Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Tracking Progress of the Mitigation Commitments of Nationally Determined Contributions (NDCs)

Presentation: Data needs and dealing with insufficient data

Dominic Sheldon

Ricardo Energy & Environment

UN environment programme | copenhagen climate centre

supported by UNOPS

UN environment programme

gef

Supported by:

Federal Ministry for Economic Affairs and Climate Action | Federal Foreign Office

on the basis of a decision by the German Bundestag

IKI INTERNATIONAL CLIMATE INITIATIVE

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Guidance: IPCC 2006 Guidelines

CHAPTER 5

TIME SERIES CONSISTENCY

during the period of overlap. In this case, the emissions or removals associated with the new method are estimated according to Equation 5.1.[2]

EQUATION 5.1
RECALCULATED EMISSION OR REMOVAL ESTIMATE COMPUTED USING THE OVERLAP METHOD

$$y_0 = x_0 \bullet \left(\frac{1}{(n-m+1)} \bullet \sum_{i=m}^{n} \frac{y_i}{x_i} \right)$$

Where:

y_0 = the recalculated emission or removal estimate computed using the overlap method

x_0 = the estimate developed using the previously used method

y_i and x_i are the estimates prepared using the new and previously used methods during the period of overlap, as denoted by years m through n

A relationship between the previously used and new methods can be evaluated by comparing the overlap between only one set of annual estimates, but it is preferable to compare multiple years. This is because comparing only one year may lead to bias and it is not possible to evaluate trends.

Figure 5.1 shows a hypothetical example of a consistent overlap between two methods for the years in which both can be applied. In Figure 5.2 there is no consistent overlap between methods and it is not *good practice* to use the overlap technique in such a case.

Other relationships between the old and new estimates may also be observed through an assessment of overlap. For example, a constant difference may be observed. In this case, the emissions or removals associated with the new method are estimated by adjusting the previous estimate by the constant amount equal to the average difference in the years of overlap.

Figure 5.1 Consistent overlap

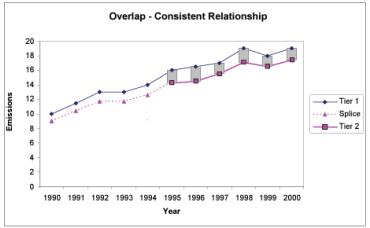
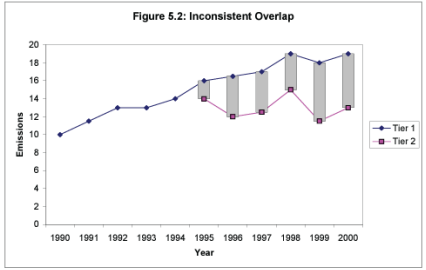


Figure 5.2 Inconsistent overlap



5.3.3.2 SURROGATE DATA

The surrogate method relates emissions or removals to underlying activity or other indicative data. Changes in these data are used to simulate the trend in emissions or removals. The estimate should be related to the statistical data source that best explains the time variations of the category. For example, mobile source emissions may be related to trends in vehicle distances travelled, emissions from domestic wastewater may be related to population, and industrial emissions may be related to production levels in the relevant industry. See Chapter 2, Approaches to Data Collection.

In its simplest form, the estimate will be related to a single type of data as shown in Equation 5.2:

EQUATION 5.2
EMISSION/REMOVALS TREND ESTIMATES USING SURROGATE PARAMETERS

$$y_0 = y_t \bullet \left(s_0 / s_t \right)$$

Where:

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Issues with data availability

Periodic data:

- Natural resource or environmental statistics, such as national forest inventories and waste statistics, may not cover the entire country on an annual basis.
    - Instead, they may be carried out at intervals such as every fifth or tenth year, or region-by-region, implying that national level estimates can only be directly obtained once the inventory in every region has been completed.
- When data are available less frequently than annual, several issues arise.
    - First, the estimates need to be updated each time new data become available, and the years between the available data need to be recalculated.
    - The second issue is producing inventories for years after the last available data point and before new data are available.
        - In this case, new estimates should be extrapolated based on available data, and then recalculated when new data become available.

