Task Force 3- Guidance for Materiality and Sampling

22 may 2017

1.1 Introduction

This guidance document is for materiality and sampling of data that is reported in accordance with the scope of the regulation and is related to Article 15 and Article 12 respectively in the delegated regulation 2016/2072.

1.2 Materiality level

The materiality level is 5% of the respective total reported for each item in the reporting period:

- Fuel consumption
- CO_2 emissions
- Cargo carried
- Distance travelled
- Transport work
- Time spent at sea

Assessing the materiality of misstatements has a quantitative and qualitative aspect;

With respect to the quantitative aspect the verifier aggregates misstatements in the reported data and compares the individual and aggregated misstatements to the total declared value in the company's report. The difference between what the verifier considers the correct total value and what is declared by the company in the annual emissions report is compared to the materiality level. If the materiality level is exceeded the impact on the reported data is material. It should be noted that misstatements can individually be minor misstatements but could exceed the materiality level once they are aggregated.

The quantitative aspect and thus the materiality level alone is not the only factor when assessing whether or not a misstatement has material effect. The qualitative aspect should be considered as well. This will depend on the size and nature of the misstatements as well as on their circumstances of occurrence. See article 17.5 of delegated regulation 2016/2072.

Examples of material misstatements

1) During the verification process of ship X of company Y, it is detected that the company Y made a typo in the application of the emission factor for fuel. The ship has consumed only heavy fuel oil for which the emission factor is 3.114 t CO_2 / tonne fuel. In calculating the emissions from this consumption, the company accidently used a factor of 2.114 t CO_2 / tonne fuel. Someone made a typing error in a table of emission factors in the system that is used to calculate the CO_2 emissions. The ship consumed 1,000 tonnes heavy fuel oil. The consequence is that the reported emissions are 2,114 tonnes instead of 3,114 tonnes. This means misstatement of 1,000 tonne CO_2. The impact of the misstatement is 1,000 / 3,114 x 100% = 32%. This means that if the company does not correct the error, the total emissions in the emissions report are 32% too low. This is considered a material misstatement, because the impact on the total reported number is above 5%. This means that the company must correct the misstatement for the verifier to provide a verification report with a positive outcome for ship X.

2) A shipping company implemented a new IT system to collect and manage the data. Through sampling the verifier noted that information on fuel consumption, distance and cargo for about 1 month of the year contains errors. If this is detected based on 1 sample voyage for the related month, the impact is larger than just the one voyage. While the error in data for the voyage may not be material, the aggregated omission will be material. For example, the ship has done X voyages in the reporting period and the verifier has sampled 20% voyages. If 2 errors have been found in the population of 20% voyages, then the size of the error = 10 errors in the reporting period. Thus, verifiers will evaluate the impact on the aggregated information of a single issue noted during the verification, and will take into consideration the likelihood of errors in the data they have not sampled.

3) During the verification process of ship X of company Y, it is detected that the company Y has not included all the bills of lading for the relevant voyages of ship X when reporting the total cargo carried. Someone adding up the total cargo carried accidently omitted some data. This would be considered a material misstatement if the impact of the missing data on the total reported number is above 5%. This means that the company must correct the misstatement for the verifier to provide a verification report with a positive outcome for ship X.

1.3 Sampling

Sampling is the application of a procedure where less than 100% of a population is checked and verified compared to all data and/or control activities/procedures that is subject to verification. This is important, because the verifier must be sufficiently confident that the results are representative enabling it to draw conclusions about the entire population from a sample. For the application of the EU MRV regulation this applies to the sampling of data from a ship and not the sampling of a ship within a fleet.

Sampling is one of the verification activities that is impacted by this risk assessment. Depending on the verifier's analysis of the level of inherent and control risks, the verifier determines whether sampling is justified, which samples it needs to take, what the sampling size and selection approach should be and which types of tests or other checks it should undertake on each sample.

1.3.1 Type of risks

Risks are classified as below:

Inherent risk: as defined in Article 2 (5).of COMMISSION DELEGATED REGULATION (EU) 2016/2072

Control risk: as defined in Article 2 (6) of COMMISSION DELEGATED REGULATION (EU) 2016/2072
Detection risk: as defined in Article 2 (7) of COMMISSION DELEGATED REGULATION (EU) 2016/2072.

