

## Article

# Probabilistic Uncertainty Consideration in Regionalization and Prediction of Groundwater Nitrate Concentration

Divas Karimanzira

Fraunhofer Institute for Optronics, System Technique and Image Exploitation IOSB, Am Vogelherd 90, 98693 Ilmenau, Germany; divas.karimanzira@iosb-ast.fraunhofer.de

**Abstract:** In this study, we extend our previous work on a two-dimensional convolutional neural network (2DCNN) for spatial prediction of groundwater nitrate, focusing on improving uncertainty quantification. Our enhanced model incorporates a fully probabilistic Bayesian framework and a structure aimed at optimizing both specific value predictions and predictive intervals (PIs). We implemented the Prediction Interval Validation and Estimation Network based on Quality Definition (2DCNN-QD) to refine the accuracy of probabilistic predictions and reduce the width of the prediction intervals. Applied to a model region in Germany, our results demonstrate an 18% improvement in the prediction interval width. While traditional Bayesian CNN models may yield broader prediction intervals to adequately capture uncertainties, the 2DCNN-QD method prioritizes quality-driven interval optimization, resulting in narrower prediction intervals without sacrificing coverage probability. Notably, this approach is nonparametric, allowing it to be effectively utilized across a range of real-world scenarios.

**Keywords:** uncertainty analysis; spatial prediction; regionalization; convolutional neural networks; high-quality definition



**Citation:** Karimanzira, D. Probabilistic Uncertainty Consideration in Regionalization and Prediction of Groundwater Nitrate Concentration. *Knowledge* **2024**, *4*, 462–480. <https://doi.org/10.3390/knowledge4040025>

Academic Editor: Jose María Merigo

Received: 8 July 2024

Revised: 11 September 2024

Accepted: 22 September 2024

Published: 25 September 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Groundwater contamination by nitrate is a significant environmental concern worldwide, with implications for human health and ecosystem integrity. Nitrate, a compound commonly found in fertilizers, can leach into groundwater systems through various anthropogenic activities, particularly agricultural practices. This contamination can lead to serious health implications, including methemoglobinemia or “blue baby syndrome” in infants, and can also result in the degradation of aquatic ecosystems, leading to eutrophication and loss of biodiversity. Groundwater processes are quite complex, and the spatial prediction of water quality parameters such as nitrate concentration is associated with large uncertainty, including variability in measurement data and complex spatial and temporal dynamics, which have been shown in several previous studies, for example, in [1–4]. A study of a ten-year period of data showed large variations in groundwater nitrate distribution in the state of Baden Württemberg in Germany and highlighted the pervasive uncertainties associated with groundwater nitrate concentration prediction, emphasizing the need for robust probabilistic forecasting models to address these challenges [4–7] with an emphasis on the importance of looking at both the value prediction performance and the uncertainty of a model when estimating groundwater nitrate concentration. For example, with respect to groundwater nitrate requirements of the European Framework Directive, providing a PI for the nitrate concentration at a site (e.g., 32–46 mg/L with 99% certainty) is more valuable than a prediction for the specific value. In this way, risk can be reduced, and planning for measures can be carried out reliably.

Traditional methods such as Bayesian inference and bootstrapping have been commonly employed to generate probabilistic predictions and quantify uncertainties in groundwater nitrate concentration models. Recently, there have been several studies on tackling

uncertainties in modeling. For example, Ransom et al., (2017) used the boosted regression tree method for the prediction of groundwater nitrate concentration and considered uncertainties using bootstrapping with multiple model runs [1].

Koch et al., (2019) extended the random forest model with geostatistics to assess uncertainty [4]. Based on the work in [8,9], which applied a modified random forest model for quantile regression consideration of uncertainty, [3] used two different residual methods based on quantile regression and local errors and clustering (UNEEC for estimating uncertainties of a random forest model).

Bayesian neural networks, which learn a distribution over weights, are currently the state of the art for estimating predictive uncertainty; however, these require significant modifications to the training procedure and are computationally expensive compared with standard (non-Bayesian) neural networks. The work in [10] proposed an alternative to Bayesian neural networks that is simple to implement, readily parallelizable, and yields high-quality predictive uncertainty estimates by using ensembles and adversarial training.

The method of quality-driven PI (QD) developed in [11] is a dedicated prediction interval (PI) generation method based on an aggregation of ensemble PIs. In [12,13], the authors extended it to a version called QD-Plus, which adds additional terms for the value prediction. The QD-Plus method is based on a multiobjective loss function, which combines quality metrics for prediction intervals and point estimates. Furthermore, it adds a penalty function, which enforces the semantic integrity of the results and stabilizes the training process of the backbone model.

All these studies mainly follow two principles in their uncertainty considerations. Either they focus on the value prediction performance and evaluate uncertainty after the model has been set, or they optimize the prediction intervals and then take the middle value. They do not integrate the uncertainty consideration in the model building.

Specifically, this study aims to (1) modify the existing 2DCNN architecture [5] to incorporate Bayesian principles using TensorFlow Probability, enabling uncertainty quantification through the integration of Bayesian layers such as `tfp.layers.Convolution2DVariational` and `tfp.layers.DenseVariational`; (2) establish prior and posterior distributions: define prior and posterior distributions to represent initial beliefs about model parameters and update these beliefs based on observed groundwater nitrate concentrations, utilizing a specified likelihood function to relate observations to model predictions; (3) employ Bayesian inference techniques, such as variational inference or Markov Chain Monte Carlo (MCMC), to estimate the posterior distribution and utilize a loss function that combines negative log-likelihood and Kullback–Leibler (KL) divergence for optimal training; (4) modify the 2DCNN structure to estimate prediction intervals (PIs) and value predictions concurrently as in [14], maintaining the complexity of the model while addressing both epistemic and aleatoric uncertainties inherent in nitrate prediction; (5) optimize prediction intervals: aim to produce tighter prediction intervals while ensuring adequate prediction interval coverage probability (PICP), thereby enhancing the reliability of the model's predictions without compromising accuracy; and (6) compare the performance of the newly developed models against the previously established 2DCNN model in terms of specific value prediction accuracy and uncertainty quantification, assessing improvements in both dimensions. The main contributions of this work are the following:

- Extended Research: Building on our previous work with a two-dimensional convolutional neural network (2DCNN) for groundwater nitrate prediction.
- Focus on Uncertainty Quantification: The enhanced model incorporates a fully probabilistic Bayesian framework for improved uncertainty quantification.
- Introduction of 2DCNN-QD: Implemented the Prediction Interval Validation and Estimation Network based on Quality Definition (2DCNN-QD) to optimize probabilistic predictions and predictive intervals (PIs).
- Improved Prediction Interval Width: Achieved an 18% reduction in prediction interval width in a model region in Germany.

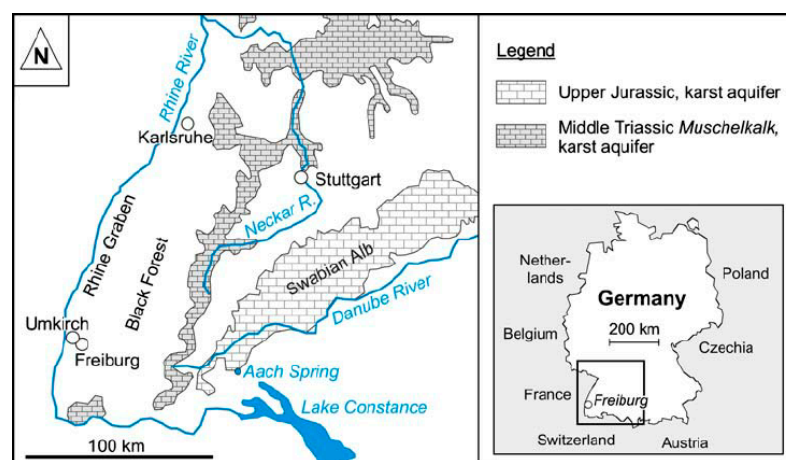
- Quality-Driven Optimization: Prioritized quality-driven interval optimization, resulting in narrower prediction intervals without compromising coverage probability.
- Non-Parametric Approach: The method is nonparametric, making it applicable across various real-world scenarios.

Section 1 introduces this study, highlighting the importance of probabilistic forecasting methods for predicting groundwater nitrate concentrations. Section 2 describes the study area and dataset, detailing the geographical location, hydrogeological characteristics, and groundwater nitrate measurement specifics. Section 3 outlines the methodology, focusing on the development of a 2D Convolutional Neural Network (2DCNN) model for probabilistic forecasting and the integration of uncertainty criteria. Section 4 presents the results and discussions, analyzing the 2DCNN model's outcomes, evaluating probabilistic forecasts, and comparing its performance with traditional methods like Bayesian inference and bootstrapping. Section 5 concludes with key findings, implications for groundwater management, recommendations for future research, and final remarks.

