

First dynamic model of dissolved organic carbon derived directly from high frequency observations through contiguous storms

Supporting Information

Timothy D Jones, Nick A Chappell, and Wlodek Tych*

Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK

*corresponding author, email: n.chappell@lancaster.ac.uk; tel: +44 1524593933; fax: +44 1524843087

Content:

Number of pages: 34

Number of sections: 4

Number of figures: 20

Number of tables: 10

Number of references: 14

■ EXPERIMENTAL (MATERIALS AND METHODS)

Section S1

Exact details of the essential infrastructure designed for each spectrolyser. Each spectrolyser consists of a xenon flash bulb and a 256 pixel array detector that measures absorbance between 200 and 732 nm. The units were all fitted with a 35 mm path-length measurement window to cover an expected range of approximately 1-10 mg/L DOC. A bespoke protective housing was developed for each spectrolyser that consisted of a vertical section of black uPVC pipe containing the interface cable and a horizontally mounted section below the minimum water-level housing the spectrolyser; the latter was installed approximately 0.20 m above the streambed. The whole assembly was secured to a galvanised steel structure. Two 86 Ah sealed leisure batteries connected in series were used to ensure a continuous power supply guaranteed in excess of 12.0 V DC; these were exchanged for fully recharged units on a weekly basis.

Section S2

Findings of the cleaning protocols developed to deliver continuous high frequency observations of DOC of a high data quality. Initially, all four spectrolysers at Brienne were cleaned manually once per week by brushing for 60 seconds with a small brush soaked with 10% hydrochloric acid (HCl). This cleaning produced step reductions of the $\text{DOC}_{\text{RIVCOL}}$ of greater than 1.5 mg/L in the algae-rich LI8 stream. The cleaning frequency was increased to twice per week to give step reductions of less than 0.7 mg/L $\text{DOC}_{\text{RIVCOL}}$ at LI8 (and all other sites). This small step was then removed from the time-series (using a routine in MATLAB) by assuming that the algal presence in the spectrolyser measurement windows responsible for the fouling had developed as an exponential function of time over the 3-4 days prior to cleaning.⁵ A subsequent 2-week trial with compressed air cleaning, namely a 4 second pulse at 5 bar triggered one minute before each 15-minute reading, failed to produce a noticeable

reduction in the small step change resulting from twice weekly cleaning. Consequently compressed air cleaning was not continued beyond the trial period. Checks on the spectrolyser readings using a special sleeve filled with deionised water demonstrated that the weekly brushing with 10% HCl was able to remove all fouling from the measurement windows at sites LI3 and LI8. This cleaning proved to be insufficient at site LI6 and LI7, possibly due to manganese precipitation.⁶ The gradual upward drift (< 0.1 mg/L/week) in LI6 and LI7 readings was mitigated by soaking the measurement window in 10% HCl for 5 minutes every week. During a few storm periods the spectrolysers were completely inundated by deposited bedload plus a few incidents of power supply failure occurred. Power failures were largely avoided by use of 24 volt DC rather than the 12 volt DC power supplies used at many other installations e.g., Grayson and Holden.⁷ The spectrolyser readings for the few periods of inundation, battery failure or sensor cleaning (e.g. 2.2% of the whole monitoring period at LI8) were replaced with a missing data flag (namely a NaN or 'Not-a-Number' value), as the RIVC algorithm is able to identify models with missing data flags in the output (here DOC concentration or load) time-series.

Section S3

Rationale for use of RIVC algorithm for model identification. The RIVC algorithm (*Refined Instrumental Variable Continuous-time Box-Jenkins identification algorithm*)¹ was considered suitable to identify new models of the DOC load observations at Llyn Brienne for several key reasons:

- 1/ In one numerical experiment RIVC can identify the structure and parameters of a range of purely static models and models that include one or more dynamic components,
- 2/ Dynamic Response Characteristics (DRCs) that have a physical interpretation^{2,3} can be derived from the identified RIVC model parameters,
- 3/ The covariance matrix for each identified model can be used directly to estimate uncertainty in the model parameter estimates; this is critical when differences in the derived DRCs between periods, sites or models are to be interpreted,

4/ An optimal model structure can be selected using a combination of this uncertainty information, heuristic measures of the parsimony (to avoid over-parameterisation) and a feasible physical interpretation. This is the Data-Based Mechanistic or DBM approach to modelling.

5/ Model estimation involves iterative pre-filtering of the signals to remove the high frequency noise inherent within environmental data (even within quality assured data) that affects identification of accurate parameter values. These pre-filters include the current (at each iteration) estimate of the dominant dynamic modes of the process, thus focusing the signal within the spectral range critical for the analysis.

6/ The routine can estimate transfer functions in continuous-time (CT-TF) that maybe more difficult to estimate compared to discrete-time TF models but are considered more accurate for systems with responses almost as fast as the monitoring time-step,⁴ and

7/ RIVC is computationally efficient, with its results permitting rapid Monte Carlo uncertainty analysis of the parameter ranges. The Instrumental Variable (IV) approach guarantees that the parameter estimates are asymptotically unbiased. In this technique the IVs are generated as the current model output given the input data (control variables).

Section S4

Identified dynamic models of DOC load in relation to hydrological response. The measure of simulation efficiency used to evaluate the RIVC models was the R_t^2 and is given as:

$$R_t^2 = \frac{\sigma_{error}^2}{\sigma_{obs}^2} \quad (\text{SI Equation 1})$$

where σ_{error}^2 is the variance in the model residuals and σ_{obs}^2 the variance in the observed data. This measure of simulation efficiency is equivalent to the simplified form of the Nash and Sutcliffe statistic and so differs from the R^2 that uses the variance in the model output as the denominator. Because ever more complicated models tend to have higher simulation efficiencies, but a greater degree of parametric uncertainty, identification of the best model needs to balance these two statistical properties. This was achieved by calculation and use of the Young Information Criterion:⁸

$$YIC = \log_e \frac{\sigma_{\text{error}}^2}{\sigma_{\text{obs}}^2} + \log_e \{NEVN\} \quad (\text{SI Equation 2})$$

The first term within SI Equation 2 is a measure of the model efficiency (SI Equation 1) while the second term is the normalised error variance norm (*NEVN*):

$$NEVN = \frac{\sigma^2}{np} \sum_{i=1}^{i=np} \frac{P_{ii}}{\hat{\theta}_i^2} \quad (\text{SI Equation 3})$$

which is a measure of the degree of over-parameterisation ($np = n + m$ is the number of estimated parameters in the θ vector; $\sigma^2 P_{ii}$ is an estimate of the variance of the estimated uncertainty on the i th parameter estimate; and $\hat{\theta}_i^2$ is the square of the i th parameter in the θ vector) that results in uncertainty in the derived model parameters.⁸ The *YIC* is a measure of whether the model has become overly complex (because of too many parameters or types of input variable) given the amount of information contained within the observed data-series. A change of +1.0 or greater in *YIC* as model order is increased may indicate that the model structure may have become too complex for the information contained in the time-series, i.e. over-parameterised,^{2,9} and the simpler model should be accepted as the optimal order.

Within this model development study, the procedure for identifying the optimal model structure using these two measures began by identifying the zero and first-order models (from those convergent models) with the highest R_t^2 that did not have complex roots (see Box et al.¹⁰). Where second-order models had a higher R_t^2 than that of the optimal first-order (or zero-order static) models, these models were then examined for a change in the magnitude of the *YIC* from first to second-order models. If a second-order model: (1) had a higher R_t^2 than the optimal first-order model, (2) its *YIC* was less than +1.0 different to the first-order model, (3) it did not have complex roots, (4) it did not exhibit oscillatory behaviour in the impulse response function (see Box et al.,¹⁰), and (5) the sign (+/-) of the two identified roots were the same, then this higher-order model was accepted as the optimal structure. This procedure was repeated for third order models, then fourth and so on and the resultant optimal models for both the February and May periods are presented in Table 1 (research article). The resultant models of parallel rainfall to

streamflow response for the same streams and periods are also presented in Table 1 (research article). To allow comparison of the model parameters between the different periods and basins, the uncertainty in the model parameters (Table 1 research article) was assessed using 1000 Monte Carlo realisations of the optimal parameter sets combined with uncertainty information identified by the RIVC routine. As a further illustration of simulation capability of optimal models, the simulated data are presented with the observed data for all streams and both simulation periods in SI Figure S18-S19.

