Final report

Project title	UDP17-II Dansk deltagelse i IEA EBC Annex 71
File no.	64017-05139
Name of the funding scheme	EUDP
Project managing company / institution	DTU Compute
CVR number (central business register)	30 06 09 46
Project partners	SSB
Submission date	7 January 2022

1. Summary

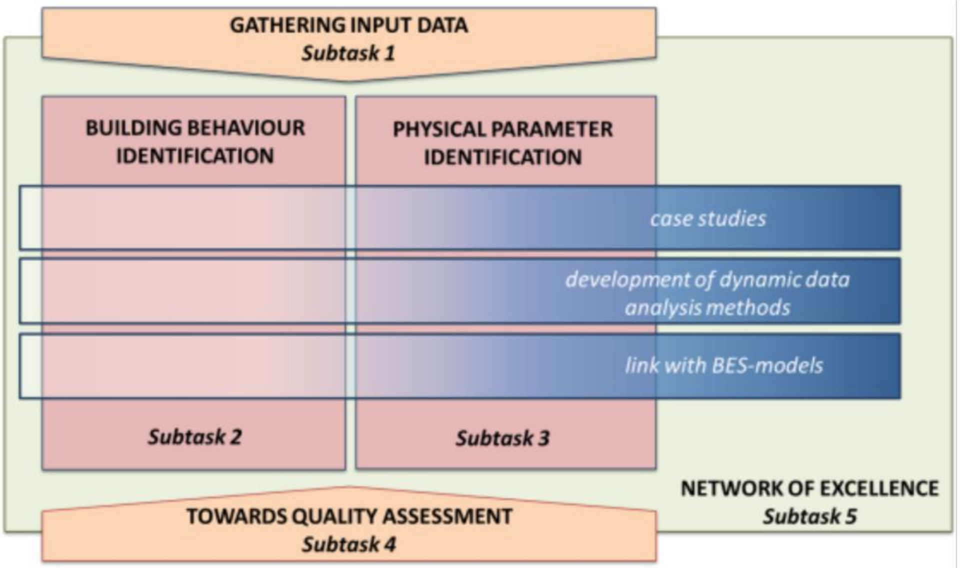
The IEA EBC Annex 71-project was limited to residential buildings, for which the development of characterisation methods as well as of quality assurance methods have been explored. Characterisation methods aim to translate the (dynamic) behaviour of a building into a simplified model that can inform predictive control, fault detection, optimisation of district energy systems, etc. Within Annex 71 we referred to this as building behaviour identification. Quality assurance methods aim to pinpoint some of the most relevant actual building performance metrics. This part is referred to as physical parameter identification. A reliable characterisation and quality assurance is strongly dependent on the availability and quality of the input data. At the same time, the expected quality and reliability of the outcome will be determined by the required accuracy to perform a quality assurance. As a result, the analysis of potential methods was steered by both the possibilities and limitations of the available input data as well as by the requested outcome to perform real quality checks. The research project was organised in five subtasks: Subtask 1 investigated the possibilities and limitations of common data bases and monitoring systems. Subtask 2 focused on the development of dynamic data analysis methods suitable for describing the energy dynamics of buildings. Subtask 3 focused on development of dynamic data analysis methods suitable for physical parameter identification of buildings. Subtask 4 investigated to what extent the methodologies developed in ST2 and ST3 can be used in a quality assessment framework. Subtask 5 had collaboration with DYNASTEE (www.dynastee.info) on dissemination and education. Finally, a building simulation validation was carried out for investigating the reliability of common building energy simulation software. The results are documented in four reports matching the subtasks topics, they are available at the Annex 71 website.

IEA EBC Annex 71-projektet var begrænset til boligbyggeri, for hvilket udviklingen af data-drevne karakteriserings- og kvalitetssikringsmetoder, blev undersøgt. Karakteriseringsmetoder har til formål at oversætte en bygnings dynamiske respons målt i data til en forenklet model, der kan anvendes i prædiktiv kontrol, fejldetektion, optimering af fjernvarmesystemer osv. Kvalitetssikringsmetoder har til formål at estimere fra data nogle af de mest relevante faktiske faktorer, der påvirker en bygnings ydeevne, f.eks. hvor isoleret bygningen er ved at estimere varmetabskoefficienten. En pålidelig karakterisering og kvalitetssikring er meget afhængig af tilgængeligheden og kvaliteten af data, hvilket førte til at analysen af potentielle metoder blev en afvejning af både

mulighederne og begrænsningerne ved de tilgængelige data, samt af det ønskede resultat i forbindelse med reelle anvendelser. Projektet var organiseret i fem dele: I del 1 blev muligheder og begrænsninger ved fælles databaser og overvågningssystemer undersøgt. I del 2 var fokus på udvikling af dynamiske dataanalysemetoder, der er egnede til at beskrive bygningers dynamiske respons. I del 3 var fokus på udvikling af dynamiske dataanalysemetoder egnet til fysisk parameteridentifikation af bygninger. I del 4 blev det undersøgt, i hvilket omfang de metoder, der er udviklet i del 2 og del 3, kan anvendes i forskellige reguleringer og lovgivninger. I del 5 blev samarbejdet med DYNASTEE (www.dynastee.info) om kommunikation og uddannelse. Desuden blev en bygningssimuleringsvalidering udført for at undersøge pålideligheden af almindeligt udbredt simuleringssoftware. Resultaterne er dokumenteret i fire rapporter, der matcher emnerne i de forskellige dele, de er tilgængelige på Annex 71 hjemmesiden.