## 2. Materials and Methods

### 2.1. Description of the Study Area and Data Availability

Baden-Württemberg (BaWü) is a state located in southwestern Germany, situated east of the Rhine River and sharing a border with France (refer to Figure 1). Spanning approximately 35,752 square kilometers, the state is home to a population of around 11.07 million residents, as per the 2019 census data. Notably, Baden-Württemberg boasts one of the largest continuous forested areas in the country, known as the Black Forest, which extends towards the west. The region also features significant agricultural zones, particularly in the upper Rhine Valley, characterized by orchards and vineyards.



**Figure 1.** The study area of Baden-Württemberg (BaWü). Generalized hydrogeological map of Baden-Württemberg, southwest Germany, showing karst aquifer systems and locations [15].

The geological landscape of Baden-Württemberg is diverse, encompassing various aquifers. The western part of the state, bordering France, is marked by highly productive porous and fractured aquifers. In the southern region, there are karstified aquifers, while the central area is dominated by a less productive fractured aquifer. Annually, an estimated 479 million cubic meters of groundwater are extracted from these aquifers.

The state's environmental authority oversees a comprehensive groundwater monitoring network comprising approximately 2200 water quality measurement points distributed across the state. This network collects and evaluates data on groundwater quality and quantity on an annual basis. Stakeholders and the public can access this valuable information through the website of the Baden-Württemberg State Office for Environment, Measurements, and Nature Conservation (LUBW).

The predictors and the data used for the studies are summarized in Table 1. The nitrate concentration data was taken from the LUBW annual groundwater data catalog and measured from 1566 monitoring wells in the upper aquifer. The most recent measured values from the year 2019 were used for the study. Table 1 shows the data type, the resolution of the data, and its source. The main source of the data was the Federal Institute for Geosciences and Natural Resources (BGR) in Hanover ([https://www.bgr.bund.de/EN/Themen/Wasser/Produkte/produkte\\_node\\_en.html](https://www.bgr.bund.de/EN/Themen/Wasser/Produkte/produkte_node_en.html), accessed on 31 March 2023) and the German Federal Agency for Cartography and Geodesy (BKG). This provided most of the hydrogeological data, such as hydrological units, soil units, surface geology, etc.

Percolation rate is a primary covariate as it is the vehicle by which the nitrates are transported into the groundwater from the surface. Land use and cover are principal components of groundwater contamination with nitrates. It controls the contamination. It is obvious that groundwater in agricultural and urban regions is more polluted than groundwater in forest or bare regions. The map of the CORINE Land cover (Copernicus Land Monitoring Service) Dataset was applied for this purpose. Furthermore, to improve the data, it was combined with the land use maps from [16] and data from Sentinel 2 [17]. The aquifer and soil type determine how fast the nitrates can be transported through them. The data are also obtained from BGR. Furthermore, the organic matter content of the soil is also used as an explanatory variable. Another important variable is the distance to the nearest surface water body. This data was derived from the global river width and depth database, calculated by [18]. The crop type data, which indicate the amount of nitrate fertilizer needed, were obtained from national-scale crop type maps for Germany based on a combined time series of Sentinel-1, Sentinel-2, and Landsat 8 data (2017, 2018, and 2019) [19]. The information on nitrogen fertilizer load is closely related to land use. Farms, industries, and animal distribution statistics are very important covariates that can explain nitrate sources. Nitrogen fertilizer contributes significantly to high crop yields, but excess amounts end up in the groundwater (see the Nitrate Surplus Map). We used data from the LUBW, which represents the amount of nitrogen fertilizer nutrients applied in croplands for this purpose. The field capacity (FK) and the usable field capacity (nFK) are used to give the statistics of the amount of soil moisture held in the soil after excess water has been drained away and the rate of downward movement has decreased, as well as the part of the water the roots can extract from the soil with their suction forces.

**Table 1.** Summary of the data used as predictor and predicted variables.

Parameter	Description	Data Period	Res [m]	Reference
Hydrological units	Hydrogeological Map of Germany 1:250,000 (HÜK250). Hydrogeological characteristics of the upper continuous aquifers	-	30	[20]
Soil units	Soil Map of Germany 1:200,000 (BÜK200) Information on soil type, soil type, spruce rock at a scale 1:200,000	-	30	[21]
Percolation rates	Mean Annual Rate of Percolation from the Soil in Germany.		30	[22]
Land cover classes	CORINE Land Cover 2018, min mapping unit: 5 ha (CLC5), Germany.	2017–2018	30	[23]
Standardized soil units	Soil Map of the Federal Republic of Germany 1:1,000,000 (BÜL1000).	-	30	[24]
Surface geology	Geological Map of Germany 1:1,000,000 (GK1000)	-	30	[25]
Hydrological regions	Hydrogeological spatial structure of Germany (HYRAUM), regions with similar hydrogeological characteristics	-	30	[26]
Land cover	Land cover classification map of Germany's agricultural area based on Sentinel-2A data from 2016	2016	20	[16]

Table 1. Cont.

Parameter	Description	Data Period	Res [m]	Reference
Soil organic matter contents	Organic matter contents in topsoils of Germany 1:1,000,000 (HUMUS1000OB)	2007	30	[27]
Hydrogeological map	Hydrogeological Map of Germany 1:250,000 (HÜK250). Hydrogeological characteristics of the upper continuous aquifers in Germany at a scale of 1:250,000	-	100	[20]
Crop types	National-scale crop type maps for Germany from combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data (2017, 2018 and 2019)	2019	10	[19]
Land cover	Germany 2019—Land cover classification based on Sentinel-2 data	2019	10	[17]
NDVI index	MODIS/Terra Vegetation Indices 16-Day L3 Global 250 m SIN Grid (250 m 16 days NDVI)	2019	230	[28]
Stream distance	Distances to the nearest surface water body derived from the global river width and depth database, calculated from (Andreadis et al., 2013).	-	100	[18]
Nitrate	Nitrate concentration form the LUBW annual groundwater data catalog	2019	-	-

Some statistics of the groundwater nitrate concentration data collected from 1566 monitoring sites in the area of study are shown in Figure 2. As indicated by the skew of 1.5, the data distribution is quite imbalanced. Lower values of nitrate are more common compared with higher values above 50 mg/L, which makes it very difficult for learning methods.

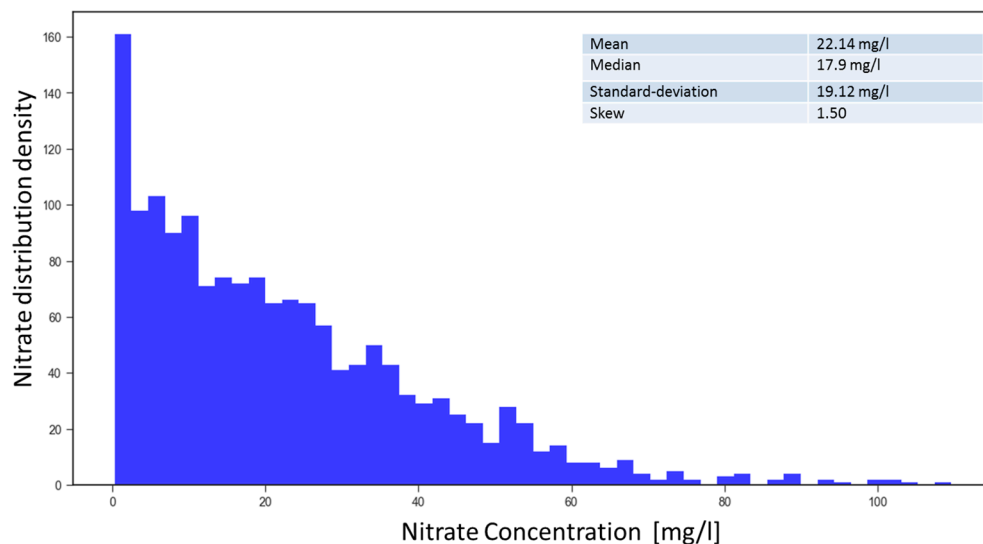
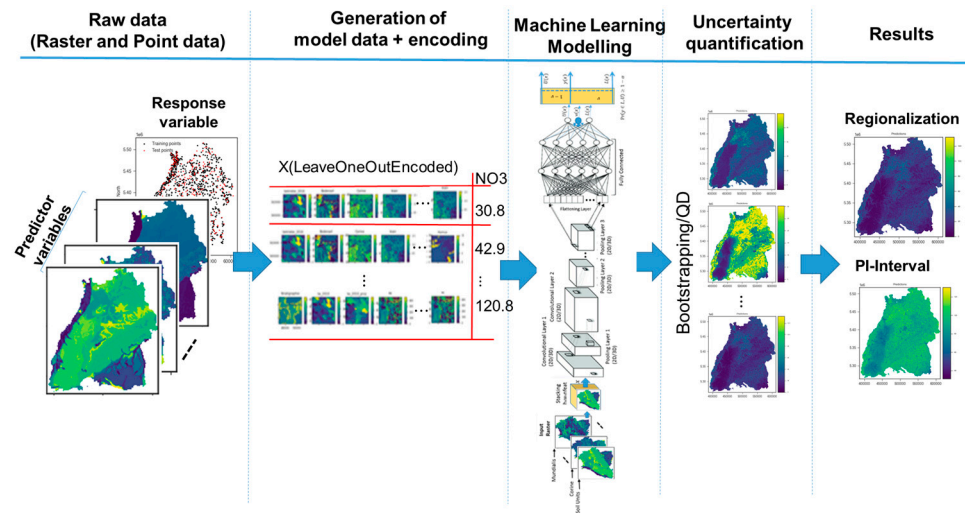


Figure 2. Frequency density of the nitrate from 2019 from 1566 measuring sites.

