated to account for the imbalanced distribution of the mapped soil erosion dataset across the classes, with the majority (81.2 %) of the data falling below $1 \text{ t ha}^{-1} \text{ yr}^{-1}$. The resulting F1 score ranges from 0 to 1, with values closer to 1 indicating a better overall alignment between the prediction and the different classes (Taha and Hanbury, 2015).

The importance of the predictor variables was evaluated through permutation importance, which reflects the influence of each variable on the final model output. This is done by randomly permuting one variable at a time and measuring the resulting increase in the mean squared error. A larger increase indicates that the variable has a greater impact on the predictive performance of the model. The results are normalized and expressed as percentages (Altmann et al., 2010).

180 2.5 Model implementation

To evaluate model predictions, we employed a leave-one-area-out validation strategy, in which one study area was excluded entirely during model training and subsequently used for validation. In each iteration, the models were trained on the remaining six areas and then used to predict soil erosion rates in the left-out area. Afterwards, the predictions were validated against the mapped erosion rates of that area. This procedure was repeated for all seven study areas, and validation metrics were calculated as weighted averages based on the relative size of the excluded area. Permutation importance values were also averaged across all seven runs. Using this validation approach enabled a comprehensive assessment of model robustness by evaluating their capacity to generalize to unseen areas with varying landscape conditions and management practices.

The predictive performance of each model depends strongly on its hyperparameters, which determine the model architecture and learning behaviour. The most important hyperparameters (e.g. number of trees or neurons) and their optimal values were described along with the respective models above (Sect. 2.3). To determine the optimal hyperparameters, we performed hyperparameter optimization prior to the model runs using a grid search (Raschka, 2020; Yu and Zhu, 2020). The grid search evaluated a predefined range of parameter settings and selected the configuration that yielded the best average performance across the seven leave-one-area-out folds. This optimal set of hyperparameters was subsequently used to train the final models. The full search space and the respective tuning scripts are provided in the accompanying repository.

195 All predictor variables were standardized and the soil-loss target was log-transformed to reduce skewness and stabilize model training. Predicted values were back-transformed to the original scale for evaluation and visualization. All models were built using the scikit-learn (Pedregosa et al., 2011) and TensorFlow (Abadi et al., 2015) Python packages.

3 Results

3.1 Pairwise correlation and multicollinearity analysis

200 To assess relationships between predictor variables and to identify potential redundancy due to multicollinearity, the Pearson correlation coefficient (r) and the variance inflation factor (VIF) were calculated. Five pairs of variables (see Fig. B1) had a strong linear correlation with Pearson coefficient close to or exceeding ± 0.7 : aspect 360 and slope length ($r = 0.69$), slope and

machining direction ($r = -0.69$), altitude and R factor ($r = -0.75$), topographic position index and divergence-convergence index 10 ($r = 0.82$), divergence-convergence index and wetness index ($r = 0.82$). However, all variables had a VIF < 5 , indicating that there is not a high level of multicollinearity between the variables (see Fig. B2). Consequently, no variables were removed for the modelling process.

3.2 Model performance

Overall, the CNN and DNN produced nearly identical error metrics when predicting soil erosion rates in previously unseen areas (Table 2). RMSE values were very similar across all models, ranging from 2.187 for the CNN to 2.221 for the SNN. The MAE showed more noticeable differences between models, with the DNN producing the lowest absolute error (0.856), followed closely by the CNN (0.859), whereas the RF had the highest value (0.971).

Table 2. Weighted evaluation metrics comparing random forest (RF), single-layer neural network (SNN), deep neural network (DNN), and convolutional neural network (CNN) models.

Metric	RF	SNN	DNN	CNN
RMSE	2.197	2.221	2.192	2.187
MAE	0.971	0.899	0.856	0.859

Table 3. F1 score evaluating the performance of the continuous model predictions (in $t\ ha^{-1}\ yr^{-1}$) within the defined soil-erosion classes, comparing random forest (RF), single-layer neural network (SNN), deep neural network (DNN) and convolutional neural network (CNN) models. The weighted average F1 scores were computed using the number of mapped grid-cells (n cells) within each class.

Soil erosion rate	n cells	F1 score			
		RF	SNN	DNN	CNN
0	655	0.000	0.028	0.010	0.023
< 0.25	110614	0.124	0.401	0.464	0.488
$0.25 - < 1$	36726	0.296	0.311	0.295	0.332
$1 - < 2$	13607	0.067	0.035	0.067	0.122
$2 - < 5$	14274	0.003	0.015	0.010	0.024
≥ 5	7073	0.000	0.000	0.000	0.000
Weighted average		0.138	0.317	0.345	0.373

The F1 scores shown in Table 3 further illustrate differences in model performance for low soil-erosion rates by measuring how well the continuous predictions align with the defined severity classes. Although F1 scores were generally low, the CNN achieved the highest weighted average value (0.373), followed by the DNN (0.345) and SNN (0.317), whereas the RF performed substantially worse in this regard (0.138). The difference between the RF and the neural networks is largely linked

to the $<0.25 \text{ t ha}^{-1} \text{ yr}^{-1}$ class, which contained the majority of mapped grid cells. Here, the CNN, DNN, and SNN achieved relatively high F1 scores of 0.488, 0.464, and 0.401, respectively, compared with 0.124 for the RF. For erosion rates equal and above $5 \text{ t ha}^{-1} \text{ yr}^{-1}$, all models had an F1 score of 0.

Across all study areas, the maximum predicted soil erosion rates ranged from $2.7 \text{ t ha}^{-1} \text{ yr}^{-1}$ predicted by the RF to $20.6 \text{ t ha}^{-1} \text{ yr}^{-1}$ predicted by the DNN. The soil-erosion predictions of the DNN, which had the lowest mean absolute error among the models, are shown in Fig. 2. The corresponding prediction maps for all models are provided in the Appendix (Fig. C1–C8). The predictions show that all models were partially able to predict the spatial distribution of erosion features in the unseen areas.

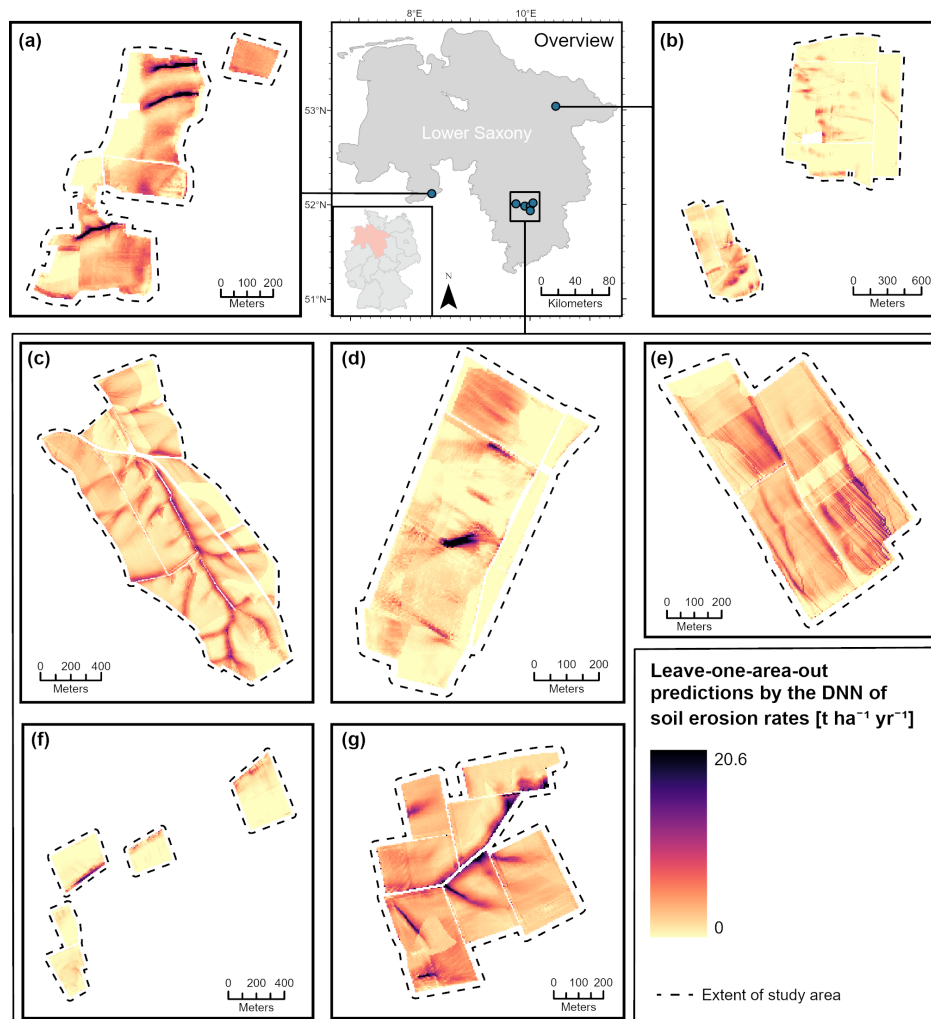


Figure 2. Predicted soil erosion rates using the leave-one-area-out approach by the deep neural network (DNN) in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt and (g) Brüggen.

225 The difference between the predicted soil erosion rate by the DNN and mapped erosion rates is shown in Fig. 3. The visualization shows that the few high erosion features within the study areas of Barum, Lamspringe, and Klein Ilde are strongly underestimated. In contrast, in the study areas of Küingdorf, Nette, Adenstedt, and Brügggen, the difference between modelled and mapped soil erosion was mostly below $5 \text{ t ha}^{-1} \text{ yr}^{-1}$, with no large over- and underestimations. Generally, locations with low mapped erosion rates ($< 1 \text{ t ha}^{-1} \text{ yr}^{-1}$) were predicted with relatively small errors, which is supported by the F1 scores (Table 3).

