

Supporting Information for “Source region geochemistry from unmixing downstream sedimentary elemental compositions”

Alex G. Lipp¹, Gareth G. Roberts¹, Alexander C. Whittaker¹,

Charles J. B. Gowing², Victoria M. Fernandes¹

¹Department of Earth Science and Engineering, Imperial College London, UK

²Centre for Environmental Geochemistry, British Geological Survey, Keyworth, UK

Contents of this file

- Figures S1 to S25

Figure S1 shows the results of ‘chequerboard’ tests. In these tests, synthetic data at sample sites downstream (white circles in Figure S1b) are generated using arbitrary elemental concentrations (e.g. Figure S1a) and the forward model described in the body text. Concentrations at the 67 downstream samples are then inverted for the composition of source regions (see color map in e.g. Figure S1a). These tests are performed using the actual sample sites considered in this study and the real drainage networks in the study region. A comparison between the ‘actual’ source region concentrations and best-fitting results from inverting the 67 sample sites is shown in adjacent panels (e.g. Figure S1c). Figure S2 shows the results from a similar test in which source region composition changes

abruptly. As discussed in the body text of the main manuscript, changes in source composition at wavelengths < 20 km are poorly resolved. In contrast, the amplitude and spatial structure of longer wavelength changes in composition are recovered. These results and the following tests are discussed in the main manuscript. Figure S3 shows the result from an synthetic inversion using the same input as Figure 5, but where random Gaussian noise equal to 5 % of the total variance was added to the data before inverting. The results show that despite this noise the spatial geochemical signal is recovered nearly as successfully as in the test without noise.

Figure S4 shows a comparison between magnesium concentrations in the full resolution G-BASE dataset and predictions from the smooth inverse model. Unsurprisingly, the full resolution G-BASE dataset contains more short wavelength variability than the predictions from the smooth inverse model. Figure S5a shows global RMS misfit between observed and predicted upstream magnesium concentrations (see Figures 7 & 9 in main manuscript) as a function of smoothing coefficient, λ . Figure S4b shows associated R^2 values. The red arrows indicate the optimal λ value used to invert for upstream concentration following the protocol described in the main manuscript (see Figure 9a). The black arrows indicate the value that yields the minimum misfit to G-BASE observations.

Figure S6 shows the results from applying a low-pass (> 25 km) Gaussian filter to the magnesium data extracted from the G-BASE survey, and to the results of the inverse model (cf. Figure 13 of the main manuscript).

Figures S7–S25 show the results from inverting the elemental concentrations of actual samples for source region chemistry. In the main manuscript we show results for Mg and a subset of results for Ca, Rb, V and Be. Figures S7–S25 shows the best-fitting inverse model, the G-BASE inventory and comparisons between these estimates of concentration for Ba, Be, Ca, Co, Cr, Fe, K, La, Li, Mn, Ni, Rb, Sr, Ti, U, V, Y, Zn and Zr. The smoothing parameters for each inverse model were determined by systematically varying λ , the optimum values for each element are given in the captions for Figures S7–S25. See body text of the main manuscript for details.

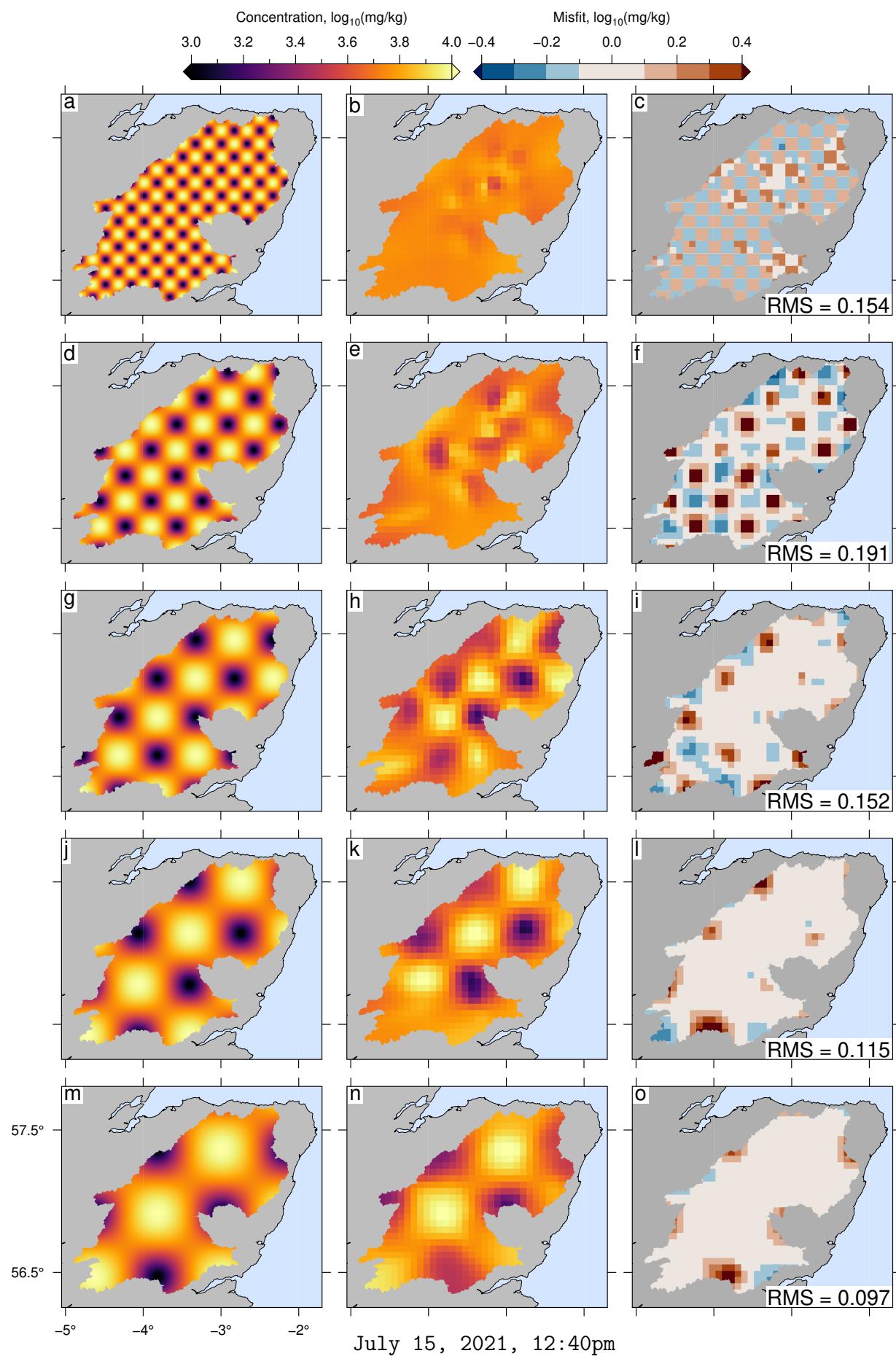
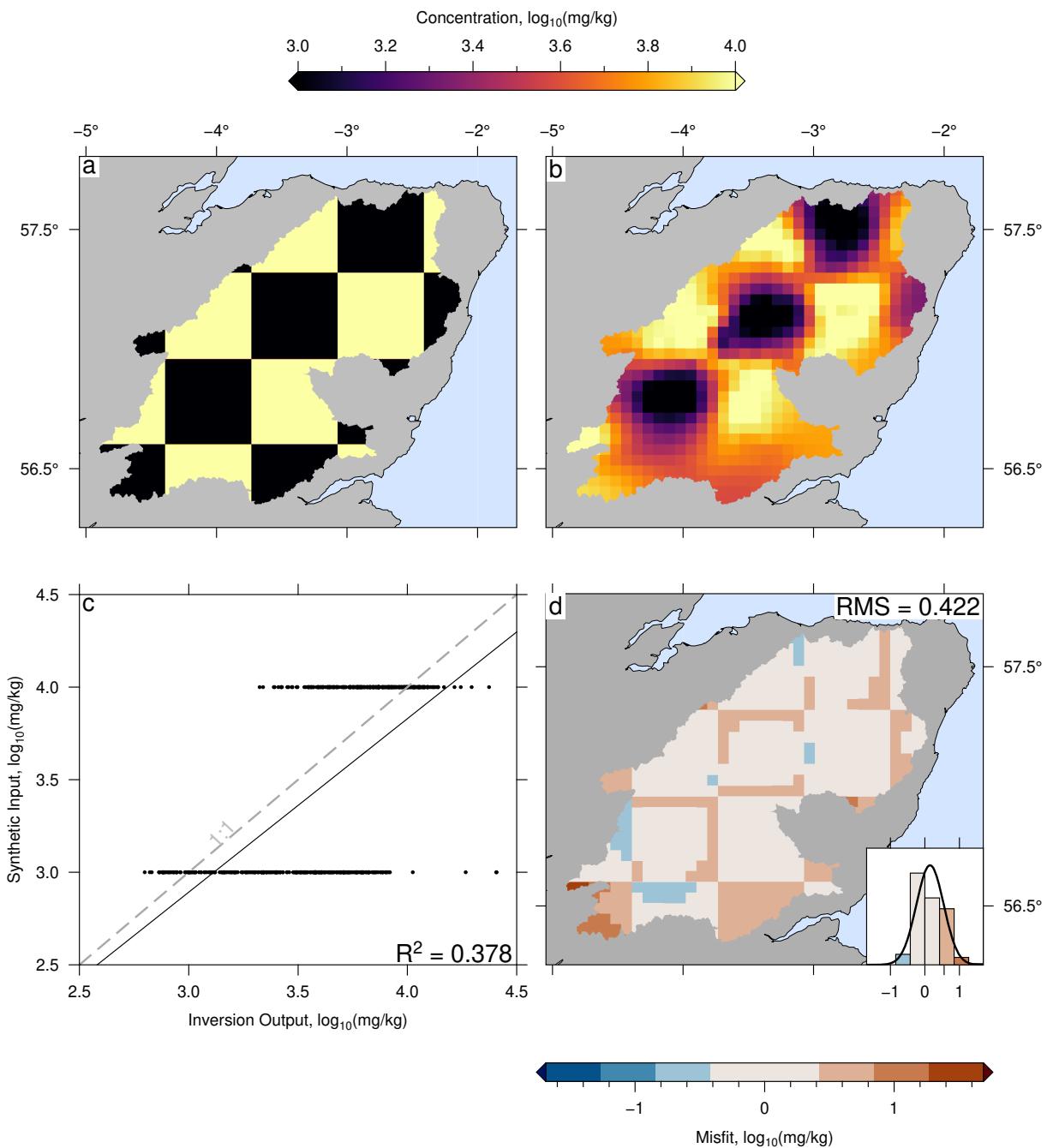


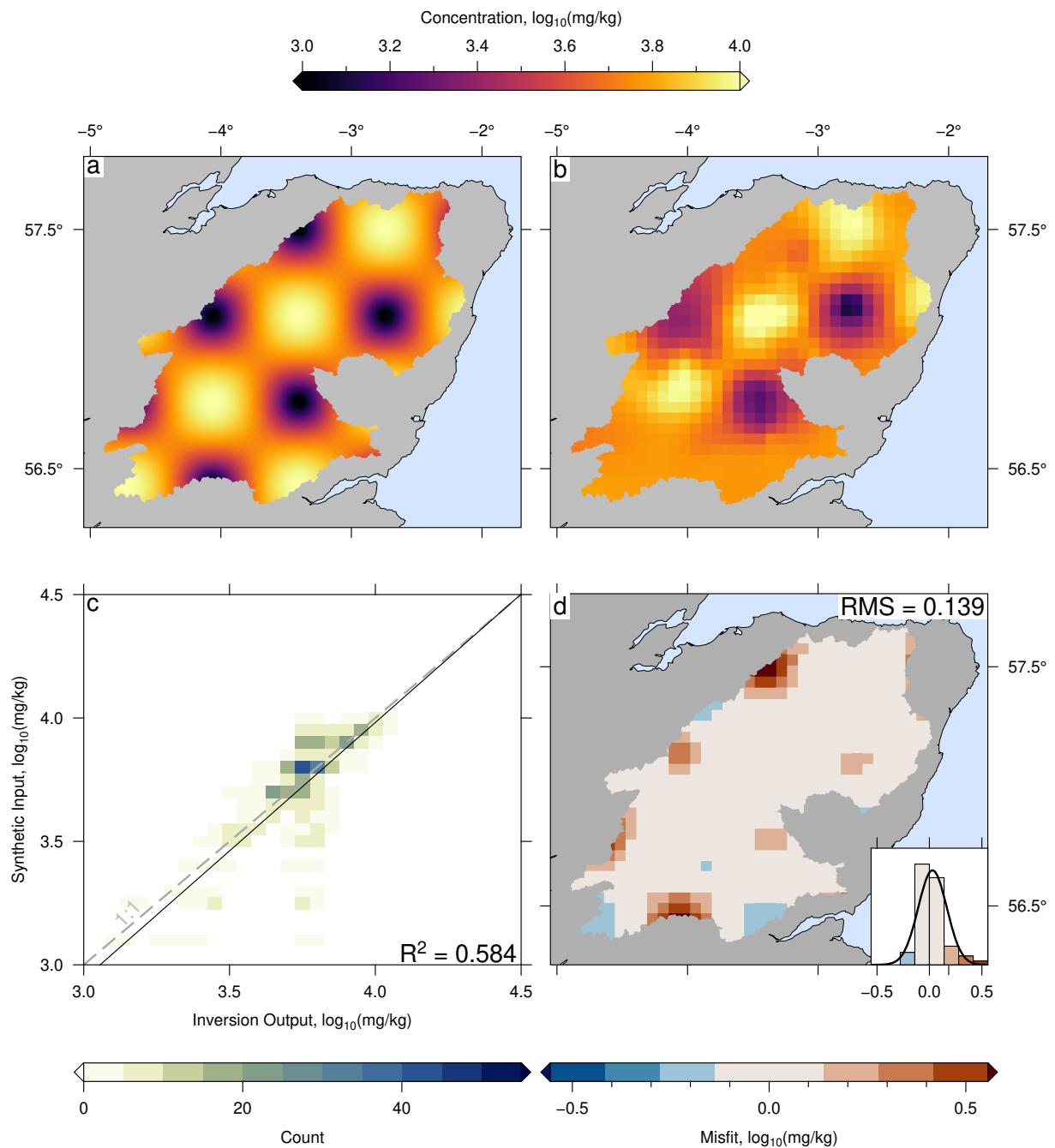
Figure S1. Testing the resolution of the inverse model using synthetic examples.

