

# A Study on the Prediction of Power Consumption in the Air-Conditioning System by Using the Gaussian Process

Chang-Yong Lee\* · Gensoo Song\*\* · Jinho Kim\*<sup>†</sup>

\*Dept. of Industrial & Systems Engineering, Kongju National University

\*\*R&D Center of Duksan A.C.M. Co. LTD.

## 정규 확률과정을 사용한 공조 시스템의 전력 소모량 예측에 관한 연구

이창용\* · 송근수\*\* · 김진호\*<sup>†</sup>

\*공주대학교 산업시스템공학과

\*\*덕산공조(주) 기업부설연구소

In this paper, we utilize a Gaussian process to predict the power consumption in the air-conditioning system. As the power consumption in the air-conditioning system takes a form of a time-series and the prediction of the power consumption becomes very important from the perspective of the efficient energy management, it is worth to investigate the time-series model for the prediction of the power consumption. To this end, we apply the Gaussian process to predict the power consumption, in which the Gaussian process provides a prior probability to every possible function and higher probabilities are given to functions that are more likely consistent with the empirical data. We also discuss how to estimate the hyper-parameters, which are parameters in the covariance function of the Gaussian process model. We estimated the hyper-parameters with two different methods (marginal likelihood and leave-one-out cross validation) and obtained a model that pertinently describes the data and the results are more or less independent of the estimation method of hyper-parameters. We validated the prediction results by the error analysis of the mean relative error and the mean absolute error. The mean relative error analysis showed that about 3.4% of the predicted value came from the error, and the mean absolute error analysis confirmed that the error is within the standard deviation of the predicted value. We also adopt the non-parametric Wilcoxon's sign-rank test to assess the fitness of the proposed model and found that the null hypothesis of uniformity was accepted under the significance level of 5%. These results can be applied to a more elaborate control of the power consumption in the air-conditioning system.

**Keywords** : Air-Conditioning System, Gaussian Process, Bayesian Statistics, Time-Series Prediction

### 1. 서론

한정된 공장 부지에 급격히 증가하는 설비들로 인해

최근의 공조 시스템은 매우 복잡한 구조를 띄게 되어 공조 시스템의 제어 및 관리가 어려워지는 문제가 발생하고, 이러한 문제점은 결과적으로 공조 시스템의 안정성과 에너지의 낭비를 초래한다[6]. 공조 시스템의 에너지 절약을 위한 기존 정책은 장비의 효율 개선을 통한 소비 전력의 절대량을 감소시키는 방식으로 주로 진행되어 왔다[7].

Received 25 January 2016; Finally Revised 26 February 2016;

Accepted 3 March 2016

<sup>†</sup> Corresponding Author : kjh@kongju.ac.kr

그러나 최근 공장 운영 측면에서 에너지 절약에 대한 관심이 높아지면서, 에너지의 효율적인 사용 및 관리를 위해 공장 에너지 관리시스템(Factory Energy Management System, FEMS)의 보급이 점차 확대되고 있다. 특히 공조 설비의 소모 전력(혹은 에너지)은 주로 시간에 따라 변하는 시계열 데이터(time-series data) 형식을 취하는데, 기존의 공조 설비 제어는 주로 기본적인 온도 제어만 수행하고 있기 때문에 시간에 따라 소모 전력이 변화하는 동적 시스템을 적절하게 제어하지 못하고 있는 실정이다.

공조기에 소요되는 전력 예측은 효율적 에너지 관리 측면에서 매우 중요하며, 소모 전력 예측을 통한 에너지 관리를 위해서는 시간에 따른 소모 전력의 변화를 예측할 수 있는 시계열 모형 구축이 필요하다. 시계열 데이터의 예측은 주어진 실험 데이터를 사용하여 종속변수의 값을 예측하는 회귀분석(regression analysis) 문제에 속하는데, 시간에 따른 전력소모량 예측의 경우는 시간(독립 변수)에 따른 소모 전력량(종속 변수)을 표현하는 회귀함수를 결정하는 문제로 귀결된다.

회귀함수를 결정하는 방법은 크게 두 가지로 요약할 수 있다. 첫 번째 방법은 특정한 유형(예를 들어 선형함수 등)에 속하는 함수로 회귀함수를 제한하는 방법이다. 이 방법은 가장 널리 사용되는 것으로 선형 및 다중회귀모형이 이 방법에 속한다. 시계열 데이터의 예측인 경우에는 다중회귀모형의 일종으로 ARMA(Autoregressive Moving Average)[11], ARCH(Autoregressive Conditional Heteroscedasticity)[4], 혹은 ARMA를 일반화한 GARMA[2] 모형 등이 사용되고 있다. 또한 최근에는 선형성과 주기성을 가진 요소가 확률 분포와 혼재되어 있는 시계열 데이터의 분석에서 선형적 주기적 요소를 제거하는 방법론에 대한 연구도 수행되었다[10, 13]. 첫 번째 방법은 선택한 특정한 유형의 함수에 대한 정당성을 확보하여야 하며, 선택된 함수에 포함되어 있는 매개변수를 추정해야 한다. 따라서 함수 선택에 대한 정당성 확보와 매개변수의 추정을 추가로 고려해야 하며, 또한 선택한 함수를 사용하여 적절하게 모델링이 되지 않는 경우에는 예측 결과 역시 좋지 않다는 단점이 있다.

두 번째 방법은 베이지안(Bayesian) 방법[9]으로 가능한 모든 함수를 사전확률(prior probability)을 부여하여 생성하고, 실험 데이터를 우도(likelihood)로 사용하여 사후확률(posterior probability)로 목적함수(즉, 예측치)를 결정하는 방법이다. 이 방법은 모든 가능한 함수를 고려하기 때문에 특정한 유형의 함수에 국한되지 않는 장점은 있으나, 무한히 많은 함수들을 현실적인 시간 내에 생성해야 하는 문제가 있다. 정규 확률과정(Gaussian process)[12]은 이러한 문제를 적절하게 해결하는 방법으로 사전확률과 우도를 사용하여 사후확률을 구하기 때문에 베이지안

추론(Bayesian inference)에 속하며, 매개변수를 사용하지 않기 때문에 비모수(non-parametric)적 방법에 속한다. 이러한 이유로 정규 확률과정은 마이크로 어레이(micro-array)를 사용한 유전자 발현[8] 혹은 로봇 팔 제어[15]와 같은 복잡하고 불규칙적으로 변하는 시계열 데이터의 예측에 활용되고 있다.

