Use of Sensitivity Analysis to Assess Reliability of Mathematical Models

Michael C. Kohn

Laboratory of Computational Biology and Risk Analysis

National Institute of Environmental Health Sciences

PO Box 12233, MD A3-06

Research Triangle Park, NC 27709

In many fields of science and engineering, it is often necessary to predict the effect of some treatment or intervention on a human subject.. It is unethical to identify health or safety effects by experiments on humans. Therefore, the biological responses to a given treatment are represented by a mathematical model, which is usually based on data from animal experiments. The parameter values for animals can be replaced by values for humans. The model's behavior under various imposed boundary conditions predicts the human response to an actual treatment. All models are simplifications—often idealizations—of the real system, and excessive simplification may reduce the model's reliability. Predictions may also be unreliable because of restricted accuracy of the data on which the model is based. One method by which the reliability of the predictions can be assessed is sensitivity analysis.

Definitions

A mathematical model comprises equations for the instantaneous state of a system given the imposed conditions. This may be a function of time. Terms used in modeling follow.

state variable—the quantities that specify the instantaneous description (state) of the system

independent variable—the dimension (typically time) over which the state of the system changes

state equations—equations that specify the state variables as functions of the independent variable (typically differential equations in a dynamic model)

boundary conditions—constraints that specify the particular solution of the state equations; *e.g.* initial values of the state variables or the rate of input to a process in the model

transient—the temporal profile of the state variables after a perturbation in the boundary conditions

steady state—a solution of the state equations when the time derivatives of the state variables are all set to zero; also called a **stationary state**

parameters—constants (as opposed to variables) in the state equations

sensitivity coefficient—a partial derivative of a state variable with respect to variations in a parameter value; these quantities may vary with time

robustness—invariance of a model's predictions with respect to small variations in its parameter values

This paper will focus on dynamic systems, their transient responses, and their steady-state solutions. Arrays will be denoted by boldface type, lower case for vectors and upper case for matrices.

System Sensitivity Theory

Early engineering models described the output $\mathbf{y}(t)$ of a system (*e.g.* an electrical current) as a "transfer" function of the input signal $\mathbf{u}(t)$ (*e.g.* a periodic voltage). The earliest application of sensitivity analysis was to predict the effect on output of variations of input as a function of frequency [1]. In the 1960s, sensitivity theory was applied to the time domain [2]. Denoting the state variables by $\mathbf{x}(t)$, the state functions are

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}, t)$$
$$\mathbf{x}(t_0) = \mathbf{x}_0$$

where $\boldsymbol{\alpha}$ is a vector of parameters with nominal values $\boldsymbol{\alpha}_0$ and t_0 is the initial time. Because of measurement and approximation errors, the true parameter values may differ from the nominal values by $\Delta\boldsymbol{\alpha}$. The resulting uncertainty in the computed state variables can be estimated by

$$\Delta\mathbf{x} = \int \left[\mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}_0 + \Delta\boldsymbol{\alpha}, t) - \mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}_0, t)\right] dt$$

The mapping of the space of parameters into the state space is uniquely determined by the functions $\mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}, t)$, but this method is computationally extremely demanding for more than a few parameters [3]. In some cases, the errors introduced by numerical integration can introduce unacceptable inaccuracies in the solution [3].

This problem is usually alleviated by defining a sensitivity function $\mathbf{S}$ that relates $\Delta\boldsymbol{\alpha}$. to $\Delta\mathbf{x}$.

$$S_{i,j} = \frac{\partial x_i}{\partial \alpha_j}$$

$S_{i,j}$ is a sensitivity coefficient. Under certain conditions $\Delta\mathbf{x}$ can be expanded about the nominal solution in a Taylor series and truncated after the first-order terms [3]. This is exact if the state

functions are linear and autonomous (not explicit functions of *t*). Otherwise, it is a reasonable approximation for small values of $\Delta\boldsymbol{\alpha}$.. The linearized sensitivity equations are

$$\dot{\mathbf{S}} = \mathbf{JS} + \frac{\partial \mathbf{f}}{\partial \boldsymbol{\alpha}}$$

$$J_{i,j} = \frac{\partial f_i}{\partial x_j}$$

where the matrix **J** is the "Jacobian" of the system. The partial derivatives in the above equations can rarely be calculated analytically, but they can be computed by finite differences with a sufficiently small $\Delta\boldsymbol{\alpha}$.

