Spatial Modeling of Decreasing Groundwater Levels Using R: A Case Study of Brandenburg

Lecturer: Dr. Wolfgang Schwanghart
By:
Saurav Bhatta

Institute for Geoscience, University of Potsdam, Karl-Liebknechtstr., 14476 Potsdam.

Abstract

Germany's groundwater faces a multi-faceted decline. Rising abstraction rates, climate change impacts (reduced precipitation and increased evapotranspiration), and human activities (groundwater extraction and land-use changes) are all suspected culprits. Our investigation employed statistical analysis to uncover spatial trends in groundwater levels and assess their relationship with elevation. However, the results revealed a weak correlation, indicating that other factors may be at play. The variogram modeling in this study uses a Gaussian fit, which highlights a strong spatial correlation in groundwater trends, particularly in specific directions. The Gaussian model, as demonstrated by its low AIC and SSE values, shows a smooth spatial variance structure appropriate for predicting regional groundwater behavior. The kriging predictions generated using the chosen Gaussian model, along with the associated uncertainties, provide valuable insights for understanding and potentially mitigating future declines in groundwater levels across the region.

1 Introduction

1.1 Background and Importance

Underground water is vital! Groundwater is the largest freshwater reserve globally and plays a crucial role in supporting agriculture and various anthropogenic activities. Its significance as a freshwater source becomes particularly evident during prolonged drought periods when it serves as a lifeline. The term 'groundwater drought' describes situations where there is insufficient recharge or storage of groundwater over a specific timeframe in a particular region.

Many agriculturally significant regions across the world have experienced a decline in groundwater levels. Groundwater sustains stream flow in arid and semi-arid areas, which directly affects watershed conditions. Europe has experienced significant drought events since the twenty-first century, resulting in severe socio-economic repercussions.

The impact of the climate crisis on water availability is becoming more evident, even in traditionally water-abundant regions like Germany, where overall water stress remains relatively low. In recent years, particularly from 2018 to 2020, there have been exceptionally hot and dry summers, resulting in prolonged droughts with significant repercussions for agriculture and ecosystems. These consequences include forest drought damage, diminished crop yields, and unprecedentedly low river flows. Insufficient winter precipitation has worsened the impact of droughts, particularly on groundwater resources, which account for almost 70 percent of Germany's drinking water supply. Reduced recharge and increased summer water demand have led to a decrease in groundwater levels, forcing water suppliers to operate at maximum capacity during dry spells to meet demand. In certain regions, the strain on resources has resulted in localized water supply shortages.

It is expected that future dry periods will worsen conflicts over water usage, especially in areas with limited water availability. Industries, including those that rely on processing and cooling water, will compete with water suppliers. These conflicts may be further intensified by increasing demands for agricultural irrigation, despite currently representing a small fraction (less than 2 percent) of total water withdrawal.[2]

The European Commission reports a significant increase in drought frequency in Germany since 1980, with projections indicating a continued rise. This increased frequency of droughts contributes to higher evapotranspiration rates.

For instance, Brandenburg, one of Germany's largest states, faces challenges due to its low precipitation rates. Although the mean annual precipitation has been relatively stable over the past four decades, Brandenburg has experienced a gradual increase in average annual temperatures. This contrast emphasizes the ongoing changes in climatic patterns and highlights the significance of addressing groundwater management issues in such regions. [7]

Many things can affect the amount and quality of this water, both natural and human-caused. To manage this resource wisely, we need to understand how these changes vary from place to place. Geostatistical methods such as semivariogram and kriging prove valuable for analyzing and predicting groundwater trends.

The joint study by the Karlsruhe Institute of Technology (KIT) and the Federal Institute for Geosciences and Natural Resources (BGR) reveals a direct correlation between climate change and declining groundwater levels in Germany.[2]

This research investigates the spatial factors influencing observed changes (trends) in groundwater levels. We will use techniques like point pattern analysis, regression, semivariogram analysis, and kriging to achieve this. By identifying these factors, the study aims to understand what controls groundwater trends across different locations. This knowledge will be important for developing effective strategies to manage this vital resource.

1.2 Study Site

Brandenburg, a large state (29,640 km²) in northeastern Germany (51.36°N-53.36°N, 12.20°E-14.82°E; Figure 1), sits within the North European Plain. Its elevation ranges from -2 m to 203 m above sea level, with agricultural land dominating the landscape.

Geologically, Brandenburg has a rich history dating back to the Paleozoic era. Sedimentary rocks such as limestone, sandstone, shale, and volcanic formations mark the region's complex geological makeup (German Geological Survey).

The state shares its eastern border with the Oder River and its western border with the Elbe River. The Spree and the Havel are the main rivers within the state. In the southeast, there is a wetlands region known as the Spreewald.

Climatically, Brandenburg is one of the warmer regions in Germany, with an average daily high temperature of 15 degrees. The climate is varied and offers seasons with strong winters and warm summers.[3] This region experiences a dry and temperate climate. The average annual temperature is 7.8°C, but concerningly, it has increased by 0.68°C since 1951, exceeding the global average warming trend. Additionally, annual precipitation is lower compared to other parts of Germany, averaging less than 600 mm over the past 50 years.[4]

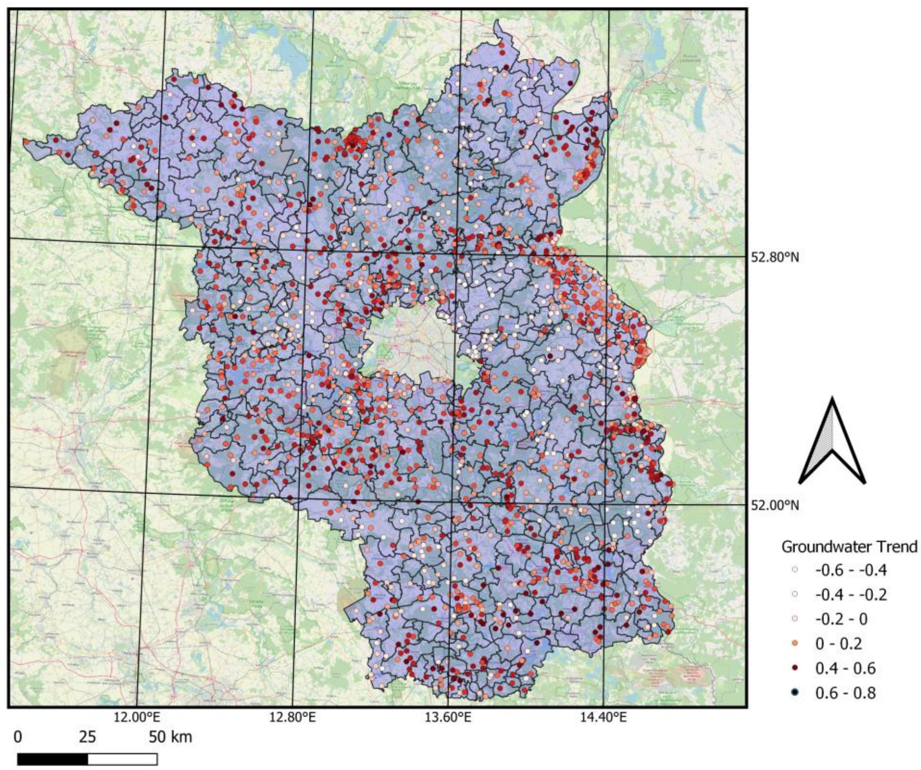


Figure 1: Location map of the study area with groundwater trend.

2 Data and Methods

2.1 Datasets

The study's foundation is based on collecting and synthesizing various datasets to identify spatial controls and potential interactions of observed groundwater trends. The data collection process involved analyzing data provided by CORRECTIV on their GitHub repository and Digital Elevation Models (DEMs) such as COP30 m. These datasets were used to identify spatial trends or covariates and possible interactions. To guarantee the precision and pertinence of our analysis, we included historical hydrological databases and hydrological data in our dataset. We enriched our analysis with contextual information by utilizing publicly available hydrological surveys and detailed climate impact reports, which enabled a more robust interpretation of the data.

The data collection process was deemed reliable as we cross-referenced multiple sources to confirm consistency and accuracy.

