Detailed Statement of Research

I would like to explain the underlying basic mathematics and engineering of many apparently diverse problems I have solved in the last few years and how we have further developed some of the basic methodologies while doing research in a diversity of application areas.

Many engineering (including financial engineering and supply chain management) problems can be modeled, for the purpose of *dynamic analysis* in continuous time, by the ordinary differential equation:

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t),\ x(0) \text{ given}, x(t), u(t) \text{ may be bounded} \qquad \text{(I)}$$

where, t is an indexing variable, $x(t)$ is the state variable of the system, $u(t)$ is the control variable, and A and B are given matrices. In the most general case, the right-hand side of (I) could be a set of nonlinear functions. This is a well-known equation, however, when there is uncertainty, due to either stochastic or fuzzy, the model becomes more realistic but can present some challenging practical difficulties in its solution such as, non-linearity, non-smoothness, non-differentiability, and large-scale. The variables $x(t)$ and $u(t)$ may be bounded, in which case, for stochastic problems, little is known. We have made some progress in discrete time domain for such problems, as explained later. For *steady-state analysis*, and for analysis in the frequency domain, the problem may become an algebraic problem. In addition to (I), for the purposes of design, we usually have a cost function that either minimizes the total cost of operations or maximizes the total net benefits from operating the system, for example, in the form of

$$\min_{u(t)} f(x(t), u(t))$$
(II)

In the case of uncertainty (depends on the problem, for example, in water reservoirs they may come from net inflows in equation (I) and prices in equation (II)), the cost function must be appropriately changed and additional probabilistic constraints may need to be added as explained later, case by case. In some of our problems, we might be only interested in knowing the effect of certain design decisions (for example, in pollution modeling, the difference between having a treatment facility or not) and this cost function may not be explicitly present.

Applications:

<u>(i) Pollution modeling:</u>

See references [1,2, and 3]. The work that started as part of my sabbatical research [1] has culminated in publications [2 and 3]. The fundamental research consisted of (a) extending the advanced first-order second moment method of Hasofer-Lind to stochastic differential equations and using it for estimating probabilities, for example, the probability of the Dissolved Oxygen level not going below the prescribed limit 4mg/l in a river. Since our original work, this type of method has been included in publicly available software for engineering communities outside the Civil Engineering field but is still not good enough for stochastic variables (for random variables the work that is available in this software is good enough; note that stochastic variables are indexed random variables); see [4] for such software, (b) a new "variance reduction" scheme that reduced considerably the number of ensembles needed for good estimation [2] and (c) a new colored noise scheme to maintain higher-order efficiency of numerical integration that were developed for deterministic processes [3]. This means, a large number of existing software meant for solving systems of differential equations, for example, in climate modeling, multibody dynamics, stock prices, and others, can be made highly efficient for more realistic noise processes. In the future we will continue with research noted in (c) to include non-gaussian colored noise, using copulas for stochastic systems [5].

(ii) Inventory type stochastic optimization:

Consider the operations of the Great-Lakes system (our computational models for Int's Joint Commission saved the U.S. and Canadian tax payers over $5 billion in 1992). Let inflows in equation (I) be considered uncertain and goals or target values for storage levels and releases in equation (II), are considered deterministic (different from a traditional inventory problem where quantity ordered (input) is deterministic but demand (output) is stochastic). In reservoir problems, as well as in any storage inventory problems, the state variables and control variables are bounded. In such cases, the methodology to solve for the statistical characteristics of bounded storage volumes in continuous time domain, especially for multi-reservoir systems, has not yet been found. We developed a methodology for long-term operations in discrete time domain (*but continuous state space*), where the statistics of x_0 is not given but must be assumed as the same as the statistics of x_{Final}, for the sake of conservation, and neither of which are known! In long-term, this is equivalent to assuming stationarity of the state variable $x(t)$. This is an extremely hard problem (for example, a single variance equation in our approximate model runs into many pages). Our earlier method using indicator functions and Taylor series expansion with deterministic release policies, although approximate, worked well in operations optimization problems. For a single reservoir, we have now extended this to consider random release policies and for policies as a function of storage volumes [6]. Extending this method to multi reservoir (or storage) systems and comparing the optimal solutions with existing methods, when prices are also uncertainty, will be part of our future research.

