Database management

Report from EPBD Concerted Action II plenary meeting, Berlin 15.-16. June, 2009

Wittchen, Kim Bjarne; van Diggelen, Leanne

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Database management

Meeting	**Berlin**
Session	**16 June 09:00-12:30**
Chair	**Kirsten Engelund Thomsen & Hans van Eck & Carmen Heinze**
Rapporteur	**Kim B. Wittchen & Leanne van Diggelen**
Reviewer	**Christina Spitzbart**
Date	
Files in PC	

LIST OF CONTENTS

1 Executive SUMMARY

Many MS have recognised the necessity or the advantages of establishing a database for collection of data from the energy performance certification (EPC) schemes. Data have started to enter into the databases, difficulties have been identified and solutions to overcoming them found. Most MS have however been very busy taking care of the implementation of the EPC schemes and have thus not had time or resources to investigate the possibilities for exploiting the information in the database for other purposes than meeting the requirements in the EPBD. Research groups and other organisations have though shown interest in the content in the database and it is evident that MS that have central registers for EPC data will be able to gain increased knowledge about the energy performance of their building stocks.

A short questionnaire was circulated in advance of the CA2 Berlin meeting in June 2009 to investigate the current status of EPC databases and database management in the MS. The questionnaire (see Annex 1) comprised only 10 questions that could be answered for MS using both calculated and/or measured EPC.

The majority of the MS (18) collect EPC data in a central register managed by an official authority. There are however MS that do not have a central register and some MS employ a private company to manage the register. Almost all MS collecting EPC data in a register have arranged that data flows directly from the experts or assessors into the database and in most cases by using an accredited EPC tool. The data being collected are in most cases all the information gathered while carrying out an energy performance certification. There are however a few MS that only collect information about the label value (energy performance id). Most of the MS do perform some kind of quality checks on data. In some cases before data are entered into the database, but in most cases as a retrospective exercise with selected (randomly or after complaints) certificates being investigated. Most of the MS have not tried to extract data from the database to re-create a full EPC and only eight MS have tried to use EPC data for other purposes than those strictly related to EPBD.

Presentations filled the first half of the double session . First there was a summary of the findings from the questionnaire survey and then the invited guest Mr Tobias Loga presented the findings from the DATAMINE project that has created a generic database structure. Then there were presentations of the situation in four MS: Sweden, Denmark, Portugal, and Germany. These four MS have different approaches and experiences regarding EPC database management. The second half of the session was arranged as a brainstorming and discussion session with the attendees assembled in six groups, each discussing a separate topic of interest. At the end of the session each group presented their suggestions for an optimum solution related to their topic.

2 Database management questionnaire survey

A well organised and structured management of the databases holding information from the energy certification (EPC) schemes in the MS is crucial for being able to utilise the content of the databases for other purposes than for storing information from the schemes.

This technical session deals with the different perspectives and ways of handling databases for the energy certification schemes.

Prior to the meeting, a short questionnaire was circulated to one key person in each MS plus Norway, Croatia, and Switzerland. Twenty-seven of the 30 countries answered the questionnaire. On the next pages a summary of all the answers is given with the following abbreviations for obtaining

the buildings energy performance: C = calculated energy performance; M = measured energy performance; B = both calculated and measured energy performance under certain circumstances. Not all answers add up to 27, in some cases MS have not answered all questions and in some cases more answers are give for one question or the answer is given for both calculated and measured energy performance. The number of answers given for each question can thus be both below and above 27.

In the following a summary of the individual questions is given.

2.1 Energy performance identification method

To be able to understand the possibility for collecting and utilising data, the energy performance identification method must be known. In 16 of those MS who answered the questionnaire, a calculated energy performance is used, while 5 MS use a measured energy performance. In 10 of the MS both calculated and measured energy performance is used and collected in the database, though often depending on the actual conditions and circumstances.

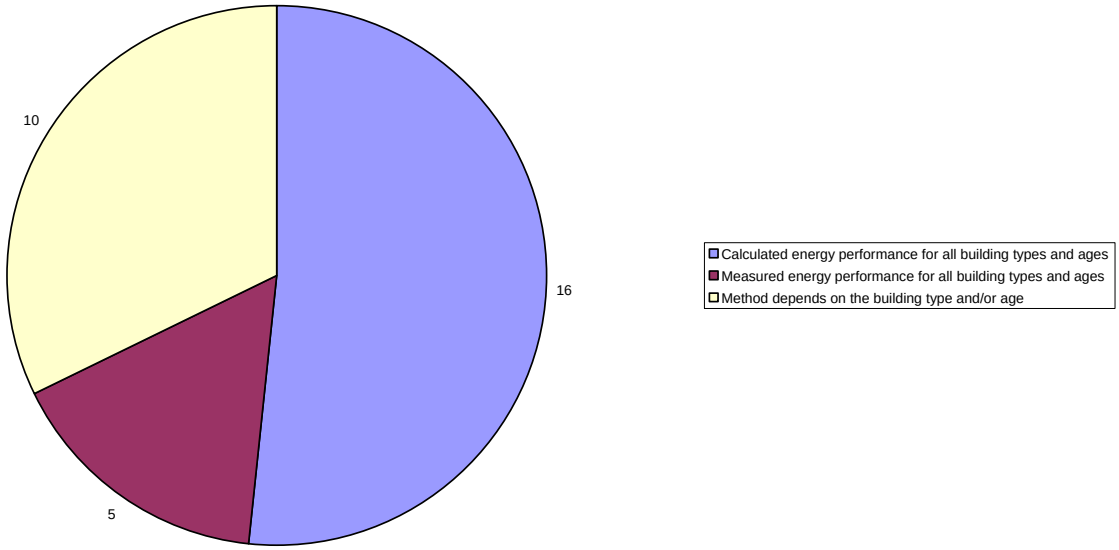


Figure 1. Energy performance identification method.

Country	Austria	Belgium (Flanders)	Bulgaria	Cyprus	Czech Republic	Denmark	Estonia	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Latvia	Lithuania	Luxembourg	Malta	Netherlands	Poland	Portugal	Romania	Slovak Republic	Slovenia	Spain	Sweden	UK (England&Wales)	Croatia	Norway	Switzerland
1. What method do you use for energy performance certification?																														
Calculated energy performance for all building types and ages	X		X	X	-	X					X			X	X	X		-	X	X	X	X	X		-			X	X	X
Measured energy performance for all building types and ages			X		-										X			-					X		-	X				X
Method depends on the building type and/or age		X			-		X	X	X	X		X	X				X	-						X	-		X			
2. How is the database for collecting building energy performance certification data organised?																														
One nation wide database			B	C		C	B		B		C	C	B		B	C			C		C	C		B		B	B	C	C	
Regional databases	C	B												C																
Multiple databases, one for each accepted certification tool																														
Multiple databases, one for each certification method (calculated/measured)																														
No collection in database								B									B			C			B							
Other	C									B				C															M	B
3. Who is responsible for the database in your country?																														
Central, official authority			B	C		C	B		B		C	C	B	C	B	C			C		C		B	B		B	B	C	C	
Regional, official authority	C	B												C																
Private company/companies	C																							B						M
Other										B												C								C
4. How are data collected in the database?																														
Automatically reported from the accredited certification tools		B		C		C							B						C		C						B			
Reported directly from the consultant/expert/assessor			B				B		B		C	C		C	B	C			C			C	B	B		B		C		M
Central secretariat transferring data from the certificates to the database																														
Other	C									B																			C	

Country	Austria	Belgium (Flanders)	Bulgaria	Cyprus	Czech Republic	Denmark	Estonia	Finland	France	Germany	Greece	Hungary	Ireland	Italy	Latvia	Lithuania	Luxembourg	Malta	Netherlands	Poland	Portugal	Romania	Slovak Republic	Slovenia	Spain	Sweden	UK (England&Wales)	Croatia	Norway	Switzerland
5. If you collect data in a database, which data do you collect?																														
All information collected during inspection (U-values, areas, efficiencies, etc), energy label		B				C					C	C	B		B	C					C	C	B	B		B	B	C		
Only energy performance, label and recommendations			C						B					C																
Energy performance	C									B																				M
Energy label												C							C											M
Only building id																														
Other	C						B			B			C			C			C										B	M
6. Quality check on data (tick all that apply)																														
Probability check on entry, e.g. acceptable value range for different parameters		C								B			C			C								B			B		C	M
Field compliance check, e.g. no text in numerical fields		C								B						C			C					B		B	B	C	B	M
Check that all requested data is available before entering the database		C	C				B		B	B			C			C			C					B		B	B		C	
All data from a certificate is rejected if crucial data is missing			C										B			C			C							B	B		C	
Generation of statistical information and cleaning of data and identifica-tion of out-of-range data after input																														
Manual cleaning of data and identification of out-of-range data after input										C														C			C			C
Other										C														C		C	C		C	C
7. Have you tried to generate a full certificate from information in the database?																														
Yes, with success													B								C					B		C	C	M
Yes, but with limited success			C			C																		B						
No	C			C			B		B	B	C	C		C	B	C			C			C					B			
8. Have you tried to use information in the database for other purposes, e.g. calculation of national saving potentials or market penetration?																														
No		B		C			B		B	B	C	C	B		B	C						C		B			B	C	B	
Yes	C		C			C								C					C		C					B				M

2.2 Organisation of the database

Databases can be organised physically in different ways ranging from one country-wide database holding all information from the certification schemes to no collection of information in a database. Most of the MS (18) that answered the questionnaire collect data in a country-wide database. In many MS the organisation of the database and the database structure is still under preparation and therefore it is important to share this kind of information at this time when things can be changed.

