TESTING THE ACCURACY OF SYNTHETICALLY-GENERATED WEATHER DATA FOR DRIVING BUILDING ENERGY SIMULATION MODELS

Larry O. Degelman

Department of Architecture, Texas A&M University
College Station, Texas 77845, USA

ABSTRACT

Weather data in formats required for annual energy simulations are not available at all locations where building designs are being evaluated. Synthetically generated weather data for these sites could be a viable option. This paper reports on the differences in weather and energy prediction results when using synthetically generated data vs. the use of recorded weather data for 50 cities worldwide. Energy simulations were performed on an office building placed in each city – first, by using recorded weather data and second, by using weather data synthetically generated by a statistically based model.

Annual energy predictions were the main focus of the study; however, other parameters of interest were also observed. These included peak heating and cooling loads and other weather-related values, such as high and low dry-bulb temperatures, humidity, annual heating and cooling degree days, and global solar radiation.

KEYWORDS

Building Simulation 2007, Weather Data, Climatic Parameters, Synthetic Weather, Statistical Models.

INTRODUCTION

Undoubtedly, detailed hourly weather records are the most desirable input data for performing building energy simulations. Key climatic elements in such data include temperature, humidity, solar radiation, and wind. Though long-term climatic norms are available for most regions around the globe, often hourly weather records are not available -- either the data are difficult to obtain, or the data have never been collected. In numerous cases, however, other weather data, such as daily max-min temperatures or monthly statistics of temperature norms and extremes, humidity, and global solar radiation, are available. For example, climatic norms are available from the World Meteorological Organization (WMO 1984) for the 30-year period 1961-1990, ASHRAE has developed the IWEC data sets (Thevenard and Brunger 2002), and NREL has developed the TMY2 (1995) data sets. The lack of hourly records for many locations, however, has led a number of researchers to develop statistically-based simulation models that will generate hourly weather data based on long-term weather records (Akasaka 1999, Degelman 1991, Erbs et al 1983, Meteonorm 2004, Zhang et al 2002). It is not within the scope of this study to test the efficacy or comprehensiveness of each of these simulation models; however, it is of interest to use one of these models as a example of what potential lies in this sort of approach when only sparse data (long-term norms and extremes) are available. The purpose of this study, therefore, is to use one of the simulation models to establish the potential performance of synthetically generated weather data when used for building energy calculations.

To maximize the breadth of the conclusions of the study, 50 sample cities were selected from six continents around the globe. The cities selected are tabulated in Table 1 and visually depicted in Fig. 1. The geographic distribution of cities ranges from near 40º South (Melbourne, Australia) to 63º North (Ostersund, Sweden.) A necessary criterion used in the selection of cities was that complete and usable weather data sets had to be available to the author for each city. These data sets are mentioned later.

Table 1 Worldwide cities used in this study

	WMO	*City Name*		*WMO*	*City Name*
1	02226	Ostersund, Swe	26	47412	Sapporo, Jpn
2	03534	Birmingham,Uk	27	47715	Tokyo, Jpn
3	03776	London, UK	28	47827	Kagoshima, Jpn
4	06700	Geneva, Szt	29	48456	Bangkok, Thl
5	07110	Brest, Fra	30	48647	Kuala-Lumpur
6	07690	Nice, Fra	31	48698	Singapore,Sgp
7	08023	Santander, Spn	32	48820	Hanoi, Vtn
8	10147	Hamburg, Ger	33	54342	Shenyang, Chn
9	10384	Berlin, Ger	34	54511	Beijing, Chn
10	11150	Salzburg, Aat	35	56778	Kunming, Chn
11	11782	Ostrava, Czk	36	58367	Shanghai, Chn
12	12375	Warsaw, Pol	37	59287	Guangzhou, Chn
13	12882	Debrecen, Hgy	38	62414	Aswan, Egy
14	15420	Bucharest, Rom	39	67775	Harare, Zbe
15	15552	Varna, Bul	40	71612	Montreal, Can
16	16066	Milan, Ity	41	71877	Calgary, Can
17	16242	Rome, Ity	42	72202	Miami, Fl
18	24959	Jakutsk, Rus	43	72469	Denver, Co
19	26629	Kaunas, Lth	44	72658	Minneapolis, Mn
20	27612	Moscow, Rus	45	76679	Mexico City, Mx
21	33837	Odessa, Ukr	46	80222	Bogota, Col
22	42809	Calcutta, Ind	47	87576	BuenosAires,Arg
23	43279	Madras, Ind	48	94120	Darwin, Aus
24	47105	Kangnung, Kor	49	94767	Sydney, Aus
25	47152	Ulsan, Kor	50	94866	Melbourne, Aus

Figure 1 Geographic distribution of the 50 cities used in the study (shown by the dots)

PROCEDURES

The procedures for this study were as follows:

1. Select 50 widely dispersed climatic regions,
2. Select a case-study building,
3. Run three energy simulations on the office building in each of the 50 cities chosen using (a) standardized IWEC weather files, (b) simulated weather from IWEC statistics, and (c) simulated weather using the 30-year statistics from the WMO database,
4. Tabulate summaries of weather and energy results from all computer runs, and
5. Prepare an error analysis summary.

