

Supplement

Table S1. Environmental covariates available.

Environmental Covariate	Grid name
Mean annual temperature ($^{\circ}\text{C} * 10$, 1979–2013)	BIO1
Mean diurnal range ($^{\circ}\text{C} * 10$, 1979–2013)	BIO2
Mean annual precipitation (mm year^{-1} , 1979–2013)	BIO12
Mean rainfall seasonality (1979–2013)	BIO15
Distance to inland water bodies (km)	DOWS
Average night-time land surface temperature ($^{\circ}\text{C}$, 2001–2017)	LSTN
Average day-time land surface temperature ($^{\circ}\text{C}$, 2001–2017)	LSTD
Average enhanced vegetation index (2000–2016)	EVI
Average MOD13Q1 band 1 reflectance (2000–2016)	MB1
Average MOD13Q1 band 2 reflectance (2000–2016)	MB2
Average MOD13Q1 band 3 reflectance (2000–2016)	MB3
Average MOD13Q1 band 7 reflectance (2000–2016)	MB7
Normalised difference vegetation index (2000–2016)	NDVI
Soil adjusted vegetation index (200–2016)	SAVI
Average annual net primary productivity (kg m^{-2} , 2000–2015)	NPPA
Variance annual net primary productivity (2000–2015)	NPPS
Slope (%)	SLOPE
Topographic index	TIM
Elevation above mean sea level (m)	MDEM

