

A process to develop operational bottom-up evaluation methods – from reference guidebooks to a practical culture of evaluation

Jean-Sébastien BROC
Ecole des Mines de Nantes (DSEE)
France
jbroc@emn.fr

Bernard BOURGES
Ecole des Mines de Nantes (DSEE)
France
bernard.bourges@emn.fr

Jérôme ADNOT
Ecole des Mines de Paris (Centre of Energy and Processes)
France
jerome.adnot@ensmp.fr

Keywords

evaluation methodology, bottom-up method, impact evaluation, theory-based evaluation, Energy Services Directive, experience capitalisation, evaluation capacity building

Abstract

Needs for evaluating energy efficiency (EE) activities are increasing, for the accounting of results and for understanding their success/failures. Indeed evaluation results should be used for both reporting past activities and improving future operations. Lack of easy to use methods is pointed out by local stakeholders as a major barrier to evaluation. Another issue is the frequent negative perception of evaluation, experienced as a control and/or a waste of time.

This paper presents a systematic process to develop bottom-up evaluation methods designed to fit to stakeholders needs: directly operational, easy to appropriate, providing useful conclusions to improve operations and to communicate about their results.

Our approach relies on the principle of experience capitalisation and on an organisation with two levels, central and on-field. It aims to create conditions for continuous improvement. Moreover it should insure involved stakeholders do actually take part in and take advantage of the evaluation process.

This methodology handles both impact and process evaluation. For the impacts, focus is on calculations transparency, data quality and reliability of the results. Regarding operation process, main issues are analysing causality between actions and results, and detecting the success and failure factors.

This work was first developed for the evaluation of local operations in France¹. The resulting methodology was tested on two case studies from the Eco Energy Plan, a local EE programme implemented in South-East of France.

Introduction

The usefulness of evaluation may be expressed through its two dimensions, summative (what are the results) and formative (how to improve and/or insure success and cost-effectiveness of operations) [European Commission 1999]. In the particular field of energy efficiency (EE) activities, reference guidebooks as the European [SRCI 2001 pp.8-10] or the IEA [Vreuls 2005 pp.4,8-10] ones highlight needs and reasons for evaluating EE programmes: quantifying and reporting results in a regulatory framework, improving cost-effectiveness, insuring best use of public funds, etc. They also emphasise numerous frameworks which increase the need for evaluation, as the Kyoto Protocol or the EU Directive on Energy End-use and Energy Services [ESD 2006].

Such evaluation needs also appear at a more local scale. Indeed, local EE activities are growing, as national governments require local authorities to involve themselves in local EE policies and as local authorities develop as well their own initiatives. This was confirmed by an inventory of local EE activities in France, which stressed a lack of evaluation practice for these activities [Broc 2005]. Contacts with stakeholders during this inventory brought out that one of the main reasons evaluations

1. within a partnership between ARMINES, the Wuppertal Institute for Climate Environment and Energy, and EDF R&D (Electricité de France)

Table 1. Evaluation objectives and expectations

	summative dimension of evaluation	formative dimension of evaluation
main evaluation objectives	<ul style="list-style-type: none"> - quantifying final results (energy savings, avoided emissions or load reduction) - feeding clear, synthetic and well-documented indicators (for decision making, for communication and for benchmarking) - justifying usefulness and cost-effectiveness of implemented operations, and pointing out local initiatives and/or contributions to national bodies 	<ul style="list-style-type: none"> - detecting success factors, to improve and support easier reproduction of operations - providing experience feedback documented enough for experience sharing and capitalisation - providing elements for critical analysis of results, especially to better understand and interpret indicators - dissemination of the results - developing local skills to rebalance tests of strength between national and local bodies, and enabling a better transparency for debates
resultant expectations	<ul style="list-style-type: none"> - quality of data used, as it is the first factor affecting reliability and therefore usefulness of evaluation results - transparency of the evaluation methods used: good understanding and interpretation of reported results, results discussion and checking, and comparisons with other operations - credibility of reported results, in order to insure recognition of local initiatives by national bodies 	<ul style="list-style-type: none"> - to show and explain to decision-makers (and also to the final public targeted) successful experiences, in order to convince and to familiarise them to these new approaches - to involve stakeholders in evaluation process: for decision making, and for both a stronger involvement of local stakeholders in EE activities (from passive to active strategies), and a change of scale (from pilot operations to widespread diffusion of best practices)

were not performed was that they did not dispose of operational evaluation methods they could easily apply.

This paper presents a methodology set up to develop such methods for ex-post evaluations. First, specifications were deduced from analysing needs and expectations of the stakeholders toward evaluation. Methodology principles and structures were then drawn to tackle these issues. Besides, a literature review and case studies were performed to provide required methodological materials. Ways to address key issues are subsequently detailed about the quantification of impacts and the analysis of programme theory. The resulting systematic process to develop operational evaluation methods is described, and ways to involve stakeholders in evaluation process are discussed. Finally, examples of application on concrete cases are briefly introduced, as well as perspectives of adaptation of this methodology to evaluation requirements for the implementation of ESD.

Defining specifications for evaluation methods at local level

WHAT DO LOCAL STAKEHOLDERS EXPECT/NEED ABOUT EVALUATION?

A previous paper [Broc 2005] presented the results from an inventory of local EE activities in France. During this inventory, many local stakeholders were contacted and asked about their

evaluation practice and needs. Main expectations were then deduced from their answers [Broc 2006a pp.122-123]^{2, 3, 4}.

ADVANTAGES AND DRAWBACKS OF THE LOCAL DIMENSION

Implementing an operation at a local scale can give both advantages or drawbacks specific to the local dimension. These are presented in the table below.

Advantages of implementing operations at a local scale

Specific advantages of local dimension may be grouped in three main lines:

- **nearness:**
 - local stakeholders are closer to the targeted publics, and particularly efficient to reach scattered targets
 - local stakeholders form networks with significant ability of mobilisation
 - local stakeholders have a better knowledge of local specificities (e.g. priority needs, special barriers)

2. These conclusions are also confirmed by other studies about involvement of local stakeholders in local energy policies in France [AGORA 2002, Godinot 2004, Bouvier 2005, Trouslot 1995].

