

Supplementary Information of

Global thermal spring distribution and relationship to endogenous and exogenous factors

Tamburello G.¹, Chiodini G.¹, Ciotoli G.^{2,3}, Procesi M.³, Rouwet D.¹, Sandri L.¹, Carbonara N.⁴, Masciantonio C.⁴

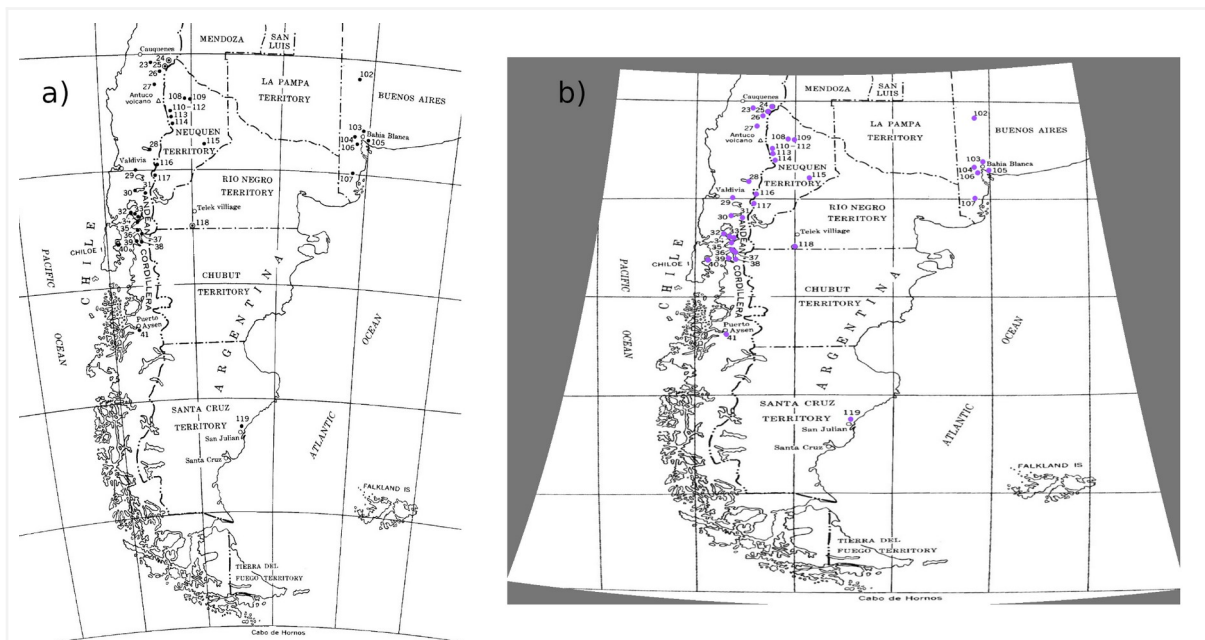
¹ Istituto Nazionale di Geofisica e Vulcanologia, sezione di Bologna, Bologna, Italy

² Consiglio Nazionale delle Ricerche, Istituto di Geologia Ambientale e Geoingegneria, Rome, Italy

