



# **MONITOOL PROJECT**

Methodology proposal for  
EQS adaptation for priority  
metals and for determining  
DGT- thresholds for other  
specific metals.  
WP nr. 4, action nr. 4

## Report/Deliverable by

José Germán Rodríguez – AZTI

María Jesús Belzunce – AZTI

Iratxe Menchaca – AZTI

Joana Larreta – AZTI

Javier Franco – AZTI

## Disclaimer

This project (nº contract: EAPA\_565/2016) is co-financed by the European Regional Development Fund through the Interreg Atlantic Area Programme. The present work reflects only the author's view and the funding Programme cannot be held responsible for any use that may be made of the information it contains.

## Table of Contents

Executive Summary .....	2
1. Introduction.....	2
1. Scope .....	2
2. Proposal 1: prediction based upon parametric regression analysis.....	3
2.1 Evaluation of the need for segmentation of information. ....	3
2.2 Calculation of a good regression model for predictions. ....	4
2.3 Using the regression model to do predictions. ....	4
2.4 Some remarks.....	6
3. Proposal 2: quantile regressions .....	6
4. Proposal 3: minimisation of mismatch.....	8
5. Proposal 4: binomial logistic regression .....	10
6. Main conclusions .....	11
7. References .....	11

## List of figures

Figure 1. Hypothetical relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured by spot sampling (IPC-MS). ....	3
Figure 2. Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured by IPC-MS (spot sampling). Blue line indicate the linear regression adjustment. Red line is the 95% upper prediction interval for new observations. Purple line is the 99% upper prediction interval for new observations.....	5
Figure 3. Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured by IPC-MS (spot sampling). Blue line indicate the linear regression adjustment. Purple line is the 99% upper prediction interval for new observations. Black line is the EQS value for waters. ....	5
Figure 4. Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured by IPC-MS (spot sampling). Blue line indicate the linear regression adjustment. Red line is the 95% upper prediction interval for new observations. Black line is the EQS value for waters. ....	6
Figure 5. Hypothetical relationship the concentration of metal X measured in DGT and the mean concentration of metal X measured by spot sampling. Lines indicate quantile regressions. Percentile 50 <sup>th</sup> is regression modelling the median (in red).....	7
Figure 6. Hypothetical relationship the concentration of metal X measured in DGT and the mean concentration of metal X measured by spot sampling. Lines indicate quantile regressions.....	7
Figure 7. Hypothetical classification of data following EQS for spot sampling (see main text). ....	8

Figure 8. Hypothetical classification of data following EQS for spot sampling (30 ng/L) and hypothetical EQS of 20 ng/L for DGT (see main text). .....	9
Figure 9. Hypothetical classification of data following EQS for spot sampling (30 ng/L) and hypothetical EQS of 30 ng/L for DGT (see main text). .....	9
Figure 10. Hypothetical relationship between the percentage of mismatched data and with different hypothetical values of EQS referred to the DGT concentration. The dashed line marks the minimum mismatch in the classification (see main text). .....	10
Figure 10. Binomial logistic regression of metal concentrations measured by DGT on probability of concentration of metal measured by spot sampling is higher than EQS for spot sampling, using a simulated data set. ....	10

## Glossary

DGT: Diffusive Gradients in Thin films. In graphs refers to the mean value of concentration of a metal measured in DGT by ICP-MS.

EQS: Environmental Quality Standard.

ICP-MS: Inductively Coupled Plasma Mass Spectrometry. In graphs refers to the mean value of concentration of a metal measured in spot samples covering an exposure period.

WFD: EU Water Framework Directive (WFD, 2000/60/EC).

## Executive Summary

This deliverable proposes few strategies for the determination of EQS in the case that DGTs are to be used for determining the concentration of metals in water. There is no single strategy, and the choice of strategy should be based on the variability of the data.

### 1. Introduction

Within the EU Water Framework Directive (WFD, 2000/60/EC), the levels of contaminants in the environment should be evaluated and these levels should be compared against Environmental Quality Standard (EQS). In the case of metals, EQS are defined for water and/or biota. There are EQS that refer to an annual average and others that refer to the annual maximum. Regarding metals in transitional and coastal waters, the EQS refers to the dissolved concentration, *i.e.* the metal concentrations measured in a water sample previously filtered through a 0.45 µm filter or subjected to any equivalent pre-treatment. Therefore, in waters it is usual to measure metals in discrete samples (spot sampling, CIS, 2009).