The differential equations for the sensitivity coefficients can be solved along with the state equations. The resulting uncertainty in the integrated values of the state variable can then be approximated by

$$\Delta\mathbf{x} = \mathbf{S}\Delta\boldsymbol{\alpha}$$

where $\Delta\mathbf{x}$ is called the "supplementary motion" of the system. This approximation is valid as long as $\Delta\boldsymbol{\alpha}$.is sufficiently small. However, for systems in which there are strong interactions among the variables, an impractically small increment $\Delta\boldsymbol{\alpha}$.would be required and unacceptable round-off error would accumulate. For steady states, the time derivatives of the sensitivity coefficients are zero, and the differential equations reduce to algebraic equations.

The reliability of a mathematical model is reflected in the numerical values of the sensitivity coefficients. A model must be sensitive to large (relative to typical experimental error) changes in parameter values. Otherwise, a wide range of values will produce substantially the same behavior and it will not be possible to verify that correct parameter values have been used in the simulation. Thus, the structure of the model will be suspect. A model must be robust with respect to small (relative to typical experimental error) uncertainties in parameter values. Otherwise, small errors in the parameter values will produce large supplementary motions, and the model will not be testable. Thus, the predictions of the model will not be reliable. These inferences hold even when the parameter value cannot be measured.

Sensitivity Analysis of Empirical Models

The functions $\mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}, t)$ for the time derivatives of the state variables are often unknown or too complex to be represented in detail. Thus, the available data may be inadequate to identify all the parameter values, and an approximate function must be obtained. One approximation that has been proposed [4] is an S-system ('S' for synergistic) in which the function is represented by empirical power laws. The derivative function is divided into two terms for processes that increase and decrease the state variable.

$$\dot{x}_i = v_i^+(\mathbf{x}) - v_i^-(\mathbf{x})$$
$$v_i^+ = \alpha_i \prod_j x_j^{g_{i,j}}$$
$$v_i^- = \beta_i \prod_j x_j^{h_{i,j}}$$

where α_i and β_i are constants (analogous to rate constants) that characterize the rates of production and consumption, respectively, of x_i and the g and h parameters are termed kinetic orders. The summation is over those variables that contribute to the derivative. Integration of the power laws can generate accurate time profiles over a wide range of values for the variables [5]. Extrapolation outside this range can lead to significant errors, however.

The steady-state solution to these equations can be obtained by setting the time derivatives to zero. A change of variable $y_i = \ln(x_i)$ gives a set of linear algebraic equations in $\mathbf{y}$ with constant coefficients which are functions of the rate multipliers and the kinetic orders. These transformed equations are easily solved for the state variable values.

An S-system readily permits calculation of sensitivity coefficients. Differentiation of the equations for $\dot{x}_i$ with respect to parameter variations gives

$$\frac{\partial \dot{x}_i}{\partial \alpha_i} = \frac{d}{dt}\left(\frac{\partial x_i}{\partial \alpha_i}\right) = v_i^+ / \alpha_i$$
$$\frac{\partial \dot{x}_i}{\partial \beta_i} = \frac{d}{dt}\left(\frac{\partial x_i}{\partial \beta_i}\right) = -v_i^- / \beta_i$$

for the sensitivities with respect to the rate multipliers and

$$\frac{\partial \dot{X}_i}{\partial g_i} = \frac{d}{dt}\left(\frac{\partial x_i}{\partial g_i}\right) = v_i^+ \ln(x_i)$$

$$\frac{\partial \dot{X}_i}{\partial h_i} = \frac{d}{dt}\left(\frac{\partial x_i}{\partial h_i}\right) = -v_i^- \ln(x_i)$$

for the sensitivities with respect to the kinetic orders. Integration of these equations gives the temporal profiles of the sensitivities.

Sensitivity analysis of S-systems in a steady state is most easily obtained from the equations transformed into logarithmic space [6]. The sensitivities of $\ln(\mathbf{x})$ to variations in $\ln(\alpha)$ or $\ln(\beta)$ are zero, and the sensitivities to the kinetic orders are $\ln(\mathbf{x})$ for g-parameters and $-\ln(\mathbf{x})$ for h-parameters. These are equivalent to relative sensitivities

$$\frac{\partial \ln(x_i)}{\partial \ln(\alpha_j)} = \frac{\partial x_i}{\partial \alpha_j} \cdot \frac{\alpha_j}{x_i}$$

which are especially convenient for comparison of variant models. As long as the solution space of $\mathbf{x}(t)$ is within the range where the approximate rate equations are valid, these sensitivities will be reliable.