### 2.2. Methodology

Our methodological approach is illustrated in Figure 3. The process begins with Step 1: Data Preprocessing, where we generate machine learning datasets from raster data. This involves executing a spatial query to extract relevant values at the locations of nitrate observation points. If any of these values are categorical variables, they are appropriately encoded for use in the model.



**Figure 3.** Flowchart showing the methodology.

Next, in Step 2: Model Development, machine learning models are constructed, trained, and validated using the dataset prepared in Step 1. We employ a robust 10-fold cross-validation technique to ensure the reliability and generalizability of the model’s performance. Additionally, we can quantify uncertainties through an optional bootstrapping approach, which enhances our understanding of prediction variability.

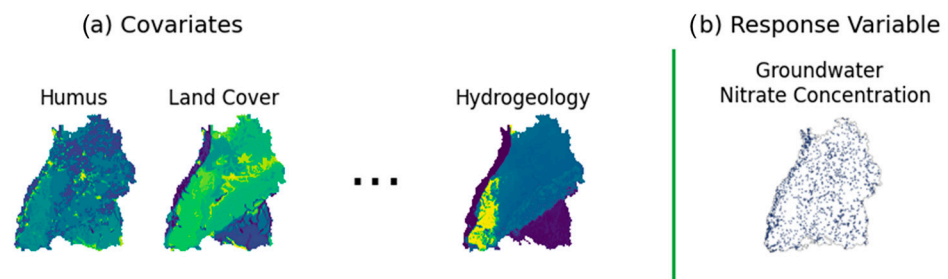
In Step 3: Feature Evaluation, we assess each observation point and calculate the importance of the features used in the model. This evaluation helps identify which input variables significantly influence nitrate concentration predictions.

Finally, in Step 4: Regionalization, we extrapolate the predicted nitrate concentrations across the entire model area. This is achieved by iterating over all grid cells of the input variables, allowing us to create a comprehensive spatial representation of nitrate concentrations throughout the region.

Each of these steps is elaborated upon in the following sections, providing a detailed understanding of our methodological framework.

### 2.2.1. Data Preprocessing

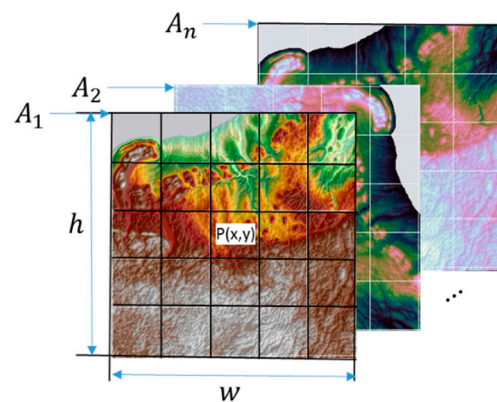
As described in Table 1, the data consist of raster files containing predictor variables and a point shape file representing groundwater nitrate levels, which served as the response variable, as illustrated in Figure 4. The initial phase involved preprocessing and preparing the data to align with the specifications of the 2DCNN model.



**Figure 4.** The data are provided in two main formats: (a) the predictor variables, represented as raster files containing covariate information, and (b) the target variable, which is the nitrate concentration, presented in the form of a point shapefile.

Convolutional neural networks work on images; hence, they automatically include information about the vicinity of  $(x, y)$  and fully leverage the spatial context of a nitrate observation point. To align the predictor variables with the response variable for training a

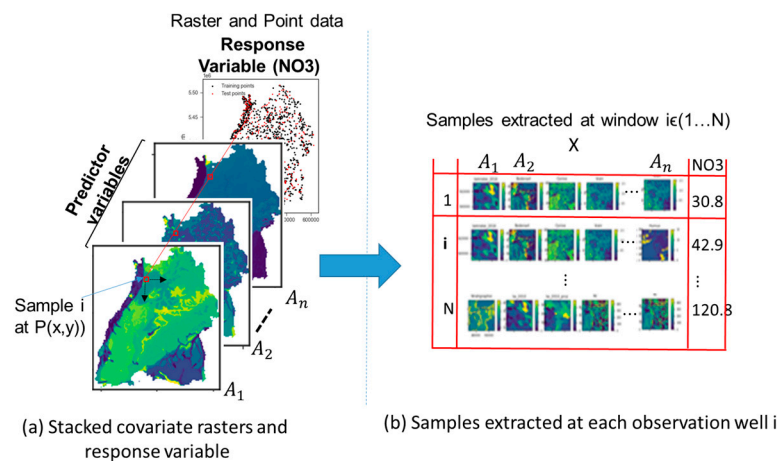
2D Convolutional Neural Network (2DCNN) model on groundwater nitrate prediction, the data preprocessing involves sampling raster patches around the points representing nitrate concentrations. This process included loading the raster files containing covariate information and the point shapefile indicating the nitrate levels, defining the patch size, and iterating over the points in the shapefile to extract raster patches centered around each point. Patches of size  $h \times w$  with  $P(x, y)$  coordinates as the center are extracted from the predictors rather than just the corresponding nitrate observation point pixel values (Figure 5). The extracted patches of the predictors ( $A_n$ ) are normalized to ensure consistent scaling, reshaped into a 4D ( $samples, h \times w \times n$ ) array format suitable for the 2DCNN model, and split into training and testing sets. By sampling raster patches of a specified size around the nitrate concentration points, the data is effectively prepared for training the 2DCNN model to predict groundwater nitrate levels based on spatially correlated covariates, facilitating the integration of spatial information into the predictive modeling process.



**Figure 5.** Spatial context in the vicinity of a nitrate observation well ( $P$ ) for  $n$  number of explanatory raster variables.

The area of the covariates influencing the occurrence of nitrate has a limited extent. In the context of groundwater nitrate prediction using a 2D Convolutional Neural Network (2DCNN), determining the influencing radius for sampling patches around nitrate measurement points is a crucial step in aligning spatial information with the predictive modeling process. This would, in turn, considerably reduce the size of the inputs supplied to the CNN model as well as the training time. There are many ways to find the zone of influence, e.g., variogramming and spatial correlation analysis, as illustrated in [29]. In this paper, the size of the window for cropping the explanatory raster images was estimated in two steps: first, using a variogram to find a rough estimate, and then setting different window sizes around the rough estimate as model hyperparameters and testing their effects on the prediction results using Bayesian optimization to obtain the best window size. By computing the variogram from the nitrate concentration data points, we gain insights into the spatial variability and correlation structure of the nitrate concentrations across the study area. Fitting a variogram model to the experimental variogram allows us to characterize the spatial correlation patterns and identify the range parameter, which signifies the distance beyond which nitrate concentrations are no longer spatially correlated. Leveraging this range parameter as a guide, we can define an influencing radius that encapsulates the spatial correlation of nitrate concentrations and dictates the distance within which patches should be sampled around each nitrate measurement point. Adhering to the influencing radius ensures that the patch size is appropriately adjusted to capture the spatial relationships inherent in the data, enabling the 2DCNN model to learn and leverage the spatial context for accurate nitrate prediction. This approach not only enhances the model's ability to capture spatial dependencies but also ensures that the sampling strategy aligns with the underlying spatial structure of the nitrate data, ultimately improving the model's predictive performance and the integration of spatial information into the modeling process.

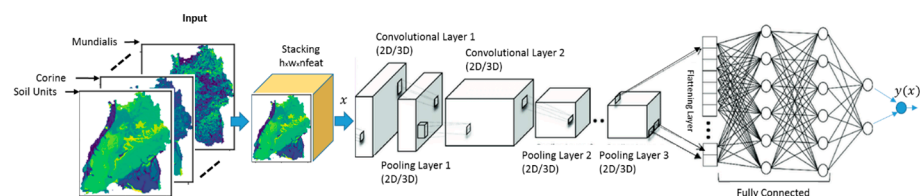
Therefore, for the spatial model, the original raster files of the covariates are cropped according to the zone around the observation point, as illustrated in Figure 6a, and put together along with the response variable to make a sample (Figure 6b). In this way, a dataset containing an array of cropped explanatory raster images with their corresponding response values was obtained. If we want to utilize all the observation points available, we will run into a problem at the borders with the points that are less than half the zone distance from the edges. Therefore, all the points that fulfilled this condition were removed from the dataset.



**Figure 6.** (a) Crops of raster images at the observation wells by a window of size  $h \times w$  and (b) cropped images representing  $n$  different input variables at the same point marked by the red square.

### 2.2.2. Backbone and Baseline Model Based on 2DCNN Architecture

In our previous study, a 2DCNN model for regionalization of groundwater nitrate concentration was developed. This model, shown in Figure 7, serves as the backbone CNN and the baseline model architecture for comparison. As described in [5], in the baseline mode, uncertainties were evaluated using the bootstrapping method.