일반적으로 공조기의 소모 전력은 공조기 구조뿐만 아니라, 외부 온도, 건조 대상, 공조기 출구와 실제 작업장 간의 거리 등 다양한 요인에 영향을 받기 때문에 특정한 유형의 함수를 선정하여 회귀분석을 수행하기에 무리가 따를 수 있다. 따라서 본 연구에서는 시간에 따라 불규칙적으로 변화하는 공조기의 소모 전력 시계열 데이터를 정규 확률과정을 적용하여 예측하고자 한다. 본 연구의 결과는 향후 예측 모형과 공조기의 소모 전력 상황을 비교 검토하고 예측 모형의 신뢰성을 검증을 통해 자동화 건조 공정에 사용되는 여러 공조기들의 에너지 소비 패턴을 예측하여 관리하는 데 적용할 수 있을 것으로 판단된다.

본 논문은 다음과 같이 구성되어 있다. 제 2장에서는 정규 확률과정을 소모 전력 시계열 데이터에 적용하는 과정을 설명하였으며, 제 3장에서는 공분산 함수의 결정과 초월매개변수의 추정에 대하여 언급하였다. 또한 제 4장에서는 실험 데이터를 정규 확률 과정에 적용한 결과와 그 의미를 분석하였고, 제 5장에서는 본 논문의 결론을 맺었다.

## 2. 정규 확률과정(Gaussian Process)

정규 확률과정은 미리 정한 확률 분포(즉, 사전확률)에 의하여 무작위로 함수를 생성하며, 생성한 함수 중에서 주어진 데이터에 적합한 함수들만 고려하여 목적함수를 결정한다. 여기서 주어진 데이터는 우도 역할을 하며 따라서 사전확률과 우도를 결합하면 정규 확률과정을 통하여 생성한 목적함수들에 대한 사후 확률 분포를 구할 수 있다. 특히 사전확률을 사용하여 임의의 함수들을 무한히 생성하고 그 중에서 데이터에 적합한 함수만 선택할 수 있기 때문에 데이터가 증가하더라도 생성하는 함수에 대한 융통성이 보장된다. 정규 확률과정을 예측에 적용하기 위해서는 데이터 개수만큼의 차원을 가지는 행렬의 역행렬을 구해야 한다. 데이터의 개수를  $n$ 이라 할 때,  $n \times n$ 행렬의 역행렬을 구하는데 걸리는 시간복잡도(time complexity)는  $O(n^3)$ 이기 때문에,  $n$ 이 큰 경우에는 계산 시간이 많이 소요되는 단점이 있다.

주어진 데이터 집합에 대하여 정규 확률과정을 사용한 예측 문제를 위해  $n$ 개의 데이터로 구성된 훈련 데이

터 집합  $D = \{(\vec{x}_i, y_i) | i = 1, 2, \dots, n\}$  을 고려하자. 여기서  $\vec{x}_i$  는 입력 데이터(독립변수, covariate)로 일반적으로 여러 성분으로 구성된 벡터이며,  $y_i$  는 출력값(종속변수)을 나타낸다. 일반적으로 정규 확률과정은 다양한 유형의 입력 데이터에 대하여 실행할 수 있으나, 본 연구는 시간에 따른 전력소모량을 예측하기 위한 것임으로 입력 데이터를 1차원 양인 시간으로 간주하고  $\vec{x}_i = t_i$  로 표현하기로 한다.

정규 확률과정은 무한개의 확률변수들의 집합(따라서 연속함수가 됨)으로, 이 집합에서 임의로 추출한 유한개의 확률변수(random variable)들이 결합 정규 분포(Joint Gaussian distribution)를 따를 때, 이 연속함수를 정규 확률과정이라고 정의한다. 임의의 정규 확률과정은 평균함수(mean function)와 공분산함수(covariance function)로 결정된다. 만약 함수  $f(t)$  가 정규 확률과정을 따른다면

$$f(t) \sim GP(m(t), k(t, t')) \quad (1)$$

으로 표현한다. 여기서  $m(t) \equiv E[f(t)]$  는 평균함수이며 일반적으로  $m(t) = 0$  으로 두고,  $k(t, t')$  는 공분산함수로  $k(t, t') \equiv E[(f(t) - m(t))(f(t') - m(t'))]$  로 정의한다. 평균함수는 시간  $t$  에서 임의로 생성된 함수값  $f(t)$  들의 평균을 나타내며, 공분산함수는 서로 다른 시간  $t$  와  $t'$  사이의 함수값  $f(t)$  와  $f(t')$  의 공분산을 의미한다. 따라서 시간  $t$  에서 함수값  $f(t)$  는 확률변수가 되며,  $f(t)$  와  $f(t')$  은 평균이

$$\vec{u} = \{m(t), m(t')\} \quad (2)$$

이고 공분산 행렬이

$$\Sigma = \begin{bmatrix} k(t, t), k(t, t') \\ k(t', t), k(t', t') \end{bmatrix} \quad (3)$$

인 결합정규분포(joint Gaussian distribution)을 따른다. 공분산함수는 서로 다른 시간  $t$  와  $t'$  에 대한 함수값을 결정하기 때문에 정규 확률과정을 통해 생성되는 함수  $f(t)$  의 성질은 공분산함수에 의해 결정된다.

또한 정규 확률과정은 주변화(marginalization) 성질을 내포하고 있다. 즉,  $f_i \equiv f(t_i)$  라 할 때, 정규 확률과정을 통해 생성한  $n$  개의 함수값  $\vec{f} = \{f_1, f_2, \dots, f_n\}$  는 결합 정규 분포를 따른다. 즉, 입력 데이터 집합을  $X = \{t_1, t_2, \dots, t_n\}$  라 하면, 정규 확률 과정을 통해 생성되는  $n$  개의 원소를 가지는  $\vec{f}$  는 아래와 같은 결합정규분포를 따른다.

$$\vec{f} \sim N(\vec{0}, K(X, X)) \quad (4)$$

여기서  $K(X, X)$  는  $n \times n$  공분산행렬로  $K(X, X) = \{k(t_i, t_j) | t_i, t_j \in X\}$  이다.

