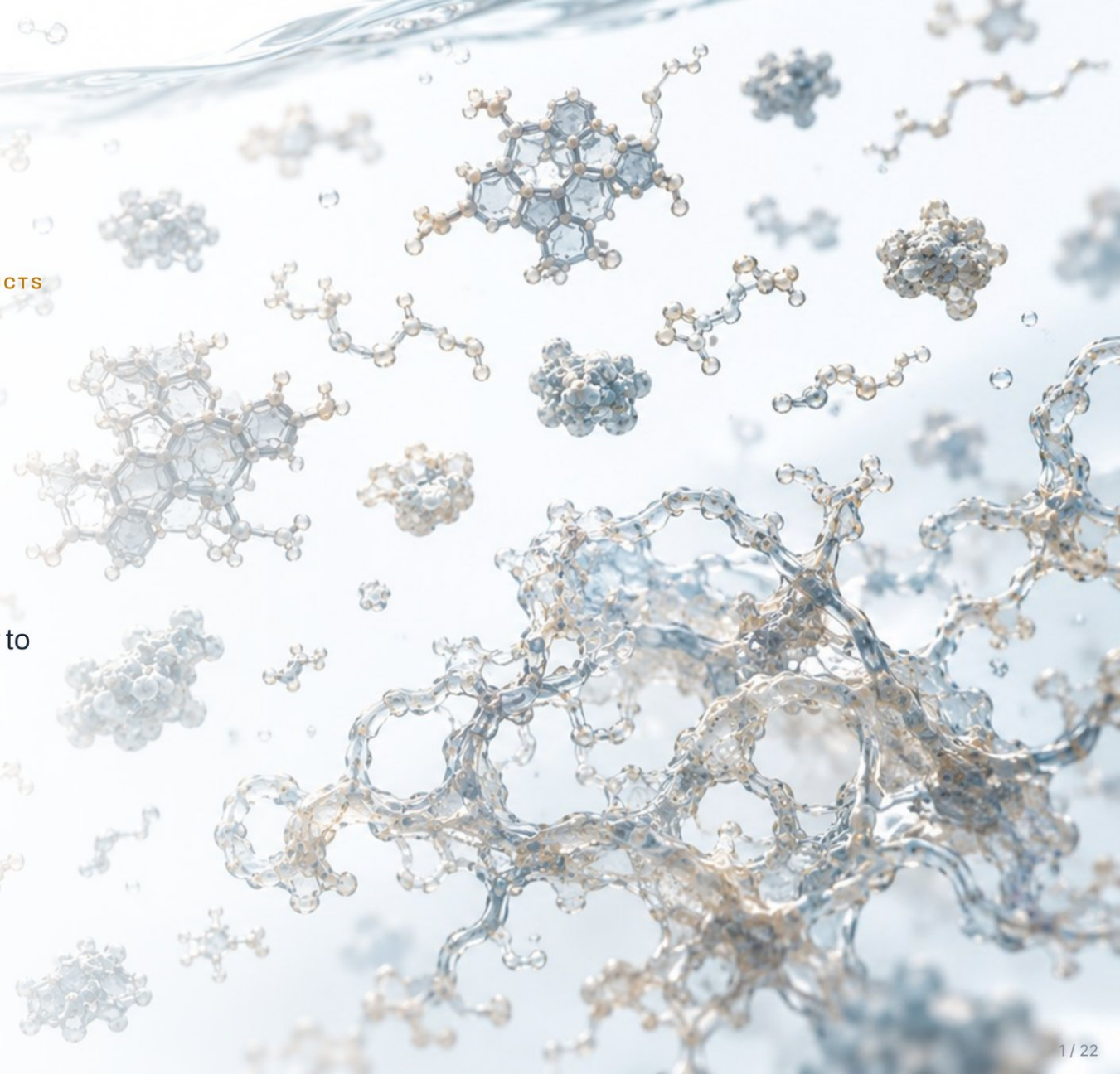
III JORNADA TÈCNICA · CÀTEDRA ATL · DISINFECTION BY-PRODUCTS

Separate to predict

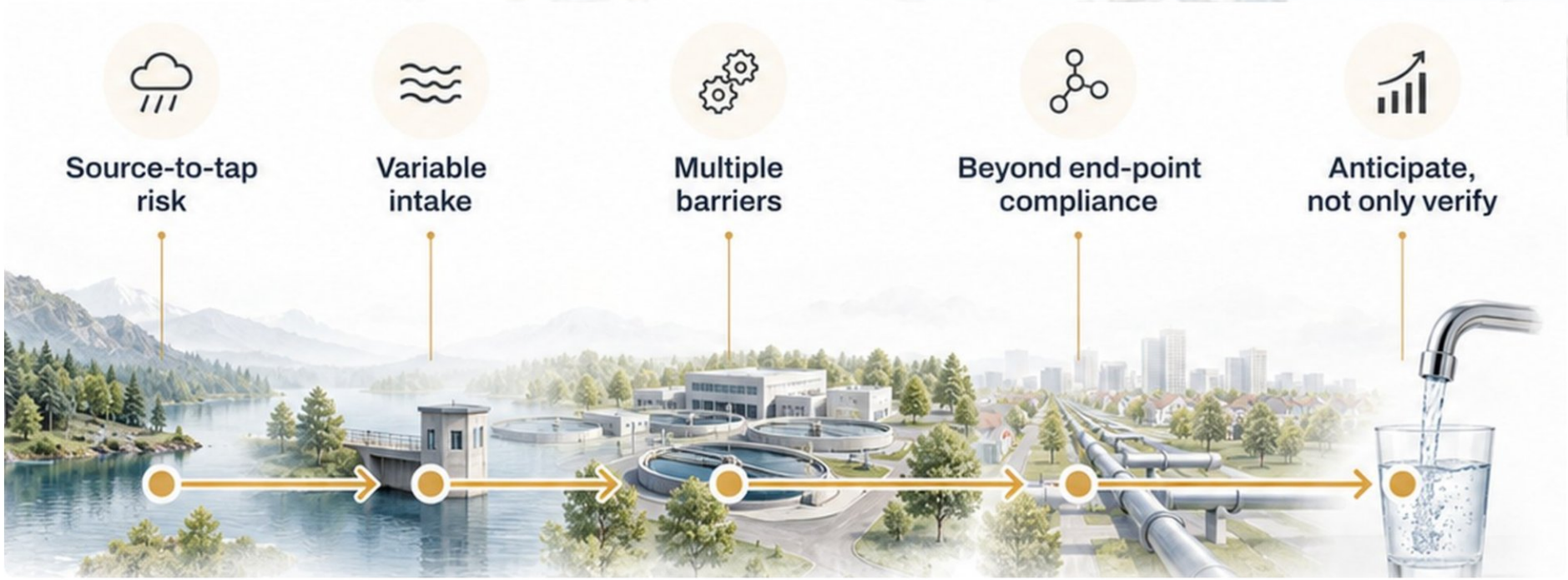
From the characterisation of dissolved organic matter to decision support in the plant.



Maria J. Martín · LEQUIA, Universitat de Girona
Barcelona, 26 October 2026

Universitat de Girona · lequia UdG ECO-INNOVATIVE ENVIRONMENTAL SOLUTIONS

Water safety plans already ask it. The plant cannot yet measure it.



Which disinfection by-products will form, and can we see it coming?

THMs and HAAs are measured at the tap, days later, as verification. No operational parameter reads their precursors before the chlorine.

The digital twin of the plant needs to know the organic matter

Data-driven knowledge
historical data · sensors · water quality

Experimental knowledge
DOM · DBPs · advanced characterisation

Expert knowledge
operation · process understanding · decision criteria

DIGITAL TWIN

DrinkIA

Anticipate

Assess risk

Decide



THE EXPERIMENTAL KNOWLEDGE · DOM AND DBPs · THREE PROJECTS

WATSON
CTM2017-83598-R · water treatment works operation and optimisation, where DrinkIA was born

SHERLOCK
PID2020-112615RA-I00 · digital tools and real-time control for drinking water plants

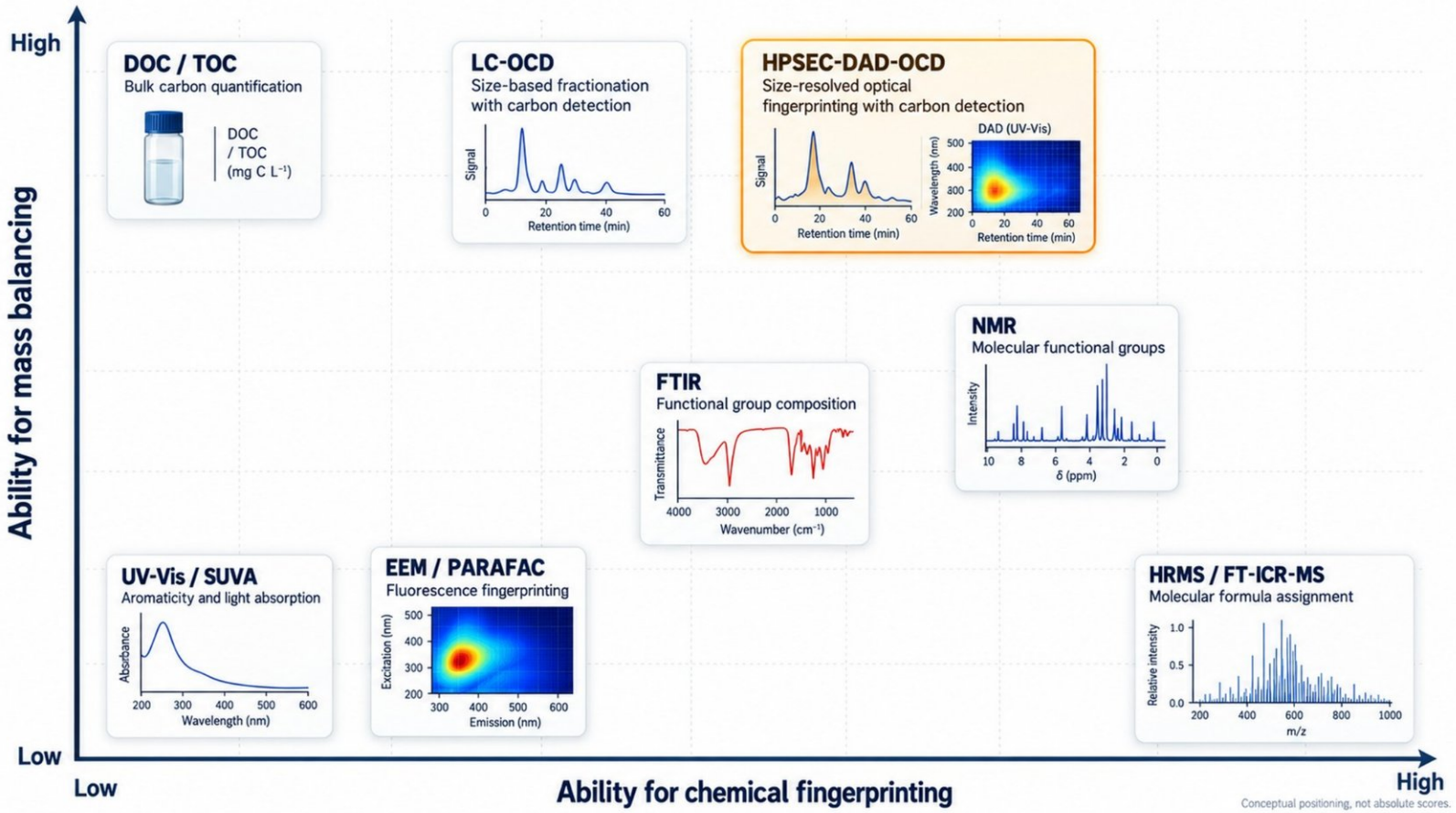
MORIARTY
PID2024-155733OB-I00 · chemical, microbiological and social risks of potable reuse

