

Presentation: Dealing with Inventory Data Gaps and time-series Consistency: Using Splicing Techniques

Training on preparation and reporting of results of national GHG inventories under the ETF of the Paris Agreement

Morogoro , Tanzania 23 – 27 September 2024

Executed by:



Funded by:



Implemented by:



Sekai Ngarize
Climate Transparency and Forestry Advisor
sngarize47@gmail.com

Presentation Outline

- Reporting Guidance outlined in Modalities, Procedures and Guidelines (MPGs- Paris Agreement Rulebook) for Time Series Consistency
- What is Time Series Consistency?
- Issues related data availability and time series consistency
- IPCC Splicing Techniques
- Practical Exercises



Executed by:



Funded by:



Implemented by:



copenhagen
climate centre

I. National inventory report of anthropogenic emissions by sources and removals by sinks of greenhouse gases

A. Definitions (17)

B. National circumstances and institutional arrangements (18-19)

C. Methods

1. Methodologies, parameters and data (20-24)
2. Key category analysis (25)
3. Time-series consistency and recalculations (26-28)
4. Uncertainty assessment (29)
5. Assessment of completeness (30-33)
6. Quality assurance/quality control (34-36)

D. Metrics (37)

E. Reporting guidance (38)

1. Information on methods and cross-cutting elements (39-46)
2. Sectors and gases (47-56)
3. Time series (Para 57-58)

Executed by:



Guidance from MPGs

26. To ensure time-series consistency, each Party **should use the same methods and a consistent approach** to underlying activity data and emission factors for each reported year.

27. Each Party **should use surrogate data, extrapolation, interpolation and other methods** consistent with **splicing techniques contained in the IPCC guidelines** referred to in paragraph 20 above to estimate missing emission values resulting from lack of activity data, emission factors or other parameters in order to ensure a consistent time series.

28. Each Party **shall perform recalculations in accordance with the IPCC guidelines** referred to in paragraph 20 above, ensuring that changes in emission trends are not introduced as a result of changes in methods or assumptions across the time series.

Decision 18/CMA.1

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



Time Series Consistency



- The time series is a central component of the greenhouse gas inventory because it provides information on historical emissions trends and tracks the effects of strategies to reduce emissions at the national level (IPCC 2006 GL).
- An inventory is not just an estimate of a single year. It includes estimates for a number of years (time series of estimates)

All emissions estimates in a time series should be estimated consistently, which means that as far as possible, the time series should be calculated using;

1. **same method** and
2. **data sources** in all years.

- Using different methods and data in a time series could introduce bias because the estimated emission trend will reflect not only real changes in emissions or removals but also the pattern of methodological refinements.
- However, it is not always possible to use the same method and data sets for the entire time series due to a lack of data

Executed by:



copenhagen climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



Federal Ministry
for Economic Affairs
and Climate Action



INTERNATIONAL
CLIMATE
INITIATIVE

on the basis of a decision
by the German Bundestag

Periodic Data

- **Some data may not be available on an annual basis.**
- 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.

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.
- 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.

Executed by:



Funded by:

copenhagen
climate centre



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



Federal Ministry
for Economic Affairs
and Climate Action



INTERNATIONAL
CLIMATE
INITIATIVE

on the basis of a decision
by the German Bundestag

Non-Calendar Year Data

It is good practice to **use calendar year data** whenever the data are available.

If calendar year data are unavailable:

Other types of annual year data (e.g., non-calendar fiscal year data e.g., April – March) can be used provided that it is used **consistently over the time series** and the collection period for the data is documented.

Similarly, different collection periods can be used for different emission and removal categories, again provided that the **collection periods are used consistently over time** and documented this is acceptable.

It is good practice to use the **same collection periods consistently** over the time series to avoid bias in the trend.

The data should be **corrected where possible to represent the calendar year**.

If uncorrected data are used, it is good practice for the inventory compiler to make consistent use of either calendar year data or fiscal year data for all years in the time series.

