NEW TECHNIQUES FOR VALIDATING BUILDING ENERGY SIMULATION PROGRAMS

D. Bloomfield[1], Y. Candau[2], P. Dalicieux[3], S. Delille[3], S. Hammond[1], K.J. Lomas[4]
C. Martin[5], F. Parand[1], J. Patronis[4], N. Ramdani[2]

[1]The Building Research Establishment, Garston, Watford, Herts WD2 7 JR, U.K.
[2] Laboratoire d' Energétique et de Thermique Industrielle de l' Est Francilien - URA CNRS 1508, University Paris 12, 94000 Créteil France.
[3]Research Centre, Elecricité de France, ADEB, 77250 Moret-sur-Loing, France.
[4]ECADAP Centre, Institute of Energy and Sustainable Development, De Montfort University, Leicester, LE1 9BH, U.K.
[5]The Energy Monitoring Company, Unit 3, Chapel Court, Newport Pagnell, Bucks MK16 OEW, U.K.

ABSTRACT
An Anglo-French collaborative project is presently under way with the dual purpose : a) to collect high quality data sets for the empirical validation of thermal simulation programs of buildings and b) to develop and test advanced data analysis techniques. The aim is to identify and remedy the causes of poor predictive -performance in thermal programs of buildings. Two data analyses are presented in this paper: one is a time-domain technique developed by the Energy Monitoring Company in UK, and the other is a frequency-domain technique developed by the University Paris 12, France. Both techniques have been applied to empirical validation results from four state-of-the-art thermal programs, and the derived error diagnoses are discussed.

1. INTRODUCTION

Detailed thermal simulation programs (DSPs) are extremely powerful, relatively inexpensive, and potentially invaluable in the design of energy efficient and environmentally superior buildings. Unfortunately, there is still considerable doubt about the absolute accuracy of their prediction.

A number of projects have sought to develop methods for validating DSPs. Three main techniques are recognised, namely: analytical verification [1, 2]; inter-model comparisons [3, 4]; and empirical validation [5, 6].

Empirical validation is particularly favoured by program developers and users because of its ability to test the programs under conditions which reflect those experienced in the 'real world'. An Anglo-French collaborative project is under way which aims to:

i. develop new empirical validation data sets for testing DSPs;
ii. develop and test advanced data analysis techniques to assist in identifying the reasons for poor performance of DSPs;
iii. test the data and techniques using state-of-the-art DSPs;
iv. produce a validation database to store, document, and distribute validation data in a clear and consistent manner; and
v. document the technical basis of the programs being tested using the Model Information System (MIS) software package developed by the UK Building Research Establishment [7].

The project is being carried out by the UK Building Research Establishment (BRE) and Electricité de France (EdF). Also participating in the project as sub-contractors are in the UK: the Environmental Computer Aided Design and Performance (ECADAP) Centre in the Institute of Energy and Sustainable Development at De Montfort University Leicester, and the Energy Monitoring Company (EMC); and in France the Laboratoire d' Energétique et de Thermique Industrielle de l' Est Francilien (LETIEF).

This paper relates to aims (i) to (iii). The performance of four DSPs are compared with measurements taken in simple test rooms. The results are analysed using two different advanced data analysis techniques and the results of applying these are discussed.

2. THE NEED FOR DETAILED ANALYSIS TECHNIQUES

The analysis of results from empirical validation of thermal programs falls into two distinct parts, addressing the questions:

- is the level of agreement between predictions and measurements adequate, and can the program therefore be deemed valid ?
- what are the reasons for any discrepancies between predictions and measurements ?

Although the role of empirical validation is often seen in terms of answering the first of these questions, it is the second question which is ultimately the most important. Only by identifying the causes of poor predictive performance can the programs be refined.

Identifying the causes of poor performance of thermal simulation programs is complicated because buildings exposed to real climates and operating conditions are dynamic systems: a prediction error at any given time may be the result of present and/or past excitations. In addition, the forces which drive a building (for example external temperature/and solar radiation) have a complex structure both within themselves and between each other: they are both auto-correlated and cross-correlated.

