ASSESSING THE VALUE OF TYPICAL METEOROLOGICAL YEARS BUILT FROM OBSERVED AND FROM SYNTHETIC DATA FOR BUILDING THERMAL SIMULATION

Ricardo Aguiar, Susana Camelo and Helder Gonçalves

INETI / ITE - Department of Renewable Energies
Estrada do Paço do Lumiar, 1649-038 Lisboa, Portugal

ABSTRACT

Typical Meteorological Years obtained from observed meteorological records have become the *de facto* data source when evaluating thermal performance of buildings. However, this data source has various drawbacks, and alternative TMY assembling methods based on statistical and stochastic models seem to have been perfected to a point where they become sound alternatives. Numerical simulations of test cells were performed for a mid-latitude temperate climate, Lisbon. Thermal performance was evaluated using as input long term observed time series (control situation), long term stochastic data, and TMY obtained by classic and by stochastic methods. It is found that stochastic data seems to provide a good source of meteorological data, indeed more flexible and adequate than those series provided by the currently usual approach.

INTRODUCTION

Typical Meteorological Year type data sets (hereafter, TMY) obtained from observed meteorological records –“classic” TMY– have become the usual data source when evaluating thermal performance of buildings through numerical simulation. However there are not many sites for which TMY adequate for this purpose exist. This can be blamed mainly on the shortness of observed hourly data in general, and of solar irradiation data in particular. Other drawbacks of this data type can be pointed. Classic TMY rely on selection of monthly series from a pool of observed data – but its limited size is seldom enough to reach the goal of exactly matching TMY and long term statistics, most remarkably when considering all meteorological parameters simultaneously. This is due to the large interannual variability of the meteorological series. For instance, Mediterranean climates display variation coefficients of monthly values is about 10% for solar irradiation, 7% for temperature and 6% for relative humidity. Also, and again due to the necessarily limited size of the observed data pools, a hierarchy of meteorological parameters and statistical characteristics of interest must be established, therefore they can’t all be represented at the same level. Finally, classic TMY are often built as “general purpose tools”, and thus correctness of mean values of parameters like precipitation and atmospheric pressure – nearly irrelevant for building simulation – are also put near the top of the selection criteria. This again can cause other parameters that interest more to the thermal simulation, like irradiation, to be less well represented.

In recent years, and for some regions/climate types, statistic and stochastic models have been developed to a degree of sophistication enough to provide an alternative way of assembling sound synthetic multivariate meteorological data series – either long term series or TMY (*e.g.* Knight *et al.*, 1991; Aguiar, 1998a). For convenience, these will be termed “stochastic” hereafter, although in fact many of the models used in time series generators are just regressions, balance, or thermodynamic equations. Stochastic data are based on necessarily simplified models of the time series occurring in Nature (note that some statistical properties are nearly irrelevant to building behaviour due to smoothing effects introduced *e.g.* by thermal inertia), but have numerous advantages in respect to classic TMY. They are free from discontinuities, gaps, and spurious values. They can be produced starting with input climatic monthly data only, widely available from Atlas and meteorological publications. And they can be made to yield at the output exactly these long term monthly data.

In previous works by the same authors [2, 3] it was shown that observed and synthetic TMY yield compatible building performance statistics for Mediterranean climates. However, these studies where not conclusive, as no reference building performance estimates (computed with “true” input data, *viz.* observed data series spanning many years) were available.

This work contributes to clarify the issue of the relative value of classic TMY and stochastic series (long term and TMY), by comparing the internal thermal conditions of a test cell evaluated through numerical simulation using as input the various data source types. A sensitivity study is performed to find how the results depend on temperature control mode – free-floating or thermostatic – and on building parameters – window area and thermal inertia of walls. The ESP-r simulation program is used [5].

METEOROLOGICAL DATA

Previously to presenting the data used, some words are convenient on the issue of data availability. When trying to assess the quality of synthetic data in general (either obtained from selection of observed data or generated by statistical/stochastic models), there are problems both from the observed and the synthetic data availability viewpoints, which will be discussed bellow in brief.

It is hard to find observed long multivariate *hourly* time series representative of the "true" climate – in fact, that is the main reason for using synthetic data. Long term time series of temperature and humidity are relatively easy to find; the shortness of long irradiation series is the main problem. The second problem is to obtain hourly data, as many of the tabulated or digitally recorded data is available for 3-hourly, 6-hourly, or daily time steps. Finally, observed data usually displays gaps and/or contains values outside the physical ranges (outliers). When assembling TMY by the usual procedures, monthly records with gaps or with too much outliers can be discarded, but this is not so when a long time series is desired to serve as control. From the side of the synthetic data there are also limitations when trying to find data series to work with, as classic TMY are not available for all sites whereto convenient observed data exist. And models for generating multivariate time series have not yet been studied in sufficient depth for all regions and climates. An exception is the case of Mediterranean climates. All these constraints considered, the authors selected the location of Lisbon for the current study. The period selected for the study was 1981 to 1990.

