International Journal of Applied Earth Observation and Geoinformation 139 (2025) 104472

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Scientists yet to consider spatial correlation in assessing uncertainty of spatial averages and totals

Alexandre M.J.-C. Wadoux [a,*], Gerard B.M. Heuvelink [b,c]

[a] LISAH, Univ Montpellier, AgroParisTech, INRAE, IRD, L'Institut Agro, Montpellier, France
[b] Soil Geography and Landscape group, Wageningen University & Research, The Netherlands
[c] ISRIC – World Soil Information, Wageningen, The Netherlands

ARTICLE INFO

Keywords:
Spatial aggregation
Change of support
MRV
Mapping
Machine learning
Geostatistics

ABSTRACT

High-resolution maps of climate and ecosystem variables are essential for supporting terrestrial carbon stocks and fluxes estimation, climate change mitigation, and ecosystem degradation assessment. These maps are usually created using remotely sensed data obtained from various types of imagery and sensors. The remote sensing data typically serve as covariates to deliver spatially explicit information using machine learning algorithms. Often the uncertainty associated with the maps is also quantified, for instance by prediction error variance maps or by maps of the lower and upper limits of a prediction interval. In addition, these products are often aggregated to regional, national, or global scales relevant to climate policy, natural resource inventory, and measurement, reporting, and verification (MRV) frameworks. Quantifying uncertainty in aggregated products is crucial as it is necessary to assess their value and evaluate whether changes and trends in aggregated estimates are statistically significant. However, we argue that such uncertainty is frequently inaccurately assessed due to the neglect of spatial correlation in map errors. This critical methodological issue has been overlooked in most large-scale mapping studies.

The Intergovernmental Panel on Climate Change (IPCC, IPCC, 2022), the United Nations Sustainable Development Goals (The United Nations, 2023), and the Paris Agreement (UNFCCC, 2015) rely on spatial information to support policy and decision-making processes to address climate change, biodiversity loss, and land degradation. This has led to the production of maps showing the spatial distribution of natural resources and climate-related variables using measurements combined with satellite imagery and modelling techniques (Phillips et al., 2019; Harris et al., 2021). Many of these maps are made at a fine spatial resolution and hence need to be spatially aggregated to infer the state of the environment for regions, countries, continents, or the entire globe. For instance, the United Nations Framework Convention on Climate Change (UNFCCC) (UNFCCC, 2023) provides detailed data on greenhouse gas emissions and removals, which are aggregated from fine-scale measurements and maps to provide a national overview. Another example is soil carbon accounting, where MRV platforms use digital soil mapping (Smith et al., 2020) to map carbon stocks and where these maps must be aggregated from the field to the project level for carbon crediting.

Estimates of spatial averages and totals are easily obtained by averaging or summing all point values of a map within the area of interest. In the scientific literature, we found many high-impact publications on natural resources and climate-related global variables that reported the uncertainty of spatial averages or totals in the same way (e.g. Nahlik and Fennessy, 2016; Xiao et al., 2019; Harris et al., 2021); that is by averaging or adding standard deviations or variances. But this is wrong and typically leads to a gross overestimation of the uncertainty.

The uncertainty of a spatial average is much smaller than the average of the spatial uncertainty because errors partially cancel out when averaging. We illustrate this with a simple example. If we roll a die and hide the outcome, then you will be quite uncertain about the outcome. It could be any number between 1 and 6 and there is only a 33% chance that it is between 3 and 4. Now, let us roll 100 dice and average their outcomes. The result still is a value between 1 and 6, but are you equally uncertain? No, you are not because you know that the average of the 100 dice will be quite close to 3.5. Your uncertainty has much decreased because high and low roll outcomes average out. The chance that the average is between 3 and 4 has dramatically increased to 99.7%. The more dice we average the less uncertain we are. With one million rolls there is a 99% probability that the average is between 3.495 and 3.505. The bottomline is that averaging reduces uncertainty, which should not come as a surprise because it is conveyed in every statistical textbook.

* Correspondence to: Laboratory of soil-agrosystem-hydrosystem interaction (LISAH), 2 place Pierre Viala, 34090 Montpellier, France.
E-mail address: alexandre.wadoux@inrae.fr (A.M.J.-C. Wadoux).

https://doi.org/10.1016/j.jag.2025.104472
Received 27 August 2024; Received in revised form 28 February 2025; Accepted 7 March 2025
Available online 23 March 2025