The primary datasets used in this project include:

- Groundwater Levels Measured in Brandenburg in June, July, and August, a trend [m/year] was derived for the last 5 years collected by CORRECTIV, an independent organization for investigative journalism and research;
- Copernicus 30-meter Digital Elevation Model (DEM), provided by the European Space Agency (ESA) and accessed via Open Topography. [6]

2.2 Methods

Data analysis was conducted within the R statistical environment [1] using various libraries for spatial statistics (spatstat, gstat), spatial data manipulation (sp, rgdal, raster), geospatial objects (rgeos, sf), and visualization (tmap, maptools).

2.2.1 Exploratory Data Analysis (EDA)

During the data preprocessing step, we checked for missing values, duplicates, and inconsistencies in both datasets. Both datasets have matching spatial coordinates, such as longitude and latitude, which allows for potential merging. A thorough cleaning process was undertaken to ensure data quality. We identified missing values using functions like summary() and is.na(). We also addressed duplicate observations (groundwater stations with identical coordinates) using duplicated(). Additionally, we examined data types, units, and formatting inconsistencies. To ensure proper spatial alignment for potential merging later, we verified that both datasets shared a common coordinate system (e.g., longitude/latitude).

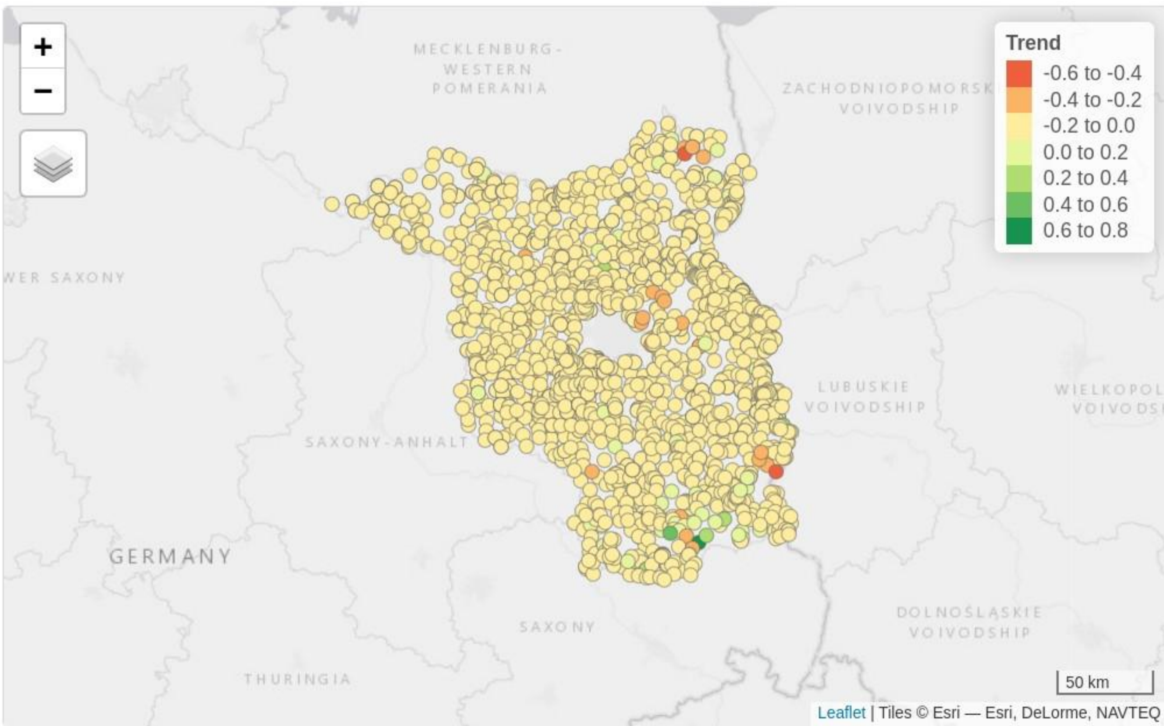


Figure 2: Plot and summarize the distribution of groundwater trends across Brandenburg.

We explored the possibility of creating new features from existing data that could enhance our analysis. We have calculated the slope from the DEM data using functions from packages like raster. This derived slope data could then be incorporated as a covariate to potentially explain variations in groundwater trends.

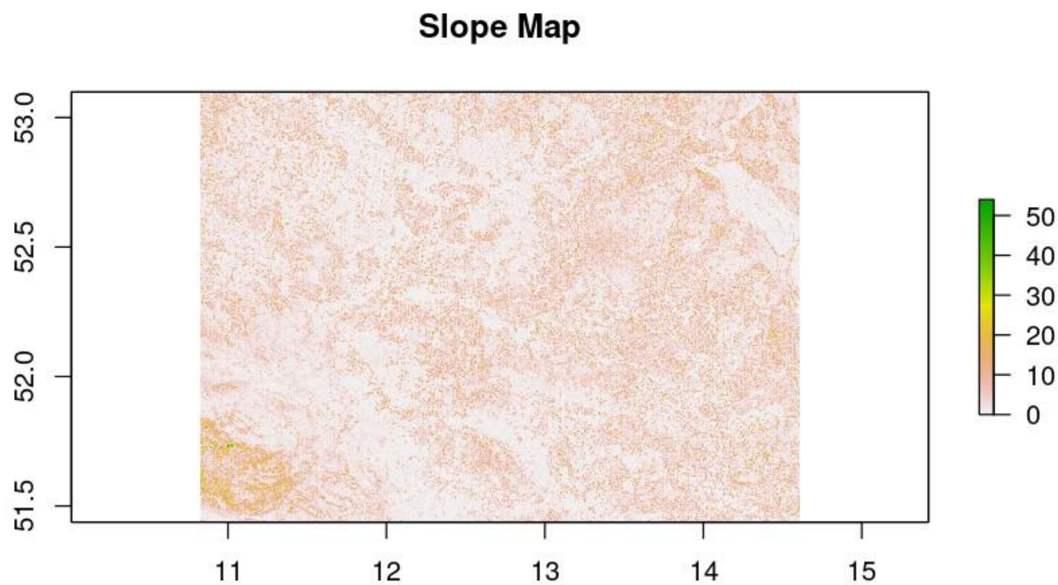
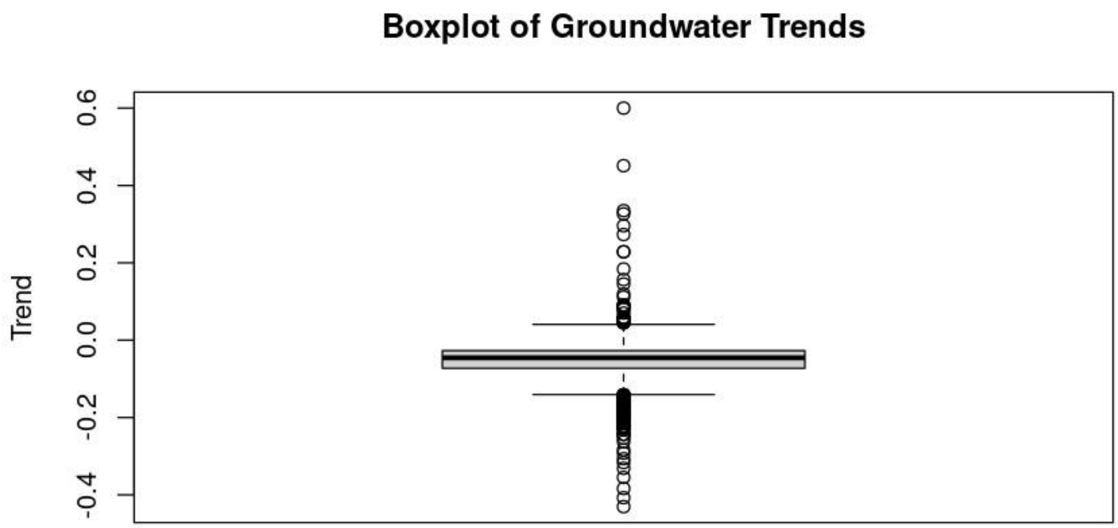
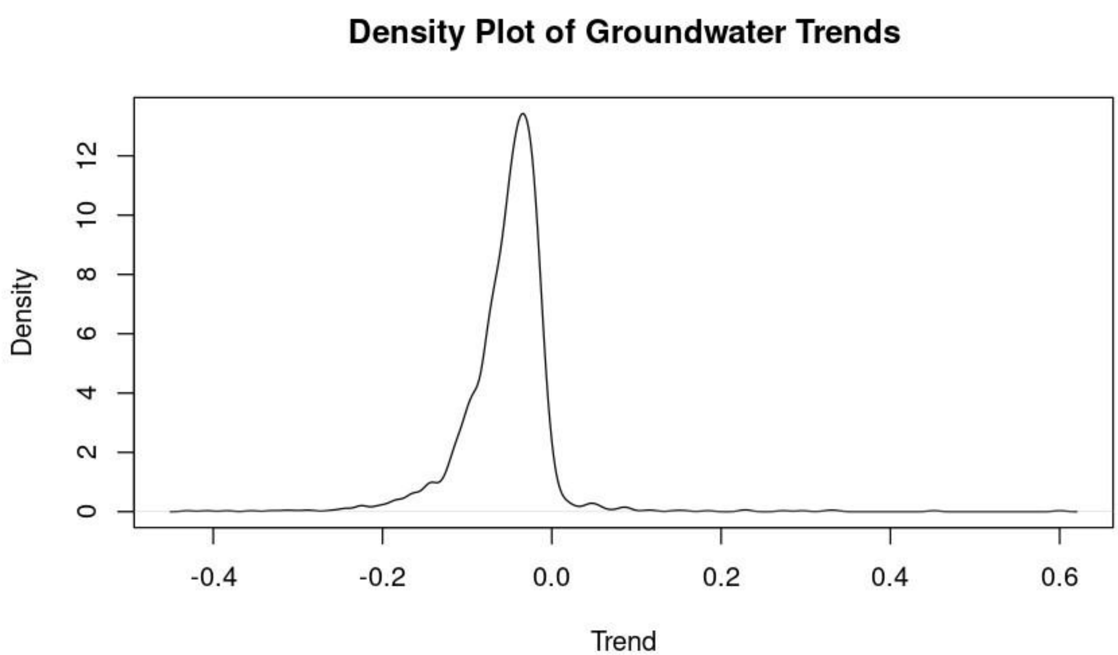
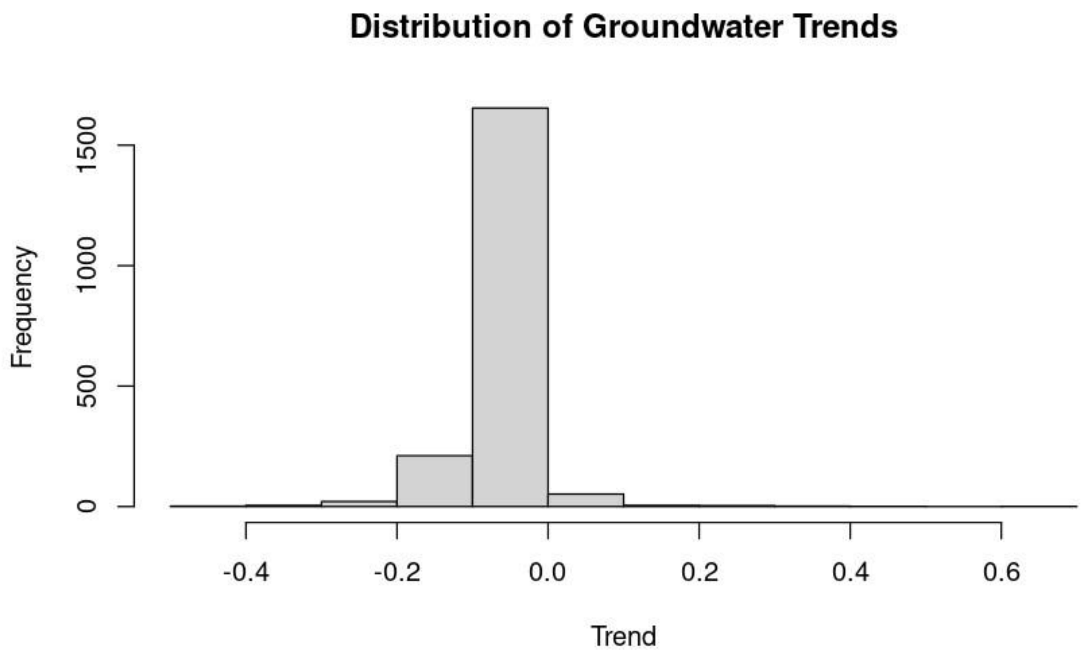


