8 Estimating the baseline

This chapter is relevant for users who are following the quantitative approach to impact assessment. Quantifying impacts by defining changes relative to a baseline scenario may not always be necessary to meet the stated objectives of the assessment. Users can assess impacts qualitatively (in Chapter 7) or track trends in key indicators over time (in Chapter 12). Attributing impacts to specific policies relative to a baseline scenario is valuable since it enables an understanding of how effective policies are, relative to what would have happened in the absence of the policy. This information enables users to meet a wider range of objectives, outlined in Chapter 2, such as improving policy design, selection and implementation, and determining whether policies have been effective.

The baseline scenario represents the events or conditions that would most likely occur in the absence of the policy being assessed. Properly estimating baseline values is a critical step, since it has a direct effect on the estimated impacts of the policy. In this chapter, users estimate baseline values for each indicator included in the quantitative assessment boundary. This chapter is relevant to both ex-ante and ex-post assessment, and provides guidance on estimating ex-ante and ex-post baseline scenarios.

Checklist of key recommendations

- Include all significant impacts in the quantitative assessment boundary, where feasible
- Define one or more appropriate indicators for each impact category included in the quantitative assessment boundary
- Define the assessment period
- Define a baseline scenario that represents the conditions most likely to occur in the absence of the policy for each indicator included in the assessment boundary
- Estimate baseline values over the assessment period for each indicator included in the assessment boundary
- Separately estimate baseline values for different groups in society, where relevant

8.1 Define the quantitative assessment boundary and period

The quantitative assessment boundary defines the scope of the quantitative assessment in terms of the range of dimensions, impact categories, specific impacts and indicators that are included in the quantitative assessment and estimated. Not all specific impacts identified in Chapter 6 need to be estimated. It is a *key recommendation* to include all significant impacts in the quantitative assessment boundary, where feasible.

FIGURE **8.1**

Overview of steps in the chapter

Define the quantitative assessment boundary and period (Section 8.1) → Choose assessment method for each indicator (Section 8.2) → Define the baseline scenario and estimate baseline values for each indicator (Section 8.3)

8.1.1 Choose which specific impacts to quantify

Users should determine which specific impacts to include in the quantitative assessment boundary and estimate, based on:

- the significance of each impact, as determined in Section 7.3, based on a combination of likelihood and magnitude
- the feasibility of estimating each impact.

Feasibility may depend on data availability, technical capacity and resources available to estimate impacts, or other factors. If it is not feasible to estimate certain impacts, the decision to exclude them from the quantitative assessment boundary should be explained and justified. Table 7.5 provides a template that can be used to report whether it is feasible to quantify each significant impact, whether the impact is included in the quantitative assessment boundary and, if it is not included, a justification for exclusion. The example in Table 7.5 shows that, out of many identified impacts, 10 specific impacts are included in the quantitative assessment boundary. This short list of specific impacts is presented in Table 8.1.

In general, users should not exclude any impacts from the quantitative assessment boundary that would compromise the relevance of the overall assessment. Users should ensure that the assessment appropriately reflects the impacts resulting from the policy and that it serves the decision-making needs of users of the assessment report. Exclusions may lead to misleading and biased results that do not accurately represent the impacts of the policy. Where possible, instead of excluding significant impacts, users should use simplified or less rigorous estimation methods to approximate each impact, or use proxy data to fill data gaps. Any significant impacts that are not quantified should be described qualitatively.

8.1.2 Choose which indicators to quantify

It is a *key recommendation* to define one or more appropriate indicators for each impact category included in the quantitative assessment boundary. The indicator(s) will be quantified in the baseline scenario and policy scenario to estimate the impact of the policy. Each indicator will generally require a different assessment method.

Section 5.2 introduces indicators and provides examples in Table 5.5. The initial indicators chosen in Chapter 5 may need to be revisited based on the outcomes of Chapters 6 and 7, since the choice of indicators should be informed by which specific impacts are significant and included in the quantitative assessment boundary.

Users can define one or more indicators for each impact category. For example, within the impact category of air quality, a user may estimate the impact of the policy on multiple indicators, such as particulate matter ($PM_{2.5}$, PM_{10}), SO_2 and nitrogen oxides (NO_x).

Some indicators for a given impact category are likely to be more feasible to quantify than others. Users should choose indicators for which it is possible to collect data and quantify impacts. If it is not possible to quantify a particular indicator, users should either select a different indicator for the same impact category or qualitatively assess any indicators and specific impacts that cannot be quantified.

The indicators selected in this step will be estimated in the baseline and policy scenarios (in Chapters 8–10), and monitored over time (Chapter 12). Table 8.1 presents indicators selected for a solar PV incentive policy.

8.1.3 Define the assessment period

It is a *key recommendation* to define the assessment period. In general, the assessment period for a quantitative assessment should be the same as the period defined in Section 7.2 for the qualitative assessment. In some cases, users may want to choose a different assessment period for the quantitative assessment, based on objectives, data availability or other reasons.

Box 8.1 provides an example from an assessment in Mexico of how the choice of assessment period can have a significant impact on the overall assessment results.

TABLE **8.1**

Example of defining the quantitative assessment boundary for a solar PV incentive policy

Chapter 5	Chapter 6 (Identify specific impacts)	Chapter 8 (Define the quantitative assessment boundary)		
Impact categories included in the assessment	Specific impacts included in the quantitative assessment boundary	Indicators to quantify	Feasible to quantify?	Included in the quantitative assessment boundary?
Climate change mitigation	Reduced GHG emissions from grid-connected fossil fuel–based power plants	GHG emissions (tCO_2e/year)	Yes	Yes
Air quality/ health impacts of air pollution	Reduced air pollution from grid-connected fossil fuel–based power plants	Emissions of $PM_{2.5}$, PM_{10}, SO_2 and NO_x (t/year); number of deaths due to air pollution	Yes	Yes
Energy	Increased renewable energy generation from increased solar generation	Solar installed capacity (MW); % solar of total installed capacity; % solar of total installed capacity of renewable energy sources	Yes	Yes
Access to clean, affordable and reliable energy	Increased access to clean, affordable and reliable electricity	Number of houses/ buildings/facilities with access to clean energy resulting from the policy	Yes	Yes
Capacity, skills and knowledge development	Increase in training for skilled workers in solar-relevant sectors	Number of new skilled trainees and workers on the ground	Yes	Yes
Jobs	Increased jobs in the solar installation, operations and maintenance sectors	Number of new jobs resulting from the policy	Yes	Yes
	Increased jobs in the solar panel manufacturing sector	Number of new jobs resulting from the policy	Yes	Yes
	Decreased jobs in fossil fuel sectors	Number of jobs reduced resulting from the policy	Yes	Yes
Income	Increased income for households, institutions and other organizations due to reduction in energy costs	Savings in annual electric bills ($/year)	Yes	Yes
Energy Independence	Increased energy independence from reduced imports of fossil fuels	Reduction in coal imports from the policy (t/year)	Yes	Yes

Abbreviations: MW, megawatt; t, tonne; tCO_2e, tonne of carbon dioxide equivalent

BOX **8.1**

Selection of assessment periods and how assessment results vary over different time periods for a policy in Mexico

A researcher at Aalto University assessed the sustainable development impacts of two climate actions in public buildings in Mexico: installing PV panels and changing fluorescent lamps to LED lamps. These actions are part of the Carbon Management Plan of the Mexican state of Jalisco. The assessment illustrates how the impacts of a policy can change over time. The net impacts of the policy may not be linear, and the nature of impacts could change from negative to positive or vice versa under different assessment periods. In such cases, it is important to assess and report both short- and long-term impacts.

Selected results of the assessment are shown in Table 8.2, and Figure 8.2 illustrates the trends in the policy's net impact over time for three selected impact categories. The assessment found that the nature and scale of impacts across short- and long-term time horizons, measured as the percentage of cumulative net impact compared with the baseline scenario, remain stable for some impact categories (GHG emissions, depletion of fossil resources, and air quality). For others (mineral resources depletion), the scale of the impact changes dramatically over time. For impact categories such as human toxicity and water ecotoxicity, the net impact changes from negative to positive when the assessment period is expanded from 5 years to 17 years. The policy had nearly all positive environmental impacts using a longer assessment period, compared with mixed results using a short assessment period.

TABLE **8.2**

Summary of environmental impacts resulting from LED lamp replacement policy over 5- and 17-year assessment periods

		Cumulative impact over 5 years				Cumulative impact over 17 years			
Impact category	Unit	Baseline scenario	Policy scenario	Net impact	% net impact	Baseline scenario	Policy scenario	Net impact	% net impact
GHG emissions	tCO_2e	239	146	**Reduction of 93**	**–39**	724	409	**Reduction of 315**	**–43**
Depletion of mineral resources	kg Cu eq	66	243	**Increase of 177**	**267**	288	315	**Increase of 27**	**9**
Depletion of fossil resources	kg oil eq	74,990	46,104	**Reduction of 28,886**	**–39**	226,106	128,755	**Reduction of 97,351**	**–43**
Freshwater consumption	m^3	531	467	**Reduction of 64**	**–12**	1,851	1,170	**Reduction of 681**	**–37**
Air quality	DALY	0.24	0.16	**Reduction of 0.08**	**–34**	0.64	0.37	**Reduction of 0.27**	**–42**
Human toxicity	DALY	0.025	0.029	**Increase of 0.004**	**15**	0.088	0.061	**Reduction of 0.027**	**–30**
Water ecotoxicity	kg 1,4-DCB	6,255	7,190	**Increase of 936**	**15**	24,739	18,549	**Reduction of 6,190**	**–25**

Abbreviations: DALY, disability-adjusted life year; kg 1,4-DCB, kilograms of 1,4-dichlorobenzene; kg Cu eq, kilograms of copper equivalent; kg oil eq, kilograms of oil equivalent

Note: Positive (good) results are shown in black and negative (bad) results are shown in red.

BOX **8.1**, continued

Selection of assessment periods and how assessment results vary over different time periods for a policy in Mexico

FIGURE **8.2**

Cumulative impact of the policy on depletion of fossil fuel resources, freshwater consumption and human toxicity

8.2 Choose assessment method for each indicator

Estimating the impacts of a policy involves comparing the outcome of the policy with an estimate of what would most likely have happened in the absence of that policy.

The impact of a policy can be quantified in three ways:

- **Scenario method** – comparison of a baseline scenario with a policy scenario for the same group or region, where separate baseline and policy scenarios are defined and estimated
- **Deemed estimates method** – a simplified approach to the scenario method, where the change resulting from a policy is estimated directly without separately defining and estimating baseline and policy scenarios
- **Comparison group method** – comparison of one group or region affected by the policy with an equivalent group or region not affected by the policy.

Ex-ante assessments can only use the scenario method or deemed estimates method. Ex-post assessments can use any method. If appropriate, users can use a different assessment method for each indicator included in the assessment boundary. The choice of method should depend on which would yield the most accurate results for a given indicator in the context of the assessment objectives, and the data and resources available.

8.2.1 Scenario method

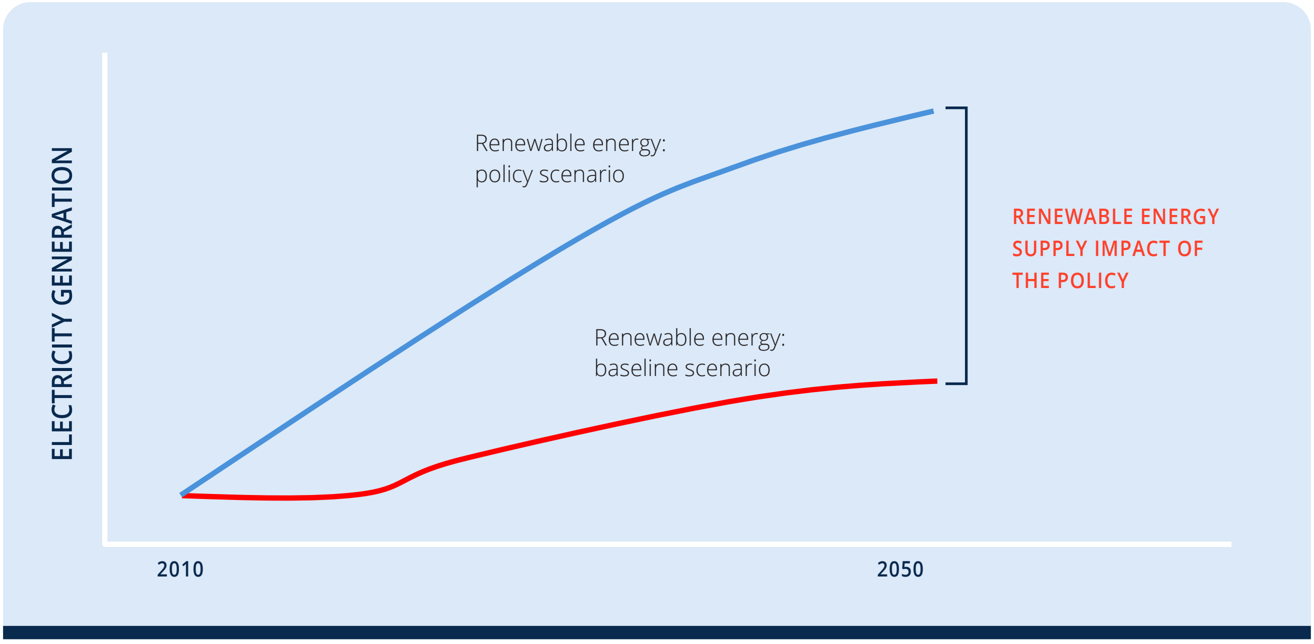
Using the scenario method, users quantify the impact of a policy by comparing two scenarios:

- the baseline scenario, which represents the events or conditions most likely to occur in the absence of the policy (or package of policies) being assessed
- the policy scenario, which represents the events or conditions most likely to occur in the presence of the policy (or package of policies) being assessed.

Figure 8.3 illustrates using the scenario method to quantify the impact of a renewable energy policy on renewable electricity generation.

FIGURE **8.3**

Example of scenario method



In the scenario method, the baseline scenario depends on assumptions relating to key impact drivers over the assessment period. Drivers include other policies that have been implemented or adopted, as well as non-policy drivers, such as economic conditions, energy prices and technological development.

Baseline scenarios can be determined ex-ante or ex-post. An ex-ante baseline scenario is a forward-looking baseline scenario, typically established before implementation of the policy, which is based on forecasts of drivers (such as projected changes in population or economic activity, or other drivers that affect the impact category), in addition to historical data. Ex-ante baseline scenarios are used for ex-ante assessment in Chapter 9.

An ex-post baseline scenario is a backward-looking baseline scenario established during or after implementation of the policy. Ex-post baseline scenarios should include updates to the ex-ante forecasts of drivers, if an ex-ante assessment was first undertaken. Ex-post baseline scenarios are used for ex-post assessment in Chapter 10.

