Transactions on Engineering and Computing Sciences - Vol. 11, No. 3
Publication Date: June 25, 2023
DOI:10.14738/tecs.113.14741.
Karimanzira, D., & Ritzau, L. (2023). Event Detection in Groundwater Sensor Networks using Artificial Intelligence. Transactions on Engineering and Computing Sciences, 11(3). 48-62.

SERVICES FOR SCIENCE AND EDUCATION

Event Detection in Groundwater Sensor Networks using Artificial Intelligence

Divas Karimanzira
Fraunhofer Institute of Optronics,
Sytem Technologies and Image Exploitation (IOSB),
Am Vogelherd 90, 98693 Ilmenau, Germany

Linda Ritzau
Fraunhofer Institute of Optronics,
Sytem Technologies and Image Exploitation (IOSB),
Am Vogelherd 90, 98693 Ilmenau, Germany

ABSTRACT

In this paper, a method for detecting spatial and temporal anomalous events in groundwater sensor networks (high-dimensional time series data), such as system faults and attacks will be developed. Unlike recently developed deep learning frameworks for anomaly detection which do not consider the dependences between the variables and apply the existing relationships to predict the expected behavior of the sensors, the method in this paper extracts the relationships between the sensors spatially and temporally and learn to detect and simultaneously explain deviations from these relationships. This challenge is solved by using graph attention neural networks and structured learning. Attention neural networks can give useful interpretability in context of the anomalies detected and allows to identify their causes. To improve robustness the method considers aleatoric and parametric uncertainties by using ensemble specific value prediction and prediction intervals without assuming any data distribution. Furthermore, the model was connected to a fully connected classifier to classify typical groundwater network anomalies. The method was applied to a study area and it could be shown that the method could capture in 92% of the cases the complex correlations between the high dimensional variables, and enabled analysts to identify the causes of the anomalies.

Keywords: Deep learning, Event Detection, Graph Attention Neural Networks, Prediction Intervals, Uncertainty consideration.

INTRODUCTION

In the digitization era, for many communities, groundwater systems are equipped large sensor networks with several sensor nodes. These sensor nodes produce a large amount of data in form of time series of different variables. In groundwater, the nodes include sensors for measuring different parameters such as groundwater flow rates, hydraulic conductivity, water level, water quality, and etc. In such a network, the sensor nodes, the sensors themselves can be related in a complex, nonlinear manner. A good example of such a complex and nonlinear cause-effect relationship is the effect of dissolved oxygen. It has substantial effects on the water

quality regulating the valence state of trace metals and by constraining the bacterial metabolism of dissolved organic species [1]. Such high complexity and dimensionality cannot be easily tackled manually and timely by humans and hence requires automated methods for detecting anomalous events and allow interpretability of the results for water operators to react for action as soon as possible. The curse of dimensionality is associated in anomaly detection with the issue that the amount of data required for generalizability of the models increases with the number of features, which result in isolated and sparse data points which can conceal the true events [2].

Groundwater monitoring process, poses many challenges to anomaly detection algorithms. The measured signals have a mean which varies over time as the groundwater monitoring process is non-stationary. Different anomaly patterns are possible as illustrated in Figure 1. One challenge in anomaly detection is the lack of training samples with labels for supervised learning. Therefore, a series of methods based on unsupervised learning have been developed. These include methods such as distance-based, density-based, and clustering-based techniques [3]. The application of these methods with their ability to capturing only linear relationships is only limited as most of the real-world problems are highly nonlinear in nature. New methods which consider nonlinearity and high dimensionality are based on deep learning. For example, Aggarwal developes an Autoencoder (AE) to predict outliers [4]. For this purpose, the AE uses reconstruction error as an outlier score. Generative Adversarial Networks with their generation and discriminating model can generate realistic-looking data by sampling from a learned data distribution. This model has been utilized by many authors for anomaly detection, e.g., [5] and [6] More recently, [7] applied an combinbation of AEs and Generative Adversarial Networks (GANs) for anomaly detection. With its ability to extract temporal dependencies the Long Short Term Memory Neural Network (LSTM) in form of AEs have found so much application in multivariate time series anomaly detection ([8], [9], [10]). Unfortunately, most of these methods do not have the capability to explicitly learn the spatial relationships between the sensor nodes and also the dependencies between the sensors, and therefore cannot sufficiently model and explain many potential contexts and interrelationships between the sensors [11],[12].

Recently, in [13],[14] graph neural networks (GNNs) have been applied successfully on learning graph-structured data. However, they cannot be applied directly to multivariate anomaly detection, because classically GNNs model each node with the same parameters [13] but in reality, every feature (sensor) is measured differently and therefore different features have different behaviours. Also, typical GNNs assume a-priori knowledge of the adjacency matrix from the graph, but in a groundwater, sensor network the relations to create an Adjacency matrix are not known in a-priori. In their paper on graph deviation network and test on WADI an SMAT Dataset, Deng et al. could show that GDN can detect anomalies and explain them very well [11]. Chao Fang et al, extended the GDN model to consider model uncertainties assuming gaussian data distribution [15].

Therefore, in this paper we will develop a method based on graph neural network and structured learning for multivariate spatial and temporal anomaly detection. The model will be equipped with attention mechanism for interpretability as attention weights can give

information about the relationship of the features to one another. Furthermore, to overcome the problem of knowing the adjacency matrix in apriori as in the typical GNNs, the adjacency matrix will be learnt along during the model training as in [16], [11],[12]. We will improve the robustness by capturing aleatoric and parametric uncertainties of the anomaly detection model by considering ensemble prediction of specific value and prediction intervals as in [17]. The method does not assume any data distribution as in [15].

The main contribution of the work is twofold. The method based on graph attention neural networks can learn spatial and temporal relationships between different sensors and sensor nodes. The GAT does not require a-priori knowledge of the adjacency matrix. Determining the adjacency matrix apriori is a major issue in real world applications as it is influenced by many factors. It learns the adjacency matrix along with the model training. Furthermore, with the attention mechanism the relations between the features and the deviations from the normal state can be explained. The method also captures aleatoric and parametric uncertainties making the model robust for anomaly detection. Furthermore, the model can classify typical groundwater anomalies, which is very useful for water managers and environmental authorities.

