# Supplementary Material for "A geothermal heat flow model for Africa based on Random Forest Regression"
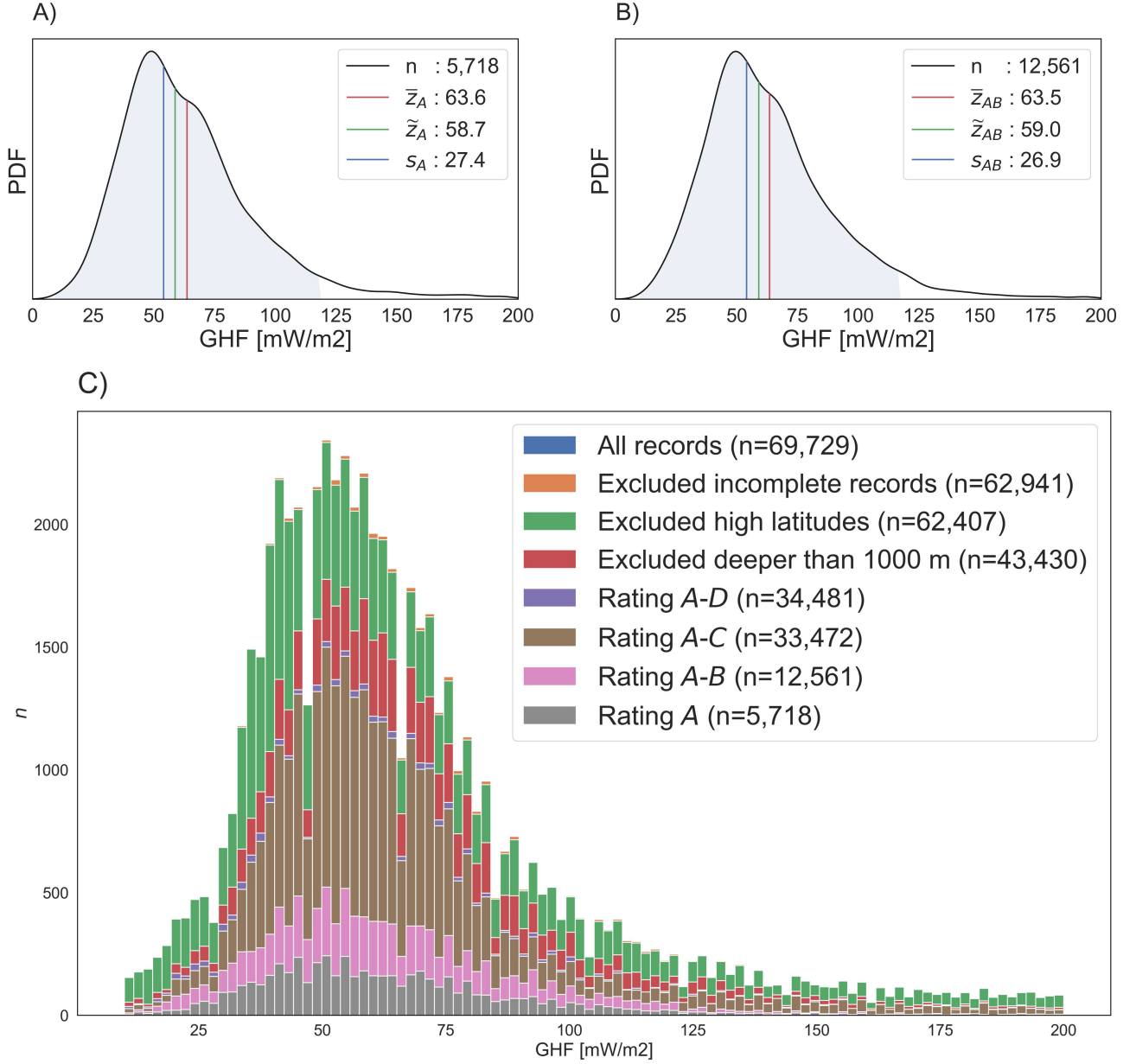
## 1 SUPPLEMENTARY FIGURES

Figure S1. (A) Density plot of global GHF measurements labeled $A$ without questionable values, (B) Density plot of global GHF measurements labeled $A$ and $B$ without questionable values, (C) Histogram of global binned GHF measurements with all records, records after removal of questionable and incomplete information, records after removal of high latitude and of deep-sea information, and records labeled with different quality in the NGHF database (Lucazeau, 2019). $\bar{z}$ = mean, $\tilde{z}$ = median, s = standard deviation.
