View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Texas A&M Repository

Predicting System Performance with Uncertainty

Bin Yan and Ali Malkawi

T.C. Chan Center for Building Simulation and Energy Studies,
University of Pennsylvania
220 South 34th Street, Philadelphia, PA 19104, United States

Email: binyan@design.upenn.edu

Abstract:
The main purpose of this research is to include uncertainty that lies in modeling process and that arises from input values when predicting system performance, and to incorporate uncertainty related to system controls in a computationally inexpensive way. We propose using Gaussian Processes for system performance predictions and explain the types of uncertainties included. As an example, we use a Gaussian Process to predict chilled water use and compare the results with Neural Network. As an initial step of our research, we examine how variation in AHU supply air temperature affects chilled water use in summer time. We briefly discuss the advantages of our proposed method and future research topics in the concluding remarks.

Keywords:
Gaussian Process, Performance prediction, System controls, Uncertainty

1. Introduction

Performing better risk and uncertainty analyses of performance predictions through the entire life cycle of a building is one of the most important challenges that engineering design faces (Augenbroe, 2002). Uncertainty and sensitivity analysis have been extensively applied in science and engineering. However, their applications to building systems are still limited.

Uncertainty can enter a model when making predictions in various contexts. One way to categorize is to consider uncertainty as that lies in modeling process and that arises from input values for predictions. Uncertainty in modeling process is seldom quantified. Most uncertainty studies focus on uncertainty in input values for predictions. The input values associated with predictions can come from estimations or measurements corrupted with noise. Therefore, it is more reasonable to assign probability distributions over their domains of plausible values than to assign fixed single-point values. In some cases, it is desired to investigate the impact of variation in inputs on outputs by allowing inputs to vary in their domains.

Monte Carlo experiment is the most widely used method of analyzing input uncertainty (Hamby, 1995). Several studies used Monte Carlo method with building simulation to study building and system design with input uncertainty (de Wit & Augenbroe, 2002; Domínguez-Muñoz et al., 2010). We found two areas where there could be improvement in current uncertainty research.

Most building simulation models are highly complex and computationally expensive. Monte Carlo experiment requires a large number of model evaluations. As the complexity of uncertainty increases, the number of simulations required increases significantly. The time cost limits the extension of uncertainty analysis.

Current studies have not covered uncertainty related to system controls in operations. Measurements in system operations are usually corrupted by sensor noise. For example, measurements of temperature, humidity, air flow and water flow. Furthermore, few systems perform as intended. Usually there is a discrepancy between intended and actual performance. A straightforward example is that there exists deviation between set-points and measured values of controlled variables such as AHU supply air temperature.

The main purpose of this research is to include uncertainty that lies in modeling process and that arises from input values when predicting system performance, and to incorporate uncertainty related to system controls in a computationally inexpensive way. In the following sections of the paper, we propose using an instance-based learning method Gaussian Process for system performance predictions. We explain the types of uncertainties included in Gaussian Processes. In order to evaluate the predicting accuracy, we test a Gaussian Process with real metered data and compared its results with another widely used machine learning method Neural Network. As an initial step of applying Gaussian Process to uncertainty analysis of system operations, we present a case study of predicting energy use with uncertain AHU supply air temperature. As a conclusion, we briefly discuss the advantages of our proposed method and future research topics.

2. Methodology

The use of Gaussian Process has grown significantly after the works of (Neal, 1995 & Rasmussen, 1996) in machine learning community. Gaussian Process regression has been successfully applied to various predicting tasks. Figure 1 summarizes the procedures of using Gaussian Processes for predictions. A Gaussian Process is built upon training data, which can be sensor or metered data of a real system, or simulated data generated from complex models. Then the model takes new inputs and makes predictions with uncertainty.

