Comparative Analysis of Spatial Interpolation Methods and Seasonal Variations of Rainfall Distribution in Lumbini Province, Nepal

Gopal Bhusal (Banner ID: 001310992)
Geoinformatics and Geospatial Analytics,
EEGS, Saint Louis University
gopal.bhusal@slu.edu

Abstract:
Accurate rainfall estimation is crucial for hydrological management and agricultural planning in Lumbini Province, Nepal, where diverse terrain and limited monitoring stations hinder measurements. This study evaluates four spatial interpolation methods—Inverse Distance Weighting (IDW), Radial Basis Function (RBF), Ordinary Kriging (OK), and Universal Kriging (UK)to determine the most reliable technique for estimating rainfall distribution. The key research question is: Which spatial interpolation method provides the most accurate rainfall estimates for Lumbini's varied terrain? This study uses rainfall data collected from 39 meteorological stations of the study location by the Department of Hydrology and Meteorology, Nepal, covering the period from 2000 to 2023, utilizing Python and ArcGIS Pro for data processing and interpolation. Each method's performance is rigorously evaluated through cross-validation, dividing the data into 75% for training and 25% for testing. The primary goal is to minimize the Root Mean Square Error (RMSE) while maximizing the coefficient of determination (R^2) to determine which method best represents rainfall patterns. The anticipated results are expected to reveal distinct rainfall patterns and seasonal trends throughout Lumbini Province. This research aims to provide valuable insights into regional rainfall variability, supporting improved understanding and management of water resources, which is critical for sustainable agricultural practices and effective disaster management in the area.

Keywords: Spatial Interpolation, Stochastic Model, Deterministic Model, Seasonal Variability

1. INTRODUCTION

Rainfall is a critical component of the hydrological cycle and directly influences water resource availability. Understanding its spatial and temporal patterns is vital for hydrology, agriculture, and disaster management. However, in regions like Nepal, where topography varies significantly, accurately estimating rainfall distribution remains a challenge due to sparse and irregular rain gauge networks. This limitation affects sustainable water resource planning, risk assessment, and development strategies (Wurbs & James, 2002).

Previous studies have explored various interpolation techniques to address rainfall distribution using limited observation data. Deterministic methods like Inverse Distance Weighting (IDW) and Radial Basis Function (RBF), as well as stochastic methods like Ordinary Kriging (OK) and Universal Kriging (UK), have shown varying degrees of accuracy depending on the region's geomorphology. While these studies demonstrated the utility of Geographic Information System (GIS) tools in rainfall distribution analysis, they often lacked comprehensive evaluation across diverse terrains and climates (Dingman, 2015), (Jones et al., 2010), (Isaaks & Srivastava, 1989).

This study focuses on evaluating the performance of four interpolation techniques—IDW, RBF, OK, and SK—using rainfall data from 40 stations in Lumbini Province, Nepal. By analyzing seasonal and spatial distributions from 2000 to 2023, this research aims to identify the most accurate technique for the region's complex topography. The study builds on existing methods to address gaps in spatial interpolation accuracy and contributes to better hydrological modeling and watershed management.

Research Questions:

1. Which spatial interpolation method performs best for rainfall distribution in Lumbini Province?
2. How does rainfall distribution vary spatially and seasonally across the province?

The study's findings are expected to enhance the use of GIS-based interpolation methods in regions with sparse data, contributing to sustainable water resource management and climate adaptation strategies.

2.METHODS

This chapter outlines the methods and methodology employed in the study. The spatial variation of rainfall distribution across the lowlands and highlands of Lumbini Province was analyzed using rainfall data collected from 39 meteorological stations within the study area. The data was preprocessed, organized, interpolated, and analyzed using Python and ArcGIS Pro. Provide an overview of the data used in your analysis. Include data sources, types of data, any preprocessing or cleaning steps, and the rationale for selecting these data sources. Be specific about variables, timeframes, and resolution, if applicable.

2.1. Study Area

The study area, Lumbini Province, is located in the southern part of Nepal and encompasses diverse topographical features, including lowlands (Terai) and highlands (Hill regions). The variation in elevation significantly influences the region's rainfall patterns. Figure 1 illustrates the geographical context of the study area, including the distribution of meteorological stations.