(iii) Electronic System Design:

Equation (I) is applicable in many electronic and mechatronic systems. A complication here is that when such systems are manufactured, due to manufacturing impurities, the expected performance may not be what is needed. Equation (II) usually refers to maximizing the yield, which is defined as the percentage of systems that pass the performance test over the total number of systems manufactured. This requires development of methods to incorporate tolerance in design, which, at the outset, may look similar to problems in inventory type problems in (ii), but in reality, is quite different. In these tolerance design problems here, the decision variables are the mean and tolerance value of components in matrices A and B that are uncertain, unlike in (ii) where the decision variable is usually considered deterministic. We developed new methods when the randomness is non-gaussian and non-symmetric in [5] for small to medium size problems (less than 10 random variables). In the future we will extend this to large-scale problems with non-gaussian dependent random variables.

(iv) Economic and Financial Systems:

A predominant theme here is that efficient nonlinear programming has been extended to include minimization of risk in addition to maximizing expected returns (benefits) in electricity generation expansion, for example, [6], refinery planning, and portfolio management. Extending this to non-gaussian dependent uncertainty is a future direction.

Both equations (I) and (II) are appropriately modified. We mention one important difference between problems in (ii) and here. While the systems are almost identical, a major difference is that, in finance, for example, an investment $u(t)$ can be in any stock $x(t)$, unlike in inventory type systems, where due to physical limitations, the connection between storage systems modeled by various $u(t)$ are much fewer. On the other hand, in finance, the end conditions are usually known (that is we start with zero value, and require a certain end value) unlike in water reservoir systems where these are results to be determined.

(v) Soft Computing Applications:

While all the above works predominantly considered stochastic (probabilistic) uncertainty it has become important also to consider another kind of uncertainty, namely the fuzzy (or possibilistic) uncertainty. [See 8, for example]. Applying these techniques in health care, network management, and fault detections are some of our future research topics. Depending on the problem, this may not explicitly need equations of types (I) and (II); for example, in disease diagnosis or fault detection we use methods of classification (supervised learning) and clustering (unsupervised learning). Fault detection method detects any significant changes of current outputs from known models of time series of various monitored device outputs indicating good working conditions. In supervised learning both inputs and outputs (known classifications) are available. The training is done using past information and testing (or predictions of classes) is done using only the inputs. In unsupervised learning, the data are simply clustered or grouped using similarity measures. Soft computing methods, in addition to traditional methods of change detections, are proposed to be used in the work involving early epilepsy detections and control.

In conclusion, while I strive to find many challenging real-world problems that require varied mathematical and software skills to solve, it is always the design questions that these problems pose that keeps me interested. Moreover, due to the diverse nature of these problems, I seek collaboration and work well with my collaborators.

References

[1] Shakshuki, E., and K. Ponnambalam, 2002. An automatic formulation of inverse free second moment method for algebraic systems, *Reliability Engineering and System Safety*, 76(1), 81-89

[2] Schoenmakers, J.G.M., A.W. Heemink, K.Ponnambalam, and P.E. Kloeden, 2002. Efficiency of variance reduction for Monte Carlo simulation of stochastic environmental models, *Applied Mathematical Modelling*, 26, 785–795.

[3] Stijnen, J., A.W (VS). Heemink, and K. Ponnambalam, 2003. Stochastic environmental quality models, *Water Resources Research*, 39 (3), 1053-61.

[4] Eldred, M.S., H. Agarwal, V. M. Perez, Jr. Wojtkiewicz, S. F., and J. E. Renaud, 2007. Investigation of reliability method formulations in DAKOTA/UQ. *Structure & Infrastructure Engineering: Maintenance, Management, Life-Cycle Design & Performance*, 3(3):199–213.

[5] Ponnambalam, K., A. Seifi, and J. Vlach, 2003. Yield optimization with correlated design parameters and non-symmetrical marginal distributions, Prof. of Int'l Symposium on Circuits and Systems (ISCAS2003), Vol 4, 736-739.

[6] Fletcher, S., and K. Ponnambalam, 1998. A constrained state formulation for the stochastic control of multi-reservoir systems, *Water Res. Res*., 34(2), 257-270.

[7] Alvarez, J., K. Ponnambalam, and V. H. Quintana, 2007. Generation and transmission expansion under risk using stochastic programming, *IEEE Transactions on Power Systems*, 22 (3), 1369 – 1378.

[8] Mousavi, S (VS). J., K. Ponnambalam, and F. Karray, 2007. Infereing rules for reservoir operations using fuzzy regression and ANFIS, *Fuzzy Sets and Systems*, 58 (10), 1064-1082.