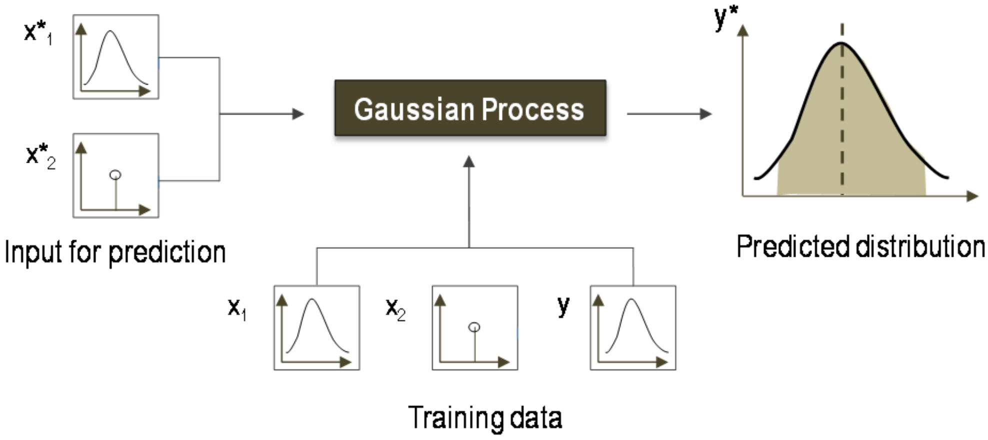


Figure 1 Diagram of predicting with uncertainty using Gaussian Process

In Gaussian Processes, the uncertainty in modeling process comes from noises in training data and distance between training inputs and inputs associated with predictions. One source of noises in training data is measurement noise. For example, it is reasonable to assume that time is noise-free, while the measurement of flow rate is usually corrupted by sensor noise. Measurement noise may exist in both training inputs and training targets. While for noise in training targets, there is another source of uncertainty. The features in an existing model might not fully explain the variance in training targets. There might be some other important features that affect outputs. Noise in targets might be reduced if we could recognize some more features and include them in the model. The variance of a prediction also depends on the distance between its input point and training inputs. Gaussian Process is an instance-based learning method. If a new input point is far from training points, the variance is large in the prediction.

Variation in input values associated with predictions leads to an extra uncertainty in predictions. In some cases, it is our interest to investigate the impact of inputs on the output by varying inputs according to appropriate distributions and examining the corresponding distributions of outputs.

We briefly list the formulas of Gaussian Process regression we use in this study. For a comprehensive introduction to Gaussian Process modeling, please refer to (Rasmussen & Williams, 2006). The following formulas do not include the noise in training inputs. In this study, training inputs are assumed to be noise-free. In our further research, we will include noise in training inputs.

When using Gaussian Process for regression, for a noise-free input $\mathbf{x}^*$, the predictive distribution is Gaussian with mean $\mu(\mathbf{x}^*)$ and variance $\sigma^2(\mathbf{x}^*)$ (Rasmussen & Williams, 2006)

$$\mu(\mathbf{x}^*) = \mathbf{k}(\mathbf{X}, \mathbf{x}^*)^{\mathrm{T}}(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y} \tag{1}$$

$$\sigma^2(\mathbf{x}^*) = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}(\mathbf{X}, \mathbf{x}^*)^{\mathrm{T}}(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\, \mathbf{k}(\mathbf{X}, \mathbf{x}^*) \tag{2}$$

where the choice of covariance function in this study is

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp\left[-\frac{1}{2}\left(\mathbf{x}_i - \mathbf{x}_j\right)^T \mathbf{W}^{-1}\left(\mathbf{x}_i - \mathbf{x}_j\right)\right]$$

and

$$\mathbf{W} = \mathrm{diag}[w_1^2, w_2^2, \ldots, w_D^2]$$

$\mathbf{k}(\mathbf{X}, \mathbf{x}^*)$ is the $N \times 1$ vector of covariance functions between training inputs $\mathbf{X}$ and the new input. $\mathbf{K}$ is the $N \times N$ matrix of covariance functions between each pair of training inputs. σ_n^2 denotes the variance of Gaussian noise in training targets $\mathbf{y}$. σ_f, σ_n and $w_1, w_2 \ldots w_D$ are hyperparamters to be trained in a Gaussian Process.