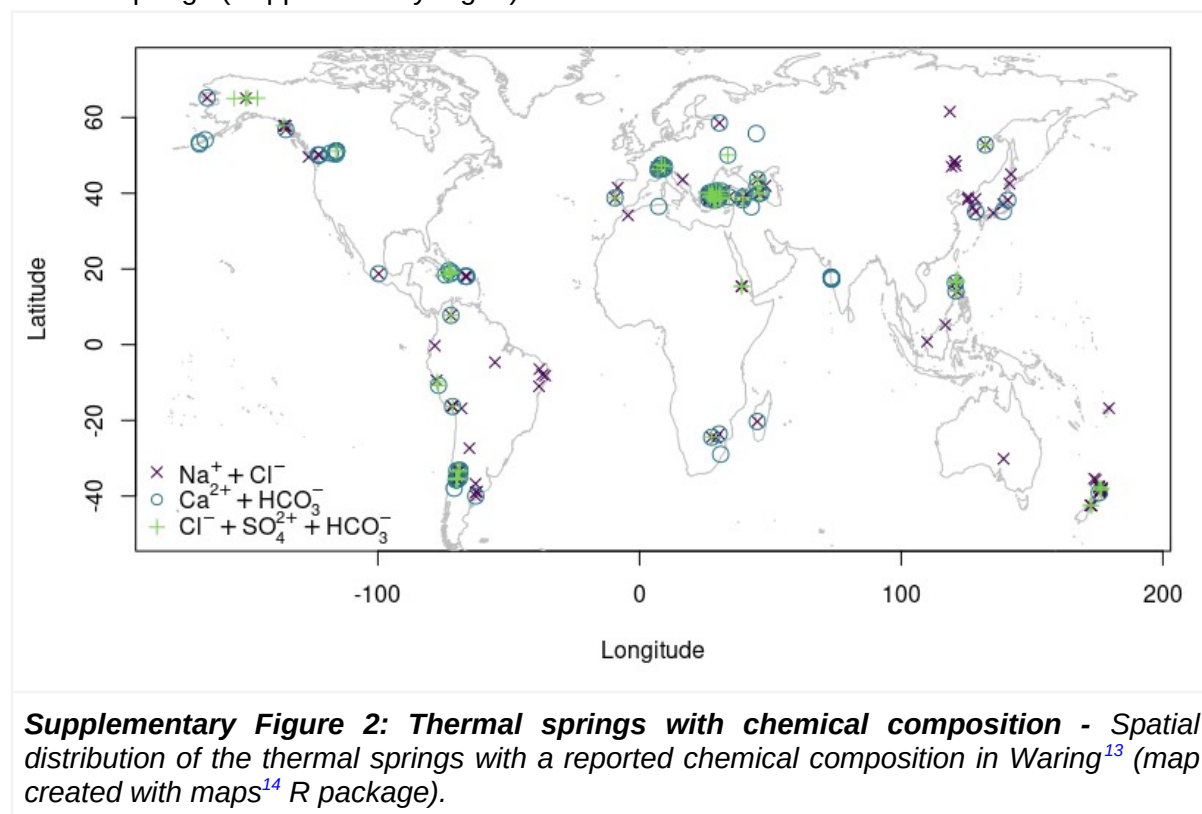
³ Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

⁴ Università degli Studi di Bologna, Bologna, Italy

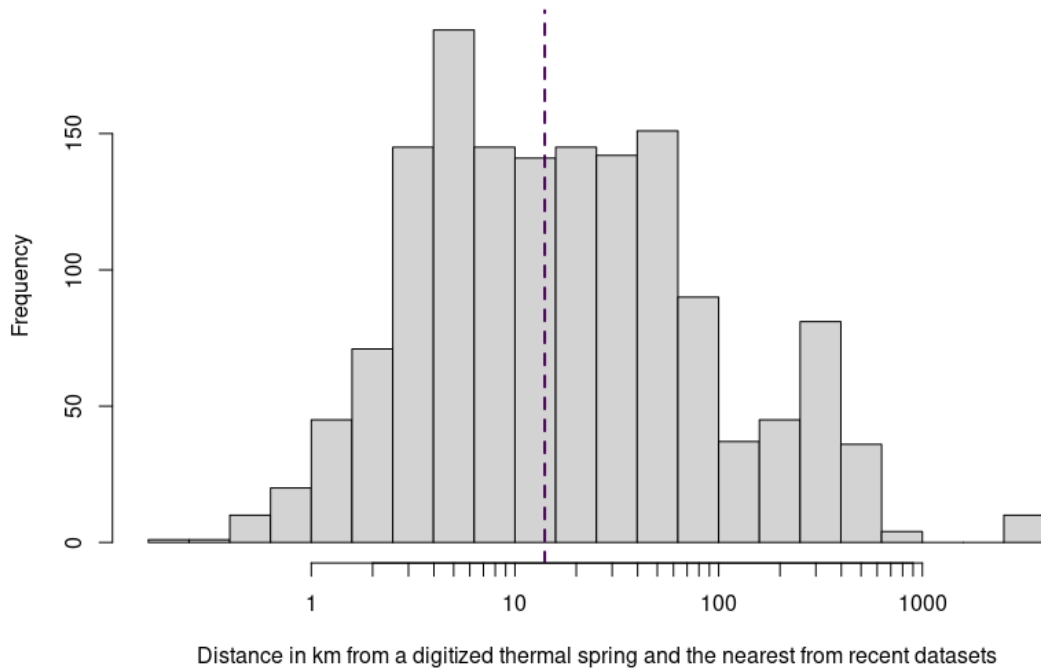
Example of georeferencing of figure 16 from Waring¹³ of Argentina (Supplementary Fig. 1). The lower latitudes generate a distortion that is corrected by fitting the X_P - Y_P -longitude and X_P - Y_P -latitude points (where X_P and Y_P are the pixel coordinates of known geographical features) with a Thin Plate Spline (TPS) algorithm. After georeferencing, the locations of the thermal springs are manually picked to obtain longitude and latitude (purple circles).