lequia UdG ECO-INNOVATIVE ENVIRONMENTAL SOLUTIONS

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA E INNOVACIÓN

By-products are decided before the chlorine: which DOM, and how to see it

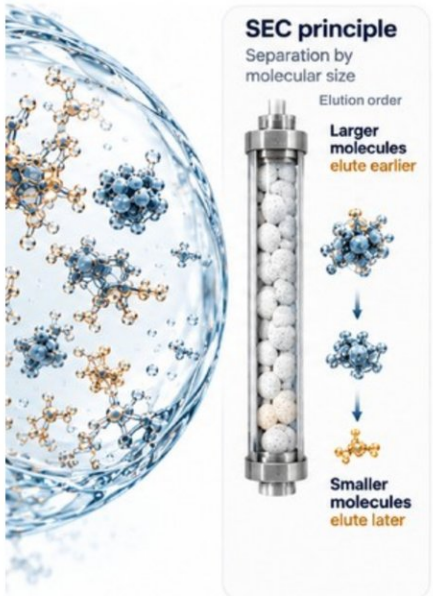
organic matter + chlorine → THMs · haloacetic acids THM4 ≤ 100 µg/L · HAA5 ≤ 60 µg/L



From how much DOC to which DOM: we started from LC-OCD and took it one step to the right.

We chose LC-OCD, and evolved it into HPSEC-DAD-OCD

Size-exclusion chromatography with carbon detection is the reference for fractionating DOM. We kept its logic and changed three things in the instrument.



LC-OCD

DOC-Labor · LC-OCD-OND



1 preparative SEC column

TSK HW-50S
250 × 20 mm

UVD: 254 nm

OCD: UV oxidation to CO_2
CO_2 detection:
infrared (NDIR)

HPSEC-DAD-OCD

Agilent 1260 Infinity II + Sievers M9 SEC



2 analytical columns in tandem

PL Aquagel-OH + OH MIXED-M
300 × 7.5 mm each · 5 / 8 µm

DAD: 190–640 nm, 2 nm step

OCD: UV/persulfate oxidation to CO_2
CO_2 detection:
membrane conductivity

Same principle

A size-exclusion column, an organic-carbon detector. What LC-OCD established in 2011 is what we keep: fractions quantified as carbon, not only as UV absorbance.

Two analytical columns in tandem

Where one preparative column gives one broad humic peak with a shoulder, the tandem separates humic substances from building blocks, in 60 min instead of 100.

A diode array instead of one wavelength

190–640 nm every 2 nm, on every fraction: each fraction gets its full spectrum, not a single A254 number. That is where the wavelengths of this talk come from.

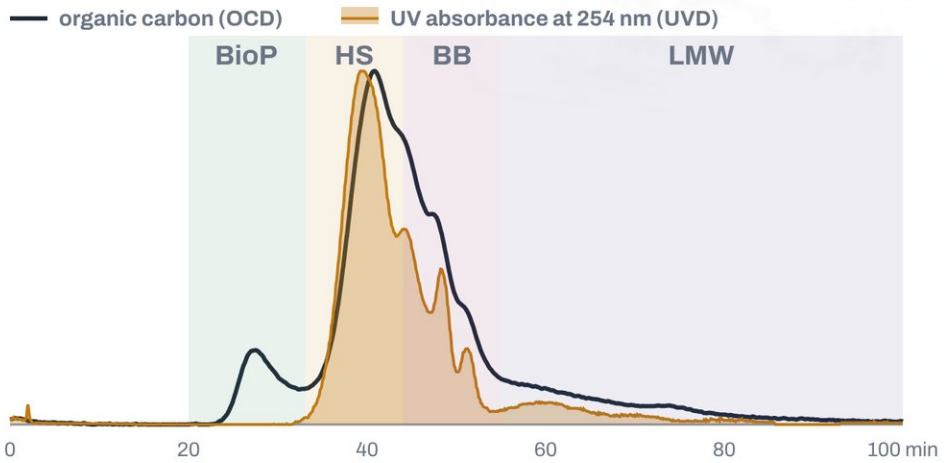
LC-OCD: Huber, Balz, Abert, Pronk (2011) Water Research 45, 879–885. HPSEC-DAD-OCD: Valentí-Quiroga et al. (2022) Water Research 227, 119314. LEQUIA instrument: Agilent 1260 Infinity II with PL aquagel-OH + PL aquagel-OH MIXED-M (300 × 7.5 mm) in series, DAD 190–640 nm, Sievers M9 SEC organic-carbon detector.

The same water, two chromatographs

Abrera intake, 6 May 2021 (DOC 2.5 mg/L, SUVA 2.4 L/(mg·m)): one bottle, sent to an LC-OCD laboratory and injected in our HPSEC-DAD-OCD.

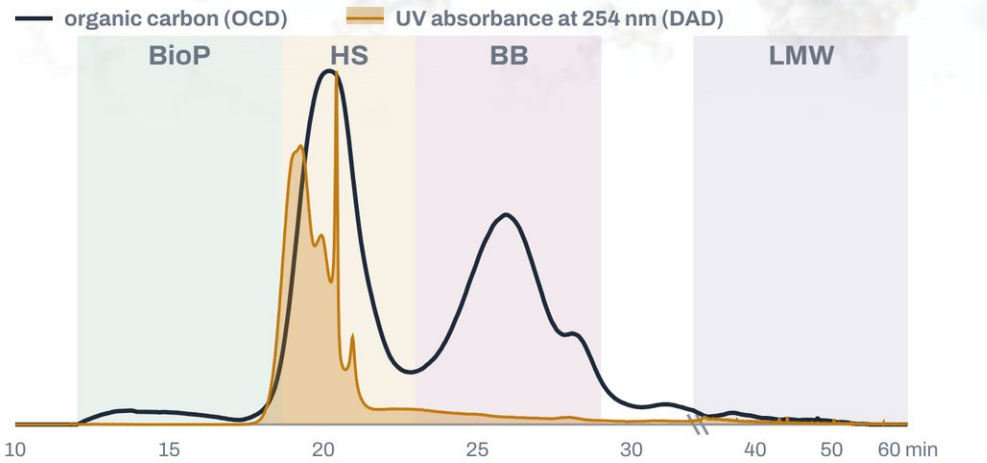
LC-OCD

one preparative column · UV at 254 nm · carbon and nitrogen · 100 min



HPSEC-DAD-OCD

two analytical columns · 190–640 nm, every fraction · carbon · 60 min



One humic peak with a shoulder becomes **humic substances** and **building blocks**, each with its full UV spectrum. **Four fractions instead of one DOC.**

BioP
SIZE > 10 kDa
SUVA ≈ 0.2
UV–VIS almost none above 230 nm

HS
SIZE 10 kDa – 700 Da
SUVA ≈ 4.5 · in LC-OCD 2.8
UV–VIS broad, from 220 nm into the visible

BB
SIZE 700 – 200 Da
SUVA ≈ 0.7 · far below the humics
UV–VIS as the humics, decaying faster

LMW
SIZE < 200 Da
SUVA low · too little signal at 254 nm
UV–VIS far UV only, below 230 nm

Same names, not the same fractions. LC-OCD fits a peak to the humics, we cut at fixed times: different separation, different fractions. Compare within one system, not across.

The by-products are not made by "the DOC" but by specific fractions, above all the humics. That is what we set out to measure.

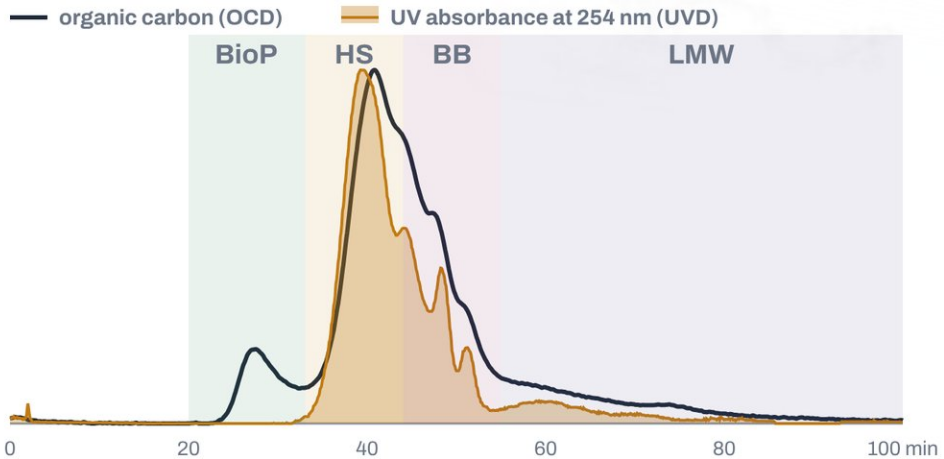
Abrera intake, 6 May 2021 · LC-OCD: DOC-Labor system, Het Waterlaboratorium (DOC and SUVA from its report) · SUVA per fraction: HPSEC-DAD-OCD, A254/DOC per window scaled to the bulk SUVA · apparent sizes: Valentí-Quiroga (2024) Table 7 · HPSEC-DAD-OCD: LEQUIA, seq. 072 · Valentí-Quiroga et al. (2022) Water Research 227, 119314

The same water, two chromatographs

Abrera intake, 6 May 2021 (DOC 2.5 mg/L, SUVA 2.4 L/(mg·m)): one bottle, sent to an LC-OCD laboratory and injected in our HPSEC-DAD-OCD.

