Huang, Y., Li, J., Li, B., Duan, R., Lin, C.-H., Liu, J., Shen, X., and Chen, Q. 2015. "A method to optimize sampling locations for measuring indoor air distributions," *Atmospheric Environment*, 102, 355-365.

A Method to Optimize Sampling Locations for Measuring Indoor Air Distributions

Yan Huang[a], Jianmin Li[a], Bingye Li[a], Ran Duan[a], Chao-Hsin Lin[b], Junjie Liu[a], Xiong Shen[a,*], and Qingyan Chen[a,c]

[a]School of Environmental Science and Engineering, Tianjin University, Tianjin 300072, China
[b]Environmental Control Systems, Boeing Commercial Airplanes, Everett, WA 98203, USA
[c]School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

[*]Corresponding email: shenxiong@tju.edu.cn

Abstract:

Indoor air distributions, such as the distributions of air temperature, air velocity, and contaminant concentrations, are very important to occupants' health and comfort in enclosed spaces. When point data is collected for interpolation to form field distributions, the sampling locations (the locations of the point sensors) have a significant effect on time invested, labor costs and measuring accuracy on field interpolation. This investigation compared two different sampling methods: the grid method and the gradient-based method, for determining sampling locations. The two methods were applied to obtain point air parameter data in an office room and in a section of an economy-class aircraft cabin. The point data obtained was then interpolated to form field distributions by the ordinary Kriging method. Our error analysis shows that the gradient-based sampling method has 32.6% smaller error of interpolation than the grid sampling method. We acquired the function between the interpolation errors and the sampling size (the number of sampling points). According to the function, the sampling size has an optimal value and the maximum sampling size can be determined by the sensor and system errors. This study recommends the gradient-based sampling method for measuring indoor air distributions.

Keywords: Gradient method, Kriging interpolation, CFD simulation, Error analysis

1. Introduction

In indoor air environments, the optimization of parameters such as air velocity, temperature, and contaminant concentrations is important for the health and comfort of occupants. To assess the detailed distributions of these parameters, two primary methods can be applied: numerical simulations by computational fluid dynamics (CFD), and in-situ measurements. CFD simulations are inexpensive, but they may not accurately predict the distributions because of the approximations used in turbulence modeling and numerical algorithms. In-situ measurements, although time-consuming and expensive, are more reliable. Furthermore, even in numerical simulations, a certain amount of experimental data is often needed for validating the computed results (Chen and Srebric, 2002; Liu et al., 2012). Therefore, it is preferable to conduct in-situ measurements.

Both optical and point-wise measurement method can be applied to acquire the distributions of indoor air parameters. The optical measurement method uses optical anemometry techniques such as particle streak velocimetry, particle tracking velocimetry, and particle image velocimetry to measure air distributions by acquiring and processing the reflected signals of particles seeded in the flow. This method can determine air velocity distributions in a local field (Cao et al., 2014). However, in indoor spaces with complex geometry, occupants and other objects may block the light from the optical anemometer, which makes it difficult to measure an entire region (Liu et al., 2012).

The point-wise method measures the air parameters with point sensors such as anemometers, thermocouples, and tracer-gas samplers (Liu et al., 2012; Li et al., 2014). In comparison with the optical method, the point-wise method is more adaptable to a complex space because the sensors can be located flexibly in spaces where the optical anemometry cannot take measurements. However, the accuracy of the interpolated air distributions based on the point sensor data is highly dependent on the sampling size and locations (Swileret al., 2006; Coetzee et al., 2012). The sampling size means the number of sampling locations and sampling methods means how the sampling locations were selected. In order to reduce the time requirement and labor costs for the measurements, the sampling size should be as low as possible, but the use of too few sampling points can result in poor spatial resolution. Laurenceau and Sagaut (2008) found that the grid method, which is the most commonly applied method in engineering fields, requires a large sampling size in order to provide good results. The grid method may not be optimal in determining the sampling locations.

This paper reports our effort in proposing a method for determining the optimal sampling location and sampling size. We have also investigated the relationship between sampling location and the accuracy of air parameter fields obtained in measuring indoor air environment.

2 Research Methods

2.1 Sampling method

Several methods are available for determining sampling locations, including grid, unstructured triangular mesh, Latin hypercube sampling, sequential, and gradient-based methods. The grid method uses equal intervals along a sampling direction (Laurenceau and Sagaut, 2008; Carvajal et al., 2010; Coetzee et al., 2012). The unstructured triangular mesh method uses an unstructured mesh to spread the sampling points such that they are adapted to the boundary of the sampling domain (Persson and Strang, 2004; Coetzee et al., 2012). The sampling domain is the plane or volume where measurements were conducted in a 2D or 3D indoor space, respectively. The Latin hypercube sampling method is an enhanced random sampling method that divides the sampling domain into cells with equal intervals and then sets one sampling point at a random position in each cell (Laurenceau and Sagaut, 2008; Nissenson et al., 2009; Coetzee et al., 2012). It is widely used in geo-statistics but does not seem to be useful for indoor air measurements. This is because indoor spaces are relatively small, and we can acquire the exact spatial coordinates easily. The sequential method sets a few initial sampling points and then adds points one by one to improve the interpolation accuracy, until the desired sampling size has been obtained (Jin et al., 2002). However, the sequential method is computationally expensive, as it sets only one point at a time, and the whole field interpolation must be calculated each time (Coetzee et al., 2012). Jouhaud et al. (2007) proposed a gradient-based sampling method that determines new sampling points in regions with a large gradient. The gradient-based method seems to be scientific, simple, and computationally inexpensive. It was therefore selected for this study.

The gradient-based method uses equal intervals along a direction if the gradient of air parameter is small. If the gradient is large, one or more points are added between the two original sampling points until the differences between two adjacent points are sufficiently small when compared with the maximum gradients in the sampling domain. However, it is difficult to determine the sampling locations for the gradient-based method because the gradient is unknown before the start of the experiment. One could estimate the gradient from experience, by identifying, for example, the regions with large velocity and temperature gradients. Such estimation may not be easy for an actual indoor space where the flow can be complex. Thus, this investigation recommends using a CFD simulation to identify the regions with large parameter gradients. The information required by CFD is typically known, such as the thermo-fluid boundary conditions. With the air distribution predicted by CFD, one can determine the sample x_i by calculating the gradient coefficient, α, as (Jouhafud et al., 2007):

$$\alpha = \frac{\left\|grad(\phi)\right\|}{\max\left\|grad(\phi)\right\|} \quad (1)$$

where, ϕ represents an air parameter such as velocity, temperature, or contaminant concentration. $\left\|grad(\phi)\right\|$ is the gradient of the flow parameter, while $\max\left\|grad(\phi)\right\|$ is the largest gradient of this flow parameter in the sampling domain.

If $\alpha > \alpha_0$, the gradient is considered to be sufficiently large to require the addition of more sampling points in the region. We chose $\alpha_0 = 0.15$ which had been recommended to be the optimal choice to avoid the use of too numerous points in sampling domain with large gradient (Jouhaud et al., 2007). By starting with a coarse grid of sampling points, one can flag a sampling point if $\alpha > 0.15$. The percentage of flagged points is:

$$\eta = \frac{w_f}{w_t} \quad (2)$$

where w_f is the number of flagged points in the sampling domain and w_t the total number of points in the domain. If $\eta \leqslant \eta_0$, then a point is refined into two points. This process is repeated for all sampling points until $\eta > \eta_0$. The recommended value of η_0 was in the range of 0.5-0.75. Here we chose η_0=0.6(Quirk, 1996).

It should be noted that in order to compare the grid and gradient-based sampling methods, the total grid number (sampling size) should be the same in both methods. However, since the gradient-based method splits points for regions with a large gradient, the starting sampling size for this method should be smaller than that for the grid method.

2.2 Kriging interpolation method

The data obtained from the sampling positions by the grid or gradient-based method can be interpolated to form field distributions. Several interpolation methods are available, such as polynomial methods, radial basis function methods, inverse distance weighting methods, Kriging methods, etc.

