



Comparison of interpolation methods in the production of rainfall-induced soil erosion maps in the urban area of Torres Novas, Portugal

Irene Vargas Pecegueiro ¹, Marco Painho ¹, and Ana Cristina Costa ¹

¹ NOVA Information Management School, NOVA University Lisbon, Campus de Campolide, 1070-312 Lisbon

Correspondence: Irene Vargas Pecegueiro (20240917@novaims.unl.pt)

Abstract. The evaluation of rainfall-induced soil erosion risk is fundamental for territorial planning and takes into account parameters such as rainfall erosivity, soil erodibility and the topographic factor. The Triangular Irregular Network (TIN) is the most frequently used interpolator in the production of digital elevation models (DEM) but is considered unsuitable by several authors for the calculation of soil erosion. Therefore, the DEM created for the city of Torres Novas, Portugal, using interpolation methods such as Inverse Distance Weighting, Ordinary Kriging, and Empirical Bayesian Kriging (EBK) were evaluated to determine which one was the most accurate. The best interpolator was EBK, from which a rainfall-induced soil erosion map was created. A map was also produced from TIN and both were compared with historical cartography. The EBK method was found to be the most effective interpolator for rainfall-induced soil erosion as well. Therefore, the authors recommend its use in future studies in the municipality of Torres Novas.

Submission Type. Analysis; Dataset.

BoK Concepts. [AM7] Spatial statistics; [AM8] Geostatistics.

Keywords. rainfall-induced soil erosion; topographic factor; triangulated irregular network; kriging.

1 Introduction

Among the geomorphological risks that affect a territory, the rainfall-induced soil erosion risk plays an important role in the evolution of the physical traits of the landscape, as it affects areas subject to excessive soil loss due to the

action of superficial drainage. This excessive loss brings negative consequences for maintaining the balance of morphogenetic and pedogenetic processes, soil quality and fertility, and the regulation of the hydrological cycle. (Cunha et al., 2021; Guduru & Jilo, 2023), for what its analysis is fundamental in territorial planning and management.

Most of the methodologies used to predict rainfall-induced soil erosion are based on empirical models, such as the Universal Soil Loss Equation, or USLE (Eq. 1), which is the most widely used in the world (Kinnell, 2010) because of its straightforward and relatively simple computational input needs compared to other models (Guduru & Jilo, 2023).

$$A = R \times K \times LS \times C \times P \quad (1)$$

A simplified version of the USLE (Eq. 2) is referenced in the National and Regional Strategic Guidelines provided in the Legal Regime of the Portuguese National Ecological Reserve (Decree order n.º 336/2019 of 2019-16-09, in its current wording) to define areas with high risk of rainfall-induced soil erosion, to which all municipalities in the country must adapt.

$$A = R \times K \times LS \quad (2)$$

In this simplified version of the Portuguese Decree, the average amount of soil loss (A) is obtained by multiplying the rainfall-runoff erosivity factor (R), the soil erodibility factor (K), and the topographic factor (LS), which depends on slope length and gradient. The values of the crop management factor (C) and the soil conservation practice factor (P) are considered constant and equal to 1.

To calculate the topographic factor, the first step is to create a digital elevation model (DEM). The Triangulated Irregular Network (TIN) is one of the most commonly

used interpolators for its production. Due to its simplicity, this interpolator is employed for a wide range of purposes, from volumetric calculations to urban area planning (Longley et al., 2015), and has become the preferred method used by the local authority of Torres Novas.

However, according to Bergonse & Reis (2015), TIN is unsuitable for the calculation of erosion as it disregards terrain curvature (strongly associated with erosive features) and has limitations in reproducing deep valleys, leading to underestimated values if the original surface is convex and overestimated values if it is concave.

Although there are no solid results in the literature regarding the performance of the various spatial interpolators available for DEM generation (Tan & Xu, 2014), some studies show that interpolators such as Inverse Distance Weighting (IDW), Spline, or Kriging provide good results depending on the purpose of the DEM and the context in which they are applied (Bergonse & Reis, 2015; Tan & Xu, 2014; Boumpoulis et al., 2023).

Considering the importance of accurate risk assessment, especially in urban areas, this study aims to evaluate the accuracy of TIN compared to other interpolators in the production of DEM for the city of Torres Novas.

The chosen interpolators were IDW, as it is also a simple and fast method to implement, Ordinary Kriging (OK) for its suitability in large areas with complex spatial variation and Empirical Bayesian Kriging (EBK), which provides more accurate error estimates than other kriging methods.

This study also seeks to assess whether the risk areas of rainfall-induced soil erosion resulting from the application of the DEM produced by TIN show significant differences when compared to those produced by the best interpolator identified.

2 Study Area

The county of Torres Novas covers an area of approximately 270 km², with a rainfall-induced soil erosion risk scenario ranging from moderate to high in more than 50% of its territory. These areas are primarily characterized by moderate to high slopes, combined with high erodibility and soils lacking vegetation (Cunha et al., 2021).

The city of Torres Novas was chosen as the study area (Fig. 1) due to its historical cartography, which includes the delineation of several risks, including erosion (Fig. 2), as well as up-to-date vector-based digital cartography at a scale of 1:2000, which allows a rigorous evaluation of the suitability of different interpolation methods.

Furthermore, the city exhibits an interesting topographical heterogeneity, ranging from higher elevations in the

western and northeastern regions, to lower floodplain areas in the east and southeast, with significant variations in slope between these zones.

