Forecasting of Cooling Coil Load using On-Site Forecasts of Outdoor Air Conditions versus Forecasts from Closest Airport

Mathieu Le Cam[1], Radu Zmeureanu[1], Ahmed Daoud[2], Andreas K. Athienitis[1]
[1] Center for Zero Energy Building Studies, Department of Building, Civil, and Environmental Engineering, Concordia University, 1515 St. Catherine W., Montreal, Quebec, Canada, H3G 1M8.
[2] Laboratoire des Technologies de l'Énergie, Institut de recherche d'Hydro-Québec, 600, Avenue de la Montagne, Shawinigan, Québec, Canada, G9N 7N5

Abstract

This paper presents the forecasting of the cooling coil load of an existing cooling system by using three different forecasts of outdoor air enthalpy. First, the air enthalpy is forecasted from measurements at the building location. Second, the air enthalpy is forecasted by Environment Canada at the closest airport. Third, the forecasts of air enthalpy at the building location are obtained from the forecasts at the airport. The results show that, for this particular case, the corrected forecasts of air enthalpy obtained from Environment Canada can be used with acceptable results, in the absence of a weather station on the site.

Introduction

Weather data such as outdoor air temperature and relative humidity are important inputs that are needed for the building energy performance modelling and forecasting. However, most buildings do not have a weather station on-site to provide such measurements. In the case of the design of new buildings the only weather data might come from measurements at the closest airport or synthetic data generated by climatic computer models.

The building simulation community usually uses "typical" or actual weather data files based on measurements in most cases at the closest airports, for the simulation of building energy performance. Investigations were carried by Crawley (1998); Huang (1998); Yang & Lu (2004) on the effect of using different types of typical weather data on the estimated thermal loads of buildings far from the airports. The variability of measured and forecasted weather data is significant and can impact the building energy performance; uncertainty and sensitivity analyses of a building energy model were performed in Goffart et al. (2016) to account for this variability. Some studies have also been performed on the use of weather data files for future climate to estimate the impact of climate change on the building energy consumption such as in Crawley (2007); Frank (2005); Robert & Kummert (2012); Wang et al. (2011).

Weather forecasts are useful for adaptive control of Heating Ventilation and Air Conditioning (HVAC) systems. Only a very few studies such as Oldewurtel et al. (2012) looked at the impact of the accuracy of weather forecasts on predictive control strategies.

The main novelty of this paper consists in the comparison of multi-step forecasts of the cooling coil load of an existing cooling system, over a time horizon of up to 12 hours, using three different forecasts of the outdoor air enthalpy:

(i) from the weather data measured at the building location on the university campus;
(ii) from the weather data forecasts by a climatic model by Environment Canada at the closest airport; and
(iii) from the forecasts of air enthalpy at the closest airport by Environment Canada, and corrected for the Loyola campus.

The term "forecasts" used in this paper indicates the estimates for future times, while the term "predictions" refers to the estimates at the current time.

The first section of this paper shows the comparison of the current weather data (air temperature and relative humidity) from the Building Automation System (BAS) trend data of the Loyola Campus of Concordia University in Montreal with the measurements at the closest airport. The forecasting method is presented in the second section. The three sets of forecasts of outdoor air enthalpy are presented in the third section. The forecasts of the cooling coil load of the existing HVAC system, which use those three different forecasts of outdoor air enthalpy, are compared with the measured cooling coil load, in the fourth section. Conclusions are presented in the last section.

Comparison of the weather data measured at Loyola campus and at the closest airport

Weather data at the Loyola campus are obtained from the BAS trend data that are measured by those sensors that control the HVAC operation, at a 15-min time-step. The quality of measurements of outdoor air temperature and relative humidity are first analysed, compared with other available sensors, and some corrections made.

The closest airport for this case study, the Pierre-Elliott Trudeau, is located in Dorval, about 8 km away from the building location on Loyola Campus and 15 km from the downtown of Montreal. The airport presents an unobstructed terrain while there is relatively dense low to midrise construction at Loyola campus. The outdoor air conditions, measured at one-hour time-step at the Trudeau airport, are available at the website of the Hydro-Quebec research centre (https://www.simeb.ca/).

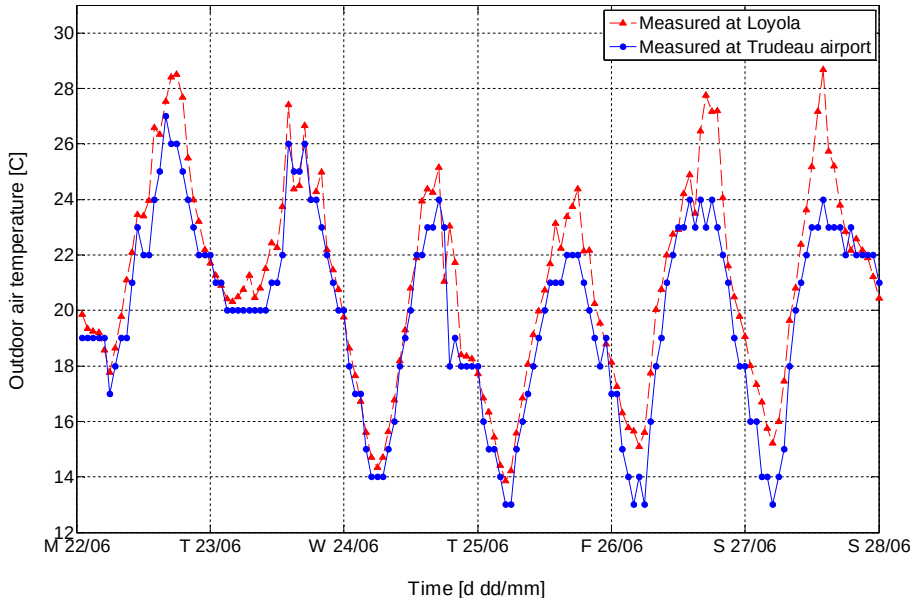

Figure 1: Measurements of dry-bulb outdoor air temperature on Loyola campus versus measurements at Trudeau airport over one week of June 2015.

