

EU-China Dialogue and Cooperation on ETS-related Policies and Measures 中欧碳市场政策对话与合作项目

TtT Intensive Training (Day 3): Data Gaps



17/09/2025



Funded by the
European Union
项目由欧盟资助



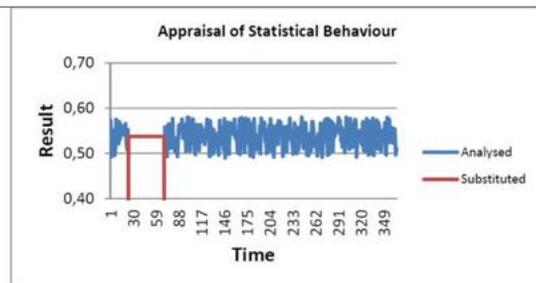
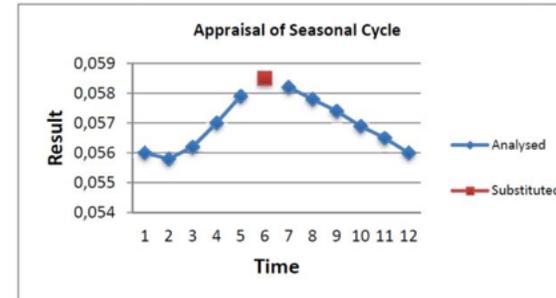
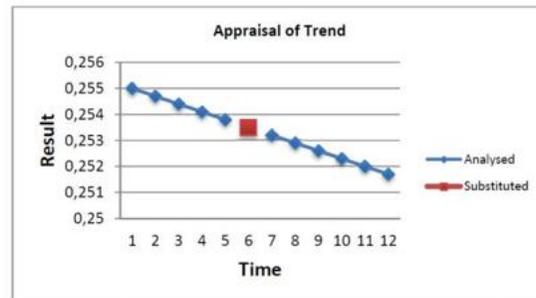
Implemented by:
项目执行方:



Background: Data Gaps

- Even with well-defined Monitoring Plans (MP) data gaps might occur
- Operator should not have “benefits” from such occurrence
> substitute missing data with “conservative” methods

Examples



Values to close the gaps should be “conservative”.

Minimum content of Annual Reports, Annex X/MRR (Article 68(3)) (Stationary Source Installations)

MRR specifies data to be reported on data gaps and surrogate data:

(11) **Where data gaps have occurred and have been closed by surrogate data** in accordance with Article 66(1):

(a) the source stream or emission source to which each data gap applies;

(b) the reasons for each data gap;

(c) the starting and ending date and time of each data gap;

(d) the emissions calculated based on surrogate data;

