Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 10-11, 2007

Comparison of Uncertainty Analyses for Auto-Associative Kernel Regression Method

Jae Yong Lee, Ho Cheol Shin and Moon Ghu Park
Korea Electric Power Research Institute, Yuseong-gu, Daejeon, 305-380, jylee@kepri.re.kr

1. Introduction

Traditionally, the calibration of nuclear instruments has been performed at each refueling cycle. However, many nuclear plants have moved toward condition-directed rather than time-directed calibration. This condition-directed calibration is accomplished by on-line monitoring. On-line monitoring employs an empirical predictive modeling technique to assess instrument channel performance. A critical issue surrounding the real application to the nuclear power plants is quantifying the uncertainty of the predictive model. This study investigates the various uncertainty quantification techniques employed in OLM algorithm which is auto-associative kernel regression(AAKR).

2. AAKR Method

Auto-Associative Kernel Regression (AAKR) is a type of similarity based model which is a nonparametric modeling technique that uses the similarity of a query vector to memory or exemplar vectors to infer the model's response [1]. AAKR is a nonparametric, data-driven modeling technique that uses historical, fault-free observations to correct faults in current observations. For a simple single-input, single-output (SISO) regression model, where the input x is used to estimate the output y, the Nardaraya-Watson estimator is: [1]

$$\hat{y}(x)=\frac{\sum_{i=1}^{n}\left[K(d,h)Y_i\right]}{\sum_{i=1}^{n}K(d,h)}.$$

where: n is the number of exemplar observations in the AAKR model

X_i and Y_i are the input and output for the i^{th} exemplar observation

x is a query input

$K(d,h)$ is a weighting or kernel function, which generates a weight (similarity) for a given difference of a query from an exemplar vector

$$K(d,h)=\frac{1}{\sqrt{2\pi h^2}}e^{-\mathbf{d}^2/2h^2}$$

h is a kernel's bandwidth

d is a distance between query input and exemplar vector which is calculated the following equation

$$d(\mathbf{X}_i,\mathbf{x})=\sqrt{(X_{i,1}-x_1)^2+(X_{i,2}-x_2)^2+\Lambda+(X_{i,p}-x_p)^2}$$

$\hat{y}(x)$ is an estimate of y, given x

For a query observation of the inputs, the AAKR estimation process can be structured into three steps. First, the distance of the query from each of the input exemplars is calculated. Next, the distances are supplied as inputs to a kernel function, which converts the distances to weights (similarities). Finally, the weights are used to estimate the output by calculating a weighted average of the output exemplars. These steps are depicted in Figure 2.1.[3]

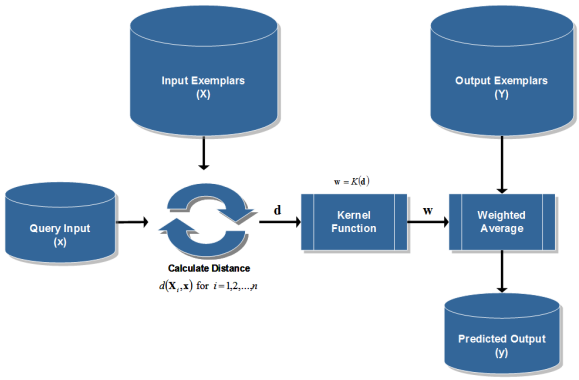

Figure 2.1 Process diagram for kernel regression prediction algorithm

3. Uncertainty Analysis Method of AAKR

There are basically two methods which are developed to analyze the uncertainty of AAKR. The analytic and Monte Carlo methods for calculating uncertainty are described in ref.[3]. Analytic uncertainty is estimated through equations derived from the model's mathematical architecture. Analytic uncertainty can be evaluated during model implementation to estimate the uncertainty for each prediction. The Monte Carlo uncertainty, however, is much more computationally intensive. As such, it is generally evaluated prior to model implementation. This uncertainty estimate is applied to each model prediction. Monte Carlo uncertainty is estimated by applying a Monte Carlo re-sampling technique. With Monte Carlo techniques, the training data is resampled multiple times and for each of these resampled datasets, a new model is constructed. The variation between all of these models is then taken as a measure of the variance portion of the total uncertainty. The uncertainty estimate is applied to the denoised residuals of the model. Figure 3.1 shows the

process diagram for Monte Carlo uncertainty. This uncertainty is used to construct a confidence interval(CI) centered at zero, the expected value of the denoised residuals. The residual coverage is then calculated as the fraction of denoised residuals contained within the confidence interval.