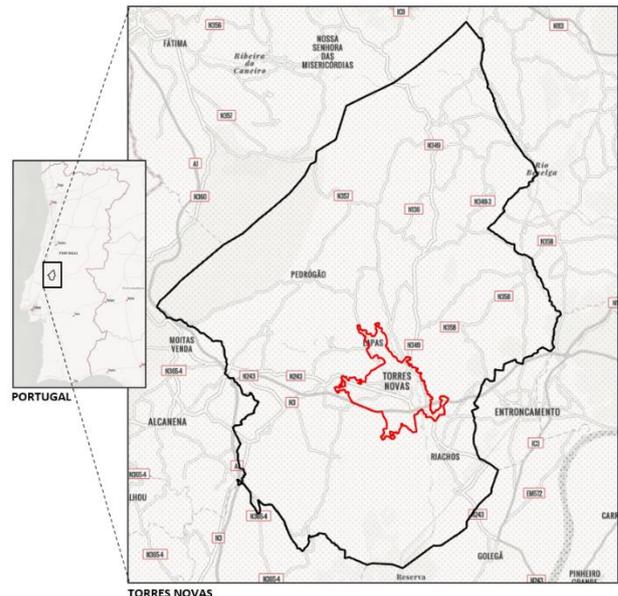


Figure 1. Study Area Location: from left to right, the municipality of Torres Novas in Portugal, followed by the study area (in red) within the municipality..



Figure 2. Urbanization Plan of the city of Torres Novas, dated March 1992 (reproduction authorized by Câmara Municipal de Torres Novas).

3 Methods and data

3.1 Analysis

The diagram in Fig. 3 illustrates the main steps of the analysis.

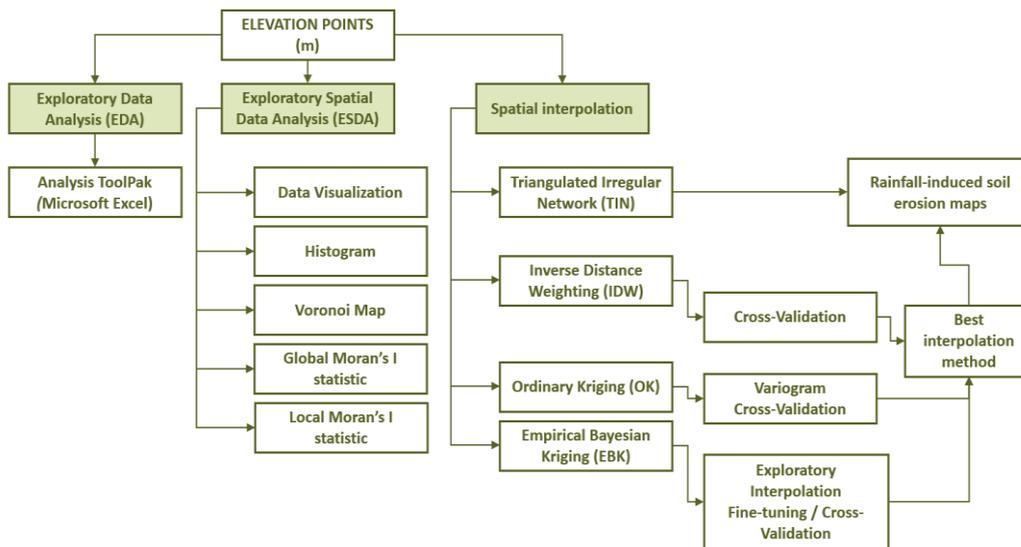


Figure 3. Diagram with the main steps of the analysis.

3.2 Data and Software Availability

The elevation points used in this study to produce the DEMs for the different interpolators were obtained from the Vector Digital Cartography (NdD1) at a 1:2000 scale for the city of Torres Novas, certified in 2024 and provided by Câmara Municipal de Torres Novas (<https://cm-torresnovas.pt/>) through a formal request form. This data cannot be redistributed due to licensing restrictions.

The Rainfall Erosivity (Panagos et al., 2015) and Soil Erodibility (Panagos et al., 2014) datasets for Europe are available for download from the European Soil Data Centre (ESDAC) of the Joint Research Centre (JRC) through a request form at <https://esdac.jrc.ec.europa.eu/> and cannot be redistributed due to licensing restrictions.

The analysis was conducted using ArcGIS Pro software, version 3.2., developed and commercialized by ESRI (<https://www.esri.com/>).

4 Results and Discussion

4.1 Exploratory spatial data analysis

The exploratory data analysis of the elevation points (Tab. 1) reveals that the city's average elevation is 63.62 meters with a standard deviation of 25.51, indicating significant variation of elevation values in relation to the average. Elevations are below 60 meters in 50% of the points, and the distribution is positively skewed (asymmetry of 0.53), suggesting a tendency towards higher elevations.

The elevation range in the city spans from 22.48 to 127.2 meters (Fig. 4), with the lowest elevations primarily

located in the east and southeast, corresponding to a floodplain along the main river of the municipality, while the highest values are found in the west and northeast.

Table 1. Descriptive statistics for elevation (m).

| | | | |
|----------|--------|----------|--------|
| Min | 22.48 | Range | 104.72 |
| Max | 127.20 | IQR | 34.15 |
| Mean | 63.62 | Q1 | 43.85 |
| Std Dev | 25.51 | Q3 | 78.00 |
| Median | 60.00 | CV | 0.40 |
| Count | 2679 | Skewness | 0.53 |
| Outliers | 0 | Kurtosis | 2.44 |

