A Weather Generator for Use in Construction Simulation Models

Adham Shahin¹,∗, Simaan M. AbouRizk², Yasser Mohamed²

¹CoSyn Technology, Edmonton, Alberta T6B 2T4, Canada
²Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta T6G 2W2, Canada

Abstract: In cold regions, weather poses uncertainty to construction projects. Project schedules can significantly deviate from original plans. Management requires reliable plans and schedules to set project activity times and baseline schedules against which the project performance will be measured; with winter weather uncertainty in cold regions, this task becomes quite challenging. In this paper, a universal weather generator for use in construction simulation purposes is presented, with the detailed approach for generating each of the different suggested weather parameters. The universal weather generator has been developed and given parameters for two cities in Alberta: Edmonton and Fort McMurray. The model parameters for both cities are given in order to facilitate future utilization of the model for academic and industrial research. To promote reusability and interoperability of the framework in future applications, the weather generator was developed using High Level Architecture (HLA), which showed great potential and suitability for this work.

Keywords: Simulation models, cold weather construction, project planning, weather simulation, decision support, weather generation

DOI: 10.7492/IJAEC.2013.008

1 INTRODUCTION

A considerable segment of Canada’s heavy construction takes place in cold weather. Construction projects are generally subject to several factors leading to uncertainty during the planning stage. Planning for cold weather construction in particular is characterized by a great deal of uncertainty in estimating activity durations and in determining the logic of performing the tasks. Cold weather can severely impact construction projects carried out in an open environment, leading to significant deviations from the scheduled finish dates. Many researchers cite weather as an influential factor in causing construction project delays (Koehn and Meilhede 1981; Laufer and Cohenca 1990). In fact, Benjamin and Greenwald (1973) suggest that approximately 50% of construction activities are affected by weather. The impacts vary from reduced productivity to complete work stoppage (Moselhi et al. 1997).

Other attempts to account for the weather impact on construction exist in the literature. Benjamin and Greenwald (1973) worked with a simulation model that simulates the construction duration by making daily work/no-work decisions based on historical weather data and the sensitivity of the construction activity to the different weather factors. Moselhi and Nicholas (1990) proposed a hybrid expert system that takes into account the impact of weather on construction planning and scheduling due to reduced labour productivity. Moselhi et al. (1995) developed a decision support system, Weather, to estimate the weather impact on activities productivities and durations. Moselhi et al. (1997) expanded the applications of Weather to include a function to account for interrupted duration due to precipitation and/or wind, the combined impact of weather on labour productivity and other interruptions.

In 2004, Shahin et al. (2004) presented a structured simulation framework that can be deployed in modeling construction processes executed in severe cold weather conditions. The framework identified the need for a universal construction weather generator that is ca-

*Corresponding author. Email: adham.shahin@gmail.com
pable for generating the weather parameters that are prevalent and impact construction projects.

This paper presents the weather generator component of the framework. This will permit development of the other components of the framework that will enable simulating and planning for the different construction processes affected by the extreme weather. The framework is developed in Simphony and also using High Level Architecture (HLA) principles to facilitate interoperability and reusability in future construction simulation and the output is validated using weather data for two Canadian cities.

2 CONSTRUCTION WEATHER GENERATOR REQUIREMENTS

To successfully simulate and generate sequences of weather for a particular location, a number of requirements should first be satisfied in the weather generator:

1. The stochastic processes underlying the different meteorological variables should be simulated based on an analysis of the historical weather data.
2. The correlations and dependencies among the meteorological variables (cross correlations) should be preserved in the generated weather sequences. For example, in summer and on a rainy day, the temperature is more likely to be below normal while the relative humidity is more likely to be above normal (Wales and AboutRzik 1996).
3. The time dependence within each weather variable (serial correlation) should also be preserved in the generated weather. For example, if a day has a significantly higher temperature than normal, then it is more likely that the following day will be higher than normal and vice versa. This will most likely be due to the heat stored in the soil.
4. The generated weather should keep the seasonal variations for each variable.

In addition to these general requirements, the weather generator should also satisfy the needs of the field in which it will be used. In the construction field, relevant weather variables are those that can affect the cost or productivity, or may cause complete work stoppage. Several weather variables are therefore needed in a weather model. Precipitation can adversely affect labor productivity, as well as the project schedule and cost or productivity, or may cause complete work stoppage. Several weather variables are therefore needed in a weather model. Precipitation can adversely affect labor productivity, as well as the project schedule and cost or productivity, or may cause complete work stoppage. Several weather variables are therefore needed in a weather model. Precipitation can adversely affect labor productivity, as well as the project schedule and cost or productivity, or may cause complete work stoppage. Several weather variables are therefore needed in a weather model. Precipitation can adversely affect labor productivity, as well as the project schedule and cost.