The Monitool project studies the use of passive samplers as an alternative or complementary system to 'classical' spot sampling. Hence, passive samplers have been used for measuring labile metal concentrations in waters. The Diffusive Gradients in Thin-films (DGT; Davison and Zhang, 1994) is the most extensively used sampler for *in situ* labile metal measurements (Menegário et al., 2017). DGT samplers accumulate metals continuously during the deployment time, usually ranging from several days to weeks. This provides time-weighted average metal concentrations and enables the achievement of lower limits of quantification compared with low-volume water samples. In addition, DGT samplers only accumulate free metal ions and easily dissociable metal complexes, operationally known as DGT-labile concentration, which has been related to observed toxicity in different types of organisms (e.g., Koppel et al., 2019). These advantages might favour the inclusion of DGT technique within monitoring programmes.

One of the limitations of using DGTs in the context of the EU WFD is that there are no defined EQS for DGTs. In this deliverable we make different proposals for the calculation of EQS, from a theoretical point of view. In WP6, instead, EQS calculations are performed based on real data.

### 1. Scope

The objective of this deliverable is to propose different strategies for the calculation of EQS to be used with DGT.

## 2. Proposal 1: prediction based upon parametric regression analysis

The main idea of this proposal is to predict the mean concentration of a metal (or the mean concentration plus a confidence interval) measured in discrete water samples by 'classical' spot sampling based on the concentration measured by DGT for new observations. This can be carried out when there is a good fit in a regression analysis between mean concentration of metal measured in water samples (spot sampling) and the concentration measured in DGT. This proposal is based upon several steps, that should be carried in a transparent way.

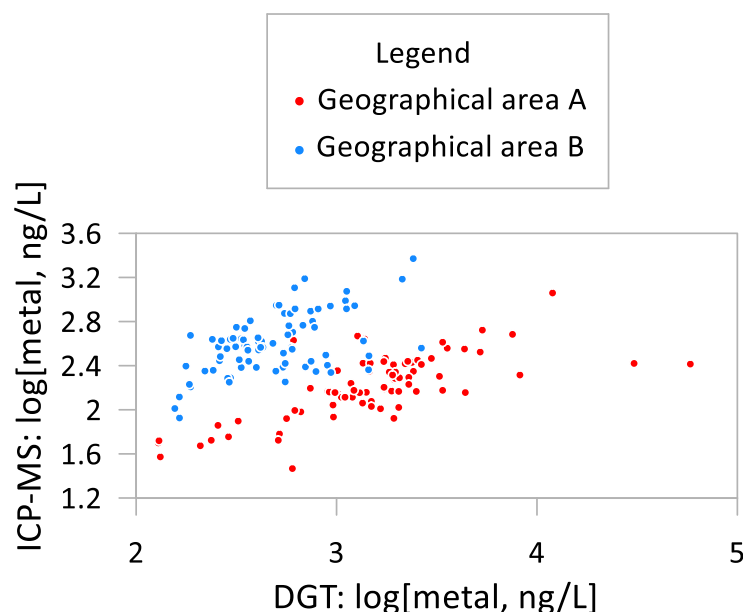
### 2.1 Evaluation of the need for segmentation of information.

The exploration of the DGT/spot sampling relationship for a metal in relation to different factors/variables (salinity, temperature, location coastal/estuarine, geographical location, etc.) should be carried out in order to know if there is a necessity of segmentation of the datasets. This exploration should be done considering the distribution of the data. For example, if a dataset follows a lognormal distribution, the exploration can be done with log transformed data. The transformations can be different for each variable.

The exploration of the relationships can be done (among others) by

- 1) Representing scatterplots using different symbology for the values of factors/parameters (e.g. Figure 1).
- 2) Representing scatterplots of the variability of residual plots from preliminary regression analyses.
- 3) Representing component effects plot from preliminary multiple regression analyses. These plots represent the portion of the fitted regression model corresponding to any single independent variable (e.g., salinity or temperature). Caution should be taken with multicollinearity.

The exploration should clarify if the information has to be segmented in different datasets or not. This exploration should be done also in proposal 2, 3, and 4 of this deliverable.



**Figure 1.** Hypothetical relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured in discrete water samples by means of spot sampling (ICP-MS).