Synthetic inverse models analogous to those shown in Figures 5–6 of the main manuscript. (a) Synthetic elemental concentration map generated using a 2D sine function with peak to trough = 10 km. This map was used to calculate composition downstream (e.g. at the 67 sample sites shown by white circles in panel b). (b) Output of best-fitting inverse model. (c) Misfit between maps of ‘observed’ and best-fitting theoretical composition. (d–f) Results when synthetic input has peak-to-trough distance = 20 km. (g–i) Peak-to-trough distance = 30 km. (j–l) Peak-to-trough distance = 40 km (see Figure 5 in main manuscript). (m–o) Peak-to-trough distance = 50 km.



July 15, 2021, 12:40pm

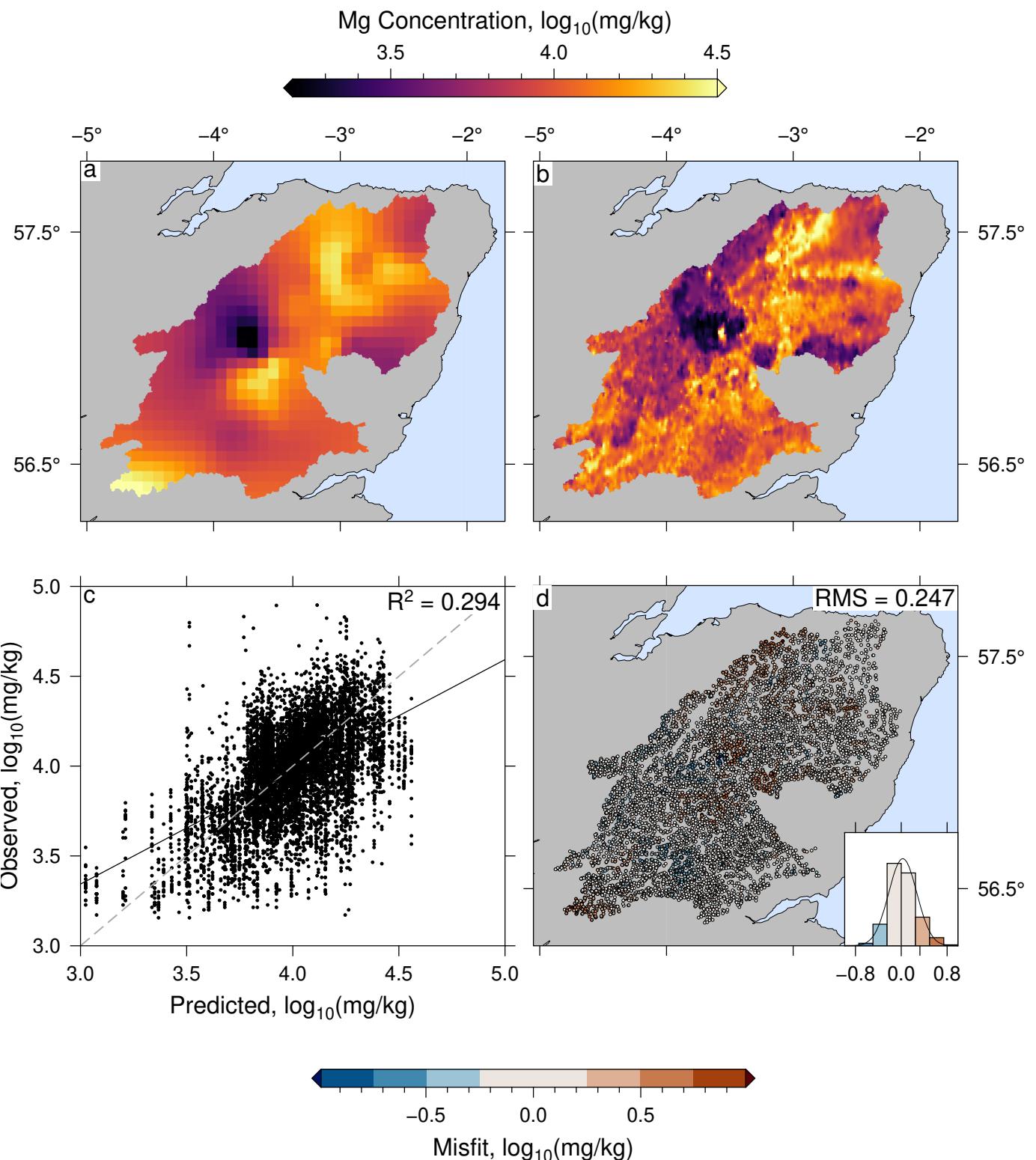
Figure S2. Predicting geochemistry in source regions: An example of solving the inverse problem. This figure is analogous to Figure 5 in main manuscript except input signal has abrupt changes in geochemistry rather than gradual changes. (a) Map of synthetic element concentrations in source regions generated as a ‘chequerboard’ with width 40 km. This map was used to calculate sediment concentrations at sample sites downstream by solving the forward problem, which were then inverted for source composition. (b) Predicted source region composition calculated by inverting synthetic compositions at the 67 sample sites (see body text). In this example smoothing parameter $\lambda = 10^{-0.5}$. (c) Cross-plot of observed and predicted source region concentrations; grid resolution of observed and predicted composition is 5×5 km (see panel b). 1:1 relationship is shown by gray dashed line; black solid line = linear regression. (d) Misfit between observed and predicted source composition. Color bar is discretised on intervals equal to global RMS misfit. Misfit is highest in regions of low sample coverage (see Figure 2c). Inset shows histogram of misfits with binwidth = global RMS misfit; best fitting normal distribution (black curve) is shown for comparison.



July 15, 2021, 12:40pm

Figure S3. Predicting geochemistry in source regions: An example of solving the inverse problem. This figure is analogous to Figure 5 in main manuscript except noise equal to 5 % of the total variance was added to the downstream data before inverting. (a) Map of synthetic element concentrations in source regions generated as a ‘chequerboard’ with width 40 km. This map was used to calculate sediment concentrations at sample sites downstream by solving the forward problem, which were then inverted for source composition. (b) Predicted source region composition calculated by inverting synthetic compositions, with added noise, at the 67 sample sites. In this example smoothing parameter $\lambda = 10^0$. (c) Heat-map of observed and predicted source region concentrations; grid resolution of observed and predicted composition is 5×5 km (see panel b). 1:1 relationship is shown by gray dashed line; black solid line = linear regression. (d) Misfit between observed and predicted source composition. Color bar is discretised on intervals equal to global RMS misfit. Misfit is highest in regions of low sample coverage (see Figure 2c). Inset shows histogram of misfits with binwidth = global RMS misfit; best fitting normal distribution (black curve) is shown for comparison.

X - 10



July 15, 2021, 12:40pm

Figure S4. Inverting real downstream sediment samples for concentration of magnesium in source regions: Comparison to full resolution G-BASE dataset. (a) Optimum upstream concentration of magnesium generated by inverting the magnesium concentration of the 67 samples gathered downstream with smoothing parameter $\lambda = 10^{0.3}$ (see Figures 4, 8a & body text for details). (b) Independent G-BASE stream sediment concentration of magnesium. (c) Cross-plot of observed (G-BASE) and predicted concentrations for each G-BASE sample. Colors show misfit discretised at intervals equal to global RMS misfit. Gray dashed line = 1:1 relationship; black line = linear regression. (d) Misfit between observed magnesium concentration and best-fitting inverse model. Inset indicates distribution of residuals and normal distribution; bin-width = global RMS misfit. Note higher residuals in regions of low coverage identified in Figure 2c.

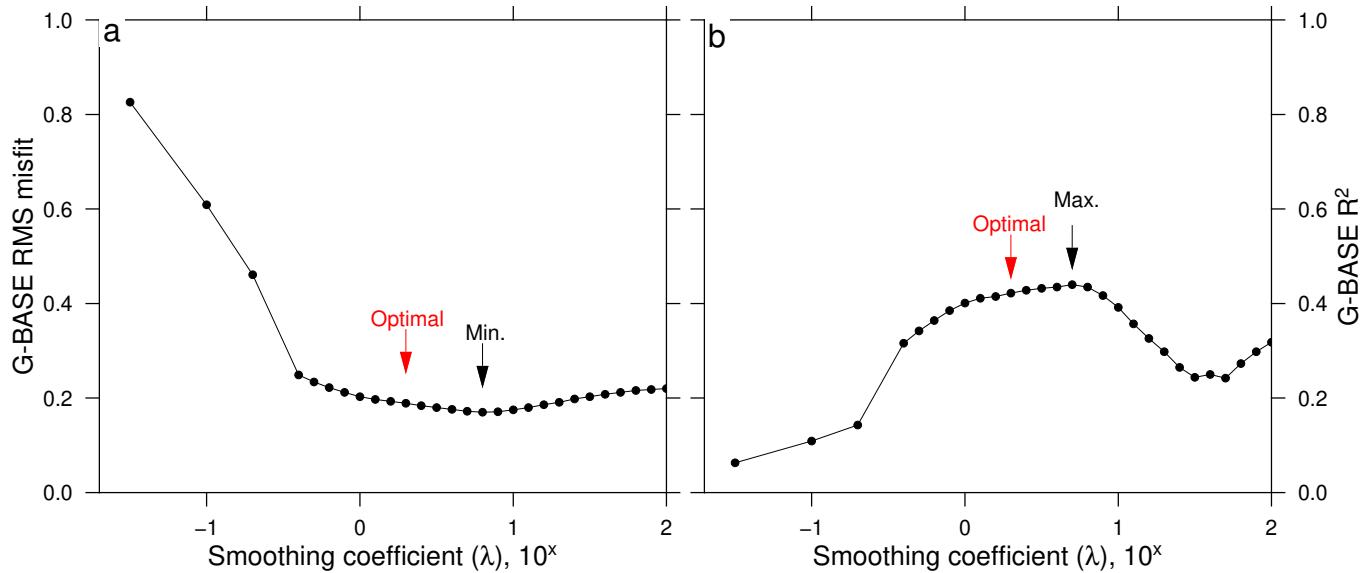


Figure S5. Comparing independent data to Mg predictions for different smoothing parameters. (a) RMS misfit between model predictions and gridded G-BASE dataset for different smoothing parameters, λ . The solution which has the minimum RMS misfit (and hence could be considered ‘best’ is highlighted with a black arrow. Those model chosen by the ‘elbow’ method used in this study is indicated by red arrow. This figure shows that our subjective approach can identify solutions close to the minima. (b) Same as panel (a) for R^2 .