Hourly irradiation data was obtained in digital support from IGIDL, located at downtown Lisbon (38.70 °N, 9.18 °W). Hourly temperature and humidity was available from Project CLIMED (Aguiar, 1998a), for a station located about 10 km apart from the irradiation station (37.77 °N, 9.13 °W). This is not though to introduce significant bias in the irradiation data. A larger drawback was that only for 1985-90 the recorded data available could be quality controlled and gap-filled to provide a reliable continuous reference hourly series (hereafter referred as "control"). Climatic estimates for the whole period could however be obtained from Aguiar (1998b).

The Lisbon classic TMY for 1981-90 was assembled specifically for building simulation purposes by the Portuguese meteorological services and civil engineering laboratory (INMG / LNEC, 1988), based on data for the same stations mentioned before. The assembling methodology applied follows the one of the EEC Typical Reference Years (*e.g.* Andersen *et al.*, 1996). In brief, every monthly series in the available pool of data is given a mark, irrespectively of the calendar year. This mark depends on closeness of its statistical parameters, *viz.* average and variance, to the statistics of the long term data for the corresponding calendar months. The "best" monthly series are then concatenated (for the current case, apparently without smoothing).

Stochastic data was generated using the CLIMED software (version 1.3). This incorporates mainly the know-how of the Projects ESRA (Aguiar and Page, 1996) and CLIMED (Aguiar, 1998b) – especially the former, as this Project focused on the Mediterranean zone and thus the model coefficients are tuned with time series data from climates similar to the one of Lisbon. The CLIMED 1.3 generator uses algorithmic chains organised in a top-down approach. Data is generated for the monthly level; then these monthly values are fed as input data for the daily level algorithms; and again the daily values are fed as input for the hourly level algorithms. At each level, the global irradiation is the master variable, and the other variables (diffuse fraction, temperature, thermal amplitude, relative humidity) are parameterised (regression models) or cross-correlated (stochastic models) with reference to it.

Microclimate tuning of the CLIMED stochastic series is allowed; the following set-ups were adopted:

i) ground albedo with the typical value of 20%;

ii) constant level horizon obstructions at 4° angular height (all radiation is diffuse for sun elevation angles bellow this value);

iii) fog occurrence implicit in the irradiation data;

iv) air turbidity corresponding to a typical seasonal profile (neither very clean or very polluted air), viz. Linke Turbidity for airmass 2 equal to 3.0 in Winter, and rising to 3.5 in Summer;

v) typical morning-afternoon asymmetry of the irradiation profiles of 3% for partially cloudy days.

The time series generator nearly exactly reproduces the climatic values that are given as input, in either of its two-modes: multi-year or TMY. A six-year series – thus with the same length of the 1985-90 period – and two TMY samples were produced. Two TMY were used so that the effects of the internal variability of the generator, due to the specific pseudo-random sequences driving the stochastic methods, could be assessed.