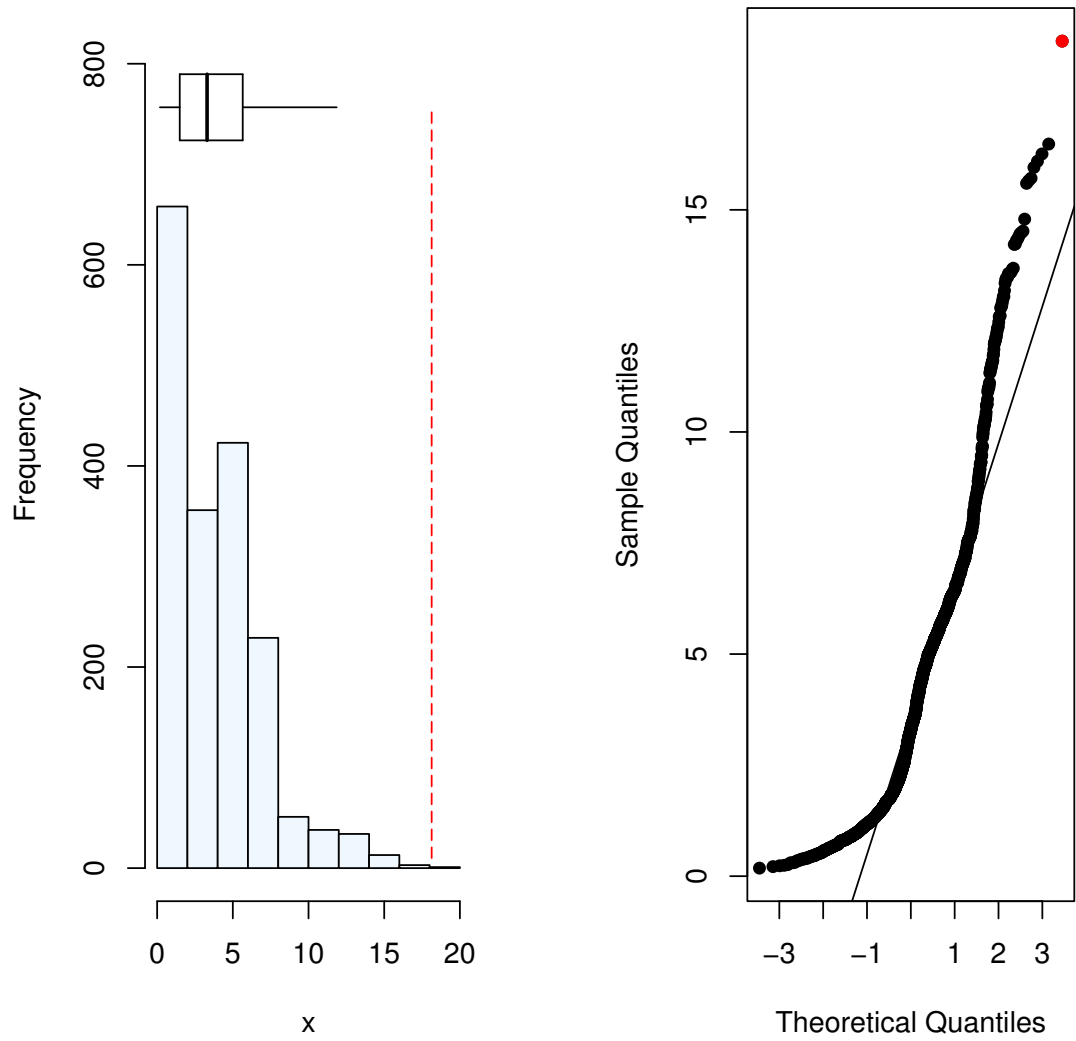


Figure S1. Histogram with boxplot and QQ plot for soil soluble Se (Se_{sol}) concentration in Malawi.

