

Supplement of

Trends in soil solution dissolved organic carbon (DOC) concentrations across European forests

Marta Camino-Serrano et al.

Correspondence to: Marta Camino-Serrano (marta.caminoserrano@uantwerpen.be)

The copyright of individual parts of the supplement might differ from the CC-BY 3.0 licence.

ICP Forests Level II plots data used for the trend analysis

Table S1. List of ICP Forests Level II plots used for the trend analysis and their dominant forest species and resulting trend calculated using the Seasonal Mann-Kendall test (NS; nonsignificant, P: positive, N: negative). Rows in green correspond to the plots where at least one time series has been used for the individual trend analysis after filtering out the breakpoints. Rows in red correspond to the plots with measurements of DOC in soil solution that have not been used for the individual trend analysis because there was not enough data (Lack data) or breakpoints were detected (BP). Collector types are tension lysimeters (TL) or zero-tension lysimeters (ZTL).

Description of the statistical methods

1) Overall trend analysis at European scale

Linear mixed-effects models (LMM) were used to detect the temporal trends in soil solution DOC concentrations at the European scale. For these models, the complete ICP Forests Level II dataset was used. Because the dependent variable (DOC concentration) was usually not normally distributed, it was log-transformed to improve normality. Different models were built per depth and per collector type (tension or zero-tension lysimeters). For each model, the variable describing the temporal effect was the year, centered on the year 2000 (year-2000), which was considered as fixed effect. Also, month (1-12) was considered as fixed effect to account for seasonality. Two random factors describing the country (*ctryint*) and plot (*plotint*) effects and one random coefficient accounting for the between plot variation of the temporal effect (*plotslp*) were considered in each LMM (Equation 1). The LMMs were further adjusted by stratification of data according to forest type in order to investigate possible differences in DOC trends between broadleaved and coniferous forests. The models were built following [Jonard et al. \(2015\)](#page-27-0).

$$
logDOC = [a + month + ctry_{int} (0, \sigma_{ci}^2) + plot_{int}(0, \sigma_{pi}^2)] + [b + plot_{slp}(0, \sigma_{ps}^2)] \cdot
$$

(year – 2000) + ε (0, σ^2) (1)

where σ_{ci}^2 , σ_{pi}^2 , σ_{ps}^2 and σ^2 are the variances of the random factors 'country' and 'plot', of the random coefficient 'plot' and of the residual term (ε) , respectively.

2) Trend analysis of individual time series

Temporal changes in terrestrial ecosystems can either be monotonic changes, or discontinuous with abrupt changes resulting in breakpoints [\(de Jong et al., 2013\)](#page-27-1). Monotonicity of time series is generally assumed when analyzing DOC data for temporal trends [\(Filella and Rodriguez-Murillo, 2014\)](#page-27-2). However, it is rarely statistically tested and, thus, potential abrupt changes in the time series may be overlooked. This issue becomes important in temporal trend analysis since a breakpoint may cause changes in the direction of the trend and could lead us, for example, to classify a time series as constant, when in reality we may have averaged out separate periods with significant changes [\(de Jong et al., 2013\)](#page-27-1). On the other hand, breakpoints may erroneously induce the detection of a significant trend in long-term time series due to artifacts.

For these reasons, we focused on the investigation of the potential long-term trends in soil solution DOC at European forests that show monotonicity. Therefore, DOC time series were first analyzed using the Breaks For Additive Seasonal and Trend (BFAST) algorithm to detect the presence of breakpoints [\(Verbesselt et al., 2010\)](#page-27-3). When a breakpoint was detected in a time series, there were two possibilities: first, one of the segments (before or after the detected breakpoint) was longer than 9 years, and, in this case, only the longest segment was used for the subsequent analysis of monotonic trends; second, the breakpoint split the time series in two segments shorter than 9 years and then the time series was not used for the analysis of monotonic trends. We used a length threshold of 9 years, which is the minimum time series length recommended for long-term trend analysis [\(Libiseller and Grimvall, 2002;](#page-27-4) [Waldner et al., 2014\)](#page-27-5). In total, 258 time series from 97 plots were selected for analysis of monotonic trends (Table S2). No clear pattern could be observed in the distribution of time series of DOC with breakpoints, which appeared to occur randomly across the study plots (Figs. 3 and 4).

