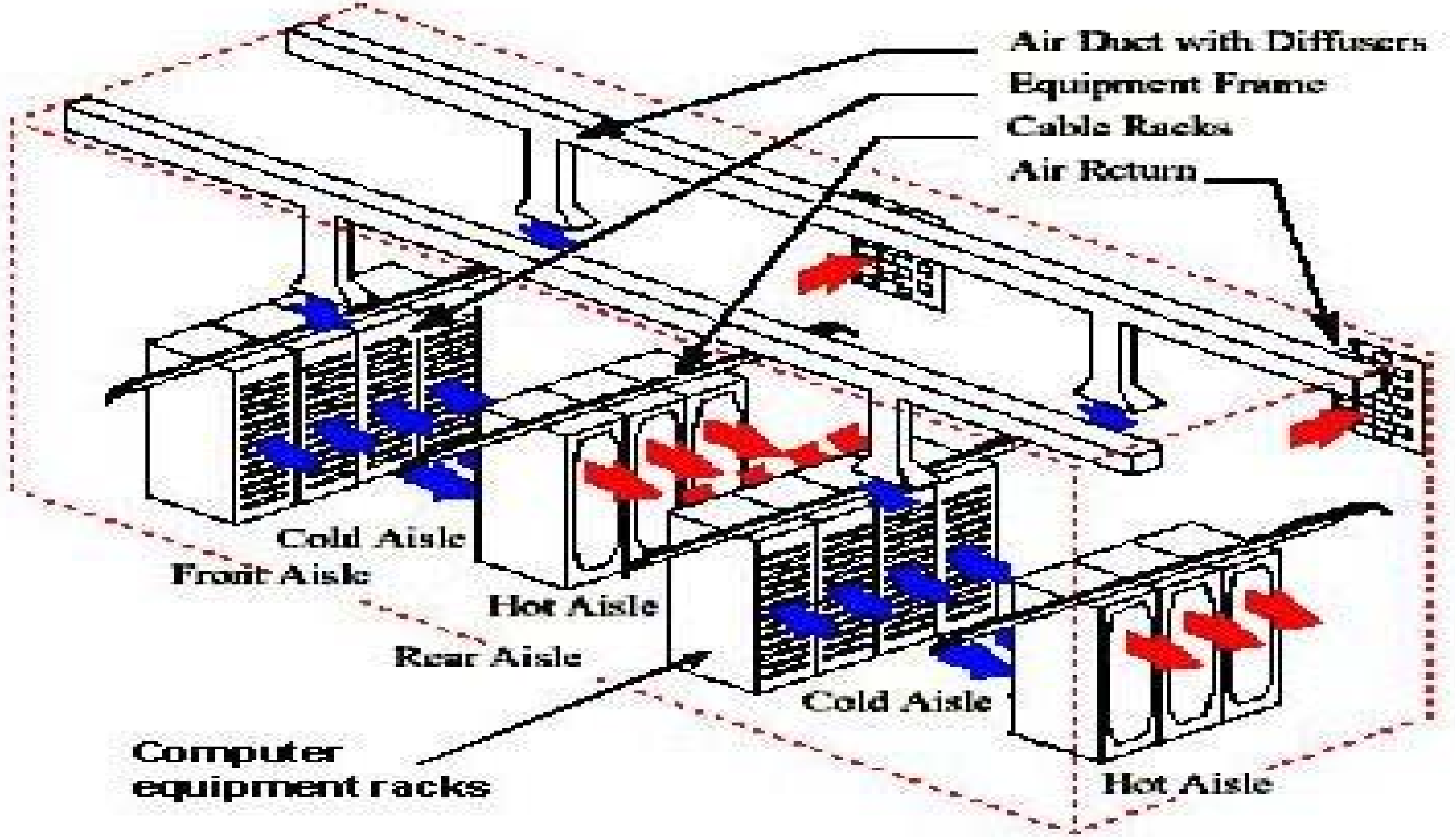
A Structural Equation Method for Modeling High-dimensional Data Center Thermal Distribution

