Calculation and Display of Sensitivity Coefficients

Yevgeniy V. Zagayevskiy and Clayton V. Deutsch

Sensitivity analysis is an essential part of mathematical modeling that describes the relationship and degree of dependence between output responses and input parameters of a given model. This paper focuses on the regression approach to sensitivity analysis and a two-step methodology is implemented. In the first step the degree of uncertainty is analyzed and in the second step the sensitivity to each input is quantified. Two programs are written for conducting a sensitivity study. `sabor.exe` *calculates sensitivity coefficients for linear and quadratic regression.* `bitornado.exe` *displays values of sensitivity coefficients graphically through a bivariate tornado chart. The proposed sensitivity analysis tool is compared to other conventional techniques in two examples 1) a trivial synthetic example of the "volume of original oil in place" and 2) a case study based on actual field data. The regression approach is found to be more accurate than conventional tornado chart and easy to implement.*

1. Introduction

While working with large datasets consisting of many variables it is important to understand the relationship between variables. Unrelated variables can be omitted from the modeling process to avoid noise and to reduce computational requirements. When a variable has a large effect on the output of the model, it should be retained and modeled carefully. The procedure of defining the relationship between input variables and the resulting variation in the output of a model is deemed a *sensitivity analysis* (Saltelli et al., 2008).

Conventionally the covariance or correlation matrices are used to assess the relationship between variables; however, when there are complex nonlinear relationships between variables such summaries can be inadequate. Tornado chart is useful graphical method of presenting sensitivity analysis results. The authors suggest using a linear or quadratic regression proxy model to assess the relationship between variables and establish sensitivity coefficients. Moreover, the uncertainty in these sensitivity coefficients is also important.

The paper is organized as follows. First the definition of a typical sensitivity analysis is given and conventional techniques are reviewed. Correlation coefficients are used to characterize the linear relationship between variables and tornado chart is used to convey the results. The regression approach to sensitivity analysis is explored and demonstrated with two examples 1) a synthetic 'volume of original oil in place' (*OOIP*) example is used to explain the methodology, and 2) a case study is used to highlight the benefits of the methodology.

2. Conventional Sensitivity Analysis

A sensitivity study explores how and to what extent variation in input variable effects variation (uncertainty) in output variable of a model and ranks input variables according to their effect values (Saltelli et al., 2008).

Prior to conducting a sensitivity study the strength of the relationship between variables can be assessed to identify weak relationships with output variables for the purpose of removing variables from the model. For instance, secondary variables (such as seismic attributes) that are poorly correlated to primary variables (such as porosity) are eliminated. Typically correlation coefficients are used to describe relationship between two variables.

$$\rho_{XY} = \frac{COV(X,Y)}{\sqrt{VAR(X) \cdot VAR(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \text{ and } |\rho_{XY}| \le 1 \tag{1}$$

where $COV(X,Y)$ or σ_{XY} – covariance between random variables X and Y; $VAR(X)$ and $VAR(Y)$ are variances of random variables X and Y respectively; σ_X *and* σ_Y are standard deviations of random variables X and Y respectively.

Covariance can be used as an alternative to the correlation and considers the standard deviation of the variables:

$$COV(X,Y) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} \tag{2}$$

where N is the number of available values of single random variable; x_i is the ith value of random variable X; $\overline{x}$ is the average value or mean of random variable X.

If more than two variables are present, the correlation or covariance is calculated for every pair of variables presented in a matrix.

For sake of the simplicity let's consider a model with single response Y and several inputs X_i, $i = 1, ..., n_I$. The relationship between values of model output and inputs can be expressed as shown in Equation (3).

$$y_i = f(x_1, x_2, ..., x_{n_I}), \quad i = 1, ..., N \tag{3}$$

The correlation matrix is a useful tool for summarizing the bivariate relationships between many variables; however, only linear relations are assessed.

The main idea of the proposed tornado chart is to show how a response is sensitive to its input variables. A tornado chart is generated for each output variable (Figure 1). On the horizontal axis the response is plotted with the mean and two fixed quantiles and associated deviations $\overline{y} + \Delta_i y$ & $\overline{y} - \Delta_i y$ are found with:

$$\overline{y} + \Delta_i y = f(\overline{x}_1, ..., x_i^{high\ quantile}, ..., \overline{x}_{n_I})$$
$$\overline{y} - \Delta_i y = f(\overline{x}_1, ..., x_i^{low\ quantile}, ..., \overline{x}_{n_I}) \tag{4}$$

where $x_i^{high\ quantile}, x_i^{low\ quantile}$ are the desired quantiles of the X_i input variables.

The tornado chart is a visual tool to represent sensitivity analysis results. A more general form of the tornado chart using the proposed regression approach is discussed.

3. Sensitivity Analysis based on Regression

To determine how the response fluctuates with the input a mathematical model to generate the response must be assumed. Recall that the model has only one output Y and several inputs X_i, $i = 1, ..., n_I$. Note that all inputs and outputs of the model are random variables. If N possible values of the random variables are known, then their distributions can be summarized by the mean value, variance, range of variation, etc. The regression approach requires sets of random variable values that are used to derive the sensitivity coefficients.

Let's consider two types of mathematical models: linear and quadratic. The linear model is simpler, but the quadratic model produces more plausible and precise results, since interactions between input variables are considered in the model. The model expressions are:

$$\hat{y} = b_0 + \sum_{i=1}^{n_I} b_i x_i \qquad \rightarrow \text{linear model} \tag{5}$$

$$\hat{y} = b_0 + \sum_{i=1}^{n_I} b_i x_i + \sum_{k=1}^{n_I} \sum_{l \geq k}^{n_I} c_{kl} x_k x_l \qquad \rightarrow \text{quadratic model} \tag{6}$$

where $\hat{y}$ is the modeled response; b_i is the ith regression coefficient corresponding to i^{th} input variable X_i; c_{kl} is the interaction term between random variables X_k and X_l.

