Electronic supplementary information (ESI) for the article

“AbspectroscoPY, a Python toolbox for absorbance-based sensor data in water quality monitoring”

for Environmental Science: Water Research & Technology.

Table A1. Follow-up discussion of Table 3, indicating how to take in account the differences in time of the three spectro::lyser units in surface water (SW), rapid sand filtrate (RSF) and ultrafilter permeate (UF) during the periods of Daylight Saving Time (DST, 03/10/2018) and Standard Time (ST, 27/11/2018).

		Time [hh:mm:ss]				
Sample	Period	$t_{s::can}$	t_{CEST}	Δt_{tot}	Δt_{SW-RSF}	Δt_{SW-UF}
SW	DST	07:16:00	08:16:00	- 01:00:00	- 01:00:00	- 01:00:00
	ST	08:06:00	08:06:00	00:00:00	- 01:00:00	- 01:00:00
RSF	DST	09:23:00	09:23:00	00:00:00		
	ST	10:42:00	09:42:00	+ 01:00:00		
UF	DST	09:21:00	09:21:00	00:00:00		
	ST	10:39:00	09:39:00	+ 01:00:00		

$t_{s::can}$ = timestamp shifted one hour forward during ST for the sensors after RSF and UF and by Δts::can for all of them

Δt_{tot} = total time difference ($t_{s::can}$ - t_{CEST})

Δt_{SW-RSF} = Δt_{tot} difference between the sensor in SW and the one after RSF ($\Delta t_{tot\ SW}$ - $\Delta t_{tot\ RSF}$)

Δt_{SW-UF} = Δt_{tot} difference between the sensor in SW and the one after UF ($\Delta t_{tot\ SW}$ - $\Delta t_{tot\ UF}$)

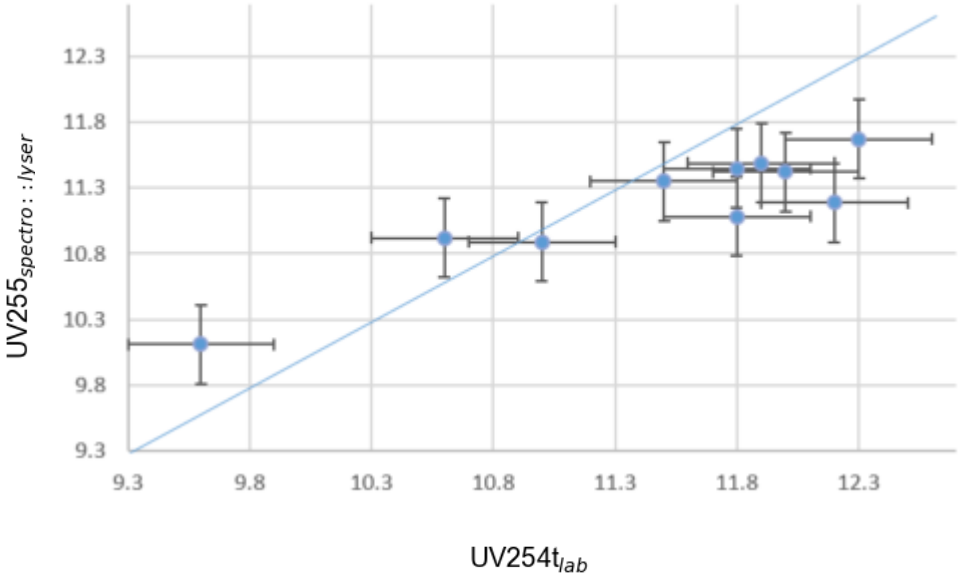

Figure A1. Scatterplot of the UV255 data (absorbance per meter) from the spectro::lyser versus the unfiltered UV254 data (absorbance per meter) from grab samples for surface water (SW), considering an instrumental error of the laboratory analyses of 0.3 absorbance per meter.