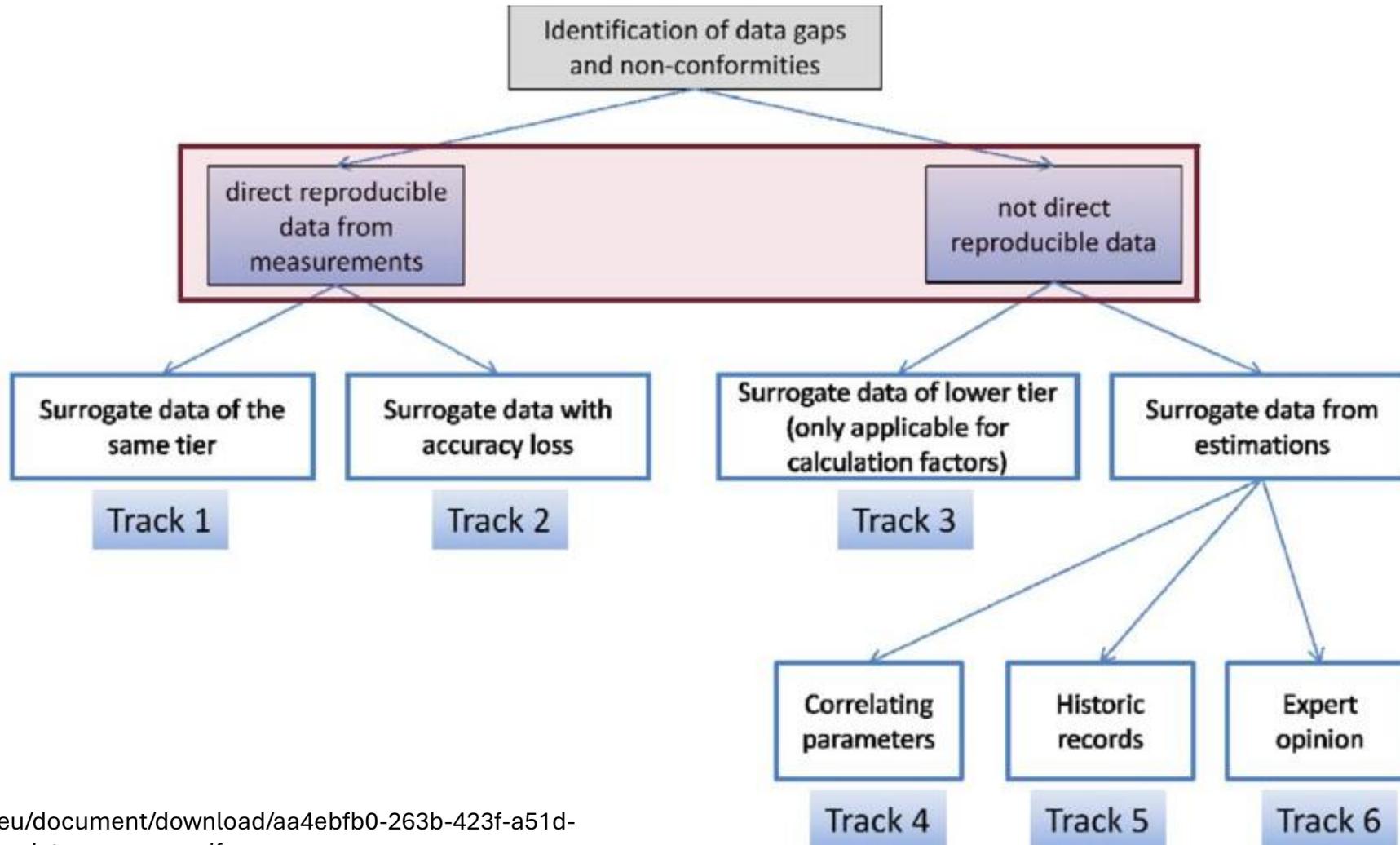
(e) **where the estimation method for surrogate data has not yet been included in the monitoring plan, a detailed description of the estimation method including evidence that the methodology used does not lead to an underestimation of emissions for the respective time period;**

Suitable approaches to generating surrogate data

- Using readings of other (redundant) instruments of same or lower “quality” (=uncertainty / tier)
- Use calculation algorithm for correcting data (e.g. for drift)
- Use a correlated parameter for calculating the required parameter
- Use historical data (average where appropriate)
- Interpolate in a trend
- Etc.

In order to make the values “conservative”, a suitable “safety margin” should be added.

Decision tree



Source:

https://climate.ec.europa.eu/document/download/aa4ebfb0-263b-423f-a51d-31f772faa6a3_en?filename=data_gaps_en.pdf



Thank you for your attention!

For further information or required clarification please contact:

Robert Gemmill: rjgemmill@hotmail.com

Machtelt Oudenes: m.oudenes@sqconsult.com

Wolfgang Eichhammer: wolfgang.eichhammer@isi.fraunhofer.de

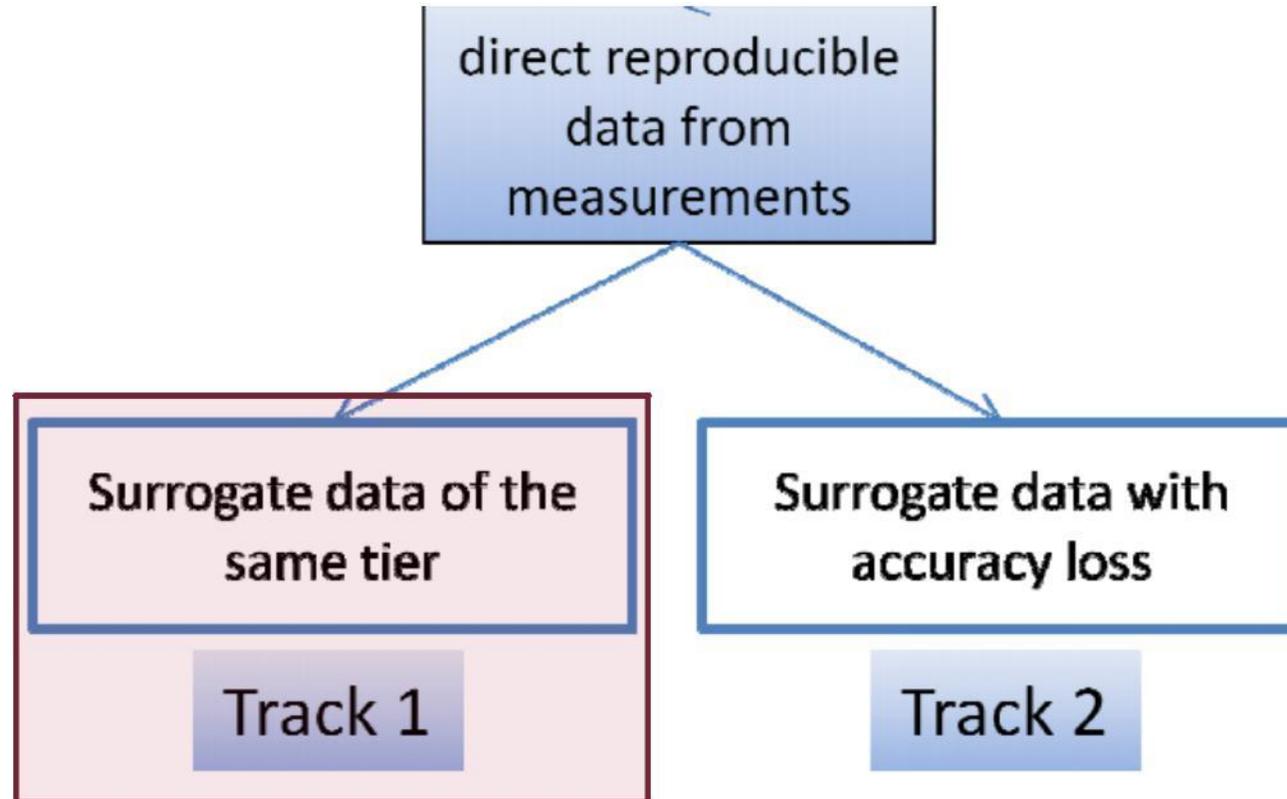
Sina Wartmann: sw@sinawartmann.org

Annex I: Details on the decision tree for data gaps

Source:

https://climate.ec.europa.eu/document/download/aa4ebfb0-263b-423f-a51d-31f772faa6a3_en?filename=data_gaps_en.pdf

Decision tree



Track 1

Reproducible data of the same quality

- Example Activity Data (AD): Redundant metering system delivering data at the same quality (e.g. subject to NLMC) is installed
- Example Calculation Factors (CF): Installation and the supplier have established procedures to Sampling and Analysis (S&A) the carbon content by accredited laboratories

$$D_r = S \text{ (no safety margin)}$$

D_r = data to be used in emissions reporting

S = surrogate data derived from a redundant system/process

Track 2

Reproducible with quality loss

- If data can be replaced by surrogate data of lower quality then a safety margin is required.

2-1: Activity data: $D_r = S + S * (U_s - U_t)$

2-2: Calculation Factor: $D_r = S + S * (U_s - U_p)$

2-3: If not quantifiable: $D_r = S + S * x\%$

D_r = data to be used in emissions reporting

S = surrogate data derived from a redundant system/process

U_s = quantified uncertainty of the secondary system including corrective measures

U_t = uncertainty of the approved tier

U_p = quantified uncertainty of the undisturbed primary system

$x\%$ = individually demonstrated safety margin

Track 2 (2)

Examples for case 2-1:

- A secondary metering system delivering data at a lower level of accuracy (e.g. not undergoing regulator calibration) is installed for the same material or fuel stream and was in operation when a data gap for the primary system was reported
- Corrective measures need to be applied retroactively as the result of a calibration has shown that the calibration function (slope or zero point) has drifted 2 %.

