

REGIONAL-SCALE TYPICAL METEOROLOGICAL YEARS FOR BUILDING STOCK MODELLING

Giulio Tonellato¹ and Michaël Kummert¹

¹ Polytechnique Montréal, Montréal, Québec, Canada
Department of Mechanical Engineering

ABSTRACT

Building stock simulation has evolved from attempting to reproduce yearly energy use to realistically representing the dynamic behaviour of the stock and its impact on peak electricity demand. As all building performance analyses, simulating the performance of a stock requires using weather data files, and “typical meteorological years” (TMY, or their variants TRY for typical reference year) are often used to obtain a synthetic data file representing an “average” year.

The different methods used to obtain representative data files (EN ISO 15927-4:2005, TMY2 and TMY3 data files in the USA, or CWEC files in Canada) have been developed with individual weather stations in mind. Representative months are individually selected for each weather station in order to represent a typical month (Morris, 2016; Pernigotto et al., 2014; Wilcox and Marion, 2008). Using these typical weather files to simulate the performance of a building stock across a large region (e.g. country, state or province) would not allow to investigate the overall dynamic behaviour of the stock, such as its impact on the peak electricity demand, because the selected months would not necessarily be consistent across the different weather stations.

This paper presents a methodology inspired by Huang (2020) to obtain population-weighted typical weather files which results in selecting the same reference months for a large set of weather stations, therefore allowing to model an entire building stock across a large geographical area with consistent weather patterns. The methodology is applied to Canada using the CWEEDS dataset provided by Environment Canada (ECC Canada, 2020).

INTRODUCTION

Worldwide efforts to fight global warming have resulted in a focus on energy efficiency standards, particularly in the building sector (UNFCCC, 2015). Since one of the main drivers of energy demand is the weather, meteorological data has been standardized over the years to represent the typical weather of a particular location so that energy efficiency measures for different buildings could be classified and categorized based on common grounds (Pernigotto et al., 2014). The most common adopted standard is the “typical year”, which is generated for a certain weather station, selecting typical months over a span of 15-30 years. Those typical years are typically referred to as “typical meteorological years” (TMY) in North America, and “test reference years” (TRY) in Europe. In Canada, the typical meteorological years published by Environment Canada follow the methodology developed in the USA with some variations on weighting factors that will be discussed below, and are referred to as “Canadian Weather for Energy Calculation” (CWECE) (Environment Canada, 2020).

When modelling buildings across a large region, every building is simulated using the closest weather station data. This ensures, using the province of Québec as an example, that buildings located in the city of Québec will be simulated with a typical weather file that better represents that location than, say, Montréal, the largest city in the province, which is more than 200 km away as the crow flies, and has about 500 heating degree days less than the city of Québec. However, with the “typical year” methodology, there is no guarantee that the typical months that are selected for the two stations (Québec city and Montréal in our example) are the same. So, the “typical year” simulations for the two cities can be considered together on an annual scale, but not on a monthly (or hourly) scale. For instance, the typical month of February for Montréal selected in the CWECE typical year is February 2011 (for the McTavish station), whereas for Québec City, the month of February 2002 appears in the CWECE typical year (using the international airport station). Their respective air external temperatures are shown in Figure 1.

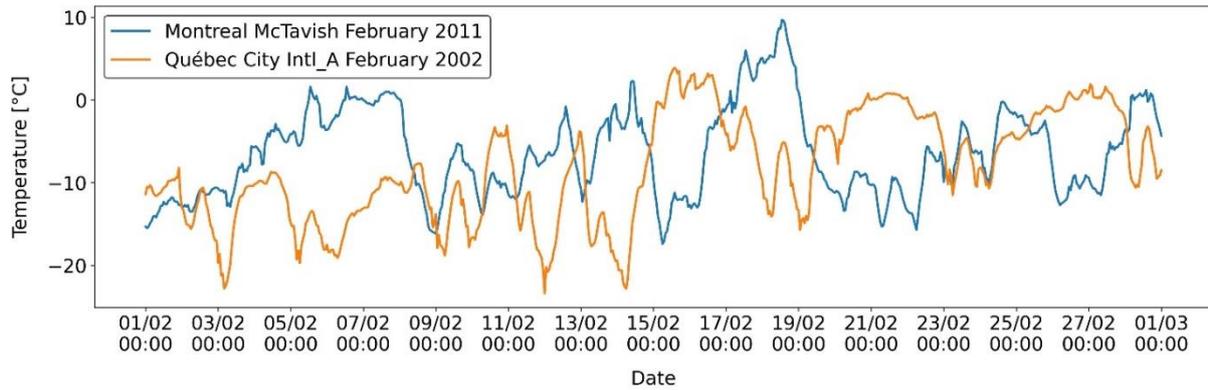


Figure 1 Air temperature registered at Montreal McTavish and Québec City International Airport stations for their respective “typical” February according to the CWEC weather file

As Figure 1 shows, using months of different years in the same simulation results in different hourly profiles and peak heating periods. In 2002, February 12 and 14 are very cold (according to Québec standards for February), while those days in are relatively mild in 2011 for Montréal.