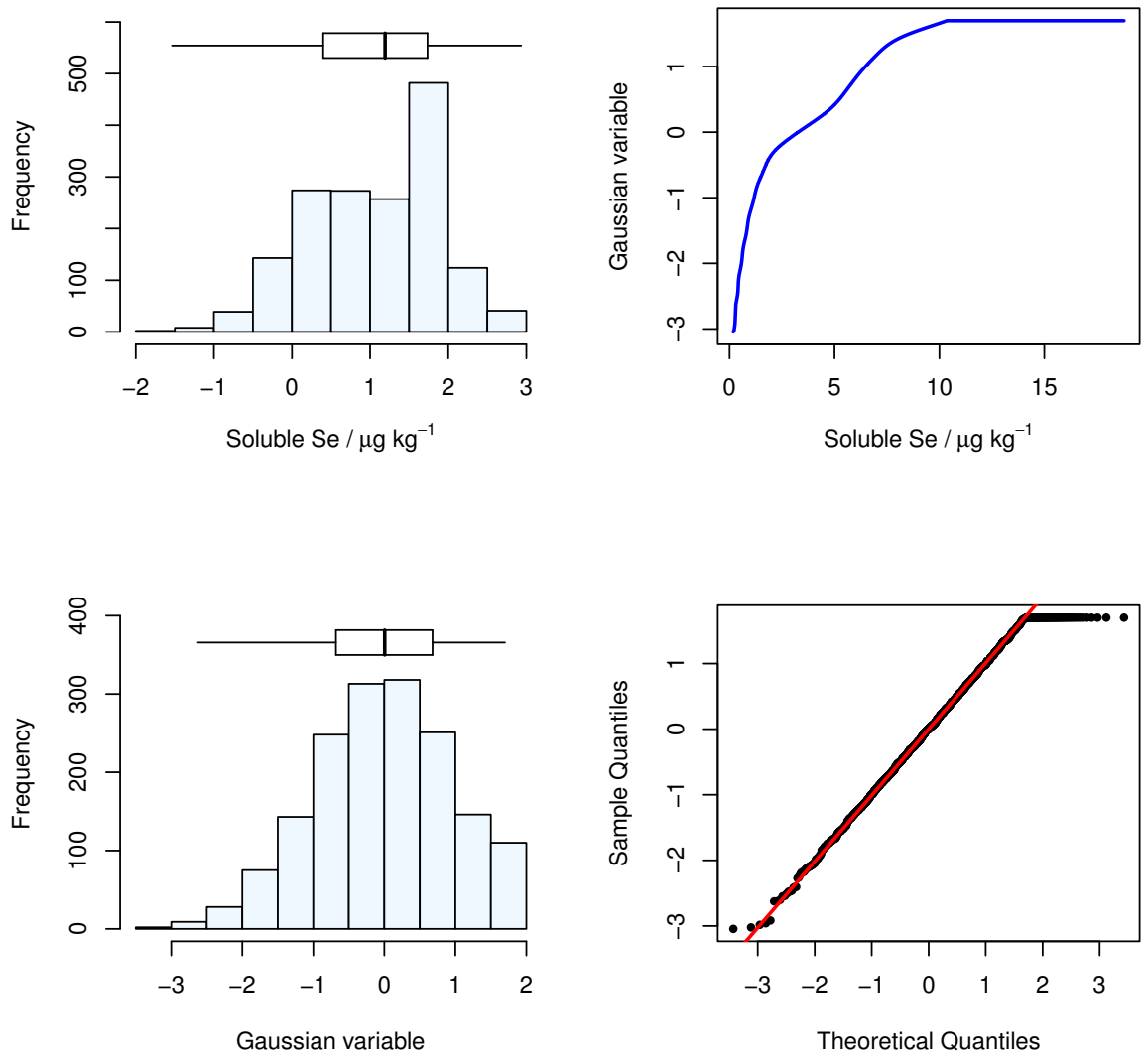


Figure S2. Histogram with boxplot and QQ plot for Se_{sol} transformed by Gaussian anamorphosis.