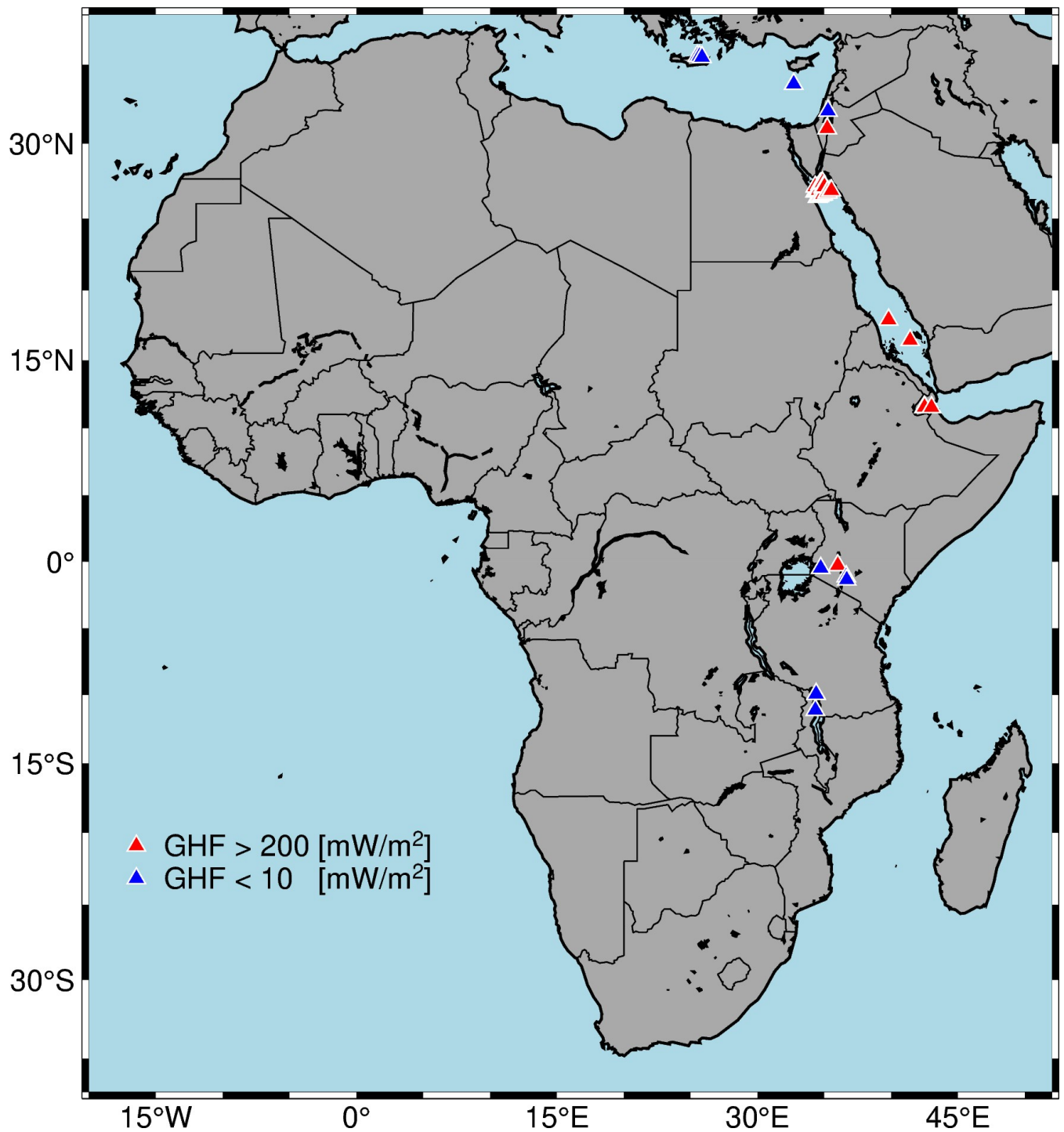
Figure S2. Red and blue triangles locate questionable GHF measurements ($> 200$ mWm$^{-2}$) and ($< 10$ mWm$^{-2}$) respectively (Lucazeau, 2019).