3. justifying usefulness and cost-effectiveness of implemented operations: either for contractual requirements, as the CPER (*Contrat de Plan Etat-Région*), a six-year agreement framework between French State and Regional Councils, or towards ratepayers

4. dissemination of the results: presenting evaluation work and results in a pedagogic way, in order to contribute to train local stakeholders to new issues they have to address

Table 2. Specificities of evaluation at local level

<p>methodological specificities</p>	<ul style="list-style-type: none"> - evaluation may be a “counterpart to decentralisation and devolution of jurisdiction” [Trousnot 1995 p.535] - local operations are very diverse, and evaluation methods have to be adjustable to these differences - local operations often involve partnerships between various stakeholders: this complicates their monitoring and requires evaluation methods to enable to take account of each point of view, and to fit to different needs and skills - evaluation has often to be managed on two levels: <ul style="list-style-type: none"> - on-field level: providing synthetic results and feedback to operation managers and partners - central level: reporting activities to supervision bodies and centralising experiences - evaluation has to be integrated in the operation process itself: <ul style="list-style-type: none"> - to strengthen mobilisation of targeted final publics (e.g. communication about gained results is a good mean for raising awareness) - to strengthen involvement of partners and other local stakeholders
<p>technical specificities</p>	<ul style="list-style-type: none"> - lack of practical culture and skills of evaluation - lack of reference data specific to a given territory (most of reference data are only available at national level) and difficulties to assess consistency between local and national data - financial and human means available for evaluation are often limited

- **flexibility**:
 - local initiatives can adapt to local contexts and take advantage of good awareness of field conditions
 - local frames let more freedom to implement actions, and therefore favour new ideas and ways to act
 - local level is more operational: operations can be started faster (e.g. decision making process is shorter, smaller budgets are easier to allocate)
- **anchoring and integrating** different policies on a given territory:
 - when the target territory is well defined, it makes easier cross-sectoral and global⁵ approaches
 - targeting a given territory may enable economies of scope

Moreover, **interactions** between stakeholders at local level⁶ may be prescriptive both in a positive or negative way, resulting either in synergies or in additional barriers.

Drawbacks of implementing operations at a local scale

Main drawbacks of local dimension are:

- unbalanced positions for negotiations between local and national (or international) bodies (e.g. negotiations between a local authority and supermarket distribution)
- smaller opportunities for economies of scale in comparison to national operations
- smaller financial means available (in comparison to national operations)

SPECIFICITIES OF EVALUATION AT LOCAL LEVEL

Evaluating local EE activities has to take account of their local dimension as expressed above. But it also includes other specificities

Methodology structure and principles

Specifications described in the above section enable to draw the structure and principles for our evaluation methodology.

GENERAL STRUCTURE BASED ON MAIN EVALUATION OBJECTIVES AND FIELDS

Analysing the inventory of French local EE activities [Broc 2005] highlighted concrete needs in terms of ex-post evaluation (see Table 1). These needs can be transcribed in concrete **evaluation objectives**⁷:

- **understanding** operation process to bring out success factors
- performing a **critical analysis** of the operation and its results, especially comparisons to global objectives this operation contributes to and to other similar operations
- making results **reliable and visible**, especially towards partners and final publics targeted
- **quantifying** final results in terms of load reductions, energy savings and avoided emissions
- defining **cost-effectiveness** of final results (e.g. in terms of c€/saved kWh)

Our methodology is then designed to answer to these objectives through a systematic approach structured in three main evaluation fields:

5. global means here integrating different public policies: e.g. activities for a given neighbourhood considering both housing and transportation issues

6. for France, these interactions concerning local energy policies have been well analysed by Bouvier [2005]

7. Another significant evaluation objective may be to assess the sum of final results of several local operations and to compare it with the evolution of global indicators (crossing bottom-up and top-down approaches). This is out of the scope of this paper.

Table 3. Three main evaluation fields for a systematic approach

<p>1) operation theory analysis:</p> <ul style="list-style-type: none"> - analysis of operation design and progress - confrontation final results / initial objectives - assessment of intermediate outputs and outcomes - involvement of partners - review of communication plan and its impacts - evaluation of market transformation (when required) - analysis of point of view differences among partners - comparison with other similar operations 	<p>2) evaluation of final results:</p> <ul style="list-style-type: none"> - energy savings - load reductions - avoided GHG emissions 	<p>3) economic assessment:</p> <ul style="list-style-type: none"> - assessment of cost-effectiveness indicators (e.g. c€/ saved kWh) - costs/benefits analysis (according to stakeholders points of view)
--	--	--

Table 4. Main principles of our methodology

Principles	Explanations
<p>1) evaluation methods have to be operational and easy to appropriate for all stakeholders involved, and to enable a progressive training to evaluation</p>	<ul style="list-style-type: none"> - a unique methodology is designed to develop several methods: general evaluation objectives are common to all local EE activities and evaluations should be consistent with each other and easy to reproduce ; but differences between operations require to adjust methods according to certain criteria such as targeted end-uses or policy instruments - evaluation methods have to be based on techniques for data collection and analysis which are already experienced - evaluation methods are designed similarly to software: <ul style="list-style-type: none"> - an interface gathers in a constant structured way all required information related to the operation and its results - two modules (one for analysing operation theory, the other for calculating results and cost-effectiveness indicators) provide calculation models and/or frameworks to perform evaluation and obtain conclusions and results from input data in order to feed the interface - tutorials provide additional advice to use evaluation modules and interface, from basic use (requiring as few input data and analysis as possible) to expert use (requiring additional input data and analysis)
<p>2) evaluation methods have to insure data quality and results transparency:</p>	<ul style="list-style-type: none"> - a systematic monitoring of evaluation provides the required information for quality indicators about data and evaluation techniques used (this is based on the Quality Assurance Guidelines approach developed by Vine [Vine 1999a pp.50-52]) - advice are provided to perform uncertainties assessment, from qualitative (order of magnitude) to quantitative (statistical confidence intervals)
<p>3) evaluation methods have to support a progressive process of experience capitalisation (see details below)</p>	

MAIN PRINCIPLES OF OUR METHODOLOGY

To fit to stakeholders needs, our methodology is also based on three main principles.

So far, experience feedback has been used to form **best practices guidebooks or databases**⁸. However, the inventory of French local EE activities brought out that such guidebooks or databases covered only a very little portion of what is actually done [Broc 2005]. Indeed, this form of experience capitalisation is informal, mainly based on a **face-to-face transmission** among insiders circles. Descriptions made of best practices do not go into details, as they are made to raise reader's interest to contact persons who managed the experiences presented.

