OFEN Project
EF - REN (90) 009

Compression of multi-year meteorological data

Final Report

Jean-Louis Scartezzini, PhD — Project manager
Marie Nygård Ferguson, PhD — Physicist
François Bochud — Physicist

SOLAR ENERGY
AND BUILDING PHYSICS LABORATORY
DEPARTMENT OF ARCHITECTURE

ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE

OFEN Project
EF - REN (90) 009

Compression of multi-year meteorological data

Final Report

SOLAR ENERGY
AND BUILDING PHYSICS LABORATORY
DEPARTMENT OF ARCHITECTURE

Abstract

Dynamic simulation programs require hourly values of solar radiation and ambient temperature forming large files, which are usually difficult to handle with available personal computers (PC). This report describes stochastic models of these variables which have been constructed to overcome this difficulty. They are based on Markov chains and autoregressive processes, determined using multi-year hourly data of both variables. A validation of the model has been carried out for five different Swiss locations. It has shown that the main statistical characteristics of these variables are reproduced by the models. A very good agreement was also obtained between results of dynamic simulations carried out using measured and synthetic data. The generalization of the method to 30 Swiss locations has been made to facilitate the transfer of these developments into practice.

Résumé

Le dimensionnement de systèmes de captage de l'énergie solaire est effectué le plus souvent sur ordinateur, par l'intermédiaire de simulations dynamiques. Celles-ci requièrent, en général, l'utilisation de fichiers de données horaires du rayonnement solaire et de la température extérieure, portant sur plusieurs années. La manipulation de ces fichiers sur ordinateur personnel est souvent difficile (volume d'information important). Afin de contourner cette difficulté, des modèles stochastiques des principales grandeurs météorologiques ont été élaborés sur la base de processus élémentaires, tels que les chaînes de Markov et les processus autoregressifs (ARMA). A l'aide de ces modèles, des séries horaires synthetiques de rayonnement et de température peuvent être produites. La validation détaillée de ces derniers a été effectuée en comparant les principales caractéristiques statistiques de ces séries avec la réalité; celles-ci se sont avérées très semblables. Afin de rendre cette méthode accessible à la pratique, les paramètres statistiques permettant son utilisation ont été déterminés pour 30 stations du réseau ANETZ, judicieusement réparties dans toute la Suisse.

TABLE OF CONTENTS

1. Introduction

This project is a continuation of previous research efforts carried out to develop stochastic models of principal meteorological data (solar radiation and ambient temperature). It was initiated to allow the transfer into practice of a new technique of synthetic data generation for dynamic simulation (hourly data files).

The method, previously described in different previous scientific reports [1, 2], is based on a stochastic approach issued from the theory of probability [3, 4]. Its most important feature is that multi-year series of hourly values of solar radiation and ambient temperature can be reduced into a few adequate statistical parameters (Markov matrixes, daily profiles). Several years of hourly meteorological data can be generated by introducing these parameters into a simple computer algorithm [1]. The main statistical characteristics of the generated data (monthly averages, distributions) are the same as for the measured data; a small personal computer (PC) is adequate for the purpose.

Following these previous developments, two main objectives have been defined for this project :

- The construction of the computer tools for processing multi-year hourly data files in order to obtain the mentioned statistical parameters.

- The processing of measured data from 30 stations of the Swiss Meteorological Institute (SMI) in order to manufacture the adequate magnetic supports for the reduced parameters (PC "floppy" diskettes).

The current report is not intending to give a complete and detailed description of the overall stochastic methodology. It will neither give a description of the content and the practical format of the "floppy" diskettes, containing the statistical parameters of the models. The reader is referred, respectively, to previous scientific reports [1, 2] and to reference [5] for these informations.

The content of this report will be limited to a short description of the stochastic models. It will be focused on the validation of the synthetic data generation approach. The latter one was carried out in two distinct ways by :

- Comparing the main statistical features of the synthetic data with measured data for some typical Swiss locations;

- Comparing the output of dynamic simulations carried out with generated and measured data on the same object.

This project has been funded by the Swiss Federal Office of Energy. Its outcome is expected to be spread to practitionners through different future knowledge transfer actions.

2. Practical computer simulation tools

2.1 State-of-the-art of current methods

Design of passive and active solar devices require the use of computer calculations and meteorological data. The range of types of programs in this field is wide : it goes from simplified methods that use mean monthly values on micro-computers to dynamic methods that need hourly meteorological data on main frames [6]. Table 1 provides the principal characteristics of these methods.

	Deterministic simulation methods		
Categories	**Dynamic Simulation**	**Correlation methods**	**Simplified methods**
Input data	hourly data files	monthly mean values	monthly mean values
Output of program	dynamic evolution of the device	average consumption	average consumption
Informatic support	personal computer	personal computer	personal computer
Computing time (one winter season)	1' to 60'	1" to 10"	1" to 10"
Facility of use	difficult	easy	easy
Potential users	researchers	praticians	praticians
Main advantages	precision, details dynamic evolution	rapidity, user friendliness	rapidity user friendliness
Main inconveniences	high computing time big data files difficult to use on PC	impossible to compute distribution and thermal comfort	impossible to compute distribution and thermal comfort

Table 1 : Current deterministic simulation techniques.

Among these methods, dynamic simulation programs usually require a longer processing time. They provide, however, outputs that take into account the time variation of the simulated system driven by the changing meteorological data. This has for example important consequences on the sizing of photovoltaic devices (PV systems) and the evaluation of extreme comfort situations experienced in passive solar buildings (eg. summer overheating).

Apart from the processing time, dynamic simulations have an other important drawback : they imply the manipulation of large data files, difficult to handle on a micro-computer. Furthermore the procedure to store and retrieve meteorological data is not always obvious : it requires the treatment of tape copies, data formatting and eventual missing information.

In the middle of the seventies, it was proposed to replace multi-year measured data by a Typical Meteorological Year (TMY) [7]. The drawback of this approach is the loss of the year to year variations of the climatic conditions, that have a significant impact on the solar system performances. Furthermore, these variations are often of great interest for the optimal sizing of a central heating plant or an energy back-up system.

Generation of synthetic meteorological data was mainly initiated this last decade. Brook and Finney proposed for instance a first model of insolation and ambient temperature [8]. They were followed by Graham et al. [9], who were able to generate daily insolation sequences using monthly mean values.

More recently Graham and Hollands [10] proposed a method to generate hourly synthetic data of solar radiation using only one monthly parameter (mean daily atmospheric transmittance). However, accounting for the strength of the method (only one parameter to start up the data generation), it is suffering from a lack of published satisfactory validation.

These considerations were the main incentives to develop stochastic models able to generate the ***horizontal global solar radiation*** and the ***ambient temperature*** using Markov chains and autoregressive processes (ARMA processes [4]). To get the adequate statistical parameters necessary to run the stochastic models, a statistically significant period of hourly data (approximatively 10 years of data [11]), must be processed. It is the only prerequisite condition regarding the original measured hourly data, giving a warranty that the generated hourly synthetic data will contain all the main statistical features of the reality.

3. Model description

The sampling rate of the simulation input data has a significant impact on the accuracy of the calculation. In general, the lower the discretization time step of the data, the better the accuracy, particularly for low inertia systems.

The proposed models are based on a time step of ***one hour***. This is the usual time step of public domain available meteorological data. This sampling rate is satisfactory to carry out dynamic simulations of passive buildings and most active solar systems. The use of the models to simulate PV systems in a very fine manner may be dubious : they usually require a shorter time step of a few minutes. However the generalization of the proposed procedure for such a time step is ***a priori*** feasible, but has not been included in the scope of this project.

In order to account for the specificity of different periods of the year, the latter has been split into 12 months. The parameters of the models have been determined for each one of these months separately. The following paragraphs describe the way these parameters have been defined.

3.1 Model of insolation

The modeled variable is the ***global horizontal solar radiation*** E_h [W/m^2]. This choice has been dictated by the fact that transposition models can be used to rebuild the global radiation in any other plane [12].

The solar radiation itself is driven by physical phenomenon too complicated to be modeled using a deterministic approach. Beside that, the deterministic trends due the daily and seasonal evolution of the solar radiation have been taken away by dividing daily and hourly insolation values by the potential solar radiation corresponding to a clear day. (I_o [MJ] and E_{ho} [W/m^2] respectively).

The main features of this model are as follows :

- The ***daily insolation ratio*** $r = \frac{I}{I_o}$ [-] (daily measured solar energy I [MJ] divided by the maximal possible solar energy I_o [MJ]) is modeled via a Markov chain.

- The ***hourly atmospheric transmittance*** $\tau = \frac{E_h}{E_{ho}}$ [-] (hourly horizontal solar radiation E_h [W/m^2] divided by the maximal possible solar radiation E_{ho} [W/m^2]) is modeled via a Markov chain.

- One 4x4 probability matrix characterizes the daily transition (among four classes of "types of day").

- Four 10x10 probability matrixes (one for each type of day) characterize the hour by hour transitions (among ten classes of transmittance).

- Five probability matrixes for each month of the year are needed to completely characterize the hourly evolution of solar radiation.

Figure 2 illustrates this model. The way the fraction of daily insolation and atmospheric transmittance are defined is given in this figure.

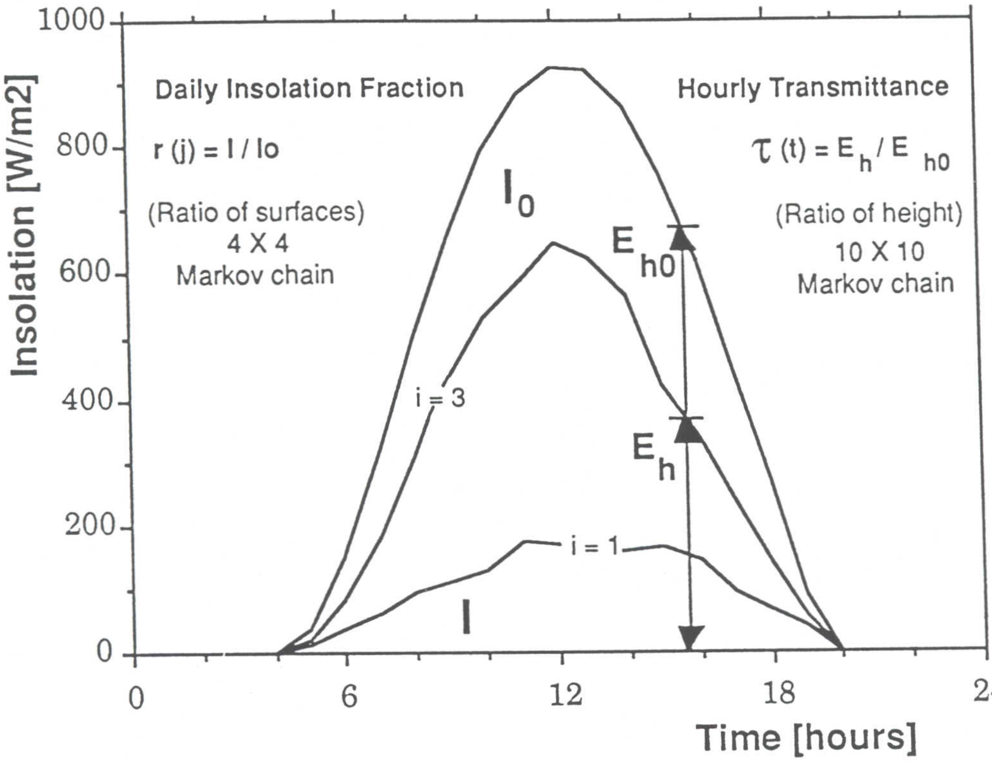


Figure 2 : Model of solar radiation. The fraction of daily insolation r [-] is defined as the ratio between the radiation received on an horizontal surface I [MJ] divided by the maximum possible radiation I_0 (clear day). The hourly transmittance is defined analogously as the ratio between the radiation received during one hour E_h [W/m^2] and the maximum possible E_{h0}.