Early attempts to diagnose the reasons for poor predictions of thermal programs relied on simple graphical comparisons between simulations and measurements. Often the dynamic effects described above were eliminated by considering only long-term average results, with consequent loss of information about the short-term performance of the program. The use of cross-correlations between simulation errors and model inputs has been extensively explored elsewhere (e.g. [8]), but this technique was often defeated by the relationships between the inputs. For example, if a program contained a fault in its processing of solar radiation, there would indeed be a cross-correlation between the prediction error and solar radiation, but, because external temperature is correlated with solar radiation, a cross-correlation will also appear between a prediction error and external temperature, and the technique cannot isolate the true source of the error.

Two techniques are described in this paper (in sections 6 and 7) which address these problems in different ways.

3. THE EMC TEST ROOMS

The first part of the collaborative project used data from the EMC test rooms (Figure 1) which were used in the recent International Energy Agency (IEA) empirical validation exercise [5, 9, 10] and previous studies e.g. [11]. The rooms have been subject to extensive investigations and quality assurance checks, including dismantling one room to check its construction; hence the construction, and data collected from the rooms are known to be reliable.

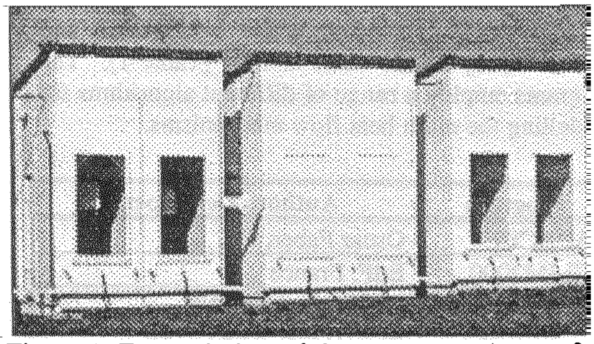

Figure 1: External view of the test rooms (rooms 0 to 5 lie from left to right).

The test site is approximately 70 Km north west of London (52.07° N, 0.63° W). Each pair of test rooms is separated by a heavy insulating party wall, and there is a roof space above. Previous publications [e.g. 9, 10, 11] have described the rooms and their characteristics in detail, so only an overview is given here. The outer shells are of stud frame construction covered by plasterboard and with a concrete slab floor. The monitored spaces were well insulated and extremely well sealed to reduce infiltration to less than 0.05 air changes per hour. Rooms 0 and 1 featured a double glazing system on their south facing facades, Rooms 2 and 3 had an opaque infill panel and Rooms 4 and 5 had a single glazing system. This paper focuses on analysing data for Rooms 0 and 1.

During the period from 13/3/1990 to 1/5/1990 (50 days) all the rooms were heated by an oil-filled electric radiator, operated at the same pseudo-random sequence (using a basic one-hour switching interval). In this sequence the state of the heater (on or off) is decided, at random, at the start of each hour. The heat output from the radiator in Rooms 1, 3 and 5 was specified as 60% radiant and 40% convective. In the remaining Rooms 0, 2 and 4 a 100% convective heater output was assumed. The dynamics of the heat source were identified as a first-order system with a time constant of 22 minutes [12].

The climate data required by the DSPs in this work were supplied at hourly intervals along with the heater power inputs - which would also become an input to the DSPs. The key geometrical and thermo-physical properties of the rooms were also defined [12]. The key building performance parameters measured were the room air and surface temperatures. It is against these measurements that the performance of the DSPs is compared.

4. THE PROGRAMS

Predictions were produced by four thermal programs. Two of these programs are in the public domain, ESP-r[13] and SERI-RES [14], CLIM2000[15] is used in-house by the EdF and APACHE[16] is sold commercially (Table 1). The programs employ a range of different algorithms for modelling the main heat flow mechanisms.

Program	Authors/Developers
APACHE	Oscar Faber, UK
CLIM2000	EdF, France
ESP-r	ESRU, Scotland UK
SERI-RES	ECOTOPE, SERI, USA

Table 1 : Programs analysed.