The polynomial method uses Taylor expansion equations to express the values at places not measured by the values at the measured points. The coefficients of the polynomial are estimated by minimizing the mean square error of the expansion equations (Shen et al., 2013). This is the most mature method for interpolation and requires the lowest number of sampling points for modeling; however, it may not as accurate as Kriging methods (Wang and Shan, 2006).

The radial basis function method estimates the value at places not sampled, by use of a basis function (such as linear, cubic, thin plate spline, multiquadric, and Gaussian functions) (Gutmann, 2001). However, the radial basis function method cannot be used to determine measurement errors (Jin et al., 2002).

The inverse distance weighting method determines the value between the measured points as the weighted sum of the measured data at the surrounding positions. The weights in this method are inversely related to the distances between the sampled point and estimated point, with a constant power or a distance-decay parameter to adjust the diminishing strength in relationship to increasing distance. However, the method cannot provide the variances of the estimated values at points not measured (Lu and Wong, 2008).

The Kriging method interpolates a value between measured points in the same manner as the inverse distance weighting method. The Kriging method estimates the value under the unbiasedness condition, and the weights are calculated by minimizing the variance of the error between the estimated and actual values (Kleijnen, 2009). This method has been applied in many different studies (Simpson et al., 1998; Jeong et al., 2005; Moral et al., 2006; Sampson et al., 2013), and all of them have reported highly satisfactory results. Therefore, the method was selected for this investigation.

The Kriging method is a statistical method which was originally proposed in the field of geo-statistics (Clark, 1977; McBratney et al., 1981; Persicani, 1995). It estimates the value between measured points $Z(x)$ as the weighted sum $\sum_{i=1}^{n} \lambda_i$ of the measured data at the surrounding positions $Z(x_1)\ldots Z(x_i)$, as illustrated in Fig. 1. Various assumptions in determining the weights lead to different Kriging variants, such as the simple, ordinary, and universal Kriging methods. The simple Kriging method assumes that the expectation of $Z(x)$ is constant and known over the entire domain. The ordinary Kriging method assumes that the expectation of $Z(x)$ is constant but unknown in the neighborhood of the estimation point, x. The universal Kriging method regards the expectation of $Z(x)$ as one that fits a linear or higher-order (quadratic, cubic, etc.) trend model of the spatial x-, y- coordinates of the measured point, x_i. This study used the ordinary Kriging method for the interpolation. The method adopts a stationary assumption, which assumes that statistical properties (such as expectation, variance, covariance, semivariogram function, etc.) only rely on the distance between the measured and not measured points (Kleijnen, 2009).

The Kriging estimator of $Z(x)$ at a point not measured, x, is acquired by a linear regression model (Goovaerts, 1997):

$$\hat{Z}(x) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{3}$$

$x, x_i \in S$, $S \in \Re^2$

where S is the sampling domain, $\hat{Z}(x)$ the estimator of $Z(x)$, $Z(x_i)$ the measured data at sample point x_i, and λ_i the weight (Goovaerts, 1997).

The difference between the estimated $\hat{Z}(x)$ and actual $Z(x)$ is called the estimation error. The Kriging method calculates $\hat{Z}(x)$ by minimizing the variance of the errors with the unbiasedness constraint. In order to obtain an unbiased estimator, the

expectation of the estimation error is equated to zero. For obtaining an optimal function estimator, the variance of the estimation error must be minimized. Therefore, the set of weights λ_i satisfies the following equations:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j C(x_i, x_j) + \mu = C(x_i, x) \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \quad (4)$$

where x_i, x_j are the sampled points, and x is the point not measured but to be estimated. A Lagrange parameter μ is introduced to calculate λ_i without affecting the equality. $C(\cdot,\cdot)$ is the covariance.

The ordinary Kriging method can also be expressed as a semivariogram function that determines the relationship between distance and the variance (*Var*) of the data:

$$\gamma(h) = \frac{1}{2} Var[Z(x) - Z(x+h)] \quad (5)$$

where h is the distance between the point x and $x+h$. The relationship between the semivariogram function and covariance function is:

$$\gamma(h) = C(0) - C(h) \quad (6)$$

where $C(h)$ represents the covariance between two points with a distance of h. Therefore, Eq. (4) can be rewritten as:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \gamma(x_i, x_j) - \mu = \gamma(x_i, x) \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \quad (7)$$

Because of the stationary assumption, $\gamma(x_i, x_j) = \gamma(h_{i,j})$, $\gamma(x_i, x) = \gamma(h_i)$. $h_{i,j}$ represents the distance between the measured point x_i and x_j, while h_i is the distance between the sampling point x_i and the estimated point x.

Thus, λ_i can be solved by Eq. (7). Once the Kriging weights (and Lagrange parameter) have been obtained, both the Kriging estimator and the Kriging variance can be determined.

In this study, we applied a spherical model to represent the semivariogram in a two-dimensional space:

$$\gamma(h) = \begin{cases} 0.5 S_2 \times \frac{h}{a}[3 - (\frac{h}{a})^2] & h < a \\ S_2 & h \geq a \end{cases} \quad (8)$$

Where S_2 is the variance of the variable across the entire region and a is the range.

209 $a = mL$ (9)

210 $$L = \sqrt{(x_{\max} - x_{\min})^2 + (y_{\max} - y_{\min})^2} \quad (10)$$

211 where m is range fraction m=0.3. L is the Euclidian distance between (x_{max}, y_{max}) and
212 (x_{min}, y_{min}) of the source zone (Tecplot, 2006).

213 Fig. 2 shows the procedure how the gradient-based method and interpolation method
214 is used to measure the parameters. In the beginning, the CFD simulation is applied to
215 obtain the distributions of the parameters to be measured. A sampling plane is
216 selected in the domain on which the sampling points are chosen. The gradient
217 distribution is determined from the simulation results and applies to calculate the
218 distribution of α in the sampling plane. The sampling plane is split into identical
219 coarse cells. If any cell with an α higher than 0.15, the cell center will be flagged as a
220 sampling point. If the number of flagged cells takes up more than 60% of the total
221 cells, the sampling points can be determined. Otherwise, the flagged cell will be
222 refined into two more cells until 60% of total cells become flagged cells.

223 2.3 Evaluation of interpolation accuracy

224 This investigation used absolute interpolation error and uniformity of errors to
225 evaluate the interpolation accuracy with different methods for the determining of
226 sampling locations.

227 The absolute interpolation error is the difference between the interpolated distribution
228 with the grid or gradient-based method and the actual distribution, i.e.

229 $$E_p = \left| \phi_{pm_F} - \phi_{pm_K} \right| \quad (11)$$
230 $m \in S$

231 where E_p represents the absolute interpolation error between the Kriging-interpolated
232 value and the actual value. P represents the air parameters such as velocity (V, V_x, V_y,
233 V_z), temperature (T), and contaminant concentrations (C). The variable ϕ_{pmk} is the
234 Kriging-interpolated air parameter P at point m, and ϕ_{pmF} is the actual air parameter P
235 at point m. Here S is the sampling plane.

236 The uniformity of errors evaluates whether or not different sampling methods would
237 lead to an uneven interpolation error distribution:

238 $$U = \sqrt{\frac{1}{n-1} \sum_{m=1}^{n} (E_{pm} - \overline{E})^2} \quad (12)$$

where U is the uniformity of errors for an air parameter, n the sampling size, E_{pm} the interpolation error of the air parameter at point m, and $\overline{E}$ the averaged interpolation error of the air parameter in the sampling plane.

3 Case analysis

3.1 Case 1: Air distribution in an office room

Figure 3 is a schematic of the office room used for the study. The air parameters concerned were air velocity, air temperature, and contaminant concentration as simulated by a tracer gas, SF_6. The room was 5.16 m in length, 3.65 m in width and 2.43 m in height. The room had an air diffuser, two desktop computers, six lights, two occupants, and several pieces of furniture. These were simplified as rectangular blocks, but they were under thermo-fluid conditions similar to those that would actually be present in an office. The SF_6 sources were placed above the simulated occupants.