MATERIALS AND METHODS

Study Area

The study area has several groundwater sensor nodes with different sensors for water level (WL) temperature (T), nitrates (NO3) measurements and others. For the demonstration and to illustrate the performance of the models, fours sensors nodes are selected. At each node, the timeseries of WL, T, NO3 are obtained which makes a total of twelve features. Data from 2019 to 2020 was used for experiments, whereby 60% was used for training, 20% for validation and the rest for testing. The data shows some typical groundwater anomalies and events such as trends or level changes.

Model Description

Sensor entities have different characteristic and their relationships in a groundwater sensor network can be complex, for example sensors measuring water level, water flow rates and water quality in a sensor node can behave similarly and it is also plausible that water quality sensors in different nodes can behave similarly. To capture all this, two components of the model are required 1) for modeling the sensor characteristics using embedding vectors and 2) a graphical representation of the sensor network.

Embedding Vector:

The first component is for capturing the characteristic behavior of each sensor, which can be accomplished by an embedding vector v_i of length d for each sensor i. The vectors v_i with the unique characteristics of the sensor will be used for learning to determine the sensor relationships in form of a graph and also in the attention mechanism. The elements of the vectors v_i will be adapted during training of the model after being initialized randomly.

Graph Representation:
In the second component, a directed node-link graph $G = (V, E)$ will be used to represent the relationships between the different sensors. In the node-link graph, the nodes will represent the sensor entities and the links will represent the causal dependencies between the sensors. The links are represented by the Adjacency matrix $A_{ij} \in R^{N\times N}$ where $E[i, j] = 1$ represent a direct link between node i and node j.

Let $C_i \in \{1,2, \cdots, N\}\backslash\{i\}$ denotes a set of candidate relationships for each sensor i to the others. The dependencies of the sensor i to the other candidates can be computed by finding the similarity between the sensor i's characteristic vector v_i and the characteristic vectors of its candidates $j \in C_i$. Hereby e_{ji}, the normalized dot product of all the candidate to one another is calculated as in Equation 1. The top k candidates with the best similarity are then selected, where k is defined by the user.

$$e_{ji} = \frac{v_i^T v_j}{\|v_i\|\|v_j\|} \text{ for } j \in C_i \quad (1)$$

$$A_{ji} = 1\{j \in TopK(\{e_{ki}: k \in C_i\})\} \quad (2)$$

The learned adjacency matrix A will be utilized in the next component of the model, which is the graph attention-based anomaly detection core.

Specific Value Prediction and Estimation of Prediction Intervals
The main aim of this component is to use unsupervised learning to predict the normal situation of the next step and then calculate the deviations of the measured sensor values of the next step from the predicted ones. The deviation will give a measure for anomalies. Therefore, the first part is a prediction component and the second part is a scoring mechanism.

To predict the sensor values at the current time $s^{(t)}$, the system uses historical information given by the time window of w. Therefore, at time t the input vector $x^{(t)} \in R^{N\times w}$ is given as follows:

$$x^{(t)} := \{s^{t-w}, s^{t-w+1}, \ldots, s^{t-1}\} \quad (3)$$

This type of model can be easily realized by sequential model structures such Long Short term Memory (LSTM) or Temporal Convolutional Neural Networks (TCN) but in our case we would like not only to do multivariate timeseries prediction but also to capture the dependencies of sensors in the sensor network. Here comes a graph attention mechanism as in [18], [19] into play. It combines a sensor i's information with its neighbors attentively using the learned graph structure represented by the adjacency matrix A. In our case, the information which describes the characteristic behavior of each sensor is in the embedding vector v_i. Therefore, this vector is incorporated unlike in the state-of-the-art graph attention mechanism (see [18],[20]). In this representation a sensor node is computed as a combination with its neighbors $z_i^{(t)}$ as given in equation 4.

$$z_i^{(t)} = ReLU(\alpha_{i,i} W x_i^{(t)} + \sum_{j \in N(i)} \alpha_{i,j} W x_j^{(t)}) \quad (4)$$

where $N(i) = \{j | A_{j,i} > 0\}$ denotes all the neighbors of sensor i obtained from the learned adjacency matrix A, $W \in R^{N \times w}$ is like the Adjacency matrix, a trainable weight matrix, and the attention coefficients $\alpha_{i,j}$ are computed as concatenates of the embedding vector v_i and the weight transformed inputs $x_i^{(t)}$:

$$g_i^{(t)} = v_i \oplus W x_i^{(t)} \quad (5)$$

From the concatenates $g_i^{(t)}$ we proceed as in [18] by applying the LeakyReLU as the nonlinear activation to compute the attention coefficient which are normalized using the softmax function as in equation 6.

$$\alpha_{i,j} = \frac{exp\left(LeakyReLU\left(a^T\left[g_i^{(t)} \oplus g_j^{(t)}\right]\right)\right)}{\sum_{k \epsilon N} exp\left(LeakyReLU\left(a^T\left[g_i^{(t)} \oplus g_j^{(t)}\right]\right)\right)} \quad (6)$$

Information about prediction intervals can be utilized to derive a Quality-Driven, Distribution-Free Loss Function for training the model as in [21]. As in [17], the loss function for minimization during training can be expressed as in Equation 7 [17].

$$Loss = \beta L_{PI} + (1 - \beta L_v) \quad (7)$$

Where v ∈ (0, 1), β is a weighting parameter for the two objectives, L_{PI} determined by equation 11 and L_v determined by Equation 12 are the losses for optimizing the PIs and specific value, respectively.

