

SUPPLEMENT 1. RAZAVI AND GUPTA ESTIMATOR (VARS)

Unlike the other total-order estimators examined in our paper, the Razavi and Gupta VARS (for variogram analysis of response surfaces [20,39]) relies on the variogram $\gamma(\cdot)$ and covariogram $C(\cdot)$ functions to compute what they call the VARS total-order (VARS-TO) index.

Let us consider a function of factors $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$. If \mathbf{x}_A and \mathbf{x}_B are two generic points separated by a distance h , then the variogram is calculated as follows:

$$\gamma(\mathbf{x}_A - \mathbf{x}_B) = \frac{1}{2} V[y(\mathbf{x}_A) - y(\mathbf{x}_B)] \quad (\text{S.1})$$

and the covariogram as

$$C(\mathbf{x}_A - \mathbf{x}_B) = \text{COV}[y(\mathbf{x}_A), y(\mathbf{x}_B)] \quad (\text{S.2})$$

Note that

$$V[y(\mathbf{x}_A) - y(\mathbf{x}_B)] = V[y(\mathbf{x}_A)] + V[y(\mathbf{x}_B)] - 2\text{COV}[y(\mathbf{x}_A), y(\mathbf{x}_B)] \quad (\text{S.3})$$

and since $V[y(\mathbf{x}_A)] = V[y(\mathbf{x}_B)]$, then

$$\gamma(\mathbf{x}_A - \mathbf{x}_B) = V[y(\mathbf{x})] - C(\mathbf{x}_A, \mathbf{x}_B) \quad (\text{S.4})$$

In order to obtain the total-order effect T_i , the variogram and covariogram are computed on all couples of points spaced h_i along the x_i -axis, with all other factors being kept fixed. Thus, Eq. (S.4) becomes

$$\gamma_{x_{\sim i}}(h_i) = V(y|x_{\sim i}^*) - C_{x_{\sim i}}(h_i) \quad (\text{S.5})$$

where $x_{\sim i}^*$ is a fixed point in the space of non- x_i . [20,39] suggest to take the mean value across the factors' space on both sides of Eq. (S.5), thus obtaining

$$E_{x_{\sim i}}[\gamma_{x_{\sim i}}(h_i)] = E_{x_{\sim i}}[V(y|x_{\sim i}^*)] - E_{x_{\sim i}}[C_{x_{\sim i}}(h_i)] \quad (\text{S.6})$$

which can also be written as follows:

$$E_{x_{\sim i}}[\gamma_{x_{\sim i}}(h_i)] = V(y)T_i - E_{x_{\sim i}}[C_{x_{\sim i}}(h_i)] \quad (\text{S.7})$$

and therefore

$$T_i = \frac{E_{x_{\sim i}}[\gamma_{x_{\sim i}}(h_i)] + E_{x_{\sim i}}[C_{x_{\sim i}}(h_i)]}{V(y)} \quad (\text{S.8})$$

The sampling scheme for VARS does not rely on $\mathbf{A}, \mathbf{B}, \mathbf{A}_B^{(i)}$... matrices, but on star centers and cross sections. Star centers are N random points sampled across the input space. For each of these stars, k cross sections of points spaced Δh apart are generated, including and passing through the star center. Overall, the computational cost of VARS amounts to $N_t = N[k((1/\Delta h) - 1) + 1]$.

SUPPLEMENT 2. FIGURES

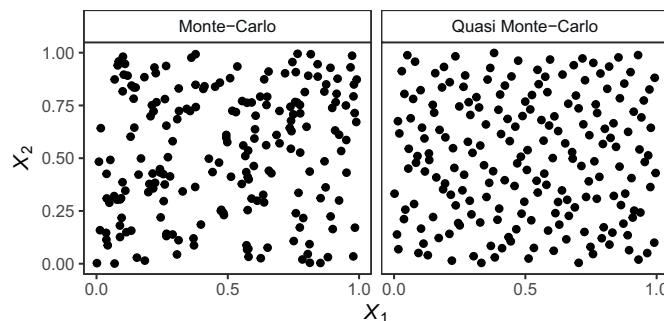


FIG. S1: Examples of Monte Carlo and Quasi Monte Carlo sampling in two dimensions, $N = 2 \times 10^2$

