1 Uncertainty Quantification Framework

This chapter consists of three parts. The first section gives an overview of UQ and introduces the basic notation. The second part describes the subdiscipline sensitivity analysis. After general remarks, this part, in turn, is divided into quantitative and qualitative GSA. I explain the most common measures for both levels of GSA and how they relate. The third part concludes with remarks on the role of correlated input parameters.

1.1 Overview of Uncertainty Quantification

Model-based forecasting includes two main steps (Smith (2014))[1]: The first step is calibration. In this step, the input parameters of the model are estimated. The second step is the prediction. The prediction contains the model evaluation at the estimated parameters to make statements about the future. These statements are made in a probabilistic way. Thereby, the uncertainty of these statements is emphasised.[2]

There are four sources of uncertainty in modern forecasting that are based on complex computational models (Smith (2014))[3]. The first source, the model uncertainty, is the uncertainty about whether the mathematical model represents the reality sufficiently.[4] The second source, the input uncertainty, is the uncertainty about the size of the input parameters of the model. The third one, the numerical uncertainty, comes from potential errors and uncertainties introduced by the translation of a mathematical to a computational model. The last source of uncertainty, the measurement uncertainty, is the accuracy of the experimental data that is used to approximate and calibrate the model.

The thesis deals with the second source of uncertainty, the input uncertainty. In my view, this is the source for which UQ offers the most and also the strongest instruments. This may result from the fact that the estimation step produces standard errors as basic measures for the variation or uncertainty in the input parameter estimates. These can then be used to compute a variety of measures for the impact of the input uncertainty on the variation in the model output.

The following explains the basic notation. It is essential to define the quantity that one wants to predict with a model. This quantity is called the quantity of interest (henceforth QoI) and is denoted by Y. For instance, the QoI in the thesis is the impact of a 500 USD tuition subsidy for higher education on average schooling years. The uncertain model

---

[1]See page ix.

[2]The general procedure of model-based forecasting can also include other steps. However, steps like model validation and model verification can also be viewed as belonging to the analysis of the so-called model uncertainty. The concept of model uncertainty is briefly explained in the next paragraph.

[3]See page 4-7.

[4]However, apparently, there are not many powerful instruments to evaluate and improve the model uncertainty except comparing statements derived from the model to the data and then improving it where appropriate.

parameters $X_1, X_2, ..., X_k$ are denoted by vector $\mathbf{X}$. The function that computes QoI Y by evaluating a model and, if necessary, post-processing the model output is denoted by $f(X_1, X_2, ..., X_k)$. Thus,

$$Y = f(\mathbf{X}). \tag{1}$$

Large-scale UQ applications draw from various fields such as probability, statistics, analysis, and numeric. These disciplines are used in a combined effort for parameter estimation, surrogate model construction, parameter selection, uncertainty analysis, LSA, and GSA, amongst others. Drawing mainly from Smith (2014)[5], I briefly sketch the first four components. The last two components, local and especially global sensitivity analysis, are discussed more extensively after that.

Parameter estimation covers the calibration step. There is a large number of estimation techniques for various types of models. The thesis uses a maximum likelihood approach, as detailed in the model section and in Appendix C.

If the run time of a model is too long to compute the desired UQ measures, surrogate models are constructed to substitute the original model f (McBride and Sundmacher (2019)). These surrogate models are functions of the model input parameters which are faster to evaluate. The functions are also called interpolants because they are computed from a random sample of input vectors, drawn from the input distribution and evaluated by the model. Typically, a surrogate model is computed by minimising a distance measure between a predetermined type of function and the model evaluations at the sample points. Therefore, the surrogate model interpolates this sample. Some specifications, like orthogonal polynomials, have properties which can simplify the computation of UQ measures tremendously (Xiu (2010)).

Another way to reduce the computation time, not directly of the model but of UQ measures, is to reduce the number of uncertain input parameters as part of a parameter selection. The decision which parameters to fix is made based on sensitivity measures. This is called screening or factor fixing (Saltelli et al. (2008))[6]. This point will be taken up again after this overview.

Uncertainty analysis is the core of the prediction step. It comprises two steps. The first step is the construction of the QoI's probability distribution by propagating the input uncertainty through the model. For instance, this can be achieved by evaluating a sample of random input parameters (as also required for the construction of a surrogate model). The second step is the computation of descriptive statistics like the probabilities for a set of specific events in the QoI range using this distribution. Both steps are conceptually simple. The construction of the probability distribution is also important for designing subsequent steps like a sensitivity analysis. For example, if the distribution is unimodal and symmetric, then variance-based UQ measures are meaningful. If the distribution has

---

[5]See page 8-10.
[6]See page 33-34.

a less tractable, for instance, a bimodal shape, then density-based measures are better suited (Plischke et al. (2013)).

