Assessing common interpolation methods (Splining and Kriging) to model and predict heavy metal contamination (lead:Pb) in the Village of Stein, on the Meuse River, Netherlands

Introduction:

Scientists of all stripes have long sought the ability to create accurate prediction models or surfaces to predict (or interpolate) data that they do not possess from an existing set of data. Kriging and Splining are two common methods of interpolation, which are used to predict unknown values at other locations. Splining is a 'deterministic' interpolation method, meaning that the method does not indicate how certain the output is. For Splining and other common deterministic interpolation methods such as Inverse Distance Weighting (IDW), the output prediction '*is what it is*' based on the existing data, with no associated confidence value or accuracy estimation.

The spline interpolation method forces a smoothed curve through a set of known input points to estimate the unknown values. The 'regularized' spline method creates a smooth, gradually changing surface with values that may lie outside the data range. The 'tension' spline method creates a less-smooth surface with values more closely constrained by the sample data range (see Fig. 2). Both methods can be refined by changing the number of points used in the calculation of each interpolated cell or by changing the weight or 'coarseness' of the output.

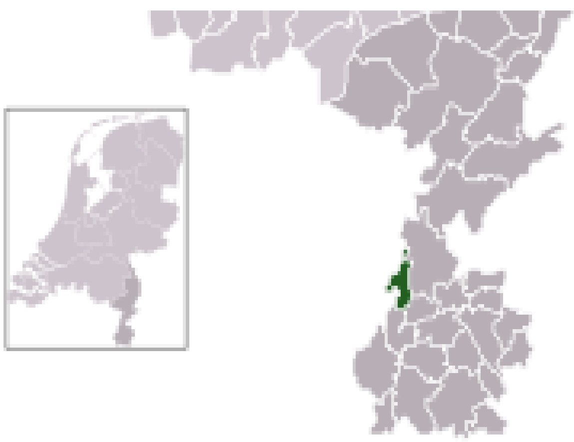

Fig. 1: Location of City of Stein, Limburg Province, NL
(https://en.wikipedia.org/wiki/Stein,_Limburg)

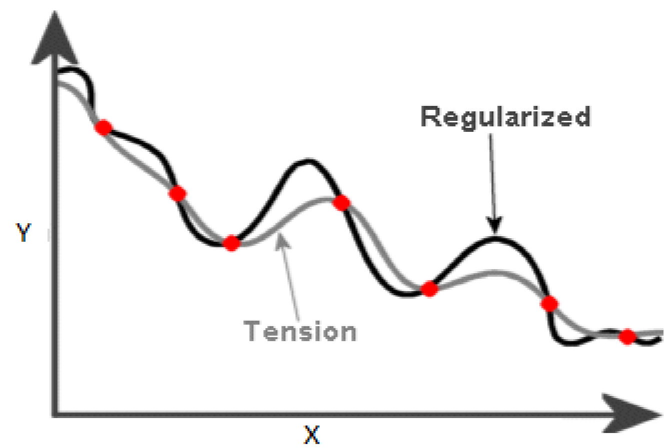

Fig. 2: generalized representation - regularized vs tension Splining methods
(https://www.neonscience.org/resources/learning-hub/tutorials/spatial-interpolation-basics)

Kriging originated in the field of mining geology and is named after South African mining engineer Danie Krige. Kriging, as opposed to Splining and IDW, is a 'geostatistical' interpolation method. Kriging offers a powerful interpolation method that can provide an optimal prediction surface while also delivering a measure of confidence of how accurate the prediction will be (see Figure 3).

1. Prediction surface:
   this surface predicts the variable that you are interpolating or kriging.
2. Error surface:
   Since geostatistics are used in the kriging method, the error surface can display the following:
   - standard error of prediction: generally higher with less input data
   - probability: the probability surface can highlight exceedances of a set threshold value
   - quantile: error surface can represent best or worst-case scenarios as 99th percentile.

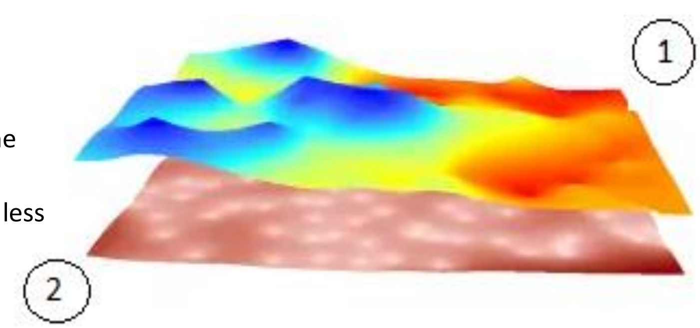

Figure 3: graphical display of the general kriging method outputs
(https://gisgeography.com/kriging-interpolation-prediction)

Kriging differs from simpler interpolation methods in that it generates estimates of uncertainty surrounding each interpolated value. This is extremely useful when the underlying dataset exhibits some degree of spatial autocorrelation, as geographic spatial data often does. This experiment uses ordinary Kriging, which assumes there is stationarity (constant mean and variance) across the entire spatial field, not often encountered in the real world.

As opposed to simpler deterministic methods such as Splining, Kriging is a multi-step process including exploratory statistical analysis of the data, semi-variogram modelling, creating a polynomial to describe the prediction surface, and exploring an error surface. It is often used in soil science, geology, mining, meteorology, and other environmental sciences. We use Splining and ordinary Kriging in this experiment to create an accurate prediction of soil lead concentrations in the City of Stein, Netherlands. We are using the well-known 'Meuse' geostatistical dataset of soil characteristics and contaminant concentrations along the Meuse River (see Figs. 1 & 4).

The Meuse dataset is important in the field of soil science, as the City of Stein in Limburg Province, Netherlands, has heavy metal contamination along the banks of the Meuse River (see Figure 4). This heavy metal contamination is known to be deposited along the river and can be further distributed and dispersed along with the frequent flooding of the Meuse River that can bring contamination from industrial areas of Belgium. The upper reaches of the Meuse River in Belgium are known as *'le Silon Industriel'* and are home to heavily polluting industries such as coal, iron and steel, and petroleum refineries. As the Meuse River crosses into southern Netherlands in the area of Stein, it can carry heavy metal contaminants (lead, cadmium, copper, zinc) along its route to the eventual outflow into the North Sea. The largest metal accumulations occur in low-lying floodplain sections of the Meuse River (Middlekoop, 2000).

