Using Statistical Methods to Investigate the Mapping from Initial Values to the Multiple Steady States in Complex Building Simulation Problems

Jinchao Yuan[1] and Leon R. Glicksman[2]
[1]Nexant Inc., Houston, TX, United States
[2]Massachusetts Institute of Technology, Cambridge, MA, United States

ABSTRACT

One of the challenges in some building simulations, especially with coupled airflow and thermal phenomena, is predicting the possible occurrence of multiple steady states under the same boundary conditions. Such behavior is introduced by the nonlinearity in the governing dynamics, and the initial values typically have some mapping rules to different final steady states. In some simple systems with one or two state variables, such rules may be analytically singled out by applying a dynamical system analysis. However, in complex multivariate systems, such mapping rules are very difficult to be determined analytically. In this study, we used a statistical analysis method to investigate the relations between the initial state variables that determine the mapping from the initial values to the steady states. The method was applied the simulation data of a natural ventilated multi-level building case that can exhibit three different steady states. Statistical summaries of both the single-zone initial value characteristics (mean and distribution), and the inter-zonal correlations between state variables were examined. Combined with existing knowledge about the case, the results revealed the formation mechanism of each possible steady state, which is related to different levels of buoyancy/wind force fine-tuning based on single zone temperatures and inter-zonal temperature correlations. The findings promoted the understanding of the governing physics of a relative complex natural ventilation system and provided potential methods to analyze the solution multiplicity related problems in such a system.

INTRODUCTION

In building simulations, multiple solutions can occur when the system is nonlinear. A typical occurrence of such nonlinear behaviors is in natural ventilation systems (e.g., Linden (1999), Li and Delsante (2001), and Heiselberg et al. (2004)), where the coupled heat transfer and fluid transport phenomena can be highly nonlinear and the system behavior can largely dependent on the history state of the system (see Linden (1999), and Hunt and Linden (2004)). The existence of such nonlinear behaviors can be a challenge to the simulation, design, and controls of building ventilation systems (e.g., Chen and Jiang (1992), Axley et al. (2002), Spindler (2004)).

Recent findings show that the nonlinear behaviors of buildings are largely related to the initial values. Starting from different initial values, the building can reach different steady states when certain conditions are met. For example, in Yuan and Glickman (2005), the initial values are found to be critically important for a typical example of competing buoyancy. There is a critical temperature, which itself represents a locally unstable steady states of the system, that divides the initial value space. Any initial temperature lower than that point would result in a wind-dominated steady states, and a initial value higher than that will result in a buoyancy-dominated steady states. For a two-zone system, the initial value dependency also exist and the problem can also be investigated using analytical or numerical methods. For example, Yang, Xu, and Li (2006) showed that two different steady states can be achieved in the initial state space. The initial value state space can be divided into two distinct attraction regions for the two steady state solutions. By applying an analytical analysis, the separation line between the two attractions regions can be approximated. Later, Yuan (2007) investigated the two-variable natural ventilation system with thermal mass that can demonstrate multiple solutions. Based on the eigen characteristics of the unstable solutions on the phase plot, the separation between the two attraction regions can be numerically computed with accuracy.

However, for multi-zone buildings with multiple state variables, the analysis on the state space separation for the attraction regions can be very difficult either analytically or numerically due to the large number of state variables. This also leads to the difficulties in understanding the physics of the system—how the system evolves from some set of initial values to a particular final steady states. In order to analyze the attraction regions for different steady states and understand the their forming conditions in the general multi-zone building simulation problems, in this study, we take an empirical statistics approach to analyze the problem: first we categorize the initial value data from a large number of randomly trial values based on the final steady states they reached; then we use statistical analysis methods to summarize the common characteristics for the initial conditions that converges to (a) particular solutions. Using these methods, the mapping relationship from the initial values to the final steady states can be probed and analyzed based on some basic understanding of the physical and dynamical system model.

