

Kinetic Model Completeness

We say a chemical kinetic model is **complete** for a particular reaction condition when it contains all the species and reactions needed to describe the chemical processes at that reaction condition to some specified level of accuracy. In other words, the kinetic model is complete when all the reactions whose rate constants it sets equal to zero really are negligible for the specified reaction conditions and error tolerance.

Note that this definition of **complete** does NOT assume that the rate constants or any other parameters in the model are correct, just that the reactions excluded from the model are indeed negligible.

If you believe you can list all possible reactions of the species in the model, a necessary condition for a model to be complete is that the rates of these reactions to form species not included in the model must be smaller than some error tolerance; this approach was introduced by R.G. Susnow et al. (1997). However, this condition is not sufficient: e.g. one of the minor neglected species could be a catalyst or catalyst poison with a very big influence on the kinetics at very low concentrations. If the kinetics are strictly linear you can show that this necessary condition is sufficient (Matheu et al. 2002, 2003); an important example of strictly linear kinetics are the ordinary master equations used to compute pressure-dependent reaction rates.

Of course, you usually do not know all the possible reactions. But for purposes of kinetic model reduction one usually assumes that the initial “full” model is complete; you are satisfied if the reduced model reproduces the full model to within some error tolerances under some reaction conditions.

Sensitivity Analyses

A. Definitions of Sensitivities

Suppose you have a model you think is complete at some reaction condition, and you have some estimates for \mathbf{k} and \mathbf{Y}_0 so you can numerically solve

$$d\mathbf{Y}/dt = \mathbf{F}(\mathbf{Y}, \mathbf{k}) \quad \mathbf{Y}(t_0) = \mathbf{Y}_0$$

to get your best prediction for the trajectory $\mathbf{Y}(t)$. (\mathbf{Y} is normally made up of many mass fractions $y_i(t)$ and a few other time-varying state variables e.g. $T(t)$.) You never know the rate (and other) parameters \mathbf{k} and the initial concentrations \mathbf{Y}_0 exactly, so you are always interested in how much the predicted trajectory would vary if these values were a little different than the values you assumed, i.e. how sensitive is your prediction to the values of these parameters? A common way to express this is to write $\mathbf{Y}(t)$ as a Taylor expansion in the parameters \mathbf{k} and \mathbf{Y}_0 :

$$y_i(t) = y_i(t; k, Y_o) + \sum (\partial y_i(t) / \partial k_n) (k_n' - k_n) \\ + \sum (\partial y_i(t) / \partial y_m^o) (y_m^o' - y_m^o) + \dots$$

The rate constants k usually depend on T and P , so if T and/or P vary with time the first term gets to be a mess. The most popular approach (used in the latest versions of CHEMKIN, but not consistently in earlier versions) is to imagine that all of the rate constants $k_n(T,P)$ are multiplied by scaling factors D_n . Our predicted trajectory $Y(t)$ corresponds to all $D_n=1$. So the preferred expansion (for T,P varying) is:

$$y_i(t) = y_i(t; D=1, Y_o) + \sum (\partial y_i(t) / \partial D_n) (D_n' - 1) \\ + \sum (\partial y_i(t) / \partial y_m^o) (y_m^o' - y_m^o) + \dots$$

The derivatives in this expansion are sometimes called the sensitivities, though – watch out! - in practice what are usually reported are the corresponding unitless normalized sensitivities:

“Normalized sensitivity of species i to reaction n ”

$$= (1/y_i(t)) (\partial y_i(t) / \partial D_n) = \partial(\ln y_i(t)) / \partial(\ln D_n)$$

“Normalized sensitivity of species i to initial concentration of species m ”

$$= (y_m^o / y_i(t)) (\partial y_i(t) / \partial y_m^o) = \partial(\ln y_i(t)) / \partial(\ln y_m^o)$$

It must be emphasized that these sensitivities depend crucially on the assumed values of the rate constants k and all the other parameters, and on $(t-t_o)$ and Y_o . If you change the simulation in any way you will change all the computed sensitivities. The (first-order) sensitivities are also very “local” in the sense that they are only useful if the perturbed trajectory is very close to the original trajectory $Y(t)$, i.e. if the perturbation in k or Y_o is very small. If you care about a big variation (e.g. if one of your rate constants is extremely uncertain) you may not be able to get away with the first-order Taylor expansion. If you need to go to higher orders or otherwise handle big variations/uncertainties in the model parameters, see papers by H. Rabitz, G.J. McRae, and Nancy Brown.

