The Canadian society for engineering in agricultural, food, and biological systems

CSAE SCGR

La société canadienne de génie agroalimentaire et biologique

Paper No. 03-118

CLIMGEN - A CONVENIENT WEATHER GENERATION TOOL FOR CANADIAN CLIMATE STATIONS

K. McKague
Ontario Ministry of Agriculture and Food, P.O. Box 666, Woodstock, ON, Canada N4S 7Z5

R. Rudra
School of Engineering, University of Guelph, Guelph, ON, Canada N1G 2W1

J. Ogilvie
School of Engineering, University of Guelph, Guelph, ON, Canada N1G 2W1

Written for presentation at the
CSAE/SCGR 2003 Meeting
Montréal, Québec
July 6 - 9, 2003

Abstract

Applying continuous-in-time rural water quality models to simulate the long-term environmental benefits of agricultural best management practices requires a parallel set of long-term continuous weather data for the area being modelled. Acquiring this data and formatting it for use as model input can be time-consuming. This has led to the development and use of weather generation procedures and tools. Few weather generators available today have been fully parameterized for Canadian weather stations. In this paper, generated values for daily precipitation, air temperature, solar radiation, wind speed and relative humidity from the weather generator ClimGen (version 3.8) were compared with corresponding historical and published 30-year station normals for a set of six southern Ontario test stations. ClimGen was found to perform as well or better than WGEN, a popular weather generator used for many US weather stations. ClimGen (version 4.1.05), which has the capability of directly importing Environment Canada station archival data files is available for downloading, evaluation, and use at www.bsyse.wsu.edu/climgen

ClimGen – A Convenient Weather Generation Tool for Canadian Climate Stations

K. McKague[1], R. Rudra[2] and J. Ogilvie[2]

[1]Resources Management Branch, Ontario Ministry of Agriculture and Food, Woodstock, Ontario, Canada N4S 7Z5; and [2]School of Engineering, University of Guelph, Guelph, ON, Canada N1G 2W1.

BACKGROUND

Weather input drives many of the processes simulated in hydrologic models including evaporation, infiltration and runoff (Arnold and Williams 1989). Similarly comprehensive crop growth models use weather data as input when simulating transpiration rates and soil microbial activity. Hydrologic models are often combined with crop growth models for the purposes of simulating the effects of agricultural practices on rural surface water and groundwater quality. Normally weather data including precipitation, air temperature, solar radiation, wind speed, and air humidity are needed on a daily basis. Since some hydrologic models require more detailed time series data, such as 30-minute or hourly storm event distribution of precipitation, the ability to generate weather data at a specified time step may be an important consideration.

Weather generators have been developed in recent years to help reduce the time required to prepare weather input data sets. Weather generators are computer programs that use existing weather records to produce a long series of synthetic daily climatic data. The statistical properties of the generated data are expected to be similar to those of the actual data for a station. Unlike historical weather data files which may be missing data due to equipment servicing or malfunction, generated weather input provides a complete data set and can be produced for any desired period of time, enhancing their use as input for continuous-in-time models.

One of the first weather generators developed for rural water quality modelling purposes is called WGEN (Richardson and Wright 1984). Numerous other weather generators have developed since then. CLIGEN, the weather generator incorporated within the WEPP (Water Erosion Prediction Project) model, is based on the weather generation methods used in WGEN (Nicks et al. 1990). CLIGEN, however, adds the capability of generating rainfall intensity and duration or "breakpoint" rainfall data necessary for the Green and Ampt infiltration model used in many of today's hydrologic and soil loss prediction models including WEPP. Other weather generators available include USCLIMATE (Johnson et al. 1996), CLIMAK (Danuso et al. 1997), and ClimGen (Stöckle et al. 1999).

Statistical parameterization of weather station datasets for use in any climate generator is a time-consuming exercise and requires access to a significant amount of historical data for a station. Weather data parameterization to-date has largely been limited to selected stations within the continental U.S. In undertaking a review of weather generator use in Ontario, Canada, only one piece of commercial software was identified as having completed weather station data analysis and parameterization for the purposes of applying a weather generator. This software called Visual HELP, assembled by Waterloo Hydrogeologic Inc. (2001), uses the weather generator WGEN and includes parameterization results for five Ontario, Canada stations. The software, however, does not include a convenient tool to allow users to parameterize historical data from any other climate station of interest beyond the five included with the software even if sufficient historical data are available for the site. Also, the WGEN software included with Visual HELP does not have the potential to generate breakpoint rainfall data.

The weather generator called ClimGen was selected for use and testing in a Canadian setting. A literature review and model testing exercise identified the advantages to using ClimGen over other available weather generators

- ClimGen includes routines that allow it to automate the task of parameterizing historical weather data from new stations of interest. All that is required is sufficient historical observed data for a station formatted in a manner that can be read by the ClimGen software.
- Daily precipitation amounts are assumed to follow a Weibull distribution in ClimGen. Selker and Haith (1990) showed the Weibull distribution to be superior to other probability distributions of daily precipitation amount.
- A spline-fitting approach is used in ClimGen. This is an improvement over the one-term Fourier series used by many of the other weather generators to model seasonal variations in climate data.
- An option to estimate solar radiation values using a temperature-based approach, developed by Bristow and Campbell (1984) is embedded in ClimGen to assist users in areas where measured daily solar radiation input data are not readily available. Only three southern Ontario Environment Canada climate stations have routinely collected historical solar radiation data.
- ClimGen can generate breakpoint precipitation in 30-minute intervals.
- Developers of the ClimGen model are available and were interested in supporting and enhancing ClimGen as a weather generation tool for use under a range of geographic locations and applications, including the Canadian setting.