The coefficients are determined by solving for the minimum squared error (*LSE*) presented in Equation (7)

$$LSE = \sum_{j=1}^{N} (y_j - \hat{y}_j)^2 \tag{7}$$

The first derivative is set to 0:

$$LSE = \sum_{j=1}^{N}(y_j - b_0 - \sum_{i=1}^{n_I} b_i x_i)^2 \Rightarrow \begin{cases} \dfrac{\partial LSE}{\partial b_0} = 0 \\[2mm] \dfrac{\partial LSE}{\partial b_i} = 0, \forall i, i = 1,...,n_I \end{cases} \tag{8}$$

$$\begin{cases} \dfrac{\partial LSE}{\partial b_0} = 0 \\[2mm] \dfrac{\partial LSE}{\partial b_p} = 0, \forall p \end{cases} \Rightarrow \begin{cases} \sum_{j=1}^{N} 2 \cdot \left(y_j - b_0 - \sum_{i=1}^{n_I} b_i \cdot x_i \right) \cdot (-1) = 0 \\[2mm] \sum_{j=1}^{N} 2 \cdot \left(y_j - b_0 - \sum_{i=1}^{n_I} b_i \cdot x_i \right) \cdot (-x_p) = 0, \forall p, p = 1,...,n_I \end{cases} \tag{9}$$

Then

$$\begin{cases} \sum_{j=1}^{N} y_j = \sum_{j=1}^{N}\left(b_0 + \sum_{i=1}^{n_I} b_i \cdot x_i \right) \\[2mm] \sum_{j=1}^{N} y_j = \sum_{j=1}^{N}\left\{ \left(b_0 + \sum_{i=1}^{n_I} b_i \cdot x_i \right) \cdot x_p \right\}, \forall p \end{cases} \Rightarrow \begin{cases} \overline{y} = b_0 + \sum_{i=1}^{n_I} b_i \cdot \overline{x}_i \quad \rightarrow \text{single equation} \\[2mm] \overline{yx}_p = b_0 \overline{x}_p + \sum_{i=1}^{n_I} b_i \cdot \overline{x_i x}_p \quad \rightarrow n_I \text{ equations} \end{cases} \tag{10}$$

Solving these $n_I + 1$ equations we obtain $\{b_0, b_1,..., b_{nI}\}$ regression coefficients. Consider the equivalent in matrix form:

$$\begin{bmatrix} 1 & \overline{x}_1 & \overline{x}_2 & ... & \overline{x}_{n_I} \\ \overline{x}_1 & \overline{x}_1\overline{x}_1 & \overline{x}_1\overline{x}_2 & ... & \overline{x}_1\overline{x}_{n_I} \\ \overline{x}_2 & \overline{x}_2\overline{x}_1 & \overline{x}_2\overline{x}_2 & ... & \overline{x}_2\overline{x}_{n_I} \\ ... & ... & ... & ... & ... \\ \overline{x}_{n_I} & \overline{x}_{n_I}\overline{x}_1 & \overline{x}_{n_I}\overline{x}_2 & ... & \overline{x}_{n_I}\overline{x}_{n_I} \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ ... \\ b_{n_I} \end{bmatrix} = \begin{bmatrix} \overline{y} \\ \overline{yx}_1 \\ \overline{yx}_2 \\ ... \\ \overline{yx}_{n_I} \end{bmatrix} \tag{11}$$

$$[A] \times [b] = [C] \tag{12}$$

$$[b] = [A]^{-1} \times [C] \tag{13}$$

The same calculations are applied to the quadratic model.

$$LSE = \sum_{j=1}^{L}(y_j - b_0 - \sum_{i=1}^{n_I} b_i x_i - \sum_{k=1}^{n_I}\sum_{l \geq k}^{n_I} c_{kl} x_k x_l)^2 \Rightarrow \begin{cases} \dfrac{\partial LSE}{\partial b_0} = 0 \\[2mm] \dfrac{\partial LSE}{\partial b_i} = 0, \forall i \\[2mm] \dfrac{\partial LSE}{\partial c_{kl}} = 0, \forall k, \forall l > k \end{cases} \tag{14}$$

From here with application of linear algebra we obtain the required regression coefficients and interaction terms. Once the coefficients have been calculated the sensitivity coefficients can be obtained from (Saltelli et al., 2008):

$$S_i = \left. \frac{\partial R}{\partial P_i} \right|_{base\,case}$$

(15)

where R is the response (output); P_i is the ith parameter (input). Usually the base case is referred to the mean values of response and parameters. Also the Equation (15) can be rewritten in the differential form using partial derivatives (Equation (16)). Later it can be transformed into expression (17) where numerical approach is applied.

$$dR = \sum_{i=1}^{n_i} \frac{\partial R}{\partial P_i} \cdot dP_i = \sum_{i=1}^{n_i} S_i \cdot dP_i$$

(16)

$$\widehat{R} = R_{mean} + \sum_{i=1}^{n_i} S_i \cdot \Delta P_i$$

(17)

The same way of thinking can be applied to the quadratic model. Its expression would be:

$$\hat{R} = R_{mean} + \sum_{i=1}^{n_i} S_i \Delta P_i + \sum_{k=1}^{n_i} \sum_{l \geq k}^{n_i} Q_{kl} \Delta P_k \Delta P_l$$

(18)

If we compare Equations (5) and (6) to (17) and (18), their uniqueness is clear. Regression coefficients correspond to the sensitivity coefficients ($S_i \equiv b_i$, $\forall l$, $i = 1, ..., n_i$), coefficient b_0 is exactly the mean value of the response ($R_{mean} \equiv b_0$), and interaction terms correspond to Q coefficients. All calculations carried out in the regression model can be used to evaluate sensitivity coefficients (Saltelli et al., 2008).

Moreover the regression approach allows us to simultaneously generate the required mathematical model (linear or quadratic) while calculating the sensitivity coefficients as well.

However, sensitivity coefficients can be misleading as they are dependent on the units of the input parameters. The proposed standardized sensitivity coefficients (*SSC*) resolve this issue (Saltelli et al., 2008). *SSC* are expressed as:

$$S_i^{\sigma} = b_i^{\sigma} = b_i \frac{\sigma_{P_i}}{\sigma_R}$$

(19)

where σ_{pi} is the standard deviation of the ith parameter; σ_R is the standard deviation of the response.

The same concept can be applied to the interaction terms to generate standardized interaction terms:

$$Q_{kl}^{\sigma} = Q_{kl} \frac{\sigma_{P_k} \sigma_{P_l}}{\sigma_R^2}$$

(20)

4. Example 1: Small Synthetic Case

Traditional sensitivity analysis methodologies (the correlation matrix between output and inputs, the tornado chart, and some statistical parameters) are to examine first, and then the proposed methodology is compared to these conventional approaches with a synthetic 'volume of original oil in place' (*OOIP*).

$$OOIP = C_{matrix} \cdot A \cdot T \cdot NTG \cdot \phi_{net} \cdot (1 - S_w) \qquad (21)$$

where C_{matrix} is the correlation matrix of input parameters; A is the area of interest (layer with accumulated oil); T is the thickness of the deposit; NTG is the net to gross ratio; ϕ_{net} is the net porosity; S_w is the water saturation.

The N = 100 values for each input variable are obtained from Monte Carlo Simulation (*MCS*) using predefined parametric distributions of each variable (Table 1). Input correlation between variables is defined in Equation (22). LU decomposition is used to generate correlated values of input random variables (Deutsch and Deutsch, 2009). *OOIP* values are calculated by using Equation (21). The histogram of each variable is presented in Figure 2.

Table 1 – Model input parameters' distribution characteristics.