1.2 Sensitivity Analysis

This section draws largely from Saltelli et al. (2004) and Saltelli et al. (2008). They define sensitivity analysis as "the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input" Saltelli et al. (2004, p. 42). This apportioning implies a ranking of input parameters in terms of their importance for the model output. Saltelli et al. (2004, p. 52) define the most important parameter as "the one that [if fixed to its true, albeit unknown value] would lead to the greatest reduction in the variance of the output Y. Therefore, a factor is not important if it influences the output Y directly but rather its variance.

Sensitivity analysis includes different objectives. These have to be determined at first because the choice of methods depends on these objectives. Typically, the main and final goal is factor prioritisation. This is the aforementioned ranking of input parameters in terms of their importance. This ranking can be used to concentrate resources on the data acquisition and estimation of a subset of parameters. The methods that meet the demands of factor prioritisation best are called quantitative. These typically require the highest computational effort.

There are multiple other objectives. The one additional objective featured in this thesis is screening/factor fixing. It is basically the same as factor prioritisation except that it only aims to identify the input parameters that can be fixed at a given value without significantly reducing the output variance. Therefore, it focuses on the lowest parameters in the potential importance ranking. The reason why one would purse factor fixing instead of factor prioritisation is computational costs. As factor fixing generates less information than factor prioritisation, less powerful methods can be applied. These methods require less computational resources and are called qualitative. Factor fixing can be used to prepare factor prioritisation for models that are more costly to evaluate. In this sense, it serves the same purpose as surrogate modelling.

Another important distinction is local versus global sensitivity analysis. It essentially refers to the applied methods. In fact, the definition by Saltelli et al. (2004) is already tailored to a global sensitivity analysis. In contrast to the given definition, "until quite recently, sensitivity analysis was [...] defined as a local measure of the effect of a given input on a given output" Saltelli et al. (2004, p. 42). This old definition differs from the definition used here in two aspects. First, it emphasises the level of output rather than its variance. Second, it describes the measure as a local one. The drawback of this approach becomes clear by considering an example of a local sensitivity measure. This measure is the so-called system derivate $D_i = \frac{\partial Y}{\partial X_i}$ (Rabitz (1989)). The derivative is typically

computed at the mean, $\overline{X_i}$, of the estimate for X_i. D_i is a so-called one-at-a-time measure because it changes only one factor. It has the following four drawbacks: First, it does not account for the interaction between multiple input parameters because it is one-at-a-time. Second, If the model derivation is not analytical, the choice of the (marginal) change in X_i is arbitrary. Third, the local derivative at $\overline{X_i}$ is only representative for the whole sample space of a random input if the model is linear in X_i. Fourth, the measure does not relate to the output variance $Var(Y)$. For these reasons, the fields, its definition and its methods have evolved beyond the notion of local sensitivity analysis. Yet, until recently, the main part of applications in different fields, such as physics (Saltelli et al. (2004))[7] and economics (Harenberg et al. (2019)), still uses local measures.

1.2.1   Quantitative Global Sensitivity Analysis

The quantitative GSA aims to determine the precise effect size of each random input parameter and its variation on the output variation. The most common measures in quantitative GSA are the Sobol' sensitivity indices. Equation (2) gives the general expression for the first order index. Let $\mathrm{Var}_{X_i}(Y|X_i)$ denote the variance of the model output Y conditional on input parameter X_i. Then,

$$S_i = \frac{\mathrm{Var}_{X_i}(Y|X_i)}{\mathrm{Var}(Y)}. \tag{2}$$

The equation becomes clearer with the following equivalent expression in Equation (3). For this purpose, let $\sim i$ denote the set of indices except i. The expectation of Y for one specific value of X_i equals the average of the model evaluations from a sample, $\boldsymbol{\chi_{\sim i}}$, of $\mathbf{X_{\sim i}}$ and a given value $X_i = x_i^*$. Then, we use $\mathbb{E}[f(X_i = x_i^*, \boldsymbol{\chi_{\sim i}})] \stackrel{\text{def}}{=} \mathbb{E}_{\mathbf{X}_{\sim i}}[Y|X_i]$ to write the first-order Sobol' index as the variance of $\mathbb{E}_{\mathbf{X}_{\sim i}}[Y|X_i]$ over all x_i^*s as

$$S_i = \frac{\mathrm{Var}_{X_i}\left(\mathbb{E}_{\mathbf{X}_{\sim i}}[Y|X_i]\right)}{\mathrm{Var}(Y)}. \tag{3}$$

The first-order index does not include the additional variance in Y that may occur from the interaction of $\mathbf{X_{\sim i}}$ with X_i. This additional variance is included in the total-order Sobol' index given by Equation (4). It is the same expression as in Equation (3) except that the positions for X_i and $\mathbf{X_{\sim i}}$ are interchanged. Conditioning on $\mathbf{X_{\sim i}}$ accounts for the inclusion of the interaction effects of X_i.

$$S_i^T = \frac{\mathrm{Var}_{\mathbf{X}_{\sim i}}\left(\mathbb{E}_{X_{\sim i}}[Y|\mathbf{X_{\sim i}}]\right)}{\mathrm{Var}(Y)} \tag{4}$$

---

[7]See page 42.

