Supplementary Material

A. Summary of dataset

A.1. Trophic state of lakes

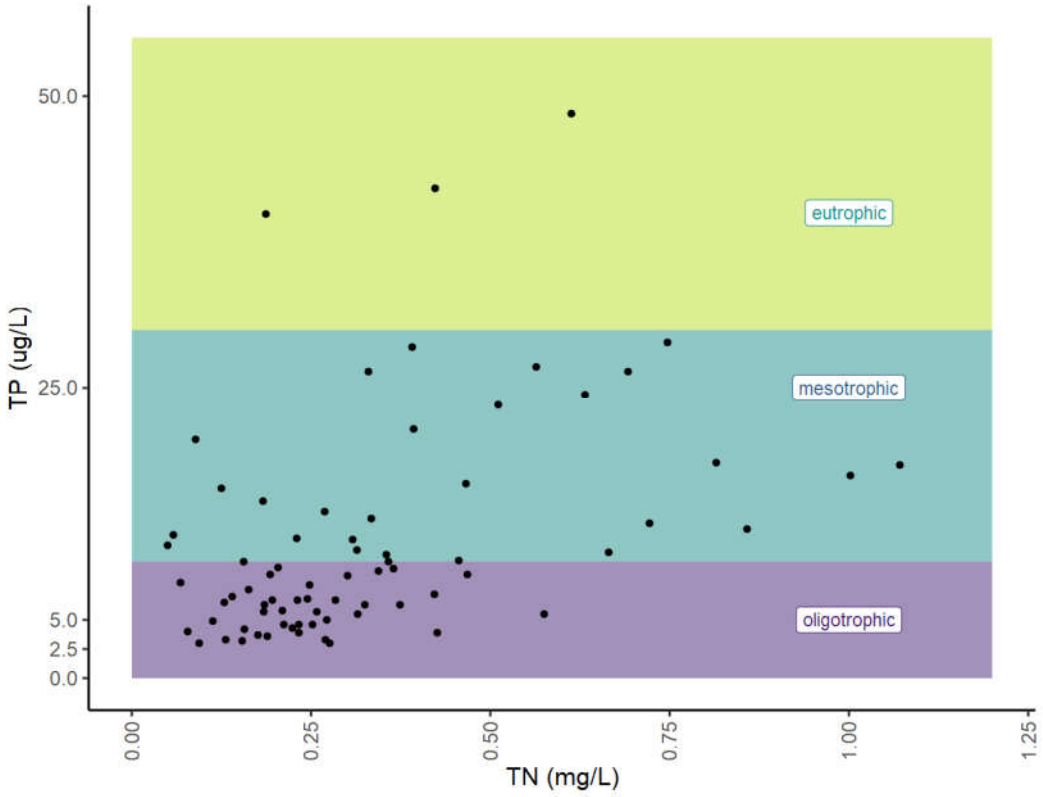

Figure S1. Repartition of the 73 lakes depending on their trophic state, based on Carlson's Trophic State Index[15],[16]. The horizontal division show the trophic state categories based on total phosphate concentration (TP). The x-axis corresponds to the concentration in total nitrogen (TN).

A.2. Summary of dataset

Table S1. Summary of the dataset from the CBA 100 lakes survey

Parameter	Units	Minimum	Q1	Median	Q3	Maximum	Missing values
RR	µM/h	0.455	1.32	2.20	4.18	47.3	0
RRn		0.0842	0.175	0.341	0.689	7.39	0
BdgT	h	1.14	2.26	2.89	3.45	17,6	0
Dissolved organic carbon (DOC)	mg/L	1.53	4.76	7.56	10.6	44.0	0
Specific UV absorbency (sUVa)	L/mg.cm	0.0031	0.0271	0.0372	0.0471	0.0939	0
Specific Visible Absorbancy (sVISa)	L/mg.cm	$1.9x10^{-4}$	$2.8x10^{-3}$	$4.3x10^{-3}$	$6.3x10^{-3}$	$1.39 x10^{-2}$	0
Specific absorbance ratio (SAR)	Abs. ratio	4.5	7.4	8.0	8.8	140	1

Parameter	Units	Minimum	Q1	Median	Q3	Maximum	Missing values
Spectral slope (SR)	Abs. ratio	0.29	0.78	0.84	0.96	2.5	2
pH		5.03	6.39	6.76	7.16	7.72	0
Alkalinity	meq/L	0.0125	0.0910	0.153	0.244	2.16	1
Temperature	°C	0.20	3.30	5.30	7.70	11.00	0
Electrical conductivity (EC)	µS/cm	6.16	18.23	27.00	57.21	243.00	0
Calcium (Ca)	mg/L	0.550	1.520	2.440	4.510	34.71	0
Magnesium (Mg)	mg/L	0.000	0.315	0.560	1.035	4.260	1
Sodium (Na)	mg/L	0.280	0.800	1.600	3.180	21.500	0
Potassium (K)	mg/L	0.060	0.200	0.420	0.880	3.1700	0
Iron (Fe)	mg/L	0.000	0.030	0.070	0.230	1.410	0
Aluminium (Al)	mg/L	0.000	0.0500	0.120	0.210	0.480	0
Sulfate (SO4)	mg/L	0.422	1.55	2.22	3.72	46.0	0
Chloride concentration (Cl)	mg/L	0.260	0.260	1.68	4.85	36.3	0
Bromide (Br)	mg/L	0.0532		0.165		0.801	0
Dissolved nitrogen (DN)	mg/L	0.04900	0.1380	0.2080	0.3150	0.8370	0
Dissolved phosphorus (DP)	µg/L	0.50	2.5	3.8	5.6	15	0
Carbon to nitrogen ratio (C:N)	Molar ratio	9.334	25.90	40.02	52.25	450.5	0
Carbon to phosphate ratio (C:P)	Molar ratio	$4.2x10^2$	$3.0x10^3$	$5.0x10^3$	$8.3x10^3$	$1.3x10^5$	0
O_2	µM	345	376	397	414	457	1
CO_2	µM	7.84	46.1	75.5	128	484	1
N_2O	nM	13.3	16.2	18.5	20.5	337	1
CH_4	nM	5.73	34.7	101	336	$3.97x10^3$	1
Bacterial abundance	count/mL	$4.12x10^4$	$3.01x10^5$	$6.05x10^5$	$7.36x10^5$	$1.10x10^6$	0

