#5367

TOWARDS AN INTERACTIVE MODEL VALIDATION FACILITY

Dr P Strachan & Professor J A Clarke
Energy Simulation Research Unit
University of Strathclyde
GLASGOW G4 0NG
Scotland

ABSTRACT

With the accelerating use of building performance prediction models in a design context, the need for comprehensive program accreditation procedures is becoming more pressing. This paper recognises the importance of the validation component of such a procedure and makes a case for containing much of the present knowledge about validation within an interactive facility centred on test cells. This facility would possess physical elements - the experimental and computing infrastructure to enable future monitoring studies - and virtual elements - the software structures and knowledge sources to support the process and represent best practice. The paper shows that many of these features have been developed in previous research projects, and in particular, have been brought together within the framework of a major European research initiative involving the testing of building components and model validation. Finally, the steps required to capitalize on the present capability are elaborated.

INTRODUCTION

At the present time, information technology (IT) - the collection, organisation and delivery of the information required to carry out tasks - is becoming central to almost all aspects of the work environment. In some sectors, such as building design where the rate of technology uptake is low, the potential is considerable; offering, for example, productivity gains through better information handling and performance gains through the ability to prototype and test at the design stage.

To effectively harness such technology, there is clearly a concomitant need for standardisation and accreditation of the IT products, especially in the case of complex performance prediction software[#] which attempts to embody the causal relationships inherent in the physical world. Just as the STEP initiative (Gielingh 1990) is addressing standardisation with respect to product description, there is a similar need for standard procedures for constructing and accrediting the software tools that will manipulate these product models. While the issue of program construction standards is the subject of several large scale projects in Europe (Charlesworth et al 1991, Sahlin 1988) and North America (Buhl et al 1990), the latter issue, program accreditation, can be recognised only implicitly in projects with other, more pressing goals (for example, the development of simulation based standards within the European standards organisation CEN, discussed in Van de Perre et al 1991).

In the case of building performance prediction, the essential ingredient of accreditation is validation, the subject of this paper. Often, building performance software is applied in situations where there exists little or no practical experience to judge whether predictions are realistic and to assess the inherent uncertainties. It is important that an estimate of reliability is available.

A distinction is drawn between models and simulation programs: a model is considered to be the overall simplified description of the building and its boundary conditions, together with the mathematical representations of the thermo-physical processes, while a simulation program is the computer code containing the tools for setting-up the models, solving the equations, and presenting the results. Model validation encompasses both the validation of the codes and the issues of model representation, i.e. the confirmation that a model's theoretical basis and software implementation is acceptably representative of the relationships observed in reality.

The hypothesis forming the basis of the paper is that the provision of a range of integrated validation tools, and the means to effectively apply them, will enable the next generation of design tool developers to turn their attention to the important issue of quality assurance in support of the effective use of the technology in practice. Further, it is argued that many of the elements of such a *validation facility* already exist within the CEC's PASSYS project (Wouters and Vandaele 1990) which will be completed by early 1992.

PROGRAM ACCREDITATION

As building performance software becomes more widely applied - in a regulatory context, in design, in education and in research - there is a growing need for procedures to ensure that the systems are fit for their intended purpose. This issue, accreditation, comprises several sub-issues as follows. Firstly, there is the need to devise, probably by the mechanism of community consultation, the criteria by which a building perfor-

[#] In this context *simulation based* design tools.

mance tool should be judged (scope, rigour and so on). Secondly, there is the issue of predictive accuracy and how this might vary as a function of the application context. Thirdly there is the need to consider the different possible abstractions of the physical reality, and the issue of the related program inputs, in terms of the impact on design tool integrity. And, finally, the form and content of the validation methodology must be elaborated.

At the present time these issues are beginning to be addressed. Within CEN, for example, simulation based building performance regulations are currently being drafted which will require the existence of an accreditation procedure. While it is anticipated that the importance of accreditation will increase rapidly in the coming years, already much work has been undertaken in the area of model validation which is, arguably, the most difficult part of the accreditation process.