For deriving the loss for optimizing the prediction intervals, several statistical measures such as the mean prediction interval width (MPIW) Equation 8 and the prediction interval coverage probability (PICP) Equation 9 can be used to evaluate the uncertainty of the model [22]. The PICP gives the proportion of observed values ($\hat{s}_i$) within the estimated PI [23].

$$MPIW = \frac{1}{n} \sum_{i=1}^{n} \left(PI_i^u - PI_i^l\right) \quad (8)$$

$$PICP = \frac{1}{n} \sum_{i=1}^{n} C \text{ , where } C = \begin{cases} 1 & \left(PI_i^u > y_i > PI_i^l\right) \\ 0 & otherwise \end{cases} \quad (9)$$

According to the high-quality definition principles, PIs should minimize MPIW subject to $PICP \geq (1 - \alpha)$. To minimize MPIW equation (8) 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 for which $PI_L \leq PI \leq PI_U$ holds

$$MPIW_{cap} = MPIW * C \quad (10)$$

Remembering that a penalty should only occur in the case where $PICP < (1 - \alpha)$ results in a one-sided loss. Combining with equation (8) and adding a Lagrangian, λ, controlling the importance of width vs. coverage gives a new loss,

$$L_{PI} = MPIW_{capt} + \lambda \frac{n}{\alpha(1-\alpha)} max(\mathbf{0}, (1 - \alpha) - PICP)^2 \quad (11)$$

$$L_v = \frac{1}{T-w} \sum_{t=w+1}^{T} \rho(v_t U_t + (1 - v_t) \cdot L_t, s_t) \quad (12)$$

ρ is a regression function, U and L being the upper and lower bound of the prediction interval respectively. The prediction interval uncertainty σ_{PI}^2 can be calculated from ensemble prediction.

Anomaly Scoring Mechanism

For scoring the predictions at a given time t of the model for anomaly, the deviation of the predicted value from the observations is calculated for each sensor i as follows:

$$e_i^{(t)} = |s_i^{(t)} - \hat{s}_i^{(t)}| \quad (8)$$

As we are dealing with groundwater sensors such as water level, pressure, flow, they are all heterogeneous and work in different ranges, we must consider robust normalization [24] of the deviations of each sensor to a specific range using the median μ_{e_i} and the inter quartile range σ_{e_i} as in equation 9.

$$a_i = \frac{e_i^{(t)} - \tilde{\mu}_i}{\tilde{\sigma}_i} \quad (9)$$

The anomalous score A(t) at time t is then computed as an aggregation of all sensor deviations using the max function

$$A(t) = \max_i a_i(t) \quad (10)$$

Sometimes not perfectly predicted values produce spikes even though the values are in the normal range. Therefore, in order to avoid misinterpretation of these values for anomaly it is purposeful to apply a moving average smoother to generate smoothed score $A_s(t)$ as in [25]. A threshold is set as the maximum $A_s(t)$ of the validation data and $A_s(t)$ of the test data is checked against this to make decision where a sample is anomalous or not.

Subsequence Classifier

Patterns of anomalies in ground water system appears in mainly four significant categories. They can be (1) spikes at a certain time point and/or location, whereby the observed value is significantly different from the other neighboring time points or locations. (2) a series of

deviating values in location or time for a continuous period of time or location; (3) a change in the baseline whereby suddenly the values change to a certain level for example due to seasonal effects of land use (cropping, different crop, etc.) (4) unusual fluctuation in observations which are troublesome in anomaly detection and can cause a lot of false positive if an anomaly detection system is very sensitive. Therefore, another component in the model is vital for classifying the anomalies in these groups. The structure of the sequence classifier is shown in Figure 2. It is comprised of two parts. The first part is an LSTM for extracting temporal features from the sequences and the second part is a fully connected neural network for the classification into 5 classes. The LSTM takes timeseries subsequences as inputs. A dataset was created for training the classifier. Typical subsequences with the five classes were extracted manually from the timeseries. The data was supplemented with synthetic subsequences to form a training dataset of 945 samples with19%, 21%, 21%, 22% and 17% of classes 1 to 5, respectively.

Baseline Model

Even though there were labeled data, we compared the model performance with one of the commonly used multivariate timeseries anomaly detector, which is based on Autoencoder technique and LSTM (LSTM-AE) [26]. It uses signal reconstruction errors as anomaly scores.

RESULTS AND DISCUSSIONS

Firstly, the classifier was analysed for its classification performance on a test dataset. The results of classifying some subsequences are shown in Figure 1. The different colors show the different categories as given in the legend.