LC-OCD

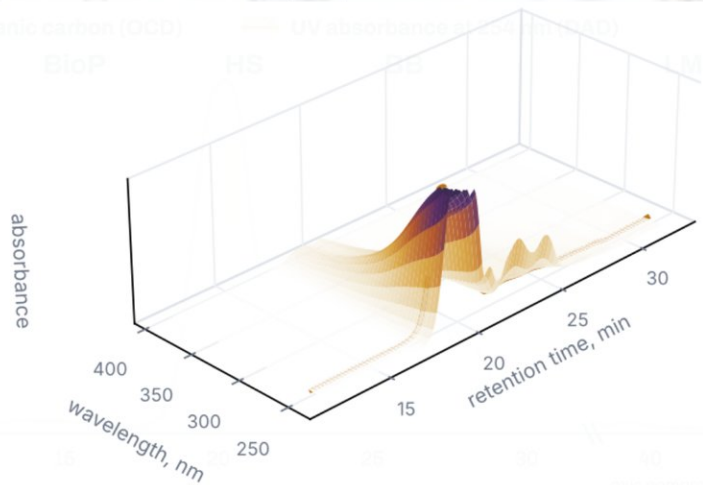
one preparative column · UV at 254 nm · carbon and nitrogen · 100 min



One humic peak with a shoulder becomes **humic substances** and **building blocks**, each with its full UV spectrum. **Four fractions instead of one DOC.**

HPSEC-DAD-OCD

two analytical columns · 190–640 nm, every fraction · carbon · 60 min



What the DAD adds: the whole surface, time × wavelength × absorbance. The 254 nm chromatogram is just one line across it.

BioP
SIZE > 10 kDa
SUVA ≈ 0.2
UV–VIS almost none above 230 nm

HS
SIZE 10 kDa – 700 Da
SUVA ≈ 4.5 · in LC-OCD 2.8
UV–VIS broad, from 220 nm into the visible

BB
SIZE 700 – 200 Da
SUVA ≈ 0.7 · far below the humics
UV–VIS as the humics, decaying faster

LMW
SIZE < 200 Da
SUVA low · too little signal at 254 nm
UV–VIS far UV only, below 230 nm

Same names, not the same fractions. LC-OCD fits a peak to the humics, we cut at fixed times: different separation, different fractions. Compare within one system, not across.

The by-products are not made by "the DOC" but by specific fractions, above all the humics. That is what we set out to measure.

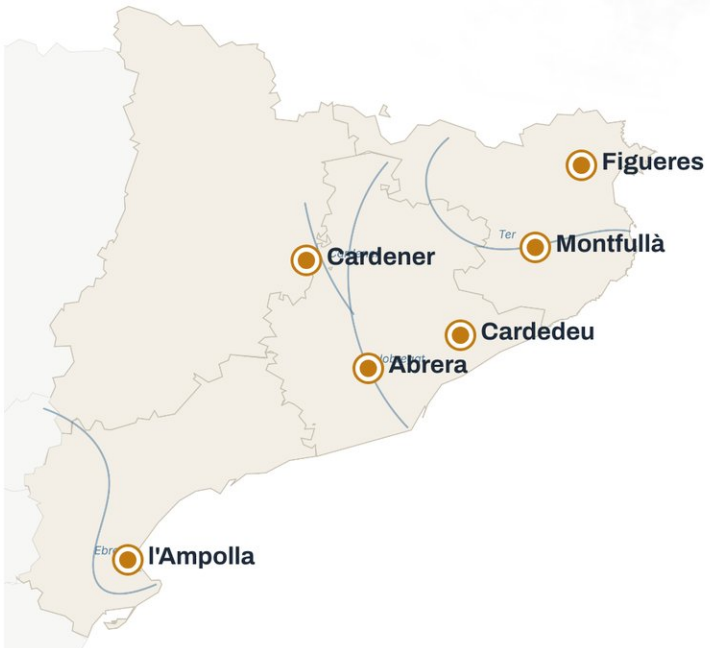
Abrera intake, 6 May 2021 · LC-OCD: DOC-Labor system, Het Waterlaboratorium (DOC and SUVA from its report) · SUVA per fraction: HPSEC-DAD-OCD, A254/DOC per window scaled to the bulk SUVA · apparent sizes: Valentí-Quiroga (2024) Table 7 · HPSEC-DAD-OCD: LEQUIA, seq. 072 · Valentí-Quiroga et al. (2022) Water Research 227, 119314

QUESTION 1

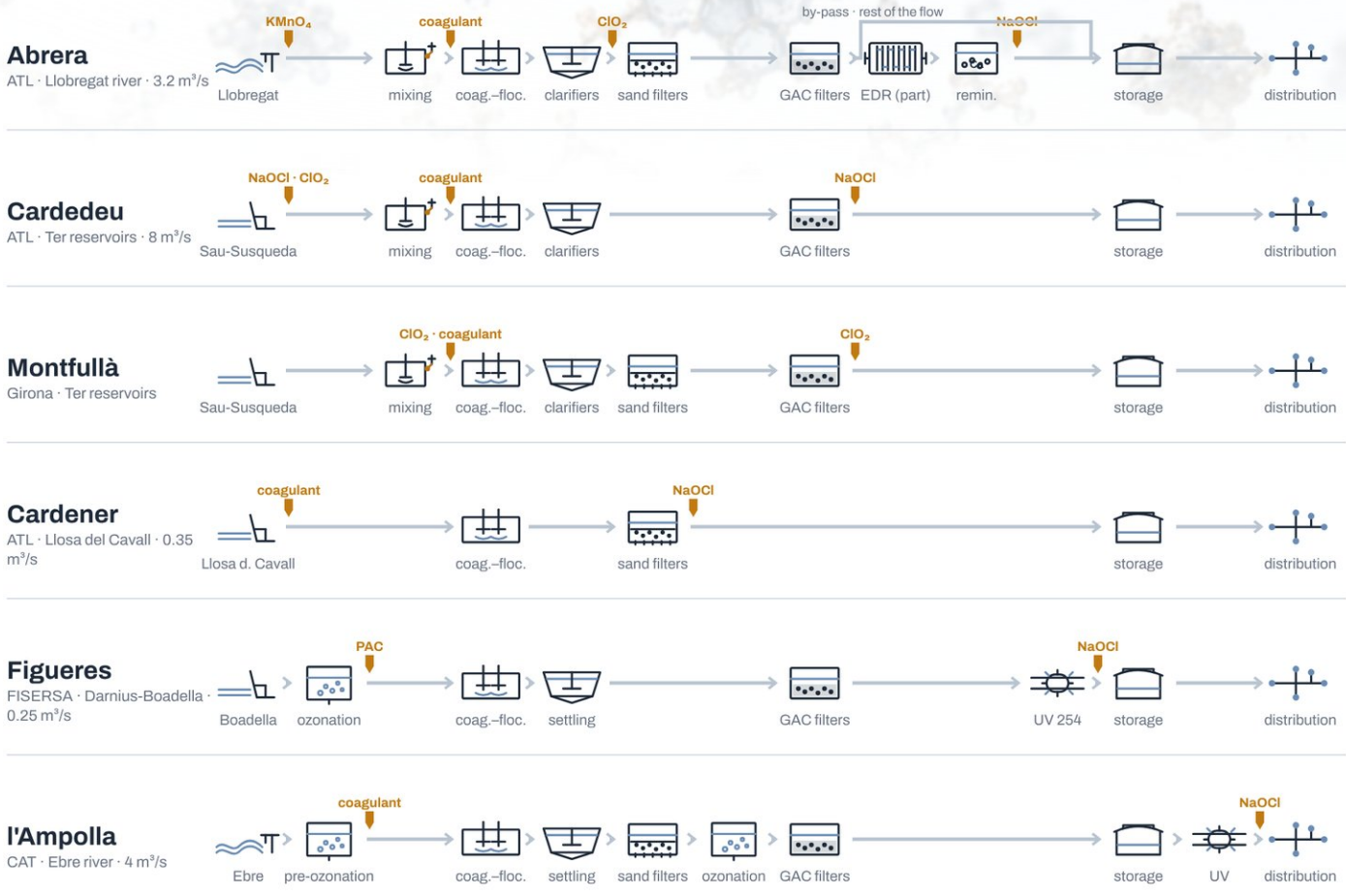
We measure along the train: what leaves at each stage, and what that does to the disinfection by-products?

Three plants, one instrument: the chromatogram at every sampling point, and the chlorine test on the same water.

Six plants, rivers and reservoirs, one instrument



Full treatment trains, 0.25 to 8 m³/s, river and reservoir sources, units aligned by type. Amber vials: reagent dosing points. The same instrument and the same four fractions at every plant.



Trains after Valentí-Quiroga et al. (2026) Fig. 1 and thesis §1 (Abrera, Cardedeu, Montfullà, Cardener), LEQUIA–UdG and Consorci d'Aigües de Tarragona (l'Ampolla) and Ferràndez-Galceran (2026) Fig. 3.2 (Figueres). Schematic map. HPSEC-DAD-OCD at LEQUIA, Universitat de Girona: Agilent 1260 · two PL Aquagel-OH columns in series · DAD 190–640 nm · Sievers M9 organic carbon detector.