**Figure 7.** The backbone neural network model based on unimodal 2DCNN architecture.

The 2DCNN model is a unimodal network based on 2D convolution [30]. In this architecture, the inputs are stacked together to form a multichannel image. We stacked our input raster images into channels to create a  $h \times w \times (M \text{ Channels})$  input structure, which goes into the convolutional layers. These initial layers are there to extract features for nitrate prediction. For these convolutional layers,  $32@3 \times 3$  filters were applied with no padding and a stride of 1 in the first convolutional layer with ReLU activation. As in most networks, this was followed by a  $2 \times 2$  max-pooling layer with a stride of 2 and a dropout layer with a dropout rate of 0.25. This was again fed into the second convolutional layer, where  $16@3 \times 3$  were used. Next, the outputs of the second convolutional layer are led through individual depth branches, flattened to a 1D array, and fed to three fully connected ReLU layers with  $n\_nodes$  neurons each. Finally, outputs are connected to a fully connected layer of size 1 with a linear activation function, which is consequently the final prediction for the target depth nitrate concentration.

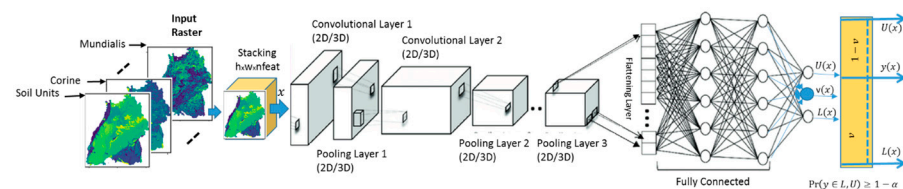
The final structure and hyperparameters of the 2DCNN were obtained using Bayesian optimization (BayesOpt) [31]. Training convolutional neural networks requires the spec-

ification of the neural network architecture, as well as options of the training algorithm, such as the learning rate, window size, and L2 regularization strength. Hyperparameter selection and tuning can be very difficult and time-consuming. Bayesian optimization is an algorithm well suited to optimizing hyperparameters of classification and regression models [31]. The advantage of Bayesian optimization is that it can be used to optimize nondifferentiable, discontinuous, and time-consuming functions. Besides the specification of the neural network architecture and deciding the options of the training algorithm, Bayesian optimization was also used to select the most important predictor variables.

As can be seen from the architecture, uncertainties are not considered in the modeling process. The resulting model is only evaluated for its uncertainty coverage. Prediction intervals (PI) are derived to determine the model uncertainty as follows: PIs are defined based on p-quantiles as the interval from the lower ( $PI_L = \frac{1-p}{2}$ ) to the upper limit  $PI_U = \frac{1+p}{2}$ ) of the predictions, in which the true value is expected with a high probability (p). For this model, the uncertainty is presented as  $p = 0.10$  prediction interval from several bootstrapping runs. The upper and the lower bounds ( $PI_L$ ) of the confidence band is computed as  $PI_L = \bar{x} \pm 1.645\sqrt{\sigma^2 + MSE}$ .  $\bar{x}$  is the mean and  $\sigma^2$  is the variance of the N-bootstrap runs, and MSE is the mean squared error of the fitted models. The bootstrapping procedure follows two steps: (1) computing a population of statistics, e.g., mean squared error, and then (2) calculating the confidence interval. A population of statistics was created by running the Bayesian optimization for hyperparameter search 100 times. Each time, a new model with different hyperparameters was found and its metrics (mse and variance) calculated. In the second step, the confidence interval was calculated using the resulting statistics.

### 2.2.3. Incorporation of a Quality Definition for Uncertainty

The structure of the 2DCNN model was modified with two components, as shown in Figure 8, to provide three outputs: one for the lower bound, the second one for the upper bound, and another one for the predicted variable.



**Figure 8.** The structure of the 2DCNN model modified for uncertainty consideration.

The first component is a chance constraint of the predicted variable  $y(x)$  with lower and upper bounds  $PI_L(x)$  and  $PI_U(x)$ , respectively, such that  $y(x)$  lies between the bounds with a probability higher or equal to a given confidence level  $1 - \alpha$ , as defined by Equation (1).

$$Pr(PI_L(x) \leq y(x) \leq PI_U(x)) \geq 1 - \alpha \tag{1}$$

The second component is a weighting factor  $v(x) \in (0, 1)$ , which provides a relative weighting of the value prediction  $y(x)$  between the lower and the upper boundaries and, in this way, enables us to compute  $y(x)$  as given in Equation (2). The advantage of expressing  $y(x)$  in terms of the upper  $U(x)$  and the lower  $L(x)$  bounds is that all the three values can be optimized jointly at the same time, and it is also ensured that the prediction value is always between the boundaries.

$$y(x) = v \cdot PI_U(x) + (1 - v) \cdot PI_L(x) \tag{2}$$

For training the 2DCNN-QD, a multiobjective minimization approach should be followed to optimize both the PIs and the predicted variable. The first objective stipulates that the PI should be as tight as possible while covering all prediction values. This can be achieved by integrating metrics that measure the quality of the PIs into the objective

function. For the evaluation of PIs, there are two important metrics: the mean prediction interval width (MPIW) Equation (3) and the prediction interval coverage probability (PICP). As given by Equation (3), the MPIW is computed as the mean of the differences in the PIs at each sample point. The PICP for  $p$ , as expressed in Equation (4), gives the proportion of observed values ( $y_i$ ) within the estimated PI [13].

$$MPIW = \frac{1}{n} \sum_{i=1}^n (PI_i^u - PI_i^l) \tag{3}$$

$$PICP = \frac{1}{n} \sum_{i=1}^n k_i, \text{ where } k_i = \begin{cases} 1 & (PI_i^u > y_i > PI_i^l) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

According to the high-quality (HQ) principle, PIs should minimize MPIW subject to  $PICP \geq (1 - \alpha)$ . With the weights and the parameters of the 2DCNN denoted by  $\theta$ , the objective function in this case can be expressed as in Equation (5)

$$\theta^* = \underset{\theta}{\operatorname{argmin}}(MPIW\theta) \text{ s.t } PICP\theta \geq 1 - \alpha \tag{5}$$

The discrete function  $k_i$ , where the gradient is always positive for all possible values, makes the optimization of Equation (5) very difficult and nonconvergent. Therefore,  $k$  is modified by smoothing factor  $s > 0$  and a sigmoid function  $\sigma$  to a smooth function  $k_s$  with form as expressed in Equation (9).

$$k_s = \sigma(s \cdot (y - PI_L)) \odot \sigma(s \cdot (PI_U - y))$$

To minimize MPIW, Equation (3) could simply be included in the loss function; however, PIs that fail to capture their data point should not be encouraged to shrink further. Therefore, a captured MPIW is introduced as the MPIW of only those points  $\sum_{i=1}^n k_i$  for which  $PI_L \leq PI \leq PI_U$  holds.

$$MPIW_{capt} = \frac{1}{\sum_{i=1}^n k_i} \sum_{i=1}^n (PI_i^u - PI_i^l) \cdot k_i \tag{6}$$

Remembering that a penalty should only occur in the case where  $PICP < (1 - \alpha)$  results in a one-sided loss. Combining with Equation (6) and adding a Lagrangian,  $\lambda$ , controlling the importance of width vs. coverage gives a new loss,

$$Loss_{PI} = MPIW_{capt,\theta} + \lambda \frac{n}{\alpha(1 - \alpha)} \max(0, (1 - \alpha) - PICP_{soft}\theta)^2 \tag{7}$$

The inclusion of both the continuous version of  $k$  in computing  $PICP_{soft}$  and discrete versions of  $k$  in the calculation of the  $MPIW_{capt}$  metric, respectively, enables the assignment of a zero score to points outside the prediction interval, while the continuous  $k$  produces continuous values that enable gradient calculations.

The second objective is to optimize the output  $v(x)$

$$L_v = \sum_{i=1}^n f(v_i \cdot PI_{U,i}(x) + (1 - v_i) \cdot PI_{L,i}(x), y_i) \tag{8}$$

where  $f$  represent the 2DCNN. Finally, the joint objective can be obtained from the two objective terms from Equations (10) and (11) to Equation (12). With  $\beta \in 0, 1$  as the control parameter, the importance of the individual objectives for PI-tightness quality and the quality of the predicted values can be controlled.

$$L_{un} = \beta L_{PI} + (1 - \beta) L_v \tag{9}$$

#### 2.2.4. Fully Probabilistic Bayesian CNN

To create a fully probabilistic Bayesian CNN model for groundwater nitrate concentration prediction using TensorFlow Probability (TFP), we replaced the standard layers in the 2DCNN architecture with Bayesian layers from TensorFlow Probability. TFP provides Bayesian variants of common neural network layers, such as ‘tfp.layers. Convolution2Dvariational’ for convolutional layers, and ‘tfp.layers. DenseVariational’ for dense layers. Prior and posterior distributions are defined to represent initial beliefs about model parameters and updated beliefs post data observation, respectively. A likelihood function is specified to relate observed groundwater nitrate concentrations to model predictions. Bayesian inference techniques like variational inference or MCMC are employed to estimate the posterior distribution. A loss function combining negative log-likelihood and KL divergence guides training to find optimal posterior distributions. The model is trained and validated on separate datasets, with uncertainty quantification facilitated by posterior distributions to generate prediction intervals.