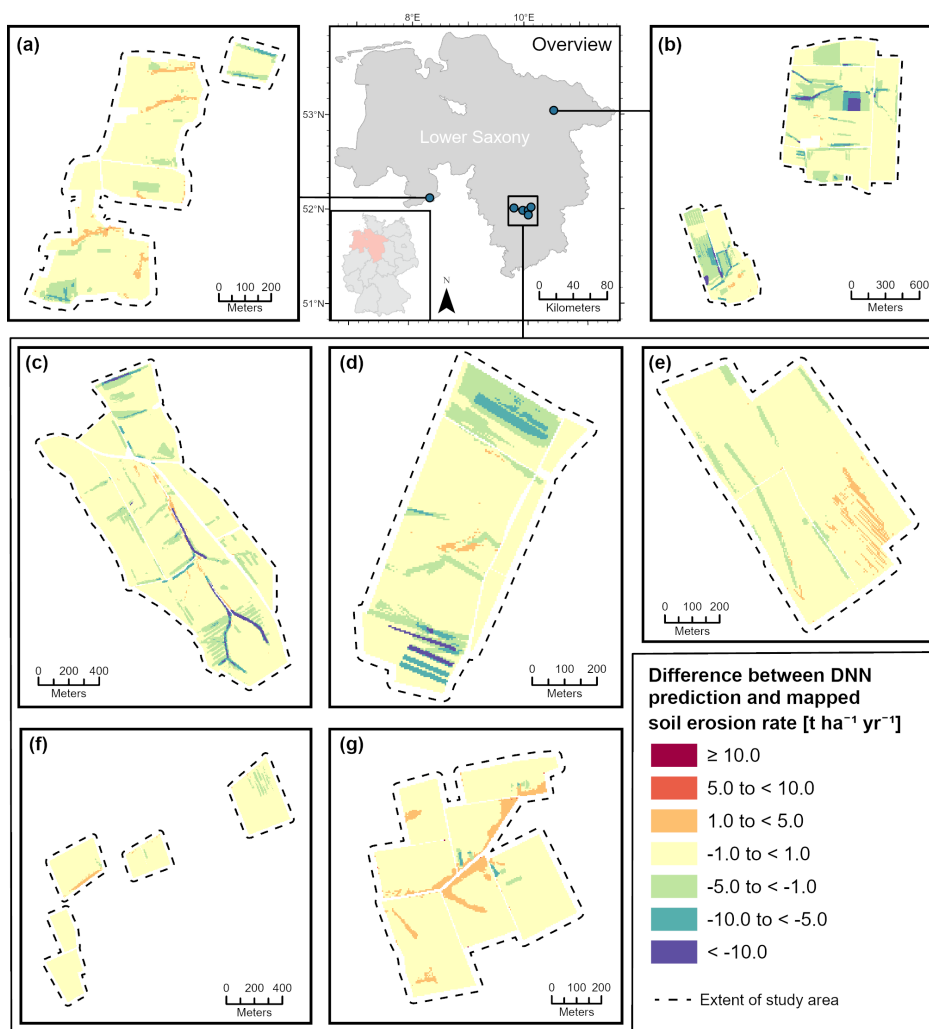


Figure 3. Difference between soil erosion rates predicted by the deep neural network (DNN) and the mapped values in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brügggen, where positive values indicate overestimation of the model and negative values indicate underestimation.

The permutation importance results are shown in Fig. 4. While machining direction vs. aspect showed consistently high importance across all models, with an average contribution of 18.2 %, the ranking of several other predictors varied noticeably between models. For instance, the CNN assigned comparatively high importance to the R factor (13.6 %), the DNN ranked aspect (180°) particularly strongly (27.0 %), and the RF attributed its highest importance to the LS factor (20.8 %). The SNN relies strongly on just two variables, machining direction vs. aspect and slope, which together account for more than 50 % of its total importance, whereas the RF, DNN and CNN distribute importance more evenly across multiple predictors.

Several variables exhibit consistently low importance (< 5 %) across all models, including slope length, the divergence–convergence index, profile curvature, plan curvature, flowline curvature and altitude. These predictors contributed to a small extent, if at all, to the model performance.

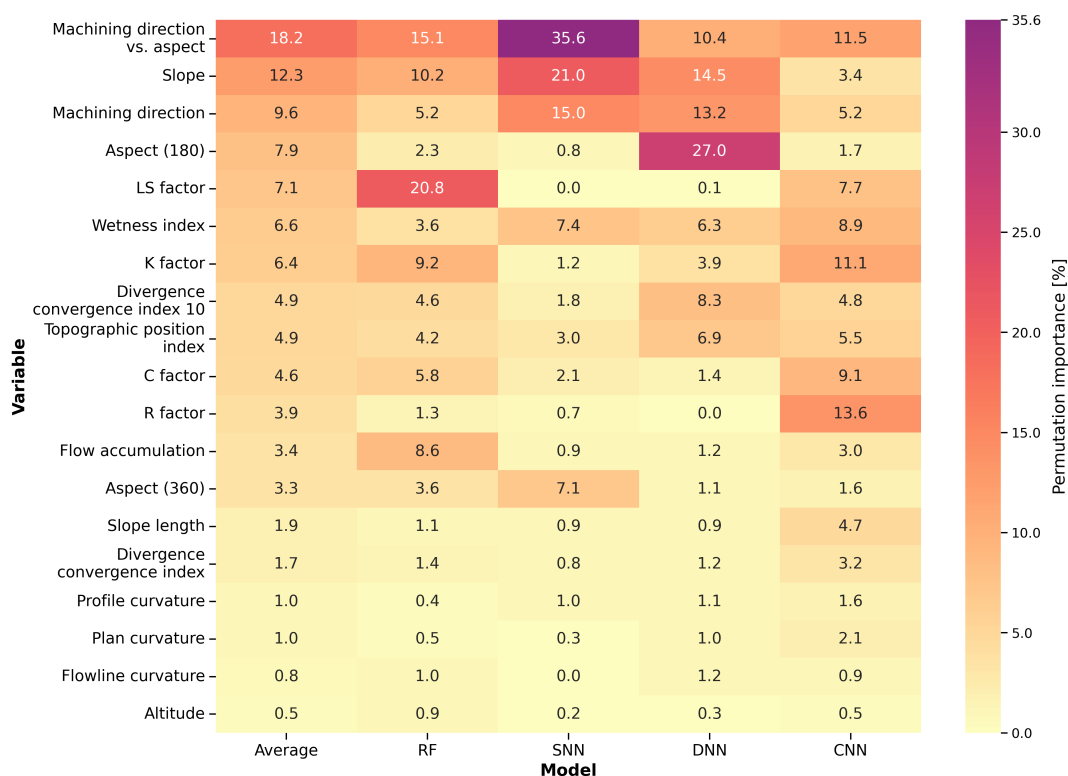


Figure 4. Heatmap of the permutation importance of all predictor variables, shown for the overall average across models and for the convolutional neural network (CNN), deep neural network (DNN), single-layer neural network (SNN) and random forest (RF). Values represent the normalized increase in mean squared error after permuting each variable.

4.1 Comparison of Model Performance

The results align with previous studies, such as Sarkar and Mishra (2018), indicating that increasing complexity through multiple hidden layers in a DNN can improve predictive performance compared to a SNN. Large neural networks have the advantage of being able to capture the non-linear interrelationships between variables and the predicted output using an extensive set of weights. This allows the neural network to model these relationships in greater detail, which, as shown in the results, can improve predictive performance. Our results also showed that predictive performance increased further with the addition of convolutional layers in the CNN, similar to the findings of Golkarian et al. (2023), although the differences compared with the DNN were only slight.

A large proportion (81.2 %) of the mapped data is below the threshold of $1 \text{ t ha}^{-1} \text{ yr}^{-1}$, and this value range therefore had a strong influence on the weighted average F1 score. As a result of underestimating these low erosion rates, the RF achieved a lower weighted average F1 score than the neural-network-based models. The neural networks, and particularly the CNN, were able to predict these low erosion rates moderately well, capturing the underlying relationships in this dominant class.

This strong class imbalance and the dominance of low soil-erosion values also affect the capability of the models to predict higher erosion rates. As shown in Table 3, all models fail to reproduce very high erosion rates ($\geq 5 \text{ t ha}^{-1} \text{ yr}^{-1}$). Since extreme values are rare in the training data, they provide too few examples for the models to learn the corresponding relationships reliably, which is a common problem in soil science for prediction tasks using machine learning (Sharififar and Sarmadian, 2023). In addition, the left-out areas have different landscape structures, land-management practices, and topographic patterns, leading to combinations of predictor variables that the models have not encountered during training. This distribution shift places many erosion features outside the range of conditions represented during training, making accurate prediction in previously unseen areas particularly challenging.

As shown in the visualizations of the predictions (e.g. Fig. 2 and Fig. 3), all models were able to reproduce parts of the spatial erosion-feature patterns in the unseen areas, although the predicted magnitudes were often too low. Only the SNN and DNN predicted very high soil-erosion rates at all, with maximum values of 18.5 and $20.6 \text{ t ha}^{-1} \text{ yr}^{-1}$, respectively. The high values predicted by the SNN, however, occurred only in a few isolated grid cells in Küingdorf, suggesting that they are more likely artefacts caused by local noise in the underlying predictor variables. In contrast, the DNN was able to predict higher values more consistently across multiple areas. The CNN output appeared notably smoother than that of the solely pixel-based models, albeit with a lower maximum predicted erosion rate. This smoothness can be attributed to the integration of local spatial context by the convolutional layers (Fu et al., 2019; Hilburn, 2023).

Overall, the CNN showed the most stable behaviour, combining low errors with the highest F1 scores. The DNN produced similarly strong results and achieved slightly lower MAE than the CNN. While the SNN exhibited the highest RMSE, the RF generally had the lowest performance, with the highest MAE and lowest F1 score. While the RF, SNN and DNN train on a pixel-by-pixel basis, the CNN potentially benefits from incorporating local spatial context through the input patches, which incorporate the neighbouring grid cells during training.

275 However, it is also important to note that the DNN, and particularly the SNN and RF models, require significantly fewer computational resources than the CNN, which can be beneficial when trained on big datasets or when the model is applied to predict soil erosion for larger areas. Therefore, depending on the specific use case, the potential trade-off between improved predictive performance and the increased complexity and computational demands of large neural networks, such as the CNN should be carefully considered.

280 The permutation-importance analysis highlights the combined influence of topographic and anthropogenic (i.e. land management) variables on soil erosion and reveals clear differences between the modelling approaches. These differences can be attributed to the respective model architectures. The RF relies on a hierarchical split structure, which causes it to prioritize variables that repeatedly improve node splits across trees, resulting in an importance pattern that differs from that of the neural networks (Breiman, 2001). The SNN relies strongly on only a few variables, indicating that its relatively simple architecture is less capable of capturing more complex interactions among the predictors (Nielsen, 2015). In contrast, the DNN distributes 285 importance more evenly across predictors, reflecting its greater capacity to model complex non-linear relationships (LeCun et al., 2015). The CNN, by applying convolutional filters over spatial neighbourhoods, prioritizes variables that contribute to local spatial patterns rather than relying solely on single-pixel attributes (LeCun et al., 2015). Consequently, each model emphasizes different predictors when optimizing performance in previously unseen areas. Despite these architectural differences, a consistent result across all models is that slope length, the divergence–convergence index, profile curvature, plan curvature, 290 flowline curvature and altitude show very low importance and contribute little to predicting soil erosion in this study.

4.2 Limitations and Future Research

Typical challenges and limitations in soil erosion modelling at the landscape scale include the high range of input variables from different domains for parametrization, leading to simplification in variable estimation and mismatches in the scale of input and output data, the application of models outside of the validation range and inherent restrictions of the used models. This 295 study addressed several of these limitations by utilizing a dataset based on a long-term soil erosion monitoring programme across multiple study areas, capturing long-term patterns of soil erosion by water within these regions. Additionally, input variables with relatively high spatial resolution were incorporated into machine learning models capable of capturing non-linear relationships at varying levels of complexity. Despite these approaches, several challenges and limitations remain. These challenges and limitations should be considered when interpreting the results and highlight key areas for future research to 300 enhance small-scale soil erosion predictions.

The results demonstrate that complex neural networks can partially reproduce erosion patterns when extrapolating to previously unseen areas. However, all models exhibited a strong tendency to underestimate soil erosion. This is mainly caused by the strong imbalance in the dataset, providing too few examples for the models to learn relationships for high erosion rates. These findings indicate the need for more comprehensive datasets that include additional study areas.