VALIDATION

The development of a validation methodology has been a gradual process, the origins of which can perhaps be traced to the work of the Solar Energy Research Institute (Judkoff et al 1983), with subsequent and important extensions being made within a UK research project (Bloomfield et al 1988) and, currently, within the CEC's PASSYS project (Ostergaard Jensen and van de Perre 1991). Because of these contributions, a contemporary validation methodology would comprise the following elements:

Review of Theory

This involves a confirmation that the theories and empirical relationships that lie at the heart of a program are appropriate in terms of their application, scope and anticipated performance. Often the debate will centre around the selection of context-dependent algorithms: for example, there may be a conflict between a solar algorithm which has general applicability and one which is more accurate for a limited geographical range.

Code Checking

This involves the systematic checking of the computer code to ensure that the selected algorithms, whether appropriate or not, are correctly implemented. The use of structured programming, appropriate levels of documentation and code checking software can aid this process.

Analytical Testing

For some heat transfer processes, it is possible to establish exact solutions for limited, simplified cases. Such tests are useful for establishing that a particular software implementation is correct within acceptable bounds of accuracy, at least for the boundary conditions and assumptions for which the analytical case is valid.

Inter-Program Comparison

Such comparisons are usually undertaken using reference building descriptions. Although such tests are not able to establish the validity of any particular program, they are useful in two respects. Firstly, they provide a 'useful' evaluation in the case where one of the programs has already been the subject of rigorous validation. And, secondly, they often highlight serious shortcomings in one or more of the programs. An important consideration in an inter-program comparison is the requirement for input equivalence in order to ensure that the results do not relate, in effect, to different problems. To date, no definitive methodology has been established in this respect although several workers have identified the problem and its consequence (Bloomfield 1989).

Empirical Validation

This involves the comparison of measured and predicted data. In principle, and given that the measured data-set is of high quality, the technique is able to operate at any level of granularity - from the whole model to the flow-path level - and to provide an indication of the reasons for divergent results. For this reason, empirical validation will inevitably form a major constituent of any accreditation procedure.

Statistical Comparisons

Finally, there are the techniques which facilitate data comparison, allow the degree of (dis)agreement to be meaningfully quantified, and help identify sources of discrepancy between measured and predicted responses. Example of such techniques include parametric sensitivity analysis, which provides information on the influence of the uncertainties in the program's input parameters, and time and frequency domain comparators, which allow the (usually) time-series data to be viewed from different mathematical perspectives.

VALIDATION PROBLEMS

While contemporary validation methodology has evolved to a sophisticated level theoretically, it is nevertheless a difficult subject to pursue in practice. Individual techniques, such as analytical testing and inter-program comparison, are not in themselves capable of validating a model. For example, by inter-program comparison it may be shown that three programs out of four agree, even though it is known by other means that the fourth program is the only 'correct' one. Another, more pernicious example, is the case where good agreement is observed between measured and predicted data because the program being tested has several compensating errors for the particular experimental configuration (Lomas 1991).

In most of the program comparison exercises to date, the test vehicle has been designed to match the lowest common denominator case of the programs being tested. If, subsequently, good agreement is obtained, it is then only valid for a totally unrealistic problem. On the other hand, in order to configure a sophisticated program to handle the simple case other problems are often introduced which may then be the cause of poor performance.

Also, validation has been pursued as a fragmented activity rather than being applied in a manner that complements the program building and testing cycle. In short, there is a discontinuity in the technological basis of the program building and program validation activities. This, in turn, has served to ensure that there is little feed-forward of important validation knowledge to the new program structures.

In particular, it is almost impossible to assess the degree of validity of any program because there is no detailed validation record in a form that is easy to assimilate. In an accreditation context, such a record would surely be mandatory.

Turning now to empirical validation, several problems can be identified:

- The present lack of high quality data-sets is perhaps the most significant deficiency. Few of the existing data-sets, even though they have been expressly collected with a view to program validation, have proven useful for the task. In the main this is due to the absence of principal parameters and the lack of documentation.
- The delay between data acquisition and analysis often provides an insurmountable barrier with respect to ensuring that the data is fit for the purpose. Unless the analysis is carried out in conjunction with the experiment, it is difficult to ensure that the resulting data-set is complete, quality assured and free of errors.
- The difficulties which arise from the uncertainties associated with occupancy and infiltration will continue to restrict the usefulness of validation data-sets captured from full scale designs.
- Finally, the recurring expense of establishing and maintaining high quality instrumentation and monitoring equipment will ensure that there are few experimental sites that are capable of producing the high quality data needed for validation.

