**Contents**

1. Problem Overview………………………………………………………………….……………………………………. 2
2. Methodology………………………………………………………………………………………………………………. 3
3. ASHRAE Report model (Untrained)….…………………………………………………………………………. 4
4. ASHRAE Report model (Trained)….……………………………………………………………………………… 5
5. Partial Least Square Regression (PLS) ….……………………………………………………………………….6
6. Support Vector Regression (SVR)….……………………………………………….…………………………….7
7. Overall Results…………………..………………………………………………………………………………………..8
8. New Method Based on SVR………………………………………………………………………………………9-10
9. Conclusion………………………………………………..………………………………………………………………..11

**Problem Overview**

Climate changes should be incorporated while designing buildings now and, in the future, to make them durable and resilient over a long period. Up to now, building envelopes are designed to withstand historical climate loads which are assumed to be static. But there is enough evidence that climate has been warming globally, thereby causing more frequent and extreme climate events. This can have significant impacts on building infrastructure, particularly the durability of building envelopes components. Therefore, there is a need to consider climate resiliency when designing new or retrofitting old building envelopes.

Performing field and laboratory tests to examine the limit of building materials under conditions like those that are expected in the future can be expensive, time consuming and requires specialized materials to test on and equipment’s to measure the performance. An alternative approach is to run the hygrothermal simulation models such as DELPHIN, WUFI, COMSOL, etc. which provide results from which to infer the durability and climate resiliency of building envelopes components. Although this mitigates the cost associated with a lab test, this approach is time-consuming since each simulation requires initial conditions setup and can takes days to run for 31 consecutive years. Moreover, the uncertainties associated with the future projected climate data results in several different sets of runs of simulation for the same year in a city. Hence computational intelligence and machine learning techniques are often looked upon.

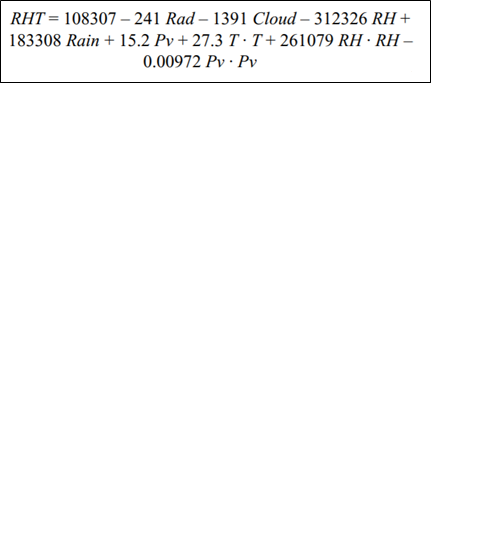
One approach for reducing the time associated with running hygrothermal simulations is by selecting Moisture Reference Years (MRYs). These are some selected years that are representative of the 31- year time series. For example, it could be the wettest plus the driest year among the series. Many methods have been developed to predict the performance of the year. Zhang et al. compared five different methods including Moisture Index (MI) Method, the PI-factor method, etc. to assess the ranking performance using several different damage functions such as Avg Mould, Max Moisture Content, DTDRH, etc. in 8 different cities in the US. However, none of the methods showed satisfactory results when compared to actual simulation results. They then constructed their own regression-based model, referred in this report as ASHRAE Report model, to predict the yearly RHT70 specifically in the North Orientation. In this report we compare the ASHRAE Report model with two purposed model, based on partial least square and support vector regression, to assess the hygrothermal performance of OSB for the North Orientation.

**Methodology**

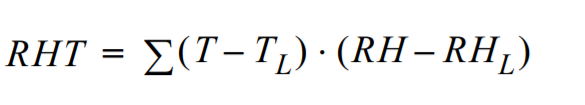
62 Years HAM Simulations consisting of 31 Historical Years (1986-2016) and 31 Future Years (2062-2092) were performed for Brick cladding in the North Orientation for each of the 3 cities -Vancouver, Calgary and Ottawa. From each city, 40 Simulations were randomly chosen to be included in the training dataset (totaling 120 simulations) while the remaining 22 simulations from each city accounted for a total of 66 test simulations.