2. Project objectives

The IEA EBC Annex 71 project focused on the development of replicable methodologies to characterize and assess the actual energy performance of buildings starting from on board monitored data of in-use buildings. The project was split into 4 subtasks and an exercise for investigation of buildings simulation software. The figure below shows the subtasks together with crossing tasks – DTU led the “Development of dynamic data and analysis methods” task:



In the following the objective of each of the subtasks are described:

Subtask 1 investigated the possibilities and limitations of common data bases and monitoring systems. This subtask is strongly related to subtasks 2 and 3 by linking the available input data – as much as possible based on existing (non-intrusive) monitoring systems and data bases – to the accuracy of the predicted outcome. A state-of-the-art survey of existing methods, their costs, timeframe and typical accuracy was made. In a second part the step from monitoring to current on board measuring methods was reviewed. Finally, the application of an on-site measured heat transfer coefficient within the global energy efficiency framework was proposed.

Subtask 2 focused on the development of dynamic data analysis methods suitable for describing the energy dynamics of buildings. Based on in-situ monitored data, prediction models were applied and optimised that can be used in model predictive control, fault detection, and design, control and optimisation of district energy systems. Necessary data acquisition, development of methodologies and accuracy and reliability of the building behaviour identification models was investigated.

Subtask 3 focused on development of dynamic data analysis methods suitable for physical parameter identification of buildings. Contrary to Subtask 2, in which the identified parameters do not necessarily have a physical meaning (or do not correspond to the actual value), parameter identification aims to characterize the actual physical parameter. Subtask 3 hence investigated which methodologies are most suitable to determine the actual energy performance indicators of buildings, such as the overall heat loss coefficient, solar aperture. As in subtask 2, the focus was on methodologies that can be used on occupied buildings, making use of (limited) monitored data.

Subtask 4 investigated to what extent the methodologies developed in ST2 and ST3 can be used in a quality assessment framework. A large survey was performed amongst possible stakeholders on interest and expectations of quality assessment methods based on in-situ measured data. The main focus was on the determination of the actual heat loss coefficient of a building in an easy, cheap and reliable way, so that it can replace the calculated design value in energy performance certifications. That way, subtask 4 made the link between the annex-participants and certification bodies, government, practitioners in the field. At the same time, subtask 4 gave the necessary boundary conditions (reliability, accuracy, cost,…) the methodologies have to fulfil to be applicable in real life quality checks.

Subtask 5 continued the collaboration with DYNASTEE (www.dynastee.info), started within Annex 58. This collaboration showed to be extremely fruitful in dissemination of the results, collecting and distributing research outcomes, and organizing conferences, workshops and training courses.

The Building Energy Simulation (BES) validation exercise investigated the reliability of common building energy simulation programs. There has been significant work undertaken in past IEA EBC Annexes on software validation, particularly inter-program comparisons (e.g. BESTEST) and empirical validation on test cells. In Annex 58, empirical validation was extended to full-scale buildings, namely the Twin Houses at Fraunhofer IBP's test site in Holzkirchen, Germany. In this research, the focus was on fabric performance with simple internal heat gain schedules. The empirical validation undertaken in IEA Annex 71 extended the scope of the experiments in the Twin Houses by including underfloor heating systems and realistic occupancy schedules.

3. Project implementation

The Annex 71 project evolved mostly as planned. Naturally, the COVID19 outbreak and following restrictions led to some unforeseen troubles, however they were overcome by running meetings and activities online. The following expert meetings were carried out in the project (note that the closing event in Salford was online):

Meeting	Place, date	Attended by
Kick off meeting	Leuven, Belgium, October 2016	49 participants
Second preparation meeting	Loughborough, UK, April 2017	61 participants
First working meeting	Chambéry, France, October 2017	62 participants
Second working meeting	Brussels, Belgium, April 2018	56 participants
Third working meeting	Innsbruck, Austria, October 2018	55 participants
Fourth working meeting	Bilbao, Spain, April 2019	59 participants
Fifth working meeting	Rosenheim, Germany, October 2019	56 participants
Sixth working meeting	On-line meeting, April 2020	50 participants
Seventh working meeting	On-line meeting, October 2020	50 participants
Eighth working meeting	On-line meeting, April 2021	56 participants
Closing event	Salford, UK, September 2021	

The project got off to a quite “smooth” start, since most of the participants had already worked together along the same lines in the Annex 58 project. The main working method was to arrange “common exercises” in which a problem was set for the participants to solve and then the results were shared. Each subtask worked independently with online meetings, but results and findings were shared at the meetings. The working method was quite productive, although it took some time before the case-studies evolved into a mature state.

3.1 Common exercises

The following is an overview of the common exercises carried out in Subtask 2 and 3, which were the main tasks in developing the modelling methodologies.

Subtask 2

Common Exercise 0: Datasets recorded in the Annex 58 Round Robin Test Box was provided to participants. The objective was to predict the temperature in the box using different modelling techniques. In particular: Prediction of overheating risk, temperature profile and overheating time (temperature exceedance). Quite good participation was achieved.

