AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Spatial Auto-correlation in Groundwater Levels

Saif Ali, Department of Social Sciences, Indraprastha Institute of Information Technology, Delhi
saifa@iiitd.ac.in
Gaurav Arora, Department of Social Sciences, Indraprastha Institute of Information Technology, Delhi
gaurav@iiitd.ac.in

Selected Poster prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Copyright 2022 by [authors]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Spatial Autocorrelation in Groundwater Levels

Saif Ali, Gaurav Arora
Department of Social Science and Humanities, Indraprastha Institute of Information Technology, Delhi (IIIITD)

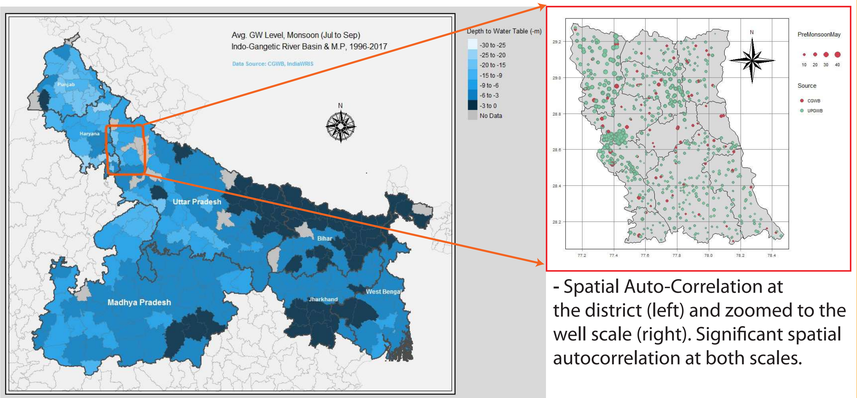
Aim and Scope

To characterize spatial auto-correlation in groundwater levels using a functional instrument called a semivariogram and use it to perform spatial prediction.
To estimate the range within which groundwater levels are spatially auto-correlated. This is useful for determining groundwater "localities" for management of groundwater resources.

Conclusions

Semivariogram was estimated using groundwater levels data for Uttar Pradesh, India in the year 2009, post-monsoon (Aug). Groundwater levels were kriged to get a high-res map for an agro-climatic sub-region of UP and were auto-correlated within a 27 km radius. There was high correlation (0.81) between predicted and observed values .

Spatial auto-correlation implies that groundwater level at a location is more similar to that at nearby locations than further away. The opposite is spatial randomness.



- Spatial Auto-Correlation at the district (left) and zoomed to the well scale (right). Significant spatial autocorrelation at both scales.

- Arises from sub-surface water flows, well interference, spatially-con tiguous climatic conditions and land use land cover (LULC)..
- Can be exploited to perform spatial prediction.
- Needs to be formally characterized in a given realization of data.
- Violates classical assumptions in econometrics if left unaccounted.

Spatial dependence model characterizes the groundwater level at each location in a domain *D* as a realization of a random variable which has a (global) spatial trend component and a (local) spatial dependence component.

A **stochastic spatial process** is a family of indefinitely many random variables indexed by location in a spatial domain *D* in R^2 expressed as *{Y(w) : w in D}* where *Y(w)* refers to the family of random variables and *w* to an arbitrary location in *D* [2].

Intrinsic spatial stationarity requires a constant mean and that the variance of the differences between two random variables separated by spatial lag *h* depend only on *h* and not on absolute positions.

$$E[Y(w)] = \mu \quad \text{for all} \quad w \text{ in } D$$

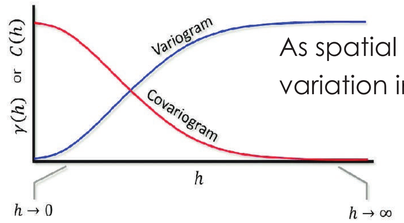
$$Var[Y(w+h) - Y(w)] = 2\gamma(h) \quad \text{for all} \quad w \text{ in } D \text{ and } h > 0$$

where *γ(h)* is called the **"semivariogram"** of the stochastic spatial process [2,3].

