Groundwater Quality and Economic Growth: An Analysis

Anoushka Kumar (2021133)
Dhruv Prakash (2022168)
Lakshay Trehan (2022267)
Nishant Singh (2022595)

Introduction

Background

- India's economic growth has placed immense pressure on its natural resources such as groundwater
- Groundwater is highly susceptible to contamination
- Understanding the impact of economic growth on groundwater quality is essential for sustainable growth

Our project:

- Explores the relationship between the concentration of groundwater indicator sulfate and economic growth
- Carried out using regression analysis on varying subsets of data

Data and Economic tests

Summary Statistics

	Year	Ginivalue	SDP_NORM	Sulfate	State
Count	9713	9531	9681	5496	9713
Unique	–	–	–	–	30
Top	–	–	–	–	Uttar Pradesh
Frequency	–	–	–	–	1397
Mean	2008.96	0.2693	362246.6	98.16	–
Std	5.47	0.0571	302153.8	125.94	–
Min	2000	0.16	1247.87	0	–
25%	2004	0.23	148760	24.21	–
50%	2009	0.26	282371	52.96	–
75%	2014	0.31	488554.2	122.16	–
95%	2018	0.37	949184	346.65	–
Max	2018	0.48	1728578	1815	–

Data:

- Mismatch in some names
- Null values (4217 in sulfate)
- Outliers (120)
- Preprocessed the data and ran regression both with and without outliers

Economic tests:

- F test and p-value
- R- square
- Regression coefficients

Simple Regression Results

Table 1: OLS Regression Results

Variable	Coefficient	Std Error	t	P $>\|t\|$
const	9.572e+01	2.693e+00	35.539	<2e-16
SDP_NORM	6.182e-06	5.236e-06	1.181	0.238

1. **MSE**: 15876.3553237917
2. **R-squared**: 0.0002541
3. **Adj. R-squared**: 7.18e-05
4. **F-statistic**: 1.394
5. **No. Observations**: 5485
6. **Df Residuals**: 5483
7. **P-Value**: 0.2378

- p-value of 0.238>0.05 indicates that the model **with outliers** does not have significance at the 95% significance level
- Small R-square (0.0002) indicates that only 0.2% of variation is explained by the model
- Better results are obtained by pre-processing the data

Simple Regression Results

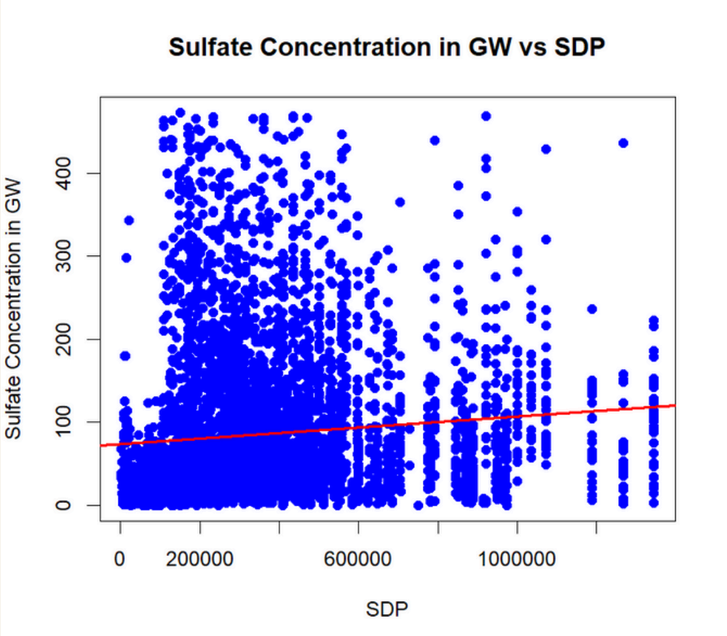
Table 2: OLS Regression Results

Variable	Coefficient	Std Error	t	P $>\|t\|$
const	$7.377e+01$	$2.180e{+}00$	33.844	$< 2e-16$
SDP_NORM	$3.333e-05$	$4.743e-06$	7.028	2.38e−12