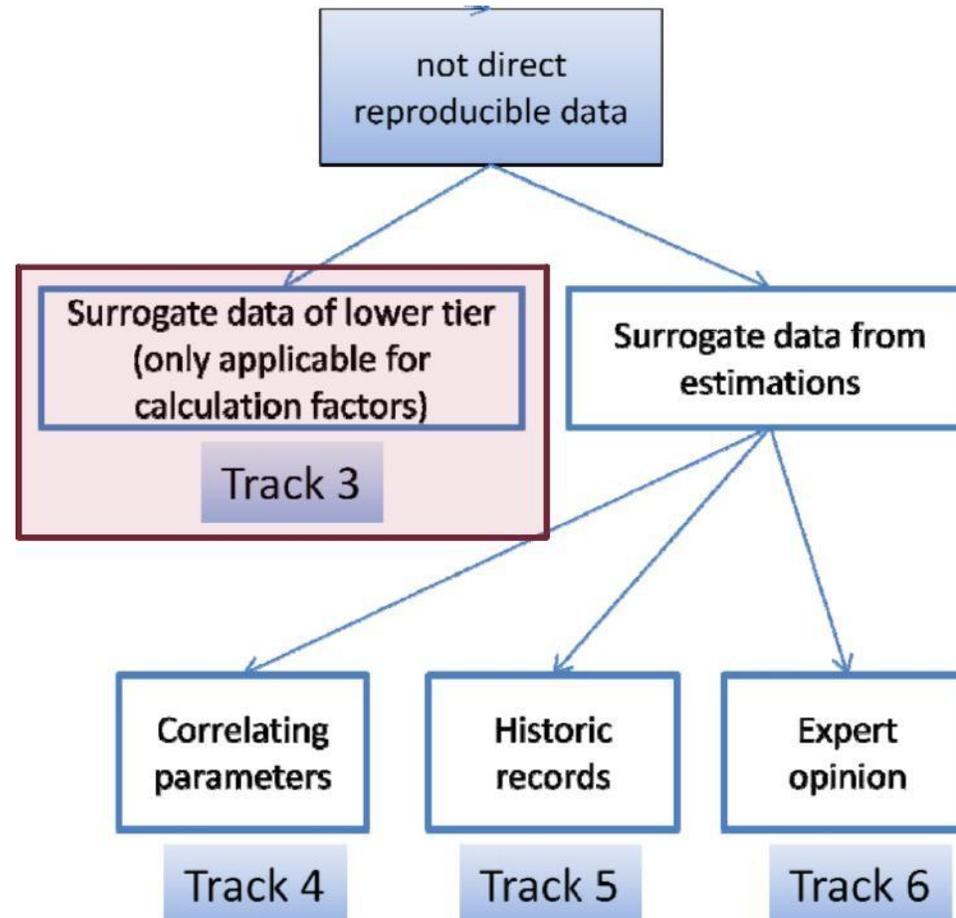
Example for case 2-2:

- Results of accredited laboratory are lost or invalid but values are available from regular control performed by own laboratory (and no retained samples available in the accredited laboratory).

Example for case 2-3:

- Results of accredited laboratory are lost or invalid and no values are available from regular control performed by own laboratory.

Decision tree (2)



Track 3 (1)

Lower tier approach (only for calculation factors!)

- Applicable when missing data has to be replaced by default values as usually requested by tier 1 or tier 2 approaches

Case 3-1a: Surrogate data given by regulation or literature $D_r = S + U_L$ or $D_r = S_U$

D_r = data to be used in emissions reporting

S = default value taken from regulation / guideline / literature

U_L = uncertainty as indicated by the same data source

S_U = default value from regulation / guideline / literature in case uncertainty already included

Track 3 (2)

Lower tier approach (only for calculation factors!)

- Applicable when missing data has to be replaced by default values as usually requested by tier 1 or tier 2 approaches

Case 3-1b: Surrogate data given by regulation or literature when missing information on uncertainty

$$D_r = S + x \% * S$$

D_r = data to be used in emissions reporting

S = default value taken from regulation / guideline / literature

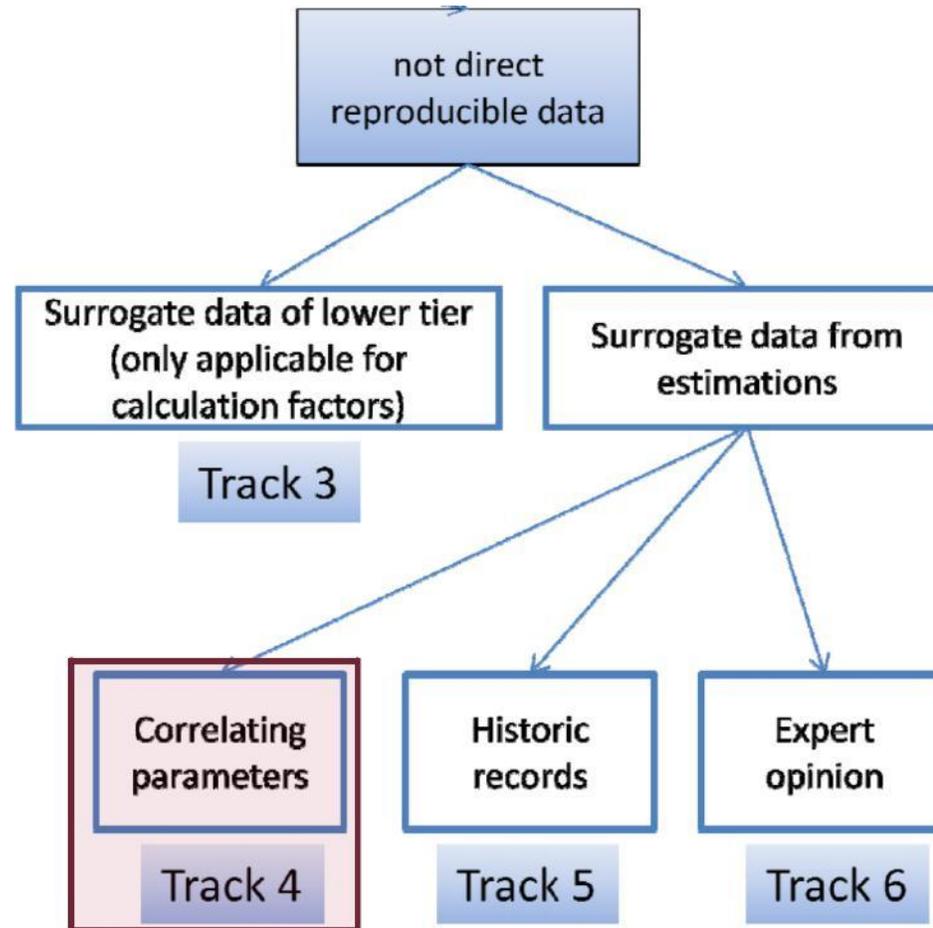
$x \%$ = individually demonstrated safety margin