One would often like to know the “Green’s function sensitivity” $\partial y_i(t) / \partial y_m(t')$ i.e. how much species i at time t would have changed if a little bit of species m was added at time t' . This can be computed as an ‘initial concentration sensitivity’ by setting $t_o=t'$ and $Y_o=Y(t')$. However, it is usually impractical to do this for many values of t' .

Finally, in addition to “concentration sensitivities”, one can also compute “rate sensitivities”, which are how much the instantaneous slope dy_i/dt would change if k or Y_o were perturbed. In addition to the ordinary integrated rate sensitivities (which depend on $t-t_o$, and capture the fact that the whole trajectory is perturbed starting at t_o) one can also

compute “instantaneous rate sensitivities” which are how much the rate at time t would change if the concentrations or rate constants were suddenly changed at time t . The instantaneous rate sensitivities are much simpler to compute than the other sensitivities, just a little algebra is required:

$$\text{“Instantaneous species } i \text{ rate sensitivity to reaction } n\text{”} = \partial F_i(\mathbf{Y}, \mathbf{k}) / \partial k_n$$

$$\text{“Instantaneous species } i \text{ rate sensitivity to species } m\text{”} = \partial F_i(\mathbf{Y}, \mathbf{k}) / \partial y_m$$

The instantaneous sensitivities do not depend on the history of the system, they will be the same numbers regardless of how the system got to its present state \mathbf{Y} .

Finally, sensitivities are partial derivatives, so it is important to know what other variables are being held constant – you can get very different numerical values depending on what you hold constant. In particular, it is important to know whether the thermochemistry is being held constant, or whether all the forward and reverse rates are individually being held constant. The former is generally preferred since you are usually more confident in the thermochemical parameter values than the kinetic parameters, and also there are usually fewer thermochemical parameters than reverse rate constants. (Though if you are highly uncertain about the thermochemistry of some of your species, you will also be interested in the sensitivity to those thermochemical parameters!). If the forward and reverse rates are treated separately, it is very important to look at sensitivities with respect to both the forward and the corresponding reverse rate simultaneously: if they are opposite in sign and nearly equal in magnitude, it means you are really sensitive to the thermochemistry, not the kinetics. (This usually happens when that reaction is fast and nearly equilibrated.)

B. Computing Sensitivities

The other sensitivities are pretty hard to compute, they all require solving rather large systems of differential equations. For example:

$$d(dy_i(t)/dD_n)/dt = d(dy_i(t)/dt)/dD_n = d(F_i(\mathbf{Y}, \mathbf{D}, \mathbf{k})/dD_n) = \sum (\partial F_i / \partial y_m) (dy_m/dD_n) + \partial F_i / \partial D_n$$

defining $S_{in}(t) = dy_i(t)/dD_n$ evaluated on the reference trajectory $\mathbf{Y}(t)$ we can rewrite this:

$$dS_{in}/dt = \sum (\partial F_i / \partial y_m) S_{mn} + \partial F_i / \partial D_n \quad S_{in}(t_0) = 0$$

For the sensitivities to initial concentrations $Z_{il}(t) = \partial y_i(t) / \partial y_l^0$ the equation is very similar:

$$dZ_{il}/dt = \sum (\partial F_i / \partial y_m) Z_{ml} \quad Z_{il}(t_0) = \delta_{il}$$

In both cases, the right hand side depends on the Jacobian $\partial \mathbf{F} / \partial \mathbf{y}$ which depends on \mathbf{Y} (and so varies with time) so it is best to solve this simultaneously with the original system of equations $d\mathbf{Y}/dt = \mathbf{F}$, to avoid having to evaluate the Jacobian repeatedly. There are several tricks which take advantage of the special simple form of the sensitivity equations (e.g. linear in \mathbf{S}), and good programs which use these tricks for computing the first-order

sensitivities, including DASAC, DASPK, and DAEPACK. For details, see for example recent papers by P.I. Barton.

