In this response, the text is accompanied by new figures generated to explain in more depth our arguments. In addition, in some sentences we recall to the figures that appeared in our original manuscript. In order do not confuse the reader, figures that appear in this response are labelled with an asterisk (i.e. FigX*). Figures and tables from the submitted original manuscript keep their original order.

Referee #1 (S. Ishi)

The referee, in his/her report formulates three relevant comments and five minor technical suggestions.

In the **first comment**, the referee asks to clarify the following sentence "A basic difference with respect to these tools lies in the assumption that peaks fit a predefined probabilistic density function" (page 4713, line 20).

The objective of this sentence is to remark that the multiway tools used for EEMs signal deconvolution do not assume that peaks fit a predefined probabilistic function.

Successively the referee requests more details about the longitudinal gradient in DOM fluorescence properties of our data set.

Effectively, the data set come from samples collected along a river main stem. However, the EEMs do not exhibit gradual changes according to this continuum. On the contrary they present variable EEM features. This heterogeneity is a consequence of the severe point source anthropogenic inputs (waste water treatment plants, industrial effluents and some urban untreated waste water effluent). We executed water and solutes mass balance in this river and we estimated that, during basal discharge conditions, the point source anthropogenic inputs contributed to approximately the 30% of water mass and 45% of DOM mass (unpublished data from authors). For instance, figure 1* evidences that the DOM total fluorescence signal increased abruptly from 25 to 28 km and from 35 to 40 km. Inevitably, these inputs anthropogenic interrupt the biogeochemical continuum along the river main stem.

Finally the **last comment** refers on how the model output addresses fluorescence components with multiple peaks.

The results of post model data analysis (i.e. table 3 in our manuscript) suggests that some peaks might describe different maxima of the same fluorophore. For instance, P1 ($\lambda_{ex}/\lambda_{em}$= 272/319 nm) and P3 ($\lambda_{ex}/\lambda_{em}$=231/339 nm) fall within the protein-like region. Both peaks intensities strongly co-vary between them ($r=0.89$, $p<0.001$). This correlation suggests that these two peaks might describe a single fluorophore with two maxima. This behaviour is usual in pure fluorescence substances such as tryptophan. Similarly, the correlation between P4 ($\lambda_{ex}/\lambda_{em}$= 250/460) and P11 ($\lambda_{ex}/\lambda_{em}$= 307/471) might represent a similar case within the humic/fulvic acid region ($r=0.74$, $p<0.001$).
On the other side, there exist other correlations that link peaks located in different portions of the ex./em. plane. For instance P1 and P3 (potential dual fluorophore in the protein-like region) co-vary with P4 and P11 (potential dual fluorophore in the humic-like region). This result suggests that the same mechanism/processes might determine the co-occurrence of these peaks.

With respect to the minor suggestions, we will adopt all these suggestions and we will update the typographic errors detected.

All these three main comments and the five minor technical suggestions can be integrated in a revised version of our manuscript.

Referee #2 (Anonymous)

Referee #2 formulates several important comments. According his/her report the comments are ordered into six main points:

Comments 1, 2, and 4: make reference to model structure and functioning.

Comment 3 makes reference to the model output and data treatment;

Comments 5 and 6 refer to the discussion of data set.

We will address to all these comments. Some of these comments are strongly related between them; therefore in some cases our answers cover two or more of these points.

We start responding to the **comments 1, 2 and 4**:

First of all the referee suggests to test the assumption that fluorescent events can fit an asymmetric Gaussian distribution (hereafter referred to as AGD). In addition he/she suggests to use a simple mixture to make it easier to understand how the number and

position of peaks are determined. Namely, he/she suggests to use the EEM from quinine sulfate to describe this step (basically the "step A" in our text). We fully adopt this proposal to address these two fundamental points.

The AGD is a simple extension of the Gaussian distribution (hereafter referred to as GD). As mentioned in our text the idea that a fluorescence event follows a GD has been proposed previously by other researchers. For instance, the formula of the GD (in one dimension) appeared in Westerhoof et al. paper (2001). However, when have a look in more detail to a simple fluorescent mixture, the peak(s) frequently show asymmetry. For instance, in figure 1 in Boehme and Coble (2000), the quinine sulfate peak shows a subtle, but visible, asymmetry along the emission axes: at λ_{Em}>450 nm the fluorescence signal decreases slower than at λ_{em}<450 nm. This evidence suggested us the need to introduce an asymmetry factor into the GD adopted by Westerhoof. Under this perspective, the AGD proposed by Kato et al. (2002; this approach is implemented in data recognition pattern problems) fulfills our needs.

Therefore, according to the suggestion of referee #2, we can use the EEM of quinine sulfate (referred to as the QS sample) to describe in more detail the "step A" (determination of number a position of potential peaks) and to test the premise that peaks fit the AGD.

QS EEM shows two clear maxima at $\lambda_{ex}/\lambda_{em}$ ~ 245/450 nm and 345/450 nm respectively and a characteristic shoulder between these two peaks (Figure 2a*). In this spectra the identification of the two local maxima is straightforward. In parallel, the contour plot of the laplacian operator ($\nabla^2 f$, Figure 2b*) evidences the presence of three local minima: two of them coincide with the two local maxima; a third local minima is at $\lambda_{ex}/\lambda_{em}$ ~ 305/450 nm and it helps to locate the position of an additional subtle peak between the two maxima. Therefore in this example, according eq. 5, L_n is a list of three coordinates:

$$L_2=\{\{245,450\},\{345,450\},\{305,450\}\}$$

Figure 2c* shows the QS modelled spectra obtained assuming that peaks fit AGD. In this example the r^2 is 0.995. Figure 3a* shows the relationship between the observed and modelled fluorescence measured at each λ_{ex} and λ_{em} pairs (black points). Points are located around the 1:1 line (the slope of the fit is 0.995±0.0026) and magnitude of

residuals do not show a trend with respect to the magnitude of the fluorescence signal (see the inset in the bottom right corner).