Common Exercise 1: the main objective is the exploration of methodologies to obtain an accurate and reliable prediction of the future, hourly energy use for space heating and domestic hot water production. This exercise did not work out! There were inherent causality issues in the proposed modelling problem. Since the gas consumption for space heating in the particular case was a highly non-linear and complicated process to model. This was pointed out early by the Danish partners. Finally, no contributions were submitted.

Common Exercise 1 bis: The main objective is the exploration of methodologies to obtain an accurate and reliable prediction of the future, hourly energy use for space heating and/or the mean indoor temperature 1 of the building. So, the objective was changed to prediction of indoor temperature. Only Two contributions were received.

Common Exercise 2: Participants were asked to identify models that are able to predict the normal operation of the building. Secondly, the models should be used to identify faulty behaviour and diagnose the cause of the faults.

Common Exercise 2 bis: Same exercise, just extended with a few different aspects.

Common Exercise 3: The exercise continues on the story-line of CE2 (and CE2_bis) in which modelling approaches were tested for their ability to detect 2 faults introduced to the Twin House BES-validation experiment. In this 3rd common exercise, we shift from simulated data to the actual measured data retrieved from the Twin House experiments.

Subtask 3

Common Exercise 1: This first Common Exercise of Subtask 3 wants to bring the Annex 71 experts together for a first exploration of possible techniques to determine the overall heat loss coefficient, H, on the basis of in-use building performance data of Twin Houses - House 1. Participants were encouraged to apply as many of their favourite analysis techniques as possible.

Common Exercise 1 bis: Participants were asked to investigate the robustness of the estimates for H with regard to available data. They could select different subsets of data: different periods and duration, but also different sensor sets. Data subsets were fed into algorithms and evaluate the estimates that result (mean and standard deviation). Then they should explain why some techniques outperform others, and why they correspond to or differ from the design/reference value? Moreover, they were invited to report any other performance indicators (solar aperture, time constants, thermal capacities, ...) that result from applying these techniques.

Common Exercise 2: For the Twin Houses, two datasets were introduced, dataset N2 referring to the Twin House with electric heaters and dataset O5 referring to the house with underfloor heating (UFH) and a heat pump as heating source. Specifically, for the ST3 CEs two basic datasets were prepared, on the basis of the original dataset. These datasets served as reference, but to investigate specific issues reference back to the original dataset.

Building simulation exercise

The datasets from the Main and Extended Experiments can be used for many different purposes such as education and training, the development of simplified reduced order models and other scientific research requiring measured data from well-specified real buildings. For these purposes usually the full dataset, containing all the data collected during the experiment, should be used. However, when the focus is a BES program validation (and/or model development) it’s recommended to follow the 2-phase blind/open validation approach (used in Annex 71), as described below, to separate input errors from program errors.

In this approach, each model validation team predicts the temperatures and heating inputs using the program(s) under investigation. The validation methodology is a two-phase blind validation, as used in Annex 58 (Strachan, et al., 2016) and similar to other previous IEA empirical validation studies. Ideally, this procedure has different persons (or even organisations) working collaboratively to improve model quality assurance and analysis techniques.

The required steps are as follows:

1) Blind validation (“Blind phase”).

1a) Modellers predict heating energy and indoor climate using the experimental specification, measured climate data and operational schedules but without knowledge of the measured heating energy consumption (in the case of known indoor climate) or indoor climate (in the case of known heating energy consumption).

1b) Modellers submit their simulation results in a modelling report with details of the programs used and assumptions made.

2) Blind stage analysis. This compares predictions against experimental data for indoor climate and heat fluxes. Inevitably at this stage, differences are due to a mix of user / modelling errors and program deviations (and potentially measurement uncertainties).

3) Re-modelling (“Open phase”). The measured data is disseminated. Modellers are encouraged to investigate differences between measurements and predictions and resubmit predictions and up-dated the report. Only changes which correct user input errors or alter a modelling assumption (with documented rationale) are allowed. It is important to ensure that model input parameters are not simply calibrated to improve agreement with measurement. In principle, this step identifies program errors by eliminating input errors.

4) The improved predictions are compared against the measurements to identify remaining flaws and identify areas where program improvements are required. When complete, validation data sets and models are archived.

3.2 Discussion of the working method

Around the common exercises (described in the following) much work and interaction were carried out, both the between the Danish partners, but very much on the international level (although mostly European).

We can mention the following important points:

- Educational activities:
    - A major contribution during the Annex was two summer schools. DTU arranged and carried out a week (2.5 ECTS points) summer school in August four consecutive years (2018-2021). We had 45 increasing to 70 (half online) participants. We will continue this successful effort in other projects.
    - Summer school in Spain. A joint effort was the summer school in Spain arranged every year, although the last two years ended up being an online webinar. DTU delivered a major contribution on statistical modelling.
- Research activities:
    - A lot of networking goes on in an Annex, so several stays of PhD students were arranged through contacts in the Annex, as well as opponent to PhD defences among the partners.
    - Deep research activities. We have analysed many cases and written papers together. The list of papers resulting directly (with other projects) from the Annex work is included in the next section.