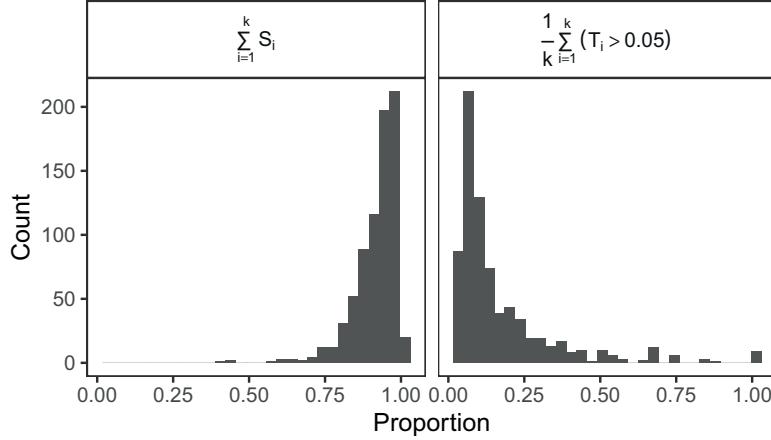


FIG. S2: Proportion of the total sum of first-order effects and of the active model inputs (defined as $T_i > 0.05$) after 10×10^2 random metaprocedure calls with $k \in (3, 100)$. Note how the sum of first-order effects clusters around 0.8 (thus evidencing the production of non-additivities) and how, on average, the number of active model inputs revolves around 10–20%, thus reproducing the Pareto principle.

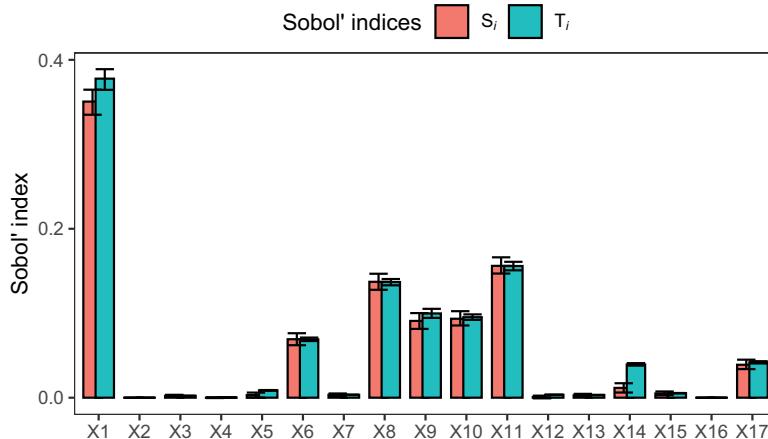


FIG. S3: Sobol' T_i indices obtained after a run of the metaprocedure with the following parameter settings: $N = 10^4$, $k = 17$, $k_2 = 0.5$, $k_3 = 0.2$, $\varepsilon = 666$. The error bars reflect the 95% confidence intervals after bootstrapping ($R = 10^2$). The indices have been computed with the Jansen [15] estimator.

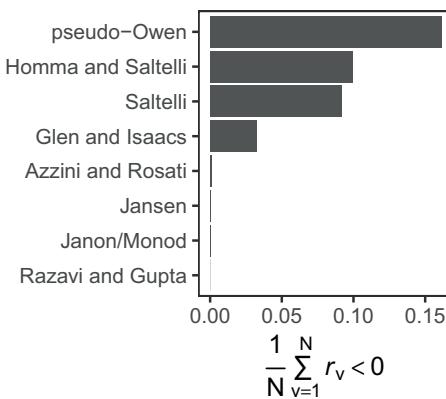


FIG. S4: Proportion of model runs yielding $r < 0$

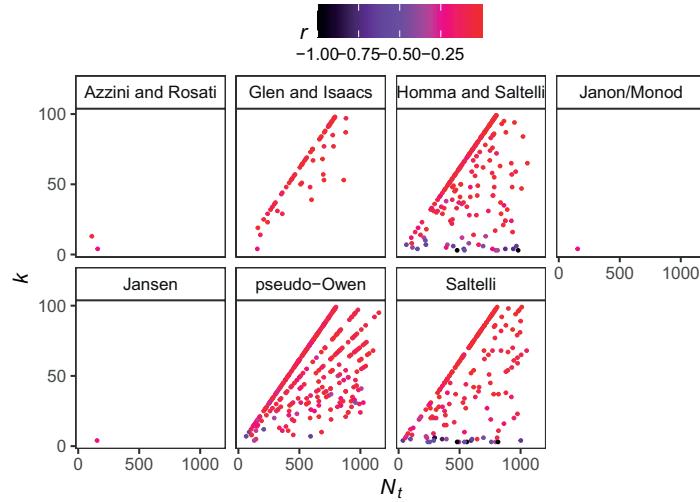


FIG. S5: Scatter of the total number of model runs N_t against the function dimensionality k only for $r < 0$