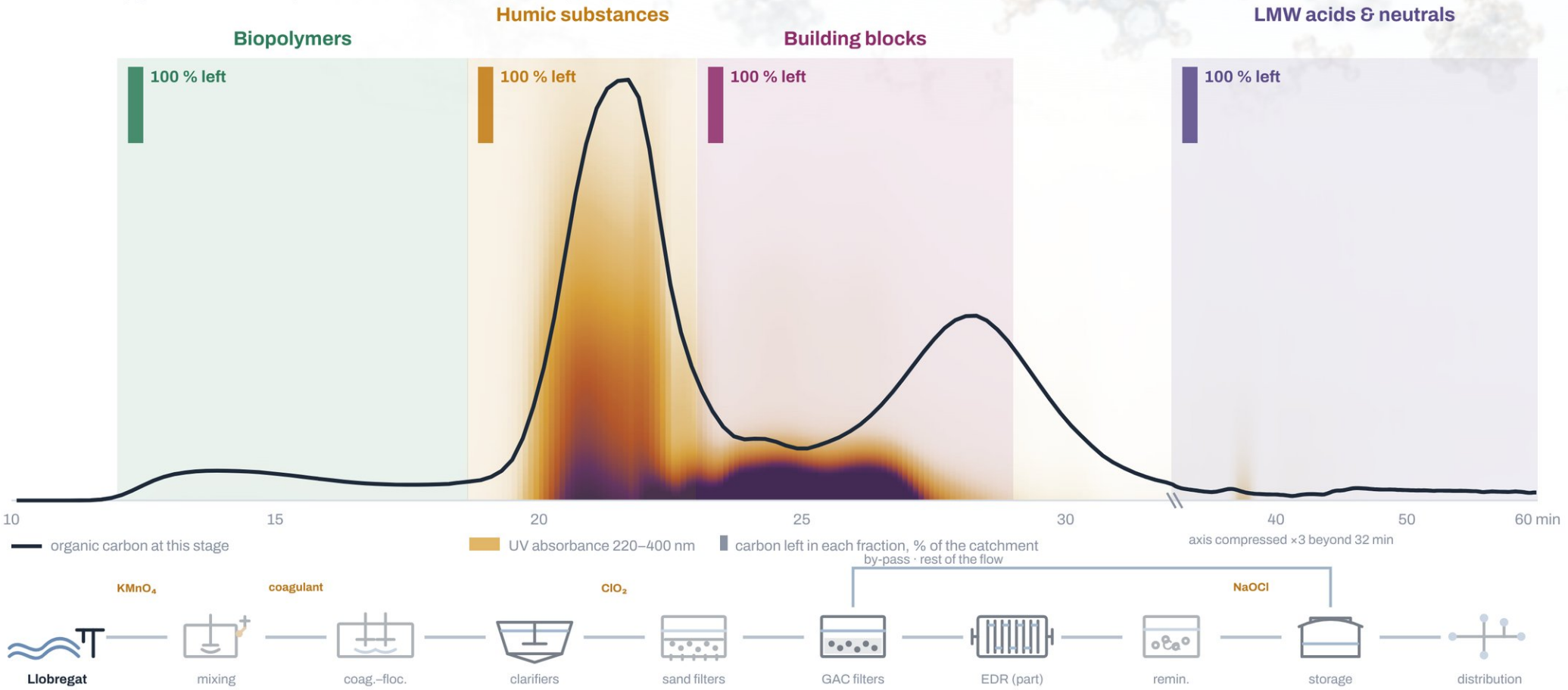
Conventional or advanced: who removes what

Abrera DWTP · Llobregat · campaign of February 2023 · five sampling points (thesis of M. Valentí-Quiroga)

99 µg/L THM4 formation potential, same sample chlorinated (48 h)

Catchment

100 % of intake carbon remaining



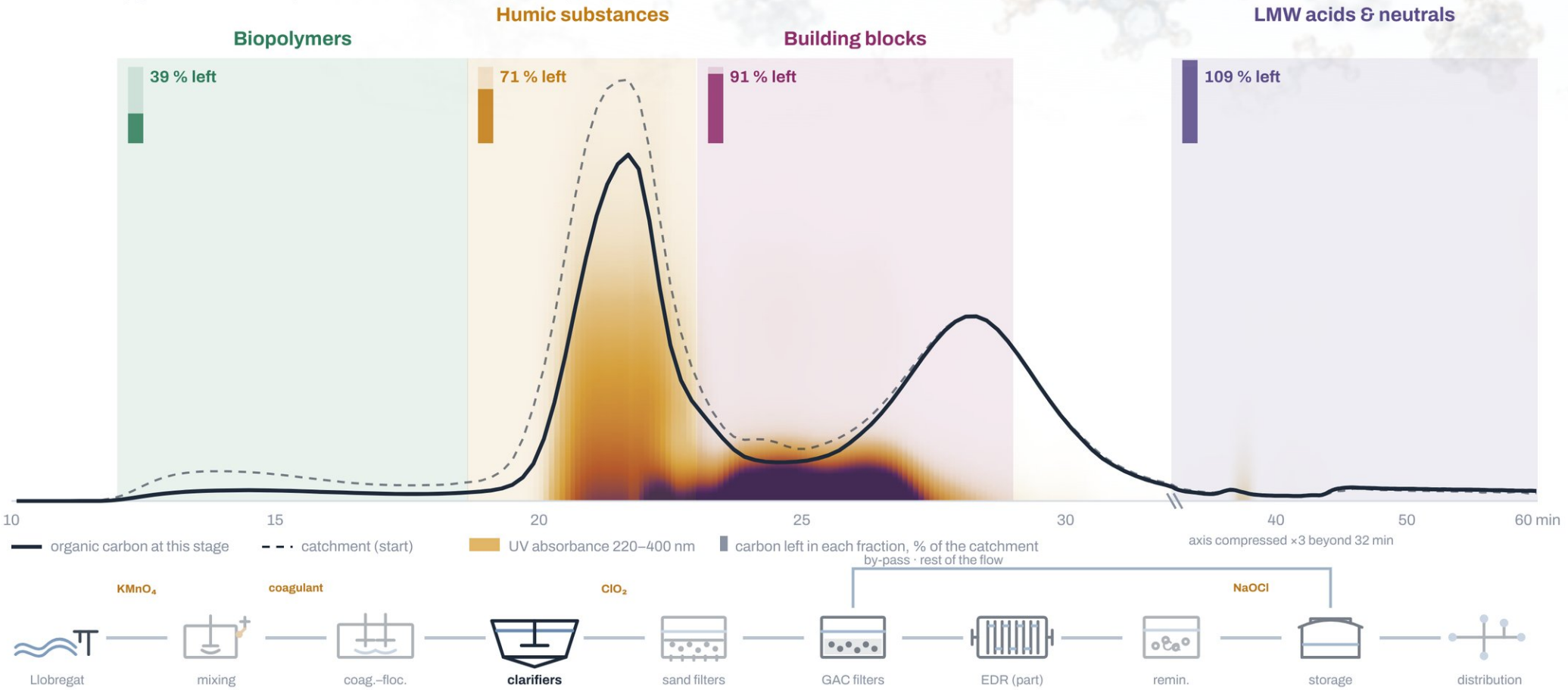
Conventional or advanced: who removes what

Abrera DWTP · Llobregat · campaign of February 2023 · five sampling points (thesis of M. Valentí-Quiroga)

Coagulation & flocculation

100 → **82** % of intake carbon remaining

99 → **56** µg/L THM4 formation potential, same sample chlorinated (48 h)



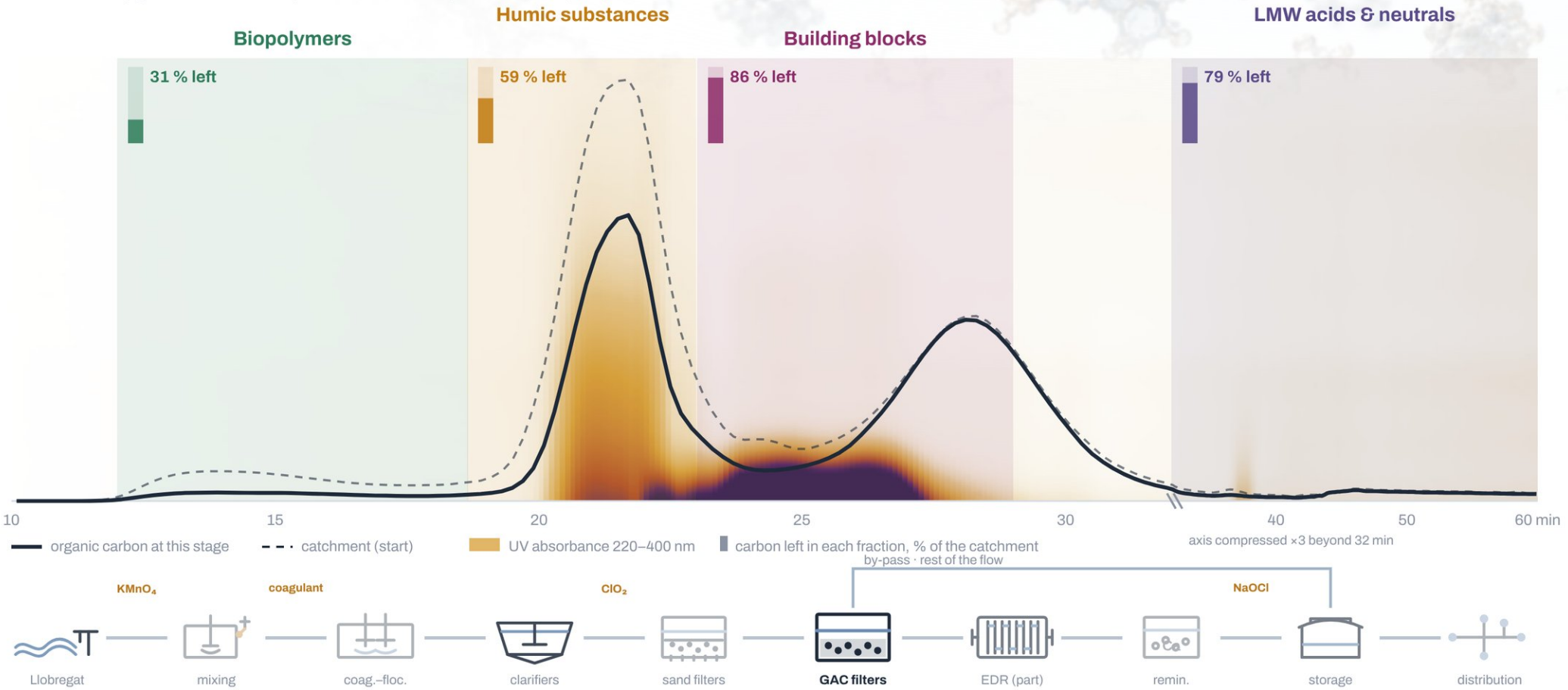
Conventional or advanced: who removes what

Abrera DWTP · Llobregat · campaign of February 2023 · five sampling points (thesis of M. Valentí-Quiroga)

99 → 69 µg/L THM4 formation potential, same sample chlorinated (48 h)