First Objective was to compare SVR & PLS with ASHRAE Report model. For this purpose, both SVR & PLS were constructed using the same variables used by theASHRAE Report model.

The ASHRAE Report model is shown below



This include Yearly Average - Shortwave Radiation (W/m2), Cloudiness (Cloud Index), Relative Humidity (in decimal), Wind Driven Rain (mm), Vapor Pressure (Pa), Temperature (°C). The response RHT70 was calculated using the formula below. The limiting value for temperature and relative humidity are and respectively.

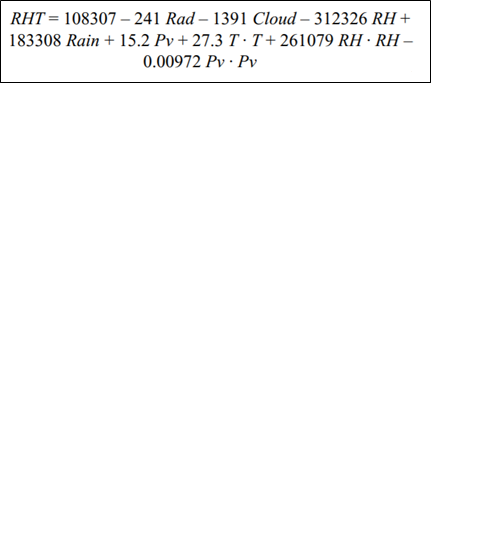


**Figure 1: Data Setup for the Training phase showing the Yearly Average Climate Variable and RHT70 of OSB**

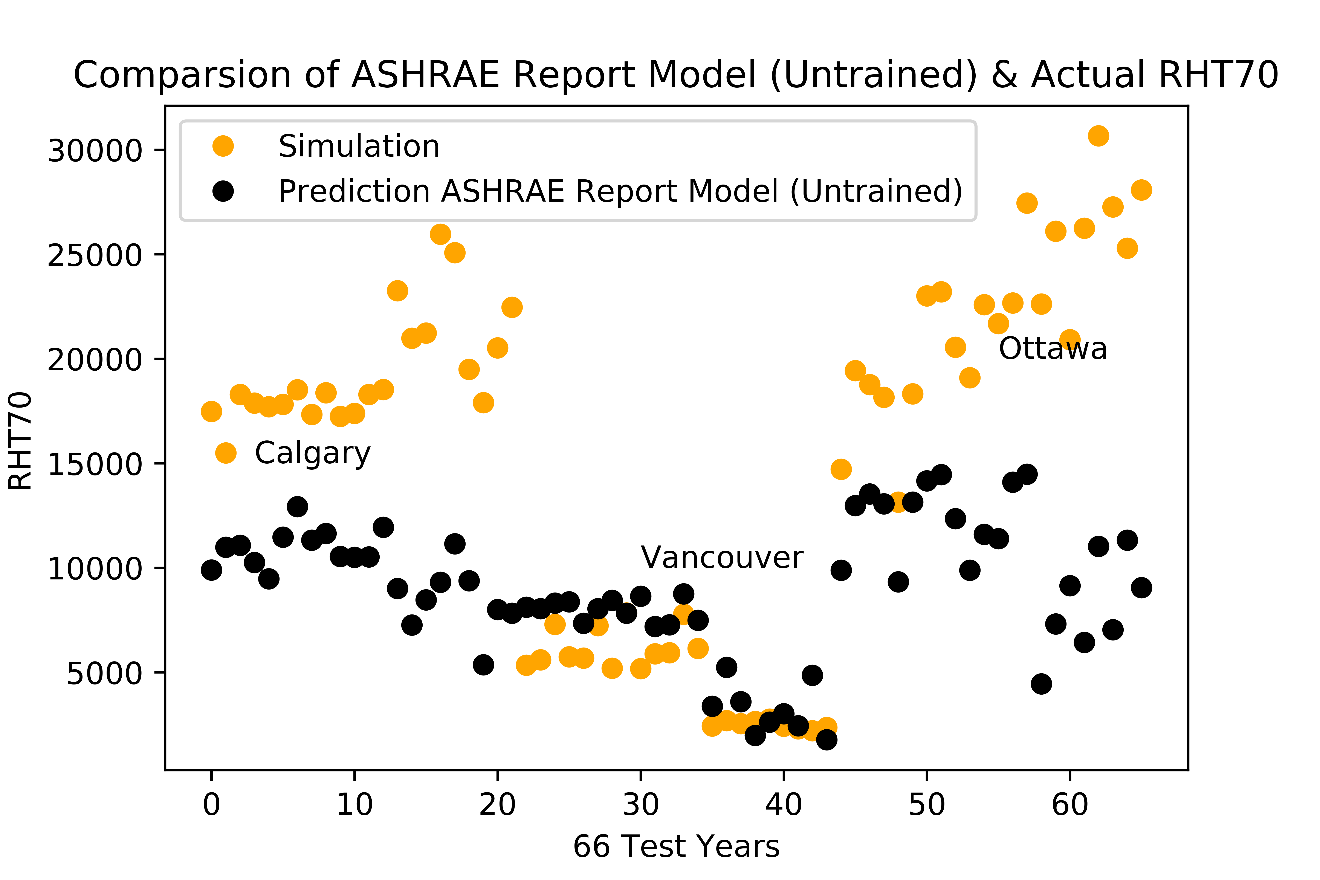


Four models were compared. PLS & SVR were trained and optimized on the current training set while ASHRAE Report model was directly used to make predictions on the test set. Additionally, ASHRAE Report model equation was reconstructed and trained on the current training set and used to make predictions to give it a fair chance.

ASHRAE Report Model **(Untrained) – Model 1**



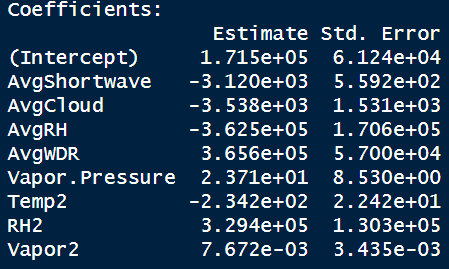
In the above equation, the appropriate climatic variables of the test set were used to calculate the RHT70. As seen in the Graph below, ASHRAE Report model tends to under predict RHT70 because it was trained on US cities where magnitude of RHT70 was lower compared to those seen in cities especially Calgary & Ottawa.