Table 3 gives an overview of the procedure used to generate synthetic hourly insolation data files using this model. The procedure itself can easily be implemented with a few lines of computer code.

Step	Operation
1.	Construct the distribution function for r knowing the previous type of day
2.	Generate the actual type of day
3.	Select the transmittance Markov chain corresponding to the new type of day
3.1.	Build the distribution function for the transmittance knowing the previous class of transmittance
3.2.	Generate the actual transmittance τ
3.3.	Calculate the clear-sky insolation E_{h0}
3.4.	Calculate the actual insolation E_h ($=\tau \cdot E_{h0}$)
4.	Repeat 3.1. to 3.4. for each hour of the whole day
5.	Repeat 1. to 4. for the next day

Table 3 : Procedure for solar radiation hourly data generation.

3.2 Model of ambient temperature

The evolution of the ambient temperature is mainly related to the insolation even though it is a function of different variables. The temperature is largely influenced by the thermal inertia of the ground. A time-lag is observed between variations in the insolation and the ambient temperature [13]. An attenuation, which makes the ambient temperature almost independent of the instantaneous value of insolation, but highly dependent on the total insolation during the particular day is also observed [1]. This is the reason why the same separation in four "types of day", has been chosen for the model of ambient temperature.

The main features of this model are the following :

- An ***average temperature profile*** T_{ei} [°C] for each type of day i is used to model the main component of the ambient temperature
- A ***secondary component*** T_R [°C], that accounts for hourly variations is modeled via a first order autoregressive process (AR(1) process)
- The slope of the temperature profile, accounting for meteorological perturbations, is random and modeled via a gaussian process
- A mathematical technique has been used to determine the starting temperature at day to day transitions
- Different profiles have been determined for each month.

The principles of this model are described in figure 4. The different components of the temperature are also indicated.

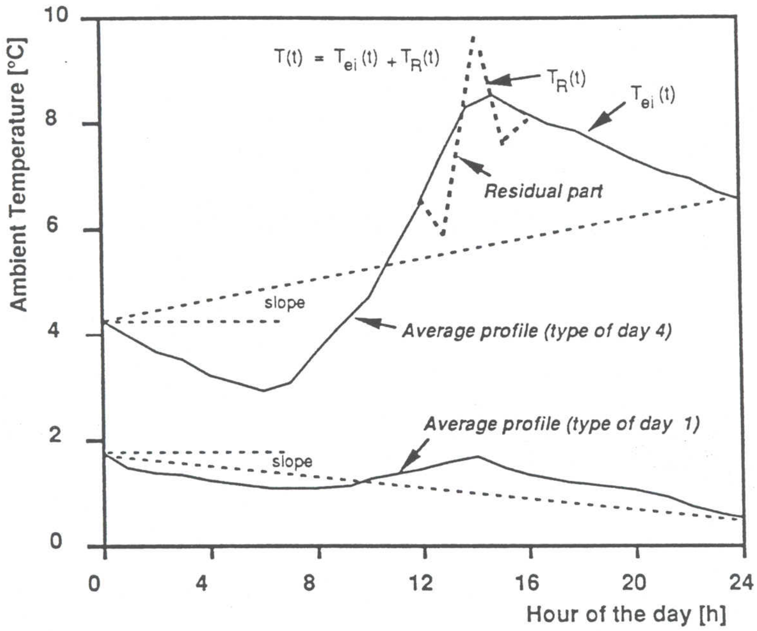


Figure 4 : Model of ambient temperature. The principal component T_{ei} [°C] (the average temperature profile) as well as the secondary component T_R [°C], are indicated in this figure.

Table 5 gives an overview of the procedure used for ambient temperature hourly data generation.

Step	Operation
1.	Generate the actual type of day
2.	Select the corresponding average profile
3.	Generate the slope (Gaussian distribution)
4.	Incline the profile according to the slope
5.	Translate the profile so it is joined to the end of the preceeding profile (maximum translation = standard deviation of the temperature)
5.1.	Generate the residual temperature
5.2.	Add the residual temperature to the profile
6.	Repeat steps 5.1. to 5.2. for every hour of the day
7.	Start again with step 1. for the next day

Table 5 : Stochastic procedure for ambient temperature hourly data generation.

4. Validation of the models

Monthly averages, standard deviations, auto- and cross-correlation functions are statistical parameters that are very relevant to judge the effectiveness of meteorological models [14]. However due to their stochastic nature and unless having exactly the same samples, it is impossible to have measured and synthetic data with identical distributions. On the other hand, the significant variations of the meteorological variables from year to year, make a rigorous comparison of the data using statistical tests (like the X^2-test) almost impossible.

The validation of the models has been made in consequence using a more empirical approach. It has been focused on the comparison of the following features :

- hourly evolution of the modeled variables
- auto- and cross-correlation of hourly values
- long term monthly averages of data
- long term monthly standard deviation of data
- distribution range of monthly average values
- outputs of dynamic simulations.

Five different Swiss locations, chosen within the five Swiss main climatic zones, were used to carry out this validation. Table 6 gives an overview of these locations.

Location	Situation	Altitude [m]	Latitude [North]	Average* Insolation [MJ/m^2]	Degree Day** (10/18°C) [K]
Pully	West of Plateau	461	46°31'	4446	2781
Sankt-Gallen	East of Plateau	779	47°26'	4194	3351
Fahy	Jura	596	47°26'	4230	3806
Montana	Alps	1508	46°19'	5155	3949
Lugano	South of the Alps	273	46°00'	4684	2090

Table 6 : Main geographical features of the five Swiss locations chosen for the validation (* Source : [15]; ** Source : [15] and [16]).

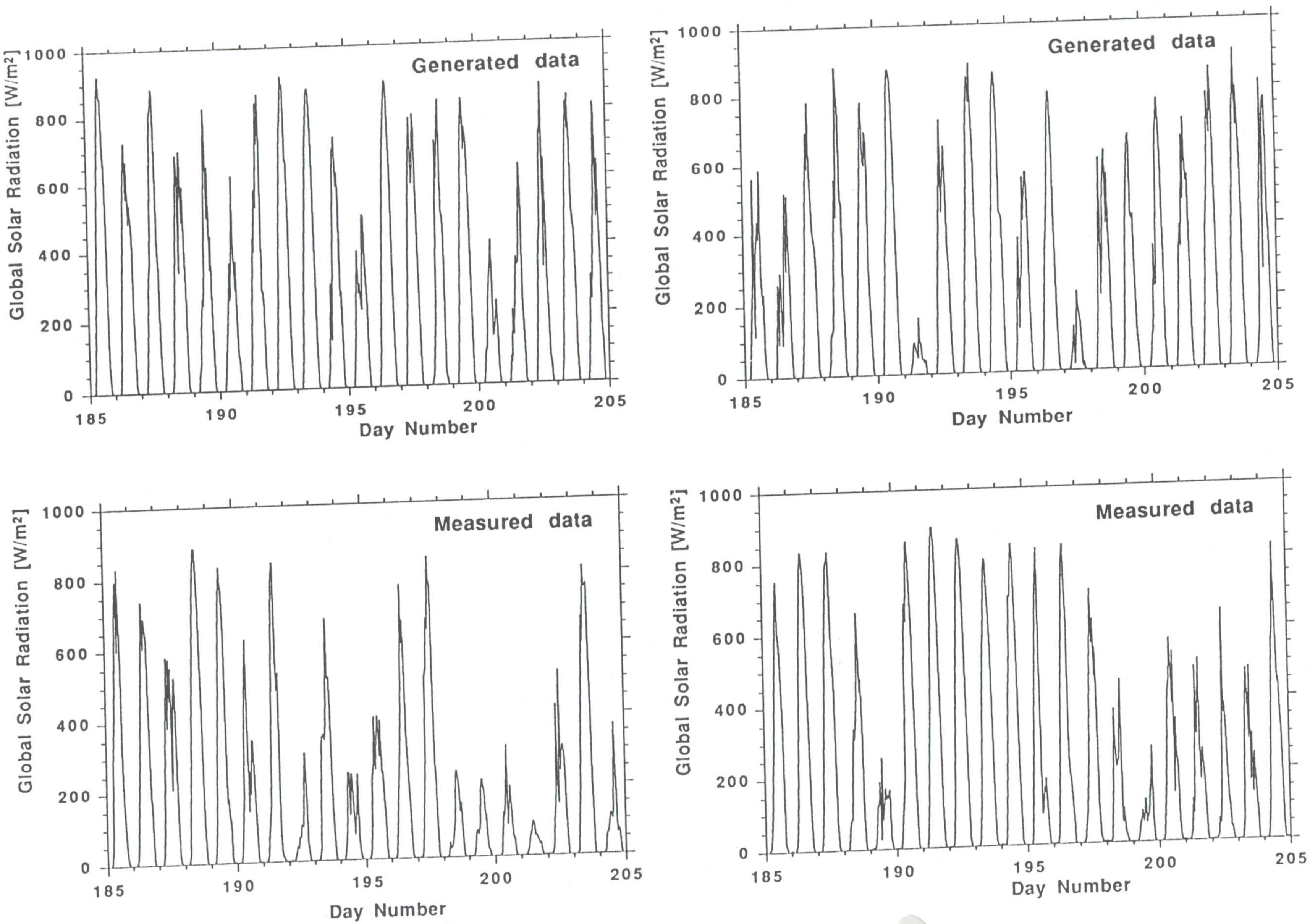
All the available measured meteorological data (generally years 1981-1989) have been used to carry out this validation. On the other hand, fourteen to twenty-eight different years of synthetic data were generated for the comparison, to be sure to have a statistically significant sample of generated data.

The following gives an overview of the main issues of the validation. In order to clarify the procedure, the different steps of the validation are presented separately.

4.1 Solar radiation

As a first step, attention has been given to the characteristics of the hourly profiles of measured and generated data. Figure 7 gives different sets of measured and synthetic data for the location of Fahy (see table 6). Two similar periods of twenty days have been chosen within the months of July and November to illustrate this comparison. Different occurrences of the same period (both measured and generated) are shown to point out the variability of the weather from year to year. The capability of the solar radiation model to reproduce this variability can be seen by comparing the data profiles illustrated in figure 7.

