ANNEX 2
METHODOLOGIES AND TOOLS USED

Contents

List of Figures and Tables 1
Acronyms and Abbreviations 2
A2.1 Evaluation Design 3
A2.2 Applying DAC Criteria 3
A2.2.1 Efficiency, Effectiveness, Relevance, Impact and Sustainability 3
A2.2.2 DAC Criteria applied to the Logical Framework 4
A2.2.3 General Evaluation Methodology 5
A2.3 Approaches to Fact Finding 6
A2.3.1 Deductive approach 6
A2.3.2 Inductive Approach 6
A2.4 Benchmarking - Method of Systematic Comparison 7
A2.4.1 General approach 7
A2.4.2 Infrastructure development 9
A2.4.3 Training 12
A2.4.4 Capacity Development 13
A2.4.5 Combined interventions 14
A2.4.6 Model Uncertainties and Weaknesses 15
A2.4.7 The Data Entry and Calculation Tool 15
A2.4.8 Approach to Rural Field Studies 18
A2.5 Hypotheses 20
A2.5.1 Hypotheses on Assistance Strategy 20
A2.5.2 Hypotheses Related to Key Evaluation Questions 20
A2.5.3 Impact on Poverty 21
A2.5.4 Affordability Issues 21
A2.5.5 Capacity Development Hypotheses 22
A2.6 Methods for Assessment of Macro and Meso Economic Impacts 22
A2.7 Quality Assurance 23

List of Figures and Tables

Fig. A2- 1: Types of Interventions and Evaluation in the Project Cycle 3
Fig. A2- 2: Results Hierarchy and DAC Criteria 5
Fig. A2- 3: Benchmarking Variables 7
Fig. A2- 4: Definition of Variables across Intervention Types 8
Fig. A2- 5: Comparison Factors Used in Assessment 9
Fig. A2- 6: Indirect Benefit of Electrification 11
Fig. A2- 7: Generic Infrastructure Costs and Benefits 12

Acronyms and Abbreviations

DAC	Development Assistance Committee (of OECD)
FTE	Full Time Equivalent (of consultant/advisor input, training output)
GDP	Gross Domestic Product
GW	Giga Watt, unit of power capacity
GWh	Giga Watt hour, unit of energy
km	Kilometre
kW	Kilo Watt, unit of power capacity
kWh	Kilo Watt hour, unit of energy
LFA	Logical Framework Analysis
MWh	Mega Watt hour, unit of energy
NGO	Non-Governmental Organisation
Norad	Norwegian Development Agency
O&M	Operation & Maintenance
PPP	Purchase Power Parity
QA	Quality Assurance
SADC	Southern African Development Community
SAPP	Southern African Power Pool
SMART	Specific, Measurable, Attainable, Relevant and Time-bound (indicators)
TOR	Terms of Reference (of the evaluation, unless specified otherwise)
USD	Currency: United States Dollar, 21 April 2007 – 1 USD = 5.96 Norwegian Kroner

A2.1 Evaluation Design

The evaluation in general is based on a case-study design which covers two of Norway's partner countries and a limited study of the results on SADC and SAPP frameworks for regional cooperation by the Norwegian assistance. The evaluation built on information produced by earlier reviews, appraisals or study reports, but used such second-hand accounts as a starting point for analysis of first-hand empirical material to achieve a deeper and broader understanding.

A benchmarking approach has been used to achieve more objective assessment by systematic comparison across all types of intervention and countries.

The analytical framework for the evaluation clarifies what types of interventions have been identified, and at what level the evaluation is directed in the project cycle. This is shown in fig. A2.1.

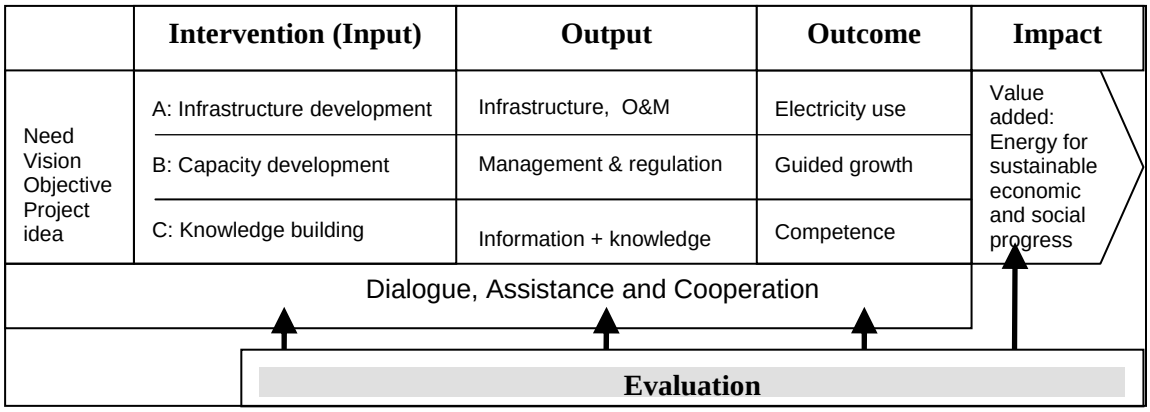

Fig. A2- 1: Types of Interventions and Evaluation in the Project Cycle

Operationalizing the requirements of the evaluation, results in the identification of four levels of evaluation:

1. *Assessments at the activity level*: Inputs provided, quality, relevance, cost-efficiency, effectiveness, institutional arrangements, relationships built
2. *Evaluation of outputs*: Actual versus planned, appropriate to the context, effective in addressing the targeted issues
3. *Assessment of outcomes*: Have outputs contributed to the desired results? In what way, quality, quantity? What are constraints and counteracting factors?
4. *Evaluation of impacts*: On poverty, industry and commerce, rural and urban life, regional and national economy. The specified rural impact studies are central to the impact evaluation. Due to their context and nature, these studies require a specifically developed approach and methodology.

A2.2 Applying DAC Criteria

The TOR is explicit about the requirement of using Development Assistance Committee (DAC) criteria for the evaluation. The DAC criteria are applied by connection to a logical framework structure.

A2.2.1 Efficiency, Effectiveness, Relevance, Impact and Sustainability

The DAC criteria are designed to promote comprehensive evaluation, and for this reason, the criteria are complementary. For example, evaluation of effectiveness may show that objectives were met, but this does not necessarily mean that the objectives were appropriate for the entire affected population, were met efficiently, are sustainable, or feed into impact. Similarly, an intervention by one agency can achieve good coverage, but may not be coordinated with other interventions. Using the DAC criteria helps to ensure that the evaluation covers all aspects of the intervention. There is often some overlap between criteria, and it is found that the same data can be included under different DAC criteria.

Deciding where to place findings is less important than ensuring that a comprehensive evaluation is carried out, using all of the criteria where relevant.

Use of the DAC criteria tends to focus on results rather than processes, even where process is more important than results. This pitfall is addressed by assessing the process separately, and by asking why-questions.

Another issue is that DAC criteria are subjective, rather than objective. Hence it depends on the perception of stakeholders, and the evaluator's interpretation. Triangulation helps to make assessments less biased, and the evaluation has attempted a degree of objectivity by using uniform valuation standards across interventions, and by benchmarking as further shown in section A2.4.

For any specific evaluation subject, it is useful to determine whether using all of the criteria is relevant. Because they were formulated for evaluation of projects or programmes, some of the DAC criteria may be less relevant for policy-based or institutional evaluations. Hence, rather than going through the motions of applying all criteria on all evaluation subjects religiously, the evaluation will make a sensible choice in each case, and apply what adds value.

To evaluate interventions against the DAC criteria, it is ideal to have measurable objectives and data for baseline and 'results' measures, but these are often missing or partial. Where data were not adequate, these were constructed from available evidence. Data constraints have been specified, where applicable, so that the reader of the report is able to assess the credibility of the methods employed, and results drawn from them.