305 The results indicate furthermore that prediction performance can improve with increasing model complexity (Table 2 and Table 3), although the difference between CNN and DNN was relatively small. Whether this generally justifies the use of a more complex architecture, such as a CNN in soil erosion modelling, must be carefully evaluated based on the aim and context of the

respective study. Future research could investigate this further by comparing a broader range of machine learning approaches, such as gradient boosting machines, or by evaluating ensemble methods. In addition, more recent neural network architectures
310 such as transformers (e.g. Liu et al. 2024) may offer more effective prediction capabilities.

5 Conclusions

By comparing different artificial neural network architectures and using a random forest model as a benchmark, this study highlights the potential of more complex neural networks for modelling complex non-linear relationships between predictor variables and soil erosion at a field-to-landscape scale. The convolutional neural network achieved the lowest RMSE and
315 highest F1 score, followed closely by the deep neural network. When applied to previously unseen areas, all models were able to partially reproduce spatial erosion patterns but tended to strongly underestimate features with high erosion rates. The permutation importance analysis further indicates the relevance of variables such as machining direction vs. aspect and slope. Furthermore, the analysis provides guidance on which predictors may be omitted in soil erosion modelling.

The results suggest that simpler neural networks may lead to lower predictive performances and that more advanced ap-
320 proaches could offer further improvements in soil erosion modelling. To build upon these findings, future research should focus on further identifying and quantifying model uncertainties, as well as improving the generalizability of these models and their scalability to larger areas.

Code and data availability. The complete dataset and Python scripts used for model training, prediction, and evaluation are available at:
<https://doi.org/10.5281/zenodo.16032628> (Barthel, 2025).

Table A1: Overview of terrain and USLE-related variables, including data sources, processing steps, and summary statistics.

Variable	Description and summary statistics	Source and data processing
Altitude	Based on digital elevation model (DEM; meters above sea level [m]) <i>Mean: 142.1, SD: 55.4, Min: 57.7, Max: 271</i>	Source: Digital Elevation Model 1 (DEM1; 1 m × 1 m native resolution) (LGLN, 2024) Processing: Resampled to 5 m × 5 m (Nearest Neighbour) in ArcGIS Pro 2.6.2, tool “Resample”
Slope	Slope steepness in the direction of maximum elevation change in a 3×3 grid-cell window [°] <i>Mean: 3.7, SD: 3.0, Min: 0, Max: 39.1</i>	Source: Based on resampled DEM (5 m × 5 m) Processing: SAGA-GIS 7.8.1; tool “Slope, Aspect, Curvature” (Method: 9-parameter 2nd-order polynomial (Zevenbergen and Thorne, 1987))
Slope length	Accumulated downslope slope length (flow-path length) [m] using the D8 algorithm <i>Mean: 54.9, SD: 73.2, Min: 0, Max: 736.5</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Slope Length”; processed for each single field block ^a
Aspect 360	Direction of maximum elevation change in a 3×3 grid-cell window [°] (0–360) <i>Mean: 150.8, SD: 97.3, Min: 0.2, Max: 360</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Slope, Aspect, Curvature”
Aspect 180	Normalized direction of maximum elevation change in a 3×3 grid-cell window [°] (0–180) <i>Mean: 91.1, SD: 47.9, Min: 0.2, Max: 180</i>	Source: Based on DEM-derived Aspect360 Processing: SAGA-GIS 7.8.1; tool “Grid Calculator”; Formula: $ifelse(Aspect360 > 180, Aspect360 - 180, Aspect360)$
Plan curvature	Curvature perpendicular to the direction of maximum elevation change in a 3×3 grid-cell window <i>Mean: 0, SD: 0.1, Min: -12.6, Max: 13.2</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Slope, Aspect, Curvature”
Profile curvature	Curvature in the direction of maximum elevation change in a 3×3 grid-cell window <i>Mean: 0, SD: 0.01, Min: -0.1, Max: 0.1</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Slope, Aspect, Curvature”
Flow line curvature	Curvature along flow lines in a 3×3 grid-cell window <i>Mean: 0, SD: 0.0001, Min: -0.01, Max: 0.01</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Slope, Aspect, Curvature”

Continued on next page

Variable	Description and summary statistics	Source and data processing
Topographic position index (TPI)	Altitude of each pixel in contrast to its neighbours (following Guisan et al. 1999) <i>Mean: -0.1, SD: 1.1, Min: -9.8, Max: 6.3</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Topographic Position Index (TPI)” (Standardize: NO; Scale: 0–100; Weighting: none)
Flow accumulation	Upslope draining area (cell count) based on multiple flow direction algorithm <i>Mean: 768.3, SD: 1738.3, Min: 25, Max: 135422.4</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Flow Accumulation (Top-Down)” (Unit: 1; Method: MFD; Convergence: 1.1); processed per field block
Wetness index	SAGA wetness index (modified topographic wetness index following Böhner et al. 2002) <i>Mean: 6.6, SD: 1, Min: 2.7, Max: 11.4</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “SAGA Wetness Index” (Suction: 10; Area: $\sqrt{\text{catchment area}}$; Slope: catchment slope)
Divergence-convergence index (DCI)	Index of water flow divergence or convergence (3×3 grid-cell area) <i>Mean: -0.1, SD: 11.4, Min: -100, Max: 89.6</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Convergence Index”
DCI 10	Index of water flow divergence or convergence (10×10 grid-cell area) <i>Mean: -0.3, SD: 15.1, Min: -61.4, Max: 80.4</i>	Source: Based on resampled DEM Processing: SAGA-GIS 7.8.1; tool “Convergence (Search Index)” (Radius: 10 cells)
Machining direction (MD)	Orientation of the machining direction (tramlines) [°] (0–180) <i>Mean: 102.8, SD: 56.3, Min: 0.4, Max: 179.9</i>	Source: Field mapping and aerial photo interpretation Processing: Digitized for main fields and headlands; converted to DEM resolution
MD vs. aspect	Angle between MD and aspect 180 [°] (0–90) <i>Mean: 48.8, SD: 27.3, Min: 0, Max: 90</i>	Source: Based on aspect 180 and MD Processing: ArcGIS Pro 2.6.2; “Raster Calculator”
LS factor	USLE topographic factor (length–slope factor) following Desmet and Govers (1996) <i>Mean: 1.7, SD: 1.4, Min: 0, Max: 27.1</i>	Source: Based on DEM Processing: SAGA-GIS 7.8.1; tool “LS-Factor, Field Based” (LS Calculation: Moore and Nieber 1989); processed per field block ^a
K factor	USLE soil erodibility factor [t h (MJ mm) ⁻¹] <i>Mean: 0.5, SD: 0.1, Min: 0.1, Max: 0.6</i>	Source: Soil Map of Lower Saxony (1:50,000) (LBEG, 2017) Processing: Converted to DEM resolution

Continued on next page

Variable	Description and summary statistics	Source and data processing
C factor	USLE crop and management factor <i>Mean: 0.1, SD: 0.04, Min: 0, Max: 0.6</i>	Source: Crop-management monitoring data (2000–2015) Processing: Calculated following Schwertmann et al. (1987); converted to DEM resolution
R factor	USLE rainfall erosivity factor [MJ mm (h ha yr) ⁻¹] <i>Mean: 78.2, SD: 4.5, Min: 71.4, Max: 91.8</i>	Source: Mean yearly R factor (2001–2017) from RADKLIM (Winterrath et al., 2018); 1 km gridded precipitation data Processing: Converted to DEM resolution

^a Field block: A contiguous agricultural reference parcel bounded by stable landscape features (roads, paths, hedges, ditches, rivers, forests, etc.)

Appendix B: Pairwise correlation and multicollinearity analysis

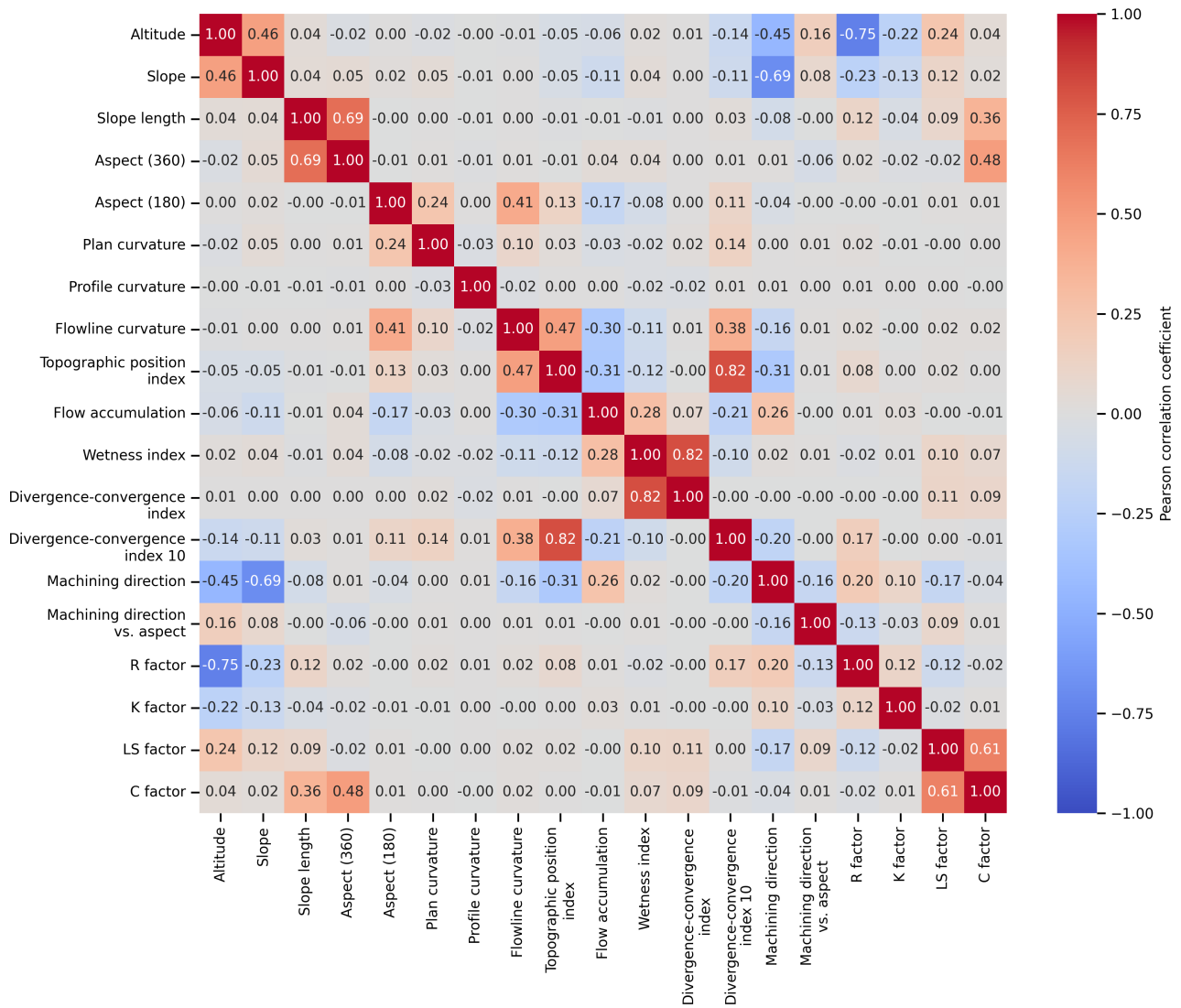


Figure B1. Pairwise linear correlation analysis of all predictor variables used in this study based on the Pearson correlation coefficient.