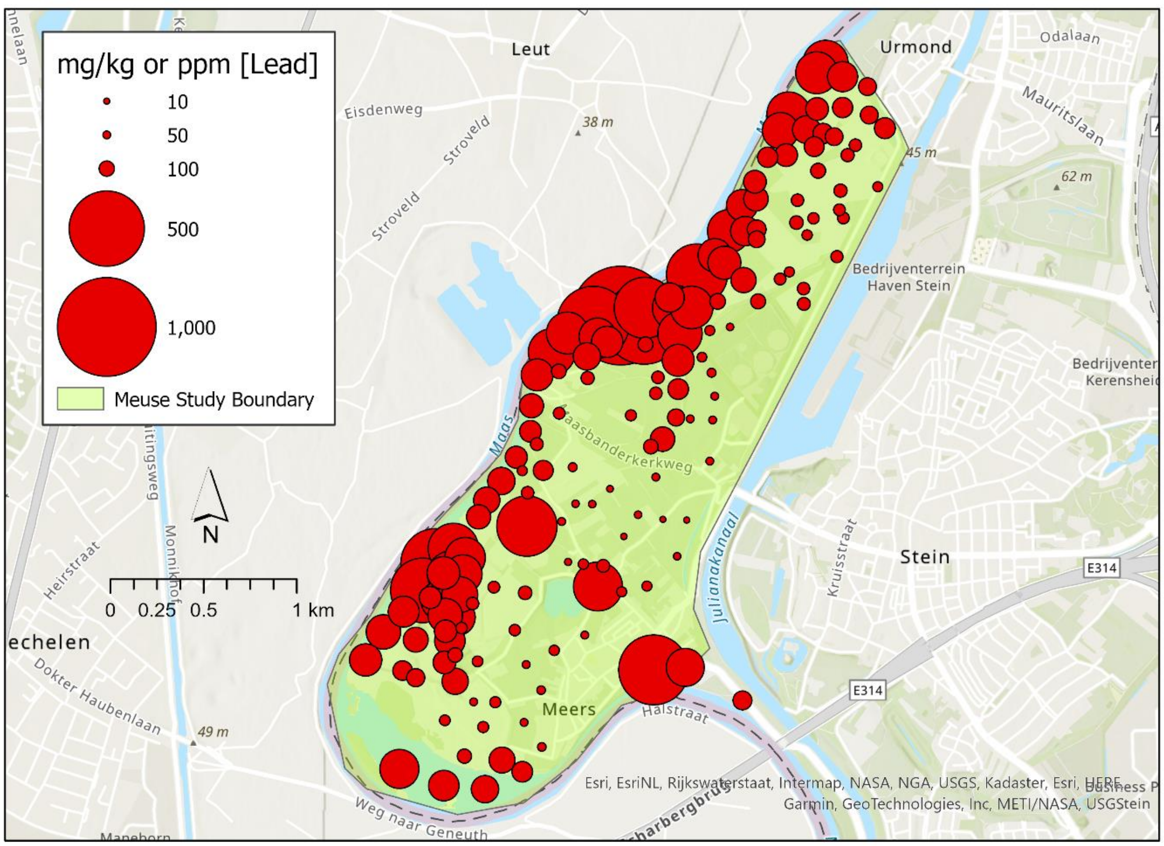

Figure 4: Observed Lead concentration ranges in the Meuse Study Area along the Meuse River banks in Stein, NL

Methods:

In order to assess deterministic (Splining) and geostatistical (Kriging) interpolation methods for our Meuse dataset, ArcGIS Pro built-in geostatistical analysis and spatial analyst tools were used (see Figure 5). The Meuse Contaminants and Meuse Boundary shapefiles were loaded into ArcGIS Pro as layers, with the contaminants dataset being subdivided into training subset (75%) and validation (25%) subset for further statistical analysis. Splining interpolation was conducted via the Spatial Analyst tool to produce two output raster layers to our map; Splining (tension type) and Splining (regularized type). For both Splining types, the lead contaminant dataset was used, the output cell size was set to 15, the Weight Value was left at the default of 0.1, and the output layer was masked to the Meuse boundary. For Splining (Tension), higher weights can result in a coarser, less-smooth output.

Ordinary Kriging was conducted utilizing the Geostatistical Wizard pack in ArcGIS, producing two output raster layers to our map; Kriging (Gaussian) and Kriging (Spherical). Both displayed (see Results) Kriging runs had the following parameters: Ordinary Kriging – Prediction; Order of Trend Removal – 2nd; Output Cell Size - 15m; Polynomial order - 2nd order; Regression Type - Linear. Most of the parameter selections in ArcGIS Pro Geostatistical Wizard were left to default settings. For example, the default parameters for the Semivariogram/Covariance Modelling (calculating sill and lags), and Searching Neighbourhood Function (direction of influence, max/min number of neighbours) were left as is. Average Standard Error (ASE) was recorded in order to compare which of the models produced the better fit for our data as part of internal validation.

The validation data subset was exported to MS Excel and explored using regression analysis to obtain important values such as RSquared (goodness of fit) and calculating the root mean square (RMSE), a measure of the standard deviation of the residuals (observed vs predicted). The validation data was used to externally validate our models and inform our decisions.