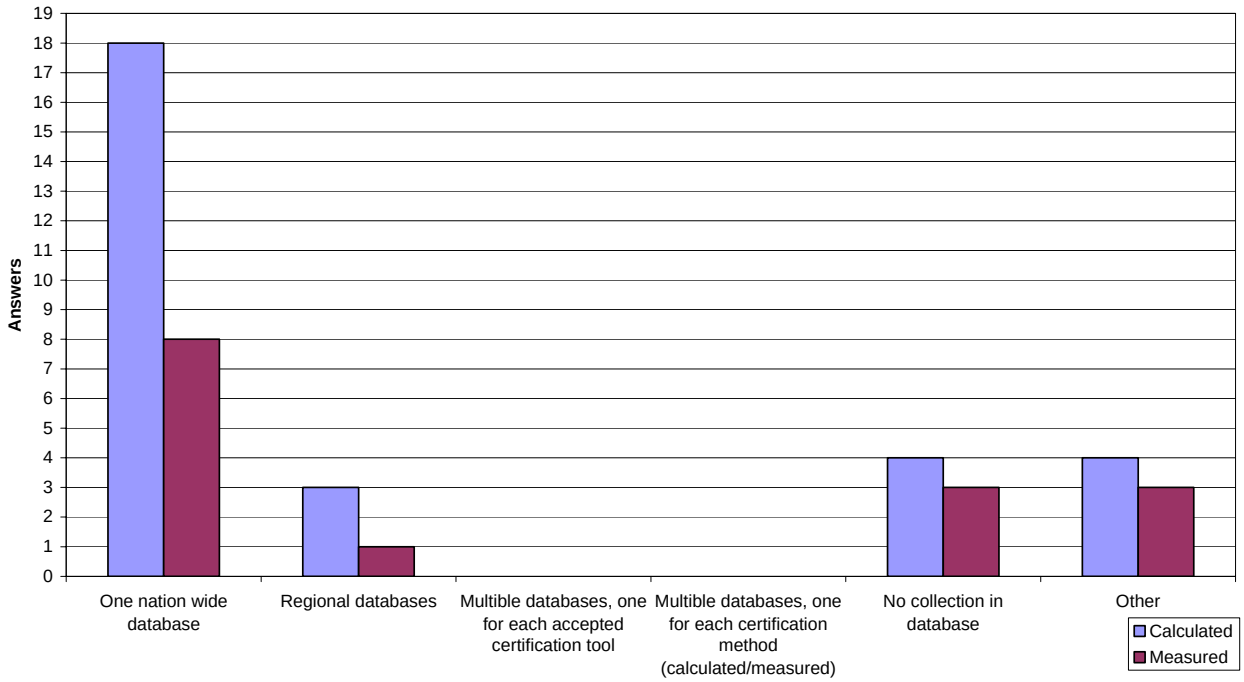


Figure 2. Organisation of databases in the MS.

2.3 Organisation responsible for the database

To get easy access to information in the database, especially for other purposes like analyses of the general energy performance of a country or region or energy saving potentials, it is important to know the organisation responsible for the database. In most of the MS (19) that answered the questionnaire there is an official authority responsible for the database, either at a national level or at a regional level. There are however three countries (Austria, Slovenia, and Switzerland) that answered that a private company managed the energy certification database (calculated and/or metered). Slovenia will have both a national database and require that the experts store the information gathered during the energy certification audits. Austria has several regional databases and one voluntary central database run by the Austrian Energy Agency.

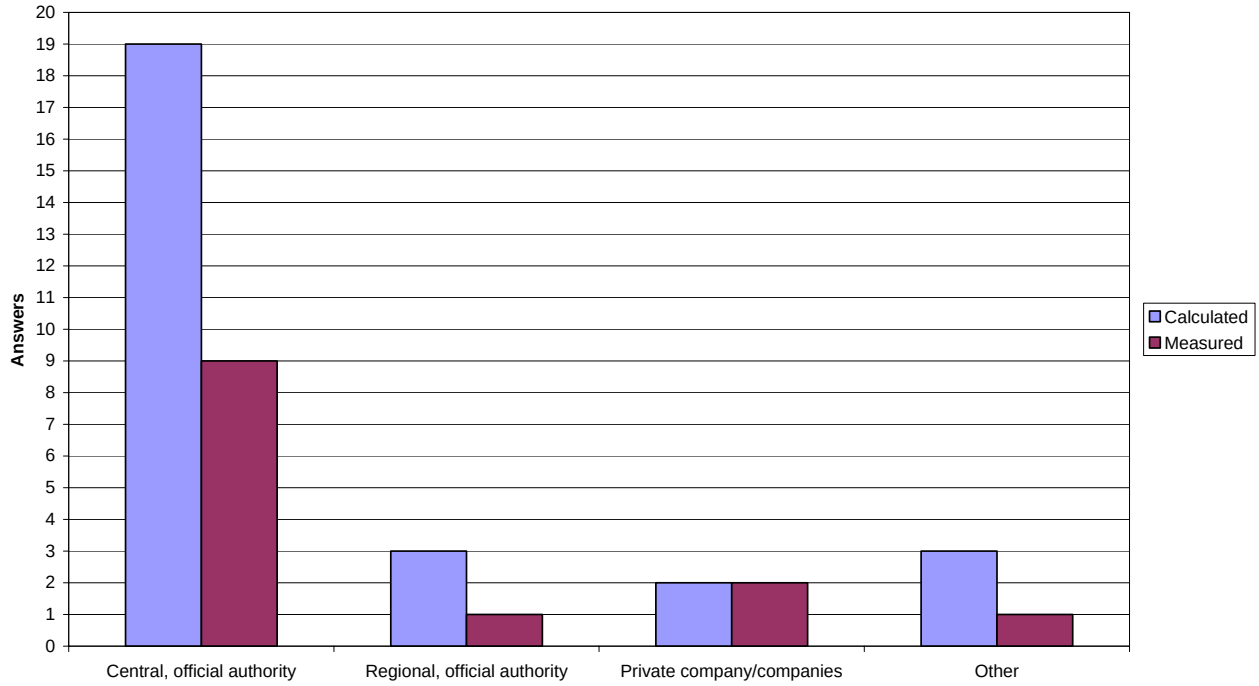


Figure 3. Responsible organisation for collection of EP data in a database.

2.4 Data collection

The way data enter the database may indicate how possible problems with data can be dealt with now or in the future. In most MS (21) data are entered directly by the assessor or expert into the database or data comes directly from accredited certification tools.

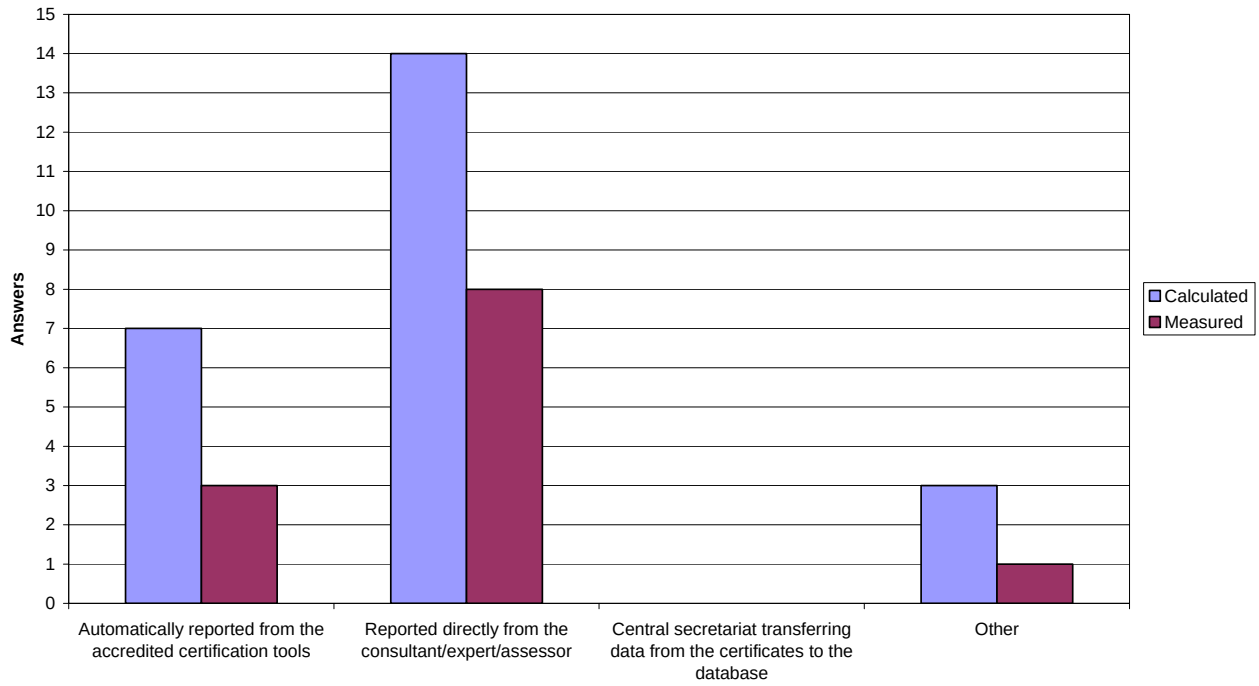


Figure 4. Different paths for data to enter the database.

2.5 What data go into the database?

To be able to use data from the energy certification schemes for other purposes than pure statistics, it is important to collect all the relevant information that is made available by the certification

process. The possibility for sophisticated statistics and analyses thus depends on the available data and the quality of these data. In most of the MS (13) that answered the questionnaire, all information collected during the certification process and audit in the buildings are stored in the database. There are however 8 MS that answered "other" or gave specific comments to this question for calculated energy performance and 3 for measured energy performance. These comments are interesting for those MS that are in the process of establishing or improving their database and are thus given in the following:

– Austria (Calculated): Technical information needed to calculate the energy performance (U-values, etc), information about the building (address, climate data, owner, etc.).

– Belgium (Flanders) (Both): All data is the aim. For new buildings, this is not yet so, but we work on it step by step.

– Estonia (Both): Date of issue of the energy performance certificate, the class (energy performance value) of weighted specific energy use from A to G (or H non-residential buildings) (kWh/m² year), name and registry number of issuer (artificial body) and information about source of heated surface of building (only for existing buildings). The heated surface can be selected among pre-determined options.

– Germany (Both): Germany does not have one official, central database. Within dena´s private database, all data of the mandatory certificate are collected in an anonymous way (postal code of the building is captured only). So the data cannot be related to a certain building or issuer (only to the postal code). The data of certificates issued according to the standards of the voluntary seal of quality are fully captured including U-values, areas, efficiency etc. These data are anonymous as well, but the data can be related to the issuer in charge (in order to allow for samples and quality follow-ups).The recommendations given within the certificate with the seal of quality can also be related to the issuer (and the postal code of the building).

– Ireland (Calculated): For residential buildings: all input data and results are collected. For commercial buildings: some input data and all results. For large public buildings: all input data.

– Lithuania (Calculated): Energy performance of buildings and all other information are stored in separate files but not as a database.

– Netherlands (Calculated): For example energy Index, calculated energy consumption, survey date, creation date, surface area, year of building, year of renovation, building type, type of ownership, heating device type and location, cooling device type, main application for commercial buildings, advisor, address, postal code, building id., date of construction, heated surface area, number of buildings, temperature, type of heating, type roof, year of renovation, capacity appliance, capacity pump.

– Norway (Both): Calculated: For so-called simple certification, all information gathered will be stored. For complex buildings with detailed building information, only information at a certain aggregated level is stored. Measured: When reported: Measured annual consumption per energy carrier.

– Switzerland (Measured): Building id, floor area, building type, construction or renovation year, energy uses of all marketed energy agents. Renewable energy uses and water are optional. No idea about what data will be kept in states databases.