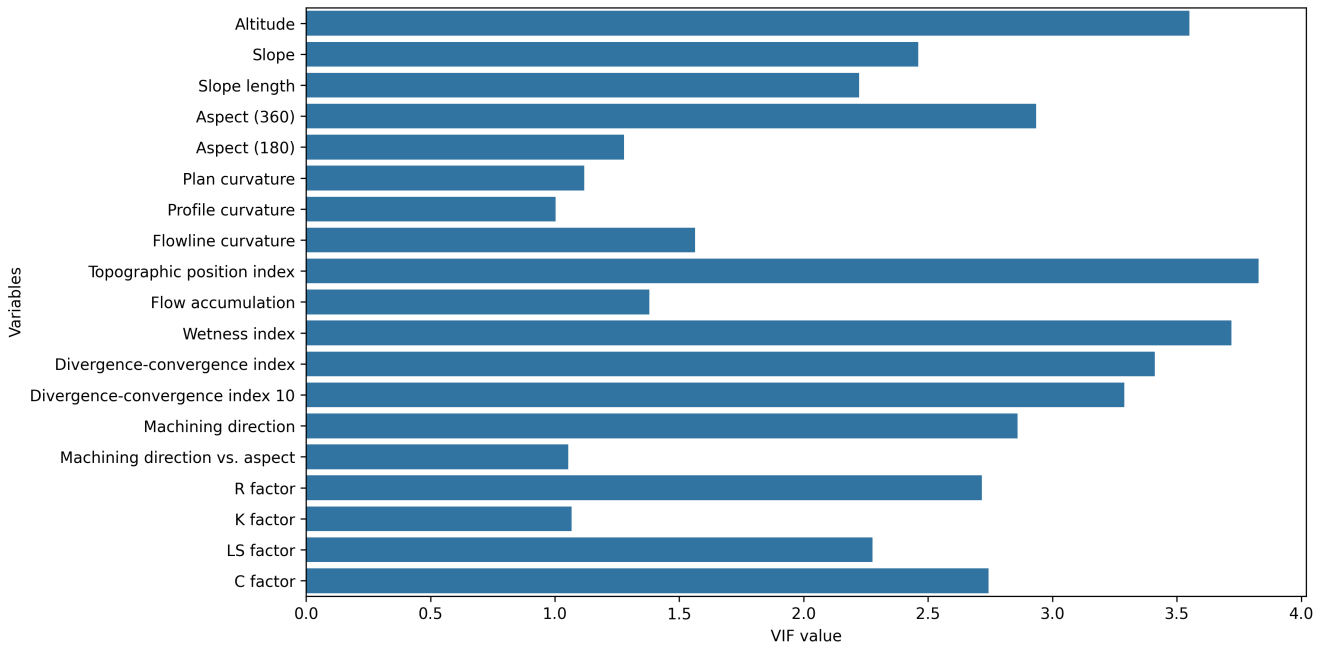


Figure B2. Variance inflation analysis of predictor variables highlights the level of multicollinearity, with higher VIF values indicating greater multicollinearity among predictors.

Appendix C: Results

C1 Raw model predictions

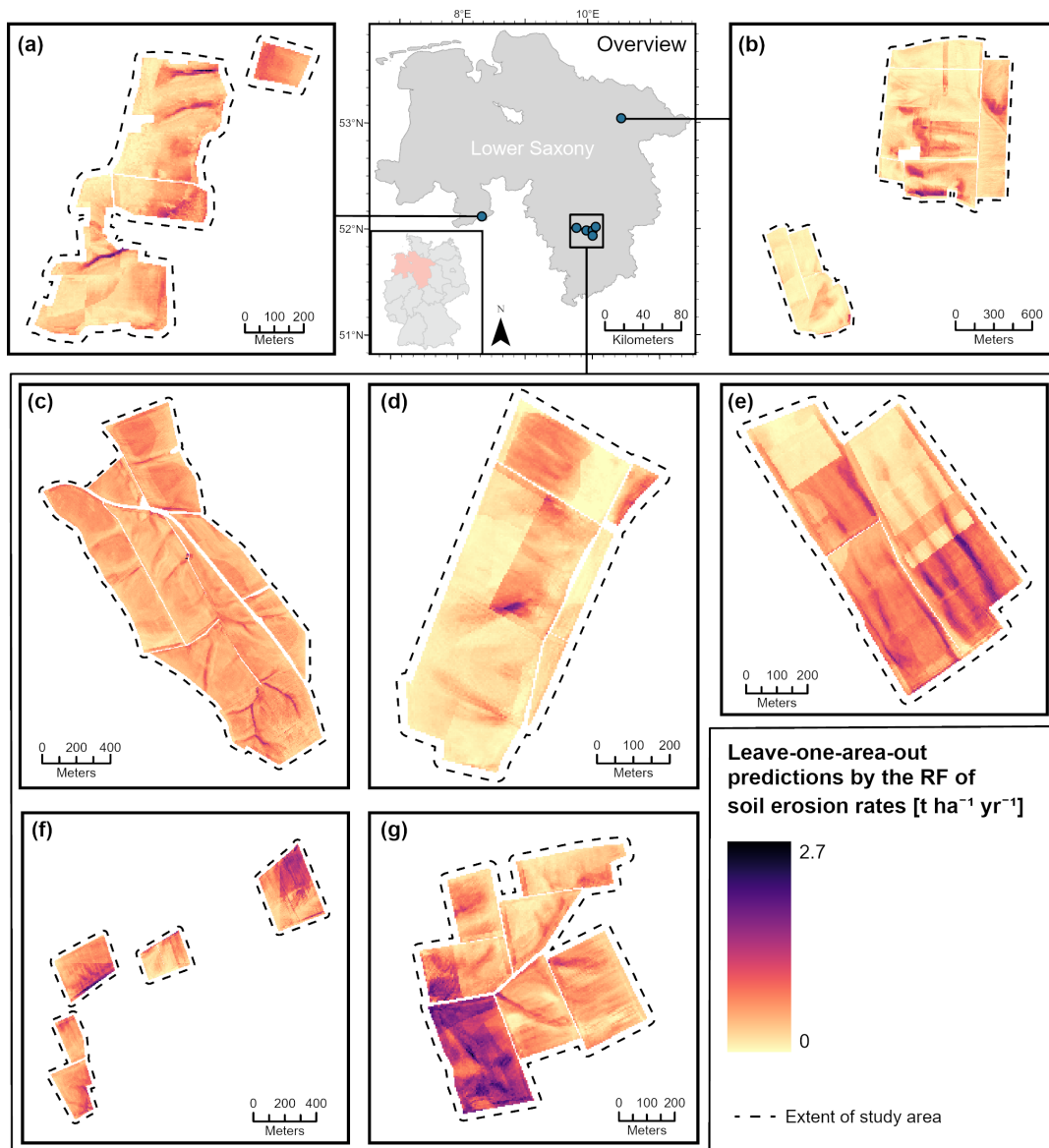


Figure C1. Predicted soil erosion rates using the leave-one-area-out approach by the random Forest (RF) in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt and (g) Brüggen.

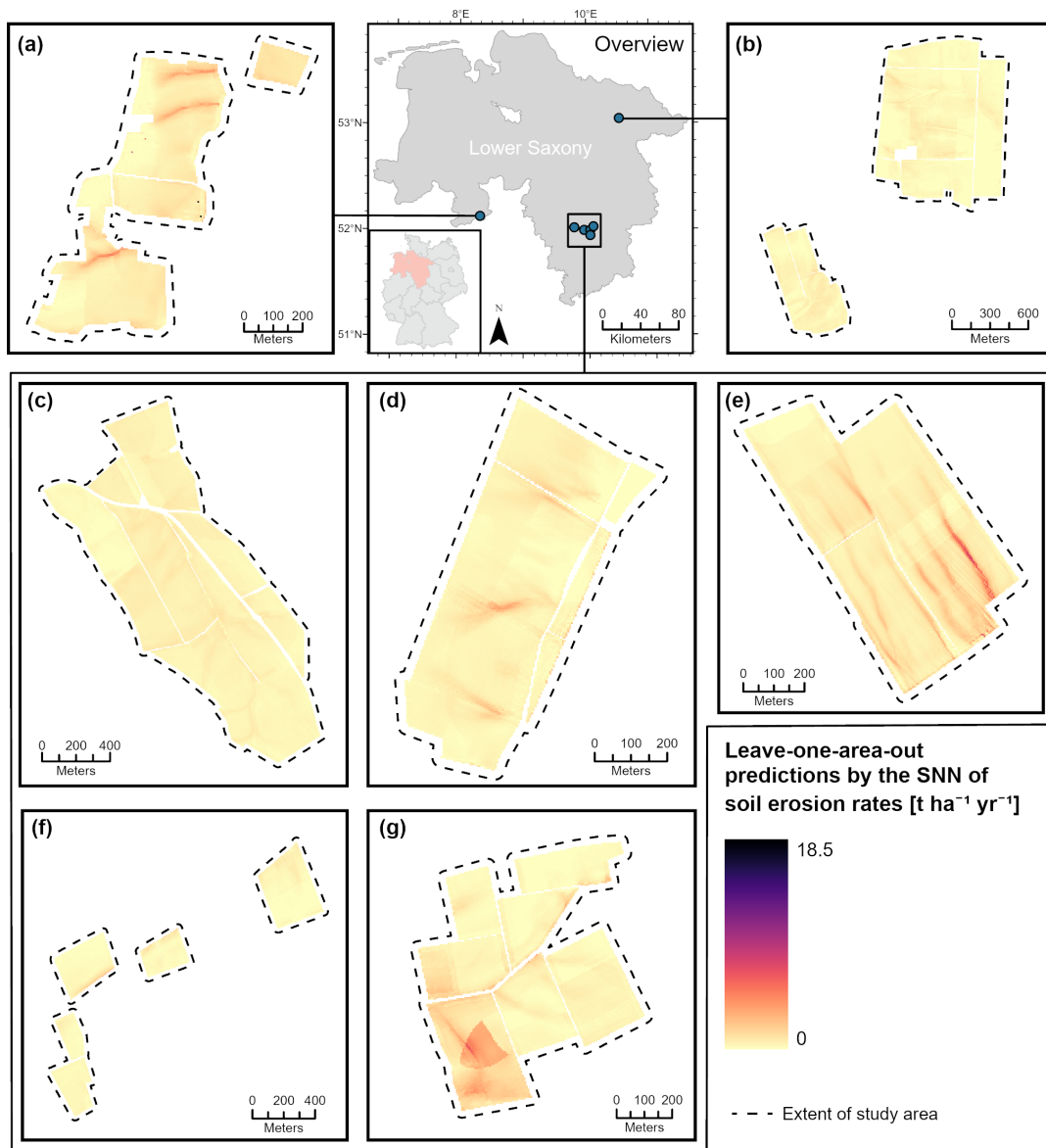


Figure C2. Predicted soil erosion rates using the leave-one-area-out approach by the single-hidden layer neural network (SNN) in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt and (g) Brüggen.

