1_C32

Investigating Uncertainty in the Relationship between Indoor Steady-State CO_2 Concentrations and Ventilation Rates

Constanza Molina, PhD **Benjamin Jones, EngD** **Andrew Persily, PhD**
Fellow ASHRAE

ABSTRACT HEADING

The steady-state concentration of occupant generated carbon dioxide (CO_2) is used in some applications as an indicator of compliance with a required ventilation rate. These applications assume that the CO_2 is at a uniform concentration in the space being considered, and that the outdoor concentration, ventilation rate, and CO_2 emission rate are all constant. Emission rates are often derived using an equation in the ASHRAE Fundamentals Handbook, which is poorly referenced and not based on the most recent understanding of the principles of human metabolism. A newer model was recently proposed that is based on these principles and considers the emission rate to be a function of age, sex, body mass and metabolic rate.

This paper presents an investigation into the uncertainty in CO_2 emission rates of US children of 4 to 19 years of age using this new model. The analysis is informed using government-published weight-for-age percentiles and sex ratios, and representative values of child metabolic rates in schools. However, since the metabolic rates are unavailable in probabilistic form, a distribution about an assumed mean value is applied. A Monte Carlo approach is used to generate cumulative distribution functions of CO_2 emission rates for infinite populations of children and for school classes of 20 students for a range of age groups. This approach shows that the individual emission rate increases with age and that there is significant variance about the mean value – the standard deviations are at least 20 % of the means. Mean emission rates for classes of 20 students also vary with age group, with a 95 % confidence interval of 4.1 to 4.3 cm^3 s^{-1} per person for 20 school age children. The uncertainty in a class mean emission rate decreases as the class size increases.

A global sensitivity analysis is used to determine that the most important model input is body mass and that metabolic rate is statistically significant. This means that standards and guidelines can use existing body mass data used by health services to determine appropriate values of mean emission rate to represent a local population.

1 INTRODUCTION

Indoor CO_2 concentrations have been used as a proxy indicator of *per capita* ventilation rates for many years, in many cases without an adequate understanding or explanation of the limitations for such applications (Persily 1997 and 2015). Among other limitations, these applications are based on assumptions that are not necessarily appreciated or confirmed by users, for example that the ventilation rate and CO_2 generation rate are constant and that the indoor

Constanza Molina is an Assistant Professor in the School of Civil Construction, Pontifical Catholic University of Chile, Santiago, Chile. **Benjamin Jones** is an Associate Professor in the Department of Architecture & Built Environment, University of Nottingham, Nottingham, UK. **Andrew Persily** is the Chief of the Energy and Environmental Division, National Institute of Standards and Technology, Gaithersburg, Maryland, USA.

concentration has reached steady-state. Guidelines, such as UK's BB101, have been published that prescribe indoor CO_2 concentrations as indicators of ventilation rate adequacy, despite these serious shortcomings in doing so (DfES 2018). In these and other applications of indoor CO_2 concentrations to characterize ventilation rates, the rate at which building occupants generate CO_2, $\dot{V}_{CO_2}$, is a key parameter. For many years, values of $\dot{V}_{CO_2}$ have been based on an equation in the ASHRAE Fundamentals Handbook (ASHRAE 2017). That equation is poorly referenced, and the discussion employs physical activity data (met levels) that are based on references predominantly from the 1960s or even earlier. More recently, a model of CO_2 generation rates has been developed that is based on principles of human metabolism and exercise physiology and explicitly accounts for age, sex, and body mass (Persily & de Jong 2017).

In addition to using these improved estimates of $\dot{V}_{CO_2}$, the use of indoor CO_2 to assess ventilation adequacy in guidelines and other applications needs to consider the uncertainty in these generation rates and how those uncertainties propagate to the reference values for CO_2 concentration. Such an analysis is needed to properly quantify the magnitude of the associated uncertainties and to identify those inputs that have the greatest impact on predictions of $\dot{V}_{CO_2}$ (i.e., a sensitivity analysis).