A DESCRIPTION OF CLIMGEN

ClimGen uses weather generation approaches that are similar to those applied in other popular weather generators. The following sections provide background on the techniques used in ClimGen to generate the key weather input variables needed for hydrologic modelling.

Precipitation

Generating precipitation data involves approaches that can assess the likelihood of both the occurrence of precipitation on a particular day as well as the amount. Rainfall intensity and duration within the rain event may also need to be generated for some applications.

Daily Precipitation Occurrence ClimGen models daily precipitation occurrence using a two-state Markov chain model to generate the number and distribution of precipitation events. The probability of a wet day following a dry day is defined in ClimGen as α, and the probability of a dry day following a wet day is defined as β. The two-state Markov chain for the combination of conditional probabilities as described by Nicks et al. (1990) is as follows:

$$P(W|D) = \alpha$$
$$P(D|D) = 1 - \alpha$$
$$P(D|W) = \beta$$
$$P(W|W) = 1 - \beta$$

where: $P(W|D)$ is the probability of a wet day given a previous dry day,
$P(D|D)$ is the probability of a dry day given a previous dry day,
$P(D|W)$ is the probability of a dry day given a previous wet day,
$P(W|W)$ is the probability of a wet day given a previous wet day.

These probabilities are calculated for each month of the year for the station being characterized by analyzing a station's historic long-term precipitation data provided by the model user.

Amount of Daily Precipitation On days when precipitation is determined to occur, ClimGen assumes the cumulative probability of precipitation amount follows a Weibull distribution as described by Selker and Haith (1990). These researchers showed the Weibull distribution to be superior to other distributions for 33 U.S. climate stations located east of the Rocky Mountains. In the case of a wet day, use of the Weibull distribution function to generate a precipitation amount in ClimGen is as follows

$$F(P) = 1 - \exp\left(-\left(\frac{P}{\beta}\right)^{\alpha}\right) \quad (1)$$

where: $F(P)$ is the cumulative probability of a precipitation amount equal or less than P, and α and β are parameters of the distribution function that are calculated on a monthly basis.

This distribution is samples for each precipitation event using the inverse method as follows:

$$P = \beta[-\ln(r)]^{1/\alpha}$$

where: r is a uniform random number between 0 and 1

Both the Markov chain transition probabilities and the precipitation distribution parameter(s) were found to vary with season in most areas (Richardson 1981). To address this, Richardson (1981) suggested a Fourier series or some other cyclical model to describe the periodic variation of these parameters. ClimGen, however, takes advantage of today's computing power to develop a spline function that is unique to each site and that was found to be more accurate than a Fourier series in mathematically representing the monthly and seasonal changes in the distribution parameters.

Intensities and Duration of Daily Precipitation Amounts The subroutines used in ClimGen to generate breakpoint precipitation data are based on a method described by Arnold and Williams (1989). This method generates ½ hour precipitation intensities using the assumption that precipitation amounts within a storm are exponentially distributed (Arnold and Williams 1989). They tested this assumption at five locations across the U.S. by comparing measured runoff with simulated runoff volumes and found it gave reasonable results. While this approach is used in the Simulator for Water Resources in Rural Basins (SWRRB) and Soil and Water Assessment Tool (SWAT) models, Arnold and Williams (1989) admit that the approach was only tested to a limited extent. It is also important to note that runoff events associated with snowmelt were not considered in their testing.

To apply the approach of Arnold and Williams (1989) the mean monthly maximum fraction of total precipitation that occurs during a ½ hour period within a storm event ($\alpha_{0.5m}$) needs to be entered into ClimGen. This can be calculated from historical station records of storm intensity. ClimGen uses these monthly values to calculate the peak ½ hour intensity for a storm event of total precipitation (R) as follows:

$$R_{0.5peak} = \alpha_{0.5m} * R \tag{2}$$

where: $R_{0.5peak}$ is the ½ hour peak precipitation amount in mm.

The number of 30-minute intervals that combine to form a precipitation event varies depending on the generated storm duration. In ClimGen, storm duration (DUR) is allowed to vary from 0.5 to 24 hours as a function of $\alpha_{0.5m}$ as follows (Arnold and Williams 1989):

$$DUR = 1440 \times \left(1 - \exp\left(\frac{-0.3}{\alpha_{0.5}}\right)\right) \tag{3}$$

The total daily precipitation generated over the storm duration is partitioned into 30-minute intervals by first selecting the point at which the maximum ½ hour intensity occurs. Arnold and Williams (1989) showed that the peak ½ hour intensity location tends to occur nearer to the beginning of the storm. They arrived at a distribution function that declines linearly from 0 to 0.5 and is constant from 0.5 to 1

With the location for the peak ½ hour intensity determined through querying this distribution with a random number, previous and subsequent 30-minute precipitation amounts are assigned outward from that point until the total amount of daily precipitation has occurred. The precipitation amounts for the individual 30-minute intervals are determined using the Weibull distribution and are related to the previous interval amount using the following equation:

$$R_{0.5i+1} = \beta R_{0.5i} + (1-\beta) R_{0.5g} \tag{4}$$

where: $R_{0.5g}$ is the ½ hour rainfall amount generated from the Weibull distribution (mm),
i is the subscript denoting the within storm location.
β is the within-storm lag coefficient

The relationship given by Eq. 4 is applied because Arnold and Williams (1989) noticed a tendency for memory within the storm structures.