Note that there are a huge number of sensitivities $O(N_{\text{species}} * N_{\text{reactions}} * N_{\text{timesteps}})$. To speed the computations and reduce the output, one can only ask for the sensitivities to a few parameters (rather than all the rate constants), or for the sensitivities of only a few species. Exactly how much this saves you will depend on the details of the numerical algorithm. The algorithms in CHEMKIN are not very efficient, and the post-processor is worse, so you generally do much better if you only ask for the output you really need. If you need to solve large sensitivity problems that CHEMKIN cannot handle, you might want to ask Prof. Barton for advice.

C. Some Uses for Sensitivities

1) Estimating effects of Varying Initial Conditions

Very often you are interested in the behavior of your system as you make (relatively small) variations in the initial conditions. For example, if you are doing an ignition problem, what happens if the initial temperature is 10 K higher? In a chemical process, what happens if you change the initial concentration of the one of the reagents by a factor of two? Are we sensitive to small variations in any of the initial concentrations?

$$\mathbf{Y}(t; \mathbf{Y}_0 + \Delta \mathbf{Y}_0) = \mathbf{Y}(t; \mathbf{Y}_0) + \mathbf{Z}(t) * \Delta \mathbf{Y}_0$$

Are any of the Z's for minor species variations greater than ϵ/Y_{max} , where Y_{max} is the expected upper bound on contaminants in the input stream, and ϵ is the acceptable upper bound on contaminants in the output stream?

2) Putting Error Bars on the Model Predictions

A first estimate of the uncertainty in the model prediction $Y(t)$ is given by

$$|\Delta \ln y_i(t)| \sim \sum |\partial(\ln y_i(t)) / \partial(\ln D_n)| * |\Delta \ln k_n| \\ + \sum |\partial(\ln y_i(t)) / \partial(\ln y_m^0)| * |\Delta \ln y_m^0|$$

The first term frequently dominates. Unfortunately, the uncertainties estimated in this way are usually pretty large, accurately reflecting our inability to precisely predict chemical kinetics. The errors in the predicted selectivities (i.e. relative yields) are usually smaller than the errors in the predicted absolute concentration profiles.

$$|\Delta \ln(y_i(t)/y_j(t))| \sim \sum |S_{in} - S_{jn}| * |\Delta \ln k_n|$$

3) Identification of the Important Uncertainties

The sums above that estimate the total uncertainties are usually dominated by a few terms. The uncertain parameters that correspond to these big terms are those that matter the most for the simulation at hand; if you want a more accurate simulation you need to tighten these error bars.

4) Design of Kinetic Experiments

In many kinetic experiments you measure a signal $S(t)$, and what you want is for $S(t)$ to be very sensitive to the rate parameter k you are trying to determine, and insensitive to the uncertainties in all the other parameters. Then you can have reasonable confidence

when you say that your observation of $S(t)$ implies a certain value for k . When you design an experiment, you are well-advised to immediately build a simulation model, and check that your observable(s) really are sensitive to the rate constant you are trying to determine.

5) Valid Range Analyses

Each kinetic model is only valid over a finite range of reaction conditions. Unfortunately, this valid range is not usually known. For our model to be complete, we require that all the ignored side reactions have negligible rates. By determining the sensitivity of these side reactions to changes in the reaction conditions we can estimate the valid range, i.e. the range over which the model will be complete. (This presumes that the rate parameters in the model are not too far off.) See J. Song et al. Chem. Eng. Sci. 2002.

Flux Analyses

It is often very helpful to examine the rates at which each of the reactions in the system are proceeding. In particular, you will want to know which reactions have high net rates. By considering the forward and reverse rates separately, you can also identify the reactions which are in quasi-equilibrium. Often only the reactions which have high net rates, high sensitivities, or which are quasi-equilibrated really matter for the kinetics. Almost certainly you cannot model the system correctly if you leave out a reaction with a high net rate or a high sensitivity.

Very often it is interesting to know which reactions contribute the most to the instantaneous formation or destruction of a key species. Often a species will be involved in dozens of reactions, but only three or four have significant net rates. CHEMKIN will make these net rate plots for you.

A problem with net rates is that all the species in quasi-steady-state, and all the fast reactions which are nearly equilibrated, will have net rates near zero. It is therefore often interesting to look at the formation and destruction rates or the forward and reverse rates separately.