

Correlation Analyses of the Temperature Time Series Data from the Heat Box for Energy Modeling in the Automobile Drying Process

Chang-Yong Lee* · Gensoo Song** · Jinho Kim*[†]

*Dept. of Industrial and Systems Engineering, Kongju National University
**R&D Center of Duksan A.C.M. Co. LTD.

자동차 건조 공정 에너지 예측 모형을 위한 공조기 온도 시계열 데이터의 상관관계 분석

이창용* · 송근수** · 김진호*[†]

*공주대학교 산업시스템공학과
**덕산공조(주) 기업부설연구소

In this paper, we investigate the statistical correlation of the time series for temperature measured at the heat box in the automobile drying process. We show, in terms of the sample variance, that a significant non-linear correlation exists in the time series that consist of absolute temperature changes. To investigate further the non-linear correlation, we utilize the volatility, an important concept in the financial market, and induce volatility time series from absolute temperature changes. We analyze the time series of volatilities in terms of the de-trended fluctuation analysis (DFA), a method especially suitable for testing the long-range correlation of non-stationary data, from the correlation perspective. We uncover that the volatility exhibits a long-range correlation regardless of the window size. We also analyze the cross correlation between two (inlet and outlet) volatility time series to characterize any correlation between the two, and disclose the dependence of the correlation strength on the time lag. These results can contribute as important factors to the modeling of forecasting and management of the heat box's temperature.

Keywords : Time series, Correlation, Volatility, Long-range Correlation

1. 서론

최근 차량 구매자의 구매 성향은 가격과 품질에 더하여 다양한 차종을 선택하는 방향으로 바뀌고 있다. 이러한 구매 성향의 변화에 따라 최근 상용 차량은 개발 시간 단축, 다품종, 그리고 고품질 생산이 요구되며, 이를 위한 생산 라인을 구축하여 소비자의 요구에 부응하는

노력을 기울이고 있다. 따라서 자동차 생산 라인 구축의 설계 및 제작에 있어, 각각의 공정 특성 분석이 반드시 고려되어야 한다.

생산 공정 중 차량 도장 공정은 제품의 미관을 높이고 기계적·화학적 성질을 개선시키는 매우 중요한 공정이다. 일반적으로 차량 도장 공정은 주로 전착 도장 공정(electro deposition process), 스프레이 도장 공정(spray painting process), 그리고 도장 건조 공정(paint drying process) 순으로 이루어진다. 이중 도장 건조 공정은 전착 도장 공

정 및 스프레이 도장 공정에서 입혀진 도료를 차량 표면에 정착시키는 공정으로 건조가 제대로 되지 않는 경우 전체 공정의 불량률 초래할 수 있다. 또한 건조 공정의 불량률은 후속 공정에 문제를 야기하거나, 외관 및 차량의 내구성 저하로 연결되기 때문에, 생산성 및 생산 비용은 물론 생산 품질에 직접적인 영향을 끼치게 된다. 이러한 이유로 건조 공정에 대한 분석은 생산 공정에서 매우 중요한 역할을 한다[4].

건조 방식에 따라 건조 공정은 복사 건조 방법과 대류 건조 방법 등 두 가지로 나뉜다. 자동차 건조 공정의 경우에는 상대적으로 피 건조체의 체적인 큰 자동차의 차체 전체를 건조시켜야 하기 때문에 복사 건조 방법보다 에너지 소비가 비교적 적은 대류 건조 방식의 일종인 열풍 건조 공정이 많이 사용된다[17]. 열풍 건조 공정에서 널리 사용되는 공조기(heat box)는 주위의 공기를 흡수한 후 코일로 가열하여 주변보다 상대적으로 높아진 더운 공기를 내보내는 역할을 한다.

공조기를 사용한 열풍 건조 공정에 소요되는 에너지 수요 예측은 효율적 에너지 관리 측면에서 매우 중요하다. 수요 예측을 통한 에너지 관리를 위해서는 에너지의 수요를 예측할 수 있는 모형의 구축이 필요하다[18]. 건조 공정에 필요한 에너지 수요 및 관리에 대한 연구는 거의 수행되지 않고 있는 실정이나, 일반적인 형태의 에너지 수요 예측 및 관리를 위한 연구를 통하여 건조 공정에 응용할 수 있다. Faranda 등[7]은 에너지 관리를 위한 조기 경보 체계를 ARMA(Auto-Regressive Moving-Average Processes)와 같은 통계적 모형을 사용하여 제안하였으며, Lee 등[11]은 예측에 필요한 변수를 최소화하는 과정을 거쳐 지지도 벡터 기계(SVM, Support Vector Machine)를 사용하여 전력 수요 예측 모형을 제안하였다. 또한 Kim 등[8]은 교차반응함수를 사용하여 에너지 상대가격의 변화에 따른 대체 수요의 변화를 효율적으로 반영하는 에너지 수요 예측 모형을 제시하였다. Kong 등[9]은 광역 단지에서 필요한 에너지 수요 예측 기법을 개발하기 위해 표준 모델을 사용하여 데이터베이스를 구축하고 인공지능 경망을 사용하여 에너지 수요 예측 기법을 개발하였다.