The methods described in this chapter apply to both ex-ante and ex-post baseline scenarios. See Figure 8.4 for an illustration of both types of baseline scenarios. Box 8.2 provides an example of applying the scenario method. Appendix A includes examples of using the scenario method for a solar PV incentive policy.

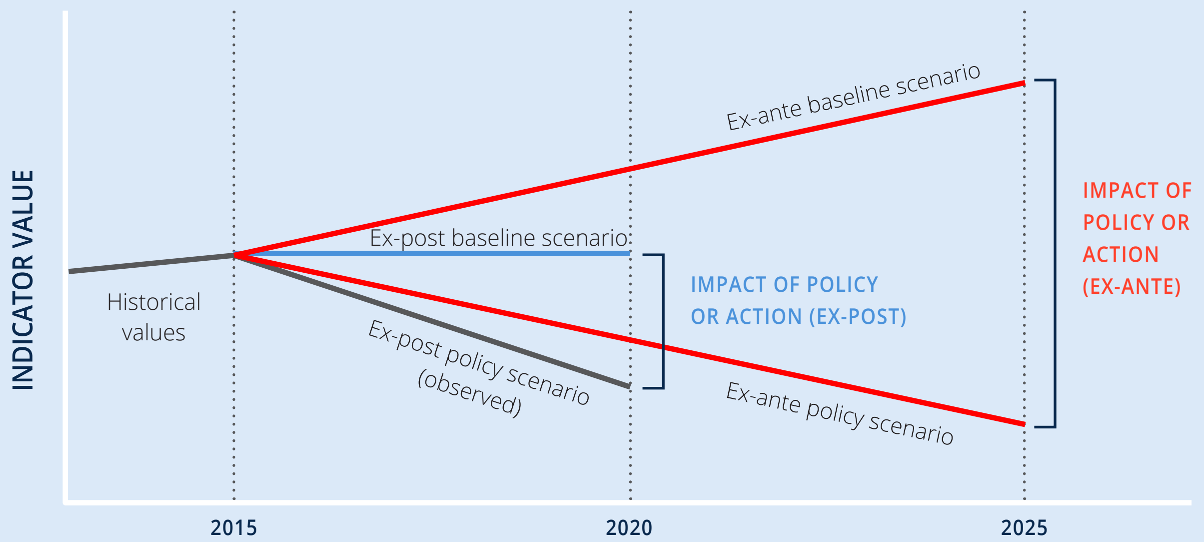
8.2.2 Deemed estimates method

The deemed estimates method (sometimes called a “deemed savings” or “unit savings” approach) is a simplified variation of the scenario method. It involves calculating the impact of a policy without separately defining and estimating baseline and policy scenarios and comparing the two. This method may be appropriate for certain common or homogeneous policies and actions where deemed estimate values are reliable, or in cases where the scenario method is not practical.

To carry out the approach, users estimate the impact by multiplying the number of projects or measures taken as a result of the policy (such as the number of solar PV systems installed) by deemed estimate values that represent the change per project or measure taken (such as the change in jobs or reduction in air pollution per megawatt of solar energy installed). For example, to estimate the energy savings from a policy to replace inefficient lightbulbs with energy-efficient lightbulbs, a user can

FIGURE **8.4**

Ex-ante and ex-post baseline scenarios



Source: Adapted from WRI (2014).

BOX **8.2**

Scenario method example – waste policy in Brazil

To quantify a range of socioeconomic benefits of an integrated solid waste management policy in Brazil, a baseline scenario was compared with four policy scenarios. The baseline scenario assumes that, without the policy, 58% of solid waste would go to sanitary landfills, most of which flare the methane produced. The remaining waste goes to open dumps, where methane vents to the atmosphere.

The four policy scenarios were as follows:

1. All waste sent to a sanitary landfill, with 50% of landfill gas (LFG) collected and flared.
2. Same as scenario 1, but LFG is used to generate electricity that displaces natural gas from the power grid.
3. Anaerobic digestion of organic waste, with electricity generation.
4. Composting of organic waste.

The calculated impacts of implementing all four policy scenarios together, relative to the baseline scenario, are as follows:

- 44,000–110,000 jobs are created.
- 0.5–1.1% of Brazil's electricity demand is saved.
- Brazil's gross domestic product (GDP) increases by $13.3–35.2 billion between 2012 and 2032.
- GHG emissions are reduced by 158–315 $MtCO_2e$.
- 2,500–4,900 premature deaths from air pollution are avoided, with a monetized value of $5.5–10.6 billion
- 550,000–1.1 million tonnes of crops are saved, worth $61–120 million.
- Total net present value of development objectives exceeds $100 billion.

Source: ClimateWorks Foundation and World Bank Group (2014).

multiply the number of lightbulbs replaced by the difference in energy use between a typical inefficient bulb and a typical replacement bulb.

Such approaches simplify the calculation and data collection required to quantify the impact of a policy. However, the calculation risks being oversimplified and inaccurate. The deemed estimates method typically holds constant many factors that could influence the indicator. The estimated impact value (or "deemed estimate") is an implicit representation of the difference between a baseline value and a policy scenario value, which may not use accurate or representative baseline or policy scenario assumptions. The deemed estimate value may assume that the maximum impact (such as energy savings) will be attained, if it does not take into account the specific conditions under which the policy is implemented. For example, using the lightbulb example, the number of hours each lightbulb is in use in the implementing country may differ from the assumptions taken from impacts in another country. These factors should be taken into consideration when calculating impacts to ensure that estimates are realistic – for example, by adjusting the number of hours of operation to represent the local context, or using a conservative estimate where there is uncertainty. Deemed estimate values can be customized to local circumstances or calculated based on local data, rather than using default factors.

Users can apply a different method for each indicator being assessed. For example, the deemed estimates method can be used for one indicator and the scenario method for other indicators. Box 8.3 provides an example of using the deemed estimates method. Appendix A includes examples of using the deemed estimates method for a solar PV incentive policy.

8.2.3 Comparison group method

The comparison group method can only be used for ex-post assessments and if an equivalent comparison group exists. To reliably and credibly implement a comparison group method, actors affected by the policy (the policy group) and actors not affected by the policy (the comparison group or control group) must be otherwise equivalent. Under ideal experimental conditions, the two groups would be randomly assigned to ensure that any differences between the groups are a result of the policy, rather than any underlying systematic differences or biases. If random assignment is not possible, other methods can be used to control for external factors, avoid "selection bias", and ensure valid comparisons (described further in Chapter 10).[23]

If an appropriate comparison group is not available, the scenario method or deemed estimates method should be used. In some cases, data obtained from a comparison group can also be used to update, calibrate or validate assumptions and data used in the scenario method or deemed estimates method. Box 8.4 provides an example of the approach.

The remainder of this chapter focuses on steps involved in applying the scenario method. Guidance

BOX **8.3**

Example of deemed estimates method

A Gold Standard (GS) study used a deemed estimates method to capture and monetize the environmental and socioeconomic net benefits associated with GS carbon projects. To quantify the improvements in health from a cookstoves project, the mortality rate was applied to the number of households with cookstoves to determine the reduction in mortality. First, the indicator was identified as the difference in indoor $PM_{2.5}$. Next, the study created an index based on the linear relationship between indoor air quality and mortality. The percentage reduction in mortality was calculated by applying $PM_{2.5}$ changes to the index. The mortality rate was then applied to the number of households with cookstoves to determine the reduction in mortality.

Source: Gold Standard (2014).

BOX **8.4**

Example of deemed estimates method

The United Kingdom Government provides analysts and policymakers at all levels of government with guidance on how to assess and review policies and projects to ensure that public funds are well spent. It views evaluation as essential to determining whether policies are effective.

The guidance, provided in *The Magenta Book*, includes approaches for using a control group to establish a baseline (i.e. counterfactual) scenario. It suggests that controlling policy allocation (i.e. which individuals or areas receive policy interventions, and when) can play a key role in successful impact evaluation by affecting whether there is a meaningful comparison group. The guidance offers several examples of how to do this:

- **Pilots.** Allow the policy to be tried and information to be collected before committing full-scale resources. Not every potential subject is exposed to the policy, and people who are not exposed can act as a control group.
- **Randomization and randomized control trials (RCT).** Allocate by lottery or other purely random mechanism which individuals, groups or local areas receive the policy. Carefully conducted, an RCT provides the clearest evidence of whether a policy has had an impact.
- **Phased introduction.** Implement the policy sequentially over a period of time. The periods when some participants have received the intervention and others have not can serve to generate a comparison group.

Source: HM Treasury, United Kingdom (2011).

[23] For more information on the applicability of the comparison group method, see Coalition for Evidence-Based Policy (2014).

on the comparison group method is provided in Chapter 10.

8.3 Define the baseline scenario and estimate baseline values for each indicator

This section provides guidance on defining the baseline scenario and estimating baseline scenario values using the scenario method. It is applicable to all ex-ante assessments and to ex-post assessments that use the scenario method.

Figure 8.5 outlines the steps in this section. Users may find it useful to follow the steps in this section separately for each impact category being estimated, since the choices made regarding methods and data are likely to be different for each impact category. In this case, users should complete the steps for one impact category at a time, then repeat the process for each impact category included in the assessment. Involving stakeholders in the selection and estimation of baseline scenarios is important to ensure credible assumptions and valid results.

Appendix A provides an example of carrying out the steps in this section for a solar PV incentive policy.

8.3.1 Select a desired level of accuracy and complexity

A range of methods and data can be used to estimate the baseline scenario. In general, users should follow the most accurate approach that is feasible in the context of the assessment objectives, capacity and resources. Because a wide variety of methods and data can be used, it is important to report the methods, assumptions and data used to estimate the baseline scenario.

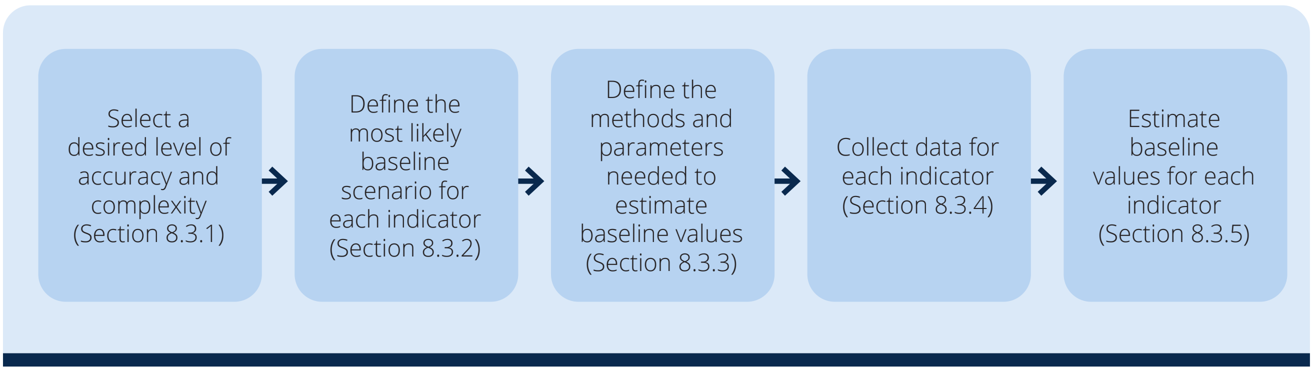
Users can choose different levels of accuracy for different impact categories included in the assessment. Users should consider the resources available for each impact category being assessed, and focus efforts on achieving higher levels of accuracy for impact categories determined to be the most relevant and significant. The availability of data, methods and models, or resources may constrain the level of accuracy, even for high-priority impacts. Users should clearly document the uncertainty – either qualitatively or quantitatively – associated with the results and explain how the methods chosen for the assessment provide an acceptable level of accuracy.

Estimation of the baseline scenario can range from simple to complex, as explained below and illustrated in Figure 8.6:

- **Constant baseline.** A constant baseline uses historical or current values as the baseline scenario. This assumes that there will be no change in the impact category in the future in the absence of the policy. This is a simple "before" and "after" comparison to indicate the impacts of the policy.
- **Simple trend baseline.** A simple trend baseline uses historical trends as the basis for the baseline scenario, and assumes that the historical trend will remain the same into the future in the absence of the policy. This can take the form of a simple linear extrapolation, exponential extrapolation or other forms of extrapolation.

FIGURE **8.5**

Overview of steps in defining the baseline scenario and estimating baseline scenario values



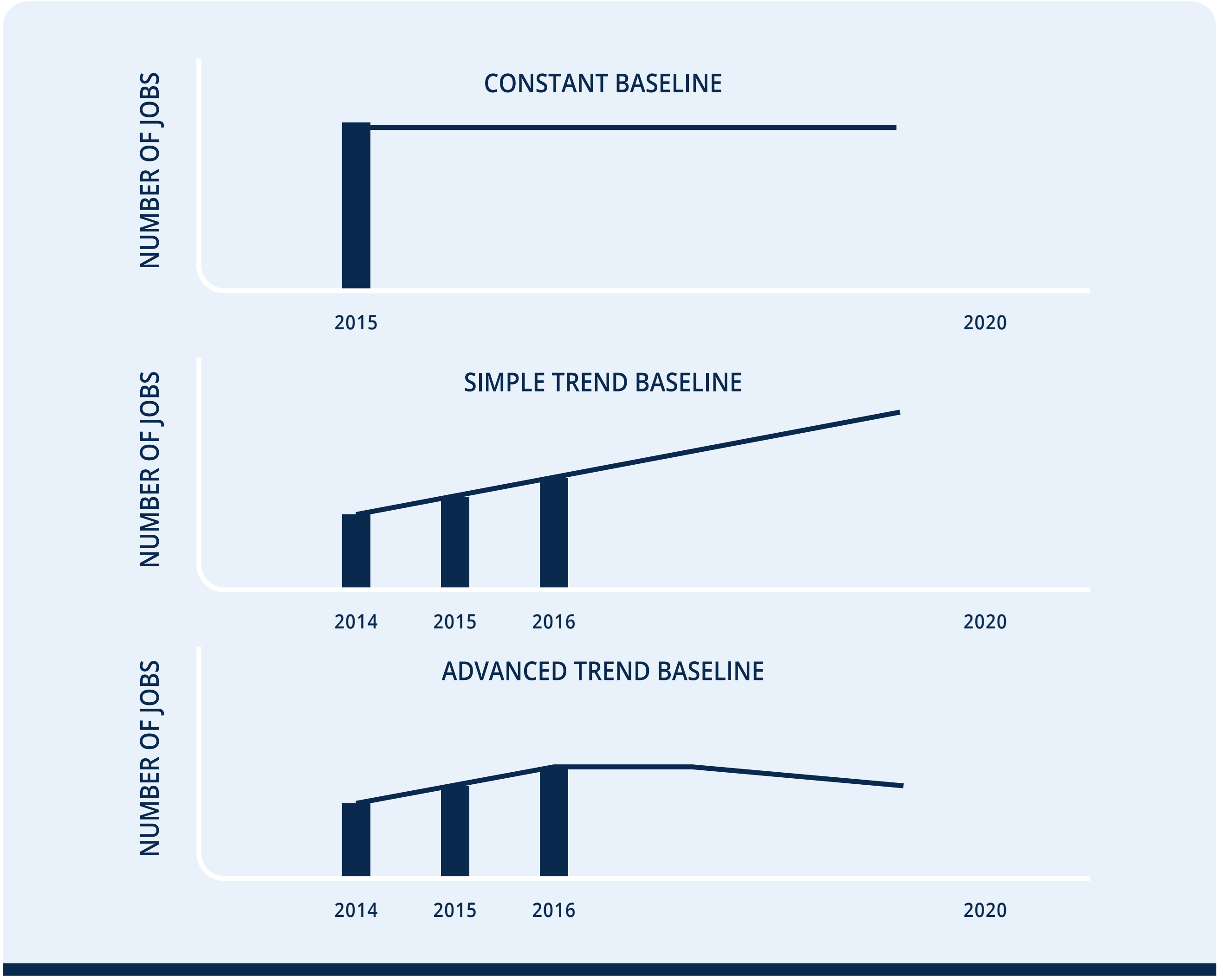
- **Advanced trend baseline.** An advanced trend baseline is a more complex approach that models the impact of many interacting elements, such as the impacts of non-policy drivers (such as macroeconomic conditions) and other policies in affecting conditions in the future.