B. Biodegradability protocol

Introduction

To characterize the biodegradability of DNOM the decline in oxygen concentration is measured during incubation of water samples containing DNOM, bacteria and nutrients (dipotassium phosphate and ammonium nitrate

During the incubation, the bacterial community goes through four phases:

- a "lag phase" during which the bacterial community is adapting to its new environment.
- an "exponential growth" phase, during which the bacterial community is growing by cell division
- a "stationary growth" phase, during which the increase in the number of cells is balanced by the number of cells deaths. The energy derived from metabolism supports cell maintenance rather than cell division. It usually happens when a substrate (DNOM) or a nutrient (ammonium nitrate or phosphate) is getting limited.
- A "death phase" when cell death exceeds cell growth.

During the growth phases, bacteria consume dioxygen, of which the concentration in the water samples declines. The slope of the oxygen concentration curve gives the respiration rate of the bacterial community and serves as a proxy of the rate at which the DNOM is degraded.

Equipment

PreSens © Oxygen Sensor Vials (SensorVial SV-PSt5-4mL) are used to measure the O2 concentrations in the sample during incubation. The vials are placed in an incubator maintaining a constant 25 °C temperature during the analysis. Oxygen sensors are situated at the bottom of each vial, which are set on a plate comprising 24 vials. PreSens sensors measure oxygen concentration by fluorescence decay: the sensors contain a fluorescent dye that is excited by a flash of light at regular intervals. The oxygen in the sample quenches the fluorescence. The fluorescence suppression is monitored by the PreSens instrument and converted to oxygen concentration using the Stern-Volmer equation[1].

Procedure

Inoculum bacteria are from a previously collected 50 L sample of raw water from NIVAs ecological monitoring station Lake Langtjern[2], north of Oslo. This site was selected as source for the inoculum bacteria, as this is a dystrophic lake, and to ease comparison with data from previous studies of biodegradability of DNOM (Gundersen et al. In prep., Mosleth Færgestad 2019, Lee Ong 2018 and Martinez Francés 2017). The 50L were filtered through a 2 μm filter to remove zooplankton and stored in a tank at 10ºC.

Day 0: Sampling

1) Sample water in a clean bucket, rinsed 3 times.
2) Filter 50 mL of water sample through 0,2 μm filter cartridge, in order to remove most of the bacteria.
3) Store in a 50 mL polyethylene bottle in a cold and dark room.

[1] Principle of Optical Oxygen and pH Sensors, Sarina Arain, PreSens
[2] https://www.niva.no/en/services/environmental-monitoring/langtjern

Day 1: Inoculum preparation

The inoculum was prepared from stored Lake Langtjern water. The following procedure was applied to each batch of inoculum:

4) Draw 100 mL from the tank in an Erlenmeyer 250 mL
5) Add 1 mL of nutrient solution
6) Close the Erlenmeyer and wrap in aluminium foil
7) Shake on shaking table or use a magnetic stirrer to homogenize the solution for 3 days.

Day 4: Incubation start

Nutrients solution

A fixed amount of nutrients is added to the water samples in order to ensure that they are no limiting factor for the biodegradation.

The concentration used here is 5 mM of dipotassium phosphate and 5 mM of ammonium nitrate.

Example: preparation of 100 mL nutrient solution

- Weight 0,114 mg K_2HPO_4
- Weight 0,040 mg NH_4NO_3
- Pour in a 100 mL graduated flask

Samples preparation

Organize the SensorVials on a plate with 24 slots as illustrated in Figure 1.

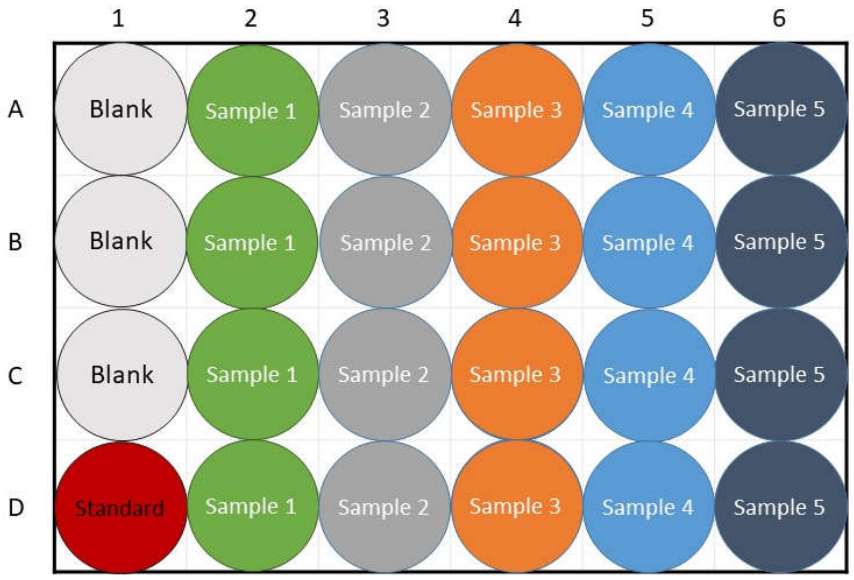

Figure S2- Organisation of a sensor plate

- **Blanks:** 5 mL Type 1 water (fresh)
- **Standard:** 5 mL of a known solution of DNOM + 50 µL nutrients + 50 µL inoculum
- Samples 1-5

- Rinse 50 mL beakers with some of the filtered water
- Add 250 µL of the inoculum
- Add 250 µL of the nutrient solution
- Measure 25 mL (use pipette) of the water sample
- Shake
- Put 5 mL of the final solution in 4 vials of the Sensors plate.
- Close with clean caps
- Seal with Parafilm

NB: 5 mL is the maximum capacity of each vial. Fill them slightly above the maximum volume to ensure that there are no bubbles in the vials.