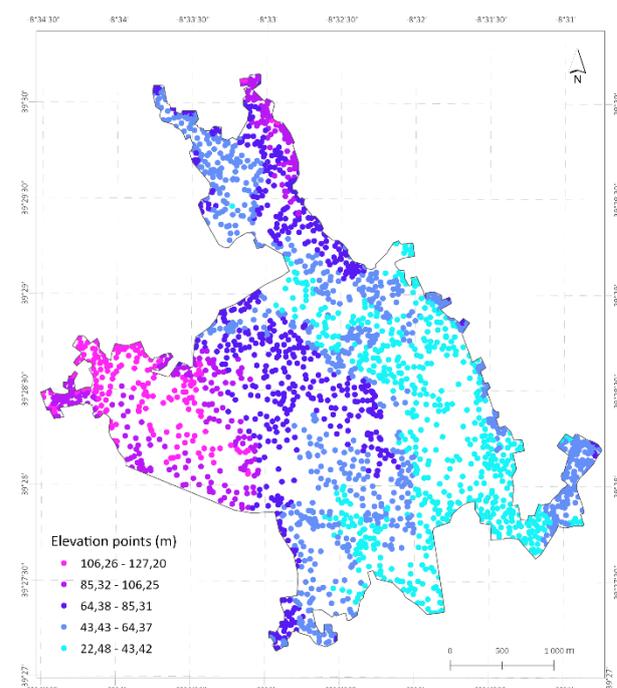


Figure 4. Elevation data posting using graduated colors.

The regional histogram (Fig. 5) also confirms the slightly positive asymmetry of the data and, by selecting different areas of the map and histogram, a trend of data aggregation can be observed.

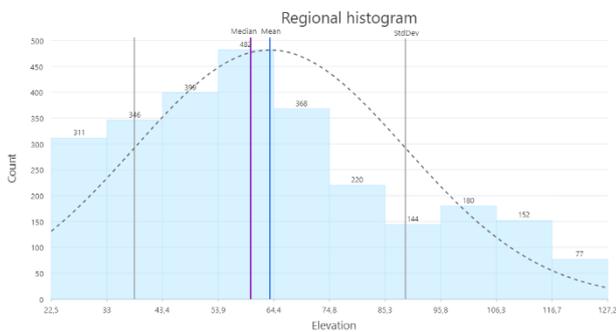


Figure 5. Regional Histogram.

The Voronoi map (Fig. 6) reveals an isotropic pattern for the overall study area.

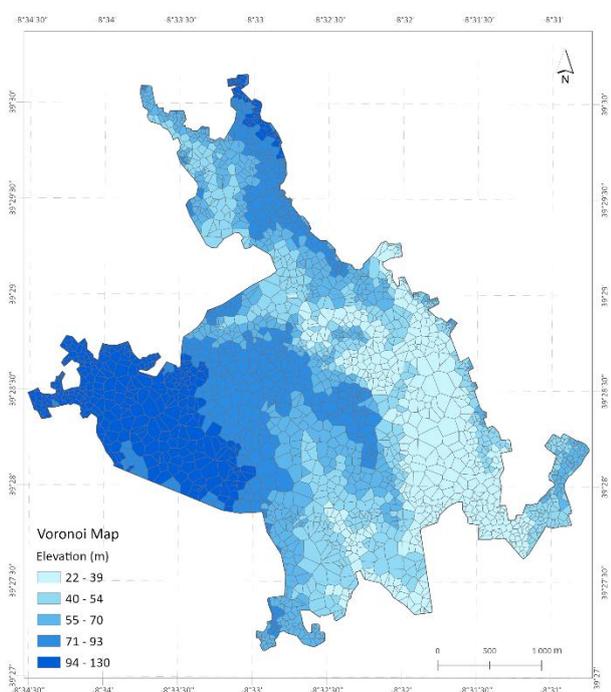


Figure 6. Simple Voronoi map of elevation.

To further analyze spatial autocorrelation, the Global Moran's I and Local Moran's I statistics were applied.

The Global Moran's I revealed a statistically significant spatial autocorrelation, with an index of 0.9645 (p -value < 0.0000). Given the z -score of 98.53, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

The Local Moran's I statistic (Fig. 7) identified 1950 local clusters (764 high-high and 1186 low-low), reflecting a positive spatial autocorrelation, and 28 high-low spatial outliers, indicating a negative spatial autocorrelation, located in areas with steep slopes. No low-high outliers

were found. The results also show 701 points, located in the central area and southwest, without significant spatial autocorrelation (p -value > 0.05).

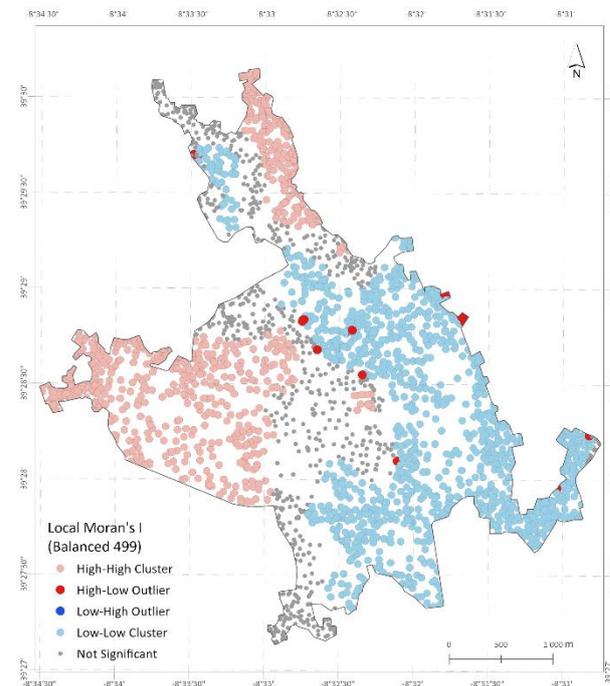


Figure 7. Local Moran's I map.

The significant spatial autocorrelation of the data and the global isotropic pattern suggest that kriging methods are likely to be more effective in producing the DEM than deterministic ones (Bergonse & Reis, 2015).