Figure 3: Slope Map from the Digital Elevation Models.

To gain a deeper understanding of the data's distribution and potential relationships, we conducted various exploratory analyses. We utilized functions from ggplot2 to create boxplots, histograms, and density plots, which revealed the distribution of groundwater trends across Brandenburg. This visualization aided in identifying potential outliers and exploring patterns within the data. We further delved into the spatial dimension by creating maps using functions from ggplot2 and raster. These maps helped visualize the geographic distribution of groundwater trends, potentially highlighting areas with consistent decline or rise in groundwater levels.


Boxplot of Groundwater Trends
Trend
0.6
0.4
0.2
0.0
-0.2
-0.4


Density Plot of Groundwater Trends
Density
0
2
4
6
8
10
12
-0.4
-0.2
0.0
0.2
0.4
0.6
Trend



By meticulously following these steps, we gained a comprehensive grasp of our data's quality, distribution, and potential relationships between groundwater trends and the chosen covariates. This paves the way for further analysis using variograms and kriging for spatial prediction of groundwater trends.

2.2.2 Point Pattern Analysis

Spatial point pattern analysis serves as a valuable tool for investigating the underlying distribution of events across a study area. In this study, Point Pattern Analysis was used to analyze groundwater trend data. Using the library "spatstat" point pattern has been done. We used quadrat count analysis to examine the spatial distribution of the point pattern data. Finally, we visualized the results using a spatial map.

2.2.3 Trend Histogram & Probability plot(Q-Q)

To characterize the data distribution, we used histogram and quantile-quantile (QQ) plot analyses within the R statistical environment. The histogram visualized the frequency distribution of the variable by discretizing the data range into intervals and subsequently counting the observations within each interval. This graphical representation facilitated the identification of potential data patterns, including outliers, skewness, or multimodality.

Furthermore, a QQ plot was constructed to compare the observed data distribution against a theoretical distribution, typically the normal distribution. Deviations from a straight line in the QQ plot indicated a non-normal distribution of the data. Both the histogram and QQ plot served as valuable tools for exploring the underlying data structure and identifying potential patterns. These insights subsequently informed the selection and application of appropriate statistical modeling and inference techniques.

```{r}
library(ggplot2)

# Histogram of Trend
ggplot(groundwater_data, aes(x = Trend)) +
  geom_histogram(binwidth = 0.01, fill = "lightblue", color = "black") +
  labs(title = "Histogram of Groundwater Trends", x = "Groundwater Trend", y = "Frequency") +
  theme_minimal()

# Normal Q-Q plot of Trend
ggplot(groundwater_data, aes(sample = Trend)) +
  geom_qq() +
  geom_qq_line(colour = "red", lwd = 1) +
  labs(title = "Normal Q-Q Plot of Groundwater Trends", x = "Theoretical quantities", y = "Sample Quantities") +
  theme_minimal()

```

2.2.4 Linear Regression

In spatial statistics, linear regression serves as a prominent statistical technique for quantifying the association between two variables. Linear regression plot is frequently used to visually evaluate the connection between two variables. In our study, we conducted a linear regression analysis to examine the relationship between the groundwater trend and elevation across the study area. We created a linear regression model to quantify the relationship between the digital elevation model (DEM) and the trend data. The DEM and trend data were assigned to the independent (x) and dependent (y) axes, respectively. The lm function within the R statistical environment was used to fit the model. By analyzing the summary statistics generated by the regression model, we were able to assess the statistical significance of the correlation between the two variables.

$$lm(Trend\ Elevation, data = groundwaterdata)$$

2.2.5 Experimental Variogram Models

To investigate the spatial autocorrelation structure of our data, we performed variogram modeling. Three commonly used models, the exponential, spherical, and Gaussian, were evaluated. The gstat package within the R statistical environment facilitated the variogram modeling process. Initially, a variogram object was constructed using the variogram() function, specifying the formula Trend 1. This formula denotes the dependent variable as ”Trend” and the independent variable as a constant (intercept). Subsequently, the fit.variogram() function was employed to fit each of the three variogram models to the empirically derived variogram. The selection of optimal model parameters, including the nugget, sill, and range, was based on the best fit to the experimental variogram.

