Article

Some methods for sensitivity analysis of systems / networks

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China; International Academy of Ecology and Environmental Sciences, Hong Kong
E-mail: zhwj@mail.sysu.edu.cn,wjzhang@iaees.org

Received 5 January 2016; *Accepted 3 March 2016*; *Published online 1 September 2016*

Abstract
A network may considerably change with certain nodes, links, flows, or parameters. To find the most important nodes, links, or other parameters to determine network structure or performance is of significant. Sensitivity analysis is originated from systems science. It explores the relationship between parametric change and systematic output, and is used to find important parameters in the system model. In principle, the sensitivity analysis used in systems science can also be extended to network analysis in which the model output means network output, network stability, network flow, network structure, or other indices, and model input means network nodes, network links, network parameters, etc. In present article, some methods for sensitivity analysis of systems / networks are described in detail.

Keywords sensitivity analysis; systems; networks; methodology.

Network Pharmacology
ISSN 2415-1084
URL: http://www.iaees.org/publications/journals/np/online-version.asp
RSS: http://www.iaees.org/publications/journals/np/rss.xml
E-mail: networkpharmacology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Sensitivity analysis explores the relationships of, e.g., structure *vs*. stability, input *vs*. output, of various systems. It helps us understand the structural properties, dynamic mechanisms, key components (or nodes, links, parameters, etc.) and other aspects of systems. A variety of methods for sensitivity analysis of system models have been proposed (McKay et al., 1979; Downing et al., 1985; Morris, 1991; Saltelli et al., 1999, 2000; Sobol, 1993; Xu et al., 2004). Networks are systems also. Like general systems, some networks can be described with mathematical models (Ferrarini, 2013, 2014, 2015). Many networks evolve naturally following some rules (Zhang, 2012a, 2015, 2016a, 2016c). To find the most important nodes, links, or other parameters that determine network structure or performance is of significant in network analysis. In present article, I will describe some methods for sensitivity analysis of systems / networks based on past studies.

2 Methods

2.1 Sensitivity analysis for systems analysis

There are two types of sensitivity analyses in systems analysis, local and global sensitivity analysis (Saltelli, 2000). In the sensitivity analysis, a system is represented by a mathematical model.

2.1.1 Local sensitivity analysis

Local sensitivity analysis (LSA) is mainly used to analyze local influence of parameters on model output. Gradients of parameters *vs.* model output can be achieved by using LSA. LSA is valuable to systems with simple mathematical representations, fewer undeterministic parameters and easily derived sensitivity equation. LSA methods include directed derivation, finite difference, and Green function.

2.1.1.1 Directed derivation

For models with less independent variables and simple structure, directed derivation is a simple and rapid method for sensitivity analysis. Suppose the initial-value problem of a system is (Han et al., 2008)

$$dy/dt=f(y, x),\ y(0)=y_0 \quad (1)$$

where $y=(y_1, y_2,\ldots, y_n)$ is output vector, $x=(x_1, x_2,\ldots, x_m)$ is the input vector, t is time. The sensitivity equation is

$$d(\partial y/\partial x_i)/dt=A\ \partial y/\partial x_i+\ \partial f/\partial x_i \quad (2)$$

or

$$dS/dt=AS+C$$

where S is the sensitivity matrix, $A=(\partial f_j/\partial y_l)$ is the Jacobi matrix, $C=(\partial f_j/\partial x_i)$ is the parametrical Jacobi matrix. Assume the network is time-invariant, i.e., $f(y, x)=0$. y can be obtained by solving the equation, $f(y, x)=0$. The static sensitivity matrix S with respect to x_i is achieved as

$$S=-A^{-1}C$$

2.1.1.2 Finite difference

In finite difference method, a perturbation of an input x_j, Δx_j, is made to obtain derivative of output to x_j. The forward difference scheme is usually used (Han et al., 2008)

$$\partial y/\partial x_i\approx(y(x^j)-y(x))/\Delta x_i$$
$$j=1, 2,\ldots, m$$

where $x^j=(x_1, x_2,\cdots, x_{j-1}, x_j+\Delta x_j, x_{j+1},\cdots, x_m)$. The more precise scheme, central-difference scheme is used also

$$\partial y/\partial x_i\approx(y(x^{j+})-y(x^{j-}))/(2\Delta x_i) \quad (3)$$
$$j=1, 2,\ldots, m$$

where $x^{j+}=(x_1, x_2,\cdots, x_{j-1}, x_j+\Delta x_j, x_{j+1},\cdots, x_m)$, $x^{j-}=(x_1, x_2,\cdots, x_{j-1}, x_j-\Delta x_j, x_{j+1},\cdots, x_m)$.