건조 공정에 소요되는 에너지의 수요 예측을 위해서는 시간에 따른 공조기의 온도 변화를 예측하는 것이 필요하다. 공조기의 온도 변화는 시계열 데이터로 표현되며, 시계열 데이터의 예측은 주로 ARMA[13], ARCH(Auto-regressive Conditional Heteroscedasticity)[6], 혹은 ARCH를 일반화한 GARCH[3] 등이 사용된다. 또한 최근에는 선형성과 주기성을 가진 요소가 확률적인 분포와 혼재되어 있는 시계열 데이터의 ARMA 분석에서 선형적 주기적 요소를 제거하는 방법론에 대한 연구도 수행되었다

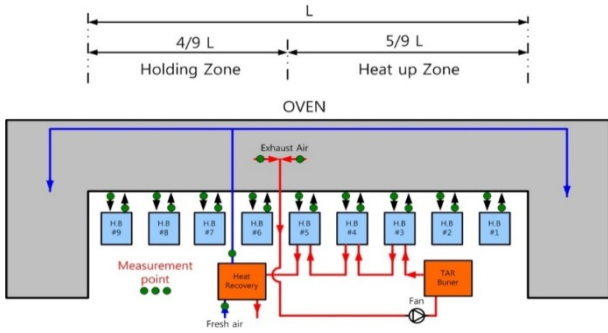
[19]. 이러한 다양한 모형들을 적용하기 위해서는 시계열 데이터 자체에 내재한 시계열적 특성을 분석하는 것이 선행되어야 하며, 공조기의 온도 변화 예측을 위해서도 공조기 온도 시계열 데이터를 다양한 각도에서 분석하여 그 특성을 파악하는 것이 선행되어야 한다.

본 연구에서는 이러한 배경을 고려하여 공조기 온도에 대한 시계열 데이터를 다양한 통계학적 방법을 적용하여 분석하여 시계열 데이터에 내재한 특성을 규명하였다. 특히 표본 분산을 적용하여 온도의 절대 변화량 사이에는 비선형 상관관계가 존재함을 보였고, 온도 변화에 대한 변동성(volatility) 시계열 데이터 사이에는 장거리 상관관계(long-range correlation)가 내재함을 규명하였다. 이러한 특성들은 향후 효율적 에너지 관리를 위한 공조기 온도의 변화를 예측하는 모형 구축에 적용할 수 있으며, 본 연구 결과의 적용에 대한 예시로 시계열 데이터 예측에 널리 사용되는 ARMA와 GARCH 모형에 필요한 매개변수 결정 방법에 대하여 개략적으로 설명하였다. 또한 본 연구의 결과를 바탕으로 향후 예측 모형과 실제 공조기의 에너지 소비 상황을 비교 및 검토를 하여 예측 모형의 신뢰성을 검증한 후, 자동차 건조 공정에 사용되는 여러 공조기들의 에너지 소비 패턴을 에너지 소비 패턴을 예측하여 관리하는 데 적용할 수 있을 것으로 판단된다.

본 논문은 다음과 같이 구성되어 있다. 제 2장에서는 공조기 온도 시계열 데이터 생성을 위한 실험 환경 및 방법을 언급하였으며, 제 3장에서는 공조기 온도 시계열 데이터의 통계적 분석을 통하여 비선형 상관관계를 규명하였다. 제 4장에서는 비선형 상관관계를 좀 더 세부적으로 분석하기 위하여 변동성 시계열 데이터를 유도하여 변동성 사이에 장거리 상관관계가 존재함을 보이고, 그 의미를 분석하였으며, 제 5장에서는 본 논문의 결론을 맺었다.

2. 실험 환경 및 방법

실험을 위한 건조 방식을 <Figure 1>에 도식화하여 나타내었다. 실제 공정에 따라 공정 길이의 차이가 있지만, 일반적으로 피 건조체를 오븐(oven)의 오른쪽에서 왼쪽으로 진행시키면서 먼저 heat up zone을 거치는 동안 피 건조체의 온도를 빠르게 상승시키고, holding zone을 거치면서 도료가 피 건조체에 정착되도록 한다. 오븐에 열풍을 공급시켜주는 공조기는 오븐 아래에 위치하여 있으며 오븐 내부의 공기를 재순환시켜주면서 오븐에 열풍을 공급한다.



<Figure 1> Schematic Diagram of Drying Process of Electro Deposition

공조기에서 토출(discharge)된 뜨거운 공기가 오븐으로 들어가면 오븐에서 건조를 시키고 온도가 떨어진 상태로 다시 공조기로 들어가게 된다. 이때 공조기에서 토출되어 오븐으로 들어가는 부분을 outlet, 오븐에서 다시 공조기로 들어가는 부분을 inlet이라 부른다. 또한 공조기의 inlet과 outlet 사이에 버너(burner)가 있어서 inlet을 통해 들어온 공기를 가열하여 outlet으로 보낸다. 공조기 inlet과 outlet의 온도 측정은 공조기와 오븐 사이의 순환 덕트에 점검구를 설치하여 계측하였으며, 온도 측정을 위한 계측기의 사양은 <Table 1>에 수록하였다. 오븐에 따라 설계되는 내부 온도가 약간의 차이는 있지만 거의 비슷한 온도의 범위 내에 있으므로, 통계 분석에 사용한 온도 측정 데이터가 공조기 입출구의 온도를 대표한다고 가정할 수 있다.

<Table 1> Specification of Temperature Measure.