Incubation

Place the Sensors Plate on the Sensor Dish Reader (SDR), in the incubator. Make sure the vial "A1" is placed on the right spot (the SDRs have the same numbering as the sensor plates). **NB: Note which plate goes on which SDR.**

1) In the SDR_v4.0.0 software, click on "Connect SDRs. In the SDR_v4.0.0 software, click on "Connect SDRs. One "SDR" tab should open for each connected SDR.

For each SDR tab:

2) Set "parameter" on "Oxygen"
3) Set "Batch No" on the batch number corresponding to the plate you placed on this SDR (see Table 1.1).
4) Set "Meas Temp" to 25 (temperature of the incubator).
5) Set "Oxygen Units" on "cO2 (µmol/L)"
6) Set "Interval" on "Min", enter "3". (The software will warn you that this change will apply to all SDR. Accept)
7) Click on "Log Measurement". Save the measurement file under the selected folder.
8) Once all the SDR tab are filled and the measurements logged, click on "Start"

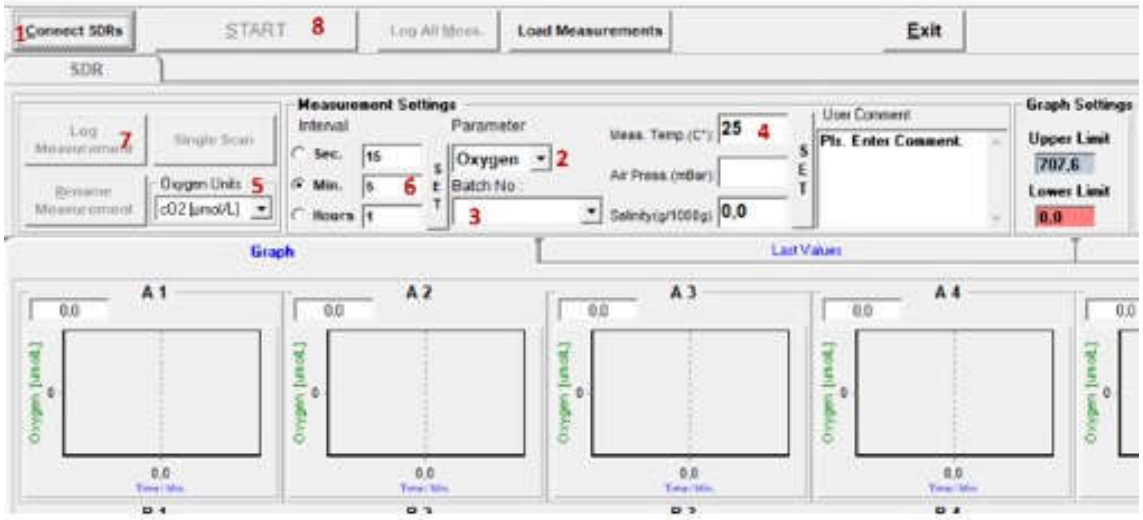

Figure S3 - Snapshot of the SDS interface

The sensors are now measuring the oxygen concentration.

Day 6: End of experiment

Incubation stop

The standard time for incubation is 30h.

At the end of the incubation, select "stop measurement" in the SDR software and export the measurement files (File/Export Measurements) in excel format. NB: each tab has to be exported separately.

Cleaning

- Remove the parafilm
- Empty the vials/caps from the sample and fill them with ethanol 15%.
  The longer the vials soak in ethanol, the better. 12 hours is perfect.
- Empty the ethanol and rinse with fresh type 1 water.
- Put in the oven to dry at 50 °C for at least 6 hours
- If the vials must be stored, wrap them in aluminium foil. Otherwise, they can be used directly after drying for a new measurement round.

Potential issues

1. Influence of temperature

The concentration of O_2 is calculated from the Sterm-Volmer equation and is therefore dependant on a temperature constant. Therefore, the value of oxygen concentration is not reliable before the temperature is stabilised.

2. No clear decrease of oxygen concentration

Sometime no clear decrease of oxygen appears. This could come from several factors:

Possible causes	Solution
Fail in the preparation of the inoculum/inoculum too old	Use an inoculum aged of 3 to 5 days

The inoculum and nutrients are heterogeneously spread in the vials	Shake twice if you must before putting in the vials
Leakage of oxygen	Check the parafilming
Biofilm in the vials	Repeat the cleaning procedure with a longer time soaking in ethanol
DNOM not labile/not enough DNOM	If there is no other reason why there is no decrease of oxygen, then it is a result in itself and indicates that there is few or no labile DNOM in the sample