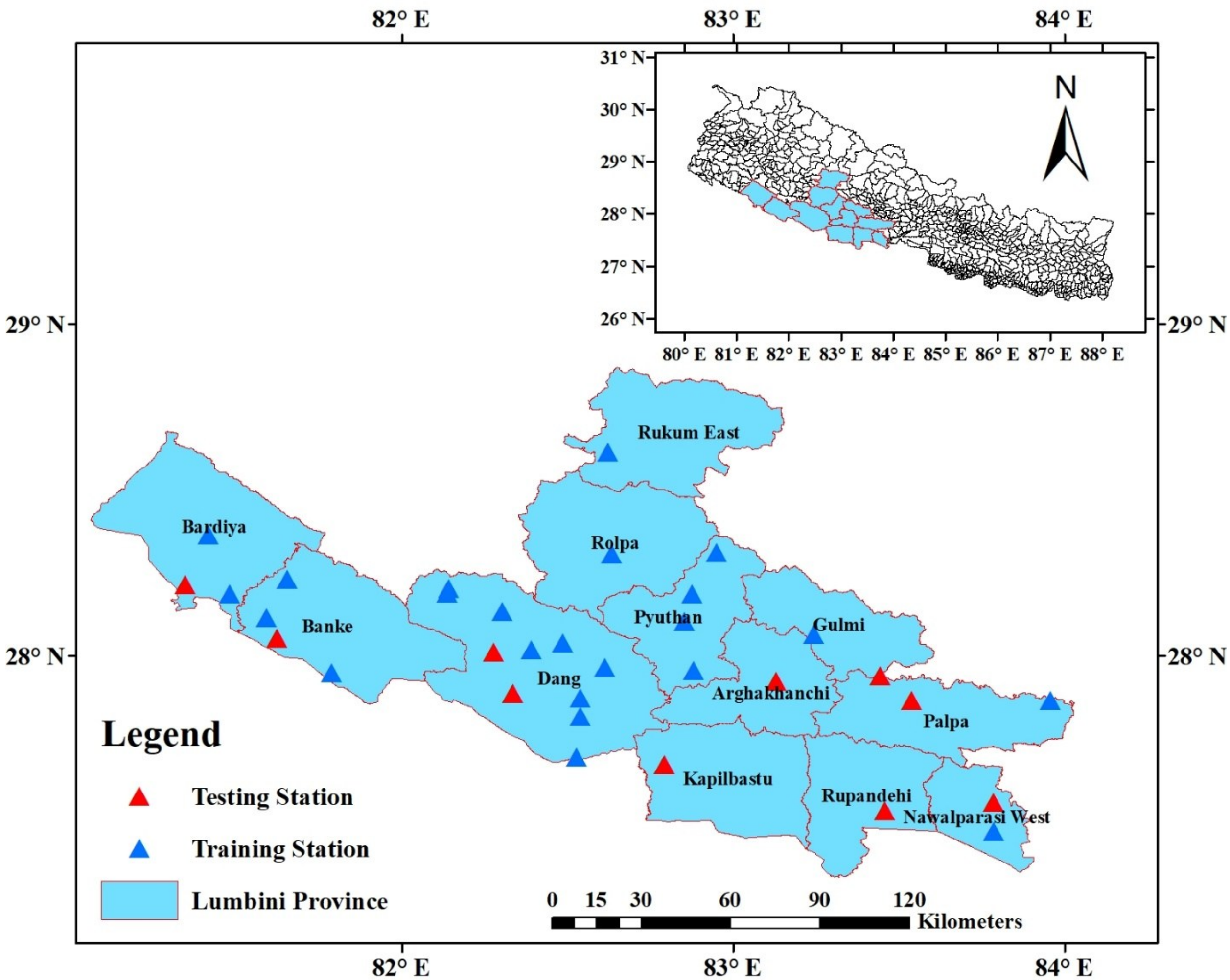


Figure 1: Map of the study area in Lumbini Province, Nepal, showing the Rainfall data stations

The table below shows the data sources and types used in this study, along with a brief description of the data collection and analysis methods.

Table 1: List of Data used in the Study

Data name	Data source	Description
Rainfall Data	Department of Hydrology and Meteorology (DHM), Nepal	24-hour cumulative rainfall data from 39 stations across Lumbini Province (2002-2023)
Station Locations	DHM, Nepal	Geographic coordinates of 39 meteorological stations across the lowlands and highlands of Lumbini Province
Shape file	Open dataset, Nepal	The boundary line of Nepal with each districts

The rainfall data was split into 75% training and 25% testing datasets. Outliers were removed using Python, and the training data was used for interpolation. Kriging, RBF, and IDW methods were applied with various parameters, and cross-validation was performed using the testing data. The optimal interpolation method was selected based on the highest R^2 and lowest RMSE values. This method was then used to analyze the seasonal variation of rainfall.