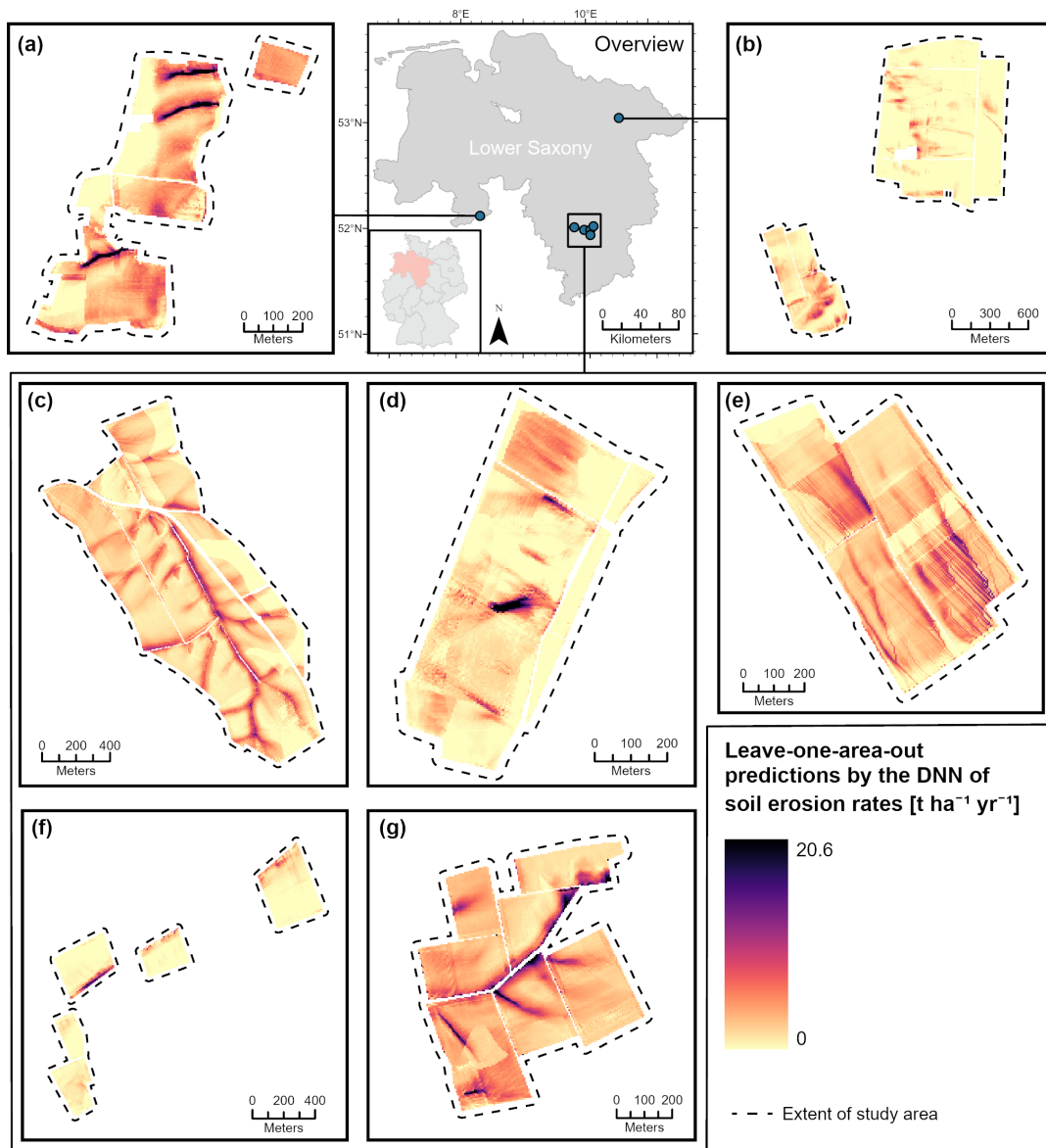


Figure C3. Predicted soil erosion rates using the leave-one-area-out approach by the deep neural network (DNN) in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt and (g) Brüngen.

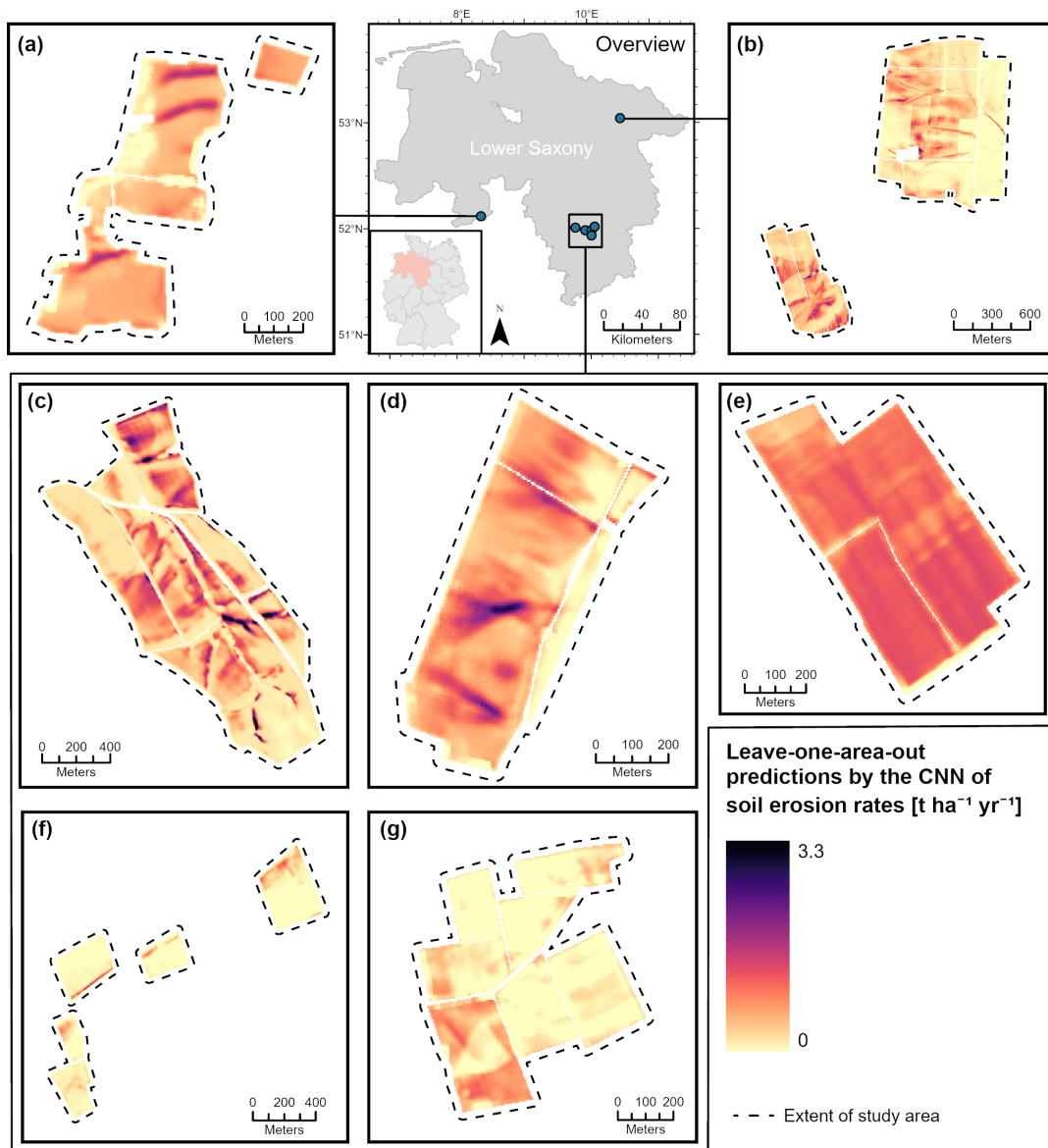


Figure C4. Predicted soil erosion rates using the leave-one-area-out approach by the convolutional neural network (CNN) in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt and (g) Brüggem.

C2 Differences between predicted and mapped soil erosion rates

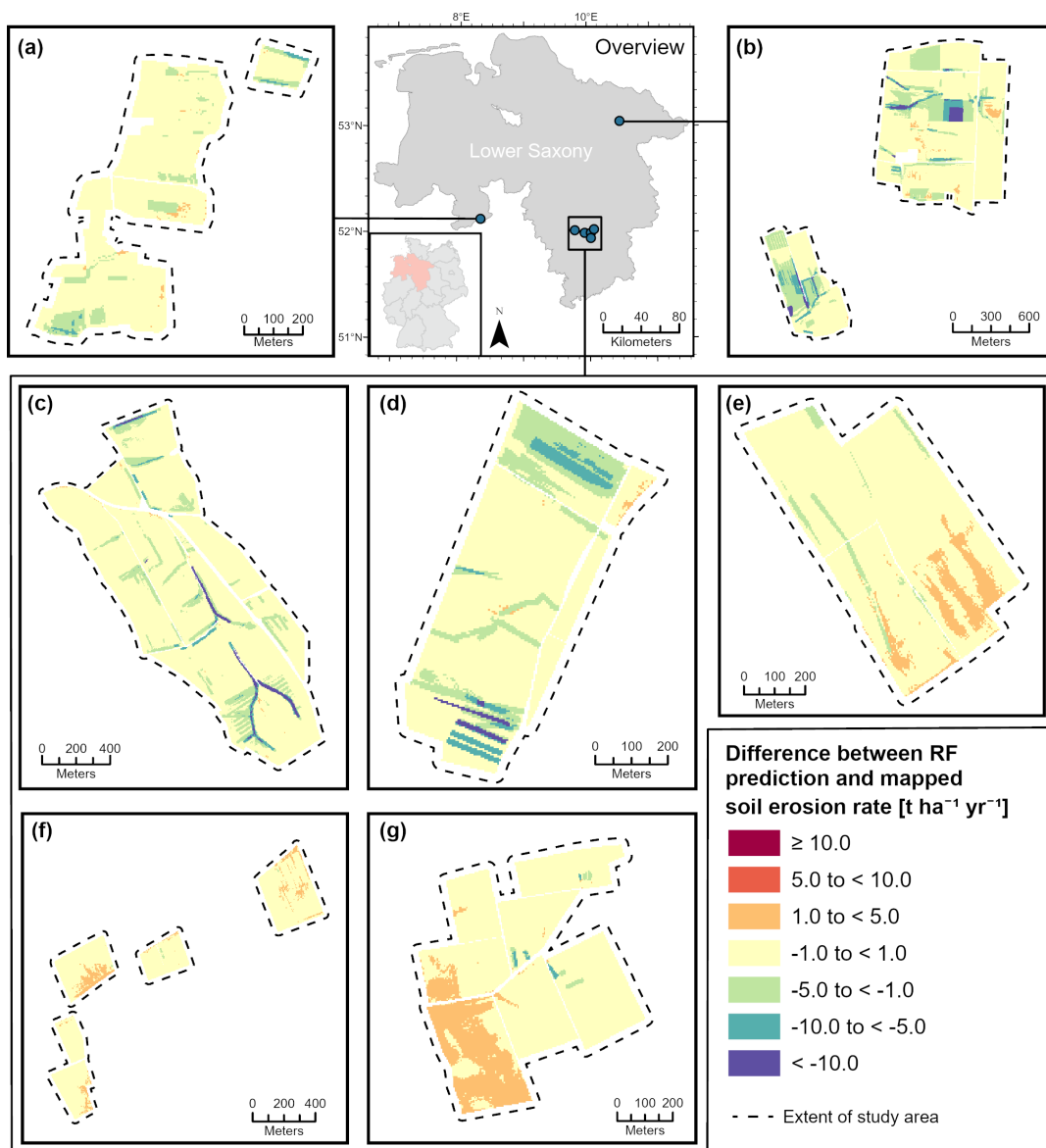


Figure C5. Difference between soil erosion rates predicted by the random forest (RF) and the mapped values in (a) Küningdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brüggen, where positive values indicate overestimation of the model and negative values indicate underestimation.

