

## Data quality

*Measuring, extrapolating, estimating or using standard values*

# Colophon

## ***Guideline 21 - Data quality***

*Measuring, extrapolating, estimating or using standard values*

## ***Carbon Footprint in logistics***

January 2021

© Connekt

## **Connekt/Topsector Logistiek**

Ezelsveldlaan 59

2611 RV Delft

+31 15 251 65 65

[info@connekt.nl](mailto:info@connekt.nl)

[www.connekt.nl](http://www.connekt.nl)

# Data quality

## Measuring, extrapolating, estimating or using standard values

This guideline deals with one of the most important questions arising from practice: what do I do if I have no good data available for a part of my chain? This may be because one of the service providers is unwilling or unable to supply data, because part of the transport takes place on another continent or because part of it is outsourced further, for example. In such a case you have a mix of good, detailed data, estimated data and no data at all.

A seemingly simple but very powerful solution has been developed to address this problem. This is known as 'data quality'.

To put it simply, this involves attaching a data-quality class to each input (data) item. An example of a class might be: 'default', i.e. a standard figure from a website or report, or 'Gold+', i.e. detailed data from an on-board computer on consumption and cargo per stop.

Attaching this kind of extra information to a figure is referred to as adding 'metadata'.

When allocation calculations are performed these metadata on input quality are taken into account and incorporated into the calculations. This means the calculated figure also has metadata. These metadata contain the mix of the input metadata, allowing you to see what the quality of the source is.

A simple example:

- input data 1: metadata default
- input data 2: metadata Gold+
- calculation: 25% data 2 and 75% data 1
- output: metadata (25% Gold+, 75% default)

When the results are analyzed you can therefore see immediately to what extent you can rely on the outcome. More importantly, you can see where best to focus your attention to obtain better information. After all, if an important component of the result has a default value as its input, it is worth making an effort to obtain a better figure.

The great thing about this approach is that it is always possible to start allocating CO<sub>2e</sub> even if little information is known to begin with. You can then make improvements step by step. This process is developed and explained in more detail below.

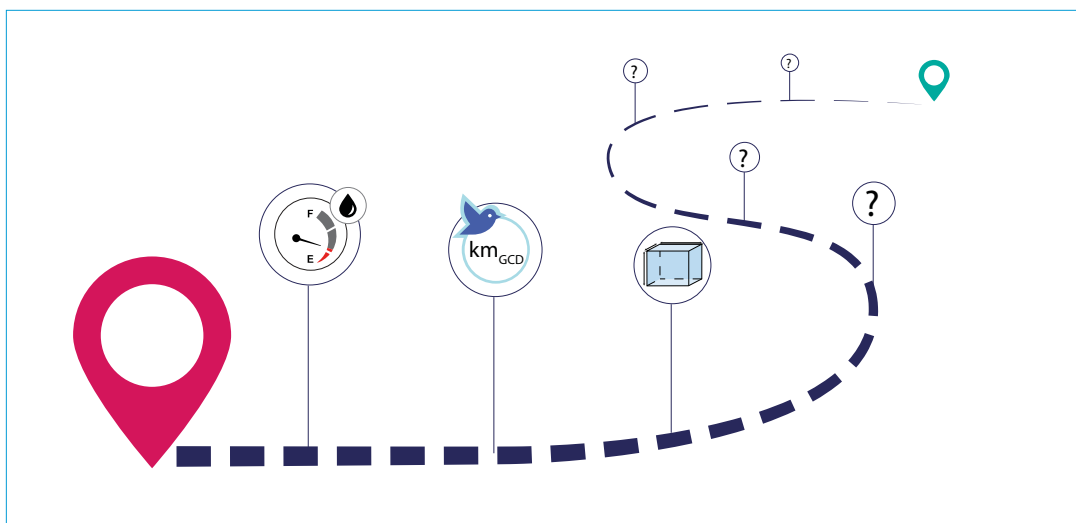
## Detailed measurements, estimated data or standard values

Guideline 1. 'Allocation' explains the principle of the allocation process: to put it simply, all the emissions from an aircraft, vessel or vehicle are allocated to the cargo transported. The CO<sub>2e</sub> emissions are derived from the amount of fuel consumed, while the consignment notes contain all the data required about the cargo.