To try and eliminate mistakes by program users, the input files were checked by a third party and any necessary corrections were made. These blind predictions were then compared with the measured performance of the test rooms. Despite this checking process, errors were discovered (partly as a result of applying the analysis techniques) in the CLIM2000 and ESP-r simulations. A second set of simulations were therefore obtained for these programs. It is these final results which are reported in this paper.

5. PRELIMINARY COMPARISONS

Predictions and analyses have been undertaken in all six rooms for the heater power input, the air and globe temperature, the surface temperatures of the back wall, floor and ceiling, and the floor and back wall heat flux. In this paper, however, results are analysed for Room 1, with results for Room 0 being briefly considered.

Prior to applying the time-series analysis techniques, the hourly predictions were compared with the hourly measurements concentrating on 17 March, a very bright day, and 18 March, a very dull day. A comparison between the measured and predicted global radiation in the plane of the glazing confirmed that no errors had been made in the timing of the meteorological data.

Comparing the measured and predicted air temperatures for Room 1 (Figure 2), it is evident that all the programs show an exaggerated response to the operation of the heater. The response in the case of ESP-r is less exaggerated, presumably because, in making the predictions, the 22 minute time-constant of the heater was taken into account. The CLIM2000 and SERI-RES predictions are very close, probably because both programs use a combined radiative/ convective network to represent heat transfer within the rooms. The predicted air temperatures for all programs lie outside the overall experimental uncertainty of ± 1.5°C.

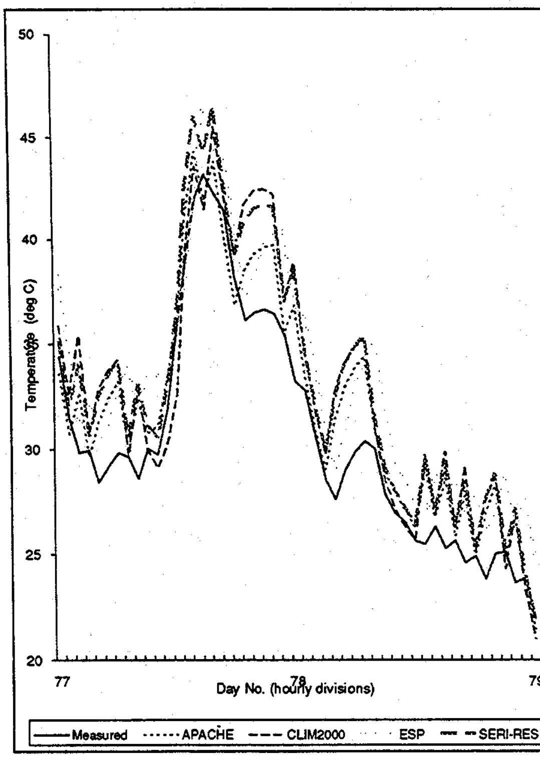

Figure 2: Air temperature in Room 1.

It is evident that CLIM2000 predicts much higher floor surface temperatures than the other programs or the measurements (Figure 3). This is probably because CLIM2000 allocates all the solar radiation transmitted into a room onto the floor.

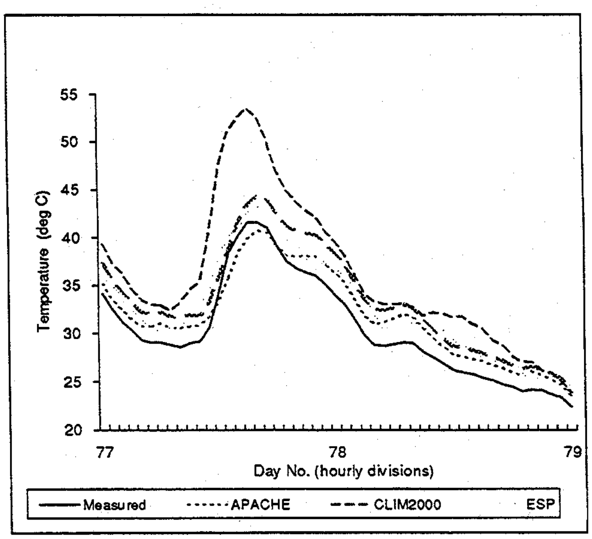

Figure 3: Floor surface temperature in Room 1.