The choice of baseline scenario depends on which is most appropriate for a given impact category and situation, as well as users' resources, capacity, access to data, and availability of appropriate models and methods. Users should choose methods and data that yield the most accurate results within a given context, based on the methodological and data options available.

A constant baseline is the simplest option and may be appropriate when indicators are considered likely to remain stable over time. A simple trend baseline is most appropriate if the change in indicator values (rather than actual indicator values) is expected to remain stable over time. In general, more advanced baselines are likely to be more accurate, since they take into account various drivers that affect conditions over time. However, more advanced baselines will only be more accurate if the data and methods available to integrate the impacts of multiple drivers are robust. Users should weigh the priority of each impact category and allocate resources accordingly when determining the complexity of the baseline scenario.

FIGURE **8.6**

Examples of constant, simple trend and advanced trend baselines



8.3.2 Define the most likely baseline scenario for each indicator

A critical step in applying the scenario method is to define the baseline scenario. It is a *key recommendation* to define a baseline scenario that represents the conditions most likely to occur in the absence of the policy for each indicator included in the assessment boundary.

Users should create a baseline scenario for each significant impact to be quantitatively assessed, where feasible. The baseline scenarios may be developed separately for each impact of interest.

The most likely baseline scenario depends on drivers that would affect the impact in the absence of the policy being assessed. Identifying key drivers for each significant impact being assessed and making reasonable assumptions about their most likely values in the absence of the policy being assessed can have a large effect on the baseline scenario, and consequently on the eventual estimate of the impact of the policy.

Drivers that affect baseline values are divided into two types:

- **other policies** – policies, actions and projects, other than the policy being assessed, that are expected to affect the impacts included in the assessment boundary
- **non-policy drivers** – other conditions, such as socioeconomic factors and market forces, that are expected to affect the impacts included in the assessment boundary

Users should ensure that baseline scenarios defined for each impact category are consistent. That is, where different impact categories are affected by common drivers or assumptions, the same values should be used for the baseline scenarios for each impact category. For example, if GDP is a common driver needed for assessing both the job impacts and the economic developments impacts of a solar PV incentive policy, users should use the same assumed GDP values for both impact categories.

Users should identify plausible baseline options and choose the option that is considered to be the most likely to occur in the absence of the policy. The choice should be made in consultation with stakeholders and experts. Possible options include:

- continuation of current technologies, practices or conditions
- discrete baseline alternatives, practices, technologies or scenarios (such as the least-cost alternative practice or technology), identified using environmental, financial, economic or behavioural analysis or modelling
- a performance standard or benchmark that indicates baseline trends.

Including other policies

In addition to the policy being assessed, there are likely to be other policies, actions or projects that affect the indicator being estimated. These may include regulations and standards, taxes and charges, subsidies and incentives, voluntary agreements, information instruments, or other types of policies and actions.

In the case of a national solar PV incentive policy, other policies that may affect the amount of solar PV installed by households and businesses in the baseline scenario include national regulations that facilitate connection of distributed generation to the electric grid (other national policies), municipal incentives to promote renewable energy at the local level (subnational policies), and utility incentives for solar PV installation (private sector actions). These other policies affect conditions in the baseline scenario and should be considered when a user is determining the incremental impact of the national solar PV policy compared with what would have happened in the absence of the policy. Appendix A provides an example of including other policies in the baseline scenario.

To identify other policies and actions to consider in the baseline scenario, users should identify key parameters in the assessment – such as the amount of solar PV installed – and identify other policies and actions that affect the same parameters.

Users should include all other policies, actions and projects in each baseline scenario that:

- have a significant effect on the impacts included in the assessment boundary
- are implemented or adopted at the time the assessment is carried out (for ex-ante assessment) or during the assessment period (for ex-post assessment).

Table 8.3 provides definitions of implemented, adopted and planned policies, and guidance on whether to include each in the baseline scenario.

Published baseline values may already include the impact of existing policies and actions in the baseline scenario. If it is not possible to include a relevant policy in the baseline scenario, users should document and justify its exclusion.

Users can establish a significance threshold or other criteria to determine which policies, actions and projects are significant and should be included. For other policies that are included, users should determine whether they are designed to operate indefinitely or are limited in duration. Users should assume that policies will operate indefinitely unless an end date is explicitly stated.

Including non-policy drivers

Non-policy drivers include a wide range of exogenous factors, such as socioeconomic factors and market forces, that may cause changes in the impact category but are not a result of the policy being assessed. Users should identify non-policy drivers based on literature reviews of similar assessments and policies, consultations with relevant experts and stakeholders, expert judgment, modelling results, or other methods.

In the case of a solar PV incentive policy, non-policy drivers that affect the amount of solar PV installed by households and businesses in the baseline scenario may include the price of solar PV systems (the less expensive they are, the more households and businesses will install them) and the price of electricity (the more expensive electricity from the grid is, the greater the incentive for households and businesses to install solar PV systems). These factors affect conditions in the baseline scenario and should be considered to determine the impact of the solar PV incentive policy compared with what would have happened in the absence of the policy.

Users should include all non-policy drivers in the baseline scenario that are not caused by the policy being assessed (i.e. that are exogenous to the assessment), and that are expected to result in a significant change in calculated impacts between the baseline scenario and the policy scenario. In ex-ante

TABLE **8.3**

Definitions of implemented, adopted and planned policies and actions

Policy status	Definition	Guidance for inclusion in the baseline scenario
Implemented	Policies that are currently in effect, as evidenced by one or more of the following: (1) relevant legislation or regulation is in force, (2) one or more voluntary agreements have been established and are in force, (3) financial resources have been allocated, (4) human resources have been mobilized.	Should be included for both ex-ante and ex-post assessments.
Adopted	Policies for which an official government decision has been made and there is a clear commitment to proceed with implementation, but implementation has not yet begun (e.g. a law has been passed, but regulations to implement the law have not yet been established or are not being enforced).	Should be included for ex-ante assessment if polices are likely to be implemented and there is enough information to estimate the impacts. Should not be included for ex-post assessment.
Planned	Policy options that are under discussion, and have a realistic chance of being adopted and implemented in the future, but have not yet been adopted or implemented.	In some cases, users may want to include planned policies for ex-ante assessment – for example, if the objective is to assess the impact of one planned policy relative to other planned policies. Should not be included for ex-post assessment.

Source: Adapted from WRI (2014).

assessments, users do not need to include drivers that are expected to remain the same under both the policy scenario and the baseline scenario. Users can establish a significance threshold or other criteria to determine which non-policy drivers are significant.

To identify non-policy drivers that should be considered in the baseline scenario, users should identify key parameters in the assessment – such as the amount of solar PV installed – and identify other policies and actions that affect the same parameters.

Published baseline values may already include the impact of non-policy drivers in the baseline scenario. If it is not possible to include a relevant non-policy driver in the baseline scenario, users should document and justify its exclusion.

Defining a range of baseline scenario options

If possible, users should identify the single baseline scenario that is considered most likely for each impact being assessed. In certain cases, multiple baseline options may seem equally likely. In such cases, users should consider estimating and reporting a range of results based on these alternative baseline scenarios. Users should conduct sensitivity analysis to see how the results vary depending on the selection of baseline options. Sensitivity analysis involves varying the parameters, or combinations of parameters, to understand the sensitivity of the overall results to changes in those parameters. It is a useful tool for understanding differences resulting from methodological choices and assumptions, and exploring model sensitivities to inputs. Sensitivity analysis is further described in Chapter 11.

Use of assumptions and expert judgment

Assumptions or expert judgment will likely be required where information is not available to make a reasonable assumption about the value of a parameter. Users may need to use proxy data, interpolate information, estimate a rate of growth, or use other types of assumptions or judgment. Users can apply their own expert judgment or consult experts. When doing so, it is important to document that other data sources were not available, and the reasons why, and the rationale for the value chosen.

8.3.3 Define the methods and parameters needed to estimate baseline values

For each indicator to be assessed, users should first identify a method (such as an equation, algorithm or model) for estimating the baseline scenario, then identify the data requirements needed to quantify the baseline value using the chosen method. When selecting the baseline scenario method, consideration should be given to the data needs and data availability under both the baseline scenario and the policy scenario, since the same method or model should be used for both scenarios.

Multiple types of data can be used to estimate the impacts of policies, including bottom-up and top-down data (see Table 8.4).

Bottom-up and top-down data may be appropriate in different contexts and are valuable for different purposes. For example, top-down data may be most appropriate for national policies, whereas bottom-up data may be better suited to smaller-scale policies. The choice of bottom-up versus top-down approaches depends on data availability and the needs of the assessment.

A wide range of tools and models can be used to quantify social, environmental and economic

TABLE **8.4**

Overview of bottom-up and top-down data

Type of data	Description
Bottom-up	Bottom-up data are measured, monitored or collected at the facility, entity or project level. Examples are energy used at a facility (e.g. using a measuring device such as a fuel meter) and production output.
Top-down	Top-down data are macro-level data or statistics collected at the jurisdiction or sector level. Examples are national energy use, population, GDP and fuel prices. In some cases, top-down data are aggregated from bottom-up data sources.

Source: Adapted from WRI (2014).

impacts. Methods range from simple equations (e.g. simple extrapolation) to complex models (e.g. simulation models, computable general equilibrium models, integrated assessment models). Simple equations may not be sufficient to represent the complexity needed to accurately estimate baseline or policy scenarios, or to capture the difference between them. Detailed models may be needed to estimate the impacts of certain policies. Detailed models may also be appropriate when the chosen impact category includes multiple interacting parameters.

A variety of methods can be used, depending on what type of data is available and the level of accuracy desired. Some methods (e.g. engineering models) calculate or model the impact of a policy for each facility, project or entity affected by the policy, then aggregate across all facilities, projects or entities to determine the total impact of the policy. Other methods may include regression analysis or other statistical methods, simulation models, computable general equilibrium models or other models.

For example, a user assessing the impact of a solar PV incentive policy on jobs could use a bottom-up approach by multiplying the estimated number of buildings that install solar PV systems by the estimated number of workers needed to install and maintain solar PV systems per building, using data provided by individual companies. Alternatively, a user could use a top-down approach by using economic models based on national employment statistics on the number of people employed in the solar energy industry and other relevant variables. Hybrid approaches that combine elements of both bottom-up and top-down approaches may also be used.

The ICAT website[24] provides examples of tools and models to support impact quantification. Users can use existing methods or models, or develop new ones (if no relevant and appropriate methods or models exist). Users should select a tool that achieves sufficiently accurate results in the context of objectives, data availability and resource constraints. Objectives may range from theoretical explorations of policy questions, to practical applications of the results in a governmental regulatory or programmatic context, to forecasting for planning purposes. These needs will determine the range of sectors that must be included in the tool, the geographic scales and time frames. For example, some users may choose simple scenarios to support their analyses, whereas others may want to use additional variables, longer time scales or more detailed time steps, or have the flexibility to incorporate changing policies or patterns and develop conditional futures. Likewise, some may be interested in assessing a small geographic region, a single sector or even a single project, whereas others may want multi-scale futures or integrated approaches.[25]

A suite of models may be available, with the choice between models depending on users' specific needs. Models will require varying levels of data input, user knowledge and expertise, and cost. Selecting the most appropriate tool will depend on users' available time and financial resources, as well as their team expertise. These considerations are illustrated in Table 8.5.

Table 8.6 provides an overview of types of economic models for quantifying economic impacts. Box 8.5 provides an explanation of one model for quantifying job and economic impacts of constructing and operating power plants, such as wind farms. Box 8.6 provides an example of a model for estimating the health and economic effects of air pollution.

[24] https://climateactiontransparency.org/icat-toolbox/sustainable-development

[25] USGCRP (2016).

TABLE **8.5**

Considerations for selecting tools to assess social, economic or environmental impacts

Level of depth/ accuracy[a]	Model capabilities	Cost	Ease of use	Data inputs
Higher	Assumptions embedded in the model are dynamic; can optimize for a specific variable or output; may produce a range of quantitative outputs	Up to tens of thousands of dollars	Highly complex; use requires trained experts, and significant time to gather input data and produce model output (several weeks or months)	Highly data-intensive; may rely on software of models for inputs
↓	↓	↓	↓	↓
Lower	Assumptions embedded in the model are static; cannot optimize for a specific variable or output; may produce limited quantitative outputs	No cost or low cost	Designed for use by the public: easy to navigate and run; requires limited time to run (several hours or days)	Not data-intensive; relies on pre-populated data and default assumptions

[a] The level of accuracy varies with the various attributes presented here. In reality, a complex, advanced model that has a high cost and requires extensive data inputs will only be as accurate as the quality of the data that go into it.

TABLE **8.6**

Overview of modelling approaches and tools for economic analysis

Method	Advantages	Disadvantages
Input–output model (also called multiplier analysis)	• Quantifies the total economic effects of a change in the demand for a given product or service • Can be inexpensive	• Static; multipliers represent only a snapshot of the economy at a given point in time • Generally assumes fixed prices • Typically does not account for substitution effects, supply constraints, and changes in competitiveness or other demographic factors
Econometric models	• Usually dynamic; can estimate and track changes in policy impacts over time • Coefficients are based on historical data and relationships, and statistical methods can be used to assess model credibility	• Historical patterns may not be best indicator or predictor of future relationships • Some econometric models do not allow foresight
Computable general equilibrium models	• Accounts for substitution effects, supply constraints and price adjustments	• Not available for all regions
Hybrid models	• Most sophisticated, combining aspects of all the above • Dynamic; can be used to analyse both short- and long-term impacts • Can be used to model regional interactions	• Can be expensive

Source: U.S. EPA (no date, a).