Fahy July


Generated data
Global Solar Radiation [W/m²]
1000
800
600
400
200
0
185
190
195
200
205
Day Number
Generated data
Global Solar Radiation [W/m²]
Day Number
Measured data
Global Solar Radiation [W/m²]
Day Number
Measured data
Global Solar Radiation [W/m²]
Day Number

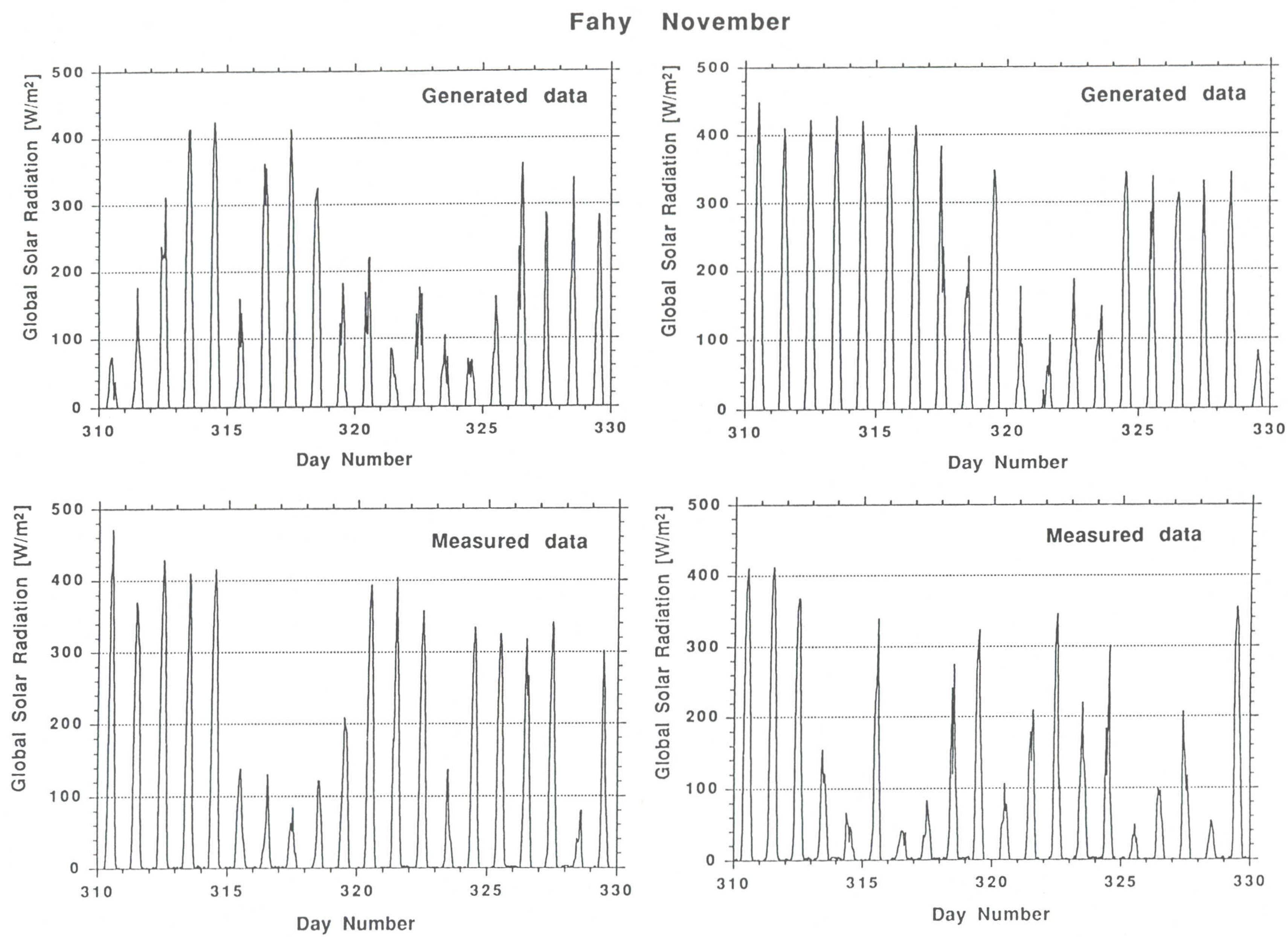
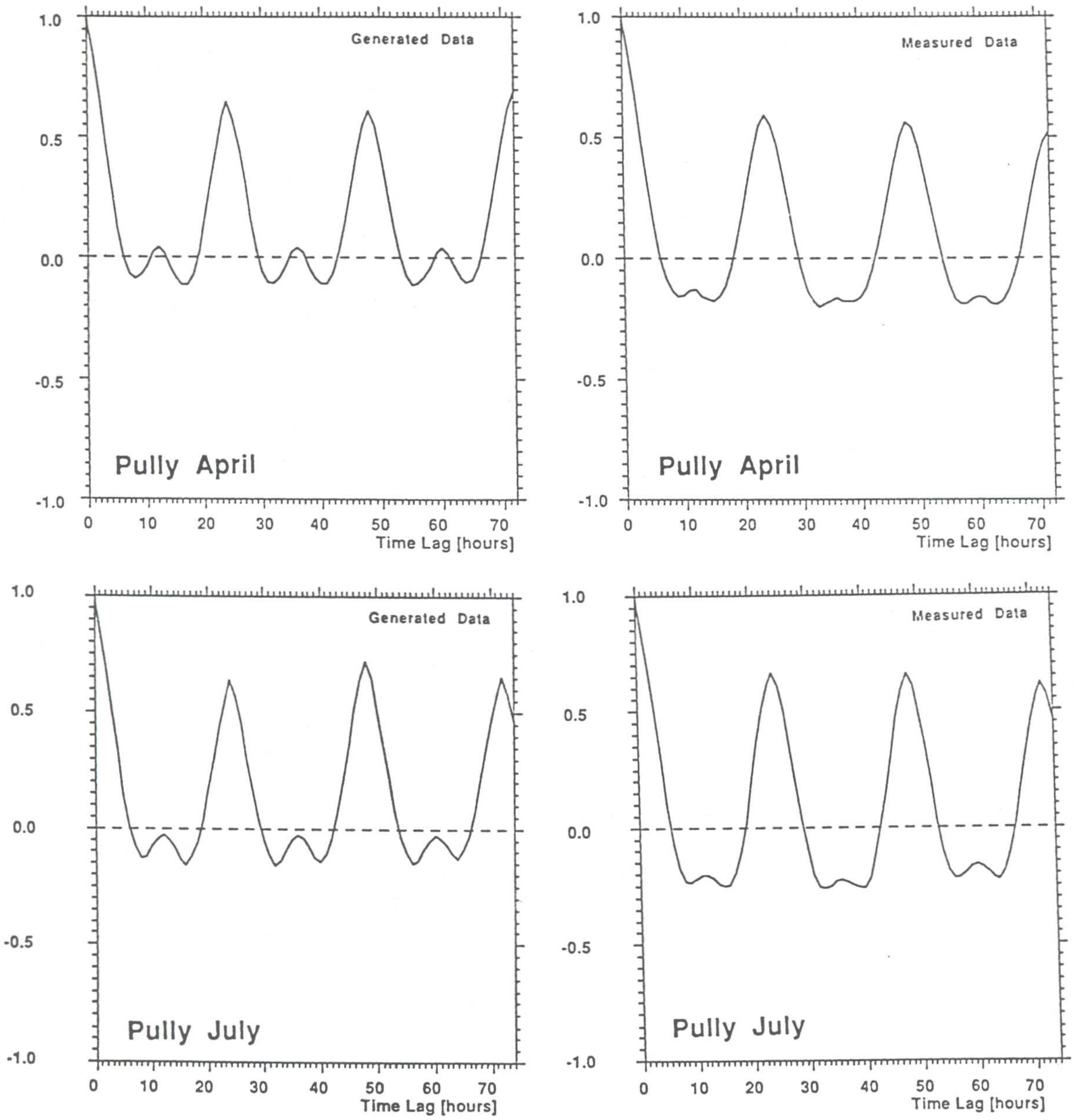


Figure 7 : Measured and synthetic data profiles of solar radiation. Location : Fahy (Jura); months : July and November.

To proceed more deeply into the analysis of the features of the hourly values, the auto-correlation functions [1] of generated and measured data have been calculated. Figure 8 shows these different functions for the location of Pully (West of Plateau) and three different months chosen within three typical periods of the year (April, July and November).



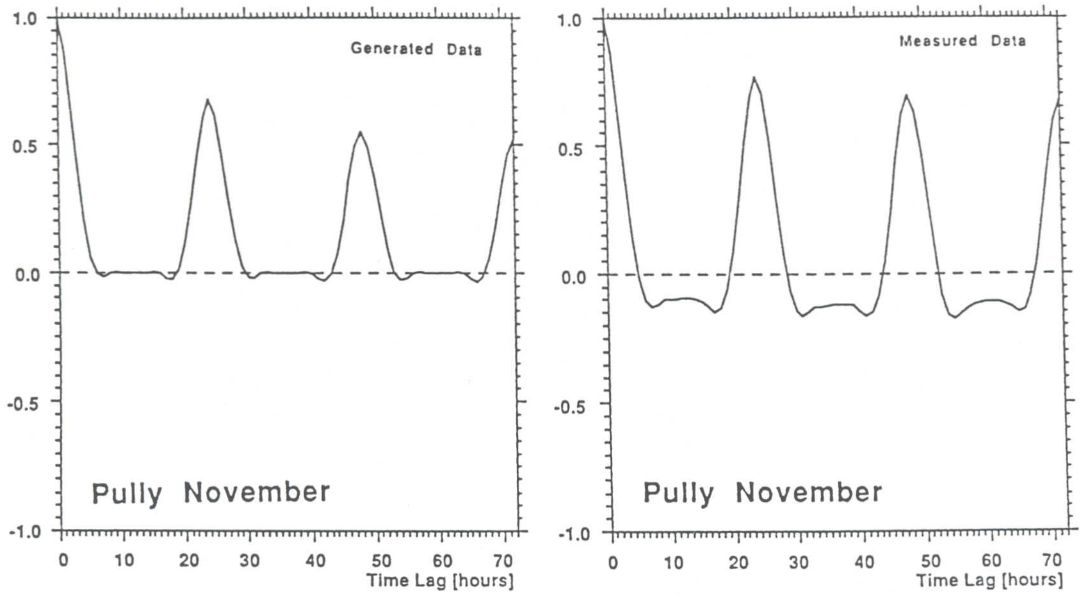


Figure 8 : Comparison of auto-correlation functions of measured and synthetic data. Location : Pully (West of Plateau); months : April, July and November.

These figures illustrate the capability of the model to reproduce the main temporal features of the solar radiation. The only notable difference between the two types of data occur at a time-lag of 12 hours (modulo 24) : it appears that synthetic data have an auto-correlation closer to zero than measured data. The reason is that night-time measured data have values that are not always equal to zero (radiation levels of 3 W/m^2, which come from artificial lights, moon light or a non-zero offset are frequent). This irrelevant discrepancy can not be assigned to the proposed stochastic model of solar radiation.