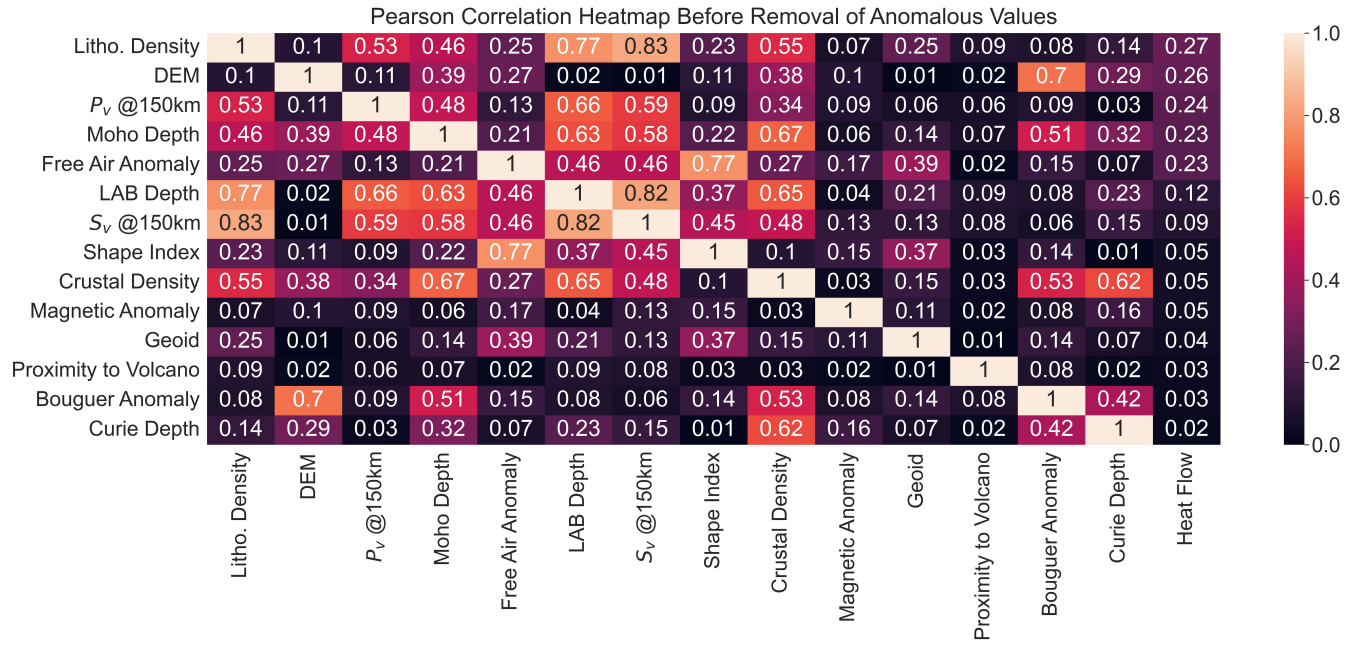
Figure S3. Pearson correlation matrix between observables and GHF reference data, before removing anomalous values; ranked by decreasing correlation with GHF.