Zhiguang Qian
Georgia Tech & University of Wisconsin-Madison

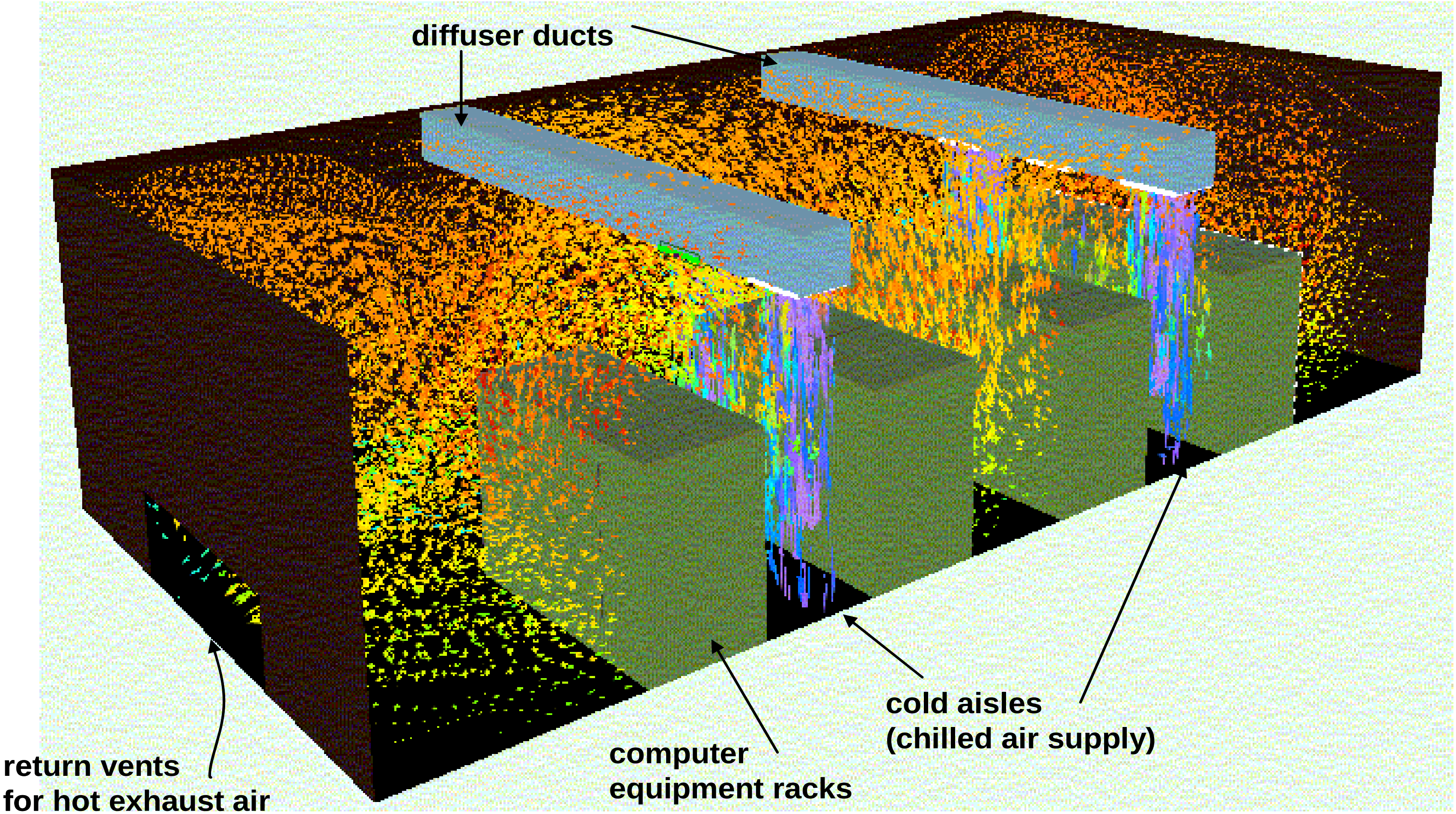
(with Yasuo Amemiya, IBM Research)