There was a tendency for all the programs to over-estimate the back wall temperature by 1°C to 6°C.

6. THE CROSS-CORRELATION/ DECONVOLUTION TECHNIQUE

The cross-correlation/deconvolution technique has been developed by the EMC. It sets out to model the prediction error in terms of the external influences which drive both the simulation model and actual building [17]. To determine a suitable form for the error model a series of criteria was developed:

i. the error model should assume minimum a-priori knowledge about the simulation error;
ii. the error model should be amenable to physical interpretation, in order to allow the characteristics of the causes of errors to be deduced; and
iii. it should be possible to derive the error model directly from sequences of input and output data.

Investigations so far have been limited to linear, time-invariant error model and in response to the requirements listed above, a multi-variate impulse response representation has been adopted.

The response of a system to an input which has the value 1 at time zero, and which is zero at all other times is its impulse response. If the system is time-invariant then the response to such an input at any other time is obtained simply by shifting the impulse response. If the system is further assumed to be linear then the response to an input of a different magnitude can be obtained by scaling the impulse response, and, most importantly, the response to a series of such inputs can be obtained simply by adding together the appropriately shifted and scaled impulse responses. Any time-series can be expressed as a succession of shifted and scaled impulses thus, once the impulse response is known, the output of the system can be rapidly calculated.

The impulse response thus provides a universal, and unique, description of any linear time-invariant system. The process of generating the response of such a system to an arbitrary input sequence is known as convolution. The problem faced here is the reverse. Given input and output sequences (in this case measured driving forces and derived simulation errors respectively), the impulse responses are to be deduced. This process is known as deconvolution, and a method for doing it has been derived [12].

Once the impulse response of the simulation error to each of the inputs driving the model has been derived, the contribution from each input can be reconstructed. This in turn allows the relative magnitudes of the error due to each input to be found, and the most important sources of error to be determined. Inspection of the impulse responses themselves then provides insight into the type of mechanism behind the error: whether it is a storage or conduction problem, and whether it is a fast or a slow thermal process.

7.THE FREQUENCY DOMAIN ANALYSIS TECHNIQUE

The frequency domain technique has been developed by LETIEF[18, 19]. The performance of the programs is assessed using two kinds of information.

Firstly the simulation errors, i.e. the discrepancy between the predictions and the actual measurements is disaggregated into a number of parts. Each part is due to one of the external influences (inputs) driving the simulation (and the building). Hence, the discrepancy between the predictions and the measurements is the output of a system described by an input/output relationship in which the inputs are the actual excitations driving the building. Using this information it is possible to detect the excitations which are responsible for the major part of the simulation error. In order to account for any correlations between the inputs, these are sorted by their order of importance, then, the part of each input which is correlated to the previous one is removed; a new set of decorrelated inputs is thus obtained. The total variance at each frequency of interest can be decomposed to find the contribution from each input by computing the cross-spectrum between the error and each input.

The second form of information is supplied by directly comparing the behaviour of the real building and the predictions of a thermal program. For linear time-invariant systems, this comparison can be based on transfer functions expressed in the frequency domain. There is no need to assume the driving forces have independent effects, as the technique is able to perform the multi-variate frequency response identification even for correlated inputs. The information obtained allows an analysis at different dynamic operating ranges (static, low, medium or high frequency).

8. ERROR DIAGNOSIS USING THE CROSS-CORRELATION/ DECONVOLUTION TECHNIQUE

When all four models are used to represent Room 1 (double glazed, panel radiator heating) the heater is the dominant source of error when predicting the room air temperature. Figure 4 shows the breakdown of the error when each model predicts the air temperature.

Figure 5 shows the impulse obtained when the errors from the model APACHE are analysed, and shows that this error occurs immediately when the heater is operated. The same is also true for CLIM2000 and SERI-RES. In the case of ESP-r the error is smaller at time-delay zero, and an even smaller error appears at a delay of 1 hour. This implies that the treatment of the heater dynamics in ESP-r goes some way towards correcting the response of the air temperature.

