Local Temperature Control in Data Center Cooling: Part II, Statistical Analysis

Keke Chen, Cullen E. Bash, David M. Auslander, Chandrakant D. Patel
Enterprise Software and Systems Laboratory
HP Laboratories Palo Alto
HPL-2006-43
March 6, 2006*

data center cooling, smart cooling, local temperature control, vent tile, time series analysis, correlation matrix

The data center plays a more and more important role with the increasing demands of internet and commodity computing. As a result of the evolution of the microprocessor and increasing demands of customers, the power density in data centers becomes much greater. Mechanical engineers face challenges in cooling system design. Recent research focuses on local control algorithms which can help to optimize local temperature distribution, and thus increase the energy efficiency of the system. This paper analyzes the relationship between rack inlet temperature and vent group configuration from the view of statistics. It also examines whether rack power can affect inlet temperature. Combined with part I in this series, the conclusion can be used as the basis of local control algorithms.

* Internal Accession Date Only
Approved for External Publication

Local Temperature Control in Data Center Cooling:

Part II, Statistical Analysis

[1]Keke Chen, [2]Cullen E. Bash, [3]David M. Auslander, [2]Chandrakant D. Patel
[1]Center for the Built Environment
University of California, Berkeley, CA
[2]Hewlett-Packard Laboratories, Palo Alto, CA
[3]Dept. of Mechanical Engineering
University of California, Berkeley, CA
Phone 510.642.2720, keke@berkeley.edu
Phone 650.236.2748, Fax 650.857.7029, cullen.bash@hp.com
Phone 510.642.4930, Fax 510.643.5599, dma@me.berkeley.edu
Phone 650.857.7140, Fax 650.857.7029, chandrakant.patel@hp.com

ABSTRACT

The data center plays a more and more important role with the increasing demands of internet and commodity computing. As a result of the evolution of the microprocessor and increasing demands of customers, the power density in data centers becomes much greater. Mechanical engineers face challenges in cooling system design. Recent research focuses on local control algorithms which can help to optimize local temperature distribution, and thus increase the energy efficiency of the system. This paper analyzes the relationship between rack inlet temperature and vent group configuration from the view of statistics. It also examines whether rack power can affect inlet temperature. Combined with part I in this series, the conclusion can be used as the basis of local control algorithms.

Key Words: data center cooling, smart cooling, local temperature control, vent tile, time series analysis, correlation matrix

INTRODUCTION

A data center is a machine room containing computing, networking and storage hardware that provides useful services. It is playing an increasingly important role in life with the development of computer technology. Modern computers have become smaller and faster. In the meanwhile, data centers are getting bigger, and they can accommodate more computers to satisfy the expanding demand for real-time computation and data exchange. Consequently, the power density of contemporary data centers has increased, and the cost of operation has escalated accordingly. Consequently, energy efficiency for data center cooling systems is of great interest.

Boucher et al. proposed a global control algorithm to optimize two global variables – supply air temperature setpoint and fan speed. However, they are limited in their scope of influence, i.e. these two variables will affect the data center in a large scale, which makes it difficult to deal with local conditions distinctively. It is possible to impact local conditions by utilizing actuated tiles that can be commanded to vary their resistance to fluid flow as discussed in Part I

of this series. To understand the properties of “smart tile”, Boucher et al. also analyzed the impact of vent tiles on the distribution of local cooling resources. They indicated that vent tile opening had a direct effect on local rack inlet temperature in close proximity to the vent and an inverse relationship on racks further away.

Previous work analyzed the relationship between vent tile configuration, rack inlet temperature and airflow from vent tiles, and summarized the correlation matrix between vent group configuration and inlet temperature for implementing local temperature control algorithms. It also mentioned that when calculating the correlation matrix, using the average temperature of steady states was more reasonable than using a single observation. As a supplement, this paper describes the details of analyzing the relationship utilizing statistical tools. Combined with the previous paper, the conclusion will serve as the basis of local temperature control algorithms.

EXPERIMENTATION

Figure 1 shows the plan view of the Data Center at Hewlett-Packard Laboratories, Palo Alto. The Grizzly area, Central area and Research area in the plot represent different sections of the data center. The Research area can be separated from the other two areas by a wall between the Research area and Grizzly area, curtains between the Research area and Central area and corresponding dampers under the plenum. There is nothing to separate the Grizzly area and Central area physically.

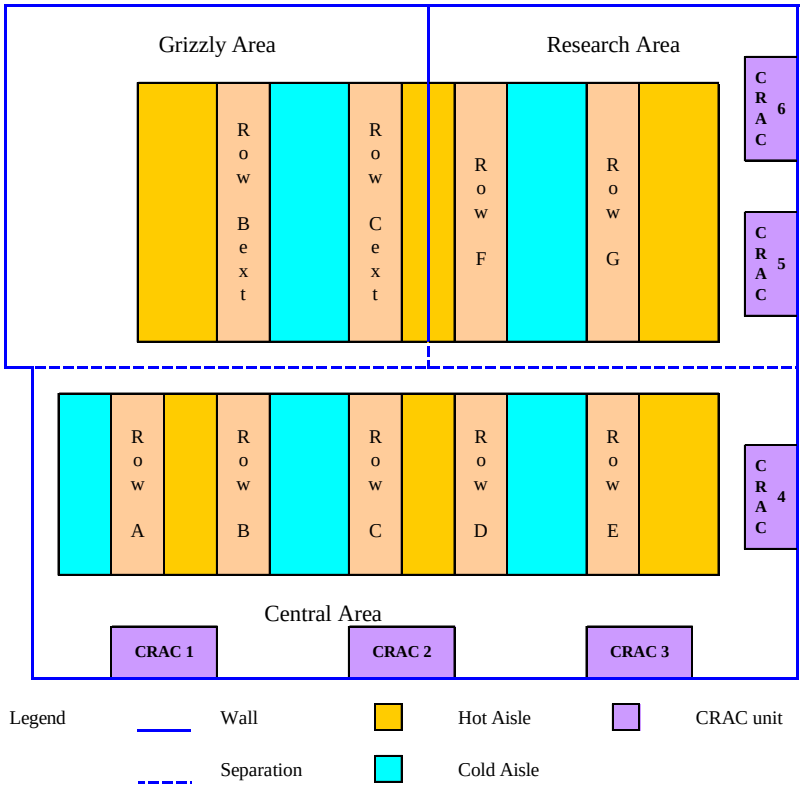


Figure 1: Plan View of the Data Center in HP Lab

Figure 2 illustrates the two rows of racks in the Grizzly area. Each row consists of several computer racks, with at least one temperature sensor installed in front of its inlet. We conducted vent-group experiments in this area while the Research area was completely separated from other areas by the wall and curtain above the floor and by the dampers under the floor. Rack power and inlet temperature were sampled every minute during the experiments according to the following procedure

- Define the base case as occurring when all the vent tiles are completely open. Close each group of vent tiles in the same row one by one, following the base case. Sample rack inlet temperatures and rack powers every minute during the period. Airflow from each tile is not measured. Keep the Research area isolated from the other areas in the HP Data Center for this series of experiments.