BOX **8.5**

JEDI model for estimating job and economic impacts from power plants

The National Renewable Energy Laboratory's Jobs and Economic Development Impact (JEDI) model is an Excel-based model that estimates the number of jobs and economic impacts from constructing and operating power plants, fuel production facilities and other projects at the local level. For example, JEDI estimates the number of construction jobs from a new wind farm. JEDI models are used by decision makers, public utility commissions, potential project owners, developers and others.

The model estimates the project costs and the economic impacts in terms of jobs, earnings (i.e. wages and salaries) and output (i.e. value of production) resulting from the project. Jobs, earnings and output are distributed across three categories: project development and on-site labour impacts, local revenue and supply chain impacts, and induced impacts. The results are more likely to better reflect the actual impacts from the specific project if the user can incorporate project-specific data and the share of spending expected to occur locally. Project-specific data include a bill of goods (costs associated with actual construction of the facility, roads, etc., as well as equipment costs, other services and fees required), annual operating and maintenance costs, the portion of expenditures to be spent locally, financing terms and local tax rates. The analysis is not designed to provide a precise forecast, but rather an estimate of overall economic impacts from specific scenarios.

The JEDI model uses an input–output methodology. It uses economic data (multipliers and consumption patterns) to estimate the local economic activity and the resulting impact from new energy generation plants. This involves aggregating national and regional economic and demographic data to calculate inter-industry linkages, the relationships between changes in demand for goods and services, and the associated economic activity at the local and regional levels. Local spending results from using local labour (e.g. concrete pouring), services (e.g. engineering, design, legal), materials (e.g. wind turbine blades) or other components (e.g. nuts and bolts).

Source: NREL (no date).

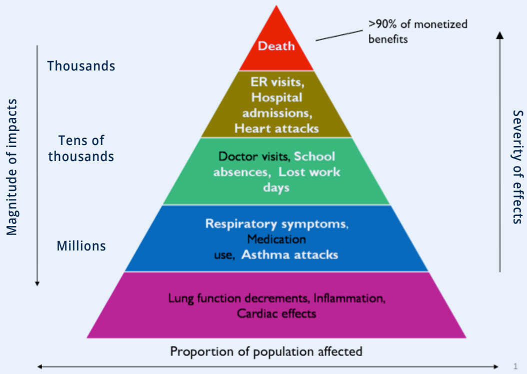
BOX **8.6**

The Benefits Mapping and Analysis Program (BenMAP) model for estimating the health and economic effects of air pollution

The United States Environmental Protection Agency's BenMAP-Community Edition (CE) tool estimates the economic value of health impacts resulting from changes in air quality – specifically, ground-level ozone and fine particles. BenMAP-CE is an open-source computer program that calculates the number and economic value of air pollution–related deaths and illnesses. The software incorporates a database that includes many of the concentration–response relationships, population files, and health and economic data needed to quantify these impacts.

Air pollution affects health through fine particles that penetrate deep into the lungs and enter the bloodstream. Health impacts from particles include premature death, non-fatal heart attacks and aggravated asthma. Ground-level ozone is an oxidant that can irritate airways in the lungs. Health impacts from ozone include premature death, aggravated asthma and lost days of school.

A "pyramid of effects" from air pollution

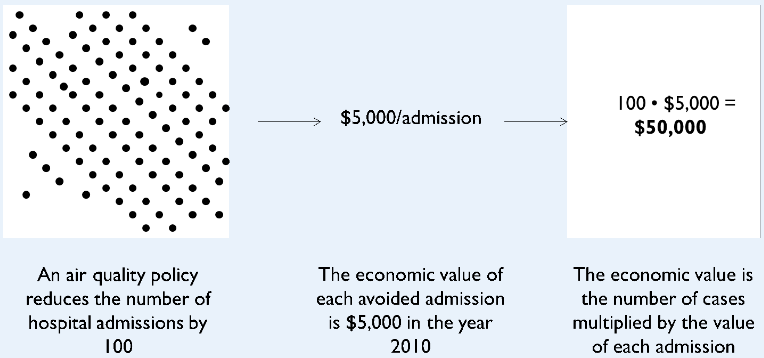


The pyramid describes how the incidence and severity of fine particle- and ozone-related health impacts are related. Health outcomes towards the bottom of the pyramid, such as asthma attacks and cardiac effects, are less severe, and affect a larger proportion of the population. Impacts towards the tip of the pyramid, such as hospital admissions and heart attacks, are more severe and affect a smaller proportion of the population. BenMAP-CE quantifies the impacts shown in white.

BOX **8.6,** continued

The Benefits Mapping and Analysis Program (BenMAP) model for estimating the health and economic effects of air pollution

BenMAP-CE calculates the economic value of air quality change using both "cost of illness" and "willingness to pay" metrics. The cost of illness metric summarizes the expenses that an individual must bear for air pollution–related hospital admissions, visits to the emergency department and other outcomes; this metric includes the value of medical expenses and lost work, but not the value that individuals place on pain and suffering associated with the event. In contrast, willingness to pay metrics account for the direct costs noted above as well as the value that individuals place on pain and suffering, loss of satisfaction and loss of leisure time. This simple example summarizes the procedure for calculating economic values using these two metrics in BenMAP-CE.



Source: U.S. EPA (no date, b).

8.3.4 Collect data for each indicator

The next step is to collect data for each indicator (and parameter, if applicable) in each baseline scenario. To estimate baseline values for each indicator, users should first decide whether to estimate new baseline values or use baseline values from published data sources. For some indicators, published values may not be available. In this case, users should estimate new values.

Users should collect data separately for different groups in society, where relevant, such as men and women, people of different income groups, people of different racial or ethnic groups, people of different education levels, people from different geographic regions, and people in urban versus rural locations.

Either using published values or estimating new values, users should report the baseline values for each indicator being estimated over defined time periods, such as annually over the assessment period, if feasible. It is important to report the methods, assumptions and data sources used. Users should also justify the choice of whether to estimate new baseline values and assumptions or to use published baseline values and assumptions. If no data source is cited, users should provide sufficient information to enable stakeholders and others tracking the impact over time to know where to look for updates to the data.

When collecting data from various data sources, users should consider whether the data source is readily available, whether data sources will be available to track indicator values over time, and how expensive or labour-intensive it will be to collect data over time. Users should use conservative assumptions to define baseline values when uncertainty is high or a range of possible values exist. Conservative values and assumptions are more likely to overestimate negative impacts or underestimate positive impacts resulting from a policy.

Parameters whose values will not change between the baseline and policy scenario may “cancel out” when the baseline and policy values are subtracted. Where that is the case, the value chosen for the parameter will not influence the final result, and fewer resources should be expended to gather the data for the parameter. Ideally, where such parameters will cancel out in the final comparison, the method should be simplified, and its description narrowed to remove parameters that are not relevant.

Option 1: Using baseline values from published data sources

In some cases, existing data sources of sufficient quality may be available to determine baseline values for indicators. Potential data sources of historical or projected data include published studies of similar policies and impact categories in the same or other jurisdictions, peer-reviewed scientific literature, government statistics, reports published by international institutions (such as the International Energy Agency, IPCC, the World Bank, and the Food and Agriculture Organization of the United Nations – FAO), and economic and engineering analyses and models.

Users should use high-quality, up-to-date and peer-reviewed data from recognized, publicly available, credible sources, if available. When selecting data sources, users should apply the data quality indicators in Table 8.7 as a guide to obtaining the highest-quality data available. Users should select data that are the most representative in terms of technologies, practices, time and geography; the most complete; and the most reliable.

In some cases, the baseline scenario itself may be the subject of published research and available for use. As above, the information should be high quality and credible. In addition, the method used should be sufficiently clear that users can generate a comparable policy scenario, with consistent methods, assumptions and data sources.

For published values, a range of data may be available, such as:

- international default values
- national average values
- jurisdiction- or activity-specific data.

In general, users should use the most accurate and representative data available.

TABLE **8.7**

Data quality indicators

Indicator	Description
Technological representativeness	The degree to which the data set reflects the relevant technologies, processes or practices
Temporal representativeness	The degree to which the data set reflects the relevant time period
Geographical representativeness	The degree to which the data set reflects the relevant geographic location (e.g. country, city, site)
Completeness	The degree to which the data are statistically representative of the relevant activity. Completeness includes the percentage of locations for which data are available and used out of the total number that relate to a specific activity. Completeness also addresses seasonal and other normal fluctuations in data.
Reliability	The degree to which the sources, data-collection methods and verification procedures used to obtain the data are dependable. Data should represent the most likely value of the parameter over the assessment period.

Source: WRI (2014), based on Weidema and Wesnaes (1996).

Option 2: Estimating new baseline values
In some cases, no published baseline data and assumptions will be available for historical or projected data, or the existing data may be incomplete, of poor quality, or in need of supplementation or further disaggregation. Users should estimate new baseline values when no relevant data are available that support the level of accuracy needed to meet the stated objectives.

To estimate new baseline values for a given indicator, users should:

1. collect historical data for the indicator

2. identify other policies and non-policy drivers that affect each indicator over the assessment period, and make assumptions for those drivers

3. estimate baseline values for each indicator, based on historical data and assumptions about drivers.

8.3.5 Estimate baseline values for each indicator

The final step in developing the baseline is to apply the method to the data collected to estimate baseline values for each indicator.

It is a *key recommendation* to estimate baseline values over the assessment period for each indicator included in the assessment boundary. Any impact included in the assessment boundary that cannot be estimated should be assessed qualitatively (as described in Chapter 7). It is a *key recommendation* to separately estimate baseline values for different groups in society, where relevant.

See Appendix A for an example of estimating the impact of a solar PV incentive policy, including estimating the baseline. The ICAT website[26] provides examples of tools and models to support impact quantification.

[26] https://climateactiontransparency.org/icat-toolbox/sustainable-development

9 Estimating impacts ex-ante

This chapter describes how to estimate the expected future impacts of a policy (ex-ante assessment). In this chapter, users estimate policy scenario values for the indicators included in the assessment boundary. The impacts of the policy are estimated by subtracting baseline values (as determined in Chapter 8) from policy scenario values (as determined in this chapter). This chapter is structured around the steps in the scenario method, but the guidance is also helpful when using the deemed estimates method (defined in Chapter 8). Users who are not quantitatively assessing impacts ex-ante can skip this chapter.

Checklist of key recommendations

- Define a policy scenario that represents the conditions most likely to occur in the presence of the policy over time for each indicator being estimated, taking into account all specific impacts included in the quantitative assessment boundary
- Estimate the net impact of the policy on each indicator by subtracting baseline values from policy scenario values, taking into account all specific impacts included in the quantitative assessment boundary
- Separately assess the impacts of the policy on different groups in society, where relevant

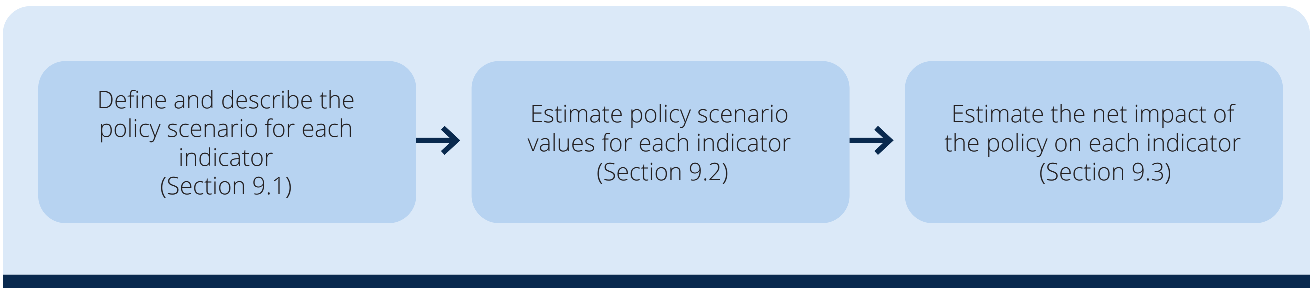
9.1 Define and describe the policy scenario for each indicator

In Chapter 8, users defined an indicator for each impact category included in the assessment boundary. For examples of indicators, see Table 5.5. The indicators will be estimated for the baseline and policy scenarios to estimate the impact of the policy. Each indicator will generally require a different assessment method. The same general assessment method(s) used to estimate the baseline value (in Chapter 8) should be used to estimate the policy scenario value for each indicator to ensure methodological consistency between the baseline and policy scenario estimations. Consistency ensures that the estimated impact reflects underlying differences between the two scenarios, rather than differences in methods. If it is not feasible or appropriate to use the same method, users should justify why different methods have been used. The ICAT website[27] provides examples of tools and models to support impact quantification.

It is a *key recommendation* to define a policy scenario that represents the conditions most likely to occur in the presence of the policy over time for each indicator being estimated, taking into account all specific impacts included in the quantitative assessment boundary. The policy scenario represents the events or conditions most likely to

FIGURE **9.1**

Overview of steps in the chapter



[27] https://climateactiontransparency.org/icat-toolbox/sustainable-development

occur in the presence of the policy (or package of policies) being assessed. The only difference between the baseline scenario and the policy scenario is that the policy scenario includes the changes caused by the policy (or package of policies) being assessed. See Figure 9.2 for an illustration of estimating impacts ex-ante. Users can estimate policy scenario values either before or after estimating baseline values.

Users should identify various policy scenario options and choose the one considered to be the most likely to occur in the presence of the policy. It is important to consult stakeholders during the selection and estimation of the policy scenario to ensure credibility. Users should describe the policy scenario for each indicator being estimated.