The three coordinates detected with the QS laplacian operator are used to further test if the GD is a valid model to fit the QS EEM. Therefore we run the finite mixture model (eqs. 1 and 2 in the manuscript assuming that peaks follow the GD model (i.e r_i={1,1} in eq.2). Gray dots in Figure 3* show the fit between observed and modelled fluorescence. The fit is still reasonably good. However the slope of the fit decreased slightly to 0.94±0.0076 and the r^2 decreased to 0.96. Therefore the residual are clearly larger than in the previous case (see the inset in the top left corner). Figure 4* compares the outputs obtained with the two distribution approaches along the emission fluorescence spectra (λex=350 nm). The plot reveals the failure of the GD to fit reasonably well the data at λ_{em}<450 nm. This example highlights that, although the classical Gaussian distribution fit reasonably well the fluorescence phenomena of a simple fluorescent substance, the introduction of an asymmetry factor in the GD improve notably the model goodness.

In comment 4, the reviewer also asks what the residuals of the fit look like and has doubts about whether r^2 values are the "most appropriate or only way to assess the fit".

In our study we used two descriptors of goodness of fit: r^2 and the Bayesian Information Criterion (BIC, eq. 7).

We consider that a model output is reasonably good if r^2>0.99. This means that the model explains more than the 99% measured variance. This threshold is identical to that used by other researchers in other deconvolution tools (Stedmon and Bro. 2008). On the other hand, BIC is used to reduce the risk of overestimating the model parameters (i.e. the number of peaks). The model likelihood (i.e. r^2) typically increases with the increase of number of parameters (in our context it is the number of potential peaks). However this might induce to overfit the data set. Therefore it is necessary to find a compromise between the goodness of fit and the number of model parameters implemented in the fitting. This aspect is crucial. The BIC parameter is used in statistics to select the optimal model dimension (Schwarz, 1978). It takes into account both the model error variance and the number of model parameters. The model with the lower BIC value is the selected model though if that model does not fit the data set with the highest r^2. In our manuscript, Figure 1 is essentially devoted to explain this process (Step B). In this example (an EEM from a natural sample is used), the selected model is not that with the

highest r^2 (r^2=0.9962 and nine peaks), but is that with the lowest BIC (and a slightly lower r^2, 0.996, and eight peaks).

In any case we consider interesting the comment of the reviewer about residuals. Therefore, we adopt the referee's suggestion of illustrating how the fit between observed and modelled data (and the associated residuals) looks like. For instance, Figure 5* is an updated version of figure 3 in our manuscript. This revised figure besides comparing some modelled EEM with that observed, it shows the fit between the modelled and observed data and the residuals.

In **comment 3**, the reviewer integrates several questions related both to the model output and how the Voronoi approach used to partition the peaks into clusters works.

As mentioned in the text the finite mixture approach deconvolves individual EEM. Therefore the model output synthesizes the information of one EEM. This output consists of a table that includes the selected statistical descriptors of each single AGD (eq. 2 in the manuscript): mean μ_i (i.e. the peak position into the $\lambda_{ex}/\lambda_{em}$ plane), deviation (σ_i), height (a_i) and asymmetry (r_i). These values allow to calculate the volume of a specific peak and thus to estimate its contribution with respect to the total fluorescence of a specific EEM.

From that point, the implementation of the model is finished and data can be explored in different ways depending of: *a*) the number of EEMs involved in the analysis and; *b*) the specific objectives of each study. According to our manuscript, "FDM analyzes individual EEMs, but it is not designed to analyze differences among EEMs or dependencies among peaks within one EEM. Therefore successive data treatment is necessary to analyse these two essential aspects" (pag 4721, line 15). We consider important to remark that this successive data analysis is disconnected from the execution of the deconvolution model and it is not our target to describe the post-model data analysis techniques. However, we consider important to explain in more detail the approach used to explore the model output.

In our data set (21 EEMs), we can see that identified peaks (a total of 153) are not randomly distributed in the excitation-emission plane. In the manuscript we write that "data are grouped around eleven centers (from P1 to P11). Successively, we implemented the Voronoi diagram method to divide the plane into 11 sub-regions (see figure 7a*).

The Voronoi diagram is a method used to divide a surface into "n" polygons frequently used in data image analysis. Each polygon contains one center. Boundary of each polygon goes through the middle of a segment that joins two adjacent centers. These segments are obtained with the Delaunay triangulation algorithm. All points that lie within a polygon are assigned to the centers of that polygon. Both the Voronoi diagram and the Delaunay triangulation are usually implemented in the mathematical software.

In our dataset, the assignation of the eleven centers is simply based on a visual inspection of the distribution o points in the excitation-emission plane. This assignation becomes more intelligible if we convert the scatter plot (Fig. 6a*) into an array to generate a contour plot (Fig. 6b*). A visual inspection of this plot allows to discern, quite clearly, ten clusters (labelled from "1" to "10" in Fig. 6b*). Contour lines also help to assign a center for each cluster. In our manuscript we further split the cluster "4" (Fig. 6b*) into two clusters (P4 and P5 respectively, Fig. 7a*). In fact in the region located at λ_{ex}<250 mn and 420<λ_{em}<500 nm the contour plot might suggest the presence of two peaks close to each other.

Obviously, the Voronoi diagram, as any other clustering approach, has some inherent limits. In this case, it is the user experience that determines the number and position of centers. Therefore we can not exclude some subjectivity in our decision to individuate eleven clusters.