Anticipated results

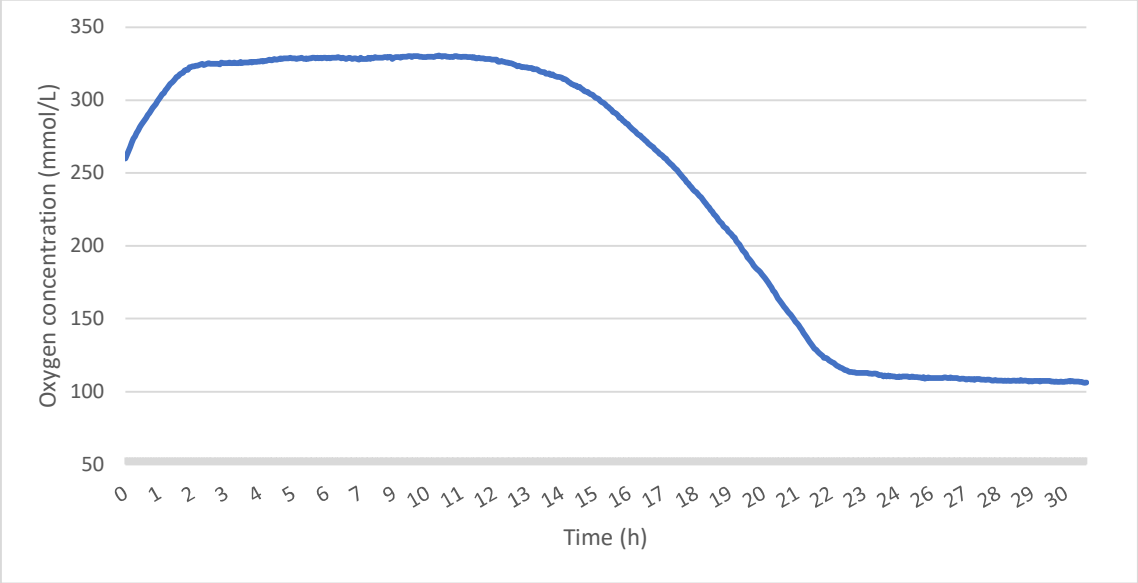

Figure S4 - Example of the oxygen concentration in a natural lake water sample

References

Francés, E. M. (2017). Biodegradability and Spectroscopic Properties of Dissolved Natural Organic Matter Fractions Linked to Hg and MeHg Transport and Uptake. (MSc Mastergradsoppgave). University of Oslo, Oslo, Norway. Retrieved from CLIMER CRIStin database.

Færgestad, E. M. (2019). *Biodegradability and spectroscopic properties of DNOM affected by mercury transport and uptake.* (MSc). University of Oslo, DUO. Retrieved from http://urn.nb.no/URN:NBN:no-73572

Ong, S. H. L., & Vogt, R. D. (2018). *Biodegradability and Mercury Fractions in Dissolved Natural Organic Matter Size Fractions*. Retrieved from Oslo, Norway: http://urn.nb.no/URN:NBN:no-66125

C. Multiple imputation process

Missing values were imputed by multiple imputation (**Error! Reference source not found.**) using the "mice" package in R[15]. Multiple imputation uses other variables of the dataset to calculate the missing value. Ten imputations were performed to reduce the uncertainty over a particular value. All statistical analysis were performed for each of the complete dataset and the results were pooled into a final set of model estimates. The pooling differs according to the type of estimates obtained from the statistical analysis applied.