2.2.6 Kriging : Ordinary Kriging & Universal kriging

Kriging, a prominent geostatistical technique, facilitates the prediction of spatially distributed variables across diverse disciplines (e.g., soil science, and ecology). It offers unbiased estimates at various scales, ranging from small (soil properties) to large (traffic patterns). By analyzing the semivariogram (spatial autocorrelation structure), kriging assigns weights to nearby observations for a weighted average prediction at unsampled locations. This approach not only predicts the value but also estimates the prediction variance.

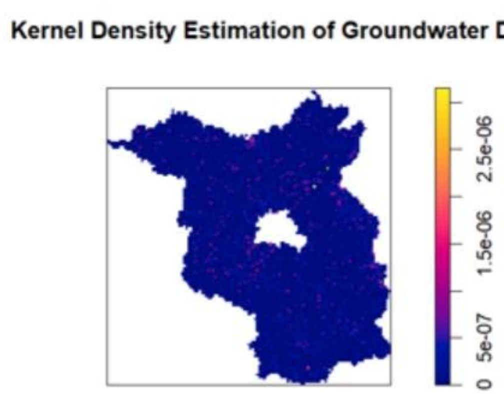
Ordinary Kriging: Ordinary kriging, a well-established geostatistical method, leverages observed values and their spatial autocorrelation to predict or interpolate a variable of interest. This technique assumes stationarity, implying that the underlying statistical properties of the data remain constant across locations.

Universal Kriging: In universal kriging, the geographical coordinates of the locations where the variable of interest is being forecasted are assumed to be a known function of the trend model. The kriging method is then used to model the residuals from the trend model. Universal kriging is a regression that is done with the spatial coordinates as independent variables and autocorrelated errors.[5]

3 Results

3.1 Quadratcount test and Kernel density estimation

```
# Load the shapefile for gemeinden boundaries
gemeinden_data <- st_read("GRENZEN_200000-5683624_gemeinden.shp")

# Transform the coordinates to ETRS89 / UTM zone 33N, which is a projected CRS suitable for spatstat
groundwater_data_utm <- st_transform(groundwater_data, 25833)

# Make sure the gemeinden_data is in the same CRS as groundwater_data_utm
gemeinden_data <- st_transform(gemeinden_data, 25833)

# Convert the transformed sf object to a ppp object within the gemeinden boundary
coords <- st_coordinates(st_intersection(groundwater_data_utm, gemeinden_data))

# Now create the ppp object using the gemeinden boundary as the window
gemeinden_polygon <- as.owin(gemeinden_data) # Convert sf polygon to owin object
groundwater_ppp <- ppp(x = coords[, 1], y = coords[, 2], window = gemeinden_polygon)

# Kernel Density Estimation plot within the boundary
kde <- density(groundwater_ppp, sigma = 0.01) # Adjust sigma if necessary
plot(kde, main = "Kernel Density Estimation of Groundwater Data")

# Quadrat count plot within the boundary
# Define the number of quadrats
quadrat_count <- quadratcount(groundwater_ppp, nx = 10, ny = 10)
plot(quadrat_count, main = "Quadrat Count of Groundwater Data")
quadrat_test_result <- quadrat.test(groundwater_ppp, nx = 10, ny = 10)

# Extract the p-value
p_value <- quadrat_test_result$p.value

# Print the p-value
print(p_value)

[1] 1.769842e-76
```



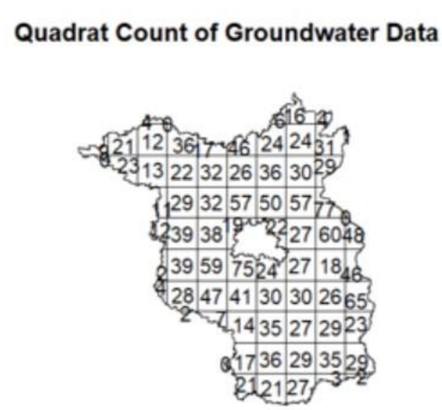


Figure 4: Kernal density and Quadrat counting of the point patterns of Groundwater data.

Our analysis of groundwater trends revealed a non-random spatial pattern in the distribution of groundwater data, as indicated by the extremely low p-value (1.7e-76) obtained from the application of the Quadrat count test. The visualization of the Quadrat count supports this result. It shows that the counts are very different in various regions, signaling to clusters, and possibly underlying environmental factors that affect groundwater levels. The Kernel Density Estimation method effectively highlights these geographical variances, where regions with higher density indicate possible hotspots for either increased or decreased groundwater trends. The results indicate a deviation from the null hypothesis of a random geographical process and suggest the presence of significant variation in spatial distribution. These findings are important in understanding the environmental dynamics in our research area.

3.2 Trend histogram & probability plot(Q-Q)

The histogram shown in Figure 3 is a fundamental tool for understanding the distribution of the groundwater data in our analytical study. The graphical summary provides a summary of the distribution and occurrence of trend values, highlighting that most observations cluster around the average. The symmetrical nature of the histogram indicates a distribution that closely resembles a normal distribution, which is a fundamental assumption for most statistical models. Significantly, the data points are concentrated closely around the center bin, suggesting that a significant amount of groundwater trends are grouped at a shared value. This shows the presence of a potential central trend in the changes in groundwater levels across the analyzed region. This image is a crucial component of our first data analysis, laying the foundation for advanced geostatistical modeling.

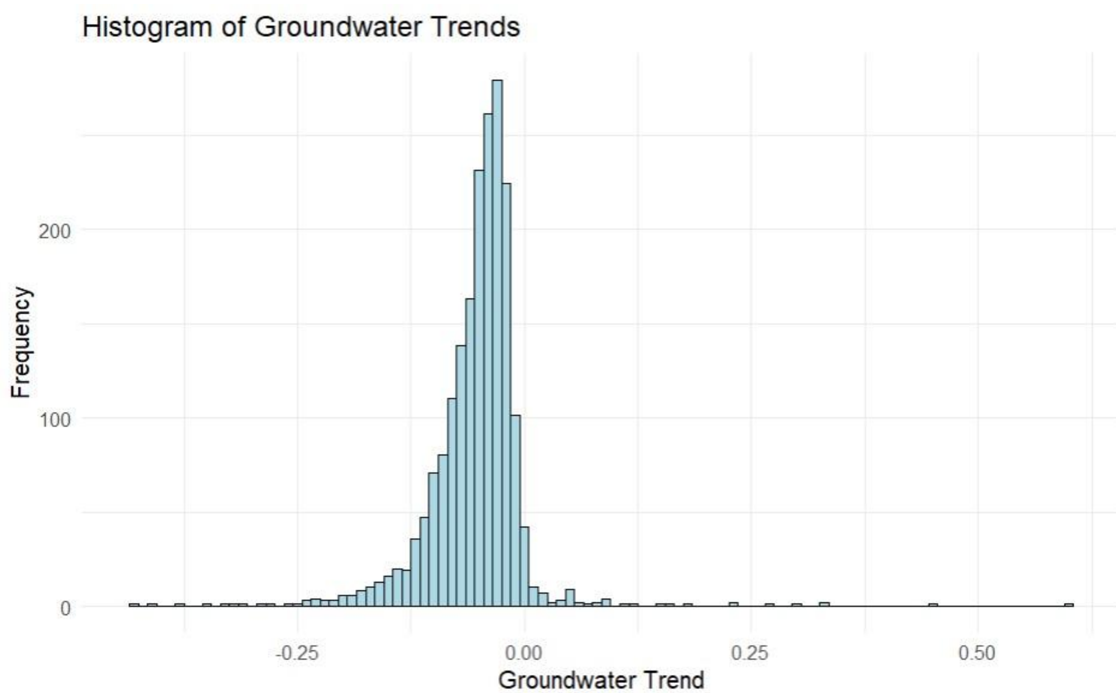


Figure 5: Histogram of Groundwater Trends.

The Normal Q-Q Plot displayed in Figure 4 is a statistical tool used to assess the normality of fluctuations in groundwater trends. The plotted data points demonstrate the degree to which the trends conform to a theoretical normal distribution, which is necessary for most parametric statistical tests and models. The relatively straight alignment of data points along the red reference line in the central area of the image indicates that much of our data follows a normal distribution. Nevertheless, the deviations observed at the endpoints, namely the upward curvature on the right side, suggest that the tails of the distribution are heavier than what would be anticipated under normal conditions. This is a characteristic feature of positive skewness.