Data logger	Graphtec GL 820
Temperature sensor	Sensor : PT100 Ω
	Accuracy : 0.15°C ± 0.002
	Range : -200°C ~ 600°C

본 연구에서 사용한 데이터는 2014년 1월 30일 울산에 위치한 건조 오븐용 공조기 9대(공조기 #1~9)에 대하여 측정한 온도 시계열 데이터이며, 초기 가동해서 정상 상태(즉, 온도가 일정한 상태)에 도달할 때까지 공조기 9대의 inlet과 outlet에서 온도를 측정하였다. 분석 결과는 공조기에 비교적 무관하였으므로 본 연구에서는 공조기 #1에 초점을 맞추어 공조기 #1의 inlet과 outlet 온도 시계열 데이터를 분석하여 그 특성을 살펴보고자 한다.

3. 시계열 데이터의 비선형 상관관계

Inlet 혹은 outlet에서 측정한 온도 시계열 데이터를 $z(1)$,

$z(2), \dots$,라 하면, 시간 t 에서 온도의 절대 변화량 $x(t)$ 는 다음과 같이 주어진다. 즉,

$$x(t) \equiv |z(t + \Delta t) - z(t)| \quad (1)$$

이다. 여기서 $\Delta t = 1$ 초이며, 절대 변화량 $x(t)$ 역시 시계열 데이터를 이룬다.

N 개의 시계열 데이터 $x(1), x(2), \dots, x(N)$ 이 주어졌을 때, 서로 다른 시간에서 시계열 데이터(즉, 서로 다른 시간에서 절대 변화량) 사이의 상관관계를 살펴보기 위하여 자기상관계수(auto-correlation coefficient)를 고려하였다. 시간 지연(time lag)이 τ 로 주어진 경우 자기상관계수 $\rho(\tau)$ 는

$$\rho(\tau) \equiv \frac{\sum_{t=1}^{N-\tau} \{x(t) - \bar{x}_t\} \{x(t+\tau) - \bar{x}_{t+\tau}\}}{\sqrt{\sum_{t=1}^{N-\tau} \{x(t) - \bar{x}_t\}^2} \sqrt{\sum_{t=1}^{N-\tau} \{x(t+\tau) - \bar{x}_{t+\tau}\}^2}} \quad (2)$$

로 주어지며, 여기서 $\bar{x}_t \equiv \frac{1}{N-\tau} \sum_{t=1}^{N-\tau} x(t)$, $\bar{x}_{t+\tau} \equiv \frac{1}{N-\tau} \sum_{t=1}^{N-\tau} x(t+\tau)$ 이다. 자기상관계수는 시계열 데이터 사이의 선형 상관(linear correlation) 정도를 살펴보는 양으로 $-1 \leq \rho(\tau) \leq 1$ 로 주어진다. 특히 $\rho(\tau) = 1$ 인 경우는 완전 상관, $\rho(\tau) = -1$ 인 경우는 완전 반상관(anti-correlation), 그리고 $\rho(\tau) = 0$ 인 경우에는 상관관계가 존재하지 않음을 의미한다.

자기상관계수는 주로 선형 상관 유무를 조사하는데 반하여, 비선형 (혹은 고차원) 상관관계를 분석하는 방법으로 표본 분산(sample variance)을 들 수 있다. 표본 분산은 표본 평균에 대한 분산으로 주어진 N 개의 시계열 데이터를 n 개의 부분집합으로 나누고 각 부분집합에 속하는 표본에 대하여 구한 n 개의 표본 평균에 대하여 분산을 계산한다. 부분집합의 개수 n 에 대한 표본 분산은

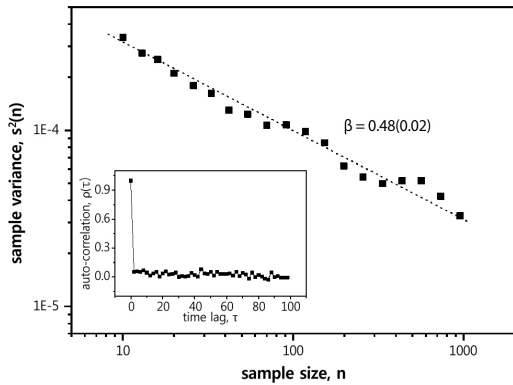
$$s^2(n) \equiv \frac{1}{n-1} \sum_{k=0}^{n-1} (\bar{x}_{m,k} - \bar{\bar{x}})^2 \quad (3)$$

으로 주어지며, 여기서 $N = m \times n$ 이며, $\bar{x}_{m,k} \equiv \frac{1}{m} \sum_{j=1}^m x(km + j)$ 그리고 $\bar{\bar{x}} \equiv \frac{1}{n} \sum_{k=1}^n \bar{x}_{m,k}$ 이다. 만약 데이터가 서로 독립이라면 중심극한정리(central limit theorem)에 의하여 표본 분산은 부분집합 개수 n 에 대하여 멱급수(power-law) 형태를 띠며

$$s^2(n) \propto n^{-\beta}, \quad \beta = 1 \quad (4)$$

이 된다[1].