A2.2.2 DAC Criteria applied to the Logical Framework

The logical framework (or logframe, LFA) is currently the most common planning tool in development assistance, and its use is also widespread in Norwegian assistance. The evaluation will attempt to link logframes to use of the DAC criteria. Experience tells us that logframes may often not exist for a specific intervention. It is then considered a useful method to construct an appropriate logical framework that is valid for the intervention, as part of the evaluation work, including the definition of performance indicators. This will make it possible to apply a systematic analytical framework to all evaluation work. The main areas usually included in the logframe are activities, inputs, outputs and objectives – also known as the results chain. The links between the DAC criteria and the results chain are discussed below.

Efficiency - A measure of how economically resources/inputs (funds, expertise, time, etc.) are converted to outputs. - Issues to consider: Political priorities of governments and agencies may cause interventions to be inefficient. The evaluation therefore will take political factors into account to look for causes of inefficiency, if necessary. What was the source of inputs? Part of the assessment of efficiency considers whether goods/inputs were purchased most efficiently in relation to source of input. One key question is whether inputs were locally purchased or imported. Financial areas to consider are: total cost of the intervention broken down by sector; costs of inputs locally and internationally; transportation costs broken down by sector and type of transportation; staff costs, broken down by local and expatriate staff; and administration costs as a percentage of intervention costs. Efficiency is mainly covered in the input and output columns or rows of the logframe.

Effectiveness – The extent to which the programme's purpose was achieved, or is expected to be achieved - is covered in the contributions of outputs to achieving outcomes. Why were interventions effective, or not? Knowing whether an intervention has met its objectives is half the story. The other half is knowing why and how this happened. In order to understand this, examining cross-cutting themes will be a useful support – for example, related to who participated and why, and the influence of local context. A key element in the assessment of effectiveness is timeliness. The phasing of interventions is often crucial to success, and the evaluation considers whether interventions were carried out in a fashion that adequately supported the purpose. Was timely provision of support, goods and services achieved, according to the perceptions of key stakeholders?

Relevance – The extent to which the objectives of a programme are consistent with beneficiaries' requirements, country needs, global priorities, and partners' and donors' priorities - is concerned with assessing whether the project is in line with local needs and priorities (as well as donor policy). In

relation to the logframe results chain, relevance can be used to evaluate the wider elements of the intervention such as the overall goal or the purpose (or objective). In assessing relevance, the evaluation will take care to assess the local context. One element in this is the cultural appropriateness of the intervention.

Impact – Positive and negative, primary and secondary, long-term effects produced by a development programme, directly or indirectly, intended or unintended – can be considered a higher-order measure and thus relates to the Purpose and Goal in the logframe. Goal statements in the logframe are often vaguely worded with few quantitative targets stated. The evaluation may therefore need to refer to the policy of the supported entity, to determine how the intervention fits with the longer-term goals of the beneficiary. A key element in the assessment of impact is explaining why events happened – that is, what were the processes involved that led to particular results? Assessment of impact may be the most challenging aspect of carrying out an evaluation. Doing it will be constrained by available time and data.

Sustainability – the continuation of benefits from a programme after major development assistance has been completed; the probability of continued long-term benefits – looks at the question whether changes introduced, or solutions provided, will have a useful life beyond the time of the actual assistance. There is no yes or no answer to sustainability. The evaluation looks for factors that are likely to contribute to or work against sustainability, and explain these.

Graph A2-2 below indicates how the evaluation's assessment relates to the log frame structure.

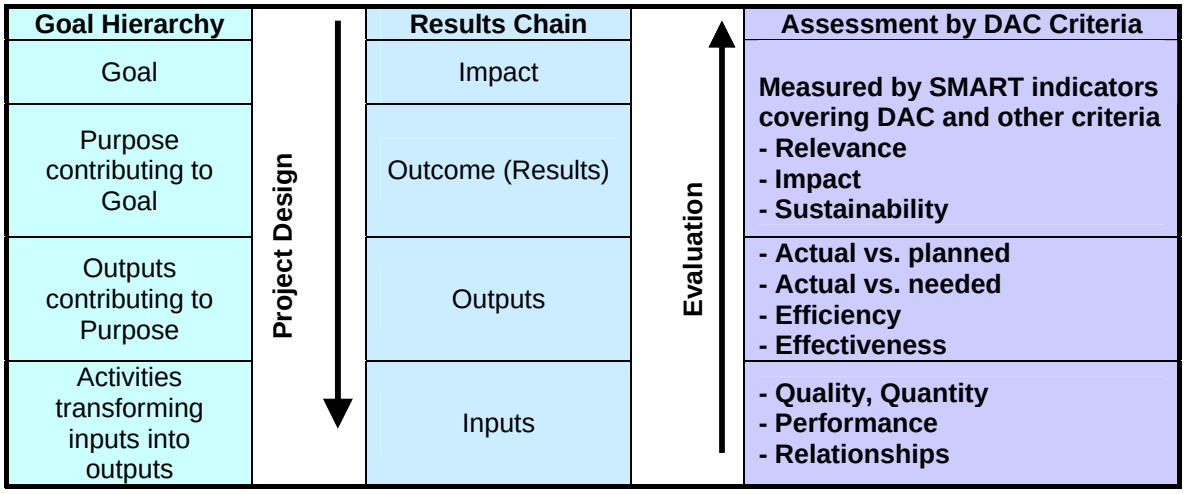

Fig. A2- 2: Results Hierarchy and DAC Criteria

A2.2.3 General Evaluation Methodology

A2.2.3.1 Preparation of Document Inventory

As a basis for getting a grasp on the universe of information available, a document inventory has been prepared, and the actual documents have been collected, and have been partially converted into electronic form to facilitate exchange between the team members.

A2.2.3.2 Desk Studies

A preliminary desk study has been done in order to determine which specific interventions it makes sense to include in the evaluation based on the premise that it is not possible, and indeed not useful to include all interventions. An attempt has been made to choose interventions that are complementary to the overall strategy. This is not to say that other, not included interventions are not complementary, but merely in order to set priorities and to evaluate what was believed to be more crucial, and more representative of the strategy followed. In the evaluation task, the chosen interventions were studied to the extent possible in-depth to extract the required information.

A2.2.3.3 Key Informant Interviews

A series of key informant interviews were held with a number of stakeholders and resource persons, in order to obtain information not available from documents. Stakeholders from the following groups were included.

- Personnel directly involved with energy policy development and implementation - management, administration, consulting, monitoring, and review of the energy programmes.
- Relevant staff in the involved line ministries and relevant institutions outside the sector
- Pertinent members of national government and municipal/village authorities;
- Women's organisations;
- NGO representatives;
- Program officers at the Norwegian embassies and other embassies, such as those of Sweden and Denmark, and other development agencies, such as the World Bank.

The purpose of the interviews was to hear informants' views on energy sector policy and performance in connection with Norwegian assistance to the energy sector. Relevant informants were also asked to assess the added value of the Norwegian assistance to institutional and operational structures.

Interviews were conducted using semi-structured checklists, where key themes and central questions were specified, and open questions were asked to probe the opinions of the respondent(s). The evaluation team made sure that a systematic approach was taken that all informants addressed the same questions, where relevant, but that there was no constraint on the answers provided. The experience was that this approach is the most efficient and effective in getting structured information as well as providing sufficient flexibility to obtain new and unexpected information.

Different draft checklists have been prepared for the different stakeholder groups to address the diverse range of topics to be covered by the evaluation, and to provide overall guidance for the field studies. To the extent possible, stakeholders received the checklists beforehand, so that preparation by the respondents was possible.

The evaluation team members generally split up when carrying out the interviews in order to maximize the number of interviews within the time available.

A2.3 Approaches to Fact Finding

The use of hypotheses as a guide to carrying out fact finding was specified by the terms of reference as the overall approach. While using this, the evaluation also found limitations when relying on hypotheses, and therefore used other methods, such as an inductive approach and Appreciative Enquiry to capture the less foreseeable.