Changes and gaps in data availability:

- A change in data availability or a gap in data is different from periodically available data because there is unlikely to be an opportunity to recalculate the estimate at a later date using better data.
- In some cases, countries will improve their ability to collect data over time, so that higher tier methods can be applied for recent years, but not for earlier years.
    - This is particularly relevant to categories in which it is possible to implement direct sampling and measurement programs because these new data may not be indicative of conditions in past years.
- Some countries may find that the availability of certain data sets decreases over time as a result of changing priorities within governments, economic restructuring, or limited resources.
- Some countries with economies in transition no longer collect certain data sets that were available in the base year, or if available these data sets may contain different definitions, classifications and levels of aggregation.

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Data types and gaps

	Input Data Type	Examples of Input Data	Typical Sources of Data	Typical Gaps
Historical Data	- Economic indicators	GDP, unemployment rates, inflation rates	National statistical offices, World Bank, IMF	Inconsistent time series, Missing data for certain years
	- Demographic data	Population size, age distribution, urbanization rates	National census bureaus, UN Department of Economic & Social Affairs	Incomplete datasets, Lack of granular data (regional, age-group)
	- Energy usage trends	Energy consumption by sector, renewable energy usage	National energy agencies, IEA	Insufficient granularity, Under-reported sectors
	- Changes in land use	Land use changes, deforestation rates, urban expansion	National environmental agencies, remote sensing databases	Time-lags in reporting, Uncertainty in measurement techniques
Historical Emissions	- Activity data	Energy production and consumption, industrial activities	National environmental agencies, industry reports	Lack of source-specific data, Inconsistent methodologies
	- Emission factors	GHG emission factors for various sectors and activities	IPCC guidelines, national research institutions	Use of default factors, Lack of country-specific data
Non-emissions Data	- Environmental data	Deforestation rates, air and water quality data	National environmental agencies, satellite imagery	Sporadic data collection, Methodological inconsistencies
	- Socio-economic data	Income levels, educational attainment, health indicators	National statistical offices, World Bank	Insufficient data on vulnerable groups, Lag in data reporting
Projected Data	- Drivers	Predicted economic growth, demographic changes, energy price fluctuations	National economic agencies, international financial institutions	Uncertainties in projections, Over-reliance on historical trends
	- Policies	Upcoming government policies, industry standards for emission reductions	Government policy documents, industry reports	Uncertainty in policy implementation, Lack of detailed policies

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Types of data

Type	Description	Example of GHG Emission Projections
Measured Data	Data obtained through direct measurement, such as emissions directly measured from a smokestack.	Using sensors to directly measure CO2 emissions from a power plant smokestack.
Modelled Data	Data derived from quantitative models, representing processes like emissions from landfills or livestock.	Using simulation models to project methane emissions from a landfill site in the next 10 years.
Calculated Data	Data obtained by multiplying activity data by an emission factor.	Calculating future vehicle emissions by multiplying the projected number of vehicles by the emission factor per vehicle.
Estimated Data	Proxy or alternative data sources used to fill data gaps when more accurate or representative data is lacking.	Estimating agricultural GHG emissions based on land use patterns and applying emission factors from similar regions.