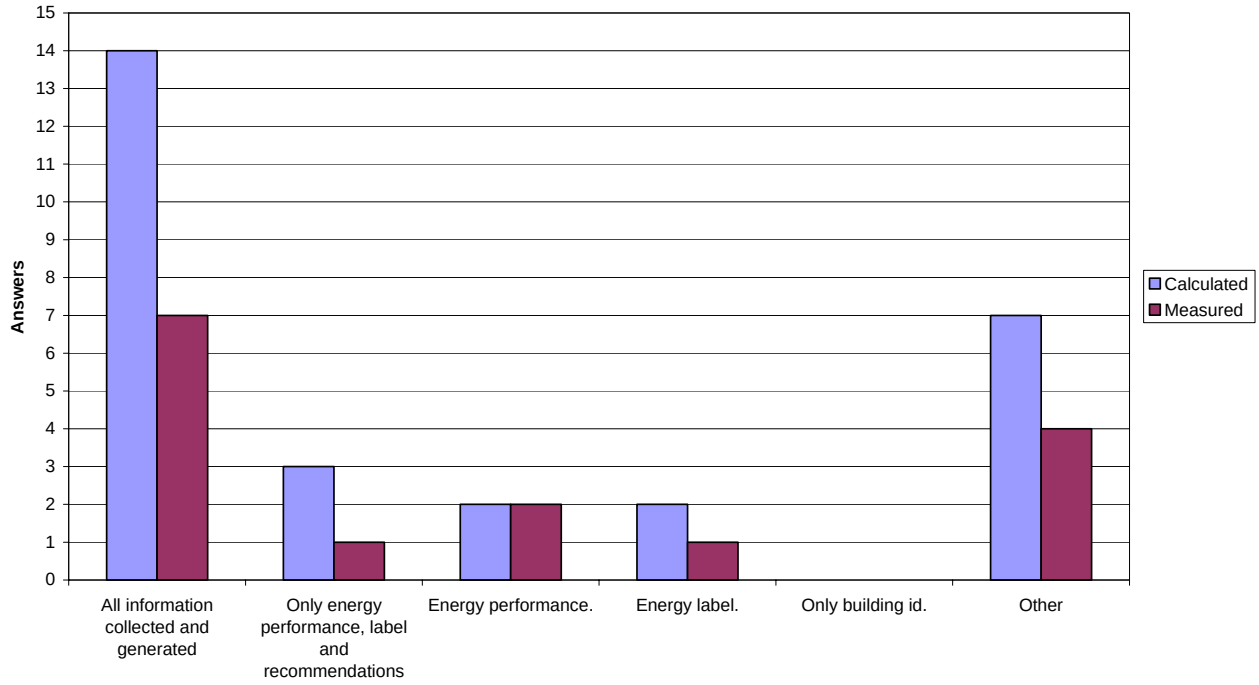


Figure 5. Data stored in energy certification databases.

2.6 Quality check of data

One issue is collection of data another is quality control of the data that actually enter the databases. A database can be nicely organised, easy to access and run queries in, but if the quality of data in the database is dubious, the database is of no use.

This question asked MS to identify what types of quality check are performed before and after data enter the database.

Some MS make a big effort to ensure that there is a high data quality. Six MS check probability on entry, field compliance and that all requested data are available before any data are let into the database. Another approach chosen by other MS or as a supplement to the above arrangements is to reject all data in a certificate if crucial data are missing. At the other end of the quality assurance scale, some MS do not perform any checks on data before entering them into the database or they have given no answer the question.

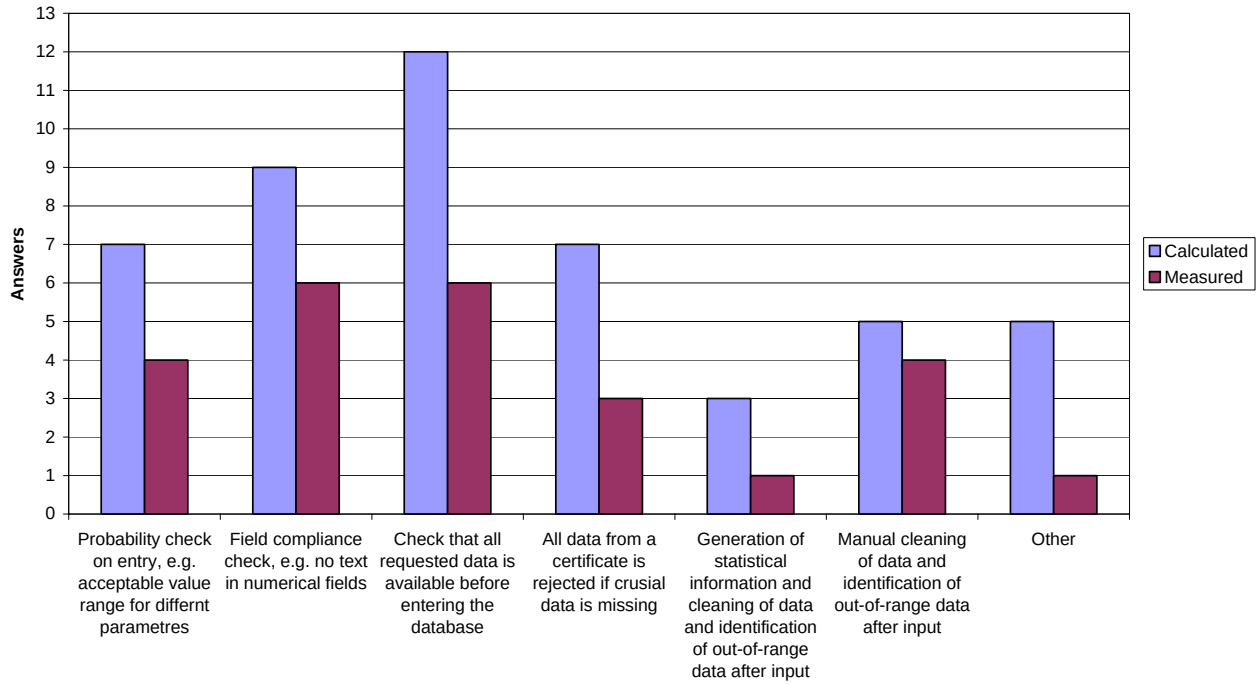


Figure 6. Quality check of data before and after entering the database.

2.7 Extraction of data

Data are collected in databases to facilitate their use for different purposes and one of these purposes could be to extract adequate information to be able to create a full certificate from information in the database. The majority of the MS (13) answering this question have not tried to create a certificate from data in the database, and among those MS that have tried to create a certificate from the database, two thirds experienced difficulties.

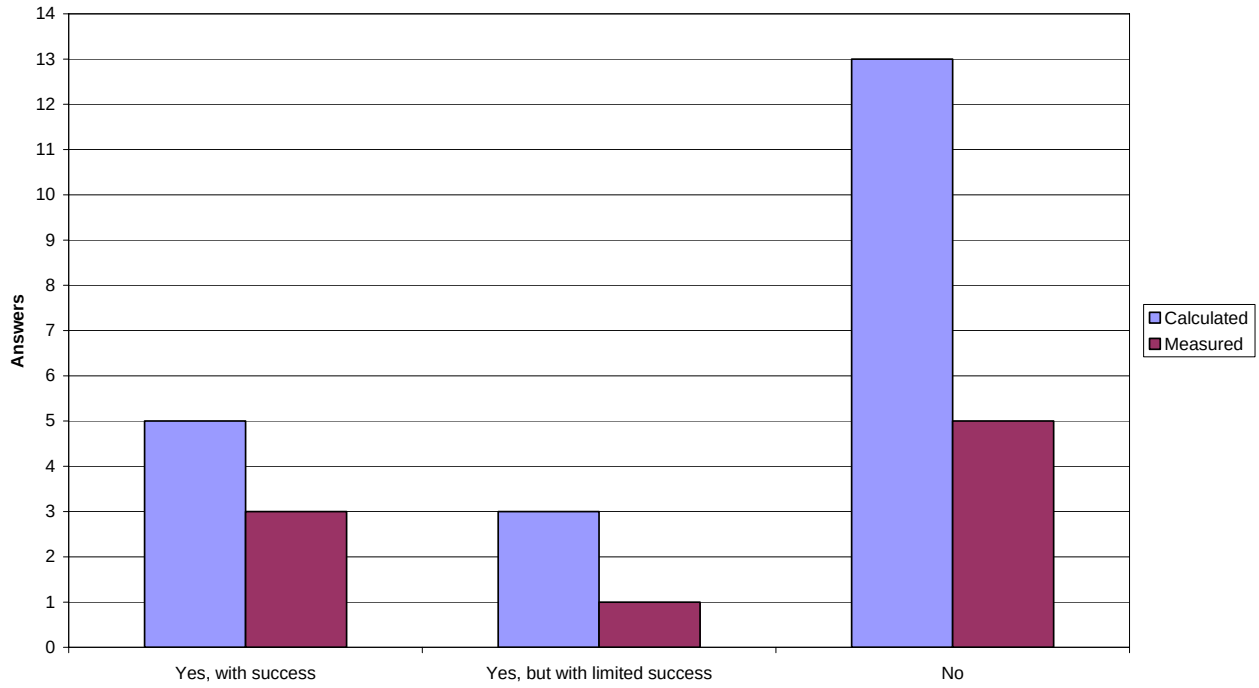


Figure 7. Have you tried to generate a full certificate from information in the database?

2.8 Other uses of data in the database

One of the new opportunities available when establishing a database with information from the energy certification procedure is to use the data for other purposes. These can be statistical analyses of the energy quality of buildings, e.g. distributed on regions; analyses of energy saving potentials and the necessary investments connected with it; market penetration of the energy certification scheme; etc. Most of the MS answering this question have not made this kind of analyses yet. However most MS acknowledge the possibility and plan to perform such studies in the future when more data are available in the databases.

