Partial Least Squares Regression (PLSR) for the Discovery of "New Biology"

The molecular and cellular biological sciences are transforming from being "data poor" to "data rich". Experimental access to the genome, the transcriptome, the proteome and the metabolome are at unprecedented levels and growing rapidly. But the availability of data does result inexorably in necessarily better data nor better science. The promise of computational molecular systems biology is to bring a deeper understanding of the principles underlying cellular function and to uncover "New Biology" in a manner analogous to the way that theory and experiment collaborate to uncover "New Physics. Partial Least Squares Regression is a multivariate modeling approach that has repeatedly proven to be surprisingly adept at uncovering novel signaling events that regulate cell fate decisions. This tutorial will introduce the student to the application of PLSR to computational molecular systems biology and will provide "hands-on" instruction of how to use PLSR in Matlab. The tutorial will be structured in the following way:

- Introduction to PLSR – Mathematical background and theory (30 minutes)
- Successful applications of PLSR to molecular systems biology – Examples and lessons learned (30 minutes)
- Applying PLSR for hypothesis generation – Experimental design and data mining (30 minutes)
- Using PLSR in Matlab – Data preparation and statistical analysis (30 minutes)

Instructor: Gerard J. Ostheimer, Ph.D.
Global Lead
Sustainable Bioenergy High-Impact Opportunity
UN Sustainable Energy for All