9.2 Estimate policy scenario values for each indicator

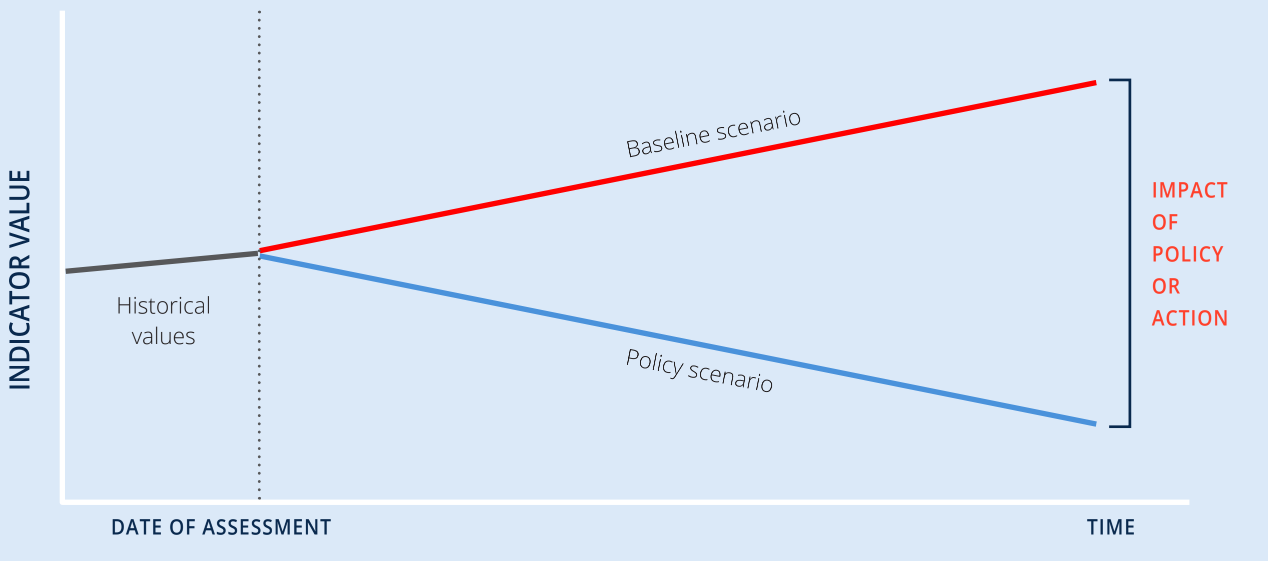
For some indicators, it is possible to directly estimate policy scenario values, without the need for additional parameters. Other assessment methods require multiple parameters to estimate policy scenario values for a given indicator. For example, estimating household cost savings from an energy efficiency policy requires data on the electricity price and the quantity of energy consumed in the baseline scenario and the policy scenario. In this example, "household cost savings" is the indicator (measured in dollars or other currency), and "electricity price" and "quantity of energy consumed" are parameters. These two parameters are not themselves indicators of interest, but are necessary to calculate the impact on the indicator of interest ("household cost savings"). Calculating the impact on each indicator therefore requires estimating policy scenario values for each parameter in the assessment method(s).

To estimate policy scenario values for each parameter, users should first identify which parameters are affected by the policy. In the example above, "quantity of energy consumed" is affected by the policy, since it is designed to save energy, whereas "electricity price" is not affected by the policy.

Parameters that are affected by the policy (such as "quantity of energy consumed") need to be estimated in the policy scenario. These parameter values are expected to differ between the policy scenario and the baseline scenario. Users should estimate policy scenario values for these parameters by developing assumptions about how the policy is expected to affect each parameter over the assessment period (described further in Section 9.3). This follows the

FIGURE **9.2**

Estimating impacts ex-ante



Source: Adapted from WRI (2014).

same general process as estimating baseline values in Section 8.3, but instead now is used to estimate policy scenario values.

Parameters that are not affected by the policy (such as "electricity price") do not need to be estimated again, since the parameter value is not expected to differ between the policy scenario and the baseline scenario.

Users should report the policy scenario values for each indicator being estimated, and the methods, assumptions and data sources used to calculate policy scenario values.

9.2.1 Guidance for estimating policy scenario values

Users can either:

- use policy scenario values from published data sources (option 1), or
- estimate new policy scenario values (option 2).

Option 1: Using policy scenario values from published data sources
In some cases, existing data sources of sufficient quality may be available to determine policy scenario values. Potential data sources of historical or projected data include published studies of similar policies and impact categories in the same or other jurisdictions, peer-reviewed scientific literature, government statistics, reports published by international institutions (such as the International Energy Agency, IPCC, the World Bank, FAO), and economic and engineering analyses and models.

Users should use high-quality, up-to-date and peer-reviewed data from recognized, publicly available, credible sources, if available. When selecting data sources, users should apply the data quality indicators in Table 8.7 as a guide to obtaining the highest-quality data available. Users should select data that are the most representative in terms of technologies, practices, time and geography, and the most complete.

For published values, a range of data may be available, such as:

- international default values
- national average values
- jurisdiction- or activity-specific data.

In general, users should use the most accurate data available.

Option 2: Estimating new policy scenario values
In some cases, no relevant published data and assumptions will be available for policy scenario values, or the existing data may be incomplete, of poor quality, or in need of supplementation or further disaggregation. Users should estimate new policy scenario values and assumptions when no relevant data are available that support the level of accuracy needed to meet the stated objectives.

Users can use a range of methods and data to estimate policy scenario values, ranging from simpler to more complex. For example, a simple method may involve an assumption that parameters will remain static (fixed) over the assessment period or involve a linear extrapolation of historical trends. A more complex approach may involve an assumption that parameters are dynamic (changing) over the assessment period; the values may be estimated using detailed modelling or equations.

Users should estimate the change in the indicator over time, based on what is considered to be the most likely scenario for each indicator. The most likely scenario can be based on evidence, such as peer-reviewed literature, modelling or simulation exercises, government statistics, or expert judgment. If scenarios or methods in existing literature are not similar enough to use directly, users may need to make adjustments to adapt the results found in literature to the assumptions made in the baseline scenario and other elements of the assessment. Users may also need to apply new methods, models and assumptions not previously used in the baseline method to estimate the expected change in each indicator as a result of the policy. However, new methods should not be used to estimate total impacts of the policy, since the same general methods used to estimate baseline values should be used to estimate policy scenario values, to ensure consistency.

Each indicator may be assumed to be static or dynamic over the assessment period. Dynamic indicators can change at a linear or non-linear rate. In many cases, dynamic models that allow for conditions to change throughout the assessment period are expected to be the most accurate, so they should be used where relevant and feasible.

To estimate policy scenario values for each indicator affected by the policy, users should consider a variety of factors (described in more detail below), such as:

- historical trends and expected values in the baseline scenario
- timing of impacts
- barriers to policy implementation or effectiveness
- policy interactions
- sensitivity of parameters to assumptions.

To the extent relevant, users should also consider:

- non-policy drivers included in the baseline scenario (see Chapter 8), which should be different between the baseline and policy scenarios if they are affected by the policy
- learning curves (economic patterns that can accelerate or slow new product development and deployment)
- economies of scale
- technology penetration or adoption rates (the pace of adoption by targeted actors, which may be slow initially then accelerate as products become more socially accepted).

Depending on the assessment, users may not need to consider each of these factors. In practice, users may also be limited by:

- the type of policy (which may require consideration of certain factors but not others)
- the assessment method – for example, simplified approaches may be limited to linear approximations
- data availability (which may limit the number of factors that can be considered)
- objectives of the assessment (which may require a more or less complete and accurate assessment)
- available resources to conduct the assessment.

In general, users should follow the most accurate approach that is feasible, and focus on achieving higher levels of accuracy for the most significant impact categories and specific impacts included in the assessment boundary.

Historical trends and expected values in the baseline scenario

Historical data can inform the expected future values of each indicator, in both the baseline scenario and the policy scenario. Understanding the historical values of the indicator as well as the expected values in the baseline scenario is useful when estimating policy scenario values.

Timing of impacts

Changes in policy scenario values depend on the timing of expected impacts. There may be a delay between when the policy is implemented and when impacts begin to occur. Impacts may also occur before policy implementation begins because of early action taken in anticipation of the policy.

Users should assume that a policy will operate indefinitely unless an end date is explicitly embedded in the design of the policy, even if there is uncertainty about whether it will eventually be discontinued.
If the policy is limited in duration, the assessment period may include some impacts that occur during the policy implementation period and some that occur after the policy implementation period.

Users should also consider whether and how the implementation of the policy is expected to change during the assessment period. Examples are tax instruments where the tax rate increases over time, performance standards where the level of stringency increases over time, or regulations with multiple distinct phases.

In addition to estimating and reporting the full impacts of the policy over the assessment period, users can separately estimate and report impacts over any other time periods that are relevant. For example, if the assessment period is 2020–2030, users can separately estimate and report impacts over the periods 2020–2025, 2025–2030 and 2020–2030.

Barriers to policy implementation, enforcement or effectiveness

The policy scenario values should represent the values most likely to occur in the presence of the policy, which depend on assumptions relating to policy implementation, enforcement and effectiveness. Depending on what is considered most likely in a particular context, users should either

(1) estimate the maximum impacts of the policy if full implementation is most likely, or (2) discount the maximum impacts based on expected limitations in policy implementation, enforcement or effectiveness that would prevent the policy from achieving its maximum potential. For example, a policy may not achieve its full potential because of governance challenges, such as a lack of capacity, interagency coordination, public participation or accountability. Users should apply conservative assumptions if there is uncertainty about the extent of policy implementation and effectiveness.

Policy interactions

The policy assessed may interact with other implemented or adopted policies included in the baseline scenario. To accurately estimate policy scenario values and the impacts of the policy, users should determine whether the policy being assessed interacts with any policies included in the baseline scenario (in either reinforcing or overlapping ways). For example, a new municipal solar PV incentive policy may overlap with an existing national renewable energy mandate and a local energy efficiency policy. Because both existing policies are included in the baseline scenario, they reduce the energy savings achieved through the new solar policy.

If interactions with policies included in the baseline scenario exist, users should estimate the magnitude of the policy interactions when estimating policy scenario values. This enables estimation of the incremental impact of the policy being assessed relative to existing policies included in the baseline scenario.[28]

Sensitivity of indicator values to assumptions

Users should use sensitivity analysis to understand the range of possible values of key indicators and parameters, and determine which scenario is most likely. Users should also understand the range of uncertainty associated with key indicators and parameters. For more information on assessing uncertainty and sensitivity analysis, see Chapter 11.

9.3 Estimate the net impact of the policy on each indicator

After estimating policy scenario values, the last step is to estimate the net impact of the policy on each indicator. It is a *key recommendation* to estimate the net impact of the policy on each indicator by subtracting baseline values from policy scenario values, taking into account all specific impacts included in the quantitative assessment boundary (see equation 9.1). This involves estimating each specific impact within an impact category, then aggregating across all of the specific impacts to determine the net impact of the policy on each impact category, where feasible.

To do this, users should follow these steps for each indicator being estimated:

1. Estimate baseline values relating to each specific impact in the quantitative assessment boundary (as described in Chapter 8).

2. Estimate policy scenario values relating to each specific impact in the quantitative assessment boundary.

3. Subtract baseline values from policy scenario values to estimate the impact of the policy for each specific impact.

4. Aggregate across all specific impacts to estimate the total net impact of the policy on a given indicator, which represents the change in the impact category, where feasible.

5. Repeat the process for each indicator in the assessment boundary.

When aggregating across impacts, users should address any possible overlaps or interactions between impacts to avoid overestimation or underestimation of the total net impact of the policy.

Users should calculate baseline values, policy scenario values and the net impact of the policy over defined time periods (e.g. annually) and cumulatively over the quantitative assessment period.

Equation 9.1: Estimating the impact of the policy on a given indicator

For a specific impact: Estimated change due to the policy = policy scenario value for the chosen indicator – baseline value for the chosen indicator

Net impact of a policy on the chosen indicator = ∑ estimated change for each specific impact included in the assessment boundary

Note: "Net" refers to the aggregation of all specific impacts included in the assessment boundary, including both positive and negative impacts.

[28] An example of assessing policy interactions is available in Del Río et al. (2013).

It is a *key recommendation* to separately assess the impacts of the policy on different groups in society, where relevant. Examples of different groups are men and women, people of different income groups, people of different racial or ethnic groups, people of different education levels, people from different geographic regions, and people in urban versus rural locations. This allows users to understand distributional impacts on different groups, and manage trade-offs in cases where policies have positive impacts on some groups and negative impacts on others.

Equation 9.1 results in a neutral estimate of impact, which may either be an increase (positive value) or a decrease (negative value). Policy scenario values may be either higher or lower than baseline scenario values, depending on the impact being estimated. For example, if estimating the impact of a policy on air pollution, the equation will yield a positive value if the policy increases air pollution and a negative value if the policy reduces air pollution. If a policy creates jobs, the equation will yield a positive value, whereas, if a policy reduces jobs, the equation will yield a negative value. Users may interpret and communicate the result as either positive or negative or an increase or decrease, depending on the impact category and the context.

If any impacts in the quantitative assessment boundary have not been estimated, users should document and justify the exclusion, and describe the impact qualitatively (as explained in Chapter 7).

See Appendix A for an example of estimating the impact of a solar PV incentive policy. Table 9.1 summarizes the ex-ante quantification results for the

TABLE **9.1**

Estimated impact of the solar PV incentive policy on all impact categories included in the assessment

Impact category	Indicator quantified	Estimated impact (cumulative impact, 2016–2025)
Climate change mitigation	GHG emissions ($MtCO_2e$) from the electricity grid	Reduction of 307 $MtCO_2e$
Air quality/health impacts of air pollution	$PM_{2.5}$ emissions (t) from the electricity grid	Reduction of 1,177,996 t $PM_{2.5}$
	PM_{10} emissions (t) from the electricity grid	Reduction of 2,437,234 t PM_{10}
	SO_2 emissions (t) from the electricity grid	Reduction of 4,265,161 t SO_2
	NO_x emissions (t) from the electricity grid	Reduction of 4,062,057 t NO_x
	Number of premature deaths per year in India resulting from air pollution from coal plants	Reduction of 32,304 premature deaths
Energy	Renewable energy installed capacity (MW)	Increase of 40,000 MW of renewable energy capacity
Access to clean, affordable and reliable energy	Increase in number of houses/buildings/facilities with access to clean energy	Increase of 5,741,889 houses/buildings/facilities with access to clean energy
Capacity, skills and knowledge development	Number of new skilled trainees and workers on the ground	Increase of 40,060 new skilled trainees and workers
Jobs	Change in jobs (number of jobs)	Net increase of 821,102 jobs
Income	Savings in annual electricity bill for households and businesses ($)	Savings of $27,855 million
Energy independence	Reduction in coal imports (t)	Reduction of 57,770,140 t of coal

solar PV incentive policy across all impact categories included in the assessment.

Users should estimate total in-jurisdiction impacts (the net change that occurs within the implementing jurisdiction's geopolitical boundary) separately from total out-of-jurisdiction impacts (the net change that occurs outside the jurisdiction's geopolitical boundary) for each indicator, if relevant and feasible.

Users should separately estimate and report the change resulting from each specific impact included in the assessment boundary, where relevant and feasible. Users can also separately report by type of impact.

When uncertainty is high (e.g. because of uncertain baseline assumptions), users should report the net impact of the policy on a given indicator as a range of likely values, rather than as a single estimate. Chapter 11 provides guidance on uncertainty and sensitivity analysis.

9.3.1 Separate reporting based on likelihood and probability, if relevant

Each impact of the policy included in the assessment may have a different likelihood of occurrence. In Chapter 7, users categorize potential impacts based on whether they are very likely, likely, possible, unlikely or very unlikely to occur. If unlikely or very unlikely effects are included in the assessment, users should consider reporting these impacts separately from the results for very likely, likely and possible impacts. Users can also separately report impacts by each likelihood category (e.g. very likely, likely, possible) if relevant and feasible.

Where likelihood is difficult to estimate, users can report a range of values for a given impact, based on sensitivity analysis for key parameters (further described in Chapter 11). Users can additionally incorporate probability into the estimation of ex-ante policy scenario values by weighting each impact by its expected probability (e.g. 100%, 75%, 50%, 25%, 0%).

Box 9.1 gives an example of a quantitative ex-ante assessment in South Africa.