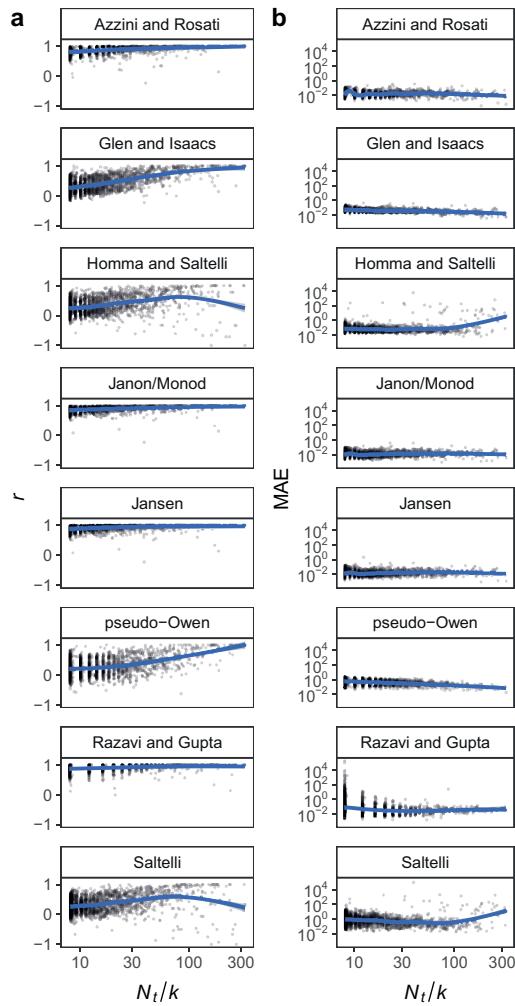


FIG. S6: Scatterplot of the correlation between \mathbf{T}_i and $\hat{\mathbf{T}}_i$ (r) against the number of model runs allocated per model input (N_t/k)

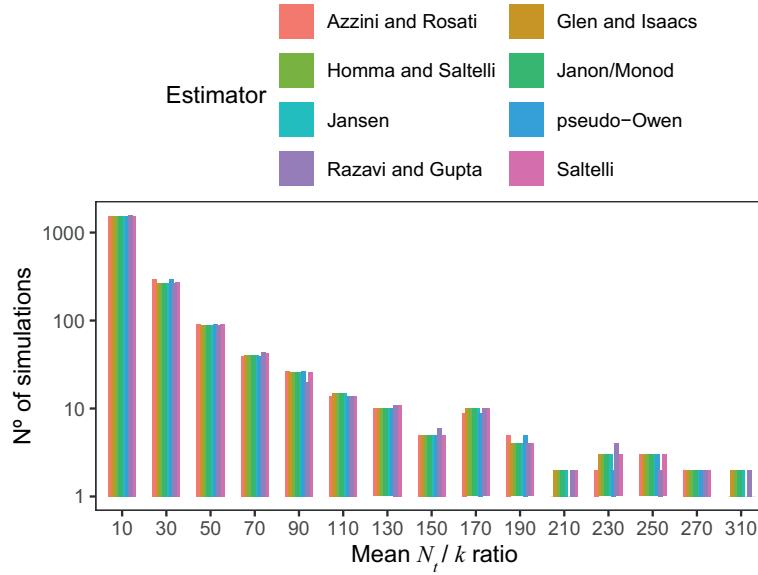


FIG. S7: Bar plot with the number of simulations conducted in each of the N_t/k compartments assessed. All estimators have approximately the same number of simulations in each compartment.