Covariance function is the central part of Gaussian Process modeling. Inputs that are judged to be close by the covariance function are likely to have similar outputs. A prediction is made by considering the covariances between the predictive case and all the training cases (Rasmussen, 1996).

To incorporate uncertain values of an input point associated with a prediction, assuming the input distribution is Gaussian $\mathbf{x}^* \sim \mathcal{N}_{\mathbf{x}^*}(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*})$, then the predictive mean $\mu(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*})$ and variance $\sigma^2(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*})$ with the noisy input are (Girard et al., 2003)

$$\mu(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*}) = \mathbf{q}^{\mathrm{T}}\boldsymbol{\beta} \tag{3}$$

$$\sigma^2(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\Sigma}_{\mathbf{x}^*}) = k(\boldsymbol{\mu}_{\mathbf{x}^*}, \boldsymbol{\mu}_{\mathbf{x}^*}) + \mathrm{Tr}[(\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}} - (\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{Q})] - (\mathbf{q}^{\mathrm{T}}\boldsymbol{\beta})^2 \tag{4}$$

where

$$\boldsymbol{\beta} = (\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}$$

$$q_i = |\mathbf{W}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}^*} + \mathbf{I}|^{-\frac{1}{2}}\sigma_f^2 \exp\left(-\frac{1}{2}(\boldsymbol{\mu}_{\mathbf{x}^*} - \mathbf{x}_i)^{\mathrm{T}}(\boldsymbol{\Sigma}_{\mathbf{x}^*} + \mathbf{W})^{-1}(\boldsymbol{\mu}_{\mathbf{x}^*} - \mathbf{x}_i)\right)$$

and

$$Q_{ij} = |2\mathbf{W}^{-1}\boldsymbol{\Sigma}_{\mathbf{x}^*} + \mathbf{I}|^{-\frac{1}{2}}\sigma_f^2 \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_i + \mathbf{x}_j}{2} - \boldsymbol{\mu}_{\mathbf{x}^*}\right)^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{\mathbf{x}^*} + \frac{1}{2}\mathbf{W}\right)^{-1}\left(\frac{\mathbf{x}_i + \mathbf{x}_j}{2} - \boldsymbol{\mu}_{\mathbf{x}^*}\right)\right) \cdot \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}}(2\mathbf{W})^{-1}(\mathbf{x}_i - \mathbf{x}_j)\right)$$

With the assumption of Gaussian input distribution and using the covariance function above, there is no need to run extra simulations to incorporate uncertain values of an input point. It can be simply derived from the analytical expressions above. This significantly reduces the time cost of uncertainty analysis.

3. Case Study

3.1. Predict Chilled Water Use

In this case study, we test Gaussian Process modeling on metered chilled water use data. Data samples are on an hourly basis. The target is chilled water (W/m^2). The features include outside air dry-bulb temperature (°C), humidity ratio (kg/kg), the hour of the day. We use $\sin\left(\frac{2\pi\cdot\mathrm{hour}}{24}\right)$ and $\cos\left(\frac{2\pi\cdot\mathrm{hour}}{24}\right)$ to represent the hour of the day. It is assumed that measurements of time, temperature and humidity ratio are noise-free, while measurements of chilled water use are noisy.

Figure 2 shows the 24-hour prediction of chilled water use by a Gaussian Process trained by 216 hourly data points. The solid line indicates the predictive mean, grey area as 95% confidence region, compared with observed values shown in red dots. Most of the predictive means are close to observed values. Noises in training targets and the distance between training inputs and test inputs account for the uncertainty in predictions.