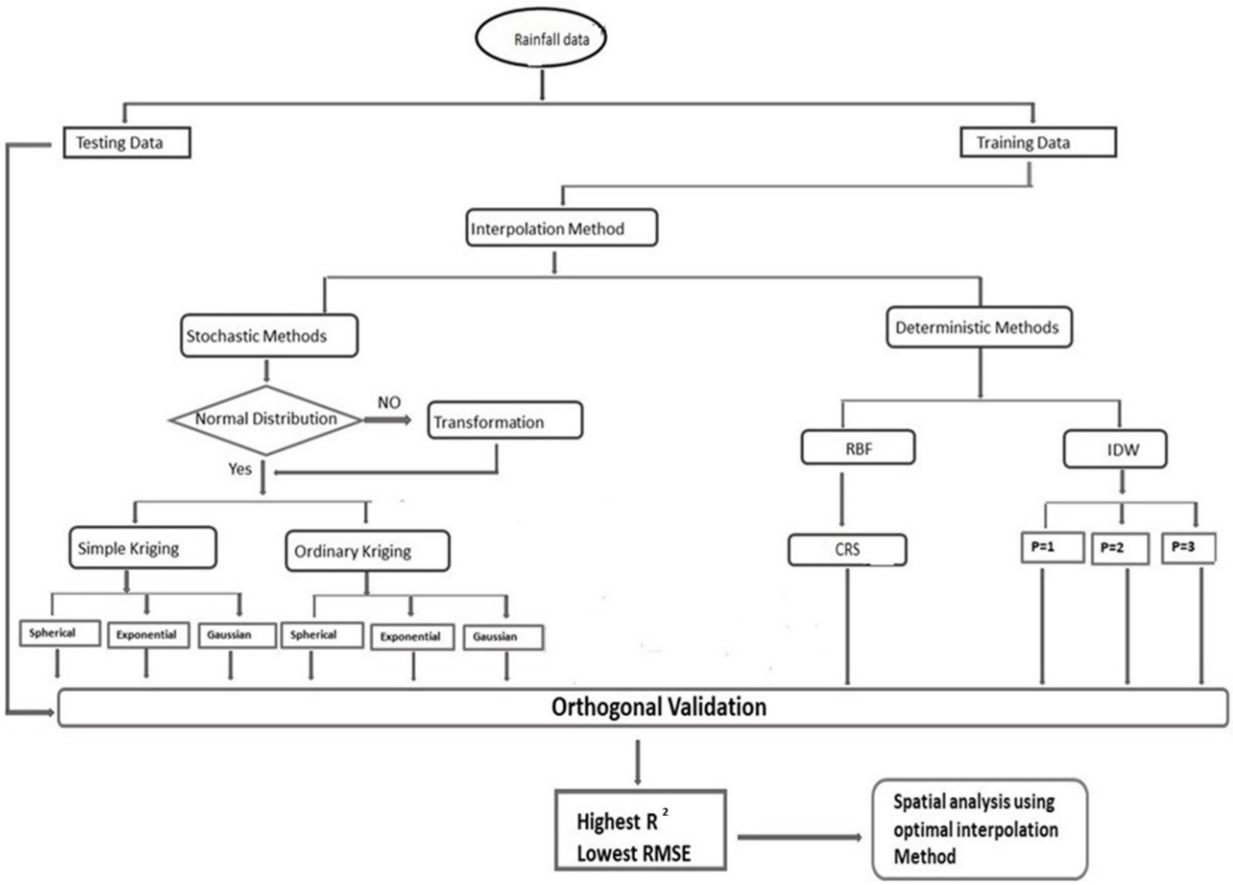


Figure 2: Workflow chart showing the methodology of the entire project

RESULTS

The study analyzed data from 39 rainfall stations spanning 22 years. A Kolmogorov-Smirnov (K-S) test confirmed that the data was not normally distributed. A Box-Cox transformation (λ=1) ensured normalization, resulting in skewness and kurtosis values of 0.0515 and 0.328, respectively.