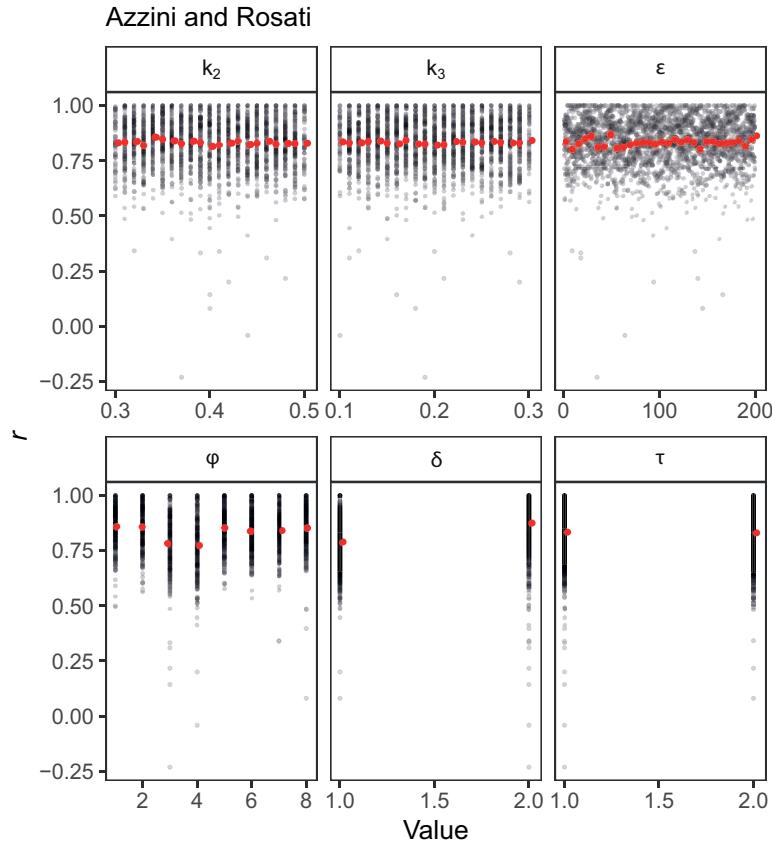


FIG. S8: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

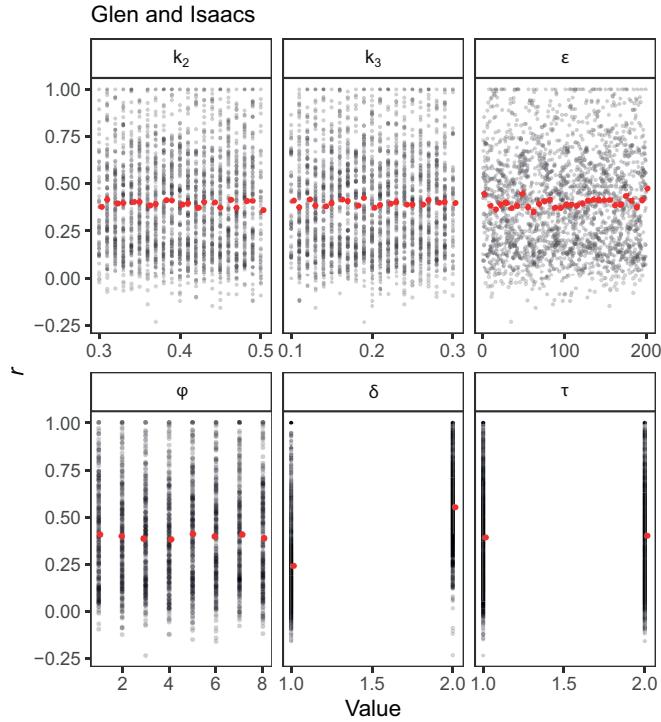


FIG. S9: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

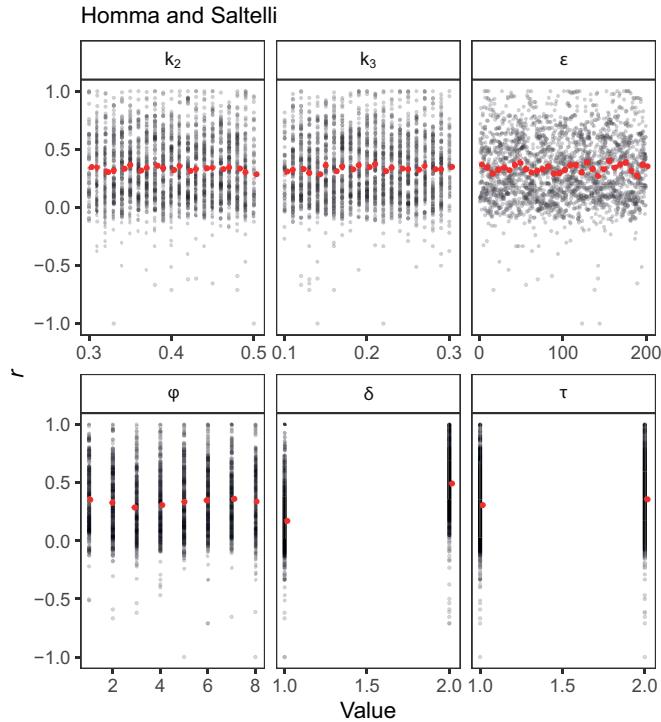


FIG. S10: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

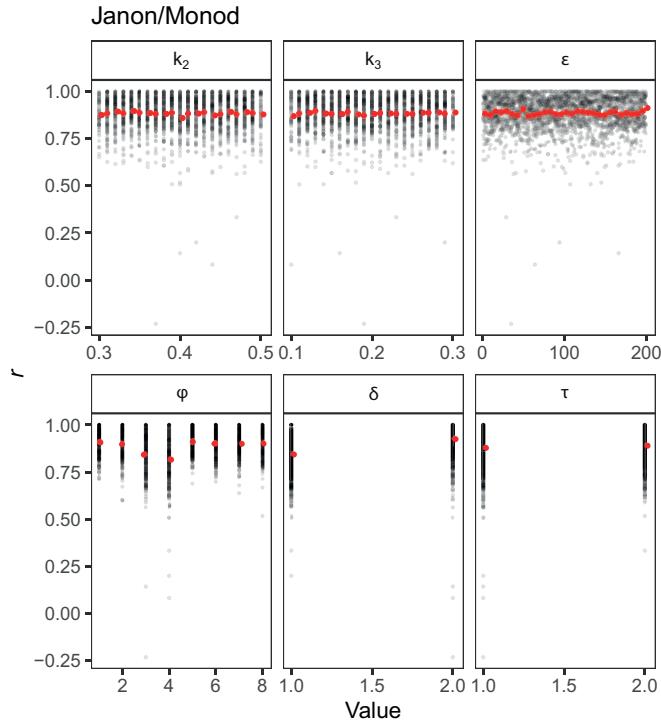