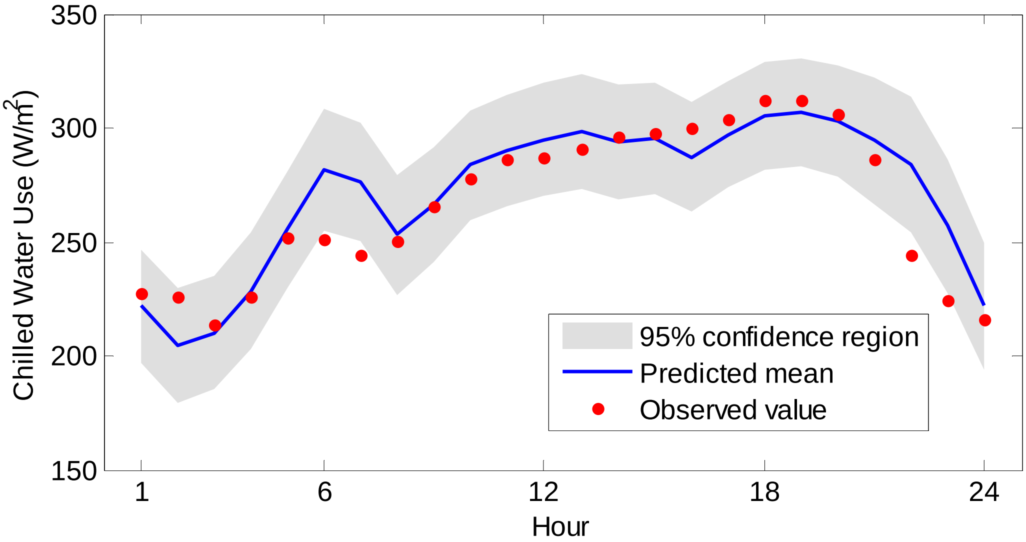


Figure 2 24-hour prediction of chilled water use

In order to evaluate the accuracy of Gaussian Process, the results are compared with those of Neural Network using the same inputs. Neural Network has been used to predict building energy use approximately since the 1990s. The reported error rates of short-term prediction (1h to 24h) can be as low as 1%-5%. Long-term prediction accuracies are also promising (Dodier & Henze, 2004). In this study, the training of neural network is implemented through Matlab (version R2011a) Neural Network Toolbox. In this model, there is one hidden layer with 15 neurons. The activation equation in the hidden layer is sigmoid, and linear in the output layer. The training algorithm is Levenberg-Marquardt backpropagation.

The coefficient of determination is used to compare how well the predictions are between Gaussian Process and Neural Network. The coefficient of determination R^2 is calculated as

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \qquad (5)$$

where the values y_i are observed values of targets, the values f_i are predicted values. For Gaussian Processes, values f_i are the predicted mean values. $\bar{y}$ is the mean value of the observed targets. The better a model is likely to predict future outcomes, the closer the value of R^2 is close to 1. A larger R^2 means a smaller sum of squared errors of prediction.

Metered chilled water use for 140 days is available for this study. Three types of prediction tasks are experimented, which are 24-hour prediction, 72-hour prediction and 9-day prediction. 10-fold cross-validation is used in model training and testing. The overall R^2 value is used for comparison. The results are shown in Figure 3. In this comparison study, it turns out which model predicts better depends on the time scale of prediction and characteristics of data.

Gaussian Process seems to outperform Neural Network on short-term prediction. According to Figure 3, Gaussian Process shows a better performance than Neural Network when predicting chilled water use 24-hour ahead. However, among the 14 cross-validations of 24-hour prediction, Gaussian Process does not outperform Neural Network all the time. When using data groups shown in

(a) GP better (b) NN better (c) GP better

Figure 4(a) and (c), the R^2 values of Gaussian Process are higher, while using data groups shown in

(a) GP better (b) NN better (c) GP better

Figure 4(b), the R^2 values of Neural Network are higher. It appears that Gaussian Processes lose advantages when chilled water use has a strong linear relationship with outside air temperature. Two modeling methods perform almost same for 72-hour prediction. The R^2 value of Neural Network is slightly higher regarding 9-day prediction.