This way of disseminating best practices is limited, especially because correspondent contacts have often a **limited availability**, which limits dissemination of best practices. Likewise, these contacts have also **skills in a particular field** but not in all issues, which do not favour cross-cutting approaches. Moreo-

ver, the **memory** of experience feedback and know-how may be lost when correspondent contacts leave.

Rakoto [Rakoto 2002, Rakoto 2004] also noticed these issues studying how experience feedback was used in companies. He suggested rationalising this process using knowledge management systems. Our methodology aims to be a way to initiate such a systematic process, applied to the field of local EE activities.

Figure 1 describes how it proposes to initiate a process of experience capitalisation.

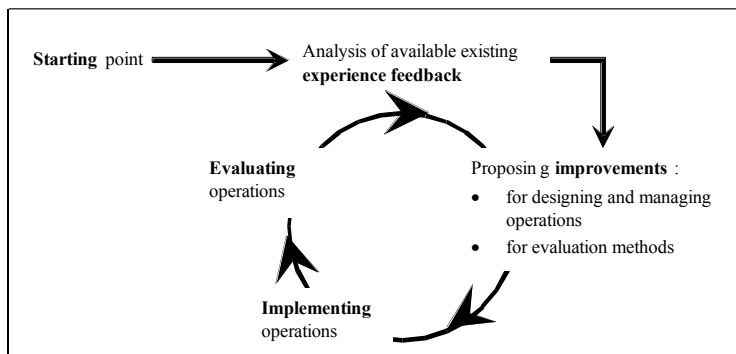


Figure 1. Progressive process for experience capitalisation.

8. Such best practices guidebooks and database are for instance built by Energie-Cités, most often with support of ADEME (French Agency for Energy Management and Environment) and European Commission. Energie-Cités' database is available at: www.energie-cites.eu (section "Resources" then "Best Practices").

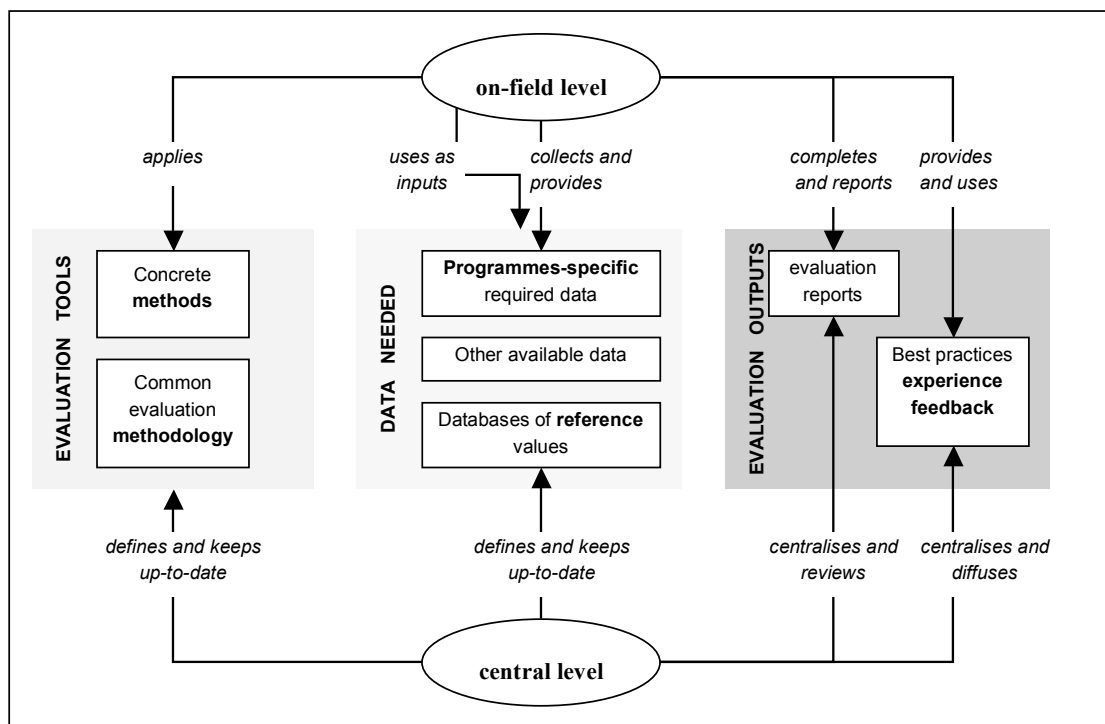


Figure 2. Evaluation system on two levels

SETTING UP AN EVALUATION SYSTEM ON TWO LEVELS

One of the key factors for experience capitalisation is operation managers being the heart of the process. However, for the particular field of local EE activities, these operation managers may be numerous and scattered. There is therefore a need to centralise information. That’s why our methodology is to set up an evaluation system on two levels (see Figure 2).

On-field evaluation is performed by (or under supervision of) operation managers. They are not evaluation experts and moreover get often only little time to devote to this. Their evaluation tools shall therefore be easy to appropriate and to apply. But these tools shall also propose further guidelines to enable going deeper into evaluation when operation managers wish so, especially in an approach of continuous improvement.

The objectives of **centralising evaluations** are:

- to gather information in order to make them available for both, decision-makers and operation managers
- to review information in order to insure evaluations are reliable and can be compared with each other
- to update evaluation methods, reference data and best practices guidebooks (for implementing operations)

This centralisation may be performed either within a national body (e.g. national agency, energy suppliers), or for a given territory (e.g. at regional scale, by a regional energy observatory/agency). Guidelines for centralisation are meant for evaluation experts, and are to register information in a systematic way, in order to provide structured and detailed experience feedback. These guidelines are also to be used to complete on-field evaluations, when the central evaluation service decides it is relevant.

In parallel, a **standard form** was designed to collect in systematic way information about local operations and their

evaluations. This form provides a frame to summarize this information. Each data holding key information is marked with a given code (e.g. A-1). These marks are common for all tools of an evaluation method (e.g. the standard form, calculation guidelines, etc.), so that each information can be easily located in any tool. It could also be used afterwards to build an information system.

Setting up an evaluation methodology

This section presents the main components of our methodology, first its methodological background, and then the key issues addressed both for impacts quantification and for theory-based evaluation.