A2.3.1 Deductive approach

Deductive reasoning works from the more general to the more specific. Sometimes this is informally called a "top-down" approach. One begins with thinking up a hypothesis about a topic of interest. Then this is narrowed down into more specific *hypotheses* that can be tested. Narrowing down even further is done by collecting *observations* to address the hypotheses that may follow. This ultimately leads to the test of the hypotheses with specific data, and a *confirmation* (or not) of the original assumption.

Formulating hypotheses up front presumes that the evaluation team had considerable knowledge of the available information prior to the evaluation. However, given the short inception period, such knowledge cannot possibly be complete, nor is there any certainty that all dispersed information sources have been fully identified. In this situation, hypotheses look for the obvious, but miss the unusual or unexpected.

A2.3.2 Inductive Approach

Almost as important will be an approach to fact-finding that is described as inductive. Inductive reasoning works the other way, moving from specific observations to broader generalizations and hypotheses. Informally, this may be called a "bottom up" approach. In inductive reasoning, we begin with specific observations and measures, begin to detect patterns and regularities, formulate some

tentative hypotheses that can be explored, and finally end up with some general conclusions. The two methods of reasoning have a very different "feel" to them. Inductive reasoning, by its very nature, is more open-ended and exploratory, especially at the beginning. Deductive reasoning is more narrow in nature and is concerned with testing or confirming hypotheses usually derived from a model or pre-existing framework for analyzing the topic. Most social research involves both inductive and deductive reasoning at some time in the process.

The evaluation used both deductive and inductive approaches. For the first, hypotheses are formulated, for the second the method used is *Appreciative Inquiry*[1]. The latter involves a open approach, without prejudice, preference or own agenda. Often, such openness brings out the unexpected and unusual.

A2.4 Benchmarking - Method of Systematic Comparison

Systematic comparison (benchmarking) is a method to improve objectivity by establishing a real standard for best practice within a group of development interventions. Historical expenditures are converted to USD using that year's PPP-exchange rate[2], which implies that all figures used are in USD-PPP. This is necessary for establishing a basis for comparison, but here it needs to be noted that the result is not perfect: PPP theory is valid in a perfect market, which in practice does not exist.

Benchmarking was used to compare

- efficiency: Intervention cost in relation to results
- effectiveness: Level of fulfilment of objectives including cross cutting issues
- relevance: Goal alignment with the needs of the target groups
- sustainability: The continuation of benefits after the end of the intervention
- impact: The sum effects of the intervention including direct and indirect, intended and unintended effects in the longer term

A2.4.1 General approach

To compare different types of interventions into different parts of the electricity value chain under different external conditions in different countries it is necessary to:

1. quantify the input (IN) to the intervention in a common currency normalised for cost level difference
2. quantify the output (OP) from the intervention; considering also the external conditions under which those output where generated; for example topography, climate, access, political conditions etc
3. quantify the historic and estimated future outcome (OC) of the intervention
4. quantify the impact (IP) of the intervention based on the outcome
5. assess the practises (PR) used in the intervention to assess effectiveness, relevance and quality
6. deal with fuzzy data and uncertainties to reach a balance between cost and accuracy. Only if all five variables can be estimated is it possible to objectively compare the interventions

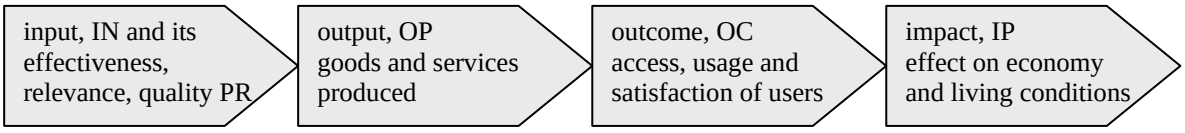

Fig. A2- 3: Benchmarking Variables

[1] The Thin Book of Appreciative Inquiry, Sue Annis Hammond, TB Publishing, ISBN 0966537319

[2] PPP (purchase power parity) refers to rates of changes of price levels, that is, inflation rates. USD-PPP is comparable over time and between different countries.

The input-output assessment quantifies the quality of the inputs (compared to best practice) and how financial resources are converted to goods and services. The outcome and impact assessment registers the intended effects of the delivered goods and services on the targeted groups and institutions.

For the different categories of intervention elements which are relevant within the power sector, we define the four variables to allow benchmarking (Fig. A2-3), within each category and across categories as defined in the following table. Definitions are chosen to achieve the maximum level of quantitative benchmarking with the available data.

Variable	Infrastructure development	Training	Capacity development
Input	The total cost of the project including cost of negative impacts, subsequent investments, O&M cost etc	The total cost of the training effort (not counting the working time lost by participants)	The total cost of all types of inputs (personnel, travel, expenses and goods)
Output	The volume of the infrastructure[3] and other services provided	The volume of the training services provided[4] (extent, level and participants)	The volume of support service[5] and other services provided (duration, level, extent)
Outcome	the electricity generation, - transport and - delivery facilitated by the intervention plus the transport volume facilitated by roads and agricultural production by irrigation	The capacity of trained people using their competence for the good of the country per year.	The volume of the useful changes/ improvements resulting from the support such as units reorganised, plans, regulations and legislation developed and adopted, institutes established, systems in operation, licences/ concessions awarded etc.
Impact	The net income from the intervention considering all costs and benefits over the infrastructure lifespan. Tangible and intangible social benefits to users and other stakeholders	The net benefits of improved quality and quantity of work over the competence lifespan	The benefits of the changes/ improvements due to reduced costs, improved services, or accelerated revenue

Fig. A2- 4: Definition of Variables across Intervention Types

Using these five variables (IN, PR, OP, OC, IP) we may calculate the following factors to be used in systematic comparison (benchmarking) of different interventions:

1. The efficiency-factor: VF = OP / IN – output / input tells us how financial resources were converted to goods and services; i.e. how much output was produced by one unit of input; for example km of line, MWh production, FTE of training per USD etc
2. The utilisation-factor: TF = OC / OP – outcome / output is a measure of how much outcome was produced by one unit of output; example: Amount of energy transported per unit of grid volume
3. The impact-factor: RF = IP / IN – impact / input tells us how much impact was generated by one unit of input (resources)
4. The normalised impact-factor: RFN = IP / OP = RF x TF – impact/ output tells us how much impact was generated by one unit of output; i.e. unit of economic benefit per output volume
5. The quality-factor: QF = Σ (practice $score_i$ vs objective $practice_i$) x $weight_i$ is the practice score added over all quality items I, is measure of effectiveness, relevance and what is/was believed to be best practice

[3] A weighted sum of goods, services and infrastructure using a common set of cost weights reflecting exogenous conditions as well as quality.

[4] The "output" from training and other CD is defined as the volume of training and of support provided (in terms of quantity and quality of personnel). The specific achievements are classified as outcome.

[5] The outcome and impact of training and other CD interventions can generally not be quantified

6. The sustainability factor: SF = OC_{t2} / OC_{t1} - shows the rate of change in outcome between the start and now

The applicability of these factors is summarised in the table below.

Factor	Meaning	Use
efficiency -factor VF	output volume in relation to normalised cost	comparison of cost efficiency among projects over all categories
utilisation-factor UF	outcome per unit of output	intermediary factor to explain difference in efficiency (infrastructure)
Impact-factor IF	impact in relation to project cost	comparison of end-user efficiency (infrastructure)
Normalised impact-factor IFN	impact in relation to the output volume	comparison of end-user efficiency for infrastructure projects compensating for exogenous conditions (infrastructure)
Quality-factor QF	weighted practice score compared to best practice at the time	comparison of effectiveness, relevance and quality of inputs over all projects
sustainability-factor SF	outcome now compared to outcome initially	comparison of sustainability (infrastructure)

Fig. A2- 5: Comparison Factors Used in Assessment

In the following sections the variables are discussed for the three elements of interventions in the power sector: Infrastructure development, training and capacity development.