Sensitivity in Metabolic Networks

The functions $\mathbf{f}(\mathbf{x}, \boldsymbol{\alpha}, t)$ in the state equations above are arbitrary, but empirical representations such as S-systems may not provide the detail needed for mechanistic understanding. When $\mathbf{x}$ is a vector of the concentrations of metabolic intermediates, the state functions are sums of the rates of the individual enzyme-catalyzed reactions in the pathway being modeled. The rate functions include kinetic constants (*e.g.* K_m and V_{max}) that can be estimated from temporal or initial velocity data obtained with the purified enzyme. The use of the relative sensitivity equations obtained for such a system has been termed metabolic control analysis [7]. The sensitivity equations have been derived for steady states [8], and attempts have been made to extend the theory to transients [9].

Metabolic control theory commences from the definition of the time derivatives of the metabolite concentrations.

$$\dot{x}_i = \sum_{j=1}^{m} \sum_{k=1}^{n} M_{j,k} v_k$$

where m is the number of metabolic intermediates, n is the number of reactions, and $M_{j,k}$ is the stoichiometry of metabolite j in reaction k whose rate is v_k. To analyze metabolic control in a steady state, the time derivatives are set to zero and the resulting linear equations can be solved.

The rate laws **v** can be differentiated with respect to each parameter to obtain relative sensitivity coefficients called "elasticities."

$$\varepsilon_{j,k} = \frac{\partial v_k}{\partial \alpha_j} \cdot \frac{\alpha_j}{v_k}$$

These quantities are fundamental properties local to the enzymes, but they determine the global regulatory properties of the reaction network. In metabolic control theory, the parameters $\boldsymbol{\alpha}$ are normally the enzyme concentrations and the enzymatic rate is assumed to vary linearly with this quantity. Two additional coefficients are defined; they are the concentration control coefficient

$$C_{i,k}^{x} = \frac{\partial x_i}{\partial v_k} \cdot \frac{v_k}{x_i}$$

and the flux control coefficient

$$C_k^J = \frac{\partial J}{\partial v_k} \cdot \frac{v_k}{J}$$

where J is the net flux through the pathway at steady state. The concentration control coefficient is the sensitivity of a steady-state metabolite concentration to the activity of an enzyme. The flux control coefficient is the sensitivity of the overall pathway flux to the activity of an enzyme.

The control coefficients are derived from a linear perturbation analysis. A new steady state achieved after an infinitesimal time following an infinitesimal parameter variation is expanded about the original state in a Taylor series, which is truncated after the first-order terms. Consequently, these sensitivity coefficients are properties of the entire network rather than of each enzyme considered in isolation. Yet, as shown above, the control coefficients are dependent on the elasticities, which are properties of the isolated enzymes.

At steady state and in the absence of conservation constraints or interactions among the

enzymes, the above sensitivity coefficients obey the following relations. The connectivity theorems [10] state

$$\mathbf{C}^x \boldsymbol{\varepsilon} = -\mathbf{I}$$
$$\mathbf{C}^J \boldsymbol{\varepsilon} = \mathbf{0}$$

where $\mathbf{C}^j$ is a diagonal matrix of flux control coefficients. Note that the control coefficients are specified by the elasticities. The summation theorems [10] state

$$\mathbf{C}^x \mathbf{I} = \mathbf{0}$$
$$\mathbf{C}^J \mathbf{I} = \mathbf{I}$$

Simultaneous solution of these linear algebraic equations yields the values of the control coefficients.

The predictions of metabolic control analysis have been tested experimentally [11]. Flux control coefficients have been determined by direct modulation of the activity a targeted enzyme [12]. Sometimes this was achieved by genetic manipulation [13]. In other cases enzymatic activity was reduced by addition of inhibitors [14]. The measured values were found to agree with the calculated values.

Examples from Physiological Modeling

A physiological model describes an experimental animal or human as a series of spaces called “compartments” corresponding to specific anatomical regions. A typical model may contain compartments for blood, fat, liver, muscle, and other tissues. Each compartment is associated with a volume and a blood perfusion rate. The agent being studied is administered into one compartment, distributed by blood flow among the other spaces, and equilibrated between the blood and tissue. The equilibrium ratio of concentrations in two compartments is called a partition coefficient. In most models, the agent is a chemical that is metabolized by enzymes in one or more tissues, and nonlinear functions for the reaction rate are included. The above processes constitute pharmacokinetics. The state equations are differential equations for the agent and its metabolites.

In more advanced models, the administered chemical or one of its metabolites participates in

other processes that are thought to be involved in the biological response. Examples are activation or inhibition of biochemical processes consequent to binding to specific proteins, alteration in gene expression by interference with the action of transcription factors, and formation of covalent complexes with proteins or DNA. These processes constitute the pharmacodynamics of the system; they introduce additional state equations and other quantities associated with these chemical species.