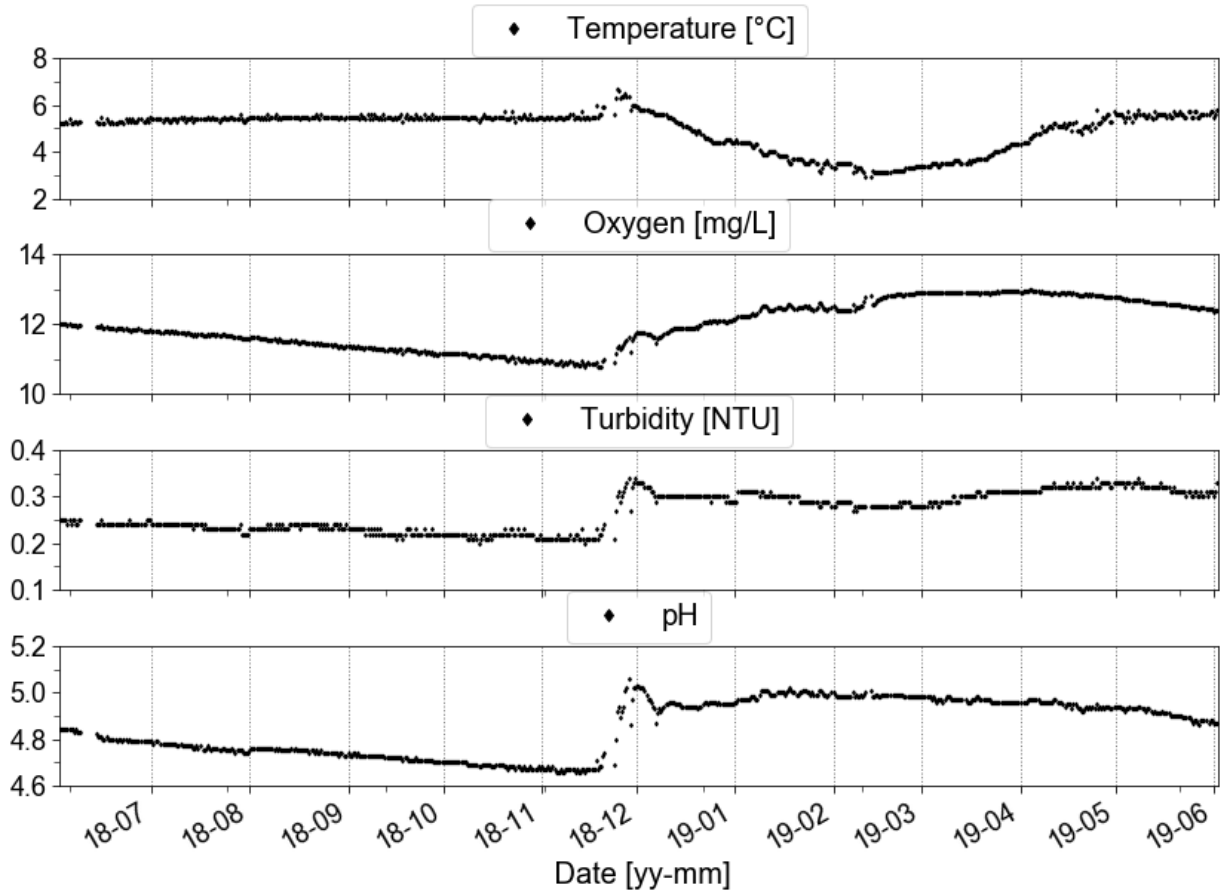

Figure A2. Sensor measurements of temperature, oxygen, turbidity and pH in surface water (SW) during the Autumn lake circulation event in Lake Neden (period P5 defined in Figure 2).