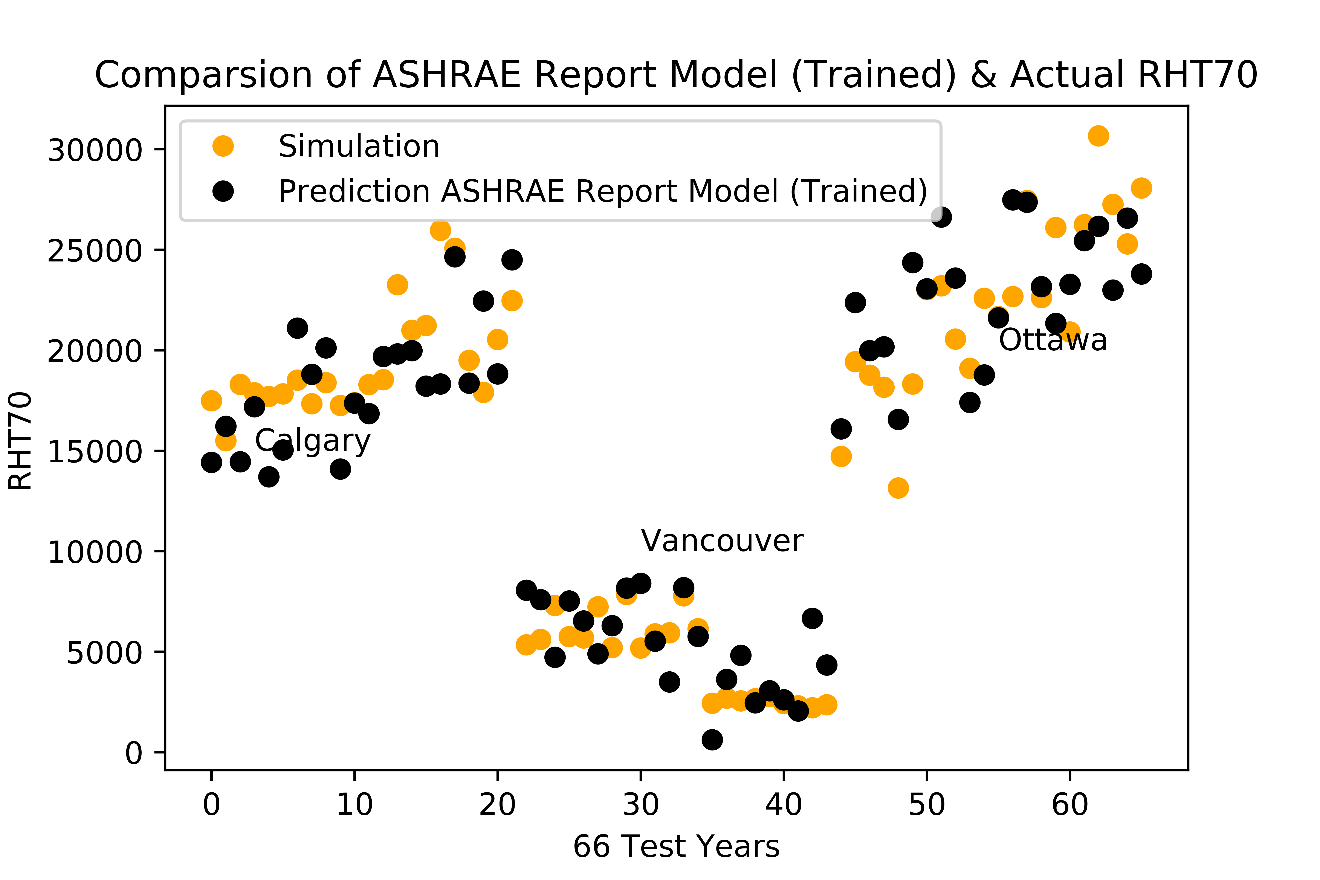


ASHRAE Report Model **(Trained) – Model 2**

Using the same variable as in the UntrainedASHRAE Report model, new Regression Equation was constructed based on the current training data. This new regression model was used to make prediction on the test set. Clearly it shows a significant improvement in predicting RHT70 compared to Untrained ASHRAE Report model.