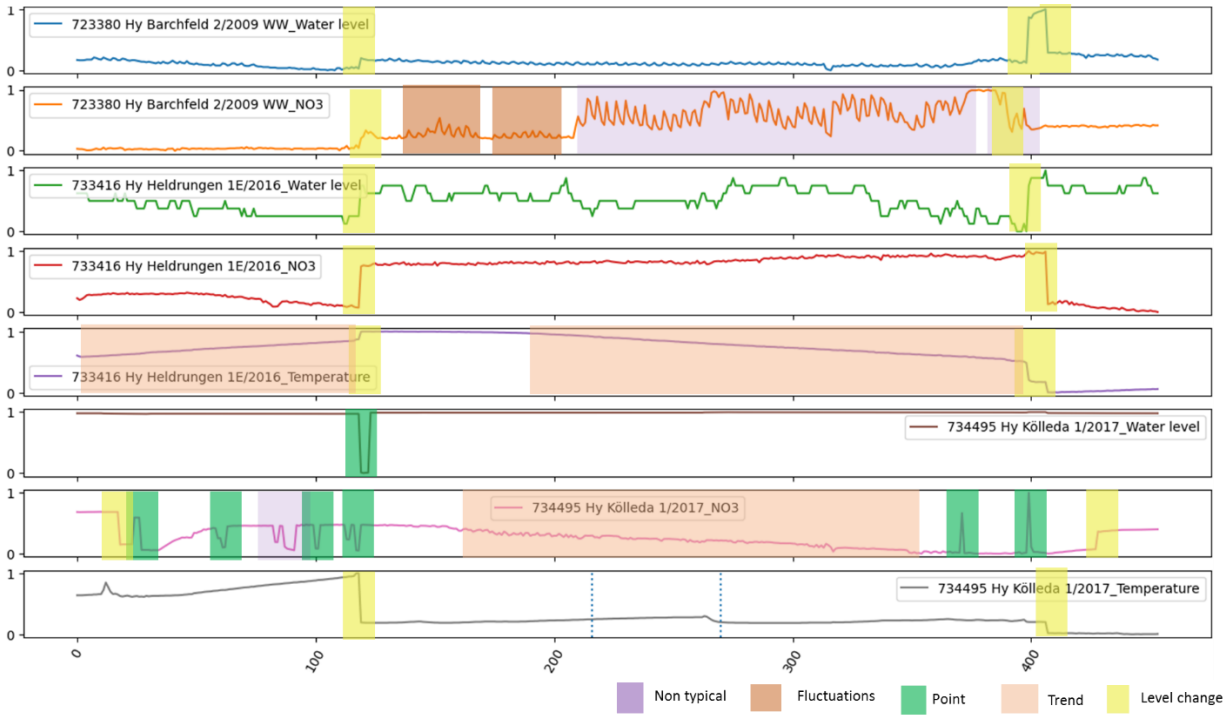


Figure 1: Results of the classification of the test dataset into the 5 classes, "non typical"," fluctuations"," point"," trend", and "level change"

The classification results of the classifier network are shown in form of a confusion matrix in Figure 1. As the name suggests, the confusion matrix simply tells us where the classifier gets confused. Several quantitative information can be interpreted out of the confusion matrix as shown in Table 1 for precision, recall and F1-Score. It can be seen that the classifier learnt very well to classify the classes "level change", "trend", and "Point" very well. The diagonal elements means number of correct classifications of each category. It can be seen that from a total of 618 samples, 504 were predicted correctly, which is an accuracy of the classifier of 82%. The class "fluctuation" was confused in some cases as "non typical" and "point" classes. It has the highest misclassification with an accuracy in prediction of about 62%.

Table 1: Quantitative measures for anomaly detection of the algorithm

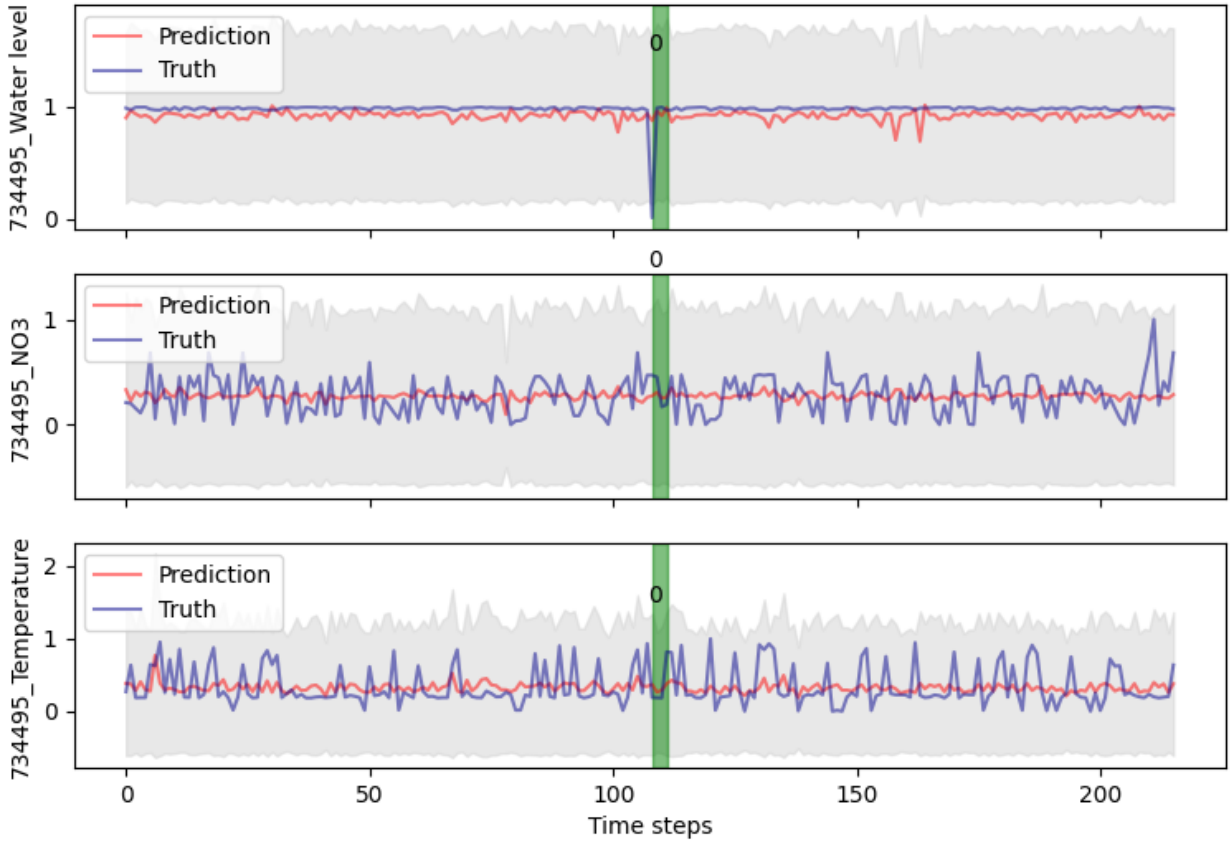
Class	Precision(%)	Recall(%)	F1-Score(%)
Non typical	76.3	78.3	77.3
fluctuation	67.5	61.3	64.2
point	82.6	86.3	84.4
trend	90.9	90.9	88.8
Level change	93.2	89.4	89.4

The model was set up as described in Section 2. Bayes optimization was applied to search for the hyperparameters which included the optimizer (adam, sdg and rmsprop), the learning rate, the data history window size (3-10) and the length of the embedding vector (32, 64, 128, 256). The best parameters were as follows: An embedding vector of length d = 128, window size = 7 and the adam optimizer with learning rate lr=0.0003 was used for training. After 48 Epochs the model stopped with early stopping.