For building stock simulations at the scale of the province, or at the scale of an electric utility (which in this case is the same), using different typical weather files for each station would result in an underestimated peak demand for the building stock, as cold periods would happen at different times for different locations. In reality, most areas within a province, state, or similar region typically experience relatively synchronized demand peaks, as weather patterns typically cover large areas. This is illustrated in Figure 2, which shows the air temperature of the same weather stations (Québec and Montréal) for February 2009 – it will be shown later that 2009 is the most “typical” month at the provincial scale. Comparing Figure 1 and Figure 2, it is clear that the combined heating demand peak of the two cities would be better modelled by using the same month for both locations.

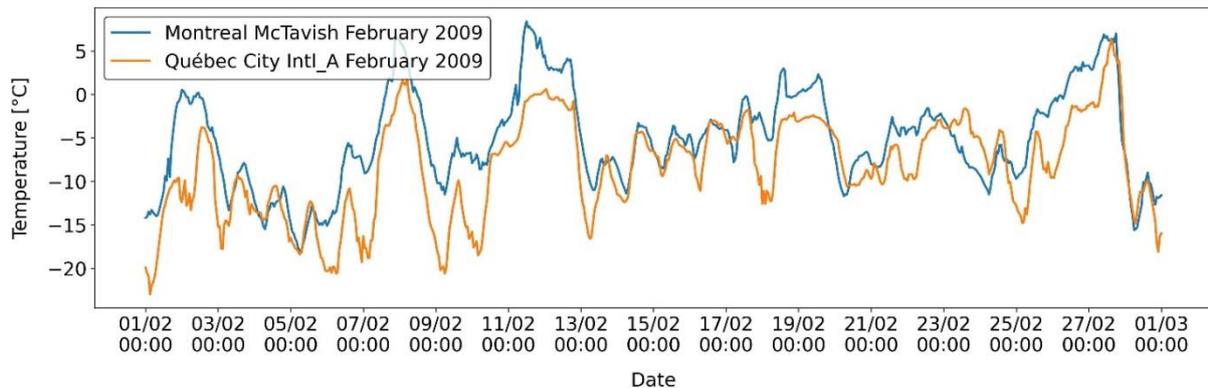


Figure 2 Air temperature registered at Montreal McTavish and Québec City International Airport stations for February 2009, the month selected by the “state-wide” approach

Therefore, a grid operator or any other stakeholder that wants to estimate demand peaks over a wide area would need a different method to handle weather data. To meet this need, Huang (2020) developed a “state-wide” typical year for California, which selects for each month the year that best matches the entirety of the weather stations, weighting each of them depending on how much population they represent. The work presented in this paper aims at demonstrating the applicability of this method to the province of Québec, Canada.

“Typical year” weather data

The “typical year” approach has been introduced by the U.S. National Climatic Data Center in 1976 (National Climatic Data Center, 1976) when they excluded years with extreme mean dry bulb mean temperatures until only one remained, calling it the “Test Reference Year” (TRY). Over the years the selection method changed, but the acronym TRY has been kept mostly in European regulations (Lund, 1996) with the standard being developed in Italy (Festa and Ratto, 1993), Denmark (Lund, 1995), Spain (Bilbao et al., 2004) and in the UK (Eames et al., 2016). The most known TRY standard is the EN ISO 15927-4:2005 (ISO, 2005), to which improvements have been suggested by Pernigotto et al. (2014).

In North America a similar year selection method to TRY has been developed with the name of “Typical Meteorological Year” (TMY) (Hall et al., 1978), later evolved in TMY2 to improve the solar models (Marion and Urban, 1995, p. 2) and in TMY3 to better manage the datasets (Wilcox and Marion, 2008). Both TRY and TMY are based on the so-called “Finkelstein-Schafer (FS) statistic” (Finkelstein and Schafer, 1971). This statistical method is used to select, for each month of the year, the most “typical month” over a certain range of years, typically between 15 and 30 years. The method consists in calculating the absolute difference between the cumulative distribution function (CDF) of a certain climate parameter in a certain day of the month and the CDF for the same parameter in the long-term period. This difference is often called “FS statistic” or “FS value”. The CDF of a certain parameter’s value “x” represents how many values are present in the dataset that are lower than “x”. The absolute differences of CDFs with the long-term are all summed up for each parameter of each month of each year. Some parameters will be weighted more than others, such as the dry bulb temperature over the wind speed. The month that will have the lowest weighted cumulative “FS value” will be the first candidate to be the most “representative month” for the location. Further examples to better understand this process can also be found in Huang (2020) (Section 9) and Pernigotto et al. (2014).

The difference between TRY and TMY consists in which climate input variables are considered and the weight of each of them. For instance, TMY considers more input variables and do not only include mean daily dry and wet bulb temperatures but also their maximum and minimum.

After 5 candidate months are selected for each calendar month, some so-called, “*persistence* criteria” can be applied to exclude the months that have long or numerous runs of consecutive days with either extreme temperatures or with extreme solar radiation (“extreme” being quantified by exceeding or being inferior to given percentiles in the long-term CDF). Many different persistence criteria exist and they are well described and compared in Pernigotto et al. (2014). The persistence criteria applied in this work will try to replicate the same procedures described in the Canadian CWEC files’ user manual (Morris, 2016).