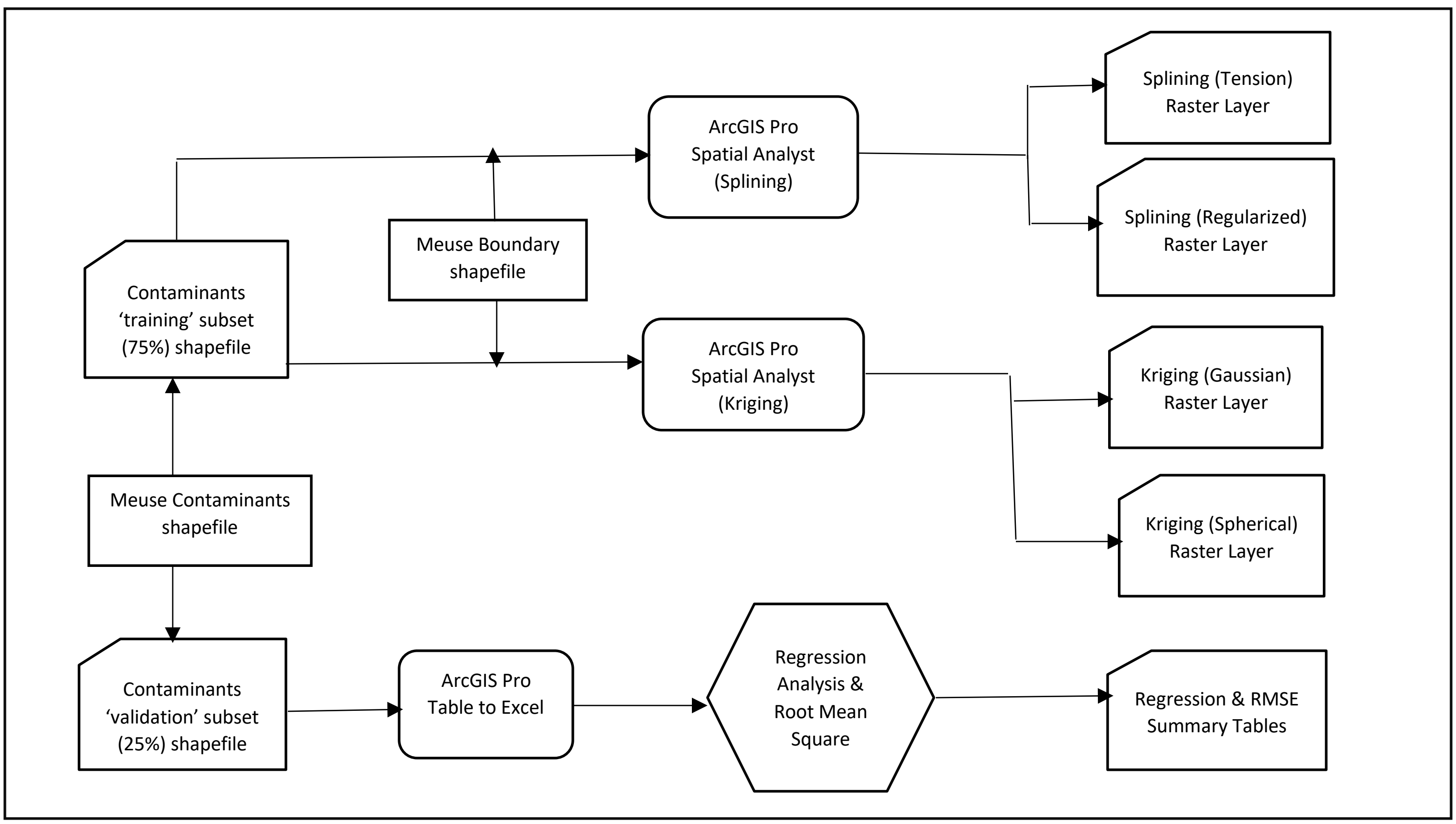

Figure 5: Workflow methodology used to examine soil lead concentrations in the Meuse dataset by Kriging & Splining interpolation methods.

Results:

Four output rasters were created by utilizing the geostatistical methods and training datasets outlined above in Figure 5. The two spline rasters (see Fig. 6 & 7) are symbolized by classified data in nine colour classes using geometric interval (defaults in ArcGIS Pro). The two Kriging rasters (see Fig. 7 & 8) are symbolized by classified data in ten colour classes using geometric interval (generally used the defaults in ArcGIS Pro).