Yuan et al. (1999) conducted experimental measurements of the air distributions in an office room. They measured air velocity, air temperature, and SF_6 concentration at nine locations in the room. However, the data points were too few to be interpolated into field distributions for evaluating the sampling methods. We performed CFD simulations of the air distributions in our room, and the computational results obtained were in good agreement with the experimental data. Because the CFD results were obtained using a very high numerical grid resolution, we can use these results to generate ”measured data” for the office room for any sampling size and at any sampling location. Because of the limited space available in this paper, we evaluate the grid and gradient-based methods only at the mid-plane in the y-direction, as shown in Fig.3. The mid-plane was applied as the sampling plane to represent a part of sampling domain.

Fig. 4 depicts the air temperature, airflow, and SF_6 concentration field at the mid y-plane as calculated by CFD. The cold air from the diffuser caused induction of the surrounding air and relatively high air velocity near the diffuser. In the main flow region, the air velocity was quite low and uniform. The results show that the displacement ventilation system created air temperature stratification in the room. The air temperature gradient was larger near the diffuser and floor. In addition, the SF_6 concentration was high in the middle of the plane and upper region of the room but low in the lower region of the room.

Fig. 5 shows the gradient coefficient calculated from the CFD results and the sampling point distribution determined by the gradient-based method. The CFD results show that the air velocity, air temperature, and SF_6 concentration gradients were high near the air supply inlet, outlet, and right wall. The SF_6 concentration gradient was also high in the middle of the plane. Thus, more grid points were used in those regions to capture the changes due to the large gradients. Table 1 show the sampling

sizes used with the grid and gradient-based methods for the three air parameters. For the same air parameter, the sizes were similar for the two sampling methods.

Table 1 lists the average interpolation errors and the uniformity of errors for the flow domain calculated by Eqs. (11) and (12), respectively, for the grid and gradient-based sampling methods. With a similar sampling size, the gradient-based method yielded a smaller surface-average interpolation error and higher uniformity than the grid method. For the gradient-based method, the absolute interpolation error for the air temperature, air velocity, and SF_6 concentration were 6.28%, 6.08%, and 3.21%, respectively, was lower than those for the grid method. This difference arose because the Kriging interpolation accuracy was sensitive to the sampling point interval in the high gradient region. The gradient-based method used more points in the high gradient region, and therefore the results were more accurate. In the low gradient regions, the Kriging interpolation accuracy was not sensitive to the distance between points, and thus the error would not be increased significantly. Furthermore, the uneven distribution of the sampling locations did not lead to an uneven error distribution in the gradient-based method.

Fig. 6 illustrates the absolute interpolation error distributions for the three air parameters obtained by the gradient-based and grid methods. The gradient-based method had higher interpolation accuracy than the grid method in the region close to the diffuser, where the gradients were high. This difference occurred because more sampling data were taken in that region. In the main flow region, the two methods had smaller interpolation errors, although fewer sampling points were used in the gradient-based method than in the grid method. This was because the gradient-based method used more sampling points in the high gradient region, and thus there were fewer sampling points in the low gradient region than with the grid method. Since the gradients of the air parameters in the main flow region were low, the use of a slightly lower number of sampling points would not cause a notable error.

It should be noted that the interpolation errors for air velocity were quite high near the floor with both methods. This is because we selected a gradient coefficient α_0 of 0.15, which did require more sampling points in the region where the gradient was high but not very high. To solve this problem, the gradient coefficient should be smaller than 0.15 (Jouhaud et al., 2007). Further investigation should be made on choosing the proper α_0. However, since the region was small, the error may not have been important. Nevertheless, it can be seen that there is a trade-off between sampling size and interpolation error.

3.2 Case 2: Air distribution in an aircraft cabin

For further demonstration and verification, this investigation used the methods to measure the air distribution in a functional aircraft cabin. The measurements were conducted in the economy-class cabin of the airliner. Because this cabin was very

long, our measurements were conducted in only four rows, as shown in Fig. 7. The four-row section of the aircraft cabin was 3.1 m long, 3.2 m wide, and 2.1 m high, and the seats were fully occupied by heated manikins. The geometry of the cabin and manikins was quite sophisticated, as was the air distribution. Air was supplied to the cabin through air slot diffusers near the ceiling on both sides and through some of the overhead gaspers. The gaspers were round nozzles mounted at the overhead board that could supply fresh air with high air velocity jets to the breathing zone of passengers for improving thermal comfort and perhaps air quality. The driving forces in the cabin included inertial forces from the air slot diffusers and gaspers, and buoyancy forces from the thermal plumes from the heated manikins and walls. The flow condition inside the cabin was in the transitional-to-turbulent (Zhang et al., 2009; Liu et al., 2013).

In order to see the influent of gaspers on the air distribution, we opened a portion of gaspers and measured the according airflow and temperature field. This study measured the mean velocity and x-, y-, z- velocity components with ultrasonic anemometers. Because the sensor size of the anemometers was 0.03 m in diameter, the finest possible sampling resolution was 0.06 m × 0.06 m, as shown in Fig. 8(a), and it was used as the baseline for assessing different sampling methods. In this experimental case, the air velocity gradient can be unknown. Gradient information can be obtained by performing a CFD simulation for the case, as was done in this investigation. The gradient information can also be estimated on the basis of personal experience, such as in the jet regions and in regions with strong thermal plumes. Using the CFD simulation, this study determined the sampling point distribution for the gradient-based method as illustrated in Fig. 8(b). The number of sampling points for the grid method was similar to that for the gradient-based method, as shown in Fig. 8(c). The sampling sizes for the baseline case, gradient-based method case, and grid method case were 302, 94, and 90, respectively.

Figure 9 illustrates the airflow pattern in the cross plane as obtained for the baseline case and with the gradient-based and grid sampling methods based on the measured results. Both methods provided a quantitative description of the jet and thermal plume in the cabin. However, the gradient-based method was able to describe the detailed interactions between the jets from the gasper and diffusers, and the method interpolated the y-velocity more accurately than did the grid method. The grid method failed to describe the airflow pattern in the upper right section of the experimental domain.

Fig. 10 shows the absolute interpolation errors for the mean velocity and the x-, y-, z-velocity components with the gradient-based and grid sampling methods, in comparison with the data for the baseline case. In the jet regions within the L-shaped box in Fig. 10, a large discrepancy is seen between the distribution generated by the grid method and that in the baseline case. This is because too few sampling points were placed near the diffuser and gasper regions, which led to higher errors in the Kriging

interpolation. In contrast, the gradient-based method used more sampling points in the jet regions, which provided better accuracy. In the main flow region, the interpolated distribution determined by the gradient-based method was similar to that determined by the grid method, even though the gradient-based method used fewer sampling points. However, in the region with strong interaction between the jet and thermal plume (indicated by the rectangular box in Fig. 10), both methods had large errors. This is because the gradient varied rapidly in the region, and thus a gradient coefficient $\alpha_0 =$ 0.15 may not be sufficiently small to capture the flow features.

Table 2 presents the averaged interpolation errors and uniformity of errors for the two sampling methods in comparison with the baseline case. With similar sampling sizes, the gradient-based method yielded a smaller interpolation errors and better uniformity (lower value) for the air velocity distributions. The measuring accuracy for velocity was 0.014 m/s and 0.019 m/s by gradient-based method and grid method, respectively. The interpolation errors for the x-, y-, and z-velocity components by the gradient-based method were lower than those by the grid method as shown on Table 2. Thus, the gradient-based method performed better than the grid method in measuring the complex airflow field in the aircraft cabin environment.

4. Discussion

This investigation also studied the relationship between sampling size and Kriging interpolation errors for the office room. Fig. 11 shows the averaged interpolation errors for air temperature, air velocity, and SF_6 concentration distributions with sampling sizes ranging from 18 to 11,104. The relationships between the interpolated errors and sampling size were obtained through curve fitting. When the sampling size was increased from several points to 2,000, the interpolation errors decreased dramatically. However, a further increase in the sampling size did not greatly reduce the errors. This result implies that a CFD simulation could be run before the experimental measurements are conducted, as we demonstrated here, in order to determine a sampling size that would provide acceptable errors. This step would identify the optimal trade-off between accuracy and effort.