The “state-wide” approach

As already introduced, Huang (2020) proposed a methodology to select a “typical year” not only for one single weather station but for a whole state, region or province. The principle is the same of the “typical year”: rating every month of every year of every weather station with an “FS cumulative value” and select, for each calendar month, the year with the lowest sum of “FS values” over all the weather stations. Since the final purpose is to have a realistic energy estimation, each weather station is weighted by the population it represents so that most populated areas, which have the largest impact on the energy use and peak demand, can play a major role in the weather selection.

In simpler words, the year selected for each calendar month will be the one that best matches the entirety of the weather stations, having a good “Finkelstein-Schafer (FS) statistic” for most populated weather stations. Further examples and tables will be shown to better understand this approach.

However, in the proposed methodology, it is unclear how the population was allocated for each weather station. In this work a method to do it has been developed.

METHODOLOGY

In Canada weather data has been stored, categorized and treated for over 30 years in the Canadian Weather Energy and Engineering Datasets (CWEEDS), publicly retrievable at (ECC Canada, 2020). The current dataset features 564 Canadian locations with hourly data starting from 1998 and it features, for all locations, satellite-based solar radiation (Morris, 2016). A meteorological reanalysis time series is used to fill data gaps longer than 3 hours while all the rest is filled with interpolation. Therefore, all the data filling and cleaning processes reported by Huang (2020) for California, are not required here to produce a “province-wide” typical year for Québec.

Since the statistical methods and the principles behind the “state-wide” approach are the same as for “typical years”, first TMY weather files will be reproduced for each of the weather stations that will be considered, and then a “province-wide” selection will be performed. Since the TMY approach has been already used to produce the CWEC typical year weather files (provided together with the CWEEDS), the statistical calculations for the “FS cumulative values” can be validated. In this way, also the province-wide selection will be relying on validated and reliable calculations.

TMY “Typical year” selection

The approach followed to calculate the “Finkelstein-Schafer (FS) statistic” for each month of each year is the same explained in the TMY3 report (Wilcox and Marion, 2008), in Pernigotto et al. (2014) and in Huang (2020). In order to replicate the CWEC weather files (Morris, 2016), the final cumulative FS values are estimated with the same weights for the same climate variables. Same as for the *persistence* criterions that are taken from the CWEEDS and CWEC files’ user manual (Morris, 2016).

To handle all the data and calculations the programming language *Python 3* is used through a *Jupyter notebook* (Kluyver et al., 2016), with the *Pandas* library (The Pandas development team, 2022) playing the major role as it can handle large dataframes with ease. The *Pvlib* library (Holmgren et al., 2018) is initially used to load the weather files in the *epw* format through the function `pvlib.iotools.read_epw` (Pvlib, 2022). Note that the CWEEDS weather files are provided in the WY3 format, to convert them to *epw* a MATLAB script from Siu and Liao (Siu and Liao, 2020) is used.

“Provincial-wide” typical year selection

To elaborate a “provincial-wide” typical year all the FS cumulative weighted values of each month of each year of each weather station are required. For each calendar month, the year having the lowest total FS statistics can be selected, but this value, in this case, is also weighted on each weather station depending on the population it represents.

The calculation steps in the “provincial-wide” selection are quite simple as the python script can just read the results collected in all other local weather stations’ “typical year” scripts. This is achieved using Microsoft Excel to collect the results and read them in Python all together. The most complex part is to understand which weather stations to select and how the population weights are calculated.

The population data can be obtained from the Canadian census data for each census subdivision in the province. Census subdivision (CSD) is the general term for municipalities (as determined by provincial/territorial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories). Municipal status is defined by laws in effect in each province and territory in Canada. In the 2021 dataset, there are 5161 CSD for Canada, and 1282 for Québec (StatsCan, 2022).

The population for Québec is 8.5 million, and the population density ranges from 0 to over 6000 people/km² for the census subdivisions – these are not equal in size nor in population, the largest by population being the city of Montréal with more than 1.76 million people and a density of 4830 people/km² and covering almost 365 km², and the largest in size being Rivière-Koksoak, with a reported population of zero. Figure 3 shows the census subdivisions in Québec with their population density.

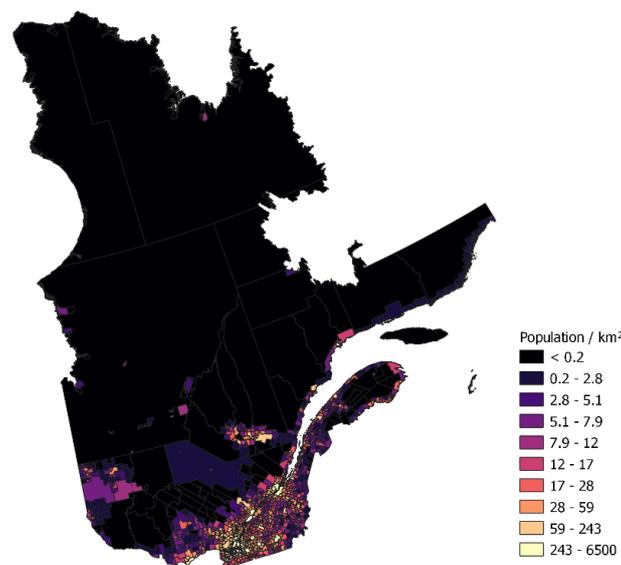


Figure 3 Census subdivisions in Québec with their population density

CWEEDS stations in Québec (or “near” Québec)