To generate daily weather conditions, precipitation is treated as the primary variable in this model, and is considered an independent variable. Maximum and minimum temperatures as well as relative humidity are regulated depending on the state of precipitation for that day; that is, wet or dry. Wind speed is generated without any correlation with the other variables and is treated as an independent variable. Frost penetration in the ground is calculated as a dependent variable. This parameter severely affects excavation activities as the excavation involves cutting through frosted soil. Although frost penetration is not a weather parameter by itself, its value is highly dependent on the temperatures recorded. It could therefore be included in the weather model as a dependent variable, the value of which will be calculated from the generated temperature variables. To summarize, the following parameters are targeted for generation by the construction weather generator: precipitation, maximum temperature, minimum temperature, maximum relative humidity, minimum relative humidity, average daily wind speed, and frost penetration in the ground.

The following sections detail the approach followed to generate the above-mentioned variables. The model parameter values for two Alberta weather stations, Edmonton International Airport and Fort McMurray Airport, are given.

3 GENERATION OF THE WEATHER VARIABLES

Historical weather data for the variables of interest were compiled and analyzed to determine the underlying stochastic processes of the meteorological phenomena to be simulated. This weather generation model is an extension of the work described by Richardson (1981). Generating daily weather variables for use in a simulation model would generally follow the flow chart outlined in Figure 1. Details of each of the flow chart components will be given in the following sections.

To generate daily weather conditions, precipitation is treated as the primary variable in this model, and is considered an independent variable. Maximum and minimum temperatures as well as relative humidity are regulated depending on the state of precipitation for that day; that is, wet or dry. Wind speed is generated without any correlation with the other variables and is treated as an independent variable. Frost penetration in the ground is calculated as a dependent variable. Its value is dependent on the maximum and minimum temperatures, as well as the soil type. In the following sections, the details of the modules to generate each of
Figure 1. Weather generation flow chart
these weather variables will be presented. The model parameters’ values were calculated for the Edmonton International Airport weather station and for the Fort McMurray Airport weather station using 42 years of historical records (1961-2002).

3.1 Precipitation Module

For the current weather generation purposes, precipitation is considered the primary variable. Its occurrence and amount can be generated in several ways. In the literature, a number of studies reported the success of using Markov chains in modelling precipitation (Wales and AbouRizk 1996; Richardson 1981; Kavvas et al. 1977; Katz 1985). For that reason, a first-order Markov chain was selected to model the precipitation component of the model. Precipitation is treated as an independent variable in this model; however, any other method that can model precipitation and generate its daily amount can be used as long as it has been successfully tested and accepted.

The Markov chain was modeled using the approach presented by Smith and Hancher (1989). A first-order, two-state (i.e., dry and wet) model was used to describe the precipitation state of the day. For the purposes of Markov chain transitional probabilities calculations, any day with a precipitation amount of 0.2 mm or more was considered a wet day. All other days were considered dry.

To calculate the weather model parameters’ values, the historical data was clustered into 12 months. To calculate the probability that a day in month m will be wet, Eq. (1) can be used.

\[
P_m(w) = \frac{n_m(w)}{N_m} (1)
\]

where \(P_m(w)\) = Probability that a day in month \(m\) will be wet; \(n_m(w)\) = Number of days in the historical records in which the state of precipitation was wet and month was \(m\); and \(N_m\) = Number of days in the historical records in which the month was \(m\).

For calculating transitional probabilities, Eq. (2) can be used.

\[
P_m(i/j) = \frac{n_{i,m}}{n_{j,m}} (2)
\]

where \(P_m(i/j)\) = Transitional probability from state \(j\) to state \(i\) for month \(m\); \(n_m(w)\) = Number of transitions from state \(j\) to state \(i\) for month \(m\) in the records; and \(N_{j,m}\) = Number of transitions from state \(j\) to any other state for month \(m\) in the records.

To fully define the state probabilities for a month, only three probabilities need to be defined: \(P_m(w)\), \(P_m(w/d)\), and \(P_m(w/w)\). The remaining transitional probabilities can be determined using Eqs. (3) and (4).

\[
P_m(d/w) = 1 - P_m(w/w) (3)
\]

\[
P_m(d/d) = 1 - P_m(w/d) (4)
\]

In reference to Figure 1, the state of precipitation for the first day (initial state of precipitation) in a simulation experiment will be determined by generating a uniform random number \((R_n)\) between 0 and 1. If \(R_n\) is less than or equal to \(P_m(w)\), then the initial precipitation state is set to “wet.” Otherwise, the first day’s precipitation state is set to “dry.” For all subsequent days in the simulation experiment, the precipitation state of the day will be conditioned based on the precipitation state of the previous day. For example, if the state of the previous day was dry, then the current day’s precipitation state will be determined by generating a uniform random number \((R_n)\) between 0 and 1. If \(R_n\) is less than or equal to \(P_m(w/d)\), then the current day’s precipitation state is set to “wet”; otherwise, the current day’s precipitation state is set to “dry.” Likewise, if the state of the previous day was “wet”, then the current day’s precipitation state will be determined by generating a uniform random number \((R_n)\) between 0 and 1. If \(R_n\) is less than or equal to \(P_m(w/w)\), then the current day’s precipitation state is set to “wet”; otherwise, the current day’s precipitation state is set to “dry.”