Sensor and system errors in the experiment must be considered when deciding on the sampling size, which should then be used to determine the thresholds of the interpolation errors. For example, Yuan et al. (1999) used thermocouples, hot-spherical anemometers, and a photoacoustic multi-gas analyzer to measure the air temperature, air velocity, and SF_6 concentration in the office room. The sensor and system errors were 0.3 K for air temperature, 0.01 m/s for air velocity, and 0.01 ppm for SF_6 concentration. These errors can be used as the minimal interpolation errors. Therefore, it is not meaningful to select $E_T < 0.3$ K, $E_V < 0.01$ m/s, and $E_C < 0.01$ ppm. One can then find the maximal sampling size for the experimental measurements through the fitting function between the sampling size and the interpolation errors. For the mid y-section of the office, this investigation determined corresponding maximal sampling sizes of

60, 660, and 272, respectively, for air temperature, air velocity, and SF_6 concentration.

5. Conclusion

This research proposes a sampling point distribution determined by the gradient-based method for measuring air parameters with point sensors, and the Kriging interpolation method for interpolating the measured data to form field distributions. The methods have been successfully applied to an office room and a section of an economy-class airliner cabin to obtain air (air temperature, air velocity, and tracer-gas concentration) distributions from point data.

On one hand, the gradient-based method can obtain the air distributions with greater accuracy than the grid method with a similar sampling size. Our error analysis shows that the gradient-based sampling method has 32.6% smaller error of interpolation than the grid sampling method. Furthermore, the errors in the gradient-based method are more uniform than those in the grid method. On the other hand, the grid method is easier to use because it does not require prior knowledge of the gradient distribution. However, with the gradient-based method, this gradient information can be obtained from a CFD simulation or from experience. Thus, the gradient-based method is recommended.

The interpolation errors can be expressed as a function of sampling size for measuring air distributions in an indoor space. The errors decrease rapidly when the sampling size is increase from a very small number to a moderate number, but a further increase in the number of sampling points does not lead to significantly better results. When the sensor and system errors are assumed to be the same as the minimal interpolation errors, the maximal sampling size for the measurements can be determined.

6. Acknowledgement

The authors are grateful for the financial support of this research by NSFC (Grant No. 51408413), the National Basic Research Program of China (the 973 Program) through Grant No. 2012CB720100 and the Center for Cabin Air Reformative Environment (CARE) at Tianjin University, China.

7. References

Cao, X., Liu, J., Jiang, N., Chen, Q., 2014. Particle image velocimetry measurement of indoor airflow field: A review of the technologies and applications. Energy and Buildings 69, 367-380.

429 Carvajal, B., Aboal, J.R., Fernández, J.A., Real, C., Carballeira, A., 2010. Influence of
430 roads and inhabited areas on metal concentrations in terrestrial mosses. Atmospheric
431 Environment 44, 3432-3441.

432 Chen, Q., Srebric, J., 2002. A procedure for verification, validation, and reporting of
433 indoor environment CFD analyses. HVAC&R Research 8, 201-216.

434 Clark, I., 1977. Practical kriging in three dimensions. Computers & Geosciences 3,
435 173-180.

436 Coetzee, W., Coetzer, R.L.J., Rawatlal, R.,2012. Response surface strategies in
437 constructing statistical bubble flow models for the development of a novel bubble
438 column simulation approach. Computers & Chemical Engineering 36, 22-34.

439 Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford
440 University Press. New York, USA.

441 Gutmann, H.M., 2001. A radial basis function method for global optimization. Journal
442 of Global Optimization 19, 201-227.

443 Jeong, S., Murayama, M., Yamamoto, K., 2005. Efficient optimization design method
444 using Kriging model. Journal of Aircraft 42, 413-420.

445 Jin, R., Chen, W., Sudjianto, A., 2002. On sequential sampling for global
446 metamodeling in engineering design. In: Proceedings of ASME 2002 International
447 Design Engineering Technical Conferences and Computers and Information in
448 Engineering Conference. American Society of Mechanical Engineers, pp. 539-548.

449 Jouhaud, J.C., Sagaut, P., Montagnac, M., Laurenceau, J., 2007. A surrogate-model
450 based multidisciplinary shape optimization method with application to a 2D subsonic
451 airfoil. Computers & Fluids 36, 520-529.

452 Kleijnen, J.P.C., 2009. Kriging metamodeling in simulation: A review. European
453 Journal of Operational Research 192, 707-716.

454 Laurenceau, J., Sagaut, P., 2008. Building efficient response surfaces of aerodynamic
455 functions with Kriging and Cokriging. AIAA Journal 46, 498-507.

456 Li, F., Liu, J., Pei, J., Chen, Q., 2014. Experimental study of gaseous and particulate
457 contaminants distribution in an aircraft cabin. Atmospheric Environment 85, 223-233.

458 Liu, W., Mazumdar, S., Zhang, Z., Poussou, S.B., Liu, J., Lin, C.H., Chen Q., 2012.
459 State-of-the-art methods for studying air distributions in commercial airliner cabins.
460 Building and Environment 47, 5-12.

461 Liu, W., Wen, J., Chao, J., Yin, W., Shen, C., Lai, D., Lin, C.H., Liu, J., Sun, H., Chen,
462 Q., 2012. Accurate and high-resolution boundary conditions and flow fields in the
463 first-class cabin of an MD-82 commercial airliner. Atmospheric Environment 56,
464 33-44.

465 Liu, W., Wen, J., Lin, C.H., Liu, J., Long, Z., Chen, Q., 2013. Evaluation of various
466 categories of turbulence models for predicting air distribution in an airliner cabin.
467 Building and Environment 65, 118-131.

468 Lu, G.Y., Wong, D.W., 2008. An adaptive inverse-distance weighting spatial
469 interpolation technique. Computers & Geosciences 34, 1044-1055.

470 McBratney, A.B., Webster, R., Burgess, T.M., 1981. The design of optimal sampling
471 schemes for local estimation and mapping of of regionalized variables—I: Theory and
472 method. Computers & Geosciences 7, 331-334.

473 Moral, F.J., Álvarez, P., Canito, J.L., 2006. Mapping and hazard assessment of
474 atmospheric pollution in a medium sized urban area using the Rasch model and
475 geostatistics techniques. Atmospheric Environment 40, 1408-1418.

476 Nissenson, P., Packwood, D.M., Hunt, S.W., Finlayson-Pitts, B.J., Dabdub, D., 2009.
477 Probing the sensitivity of gaseous Br_2 production from the oxidation of aqueous
478 bromide-containing aerosols and atmospheric implications. Atmospheric Environment
479 43, 3951-3962.

480 Persicani, D., 1995. Evaluation of soil classification and kriging for mapping
481 herbicide leaching simulated by two models. Soil Technology 8, 17-30.

482 Persson, P.O., Strang, G., 2004. A simple mesh generator in MATLAB. SIAM Review
483 46, 329-345.

484 Quirk, J.J., 1996. A parallel adaptive grid algorithm for computational shock
485 hydrodynamics. Applied Numerical Mathematics 20, 427-453.

486 Sampson, P.D., Richards, M., Szpiro, A.A., Bergen, S., Sheppard, L., Larson, T.V.,
487 Kaufman, J.D., 2013. A regionalized national universal kriging model using Partial
488 Least Squares regression for estimating annual $PM_{2.5}$ concentrations in epidemiology.
489 Atmospheric Environment 75, 383-392.

490 Simpson, T.W., Mauery, T.M., Korte, J.J., Mistree, F., 1998. Comparison of response
491 surface and kriging models for multidisciplinary design optimization. American
492 Institute of Aeronautics and Astronautics 98, 1-16.