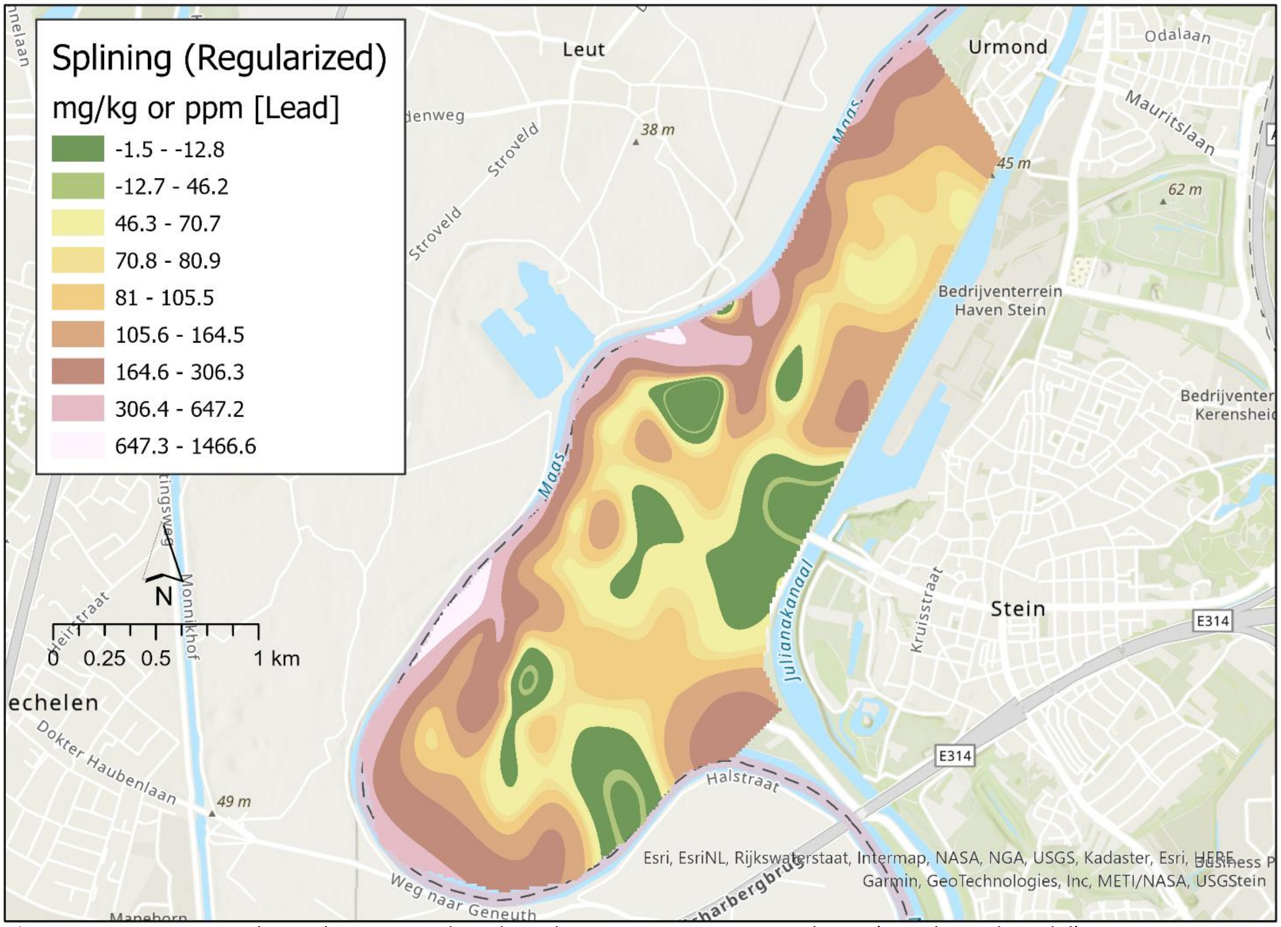

Figure 6: Raster output layer showing predicted Lead concentrations using Splining (Regularized model), Meuse R., Stein NL

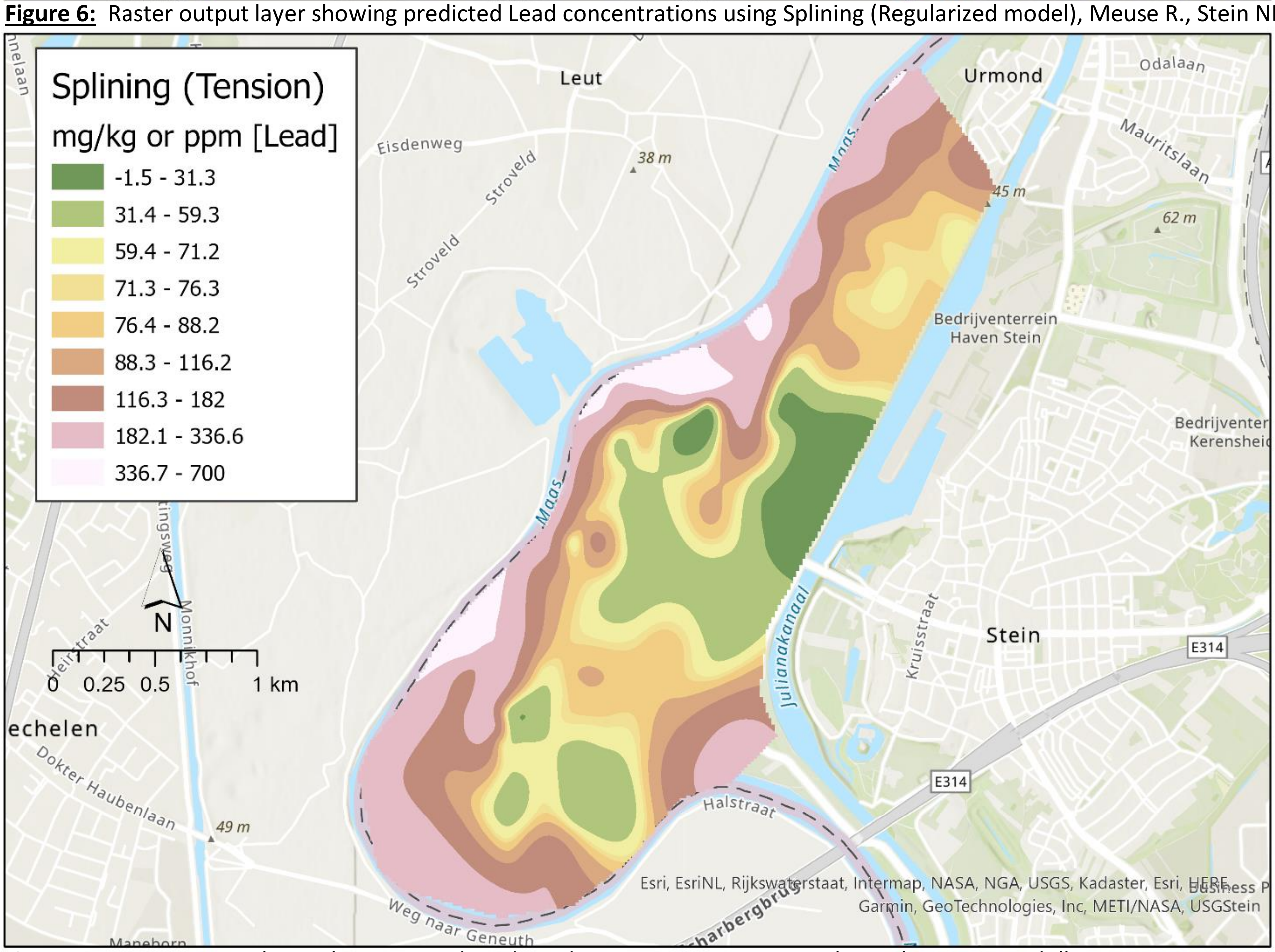

Figure 7: Raster output layer showing predicted Lead concentrations using Splining (Tension model), Meuse R., Stein NL