METHODOLOGICAL MATERIALS USED

Selection of main bibliographic references was based on three criteria:

- **recognition and/or use** by international or national institutions and/or by energy efficiency professionals (utilities, ESCo, regulatory or other public bodies)
- **up-to-date contents** (based on current state-of-art and/or recent experiences)
- **wide methodological scope** (covering most of main evaluation issues)

Other bibliographic inputs have been used, especially for specific evaluation issues (e.g. [Vine 1992] for persistence of savings). But the following references provided the basic material for our methodology.⁹

9. the IEA DSM Programme: International Energy Agency Demand-Side Management Programme, see <http://dsm.iea.org>

Table 5. Main references used to build our methodology.

Reference	Contents
International Performance Measurement and Verification Protocol [IPMVP 2002]	Prepared for the US Department of Energy (DoE), IPMVP stands as an international reference for M&V (Measurement and Verification) of energy savings. A summary of IPMVP can be found in [Appendix B of SRCI 2001 pp.B50-B53]. Main inputs from IPMVP are: <ul style="list-style-type: none"> - 4 M&V options to quantify gross energy savings - description of measurement techniques and issues
Guidelines for the Monitoring, Evaluation, Reporting, Verification, and Certification of energy-efficiency projects for climate change mitigation [Vine 1999a]	Prepared for the US Environmental Protection Agency and US DoE, the MERVC guidelines were developed to evaluate CDM (Clean Development Mechanism) and JI (Joint Implementation) projects within Kyoto protocol framework. A summary can be found in [Vine 1999b]. Main inputs from MERVC guidelines are: <ul style="list-style-type: none"> - global approach from ex-ante registration to ex-post verification - methodological approach to evaluate net energy savings (estimating gross results, then a baseline, and finally net results) - guidelines to estimate ex-ante an initial baseline (through performance benchmarks), and then to re-estimate it ex-post (through surveys, discrete choice or regression models), in order to better capture free-rider effect - quality assurance guidelines, to insure quality of performed evaluations and thus reliability of results - reporting forms
A European ex-post evaluation guidebook for DSM and EE service programmes [SRCI 2001]	Prepared for the European Commission within a SAVE project, the European guidebook stands as an international reference for evaluation preparation and planning. A summary of its approach can be found in [Birr-Pedersen 2001]. Main inputs from the European guidebook are: <ul style="list-style-type: none"> - guidelines for evaluation planning - synthetic description of the basic concepts related to the evaluation of EE programmes (especially net-to-gross adjustment factors) - description of the main techniques for data collection and energy savings calculation
California Energy Efficiency Protocols and Evaluation Framework [TecMarket Works 2004, TecMarket Works 2006]	Prepared for the California Public Utilities Commission (CPUC), both manuals provide detailed guidelines for energy efficiency programmes evaluation to the evaluation experts and programme managers in charge of it. The Californian Evaluation Framework covers all evaluation issues, forming a very complete state of the art of evaluation practices in United States. The protocols present official requirements and recommended process for evaluation of EE programmes under CPUC regulation. Main inputs from Californian manuals: <ul style="list-style-type: none"> - systematic approach of evaluation work, divided in main issues: M&V and impact evaluation, process evaluation, market transformation evaluation, uncertainty, sampling, cost-effectiveness - detailed review of all evaluation key issues (with a significant bibliography) - detailed requirements and organisation of an evaluation scheme in a given regulatory framework
Evaluation guidebook for evaluating energy efficiency policy measures & DSM programmes [Vreuls 2005]	Prepared within task 1 of IEA DSM Programme, the objective of the IEA guidebook is “ <i>to provide practical assistance to administrators, researchers, and policy makers who need to plan assessments and to evaluators who carry out evaluations of energy efficiency programmes</i> ”, especially programmes related to Kyoto greenhouse gas targets. Main inputs from IEA guidebook: <ul style="list-style-type: none"> - a methodological approach based on seven key elements: statement of policy measure theory, specification of indicators, development of baselines for indicators, assessment of output and outcome, assessment of energy savings, calculations of cost-effectiveness, and choice of level with regard to the evaluation effort - application of this framework to four main types of EE programmes (regulation, information, economic incentives, voluntary agreements) and to combinations of these types - detailed and standardised description of main evaluation experiences of eight countries (Belgium, Canada, Denmark, France, Italy, Republic of Korea, The Netherlands, Sweden)
Evaluation and comparison of utility's and governmental DSM-programmes for the promotion of condensing boilers [Haug 1998]	Contrary to previous references, this SAVE project is not general but specific to a particular energy efficient solution (condensing boilers). It was selected because it presents a well-detailed evaluation approach and because it represents an example of specific material to be used developing a specific method. Main inputs from this particular evaluation are: <ul style="list-style-type: none"> - a concrete and detailed evaluation framework, especially to study factors influencing the success of a programme - significant material about market transformation issues, which can be used for other cases of promotion of efficient appliances - significant technical material useful for developing method for other energy efficiency programmes related to heating

Table 6. A four steps calculation process

step 1	calculating the unitary gross annual results : main issues raised for this step are the baseline , the correction factors and the choice of a calculation method (see the following section)
step 2	calculating the total gross annual results : main issue raised for this step is to define a method to account for the number of participants, especially when this accounting is not direct (e.g. when participants are not directly registered, their number may be assessed through sales data)
step 3	calculating the total net annual results : main issue raised for this step is to adjust the baseline or to apply adjustment factors
step 4	calculating total net results over time : main issue raised for this step is to assess the lifetime of the results, taking account of their persistence

KEY ISSUES FOR IMPACTS QUANTIFICATION

Making it clear what are gross and net results

Gross results are the results gained from the **point of view of final end-users**. These results correspond to the changes between the situations before and after taking part in an EE programme. Situation before may be either the energy consumption in the period preceding the programme, or the changes which would have occurred in absence of the programme (e.g. buying a present standard equipment instead of a present efficient one). Defining such a before situation means **choosing a baseline**. Besides, before and after situations have to be compared in similar conditions (e.g. similar weather conditions). This requires the application of **correction** (or normalisation) **factors** (e.g. heating degree days).