Performance Analysis of Interpolation Methods

Different stochastic and deterministic model are cross-validated using the training data.

A. Cross Validation using Inverse Distance Weighting (IDW)

Cross-validation was performed on IDW models with powers 1, 2, and 3. The results indicate that IDW with power 1 provided the best fit, achieving an RMSE of 35.68 and an R^2 value of 0.54.

Table 2: Spatial Interpolation Data by IDW Method

Model	Power	RMSE	R^2
IDW	1	35.68	0.54
IDW	2	40.63	0.36
IDW	3	38.57	0.46

B. Cross Validation Using Radial Basis Function (RBF)

The Completely Regularized Spline (CRS) model was employed for RBF interpolation.

Table 3: Spatial Interpolation Data by RBF Method

Model	Semivariogram	RMSE	R^2
RBF	CRS	29.38	0.57

C. Cross-Validation of Stochastic Models Using Ordinary(OK) and Simple Kriging (SK)

Ordinary Kriging (OK), the exponential semivariogram performed best with an RMSE of 26.35 and R^2 of 0.61, outperforming the Gaussian model due to a smaller nugget effect. For Simple Kriging (SK), the Gaussian semivariogram was optimal with an R^2 of 0.36 and RMSE of 37.62. The nugget-to-sill ratio (0.32) indicates moderate correlation, aligning with findings by (Adhikary ,2014) for New Delhi, India. Spatial rainfall variation is influenced by factors such as temperature, winds, ocean currents, and terrain, which contribute to the moderate correlation observed.

Table 4: Spatial Interpolation Data by OK

Semi variogram	Spherical	Exponential	Gaussian
Nugget (Co)	14.41	16.81	15.63
Sill ($Co + C$)	33.5	52.83	41.26
Nugget/Sill	0.43	0.31	0.37
R^2	0.29	0.61	0.36
RMSE	31.25	26.35	38.26

Table 5: Spatial Interpolation Data by SK Method

Semi variogram	Spherical	Exponential	Gaussian
Nugget (Co)	19.48	16.82	16.98
Sill ($Co + C$)	37.53	40.59	43.24
Nugget/Sill	0.51	0.41	0.39
R^2	0.31	0.28	0.36
RMSE	40.86	45.52	37.62

Validation of Best Cross Validated Methods

Cross-validation revealed that the IDW (power 1), RBF (CRS), OK (exponential), and SK (Gaussian) models yielded the highest R^2 and lowest RMSE for their respective methods. These models were applied to validation datasets (25% of the total) to determine the most optimal method for rainfall mapping. Among all methods, Ordinary Kriging (OK) with the exponential semivariogram performed best, achieving the highest R^2 (0.48) and lowest RMSE (26.35), making it the most suitable for predicting rainfall distribution in the study area

Table 6: Validation Metrics for Testing dataset

Model	Power/Semivariogram	R^2	RMSE
OK	Exponential	0.48	26.35
SK	Gaussian	0.25	37.25
IDW	1	0.36	35.68
RBF	CRS	0.41	29.38

Visualization of Prediction

The spatial distribution of rainfall was visualized using the OK-Exponential model.

- Low Rainfall: Areas with rainfall between 20,000–35,000 mm are predominant in the region (weight = 1).
- Moderate Rainfall: Areas receiving 35,000–46,000 mm of rainfall are concentrated in specific zones (weight = 2).
- High Rainfall: The northeastern areas exhibit the maximum rainfall (weight = 3), reflecting the influence of topography and climatic factors.