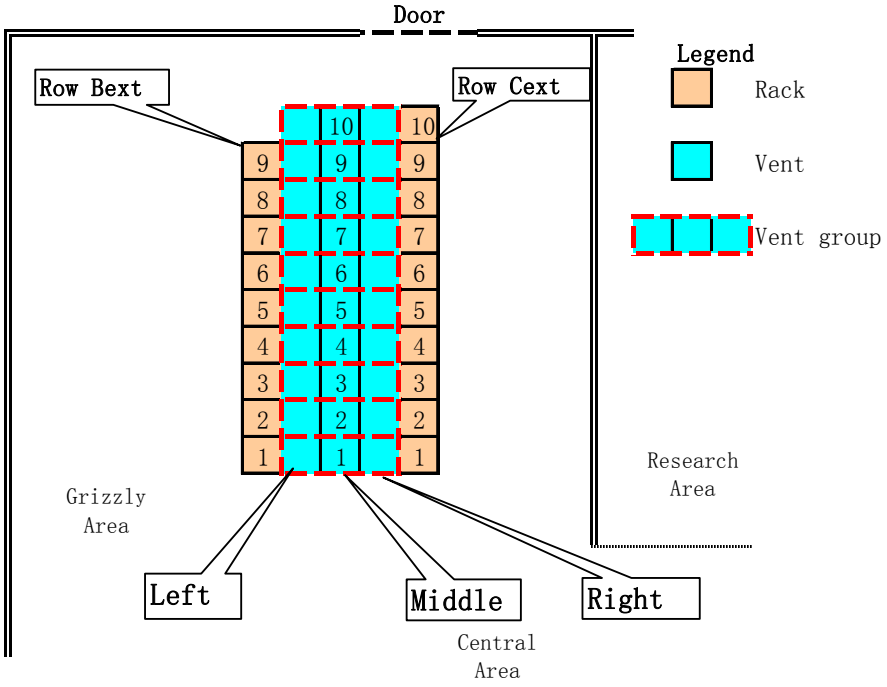
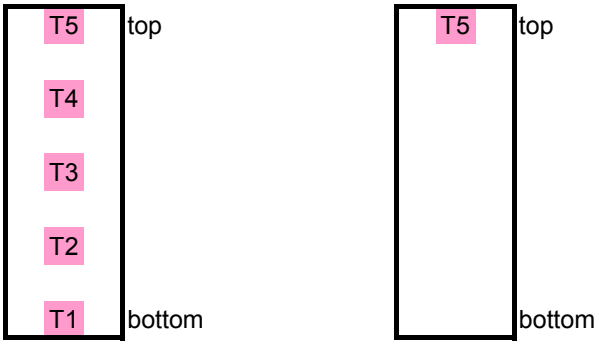


Figure 2: The Layout of Grizzly Area in HP Data Center



Rack Bext1~Bext9, Cext1~Cext6 **Rack Cext7~Cext10**

Figure 3: Location of Temperature Sensor at Each Rack (facing the racks)

The objectives of this paper are:

1. Analyze the relationship between vent group configuration and inlet temperature. Investigate whether rack power will affect inlet temperature as well.
2. Understand theproperties of the system, i.e., response time, lags.
3. Identify ARMAX (autoregressive moving-average with extraneous input) model: model the impact of input on output, access the residuals with the appropriate ARMA (autoregressive moving-average) model.
4. Give recommendations/suggestions for follow-up experiments.

RESULTS/DISCUSSION

1. Preliminary Analysis

1) Check Inlet Temperature and the Influence of Vents

Figure 4 is an example of the relationship between vent group and rack inlet temperature. We can represent the temperature data of sensor "B1T2" (The temperature sensor T2 mounted on Bext1 of Grizzly area. See Figure 2 and 3) and the status of vent group 1 in a time plot, as shown in Figure 4. For each vent group, we use 0 as the status of "completely open", 1 as the status of "completely closed". The inlet temperature of each location is determined by global CRAC settings and local vent group configurations. Its trend doesn't necessarily follow any single factor.

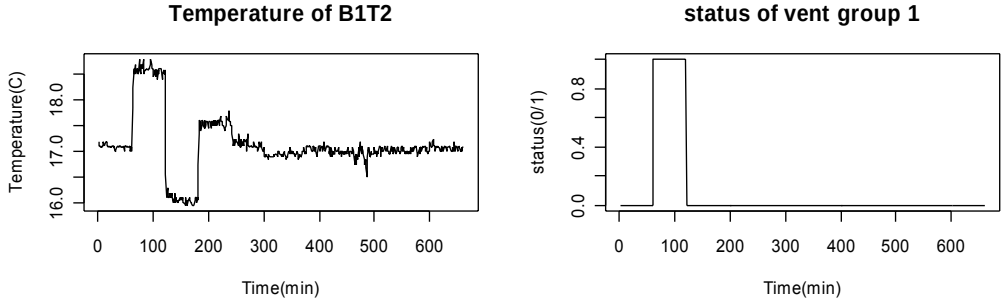


Figure 4: Time Plot for Temperature of B1T2 and Status of Vent Group 1

An autocorrelation function can be used to show the non-stationary properties of a time series. Suppose a time series X(t) runs throughout time, and is observed for t=1,…,n. Then the autocovariance at time lag τ is

$$\gamma_{XX}(\tau) = \mathrm{cov}(X(t+\tau), X(t)) = E[\{X(t+\tau)-\mu\}\{X(t)-\mu\}] \qquad \text{(a)}$$

Correspondingly, autocorrelation at lag τ is

$$\rho_{XX}(\tau) = corr(X(t+\tau), X(t)) = \frac{\gamma_{XX}(\tau)}{\gamma_{XX}(0)} = \frac{\mathrm{cov}(X(t+\tau), X(t))}{\sqrt{\mathrm{var}(X(t+\tau))\,\mathrm{var}(X(t))}} \qquad \text{(b)}$$

This can be extended for multiple time series. Suppose two series $X_i(t)$, $X_j(t)$ are observed for t=1,…,n. Then cross covariance at lag τ is given by

$$\gamma_{ij}(\tau) = \mathrm{cov}(X_i(t+\tau), X_j(t)) = E[\{X_i(t+\tau)-\mu_i\}\{X_j(t)-\mu_j\}] \qquad \text{(c)}$$

And cross correlation at lag τ is given by

$$\rho_{ij}(\tau) = corr(X_i(t+\tau), X_j(t)) = \frac{\gamma_{ij}(\tau)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}} \qquad \text{(d)}$$

The autocorrelation function is the variable indicating the correlation of a time series with a certain lag, and the cross correlation function is the variable indicating the correlation of two time series with a certain lag.

Figure 5 shows the autocorrelation and cross correlation plot of inlet temperature and vent group status, which can be calculated by equation (b) and (d). Temperature expresses some long-term trend behavior. The controlled variable, vent group status, behaves in a similar way. In both the autocorrelation plot and cross correlation plot, the pair of blue broken lines represents a 95% confidence interval of zero correlation. In other words, if the correlation is significantly different than 0, its line will fall outside the limit bounded by those broken lines. Figure 5 indicates that vent group 1 is correlated to the temperature of B1T2, and vent group 6 isn't. This conclusion is consistent with the physics: rack inlet temperature will be affected by nearby vent groups.

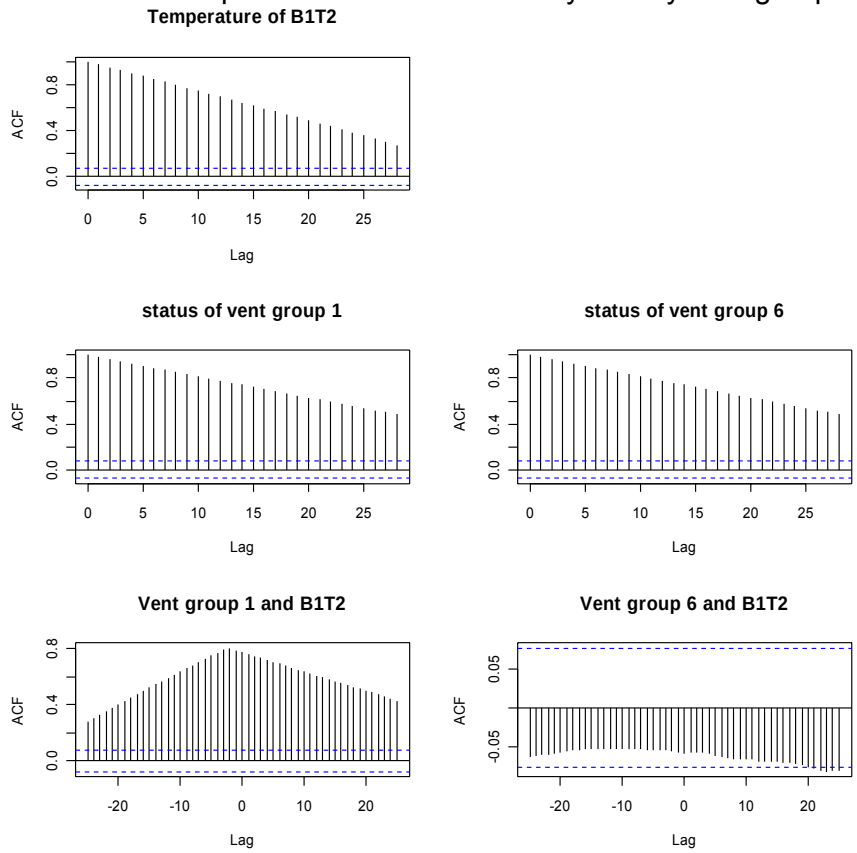


Figure 5: ACF and CCF Plot for Temperature of B1T2 and Vent Group 1

Spectral analysis in frequency domain can also be useful when the time series incorporate some periodical trends [6].