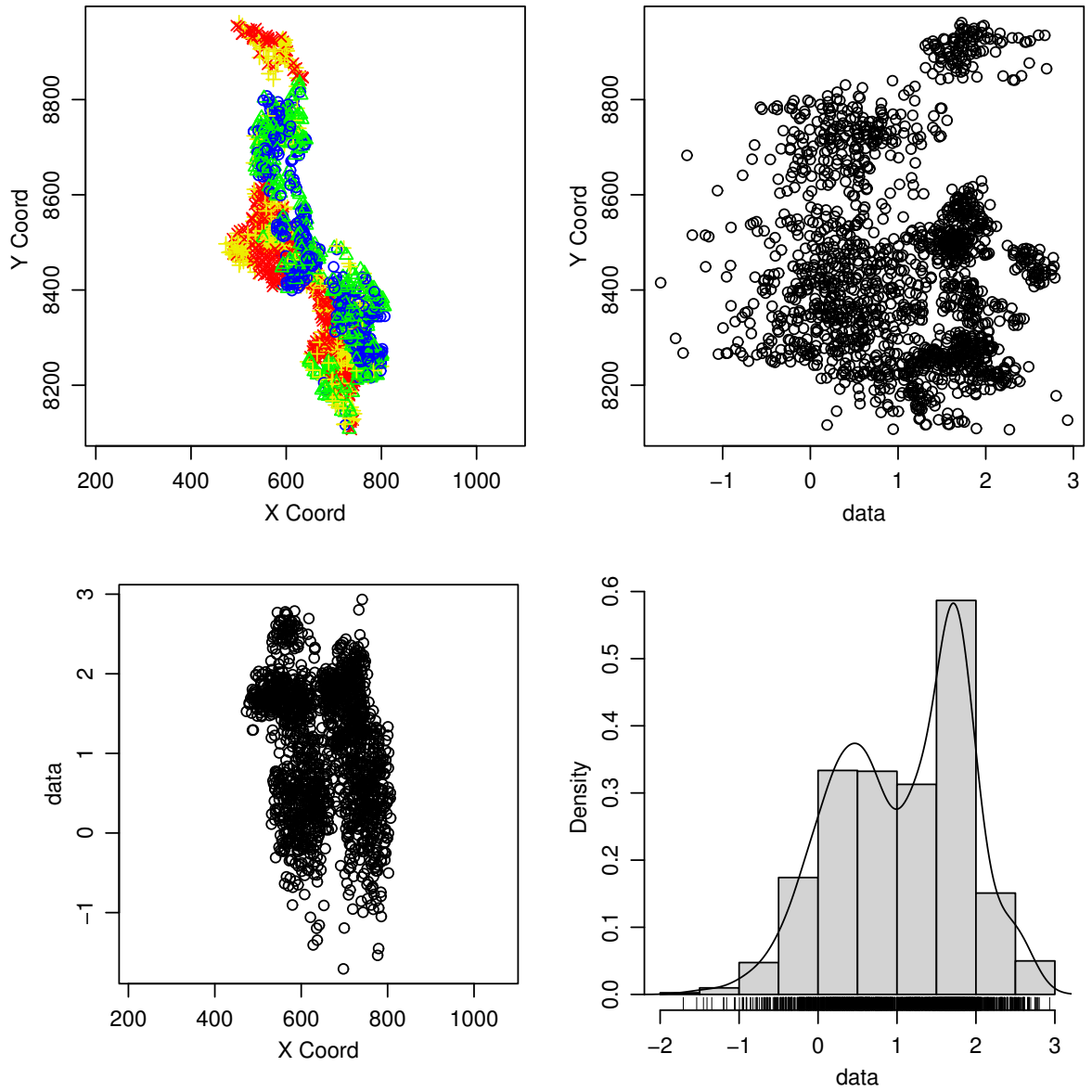


Figure S3. Post plot Se_{sol} transformed by natural logarithms. The blue, green, yellow and red symbols in top left panel correspond to 1st, 2nd, 3rd and 4th quartiles.