Air Temperature and Solar Radiation

Meteorological elements such as air temperature and solar radiation are not as difficult as precipitation to model statistically because the high proportion of observations with values of "zero", common with daily precipitation records, is not associated with these climate observations (Richardson 1981). There is, however, some dependency between precipitation and temperature and solar radiation that needs to be accounted for in weather generation. Analysis of measured daily air temperatures relative to the occurrence of precipitation showed that both wet days and wet days which follow a wet day tend to have lower mean temperatures than dry days or dry days which follow a dry day (Richardson 1981). Similarly, solar radiation levels have been observed to be lower on a wet day, due to increased cloud cover, compared to a dry day when clear skies are more probable (Richardson 1982).

After recognizing that on a given day both air temperature and solar radiation depended on whether there was any precipitation, Richardson (1981) developed a generation approach which assumed the daily maximum and minimum temperature and the level of solar radiation would occur as part of a continuous multivariate stochastic process. This technique is used in

WGEN, CLIGEN, CLI90 as well as ClimGen. As a result of applying this method, the daily means and standard deviations of the cumulative temperature and solar radiation distributions are influenced by whether the day is a wet day or dry day (Richardson 1981). In WGEN, CLIGEN and CLI90, a Fourier series is used to smooth seasonal variations in means and standard deviations. ClimGen, however, uses a spline-fitting procedure to adjust for variations in season means and standard deviations. Similar to other weather generation tools, the residual elements of the function are obtained by removing the periodic mean and standard deviation using the following equations (Richardson 1981):

$$\chi_{p,i}(j) = \frac{X_{p,i}(j) - \overline{X}_i^o(j)}{\sigma_i^o(j)} \qquad \text{for} \qquad Y_{p.i} = 0 \tag{5}$$

OR

$$\chi_{p,i}(j) = \frac{X_{p,i}(j) - \overline{X}_i^1(j)}{\sigma_i^1(j)} \qquad \text{for} \qquad Y_{p.i} > 0 \tag{6}$$

where: $\overline{X}_i^o(j)$ and $\sigma_i^o(j)$ are the mean and standard deviation for a dry day ($Y_{p,i} = 0$),
$\overline{X}_i^1(j)$ and $\sigma_i^1(j)$ are the mean and standard deviation for a wet day ($Y_{p,i} > 0$),
$\chi_{p,i}(j)$ is the residual component for the variable *j* (i.e. maximum temperature (*j*=1), minimum temperature (*j*=2), or solar radiation (*j*=3)).

Removing the mean and standard deviation from the series of observed temperature and solar radiation values leaves a set of residual values ($\chi_{p,i}(j)$) that have a mean of zero and a standard deviation of one. Not only are these residual values for maximum and minimum air temperature and solar radiation dependent on the precipitation status of the day, they are also interdependent of each other. As a result, a matrix of serial correlation and cross-correlation coefficients are calculated and used to mathematically describe this time dependence and interdependence.

The above technique can be used to generate air temperature and solar radiation if historical data are available to develop the various matrices of correlation coefficients. Long-term observed solar radiation data, however, is rarely available for a site. Recognizing this, ClimGen developers also included a solar radiation estimation technique developed by Bristow and Campbell (1984) for estimating daily solar radiation data from a station's air temperature data. Bristow and Campbell (1984) showed that the difference between maximum and minimum daily temperatures has a high correlation with the amount of solar radiation received on any day. Overcast days tend to have a smaller difference between maximum and minimum temperature (ΔT) for the day. Conversely, on days with clear skies, the temperature extremes are relatively large. Bristol and Campbell (1984) found that these temperature extremes could be related to the atmospheric transmission of solar radiation, and if the extraterrestrial irradiance was known, solar radiation estimates could be calculated from temperature data. ClimGen developers have found, however, that generation of solar radiation from correlated long-term historical data produces better solar radiation generation results than estimation approaches.

Wind Speed

Wind speed is generated by ClimGen without any correlation with other variables. Similar to daily precipitation, average daily wind speed (U) is represented using a Weibull distribution as follows:

$$F_m(U) = 1 - \exp\left(-\left(\frac{U}{\beta}\right)^{\alpha}\right) \tag{7}$$

where: $F_m(U)$ is the cumulative probability distribution of average daily wind speed in month,
β is a scale parameter determined from the observed wind data,

α is a shape parameter determined from the observed wind data

This distribution is sampled for each day of weather generation using the inverse method:

$$U = \beta[-\ln(r)]^{1/\alpha} \tag{8}$$

where: r is a random number ranging between 0 and 1

Air Humidity

To parameterize station data for the purposes of generating air humidity data, ClimGen begins by using measured minimum and maximum relative humidity (RH_{min} and RH_{max}), and corresponding measured maximum and minimum air temperature data (T_{max} and T_{min}) to calculate the daytime and nighttime dew point temperatures (T_{dd} and T_{dn}). This is achieved as follows:

$$e^0(T_{dd}) = e^0(T_{max}) \frac{RH_{min}}{100} \tag{9}$$

$$e^0(T_{dn}) = e^0(T_{min}) \frac{RH_{max}}{100} \tag{10}$$

where: $e^0(T)$ is the saturation vapour pressure (kPa) determined at specified temperature (T) (˚C).
Dew-point temperatures are obtained by inverting Eq. 11.

$$e^0(T) = 0.6108 \exp\left[\frac{17.27T}{T + 237.3}\right] \tag{11}$$

A linear regression between daytime and nighttime dew point temperatures is calculated during the parameter optimization phase of ClimGen.