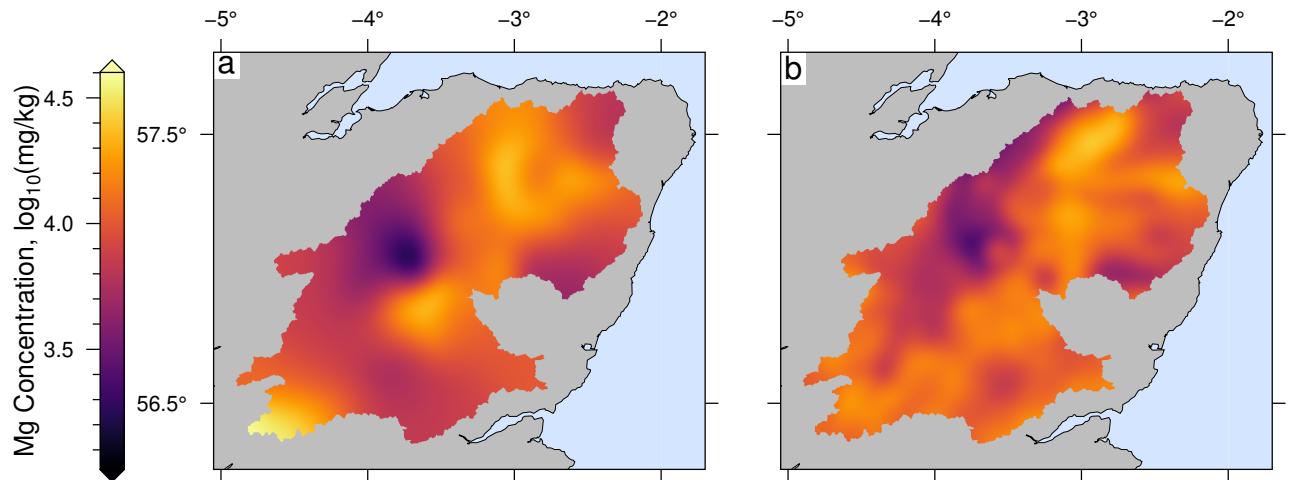
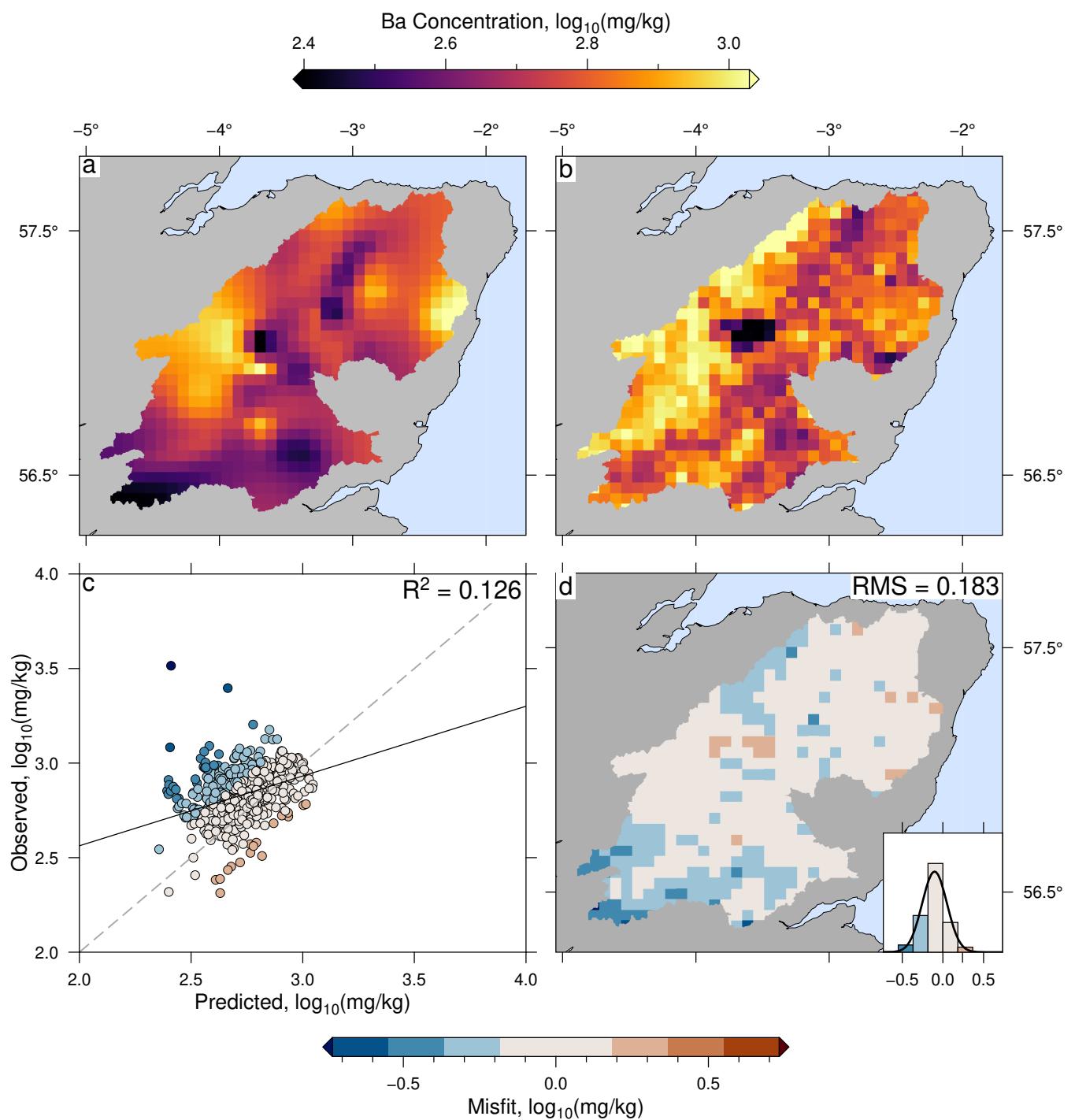


Figure S6. Low-pass filtering of Magnesium. (a) Best-fitting inverse result for Mg filtered using a 2D Gaussian filter of wavelength 25 km. See Figure 7b of the main manuscript for unfiltered results. (b) G-BASE Mg data filtered using same filter as panel (a). Filtered results for other elements are given in Figure 13 of the main manuscript.



July 15, 2021, 12:40pm

Figure S7. Inverting real downstream sediment samples for concentration of barium in source regions. (a) Optimum upstream concentration of barium generated by inverting the barium concentration of the 67 samples gathered downstream with smoothing parameter $\lambda = 10^{-0.1}$. (b) Independent G-BASE stream sediment concentration of barium gridded to same resolution as panel (a). (c) Cross-plot of observed (G-BASE) and predicted concentrations for each grid cell (5 km resolution). Colors show misfit discretized at intervals equal to global RMS misfit. Gray dashed line = 1:1 relationship; black line = linear regression. (d) Misfit between observed barium concentration and best-fitting inverse model. Inset indicates distribution of residuals and normal distribution; binwidth = global RMS misfit.