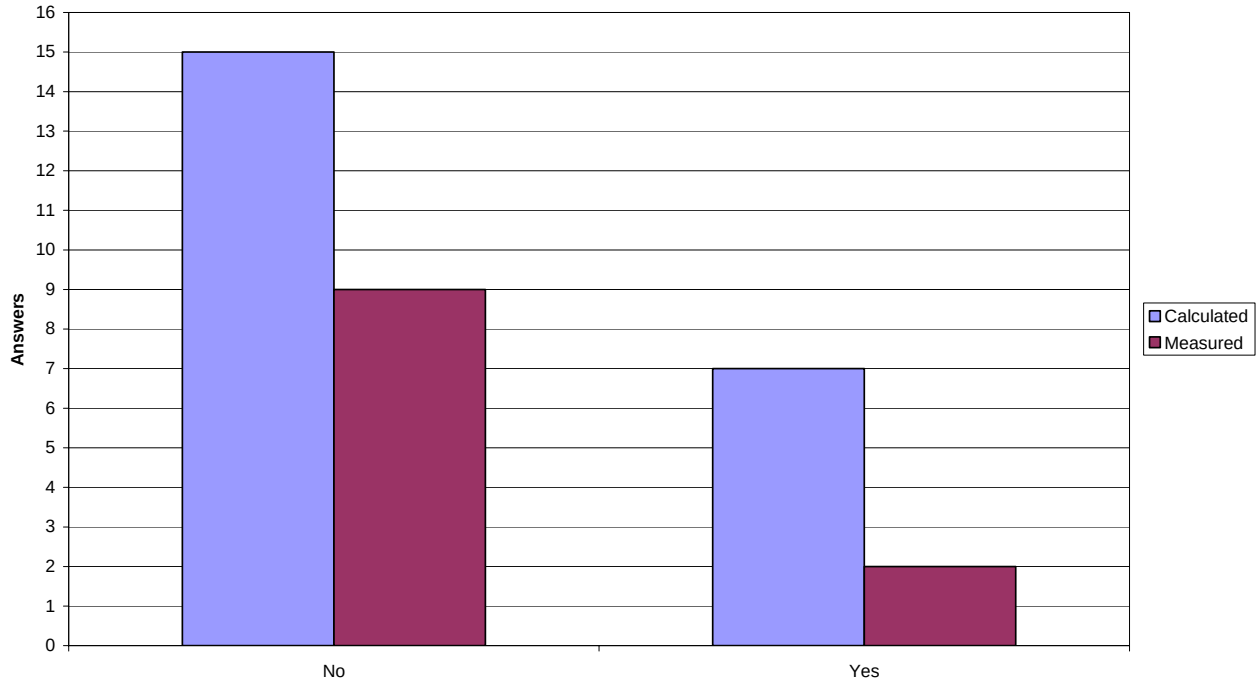


Figure 8. Have you tried to use information in the database for other purposes, e.g. calculation of national saving potentials or market penetration?

The detailed answers from the MS answering "Yes" to the question about having tried to use EPC data for other purposes are listed below:

- Austria: Information stored in the database allows statistical analysis and therefore conclusions on the success of legislative or other measures. These analyses are conducted by the regional authorities. Within the IEE project ESAM (Energy Strategic Asset Management) a tool has been developed that uses the information from energy performance certificates stored in the database for an asset management tool that lists buildings according to their energy saving potential.
- Bulgaria: We use the information for the preparation of the national programme in buildings.
- Denmark: Calculation of country-wide energy saving potentials for existing buildings grouped by their different use and construction age.
- Italy: Statistics for incentives planning, statistics on refurbishment type's frequency, extent and location.
- Netherlands: The data is used (or planned) for calculation of the results of the convenant MmM (= treaty about energy saving in buildings with commercial companies and the Dutch government), subsidy for home/buildings owners who invest in their home/building, differentiating in pricing for social housing (Woningwaarderingsstelsel), market penetration.

– Poland: Until now, only a limited output, like number of certificates issued, percent for each rate, etc. A business information tool to work on the database to extract all types of reports, score-cards and dashboards is now under preparation.

– Sweden: We only just started to evaluate statistics and possibilities, but several studies have been initiated regarding saving potentials and quality assurance.

– Switzerland: We did use and still intend to use the database for improving our knowledge on energy use in buildings, in particular for uses other than heating. This information will be then used to improve the certificate, in particular to provide a more realistic A-G scale.

2.9 Do you meet difficulties in database management and what sort of problems do you have?

The individual answers to this question are shown in the table below.

Table 1. Answers from MS to the question: Do you meet difficulties in database management and what sort of problems do you have?

Austria	Different regional databases instead of one central database. No automatically quality control possible. No central statistical analyses possible.
Belgium (Flanders)	Amount of data, database is growing too fast. Good way to open up the data. Transfer of all the data.
Bulgaria	Information is collected in Excel files and the work with this database is very slowly.
Cyprus	Database not built yet
Czech Republic	
Denmark	Some numbers are stored in text fields, which makes it difficult to sort and calculate.
Estonia	So less information for proper analysis. Transferring the information to the database is not completed and this work was started in February 2009.
Finland	-
France	-
Germany	Quality of data. Data are submitted on a voluntary basis, so the database does not cover the data of all mandatory certificates
Greece	-
Hungary	The process started a few weeks ago, till date no experience
Ireland	-
Italy	Dis-homogeneity of information in the different databases
Latvia	Will be specified after system testing phase
Lithuania	-
Luxembourg	-
Malta	
Netherlands	There are no difficulties. The systems functions.
Poland	-
Portugal	Some difficulties, although limited, like assuring a uniform input of information by the experts in the database
Romania	The database management is under development and implementation.
Slovak Republic	-
Slovenia	-
Spain	-
Sweden	Mainly on quality assurance of data and recommendations, but also simple procedures.
UK (England&Wales)	Some commentators would like to see more access to the database.
Croatia	No problems, database is in preparation, no certificate issued until now
Norway	-
Switzerland	Getting the information.

2.10 How would you improve your database management system in the future?

The individual answers to this question are shown in the table below.

Table 2. Answers from MS to the question: How would you improve your database management system in the future?

Austria	Central quality control (probability check, identification of out of range data). Make regional databases compatible so that central quality control and statistical analyses are possible.
Belgium (Flanders)	New tool for statistics and calculations to implement. Maybe in future data warehouse system in which energy performance databases are loaded.
Bulgaria	All information of the certification of the buildings we will organise in a National information system on the base of the GIS at the end of the year.
Cyprus	Database not built yet
Czech Republic	-
Denmark	A project is being launched now to improve the database and the screening of data. The project is expected to be finalised late summer 2009 and will include a re-input of all certificates found in the database today.
Estonia	There would be more proper information: heated surface of building, type of building, weighted specific energy consumption value, suggested energy saving measures.
Finland	-
France	-
Germany	XML-Standard More automatic quality checks (in case of the voluntary seal of quality)
Greece	-
Hungary	Professional group has a clear idea; however it does not coincide with the plans of the authorities.
Ireland	At present the public buildings is in a smaller standalone d/b, it is intended to incorporate this into the NAS d/b
Italy	Better coordination and homogeneity among different databases
Latvia	System to be elaborated
Lithuania	We plan to improve our database with more narrow classifications of buildings.
Luxembourg	-
Malta	-
Netherlands	To create a more extended back-up system for the database. To add the mentioned energy saving measures, the type of Hot Water Device', 'Type Ventilation' and 'Type Renewable Energy'.
Poland	-
Portugal	Increase the number of variables that are inputted based on pre-selected options (not as free text), including some typical/standard improvement measures.
Romania	To develop an online secure system in order to facilitate data input directly by the consultant/energy auditor.
Slovak Republic	-
Slovenia	-
Spain	-
Sweden	Simplifying procedures, adjusting database interface for specified building types/ owners, adding the possibility of including calculated values for example.
UK (England&Wales)	-
Croatia	First of all we have to start with estimated database
Norway	-
Switzerland	Depends on what happens at the states level.

3 Programme of the session

During the first part of the technical session (16 June 09:00–12:30) there was a short recap of the questionnaire survey and country presentations from Portugal; Germany, Sweden and Denmark. In addition to this, Mr Tobias Loga from the Institute for Housing and Environment in Darmstadt, Germany, gave a presentation on the findings and recommendations of the European DATAMINE project. In the second part of the session there was an interactive session arranged by Core Theme 5 where the following subjects were suggested for discussion:

– What data to collect from the certification schemes and why?
– How should data be transferred from the certification scheme to the database?
– Quality check of data?
– How could the running cost for operating the system be optimised?
– How can you use the information for other purposes?
– Public access to data?

4 Content of Presentations

First Mr Tobias Loga presented the DATAMINE project and afterwards the situation in four MS namely Sweden, Denmark, Portugal and Germany was presented. Each of the presentations was followed by a couple of short questions and answers.

4.1 The DATAMINE project

Invited guest Tobias Loga from the Institute for Housing and Environment in Darmstadt, Germany gave the presentation "DATAMINE - Collecting DATA from Energy Certification to Monitor Performance Indicators for New and Existing buildings".

The goal of the DATAMINE project was to set up a structure to collect data from EPC like procedures to be able to facilitate cross country comparison of energy performance. To be able to do that, development of a harmonised data structure was essential holding all necessary information yet not restricting input in such a way that only a limited number of datasets could enter the database. The starting point was thus identification of where data can be used and which data to collect to be able to facilitate achieving the goals. First a harmonised data structure was developed for establishment of a model for data collection. Then carrying out national model projects to verify and fill the database and finally performing a cross country comparison over the steps in DATAMINE.

The intention of DATAMINE was not a monitoring system for all of Europe, but rather a bottom up approach offering a common database structure to facilitate cross-country/region comparisons of building energy performance.

The structure in the DATAMINE database is based on information divided into the following categories:

– Energy performance certificate data,
– General data of the building,
– Building envelope data,
– System data,
– Standard calculation of the energy demand (Asset Rating),
– Basic parameters gained from Operational Rating,
– Summary of energy consumption and energy generation (calculated and measured),
– National/regional definitions of primary energy, carbon dioxide emissions and benchmarks.

To meet all requirements for a standard calculation and all national specifications for collecting EPC data, the DATAMINE database ended up having 255 input fields. It is not necessary to give data in all fields; there are for instance four different definitions of the floor area (conditioned gross floor area; conditioned floor area; conditioned useful floor area; and conditioned living area) and only one of these needs to be filled in for each dataset. Many input fields offer selection between different options, e.g. the main utilization type of the building or the construction period (Figure 9).

The participants of the DATAMINE project collected datasets for a large number of different building types and made their own evaluations in accordance with the need for information of the involved national key actors. Since the data structure was the same for all 12 countries, the databases could easily be merged in order to make cross-country comparisons. The figure below shows the number of datasets and types of buildings that have been handled in the project.

Total number of collected datasets			**19095**
Certificate types		**Utilisation types**	
whole buildings	10927	residential buildings	17727
building parts	0	offices	215
apartments	8168	education	612
Rating types		higher education	137
only asset rating	17542	hospitals	68
only operational rating	1112	hotels and restaurants	15
both asset and oper. rating	421	others	378
Considered energy uses		**Buildings constructed ...**	
heating	19053	1900 or earlier	160
hot water	18679	from 1901 to 1940	352
cooling / air conditioning	524	from 1941 to 1980	8616
lighting	10614	from 1981 to 2000	3920
others	10504	since 2001	4313
Contribution of the Model Projects			
MP 1 Germany	515	MP 7 Belgium	113
MP 2 Poland	133	MP 8 Austria	6715
MP 3 United Kingdom	302	MP 9 Slovenia	100
MP 4 The Netherlands	10109	MP 10 Spain	50
MP 5 Italy	188	MP 11 Ireland	126
MP 6 Greece	250	MP 12 Bulgaria	494

Figure 9. Datasets for different building types collected in the DATAMINE project.

It is possible to perform different analyses using the tools developed during the DATAMINE project. Among these tools are standard calculations of energy performance and cross-country/region comparison of the results (see figures below).