3.2.1 Common exercises

Generally, the common exercises proved to be a good way of collaborating in the Annex. They allowed everyone to work on the same problems, which serves well that all are on the same pace and that results can be compared directly across modelling techniques and participants. Ideally, the sources variability can separated across factors such as: modelling technique, data input decisions, pre-processing, etc. However, this require very well-designed exercises, which was not always the case. One particular problem with the exercises carried out during Annex 71, was the representativity and replicability. Clearly, when carrying out data-driven modelling, which was the case in all exercises, it is very important that the case selected, i.e. the particular building monitored, the conditions etc., are representative for important fractions of the real cases. It is a very difficult point to live up to when dealing with buildings and energy, and in most of the common exercises the case was not highly representative. It was what happened in the process, many times calls for useful datasets were send, and very few responses were received. Main reason is, that the data needed to be detailed, with detailed information about the building and the conditions under which the experiment were carried out. So, this should be kept in mind when running such projects, that common exercises must have the appropriate cases and data available. Generally, the common exercises were, however, a success, and in the end of the project the cases were matured and led to the basis of the joint work presented in the four final reports. The Danish partners participated in most of the exercises and developed new models and techniques, which led to journal publications.

3.2.2 Summer school at DTU

A major contribution from the Danish project has been the PhD summer-school arranged since 2018 every year. The following is a description of the content of the weekly long program.

Description:

Focus is on data from buildings and aggregated buildings in districts etc., and on load, solar and wind power forecasting. Finally, the program ends with a day devoted to MPC and optimization.

Some particular examples of applications and techniques:

When including the effect of wind and solar in building energy models, without detailed information about the building and its usage, it is important to describe non-linear effects and interactions. For this kernel and basis function (e.g. spline) techniques are indispensable in both discrete and continuous time models.

Occupancy in buildings leads to different patterns of stochasticity in observed data. Even though it is “random” there are still systematic patterns (e.g. varying levels of noise), which must be taken into account in models. Included in the course are very useful techniques for modelling: non-parametric diurnal curves (Fourier series basis functions), disaggregating levels of system noise and observation noise from sensors (noise level functions in grey-box models) and, non-linear effects of solar radiation (Conditional parametric models with base splines).

Tracking changing model parameters over time in order to adapt to changes in systems and use, using recursive estimation techniques such as Kalman filters and recursive least squares.

Including weather forecasts as model inputs to create optimal forecast models.

Optimized control and operation of energy systems using model predictive control. Examples are optimizing heat pump operation with respect to varying price and optimal charging of batteries.

Each day will be max. 3 hours of lectures (e.g. 2 x 1.5 hours) and computer exercises for the rest. Computer exercises will be related to both building and district level (also building components, solar PV and wind).

Online participation will be via Zoom: Lectures will be streamed, group chat and individual help for exercises will be arranged.

4. Project results

4.1.1.1 Final reports

A final report for each subtask (with 1 and 4 compiled together in one report) were the main joint outcome of the Annex. In the following the main conclusions of each report are presented, all reports are available on the Annex website.

Subtask 1 and 4

This report has two main aims, to examine the inputs and outputs of the analytics which formed the main work of Annex 71. The inputs ranged from smart meters, BIM, weather platforms and on-board data from domestic systems. The current methods of the measuring an HTC on site were also considered and it was noted that this is also a rapidly developing in academic and industry.

Although one of the significant outputs of this Annex was of the course the ability to measure the HTC of a dwelling, it was also important to examine which parties might want this, and how it could also be used for several different purposes, ranging from identification of energy performance gaps, assisting in the accurate certification of dwellings and energy performance guarantees.

A wide-ranging review of the current in situ testing methods that are used for HTC measurement, examining the existing methods and cutting-edge methods, this work was compelling in that it showed that work is accelerating in this area which could grow into an industry.

It was found that there is a significant interest from industry and policy makers to use methods such as the ones that are being developed within this Annex, and some of the current methods reviewed in this document. However, there are some obstacles which are currently making this difficult: There lacks a common approach to the validation of the methods being developed, to provide levels of uncertainty and operating parameters (such as time taken for the test). This is required to allow for a degree of certainty in the methods, especially when we consider that more than one method could be used, and some being more capable with one building type than another. There is also an issue around the ability to actually use an HTC figure, most of the regulatory energy models around the EU and UK for instance do not allow for the HTC to be directly entered as a measured value. Some states have changed this around the air permeability figure which is now allowed to be manually entered. Future Work To address the obstacles that are presented further work is needed. A method of validating the uncertainty and effectiveness of the methods provided in the annexe is required, which may assist in the removal of the second obstacle of the lack of ability to amend regulatory certified values around HTC.

Subtask 2

IEA ECB Annex 71 focusses on an accurate characterization of the as-built energy performance of buildings based on in-situ measurements. This characterization may serve to identify the performance gap with focus on physical parameters or looks into the energy behaviour of the building. Subtask 2 focuses on the latter and assesses data analysis methods suitable for describing and predicting the energy dynamics of buildings. Knowledge on the energy behaviour of buildings is important in the ongoing energy transition. The shift towards renewable energy sources introduces a new paradigm where not only the amount of energy use but also the time of usage becomes significant and requires matching the energy demand with the intermittent production of renewable energy sources. During the operational phase, knowledge on the expected energy behaviour can be used to determine correct functioning of the building and its installations.

The activities in Subtask 2 were organized through setting up common exercises in which participants could contribute on a particular topic. The common exercises first explored the existing modelling techniques for building behaviour identification. The results of this analysis are reported in Chapter 2. Subsequently, two applications in which building behaviour identification plays an important role were identified. Fault Detection and Diagnosis (FDD) was chosen as a first application in which the behavioural models are used to detect and diagnose errors in the operation of the building and its systems. The results of this analysis are reported in Chapter. Chapter 4 focused on identifying models for Model Predictive Control (MPC) applications. The main findings on these two applications are summarized in the next 2 sections.