Track 3 (3)

Example for 3-1a:

- A data gap caused by missing/invalid analyzes compared to the required minimum amount is filled by default values after proofing that these default values deliver a good estimate (e.g. low variability, standard commodity) e.g. from IPCC guidelines

Decision tree (2)



Track 4 (1)

Estimation: correlating parameters

- Careful assessment of the rationale behind the identified correlations required (e.g. on the statistical parameter R^2).
- Case 1: Gap between surrogate data and real data obtained by primary systems can be assessed.
- Case 2: No primary data but scientifically proven facts (e.g. heating degree days and energy consumption of a district heating plant)
- **Example Activity Data AD**: Fuel input to energy output, energy demand to air temperature, waste streams to production
- **Example Calculation Factor CF**: Heat value to net calorific value, density to heat value, density and emission factor, net calorific value to emission factor

Track 4 (2)

Estimation: correlating parameters

Case 4-1: Installation-specific surrogate data based on correlating parameters

$$D_r = S + 2 * \sigma$$

D_r = data to be used in emissions reporting

S = surrogate data delivered by correlation function

σ = standard deviation of historic simultaneous monitoring

Case 4-2: Installation-specific surrogate data based on proven correlation without records of simultaneous monitoring

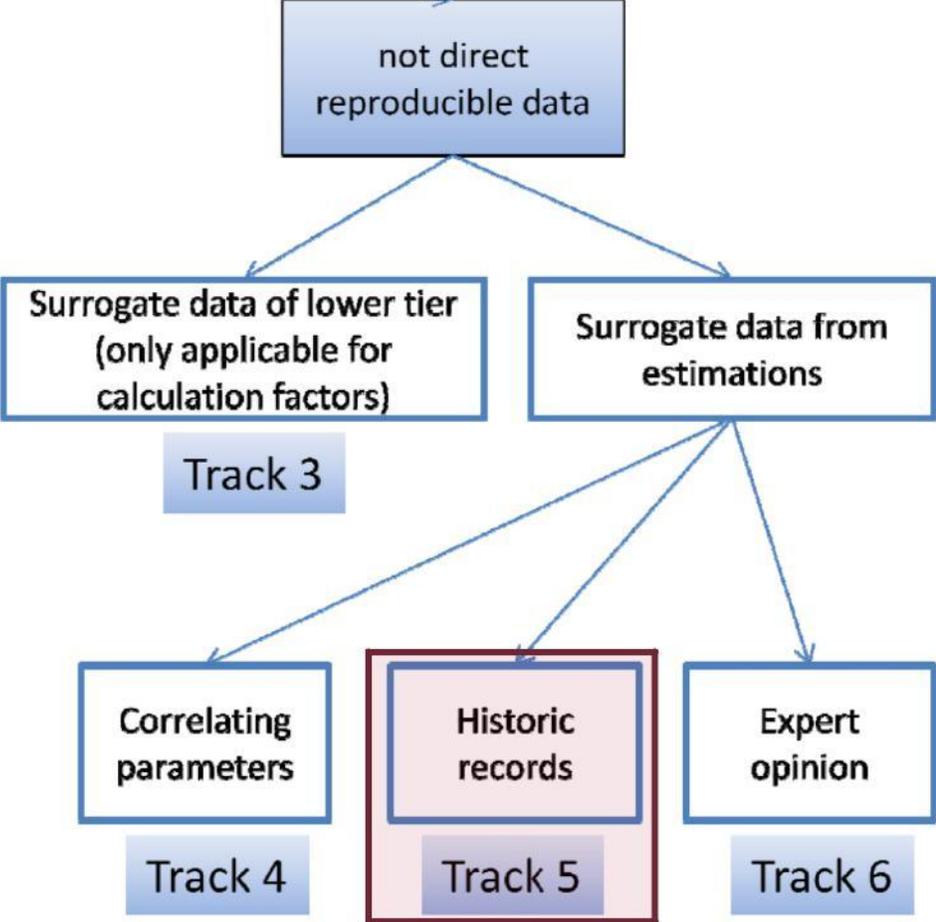
$$D_r = S + x \% * S$$

D_r = data to be used in emissions reporting

S = surrogate data derived from correlating parameter

$x \%$ = individually demonstrated safety margin