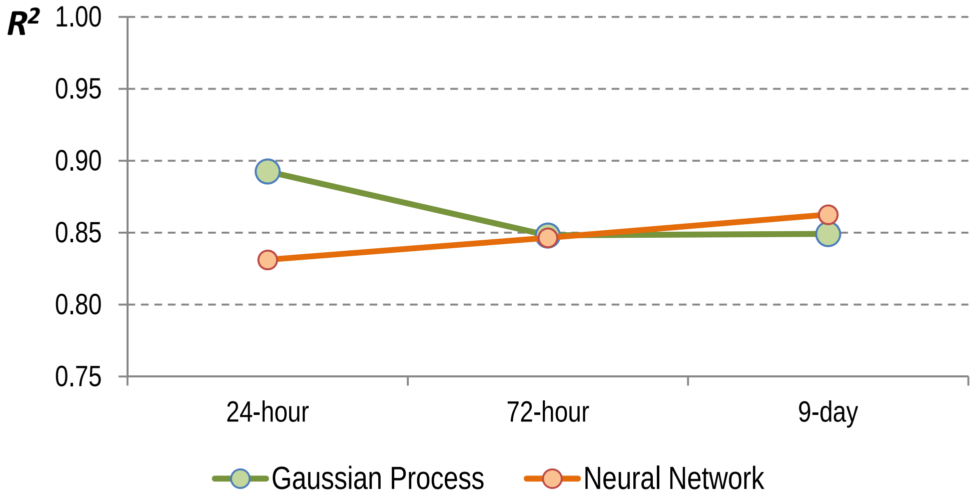
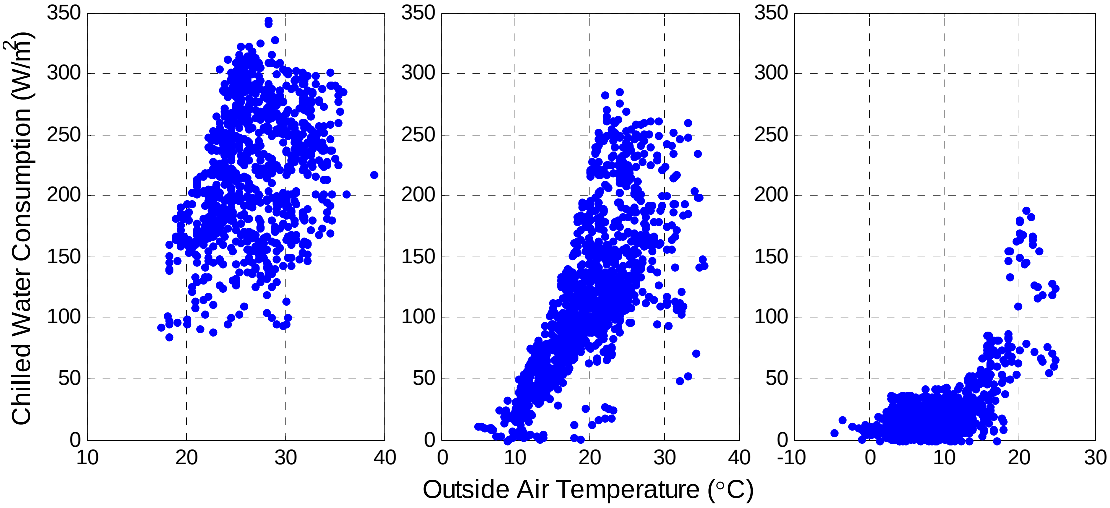


Figure 3 Comparison of $\boldsymbol{R^2}$ values of Gaussian Processes and Neural Networks



(a) GP better (b) NN better (c) GP better

Figure 4 Data groups on which Gaussian Processes and Neural Networks have different predictive accuracy

It can be concluded from the cross-validations above that the predictive accuracy of Gaussian Processes is close to the widely used Neural Networks. In some cases, such as several short term prediction tasks, Gaussian Processes even show some advantages. More careful design for comparative studies might be necessary in order to be confident that the observed R^2 of these two methods reflects their real accuracy in performance rather than a random fluctuation. Furthermore, it is unreasonable to generalize the conclusion of this experiment, which is based on a particular dataset, to other datasets. However, this experiment still enables us to get an idea of how well Gaussian Processes will perform on other datasets with similar characteristics, which seems very promising.