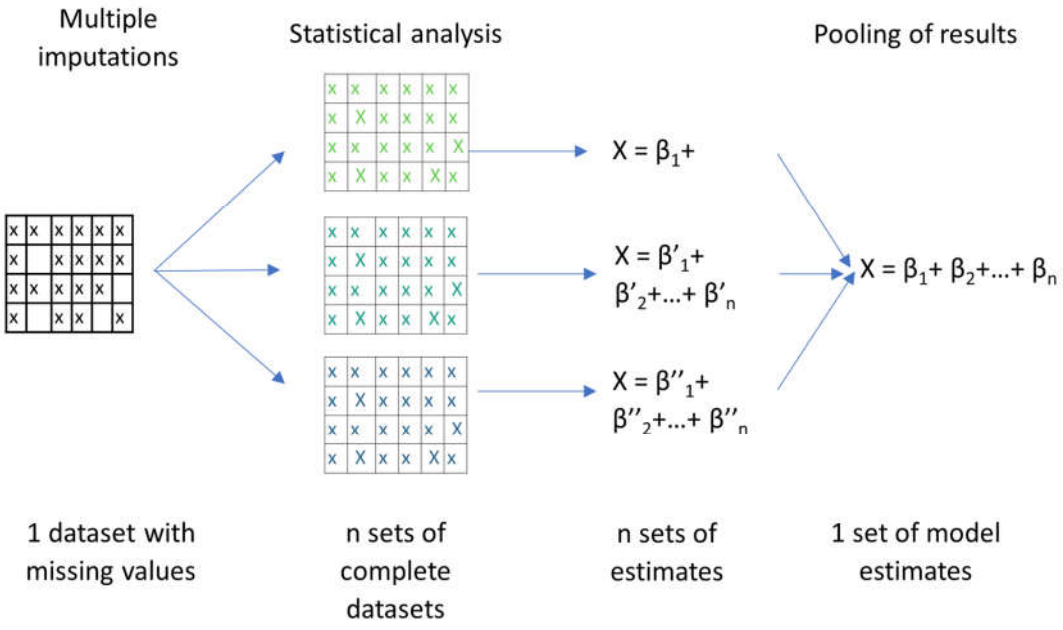

Figure S5. Multiple imputation process

D. Correlation analysis

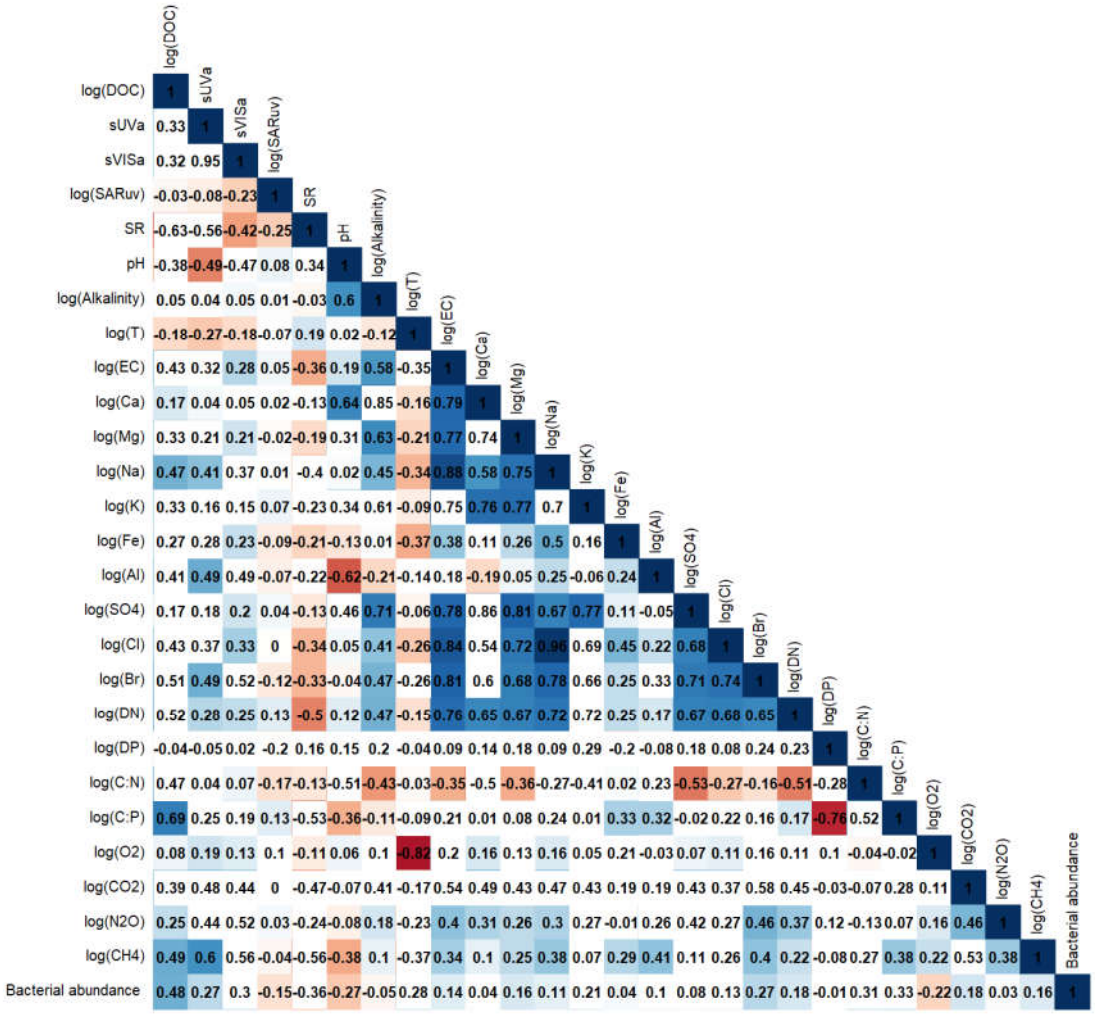

Figure S6 - Correlation plot for the 27 covariates. Coloured squares indicate significant correlation coefficients (p-value < 0.05)- Shades of blue and red indicate positive and negative correlations, respectively.

E. Lasso regression coefficients and model residuals

Table S2 - Lasso regression estimates (pooled for n > 25)

Parameter	Log(RR)	Log(RRn)	Log(BdgT)
Intercept	-0,0026	0,0031	-0.0021
Log(DOC)		-0,12	0,29
SUVA	-0.16	-0.15	
SARuv	-0.038	-0.035	-0.088
pH			0,38
Log(EC)		-0,17	
Log(Fe)	-0,063	-0,078	-0,053
Log(DP)	-0,13	-0,13	0,019
Log(C:N)	0,34		-0,013
Log(O_2)			0,10
Log(CO_2)		-0,047	
Log(N_2O)			0,011
Cells	0.073		-0.17

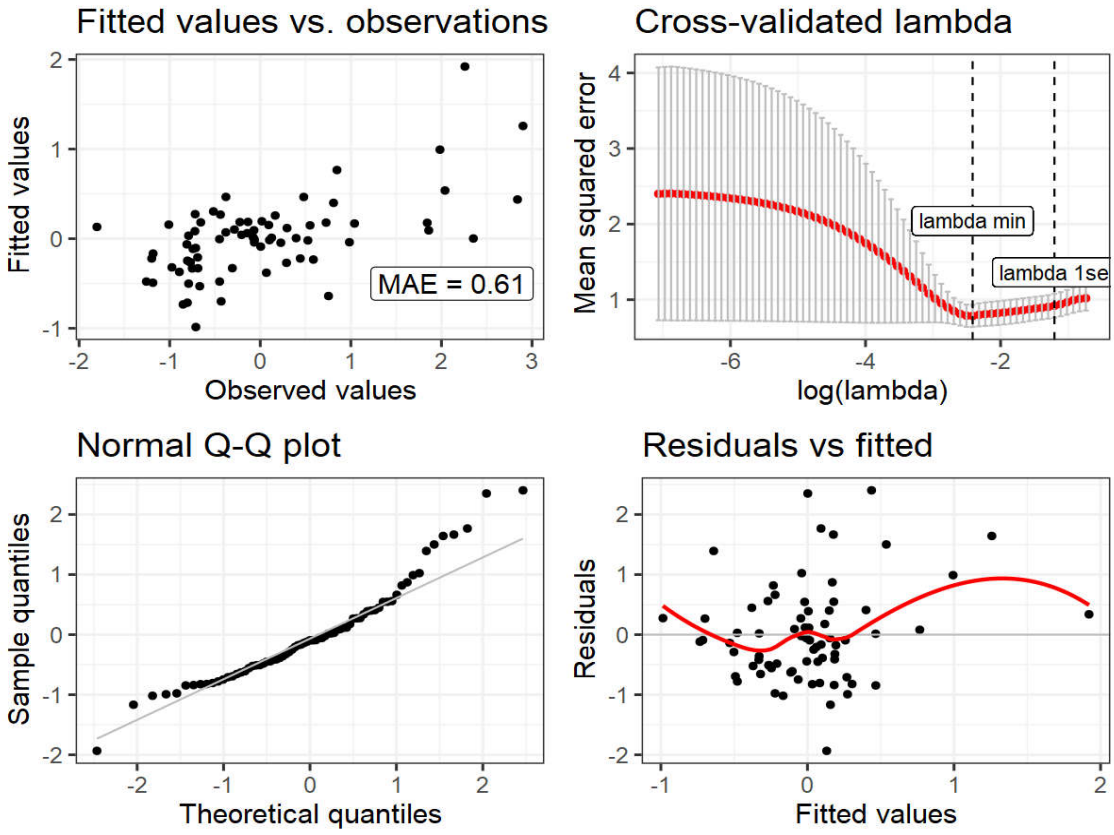

Figure S7 – Residual plots for lasso regression with log(RR) as response variable