1. **MSE**: 8421.66829947241
2. **R-squared**: 0.009586
3. **Adj. R-squared**: 0.009392
4. **F-statistic**: 49.39
5. **No. Observations**: 5105
6. **Df Residuals**: 5103
7. **P-Value**: 2.375e-12

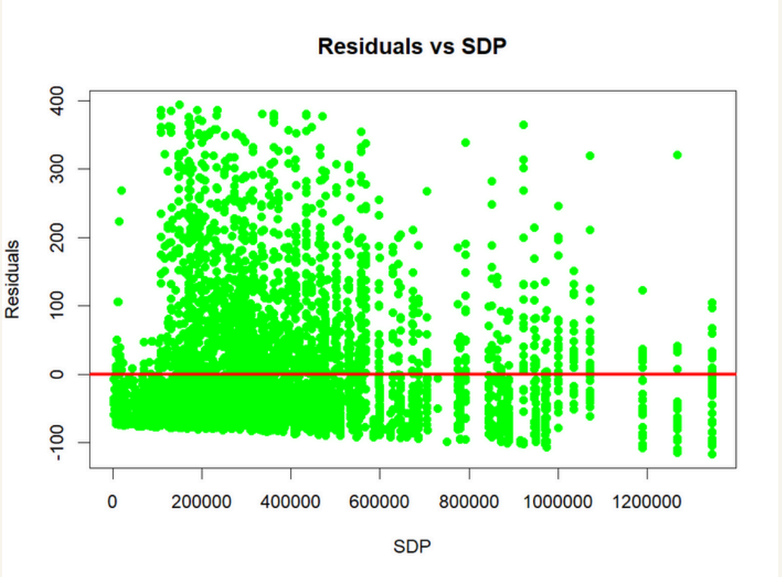
- Positive coefficients indicate a positive association with the groundwater quality indicator
- Small p-value associated with the F- test post-processing of data indicates a high significance of the model
- R- square of 0.009 indicates that only 0.9% of the variance can be explained using this model

Residuals

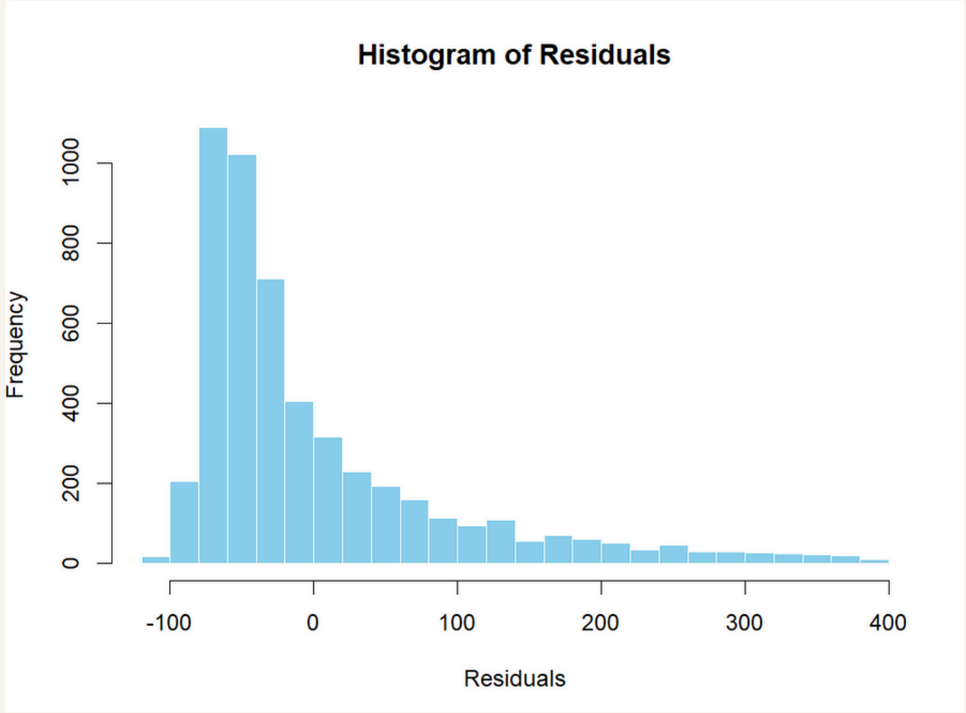


- Red line: regression line
- Points are distributed about the regression line
- A clear linear relationship is not observable, which indicates that a linear model might not be the best fit