Net results are the results gained thanks to an EE programme, from the **point of view of society**. This means net results only account for the portion of the results which would not have been gained without the programme. The difference between gross and net results can be taken into account, either in **adjusting the baseline** (e.g. using control groups or market modelling), either in **applying adjustment factors** (such as free-rider and spill-over effects).

For local operations, an interesting option is to **consider what occurs at national level as a control group**. For instance, for a local operation promoting CFL (Compact Fluorescent Light bulb), the baseline can be the national evolution of CFL sales¹⁰. An advantage of this alternative is that it may enable to take into account the effects of national programmes. However, it has to be controlled whether the comparison between national and local evolutions is not biased (e.g. due to local specificities).

Defining a four steps calculation process

To make the calculation of the results easier, it is possible to perform it as a step-by-step process.^{11,12}

- Setting up an operational evaluation method for calculating impacts results then in two main tasks: specifying a calculation method (for step 1) and an accounting method (for steps 2 and 3). Persistence studies (for step 4) are often to

10. If the local operation cover a significant territory, the national evolution has to be understood as the evolution for the rest of the country (not covered by the operation).

11. unitary results means here either the results per participant or per equipment

12. Persistence of results means the efficient solution is still in place and operational (retention), with a certain level of performance compared to the first year of implementation (performance degradation) [Vine 1992, Wolfe 1995].

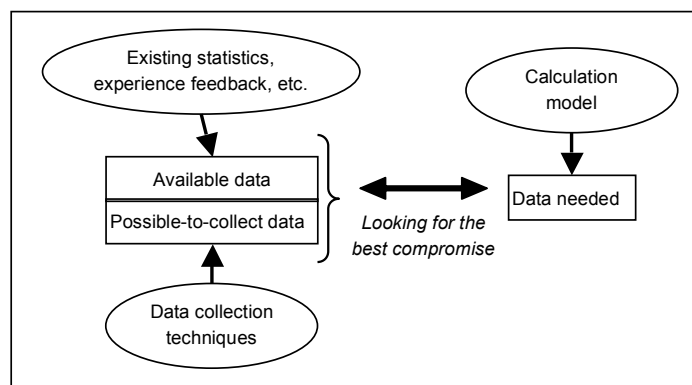


Figure 3. Iterative process for choosing a calculation method.

be managed at the central level to be more cost-effective [Skumatz 2005].

Choosing a calculation method

A calculation method is composed of three main elements: a **calculation model**, **data collection techniques** to feed this model, and **reference ex-ante values** which complete the data collected ex-post. Choosing a calculation method is then an iterative process. A particular calculation model will require a given set of data. Looking for what data are already available or not, it can be deduced what data should be collected. In parallel, the different possible collection techniques¹³ can be considered, taking into account their technical¹⁴ and practical¹⁵ feasibility, and their costs. After that, available and possible-to-collect data are compared to the data needed. Finally, the choices of calculation model and data collection techniques are adjusted until available and possible-to-collect data fit to data needed.

Two calculation approaches

Two main approaches may be used to calculate unitary energy savings, whether energy consumption data are directly available or not:

1. When energy consumption data are **directly available**, at least for a sample of participants, the general calculation formula compares directly energy consumption before and after the implementation of the evaluated efficient solution.

13. A list of these techniques can be found at: <http://www.tecmarket.net/data.htm>

14. e.g. it may not be possible to install a sub-meters, or to disaggregate data from existing metering

15. e.g. some data may be difficult to access for privacy (e.g. commercial data) or confidentiality (personal data) reasons

Correction or normalisation factors may be applied when necessary as explained before. Energy consumption data can be collected either from energy bills, meter readings or energy end-use measurements.

Equation 1. Direct formula

<p>unitary gross annual energy savings = ([annual energy consumption]₀ - [annual energy consumption]₁) +/- normalisation factors where: - 0 and 1: situation respectively before and after implementation of an efficient solution</p>

2. When energy consumption data are not directly available, formula can be broken down in intermediate parameters (for which data are easier to access/assess) to calculate energy consumption afterwards. For instance, energy consumptions may be broken down in two terms, load and duration of use:

Equation 2. Broken down formula.

<p>unitary gross annual energy savings = ([P*ALF*D]₀ - [P*ALF*D]₁) where: - 0 and 1: situation respectively before and after implementation of an efficient solution - P: average rated power - ALF: average load factor - D: average annual duration of use</p>
--

Other decomposition may be used (e.g. average consumption per cycle and number of cycles for washing machines). And as for energy consumption, the intermediate parameters may be related to additional parameters (e.g. number of persons per dwelling for the number of cycles).

Quality control and uncertainty analysis

Qualification of the input data

Assessing quality of input data results in linking their values either with a **range of variations** (between a minimum and a maximum), or with a **confidence interval** (and its confidence level). In practice, the latter is difficult to apply. And sometimes, even the former option may not be possible. In that case, it may be used a qualitative appreciation (e.g. approximate, medium, good).

Qualification of the data sources

Data sources can be distinguished either they were assessed ex-ante or defined ex-post. **Ex-ante** data are reference values deduced from other existing experiences, studies or statistics. **Ex-post** data are defined from characteristics, measurements, monitoring or surveys specific to the operation evaluated. Usually, an ex-post data will be more accurate than an ex-ante data, as it takes into account the operation specificities. However, the opposite is also possible, especially when ex-post data are defined from non-representative samples.

Quality Assurance Guidelines (for the evaluation)

No calculation model can be assumed to be more accurate than the others¹⁶. Evaluation reliability depends more on the input data and the assumptions done than on the model itself. That's why Vine [Vine 1999a pp.50-52] proposed to use **quality assurance guidelines** to assess the application of an evaluation method. Our methodology has included this approach, adjusted to the specificities of evaluation local operations.

Three progressive levels for addressing uncertainties

At the end, the aim is to present the results with information about their **uncertainties**. So far, such information has never been found, among all experience feedback of local EE operations we could find. It would therefore be unrealistic to expect results with confidence interval right away. So our methodology proposes a **progressive approach**, with three levels of information. First level is at least an **order of magnitude**, based on qualitative appreciations and/or assessments of minimum and maximum values. Second level is a **sensitivity analysis**, using the variation ranges of each parameter to define pessimistic and optimistic scenarios, leading to more accurate minimum and maximum values. Third level is either a **statistical approach** leading to confidence intervals (at given confidence levels), or Monte Carlo simulations, which analyses the distribution of the results from random sensitivity analysis.