FIG. S11: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

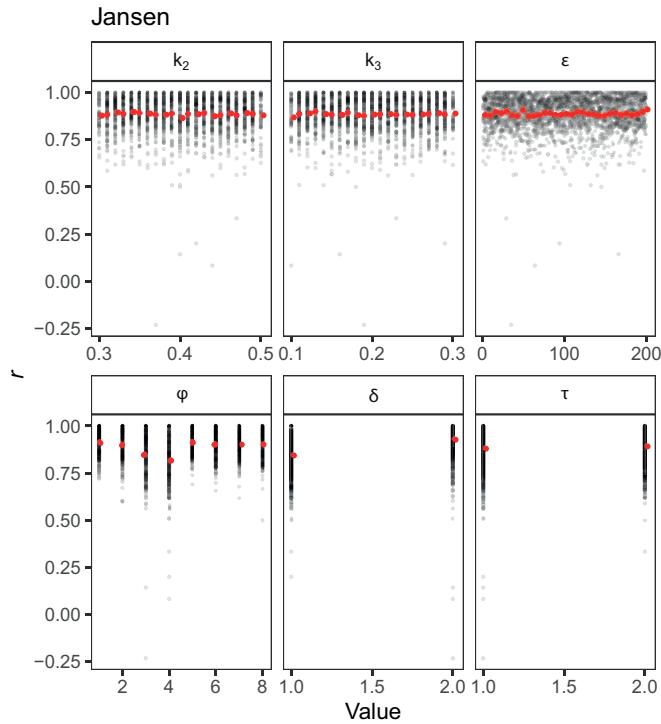


FIG. S12: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

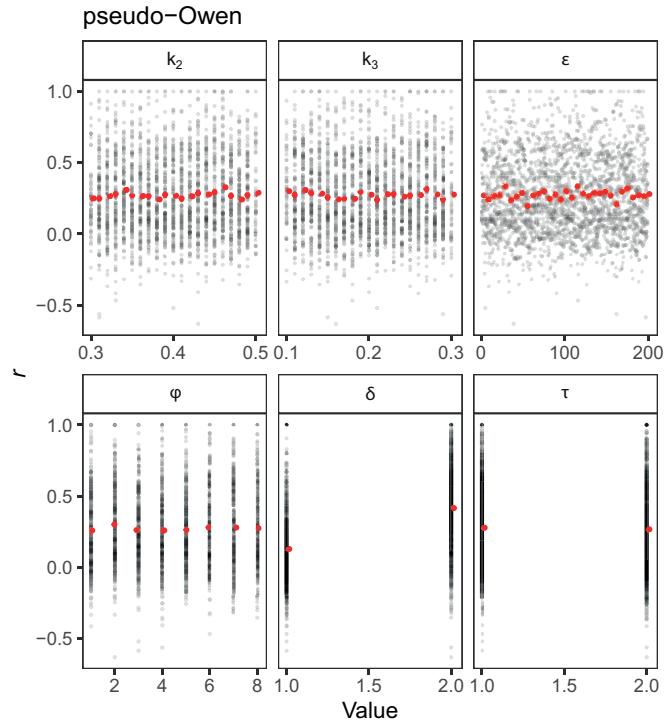


FIG. S13: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

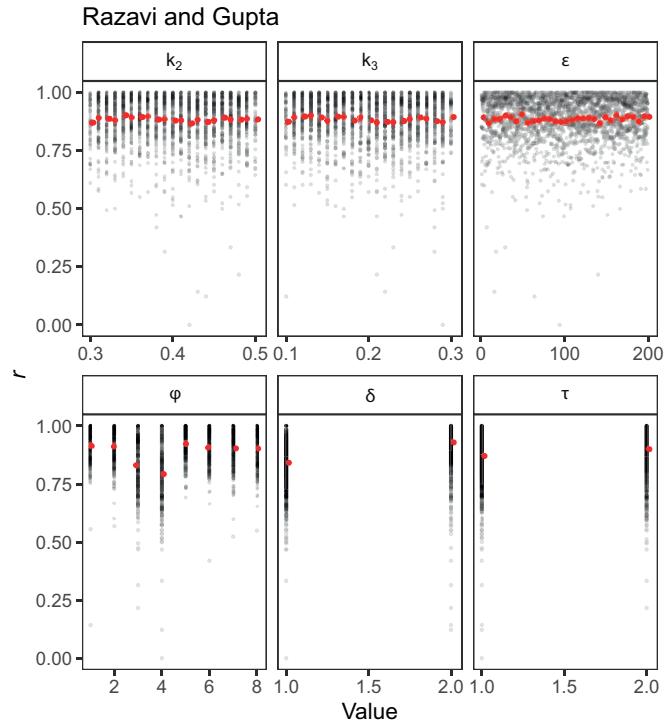


FIG. S14: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