The chemical composition of major cations and anions is reported in Waring¹³ for very few thermal springs (Supplementary Fig. 2).

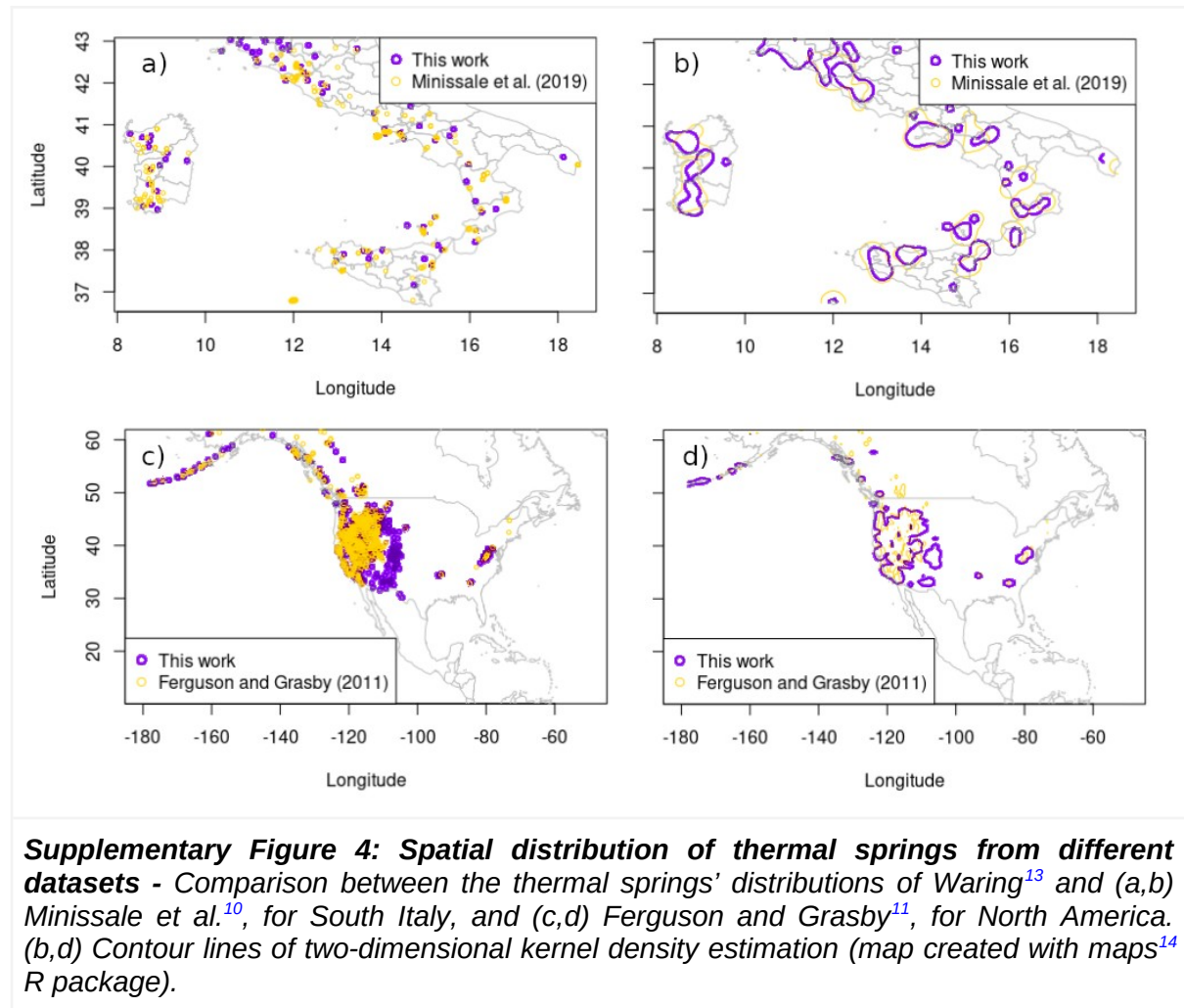


Frequency histogram of calculated distances from each digitized thermal spring and the nearest obtained from Minissale et al.¹⁰ and Ferguson and Graspy¹¹ (Supplementary Fig. 3). The vertical purple dashed line is the median value of 14 km. The high values correspond to thermal springs reported in Waring¹³ but are missing in the more recent dataset (e.g., the springs in Colorado, New Mexico and Arizona that are missing in Ferguson and Graspy¹¹).