GAC filtration

100 → 70 % of intake carbon remaining



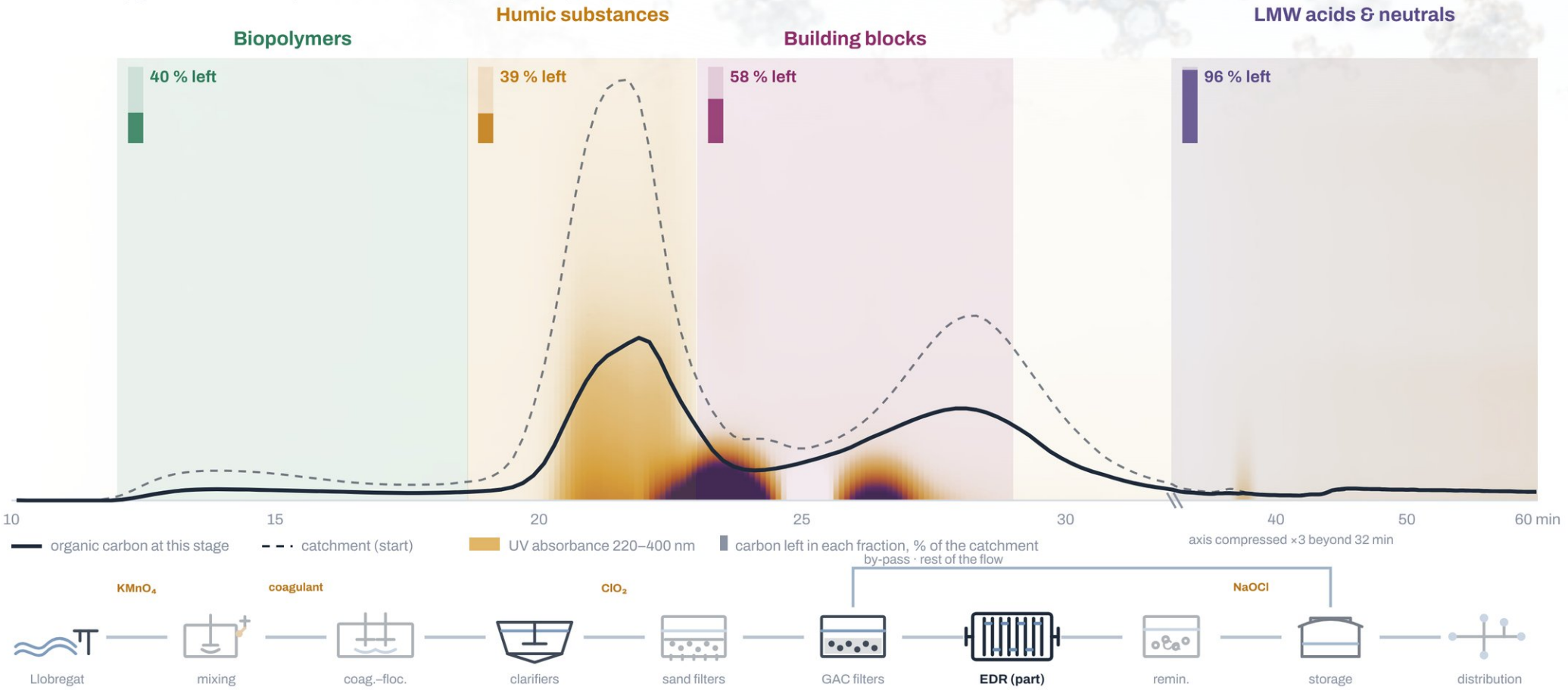
Conventional or advanced: who removes what

Abrera DWTP · Llobregat · campaign of February 2023 · five sampling points (thesis of M. Valentí-Quiroga)

99 → 38 µg/L THM4 formation potential, same sample chlorinated (48 h)

Electrodialysis reversal

100 → 52 % of intake carbon remaining



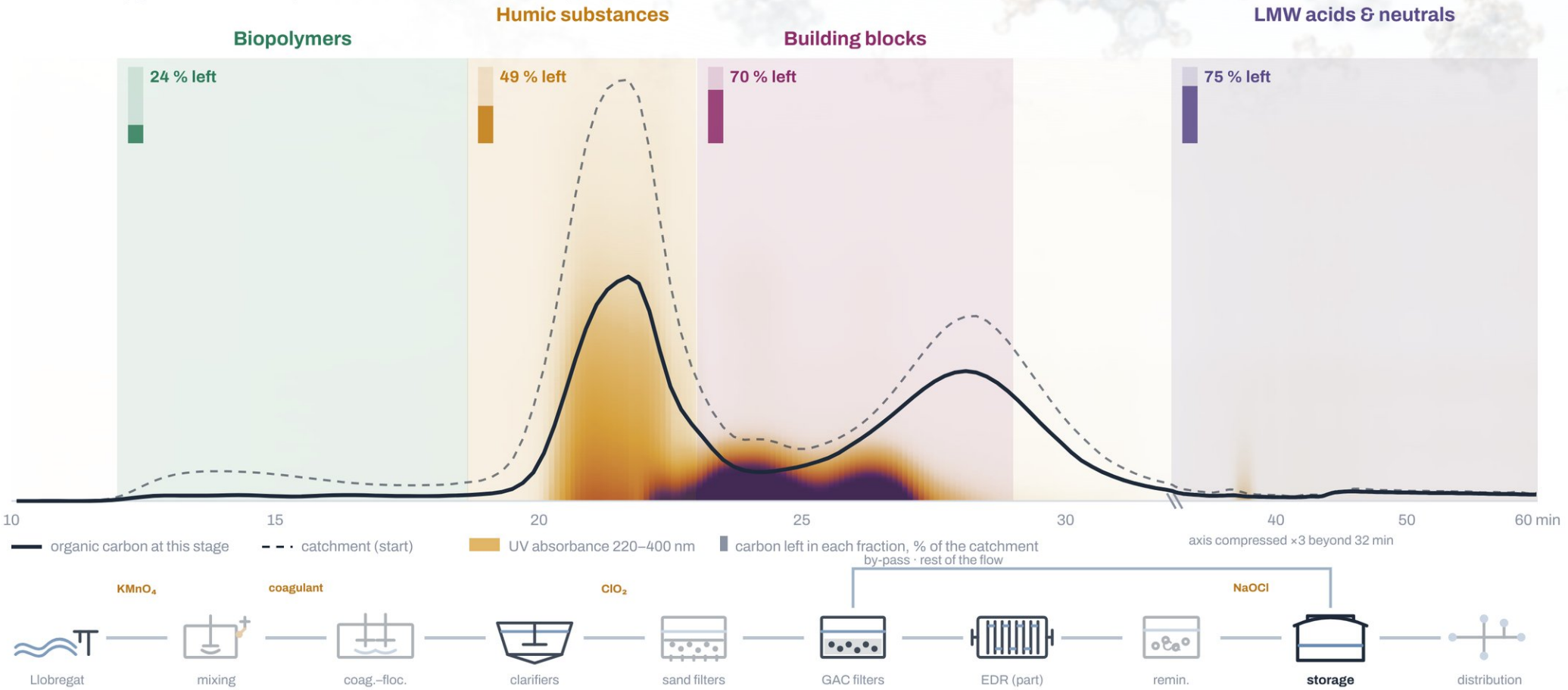
Conventional or advanced: who removes what

Abrera DWTP · Llobregat · campaign of February 2023 · five sampling points (thesis of M. Valentí-Quiroga)

Treated water (50 % blend)

100 → **58** % of intake carbon remaining

99 → **54** µg/L THM4 formation potential, same sample chlorinated (48 h)



Do the usual units remove enough, or always the same profile?

Per process step, % removed of what enters the step: it depends on the water and on the steps before it. The darker, the more removed.

		CARBON REMOVED, PER FRACTION · OCD				FORMATION POTENTIAL		WHAT THE STEP REMOVES
		Biopolymers	Humic substances	Building blocks	LMW acids & neutrals	THM4-FP	HAA-FP	
	Coagulation Abrera	−61 %	−34 %	−11 %	*n.d.*	−44 %	−81 %	biopolymers and humics; building blocks barely
	Coagulation Cardedeu	−33 %	−13 %	−6 %	*n.d.*	−8 %	−13 %	
	Coagulation Montfullà	−20 %	−18 %	−1 %	*n.d.*	−24 %	−47 %	
	Coagulation l'Ampolla	−40 %	−18 %	−10 %	−7 %	−17 %	−21 %	
	Sand filtration Montfullà	−14 %	+2 %	+9 %	*n.d.*	−10 %	*n.r.**	depends on what comes before: more after pre-ozonation
	Sand filtration l'Ampolla	−38 %	−17 %	−12 %	−3 %	−23 %	−32 %	
	GAC filtration Abrera	−22 %	−17 %	−1 %	*n.d.*	*n.r.**	*n.r.**	a modest share of every fraction
	GAC, newest filter Montfullà	+1 %	−10 %	−20 %	*n.d.*	−7 %	−56 %	
	GAC filtration l'Ampolla	−46 %	−17 %	−11 %	−3 %	−20 %	−18 %	
	Pre-ozonation l'Ampolla	+11 %	−4 %	−6 %	0 %	−19 %	−1 %	carbon barely changes; formation potential still falls
	Intermediate ozonation l'Ampolla	0 %	−4 %	0 %	+18 %	−9 %	−19 %	
	Electrodialysis (EDR) Abrera	*n.d.*	−34 %	−33 %	*n.d.*	−45 %	−26 %	REMOVE BY CHARGE humics and building blocks alike
	Ion exchange resin Abrera, bench	−1 %	−69 %	−32 %	−10 %	−76 %	*NA*	

n.d. not determined · ***NA*** not analysed · ***n.r.**** not reported: single-campaign increase above +20 % (atypical) · ***+x %*** the step increases it

Same unit, different water and train: coagulation takes 13–34 % of the humics and 8–44 % of THM4-FP.
To understand it, measure it fraction by fraction. And one profile repeats: the humics. All the humics alike?

Abrera, Cardedeu, Montfullà: Valentí-Quiroga (2024) Tables A6–A8, one campaign, 48 h test, HAA5; ion exchange at bench scale. l'Ampolla: LEQUIA–UdG and Consorci d'Aigües de Tarragona, medians of 12 monthly campaigns, fractions from LC-OCD composition (8 campaigns) and DOC change, 7 d test, HAA9. EDR relative to its feed, the GAC effluent.

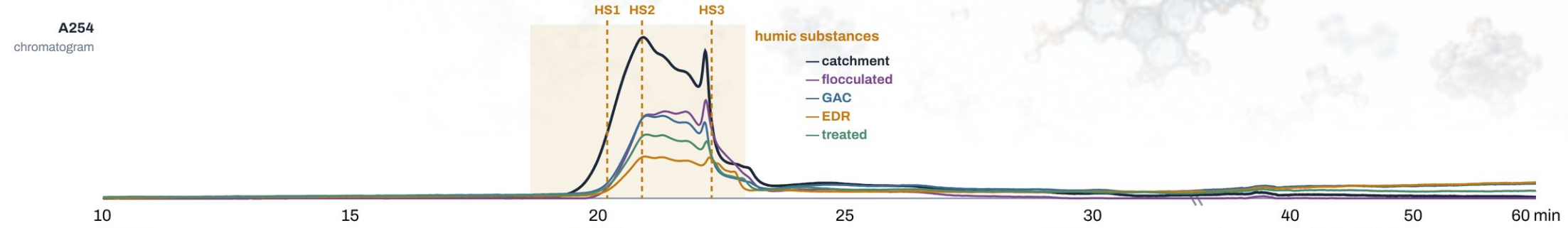
QUESTION 2

Are all the humics alike, or do some matter more for the by-products?