CASE-STUDY BUILDING

The author had previously occupied a building that is ideally suited to the purposes of this study. The building is an office/laboratory/classroom on the campus of Nagoya University. (See Figs. 2 & 3.) It has very modest internal lighting loads, so it is naturally a good candidate for studying the effects of external weather conditions.

The case study building characteristics are as follows:

- Size: 10-story, 62m by 30m, U-shaped plan, Gross floor area: 15,980 m^2.
- Wall: 25 cm thick (10 cm brick, 5 cm rigid insulation, 10 cm block); U.F. 0.58 W/m^2K, per Table 4, Chap 25 (ASHRAE 2005).
- Windows: Single pane clear glass, aluminum frames; U.F. of 6.49 W/m^2K, per Table 4, Chap. 31 (ASHRAE 2005).
- Lighting power density: Ranges from 1.8 to 18.2 W/m^2, resulting in 10.5 W/m^2 on a weighted floor area basis.
- Typical office: 3.57m by 7.4 m, or 26 m^2.
- Motion sensors control lights (ON/OFF control) in circulation areas and restrooms.

The internal lighting power densities comply with the ASHRAE Standard 90.1-2004 criteria, though this was not a requirement for this study. No attempt was made to alter the building's thermal properties to adapt to the various climates. To assure that the results would not be skewed by important weather conditions occurring during unoccupied periods, the building occupancy was assumed to be seven days per week.

Figure 2 South face of case study building

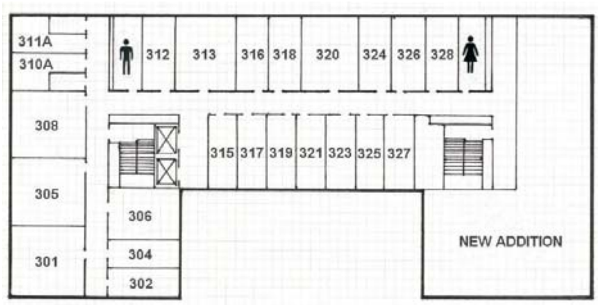

Figure 3 Typical office floor plan in case study building. Approximately 268 offices total.

The weather simulation model used in this study cannot be described in detail here, but it has been reported by the author in earlier publications shown in the reference list (Degelman 1991, 2004.) In brief, the model accepts as inputs: the site latitude, longitude and elevation; monthly means and standard deviations of temperatures for monthly average, average maximum, and average dew-point; and monthly values for average wind speed and average daily global insolation. This model has been undergoing improvements for the past 10 years.

SIMULATION RESULTS

The results are examined in terms of how well the modeling of synthetic weather data performs on: (1) an annual basis for weather parameters, (2) a monthly basis for weather parameters, (3) an annual basis for building energy use, and (4) peak building loads encountered through the year of simulation. The reference base for the simulations is considered to be the IWEC weather files, since these have had extensive peer-reviewed research behind their development. So, in most cases, the IWEC results will be regarded as the "truth" condition. All other computer runs are regarded as "simulations."

A look at existing weather data sets.

Before examining whether simulated weather data closely align with real weather data, it is instructive to look into whether existing weather records are consistent in themselves. If fact, it can be shown that there are marked differences in the year-to-year weather records within a single data set, and even larger differences between different data sets. The data sources used in this study are the International Weather for Energy Calculations, IWEC (Thevenard and Brunger 2002), and the World Meteorological Organization (WMO 1984.) The IWEC records are based mainly on years 1982 through 1999; while the WMO summaries are based on the 30-year record from 1961 through 1990.

A summary chart of monthly records for all 50 cities would be too large for this paper, so only the records for Beijing are used in this first discussion. The following graphs show the annual and monthly data for several weather parameters important to energy calculation models. IWEC data files are a full 8760 hours of annual weather data selected specifically for energy calculations in buildings. On the other hand, WMO data are 30-year statistical summaries of only 100 values that need to be put into a simulation model to synthetically generate what could be regarded as "typical" weather data. As stated above, IWEC files are used as the reference base against which all other results are measured. The following analyses include calculation of the MBE (mean bias error), the RMSE (root mean square error), and the R-squared from regression calculations. In all but a few cases, the comparisons are always made against the IWEC data.