Decision tree (2)



Track 5 (1)

Estimation: Historic records

- Applied when an estimation to fill a data gap can be derived from long-term historic records (e.g. trends or seasonal behaviour) while no other information is available for a proper estimate.
- It needs to be demonstrated that conditions did not change and therefore the historic trend or behaviour delivers a reasonable estimate.
- In most cases track 3 for calculation factors is applicable, but if e.g. the amount of available historical data is not high enough or “exotic” material streams are used not enabling any comparison to other installations a standard deviation as in track 3 cannot be reasonably determined.

Track 5 (2)

Estimation: Historic records

Case 5-1a: Surrogate data derived from statistical behaviour

$$D_r = S + 2 * \sigma$$

D_r = data to be used in emissions reporting

S = surrogate data derived from statistics of historic records

σ = standard deviation of historic records

Case 5-1b: Installation-specific surrogate data based on historic records with limited data set (<20 data points) > only valid for calculation factors

$$D_r = S (\text{max})$$

D_r = data to be used in emissions reporting

S = Maximum value of historic data set

Track 5 (3)

Estimation: Historic records

Case 5-2: Surrogate data where a standard deviation cannot be reasonably determined

$$D_r = S + x\% * S$$

D_r = data to be used in emissions reporting

S = surrogate data derived from statistics of historic records

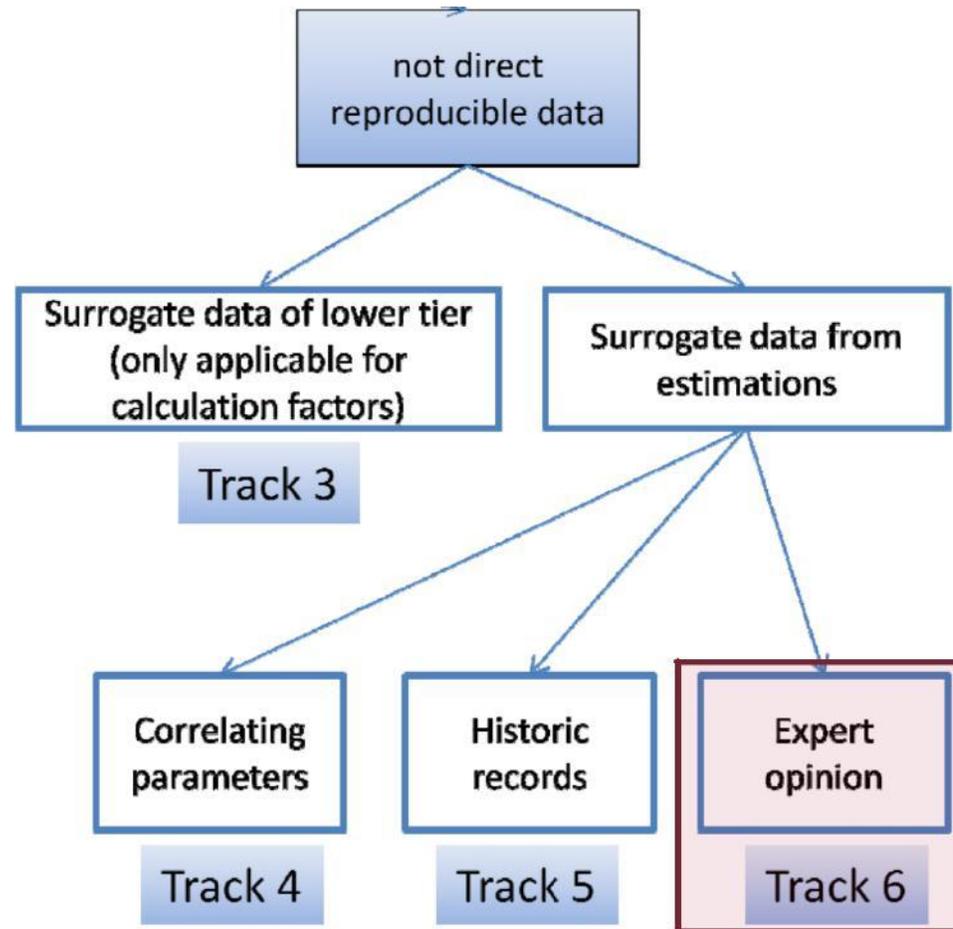
$x\%$ = individually demonstrated safety margin by the operator