Squared coherency is defined by

$$R_{ij}^2(\lambda) = \frac{f_{ij}^2(\lambda)}{f_{ii}(\lambda) f_{jj}(\lambda)} \qquad \text{(e)}$$

Where $f_{ij}(\lambda)$ is the cross spectrum of $X_i(t)$ and $X_j(t)$, $f_{ii}(\lambda)$ and $f_{jj}(\lambda)$ is the power spectrum of $X_i(t)$ and $X_j(t)$, respectively.

$$f_{ij}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{ij}(h) \qquad \text{(f)}$$

$$f_{ii}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{X_i}(h) \qquad \text{(g)}$$

$$f_{jj}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{X_j}(h) \qquad \text{(h)}$$

Notice that $0 \le \left| R_{ij}(\lambda) \right| \le 1$, and we refer $R_{ij}(\lambda)$ as the "correlation coefficient" in frequency domain.

Figure 6 is the spectral analysis of B1T2 with vent group 1 and 6, respectively (broken lines represent 95% confidence interval). The coherency of vent group 6 and B1T2 doesn't show any obvious correlation at a particular frequency (second plot in the second row), while that of vent group 1 and B1T2 (first plot in the second row) ranges from 0.2 to 0.7 but doesn't indicate correlation at any particular frequency. This means that vent group 1 and B1T2 are somewhat correlated, while vent group 6 and B1T2 are not. Phase doesn't have explicit meaning in this problem. From the analysis in frequency domain (the two plots at the top of figure 6), we also notice that inlet temperatures do not show seasonal behavior. The power spectrum of temperature concentrates in low frequency, and mainly results from the variation of vent group configuration.

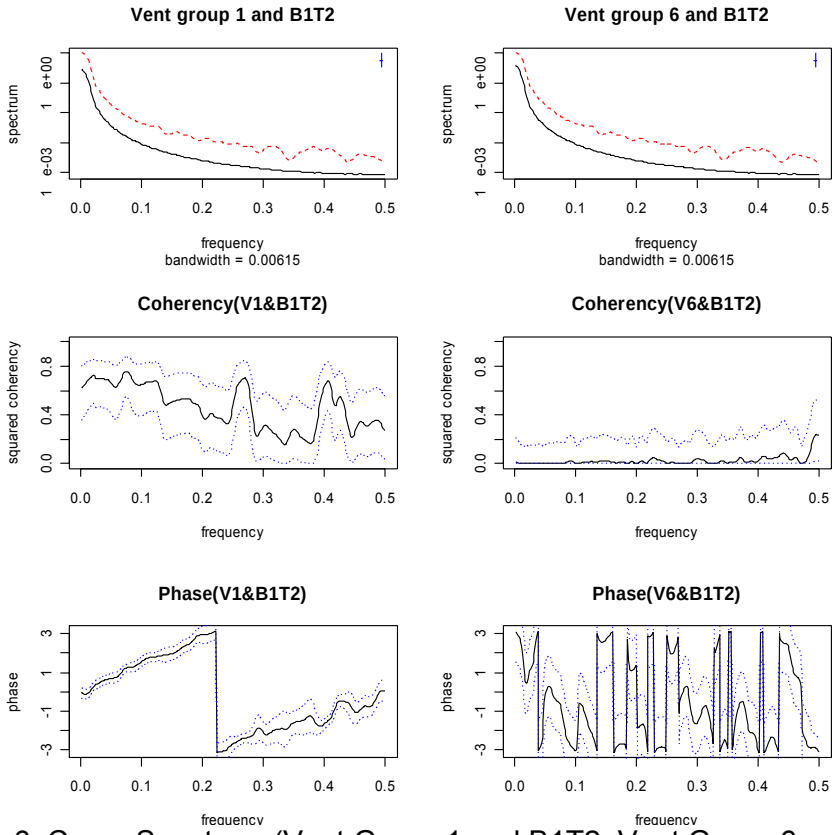


Figure 6: Cross-Spectrum (Vent Group 1 and B1T2; Vent Group 6 and B1T2)

2) Check the Relationship between Rack Power and Inlet Temperature

What interests us is the influence of vent tiles on rack inlet temperature. However, the power of each computer rack might affect inlet temperatures. Figure 7 includes the time plot for the power of rack Bext1, Bext9 and Cext7, and corresponding ACF. Among all of the rack-power series, some of them express embedded trend/seasonal behavior over the period, while others oscillate around a constant value. Their nonstationary behaviors serve as a potential factor which could influence rack inlet temperatures.

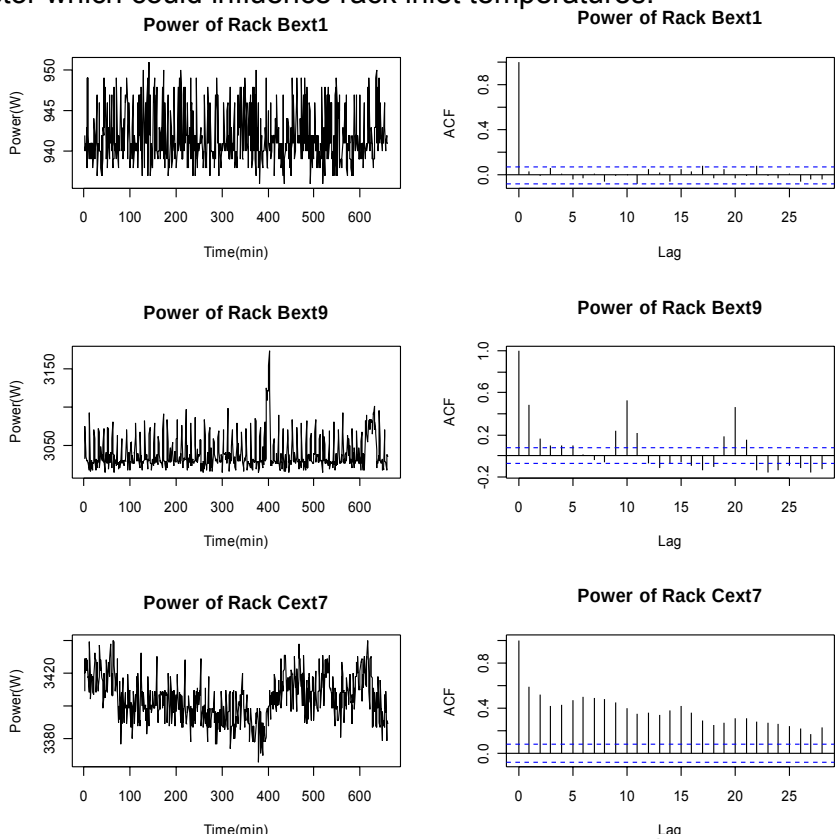


Figure 7: Rack Power (Time Plot; ACF)

One way to check whether the power has influence on temperature is to calculate the cross correlation coefficient between these two variables. If two time series $X_i(t)$, $X_j(t)$ (t=1,…,N) are uncorrelated, let $\rho_{ij}(k)$ be the cross correlation coefficient between $X_i(t)$ and $X_j(t)$ at lag k, then approximately we have

$$E(\rho_{ij}(k)) \approx -1/N \qquad \text{(i)}$$

$$Var(\rho_{ij}(k)) \approx 1/N \qquad \text{(j)}$$

and $\rho_{ij}(k)$ is asymptotically normally distributed under weak conditions [7]. We can plot an approximate 95% confidence interval $-1/N \pm 1.96/\sqrt{N}$, which are often further approximated to $\pm 1.96/\sqrt{N}$. Observe that $\rho_{ij}(k)$, which falls outside these limits, are significantly different from zero at the 5% level. So if there are too many outliers, we might consider the high possibility of a violated null hypothesis. Alternatively, $X_i(t)$ and $X_j(t)$ are correlated.