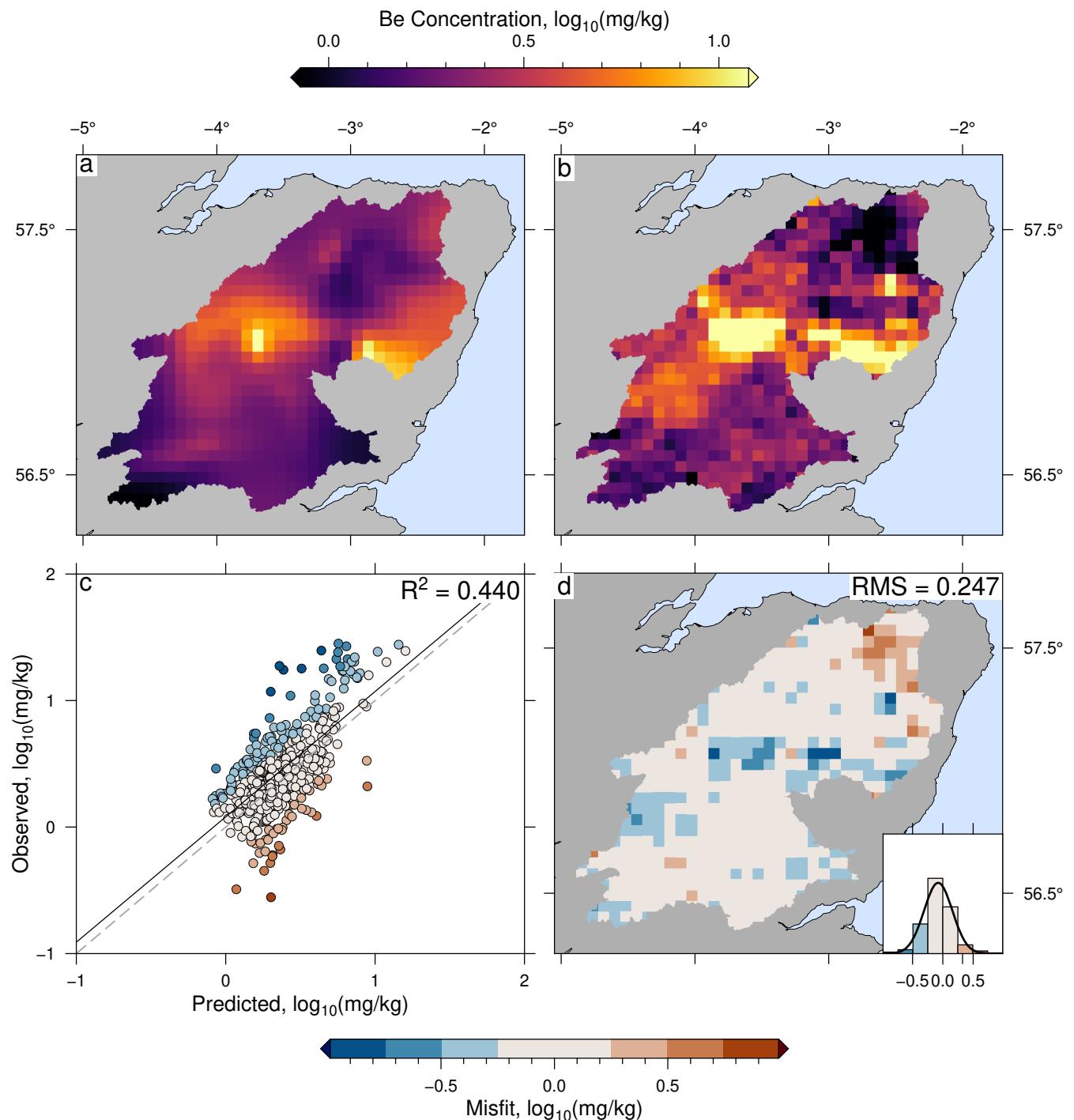


Figure S8. Same as Figure S7 for Be, $\lambda = 10^{0.3}$

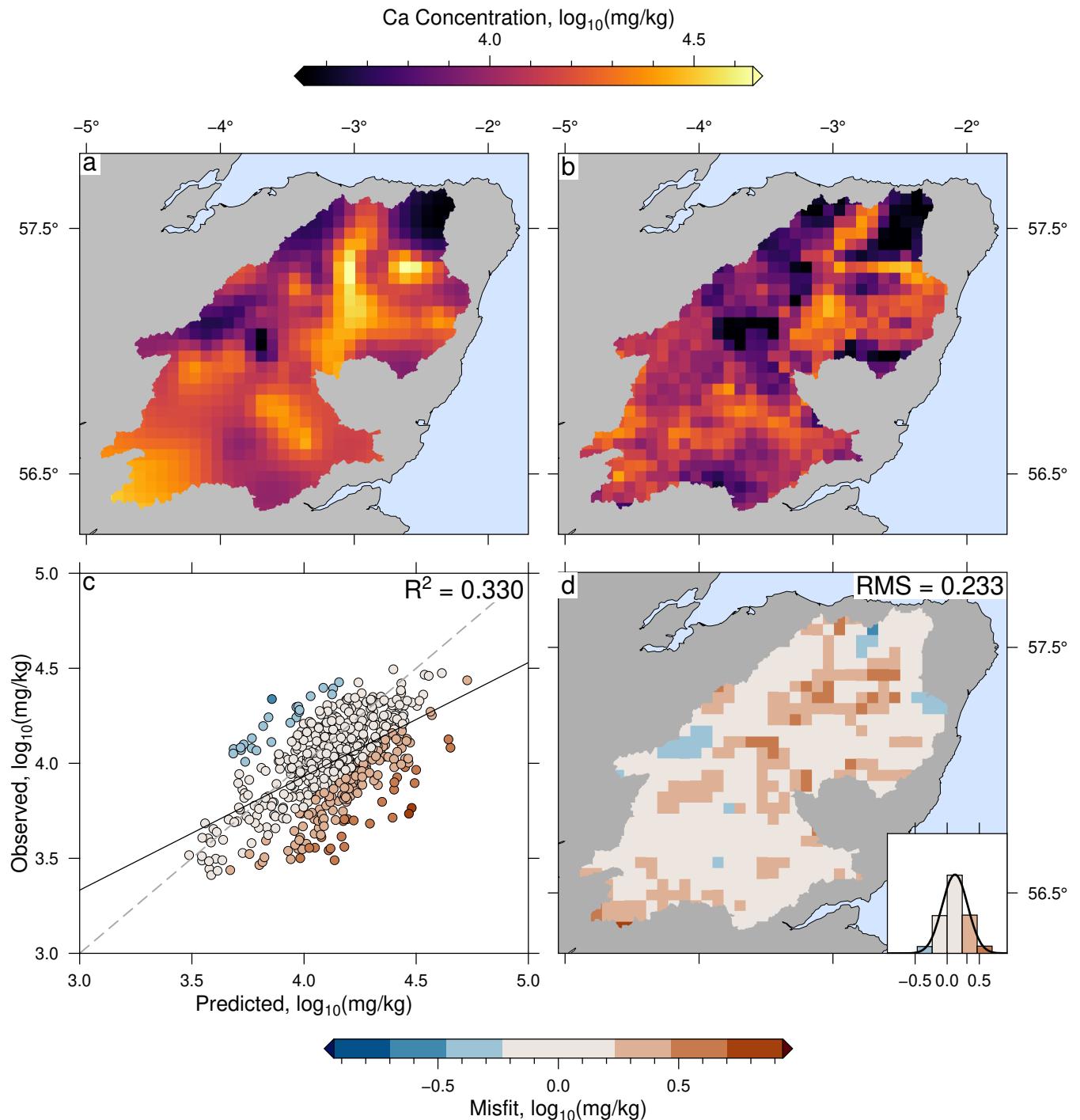


Figure S9. Same as Figure S7 for Ca, $\lambda = 10^{-0.3}$

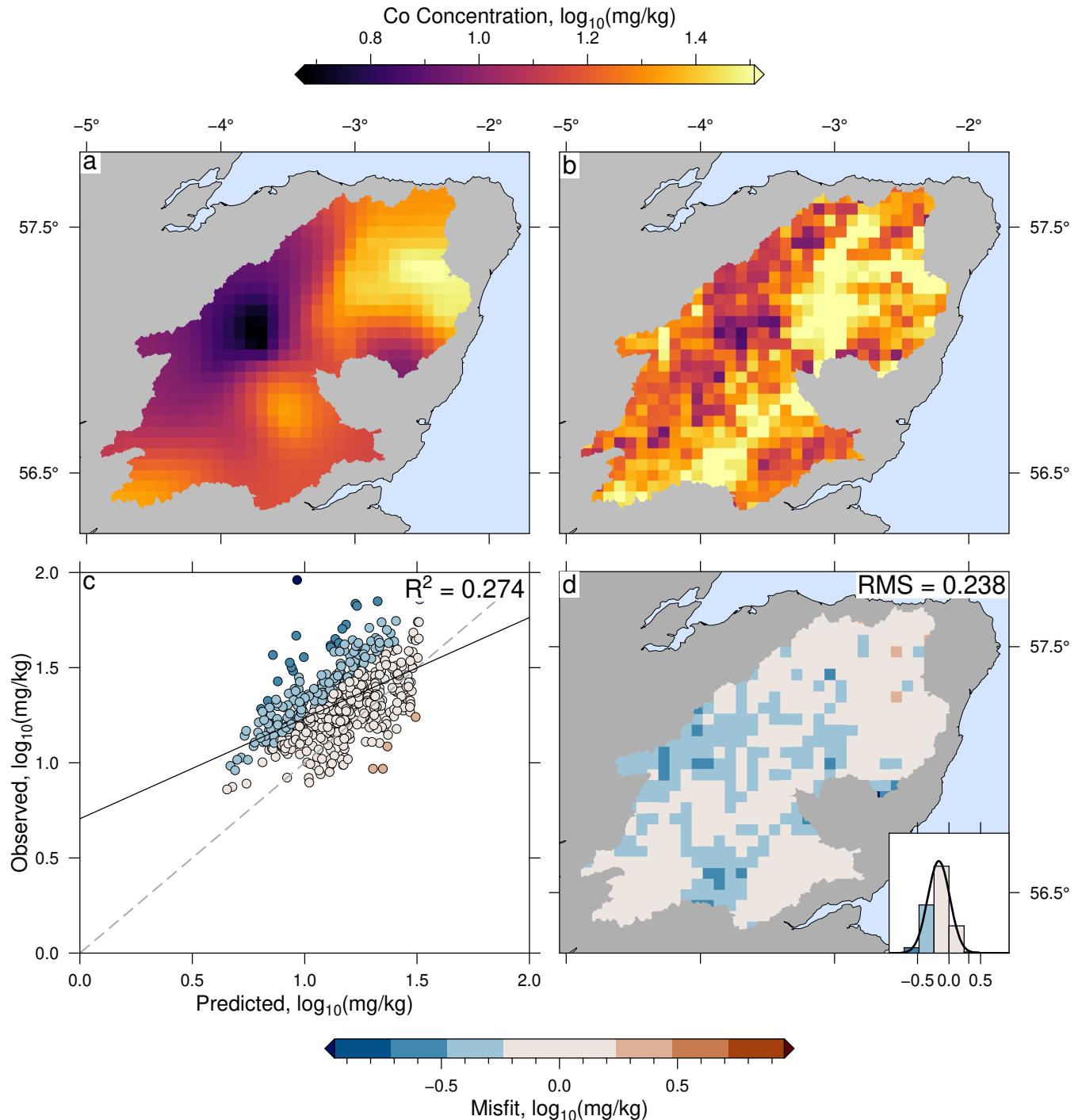


Figure S10. Same as Figure S7 for Co, $\lambda = 10^{0.7}$

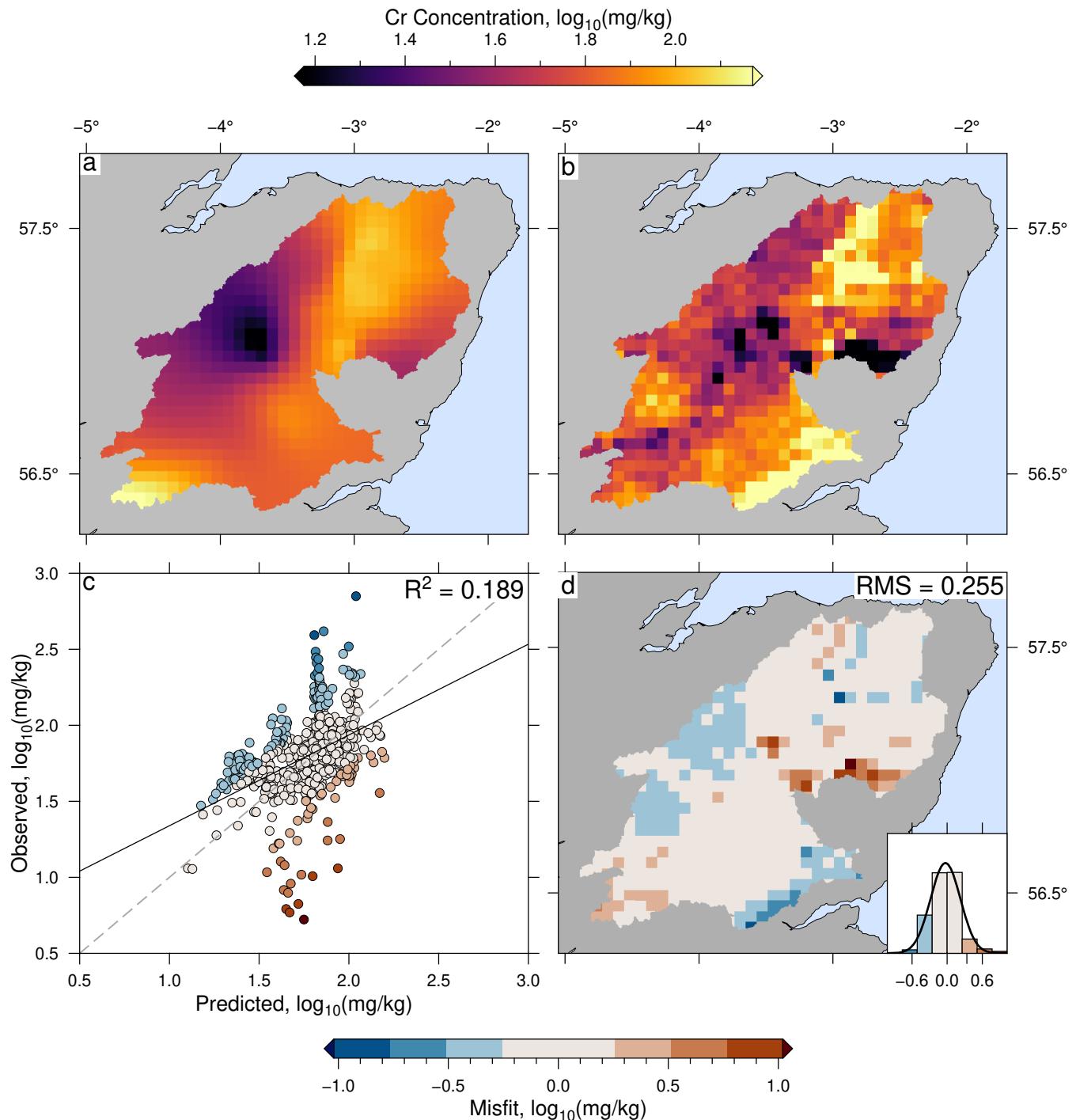