16. see for instance [Ridge 1994, Schiffman 1993] for a comparison between statistical and engineering models.

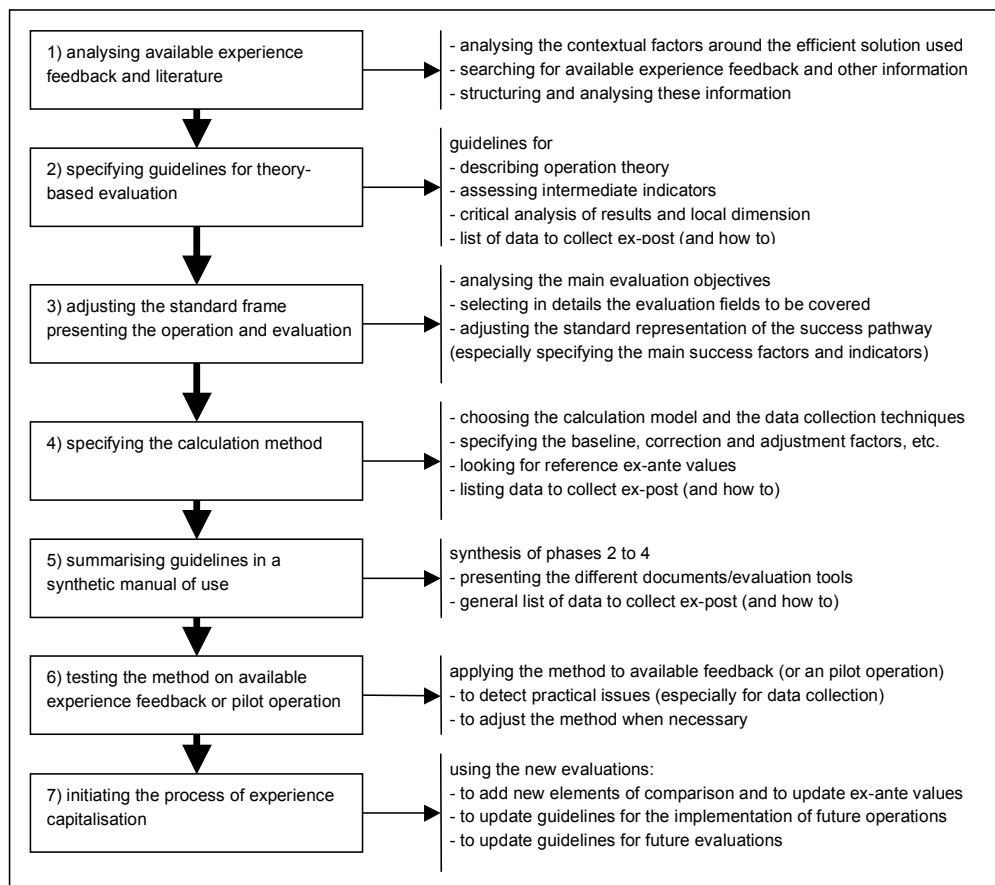


Figure 4. Systematic process for developing operational methods.

KEY ISSUES FOR THEORY-BASED EVALUATION

Systematic description of the operation

First step in analysing the operation theory is to describe it. The standard frame designed to report evaluations enables a systematic description, making this step easier. The frame is particularly used to **make the objectives** of the operation **more explicit**, and thus to deduce the priority objectives for the evaluation. It also brings out the main **contextual factors and assumptions** of the operation, especially the barriers the operation is assumed to address.

Expressing operation theory into concrete intermediate indicators and success factors

From the systematic description of the operation, the operation theory is then expressed into concrete intermediate indicators and/or success factors possible to monitor. Usually, such an approach is used to represent the operation theory as a linear (or step-by-step) process (see for instance [Joosen 2005]). However this approach may either require a large amount of data (trying to be exhaustive), or miss some key factors of success (trying to be straightforward). Our methodology suggests a compromise by representing the operation theory as a **non-linear combination of success factors or intermediate steps**, selecting the most relevant factors/steps from existing experience feedback. The aim is also to bring out possible interactions between success factors and/or intermediate steps. The final result of this evaluation work is the representation of the operation theory as a **success pathway**.

Critical analysis of results

As evaluation results are always relative, our methodology provides guidelines for their critical analysis. The two main guidelines are to take into account the possible **differences of points of view** among stakeholders (e.g. between a public agency, a utility, the retailers and the participants) and to **compare** the operation evaluated with similar ones. This comparison is to focus not only on the results but also, and in priority, on the choices made while designing the operation (e.g. what communication means are used, what partnerships were set up).

Critical analysis of local dimension

As our methodology is to evaluate local operations, it is essential to provide guidelines to consider the local dimension of the operations analysed. These guidelines are mainly based on the **specificities** of the local dimension described in section “*Advantages and drawbacks of the local dimension*”. The approach is to review each specificity, whether it was an effective advantage or drawback for the particular operation.

Systematic process for developing operational methods

DESCRIPTION OF THE PROCESS

The final result of our methodology is a process for developing operational methods, whose aim is to address in a systematic way the key evaluation issues described in above sections, in order to complete the evaluation tools forming an evaluation method.

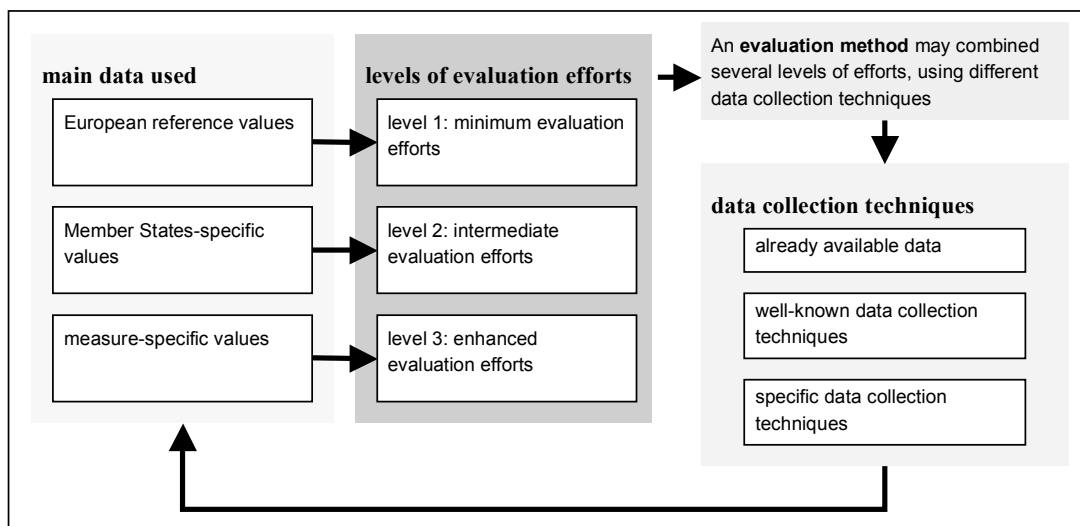


Figure 5. Level of evaluation efforts related to data collection techniques.