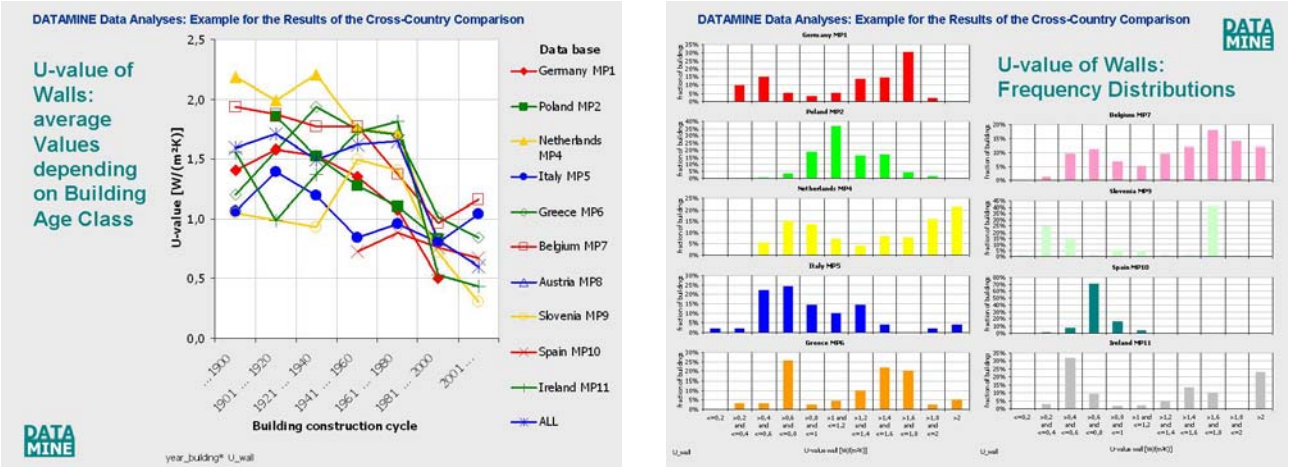


Figure 10. Example analyses, given the precondition that all countries use the same data structure.

Different analyses were performed for this merged database, e.g. comparisons of the U-values by building age class, the U-values of walls by frequency distributions (see Figure 10), or the calculated energy need for heating by the building's thermal envelope (see Figure 11). Also analyses for used systems or energy carrier types were performed.

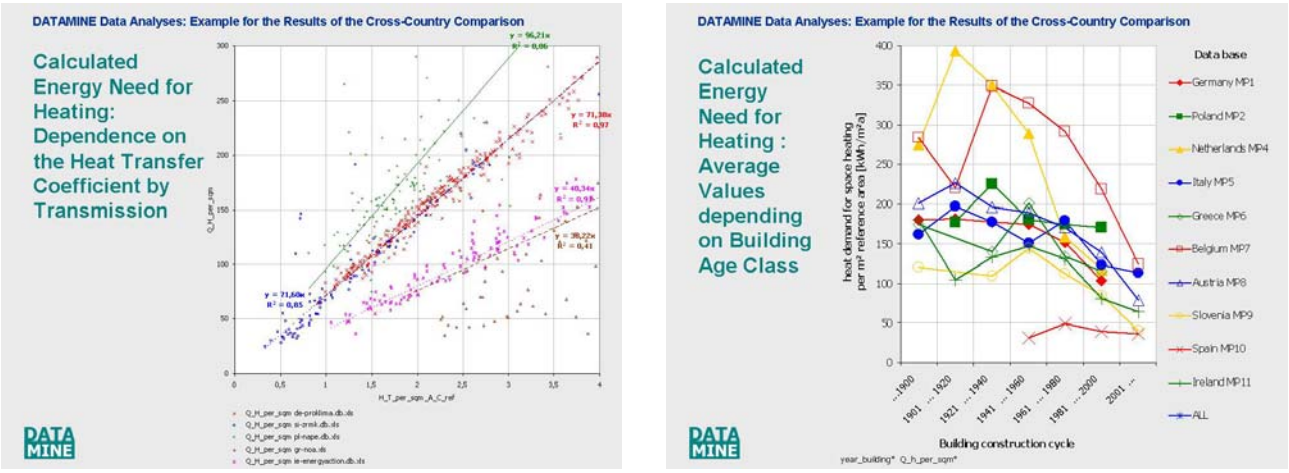


Figure 11. Analyses of the energy performance calculated by the respective national EPC procedures and comparison between different countries / regions.

One of the main results of the DATAMINE project was a set of recommendations for national key actors who intent to build up or improve energy efficiency monitoring activities in the building sector of their country. The DATAMINE partners recommended:

- to use the common DATAMINE data structure in order to make monitoring databases comparable
- to use national energy certificate databases for making statistical analyses
- to provide house owners with the complete information about their buildings in order to facilitate an EPC updating and also to enable representative surveys for regions or whole countries.
- to develop concepts for the monitoring of the national building stock using EPCs but also other available data sources
- to carry out representative surveys in order to derive the information which is not yet available
- to develop building typologies which can be used to identify potential energy savings
- to continuously monitor the sectors „new buildings" or „rented houses" by use of EPCs since EPCs are issued for practically all buildings of these sectors and therefore can be used as a census

– to use EPC data as an opportunity to answer specific questions by monitoring, e.g.
  - portfolio analysis (e.g. of the building stock of housing companies or municipalities),
  - monitoring of support programmes,
  - improvement of energy balance calculation methods (for example by comparing the calculated and the measured consumption, different calculation methods etc.),
  - quality control of energy certificates.

Summing up, DATAMINE is:
– a concept providing data to be collected in order to monitor the energy performance of the existing building stock,
– a convention ("language") for exchanging information between countries about a building's energy-related features,
– a scheme for evaluation of national databases and for international comparison.

A full set of reports:
– Harmonised Data Structure (Excel table with data field names and descriptions)
– Executive Summary
– Final Report:
  - monitoring targets
  - collection method types
  - description of the data structure
  - description and results from 12 model projects
  - cross-country comparison
  - recommendations.
– Abstracts for each of the 12 Model Projects (in English and in national languages).

All information listed above is available at the DATAMINE website at www.meteo.noa.gr/datamine.

A full list of the participants in the DATAMINE project is available at the project web site and in the DATAMINE presentation material placed at the Mayetic project centre under Core Theme 1 for the Berlin CA2 meeting.

Knowledge gained in the DATAMINE project will be utilised and further elaborated in a new IEE project TABULA (Typology Approach for Building Stock Energy Assessment). TABULA will establish the structure for identifying harmonised building typologies in European countries. Among the results will be a set of representative buildings (energy performance and frequencies), a model for imaging the whole building stock and thus impact of measures, and simulation of different modernisation strategies on a broad scale.

TABULA will comprise the following steps:
– collect information about already existing Building Typologies (structure, application fields);
– find a common structure for national Building Typologies;
– fill the structure with data of typical buildings and heating systems for each participating country;
– find possible ways for determining and continuously updating frequencies of building types and system types;
– apply national building typologies in the area of energy advice and building stock modelling;
– expand the approach to other countries not directly involved in the TABULA project.

In the long run it is the intention that the identified building typologies will become a publicly available data source for forecasting and evaluating the energy savings and the carbon dioxide emission reduction for each European country or region. The main deliverable from TABULA will be:

– National Building Typology Brochures for each country with example of buildings, overview of the energy performance of typical buildings and the potential energy saving by refurbishment measures (target group: national experts);
– Building Typology web-tool including datasets of buildings and system types from all countries plus measures and online standardised calculations (a simple energy performance calculation method).

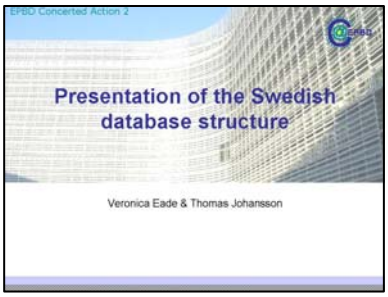
A full list of the participants in the TABULA project is available at the project web site at www.building-typology.eu and in the DATAMINE presentation material placed at the Mayetic project centre under Core Theme 1 for the Berlin CA2 meeting.

4.1.1 Questions to Tobias Loga

– Are the data structure and the data publicly available?
  The data structure, the results of the analyses and the tools can be downloaded from the DATAMINE website. For reasons of data privacy, the database itself is not publicly available.

– Is one of the outcomes recommendations on databases?
  If a central EPC data collection does not yet exist in a country or is only restricted to a small amount of information, the DATAMINE data structure can be used as a pattern in order to identify the relevant energy performance indicators necessary for future monitoring activities and which should be collected as a minimum.
  For already existing national EPC databases, it is recommended to programme an export feature in order to transform the datasets to the DATAMINE data structure. This will later offer the opportunity to easily perform comparisons with other countries.
  Generally it is recommended to not only concentrate on data from EPCs since they only cover a certain part of the building stock. Other sources of building data should also be integrated in the database.

4.2 Sweden, Veronica Eade & Thomas Johansson



In Sweden, the database structure has been the same from day one of the EPC scheme, but even so it has evolved over time. Boverket owns the database and the major part of the input comes from the experts and municipalities who perform certification in the EPC scheme. However, data are combined with information from other sources i.e. the national building stock register, the metrological office and SWEDAC (information about accredited bodies to issue EPC). Besides Boverket, information in the database can be used for all kinds of statistical analyses by different bodies like municipalities, the Swedish Energy Agency (via a statistics system), building owners (in the future) and for research (direct access, but normally via the statistics utility) purposes. The data flow is illustrated in figure 12.

Data can be transported to and from the database using three different tools, namely using an XML web interface; using the Swedish EA tool for certification (fi2); or using the Microsoft InfoPath. The benefit of using InfoPath is that it works for the energy expert both on- and off-line.

The database holds all information about the building that enables an energy performance calculation to be performed. In addition, information about the building owner, the building id, and the use of the building is being collected. As something special, the assessor also collects information about the indoor climate as a radon gas measurement (as a preparation for the future) and a ventilation check. Last, but not least, all suggested profitable energy saving measures are being collected.

The Swedish EPC scheme does not require an inspection of the building and information about inspection or not is also gathered in the database. Until now only 40 % of the certified buildings have been physically inspected.