3.2. Examine the Impact of AHU Supply Air Temperature on Chilled Water Use

In this case study, we examine the impact of variation in AHU supply air temperature on chilled water use. The system under study is an AHU VAV system with terminal reheat for an office building, which runs 24 hours a day. One summer month of measured hourly AHU supply air temperature is available for study. Its histogram is shown in Figure 5.

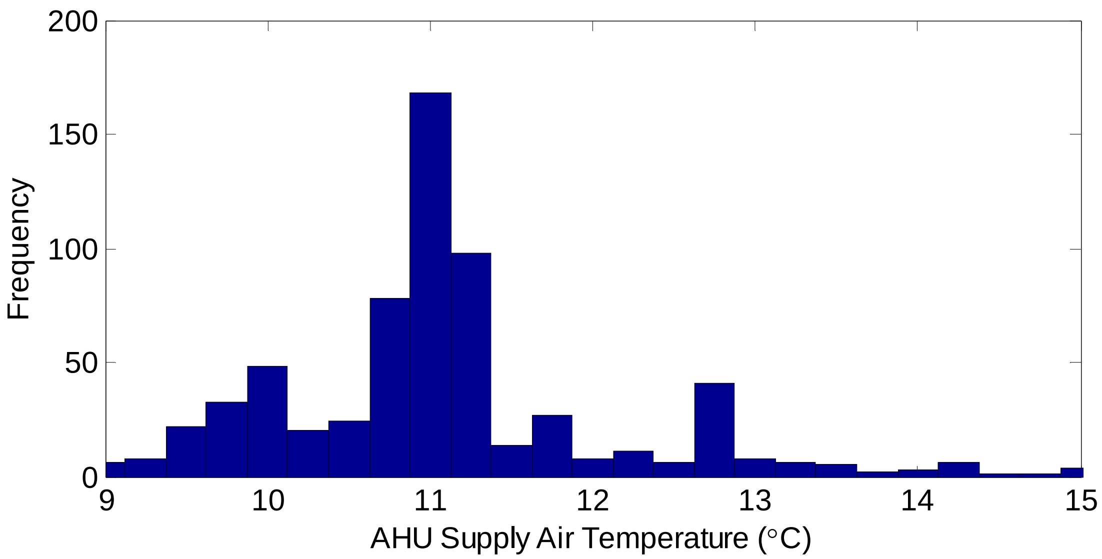


Figure 5 Histogram of measured AHU supply air temperature

The set-point of AHU supply air temperature is 11.1°C. The mean value of measured hourly AHU supply air temperature equals the set-point. However, a standard deviation of 1.1°C is observed. The AHU supply air temperature varies from 9°C to 15°C. Poor PID control, insufficient or excessive supply of chilled water might account for the deviation from set-point.

The wide range of variation in actual AHU supply air temperature directly affects system energy use. One conventional way to examine the extent of the impact is to perform a Monte Carlo experiment, generating random AHU supply air temperature from its probability distribution and running simulations over all combinations of the inputs. We propose a different method, using a Gaussian Process to build a surrogate model based on data points available and plugging the input distribution into equations (3) and (4) to get the predictive distribution of energy use directly.

Some constraints exist when deriving the predictive distribution of output from analytical expressions of equations (3) and (4). First, the distributions of the inputs to be examined are assumed to be Gaussian. For other distributions, approximate or exact analytical expressions are also possible, but will be different. Second, training sets need to cover most of the input domain to be examined in the study. Otherwise, prediction accuracy will be compromised and the uncertainty introduced by modeling process will be dominant. Third, the predictive distribution includes the uncertainty of modeling process. Comparison with the predictive distribution derived from noise-free inputs is necessary. Lastly, the computational cost of Gaussian Process is $O(n^3)$, where n is the number of training points. If the number of training points needed for the model is large, the advantage of using Gaussian Processes is less prominent.