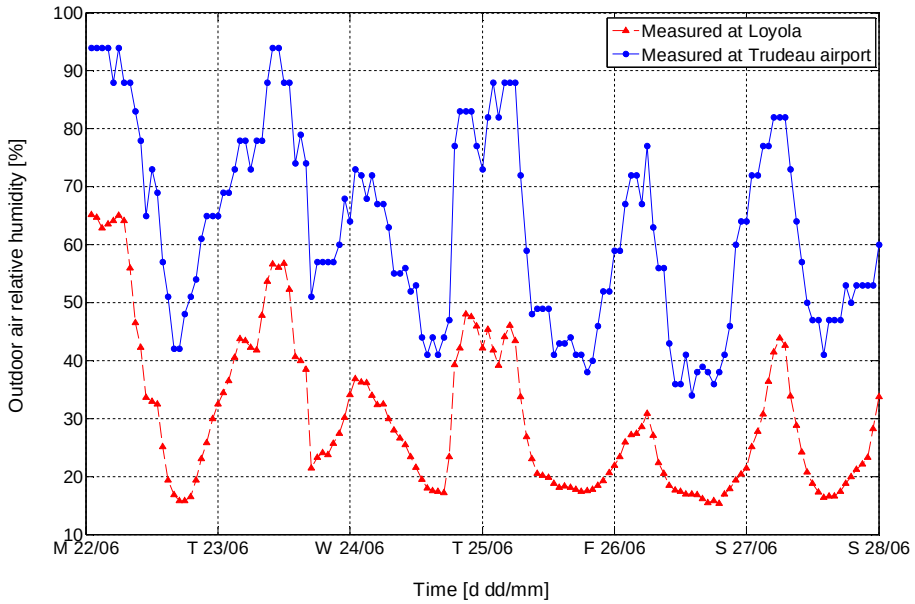

Figure 2: Measurements of outdoor air relative humidity on Loyola campus versus measurements at Trudeau airport over one week of June 2015.

Figure 1 compares the outdoor air temperature at Loyola campus over one week of June 2015 with the measurements at the Trudeau airport as an example; the difference has a Root Mean Square Error (RMSE) of 1.7°C.

A linear regression model (Equation (1)) estimates the dry-bulb outdoor air temperature at Loyola campus from the measurements at Trudeau airport from June to August 2015; with a coefficient of determination $R^2 = 0.91$ and a RMSE of 1.3°C. The model shows that the air temperature measured at Loyola campus is greater than the one measured at the airport with an intercept term of 2.0°C; this could be explained by the heat island effect due the concrete buildings and pavement around the location of outdoor air temperature sensor. A greater difference between the two measurements is noticed in the afternoon, from about 14:00 to 19:00, when the air temperature is greater than 24°C.

$$T_{Loyola} = 0.95 \cdot T_{airport} + 2.0 \qquad (1)$$

The comparison of the outdoor air relative humidity on Loyola campus and at the airport over the same week of June is presented in Figure 2. The regression model that estimates the relative humidity at Loyola campus from the measurements at Trudeau airport (Equation (2)) has $R^2 = 0.89$ and RMSE of 5.7%.

$$RH_{Loyola} = 1.09 \cdot RH_{airport} - 18.5 \qquad (2)$$

The outdoor air enthalpy is calculated at Loyola campus by using the measurements of outdoor dry-bulb air temperature and relative humidity, while at the Trudeau airport by using the measurements of dry-bulb and wet-but air temperatures. A linear regression model (Equation (3)) estimates the enthalpy of outdoor air at Loyola campus from the measurements at Trudeau airport from June to August 2015, with $R^2 = 0.82$ and RMSE of 3.5 kJ/kg.

The calculated outdoor air enthalpy at Loyola and at the Trudeau airport are compared in Figure 3.

$$h_{Loyola} = 0.73 \cdot h_{airport} + 9.8 \qquad (3)$$

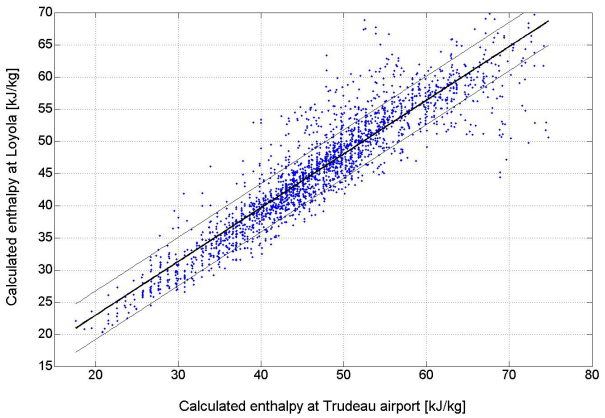

Figure 3: Calculated outdoor air enthalpy at Loyola versus calculated enthalpy at the Trudeau airport over summer 2015.

Forecasting method

The forecast of the target variable (i.e., the cooling coil load) is obtained with a data-driven model based on Support Vector Machines (SVM) introduced in Vapnik (1995).

The SVM method approximates an unknown function (f) by mapping, with a nonlinear function (ϕ), the observations of the training set ($\boldsymbol{x}$) into a higher dimension feature space. A linear regression problem is then solved in this feature space.

$$f(\boldsymbol{x}) = \langle \boldsymbol{w}, \phi(\boldsymbol{x}) \rangle + b \qquad (4)$$

where, $\boldsymbol{w}$ is the matrix of regression coefficients, b is the intercept term, $\boldsymbol{x}$ is the matrix of regressors and $\langle .,. \rangle$ is the dot product. An optimization problem was formulated by Vapnik (1995) to identify the function f (Equation (5)), which is expressed by the regression coefficients ($\boldsymbol{w}$) and intercept (b), that predicts all the observations of the target vector ($\boldsymbol{y}$) with a precision $\boldsymbol{\epsilon}$.

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|\boldsymbol{w}\|^2 + C.\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$
$$\text{Subject to: } \begin{aligned} &\langle \boldsymbol{w}, \phi(\boldsymbol{x_i}) \rangle + b - y_i \leq \epsilon + \xi_i, \\ &y_i - \langle \boldsymbol{w}, \phi(\boldsymbol{x_i}) \rangle - b \leq \epsilon + \xi_i^*, \\ &\xi_i, \xi_i^* \geq 0, i = 1, \dots, l. \end{aligned} \qquad (5)$$

where, y_i is an observation of the target vector, ξ is a slack variable, and $\boldsymbol{\epsilon}$ and C are parameters to be set by the user.

The regression function is reformulated in Equation (6) and the regression coefficients ($\boldsymbol{w}$) are given in equation (7).

$$f(\boldsymbol{x}) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\, k(x_i, x) + b \qquad (6)$$

$$\boldsymbol{w} = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\, x_i \qquad (7)$$

where, α_i and α_i^* are the Lagrange multipliers, and k is the kernel function.