Examples of the types of risk that determine the sampling are shown in the table below:

Factor	Explanation
Inherent risk	**Inherent risks are risks linked to the data flow activities** themselves assuming that there are no related control activities to mitigate these risks, and without considering the company's control environment. The risks are thus purely related to the size and characteristics of the company's data flows. The purpose of the company's control system is to mitigate its inherent risks. Examples of potential sources of inherent risk: • complexity and number of emissions sources and fuels used • significant manual transfers and input of data concerning fuel consumption etc. • complex data management systems for collecting data and quantifying emissions (e.g. multiple spread sheets related/ linked to each other) or changes in data management • Inconsistent or complex monitoring methodologies and reporting policies for example incorrect use of emissions factors, incorrect identification of voyages in scope, incorrect use of selected fuel consumption monitoring method. • Unit conversions when consolidating information for example volume to mass • Measuring equipment failure • IT system failure

Control risk	**Control risks are to a large extent determined by the Company's control environment**, i.e. the way and the stringency with which inherent risks are addressed and mitigated within the ship operations. Examples of potential sources of control risks: • automated controls in the IT system that are missing or not functioning properly • internal audits that have not been correctly performed • No monitoring / maintenance of measuring equipment as required by maker. • there is no separation of data input from data checking (i.e. the checking is done by one person which means there is no proper segregation of duties) • internal data reviews and the checking of the manual transfers of data that are not carried out, or not carried out to the rigour required in view of the inherent risk level • the person responsible for the control activities is not or not sufficiently knowledgeable regarding the task concerned
Relevant control activities	**When both the inherent risks and control risks are high,** the verifier should apply more detailed and robust verification activities and has to select a larger sample to lower the detection risk. Determination of the sample size for testing the control system as presented in the assessed monitoring plan depends on the frequency of the internal control tests and the control activities, and the number of items that need to be controlled. **The frequency of the control activity** means how many times a control activity is being carried out by the company, for example how frequent is the data cross checked or how frequent the monitoring plan is checked for relevance. **The number of items** refers to the number of data points and data flows that are being controlled by the control activities, e.g. how many measurement points are being used, how many documents there are in the documentation management system etc.
Detection Risk	The detection risk will depend upon the procedures, processes and systems used by the verification body to test the data and the subsequent risk not to detect a material misstatement.

Verification opinion with reasonable assurance	**Where the verifier identifies a misstatement or non-conformity** during sampling, it should request the company to explain the root cause(s) of that misstatement or non-conformity. Based on the outcome of that assessment the verifier should determine whether additional verification activities are needed, and whether the sampling size needs to be increased (usually the case).

Factors that impact the sample size are shown in the table below:

Factor	**Explanation**
Inherent risks and control risks	If major weaknesses are identified during the testing of control activities, the verifier will conclude that the confidence obtained from that control activity is low and therefore that the risk of material misstatement is high. In that case the verifier will aim for a larger test sample to give it the necessary confidence that all possible misstatements will be detected. If no major weaknesses are found in the testing of the control activities, the confidence obtained from applying tests on the system and the control activities will be high meaning that the verifier is confident that it may trust the system and therefore aim for a smaller test sample. In both cases the verifier's professional judgement is applied to the percentage of the population that is sampled to give it the necessary confidence that all possible misstatements will be detected.
The results of analytical procedures	Fluctuations and trends in data, deviations from previous years, data gaps, outliers, as well as unexpected data without explanation from the company will require special attention and affect the number of data points to be sampled.
The requirement to deliver a verification opinion with reasonable assurance	The sampling and the sampling results need to enable the verifier to provide an opinion with reasonable assurance suggesting a higher rather than lower percentage of the population being included in the sample.

1.3.2 Types of Sampling

The verifier has the option to choose between statistical and non-statistical sampling using its professional judgment. Professional judgment will also be used in the planning, performing, and evaluating of sampling, and the sample evidence obtained in relation to other verification evidence.

This choice between the statistical and non-statistical is often based on several considerations, such as the number of emission source streams and data points per emission source stream, the variation between those data points, and the degree the sample allows a conclusion over the entire population of data or control activities. The verifier uses its professional judgment to assess

factors such as the characteristics of the data, the control activities or the procedures for control activities, and the risks in relation to these characteristics to determine the appropriate sample size.

Sampling risk is the risk that the verifier's conclusion based on a sample may be different from the conclusion if the entire population were subjected to the same verification procedure.

1.3.2.1.Non-statistical Sampling

Any sampling procedure that does not permit the numerical measurement of the sampling risk is a non-statistical sampling procedure, even if the verifier rigorously selects a random sample, instead judgment is used to select the sample items.