HOW TO INVOLVE CONCERNED STAKEHOLDERS IN THE PROCESS

Involving the stakeholder in the process for developing evaluation methods is a key factor to insure the methods are operational and they will use it, and so be involved in the evaluation process. This was tried while applying our methodology on two concrete cases (local campaigns for promoting CFL, and raising awareness actions in office buildings). The personal contacts were very good, and operation managers as well as central bodies were very interested in the evaluation process, even before our first contacts. So these applications were successful and provided interesting feedback and lessons learned. However, the issue of involving stakeholders should be further examined, especially for less favourable cases.

Conclusions and perspectives

As mentioned above, our methodology was tested on two concrete case studies of local operations implemented by EDF in South-East of France. The first one was a local campaign to promote CFL. A complete evaluation method was set up and applied, providing detailed results [Broc 2006a]. The second case study was about raising awareness actions in office buildings. Existing experience feedback was not rich enough to set up a complete evaluation method. The part on theory-based evaluation could be fulfilled, but not the part on impact evaluation. Still, this case study enabled us to obtain interesting results [Broc 2006b]. Presenting both, the methodology and these case studies, was not possible in a single paper. So the analysis of the case studies will be the subject of further publications.

In parallel, this methodology is currently adjusted to evaluate larger EE activities, especially EE activities to be reported within the new European Directive on energy end-use efficiency [ESD 2006]. Changing of scale and fitting to particular regulatory requirements raise additional evaluation issues, for instance:

- specifying harmonised accounting rules for Member States with different experiences about evaluation

- double counting of results between EE activities with similar targets

One of the main evolutions of our methodology is to consider three levels of evaluation efforts¹⁷.

These new developments of our methodology are currently studied and discussed within the Intelligent Energy Europe project entitled EMEEES (Evaluation and Monitoring for the EU Directive on Energy end-use Efficiency and Energy Services), coordinated by the Wuppertal Institute for Climate Environment and Energy¹⁸.

References

- [AGORA 2002] AGORA. *ALES - Autorités Locales et Effet de Serre*. contrat pour l'ADEME, décembre 2002.
- [BIRR-PEDERSEN 2001] Preben Birr-Pedersen. *European ex-post evaluation guidebook for DSM and EE services programmes*. Proceedings of the ECEEE 2001 Summer study, pp.116-118, panel 1, paper 1.098.
- [BOUVIER 2005] Guillaume Bouvier. *Les Collectivités locales et l'électricité - Territoires, acteurs et enjeux autour du service public local de l'électricité en France*. Thèse CIFRE de Doctorat en géographie (mention géopolitique), réalisée à la Direction des Affaires Publiques d'EDF, Institut Français de Géopolitique - Université Paris 8, 17 juin 2005.
- [BROC 2006a] Jean-Sébastien Broc. *L'évaluation ex-post des opérations locales de Maîtrise de la Demande en Energie - Etat de l'art, méthodes bottom-up, exemples appliqués et approche du développement d'une culture pratique de l'évaluation*. Thèse de Doctorat en Energétique de l'Ecole de Mines de Paris, réalisée à l'Ecole des Mines de Nantes, 8 décembre 2006.
- [BROC 2006b] Jean-Sébastien Broc, Bertrand Combes, Bernard Bourges, Jérôme Adnot, Sandrine Hartmann,

17. This approach is inspired from recent guidebooks, such as the Danish [SRCI 2003], the IEA [Vreuls 2005 pp.43-47], and the IPCC [IPCC 2006 pp.1.7-1.8] ones.