The kernel function is selected to represent the distribution of input values on the training set, according to Cherkassky & Ma (2004). Radial basis functions are used in most studies as kernel functions:

$$k(x_i, x_j) = exp\left(-\gamma . \left\| x_i - x_j \right\|^2\right) \qquad (8)$$

where, γ is the "width" parameter that reflects the range of variation of all the regressors in the training set. More details are given in Le Cam (2016).

For each SVM model, one for each regressor and one for the target variable, three parameters (ϵ, C, γ) are identified. The forecasting SVM model is trained to estimate the target variable y with a precision ϵ. The positive constant C that is used in the objective function determines the trade-off between the complexity of the model and the amount up to which training errors greater than ϵ are tolerated. γ is the "width" parameter that reflects the range of variation of all the regressors in the training set. More details and explanations about the use of support vector machines for regression are given in Smola & Schölkopf (2004).

To identify the three parameters, γ and ϵ are varied from 2^{-12} to 2, and C from 2^{-5} to 2^5. For each combination of the three parameters, the SVM model is trained over ten days and tested over the following twelve hours. The training and testing are repeated over ten different samples of the previous month dataset. The ten samples have the same training and test lengths; however, they differ only by the random starting time and date. The root-mean-square error (RMSE) between the observations of the target variable and forecasted values is calculated over each test subset. Finally, the combination of the three parameters (ϵ, C, γ) that gives the minimum averaged RMSE over the ten testing subsets is selected.

Table 1: Best combination of parameters (ϵ, C, γ) for the SVM model of the cooling coil load with a twelve-hour prediction horizon.

Parameters	$\boldsymbol{\epsilon}$	$\boldsymbol{C}$	$\boldsymbol{\gamma}$
SVM model of cooling coil load	2^{-7}	2^{-1}	2^{-1}

The forecasting model uses time series of some regressors to forecast the future hourly values of the target variable over the following twelve hours. A filtering technique described in Le Cam (2016) is used to rank and select the regressors of the forecasting model, based on a cross-correlation analysis. The regressors are selected from a list of measured variables extracted from the BAS and derived variables as well. The cross-correlation coefficients of the cooling coil load with the independent variables are calculated. The variables whose cross-

correlation coefficients are greater than 0.7 are selected as regressors. The cross-correlation analysis also led to the selection of the time lag of the relevant past values of the regressors to the future value at the next time step. The following regressors are selected for the forecasting of cooling coil load: the target variable itself, the outdoor air enthalpy, the mixed air temperature in the air-handling units (AHUs), and the supply air flow rate of AHU. Since the supply air temperature and relative humidity leaving the AHU is almost constant, those variables are not selected as regressors.

The cooling coil load is calculated from measurements as follows (Equation (9)):

$$\dot{Q}_{CCL}[kW] = \dot{V}_{SA} \cdot \rho_A \cdot 10^{-3} \cdot (h_{MA} - h_{CD}) \qquad (9)$$

where, h_{CD} $[kJ/kg]$ and h_{MA} $[kJ/kg]$ are the cold deck and mixed air enthalpies, respectively calculated as averages of the two AHUs. $\dot{V}_{SA}$ $[L/s]$ is the volumetric air flow rate supplied by both AHUs of the Genome building; the density of the air ρ_A is assumed constant at 1.2 $[kg/m^3]$ at 20 $[°C]$ and 101.3 $[kPa]$.

The iteration strategy presented by Taieb et al. (2010) for multi-step-ahead forecasting is used in this study. For instance, when the current time is t, the model forecasts the target variable at time t+1 using the current and previous measurements of the target variable and of additional regressors at times t, t-1, t-n etc. From the same current time t, the forecast at time t+2 is performed using the previous measurements at times t, t-1, t-n+1 etc. and the first forecast of the target variable and regressors at time t+1, and so on. For each additional regressor a different forecasting model is developed.

Forecasts of outdoor air enthalpy on Loyola Campus

Use of forecasts from measurements at Loyola campus

A data-driven model, based on SVM method, is used to forecast the outdoor air enthalpy from the current and past measurements of dry-bulb air temperature and relative humidity extracted from BAS trend data of Loyola campus. The three parameters of the SVM model are given in Table 3.

Use of forecasts at Trudeau airport

The outdoor conditions forecasted by Environment Canada at the Trudeau airport are gathered from their website (http://dd.weatheroffice.ec.gc.ca/nowcasting/matrices/). Different types of forecasts are available. The forecasts used in this case study are obtained through the Integrated NowCasting System (INCS) presented in Landry et al. (2005), which contains numerical weather prediction models, statistical models and nowcasting models. The INCS model uses different types of observations: surface, radar and lightning data. These observations are used as inputs to different forecasting models. The radar and lightning observations are extrapolated with the McGill Algorithm for Precipitation nowcasting by Lagrangian Extrapolation (MAPLE). A numerical weather prediction model, the Regional Deterministic Prediction System (RDPS), is used to compensate for the missing measurements. A statistical model (PubTool) performs a probabilistic forecast of some weather elements over the next twelve hours. All those observations and forecasts are processed by the INCS system through a rule-based system to produce the matrix for each observation site. In this paper, only the forecasts of dry and wet bulb outdoor air temperatures are used, from which the outdoor air enthalpy is forecasted at the airport.

The scribe nowcasting matrices contain observed values of the current weather conditions and past six hours for over 450 weather stations in Canada. This system generates matrices of 19 hours of data for a selected number of Canadian observations sites. The data include seven hours of synthetized observations followed by twelve hours of forecasts for each observation site. The observations and forecasts of the matrices are updated hourly.

The forecasting results of the INCS model for the airport and of the SVM model from measurements at Loyola campus are compared over two examples. The first example is for May 25, 2016 starting at 00:00 (Figure 4).