The application of indoor CO_2 concentrations to evaluate ventilation rates and the analyses presented in this paper use the following equation to relate steady-state concentrations to per capita ventilation rates.

$$Q_o = \dot{V}_{CO_2}\left(C_{i,ss} - C_o\right)^{-1} = \dot{V}_{CO_2} C_{e,ss}^{-1} \qquad (1)$$

Here, Q_o is the outdoor air ventilation rate *per capita* ($m^3\ s^{-1}$ per person), $\dot{V}_{CO_2}$ is the generation rate of CO_2 *per capita* ($cm^3\ s^{-1}$ per person), $C_{i,ss}$ is the steady-state indoor CO_2 concentration (ppm*), and C_o is the outdoor CO_2 concentration (ppm). The difference between $C_{i,ss}$ and C_o is known as the *excess* CO_2 steady-state concentration $C_{e,ss}$ (ppm), which is generally a more useful metric because C_o is highly variable in practice.

It should be noted that ASHRAE (2019) 62.1 (ventilation for acceptable indoor air quality), which is used to support the design of ventilation systems in US school classrooms, does not contain a CO_2 concentration limit. However, the UK's BB101 does contain such a limit* and implicitly assumes a value of $\dot{V}_{CO_2}$ = 5.5 $cm^3\ s^{-1}$ per person.

The aim of this paper is to investigate uncertainty in predicted $\dot{V}_{CO_2}$ of school age children in the US and associated uncertainty in the relationship between steady-state CO_2 concentration and *per capita* ventilation rates. To do this, we use Persily & De Jong's model of $\dot{V}_{CO_2}$ and a statistical framework to investigate the probability space. A standard Global Sensitivity Analysis identifies those inputs that influence the model's predictions the most and describes the nature of their relationships.

2 METHOD

A MATLAB (MathWorks, 2019) code was developed to exercise the model of $\dot{V}_{CO_2}$ developed by Persily & de Jong (2017) using a statistical framework that randomly selects inputs and tests the sensitivity of model predictions to those inputs. We used units that are generally used in practice so that the results can be directly related to existing standards and guidelines.

2.1 Model of CO_2 generation

This section summarizes Persily & de Jong's model of $\dot{V}_{CO_2}$, but a fully referenced description is given elsewhere

* ppm, which is equivalent to μL of CO_2 per L of air, is used in this paper for CO_2 concentrations.
* Maximum CO_2 concentration of 1000 ppm or 1500 ppm during the occupied period in spaces where mechanical ventilation or natural ventilation is used, respectively.

(Persily & de Jong, 2017). The generation rate of CO_2 *per capita* ($cm^3\ s^{-1}$ per person) is given by

$$\dot{V}_{CO_2} = \mathrm{BMR}\, M\, \mathrm{RQ}\, \dot{V}_{O_2}\, T\, P^{-1} \quad (2)$$

Here, BMR is the essential energy a person requires to sustain life, known as the *basal metabolic rate* ($J\ s^{-1}$ per person). It is assumed to be a linear function of body mass requiring sex and age dependent gradients and intercepts (see Table 1 in Persily & de Jong, 2017). M is a dimensionless metabolic rate that describes the ratio of a person's energy demand required to complete a specific physical activity relative to their BMR. RQ is the ratio of the volumetric rate at which CO_2 is produced to the rate at which oxygen is consumed, known as the *respiratory quotient*. For well-nourished people in normal weight range, its primary determinant is diet. Fractions of dietary carbohydrate, fat, protein, and alcohol by sex are derived from the US National Health and Nutrition Examination Surveys (Wright & Wang, 2010). They are assumed to be constant because uncertainties in these values are not given, and we also assume that children do not consume alcohol. $\dot{V}_{O_2}$ is a person's rate of oxygen consumption ($cm^3\ J^{-1}$) and implicitly assumes that 1 kcal (4.2 J) of energy use requires 206 cm^3 of oxygen consumption. Finally, T is the ratio of the air temperature to 273.15 K, and P is the ratio of the air pressure to 101.325 kPa. We assume indoor temperatures and pressures are 293.15 K and 101.325 Pa, respectively, so T and P are 1.07 and 1.00, respectively, throughout this analysis.

2.2 Sources of demographic and activity data

In the US, children typically start elementary school in their 5th year of life and conclude high school in their 19th year, so school aged children are assumed to be > 4 and < 19 years old. In these analyses, ages are uniformly sampled between these limits to model a school where all school age groups contain the same number of students. The school system varies across the country, but we assume that elementary school concludes, and high school starts, in a child's 14th year. Three age ranges are considered: the first is all children, the second is elementary school aged children, and the third is high school aged children. The analysis does not account for the presence of one or more adult teachers, but their influence will be considered by future work.