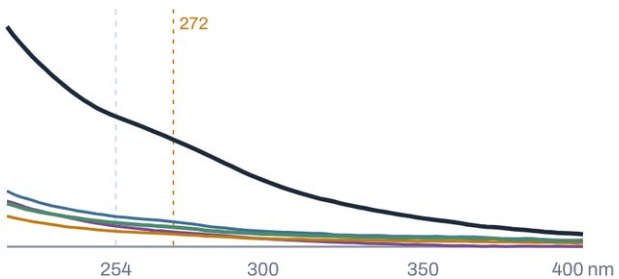
Three cuts inside the same humic peak: what each treatment stage leaves of them, and how each reacts with chlorine.

Do they remove all the humics alike? Three cuts, three responses

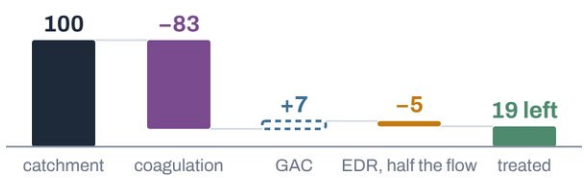
Abrera, Llobregat water, no chlorine. Top: the A254 chromatogram at each stage. Below: the full UV spectrum at three times inside the humic peak, stage by stage.



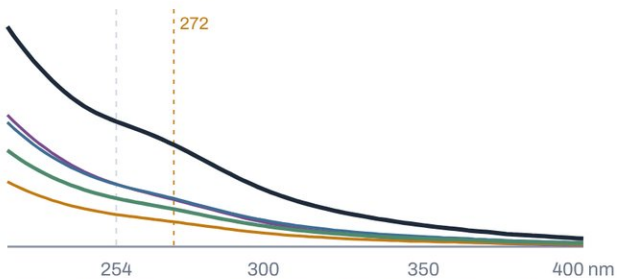
HS1 20.2 min
leading edge · larger molecules



A254 · CATCHMENT = 100 · WHAT EACH STEP TAKES



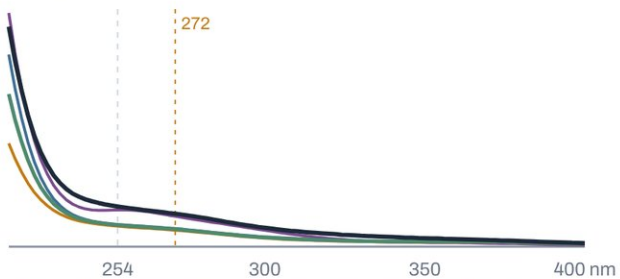
HS2 20.9 min
the apex of the humic peak



A254 · CATCHMENT = 100 · WHAT EACH STEP TAKES



HS3 22.3 min
trailing shoulder · smaller molecules



A254 · CATCHMENT = 100 · WHAT EACH STEP TAKES



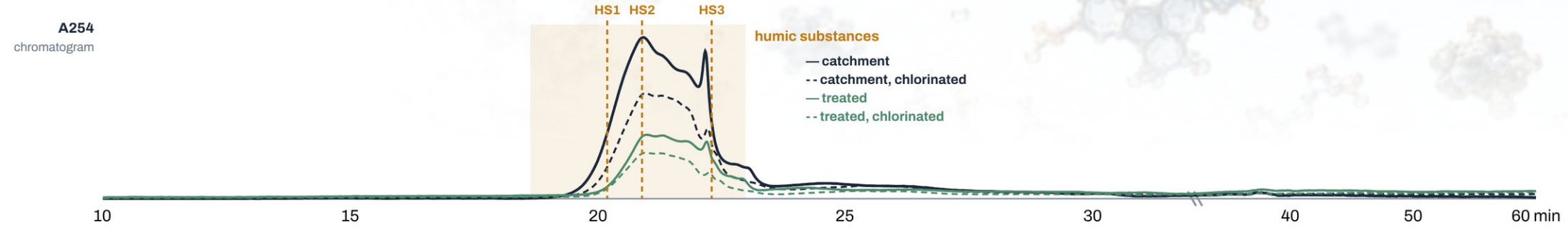
Catchment · Flocculated · GAC · EDR · Treated (50 % blend)

Three cuts of the same peak, three different responses to the same train.

Raw DAD, no smoothing · spectra at 20.2, 20.9 and 22.3 min · Abrera, February 2023 (SK090–SK094) · spectra scaled to the catchment maximum · bars: A254 ±0.15 min around each cut; treated = 50 % GAC + 50 % EDR; dashed = increase.

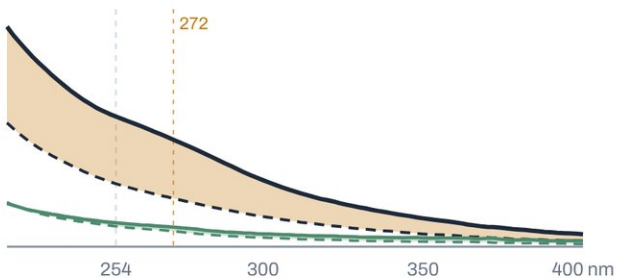
Do they react with chlorine alike? The same three cuts, chlorinated

Abrera. The same three cuts, catchment and treated water, each before and after the formation-potential test. Shaded: what the chlorine took.

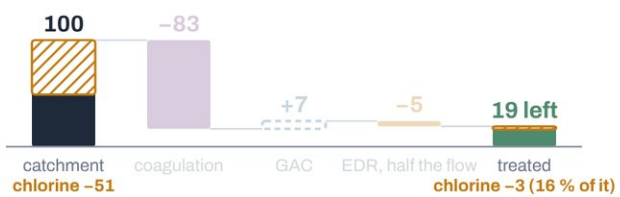


HS1 20.2 min

leading edge · larger molecules

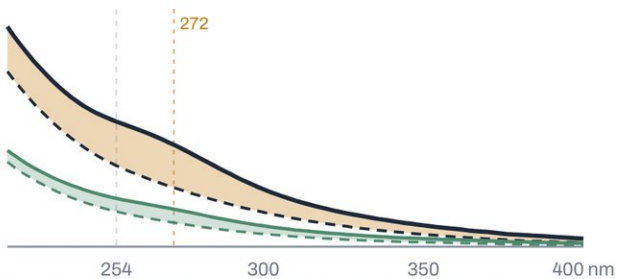


A254 · CATCHMENT = 100 · HATCHED: TAKEN BY CHLORINE



HS2 20.9 min

the apex of the humic peak

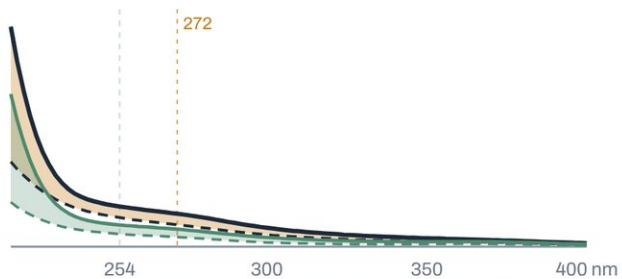


A254 · CATCHMENT = 100 · HATCHED: TAKEN BY CHLORINE



HS3 22.3 min

trailing shoulder · smaller molecules



A254 · CATCHMENT = 100 · HATCHED: TAKEN BY CHLORINE



Catchment · After the chlorine test · Treated (50 % blend) · After the chlorine test · What the chlorine took

Three cuts, three reactivities, and they change along the train: in the treated water the shoulder is the most reactive.

Raw DAD, no smoothing · spectra at 20.2, 20.9 and 22.3 min · Abrera, February 2023, SK090/SK094 and their chlorinated pairs (seq. 191 / 194) · spectra scaled to the catchment maximum · bars: A254 ±0.15 min around each cut.

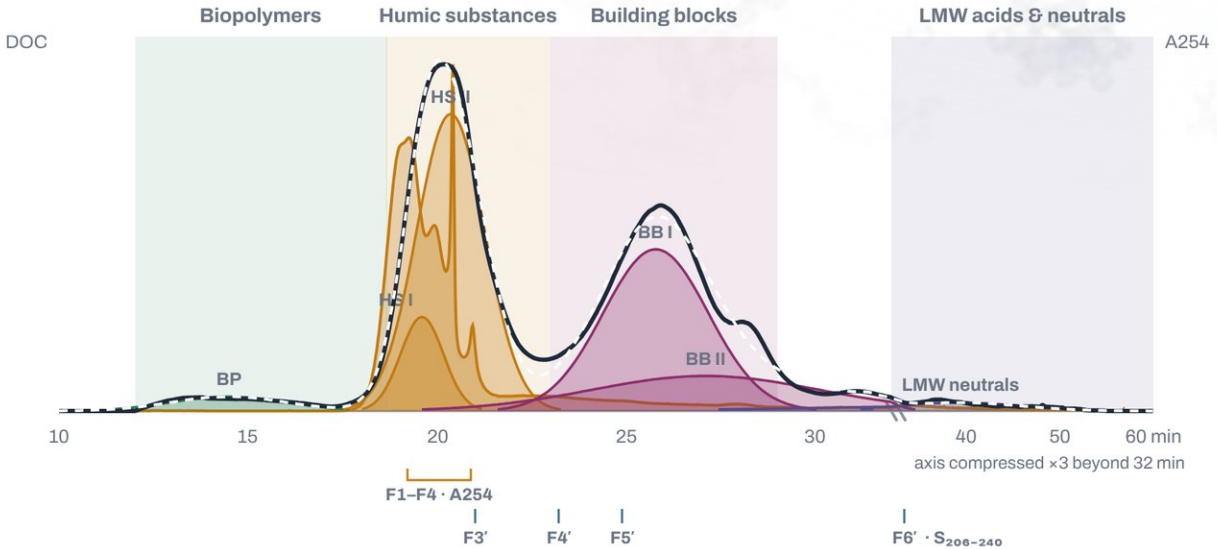
QUESTION 3

We separate to predict: can the fractions predict the DBPs?

Carbon and spectra per fraction, against the formation-potential test of the same water.

Fractionate, and the DBPs become predictable

Twenty-one waters, three plants: chromatogram and formation-potential test on the same bottle; a linear model per detector.