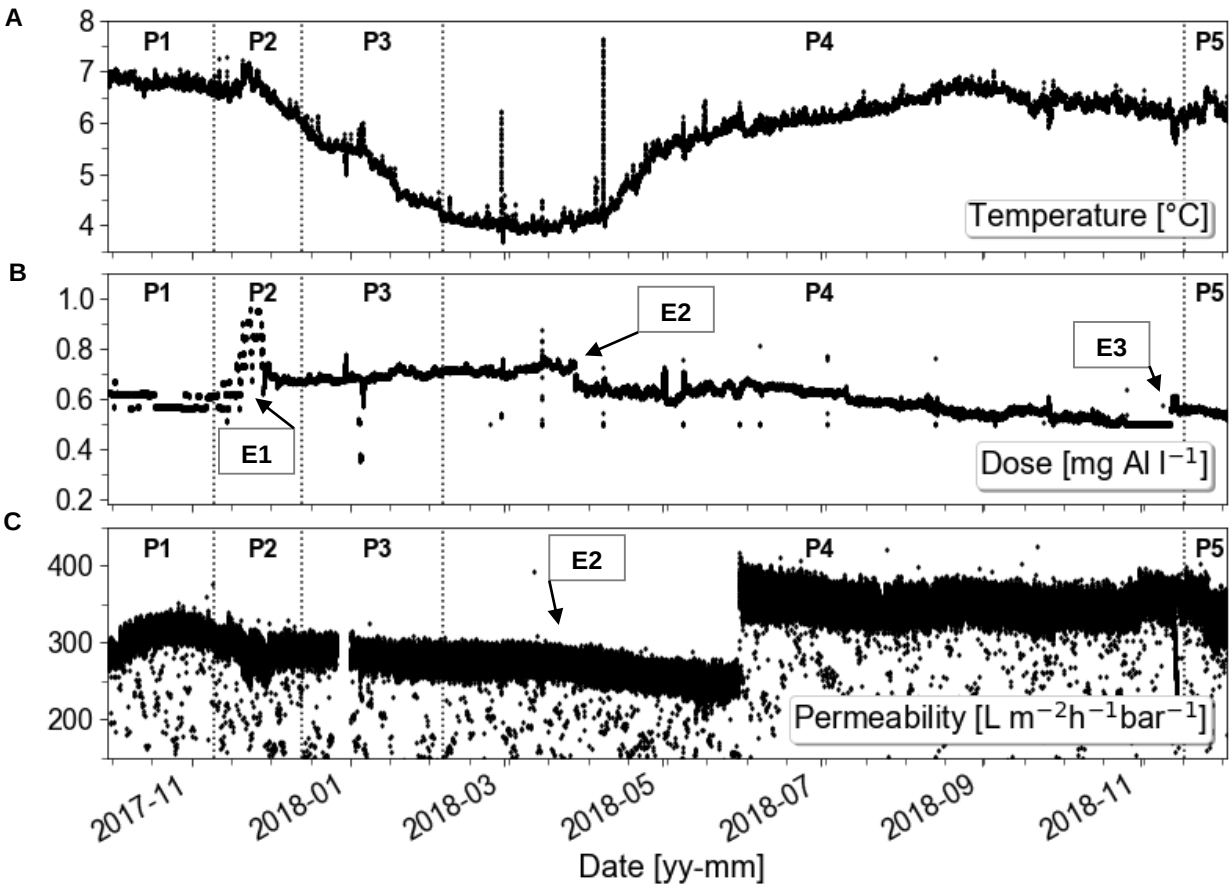

Figure A3. Sensor measurements from the ultrafilter (UF) over 13 months in the DWTP: (A) temperature in the feed water; (B) optimal dose as calculated by the coagulation controls system; (C) UF membrane permeability. Compare the three events with Figure 2.

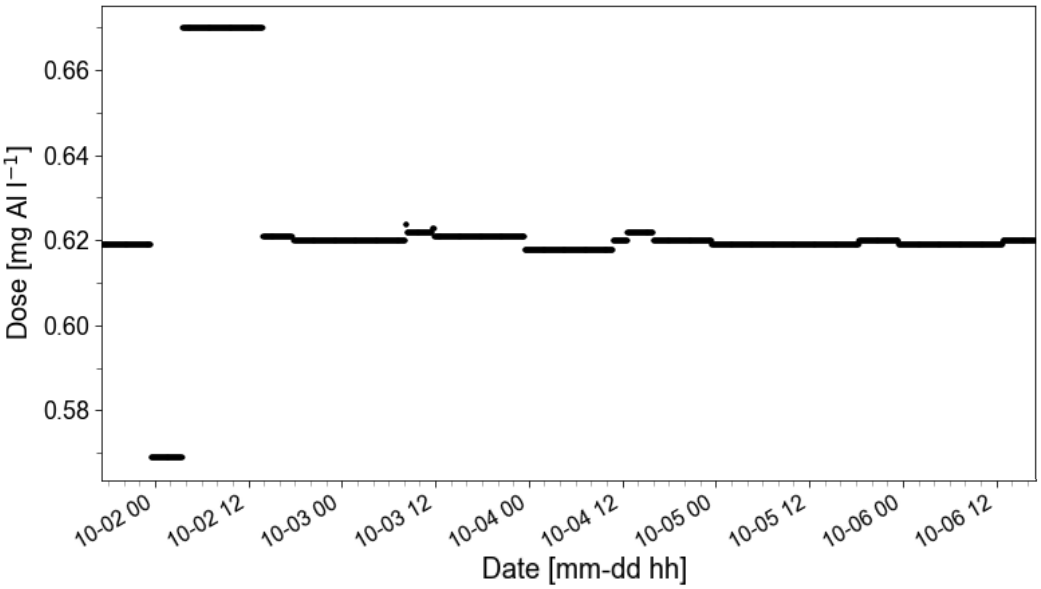

Figure A4. Optimal coagulant dose as calculated by the coagulation control system in figure A3 for comparison with Figure 3 (October 2017).