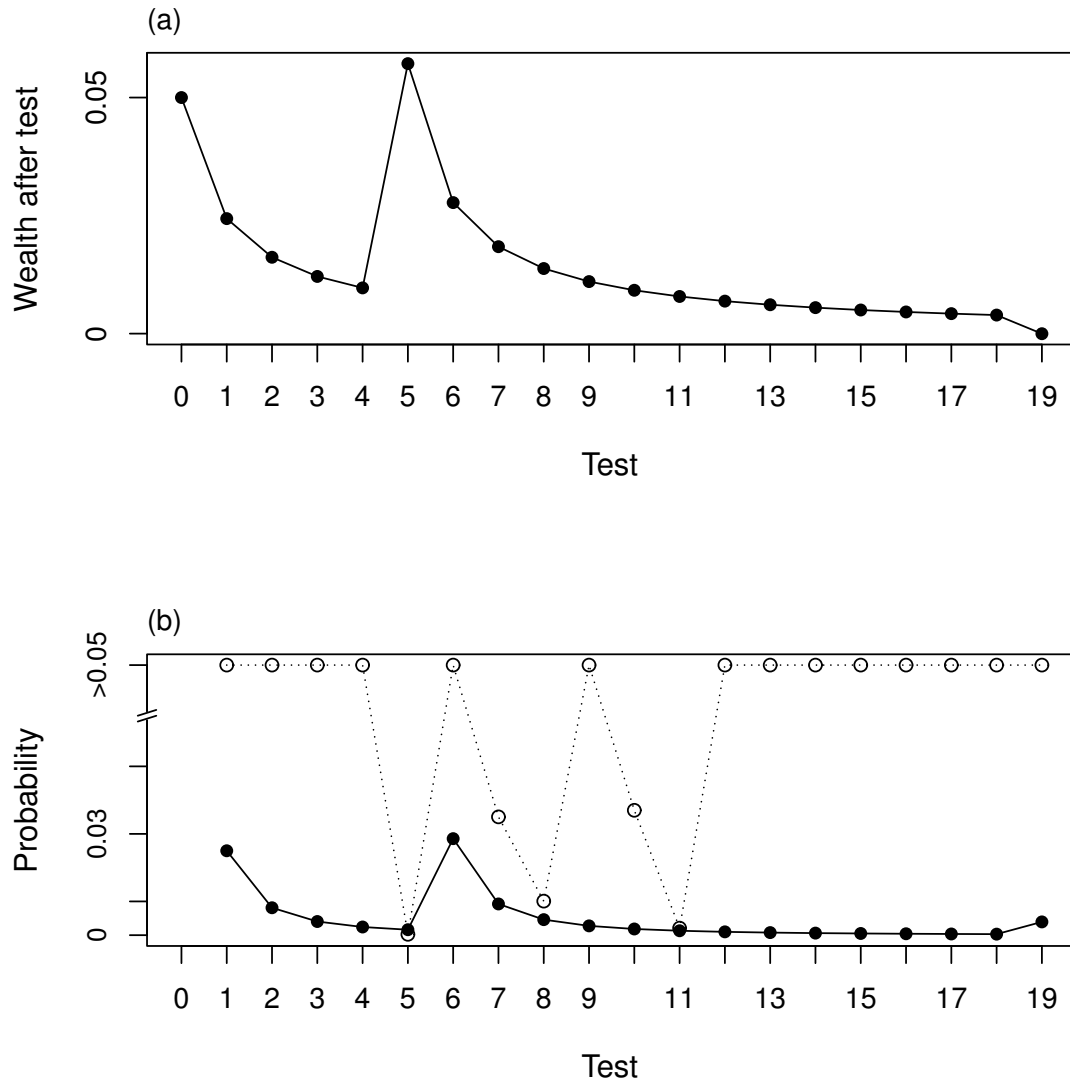


Figure S4. Ordered tests for covariate selection for $S_{e_{sol}}$. This for all covariates in Table 2. The graph (a) shows α -wealth over the sequence of tests and the lower (b) shows the p -values for successive tests (open symbols) and the corresponding threshold values with marginal false discovery rate control.