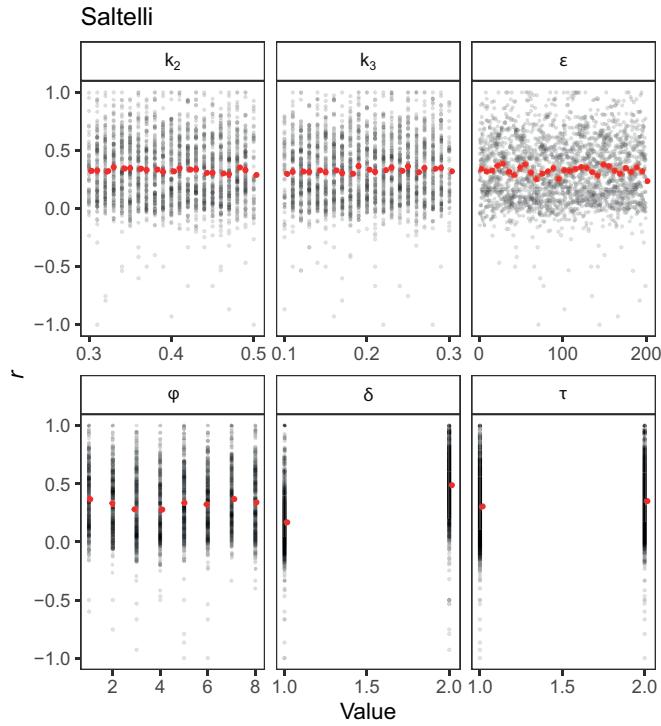


FIG. S15: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

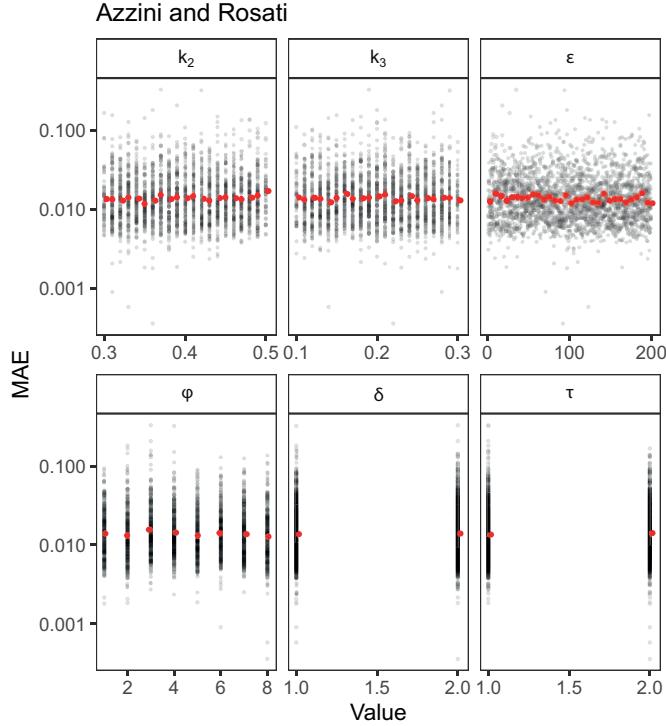


FIG. S16: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

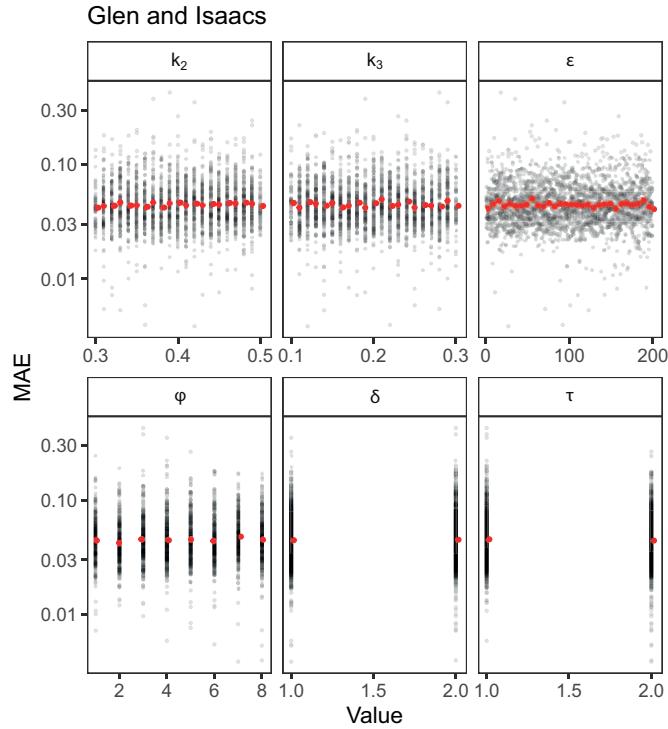


FIG. S17: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