Some road hauliers have an FMS and/or TMS that allows trucks with on-board computers to precisely measure consumption per stop. For each stop it is known what cargo is unloaded or loaded, how heavy or large the cargo is, who the customer is, etc. These systems generate a wealth of (digital) data, allowing allocation to be performed with great precision.

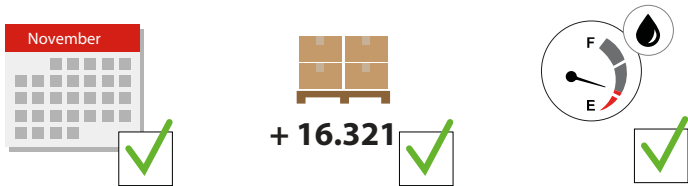
Such a wealth of data is not available everywhere at all times. Some countries are still a long way behind in terms of digitization, which may be of relevance, for example, to international logistics chains with legs located in other parts of the world. In the field of logistics many activities are outsourced (sometimes several times) to smaller parties who have less sophisticated systems or whose processes are barely digitized. Sometimes data are only available for an entire fleet or a longer period of time. Not all subcontractors are able or willing to perform this allocation for their customers.

In such cases is it possible to use estimated data or even standard values? How does it work if some of the data in a chain have been measured extremely precisely, while other data have been estimated?



And how can you then compare two figures for two different chains?

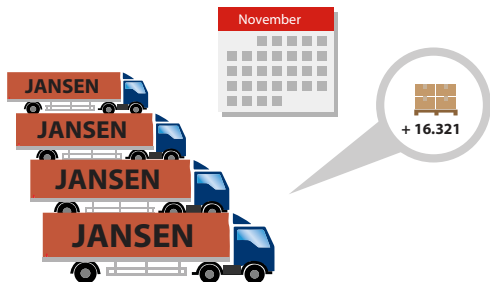
Measuring, extrapolating, estimating and using standard values: cargo data



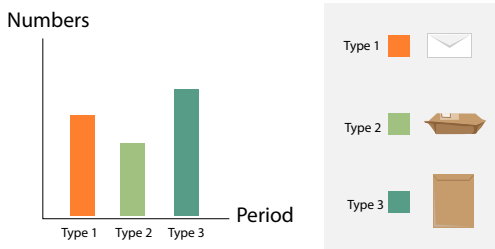
*In most cases there is good knowledge of the basic cargo data: quantity of cargo, origin and destination, mode, customer, etc. Generally speaking, the amount of fuel consumed in each period is also recorded.*



*Some transporters know everything in detail: precise data about the cargo, linked to the amount of fuel consumed per license plate, per stop or trip.*



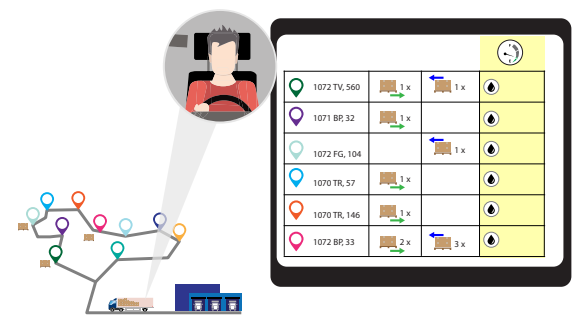
*In the case of transport that is regularly repeated, sometimes all that is known is how many times the same quantity of cargo has been transported between locations. Sometimes only the quantity (tons, number, etc.) is known per week, month, quarter or year.*



*In other cases only the number of units may be known, e.g. in the case of post. The transporter only knows how many items of post are inside the bags.*

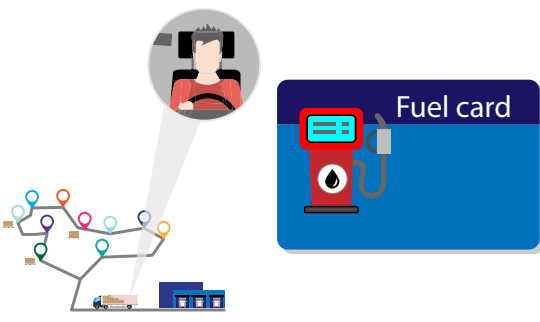
In the case of fuel there is an even greater variety of data. Guideline 4. 'Fuel' explains in more detail how you can deal with fuel data. The main points are repeated below.