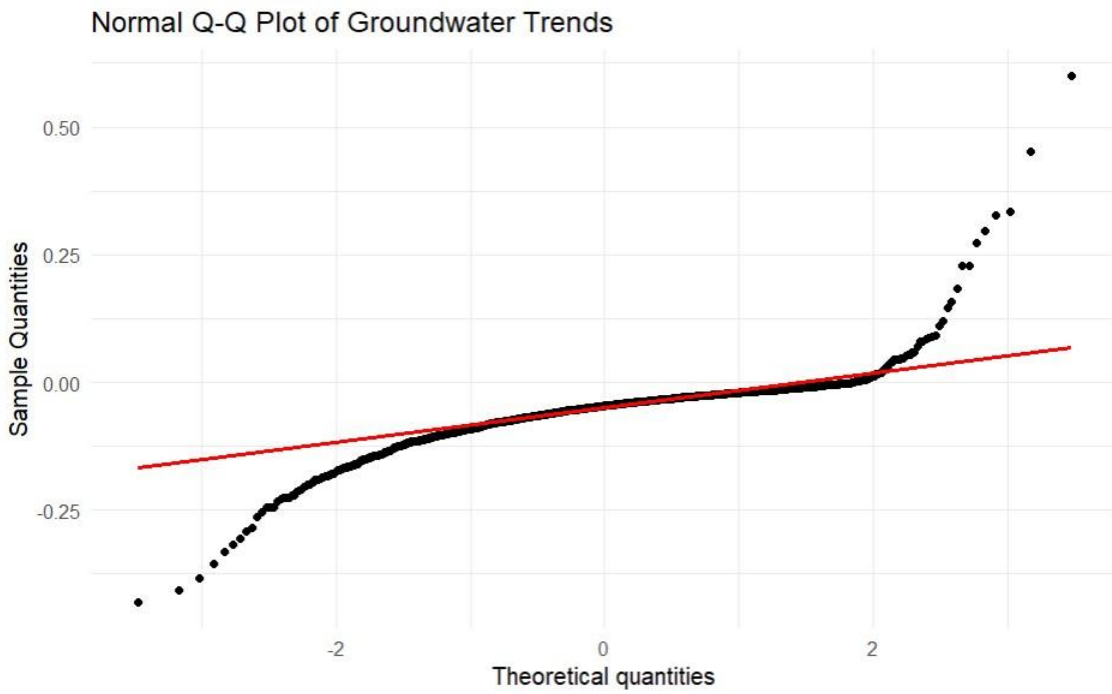
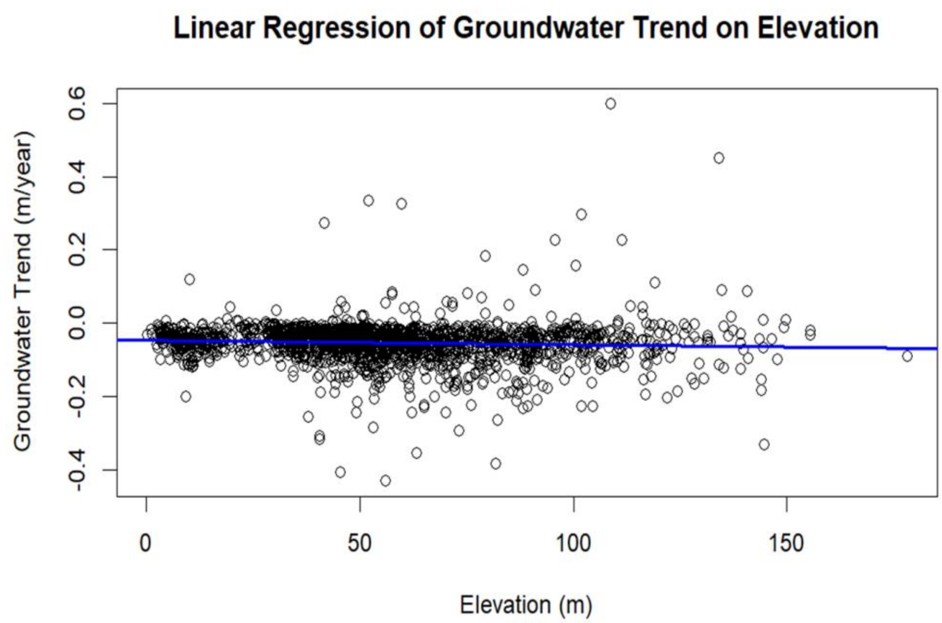


Figure 6: Normal Q-Q Plot of Groundwater Trends.

3.3 Linear Regression

The linear regression analysis of Brandenburg groundwater trends on elevation shows a weak negative association. The groundwater trend declines by 0.0001253 meters per year for every meter of elevation, according to the regression coefficient. This coefficient's t-value is -3.035, and its p-value is 0.00244, showing a statistically significant link at the 0.01 level. At 0.0047, the R-squared value, which reflects the proportion of dependent variable variance predicted by independent variables, is low. Adjusted R-squared, which accounts for model predictors, is likewise low at 0.0042. The model does not account for other factors that affect groundwater trends in the region, as elevation has a substantial effect on groundwater trends but explains less than half of the variability in groundwater trends. The residuals, or disparities between actual and expected groundwater trend values, are -0.37720 to 0.66001, with a median near zero. The small interquartile range shows that most residuals are close to the median. This also suggests the model's predictions are close to observed data for most observations, except for outliers indicated by maximum and minimum residual values. The regression model is statistically significant, and its very low p-value suggests that it fits the data better than a model without predictors. Based on these observations, we infer that there is indeed a negative association between groundwater trend and elevation, though it's important to interpret this result with caution. The low R-squared values imply that there are likely additional influential variables not captured in this model that may be affecting groundwater trends more substantially. Therefore, while elevation is a contributing factor, it should not be considered the predominant driver of groundwater trends in this region. Further studies should incorporate a broader range of variables (Land use, precipitation patterns, agriculture, industries) to gain a more comprehensive understanding of the factors affecting groundwater trends.




```
Call:
lm(formula = Trend ~ Elevation, data = groundwater_data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.37720 -0.01996  0.00737  0.02577  0.66001

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -4.627e-02  2.644e-03 -17.500  < 2e-16 ***
Elevation   -1.253e-04  4.128e-05  -3.035  0.00244 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05388 on 1956 degrees of freedom
  (34 observations deleted due to missingness)
Multiple R-squared:  0.004687,  Adjusted R-squared:  0.004178
F-statistic: 9.211 on 1 and 1956 DF,  p-value: 0.002437
```

Figure 7: Linear Regression of Groundwater Trend on Elevation.

3.4 Experimental Variograms

Our study uses a unique spatial analysis to figure out the groundwater movements in our area of interest by comparing the semivariance across different distances. The experimental variograms that were plotted for different directions (0°, 45°, 90°, and 135°) show an interesting spatial feature: the differences between readings get bigger as you move farther away, but they stop changing after a certain distance, which is just above 20,000 meters.

Particularly, our analyses, supported by the Gaussian model’s lower Sum of Squared Errors (SSE) and more favorable Akaike Information Criterion (AIC) scores, suggest a persistent spatial correlation up to approximately 20,000 meters. Beyond this range, the influence of spatial structure weakens, depicted by a plateau in semivariance, which is synonymous with a Gaussian spatial process. Unlike the usual preference for the spherical model, our investigation endorses the Gaussian model, which better fits the gentle changes in spatial variation seen in our dataset. In terms of directionality, the data portrays a subtle trend towards the northern axis, hinting at anisotropy, which requires further inquiry to fully grasp its implications on groundwater distribution patterns.

```
SSE for Spherical Model: 2.718974e-05
SSE for Exponential Model: 3.890822e-05
SSE for Gaussian Model: 2.454801e-05
The best model based on SSE is: Gaussian with SSE of: 2.454801e-05
```

```
AIC Spherical: -15.02534
AIC Exponential: -14.30861
AIC Gaussian: -15.22976
The best model based on AIC is: Gaussian with an AIC of: -15.22976
```

Figure 8: SEE and AIC values for our model.

