A Gradient-based Sensitivity Analysis Method for Complex Systems

Ingrid Kovacs\textsuperscript{1}, Alexandra Iosub\textsuperscript{2}, Marina Țopa\textsuperscript{1}  
\textsuperscript{1}Technical University of Cluj-Napoca, Romania  
\textsuperscript{2}Gheorghe Asachi Technical University of Iași, Romania  
\textsuperscript{3}Infineon Technologies AG, Neubiberg, Germany  

\texttt{ingrid.kovacs@staff.utcluj.ro}

\begin{abstract}
Most sensitivity analysis methods from literature impose specific experiment plans (design-of-experiments). Moreover, the size of the experiment plan, which is in conjunction with the number of system evaluations, increases with the number of factors that may affect the systems' behavior. This paper introduces a gradient-based global sensitivity analysis method, which overcomes these limitations. First, its performance is compared against six sensitivity analysis methods on sets of polynomial test functions. The comparison is carried out by means of the number of system evaluations implied and the reported factor ranking list. The proposed method proved to have comparable accuracy to the best of the six known methods—the EFAST method—while allowing a lower number of system evaluations. These two methods are further applied on an electronic system, an E-Bike application. In this case, the proposed method employs the verification plan of the EFAST method, as well as a standard Monte Carlo experiment plan with about one third of the system evaluations of the EFAST. Even with a much lower number of system evaluations, the proposed method yields the same result as the EFAST method in terms of the factor ranking list.

\textbf{Keywords}—Sensitivity analysis, variance decomposition, local regression
\end{abstract}

I. INTRODUCTION

Verification of modern electronic systems is hard pressed to keep pace with the fast increase of their complexity and the resulting even faster increasing number of factors which can influence their behavior. The responses of such complex systems may be influenced by hundreds of input parameters, called factors. These factors include design and process parameters, configuration settings and operating conditions such as supply voltage and temperature. In this context, system responses include the noise level, the power consumption, frequency characteristics, etc.

The verification of such complex systems is time-consuming and costly and a solution to shorten it is to reduce the analysis to a smaller number of factors. Although there may be hundreds of factors with potential impact on the system response(s), the parsimony principle states that only a few of them are responsible for most of the effect in a response.

As the relationship between factors and responses is often unknown or uncertain due to the complexity, one may apply Sensitivity Analysis (SA) [1-2] to determine the effect of a given factor on the response and determine the subset of factors most influential on the response. This way, one can concentrate further analysis only on these few most important factors, significantly reducing the verification time and analysis complexity. Thus, SA is usually the first phase of an experimental study on a system.

The choice of the SA method should be done depending on the types of factor-response relationships involved and the execution cost implied, especially for complex systems which include non-linear effects.

In [3], the SA methods are categorized into three groups: local SA methods, global SA methods and screening methods. Local methods are the ones in which the variation of the response is observed when only one factor is varied and the rest are held constant. Such examples are the One-Factor-at-a-Time (OAT) method [4] and the partial derivatives [5]. Note that for local SA, the factor-response relationship is assumed to be linear and the correlation between parameters is neglected.

In the global SA approaches, the response is observed when all factors are varied simultaneously [6-7]. The most common statistical procedures for global SA are the Analysis of Variance (ANOVA) [8] and the standardized regression coefficient (SRC) method [9]. However, these become very complex for a large number of factors or are valid only for linear factor-response relationships.

The statistical linear regression [10] is another global SA technique which aims to estimate the factor-response relationship, but the accuracy of the results is conditioned by the fitness of the model. Another direction for SA is the variance-based methods [11-12]. However, they impose particular experiment plans and the number of system evaluations increases with the number of factors, which can lead to high execution costs in terms of the computational time required for the system evaluations. The screening methods are recommended for systems which imply a high computational cost [11].

SA methods were applied in several domains such as building energy models and environmental modeling [13-15]. However, SA has not been often applied to electronic systems, although their increased complexity would warrant its use.