Monthly average values of insolation level, as well as standard deviations, have been determined on a long term basis to complete the validation. Figure 9 shows the results of this analysis for the case of the location of Sankt-Gallen (East of Plateau) and the months of April, July and November. Twenty eight occurrences of each months have been used to carry out this comparison. A Gaussian function has been fitted to the distribution of the twenty eight monthly insolation averages and compared to the range of distribution of the eight monthly averages obtained from measured data (years 1982-1989). It appears that both monthly values have the same scattering range, showing that the model can reproduce year to year variability with a reasonable fidelity.

The overall average calculated from both samples of monthly insolation mean values (twenty eight synthetic months and eight measured) are very close. It will be shown later that they are close enough to warranty that no bias will be introduced in the dynamic simulations carried out with synthetic data.

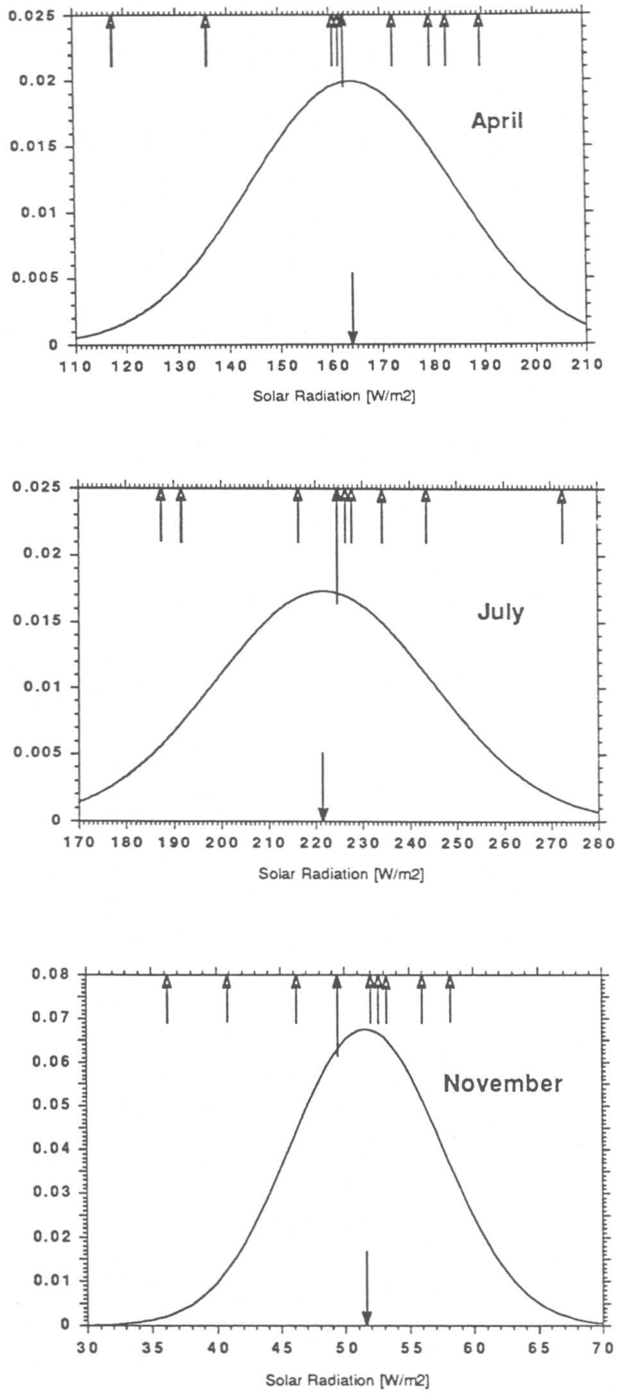


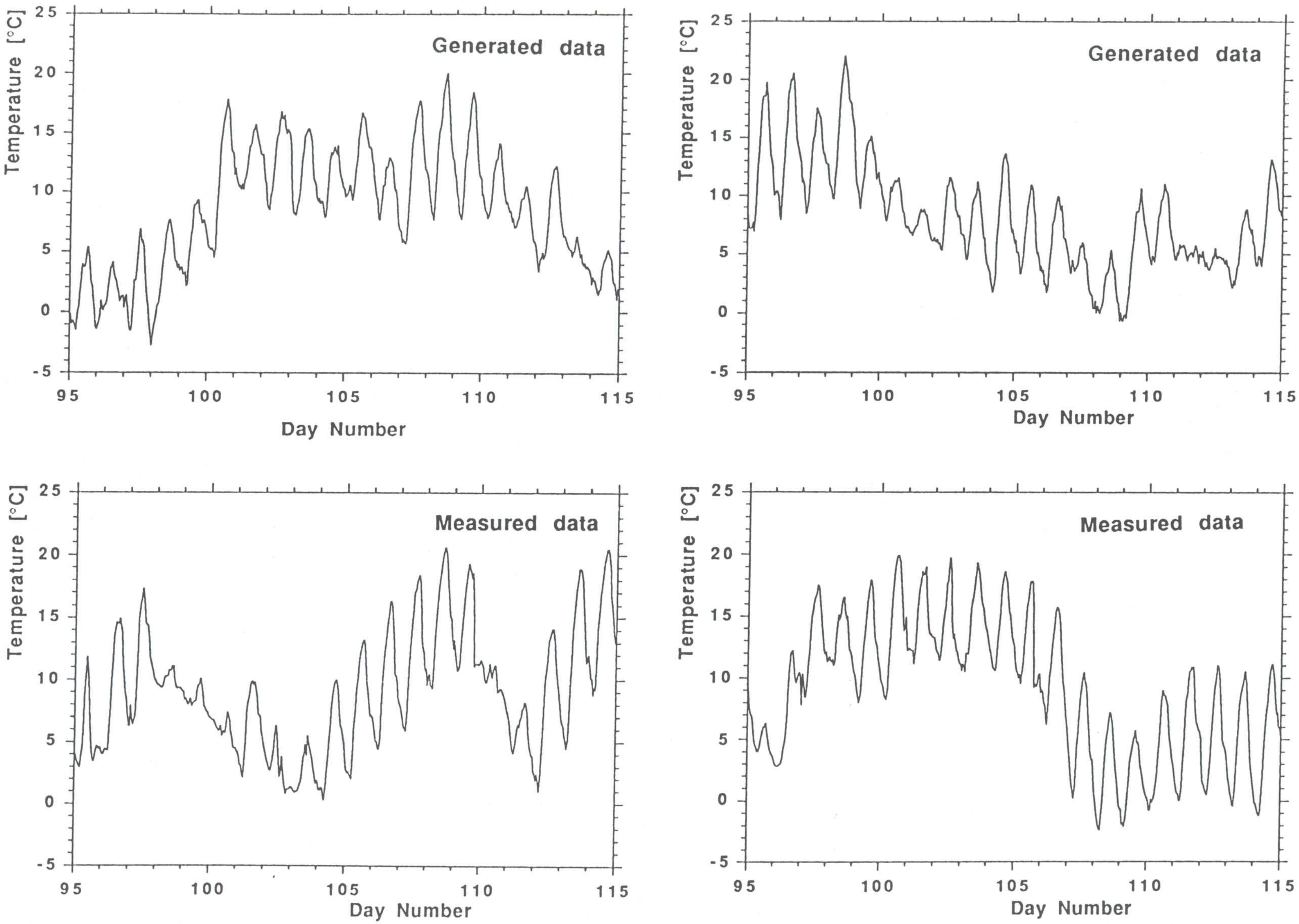
Figure 9 : Distribution ranges of monthly average values of synthetic and measured solar radiation data.
Location : Sankt-Gallen (East of Plateau); months : April, July and November.
(The white arrows show the monthly averages of measured data, the blanck upper arrow shows the overall average of measured and the other one shows the overall average of synthetic data).

4.2 Ambient temperature

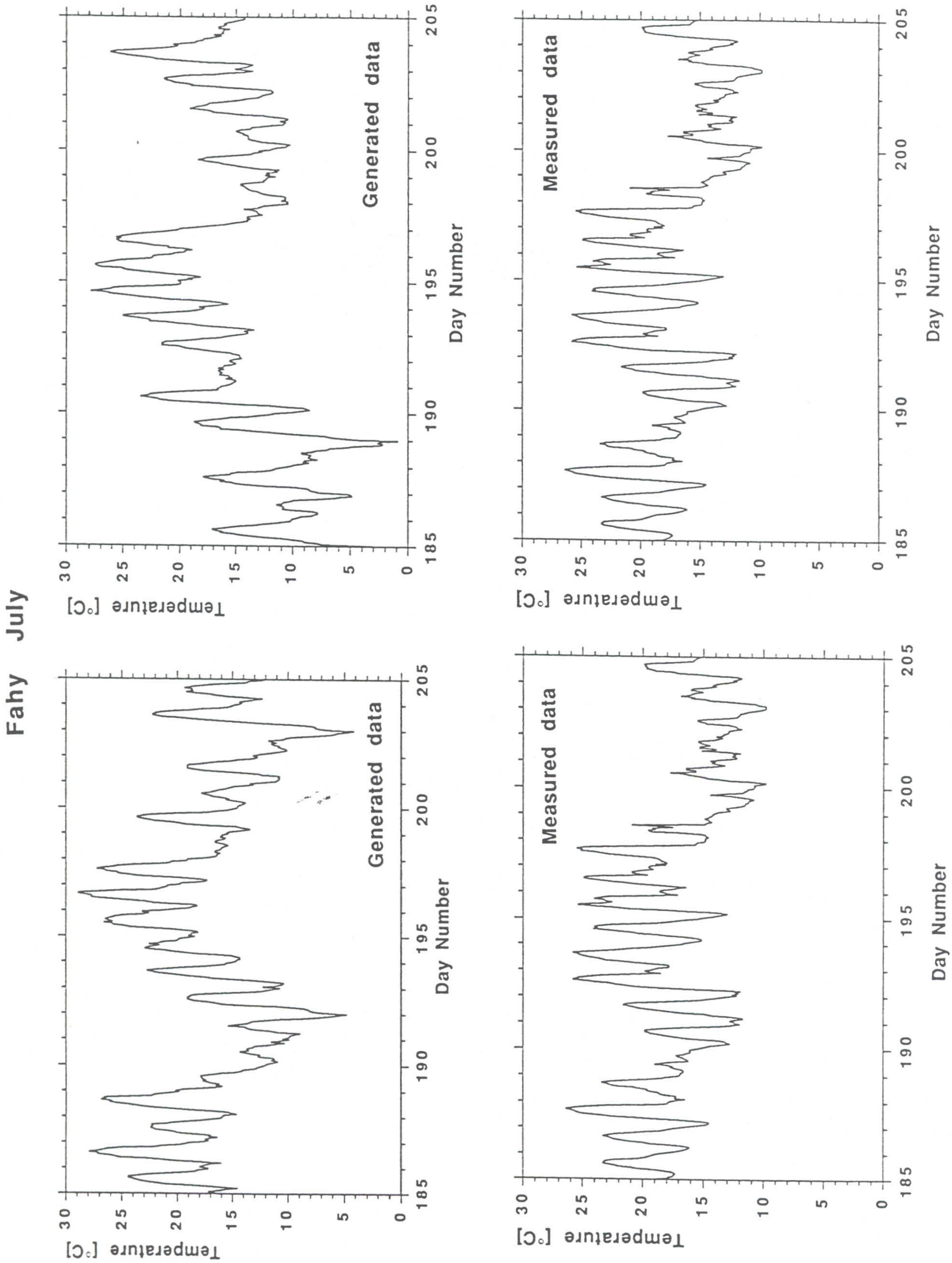
A similar validation procedure has been applied to the model of ambient temperature. Figure 10 shows different hourly profiles of this variable for the location of Fahy during three different months (April, July and November).