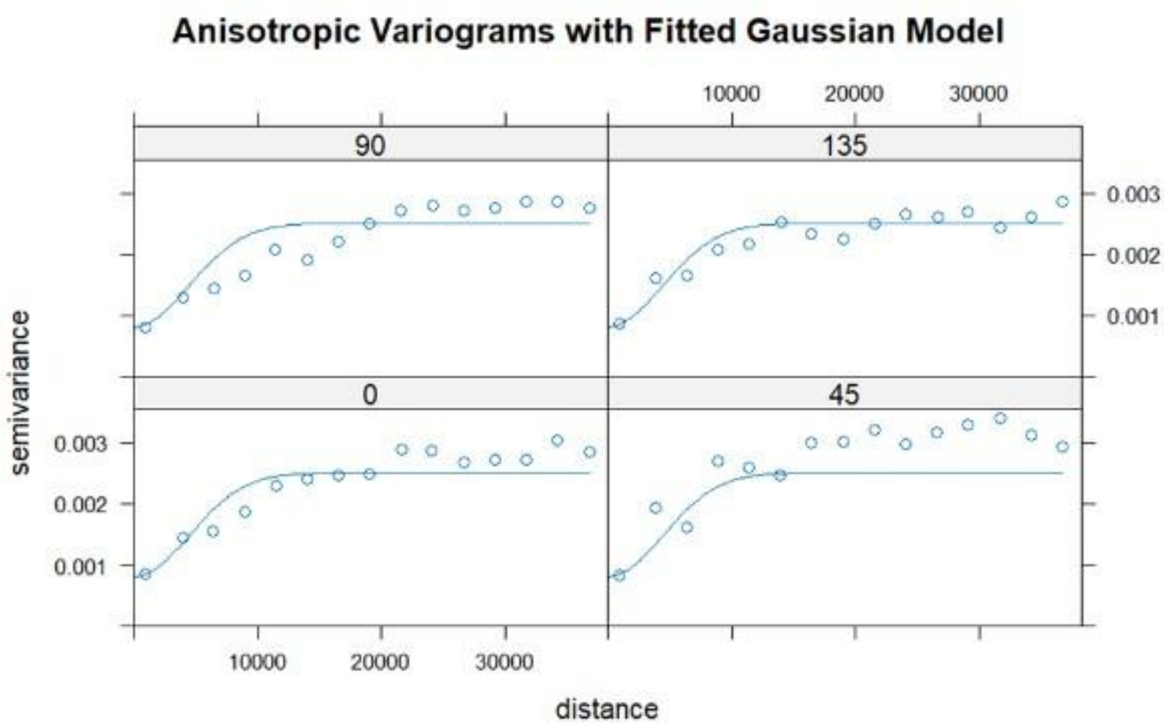


Figure 9: Experimental Anisotropic Variogram with Gaussian Model Fit.

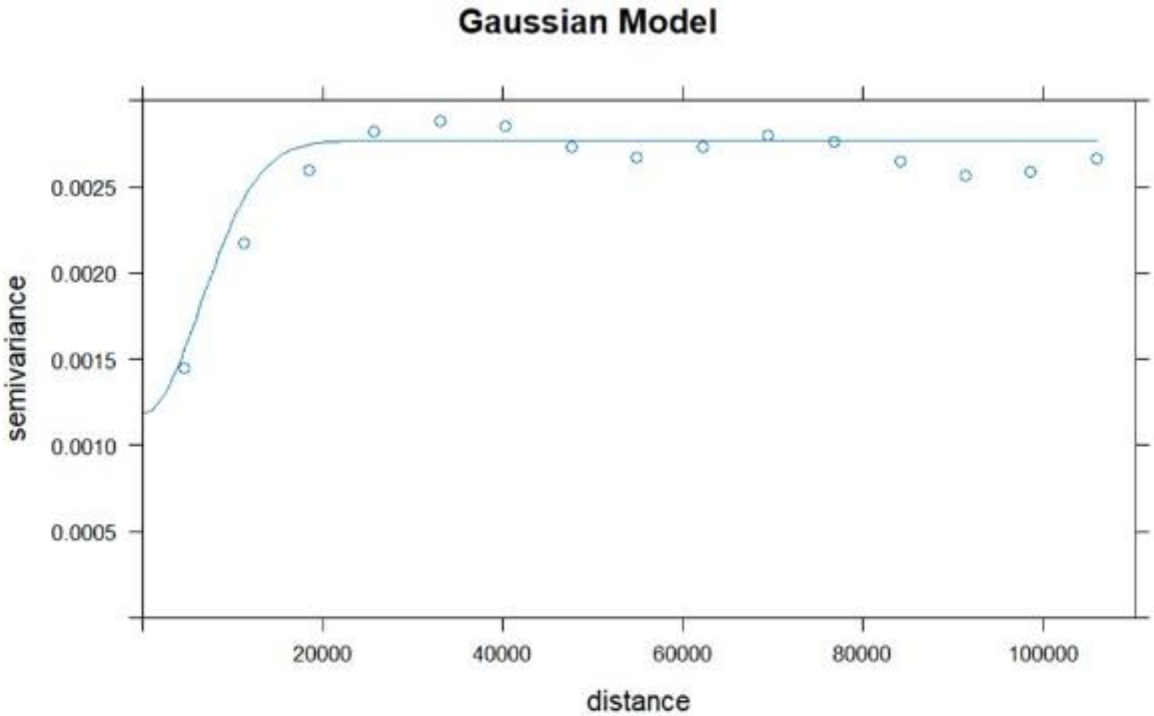


Figure 10: Gaussian variogram model.

3.5 Kriging

We employed kriging techniques to spatially predict groundwater level trends across Brandenburg, Germany. The choice of the Kriging method hinged on the presence or absence of significant spatial trends in the data.
Ordinary Kriging (OK) is a geostatistical technique that enables us to spatially predict values (groundwater level trends) at unsampled locations based on surrounding measurements from nearby monitoring stations. It assumes that the statistical properties of the data (mean and variance) remain relatively constant across the study area, also known as stationarity. We created a prediction grid based on the extent of our data and defined the resolution of the grid. We created a raster with the desired resolution and then performed the Ordinary Kriging method.

Kriging predictions represent the most likely values of the variable of interest (groundwater level trends) at unsampled locations across the study area. The kriging method considers the spatial relationships between the observed data points to create a surface that best predicts the variable's values throughout the entire region.