493 Shen, X., Zhang, G., Bjerg, B., 2013. Assessments of experimental designs in
494 response surface modelling process: Estimating ventilation rate in naturally ventilated

livestock buildings. Energy and Buildings 62, 570-580.

Swiler, L.P., Slepoy, R., Giunta, A.A., 2006. Evaluation of sampling methods in constructing response surface approximations. In: Proceeding of 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, number AIAA-2006-1827, Newport, RI, Vol. 201.

Tecplot 360 users manual, 2006. Tecplot, Inc., Bellevue, WA.

Wang, G.G., Shan, S., 2006. Review of metamodeling techniques in support of engineering design optimization. Journal of Mechanical Design 129, 370-380.

Yuan, X., Chen, Q., Glicksman, L.R., Hu, Y., Yang, X., 1999. Measurements and computations of room airflow with displacement ventilation. ASHRAE Transactions 105, 340-352.

Zhang, Z., Chen, X., Mazumdar, S., Zhang, T., Chen, Q., 2009. Experimental and numerical investigation of airflow and contaminant transport in an airliner cabin mockup. Building and Environment 44, 85-94.

Table 1

The sampling sizes, averaged absolute interpolation errors, and uniformity of errors with the grid and gradient-based sampling methods.

	Sampling method	Sampling size	$\overline{E}$	U
T (K)	Gradient-based	292	0.170	0.392
	Grid	292	0.181	0.397
V (m/s)	Gradient-based	211	0.015	0.028
	Grid	215	0.016	0.034
C (ppm)	Gradient-based	237	$1.084*10^{-2}$	0.016
	Grid	237	$1.120*10^{-2}$	0.022

Table 2

The averaged interpolation errors and uniformity of errors for the gradient-based and grid sampling methods compared with the baseline case.

	Sampling Method	Sampling size	$\overline{E}$	U
V (m/s)	Gradient-based	94	0.0144	0.0187
	Grid	90	0.0191	0.0196
V_x (m/s)	Gradient-based	94	0.0107	0.0150
	Grid	90	0.0151	0.0157
V_y (m/s)	Gradient-based	94	0.0078	0.0101
	Grid	90	0.0101	0.0111
V_z (m/s)	Gradient-based	94	0.0142	0.0203
	Grid	90	0.0193	0.0213

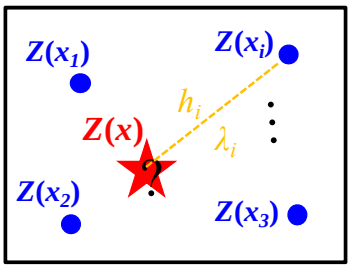
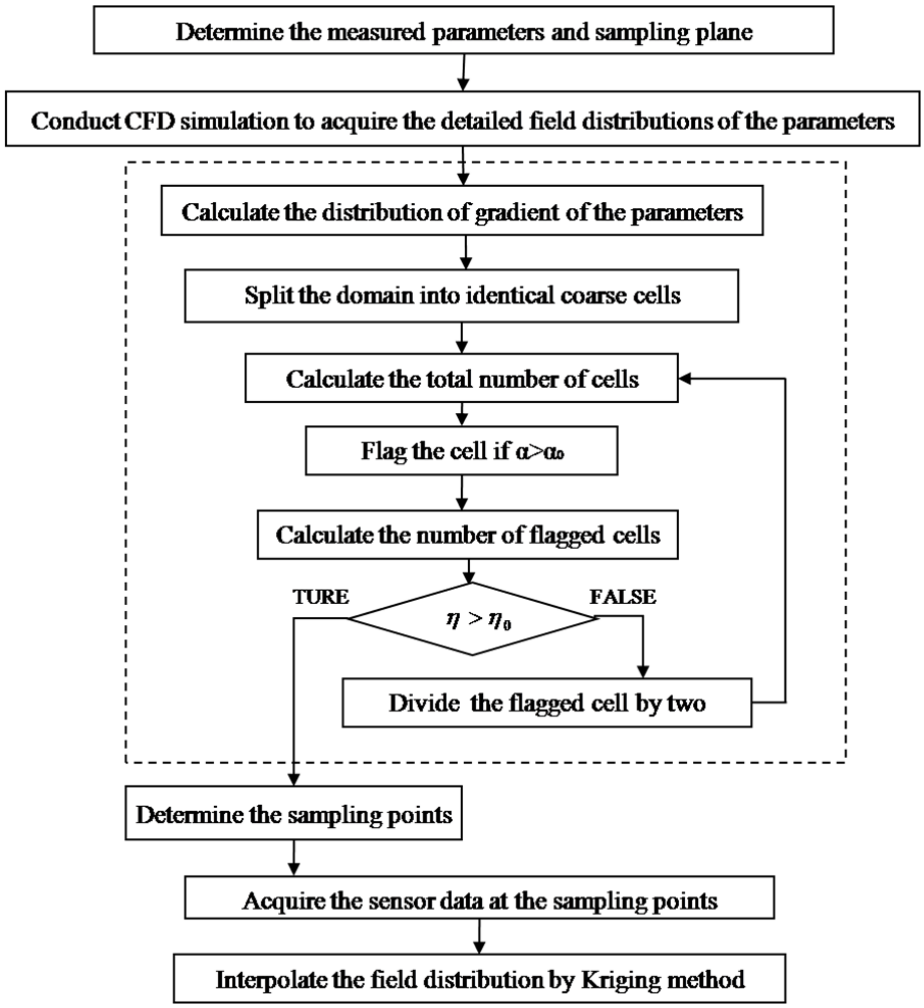
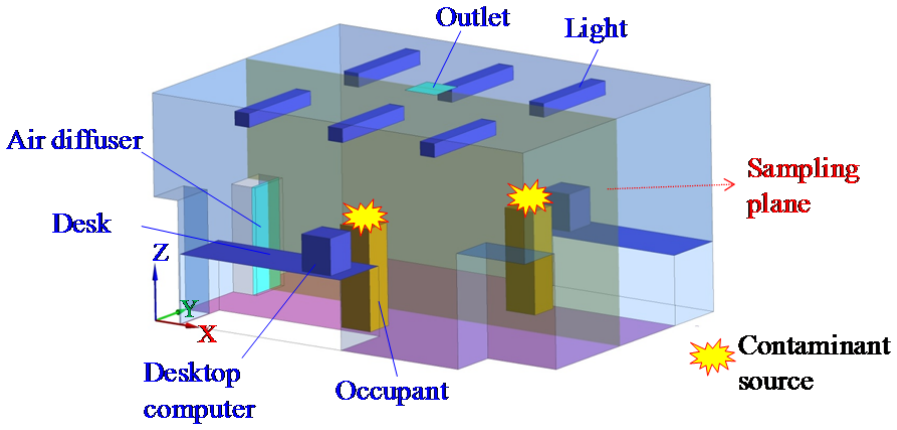


Fig. 1. Theory of the Kriging interpolation.



523

524 **Fig.2.** Flow chart of the gradient-based method and the Kriging interpolation method
525 with α_0= 0.15 and η_0=0.6.



526

527 **Fig. 3.** Schematic of the office and the sampling plane where the comparison was
528 made.