To generate a precipitation amount for wet days, a two-parameter Gamma distribution was used. Richardson (1981) used an exponential distribution, while Wales and AbouRizk (1996) used a two-state Gamma distribution. The Gamma distribution was selected due to its greater flexibility; this flexibility is due to its use of two parameters to describe the distribution. A separate Gamma distribution was defined for each month of the year to model the variation in precipitation amounts throughout a year. The Gamma distribution can be described by the formula in Eq. (5).

\[
f(x) = \frac{1}{\Gamma(\alpha)\beta^2} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) \quad x > 0 (5)
\]

The mean \(\mu\) and the variance \(\sigma^2\) of the Gamma distribution are related to the \(\alpha\) and \(\beta\) parameters of the distribution by Eqs. (6) and (7). To calculate the Gamma distribution parameters values for each month, the mean and variance of the historical records are calculated for each month of the year and then matched to the Gamma distribution using Eqs. (6) and (7).

\[
\mu = \alpha\beta (6)
\]

\[
\sigma^2 = \alpha\beta^2 (7)
\]

For generating the precipitation amounts on wet days, it should be noted that the procedure used for sampling from a Gamma distribution depends on the value of \(\alpha\). For values of \(\alpha\) between 0 and 1, the procedure proposed by Ahrens and Dieter (1974) was used. For values of \(\alpha\) greater than 1, the procedure proposed by Cheng (1977), which is shown in the wind speed module, can be used. Figure 2 illustrates pseudo-code used in their acceptance-rejection procedure. Table 1 lists the precipitation variables and transitional prob-
Figure 2. Pseudo code for Ahrens and Dieter (1974) Gamma variate procedure

abilities for the Edmonton International Airport and Fort McMurray Airport weather stations, calculated using the historical weather records for years from 1961 to 2002.

The historic time-series of each variable was reduced to a time-series of residual elements by removing the daily mean and standard deviation. Eqs. (8) and (9) (Richardson 1981) were used to determine the residual elements of each series. The elements’ time dependence (i.e. serial correlation within each variable and cross correlation between each pair of variables) was also determined in the process.

$$x_d(i) = \frac{X_d(i) - \overline{X_d}(i)}{\sigma_d(i)}, \quad \text{if } Pr_d = 0$$  \hspace{1cm} (8)

$$x_d(i) = \frac{X_d(i) - \overline{X_d}(i)}{\sigma_d(i)} + 1 \quad \text{if } Pr_d > 0$$  \hspace{1cm} (9)

where $x_d(i)$ = Residual element of parameter $i$ for day $d$ in the records; $X_d(i)$ = Value of parameter $i$ for day $d$ in the records; $\overline{X_d}(i)$ = Periodic mean of parameter $i$ for a dry day $d$ in the records; $\sigma_d(i)$ = Periodic standard deviation of parameter $i$ for a dry day $d$ in the records; $\overline{X_d}(i)$ = Periodic mean of parameter $i$ for a wet day $d$ in the records; $\sigma_d(i)$ = Periodic standard deviation of parameter $i$ for a wet day $d$ in the records; and $Pr_d$ = Precipitation amount for day $d$ in the records.

The daily mean and standard deviation for each variable are determined for each day of the year for both “wet” and “dry” conditions using the available historical weather data. The daily means and standard deviations are then smoothed using the Fast Fourier Transform (FFT) method implemented in MATLAB® version 6.5. The daily means and standard deviations for the parameter under consideration are listed in a data file in the form of a single column composed of 365 rows. The code simply reads the target variable’s data and returns a smoothed array of 365 data points, which will then be used to determine the residual elements from Eqs. (8) and (9). Figure 3 shows a plot of the mean daily maximum temperatures for wet days along with a plot of their smoothed values. Residual elements will be the basis of the weather generation scheme used. Figure 4 shows the maximum temperature residual series for a sample year (1967).

Finally, this work uses the weakly stationary generating process, suggested by Matalas (1967), to generate the weather data for the four parameters under consideration. This generates residual elements of the weather parameters by considering the residual element from the previous day plus a random component. The actual weather parameter values are determined by back-substituting in Eqs. (8) and (9). The weakly stationary generating process is defined by Eq. (10) (Matalas 1967) for $n$ weather parameters; $n$ is a positive integer that stands for the number of weather parameters considered in the process. In our case $n$ is equal to four parameters, and the weather parameters that we considered are max temperature, min temperature, max relative humidity, and min relative humidity.

$$x_d = Ax_{d-1} + B\varepsilon_d$$  \hspace{1cm} (10)

where $x_d = (n \times 1)$ matrix of residual elements for day $d$ for parameters 1 to $n$; $x_{d-1} = (n \times 1)$ matrix of residual
elements for day $d - 1$ for parameters 1 to $n$: $\mathbf{A}$ and $\mathbf{B} = (n \times n)$ matrices defined so that the correlations within and among the residual series are preserved; and $\mathbf{e}_d = (n \times 1)$ matrix of random components sampled from a standard normal distribution with a mean of 0 and a standard deviation of 1.