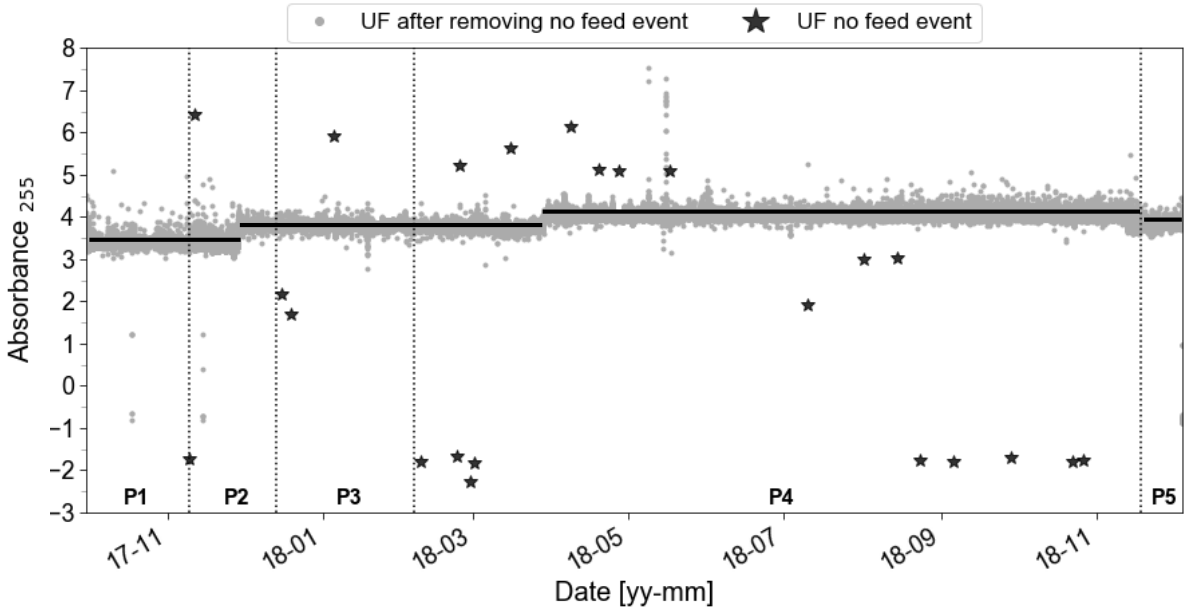

Figure A5. Preprocessed spectro::lyser time series showing UV absorbance at 255 nm (absorbance per meter) in the ultrafilter permeate (UF) (September 2017 - December 2018) sampled every 3 min. Events caused by absence of feed water to the UF (stars, compare with Figure 4) have been removed using the function *outlier_id_drop*. The horizontal lines were superimposed to the UV absorbance time series to compare it with Figure A7. Compare the three events with Figure 2.