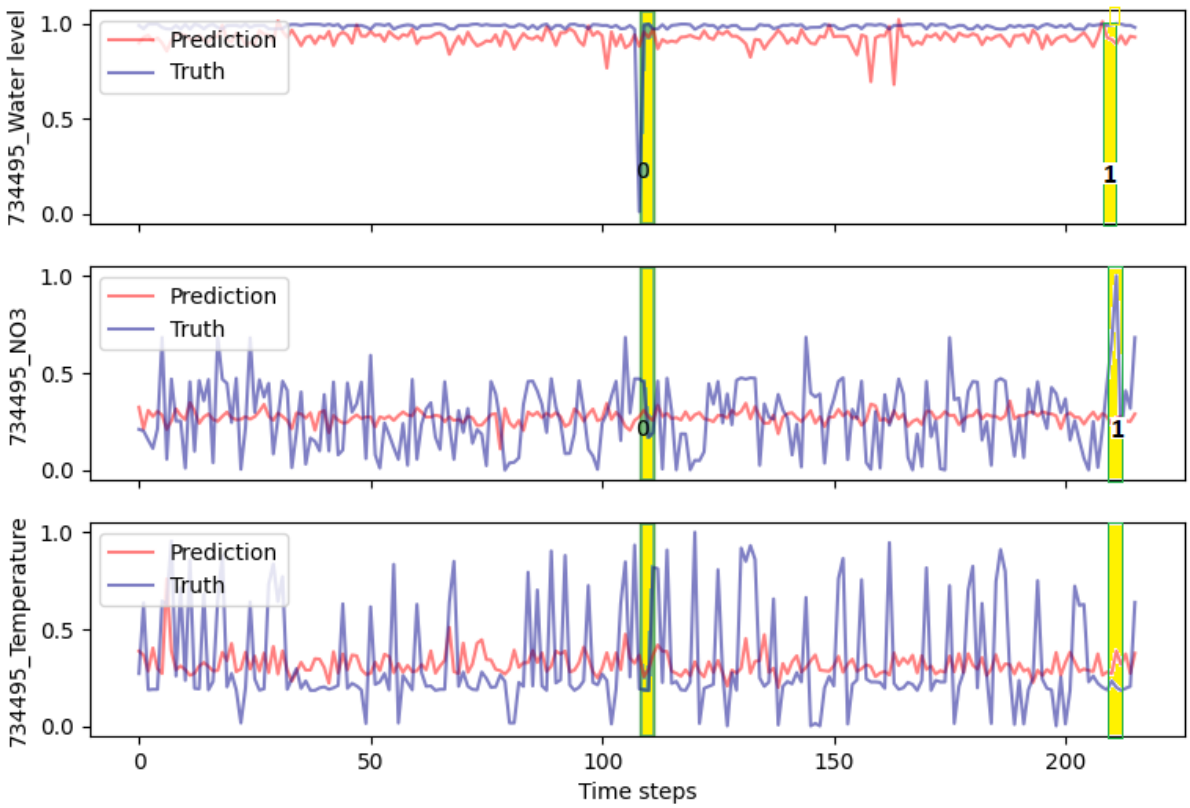
Due to lack of labeled data, the system is evaluated by plausibility checks from expert knowledge. There are typically known types of anomalies which appear in groundwater sensor networks such as single value anomalies, ripples in sensor measurements which are usually wrongly declared anomalies, continuous anomalies in form level change or trend. So, the expert knowledge will be used to evaluate the results of the models. Other knowledge, which can be used to evaluate the model is the knowledge about the relationships and dependencies between sensors, and sensor nodes. For example, oxygen and temperature affects the development of nitrate concentration, and also that the sensors in the same region should behave similarly. All this can be easily answered by the analysis of the learned graph structure, learned graph attention weights between the sensors and the clustering of the embeddings of the characteristics of the sensor plotted in t-SNE diagrams [27].

For the whole groundwater system under consideration, several temporal regions of anomalies were identified and are presented in the following figures by the shaded parts. There are single and continuous anomalies. In the following the plausibility of the findings will be elaborated and explained using the results of the individual sensors, sensor groups, t-SNE plot and the attention weights. In the test period 1 point anomaly was detected as show in Figure 2 a and b. As can be seen in Figure 1a, this anomaly was due to the water level sensor at node 734495. Such anomaly can result from sensor failure or uncertain transmission error. It can also be seen

that the ripples in the nitrate and the temperature sensor measurements are correctly identified as not anomalies compared to the results of the baseline LSTM-AE (Figure 2 b). The LSTM-AE does not consider uncertainties in its structure; therefore, it is over sensitive to the ripples in the nitrate sensor measurements in comparison to the method introduced here which considers both aleatoric and parametric uncertainties by using ensemble specific value prediction and prediction intervals without assuming any data distribution.



Our Method



(b) LSTM-AE

Figure 2: Point anomaly detected in the test period 1 resulting from the water level and nitrate sensors a) our method and b) LSTM-AE identifies the ripple in nitrate sensor as an anomaly.

In the test period 2, several continuous anomalies can be seen as shown in Figure 3 to Figure 5. It can be seen that the main contributors of the anomaly in region 0 are the sudden change in water levels at node 733416 and 734495, increase in nitrate levels at node 733416 and drop in the temperature at node 734495. Physically these changes are understandable as the sudden drop in water level can cause drop in temperature and increase in nitrate levels at the nodes. This behavior and dependence is also confirmed by the forced directed graph in Figure 6. It can be seen that the heaviest weights connect the water level at 733416 and the nitrates and temperature at 733416 and 734495. In anomaly time period 1, the sudden increase in the water level at 723380 and 733416 can be seen. This water level increase is in connection with the decrease in nitrate concentration at these nodes and the decrease in temperature. Anomaly in region 2 of the second test period is due to the increase in nitrate concentration at node 734495.