Once the above step is completed, ClimGen then calculates from the historical daily data the daily maximum vapour pressure deficit (VPD_{max}). This maximum difference between the saturation vapour pressure and the actual vapour pressure is typically obtained at the time of minimum relative humidity and maximum temperature as follows (Nelson 2002):

$$VPD_{max} = e^0(T_x)\left(1 - \frac{RH_{min}}{100}\right) \qquad (12)$$

The daily maximum vapour pressure deficit can also be estimated from temperature using the following relationship:

$$VPD_{max} = \frac{e^0(T_{max}) - e^0(T_{min})}{1 - a\left[e^0(T_{max}) - e^0(T_{min})\right]} \qquad (13)$$

where: a is the aridity factor

The aridity factor is optimized in ClimGen by combining Eqs. 12 and 13. Once the aridity factor is optimized, VPD_{max} from Eqs. 11 and 12 are correlated through linear regression. ClimGen developers suggest a minimum of two years of daily humidity and air temperature data is sufficient to parameterize ClimGen for air humidity generation (Stöckle et al. 1999).

To generate air humidity data, ClimGen uses Eq. 13 to calculate VPD_{max}. It also uses the optimized aridity factor for the station, the linear regression slope and intercept values as well as generated T_{max} and T_{min} values for the day. It then uses the calculated values for VPD_{max} in Eq. 12 to determine RH_{min}. Once a value for RH_{min} has been calculated, Eq 9, the linear regression results between daytime and nighttime dew point temperatures, and Eq. 10 can be used to arrive at a value for RH_{max}.

APPLYING CLIMGEN IN SOUTHERN ONTARIO

Stöckle et al. (1999) recommend that a minimum of 20 years, and preferably 30 to 35 years of measured historical meteorological data, be available as the basis for completing the parameterization of many of the meteorological weather elements to be produced by a weather generator. Given this requirement for a range of meteorological measurements together with a long period of continuous recording, the number of Ontario (Environment Canada) stations that could potentially be used for parameterization is significantly reduced. This section describes how suitable stations for testing ClimGen's performance in southern Ontario were identified and summarizes the results obtained.

Selection of Southern Ontario Test Stations

Information provided in Environment Canada's Ontario Climatic Station Catalogue, (Environment Canada – Atmospheric Environment Service 1989) was used to identify Environment Canada weather stations best suited to weather generation. Only three southern Ontario locations (Ottawa, Toronto, and Guelph) were identified as routinely collecting all of the necessary weather data. However, because ClimGen can estimate solar radiation from daily precipitation and temperature data, by including stations that only lack this record, the number of Environment Canada stations suitable for parameterization increased to eleven. These eleven stations are reasonably well distributed across southern Ontario (Figure 1). However, when the main climate regions of southern Ontario as defined by Brown et al. (1980) are considered, these eleven stations do not provide full coverage. A second set of stations indicated by a "triangle" symbol in Figure 1 include stations missing solar radiation plus one of the other preferred measurements - typically relative humidity, wind speed, or precipitation intensities. These ten stations could be combined with the other eleven stations to provide more complete coverage of the major climate and crop production regions of Ontario.

Documentation for the weather generators WGEN and CLIGEN were reviewed to determine the number of stations whose data were parameterized to provide a full coverage weather generation tool for the American states. The approach American researchers used in identifying stations for CLIGEN parameterization was to have one parameterized station for each grid of one degree latitude by one degree longitude. If similar criteria were used for southern Ontario, twelve to thirteen stations would be appropriate.

To fully parameterize all of the climate stations identified in Figure 1 exceeded the time available in this study. A sub-set of these stations was selected to test ClimGen's ability to generate southern Ontario weather data. A station's geographic location as well as the range and form of available historical data were key factors considered when selecting the test stations. Figure 2 shows the set of pilot weather stations that were selected. Of the stations chosen, only one station (Ottawa) possessed the complete data set needed. The Guelph station record contains all the necessary weather measurements however, long-term maximum and minimum relative humidity and rainfall intensity data were not in a useable format for the ClimGen model. To compensate, rainfall intensity data were obtained from the nearby Waterloo-Wellington station, while historical relative humidity data were obtained from the Elora research station. The London and Windsor stations were missing solar radiation data only. The Harrow station was missing long-term maximum/minimum relative humidity data as well as solar radiation data. Solar radiation data for Harrow, however, were provided with the data set obtained from the University of Guelph's Land Resource Science Department (Parkin et al. 1999). These data are based on readings from the nearby Detroit, Michigan station. Finally, the Stratford site was missing maximum and minimum relative humidity, wind speed, and solar radiation data.

Parameterizing Southern Ontario Test Station Weather Data Sets

Climate data not already compiled by Parkin et al. (1999) were acquired from Environment Canada's Atmospheric Environment Service for the six test stations identified in Figure 2. Thirty-five or more years of historical weather were purchased, beginning in 1955, and including daily precipitation, 30-minute rainfall intensity, maximum and minimum temperature, maximum and minimum relative humidity, and hourly wind speed data. Table 1 summarizes the period and origin of data that in the end was used as the basis for parameterization of the test station data sets through ClimGen.