The paper proposes a gradient-based global SA method developed using the concept of local regression [10] for cumulating the local effects in order to unveil global
sensitivities. First, it is applied on second-order polynomial functions and its accuracy of determining the most important factors is compared to six SA methods from literature: three variance-based methods [11-12], the Morris OAT [12] and two entropy-based methods proposed by the authors in [16]. Then, the proposed method is applied on an E-bike application and its ranking of most important factors is compared to the ranking of the well-known EFAST SA method. In each case, the number of system evaluations involved is also reviewed.

The paper is organized as follows: Section II summarizes the related work, followed by the description of the proposed method in Section III. Section IV presents a comparison of the proposed method to several SA methods on synthetic test functions. The SA results on an electronic application are presented in Section V. Conclusions are drawn in Section VI.

II. RELATED WORK

Different SA approaches are applicable on different types of models. A common and simple first step is to inspect the scatter plots [9] and qualitatively determine nonlinearities, non-monotonicities and correlations between the factors and responses. For linear models, methods such as the Partial Correlation Coefficients are suggested [9]. In the case of non-linear but monotonic models, the Partial Rank Correlation Coefficient and Standardized Rank Regression Coefficients are a suitable choice [9]. For non-linear and non-monotonic models, the variance-based methods should be used, while for systems which imply a high computational cost, screening methods such as the Morris OAT [11] are recommended.

A. The variance-based methods

The variance-based SA methods [11-12] aim to estimate the amount of variance added by an input factor \( X_i \) to the unconditional output variance \( V(Y) \). Furthermore, \( V(Y) \) can be decomposed into conditional variances as in (1):

\[
V(Y) = V(E(Y | X_i)) + V(Y | X_i)
\]

where \( E(V|Y|X_i) \) is the expectation value over the whole interval of variation of the input factor \( X_i \), where \( K \), where \( K \) denotes the total number of factors and \( V(E|Y|X_i) \) is the variance of the conditional expectation, which is simply denoted \( V_i \) and is often called ‘main effect’, i.e. the sensitivity of \( Y \) to \( X_i \). Then, the first order sensitivity index \( S_i \) is:

\[
S_i = \frac{V(E(Y | X_i))}{V(Y)}
\]

The first order sensitivity index measures only the main effect of a factor on the response and it does not take into account the interactions between input factors. Note that two factors are considered to be in interaction if their total effect on the response is not the sum of their first order effects. In order to determine the higher order effects (interactions and nonlinear effects), the so-called total order effect indices need to be computed:

\[
S_n = \frac{E[V(Y | X_n)]}{V(Y)} = \frac{V(Y) - V(E(Y | X_n))}{V(Y)} = 1 - \frac{V(E(Y | X_n))}{V(Y)}
\]

where \( X_n \) denotes all factors except \( X_i \).

Examples of these methods are: Fourier Amplitude Sensitivity Test (FAST), Extended Fourier Amplitude Sensitivity Test (EFAST) and Sobol’ indices [11-12].

Note that the number of system evaluations required by the variance-based methods to perform the SA is increasing as the number of factors increases. Moreover, each method imposes a specific experiment plan. Providing accurate conclusions with a minimum number of model evaluations is not a simple task, especially for complex systems.

B. The Morris OAT method

This approach aims to assess the impact on the output when changing one factor value at a time [1], [11]. The theory on how to produce the Morris mean \( \mu \) and its related standard deviation \( \sigma \) for each factor can be found in [1].

A high mean indicates a factor with an important total effect and a high standard deviation means an interaction with other factors or non-linear effects on the output.

C. The entropy-based methods

The authors proposed in [16] two methods based on the measure of statistical entropy: the Entropy Simple method, which determines main effects of factors and the Entropy Pair method, which is capable of determining also interactions between factors.

The core idea of these methods lies in transforming the information provided by scatter plots into quantifiable measures of the impact of the factor on the response by using the statistical entropy [17]. Recall that a random factor-response distribution reveals a low impact of the factor on the response (thus high entropy), while the order of the distribution, i.e. a pattern or deformation from randomness reveals an important impact.