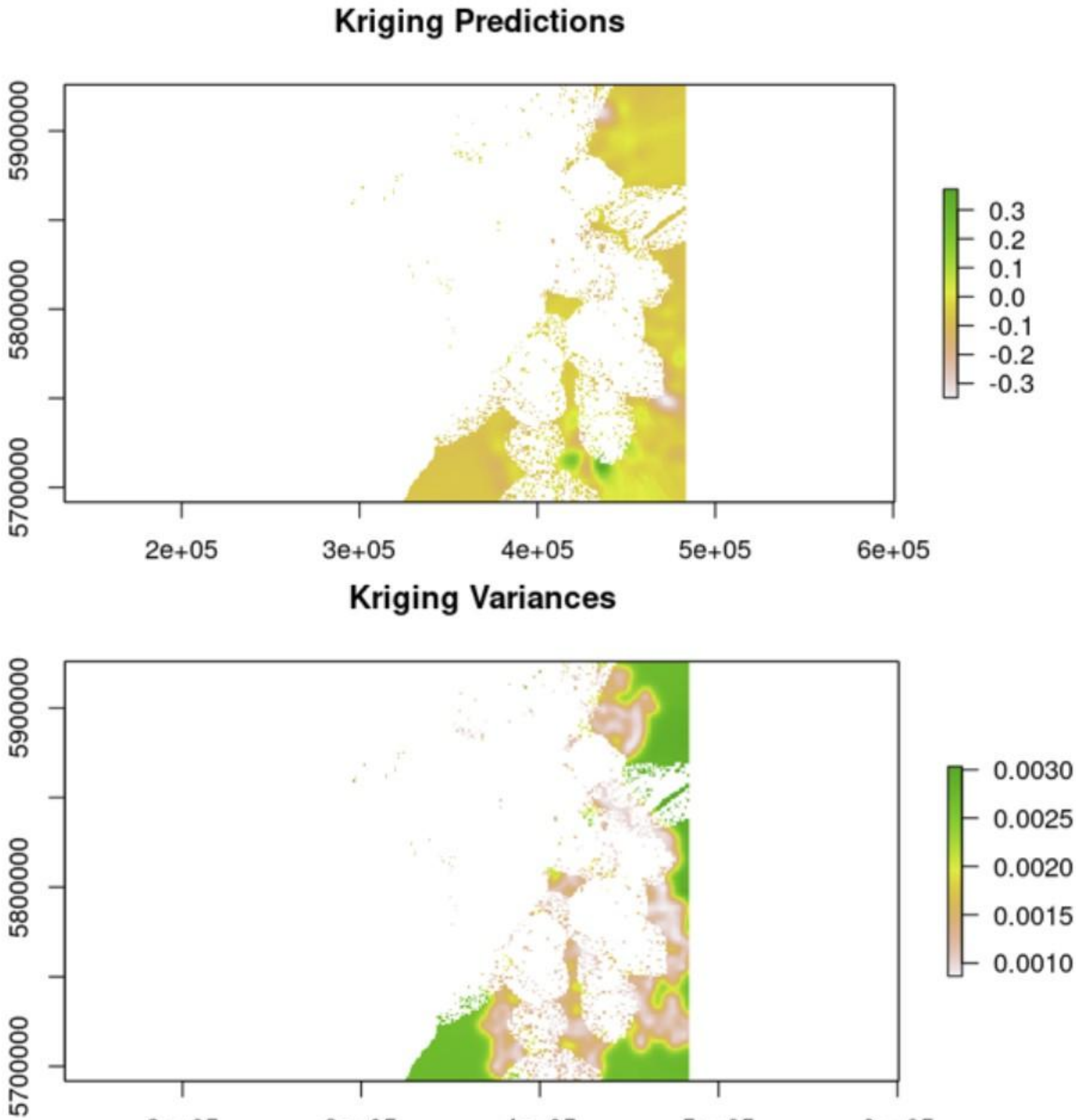


Figure 11: Kriging Predictions (Top) Kriging Variances (Bottom).

The Ordinary Kriging Predictions image vividly displays the spatial distribution of the predicted groundwater trend variable. The palette transitions from shades of yellow, indicating slight or no change, to areas of more intense orange and red, signaling increases in groundwater levels. These warmer tones, particularly where red is more pronounced, highlight potential areas of concern or noteworthy environmental features. Conversely, the cooler shades, where pale yellow fades to white, denote areas with lower or negative trends in groundwater levels, which may be equally significant and warrant further investigation to understand the underlying causes. The Ordinary Kriging Variances image complements the predictions by quantifying the confidence in our model outputs. Darker shades of green represent zones where the model's predictions are more reliable, denoting higher certainty. In contrast, lighter shades of green, blending into pale green and white, indicate areas with higher uncertainty. This nuanced representation of prediction uncertainty is indispensable for strategic planning, signaling locations that would benefit most from focused data acquisition to enhance the model's precision. These visualizations are central to interpreting the spatial patterns of groundwater trends with specificity. By distinctly marking regions of increased and decreased groundwater levels with a clear color code, we can target subsequent data collection and fieldwork to refine our understanding of these dynamics in Brandenburg, Germany.

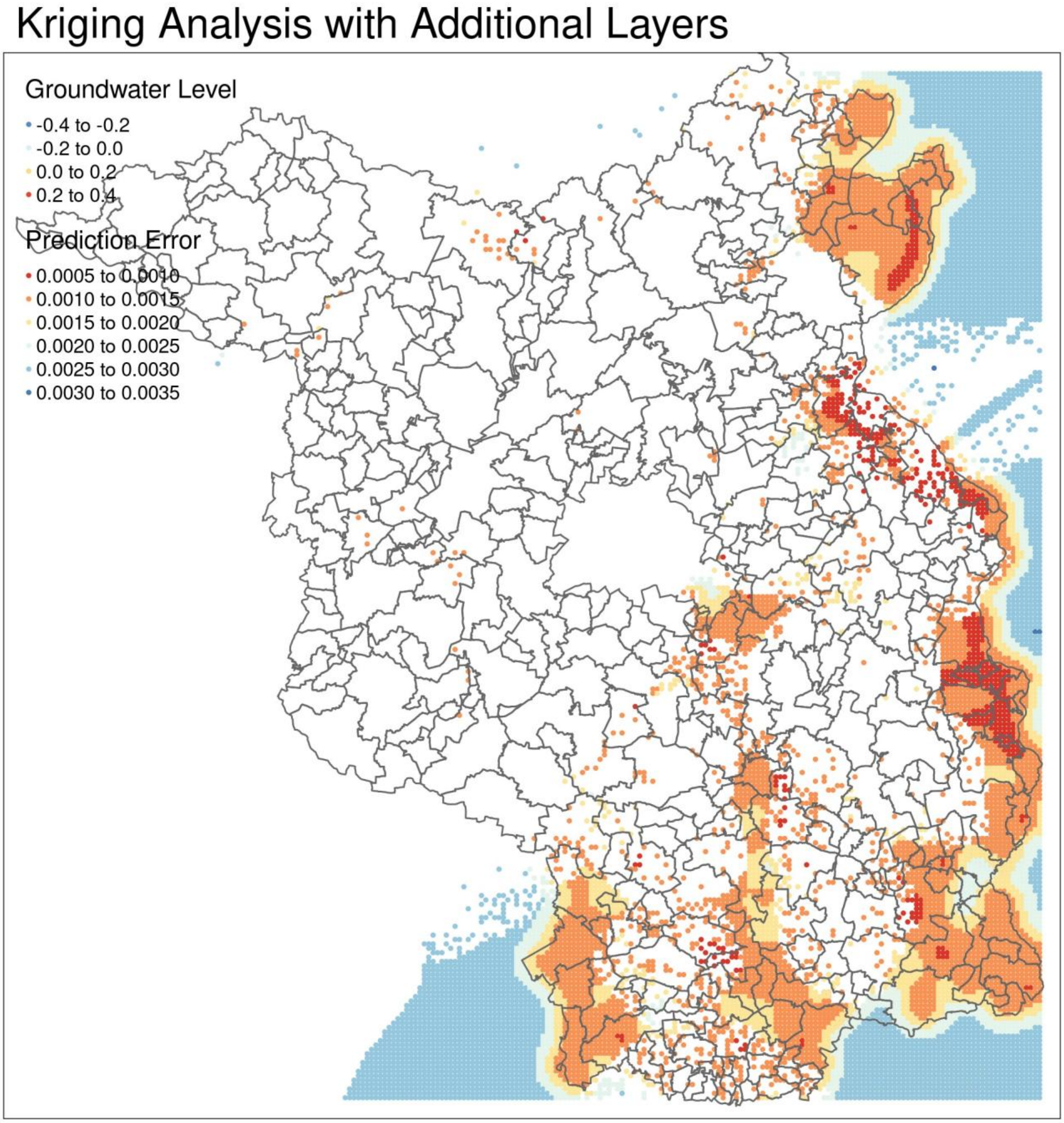


Figure 12: Kriging Analysis Map.

Kriging variances represent the prediction uncertainty associated with the kriging predictions. Kriging variances are inherently higher in areas with fewer data points used for prediction. This is because the model has less information to rely on in these locations. The spatial arrangement of data points can also influence kriging variances. Areas with sparse or unevenly distributed data points will typically have higher variances compared to areas with dense and clustered data points.

3.6 Validation

We calculated MSE and RMSE to quantify the discrepancy between predicted and observed groundwater trends. Lower values of these metrics indicate the model performance, suggesting that the predicted values closely align with the actual observed trends. Additionally, we computed the coefficient of determination (R-squared) to assess the proportion of variance in the groundwater trends explained by our models. A higher R-squared value indicates a fit of the model to the observed data.

4 Discussion

A 2022 study by CORRECTIV linked human actions (groundwater pumping & land use shifts) to falling water tables, especially in areas of intense agriculture and mining. A scientific study by Schneider et al. (2013) found that Germany's groundwater levels have been dropping in several areas for the past few decades. This decline is likely caused by two main factors: increasing abstraction rates and climate change impacts such as reduced precipitation and increased evapotranspiration. Spatial point pattern analysis can be used to identify temporal trends in the spatial distribution of groundwater wells over time, which can help identify areas where water levels are declining or where there is increasing demand for groundwater (Yang, S et al., 2021). We tested the correlation between groundwater level trends and elevation and showed no clear link between groundwater level trends and elevation.