A2.4.2 Infrastructure development

Inputs: To get the total costs of an intervention the life-cycle cost over the technical lifetime of the project is calculated from

- initial project investments
- later investments: Calculated based on estimated infrastructure development and average cost per unit
- depreciation of investment (if relevant)
- operation and maintenance cost (O&M cost) for the infrastructure volume built in the project (originally and later): Calculated on the basis of number of units in operation and average cost per unit.
- cost of grid losses if relevant (in the calculations here only net energy was considered)
- indirect cost of negative impacts such as environmental impacts

Outputs: The output is the total grid volume being built in the project; also taking exogenous conditions (cost drivers) into consideration.

Example: A project resulting in increased hydro power production of 2 MW, a high voltage distribution line of 20 km, one high to low voltage station, 5 km of low voltage distribution and 500 customer connections gets a grid volume (GV) as follows:

$GV = 2MW * W_{HP} + 20km * W_{RDl} + 1station * W_{RDs} + 5km * W_{LDls} + 500connections * W_{LDc}$

Where:

W_{HP} = cost weight per MW of hydro production capacity

W_{RDl} = cost weight per km of RD line

W_{RDs} = cost weight per LD station

W_{LDls} = cost weight per km of LD line

W_{LDc} = cost weight per LD customer connection

To differentiate the grid volume according to exogenous conditions we also split the MW of production capacity and the km of lines into easy, average and difficult conditions. The relative cost weights can then be considered universal. The weights are adapted to show output in USD.

If correct, the cost weights make it possible to calculate the "grid volume" provided and compare the output of projects with different structures and contents. The weights used are based on a comprehensive database of more than 100 European electricity providers. The weights are averages so the weight per MW of generation would reflect both small and big stations. Cost weights from the utilities involved in the Nepal and Mozambican interventions were also collected. The accuracy of these figures are questionable so a, sensitivity analysis is included to check the robustness of results,.

The indicator for efficiency is then calculated as actual total cost divided by grid volume (normalised unit cost). We may expect a big station to appear more efficient than a micro station, due to economies of scale.

Outcome: The outcome of an electricity infrastructure intervention is ultimately the amount of electricity coming to good use for the intended end users. To differentiate this a little further we may distinguish between the following levels of outcome:

1. The amount of electricity produced from a generating plant, E_G
2. The amount of electricity being transported over a grid at transmission, regional distribution and local distribution level, E_{TR}, E_{RD}, E_{LD},
3. The amount of electricity consumed by private household, E_{PP}
4. The amount of electricity consumed by businesses and public institutions, E_{PB}

A measure for the outcome can then be

- the electricity consumed
- the electricity transported
- the electricity produced
- the weighted sum of these; the weighting should be done according to the average cost of energy at the different grid levels.

In many cases grid development will continue after the project has finished to add new customers etc. The normalised cost of such development will be included in the life-cycle cost of the intervention (see input) as the total delivery/transport etc., is included in the outcome and impact. Additional outcomes include the transport volume of roads, additional farm production due to irrigation etc.

Impact: To quantify impacts the net benefits of the intervention are calculated and distributed among the beneficiaries. In infrastructure interventions the benefits of both the recipient country, the poor and the company owning and operating the infrastructure are considered.

- for the company, sales revenue from the electricity business and reduced cost of electricity generation
- for the country the grants given and the value added due to increased electricity use, irrigation and other infrastructure established as part of the intervention
- for the poor direct benefits are calculated based on their part of electricity consumption and indirect benefits as their part of GDP.

The income comes from electricity sold to the grid, electricity sold to local customers and in some cases wheeling income from electricity transport for third parties. In all cases the tariff revenue is the part of total revenue attributable to the infrastructure built in the project. Generation costs may be reduced when inefficient diesel plants or expensive imports are replaced by locally produced hydro power (see below).

The total revenue is found by adding annual revenue over the technical lifetime correcting for inflation.

The value added due to increased use of electricity is estimated as follows:

- savings for lighting and radio use: (price of kerosene per kWh, and cost of dry cell batteries – price of electricity per kWh) x GWhs used for lighting and other domestic use
- savings compared to diesel power: (price of diesel power per kWh – price of hydro power per kWh) x GWhs replaced diesel production
- indirect benefits
    - industry and agriculture – increased output and efficiency of production
    - education – increased earnings due to better education

- o health – reduced cost due to fewer fatalities
- o other (for example street lightning)

The following known benefits were not considered due to the lack of reliable data:

- benefits of other infrastructure elements such as roads, water and sewage
- indirect savings: Rural electrification also has indirect savings by facilitating economic development in rural areas and thereby stemming the costly movement of people to rural areas

The indirect benefits will be calculated as:

increased GWhs consumed x GDP per GWh x part of GDP-increase attributed to electricity.

The indirect economic benefits attributed to increased use of electricity was calculated by Mulder and Tembe in their 2006 study of rural electrification in Mozambique and found to be 0.98 USD-PPP/GWh. The present evaluation assumes that a conservative 30% of this GDP increase can be directly attributed to electricity, and the sensitivity of this value is analysed. This figure is also adapted to Nepal using the relationship between GDP per GWh for the two countries and giving the following estimates for indirect benefits:

Country	indirect economic benefits / electricity consumption mill USD/kWh	part indirect benefits assumed to be attributable to electricity	indirect benefit attributed to electricity use (mill USD-PPP/GWh)
Mozambique	0.98	30%	0.29
Nepal	0.97	30%	0.29

Fig. A2- 6: Indirect Benefit of Electrification

The value added due to irrigation schemes is calculated as hectares of irrigated area x added production value per hectare.

The total net economic benefits are found by addition over the technical lifetime taking changing prices of energy as well as inflation into consideration.

The economic benefits are attributed to beneficiaries as follows:

- government: the value of the infrastructure received (value of grant)
- owners of the infrastructure and the company: Net estimataswed result from el. sales as calculated above
- industry: value added from electricity use as calculated above
- non-poor:
  - + fraction of the value added for industry as calculated above
  - + fraction of consumption on regular domestic tariffs in local areas x sum savings on kerosene
  - + fraction of irrigated land for non-poor x sum added production value
  - + fraction of benefits from other initiatives for non-poor x sum added value from other initiatives
- poor:
  - + fraction of the value added for industry as calculated above
  - + fraction of consumption on social tariffs x sum savings on kerosene
  - + fraction of irrigated land for poor x sum added production value
  - + fraction of benefits from other initiatives for poor x sum added value from other initiatives
- the net impact for Norwegian stakeholders will only be evaluated in qualitative terms

The costs and benefits are summarised in figure Fig. A2-7.