We build a Gaussian Process using time, outside temperature and humidity, and AHU supply air temperature as training inputs, chilled water use as training targets. The data is on an hourly basis for nearly one month. The training inputs are treated as noise-free, while training targets as noisy. The training R^2 is 0.9808. Then for each point, we use $\mathcal{N}(11.1, 1.1^2)$ as the new AHU supply air temperature distribution. The new predictive distributions of hourly chilled water use are calculated according to equations (3) and (4). An extra uncertainty in predictions is introduced by variance in AHU supply air temperatures.

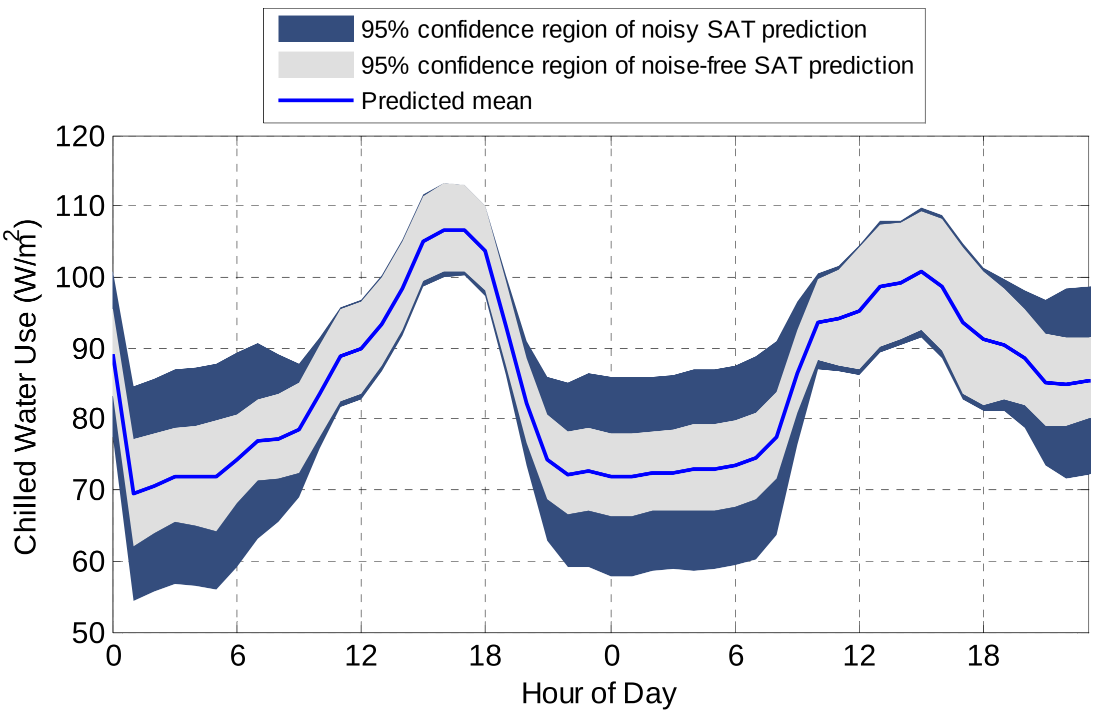


Figure 6 Predictive distributions of hourly chilled water use which include the uncertainty introduced by the variance in AHU supply air temperatures

Figure 6 shows the predictive distributions of chilled water use for 48 hours. The results are compared with the predictive distributions derived from noise-free AHU supply air temperature, which is assumed to be 11.1°C all the time. With a variance of 1.1^2 in AHU supply air temperature, the predictive means stay almost same, while the 95% confidence regions expand in some time periods. The dark blue area is the extra uncertainty introduced by variance of AHU supply air temperature.

We can see from Figure 6 that, during working hours, the variation in AHU supply air temperature almost has no effect on chilled water use. In summer during working hours, the amount of chilled water needed to process the cooling load does not change with AHU supply air temperature. When cooling load is high, higher AHU supply air temperature results in larger supply air flow rate, and the chilled water needed to process the air remains the same. And vice versa. During nighttime, the outside air temperature drops and internal load is minimal. As cooling load decreases, supply air flow rate is fixed at its minimum. Therefore, increasing AHU supply air temperature reduces chilled water use. A low AHU supply air temperature will increase chilled water use, and more reheat is necessary to compensate the excessive cooling. Around 1°C standard deviation in AHU supply air temperature accounts for a standard deviation as large as 5-8% of the predictive mean values of chilled water use during some night hours.