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

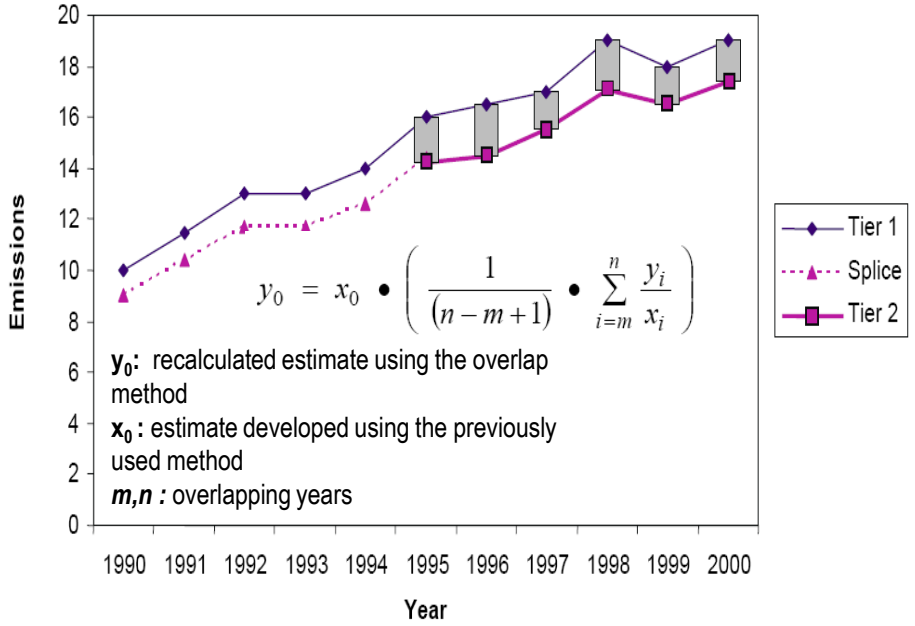
Splicing technique: Overlap

Use case: The overlap method is used when implementing a new data collection method without historical data.

Method: It creates a consistent time series by aligning older estimates with new methods during years where data overlap.

Outcome: This adjusted series ensures a smooth transition to the new method without losing historical continuity.

The formula adjusts original emissions estimates to align with a new method by averaging the ratios of new to old estimates during years where both data sets overlap.

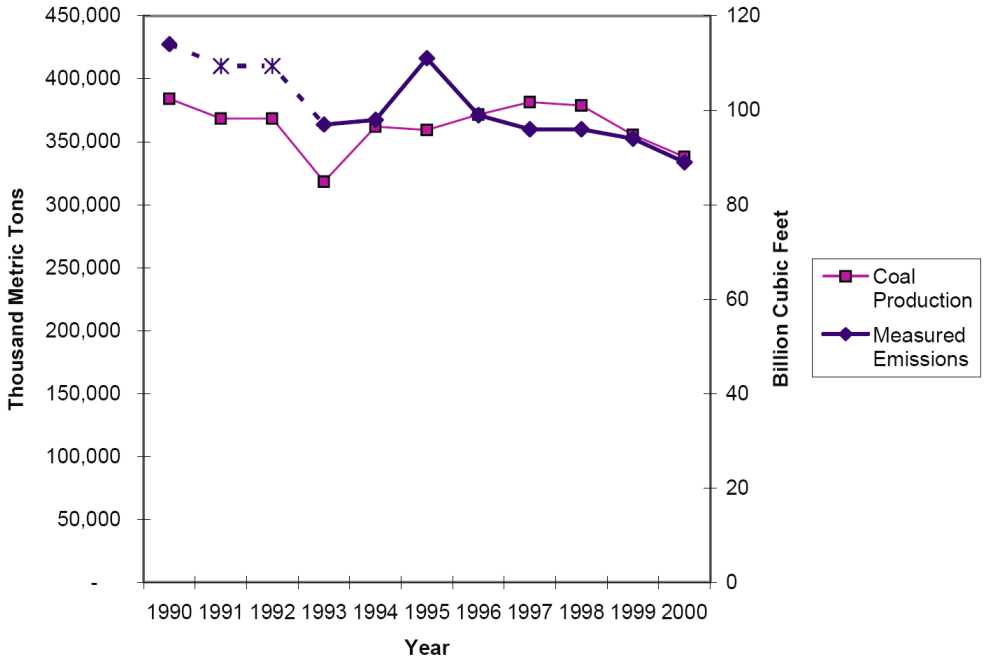


Splicing technique: Surrogate

Use Case: The surrogate method is applied when direct data is unavailable.

Method: It uses related activity or indicators to estimate emissions, drawing on statistical correlations to fill data gaps.

Outcome: This approach enhances time series reliability by allowing for more accurate estimations without direct data.