Data Center System

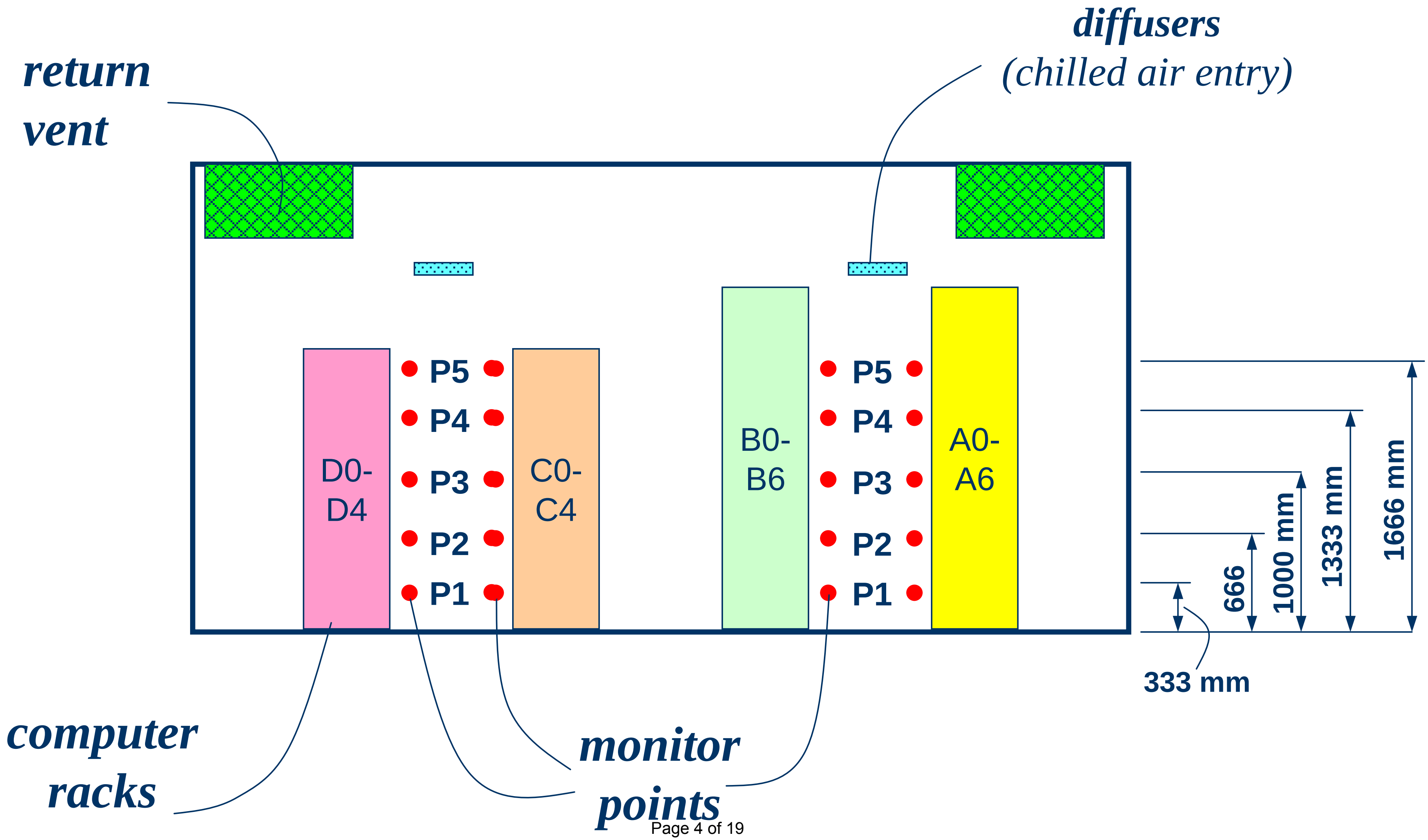


Courtesy of IBM T. J. Waston Research Center

Flotherm: Computational Fluid Dynamics (CFD) Based Computer Experiment



A Data Center with Four Rack Rows



Configuration Variables

x_1	Rack temperature rise (C)	10	15	20		
x_2	Rack power (KW)	4	12	22	28	36
x_3	Diffuser height	10 ft	Ceiling			
x_4	Diffuser location/config.	Alt1	Alt2			
x_5	Diffuser angle	0	30			
x_6	Diffuser flow rate (%)	100	80	60		
x_7	Ceiling height (ft)	12	17	22		
x_8	Hot air return vent loc.	Bot-Per	Top-Per	Bot-Par	Top-Par	
x_9	Remove/mixed power	Uniform	Alt-Zero	Alt-Half		

Data

- 148 runs generated in SAS/QC
- 4 outlier (unsuccessful) runs
- Use 144 runs
- 100 dimensional response for each run

3-Fold Latent Variable Model for Dimension Reduction

Step 1: Physical summary model for 100 responses

Step 2: Factor analysis model for physical model parameters

Step 3: Structural model relating factors to configuration variables $\mathbf{x_1}, \ldots, \mathbf{x_9}$

Step 1: Physical Summary Model

- No significant difference among horizontal positions within each rack $i = A, B, C, D$
- Linear trend for temperature measurements over different vertical heights $h = 1, \ldots, 5$
- Rack-wise linear model:

$$z_{ijh} = \alpha_i + \beta_i h + \varepsilon_{ijh}$$

- $Var(\varepsilon_{ijh}) = \sigma_i^2$
- 100-dim response is summarized by 12 parameters: α_i, β_i, $\log \sigma_i, i = A, B, C, D$