Inlet 온도 n 시계열 데이터를 사용하여 시간 지연 τ 와 표본 크기에 대하여 자기상관계수와 표본 분산을 구하였으며, 그 결과를 <Figure 2>에 나타내었다. <Figure 2>의 삽입 그림(inset)에서 볼 수 있듯이 자기 상관계수는 시간 지연이 커짐($\tau > 1$)에 따라 급속히 감소하여 $\rho(\tau) \approx 0$ 에 이른다. 이것은 고려하는 시계열 데이터 사이에 선형적인 상관관계가 거의 없음을 나타낸다. 그러나 표본 분산의 결과를 보면 $\beta \approx 0.48 \pm 0.02$ 로 $\beta = 1$ 보다 작은 값을 가진다. 이것은 표본 분산이 n 이 커짐에 따라 데이터들이 상관관계가 없는 독립인 경우보다 느리게 감소함을 의미하며, 따라서 유의한 상관관계가 시계열 데이터에 존재함을 나타낸다. 이러한 결과를 통해 볼 때 고려하는 시계열 데이터 사이에는 선형 상관관계는 거의 존재하지 않으나 비선형(혹은 고차원) 상관관계는 존재함을 알 수 있다. Outlet의 경우에도 유사한 결과를 얻을 수 있었다.



<Figure 2> Log-log Plot of the Sample Variance, $s^2(n)$, for Various Sample Sizes n . The Estimated β is $\beta = 0.48 \pm 0.02$ (Dotted Line). β is the Estimate using Lease-Square fit with the error in the Parenthesis. The Inset Represents the Auto-Correlation Function for with Respect to the Time Lag τ

4. 변동성에 대한 상관관계 분석

4.1 장거리 상관관계 분석

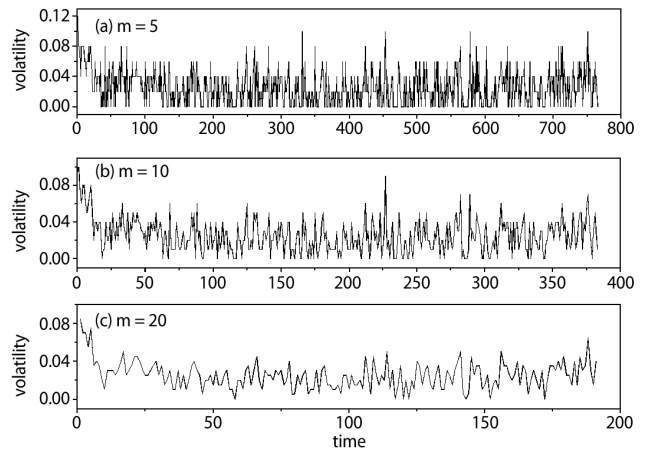
공조기 온도 시계열 데이터에 대한 절대 변화량들 사이에 비선형 상관관계가 존재함에 기초하여 이를 좀 더 면밀히 조사하기 위하여 변동성(volatility)[12, 14]에 대한 분석을 수행하였다. 변동성은 주로 주식 시장에서 주가의 변화를 연구하는데 사용되는 양으로 주식 시장의 섭동(fluctuation)을 양적으로 표현한 것이다. 변동성은 주식 시장으로 유입되는 정보의 양과 밀접한 관계가 있다. 즉,

주식 시장으로 유입되는 정보가 많을수록 거래량은 일반적으로 늘어나며, 따라서 큰 주가 변동을 야기한다. 특히 변동성은 Black-Scholes 모형[2]이나 Cox-Ross-Rubinstein 모형[5]과 같은 옵션가격 모형(option pricing model)의 설정에 중요한 역할을 한다.

변동성을 계량적으로 표현하기 위해 널리 사용하는 방법으로 조절 가능한 윈도우 크기(window size)에 대하여 절대 변화량에 대한 평균치를 들 수 있다. 즉, 절대 변화량으로 구성된 N 개의 시계열 데이터 $x(1), x(2), \dots, x(N)$ 를 표본 분산의 경우와 유사하게 n 개의 부분집합으로 나누어 각 부분집합이 m 개의 데이터로 구성되도록 즉, $N = m \times n$ 이 되도록 한다. 이렇게 하면 윈도우 크기가 m 인 경우에 대하여 n 개의 변동성을 구할 수 있다. 즉, 윈도우 크기가 m 인 경우 k 번째 변동성 $v_m(k)$ 는

$$v_m(k) \equiv \frac{1}{m} \sum_{t=mk}^{mk+m-1} x(t), \quad k=0, 1, \dots, n-1 \quad (5)$$

로 표현할 수 있다. 따라서 n 개의 변동성 $v_m(1), v_m(2), \dots, v_m(n-1)$ 은 절대 변화량을 사용하여 유도한 시계열 데이터가 된다. <Figure 3>은 윈도우 크기 $m = 5, 10, 20$ 에 대하여 구한 변동성 시계열 데이터를 나타낸 것이다. <Figure 3>을 통해 알 수 있듯이 윈도우가 클수록 많은 절대 변화량에 대한 평균을 취하기 때문에 지역적 섭동(local fluctuation)은 줄어들고, 시계열 데이터 개수는 감소한다.