Input parameter	Definition	Distribution type	Units	Distribution parameters
A	Area	Triangular	m^2	a = 2, b = 4, c = 6
T	Thickness	Gaussian (Normal)	m	μ = 10, σ = 1
NTG	Net to gross ratio	Uniform	-	a = 0.6, b = 0.8
ϕ_{net}	Net porosity	Triangular-like	-	a = 0.15, b = 0.25, c = 0.35, h = 1
S_w	Water saturation	Triangular	-	a = 0.15, b = 0.2, c = 0.3

$$C_{input} = \begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.3 & 0.25 & -0.4 \\ 0.0 & 0.3 & 1.0 & 0.4 & -0.5 \\ 0.0 & 0.25 & 0.4 & 1.0 & -0.6 \\ 0.0 & -0.4 & -0.5 & -0.6 & 1.0 \end{bmatrix} \begin{matrix} \leftarrow & A \\ \leftarrow & T \\ \leftarrow & NTG \\ \leftarrow & \phi_{net} \\ \leftarrow & S_w \end{matrix} \qquad (22)$$

`mcsooip.exe` is used to run *MCS*, which also computes the output *OOIP* and correlation matrix between all generated values. The correlation matrix between the simulated values is shown in Equation (23) and shows reasonable reproduction.

$$C_{output} = \left[\begin{array}{ccccc|c} 1.00 & -0.09 & 0.09 & 0.12 & -0.15 & 0.60 \\ -0.09 & 1.00 & 0.19 & 0.25 & -0.31 & 0.37 \\ 0.09 & 0.19 & 1.00 & 0.45 & -0.42 & 0.57 \\ 0.12 & 0.25 & 0.45 & 1.00 & -0.65 & 0.77 \\ -0.15 & -0.31 & -0.42 & -0.65 & 1.00 & -0.64 \\ \hline 0.60 & 0.37 & 0.57 & 0.77 & -0.64 & 1.00 \end{array} \right] \begin{matrix} \leftarrow & A \\ \leftarrow & T \\ \leftarrow & NTG \\ \leftarrow & \phi_{net} \\ \leftarrow & S_w \\ \leftarrow & OOIP \end{matrix} \qquad (23)$$

The output correlation matrix C_{output} shows which input parameters are more correlated to the response. All variables except S_w are in direct relation with *OOIP* as define in Equation (21). Water saturation S_w and *OOIP* have inverse relationship.

For this example, the order of importance in the input variables is ϕ_{net}, S_w, *A, NTG,* and *T*, determined by considering the absolute value of the correlation coefficients.

In addition to simple ranking of the input variables, conventionally derived tornado chart from `tornado.exe` can be examined (Deutsch et al., 2002). Conventional tornado chart is obtained in a following manner what is different from the proposed approach. To determine the length of the bar on the tornado chart for

each input variable their low and high quantiles of variation are chosen to define consequent variation range of *OOIP*. Single input variable is varied between the desired low and high quantiles while all other inputs are kept at their mean values. The calculations are repeated for each input variable. Then response variations are plotted vertically from the highest variation to the lowest one with input variable name aside caused the variation. Consider the 10% probability (p_{10}) and the 90% probability (p_{90}) limits for the synthetic example (Figure 3). *OOIP* appears to be the most sensitive to the *Area*, and least sensitive to the *Water saturation*. A different order of importance is obtained with this conventional tornado chart: *A*, ϕ_{net}, *T*, *NTG* and S_w. Most notably, S_w is ranked second for linear relationship strength with *OOIP* and, however, it is least influential to the response.

Finally, sensitivity analysis using the regression approach is compared to the results of conventional methods. The executable `sabor.exe` computes 1) the regression (or sensitivity) coefficients, 2) their standardized values for the linear and quadratic models, 3) the interaction terms and 4) their standardized values for the quadratic model using Equations (7) - (20). The program also computes the values of the response *OOIP* obtained from both mathematical models, i.e. regressed values of *OOIP*. Regressed and original values of the response are plotted against each other in a scatter plot to assess the accuracy of the model. The Figure 4 shows scatter plots of both models. The estimation accuracy is assessed through a measure of prediction power, which is a ratio of standard deviations of regressed values of the response to actual ones in percentage (Equation (24)).

$$\text{Prediction power} = \frac{\sigma_{regressed\ response\ values}}{\sigma_{actual\ response\ values}} \cdot 100\% \tag{24}$$

The regression or sensitivity coefficients and their standardized values of the linear and quadratic models derived from the executable `sabor.exe` are tabulated in Table 2. The results are summarized in a bivariate tornado (Figure 5) generated from `bitornado.exe`. The quadratic model produces better estimation results for this example. The standardized interaction terms of the quadratic model are shown in Figure 5 as well. The linear and quadratic models are similar. So, the rank of input parameters from both models is the same. The rank of input variables according to the sensitivity coefficients is ϕ_{net}, *NTG*, S_w, *A*, and *T*. The rank according to the standardized sensitivity coefficients is *A*, ϕ_{net}, *T*, *NTG*, and S_w. The ranks are different, because standardized coefficients comprise the variation of input variables and response beside their relationship strength.

The uniqueness of the ranks of input parameters from conventional and bivariate tornado charts (listed in accordance with *SSC*s) is clear.

Table 2 – Sensitivity coefficients and their standardized values from sabor.exe.

Model Type	Prediction power	Coefficients	A	T	NTG	ϕ_{net}	S_w	OOIP mean value, m³
Linear	96.97%	Sensitivity	1.49	0.53	8.25	23.37	-5.54	5.4940
		Standardized Sensitivity	0.53	0.23	0.22	0.50	-0.08	
Quadratic	99.97%	sensitivity	1.40	0.56	7.74	21.59	-7.02	5.4940
		Standardized sensitivity	0.50	0.24	0.21	0.47	-0.10	

Knowing sensitivity coefficients numerical models can be constructed. For example the linear model of *OOIP* has following form (Equation (25)):

$$OOIP = 5.49 + 1.49 \cdot (A - 3.87) + 0.53 \cdot (T - 9.97) + 8.25 \cdot (NTG - 0.69) +$$
$$+ 23.37 \cdot (\phi_{net} - 0.25) - 5.54 \cdot (S_w - 0.22) \tag{25}$$

Other statistical parameters, such as variance and coefficient of variation, can be used to assess sensitivity. Ranked model inputs using these statistical parameters are shown in Table 3. All rankings discussed are summarized in Table 4.

Table 3 – Variance and specific standard deviation values of each random variable.

Parameter/Response	A	T	NTG	ϕ_{net}	S_w	OOIP	Rank
Variance	0.6492	0.7362	0.0031	0.0021	0.0008	4.8004	T, A, NTG, ϕ_{net}, S_w
Mean	3.8710	9.9749	0.6939	0.2549	0.2185	5.4940	-
Coefficient of variation (CV)	0.2081	0.0860	0.0805	0.1815	0.1334	0.3988	A, ϕ_{net}, S_w, T, NTG

Table 4 – Possible measure tools as a basis for the sensitivity analysis.