In Figures 4 through 6, all 17 years that were used to create the IWEC file are plotted (i.e., each year from 1982 to 1999.) This is to reveal the spread of values that may be expected from year to year. These lines are in light tone and unlabeled. Each graph also shows both the 30-year (1961-90) statistical results from the WMO data compared to the monthly values obtained from IWEC files (years 1982-99). These are in a heavier tone to make them distinct. Only a sample of the output comparisons is shown here, mainly to ensure inclusion of the main parameters influencing heating and cooling loads; i.e., dry-bulb temperatures, humidity and solar radiation. The first observation for Beijing is for average monthly temperatures, shown in Figure 4.

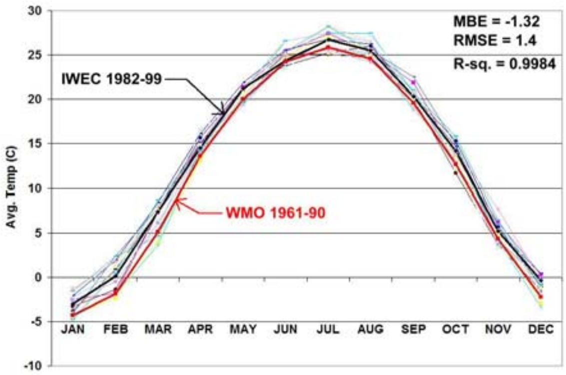

Figure 4 Monthly average temperatures from IWEC and WMO records for Beijing.

A quick glance at Figure 4 reveals that the IWEC values lie close to the average of the spread of yearly data points, which was their intent. The 30-year WMO records, which are unrelated to the IWEC records, show a lower average temperature than the IWEC records (MBE= -1.32C).

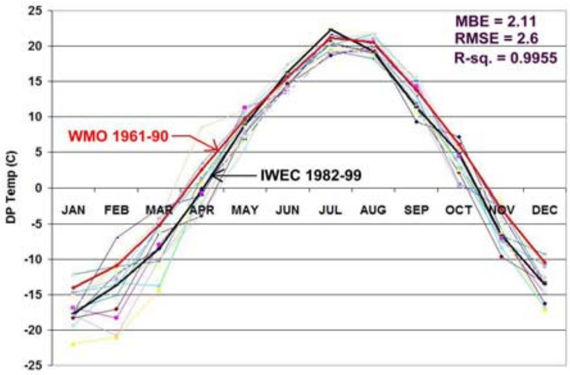

Figure 5 Average daily dew-point temperatures from IWEC and WMO records for Beijing.

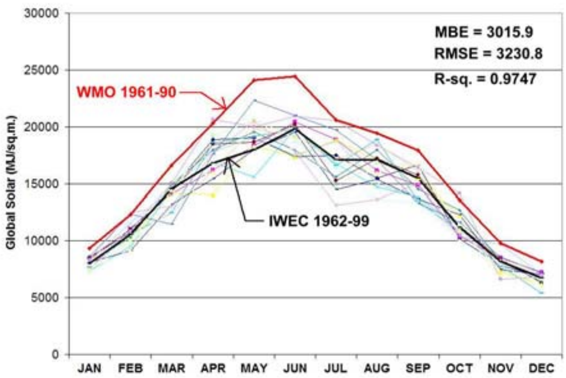

Figure 6 Average monthly global insolation from IWEC and WMO records for Beijing.

The solar results show a wide departure between IWEC and WMO records. One explanation for this is that the weighting factors used in selecting the months in the IWEC files may have been weighted heavier on the temperature values than on solar.

Plots of the wind speed parameter showed the most radical deviations through successive years. This is expected since the IWEC selection procedure places least importance on wind speed; thus, less weight.

The foregoing was an examination of weather data source differences for a single city. Next, we should look at consistencies, or lack of, on a global scale by examining all 50 cities used in this study. Annual heating degree-day (HDD) comparisons are displayed in Figure 7. Here, both the IWEC and WMO data are compared to a published reference standard, ASHRAE Standard 169 (2006.) This is a bar chart with a comparison of HDD18 values from Standard 169, IWEC data and the WMO data. Figure 8 is a plot of both IWEC and WMO against the ASHRAE source reference.

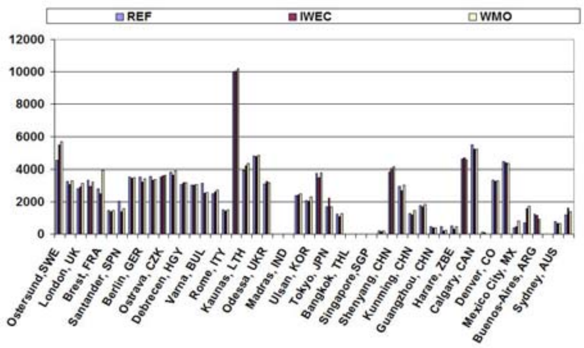

Figure 7 Annual heating degree-days (base 18C) for the 50 cities worldwide

In Figure 8, the IWEC data provide a better match (MBE=15.5) to the degree-days published by ASHRAE than do the WMO data (MBE=99.7), most likely because the periods of record for IWEC and ASHRAE are similar and more recent.