This subjectivity can generate benefits and shortcomings. For instance, the decision to split the points enclosed into the region λ_{ex}<250 mn and 420<λ_{em}<500 nm into two peaks (P4 and P5, Fig. 7a*) appears arbitrary because it exists only a slight separation between these two groups. However, this decision helps to evidence that downriver, the peaks that appear in this portion of the excitation-emission plane, shift toward larger emission wave length. This shift might suggest changes in DOM molecular weight along the river system (from relatively small compounds with low concentration to larger and more concentrated ones - Ishii and Boyer, 2012). On the other hand, an example of shortcoming might be considered analyzing points enclosed into the cluster P2 (centered at $\lambda_{ex}/\lambda_{em}$ =290/356nm, protein-like, Fig. 7a*). In this cluster a sub-peak at $\lambda_{ex}/\lambda_{em}$ ~300/370nm is detectable (see the red asterisk in Fig. 7a*). Therefore, if points that generate this sub-peak are integrated into P2, we can not relate those data to some specific samples and we might be missing some interesting information. In this case, if

the researcher is interested to identify exactly these samples, he has the option to create a new peak *ad hoc*, centered at $\lambda_{ex}/\lambda_{em}$ =300/370nm.

Alternatively, hierarchical clustering approach (HCA) can help to identify automatically (and more objectively) the number of clusters and data classification. Figure 7b* shows the result of this clustering technique: nine clusters emerged. To obtain this clustering configuration we implemented the Euclidean distance function and the "agglomerative" function for clustering. The "silhouette" test is used for identifying the optimal number of clusters (Rousseeuw, 1987). Eight clusters out of nine coincided with those clusters identified visually from the Voronoi diagram. The difference lie in the cluster located in the portion λ_{ex}<275 mn and 420<λ_{em} <470 nm, the red rectangle in Fig. 7b*). In this sub-region, we visually identified three clusters (P4, P5 and P7, the red rectangle Fig.7a*). The shortcoming of the HCA approach consists on the fact that in some EEMs appear two peaks in the same cluster. For instance, in sample S8 (Fig. 5*, or in more detail, Fig. 8*) appear clearly two peaks centered at $\lambda_{ex}/\lambda_{em}$~270/435 nm and $\lambda_{ex}/\lambda_{em}$ ~240/430 nm respectively. If we adopt, in an uncritical manner, the clustering obtained with the HCA, these two peaks should coincide frequently and treated as a unique peak with two maxima. Nevertheless these two peaks coincided in few EEMs only (namely, S7, S8, S9, S10, see Fig.5c). Therefore this cluster might be appropriate in a statistical perspective but might be an artifact in a biogeochemical context. The peak at $\lambda_{ex}/\lambda_{em}$~270/435 nm (i.e. P7) has been detected in wastewater in another study (Saadi et al., 2006) and it appears to have a different origin to that detected at $\lambda_{ex}/\lambda_{em}$ ~240/430 nm (of humic/fulvic origin, i.e. P5). In consequence, if we use information from literature, these two peaks should be treated separately. Then, it seems more reasonable to integrate the peak $\lambda_{ex}/\lambda_{em}$~270/435 into a new cluster centered at $\lambda_{ex}/\lambda_{em}$~269/433 nm (P7 according our visual clustering in Fig. 7a* and table 2).

In summary, a simple visual inspection of peaks coordinates in the excitation-emission plane allows individuating a minimum of ten clusters. The contour plot of the excitation-emission plane allows locating the center of each cluster (Fig. 6b*). By contrast, the hierarchical clustering approach individuates nine clusters (Fig. 7b*). Eight clusters out of nine coincided with those we identified visually suggesting that the visual clustering is reasonably accurate. Furthermore, if we integrate information from literature and we plan to gain details, one of the clusters individuated with HCA

(namely, cluster 4) is further split into three clusters (P4, P5 and P7) to obtain a total number of eleven centers. According the position of the centers, the Voronoi diagram is implemented to delimit the eleven clusters. This clustering approach limits the problem to find, in a specific EEM, two peaks in the same cluster.

Comments 5 and 6 focus on the discussion of data set.

Namely, the reviewer suggests comparing modelled EEMs with those observed. Figure 3 in the manuscript is devoted to this comparison. Alternatively, figure 5* can further improve this comparison because it incorporates the fit between observed and modelled EEM and associated residuals.

In the last comment, the reviewer considered our discussion of the results superficial. We partially agree with this comment. We remark that our text is a "technical note". According the directive of Biogeosciences, a "technical note" "report new developments, significant advances and novel aspects of experimental and theoretical methods and techniques". In consequence this manuscript should not be judged as a "research paper". Furthermore a technical note "should be short (a few pages only)." Therefore, in our manuscript, the data set is simply used to show how the model fits a reduced set of EEMs and to provide a general view how outputs look like. Inevitably results are described concisely. It is not our objective to provide an exhaustive biogeochemical interpretation of our results. To analyze appropriately these results we should to add new information/parameters that inevitably might enlarge our text, and moves the reader towards a different story.

Beside this clarification, we consider that we can enrich our data discussion in the context of the model introducing only the information of dissolved organic carbon concentration (DOC) measured in water samples.