The format of the data provided by both Environment Canada and through the University of Guelph’s Land Resource Science Department did not match the format that could be read directly by ClimGen. A significant amount of time was spent reformatting and converting the data to the units expected by ClimGen. To make future applications of ClimGen easier in Canada, a format conversion tool was added to ClimGen (versions 3.8 and 4.1.05). It converts raw data received from Environment Canada automatically into ClimGen's universal environment database (UED) format.

To enable ClimGen to generate breakpoint precipitation data, an analysis of the 30-minute rainfall intensity data was completed for each pilot site to determine a value for the mean monthly peak ½ hour fraction of total rainfall. Table 2 summarizes the results of this analysis. Note that the 30-minute intensities for the Guelph station were derived from the nearby Waterloo-Wellington airport station data. Also note that not all stations had a full set of data for the winter months because precipitation typically comes as snowfall in those months and intensities are not recorded. No intensity data were available for Harrow for the months of December through March. Values for the mean peak ½ hour fraction of total rainfall for these months, as shown in Table 2, were assumed to equal the values calculated for the Windsor station. No precipitation intensity data were available to calculate a rainfall fraction value for February in Ottawa. It was estimated to be 0.10 after reviewing the results from the other stations for the same month. The rainfall fraction values presented in Table 2 tend to show good consistency within months among stations.

EVALUATION OF CLIMGEN'S WEATHER GENERATION PERFORMANCE

Weather data were generated for each of the six Ontario test stations using ClimGen (version 3.8). An approach similar to that described by Nicks et al. (1990) was used to evaluate the generated data sets. For each of the six pilot weather stations parameterized, ten unique ten-year simulations of daily weather data were completed. Monthly means of the weather data obtained from the ten simulation runs were computed. These monthly means were then compared with the monthly means calculated from the actual **historical** data used to derive the station's weather generation parameters. The generated monthly means were also compared with the station's monthly 30-year **normal** means as published by Environment Canada. This analysis comparing the historical data set means and the published 30-year normal means with the ClimGen-generated means was undertaken for all weather variables generated. The analysis involved a two-sided t-test at a 0.05 probability level (Nicks et al. 1990). Table 3 summarizes the outcome of the series of statistical comparisons for all the test stations.

A review of Table 3 suggests that ClimGen performs as well as, or better than WGEN in simulating the range of monthly average precipitation and temperature values at the test stations. The best evidence of this comes when one compares the statistical analysis results from WGEN precipitation output for the London station with the corresponding statistical results using data generated by ClimGen for the same station. There was no significant difference identified between the generated values and historical or normal values for any month when ClimGen was used. With WGEN, however, there was one month when comparing the historical data set with generated data and two months when comparing published normals to the generated data where WGEN gave significantly different monthly precipitation amounts. Similarly, if Detroit's WGEN and Windsor's ClimGen summaries are compared the ClimGen model appears to perform better, particularly in the area of temperature simulation (See Table 3). Solar radiation data for Windsor, were not available to allow a comparison of the ability of the two models to properly generate this weather measurement.

The monthly average daily solar radiation data generated by ClimGen did not match particularly well with observed data. When generating the data, ClimGen was applied in a manner that forced it to generate solar radiation data from temperature and precipitation data (refer to Air Temperature and Solar Radiation section). This was done to see how well this method matches measured solar radiation data. Perhaps the model would have performed better if it had been run in the mode that calculates the residual adjustment parameters for solar radiation. With so few climate stations possessing long-term solar radiation data, however, it is often impossible to apply this option for a location.

Nicks et al. (1990) did not test WGEN for its ability to estimate monthly average relative humidity and wind speed values. In this study, ClimGen was found to do a good job of predicting average daily wind speeds but a rather poor job of predicting average values of relative humidity (RH). The model's poor performance for RH may have been due to the way the statistical evaluation was performed. ClimGen generates both the maximum and minimum RH for a day. For the statistical evaluation the average RH for the day was required to enable a comparison with other weather generators. The average RH was calculated as the average of the maximum and minimum RH for the day assuming RH values change linearly over the course of a day. This is not necessarily the case so a test was conducted to assess the degree of error possibly associated with this assumption. Detailed 15-minute RH data from Elora were used to determine if simply averaging the maximum and minimum RH values for a day would give a significantly different average value from that obtained by averaging all of the 15 minute values measured over the course of the same day. Results showed that on some days the average RH was significantly different from the average calculated using the maximum and minimum RH for the day. However, when the generated average RH values were compared to the adjusted average RH values for the Elora station, it did little to improve ClimGen's statistical score.

Table 3 shows that wind speed data generated by ClimGen did not compare well with the published normal wind speeds for Ottawa. It is important to note, however, that the published normal wind speed data came from Ottawa's airport weather station and not the Ottawa NRC weather station used to parameterize the location. Published normal data for the Ottawa NRC site were not available so the next closest station with published normals was chosen when comparing published normals to generated means. Historical data matched well with the generated data.

Weather generation tools have been criticized for their inability to generate the extreme conditions, particularly precipitation patterns, over a generation period. A simple analysis of the generated data was performed to determine if ClimGen overcomes some of this criticism through the use of both the Weibull distribution function (Selker and Haith 1990) and a spline function to represent the fluctuations among seasons. The analysis consisted of identifying the top and bottom 3%, 6%, 9%, and 12% of annual precipitation levels generated over 100 years and comparing this with the same return periods observed in the actual historical data set used to develop the parameterization values for the weather station. The difference between the actual data and the measured data was used to determine the percent error in the generated data. A summary of the results of this analysis is presented in Table 4. Note that the same analysis was completed on the Detroit WGEN data using 4%, 8%, and 12% breakpoints in extremes because only 25 years of historical weather data were available for Detroit.