## 2.2 Calculation of a good regression model for predictions.

The calculation of a regression model required a carefully exploration of the data, including (among others):

- a) Missing values: cases with missing values should be removed from database.
- b) Outliers: exploration should be carried out to identify possible univariate and multivariate outliers.
- c) Influencing points exploration.

After the exploration some values can be removed for the calculation of the regression model. Within the same data, different regression models can be calculated. The following indicators/actions can be used in order to evaluate which model is more appropriated for predictions:

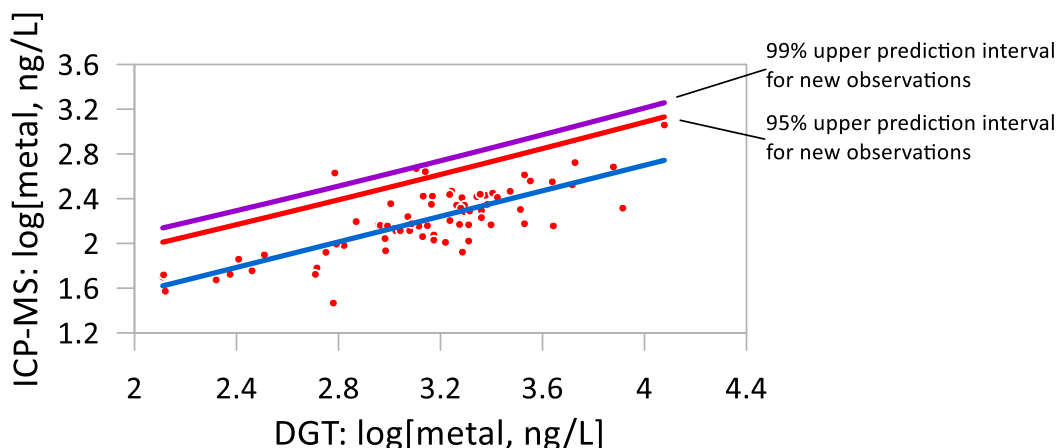
- a) Adjusted R-squared and predicted R-squared.
- b) P-values for the independent variables.
- c) Simplicity/Complexity of the model.
- d) Prediction intervals.
- e) Exploration of Residual Plots.

## 2.3 Using the regression model to do predictions.

When there is a regression model that appears to produce unbiased predictions, it is possible to predict new observations. This prediction can be done only for the range of values included in the model (although it can also be projected beyond the observation range, in justified cases). Since the new observations will be refer to observations in DGT concentration values, these will be considered as independent variable (although the opposite may also be the case, depending on whether one of the variables is considered to explain the other or one lacks error). The prediction interval is the range where a single new observation is likely to fall. To ensure that the proposal of EQS for DGT is at least as protective as the available EQS for marine water, the 95% or the 99% prediction interval can be used as EQS for DGT. This gives a stricter EQS (i.e., the probability of non-compliance is higher if DGT EQS is used than water EQS), and therefore the mean predicted value can also be used without adding the confidence interval.

In the Figure 2 it is shown a hypothetical log-log linear relationship between the concentration of metal X measured in DGT and the mean concentration of metal X measured in water by IPC-MS (spot sampling). The 99% and 95% upper prediction intervals for new observations are shown.



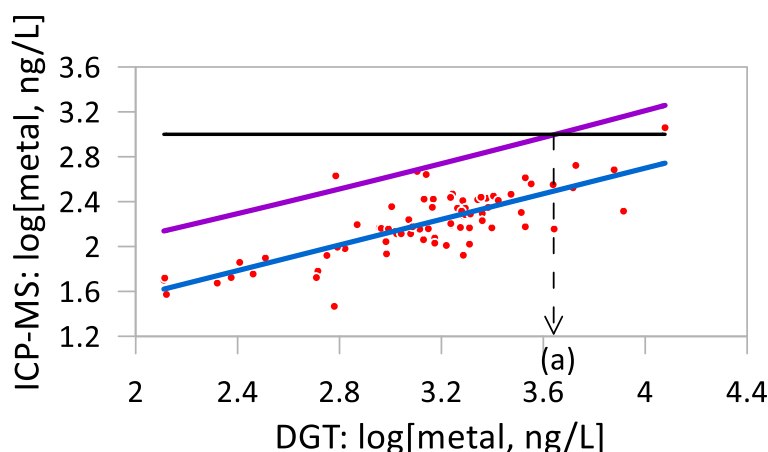


**Figure 2.** Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured in water by IPC-MS (spot sampling). Blue line indicates the linear regression adjustment. Red line is the 95% upper prediction interval for new observations. Purple line is the 99% upper prediction interval for new observations.