Canadian Weather Energy and Engineering Data Sets (CWEEDS files) are provided by Environment Canada for 564 locations in Canada, of which 81 in Québec (ECC Canada, 2020). However, in some cases the closest weather station for a location with a significant population density in Québec will be located in another province – this is for example the case for Gatineau, for which the closest CWEEDS weather station is the Ottawa International Airport. For this reason, a set of 134 CWEEDS stations were considered in this study. They are shown in Figure 4. Most stations include data for 20 years (1998–2017), but some of them only include partial data. The broader objective of this study is to develop a “province-wide” typical weather data file, and the candidate years for typical months would have to be restricted to the smallest common range amongst all stations, which would be 2002–2007 for the full data set. To avoid that problem, only stations with the full date range are considered in the following (94 stations).

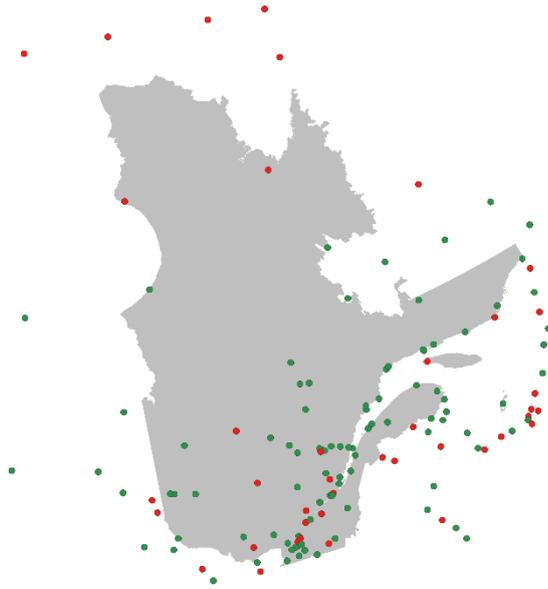


Figure 4 134 CWEEDS stations were selected “near” the province of Québec to represent the potential closest station to every census subdivision. Green dots represent the 94 stations with the full 20 years of data (1998-2017), red dots represent stations with partial data (which have been excluded, as discussed in the text)

Voronoi polygons were created in QGIS (QGIS.org, 2022) to represent a “catching area” of each weather station. The Voronoi polygons associated with a given set of reference points in a plane (called “seeds”, in our case the points representing the weather stations) correspond to the regions consisting of all points in the plane which are closer to the given reference “seed” (the basic concept can be found in the relative Wikipedia page (“Voronoi diagram,” 2022)). The polygons are shown in Figure 5.

Each polygon represents the region where all points are closer to a given weather station (the red dot inside each polygon). It is apparent that some regions in Québec are covered by the Voronoi polygons of weather stations in other provinces (e.g. Ontario for the South-West part of Québec). The extreme North is not covered by weather stations with complete date range, since most Nordic stations rely on satellite-derived solar radiation data available after 2005.

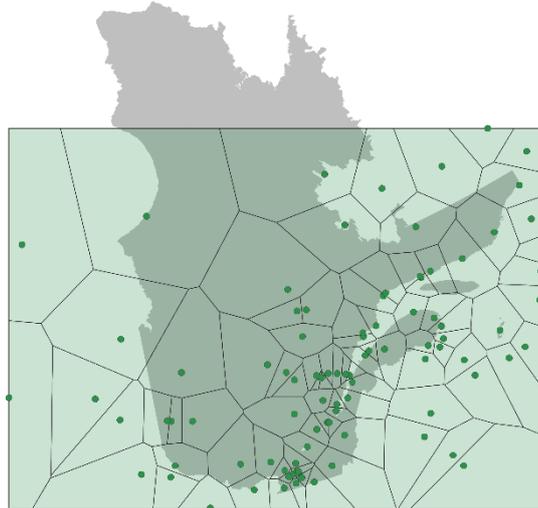


Figure 5 Voronoi polygons around the weather stations in the Québec province, Canada

To obtain weighting factors for the CWEEDS weather stations, each Voronoi polygon was associated with the Québec population covered by its area. To obtain an estimate of that population, the population density obtained from the census subdivisions was first rasterized over a 100 m mesh, and the total population was then calculated for each Voronoi polygon. This results in 65 weather stations associated with non-zero population (from 18 people for Île aux Perroquets to 1.385 million people for Montréal-McTavish). 29 of the 94 stations originally considered “near Québec” are associated with zero population (i.e., no population is found in their Voronoi polygon); they are removed from the list.