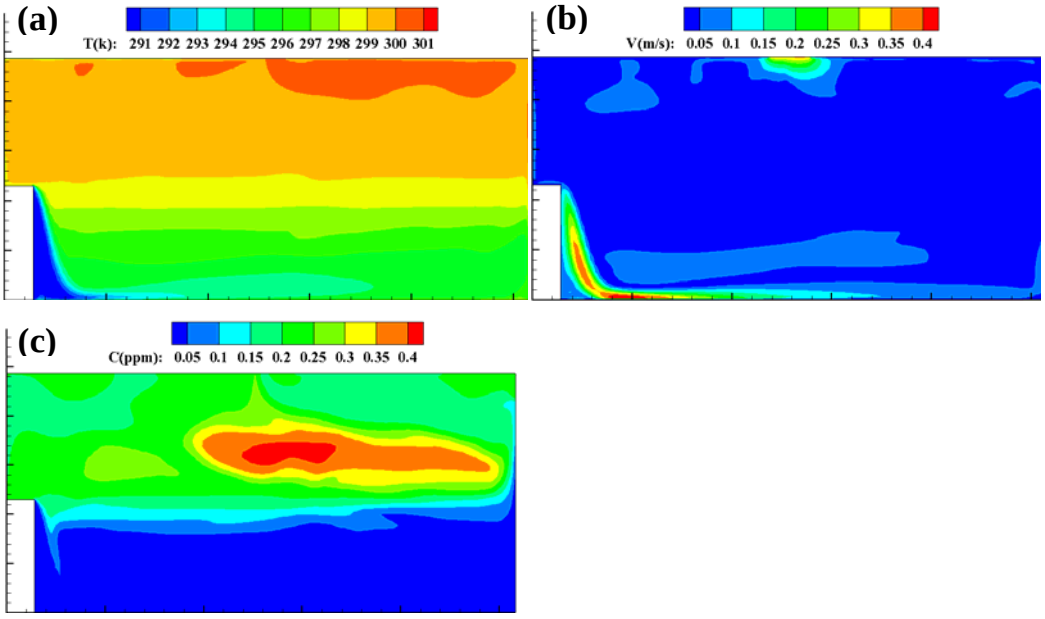


Fig. 4. The CFD-calculated (a) air temperature, (b) airflow, and (c) SF_6 concentration field.

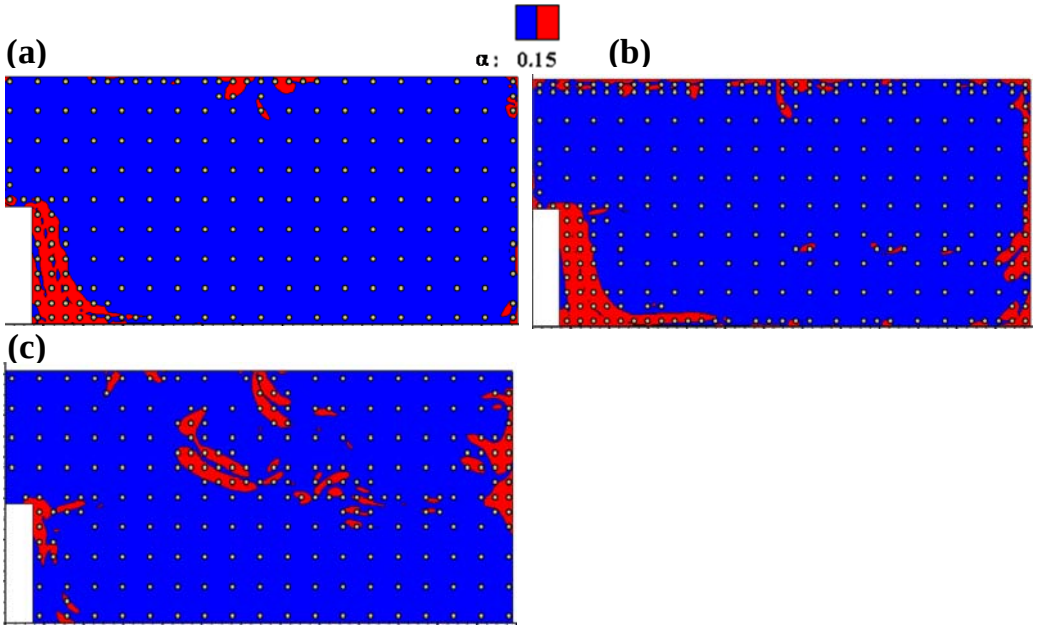


Fig. 5. The gradient coefficient (contours) and sampling point distribution (small circles) determined for the office room (red regions had $\alpha > 0.15$ and blue regions $\alpha \leqslant 0.15$) for (a) air temperature, (b) air velocity, and (c) SF_6 concentration.

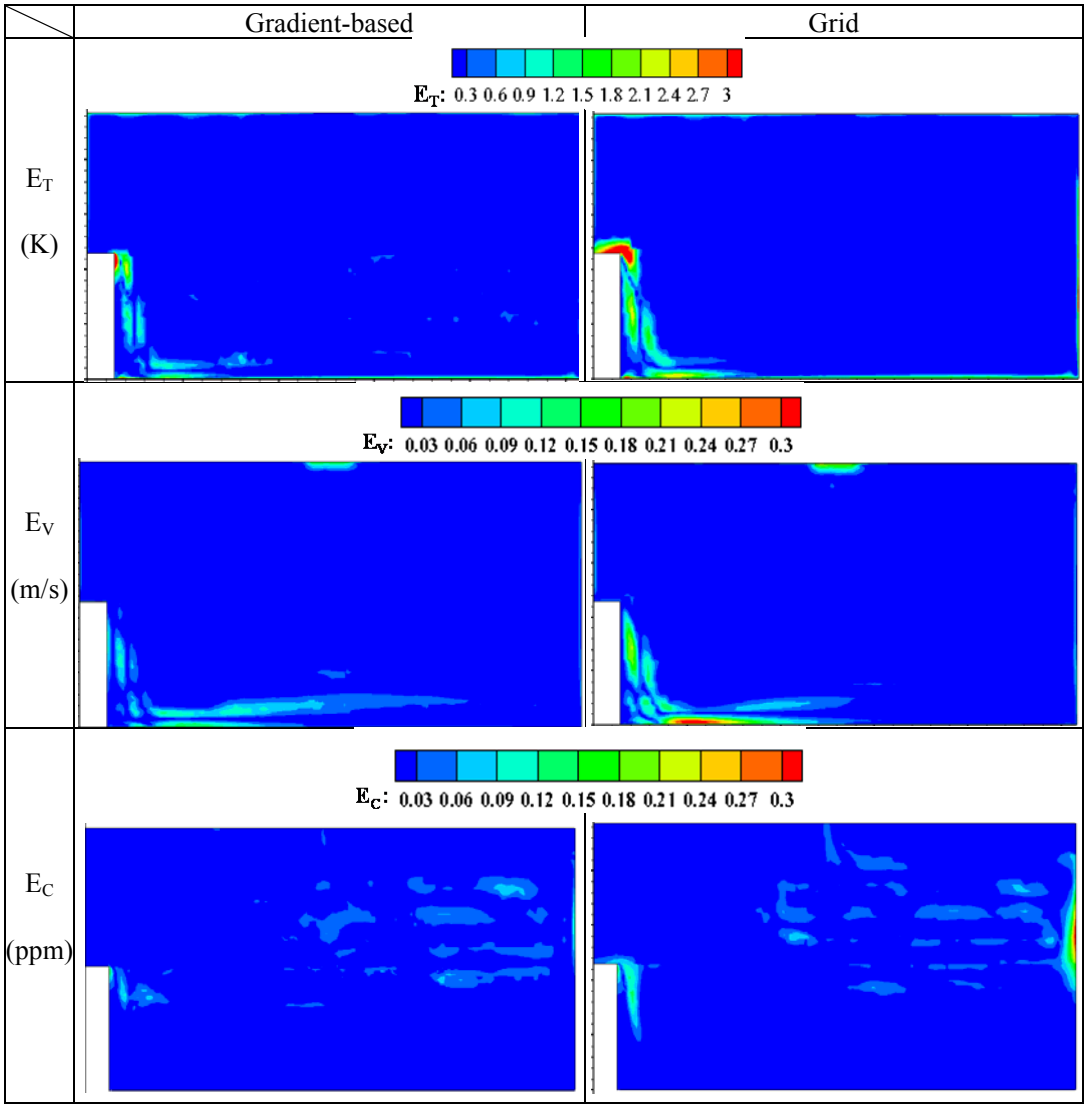


Fig. 6. The distributions of Kriging-interpolated errors in air temperature (E_T), air velocity (E_V), and SF_6 concentration (E_C) on the mid-y plane with the gradient-based method (left) and grid method (right).