Eq. (10) implies that the residuals of maximum temperature, minimum temperature, maximum relative humidity, and minimum relative humidity are normally distributed and that the serial correlation within each parameter can be described by a first-order linear autoregressive model (Matalas 1967). Matrices $\mathbf{A}$ and $\mathbf{B}$ can be determined from Eqs. (11) and (12) (Matalas 1967).

$$\mathbf{A} = \mathbf{M}_1 \mathbf{M}_0^{-1}$$ (11)

$$\mathbf{B} \mathbf{B}^T = \mathbf{M}_0 - \mathbf{M}_1 \mathbf{M}_0^{-1} \mathbf{M}_1^T$$ (12)

where $\mathbf{M}_0 = (n \times n)$ lag 0 covariance matrix of the residual series; and $\mathbf{M}_1 = (n \times n)$ lag 1 covariance matrix of the residual series.

The variances of the residual series were found to equal approximately 1; consequently, $\mathbf{M}_0$ and $\mathbf{M}_1$ are $(n \times n)$ lag 0 and lag 1 cross correlation coefficients of the residual series. The components $\mathbf{M}_0$ and $\mathbf{M}_1$ are defined by the matrices in Eqs. (13) and (14).

$$\mathbf{M}_0 = \begin{bmatrix}
1 & \rho_0(1,2) & \rho_0(1,3) & \rho_0(1,4) \\
\rho_0(2,1) & 1 & \rho_0(2,3) & \rho_0(2,4) \\
\rho_0(3,1) & \rho_0(3,2) & 1 & \rho_0(3,4) \\
\rho_0(4,1) & \rho_0(4,2) & \rho_0(4,3) & 1
\end{bmatrix}$$ (13)

$$\mathbf{M}_1 = \begin{bmatrix}
\rho_1(1) & \rho_1(1,2) & \rho_1(1,3) & \rho_1(1,4) \\
\rho_1(2,1) & \rho_1(2,2) & \rho_1(2,3) & \rho_1(2,4) \\
\rho_1(3,1) & \rho_1(3,2) & \rho_1(3,3) & \rho_1(3,4) \\
\rho_1(4,1) & \rho_1(4,2) & \rho_1(4,3) & \rho_1(4,4)
\end{bmatrix}$$ (14)

where $\rho_0(i,j) = \text{Lag 0 cross correlation coefficient between the residual series for parameters } i \text{ and } j$; $\rho_1(i,j) = \text{Lag 1 cross correlation coefficient between the residual series for parameters } i \text{ and } j \text{ with parameter } j \text{ lagged one day with respect to parameter } i$; and $\rho_1(i) = \text{Lag 1 serial correlation coefficient for the residual series for parameters } i$.

For the proposed universal construction weather generator under consideration, four parameters are being generated (maximum temperature, minimum temperature, maximum relative humidity, and minimum relative humidity). The values of $\mathbf{M}_0$ and $\mathbf{M}_1$ for the Edmonton International Airport weather station were calculated using the historical records between the years of 1961 and 2002 and are shown in Eqs. (15) and (16).

$$\mathbf{M}_{0E} = \begin{bmatrix}
1 & 0.673 & 0.004 & -0.410 \\
0.673 & 1 & 0.0317 & -0.004 \\
0.004 & 0.0317 & 1 & 0.489 \\
-0.41 & -0.004 & 0.489 & 1
\end{bmatrix}$$ (15)

$$\mathbf{M}_{1E} = \begin{bmatrix}
0.643 & 0.471 & 0.085 & -0.189 \\
0.634 & 0.632 & 0.105 & -0.029 \\
0.013 & 0.093 & 0.613 & 0.428 \\
-0.206 & -0.003 & 0.319 & 0.526
\end{bmatrix}$$ (16)

The same calculations were carried out for the Fort McMurray Airport station for the years 1961 to 2002, and the $\mathbf{M}_0$ and $\mathbf{M}_1$ are shown in Eqs. (17) and (18).

$$\mathbf{M}_{0F} = \begin{bmatrix}
1 & 0.442 & 0.008 & -0.270 \\
0.442 & 1 & 0.012 & 0.020 \\
0.008 & 0.012 & 1 & 0.476 \\
-0.270 & 0.020 & 0.476 & 1
\end{bmatrix}$$ (17)

$$\mathbf{M}_{1F} = \begin{bmatrix}
0.378 & 0.148 & 0.052 & -0.121 \\
0.213 & 0.103 & 0.024 & -0.020 \\
0.019 & 0.042 & 0.524 & 0.417 \\
-0.090 & 0.020 & 0.291 & 0.451
\end{bmatrix}$$ (18)

Once the coefficients of the $\mathbf{M}_0$ and $\mathbf{M}_1$ matrices are calculated, solving for the $\mathbf{A}$ matrix is straightforward; however, solving for the $\mathbf{B}$ matrix is somewhat more complex. The $\mathbf{B}$ matrix can be calculated using the procedure suggested by Young (1968), in which matrix $\mathbf{B}$ is assumed to be a lower triangular matrix. In addition, matrix $\mathbf{C}$ is needed in the procedure to solve for matrix $\mathbf{B}$. To solve for matrix $\mathbf{B}$, the coefficients were derived and can be calculated from the set of equations described in Young (1968). Using these equations, the
A and B matrices for the Edmonton International Airport weather station for the years from 1961 to 2002 were calculated and are shown in Eqs. (19) and (20).

