Introduction to Panel 8 Monitoring and evaluation: building confidence and enhancing practices

Panel leader: **Kirsten Dyhr-Mikkelsen**
Ea Energy Analyses
Denmark
kdm@eaea.dk

Panel leader: **Jean-Sébastien Broc**
B R & C
France
brocjs@yahoo.fr

Panel leader: **Bernard Bourges**
Ecole des Mines de Nantes
France
bernard.bourges@mines-nantes.fr

Energy efficiency investments are growing: what about evaluation?

In 2013, the IEA placed energy efficiency high on the agenda declaring it was "the world's first fuel"[1], but also highlighting it is a "hidden fuel". The IEA Executive Director, Maria van der Hoeven, indeed said that energy efficiency needs a face[2], in order to become visible and more convincing for decision makers and investors. She added that this face should be built on evidence, in other words data and evaluation results.

Recent years have seen an increasing recognition of investments in energy efficiency. And the National Energy Efficiency Action Plans reported by the EU Member States in 2014 clearly showed a significant progress in structuring energy efficiency strategies, in analysing barriers (sometimes, even energy efficiency markets) and in presenting their policies and programmes.

Even if there are no official records or statistics about it, it is very likely that more evaluation work is being carried out currently than just a few years ago. However, are the current evaluation efforts in line with the overall ambitions of scaling up energy efficiency investments? And are evaluation methods and practices able to address the issues raised by the upgraded policies and programmes or by the emerging new generation of measures or data? These are general and transversal questions that this panel intends to raise during the Summer Study.

This panel introduction presents more specifically the main issues and questions raised in the papers of the panel.

1. http://www.iea.org/newsroomandevents/pressreleases/2013/october/name,43788,en.html; http://www.iea.org/Textbase/nptable/2013/EEMR2013_f3_4.pdf

2. Opening speech of IEPPEC2014 in Berlin, September 2014.

Different levels of policies mean different evaluation approaches

Evaluating energy efficiency policies is a broad topic, as shown by the papers of this panel. And this topic is even getting wider, because policy makers intend to improve the design process, and because the scope of these policies is also expanding.

The scope of Vringer et al. (8-216-15) concerns national energy efficiency policies implemented on the same sectoral target in a given country. Their work is an independent assessment of the different Dutch policies targeting energy efficiency in buildings. The quantitative side of their evaluation is limited to the global impact of all policies together. Although a significant decline of energy consumption is expected (by 82 PJ to 521 PJ, between 2008 and 2020), this pace is not fast enough to achieve the target for 2020. Lack of detailed data and high degree of interactions between the various policy instruments do not enable to establish the effect of each isolated policy. Nevertheless, the qualitative study of policies applied to different segments of the building sector is particularly relevant. It emphasises significant improvements to be brought to the respective instruments and required consistency of energy taxes. Both quantitative and qualitative conclusions of this work are likely to apply *mutatis mutandis* to most European countries.

The scope of Mulholland et al. (8-404-15) deals with policies developed for another sector (transportation) in a different country (Ireland). It illustrates a totally different evaluation approach, based on coupling of two bottom-up techno-economic models to evaluate how a suite of policy-measures can achieve a top-down target. The two models used are i) a full energy systems optimisation model and ii) a car stock simulation model. Their combination makes it possible to determine a more feasible policy pathway after an important bottleneck has been identified.

The scope of Fleiter et al. (8-468-15) is about a European policy implemented in the 28 EU Member States. They performed an ex-ante and bottom-up assessment of the EcoDesign Directive, to determine the potential savings until 2020 and 2030 for the industry, tertiary and residential sectors. This contribution is very valuable for the energy efficiency community. The effects of the EcoDesign Directive are sometimes overlooked (compared to energy efficiency in buildings for example), while the corresponding implementing measures are key pieces for the EU energy efficiency policy framework. Conducting an EU-wide evaluation is a big challenge. And this ex-ante assessment also raises key issues about the actual implementation of these measures. Some of these issues are addressed in the papers presented in the next section.

From the lab to the market: compliance is the key!

One of the key issues related to regulatory measures is compliance. When dealing with minimum energy performance requirements for appliances, this can be looked especially at two stages: the laboratory test of performances and performance in the market.