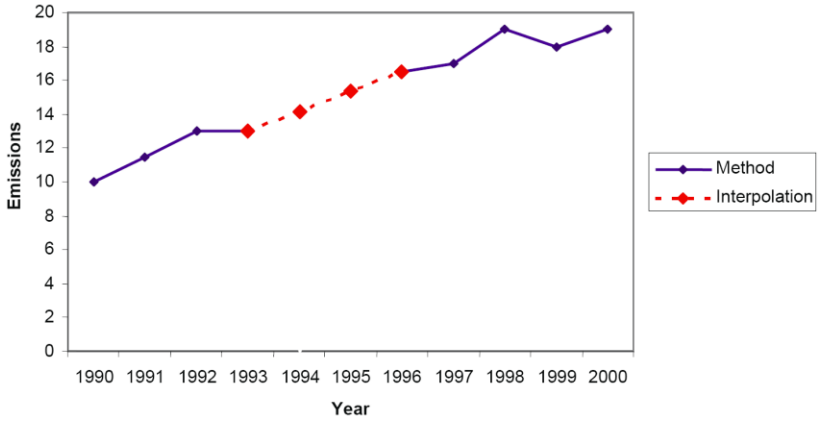


Splicing technique: Interpolation

Use case: The interpolation technique is used when you have emissions data for some years but not for others.

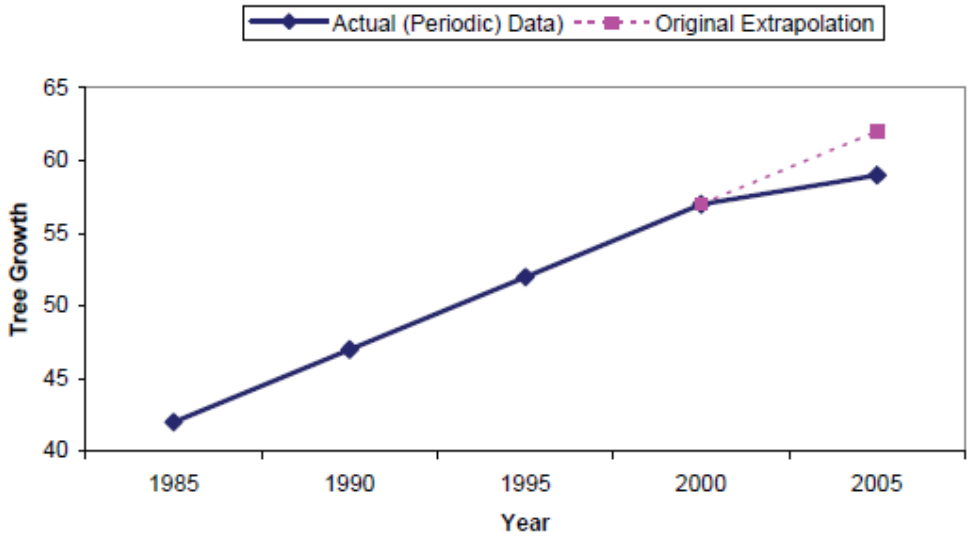
Method: It estimates the missing data by drawing a straight line — linear interpolation — from known data points, assuming emissions change at a consistent rate.

Outcome: This helps to create a smoother, more complete emissions trend line.



Splicing technique: Extrapolation

- **Use case:** Extrapolation is used when data at the start or end of a series is missing.
- **Method:** It extends the trend from known data to estimate these points, using a consistent pattern seen in the data.
- **Outcomes:** This method fills in gaps temporarily, acknowledging that these estimates grow less certain the farther they extend from known data, and are best updated when more information becomes available.



Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Summary of splicing techniques

Approach	Applicability	Comments
Overlap	Data necessary to apply both the previously used and the new method must be available for at least one year, preferably more.	• Most reliable when the overlap between two or more sets of annual estimates can be assessed. • If the trends observed using the previously used and new methods are inconsistent, this approach is not good practice.
Surrogate Data	Emission factors, activity data or other estimation parameters used in the new method are strongly correlated with other well-known and more readily available indicative data.	• Multiple indicative data sets (singly or in combination) should be tested in order to determine the most strongly correlated. • Should not be done for long periods.
Interpolation	Data needed for recalculation using the new method are available for intermittent years during the time series.	• Estimates can be linearly interpolated for the periods when the new method cannot be applied. • The method is not applicable in the case of large annual fluctuations.
Trend Extrapolation	Data for the new method are not collected annually and are not available at the beginning or the end of the time series.	• Most reliable if the trend over time is constant. • Should not be used if the trend is changing (in this case, the surrogate method may be more appropriate). • Should not be done for long periods.
Other Techniques	The standard alternatives are not valid when technical conditions are changing throughout the time series (e.g., due to the introduction of mitigation technology).	• Document customised approaches thoroughly. • Compare results with standard techniques.