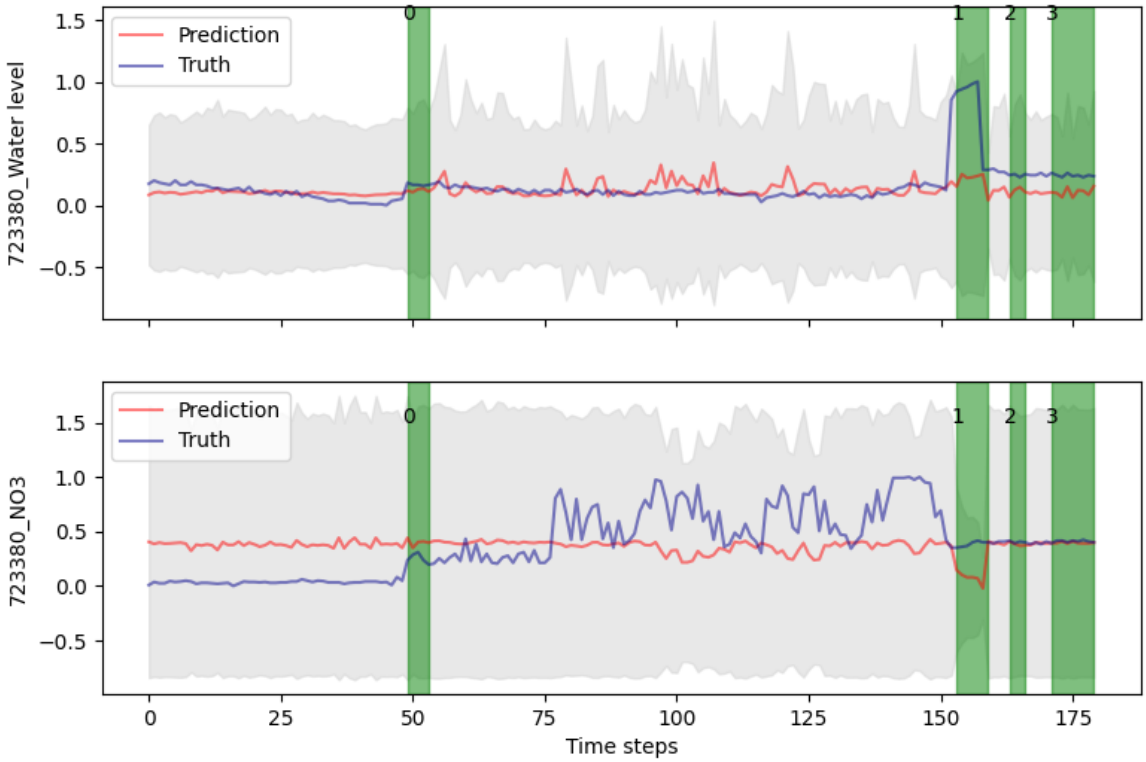


Figure 3: Temporal regions of anomalies detected from the groundwater sensor networks in the test data period 2 at node 723380

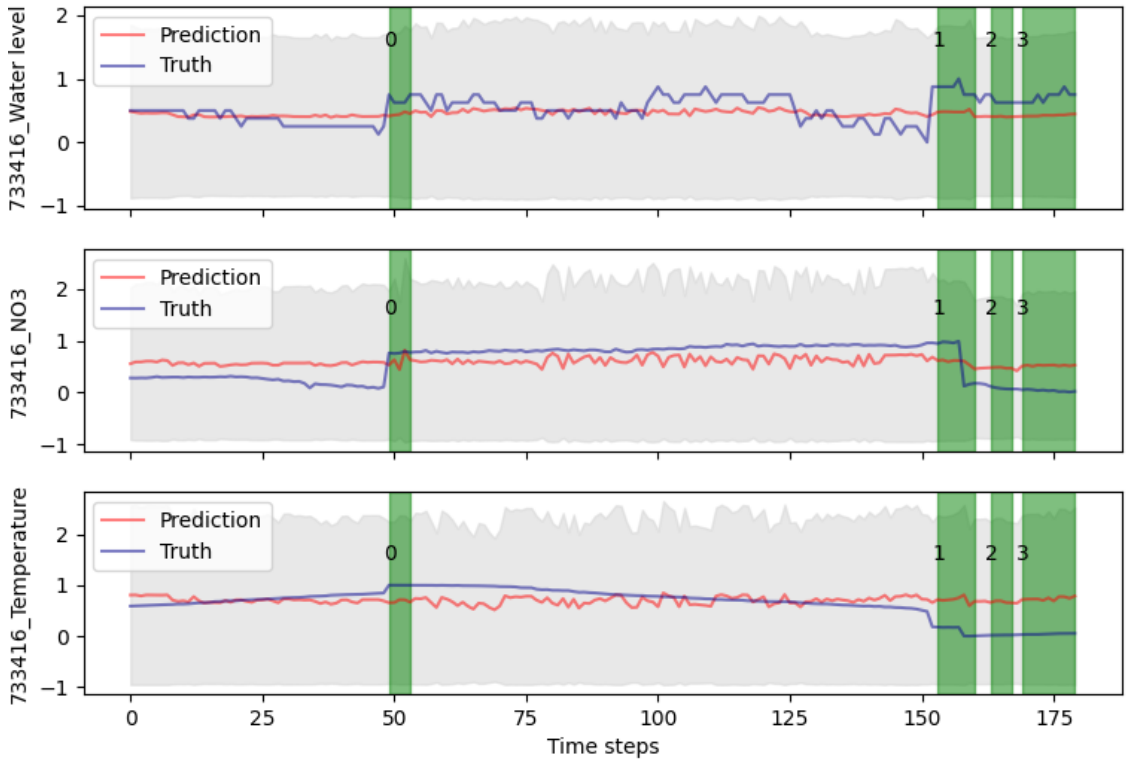


Figure 4: Temporal regions of anomalies detected from the groundwater sensor networks in the test data period 2 at node 733416.