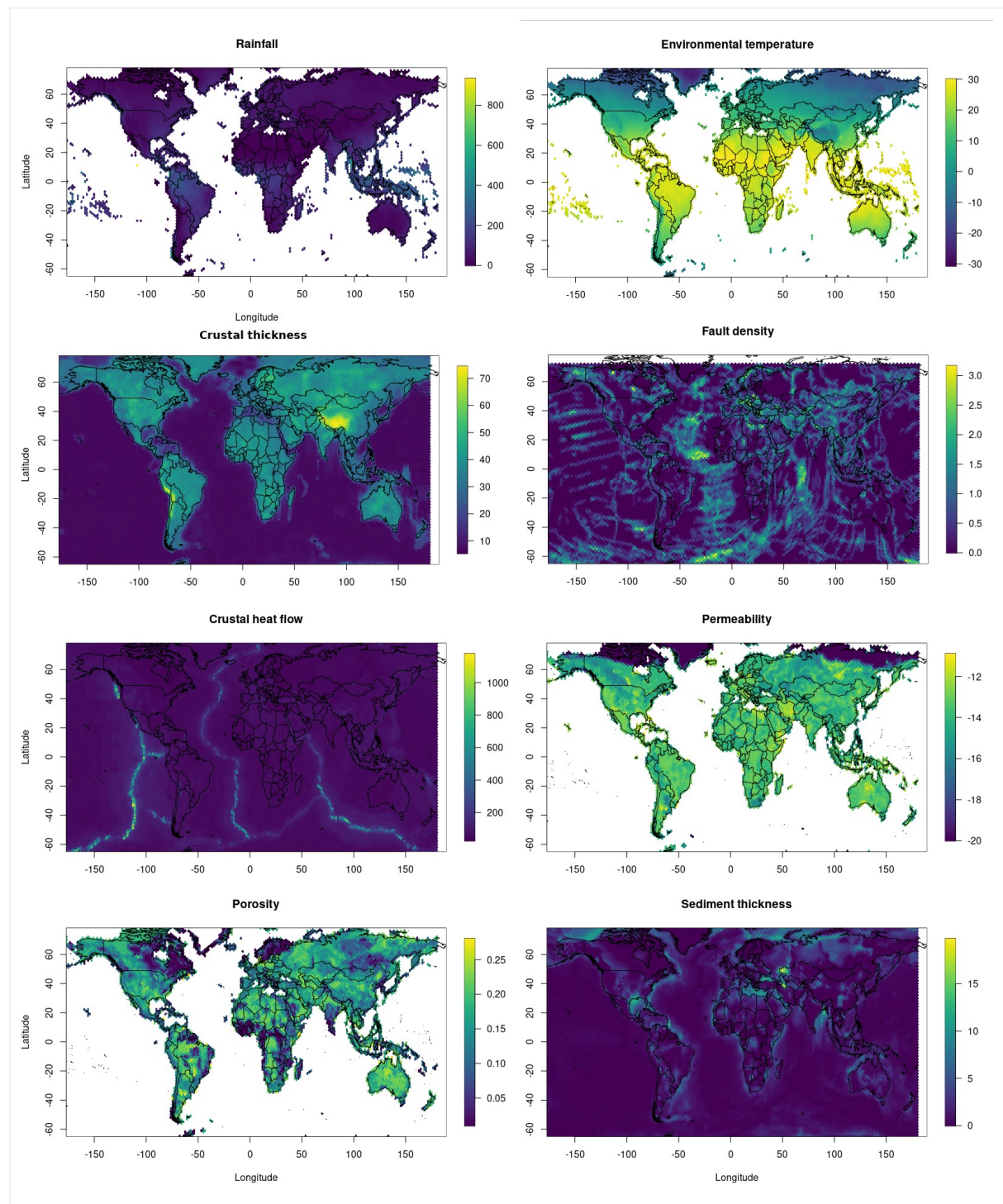


Supplementary Figure 3: Distances between thermal springs from different datasets
- Frequency histogram of distances between each digitized thermal spring from Waring¹³ and the nearest thermal spring from Minissale et al.¹⁰ and Ferguson and Graspy¹¹.