Overall, the results presented in Table 4 suggest that ClimGen's algorithms are perhaps slightly better than those used in WGEN. However, for the London station, WGEN was actually better than ClimGen in generating the upper rainfall extremes. Both of the models performed poorly at estimating the lower rainfall extremes.

For most sites, ClimGen underestimated both extremes as indicated by the negative percentage difference between generated and observed data for the upper extremes and the positive percentage difference between generated and observed data for the lower extremes. In most cases, however, the differences were often less than five percent. ClimGen seemed to perform particularly well on the Ottawa data set.

USING CLIMGEN OUTPUT IN CONTINUOUS-IN-TIME RURAL WATER QUALITY MODELS

The analysis of ClimGen output undertaken in this investigation suggests that ClimGen can prepare weather input files for continuous-in-time rural water quality models that are representative of historical weather data. ClimGen was also found to perform at least as well as one other popular weather generation tool, WGEN. A key advantage of ClimGen over other available weather generation tools is that it also automates the process of parameterizing historical weather station data for weather generation purposes. As well, ClimGen (version 4.1.05) includes the capability to read archival data files for any Environment Canada Station, further expediting the parameterization process for Canadian locations where adequate historical weather data exists.

Generated weather data output provided by ClimGen is routinely stored in a universal environment database (UED) format. Few rural water quality models can directly read this file format. ClimGen, however, includes the option of exporting the generated data into a text format, allowing users to manipulate the weather files using common software tools to conform to their particular model requirements. It also includes an option to export the generated daily and breakpoint data into files. ClimGen can format these files in a manner that can be read directly by Root Zone Water Quality Model (RZWQM) and the CropSyst model. For other models employing the Green and Ampt infiltration model, manipulation of the breakpoint precipitation files in standard text manipulating software (e.g. MS Excel spreadsheet) may be necessary to match the model's input requirements.

A copy of the most recent version of ClimGen is available and may be downloaded from the ClimGen website. The website address is www.bsyse.wsu.edu/climgen. Those interested in using the model are encouraged to visit the site, download and test the model in their own setting, and provide feedback to the developers on its use.

ACKNOWLEDGEMENTS

The authors wish to express their appreciation and thanks to Dr. Roger Nelson, Biological Systems Engineering Department, Washington State University, Pullman, WA, USA. His familiarity with the ClimGen model, computer programming skills, and willingness to assist in further developing ClimGen for use as a weather generation tool for Canadian climate stations were of great value in undertaking this assignment.

REFERENCES

- Arnold, J.G. and J.R. Williams. 1989. Stochastic generation of internal storm structure at a point. *Transactions of the ASAE*. 32(1): 161-167.
- Bristow, K.I. and G.S. Campbell. 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agriculture and Forest Meteorology*. 31: 159-166.
- Brown, D.M., G.A. McKay, and L.J. Chapman. 1980. *The Climate of Southern Ontario*. Toronto, ON: Minister of Supply and Services Canada.
- Danuso, F. et al. 1997. CLIMAK reference manual. DPVTA, University of Udine, Italy.

- Environment Canada – Atmospheric Environment Service. 1989. *Climatological Station Catalogue, Ontario*. Downsview, ON: Environment Canada – Atmospheric Environment Service.
- Johnson, G.L., C.L. Hanson, S.P. Hardegree, and E.B. Ballard. 1996. Stochastic weather simulation: overview and analysis of two commonly used models. Journal of Applied Meteorology. 35: 1878-1896.
- Nelson, R. 2002. Description of ClimGen, a Weather Generation Program. http://www.bsyse.wsu.edu/climgen/documentation/description.htm. Accessed May 20, 2003.
- Nicks, A.D., C.W. Richardson, and J.R. Williams. 1990. Evaluation of the EPIC model weather generator. In *EPIC – Erosion/Productivity Impact Calculator, 1. Model Documentation, USDA Technical Bulletin No. 1768*, Eds. A.N. Sharpley and J.R. Williams. 105-124. Washington, DC, U.S.A.: Government Printing Office
- Parkin, G.W., C. Wagner-Riddle, D.J. Fallow, and D.M. Brown. 1999. Estimated seasonal and annual water surplus in Ontario. *Canadian Water Resources Journal.* 24(4): 277-292.
- Richardson, C.W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*. 17:182-190.
- Richardson, C.W. 1982. Dependence structure of daily temperature and solar radiation. *Transactions of the ASAE.* 25: 735-739.
- Richardson, C.W. and D.A. Wright. 1984. *WGEN, A Model for Generating Daily Weather Variables, USDA ARS Bulletin No. ARS-8.* Washington, DC, U.S.A.: Government Printing Office.
- Selker, J.S. and D.A. Haith. 1990. Development and testing of single-parameter precipitation distributions. *Water Resources Research*. 26(11): 2733-2740.
- Stöckle, C.O., G.S. Campbell, and R. Nelson. 1999. ClimGen manual. Biological Systems Engineering Department, Washington State University, Pullman, WA.
- Waterloo Hydrogeologic Inc. 2001. Winter/Spring 2001 Software and Services Catalog. Waterloo, ON: Waterloo Hydrogeologic Inc.