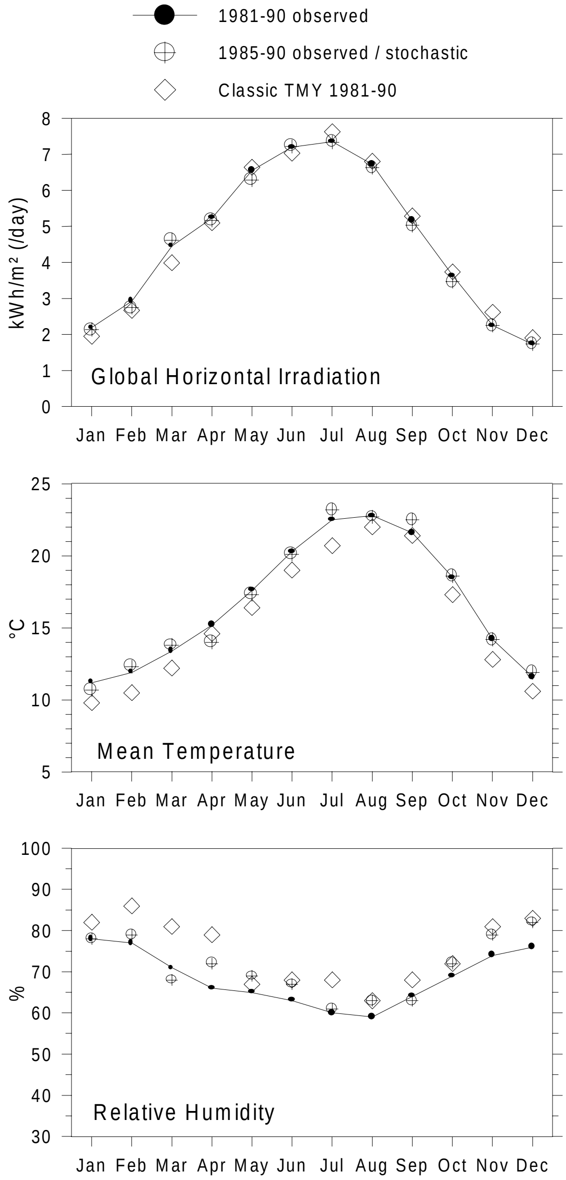

Fig. 1 – Climate of Lisbon according to various data sets.

Table 1 – Yearly mean bias and root-mean-square errors for the 1985-90 and classic TMY data (in respect to the 1981-90 climatic estimates).

	Global irradiation		Temperature		Relative humidity	
	mbe	rmse	mbe	rmse	mbe	rmse
1985 to 1990	-2%	3%	0%	3%	4%	6%
classic TMY	0%	5%	-7%	7%	9%	10%

Wind speed has not been discussed, as it was not considered for this study. This was decided for a number of reasons: stochastic models for daily and hourly wind speed and direction are not yet fit for the type of comparisons intended; wind is known to be very dependent on local conditions (building dimensions, orientation, surrounding terrain type, obstructions, etc.); in conjunction with this, observed data corresponded to non representative conditions (airport).

Figure 1 shows the seasonal profiles of global horizontal irradiation, temperature and relative humidity for Lisbon. This area is situated near the Atlantic Ocean and benefits from its moderating influence; Winters are mild and cloudy, with occasional runs of good, clear, cold weather; Summers often display long "hot spells" with cloudless days and nights. Monthly mean clearness index values (*i.e.* ratio of irradiation reaching the surface to what is available outside the atmosphere) are high, in the typical range 0.45 to 0.68; the fraction of diffuse radiation ranges from around 50% in Winter months to around 25% during the Summer. Temperature shows marked seasonality, but very rarely descending to negative values. Humidity varies in a narrow range around 70% to 85% due to the presence of the ocean. Wind speed is generally low.

The impact of using the six years of available hourly data instead of the ten years corresponding to the reference period. can be appreciated also in Fig. 1. It is verified that both sets of monthly means are very close. To the exception of a too high value in March, the radiative climate is very similar to the one of 1981-90. The seasonal temperature profiles are also very similar. Although for July and September the 1985-90 values are in excess, the annual averages are in fact the same. Relative humidity values are close, but somewhat higher for 1985-90 at some months. Table 1 quantifies these differences in yearly terms. It is recalled that the stochastic data shares nearly exactly the same 1985-90 averages as the control series.

Finally, Fig. 1 and Table 1 also enable to inspect the climatology of the classic TMY. Irradiation is reproduced a too low in March and too high in July and November, otherwise it is very close to the "true" 1981-90 climate. In fact, the yearly average irradiation is the same for both data sets. However, temperature means are always too low, except for September. Absolute humidity (not shown) is nearly the same for both data sets, but as temperature is underestimated, relative humidity here turns out too high in comparison with 1981-90.

Therefore, the most important features to consider later when interpreting the simulation results here are (i) the 2% yearly irradiation deficit displayed by the control series; (ii) the 7% yearly temperature deficit displayed by the classic TMY data set.

SIMULATION CONDITIONS

A test cell representative of a full-scale parallelepiped room was used as reference model (height: 3 m, width 5 m, length 4 m). A single glazing area faces south; no control strategy for the solar gains was implemented. Two situations were considered: a cell representing an intermediate level room, namely with no net flow of energy through pavement and ceiling; and a top-floor room with energy exchange through the ceiling. However, the results obtained were similar for both cases, thus only for the first situation (intermediate floor room) the results will be presented in some detail.

The thermal model behaviour has been numerically simulated by ESP-r also for two different levels of heat storage capacity and three different solar heat gain solutions. The external walls and floor determine the two levels of heat storage capacity, as the ceiling constructive solution has been kept the same. In Table 1 the thermal conductivity U, heat capacity C and time constant, $\tau = C / U$, are specified.

Table 1 – Constructive solutions (thermal inertia).