Measure	Rank				
	1	2	3	4	5
Correlation matrix	ϕ_{net}	S_w	A	NTG	T
Sensitivity coefficients from linear model	ϕ_{net}	NTG	S_w	A	T
Sensitivity coefficients from quadratic model	ϕ_{net}	NTG	S_w	A	T
Variance	T	A	NTG	ϕ_{net}	S_w
Tornado chart	A	ϕ_{net}	T	NTG	S_w
Standardized sensitivity coefficients from linear model	A	ϕ_{net}	T	NTG	S_w
Standardized sensitivity coefficients from quadratic model	A	ϕ_{net}	T	NTG	S_w
Coefficient of variation (CV)	A	ϕ_{net}	S_w	T	NTG

The ranks obtained using standardized sensitivity coefficients and conventional tornado chart are identical. Thus, *SSC*s describe how the response is sensitive to a certain model parameter. Also, coefficient of variation can be used as a weak measure of output sensitivity on variation of an input. Rank based on variance or the coefficient of variation is interesting but does not relate to sensitivity directly; however, such a ranking does capture the uncertainty of individual parameters and is useful.

Sensitivity analysis combines the importance of input variables relative to the response and the variation ranges of these inputs. It can be seen from combining ranks based on correlation coefficients and variance and comparing with rank based on standardized sensitivity coefficients or conventional tornado chart. Thus, sensitivity of the response on a certain input parameter is a function of their relationship strength and variation ranges (Equation (26)).

$$\text{Sensitivity} = f(\text{relationship strength,input variable \& response variation ranges}) \text{ or } S^{\sigma} = f(S \text{ or } \rho, \sigma^2) \qquad (26)$$

5. Example 2: Real Dataset

The proposed sensitivity analysis using regression approach is applied to a real data set extracted from (Fenik, 2010). The data contains two dynamic response variables, cumulative steam/oil ratio (*CSOR*) and cumulative oil production (*COP*) after 10 years of oil extraction. Seven input static variables are available: connected hydrocarbon volume (*CHV*); volume of original oil in place (*OOIP*); sand permeability (K_{sand}); sand porosity (ϕ_{sand}); shale permeability (K_{shale}); shale porosity (ϕ_{shale}) and; shale fraction (*SF*). Exact relationships between the responses and input parameters are not known. The correlation matrix between all the variables is shown in Figure 6. The regressed values of *CSOR* and *COP* from `sabor.exe` against their actual values are plotted in Figure 7 for linear and quadratic models. The results show that quadratic models have higher estimation accuracy. But caution should be promoted to avoid over fitting. Bivariate tornado charts are shown in Figure 8 and Figure 9. It is evident that *CSOR* mostly depends on shale fraction and volume of oil in place, and the relationships are direct. Other parameters are not so important and can be omitted from future modeling. A larger number of input parameters is required for the *COP* models. Shale fraction, volume of oil in place, shale and sand permeability, and shale porosity

are essential factors in the estimation of *COP*. There is a large uncertainty in the standardized sensitivity coefficients of shale fraction and volume of oil, likely due to their large variation. Shale porosity, sand porosity and connected hydrocarbon volume play much a very small role the responses.

6. Conclusion

The proposed method includes an assessment of uncertainty, which makes it attractive compared to a typical correlation matrix approach. Moreover, the regression approach also simultaneously constructs the linear and quadratic regression models. The standardized regression coefficients are identical to the standardized sensitivity coefficients and explain how the output is sensitive to each input parameter of the model. That is entire sensitivity analysis is based only on evaluation of the sensitivity coefficients and their standardized values, which are the functions of sensitivity coefficients and variable variations.

Thus, standardized sensitivity coefficients are the quantitative measure of the sensitivity of the output on its inputs that should be treated as the results of the sensitivity study based on the regression approach.

References

Deutsch, C.V. and Journel, A.G., 1998, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 2nd Ed., 369 pp.

Deutsch, C.V., Zanon, S. and Leuangthong, O., 2002, Procedures and Guidelines for Assessing and Reporting Uncertainty in Geostatistical Reservoir Modeling, *Centre for Computational Geostatistics* 4, 18-22.

Deutsch, J.L. and Deutsch, C.V., 2009, Latin Hypercube Sampling with Multidimensional Uniformity, *Centre for Computational Geostatistics* 11, 125-2 – 125-4.

Fenik, D.R., 2010, Practical Use of Multiple Geostatistical Realizations in Petroleum Engineering, *M.Sc. Thesis*, University of Alberta, 102 pp.

Helton, J.C., Sallaberry, C.J. and Storlie C.B., 2006, Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis, Sandia National Laboratories, the USA, 23-28.

Saltelli, A.M., Ratto, M. and Andres, T., 2008, *Global Sensitivity Analysis. The Primer*, John Wiley & Sons, Ltd., West Sussex, England, 292 pp.

Figures

$$\bar{y} - \Delta_i y \qquad\qquad \bar{y} \qquad\qquad \bar{y} + \Delta_i y$$

x_1

x_k

x_{n_i-1}
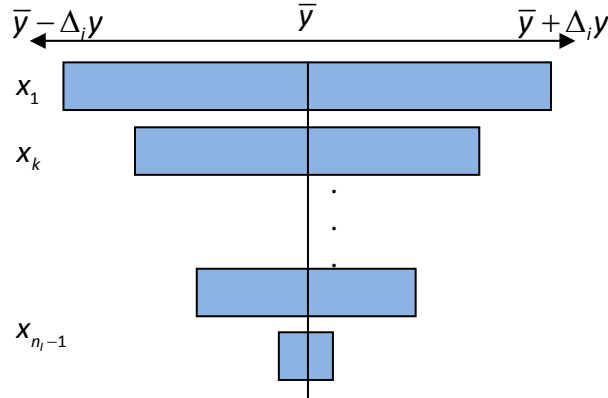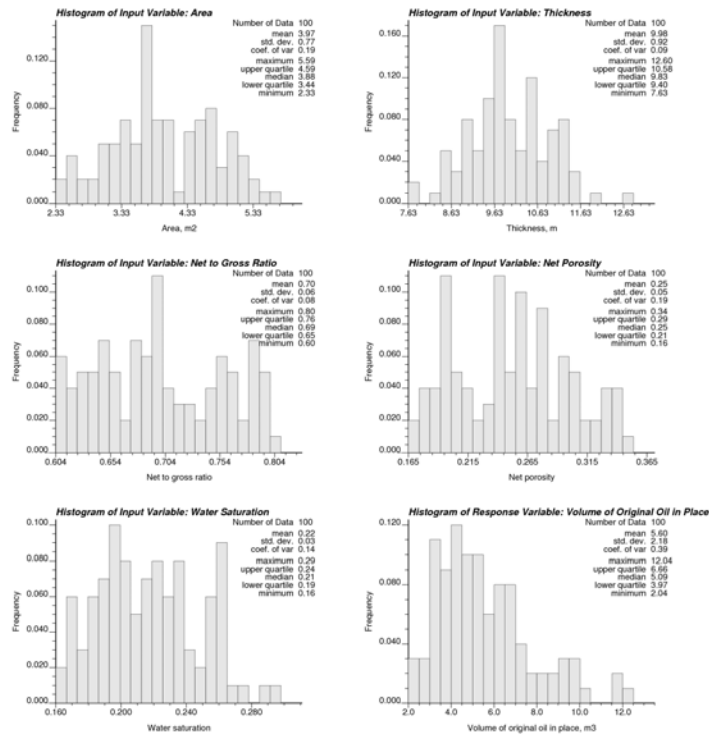
Figure 1 – Schematic tornado chart plot.