The result can be seen in Figure 6. Each CWEED station is colored with the total population covered by the polygon (stations with zero population are not shown). Most stations located outside the province of Québec are associated with low population counts, except for Ottawa International airport (the yellow dot in the lower part of the map).

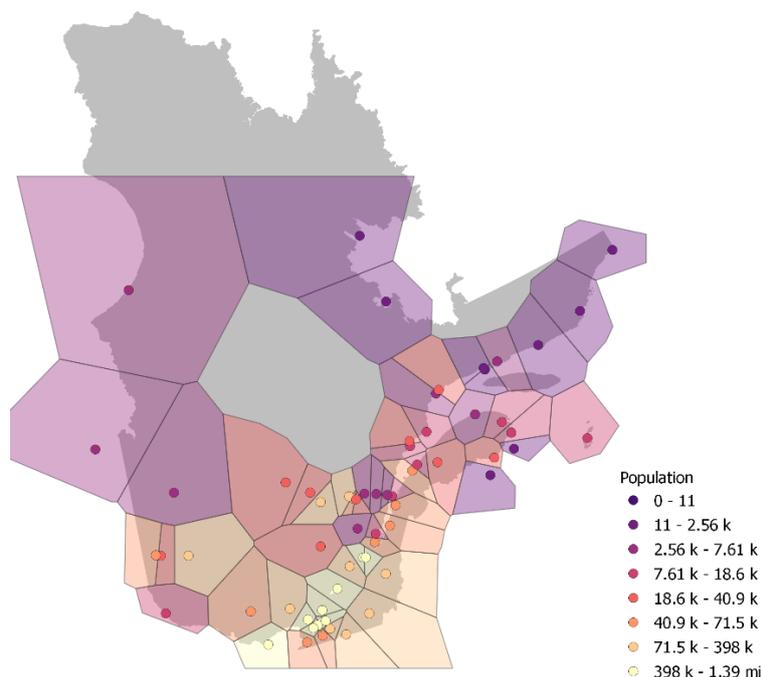


Figure 6 CWEEDS stations “near Québec” with their Voronoi polygons

Ranking of Québec CWEEDS station by population

The first 20 “most populated” weather stations are listed in Table 1, sorted by population. The first 6 stations are located within (or close to) the greater metropolitan area of Montréal (Census Metropolitan Area). They represent slightly over 50 % of the population and are located within 60 km of each other. Adding 6 stations (Québec city region, Gatineau, Trois-Rivières, Sherbrooke, and Saint-Jean-Sur-Richelieu) allows to reach almost 80 % of the population. These 20 weather stations will be considered for the “provincial-wide” selection, accounting for the 90 % of the population. Each population percentage will be used as the weighting factor.

Table 1 CWEEDS stations with their associated population percentage

No	Station name	Prov.	Climate ID	Lat/°	Lon/°	% Pop.	% Cumul.
1	MCTAVISH	QC	7024745	45.5	-73.58	15.5	15.5
2	MONTREAL INTL A	QC	7025251	45.47	-73.74	9.9	25.4
3	MONTREAL-ST-HUBERT	QC	7027329	45.52	-73.42	8.5	33.9
4	L'ASSOMPTION	QC	7014160	45.81	-73.43	7.5	41.3
5	MONTREAL MIRABEL INTL A	QC	7034900	45.68	-74.03	6.0	47.4
6	STE-ANNE-DE-BELLEVUE 1	QC	702FHL8	45.43	-73.93	5.7	53.0
7	NICOLET	QC	7025442	46.23	-72.66	5.5	58.6
8	QUEBEC INTL A	QC	7016293	46.8	-71.38	5.0	63.6
9	STE-FOY (U. LAVAL)	QC	701Q004	46.78	-71.29	5.0	68.6
10	OTTAWA INTL A	ON	6106001	45.32	-75.67	4.6	73.2
11	LENNOXVILLE	QC	7024280	45.37	-71.82	4.0	77.2
12	L'ACADIE	QC	702LED4	45.29	-73.35	2.4	79.6
13	BEAUCEVILLE	QC	7028754	46.21	-70.78	2.1	81.7
14	FRELIGHSBURG	QC	7022579	45.03	-72.85	1.8	83.6
15	JONQUIERE	QC	7063370	48.43	-71.14	1.5	85.1
16	ST-JOVITE	QC	703GDKB	46.08	-74.56	1.3	86.4
17	DESCHAMBAULT	QC	7011983	46.69	-71.97	1.2	87.6
18	ROBERVAL A	QC	7066686	48.52	-72.27	0.83	88.4
19	VAL-D'OR	QC	7098603	48.06	-77.79	0.81	89.2
20	MANIWAKI AIRPORT	QC	7034482	46.27	-75.99	0.78	90.0

RESULTS

First the typical year selection will be discussed and compared to the CWEC weather files in order to have a validation of the calculations performed, and then the provincial-wide year selection will be shown.