Monotonic trend analyses were carried out using the Seasonal Mann Kendall (SMK) test for monthly DOC concentrations [\(Hirsch et al., 1982;](#page-27-6) [Marchetto et al., 2013\)](#page-27-7). Partial Mann Kendall (PMK) test was also used to test the influence of monthly precipitation as a covariable, i.e., to test if the trend detection might be due to a DOC dilution/concentration effect [\(Libiseller and Grimvall, 2002\)](#page-27-4). For the SMK and PMK tests, the trend slopes were estimated following [Sen \(1968\)](#page-27-8), as the median of all the slopes determined by all pairs of sample points. The SMK and PMK account for seasonality of the time series by computing the test on each of the seasons (in our case months) separately. The resulting slopes were also tested against the slopes calculated by BFAST. Finally, the individual slopes calculated according to [Sen \(1968\)](#page-27-8) for each time series using the SMK or PMK method were standardized by dividing them by the median DOC concentration over the sampling period to avoid the influence of the magnitude of DOC concentration in the between-site comparison. The resulting standardized slopes (relative slopes) were used for the subsequent statistical analysis.

For this study, five depth intervals were considered: the organic layer (0 cm), topsoil (0-20 cm), intermediate (20-40 cm), subsoil (40-80 cm) and deep subsoil (> 80 cm). The slopes of each time series were then aggregated to a unique slope per depth interval in each plot (hereafter called "plot-soil depth combination") and classified by the direction of the trend as significantly positive (P, $p < 0.05$), significantly negative (N, $p < 0.05$) and not significant (NS, $p \ge 0.05$). When there was more than one collector per depth class, the median of the slopes was used when the direction of the trend (P, N or NS) was similar. When the different trends at the same plot-soil depth combination were either P and NS, or N and NS, it was marked as "Weighted positive" and "Weighted negative". The five plot-soil depth combinations for which the calculated slopes showed opposite trend directions were discarded. All aggregated trend slopes came from time series measured using the same collector type. After aggregation per plot-depth combinations, 191 trend slopes from 97 plots were available for analysis (Table S2).

Trends for soil solution parameters (NO₃, Ca, Mg, NH₄⁺, SO₄⁻², total dissolved Al, total dissolved Fe, pH, electrical conductivity), precipitation and temperature were calculated using the same methodology as for DOC: individual time series were analyzed using the SMK test and the relative slopes were calculated and aggregated to plot-soil depth combinations.

Finally, we performed a multivariate statistical analysis to investigate the main factors explaining differences in DOC trends among the selected plots. We applied Structural Equation Models (SEM) to test whether deposition variables had an effect (direct, indirect or total) on the relative trends slopes of DOC through different pathways [\(Grace et al., 2010\)](#page-27-9). For the SEMs, we assumed that there is no effect of soil depth on the DOC trends (see next section in Supplementary Material). We applied three SEM models: 1) for all the slopes in DOC, 2) only for the forests with low or medium total N deposition, and, 3) only for the forests with high total N deposition. For each case, we searched for the most parsimonious adequate model using the Akaike information criterion (AIC) and R^2 . The significance level (*p* value) of the total, direct and indirect effects were calculated using the bootstrap (with 1200 repetitions) technique [\(Davison et al., 1986\)](#page-27-10). Dependent variables were log-transformed to improve normality of the continuous variables and then standardized before performing the SEMs. All the statistical analyses were performed in R software version 3.1.2 [\(R Core Team,](#page-27-11) [2014\)](#page-27-11) using the "rkt" [\(Marchetto et al., 2013\)](#page-27-7), "bfast01" [\(de Jong et al., 2013\)](#page-27-1) and "sem" [\(Fox et al., 2013\)](#page-27-12) packages, except for the LMMs that were performed using SAS 9.3 (SAS institute, Inc., Cary, NC, USA).