2.1.1.3 Green function

The differential equation of Equation (1) with respect to the initial value y_0 is (Han et al., 2008)

$$dS(t,t_1)/dt=A(t)\ S(t,t_1) \quad (4)$$

where t, t_1: perturbation time and observation time; $x(t,t_1)$: sensitivity matrix. $S(t,t_1)=(\partial c_i(t)/\partial c_j^0(t_1))$, $S(t_1,t_1)=1$,

$t \geq t_1$. The solution of sensitivity matrix is made of two parts, the general solution of homogeneous equation of equation (2), and the special solution of nonhomogeneous equation (2). The general solution of homogeneous equation of equation is obtained by using equation (3). And the special solution of nonhomogeneous equation is

$$T(t_1,t_2)=\int_{t1}^{t2} S(t_2,s)\, F(s)ds$$

Local sensitivity analysis (LSA) is used to test model sensitivity to the change of a single specified parameter only, and the remaining parameters are fixed. LSA does not consider the influence of interactions between model parameters on model output.

2.1.2 Global sensitivity analysis (GSA)

Global sensitivity analysis (GSA) tests the joint influence of multiple parameters on model output, and analyze model sensitivity of both a single parameter and between-parameter interactions. It can be used in nonlinear, non-overlapping, or non-monotonous models.

2.1.2.1 Qualitative GSA

With less calculation, Qualitative GSA is used to rank parameters according to their sensitivities.

(1) Multivariable regression method

In this method, Latin hypercube sampling is used (McKay et al., 1979; Downing et al., 1985). It divides the cumulative probability distribution (i.e., the interval (0, 1)) of a parameter into multiple non-overlapping intervals of equal length along *y*-axis, and thus more effective than random sampling method. Each interval corresponds to an interval in *x*-axis. Randomly sampling a point in an interval of *y*-axis, we can obtain a corresponding parametrical value in an interval of *x*-axis. For the model with *n* parameters, each parameter has a cumulative probability distribution and *m* intervals are generated. In total of n^m sampling combinations are thus produced.

The procedure is that, the arrange values of *n* parameters into a matrix of $n \times m$, and randomize elements of each column, the m input of *n* parameters are thus obtained. Each row of the matrix can be taken as the input values of a parameter, and use them as model input to get model output. Finally, build multivariable regression between model outputs and inputs and use regression coefficients or partial correlation coefficients as the sensitivity values of corresponding parameters. Stepwise regression is more reliable in this method (Qi et al., 2016; Zhang, 2016b).

(2) Morris method

Morris method (Morris, 1991; Xu et al., 2004) maps the domain of each parameter into [0, 1] and discretizes the interval, such that each parameter takes values in $\{0,1/(p-1),2/(p-1),...,1\}$, where *p* is the number of sampling points of the parameter. Each parameter takes values randomly from *p* sampling points and obtains a vector $X=(x_1, x_2,\ldots, x_k)$, where *k* is the number of parameters.

Consider a matrix, $B=(b_{ij})_{(k+1)\times k}$, ($b_{ij}=1$, if $i<j$; $b_{kk}=0$, $b_{kj}=1$, if $j<k$; $b_{k+1\times j}=1$; otherwise, $b_{ij}=0$) and the change $\Delta=s/(p-1)$, where *s* is the constant. Use the adjacent two rows in the matrix ΔB as the model input and obtain two outputs y_1 and y_2. The sensitivity of parameter *i* can be calculated by $\Delta_i(X)=(y_1-y_2)/\Delta$. Take *k* pairs of adjacent rows as mode inputs, we may obtain sensitivities of *k* parameters (Xu et al., 2004).