We observe that DOC concentration and DOM total fluorescence strongly co-vary through the river continuum (r^2=0.695, d.f.=20, $p<0.001$, Fig. 9a*). This relationship suggests that fluorescent DOM should be a relevant component of total DOM. Therefore the question that might arise at this step is: "What portion of the EEMs co-vary more significantly with DOC?". To address to this question we firstly estimate the volume of each peak indentified by deconvolution in each EEM (eq. 2 is used to calculate the volume of each peak). Successively, we execute a step-wise linear multiple regression to extract those peaks that more significantly co-vary with DOC. This analysis reveals that DOC strongly co-vary with six peaks: P2, P5, P8, P9, P10 and P11

(r^2=0.96, d.f.=15, p<0.001). Peaks P8, P9 (humic-like, p<0.00057 and p<0.0006 respectively) and P2 (protein-like, p<0.006) are the fluorescent events which are more significantly related to DOC concentration. Additionally, the analysis shows that the emergence of P8 is associated with the beginning of the increase of DOC from 20 km. Successively, P2 and P10 are the two peaks that acquired more relevance in coincidence of the higher DOC concentrations (from 25 to 36 km Fig. 8b*). Note that this analysis implements the volume of the fluorescent events obtained with the model and could not be executed with the traditional peak-picking.

References

Boheme, J.R. Coble, P.G.: Characterization of coloured dissolved organic matter using high energy laser fragmentation, Environ. Sc. Technol., 34, 3283-3290, 2000.

Kato ,T., Omachi , S. Aso, H.: Asymmetric Gaussian and Its Application to Pattern Recognition In Structural, Syntactic, and Statistical Pattern Recognition. Lecture Notes in Computer Science 10.1007/3-540-70659-3 T.Caelli et al. (Eds.): SSPR&SPR, LNCS 2396, pp 405-413, 2002.

Ishii S.K. and Boyer T.H.: Behavoir or reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineering systems: a critical review, Environ. Sc. Technol., *46*, 2006-2017, 2012.

Saadi I., Borisover, M., Armon R., Laor Y.: Monitoring of effluent DOM biodegradation using fluorescence, UV and DOC measurements, Chemosphere, 63, 530–539, 2006.

Stedmon C.A.,, Markager, S., Bro, R. 2003. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. Marine Chemistry 82: 239– 254

Stedmon, C.A.; Bro, R.: Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnol. Oceanogr-Meth., 6, 572-579, 2008.

Schwarz G.: Estimating the dimension of a model. Ann. Stat., 6, 461-464, 1978.

Rousseeuw, P.J., Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20, 53–65, 1987.

Westerhoff, P. Chen, W. Esparza, M.: Fluorescence Analysis of a Standard Fulvic Acid and Tertiary Treated Wastewater, J. Environ.. Qual., 30: 2037–2046, 2001.

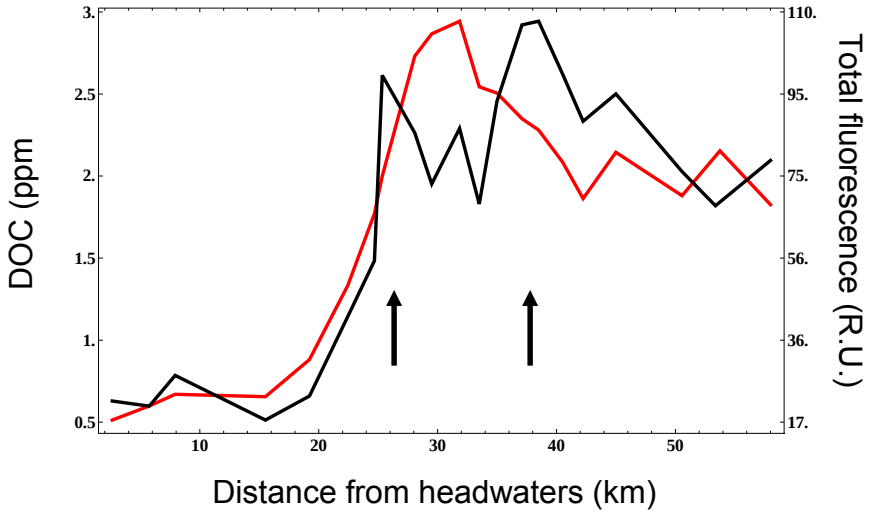

Figure1* DOC (red line) and DOM total fluorescence (black line) along the Tordera river main stem. Arrows highlight abrupt changes in total fluorescence.