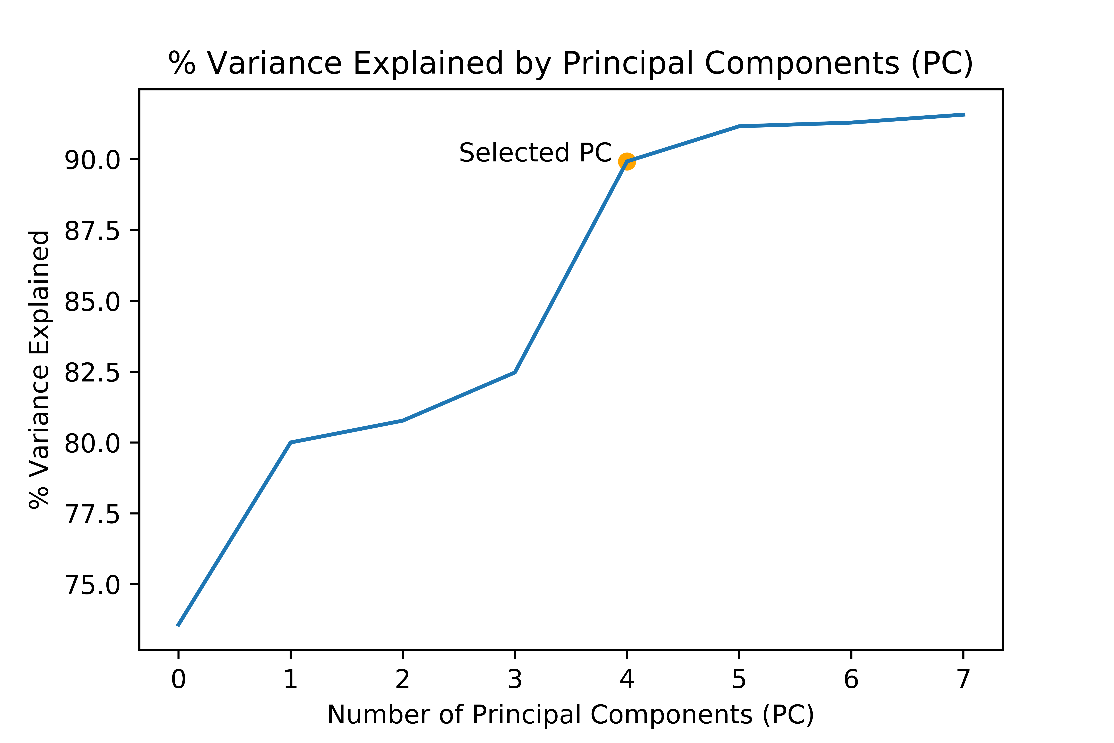
**Figure 2: Regression Coefficient associated with the Climate Variables**