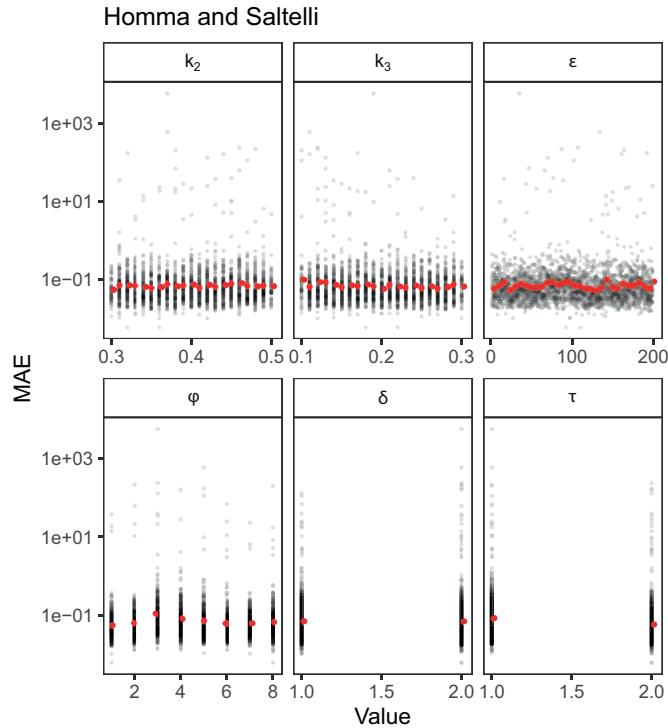


FIG. S18: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

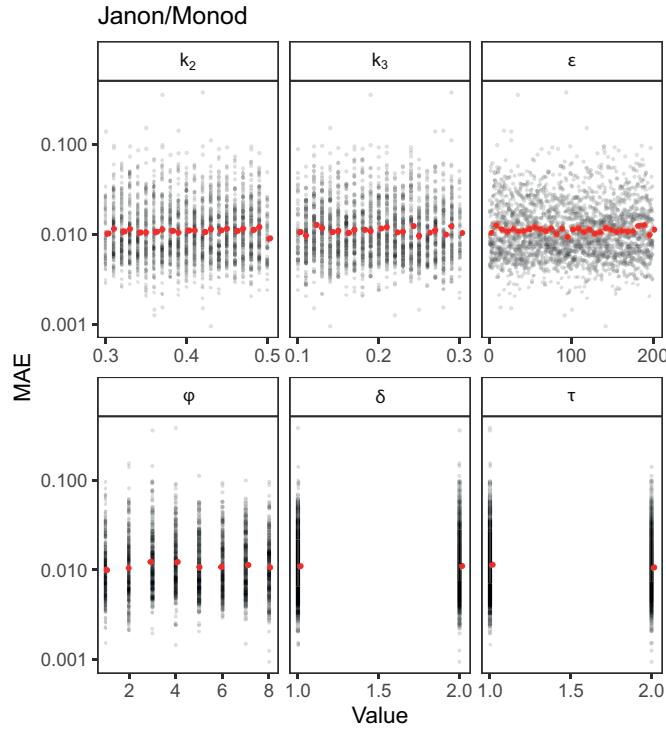


FIG. S19: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

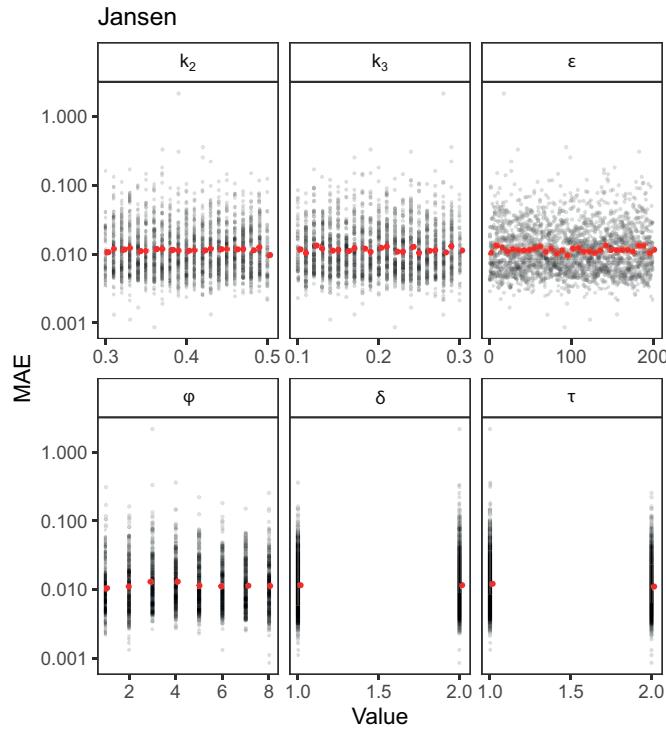


FIG. S20: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

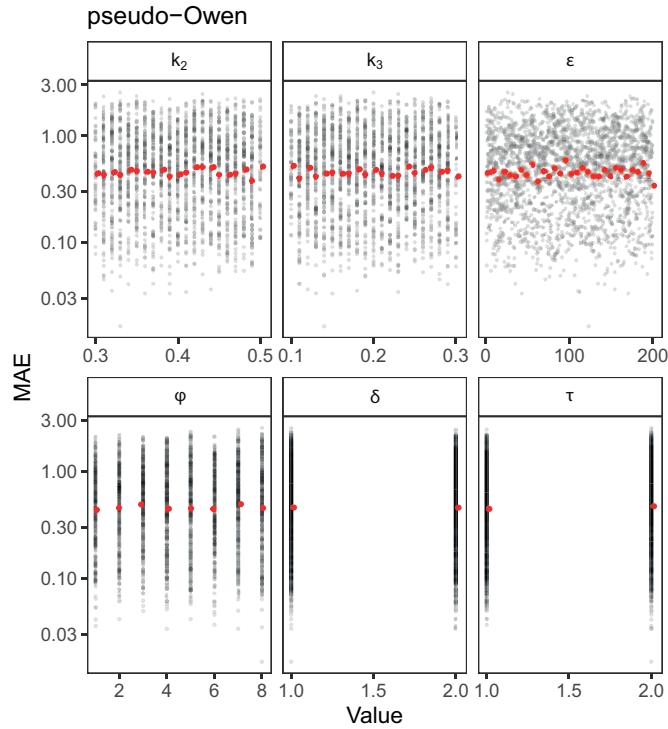