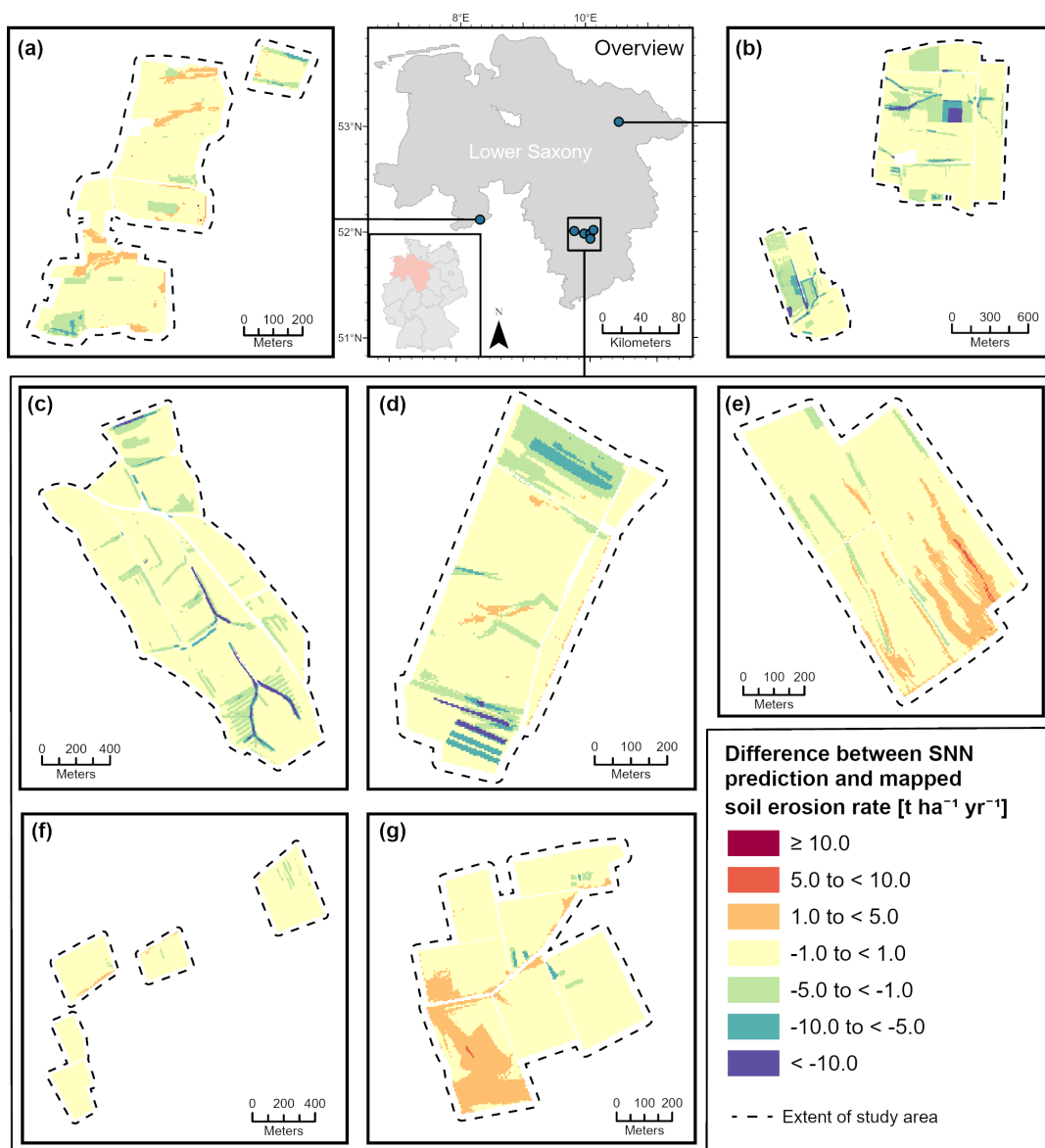


Figure C6. Difference between soil erosion rates predicted by the single-hidden layer neural network (SNN) and the mapped values in (a) Küngdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brüggen, where positive values indicate overestimation of the model and negative values indicate underestimation.

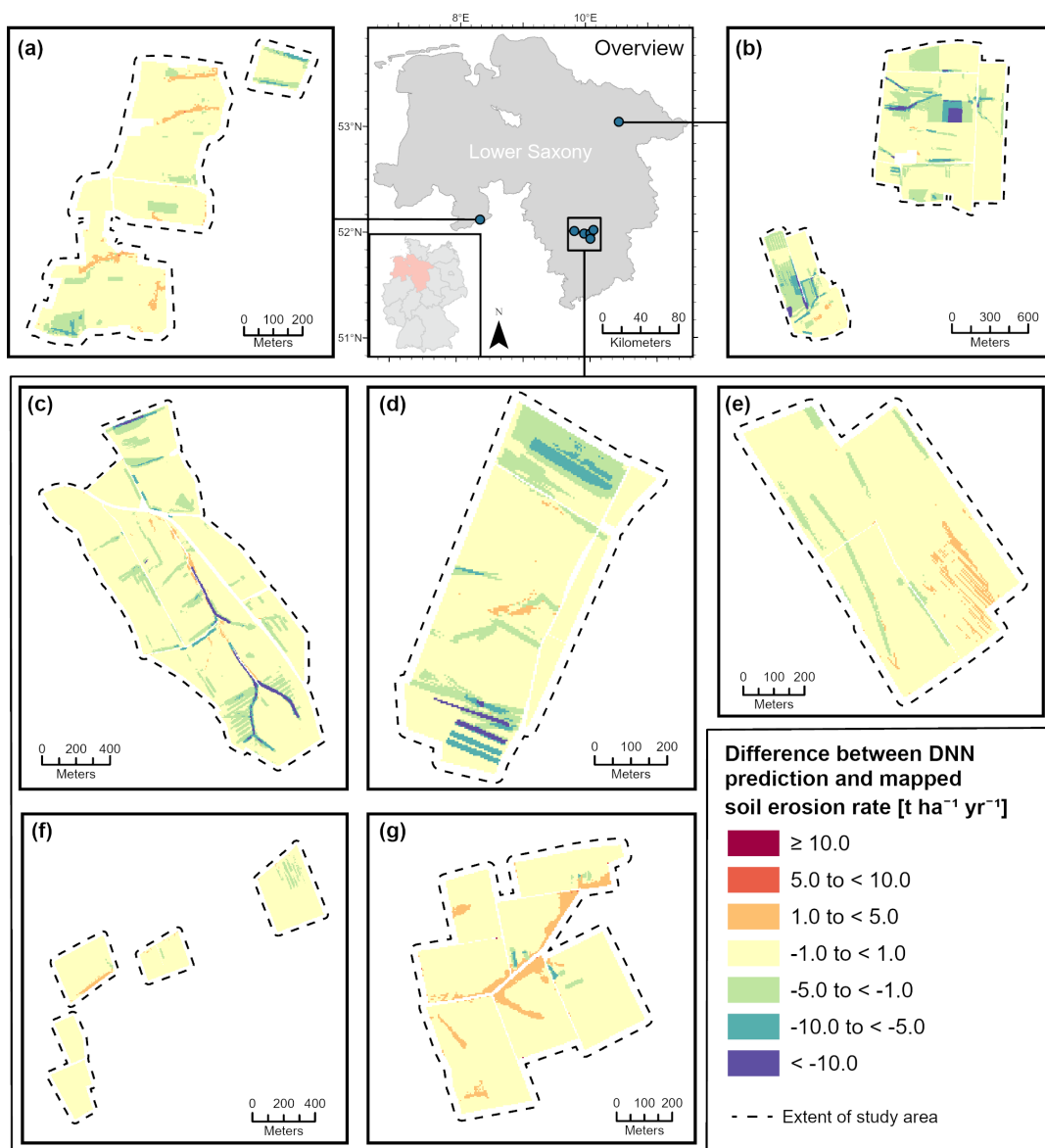


Figure C7. Difference between soil erosion rates predicted by the deep neural network (DNN) and the mapped values in (a) Küingdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brüggen, where positive values indicate overestimation of the model and negative values indicate underestimation.