Values of elements in ΔB are not stochastic. Thus in practical application, we can use a stochastic process to ensure the stochasticity of elements' values (Xu et al., 2004) as the following: suppose $D^*_{k\times k}$ is a diagonal matrix, in which any element has the same chance as 1 and -1. Suppose $J_{m\times k}$ is the unit matrix. Each column of elements of matrix $C_{m\times k}=(1/2)[(2B-J_{m\times k})D^*+J_{m\times k}]$ are equal to corresponding elements in *B* (or 1 change to 0, or 0 change to 1 in *B*); suppose X^* is the base vector of *X*. Each parameter in *X* randomly takes values in

{0,1/(p-1),2/(p-1),...,1}, and suppose $P^*_{k\times k}$ is the random permutation matrix in which each row and each column contains one 1 only and others are zeros. Let $B^*=(J_{m\times 1}X^*+(\Delta/2)[(2B-J_{m\times k})D^*+J_{m\times k}])P^*$, B^* is the stochastic matrix of B. Use the adjacent two rows in the matrix B^* as the model input and obtain two outputs y_1 and y_2. The sensitivity of parameter i can be calculated by $\Delta_i(X)=(y_1-y_2)/\Delta$. Replications can be made to obtain the mean sensitivity and its standard deviation of each parameter. A less standard deviation means a little interaction between the parameter and other parameters.

Morris method can be used to freeze less sensitive parameters and further made to make quantitative GSA.

2.1.2.2 Quatitative GSA

Quantitative GSA can quantitatively calculate the contribution proportion of uncertainty of model parameters to uncertainty of model output. It holds that variance of model output represents the uncertainty of model output. Before making quatitative GSA, qualitative GSA is always used to filter the less sensitive parameters from the model.

(1) Sobol method

A representative method of quatitative GSA is Sobol method (Sobol, 1993; Xu et al., 2004; Han et al., 2008). Sobol method is a Mont Carlo method based on variance.

Suppose the model is $y=f(x)$, $x=(x_1, x_2,\ldots, x_k)$. x_i follows uniform distribution of [0,1], and $(f(x))^2$ is integrable. First, define a k-dimensional unit cube as the space domain of parameters, $\Omega=\{x|0\le x_i\le 1;i=1,2,\ldots,k\}$. The core procedure of Sobol method is to decompose the function $f(x)$ into the sum of sub-terms

$$f(x)= f_0+\sum_{i=1}^{k} f_i(x_i)+ \sum_{i<j} f_{ij}(x_i,x_j)+\ldots+f_{1,2,\ldots,k}(x_1, x_2,\ldots, x_k) \qquad (5)$$

The decomposition above is not unique. If

$$\int_0^1 f_i(x_i)dx_i=0,\ \forall x_i,\ i=1,2,\ldots,k$$
$$\int_0^1\int_0^1 f_{ij}(x_i,x_j)dx_idx_j=0,\ \forall x_i,x_j,\ i<j$$
$$\int_\Omega f_{1,2,\ldots,k}(x_1, x_2,\ldots, x_k)dx_1dx_2\ldots dx_k=0$$

the decomposition (5) is unique, and

$$\int_\Omega f_{i1,i2,\ldots,is}f_{j1,j2,\ldots,jl}dx=0,\ (i_1,i_2,\ldots,i_s)\ne(j_1,j_2,\ldots,j_l),\ k=s+l$$

Calculate

$$f_0=\int_\Omega f(x)dx$$
$$f_i(x_i)=-f_0+\int_0^1\ldots\int_0^1 f(x)dx_{-i},\ i=1,2,\ldots,k$$
$$f_{ij}(x_i,x_j)=-f_0-f_i(x_i)-f_j(x_j)+\int_0^1\ldots\int_0^1 f(x)dx_{-ij},\ i<j$$

where x_{-i} and x_{-ij} represent the parameters in exception of x_i and x_i together with x_j, respectively. Similarly, we can get $f_{1,2,\ldots,k}(x_1, x_2,\ldots, x_k)$. The total variance of model output, i.e., the effect of all parameters on model output, is

$$V=\int_\Omega (f(x))^2dx-(f_0)^2$$

The effect of a single parameter on model output is represented by its partial variance

$$V_i=\int(f_i)^2dx_i$$

And the joint effect of several parameters on model output is represented by the partial variance

$$V_{i1,i2,\ldots,is}=\int_0^1\ldots\int_0^1(f_{i1,i2,\ldots,is})^2dx_{i1}dx_{i2}...dx_{is}$$