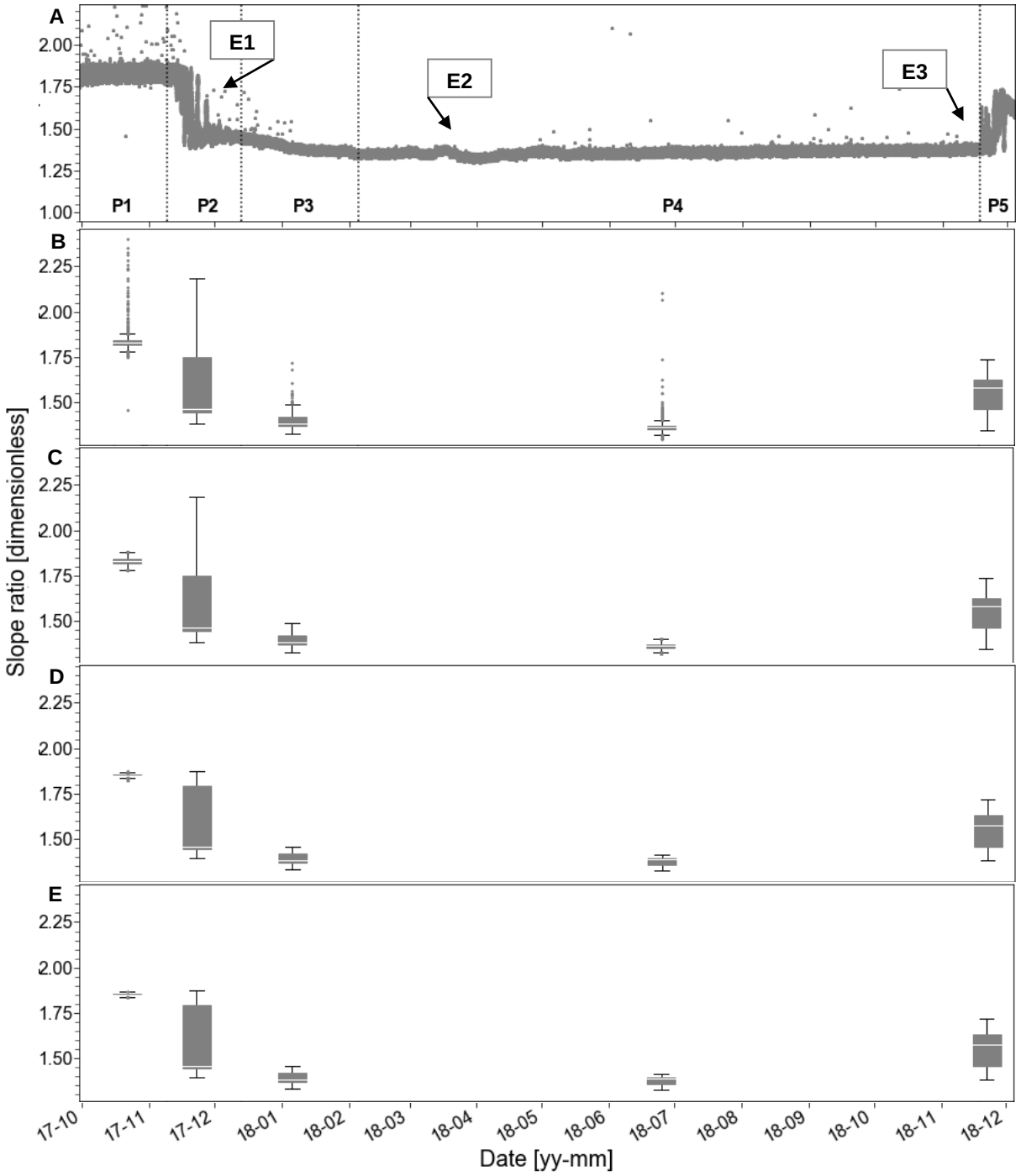

Figure A6. Slope ratio results obtained applying the function *abs_slope_ratio* to the surface water attenuation time series before baseline correction and median smoothing (A) and used to compute boxplots for outlier identification and removal using the function *outlier_id_drop_iqr* within the AbspectroscoPY toolbox: (B) before baseline correction and median smoothing before outlier removal; (C) before baseline correction and median smoothing after outlier removal; (D) after baseline correction and median smoothing before outlier removal; (E) after baseline correction and median smoothing after outlier removal. Compare the three events with Figure 2.

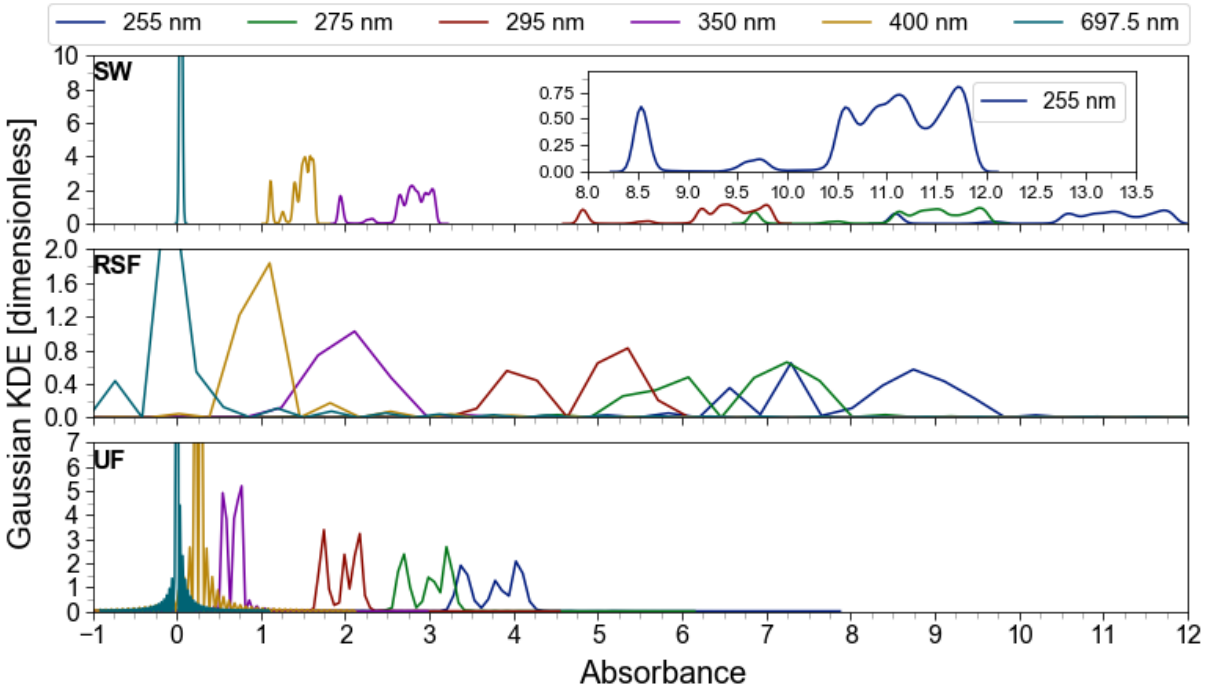

Figure A7. The distribution of spectro::lyser absorbance measurements (absorbance per meter) at specific wavelengths (255 nm, 275 nm, 295 nm, 350 nm, 400 nm, 697.5 nm) in surface water (SW), rapid sand filtrate (RSF) and ultrafilter permeate (UF) in September 2017 - December 2018. Data were plotted using the function *kdeplot* in the AbspectroscoPY toolbox. The inset shows a closeup of SW at 255 nm.