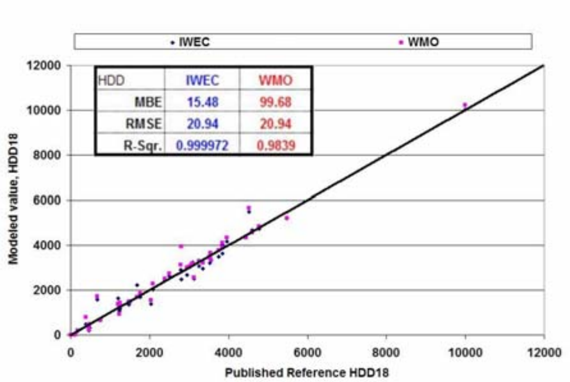

Figure 8 Regression of IWEC and WMO data against ASHRAE published heating degree-days (base 18C) for the 50 cities worldwide

So, the records show that there are wide variations in year-to-year weather data and also in values obtained from different weather data sets. This gives rise to the question of just what a weather simulator should emulate. Generating a year of 8760 hours of weather data for an energy simulation model is still (and always will be) a single "possible" year. The simulation model utilized here attempts to generate a full range of values from the lows to the highs so that design values will be encountered during the simulation. It does this by varying parameters using means and standard deviations, so average values are also met, thus enabling energy calculations. The year of weather generated has never happened in the past, nor will it ever happen in the future. The same can be said of IWEC data, since they consist of selected months spliced into what could be referred to as a "typical year" of weather. IWEC, or any other typical year methodology, has the disadvantage of possibly missing peak values in a number of weather variables for any given month of the year, but the statistical simulation model will always encounter the extreme high and low values while maintaining the average conditions throughout.

Performance of simulated weather data

The next step in this study is to view how well simulated data align with the real weather records. This will be examined in two stages – first, looking at aggregated values such as heating and cooling degree-days and annual solar radiation, and second, by looking at extreme high and low temperatures. The graphs show comparisons of simulated results against the IWEC source statistics and the WMO data set against the IWEC files. The presence of WMO comparisons is provided solely to reveal what levels of uncertainty might be present when one must choose from a statistical source where IWEC (or other "typical year") files have not been developed. It should be pointed out that where the simulation model is used, the input statistics were from the same period of record as that used by the IWEC files, so these results truly test the ability of the simulator to match the records of the data source. The simulator would equally match the WMO records if that data set were used as the source.

Figure 9 shows the comparison of monthly heating degree-days for Beijing. In this figure, the curve for the IWEC is almost indiscernible from the curve for the simulated results. Its MBE is 1.08, while the WMO curve shows a much greater departure (MBE=27.8.) Figure 10 is similar, but it is for monthly cooling degree days for Beijing.

Simulated results provide a fairly easy match for aggregated data (like degree-days), either monthly or annual; but since the simulation methodology is based on statistical principles, it is more of a challenge to meet individual monthly values and peak highs and lows. We will now look at results of the other parameters such as, average temperatures, solar and wind. These comparisons are made between the base reference (IWEC files) and the simulated results using the same period of records (1982-99) for the statistical input values. Results are shown in Figures 11 and 12. As these figures show, there is an almost indiscernible difference in the plots in the graphs. The MBE and R-value are included on each graph.

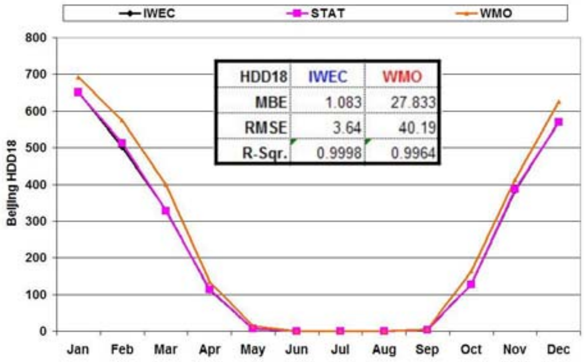

Figure 9 Comparison of simulated and WMO monthly heating degree-days to the IWEC source

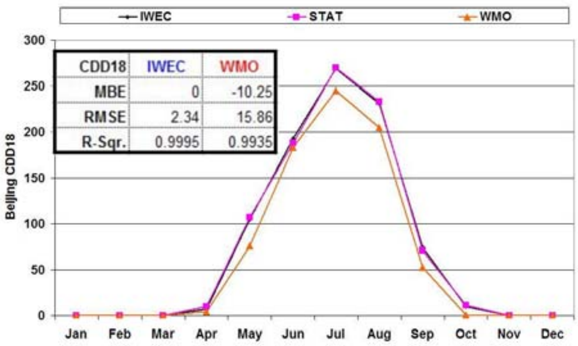

Figure 10 Comparison of simulated and WMO monthly cooling degree-days to the IWEC source