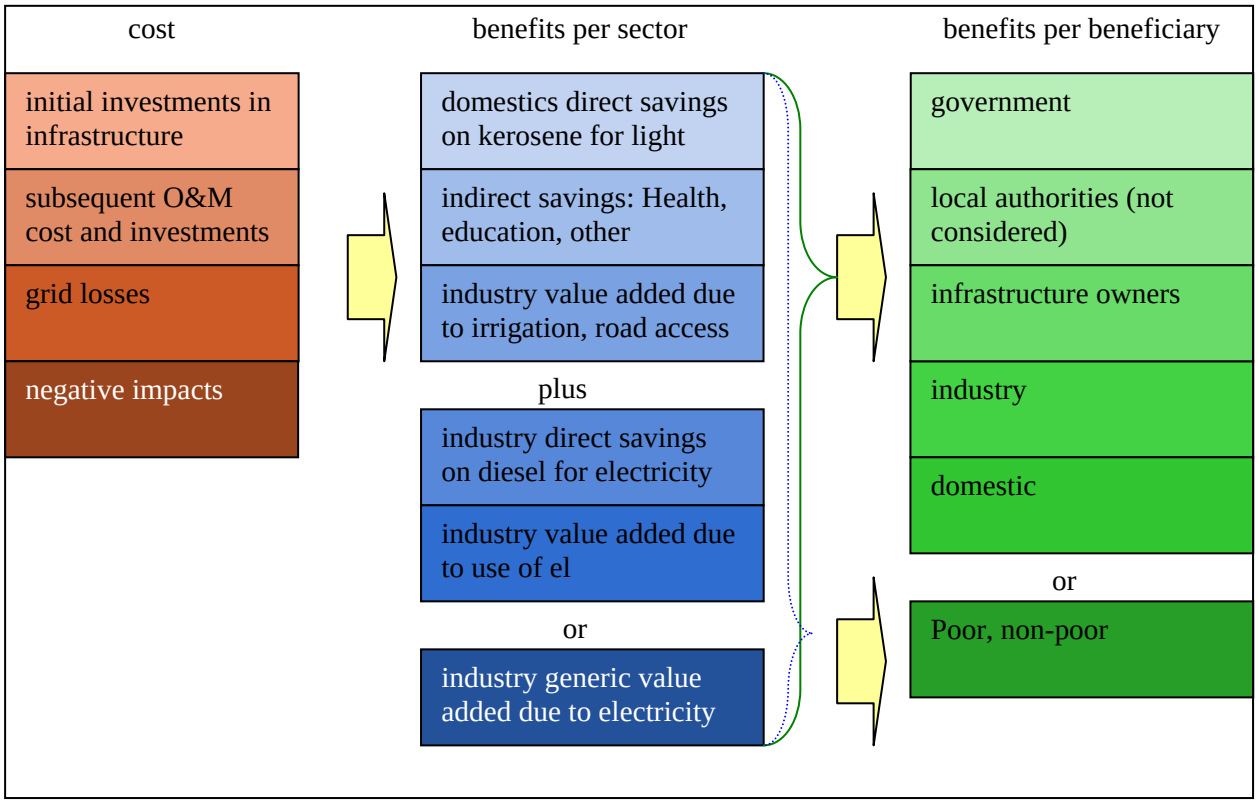

Fig. A2- 7: Generic Infrastructure Costs and Benefits

Sustainability: A measure of the sustainability/quality of the infrastructure established is the outcome as a function of time.

Quality: Quality is measured as the deviation of intervention practices from the best practice as seen today and the best practice as defined in project specifications. The latter is the value reflected throughout the evaluation, as it is deemed the fair procedure. The generic evaluation format for infrastructure interventions contains the following main parts:

- Selection, preparation and planning the intervention
- Implementation of the project
- Start-up and operation of the infrastructure provided
- Preparation for further development of the infrastructure
- Follow up work

A2.4.3 Training

Inputs: Generally the input is defined as the total effort measured in expenses for an intervention. The Norad input may however be only part of this. To get the total costs of an intervention project expenditures for each year is first converted to USD using that year's PPP-exchange rate and then brought up to the present by correcting for inflation.

Outputs: The output is the added competence resulting from the project, also taking exogenous conditions (cost drivers) into consideration. To illustrate this we consider a training project giving 10 engineers with a bachelor degree, a master in power engineering (the equivalent of two years of university level education): and two an additional phd (the equivalent of six years of university level education):

the output (OP) = $10 \times 2 \times W_{Um} + 2 \times 4 \times W_{Uphd}$

where:

W_{Um} = cost weight per year per person of University master level training

W_{Uphd} = cost weight per year per person of University phd level training

Outcome: The outcome of a training intervention is the added competence of companies, public administration institutions and training facilities/ schools aimed at the power sector. Possible indicators for measuring this may be:

1. The number of alumnae (FTEs) from the training effort working locally in the sector
2. The number of professionals of national origin working in the sector
3. The part of total value creation in the sector carried out by local companies

Outcomes of training are only assessed qualitatively.

Impact: Two types of impact are considered:

- savings compared to paying for the same education = the grant provided by Norad
- value added due to earlier increased revenue from the power business

The value added due to X years earlier revenue from the power business is found as:

- present value of current revenues - present value of revenues delayed by X years If the speed-up time can be assessed accurately the net income and value added may be attributed to beneficiaries as follows:
- government: the tax part of el. sector revenue difference calculated above
- owners of the infrastructure and the company: Dividend part of the difference calculated above
- industry: value added from earlier electricity use
- non-poor: fraction of el. consumption x difference in kerosene saving due to earlier use of electricity + fraction of industry added value
- non-poor: fraction of el. consumption x difference in kerosene saving due to earlier use of electricity + fraction of industry added value

If the speed-up time cannot be assessed accurately, the impact must be assessed in a more qualitative way. The calculation above may then be used to show how much impacts needs to be speeded up to cover expenses.

Sustainability: Sustainability of training interventions can only be assessed qualitatively

Quality: Quality is measured as the deviation of intervention practices from the best practice as seen today and the best practice as defined in project specifications. The generic evaluation format for training interventions contains the following main parts:

- Selection, preparation and planning the intervention
- Implementation of the project
- Preparation for further training efforts
- Follow up work

A2.4.4 Capacity Development

Inputs: Generally, the input consists of consultancy services and administration support. It is defined as the total effort measured in expenses for an intervention. The Norad input may however be only part of this. To get the total costs of an intervention project expenditures for each year is first converted to USD using that year's PPP-exchange rate and then brought up to the present by correcting for inflation.

Outputs: The output is the capacity added to the recipient organisation, also taking the quality of the capacity into consideration. To illustrate this we consider a project aimed at improving the capacity of a government agency using one person with 18 years of relevant experience in 1 year, two persons of 12 years of relevant experience in 0.5 years and one person of 4 years of relevant experience in 1 year.

the output, $OP = [(1 \times 1 \times W_T) + (0.5 \times 2 \times W_H) + [(1 \times 1 \times W_A)] \times W_C$

where:

W_T = is the cost weight of top level advisors (relevant experience of more than 15 years)

W_H = is the cost weight of high level advisors (relevant experience of more than 10 years)

W_A = is the cost weight of average level advisors (relevant experience of more than 3 years)

W_C = is the utilisation factor of the competence provided given the limitations of the receivers

Outcome: The outcome of a competence building intervention is the added competence of the target organisation(s). Possible indicators for measuring this may be:

1. The efficiency and effectiveness of the target organisation; for example estimated by practise comparison to similar organisations elsewhere or by qualitative assessment
2. Concrete results of the assistance such as new laws, regulations, systems, management procedures etc
3. other?

Outcomes of capacity development interventions are only assessed qualitatively.

Impact: Three types of income are considered:

- savings compared to buying assistance from international consultants
- value added due to earlier revenue from (parts of) the power business
- in some cases also reduced cost may also be measurable (like for Hydrolab)

Savings compared to international consultants is found as:

- number of days assistance provided x daily rate of international consultants

The value added due to X years earlier revenue from the power business is found as:

- present value of current revenues - present value of revenues delayed by X years

If the speed-up time can be assessed accurately the net income and value added may be attributed to beneficiaries as follows:

- government: the tax part of el. sector revenue difference calculated above
- owners of the infrastructure and the company: Dividend part of the difference calculated above
- industry: value added from earlier electricity use
- non-poor: fraction of el. consumption x difference in kerosene saving due to earlier use of electricity + fraction of industry added value
- non-poor: fraction of el. consumption x difference in kerosene saving due to earlier use of electricity + fraction of industry added value

If the speed-up time cannot be assessed accurately the impact must be assessed in a more qualitative way. The calculation above can then be used to show how much impacts needs to be speeded up to cover expenses.

Sustainability of institutional support interventions can only be assessed qualitatively.

Quality is measured as the deviation of intervention practices from the best practice as seen today and the best practice as defined in project specifications. The generic evaluation format for training interventions contains the following main parts:

- Selection, preparation and planning the intervention
- Implementation of the project
- Start up and operation based on the support effort
- Preparation for further support efforts
- Follow up work

A2.4.5 Combined interventions

Many interventions cover different aspects within all the three categories mentioned above, In such cases each dimension is assessed separately to derive at numerical figure for the five variables and are subsequently added to arrive at the sum figures for the intervention.

A2.4.6 Model Uncertainties and Weaknesses

The main uncertainty in the systematic comparison is the quality of input data. It is noted that delivered inputs are not commonly reported in sufficient detail, clarity and accuracy in projects.

For infrastructure projects model uncertainty is also related to the weights per unit of infrastructure and the part of observed GDP growth to be assigned to the increase in electricity consumption. Neither of these problems have significant impact on the results.