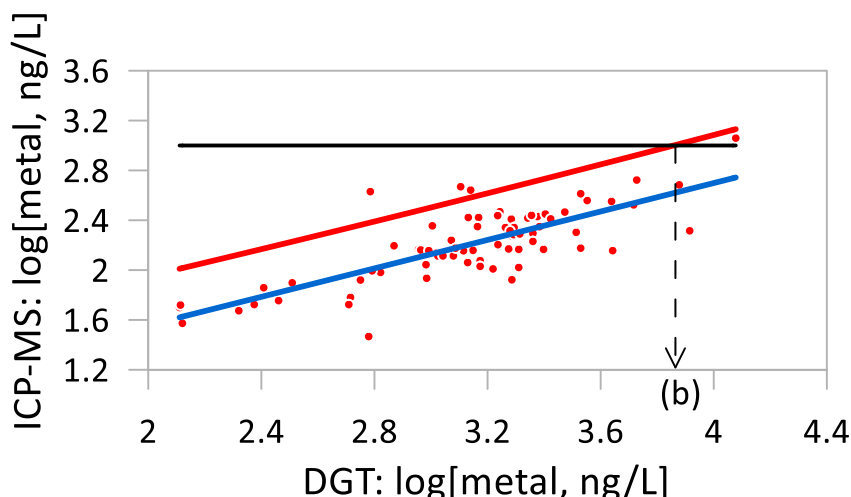
If there is a hypothetical EQS for metal in water of 1000 ng/L, this equals to a  $\log_{10}$ -transformed value of 3. The point where this value crosses with the 99% upper prediction interval for new observations correspond to the (a) value of Figure 3. This (a) value corresponds to the metal concentration measured with DGT (for new observations) which would imply that the mean metal concentration measured by the 'classical' spot sampling and filtering method would be lower than the EQS (i.e., the EQS for spot sampling and filtering) with a probability of 99%. In this case the (a) value is 3.65616, that is equal to a concentration of 4531 ng/L measured in DGT.

In a similar way, the (b) value in Figure 4 equals to a 7453 ng/L measured in DGT.

In other words, it can be proposed as EQS for the metal in DGT as 4531 ng/L (with a 99% of confidence) and 7453 ng/L (with a 95% of confidence).



**Figure 3.** Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured in water by IPC-MS (spot sampling). Blue line indicates the linear regression adjustment. Purple line is the 99% upper prediction interval for new observations. Black line is the EQS value for waters.



**Figure 4.** Hypothetical linear relationship between the logarithm of the concentration of metal X measured in DGT and the logarithm of the mean concentration of metal X measured in water by IPC-MS (spot sampling). Blue line indicates the linear regression adjustment. Red line is the 95% upper prediction interval for new observations. Black line is the EQS value for waters.