Ohno et al. (8-346-15) raise the issues of testing competency. Testing laboratories can verify their measurement uncertainties and gain confidence in their measurements only by inter-laboratory comparisons. As a consequence, IEA 4E Annex SSL carried out a large scale international comparison among 110 laboratories to develop reliable and repeatable testing method for LED. The comparison led to a clarification of the levels of agreement of measurements of photometric, electrical, and colorimetric quantities of LEDs, which will help governments and manufacturers around the world to learn the uncertainties of the measurements of SSL products and ensure that new products sold to consumers and companies are of high quality and meet the claimed performance. The results also revealed some cases of extremely large errors (e.g., 25 % error in luminous flux). Identification of these large deviations for a few laboratories demonstrates the importance of proficiency testing, as these laboratories would not have been aware of their particular problems without having participated in this inter-laboratory comparison.

Blomqvist and Fjordbak Larsen (8-266-15) and Jones (8-076-15) propose two distinct approaches and points of view about market surveillance. The aim of the 'Effect' project, presented by Blomqvist and Fjordbak Larsen, was to assess the achieved benefits of market surveillance carried out 2011–2013 and to estimate the potential energy loss due to non-compliant products on the Nordic market (Iceland, Finland, Norway, Sweden, and Denmark). A refined calculation approach was developed to estimate how big a deviation in annual consumption the non-compliant appliances introduce, compared to a standard purchase. Their conclusion is that approximately €28 million can be saved in the Nordic countries alone through collaborative market surveillance, through an investment of around €2.1 million, equal to a return on investment of 13. Scaled-up to EU level this corresponds to €255 million. The overall energy non-compliance rate was 6.3 % at a typical level of 35 kWh/year for the dominant non-compliant appliances, but with a wide spread from 1.6 to 700 kWh/year in non-compliance.

Jones et al. explore how Civil Society Organisations (CSO) can contribute to market surveillance and identify non-compliances to Ecodesign and Energy Labelling Directives. A European network of CSO's has been established within the framework of the MarketWatch project. A first round retail observations of 67,000 products within 7 categories was implemented; high non-compliance rates (higher than 50 %) are observed for some criteria. The paper discusses how such observation by CSO's can stimulate better legislation and better application of existing legislations, particularly in those countries where Market Surveillance Authorities face legal barriers.

What do information and audits deliver?

Complementary to regulatory measures, information measures are also an important and very common part of energy efficiency strategies. But they are among the most challenging to evaluate, especially in terms of impacts, as they are often one component of a broader set of policy instruments targeting the same actions.

Schleich et al. (8-250-15) used a unique opportunity to compare the adoption rates of energy efficiency actions of large samples of SMEs (Small and Medium-sized Enterprises) in Germany, one sample with SMEs having taken part in an energy audit programme and one sample of non-participant SMEs. This comparison offers a valuable experience feedback about both, the statistical method used and the findings on the effects of the audits. Their results reinforce the conclusions that energy audits help overcome more organizational barriers (like intra-company priority setting or lack of strategic importance) than the lack of information about technology availability.

Loch et al. (8-068-15) present a complementary experience feedback about the evaluation of a regional German energy audit schemes for households. They used a more classical approach of surveys among participants. Their results show a difference in the forms of the effects of audits for households, compared to what Schleich et al. observed for audits for SMEs. For households, the main effect seems to be improved technical specifications of the energy efficiency projects, resulting in the implementation of different types or more efficient actions. Another interesting finding is that the age distribution of participants is similar to the one of the population of home owners. In particular, more than 20 % of the surveyed participants were older than 70 years. This result contradicts commonly held beliefs that older people would not consider home improvements.

Barthel et al. (8-115-15) try to evaluate the impacts of another type of information dissemination regarding energy efficiency. An original approach of a web-based information tool about energy efficient products (Euro-Topten) is reviewed in order to analyse how purchasing energy efficient products is actually stimulated. Results are based on two on-line surveys conducted among users of the web-sites. Information proposed is considered as highly relevant by users, both consumers and professionals. Suggestions for further improvement of the sites were made. Moreover, avoided energy consumption has been estimated showing that consumer information should have its place in the energy efficiency policy mix.

While the three papers above deal with information and behaviours related to buying or installing or retrofitting an equipment or a building, the next section deals with behaviours related to consuming energy.

Behaviours & evaluation

As for information measures, approaches aiming at changing behaviours of energy consumption are very challenging to evaluate, as the causality between the activities of the programmes and changes in energy consumption is often very difficult to establish. However, the complexity of the issue should not prevent this kind of approaches first to be developed, and then to be evaluated. At the opposite, this is clearly a field where research is definitely needed. And the three papers presented below offer contributions in this sense.