잡음(noise)을 고려한 경우의 관측값  $\vec{y} = \{y_1, y_2, \dots, y_n\}$  은

$$\vec{y} = \vec{f} + \vec{\epsilon}, \vec{\epsilon} \sim N(\vec{0}, \sigma_n^2 I) \quad (5)$$

가 되며, 잡음  $\vec{\epsilon}$  는 일반적으로 평균이 0이고 분산이  $\sigma_n^2$  인 정규분포를 따르고, 서로 독립임을 가정한다. 따라서 잡음을 고려한 관측값  $\vec{y}$  는  $\vec{y} \sim N(\vec{0}, cov(\vec{y}))$  를 따르고,

$$cov(\vec{y}) = K(X, X) + \sigma_n^2 I \quad (6)$$

로 주어진다.

$n$  개의 측정된 훈련 데이터(training data) 집합을  $\{X, \vec{y}\} = \{(t_i, y_i) : i = 1, 2, \dots, n\}$  라 하고, 예측을 위한  $n_*$  개의 시험 데이터(test data) 집합을  $\{X_*, \vec{f}_*\}$  라 하면, 정규 확률과정과 식 (5)에 의하여  $\vec{y}$  와  $\vec{f}_*$  의 확률분포는 결합정규분포를 따르며,

$$\begin{pmatrix} \vec{y} \\ \vec{f}_* \end{pmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \quad (7)$$

로 주어진다. 여기서  $K(X_*, X_*)$  는  $n \times n_*$  공분산 행렬로 행렬의 원소는 모든 가능한 훈련 데이터와 시험 데이터 쌍에 대한 공분산으로 구성되며,  $K(X_*, X) = K(X, X_*)^T$  이고,  $K(X_*, X_*)$  는 시험 데이터로만 구성된 공분산 행렬이다. 식 (7)과 결합정규분포의 조건부 확률분포의 성질인

$$\begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix} \sim N\left(0, \begin{pmatrix} A & C^T \\ C & B \end{pmatrix}\right) \text{ 이면} \\ \vec{b} | \vec{a} \sim N(CA^{-1}\vec{a}, B - CA^{-1}C^T) \quad (8)$$

을 이용하면  $\vec{y}$  가 주어진 경우,  $\vec{f}_*$  에 대한 조건부 확률을 구할 수 있다. 식 (8)을 식 (7)의 경우에 적용하면,  $X_*, X, \vec{y}$  가 주어진 조건하에  $\vec{f}_*$  의 확률 분포는

$$\vec{f}_* | X_*, X, \vec{y} \sim N(\vec{f}_*, cov(\vec{f}_*)) \quad (9)$$

로 주어지며, 여기서

$$\vec{f}_* \equiv K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}\vec{y} \quad (10)$$

이며,

$$cov(\vec{f}_*) \equiv K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1}K(X, X_*) \quad (11)$$

이다. 따라서 예측값  $\vec{f}_*$ 는 정규 확률분포로 주어지고, 이 값을 예측하기 위해서는 공분산 함수로 구성된 행렬  $K(X, X)$ 을 결정해야 한다.

### 3. 공분산 함수와 초월매개변수 추정

#### 3.1 공분산 함수 선택과 초월매개변수

공분산함수는 정규 확률과정에서 임의로 생성하는 사전 함수(prior function)의 성질을 규정하는 역할을 한다. 공분산의 성질에 의하여 정규 확률과정에서 사용할 수 있는 공분산 함수는 대칭적 성질(즉,  $k(t, t') = k(t', t)$ )을 가지고 있어야 한다. 여러 유형의 공분산 함수가 사용되고 있는데, 대표적인 것을 보면 제곱 지수형(squared exponential), 다항식형(polynomials), 신경망(neural network) 등을 들 수 있다[1, 13]. 각 공분산 함수는 고유한 매개변수를 가지고 있으며 사전확률의 매개변수임을 강조하기 위하여 초월매개변수(hyper-parameter)라 부르고, 일반적으로 우도가 가질 수 있는 매개변수와 차별적으로 사용된다. 표준적인 공분산함수로 제곱지수형이 가장 널리 사용되고 있는데[12], 식 (5)와 같이 잡음을 고려하는 경우의 제곱 지수형은

$$k(t_p, t_q) = \sigma_f^2 \exp\left\{-\frac{|t_p - t_q|^2}{\sigma_l^2}\right\} + \sigma_n^2 \delta_{pq} \quad (12)$$

로 표현되며 3개의 초월매개변수  $\vec{\theta} = \{\sigma_f, \sigma_l, \sigma_n\}$ 를 가지고 있다.  $\sigma_f$ 은 신호 분산(signal variance)로 공분산의 크기에 해당하고,  $\sigma_l$ 은 특정 길이 척도(characteristic length-scale), 그리고  $\sigma_n$ 은 잡음 분산(noise variance)에 해당한다. 특히  $\sigma_l$ 은 서로 다른 시간 사이에 공분산함수의 상관 정도를 나타내며, 입력 값에 따라 함수의 변동량을 조절하는 역할을 한다. 예를 들어  $\sigma_l$ 이 작을수록 함수는 빠르게 변하며,  $\sigma_l$ 이 매우 큰 경우에는 공분산함수는 시간에 거의 무관함을 의미한다. 본 연구에서는 가장 널리 사용되고 있는 제곱 지수형을 사용한다.