<Figure 3> Plots of Time Series of the Volatility for the window Sizes (a) $m = 5$, (b) $m = 10$, and (c) $m = 20$

다양한 윈도우 크기에 대하여 구한 변동성 시계열 데이터에 대한 상관관계를 조사하기 위하여 Detrended Fluctuation Analysis(DFA)[10, 15]를 적용하였다. DFA는 일반적인 평균자승평방근(root-mean-square) 분석을 수정한 것

으로 시계열 데이터에 내재할 가능성이 있는 지역적 경향(local trend)를 효율적으로 제거하고 장거리 상관관계(long-range correlation)를 규명하기 위하여 널리 사용되는 방법이다. DFA는 시계열 데이터에 존재하는 장거리 상관관계를 발견할 수 있을 뿐만 아니라 지역적 경향을 제거할 수 있기 때문에 인위적인 요인으로 인한 허위 상관을 제외할 수 있는 장점이 있다. 또한 DFA는 심리학적 신호 및 다양한 금융 데이터의 장거리 상관관계 규명에 적용되고 있다[16]. 주어진 시계열 데이터 $y(1), y(2), \dots, y(N)$ 에 대하여 지역적 경향을 제거하는 상자 크기(box size)가 T 일 때, $M = N/T$ 개의 detrended fluctuation function $F_k(T)$ 는

$$F_k(T) = \sqrt{\frac{1}{T} \sum_{i=kT+1}^{kT+T} \{y(i) - z_{k,T}\}^2}, \quad k=0, 1, \dots, M-1 \quad (6)$$

로 주어진다. 여기서 $z_{k,T}$ 는 $y(kT+1), y(kT+2), \dots, y(kT+T)$ 를 최소자승법(least square)을 사용하여 적합한 것으로 상자 크기가 T 인 k 번째 지역적 경향을 나타낸다. M 개의 $F_k(T)$ 에 대한 평균을 $\langle F(T) \rangle$ 라 두면(즉, $\langle F(T) \rangle = \frac{1}{M} \sum_{k=0}^{M-1} F_k(T)$ 라 하면), 데이터의 섭동을 박스 크기 T 에 대한 함수로 표현할 수 있다. 특히 $\langle F(T) \rangle$ 가 상자 크기 T 에 대하여 축척(scaling)에 무관한 성질을 가지고 있는 경우에는

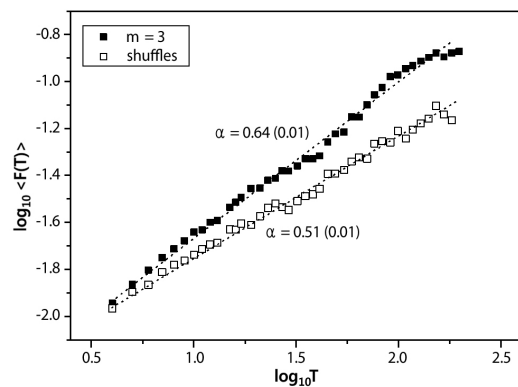
$$\langle F(T) \rangle \propto T^\alpha \quad (7)$$

로 표현할 수 있다. 여기서 α 는 축척 지수로 양의 값을 가지며, 지수 α 의 값을 추정하기 위해서는 위의 식 (7)의 양변에 로그(log)를 취한다. 축척 지수 α 의 값을 사용하여 고려하는 시계열 데이터의 특성을 구별할 수 있다. 고려하는 시계열 데이터들이 통계적으로 서로 독립이어서 데이터 사이에 상관관계가 존재하지 않으면 $\alpha = 0.5$ 이며, $0.5 < \alpha < 1$ 인 경우에는 데이터들 사이에 지속적인 장거리 상관관계가 존재하고, $0 < \alpha < 0.5$ 인 경우에는 장거리 반상관관계가 존재함을 의미한다. 특히 $\alpha = 1$ 인 경우에는 소위 $1/f$ 노이즈(noise)라 불리며, $\alpha > 1$ 인 경우에는 상관관계가 존재하나 멱급수적은 아님으로 장거리 상관관계가 존재하지 않음을 의미한다[10].

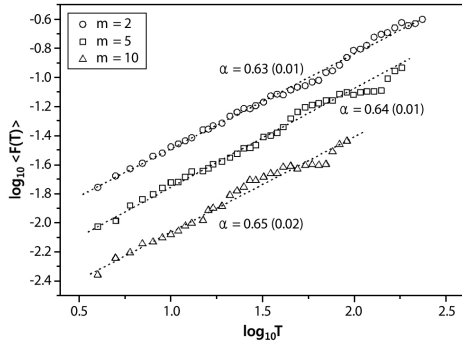
윈도 크기가 $m=3$ 인 변동성 시계열 데이터 $v_m(k)$ [식 (5)]에 DFA를 실행하였으며, 그 결과를 <Figure 4>에 나타내었다. <Figure 4>에서 볼 수 있듯이 DFA의 축척 지수는 $\alpha = 0.64$ 으로 $0.5 < \alpha < 1$ 임으로 변동성 시계열 데이터 사이에는 장거리 상관관계가 존재함을 알 수 있다. 장거리 상관관계가 허위(artifact)가 아님을 입증하기 위

하여 주어진 변동성 시계열 데이터의 순서를 임의로 섞어서 구한 (즉, 시계열 데이터에서 시간적 요소를 제거한) 데이터에 대하여 같은 DFA를 실행하였으며 이 경우에는 $\alpha = 0.51$ 을 얻었다. 임의로 섞은 데이터에 대한 축척 지수는 $\alpha \approx 0.5$ 임으로 상관관계가 존재하지 않음을 알 수 있다. 즉, 시간을 고려하지 않은 데이터에 대해서는 장거리 상관관계가 존재하지 않고, 시간이 고려된 변동성 시계열 데이터에는 장거리 상관관계가 존재함을 확인할 수 있다. 장거리 상관관계의 존재는 변동성의 지속적인 간헐성(burstiness)를 의미한다. 일단 간헐성이 발생하면 이 간헐적인 변동성은 또 다른 간헐적인 변동성을 야기하고 시간이 흐름에 따라 지속적으로 발생한다.