Square equation (5) and thereafter make integration, we have

$$V=\sum_{i=1}^{k}V_i+\sum_{i<j}V_{ij}+\ldots+V_{1,2,\ldots,k}$$

and

$$S_{i1,i2,\ldots,is}=V_{i1,i2,\ldots,is}/V \qquad (6)$$

where S_i is the sensitivity of parameter i, known as 1st sensitivity; $S_{i1,i2,\ldots,is}$ is the joint sensitivity of interactive parameters $i_1,i_2,\ldots$, and i_s, known as sth sensitivity.

The integrals above can be solved using Monte Carlo method, i.e.

$$f_0=\sum_{i=1}^{n}f(x_i)/n$$

$$V=\sum_{i=1}^{n}(f(x_i))^2/n-(f_0)^2$$

$$V_i=\sum_{j=1}^{n}f(x_{ij}^{(1)},x_{-ij}^{(1)})\ f(x_{ij}^{(1)},x_{-ij}^{(2)})/n-(f_0)^2$$

…

(2) Extended Fourier Amplitude Sensitivity Test (EFAST)

Extended Fourier Amplitude Sensitivity Test was proposed by Saltelli et al. (1999). In this method, the spectrum of Fourier series is obtained by Fourier transformation. The variances of model output aroused from a single parameter and parameter interactions are calculated from frequency spectrum curve (Tarantola et al., 2002;Xu et al., 2004).

The model $y=f(x)$, $x=(x_1, x_2,\ldots, x_k)$. can be transformed to $y=f(s)$ using a certain search function. The Fourier transformation is

$$y=f(s)=\sum_{j=-\infty}^{\infty}(A_j\cos(js)+B_j\sin(js))$$

where s is the independent parameter of all parameters, and

$$A_j=(1/(2\pi))\int_{-\pi}^{\pi}f(s)\cos(js)ds$$
$$B_j=(1/(2\pi))\int_{-\pi}^{\pi}f(s)\sin(js)ds$$
$$j\in Z=\{-\infty,\ldots,-1,0,1,\ldots,\infty\}$$

Frequency spectrum curve of Fourier series is $\wedge_j=A_j+B_j$, where $A_{-j}=A_j$, $B_{-j}=B_j$, $\wedge_{-j}=\wedge_j$. The variance of model

output aroused from uncertainty of x_i is

$$V_i=2\sum_{p=1}^{\infty}\wedge_p w_i \quad (7)$$

where w_i is a specified frequency. The total variance is

$$V=2\sum_{p=1}^{\infty}\wedge_p \quad (8)$$

Take s in the $[-\pi,\pi]$ with the same interval, and input each sampled parameter to the model. Run the model many times, and A_j and B_j can be obtained from

$$A_j=\sum_{Sk=1}^{Ns} f(s_k)\cos(js_k)$$

$$B_j=\sum_{Sk=1}^{Ns} f(s_k)\sin(js_k)$$

$$j\in Z'=\{-(N_s-1)/2,\ldots,-1,0.,1,\ldots,(N_s-1)/2\}$$

where N_s is the number of samples, $N_s=2Mw_{max}+1$. Finally, based on equation (6)-(8), we can get sensitivity of each parameter.

EFAST costs much lower than Sobol. Both EFAST and Sobol require parameters are irrelevant from each other.

2.2 Sensitivity analysis for networks

The above methods can be used in network analysis also. In network analysis, the model output in above methods means network output, network stability, network flow, network structure, or other indices we specified, and model input means network nodes, network links, network parameters, etc.

Zhang (2012a, b) proposed a series of methods for determination of crucial nodes, which can be used in sensitivity analysis also, for example, Node Perturbation index (NP) (Zhang, 2012b)

$$NP=dN/dn/N$$

or

$$NP=dN/dn$$

where N: measure of network structure; n: state value or proportion of a known node in the network. There are many measures of network structure, i.e., total links, total number of nodes, network flow (Latham, 2006), degree distribution (Zhang, 2011; Zhang and Zhan, 2011), aggregation index, coefficient of variation, entropy (Zhang and Zhan, 2011; Zhang, 2012a), and other measures (Paine, 1992; Power et al., 1996; Dunne et al., 2002; Montoya and Sole, 2003; Allesina et al., 2005; Barabasi, 2009). More methods that can be used in sensitivity analysis of networks, e.g., adjacency matrix index, flow change index, etc., have been discussed in Zhang (2012b) and other references.