Steps 2 and 3: Structural Equation System for 12 Physical Parameters

$$\mathbf{y}_{12\times 1} = (\alpha_A, \beta_A, \log\sigma_A, \ldots \alpha_D, \beta_D, \log\sigma_D)^t$$

- Factor analysis (measurement) model

$$\mathbf{y} = \lambda + \Lambda\mathbf{f} + \mathbf{e}$$

- Structural (path) model for **f**

$$\mathbf{f} = \mathbf{G}(\mathbf{x}) + \mathbf{u}$$

Hierarchical Temperature Model

For $n = 1, \ldots, 144$,

$$
\begin{aligned}
z_{ijhn} &= \alpha_{in} + \beta_{in} h + \varepsilon_{ijhn}, \;\; i = 1, \ldots, 4, \;\; j = 1, \ldots, 4(6), \;\; h = 1, \ldots, 5, \\
\mathbf{y}_n &= \lambda + \Lambda \mathbf{f}_n + \mathbf{e}_n, \qquad \mathbf{y}_n = (\alpha_{An}, \beta_{An}, \log \sigma_{An}, \ldots \alpha_{Dn}, \beta_{Dn}, \log \sigma_{Dn})^t, \\
\mathbf{f}_n &= \mathbf{G}(\mathbf{x}_n) + \mathbf{u}_n. \qquad (1)
\end{aligned}
$$

- $\mathbf{u}_n \sim N(0, \Sigma_{\mathbf{uu}}), \;\; \mathbf{e}_n \sim N(0, \Sigma_{\mathbf{ee}}), \;\; \mathbf{u}_n \perp \mathbf{e}_n$.
- The likelihood of z_{ijhn}'s is not simple.
- **Important observation**: given $\mathbf{y}_n$, z_{ijhn} follows a linear model.
- The OLS estimator $\widehat{\mathbf{y}}_1, \ldots, \widehat{\mathbf{y}}_{144}$ is a sufficient statistics for model (1).

A Two-Stage Approach for Model Fitting

Stage 1: Estimate the physical parameter $\mathbf{y}$
by its OLS estimator $\widehat{\mathbf{y}} = (\widehat{\alpha}_A, \widehat{\beta}_A, \log\widehat{\sigma}_A, \ldots \widehat{\alpha}_D, \widehat{\beta}_D, \log\widehat{\sigma}_D)^t$
and quantify the estimation error η. $\widehat{\mathbf{y}} = \mathbf{y} + \eta$.

Stage 2: Apply structural equation analysis to $\widehat{\mathbf{y}}$.

Structural Equation Analysis of $\widehat{\mathbf{y}}$

$$\begin{aligned} \widehat{\mathbf{y}}_n &= \lambda + \Lambda \mathbf{f}_n + \mathbf{e}_n + \eta_n, \\ \mathbf{f}_n &= \mathbf{G}(\mathbf{x}_n) + \mathbf{u}_n. \end{aligned} \tag{2}$$

- $\mathbf{u}_n \sim N(0, \Sigma_{\mathbf{uu}}), \ \mathbf{e}_n \sim N(0, \Sigma_{\mathbf{ee}}), \ \eta_n \sim N(0, \Sigma_{\eta\eta}).$
- $\mathbf{u}_n \perp \mathbf{e}_n, \ \ \eta_n \perp \mathbf{u}_n, \ \eta_n \perp \mathbf{e}_n.$
- **Pseudo-likelihood estimation** (Gong and Samaniego 1981):
  1. Estimate $\Sigma_{\eta\eta}$ by the method of moment.
  2. Replace $\Sigma_{\eta\eta}$ by the moment estimator in $l(\widehat{\mathbf{y}}_1, \ldots, \widehat{\mathbf{y}}_{144})$; estimate other parameters by maximizing the resultant "imputed" likelihood function.