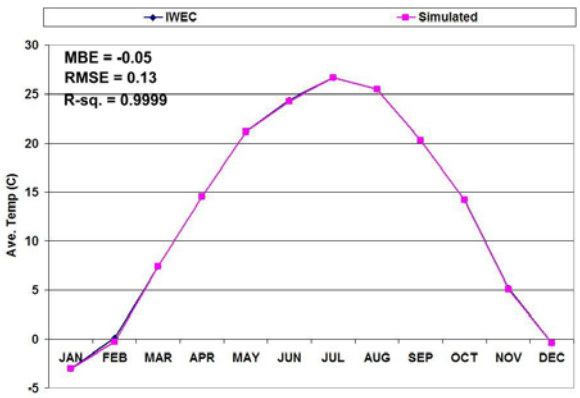

Figure 11 Comparison between simulated and IWEC monthly dry-bulb averages for Beijing

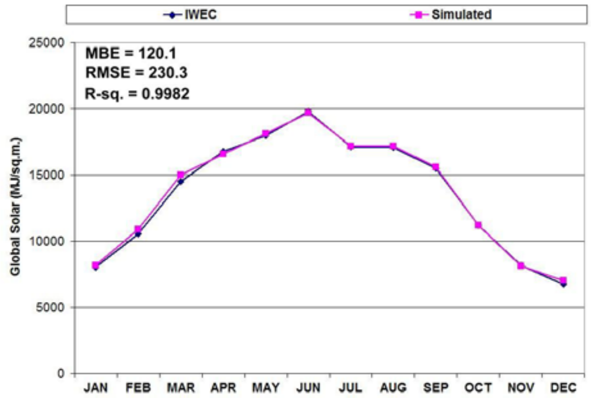

Figure 12 Comparison between simulated and IWEC monthly global solar for Beijing

Peak temperature performance evaluations

This section looks at maximum and minimum temperatures simulated from the IWEC and WMO sources. Figures 13 and 14 show comparisons of the max-min values for all 50 cities. In these graphs, the published ASHRAE "mean of annual extremes" was used as the "reference" maximum and minimum temperatures. Compared to this are simulation results using IWEC statistics and WMO statistics as input to the model. The MBE, RMSE and R-square are computed for both the IWEC and WMO.

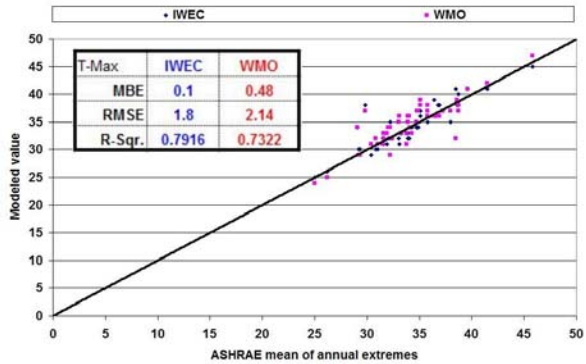

Figure 13 Maximum temperature comparison, 50 cities

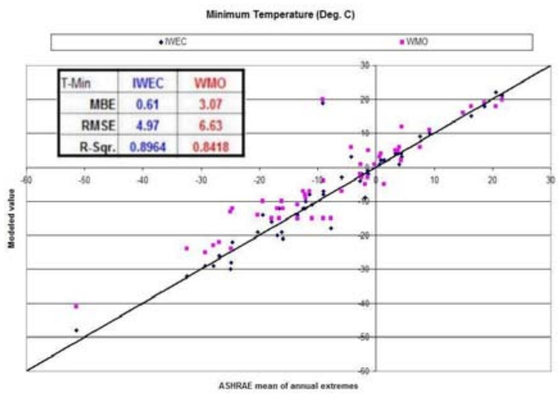

Figure 14 Minimum temperature comparison, 50 cities

Annual performance for energy simulations

This section shows graphs and charts that contrast two computer runs against the computer run results from an IWEC file: (1) a simulation using IWEC statistics, and (2) a simulation using the WMO statistics. Figure 15 shows annual cooling energy and Figure 16 shows annual heating energy.