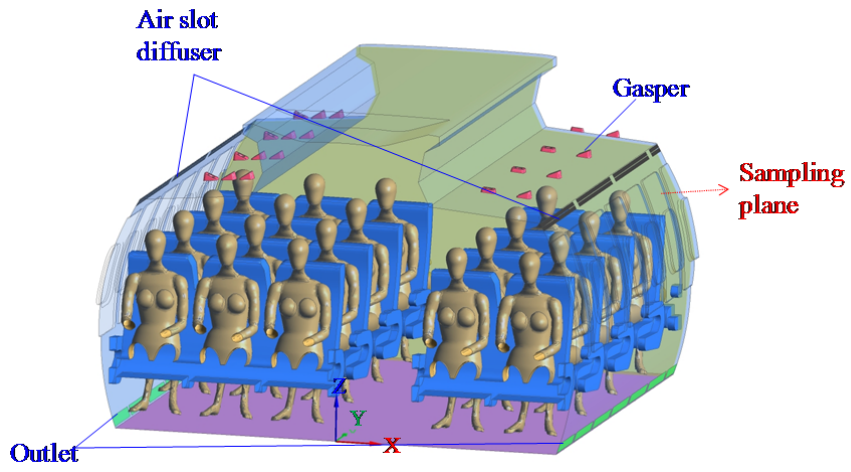


Fig. 7. Schematic of the four-row section of the economy-class airliner cabin and the cross-section (yellow color) in which the comparison was made in this study.

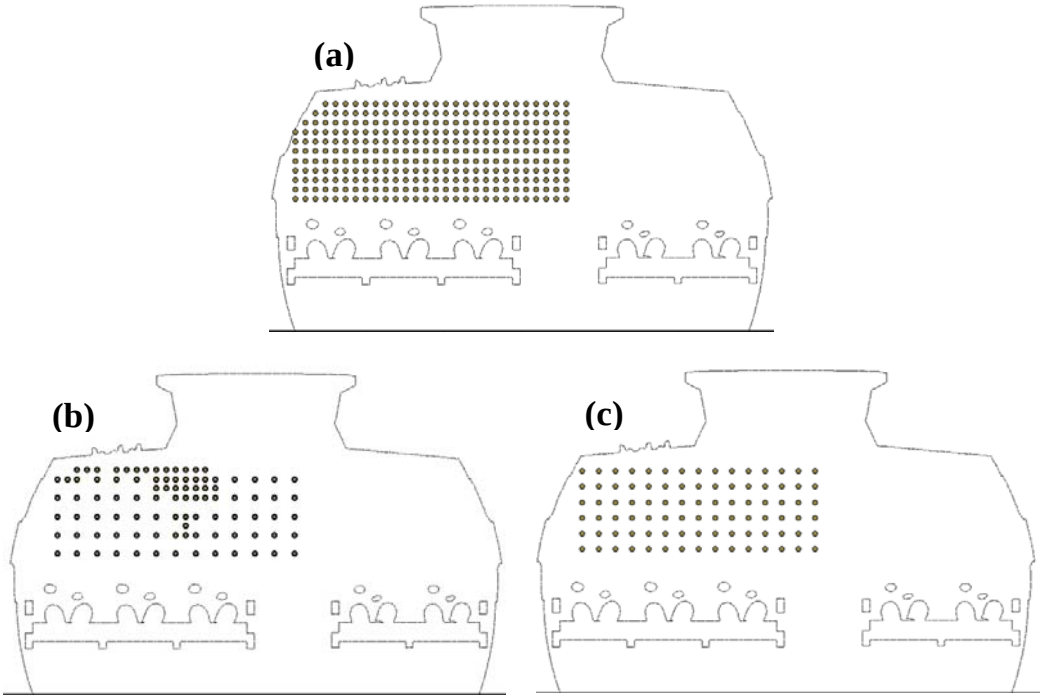


Fig. 8. Sampling point distribution for (a) the baseline method, (b) the gradient-based method, and (c) the grid method in the cross-section of the airliner cabin.