\[
M_{AE} = \begin{bmatrix}
0.625 & 0.048 & 0.064 & 0.037 \\
0.518 & 0.284 & 0.005 & 0.183 \\
0.086 & 0.020 & 0.507 & 0.215 \\
-0.014 & 0.006 & 0.085 & 0.479
\end{bmatrix}
\]
\[
M_{BE} = \begin{bmatrix}
0.760 & 0 & 0 & 0 \\
0.316 & 0.630 & 0 & 0 \\
-0.084 & -0.089 & 0.761 & 0 \\
-0.422 & 0.221 & 0.284 & 0.641
\end{bmatrix}
\]

\[
M_{AF} = \begin{bmatrix}
0.368 & -0.014 & 0.077 & -0.058 \\
0.221 & 0.004 & 0.004 & 0.037 \\
0.085 & -0.005 & 0.406 & 0.246 \\
0.020 & 0.002 & 0.095 & 0.411
\end{bmatrix}
\]

\[
M_{BF} = \begin{bmatrix}
0.923 & 0 & 0 & 0 \\
0.393 & 0.894 & 0 & 0 \\
-0.016 & -0.005 & 0.827 & 0 \\
-0.252 & 0.135 & 0.304 & 0.784
\end{bmatrix}
\]

All the components of Eq. (10) are now fully defined. With reference to Figure 1, the maximum temperature, minimum temperature, maximum relative humidity, and minimum relative humidity can be generated by applying the weekly stationary generating

---

**Figure 3.** Maximum Temperature Means (“wet” days)

**Figure 4.** Maximum Temperature Residual Series (1967)
process using the previous day’s residual elements and Eq. (10). Nevertheless, the previous residual elements should only be initialized for the first day in the simulation run. The initialization process is accomplished by sampling the previous day’s residuals from a normal distribution with a mean of 0 and a variance of 1. The weakly stationary generating process is then applied seven consecutive times. This is done to reduce any transitional start-up problems. In all cases, the values of the four parameters to be generated are determined by back-substitution in Eqs. (8) or (9), depending on the precipitation status.

3.2 Wind Speed Module

Wind speed is generated without any correlation to other variables. The variable of interest was the average daily wind speed. Environment Canada keeps historical records of the wind speed for the 24 hours of the day recorded every hour. The data was downloaded from Environment Canada’s website using an online query. There were therefore 24 readings for each historical day; these readings were then averaged, resulting in the daily average wind speed. All these calculated daily averages were grouped into 12 groups based on the 12 months of the year. The data for each group was fit to a 2-parameter Gamma distribution. Three “goodness of fit” tests, Chi-square, Kolmogorov-Smirnov, and Anderson-Darling, were applied to the fitted distributions.

Table 2 lists the α and the β parameters of the Gamma distribution for the 12 months of the year along with the three “goodness of fit” test results for the Edmonton International Airport and Fort McMurray Airport weather stations for the years from 1961 to 2002. Note in the Fort McMurray results, neither the data from December nor that from January conformed to a Gamma distribution. A better option was to fit the data to a Beta distribution. For December, a Beta (1.719, 3.561, 0, 25.709) and for January, a Beta (1.947, 4.526, 0, 27.918) would pass two “goodness of fit” tests.

For sampling from the Gamma distribution, the α parameters for all 12 months are greater than 1; the procedure outlined in Figure 2 cannot, therefore, be used. Figure 5 shows pseudo-code for Cheng (1977) acceptance-rejection procedure, which is suitable for situations where α is greater than 1. In the simulation run, an average wind speed for the day is sampled from the Gamma distribution of the appropriate month.

3.3 Frost Penetration Module

Frost penetration in the ground has a significant effect on many construction activities. Although frost penetration is not a weather parameter by itself, its value is highly dependent on the temperatures recorded. It could therefore be included in the weather model as a dependent variable, the value of which will be calculated from the generated temperature variables.

Eq. (23) (Sego 2005) shows the governing differential equation in one dimension required for calculating the frost penetration in the ground.

\[ K \frac{\partial^2 T}{\partial z^2} + L = C \frac{\partial T}{\partial t} \]  

where \( K \) = conductivity, the quantity of heat flow through a unit area of substance of unit thickness in unit time under a unit temperature gradient \( (J \cdot s^{-1} \cdot m^{-1} \cdot K^{-1}) \); \( C \) = volumetric specific heat; \( L \) = volumetric latent heat, the quantity of heat liberated when a unit volume of soil undergoes phase change without a temperature change \( (kJ \cdot m^{-3}) \); \( \partial^2 T/\partial z^2 \) = second derivative of temperature with respect to depth; \( \partial T/\partial t \) = first derivative of temperature with respect to time; \( t \) = time (day); \( z \) = depth; and \( T \) = temperature.

Conductivity \( K \) depends on the mineral composition of the soil and the soil state, whether frozen, written as \( K_f \), or unfrozen, written as \( K_u \). Values of \( K_f \) and \( K_u \) for different soil conditions (coarse grained soils, fine grained soils, and peats) can be obtained using thermal conductivity charts in Sego (2005) that relate the moisture content and the dry density of soil to the thermal conductivity. Latent heat \( L \) can be calculated using Eq. (24) (Sego 2005).

\[ L = 334 \cdot \rho_d \cdot m_e \]  

where \( \rho_d \) = dry density of soil \( (kg \cdot m^{-3}) \); and \( m_e \) = moisture content of soil.