Benefits calculated are comparable if the PPP-conversion is correct. Efficiency is comparable if the weights assigned to the various output items (km, MW production capacity, FTE expert assistance etc) is correct. Effectiveness is comparable if the same judgement is applied to all projects. The latter is done by scoring and weighting a large number of criteria throughout the project cycle, where an error of judgment of individual scoring has little impact on the total score.

The cost weights used for estimating the O&M cost and investments after the (initial) project are estimated on the basis of international averages. Comparability between countries would be improved, if national cost weights for the countries in question could be used, but the units costs collected from the local utilities were sometimes lacking and some times not of sufficient quality The impact of using the same weights for Nepal and Mozambique could lead to an overestimation of cost in Nepal and an underestimation in Mozambique, but significant errors may only appear in the two projects where most of the life-time is in the future; namely Jhankre and AMG.

Sensitivity analysis carried out attempted to put uncertainties into perspective. Variables, such as cost weights, values of direct and indirect benefits, and household electricity consumption data were altered, and the effect on the results is observed. The overall result is that “best estimates” used may be considered valid, and results are fairly robust.

A weakness of the benchmarking process is that the number of projects being benchmarked here is small. To improve the model and its cost weights more projects (with better inputs) should be added. This would lead to the following improvements:

- better cost weights per unit of infrastructure → more accurate comparison of efficiency and impact
- better figures to estimate indirect economic impacts → more accurate comparison of impact for society and for the poor.

A2.4.7 The Data Entry and Calculation Tool

To enter all data, and carry out all calculations, a tool has been developed on the basis of Excel. This makes it possible to avoid manual calculation work, and results reports. The tool is used for the assessment of all interventions subject to evaluation, and it generates various results reports and comparison tables. It is not suitable for complete presentation in printed form, but is instead attached to the report in electronic form. For the purpose of illustration, some of the worksheets are pictured on following pages. The complete worksheets are made available in Excel in electronic form (DI.zip, ~1.9 MB).

To enter all data, and carry out all calculations, a tool has been developed on the basis of Excel. This makes it possible to avoid manual calculation work, and results reports. The tool is used for the assessment of all interventions subject to evaluation and with sufficient evidence. It generates various results reports and comparison graphs. It is not suitable for complete presentation in printed form, but is instead attached to the report in electronic form. For the purpose of illustration, some of the worksheets are pictured in the following.

1.1 Investment financial figures given in mill USD

item	sum	year 1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
reservoir		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
intake, waterway, headrace, tailrace		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
power station 1)		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
hydro generation	5.633	0.939	0.939	0.939	0.939	0.939	0 939	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
irrigation		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
grid	1.749	0.291	0.291	0.291	0.291	0.291	0 291	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
training		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
capacity development																
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
other		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SUM	7.382	1.23	1.23	1.23	1.23	1.23	1.23	0	0	0	0	0	0	0	0	0
part of cost foreign sources	1.043	0.174	0.174	0.174	0.174	0.174	0.174	0	0	0	0	0	0	0	0	0
part of cost from Norway	1.043	0.174	0.174	0.174	0.174	0.174	0.174	0	0	0	0	0	0	0	0	0
part of cost local taxes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
part of cost from NORAD	7.382															

1.3 O&M cost and later investments financial figures given in mill USD 2006

item	sum	year 1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
generation																
grid		0.234	0.237	0.241	0.244	0.247	0 251	0.254	0.258	0.261	0.265	0.268	0.271	0.275	0.278	0 282
roads																
other (later investments)		0.275	0.275	0.275	0.275	0.275	0 275	0.275	0.275	0.275	0.275	0.275	0.275	0.275	0.275	0 275
sum		0.509	0.512	0.516	0.519	0.523	0 526	0.53	0.533	0.536	0.54	0.543	0.547	0.55	0.554	0 557

Fig. A2- 8: Example of Data Entry: Initial Investment and O&M Cost

	best practice according to TOR and objectives (modify if different from best practice to-day)	correct target practice	importa nce weight	actual practice	score vs target practice
		not=0 some=1 quite=2 very=3	not=0 some=1 quite=2 very=3		
1	**Selection, preparation and planning**				0
1.1	the feasibility study included a realistic evaluation of the needs based on local knowledge	3	3	UMN has studied rural development in an overall context. HEP development has successively been chosen as a prerequisite	9
1.2	the feasibility study assessed also alternative ways to get he wanted results		0	there are no realistic alternatives	0
1.3	the feasibility study gave a realistic estimate for inputs	1	2	optimistic progress schedule, despite lack of geological data or sample drilling	2
1.4	the feasibility study / needs assessment gave a realistic estimate of the full impacts	2	2	full impact estimation was done to a limited degree	4
1.5	the objectives and goals were clearly defined in line with diagnosis	2	1	objectives definition is too broad to provide relevant indicators	2
1.6	the time budget was realistic	3	1	time overrun could not have been foreseen	3
1.7	the resource budget was realistic	1	1	it was severely underestimated	1
1.8	the plans dealt with measures and mitigations wi h a view to he full infrastructure value chain	3	3	the project addressed overall development and irrigation for agricultural development	9
1.14	.....				
	total score		**13**		2.3076923
2	**Implementation**				
2.1	international bidding for tenders whenever relevant was conflicting with other objectives		0	the concept was to use companies built up for the purpose, and second-hand machinery	0
2.2	tenders were selected with a view to total discounted cost as well as sustainability	2	0	not important due to 2.1	0
2.3	implementation was organised to assure sustainability by local resources	3	1	Committees were set up and met to coordinate hroughout project	3
2.4	on-the-job training was done to utilise local personnel	3	2	it was an important part of the concept	6
2.5	technology was selected with a view to total discounted cost as well as sustainability	2	2	by reducing initial investments and building the capacity to maintain and operate	4
2.6	gender issues were not in focus at he time	3	1	the focus was on development of the local community in general	3
2.7	locally produced components were used whenever possible	3	2	it was an important part of the concept	6
2.23	.....				
	total score		**8**		2.75
3	**From implementation to operation**				
3.1	competent local personnel was selected at all levels	3	2	it was an important part of the concept	6
3.2	coaching efforts were made to assure the quality of operation	3	2	it was an important part of the concept	6
3.3	the work of intructors was of high quality'	2	1	it was an important part of the concept, but there is no specific evidence	2
3.4	the input from twining personnel was of high quality		0	no twinning	0
3.6	.....				
	total score		**5**		2.8
4	**Preparation for further development**				
4.1	a realistic plan for the operation and further development was in place	3	3	this is reflected in cept papers and income calculations done	9
4.2	a competent organisation took over responsibilities	3	3	BPC remained responsible as the only competent organisation existing at that time	9
4.3	the activities, results and lessons learned for all stakeholders were reported	3	1	reporting did not follow a standard format, not considered important at the time	3
4.4	the final report documented the full inputs, outputs and forecasts for outcomes	2	1	final report contained what was considered important at the time	2
4.5	all disbursements were documented and audited according to NORAD rules	3	1	it was properly reported. Audit reports not seen	3
	total score		**9**		2.8888889
5	**Follow up work**				
5.1	The operation and further development of the outputs was monitored in the years after the project ended	3	1	Various further studies were done, i.e. Nafziger 1994	3
5.2.	proper action was taken to assure sustainability of outputs when necessary	3	2	it was the common BPC operation routine	6
	total score		**3**		3
	Overall total score		**38**		2.6578947

The example is incomplete

Fig. A2- 9: Example - Input Form for Quality Scores

Illustration: The value of outputs is entered in the column of the year in which it is produced.