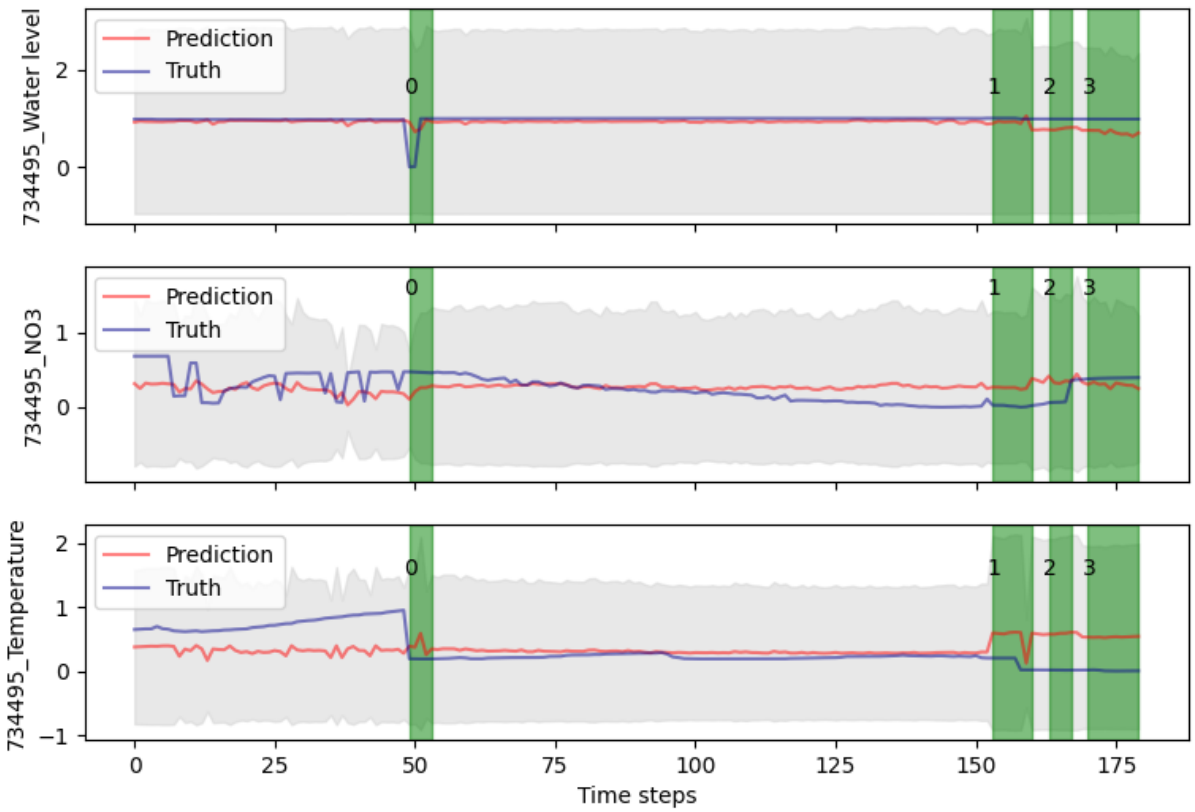


Figure 5: Temporal regions of anomalies detected from the groundwater sensor networks in the test period 2 at node 734495.

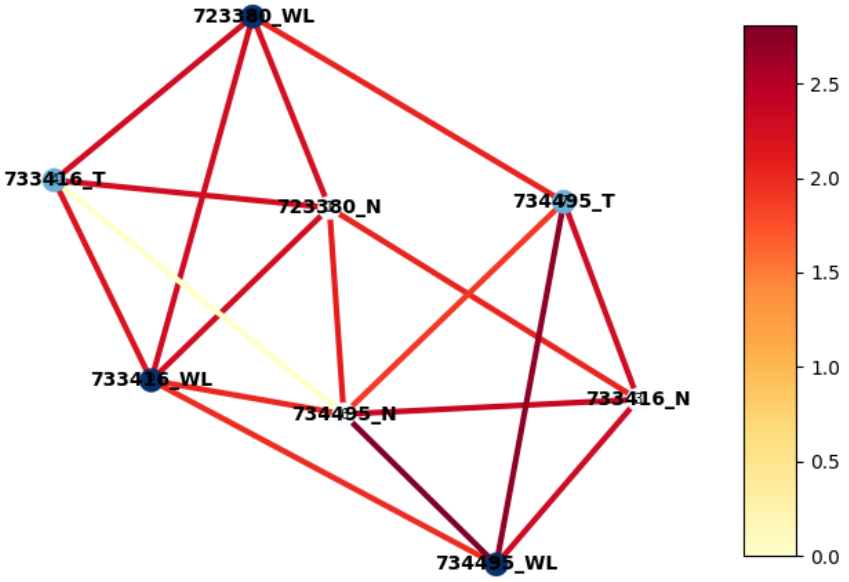


Figure 6: Graph structure with attention weights. The darker the color the heavier the weight showing the similarities between the entities.

Anomalies are rare events and therefore the datasets are unbalanced. We use standard metrics (precision, recall and F1-Score) which consider this aspect for evaluation of the performance of the model on the test dataset. Our method produced plausible results and high-performance results. In terms of accuracy, a precision of 95% on the dataset from our model region. Compared to the LSTM model with F1-score of 0.37, it had a 54% better F1-score.

Besides the high accuracy, the methods produce explainable results through its sensor embeddings and learned graph structure. The sensor embedding shows the similarity between the sensors. Several plots can be made which can help to explain the anomalies. For example, a forced directed graph is plotted in Figure 6 to show the connections between the sensors and their weights. The color an of the lines gives the similarities between the sensors. For example, there is a strong correlation between the Nitrate and the water level sensors at position 734495.

CONCLUSION

In this paper a method for anomaly and event detection in groundwater sensor networks was presented. It is based on unsupervised learning to detect anomalies in unlabeled data. Its graph structured learning enables it to find the relationships and dependencies between sensors spatially and temporary. In this way dissimilarities in sensor or regional clusters can be easily detected for example in groundwater sudden rise of nitrate on a spot. With temporal feature extraction the spread in time to other regions can also be detected, which is a very important feature for early warning systems for the environmental agencies and water managers.

The results show that the system can detect most of the known anomalies in groundwater sensor networks. It is also able to support in explaining most of the detected anomalies through its attention weights. The method is also robust by considering aleatoric and parametric uncertainties by ensemble specific value prediction and prediction intervals, which makes it a very important tool for real world problems as could be shown for the groundwater sensor networks.

