

# REVISION OF STATISTICAL METHODS REPORT-PLC-8 2021



# LATEST VERSION

---

## STATISTICAL ASPECTS IN RELATION TO BALTIC SEA POLLUTION LOAD COMPILATION

Task under HELCOM PLC -7 project

Søren E. Larsen<sup>1</sup>

Lars M. Svendsen<sup>2</sup>

<sup>1</sup> Aarhus University, Department of  
Bioscience

<sup>2</sup> Aarhus University, DCE - Danish Centre for  
Environment and Energy

No. 137

2019



# CONTENTS:

---

1. Introduction 5
  2. Data gaps and outliers 8
  3. Uncertainty of inputs (yearly input from a specific country or area) 11
  4. Hydrological normalization of nutrient inputs 17
  5. Trend analysis, change points and estimation of change 21
  6. Testing fulfilment of BSAP reduction targets 26
  7. Step by step analysis illustrated by HELCOM data examples 31
  8. Discussion and recommendations 34
  9. References 36
- Annex 1: Mathematical description of the Mann -Kendall trend test 38
- Annex 2: List of 95% percentiles and 97.5 percentiles of the t-distribution for the different possible combinations of degrees of freedom (df) 41



# INTRODUCTION

---

Will be revised according to the new and revised contents of the revised report.



# DATA GAPS AND OUTLIERS

---

Will not do a major revision of this chapter, mainly look for adding more text so concepts are more fully explained.

Add a few examples.

Unless , some have suggestions for subjects they think need to be addressed in this chapter.



# UNCERTAINTY OF INPUTS

---

Will elaborate on the DUET\_H/WQ method with more realistic examples.

Will furthermore elaborate on the "Danish" method, which I have presented a couple of times at earlier meetings.

Compare the two methods, by using examples.

Work on expanding these tables, with larger rivers and other load calculation methods:

TN (Linear interpolation method):

0-50 km <sup>2</sup> :	Bias: -1% to -3%;	Precision: 1-3%
50-200 km <sup>2</sup> :	Bias: -0.7% to -3%;	Precision: 1-3%
>200 km <sup>2</sup> :	Bias: -1% to -4%;	Precision: 2-5%



---

TP (Linear interpolation method)

0-50 km <sup>2</sup> :	Bias: -16% to -27%;	Precision: 18-37%
50-200 km <sup>2</sup> :	Bias: -2% to -5%;	Precision: 9-13%
>200 km <sup>2</sup> :	Bias: -2% to -4%;	Precision: 3-8%

Tables are used for the "Danish" method. Need help to expand these tables.

Methods: Linear interpolation: Sweden, Germany, Estonia, Denmark.

Monthly mean: Finland (changed?), Poland, Russia (change?), Latvia, Germany.

Lithuania?

Range of catchment areas.



---

The number for the precision is for yearly load on one single stream station, and include uncertainty from laboratory for concentrations and flow measurements, seasonal variation in concentrations and riverflow and method uncertainty from load calculation.

Bias is from the load calculation method. Laboratory bias? Flow measurement bias?





# HYDROLOGICAL NORMALIZATION

---

My plan is to introduce a new method for flow-normalization.

The new method is based on statistical time series analysis.

Will test the new method will on a number different datasets from the Helcom database.

Will not go further into this at the moment – need to see how the method perform on HELCOM data.



# TREND ANALYSIS, CHANGE POINTS AND ESTIMATION OF CHANGE

Not very much to revise here: more explaining text here and there, perhaps a few more examples as well. But I will for sure cover these three subjects:

- Model for the first part of the time series?
- Trends in discharge, how to adjust for that. I have a few ideas for this, which I need elaborate on.
- Non-linear trend models. If needed and if they give better fit without being an overfit.



# TESTINGFULFILMENT OF BSAPREDUCTIONTARGETS

---

This chapter only need a small revision – more text, more examples, especially.

See if the description of the statistical method for the analysis of the last year estimate against the reference period can be improved, text, examples.



# DATA EXAMPLES

---

We will include even many more new and up to date examples, and provide as many calculations steps as possible.

At least one example on all methods/ calculations in the report.



# DISCUSSION RECOMMENDATIONS

---

Will be updated according to the new version of the method chapters.



# ANNEXES

---

In the latest version we have included two annexes:

- Mathematical description of the Mann-Kendall trend test (needed?)
- List of percentiles of the t-distribution

Two new planned:

- List of mathematical symbols/ notation
- Simple mathematical operations – like rules for logarithmic functions

Wishes for additional annexes?

- SAS-programs, R-programs, other?



# TO SUM UP:

---

So generally:

- More text
- More examples

Specific :

- Major revision of the chapter on uncertainty
- Major revision on hydrological normalization
- Annex with mathematical symbols
- Annex with simple mathematical operations/ calculation rules
- Software programs/ coding?

Deadline: 1st draft – late spring, early summer 2021.



” “All models are wrong, but some are useful” -  
George Box







AARHUS  
UNIVERSITET