The output of the model is represented as a distribution object, specifically a OneHot-Categorical distribution. This choice of output distribution allows for the direct use of the negative-log-likelihood as the loss function during training. By modeling the output as a distribution object, the model can inherently capture aleatoric uncertainty, which relates to the inherent randomness or variability in the data that cannot be reduced even with additional information.

Furthermore, in a BNN model, each weight and bias parameter is associated with both a mean and a variance, reflecting the uncertainty in the parameter estimates. Through techniques like Bayes-by-Backprop, the model learns these mean and variance values, enabling the representation of epistemic uncertainty, which pertains to uncertainty arising from a lack of knowledge about the true model structure or parameters.

It is worth noting that the total number of parameters in a BNN is typically double that of a standard convolutional neural network (CNN) model. This increase in parameters is attributed to the fact that each weight and bias parameter in the BNN now has both a mean and a variance, thereby doubling the parameter count. This capability to capture both aleatoric and epistemic uncertainty distinguishes the Fully Probabilistic Bayesian Convolutional Neural Network model. By incorporating aleatoric uncertainty related to data variability and epistemic uncertainty associated with model parameter uncertainty, the model provides a comprehensive representation of uncertainty, making it a powerful tool for probabilistic forecasting and decision making in groundwater quality management.

#### 2.2.5. Evaluation Metrics

Comparing the results of the Bayesian 2DCNN model to deterministic method (2DCNN) and Prediction Interval Validation and Estimation Network with Quality Definition (2DCNN-QD) involves assessing metrics such as Continuous Ranked Probability Score (CRPS), mean prediction interval width (MPIW), and prediction interval coverage probability (PICP).

CRPS (Equation (10)) measures the accuracy of probabilistic forecasts, with lower values indicating better performance in capturing the predictive distribution. The Bayesian CNN model is expected to have a higher CRPS compared with deterministic methods due to its probabilistic nature and consideration of uncertainties.

$$CRPS = \frac{1}{199N} \sum_{m=1}^N \sum_{n=-99}^{99} (P(y \leq n) - H(n - Y_m))^2 \quad (10)$$

MPIW (Equation (3)) quantifies the width of the prediction intervals provided by the model. Bayesian CNN models typically result in wider prediction intervals compared with deterministic methods, reflecting the model’s uncertainty-aware nature.

PICP (Equation (4)) assesses the coverage probability of prediction intervals, with higher values indicating better coverage. Bayesian CNN models tend to exhibit higher

PICP values compared with deterministic methods, as they aim to provide reliable and well-calibrated uncertainty estimates.

#### 2.2.6. Dealing with an Imbalanced Dataset

To address the issue of class imbalance in the dataset of groundwater nitrate concentration, the Synthetic Minority Oversampling Technique for Regression (SMOGR) was utilized to create a balanced dataset. SMOGR is an extension of the SMOTE (Synthetic Minority Oversampling Technique) method, specifically designed for regression tasks. In the context of groundwater nitrate concentration prediction, the imbalance in the dataset refers to the unequal distribution of samples across different concentration levels. For instance, there are fewer samples representing high nitrate concentrations compared with low or moderate concentrations. SMOGR works by oversampling the minority class (in this case, high nitrate concentrations) and undersampling the majority class to create a balanced dataset. Unlike SMOTE, which focuses on creating synthetic examples along the line segments connecting minority class instances, SMOGR generates synthetic examples using a regression model to capture the underlying distribution of the data. Hereby SMOGR fit a regression model to the original dataset to capture the relationship between input features and the target variable (nitrate concentration). The regression model is used to calculate the residuals (the differences between the predicted and actual nitrate concentrations) for each sample in the dataset. Synthetic samples are generated for the minority class by perturbing the original samples in the direction of their residuals. This process creates new synthetic instances that reflect the distribution of the minority class more accurately. The majority class is undersampled to reduce the imbalance in the dataset while maintaining its overall distribution. By combining the original and synthetic samples, a balanced dataset is created with an equal representation of different nitrate concentration levels.

The advantages of SMOGR are that it addresses the class imbalance issue in regression tasks by generating synthetic examples that reflect the distribution of the minority class more accurately. The use of regression modeling in SMOGR allows for a more nuanced approach to oversampling, considering the underlying relationships in the data. Once the balanced dataset was created using SMOGR, it was used to train the models (2DCNN, Bayesian 2DCNN, and 2DCNN-QD) for groundwater nitrate concentration prediction. The performance of the models was evaluated based on standard regression metrics to assess the effectiveness of balancing the dataset using SMOGR.

#### 2.2.7. Experimental Setup

For the experiments, the models were implemented in Python 3.10 with Kera and TensorFlow v2.0. A CPU computer was used for both training and inference. The three models (2DCNN, Bayesian 2DCNN, and 2DCNN-QD) were trained, validated, and tested using known nitrate concentrations at monitoring sites. The latest available measured values at the monitoring sites from 2019 were used for this purpose. Firstly, all the nitrate data were cleaned, removing outliers, and then preprocessed as described in Section 2.2.1 for CNN conformity. In total, 1824 data samples were created. The categorical variables were target encoded using the leave-one-out encoder.

Bayesian optimization was used for tuning and selection of the hyperparameters of the CNN models. Tenfold cross-validation was used during the Bayesian optimization, where the data were randomly partitioned into 10 subsets. The model fitted to the remaining 9 subsets is then validated using each subset in turn. In this way, models with better generalizability could be established. For each CNN model type, the optimization parameters included window size, input features, batch size, learning rate, number of nodes in the layers, and number of layers.

The mean absolute error (MAE) Equation (11), the mean squared error (MSE) Equation (12), the coefficient of determination ( $R^2$ ) Equation (13), and the model bias Equation (14) were used as metrics to evaluate model specific value prediction performance. The MAE gives a very good idea of the prediction accuracy, but it does not show whether or not the model

tends to overestimate or underestimate the predictions. This is where the bias comes into play. It allows for the evaluation of prediction accuracy as well as whether the model tends to overestimate or underestimate the values of the variable of interest. The better the prediction, the closer the bias is to zero. It should be noted that the bias does not account for the variability of the predictions. On this issue, a useful metric is the MSE. It provides an indication regarding the dispersion or variability of the prediction accuracy. The  $R^2$  gives us the proportion of the variance in the dependent variable that is predictable from the independent variables and indicates the covariance in the model's prediction. Therefore, these metrics were used in combination to evaluate the performance in terms of accuracy of the models in this paper.

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \tag{11}$$

$$MSE = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \tag{13}$$

$$Bias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{n} \tag{14}$$

where  $n$  is the number of observations,  $y_i$  is the value of the  $i$ th observation in the validation/test dataset,  $\bar{y}_i$  is the mean value of the validation/test dataset, and  $\hat{y}_i$  is the predicted value for the  $i$ th observation.

For the performance towards uncertainty, the quality of the PIs produced by the models were evaluated using the mean prediction interval width (MPIW), the prediction interval coverage probability (PICP), and CRPS. These metrics are described by Equations (3), (4), and (10), respectively.