Some of the cross correlation plots are shown in figures 8 and 9. In this project, there are 79 temperature series and 19 power series. We calculated all of the cross correlation coefficients between these series by command “ccf” in a statistical package – R [8]. By default in R, the maximum number of lags is given as $10 \times \log_{10}(\frac{N}{m})$, where N is the number of observations and m is the number of series. Then in the current context, we have

$$\max.lag = 10 \times \log_{10}(\frac{660}{2}) \approx 25 \qquad \text{(k)}$$

So for each pair of power and temperature, we will get $25 \times 2 + 1 = 51$ coefficients. Recall that $\rho_{ij}(k)$ is asymptotically normally distributed under weak conditions as mentioned above. If power has no influence on temperature, we expect at most $51 \times 5\% \approx 2$ observations of each of these results to be outside the 95% confidence interval of $\rho_{ij}(k)$. So from a total $79 \times 19 = 1501$ pairs of power and temperature, we discover only 243 of them have at least 3 observations outside the 95% confidence interval, which indicates power and temperature might be “correlated” at a certain lag. Figure 8 and 9 show only part of these 243 cross correlation plots. The main title indicates the component rack and temperature sensor, while the subtitle describes the number of cross correlation coefficients which are outside the confidence interval.

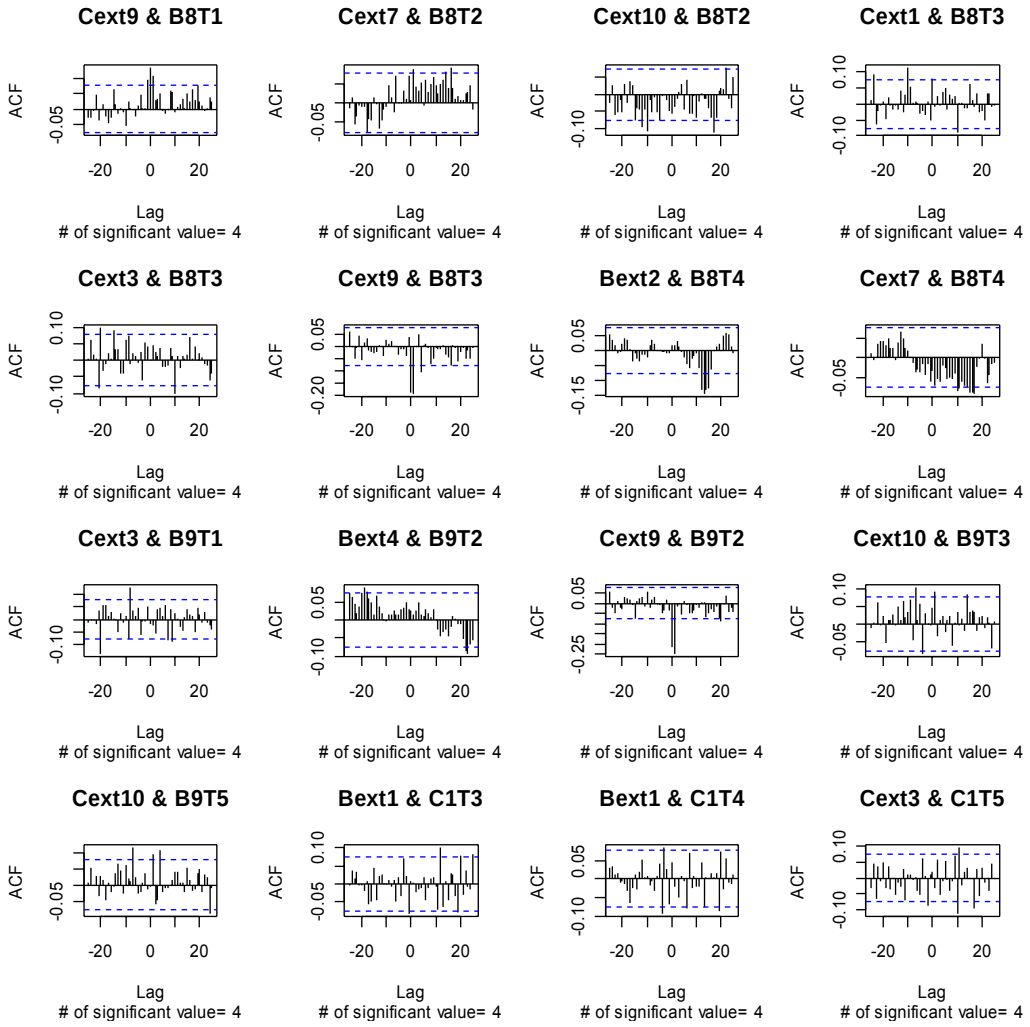


Figure 8: Cross Correlation between Power and Temperature (1)

According to the data, there is no evidence to suggest that temperature series are correlated to power series for this particular data set. Care must be taken when generalizaing these results.

(1) Most power-temperature pairs ($\frac{1501-243}{1501}=\frac{1258}{1501}=83.81\%$) indicate that there is no correlation between power and temperature.

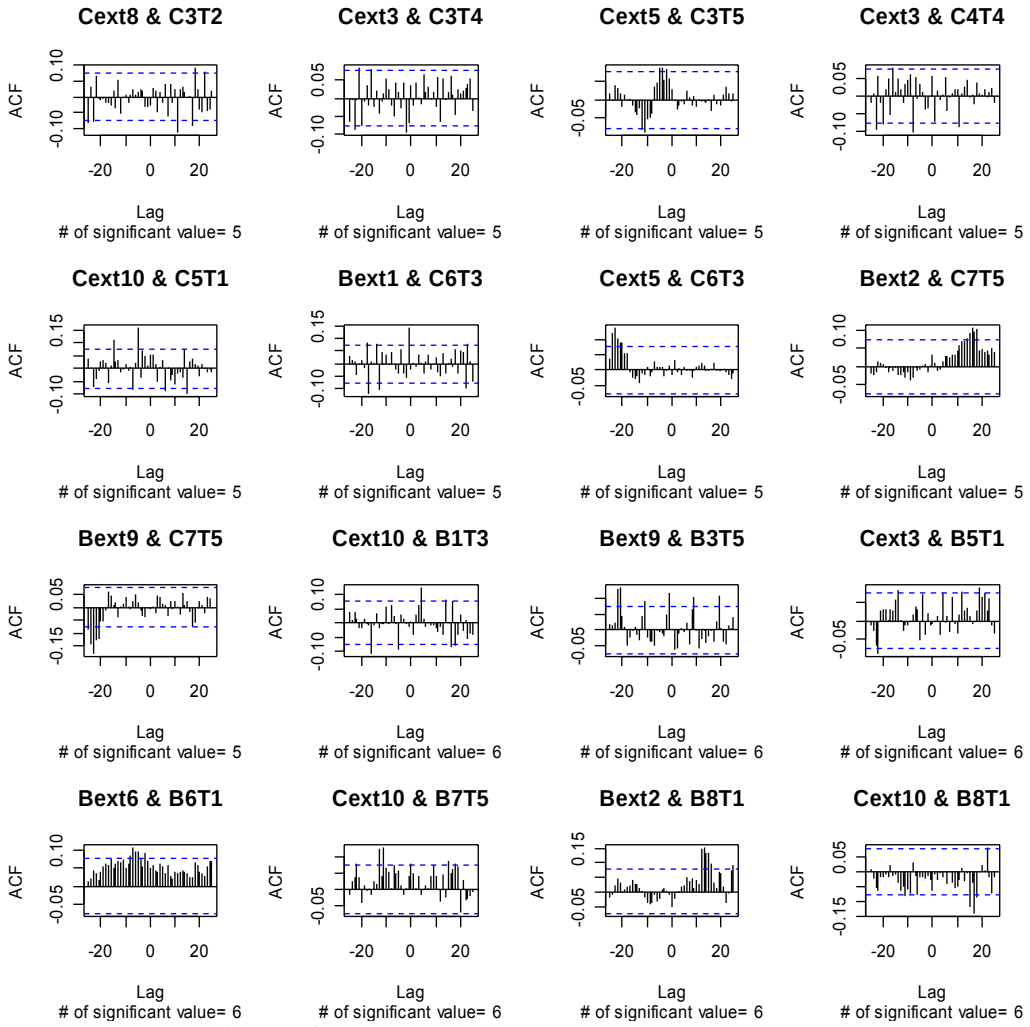


Figure 9: Cross Correlation between Power and Temperature (2)

(2) The correlation is very weak or nonexistent in these 243 exceptions, since most of the “outliers” are just marginally across the boundary.