TMY typical years results and validation

After applying the *persistence* methods, every typical month of the 20 “most populated” weather stations, was selected. Generally, for all the calendar months of all the 20 weather stations, the CWEC years are always in the top five-month candidates, with few exceptions that can be explained with the exclusion by the author of the months that required a high amount of data filling to produce the CWEC weather files. It has been remarked also a mismatch between the *persistence* criteria applied in this work and the ones applied for the generation of CWEC weather files. However, this result was expected since the final report for the generation of the CWEC weather files (Morris, 2016) was not clear on the exact procedures that were followed for the *persistence* methods to exclude some of the candidate months. However, the goal of this step was just to develop and process the Finkelstein-Schafer (FS) statistic which, according to the results obtained, can be considered validated.

“Province-wide” typical year generation

Once all the cumulative FS weighted values are imported from any single python script of the 20 “most populated” weather stations, it is possible to weight them by population and sum them up to be able to understand which months can better represent the whole province of Québec.

A summary of FS values for January is shown in Figure 7 with a color scheme using dark green for the best statistical fit and gradual colors towards red for worse fits. The numbers shown do not yet account for the population weighting factors.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
QC_MCTAVISH	2,20	1,99	1,79	3,05	5,21	4,37	5,48	2,63	4,00	1,87	2,23	3,67	3,04	1,95	1,50	2,04	1,95	2,85	2,86	4,08
QC_MONTREAL_INTL_A	2,14	1,60	1,67	3,25	5,10	4,28	5,92	2,74	3,95	1,66	2,15	3,95	3,13	2,16	1,54	2,25	1,89	3,02	2,85	3,73
QC_MONTREAL-ST-HUBERT	1,64	1,82	1,65	3,38	4,84	4,26	5,40	2,98	4,02	1,83	1,87	3,90	3,28	2,15	1,98	1,76	1,68	2,85	2,60	4,02
QC_LASSOMPTION	1,94	1,81	1,58	2,48	4,68	3,96	5,00	2,51	4,36	1,92	2,14	4,09	3,22	1,95	1,35	1,62	1,70	3,28	3,10	3,99
QC_MIRABEL_INTL_A	2,40	1,92	1,81	2,73	4,71	4,01	5,37	2,77	3,85	1,97	2,35	4,26	3,70	2,13	1,56	1,94	1,73	2,97	2,83	3,54
QC_STE-ANNE-DE-BELLEVUE	1,99	1,74	1,68	3,12	5,14	4,43	5,48	2,42	4,04	1,67	2,16	3,90	3,29	2,02	1,60	1,94	1,87	2,99	2,72	3,76
QC_NICOLET	1,96	2,15	1,68	2,67	3,90	4,30	4,57	2,54	4,40	1,87	2,58	4,12	3,60	2,05	1,53	1,51	1,93	3,06	3,41	3,91
QC_QUEBEC_INTL_A	2,18	2,07	2,09	3,08	4,20	3,98	4,50	3,42	4,20	1,75	2,28	3,79	4,09	1,86	1,51	1,35	1,89	2,72	3,45	3,59
QC_STE-FOY (U. LAVAL)	2,93	2,00	2,24	2,98	3,82	4,04	4,41	2,95	3,78	2,07	2,15	3,79	3,89	1,83	1,46	1,59	2,10	2,75	3,52	3,80
ON_OTTAWA_INTL_A	2,06	2,07	2,08	3,02	5,13	4,46	5,08	3,00	4,27	1,96	2,52	3,57	3,34	1,80	1,90	1,89	2,02	3,68	2,39	3,79
QC_LENNOXVILLE	2,14	1,79	1,88	3,67	4,74	4,56	4,68	1,89	4,58	1,43	2,02	3,47	3,01	2,55	2,64	1,31	1,73	3,12	2,19	4,04
QC_LACADIE	1,98	1,75	1,82	3,26	4,72	4,59	5,34	2,53	4,16	1,76	2,21	3,80	3,18	2,35	1,94	1,69	1,88	3,03	1,92	4,44
QC_BEAUCEVILLE	2,62	2,17	1,90	3,06	4,27	4,20	4,22	2,68	3,94	1,99	2,21	3,70	3,16	2,18	2,28	1,59	2,39	2,98	2,89	3,71
QC_FRELIGHSBURG	2,10	1,75	3,24	3,34	5,11	4,76	5,01	2,27	4,51	1,58	2,16	3,29	2,79	2,74	2,78	2,15	2,07	3,55	2,36	3,93
QC_JONQUIERE	1,96	1,59	1,95	3,26	2,86	4,03	3,81	3,28	3,53	1,89	2,79	4,62	4,00	1,80	1,35	1,74	1,74	3,21	3,94	3,40
QC_ST-JOVITE	2,89	1,98	2,18	3,08	4,34	4,21	4,33	2,66	3,98	1,83	2,25	4,40	3,60	2,23	2,04	2,15	1,85	3,65	3,05	3,55
QC_DESCHAMBAULT	2,58	2,12	2,23	2,92	4,27	4,22	4,29	3,27	4,26	2,27	2,47	3,93	4,20	2,33	1,49	1,46	1,87	2,48	3,21	3,89
QC_ROBERVAL_A	1,89	1,59	1,91	2,85	2,46	3,46	3,94	3,37	3,85	2,03	2,22	4,49	3,21	1,79	1,37	1,49	1,91	2,86	3,77	3,31
QC_VAL-DOR	2,67	1,65	1,95	2,94	2,91	3,31	4,96	3,45	5,15	1,80	2,55	4,69	2,59	2,55	4,35	1,55	2,25	3,46	2,54	3,95
QC_MANIWAKI_AIRPORT	2,88	2,04	2,03	2,98	4,63	4,03	4,21	2,54	4,76	1,70	2,71	4,33	3,12	2,33	1,91	1,87	1,67	3,87	3,02	3,39

Figure 7 “Province-wide” selection for January.
Reported values still have to account for population factors.