For most verifications, the non-statistical approach will be appropriate, since for system audits, addressing questions such as "are the proper control activities installed, implemented and maintained", are important and highly relevant. This also applies to the verifier's analysis of the nature and cause of errors as well as its conclusion on the mere absence or presence of errors. The verifier can in this case choose a fixed sample size of items to be tested for each key control activity if the size of the sample is increased if errors are identified. Nonetheless, professional judgment remains critical in determining the relevant factors to consider. However, if a non-statistical approach is being used, the results of the sampling do not allow extrapolation to the entire population.

What impacts the verifier's detection risk and therefore the sample size?

Verification Risk (VR) = Inherent risk (IR) x Control risk (CR) x Detection Risk (DR)

The combined inherent and control risk can be determined, an example of a 3-tier risk approach is shown below, however this could be greater. A verifier could decide to use a risk approach using 5 levels: limited, low, medium, high, extreme or a quantitative methodology (risk quoted from 0 to 10 for instance). It is up to the verifier to decide.

		Control risk		
		Low	**Medium**	**High**
Inherent risk	**Low**	Low	Medium	Medium
	Medium	Medium	Medium	High
	High	Medium	High	High

1.3.2.2 Statistical Sampling

With statistical sampling, sample items are selected in a way that each sampling unit has a known probability of being selected. The verifier will use probability sampling and selection methods, i.e. random, systematic or stratified sampling, to select the items to be reviewed during verification. Probability sampling provides an objective method of determining the sample size and selecting the items to be examined. A number of sampling techniques come into perspective that assists the verifier in its conclusion on the number of misstatements in the sample and the misstatements in the entire population of data.

1.3.2.3 Sample selection

Apart from the distinction between statistical and non-statistical sampling, the verifier will also choose between the following sampling approaches:

1. ***Random selection*** - Applied through random number generators, for example, random number table
2. ***Systematic selection*** - The number of sampling units in the population is divided by the sample size to give a sampling interval, for example 50, and having determined a starting point within the first 50, each 50th sampling unit thereafter is selected.
3. ***Value-weighted selection*** - Sample size, selection and evaluation results in a conclusion in value amounts (e.g. tons of fuel consumed)
4. ***Haphazard selection*** - "The auditor selects the sample without following a structured technique. Although no structured technique is used, the auditor would nonetheless avoid any conscious bias or predictability (for example, avoiding difficult to locate items, or always choosing or avoiding the first or last entries on a page) and thus attempt to ensure that all items in the population have a chance of selection. "
5. ***Block selection*** - Selection of a block(s) of contiguous items from within the population. Block selection cannot ordinarily be used in audit sampling because most populations are structured such that items in a sequence can be expected to have similar characteristics to each other, but different characteristics from items elsewhere in the population. Although in some circumstances it may be an appropriate audit procedure to examine a block of items, it would rarely be an appropriate sample selection technique when the auditor intends to draw valid inferences about the entire population based on the sample.

1.4 Examples on Sampling Procedure

Case 1: Non-Statistical Sampling example based on sampling in scope voyages data

You would like to take a sample from the voyages which are in scope a ship performed in the reporting period.

Scenario 1: The ship made 15 voyages a year.

Scenario 2: The ship made 50 voyages a year.

Scenario 3: The ship made 150 voyages a year.

Scenario 4: The ship made 400 voyages a year.

Scenario 5: The ship made 800 voyages a year.

How many samples do you need to take in each scenario?

Assumption taken: Inherent risk is high and control risk is medium. Then the combined inherent and control risk is high. This means the verifier needs to increase the sample size to decrease the detection risk.

Assumption taken: Inherent risk is low and control risk is low, then the combined inherent and control risk is low. This means the verifier can decrease the sample size as a higher detection risk can be accepted.

Scenario 1: 15 voyages a year.

		Sample size
Combined inherent & control risk	Low	5
	Medium	11
	High	15

Scenario 2: 50 voyages a year.

		Sample size
Combined inherent & control risk	Low	9
	Medium	20
	High	33

Scenario 3: The ship made 150 voyages a year.

		Sample size
inherent & control risk	Low	13
	Medium	29
	High	48

Scenario 4: The ship made 400 voyages a year.

		Sample size
Combined inherent & control risk	Low	16
	Medium	36
	High	60

Scenario 5: The ship made 800 voyages a year.