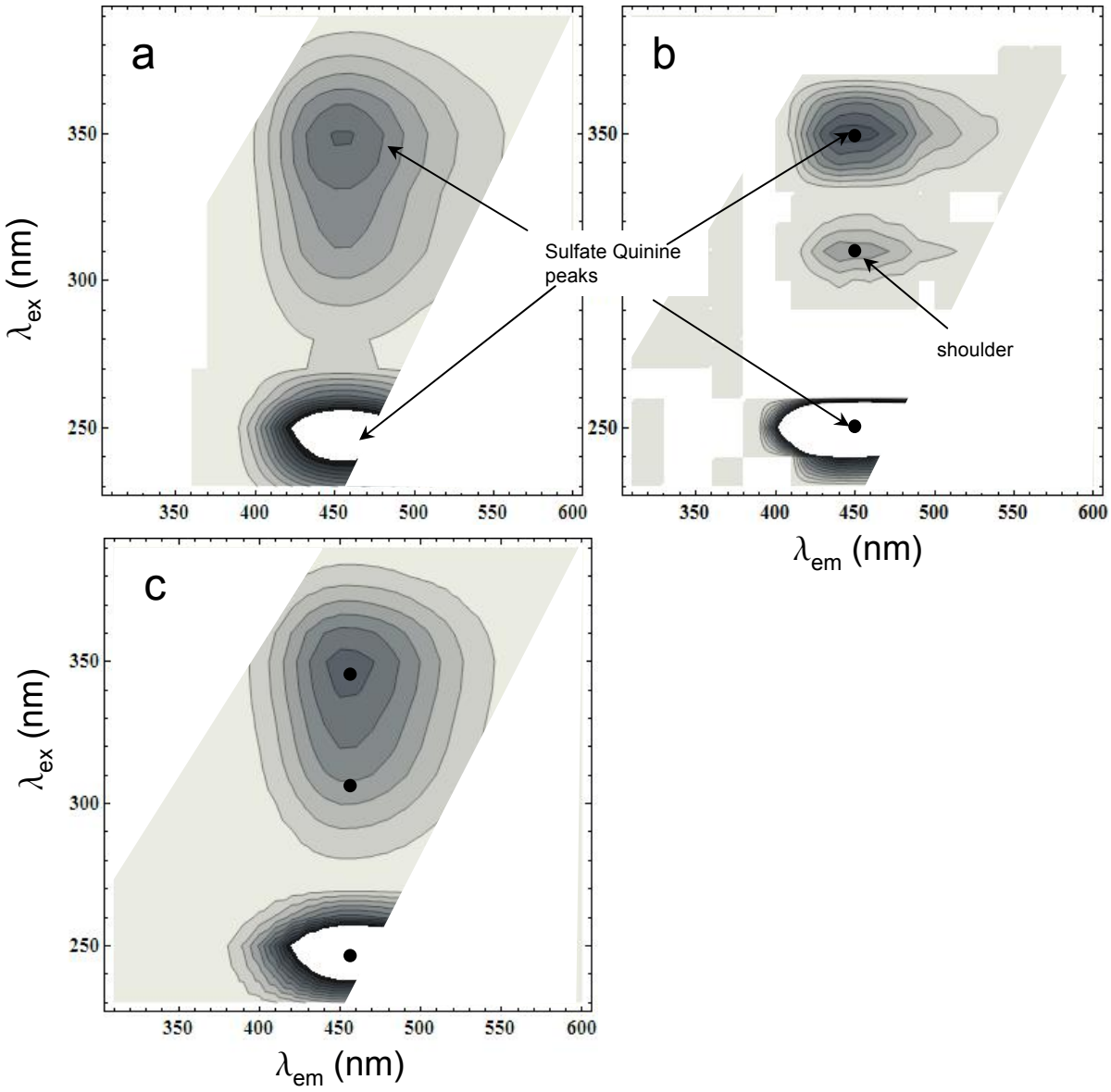

Figure2* Contour plot for quinine sulfate (QS) EEM (a), its laplacian $\nabla^2 f$, (b) and the modelled EEM (c).

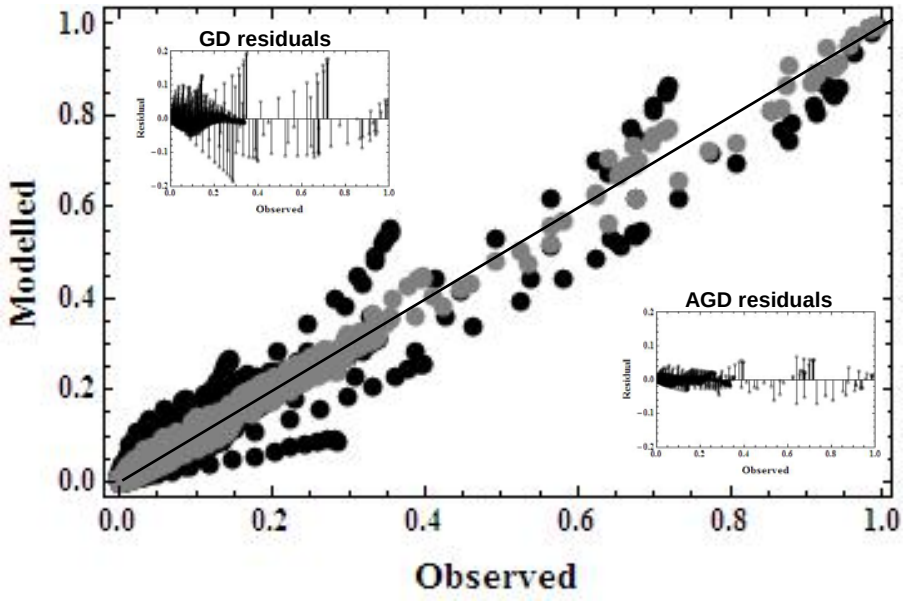

Figure 3* Comparison of the modeled QS EEM obtained with the FDM model (eqs. 1 and 2) with the real EEM. The plot shows the relationship between each individual fluorescence pairs that compose the observed and modeled EEM. Black and gray points show the relationship obtained assuming the asymmetric Gaussian distribution or the classical normal Gaussian distribution (i.e. forcing r_i=1 in eq. 2) respectively. The goodness of the two fit are: r^2= 0.995 (asymmetric Gaussian) and r^2=0.96 (normal Gaussian). The solid line shows the 1:1 line. The two insets show the residuals with respect to the observed fluorescence signal.

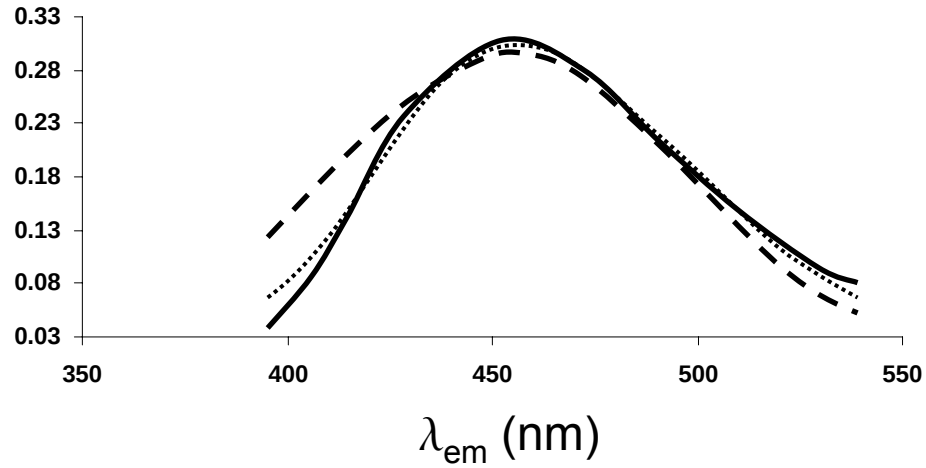

Figure 4* Emission fluorescence spectra (λex=350 nm) for the QS sample (solid line) and model output with the AGD (dotted line) and normal GD (dashed line)