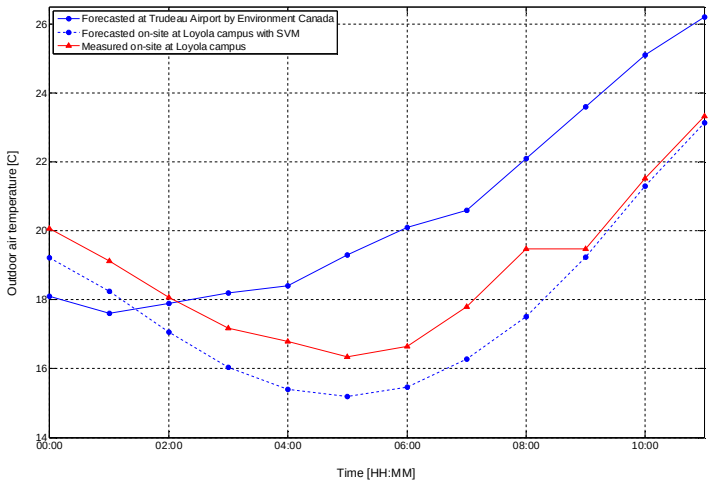

Figure 4: Comparison of forecasted and measured outdoor air temperature at Loyola campus with the forecasts at the Trudeau airport over 12 hours on May the 25th, starting at 00:00.

The second example is for the same day starting at 08:00 to analyse the performance of the model at the peak temperature of the day. The forecasted values at the airport are presented in blue circles and continuous line against the measurements at Loyola campus in red triangles. The forecasts of air temperature using on-site measurements corresponds to the dashed blue line. As expected, the forecasts of air temperature at Loyola using on-site measurements are closer to the measurements at Loyola campus on these two examples with RMSE = 1.1°C, compared to 2.6°C when the uncorrected forecasts at the Trudeau airport are compared with the measurements at Loyola; with RMSE of 1.4°C and 5.4°C, respectively (Table 2).

The forecasts of outdoor air enthalpy are also compared on the same examples of 25th of May 2016 starting at 00:00 (Figure 5) and at 08:00 (Figure 6), against the measured values at Loyola. As expected the forecasts based on Loyola measurements give better results.

Use of corrected forecasts at Trudeau airport for Loyola campus

A third forecast of the outdoor air enthalpy at Loyola campus is obtained using the forecasts of Environment Canada at the Trudeau airport that are corrected for the Loyola campus using the regression model (Equation (3)). The comparison between the corrected enthalpy and the SVM results at Loyola campus is presented in Figure 5 and Figure 6.

The results of the SVM model and the INCS model used by Environment Canada for the forecast of the outdoor air conditions is presented in Table 2. The forecasts of outdoor air enthalpy based on the on-site measurements show a better performance. A RMSE of 1.6 kJ/kg is calculated between the forecasted and measured values at Loyola; compared to 8.5 kJ/kg for the INCS model at Trudeau airport, for instance on the 25th of May 2016 starting at 00:00.

The forecast of outdoor air enthalpy provided at the Trudeau airport by Environment Canada is slightly improved when it is corrected with the linear regression model for Loyola campus. A RMSE of 7.6 kJ/kg is calculated between the forecasted and measured value at Loyola compared to 8.5 kJ/kg for the forecast at the Trudeau airport, on the 25th of May 2016 starting at 00:00 (Table 2).

Table 2: RMSE of the difference between the forecasts and measurements on May 25, 2016.

Variable	**Model**		
	SVM	**INCS uncorrected for Loyola**	**INCS corrected for Loyola**
Outdoor air temperature at Loyola campus [°C] *starting at 00:00 AM*	1.7	2.6	-
Outdoor air temperature at Loyola campus [°C] *starting at 08:00 AM*	1.4	5.4	-
Outdoor air enthalpy at Loyola campus [kJ/kg] *starting at 00:00 AM*	1.6	8.5	7.6
Outdoor air enthalpy at Loyola campus [kJ/kg] *starting at 08:00 AM*	3.1	8.5	6.1

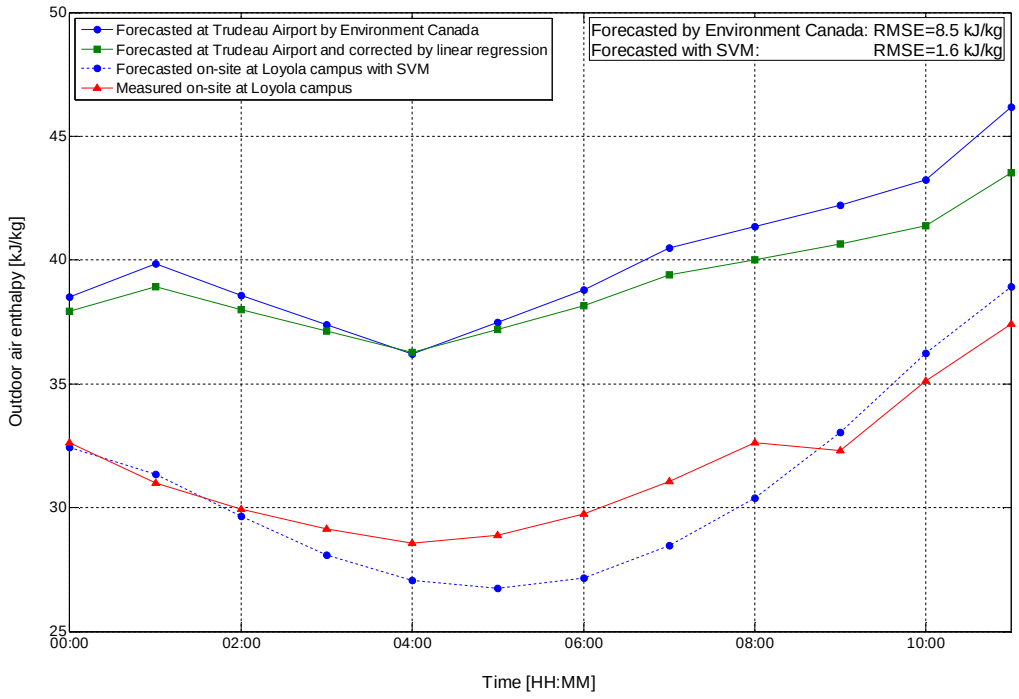

Figure 5: Comparison of forecasted and measured air enthalpy on site as well as forecasted at the Trudeau airport over 12 hours on May the 25th, starting at 00:00.