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



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

Executed by:



copenhagen climate centre

Funded by:



Implemented by:



Partnership on Transparency in the Paris Agreement



Ministry of the Environment Government of Japan

Supported by:



on the basis of a decision by the German Bundestag



INTERNATIONAL CLIMATE INITIATIVE



Splicing Techniques

- Splicing: combining or joining of more than one method or data series to form a complete time series
 - Methodological change and refinement
 - Data gaps
- The 2006 IPCC Guidelines provide several splicing techniques
 - Overlap
 - Surrogate
 - Interpolation
 - Extrapolation
- Selecting a technique requires an evaluation of the specific circumstances and a determination of the best option for the particular case

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



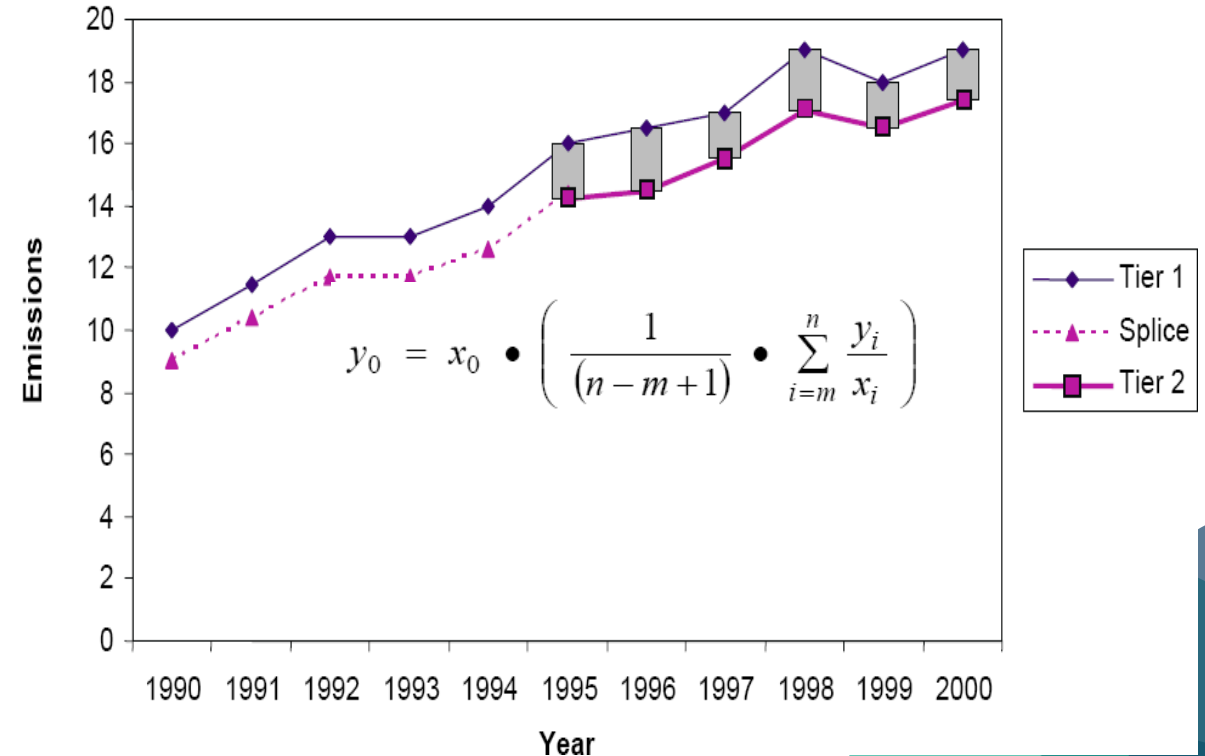
Overlap Technique

Use case: The overlap method is used when a new method is introduced but historical data are not available for early years in the time series (e.g. implementing a higher tier methodology)

Method: Develop a time series based on the relationship (or overlap) observed between the previously used and new method during the years when both can be used

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.