윈도 크기에 따른 장거리 상관관계의 유무 및 정도를 조사하기 위하여 서로 다른 윈도 크기의 변동성 시계열 데이터에 대하여 DFA를 수행하였으며 그 결과를 <Figure 5>에 나타내었다. <Figure 5>에서 볼 수 있듯이 윈도 크기에 무관하게 장거리 상관관계가 존재하며, 윈도 크기가 클수록 축척 지수는 증가하나 큰 차이는 없음을 알 수 있다. 또한 상자 크기 T 가 클수록 섭동이 심한데 이것은 섭동성은 안정적이지 않고 상자 크기에 의존함을 의미한다. 상자 크기는 지역적 경향을 제거하는 시간 범위에 해당함으로 시간에 의존하는 섭동성은 ARCH 혹은 GARCH 등과 같은 시계열 모형을 사용하여 설명할 수 있음을 의미한다. 그 이유는 이러한 모형들은 지역적으로는 정상적(stationary)이지 않고 섭동이 있으나 전역적으로 정상적인 시계열 데이터 모형에 적합하기 때문이다. Outlet 온도 시계열 데이터의 경우에도 inlet와 유사한 특성을 보였다.



<Figure 4> Plots of $\log_{10} \langle F(T) \rangle$ Versus $\log_{10} T$ of the Volatility for window Size $m = 3$. Estimated α is $\alpha = 0.64 \pm 0.01$, Indicating a Long-Range Correlation. The Plot for a Randomly Shuffled Volatility of the Same window Size has $\alpha = 0.51 \pm 0.01$, Indicating no Correlation. The α 's are Estimates Using Least-Square Fits with the Errors in the Parentheses



<Figure 5> Dependence of the Exponent α on the Window Size m by Using Plots of $\log_{10}\langle F(T) \rangle$ Versus $\log_{10}T$

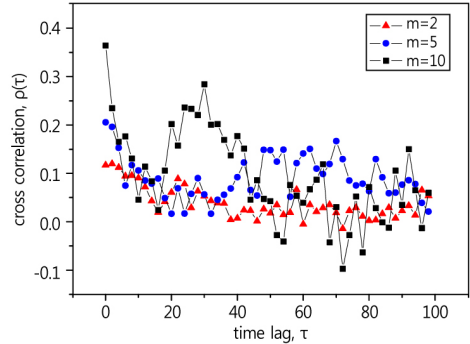
위의 분석을 통해 밝힌 공조기 온도 데이터 사이에 존재하는 장거리 상관관계는 어느 한 시점에서 공조기 온도가 오랜 기간에 걸쳐 다른 시점의 공조기 온도에 영향을 미치는 것을 의미한다. 즉, 시간 t 에서 공조기 온도 $x(t)$ 는 인접한 시간 $t+1$ 혹은 $t+2$ 에서 온도 $x(t+1)$ 과 $x(t+2)$ 뿐만 아니라 시간 차이가 매우 큰(즉, $\tau \gg 1$) $t+\tau$ 시간에서 공조기 온도 $x(t+\tau)$ 에도 영향을 미침을 의미한다. 따라서 추 후 공조기 온도 예측을 위한 모형을 설정할 때 이러한 특성을 고려해야 함을 제시한다.

4.2 Inlet과 outlet 온도 사이의 상관관계 분석

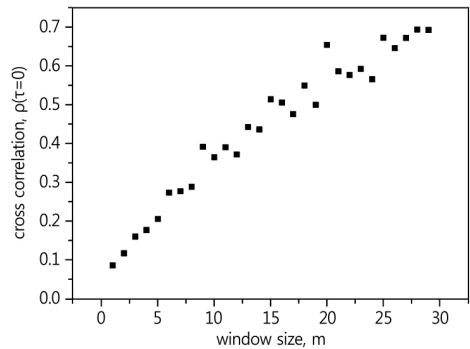
공조기가 오븐에 공급하는 열(heat)의 시간에 따른 변화를 분석하기 위하여 inlet과 outlet에 대한 변동성 시계열 데이터를 사용하여 inlet과 outlet 사이의 상호상관관계(cross correlation) 분석을 수행하였다. 상호상관관계 분석은 제 2장의 자기상관관계(auto-correlation)와 유사하며, 차이점은 서로 다른 시계열 데이터에 대한 분석이라는 점이다. 즉, N 개의 데이터로 구성된 2개의 시계열 데이터 $x(1), x(2), \dots, x(N)$ 과 $y(1), y(2), \dots, y(N)$ 이 주어졌을 때, 시간 지연(time lag)이 τ 인 경우의 상호상관계수 $\rho(\tau)$ 는

$$\rho(\tau) \equiv \frac{\sum_{t=1}^{N-\tau} \{x(t) - \bar{x}_t\} \{y(t+\tau) - \bar{y}_{t+\tau}\}}{\sqrt{\sum_{t=1}^{N-\tau} \{x(t) - \bar{x}_t\}^2} \sqrt{\sum_{t=1}^{N-\tau} \{y(t+\tau) - \bar{y}_{t+\tau}\}^2}} \quad (8)$$

로 주어진다. 또한 상호상관계수 $\rho(t)$ 의 값에 따른 특성은 자기상관계수의 경우와 동일하다. <Figure 6>는 변동성을 구하기 위한 윈도우 크기 m 과 시간 지연 τ 에 따른 상호상관계수의 변화를 나타낸 것이다. 모든 윈도우 크기에 대하여 시간 지연이 커짐에 따라 상호상관계수는 작아짐을 알 수 있다. 이것은 시간 지연이 클수록 inlet과 outlet 온도 사이에 선형 상관관계가 감소하는 경향을



<Figure 6> Plot of the Cross Correlation Coefficient Versus Time Lag τ between Volatilities from Inlet and Outlet with Different Window Sizes m



<Figure 7> Plot of Cross Correlation Coefficient Versus Different Window Sizes

의미한다. 또한 윈도우 크기가 커짐에 따라 상호 상관 계수의 섭동(fluctuation)이 증가함을 알 수 있다.