Track 5 (4)

Estimation: Historic records

- **Example Activity Data AD:** Installation has regular maintenance shut-down during summer. Auxiliary installations are fuelled by a neighbouring installation (outside ETS). The consumption is always within the same range. During a reporting period it is not recognized that the meter is not functioning and no other records are available.
- **Example Calculation Factors CF:** The carbonate content of clay taken from a single mining area shows a constant increase over time following the exploitation of a geological structure. The analysing of samples failed for some months, leaving a data gap, while the continuation of the concentration curve has been observed before and after that gap.

Decision tree (2)



Track 6

Estimation: Expert opinion

- Tracks presented above are not applicable
- Estimates shall be made by engaging independent professional experts (other than the verifier)
- Expert opinion should be prepared including the aspect of conservativeness and needs to justify why no higher emissions can be expected

Example: *An operator is missing activity data from a flare source stream. The data gap took place during an unplanned shutdown where varying amounts of gas was flared. Historical values from a similar shut down could be looked at, but differences in process conditions have to be taken into account. Data from other parts of the process can be used for additional information.*

Annex II: Case studies

Source:

https://climate.ec.europa.eu/document/download/aa4ebfb0-263b-423f-a51d-31f772faa6a3_en?filename=data_gaps_en.pdf

Case 1

- *The main meter malfunctions and needs to be replaced. Until replaced, the operator replaces the data with data from a secondary meter which achieves an uncertainty of 3.7% instead of tier 4 achieved by the main meter.*
- *How should the operator close the data gap?*

Case 1: model answer

- *How should the operator close the data gap?*

$$D_r = S + S * (U_s - U_t) = S + S * (3.7\% - 1.5\%)$$

Surrogate data with
accuracy loss

Track 2

- *Alternative: If there is also metering at the supplier's site, gap might be closed based on the invoices (commercial transaction > meter under NLMC)*
 - This might achieve tier 4
 - Note that evidence (NLMC certificate, etc.) would be needed

Case 2

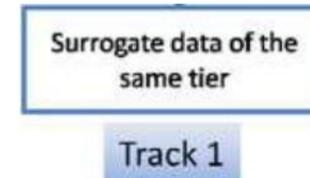
- *Operator has to apply tier 3 (sampling & analysis) for the carbon content of a material with a frequency of analysis of 4 times per year*
- *One sample sent to the accredited laboratory has shown to be contaminated making results invalid*
- *What should the operator do to close the data gap?*

Case 2: model answer

- *What should the operator do to close the data gap?*

1. *Are there retained samples?*

$$D_r = S$$



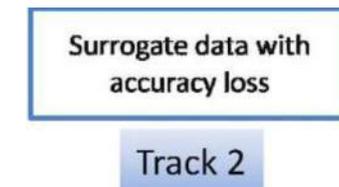
2. *Are there results from the own laboratory*

$$D_r = S + S * (U_s - U_p)$$

U_s = quantified uncertainty of the secondary system including corrective measures

U_t = uncertainty of the approved tier

U_p = quantified uncertainty of the undisturbed primary system



Case 2: model answer

- ***What should the operator do to close the data gap?***

3. Is a default value available (e.g. tier 2a)

$$D_r = S + U_L$$

S = default value taken from regulation / guideline / literature
U_L = uncertainty as indicated by the same data source

Surrogate data of lower tier
(only applicable for calculation factors)