	U W m^{-2} K	C kJ m^{-2} K	τ h
Lower Thermal Inertia			
insulated (2 cm) double brick wall	0.76	177.6	234
brick slab floor	0.69	125.5	182
ceiling thermally insulated	0.68	169.9	293
Higher Thermal Inertia			
insulated (4 cm) single concrete wall	0.76	423.6	614
concrete slab floor	0.64	378.4	591
ceiling thermally insulated	0.68	169.9	293

The glazing areas considered correspond to ratios of glass to floor area of 20%, 40% and 60%. For the internal conditions a constant infiltration rate was assumed, 0.6 ACH, and the non-existence of internal gains (occupation, light, and equipment). Studies were carried out with thermostatic control with setup temperatures 24 ºC for the hot season and 18 ºC to the cold season; and without any temperature control, *i.e.* in free-floating temperature.

RESULTS

The results of the control simulation 1985-90 for the free-floating simulation mode are presented in Fig. 2 in terms of monthly discomfort degree-days. The profiles have similar shape for all test cell configurations. The magnitudes are highly sensitive to the window area, but less sensitive to the thermal inertia of the constructions (note however that a reasonably high insulation was already assumed for the lower thermal inertia cells).

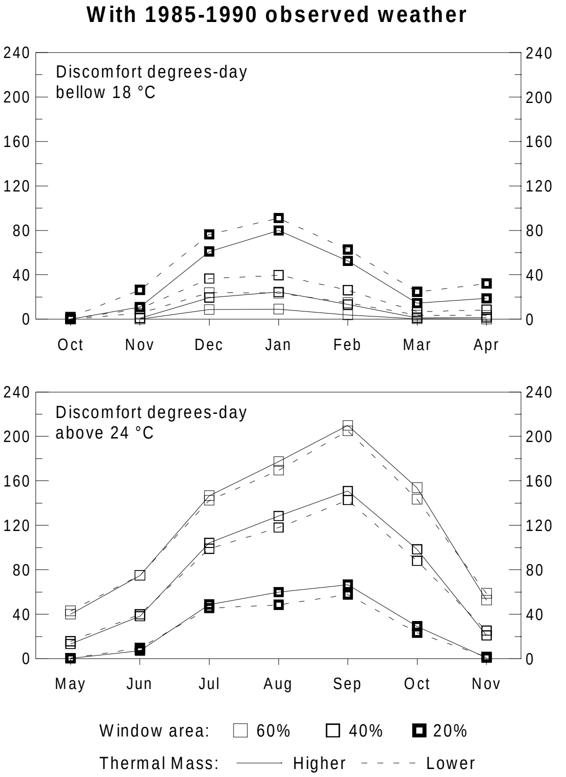

Fig. 2 – Free-floating simulation mean discomfort analysis for the control meteorological input.

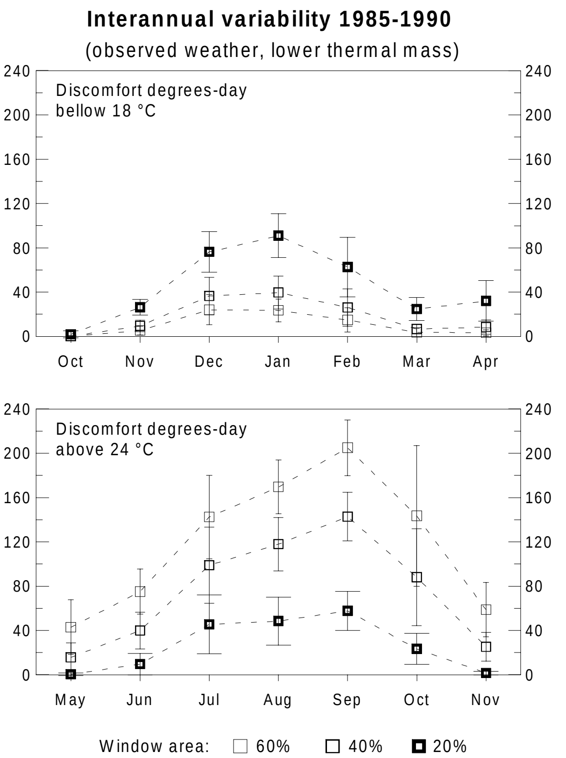

Fig. 3 – Variability of discomfort for free-floating simulations under the control meteorological input.

Inspecting Fig. 3 the interannual variability of these results can be evaluated. The bars displayed show the 67% confidence level, *i.e.* one standard deviation of single year simulations above / bellow the mean value for the six years. This interannual variability is typically in the range 20% to 70% of the mean value, for the more discomfortable months. For instance, for lower thermal inertia cells with 20% window areas, individual January discomfort estimates typically deviate ± 20 kWh in respect to an average 91 kWh; individual August discomfort estimates typically deviate ± 22 kWh in respect to an average 49 kWh. For 60% window areas, the corresponding figures are ± 20 kWh in respect to 24 kWh for January; and ± 24 kWh in respect to 170 kWh for August. October data are yet more impressive, due to the characteristic variability of Autumn weather.