BOX **9.1**

Quantitative ex-ante impact assessment in South Africa

A landfill in Garden Route District Municipality in South Africa was recently closed because of capacity constraints, and will be replaced by a new regional waste management and landfill facility. The new landfill will not accept organic waste materials. To inform the municipality's new organic waste management plan, the South Africa Low Emission Development (SA-LED) programme supported the municipality in conducting an ex-ante assessment of the sustainable development impacts of different organic waste management options. The assessment focused on different approaches to managing abattoir waste, which is a major component of organic waste in the district. The findings are expected to inform broader organic waste management policy in the region.

Defining the baseline and policy scenarios: The baseline scenario assumed that the new regional landfill would be built without an abattoir waste management facility, and the abattoir waste would go to other regional landfills, or be discarded at the community or household level. The policy scenario assumed that the new waste management facility includes an abattoir waste management facility that uses anaerobic digestion. The study quantified the impact of building the facility with an abattoir waste management facility compared with the baseline scenario.

BOX **9.1,** continued

Quantitative ex-ante impact assessment in South Africa

Determining impact categories and indicators to assess: Table 9.2 provides examples of impact categories and indicators that were assessed.

TABLE **9.2**

Examples of assessed impact categories and indicators

Impact category	Indicator
Climate change mitigation	• Amount of CO_2e avoided (t/year)
Economic development	• Earnings gained from the project (ZAR/year) • GDP gained from the project (ZAR/year)
Jobs	• Number of short-term jobs created, disaggregated by direct (on-site) and indirect (supply chain) jobs • Number of long-term operations and maintenance (O&M) jobs created, disaggregated by direct and indirect jobs
Water saving	• Amount of water saved (t/year)
Waste generation	• Change in amount of waste sent to landfill (t/year)
Women employment	• Number of full-time, trained women employees
Youth employment	• Number of full-time, trained employees under 35 years old
Land use	• Years of landfill life saved (years)

Identifying and assessing specific impacts: Based on the included impact categories, the study identified specific impacts of the abattoir waste management facility. Each specific impact was qualitatively assessed, including its likelihood and magnitude, to determine whether it was significant. With the exception of water savings, all impacts in Table 9.2 were found to be significant. Because of data limitations, impacts on women employment and youth employment were assessed qualitatively rather than quantitatively.

BOX **9.1**, continued

Quantitative ex-ante impact assessment in South Africa

To quantify the baseline scenario, policy scenario and net impacts, the assessment used recent studies, including a municipal waste characterization study performed by SA-LED, and tools such as the International Jobs and Economic Development Impacts (I-JEDI) tool and the United States Environmental Protection Agency's Waste Reduction Model (WARM) tool. The quantitative results are shown in Table 9.3.

TABLE **9.3**

Selected quantitative results for the waste management policy

Indicator	Change
Change in GHG emissions from diverting waste to anaerobic digester	Reduction of 5,718 tCO_2e/year
Change in earnings gained from diverting waste to biopower	Increase of 2,284,016 ZAR/year
Change in GDP gained from diverting waste to biopower	Increase of 3,907,917 ZAR/year
Number of direct one-time construction jobs created in a single year	Increase of 31 jobs
Number of indirect one-time construction jobs created in a single year	Increase of 22 jobs
Number of direct long-term O&M jobs created from diverting waste to biopower	Increase of 1 job
Number of indirect long-term O&M jobs created from diverting waste to biopower	Increase of 1 job
Change in tonnes of waste sent to landfill	Reduction of 9,697 t/year
Change in lifespan of new regional landfill site	Increase of 3 years

10 Estimating impacts ex-post

Ex-post assessment is the process of estimating historical impacts of policies. It is a backward-looking assessment of impacts achieved to date. In this chapter, users estimate the impact of the policy by comparing observed policy scenario values of an indicator (based on monitored data) with ex-post baseline values (described in Chapter 8). Unlike ex-ante assessment, which involves forecasted values, ex-post assessment involves monitored or observed values. The impact of the policy (ex-post) is estimated by subtracting baseline values from policy scenario values. Users who are not quantitatively assessing impacts ex-post can skip this chapter. Sections 10.1–10.4 apply to users following the scenario method, while Section 10.5 applies to users following the comparison group method.

Checklist of key recommendations

- Recalculate baseline values (as described in Chapter 8) every time an ex-post assessment is undertaken
- Estimate the net impact of the policy on each indicator in the quantitative assessment boundary by subtracting baseline values from policy scenario values, taking into account all specific impacts included in the quantitative assessment boundary
- Separately assess the impacts of the policy on different groups in society, where relevant
- For users following the comparison group method, identify an equivalent comparison group for each impact category in the assessment boundary, and collect data from the comparison group and the policy group over the assessment period for each indicator included in the assessment boundary

10.1 Update baseline values or ex-ante assessment (if relevant)

Figure 10.2 illustrates ex-post estimation of impacts. In contrast to ex-ante policy scenario values, which are forecasted based on assumptions, ex-post policy scenario values are based on data collected

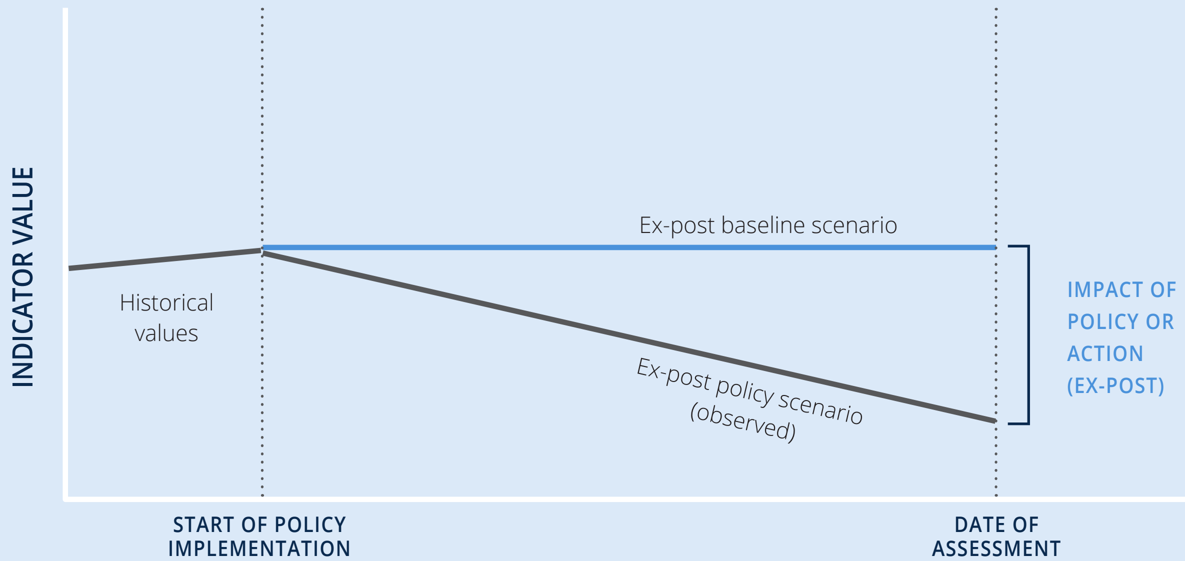
FIGURE **10.1**

Overview of steps in the chapter

Update baseline or ex-ante assessment, if relevant (Section 10.1) → Choose assessment method for each indicator (Section 10.2) → Estimate policy scenario values for each indicator (Section 10.3) → Estimate net impact of the policy on each indicator (Section 10.4) → Use comparison group method, if relevant (Section 10.5)

FIGURE **10.2**

Estimating impacts ex-post



Source: Adapted from WRI (2014).

during the time the policy was implemented. Users carrying out an ex-post assessment may estimate ex-post policy scenario values either before or after estimating ex-post baseline values.

It is a *key recommendation* to recalculate baseline values (as described in Chapter 8) every time an ex-post assessment is undertaken. The ex-post baseline scenario should include all other policies with significant impacts that were implemented either before the implementation of the policy being assessed or after the implementation of the policy being assessed, but before the ex-post assessment.

The baseline scenario should also be recalculated to include updates to all non-policy drivers, based on their observed values over the assessment period. Non-policy drivers should be considered in the baseline scenario if they are exogenous to the assessment – that is, if they are not affected by the policy being assessed.

If an ex-ante assessment for the policy was previously carried out, the same method can be used for the ex-post assessment, by replacing the forecasted indicator values (ex-ante) with observed indicator values (ex-post). Alternatively, users can apply a different method to estimate policy scenario values. Users should choose the method that yields the most accurate results. If both an ex-ante and an ex-post assessment are carried out for the same policy at different times, each assessment will likely yield different estimates of the impacts of the policy, since the observed (ex-post) indicator values will likely differ from assumptions forecasted in the ex-ante scenario.

10.2 Choose assessment method for each indicator

This section provides a list of ex-post assessment methods that can be used to estimate the impacts of a policy (see Table 10.1). The list is not exhaustive, and users can classify methods differently depending on the individual context. Users can also use a combination of the approaches listed in Table 10.1. The ICAT website[29] provides specific examples of tools and models to support impact quantification.

[29] https://climateactiontransparency.org/icat-toolbox/sustainable-development

Users should select methods based on a combination of factors, such as data availability; the type of policy and sector; the number of actors influenced by the policy; the number of interacting policies; and the capacity, resources and expertise available for each method.

Users should ensure consistency in the methods used to estimate baseline values and policy scenario values for each indicator, to ensure that the estimated impact reflects underlying differences between the two scenarios, rather than differences in method. If it is not feasible or appropriate to use the same method in a given situation, users should justify why different methods have been used.

When selecting methods to estimate impacts ex-post, users should determine the desired level of accuracy to be achieved. In general, users should follow the most accurate approach that is feasible.

TABLE **10.1**

Examples of ex-post assessment methods

Method	Description
Collection of data from affected participants, facilities or actors	Indicator values in the policy scenario are determined using data collected from affected participants, facilities or other actors. Data-collection methods may include monitoring of parameters (e.g. metering of energy consumption), collection of expenditure or billing data (e.g. purchase records), or sampling methods.
Deemed estimates method	The *change* in indicator values (rather than the policy scenario value of indicators) is estimated using previously estimated effects of similar policies. This involves collecting data on the number of actions taken (e.g. number of buildings that install rooftop solar PV) and applying default values for the estimated impact or other relevant parameter per action taken (e.g. average reduction in grid-connected electricity use per building that installs solar PV). The deemed estimate may be based on published studies, equipment specifications, surveys or other methods. Deemed estimates are used as a lower-cost method for policies that are homogeneous across policy contexts, such that deemed estimates from other contexts are representative of the policy being assessed. Deemed estimates can be complemented by sampling the affected participants or sources to determine whether the estimates are sufficiently accurate and representative. In this approach, the impact is estimated directly, without subtracting baseline values from policy scenario values. Baseline values may be estimated as a subsequent step by adding or subtracting the deemed estimates from observed policy scenario values.
Monitoring of indicators	Indicator values in the policy scenario are monitored using sector or subsector activity changes. In this case, the user may have limited or no information on end use or stock statistics, but may have information on changes in relevant indicators for a sector (e.g. transportation, buildings) or subsector (e.g. space heating in buildings). Policy scenario indicator values should be compared with baseline indicator values to estimate the change.
Economic modelling	The *change* in indicator values (rather than the policy scenario value of indicators) is estimated by using econometric models, regression analysis, extended modelling such as input–output analysis with price elasticities, or computable general equilibrium models. These types of models are most appropriate for estimating economic impacts or estimating other types of impacts from fiscal policies, such as taxes or subsidies. Economic models may specify that a dependent variable (the indicator being assessed) is a function of various independent variables, such as the policy being assessed, other policies and various non-policy drivers (e.g. prices, price elasticities of fuels, economic activity, population). By doing so, models can control for various factors that affect the impact category other than the policy being assessed.

Source: Adapted from WRI (2014).

10.3 Estimate policy scenario values for each indicator

Ex-post policy scenario values are based on data collected during the time the policy is implemented. Users should first assess whether the specific impacts identified in Chapter 6 actually occurred. This may include assessing the degree of policy implementation to ensure that the policy was implemented as planned, including assessing the extent of enforcement and non-compliance, if relevant and feasible.

Users should then update the impacts identified, based on observed data, before estimating each impact. To estimate certain impacts, users may find it useful to conduct surveys with consumers or businesses affected by the policy, or use results from similar policy assessments, if the conditions are similar enough for valid comparisons.

Users should report the policy scenario values for each indicator being estimated, and the methods, assumptions and data sources used to calculate policy scenario values.

10.4 Estimate net impact of policy for each indicator

The last step is to estimate the net impact of the policy. It is a *key recommendation* to estimate the net impact of the policy on each indicator by subtracting baseline values from policy scenario values, taking into account all specific impacts included in the quantitative assessment boundary (see equation 10.1). This involves estimating each specific impact within an impact category, then aggregating across all the specific impacts to determine the net impact of the policy on each impact category, where feasible.

To do so, users should follow these steps for each indicator being estimated:

1. Estimate baseline values relating to each specific impact in the quantitative assessment boundary (as described in Chapter 8).
2. Determine policy scenario values relating to each specific impact in the quantitative assessment boundary.
3. Subtract baseline values from policy scenario values to estimate the impact of the policy for each specific impact.
4. Aggregate across all specific impacts to estimate the total net impact of the policy on a given indicator, which represents the change in the impact category, where feasible.
5. Repeat the process for each indicator in the assessment boundary.

When aggregating across impacts, users should address any possible overlaps or interactions between impacts to avoid overestimation or underestimation of the total net impact of the policy.

Users should calculate baseline values, policy scenario values and the net impact of the policy over defined time periods, such as annually and cumulatively over the quantitative assessment period.

Equation 10.1: Estimating the impact of the policy on a given indicator

For a specific impact: Estimated change due to the policy = policy scenario value for the chosen indicator – baseline value for the chosen indicator

Net impact of a policy on the chosen indicator = ∑ estimated change for each specific impact included in the assessment boundary

"Net" refers to the aggregation of all specific impacts included in the assessment boundary, including both positive and negative impacts.

It is a *key recommendation* to separately assess the impacts of the policy on different groups in society, where relevant. Examples of different groups are men and women, people of different income groups, people of different racial or ethnic groups, people of different education levels, people from different geographic regions, and people in urban versus rural locations. This allows users to understand distributional impacts on different groups, and manage trade-offs in cases where policies have positive impacts on some groups and negative impacts on others.

Equation 10.1 results in a neutral estimate of impact, which may either be an increase (positive value) or a decrease (negative value). Policy scenario values may be either higher or lower than baseline scenario values, depending on the impact being estimated and the nature of the policy. Users may interpret and communicate the result as either positive or negative or an increase or decrease, depending on the impact category and the context.

If any impacts in the assessment boundary have not been estimated, users should document and justify the exclusion, and describe the impact qualitatively (as described in Chapter 7).

See Appendix A for an example of estimating the impact of a solar PV incentive policy.