Figure 2 – Histograms of input random variables and response of small synthetic example 'OOIP'. Input parameters are: area (A) – triangular distribution, thickness (T) – normal distribution, net to gross ratio (NTG) – uniform distribution, net porosity (Fnet) – triangular-like distribution, and water saturation (Sw) – triangular distribution.

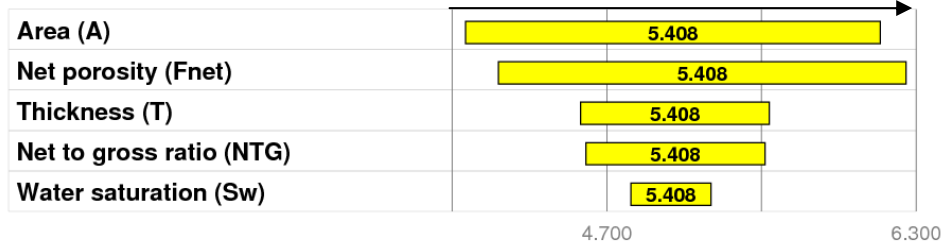|	
Area (A)	5.408
Net porosity (Fnet)	5.408
Thickness (T)	5.408
Net to gross ratio (NTG)	5.408
Water saturation (Sw)	5.408

4.700   6.300

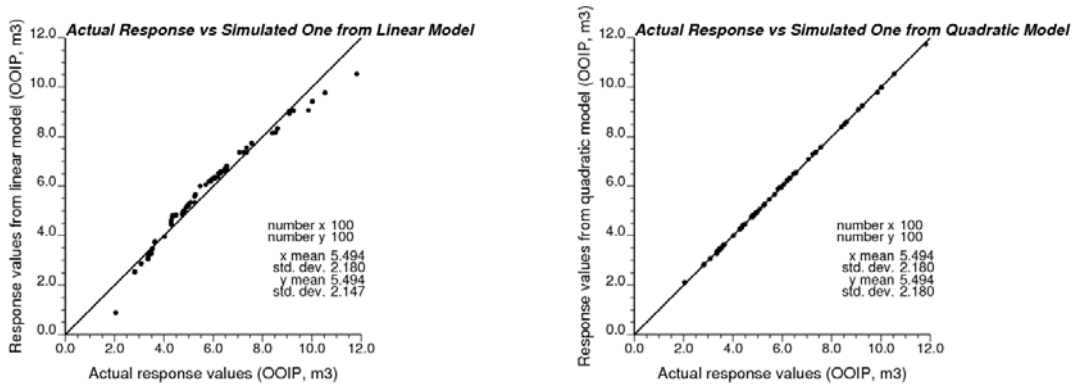**Figure 3** – The tornado chart for output *OOIP* of the simulated Example 1.

Figure 4 – Scatter plot of regressed values of response *OOIP* from mathematical models (from linear model – on the left, from quadratic model – on the right) versus actual response *OOIP* values.
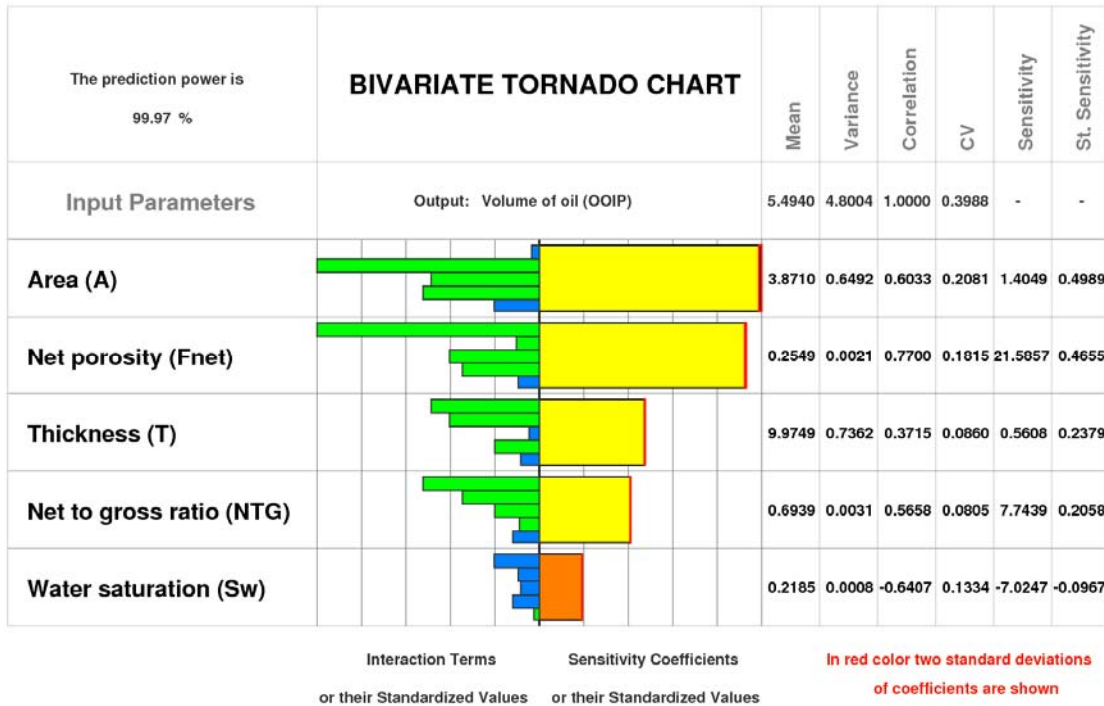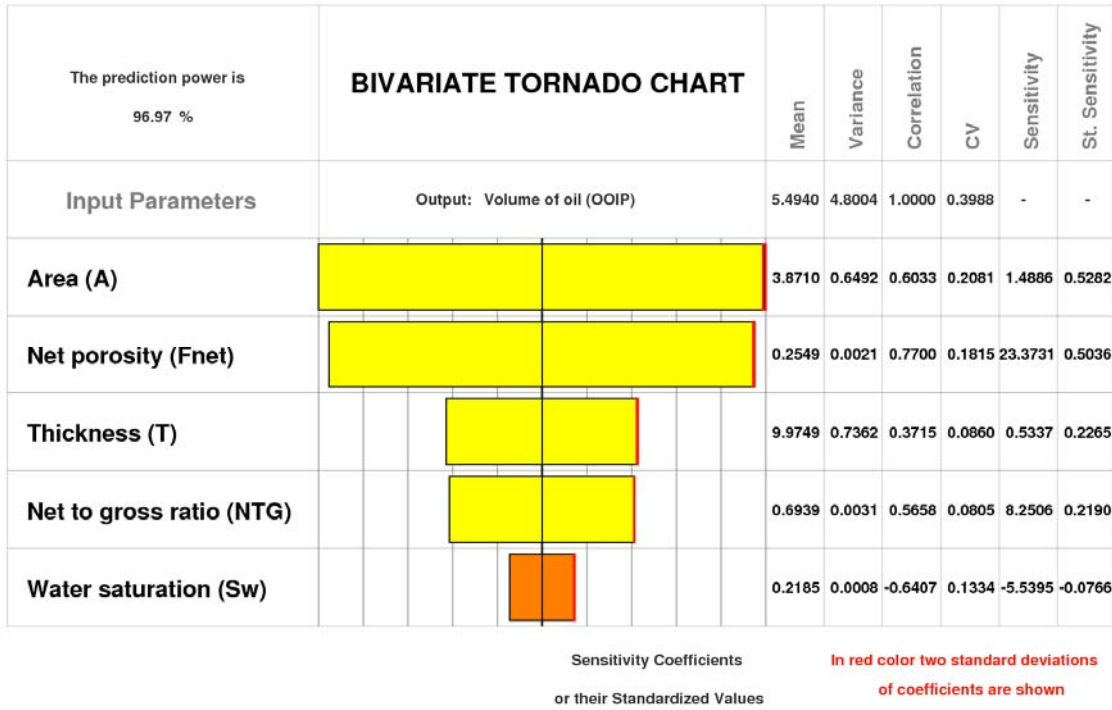
Figure 5 – Bivariate tornado chart obtained from the linear mathematical model (top) and quadratic model (bottom) for response *OOIP* from Example 1. Here the standardized sensitivity coefficients with two-standard-deviation confidence interval as a red box are shown on the right side for both charts (yellow color corresponds to positive value, orange – to negative), the standardized interaction terms of quadratic model are shown on the left side (green – positive, blue - negative). The input parameters are listed in the descending order according to their standardized sensitivity coefficients. Some other statistical parameters of random variables are shown on the right side in table form. The prediction power of a model is in the top-left corner.