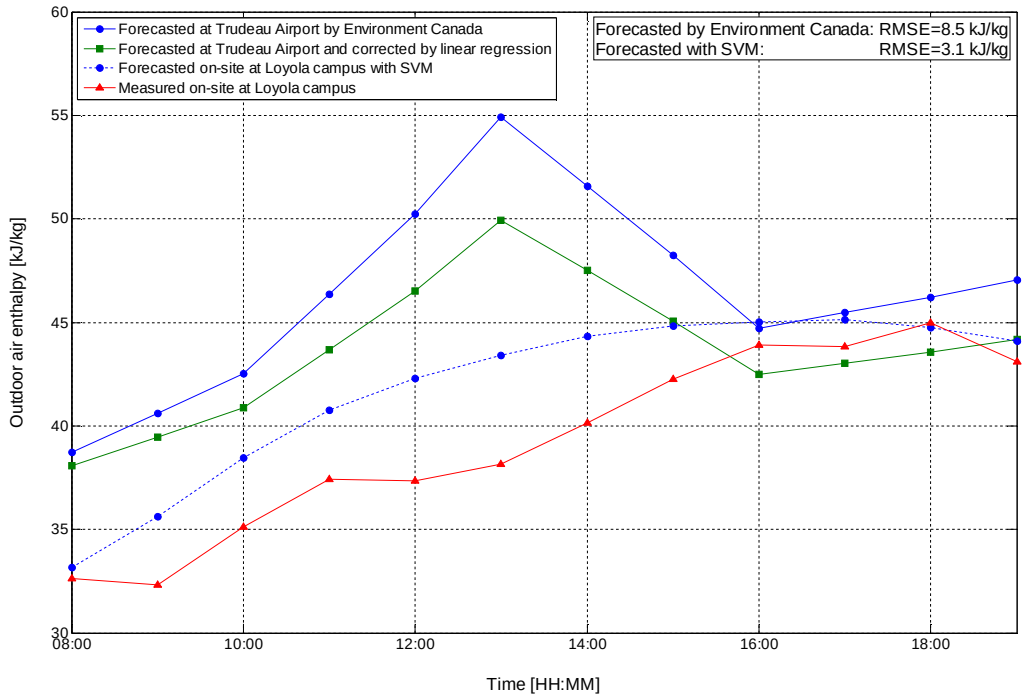

Figure 6: Comparison of forecasted and measured air enthalpy on site as well as forecasted at the Trudeau airport over 12 hours on May the 25th, starting at 08:00.

In conclusion, as expected the SVM-based model that uses history of on-site measurements at Loyola campus produces closer forecasts of outdoor air temperature and enthalpy at the building location than the INCS model, which was developed for the airport location. The INCS model includes the forecasts of several indices of the meteorological conditions for the Trudeau airport. The corrected forecast of outdoor air enthalpy is slightly improved by using the regression model (Equation (3)). The forecast of outdoor air enthalpy is expected to have an impact on the forecast of HVAC system thermal behaviour. This assessment is presented in the following section.

Impact on the forecast of the cooling coil load of the building

This section presents the impact of using the three different forecasts of the outdoor air enthalpy, discussed in the previous section, as regressors for the forecast of the cooling coil load of the Genomic research centre on the Loyola campus.

The paper focuses on the forecast of the total cooling coil load; the separation of the ventilation from the latent load is beyond the scope of this paper. However, in a future work the sensible and latent load could be defined as target variables to forecast.

The future value of the cooling coil load of the Genomic research centre is forecasted using SVM models and as regressors the current and past values of the cooling coil load, the outdoor air enthalpy, the mixed air temperature in both AHUs and the supply air flow rate. Three additional forecasting models are developed for each regressor; they are based on SVM models using only the current and past values of each regressor. The three parameters of the SVM models are presented in Table 3 for each regressor.

The measured profiles of three regressors in Figure 7 are presented in red triangles while the forecasted values are in blue circles. The performance of the forecasting models for each regressor is listed in Table 4, in terms of RMSE and coefficient of variation of the RMSE (CV(RMSE)).

Table 3: Best combination of parameters (ϵ, C, γ) for the SVM model of the regressors with a twelve-hour prediction horizon.

Parameters	ϵ	C	γ
SVM model of outdoor air enthalpy	2^{-4}	2^{-6}	2^{-2}
SVM model of mixed air temperature in AHU#1	2^{-9}	2^{-3}	2^{-2}
SVM model of mixed air temperature in AHU#2	2^{-9}	2^{-10}	2^{5}
SVM model of supply air flow rate	2^{-4}	2^{-6}	2^{5}

Table 4: Forecasting performance of the models for the regressors.

Variable	**RMSE**	**CV(RMSE) [%]**
Mixed air temperature in AHU#1 [°C]	1.0	4.3
Mixed air temperature in AHU#2 [°C]	1.0	4.3
Supply air flow rate [L/s]	3450	20.5

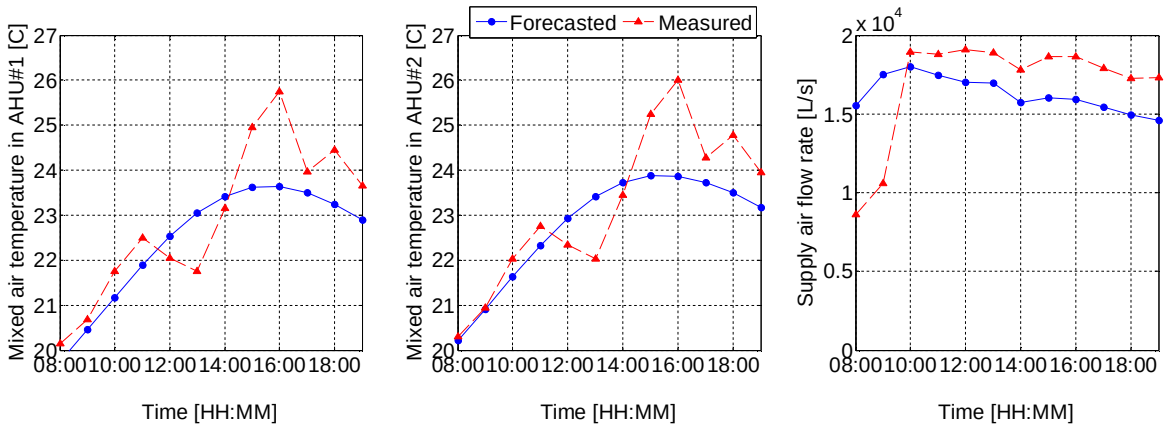

Figure 7: Forecast of the mixed air temperature and supply air flow rate used as regressors for the cooling coil load, on the 25th of May 2016.