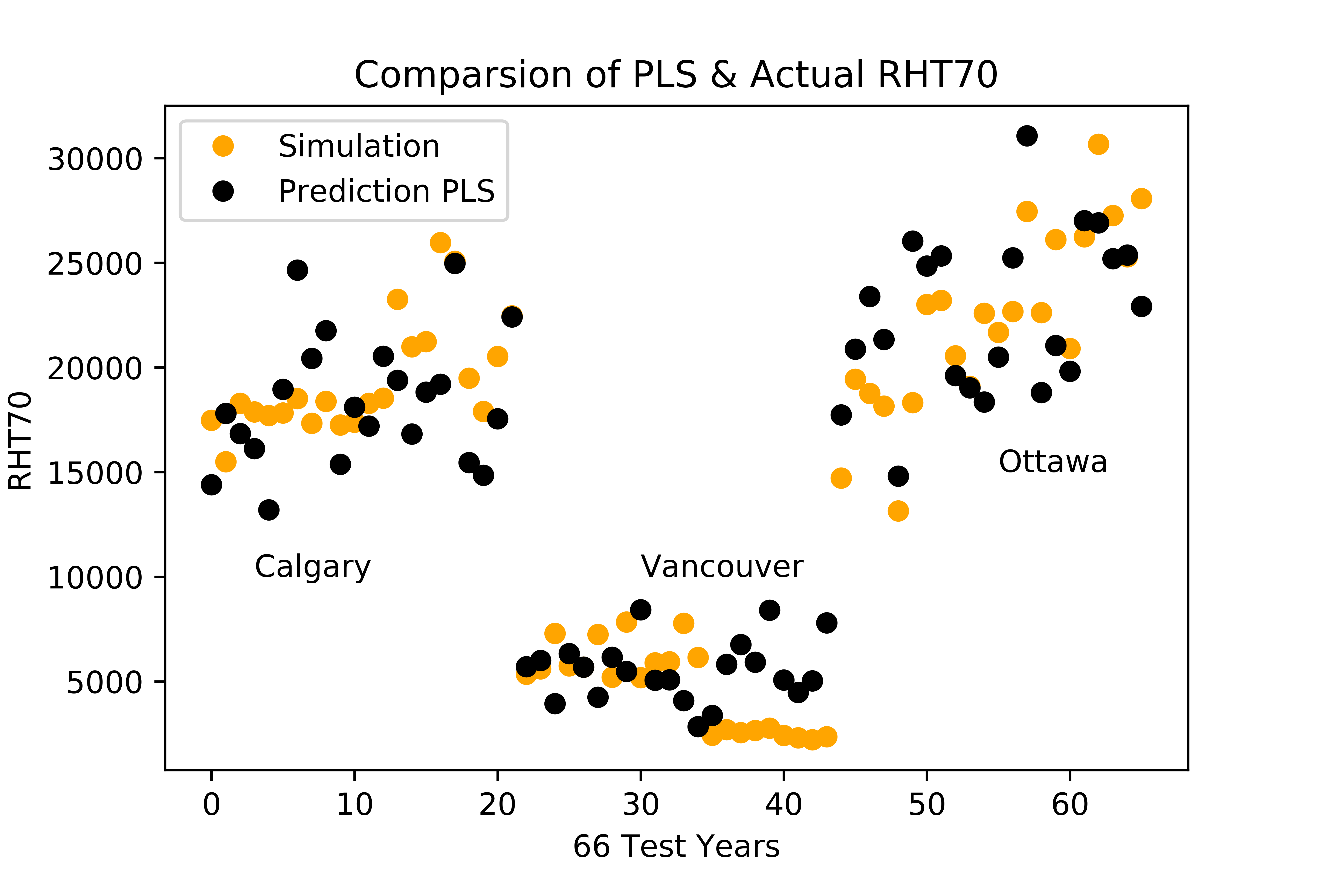




**Partial Least Square Regression – Model 3**

Due to possible correlation between climate variables, PLS regression was considered. Data was centered and scaled before been used to train the model. The optimal number of Principal Components was chosen based on the percentage of Variance explained in the response. In this case it was found to be 4 PCs. Then the corresponding model with 4 PCs was used to make prediction on the test set. The performance seems to be on par with the Trained ASHRAE Report model.