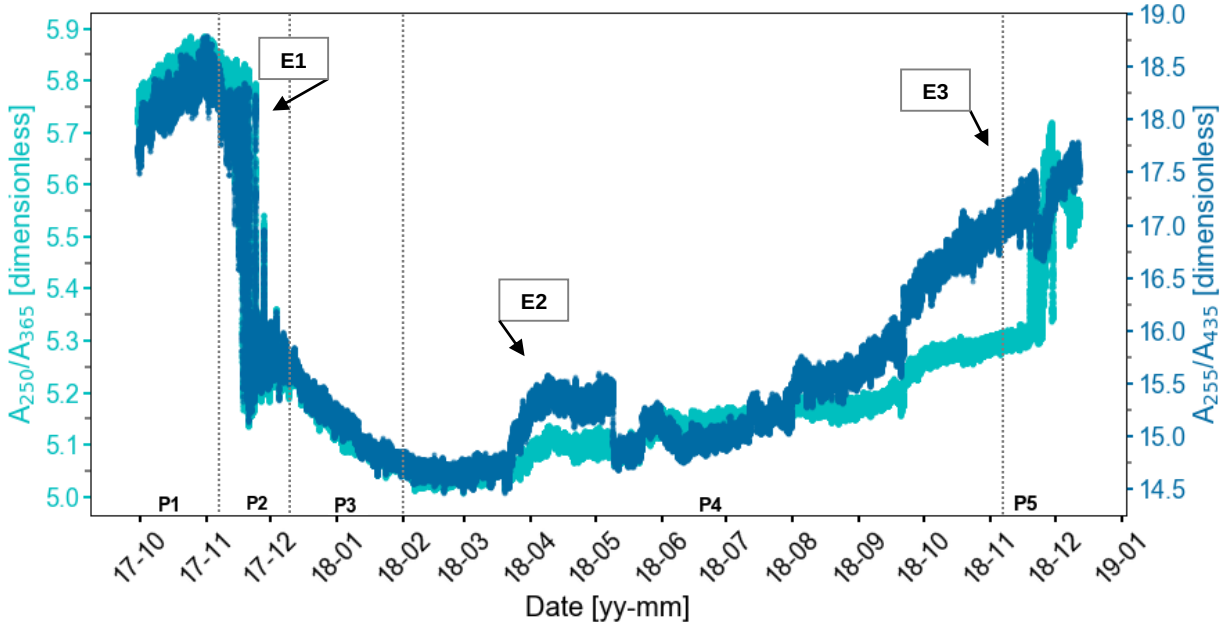

Figure A8. The ratio of the spectro::lyser absorbance data at 250 nm to 365 nm (cyan) and at 255 nm to 435 nm (blue) in surface water (SW). Compare the three events with Figure 2.

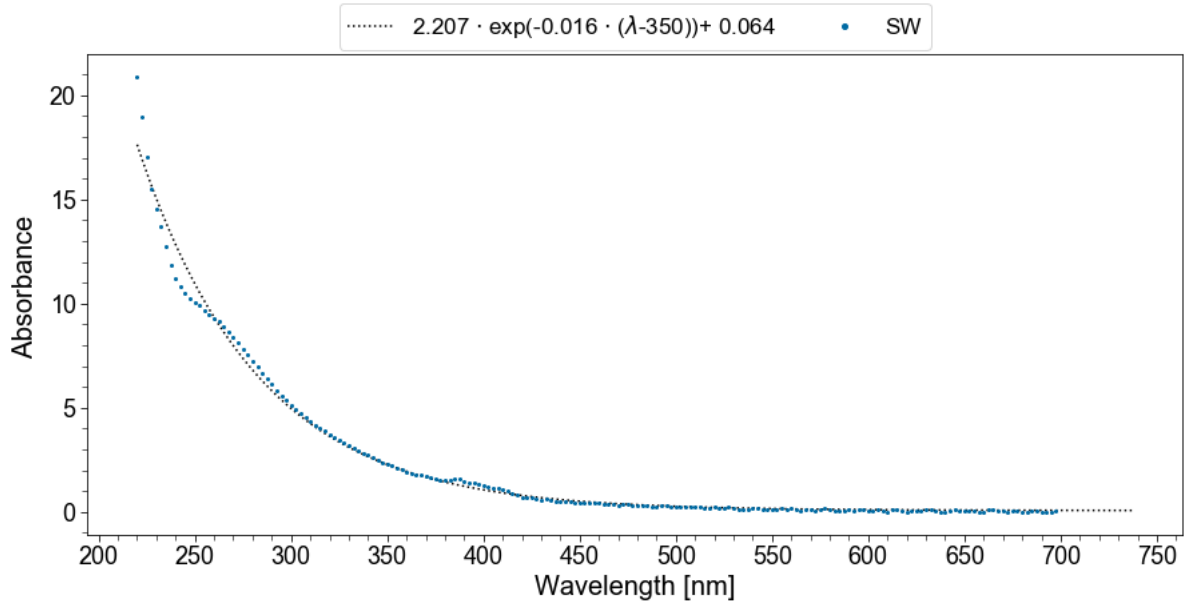

Figure A9. Example of an exponential curve (black dotted line) fit to the spectro::lyser absorbance spectrum (absorbance per meter) from surface water (SW) in mid-April 2018. The reference wavelength (λ_0) was 350 nm in the *abs_fit_exponential* function.