Solving for Eq. (23) is very complex; however, Stefan’s solution offered a relatively simple approximation of the problem, which is widely accepted (Sego 2005). Stefan’s solution is presented in Eq. (25). Sego (2005) gives the material properties that are needed for different kinds of soils - sand, gravel, clay - in applying Eq. (25).

\[ Z_0 = \left[ \frac{2K_f T_s t}{L} \right]^{0.5} = \left[ \frac{2K_f (GFI)}{L} \right]^{0.5} \]  

where \( Z_0 \) = depth of frost penetration \( (m) \); \( T_s \) = surface temperature \( (°C) \); and GFI = ground freezing index = \( N_f \cdot AFI \), where \( N_f \) = N-factor relating the air freezing index to the ground freezing index; \( AFI = \Sigma(TM) \), which is the air freezing index used for the days when TM is below 0°C; and TM = mean daily temperature, which is the average of the maximum and minimum daily temperatures.

In Figure 6, which shows a typical smoothed yearly temperature curve, the day of the year for points A and B should be identified. This is accomplished by averaging the smoothed curves for the maximum and minimum temperatures resulting in the average temperature smoothed curve. The points of intersection with the zero temperature level are identified as points A and B. For calculating the incremental daily frost penetration depth, it is assumed that freezing can only occur between points A and point B (freezing period) and that thawing can only occur outside of this period.
Table 2. Gamma distribution parameters and goodness of fit results

<table>
<thead>
<tr>
<th>Month</th>
<th>Edmonton International Airport</th>
<th>Fort McMurray Airport</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>January</td>
<td>4.658</td>
<td>2.503</td>
</tr>
<tr>
<td>February</td>
<td>5.431</td>
<td>2.134</td>
</tr>
<tr>
<td>March</td>
<td>5.634</td>
<td>2.212</td>
</tr>
<tr>
<td>April</td>
<td>6.551</td>
<td>2.2</td>
</tr>
<tr>
<td>May</td>
<td>5.932</td>
<td>2.487</td>
</tr>
<tr>
<td>June</td>
<td>6.387</td>
<td>2.208</td>
</tr>
<tr>
<td>July</td>
<td>6.303</td>
<td>2.086</td>
</tr>
<tr>
<td>August</td>
<td>6.264</td>
<td>1.976</td>
</tr>
<tr>
<td>September</td>
<td>5.626</td>
<td>2.352</td>
</tr>
<tr>
<td>October</td>
<td>5.769</td>
<td>2.261</td>
</tr>
<tr>
<td>November</td>
<td>4.913</td>
<td>2.41</td>
</tr>
<tr>
<td>December</td>
<td>4.715</td>
<td>2.475</td>
</tr>
</tbody>
</table>

Figure 5. Pseudo code for Cheng (1977) Gamma variate procedure ($\alpha$ > 1, $\beta$ > 0)

done = False
$a = (2^a - 1) \cdot \beta^3$
$b = a - \log(4)$
$g = \alpha + \alpha^{-1}$
d$= 1 + \log(4.5)$

While Not done
  $u_1 = \text{Uniform}(0,1)$
  $u_2 = \text{Uniform}(0,1)$
  $v = a \cdot \log \left( \frac{u_1}{1 - u_1} \right)$
  $y = \alpha \cdot e^v$
  $z = u_1^2 \cdot u_2$
  $w = b + g \cdot v \cdot y$
  If ($w + d - 4.5 \cdot z > 0$) Then
    done = True
  Else
    If $w > \log(z)$ Then done = True
  End If
End While
Gamma Variate = $\beta(y)$

For the two weather stations under consideration, the calculated freezing period limits are October 31 (A) to April 3 (B) for Edmonton International Airport and October 26 (A) to April 8 (B) for Fort McMurray Airport. The assumption that freezing can only occur during the freezing period and thawing can only occur during the non-freezing period is an approximation; however, the margin for error in the frost depth created as a result of such an approximation is negligible.

Eq. (26) can be used for calculating the incremental daily frost depth increase ($\Delta Z_{+ve}$) during the freezing period. Outside of the freezing period, the incremental daily frost depth decrease ($\Delta Z_{-ve}$) will be calculated using Eq. (27).

$$\Delta Z_{+ve} = \begin{cases} 
\left[ \frac{2K_f \times N_f \times TM}{L} \right]^{0.5} & \text{if TM < 0,} \\
0 & \text{otherwise}
\end{cases}$$  (26)

$$\Delta Z_{-ve} = \begin{cases} 
\left[ \frac{2K_u \times N_u \times TM}{L} \right]^{0.5} & \text{if TM > 0,} \\
Z > 0, \text{ and thawing period} & \text{otherwise}
\end{cases}$$  (27)
where \( \Delta Z_{\text{+ve}} \) = Incremental daily increase in frost depth during freezing period; \( \Delta Z_{\text{-ve}} \) = Incremental daily decrease in frost depth during thawing period; and \( N_t \) = N-factor relating the air thawing index to the ground thawing index. A number of studies were conducted in Alaska and Canada that established the N-Factors for different ground surface conditions and these values can be found in Sego (2005).