The advantage of the entropy-based methods is that they do not impose a specific experiment plan for the SA. Also, the number of system evaluations does not increase with the number of factors. Moreover, there is no limitation about the orthogonality of the factors.

III. THE PROPOSED METHOD

The core idea of the method is to derive the global SA result by cumulating results yielded by local SAs which employ linear regression [10] on data subsets.

The factor-response function can be seen as a hyper-surface in the multidimensional space, usually with a very sophisticated shape. This multidimensional space can be very difficult to estimate, but for a small region of the factor space, the surface can be estimated with a plane. The orientation of the plane, tangent to the surface, is given by the gradient of the surface in that region. The orientation of the gradient provides information about the factors that contribute to the variation of the response.

For simplification, let us illustrate the idea by an example as in Fig. 1, where \( X \) represents a factor and \( Y \) the response. The classical regression approach (based on low-order
polynomials) is usually valid for regions where the response does not vary too much. However, for complex systems, the response may depend non-monotonically on the factors, meaning that higher order effects may remain unrevealed.

In Fig. 1 the dots represent the measured data, while the real response-factor relationship is illustrated with the continuous line. Using linear regression, one can estimate a line that passes through the dots and whose inclination provides information about the impact of the factor on the response. Using this classical regression approach, one would conclude that the factor has little impact, although it actually has a greater one.

Using local regression, one can estimate sophisticated effects, because for small variations of the variable, any function can be approximated by a line. We use the local regression approach for cumulating the effects and computing global sensitivities.

Fig. 2 illustrates the steps of the gradient-based SA method in detail. In the first step, the experiment is planned and the responses are measured, which is common to any SA approach. The experiment plan can be made based on possible previous knowledge about the system, but usually Monte Carlo is preferred, if no other analysis is already available. Let $N$ denote the number of simulations/measurements performed.

In the second step, for each point $i$, $i=1,..., N$, the subset of $q$ closest points is determined for the local regression. Fig. 3 is an illustration for a better understanding. In the third step, we perform a linear regression on the subspace and compute the regression coefficients $\{\beta_1, ..., \beta_K\}$ for each subset $i$.

After computing all partial regression coefficients for the subsets, one computes the simple and the higher order effects of $X_i$. The simple effect of $X_i$ ($Se_i$) is computed as the mean of the absolute partial regression coefficients from the $N$ subsets, while the higher order effect ($He_i$) is their variance, as described in (4-5). In the last step, the most important factors are determined. Note that higher $Se_i$ and $He_i$ means greater impact.

![Image](image1.png)

**Fig. 1.** Example of local linear regression for a non-monotonic function

**Fig. 2.** The steps of the Gradient SA method

\[ Se_i = \mu_i = \frac{1}{N} \sum_{i=1}^{N} |\beta_i| \]  \hspace{1cm} (4)

\[ He_i = \sigma_i^2 = \frac{1}{N} \sum_{i=1}^{N} (\beta_i - \mu_i)^2 \]  \hspace{1cm} (5)

Table I summarizes a comparison of the SA methods from the points of view of the class they are based on, the computed sensitivity indices and the execution cost implied. The advantages of the proposed method, compared to the state-of-the-art methods, are that it does not impose a specific experiment plan and the required number of system evaluations does not increase with the number of factors.

In Table I: $K$ - no. factors, $N$ - sample size, $M$ - no. Fourier coefficients, $OM$ - set of predefined frequencies, $\omega_{max} = \max\{\omega_1, ..., \omega_K\}$, $r$ - no. elementary effects, $Nr$ - search curves, and $N_{MC}$ - the sample size of the Monte Carlo simulation, if no other experiment plan is available.

In the next section we present a comparison of the discussed SA methods on custom test functions, where we discuss the accuracy of determining the important factors and the execution cost implied by each analysis.
In order to test the methods’ capability in identifying the set of important factors, we tested the methods via simple polynomial functions as in (6), which include different types of factor effects (main, quadratic and first order interactions), while the coefficients $\beta_i$ and $\beta_{ij}$ indicate the importance of factor $X_i$.