Track 3

4. Use historic records

$$D_r = S + 2 * \sigma$$

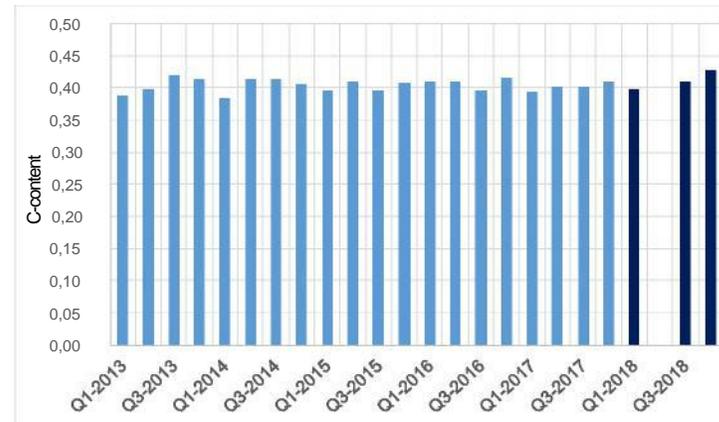
D_r = data to be used in emissions reporting

S = surrogate data derived from statistics of historic records

σ = standard deviation of historic records

Historic records

Track 5



Case 3

- *A small peak-load district heating plant is exporting hot water to the district heating network provider. It measures the natural gas consumption and the heat exported (commercial transaction).*
- *The gas meter broke down and data for two weeks in December was lost.*
- *How should the operator close the data gap?*

Case 3: model answer

- *How should the operator close the data gap?*

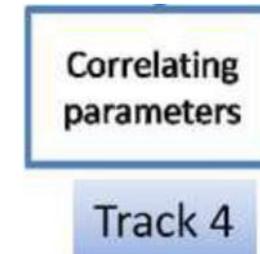
1. *Are there measurements by the gas supplier?*

$$D_r = S$$



2. *Can the correlation "heat-fuel input" be used?*

$$D_r = S + 2 * \sigma$$



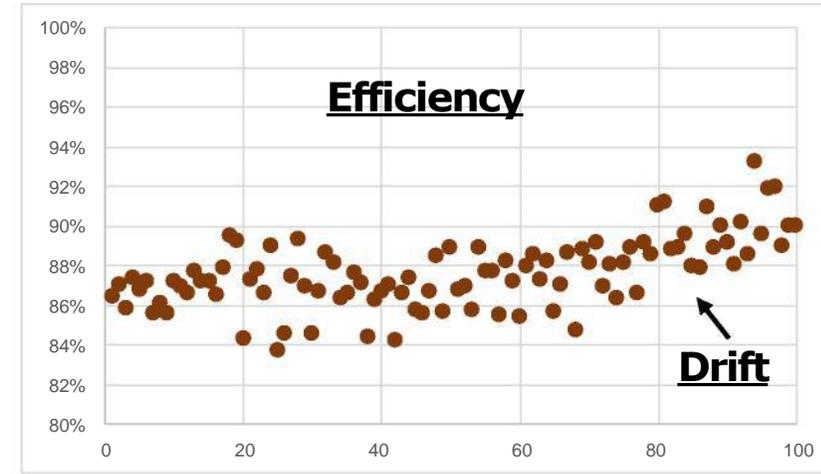
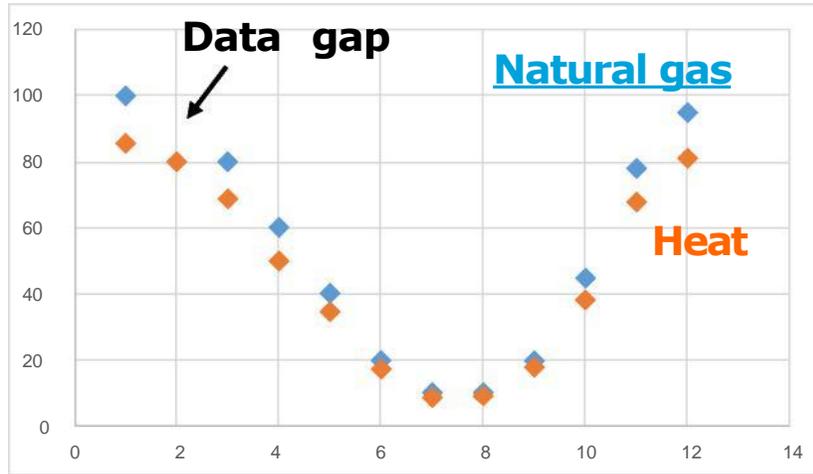
D_r = data to be used in emissions reporting

S = surrogate data delivered by correlation function

σ = standard deviation of historic simultaneous monitoring



Case 3: model answer



$$D_{r,NG} = \frac{S_{heat}}{S_{eff} - d - 2 * \sigma_{eff}}$$

$D_{r,NG}$ = data to be used in emissions reporting (natural gas)

S_{heat} = surrogate data for heat (during the data gap)

S_{eff} = surrogate data for efficiency (other than during data gap)

d = the drift of the efficiency

σ_{eff} = the standard deviation of historic efficiency