A physiological model of the disposition in rats of inhaled carbon tetrachloride included compartments for blood, fat, liver, and consolidated compartments for slowly (muscle, skin, and bone) and rapidly (remainder of body) perfused tissues [15]. Metabolism was assumed to be confined to the liver, and apparent V_{max} and K_m values were estimated by fitting data for gas uptake from a closed chamber. To identify the physiological parameters that control the disposition of the chemical this model was subjected to sensitivity analysis. The predicted chamber concentration was sensitive to about half of the parameters, and many of the relative sensitivity coefficients displayed striking time dependence.

The sensitivity to the volume of slowly perfused tissues was high early in the time course, reflecting its large volume into which the inhaled gas could be distributed. The sensitivity to the fat volume was high late in the time course, because at equilibrium most of the body burden was in this tissue, owing to its high partition coefficient. Other compartment volumes had small sensitivities. The blood perfusion rate of fat was the only flow rate (aside from overall cardiac output) to which the model's predictions were sensitive. The sensitivity with respect to the ventilation rate was significant mainly at early time points.

The model was sensitive to partition coefficients for blood:air, fat, and slowly perused tissues but not for liver or rapidly perfused tissues. As for compartment volumes, the sensitivity to the partition coefficient for slowly perfused tissues was high at early time points, and the sensitivity to the fat partition coefficient was high at late time points. The sensitivities to the metabolic parameters increased as the chamber concentration decreased over time. This behavior is due to the decreasing enzymatic elasticity (see above) with increasing substrate concentration, *i.e*. the

reaction rate becomes less sensitive to the substrate concentration as the enzyme approaches saturation.

Sensitivity analysis was performed on a similar physiological model for the disposition of inhaled 2-butoxyethanol [16]. In this model, the conversion of the parent to the metabolite was represented by several steps. The simulated experiments involved several hours of inhalation of the parent chemical followed by several hours during which the chemical and its metabolite 2-butoxyacetic acid were cleared from the blood. Time profiles were computed for the sensitivities of the blood concentrations of parent and metabolite with respect to the dissociation constant of 2-butoxyethanol for plasma protein, the metabolic parameters, and the metabolite excretion parameters.

The blood concentration of parent was sensitive only to the V_{max} and K_m of the initial step of metabolism, and the sensitivity increased as the chemical was cleared from the blood. The metabolite blood concentration was sensitive only to the partition coefficient for slowly perfused tissues, reflecting the large volume of this space that is accessible to this polar chemical species. The sensitivity of metabolite production decreased during the clearance period, corresponding to the falling concentration of the precursor chemical.

A complex physiological model was developed for the disposition of intravenous and oral doses of 2,4,4-trimethyl-2-pentanol in male rats [17]. This model included compartments for blood, fat, liver, kidney, gastrointestinal tract, and consolidated compartments for slowly and rapidly perfused tissues. Metabolism by oxidation and glucuronidation was included in liver and kidneys. The parent chemical binds to a plasma protein found only in male rats, and the liganded protein is partly excreted in urine and partly taken up by kidney proximal tubule epithelial cells where it accumulates. The model includes enhanced production and re-absorption of the protein and reduced proteolysis in kidney lysosomes consequent to ligand binding. A sensitivity analysis was performed to identify those adjustable parameters to which the blood and kidney concentrations of the chemical and the accumulation of the binding protein in the kidney were most sensitive.

The pathways represented in this model are complex. In addition to equations for delivery of parent chemical to tissues and metabolic clearance, the model includes equations for induction of the protein, ligand binding, renal uptake, and proteolytic degradation. Because the state variables depend on so many interacting factors, the relative sensitivities of the blood and kidney concentrations to the parameters are similar. The variation is greatest at low dose where the numerical values of the sensitivity coefficients were smallest. The sensitivities increase in value as the dose increases, reflecting the increased rates of the regulatory processes.

Conclusions

It appears that simple models such as that for 2-butoxyethanol disposition are sensitive to a small subset of the model's parameters, whereas a complex model such as that for 2,4,4-trimethyl-2-pentanol is more uniformly sensitive to the parameters. Complexity is only one aspect; the identity of the state variable whose sensitivity is of interest is also important. In the 2-butoxyethanol model, the blood concentrations of parent were sensitive only to the kinetic parameters for its metabolic clearance and not to the parameters for further metabolism. Clearly, in the absence of metabolism 2-butoxyethanol would rapidly equilibrate with the ambient air and all sensitivities would be zero. Thus, the blood concentration is set solely by the kinetic parameters for the enzyme that consumes this chemical.