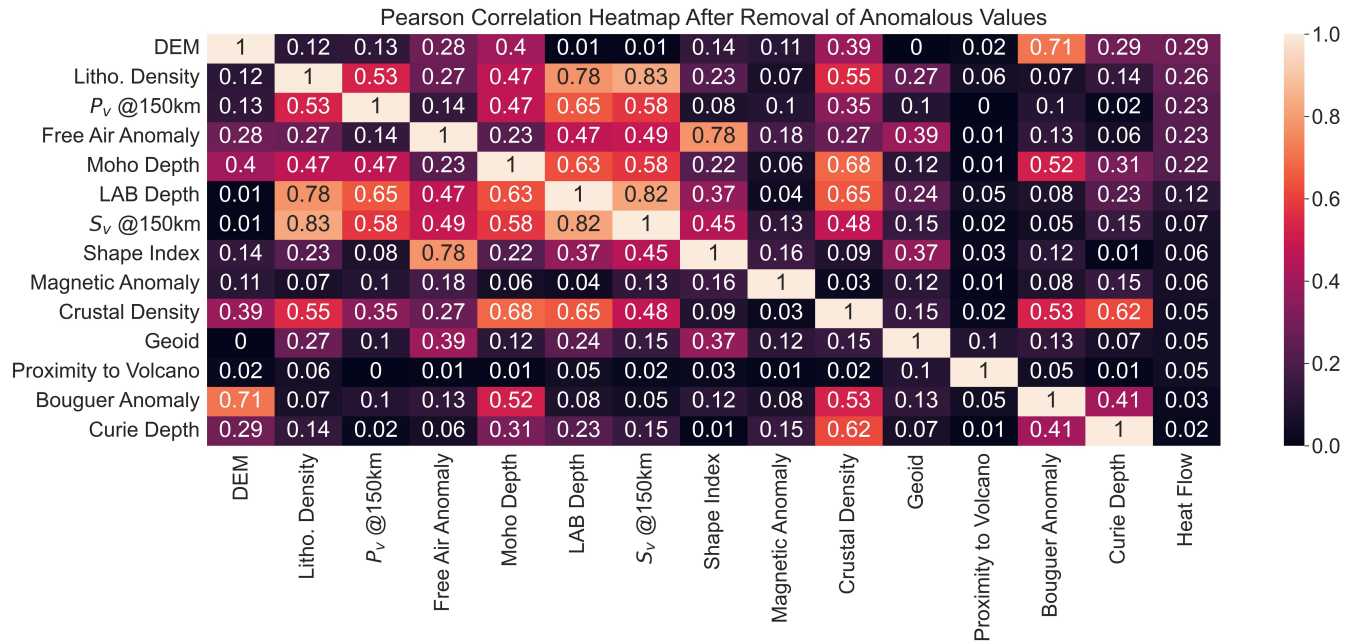


Figure S4. Pearson correlation matrix between observables and GHF reference data, after removing anomalous values; ranked by decreasing correlation with GHF.
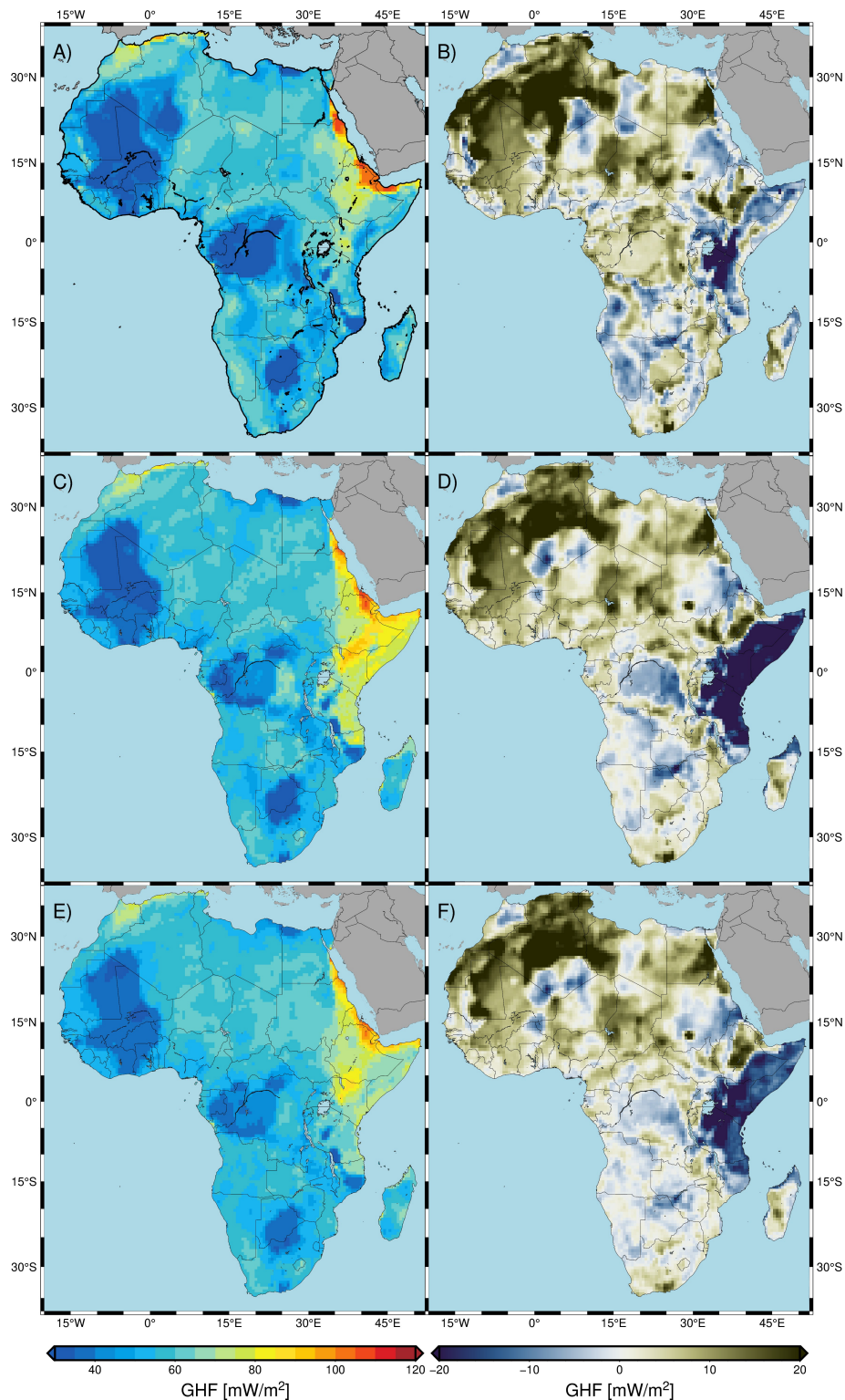
Figure S5. AFQ models trained with reference GHF data labeled *A* and different numbers of observables: (A) AFQ model trained with four observables, (B) Residual between optimal AFQ model trained with eleven observables and AFQ trained with four observables, (C) AFQ model trained with nine observables, (D) Residual between optimal AFQ model trained with eleven observables and AFQ trained with nine observables, (E) AFQ model trained with all sixteen observables, (F) Residual between optimal AFQ model trained with eleven observables and AFQ trained with all sixteen observables.