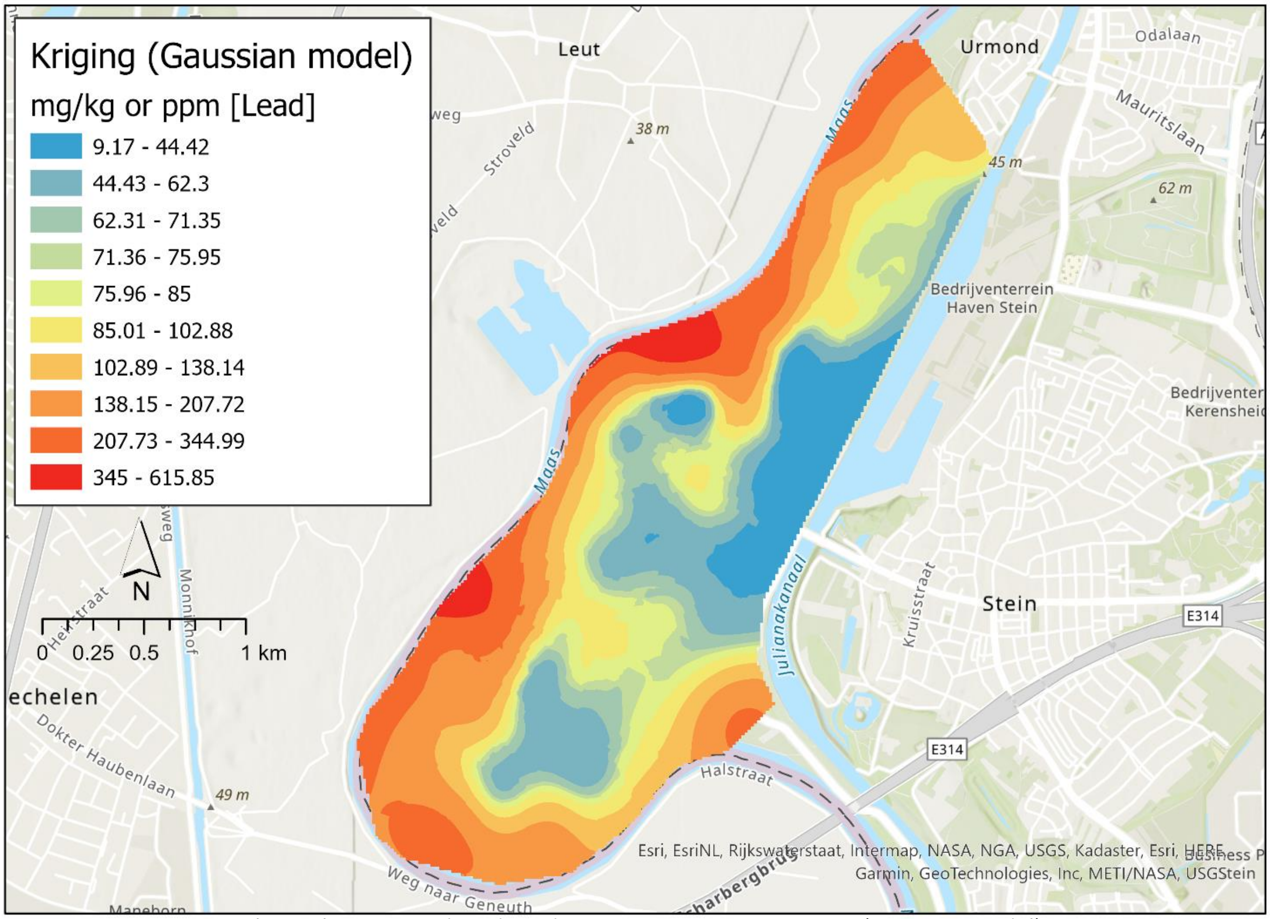

Figure 8: Raster output layer showing predicted Lead concentrations using Kriging (Gaussian model), Meuse R., Stein NL