Body mass is derived from growth charts of the US Centers for Disease Control and Prevention for children and young adults aged 2 years to 20 years old (CDC, 2020). Data are given at 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, and 97th centiles for both sexes at 1-month intervals and are modelled using a 2D spline by MATLAB's `interp2` function to produce a continuous function.

Sex ratios are obtained from the US Census Bureau (2018) for age intervals of < 15 years and 15 to 19 years. We apply ratios by age group rather than using a population mean because there are more males in the under 20 age band than there are in the higher age bands. The ratios are obtained from numbers of each sex given to the nearest thousand.

Metabolic rates are determined from the National Collaborative on Childhood Obesity Research (NCCOR, 2017) youth compendium. We use activity 55400X: *Quiet Play/Schoolwork Television (sitting) Schoolwork*, which gives M = 1.6 for children under 10 years of age, M = 1.5 for ages 10 to 15, and M = 1.4 for over 15 years of age. We assumed M to be log-normally distributed with a standard deviation equal to 10 % of its mean. Because there is no evidence for either of these assumptions, we tested their significance using the sensitivity analysis described in Section 2.4.

2.3 Sampling method

The sampling method follows that described by Jones *et al.* (2015) and uses a Monte Carlo (MC) approach and Latin Hypercubes to interrogate the probability space more efficiently. The model requires input variates that are specified deterministically or are described by continuous probability distributions. They are applied to the $\dot{V}_{CO_2}$ model and, by systematically varying the variates and running multiple simulations, distributions of $\dot{V}_{CO_2}$ are generated that quantify the uncertainty in $\dot{V}_{CO_2}$. There are four probabilistic inputs: age, body mass, sex, and metabolic rate as described in Section 2.2.

Predictions of $\dot{V}_{CO_2}$ are obtained for a set of 1000 samples of 20 students and a mean $\dot{V}_{CO_2}$ is obtained for each set. After 100 sets (0.1 million samples), the means are tested for normality using a one-sample Kolmogorov-Smirnov test, and sampling is stopped when this test is found to be true (p-value < .01). The number of samples in each set is dependent on the cohort size. For a population sample, a set size of 1000 is chosen arbitrarily, although it has no effect on the predictions. Nevertheless, when considering a classroom of children in this analysis, a class size of 20 is informed by the National Teacher and Principal Survey (NCES, 2018). It is important to note that, due to the variance in the emission rates, the value of class size does affect the predictions.

2.4 Sensitivity Analysis

A global sensitivity analysis is used to test the dependence of $\dot{V}_{CO_2}$ on the four inputs. We follow the method of Jones *et al.* (2015) and Das *et al.* (2014), which tests for linear, monotonic, and non-monotonic relationships[*]. The input and output data are not transformed, and all outliers are retained. Coefficients and p-values are obtained for each test and the inputs are ranked according to the magnitude of the coefficient. The p-values are used to determine whether a result is statistically significant at a 5 % level.

A fundamental requirement of the sensitivity analysis is that there is no multicollinearity. Body mass and age are themselves correlated (the former is a function of the latter), so the sensitivity of the model cannot be assessed against both inputs. Sex is a categorical variable, although body mass is dependent on it, so only body mass and metabolic rate are tested.

3 RESULTS AND DISCUSSION

To consider uncertainty in individual $\dot{V}_{CO_2}$ for an infinite population of U.S. children, exactly $n = 10^5$ samples (100 sets) were required per age range to achieve normality in the means. The inputs for the 5th to 19th year of life range (all school years) are given in Figure 1 and show the uniform distribution of ages, the slight bias towards the male sex, a bimodal distribution of body mass (discernable by sex), and a log-normal distribution of metabolic rates.