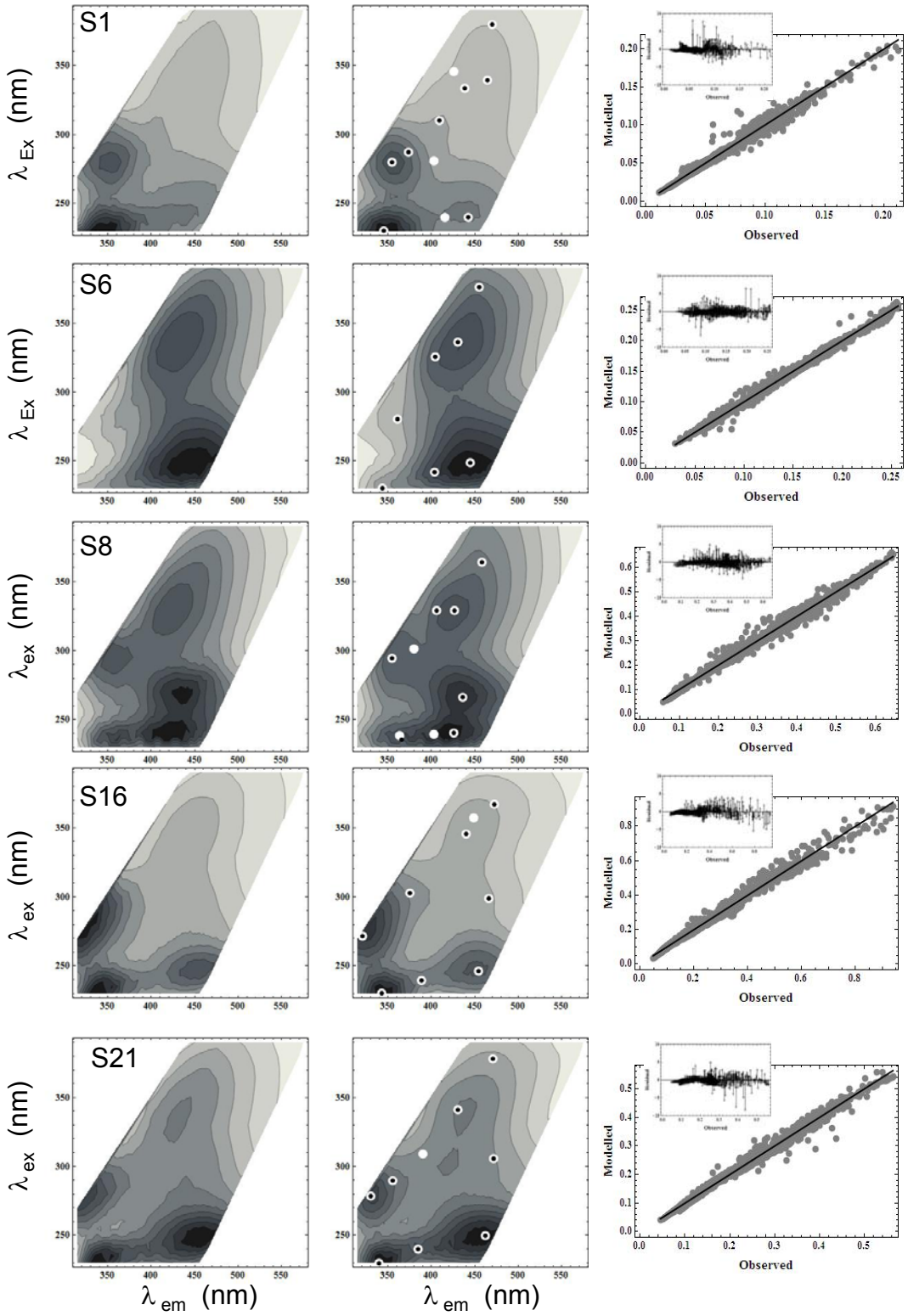

Figure 5*. Examples of observed (left) and modelled (center) EEMs from la Tordera river from five sampling sites. Contours indicate the 10% of the intensity peak. Large white and small black dots in contours plots show location of potential and selected peaks respectively. Scatter plots on the right show the fit between modelled and observed EEMs. Each dot represents a single fluorescence pair that compose the observed and modelled EEM. The inset shows the plot of standardized residuals respect to the observed data.