References

Allesina S, Bodini A, Pascual M. 2005. Functional links and robustness in food webs. Philosophical Transactions of the Royal Society B, 364(1524): 1701-1709

Barabasi AL. 2009. Scale-free networks: a decade and beyond. Science, 325: 412-413

Downing DJ, Gardner RH , Hoffman FO. 1985. An examination of response-surface methodologies for uncertainty analysis in assessment models. Technometrics, 27: 151-163

Dunne JA, Williams RJ, Martinez ND. 2002. Food-web structure and network theory: the role of connectance and size. Ecology, 99(20): 12917-12922

Ferrarini A. 2013. Exogenous control of biological and ecological systems through evolutionary modelling. Proceedings of the International Academy of Ecology and Environmental Sciences, 3(3): 257-265

Ferrarini A. 2014. Local and global control of ecological and biological networks. Network Biology, 4(1): 21-30

Ferrarini A. 2015. Evolutionary network control also holds for nonlinear networks: Ruling the Lotka-Volterra model. Network Biology, 5(1): 34-42

Han LS, Li XY, Yan DK. 2008. A discussion on several mathematical methods of sensitivity analysis. China Water Transport, 8(4): 177-178

Latham LG. 2006. Network flow analysis algorithms. Ecological Modelling, 192: 586-600

McKay MD, Beckman RJ, Conover WJ. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21: 239-245

Montoya JM, Sole RV. 2003. Topological properties of food webs: from real data to community assembly models. Oikos, 102: 614-622

Morris MD. 1991. Factorial sampling plans for preliminary computational experiments. Technometrics, 33: 161-174

Paine RT. 1992. Food-web analysis through field measurement of per capita interaction strength. Nature, 355: 73-75

Power ME, Tilman D, Estes JA, et al. 1996. Challenges in the quest for keys. Bioscience, 46: 609-620

Qi YH, Liu GH, Zhang WJ. 2016. A Matlab program for stepwise regression. Network Pharmacology, 1(1): 36-41

Saltelli A, Chan K, Scott M .2000. Sensitivit y Analysis, Probability and Statistics Series. John Wiley, New York, USA

Saltelli A, Tarantola S, Chan KPS. 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics, 41: 39-56

Sobol IM. 1993. Sensitivity Estimates for nonlinear mathematical models. Mathematical Model and Computational Experiment, 1: 407-414

Tarantola S, Giglioli N, Jesinghaus J, et al. 2002 .Can global sensitivity analysis steer the implementation of models for environmental assessments and decision-making? Stochastic Environmental Research and Risk Assessment, 16: 63-76

Xu CG, Hu YM, Chang Y, et al. 2004. Sensitivity analysis in ecological modeling. Chinese Journal of Applied Ecology, 15(6): 1056-1062

Zhang WJ. 2011. Constructing ecological interaction networks by correlation analysis: hints from community sampling. Network Biology, 1(2): 81-98

Zhang WJ. 2012a. Computational Ecology: Graphs, Networks and Agent-based Modeling. World Scientific, Singapore

Zhang WJ. 2012b. Several mathematical methods for identifying crucial nodes in networks. Network Biology,

2(4): 121-126

Zhang WJ. 2015. A generalized network evolution model and self-organization theory on community assembly. Selforganizology, 2(3): 55-64

Zhang WJ. 2016a. A random network based, node attraction facilitated network evolution method. Selforganizology, 3(1): 1-9

Zhang WJ. 2016b. Screening node attributes that significantly influence node centrality in the network. Selforganizology, 3(3): 75-86

Zhang WJ. 2016c. Selforganizology: The Science of Self-Organization. World Scientific, Singapore

Zhang WJ, Zhan CY. 2011. An algorithm for calculation of degree distribution and detection of network type: with application in food webs. Network Biology, 1(3-4): 159-170