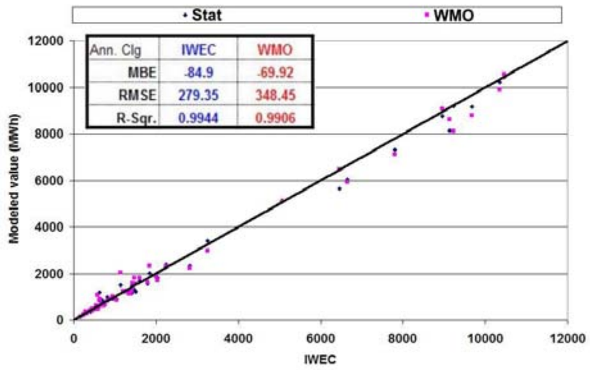

Figure 15 Annual cooling energy comparisons, 50 cities

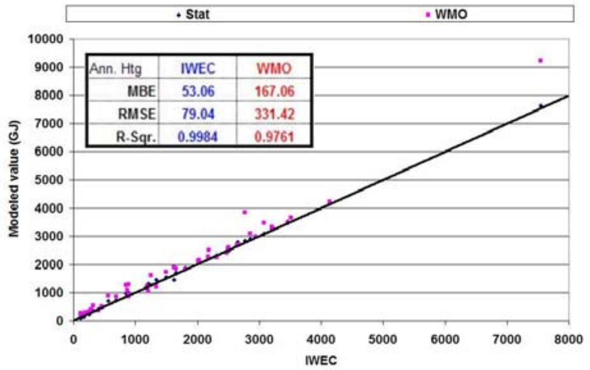

Figure 16 Annual heating energy comparisons, 50 cities

The simulation results include a whole-building performance factor, designated as Building Energy Performance Summary (or BEPS). This represents the total source-line energy divided by the building's gross floor area. The units are MJ/sq.m., which is a measure of the building's overall energy efficiency. The BEPS results for the 50 cities are shown in Figure 17. Peak cooling load comparisons are shown in Figure 18. Once again, three annual simulations are performed for each city: one using the IWEC data, the other two being simulations using statistics from IWEC and WMO data. The latter are contrasted against the actual IWEC run.

Error discussion

For each of the graphs, an error analysis was performed to determine the closeness of fit between data simulated and the source data statistics. These have been shown in each graph as MBE (mean bias error), RMSE (root mean squared error), and R-sq. (the square of the correlation coefficient, or the fraction of explained variance.)

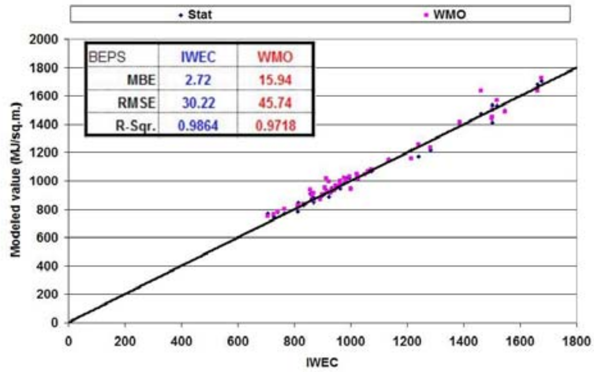

Figure 17 Annual BEPS comparisons for 50 cities

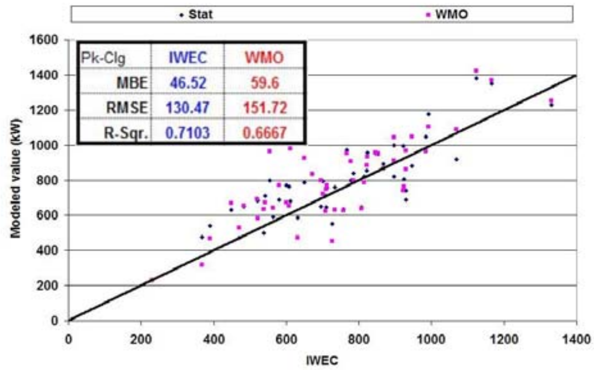

Figure 18 Peak cooling load comparisons, 50 cities

It is difficult to draw conclusions with such a large number of plots, so several of the parameters have been condensed to just three graphs. These are shown in Figures 19, 20, and 21, categorized as to weather, energy, annual, monthly, and peak values. For these graphs, the values for "IWEC MBE" and IWEC RMSE" are errors produced by the simulation using IWEC statistics compared to the IWEC file.

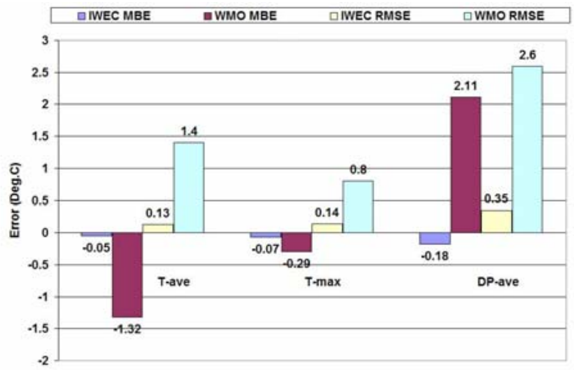

Figure 19 Monthly Tave, Tmax, and DPave errors for Beijing for IWEC and WMO simulations

The purpose of this exercise was to illustrate the potential errors that might be expected when using a readily accessible data source (such as WMO) when IWEC files are not available. Figures 19 and 20 show temperature comparisons for a single city (Beijing), and for all 50 cities, respectively.