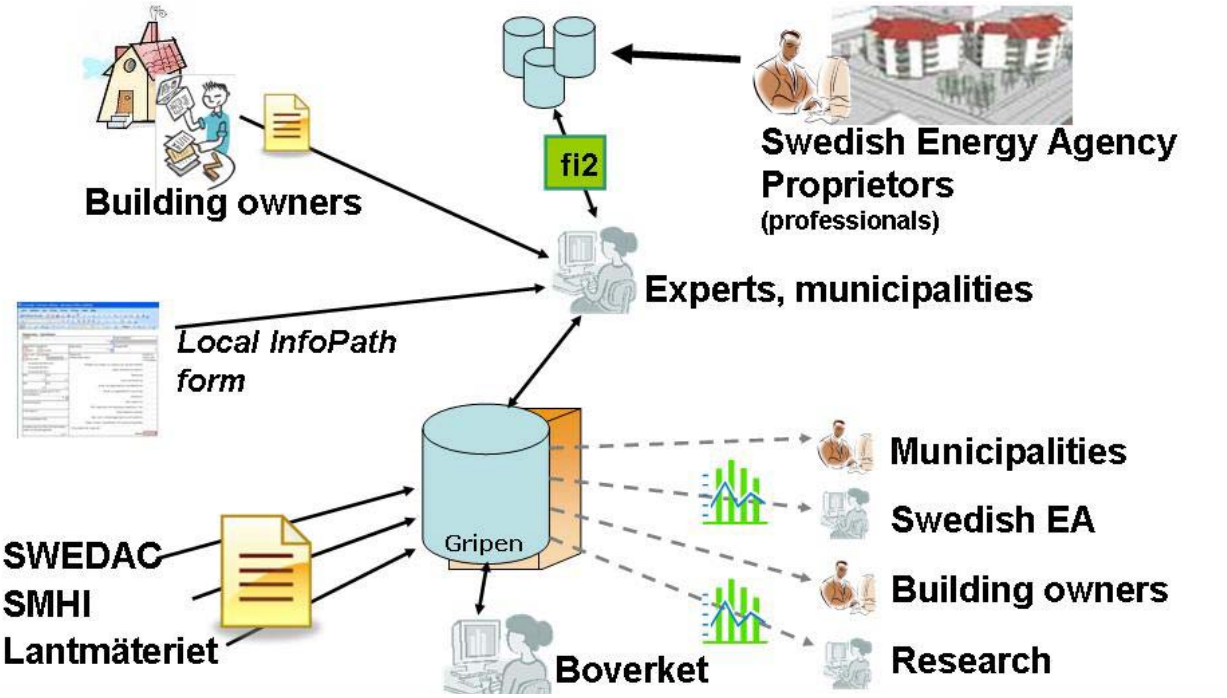


Figure 12. Data flows in the Swedish EPC scheme.

Sweden has not used data for other purposes yet. Sweden does however see many opportunities and research projects call for possibilities to exploit information in the database. Boverket is constantly asked why this information is being collected, and it is nice to give other reasons than a requirement from EU.

4.2.1 Questions to Veronica Eade and Thomas Johansson

– Does the Swedish EPC scheme rely on calculations?
  Energy performance is calculated in the standard tools, but measured data is also stored.

– Where does input data come from?
  Input data come from the assessor and calculation of the EPC based on this input.

4.3 Denmark, Kim Wittchen

The Danish EPC scheme is the second generation and the information on EPC databases can thus be divided in two parts. Historically the EPC scheme was divided into two separate schemes.

One scheme was for small buildings and owner occupied flats, which should be certified when they were sold. This certification was based on calculated EP, but two companies supplied the certification tool and data was stored in two different databases, one for each tool. The other scheme was for large buildings (+1500 m²) and these buildings should be certified every year. This certification scheme was based on operational (measured) rating and had a separate database as the structure differed completely from the calculation approach.

The current EPC scheme (implemented after the EPBD) also has two levels of buildings, small and large (+1000 m²), but both levels are certified by means of calculated rating. The data needed to perform an energy performance calculation is thus similar for both building categories and data are stored in the same central database.

Calculation of the energy performance certificate is made with the same calculation engine as is being used for calculation of energy performance of new buildings when applying for a permit to build. When it is completed, certification of a new building is thus a quality check in terms of energy performance. The EPC assessor uses the same calculation engine, but with a different user interface. Any company can make its own user interface for the calculation engine and so tailor it to be used in the certification scheme.

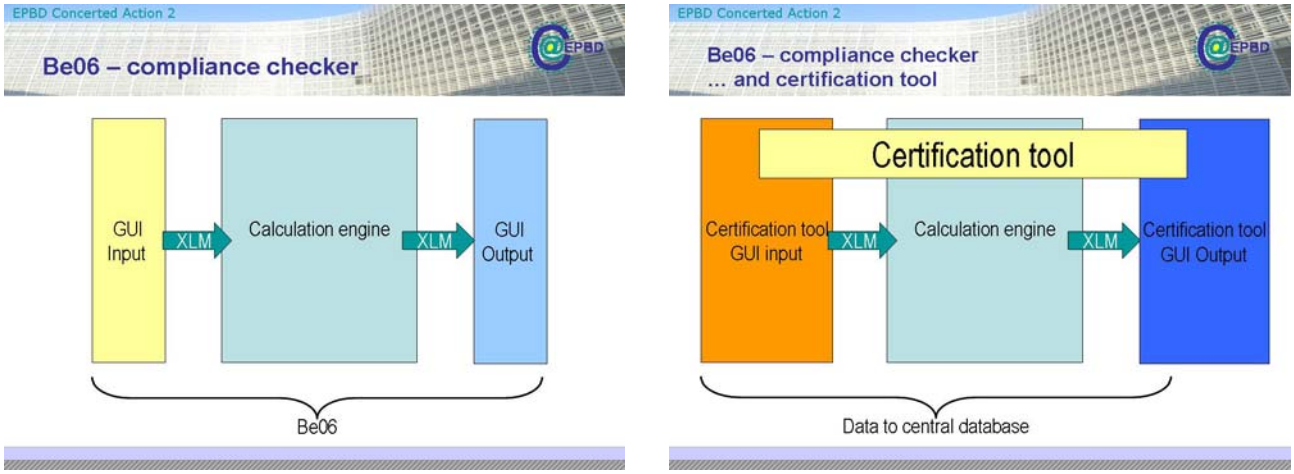


Figure 13. The Danish compliance checker, Be06, and the Danish EPC tools use the same calculation engine. Most data is thus generated in the calculation engine and well defined.

The Danish Energy Agency owns the database, and daily operation is maintained by a secretariat. The secretariat also carries out quality control and suggests thresholds for allowing data into the database.

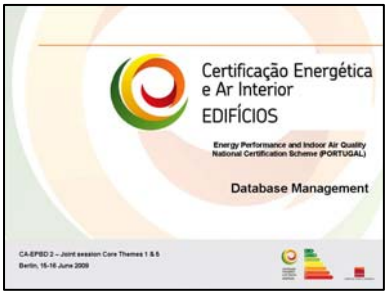
Data in the database are grouped into three general categories:
- General data to identify building and the expert,
- Data needed to perform a calculation of the energy performance,
- Results of certification and energy savings proposals.

During 2009 energy performance certificates showing all their information will be made publicly available on the Internet from a public information server at www.ois.dk.

The data is organised in a standard SQL database and the Danish Building Research Institute have direct access to the database. Analyses are therefore an ongoing activity to facilitate:
– Identification of energy savings potential for different building types;
– Analyses of potential cost of the savings;
– Support to political initiatives for promoting energy savings.

Additional benefits of these analyses have been identification of problems with the database like missing/undefined parameters; outliers; wrong value properties (numbers/text); parameters stored in wrong columns (consultants errors); etc. Furthermore it has been possible to identify parameters that would become a good supplement to the database and pave the way for new analyses.

4.4 Portugal, Paulo Santos



The EPC always refers to calculated values coming from the asset rating analysis made by the experts. For public buildings data about the energy consumption is collected additionally. There is one nation wide database, under the responsibility of ADENE, who is the managing authority. Input to the database is made directly by the expert as part of the process of issuing an EPC. Inputs and issuing the EPC is one single on-line process like a typing machine. An XML interface has been established and can transfer data directly to the database. Input can also be made while offline and be uploaded later. Most of the relevant information collected during inspection is stored in the database (up to 200 entries).

The certificate itself is issued from a central entity and returned to the assessor as a PDF file which can be printed and delivered to the customer.

The cost of setting up the database and the management system was about 1.2 M€. The database is structured in a way that improvements can be implemented without additional costs. Information is distributed on five servers each with different purposes.

There are many advantages attached to having all information available in one central register such as:
– Easy overview of the number of experts;
– Direct access to the number of certificates issued;
– Geographic distribution of number of certificates and certificate label;
– Primary energy consumption;
– Estimation of CO_2 emissions;
– Energy demand distributed on different types of use;
– Average U-values for constructions;
– Types of systems used;
– Solar collectors installed;
– Ventilation rates;
– Lighting (non-residential buildings only),
– etc…

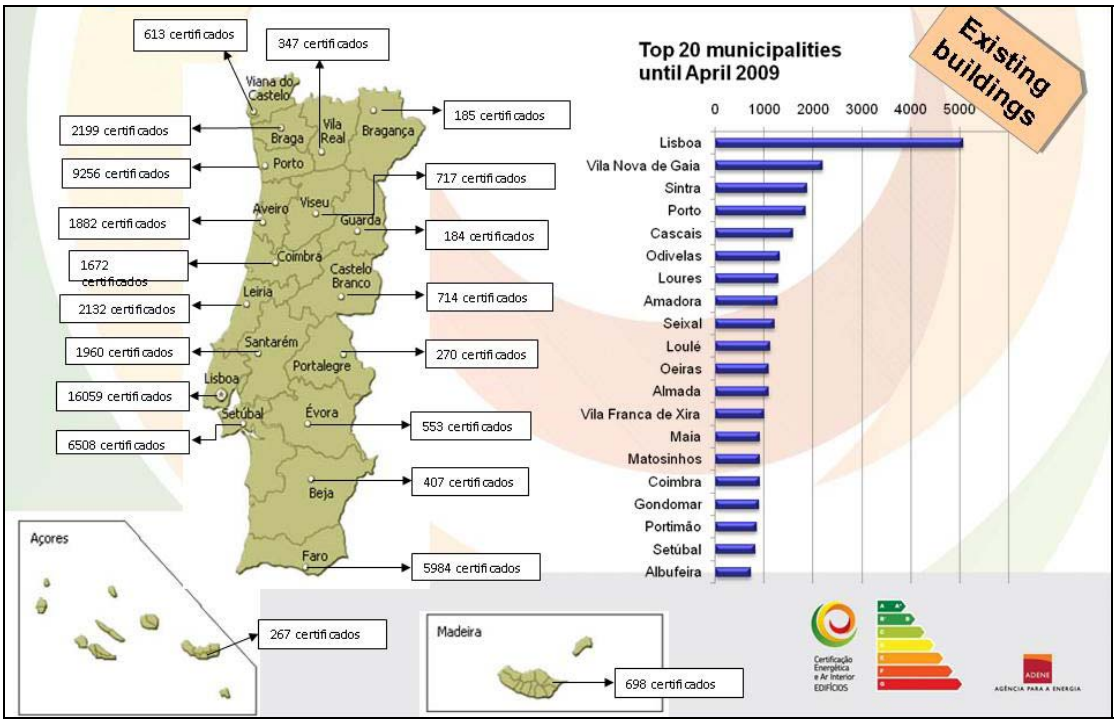


Figure 14. Example: Geographical distribution of issued EPC certificates in the Portuguese regions.