Table S2. Summary of number of time series used in the study

	Entire dataset	Without breakpoints
All time series	1480 (173 plots)	--

Depth effect on the individual trends in soil solution DOC

Trends in soil solution from different soil depth intervals were mixed for the Pearson's chisquared test performed for Fig. 5 and the Structural Equation Models (SEM) (Fig. 6), as the number of cases available for each depth are insufficient to compute the statistics if we separate per soil depth interval. To check if the trends calculated at different depths were actually independent from the soil depth interval, we performed a Pearson's chi-squared test and found that the differences in trends among soil depth intervals were not statistically significant $\chi^2(8, N = 174) = 10.94$, $p = 0.21$) (Fig. S1). Therefore, we assumed that there is no difference in trends among soil depth layers and performed the subsequent statistical analysis mixing the trends from different soil depths.

Figure S1. Percentage of non-significant, positive and negative trends per soil depth interval (O: organic layer, M02: mineral soil 0-20 cm, M24: mineral soil 20-40 cm, M48: mineral soil 40-80 cm, M8: mineral soil > 80 cm).

However, a real difference in DOC trends between soil depths may be obscured by the fact that datasets differ between different depths (not all the sites have DOC time series that could be analyzed for trends at all the soil depth intervals) and thus, we cannot rule out that there exists a difference in trends per soil depth. Although the number of sites with DOC trends analyzed at more than three soil depths (including the organic layer) is not enough to apply the same statistics for this subset, we visually compared the 11 sites with this information available and found that, at first sight, it was confirmed that there is no a real difference in trends between soil depth intervals (Fig. S2).

Figure S2. Direction of the trend (non-significant, positive and negative) per soil depth interval (O: organic layer, M02: mineral soil 0-20 cm, M24: mineral soil 20-40 cm, M48: mineral soil 40-80 cm, M8: mineral soil > 80 cm) for the 11 plots with DOC measured at least at 3 soil depth intervals including the organic layer. The size of the circle is proportional to the magnitude of the trend slope.

$\bf{{Structural~equation~model~with~trends~in~}{SO_4}^2 \cdot \bf{and~} NO_3 \cdot \bf{deposition}}$

The same structural equation models (SEM) represented in Fig. 6 were performed using the trends in SO_4^2 and NO_3^- deposition (% yr⁻¹) instead of the mean values of SO_4^2 and $NO_3^$ throughfall deposition (kg ha⁻¹ yr⁻¹). The SEMs for all the cases and for cases with low and medium inorganic N deposition are shown in Fig. S3.

Figure S3. Diagram of the structural equation model (SEM) that best explains the maximum variance of the resulting trends of DOC concentrations in soil solution for: A) all the cases and B) cases with low or medium inorganic N deposition, with trends in SO_4^2 and $NO_3^$ deposition $(\%$ yr⁻¹) with direct effects and indirect effects through effects on mean annual stem volume increment (growth) in m^3 ha⁻¹ yr⁻¹). P-values of the significance of the corresponding effect between brackets. Green arrows indicate positive effects and red arrows indicate negative effects.

Comparison of methods of individual trend analysis

We applied the BFAST analysis to select the monotonic time series in order to assure that the overall detected trends were not influenced by breakpoints in the time series. Time series with breakpoints represented more than 50% of the total time series aggregated by soil depth interval (245 out of 436). In total, 191 plot-soil depth combinations from 97 plots were analyzed after filtering out the time series showing breakpoints and 94% of the analyzed plotdepth combinations showed consistent trends among replicates collected at the same depth. In contrast, when also considering the time series with breakpoints, the trends calculated for plot-depth combinations agreed only in 75% of the cases implying that the proportion of contradictory trends within plot-depth combinations increased from 6% in the dataset without breakpoints to 25% in the entire dataset (Fig. S4). For both datasets, the majority of the trends were not statistically significant (44% and 41%, for the dataset with and without breakpoints, respectively). In other words, filtering the time series for breakpoints reduced the within-plot variability, while most of the plots showed similar aggregated trends per plot-depth combinations. For this reason, the results discussed in this paper correspond only to the trends of monotonic (breakpoint filtered) time series of soil solution DOC concentrations.

Figure S4. Percentage of plot-soil depth combinations for which negative (N), non-significant (NS), positive (P), negative and non-significant (Weight_N) and positive and non-significant (Weight_P) trends of DOC concentrations were found using SMK (seasonal Mann-Kendall)

tests when 1) all the 436 time series were used, 2) only 191 time series without breakpoints (detected using the BFAST (*Breaks For Additive Seasonal and Trend*) analysis) were used.