Acknowledgements

We acknowledge all the colleagues who supported in data preparation and the agencies who provided the data used in the studies.

References

[1]. Seth Rose, Austin Long (1988), Monitoring Dissolved Oxygen in Ground Water: Some Basic Considerations, Groundwater Monitoring & Remediation Volume 8, Issue 1

[2]. Erfani SM, Rajasegarar S, Karunasekera S, Leckie C. High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning. Pattern Recogn. 2016;58:121–34.

[3]. Aggarwal CC, Philip SY. An effective and efficient algorithm for high-dimensional outlier detection. VLDB J. 2005;14(2):211–21.

[4]. Aggarwal, C. C. 2015. Outlier analysis. In Data mining, 237–263. Springer.

[5]. Li, D.; Chen, D.; Jin, B.; Shi, L.; Goh, J.; and Ng, S.-K. 2019. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. In International Conference on Artificial Neural Networks, 703–716. Springer.

[6]. Sabuhi, Mikael & Zhou, Ming & Bezemer, Cor-Paul & Musilek, Petr. (2021). Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3131949.

[7]. Rafiee, Laya and Fevens, Thomas (2020),Unsupervised anomaly detection with a GAN augmented autoencoder, International Conference on Artificial Neural Networks, 479--490}, 2020. Springer

[8]. Zhou, Xiaokang & Hu, Yiyong & Liang, Wei & Ma, Jianhua & Jin, Qun. (2020). Variational LSTM Enhanced Anomaly Detection for Industrial Big Data. IEEE Transactions on Industrial Informatics. PP. 1-1. 10.1109/TII.2020.3022432.

[9]. Z. Tian, C. Luo, J. Qiu, X. Du and M. Guizani, "A Distributed Deep Learning System for Web Attack Detection on Edge Devices," in IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1963-1971, March 2020

[10]. Mobtahej, Pooyan & Zhang, Xulong & Hamidi, Maryam & Zhang, Jing. (2022). An LSTM-Autoencoder Architecture for Anomaly Detection Applied on Compressors Audio Data. Computational and Mathematical Methods. 2022. 1-22. 10.1155/2022/3622426.

[11]. Deng, Ailin & Hooi, Bryan. (2021). Graph Neural Network-Based Anomaly Detection in Multivariate Time Series. Proceedings of the AAAI Conference on Artificial Intelligence. 35. 4027-4035. 10.1609/aaai.v35i5.16523.

[12]. Tian, Z., Zhuo, M., Liu, L. *et al.* Anomaly detection using spatial and temporal information in multivariate time series. *Sci Rep* **13**, 4400 (2023). https://doi.org/10.1038/s41598-023-31193-8

[13]. Zou, Xu & Zheng, Qinkai & Dong, Yuxiao & Guan, Xinyu & Kharlamov, Evgeny & Lu, Jialiang & Tang, Jie. (2021). TDGIA:Effective Injection Attacks on Graph Neural Networks.

[14]. Bachiller, Pilar & Rodriguez-Criado, Daniel & Jorvekar, Ronit & Bustos, Pablo & Faria, Diego & Manso, Luis. (2022). A graph neural network to model disruption in human-aware robot navigation. Multimedia Tools and Applications. 81. 10.1007/s11042-021-11113-6.

[15]. Ray, S., Lakdawala, S., Goswami, M., & Gao, C. (2021). Learning Graph Neural Networks for Multivariate Time Series Anomaly Detection. ArXiv, abs/2111.08082.Aykas, Dogan & Mehrkanoon, Siamak. (2021). Multistream Graph Attention Networks for Wind Speed Forecasting.

[16]. Simhayev, Eli, Gilad Katz, and Lior Rokach.(2020) "PIVEN: A Deep Neural Network for Prediction Intervals with Specific Value Prediction." arXiv preprint arXiv:2006.05139 (2020).

[17]. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, andY. Bengio, "Graph attention networks," 2017.

[18]. Yukai Ding, Yuelong Zhu, Jun Feng, et al. Interpretable spatio-temporal attention LSTM model for flood forecasting. Neurocomputing 2020(403): 348-359.

[19]. Kong, X., Xing, W., Wei, X., Bao, P., Zhang, J., & Lu, W. (2020). STGAT: Spatial-Temporal Graph Attention Networks for Traffic Flow Forecasting.IEEE Access,8, 134363-134372.

[20]. Tim Pearce, Mohamed Zaki, Alexandra Brintrup, and Andy Neely (2018). High-quality prediction intervals for deep learning: A distribution-free, ensembled approach. arXiv preprint arXiv:1802.07167, 2018.

[21]. Rahmati O et al 2019 Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and UNEEC methods Sci. Total Environ. 688 855–66

[22]. Dogulu, N., Lopez, P. L., Solomatine, D. P., Weerts, A. H., and Shrestha, D. L.: Estimation of predictive hydrologic uncertainty using the quantile regression and UNEEC methods and their comparison on contrasting catchments, Hydrol Earth Syst Sc, 19, 3181-3201, 10.5194/hess-19-3181-2015, 2015.

[23]. Dekking, Frederik Michel; Kraaikamp, Cornelis; Lopuhaä, Hendrik Paul; Meester, Ludolf Erwin (2005). A Modern Introduction to Probability and Statistics. Springer Texts in Statistics. London: Springer London. doi:10.1007/1-84628-168-7. ISBN 978-1-85233-896-1.

[24]. Hundman, K.; Constantinou, V.; Laporte, C.; Colwell, I.; and Soderstrom, T. 2018b. Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, 387–395.

[25]. Park, D.; Hoshi, Y.; and Kemp, C. C. 2018. A multimodal anomaly detector for robot-assisted feeding using an lstmbased variational autoencoder. IEEE Robotics and Automation Letters 3(3): 1544–1551.

[26]. Roweis, Sam; Hinton, Geoffrey (January 2002). Stochastic neighbor embedding (PDF). Neural Information Processing Systems.