These are remarkably high numbers, especially when it is considered that the interannual variability in the meteorological monthly data is only of the order of 10% for irradiation and 7% for temperature. Thus small variations in the mean level of the meteorological forcing seem to reflect non-linearly on the thermal behaviour of the simulated cells – an important fact to recall further down this paper, when analysing the relative performance of the various synthetic data sets.

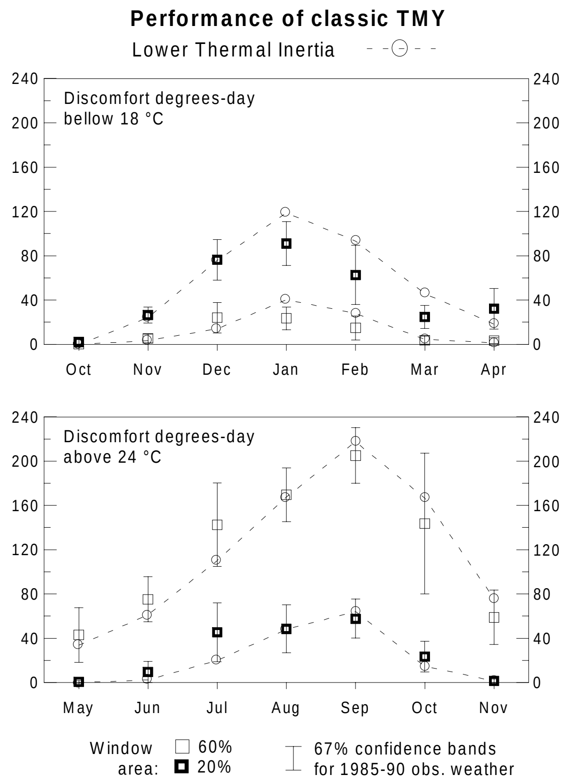

Fig. 4 – Performance of classic TMY for free-floating simulations.

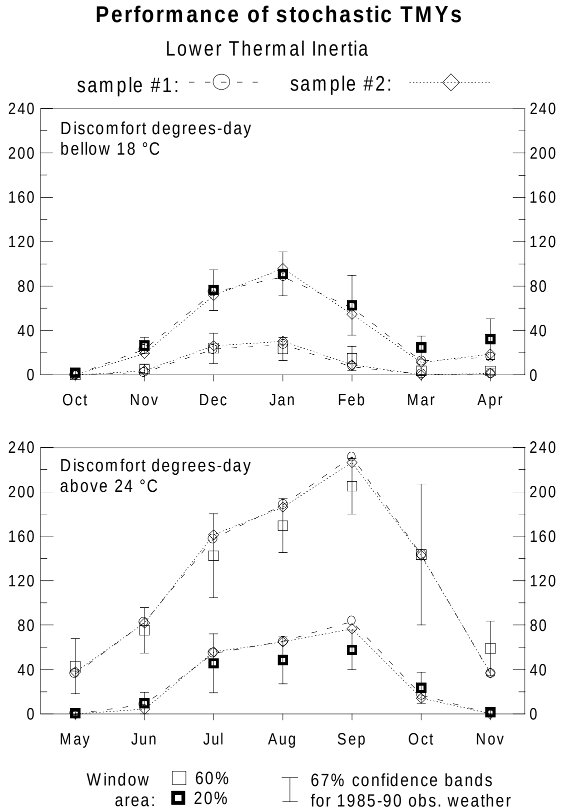

Fig. 5 – Performance of stochastic TMY data for free-floating simulations.

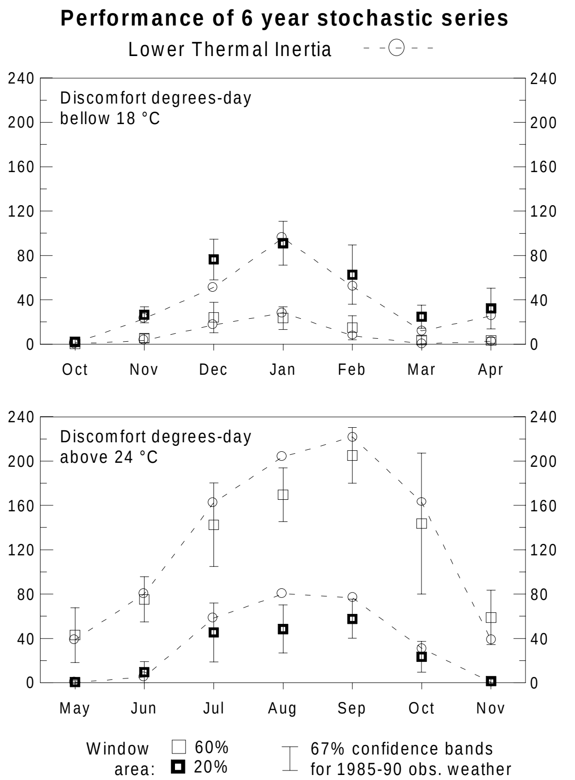

Fig. 6 – Performance of six year stochastic data for free-floating simulations.