### 3. Results

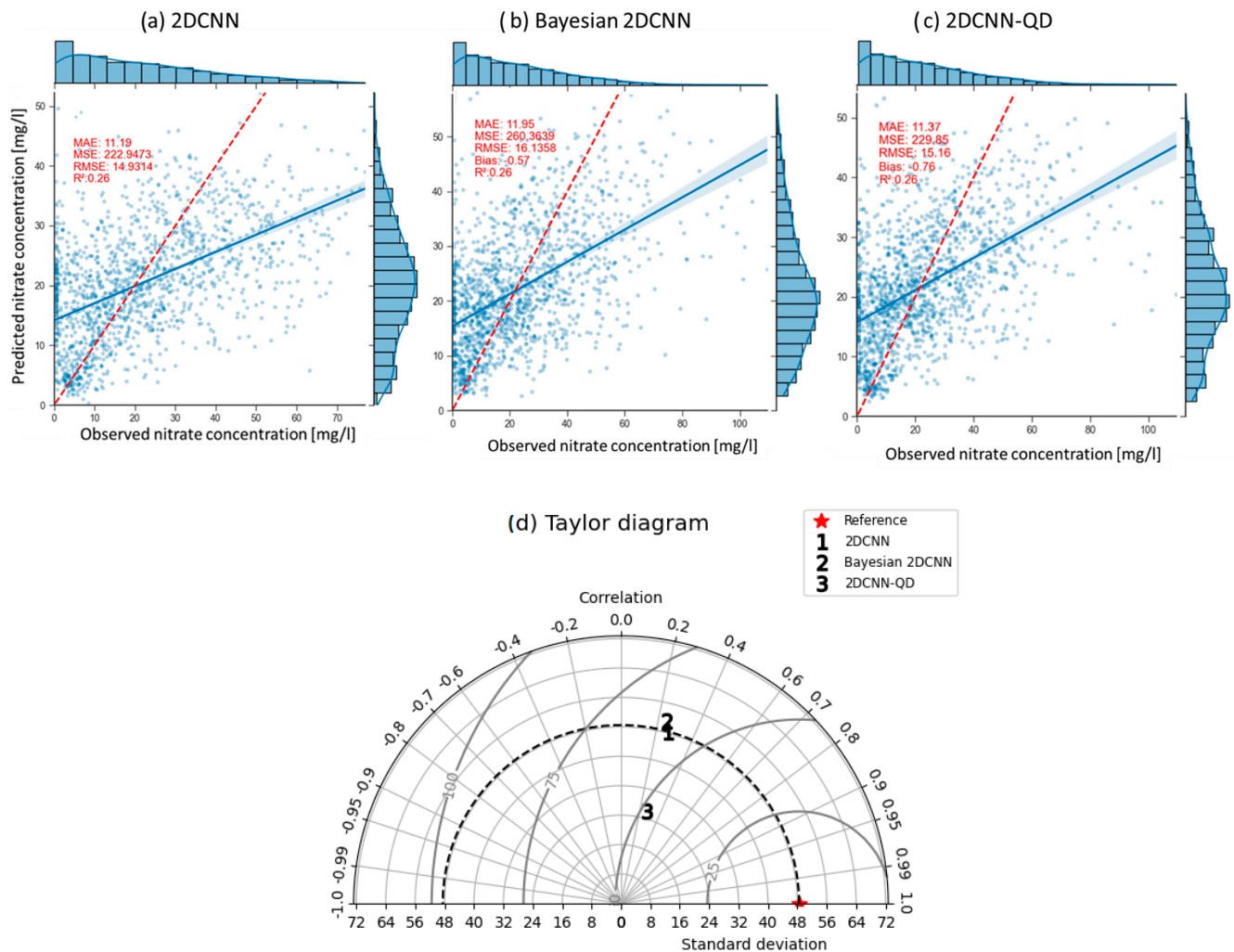
The models were evaluated on their performance based on six metrics (MAE, MSE, R2, Bias, CRPS, PICP, and MPIW) after the Bayesian optimization with 10-fold cross-validation, and the results are shown in Table 2. The training and inference times of the different models are also listed in the table as a proxy for complexity of the model.

**Table 2.** Results of the models for accuracy and quantification of uncertainty. Best values are shown in bold.

Model	2DCNN	Bayesian 2DCNN	2DCNN-QD
MAE (std) (mg/L)	<b>11.19 (0.43)</b>	11.95 (0.21)	11.37 (0.36)
MSE	<b>222.95</b>	260.36	229.85
RMSE	<b>14.93</b>	16.13	15.16
$R^2$	<b>0.26</b>	0.24	<b>0.26</b>
Bias	-0.76	-0.57	-0.43
CRPS	0.10	0.15	0.12
PICP	0.87	<b>0.94</b>	0.92
MPIW (mg/L)	48.3	51	<b>26.1</b>

Scatterplots and a Taylor diagram of the nitrate predictions by the models are shown in Figure 9a–d. Scatterplots provide insight into the degree of fitness of the models, the model bias, and the variance of predictions shown in Table 2. The cloud of all the models' predictions shows scatter, especially in the high values, which is consistent with the negative biases of the models of  $-0.76$  mg/L and  $-0.43$  mg/L, respectively (Table 2). The 2DCNN produces better accuracy results in terms of the MAE, MSE, and RMSE of 11.19 mg/L, 222.95, and 14.93, respectively. The 2DCNN-QD clearly has the best behavior towards uncertainties with an MPIW of about 26 mg/L compared with the 48.3 mg/L of the 2DCNN

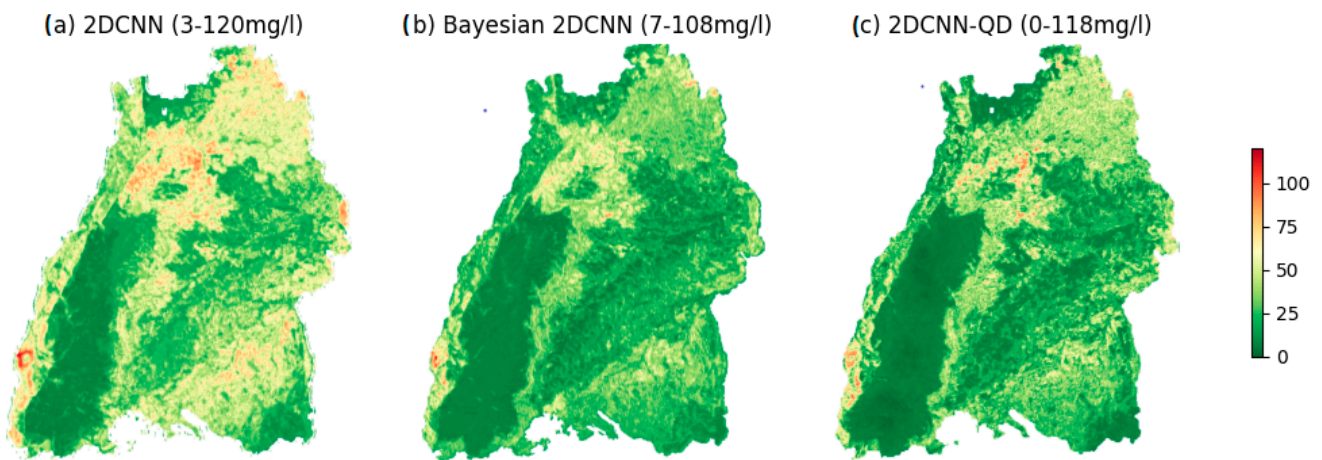
and 51 mg/L of the Bayesian 2DCNN. Figure 9d shows a better comparison of the models using a Taylor diagram.



**Figure 9.** Measured vs. predicted groundwater nitrate concentrations for (a) 2DCNN, (b) Bayesian 2DCNN and (c) 2DCNN-QD models, and (d) Taylor diagram. The red line represents the ideal predictions and the blue lines the regression line of the point cloud.

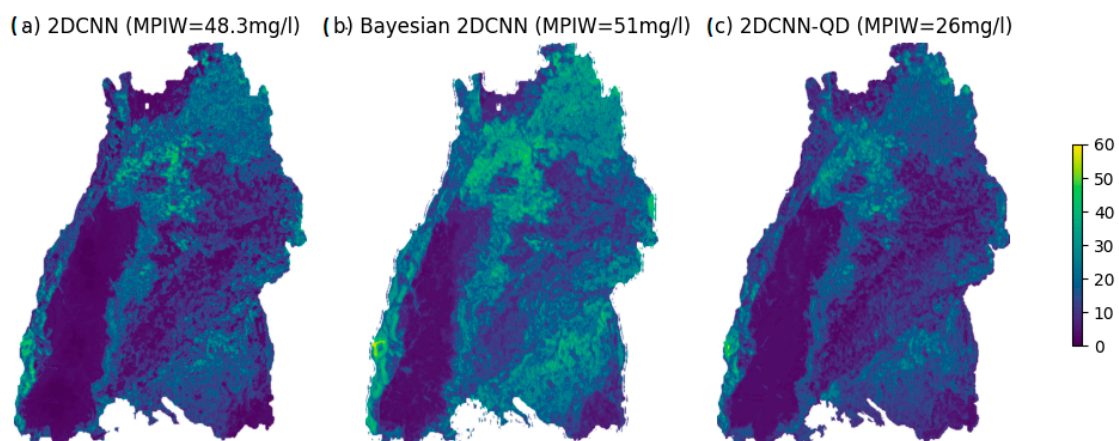
After training and validation, the models were used to predict the nitrate concentration for the whole grid, including unknown regions. A sliding window of the size determined by the Bayesian optimization algorithm is run through all the raster cells to produce a target dataset, with each sample containing the eight explanatory rasters identified by the Bayesian optimization according to their importance. In this way, a map of nitrate distribution can be produced by regionalization. The regionalization results of the three models, 2DCNN, Bayesian 2DCNN, and 2DCNN-QD, are shown as raster maps in Figure 10a–c. It can be seen that all the methods produce plausible results for regionalization with slight differences. If the sampling locations are overlaid on the prediction surface, the spatial pattern of the observed nitrate concentration concurs very well with the predictions. Higher nitrate concentrations are identified in the northern region, where most of the agricultural activities are conducted, and in some regions with porous aquifers. The lowest groundwater nitrate concentration is identified appropriately in the regions with karstified aquitards, e.g., in the Black Forest region, where fewer human influences occur. High nitrate concentrations can also be found in the middle Neckar-Taube-Gäuplatten, near the southern edge of the Black Forest, in the northern part of the Swabian Keuper-Lias Plains, in the middle of the Donau-Iller-Platte, the bordering northern part of the Voralpien Huegel and Moorland,

in the southern and northern parts of the Oberrhein-Tieflands, and in the north of the Odenwaldes. These areas are also mentioned as having high nitrate concentrations in the LfU report 2001.