		Sample size
Combined inherent & control risk	Low	16
	Medium	36
	High	60

Please note that sample sizes in non-statistical sampling are based on professional judgement. Verifiers can identify different levels of sample sizes based on their professional judgement.

Case 2: Statistical Sampling based on sampling in scope voyages data

Data could be verified based on the approach below;

(1) If the number of voyages for significant emissions (accounting for more than 3.3% of the total emissions reported) is less than 30, 100% sampling should be carried out.

(2) If the number of voyages for significant emissions is more than one per month and is more than 30 per year, and if verification is possible for more than 50% of the time in the total number of voyages, then voyages for six months out of a year shall be selected in the descending order of reported emissions and 100% sampling of these shall be carried out.

(3) If no voyages contribute to the significant emission (less than 3.3%) or if even selection and verification of the 50% mentioned above is difficult due to time consideration, then the following table for the minimum sample size should be taken. The verifier will evaluate the inherent risk and control risk and control measures taken by the company. The High, Medium and Low residual risk shall be evaluated. The verifier may increase the minimum sample size based on his professional judgement, professional skepticism and the result of the risk analysis.

Number of voyages of significant emission	% of the Total voyages
0 - 12	100
13 - 30	50
31 – 50	34
51 – 90	21
91 – 150	14
151 – 280	9
281 – 500	5
501 – 1200	3
1200-1500	2

The above table is based on the following assumptions:

a) Optimum use of the time allocated by the verifier for each element of the verification process.
b) Consideration is given to the relationship between the cost of obtaining evidence and the usefulness of the data and information obtained.
c) The table can be revised after due diligence i.e. based on analysis of the time allocated for each vessel, cost vs. sample data usefulness for evaluating materiality of the sampled data.
d) The revision of the above table can be done after the first reporting period i.e. after 31st Dec. 2018.

(4) A random selection method can be implemented when selection and verification of the 50% is difficult. Sampling work may become complicated, so systematic sampling method may be used for sampling.

When sampling 30 voyages from the total voyages, systematic sampling method (equal interval sampling) can be used. Equal interval sampling is performed as described below. Line up the annual voyages sequentially starting from the beginning of the reporting period and assign numbers to the voyages. Take the start number as 1. Divide the total number of voyages by 30 and fix the sampling interval. An example of the number of voyages = 258 is shown below;

(Number of voyages in 1 year)/ (Number of voyages sampled) = 258/30 = 8.6

- ✓ Omitting the fractional part, we take sampling interval d as 8.
- ✓ Take the number for performing the first sampling as `a`.
- ✓ The second sampling number is `a`＋1 x 8
- ✓ The third sampling number is `a`＋2 x 8
- ✓ The fourth sampling number is `a`＋3 x 8
- ✓ The n^{th} sampling number is `a`＋（n－1） x d
- ✓ Do this sequentially until (30-1) = 29.

The general rule is to determine the starting point `a` after generating a random number between 1 and N, but a simple method may be implemented for the first number. That is, select an arbitrary number within the numbers of the sampling interval (1 to 7) and take it as the starting point.

During systematic sampling, “care is needed to ensure that the population list does not have periodicity.” Under this condition, the examples obtained by systematic sampling may be treated almost similarly to the examples obtained by random sampling.

- Confirm that the value shown on the sampled voyages coincides with the emission value reported by the company.
- Furthermore, estimate the total for one year from the 30 sampled voyages. If this value is within 5% of the total submitted by the company, end the sampling.

The calculation method is shown as following

Calculate the total for one year estimated from 30 extracted voyages as:

T total = ((Total voyages for 1 year) / (Number of sampled voyages) x (Total number of sampled voyages)

＝（258/30） x (Total sampled voyages) = 8.6 x (Total sampled voyages)

In this way, compare the calculated value of T total with the total (Ｔ（＾）) for one year submitted by the company.

(5) If ((T total - Ｔ（＾））/Ｔ（＾）) x100 is less than 5%, treat the sampling as complete.

If 5% is exceeded, the sampled number is probably inadequate; therefore, add to the sampled number and perform the sampling and verification for the second time.

(6) Take the sampling number for the second time as a multiple of 30. That is, take 30 x 2 times = 60. If the sampling number is taken as two times, improved accuracy of $\sqrt{2}$ = 1.4 times may be

anticipated. However, even if 60 items are sampled, if the estimated value of the total is not within 5% of the annual total value, increase the sampling number further.