공분산 함수는 초월매개변수를 포함하고 있기 때문에 초월매개변수를 추정해야 하는데, 일반적으로 사용하는 추정 방법은 크게 2가지가 있다. 첫 번째 방법은 주변우도(marginal likelihood)라 불리는  $p(\vec{y} | X, \vec{\theta})$ 를 최대로 하는 초월매개변수 값을  $\vec{\theta}$ 의 추정치로 정하는 방법이고, 두 번째 방법은 교차검정(cross-validation)을 사용하는 방법이다. 본 논문에서는 주변우도와 교차검정 등 두 가지 방법을 적용하여 초월매개변수를 추정하고 그 결과에 대하여 논하고자 한다.

#### 3.2 주변 우도를 사용한 초월매개변수 추정

주변 우도 방법은 초월매개변수가 포함된 함수의 우도를 최대화하여 초월매개변수를 추정하는 것이다. 이러한 접근은 베이지안 방법이 아닌 일반적인 빈도에 의한 통계의 최대우도법(maximum likelihood estimation)과 유사하다. 식 (12)에 의해  $\vec{y}$ 에 대한 주변로그우도(marginal log likelihood, 혹은 evidence)는

$$\log p(\vec{y} | X, \vec{\theta}) = -\frac{1}{2} \vec{y}^T (K + \sigma_n^2 I)^{-1} \vec{y} - \frac{1}{2} \log |K + \sigma_n^2 I| - \frac{n}{2} \log 2\pi \quad (13)$$

과 같이 주어지며, 여기서 공분산함수 행렬  $K$ 은  $\vec{\theta}$ 의 함수로  $K = K(\vec{\theta})$ 이다.

초월매개변수를 추정하기 위하여 위의 식 (13)을 각 초월매개변수  $\theta_j (j = 1, 2, 3)$ 에 대하여 편미분하면

$$\frac{\partial}{\partial \theta_j} \log p(\vec{y} | X, \vec{\theta}) = \frac{1}{2} \text{tr} \left\{ (\alpha \alpha^T - K_y^{-1}) \frac{\partial K_y}{\partial \theta_j} \right\} \quad (14)$$

가 된다. 여기서  $K_y \equiv K + \sigma_n^2 I$ 이며  $\alpha \equiv K_y^{-1} \vec{y}$ 이고, 편미분하는 과정에서  $\frac{\partial}{\partial \theta_j} K_y^{-1} = K_y^{-1} \frac{\partial K_y}{\partial \theta_j} K_y^{-1}$ 와  $\frac{\partial}{\partial \theta_j} \log |K_y| = \text{tr} \left( K_y^{-1} \frac{\partial K_y}{\partial \theta_j} \right)$ 를 사용하였다.

주변로그우도를 수치적으로 계산하는데 가장 시간이 많이 걸리는 부분은  $n \times n$  행렬  $K_y$ 의 역행렬  $K_y^{-1}$ 를 구하는 부분이다. 일반적인 방법 (예를 들어 Gauss-Jordan 방법)을 사용하면 시간복잡도(time complexity)는  $O(n^3)$ 가 되나, 공분산 함수의 성질을 사용하면 보다 효율적으로 역행렬을 계산할 수 있다. 공분산 함수로 구성된 공분산 행렬  $K(X, X)$ 는 대칭행렬이며 양의 준정부호행렬(positive semi-definite)임으로 공분산 행렬의 고유값(eigenvalue)는 모두 양수가 되며, 따라서  $K = LL^T$ 가 성립하는 하삼각행렬(lower triangular matrix)  $L$ 이 유일하게 존재한다. 이러한 분해를 콜레스키 분해(Cholesky decomposition)라고 부르며 추후 실제적인 수치계산에 유용하게 사용된다. 특히 콜레스키 분해는 수치적으로 매우 안정적이며  $O(n^3)/6$ 의 시간이 요구됨으로  $K(X, X)$ 의 역행렬을 직접 구하는 것보다 빠르다[3].

#### 3.3 LOO 교차검정(Cross-Validation)을 사용한 초월매개변수 추정

교차검정 방법을 사용한 초월매개변수 추정은 k-fold

교차검정법에서  $k=n$  경우인 LOO(leave-one-out) 교차검정 방법을 사용한다.  $i$  번째 데이터를 제외한 훈련 데이터가 주어진 조건하에  $i$  번째 데이터에 대한 조건부 확률은 식 (16)에서

$$\log p(y_i | X, \vec{y}_{-i}, \vec{\theta}) = -\frac{1}{2} \log \sigma_i^2 - \frac{(y_i - \mu_i)^2}{2\sigma_i^2} - \frac{1}{2} \log 2\pi \quad (15)$$

로 주어지며, 여기서  $\vec{y}_{-i}$  는  $i$  번째 훈련 데이터를 제외한 값들을 나타낸다. 또한  $\mu_i$  와  $\sigma_i^2$  는 식 (5) 와 식 (6) 에서  $\vec{f}_* = y_i$ , 그리고 훈련 데이터로  $(X_{-i}, y_{-i})$  를 사용한 경우에 해당한다. 따라서 식 (18) 을 사용하면 LOO 교차검정의 예측로그확률(predictive log probability)은

$$L_{LOO}(X, \vec{y}, \vec{\theta}) = \sum_{i=1}^n \log p(y_i | X, \vec{y}_{-i}, \vec{\theta}) \quad (16)$$

가 된다.

식 (19) 를 계산하기 위해서는  $i=1, 2, \dots, n$  각각에 대하여 공분산 행렬의 역행렬인  $[K(X_{-i}, X_{-i}) + \sigma_{n-1}^2 I]^{-1}$  을 구해야한다. 이것은 계산량이  $O(n^4)$  로 시간이 매우 많이 걸리며 데이터가 많은 경우 현실적으로 가능하지 않을 수 있다. 그러나 서로 다른  $i$  에 대한  $[K(X_{-i}, X_{-i}) + \sigma_{n-1}^2 I]^{-1}$  의 계산은  $[K(X, X) + \sigma_n^2 I]^{-1}$  에서  $i$  번째 행과 열을 제외한 결과와 동일하므로 “분해를 사용한 역행렬 계산” 방법[14]을 적용하면 효율적으로 계산할 수 있다. 그 결과로  $\mu_i$  와  $\sigma_i^2$  는

$$\mu_i = y_i - \frac{[K_y^{-1} \vec{y}]_{ii}}{[K_y^{-1}]_{ii}} \quad \text{와} \quad \sigma_i^2 = \frac{1}{[K_y^{-1}]_{ii}} \quad (17)$$

로 주어지며, 여기서  $K_y \equiv K(X, X) + \sigma_n^2 I$  이다.