References
1. Bárdossy, A. (1997). Introduction to geostatistics. Institute of Hydraulic Engineering, University of Stuttgart.
2. Cressie, N., Cressie, A., & Sons, J. W. . (1991). Statistics for spatial data. J. Wiley. Retrieved from https://books.google.co.in/books?id=k_JQAAAAMAAJ
3. Cressie, N., & Moores, M. T. (2021). Spatial statistics. arXiv preprint arXiv:2105.07216.

Spatial Characterization: the Semivariogram is a measure of spatial *dissimilarity*. It is the inverse of the "covariogram" which is a measure of spatial similarity.

$$\gamma(h) = C(0) - C(h)$$

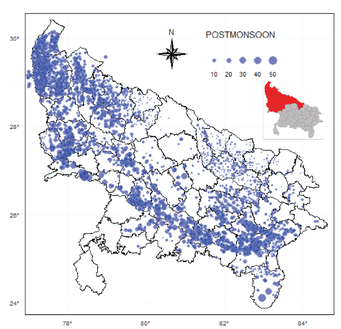


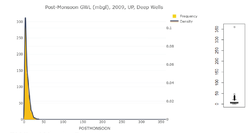
As spatial lag increases, variation in realizations increases.

Semivariogram Estimation and Model Fitting

We estimate a semivariogram from the Western half of the state of Uttar Pradesh. This sub-region is considered distinctly from the Eastern half because of greater uniformity of agro-climatic and hydro-geological conditions and groundwater levels in this region can be reasonably considered as being the realizations of a single stochastic spatial proces.

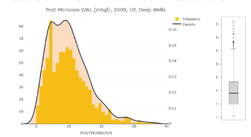
Groundwater Level (mbgl), 2009, Post-Monsoon, Uttar Pradesh (Deep Wells, NA omitted)





GWL - Whole State

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.03	3.0	5.4	7.5	10.0	360.0	26437



GWL - Western Half

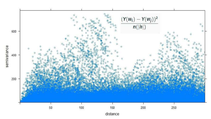
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.4	5.0	9.1	10.0	13.4	36.3

Step 1: Variogram cloud

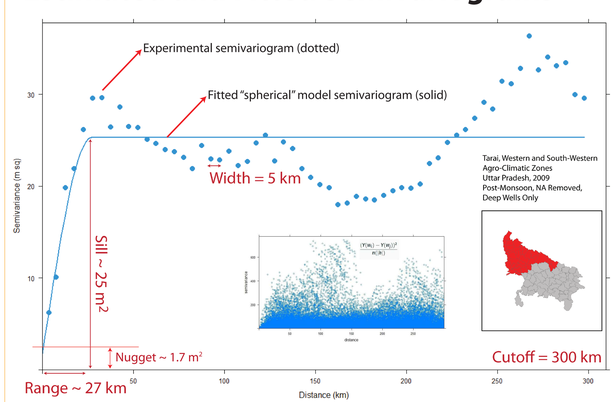
Classical estimator for an isotropic semivariogram: (N. A. Cressie, 1993).

$$2\hat{\gamma}(h) = \sum_{|h|} \frac{(Y(w_i) - Y(w_j))^2}{n(|h|)} \quad (5)$$

$|h|$ - all pairs of wells w_i, w_j distance *h* apart
$n(|h|)$ number of such wells



Estimated and Fitted Semivariograms



Step 2: Experimental Semivariogam

The experimental semivariogram is obtained by binning the points in the cloud and taking expected values within each bin.

Step 3: Fitted Model Semivariogam

A closed form model is fitted to the experimental semivariogram using either MLE, restricted MLE or OLS fitting criteria (Cressie, 1993).

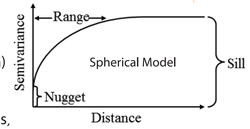
A fitted model can be used to perform spatial prediction as it provides a measure of variance at arbitrary values of spatial lag *h*.