(3) Inlet temperature is the response variable, while rack power is an independent variable which certainly doesn’t depend on inlet temperature (under normal conditions, power depends on the computing load of computers). Thus, if the cross correlation plot indicates some significant values around a certain non-negative lag, rack power may affect inlet temperature in physics. Some of the relations with negative lags became meaningless, i.e. “Cext5 & C6T3” in Figure 9.

(4) Some relations may be significant. However, if the rack and temperature sensor are far away from each other in the layout, a high correlation between them is suspect, i.e. “Bext2 & B8T4” in figure 8.

(5) In some cases, the “outliers” periodically appear in the plot, i.e. “Bext9 & B3T5” in Figure 9, which might result from the seasonal behavior embedded in the series. They are also marginally across the boundary.

As a check, we can calculate the cross-spectrum between power and temperature in frequency domain. Figure 10 includes coherency and phase spectrum from two cases: Bext2 and B8T4; Bext2 and B2T1. The coherency of the first case ranges from 0 to 0.5, which shows a weak correlation between power and temperature. And the second one indicates no clue about this correlation (it is not even part of the 243 cases mentioned above). The lack of coherency in figure 10 agrees with the result from the time domain analysis of figure 9.

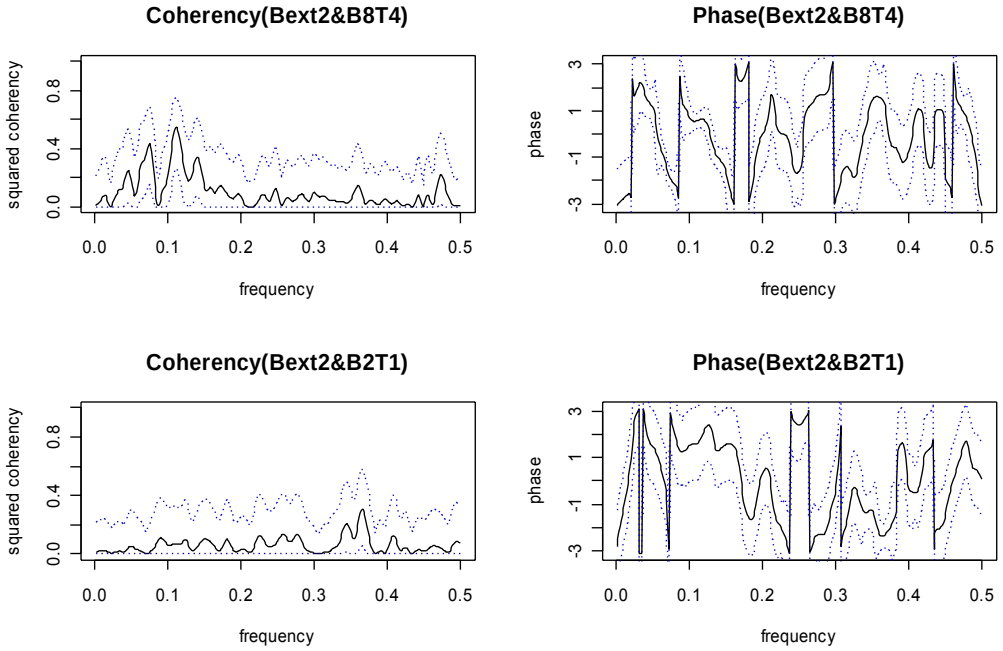


Figure 10: Cross-Spectrum (Bext2 and B8T4; Bext2 and B2T1)

Generally speaking, the influence of rack power on inlet temperature is very weak or negligible. Therefore, for this data set, we don't incorporate it in the final model.

2. System Identification

As mentioned above, there are 79 temperature sensors in this area. For simplicity, we only carried out the analysis on one of them. A conclusion can be drawn when similar analysis is performed on other series. Use the temperature sensor "B1T2" as the example. Based on preliminary analysis in both time and frequency domain, we consider the following model,

$$Y = X\beta + W \qquad (l)$$

Where Y is temperature, X is the status of the vent group, β is the coefficient of vent status, and W is the error which can be determined through appropriate ARMA model.

The above model can be written into

$$y(t) = \beta_0 + \sum_{j=1}^{n_a} \sum_{k=k_1}^{k_2} \beta_{jk} x_j(t-k) + w(t) \qquad (m)$$

Where n_a represents the number of the vent group (n_a=10); k_1, k_2 represents the minimum and maximum time lag of x_j. Notice that in a simple situation in which we do not consider the transient state, only $X_1,\ldots,X_{10}$ are in the model, then β_i (i=1,…,10) represents the influence of each group of vent tiles. For instance, β_1 is the influence of vent group 1 while holding other vent groups unchanged.

1) Access the properties of the system

Clearly, temperatures will respond to the change of vent group with a positive time lag. And the influence of a vent group only lasts a certain period. We now identify the parameter k_1 and k_2. Set k_1=0, k_2=9 and fit the model by ordinary least square (OLS) [9]. The adjusted R-square [10] for this model is 0.9874, which indicates a fairly good fit for the model. Next we select a criterion to reach an optimized model. Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are considered [11][12].

AIC: (Akaike 1974). Can be used to compare models. A smaller value is better.

$$AIC = -2\log L(\hat{\theta}) + 2K \qquad \text{(n)}$$

Where K is the number of parameters.

BIC: (Schwarz 1978). Can also be used to compare models. A smaller value is better.

$$BIC = -2\log L(\hat{\theta}) + \log(n)K \qquad \text{(o)}$$

Where K is the number of parameters and n is the number of observations.

Both criteria are tried. AIC tends to give out a model with too many parameters in the current context. We used BIC and the selected β_{jk} are expressed by a stem-and-leaf plot in terms of j (vent group label) in the stem and k (time lag label) in the leaf, as shown in Table 1.

j	1	2	3	4	5	6	7	8	9	10
k	3 4 5	3 5 7	2 3 4	2 3	1 2 3	2 3 8	2 3 5 7 8	3 8	3 8	3 8

Table 1: Stem-and-leaf of selected β_{jk}

Recall that in β_{ik}, j represents vent group, k represents time lag. Suppose the temperature of a specific location is only influenced by vent group 1 for 3 minutes, starting from time lag 1 minute. Then in the stem-and-leaf plot, the only stem would be j=1 and k=1, 2, 3. Thus, here the stem-and-leaf plot is used to indicate the vent group that has significant influence on inlet temperature statistically and the duration of transient state from each group. For example, in table 1, vent group 1 only has influence on “B1T2” at time lag minutes 3, 4, and 5; vent group 9 has influence on “B1T2” at time lag minute 3 and 8. How-

ever, this model is selected by a specific statistical criterion. Physically, it doesn't make sense that vent group 9 will have an effect on the sensor at two discontinuous points in time (minutes 3 and 8).. Thus it is possible that, although some vent groups contribute to the model statistically, they don't have significant influence on temperature physically. From Table 1, it seems that k_1 and k_2 depend on j. We learn that temperatures will respond to the change of vent group with approximately a 1 minute lag, and after about 8 minutes from the time vent group status changes, the influence on temperature will diminish to a negligible level statistically. Currently any setting of vent group configuration was kept for 1 hour during the experiment. The model supports a possible shorter period.

Further, we tried a model which ignored the transient status in the problem and can be written into $y(t) = \sum_{j=1}^{n_a} \beta_j x_j(t) + w(t)$. The adjusted R-square for this model is 0.886, which is also good enough within the current context. This example indicates that it may be acceptable to ignore all the transient states and only average the temperature within each period when the vent group configuration is maintained.

The conclusion that transient state lasts for about 8 minutes is based on the assumption that it won't last over 9 minutes. In this particular case, we can accept this fact in terms of the high R-square value of the model. However, if we want to get a more reliable conclusion, we need to analyze all the relationships of each vent group-temperature pair, since the temperatures of some locations may respond to the vent group configuration quickly, while others may not. If the fact that transient states won't last for more than 10 minutes is true for each location, we can simply regard 10 minutes as the period of transient state. Table 2 indicates high R-square for each location, generated by two methods: "R^2 (1)" represents the R-square value for the model of averaging all the temperature records and ignoring the transient state, and "R^2 (2)" represents the R-square value for the model of averaging the temperature records of "steady state" (remove the records of the first 10 minutes and last 5 minutes of each test). Simply averaging the temperature records of "steady state" can help to improve the models, which are accurate enough in the context. A parsimonious model for each vent-temperature pair is not necessary, since it won't give out significantly different results.