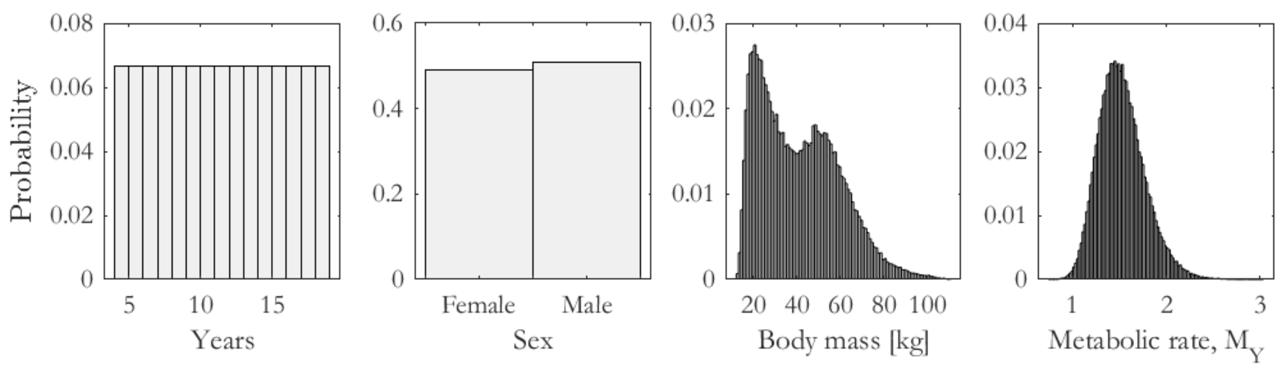

Figure 1 Histograms of model inputs for the 5th to 19th year of life range.

Cumulative distributions of predicted $\dot{V}_{CO_2}$ for all three age ranges are given in Figure 2, and descriptive statistics are given in Table 1. The distributions look log-normal, although this null hypothesis is rejected by both Kolmogorov-Smirnov and Lilliefors tests because the tails are too long. They show that $\dot{V}_{CO_2}$ increases with age, and there is significant variance about the mean value (μ), i.e., having standard deviations (σ) of at least 20 % of the means, also known as coefficient of variation, C_V. The limits of $\dot{V}_{CO_2}$ for children in their 5th to 19th year of life are $2.6 \leq \dot{V}_{CO_2} \leq 6.7$ cm³ s-

[*] The code is available under a creative commons license from DOI: 10.13140/RG.2.2.30311.39844.

^{1} per person with 95 % confidence. Therefore, in a classroom with a single occupant of unknown demographics where the specified *per capita* airflow rate is Q_o = 0.01 m^3 s^{-1} (or 10 l s^{-1}), the 95 % confidence interval in $C_{e,ss}$ is between 250 and 670 ppm. This shows that there is significant uncertainty in both $\dot{V}_{CO_2}$ and $C_{e,ss}$ for this unlikely scenario. Similarly, the 95 % confidence intervals for children in their 5th to 13th year of life (elementary school) range of $2.5 \leq \dot{V}_{CO_2} \leq 6.0$ cm^3 s^{-1} per person and $3.4 \leq \dot{V}_{CO_2} \leq 7.3$ cm^3 s^{-1} per person for the 14th to 19th year of life (high school) show significant uncertainty.

Table 1. Statistical summary of individual *per capita* CO_2 emission rates for different age ranges

Year of life	$\dot{V}_{CO_2}$ (cm^3 s^{-1} per person)				
	Mean, μ	**Standard deviation, σ**	**Median, P_{50}**	**95 % CI ($P_{2.5}$, $P_{97.5}$)**	**Coefficient of variation, C_V**
5th to 19th	4.2	1.1	4.1	(2.6, 6.7)	0.26
5th to 13th	3.9	0.9	3.8	(2.5, 6.0)	0.23
14th to 19th	5.0	1.0	4.9	(3.4, 7.3)	0.20