No detector: bulk one number for the whole sample

$THM4\text{-}FP = \beta \cdot DOC \mid \beta \cdot A254 \mid \beta \cdot SUVA$

Carbon detector (OCD) carbon of each fraction

$THM4\text{-}FP = \beta_1 \cdot BP + \beta_2 \cdot HS_I + \beta_3 \cdot HS_{II} + \beta_4 \cdot BB_I + \beta_5 \cdot BB_{II} + \beta_6 \cdot LMW$

Diode array (DAD) spectrum of each fraction

$THM4\text{-}FP = \Sigma\, \beta_i \cdot A254_i \mid \Sigma\, \beta_i \cdot slope_i \mid hybrid$

absorbance and spectral slopes (206–240, 350–380 nm) per fraction

THE DBP FORMATION-POTENTIAL TEST, ON THE SAME BOTTLE

Chlorinate
NaOCl, excess

→

Incubate
24 · 48 · 72 h

→

Quench

→

Measure
THM4 · HAA

Valentí-Quiroga et al. (2022) Water Research 227, 119314 · thesis chapter IV.

How well each model fits: R^2adj of the formation potential at 48 h

Per detector and per THM species. < 0.80 | 0.80–0.90 | 0.90–0.95 | ≥ 0.95

	LLOBREGAT · 11 waters					TER – CARDENER · 10 waters			
	tTHM	$CHCl_3$	$CHBrCl_2$	$CHBr_2Cl$	$CHBr_3$	tTHM	$CHCl_3$	$CHBrCl_2$	$CHBr_2Cl$
BULK · no chromatograph									
A254	0.92	0.70	0.84	0.89	0.83	0.93	0.93	0.89	0.59
DOC	0.85	0.78	0.80	0.81	0.69	0.89	0.87	0.88	0.63
SUVA	0.77	0.61	0.68	0.74	0.70	0.92	0.86	0.97	0.87
CARBON DETECTOR · OCD									
Carbon per fraction	0.97	0.95	0.99	0.96	0.87	0.99	0.98	0.99	0.97
DIODE ARRAY · DAD									
A254 per fraction	0.96	0.79	0.88	0.96	0.90	0.95	0.95	0.91	0.82
Slopes per fraction ($S_{206-240}$, $S_{350-380}$)	0.98	0.84	0.92	0.97	0.94	0.98	0.97	0.98	0.88
Hybrid A254 + slope	0.96	0.95	0.95	0.97	0.98	0.99	0.99	0.99	0.99

Valentí-Quiroga, Daunis-i-Estadella, Emiliano, Valero, Martín (2022) Water Research 227, 119314; thesis chapter IV, Figure 14 (48 h column), simplified MLR models after stepwise selection, assumptions tested. Species ordered by abundance in each basin.

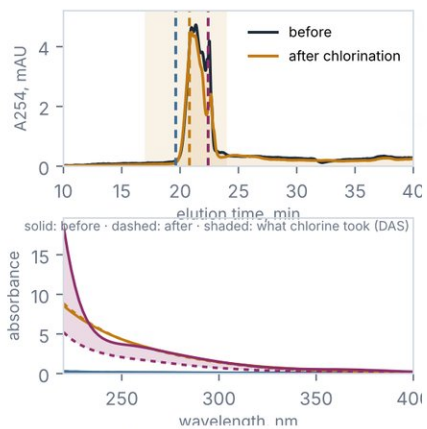
QUESTION 4

Can we predict from simpler measurements?

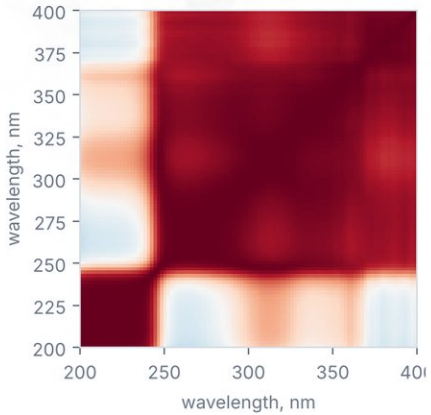
Which wavelengths, chosen with the chromatograph; then whole water on a laboratory spectrophotometer.

From 101 wavelengths to four, the same four at every plant

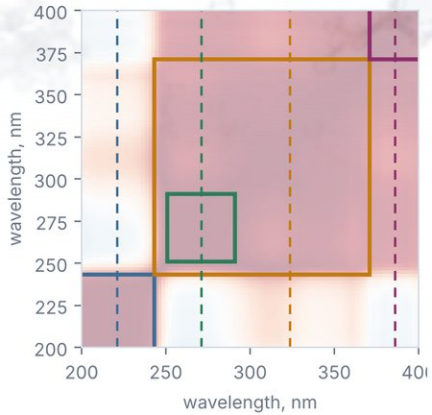
Seventeen waters, three plants, every stage: chlorinated in the formation-potential test and injected again.



1 · Cut and subtract. Before – after the chlorine, at each time of the humic peak.



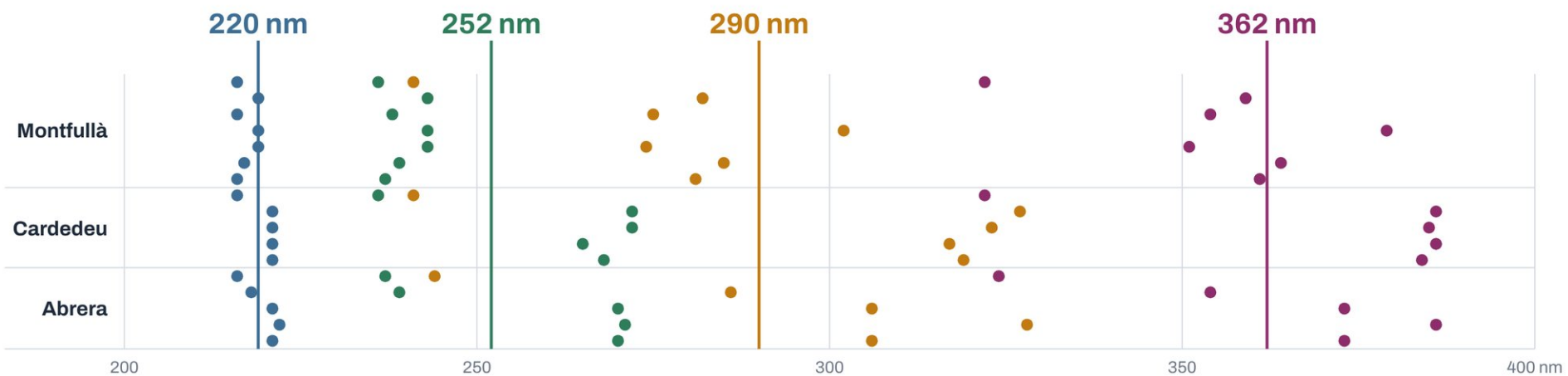
2 · Correlate. Which wavelengths rise and fall together.



3 · Cluster. Four groups; keep one wavelength per group.

4 · REPEAT FOR THE 17 WATERS

One dot per cluster and sample, below. The mean is the chosen wavelength: **220 · 252 · 290 · 362 nm**, the same four whatever the water and the stage.

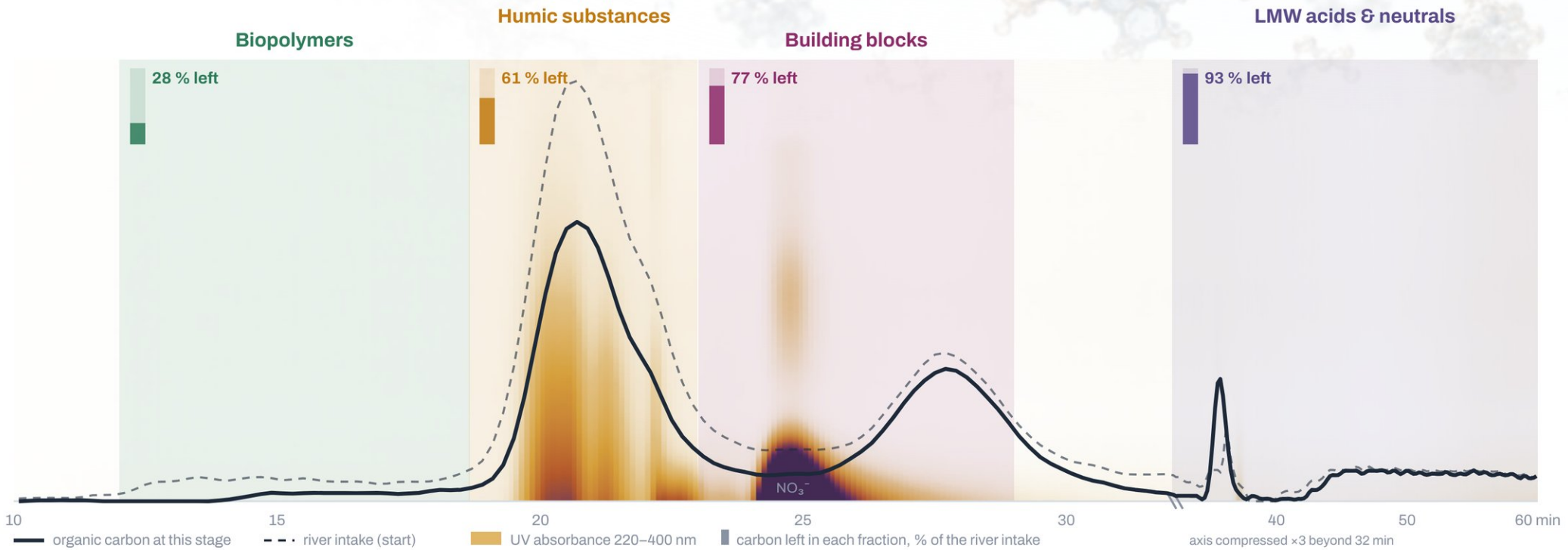


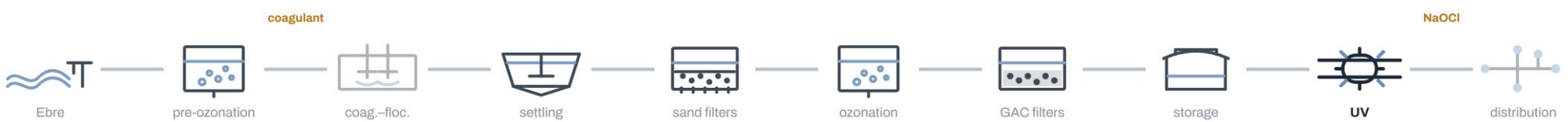
The plant the model never saw: l'Ampolla, on the Ebre

l'Ampolla DWTP (CAT) · lower Ebre · one of twelve monthly campaigns · ten sampling points

UV disinfection

100 → **60** % of intake carbon remaining

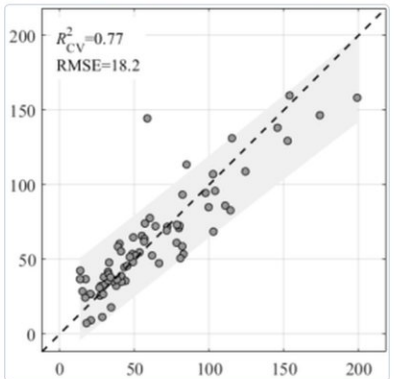




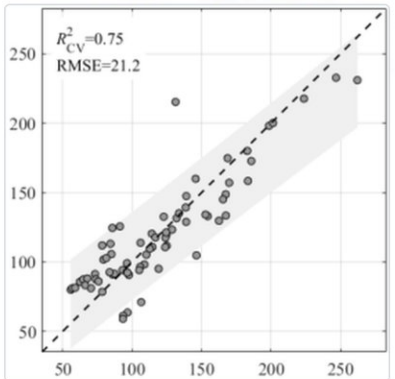
Observed against predicted at a plant the model never saw

l'Ampolla, 80 samples along the train. **Whole water, no chromatograph: four absorbances read on a laboratory bench-top spectrophotometer.** Top: A254 alone. Bottom: the four wavelengths. 10-fold cross-validation.