#### 2.4 Some remarks.

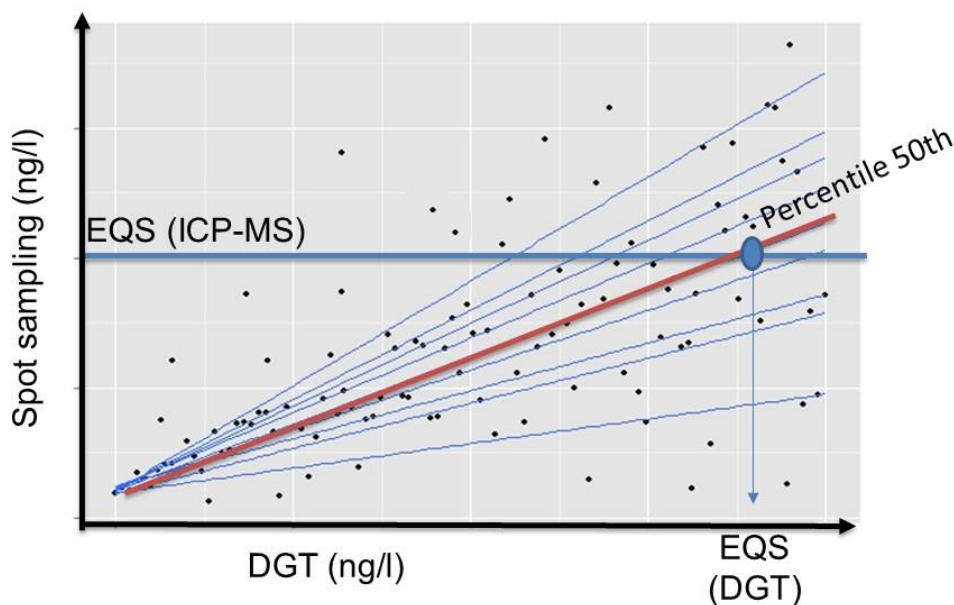
- In those cases where the relationship between the results given by both methodologies (DGT and spot sampling) could be significantly explained by other variables or factors (e.g., salinity or pH), more complex predictive models are needed, as long as they can predict with confidence intervals for new observations.
- Different models can be used, one can opt for a model with a better overall fit or one can opt for a model with a better fit in the data area where the EQS is located.
- In the case that a confidence interval is used, it should be aware that the resulting  $EQS_{DGT}$  will be stricter than the  $EQS_{marine\ water}$ . In addition, the degree of mismatch (in the case of comparing both methodologies, i.e. DGT and spot sampling) is likely to be greater if the mean predicted value is used without adding the interval.

### 3. Proposal 2: quantile regressions

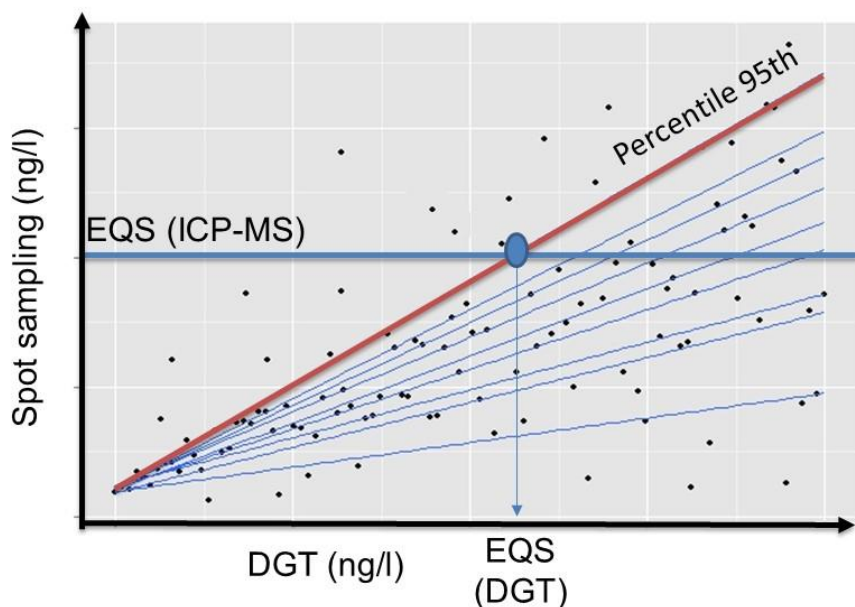
In some cases, the dispersion of the data makes difficult to construct a parametric fit model. The Figure 5 shows a hypothetical example in which there is a direct relationship between the concentration measured by means of spot sampling and the concentration measured with DGT, but in which there is a high dispersion. In red it is shown the quantile regression that estimates the conditional median of the variable response (in this case the concentration of a metal measured in water by means of spot sampling). Using this regression, it is possible to calculate the  $EQS_{DGT}$  as indicated in the Figure 5.

In a similar way to that indicated in the previous section, the  $EQS_{DGT}$  can be calculated with a confidence margin. In this case, for example, the regression modelling the 95<sup>th</sup> percentile (or other) can be used instead of the regression modelling the median (Figure 6). In the case that a percentile higher than the median is used, it should be aware that the resulting  $EQS_{DGT}$  will be stricter than the water  $EQS_{marine}$ .

water (or  $EQS_{(ICP-MS)}$  or  $EQS_{(spot\ sampling)}$  in Figures). In addition, the degree of mismatch (in the case of comparing both methodologies, i.e. DGT and spot sampling) is likely to be greater if the mean predicted value is used without adding the interval.