- Red line: line at y=0
- Points are randomly distributed about the line
- No clear trend, which indicates that the model is adequate at capturing the relationship between X and Y



Residuals



- sum of residuals: -9.75548530846027e-11

- the sum is an extremely small value, approaching 0, thus verifying the key assumption that ∑ûi = 0.

Kuznet Curve and Outliers

$$GWQ = \beta 0 + \beta 1(SDP) + \beta 2(SDP2) + \beta 3(SDP3) + \beta 4(GINI) + u$$

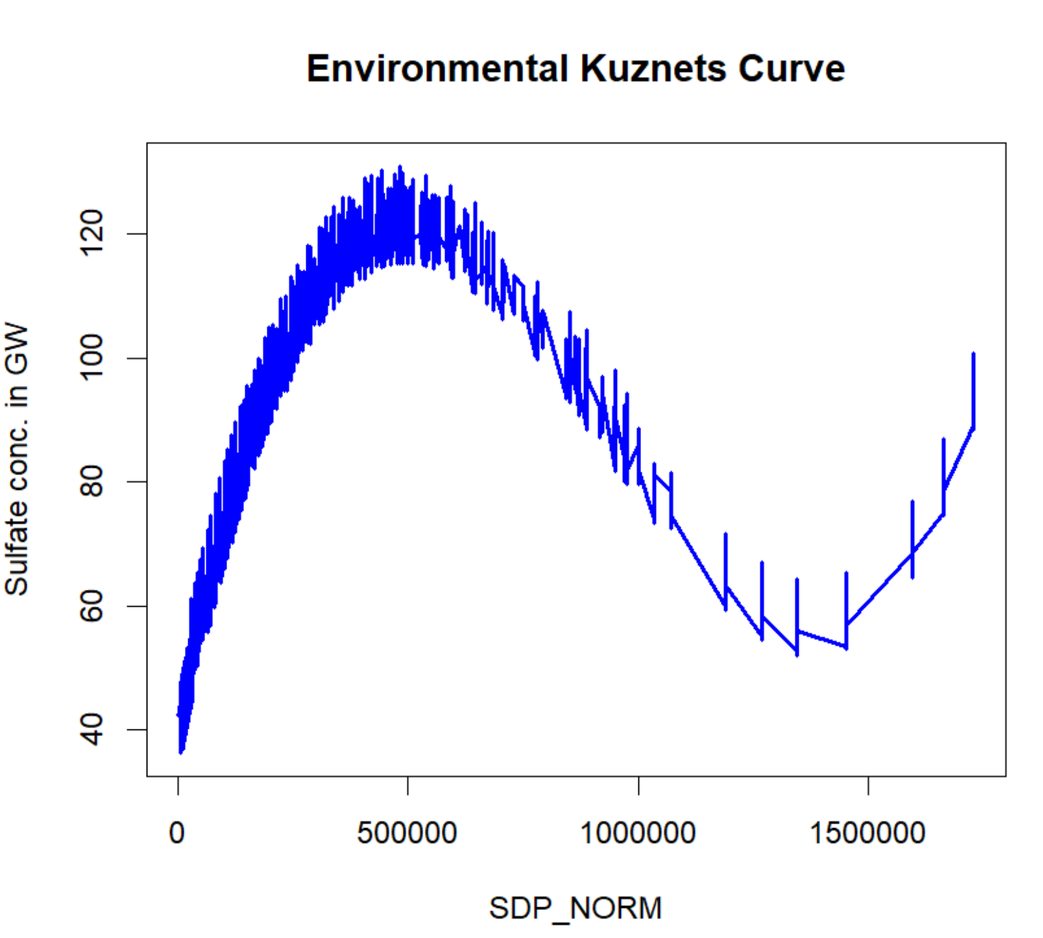
Table 4: Kuznets Curve Regression Results