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j \quad (6)$$

Knowing a-priori the factor-response relationship, one can test the methods’ accuracy in determining the important factors of the system and highlighting also the execution cost can serve as a guideline when selecting the SA method for the study of a real system, for which the factor-response relationship is not known a-priori.

Based on the results from [18], we extended the validation approach on synthetic functions and also included the proposed method in the analysis. Thus, we consider 60 polynomial functions as in (6), each with a number of 30 factors. From these factors, we selected a set of target (important) factors and attributed 90% of the effects to them and distributed the rest of the effects to the other, non-important factors.

In order to reproduce real scenarios, we added random noise to the responses and tested the capability of the methods to determine the important factors even if the system was affected by noise.

Noise scenarios are common for measurement data, not simulation ones. A response may be affected by noise because of other factors which have not been considered for the analysis.

For the noise scenarios, we added a random Gaussian noise to the response and considered the measure of Signal-to-Noise Ratio (SNR) to find the variance of the noise as in (7):

$$SNR = \frac{\sigma_y^2}{\sigma_{\text{noise}}^2} \quad (7)$$

where $\sigma_y^2$ is the variance of the response and $\sigma_{\text{noise}}^2$ is the variance of the noise. Then, the responses affected by noise can be described as:

$$y_{\text{noise}} = y + \sigma_{\text{noise}} \cdot \eta_i \quad (8)$$

where $\eta_i$ are normally distributed pseudorandom numbers.

We consider seven SNR values $\text{SNRdB} = \{25, 20, 15, 10, 5, 0, -5\}$ dB for the analysis; note that $\text{SNR}=25$ dB has the same effect as if no noise was added.

As a measure of accuracy of the methods performance, we define the pass rate, which computes the percentage of factors selected a-priori to be important and which can be found also in the top five most important factors of a method. The top five is also a common number of important factors to which a SA is reduced.

For each polynomial function, we selected 2-4 important factors, so it is a satisfactory result if the factor is found in the top five most important factors.

Fig. 4 illustrates a comparison of several SA methods’ performance in terms of the accuracy. We compared the accuracy of the proposed Gradient SA method to three variance-based methods, the Morris OAT and the entropy-based methods proposed by the authors in [16].

Note that the FAST method is capable of determining only main effects of factors, so it was tested only for this type of factor-response relationships.

The proposed Gradient SA method gains slightly better accuracy than the EFAST variance-based method and the Pair entropy-based method and it outperforms the Entropy Simple, the Morris and the Sobol methods.

Another analysis implies increasing the number of factors to 50 and test the accuracy of the methods. The results were similar as in the case of 30 factors.

We summarized also the execution cost implied by each method for different number of factors in Table II. For most of the methods, an increase of the number of factors implies also an increase of the number of experiment runs. The advantage of the proposed method is emphasized here, as the number of necessary experiment runs does not increase with the increase of the number of factors.

The application of different SA methods on test functions serves as a starting point in inspecting the compromise involved by the high accuracy in determining the important factors and the execution cost involved by the analysis.

The conclusions provide valuable information when selecting SA methods for a real system, where we have no information about the factor-response relationships. Another finding is that the important factors are easier to be determined from a lower number of factors.
![Image](image.png)

Fig. 4. Accuracy comparison of the SA methods; 30 factors

<table>
<thead>
<tr>
<th>Method</th>
<th>Execution cost (no. runs)</th>
<th>30 factors</th>
<th>50 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobol</td>
<td>1830</td>
<td>3030</td>
<td></td>
</tr>
<tr>
<td>FAST</td>
<td>23081</td>
<td>87241</td>
<td></td>
</tr>
<tr>
<td>EFAST</td>
<td>1950</td>
<td>3250</td>
<td></td>
</tr>
<tr>
<td>Morris</td>
<td>310</td>
<td>510</td>
<td></td>
</tr>
<tr>
<td>Entropy-based</td>
<td>300</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>Gradient</td>
<td>300</td>
<td>300</td>
<td></td>
</tr>
</tbody>
</table>

### V. EVALUATION OF THE SA METHODS ON THE E-BIKE APPLICATION

The methods are applied on an electronic system, which is an E-Bike application described in [19-20]. An E-Bike is a regular bicycle with an integrated electric motor to provide additional assistance.