**Figure 5.** Hypothetical relationship between the concentration of metal X measured in DGT and the mean concentration of metal X measured in water by means of spot sampling. Lines indicate quantile regressions. “Percentile 50<sup>th</sup>” is regression modelling the median (in red).



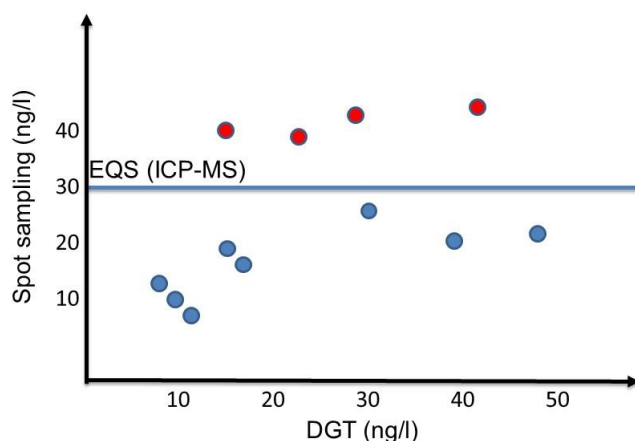
**Figure 6.** Hypothetical relationship between the concentration of metal X measured in DGT and the mean concentration of metal X measured in water by means of spot sampling. Lines indicate quantile regressions. “Percentile 95<sup>th</sup>” is regression modelling the Percentile 95<sup>th</sup> (in red).

#### 4. Proposal 3: minimisation of mismatch

This categorical approach is based on the proposal of Phillips *et al.*, 2018. These authors propose the following method to calculate a boundary between classes of a nutrient (e.g. good/moderate boundary) when there is a relationship between the concentration of that nutrient and a response of the biota in which the boundary (good/moderate) is previously defined. This has been proposed within the scope of the Water Framework Directive.

In our case it is a matter of looking for the limit at which there is less mismatch between the two considered methodologies.

In a first step, the data are classified according to the methodology for which the EQS is established (i.e., spot sampling). For example, in Figure 7, four data have higher concentration than the EQS (spot sampling) and eight data have lower concentration.



**Figure 7.** Hypothetical classification of data following EQS for spot sampling (see main text).

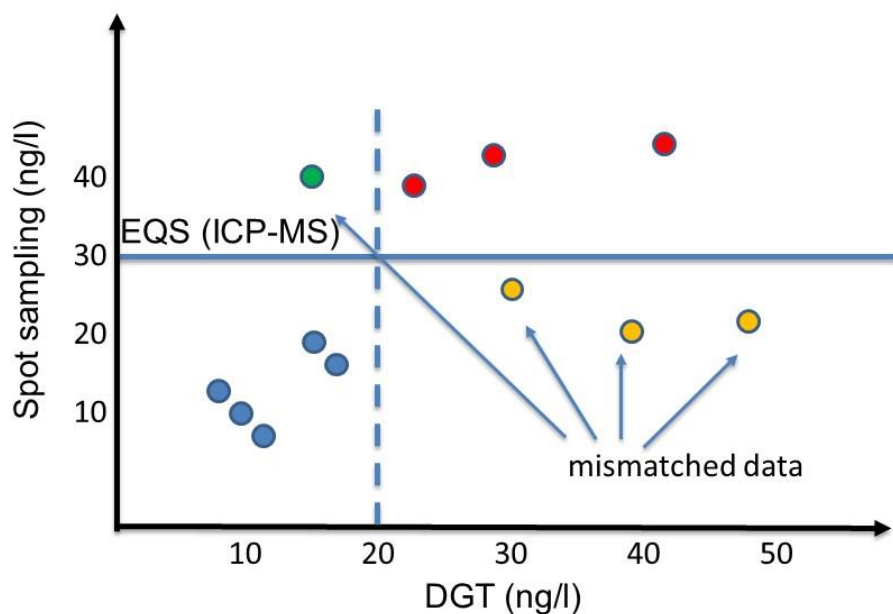
Once the previous step has been carried out, it is necessary to calculate, for all possible hypothetical concentrations measured in DGT, the degree of mismatch between both EQS (i.e., the  $EQS_{spot\ sampling}$  and the hypothetical  $EQS_{DGT}$ ). For example, with a hypothetical  $EQS_{DGT}$  of 20 ng/L there is a classification disagreement in 4 data and agreement in 8 data (Figure 8). For example, with a hypothetical  $EQS_{DGT}$  of 30 ng/L there is classification disagreement in 6 data and agreement in 6 data (Figure 9).