Conclusions on FDD

The goal of this activity was to evaluate and test the contribution of prediction models - obtained through building behaviour characterization – to automated fault detection and diagnosis. It was identified based on a literature review that FDD in commercial applications still recides on the individual component or system level, with individual processes being monitored by dedicated sensors. While building energy management systems are finding their way into the market, only recently methodologies that exploit overarching data are being developed.

The work carried out in this activity contributes to that development in three significant areas. First, a conceptual framework has been described that defines and organizes different types of faults. For each type of faults, different detection methods are proposed. The development and layout of this framework is inspired by the observations in the common exercises that demonstrated that despite detailed and accurate modelling techniques some type of faults could not be detected. For example, when using input-output models for anomaly detection that use the heating power as an input, it becomes impossible to detect errors in the control of that heating power (e.g. a thermostat malfunction). Consequently, as discussed in the framework, the type of errors to be detected should be properly matched to the modelling approach, taking into account the physical behaviour of the building.

Second, an overview was made on statistical methods used to detect the actual faults. The methods discussed focus on the detection of faults by comparing the predicted behaviour against the actual measurements. With this overview, guidelines are provided to move beyond the need for modeller/operator interpretation and move towards an automated detection process.

Third, by demonstrating the application of the identification of prediction models and the fault detection process for both simulation and actual measurement data for the same case study, this activity indicates that detailed building energy simulation models can play a significant role in the further development and research on automated fault detection and diagnosis methods. Further research is however needed to generalize these findings are a wider set of building and addressing a broader range of anomalies.

Conclusions on MPC

As a second application, it was chosen to determine how models used in Model Predictive Controllers (MPC) could be identified and what the impact of their predictive power was on achieving the goals set for the MPC. MPC was chosen as an application as it is a promising method for integrating Renewable Energy Sources (RES) and smart technologies in buildings. The developed common exercise aims at evaluating the performance of building behaviour models in MPC for the Holzkirchen Twin House O5 building and to demonstrate the opportunities of MPC. This MPC aimed at minimizing the heating system's energy cost while maintaining indoor thermal comfort.

An important question while developing a model for building energy assessment is: which model is better suited for the application at hand? Therefore, in this study we searched for a suitable Key Performance Indicator (KPI) to score the performance of predictive models for a popular building optimization application. This application is called Model Predictive Control (MPC), which has been proven successful in optimizing building's energy use while maintaining thermal comfort. The building is heated by an underfloor heating system, which is coupled to an air-water heat pump. A varying electricity price of the occupants in many cases. MPC employs a predictive model of the building to optimize its load profile was applied to mimic the dynamic behaviour of RES generation over a time horizon (Drgoňa et al., 2020).

In the exercise, an OpenIDEAS simulation model developed in Modelica serves as the emulator in which the MPC was implemented. Different modelling teams were asked to identify behavioural models based on data that were generated by the emulator. Subsequently, the provided models were implemented in the emulator. To do so, two options were provided: in option 1 partners provided their predictive models that were implemented in the MPC of the developed framework while in option 2 the participants could develop their own MPC and communicate to the building part in the emulator through an API. The exertion of option 2 revealed the need for an in-depth report of the emulator along with its inputs and outputs. Detailed report of the emulator becomes more relevant when one wants to compare controllers that have been implemented in different environments. Models which were studied for this exercise are amongst the most common modelling techniques used for developing a predictive model in the context of MPC; namely Grey-box RC models, AutoRegressive models with eXogenous inputs (ARX), State Space models (SS) and Artificial Neural Networks (ANN). At first, models' quality was reported based on their one-step ahead forecast but these reported accuracies did not reflect the performance of the resulting controller very well. As the predictive model is required to provide the controller with forecast of building's thermal behaviour throughout a time horizon, we opted to look into Multi-Step ahead Prediction Error (MSPE) as a performance indicator.). Evaluating MSPE of different models shed some light on the poor performance of controllers, which incorporated these predictive models. For example, ANN has the best prediction performance in terms of one-step ahead accuracy but the MPC, which uses ANN as its predictor, yielded the highest energy cost compared to other controllers, which deployed other modelling techniques. This could be explained by looking into MSPE of different models where we can see that ANN yielded the poorest prediction performance and therefore the MPC, which employed ANN model, led to the worst KPIs compared with other MPCs in this exercise. It was concluded that for scoring building energy assessment methods that are to be used in applications such as MPC, modellers should consider MSPE instead of one-step ahead prediction error. It has also been shown that MPC outperforms a well-tuned Rule-Based Controller (RBC) in this case by 22.7%.

Subtask 3

This document reported the work performed within Subtask 3 'Physical Parameter Identification' of the IEA EBC Annex 71 project. In this subtask, we investigated methodologies to identify the thermal performance of the fabric of residential buildings that are in use, and for which we have only limited monitoring data available. We started with deducing the full heat balance equation of a building to highlight the assumptions that are made in simplified models to determine the building's Heat Transfer Coefficient (HTC). We then, step by step, analysed the impact of different approaches to determine the input parameters to solve the heat balance equation with simple statistical tools. Both static (averaging and single and multiple linear regression) and dynamic (ARX and state space modelling) methods were investigated.