Location	R^2 (1)	R^2 (2)	Location	R^2 (1)	R^2 (2)	Location	R^2 (1)	R^2 (2)
1	0.891	0.959	28	0.902	0.964	55	0.894	0.914
2	0.886	0.989	29	0.747	0.809	56	0.845	0.912
3	0.876	0.955	30	0.832	0.894	57	0.944	0.953
4	0.777	0.797	31	0.97	0.996	58	0.9	0.922
5	0.804	0.821	32	0.96	0.97	59	0.899	0.927
6	0.852	0.947	33	0.915	0.955	60	0.871	0.882
7	0.919	0.955	34	0.868	0.894	61	0.921	0.967
8	0.849	0.856	35	0.978	0.985	62	0.913	0.955

9	0.775	0.803	36	0.972	0.982	63	0.775	0.87
10	0.938	0.966	37	0.969	0.983	64	0.905	0.919
11	0.927	0.994	38	0.977	0.987	65	0.914	0.913
12	0.954	0.977	39	0.988	0.996	66	0.935	0.969
13	0.815	0.842	40	0.988	0.994	67	0.957	0.969
14	0.938	0.956	41	0.876	0.927	68	0.737	0.823
15	0.905	0.932	42	0.983	0.995	69	0.768	0.775
16	0.965	0.998	43	0.942	0.952	70	0.842	0.841
17	0.898	0.934	44	0.943	0.947	71	0.968	0.981
18	0.758	0.803	45	0.941	0.951	72	0.877	0.936
19	0.7	0.798	46	0.844	0.901	73	0.843	0.856
20	0.799	0.875	47	0.825	0.879	74	0.942	0.968
21	0.964	0.994	48	0.84	0.875	75	0.937	0.948
22	0.95	0.98	49	0.833	0.855	76	0.936	0.98
23	0.867	0.932	50	0.843	0.858	77	0.905	0.946
24	0.783	0.899	51	0.703	0.744	78	0.945	0.979
25	0.764	0.833	52	0.787	0.811	79	0.965	0.967
26	0.976	0.997	53	0.819	0.846			
27	0.926	0.964	54	0.867	0.892			

Table 2: Adjusted R-square Value for Each Model by Two Methods

2) Error Determination

We continue to analyze the parsimonious model of “B1T2”, which is selected by BIC above. In detail,

$$Y(t)=\sum_i\sum_j \beta_{ij}X_i(t-j)+W(t)\,,\ \phi(B)W(t)=\theta(B)Z(t)\,,\ Z(t)\sim WN(0,\sigma^2)$$

Where j is the lag in the final model selected by BIC criterion. The error W(t) is estimated by ARMA model. $WN(0,\sigma^2)$ is Gaussian white noise.

Figure 11 plot the ACF and PACF for the residuals from this model.

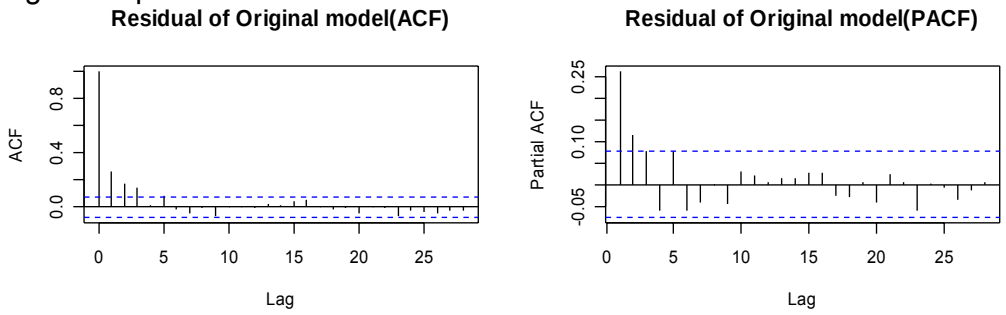


Figure 11: ACF and PACF for Residuals from the Model Selected by BIC

Notice that the AR(p) process cuts off at lag p in PACF, while the MA(q) process cuts off at lag q in ACF [13]. Figure 11 suggests that the maximum order for AR and MA are approximately 3 and 3, respectively. With the command “arima” in R, we can evaluate all the models whose p and q are not larger

than 4 (here select a larger number for more flexibility). Table 3 shows the AIC value for each model at corresponding p, q.

		q				
		0	1	2	3	4
p	0	-1696.98	-1734.07	-1742.14	-1753.25	-1751.73
	1	-1743.53	-1752.53	-1750.68	-1753.22	-1751.33
	2	-1750.03	-1750.65	-1751.39	-1752.45	-1751.46
	3	-1752.77	-1756.36	-1757.12	-1756.46	-1755.11
	4	-1753.16	-1754.93	-1755.94	-1754.85	-1753.36

Table 3: Access ARMA Model by AIC

We find the minimum AIC at p=3, q=2. The estimates of coefficient and corresponding standard errors are listed in Table 4.

	Coef.	Std. Err		Coef.	Std. Err		Coef.	Std. Err
ar1	-0.790	0.175	X3(t-2)	0.647	0.054	X7(t-3)	-0.614	0.075
ar2	-0.150	0.189	X3(t-3)	-0.467	0.067	X7(t-5)	0.170	0.041
ar3	0.297	0.052	X3(t-4)	0.282	0.045	X7(t-7)	-0.165	0.048
ma1	1.047	0.182	X4(t-2)	0.601	0.067	X7(t-8)	0.220	0.062
ma2	0.514	0.220	X4(t-3)	-0.566	0.068	X8(t-3)	-0.486	0.063
intercept	17.081	0.013	X5(t-1)	-0.097	0.042	X8(t-8)	0.406	0.063
X1(t-3)	0.808	0.047	X5(t-2)	0.493	0.078	X9(t-3)	-0.513	0.073
X1(t-4)	0.447	0.052	X5(t-3)	-0.548	0.073	X9(t-8)	0.459	0.072
X1(t-5)	0.236	0.047	X6(t-2)	0.364	0.069	X10(t-3)	-0.431	0.081
X2(t-3)	-1.134	0.044	X6(t-3)	-0.550	0.077	X10(t-8)	0.403	0.081
X2(t-5)	0.226	0.047	X6(t-8)	0.141	0.037			
X2(t-7)	-0.116	0.034	X7(t-2)	0.277	0.054			

Table 4: Parameter estimation of final model

Check the final model from different charts in Figure 12. Fitted values are pretty close to the observation. Note that the noise is very close to Gaussian white noise, when observing the ACF, CCF and power spectrum plot of the residuals.