SI Table S1. Characteristics of the monitored experimental basins near the Llyn Brianne reservoir in upland Wales, United Kingdom

stream name	Nant y Craflwyn	Nant Esgair Garn	Nant Rhesfa	Nant Trawsnant
basin name	LI3	LI6	LI7	LI8
land cover	coniferous plantation	improved moorland	improved moorland	coniferous plantation
area (km ²) ¹	0.76	0.57	0.49	1.23
catchment slope (m m ⁻¹) ²	0.107	0.122	0.221	0.109
podzol coverage ⁴	47%	45%	33%	44%
histosol/gleysol coverage ⁴	53%	55%	67%	57%

¹ 3-dimensional basin area upstream of the flumes, obtained using ArcGIS; ² from Weatherley & Ormerod (1987)¹¹; ³ from Littlewood (1989)¹²; ⁴ adapted from Reynolds and Norris (1990)¹³ and presented as FAO-Unesco (1990)¹⁴ soil units

SI Table S2. The most efficient twenty rainfall-DOC load CT-TF models for the LI3 basin for 5th to 18th February 2013 period, arranged in descending order according to their R_t^2 (efficiency measure term). The term *den* is the number of transfer function denominators (recession or α coefficients), *num* are the number of transfer function numerators (gain or β coefficients), *delay* is the pure time delay between rainfall and runoff response, and *YIC* is the Young Information Criterion. Terms *BIC* (Bayesian Information Criterion), *S2* the model residual variance and *condP* is the condition coefficient of the parameters covariance matrix; all are included here for reference. The optimal model is highlighted in bold

den	num	Delay	YIC	R_t^2	BIC	S2	condP
2	2	3	-5.597	0.9043	-7777	0.0016	0.9037
2	2	2	-5.379	0.9031	-7777	0.0017	0.9031
2	2	4	-5.743	0.9020	-7748	0.0017	0.9020
2	2	1	-5.092	0.9004	-7749	0.0017	0.9004
2	2	5	-5.824	0.8977	-7688	0.0017	0.8977
2	2	0	-4.744	0.8956	-7699	0.0018	0.8956
2	2	6	-5.850	0.8909	-7603	0.0019	0.8909
2	2	7	-5.828	0.8815	-7495	0.0020	0.8815
1	1	2	-9.609	0.8802	-7532	0.0020	0.8802
1	1	1	-9.622	0.8800	-7536	0.0021	0.8800
1	2	2	-5.043	0.8795	-7517	0.0021	0.8795
1	2	1	-2.595	0.8795	-7524	0.0021	0.8795
2	1	0	-2.697	0.8793	-7529	0.0021	0.8793
1	2	3	-6.100	0.8787	-7502	0.0021	0.8787
1	2	0	-3.143	0.8787	-7523	0.0021	0.8787
1	1	3	-9.557	0.8782	-7504	0.0021	0.8782
1	1	0	-9.597	0.8776	-7519	0.0021	0.8776
1	2	4	-6.7340	0.8768	-7476	0.0021	0.8768
1	2	5	-7.151	0.8738	-7439	0.0022	0.8738
1	1	4	-9.464	0.8737	-7453	0.0022	0.8737

SI Table S3. The most efficient twenty rainfall-DOC load CT-TF models for the LI3 basin for the 26th May to 5th June 2013 period, arranged in descending order according to their R_t^2 (efficiency measure term).

den	num	delay	YIC	R_t^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>6</u>	<u>-9.5426</u>	<u>0.9766</u>	<u>-7444</u>	<u>0.00016</u>	<u>0.9766</u>
2	2	7	-9.6046	0.9761	-7421	0.00016	0.9761
2	2	5	-9.2312	0.9743	-7371	0.00017	0.9743
2	2	8	-9.4418	0.9730	-7309	0.00018	0.9730
2	2	4	-8.7262	0.9692	-7223	0.00021	0.9692
2	2	9	-9.1420	0.9675	-7143	0.00022	0.9675
2	2	3	-8.1161	0.9614	-7036	0.00026	0.9614
2	2	10	-8.7484	0.9599	-6956	0.00027	0.9599
2	2	2	-7.4720	0.9511	-6840	0.00033	0.9511
2	2	1	-6.8231	0.9386	-6650	0.00042	0.9386
2	2	0	-6.1837	0.9240	-6474	0.00052	0.9240
1	1	4	-9.6267	0.8951	-6183	0.00071	0.8951
1	1	5	-9.6201	0.8948	-6174	0.00072	0.8948
1	2	10	-8.5371	0.8945	-6131	0.00072	0.8945
1	2	0	-7.0658	0.8940	-6195	0.00072	0.8940
1	2	1	-6.3957	0.8939	-6187	0.00072	0.8939
1	2	2	-5.1406	0.8931	-6174	0.00073	0.8931
1	2	9	-8.3813	0.8925	-6122	0.00073	0.8925
1	2	3	2.1347	0.8919	-6158	0.00073	0.8919
1	1	3	-9.5702	0.8919	-6164	0.00074	0.8919

SI Table S4. The most efficient twenty rainfall-DOC load CT-TF models for the LI6 basin for the 5th to 18th February 2013 period, arranged in descending order according to their R_t^2 (efficiency measure term).

<i>den</i>	<i>num</i>	<i>delay</i>	<i>YIC</i>	R_t^2	<i>BIC</i>	<i>S2</i>	<i>condP</i>
<u>2</u>	<u>2</u>	<u>1</u>	<u>-4.1951</u>	<u>0.8382</u>	<u>-6831</u>	<u>0.0036</u>	<u>0.8382</u>
2	2	0	-3.9160	0.8377	-6834	0.0036	0.8377
2	2	2	-4.4167	0.8361	-6808	0.0037	0.8361
2	2	3	-4.5796	0.8313	-6766	0.0038	0.8313
2	2	4	-4.6618	0.8241	-6707	0.0039	0.8241
1	1	0	-8.4026	0.8232	-6744	0.0039	0.8232
1	2	0	-1.8493	0.8231	-6736	0.0039	0.8231
1	2	1	-4.2663	0.8224	-6724	0.0040	0.8224
1	1	1	-8.3533	0.8216	-6726	0.0040	0.8216
1	2	2	-5.2818	0.8203	-6703	0.0040	0.8203
1	1	2	-8.2667	0.8172	-6689	0.0041	0.8172
1	2	3	-5.8669	0.8165	-6670	0.0041	0.8165
2	2	5	-4.6044	0.8151	-6640	0.0041	0.8151
1	2	4	-6.2117	0.8107	-6625	0.0042	0.8107
1	1	3	-8.1475	0.8097	-6633	0.0042	0.8097
2	2	6	-4.4900	0.8042	-6562	0.0044	0.8042
1	2	5	-6.3934	0.8026	-6567	0.0044	0.8026
1	1	4	-7.9950	0.7991	-6559	0.0045	0.7991
1	2	6	-6.4549	0.7918	-6495	0.0046	0.7918
2	2	7	-4.3622	0.7906	-6473	0.0047	0.7905

SI Table S5. The most efficient twenty rainfall-DOC load CT-TF models for the LI6 basin for the 26th May to 5th June 2013 period, arranged in descending order according to their R_i^2 (efficiency measure term).