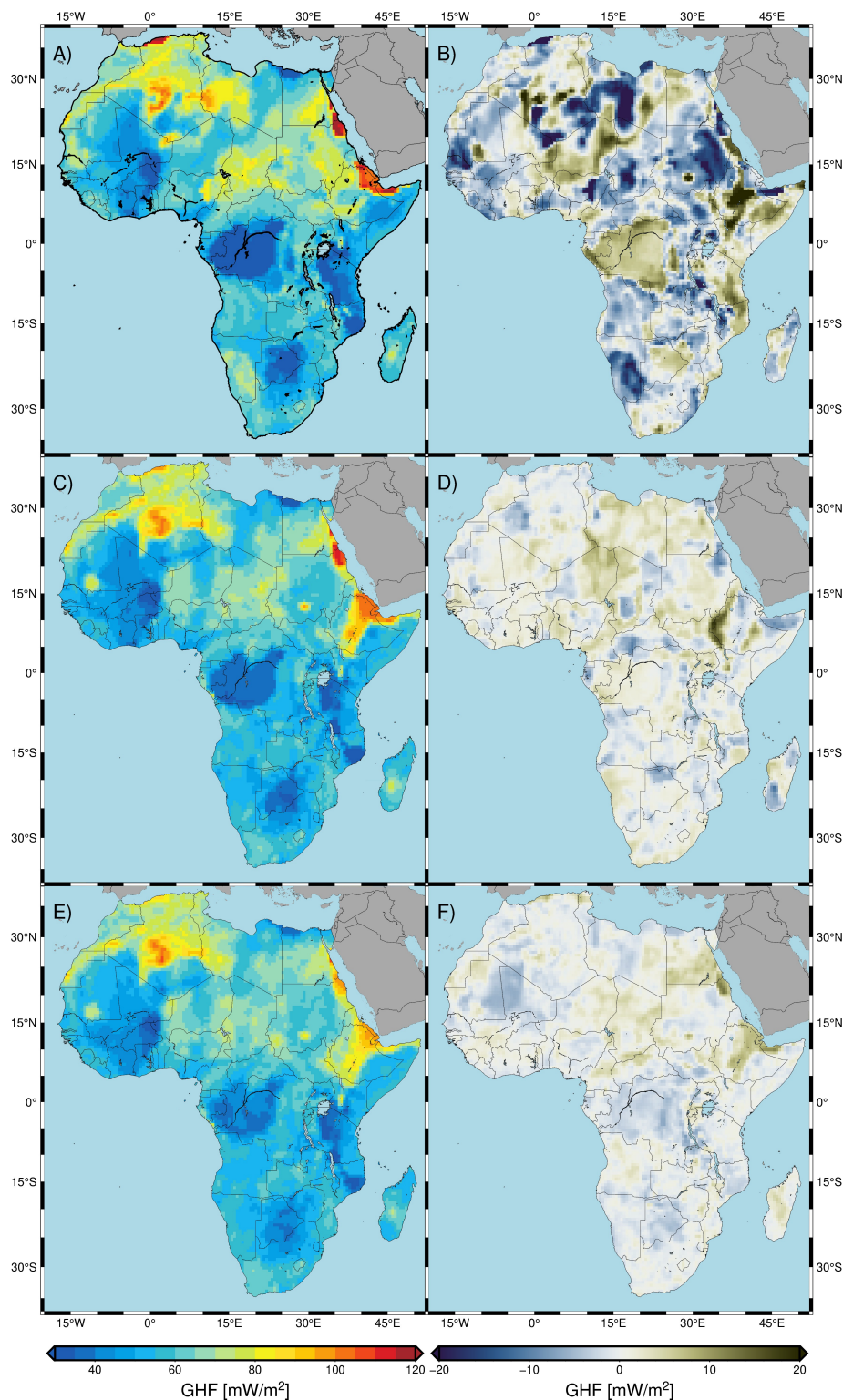
Figure S6. AFQ models trained with reference GHF data labeled *A* and *B* and different numbers of observables: (A) AFQ model trained with four observables, (B) Residual between optimal AFQ model trained with eleven observables and AFQ trained with four observables, (C) AFQ model trained with nine observables, (D) Residual between optimal AFQ model trained with eleven observables and AFQ trained with nine observables, (E) AFQ model trained with all sixteen observables, (F) Residual between optimal AFQ model trained with eleven observables and AFQ trained with all sixteen observables.