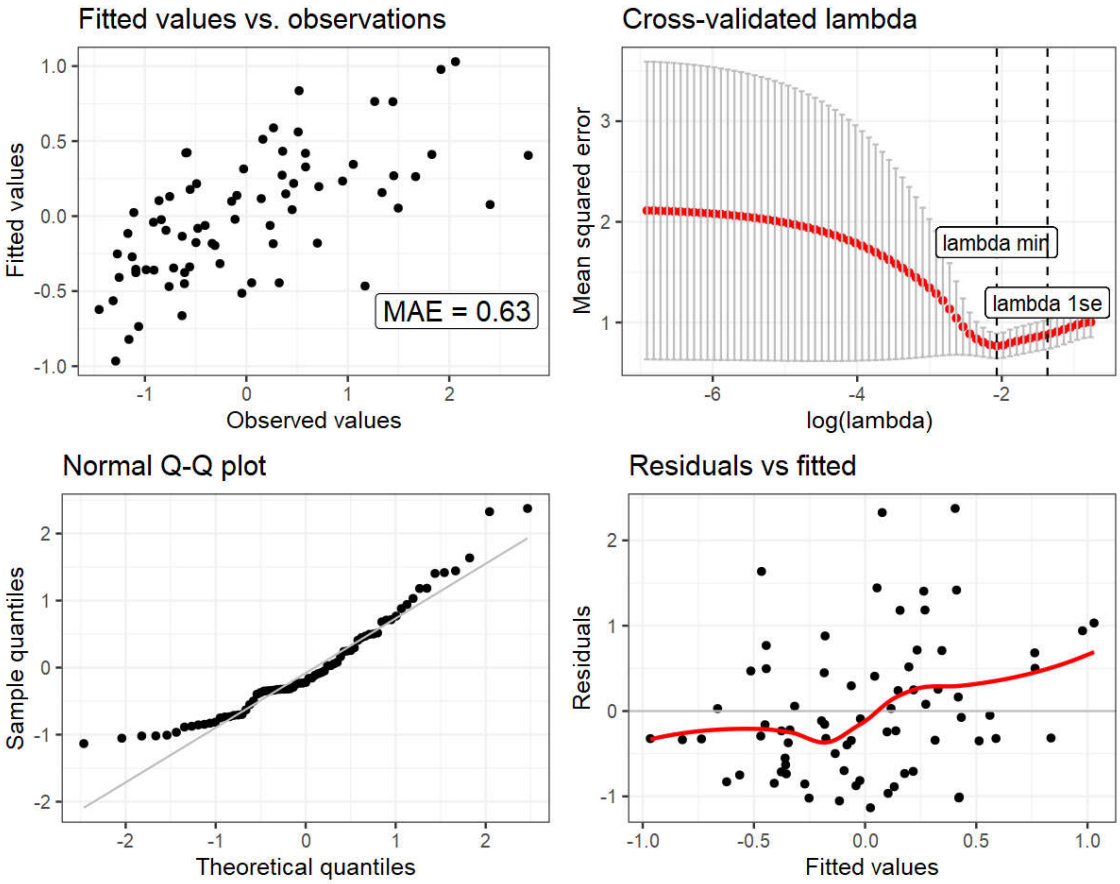

Figure S8 - Residual plots for lasso regression with log(RRn) as response variable