Computing these measures requires many function evaluations, even if an estimator is used as a shortcut (Saltelli et al. (2004))[8]. The more time-intense one function evaluation is, the more utility provides the factor fixing based on qualitative measures.

1.2.2 Qualitative Global Sensitivity Analysis

Qualitative GSA deals with the computation of measures that can rank random input parameters in terms of their impact on the function output and the variability thereof. This is done to a degree of accuracy that allows distinguishing between influential and non-influential parameters. If the measures for some input parameters are negligibly small, these parameters can be fixed so that the number of random input parameters decreases for a subsequent quantitative GSA. This section explains the qualitative measures and the trade-off between computational costs and accuracy.

The most commonly used measures in qualitative GSA is the mean EE, μ, the mean absolute EEs, μ^*, and the standard deviation of the EEs, σ. The EE of X_i is given by one individual function derivative with respect to X_i. The "change in", or the "step of" the input parameter, denoted by Δ. The only restriction is that $X_i + \Delta$ is in the sample space of X_i. The Elementary Effect, or derivative, is denoted by

$$d_i^{(j)} = \frac{f(\mathbf{X}_{\sim \mathbf{i}}^{(\mathbf{j})}, X_i^{(j)} + \Delta^{(i,j)}) - f(\mathbf{X})}{\Delta^{(i,j)}}, \tag{5}$$

where j is an index for the number of r observations of X_i. Note, that the EE, $d_i^{(j)}$ is equal to the aforementioned local measure, the system derivate $S_i = \frac{\partial Y}{\partial X_i}$, except that the value Δ has not to be infinitesimally small. To offset the third drawback of d_i or S_i, that base vector X_i does not represent the whole input space, one computes the mean EE, μ_i, based on a random sample of X_i from the input space. The second drawback, that interaction effects are disregarded, is also offset because elements $X_{\sim i}$ are also resampled for each new X_i. This measure is given by

$$\mu_i = \frac{1}{r} \sum_{j=1}^{r} d_i^{(j)}. \tag{6}$$

Thus, μ_i is the global version of $d_i^{(j)}$. Then, the standard deviation of the EEs writes $\sigma_i = \sqrt{\frac{1}{r} \sum_{j=1}^{r} (d_i^{(j)} - \mu_i)^2}$. The mean absolute EE, μ_i^* is used to prevent observations of opposite sign to cancel each other out:

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^{r} \left| d_i^{(j)} \right|. \tag{7}$$

---

[8]See page 124 -149.

Step $\Delta^{(i,j)}$ may or may not vary depending on the sample design that is used to draw the input parameters.

One last variant is provided in Smith (2014)[9]. That is, the scaling of μ_i^* by $\frac{\sigma_{X_i}}{\sigma_Y}$. This measure is called the sigma-normalized mean absolute EE:

$$\mu_{i,\sigma}^* = \mu_i^* \frac{\sigma_{X_i}}{\sigma_Y}. \tag{8}$$

This scaling accounts for the variability in X_i. Otherwise, the ranking would be influenced by differences in the variation of the input parameters. The reason is that the input space constrains Δ. If the input space is larger, then the function evaluation is divided by a larger Δ. If the model is non-linear, this may lead to extreme differences between parameters with larger and smaller σ. On the other hand, a GSA measure should reflect the uncertainty, or variation, in X_i given by σ_x. Therefore, $\mu_{i,\sigma}^* e$ is a better measure for the μ_i^* as a normed mean absolute derivative, but not as a measure of the propagated uncertainty from X_i to Y.

From the aforementioned set of drawbacks of the local derivate $D_i = \frac{\partial Y}{\partial X_i}$, two drawbacks are remaining for the EE method. The first drawback is the missing direct link to the variation in $Var(Y)$. The second drawback is that the choice of Δ is somewhat arbitrary if the derivative is not analytic. To this date, the literature has not developed convincing solutions to these issues.