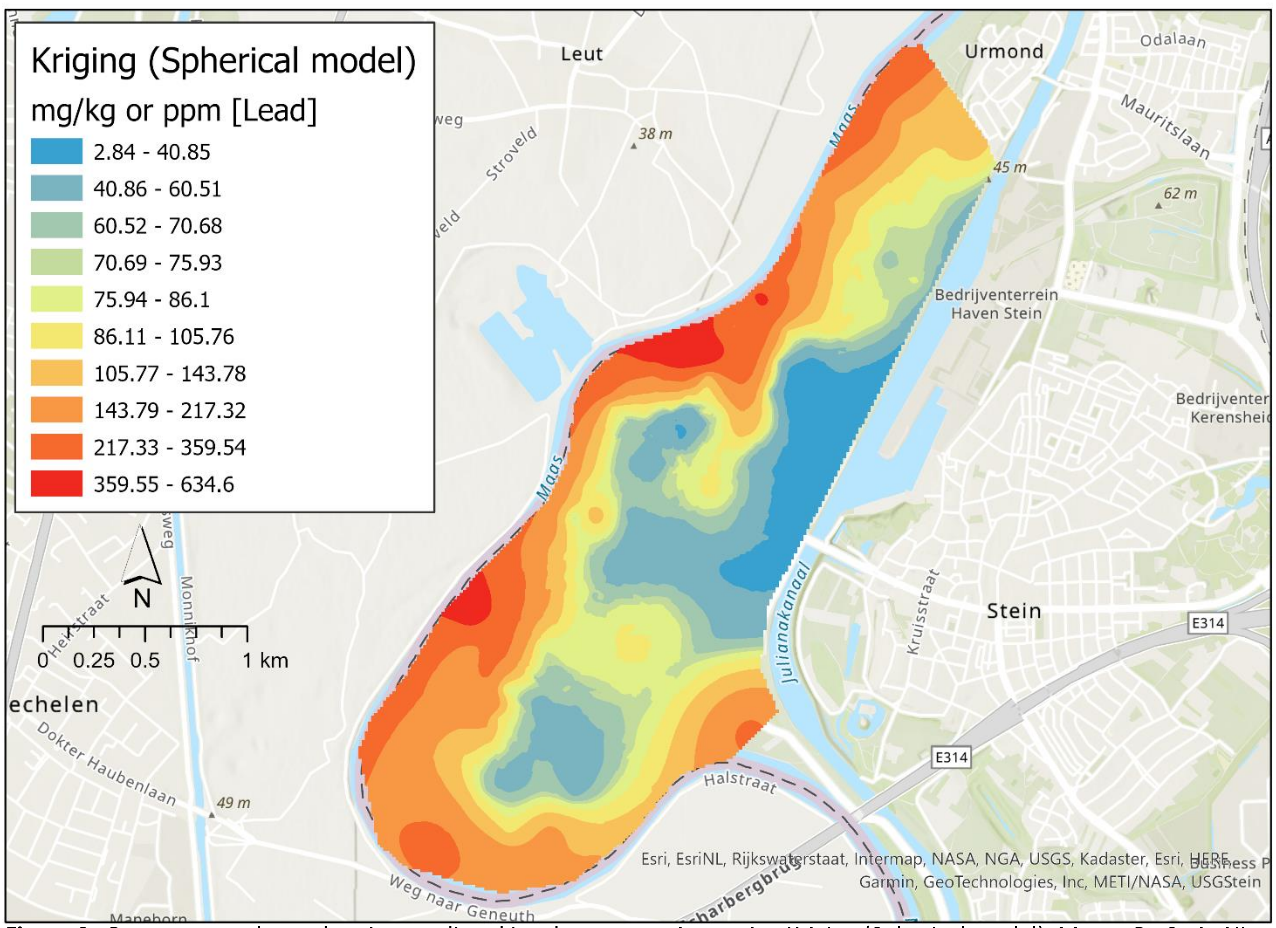

Figure 9: Raster output layer showing predicted Lead concentrations using Kriging (Spherical model), Meuse R., Stein NL

Comparison of regression statistics and calculated root mean square error (RMSE) values for each of the interpolation methods (see Table 1) using the validation dataset shows that Kriging (Gaussian model) has the highest values for RSquare and Multiple R at 0.3836 and 0.6194, respectively. The calculated RMSE was also lowest for the Kriging (Gaussian) method at 90.960, meaning there is less overall error associated with this method. A comparison between the four interpolation methods indicates that Kriging (Gaussian) is the better fit to our dataset. Between the two ordinary Kriging methods, the Gaussian model is significantly better than the Spherical model, with higher RSquare and Multiple R values, along with lower RMSE. Kriging (Gaussian) also had the lowest ASE value among the models indicating a comparatively better model.

The Splining (Tension) type is a better fit than the regularized type with higher RSquared and lower RMSE. All regression statistics for the validation dataset were statistically significant with p values significantly less than 0.05 (> $1x10^{-10}$). RSquared shows the 'goodness of fit' of how well the data fits the regression model. RMSE tells how concetrated the data is around the line of best fit, with lower numbers being more desirable. Both Kriging methods were run utilizing the 'removal of trend' function in ArcGIS Pro in order to remove the 2^{nd}-order trend (quadratic shape of lead distrubution in the study area) apparent in our dataset.

Table 1: Comparison (external validation) of regression statistics and calculated root mean square error (RMSE) values vs interpolation methods

Comparison of Regression Stats & RMSE vs Interpolation Method	*Interpolation Method*			
	Ordinary Kriging Exponential Kernel Spherical Model	Ordinary Kriging Exponential Kernel Gaussian Model	Splining Regularized	Splining Tension
Regression Statistics				
Multiple R (Pearson's r, correlation coefficient)	0.5857	**0.6194**	0.5992	0.6142
R Square (Coefficient of determination)	0.3431	**0.3836**	0.3590	0.3773
Root Mean Square Error (RMSE)	93.215	**90.960**	98.872	93.085

Discussion:

From our Results section above, the Kriging (Guassian) model is the interpolation method best suited to model our Meuse dataset experiment and predict unknown values in between the known data points. While Kriging can be used to create very accurate prediction surfaces, ordinary Kriging is its simplest form and subject to some assumptions that are not always met in applications relevant to environmental health and science, including pollution distributions as seen in some dataset. Ordinary Kriging assumes stationarity for a dataset, which requires isotropic distribution of data points in order to be a suitable method. Our dataset clearly exhibits a 2^{nd} order trend and this was removed in the Kriging process in ArcGIS Pro. A Kriging run without 'removal of trend' tool was also run to back-check our assumptions, showing a slightly lower RSquared and higher RMSE than our selected model indicating that our model parameterization choices were appropriate for this experiment.

In addition to the visual aspects of our expriment shown in the four raster outputs, the statistical analysis of the validation dataset clearly shows that Kriging (Gaussian) and Splining (Tension) are the two best fit models for our dataset, as inidcated by higher Multiple R and RSqaure values and lower RMSE values (see Table 1). Considering the distribution of our raw lead sample data and concentrations, it is conceivable to see why the two methods named above are the best fit of the four models for our data. Spline (Tension) method, for example, adjusts the best fit line to adhere tighter to actual data points (see Fig 2). As our data appears to trend towards higher concentrations of lead on the eastern bank of the Meuse, this explains why Spline (Tension) is the better deterministic method. It should be noted that a Kriging (Gaussian) model (our best fit model of the four tested) run *with* correction for anisotropy resulted in slightly higher RSqaured (0.387), Pearson's r (0.622), and slightly lower RMSE (90.76) values as compared to Kriging (Gaussian) alone. Since anisotropy is encountered when the correlation between two points depends on distance and orientation, correcting for this phenomenon improves the model accuracy marginally.

Since ordinary Kriging relies on stationarity in order to be accurate, there are other Kriging methods that can be used to account for complexity of environmental monitoring experiments. Universal Kriging, Co-Kriging, and block Kriging are more complex types of Kriging that allow for some variation in isotropy and stationarity across the spatial field. In our experiment, ordinary Kriging (Gaussian) was the best model for our dataset, however, it could be argued that the simpler Splining (Tenison) type method could be used with almost identical accuracy and output. Less complexity and cost can be a factor in many sectors, so it is not always better to chose more complex models that likely require more time and effort to run. If the pre-requisites (isotropy, stationarity) are not met for the more complex types of interpolation, there is often no benefit over running a more simplistic deterministic or other interpolation model.