This explorative study was performed on five case studies, and for each input parameters (solar gains, the net heat input, infiltration and ventilation heat losses, occupant and appliance heat gains, weather data and indoor temperature) the impact of a modelling approach on the HTC-estimates was evaluated. Based on this information, the optimal approach was determined for each case study building and applied to determine the HTC as precise as possible. Since for all case study buildings, reference values were available, we could compare the different modelling outcomes with the target values. The results of the different static and dynamic statistical methods showed to be rather consistent, but deviated often significantly (20% or more) from the target value. This analysis allowed us to deduce general guidelines on how to deal with the different input parameters for specific cases.

In a final step, a blind exercise was performed on five inhabited dwellings in the UK for which indoor temperature measurements and smart meter readings were available. Based on the guidelines, an optimal approach was selected and applied to all five dwellings. Different participants participated in this blind test. It was found that the outcomes of the static methods were very consistent, while for the dynamic methods, which allow more freedom to the user, larger deviations were observed between the HTC's as estimated by the different participants. Hence, the static methods can be considered more robust in application, but overall both static and dynamic measurements resulted in similar estimates, which for some of the buildings were in close agreement with the target values (co-heating test results), while for other buildings deviations up to almost 50% were found. Analysis of the impact of the assumptions on the predicted outcomes revealed that assumptions on almost all parameters (measurement time and period, internal heat gains, temperature averaging,…) significantly impact the outcome.

Overall, it can be concluded that the statistical tools investigated in this work show promise to determine the building's HTC based on limited on-site monitored data. Care has to be taken, however, on the assumptions made with regard to the input parameters in the heat balance equation. A further in-depth analysis on more case studies is advisable to turn the methods into reliable tools to be used in actual performance assessment. In doing so, specific attention should go to an in-depth analysis of the uncertainty and repeatability before moving to large scale applications.

Building simulation validation

As a basis for this validation exercise for Building Energy Simulation (BES) programs a well-documented measurement dataset with synthetic users was created. This dataset has been made publicly available (Kersken & Strachan, 2020) and can be used for further validations, teaching and educational purposes and further research, especially focusing on the modelling of wet and dry screed underfloor heating systems, air source heat pump systems including domestic hot water, profiled internal heat and moisture gains and internal air flows through open doors and trap doors. There are no other comparable whole-house datasets with such detailed specifications publicly available. So, this new dataset itself is a valuable contribution to research and improvement of confidence in building simulation, as well as an important resource for model developers. The dataset is suitable for whole building modelling, but subsets could also be used to check sub-systems such as dynamic occupancy profiling or underfloor heating systems modelling. As was the case for the IEA Annex 58 datasets, future research publications are expected, including reporting of the sensitivity analyses currently being progressed by the experimental team and modellers The datasets created were also used by Subtasks 2 and 3 of this Annex as development cases ("Common Exercises") for their work on model predictive control (MPC), fault detection and diagnosis (FDD) and regarding the identification of building standard metrics, particularly the heat loss coefficient (HLC), from transient field data. To evaluate the teams' results handed in for the validation study two different metrics were chosen to be able to determine the results' bias (by mean deviation) and the dynamics' fit (by Spearman's ranked correlation) separately. For all experimental periods the relevant validation goal is analysed for the entire buildings' mean, for ground floor and attic and room-wise. The detailed and realistic validation experiment created within this IEA EBC Annex 71 contains numerous aspects that need to be / can be considered when modelling. These start with thermal bridges through the envelope and between the individual rooms and the cellar, internal air flow with operated internal doors, infiltration and natural ventilation by an operated window. Considering the electrical convector heaters, the dry and the wet screed underfloor heating, three heating systems need to be modelled together with a balanced mechanical ventilation system. Real outside weather conditions including time varying ground reflectance due to snow cover pose an additional challenge. A large number of time dependent inputs like set temperatures, heat and moisture inputs, ventilation rates, window and door operation schedules, etc. have to be included into the model, considering different operation modes for the six periods of the experiment.

This complexity, allowing for the validation of many modelling aspects under realistic but still well-known conditions on the one hand is an additional challenge on the other hand. All the different modelling aspects interact with the deviations and dynamics in multiple rooms. These interactions pose a difficulty when trying to identify one or more particular causes for observed differences between modelling and predictions. This makes it difficult for an individual modelling team to determine which aspects of the modelling need improvement. From the perspective of the analysts attempting to assess the performance of all the modelling teams' predictions, it is difficult to isolate the various approaches used by the modelling teams as a contributory factor to better or poorer predictive performance. Arguably, the limit of feasible complexity has been reached in this validation experiment. It is recommended that future validation studies should be designed with a reduced complexity. This does not necessarily mean reduced realism, but the experimental design should ensure a reduction in the number of parallel modelling changes at each step of the experimental schedule, by focusing more on specific effects with realistic boundary conditions. Most previous empirical validation studies used test rooms and test cells which are as simple as possible outside a laboratory context. IEA Annex 58 was a step up in terms of complexity by focusing on the envelope and solar aspects of a whole house, but in a simplified manner with no systems or variable occupancy profiles. It was shown that, with care, modelling could be successful and the Annex provided useful empirical validation datasets. IEA Annex 71 increased complexity with synthetic occupancy and systems. It was found that modelling, with much effort, could produce reasonable agreement with experimental data, but that the complexity (e.g. the number of dynamic model inputs required using measured data, which is not usual in design use) meant that user input errors became significant. It is not considered feasible to construct a suitable empirical validation experiment that increases complexity to a fully realistic building with real occupants. Another aspect related to the experimental design is the observed fact that the Open phase usually shows an improvement compared to the Blind phase but still shows significant deviations. This is another indicator that the chosen experimental design approaches a level of complexity that is challenging to meet even for experienced modellers. This should not be taken as an indication that BES tools aren't able to accurately represent complex situations: it is important to be aware that the modelling requirement in this experiment involves a high number of measured time-varying inputs, different modes of operation and provides a lot of necessary details in the experimental specification that of course all come with an uncertainty.