Figure S11. Same as Figure S7 for Cr, $\lambda = 10^{0.8}$

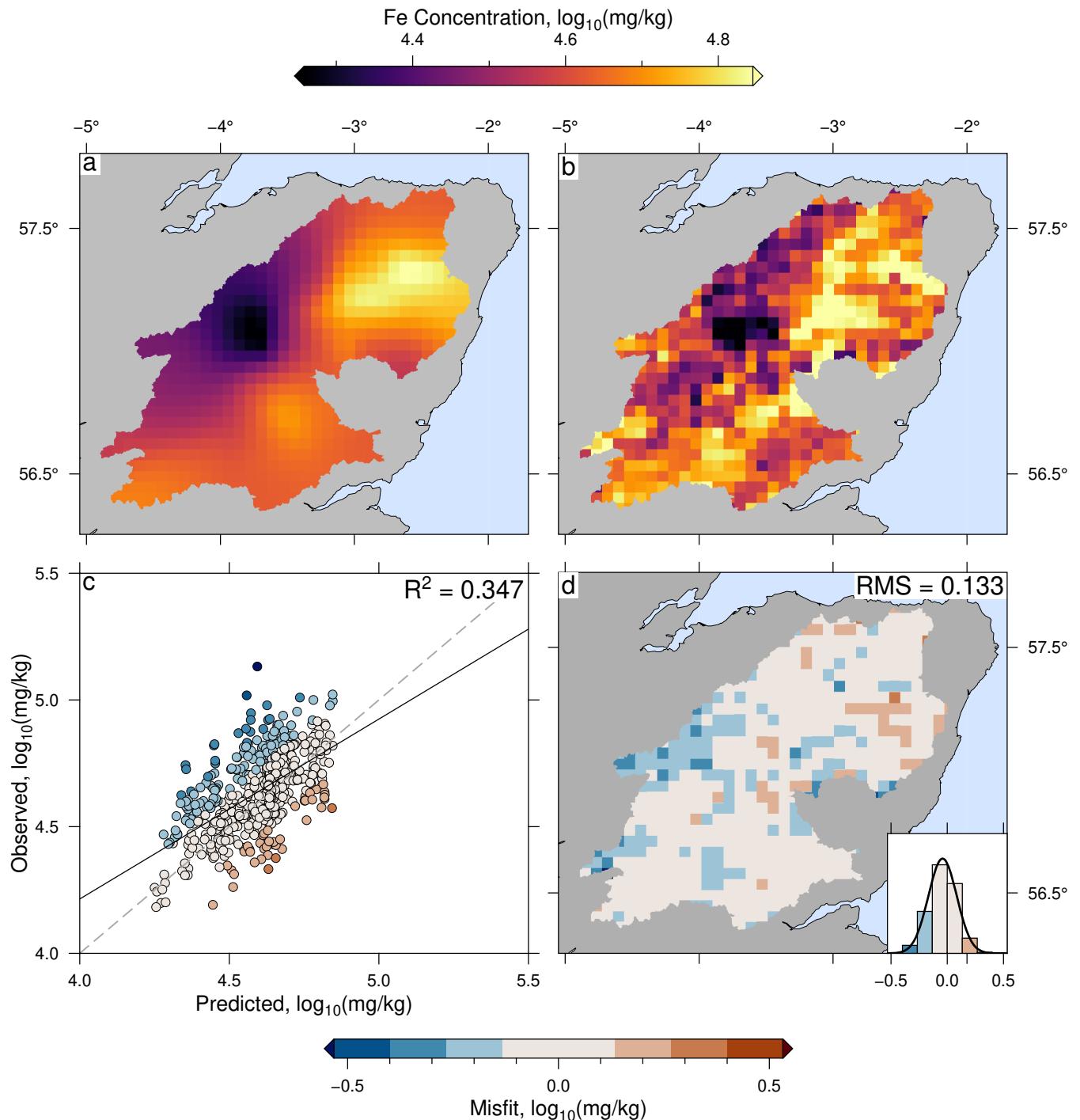


Figure S12. Same as Figure S7 for Fe, $\lambda = 10^{0.8}$

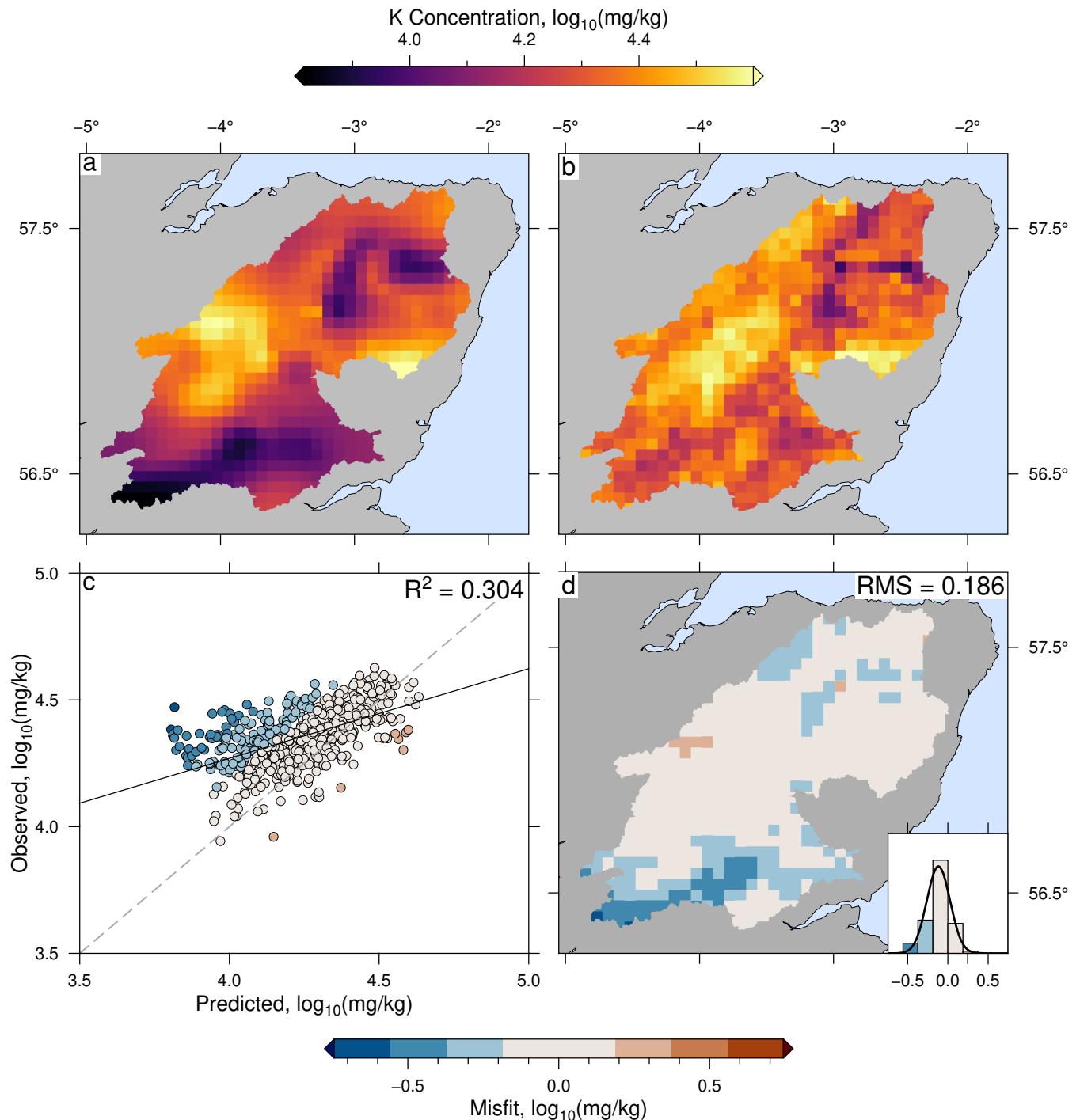


Figure S13. Same as Figure S7 for K, $\lambda = 10^0$

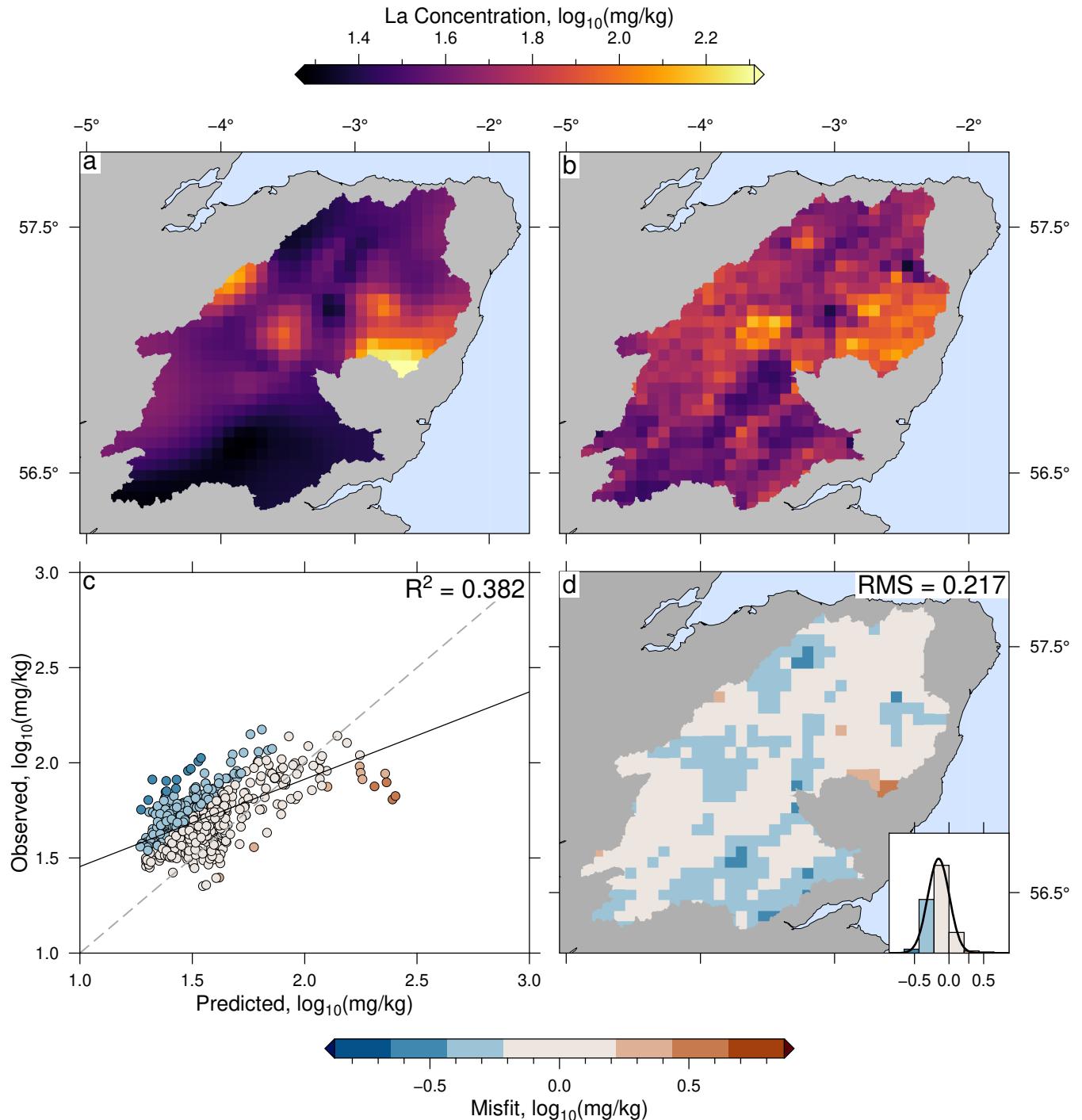


Figure S14. Same as Figure S7 for La, $\lambda = 10^{0.2}$

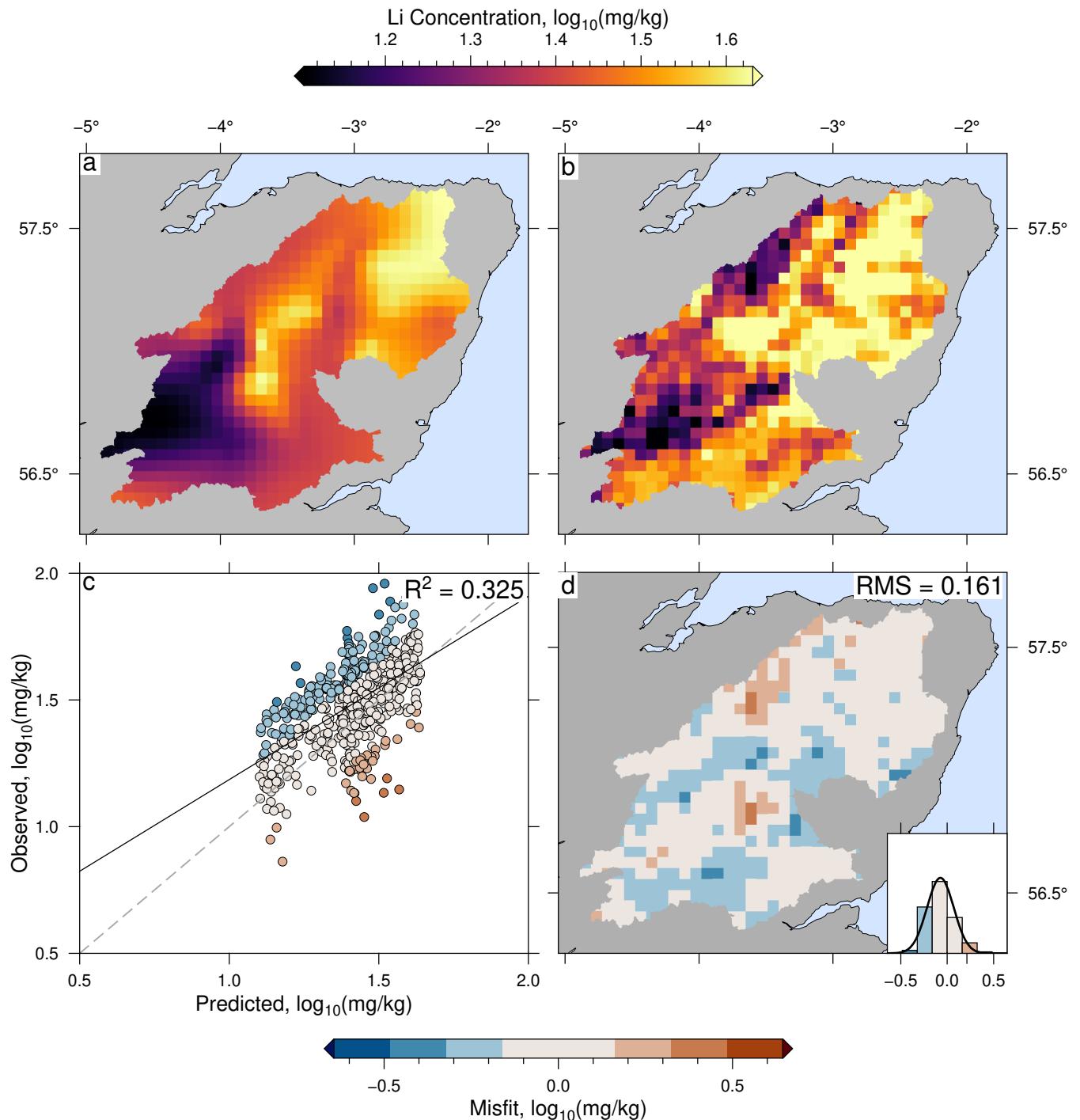