This figure shows the capability of the model to reproduce the characteristics of summer temperature profiles, having a well defined 24 hours periodicity, as well as winter profiles with almost no daily trend (persistent status clouds cover).

Fahy April



Generated data
Temperature [°C]
Day Number
Generated data
Temperature [°C]
Day Number
Measured data
Temperature [°C]
Day Number
Measured data
Temperature [°C]
Day Number


Fahy July
Generated data
Generated data
Measured data
Measured data
Temperature [°C]
Day Number
0
5
10
15
20
25
30
185
190
195
200
205

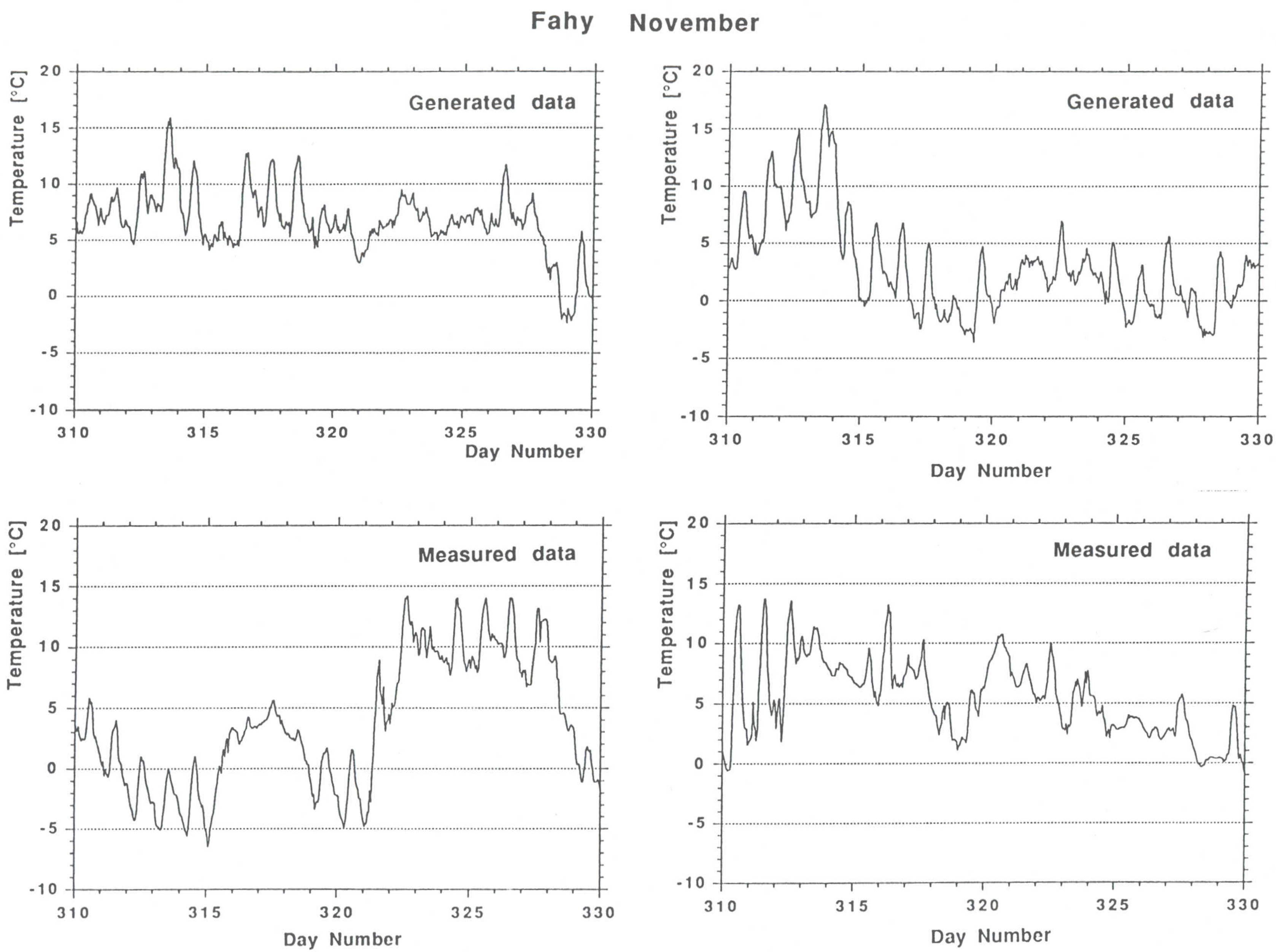


Figure 10 : Measured and synthetic data profiles of ambient temperature. Location : Fahy (Jura); months : April, July and November.

Figure 11 confirms this particularity of the stochastic temperature model. Auto-correlation functions of two different months (July and December), calculated for the location of Pully, are shown on this figure. This city is characterised by a heavy stratus clouds cover in December, explaining the difference between the auto-correlation function determined for this month and the one obtained for July. Again, the capability of the model to handle extremely different weather behaviour is confirmed by the present analysis.

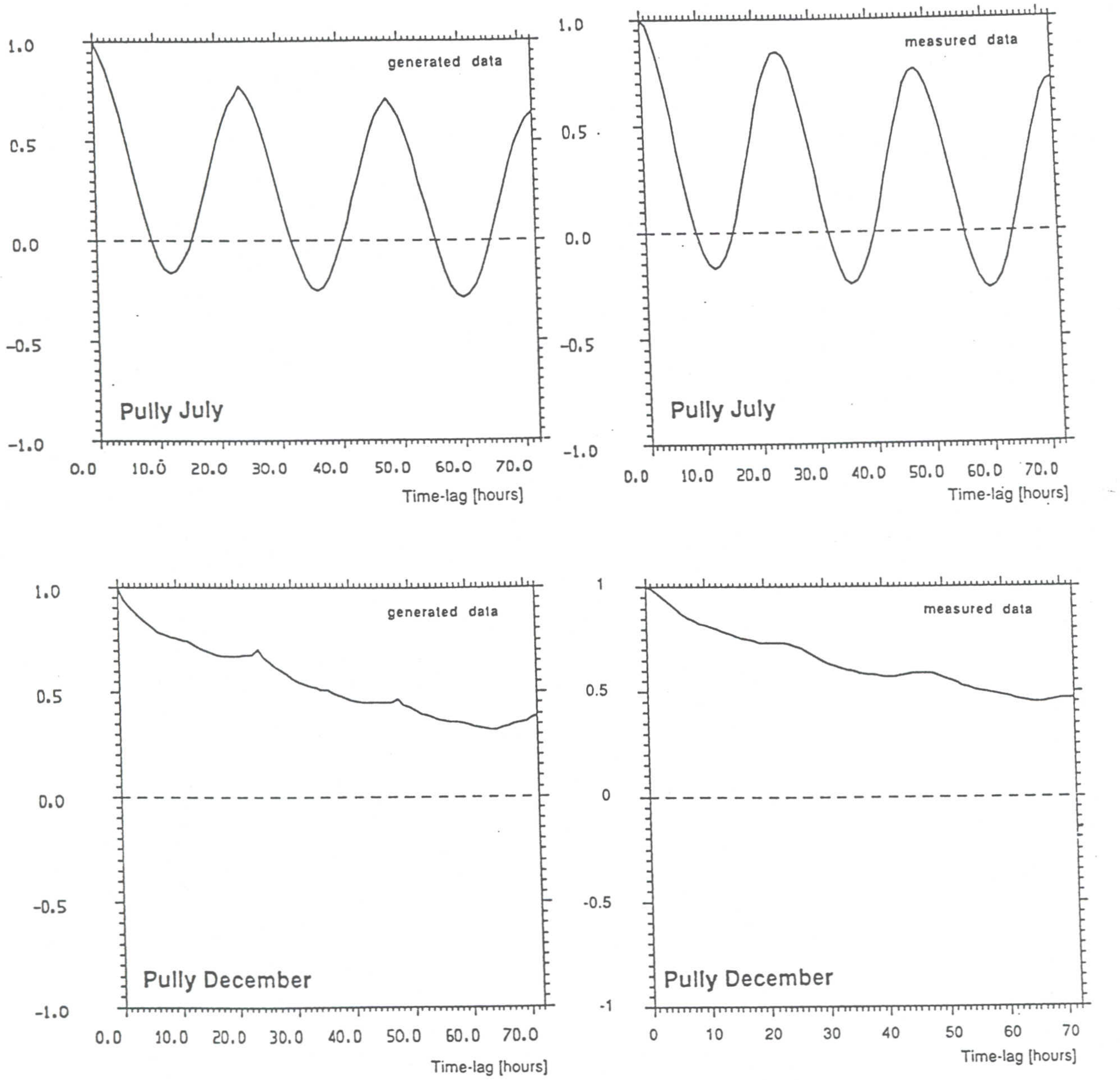


Figure 11 : Auto-correlation functions of ambient temperature calculated for two extremely different months.
Location : Pully (West of Plateau); months : July and December.

Monthly averages of ambient temperatures have also been studied in detail. A comparison of the mean values obtained from twenty eight occurrences of synthetic data and nine measured months (year 1981-1989) have also been carried out. Figure 12 shows the distribution range of both types of data (months of April, July and November) obtained for the location of Fahy (the gaussian distribution is fitted using the twenty eight monthly averages of the synthetic data). It can be seen from this figure that the distribution ranges of the two types of data are similar. The overall mean values, calculated for each month from the twenty eight and nine monthly temperature averages, are also shown on this figure. All the calculated global averages are very close.