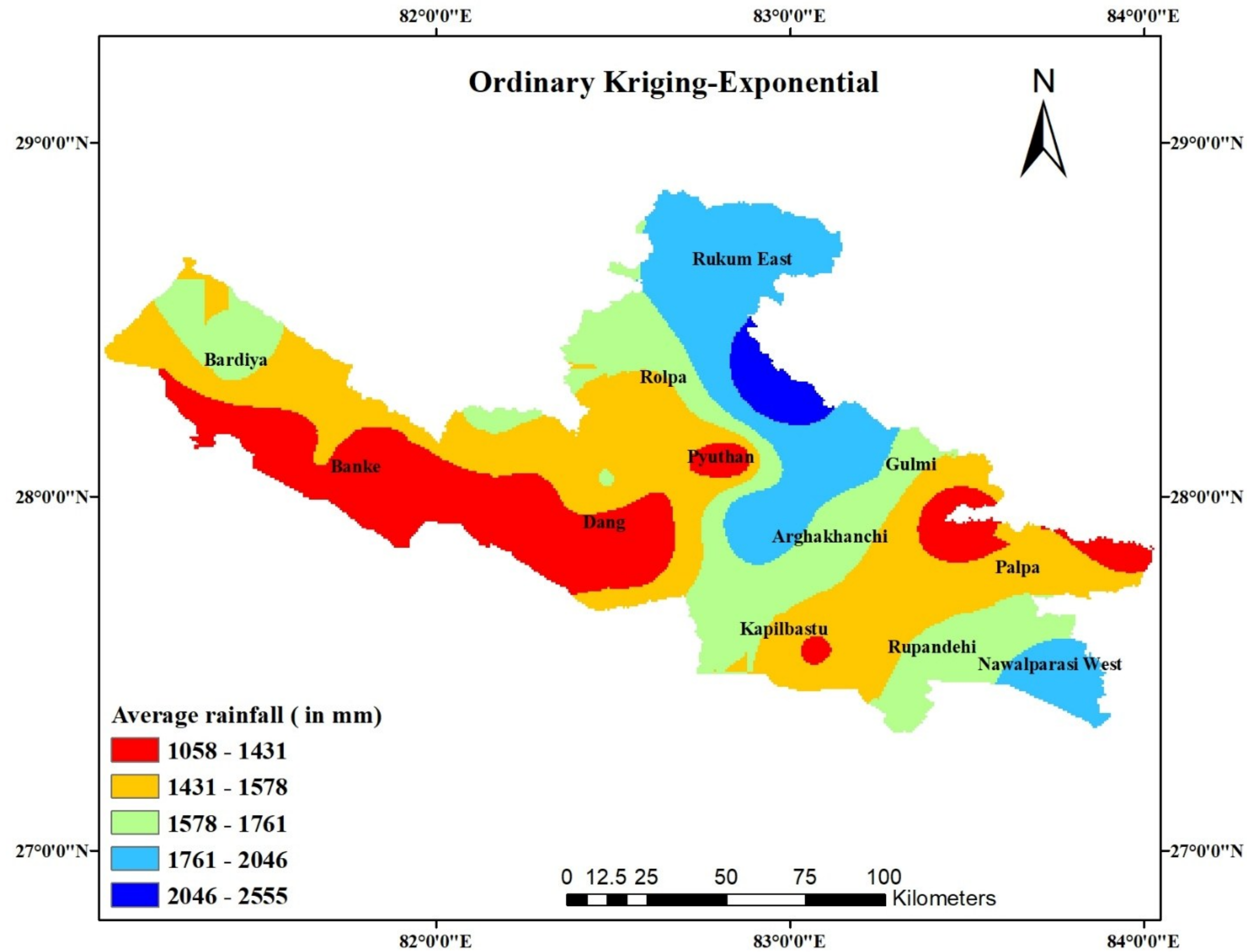
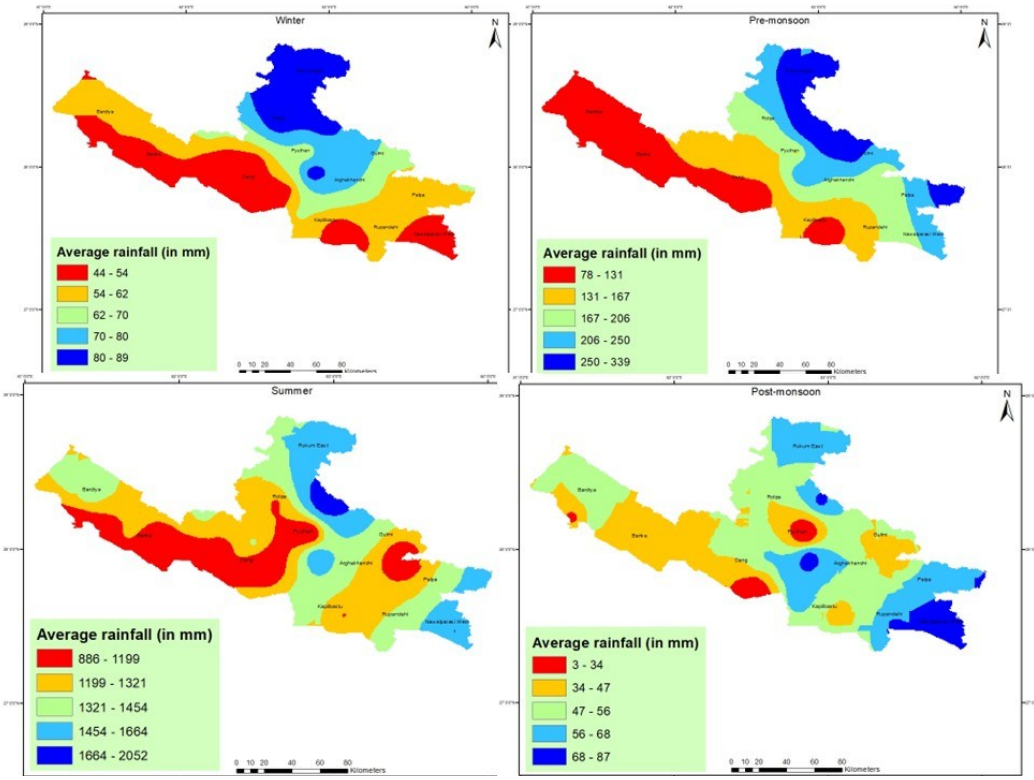