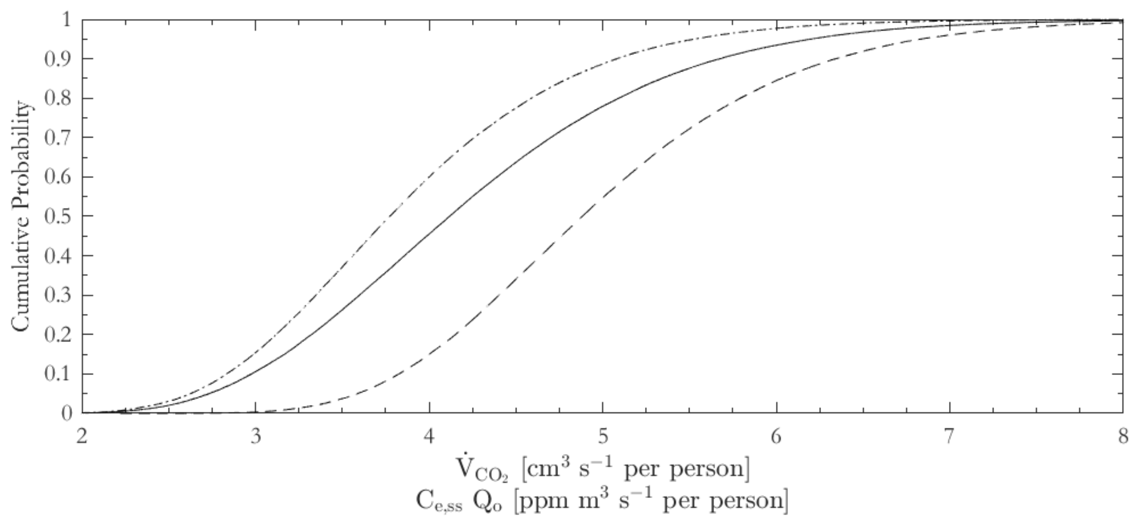

Figure 2 Cumulative distribution functions of individual $\dot{V}_{CO_2}$ and the product of the excess steady-state carbon dioxide concentration and outdoor airflow rate *per capita*. —— 5th to 19th year of life; -·-·-·- 5th to 13th year of life; - - - - - 14th to 19th year of life.

As presented in Section 1, $\dot{V}_{CO_2}$ = 5.5 cm^3 s^{-1} is used to determine a relationship between $C_{e,ss}$ and Q_o by some documents (DfES 2018). This value of $\dot{V}_{CO_2}$ is a mean for all occupants present in a zone of interest, which in the context of a school might be a classroom. The values given in Table 1 and Figure 2 should not be used to indicate the mean $\dot{V}_{CO_2}$ of a class of children because they are for individuals. Therefore, to explore the uncertainty in the mean $\dot{V}_{CO_2}$ for a classroom, the model is re-run considering a class of 20 students; see Section 2.3. 2×10^4 samples were required to form normal distributions of mean $\dot{V}_{CO_2}$ where the mean and median (P_{50}) are identical, and the standard deviation indicates the variance about the mean. Table 2 gives a statistical summary of mean $\dot{V}_{CO_2}$ for a class of n = 20 children ages 4 through 19 and aggregated for all school ages, for elementary school ages, and for high school ages. It shows that the mean $\dot{V}_{CO_2}$ (μ) and its standard deviation (σ) generally increase with age and that the variance is approximately 1 % of the mean. It also shows that a mean $\dot{V}_{CO_2}$ of 5.5 cm^3 s^{-1} is unlikely to be appropriate for school children of any age for the metabolic rates we applied (see Section 2.2). Instead, the mean values given in Table 2 could be considered and expanded for other activities that yield different metabolic rates. When the ages of the children are unknown, then the mean value of $\dot{V}_{CO_2}$ = 4.2 cm^3 s^{-1} per person for all school ages should be used to account for its uncertainty. Assuming a *per capita* airflow rate of Q_o = 0.01 m^3 s^{-1} (10 l s^{-1}) for a school classroom containing 20 school age children yields $C_{e,ss}$ = 4.2 / 0.01 = 420 ppm.

Table 2. Statistical summary of mean *per capita* CO_2 emission rates for a class of *n* = 20 children (cm^3 s^{-1} per person *or* ppm m^3 s^{-1} per person)

Year of life:	5th	6th	7th	8th	9th	10th	11th	12th	13th
μ, P_{50}	3.1	3.3	3.4	3.6	3.9	3.9	4.1	4.3	4.5
σ	0.021	0.024	0.021	0.026	0.030	0.028	0.029	0.029	0.033
Year of life:	**14th**	**15th**	**16th**	**17th**	**18th**	**19th**	**5th to 19th**	**5th to 13th**	**14th to 19th**
μ, P_{50}	4.8	5.0	4.9	5.0	5.1	4.8	4.2	3.9	5.0
σ	0.040	0.043	0.039	0.040	0.051	0.037	0.047	0.041	0.042