There are two possibilities of mismatch:

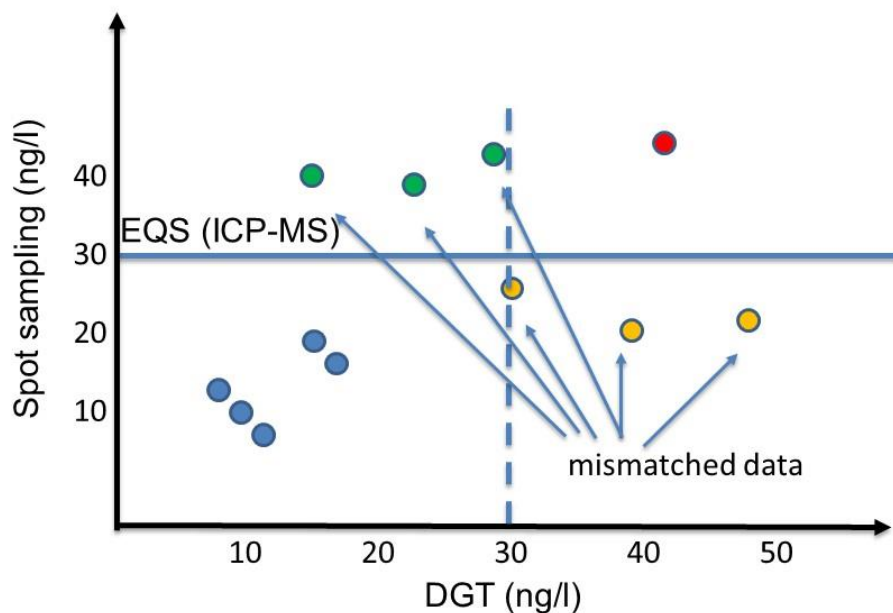
Data with  $[DGT] > [EQS_{DGT}]$  and  $[Spot\ sampling] < [EQS_{spot\ sampling}]$  (in orange in Figure 8 and Figure 9).

Data with  $[DGT] < [EQS_{DGT}]$  and  $[Spot\ sampling] > [EQS_{spot\ sampling}]$  (in green in Figure 8 and Figure 9).

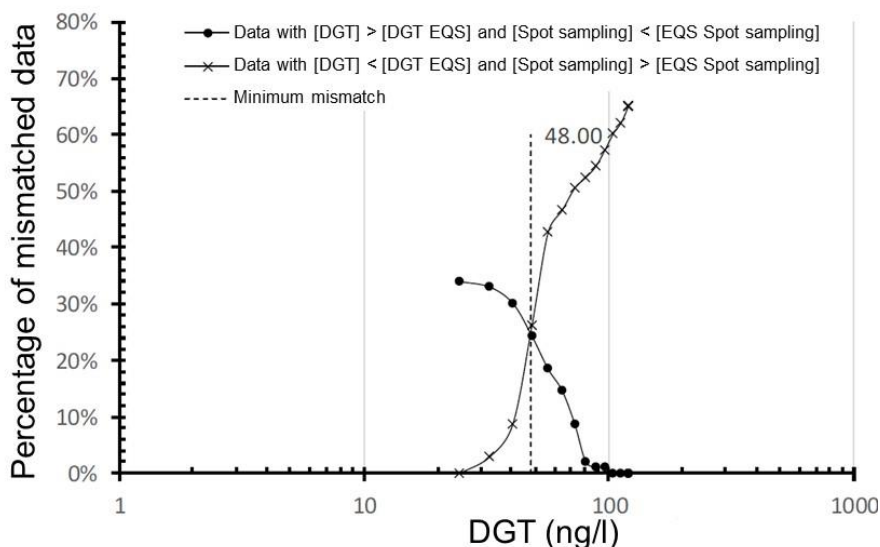
The DGT concentration involving the least mismatch can be proposed as EQS. This calculation can be automated with the software proposed by Phillips *et al.*, 2018. For example, in the hypothetical case represented in the Figure 10, the DGT concentration with the lowest mismatch is 48 ng/L.