SIMULATION METHODS AND DATA

Mathematical models and search methods

The simulation conducted here is a coupled dynamic multi-zone airflow and thermal simulation. In order to generalize the model, we assume that the building can be represented by N zones in a multi-zone model frame work. By assuming that each zone has a uniform temperature T_i, the system's state equations can be written as

$$\begin{aligned} \rho_i c_p V_i \frac{dT_i}{dt} &= c_p \sum_{j:q_{ji}>0} \rho_j q_{ji} T_j - \rho_i c_p \sum_{j:q_{ij}>0} q_{ij} T_i + S_i, \\ \text{for } \forall\, i &= 1 \cdots N. \end{aligned} \tag{1}$$

where q_{ij} is the volumetric flow rate from zone i to zone j. The flow rate q_{ij} is usually a function of the zone temperatures and other known parameters. In general, if we denote the state vector $[T_1, T_2, \cdots, T_n]'$ as $\mathbf{T}$, the equation above can be generally written as

$$\frac{d\mathbf{T}}{dt} = f(\mathbf{T}, \mathbf{q}), \tag{2}$$

where $\mathbf{q}$ is the flow rate vector that is formed by all the flow rates q_{ij}. It is a function of the zone temperatures $\mathbf{T}$ and other parameters such as the flow network resistance and wind pressure. However, at each particular time, the flow rate $\mathbf{q}$ can be solved by a multi-zone airflow network model (see the theoretical background in Walton and Dols (2005)) for a fixed $\mathbf{T}$.

The underlying equations solved are the coupled time-dependent energy balance equation (see Hensen (1999), Li and Delsante (2001), Yuan and Glicksman (2008)) for each zone and the pressure-flow balance for each time steps (see Walton and Dols (2005)). The equations are coupled and they are solved using a "ping-pong" coupling strategy described by Hensen (1999).

Since it is known that the system could potentially reach different solutions from differen initial temperature, we conduct solution search by starting the simulation from differen initial values. By observing the final steady states the system finally converges from these random initial value, we can discover how many stable steady states the system can have. Figure 1 shows the simulation method in a two-variable case in the phase plot. Starting from 10 different randomly selected initial values (the filled points), the system are observed to converge to two

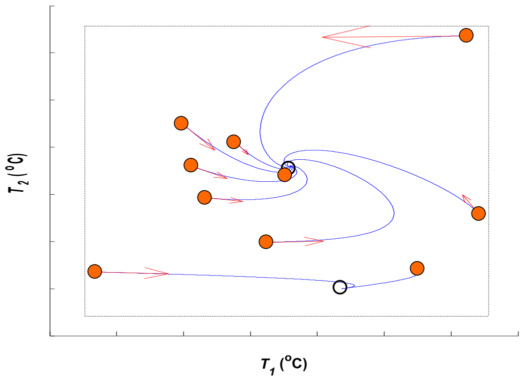


Figure 1: Random initialization for a 2-D system

different steady states (the unfilled points) based on different initializations.

In the solution for the 7-zone building (which will be introduced shortly below) in this study, we structure the initial values selection by drawing randomly initial values from a certain probability distribution. In this study, the selected distribution is uniformly distributed between 5 to 35°C, a reasonable range for summer and shoulder season temperature, where natural ventilation can be used. Since we have 7 dimensions (7 zones) in the current simulation, the 7 state variables (temperatures of Zones 1 to 7) forms a joint uniformed distribution from 5 to 35 °C. The 1024 samples are drawn in an identical distribution and they are independent to each other.

Case introduction

We will apply the dynamical simulation and the search methods to a sample multiple-zone building to test the models and methods. A hypothetical but reasonably configured building (Figure 2) with 7 zones (consecutively labeled as 1 to 7) is constructed for this purpose. The building has three stories, with a chimney type natural ventilation design. Each story has an occupied office space, whose width is 10m and the depth (perpendicular to the paper) is 20m. The floor and roof height is 3m and the heat generation in each occupied office is assumed to be 6000W (30W/m^2).

The building has a top window on the roof and three exterior windows on the office side, one on each floor. The top window on the roof is 1.5m^2. The exterior windows on each floor have an area of 1.8m^2. We assumed that the center height of the exterior window(s) on each floor is at the center height of the exterior wall on that floor (i.e. the window center's relative elevation is 1.5m). Further, we assumed that only one-way flow occurs at all the exterior and internal openings in this study. The internal openings include the vertical connections between the office and the