Factor Analysis Model

$$\alpha_A = f_1 + e_1 \qquad b_A = f_2 + e_2 \qquad \log\sigma_A = f_3 + e_3$$

$$\begin{aligned}
\alpha_B &= \mathbf{0.97}\,f_1 - 0.47\,f_2 + 0.78\,f_3 + e_4 \\
\beta_B &= \mathbf{1.08}\,f_2 - 0.21\,f_3 + e_5 \\
\log\sigma_B &= -0.01\,f_1 - 0.10\,f_2 + \mathbf{1.07}\,f_3 + e_6 \\
\alpha_C &= \mathbf{1.01}\,f_1 - 0.64\,f_2 + 1.50\,f_3 + e_7 \\
\beta_C &= \mathbf{1.31}\,f_2 - 0.45\,f_3 + e_8 \\
\log\sigma_C &= -0.02\,f_1 - 0.04\,f_2 + \mathbf{0.94}\,f_3 + e_9 \\
\alpha_D &= \mathbf{1.05}\,f_1 - 0.30\,f_2 + 0.45\,f_3 + e_{10} \\
\beta_D &= \mathbf{1.18}\,f_2 - 0.14\,f_3 + e_{11} \\
\log\sigma_D &= -0.02\,f_1 + 0.02\,f_2 + \mathbf{0.96}\,f_3 + e_{12}
\end{aligned}$$

Structural Model

$$\mathbf{f} = \mathbf{G}(\mathbf{x}) + \mathbf{u}$$

- Full quadratic-interaction model
- Drop nonsignificant second-order terms meaningfully
- The same terms kept for f_1 and f_2
  - Quadratic: x_2, x_6
  - Cross product: (x_1, x_2), (x_1, x_6), (x_6, x_7)
  - Cross interaction: (x_3, x_2), (x_3, x_6),(x_3, x_7), (x_4, x_1),(x_4, x_2),(x_4, x_6),(x_4, x_7),(x_5, x_1),(x_5, x_2), (x_5, x_6) ,(x_5, x_7),(x_8, x_2),(x_8, x_7),(x_9, x_1),(x_9, x_2),(x_9, x_7)
  - Interaction: (x_3, x_4), (x_3, x_5),(x_4, x_8), (x_4, x_9), (x_5, x_9)

Model-based Prediction

- Temperature at height h for j^{th} position in i^{th} rack

$$z_{ijh} = \alpha_i + \beta_i h + \varepsilon_{ijh}$$

- With $\beta_i > 0$, the max temperature at $h = 5$
- Predict a 95% upper bound for the max temperature

$$L_i = \alpha_i + 5\beta_i + 1.67\sigma_i$$

- Model for α_i,β_i, σ_i involves errors **e** and **u**
- Compute 95% upper prediction limit UL_i for L_i.

Configuration Decisions

- Understand difference among four racks
- Concern for a rack with highest temperature

$$\max UL = \max(UL_A, UL_B, UL_C, UL_D)$$

- Consider various “what if” scenarios for configurations
- E.g., Given physical and usage requirements

$$x_2 = 28 \qquad x_4 = 10\,\text{ft} \qquad x_7 = 17 \qquad x_9 = \text{Uniform}$$

choose the levels of $x_1 \quad x_3 \quad x_5 \quad x_6 \quad x_8$

10 Cases with Lowest Temperature

$\mathbf{x_1}$ C	$\mathbf{x_3}$ ft	$\mathbf{x_5}$ degree	$\mathbf{x_6}$ %	$\mathbf{x_8}$	UL_A	UL_B	UL_C	UL_D	maxUL
10	10	30	0.8	Bot-Per	12.50	12.85	12.04	11.43	**12.85**
10	10	30	0.6	Bot-Per	13.74	14.30	13.23	12.51	14.30
10	10	30	1.0	Bot-Per	14.06	14.45	13.79	13.16	14.45
10	ceil.	30	0.8	Bot-Per	14.05	14.51	13.65	12.96	14.51
10	ceil.	30	1.0	Bot-Per	14.45	14.95	14.26	13.62	14.95
10	10ft	30	0.8	Top-Per	15.46	15.91	15.16	14.45	15.91
10	10ft	30	0.8	Bot-Par	15.53	16.15	15.00	14.19	16.15
10	ceil.	30	0.6	Bot-Per	16.52	17.17	16.03	15.20	17.17
10	10ft	30	0.6	Top-Per	16.73	17.39	16.38	15.58	17.39
10	10ft	30	0.8	Top-Par	16.77	17.49	16.08	15.19	17.49

Interpretation

- Rack differences are small
- Inside racks B & C tend to have higher temperatures than outside racks A & D
- Use $x_1 = 10 \quad x_3 = 10\,\text{ft} \quad x_5 = 30\,\text{degree} \quad x_8 = \text{Bot-Per}$
- x_6 has non-linear effect, and $x_6 = 0.8$ gives the best configuration

Summary

- Efficient and informative experimental design
- Latent variable modeling of multivariate response, incoporating physical structure and various sources of variability
- Prediction of practically relevant quantities based on a statistical model
- Decision making based on statistical analysis of computer experiment