4.2 Triangulated Irregular Network (TIN)

The TIN predicted surface (Fig. 8) was created using the 'Create TIN' tool, available with the 3D Analyst extension of ArcGIS Pro, and then exported to a raster, with a cell size of 2 meters. This cell size was chosen for all the interpolators to provide more detail, as the base and historical cartography are at a 1:2000 scale.

The prediction errors associated with the TIN were obtained using the 'Extract Values to Points' tool, available with the Spatial Analyst extension of ArcGIS Pro, and simple map algebra. The resulting Mean Error and Root Mean Square Error (RMSE) values were 0.001 and 0.11, which was expected as this is an exact interpolator.

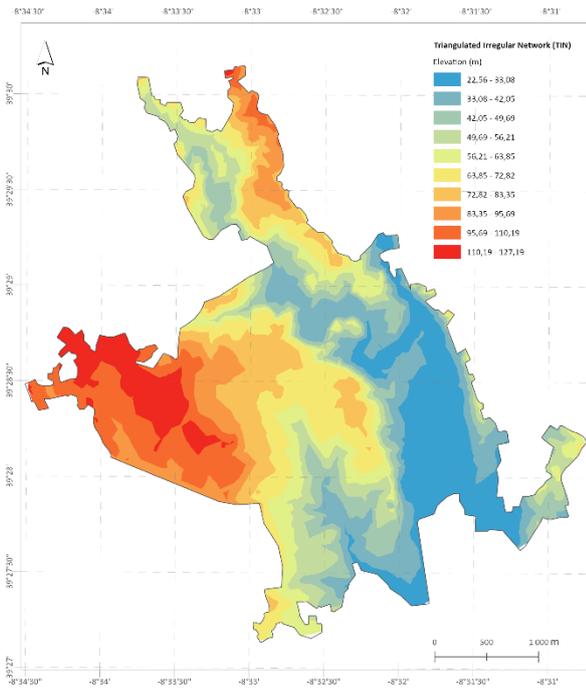


Figure 8. TIN predicted surface.

4.3 Inverse Distance Weighting (IDW)

To obtain the most accurate surface from the IDW interpolator, the ‘Geostatistical Wizard’ tool, available with the Geostatistical Analyst extension of ArcGIS Pro, was used. Cross-validation was performed by adjusting different settings for the neighborhood parameter (Tab. 2) and the configuration with the lowest RMSE was chosen, as it provided the most accurate model.

Table 2. Local neighborhood definitions and prediction errors statistics for IDW.

| Sector Type | Max. neighbors | Min. neighbors | Mean Error | RMSE |
|-------------|----------------|----------------|--------------|--------------|
| 1 sector | 15 | 10 | 0.293 | 3.246 |
| | 8 | 5 | 0.249 | 3.004 |
| 4 sectors | 15 | 10 | 0.295 | 3.726 |
| | 8 | 5 | 0.255 | 3.381 |
| 4 sec. 45° | 15 | 10 | 0.292 | 3.723 |
| 8 sectors | 15 | 10 | 0.309 | 4.073 |
| | 8 | 5 | 0.264 | 3.688 |

The resulting IDW predicted surface (Fig. 9) was then exported to a raster with a cell size of 2 meters.

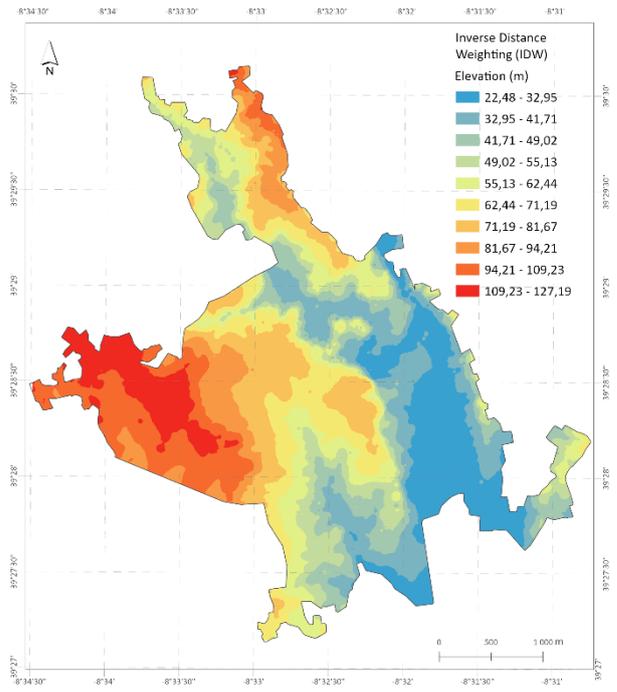


Figure 9. IDW predicted surface.

A contour line map was also produced (Fig. 10), confirming that there is no global anisotropic pattern in the data.

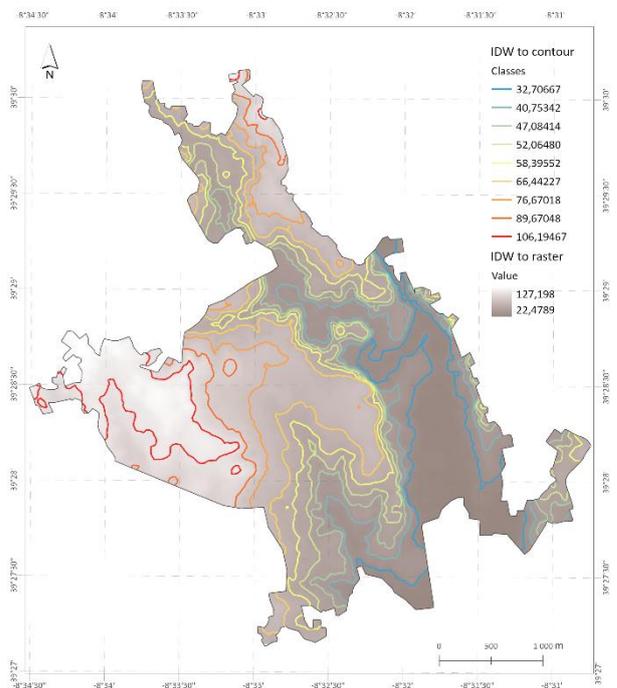


Figure 10. Contour line map for IDW.