Figures 4 to 6 and Table 3 show comparative results of discomfort estimation for the control simulation vs. the classic TMY, the stochastic TMY samples, and the six-year stochastic series. Only the cases of the lower thermal inertia test cells with window areas of 20% and 60% of pavement area are displayed, for clearness of presentation.

The classic TMY significantly overestimates discomfort for January-March (e.g. 258 degrees-day instead of the control value 178 degrees-day for a 20% window area), but for the hot season only underestimates significantly for July (e.g. 20 degrees-day instead of the control value 46 degrees-day for a 20% window area)

Table 3 – Yearly mean bias error of discomfort estimates, for the various test cell configurartions. Lowest bias cases are underlined.

Window area	20%	40%	60%
Discomfort bellow 18°C – Lower thermal inertia			
Control (degrees-day)	**320**	**128**	**75**
Classic TMY	+19%	+24%	+22%
6 yr. stochastic	-36%	-43%	-42%
Stoch. TMY #1	<u>-14%</u>	-18%	-19%
Stoch TMY #2	-14%	<u>-9%</u>	<u>-6%</u>
Discomfort bellow 18°C – Higher thermal inertia			
Control (degrees-day)	**239**	**62**	**23**
Classic TMY	+27%	+38%	+34%
6 yr. stochastic	-37%	-16%	+50%
Stoch. TMY #1	<u>-12%</u>	-13%	<u>-11%</u>
Stoch TMY #2	-15%	<u>+10%</u>	+51%
Discomfort above 24°C – Lower thermal inertia			
Control (degrees-day)	**188**	**570**	**982**
Classic TMY	-19%	-9%	-6%
6 yr. stochastic	+54%	+21%	+12%
Stoch. TMY #1	+22%	<u>+5%</u>	<u>0%</u>
Stoch TMY #2	<u>+15%</u>	<u>+5%</u>	+3%
Discomfort above 24°C – Higher thermal inertia			
Control (degrees-day)	**214**	**576**	**953**
Classic TMY	-20%	-8%	-6%
6 yr. stochastic	+47%	+20%	+14%
Stoch. TMY #1	+22%	<u>+7%</u>	<u>+2%</u>
Stoch TMY #2	<u>+17%</u>	<u>+7%</u>	+4%

The six-year stochastic time series underestimates Winter discomfort and overestimates Summer discomfort, by absolute amounts similar to those of the classic TMY. However, while in the case of the classic TMY there is some cancelation of bias from one month to another, in the case of the six-year stochastic the deviations are always of the same sign, so the yearly bias turns out double from the one obtained using the classic TMY. For the example case of the 20% window area cells, which yields 320 degrees-day for the control series, yearly bias are -117 degrees-day for the six year stochastic series and +61 degrees-day for the classic TMY.

In contrast, both stochastic TMY samples perform better than either the six-year stochastic series or the classic TMY. The Winter discomfort is well reproduced, and although Summer discomfort is still overestimated, the yearly bias are about two-thirds of those obtained with the classic TMY. For the same numerical case presented above, the bias are –44 and -46 degree-days for the stochastic TMY.

Figure 7 shows the results obtained for thermostatic simulations for the control simulation *vs.* the various synthetic data sources, for the case of the heavier thermal inertia intermediate floor cells (only yearly values). The results have shown again that the sensitivity of thermal behaviour to window dimensions was larger than the sensitivity tied to the thermal inertia. For instance, the energy demand values for the cooling season are about 50% in excess of those for the cooling season for the smallest window area, but increase ten-fold for the case of the largest window area. Any of the synthetic series gives a estimates of energy requirements within the same order of magnitude of those obtained with the control time series.

The heating requirements estimated with the classic TMY are overestimate by about 25% (e.g. 43 kWh in a target 140 kWh for the 40% window area). The results obtained with the stochastic data sets change sign with the specific cell configuration analysed. Using the six-year stochastic series yields –10% for the smaller window, but +10% for the medium size window and +38% for the largest window. The stochastic TMY sample #1 underestimates by about –13% for all window sizes. The stochastic TMY sample #2 underestimates by –16% for the smaller window, and overestimates by just +1% for the medium size window but by +21% for the largest window. For the case of the cooling needs, all the synthetic data sets consistently under- or overestimate. They all improve performance in percentual terms from smaller to larger window size, but normally because the reference energy requirement is growing. For the 40% window size, which requires 1107 kWh, using the classic TMY underestimates by -8%; using the six-year stochastic series overestimates needs by +10%; and using any of the stochastic TMY overestimates needs by +5%.