den	num	delay	YIC	R_i^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>9</u>	<u>-8.5043</u>	<u>0.953156</u>	<u>-10889</u>	<u>0.00031</u>	<u>0.9532</u>
2	2	8	-8.3967	0.951931	-10861	0.00032	0.9519
2	2	10	-8.4597	0.951415	-10832	0.00032	0.9514
2	2	7	-8.1402	0.947516	-10749	0.00035	0.9475
2	2	6	-7.7567	0.93984	-10571	0.00040	0.9398
2	2	5	-7.3065	0.928993	-10352	0.00047	0.9290
1	1	8	-10.1254	0.916168	-10120	0.00055	0.9162
2	2	4	-6.8298	0.915201	-10118	0.00056	0.9152
1	1	9	-10.0263	0.91437	-10084	0.00057	0.9144
1	1	7	-10.1185	0.91429	-10097	0.00057	0.9143
1	2	7	-2.7225	0.913587	-10078	0.00057	0.9136
1	2	8	-6.4964	0.913364	-10068	0.00057	0.9134
1	2	6	-5.5925	0.913143	-10079	0.00057	0.9131
2	1	6	-4.2527	0.91297	-10076	0.00057	0.9130
1	2	5	-7.0416	0.912193	-10071	0.00058	0.9122
1	2	9	-7.6697	0.911977	-10039	0.00058	0.9120
2	1	5	-5.5807	0.911019	-10053	0.00059	0.9110
1	2	4	-7.7423	0.910648	-10055	0.00059	0.9106
1	1	10	-9.829	0.908961	-9993	0.00060	0.9090
1	1	6	-10.0191	0.908749	-10019	0.00060	0.9087

SI Table S6. The most efficient twenty rainfall-DOC load CT-TF models for the LI7 basin for the 5th to 18th February 2013 period, arranged in descending order according to their R_t^2 (efficiency measure term).

den	num	delay	YIC	R_t^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>1</u>	<u>-3.7545</u>	<u>0.8384</u>	<u>-7366</u>	<u>0.0023</u>	<u>0.8384</u>
2	2	0	-3.4120	0.8380	-7370	0.0023	0.8380
2	2	2	-4.0142	0.8363	-7344	0.0024	0.8363
2	2	3	-4.2109	0.8317	-7303	0.0024	0.8317
1	1	0	-8.5496	0.8253	-7293	0.0025	0.8253
1	2	0	-1.2903	0.8252	-7285	0.0025	0.8252
1	2	1	-4.1240	0.8248	-7275	0.0025	0.8248
2	2	4	-4.3660	0.8244	-7244	0.0025	0.8244
1	1	1	-8.5131	0.8240	-7276	0.0025	0.8240
1	2	2	-5.2101	0.8233	-7257	0.0025	0.8233
1	2	3	-5.8341	0.8205	-7231	0.0026	0.8205
1	1	2	-8.4423	0.8201	-7243	0.0026	0.8201
2	2	5	-4.1384	0.8168	-7185	0.0026	0.8168
1	2	4	-6.2176	0.8162	-7195	0.0026	0.8162
1	1	3	-8.3443	0.8137	-7193	0.0027	0.8137
1	2	5	-6.4425	0.8101	-7148	0.0027	0.8101
2	2	6	-4.0866	0.8088	-7125	0.0028	0.8088
1	1	4	-8.2193	0.8045	-7127	0.0028	0.8045
1	2	6	-6.5526	0.8019	-7089	0.0029	0.8019
2	2	7	-5.8258	0.7944	-7030	0.0030	0.7944

SI Table S7. The most efficient twenty rainfall-DOC load CT-TF models for the LI7 basin for the 26th May to 5th June 2013 period, arranged in descending order according to their R_i^2 (efficiency measure term).

den	num	delay	YIC	R_i^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>8</u>	<u>-8.6520</u>	<u>0.9602</u>	<u>-12220</u>	<u>0.00012</u>	<u>0.9602</u>
2	2	9	-8.6425	0.9598	-12198	0.00012	0.9598
2	2	7	-8.5457	0.9586	-12172	0.00012	0.9586
2	2	10	-8.5187	0.9574	-12114	0.00012	0.9574
2	2	6	-8.3475	0.9551	-12069	0.00013	0.9551
2	2	5	-8.0714	0.9501	-11934	0.00015	0.9501
2	2	4	-7.7046	0.9437	-11776	0.00017	0.9437
2	2	3	-7.2795	0.9355	-11598	0.00019	0.9355
2	2	2	-6.8411	0.9255	-11410	0.00022	0.9255
1	2	10	-9.1341	0.9191	-11248	0.00024	0.9191
1	2	9	-8.9104	0.9169	-11218	0.00024	0.9169
1	2	1	-8.0204	0.9147	-11241	0.00025	0.9147
1	2	8	-8.5654	0.9147	-11190	0.00025	0.9147
1	2	0	-8.3931	0.9144	-11244	0.00025	0.9144
1	2	2	-7.4198	0.9142	-11226	0.00025	0.9142
2	2	1	-6.3355	0.9138	-11220	0.00025	0.9138
1	2	3	-6.3544	0.9133	-11204	0.00025	0.9133
1	1	5	-10.2063	0.9132	-11195	0.00025	0.9132
1	2	7	-8.0332	0.9130	-11171	0.00026	0.9130
1	2	4	-3.2084	0.9124	-11183	0.00026	0.9124

SI Table S8. The most efficient twenty rainfall-DOC load CT-TF models for the LI8 basin for the 5th to 18th February 2013 period, arranged in descending order according to their R_r^2 (efficiency measure term).