There was a good agreement between results using the three methods: BFAST, SMK, and PMK. The direction and significance of the trend agreed for 84.5% of the time series analyzed. For the majority of the remaining time series for which the trends did not agree, BFAST did not detect a trend when SMK and PMK did, thus, the latter two methods seemed more sensitive for trend detection than BFAST. Trends computed with SMK and PMK agreed well. The direction of the trend for SMK and PMK only differed for the intermediate soil layer (20-40 cm), as a result of the two extra sites for which SMK tests were performed, but not the PMK, that showed a marked positive trend $(1.1 \text{ and } 2 \% \text{ yr}^{-1})$. However, when using exactly the same set of sites, the trend did not differ between the two methods.

Implications of using standardized DOC slopes versus absolute DOC slopes.

The standardized (relative) slopes of DOC concentrations were used for the study of the factors affecting the soil solution DOC trends (Fig. 5 and 6). The main reason for this decision was that using the median DOC concentration as a reference (as we did with the standardization) allowed us to determine whether the absolute trend in DOC was quantitatively large or not from an ecological perspective, because the absolute trend slope will be highly dependent on the initial DOC concentrations of the site.

The absolute trend slopes show the real magnitude and significance of the trend, but do not allow for comparison among sites or horizons. Since the aim of this study is to test whether there is a general DOC trend and to compare sites across Europe, we decided that using the relative slope was more consistent.

Moreover, due to limitations of the statistical analysis, we worked with time series per "plotsoil depth combinations", which means that different soil layers were mixed in the statistical analysis. Again, the standardization of the slopes of DOC concentrations allowed us to compare trends among different soil horizons by removing the effect of the decreasing soil solution DOC concentrations with soil depth. Otherwise, using the absolute trends would introduce a bias when we try to explain the DOC trends in relation with other parameters, because the trend slope would be highly dependent on the actual DOC concentrations, which, in turn, are very variable, not only among sites, but also among soil depths.

The influence of the DOC concentration levels was checked before deciding to use the standardized slopes (Fig. S5). It seemed that there was no relationship between the DOC trend slopes (relative and absolute) and the median DOC concentrations, with positive and negative trends occurring at both low and high DOC concentrations and, thus, we decided that using the standardized slopes will not hide any effect of the median DOC concentrations on the direction of the DOC trends.

This decision, however, has a drawback: the strength of the trend is clearly influenced by the DOC concentration levels. The fact that we used the standardized slope of DOC implied that it may be identical for two sites with very different mean DOC concentrations. DOC concentration decreases with depth and is lower in the deep mineral soil than in the upper mineral soil (Table S3) and by standardizing the slope, the magnitude of the trend was exaggerated in lower soil layers where both the absolute slope of DOC and the median DOC concentration are low (Table S3). This issue is well illustrated in Fig. S5, that shows how the highest standardized slopes are usually at low DOC concentrations (mostly in mineral soil layers), while the highest absolute slopes are at higher DOC concentrations (mostly in organic and upper soil layers).

In other words, in quantitative terms DOC trends are much higher in the organic layer than in the mineral soils but, in relative terms, DOC is increasing in the same proportion (Table S3). Because the aim of this study is to explain the high heterogeneity of DOC trends found across Europe, instead of the quantification of the trends at local scale, the relative trends were discussed throughout the manuscript. Consequently, our results should be interpreted with caution, keeping in mind that the relations between DOC trends and explaining factors are discussed only from a relative point of view.

Nevertheless, the statistical analyses (LMM, SMK, PMK and BFAST) were done on the absolute value and the resulting Sen's slopes were then standardized. Thus, the fact that trends are expressed in relative terms has consequences on the interpretation of the results, but has no influence on the statistical test itself (carried out on the absolute values of DOC), that is, on the significance and direction of the trends.

Table S3. Comparison of median relative trend slope (rslope in % yr-1) and absolute trend slope (abs slope in mg L^{-1} yr⁻¹) of DOC concentrations in soil solution and their interquartile range using the Seasonal Mann-Kendall test (SMK). (O: organic layer, M02: mineral soil 0- 20 cm, M24: mineral soil 20-40 cm, M48: mineral soil 40-80 cm, M8: mineral soil > 80 cm.)