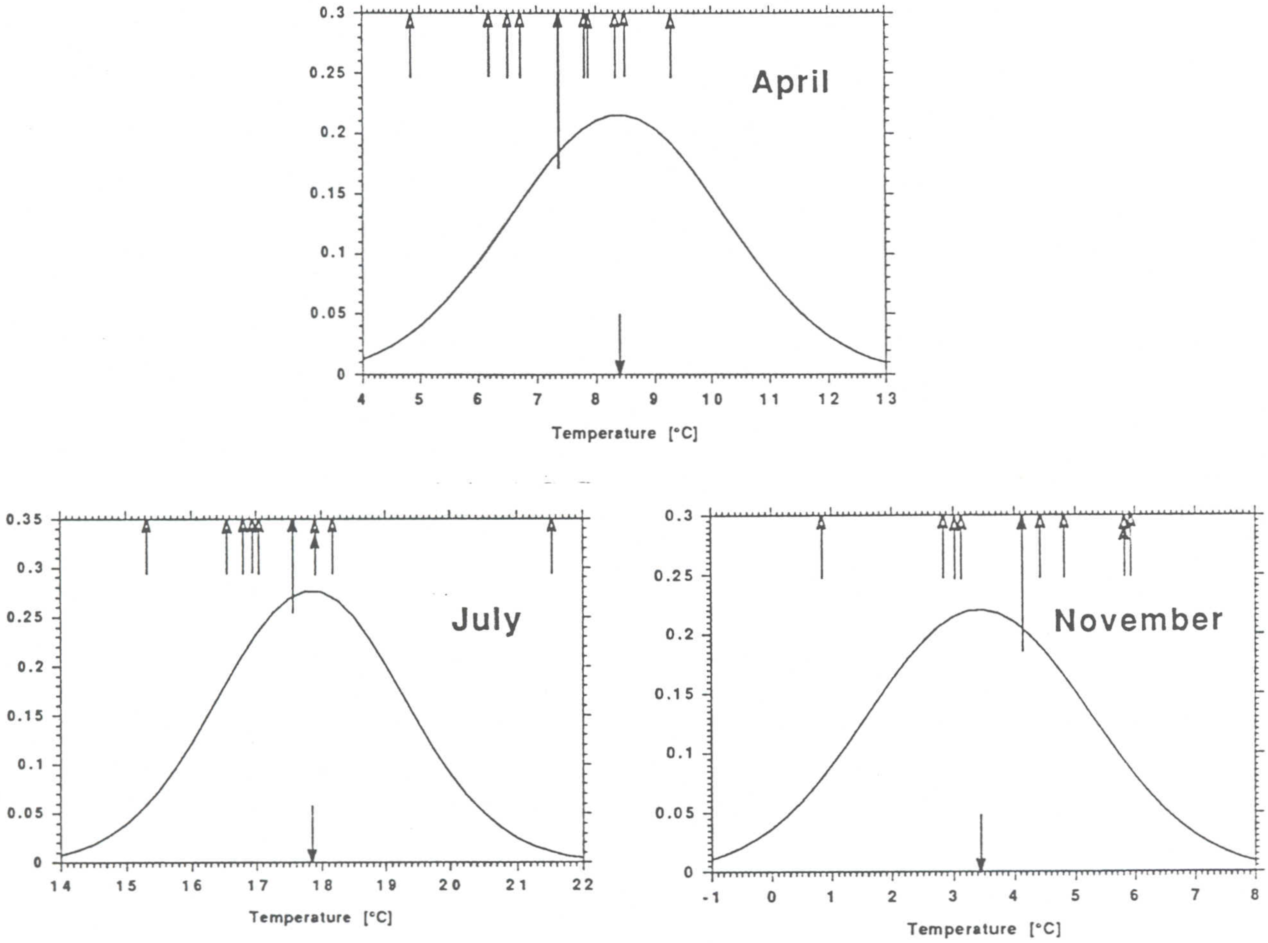


Figure 12 : Distribution ranges of monthly average ambient temperature obtained from synthetic and measured data.
Location : Fahy (Jura); months : April, July and December.
(The white arrows show the monthly averages of measured data, the blanck upper arrow shows the overall average of measured and the other one shows the overall average of synthetic data).

The mean values obtained for all the months of the year are illustrated in figure 13 (location of Fahy). It can be seen that both types of monthly average temperatures (measured and synthetic) correspond well with each other. This is also true for the standard deviations of these monthly values, which are shown on the same figure (vertical segments).

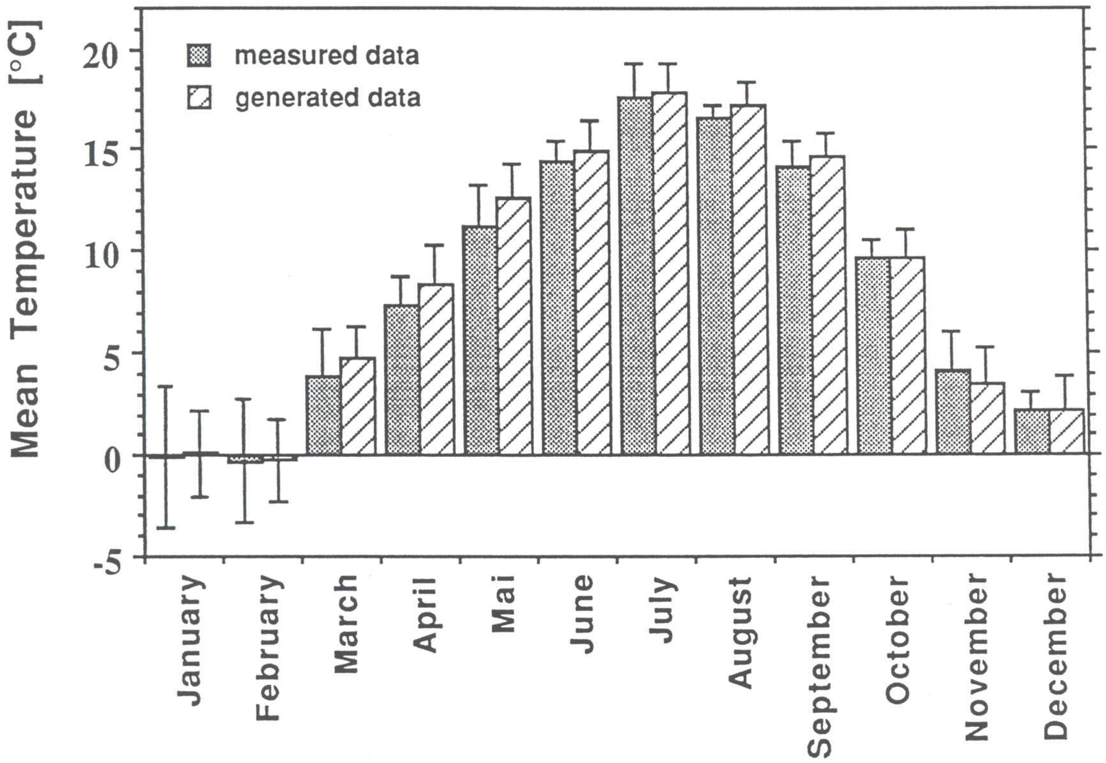


Figure 13 : Monthly averages of temperature obtained for all the year round (calculation over twenty eight synthetic and nine measured occurrences for each month). Location : Fahy (Jura).

4.3 Solar radiation and ambient temperature interdependence

The cross-correlation between the two modeled stochastic variables has been determined in order to check the ability of the models to reproduce their interdependence.

Figure 14 shows the functions obtained for two extremely different months (July and December) in the location of Pully (West of Plateau). The figure corresponding to the summer month is typical : it shows the well defined daily correlation between both variables, as well as the time lag of 1 to 3 hours between solar radiation and air temperature maxima (effect of ground and atmosphere thermal inertia).

The figure for the month of December shows the effect of the stratus clouds : the correlation is weaker for 24 hours time lags. The generated data, based on types of day, do not exactly reproduce this feature (they still have a 24 hours periodicity). This is inherent to the chosen approach (definition of types of days) and the use of Markov chains (having no memory), which are slightly in contradiction, for this month, with the strong persistence of the weather (important effect of memory).

This weakness of the model does not have any significant impact on the possibility of using it to carry out dynamic computer simulations, as will be shown in the next paragraph.

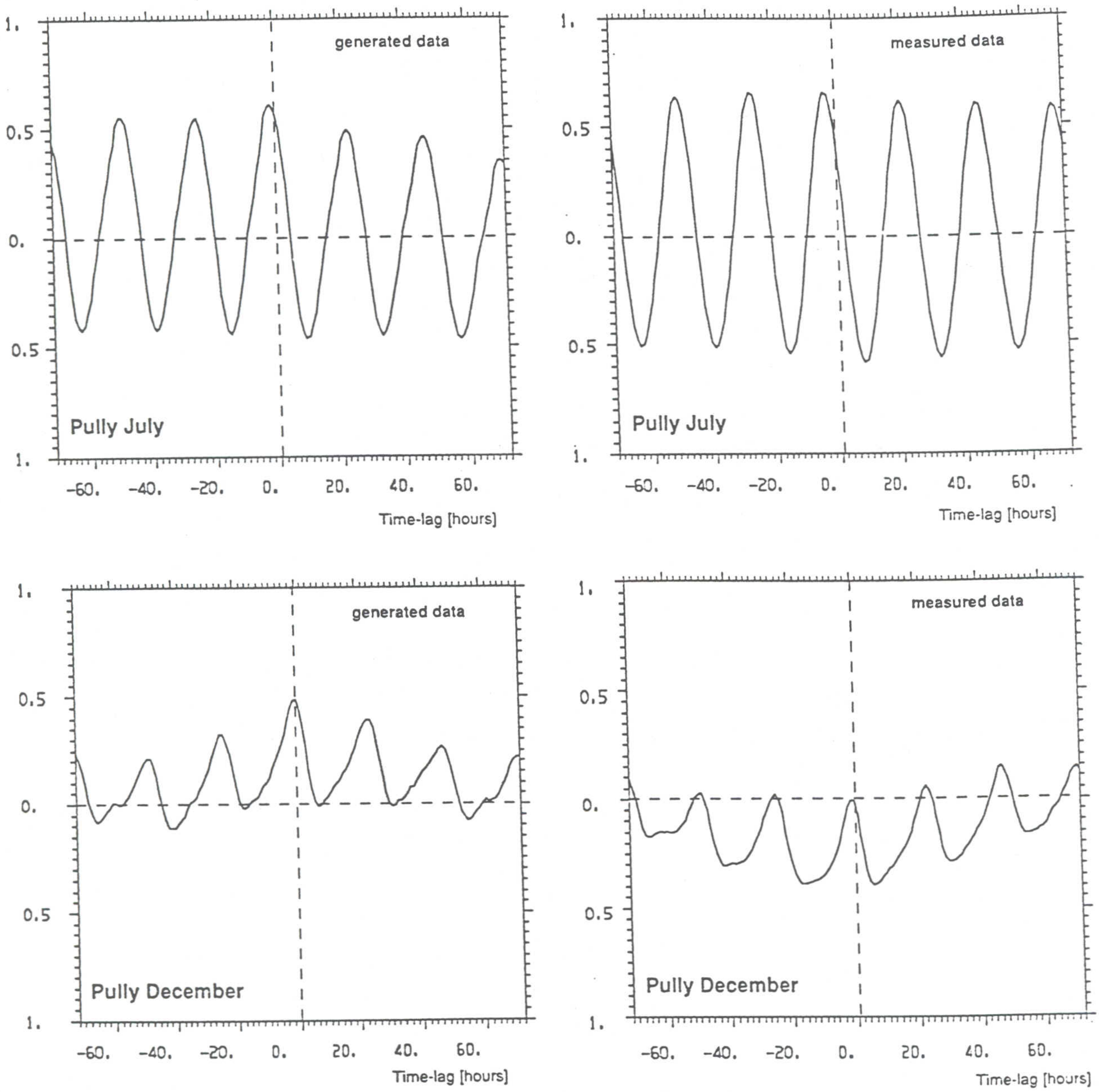


Figure 14 : Cross-correlation of solar radiation and ambient temperature determined for two extremely different months.
Location : Pully (West of Plateau); months : July and December.