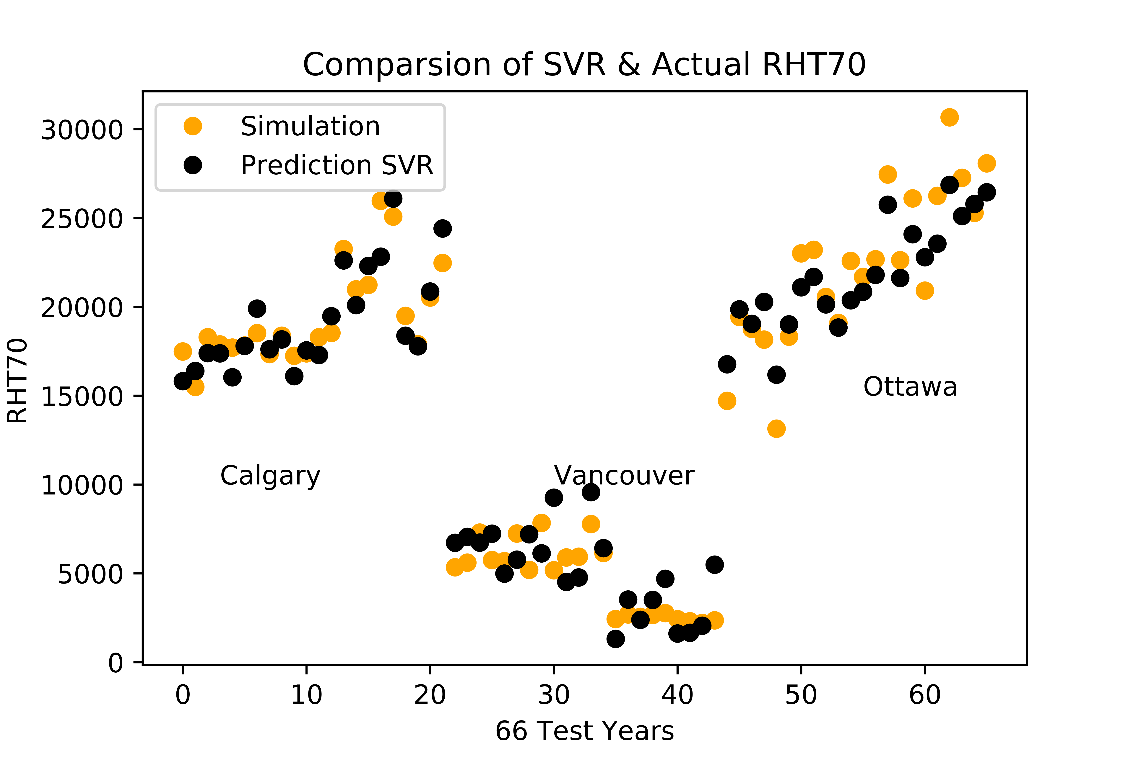




**Support Vector Regression – Model 4**

Since SVR is nonparametric i.e. it does not assume an underlying normal distribution of data, it was also considered. To optimize the parameter associated with SVR such as epsilon, gamma & Cost, grid search was performed with epsilon fixed at 0.1. Different combination of gamma & Cost was varied with 10-fold cross validation performed on the training set and the error calculated. Parameters combination which yielded the lowest error were then used for predicting on the test set.

Note data was scaled and centered before optimizing and training the model. Also, RBF kernel was used which is a simple function that can model systems of varying complexity. It is an extension of the linear kernel. The parameter gamma is associated with this kernel.



**Overall Results**

**Table 1: Performance Comparison of Different Model for Predicting RHT70 in Test Dataset**

|  |  |  |
| --- | --- | --- |
| Model | R2 | RMSEP |
| ASHRAE Report model (Untrained) | 39% | 9400 |
| ASHRAE Report Model (Trained) | 89% | 2684 |
| Partial Least Square Regression | 89% | 2713 |
| Support Vector Regression | 96% | 1555 |

It was clear from the above the results that SVR outperformed both PLS & ASHRAE Research Project Report model in accurately predicting the RHT70 in all three cities. It was also observed that since the ASHRAE Research Project Report model was constructed/ trained with data from some cities in the US, it was under predicting the RHT70 in Canadian cities. Even after training it on Canadian cities, SVR still seems to outperform the ASHRAE model.

New Method Based on SVR

To maximize the potential of the SVR technique, a new SVR model was constructed that was not restricted to the variables in the ASHRAE Report Model. Yearly Average - Shortwave Radiation (W/m2), Cloudiness (Cloud Index), Relative Humidity (in decimal), Wind Driven Rain (mm), Wind Speed (m/s), Vapor Pressure (Pa), Temperature (°C) were considered.

Note for variables like WDR & Solar Radiation taking the average of the year may not be the true reflection since there can be lot of nonevents (Zeros) that influence the average calculation. Therefore, the average of these 2 variables were calculated as shown below.