Measuring

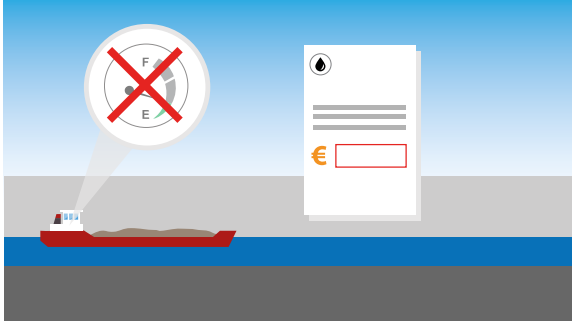


Location	Stop	Direction	Fuel
1072 TV, 560	1 x	1 x	1 x
1071 BR, 32	1 x		1 x
1072 FG, 104		1 x	1 x
1070 TR, 57	1 x		
1070 TR, 146	1 x		1 x
1072 BR, 33	2 x	3 x	3 x

Modern on-board computers in trucks indicate how much fuel has been consumed, if necessary per trip or stop.

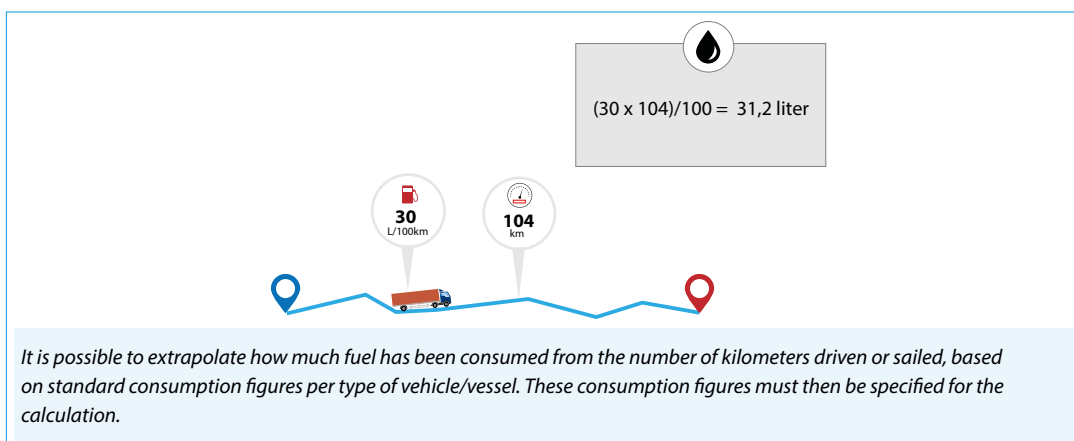


However, the totals per fuel card also represent a measured amount, over a certain period of time.