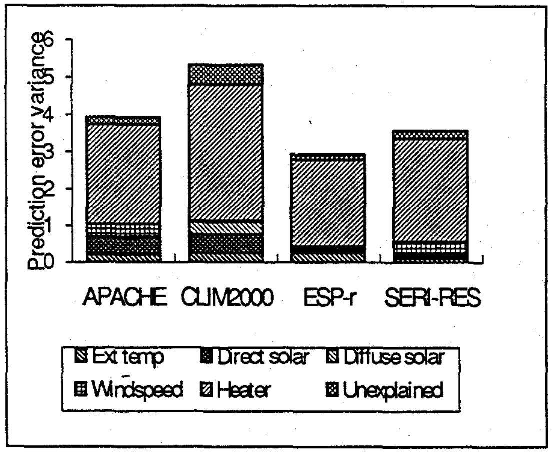

Figure 4 : Breakdown of air temperature prediction errors in Room 1.

The analyses for Room 1 also revealed that the coupling of the panel radiator to the massive floor of the building was well represented by ESP-r, with progressively poorer representations in SERI-RES, APACHE, and CLIM2000. Only small errors were observed in the coupling between the heater and the back wall.

To test whether the problem lies with the treatment of the radiative or the convective part of the panel radiator output, data from Room 0 (heated by a convector) was subjected to the same analysis. The dependence of the simulation error on heater operation has been found to be even more acute than for Room 1. Therefore it was concluded that the modelling of the convective portion of the heater output was the primary cause of the errors observed.

When floor surface temperatures and heat fluxes are predicted, a significant part of the resulting errors cannot be explained by the linear time-invariant model used here. This is most likely to be due to the time varying behaviour as the sun moves across the test room over the course of a day.

In CLIM2000 there is an additional problem with the response of the floor temperature and heat flux to the solar radiation. This confirms that the over-prediction observed is due to the way in which incoming solar radiation is apportioned by CLIM2000.

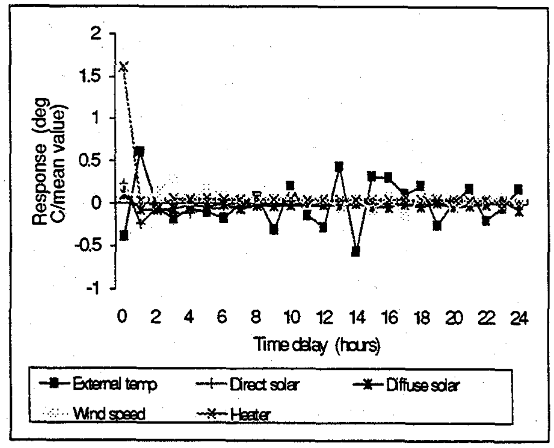

Figure 5 : APACHE air temperature predictions error impulses in Room 1.

The analyses have reached the foregoing conclusion by considering only the performance of the two double glazed rooms (Rooms 0 and 1). However, data from two rooms without glazing (Rooms 2 and 3) was also available, and this has been used to confirm some of the above conclusions. In other facilities, however, it may not be possible to collect this additional data. In this situation the type of analysis used here would be especially valuable.

9. ERROR DIAGNOSIS USING THE FREQUENCY DOMAIN ANALYSIS TECHNIQUE

Figure 6 shows the disagreggation of the air temperature prediction errors in Room 1 using the frequency domain analysis technique. This shows that the heater is the dominant source of the four programs errors in the fast dynamics band (0.03 h^{-1} - 0.055 h^{-1}). Also shown is that the CLIM2000 and APACHE air temperature prediction errors are very sensitive to solar radiation in the frequency range 0.03 h^{-1} - 0.055 h^{-1}, i.e., in the vicinity of 1 day^{-1} frequency.