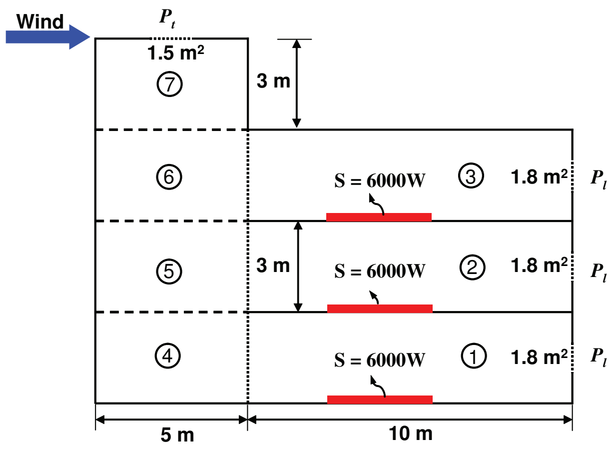


Figure 2: The geometries of the simulated building

chimney areas and horizontal openings between the adjacent zones 4, 5, 6, and 7 on Figure 2.

The building is subject to a combined wind and buoyancy force. The buoyancy force is introduced by the temperature difference between the interior and the outdoor air. The outdoor air temperature is assumed to be to 20°C, a reasonable temperature for natural ventilation. The wind force is caused by the wind pressure difference between the top (roof) window and the leeward office windows. In this case, we specified the wind pressure coefficient at the roof (for P_t) as -0.3 (a lift force) and the wind pressure coefficients at the office windows (P_l) as -0.5. Therefore, the net wind pressure is

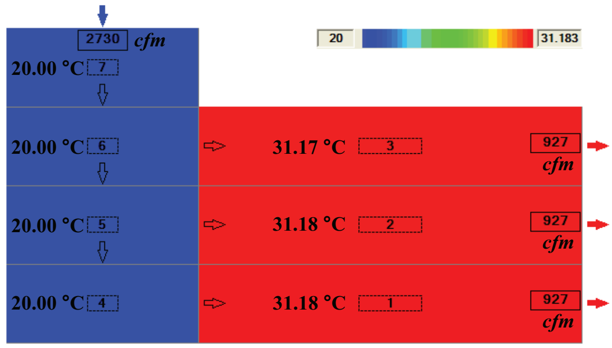
$$P_w = P_t - P_l = (-0.3+0.5)\cdot\frac{1}{2}\rho_a v_{ref}^2 = 0.1\rho_a v_{ref}^2. \quad (3)$$

Please note that the building is set up as a simulation case only and it may not be a case that precisely matches every single detail of a physical model. The main purpose of the illustration of the case is to demonstrate one numerical example where multiple steady states can occur when coupled building airflow and thermal simulation are performed.

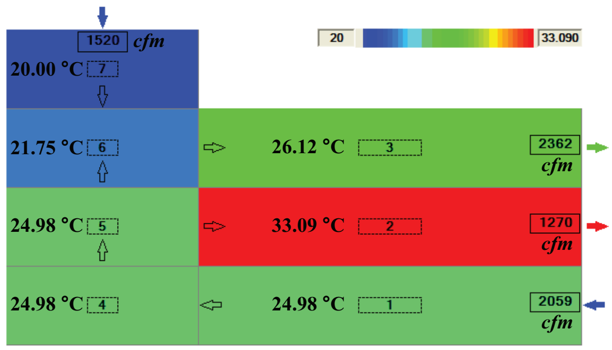
Multiple solutions found

Starting from the 1024 random selected initial values from the initial state space, the system shown in Figure 2 finally reached three steady states. The solution values (temperatures and airflows) can be visualized in Figure 3.

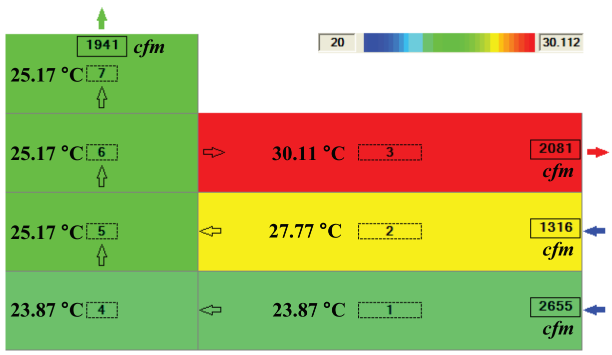
Physically, the three solutions stand for three different competition stages of the wind and buoyancy. The first steady state (SS1) stands for a pure wind-dominated steady state. The second steady state (SS2) is also wind-dominated. However, some buoyancy flows are formed in the lower level of the building. Therefore, the upper offices (level 2 and 3, or Zones 2 and 3) have outward