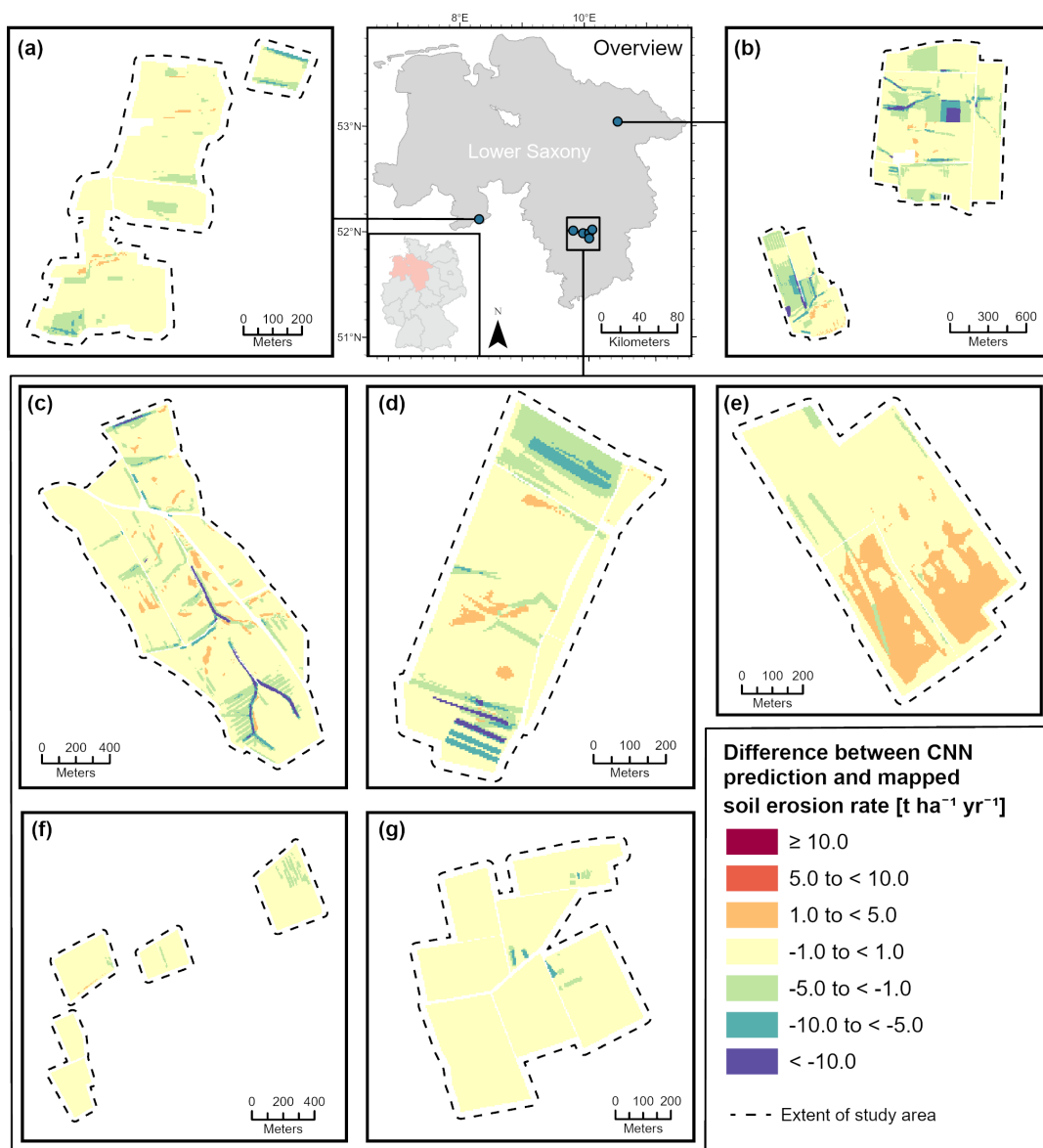


Figure C8. Difference between soil erosion rates predicted by the convolutional neural network (CNN) and the mapped values in (a) Küngsdorf, (b) Barum, (c) Lamspringe, (d) Klein Ilde, (e) Nette, (f) Adenstedt, and (g) Brüggen, where positive values indicate overestimation of the model and negative values indicate underestimation.

330 *Author contributions.* NB: conceptualization, methodology, investigation, visualization, writing (original draft), writing (review and editing).
SO: conceptualization, data curation, writing (review and editing). BB: conceptualization, project administration, funding acquisition, writing
(review and editing). BSK: conceptualization, data curation, visualization, writing (review and editing).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. The research and collection of the associated data were made possible through the continuous funding provided by the
335 Lower Saxony State Authority for Mining, Energy and Geology (LBEG). We would also like to acknowledge the many years of work by all
those who have contributed to the creation of this long-term dataset through fieldwork since 2000. In particular, we thank Frank Beisiegel
and Heiko van Wensen for their continuous monitoring efforts since the beginning of the programme. We thank Angie Faust for proofreading
this work.

References

- 340 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <https://www.tensorflow.org/>, 2015.
- 345 Alewell, C., Borrelli, P., Meusburger, K., and Panagos, P.: Using the USLE: Chances, challenges and limitations of soil erosion modelling, *International soil and water conservation research*, 7, 203–225, <https://doi.org/10.1016/j.iswcr.2019.05.004>, 2019.
- Altmann, A., Toloşi, L., Sander, O., and Lengauer, T.: Permutation importance: a corrected feature importance measure, *Bioinformatics*, 26, 1340–1347, <https://doi.org/10.1093/bioinformatics/btq134>, 2010.
- Anache, J. A., Flanagan, D. C., Srivastava, A., and Wendland, E. C.: Land use and climate change impacts on runoff and soil erosion at the hillslope scale in the Brazilian Cerrado, *Science of the Total Environment*, 622, 140–151, <https://doi.org/10.1016/j.scitotenv.2017.11.257>, 2018.
- 350 Avand, M., Mohammadi, M., Mirchooli, F., Kavian, A., and Tiefenbacher, J. P.: A new approach for smart soil erosion modeling: integration of empirical and machine-learning models, *Environmental Modeling & Assessment*, 28, 145–160, <https://doi.org/10.1007/s10666-022-09858-x>, 2023.
- 355 Barthel, N.: Zenodo [code], <https://doi.org/10.5281/zenodo.16032628>, 2025.
- Batista, P. V., Möller, M., Schmidt, K., Waldau, T., Seufferheld, K., Htitiou, A., Golla, B., Ebertseder, F., Auerswald, K., and Fiener, P.: Soil-erosion events on arable land are nowcast by machine learning, *Catena*, 256, 109 080, <https://doi.org/10.1016/j.catena.2025.109080>, 2025.
- Böhner, J., Köthe, R., Conrad, O., Gross, J., Ringeler, A., and Selige, T.: Soil regionalisation by means of terrain analysis and process parameterisation, in: *Soil Classification 2001*, edited by Micheli, E., Nachtergaele, F., and Montanarella, L., no. 7 in European Soil Bureau Research Report, pp. 213–222, Office for Official Publications of the European Communities, Luxembourg, https://esdac.jrc.ec.europa.eu/ESDB_Archive/eusoils_docs/esb_rr/n07_ESBResRep07/601Bohner.pdf, eUR 20398 EN, 2002.
- 360 Borrelli, P., Van Oost, K., Meusburger, K., Alewell, C., Lugato, E., and Panagos, P.: A step towards a holistic assessment of soil degradation in Europe: Coupling on-site erosion with sediment transfer and carbon fluxes, *Environmental research*, 161, 291–298, <https://doi.org/10.1016/j.envres.2017.11.009>, 2018.
- 365 Borrelli, P., Alewell, C., Alvarez, P., Anache, J. A. A., Baartman, J., Ballabio, C., Bezak, N., Biddoccu, M., Cerdà, A., Chalise, D., et al.: Soil erosion modelling: A global review and statistical analysis, *Science of the total environment*, 780, <https://doi.org/10.1016/j.scitotenv.2021.146494>, 2021.
- Botschek, J., Billen, N., Brandhuber, R., Bug, J., Deumlich, D., Duttmann, R., Elhaus, D., Mollenhauer, K., Prasuhn, V., Röder, C., et al.: Bodenerosion durch Wasser–Kartieranleitung zur Erfassung aktueller Erosionsformen, DWA-Regelwerk, Merkblatt DWA-M, 921, 112, 2021.
- 370 Breiman, L.: Random forests, *Machine Learning*, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- Capelle, A.: Die erosionsgefährdete Landesfläche in Niedersachsen und Bremen, *Z. Kult. Landentwicl*, 31, 11–17, 1990.
- Capelle, A. and Lüders, R.: Die potentielle Erosionsgefährdung der Böden in Niedersachsen, *Göttinger Bodenkundliche Berichte*, 83, 107–375 127, 1985.

- Chinchor, N.: MUC-4 Evaluation Metrics, in: Fourth Message Understanding Conference (MUC-4): Proceedings of a Conference Held in McLean, Virginia, June 16-18, 1992, 1992.
- Daoud, J. I.: Multicollinearity and regression analysis, in: Journal of Physics: Conference Series, vol. 949, p. 012009, IOP Publishing, <https://doi.org/10.1088/1742-6596/949/1/012009>, 2017.
- 380 De la Rosa, D., Mayol, F., Moreno, J., Bonsón, T., and Lozano, S.: An expert system/neural network model (ImpelERO) for evaluating agricultural soil erosion in Andalusia region, southern Spain, *Agriculture, Ecosystems & Environment*, 73, 211–226, [https://doi.org/10.1016/S0167-8809\(99\)00050-X](https://doi.org/10.1016/S0167-8809(99)00050-X), 1999.
- Desmet, P. J. and Govers, G.: A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units, *Journal of soil and water conservation*, 51, 427–433, <https://doi.org/10.1080/00224561.1996.12457102>, 1996.
- 385 DIN 19708:2022-08: Bodenbeschaffenheit- Ermittlung der Erosionsgefährdung von Böden durch Wasser mit Hilfe der ABAG, Tech. rep., Beuth Verlag GmbH, Berlin, 2022.
- DVWK: Bodenerosion durch Wasser: DVWK-Merkblatt 239: Kartieranleitung zur Erfassung aktueller Erosionsformen, Wirtschafts- und Verl.-Ges. Gas und Wasser, Bonn, 1996.
- Ebrahimi-Khusfi, Z., Nafarzadegan, A. R., and Dargahian, F.: Predicting the number of dusty days around the desert wet-lands in southeastern Iran using feature selection and machine learning techniques, *Ecological Indicators*, 125, 107499, <https://doi.org/10.1016/j.ecolind.2021.107499>, 2021.
- 390 Fiener, P., Dostál, T., Krása, J., Schmaltz, E., Strauss, P., and Wilken, F.: Operational USLE-based modelling of soil erosion in Czech Republic, Austria, and Bavaria—Differences in model adaptation, parametrization, and data availability, *Applied Sciences*, 10, 3647, <https://doi.org/10.3390/app10103647>, 2020.
- 395 Fu, B., Zhao, X., Li, Y., Wang, X., and Ren, Y.: A convolutional neural networks denoising approach for salt and pepper noise, *Multimedia Tools and Applications*, 78, 30707–30721, <https://doi.org/10.1007/s11042-018-6521-4>, 2019.
- Garosi, Y., Sheklabadi, M., Conoscenti, C., Pourghasemi, H. R., and Van Oost, K.: Assessing the performance of GIS-based machine learning models with different accuracy measures for determining susceptibility to gully erosion, *Science of the Total Environment*, 664, 1117–1132, <https://doi.org/10.1016/j.scitotenv.2019.02.093>, 2019.
- 400 Gholami, V., Sahour, H., and Amri, M. A. H.: Soil erosion modeling using erosion pins and artificial neural networks, *Catena*, 196, 104902, <https://doi.org/10.1016/j.catena.2020.104902>, 2021.
- Ghorbanzadeh, O., Shahabi, H., Mirchooli, F., Valizadeh Kamran, K., Lim, S., Aryal, J., Jarihani, B., and Blaschke, T.: Gully erosion susceptibility mapping (GESM) using machine learning methods optimized by the multi-collinearity analysis and K-fold cross-validation, *Geomatics, Natural Hazards and Risk*, 11, 1653–1678, <https://doi.org/10.1080/19475705.2020.1810138>, 2020.
- 405 Ghosh, A. and Maiti, R.: Soil erosion susceptibility assessment using logistic regression, decision tree and random forest: study on the Mayurakshi river basin of Eastern India, *Environmental Earth Sciences*, 80, 328, <https://doi.org/10.1007/s12665-021-09631-5>, 2021.
- Golkarian, A., Khosravi, K., Panahi, M., and Clague, J. J.: Spatial variability of soil water erosion: Comparing empirical and intelligent techniques, *Geoscience Frontiers*, 14, 101456, <https://doi.org/10.1016/j.gsf.2022.101456>, 2023.
- Guerra, C. A., Rosa, I. M., Valentini, E., Wolf, F., Filipponi, F., Karger, D. N., Nguyen Xuan, A., Mathieu, J., Lavelle, P., and Eisenhauer, N.: Global vulnerability of soil ecosystems to erosion, *Landscape ecology*, 35, 823–842, <https://doi.org/10.1007/s10980-020-00984-z>, 2020.
- 410 Guisan, A., Weiss, S. B., and Weiss, A. D.: GLM versus CCA spatial modeling of plant species distribution, *Plant ecology*, 143, 107–122, <https://doi.org/10.1023/A:1009841519580>, 1999.