Figure 7 contains the gains and phases of the transfer functions between the heating power and the air temperature for the four program predictions and the actual data. This shows that all four programs over-estimated the effect of the heater on all the temperatures measured in Room 1. Conversely, the effect of the outdoor air temperature on internal temperatures was under-estimated by all the programs. Nevertheless, the phases of the transfer functions were comparable. These results would suggest that the dynamics of Room 1 are adequately described by the programs, but that one or both of the following errors are present :

i. The models tend to over-estimate the insulating capacity of the building compared with that observed on the field.
ii. Horizontal and/or vertical stratification exists in the test-rooms which introduces a systematic shift between the measured temperatures and the predictions. The air inside the rooms had not been stirred and therefore such stratification is possible. Detailed measurements (e.g. a number of air temperatures and more surface temperatures etc.,) would help in checking this hypothesis.

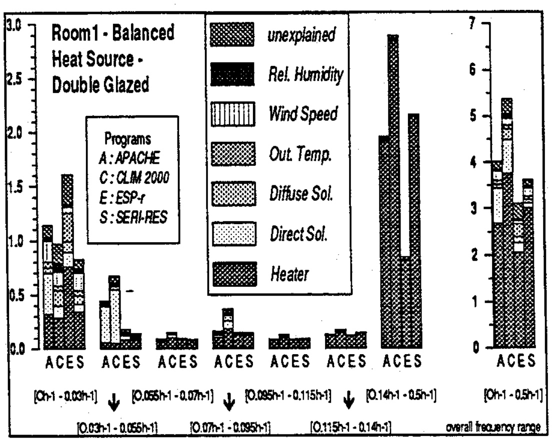

Figure 6: Breakdown of air temperature prediction errors in Room 1 for different time scales using the frequency domain analysis technique.

From an analysis of the phases of transfer functions, it seems that a physical phenomenon occurs in the test rooms which introduces an 1 hour time-delay in the actual effect of the heater on all the temperatures, but no delay in the actual effect on surface heat fluxes. For the ESP-r program, the discrepancy between the predicted heat fluxes and temperatures and the corresponding measurements is larger in the frequency range $0h^{-1}$ - $0.03\ h^{-1}$ (very slow dynamics with period of variations wider than 34 h). This is true since the ESP-r program is the only one which uses complex coefficients to describe the air-wall exchanges; it may be that the differences observed are rooted in this specific feature. The ESP-r modelling error (whether it concerns the air temperature, the surface temperatures or the heat fluxes) was the least responsive to the wind speed. This strengthened the belief that the indoor air-wall exchange coefficients were probably responsible for the observed discrepancies.

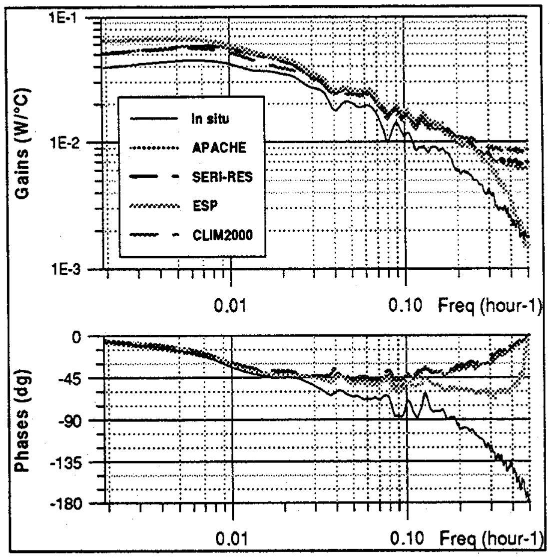

Figure 7 : Comparison of programs predictions and experimental data in Room 1 by using heater sequence - air temperature transfer functions.

CLIM2000 over-estimated the influence of solar radiation on the floor temperatures. Because the floor was the most massive part of the room, this introduced a large phase delay between the predicted and measured responses of floor temperature to solar radiation.

The influence of solar radiation on the floor was under-estimated by APACHE. This error led to discrepancies which were exactly opposite to those observed for CLIM2000, and of smaller amplitude. The reasons for the CLIM2000 and APACHE errors could be: (i) erroneous solar transmission coefficients, or (ii) incorrect distribution of the solar radiation over the indoor surfaces, or (iii) an unsatisfactory glazing model or any combination of these.