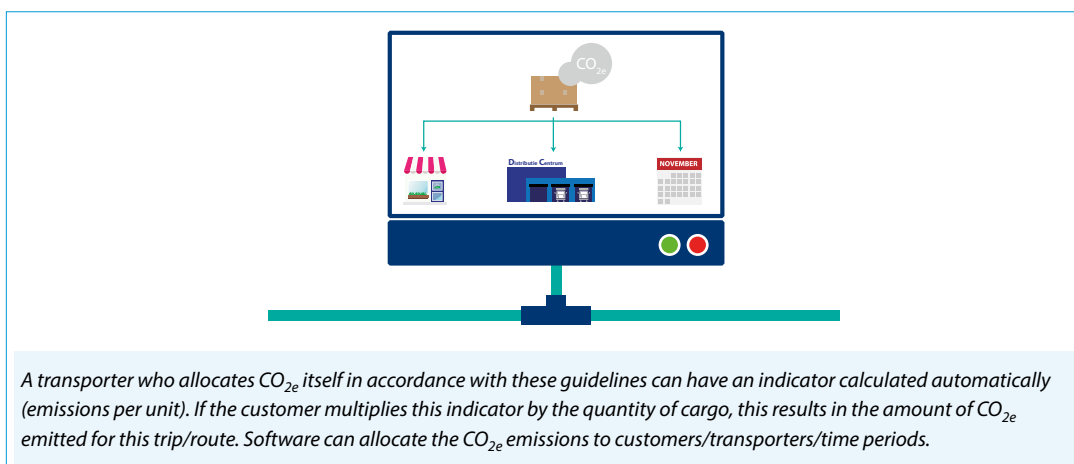
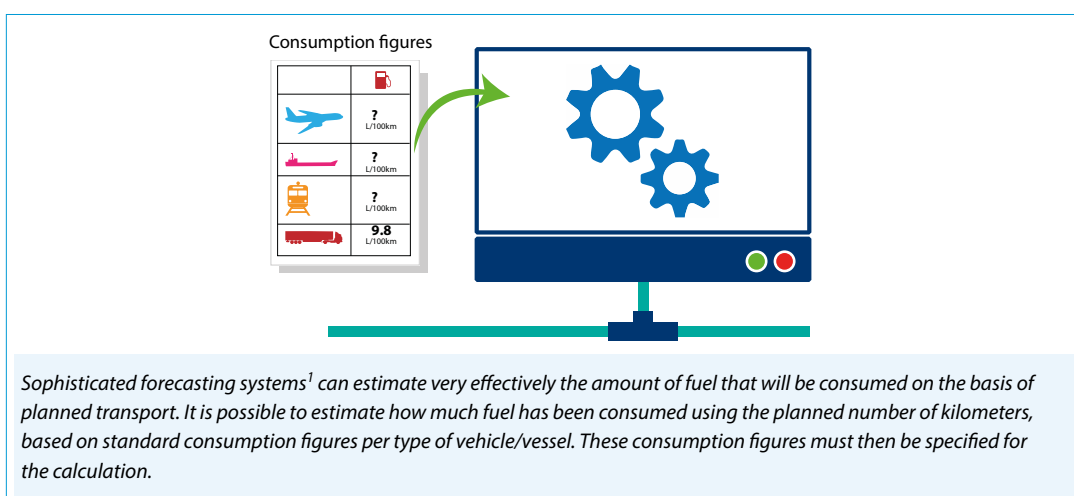


Fuel purchase invoices are also a source of data for measuring the amount of fuel consumed. This applies in the case of inland vessels with no fuel gauge, for example. Fuel data based on invoices or annual statements is still valuable (accurate) information, although aggregating it can be a rather laborious task.