Variable	Estimate	Std. Error	t value	P $>\|t\|$
(Intercept)	$1.741e+00$	$7.464e+00$	0.233	0.81558
SDP_NORM	$4.217e-04$	$3.136e-05$	13.448	$2e-16$
ginivalue	$7.288e+01$	$2.283e+01$	3.193	0.00142
SDP_sq	$-6.510e-10$	$6.170e-11$	-10.550	$2e-16$
SDP_cube	$2.862e-16$	$3.406e-17$	8.403	$2e-16$

1. **MSE**: 8091.5022825187
2. **R-squared**: 0.04841
3. **Adj. R-squared**: 0.04767
4. **F-statistic**: 64.87
5. **No. Observations**: 5105
6. **Df Residuals**: 5100
7. **P-Value**: < 2.2e-16

- Outliers= 121
- Influential Observations = 126
- p-value: < 2.2e-16 shows the significance of the model
- Positive coeff. of SDP, SDP_cube and Gini show positive association with the sulfate conc. in GW
- Negative coeff of SDP_sq shows its inversely related to sulfate conc. in GW
- P-values of coefficients show their significance

Kuznets Curve



- N shaped EKC curve due to trends in coefficients

Year-wise and Region-wise Results

Region	R-squared	F-statistic	p-value
North	0.09104	32.78	2.2e-16
South	0.05609	13.41	1.264e-10
Northeast	0.1133	12.17	2.558e-09
Central	0.03136	11.26	5.431e-09
West	0.02061	4.297	0.001905
East	0.0581	6.399	5.274e-05

- Small p values for all regions indicate that the model is significant for all regions
- Significance varies from region to region, lowest for West and highest for North

- Relation between economic growth and sulfate conc. in GW varies with year
- Magnitude of coefficient indicates the amount of change
- Sign indicates the direction and nature of the relationship

Year	SDP_NORM	Gini	SDP_sq	SDP_cube
2000	1.099e-03	-7.231e+01	-5.398e-09	6.517e-15
2003	-7.890e-02	-2.533e+02	6.515e-07	-1.371e-12
2007	8.656e-04	1.020e+02	-1.741e-09	9.744e-16
2010	3.810e-04	-2.958e+01	2.245e-10	-9.317e-16
2014	4.191e-04	3.999e+01	-3.744e-10	8.914e-17
2017	2.799e-04	1.942e+01	-1.481e-10	2.194e-18

Limitations

- **Data Limitations:** the analysis relies on the data sources, which may have limitations in terms of accuracy, completeness and consistency.
- Factors like **regulations**, **government policies**, **institutions** also have an affect on the GWQ which we have not considered.
- GWQ is measured by the presence of various minerals in water but we are just analyzing the trend of the presence of a single salt (i.e. Sulfate) in ground water of different districts in India.

References

[1] Wang, H., & Zhang, Q. (2019). Research Advances in Identifying Sulfate Contamination Sources of Water Environment by Using Stable Isotopes. International journal of environmental research and public health, 16(11), 1914. https://doi.org/10.3390/ijerph16111914

[2] Frost, J. (n.d.). Interpreting the F-test of overall significance in regression analysis. Statistics by Jim. Retrieved from https://statisticsbyjim.com/regression/interpret-f-test-overall-significance-regression/

[3] Wooldridge, J. M. (2012). Introductory econometrics: A modern approach (7th ed.). Cengage Learning.

[4] Patrícia Hipólito Leal and António Cardoso Marques (2022) The evolution of the environmental Kuznets curve hypothesis assessment. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9668524/

[5] Alexandra Allard, Johanna Takman (2018) The N-shaped environmental Kuznets curve
Retrieved from: https://link.springer.com/article/10.1007/s11356-017-0907-0

[6] Sharma, M. K., & Kumar, M. (2020). Sulphate contamination in groundwater and its remediation: an overview. Environmental monitoring and assessment, 192(2), 74. https://doi.org/10.1007/s10661-019-8051-6