The frost penetration module has to be initialized. Since the calculations are based on incremental daily freezing or thawing, the purpose of initialization is to set the initial startup frost depth at a reasonable value on the project start date. In non-permafrost regions, it can be safely assumed that there is no frost at point A. The frost penetration module would therefore always start from a point A that is prior to the construction project’s start date. At that point in the simulation experiment, the weather model would generate weather solely for the frost penetration module’s calculations. This would continue until the construction project’s start date, after which time the simulation model and the weather generator will advance at a normal pace. The purpose of the proposed initialization was to ensure that a reasonable initial frost depth was assumed for the first day of the construction project. After this point in time, the incremental daily calculations would be performed as outlined.

4 HIGH LEVEL ARCHITECTURE AND ITS WEATHER SIMULATION APPLICATION

In today’s complex simulation models, a simulation model can be composed of multiple simulations, each simulating the role of a specific aspect of the environment it is modelling. In many cases, parts of those simulations may have already been developed for another application, and the task is to link the simulation models to simulate the environment of interest. Unfortunately, this task is not as simple as it sounds. For example, it may be necessary to make extensive modifications to the simulation models in order to adapt the simulation components to one another. In many cases, it might prove more convenient to redevelop the simulation component from scratch than to make the necessary modifications.

HLA provides a framework that allows several computer simulations (federates) to be combined into a larger simulation model (federation). For example, for a tunnelling federation to be developed using the tunnelling construction application, a number of supporting tunnelling federates may be required. A tunnel face federate simulates the excavation; a transportation federate simulates dirt removal trains, and a soil removal federate is responsible for operating the crane that hoists the muck cars at the shaft location to get rid of the excavated soil. Each of these federates can be developed separately; however, for seamless communication between federates, each federate must conform to the HLA framework specifications.

There are two important properties of HLA-compliant simulations: reusability and interoperability. The simulation models, or “federates,” can be reused with a different simulation scenario or even with an entirely different simulation application, and the reusable simulation component can be combined with the other simulation components into a “federation” without the need for recoding. This means that component simulations that are running on a number of distributed platforms can be combined together even if the platforms are different in type.

HLA helps combine the distributed simulation components together in a single simulation execution. It also helps extend the functionality of the combined simulation model by introducing other simulation compo-
4.1 Components of an HLA Compliant Federation

An HLA-compliant federation is defined by three main documents:

1. Federation rules: A collection of principles to ensure the proper interaction of federates during a simulation. It also outlines the responsibilities of the federates and simulation.
2. Object Model Template (OMT): A meta-model for all Federation Object Models (FOM), which establishes their allowed structure.
3. Interface Specifications: Specifies a standardized interface between the simulation federates and the Run Time Infrastructure (RTI).

HLA is an architecture, not a software implementation; however, the RTI and the developed federates are coded software that should conform to the HLA specifications. In HLA, a federation needs to have access to the following components (Kuhl et al. 1999).

1. Run Time Infrastructure (RTI): A software that conforms to the HLA specifications. It provides all the software services that are needed to support the HLA federation execution. It is considered to be the backbone through which all the communications and interactions between the federates go through.
2. Federation Object Model (FOM): A single, common object model per federation, which is developed for the data exchange between the federates (e.g., objects, interactions).
3. A collection of federates comprising the federation.

4.2 HLA and the Weather Simulation Application

HLA, as described earlier, has two appealing properties for our purposes: reusability and interoperability. A universal weather generator, developed as a stand-alone HLA federate, will enable integration with other construction federates (for example, tunnelling or pipeline installation) without recoding, relieving researchers of the overhead effort involved in using a weather generator in their work.

The Universal Weather Generator federate was developed for the version 2 HLA framework to run with the current version of the RTI developed at the University of Alberta. This weather generator generates all the weather parameters previously identified in this paper and was combined with an HLA-compliant general construction federate responsible for reading the weather parameters values generated by the weather federate for use in the context of a construction process simulation. The federation was then successfully tested with the two federates running on separate computers. Figure 7 shows the output interface of the general construction federate.

5 VALIDATION OF THE UNIVERSAL WEATHER GENERATOR

Tests were made to validate the model assumptions for the historical weather data for the two weather stations under consideration. Next, statistical tests of the generated weather output for both stations were carried out to ensure that the similarity between the historical weather and the generated weather is statistically adequate.
Table 3. Moments of residual data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Edmonton International Airport</th>
<th>Fort McMurray Airport</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum temperature</td>
<td>$\mu = -0.00054$</td>
<td>$\sigma = 1.0036$</td>
</tr>
<tr>
<td>Minimum temperature</td>
<td>$\mu = -0.0017$</td>
<td>$\sigma = 1.0049$</td>
</tr>
<tr>
<td>Maximum rel. humidity</td>
<td>$\mu = 0.3$</td>
<td>$\sigma = 0.886$</td>
</tr>
<tr>
<td>Minimum rel. humidity</td>
<td>$\mu = 0.1538$</td>
<td>$\sigma = 1.022$</td>
</tr>
</tbody>
</table>