Table 1 A summary of the source of historical weather data input used to parameterize the southern Ontario test stations

Climate Variable	Source and Period of Record [1]					
	Harrow	*Windsor*	*London*	*Stratford*	*Guelph*	*Ottawa*
Daily precipitation	LRS 1955-1989	EC 1955-1989	EC 1955-1996	EC 1960-1996	LRS 1955-1989	LRS 1955-1989
30-minute Rainfall Intensities	EC 1967-1989	EC 1960-1989	EC 1960-1996	EC 1966-1996	EC (Wat-Well)[2] 1970-1989	EC 1960-1989
Maximum Temperature	LRS 1955-1989	EC 1955-1989	EC 1955-1996	EC 1960-1996	LRS 1955-1989	LRS 1955-1989
Minimum Temperature	LRS 1955-1989	EC 1955-1989	EC 1955-1996	EC 1960-1996	LRS 1955-1989	LRS 1955-1989
Maximum Relative Humidity	EC (Windsor) 1955-1971 1982-1989	EC 1955-1971 1982-1989	EC 1955-1996	EC (London) 1960-1996	LRS (web) (Elora) 1990-1999	EC 1955-1969 1977-1989
Minimum Relative Humidity	EC (Windsor) 1955-1971 1982-1989	EC 1955-1971 1982-1989	EC 1955-1996	EC (London) 1960-1996	LRS (web) (Elora) 1990-1999	EC 1955-1969 1977-1989
Wind Speed	LRS 1955-1989	EC 1955-1989	EC 1955-1996	EC (London) 1960-1996	LRS 1955-1989	LRS 1955-1989
Solar Radiation	LRS (Detroit) 1955-1989	ClimGen Estimated	ClimGen Estimated	ClimGen Estimated	LRS 1955-1989	LRS 1955-1989

Notes [1] EC = Environment Canada, LRS = Parkin et al., (1999), (web) = data from University of Guelph, Land Resource Science (LRS) website
[2] Wat-Well = Waterloo-Wellington airport weather station.

Table 2: Mean peak half hour rainfall fractions for test stations

Test Station	Month											
	JA	**FE**	**MR**	**AP**	**MA**	**JN**	**JL**	**AU**	**SE**	**OC**	**NO**	**DE**
Harrow	0.07	0.12	0.27	0.34	0.45	0.55	0.56	0.58	0.50	0.42	0.35	0.16
Windsor	0.07	0.12	0.27	0.36	0.47	0.58	0.58	0.67	0.58	0.40	0.35	0.16
London	0.05	0.08	0.14	0.30	0.47	0.65	0.61	0.58	0.69	0.38	0.23	0.08
Stratford	0.05	0.02	0.1	0.33	0.47	0.50	0.58	0.55	0.53	0.37	0.17	0.06
Guelph	0.01	0.01	0.16	0.31	0.46	0.53	0.54	0.59	0.45	0.43	0.17	0.16
Ottawa	0.04	0.10	0.15	0.31	0.42	0.55	0.61	0.57	0.52	0.40	0.13	0.08

Table 3 **Summary of results of the statistical comparison of generated Ontario weather station data with historical and published normal data[1]**

Test Station	Number of Months the Generated Mean is Significantly Different from the Historical Data Mean and Published Normal Mean									
	Precipitation		Temperature		Solar Radiation.		Relative Humidity		Wind Speed	
	Hist[4]	Nrm[5]	Hist[4]	Nrm[5]	Hist[4]	Nrm[5]	Hist[4]	Nrm[5]	Hist[4]	Nrm[5]
Harrow	0	0	0	0	0	-	-	-	0	-
Windsor	0	0	0	0	-	-	9	6	0	1
London	0	0	0	0	-	-	10	12	0	0
Stratford	0	0	0	0	-	-	-	-	-	-
Guelph	0	0	0	1	10	11	10	-	0	-
Ottawa	0	0	1	0	3	9	11	9	0	12
Detroit, MI WGEN[2]	0	0	3	6	0	11	-	-	-	-
London WGEN[3]	1	2	0	0	-	-	-	-	-	-

Notes: [1] p = 0.05
[2] Source: Nicks et al. (1990) – WGEN evaluation results, Detroit station.
[3] Results from WGEN model supplied with Visual HELP (Waterloo Hydrogeologic Inc. 2001)
[4] Refers to the comparison with the historical data, used to parameterize the station.
[5] Refers to the comparison with the published normals for the station.

Table 4 **Analysis of the ability of ClimGen to represent extreme annual precipitation patterns**