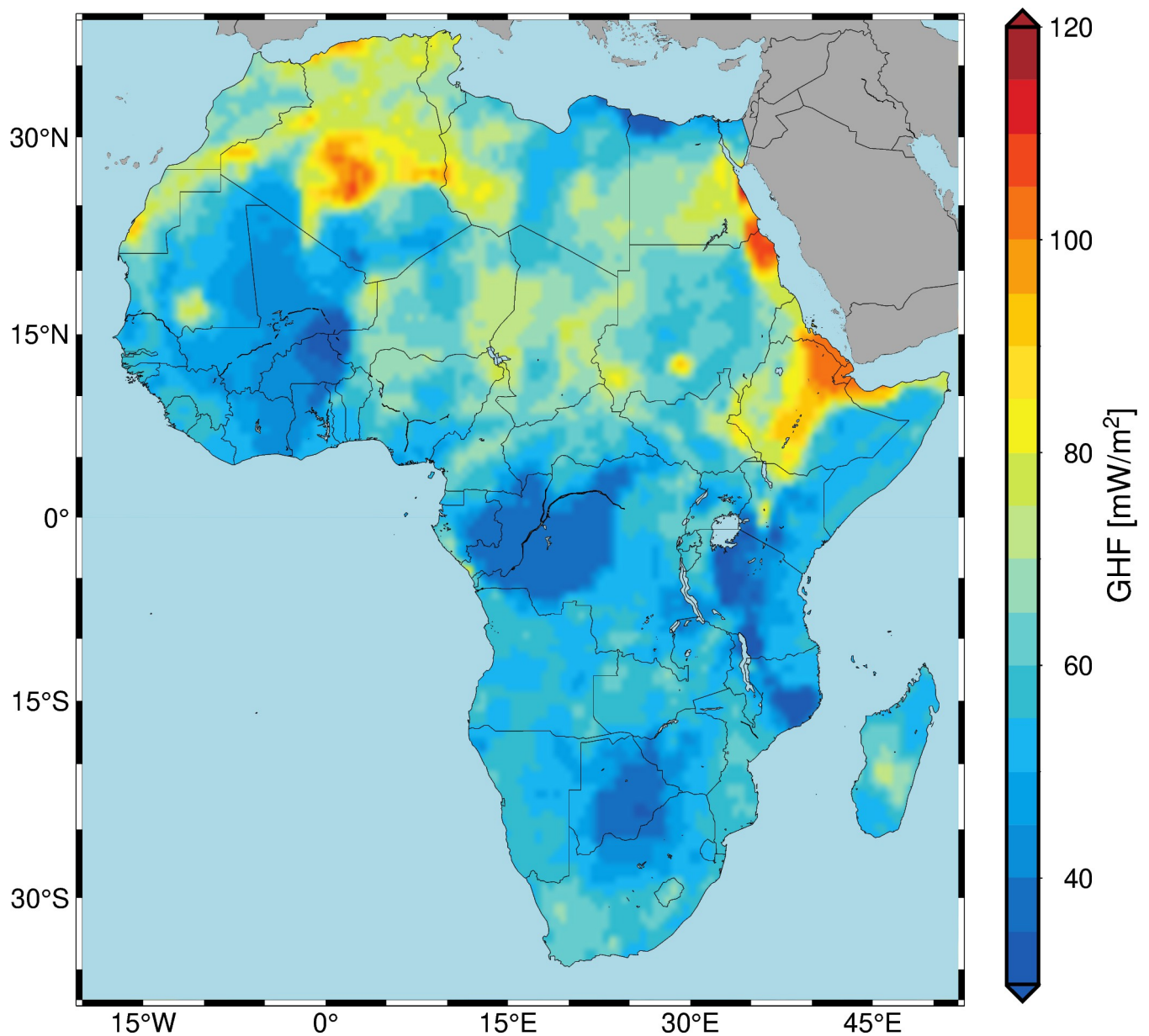
Figure S7. Optimal AFQ model trained with eleven observables. Same as in Figure 6 in the main part of the paper but without any additional overlain information. This model has been used to compute the residuals in Figures S5 and S6.