Quality check of data is made in 1 to 4 % of all issued EPCs (after input). Which certificates to check is chosen by random, but non-random selection is also used, e.g. in case of complaints from the customers.

Some improvements have already been made and more are planned. At the beginning of the EPC scheme there were weekly manual counts of the content in the database. Since November 2008 pre-defined reports taken from the ADENE intranet have been available. And during 2009 it is planned to enable search in the database with a business intelligence tool (estimated cost between 30 and 50 k€). The use of a business intelligence tool for customised search of information is highly recommended as it facilitates pre-defined reports; dash boards; score cards; ad hoc queries in pivot tables; etc.

Public access to searches in the database has been made possible, but only limited information like the label is shown. The key to search for an EPC can be EPC number; the date of issue; the address; name of the municipality; name of the region; or the name of the expert.

Difficulties have been identified and managed in the Portuguese EPC scheme:
– The balance between detail of data and time needed to input the information;
– Lack of uniform or pre-defined input of information in some fields, e.g. energy saving measures;
– Not having database reporting tools from the very beginning.

Improvements have thus been identified as:
– Recommendations that should be given as pre-defined options (to allow monitoring),
– Implementation of a business intelligence tool to ease analyses of data in the database,
– Integration in and communication with other databases that is essential to facilitate more analyses.

Recommendations:

– Be prepared to respond to all kinds of requests from the public, research organisations, public authorities, etc. (ad hoc reporting),
– Explore the database on a regular basis to produce public information and validate data quality,
– Allow simple public search of the database for valid EPCs.

4.4.1 Questions to Paulo Santos

– What do you check in the1-4 % of the certificates that are selected for quality checks?
  This is the normal EPC quality control of the experts. Certificates are selected randomly or based on complaints. We do also perform visual checks of input data. There are no penalties for weak quality at the moment.

– Data security, is that a problem?
  We are very concerned about that issue. Changes have to be made if anybody is to be allowed to see all EPC data. We have tried to make information as anonymous as possible.

4.5 Germany, Thomas Kwapic

There is no official database in Germany for collecting information about data from the EPC scheme. All information about databases given in this presentation is related to the voluntary database created by dena.

The reason for not having a central database for the EPC scheme in Germany is that the German government aims at reducing the costs of implementing the EPBD, reducing bureaucracy and the fact that there is no obligation in the EPBD for collecting EPC information in a database.

In this context, dena has introduced two private, voluntary databases. One database is for energy performance certificates and one is for EPC experts. Dena's motivation for doing so is to give building owners a searchable database of issuers in their neighbourhood and to help issuers to promote their service. And last, but not least, to be able to collect information about the building stock in Germany and evaluate the effects of national activities (promotion, market campaigns, and minimum requirements) and in the long run to generate a technical basis for the introduction of a quality assurance system for EPC certificates.

The dena database contains two types of certificates. All regular certificates can be submitted on a voluntary basis. All certificates with dena´s seal of quality have to be submitted to dena´s database.

The seal of quality is the result of a private initiative launched by dena. It contains a definition of qualification standards for the assessors and gives uniform standards for the issuing process. Certificates must be issued according to these standards and comply with the requirements of the national legislation.

Certificates submitted to the database on a **voluntary basis** covers residential and non-residential buildings and can be either calculated or consumption based certificates. All data of the certificate are transmitted in an anonymous way to the database. This also mean that no Information on the responsible expert is being stored and nor is the address of the building (postal code only).

Data are transmitted automatically using calculation software which communicates with dena´s software which in turn transmits data to dena´s database (via the Internet). Data include upgrading of recommendations that are fully captured. However, the recommendations are stored as free text entries (no tick boxes or default selections) and therefore cannot be used for statistical purposes.

In principle a certificate could be created using captured data, but dena has not realised that yet.

Certificates with **dena´s seal of quality** covers residential buildings only and are solely based on calculated rating. Data have to be transmitted using calculation software that communicates with dena´s software that transmits data to dena´s database. As for the voluntary scheme, all data are transmitted (partly in an anonymous way) followed by an automatic check (i.e. check on plausibility) and certificates are rejected if crucial data are missing or not plausible. In this case data can be related to an expert (to allow for random samples).

The expert must give recommendations on improvements of the heating technology and building envelope. If the expert only comments on the envelope or the technology, the certificate cannot be issued. Additionally there must be documentation showing refurbishment recommendations in two levels of effort. The first level is made up of those recommendations that are single measures that are simple to realise. The second level consists of those recommendations that should be implemented in conjunction with a comprehensive refurbishment of the building.

The automatic quality check of data is defined by dena as ranges of possible data. Additionally there are automated checks of certain criteria and comparison with plausible parameter. Automated checks are e.g. activated for year of construction, year of installation engineering, surface-volume ratio (A/V), areas of building components, and U-values on specific components.

Until now data in dena's database have been used for research and analyses and will also be used in the new IEE project TABULA (see summary of the DATAMINE presentation earlier in this session). The objective of these analyses will be to collect information about status of the existing building stock, identify weak points (in relation to building type and age) etc.

Additional analyses are possible and in the process of being developed, like statistical analyses on when EPC´s are being issued (and for which type of EPC), which recommendation are the most commonly used by the experts.

There have been difficulties in establishing and running the database and some of these are related to the communication into the database. One of the main issues has been to find duplicates in the database. One reason for this problem is that the issuer is allowed to upload an EPC more than once (in case of changes). Other reasons are related to problems with saving the document or technical problems during submission process. The great challenge is to identify and eliminate the duplicates to create a valid basis for the analysis of the database.

Dealing with duplicates has been a task in itself. Some of the methods being used in this process can be summarised as:

– Has a certificate for a building with exactly the same address already been registered? (note: Address is saved in an anonymous way via hash-key)
– Cross check of certain criteria (very unlikely to be exactly the same for different buildings) i.e.: (Primary) energy demand, U-values, Volume, Areas.

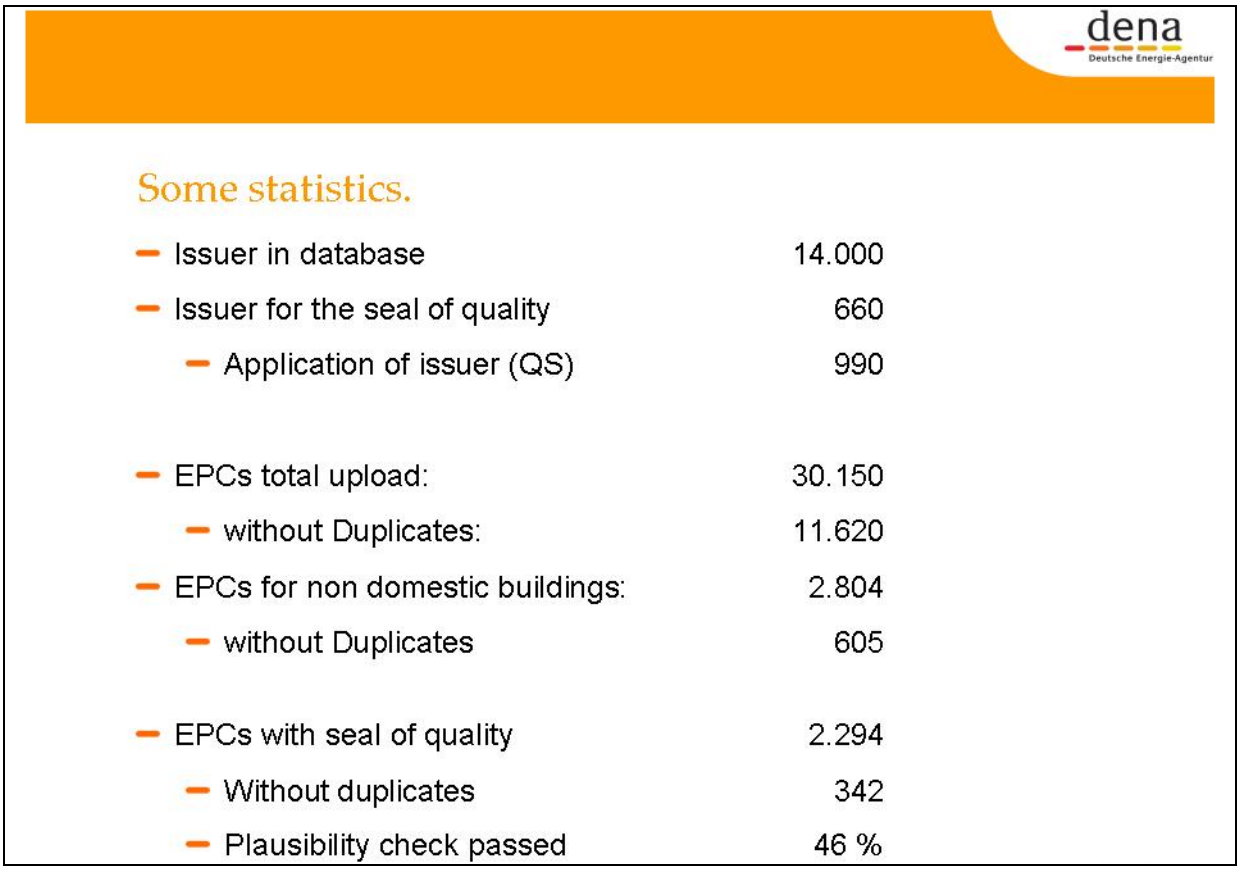


Figure 15. Statistics on presence of duplicates in the dena database.

Potential improvements of the German situation have been identified. In the current market approximately 20 software providers are active. It is the intention to implement a uniform XML standard to ease the communication between these software providers and dena's database. It is also the objective to generate convertible datasets for energy performance certificates (regardless of the software provider).

4.5.1 Questions to Thomas Kwapic

– Is there no mandatory DB in Germany?
  No, we will have an amendment to the law soon, and no central database is mentioned in this. The next change of the law will take place in 3 years.

– What is the reason for having so many duplicates in the database?
  We don't want to have high obstacles to enter information into the database and this causes the problem. We have developed an algorithm to remove duplicates and it is constantly being improved.

5 Main discussion and outcomes

With the information given in the four presentations the delegates could gather in smaller groups and discuss different subjects relating to database management. The chair of the second part suggested five topics, and the delegates had the opportunity to suggest additional topics. Every delegate voted for the topic he or she was most interested in. All participants who decided on the same topic formed a discussion group. In this way six discussion groups were established. Three groups discussed the quality of data, one group the integration of the EPC database with other systems

and tools, another group discussed the different possibilities for using the data stored in databases and the last group public discussed access to the database management system.

All discussion groups were asked to prepare a short presentation on a flip chart or to answer some questions in a brief interview about the topic. In the next paragraphs, the outcome of the discussion groups on the four topics is given.

5.1 Integration between different systems

Within the presentations it was mentioned that on the one hand the EPC databases should be integrated in the whole EPC issuing process but also with other related initiatives. On the other hand MS faced difficulties when linking different already self-contained structures. One discussion group started working on this topic by brainstorming on the question how the database should be integrated with other initiatives.