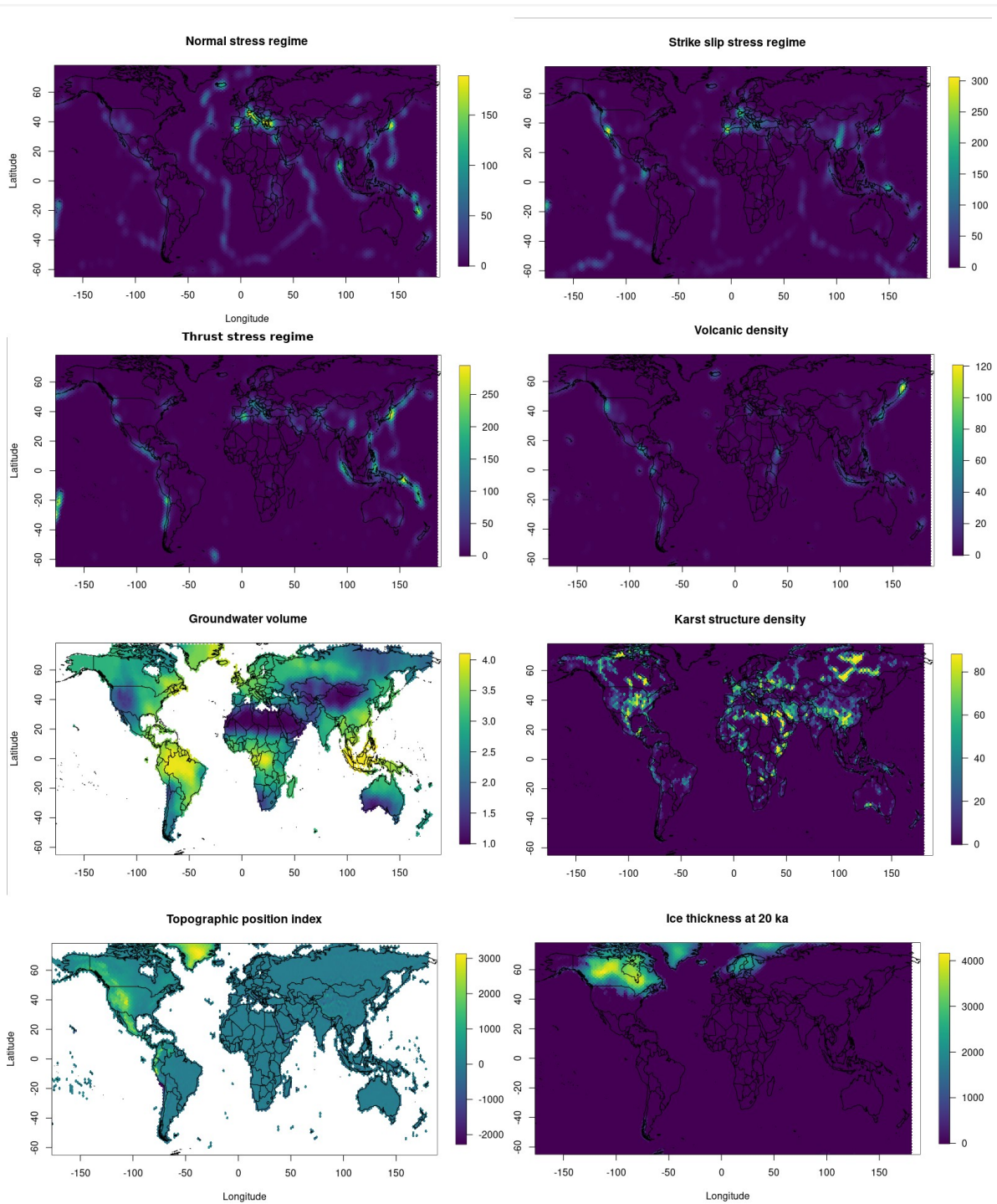
Comparison between the position of the digitized thermal springs from Waring¹³ and a more recent dataset for Italy¹⁰ and North America¹¹. We also compared the main clusters of thermal springs by calculating the kernel density