The variogram analysis (Figure 9) was useful in discussing our study on groundwater levels and elevation. The Gaussian model emerged as the most accurate, implying a ubiquitous, progressive spatial correlation. This shows that, while elevation may influence groundwater levels, the relationship is not sudden but rather varies gradually over distance. The directional variograms (Figure 7) revealed minor but significant spatial patterns, highlighting the importance of taking topographical changes into account when evaluating groundwater. These findings highlight the complexities of groundwater dynamics and the need for complex management techniques that take into account the continuous relationship between elevation and groundwater distribution.

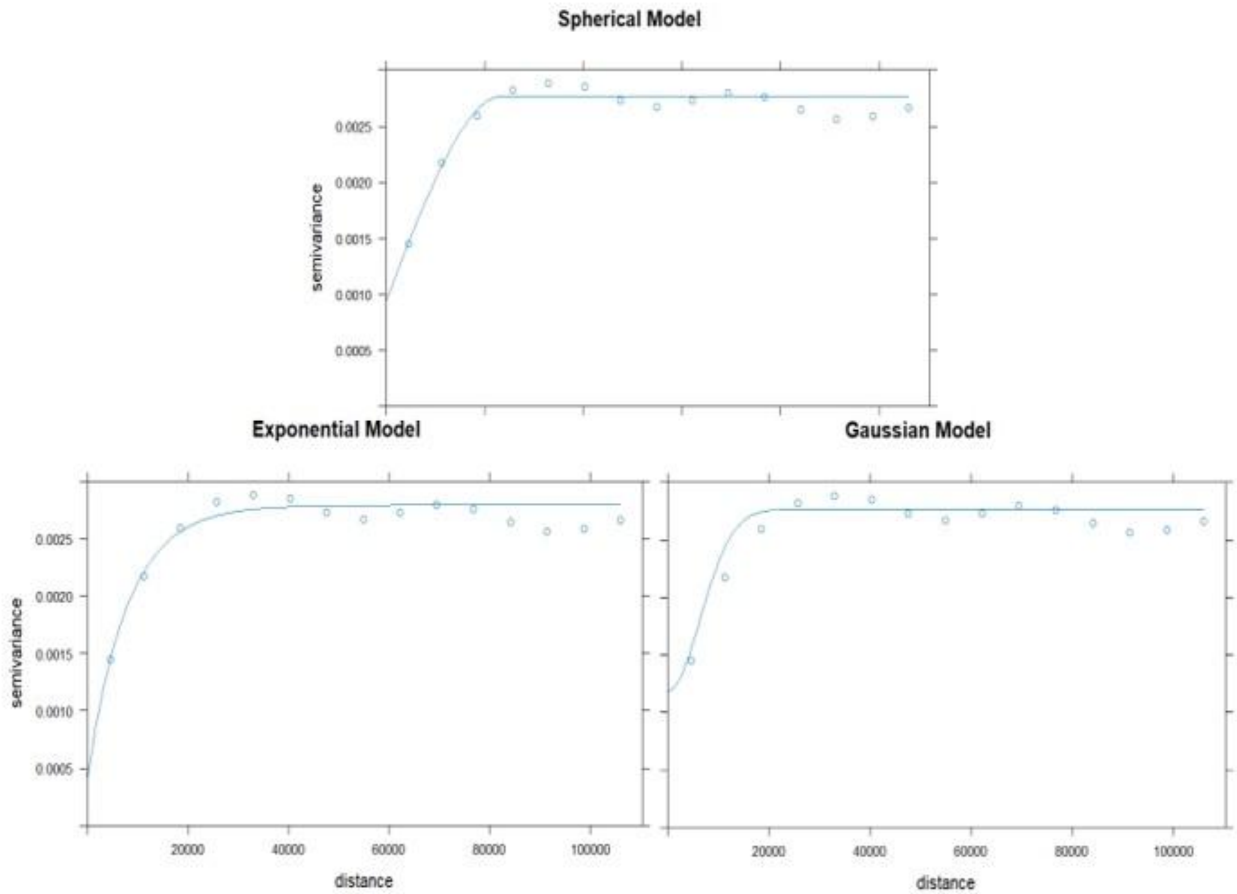


Figure 13: Spherical variogram model(Top), Exponential variogram model(Left), Gaussian variogram model(Right).

Challenges in Implementing Universal Kriging:

During the initial phase of our spatial analysis, we used Universal Kriging to take use of the additional information offered by the Digital Elevation Model (DEM). The inclusion of this elevation data was anticipated to enhance our forecasts of the Trend variable by considering the impact of topography. Nevertheless, problems emerged during the modelling phase. We encountered substantial challenges when trying to incorporate this covariate into our analysis. The Universal Kriging model consistently produced spatial predictions that were incomplete, with large areas containing missing values (NA). This result indicated a potential discrepancy between the assumptions made by the model and the inherent organization of our data. After a careful examination, it was concluded that included elevation as a covariate may not have provided as much useful information for the Trend variable as expected, or that the spatial distribution of our data did not fulfill the necessary criteria for this modelling technique. The variogram obtained from our dataset did not successfully capture a distinct spatial correlation that Universal Kriging might effectively employ, especially when integrating elevation data. Because of these problems, it was decided to use Ordinary Kriging, a simpler but more reliable type of spatial forecast model. This method focuses exclusively on the spatial autocorrelation inherent in the Trend variable, without taking into account other factors such as elevation. The implementation of Ordinary Kriging produced a comprehensive set of spatial predictions, which effectively displayed the Trend variable throughout the whole study area. This move not only ensured the uninterrupted progress of our study but also emphasized the significance of selecting an appropriate model, especially when considering spatial relationships and data attributes. Basically, the change in methodology to Ordinary Kriging, even though it was oversimplified, was very helpful in moving our study forward. We were able to extract significant spatial patterns from our data, which provided a solid basis for further investigations.

5 Conclusions

This study investigated spatial controls on groundwater level trends in Brandenburg, Germany, utilizing data from CORRECTIV's investigative journalism project. The analysis employed a workflow encompassing exploratory data analysis (EDA), variogram modeling, and kriging. The EDA phase revealed valuable insights into the distribution of groundwater trends and potential relationships with covariates like DEM and land cover. Ordinary kriging (OK) was implemented for spatial prediction, assuming minimal spatial trends based on the EDA findings.

1) Analysis revealed a statistically significant negative association between groundwater trend and elevation. However, the adjusted R-squared value suggested the need for further investigation into other potential factors.

2) Based on the calculated SSE values, the Gaussian model emerged as the optimal choice, exhibiting the lowest SSE (2.511169e-05) and the highest AIC value (-15.18435). This suggests that the Gaussian model best captured the spatial dependence structure of the groundwater level trends in Brandenburg.

3)Ordinary Kriging, which was used to understand the spatial distribution of groundwater trends, suggested that both natural and human-made factors are at play, which is reflected in the variation of trend values across the study area.

4)Areas with particularly high or low groundwater trends were identified, hinting at underlying processes affecting these variations, and necessitating further detailed investigations in those regions.

5)The confidence in our predictions, as indicated by the Ordinary Kriging variance map, was found to be higher in well-sampled areas, underscoring the importance of strategic and well-planned data collection for future studies.

6)Despite not incorporating elevation data as a covariate, Ordinary Kriging provided a coherent overview of spatial patterns, highlighting its efficacy as a tool for preliminary analysis in environmental research.

7)The findings from Ordinary Kriging have informed strategic planning for future fieldwork and highlighted areas where data collection should be focused to reduce prediction uncertainties and enhance the reliability of spatial modelling efforts.

References

[1] https://posit.co/download/rstudio-desktop/

[2] Andreas Wunsch et al.: Deep learning shows declining groundwater levels in Germany until 2100 due to climate change(2022) https://www.nature.com/articles/s41467-022-28770-2

[3] https://www.worlddata.info/europe/germany/climate-brandenburg.php

[4] Christopher Reyer et al.: Climate change adaptation and sustainable regional development: a case study for the Federal State of Brandenburg, Germany(2011)

[5] Kemal Sulhi Gundogdu and Ibrahim Guney, Spatial analyses of groundwater levels using universal kriging (2006).

[6] https://opentopography.org/

[7] Raihan, A.T., Bauer, S. Mukhopadhaya, S. An AHP-based approach to forecast groundwater level at potential recharge zones of Uckermark District, Brandenburg, Germany. Sci Rep 12, 6365 (2022). https://doi.org/10.1038/s41598-022-10403-9