The need then is for an approach to validation which brings together the various components of the process - both the physical (experimental and computing infrastructure) and the virtual (software structures and knowledge sources) together with the resources to effectively apply them. The necessity for the experimental facility as an integral part of the overall system is of particular importance in view of the contemporary shortage of high quality data-sets that can be used for empirical validation.

A VALIDATION ENVIRONMENT

From the foregoing, it can be concluded that there is a need for a facility which contains both physical and virtual elements. This section details the individual constituents of such an environment, although it must be remembered that the most important aspect is their combination.

The Physical Elements

It is envisaged that the prerequisites of configurability and variable focus are best achieved by a facility which comprises the following:

Test Cells

Such experimental devices offer an economic and convenient intermediate step between the laboratory and real buildings. While laboratory methods are well adapted for the determination of heat loss characteristics (guarded hot plate and hot box) and solar factors (spectrophotometry), these are unable to duplicate real climatic effects and the many secondary interactions. Real building measurements, for their part, are expensive and fraught with the problems of achieving high quality data. Even assuming that such problems are tractable, the difficulties of formulating the results for general application remain. Test cells on the other hand ensure applicability and generality by allowing testing in real climate conditions and by providing a controlled internal environment in which a range of experiments can be conducted.

The perceived advantages of a test cell approach are:

- They are room-sized and provide a realistic intermediate step between the laboratory and real building situations. Importantly, the effects of occupancy and infiltration can be removed or controlled.
- It is possible to match cells and thereby facilitate side-by-side experiments.
- A high level of instrumentation and control is possible.
- They provide opportunities for collaborative research and, because of the possibility of site replication, they can be made to represent a wide climatic diversity.

Test Components

The large variation in impact of the various heat flow paths according to building type means that a program cannot be validated for one design and then be expected to perform well for all other designs. There is, therefore, a requirement that the experimental facility is configurable, so that a wide range of constructions and their modes of operation can be tested.

Sensors and Data Acquisition System

In order to obtain a good definition of the heat transfer processes, a comprehensive instrumentation set and sophisticated data acquisition system are required. The former is necessary as a means of providing all program inputs and to provide sufficient information on the building performance to reduce the possibility of not recognising compensating errors in programs. The latter should be capable of acquiring data at high sample rates for a large number of channels, with suitable data storage capabilities, and permit the on-line display of sensor data.

Heating and Cooling Control System

In order to fully characterise the design features under test, it is necessary to employ a wide range of operating strategies, from steady state to short period dynamic fluctuations. The control system should be flexible enough to allow programmable switching, preferably of both heating and cooling.

Data Control System

Data handling capabilities, involving a flexible pre-processing package and a rapid mechanism for data transfer to a workstation environment, is an important topic in view of the large data quantities generated. The data control should include data checking, procedures for establishing the reliability of the data, and documentation.

Workstation

Recent improvements in the cost/ performance ratio and data storage capabilities of workstations, makes them the ideal medium for the storage and manipulation of the data-sets and models involved in the validation activity.

<u>The Virtual Elements</u>

Many items can be placed in this category. It is important to note that the real power of such tools is derived from their ability to operate one with the other, in addition to their accessibility.

High Quality Data Sets

A database of previously collected high quality data-sets, including full documentation on their contents and previous analysis history.

Simulation Programs

State-of-the-art building performance programs, increasingly validated against previous data-sets. These should include documented source code so that the full range of the validation methodology can be employed.

Sensitivity Analysis

Such analysis tools have proved to be particularly useful for experimental design. They can also be used to assess the effects of the uncertainties in the input parameters when comparing measured and predicted data.

Statistical Tools

These might, for example, range from simple statistical measures of time-series data to time and frequency domain comparators of the measured and simulated variables (Palomo and Madsen 1991).