The 15 factors included in the analysis are listed in Table III. Note that the factors are of different nature, including system architecture properties, component properties, system inputs and operating conditions. Two responses were considered: the acceleration time (denoted as AccTime) and the Efficiency in steady state (denoted as PowerEfficiency).

In order to validate the proposed Gradient SA method, we considered the results of the EFAST method as reference (similar as in [16]) and performed a cross-validation of the returned important factors.

The EFAST method is known to have high accuracy, but it comes with the disadvantage of a high execution cost, which translates into 975 simulations and approximately 112 hours of execution time. The reason for the high execution cost is that the application uses a direct driven motor with a large mechanical time constant.

The first step for validation was to apply the Gradient method on the experimental runs of the EFAST method and to compare the top of the most important factors returned by the two methods. It is a considerable advantage that the Gradient SA method does not impose its own experiment plan and the SA can be applied on the results of the EFAST experiment plan. As the system contained a number of 15 factors, we considered the top three most important factors for both responses.

As a second step, we applied the Gradient method on a lower number of simulations and compared the top of returned factors to the top of the EFAST method. In this case, for the Gradient SA experiment plan we considered a uniform Monte Carlo design with 300 runs. The purpose of this analysis was to determine if the Gradient SA has comparable accuracy with the EFAST method even with a lower number of simulations.

Note that the Monte Carlo design with 300 runs was chosen because this is the common number of simulations performed at early stages of IC verification. Moreover, from previous Gradient SA analysis on the same system, this was the minimum number of simulations for which accurate results could be obtained.

Table IV summarizes the results of the SA analysis for the EFAST method (first column) and the Gradient method applied on the simulation runs of the EFAST method (second column). Note that $S_{ij}$ indicates the total sensitivity index of the EFAST method. The proposed Gradient method identifies the same important factors as the EFAST method, considered as reference.

One subsequent approach was to determine the rate in which the top of the most important factors is preserved when considering a lower number of simulations. The last column of Table IV presents the results. Even with a much lower number of simulations, the Gradient SA method determines the same important factors.

### TABLE II. THE EXECUTION COST

<table>
<thead>
<tr>
<th>Method</th>
<th>Execution cost (no. runs)</th>
<th>30 factors</th>
<th>50 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sobol</td>
<td>1830</td>
<td>3030</td>
<td></td>
</tr>
<tr>
<td>FAST</td>
<td>23081</td>
<td>87241</td>
<td></td>
</tr>
<tr>
<td>EFAST</td>
<td>1950</td>
<td>3250</td>
<td></td>
</tr>
<tr>
<td>Morris</td>
<td>310</td>
<td>510</td>
<td></td>
</tr>
<tr>
<td>Entropy-based</td>
<td>300</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>Gradient</td>
<td>300</td>
<td>300</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE III. THE FACTORS OF THE E-BIKE APPLICATION

