Data cleaning:

Data from the TAMSA was obtained from CIMMYT Research Data repository (<https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10548038>) {citation available on the website}. The variables of interest were extracted from the different parts of the survey. Units conversion were performed (mainly weight to kg and area to ha). Each variable was screen for outliers. Values above a certain threshold or the 5 highest percentiles were discarded. All the data cleaning workflow is available as a bookdown {citation} document at <https://github.com/ALanguillaume/TAMASA> [temporary address].

Sensitivity analysis:

To perform the sensitivity analysis we relied on an iterative sampling scheme as illustrated in Figure-XX. Each of the p variables used for our analysis underwent the same scheme N times. For the variables with a number of available data points greater than 5, a Latin hypercube sampling scheme was implemented to get the n most different combinations of those variables quantiles. This aims at optimally sampling the parameter space at each simulation run. In parallel, the mean and standard deviation of each selected variables was calculated. For each selected variable an appropriate gamma distribution was derived using a shape/scale parametrization such that shape = \frac{\mu^{2}}{\sigma^{2}} and scale = \frac{\sigma^{2}}{\mu}. n values corresponding to the n quantiles derived from the Latin hypercube were drawn from each gamma distribution. Gamma distributions were used to reflect the skewness of the majority of variables see Figure-XX. In case, the number of available data points was lower than 5, existing values were simply sampled n times. Once either the Latin hypercube or the random sampling schemes performed the resulting simulated data was feed as input to the model. The complete sampling process was then repeated before each of the N model runs.

Software:

All data handling and simulation was implemented using the R programming language {citation up to date}. Latin hypercube sampling was done using the lhs r package{citation on CRAN}.