Maps showing the dataset used for the Random Forest analysis. The values per cell are reported in Supplementary data 2.

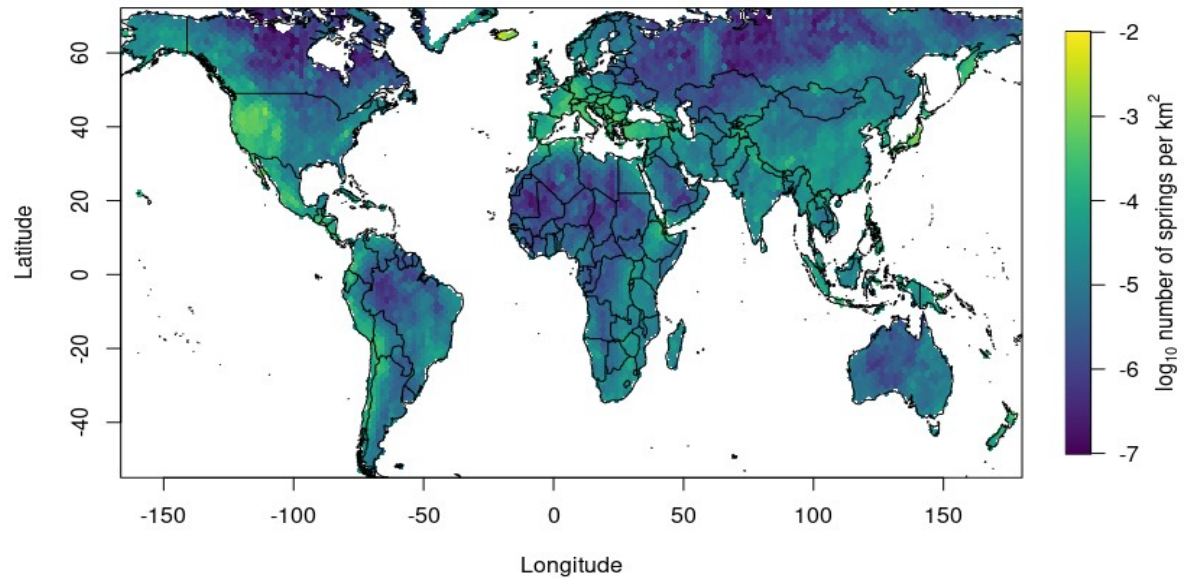


Supplementary Figure 5: Geological factors for machine learning analysis - Maps of the input parameters used for the Random Forest Analysis (map created with maps¹⁴ R package).