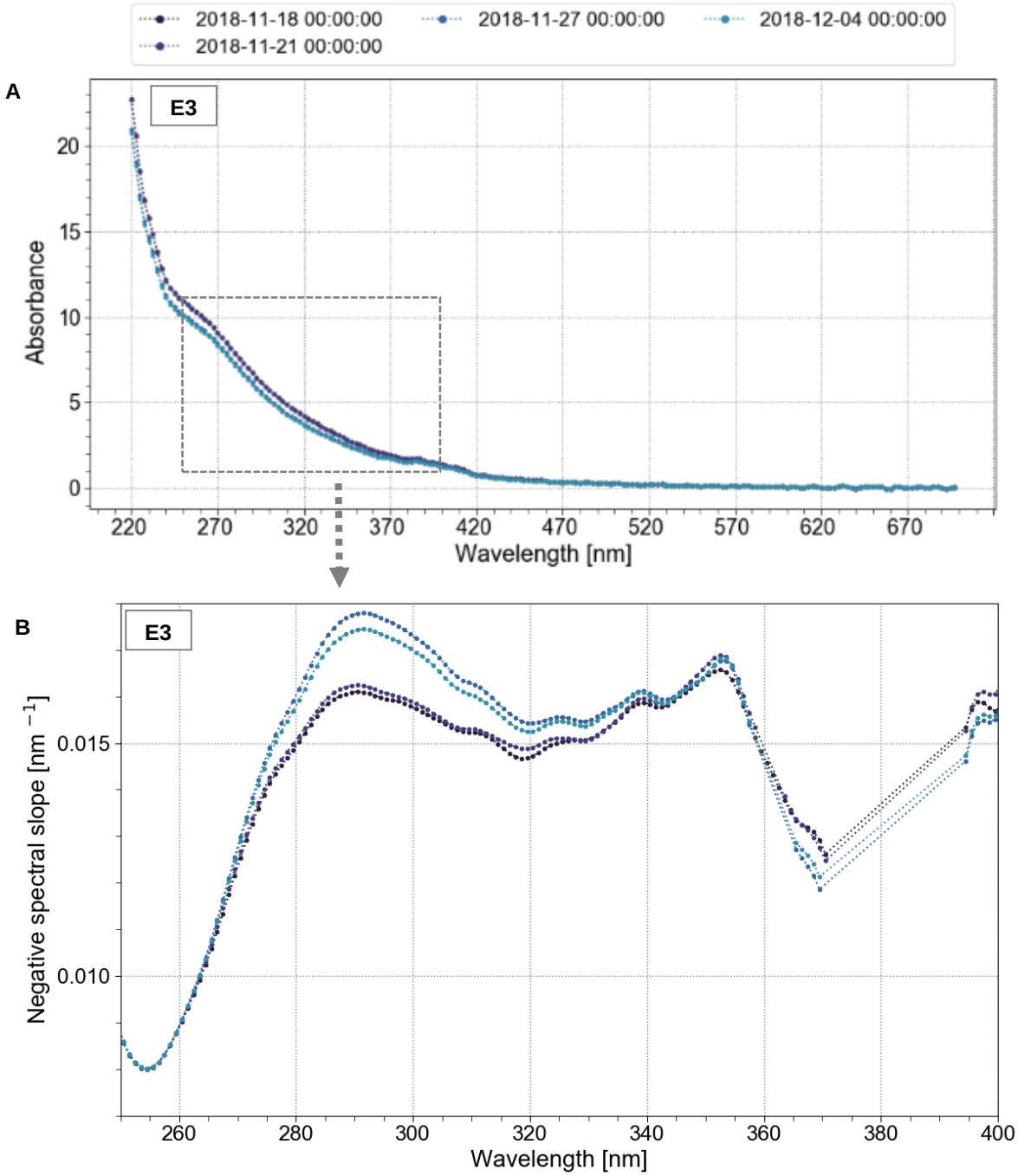

Figure A10. Absorbance spectra (upper plot, absorbance per meter) and spectral slope as a function of wavelength (lower plot) calculated using the *abs_spectral_curve* function. The selected dates cover the period before, during and after the Autumn lake circulation event (November-December 2018, event 3, Figure 2).