In an attempt to establish a closer link between EE-based measures and Sobol' indices, Kucherenko et al. (2009) made two conclusions: The first conclusion is that there is an upper bound for the total index, S_i^T, such that

$$S_i^T \leq \frac{\frac{1}{r}\sum_{j=1}^r d_i^{2(j)}|}{\pi^2 \sigma_Y}. \tag{9}$$

This expression makes use of the squared EE. In light of this characteristic, the use of σ_i as a measure that aims to include the variation of d_i^j appears less relevant. Nevertheless, this rescaling makes the interpretation more difficult. The second conclusion is that the Elementary Effects method can lead to false selections for non-monotonic functions. This is also true if functions are non-linear. The reason is linked to the aforementioned second drawback, the arbitrary choice of step Δ. More precisely, depending on the sampling scheme, Δ might be random instead of arbitrary and constant. In both cases, Δ can be too large to approximate a derivative. If, for example, the function is highly non-linear of varying degree with respect to the input parameters $\mathbf{X}$, $\Delta > \epsilon$ can easily distort the results. Especially if the characteristic length of function variation is much smaller than Δ.

---

[9]See page 332.

In my view, a brief, intuitive summary of the link from EEs to Sobol' indices can be as follows: Sobol' indices are variance-based. Therefore, they deal with the squared distance of random draws from the input distribution to the mean, both evaluated by function f. The EE-based means are derivative-based, linear and they do not use the mean as a reference point. This last property implies that EE-based measures are derived from partly larger (untransformed) differences because step Δ has not to be taken from the mean. This property can cause large differences and deflate the EE-based measures vis-á-vis the Sobol' indices if f is non-linear. However, given the number of samples is large, $\mu_i^* = 0$ implies $S_t = 0$. This is because the Elementary Effects also contain differences from random draws to draws that are close to the mean. Yet, it remains unclear how to make inferences from $\mu_i^* > 0$ about Sobol' indices in the context of factor fixing.

1.3   Correlated input parameters

So far, all measures described in this chapter assumed uncorrelated input parameters. Typically, this assumption is not met by practical applications as joint estimations of parameters tend to produce correlated estimates. This gives reason to expand sensitivity measures to a more general setting.

Today, several recent contributions deal with the extension of the Sobol' sensitivity measures to the setup with correlated inputs. For instance, estimators for two complementary sets of measures have been developed by Kucherenko et al. (2012) and Mara et al. (2015). On the other hand, the only contribution to the computation of EE-based screening measures for the correlated setup has been made by Ge and Menendez (2017). Some authors, e.g. Saltelli et al. (2004)[10] even negate the necessity for expanded Elementary Effects due to overcomplexity. Obviously, this can lead to false results.

As the computation time of the Keane and Wolpin (1994) model is considerable[11], the thesis reviews Ge and Menendez (2017) and computes measures derived from this contribution to the occupational choice model. The aim is to fix parameters as preparation for a potential quantitative GSA. The measures correspond to $\mu_{i,\sigma}^*$ and μ_i^*. The measure in Kucherenko et al. is discarded because their results are not valid in the presence of correlations.[12]

The next chapter reviews uncertainty analyses and quantitative GSAs for economic models in the literature.

---

[10]See page 46.

[11]On my machine it takes approximately six seconds to compute the QoI. Given that Sobol' indices can require a five digits number of evaluations per input parameter, the computation time is not small.

[12]The fundamental reason is that the unique first- and higher-order Sobol' indices do not add to one if input parameters are correlated.

References

Ge, Q. and M. Menendez (2017). Extending morris method for qualitative global sensitivity analysis of models with dependent inputs. *Reliability Engineering & System Safety 100*(162), 28–39.

Harenberg, D., S. Marelli, B. Sudret, and V. Winschel (2019). Uncertainty quantification and global sensitivity analysis for economic models. *Quantitative Economics 10*(1), 1–41.

Keane, M. P. and K. I. Wolpin (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Review of Economics and Statistics 76*(4), 648–672.

Kucherenko, S. et al. (2009). Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation 79*(10), 3009–3017.

Kucherenko, S., S. Tarantola, and P. Annoni (2012). Estimation of global sensitivity indices for models with dependent variables. *Computer physics communications 183*(4), 937–946.

Mara, T. A., S. Tarantola, and P. Annoni (2015). Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental modelling & software 72*, 173–183.

McBride, K. and K. Sundmacher (2019). Overview of surrogate modeling in chemical process engineering. *Chemie Ingenieur Technik 91*(3), 228–239.

Plischke, E., E. Borgonovo, and C. L. Smith (2013). Global sensitivity measures from given data. *European Journal of Operational Research 226*(3), 536–550.

Rabitz, H. (1989). Systems analysis at the molecular scale. *Science 246*(4927), 221–226.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons.

Smith, R. C. (2014). *Uncertainty Quantification: Theory, Implementation, and Applications*. Philadelphia: SIAM-Society for Industrial and Applied Mathematics.

Xiu, D. (2010). *Numerical methods for stochastic computations: a spectral method approach*. Princeton university press.