Executed by:



Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



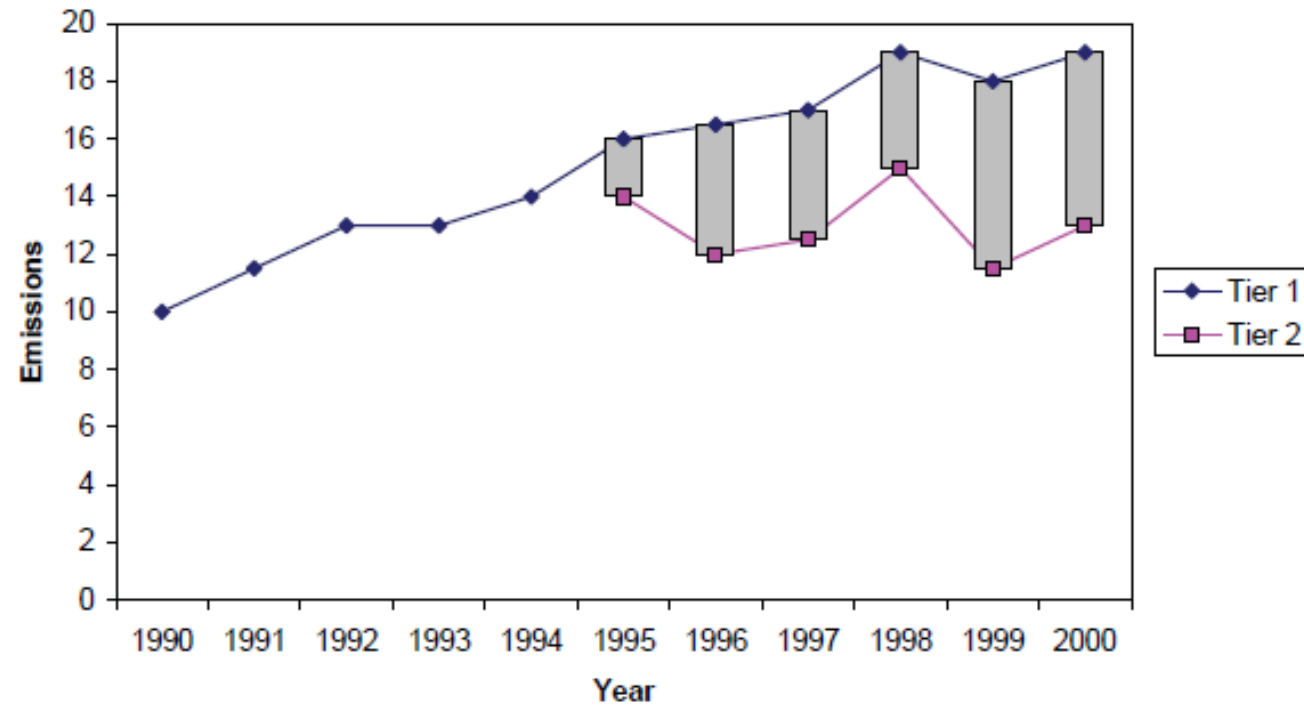
Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag

Overlap: Inconsistent Relationship



Executed by:



Funded by:



Implemented by:



copenhagen
climate centre

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.

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

y_0 : recalculated estimate using the overlap method
 x_0 : estimate developed using the previously used method
 m, n : overlapping years