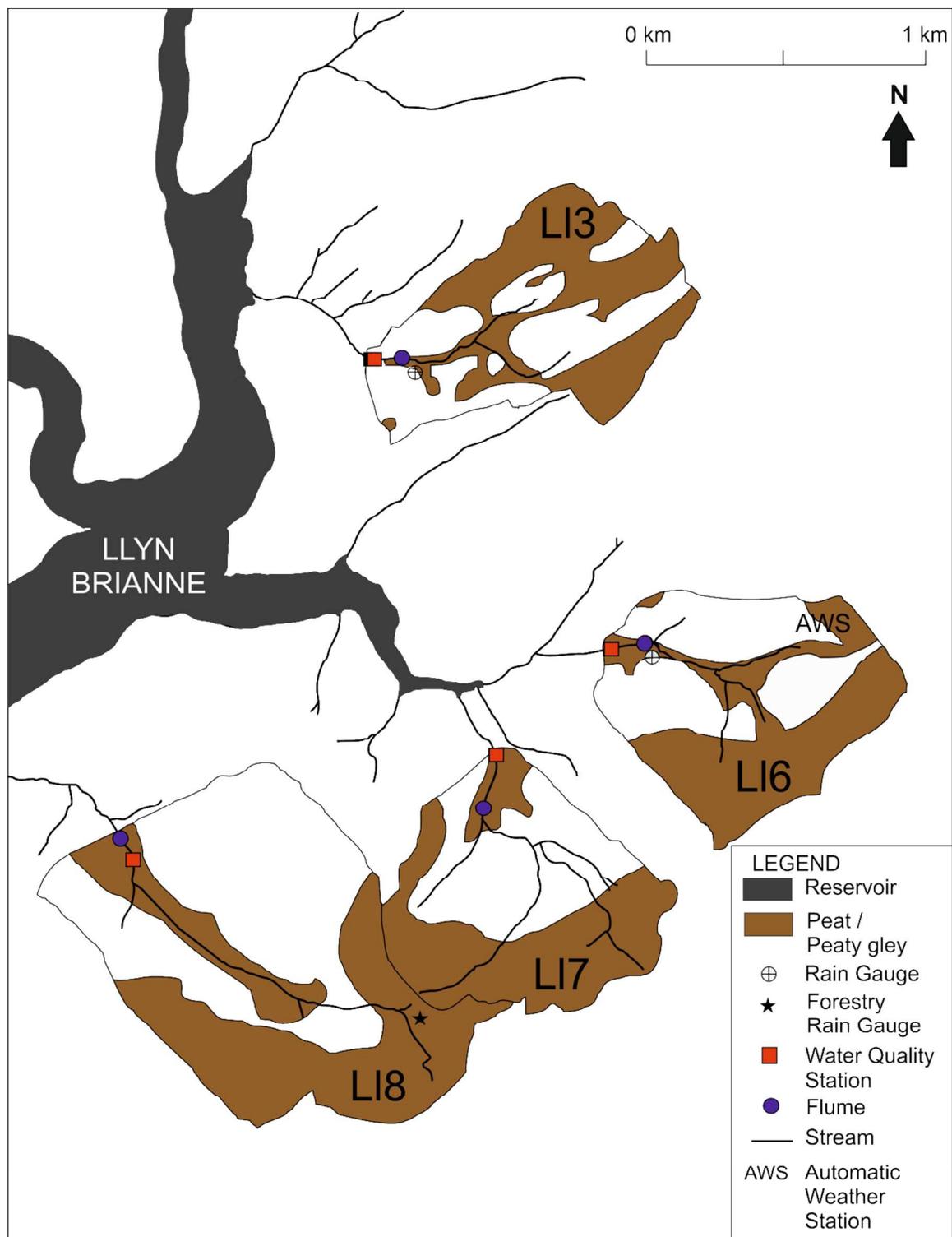
den	num	delay	YIC	R_r^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>2</u>	<u>-5.2748</u>	<u>0.8691</u>	<u>-6895</u>	<u>0.0034</u>	<u>0.8691</u>
2	2	3	-5.4135	0.8682	-6879	0.0034	0.8682
2	2	1	-5.0868	0.8678	-6890	0.0034	0.8678
2	2	4	-5.5048	0.8650	-6843	0.0035	0.8650
2	2	0	-4.8485	0.8645	-6867	0.0035	0.8645
2	2	5	-5.5524	0.8595	-6787	0.0037	0.8595
2	2	6	-5.5594	0.8516	-6714	0.0039	0.8516
2	2	7	-5.5291	0.8416	-6627	0.0041	0.8416
2	2	8	-5.4152	0.8302	-6534	0.0044	0.8302
1	1	0	-8.9631	0.8267	-6581	0.0045	0.8267
1	1	1	-8.9416	0.8258	-6568	0.0045	0.8258
1	2	0	-3.8910	0.8256	-6566	0.0045	0.8256
1	2	1	-5.2109	0.8253	-6557	0.0045	0.8253
1	2	2	-5.9782	0.8245	-6544	0.0046	0.8245
1	2	3	-6.4893	0.8230	-6526	0.0046	0.8230
1	1	2	-8.8931	0.8229	-6540	0.0046	0.8229
1	2	4	-6.8401	0.8206	-6503	0.0047	0.8206
1	1	3	-8.8192	0.8176	-6497	0.0047	0.8176
2	2	9	-5.2239	0.8175	-6439	0.0048	0.8175
1	2	5	-7.0758	0.8172	-6473	0.0048	0.8172

SI Table S9. The most efficient twenty rainfall-DOC load CT-TF models for the LI8 basin for the 26th May to 5th June 2013 period, arranged in descending order according to their R_r^2 (efficiency measure term).

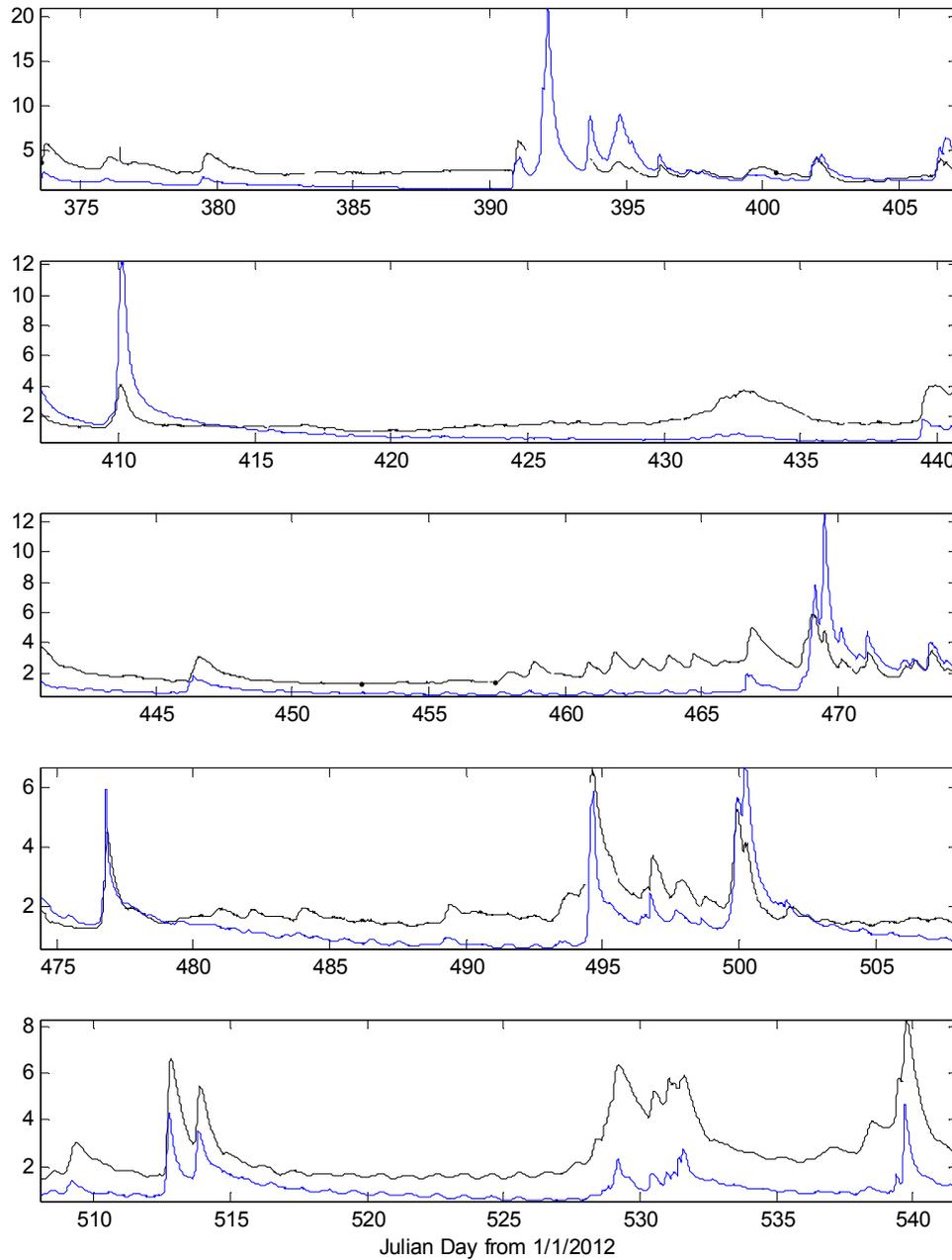
den	num	delay	YIC	R_r^2	BIC	S2	condP
<u>2</u>	<u>2</u>	<u>8</u>	<u>-8.5245</u>	<u>0.9715</u>	<u>-10961</u>	<u>0.00029</u>	<u>0.9715</u>
2	2	7	-8.2544	0.9705	-10921	0.00031	0.9705
2	2	9	-8.6223	0.9703	-10895	0.00031	0.9703
2	2	6	-7.8330	0.9674	-10791	0.00034	0.9674
2	2	10	-8.5630	0.9666	-10730	0.00035	0.9666
2	2	5	-7.3023	0.9623	-10601	0.00039	0.9623
2	2	4	-6.6980	0.9554	-10381	0.00046	0.9554
2	2	3	-6.0509	0.9470	-10153	0.00055	0.9470
2	2	2	-5.3523	0.9373	-9930	0.00065	0.9373
1	1	8	-10.7584	0.9265	-9685	0.00076	0.9265
2	2	1	-4.6002	0.9264	-9720	0.00076	0.9264
1	1	7	-10.7517	0.9261	-9687	0.00077	0.9261
1	1	9	-10.7027	0.9246	-9645	0.00078	0.9246
1	2	6	-3.1153	0.9244	-9656	0.00078	0.9244
1	2	5	-6.0382	0.9244	-9662	0.00078	0.9244
1	2	7	-4.9610	0.9243	-9645	0.00078	0.9243
2	1	6	-1.5255	0.9242	-9652	0.00078	0.9242
1	2	4	-7.0960	0.9240	-9663	0.00079	0.9240
1	2	8	-6.7378	0.9240	-9633	0.00079	0.9240
1	2	9	-7.6524	0.9237	-9621	0.00079	0.9237