Table 2 can also be used to determine a 95 % confidence interval in the mean $\dot{V}_{CO_2}$ where the limits are $\mu \pm 2\sigma$. Therefore, for the same school classroom with the same specified *per capita* airflow rate, the 95 % confidence interval in mean $\dot{V}_{CO_2}$ is $4.1 \leq \dot{V}_{CO_2} \leq 4.3$ cm^3 s^{-1} per person, which corresponds to a 95 % confidence interval in $C_{e,ss}$ of $410 \leq C_{e,ss} \leq 430$ ppm. This shows that the difference between μ and $\pm 2\sigma$ of $C_{e,ss}$ is only 10 ppm, which is only about 2 %. Furthermore, the magnitude of σ decreases as *n* increases, and so the mean value of the mean $\dot{V}_{CO_2}$ can be used with confidence for most cases. An exception might be a classroom with only a small number of occupants, because Figure 2 and Table 1 show that the uncertainty in $\dot{V}_{CO_2}$ per person is high in children. The presence of one or more adult teachers may also significantly affect the value of mean $\dot{V}_{CO_2}$ in a classroom with relatively few students, which will be considered by future work.

This analysis only represents a preliminary assessment of the problem of using a single, mean CO_2 emission rate, particularly in guidance documents and standards, because it only accounts for parametric uncertainty and does not consider model uncertainty. The model performance can be evaluated by comparing this model's outcome with observed data. Other parameters included in Equation 2 could be important but are currently included only as heuristics. For example, RQ is considered constant but could be a function of body mass given that is calculated from fractions of dietary carbohydrate, fat, protein, and alcohol. Similarly, BMR, M, and mean $\dot{V}_{CO_2}$ should be investigated by future analyses, as well as transient effects of occupancy variations, changes in occupant activity levels over time, and the presence of adults. The final model could then be incorporated within a tool for use by practitioners, such as the NIST online CO_2 Metric Analysis Tool (NIST 2020).

The relative importance of the inputs can be determined using the sensitivity analysis outlined in Section 2.4. Figure 3 shows the model inputs plotted against its predictions. The age of the children is included for comparison even though it is itself correlated with body mass. Figure 3 shows that all the relationships are approximately linear. All statistical tests of the sensitivity analysis indicate that the model is most sensitive to body mass, and that both body mass and metabolic rate are statistically significant at 5% level. Therefore, the metabolic rate cannot be ignored. This is interesting because Section 2.2 identified that there is no knowledge of the uncertainty in child metabolic rates, so this data should ideally be obtained by future research. Fortunately, there are good data to describe uncertainty in the body mass of children because they are used by the health services of most countries to monitor the health of their populations. This means that standards and guidelines can use existing body mass data to determine appropriate values of mean $\dot{V}_{CO_2}$ to represent a local population and to discretize the data, especially where uncertainty in body mass is high, such as the 5th to 13th year of life range.

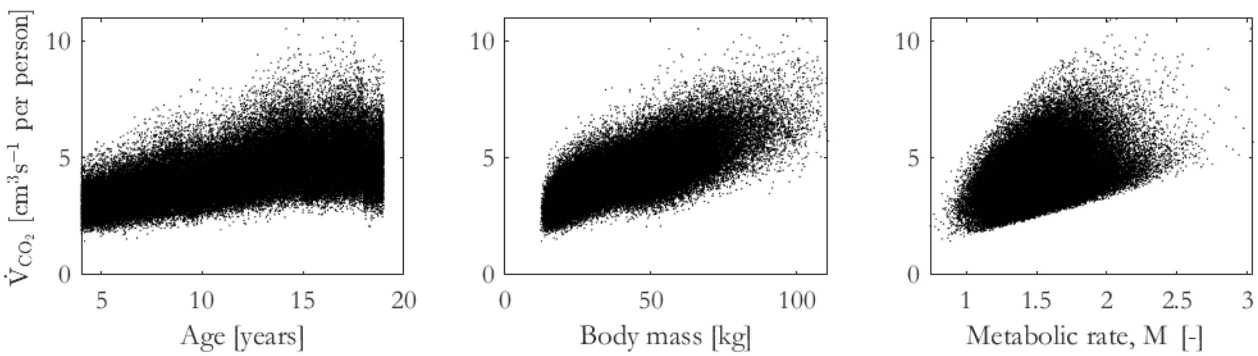

Figure 3 A visual sensitivity analysis by comparing model inputs to its output for the 5th to the 19th year of life.