4.4 Ordinary Kriging (OK)

Using the ‘Geostatistical Wizard’ tool again, this time for the OK, the semivariogram model that best fit the data was first determined. Of the two models tested (Tab. 3), the Sill, representing the value where the variogram

stabilizes, is higher in the Spherical model, while the range, representing the distance beyond which spatial autocorrelation no longer exists, is higher in the Exponential model. As for the nugget effect, related to variability at very short distances, the value in the Spherical model is low, suggesting reduced measurement errors, whereas in the Exponential model, it is zero.

Table 3. Experimental variogram parameters.

| Model | Spherical | Exponential |
|--------------|-----------|-------------|
| Nugget | 5.525 | 0 |
| Range | 5258.949 | 6071.682 |
| Anisotropy | False | False |
| Partial Sill | 1130.730 | 1009.941 |
| Model | Spherical | Exponential |

As both methods yielded similar results, the semivariogram model chosen was the Spherical, as it followed the averaged values more closely, particularly the first ones (Fig. 11).

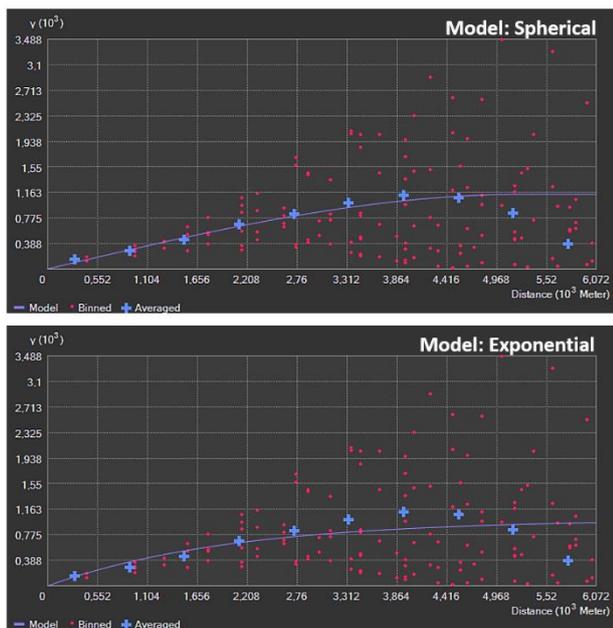


Figure 11. Semivariogram models for OK.

Cross-validation was then performed, but the configurations with different neighborhood values did not result in changes to the outcomes, so only the configurations for different sectors are presented (Tab. 4). Based on the resulting prediction errors, the configuration with the lowest RMSE was chosen, as it provided the most accurate model.

Table 4. Prediction errors statistics for OK, using 5 maximum neighbors and 2 minimum neighbors for all sectors

| Sector Type | Mean Error | RMSE | Root-Mean-Square Standardized error (RMSSE) |
|--------------------|--------------|--------------|---|
| 1 sector | 0.117 | 2.558 | 0.561 |
| 4 sectors | 0.047 | 2.438 | 0.552 |
| 4 sect. 45° | 0.049 | 2.438 | 0.552 |
| 8 sectors | 0.055 | 2.431 | 0.552 |

The resulting OK predicted surface (Fig. 12) was then exported to a raster with a cell size of 2 meters.

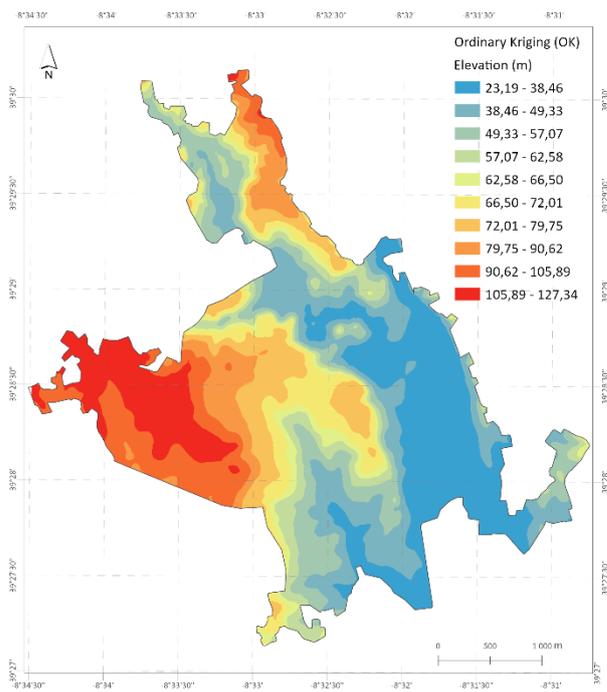


Figure 12. OK predicted surface.

4.5 Empirical Bayesian Kriging (EBK)

The best-fitting model for the EBK was found using the 'Exploratory Interpolation' tool, available with the Geostatistical Analyst extension of ArcGIS Pro. The tool identified 'EBK Advanced' as the best result (Tab. 5).

Table 5. Interpolation models ranks and cross-validation statistics

| Model | Rank | RMSE | Mean Error |
|-----------------|------|-------|------------|
| EBK - Advanced | 1 | 2.061 | 0.081 |
| OK - Optimized | 2 | 2.178 | 0.101 |
| EBK - Default | 2 | 2.191 | 0.071 |
| IDW - Optimized | 2 | 2.864 | 0.239 |
| IDW - Default | 2 | 3.246 | 0.293 |
| OK - Default | 2 | 4.314 | -0.054 |

Based on the model produced by the Exploratory Interpolation tool, the EBK was fine-tuned using the values shown in Tab. 6.