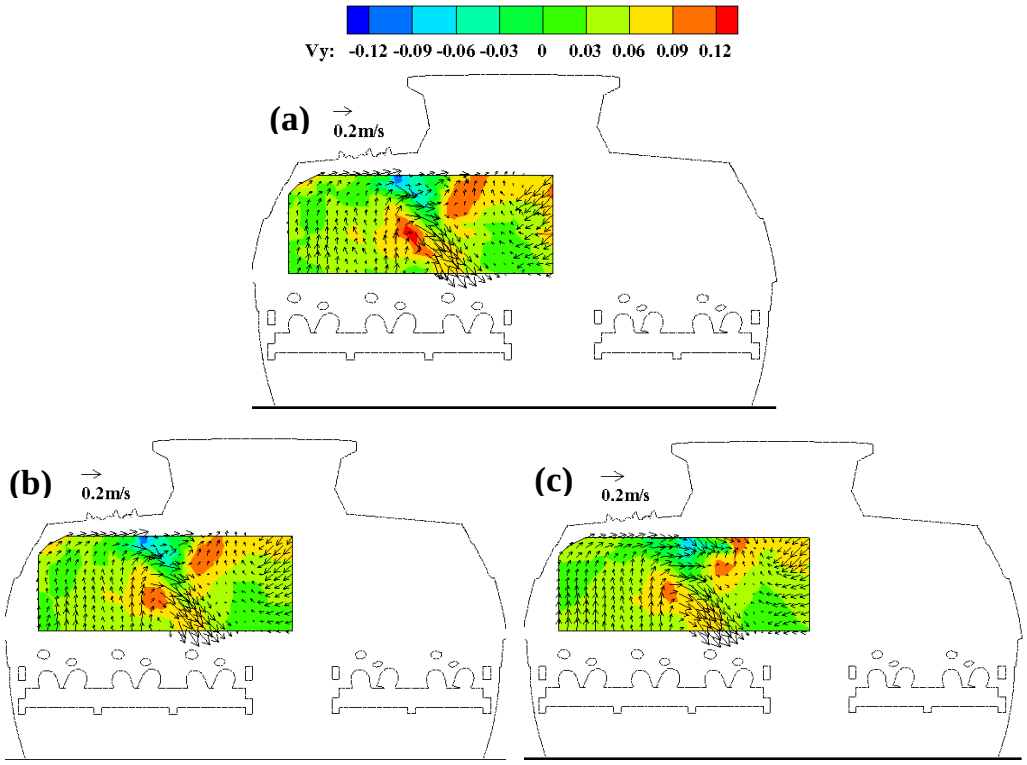
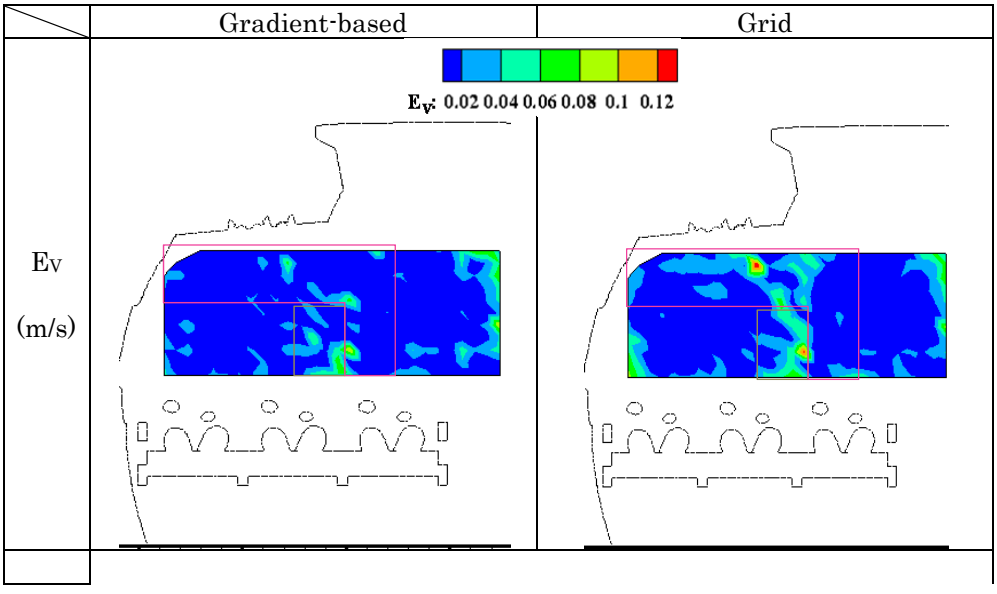
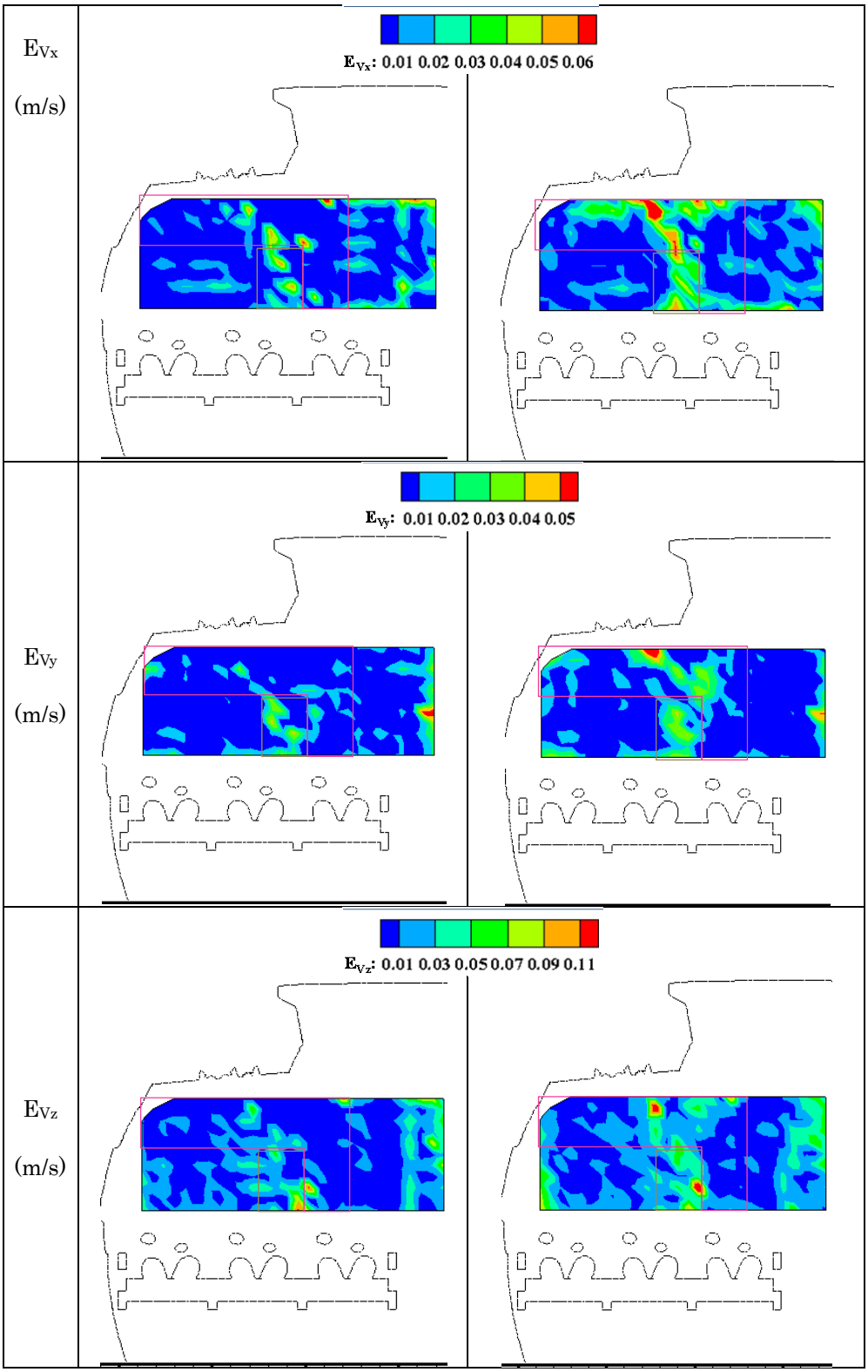


Fig. 9. Airflow pattern with (a) baseline, (b) gradient-based, and (c) grid sampling distributions in the cross-section of the airliner cabin based on the experimental measurements.





562 **Fig. 10.** The distributions of the Kriging-interpolated errors with the gradient-based
563 (left) and grid (right) sampling methods for the mean air velocity (E_V), x-velocity
564 component (E_{Vx}), y-velocity component (E_{Vy}), and z-velocity component (E_{Vz}) in the
565 cross-section of the cabin.

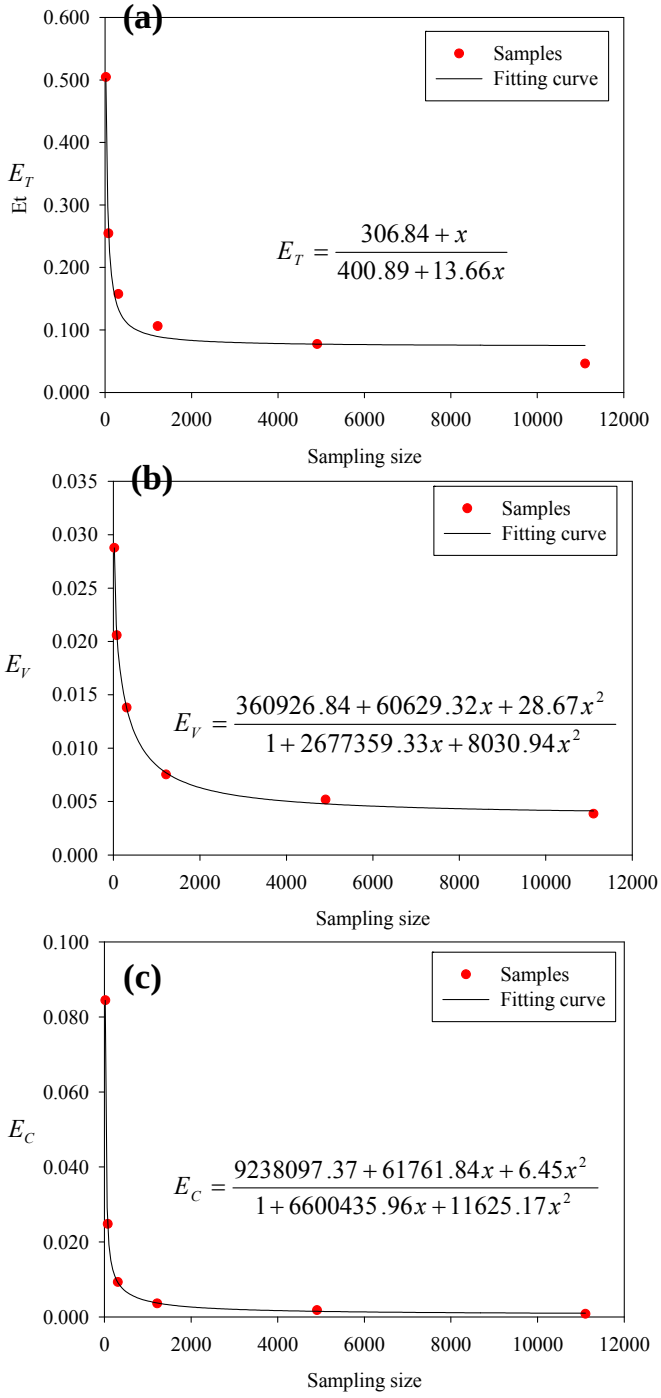


Fig. 11. The relationship between the average interpolation errors by the Kriging method and sampling size ranging from 18 to 11,104: (a) air temperature (E_T), (b) air velocity (E_V), and (c) SF_6 concentration (E_C).