As mentioned, with this complexity, it is also a big challenge to identify the single effect or submodel that causes a certain deviation. Additionally, it should be mentioned that it is very challenging for the experimental team to perform such a complex experiment as a continuous, uninterrupted time series, as is required for modelling such a time series accurately.

In spite of the difficulties just described in general it can be stated that most modelling teams / programs were able to reach a good degree of accuracy in their simulations. Interestingly, the results for the electrical convectors equal the accuracy of the results for the more complex underfloor heating. The additional complexity of the underfloor systems is most probably compensated by strong stratification created by the convectors. The results for the User-1 period are 85more accurate than for the User-2 period, because the User-2 period has the more complex synthetic user profiles. For these User periods some teams for various reasons did not use the provided supply temperatures and flow rates but they used the rooms' set temperatures together with an individual control. Regarding the room air temperatures (the validation goal of the User periods), there is no clear advantage to either of the approaches adopted. In the Extended Experiment the O5 house features internal moisture gains instead of the underfloor heating. It is apparent that the results for this Extended Experiment, also regarding the relative humidity, are quite good. This is also true for the N2 building where no change between the two experiments was made. The reason for this is not certain – possibly due to modelling teams having results from the Main Experiment, or possibly because the weather conditions were different between the two experiments.

Concerning the modelling, it was interesting to compare program capabilities and the approaches adopted. For many of the heat transfer paths, the modelling approach was similar, e.g. for mechanical ventilation, fabric heat conduction and solar processes. However, there were significant differences in approach and assumptions for thermal bridges, stratification, albedo (due to snow cover), underfloor heating system modelling and infiltration leakage paths. It was noted that modellers often undertook detailed calculations of external shading (for example) which had a minor impact on predictions, but simplified or ignored some of the other important factors (e.g. stratification) which are harder to model. A recommendation from this study is that the modelling of thermal bridges, stratification, albedo, underfloor heating system modelling and infiltration leakage paths should be researched and improved in BES programs. There are still some unresolved aspects of the research. There are some experimental uncertainties, such as the in- situ construction U-values, which could be addressed by further data analysis, and the leakage distribution and interzonal air exchange, which could be addressed by analysis of the existing (and provided) tracer gas measurements or further experimental room-by-room pressurization testing. Regarding modelling predictions, causes of discrepancies between measurements and predictions could be investigated by additional sensitivity analyses. The extensive dataset gathered will be invaluable in further research to investigate these points, and additional validation research is possible, for example in analysing the heat pump performance and other topics as mentioned at the start of the Conclusion section. In near future the experimental team conducting this validation study will use sensitivity analyses, provided by some of the modelling teams, to compile an overview of the impact on the simulation results when one factor at a time (OAT) is changed. This is also a lesson learned for future application. Aspects that are expected to be critical should be modified OAT in the experimental design. Other aspects, identified to be often subject to poor numerical representation like interzonal air exchange and stratification, might by analysed by means of detailed and focused experiments or computational fluid dynamics (CFD). For building energy simulation this does only add a value when the information from the CFD is used to derive or improve simplified models that are feasible to be incorporated into BES tools.

Compared to the previous validation of IEA EBC Annex 58 the experimental setup chosen in this validation was comparable but with more realistic boundaries. As this required more details in the modelling this validation's results are a little less accurate in terms of quantifiable metrics, as can be expected. However, the visual inspection of the simulation results with the measurements still show a good agreement. One important finding is that results for cases with underfloor heating are accurate; partly exceeding the quality of predictions in

Annex 58. Despite the high level of complexity and realism and the difficulties that come along with it, the dataset is still very useful for education or training and to test and investigate various sub-models.

4.1.1.2 Danish partners results

The main research output directly resulting from the Annex project (together with other projects) from the Danish partners is presented in the following.

Annex Reports (joint work):

Four joined reports were the main outcome of the IEA EBC Annex 71:
ST1-4 Challenges and general framework
ST2 Building behaviour identification
ST3 Physical parameter identification
BES validation exercise
The Danish participation was in ST2, ST3 and the BES validation exercise.

Journal papers (Danish partners):

Morten Brøgger, Peder Bacher, Kim B. Wittchen. *A hybrid modelling method for improving estimates of the average energy-saving potential of a building stock.* Energy and Buildings, Volume 199, 2019, Pages 287-296, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2019.06.054.

Rasmussen, C., Bacher, P., Calì, D., Nielsen, H.A. and Madsen, H., 2020. *Method for scalable and automatised thermal building performance documentation and screening.* Energies, 13(15), p.3866.