Transport sector example

1. **Identify Data Availability**:
    1. Start by assessing the availability of data for vehicle types, fuel consumption, and distance traveled annually.
2. **Determine Data Gaps**:
    1. Identify missing data points, such as specific years, vehicle types, or fuel usage.
3. **Select Splicing Technique**:
    1. **If data are missing for initial years but available later**:
        1. Use the **Overlap Technique** if there's a method change and overlapping data exist.
        2. Otherwise, **Trend Extrapolation** might be suitable to project backward.
    2. **If data are periodically available**:
        1. Consider **Surrogate Data** if related indicators (e.g., fuel sales data) can approximate missing values.
        2. **Interpolation** can be used if you have data points before and after the gap, assuming stable trends.
    3. **For non-calendar year data**:
        1. Adjust collection periods to match the rest of the series to maintain consistency.
4. **Implement Technique**:
    1. Apply the selected technique based on the decision tree pathway.
5. **Validate and Adjust**:
    1. Compare estimated data with known data or surrogate indicators to validate accuracy.
    2. Adjust methodology as necessary based on validation results.

- Example Dataset: Transport Sector GHG Emissions
- Data includes vehicle kilometers traveled (VKT) and fuel efficiency (FE) by vehicle type per year.
- Gaps may exist in early years, for certain vehicle types, or in FE data due to methodological changes.

Agriculture sector – Livestock example

1. **Identify Data Availability**:
    1. Assess the availability of livestock population, feed type, and manure management practices over the years.
2. **Determine Data Gaps**:
    1. Identify missing data points for specific livestock categories, feed conversion efficiencies, or manure management systems for certain years.
3. **Select Splicing Technique**:
    1. **If initial years' data are missing**:
        1. Use **Overlap Technique** if methodologies changed and some overlapping data exist.
        2. **Trend Extrapolation** might be suitable to estimate emissions backward from available data.
    2. **If data are available at intervals**:
        1. **Surrogate Data** can be used if related indicators (e.g., feed sales, land use changes) can approximate missing emissions data.
        2. **Interpolation** is suitable if data exist before and after the gap, assuming trends are stable.
    3. **For irregular data collection periods**:
        1. Adjust the periods to align with the series to maintain consistency.
4. **Implement Technique**:
    1. Apply the chosen technique based on the pathway identified in the decision tree.
5. **Validate and Adjust**:
    1. Validate estimated data against known data points or use surrogate indicators to check accuracy.
    2. Adjust the method as necessary based on these validations.

Example Dataset: Livestock GHG Emissions

- Data includes the number of animals, types of feed, and manure management practices, with gaps in early years or for specific livestock categories.

Industrial sector – Cement example

1. **Data Availability**: Assess data on cement production volume, fuel types used, and emission factors over time.
2. **Data Gaps Identification**: Find missing annual production data or emission factors for certain years.
3. **Splicing Technique Selection**:
    1. **Initial Data Gaps**: Use **Overlap Technique** for method changes with some overlapping data, or **Trend Extrapolation** for backward estimates from available data.
    2. **Periodic Data**: Apply **Surrogate Data** with indicators like economic output of the construction sector, or **Interpolation** for stable trends between known data points.
4. **Implementation**: Apply the selected technique.
5. **Validation and Adjustment**: Validate estimates with known data or indicators, adjusting methods as needed.

- Example Dataset: Annual cement production volume and emission factors, with gaps in early production data or emission factors for specific process changes.

Ministry of Environment
Greenhouse Gas Inventory and Research Center

Partnership on Transparency in the Paris Agreement

CBIT-GSP
CLIMATE TRANSPARENCY

Food and Agriculture Organization of the United Nations

Thank you!

Dominic Sheldon – Dominic.Sheldon@Ricardo.com
Ricardo

UN environment programme

gef

Supported by:
Federal Ministry for Economic Affairs and Climate Action
Federal Foreign Office
on the basis of a decision by the German Bundestag

IKI INTERNATIONAL CLIMATE INITIATIVE