(a) Fully wind-dominated, SS1



(b) Top-down wind and bottom induction, SS2



(c) Buoyancy-dominated, SS3

Figure 3: Three possible stable steady states obtained in a multi-zone building model

flow caused by dominant wind force; and the lowest level office (level 1, or Zone 1) has incoming flow caused by dominant buoyancy force in the lower level of the chimney. The third one is an overall buoyancy dominated flow pattern, with an upward flow coming out of the top atrium chimney. However, there are still some outgoing flows from the office on the third floor, as a result of the downward wind force.

All the three solutions are locally stable, due the dynamical system simulation method used in the simulation. This means that they can resist small perturbations and will not diverge to other steady state solutions under infinitesimal disturbances.

STATISTICAL ANALYSIS

In order to understand how the system evolves from an initial value to the final steady state, we categorize the 1024 randomly selected initial values by the final steady states that they converged to. The sample sizes of the three categories are listed in Table 1.

Table 1: Sample size for the three categories

	To SS1	To SS2	To SS3	Total
Samples	666	88	270	1024

In theory, the three categories should contain three different mapping relations from the initial values to the final steady states. We will apply a statistical analysis on each data category to explore their characteristics and the potential mapping from these initial values to the corresponding steady states. To do this we will proceed in three steps:

First, we will examine the mean initial temperature of each zone for each of the categories, trying to find out the average temperature for the zones to reach different steady states. Secondly, we will examine the statistical distribution of the initial temperatures of each zone in each category. Thirdly, we will examine the inter-zonal relation through the correlations between two zones, since it is generally true that in a complex system, some variables are dependent on each others.

Through these three steps, we also expect to further understand the physics of the simulated systems such as the buoyancy and wind force requirements for the system to converge to different steady states. As an additional step, we will finally examine how well the three steady states can be numerically classified (separated) based on the final steady states that they converge to, which may be used in some inverse model based methods.

Mean initial zone temperatures

The mean temperature is the simplest value to examine the general characteristics of categorized data. For each of the zones, we average the temperature of every sample in the same category and Figure 4 shows the mean zone temperatures of the 7 zones for the three categories. In general, the mean initial temperature values for all the 7 zones are the lowest for the solutions that converge to SS1. Physically, since SS1 is a pure wind-dominated steady states, the low mean initial temperatures indicate that in the initial stage, the space temperature of building should be low enough in order not to build up too much buoyancy force.

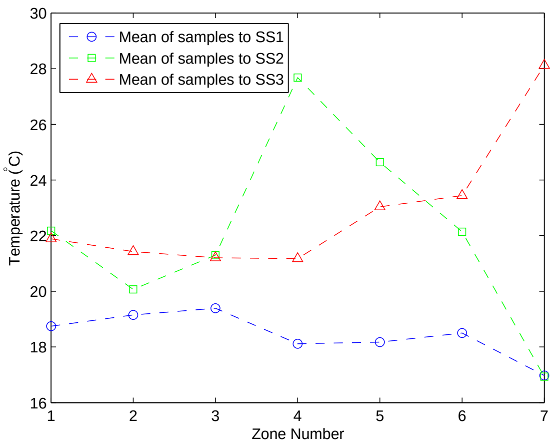


Figure 4: Mean initial zone temperatures for three different steady states

For the solutions that converge to SS2 and SS3, the average initial zone temperatures are much higher that those for SS1. In particular, when comparing the mean initial zone temperature for the SS2 and the SS3 categories, a few meaningful observations were found. First, the average temperature for Zone 7, which is very low for the SS2 category, indicates that the buoyancy force on the upper chimney of the building should not be so high (statistically) in order to achieve SS2. This is opposite to the SS3 category, where the average initial temperature for Zone 7 is much higher—the highest among these three categories and much higher than the other two.

On the other hand, for SS2, the average temperature of the lower zones (Zones and 4), especially Zone 4, has very high average initial zone temperature. This indicates that for such a case (SS2, upper two levels wind dominated and lower one level buoyancy dominated), the lower level needs to accumulate enough buoyancy force in order to achieve the buoyancy in the lower level.