Morgenstern et al. (8-182-15) investigate how energy audits can help to assess the energy savings potential of given behavioural actions in complex non-domestic buildings such as hospitals. As sub-metering is generally lacking, energy audits are found to be a workable tool for the identification of potential savings through behaviour changes of healthcare staff, aiming to eliminate redundant energy use from standard procedures. An interesting approach of uncertainties for data inputs – a crucial point to evaluate the potential – is proposed, not restricted to hospitals. For specific hospital processes, the authors emphasise the need for interdisciplinary team work associating clinicians and technicians to energy efficiency professionals. Limitations of the technical potential for behaviour change are discussed in the context of high-pressure environment of hospitals (care and cost constraints).

Mourik et al. (8-393-15) present a stimulating discussion paper. They highlight that approaches to behaviour changes are evolving to take into account experience feedback and research results. This put into question the way the success of a programme has to be defined, and then monitored and evaluated. Their proposal is to look for more participatory approaches and for combining or connecting more impact and process evaluations.

Stieß et al. (8-039-15) investigate at an original approach for demand-side management through comparative feedback. In a research project, power efficiency classes for individual electric end-uses were explored as a means to inspire 98 residential customers to become more energy efficient in combination with an information campaign and energy audits. Using focus groups, Stieß et al. assessed the response to this new concept for annual comparative feedback for households.

The challenge of overcoming the self-selection bias when evaluating opt-in programmes is explored by Gaffney et al. (8-474-15). Randomised controlled treatment assignment is not always possible and in those cases other approaches to constructing the counterfactual situation are required. While audit and information programs have existed for decades, evaluation of these programs using advanced consumption data analysis methods is still not fully developed and work remains to assess the effectiveness of various techniques to quantify and mitigate self-selection effects.

Statistical methods: what can energy data tell us?

The development of information technologies and systems makes it possible to collect and process much larger amounts of data. However data alone do not bring by themselves a better understanding of what is observed or monitored. They require new methodologies to process them and to get out of them a true added value. Data mining is indeed sometimes referred to as the new resource for added value, or even the source of a new industrial revolution. The three following papers provide useful insights about what could be expected from these on-going changes, and also about how the analysis of evaluation objectives remain a useful safeguard against data turning into an objective by itself whereas it should remain a means.

Wade and Eyre (8-099-15) perform a review of evaluation practices in order to question the relevance of defining a particular type of method (based on Randomised Control Trials) as the gold standard. They show that what would be ideal in theory may not be relevant in practice, when taking into account usual constraints such as difficulties to define a control group. Moreover, this could also focus the attention on methods for impact evaluations while methods for process evaluations are as important. The review done by Wade and Eyre concludes about stimulating research questions or priorities for evaluation efforts, such as developing the peer-review of evaluations.

Metoyer et al. (8-469-15) present interesting experience feedback on comparing the use of monthly and "real-time" metering to assess the impact of the California Whole House Retrofit programme. This comparison shows some significant differences in the results, emphasising the need for a clear understanding of the advantages and limitations of both approaches. "Real-time" metering data appear promising in investigating effects on end-uses with a strong initial correlation to weather parameters. This may bring a better understanding of the effects on the load curve, which can be of upmost importance when the programmes aim at reducing peak loads.

Majcen et al. (8-080-15) had the opportunity to compare the data from Energy Performance Certificates (EPC) and from energy bills on a large sample of social housing in the Netherlands. They show that EPC's are highly biased, generally overestimating actual energy use, but rather underestimating it for high performance dwellings. Actually, EPC's provide a good ranking of energy performance of different dwellings, but the theoretical EPC energy consumption cannot be a satisfactory base for policy evaluation or individual economic decision support such as payback period. Ways to improve the theoretical energy consumption estimated by EPC are explored through regression analysis of theoretical and actual gas uses with some predictor parameters characterising dwelling and occupant behaviour. A specific data set coupled with a survey among inhabitants is used for this purpose, enabling to identify the most influencing parameters. Importance of behavioural characteristics and their link to dwellings characteristics are highlighted (inhabitants behave differently in differently performing houses), suggesting to define different EPC standard behaviours according to dwelling categories.