The year that is selected to be the most representative for the Québec province, 2012, is highlighted with a darker black border. It can be noticed that the 2012 column features many really good statistical fits with most of the stations, especially the 9 most populated ones (almost 70 % of population).

Results for all calendar months are shown in Figure 8 with their own FS value summations after applying the population weighting factors.

	January	February	March	April	May	June	July	August	September	October	November	December
Year Selection	2012	2009	2011	2013	2014	2011	2015	2010	2001	2000	2003	2009
Total FS weighted	1.53	1.38	1.45	1.45	1.51	1.25	1.36	1.16	1.26	1.22	1.49	1.22

Figure 8 “Provincial-wide” year selection for each calendar month

In the case of January, the month was also the 1st ranked for the most populated McTavish station, but it is not the case for most of the other calendar month, therefore the selection is not that predictable at a first guess. However, it is also true that most populated areas will largely influence the selection, with the result that the typical year will most likely be “more typical” for the Montréal, Québec city, Sherbrooke and Gatineau areas rather than other locations in Québec.

20 weather stations were used here, representing the 90 % of the population, but the calculations were also repeated with the 15 “most populated” stations (85.1 % of population) and with 10 stations (73.2 % of the population) to see how many stations are recommended to be used.

Using 10 stations the years selected were the same except for February for which 2006 was preferred with a very narrow margin of less than 0.02 points over 2009. With 15 stations, 2009 was instead selected for February (same as with 20 stations), with a margin of less than 0.01 points: 11th, 12th, 13th and 14th stations were in fact better represented by 2009 and changed the selection. With 20 stations the trend was confirmed in favor for 2009 with 16th, 17th, 18th and 19th stations having a better match with 2009 rather than 2006.

It is true that most populated areas have a major impact, but it can also be remarked that sometimes also less represented stations can make the difference in certain cases. However, processing all the possible weather stations could require a considerable amount of work, since all the locations would require TMY calculations. It is therefore recommended to account at least for the 85 % of the total population or to consider all weather stations that represent at least 1.5 % of the population.

It is possible to remark that, even if for some stations there is not the best possible statistical fit, the years selected have a fairly good “typical” representation of most of the weather stations considered. The final result is consequently deemed to be suitable for energy simulation purposes and for peak demand estimations over the whole province.

CONCLUSION

Typical weather years known as “typical meteorological years” (TMY) or “test reference years” (TRY) are widely used for building energy modelling, as they offer a convenient solution to simulate one year that is representative of the climate over a longer period, typically 15 to 30 years. In Canada, Environment Canada has been publishing “Canadian weather for energy calculation” data files which follow the US methodology for TMY with some variations. These typical years are obtained separately for each weather station, and the methodology often selects different typical months for different stations. Simulations of buildings across a large region with different typical weather files can be analyzed together at a yearly (or perhaps monthly) scale, but not on an hourly scale or below, therefore preventing such results from being used to assess a utility-scale peak demand for example. To answer this need coming from electric grid and energy stakeholders, Huang (2020) proposed a “state-wide” typical year selection for California in which typical months are chosen, as for TMY, with the Finkelstein-Schafer (FS) statistic, but to best represent the entirety of the weather stations, weighting them by population. In this work, this methodology is applied to the province of Québec, Canada. First, the TMY typical year selection has been developed in *Python* and it has been validated through a comparison with the files coming from the Canadian Weather year for Energy Calculation (CWEC) (ECC Canada, 2020). Then, the “province-wide” approach has been applied to the province of Québec, Canada, using the results of the FS statistic of the TMY scripts and allocating the census population to each weather station thanks to Voronoi polygons. Results show that it is possible to select typical months for all significant weather stations together without losing the representativity for any particular one. Even if “most populated” weather stations had a major impact due to the weighting factors, applying the same method with the 10, 15, and 20 most populated weather stations, showed that in some cases even less populated weather stations can change the decision of the typical month. For this reason, it is recommended to use at least the 85 % of the total population or to consider all weather stations with at least 1.5 % of the province’s inhabitants. Further work could refine this “provincial-wide” method, analyzing the impact of different weighting factors on the final selection. It would also be useful to obtain a clarification from Environment Canada on the exact persistence criteria used in developing the published CWEC data files. Finally, it is recommended to apply the same methodology to all Canadian provinces to provide to the building simulation community an alternative to “per-station” typical files.