10. COMPARISON OF THE TWO ANALYSIS TECHNIQUES

Although both techniques assume driving forces to have linear effects which are invariant with time, they are significantly different in their approach. The EMC "cross correlation/ deconvolution" technique disaggregates errors in the time domain, which makes it easier to understand and interpret results. The LETIEF technique disaggregates errors in the frequency domain and is very powerful in obtaining explicit information about the dynamics of the system.

Figure 8 shows that despite their differences, very similar diagnoses emerged from the two methods, regarding the sources of errors in the predictions of the four programs. The fact that the two techniques lead to broadly the same conclusions suggests:

i. that the two methods have been implemented correctly; and
ii. that they can correctly identify sources of errors in programs' algorithms and input data, where measured data are available for the related processes.

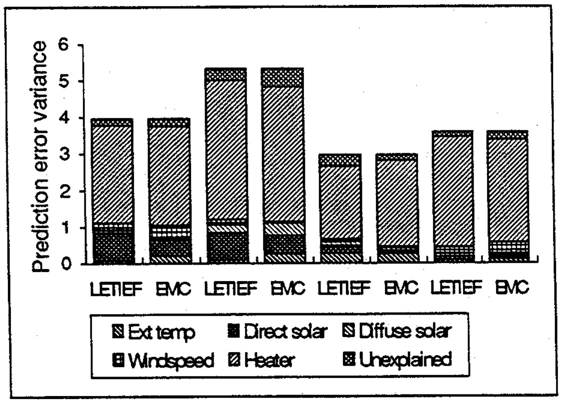

Figure 8 : Comparison of error disaggregation using the EMC and LETIEF techniques.

Differences between the measurements and predictions were identified as being due to:

i. the modelling of the heater and its dynamics and more particularly in modelling the convective part of the heat source;
ii. the distribution of solar radiation in CLIM2000 and perhaps APACHE;
iii. the modelling of stratification in the rooms; and
iv. the modelling of internal convective heat transfer coefficients in ESP-r.

Points (i) and (ii) were also noted by simple comparisons of the time-series data (section 5).

Although heater modelling was to blame for about 50% of the prediction errors, it is evident that neither technique could, on its own, accurately identify the heat transfer mechanisms which were to blame for the remaining errors. The disaggregation of the effect of these individual processes (radiation, convection, conduction) is only possible for certain if one or more of such processes are measured. For example, the two techniques would be capable of disaggregating the errors due to radiant heat transfer from the heater if the measured data were available for radiant heat flow from the heater (the localisation of the error to the convective part was achieved (section 8) by recourse to data from another room).

At present, mechanism level measurements are very difficult to make. Therefore, there is a need for further development of the techniques to identify errors in the models of various mechanisms of heat exchange, namely, radiation (short-wave and longwave), convection and conduction.

In the absence of a method for measuring and/or disaggregating mechanism level errors, the present techniques can be reliably used in conjunction with knowledge of programs' theoretical basis. By comparing the programs' assumptions and identifying their similarities and differences in modelling heat transfer mechanisms, it may be possible to infer the likely sources of any errors. As a further test of the effectiveness of the two error analysis techniques, it is planned to use them for the same set of data after improvements have been made to the program algorithms.

11. CONCLUSIONS

Measurements made in two double-glazed test rooms heated in a pseudo-random way have been compared with the predictions of four thermal simulation programs: APACHE; CLIM2000; ESP-r; and SERI-RES. Their predictions have been assessed using conventional methods and two advanced data analysis techniques. The following conclusions may be drawn.

i. The two error analysis techniques described in this paper proved to be capable of uncovering errors in both program algorithms and user inputs.
ii. The fact that the two techniques lead to the same conclusions is a strong indication of their reliability. The two techniques have, however, individual strengths in dealing with different problems. The EMC technique is useful for analysing time-based processes, whereas the LETIEF technique can be used for identifying errors at different frequencies.
iii. The main feature of the two techniques is that they disaggregate errors due to different sources of energy. This allows the analyst to focus on the algorithms associated with these sources in order to identify any possible errors.
iv. Most of the programs tested could not accurately model the heater and its dynamics. This inability, in some cases, accounted for about 50% of errors in program predictions. There is a need for an improved method for modelling heat sources of various types. Such a method should take into account the radiant and convective components of heat input and the storage effects of the heater if appropriate.
v. The prediction of solar radiation entering a zone and its distribution to the enclosing surfaces was found to be another likely source of error in some programs.