4.4 Dynamic simulation

The analysis, carried out so far was devoted to the comparison of the statistical features of the measured and generated data. However, even if this analysis is necessary for checking the presence of eventual bias between the two types of data, it does not give any information about the utilizability of such synthetic data to carry out dynamic simulations. On the basis of this, the following tests have been made.

Dynamic simulations were carried out using the simulation program PASSIM [17]. Nine years of measured and synthetic data were employed as input to simulate the dynamic time evolution of a south oriented direct gain office room.

Figure 15 gives a schematic view of this object. The room is located in the corner of a building : the east and the south walls are facing the external air. The window opening area is equal to 20% of the floor area. The auxiliary heating device is an air convector, characterized by an instant response time and a maximum power of 1500 W.

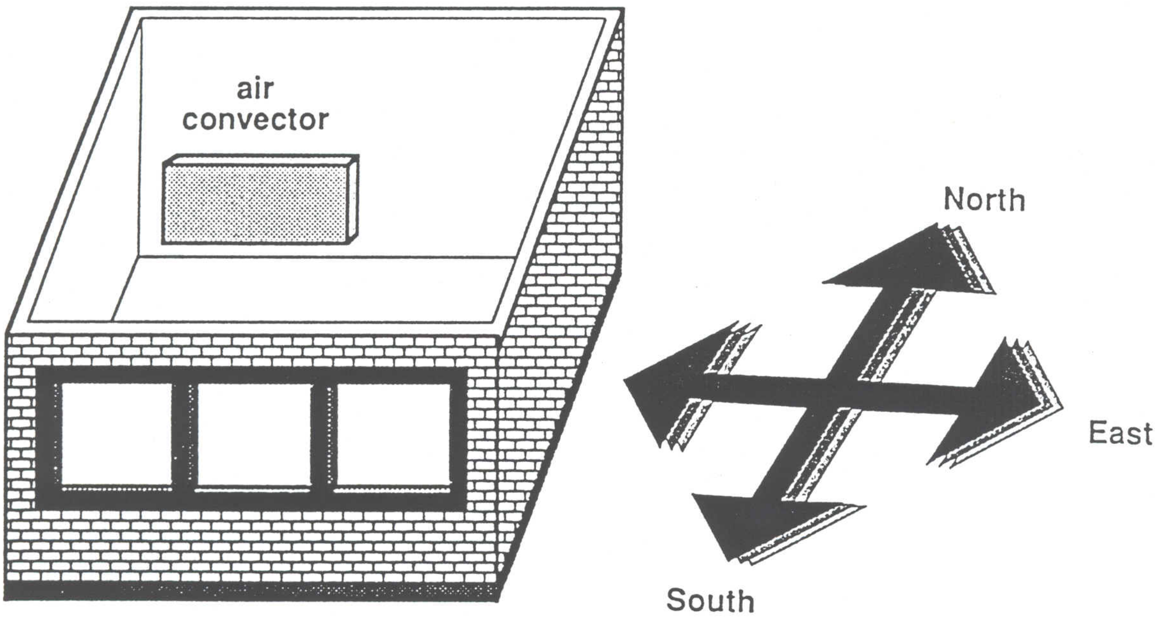


Figure 15 : View of the simulated room with 20% window opening area.

Volume :	75 m^3	Glazing :	4.9 m^2
Floor :	30 m^2	Loss coeff. :	30 W/K
Capacity :	15 MJ/K		

Response time of the Air Convector : 0 s

The room was simulated using a 33-node thermal model accounting for conductive, convective and radiative transfers by separate conductances. Three output variables were analysed in details on a monthly basis :

- The solar gains
- The auxiliary heating needs
- The indoor air temperature.

Table 16 gives an overview of the main simulation results. Monthly heating needs and solar gains, calculated using nine years of measured and synthetic data, are reported on this table.

The relative differences between both types of results are shown. These differences are smaller than 6% and 10% for the heating needs and the solar gains, respectively, determined for the colder winter months (November through February). The discrepancy for the heating needs can reach 25% for the month of October, characterised however by a very low energy consumption.

On a yearly basis, these differences are smaller than + 2.4% and + 0.7% for the heating needs and the solar gains respectively. The variability of the meteorology from year to year is expected to explain the biggest differences experienced for the monthly values.

Month	Heating needs [MJ]			Solar gains [MJ]		
	Measured Input	Generated Input	Difference [%]	Measured Input	Generated Input	Difference [%]
January	933	910	-2	646	710	+10
February	711	755	+6	764	745	-2
March	390	448	+15	1048	956	-9
April	225	194	-14	930	937	+1
Mai	119	107	-10	807	800	-1
June				782	756	-3
July				889	890	0
August				1043	987	-5
September				1136	1069	-6
October	81	101	+25	1080	1155	+7
November	479	509	+6	801	884	+10
December	784	788	+1	634	595	-6
Annual	**3722**	**3812**	**+2.4**	**10560**	**10484**	**-0.7**

Table 16 : Comparison of dynamic simulations carried out with nine years of measured and synthetic input data files. Location : Pully (West of Plateau). Measured data : 1981-1989; synthetic data : 9 occurrences.

Figure 17 illustrates the importance of the weather variability on a year to year basis for monthly values. Heating needs calculated using both approaches for the whole heating season are shown in this figure with their respective standard deviation. As expected the average monthly values are enclosed within the limits defined by the standard deviation. This is true for all months and shows the ability of the model to correctly reproduce this variability.

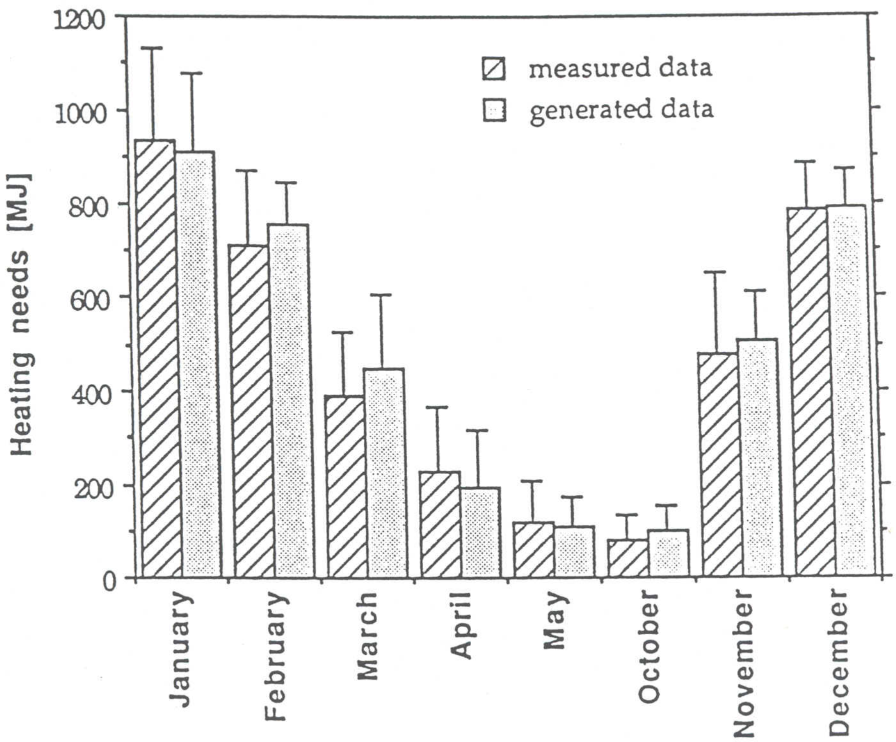


Figure 17 : Monthly heating needs and standard deviations calculated using nine years of measured and synthetic data.
Location : Pully. Period : October to May, Measured data 1981-1989, Synthetic data : 9 occurrences.

A similar analysis was made for the solar gains calculated over the entire year. Figure 18 shows the monthly average values of these gains, as well as their respective standard deviation. A similar conclusion can also be drawn regarding the closeness of the monthly values and their variability on a year to year basis.

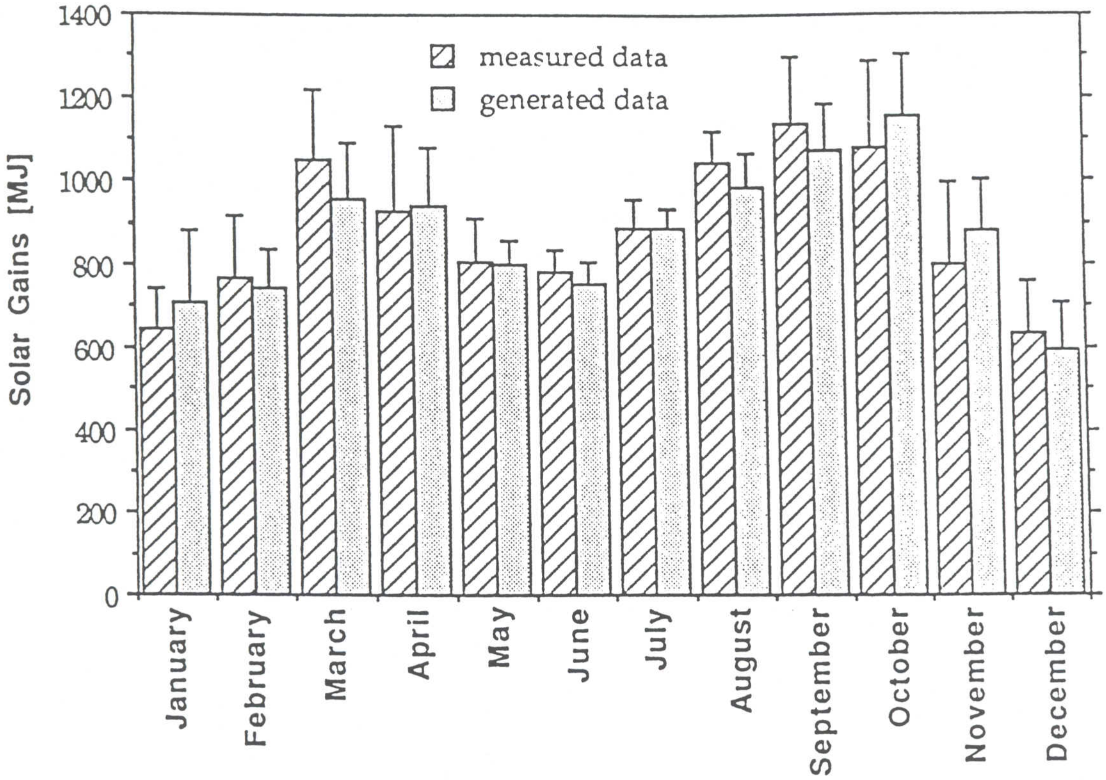


Figure 18 : Monthly solar gains and standard deviations calculated using nine years of measured and synthetic data.
Location : Pully (West of Plateau). Period : Overall year measured data 1981-1989, Synthetic data : 9 occurrences.