FIG. S21: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

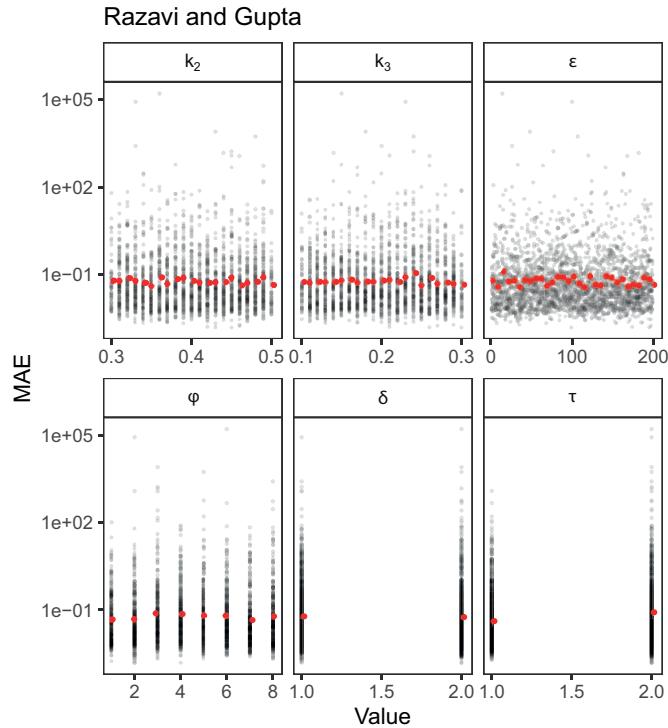


FIG. S22: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).

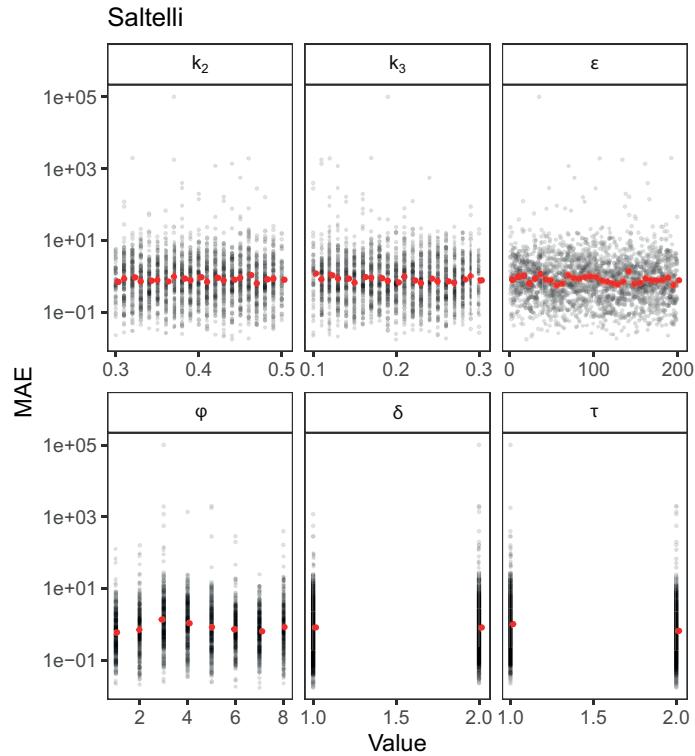


FIG. S23: Scatterplots of the model inputs against the model output. The red dots show the mean value in each bin (we have set the number of bins arbitrarily at 30).