2 Output

item	1992	year 1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
reservoir - Mm3	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
power station - installed MW	5.1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
irrigation - hectares	282	283	284	286	287	288	289	290	292	293	294	295	296	298	299	300
TR lines, difficult - km 1)																
TR lines, average - km																
TR lines, easy - km																
TR/RT stations																
RT lines, difficult - km																
RT lines, average - km																
RT lines, easy - km																
RT/RD stations																
RD lines, difficult - km	180	197	213	230	247	264	280	297	314	331	347	364	381	398	414	431
RD lines, average - km																
RD lines, easy - km																
RD/LD stations	175	187	200	212	224	237	249	261	274	286	298	311	323	335	348	360
LD lines 2), difficult - km																
LD lines, average - km	62	66	71	75	79	84	88	92	97	101	105	110	114	118	123	127
LD lines, easy - km																
customer conn. 3) metered domestic	1,356	1832	2308	2784	3260	3736	4212	4688	5163	5639	6115	6591	7067	7543	8019	8495
customer conn. non-meter domestic	8,230	8427	8624	8822	9019	9216	9413	9610	9808	10005	10202	10399	10596	10794	10991	11188
customer connections industry	74	85	97	108	120	131	143	154	166	177	189	200	212	223	235	246
		0			0											
new roads - km		0			0											
upgraded roads - km		0			0											
basic training of workers - FTE 4)	3.5															
training at skilled level - FTE	4															
training at bachelor level - FTE	1															
training at master level - FTE																
training at graduate level - FTE																
top level support provided 5) - FTE																
high level support provided - FTE																
average level support provided - FTE																
description	The output includes the hydro power station, distribution grid, connection to the transmission grid and customer connections. Also 300 hectares of agricultural land have received irrigation and local personnel were trained both for construction and operation tasks.															

1) conditions of topography, congestion, forest, strong winds, corrosion and lack of roads causing significant increased cost of construction and mainte
- difficult = difficult access, significant impact on design & material (cost level 120-160%)
- average = good access, some restrictions on design (cost level 80-120%)
- easy = access by road, easy soil and no restrictions to design (cost level 50-80%)
2) LD lines includes all lines up to the customer delivery point
3) Customer connection includes only the meter and the connection point between the LD grid and the customers internal grid
4) FTE = full time equivalents, example: If 4 persons were trained for 3 months the resulting traning effort is 1 FTE
5) top level support = professional with more than 15 years of relevant, hands on experience, high level 10 years, average level 3 years

Fig. A2- 10: Example Input form – Quantity and Type of Outputs Achieved

A2.4.8 Approach to Rural Field Studies

Data gathering in the rural areas required a number of additional methods and tools. The studies attempted to measure four main aspects of the outcomes and impact of the interventions: (1) use and importance of electricity in households; (2) importance of electricity for industry, trade and commerce, and other economic activities; (3) changes in the public services due to electricity (including other changes at the community level) and (4) electricity's interrelation with other, accompanying interventions that have taken place. This fourth aspect involved to assess under what conditions the improved access to electricity has had particularly beneficial (or limited) effects. A number of tools and methods were employed to collect data at the rural locations of the projects in question. The combination of all tools, instruments and methods, attempted to achieve collection of the needed data with acceptable reliability, include sources and methods for verifying the data, compare with situations without intervention such as before and without electricity, use existing baseline data and generally standardise results to make a degree of comparison feasible between Nepal and Mozambique, including quantitative and qualitative aspects.

A2.4.8.1 Social Mapping and Wealth Ranking

Social mapping and wealth ranking are standard methods in participatory research. This was used in the field sites in Nepal. In Mozambique, it was found more appropriate to do the ranking by observation and individual discussion of the interviewers of type/condition of housing and assets owned, with the following purposes:

- To enter into a dialogue with formal and informal leaders and resource persons in the community, to explain the purpose of the fieldwork and obtain their cooperation;
- To obtain an overview of the communities, including locations of houses with and without electricity, schools, institutions, irrigation, etc.
- To obtain a rough overview of poverty and wealth distribution in the community.

S*ocial mapping* was done with small groups of formal and/or informal leaders. Information obtained was then used to add significant issues to the semi-structured interview guides used for focus group discussions, particularly in Nepal, where pertinent issues were further investigated from the point of view of different groups (e.g. women).

Wealth ranking was then done, using the social map drawn by the group. Group members were asked to define what characterizes the poorest members of the community, and then point out where they live, and which caste (Nepal) they belong to. The same was then done for the richest members. For the remaining community members, the group was asked how they would divide them – e.g. into two or three more groups, based on criteria that were discussed and agreed upon.

Note: Wealth ranking is based on the understanding, that poverty is a relative concept, and depends on local perceptions valid at the time when the exercise is done. Such perceptions may change over time. Therefore, it is not useful to define all criteria used in wealth ranking for the purpose of later use. Rather, sub-sequent studies of the wealth or poverty status would need to undertake a new wealth ranking exercise, including the definition of criteria used at the time of the future study.

A2.4.8.2 Quantitative Method

Household surveys were used in both countries to address questions with relevance to energy and electricity use, and the impact of this, and indirect effects on other aspects of everyday life (such as time saved on water collection, firewood collection, etc). A standard household questionnaire has been developed, that also included open questions. This was adapted to the relevant local situation, and translated into the local language where needed, before use in the field study.

A2.4.8.3 Qualitative Methods

Focus Group Discussions were held with a number of different stakeholders in the communities, to give in-depth information to supplement household survey data. Draft checklists have been developed that were used in the field.

Observations: Formal (*using checklists*) and informal (*being generally aware of everything that is related to the evaluation)* observations were important tools for the researchers, mainly to classify the quality of life of the household members. A lot of important information, including the understanding of the context where the intervention is implemented, was yielded from this method.

Transect walks: This method enables the interviewer to talk to a respondent while walking through his or her village, pointing out and discussing important issues relating to how the electricity lines are distributed, who benefits the most, who misses out on the benefits of the interventions, and – why it is like this.

With regard to emery-use related data, i.e. consumption data of various forms of energy, the concept of household energy balance is applied. This allows for checking consumption data by analysing if any reduction or increase in one form of energy is compensated by another form of energy, while causes for a resulting imbalance are analysed and discussed.

A2.5 Hypotheses

A number of hypotheses have been prepared in the inception phase, as a guide to fact finding. Some of these have been discarded as inappropriate or irrelevant in the course of the evaluation. Hypotheses actually used are grouped and listed below.

A2.5.1 Hypotheses on Assistance Strategy

- Mozambique: Norwegian assistance has been very focused and limited to a small number of institutions, for reasons to be identified and assessed.
- Nepal: Norwegian assistance has developed from a narrow hydropower investment approach to a strategic and broad sector approach. This has resulted in extraordinary and broad sector competence in the public as well as the private sector.

A2.5.2 Hypotheses Related to Key Evaluation Questions

Key evaluation question 1: What have been the results of Norwegian assistance to the partner country/ region, its power sector, institutions and participating staff and when applicable to local communities, industries and households? The result assessments should include the effects of technical and financial assistance in production and transmission/distribution, on the development of industry and trade on local, national or regional level, effects related to import/export of energy. In addition to the effects on economic development and poverty reduction the result analysis should cover capacity/competence building effects and effects on resource and environmental management, including the cost-effectiveness.

Hypotheses:

- If Norwegian assistance had not been provided, then a significant number of personnel would not have received higher level training of relevance
- Norwegian grant assistance to transmission investments has been significant in bringing more power to rural areas
- Sector assistance to widening the knowledge base (by conducting studies and research) may have led to more informed decisions, and avoidance of costly mistakes.

Key evaluation question 2: What have been the reasons for success or failures, and the risks involved? The assessments should cover the full period from the planning of an intervention and up to the end of 2006. Changes that take place over time can be related to other factors than the intervention which are studied. The evaluation team should check if other factors or processes than the Norwegian assistance may have influenced the results, including changes in partner countries policies, institutional arrangements, economic situation and ability to absorb assistance.

Hypotheses:

- Limited capacity to absorb intervention inputs hindered institutional development.
- Policy changes outside the sphere of influence of Norway limited the effectiveness of its assistance
- Successful interventions depend to a significant extent on a flexible project management who adjust their plans (and document such changes) according to shifting policies, sector restructuring and local initiatives.