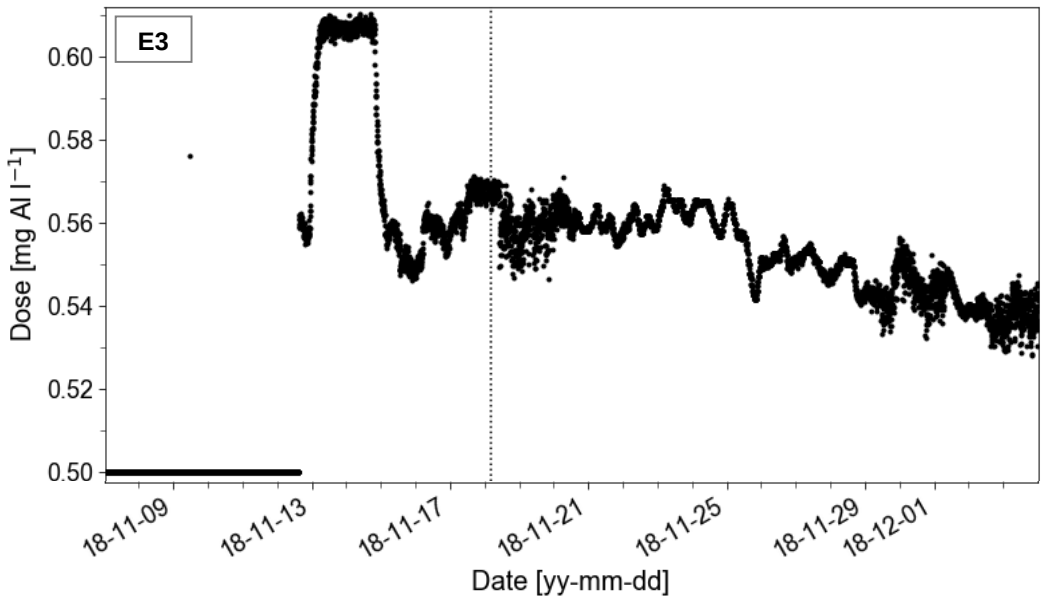

Figure A11. Optimal coagulant dose as calculated by the coagulation control system in figure A3 during event 3 for comparison with Figure 6.

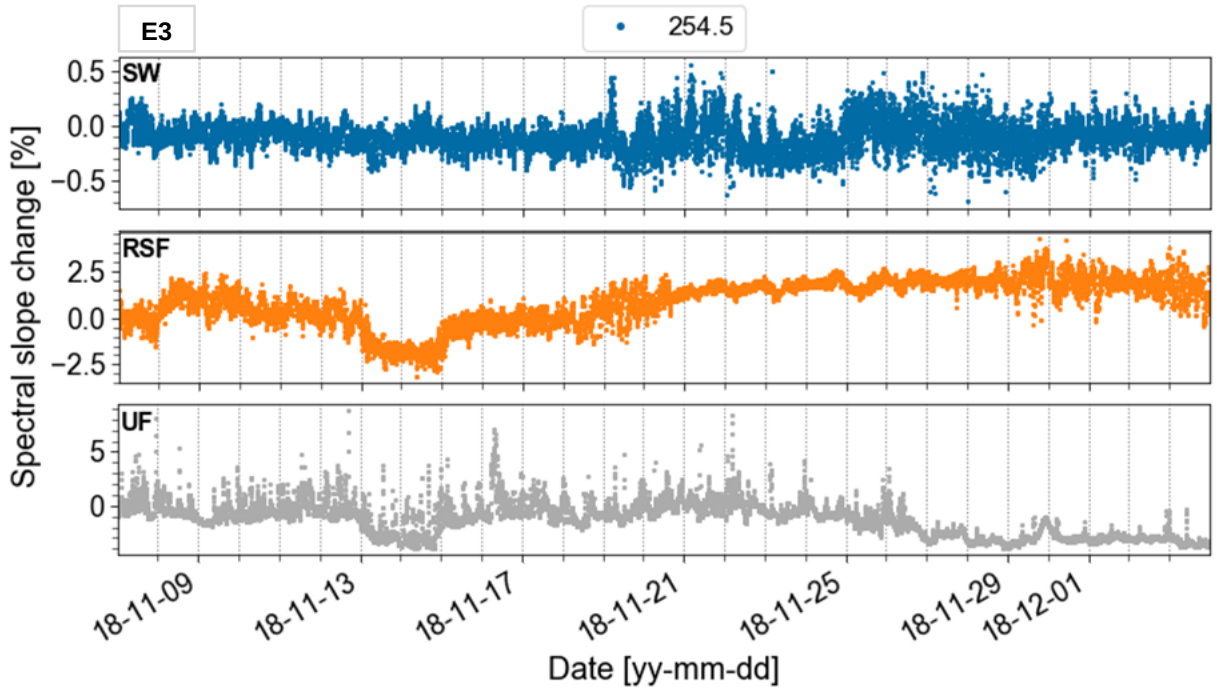

Figure A12. Time series of spectral slope percent changes in spectro::lysers absorbance data at 254.5 nm. The plots show surface water (SW), rapid sand filtrate (RSF) and ultrafilter permeate (UF) in the period before, during and after the Autumn lake circulation event (November-December 2018, event 3, Figure 2).