**Figure 8.** Hypothetical classification of data following EQS for spot sampling (30 ng/L) and hypothetical EQS of 20 ng/L for DGT (see main text).



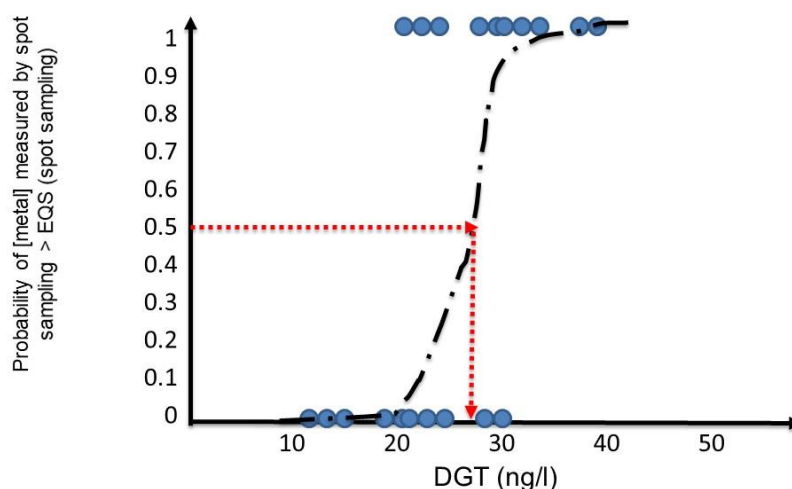
**Figure 9.** Hypothetical classification of data following EQS for spot sampling (30 ng/L) and hypothetical EQS of 30 ng/L for DGT (see main text).



**Figure 10.** Hypothetical relationship between the percentage of mismatched data and with different hypothetical values of EQS referred to the DGT concentration. The dashed line marks the minimum mismatch in the classification (see main text).

## 5. Proposal 4: binomial logistic regression

This approach is like the previous one, classifying the data into two categories: those that exceed the EQS established for spot sampling and those that do not. A logistic relationship is used to calculate the concentration in DGT from which the probability of exceeding the EQS for spot sampling is 50% (Figure 11). A lower probability can also be used, which would result in a lower EQS for DGTs. In the case that a lower probability is used, it should be aware that the resulting  $EQS_{DGT}$  will be stricter than the water EQS. In addition, the degree of mismatch (in the case of comparing both methodologies, i.e., DGT and spot sampling) is likely to be greater if the mean predicted value for a probability of 50% is used without adding the interval.



**Figure 11.** Binomial logistic regression of metal concentrations measured in DGT on probability of concentration of metal measured in water by spot sampling is higher than EQS for spot sampling, using a simulated data set.

## 6. Main conclusions

There are several strategies for calculating EQS for use with DGT. In those cases where there is a very close relationship between the results obtained with the 'classical' spot sampling method and the DGT-based method, parametric regressions can be used. When the relationship between the two methods is not very close, other methods can be used.

It is important to consider whether it is desired to calculate an average value or an average value plus a confidence interval.

## 7. References

CIS, 2019. Common implementation strategy for the Water Framework Directive (2000/60/EC)-Guidance Document No. 19: Guidance on surface water chemical monitoring under the Water Framework Directive.

Davison, W., Zhang, H., 1994. In situ speciation measurements of trace components in natural waters using thin-film gels. *Nature* 367, 546-548.

Phillips, G., Kelly, M., Teixeira, H., Salas, F., Free, G., Leujak, W., Lyche Solheim, A., Varbiro, G., Poikane, S. 2018. Best practice for establishing nutrient concentrations to support good ecological status, EUR 29329 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-92906-9, doi:10.2760/84425, JRC112667.

Koppel, D.J., Adams, M.S., King, C.K., Jolley, D.F., 2019. Diffusive Gradients in Thin Films Can Predict the Toxicity of Metal Mixtures to Two Microalgae: Validation for Environmental Monitoring in Antarctic Marine Conditions. *Environmental Toxicology and Chemistry* 38, 1323-1333.

Menegário, A.A., Yabuki, L.N.M., Luko, K.S., Williams, P.N., Blackburn, D.M., 2017. Use of diffusive gradient in thin films for in situ measurements: A review on the progress in chemical fractionation, speciation and bioavailability of metals in waters. *Analytica Chimica Acta* 983, 54-66.