The procedure of comparison was extended to variables depending on the dynamic behaviour of the passive solar system. As a first step, the maximum heating power (maximum maximurum), required to the auxiliary heating plant during the month of December, and calculated using both types of input data was examined. This figure is very important for the sizing of the heating equipments.

Figure 19 illustrates the distribution of the maximal required power to the heating plant, determined using 9 years of measured and 9 years of synthetic data. The scattering of this variable between 625 and 800 W is due to the variability of the weather on a year to year basis. The distribution obtained using the two different simulation approaches are similar. The relative difference between the extreme values is smaller than 10% (730 W instead of 805 W). According to these results, it can be noted that the proposed models can reasonably be used for sizing the heating installation. It confirms the utilizability of the method as a decision and design tool.

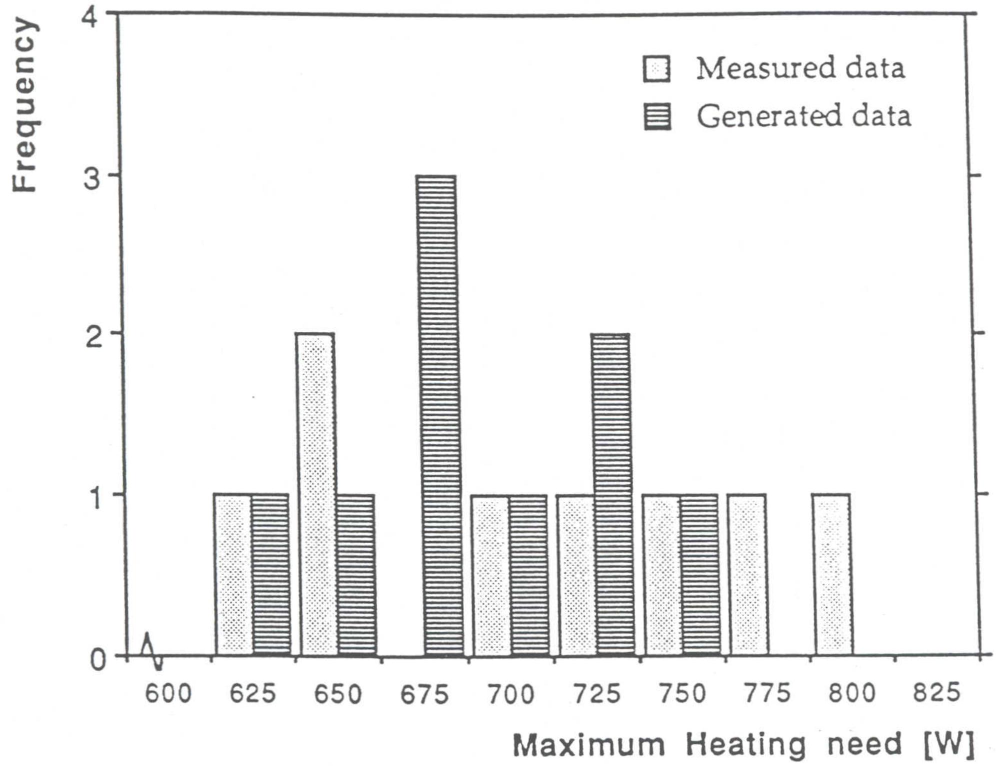


Figure 19 : Distribution of maximal required power to the heating plant calculated using nine years of measured and synthetic input data. Location : Pully (West of Plateau); month : December.

The validation procedure has been also applied to the indoor temperatures. The comparison of simulation results obtained using both approaches confirmed the sufficiency of the proposed method to determine indoor comfort situations (e.g. summer overheating). Figure 20 gives a comparison of the monthly average temperatures (as well as their standard deviation) observed using both types of data.

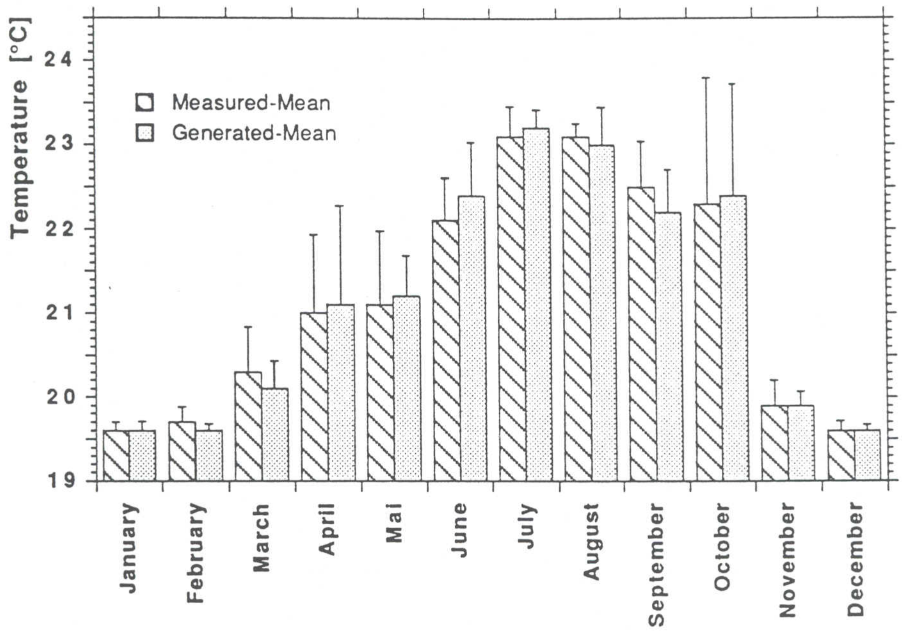


Figure 20 : Monthly indoor temperature and standard deviations calculated using nine years of measured and synthetic data. Location : Pully, period : overall year; measured data 1981-1989.

Even if they should be extended to an infinity of simulated objects and climates, the validation presented here gives a good overview of the possibility of application of the proposed models. It should not be forgotten that many other parameters may influence the accuracy of a dynamic simulation (users behaviour, randomness of physical parameters). They usually have a stronger impact than the possible inaccuracies introduced by using synthetic data instead of measured (usual accuracy of a dynamic simulation is expected to be 10%).

5. Conclusion

Dynamic computer simulations require meteorological hourly data files which are usually difficult to handle on microcomputers.

Stochastic models of solar radiation and ambient temperature have been developed to overcome this difficulty. These models use elementary stochastic processes, like Markov chains and autoregressive processes. Even though they are not phenomenological (they don't describe the physical phenomenons themselves) they are able to reproduce the hourly behaviour of the meteorological variables very similar to the reality.

To obtain these results both variables were considered at two different levels :

- the daily level
- the hourly level.

For each of them, two components were defined and modeled separately. These components are the following :

Insolation model

- daily insolation fraction (defining the type of day)
- hourly atmospherical transmittance

Ambient temperature model

- mean daily profile and slope for each type of day
- hourly residual temperature.

A validation of the proposed models was made for five Swiss locations, chosen within the five main climatic zones of the country. Different statistical features of measured and synthetic data, generated with these models, were compared. This includes the following characteristics :

- hourly evolution of the modeled variables
- auto- and cross-correlation of hourly data
- long term monthly average of the variables
- standard deviation of the monthly values
- distributions ranges of the monthly values.

Nine years of measured (1981-1989) and twenty eight years of generated data were generally used to carry out this validation.

To complete the validation, the comparison of dynamic simulations carried out with both types of data, was also made. This comparison has been focused on the main outputs of these simulations including :

- the solar gains
- the auxiliary heating needs
- the maximal required heating power
- the indoor temperature.

Monthly values of these variables, calculated using nine years of measured and synthetic data, were used to carry out the comparison of dynamic simulations. It was noted that the results fall well within the variations due to the year to year variability, confirming the applicability of the proposed models to design passive solar systems.

The statistical parameters, necessary for using the models in thirty different Swiss locations [5], have been calculated, achieving a data storage compression factor of thirty. This is expected to facilitate the transfer of this development into practice, planned within the framework of future knowledge actions initiated by the Swiss Confederation.

References

[1] J.-L. Scartezzini, F. Bottazzi and M. Nygård Ferguson, "Application des méthodes stochastiques : Dimensionnement et régulation", Rapport détaillé, Projet NEFF 349 / FN 2.331-0.86, LESO-PB, EPFL (1989).

[2] J.-L. Scartezzini, F. Bottazzi and M. Nygård Ferguson, "Applying stochastic methods to building thermal design and control", Synthesis Report, LESO-PB, EPFL (1989).

[3] Kemeny and Snell, "Finite Markov Chains", Springer Verlag, New York, NY (1986).

[4] G.O. Box and G.M. Jenkins, "Time Series Analysis : Forecasting and Control", Holden Day, San Francisco, CA (1970).

[5] F. Bochud, M. Nygård Ferguson and J.-L. Scartezzini, "Description des fichiers", LESO-PB / EPFL, (1990).

[6] P. Brejon, "Les outils de conception thermique des bâtiments en France", Proc. des 3èmes Journées Internationales de Thermique, 2, 459, 13-16 April, Lyon, France (1987).

[7] S.A. Kein, P.I. Cooper, T.L. Freeman, D.M. Beekman, W.A. Beckman and J.A. Duffie, "A Method of Simulation of Solar Processes and its Applications", Solar Energy, Vol. 17 (1), pp 29-37 (1975).

[8] M.J. Brook and B.A. Finney, "Generation of Bivariate Solar Radiation and Temperature Time Series", Solar Energy, Vol. 39 (6), pp 533-540 (1987).

[9] V.A. Graham, K.G.T. Hollands and T.E. Unny, "A Time Series Model for K_t with Application to Global Synthetic Weather Generation", Solar Energy, Vol. 40 (2), pp 83-92 (1988).

[10] V.A. Graham and K.G.T. Hollands, "A Method to Generate Synthetic Hourly Solar Radiation Globally", Solar Energy, Vol. 44 (6), pp 333-341 (1990).

[11] R.W.R. Muncey, "Heat Transfer Calculation for Buildings", Applied Science Publishers, London (1979).

[12] T.M. Klucher, "Evaluation of Models to Predict insolation on Tilted Surfaces", Solar Energy, Vol. 23 (2), pp 111-114 (1979).

[13] H. Madsen, "Statistically Determined Dynamical Models for Climate Processes", Licentiatafhandling DTH, Lyngby, Danemark (1985).

[14] B.J. Brinkworth, "Autocorrelation and Stochastic Modelling of Insolation Sequence", Solar Energy, Vol. 19 (4), pp 343-347 (1977).

[15] P. Bremer and M. Heimlicher, "Meteonorm, Energie Solaire, Données", OFEN, Berne (1986).

[16] A. Faist et al., "Le Soleil, Chaleur et Lumière dans le bâtiment", SIA/OFEN D056, Zurich (1990).

[17] N. Morel, "Passim Version 3", Manuel d'utilisation, GRES-EPFL, Lausanne, Switzerland (1984).