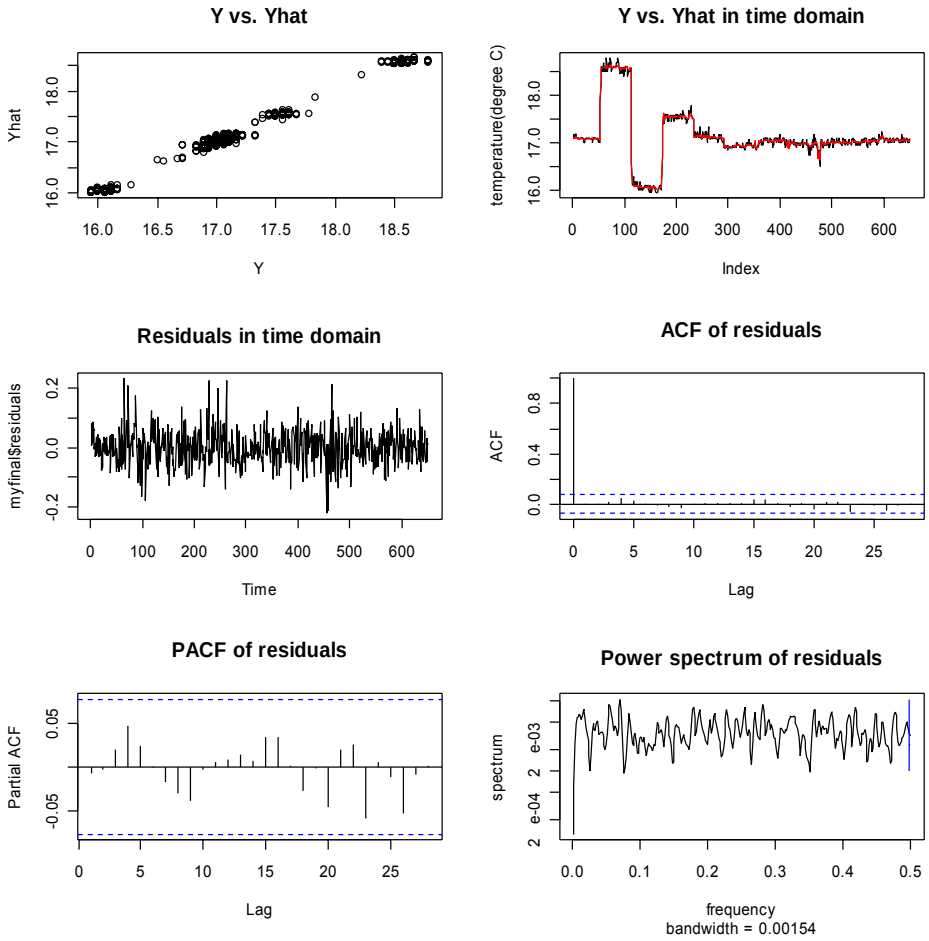


Figure 12: Some Plots of Final Model

CONCLUSION

This paper investigates the relationship between vent group configuration and rack inlet temperature in the HP Data Center with time series analysis. We used the techniques in both time domain and frequency domain to evaluate how the temperature would be affected by a controlled variable (vent group configuration) and a potential factor (rack power).

- Most inlet temperature variation results from different vent group configurations, while keeping the global settings constant.

- Rack power shows some nonstationary behaviors during the experiment. However, it doesn't affect rack inlet temperature significantly.

- For the specific example, we determine the final ARMAX(p,q) model with p=3, q=2 and X is selected by BIC. In other words, after removing the mean and influence from the vent groups, the temperature can be appropriately modeled by ARMA(3,2) process.

- Since fitting the ARMA model is actually not of direct interest in this project, we also discover it is appropriate to use a simpler model to represent the final sophisticated model, by averaging the temperature of

"steady state". The transient effects do not last longer than 10 minutes, which can be used for further experiments, i.e. shorter periods for each test are acceptable.

Potentially, there are other applications of time series analysis in data center cooling, especially when we are particularly interested in two time series. For example, we might want to know the relationship between the return air temperature of CRAC units and total power of a row of racks. Then time series analysis can be carried out based on the measurements. Certainly the relationship may change from time to time. We have to update the results with the latest measurements.

ACKNOWLEDGMENTS

Many thanks to Jack Cattolico and Wayne Mack of Hewlett-Packard Laboratories for assistance with experiments and data collection. Also thanks to Ratnesh Sharma who helped a lot with temperature network and data logging. Finally, thanks to Kathe Gust for reviewing the report.

REFERENCES

Chen, K., et. al., "Local Temperature Control in Data Center Cooling: Part I, Correlation Matrix", HP Labs Technical Report #xxx

Bash, C. E., Patel, C. D., Sharma, R. K., 2003, "Efficient Thermal Management of Data Centers – Immediate and Long-Term Research Needs", HVAC&R Research, Vol.9, No.2, Apr 2003

Bloomfield, P., "Fourier Analysis of Time Series: An Introduction", John Wiley & Sons, 1976

Boucher, T. D., Auslander, D. M., Bash, C. E., Federspiel, C. C., Patel, C. D., 2004, "Viability of Dynamic Cooling Control in a Data Center Environment"

Chatfield, C., "the Analysis of Time Series: An Introduction", Fourth Edition, Chapman and Hall, 1989

Chen, K., Auslander, D. M., Bash, C. E., Patel, C. D., 2005, "Local Temperature Control in Data Center Cooling: Part I, Correlation Matrix"

Friedrich, R., Patel, C. D., 2002, "Towards Planetary Scale Computing – Technical Challenges for Next Generation Internet Computing", THERMES 2002, Santa Fe, New Mexico

Patel, C. D., Bash, C. E., Sharma, R. K., Beitelmal, A and Friedrich, R. J., "Smart Cooling of Data Centers", Proceedings of IPACK'03, The Pacific Rim/ASME International Electronic Packaging Technical Conference and Exhibition, Kauai, Hawaii, USA, IPACK-35059, July 2003

Peter J. Brockwell, Richard A. Davis, "Introduction to Time Series and Forecasting", Second Edition, Springer, 2002

Stone, C. J., “A Course in Probability and Statistics”, Duxbury Press, 1996

W.N. Venables, B.D. Ripley, “Modern Applied Statistics with S-PLUS”, Third Edition, Springer, 1999

http://cran.r-project.org/

NOMENCLATURE

Item	Type	Example
Observations, time series	Lowercase Roman	$x_1, x_2; z_t, z(t)$
Random variables, stochastic processes	Uppercase Roman	$X_1, X_2; Z_t, Z(t)$
Estimates	Lowercase Roman, lowercase Greek with caret	$\overline{x}, s^2; \hat{\theta}, \hat{\alpha}_1, \hat{\alpha}_2$
parameter	Lowercase Greek	$\theta, \alpha_1, \alpha_2$

$X(t)$	Continuous time series
X_t	Discrete time series ($X(t), t = 0, \pm 1, \pm 2, \cdots$)
$c_{XX}(k)$	Sample autocovariance of X at lag k
$\gamma_{XX}(k)$	Theoretical autocovariance of X at lag k
$c_{XY}(k)$	Sample cross-covariance of X and Y at lag k
$\gamma_{XY}(k)$	Theoretical cross-covariance of X and Y at lag k
$r_{XX}(k)$	Sample autocorrelation of X at lag k
$\rho_{XX}(k)$	Theoretical autocorrelation of X at lag k
$r_{XY}(k)$	Sample cross correlation of X and Y at lag k
$\rho_{XY}(k)$	Theoretical cross correlation of X and Y at lag k
$f_{XX}(\lambda)$	Power spectrum of X at frequency λ
$f_{XY}(\lambda)$	Cross spectrum of X and Y at frequency λ
$R_{XY}(\lambda)$	Coherency of X and Y at frequency λ
μ	Mean of time series

ACF	autocorrelation function
ACVF	autocovariance function
AR	autoregressive
ARMA	autoregressive moving-average
ARMAX	autoregressive moving-average with extraneous input
CCF	cross correlation function

CORR	correlation
COV	covariance
CCVF	cross covariance function
E	expectation
MA	moving average
OLS	ordinary least square
PACF	partial autocorrelation function
VAR	variance

Major Concepts

- Time Series: a random or nondeterministic function x of an independent variable t. Since different sections of a time series resemble each other only in their average properties, it is necessary to describe these series by probability laws or models. Thus, possible values of the time series at a given time t are assumed to be described by a random variable X(t) and its associated probability distribution. The observed value x(t) of the time series at time t is then regarded as one of the infinity of values which he random variable X(t) might haven taken at time t. Time series which occur in practice are either discrete or continuous.

- Objective of Time Series Analysis: set up a hypothetical probability model, estimate parameters, check for goodness of fit to data, and use the fitted model to enhance our understanding of the mechanism generating the series. Once a satisfactory model has been developed, it may be used in a variety of ways depending on the particular field of application (filtering, smoothing, forecasting).

- Stochastic Process: the ordered set of random variables {X(t)} and its associated probability distributions.

Let {X(t)} be a time series with $E(X^2(t)) < \infty$.

- Mean Function: $\mu_X(t) = E(X(t))$
- Covariance Function:

$$\gamma_X(r,s) = \mathrm{cov}(X(r), X(s)) = E[(X(r) - \mu_X(r)(X(s) - \mu_X(s)]$$

- Stationarity

Strictly stationary: the joint distribution of {X(t+τ)} does not depend on τ for t, τ = 0, ±1, ±2, …

Weakly stationary: (i) $\mu_X(t)$ is independent of t; (ii) $\gamma_X(t+h,t)$ is independent of t for each h.

Since the statistical properties of stationary series do not change with time, these properties can be conveniently summarized by computing certain functions from the data. Let {X(t)} be a stationary time series

- Autocovariance Function (ACVF)

$$\gamma_X(h) = \mathrm{cov}(X(t+h), X(t))$$

- Autocorrelation Function (ACF)

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \mathrm{cor}(X(t+h), X(t))$$

Let $x_1,\ldots,x_n$ be observations of a time series.