## Extrapolating

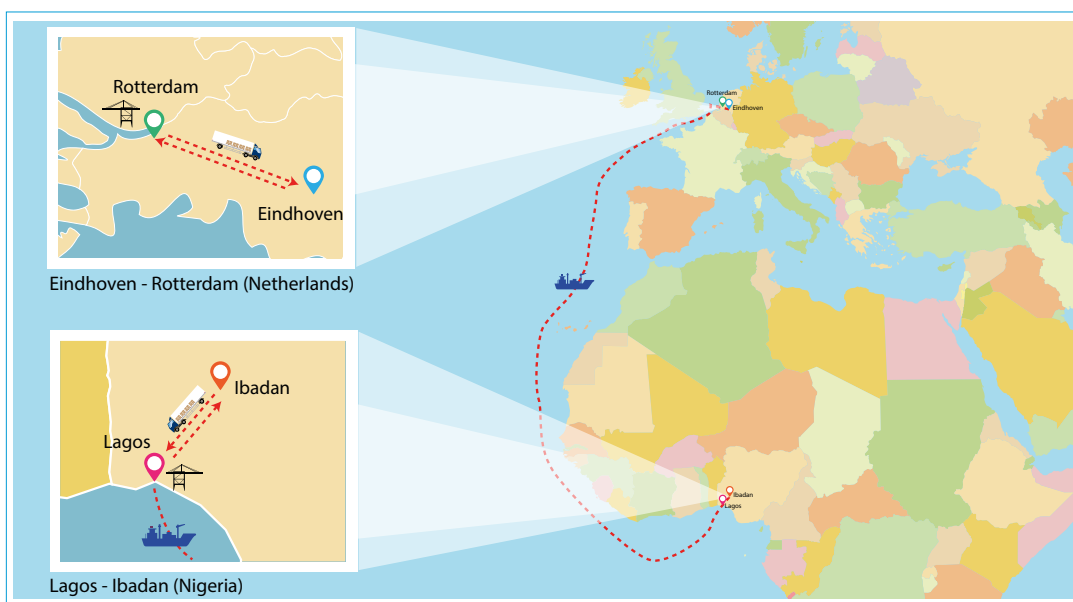


## Estimating



1 Such as EcotransIT

## Standard values



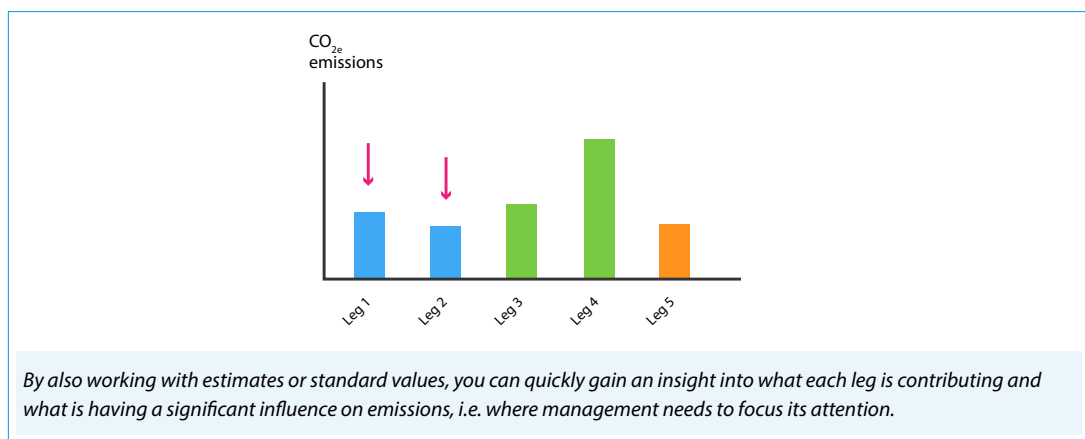
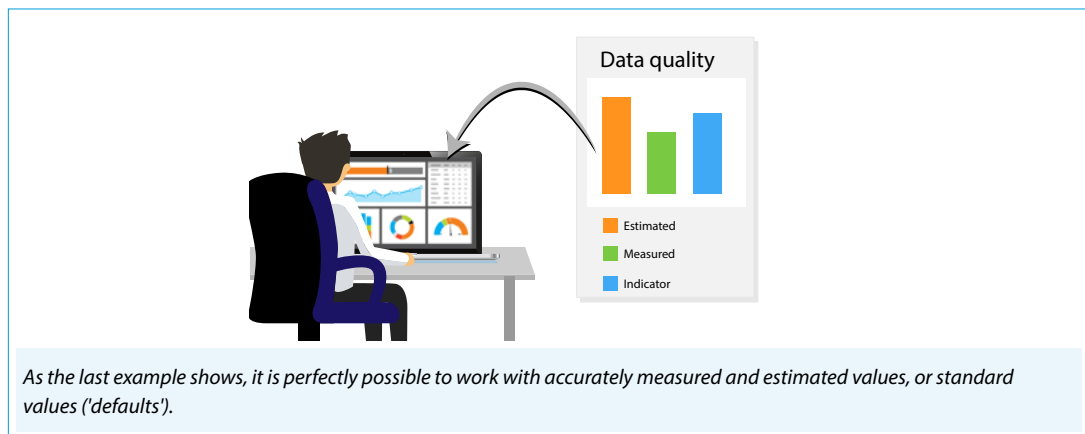
Route	km <sub>gcd</sub>	Consumption or emissions	Number	Kilometers	Liters of diesel	per 40 ft container	Total
Eindhoven - Port of Rotterdam	112	90 Measured fuel consumption per container incl. empty kilometers	40 trips		3,600	290.7 kg CO <sub>2e</sub> per container	11,628 kg CO <sub>2e</sub>
Port of Rotterdam transshipment		8 Measured emissions per transshipment 40 ft container	40 x transshipment			8 kg CO <sub>2e</sub> per container	320 kg CO <sub>2e</sub>
Port of Rotterdam - Port of Lagos	5,065	77 Standard value for container shipping* CO <sub>2e</sub> /TEU <sub>km</sub> (sailed)	40 x 2 TEU	7,708 estimated km sailed		1,187 kg CO <sub>2e</sub> / container	47,481 kg CO <sub>2e</sub>
Lagos transshipment		12 Standard value for emissions per transshipment 40 ft container	40 x transshipment			12 kg CO <sub>2e</sub> per container	480 kg CO <sub>2e</sub>
Lagos - Ibadan	108	42 Standard value for truck (liters/100 km)	40 trips	260 estimated km driven	4,368	353 kg CO <sub>2e</sub> per container	14,109 kg CO <sub>2e</sub>
<b>Cargo</b> 40 ft container 25 tons per container 40 containers 1,000 tons total							<b>1,850 kg CO<sub>2e</sub> per container</b> <b>74,018 total kg CO<sub>2e</sub> of which 62,070 on basis of indicators</b>
For the emission factor for diesel the figure indicated at <a href="http://www.co2emissionfactoren.nl">www.co2emissionfactoren.nl</a> is used							<b>74 kg CO<sub>2e</sub> per ton</b> of which 62 kg CO <sub>2e</sub> on basis of indicators