The prediction of solar radiation entering a zone and its distribution to the enclosing surfaces was found to be another likely source of error in some programs.

REFERENCES

1. Bland, B.H. "Conduction in dynamic thermal models: Analytical tests for validation", Building Services Engineering Research and Technology, 13(4), pp 197-208 (1992).
2. Martin, C., Seale C. and Eppel H., "Analytical testing of detailed thermal simulation programs", Proc. BEPAC 94 Conf., York, UK, (1994).
3. Lomas, K.J. "Aplicability Study 1, Executive Summary, De Montfort University Leicester", Report to ETSU under contract E/5A/CON/1213/1784, 73-pp (1992).
4. Judkoff, R. and Neymark, J., "Building Energy Simulation Test (BESTEST) and Diagnostic Method", International Energy Agency Task Solar Heating and Cooling Task 12/B Energy Conservation in Buildings and Community Systems Annex 21C, NREL/TP-472-6231, US National Energy Laboratory, Golden, CO (1994).
5. Lomas, K.J., Eppel, H., Martin, C. and Bloomfield D., "Empirical validation of thermal building simulation programs using test room data", IEA, Volume 1: Final report, September 1994, BRE, Watford, UK
6. Baker, P.H. et al., PASSYS II - "Final report of the UK consortium"., Energy Technology Support Unit, Rep. No. ETSU S/DS/00107/REP (1992).
7. Jiang, Y., Bloomfield, D. and Parand, F., "Model Information System and its application", Proceedings, CLIMA 2000, September 93, UK.
8. Lomas, K.J. and Bowman, N.T., "An investigation into analytical and empirical validation techniques for dynamic thermal models of buildings", A SERC/BRE Collaborative Research Project, GR/C/62871, Vol. IV, (1987).
9. Lomas, K.J., Martin, C., Eppel H., Watson M., and Bloomfield D., "Empirical validation of thermal building simulation programs using test room data", International Energy Agency, Volume 2: Empirical validation package, September 1994, BRE, Watford, UK.
10. Lomas K.J. (Editor), Empirical validation of thermal building simulation programs using test room data, International Energy Agency, Volume 3: Working reports, September 1994, BRE, Watford, UK.
11. Martin C., and Watson M., Empirical validation of the model SERI-RES using data from test rooms. Building and Environment. Vol 28, No 2, p175-187, 1993.
12. Martin C., and Watson M., Site Handbook : Empirical validation datasets 202, 203, 206 and 400 from the EMC test rooms, April 1993, EMC report.
13. ESRU, 1993, "ESP-r User Guide Version 8 Series, ESRU Manual U93/2, University of Strathclyde, Glasgow, UK.
14. Palmiter, L. and Wheeling T., "Solar Energy Research Institute residential energy simulator, version 1.0", Solar Energy Research Institute, Golden, CO, USA, pp-356 (1983)
15. Bonncau D., Covalet D., Gautier B., Rongére F., Manuel de prise en main de CLIM2000 X.EDF-HE12 W 3391 B.
16. APACHE version 7.2, "Apache User Guide, August 94", Facet Ltd., Marlborough House, Upper Marlborough Road, St. Albans, Hetfordshire AL1 3UT.
17. Martin, C., "A new technique for determining the cause of discrepancies between simulation model predictions and measured data". Proceedings 1993 CIBSE Conference. p378-388, 1993.
18. Ramdani, N. "Validation Expérimentale et Analyse de Signaux : Développement d'une méthodologie de comparaison modèle/mesures en thermique du bâtiment". Doctorat Thesis, University of Paris XII, 1994.
19. Ramdani, N., Candau, Y. and Dalicieux, P. "Linear Model Identification as a Tool for Empirical Model Validation of Dynamic Building Energy Simulation Models", SSB'94 Prceedings, 4th International Conference on System Simulation in Buildings,University of Liège, December 5-7, 1994.