Figure 8. Maximum temperature residuals NPP (Edmonton international airport)

Figure 9. Maximum temperature residuals NPP (Fort McMurray airport)
Figure 10. Autoregressive model vs. serial correlation for minimum temperature
(Edmonton international airport)

Figure 11. Autoregressive model vs. serial correlation for minimum temperature
(Fort McMurray airport)

Table 4. Comparison of minimum temperatures at Edmonton international and Fort McMurray airports

<table>
<thead>
<tr>
<th>Month</th>
<th>Edmonton International Airport</th>
<th>Fort McMurray Airport</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual µ</td>
<td>Actual σ</td>
</tr>
<tr>
<td>October</td>
<td>-2.437</td>
<td>1.531</td>
</tr>
</tbody>
</table>

Note: Italics indicate temperatures were not within 99% confidence interval.
5.1 Test of Assumptions

For the weather parameters generated using the weakly stationary generating process, i.e. maximum temperature, minimum temperature, maximum relative humidity, and minimum relative humidity, the first model assumption to be tested is that the residual series of the parameters are normally distributed. To test this, the moments of the residual series were first calculated, and then compared to the normal distribution values. In addition, the normal probability plots for the residuals were plotted to assess the normality assumption. For brevity the results for only the maximum temperature parameter are shown.

Table 3 shows the mean, standard deviation, skewness, and kurtosis for the residual data taken from the Edmonton International Airport and Fort McMurray weather stations. Figure 8 shows the normal probability plots for the residual series of the weather parameters used for the Edmonton International Airport, and Figure 9 shows the normal probability plots for the residual series for the Fort McMurray Airport weather parameters.

In general, for the normality of the residuals, the mean and standard deviation are very close to the standard normal distribution values, however, some skewness was detected and the residuals appeared to have flatter kurtosis. For the normal probability plots, the normality assumption was generally not violated.

The second assumption to be tested is that the serial dependence of the residual series approximates a first-order linear autoregressive model. Serial dependence of a first-order autoregressive model can be defined as shown in Eq. (28) (Richardson 1981):

$$\delta_k = \delta_1^k$$

(28)

where $$\delta_k$$ = Serial correlation with lag of $$k$$ days.

For brevity, Figure 10 shows the comparisons between the first-order model and the serial correlation of the residual series for only the minimum temperature for the Edmonton International Airport weather station; Figure 11 shows the same plot for the Fort McMurray Airport weather station. Based on these, it can be concluded that the serial dependence of the residual series did not significantly deviate from the first-order autoregressive assumption and can be approximated by a first-order autoregressive model. In the next section, the output of the weather generator will be tested statistically to ensure that the weather generator produces an acceptable output.

5.2 Test of Assumptions

To test the output of the weather generator statistically, 30 years of simulated data were generated for the Edmonton International Airport weather station and for the Fort McMurray Airport weather station, respectively. For the generated data, the number of wet days per month, the amount of daily precipitation, the maximum temperature, the minimum temperature, the maximum relative humidity, the minimum relative humidity, and the average daily wind speed were computed. The monthly means of the weather parameters’ values were then calculated for each month of the year. The historical monthly means of the weather parameters were also calculated for the above mentioned weather parameters. Means and standards deviations for each month were calculated for the historical data.

Confidence intervals for the historical weather parameters were constructed at the 99% level for each of the 12 months. Finally, the simulated means were compared to see if they fell within the constructed confidence interval. Table 4 summarizes the above calculations for the Edmonton International Airport and Fort McMurray Airport weather stations. For brevity only the output testing results for minimum temperatures are shown.

Most of the generated means were within the 99% confidence interval for minimum temperatures for each month, and, since the number of failing tests is limited, the conclusion is that the models for these two weather stations were successful. It is believed that the weather generators for the two cities will operate efficiently in the construction simulation models.

6 CONCLUSION

This paper presented the development and successful implementation of a weather generator that can be used as an important building block for simulating construction processes that take place in severe cold weather; extending the analytical weather generator models in order to generate the different weather parameters significantly affecting construction enables researchers to study and model the impact of weather on the different construction processes. This paper began by identifying the weather parameters that are important to the construction industry. Details of the steps that are needed to successfully simulate the weather were given. The model output was validated against two weather data series from two northern Canadian cities, Edmonton, Alberta, Canada, and Fort McMurray, Alberta, Canada and was found to be accurate.

To promote further use of the model, the weather generator was implemented as an HLA federate that can be used in a construction project federation, which enables distributed simulation. This development is anticipated to promote the reuse and reapplication of the previously developed framework and weather generator to assess the impact of cold weather on various construction processes. As an integral component of a framework for developing construction process simulation models that take into account weather uncertainty, the weather generator will facilitate incorporation of the high level of weather uncertainty seen in cold regions’ winter weather into the construction simulation
models. This weather generator can also be used with other industries where weather is seen as a significant factor contributing to uncertainty.

ACKNOWLEDGEMENTS
Funding for this research was partly provided by the University of Alberta, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Alberta Ingenuity Fund.

REFERENCES