<table>
<thead>
<tr>
<th>Components</th>
<th>Factors Label [units]</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Sensor</td>
<td>Gain of the sensor amplifier Offset voltage of the sensor amplifier</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LevelNoise [-]</td>
<td>Adjustment factor for the level of the noise floor of the amplifier</td>
</tr>
<tr>
<td></td>
<td>RoLPF [Ω]</td>
<td>Resistance and capacitance of the motor Low Pass Filter</td>
</tr>
<tr>
<td></td>
<td>Rs [Ω]</td>
<td>Shunt resistor</td>
</tr>
<tr>
<td>Angle Sensor</td>
<td>OffSin [V]</td>
<td>Offset in the sine sensor voltage Synchronicity error</td>
</tr>
<tr>
<td></td>
<td>ASin [-]</td>
<td>Synchronicity error</td>
</tr>
<tr>
<td></td>
<td>PhiY [°]</td>
<td>Mechanical misalignment</td>
</tr>
<tr>
<td>Motor</td>
<td>Rs [Ω]</td>
<td>Motor resistance</td>
</tr>
<tr>
<td></td>
<td>Ls [H]</td>
<td>Motor inductance</td>
</tr>
<tr>
<td></td>
<td>Ke [V/°/rad/s]</td>
<td>BackEMF voltage constant</td>
</tr>
<tr>
<td>Operating Conditions</td>
<td>Wref [V]</td>
<td>Speed reference</td>
</tr>
<tr>
<td></td>
<td>HumanInertia [kgm²]</td>
<td>Human moment of inertia</td>
</tr>
<tr>
<td>Inverter</td>
<td>Inverter/Supply [V]</td>
<td>Supply of the inverter</td>
</tr>
</tbody>
</table>

### TABLE IV. TOP MOST IMPORTANT FACTORS; AccTime RESPONSE

<table>
<thead>
<tr>
<th>EFAST (975 runs)</th>
<th>Gradient (975 runs)</th>
<th>Gradient (300 runs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor</td>
<td>S_{ji}</td>
<td>factor</td>
</tr>
<tr>
<td>Wref</td>
<td>0.629</td>
<td>HumanInertia</td>
</tr>
<tr>
<td>Ke</td>
<td>0.228</td>
<td>Ke</td>
</tr>
<tr>
<td>GainA</td>
<td>0.073</td>
<td>GainA</td>
</tr>
</tbody>
</table>

Note that the rest of the factors (4-15) would appear in the top only by chance and their ranking is different from one scenario to the other. When deciding upon the important and unimportant factors after the SA analysis, one has to inspect also the sensitivity indices returned by the method. Usually, a
clear boundary can be determined in the sensitivity indices values, which in the present case is after the third factor.

The same approach is applied for the Power Efficiency response. Table V presents the top three most important factors with the Gradient method applied on the simulation results of the EFAST method’s experiment plan (second column). Note that the top of the factors is the same for both methods.

Then, we apply the Gradient SA approach on a Monte Carlo design with 300 runs and the results are illustrated in the last column of Table V. The top three most important factors of the Gradient method remains the same.

By applying two SA methods and cross-validating the results, we were able to determine the most influential factors on the E-Bike’s responses of interest.

### VI. CONCLUSIONS

A novel gradient-based global SA method was introduced which relies on the local regression approach. Its performance was compared to other SA methods in terms of accuracy of determining factor–response relationships and the execution cost implied by the analysis.

A first step was to apply the method on polynomial test functions, where the true factor-response relationship was known and the accuracy of the methods in identifying those relationships was easy to test. At this step, we considered three variance-based SA methods from literature, the Morris OAT method and the entropy-based methods proposed earlier by the authors. The proposed Gradient SA had better results than the well-known EFAST variance-based SA method, with the advantage of a lower execution cost. Moreover, it had comparable results to the earlier proposed entropy SA methods.

A second step was to apply the method on an E-bike application. For this analysis, we considered the results of the well-known EFAST method as reference because of its high accuracy proved also on the polynomial test functions. As in this case we did not know the true set of factors impacting the response, we compared the rankings of the factors returned by the EFAST and Gradient SA methods.

Taking advantage of the fact the Gradient SA method does not impose a specific experiment plan, we applied it also on the experiment results of the EFAST method. The conclusion was that the top three most important factors was the same for both methods and responses.

A subsequent analysis was to consider a uniform Monte Carlo design with a much lower number of runs and apply the Gradient SA method on it. Even with three times less number of runs, compared to the EFAST method, the Gradient SA method succeeded in identifying the same important factors as the EFAST method.

### ACKNOWLEDGMENT

This work was co-funded by the European Regional Development Fund through the Operational Program “Competitiveness” PO.C -1.2.3-G-2015, project P_40_437, contract 19/01.09.2016.

### REFERENCES