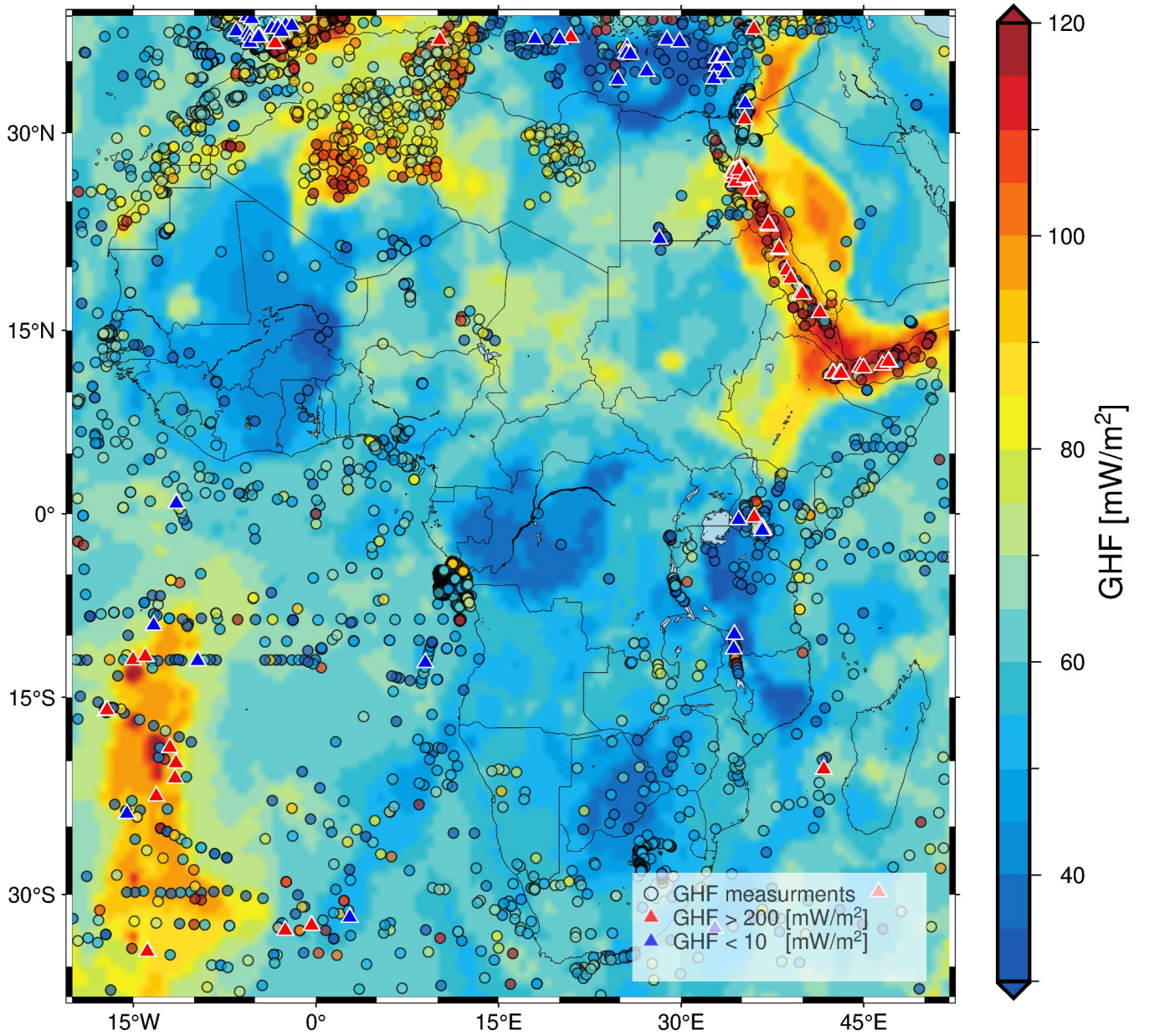
Figure S8. Same GHF model as in Figure S7 (i.e., trained with eleven observables and GHF reference data over continental Africa), but now also indicating GHF predictions over oceanic areas and the Arabian peninsula. Circles represent GHF data labeled $A$, $B$, $C$, $D$ and $Z$, as well as deep-sea measurements derived from the global compilation of heat flow databases (Lucazeau, 2019). Red and blue triangles locate questionable GHF measurements ($> 200$ mWm$^{-2}$) and ($< 10$ mWm$^{-2}$) respectively.

## REFERENCES

Lucazeau, F. (2019). Analysis and mapping of an updated terrestrial heat flow data set. Geochemistry, Geophysics, Geosystems 20, 4001–4024. doi:https://doi.org/10.1029/2019GC008389