SI Table S10. Parameters of the optimal second-order CT-TF models (see Equation 3 research article), plus parameters of the two first-order models after decomposition by partial fraction expansion (see Equation 4 research article). As an illustration, Equation 5 in the research article shows the values of the α_f , α_s , β_f , and β_s parameters for the model of stream DOC_{LOAD} in the LI3 basin during the selected February period (i.e., first line of the table below):

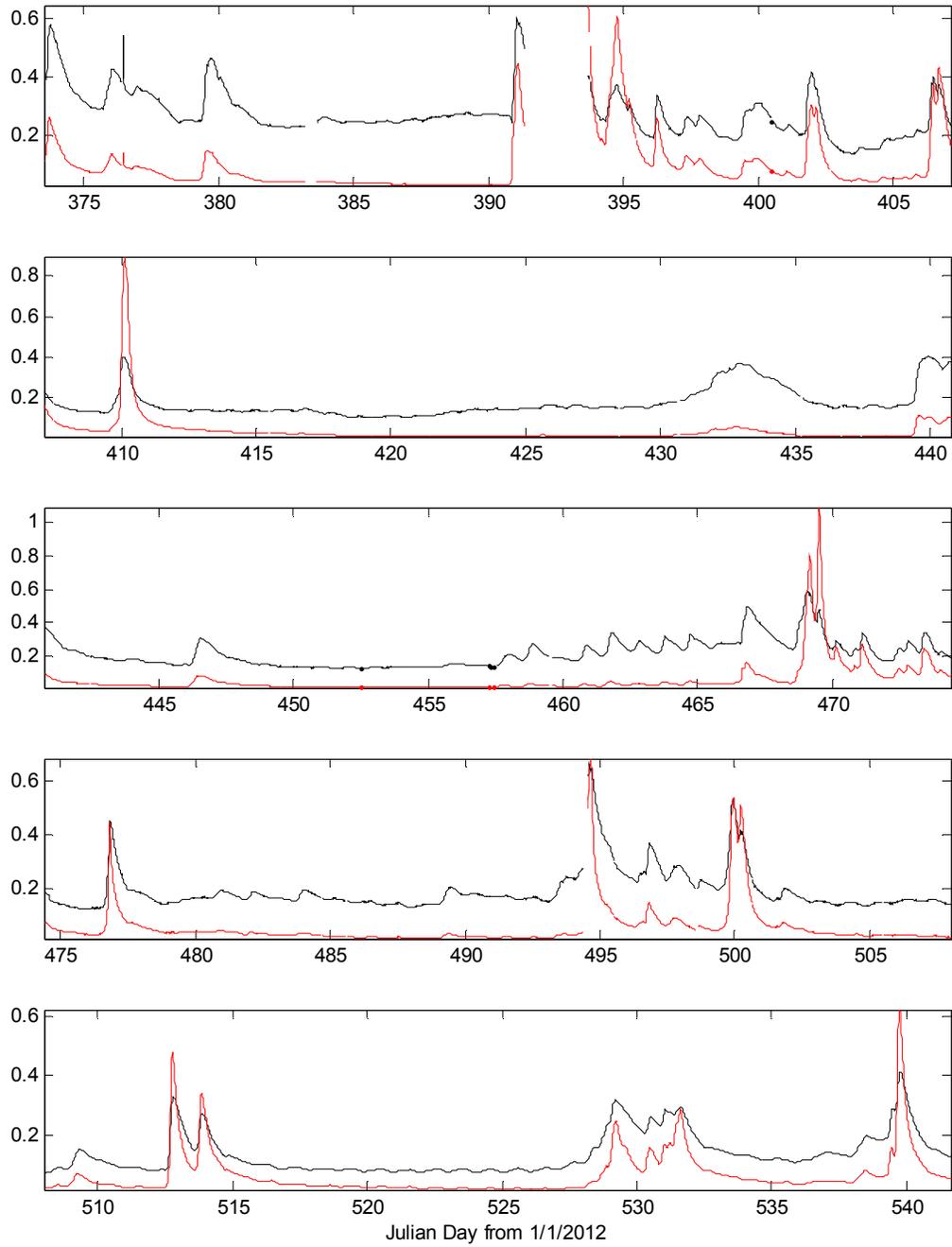
site	DOC load/Q	α_1	α_2	β_0	β_1	α_f	α_s	β_f	β_s
LI3	cold load	0.0559	0.000275	0.0492	0.000439	-0.0504	-0.00546	0.0454	0.00379
	cold Q	0.0546	0.000222	0.0110	0.000156	-0.0501	-0.00443	0.0086	0.00235
	warm load	0.0520	0.000141	0.0365	0.000255	-0.0492	-0.00286	0.0365	0.00026
	warm Q	0.0432	0.000060	5.7553	0.053891	-0.0418	-0.00143	4.6242	1.13100
LI6	cold load	0.0803	0.000562	0.0693	0.000775	-0.0724	-0.00776	0.0657	0.00365
	cold Q	0.0731	0.000339	0.0184	0.000226	-0.0682	-0.00497	0.0162	0.00212
	warm load	0.1118	0.000303	0.0771	0.000472	-0.0028	-0.10902	0.0747	0.00243
	warm Q	0.0928	0.000087	11413	91.28451	-0.0918	-0.00095	10528	885.482
LI7	cold load	0.0585	0.000316	0.0445	0.000370	-0.0525	-0.00601	0.0423	0.00219
	cold Q	0.0569	0.000214	12253	135.1786	-0.0528	-0.00406	10502	1750.86
	warm load	0.0784	0.000407	0.0330	0.000432	-0.0728	-0.00588	0.2930	0.00043
	warm Q	0.0233	0.000015	3706.9	16.98283	-0.0226	-0.00067	3047.7	659.258
LI8	cold load	0.0545	0.000262	0.0563	0.000564	-0.0492	-0.00532	0.0502	0.00603
	cold Q	0.0757	0.000281	21074	395.2800	-0.0039	-0.07178	16465	4608.85
	warm load	0.0234	0.000015	0.0362	0.000066	-0.0228	-0.00067	0.0343	0.00189
	warm Q	0.0327	0.000018	6774.0	47.14361	-0.0322	-0.00057	5403.1	1370.90