Figure S15. Same as Figure S7 for Li, $\lambda = 10^{0.5}$

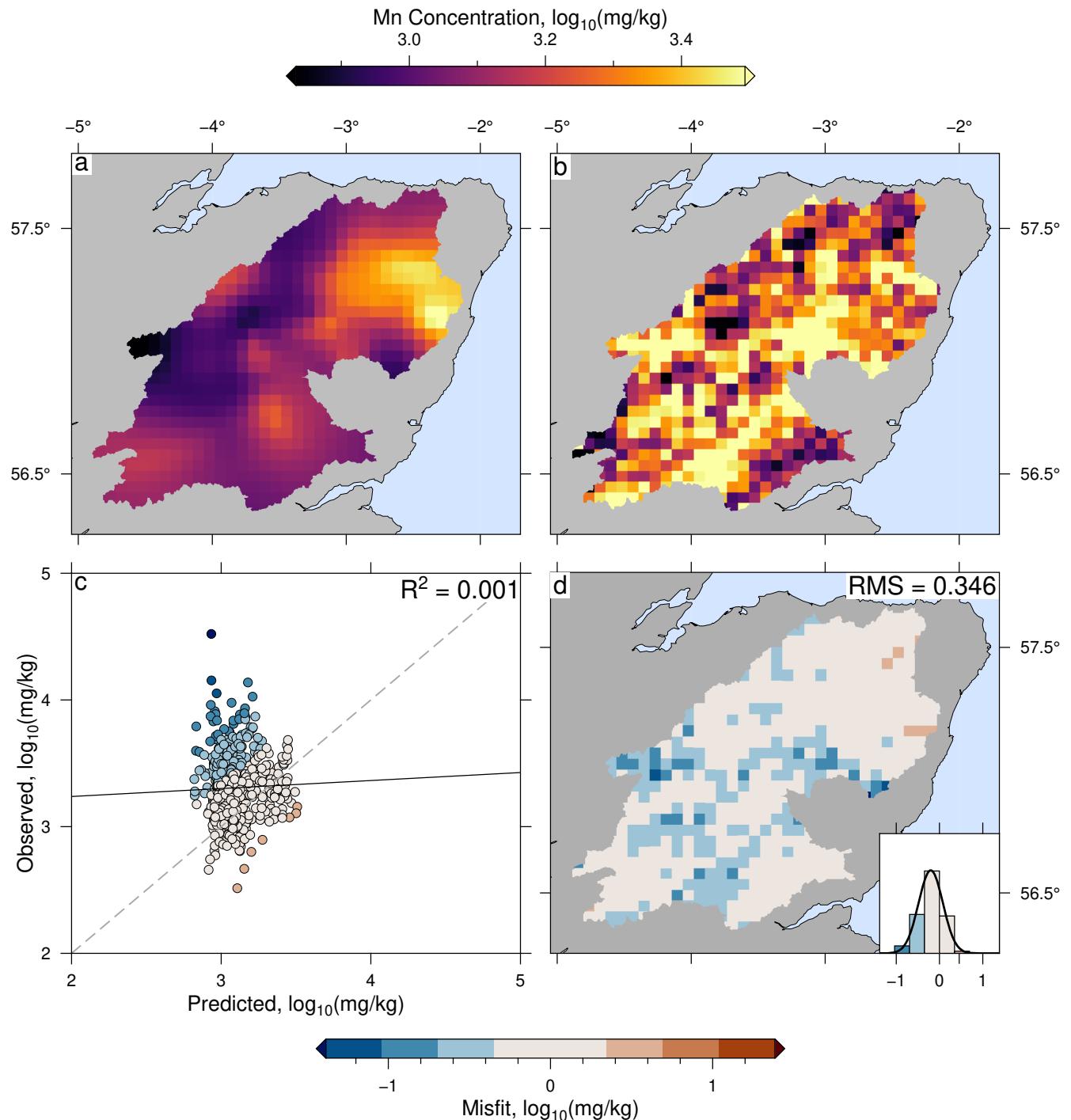


Figure S16. Same as Figure S7 for Mn, $\lambda = 10^{0.7}$

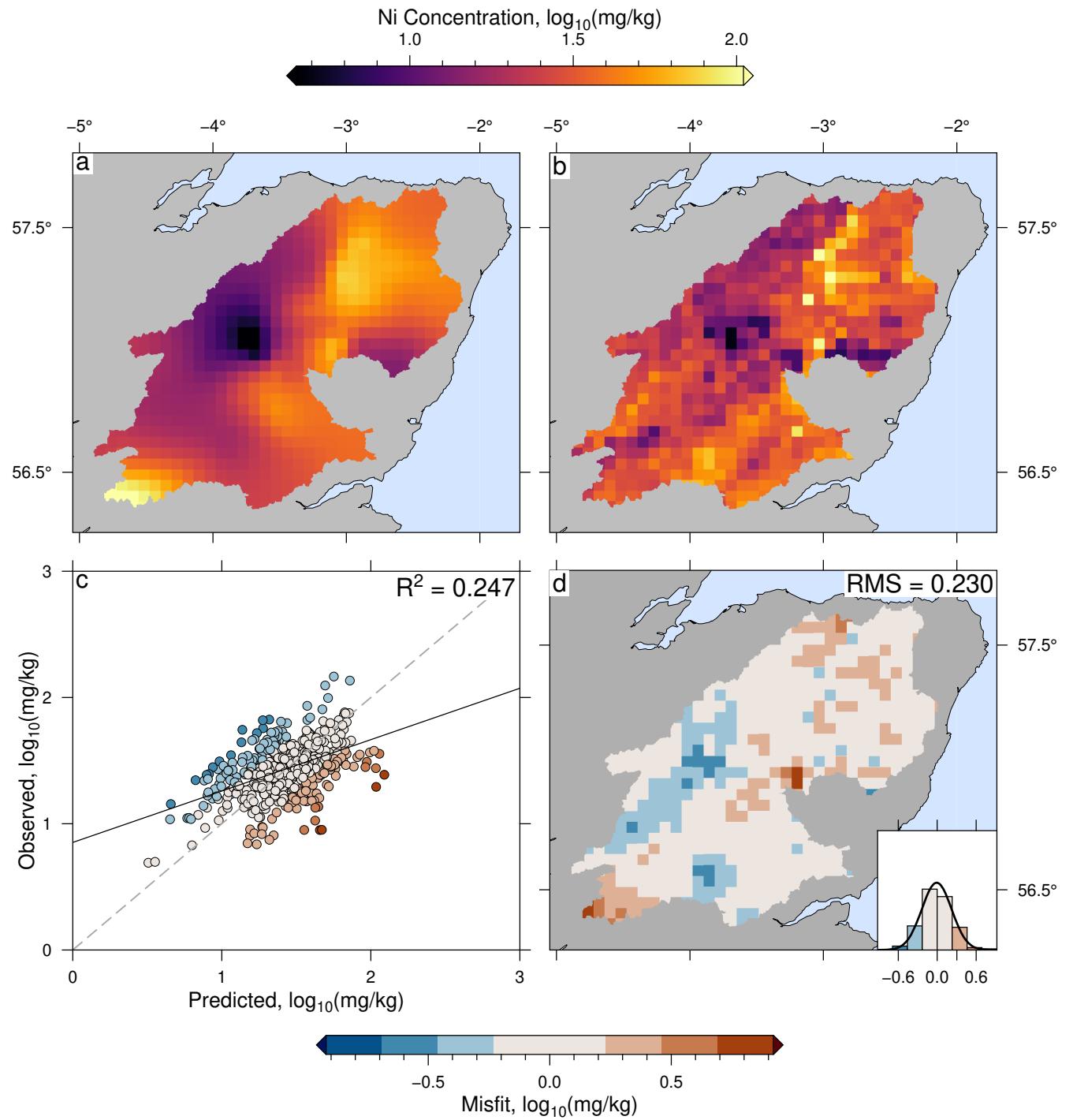


Figure S17. Same as Figure S7 for Ni, $\lambda = 10^{0.5}$

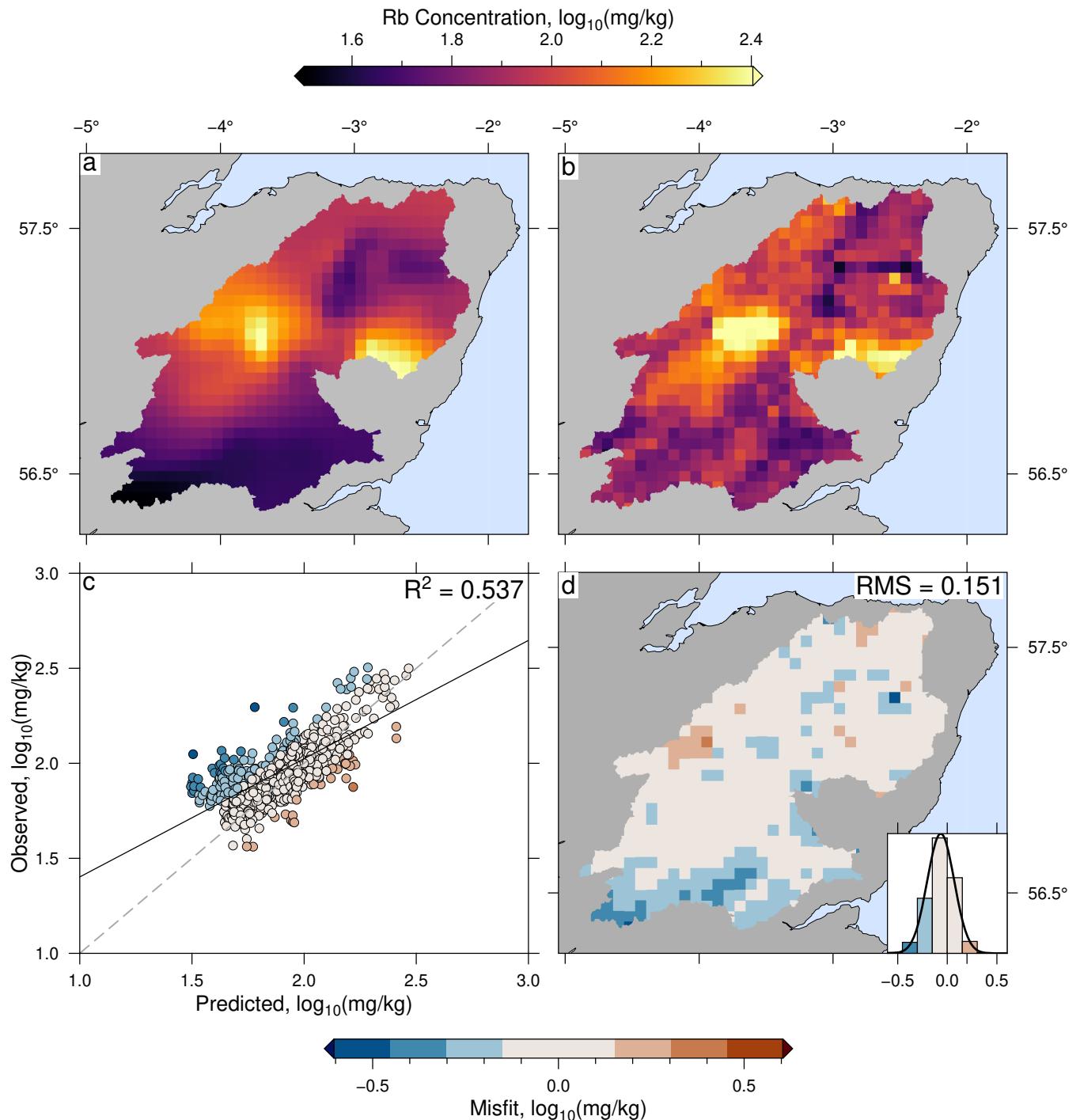


Figure S18. Same as Figure S7 for Rb, $\lambda = 10^{0.4}$