For a confidence level of $1-\alpha$, the CI may be written as:[2]

$$E(\hat{y}) \pm t_{n-p,\alpha/2} \sqrt{Var(\hat{y}) + Bias^2}$$

where: n is the number of training observations

p is the number of variables used to infer y

$t_{n-p,\alpha/2}$ is the t-statistic which approximates the normal distribution for $n-p$ degrees of freedom and confidence level $1-\alpha$

$E(\hat{y})$ is the expected model prediction of y

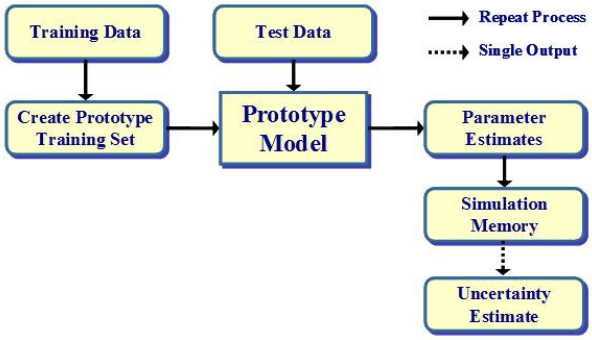

Figure 3.1 Process Diagram for Monte Carlo Uncertainty Estimation

4. Results and Discussion

In order to check the applicability of AAKR method and its uncertainty analysis method, the RCS flow loop The real measurement data are used. The data sets are composed of 10,000 of training data, 30,000 of test data and 30,000 of validation data. The sampling time is 30 seconds.

With these data, prediction accuracies, uncertainties which are calculated by using two methodologies and detectabilities are presented. The detectability is defined as follows,[3]

$$D_i = \frac{U_i}{span\ (x)} \cdot \frac{1}{1 - S_{A,i}}$$

where: U_i is sensor i's model uncertainty (95% CI)

span(x) is sensor x's span

$S_{A,i}$ is sensor i's auto-sensitivity.

Per Table 4.1, the prediction accuracies are less than 0.04% which are very accurate because those signals are well correlated. The analytic uncertainties are about 0.1% which are bigger than Monte Carlo Uncertainties. This means analytic method is more conservative than Monte Carlo method. The detectabilities are less than 0.4%. These are well below 1% which is a typical instrument drift limit of 1%.

Table 4.1 : Results of uncertainty analysis for AAKR

	Signal 1	Signal 2	Signal 3
Signal Noise Estimate (%span)	0.14	0.14	0.129
Correlation Coefficients of the Signals			
Signal 1	1	0.839	0.814
Signal 2	0.839	1	0.801
Signal 3	0.814	0.801	1
Accuracy (%span)	0.0335	0.0299	0.0273
Auto-Sensitivity	0.652	0.694	0.642
Cross-Sensitivity	0.147	0.175	0.128
Detectability (%span)	0.316	0.361	0.284
Analytic Uncertainty (%span) and its Coverage	0.11	0.11	0.102
	1	1	1
MonteCarlo Uncertainty (%span) and its Coverage	0.0306	0.0288	0.0275
	0.799	0.924	0.959

5. Conclusion

Per the reasonable uncertainties and detectabilities, AAKR method and its uncertainty analysis method which are presented in this paper are applicable to the RCS flow estimation for power plants.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Dr. J. Wesley Hines at The University of Tennessee, Knoxville for his guidance and providing useful information on OLM.

REFERENCES

[1] D.Garvey, J.Garvey, R.Seibert, S.Arndt and J. Hines, Application of On-Line Monitoring Techniques to Nuclear Plant Data, 5th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technology (NPIC&HMIT '05), Albuquerque, NM, Nov., 2006.

[2] B. Rasmussen, Prediction Interval Estimation Techniques for Empirical Modeling Strategies and their Applications to Signal Validation Tasks, Ph.D. dissertation, Nuclear Engineering Department, The University of Tennessee, Knoxville, TN,2003

[3] J .Hines, R. Seibert, Technical Review of On-line Monitoring Techniques for Performance Assessment Volume II: Theoretical Issues, NUREG/CR-6895 (Draft), U.S. Nuclear Regulatory Commission, Washington, D.C.