식 (20) 을 사용하면

$$\frac{\partial L_{LOO}}{\partial \theta_j} = \frac{\sum_{i=1}^n \left( \alpha_i [Z_j \alpha]_i - \frac{1}{2} \left( 1 + \frac{\alpha_i^2}{[K_y^{-1}]_{ii}} \right) [Z_j K_y^{-1}]_{ii} \right)}{[K_y^{-1}]_{ii}} \quad (18)$$

가 된다. 여기서  $\alpha \equiv K_y^{-1} \vec{y}$  이며  $Z_j \equiv K_y^{-1} \frac{\partial K_y}{\partial \theta_j}$  이다. 따라서

$\frac{\partial}{\partial \theta_j} L_{LOO}(X, \vec{y}, \vec{\theta}) \big|_{\theta_j = \theta_j^*} = 0$  를 만족하는  $\theta_j^*$ , ( $j=1, 2, 3$ ) 을 구하여 초월매개변수의 추정치로 사용할 수 있다.

### 3.4 켈레기울기방법(Conjugate Gradient Method)을 사용한 초월매개변수 추정

식 (14) 와 식 (18) 의 계산을 위해  $f(\vec{\theta})$  를 각각  $f(\vec{\theta}) = -L_{LOO}$  그리고  $f(\vec{\theta}) = \log p(y | X, \theta)$  라 두면, 동일한 일반화 켈레기울기방법(generalized conjugate gradient method)[5]을 사용하여 초월매개변수를 추정할 수 있다. 일반화 켈레기울기 알고리즘의 순서도를 의사(擬似) 코드(pseudo code)로 표현하면 <Figure 1> 와 같다.

<p>Step 1 : Set <math>k=0</math>, choose <math>\vec{\theta}_{(0)}</math> arbitrary, and let <math>\vec{d}_{(0)} = -\nabla f(\vec{\theta}_{(0)})</math></p> <p>Step 2 : For <math>k &gt; 0</math>, update <math>\vec{\theta}_{(k+1)} = \vec{\theta}_{(k)} + t_{(k)} \vec{d}_{(k)}</math>, where <math>t_{(k)} = -\frac{\vec{d}_{(k)}^T \nabla f(\vec{\theta}_{(k)})}{\vec{d}_{(k)}^T A \vec{d}_{(k)}} = \frac{\vec{d}_{(k)}^T (A \vec{\theta}_{(k)} + \vec{b})}{\vec{d}_{(k)}^T A \vec{d}_{(k)}}</math></p> <p>Step 3 : Update direction vector <math>\vec{d}_{(k+1)} = -\nabla f(\vec{\theta}_{(k+1)}) + \beta_{(k)} \vec{d}_{(k)}</math>, Where <math>\beta_{(k)} = -\frac{\nabla f(\vec{\theta}_{(k+1)})^T A \vec{d}_{(k)}}{\vec{d}_{(k)}^T A \vec{d}_{(k)}}</math></p> <p>Step 4 : If <math>k=n</math>, then terminate; Otherwise go to Step 2.</p>
--

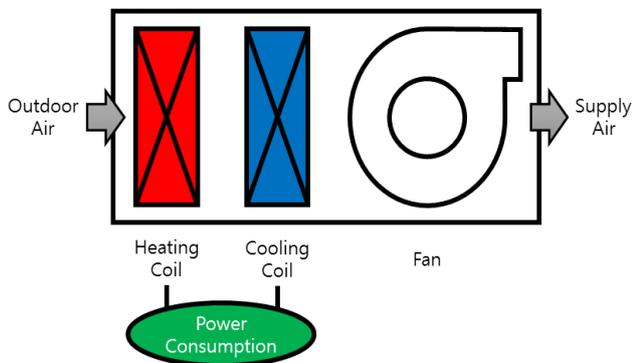
<Figure 1> Pseudo Code for the Generalized Conjugate Gradient Method

## 4. 실험 환경 및 결과

### 4.1 실험 환경

본 연구에서 사용한 산업용 공조기는 일반적으로 공장의 유틸리티실에 위치하고 <Figure 2> 과 같이 크게 히팅 코일(heating coil), 쿨링 코일(cooling coil), 그리고 송풍기(fan)로 구성되어 있다. 온도조절은 보일러를 통해 생성된 스팀이 히팅 코일에 들어가 온열을 발생시키고 냉동기를 통해 생성된 냉수가 쿨링 코일에 들어가 냉열을 발생시켜 공조기 토출 온도를 제어한다. 따라서 <Figure 2> 의 왼쪽으로부터 들어온 외부 공기는 히팅 코일과 쿨링 코일을 거치면서 일정한 온도를 유지하게 되고, 송풍기와 덕트(duct)를 통하여 작업장으로 나가게 된다. 본 연구에서 사용한 데이터는 공조기의 입구(외기) 온도, 히팅 코일 통과 후 온도, 냉각코일 통과 후 온도, 작업장에

공급되는 공기 온도를 <Table 1>과 같은 센서(sensor)와 데이터 로거(data logger)를 통해 수집하였다. 소모 전력을 히팅 및 쿨링 코일에 소모되는 에너지를 전력으로 환산한 값으로 정의하고, 매 시간 소모 전력을 측정하였다. 따라서 본 연구에서 사용한 시계열 데이터는 매시간 측정된 소모 전력이 된다. 기존 공조 설비 분야에서는 주로 최종 소비전력을 측정하고 설비 자체의 개선을 통해 소비전력을 낮출 수 있는 방안 위주로 연구를 진행하였으며, 소모 전력을 예측하는 모형 혹은 방법에 대한 연구는 거의 전무한 것으로 파악된다.