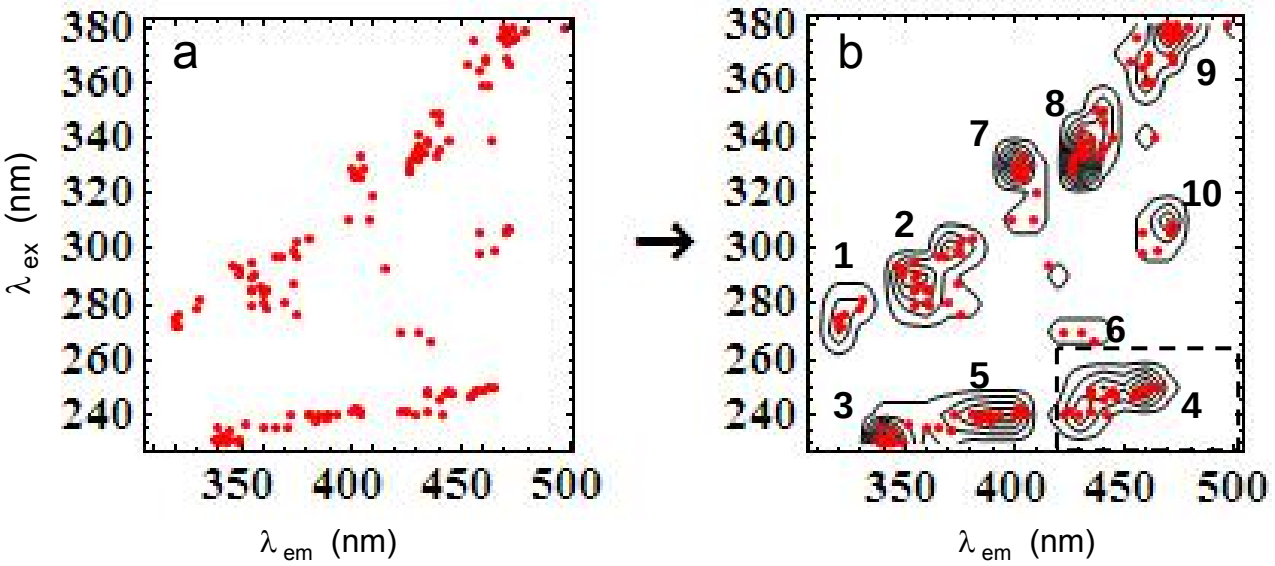

Figure 6*. a) Distribution of peaks (red dots) within the excitation vs. emission plane identified in the data set after the deconvolution; b) Contour plot generated by converting the two-dimensional data into an array. Contours help to visualize a discrete number of clusters and position of the centers of each cluster. In this case, ten cluster emerged from the contour plot. Dashed rectangle located at λ_{ex}<250 mn and 420<λ_{em}<500 nm highlight the cluster 4 that has been further divided into two clusters close to each other (P4 and P5, see figure 6a*).

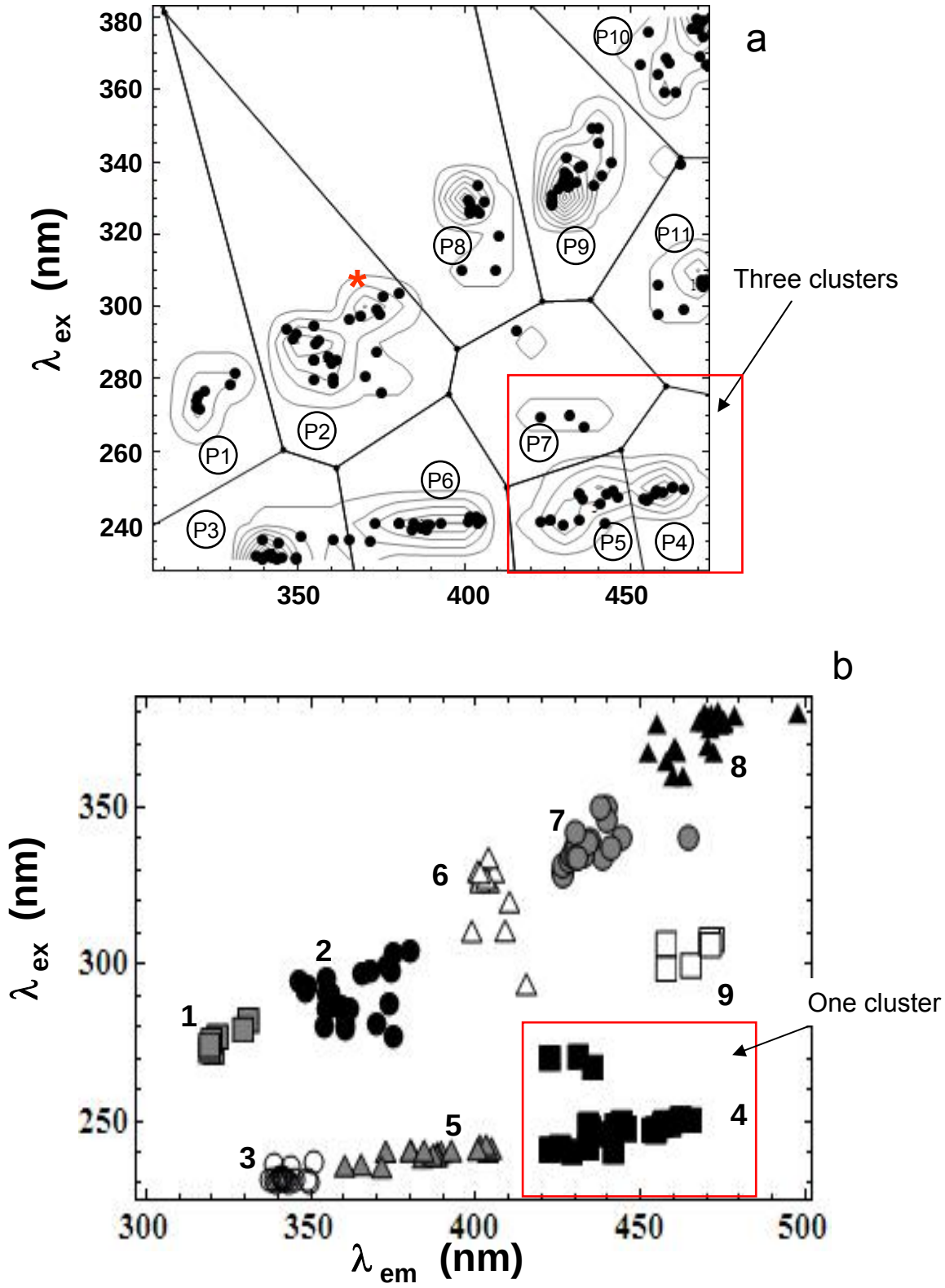

Figure 7*. a) Clusters identified in the dataset according to a hierarchical clustering approach based on Euclidean distance and agglomerate dissimilarity function. The approach identify nine clusters. Data points that integrate a cluster have the same marker; b) Clusters identified after the identification of eleven centers based on a visual inspection as explained in the manuscript. The red rectangles in the two plots highlight the portion of the ex-em plane with different clustering results. The red asterisk (*) indicates the position of a potential undetected peak.