**Variable Selection**

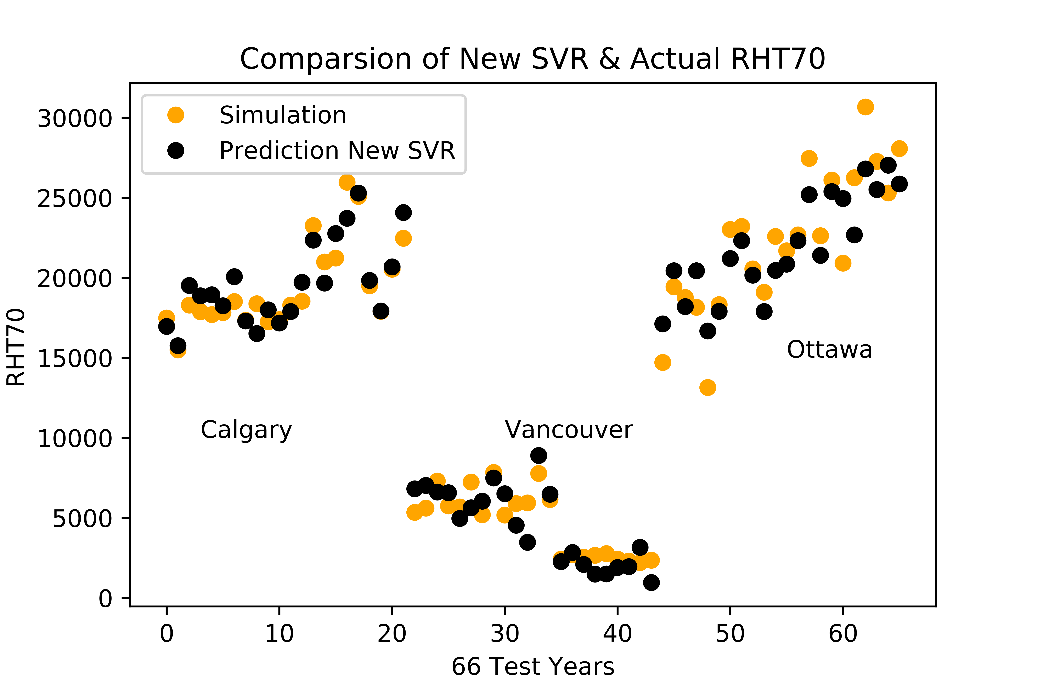
If there are *p* potential predictors, then there are 2*p* possible models. Generally, all the possible models are fitted and chosen based on some criteria. Here we use Bayes Information Criterion (BIC) that does not evaluate BIC for all possible models but uses a method like stepwise selection to compare models. BIC has advantage that is penalizes large models heavily and result in smaller models. Smaller BIC value means a better model.

Note: The BIC for a model is usually written in the form [-2logL + k\*p], where L is the likelihood function, p is the number of parameters in the model, and k is log(n) for BIC.

8 Possible Explanatory Variable included Yearly Average - Shortwave Radiation (W/m2), Cloudiness (Cloud Index), Relative Humidity (in decimal), Wind Driven Rain (mm), Wind Speed (m/s), Wind Direction (degrees), Vapor Pressure (Pa), Temperature (°C) were considered.

Variable Selection was performed using the Bayes Information Criterion (BIC) using the stepwise selection as “Both” i.e. at each stage in the process, a variable my be added or removed. This resulted in the following 6 variables:

Yearly Average - Shortwave Radiation (W/m2), Relative Humidity (in decimal), Wind Driven Rain (mm), Wind Speed (m/s), Temperature (°C) & Wind Direction (degrees).Using these 6 Variable, new SVR model were constructed and optimized to make prediction on the test dataset. The performance gains although minimal, was still the best among all the models considered.



**Table 2: Performance Comparison of Different Model for Predicting RHT70 in Test Dataset**

|  |  |  |
| --- | --- | --- |
| Model | R2 | RMSEP |
| ASHRAE Report model (Untrained) | 39% | 9400 |
| ASHRAE Report model (Trained) | 89% | 2684 |
| Partial Least Square Regression | 89% | 2713 |
| Support Vector Regression | 96% | 1555 |
| Support Vector Regression (New) | 97% | 1508 |

Conclusion

The objective of this study was to compare different machine learning approaches that could be used to select the Moisture Reference Years (MRYs). The ASHRAE Report model was used alongside SVR & PLS to predict the RHT70 of OSB while being constraint to the same explanatory climate variables.The result showed SVR outperformed both PLS & ASHRAE Research Project Report model in accurately predicting the Severity index in all three cities. It was also observed that since the ASHRAE Research Project Report model was constructed/ trained with data from some cities in the US, it was under predicting the Severity index in Canadian cities. Even after training it on Canadian cities, SVR still seems to outperform the ASHRAE model.A new model based on SVR was also constructed by calculating a more precise yearly average of WDR & Solar Radiation and using 6 climate variables selected through the BIC criteria. This model resulted in the tightest fit across the 3 cities among all the models.

Although this report considers the most successful model based on closeness to the actual data point, there is still a need to test all the 3 models on a larger dataset and compare them by the ranking of years in terms of RHT70 rather than accuracy of the RHT70.