<Figure 2> Schematic Diagram of a Air Handling Unit(AHU)

<Table 1> Specification of Temperature Measure

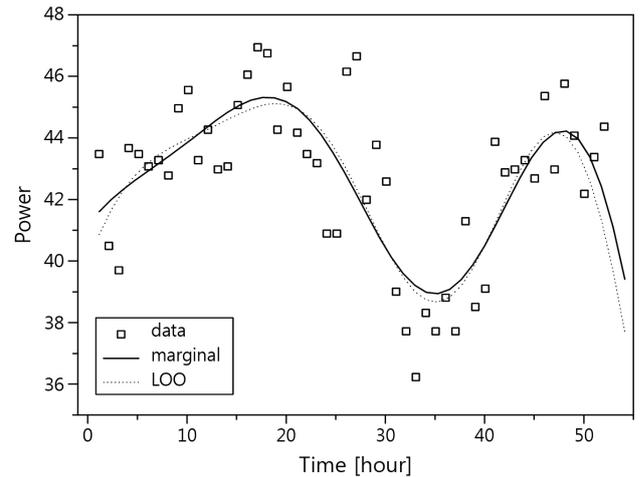
Temperature Measure	Specification
Data Logger	GRAPHTEC midi Logger GL-820
Temperature Sensor	RTD sensor 3 lines(PT-100)

#### 4.2 소모 전력 예측 결과

제 3장에서 언급한 두 가지 방법(주변우도와 교차검정)을 사용하여 공분산 함수를 위한 초월매개변수를 추정하였고, 그 결과를 <Table 2>에 요약하였다. <Table 2>에서 볼 수 있듯이 두 방법으로 추정한 초월매개변수 중  $\sigma_l$ 와  $\sigma_n$ 는 거의 일치하는 반면,  $\sigma_f$ 는 큰 차이를 보였다. 추정한 초월매개변수를 제곱 지수형 공분산 함수에 적용하여 시간에 따른 소모 전력을 예측하였으며, 그 결과를 <Figure 3>에 나타내었다. <Figure 3>의 네모상자(□)는 시간에 따라 측정된 소모 전력 데이터를 나타내며, 실선과 점선은 각각 주변우도와 교차검정을 사용하여 추정한 초월매개변수를 적용하여 구한 소모 전력 예측치의 평균값을 나타낸다. <Figure 3>을 통해 볼 때, 예측 결과는 실제 실험 데이터의 불규칙적인 시계열 경향을 잘 설명하고 있음을 알 수 있으며, 또한  $\sigma_f$ 의 추정 방법 및 추정치에 비교적 무관함을 알 수 있다.

<Table 2> Summary of Estimated Parameters by Using the Marginal Likelihood and the LOO Cross Validation

	$\sigma_f$	$\sigma_l$	$\sigma_n$
Marginal Likelihood	31.147867	1.568892	1.782392
LOO Cross Validation	11.731334	1.205452	1.763017



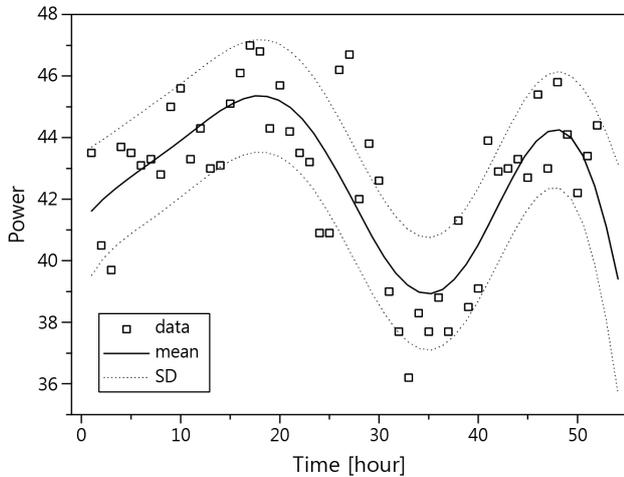
<Figure 3> The Prediction of Power Consumption Versus Time by Using the Marginal Likelihood and the LOO Cross Validation

$\sigma_f$ 의 추정방법에 따라 추정치가 큰 차이가 있음에도 불구하고 예측 결과는 추정방법에 비교적 무관한 것은 다음과 같이 두 가지 측면에서 이해할 수 있다. 첫 번째는 추정치 계산과 관련된 것으로, 주변우도와 교차검정을 사용하여 초월매개변수를 추정하기 위해서는 식 (14)와 식 (18)의 해를 구해야 한다. 식 (14)와 식 (18)은 일반적으로 많은 지역 최적해를 가지고 있고, 본 연구에서 사용한 켈레기울기방법은 그 중에서 하나의 지역 최적해를 찾아주는 방법이다. 따라서 <Table 2>는 두 방법이 서로 다른 지역 최적해를 찾은 결과일 수 있다. 또한 정규 확률과정에서는 많은 지역 최적해 중에서 특정한 지역 최적해를 찾아서 초월매개변수의 추정치로 사용하는 것이 실제적으로는 큰 문제가 되지 않음이 밝혀졌다[12].

두 번째는 초월매개변수 추정을 위해 실제 데이터를 사용했기 때문에 추정 방법에 비교적 무관한 결과를 얻을 수 있다. 즉, 사전확률에 필요한 초월매개변수를 추정하기 위한 두 방법 모두 실제 데이터를 사용했기 때문에 초월매개변수 값은 실험 데이터에 영향을 받을 수밖에 없다. 즉, 두 방법 모두 데이터에 적합하도록 초월매개변수를 추정했기 때문에 추정 방법이 다르므로 인해 추정치는 다를 수 있으나, 베이저안 관점에서 보면 사후확률은 거의 차이가 없다. 또한 공분산함수는 사전확률을 구

성하는 요소이며, 임의의 함수를 생성할 때 서로 다른 시간 사이의 함수 값에 대한 분산을 나타낸다. 따라서 비록 생성된 함수들은 초월매개변수의 추정치에 따라 다르지만, 예측치는 생성된 함수 중에서 우도에 해당하는 데이터를 만족하는 함수들로 구성된다. 이러한 이유로 인하여 사후확률을 통한 예측치는 초월매개변수의 추정에 비교적 영향을 덜 받음을 알 수 있다.

