



Global Climate & Energy Project  
**STANFORD UNIVERSITY**

## Process Informatics

### Investigators

David M. Golden, Professor, Mechanical Engineering

### Objective

“PriMe” (Process Informatics Model) is a cooperative activity aimed at the development of predictive reaction models for combustion. The primary motivation for this effort is to establish and demonstrate the community approach to kinetic-model development and, perhaps most importantly, establish a means for reaching community consensus on the models and data.

### Background

Chemical reaction models will never be complete, yet there are large amounts of information and methods required to produce such models. These data are scattered over different sources and are not properly evaluated. Most importantly, these data cannot be applied directly to practical problems, rather they have to be “transformed” into useful models. Such models, however, cannot be created by simple “compilation” of the data. Chemical reaction model building is a time-consuming activity that requires expert knowledge. At present, most treat this activity as an art. The goals in this activity are to convert such model building into a science, to automate the methodology, and to make the most current results available in a convenient form for the user.

Immediate needs for predictive reaction models presently exist in combustion engineering, the petrochemical industry, and pharmaceuticals. As the quantity and quality of information in biological fields increases, there will also soon be a need for predictive dynamic models of biological systems. As scientific computing becomes more powerful and more readily accessible to industry, the industrial interest in process simulation continuously grows.

Nearly all of the energy currently used in the industrialized world comes from burning fossil fuels, and chemistry is the essence of combustion systems, from internal combustion engines to gas turbines. Knowledge of the chemical mechanism is at the center of device design to limit combustion-generated environmental pollution. Societal demands for cleaner and more efficient combustion are rapidly bringing the chemical aspects of combustion processes to the forefront.

Combustion systems are represented by complex dynamic models that serve as input for the design of equipment so as to maximize the efficiency of fuel utilization while minimizing unburned hydrocarbons, greenhouse gases and pollutants such as nitrogen oxides, sulfur oxides, soot or air toxics. The complexity of these models originates from high dimensionality, nonlinearity of the underlying differential equations, multi-dimensional correlations, and multifaceted dependence on temperature and pressure. This complexity is beyond simple analysis. At the same time, because of the important societal dependence on the knowledge of combustion processes, decades of research have been devoted worldwide to the subject matter. As a result, a large amount of knowledge has accumulated and the levels of fundamental theory and numerical solution have been advanced to the stage of practical application.

However, the history of combustion chemistry model development has been sporadic and piecemeal. Models are often constructed from a single investigation and have little chance of allowing reliable extrapolations to other conditions. Attempts to correct such individual models lead to a larger number of models, each failing in one respect or another. Thus, although specific

properties of a combustion process, such as ignition, flame speed, temperature, and both desirable and undesirable combustion products, are subject to calculation through the model, accuracy is severely limited by the reliability of the input rate parameters. A consistent, reliable set of values for these input parameters is best obtained by a systematic constrained optimization, based on an evaluated database of elementary reaction rate coefficients. This procedure adjusts key input rate parameters within their error ranges to obtain the best self-consistent model predictions of the experimental data.

A first effort at developing the new paradigm for dynamic models of complex chemical systems has been demonstrated with a quantitative chemical model for natural gas combustion, "GRI-Mech". That model was optimized using the method of "Solution Mapping". The approach starts with the proposition that although measured combustion properties, such as ignition, temperature, flame speed and concentrations of species are predictable by models containing the chemistry of many basic reactions, the combined errors in the input quantities and in the experiments require a systematic optimization procedure in order to produce a consistent reliable mechanism. Thus the result is more than a collection of individual reaction pathways, it is an integrated model that accounts for the uncertainties.

## Approach

The Process Informatics infrastructure will have two principal components: a *Database* and a collection of *Tools*. The Process Informatics *Database* will represent the most currently complete set of knowledge available in a given field. In the field of combustion, it will contain experimental data, on both combustion systems and on elementary reactions, molecular properties determined from quantum chemical calculations, reaction rates obtained from reaction-rate theories, and similar information. The Process Informatics *Tools* will be of two general kinds, those enabling the collection, transfer, organization, display, and mining of the data —*i.e.*, computer science tools, and those enabling processing and analysis of the data along with assembly of the data into models—*i.e.*, scientific and numerical tools.

The two principal customers of the Process Informatics System are the *data provider* and the *model user*. During the development stage, there will also be a *model builder*, whose role will eventually be automated—and providing the means for this automation will be a prime objective.

A *Data Provider* (Experimenter, Theorist) makes a request to deposit new observations or new computational results. The protocol assures completeness of the data submission. The deposited data are immediately analyzed for consistency with the database and the results are reported both to the data provider and to the *scientific council* (see below). Upon approval of the council, the database is modified. In other words, the database will be fluid and will be continuously modified and these modifications will be documented.

A *Model User* (Design Engineer, CFD Researcher) requests a kinetic model (or a simulation with such a model) and specifies the conditions of interests, the desired level of accuracy, and the mathematic form of the model (detailed, reduced, parameterized, etc). The system checks for the existence of such a model and if none is available, then one is generated.

The goal is not merely a collection of tools, but a shift in the paradigm of the scientific process: building targeted knowledge by the entire community and providing the wealth of information in its entirety to every user.