ArcGIS Pro Geostatistical Wizard offers a large degree of parameterization for the more complex Kriging processes in order to account for real-world variability and complexity. The following parameters are modifiable in ArcGIS Pro and can be used to produce a better fit model if the modeller knows the characteristics of the data: trend removal order polynomial, transformation type, kernel function, number of lags, lag size, anisotropy correction, smoothing model type, searching neighbourhood (max and min neighbours, neighbour weightings). Kriging generally requires that the modeller make a deep investigation of the underlying data before choosing appropriate modifiable parameters and methods.

From our Meuse dataset for lead and other heavy metal contamination in Stein, NL, (visualized in Fig. 10), and the known serious environmental and health hazards caused by lead and heavy metals, society needs an accurate, realtively simple way to model pollution and other parameters in order to protect the environment and ecosystem health. In aquatic ecosystems, the sediments are the main sink and source of heavy metals, serving a significant role in the transportation and storage of noxious metals (Zhang, 2014). The Meuse River contamination brought down from the industrial areas of Belgium is further impacted by frequent flooding in the Meuse watershed. In 2021, the Meuse River flood was the highest in over 100 years (see maximum flood extent layer in Fig 10), directly impacting areas known to be highly contaminated with heavy metals. Zhang et al (2014) found that hydrological kinetic conditions, including physical disturbance (flooding) can influence the remobilization and resuspension of heavy metals. The 'flood extent' overlay below shows the need for effective modelling to identify hotspots along with other natural processes in order to help determine the best way to remediate pollutants, through excavation and treatment, in-situ remediation, or seal-in-place strategies.

Intermittent flooding along this contaminated river can deposit more contaminants deeper along the low-lying floodplains and surrounding areas (Tu, 2014). As heavy metals can be bound in the soil matrix, adding rushing water to the mix can mobilize the metal ions and potentially cause the metal ions to form bonds or species that are more bioavailable, i.e. more dangerous to life. However, sediments are the main source and sink of metal contaminants, and lead is generally found to be less mobile than some other metals because it is strongly retained by iron-oxides, organic matter and clay minerals (Zheng, 2017). Non-essential heavy metals, including lead, deserve speical attention since they do not have nay essential function in living organisms, are highly toxic at low exporsure levels, and are considered a main threat to all lfie forms (Henao, 2021).

Accurate geospatial prediction modelling is one process that can help identify, isolate, and help remediate contaminated areas. While Kriging (Gaussian) was best fit, our four models showed relatively similar visual raster outputs and external validation results.

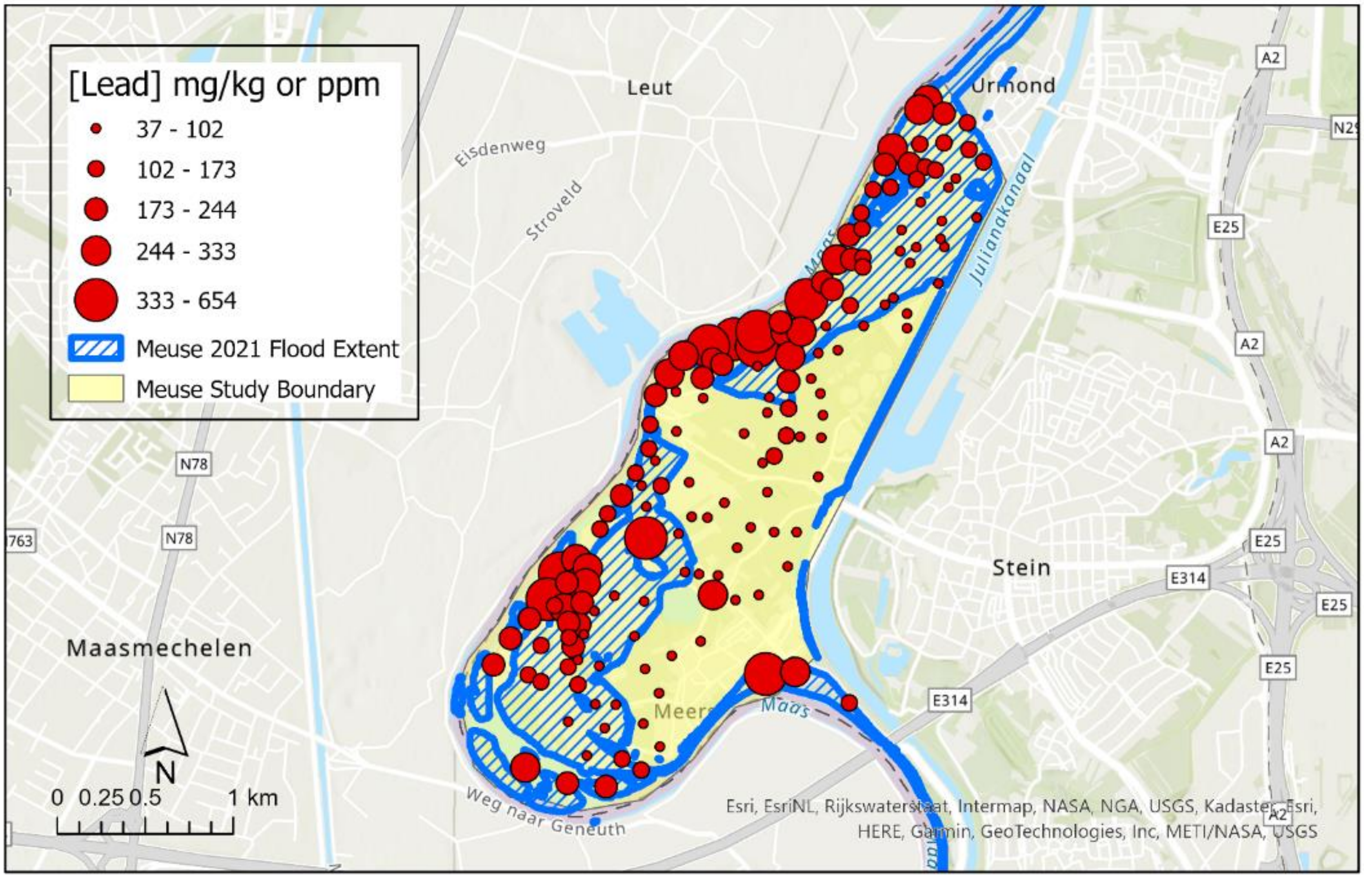

Figure 10: Map showing observed Lead concentrations along with 2021 Flood Extent boundary layer, Meuse R., Stein NL