Initial zone temperature distributions

Other than the mean zone temperature, the distributions of the single zone temperatures for each category are also of importance. Physically, if there is a very tight distribution of a single variable, it could mean that the particular variable can have some strong influences on the final outcome. Although strictly speaking, the above statement needs to be combined with sample size, the examination of the distribution of the zone temperature is still very important to reveal the important factors and to understand the underlying governing physics.

Figure 5 shows the box plot of the single zone temperatures for different cases.

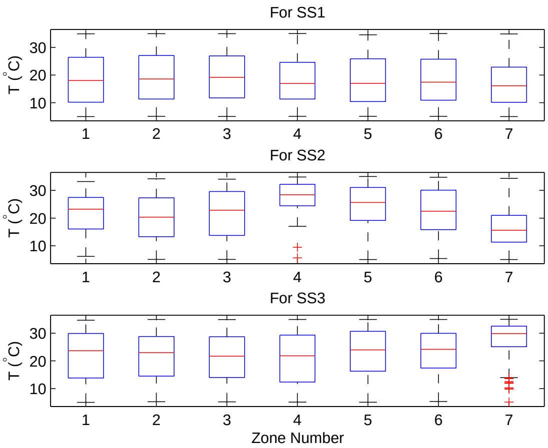


Figure 5: Zone temperature distribution

In the plot, we can find that temperature distributions of Zone 7 for all the three categories are quite distinctive between Zones 1&2 and Zone 3. This means that is in general the initial temperature of Zone 7 should be kept to a high value and in a relatively narrow band in order to reach this steady state.

For SS2, the temperature of Zone 4 is also important, in that the distribution also has a relatively narrow distribution (and most of Zone 4 sample values have a high initial temperature) in order to reach SS2. This patten of Zone 4 for SS2 is quite different from those for SS2 and SS1, where Zone 4 temperature distribution is more spreading (or flat).

Some of relations (mostly for shown as a group) for the zone temperatures can also be observed in the boxplot. For example, the Zone 4 and Zone 7 temperatures for the SS2 category (the initial value samples that converge to SS2) tend to have an opposite trend. The relationship between the temperatures of two zones can be more accurately described by correlation coefficients between them, in that the correlation coefficient provides more accurate information about relationship in individually paired samples, which cannot be read from the box plot. This will be the topic of the next subsection.

Initial zone temperature correlations

The mean and distribution are only focused on a single variable and see what kind of information can be analyzed from there. However, in a complex system, the initial temperature of different zones can be correlated such as the Zone 4 and Zone 7 temperatures distribution for SS2 described in the previous subsection. In this section, we will use another descriptive statistical measure–the correlation coefficient between two random variables—to measure the degree of dependency of the between any of the two state variables in the categorized initial samples.

Tables 2 to 4 show the correlation matrixes of zone temperatures for the three different categories. They are all symmetric matrixes and only the upper triangle part was shown to reduce redundancy. The correlations between each two zones (one indicated by the row index number, and the other indicated by the column index number) are between -1 and 1. A correlation of 1 means the strongest positive correlation between the two variables. Physically, this means that these two variables tend to move in the same trend. A correlation of -1 means the strongest negative correlation, which physically means that the two variables tend to move in complete opposite trends. A correlation of zero means no statistical dependency between the two variables, although strictly speaking, it is not equivalent to independency.

Table 2: Correlation matrix for samples that reach SS1

Zone	1	2	3	4	5	6	7
1	1.0000						
2	-0.0098	1.0000					
3	-0.0975	0.0347	1.0000				
4	-0.1428	0.0554	0.0467	1.0000			
5	-0.1433	-0.0807	0.0727	-0.1069	1.0000		
6	-0.0820	-0.0223	0.0030	-0.0815	-0.0423	1.0000	
7	-0.0704	-0.1038	-0.0383	-0.0168	-0.1157	-0.2106	1.0000

Table 3: Correlation matrix for samples that reach SS2

Zone	1	2	3	4	5	6	7
1	1.0000						
2	-0.0131	1.0000					
3	-0.0049	-0.0531	1.0000				
4	-0.0986	0.0242	-0.0416	1.0000			
5	0.2573	-0.0520	0.0300	-0.1896	1.0000		
6	-0.0839	-0.1136	0.0287	0.0200	-0.3571	1.0000	
7	-0.4227	-0.0475	-0.0174	-0.2987	-0.1721	-0.2658	1.0000