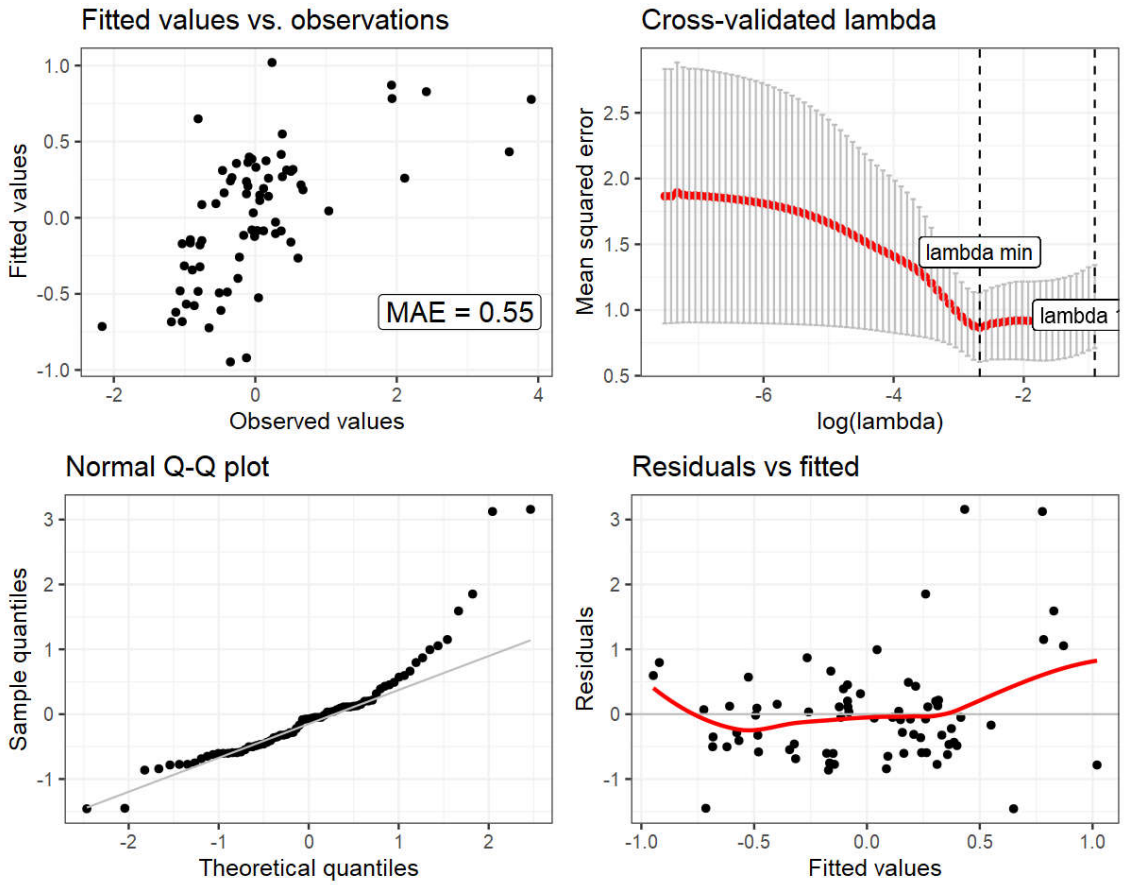

Figure S9 - Residual plots for lasso regression with log(BdgT) as response variable

F. Multiple Linear regression

Table S3 - Linear model estimates and p-values.

	RR		RRn		BdgT	
parameter	estimate	p-value	estimate	p.value	estimate	p.value
Intercept	-0.070	0.94	0.0097	0.091	-010	0.92
Log(DOC)			-0.18	0.086	0.57	0.00020
SUVA	-0.26	0.015	-0.22	0,048		
SARuv	-0.16	0.13	-0.23	0.020	-0.23	0.033
pH					0.41	0.00076
Log(EC)			-0.12	0.34		
Log(Fe)	-0.16	0.12	-0.23	0.019	-0.023	0.052
Log(DP)	-0.26	0.017	-0.34	0.00092	-0.025	0.83
Log(C:N)	0.34	0.0021			-0.18	0.19
Log(O_2)					0.17	0.12
Log(CO_2)			-0.13	0.30		
Log(N_2O)					-0.0061	0.96
Cells	0.17[7]	0.11			-0.31	0.011

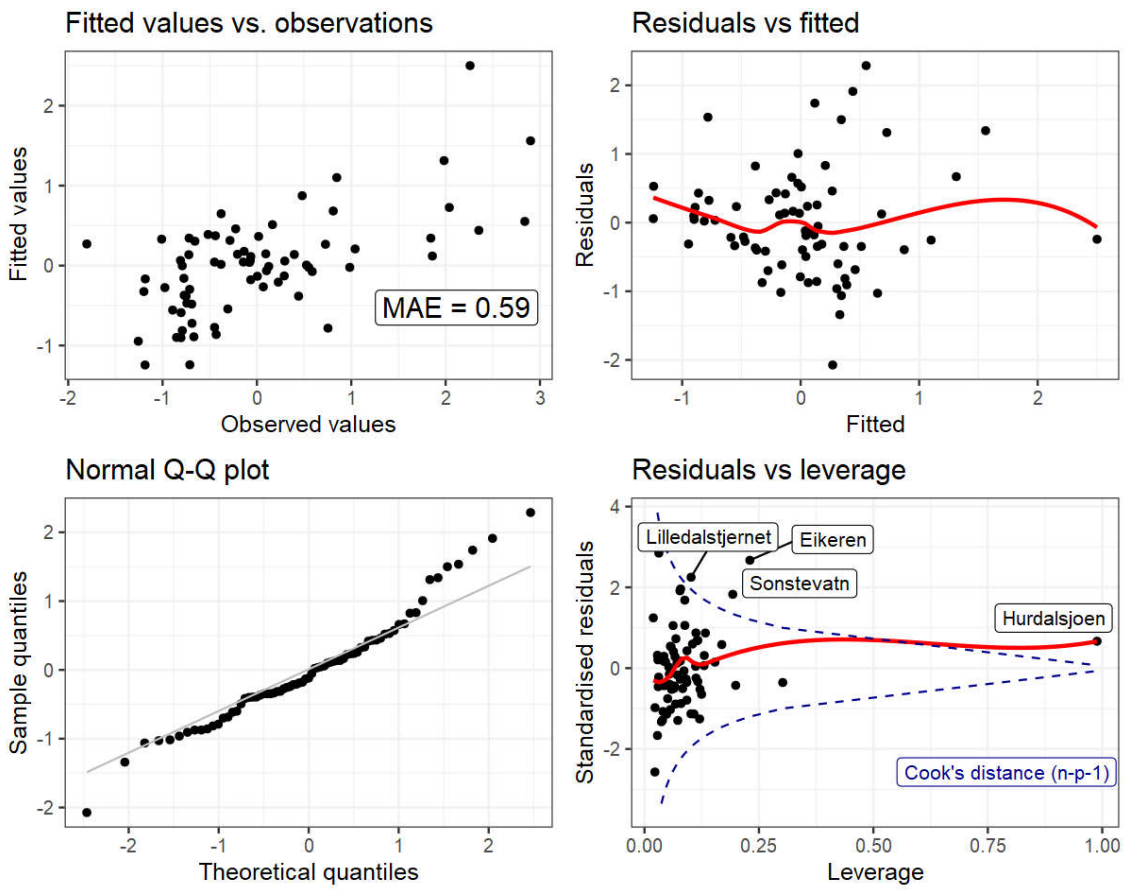

Figure S10 - Residual plots for linear model with log(RR) as response variable.