SUPPLEMENTARY MATERIAL

The “province-wide” typical year for the province of Québec, Canada, will be provided in GitHub at this address <https://github.com/GiulioTonellato/Canadian-weather-files>. Future projects also involve the publication of the python scripts to produce TMY weather files for single weather stations based on the CWEC methodology.

REFERENCES

Bilbao, J., Miguel, A., Franco, J.A., Ayuso, A., 2004. Test Reference Year Generation and Evaluation Methods in the Continental Mediterranean Area. *Journal of Applied Meteorology* (1988-2005) 43, 390–400.

- Eames, M.E., Ramallo-Gonzalez, A.P., Wood, M.J., 2016. An update of the UK's test reference year: The implications of a revised climate on building design. *Building Services Engineering Research and Technology* 37, 316–333.
- ECC Canada, 2020. Canadian Weather Energy and Engineering Datasets (CWEEDS) and Canadian Weather Year for Energy Calculation (CWEC) - 2020 update.
- Environment Canada, 2020. Canadian Weather Energy and Engineering Data Sets (CWEEDS files) and Canadian Weather for Energy Calculations (CWEC files) [WWW Document]. URL https://climate.weather.gc.ca/prods_servs/engineering_e.html
- Festa, R., Ratto, C.F., 1993. Proposal of a numerical procedure to select Reference Years. *Solar Energy* 50, 9–17. [https://doi.org/10.1016/0038-092X\(93\)90003-7](https://doi.org/10.1016/0038-092X(93)90003-7)
- Finkelstein, J.M., Schafer, R.E., 1971. Improved Goodness-Of-Fit Tests. *Biometrika* 58, 641–645. <https://doi.org/10.2307/2334400>
- Hall, I.J., Prairie, R.R., Anderson, H.E., Boes, E.C., 1978. Generation of a typical meteorological year (No. SAND-78-1096C; CONF-780639-1). Sandia Labs., Albuquerque, NM (USA).
- Holmgren, W.F., Hansen, C.W., Mikofski, M.A., 2018. pvlib python: a python package for modeling solar energy systems. *Journal of Open Source Software* 3, 884. <https://doi.org/10.21105/joss.00884>
- Huang, J., 2020. Update of California Weather Files for Use in Utility Energy Efficiency Programs and Building Energy Standard Compliance Calculations (No. PGE0450). Pacific Gas and Electric Company, Oakland, CA, USA.
- ISO, 2005. ISO 15927-4:2005. Hygrothermal performance of buildings -- Calculation and presentation of climatic data - Part 4: Hourly data for assessing the annual energy use for heating and cooling. International Organization for Standardization, Geneva, CHE.
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J.B., Grout, J., Corlay, S., 2016. Jupyter Notebooks-a publishing format for reproducible computational workflows.
- Lund, H., 1996. Design Reference Years and test Reference years in Europe, Turkey and Israel. Citeseer.
- Lund, H., 1995. The Design Reference Year user's manual, Thermal Insulation Laboratory. Technical University of Denmark, Lyngby.
- Marion, W., Urban, K., 1995. User's Manual for TMY2s (Typical Meteorological Years)-Derived from the 1961-1990 National Solar Radiation Data Base. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Morris, R., 2016. Final Report – Updating CWEEDS Weather Files. Environment and Climate Change Canada, Toronto, ON, CAN.
- National Climatic Data Center, 1976. Test Reference Year (TRY) – Tape Reference Manual TD-9706. Asheville, NC: National Climatic Data Center.
- Pernigotto, G., Prada, A., Gasparella, A., Hensen, J.L.M., 2014. Analysis and improvement of the representativeness of EN ISO 15927-4 reference years for building energy simulation. *Journal of Building Performance Simulation* 7, 391–410. <https://doi.org/10.1080/19401493.2013.853840>
- Pvlib, 2022. pvlib.iotools.read_epw documentation (Pvlib library) [WWW Document]. URL https://pvlib-python.readthedocs.io/en/stable/reference/generated/pvlib.iotools.read_epw.html#pvlib.iotools.read_epw (accessed 6.3.22).
- QGIS.org, 2022. QGIS Geographic Information System. QGIS Association.
- Siu, C.Y., Liao, Z., 2020. Method for converting CWEEDs weather files to EPW format for multiyear simulation of building thermal dynamics. *MethodsX* 7, 101016. <https://doi.org/10.1016/j.mex.2020.101016>
- StatsCan, 2022. Population and dwelling counts: Canada and census subdivisions (municipalities) - Table: 98-10-0002-01 (2021 census) [WWW Document]. URL <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810000201> (accessed 5.31.22).
- The Pandas development team, 2022. Pandas 1.4.3. <https://doi.org/10.5281/zenodo.6702671>
- UNFCCC, 2015. COP 21 Agenda : Updated compilation of information on mitigation benefits of actions, initiatives and options to enhance mitigation ambition.
- Voronoi diagram [WWW Document], 2022. . Wikipedia. URL https://en.wikipedia.org/w/index.php?title=Voronoi_diagram&oldid=1088953309 (accessed 6.4.22).
- Wilcox, S., Marion, W., 2008. Users manual for TMY3 data sets (NREL/TP-581-43156). National Renewable Energy Laboratory, Golden, CO, USA.