정규 확률과정의 특징 중 하나는 각 독립변수에 따른 예측이 독립변수 값에 의존하는 매개변수를 가지는 확률 분포로 주어진다. 즉, 정규 확률 과정에서는 식 (9)에서 볼 수 있듯이 예측  $f_*$ 가 독립변수인 시간에 따라 평균과 분산이 다른 정규 분포를 따르기 때문에 예측치의 확률분포를 사용하여 예측치와 관련된 다양한 통계량 (예를 들어 신뢰구간 등)을 구할 수 있다. <Figure 4>는 소모 전력 예측치의 95% 신뢰구간을 평균과 표준편차를 사용하여 그림으로 나타낸 것이다. <Figure 4>를 통해 볼 때, 실험 데이터는 예측치에 대한 95% 신뢰구간에 거의 모두 포함되어 있음을 알 수 있고, 이것은 본 연구에서 구한 모형의 유의성을 입증하는 결과라 할 수 있다.



<Figure 4> The Prediction of Power Consumption Versus Time with the Confidence Interval by Using the Mean and the Standard Deviation

### 4.3 예측 결과의 유의성 검정

정규 확률과정을 통한 예측 결과의 유의성을 평균상대오차와 평균절대오차 등 예측 오차를 통하여 조사하였다. 평균상대오차(mean relative error, MRE)와 평균절대오차(mean absolute error, MAE)는 각각

$$MRE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{A_t}, \quad MAE = \frac{1}{n} \sum_{t=1}^n |A_t - F_t| \quad (19)$$

로 주어지는데, 여기서  $A_t$ 와  $F_t$ 는 각각 시간  $t$ 에서의 테스트 데이터 값과 예측치를 나타내고, 그리고  $n$ 은 데이터 개수이다. 본 실험에서 구한 데이터에서 훈련 데이터를 제외한 데이터 중에서 훈련 데이터의 개수와 동일한 개수로 구성된 테스트 데이터 집합 10개를 임의로 추출하여 평균상대오차와 평균절대오차를 계산하였으며, 그 결과를 <Table 3>에 나타내었다.

<Table 3>에서 볼 수 있듯이 평균상대오차는 약 0.034로 이것은 예측치의 약 3.4%가 예측오류임을 의미하며, 데이터 집합에 무관하게 비교적 일정한 값을 알 수 있다. 평균절대오차의 경우에도 데이터 집합에 따른 편차가 크지 않으며, 평균값은 약 1.430이다. 이 값은 정규 확률과정을 통해 구한 예측치 확률분포의 표준편차인 1.813보다 작은 값임으로, 평균절대오차가 예측 모형의 표준편차 내에 포함됨을 의미한다. 이와 같은 예측 결과를 통해 볼 때, 정규 확률과정을 적용한 예측은 비교적 만족스러움을 알 수 있다.

<Table 3> The Result of Prediction Errors

Data set	MRE	MAE
1	0.033	1.388
2	0.034	1.446
3	0.038	1.590
4	0.030	1.259
5	0.035	1.501
6	0.037	1.603
7	0.030	1.302
8	0.034	1.411
9	0.033	1.393
10	0.034	1.407
Average	0.034	1.430

추정 모형에 대한 적합도를 아래와 같은 가설검정을 통하여 고려하였다. 정규 확률과정에서 예측치는 독립변수인 시간에 따라 변하는 평균과 분산을 가지는 정규 확률분포로 주어지기 때문에 예측치를 정규 확률분포의 평균으로 간주하면 동일한 시간에 대하여 같은 개수의 훈련 데이터와 예측치를 사용하여 가설검정을 수행할 수 있다. 또한 각 예측치는 동일한 평균과 분산을 가지는 확률분포에서 추출한 표본이 아니며 훈련 데이터가 예측치의 신뢰구간에 대부분 포함되기 때문에, 비모수 가설 검정 중에서 예측치를 중심으로 훈련 데이터가 고르게 분포됨을 Wilcoxon의 부호-순위 검정(singed-rank test)을 사용하여 추정 모형에 대한 적합도 검정을 실시할 수 있다. 이때 귀무가설은 “훈련 데이터와 예측치의 차이가 0을 중심으로 고르게 분포되어 있다”이다.  $n$ 개의 훈련 데이터  $y_{1,i}$ 와 예측치  $y_{2,i}$ 에 대하여  $i$ 번째 데이터 쌍의 절대값 순

위를  $R_i$ 라 하면, 귀무가설 하에 검정통계량  $W$ 는

$$W = \sum_{i=1}^n \text{sgn}(y_{1,i} - y_{2,i}) R_i \quad (20)$$

로 주어지며, 여기서  $\text{sgn}(y_{1,i} - y_{2,i})$ 는  $y_{1,i} - y_{2,i}$ 의 부호를 나타낸다.

표본의 개수가  $n \geq 15$ 인 경우, 귀무가설 하에 검정통계량  $W$ 는 근사적으로 평균이 0이고 분산이  $n(n+1)(2n+1)/6$ 인 정규분포를 따른다. 즉,

$$W \sim N(0, n(n+1)(2n+1)/6) \quad (21)$$

이다.  $n = 50$ 인 데이터를 사용하여 계산한 검정통계량은  $W_0 = 1.457$ 이며, 유의 수준  $\alpha = 0.05$ 에서 기각역은  $|W_0| > 1.645$ 임으로 귀무가설을 채택할 수 있다. 즉, 훈련 데이터와 예측치의 차이는 0을 중심으로 고르게 분포되어 있고, 또한 훈련 데이터가 예측치의 신뢰구간에 대부분 포함되기 때문에 추정 모형은 적합하다고 할 수 있다.