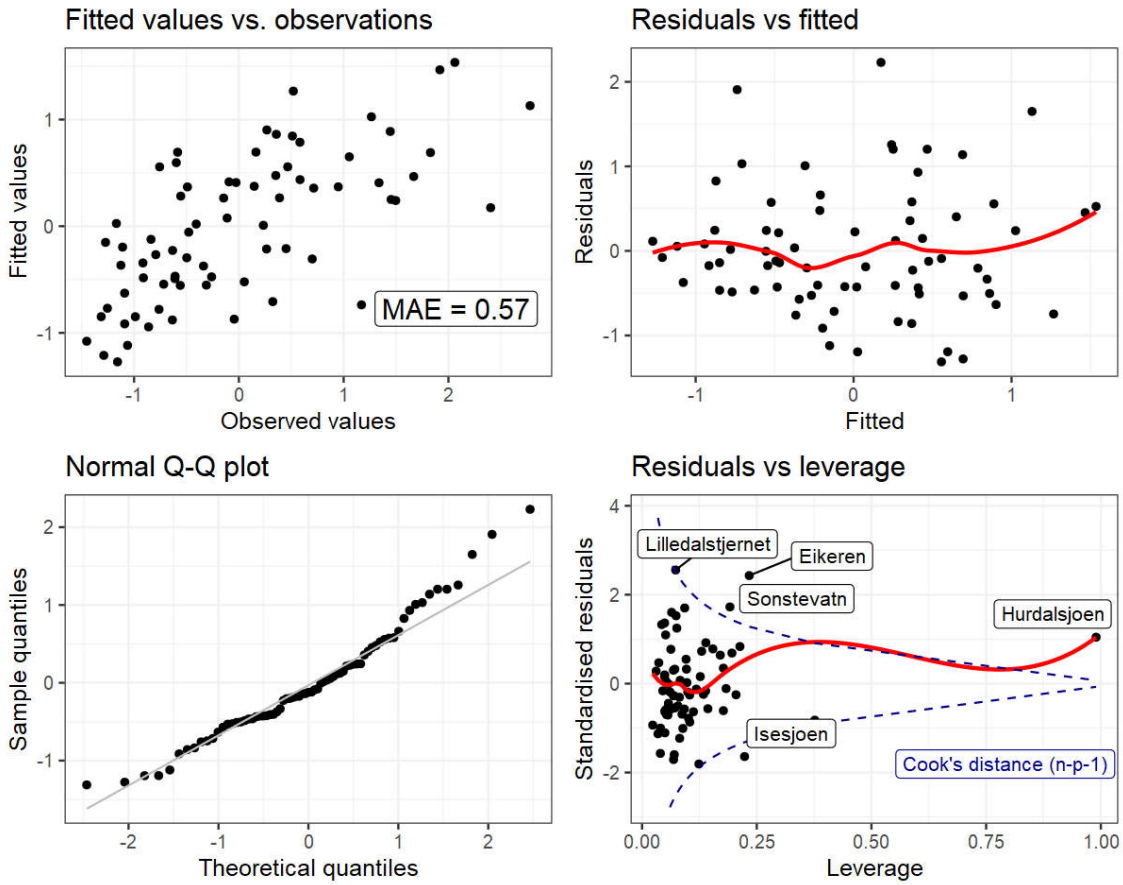

Figure S11 - Residual plots for linear model with log(RRn) as response variable.

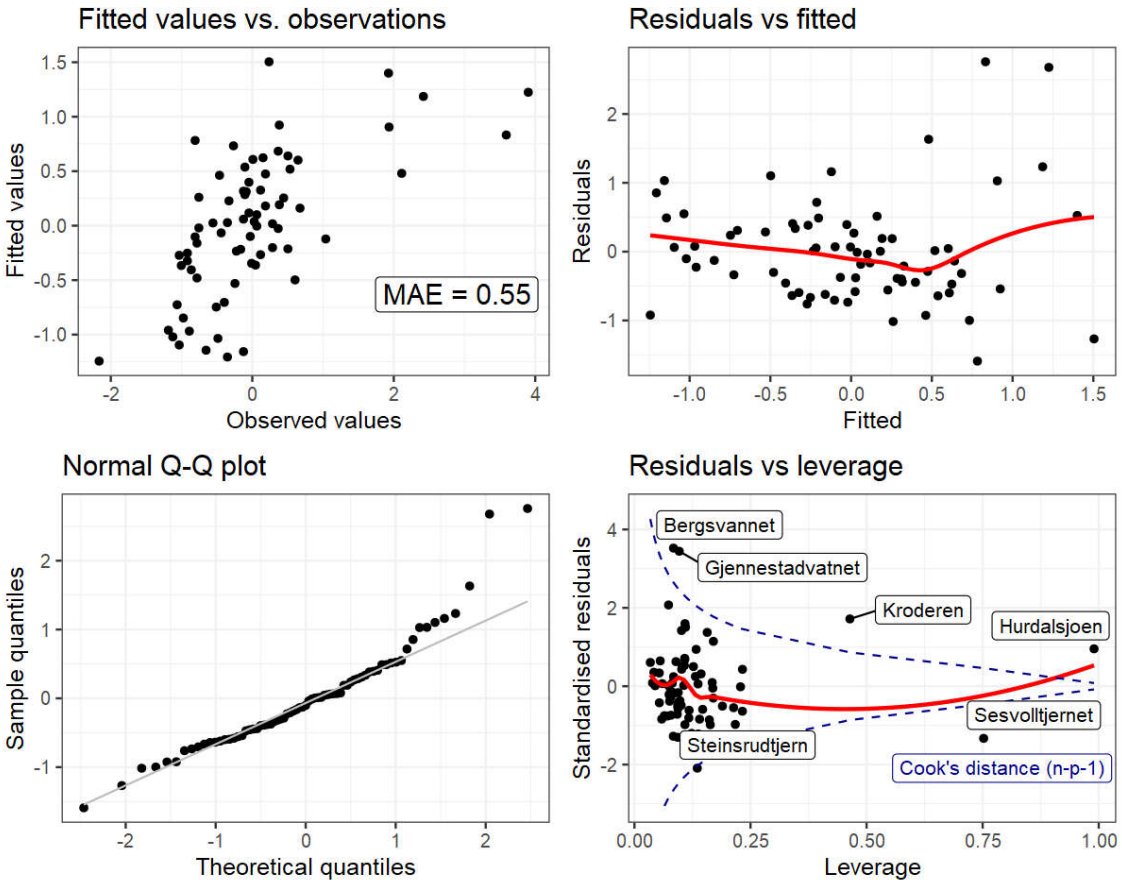

Figure S12 - Residual plots for linear model with log(BdgT) as response variable.