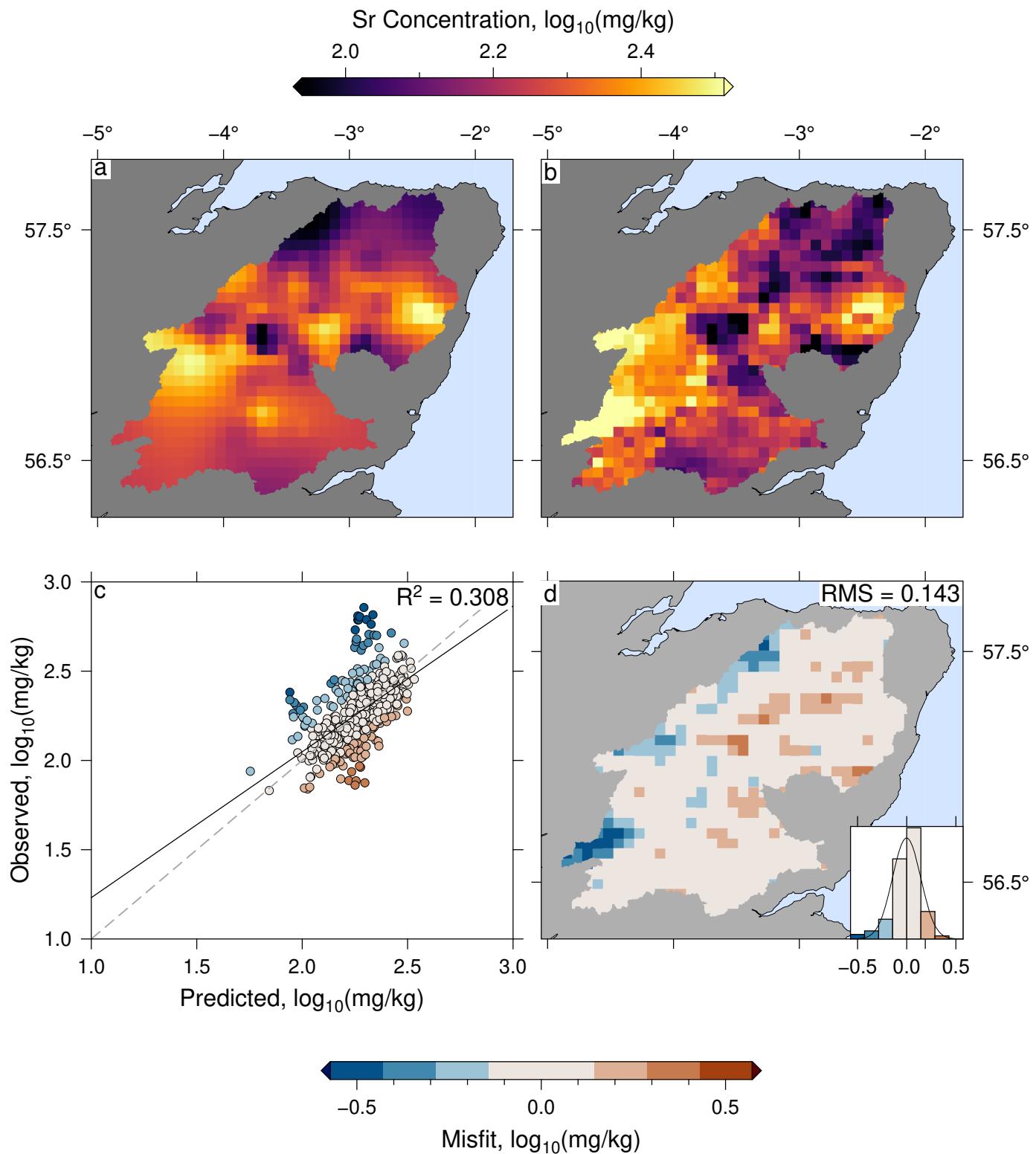


Figure S19. Same as Figure S7 for Sr, $\lambda = 10^{-0.1}$

July 15, 2021, 12:40pm

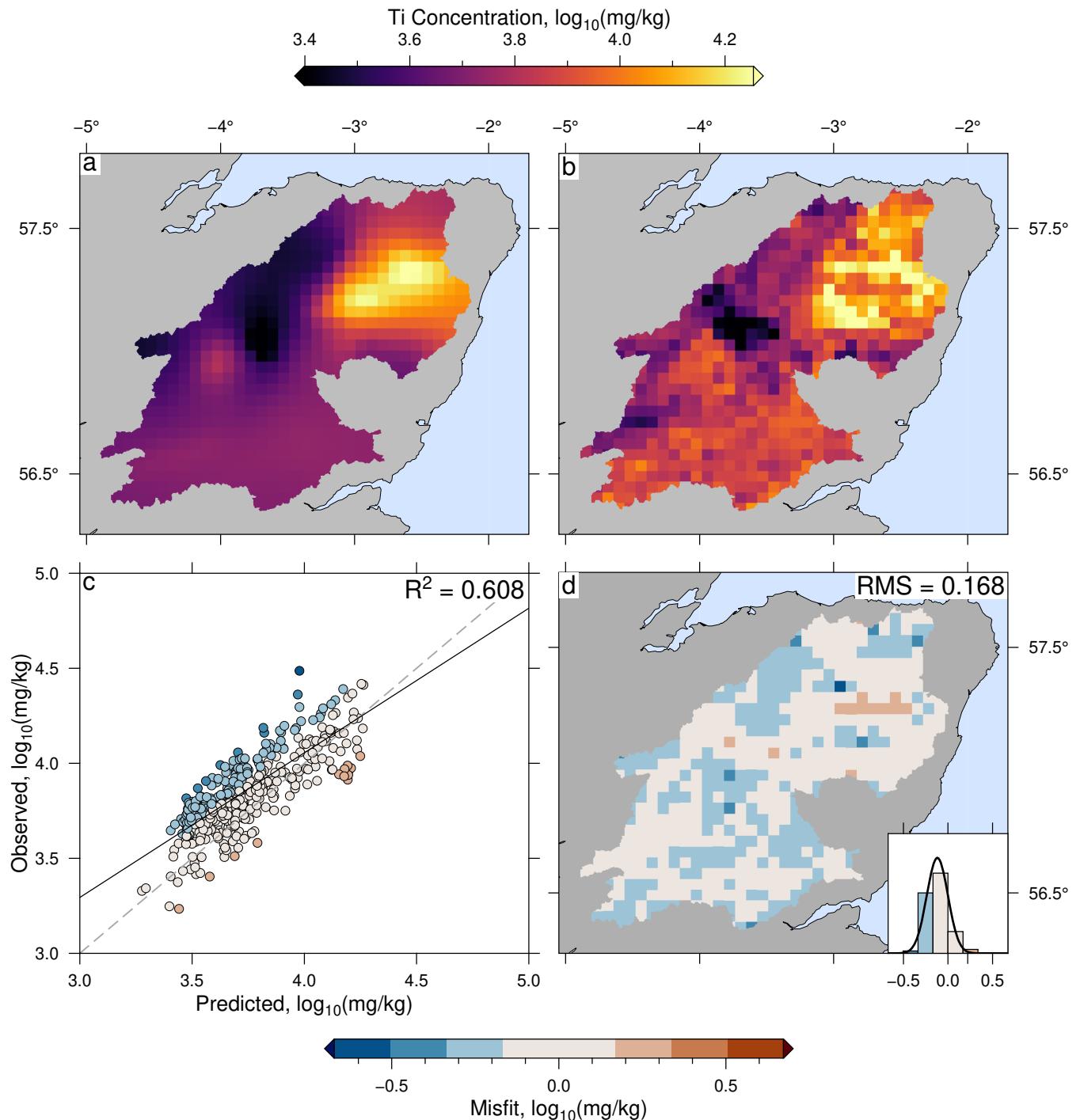


Figure S20. Same as Figure S7 for Ti, $\lambda = 10^{0.7}$

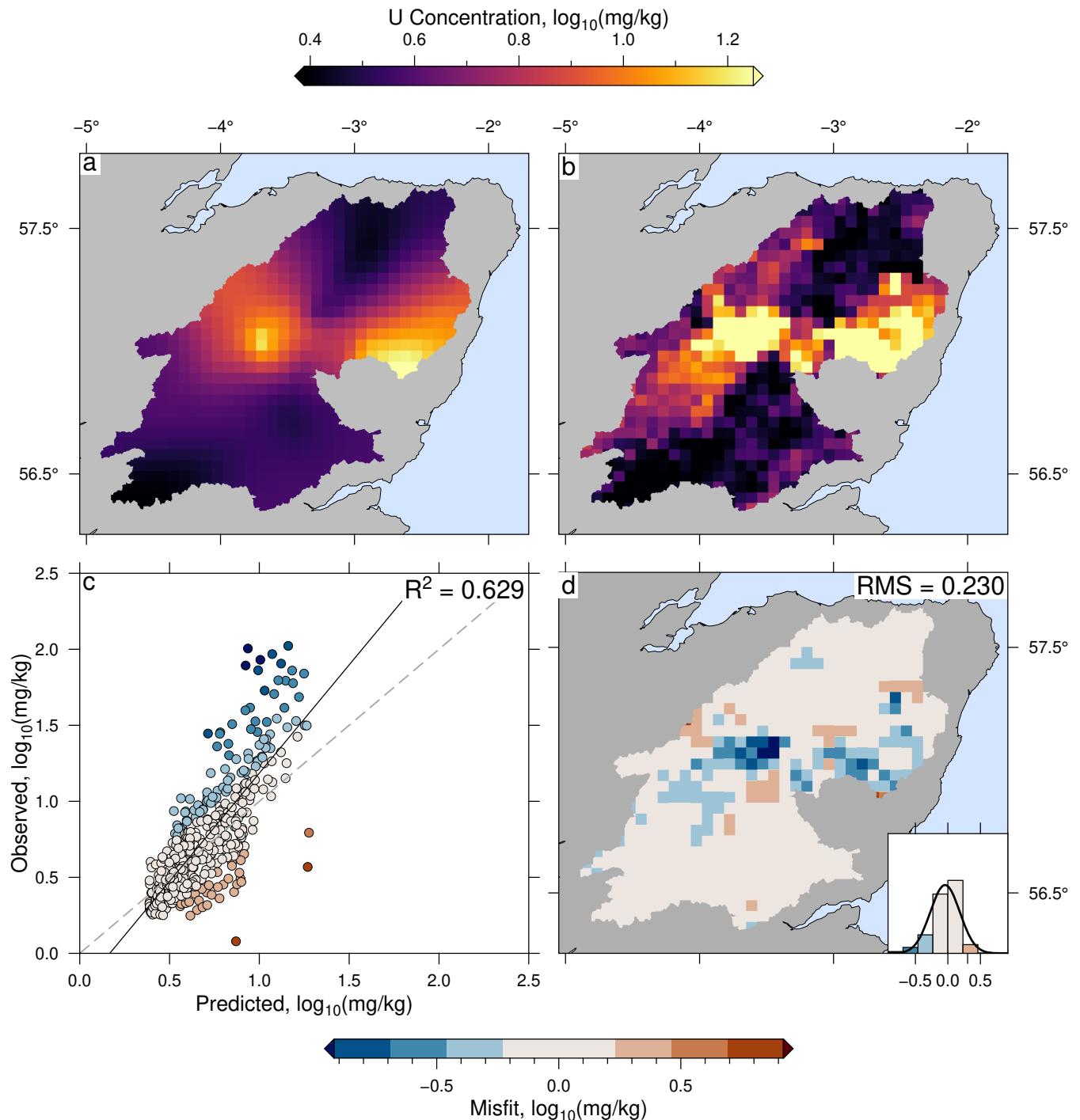


Figure S21. Same as Figure S7 for U, $\lambda = 10^{1.0}$

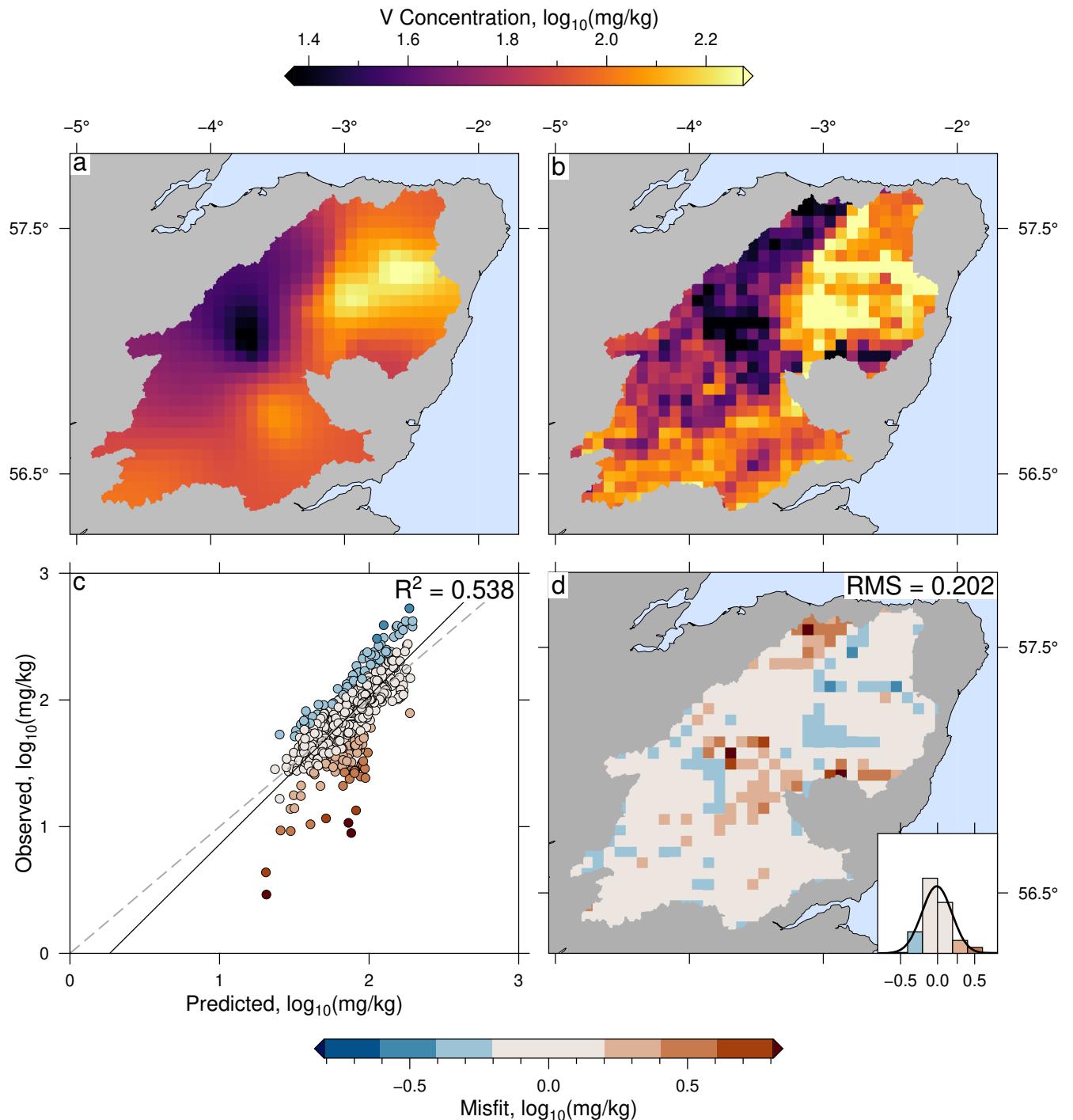