$$C = \begin{bmatrix} 1.00 & -0.92 & -0.48 & -0.63 & -0.25 & -0.16 & -0.11 & -0.08 & 0.63 \\ -0.92 & 1.00 & 0.48 & 0.63 & 0.21 & 0.09 & 0.14 & 0.03 & -0.64 \\ -0.48 & 0.48 & 1.00 & 0.76 & 0.10 & 0.04 & -0.07 & 0.01 & -0.75 \\ -0.63 & 0.63 & 0.76 & 1.00 & 0.10 & 0.05 & -0.01 & 0.06 & -0.99 \\ -0.25 & 0.21 & 0.10 & 0.10 & 1.00 & 0.54 & 0.14 & 0.22 & -0.11 \\ -0.16 & 0.09 & 0.04 & 0.05 & 0.54 & 1.00 & 0.05 & 0.04 & -0.02 \\ -0.11 & 0.14 & -0.07 & -0.01 & 0.14 & 0.05 & 1.00 & 0.49 & 0.01 \\ -0.08 & 0.03 & 0.01 & 0.06 & 0.22 & 0.04 & 0.49 & 1.00 & -0.05 \\ 0.63 & -0.64 & -0.75 & -0.99 & -0.11 & -0.02 & 0.01 & -0.05 & 1.00 \end{bmatrix} \begin{matrix} \leftarrow & CSOP \\ \leftarrow & COP \\ \leftarrow & CHV \\ \leftarrow & OOIP \\ \leftarrow & K_{sand} \\ \leftarrow & \phi_{sand} \\ \leftarrow & K_{shale} \\ \leftarrow & \phi_{shale} \\ \leftarrow & Shale\ fraction \end{matrix}$$

Figure 6 – Correlation matrix between variables for Example 2. Seven static input variables: connected hydrocarbon volume, volume of original oil in place, sand permeability, sand porosity, shale permeability, shale porosity, shale fraction, and two dynamical responses: cumulative steam/oil ratio, cumulative oil production.