The impact of the outdoor air enthalpy on the forecast of the cooling coil load is illustrated with an example on May 25, 2016 from 08:00 to 19:00.

The forecast of the cooling coil load shows a slightly better performance when using the outdoor air conditions forecasted from measurements on-site (Figure 8) with RMSE = 42 kW and CV(RMSE) = 23.6 %; compared to 56 kW and 31.1 % when using the uncorrected forecasts at the airport (Figure 9); and 53.9 kW and 29.4% when using the corrected forecast at the airport (Figure 10). The results of Figure 8 to Figure 10 are obtained with the SVM model parameters which are identified for 12 hours. When the focus is on the first five hours of the forecasts, the forecasts of the cooling coil load are slightly better when using the outdoor enthalpy forecasted at Trudeau airport.

The measured profile of the cooling coil load is presented by the red triangles on Figure 8 to Figure 10 with an envelope profile in red dash lines. It corresponds to the error propagated from the uncertainty in the measurements of the supply air flow rate, the temperature at the mixing and supply air conditions, and the relative humidity at the supply, return and outdoor conditions (Equation (10)).

$$U_{\dot{Q}_{CCL}} = \sqrt{\sum_i \left(\frac{\partial \dot{Q}_{CCL}}{\partial x_i} * U_{x_i}\right)^2} \qquad (10)$$

Where x_i is the i[th] measured variable used to calculate the cooling coil load ($\dot{Q}_{CCL}$) and U_{x_i} is the uncertainty in the measurements of the variable x_i.

The forecasted profiles of the cooling coil load are also given with an error bar on Figure 8 to Figure 10. This forecasting error is estimated by a cross-validation: 100 subsets of data are selected on the period from the 1st to 15th of May 2016. Each subset includes a seven-day training period followed by a twelve-hour test set. The starting date and time of the subsets are randomly selected. The probability distribution function of the forecasting error at each time-step of the twelve-hour forecasting horizon for the hundred subsets is presented on Figure 11 when using as regressor the forecast of outdoor air enthalpy at the building location. A Laplace distribution function (Equation (11)) is fitted to the forecasting errors with a location parameter (μ) of -2.3 and a scale parameter (b) of 48.2 (estimated for $x = \mu$). Lin & Weng (2004) showed that a Laplace distribution presents a better fit to the forecasting errors than a Gaussian distribution. For this case, the cooling coil load is forecasted at each time-step with an error of $\pm 108.6\ kW$ around the forecasted value with a confidence level of 95%.

$$f(x|\mu, b) = 1/2b \cdot \exp(-|x - \mu|/b) \qquad (11)$$

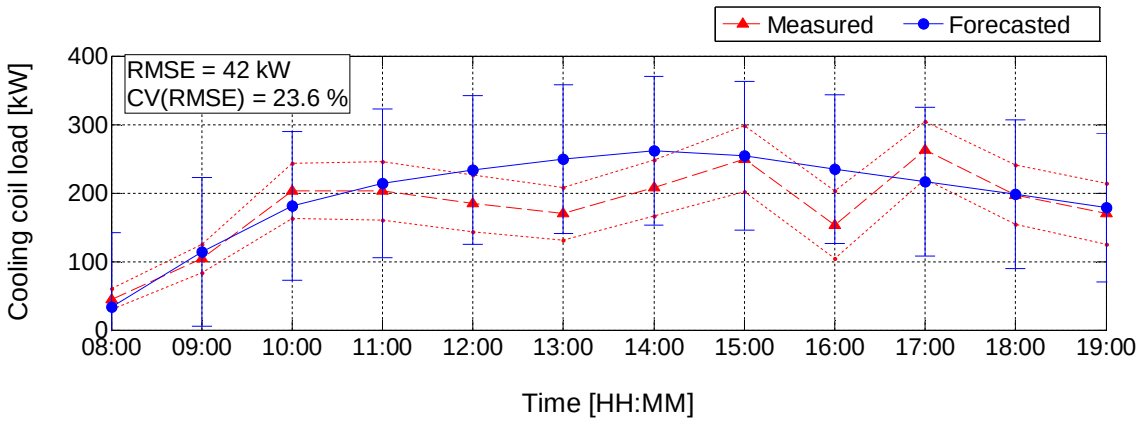

Figure 8: Forecast of the cooling coil load using the outdoor enthalpy forecasted by SVM model at Loyola campus on May 25, 2016 over a 12-hour forecasting horizon.

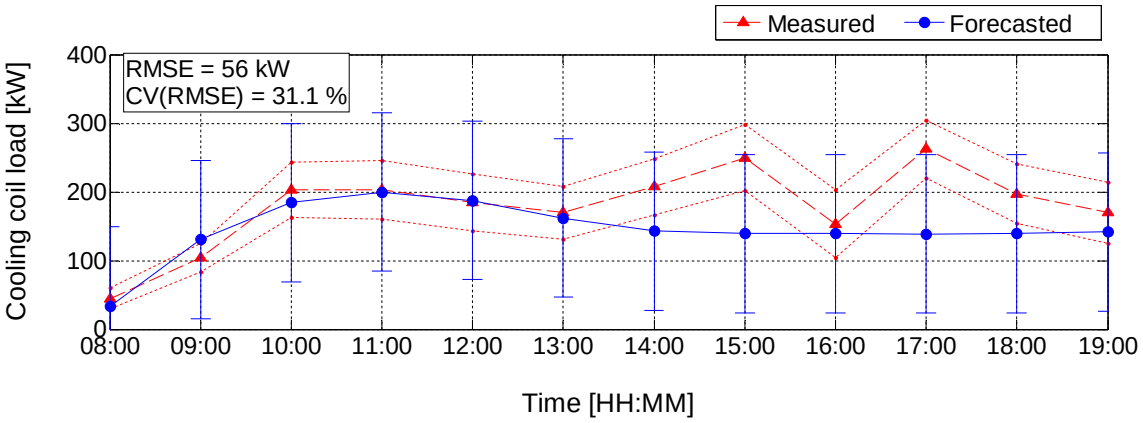

Figure 9: Forecast of the cooling coil load using the outdoor enthalpy forecasted by INCS at Trudeau airport on May 25, 2016 over a 12-hour forecasting horizon.