Rasmussen, C., Frölke, L., Bacher, P., Madsen, H., & Rode, C. (2020). *Semi-parametric modelling of sun position dependent solar gain using B-splines in grey-box models.* Solar Energy, 195, 249-258.

Thilker, C. A., Bacher, P., Bergsteinsson, H. G., Junker, R. G., Cali, D., & Madsen, H. (2021). *Non-linear grey-box modelling for heat dynamics of buildings.* Energy and Buildings, 252, 111457.

Conference papers (Danish partners):

Wittchen, K. B., Jensen, O. M., Real, J. P., & Madsen, H. (2020). Analyses of thermal storage capacity and smart grid flexibility in Danish single-family houses. In BuildSim-Nordic 2020 Selected papers (pp. 131-138). SINTEF. BuildSim-Nordic 2020 Vol. 5

Rasmussen, C., Hviid, C.A., Bacher, P., Calí, D. and Madsen, H., 2021. Estimating Building Airtightness from Data–A Case Study. In E3S Web of Conferences (Vol. 246, p. 10004). EDP Sciences.

Jaume Palmer Real, Christoffer Rasmussen, Davide Calì, Henrik Madsen: Revealing the hidden dynamics of the energy signature model. Accepted for the BS2021 conference, September 1-2-3, 2021 Bruges, Belgium

Thilker, C.A., Bergsteinsson, H.G., Bacher, P., Madsen, H., Calì, D. and Junker, R.G., 2021. Non-linear Model Predictive Control for Smart Heating of Buildings. In E3S Web of Conferences (Vol. 246, p. 09005). EDP Sciences.

Erfani Beyzaee, A., Yu, X., Kull, T. M., Bacher, P., Jafarinejad, T., & Saelens, D. (2021). Analysis of the impact of predictive models on the quality of the model predictive control for an experimental building. Proceedings of Building Simulation 2021.

Other publications (Danish partners):

Book chapter

Thilker, C. A., Junker, R. G., Bacher, P., Jørgensen, J. B., & Madsen, H. (2012). Model Predictive Control Based on Stochastic Grey-Box Models. Towards Energy Smart Homes, 329–380. https://doi.org/10.1007/978-3-030-76477-7_11

Søren Østergaard. “Intelligente bygninger” i HVAC magasinet August 2020 Årgang 56.

5. Utilisation of project results

Subtask 2 results used in multiple current research projects with control of buildings for enabling energy flexibility. Subtask 3 results used in multiple projects on building energy performance assessment:

- Flexible Energy Denmark (Innovation Fund Denmark, No. 9045-00017B)
- Smart City Accelerator+ (EU Interreg, No. 20293290)
- Top-Up (Innovation Fund Denmark and ERA-NET, No. 9045-00017B)
- IEA Annex 83 (EUDP, No. 64020-1007)
- Several other running at DTU Compute.

The research results will definitely be used in many applications, both in companies and in public institutions. Current activities include Center Denmark, which will host much smart meter data, which can be a source of data for applying the developed methods, and Energistyrelsen is working on a building hub, where building data will be collected. All these activities will eventually lead to better information on building energy performance, which will be crucial and valuable information in the energy transition, where buildings energy consumption must be lowered substantially. Same goes for energy flexibility.

6. Project conclusion and perspective

The project has shown that it is possible to assess energy performance based on measurements from in-use buildings. It is proven by the common exercises that multiple factors impact the results, especially data preprocessing and choice of modelling technique. It is therefore concluded that standardization and uniform data formats must be developed to ensure consistency of the results. It’s not straight forward to do these tasks and the knowledge from this project must be used in future efforts to realize standards. On the other hand many of the modelling approaches can be directly used in energy services, hence the value of the work will be realized before standards are developed.

The First steps of gathering and evaluating different modelling methods for MPC and fault detection was performed. Gathering many researchers to work on the same problem through the common exercises was a strong aspect, since experiences could be shared directly. This led to the understanding of the problems faced in modelling buildings - one particular aspect is the diversity of buildings. They are machines build in many steps over time. They are not a standardized like cars! So, before the methodologies can be realized and scaled, data and system information must be organized. This is already happing with BIM, however this is mostly available for newer buildings – to realize optimization and control in older buildings simple and robust methods must be developed.

The next steps are to continue the work and efforts with the established network through the international Annex. Many activities are ongoing, for example the BOBTEST framework (see link in Appendix), which focus on developing virtual environments for control and operation testing. This emphasize a particularly important point, which is, that real experiments are very costly in terms of time and complexity, hence a virtual test setup will be a great way to develop basic concepts of control methods, and then after testing them there, they can be moved to reality and tested.

The summer school and training activities must also continue to harvest the full value of the project. We build up the summer school in the project and we will continue arranging through projects in the future. This will give the acquired knowledge and experience to new engineers, who can bring it to the next level and build useful services with many applications in the future.

7. Appendices

For publications from the DK partners see the list in Section 5. Furthermore, a summary report will be available online.

Some links:

- **Annex 71 website:** https://bwk.kuleuven.be/bwf/projects/annex71/ and https://iea-ebc.org/projects/project?AnnexID=71
- **Summer school:** https://drive.google.com/file/d/1jJXH9i39jNe1w82Mn19GLb5bZGHxjJwb/view
- **BOPTEST:** https://www.energy.gov/eere/buildings/boptest-building-operations-testing-framework
- **Video presentation of Annex 71:** https://www.youtube.com/watch?v=lXN4JDlzuZc