Table 6. Parameters for EBK advanced

| Parameters | |
|-----------------------|--------------------|
| Subset size | 200 |
| Overlap factor | 2 |
| Number of simulations | 300 |
| Transformation | Empirical |
| Semivariogram Type | K-Bessel Detrended |
| Radius | 75.159 |

Through the analysis of specific locations on the map (Fig. 13), it was found that the semivariograms in the western area are more accurate, whereas in the eastern area, with fewer and more homogeneous data, they are less precise. However, the standard deviation values in several locations are lower than the average standard error of EBK Advanced, therefore, the model is well-adjusted.

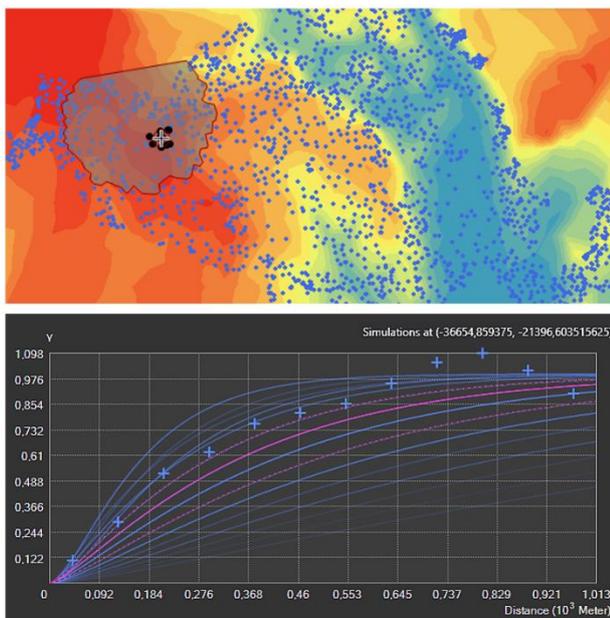


Figure 13. Example of semivariogram analysis.

Subsequently, different cross-validation configurations were tested (Tab. 7), and the option of 4 sectors with a neighborhood between 8 and 5 was chosen to generate the DEM from the EBK, as it yielded the lowest estimation errors.

The resulting EBK predicted surface (Fig. 14) was then exported to a raster with a cell size of 2 meters.

Table 7. Local neighborhood definitions and prediction errors statistics for EBK.

| Sector Type | Max. neigh. | Min. neigh. | Mean Error | RMSE | RMSSE |
|--------------|-------------|-------------|------------|-------|-------|
| 1 sector | 15 | 10 | 0.081 | 2.061 | 0.890 |
| 4 sectors | 15 | 10 | 0.064 | 2.050 | 0.881 |
| 4 sector 45° | 8 | 5 | 0.055 | 2.039 | 0.884 |
| 8 sectors | 15 | 10 | 0.060 | 2.051 | 0.877 |
| 8 sectors | 8 | 8 | 0.062 | 2.053 | 0.870 |

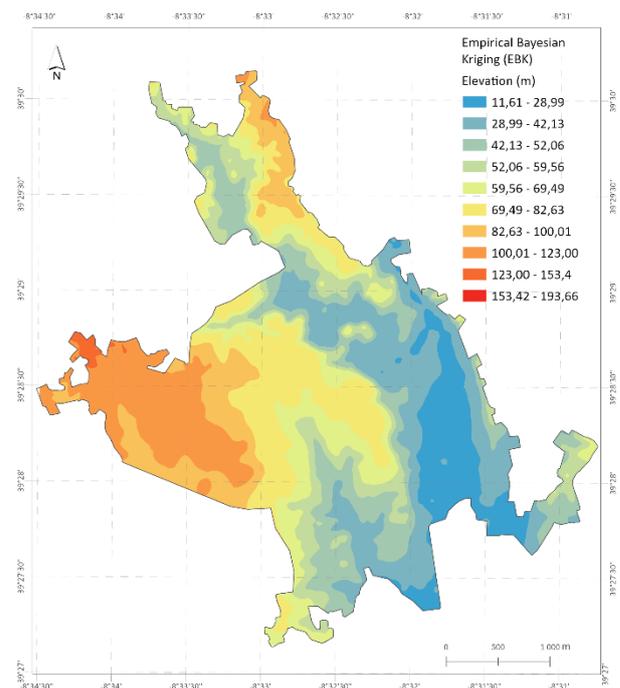


Figure 14. EBK predicted surface.

4.6 Rainfall-induced soil erosion maps

The comparison of prediction error statistics for the different interpolators (Tab. 8) shows that the interpolator that presented the lowest RMSE was the EBK. However, EBK required the most processing time, taking about 5 hours to produce a 2m cell size raster, while the others took less than 2 minutes.

Table 8. Prediction errors statistics comparison.

| Method | Mean Error | RMSE | Min. estimated elevation | Max. estimated elevation |
|--------|------------|-------|--------------------------|--------------------------|
| IDW | 0.249 | 3.004 | 22.48 | 127.19 |
| OK | 0.049 | 2.438 | 23.19 | 127.34 |
| EBK | 0.055 | 2.039 | 11.61 | 193.66 |

From the DEM produced by the TIN and the EBK, and the precipitation erosivity and soil erodibility data available in the JRC/ESDAC, two rainfall-induced soil erosion maps were generated (Fig. 15 and 16) according to the simplified version of the USLE (Eq. 2).

Comparing the resulting maps with the historical cartography of the city (Fig. 17), it is clear that the EBK produces much more accurate rainfall-induced soil erosion maps, modelling areas that are more aligned with the city's reality.