Source: IPCC 2019 Refinement

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



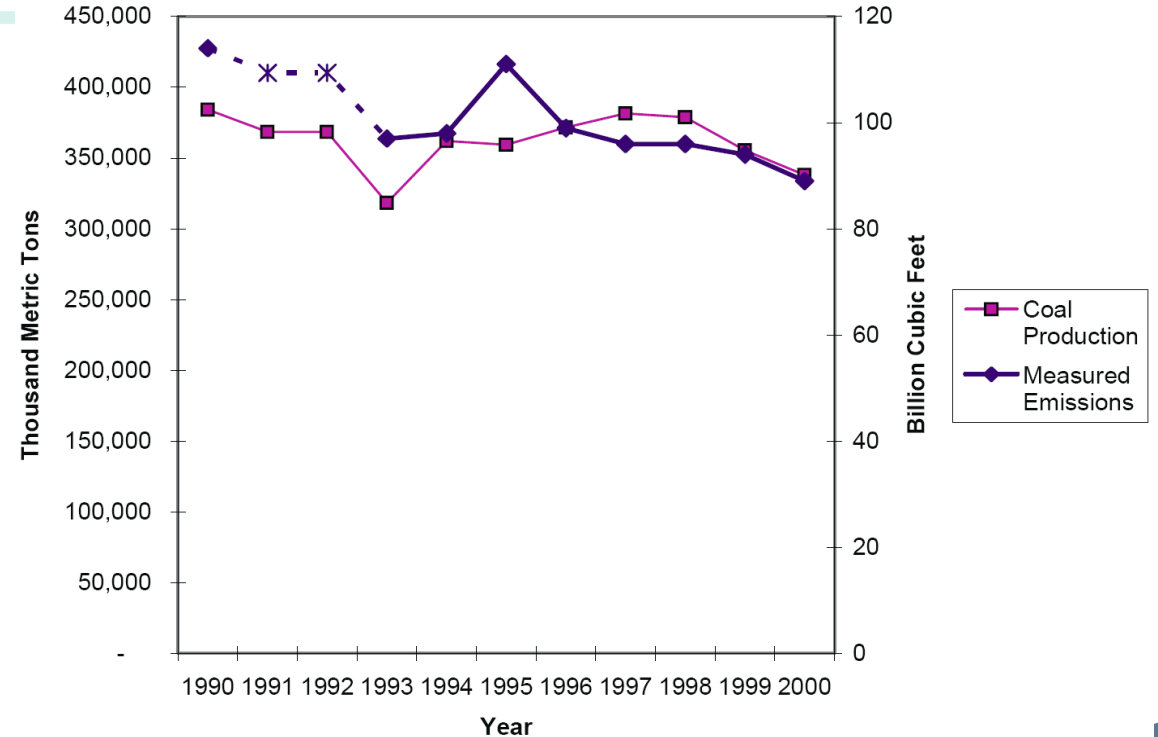
Splicing technique: Surrogate

Use Case: The surrogate method relates emissions or removals to underlying activity or other indicative data

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

Data (statistical) that is related to the emission (emissions may be proportional to production, vehicle distances travelled and population etc.)

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

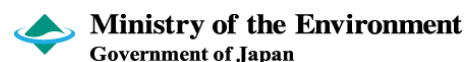


$$y_0 = y_t \bullet \left(s_0 / s_t \right)$$

y : emission/removal estimate in years 0 and *t*
s : surrogate statistical parameter in years 0 and *t*

Although the relationship between emissions/removals and surrogate can be developed on the basis of data for a single year, the use of **multiple years** might provide a better estimate.

Executed by: Funded by: Implemented by:



Splicing technique: Surrogate

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
- Industrial emissions may be related to production levels in the relevant industry

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

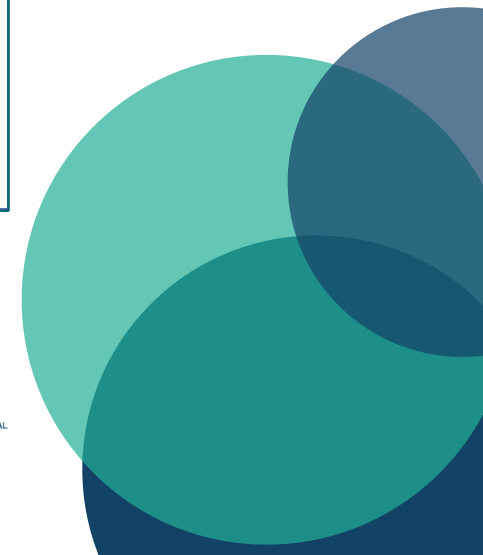
Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



Examples of Surrogate Data

TABLE 5.0 (NEW)
EXAMPLES OF SURROGATE DATA BY SECTOR

Energy	IPPU	AFOLU	Waste
<ul style="list-style-type: none"> • Gross-domestic product • Population statistics • Vehicle fleet • Fuel sales data (taking into account import/export) • Annual Income 	<ul style="list-style-type: none"> • Commodity Production statistics • Gross-domestic product (of each specific category where available) • Plant-specific parameters 	<ul style="list-style-type: none"> • Crop sales data (taking into account import/export) • Crop productivity and harvested area • Milk production data • Animals slaughtered • Gross-domestic product of each specific category • Fuelwood consumption data (taking into account import/export) 	<ul style="list-style-type: none"> • Gross-domestic product • Population statistics • Annual Income • Protein intake data • Ratio of domestic/ industrial wastewater