Based on detailed measured data, Berry and Whaley (8-134-15) analyse how near-zero energy homes maintain their performance over time. A sample of nine near-zero energy homes from an Australian suburban estate has been monitored over a 4-year period, including the main end-uses and PV generation. Measured data shows no pattern of change in energy performance, except for a relatively small lighting use increase.

Evaluation framework and capacity building

While the evaluation literature is already rich in terms of papers presenting evaluation results or discussing evaluation methodologies, there are much fewer papers about capacity building, and especially about how to build and develop an evaluation system or framework.[3] Therefore, the three papers presented below are very valuable contributions to the energy efficiency community.

Bosseboeuf et al. (8-113-15) provide detailed experience feedback on the dissemination of the use of energy efficiency indicators in Latin American and North Africa countries. Their analyses are based on both quantitative assessments of the achievements in terms of data collection and processing, and qualitative feedback about success factors and difficulties encountered. This shows that these dissemination projects will very likely help further international comparisons and experience sharing. And that they could also give confidence in more countries to enter into similar processes, knowing what should be adapted to their specific context.

Michaelis and McQualter (8-226-15) report a complementary experience feedback about capacity building for policy evaluation. The Australian Government is supporting the introduction of minimum energy performance standards and energy labels in Vietnam between 2012 and 2015. Integral to the support is to develop evaluation capacity within the Vietnamese Ministry of Industry and Trade (MOIT). The paper reports on the results of the capacity building activity and reflects on the challenges, successes, and lessons learned about building evaluation capacity in developing countries. MOIT's interest in building their evaluation capacity at an early stage in the program implementation is likely to increase the success of the standards and labelling program in the future.

Epstein et al. (8-330-15) are on a quest to improve the reputation of energy efficiency projects. In their paper, they discuss the rationale for technically sound and consistent evaluation protocols, focusing on the US systems and recent methodology evolution and enhancements. Incorporation of appropriate, cost-effective evaluation protocols can further energy efficiency projects as they will be viewed as progressively dependable. As a consequence, the efficiency industry will successfully remain a mainstay in the quest to resolve energy and climate challenges.

Hot topics: quality, costs and jobs

Most of the evaluation work about energy efficiency policies and programs is about assessing the impacts in terms of energy or power savings, and/or about assessing how the policies or programs actually work and deliver. An increasing attention has recently been paid to complementary issues, such as in the following papers.

Austria is one of the most advanced energy efficiency service (EES) markets in Europe. DECA, the Austrian Association of Energy Service Providers, representing 25 energy service companies across different branches of origin (utilities, technology suppliers, building service companies, energy consultants etc.) set itself the task to develop a quality assurance system for EES. The main dimensions used to describe the quality of EES are the quality of the service provider, the quality of the service, and the quality of the EES order. For all three dimensions, sets of quality criteria have been identified and evaluated. Leutgöb et al. (8-401-15) present the initial findings of this work and discusses different options for the institutional framework.

Brunetière et al. (8-277-15) explore how the observed scattering of costs within a given category of actions can help a better modelling of the investments in energy efficiency, providing a basis for a comparison between what would happen based on the economic rationality and what actually happened on the market. This has been applied to the case of residential refurbishment actions within the French white certificates scheme. This contribution is very valuable for the energy efficiency community as the papers really entering the details of and discussing actual energy savings costs are still rare.

Finally, Gaffney et al. (8-439-15) address a topic that is high on the political agenda in many countries, namely stimulating employment and economic impacts through investments in energy efficiency and renewable energy. Using the Renewable Energy Efficiency Mapping (REEM) framework developed by Economic Development Research Group and the Regional Economic Models, Inc., (REMI) Policy Insight model, Gaffney et. al. estimate the number of direct, indirect, and induced jobs; the annual and cumulative outcomes for income; gross state revenue; and gross state product of ARRA a subset of a US stimulus package designated for California. The baseline incorporates key economic drivers such as the mix of businesses, population growth and other impacts, such as those arising from the recession. The results indicate that the funding provided through the ARRA programs can be expected to generate US$1.3 billion in increased personal income and US$2 billion in gross state product by 2026. The added employment and economic activity from these program investments are forecast to increase state revenue from taxes and fees by nearly US$243 million.

3. One of the rare papers of this kind is Vine, E.L., Rhee, C.H., Lee, K.D., 2006. Measurement and evaluation of energy efficiency programs: California and South Korea. *Energy Policy*, 31, 1100–1113.