\* The standard value for container shipping to Lagos is trade-lane-dependent. There is a considerable imbalance in the shipping to Lagos, which is why this figure is almost twice as high as the standard value for Rotterdam-Shanghai, for example: this is 47 grams of CO<sub>2e</sub>/TEU<sub>km</sub>.

In practice, especially in long international supply chains, all variants may be used at the same time. The basis applied for each fuel or CO<sub>2e</sub> figure needs to be specified.

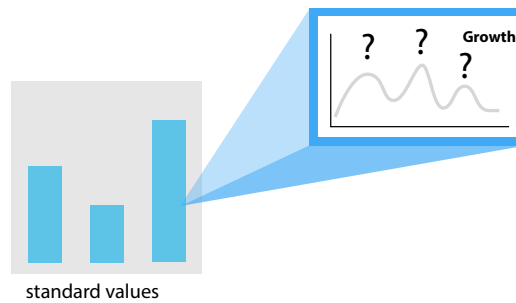


## Data quality



It is therefore important to understand that different kinds of figures have different significance.

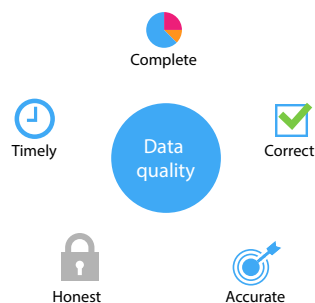




*Standard values do not provide any insight into the actual situation and do not reveal whether a company is improving over time.*



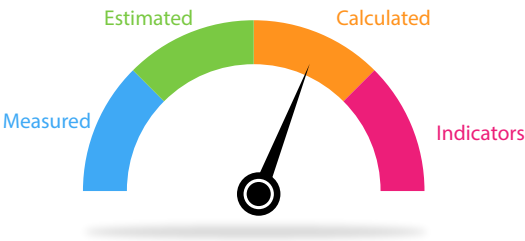
*The more measurements are performed, and the greater their accuracy, the greater the value of the result in terms of operational control.*



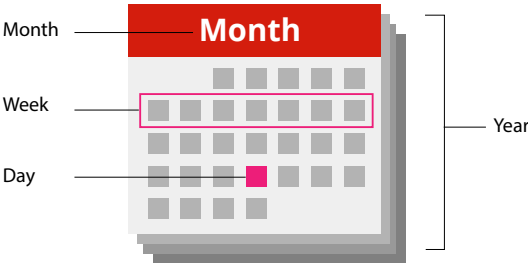
*But how can you tell from a figure whether it has been measured accurately or extrapolated from a standard value? The solution is to assign a data quality to basic data and incorporate this characteristic of the data into the calculation.*

Quality: time period, measuring or estimating, individual or fleet


A number of factors determine quality for both cargo and fuel.



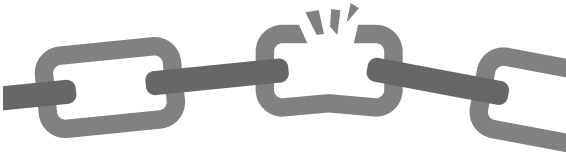
Are the data measured, estimated/extrapolated or based on standard values?



How detailed are the measurements in terms of time? Day, week, month, year?



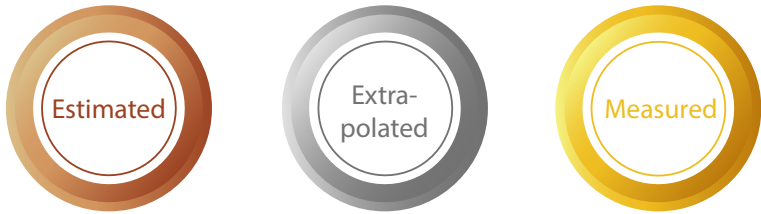
Is each aircraft, vessel or vehicle measured separately or is everything lumped together for a fleet?