Figure S22. Same as Figure S7 for V, $\lambda = 10^{0.7}$

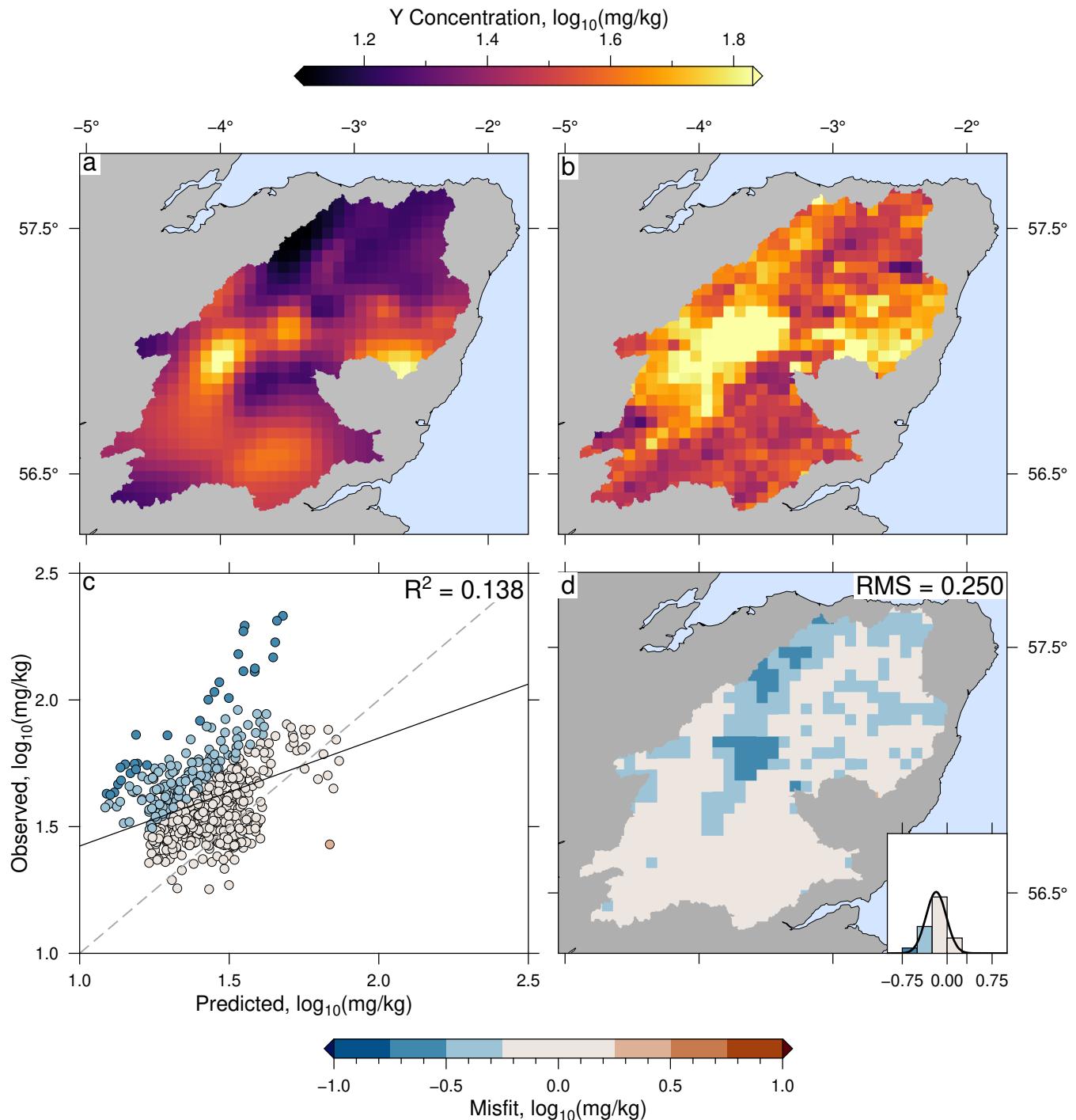


Figure S23. Same as Figure S7 for Y, $\lambda = 10^{0.2}$

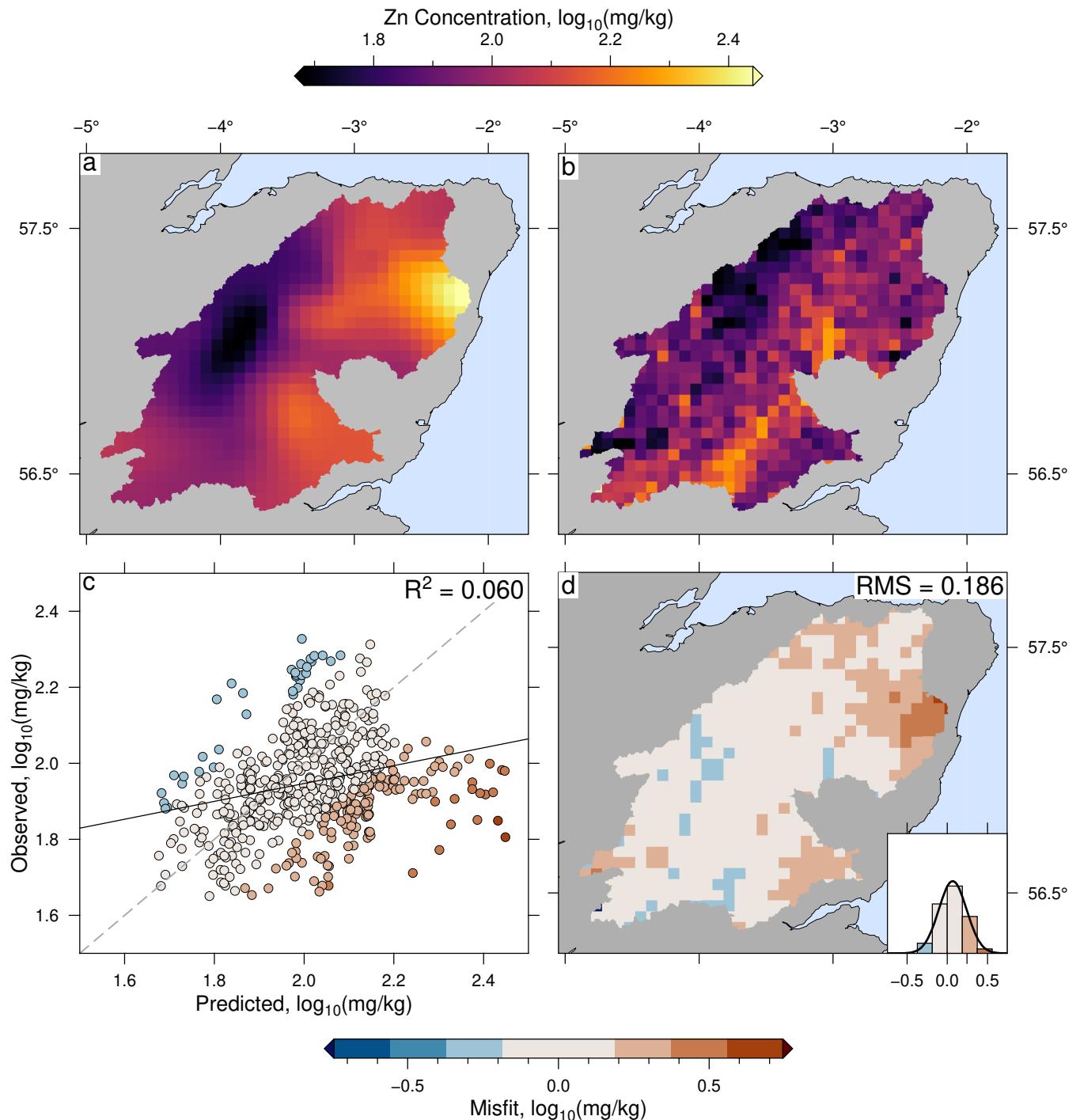


Figure S24. Same as Figure S7 for Zn, $\lambda = 10^{0.6}$

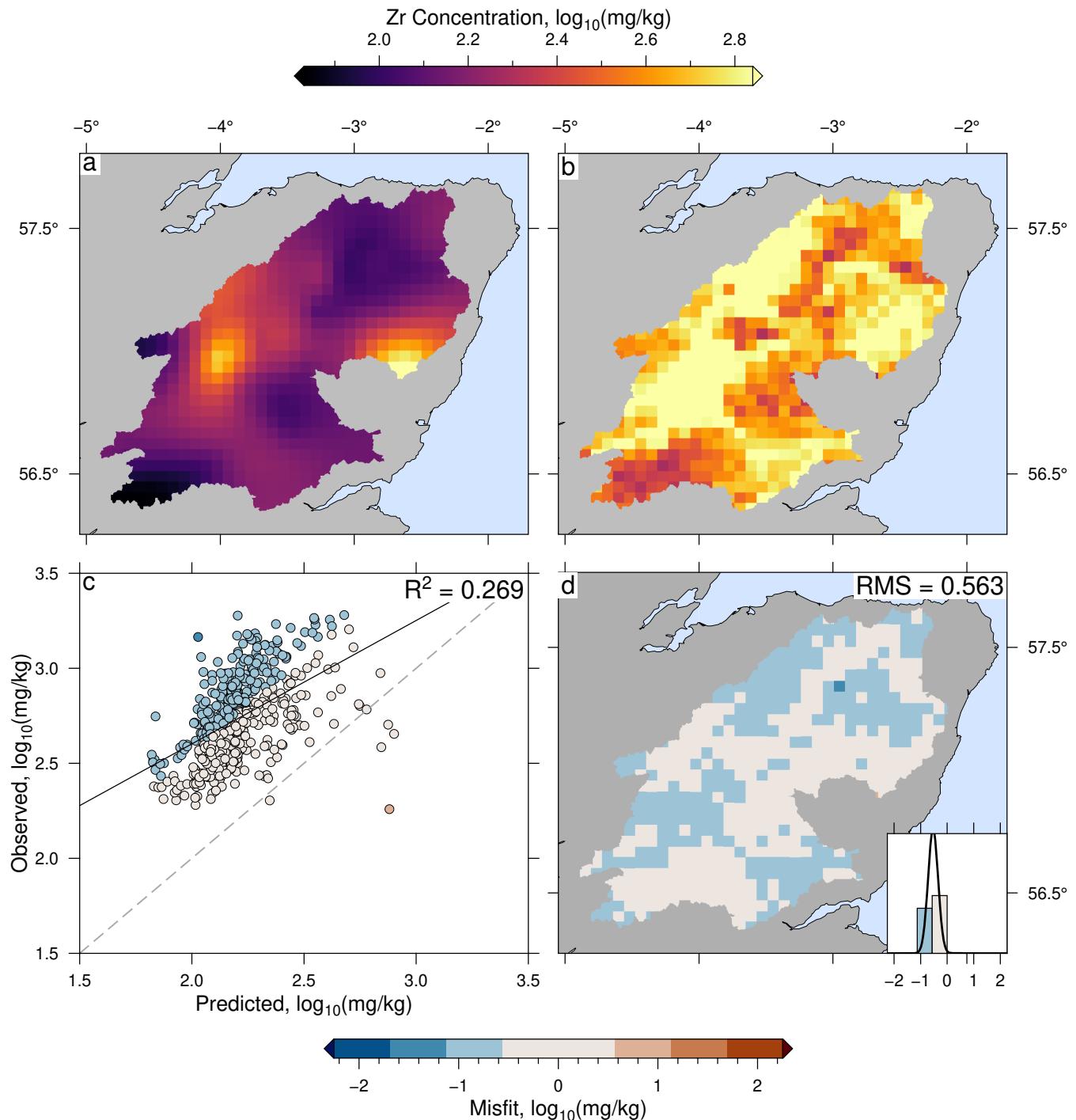


Figure S25. Same as Figure S7 for Zr, $\lambda = 10^{0.7}$