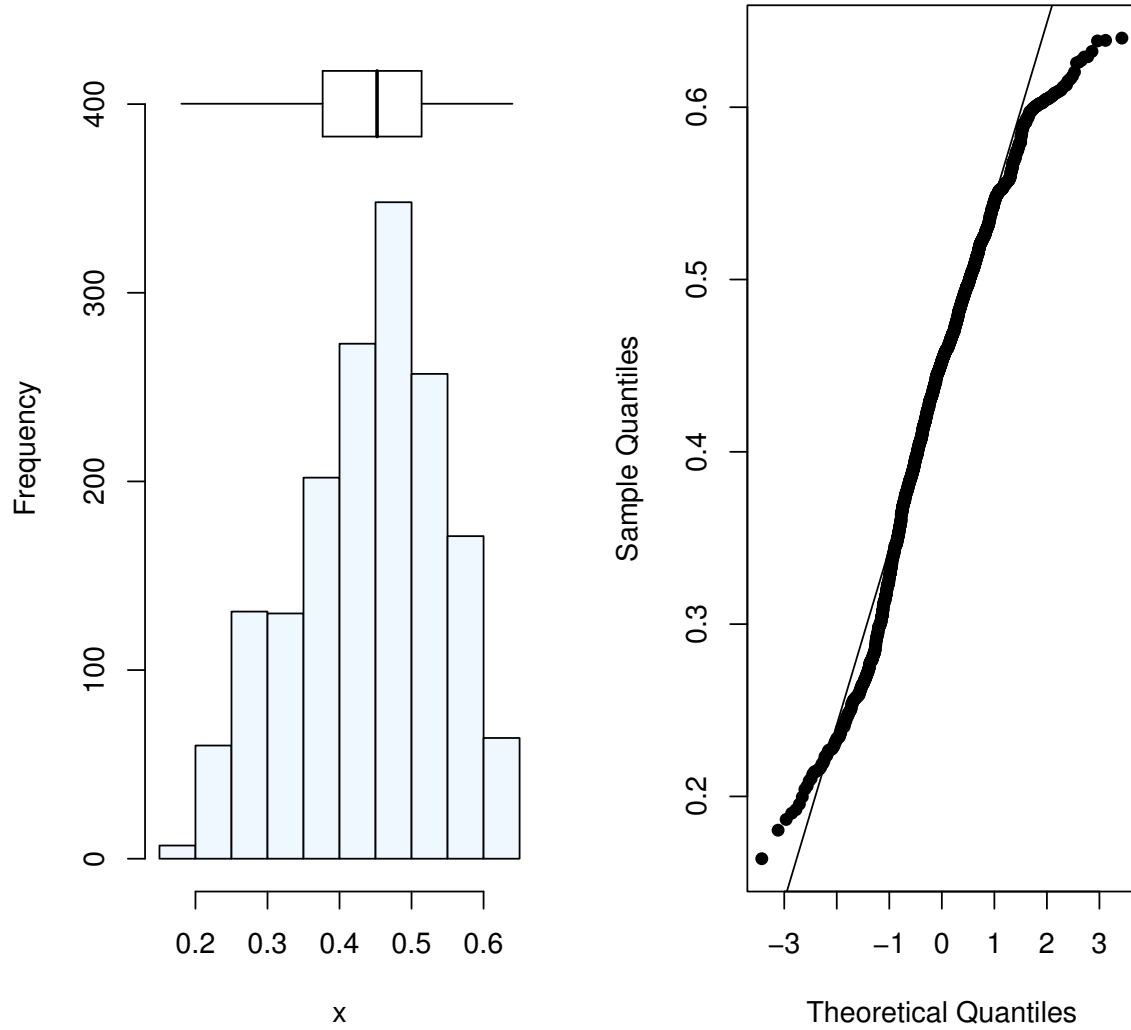


Figure S5. Histogram with boxplot and QQ plot for the prediction error variance from cross validation for the robust REML E-BLUP.

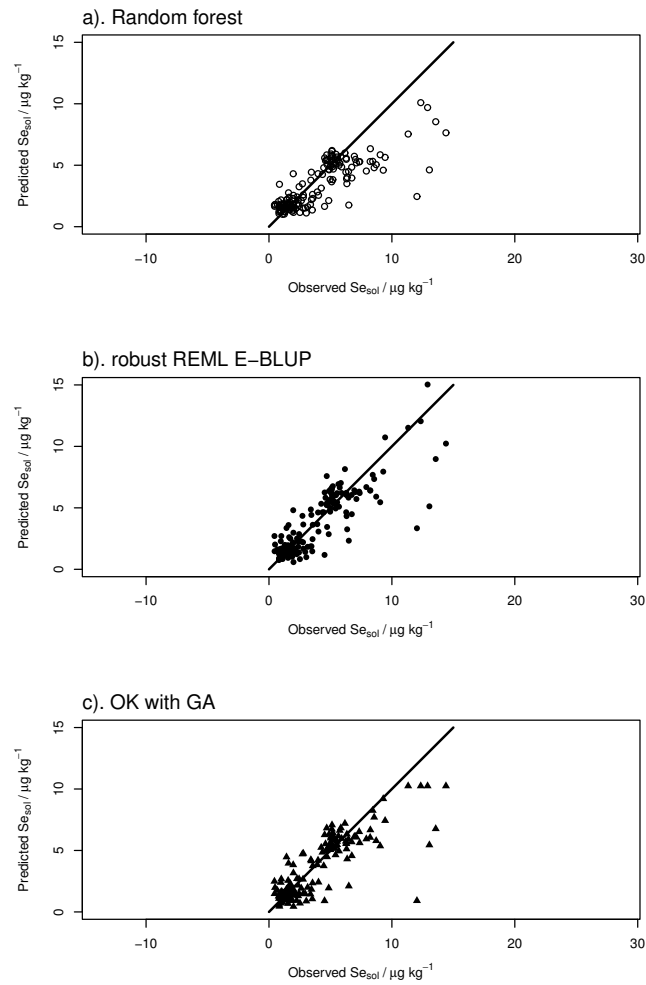


Figure S6. Scatter plot for the predictions of Se_{sol} (a) random forest, (b) robust REML E-BLUP, and (c) ordinary kriging with Gaussian anamorphosis transformed Se_{sol} .