Users should estimate total in-jurisdiction impacts (the net change that occurs within the implementing jurisdiction's geopolitical boundary) separately from total out-of-jurisdiction impacts (the net change that occurs outside the jurisdiction's geopolitical boundary) for each indicator, if relevant and feasible.

Users should separately estimate and report the change resulting from each specific impact included in the assessment boundary, where relevant and feasible. Users can also separately report by type of impact.

When uncertainty is high (e.g. because of uncertain baseline assumptions), users should report the net impact of the policy on a given indicator as a range of likely values, rather than as a single estimate. Chapter 11 provides guidance on uncertainty and sensitivity analysis.

10.4.1 Combining ex-ante and ex-post assessments

Ex-ante and ex-post assessment may be combined in a "rolling monitoring" approach. Under this approach, the forecast provided by the ex-ante assessment is continually overwritten with the results from ex-post assessment, which allows comparison of the original expectations and the final results. By combining ex-ante and ex-post data, rolling monitoring can demonstrate the impacts that have been initiated up to a certain date (through ex-ante assessment), the impacts that have been achieved up to a certain date (through ex-post assessment), and the impacts that have been achieved (ex-post) compared with the ex-ante estimates.

10.5 Use the comparison group method to estimate impacts (if relevant)

This section provides guidance on using the comparison group method to estimate the impacts of a policy.

As outlined in Chapter 8, users can use the comparison group method to define the baseline scenario when carrying out an ex-post assessment. The comparison group method cannot be used for ex-ante assessments, since comparative data for the comparison group and policy group during policy implementation cannot be obtained before policy implementation.

The comparison group method involves comparing one group or region affected by a policy with an equivalent group or region that is not affected by that policy. For users following the comparison group method, it is a *key recommendation* to identify an equivalent comparison group for each impact category in the assessment boundary, and collect data from the comparison group and the policy group over the assessment period for each indicator included in the assessment boundary. Any impacts in the assessment boundary that have not been estimated should be documented and described qualitatively, with justification.

Figure 10.3 provides an overview of key steps.

FIGURE **10.3**

Overview of steps for using the comparison group method

10.5.1 Identify the policy group and comparison group

The first step is to identify the policy group (the group or region affected by the policy) and the comparison or control group (an equivalent group or region not affected by the policy). The policy group and comparison group may be groups of people, facilities, companies, jurisdictions, sectors or other relevant groups.

Ideally, the policy group and the comparison group should be equivalent in all aspects except for the existence of the policy for the policy group and absence of the policy for the comparison group. The most robust way to ensure that two groups are equivalent is to implement a randomized experiment – for example, by randomly assigning one subset of entities to participate in a programme and the other subset to not participate in the programme.

"Equivalent" means that the comparison group should be the same as, or similar to, the policy group in terms of:[30]

- **geography** – for example, facilities in the same city, subnational region or country
- **time** – for example, facilities built within the same time period
- **technology** – for example, facilities using the same technology
- **other policies** – for example, facilities subject to the same set of policies and regulations, except for the policy being assessed
- **non-policy drivers** – for example, facilities subject to the same external trends, such as the same changes in economic activity, population and energy prices.

When identifying a potential comparison group, users should collect data from both the policy group and the comparison group before the policy is implemented to determine whether the groups are equivalent. Users should ensure that the entities in the comparison group are not directly or indirectly affected by the policy.

If the groups are similar but not equivalent, statistical methods can be used to control for certain factors that differ between the groups (for examples, see Box 10.1). If the groups are not sufficiently equivalent, the comparison group method will yield misleading results, so users should follow the scenario method instead (described in Chapter 8).

10.5.2 Collect data from the policy group and comparison group

Users should collect data from both the policy group and the comparison group for each indicator included in the assessment boundary. Users should collect data from both groups at multiple points in time to account for changes that occur over time. At a minimum, users should collect data from both groups before and after the policy is implemented (in the policy group), so that the two groups can be compared during both the pre-policy period and the policy implementation period.

Either top-down or bottom-up data (see Section 8.3.3) may be used. To collect bottom-up data, representative sampling may be used to collect data from a large number of individual entities or facilities. Appropriate statistical sampling procedures should be used, and the sample size should be large enough to draw valid statistical conclusions.

10.5.3 Estimate the impact of the policy

After data are collected, users should determine values without the policy (from the comparison group) and values with the policy (from the policy group). In rare cases where the policy group and comparison group are equivalent, the outcomes of each group can be compared directly. A statistical test (such as a t-test) should be employed to ensure that the difference in values cannot be attributed to chance. If the difference between the two groups is statistically significant, the difference can be attributed to the existence of the policy, rather than to other factors.

In most cases, differences are expected to exist between the groups. If material differences exist that may affect the outcome, users should use statistical methods to control for variables other than the policy that differ between the non-equivalent groups. Such methods are intended to address selection bias and isolate the impact of the policy being assessed. See Box 10.1 for examples of methods that may be used.

[30] Adapted from WRI (2014).

BOX **10.1**

Examples of statistical methods for estimating impacts and controlling for factors that differ between groups

Multiple regression analysis involves including data for each relevant driver that may differ between the groups (e.g. economic activity, population, energy prices) as explanatory variables in a regression model, as well as proxies for other relevant policies (other than the policy being assessed) that may differ between the two groups. If the expanded regression model shows a statistically significant effect of the policy being assessed, the policy can be assumed to have an effect on the policy group, relative to the comparison group. Statistical significance refers to the certainty that the difference between two outcomes is unlikely to be a result of random chance.

Difference-in-difference methods compare two groups over two periods of time: a first period when neither the policy group nor the comparison group implements a given policy, and a second period when the policy group implements the policy and the comparison group does not. This method estimates the difference between the groups before policy implementation (A1 – B1 = X), the difference between the two groups after policy implementation (A2 – B2 = Y), and the difference between the two differences (Y – X) as a measure of the change attributable to the policy.

Matching methods are statistical approaches for making two groups (a policy group and a comparison group) more equivalent, when random assignment is not possible.

Source: Adapted from WRI (2014).

11 Assessing uncertainty

This chapter provides an overview of concepts and procedures for understanding and evaluating the uncertainty of the assessment. Uncertainty can be assessed either qualitatively or quantitatively. This chapter is relevant to both qualitative and quantitative assessment of impacts.

Checklist of key recommendations

- Assess the uncertainty of the assessment results, either qualitatively or quantitatively
- For quantitative assessments, conduct a sensitivity analysis for key parameters and assumptions in the assessment

11.1 Introduction to uncertainty analysis and sensitivity analysis

Understanding uncertainty is important for properly interpreting and communicating the results of the assessment. Uncertainty analysis refers to a systematic procedure to quantify and/or qualify the uncertainty associated with the impact assessment results. Identifying, documenting and assessing uncertainty can help users and stakeholders understand the level of confidence they can have in the results and identify the areas of the assessment that contribute most to uncertainty. Users should identify and track key uncertainty sources throughout the assessment process. Identifying, assessing and managing uncertainty are most effective when done during, rather than after, the assessment process.

Sensitivity analysis is a useful method to test the robustness of the assessment results. It involves varying the value of key parameters (or combinations of parameters) to determine the impact of such variations on the overall results. Key parameters are those that are highly variable, highly uncertain or most likely to significantly affect assessment results. Sensitivity analysis can be conducted in combination with uncertainty analysis to prioritize efforts for improving data. If a parameter is determined to be highly uncertain and sensitive, users should prioritize collecting better data for that parameter. If a parameter is certain and insensitive, there is less need for improving data quality. Figure 11.2 illustrates how to prioritize data improvement based on uncertainty and sensitivity.

Understanding uncertainty can help users understand whether to apply conservative assumptions. As explained in Chapter 3, accuracy should be pursued as far as possible, but, once uncertainty cannot be reduced to an acceptable level, conservative estimates should be used.

FIGURE **11.1**

Overview of steps in the chapter

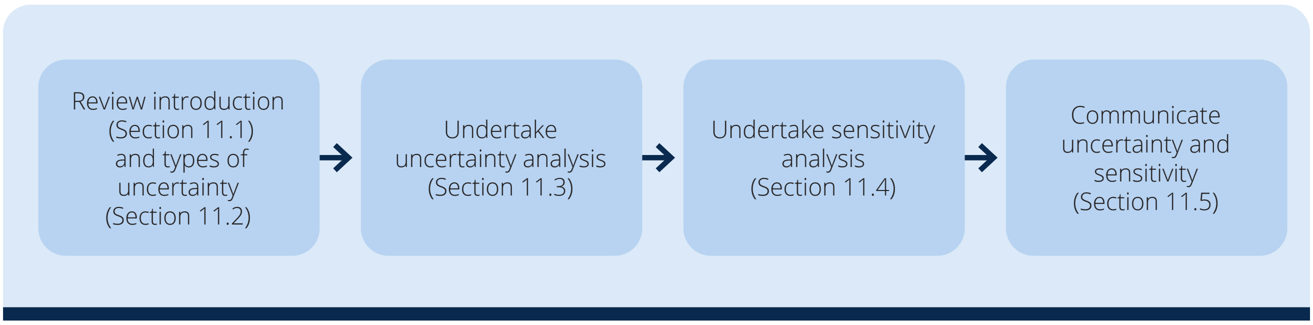
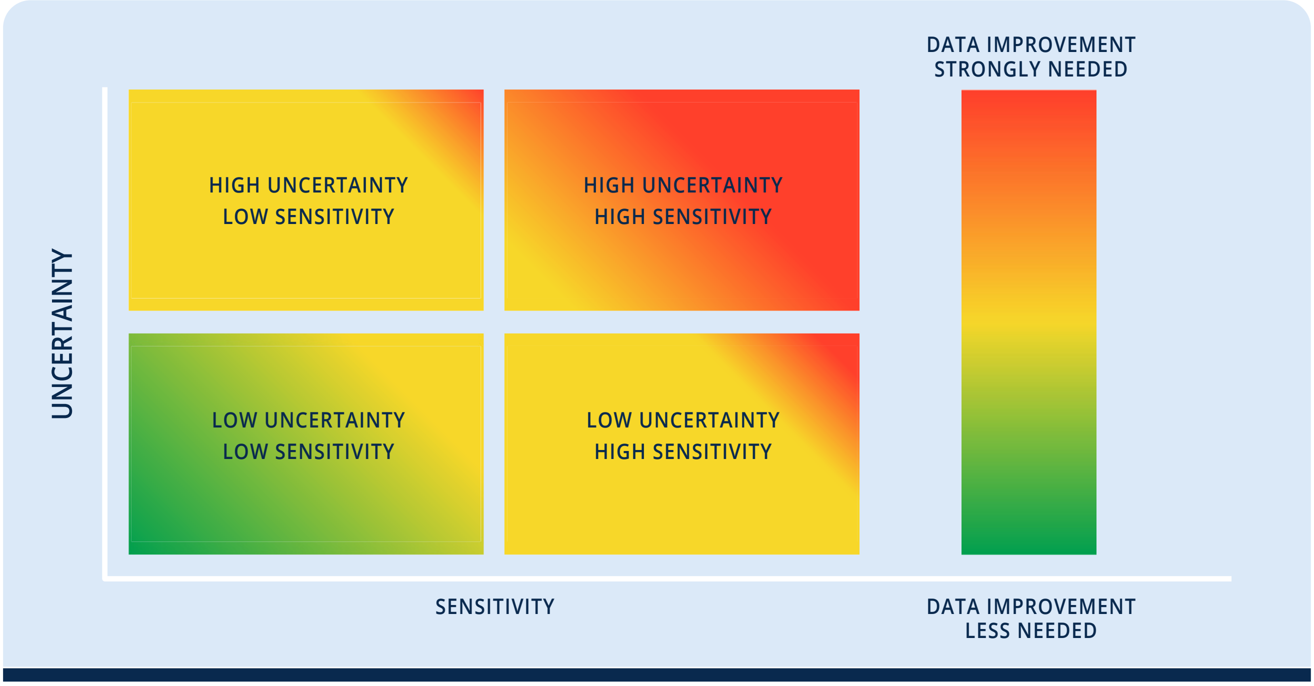


FIGURE **11.2**

Identifying where data improvement is needed in relation to uncertainty and sensitivity



11.2 Types of uncertainty

This chapter classifies uncertainty into three categories according to the source of uncertainty: parameter uncertainty, scenario uncertainty and model uncertainty. The categories are not mutually exclusive, but they can be evaluated and reported in different ways. Table 11.1 summarizes each type of uncertainty.

11.2.1 Parameter uncertainty

Parameter uncertainty represents the imperfect knowledge of true parameter values in an assessment method or model. It may arise from insufficient data, measurement errors, inaccurate approximation, or geographical and temporal variability. For example, wind speed may be used as an input parameter to model the dispersion and concentration of $PM_{2.5}$. The test equipment will deliver wind speeds with a certain uncertainty. Meanwhile, wind speed may vary every second, but only limited numbers of values (e.g. one value per hour) will be used to model the dispersion of $PM_{2.5}$.

If parameter uncertainty can be determined, it can typically be represented as a probability distribution of possible values that include the chosen value used in the assessment. Individual parameter uncertainties can be propagated to provide a quantitative measure of the uncertainty of the assessment results, which may be represented in the form of a probability distribution.

11.2.2 Scenario uncertainty

Ex-ante assessments involve baseline scenarios and policy scenarios that describe how conditions are expected to develop in the future, while ex-post assessments involve baseline scenarios that describe how conditions would have developed in the past if a policy were not implemented. These scenarios are based on a set of uncertain assumptions, which creates scenario uncertainty. To identify the influence of these assumptions on the results, users should undertake a sensitivity analysis for key parameters in the assumptions (described in Section 11.4).

TABLE **11.1**

Types of uncertainty

Type of uncertainty	Description
Parameter	Uncertainty regarding whether a parameter value used in the assessment accurately represents the true value of the parameter
Scenario	Uncertainty of the calculated result due to various assumptions made in the baseline and policy scenarios
Model	Imperfect representation of modelling approaches, equations or algorithms to reflect the real world

Source: Adapted from WRI (2014).

11.2.3 Model uncertainty

Simplifying the real world into a numerical model introduces inaccuracies, and different models are likely to yield different results. For example, various life cycle impact assessment models can be used to assess the environmental impacts associated with producing solar PV panels. Each model is likely to yield different results, leading to model uncertainty. The extent of uncertainty can be estimated by comparing the results of different models. Users should acknowledge model uncertainties and report model limitations qualitatively.

11.3 Uncertainty analysis

The two primary approaches to assessing uncertainty are:

- qualitative uncertainty analysis
- quantitative uncertainty analysis.

It is a *key recommendation* to assess the uncertainty of the assessment results, either qualitatively or quantitatively. Only qualitative uncertainty analysis is relevant to assessing the uncertainty of a qualitative impact assessment. Either approach can be used to assess the uncertainty of a quantitative impact assessment. Quantitative uncertainty analysis can provide more robust results than qualitative analysis. Reporting quantitative uncertainty estimates also gives greater clarity and transparency to stakeholders.