Source: IPCC 2019 Refinement

Executed by:



Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE

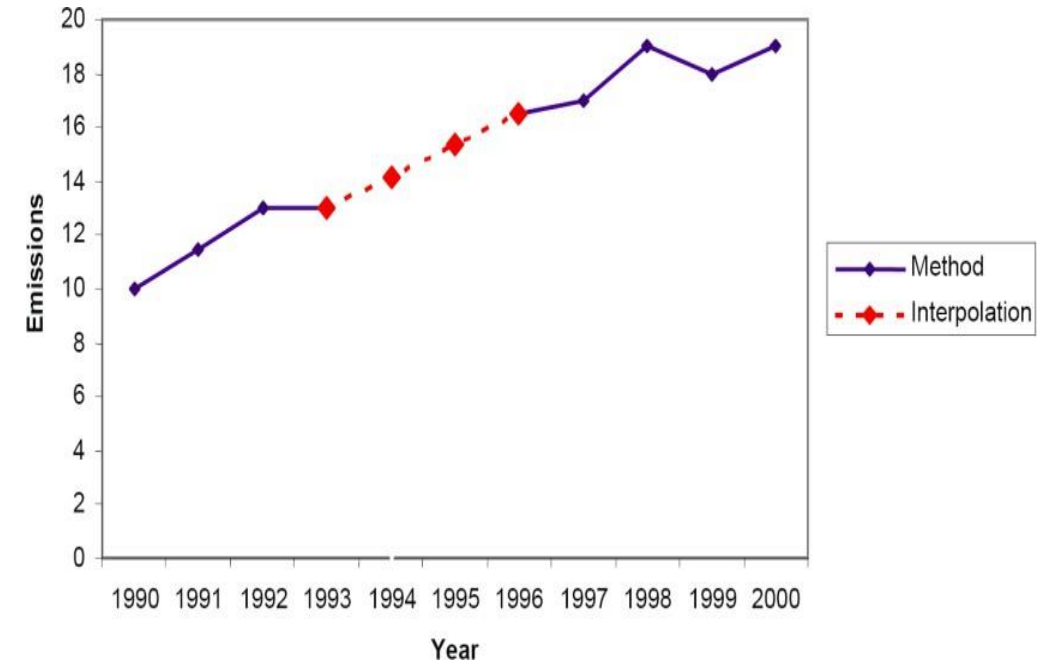


Interpolation

Use case: The interpolation technique is used when detailed statistics is collected every few years, or it is impractical to conduct detailed surveys on an annual basis. In addition, if information on the general trends or underlying parameters is available, then the surrogate data can be used

Method: Estimates for the intermediate years in the time series can be developed by interpolating between the detailed estimates when:

- Overall trend appears stable
- Actual emissions are not substantially different from the values estimated by interpolation
- Therefore, the methodology 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.



$$Y_t = Y_{start} + \frac{(T_t - T_{start})}{(T_{end} - T_{start})} * (Y_{end} - Y_{start})$$

It is good practice to compare interpolated estimates with surrogate data as a QA/QC check.

Executed by:

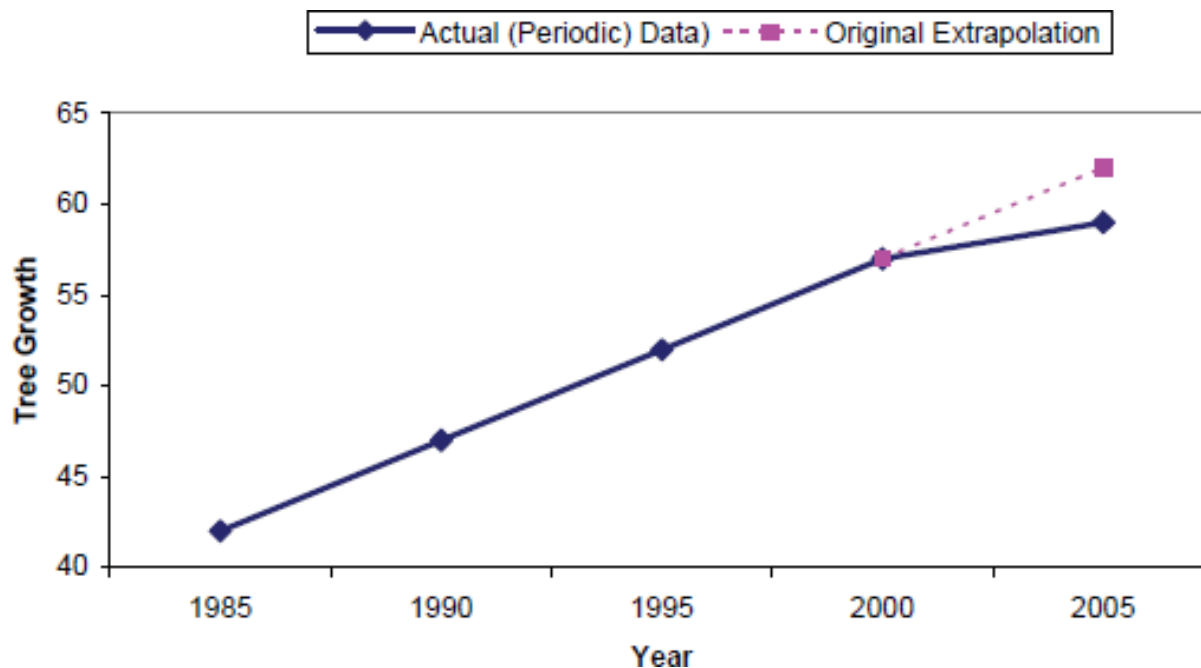
Funded by:

Implemented by:

Supported by:

Extrapolation

- **Use case:** Extrapolation is used when data for the base year or the most recent year are not available
- **Method:** It extends the trend from known data to estimate these points, using a consistent pattern seen in the data.
- The data can be extrapolated assuming that the trend in emissions/removals remains constant over the period of extrapolation
- **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.



$$Y_t = Y_{t-1} + (Y_{t-1} - Y_{t-2})$$

Extrapolation should not be used if the change in trend is not constant over time and should also not be used over long periods of time.

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE

Summary of Splicing Techniques

Source: IPCC 2019 Refinement

Splicing Technique	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.	<ul style="list-style-type: none"> • 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 <i>good practice</i>.
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.	<ul style="list-style-type: none"> • 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.	<ul style="list-style-type: none"> • 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.	<ul style="list-style-type: none"> • 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 applied for long periods.
Non-Linear Trend Analysis (Interpolation/ Extrapolation)	In cases where time series consistency is best represented by multiplicative (exponential) rather than additive (linear) relationships	<ul style="list-style-type: none"> • Most reliable for trend analysis of model outputs. • Applicable in the case of large annual fluctuations with observed high standard deviations (see Box 3.0a, Chapter 3, Volume 1 of the <i>2019 Refinement</i> for guidance on standard deviation values).
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).	<ul style="list-style-type: none"> • Document customised approaches thoroughly. • Compare results with standard techniques.

Executed by:



Funded by:



Implemented by:



copenhagen climate centre

Quality of time Series and Documentation

Comparison of the results of multiple approaches where it is possible to use more than one approach

- Plotting and comparing the results of splicing techniques on a graph is useful
- If alternative splicing methods produce different results, should consider which result is most realistic

Comparison of recalculated estimates with previous estimates can be a useful check on the quality of a recalculation

- However, higher tier methods may produce different trends than lower tier methods because they more accurately reflect actual conditions

All recalculations and measures taken to improve time series consistency should be documented and reported

- Reason of the recalculation
- Effect of the recalculation on the time series
- Splicing techniques used

Executed by:



copenhagen
climate centre

Funded by:



Implemented by:



Partnership on Transparency
in the Paris Agreement



Ministry of the Environment
Government of Japan

Supported by:



on the basis of a decision
by the German Bundestag



INTERNATIONAL
CLIMATE
INITIATIVE



Thank you for your attention!

For more information:
<https://climate-transparency-platform.org>

Sekai Ngarize
Climate Transparency and Forestry Advisor
sngarize47@gmail.com

Executed by:



Funded by:



Implemented by:



copenhagen
climate centre