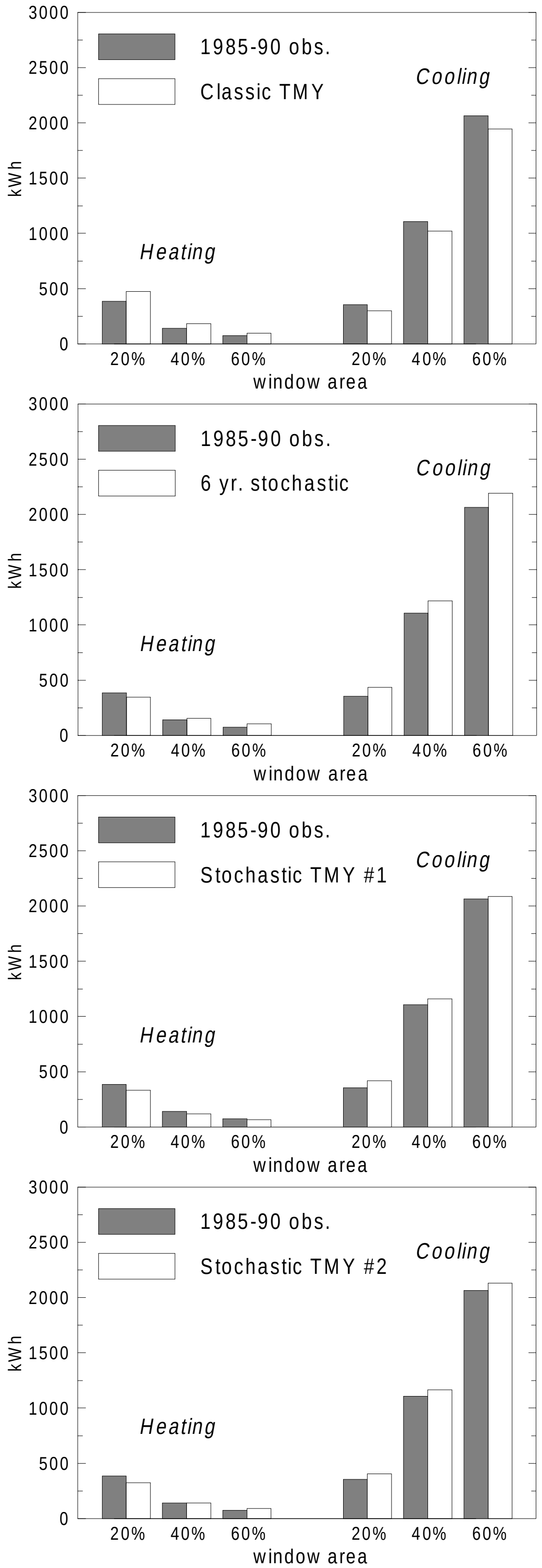

Fig. 7 – Yearly performance of various synthetic data sources for thermostatic simulations of the higher mass intermediate floor test cell.

DISCUSSION

The most salient feature of the results presented is that both the classic and stochastic TMY reasonably substitute for the real long term weather in the thermal simulations performed.

However, the stochastic TMY generally can out-perform the classic TMY data set. It can be argued that the control simulations themselves can be biased, as they correspond only to six out of the ten years which the classic TMY is deemed to represent. This brings us back to Fig. 1 wherein the various climatologies are presented. There one can see that the 1985-90 climate is very close to the 1981-90 climate – except for March, which is nearly irrelevant month for comfort and energy computations in the current case.

In fact, taking into account that the 1985-90 climate features a –2% deficit of irradiation in respect to the period to which the classic TMY corresponds, the excess discomfort and enhanced heating predicted with this data set for Winter (and opposite for Summer) would only point the other way, i.e. the classic TMY should be slightly beneficiated in the current inter-comparison of synthetic data sets.

Analisys of the seasonal profiles of meteorological data and discomfort degrees-day data reveals that the months when the classic TMY performs less satis-factorily are coincident with those for which mean temperature is underestimated in this data set. In other words, poor representation of some climatic temperature means in the classic TMY seems to be responsible for lower performance in those periods. These differences generally amount to less than 1.5 °C, but as analysed before, they seem to reflect non-linearly in the thermal performance of the buildings. A mechanism that could be implied in this effect is the frequency and duration of hot or cold spells, which is known to vary in a non linear way with mean temperature and/or irradiation levels.