Take the sampling number for the third time as 30 x 3 times = 90.

(7) Repeat the procedure above similarly from here onward. Increase the sampling number until the estimated value of the total obtained from sampled data falls within 5% of the total value for the year.

Case 3: Statistical Sampling based on sampling the numbers of Bunker Delivery Notes (BDNs) used in the reporting period or another data set such as tank sounding readings

Data could be verified based on the approach below;

(1) if detection risk is low then divide by 2 the sampling size

(2) if detection risk is medium then use the sampling size

(3) if detection risk is high then multiply by 2 the sampling size

Size of dataset	Sampling size to reach a 5% materiality level
2 to 8	3
9 to 15	3
16 to 25	5
26 to 50	8
51 to 90	13
91 to 150	20
151 to 280	32
281 to 500	50
501 to 1200	80
1201 to 3200	125
3201 to 10000	200
10001 to 35000	315
35001 to 150000	500

Example :

Ship X from company Y bunkers twice a week. How many BDNs do you need to sample ?

Assumption taken: Inherent risk is high and control risk is medium. Then the combined inherent and control risk is high. This means the verifier needs to increase the sample size to decrease the detection risk.

As detection risk is high, sample size need to be increased. Over the year ship X bunkers 104 times, so the sample size is 40 BDNs.

1.5 Process flow

The following steps that the verifier should consider when it tests a sample of a data set are shown in the diagram below:

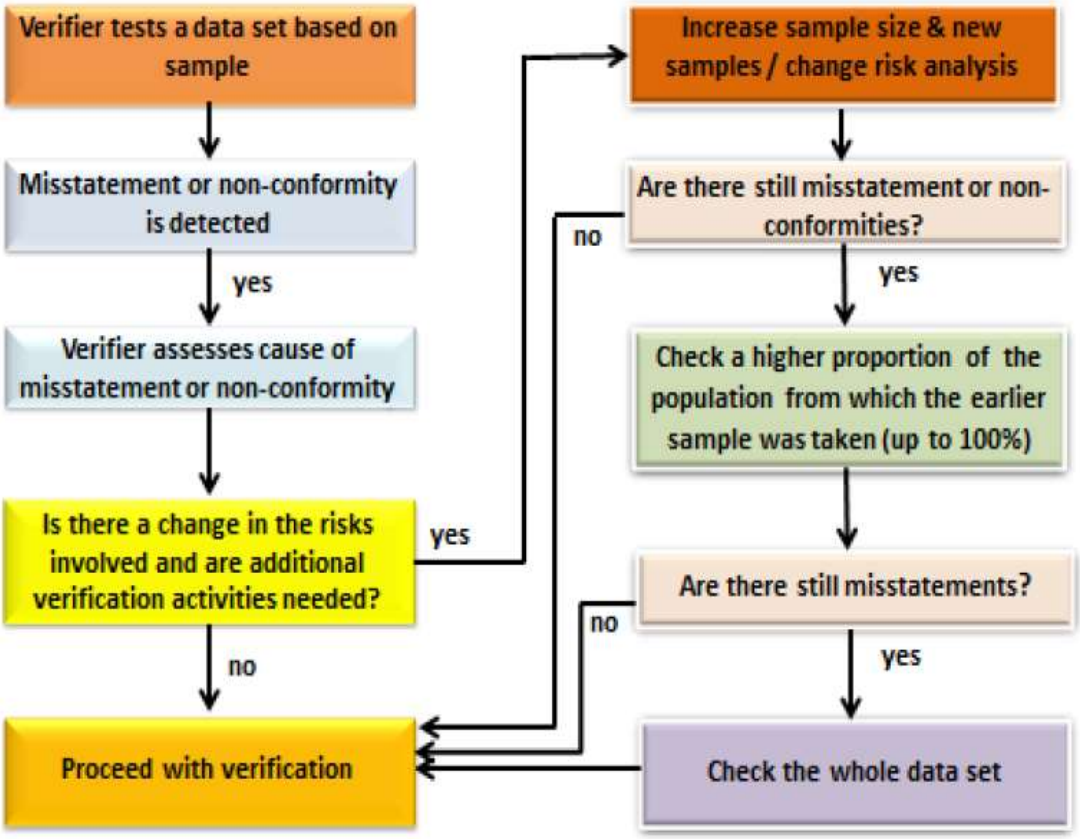

Figure 1: decision tree for sample size