Figure 7 – Scatter plot of regressed values of response *CSOR* from mathematical models (from linear model – on the top-left, from qu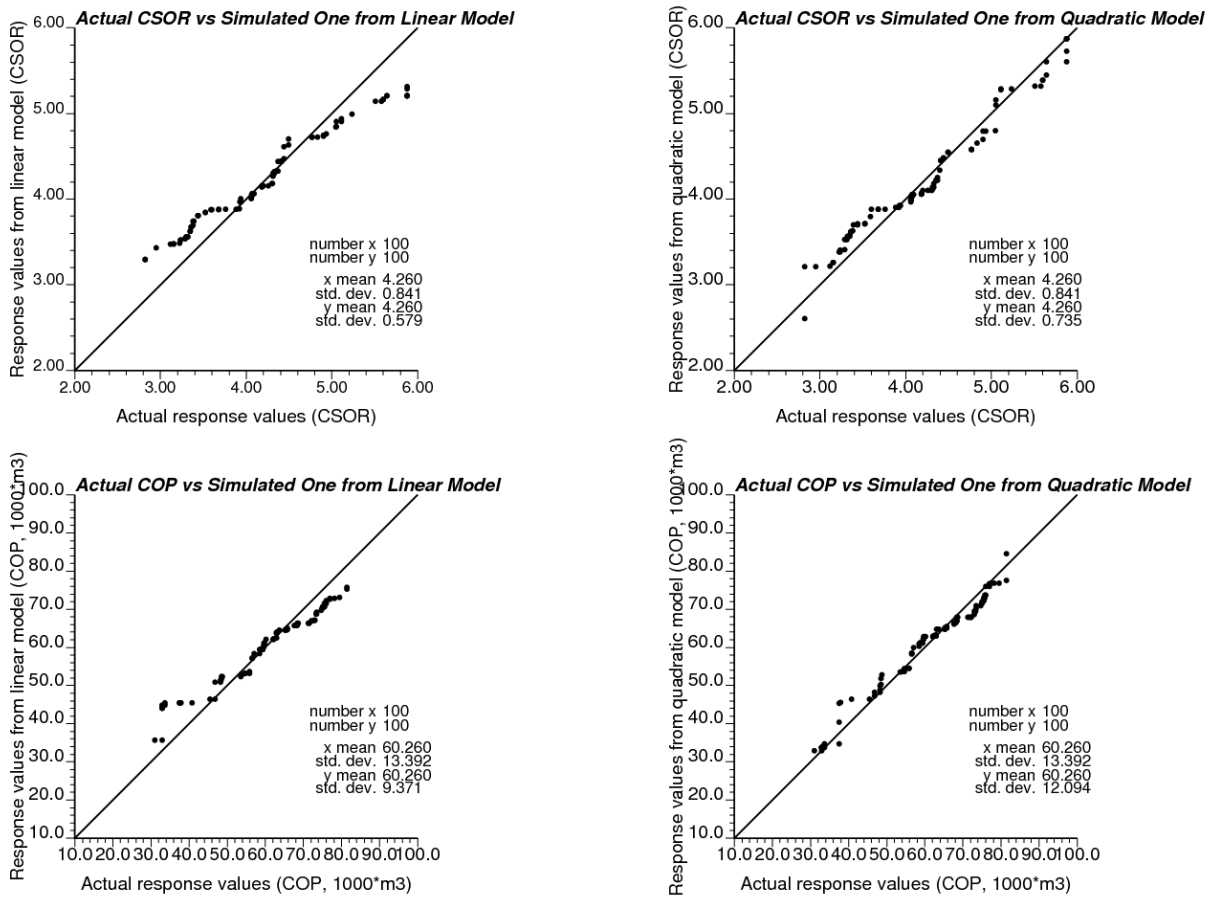adratic model – on the top-right) versus actual response *CSOR* values and scatter plot of regressed values of response *COP* from mathematical models (from linear model – on the bottom-left, from quadratic model – on the bottom-right) versus actual response *COP* values.
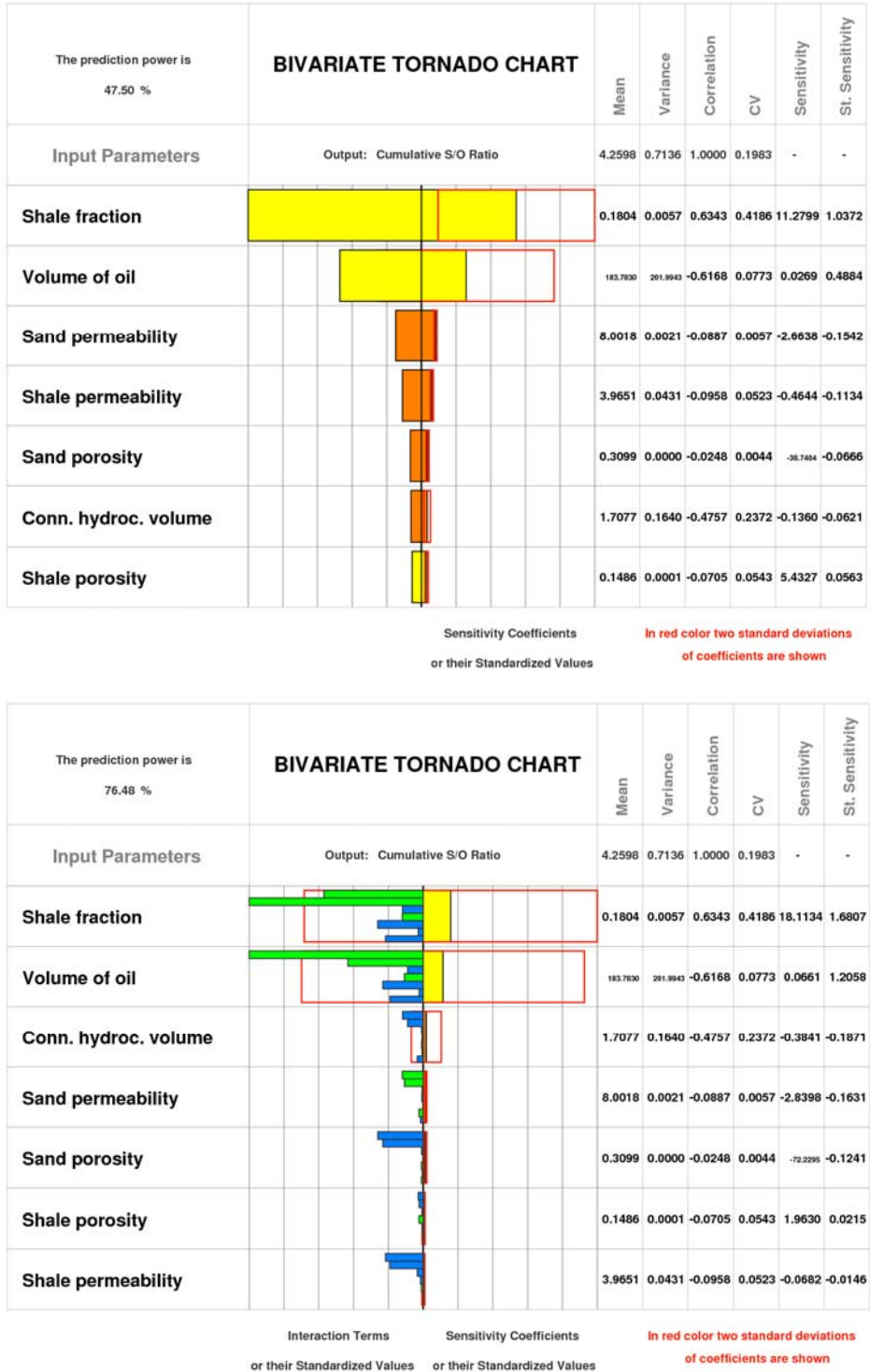
	Mean	Variance	Correlation	CV	Sensitivity	St. Sensitivity
BIVARIATE TORNADO CHART — The prediction power is 47.50 %						
Input Parameters — Output: Cumulative S/O Ratio	4.2598	0.7136	1.0000	0.1983	-	-
Shale fraction	0.1804	0.0057	0.6343	0.4186	11.2799	1.0372
Volume of oil	183.7830	201.9943	-0.6168	0.0773	0.0269	0.4884
Sand permeability	8.0018	0.0021	-0.0887	0.0057	-2.6638	-0.1542
Shale permeability	3.9651	0.0431	-0.0958	0.0523	-0.4644	-0.1134
Sand porosity	0.3099	0.0000	-0.0248	0.0044	-38.7484	-0.0666
Conn. hydroc. volume	1.7077	0.1640	-0.4757	0.2372	-0.1360	-0.0621
Shale porosity	0.1486	0.0001	-0.0705	0.0543	5.4327	0.0563