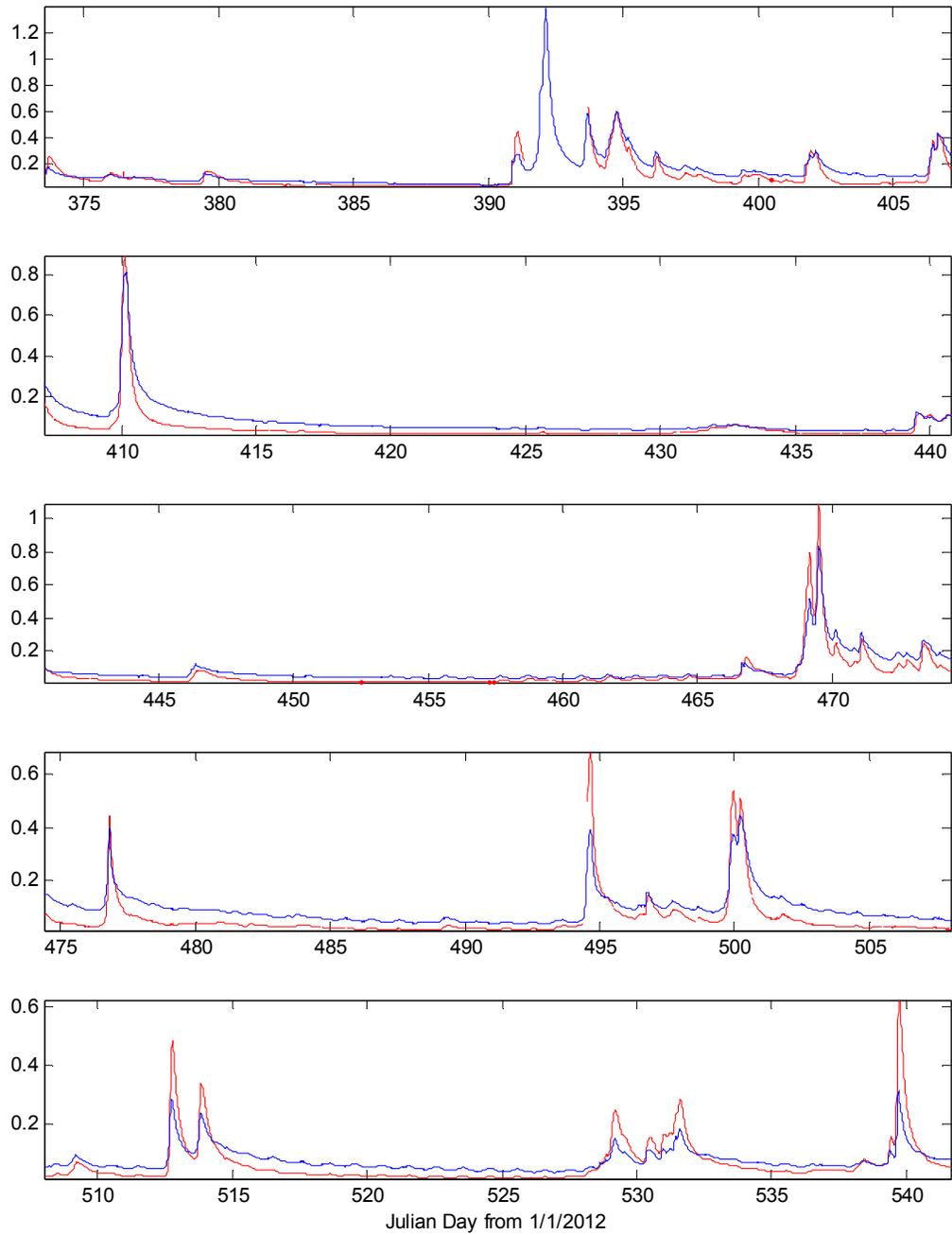
SI Figure S1. Location of monitoring equipment and organic-rich soils within the LI3, LI6, LI7 and LI8 basins near the Llyn Brianne reservoir in upland Wales, United Kingdom



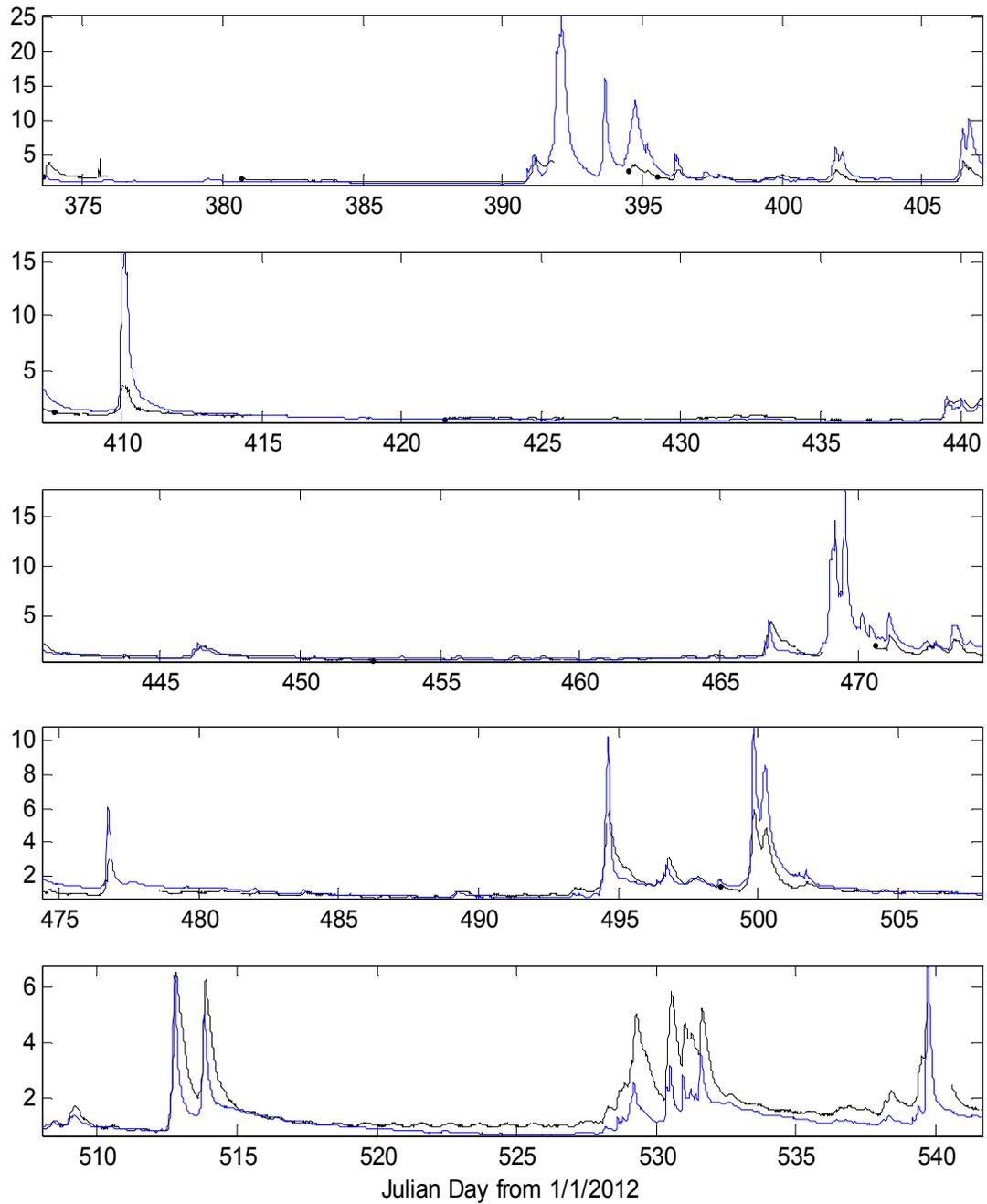
SI Figure S2. Full time series (8th January to 25th June 2013) of DOC concentration (mg/L per 15 min period; black line) alongside streamflow in ML/15min (scaled /40; blue line) at LI3