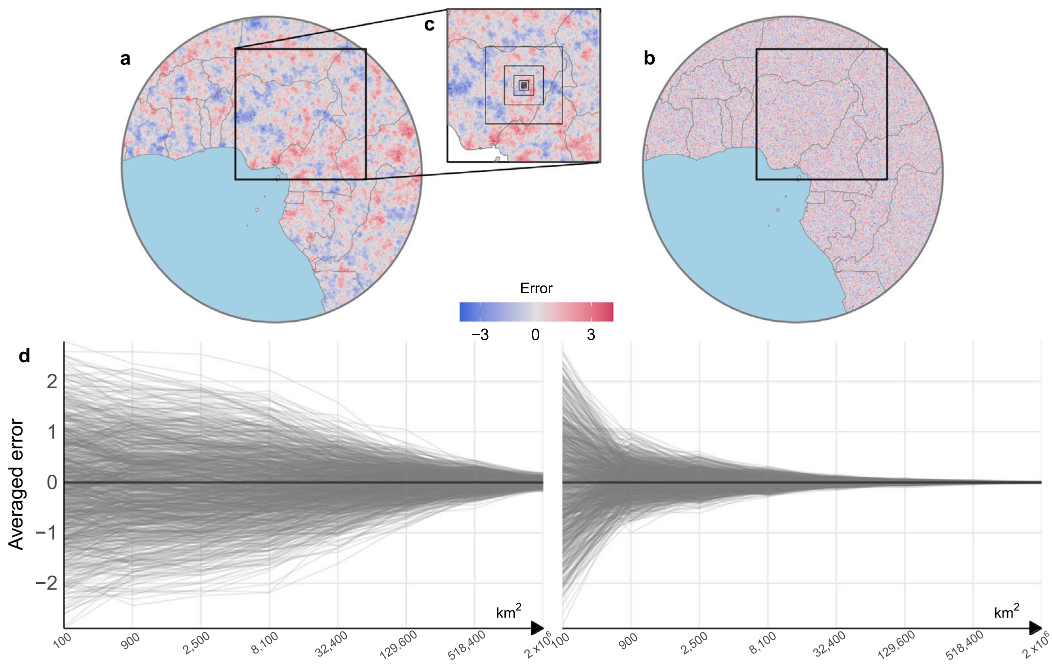

Fig. 1. Illustration of uncertainty reduction for a simulated map error of a climate or ecosystem variable in Western Africa. Two maps showing **a** a case with spatially correlated errors and **b** a case with uncorrelated errors. **c** Eight square areas over which the aggregation of uncertainty is performed, for sizes of 100, 900 2500, 8100, 32,400, 129,600, 518,400 and 2×10^6 km^2. **d** Average error for the eight square areas for the correlated and uncorrelated case over 1000 simulated error maps. The maps shown in **a–b** refer to one of the 1000 simulations, each of the 1000 grey lines in the left and right panels of **d** corresponds with one simulation.

In the dice example all rolls were statistically independent: the outcome of one roll did not influence that of another. But when we compute spatial averages the errors at neighbouring locations will often be positively correlated. This reduces the cancelling out effect. In other words, the uncertainty reduction depends on the degree of spatial dependence in the map error. Unfortunately, this fact is often overlooked in the scientific literature; most studies that quantify uncertainty of spatial aggregates do not properly account for spatial correlation of prediction errors. Disregarding spatial dependence of map errors in spatial aggregation leads to a seriously misleading estimate of uncertainty of the average or total.

This is illustrated with a synthetic example in Fig. 1. We simulated 1000 realizations of two map errors, one with a strong spatial correlation and one with uncorrelated errors. The realizations with strong spatial correlation were obtained from 1000 sequential Gaussian simulations with an exponential variogram with a sill of 1, a nugget of 0 and a distance parameter of 10 km. The values of each realization were then randomly shifted to obtain the realization of the uncorrelated case. The simulated map errors were then averaged for different window sizes. Fig. 1d shows that uncertainty decreases as we average over larger areas and that the uncertainty reduction strongly depends on the degree of spatial correlation. Between the aggregation for sizes 100 and 129,600 km^2, the variance of the average error from all simulations is reduced by 91.6% for the correlated case while it is reduced by 99.9% for the uncorrelated case.

Spatial dependence in map errors can be accounted for with relatively simple geostatistical techniques such as block kriging — the theory of which was developed in the 1970s (Journel and Huijbregts, 1976). Proper estimates of aggregate uncertainty can also be obtained using Monte Carlo integration of the point support uncertainty (Wadoux and Heuvelink, 2023), and can therefore be implemented for any model that reports point support uncertainty (for example, quantile regression forests (Meinshausen, 2006)). If a probability sample of map errors within the area of interest is available, uncertainty of the average or total can also be obtained with design-based statistical inference (Brus et al., 2011). The solutions are there, both in terms of methods and open source software implementations, and there is nothing that stops the scientific community from using them.

CRediT authorship contribution statement

Alexandre M.J.-C. Wadoux: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Gerard B.M. Heuvelink:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101059012.

Data availability

No data was used for the research described in the article.

References

Brus, D.J., Kempen, B., Heuvelink, G.B.M., 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62 (3), 394–407.
Harris, N.L., Gibbs, D.A., Baccini, A., Birdsey, R.A., De Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M.C., Herold, M., Houghton, R.A., et al., 2021. Global maps of twenty-first century forest carbon fluxes. Nat. Clim. Chang. 11 (3), 234–240.
IPCC, 2022. In: Pörtner, H.O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Journel, A.G., Huijbregts, C.J., 1976. Mining Geostatistics. Academic Press, London.
Meinshausen, N., 2006. Quantile regression forests. J. Mach. Learn. Res. 7 (6), 983–999.
Nahlik, A.M., Fennessy, M.S., 2016. Carbon storage in US wetlands. Nat. Commun. 7 (1), 1–9.
Phillips, H.R., Guerra, C.A., Bartz, M.L., Briones, M.J., Brown, G., Crowther, T.W., Ferlian, O., Gongalsky, K.B., Van Den Hoogen, J., Krebs, J., et al., 2019. Global distribution of earthworm diversity. Science 366 (6464), 480–485.
Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H., Van Egmond, F., McNeill, S., Kuhnert, M., et al., 2020. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. Global Change Biol. 26 (1), 219–241.
The United Nations, 2023. The Sustainable Development Goals Report Special edition. Technical Report, The United Nations.
UNFCCC, 2015. Adoption of the Paris Agreement. Technical Report, FCCC/CP/2015/10/Add.1.
UNFCCC, 2023. National inventory submissions 2023.
Wadoux, A.M.J.-C., Heuvelink, G.B.M., 2023. Uncertainty of spatial averages and totals of natural resource maps. Methods Ecol. Evol. 14, 1320–1332.
Xiao, D., Deng, L., Kim, D.-G., Huang, C., Tian, K., 2019. Carbon budgets of wetland ecosystems in China. Global Change Biol. 25 (6), 2061–2076.