18. For more details, see the paper presented by Stefan Thomas at this ECEEE 2007 Summer Study.

- Marie-Isabelle Fernandez. *Raising awareness for energy efficiency in the service sector: learning from success stories to disseminate good practices*. Proceedings of the IECEB 2006 conference, *Improving Energy Efficiency in Commercial Buildings*, Frankfurt am Main, Germany, 26-27 April 2006.
- [BROC 2005] Jean-Sébastien Broc, Bernard Bourges, Jérôme Adnot, Sandrine Hartmann. *Local energy efficiency and demand-side management activities in France*. Proceedings of the ECEEE 2005 Summer Study, volume I, panel 1, paper 1,202, pp.183-194, *Energy savings: what works and who delivers?* Mandelieu La Napoule, France, 30 May - 4 June 2005.
- [ESD 2006] ESD. *Directive 2006/32/EC of the European Parliament and of the Council of 5 April 2006 on energy end-use efficiency and energy services and repealing Council Directive 93/76/EEC*. European Union, 2006/32/CE, Brussels, 5 April 2006.
- [EUROPEAN COMMISSION 1999] European Commission. *Paper I - Defining criteria for evaluating the effectiveness of EU environmental measures*. Towards a new EU framework for reporting on environmental policies and measures (Reporting on environmental measures - 'REM')
- [GODINOT 2004] Sylvain Godinot, Jacques Ravailault, Laurent Comélieu. *Collectivités locales et effet de serre : étude sur la mise en oeuvre locale de la lutte contre le changement climatique*. Convention de recherche pour le compte de la MIES, ADEME, Direction de l'Action Régionale, février 2004.
- [HAUG 1998] J. Haug, B. Gebhardt, C. Weber, M. Van Wess, et al. *Evaluation and comparison of utility's and governmental DSM-Programmes for the promotion of condensing boilers*. SAVE contract XVII/4.1031/Z/96-136, IER (Institut für Energiewirtschaft und Rationelle Energieanwendung), Stuttgart, Germany, October 1998.
- [IPCC 2006] IPCC, Simon Eggleston, Leandro Buendia, Kyoko Miwa, Todd Ngara, Kiyoto Tanabe. *2006 IPCC Guidelines for National Greenhouse Gas Inventories - volume 2 : Energy*. Prepared by the National Greenhouse Gas Inventories Programme, Published by IGES on behalf of the IPCC, Japan. (available at: www.ipcc-nggip.iges.or.jp)
- [IPMVP 2002] IPMVP. *International Performance Measurement & Verification Protocol - Volume I : Concepts and Options for Determining Energy and Water Savings*. report DOE/GO-102002-1554 for the US Department of Energy, Efficiency Valuation Organization, Revised edition of March 2002. (available at: www.evo-world.org)
- [JOUSEN 2005] Suzanne Joosen, Mirjam Harmelink. *Guidelines for the ex-post evaluation of policy instruments on energy efficiency*. Report EIE-2003-114, AID-EE project for the European Commission within the Intelligent Energy for Europe (EIE) programme, August 2005. (available at: www.aid-ee.org)
- [RAKOTO 2004] Holitiana Rakoto. *Intégration du Retour d'Expérience dans les processus industriels - Application à Alstom Transport* Thèse CIFRE de Doctorat en Systèmes Industriels, réalisée au Département Qualité d'Alstom Transport SA, Institut National Polytechnique de Toulouse, 15 octobre 2004.
- [RAKOTO 2002] Holitiana Rakoto, Philippe Clermont, Laurent Geneste. *Elaboration and Exploitation of Lessons Learned*. Proceedings of the IFIP 17th World Computer Congress - TC12 Stream on Intelligent Information Processing, IFIP Conference Proceedings; Vol. 221
- [RIDGE 1994] Richard Ridge, Dan Violette, Don Dohrmann. *An Evaluation of Statistical and Engineering Models for Estimating Gross Energy Impacts*. Final report of Pacific Consulting Services for the California Demand Side Management Advisory Committee, Berkeley, California, June 1994. (available at: www.calmac.org)
- [SCHIFFMAN 1993] Dean A. Schiffman, Robert F. Engle. *Appendix Z Simulation Study: Comparison of Alternative Methods for Measuring the Gross and Net Energy Impacts of Demand-Side Management Programs (with Addendum)*. Report for San Diego Gas and Electric, August 1993. (available at: www.calmac.org)
- [SKUMATZ 2005] Lisa Skumatz, John Gardner. *Best Practices in Measure Retention and Lifetime Studies: Standards for Reliable Measure Retention Methodology Derived from Extensive Review*. proceedings of the 2005 International Energy Program Evaluation Conference, Session 4B Persistence - the issue that never goes away, pp.483-493, *Reducing Uncertainty Through Evaluation*, New York, 17-19 August 2005.
- [SRCI 2003] SRCI, AKF, Elkraft system. *Handbog i evaluering af energispareaktiviteter (Handbook for evaluating energy saving projects)*. January 2003.
- [SRCI 2001] SRCI, NOVEM, Electricity Association, MOTIVA, et al. *A European Ex-Post Evaluation Guidebook for DSM and EE Service Programmes*. SAVE Project No. XVII/4.1031/P/99-028, April 2001.
- [TECMARKET WORKS 2006] TecMarket Works. *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. Report prepared for the California Public Utilities Commission, April 2006. (available at: www.calmac.org)
- [TECMARKET WORKS 2004] TecMarket Works, Megdal & Associates, Architectural Energy Corporation, RLV Analytics, et al. *The California Evaluation Framework*. Report prepared for the Southern California Edison Company as mandated by the California Public Utilities Commission, K2033910, Revised September 2004. (available at: www.calmac.org)
- [TROUSLOT 1995] Franck Trouslot. *L'évaluation des actions publiques à l'échelon local : illustration et analyse critique à partir de l'exemple de la politique de maîtrise de l'énergie en Poitou-Charentes*. Thèse de Doctorat en Sciences Economiques, Université de Poitiers, Faculté de Sciences Economiques, décembre 1995.
- [VINE 1999a] Edward Vine, Jayant Sathaye. *Guidelines for the Monitoring, Evaluation, Reporting, Verification and Certification of Energy-Efficiency projects for Climate Change Mitigation*. Report prepared for the U.S. Environmental Protection Agency, LBNL-41543, March 1999.
- [VINE 1999b] Edward L. Vine, Jayant A. Sathaye. *An Overview of Guidelines and Issues for the Monitoring, Evaluation, Reporting, Verification and Certification of*

Energy-Efficiency Projects for Climate Change Mitigation.
Proceedings of the ECEEE 1999 Summer Study.

- [VINE 1992] Edward L. Vine. *Persistence of energy savings: What do we know and how can it be ensured?* *Energy*, **17** (11), pp. 1073-1084.
- [VREULS 2005] Harry Vreuls, Wim De Groot, Peter Bach, Richard Schalburg, et al. *Evaluating energy efficiency policy measures & DSM programmes - volume I : evaluation guidebook*. Report for the IEA-DSM task IX. October 2005. (available at: <http://dsm.iea.org>)
- [WOLFE 1995] Amy K. Wolfe, Marilyn A. Brown, David Trumble. *Measuring persistence : a literature review focusing on methodological issues*. Report n° ORNL/CON-401 by the Oak Ridge National Laboratory for the US-DOE, Washington DC, march 1995.

Acknowledgements

The work presented in this paper was done during a PhD thesis jointly supervised in Ecole des Mines de Nantes and Ecole des Mines de Paris. A part of this PhD consisted in a research contract, entitled EVADEM, for EDF R&D performed in partnership with the Wuppertal Institute for Climate Environment and Energy. Current works to go further are mainly led within the Intelligent Energy Europe project entitled EMEEES (Evaluation and Monitoring for the EU Directive on Energy end-use Efficiency and Energy Services), coordinated by the Wuppertal Institute for Climate Environment and Energy.

The authors would like to thank Paul Baudry and Sandrine Hartmann (EDF R&D), Marie-Isabelle Fernandez and Bertrand Combes (Regional Delegation of EDF PACA), Stefan Thomas and Wolfgang Irrek (Wuppertal Institute) for the successful collaboration within EVADEM. We would also like to thank Stefan Thomas (again), Harry Vreuls (SenterNovem) and all EMEEES partners for their comments within EMEEES.