## 5. 요약 및 결론

본 논문에서는 공조기의 효율적 에너지 관리의 필요성과 이에 따른 공조기에 소요되는 전력 예측의 중요성에 입각하여, 공조 시스템에서 측정된 소모 전력 데이터를 사용하여 불규칙적인 소모 전력을 예측하는 모형을 제안하였다. 제안한 예측 모형은 정규 확률과정을 사용한 것으로 기존의 선형 회귀 등과 같이 회귀 모형을 위한 함수 유형을 제한하는 방법의 단점을 보완한 것이다. 특히 제안한 모형은 베이지안 관점에서 사전확률로 정규 확률과정을 통해 생성한 함수를 사용하고 실험 데이터를 우도로 하여 예측치를 사후확률의 형태로 구하는 모형이다. 정규 확률과정을 적용한 예측 모형은 시간에 따라 불규칙적으로 변화하는 공조기의 소모 전력 시계열 데이터 예측에 적용하기 적합한 모형이며, 공분산 함수로 제공 지수형을 사용하였고, 이에 따른 초월매개변수를 주변우도와 교차검정 방법을 적용하여 추정하였다.

제안한 예측 모형을 실험 데이터를 적용하여 구한 예측 결과는 실제 실험 데이터의 시계열적 경향을 비교적 잘 설명하고 있음을 알 수 있었으며, 또한 예측 결과는 초월매개변수의 추정 방법 및 추정치에 비교적 무관함을 알 수 있었다. 특히 제안한 예측 모형을 통해 예측치에 대한 확률분포를 구할 수 있으므로 이를 통해 예측치와 관련된 다양한 통계량을 구할 수 있었다. 그 예로 예측치에 대한 신뢰 구간을 구한 결과, 실험 데이터는 예측치에 대한 95% 신뢰구간에 거의 모두 포함되어 있음을 알 수

있고, 이것은 본 연구에 적용한 모형의 유의성을 입증하는 결과라 할 수 있었다.

본 연구에서는 시간을 독립변수로 간주하고 정규 확률과정을 사용하여 소모 전력을 예측하였다. 그러나 본 연구에서 고려한 공조 시스템은 시간 외에 송풍기, 히팅 및 쿨링 코일 등의 제어 가능한 변수들이 있다. 따라서 완전한 에너지 관리를 위해서는 이러한 변수들을 추가로 포함하여 소모 전력을 예측해야 한다. 그러나 다변량(multivariate)이며 각 변수들의 특성이 서로 상이한 경우의 공분산 함수에 대한 연구는 아직 미비한 실정이며, 특히 본 연구의 경우에는 송풍기의 속도, 히팅 및 쿨링 코일의 온도 등 각 변수 값의 단위가 다르기 때문에 이러한 특성을 고려한 공분산 함수에 대한 연구가 선행되어야 한다. 따라서 향후 시간 외의 제어 가능한 모든 변수를 포함하는 공분산 함수에 대한 연구를 수행해야 할 필요성이 있다고 판단된다.

본 연구의 결과는 향후 시계열 데이터 예측 모형과 공조기의 소모 전력 상황을 비교 검토하고, 예측 모형의 신뢰성 검증을 통해 자동차 건조 공정에 사용되는 여러 공조기들의 에너지 소비 패턴을 예측하여 관리하는 데 적용할 수 있을 것으로 판단된다.

## Acknowledgement

This work was supported by the Power Generation and Electricity Delivery Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry and Energy, Republic of Korea (No. 2013 1020400850).

## References

- [1] Andrieu, C., Freitas, N., Doucet, A., and Jordan, M., An Introduction to MCMC for Machine Learning, *Machine Learning*, 2003, Vol. 50, No. 1, pp. 5-43.
- [2] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 1986, Vol. 31, No. 3, pp. 307-327.
- [3] Chalupka, K., Williams, C., and Murray, I., A Framework for Evaluating Approximation Methods for Gaussian Process Regression, *Journal of Machine Learning Research*, 2013, Vol. 14, No. 1, pp. 333-350.
- [4] Engle, R., Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 1982, Vol. 50, No. 4, pp.

- 987-1007.
- [5] Golub, G. and Van Loan, C., *Matrix Computations*, The John Hopkins University Press, third edition, 1996.
- [6] Han, D. and Kim, J., Implementation of the Real-time Fault Detection and Diagnosis System of Air Handling Unit. Proceedings of the SAREK 2002 Summer Annual Conference, 2002, pp. 301-306.
- [7] Han, D. and Youn, H., Building Energy Control Algorithms by Using Outdoor Air Temperature Prediction, Proceedings of the SAREK 2002 Summer Annual Conference, 2002, pp. 345-350.
- [8] Kim, J. and Kim, T., Screening and Clustering for time-course Yeast Microarray Gene Expression Data using Gaussian Process Regression, *The Korean Journal of Applied Statistics*, 2013, Vol. 26, No. 3, pp. 389-399.
- [9] Lee, P., *Bayesian Statistics : An Introduction*, 4th Edition, John Wiley and Sons, Chichester, 2012.
- [10] Lim, S., Cho, S., and Lee, C., Model for the Spatial Time Series Data, *Journal of the Korean Society for Quality Management*, 1996, Vol. 24, No. 1, pp. 137-145.
- [11] Mills, T., *Time Series Techniques for Economists*, Cambridge University Press, New York, 1990.
- [12] Rasmussen, C. and Williams, C., *Gaussian Processes for Machine Learning*, the MIT Press, Cambridge, Massachusetts 2006, ch. 5.
- [13] Suh, J., Lee, S., Oh, H., Koo, J., Lim, T., and Cho, J., ARMA-PL : Tackling Nested Periods and Linear Trend in Time Series Data, *Journal of the Society of Korea Industrial and Systems Engineering*, 2010, Vol. 33, pp. 112-126.
- [14] Sundararajan, S. and Keerthi, S., Predictive Approaches for Choosing Hyperparameters in Gaussian Processes, *Neural Computation*, 2001, Vol. 13, No. 5, pp. 1103-1118.
- [15] Vijayakumar, S., D'Souza, A., Shibata, T., Conradt, J., and Schaal, S., Statistical Learning for Humanoid Robots, *Autonomous Robot*, 2002, Vol. 12, No. 1, pp. 55-69.

#### ORCID

Chang-Yong Lee | <http://orcid.org/0000-0003-1778-6532>

Gensoo Song | <http://orcid.org/0000-0002-3185-2965>

Jinho Kim | <http://orcid.org/0000-0003-4243-7318>