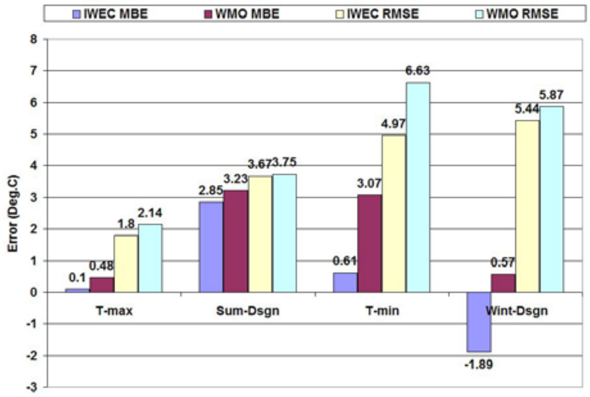

Figure 20 Tmax, Tmin, and design temp errors for all 50 cities for IWEC and WMO simulations

The last bar chart, Figure 21, compares energy simulation results, first using IWEC files, then comparing simulated IWEC data and WMO data. Since values of MBE and RMSE are rather large, these lead to misleading conclusions, so these values are expressed as percent error, rather than using the absolute values.

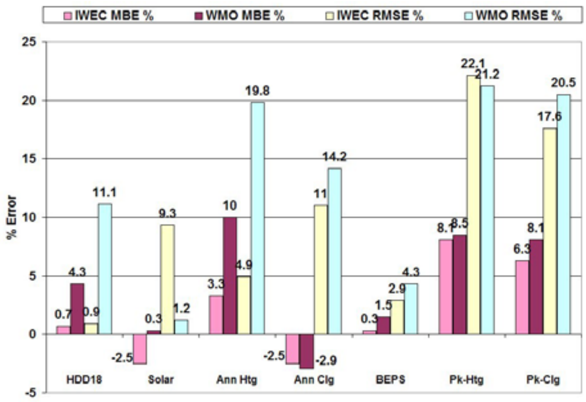

Figure 21 Percent error for IWEC and WMO annual weather and building energy use, 50 cities

SUMMARY AND CONCLUSIONS

This study attempted to evaluate the accuracy of simulated weather data when performing annual energy analyses for buildings. A standardized IWEC data file was used for each of 50 cities worldwide, followed by simulations using an IWEC statistical summary and a WMO climatic summary data set. For the three computer runs, comparisons were made on the Tave, Tmax, Tmin, HDD18, CDD18, heating and cooling peak loads, and the whole-building energy performance (BEPS.)

Errors and correlation coefficients from the foregoing graphs are summarized here in Table 2, for the city of Beijing. Using IWEC statistics, the simulated monthly results show very little departure from the IWEC reference file. Figure 19 revealed a simulation MBE of -0.05C for the Tave, -0.07C for the Tmax, and -0.18 for the DPave. When using a different data set (the WMO climatic summary), the departures are somewhat larger, but the dry-bulb temperatures are still within 1.4C of the recorded values, and the dew-point temperatures are within 2C of the real weather. These differences are not unexpected since the WMO and IWEC data are from completely different periods of record.

Table 2 Error comparisons for Beijing monthly data using two types of simulations; IWEC as reference

Weather Parameter	IWEC sim			WMO sim		
	MBE	RMSE	R^2	MBE	RMSE	R^2
Tave (C)	-0.05	0.13	1.000	-1.32	1.4	.998
Tmax (C)	-0.07	0.14	1.000	-0.29	0.8	.996
DPave (C)	-0.18	0.35	.9995	2.11	2.6	.996
HDD18 (%)	0.5	1.6	1.000	12.4	17.9	.996
CDD18 (%)	0	3.2	.9995	-13.8	21.4	.994
Solar (%)	0.9	1.7	.998	22.1	23.7	.975
Wind (%)	0.4	1.3	.996	15.7	20.6	.675

From Table 2, it is evident that when a statistical simulation model is used for a single city, it can generate average monthly temperatures with an average error of about 0.13C. Errors in predicting the highest monthly temperatures average around 0.14C, while dew-point temperature errors are around 0.35C. The mean predicted temperatures are about 0.1C too low with the particular model tested. As for aggregated data, errors in heating degree day predictions average around 1.6%, and cooling degree days at around 3.2%. Monthly global solar radiation data shows an average error of about 1.7%, and wind speeds have an average monthly error of 1.3%. The correlation coefficients between simulated and real weather are very high, all being above 0.996.

When another weather data set (WMO) is used, the dry-bulb values still average around 1.5C difference. Monthly dew-point temperature differences average around 2.6C. The other aggregated data (degree days, solar and wind) are within 17 to 24% of real data for the monthly predictions. Using WMO data as a second reference point does not test the simulation model; rather, it gives a benchmark as to what deviations do exist in real weather patterns and what differences one can expect when using different periods of weather records. These differences, therefore, should not be referred to as errors.