Alternative Algorithms

These can be embedded into selected simulation programs for inter-comparison studies and investigations of performance against analytical solutions and empirical data-sets. Incorporating different algorithms within otherwise identical code is a powerful means of isolating the effects of the particular heat transfer process. In the longer term, the Energy Kernel System (Charlesworth et al 1991) will provide a suitable basis for such a real-time comparison of competing algorithms.

Analytical Solutions

A series of coded analytical solutions held on-line in a database.

Reference Data

Typically, this might include pre-constructed models of reference buildings and test cells and on-line knowledge sources covering climatic data, material properties and the like.

Design of Experiments

Where new empirical data is required for validating a particular aspect of model performance, a detailed study can be undertaken using sensitivity analysis and energy balance techniques as contained within software structures already accredited for use with the facility. This would ensure that the experiment to be conducted is optimal in terms of the specification of instrumentation requirements and control strategy. In this way it would be possible to construct experiments based on existing test cell components or, alternatively, new components tailored for a particular validation experiment.

In practice it would be possible to undertake 'blind' validation in which the program is run prior to any comparison with the experiment data. This is an important device when trying to establish confidence in the predictive capabilities of a program. Alternatively, in research mode it would be advantageous if it were possible to fine tune algorithms by observing their performance against measurements as their parameters are subjected to controlled adjustments.

Miscellaneous Tools

These might, for example, include spread-sheet programs for data tabulation and manipulation; graphics tools for visualisation; and code checking and debugging software.

ADVANTAGES

It is suggested that such an *interactive validation environment* would lead to a marked improvement in validation capabilities. For the non-empirical elements of validation, the perceived advantages are:

- Improved efficiency through ready access to the tools of validation and a range of on-line knowledge sources in support of the process.
- Better continuity and harmony because future programs could be tailored to suit the standards of the environment rather than the validation elements needing to be configured to suit the program. After a program was made ready, the various validation tools could be applied in a structured manner.
- A validation log would ensure that the validation activity, as conducted, was well documented. This, in turn, would ensure that the quality assurance record of future design tools was captured in a uniform and comparable manner. Assessments of program performance would be enabled and it would be possible to easily re-implement checks after each program modification.
- The facility would serve as a repository for research knowledge concerning the validation process itself, and help identify priority areas for further research.
- Regular program checking would be enabled via a range of standard test procedures which could be invoked routinely, thereby detecting any errors introduced during the development process itself. This has been a particular problem in the past where validation results have been reported on earlier versions of a program and it is not known whether developments in the intervening period have removed or exacerbated the problem.

With regard to empirical validation there are many benefits to be derived from a closer link between the modeller and the experimentalist, where real-time activities (or at least data use with short delay), would be effectively enabled. These potential benefits are:

- Any divergence between model predictions and experiment can be acted upon in a rational manner. For example equipment might firstly be checked, a sensitivity analysis might be conducted, statistics might be extracted from the predicted/ measured time-series prior to divergence, the experiment might be adapted to examine some related issue and so on. In any event the process of validation is placed on a logical footing and the research activity is enhanced.
- Efficiency gains deriving from resource sharing and from the experimental and modelling knowledge to hand.
- By minimising the delay between data collection and analysis, the model proving time is reduced. Also, since test cell experiments can have a duration of 30 days or more, the possibility of the timely detection of sensor error is improved.
- Influencing the experimental plan is enabled, whereby the experiment and/or model can be adapted or extended in response to any uncovered deficiencies. For example, when testing a solar algorithm, it may be necessary to prolong an experiment if adverse weather conditions prevail.
- Because the data user is involved at the time of data collection, the acquired data will be of a higher quality. The user would also be responsible for documentation and archiving of the dataset to ensure its future applicability.

POTENTIAL USERS

There are several potential users of such a facility:

Accreditation Agency

Accreditation agencies will be interested in ensuring that the methodology and standard tests are uniform at least within a given national boundary. The envisaged facility would provide a mechanism to achieve this goal. In addition it would provide a network of centres of excellence which could pursue extensions to the methodology and improvements to the tests.