Sensitivity Coefficients or their Standardized Values

In red color two standard deviations of coefficients are shown

	Mean	Variance	Correlation	CV	Sensitivity	St. Sensitivity
BIVARIATE TORNADO CHART — The prediction power is 76.48 %						
Input Parameters — Output: Cumulative S/O Ratio	4.2598	0.7136	1.0000	0.1983	-	-
Shale fraction	0.1804	0.0057	0.6343	0.4186	18.1134	1.6807
Volume of oil	183.7830	201.9943	-0.6168	0.0773	0.0661	1.2058
Conn. hydroc. volume	1.7077	0.1640	-0.4757	0.2372	-0.3841	-0.1871
Sand permeability	8.0018	0.0021	-0.0887	0.0057	-2.8398	-0.1631
Sand porosity	0.3099	0.0000	-0.0248	0.0044	-72.2295	-0.1241
Shale porosity	0.1486	0.0001	-0.0705	0.0543	1.9630	0.0215
Shale permeability	3.9651	0.0431	-0.0958	0.0523	-0.0682	-0.0146

Interaction Terms or their Standardized Values

Sensitivity Coefficients or their Standardized Values

In red color two standard deviations of coefficients are shown

Figure 8 – Bivariate tornado chart obtained from the linear mathematical model (top) and quadratic model (bottom) for response *CSOR* from Example 2. Here the standardized sensitivity coefficients with two-standard-deviation confidence interval as a red box are shown on the right side for both charts (yellow color corresponds to positive value, orange – to negative), the standardized interaction terms of quadratic model are shown on the left side (green – positive, blue - negative). The input parameters are listed in the descending order according to their standardized sensitivity coefficients. Some other statistical parameters of random variables are shown on the right side in table form. The prediction power of a model is in the top-left corner.
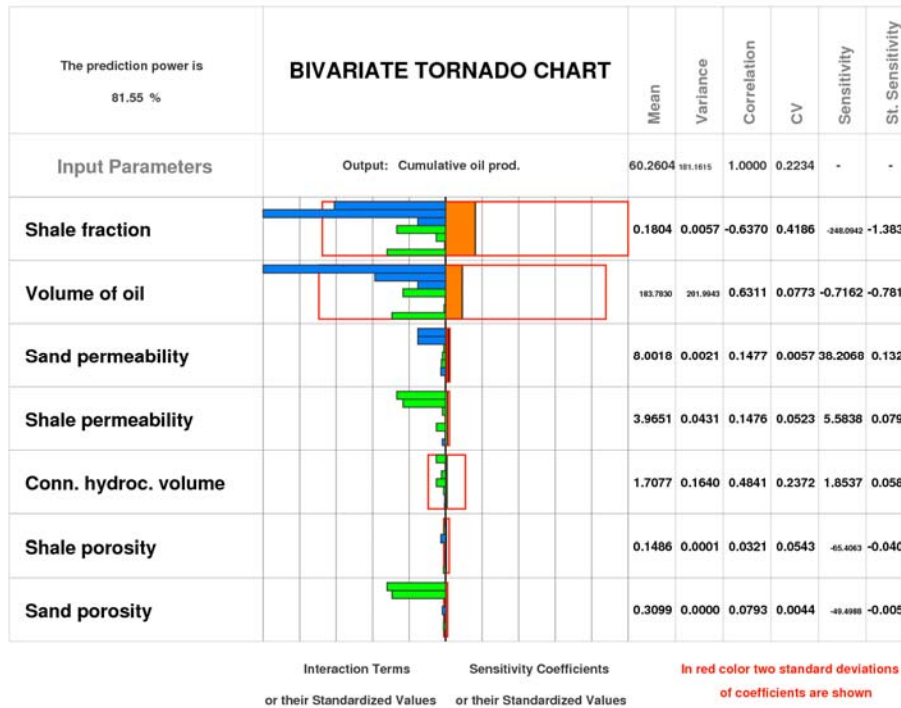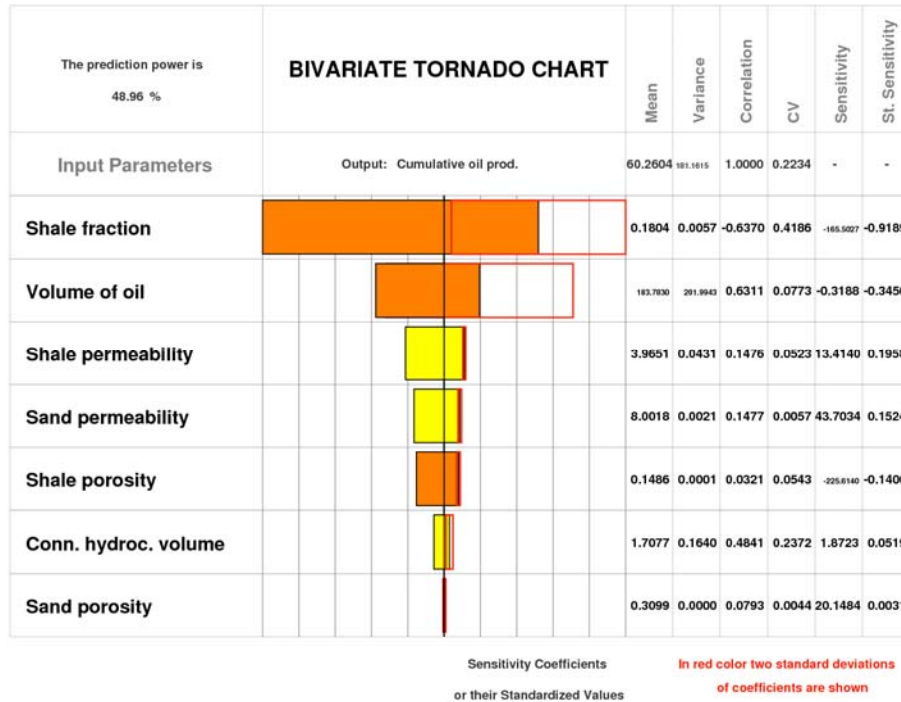
Figure 9 – Bivariate tornado chart obtained from the linear mathematical model (top) and quadratic model (bottom) for response *COP* from Example 2. Here the standardized sensitivity coefficients with two-standard-deviation confidence interval as a red box are shown on the right side for both charts (yellow color corresponds to positive value, orange – to negative), the standardized interaction terms of quadratic model are shown on the left side (green – positive, blue - negative). The input parameters are listed in the descending order according to their standardized sensitivity coefficients. Some other statistical parameters of random variables are shown on the right side in table form. The prediction power of a model is in the top-left corner.