Figure 3: The spatial distribution of rainfall using the optimal interpolation (Ordinary Kriging)

Seasonal Rainfall Distribution

Using the best interpolation method (OK), the seasonal rainfall distribution over 22 years in Lumbini Province was analyzed across four seasons:

- Winter (Dec-Feb): Rainfall ranged from 892–2423 mm, highest in the northeast, decreasing southwest.
- Pre-Monsoon (Mar-May): Rainfall varied between 1726–9713 mm, with a similar north-to-south decreasing trend.
- Summer Monsoon (Jun-Sep): Rainfall ranged from 20,028–44,534 mm, peaking in the northeast and tapering southwest.
- Post-Monsoon (Oct-Nov): The driest season, with rainfall between 66–1919 mm, following the same spatial pattern.



1

Figure 4: The Seasonal spatial distribution of rainfall using the optimal interpolation (Ordinary Kriging)

CONCLUSIONS

This study aimed to analyze the spatial distribution of rainfall in Lumbini Province using various interpolation techniques, focusing on IDW, RBF, OK, and SK methods. Ordinary Kriging (OK) was identified as the optimal method due to its superior performance in validation and cross-validation, achieving the lowest RMSE and highest R^2. Seasonal analysis revealed a decreasing rainfall trend from the north to the south, with the monsoon season exhibiting the highest precipitation.

The findings contribute to a better understanding of spatial rainfall patterns, supporting improved water resource management and hydrological modeling. This research highlights the potential for applying advanced geostatistical techniques to regional rainfall distribution studies, offering insights relevant to flood management, agriculture, and climate monitoring.

Limitations:

- The correlation (R^2) and RMSE metrics showed moderate accuracy, indicating room for improvement in model performance.

- Limited availability of high-resolution data reduced the precision of spatial rainfall distribution results.

- The analysis was constrained by the lack of additional factors such as temperature, moisture, or soil properties, which may have improved interpolation accuracy

Future work:

- Explore more sophisticated methods like Bayesian Kriging or hybrid interpolation models to improve accuracy.

- Incorporate additional meteorological and geological variables to better understand rainfall distribution dynamics.

- Utilize high-resolution temporal and spatial datasets to enhance the reliability of spatiotemporal mapping.

REFERENCES

Smith, J. (2022). Geospatial Methods in Environmental Research. New York, NY: Academic Press.

Dingman, S. L. (2015). *Physical hydrology*. Waveland press.

Jones, L., Ludi, E., & Levine, S. (2010). Towards a characterization of adaptive capacity: a framework for analyzing adaptive capacity at the local level. *Overseas Development Institute, December*.

Isaaks, E. H. (1989). Applied Geostatistics.

Adhikary, P. P., Dash, C. J., Kumar, G., & Chandrasekharan, H. (2014). Characterization of groundwater quality for irrigation and drinking purposes using a modified groundwater quality index. *Indian Journal of Soil Conservation*, *42*(3), 260-267.

https://www.dhm.gov.np/

https://opendatanepal.com/dataset/new-political-and-administrative-boundaries-shapefile-of-nepal