- Sample Mean

$$\bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

- Sample Autocovariance Function

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x}), -n < h < n$$

- Sample Autocorrelation Function

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}, -n < h < n$$

Suppose two series $X_i(t)$, $X_j(t)$ are observed for t=1,…,n.

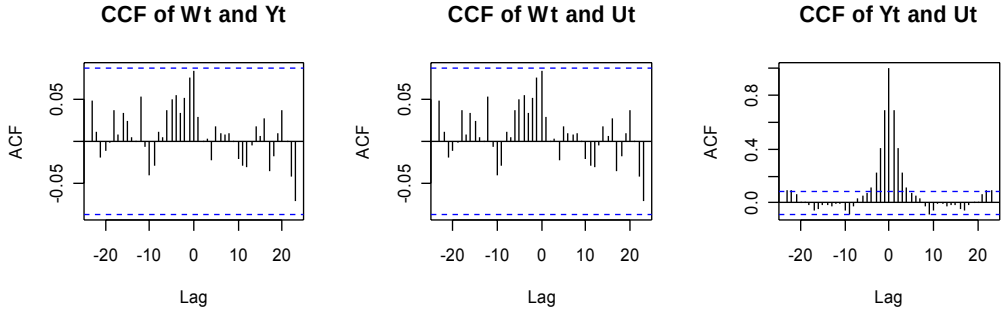
- Cross Covariance Function (CCVF)

$$\gamma_{ij}(t) = \mathrm{cov}(X_i(t+\tau), X_j(\tau)) = E[\{X_i(t+\tau) - \mu_i\}\{X_j(\tau) - \mu_j\}]$$

- Cross Correlation Function (CCF)

$$\rho_{ij}(t) = corr(X_i(t+\tau), X_j(\tau)) = \frac{\gamma_{ij}(t)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}}$$

- Cross correlation function is useful to explore the correlation between two time series. Suppose $\{W_t\} \sim WN(0,1)$ is Gaussian white noise, $Y_t = 0.7Y_{t-1}$ is AR(1) process, $U_t = 0.2Y_t$. Notice that W_t and Y_t are uncorrelated, while Y_t and U_t are correlated. Following plot carries this information.



- Spectral Representation: decompose {X(t)} into a sum of sinusoidal components with uncorrelated random coefficients. It is an analogue of Fourier representation of deterministic functions.

- Spectral Analysis: the analysis of stationary processes by means of their spectral representation. It is equivalent to “time domain” analysis based on the autocovariance function, but provides an alternate way of viewing the process, which for some applications may be more clear. For example, in the design of a structure subject to a randomly fluctuating load, it is important to be aware of the presence in the loading force of a large sinusoidal component with a

particular frequency to ensure that this is not a resonant frequency of the structure.

- Spectral Density: $f_X(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_X(h)$

Reversely, $\gamma_X(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f_X(\lambda) d\lambda$

- ARMA(p,q) Process: Autoregression with order p and moving average with order q. Let $\{X_t\}$ is stationary. For every t,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

Where $\{Z_t\} \sim WN(0, \sigma^2)$ (Gaussian white noise).

The equation can be express in the more concise form

$$\phi(B) X_t = \theta(B) Z_t$$

Where $\phi(\cdot)$ and $\theta(\cdot)$ are the pth and qth-degree polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

and

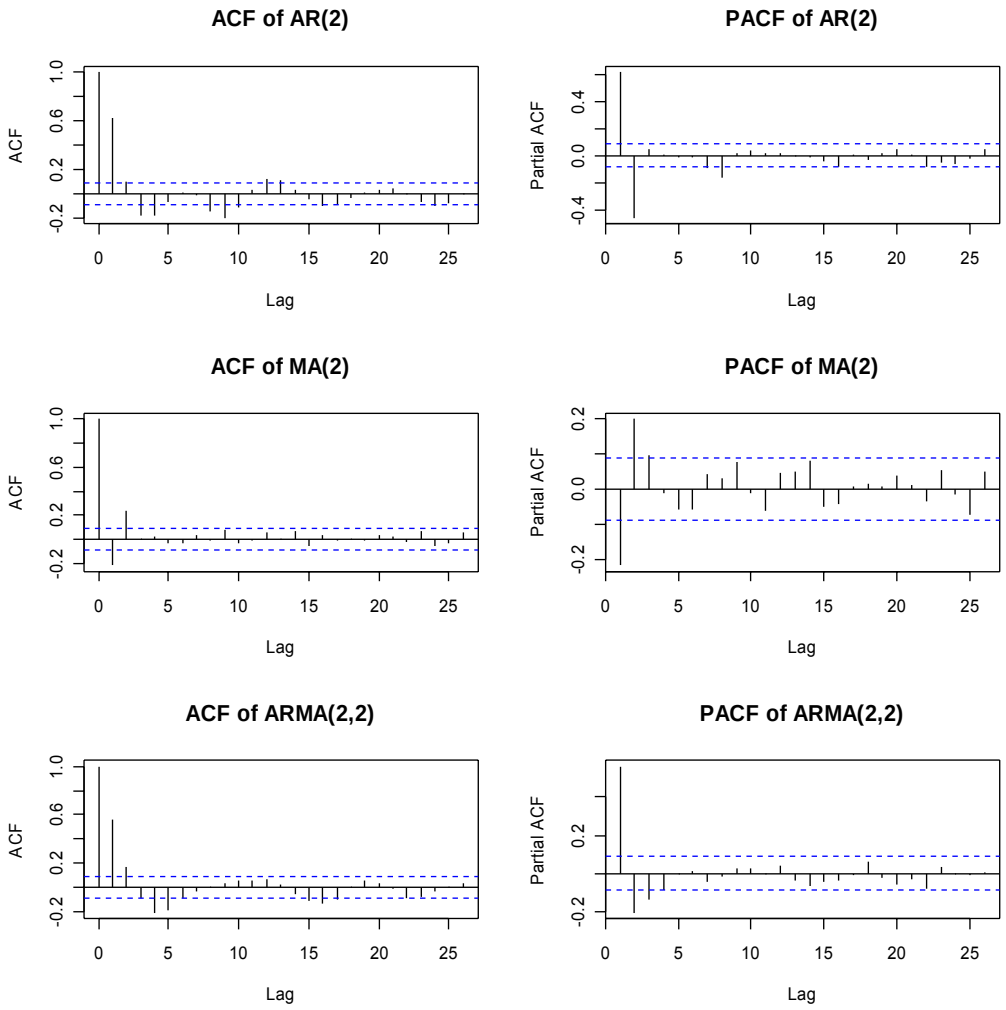
$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

and B is the backward shift operator ($B^j X_t = X_{t-j}$, $B^j Z_t = Z_{t-j}$, $j = 0, \pm 1, \pm 2, \cdots$)

- Properties of ARMA(p,q) in ACF, PACF

	ACF	PACF
AR(p)	Exponential decay	Cut off at lag p
MA(q)	Cut off at lag q	Exponential decay
ARMA(p,q)	Exponential decay	Exponential decay

Example: AR(2), MA(2), ARMA(2,2)



- ARMAX: Autoregressive Moving-Average with Extraneous input.

$$Y_t = \mu + \sum_{i=0}^{m} \beta_i X_{t-i} + W_t$$

Where Y is the response time series, X is the input time series, μ is the mean, β is the coefficient and $\{W_t\}$ is ARMA(p,q) process. Alternatively, after removing the mean and influence from input, the residual is a stationary time series and can be described by ARMA(p,q) process.

- Stem-and-leaf plot: A display that organizes data to show its shape and distribution.

In a stem-and-leaf plot each data value is split into a "stem" and a "leaf". The "leaf" is usually the last digit of the number and the other digits to the left of the "leaf" form the "stem". The number 123 would be split as:
stem 12
leaf 3

Suppose we have the following data
35, 36, 38, 40, 42, 42, 44, 45, 45, 47, 48, 49, 50, 50, 50

Defining the tens digit as the stem and the units digit as the leaf, we can reach

Stem	Leaf
3	5 6 8
4	0 2 2 4 5 5 7 8 9
5	0 0 0

It can be clearly seen in the diagram above that the data clusters around the row with a stem of 4.

The stem values could represent the intervals of a histogram, and the leaf values could represent the frequency for each interval.

One advantage to the stem-and-leaf plot over the histogram is that the stem-and-leaf plot displays not only the frequency for each interval, but also displays all of the individual values within that interval.