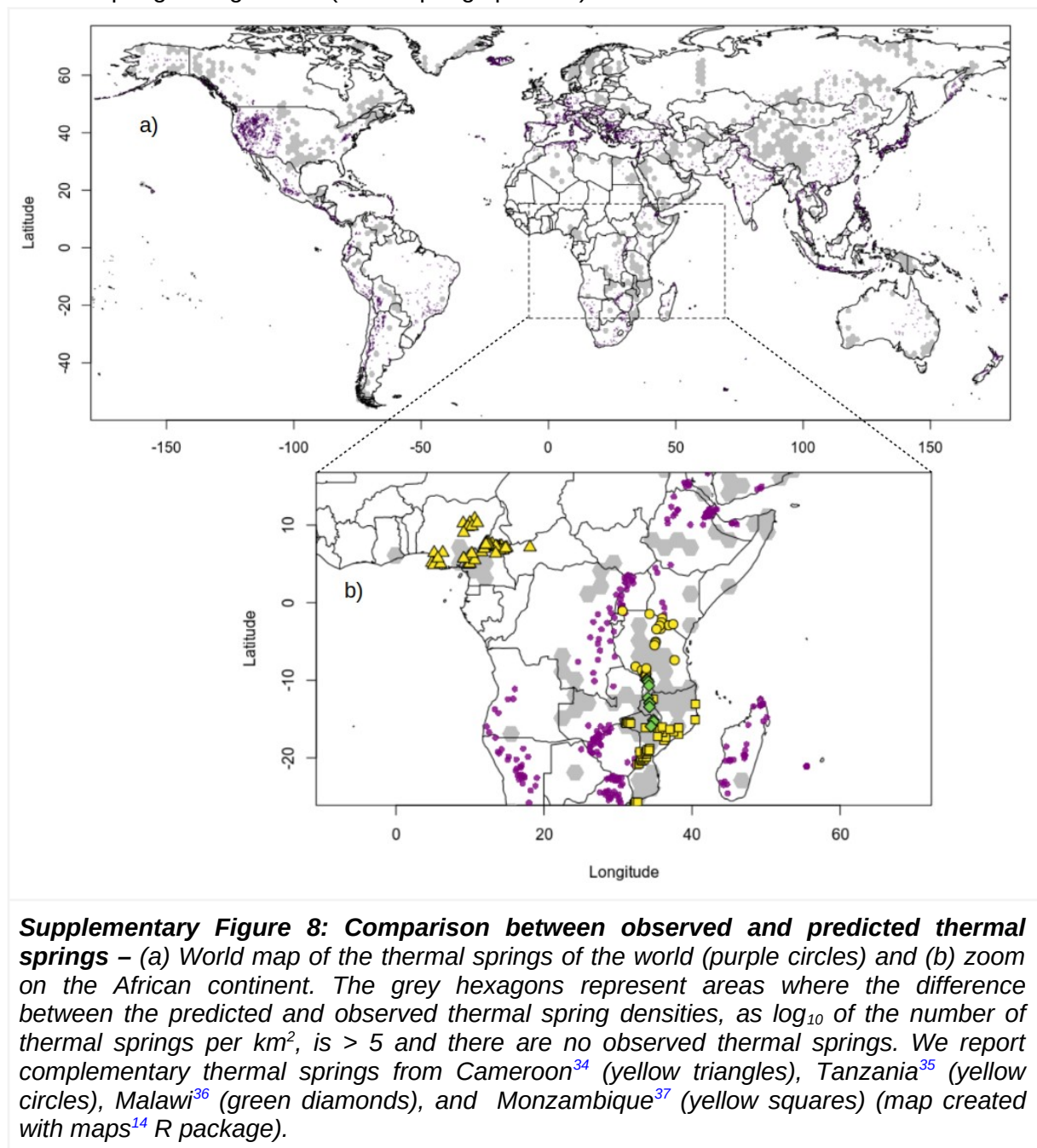
Supplementary Figure 6: Geological factors for machine learning analysis - Maps of the input parameters used for the Random Forest Analysis (map created with maps¹⁴ R package).

In Supplementary Fig. 7 we show the predicted number of thermal spring areas by the random forest algorithm. The map shows a lower contrast than the prediction of the number of thermal springs (Fig. 7).



Supplementary Figure 7: Predicted number of thermal spring areas - Random forest prediction of the distribution of the number of thermal spring areas across the globe. The value represents the averaged 500 random forest results divided by each hexagon's area in km^2 (map created with maps¹⁴ R package).

In the plot in Supplementary Fig. 8, we show the hexagonal areas (in grey) where there are no observed springs and the difference between the predicted and the observed number of thermal springs is significant ($> 10^{-5}$ springs per km^2).



In the subplot in Supplementary Fig. 8, we zoom on the Eastern African Rift to show how the Random Forest algorithm predicted a significant number of missing thermal springs. We searched for literature listing such missing springs and found them in Cameroon³⁴, Tanzania³⁵, Malawi³⁶, and Mozambique³⁷. This example shows the advantages of this dataset and our analysis for future database preparations. For completeness, we list in Supplementary Data 4 the coordinates of the thermal springs added from the above literature. We suggest including these thermal springs to the dataset available in Supplementary Data 1 for future studies on thermal springs' global distribution.