Users should select an approach based on the objectives of the assessment, the level of accuracy needed to meet stated objectives, data availability, and capacity and resources. Depending on the methods used and data availability, users may not be able to assess the uncertainty of all parameters in the assessment method(s). Users should assess the uncertainty of all parameters for which this assessment is feasible. Where quantitative uncertainty analysis is not possible or appropriate, uncertainty should be assessed and described qualitatively.

11.3.1 Qualitative uncertainty analysis[31]

Qualitative uncertainty analysis can be done in a variety of ways. This section outlines a structured approach, which involves characterizing the level of confidence of the results based on:

- the quantity and quality of evidence (robust, medium or limited)
- the degree of agreement of the evidence (high, medium or low).

The level of confidence is a metric that can be expressed qualitatively to indicate certainty in the validity of a parameter value or result. (The qualitative confidence level described in this section is distinct from statistical confidence and should not be interpreted in statistical terms.)

[31] This section is adapted from IPCC (2010).

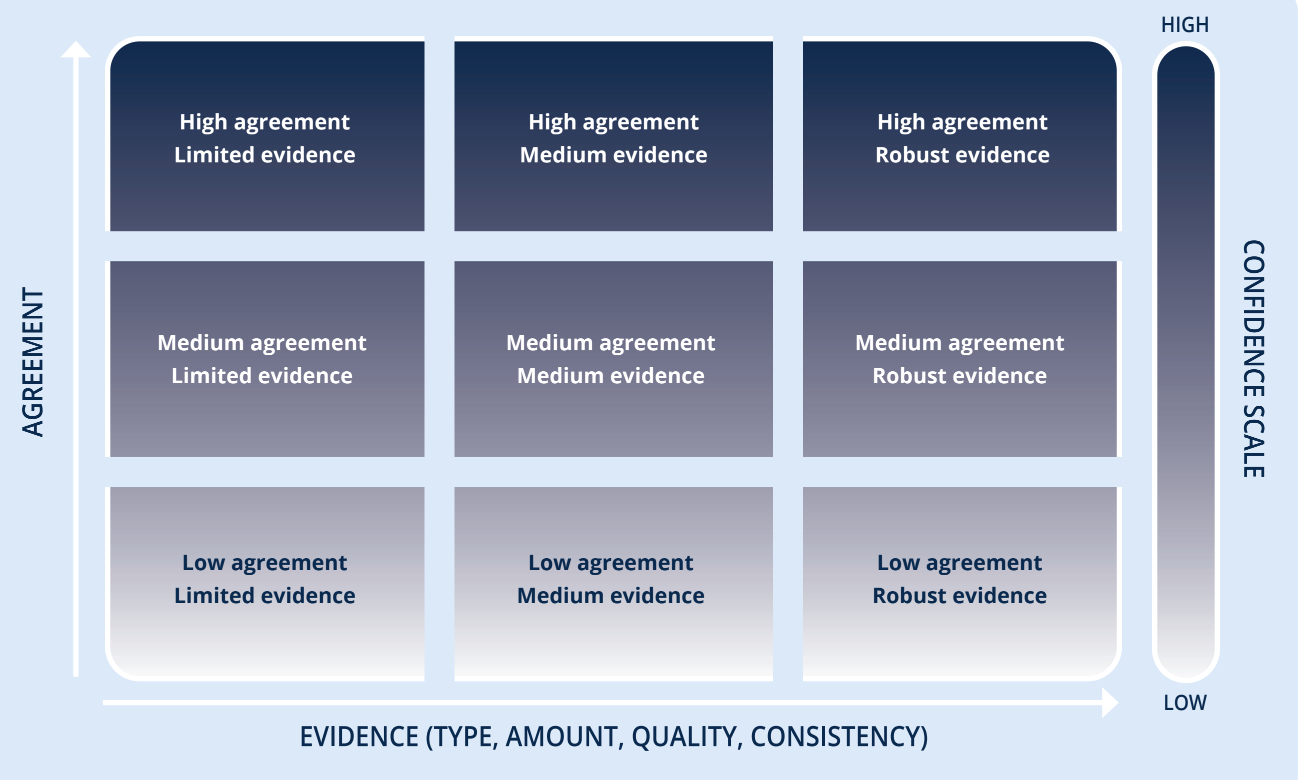
When characterizing parameter uncertainty, evidence refers to the sources available for determining a parameter value. Evidence should be assessed with regard to both its quantity and quality. Quantity and quality of evidence can be classified as robust, medium or limited. Evidence should be considered robust when there is a large quantity of high-quality evidence. Evidence should be considered medium when there is a medium quantity of medium-quality evidence. Evidence should be considered limited when there is a small quantity of low-quality evidence. High-quality evidence adheres to principles of research quality. Low-quality evidence shows deficiencies in adhering to principles of research quality. Medium-quality evidence is a mix of high-quality and low-quality evidence.[32]

The degree of agreement of evidence is a measure of consensus or consistency across available sources for a parameter value or result. The degree of agreement can be classified as high, medium or low. As a rule of thumb, high agreement means that all sources had the same conclusion; medium agreement means that some sources had the same conclusion; and low agreement means that most of the sources had different conclusions. This step is not applicable if only one source is available.

A level of confidence provides a qualitative synthesis of the user's judgment about the result, integrating both the evaluation of evidence and the degree of agreement in one metric. Figure 11.3 depicts summary statements for evidence and agreement, and their relationship with confidence; confidence increases as evidence and agreement increase. The level of confidence can be considered very high, high, medium, low or very low. In the best case (very high confidence), the evidence found should be sourced from multiple credible, independent institutions. Presentation of findings with "low" and "very low" confidence should be reserved for areas of major concern, and the reasons for their presentation

FIGURE **11.3**

Summary statements for evidence and agreement, and their relationship with confidence



Source: WRI (2014), adapted from IPCC (2010).

[32] Adapted from DFID (2014).

should be explained. The confidence level of individual parameters, models and scenarios should be aggregated to provide a level of confidence for the overall assessment, if feasible.

11.3.2 Quantitative uncertainty analysis

If feasible, users should carry out a quantitative uncertainty analysis to characterize the uncertainty of key parameters. This involves estimating the uncertainty of individual parameters (single parameter uncertainty), then aggregating the uncertainties for a given indicator as a whole (propagated parameter uncertainty). Propagated parameter uncertainty is the combined effect of each parameter's uncertainty on the total result.

Users should estimate uncertainty at a specified confidence level, preferably 95%. Users should use the best available estimates from a variety of methods and approaches, such as a combination of measured data, published information, model outputs and expert judgment.

Approaches to quantifying the uncertainty of individual parameters include the following:

- Default uncertainty estimates for parameters reported in literature.
- Probability distributions and standard deviations.
    - » This method is feasible and preferred when a large amount of data is available for a given parameter. In such cases, it is possible to generate a probability distribution and other statistical values, such as standard deviations, which can be propagated to the uncertainty of the final output.
- Uncertainty factors for parameters reported in literature.
    - » One application of uncertainty factors is in environmental assessments relating to risk and safety. For example, when assessing the toxicity impact of a certain chemical, experiments may be conducted on a small group of people. To extrapolate the test results to a larger group, an uncertainty factor is applied to ensure maximum protection and safety. This method is especially relevant when conservative methods are applied.
- Pedigree matrix approach from life cycle assessment (based on qualitative data quality indicators in Table 8.7).
    - » This method provides a way to quantify uncertainties based on a qualitative assessment of data. Five criteria are provided in Table 8.7 to assess data quality from different perspectives. For each criterion, a value is assigned by the practitioner to describe the data quality. These values can then be translated into the standard deviation of the data set.[33]
- Survey of experts to generate upper- and lower-bound estimates.
- The user's expert judgment (based on as much data as available) or other approaches.

Once the uncertainties of individual parameters have been estimated, they may be aggregated to provide uncertainty estimates for the entire assessment for an indicator. Approaches to combining uncertainties include:

- error propagation equations – an analytical method used to combine the uncertainty associated with individual parameters from a single scenario. Equations involve estimates of the mean and standard deviation of each input
- Monte Carlo simulation – a form of random sampling used for uncertainty analysis that shows the range of likely results based on the range of values for each parameter and probabilities associated with each value. To perform Monte Carlo simulation, input parameters must be specified with probability distributions. The input parameters are varied at random but restricted by the given probability distribution for each parameter. Repeated calculations produce a probability distribution of the predicted output values, reflecting the propagated uncertainty of the various parameters. This method gives comprehensive results, but is more resource- and time-intensive. Simple Monte Carlo simulations can be done using the Crystal Ball tool in Microsoft Excel.

[33] For more information, see Weidema and Wesnaes (1996).

Further references on quantitative uncertainty analysis
For more detailed guidance on the methods outlined in this section, see the following references:

- IPCC (2000). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*[34]
- IPCC (2006). Chapter 3, Uncertainties. In *2006 IPCC Guidelines for National Greenhouse Gas Inventories*, vol. 1[35]
- World Resources Institute (WRI) and World Business Council for Sustainable Development (WBCSD) (2003). *Aggregating Statistical Parameter Uncertainty in GHG Inventories: Calculation Worksheets*[36]
- WRI and WBCSD (2003). *GHG Protocol Guidance on Uncertainty Assessment in GHG Inventories and Calculating Statistical Parameter Uncertainty*[37]
- WRI and WBCSD (2011). *Quantitative Inventory Uncertainty*[38]
- WRI and WBCSD (2011). *Uncertainty Assessment Template for Product GHG Inventories*.[39]

11.4 Sensitivity analysis

A sensitivity analysis involves varying the value of key parameters (or combinations of parameters) to determine the impact of such variations on the overall results. Sensitivity analysis is a useful tool to understand differences resulting from methodological choices and assumptions, and to explore model sensitivities to input parameters.

For quantitative assessments, it is a *key recommendation* to conduct a sensitivity analysis for key parameters and assumptions in the assessment. Sensitivity analysis is expected to be most relevant for quantitative impact assessments, but may also be useful for certain qualitative impact assessments.

To conduct a sensitivity analysis, users should adjust the value of key parameters to determine the impact of such variations on the overall results. Since an assessment may include many impact categories and involve many parameters, users should conduct sensitivity analysis only on key parameters.

Users should consider reasonable variations in parameter values. Not all parameters need to be subjected to both negative and positive variations of the same magnitude, but they should be varied based on what is considered reasonable. Past trends may be a guide to determining the reasonable range. As a general rule, variations in the sensitivity analysis should at least cover a range of +10% and –10% (unless this range is not deemed reasonable under the specific circumstances).

Sensitivity analysis can be conducted in several ways. One simple method is to assess the relative sensitivity for one parameter at a time, according to equation 11.1.

Equation 11.1: Assessing the sensitivity of a parameter

$$S = \frac{\Delta output / output}{\Delta input / input}$$

In the equation, *S* represents the relative sensitivity of the assessment output to the specific input parameter. *Input* and *output* represent the original values. Δ*input* is the marginal change in the input parameter, which should represent a reasonable expected change. Δ*output* is the corresponding marginal change in the output parameter. Using this equation, users can compare the sensitivity of the output in response to different input parameters.

See Box 11.1 for an example of applying equation 11.1 to determine which of various parameters is most sensitive.

34 Available at: www.ipcc-nggip.iges.or.jp/public/gp/english.

35 Available at: www.ipcc-nggip.iges.or.jp/public/2006gl.

36 Available at: www.ghgprotocol.org.

37 Available at: www.ghgprotocol.org.

38 Available at: www.ghgprotocol.org.

39 Available at: www.ghgprotocol.org.

BOX **11.1**

Example of sensitivity analysis

Table 11.2 illustrates a sensitivity analysis of three key parameters for a solar PV incentive policy. It is assumed that there are 186,306,371 grid-connected households in India, with an annual consumption of 900 kilowatt-hours (kWh) electricity per year per household. In the original policy scenario, 10% of existing grid-connected households are expected to adopt rooftop solar PV systems and will be able to rely on solar for the entire household electricity demand. The other 90% of grid-connected households will rely on a combination of grid-connected electricity and back-up diesel generators for electricity, assuming that 90% (810 kWh) is supplied by the grid and 10% (90 kWh) is supplied by a diesel-fuelled power generator when blackouts occur.

The three chosen parameters for sensitivity analysis are annual electricity consumption per household, the percentage of households that will adopt solar PV, and the percentage of electricity supplied by grid for the households that use combined electricity supply, assuming that the remaining electricity demand is met by diesel-fuelled power generators. Table 11.2 illustrates a scenario in which each parameter value is set to a reasonable assumption. The table also shows calculation of the output – in this case, changes in emissions for each scenario. This example specifically focuses on PM_{10}. Combined, this information enables calculation of relative sensitivity. The input, output and sensitivity analysis results are presented below.

TABLE **11.2**

Sensitivity analysis of estimated PM_{10} emissions

Parameter	Annual electricity consumption	Percentage of households that adopt solar PV	Percentage of electricity supplied by grid
Input			
Original value (kWh)	900	10%	90%
Scenario value (kWh)	1,800	80%	50%
Δinput/input	100%	700%	–44%
Output: emissions reduction			
Original value (t PM_{10})	300,817	300,817	300,817
Scenario value (t PM_{10})	601,635	71,886	171,695
Δoutput/output	100%	–76%	–43%
Sensitivity analysis result			
Relative sensitivity	100%	–11%	97%

This sensitivity results show that, of the three parameters, PM_{10} emissions are more sensitive to annual electricity consumption and percentage of electricity supplied by the grid, and less sensitive to percentage of households that adopt solar PV. This information can be used to prioritize future data-collection efforts.

11.5 Communicating uncertainty and sensitivity

Reporting information about uncertainty helps users and stakeholders assess the accuracy and uncertainty of the reported results, to inform how the information should be used. It is important to properly communicate the results, since the estimate of policy impact may not be very accurate, depending on the methods, assumptions and data sources that were used to assess the impacts.

Uncertainty can be reported in many ways, including qualitative descriptions of uncertainty sources and quantitative representations, such as error bars, histograms and probability density functions. Users should provide as complete a disclosure of uncertainty information as possible.

Users should report a quantitative estimate or qualitative description of the uncertainty of the results. They should also report the range of results from sensitivity analysis for key parameters and assumptions.

Users should report the range of possible outcomes based on different parameter values (representing upper and lower bounds of plausible values) to indicate the level of uncertainty. When uncertainty is high, users should consider reporting a range of values around the average or most likely value, rather than only a single value. Users should transparently report the full range of likely values, rather than reporting only upper-bound or lower-bound values.

Users should also use an appropriate number of significant figures, depending on the uncertainty of the results, to avoid overstating the precision of the results.

Users should make a thorough yet practical effort to communicate key sources of uncertainty in the results, including key parameters and assumptions that have high uncertainty. If feasible, users should report both qualitative and quantitative uncertainty information. They should also describe their efforts to reduce uncertainty in future revisions of the assessment, if applicable.