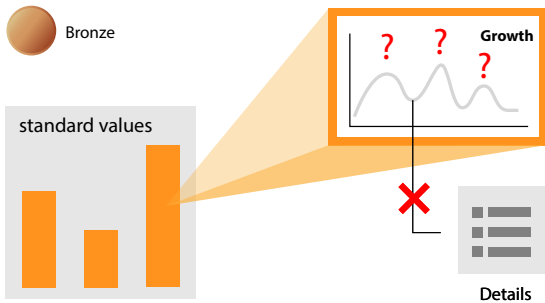
The least accurate factor determines the best that can be achieved: if a lot of details are known about the cargo (precise timings of each order), but the amount of fuel is only known for a fleet of trucks per year, the allocation relates to a year.

### Gold, Silver, Bronze

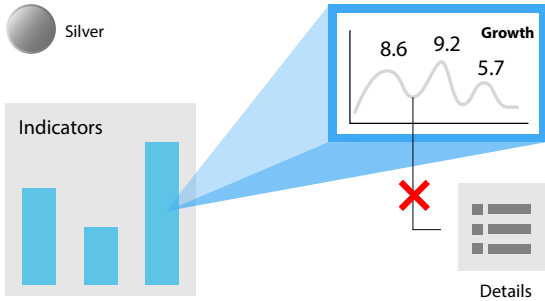
A useful way to work with data quality is on the basis of 3 quality levels.



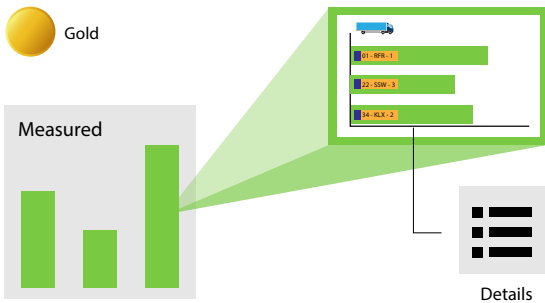
**Bronze:** estimate or standard value.  
**Silver:** year or month, measured (aggregated per month) or extrapolated.  
**Gold:** measured in detail per vehicle or vessel.  
The extra category **Gold+** has also been created for highly detailed measured data, such as daily data generated by FMS/TMS.



This breakdown highlights the key differences: Bronze: if a company repeatedly analyzes itself in the same way, the (relative) development over time can be tracked.




Measured values, even if they relate to a number of vehicles in combination, offer a much greater insight. Monthly data allow seasonal influences to be seen.




Gold: measurement is the key to knowledge; the details become visible. If measurement is performed per vehicle or vessel, the real details emerge.

Data quality for fuel (or energy)


The following breakdowns and notations are used for fuel.




**Bronze**



**Silver**



**Gold**



**Gold +**

Indicators  
year / month

Measured  
per period **or** per  
license plate/location

Measured  
per period **and** per  
license plate/location

Measured  
per trip

Data quality for fuel or energy			
Period <sup>A</sup>	Fuel / energy total	Fuel / energy per license plate	Fuel / energy per location
Year	B <sub>y</sub>	B <sub>y</sub>	B <sub>y</sub>
Month	B <sub>m</sub>	B <sub>m</sub>	N/a
Jear	S <sub>y</sub>	S <sub>y</sub>	S <sub>y</sub>
Month	S <sub>m</sub>	G <sub>m</sub>	G <sub>m</sub>
Week	N/a	G <sub>w</sub>	N/a
Trip	N/a	G <sub>t</sub>	N/a

**B - Bronze:**  
Estimates based on default values and indicators.  
The fineness of the period is indicated as a subscript next to the B (year, month).

**S - Silver:**  
Average measured values per period (year, month), such as total fuel in period/number of kilometers traveled in period.  
The fineness of the period is indicated as a subscript next to the S (year, month).

**G - Gold:**  
Measured values per aircraft, vessel or vehicle or location per period (month, week trip).  
The fineness of the period is indicated as a subscript next to the G (month, week, trip).

A location is a transshipment or storage point where energy is consumed.

Data quality for cargo

If the cargo documents are available in detail, this automatically results in the quality  $G_t$ . Whether this information can be put to use in the allocation depends on the data quality for the fuel.