**Figure 10.** Regionalization results of the three models, (a) 2DCNN, (b) Bayesian 2DCNN, and (c) 2DCNN-QD.

Figure 11a–c shows the uncertainty band of the 2DCNN, Bayesian 2DCNN, and the 2DCNN-QD models. It can be seen that the uncertainty band of the 2DCNN and the Bayesian 2DCNN is quite wide compared with that of the 2DCNN-QD model. The 2DCNN-QD method is able to produce both PIs with tighter uncertainty bounds and specific value predictions using a loss function that expresses the value prediction as a function of the upper and lower bounds. The method guarantees the best values of the two metrics, PICP and MPIW. With the method, it is assured that the prediction interval will cover the true value; hence, the higher the PICP, the better. However, the second metric, MPIW, is required to measure the average size of all prediction intervals. This is because only using PICP is not enough, since 100% PICP can be achieved by setting the prediction interval to infinity. Therefore, the second metric is the mean prediction interval width (MPIW). In general, we want the model to have a high PICP while maintaining a low MPIW.



**Figure 11.** Uncertainty band of the regionalization by the two developed methods: (a) obtained through bootstrapping of 2DCNN and (b) Bayesian and (c) prediction intervals of the deep learning method, which produces both PIs and specific value predictions (2DCNN-QD) with MPIWs of 48, 59, and 26 mg/L, respectively.

#### 4. Discussion

All three models exhibit similar behavior in terms of observed patterns in relation to hydrogeology as follows:

**Nitrate Concentration and Agricultural Activity:** The models indicate higher nitrate concentrations in the northern regions of Baden-Württemberg, which correlates well with areas of intensive agricultural activity. This is consistent with hydrogeological principles, as agricultural practices, especially the application of fertilizers, significantly contribute to nitrate leaching into groundwater. The northern regions, characterized by porous aquifers, facilitate this leaching process, allowing nitrates to infiltrate and accumulate in the groundwater.

**Influence of Aquifer Types:** The presence of porous aquifers in certain areas implies the potential for higher nitrate concentrations. These aquifers can transmit water and contaminants more efficiently than less permeable formations, leading to elevated nitrate levels. The spatial distribution of high nitrate concentrations in regions such as the middle Neckar-Taube-Gäuplatten and northern parts of the Swabian Keuper-Lias Plains reflects the underlying hydrogeological conditions that promote nitrate mobility.

**Karstified Aquitards and Low Nitrate Concentrations:** Conversely, the lowest nitrate concentrations are identified in karstified aquitards, particularly in the Black Forest region. The karst systems, with their complex geological formations, often have reduced permeability, limiting the movement of contaminants into the groundwater. This finding underscores the protective role that certain hydrogeological features play in mitigating nitrate pollution, as these areas experience fewer human influences and agricultural runoff.

**Consistency with Previous Reports:** The observed patterns align with findings from the LfU report (2001), which also identified high nitrate concentrations in similar regions. This consistency reinforces the reliability of the models and their ability to reflect historical and ongoing trends in groundwater quality.

**Geographical Variability:** The model results highlight geographical variability in nitrate concentrations, with specific hotspots identified in areas such as the Donau-Iller-Platte and the northern parts of the Oberrhein-Tieflands. These regions, characterized by a combination of agricultural land use and favorable hydrogeological conditions, are critical for understanding the dynamics of nitrate contamination in groundwater.

In terms of uncertainty estimation, 2DCNN-QD demonstrates a lower mean prediction interval width (MPIW) of 26.1 compared with Bayesian CNN (MPIW: 59.0) and the deterministic method (MPIW: 48). The narrower prediction intervals produced by 2DCNN-QD are attributed to its innovative Quality Definition mechanism, which helps refine and optimize the prediction intervals based on data quality and model performance.

2DCNN-QD incorporates Quality Definition to enhance the accuracy and precision of the prediction intervals, ensuring that the intervals are tighter while maintaining high coverage probability (PICP: 0.92). By leveraging Quality Definition, 2DCNN-QD effectively adjusts the width of the prediction intervals to reflect the quality of the data and the model's performance, resulting in more informative and reliable uncertainty estimates.

While Bayesian CNN models may exhibit wider prediction intervals to capture uncertainties comprehensively, 2DCNN-QD's emphasis on quality-driven interval optimization allows it to provide narrower prediction intervals without compromising coverage probability. This characteristic makes 2DCNN-QD a valuable tool for applications where precise uncertainty quantification and narrower prediction intervals are essential for decision-making and risk assessment in groundwater quality management.

The unique feature of 2DCNN-QD in producing narrower prediction intervals while maintaining high coverage probability showcases its potential to offer accurate and reliable probabilistic forecasts that balance uncertainty quantification and prediction precision. This ability to tailor prediction intervals based on data quality and model performance highlights the effectiveness of Quality Definition in enhancing the performance of probabilistic forecasting models like 2DCNN-QD in groundwater nitrate concentration prediction tasks.

The findings suggest that 2DCNN-QD represents a significant advancement in probabilistic forecasting for groundwater nitrate concentration, offering a unique approach that leverages Quality Definition to optimize prediction intervals. By outperforming conventional methods in terms of interval width and coverage probability, 2DCNN-QD showcases its potential to enhance decision-making processes, risk assessment strategies, and environmental management practices in the context of groundwater quality monitoring and management.

Future work includes enhanced model integration, i.e., exploring the integration of multiple models, such as combining 2DCNN-QD with Bayesian CNN or ensemble methods, to leverage the strengths of different approaches and improve prediction accuracy and uncertainty quantification; investigating the incorporation of additional data sources, such as hydrogeological parameters, climate data, or land use information, to enhance the predictive power of models and capture more complex relationships in groundwater nitrate concentration prediction; developing dynamic models that can adapt to changing environmental conditions and incorporate real-time data for adaptive forecasting of groundwater nitrate concentrations, enabling more responsive and accurate predictions; and conducting spatial and temporal analyses of groundwater nitrate concentrations to identify trends, hotspots, and seasonal variations, providing valuable insights for targeted monitoring and mitigation strategies.

## 5. Conclusions

Nitrate prediction is associated with uncertainties, as several previous studies have shown. The key to applying deep learning methods to solving regression problems in such domains is improving their robustness. In a previous study, a 2DCNN model was developed for nitrate prediction. Despite having impressive prediction value accuracy in terms of MAE, RMSE, etc., the 2DCNN model produces poor uncertainty estimates, shown using the bootstrapping method. Since overly confident yet incorrect predictions may be harmful, precise uncertainty quantification is integral for practical applications of such networks. Therefore, in this paper, the 2DCNN model from previous studies was extended to incorporate uncertainty considerations in the modeling and quantify uncertainty using prediction intervals (PIs) and quality definition. Quantifying uncertainty offers several benefits for water managers, such as a reduction in risk and the ability to plan in a more reliable manner.

This study has provided valuable insights into the comparative performance of the Prediction Interval Validation and Estimation Network with Quality Definition (2DCNN-QD) and a Bayesian Convolutional Neural Network (CNN) model in the probabilistic prediction of groundwater nitrate concentration. A loss function applied to the 2DCNN-QD model combines three objectives: (1) Maximizing the coverage (number of observations within the lower and upper prediction intervals), which should be approximately  $1-\alpha$ , where  $\alpha$  is the desired significance level; (2) minimizing the PI; and (3) minimizing the squared error of the point predictions. The evaluation metrics, including mean prediction interval width (MPIW), prediction interval coverage probability (PICP), and Continuous Ranked Probability Score (CRPS), have shed light on the strengths and limitations of both models in capturing uncertainties and improving prediction accuracy.

Overall, the results indicate that 2DCNN-QD excels in producing narrower prediction intervals while maintaining high coverage probability, thanks to its innovative Quality Definition mechanism. On the other hand, the Bayesian CNN model showcases robust uncertainty quantification capabilities, providing reliable probabilistic forecasts with broader prediction intervals.

In practical applications of groundwater nitrate prediction, the choice between 2DCNN-QD and the Bayesian CNN model depends on the specific requirements of the task. 2DCNN-QD's ability to tailor prediction intervals based on data quality and model performance makes it suitable for scenarios where precise uncertainty quantification and narrower prediction intervals are crucial. Conversely, the Bayesian CNN model's comprehensive un-

certainty representation and probabilistic forecasts are advantageous for decision-making processes that require a thorough understanding of uncertainties.

Future research directions could explore hybrid approaches that leverage the strengths of both 2DCNN-QD and Bayesian CNN models to enhance prediction accuracy and uncertainty quantification further. Additionally, the integration of additional data sources, such as hydrogeological parameters and climatic variables, could improve the performance of probabilistic forecasting models in groundwater quality management.

This study contributes to advancing the field of probabilistic groundwater nitrate prediction by evaluating state-of-the-art models and providing insights into their comparative performance. The findings offer valuable guidance for researchers and practitioners seeking to improve probabilistic forecasting accuracy and uncertainty quantification in groundwater quality monitoring and management.