It is striking in the results for the intermediate floor cell presented and discussed above, that the six years stochastic series performs worse than the other data, in particular the stochastic TMY data. However, for the top-floor simulations, not discussed here, the six-year stochastic series did perform at a similar level, if not better.

CONCLUSIONS

The ability of various types of synthetic meteo-rological data sets to substitute for the “true” long term series when estimating the thermal performance of buildings was evaluated.

The mid-latitude temperate location of Lisbon was selected for the study; multivariate 1985-90 data was obtained to serve as control input. Top floor and intermediate floor test cells were simulated with the ESP-r software. The building parameter space was explored by variation of the sole south-facing window and by using constructive solutions with two very different thermal inertia properties. Performance was evaluated by degree-days of discomfort in free-floating mode as well as by heating / cooling requirements in thermostatic mode.

The synthetic data sets used where: (i) a TMY obtained by classic methods, *i.e.* concatenation of observed monthly series best fitting the properties of long term data; (ii) long term series generated by stochastic and statistical models; (iii) TMY generated by stochastic and statistical models.

All the synthetic data sets tested could substitute for the "true" climate (long term observed series) in a reasonable way, the details varying somewhat with the specific cell configuration and temperature control mode. In most cases the stochastic TMY data was found to be the best alternative meteorological input.

When results were analysed at the monthly level, it appears that what give stochastic data its edge over classic TMY data is the ability to exactly represent monthly mean meteorological values. For instance, for certain months classic TMY data can yield nearly perfect results when compared with long term data, but for others they under- or overestimate significantly. With stochastic TMY data the results are more regular, i.e. under- or overestimating slightly at each month, and finally yielding somewhat better yearly performance estimates.

It is remarked that the (small) thermal behaviour differences predicted with the stochastic data, in respect to the simulations with multi-year observed data, could in principle be eliminated; or at least reduced very much. One way how this could be done would be by tuning the statistical and stochastic models with data for the site under inspection (*e.g.* the diffuse fraction parameterisations or the amount of cross-correlation of the pair irradiation / thermal amplitude). But this would have the drawback of tying the stochastic data to the previous existence of observed hourly data for that site – which, by the way, is the situation for the classic TMY technique. In contrast, using models for a whole region / climate, stochastic data generators have the flexibility to supply time series data for any site for which climate data can be found.

Anyway, these findings seem to support the possibility of using just one-year simulations for obtaining good estimates of the thermal performance of buildings – an almost *de facto* situation for system simulators nowadays, but by no means an evident assumption.

The test simulations performed are certainly not general enough to provide definitive conclusions. Although the sensitivity to window area and thermal inertia was inspected, results are prone to depend also on other parameters, like building occupation, internal gains, window orientation or radiation shielding strategies. Also, only very few sites, with well-known and deeply studied climates, were yet inspected in this kind of comparative studies.

However, the current work certainly constitutes one more strong piece of evidence to support the views of a bright future for stochastic data in system simulation, as a flexible and accurate alternative to the seldom available multivariate observed meteorological series.

REFERENCES

Aguiar, R. CLIMED – Climatic Synthetic Meteorological Series for the Mediterranean Area. Final Project Reports. EU – DG XII JOULE III Project No. JOR3CT960042, Ed. INETI / ITE – DER, Lisbon (1998a).

Aguiar, R. Meteorological Data for Renewable Energies and Rational Use of Energy in Portugal. Final Report of ALTENER Project No. XVII/4.1013/ Z/92-096. INETI / ITE – DER, Lisbon (1998b).

Andersen, B., S. Eidorff, L. Hallgreen, H. Lund, E. Pedersen, S. Rosenørn and O. Valbjørn, Danish Test Reference Year, TRY – Meteorological data for HVAC and Energy. Thermal Insulation Laboratory, Report No. 174 (1986).

Camelo, S., Aguiar, R., Gonçalves, H. Applicability of Synthetic Typical Meteorological Years in Building Thermal Performance. In: Proceedings of PLEA'98, Lisbon. Paper # 88 (1998).

Clarke, J., Hand, J., Strachan, P., ESP - a Building and Plant Energy Simulation System. ESRU, University of Strathclyde (1990).

INMG / LNEC, Dados Climáticos para Estudos em Edifícios. Doc. 02 – Ano Climático de Referência. Lisbon (1988).

Knight, K., S. Klein, Duffie, J., A methodology for the synthesis of hourly weather data. Solar Energy 46(2), 109-120, Pergamon Press (1991).

Mendes, C., Guerreiro, R., Pina dos Santos, Vasconcelos Paiva, Dados Climáticos de Referência para Estudos em Edifícios. LNEC and INMG, Lisbon (1988).