Component Manufacturer

This user type will wish to obtain the performance characteristics of a new component and, by the use of an accredited program, scale the results to other design configurations and climatic contexts. This gives rise to the concept of 'practical validation' whereby confidence in a program is firstly gained by comparing predictions with the performance of a component mounted on the test cell. Good agreement, with a comprehensive test sequence that stresses the component through its full thermal range, leads to confidence in the program which is then used to extrapolate the test cell results to real buildings. Clearly, in the case of poor agreement between measured and predicted data, the need for further model development work is highlighted and the mis-application of the program is prevented.

In practice, existing programs could be used to help design a suitable experiment, in terms of the anticipated behaviour of the components to be tested, the levels of instrumentation and accuracy and the control strategy to be employed.

Vendor

This user type will typically wish to compare some target program predictions with those from other credible programs in order to demonstrate that the system meets with the minimum requirements deemed acceptable. A similar objective might also be pursued by the use of historical data-sets which embody a range of environmental control regimes, component types and climatic influences.

Researcher

This user type would wish to test the validity of a particular algorithm or a program as a whole. This may entail an inter-program comparison in which the algorithm is to be compared with alternative formulations as present within another respected system. Alternatively, an empirically based study may involve an experiment using two matched cells in a side-by-side experiment to isolate the effects of a single parameter. A third approach may involve the introduction of a specially designed component to the test cell which accentuates the cell's sensitivity in terms of the process being studied. An example of this would be the introduction of a component with a deliberate thermal bridge in order to test the three dimensional conduction capabilities of the target program. Using the support systems available - empirical data-sets, source code control software, staff expertise and so on - new algorithms could be easily incorporated into existing programs to allow assessment of the impact.

Model User

This user type could use the facility to check the validation history of a program for the particular application envisaged. In some cases reference design descriptions might be used during a training phase in order to confirm, for example, that a program's operational syntax is correctly understood (by comparing achieved predictions with standard result-sets for example). In other cases this user type might wish to obtain a second opinion or estimate the effects of the uncertainties inherent in the design parameters.

EXISTING FACILITIES

Many of the elements of the proposed facility - both physical and virtual - are to be found in the many validation projects completed in recent years or ongoing. However, it is only with the advent of the CEC's PASSYS project (Wouters and Vandaele 1990) that the potential exists to bring these elements together in the manner proposed in this paper.

PASSYS, a European concerted action in the field of Passive Solar Architecture, has concentrated on the use of test cells with high levels of instrumentation and control. Because these test cells are standardised and distributed over 10 EEC member countries, they provide a unique facility for future European model validation activities.

The PASSYS facilities already possess advanced data acquisition capabilities, Unix[#] workstations, high quality empirical data-sets, on-line simulation and data analysis packages and, because of the international dimension, a harmonised component testing and model validation methodology. In addition, teams from each country have been involved in the application of the validation methodology, giving rise to a large pool of expertise.

From a modelling perspective, the PASSYS cells offer the necessary benefits of realistic dimensions, configurability and tight experimental control. They have two main limitations:

- The large thickness of the walls gives rise to significant two- and three-dimensional effects, not conventionally treated in dynamic simulation programs. This does provide opportunities for research, however, since edge losses are not insignificant even in full size buildings (Hassid 1991).
- The cells themselves are highly insulated and therefore have low internal surface energy flows, making them unsuitable for the study of some heat transfer processes. However, the configurable south wall components can have high loss coefficients; in this case associated fluxes would be dominant.

The possibility therefore exists to evolve the PASSYS sites into distributed test centres which can offer component testing to industry and, by virtue of a growing number of 'approved' simulation programs, the rapid assessment of the replication potential of new technology. The validation work of the centres would therefore have at least three immediate clients: the researchers and model developers who are concerned with accuracy and applicability, and the industrialists who will require product performance information in relation to the anticipated national markets and beyond.

It is important to note, however, that the proposed facility is not in any way restricted to a single cell type. It is equally possible to set up such a facility at alternative test cell locations or other experimental buildings.

CONCLUSIONS

There are indications that the emergence of simulation based building energy standards is accelerating. This, in turn, will provide the incentive to establish software accreditation methodologies. It has been argued that the implementation of accreditation can best be achieved by establishing a network of sites, containing test cells, where the necessary techniques and tools are located and where a detailed record can be kept of the accreditation history of the advanced modelling systems.