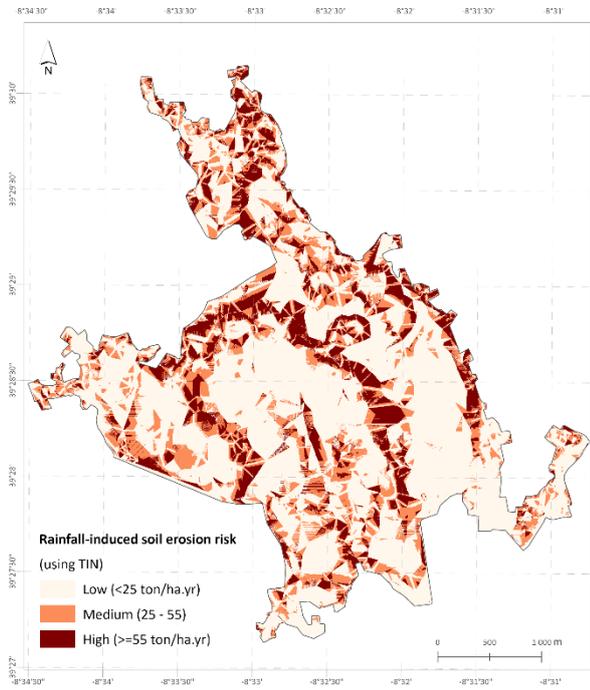


Figure 16. Rainfall-induced soil erosion using the TIN predicted surface.

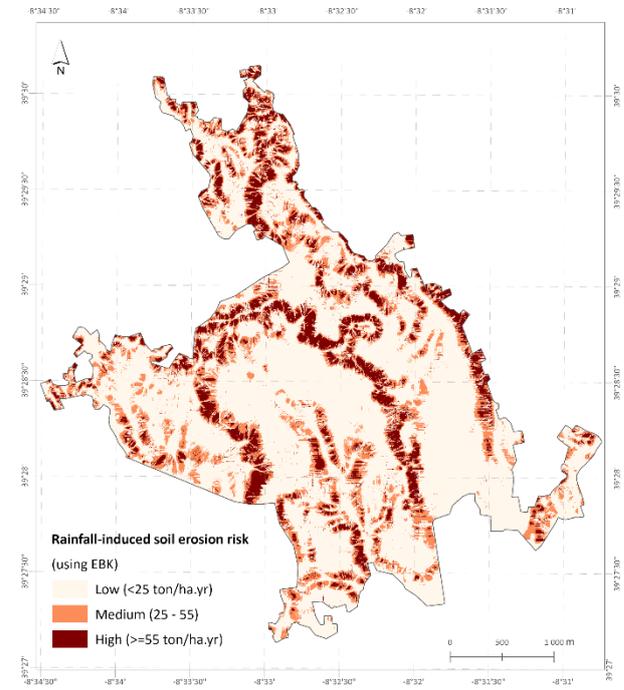


Figure 17. Rainfall-induced soil erosion using the EBK predicted surface.

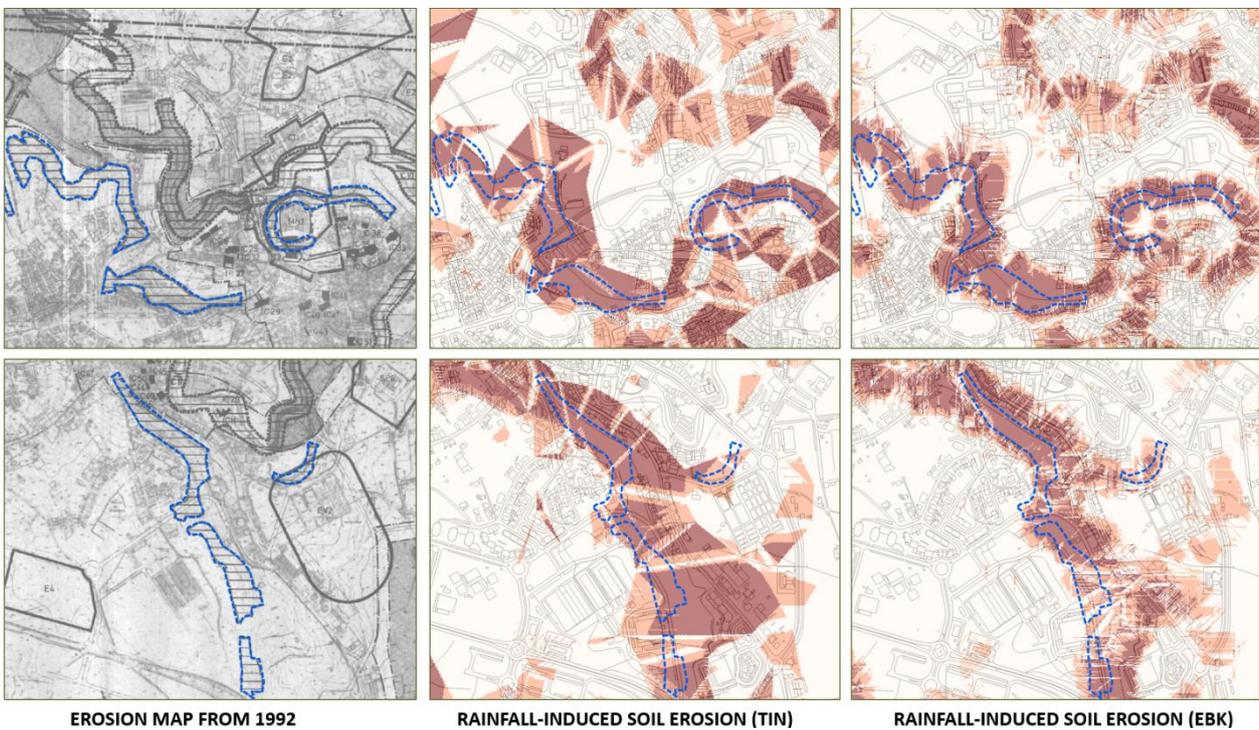


Figure 15. Comparison of the erosion areas defined in the Urbanization Plan of the city of Torres Novas (from 1992) and the rainfall-induced soil erosion maps produced using the TIN and EBK predicted surfaces

The improved accuracy of the EBK can be attributed to its ability to better handle terrain heterogeneity, such as those observed in the study area, provided the model is properly parameterized. In contrast, the TIN is more sensitive to abrupt slope variations, which may affect its performance in areas with significant topographical differences.