Ideally, higher RSquared scores and lower RMSE than we achieved in our experiment would be preferred for deciding on a best model or whether a model is adequately predictive. A deep dive on the parameterization in the Kriging or Splining methods in ArcGIS Pro could create better fit models given enough time and resources. However, some fields of study such as environmental systems tend to have a greater amount of unexplainable variation where the RSquared value is bound to be lower, and the RMSE higher. Fortunately, a relatively low RSquared value combined with statistical significance (very low p value <0.05) can still provide important conclusions about relationships. To get a full picture, one must consider RSquared and RMSE values in combination with other statistics, residual plots, and in-depth knowledge of the subject area.

According to Occam's Razor, when one is challenged with competing explanations, and they are equal in all other respects, selecting the one with the fewest assumptions is likely the best approach (Trevors, 2008). Occam's Razor asserts that assumptions shall not be multiplied beyond need. This approach has been proven to be generally true in many discplines over time, unless more complexity has been shown to be required (Trevors, 2008).

With respect to our experiment, Kriging (Gaussian) was determined to be the best fit model, however when considering Occam's Razor, a strong argument can be made for choosing Splining (Tension) in certain situations. Due to the fact that Splining is less complex and has fewer assumptions than the Kriging model (at least in ArcGIS Pro), yet resulted in very similar validation scores, it may also be a prudent choice. Simplicity and model comprehensibility are also often identified as desired outcomes in their own right.

Conclusions:

Many geostatistical methods exist that help study and model myriad complex environmental and health issues. More and more complex models are being developed over time in the pursuit of accurate information related to extremely complex systems such as climate change and global weather patterns. Scienctists of all stripes are increasingly seeking to quanitfy, model and ultimately solve the totality of all anthropogenic impacts on the greater environment, including pollution, reasource consumption, and climate change (Trevors, 2008). Our experiment identified the most appropriate interpolation model between Kriging and Splining for the Meuse dataset, along with some caveats discussed above. Depending on the complexity of the system being studied, there is likely a suitable method to model desired parameters to such a degree that allows better planning, conservation, and infrastructure funding decisions. Soil and water contamination with industrial heavy metals is one piece of the giant puzzle that is human-induced damage to the enviroment. It is largely up to scientists to provide decision makers with possible solutions to extremely small-scale local issues as well as global-scale problems.

Relatively fast, efficient, accurate methods can be developed along with increased computing horsepower for studying issues such as heavy metal contamination of the Meuse Rivershed in Belgium and the Netherlands. To educate the general public and the politicians that pull the purse strings, it is likely a prudent step to employ Occam's Razor to all geostatistical methods and models. Problems with model complexity can result in both under-fitting, ie. removing too many important variables, and over-fitting, ie. ahering exactly to the training data. Under-fitting a model results in a model unable to properly determine correlation between variables. Over-fitting results in a model that is poorly able to generalize and predict new data thus defeating the purpose of having a model.

A complex model devloped with improper assumptions and variables is not a useful tool, in fact it may provide wrong answers to the right question, or right answers to the wrong question. In order to get maximum buy-in and public attention on a partuclar issue, we should ensure that models are as simple and as comprehensible *as they can be, by making them only as complex as they need to be.*

References:

Middlekoop, H. **(2000)**. Heavy-metal pollution of the river Rhine and Meuse floodplains in the Netherlands. *Netherlands Journal of Geosciences*, 79(4): 411-428. https://doi.org/10.1017/S0016774600021910

Zheng et al. **(2017)**. Toxicity threshold of lead (Pb) to nitrifying microorganisms in soils determined by substrate-induced nitrification assay and prediction model. *Journal of Integrative Agriculture 2017*, 16(8): 1832–1840. doi: 10.1016/S2095-3119(16)61586-1

Henao, S.G. **(2021)**. Heavy Metals in Soils and the Remediation Potential of Bacteria Associated with the Plant Microbiome. *Frontiers in Environmental Science*, 9:604216. doi: 10.3389/fenvs.2021.604216

Tu, Min et al. **(2005)**. Extreme floods in the Meuse River over the past century: aggravated by land-use changes? *Physics and Chemistry of the Earth*, 30(4-5): 267-276. https://doi.org/10.1016/j.pce.2004.10.00

Trevors, J.T. et al. **(2008)**. Occam's Razor and Human Society. *Water Air & Soil Pollution*, 205(Suppl 1): S69-70. DOI 10.1007/s11270-008-9736-6

Zhang, C et al. **(2014)**. Effects of sediment geochemical properties on heavy metal bioavailability. *Environment International,* 73: 270-281. https://doi.org/10.1016/j.envint.2014.08.010

ArcGIS and other resources consulted:

https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/comparing-interpolation-methods.htm

https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/how-Kriging-works.htm

https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/what-are-the-different-Kriging-models-.htm

https://pro.arcgis.com/en/pro-app/2.7/help/analysis/geostatistical-analyst/an-introduction-to-interpolation-methods.htm

https://pro.arcgis.com/en/pro-app/2.7/help/analysis/geostatistical-analyst/understanding-how-to-remove-trends-from-the-data.htm

https://www.ibm.com/cloud/learn/overfitting

https://www.publichealth.columbia.edu/research/population-health-methods/Kriging-interpolation

CODG 212 – Spatial Statistical Methods: Winter 2022 Course Notes

Meuse River 2021 Flood Extent data:
https://data.4tu.nl/articles/dataset/Maximum_flood_extents_Limburg_floods_July_2021/16817389

Leaflet API stand-alone webmap file attached.