There is a clear need for further empirical validation data-sets. Given this need, and the difficulties

[#] Unix is a trade mark of AT&T.

inherent in constructing suitable experimental facilities, it remains for the model validators to make most use of existing facilities - at present the best data is being derived from test cells. It is believed that the PASSYS facilities provide a good basis for the development of closer links between model developers and experimentalists, with their powerful combination of highly specified, matched test cells distributed throughout Europe and state-of-the-art simulation knowledge. However, the validation facility could equally well be based at other locations where there exists a conjunction of a high quality outdoor experimental facility and appropriate analysis capabilities.

The proposed developments would provide a useful facility for model validation, both fundamental validation of the performance of the simulation program against measured data, and practical validation, where the aim is to develop confidence in the modelling of particular building components. This paper has stated the advantages of the interactive facility, and has described the potential users.

Looking to the longer term, it is possible to foresee a time when simulation programs will be linked to intelligent buildings to predict the building response in advance and to enable appropriate control procedures. It is clear that many of the techniques discussed in this paper will then be routinely applied, not least the real-time link between simulation and monitoring.

REFERENCES

Bloomfield D, Bland B, Gough M, Keeble E, Pinney A, Steffanizzi P, Wilson A, Bowman N, Lomas K, Allen J, Whittle J, Irving A, Allen E and Bean M 1988, "Final Report of the BRE/SERC Collaborative Research Project: An Investigation into Analytical and Empirical Validation Techniques for Dynamic Thermal Models of Buildings", Building Research Establishment, Garston, UK.

Bloomfield D 1989, "Evaluation Procedures for Building Thermal Simulation Programs" In *Proc. Building Simulation '89*, Vancouver, Canada, pp217-222.

Buhl W F, Erdem A E, Nataf J-M, Winkelmann F C, Moshier M A, Michael A and Sowell E F 1990, "The U.S. EKS: Advances in the SPANK-based Energy Kernel System" In *Proc. Third International Conference on System Simulation in Buildings* University of Liege, Thermodynamics Laboratory, Liege.

Charlesworth P, Clarke J A, Hammond G, Irving A, James K, Lockley S, Mac Randal D, Tang D, Wiltshire T J and Wright A J 1991, "The Energy Kernel System" In *Proc. of BS'91*, Nice, France, Aug. 20-22.

Gielingh W 1990, "Computer Integrated Construction: A major STEP forward" *Computer Integrated Construction* 1(3), Wix McLelland Ltd.

Hassid S 1991, "Algorithms for Multi-Dimensional Heat Transfer in Buildings" In *Proc. of BS'91*, Nice, France, Aug. 20-22.

Judkoff R, Wortman D, O' Doherty B and Burch J 1983, "A Methodology for Validating Building Energy Analysis Simulations" *Report TR-254-1508* Solar Energy Research Institute, Colorado, USA.

Lomas K J 1991, "Dynamic Thermal Simulation Models of Buildings: New Method for Empirical Validation" *Building Serv. Eng. Res. Technol.* 12(1), pp25-37.

Ostergaard Jensen S and van de Perre R 1991, "Tools for Whole Model Validation of Building Simulation Programs: Experience from the CEC Concerted Action PASSYS" In *Proc. of BS'91*, Nice, France, Aug. 20-22.

Palomo E and Madsen H 1991, "Methods to Compare Empirical Measurements and Simulations" In *Proc. of BS'91*, Nice, France, Aug. 20-22.

Sahlin P 1988, "MODSIM: A Program for Dynamic Modelling and Simulation of Continuous Systems" In *Proc. of the 30th Annual Meeting of the Scandinavian Simulation Society*, Swedish Institute of Applied Mathematics, Stockholm.

Van de Perre R, Ostergaard Jensen S, Bloomfield D and Agnoletto L 1991, "Simulation Based Building Energy Standards: A Case Study, Overheating Risk Assessment. International Workshop on Computers and Building Standards, VTT, Espoo, Finland, 27-29 May.

Wouters P and Vandaele L (eds) 1990, "The PASSYS Test Cells" *PASSYS document number 114-89-PASSYS-INS-033* Commission of the European Communities, DGXII.