공조기 입구 데이터와 출구 데이터의 상호 상관관계 분석을 통해 볼 때, 입구와 출구 온도 사이의 상관관계는 시간 지연이 클수록 선형적으로 감소함을 알 수 있다. 이러한 결과는 공조기 입구를 통해 들어온 뜨거운 공기가 오븐에 머무는 시간이 길수록 입구 온도의 변동과 출구 온도의 변동은 비교적 무관함을 의미한다. 이러한 특성은 입구 및 출구 온도에 대한 예측 모형은 시간 지연에 선형적인 관계를 가지도록 구축되어야 함을 의미한다.

또한 <Figure 6>에서 볼 수 있듯이 시간 지연이 없는 경우(즉, $\tau=0$)에는 윈도우 크기가 클수록 상호상관계수는 커짐을 알 수 있다. 이를 좀 더 자세히 분석하기 위하여 시간 지연이 없는 경우에 대하여 여러 윈도우 크기를 사용하여 구한 inlet과 outlet의 변동성 사이의 상호상관계수를 구하였고, 그 결과를 <Figure 7>에 나타내었다. 시간 지연을 고려하지 않은 경우의 상호상관계수는 윈도우 크기가 증가함에 따라 커짐을 알 수 있다. 그 이유는 넓은 윈도우를 사용할수록 지역적 변화량 보다 전역적인 변화량이 변동성에 주된 영향을 미치고, 전역적으로 볼 때 inlet 온

도의 변화는 outlet 온도 변화와 관련성이 높기 때문에 추측된다.

4.3 분석 결과의 적용 예시

공조기 온도 시계열 데이터의 분석 결과는 시계열 예측 모형 구축에 적용할 수 있다. 서론에서 언급하였듯이 시계열 데이터 예측을 위해 널리 사용되는 모형으로 ARMA와 GARCH 등을 들 수 있다. ARMA 모형은 자기회귀 항 (autoregressive term)과 이동 평균 항(moving average term)을 결합하여 시계열 데이터를 예측하는 모형으로,

ARMA(p, q)로 표현하고 2개의 매개변수 p 와 q 를 가지고 있다. ARMA(p, q)모형에서 시간 t 에 대한 예측치 $x(t)$ 는 다음과 같이 주어진다.

$$x(t) = c + \epsilon_t + \sum_{i=1}^p \phi_i x(t-i) + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad (9)$$

여기서 c 는 상수이고, ϵ_t 는 오차로 서로 독립이며 $\epsilon_t \sim N(0, 1)$ 따른다. 또한 ϕ_i 와 θ_i 는 일반적으로 주어진 데이터를 사용하여 최소자승법으로 추정한다. 매개변수 p 와 q 각각은 $x(t)$ 를 예측하기 위하여 사용하는 과거 시간 $t-i$ 에서 예측치 $x(t-i)$ 와 오차 ϵ_{t-i} 의 개수에 해당한다.

데이터 사이에 장거리 상관관계가 존재하지 않는 경우에는 시간 t 에서 예측치는 과거 시간 $t-i$ 에서 데이터와 상관관계가 적다. 따라서 일반적으로 부분 자기 상관 함수(partial autocorrelation function)와 자기 상관 함수 (autocorrelation function) 값이 급격하게 감소하다가 완만해지는 시점인 소위 “팔꿈치(elbow)” 지점을 p 와 q 로 선정한다. 그러나 본 연구를 통해 규명하였듯이 공조기 온도 데이터 사이에는 장거리 상관관계가 존재함으로(부분) 자기 상관 함수는 팔꿈치 현상을 보이지 않으며 완만히 감소할 것으로 예상된다. 따라서 이러한 경우, (부분) 자기 상관 함수의 값이 최대값의 반(半) 정도가 되는 시점을 p 와 q 값으로 일반적으로 정할 수 있다.

GARCH 모형은 오차 ϵ_t 를 확률적 오차(stochastic error) z_t 와 시간에 의존하는 편차(time-dependent deviation) σ_t 의 곱인 $\epsilon_t = z_t \sigma_t$ 으로 간주하는 모형으로, ARMA와 마찬가지로 매개변수 p 와 q 를 가지고 있다. GARCH(p, q) 모형에서 시간 t 에 대한 예측치 $x(t)$ 는 다음과 같이 주어진다.

$$x(t) = a_0 + \sum_{i=1}^p a_i x(t-i) + \epsilon_t \quad (10)$$

여기서 a_0 는 상수이며, a_i 는 최소자승법을 사용하여 추정한다. 또한 $\epsilon_t = z_t \sigma_t$ 이며, $z_t \sim N(0, 1)$, 그리고

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2 \quad (11)$$

로 주어진다. 여기서 a_0 는 상수이며, α_0 와 β_i 는 역시 최소자승법을 사용하여 추정한다. 데이터 간에 장거리 상관관계가 존재하는 경우, GARCH(p, q)모형에서 매개변수 p 와 q 의 결정은 ARMA(p, q)모형의 경우와 유사하게 (부분) 자기 상관 함수의 값이 최대값의 반 정도가 되는 시점을 p 와 q 의 값으로 정한다.