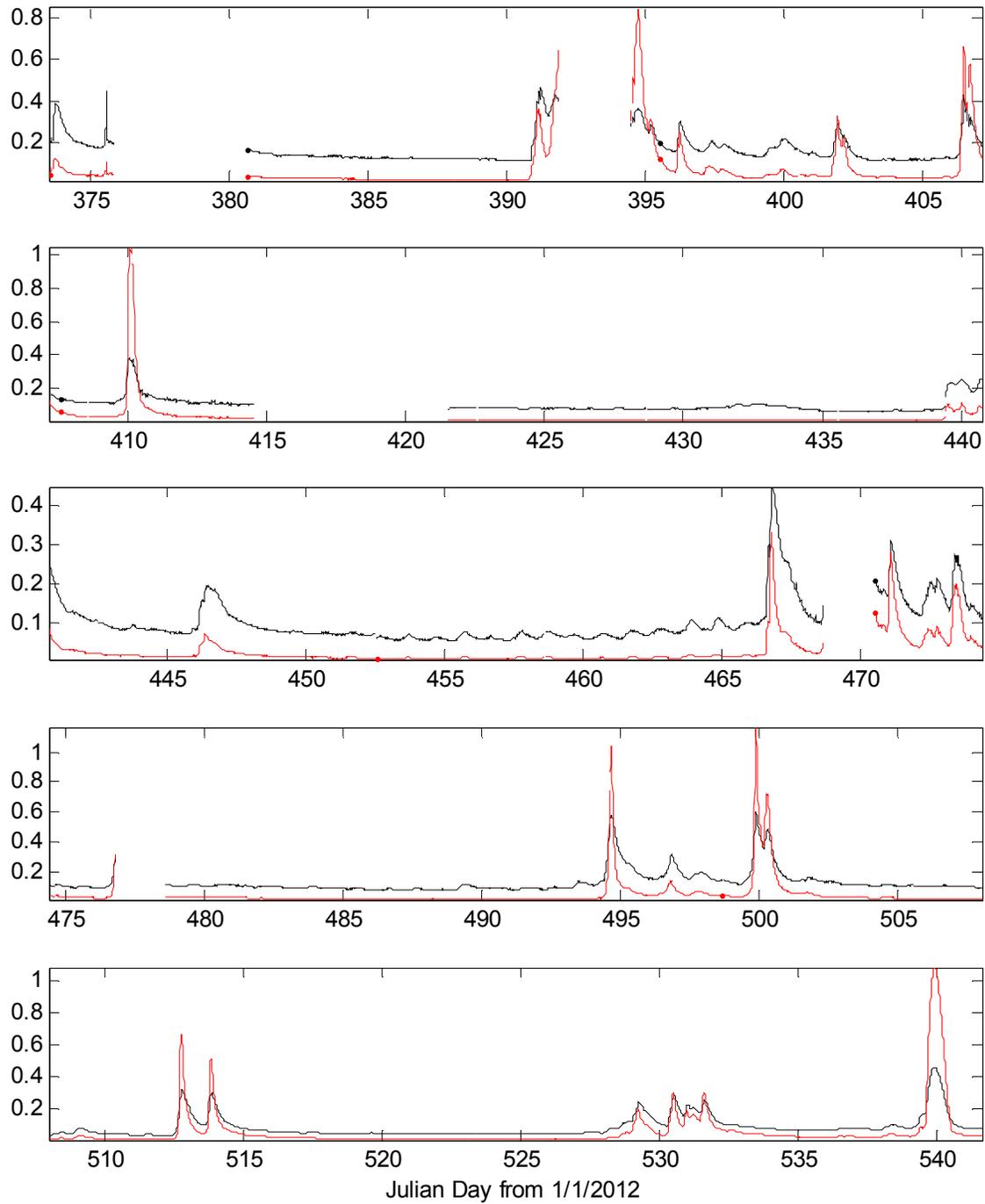
SI Figure S3. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside DOC concentration in mg/L (scaled /10; black line) at LI3



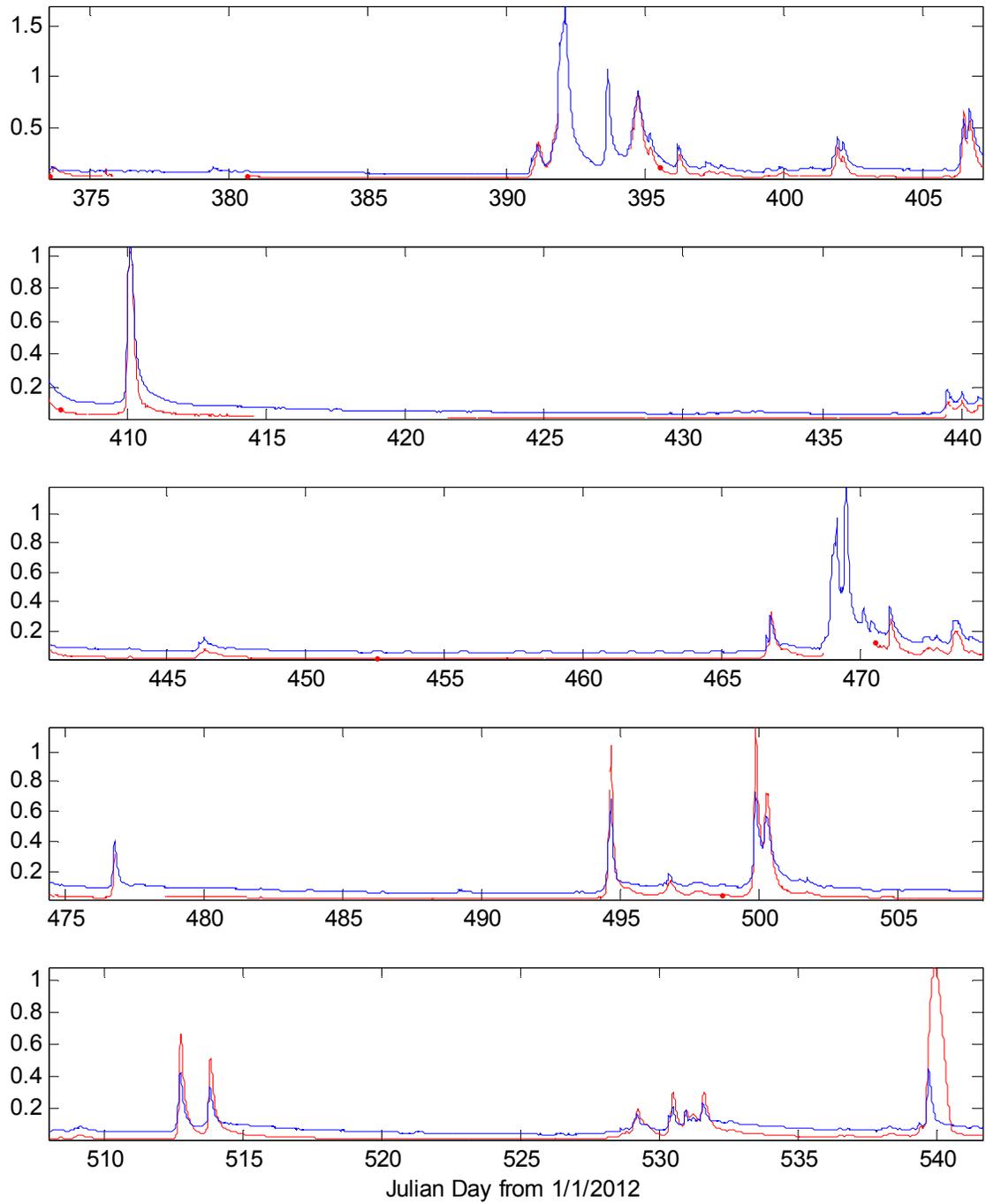
SI Figure S4. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside streamflow in ML/15min (scaled /300; blue line) at LI3