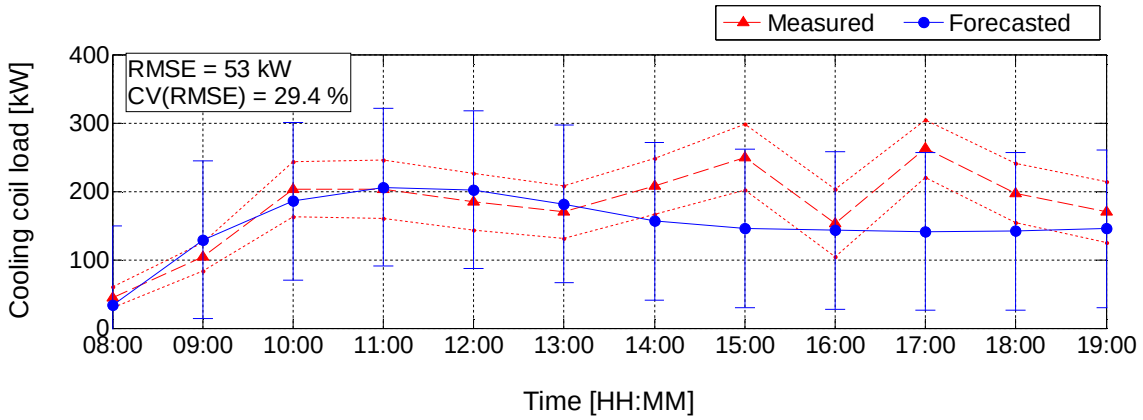

Figure 10: Forecast of the cooling coil load using the outdoor enthalpy forecasted by INCS at Trudeau airport and corrected for the Loyola campus, on May 25, 2016 over a 12-hour forecasting horizon.

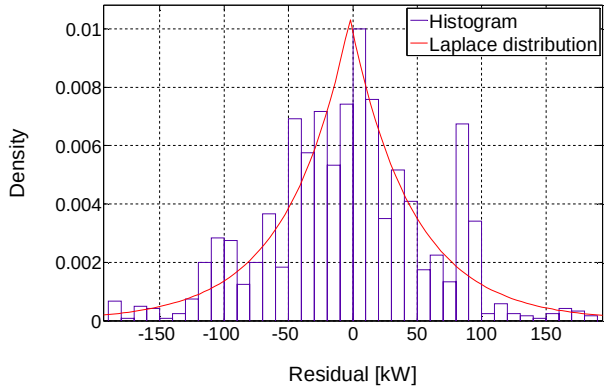

Figure 11: Validation error of the cooling coil load over 100 cross-validations from May 1st to May 15th 2016 using the outdoor enthalpy forecasted by SVM model at Loyola campus with a 12-hour forecasting horizon.

The forecasting horizon is reduced, for instance, to five hours, the parameters of the forecasting models are identified and presented in Table 5.

Table 6 presents the forecasting performance of the three models (RMSE and CV(RMSE)) over a five-hour forecasting horizon on the same day, May 25 from 08:00 to 12:00. The three models show similar performances on this example. The corrected and uncorrected versions of INCS model show identical results in this case because the outdoor enthalpy at the Trudeau airport and corrected for Loyola are very similar on the training set.

The forecasting approach presented in Le Cam (2016) was originally developed for a short-term of six-hour prediction horizon. The prediction horizon can be changed; it requires to identify the optimal values of the parameters of SVM model (ϵ, C, γ) for the desired time span. Other forecasting method might be more accurate for a prediction horizon longer than twelve hours.

In the absence of weather station at or near the building of analysis, the corrected forecasts of outdoor air enthalpy obtained from the forecasting climatic model from Environment Canada can be used with acceptable good results for the forecast of the cooling coil load.

This forecasting approach could handle other scenarios for non-daytime application such as residential houses. The filtering method used in this study selects the relevant regressors to the target variable to forecast for a non-daytime scenario as well. The forecasting approach could be improved by integrating additional regressors accounting for solar and internal loads, if measurements of solar radiation and occupation are available. New forecasting models would be developed for additional regressors by following the steps presented in the paper.

Table 5: Best combination of parameters (ϵ, C, γ) for the SVM model of the cooling coil load with a five-hour prediction horizon.

Parameters	ϵ	C	γ
SVM model of cooling coil load	2^{-2}	2^{-3}	2^{2}
SVM model of outdoor air enthalpy	2^{-4}	2^{-7}	2^{2}
SVM model of mixed air temperature in AHU#1	2^{-2}	2^{-9}	2^{5}
SVM model of mixed air temperature in AHU#2	2^{-5}	2^{-10}	2^{5}
SVM model of supply air flow rate	2^{-3}	2^{-3}	2^{-4}

Table 6: Forecasting performance of the models for the cooling coil load on May 25, 2016 from 08:00 to 12:00 over a five-hour forecasting horizon.

Model	**SVM**		**INCS uncorrected for Loyola**		**INCS corrected for Loyola**	
	RMSE [kW]	**CV(RMSE) [%]**	**RMSE [kW]**	**CV(RMSE) [%]**	**RMSE [kW]**	**CV(RMSE) [%]**
Cooling coil load	43	28.9	45	30.5	45	30.5

A sensitivity analysis is performed to estimate the impact of the outdoor air enthalpy on the forecast of the cooling coil load compared to the three other regressors: the mixed air temperature in each AHU and the supply air flow rate.

The method presented in Goffart et al. (2016) is applied to generate coherent stochastic profiles of the regressors based on statistics features extracted from observations. From the measurements of the month of July 2016, for each regressor, the hourly-averaged daily profile is calculated from the measurements. The daily profiles of residuals are calculated by subtracting the hourly-averaged from the daily profiles of July 2016. From the matrix of residuals, the cumulative distribution function and the autocorrelation of each regressor are extracted. Coherent stochastic profiles of the regressors are generated using equation (12), where $\bar{x}$ is the calculated hourly-averaged daily profile and ϵ is a random variable that presents the same cumulative distribution and autocorrelation as the measurements.

$$x = \bar{x} + \epsilon \tag{12}$$

In the example of Figure 12, the statistics features extracted from the observations of outdoor air enthalpy on Loyola campus over July 2016 are presented. Figure 12 a) corresponds to the hourly-averaged daily profile; b) is the autocorrelation; and c) presents the cumulative distribution function for each hour of the day.