5. 결론

본 연구에서는 자동차 생산에 필요한 에너지의 효율적 관리를 위하여 자동차 건조 공정을 위한 공조기의 온도 시계열 데이터에 내재한 특성을 다양한 통계학적 방법을 적용하여 분석하고 규명하였다. 공조기 온도 시계열 데이터의 절대 변화량들 사이에 시간 지연에 따른 선형 상관관계는 무시할 수 있을 정도로 작으나, 비선형 상관관계는 존재함을 표본 분산을 사용하여 밝혔다. 절대 변화량에 대한 비선형 상관관계를 보다 구체적으로 분석하기 위하여 주식 시장에서 주식의 섭동(fluctuation)을 양적으로 표현하기 위해 사용하는 변동성을 적용하여 다양한 각도에서 분석을 수행하였다. 특히 다양한 윈도우 크기에 대하여 구한 변동성 시계열 데이터에 대하여 DFA를 적용하였으며, 그 결과 변동성 시계열 데이터 사이에는 장거리 상관관계가 존재함을 규명하였다. 장거리 상관관계는 주식 시장, 인터넷 트래픽 등에서 나타나는 현상으로 변동성의 간헐성이 오랜기간 지속적으로 발생함을 의미한다. 또한 inlet과 outlet의 온도 변화에 대한 상호상관관계 조사를 통하여 윈도우 크기에 따른 상관관계의 변화를 분석하였다.

공조기 온도 데이터에 대한 분석 결과는 열풍 건조 공정에 소요되는 에너지 수요 예측을 통한 효율적 에너지 관리와 절감을 위한 선행 자료로 그 가치가 있다. 이러한 측면에서 볼 때 본 연구의 결과는 건조 공정에 필요한 에너지의 수요 예측을 위한 모형 구축에 활용될 수 있다. 또한 본 연구에서 사용한 방법 외에 다양한 방법(예를 들어 다차원 프랙탈 분석)들의 적용은 향후 공조기 온도 변화를 이해하고 에너지 수요 예측 모형을 구축하는데 적용될 수 있을 것으로 예상된다.

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] Beran, J., *Statistics for Long-Memory Processes*, London, England : Chapman and Hall/CRC, 1994.
- [2] Black, F. and Scholes, M., The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 1973, Vol. 81, p 637-654.
- [3] Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 1986, Vol. 31, p 307-327.
- [4] Choi, J.R., Hur, N.K., Kim, D.C., and Kim, H.S., A Numerical Study on the Improvement of the Performance of a Vehicle Paint Drying Process. *Korean J Air-conditioning and Rdf Eng*, 2012, Vol. 24, p 868-875.
- [5] Cox, J., Ross, S., and Rubinstein, M., Option pricing : A simplified approach. *Journal of Financial Economy*, 1979, Vol. 7, p 229-263.
- [6] Engle, R., Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 1982, Vol. 50, p 987-1007.
- [7] Faranda, D. and Dubrull, B., Statistical early-warning indicators based on Auto-Regressive Moving-Average Processes, 2014, arXiv : 1402.2885v1.
- [8] Kim, I.M., Kim, C.S., and Park, S.K., Forecasting the Energy Demand Responses to Relative Price Changes. *Economic Studies*, 2012, Vol. 59, p 199-228.
- [9] Kong, D., Kwak, Y., and Lee, B., Huh, J.,. A Methodology of Databased Energy Demand Prediction Using Artificial Neural Networks for a Urban Community. *Society of Korean Solar Energy*, 2009, Vol. 29, p 184-189.
- [10] Lee, C., Detection of a long-range correlation with an adaptive detrending method. *Phys Rev E*, 2012, Vol. 86, 011135-5.
- [11] Lee, H., Park, K., and Shin, H., Electricity demand forecasting based on machine learning algorithms. Society of Korean management Science/Society of Korean Industrial Engineering Spring Joint Conference Proceedings, 2011, p 521-546.
- [12] Liu, Y., Gopikrishnan, P., Cizeau, P., Meyer, M., Peng, C.K., and Stanley, H.E., Statistical properties of the volatility of price fluctuations. *Physical Review E*, 1999, Vol. 60, p 1390-1400.
- [13] Mills, T.C., *Time Series Techniques for Economists*. New York : Cambridge University Press, 1990.
- [14] Pagan, A., The econometrics of financial markets. *J Empirical Finance*, 1996, Vol. 3, p 15-102.
- [15] Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., and Goldberger, A.L., Mosaic organization of DNA nucleotides. *Physical Review E*, 1994, Vol. 49, p 1685-1689.
- [16] Peng, C.K., Havlin, S., Stanley, H.E., Goldberger, A., Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, 1995, Vol. 5, p 82-87.
- [17] Song, G.S., *Development of Air-Conditioning System for Energy Saving type vehicle drying*. Report of DUKSAN Co. Ltd., 2013.
- [18] Suganthi, L. and Samuel, A.A., Energy models for demand forecasting-A review. *Renewable and Sustainable Energy Reviews*, 2012, Vol. 16, p 1223-1240.
- [19] 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, p 112-126.