Figure S5. A) Standardized trends (relative DOC slope) versus median DOC concentrations. B) Absolute trends (absolute slope DOC) versus median DOC concentrations. The different colors represent the different soil layers.

Information on the soil solution chemistry at the studied ICP Forests Level II plots

Table S4. Median soil solution DOC concentrations (mg L^{-1}), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S5. Median soil solution pH, 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S6. Median soil solution conductivity (μ S cm⁻¹), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S7. Median soil solution Ca (mg L^{-1}), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S8. Median soil solution Mg (mg L^{-1}), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

M24	0.44	0.25	0.927	644	0.00	0.00		
M48	0.76	0.49	3.725	227	0.55	0.35	0.91	221 34 I
M8	0.85	0.37	1.3425	84	0.00	0.00		

Table S9. Median soil solution $S-SO₄²$ (mg L⁻¹), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S10. Median soil solution N-NO₃ (mg L^{-1}), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the

entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

Table S11. Median soil solution Al (mg L^{-1}), 25% and 75% percentiles and number of observations (n) for the different forest types, soil depth intervals and collector types with the entire dataset (with breakpoints) and with the dataset without time series showing breakpoints (without breakpoints).

References

Davison, A. C., Hinkley, D. V., and Schechtman, E.: Efficient Bootstrap Simulation, Biometrika, 73, 555-566, 1986.

de Jong, R., Verbesselt, J., Zeileis, A., and Schaepman, M. E.: Shifts in Global Vegetation Activity Trends, Remote Sens-Basel, 5, 1117-1133, 2013.

Filella, M. and Rodriguez-Murillo, J. C.: Long-term Trends of Organic Carbon Concentrations in Freshwaters: Strengths and Weaknesses of Existing Evidence, Water-Sui, 6, 1360-1418, 2014.

Fox, J., Nie, Z., and Byrnes, J.: sem: Structural Equation Models., 2013.

Grace, J. B., Anderson, T. M., Olff, H., and Scheiner, S. M.: On the specification of structural equation models for ecological systems, Ecol Monogr, 80, 67-87, 2010.

Hirsch, R. M., Slack, J. R., and Smith, R. A.: Techniques of Trend Analysis for Monthly Water-Quality Data, Water Resour Res, 18, 107-121, 1982.

Jonard, M., Furst, A., Verstraeten, A., Thimonier, A., Timmermann, V., Potocic, N., Waldner, P., Benham, S., Hansen, K., Merila, P., Ponette, Q., De La Cruz, A. C., Roskams, P., Nicolas, M., Croise, L., Ingerslev, M., Matteucci, G., Decinti, B., Bascietto, M., and Rautio, P.: Tree mineral nutrition is deteriorating in Europe, Global Change Biol, 21, 418- 430, 2015.

Libiseller, C. and Grimvall, A.: Performance of partial Mann-Kendall tests for trend detection in the presence of covariates, Environmetrics, 13, 71-84, 2002.

Marchetto, A., Rogora, M., and Arisci, S.: Trend analysis of atmospheric deposition data: A comparison of statistical approaches, Atmos Environ, 64, 95-102, 2013.

R Core Team: R: A language and environment for statistical computing., R Foundation for Statistical Computing, Vienna, Austria, 2014.

Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, Journal of the American Statistical Association, 63, 1379-1389, 1968.

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D.: Detecting trend and seasonal changes in satellite image time series, Remote Sens Environ, 114, 106-115, 2010.

Waldner, P., Marchetto, A., Thimonier, A., Schmitt, M., Rogora, M., Granke, O., Mues, V., Hansen, K., Karlsson, G. P., Zlindra, D., Clarke, N., Verstraeten, A., Lazdins, A., Schimming, C., Iacoban, C., Lindroos, A. J., Vanguelova, E., Benham, S., Meesenburg, H., Nicolas, M., Kowalska, A., Apuhtin, V., Napa, U., Lachmanova, Z., Kristoefel, F., Bleeker, A., Ingerslev, M., Vesterdal, L., Molina, J., Fischer, U., Seidling, W., Jonard, M., O'Dea, P., Johnson, J., Fischer, R., and Lorenz, M.: Detection of temporal trends in atmospheric deposition of inorganic nitrogen and sulphate to forests in Europe, Atmos Environ, 95, 363- 374, 2014.