Here we take the correlations in the SS2 category (Table 3) for example. A strong negative correlation is observed between Zone 1 and Zone 7. This means that trends of the initial temperatures for these two zones should be opposite to each other in order to form a proper buoyancy and wind competition level for a flow pattern corresponding to SS2 shown in Figure 3(b). There is also strong correlation shown between Zones 4 and 7 (where some of the

Table 4: Correlation matrix for samples that reach SS3

Zone	1	2	3	4	5	6	7
1	1.0000						
2	-0.1132	1.0000					
3	-0.0433	0.0051	1.0000				
4	0.1655	0.0676	-0.1057	1.0000			
5	-0.0269	-0.0770	-0.1975	0.0296	1.0000		
6	-0.0463	0.0317	-0.0094	-0.1010	-0.1717	1.0000	
7	-0.2927	-0.0962	0.0650	-0.1686	-0.2136	-0.1969	1.0000

bulk negative correlation can be shown in the boxplot for the SS2 category in Figure 5).

Other correlations such as those between Zones 5 and 7, Zones 6 and 7 are also strong and negative, which means these zones should move to the opposite direction to fine tune a reasonable wind and buoyancy competition level for SS2 to develop. The right amount of buoyancy is critical in this case, since SS2 is the first state that shows some level of buoyancy development in the lower level.

For SS3 (see Table 4), similar correlation relationships between Zones 5, 6, and 7 also exist and they are also negative, but in a smaller magnitude than those shown for the SS2 category (in Table 3. This is because SS3 is a state where buoyancy force further dominates. Therefore, the opposite trend of these zone temperatures (Zones 5, 6, 7) no longer need to be so strong to maintain a relative low buoyancy level to confine the buoyancy pattern only in the lowest level.

Linear and quadratic classification study

In some building simulations and control applications (e.g. in developing control strategies), inverse modeling can also be important in determining the final states of the building by learning from some existing data or status. In order to examine how the obtained simulation data can potentially be used in an inverse model, in this subsection we conducted a classification study to classify the initial values into different categories (SS1, SS2, and SS3 in this case) based on the statistical learning on the samples and the steady state outcomes listed in Table 1.

Here we presented the results for two simple classifying models: the linear classifier (linear discriminant analysis) and nonlinear (quadratic) classifier. For both models, the classifiers are first trained based on the 1024 initial value samples listed in Table 1; then randomly generated new initial values will be used to test the accuracy of the trained classifiers.

The linear classifier uses only linear combinations of the dependent variable to classify the data. For the 1024 initial value samples used in this simulation, the error rate of the training model is 23%. Using another set of independent samples to test the trained model, we obtained an error rate of 25% on the test data set.

Realizing that the the boundaries between different attraction regions are typically nonlinear, we then used a quadratic classifier to examining the potential of classifying the model with nonlinear boundaries. Using the quadratic classifier, the training data set shows an error rate of 17%, and the classification error rate is 20% using a set of newly generated independent test samples.

From this study, it is found that numerically, the initial values can be potentially classified using numerical learning and training methods, with an acceptable error rate. The nonlinear (quadratic) classifier performs slight better than the linear classifier in this presented case, which may due to the nonlinear nature of the "true" region separation for different steady states in the state space.

CONCLUSION

In this study, we applied a statistical analysis to the initial value seeds that converge to different steady states in a dynamical multi-zone simulation with coupled airflow and thermal solution.

We found that the individual initial zone temperature means and distributions have distinctive patterns for different steady states. The patterns agree with the underlying physics of the forming mechanism of the studied steady states. The patterns in the mean temperatures of the sample initial values reveal that different levels of initial buoyancy should be initialized in order to achieve the three different flow patterns. Further, the patterns in the statistical distributions of the zone temperatures also reveal that the buoyancy force requirement in different zones, especially in the top chimney Zone 7 and the lower Zone 4, have significant impact in determining which steady states the system will reach from an initial value. For example, in forming the buoyancy dominated steady state SS3, the atrium Zone 7 temperature is a strong decision variable in that most of the initial value samples that converged to SS3 have high zone 7 temperatures, and the distribution has a relatively narrow and distinctive band. Similar conclusions can be reached for the lower atrium temperature of Zone 4 in reaching steady state SS2. In general, Zone 4 needs to have a high temperature and needs to be confined to a relative narrow band in order to form the adequate initial buoyancy level for SS2.