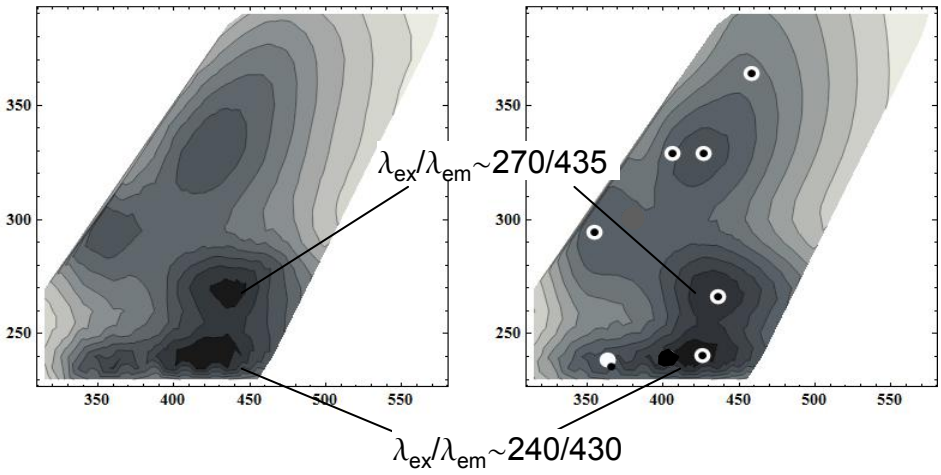

Figure 8*. Observed (left) and modelled (right) EEM of sample S8. Arrows highlight the position of two peaks that in the manuscript are labeled as “P5” and “P7”. These two peaks coincided in samples S7, S9 and S10 but not in the other EEMs.

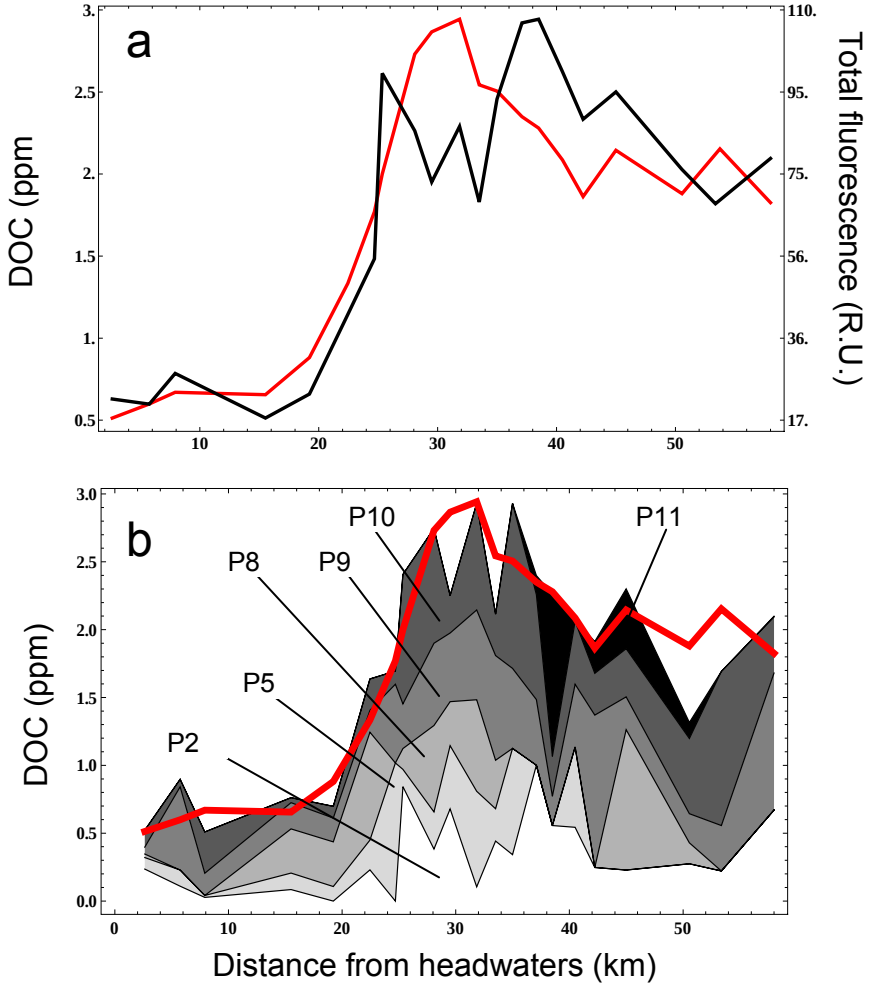

Figure 9*. a) DOC (red line) and DOM total fluorescence (black line) along the Tordera river main stem. DOC and total fluorescence significantly co-vary (r^2=0.695. d.f.=20, p<0.001); b) Comparison of the observed DOC spatial dynamic (red line) with that modelled with six peaks (P2, P5, P8, P9, P10, P11) according the step-wise linear multiple regression (r2=0.969, d.f.=14, p<0.001). Different gray tonality depicts the contribution of each peak. See text for more details.