- Hilburn, K. A.: Understanding spatial context in convolutional neural networks using explainable methods: Application to interpretable GREMLIN, *Artificial Intelligence for the Earth Systems*, 2, 220 093, <https://doi.org/10.1175/AIES-D-22-0093.1>, 2023.
- 415 Hoerl, A. E. and Kennard, R. W.: Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, 12, 55–67, 1970.
- Igwe, P., Onuigbo, A., Chinedu, O., Ezeaku, I., Muoneke, M., et al.: Soil erosion: A review of models and applications, *International Journal of Advanced Engineering Research and Science*, 4, 237 341, <https://doi.org/10.22161/ijaers.4.12.22>, 2017.
- Issaka, S. and Ashraf, M. A.: Impact of soil erosion and degradation on water quality: a review, *Geology, Ecology, and Landscapes*, 1, 1–11, <https://doi.org/10.1080/24749508.2017.1301053>, 2017.
- 420 Jaafari, A., Janizadeh, S., Abdo, H. G., Mafi-Gholami, D., and Adeli, B.: Understanding land degradation induced by gully erosion from the perspective of different geoenvironmental factors, *Journal of Environmental Management*, 315, 115 181, <https://doi.org/10.1016/j.jenvman.2022.115181>, 2022.
- Khosravi, K., Rezaie, F., Cooper, J. R., Kalantari, Z., Abolfathi, S., and Hatamiafkoueieh, J.: Soil water erosion susceptibility assessment using deep learning algorithms, *Journal of Hydrology*, 618, 129 229, <https://doi.org/10.1016/j.jhydrol.2023.129229>, 2023.
- 425 Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, <https://doi.org/10.48550/ARXIV.1412.6980>, 2014.
- Kingsford, C. and Salzberg, S. L.: What are decision trees?, *Nature biotechnology*, 26, 1011–1013, <https://doi.org/10.1038/nbt0908-1011>, 2008.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E.: ImageNet classification with deep convolutional neural networks, *Communications of the ACM*, 60, 84–90, <https://doi.org/10.1145/3065386>, 2017.
- 430 Kumar, M., Sahu, A. P., Sahoo, N., Dash, S. S., Raul, S. K., and Panigrahi, B.: Global-scale application of the RUSLE model: a comprehensive review, *Hydrological Sciences Journal*, 67, 806–830, <https://doi.org/10.1080/02626667.2021.2020277>, 2022.
- LBEG: Bodenkarte von Niedersachsen 1:50 000 (BK50): Blattschnittfreie Vektordaten, Geodataset, <https://nibis.lbeg.de/geonetwork/srv/api/records/611135b8-7168-4960-ad9d-3103ee96dcc6>, landesamt für Bergbau, Energie und Geologie (LBEG), 2017.
- LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, *nature*, 521, 436–444, <https://doi.org/10.1038/nature14539>, 2015.
- 435 LGLN: Digital Elevation Model (DEM1), <https://arcg.is/aH4Cy>, data provider: Landesamt für Geoinformation und Landesvermessung Niedersachsen (LGLN), licensed under CC BY 4.0, 2024.
- Licznar, P. and Nearing, M.: Artificial neural networks of soil erosion and runoff prediction at the plot scale, *Catena*, 51, 89–114, [https://doi.org/10.1016/S0341-8162\(02\)00147-9](https://doi.org/10.1016/S0341-8162(02)00147-9), 2003.
- Liu, C., Fan, H., and Wang, Y.: Gully erosion susceptibility assessment using three machine learning models in the black soil region of Northeast China, *Catena*, 245, 108 275, <https://doi.org/10.1016/j.catena.2024.108275>, 2024.
- 440 Moore, I. D. and Nieber, J. L.: Landscape assessment of soil erosion and nonpoint source pollution, *Journal of the Minnesota Academy of Science*, 55, 18–25, 1989.
- Morgan, R., Quinton, J., Smith, R., Govers, G., Poesen, J., Auerswald, K., Chisci, G., Torri, D., and Styczen, M.: The European Soil Erosion Model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments, *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Group*, 23, 527–544, [https://doi.org/10.1002/\(SICI\)1096-9837\(199806\)23:6<527::AID-ESP868>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1096-9837(199806)23:6<527::AID-ESP868>3.0.CO;2-5), 1998.
- 445 Nair, V. and Hinton, G. E.: Rectified linear units improve restricted boltzmann machines, in: *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, 2010.
- Nearing, M. A., Foster, G. R., Lane, L., and Finkner, S.: A process-based soil erosion model for USDA-Water Erosion Prediction Project 450 technology, *Transactions of the ASAE*, 32, 1587–1593, 1989.

- Ng, A. Y.: Feature selection, L_1 vs. L_2 regularization, and rotational invariance, in: Proceedings of the twenty-first international conference on Machine learning, p. 78, <https://doi.org/10.1145/1015330.1015435>, 2004.
- Nielsen, M. A.: Neural networks and deep learning, vol. 25, Determination press San Francisco, CA, USA, 2015.
- O'Brien, R. M.: A caution regarding rules of thumb for variance inflation factors, *Quality & quantity*, 41, 673–690, <https://doi.org/10.1007/s11135-006-9018-6>, 2007.
- 455 Panagos, P., Ballabio, C., Himics, M., Scarpa, S., Matthews, F., Bogonos, M., Poesen, J., and Borrelli, P.: Projections of soil loss by water erosion in Europe by 2050, *Environmental Science & Policy*, 124, 380–392, <https://doi.org/10.1016/j.envsci.2021.07.012>, 2021.
- Parsons, A. J.: How reliable are our methods for estimating soil erosion by water?, *Science of the Total Environment*, 676, 215–221, <https://doi.org/10.1016/j.scitotenv.2019.04.307>, 2019.
- 460 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, 12, 2825–2830, 2011.
- Pieri, L., Bittelli, M., Wu, J. Q., Dun, S., Flanagan, D. C., Pisa, P. R., Ventura, F., and Salvatorelli, F.: Using the Water Erosion Prediction Project (WEPP) model to simulate field-observed runoff and erosion in the Apennines mountain range, Italy, *Journal of hydrology*, 336, [84–97](https://doi.org/10.1016/j.jhydrol.2006.12.014), <https://doi.org/10.1016/j.jhydrol.2006.12.014>, 2007.
- 465 Plambeck, N. O.: Reassessment of the potential risk of soil erosion by water on agricultural land in Germany: Setting the stage for site-appropriate decision-making in soil and water resources management, *Ecological Indicators*, 118, 106732, <https://doi.org/10.1016/j.ecolind.2020.106732>, 2020.
- Raschka, S.: Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning, <https://arxiv.org/abs/1811.12808>, 2020.
- 470 Renard, K. G.: Predicting soil erosion by water: a guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE), US Department of Agriculture, Agricultural Research Service, 1997.
- Rohr, W., Mosimann, T., Bono, R., Rüttimann, M., and Prasuhn, V.: Kartieranleitung zur Aufnahme von Bodenerosionsformen und-schäden auf Ackerflächen, *Legende, Erläuterungen zur Kartiertechnik, Schadensdokumentation und Fehlerabschätzung. Materialien zur Physiso-geographie*, 14, 1990.
- 475 Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning representations by back-propagating errors, *nature*, 323, 533–536, <https://doi.org/10.1038/323533a0>, 1986.
- Saha, S., Sarkar, R., Thapa, G., and Roy, J.: Modeling gully erosion susceptibility in Phuentsholing, Bhutan using deep learning and basic machine learning algorithms, *Environmental Earth Sciences*, 80, 295, <https://doi.org/10.1007/s12665-021-09599-2>, 2021.
- Sahour, H., Gholami, V., Vazifedan, M., and Saeedi, S.: Machine learning applications for water-induced soil erosion modeling and mapping, *Soil and Tillage Research*, 211, 105032, <https://doi.org/10.1016/j.still.2021.105032>, 2021.
- 480 Sarkar, T. and Mishra, M.: Soil erosion susceptibility mapping with the application of logistic regression and artificial neural network, *Journal of Geovisualization and Spatial Analysis*, 2, 8, <https://doi.org/10.1007/s41651-018-0015-9>, 2018.
- Schmidt, J., Werner, M., and Michael, A.: Application of the EROSION 3D model to the CATSOP watershed, The Netherlands, *Catena*, 37, 449–456, [https://doi.org/10.1016/S0341-8162\(99\)00032-6](https://doi.org/10.1016/S0341-8162(99)00032-6), 1999.
- 485 Schober, P., Boer, C., and Schwarte, L. A.: Correlation coefficients: appropriate use and interpretation, *Anesthesia & analgesia*, 126, 1763–1768, <https://doi.org/10.1213/ANE.0000000000002864>, 2018.
- Schwertmann, U., Vogl, W., and Kainz, M.: *Bodenerosion durch wasser*, Ulmer Verlag, 64 p, 1987.

- Sharififar, A. and Sarmadian, F.: Coping with imbalanced data problem in digital mapping of soil classes, *European Journal of Soil Science*, 74, e13 368, <https://doi.org/10.1111/ejss.13368>, 2023.
- 490 Steinhoff, B., Bug, J., and Mosimann, T.: Einsatz eines mobilen GIS zur Kartierung von Bodenerosion durch Wasser, *Neue Horizonte für Geodateninfrastrukturen—Open GeoData, Mobility D*, 3, 27–32, 2013.
- Steinhoff-Knopp, B. and Burkhard, B.: Soil erosion by water in Northern Germany: long-term monitoring results from Lower Saxony, *Catena*, 165, 299–309, <https://doi.org/10.1016/j.catena.2018.02.017>, 2018.
- Taha, A. A. and Hanbury, A.: Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, *BMC medical imaging*, 495 15, 1–28, <https://doi.org/10.1186/s12880-015-0068-x>, 2015.
- Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Lengfeld, K., Walawender, E., Weigl, E., and Becker, A.: Radar-based gauge-adjusted one-hour precipitation sum climatology Version 2017.002: Gridded Precipitation Data for Germany (v2017.02), Dataset, https://doi.org/10.5676/DWD/RADKLIM_RW_V2017.002, produced by Deutscher Wetterdienst (DWD), 2018.
- Wischmeier, W. H. and Smith, D. D.: Predicting rainfall erosion losses: a guide to conservation planning, 537, Department of Agriculture, 500 Science and Education Administration, 1978.
- Wythoff, B. J.: Backpropagation neural networks: a tutorial, *Chemometrics and Intelligent Laboratory Systems*, 18, 115–155, [https://doi.org/10.1016/0169-7439\(93\)80052-J](https://doi.org/10.1016/0169-7439(93)80052-J), 1993.
- Yu, T. and Zhu, H.: Hyper-parameter optimization: A review of algorithms and applications, *arXiv preprint arXiv:2003.05689*, <https://doi.org/10.48550/arXiv.2003.05689>, 2020.
- 505 Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography, *Earth surface processes and landforms*, 12, 47–56, <https://doi.org/10.1002/esp.3290120107>, 1987.