The correlation relationship between the initial zone temperature is also found to be an important factor in forming different levels of initial buoyancy and wind competitions to reach different steady states. For the SS2 category, the initial temperatures of the lower Zone 4 and the top chimney Zone 7 are strongly negatively correlated. Other correlations between Zones 5, 6, 7 are also important to shaping the buoyancy level for SS2. These correlation relationship reveals the requirement for the initial zone temperatures to coordinate together to fine tune a reasonable level of buoyancy domination in the lower zones to draw buoyancy flows and a reasonable wind domination in the upper zones in order to allow wind flows going outward. For SS3, a steady states which has a higher buoyancy dominance, similar correlation relationships also exist for its initial values. However, the level of correlation (the absolute values) is much less than that for SS2.

We also investigated the possibility of classifying the initial values based on trained linear and nonlinear classification models on the simulation data. It is found that a classification rate of about 75–80% can be obtained using either a simple linear or quadratic classification model.

The case presented in this study can serve as an example that multiple steady states can exist in a relatively complex building in coupled thermal and airflow simulations. Further, the study also provided a way to identify possible steady states through a numerical search method and to investigate the dynamics of these states through statistical analysis in these simulation problems.

REFERENCES

Axley, James W., S. Emmerich, S. Dols, and George N. Walton. 2002. "An Approach to the Design of Natural and Hybrid Ventilation System for Cooling Buildings." *Proceedings of Indoor Air*.

Chen, Qingyan, and Zheng Jiang. 1992. "Significant questions in predicting room air motion." *ASHRAE Transactions* 98 (1): 929–939.

Heiselberg, Per, Yuguo Li, A. Andersen, M. Bjerre, and Z. Chen. 2004. "Experimental and CFD evidence of multiple solutions in a naturally ventilated building." *Indoor Air* 14:43–54.

Hensen, J. L.M. 1999. "A comparison of coupled and de-coupled solutions for temperature and air flow in a building." *ASHRAE Transactions* 105 (2): 962–969.

Hunt, G. R., and Paul F. Linden. 2004. "Displacement and mixing ventilation driven by opposing wind and buoyancy." *Journal of Fluid Mechanics* 527:27–55.

Li, Yuguo, and A. Delsante. 2001. "Natural ventilation induced by combined wind and thermal forces." *Building and Environment* 36:59–71.

Linden, Paul F. 1999. "The fluid mechanics of natural ventilation." *Annual Review of Fluid Mechanics* 31:201–238.

Spindler, Henry C. 2004. "System Identification and Optimal Control for Mixed-mode Cooling." Ph.D. diss., the Massachusetts Institute of Technology, Cambridge, MA, USA.

Walton, George N., and Stuart W. Dols. 2005. *CONTAM 2.4 User Guide and Program Documentation*. National Institute of Standard and Technology. NISTIR 7251.

Yang, Lina, Peng Xu, and Yuguo Li. 2006. "Nonlinear dynamics analysis of natural ventilation in a two-zone building: Part A—Theoretical analysis." *HVAC&R Research* 12 (2): 231–255.

Yuan, Jinchao. 2007, September. "Transition dynamics between the multiple steady states in building ventilation systems: From theory to Applications in optimal ventilation controls." Ph.D. diss., the Massachusetts Institute of Technology.

Yuan, Jinchao, and Leon R. Glickman. 2005. "Multiple steady states in a combined buoyancy and wind driven natural ventilation system: necessary conditions and initial values." *Proceedings of Indoor Air*. Beijing, China, 1207–1212.

Yuan, Jinchao, and Leon R. Glicksman. 2008. "Multiple steady states in combined buoyancy and wind driven natural ventilation: The conditions for multiple solutions and the critical point for initial conditions." *Building and Environment* 43 (1): 62–69.

NOMENCLATURE

P	pressure (Pa)
q	flow rate (m^3/s)
S	heat source strength (W)
T	temperature (°C)
V	volume (m^3)
v	velocity (m/s)
ρ	density (kg/m^3)