In the 2,4,4-trimethyl-2-pentanol model, the blood and kidney concentrations of this chemical depend on parameters for absorption, distribution to tissues, metabolism, and induction of binding protein. As the renal accumulation of binding protein also depends on the amount of the parent chemical in blood and tissues as well as hepatic production and renal re-absorption, it is sensitive to the same parameters. Because the equations for these effects are not separable, the sensitivity coefficients for these variables with respect to each parameter are similar. Thus, the ability of sensitivity analysis to identify the critical parameters depends on selection of an appropriate state variable that captures the essential features of the biological response.

REFERENCES

[1] Bode HW, Network Analysis and Feedback Amplifier Design. Van Nostrand–Reinhold, Princeton, NJ, 1945.

[2] Tomovic R, Sensitivity Analysis of Dynamic Systems. McGraw-Hill, New York, 1963; Cruz JB and Perkins WR, A new approach to the sensitivity problem in multivariable fedback system design. IEEE Trans Autom Contr **9:** 216–223, 1964; Kokotovic PV and Sannuti P, Singular perturbation method for reducing the model order in optimal control design. IEEE Trans Autom Contr **13:** 377–384, 1968; Kreindler E, Formulation of the minimum trajectory sensitivity problem. IEEE Trans Autom Contr **14:** 206–207, 1969.

[3] Frank PM, Introduction to System Sensitivity Theory. Academic, New York, 1978.

[4] Savageau MA, Biochemical systems analysis. III. Dynamic solutions using a power law approximation. J Theor Biol **26:** 215–226, 1970.

[5] Voit EO and Savageau MA, Accuracy of alternative representations for integrated biochemical systems. Biochemistry **26**(6869–6880), 1987.

[6] Voit EO, Canonical modeling: review of concepts with emphasis on environmenntal health. Environ Health Prespect **108**(Suppl. 5)**:** 895–909, 2000.

[7] Kacser H and Burns JA, The control of flux. Symp Soc Exp Biol **27:** 65-104, 1973; Heinrich R and Rapoport TA, A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur J Biochem **42:** 89–95, 1974.

[8] Burns JA, Cornish-Bowden A, Groen AK, Heinrich R, Kacser H, Porteous JW, Rapoport SM, A RT, Stucki JW, Tager JM, Wanders RJA and Westerhoff HV, Control of metabolic systems. Trends Biochem Sci **10:** 16, 1985.

[9] Acerenza L, Sauro HM and Kacser H, Control analysis of time-dependent metabolic systems. J Theor Biol **137:** 423–444, 1989; Kohn MC and Chiang E, Metabolic network sensitivity analysis. J Theor Biol **98:** 109-126, 1982.

[10] Heinrich R and Schuster S, The Regulation of Cellular Systems. Chapman & Hall, New York, 1996.

[11] Groen AK, Wanders RJA, Westerhoff HV, van der Meer R and Tager JM, Quantification of the contribution of various steps to the control of mitochondrial respiration. J Biol Chem **257:** 2754–2757, 1982; Torres NV, Mateo F, Meléndez-Hevia E and Kacser H, Kinetics of metabolic pathways. A system in vitro to study the control of flux. Biochem J **234:** 169–174, 1986; Delgado J and Liao JC, Metabolic control analysis using transient metabolite concentrations. Determination of metabolite concentration control coefficients. Biochem J **285:** 965–972, 1992.

[12] Flint HJ, Tateson RW, Barthelmess IB, Porteous DJ, Donachie WD and Kacser H, Control of the flux in the arginine pathway of Neurospora crassa. Modulations of enzyme activity and concentration. Biochem J **200:** 231–246, 1981.

[13] Niederberger P, Prasad R, Miozzari G and Kacser H, A strategy for increasing an in vivo flux by genetic manipulations. The tryptophan system of yeast. Biochem J **287:** 473–479, 1982.

[14] Small JR, Flux control coefficients determined by inhibitor titration: the design and analysis of experiments to minimize errors. Biochem J **296:** 423–433, 1993.

[15] Evans MV, Crank WD, Yang HM and Simmons JE, Applications of sensitivity analysis to a physiologically based pharmacokinetic model for carbon tetrachloride in rats. Toxicol Appl Pharmacol **128:** 36–44, 1994.

[16] Lee KM, Dill JA, Chou BJ and Roycroft JH, Physiologically based pharmacokinetic model for chronic inhalation of 2-butoxyethanol. Toxicol Appl Pharmacol **153:** 211–226, 1998.

[17] Kohn MC and Melnick RL, A physiological model for ligand-induced accumulation of α_{2u}-globulin in male rat kidney. Role of protein synthesis and lysosomal degradation in the renal dosimetry of 2,4,4-trimethyl-2-pentanol. Toxicology **136:** 89–105, 1999.