Table 3 summarizes the errors for the annual runs for all 50 cities. Temperature differences are expressed in degrees C, and aggregated value differences are expressed in percentages. It shows performance in the prediction of max-min temperatures, heating degree days, annual global solar radiation, annual heating and cooling energy, peak heating and cooling loads, and finally the overall building energy performance metric (BEPS.) The main conclusion from this table is that when a simulation model uses a set of statistics from a given period of weather records, it can produce results in fair proximity to those coming from calculations using the real weather records. Further, it is also evident that the simulation results easily provide a closer match than those that would be produced by calculations using a different weather data record, to wit, IWEC vs. WMO records.

Table 3 Error comparisons of annual data for all 50 cities using two types of simulations vs. IWEC files

Weather Parameter	IWEC sim			WMO sim		
	MBE	RMSE	R^2	MBE	RMSE	R^2
Tmax (C)	0.10	1.8	.792	0.48	2.14	.732
Tmin (C)	0.61	4.97	.896	3.1	6.63	.842
HDD18 (%)	0.7	0.9	1.000	4.3	11.1	.984
Solar (%)	0.3	1.2	.998	-2.5	9.3	.851
Ann Htg (%)	3.3	4.9	.998	10.0	19.8	.976
Ann Clg (%)	-3.4	11.0	.994	-2.9	14.2	.991
Pk Htg (%)	8.1	22.1	.931	8.5	21.1	.923
Pk Clg (%)	6.3	17.6	.710	8.1	20.5	.667
BEPS (%)	0.3	2.9	.986	1.5	4.3	.972

Across all 50 cities, simulated results of annual maximum temperatures average 0.1C above values found on the IWEC data files, and annual minimums average 0.6C above. The differences between the WMO and IWEC records are about 5 times that. Aggregated values of degree-days and solar radiation yield only 0.3 to 0.7 percent difference compared to the IWEC data. Energy calculations revealed by the heating and cooling summaries seem to incur differences almost equally between simulated values and WMO data, though the tendency is still toward sightly less deviance in the simulated weather results. The reasons for this can be quite obscure, since energy calculations take into account lighting and occupancy schedules that may either add to or subtract from the thermal loads depending on whether these loads are supporting or opposing the weather-induced loads. A building also has significant heat storage and time lags that influence its thermal response independent of the weather.

RECOMMENDATIONS

This study has looked at one particular simulation model and has made comparisons to two weather data sets: IWEC and WMO. Several other countries, such as Brazil, China, France, Hong Kong, and Japan, have also produced typical year weather data. It would add value to this topic if these other data sets would be included. Also beneficial would be the testing of a few other weather simulation models. This study merely exposed some of the accuracies and/or inaccuracies in results when using a single simulation model.

REFERENCES

Akasaka, H. 1999. "Typical Weather Data and the Expanded AMEDAS Data", Solar Energy, Vol. 25, No.5.

ASHRAE 2005. Handbook of Fundamentals 2005, American Society of Heating, Refrigerating, and Air-conditioning Engineers, Inc., Atlanta, GA.

ASHRAE 2006. ANSI/ASHRAE Standard 169-2006 Weather data for building design standards, ASHRAE, Inc., January, 64 pp.

Degelman, L. 1991. "A statistically-based hourly weather data generator for driving energy simulation and equipment design software for buildings", *Proc. Building Simulation '91*, Int'l Building Performance Simulation Assoc. (IBPSA), 20-22 Aug., pp 592-599.

Degelman, L. 2004. "Simulation and uncertainty: weather predictions", ***Advanced Building Simulation***, (Malkawi and Augenbroe, Ed.), Spon Press, New York and London, Chap. 3, pp 60-86.

Erbs, D.G., Klein, S.A., and Beckman, W.A. 1983. ''Estimation of Degree-Days and Ambient Temperature Bin Data from Monthly Average Temperatures,'' ASHRAE J., **25**, pp. 60–65.

Meteonorm 2004, Global Meteorological Database for Solar Energy and Applied Meteorology, Version 5.1, Meteotest, Born, Switzerland.

Thevenard, D.J. and A.P. Brunger, 2002. "The Development of Typical Weather Years for Int'l Locations: Part II, Production", (Technical Paper 4596, RP-1015). ASHRAE Trans., Vol 108(2).

TMY2 1995. User's Manual for TMY2s - Typical Meteorological Years, NREL/SP-463-7668, National Renewable Energy Laboratory, Golden, CO, June, 47 pp.

WMO 1984. World Meteorological Organization, 1984: Technical Regulations, Vol. I. WMO Publication No. 49. Geneva, Switzerland.

Zhang Q., Huang, J., and Lang S. 2002. "Development of typical year weather data for Chinese locations", ASHRAE Trans. v. 108 (2), June.