A hundred stochastic profiles are generated with this approach. The first order sensitivity indices of the forecasting model of the cooling coil load of the building are calculated with the method presented in Goffart et al. (2016). It is a variance-based measure of the sensitivity as stated in equation (13), where y is the forecasted cooling coil load , x_i is the i[th] regressor, $X_{\sim i}$ is the matrix of regressors but x_i, $var(.)$ is the variance of the argument (.), $var_{x_i}(.)$ is the variance of the argument (.) taken over x_i, and $E_{X_{\sim i}}(.)$ is the mean of the argument (.) taken over $X_{\sim i}$, more details in Saltelli et al. (2010).

$$S_i = var_{x_i}\left(E_{X_{\sim i}}(y|x_i)\right) / var(y) \tag{13}$$

According to the sensitivity analysis, the SVM model forecasting the cooling coil load of the building presents the most sensitivity to the outdoor air enthalpy with a main effect index of 9.3% compared to 1.1%, 2.5% and 1.4%, respectively for the mixed air temperature in AHU#1, in AHU#2 and for the supply air flow rate.

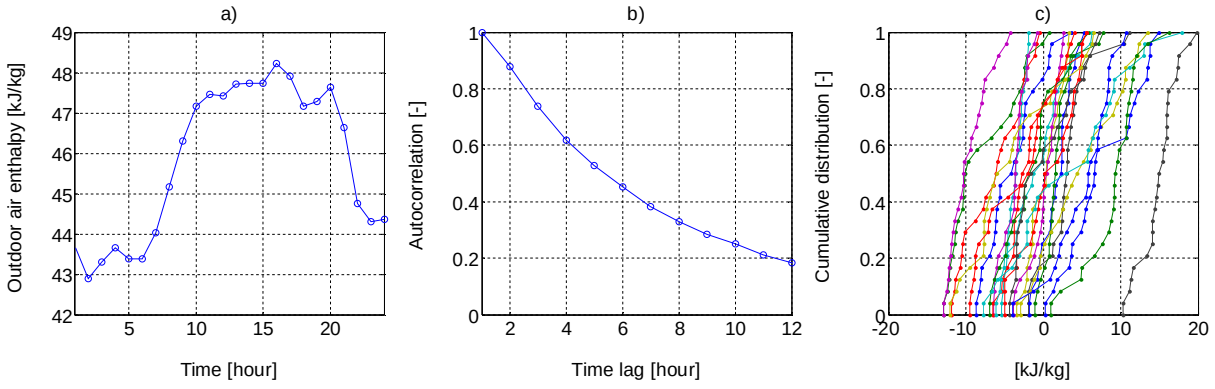

Figure 12: Extraction of statistics from measurements of outdoor air enthalpy over July 2016.

Conclusion

This study compared the forecasts of cooling coil load of an existing HVAC system of a building located on the Loyola campus of Concordia University in Montreal. The results show that for this case study, if the forecast is carried out on a time horizon of 12 hours, the data-driven SVM model that used the measurements on the campus, gave slightly better forecasts. However, if the forecasting horizon is five hours, the corrected forecasts of outdoor air enthalpy obtained from the forecasting climatic model from Environment Canada can be used with acceptable results for the forecast of the cooling coil load rather than installing a weather station on-site. In the context of demand response, more investigation is required on the error band that was estimated over the 12-hour forecasting horizon. Further work would look at the evolution of the error band at each time-step of the forecasting horizon.

Acknowledgement

The authors acknowledge the financial support from NSERC Smart Net-Zero Energy Building Strategic Research Network, NSERC/Hydro Quebec Industrial Research Chair, and the Faculty of Engineering and Computer Science of Concordia University.

Nomenclature

T_X	Outdoor air temperature at the location X (°C).
RH_X	Outdoor air relative humidity at the location X (%).
h_X	Outdoor air enthalpy at the location X (kJ/kg).
$\dot{Q}_{CCL}$	Cooling coil load on the air side of both AHUs (kW).
$\dot{V}_{SA}$	Supply air flow rate of both AHUs (L/s).
h_{MA}	Mixed air enthalpy (kJ/kg).
h_{CD}	Cold deck enthalpy (kJ/kg).
ρ_A	Density of the air (kg/m^3).
$U_{\dot{Q}_{CCL}}$	Propagated uncertainty in the derived cooling coil load (kW).
U_{x_i}	Uncertainty in the measurements of variable x_i.
S_i	First-order sensitivity index due to variable i (%).

References

Crawley, D. B. (1998). Which weather data should you use for energy simulations of commercial buildings?/Discussion. *ASHRAE Transactions, 104*, 498.

Crawley, D. B. (2007). *Creating weather files for climate change and urbanization impacts analysis.* Paper presented at the Proceedings of Building Simulation 2007, Beijing, China.

Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural networks, 17*(1), 113-126.

Frank, T. (2005). Climate change impacts on building heating and cooling energy demand in Switzerland. *Energy and buildings, 37*(11), 1175-1185.

Goffart, J., Mara, T., & Wurtz, E. (2016). Generation of stochastic weather data for uncertainty and sensitivity analysis of a low-energy building. *Journal of Building Physics*, 1744259116668598.

Huang, J. (1998). The Impact of Different Weather Data on Simulated Residential Heating and Cooling Loads. *ASHRAE Transactions, 104*, 516-527.

Landry, C., Parent, R., Deschenes, J.-F., Giguère, A., Hardy, G., & Verret, R. (2005). *Scribe Nowcasting - An Integrated Nowcasting Sub-System.* Paper presented at the Symposium on nowcasting and very short range forecasting, Toulouse, France.

Le Cam, M. (2016). *Short-term forecasting of the electric demand of HVAC systems.* Ph.D. thesis, Concordia University, Montreal, Canada.

Lin, C. J., & Weng, R. C. (2004). Simple probabilistic predictions for support vector regression. *National Taiwan University, Taipei*.

Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., et al. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings, 45*, 15-27.

Robert, A., & Kummert, M. (2012). Designing net-zero energy buildings for the future climate, not for the past. *Building and Environment, 55*, 150-158.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications, 181*(2), 259-270.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing, 14*(3), 199-222.

Taieb, S. B., Sorjamaa, A., & Bontempi, G. (2010). Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomputing, 73*(10), 1950-1957.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York: Springer Science & Business Media.

Wang, X., Chen, D., & Ren, Z. (2011). Global warming and its implication to emission reduction strategies for residential buildings. *Building and Environment, 46*(4), 871-883.

Yang, H., & Lu, L. (2004). Study of Typical Meteorological Years and Their Effect on Building Energy and Renewable Energy Simulations. *ASHRAE Transactions, 110*(2).