The example above shows how to use a Gaussian Process to study the uncertainty introduced by input variation. With the assumption that the input distribution is Gaussian, the predictive distribution can be calculated directly without Monte Carlo experiment. It is necessary that the training set should cover the most of the input domain. Otherwise, the uncertainty introduced by modeling process itself would be too large. Usually this is not an issue if data is generated from simulation. It might be challenging when building a Gaussian Process based on observations from actual performance. The example above uses measured AHU supply air temperature, while the chilled water use is simulated by EnergyPlus since no metered data is available.

4. Concluding Remarks and Further Work

This paper introduces predicting system performance through Gaussian Processes, which include uncertainty that lies in modeling process and arises from input values. Instead of building a model based on physical principles and using metered data for calibration, Gaussian Processes can directly use observed system performance to build a statistical model for further analysis. It avoids configuring numerous parameters required, which are difficult to know and estimate. Gaussian Processes can also serve as surrogate models for computationally expensive simulations. The outputs are predictive distributions with mean and variance. With the assumption that the input distribution is Gaussian, the uncertainty introduced by input variation can be calculated directly without Monte Carlo experiments.

In the first case study, we use a Gaussian Process to predict chilled water use based on time information, outside temperature and humidity. In the second case study, we examine how the variation in AHU supply air temperature affects chilled water use in summer time. As an initial step of our research, we still rely on simulated data to explore the application of Gaussian Processes, in order to develop and validate the methodology. In the future work, it will be valuable to apply Gaussian Processes to measured data of actual system performance. In addition to AHU supply air temperature, it will be interesting to study the uncertainty introduced by non-ideal control of air mixing in AHU, air flow rate and reheat in VAV terminal units, and their effect on electricity, heating and cooling energy use. Most important, we want to explore how to utilize the information of uncertainty provided by Gaussian Processes in addition to sensitivity analysis. One promising application that comes to our mind is fault detection and diagnosis, especially when our focus is on the uncertainty introduced by non-ideal control. Since Gaussian Process not only gives the predicted values, but also a measure of how confident about predictions, this extra information could probably help increase credibility in fault detection and diagnosis, especially to reduce false alarm occurrences. This will be our potential research topic in the future.

References:

Augenbroe, G. (2002), *Trends in building simulation*, Building and Environment 37 (8-9): 891-902.

de Wit, S. and Augenbroe, G. (2002), *Analysis of uncertainty in building design evaluations and its implications*, Energy and Buildings 34 (9): 951-8.

Dodier, R.H. and Henze, G.P. (2004), *Statistical analysis of neural networks as applied to building energy prediction*, Journal of Solar Energy Engineering 126 : 592.

Domínguez-Muñoz, F., Cejudo-López, J.M. and Carrillo-Andrés, A. (2010), *Uncertainty in peak cooling load calculations*. Energy and Buildings 42 (7): 1010-8.

Girard, A., Rasmussen, C. E., Quinonero-Candela, J. and Murray-Smith, R. (2003), *Gaussian process priors with uncertain inputs: application to multiple-step ahead time series forecasting*, In Becker, S., Thrun, S., and Obermayer, K., editors, Advances in Neural Information Processing Systems 15, MIT Press.

Hamby, D. M. (1995), *A comparison of sensitivity analysis techniques*, *Health Physics* 68 (2): 195-204.

Neal, R.M. (1995), *Bayesian Learning for Neural Networks* PhD thesis, Dept. of Computer Science, University of Toronto.

Rasmussen, C.E. (1996). *Evaluation of Gaussian Processes and other Methods for Non-Linear Regression* PhD thesis, Dept. of Computer Science, University of Toronto.

Rasmussen, C.E. and Williams, C.K.I. (2006), *Gaussian Process for Machine Learning*. MIT Press.