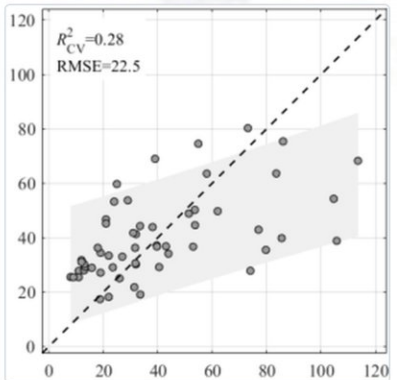
TCM · A254 alone R²CV 0.77



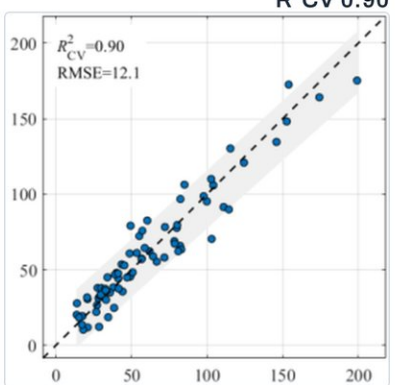
THM4 · A254 alone R²CV 0.75



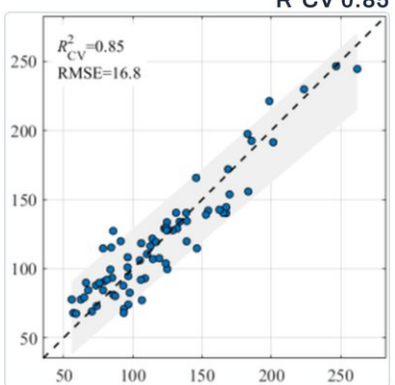
HAA9 · A254 alone R²CV 0.28



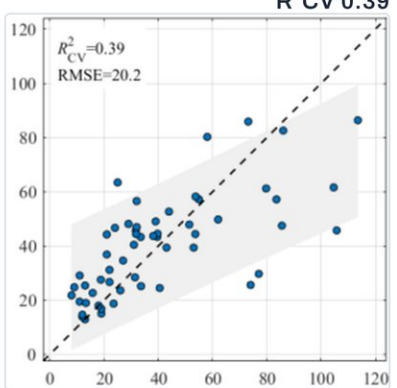
TCM · four wavelengths R²CV 0.90



THM4 · four wavelengths R²CV 0.85



HAA9 · four wavelengths R²CV 0.39



0.75 → 0.85 R²CV · THM4

chloroform 0.77 → 0.90

0.28 → 0.39 R²CV · HAA9

modest: part of the HAA precursors do not absorb UV

> **Why four and not one.** A254 reads how much. The four read what kind: 290 and 362 nm separate the bromine-reactive humics from the rest.

Chosen on the chromatograph, they work without it: whole water, a laboratory spectrophotometer, four numbers, a plant the model had never seen.

Valentí-Quiroga et al. (2026) Water Research X 31, 100545, Figure 6 · grey band: 90 % prediction interval.

REFERENCES

Where the numbers in this talk come from

1 **Mapping the UV response of natural organic matter to chlorination: from fraction-resolved fingerprints to operational monitoring of DBP formation.**
Water Research X 31, 100545 · doi 10.1016/j.wroa.2026.100545
Valentí-Quiroga M, Cabrera-Codony A, Daunis-i-Estadella P, Mulet M, Fargas-Marquès A, Monclús H, Martín MJ (2026)
four wavelengths · independent validation at l'Ampolla

2 **In-depth analysis of natural organic matter fractions in drinking water treatment performance: fate and role of humic substances in trihalomethanes formation potential.**
Science of the Total Environment 955, 176600 · doi 10.1016/j.scitotenv.2024.176600
Valentí-Quiroga M, Cabrera-Codony A, Emiliano P, Valero F, Monclús H, Martín MJ (2024)
who removes what · HS1/HS2/HS3

3 **NOM fractionation by HPSEC-DAD-OCD for predicting trihalomethane disinfection by-product formation potential in full-scale drinking water treatment plants.**
Water Research 227, 119314 · doi 10.1016/j.watres.2022.119314
Valentí-Quiroga M, Daunis-i-Estadella P, Emiliano P, Valero F, Martín MJ (2022)
the method

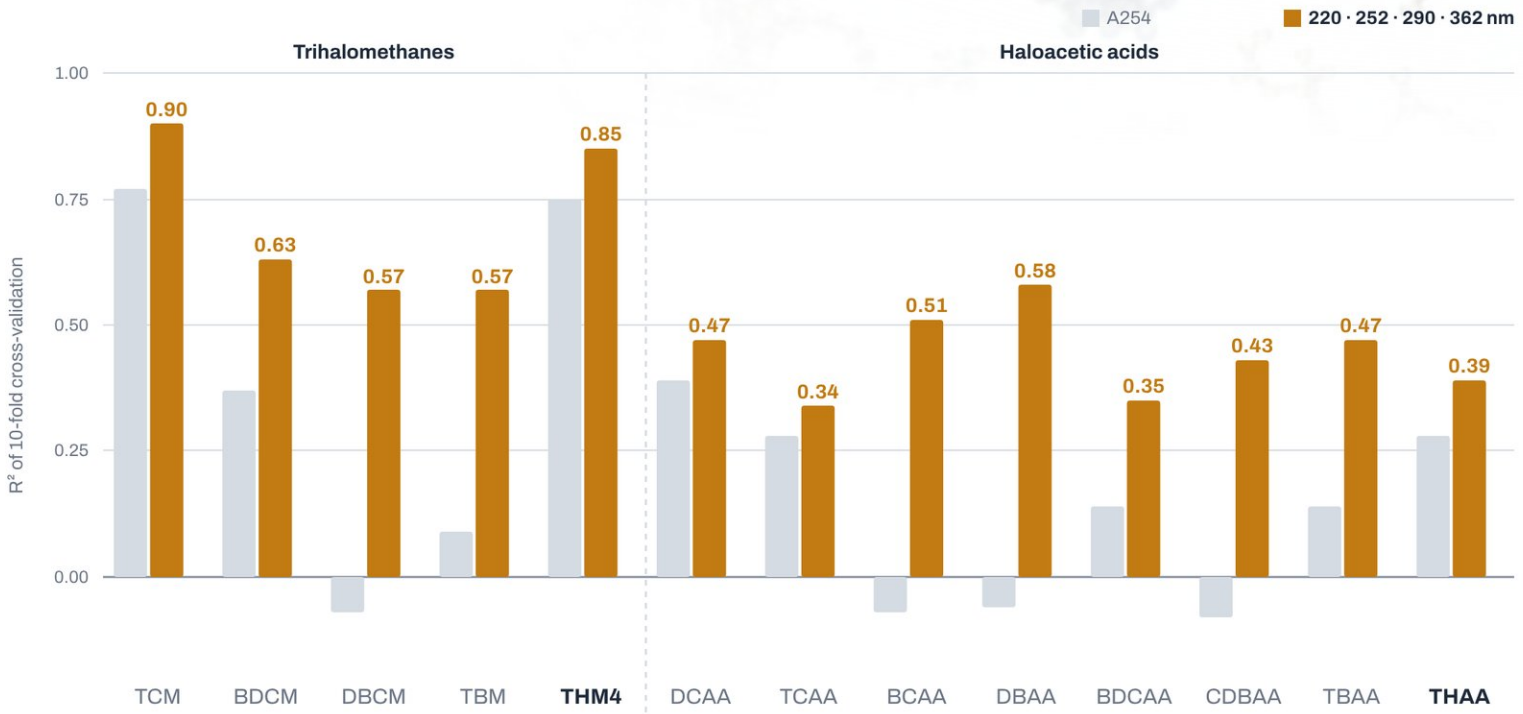
4 **Formation of disinfection by-products in drinking waters: a combination of advanced analytical techniques and modelization.**
PhD thesis, Universitat de Girona · Càtedra ATL drinking-water prize 2026 · hdl.handle.net/10803/692652
Valentí-Quiroga M (2024)
campaign data, Tables A6–A8, Table 21

5 **Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography – organic carbon detection – organic nitrogen detection (LC-OCD-OND).**
Water Research 45, 879–885 · doi 10.1016/j.watres.2010.09.023
Huber SA, Balz A, Abert M, Pronk W (2011)
LC-OCD and its fractions

Universitat de Girona · lequia UdG ECO-INNOVATIVE ENVIRONMENTAL SOLUTIONS

…and species by species: where one wavelength is blind

Same 80 samples, **whole water on a laboratory spectrophotometer**. Four wavelengths against A254, species by species.



0.85 R^2CV · THM4 · four λ

0.75 with A254 alone

0.47–0.58 brominated species

where one wavelength sees nothing (below 0.10)

Where one wavelength is blind, four see. In-line UV sensors already cover 200–400 nm: taking the four there is the step that follows.

Valentí-Quiroga et al. (2026) Water Research X 31, 100545, Table 2 · 10-fold cross-validation.

SEPARATE TO PREDICT · CLOSING

“The analytical technique is not the destination.
It is how we learn what the plant should pay attention to.”

Fractionate, and the DBPs become predictable. Look closer at the humics, and **four wavelengths** are enough, read on whole water with a laboratory spectrophotometer. That is the parameter the plan was missing, and what DrinkIA takes into operation.

WITH

Universitat de Girona · lequia UdG ECO-INNOVATIVE ENVIRONMENTAL SOLUTIONS · ATL Ens d'Abastament d'Aigua Ter-Llobregat · Consorci d'Aigües de Tarragona · Cicle de l'Aigua del Ter · Fisersa

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA E INNOVACIÓN

WATSON CTM2017-83598-R · ShERLOcK PID2020-112615RA-I00 · MORIARTY PID2024-155733OB-I00