Data quality for cargo



Period <sup>A</sup>	Aggregated over cargo documents	CMR or cargo document
Year	$B_y$	N/a
Month	$B_m$	N/a
Year	$S_y$	N/a
Month	$S_m$	N/a
Week	$S_w$	N/a
Trip	N/a	$G_t$

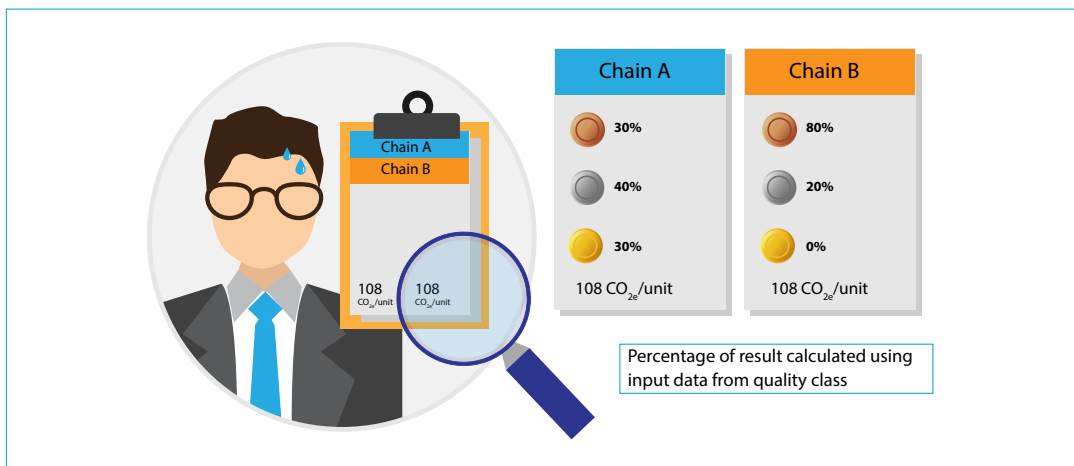
**B - Bronze:**  
Estimates of transport volume per period.  
The fineness of the period is indicated as a subscript next to the B (year, month).

**S- Silver:**  
Measured values for transport volume per year ( $S_y$ ).  
Measured values for transport volume per month ( $S_m$ ).  
Measured values for transport volume per week ( $S_w$ ).  
The fineness of the period is indicated as a subscript next to the S (year, month, week).

**G - Gold:**  
Measured values for transport volume per trip ( $G_t$ ).  
The fineness of the period is indicated as a subscript next to the G (trip).

## Auditors

Following the allocation calculation, the data quality of the allocation is indicated per (smallest) element of the allocation. By incorporating this allocation quality into the data, you ensure that the insight obtained retains its significance during analyses.



For example, it remains possible to see that an emissions figure for a chain is made up of:

- 30% of Gold quality data,
- 40% of Silver quality data,
- 30% of Bronze quality data.

If another chain has a similar emissions figure, but the data quality is only:

- 20% of Silver quality data and
- 80% of Bronze,

the comparison must be viewed in an entirely different light.

In the latter case 80% of the 108 kilograms (86.4 kg) has been calculated on the basis of standard values in the underlying data.

With chain A much more has been measured or extrapolated, which means the figure is more reliable. For data controllers, such as auditors, this information is extremely important: to what extent can the figure be relied on?

# Carbon Footprint guidelines

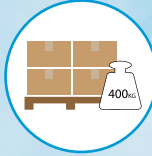
0. Measuring, calculating, allocating and reducing



1. Allocating



2. Cargo



3. Origin and destination



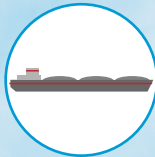
4. Fuel



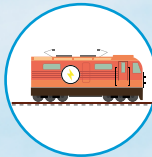
5. Inland shipping - containers



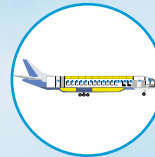
6. Inland shipping - bulk



7. Freight transport by rail



8. Air freight



9. Maritime and short sea shipping



10. Transshipment



11. Storage



12. Parcel transport and post



13. General road transport



14. Perishable and temperature controlled



15. Outsourced transport



16. Repositioning and empty kilometers



17. (Inter)national supply chains



18. Benchmarking



19. Intermediaries and platforms



20. Auditors and accountants



21. Data quality



22. The relationship between social goals and corporate goals