Station Name	Precipitation (mm)	Upper Extreme				Lower Extreme			
		3%	6%	9%	12%	3%	6%	9%	12%
Harrow	Measured	1179	1146	1120	1106	578	617	632	652
	Generated	1150	1117	1096	1083	604	639	661	679
	Error (%)	-2.45	-2.52	-2.08	-2.10	4.43	3.69	4.57	4.20
Windsor	Measured	1140	1127	1107	1096	581	597	614	635
	Generated	1113	1080	1057	1041	640	667	681	690
	Error (%)	-2.34	-4.19	-4.53	-5.05	10.15	11.71	10.88	8.81
London	Measured	1273	1238	1200	1179	535	605	636	678
	Generated	1217	1194	1165	1147	730	737	742	753
	Error (%)	-4.40	-3.60	-2.89	-2.65	36.48	21.79	16.69	10.95
Stratford	Measured	1366	1357	1349	1342	688	750	777	792
	Generated	1419	1377	1333	1304	821	852	869	882
	Error (%)	3.85	1.48	-1.2	-2.84	19.31	13.62	11.80	11.41
Guelph	Measured	1189	1156	1130	1106	663	695	709	718
	Generated	1093	1073	1056	1046	657	664	675	686
	Error (%)	-8.03	-7.23	-6.53	-5.44	-0.89	-4.49	-4.74	-4.55
Ottawa	Measured	1175	1136	1106	1090	653	674	687	694
	Generated	1172	1130	1090	1063	655	665	679	689
	Error (%)	-0.23	-0.53	-1.44	-2.44	0.23	-1.46	-1.21	-0.74
WGEN (Visual HELP)		3%	6%	9%	12%	3%	6%	9%	12%
London	Measured	1273	1238	1200	1179	535	605	636	678
	Generated	1285	1228	1196	1174	730	747	757	765
	Error (%)	1.01	-0.78	-0.35	-0.42	36.50	23.42	18.95	12.75
WGEN (EPIC)		4%	8%		12%	4%	8%		12%
Detroit	Measured	1109	1033		1004	534	556		590
	Generated	1198	1176		1158	684	694		703
	Error (%)	8.04	13.92		15.31	28.13	24.74		19.19

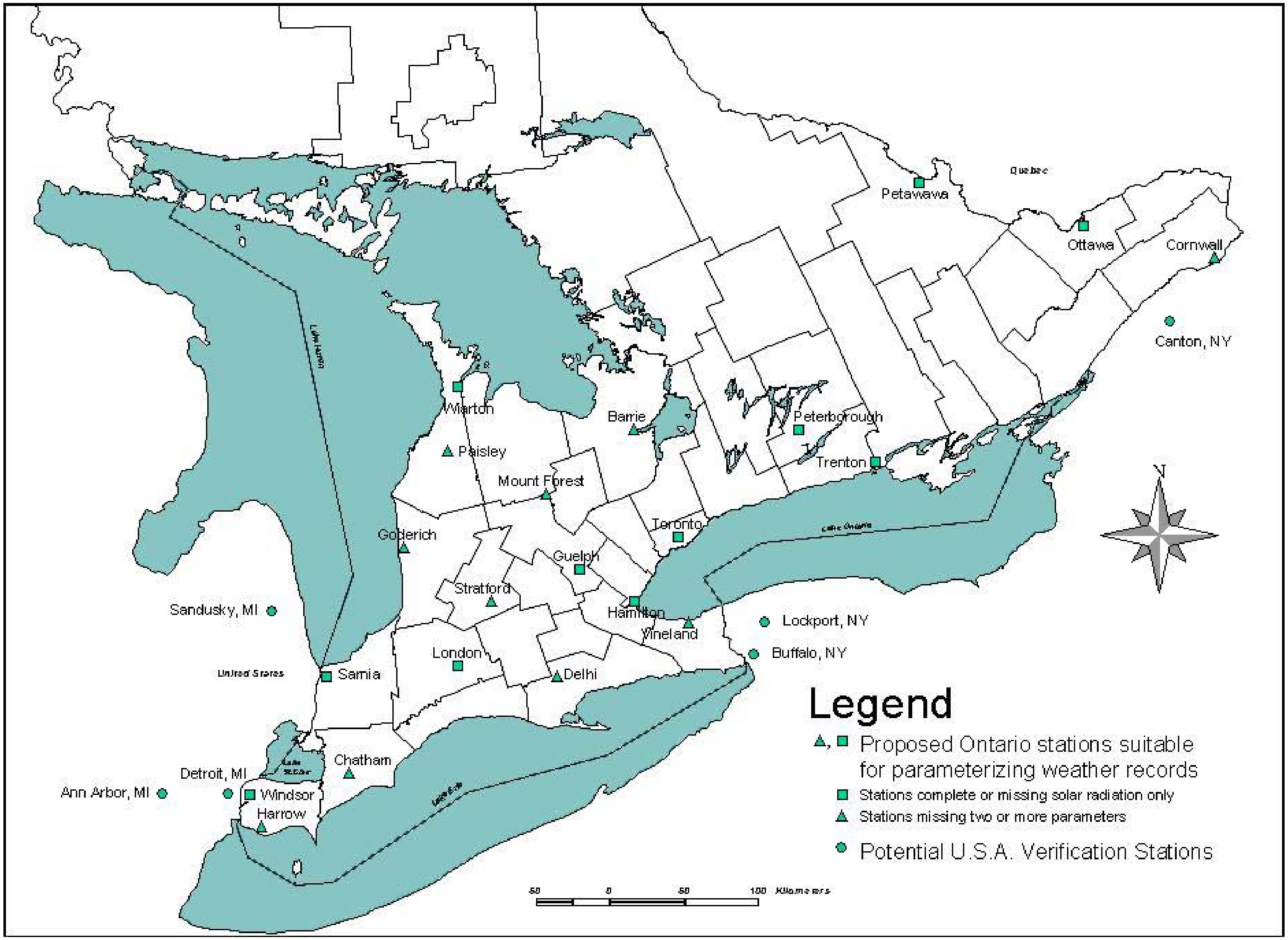

Figure 1 **Southern Ontario climate stations collecting the type and amount of data needed to apply weather generation techniques**

Figure 2 Southern Ontario pilot stations selected for testing ClimGen's weather generation techniques