On their flip chart picture they placed the database as the foundation of the whole system. At the top the implementation of refurbishments was placed. On one hand data stored in a database should increase the number of refurbishments by a strong link with quality assessment, energy consultation and information campaigns. A transfer standard (e.g. xml) and transfer file could allow using data collected for different purposes (e.g. issuing of the EPC, energy consultation, adaptation of recommendations, etc.). Quality assurance plays a major role within the whole process surrounding the EPC schemes and should therefore be firmly linked to the database. Also the energy consultation, the accreditation system and the monitoring must have links to the quality system. This results in a qualified certification of buildings and the implementation of refurbishments. On the other hand the information from implemented refurbishments is fed back into the database and can be used to design strategies for information campaigns or to evaluate the implementation of the EPBD and to improve it.

5.2 Quality of data

Three groups decided to discuss the quality of data. Not every group prepared a presentation or flip chart but all shared their ideas and problems related to the quality of data. The first group gave a presentation on a flip chart about different types of quality checks.

The quality of the data from the energy performance certificates can be checked at different stages of the whole process of certification. Prior to this process, a plausibility check needs to be integrate in the database e.g. by defining ranges and limits. There should also be verification procedures integrated in software tools used for certification. During the certification process the energy advisers themselves should check the plausibility of the data they use. A software tool that allows cross checks could be very useful.

After the certification the quality of data could be checked by self control or by accredited bodies, similar to ISO 9000 and by accomplishing manual verification. If a ‘wrong’ certificate is identified in this process, then this certificate has to undergo a full quality check and has to be calculated again by the experts in situ.

In the database management system the duplicity of data has to be checked automatically and handled in an intelligent way.

During the whole certification process, prior, during and after the certification one has to evaluate the consistency of the database continuously.

The other two groups gave a short interview to the chair of the meeting and said:

– The possibilities depend on the use of a central system. A quality check is easier if there is only one database. However it is difficult to correctly integrate data in one central database if data related to one property comes from different sources. When a database is directly linked with the software tools used for issuing the EPC, it is much easier to check the quality of the data.

– In non-residential buildings it is difficult to check the quality of data, because of the complexity of the building which results in a higher number of data. Checks on some main parameters only may be a possibility.

– The method of collecting data is very important. Type and amount of data differ depending on the EPC method in use (calculated or measured energy performance). Also the skill of the assessors is important.

5.3 Public access

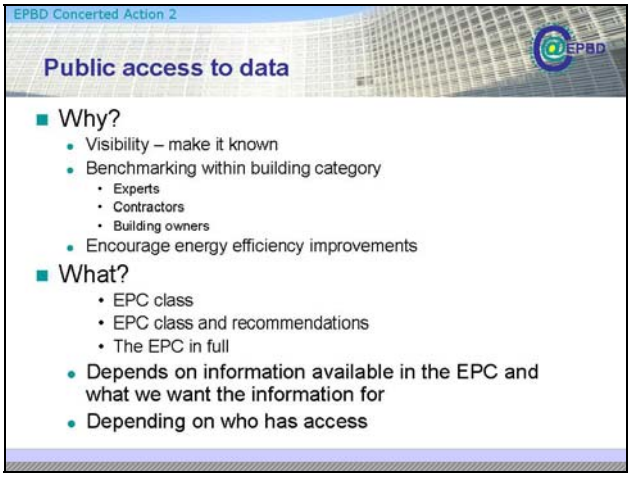
The participants from Scandinavia discussed public access of data. They prepared a Power point presentation and presented it to the other participants.

They discussed why public access to the databases should be provided and what obstacles have to be faced. One argument for providing public access is that the visibility of the database and the EPC in general increases. With public access it is possible to use the energy certificates as a benchmarking system as well as services for suppliers of building materials for them to analyse their marked potential in providing energy saving measures. Public access could also help the consumers to better understand their own certificate.

The minimum information provided by a public database is the EPC rating. But also the recommendations on the certificate and services aspects are very important. The differences between calculated and measured data have to be clear in the public database.

One important issue is how MS can realise public access to their database. The best possibilities for the discussion group seemed to be internet access to the database or development of internet tools to make some parameters visible for the public.

One negative aspect of public access to the database is data privacy when personal and private information is affected. Another one is that consumers will come up with a lot of questions once they have access to the database. They also have the formal rights to complain. This requires some kind of support for the database.



5.4 Use of data

Data stored in a database can be used for different purposes, which is a great advantage for many parties. A more intensive use of the huge amount of data is desired by the whole discussion group. The data can for example be linked to a geographical information system for analysis of the different regions and climate conditions. It could be possible to arrange competition between regions or produce better information campaigns.

The discussion group exchanged experiences and discussed the use of data to address and involve different target groups.

For example, the government could base a national energy strategy on the results of analyses of the database. The energy saving potential of the building stock could be analysed and activities could focus accordingly on those regions with buildings having the highest energy saving potential. Also MS could develop more precise climate change projects and inform the EU on their outcomes when using data from EPC databases. A common obstacle for performing these analyses is the gap between the information and data needed and what information is available.

Subsidy providers could better adjust their information campaigns and marketing tools towards their target groups. Certain analyses may allow them to focus on the target groups who need the subsidy urgently and who have the highest potential.

Technology solution providers, for example manufacturers of insulation products, could analyse how big their market is and who belongs to their target group. In case the database is available to them, they could acquire new business opportunities and market potential.

Public access to the database would in turn increase the general awareness. Consumers could compare their certificate with their neighbours. Another option is to assemble public reports based on the database.

The local government could use the data for designing their local campaigns. Evaluation of their strategies and plans to improve the building stock and growing of competition could be the most important topics for them.

6 Conclusion of topic

There was general consensus that the presence of central databases to collect information gathered while issuing the energy performance certificates opens a world of possibilities for numerous analyses on all kinds of topics. Data from the EPC cannot only provide information about the num-

ber and quality of issued EPC, but can also be used in estimating the potential energy savings in MS, regions or counties for the benefit of the authorities.

A more widespread deployment of statistical information from the EPC schemes will probably lead to a higher degree of acceptance by the public.

There is still one issue that MS need to address when discussing publication of information from the EPC schemes and that is data confidentiality. It is not a straightforward matter to publish all information collected from EPC and the MS must identify possible sensible information and find ways to make it anonymous before publication.

7 Future directions

Many interesting viewpoints were highlighted during the session and it was suggested to continue the topic, but with a slight twist towards possible other uses of EPC data at the next plenary meeting. The following text has been suggested to EACI for a technical session at the next CA2 plenary meeting.

Data collected and codified while issuing EPC can be used for many other purposes such as:
– calculation of potential energy savings on national/regional level;
– proof of EPC market penetration;
– proof to EU of EPBD implementation;
– regional competition on marked penetration and/or implementation of energy savings;
– promotion of top level certified buildings;
– savings;
– prioritisation and targeting of buildings for energy retrofit;
– possible links to incentive schemes;
– collection of building stock knowledge in general;
– etc.

To be able to exploit the data in the most effective ways and thus increase the value of the EPC schemes, it is crucial that all relevant data are collected and codified in compatible formats to enable merging from regional or other different sources. Examples of external data could be:
– district plans,
– number of persons living in the building,
– ownership,
– number of bedrooms,
– etc.

The core theme leaders will identify, evaluate and discuss the possibilities for utilisation of EPC data for other purposes than issuing the certificate and in this way adding value to the EPC scheme. The key protocol challenge of accessibility to data will be discussed as:
– who should have access to data and to which data?
– how can access be granted and/or restricted?
– how can data be activated for other purposes without compromising data confidentiality?

Annex 1 - Questionnaire on database management

Questionnaire on *Database management*

Purpose: This questionnaire collect information about management of databases in the certification process is handled. Your answers will be used as preparation for the Concerted Action 2 joint Core Theme 1 and 5 Technical Session on 16. June 09:00 - 12:30.

There are ONLY 10 questions !

Please return your answers to Kim B. Wittchen (kbw@SBi.dk) no later than April 27, 2009.

Please indicate your name, affiliation, country, and e-mail address

Name:	
Affiliation:	
Country:	
E-mail:	

1. What method do you use for energy performance certification ?

Tick one!

- ☐ Calculated energy performance for all building types and ages.
- ☐ Measured energy performance for all building types and ages.
- ☐ Method depends on the building type and/or age. **Please specify!**

2. How is the database for collecting building energy performance certification data organised ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	*Measured*	
☐	☐	One nation wide database.
☐	☐	Regional databases.
☐	☐	Multible databases, one for each accepted certification tool.
☐	☐	Multible databases, one for each certification method (calculated/measured).
☐	☐	No collection in database. *If you do not collect data in a database, you can skip the rest of the questions.*
☐	☐	Other, **please specify ...**

3. Who is responsible for the database in your country ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	*Measured*	
☐	☐	Central, official authority.
☐	☐	Regional, official authority.
☐	☐	Private company/companies.
☐	☐	Other, **please specify ...**

4. How are data collected in the database ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	*Measured*	
☐	☐	Automatically reported from the accredited certification tools.
☐	☐	Reported directly from the consultant/expert/assessor.
☐	☐	Central secretariat transferring data from the certificates to the database.
☐	☐	Other, **please specify ...**

5. If you collect data in a database, which data do you collect ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	*Measured*	
☐	☐	All information collected during inspection (U-values, areas, efficiencies, etc), energy label (e.g. A to G) and recommendations.
☐	☐	Only energy performance, label and recommendations.
☐	☐	Energy performance.
☐	☐	Energy label.
☐	☐	Only building id.
☐	☐	Other, **please specify ...**

6. Quality check on data (tick all that apply)

Tick all that apply !

Calculated	Measured	
☐	☐	Probability check on entry, e.g. acceptable value range for differnt parametres.
☐	☐	Field compliance check, e.g. no text in numerical fields.
☐	☐	Check that all requested data is available before entering the database.
☐	☐	All data from a certificate is rejected if crusial data is missing.
☐	☐	Generation of statistical information and cleaning of data and identification of out-of-range data after input.
☐	☐	Manual cleaning of data and identification of out-of-range data after input.
☐	☐	Other, **please specify ...**

7. Have you tried to generate a full certificate from information in the database ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	Measured	
☐	☐	Yes, with success.
☐	☐	Yes, but with limited success.
☐	☐	No.

Comments:

8. Have you tried to use information in the database for other purposes, e.g. calculation of national saving potentials or market penetration ?

Tick one in each collumn if you have both calculated and measured energy performance!

Calculated	Measured	
☐	☐	No.
☐	☐	Yes! **Please explain ...**

9. Do you meet difficulties in database management and what sort of problems do you have?

Describe:

10. How would you improve your database management system in the future?

Describe:

END OF QUESTIONNAIRE !

Please send you answers to kbw@SBi.dk no later than April 27.