5 Conclusions

This essay aimed to evaluate the accuracy of TIN in comparison to other interpolation methods in the production of rainfall-induced soil erosion maps for the city of Torres Novas.

The elevation points used to generate the DEMs were obtained from the Vector Digital Cartography (NdD1) at a 1:2000 scale for the city of Torres Novas, certified in 2024. As elevation is a continuous variable, strong spatial correlation and spatial dependence were expected in the data, particularly in the western part of the city, with higher elevations, and in the eastern part, with lower altitudes corresponding to a floodplain along the municipality's main river.

DEMs using four different interpolation methods were produced: TIN, commonly used in this type of studies for its simplicity and rapidness, IDW and two Kriging methods – Ordinary Kriging and EBK. The interpolated surfaces that were produced do not differ significantly from each other in a global form, corresponding well to the city reality. This likeness among the interpolators is probably due to the fairly large number of points that produce a tendency to equity. Excluding TIN, EBK was the interpolator that presented fewest estimation errors.

From the comparison between the rainfall-induced soil erosion maps produced by TIN and EBK and the existing historical cartography for the city, EBK stands out as the best interpolator, producing more accurate and detailed areas. Although TIN creates a reliable map, it resulted in some overestimation of values in the areas with high erosion risk, which did not occur with EBK.

Producing rainfall-induced soil erosion maps that are closer to reality allows for better decision-making in the scope of territorial planning and management. Therefore, it is recommended the use the EBK interpolator in future studies of rainfall-induced soil erosion in the municipality of Torres Novas. However, it is important to consider that there are some limitations regarding the processing time of this interpolator, which is longer, and the need for proper model tuning, as it is sensitive to incorrect specification of the semivariogram model.

Declaration of Generative AI in writing

The authors declare that they have not used Generative AI tools in the preparation of this manuscript, nor for generating scientific content, research data, or substantive conclusions. All intellectual and creative work, including the analysis and interpretation of data, is original and has been conducted by the authors without AI assistance.

Acknowledgements

The authors acknowledge the support by national funds through FCT (Fundação para a Ciência e a Tecnologia), under the project - UIDB/04152/2020 (DOI: 10.54499/UIDB/04152/2020) - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS).

Author contributions. **Irene Vargas Pecegueiro:** Conceptualization, Methodology, Validation, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Visualization. **Marco Painho:** Writing - Review & Editing, Supervision. **Ana Cristina Costa:** Writing - Review & Editing, Supervision.

Competing interests. The authors declare that they have no conflict of interest.

References

- Bergonse, R., & Reis, E. (2015). Reconstructing pre-erosion topography using spatial interpolation techniques: A validation-based approach. *Journal of Geographical Sciences*, 25(2), 196–210. <https://doi.org/10.1007/s11442-015-1162-2>
- Boumpoulis, V., Michalopoulou, M., & Depountis, N. (2023). Comparison between different spatial interpolation methods for the development of sediment distribution maps in coastal areas. *Earth Science Informatics*, 16(3), 2069–2087. <https://doi.org/10.1007/s12145-023-01017-4>
- Cunha, L., Leal, C., Santos, N., Tavares, A., Santos, P., Ligeiro, A., & Pecegueiro, I. (2021). CARTAS DE RISCOS DO CONCELHO DE TORRES NOVAS. <https://www.researchgate.net/publication/360167847>
- Guduru, J. U., & Jilo, N. B. (2023). Assessment of rainfall-induced soil erosion rate and severity analysis for prioritization of conservation measures using RUSLE and Multi-Criteria Evaluations Technique at Gidabo watershed, Rift Valley Basin, Ethiopia. *Ecology & Hydrobiology*, 23(1), 30–47. <https://doi.org/10.1016/j.ecohyd.2022.09.002>

- Kinnell, P. I. A. (2010). Event soil loss, runoff and the Universal Soil Loss Equation family of models: A review. *Journal of Hydrology*, 385(1–4), 384–397. <https://doi.org/10.1016/j.jhydrol.2010.01.024>
- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic information science and systems* (Fourth edition). Wiley.
- Panagos, P., Ballabio, C., Borrelli, P., Meusburger, K., Klik, A., Rousseva, S., Tadić, M. P., Michaelides, S., Hrabalíková, M., Olsen, P., Aalto, J., Lakatos, M., Rymaszewicz, A., Dumitrescu, A., Beguería, S., & Alewell, C. (2015). Rainfall erosivity in Europe. *Science of The Total Environment*, 511, 801–814. <https://doi.org/10.1016/j.scitotenv.2015.01.008>
- Panagos, P., Meusburger, K., Ballabio, C., Borrelli, P., & Alewell, C. (2014). Soil erodibility in Europe: A high-resolution dataset based on LUCAS. *Science of The Total Environment*, 479–480, 189–200. <https://doi.org/10.1016/j.scitotenv.2014.02.010>
- Tan, Q., & Xu, X. (2014). Comparative Analysis of Spatial Interpolation Methods: An Experimental Study. *Sensors and Transducers*, 165, 155–163. <https://www.researchgate.net/publication/279902313>