The important parameters of a fitted semivariogram are the nugget, sill and range which are defined below-left and visualized in the figure to the left and below.

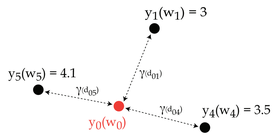
Spherical Semivariogram Model (Cressie, 1993)

$$\gamma(\mathbf{h}) = \begin{cases} 0, & \mathbf{h} = \mathbf{0}, \\ c_0 + c_s[(3/2)(\|\mathbf{h}\|/a_s) - (1/2)(\|\mathbf{h}\|/a_s)^3], & 0 < \|\mathbf{h}\| \le a_s, \\ c_0 + c_s, & \|\mathbf{h}\| \ge a_s, \end{cases}$$

c_0 - Nugget effect (micro-scale variation)
c_s - Partial Sill (max. variance)
a_s - Range (distance of max. variance)
h - Spatial lag (separation between wells,



Spatial Prediction: the Kriging Estimator



$$y_0 = \sum \lambda_i y_i$$

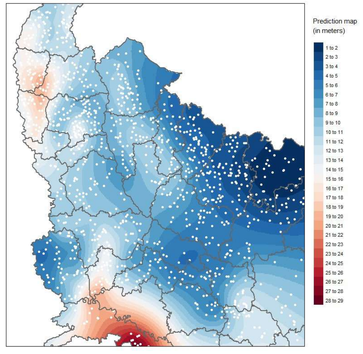
$$\lambda_i = \gamma(d_{0i})^2 / \sum \gamma(d_{0i})^2$$

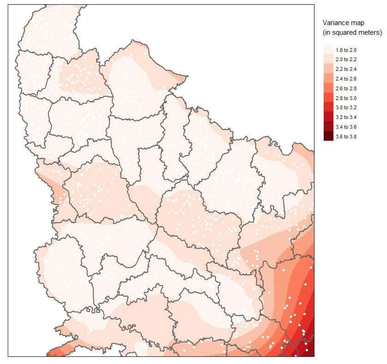
Predict unknown value as a weighted sum of neighboring values using semivariogram, subject to:

1. **Unbiasedness**
$E[\hat{Y}(w0)] = E[Y(w0)] = \mu$
Satisfied if $\sum \lambda_i = 1$ and the mean is stationary i.e $E[Y(wi)] = \mu$ for all *i*, where μ is a constant.
2. **Minimum Variance**
Find weights that minimize prediction variance, σ^2 subject to unbiasedness constraints
$\sigma^2 = Var[\hat{Y}(w0) - Y(w0)]$

Kriging Vs Deterministic Spatial Interpolation

- Accounts for data redundancy due to spatial clustering.
- Provides confidence interval for estimates.

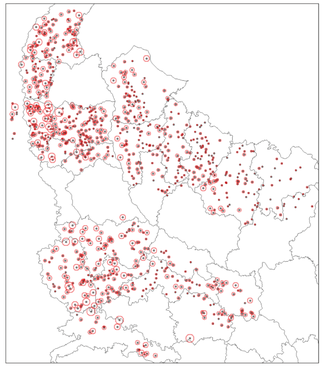




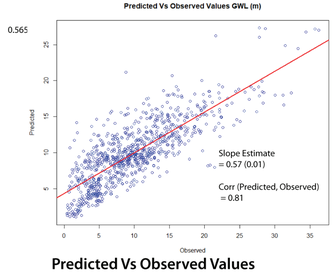
Above Left: spatial prediction of groundwater levels using administrative data for UP, 2009. Right: Prediction variance. White dots represent the wells where observations were made by the monitoring agency. Source of data: UPGWD

Variogram Model Cross Validation

Spherical Variogram Model, Range ~ 27 km, Sill ~ 25 m², Nugget ~ 1.7 m²



Spatial Distribution of Residuals,



Predicted Vs Observed Values