4 CONCLUSIONS

There is significant uncertainty in predictions of $\dot{V}_{CO_2}$ for school age children. Its magnitude is heavily dependent on child body mass, and the rate of change and uncertainty in body mass is most significant in children between the ages of 4 and 14 years old. The uncertainty in a value of mean $\dot{V}_{CO_2}$ for a class of children decreases as the class size increases. However, there is clear variation in values of mean $\dot{V}_{CO_2}$ for classes of children with different ages, so it is important not to use a one-size-fits-all value of mean $\dot{V}_{CO_2}$, especially for children in the 4 to 14 age range.

Future work is required to corroborate the model and determine model uncertainty and validity ranges, to consider the transient effects of occupancy variations, changes in occupant activity levels over time, and the effects of adult teachers on mean $\dot{V}_{CO_2}$ for a range of class sizes.

DISCLAIMER

Certain commercial entities, equipment, or materials may be identified in this document to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

ACKNOWLEDGMENTS

The authors would like to acknowledge Oliver Brench whose research dissertation informed this work.

NOMENCLATURE

BMR = basal metabolic rate (J s^{-1} per person)
C = carbon dioxide concentration (ppm)
CO_2 = carbon dioxide
M = dimensionless metabolic rate
O_2 = oxygen
P = percentile
P = ratio of air pressure to that at 101.325 kPa
Q = airflow rate (m^3 s^{-1})
T = ratio of air temperature to that at 273.15 K
$\dot{V}$ = generation rate (cm^3 s^{-1})

μ = mean
σ = standard deviation

Subscripts

e = excess
i = indoor
o = outdoor
ss = steady-state

3 REFERENCES

ASHRAE. 2017. Fundamentals Handbook, Atlanta, GA, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE. 2019. Ventilation for Acceptable Indoor Air Quality. Atlanta GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, (ANSI/ASHRAE Standard 62.1-2016).

CDC. 2020. Centers for Disease Control and Prevention. National Centre for Health Statistics. Data Table of Weight-for-age Charts.

Das P *et al.* 2014. Using probabilistic sampling-based sensitivity analyses for indoor air quality modelling. *Building & Environment*. 78:171-82.

DfES. 2018. Guidelines on ventilation, thermal comfort and indoor air quality in schools, version 1. Building Bulletin 101. Manchester, UK: Education & Skills Funding Agency.

DfES. 2018. Building bulletin 101, 2018. A design guide: Ventilation of school buildings. Department for Education and Skills, London, UK.

Jones B *et al.* 2015 Assessing Uncertainty in Housing Stock Infiltration Rates and Associated Heat Loss: English and UK Case Studies. *Building and Environment*: 92:645-56.

MathWorks. 2019. MATLAB Version 9.6.0.1011450 (R2019a). The MathWorks Inc.

NCES. 2018. National Centre for Educational Statistics. National Teacher and Principal Survey. Available from: https://nces.ed.gov/surveys/ntps/tables/ntps1718_fltable06_t1s.asp. Accessed: 15th May 2020.

NCCOR. 2017. Youth Compendium of Physical Activity. Washington, DC: National Collaborative on Childhood Obesity Research. Available from: https://www.nccor.org/tools-youthcompendium/. Accessed: 22nd May 2020.

NIST. 2020. NIST Multizone modeling. CO_2 metric Analysis Tool. Available from: https://pages.nist.gov/CONTAM-apps/webapps/CO2Tool/#/ Accessed: 27th May 2020

Persily, A.K. 1997. Evaluating Building IAQ and Ventilation with Indoor Carbon Dioxide. *ASHRAE Transactions*: 103 (2), 193-204.

Persily, A.K. 2015. Indoor Carbon Dioxide Concentrations in Ventilation and Indoor Air Quality Standards, 36th AIVC Conference Effective Ventilation in High Performance Buildings, Madrid, Spain, Air Infiltration and Ventilation Centre, 810-819.

Persily A., de Jonge L. 2017. Carbon dioxide generation rates for building occupants. *Indoor Air*. 27:868-879

US Census Bureau. 2018. Age and Sex Composition in the United States. Table 1: Population by Age and Sex.

Wright JD, Wang C-Y. 2010. Trends in Intake of Energy and Macronutrients in Adults from 1999-2000 through 2007-2008. Hyattsville MD: Centers for Disease Control and Prevention, National Center for Health Statistics.