Key evaluation question 3: What have been the challenges and opportunities for Norwegian assistance when other donors or commercial interests have been involved directly or indirectly? In some cases private Norwegian stakeholders have been involved together with Norwegian public interests or different public institutions have been partners. The results of such interactions processes should be covered with focus on how the results have been influenced by processes as harmonization and privatization.

Hypotheses:

- If Norway had not provided assistance to power related legislation then private-sector participation in sector growth would still not be possible.
- Norway's absorption of the up-front risk (i.e., financing the feasibility studies) leveraged private capital, ultimately resulting in increased generating capacity.

Key evaluation question 4: What have been the results for Norwegian Stakeholders? The evaluation should also analyse the results for the Norwegian stakeholders and cover not only the Norwegian institutions directly involved in the partnerships, but clarify the economic linkages to institutions, companies or individual consultants in other steps in the results-chain.

Hypotheses:

- Norwegian stakeholders benefited considerably from Norwegian development assistance by maintaining a sizable business volume.
- Norwegian stakeholders benefited considerably from Norwegian development assistance by having been enabled to maintain cutting edge know-how.

Key evaluation question 8: How and to what degree has the assistance covered cross-cutting issues?

- Gender issues have been included in the planning, implementation (capacity building and infrastructure) and evaluations at all levels.
- Environmental assessments have been systematically carried out and such concerns have been considered throughout the planning and implementation.

A2.5.3 Impact on Poverty

Hypotheses:

- Electricity helps in reducing poverty by reducing the vulnerabilities and increasing the ability to cope with future events (like drought, incidence of disease, illness, personal injury, and emergencies due to natural calamities) and market failures (like inability to store the products and medicines, processing, communicating market information).
- Electricity can be a catalyst for empowering the socially excluded groups to develop their organizations. Social capital thus developed can be useful to reduce poverty and associated vulnerability.
- Electricity has helped in reducing poverty because it has helped in increasing the production and income (as measured in production of grain, income and employment status, industrial production) and in providing more safety and security.
- On balance, considering all impacts, the provision of hydropower energy is positive for the poor, when measured against the alternatives it replaces.

A2.5.4 Affordability Issues

Testing of "affordability hypotheses" was a part of the field studies, as suggested in the state-of-the-art study. Even though the statements were formulated in the context of Africa, use of them was made also in Nepal, to permit comparative analysis.

Hypotheses:

- Consumers who are connected to a network do not have a major affordability problem,
- Users who are not connected often do have a major affordability problem;
- The majority of the population is not connected to a utilities network;
- In rural areas, very few are connected;
- The population not connected tends to be among the poorest income groups, facing an affordability problem for utility services;

- The main stumbling block to affordability is not the initial connection charge, but the monthly bills.
- Poorer households electricity use is mainly an affordability issue.
- There are many factors/events affecting the (income) poverty status of a family that are more significant than access to electricity. Therefore, if such factors/events occur, attribution of change to electricity is unlikely to be possible.
- In Africa the idea of relying on private operators concerned with full cost recovery without any subsidy component would be incompatible with network coverage expansion efforts.

A2.5.5 Capacity Development Hypotheses

- The provision of training alone may be an insufficient input to institutional capacity development.
- Assessing capacity and capacity development needs and available resources are key to effective capacity development efforts.
- When developing institutional capacity, motivational aspects of interventions may be crucial
- Local ownership and leadership of the capacity development agenda may be key to its success.

A2.6 Methods for Assessment of Macro and Meso Economic Impacts

In addition to the direct effects which was analysed project by project. the *overall* effects of increased electricity supply were assessed at the macro and meso level. A main problem for this assessment was that available methods and statistics did not permit an analysis that determined the specific effect of *Norwegian* aid to the sector. The approach taken was first to assess, the contribution of *overall* aid to the sector and then approximate the effect of the Norwegian assistance by looking at the Norwegian share of the total of interventions to the sector.

For Mozambique, the evaluation considered **macro and meso economic impacts** in several ways, however constrained by data availability. The macro (national level) and meso (sectoral) impacts were analysed in four ways a) by considering the power sector's share of total GDP based on national accounts data, b) the possibly dampening effect on imports of petroleum products from increased electricity production. This was done simply by comparing indices for constant price GDP, electric power consumption (kWh) and demand for oil and oil products (volume) over the years 1980 to 1998 and c) by considering the relation between overall economic growth and the growth of the power sector. Because of the limited data availability it was not possible to do a quantitative analysis of this relation and tentative conclusions were reached through examining other quantitative analyses of that relation for African and Asian countries, partly drawn from the Norad state of the art study. d) some indication of the importance of electricity supply was also arrived at from a company level study carried out by the Confederation of Mozambican Business Associations in 2006, which indicated businessmen's perceived importance of electricity supply for their business.

For the analysis of Socio Economic Impact at the national level in Mozambique we used a regression analysis of factors influencing household consumption (including availability of electricity. A straightforward OLS (Ordinary Least Squares) regression was used with the variables that were believed could be of relevance for household consumption to make as sure as possible that variation in incomes was due to availability / non-availability of electricity and not to other factors.

The dataset came from the official Mozambican household income and expenditure survey 2002/2003 conducted by Mozambican statistical bureau (INE). and was formerly used by DNEAP for income / poverty analyses). Quality control of the data has been undertaken by the INE and sets some limits to the further controls that may be made. The sampling for the official household income and expenditure survey was done to enable significant statistical conclusions only at the **provincial** level. Thus, analysis of **community** level data would not give significant results and there would be no advantage (in terms of statistically significant conclusions) to analyse and present these data on the 'community level'.

In order to measure the effect of electricity supply on a most important public service sector, education, a dataset by province containing pass rates, teacher quality indicators, class size data and enrolment for the years 1980 to 2006 was obtained from the Ministry of Education. The analytical method used was a standard panel data regression based on these variables and 'time' with pass rates as a dependent variable and including an electricity 'dummy' indicating in what year major parts of the given province was hooked up to the national grid.

For Nepal, no specific macro economic assessment method was applied, due to a lack of reliable data, and the fact that evaluated interventions are believed to have a limited discernible impact at the national level.

A2.7 Quality Assurance

The evaluation applied the DAC Evaluation Standards[6], which assure that:

- The scope of the evaluation itself is clearly reflected in the report. In other words: What is being evaluated is specified and set out in sufficient detail.
- The evaluation report describes the intervention logic of the assistance provided.
- QA will check that the DAC evaluation criteria efficiency, effectiveness, relevance, sustainability and impact, are used consistently for all objects of evaluation, and that any deviation from this is explained.
- Various contextual quality elements will be assured:
  - Development and policy context is described;
  - Institutional context and institutional arrangements are made transparent;
  - The socio-political context, and its influence and the evaluated intervention is addressed.
  - Related to data used in regression analysis (macro economic impact, Mozambique), the evaluation believes to have use the best data available, and is not in a position to perform any specific quality assurance process.

Further QA elements and notes are:

- Key data used in benchmarking have been subjected to sensitivity analysis, and an assessment is made of the robustness of results on that basis.
- Triangulation of data: This was used to improve the validity of findings by cross-checking information from two sources with a third source, usually an informant.
- The evaluation methodology first level of quality assurance has been through the interaction with the client and reference group in the discussions on the Inception Report and later, by obtaining comments to the Draft Report. Comments were analysed, and where found relevant, addressed in the Final Report.
- An internal quality assurance team has been assigned, which discussed with the core team at various points, read and commented preliminary drafts, and addressed issues becoming apparent. Scanteam has applied its internally established procedures in this.
- Language washing and verifying that the various formal requirements of the report are in place has been ensured by Scanteam management.
- The terms of reference for the evaluation have been elaborated in accordance with the DAC standard. By complying with TOR requirements, all quality aspects not specifically mentioned are observed.
- The reliability of the regression results (annex 4: macro economic assessment) are mostly dependent on the data used and the method used. Clearly the data availability sets some limits, while the method used is straightforward and standard, and therefore fairly reliable.

6 DAC Evaluation Quality Standards, March 2006