**Funding:** The work is part of the Project “NiMo”—Nitrate Monitoring 4.0 Intelligent Systems for sustainable reduction of nitrates in groundwater, which is funded by the Federal Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection (67ki2048).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Due to their proprietary nature, supporting data cannot be made openly available. Further information about the data and conditions for access are available upon request from the author.

**Acknowledgments:** I acknowledge the contributions of all colleagues, institutions, or agencies that aided the efforts of the authors, especially the Landesanstalt für Umwelt Baden-Württemberg (LUBW) for providing us with data on nitrate concentration and agricultural land use information. Additionally, I extend my appreciation to the reviewers for their insightful comments, constructive feedback, and valuable suggestions that have helped strengthen the manuscript and improve its clarity and coherence. Their expertise and thorough review were instrumental in refining the content and ensuring the credibility of the research outcomes.

**Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

## References

1. Ransom, K.M. A hybrid machine learning model to predict and visualize nitrate concentration throughout the Central Valley aquifer, California, USA. *Sci. Total Environ.* **2017**, *601–602*, 1160–1172. [[CrossRef](#)] [[PubMed](#)]
2. Knoll, L.; Häußermann, U.; Breuer, L.; Bach, M. Spatial Distribution of Integrated Nitrate Reduction across the Unsaturated Zone and the Groundwater Body in Germany. *Water* **2020**, *12*, 2456. [[CrossRef](#)]
3. Rahmati, O. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods. *Sci. Total Environ.* **2019**, *688*, 855–866. [[CrossRef](#)] [[PubMed](#)]
4. Koch, J.; Stisen, S.; Refsgaard, J.C.; Ernsten, V.; Jakobsen, P.R.; Højberg, A.L. Modeling depth of the redox interface at high resolution at national scale using Random Forest and Residual Gaussian simulation. *Water Resour. Res.* **2019**, *55*, 1451–1469. [[CrossRef](#)]
5. Divas, K.; Jonas, W.; Andreas, W.; Linda, R.; Tanja, L.; Marc, O. Application of machine learning and deep neural networks for spatial prediction of groundwater nitrate concentration to improve land use management practices. *Front. Water* **2023**, *5*, 1193142, ISSN 2624-9375. [[CrossRef](#)]
6. Zhang, Y.; Zhao, Y.; Liu, Y. Machine Learning-Based Nitrate Concentration Prediction in Groundwater: A Case Study in Northern China. *J. Hydrol.* **2021**, *597*, 125864.
7. Refsgaard, J.C.; van der Sluijs, J.P.; Højberg, A.L.; Vanrolleghem, P.A. Uncertainty in the environmental modelling process—A framework and guidance. *Environ. Model. Softw.* **2007**, *22*, 1543–1556. [[CrossRef](#)]
8. Meinshausen, N. QuantregForest: Quantile Regression Forests. 2017. Available online: <https://CRAN.R-project.org/package=quantregForest> (accessed on 18 June 2024).
9. Knoll, L.; Breuer, L.; Bach, M. Large scale prediction of groundwater nitrate concentrations from spatial data using machine learning. *Sci. Total Environ.* **2019**, *668*, 1317–1327. [[CrossRef](#)] [[PubMed](#)]
10. Lakshminarayanan, B.; Pritzel, A.; Blundell, C. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2017; pp. 6402–6413.

11. Pearce, T.; Zaki, M.; Brintrup, A.; Neely, A. High-quality prediction intervals for deep learning: A distribution-free, ensembled approach. *arXiv* **2018**, arXiv:1802.07167.
12. Salem, T.S.; Langseth, H.; Ramampiaro, H. Prediction intervals: Split normal mixture from quality-driven deep ensembles. In Proceedings of the Conference on Uncertainty in Artificial Intelligence, Virtual, 29 July–1 August 2020; pp. 1179–1187.
13. Tagasovska, N.; Lopez-Paz, D. Single-model uncertainties for deep learning. In Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; pp. 6414–6425.
14. Simhayev, E.; Katz, G.; Rokach, L. PIVEN: A Deep Neural Network for Prediction Intervals with Specific Value Prediction. *arXiv* **2020**, arXiv:2006.05139.
15. Goldscheider, N.; Goeppert, N. A conversation with Werner Käss (Germany) about his contributions to tracer hydrogeology and characterisation of mineral waters and spas. *Hydrogeol. J.* **2009**, *17*, 1543–1546. [[CrossRef](#)]
16. Preidl, S.; Lange, M.; Doktor, D. Land cover classification map of Germany's agricultural area based on Sentinel-2A data from 2016. *PANGAEA* **2020**. [[CrossRef](#)]
17. Riembauer, G.; Weinmann, A.; Xu, S.; Neteler, M. Germany-wide sentinel-2 based land cover classification and change detection for settlement and infrastructure monitoring. In *Proceedings of the 2021 Conference on Big Data from Space—18–20 May 2021*; Soille, P., Loekken, S., Albani, S., Eds.; Publications Office of the European Union: Luxembourg, 2021. [[CrossRef](#)]
18. Andreadis, K.; Schumann, G.; Pavelsky, T. A simple global river bankfull width and depth database. *Water Resour. Res.* **2013**, *49*, 7164–7168. [[CrossRef](#)]
19. Lukas, B.; Marcel, S.; Dirk, P.; Claas, N.; Stefan, E.; Patrick, H. National-Scale Crop Type Maps for Germany from Combined Time Series of Sentinel-1, Sentinel-2 and Landsat 8 Data (2017, 2018 and 2019). 2021. Available online: <https://zenodo.org/record/5153047#.YWFyXn1CREZ> (accessed on 4 June 2024).
20. BGR; SGD. Hydrogeological Map of Germany 1:250,000 (HÜK250). Federal Institute for Geosciences and Natural Resources (BGR) and German State Geological Surveys (SGD), Hannover. 2019. Available online: <https://www.bgr.bund.de/huek200> (accessed on 21 September 2022).
21. BGR; SGD. Soil Map of Germany 1:200,000 (BÜK200), Federal Institute for Geosciences and Natural Resources (BGR) and German State Geological Surveys (SGD), Hannover. 2020. Available online: <https://www.bgr.bund.de/buek200> (accessed on 21 August 2022).
22. BGR; SGD. Mean Annual Rate of Percolation from the Soil in Germany (SWR1000), Federal Institute for Geosciences and Natural Resources (BGR), Hannover. 2003. Available online: <https://services.bgr.de/boden/swr1000> (accessed on 8 August 2022).
23. BKG; SGD. WMS CORINE LAND COVER 5 HA—Status 2018, The Federal Agency for Cartography and Geodesy (BKG), Frankfurt am Main. 2021. Available online: <https://gdz.bkg.bund.de/index.php/default/corine-land-cover-5-ha-stand-2018-clc5-2018.html> (accessed on 21 August 2022).
24. BGR; SGD. Soil Map of the Federal Republic of Germany 1:1,000,000 (BÜK 1000), Federal Institute for Geosciences and Natural Resources (BGR), Hannover. 2013. Available online: <https://www.bgr.bund.de/buek1000> (accessed on 8 August 2022).
25. BGR; SGD. Geological Map of Germany 1:1,000,000 (GK1000): Federal Institute for Geosciences and Natural Resources (BGR), Hannover. 2002. Available online: <https://services.bgr.de/geologie/gk1000> (accessed on 8 August 2022).
26. BGR. Field capacity of German soils to a depth of 1 m (FK10dm1000\_250). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo. 2015. Available online: <https://produktcenter.bgr.de/terraCatalog/OpenSearch.do?search=5a75335e-ba46-452f-8792-ac7e9b49da88&type=/Query/OpenSearch.do> (accessed on 2 November 2018).
27. BGR. Organic Matter Content of Top-Soils in Germany 1:1,000,000 (HUMUS1000OB). Federal Institute for Geosciences and Natural Resources (BGR), Hannover. Soil Information System FISBo 2007. Available online: [https://www.bgr.bund.de/DE/Themen/Boden/Projekte/Stoffeigenschaftenabgeschlossen/Flaechenrepraesentative\\_Auswertungen/Corg.html](https://www.bgr.bund.de/DE/Themen/Boden/Projekte/Stoffeigenschaftenabgeschlossen/Flaechenrepraesentative_Auswertungen/Corg.html) (accessed on 11 June 2018).
28. Didan, K. MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V061 [Data set]. NASA EOSDIS Land Processes DAAC. 2021. Available online: <https://lpdaac.usgs.gov/products/mod13q1v061/> (accessed on 8 August 2022).
29. Padarian, J.; Minasny, B.; McBratney, A.B. Using deep learning to predict soil properties from regional spectral data. *Geoderma Reg.* **2019**, *16*, e00198. [[CrossRef](#)]
30. Bengio, Y.; Lecun, Y. Convolutional Networks for Images, Speech, and Time-Series. In *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 1997.
31. Nogueira, F. Bayesian Optimization: Open Source Constrained Global Optimization Tool for Python. 2014. Available online: <https://github.com/fmfn/BayesianOptimization> (accessed on 8 March 2024).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.