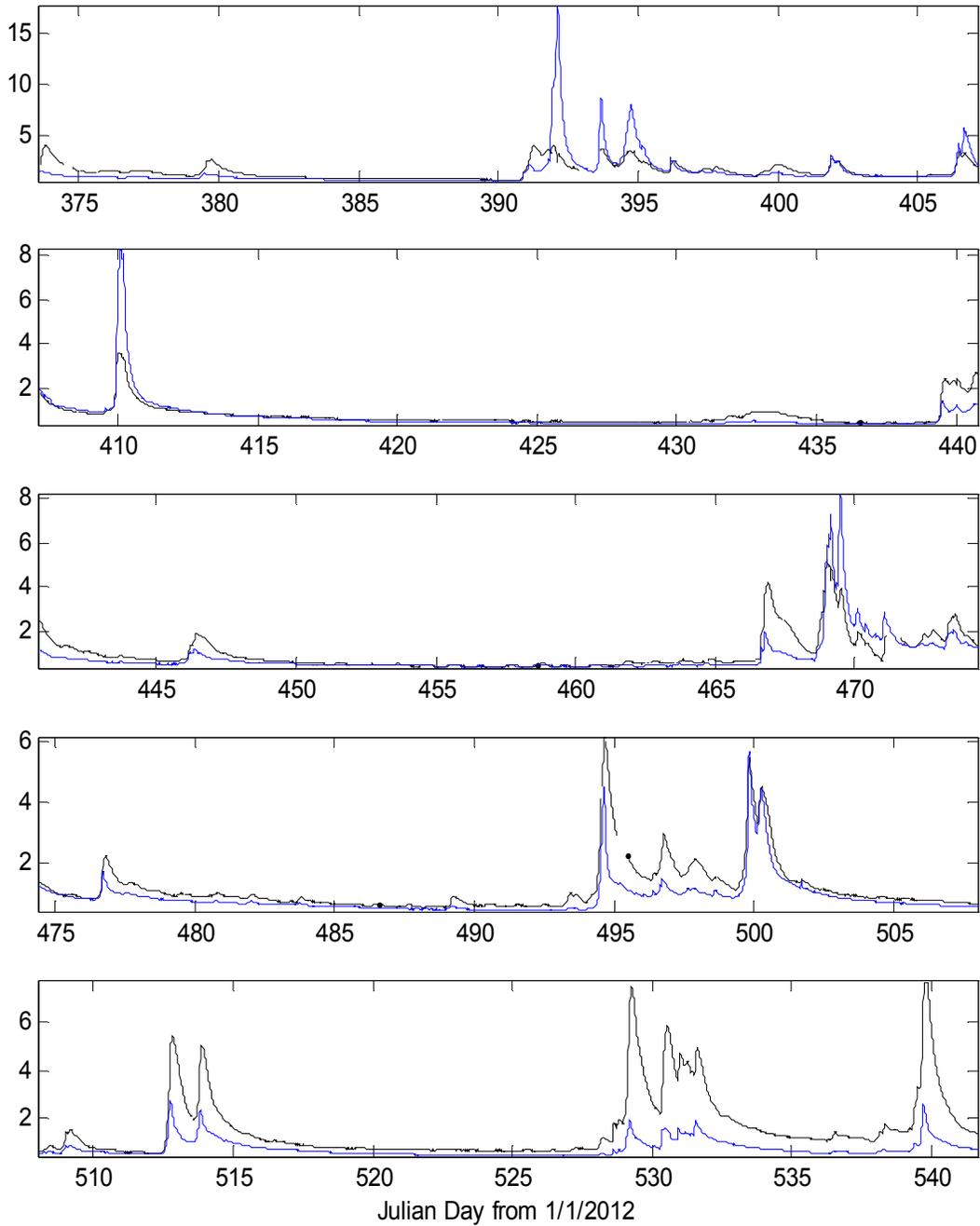
SI Figure S5. Full time series (8th January to 25th June 2013) of DOC concentration (mg/L per 15 min period; black line) alongside streamflow in ML/15min (scaled /30; blue line) at LI6



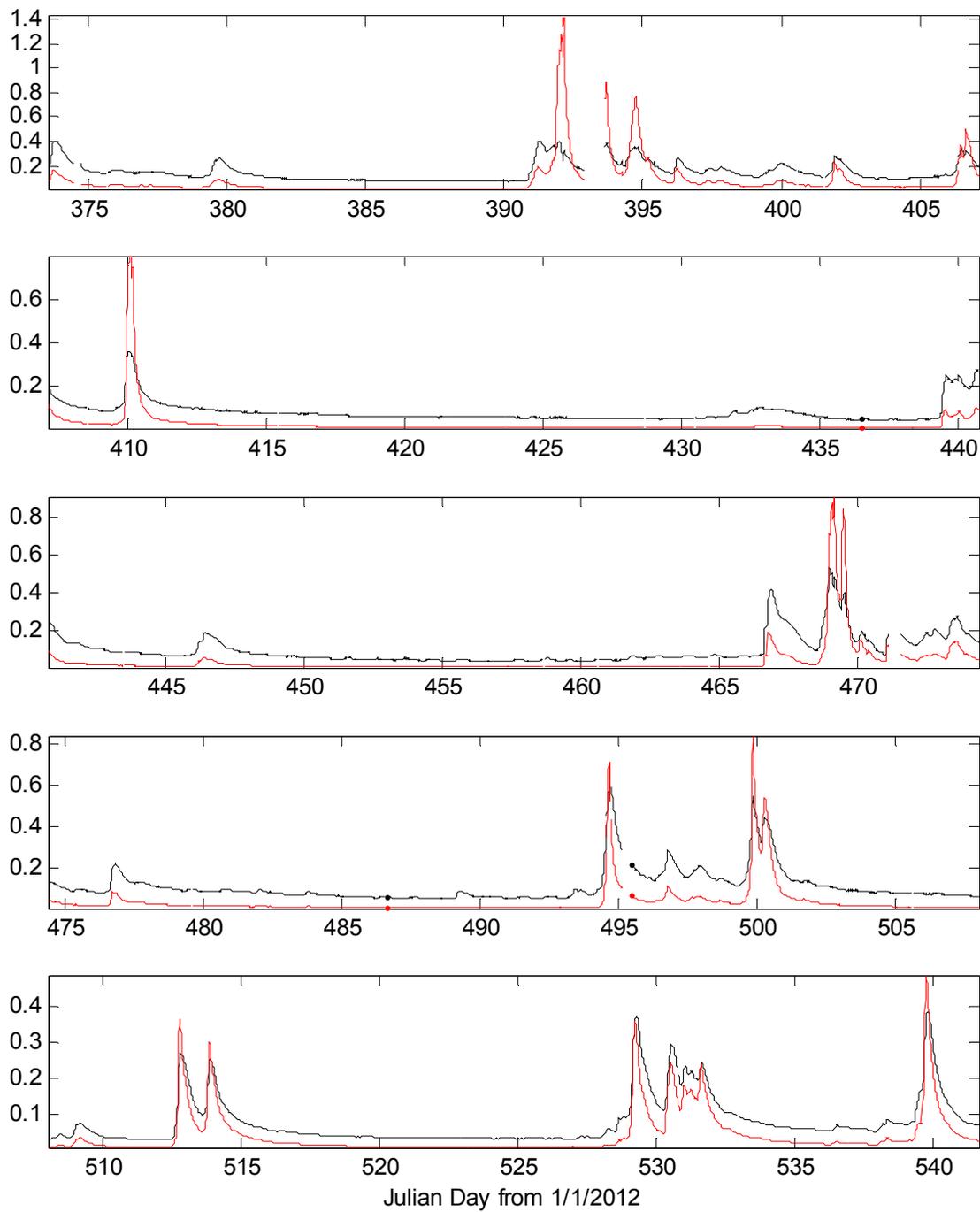
SI Figure S6. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside DOC concentration in mg/L (scaled /10; black line) at LI6



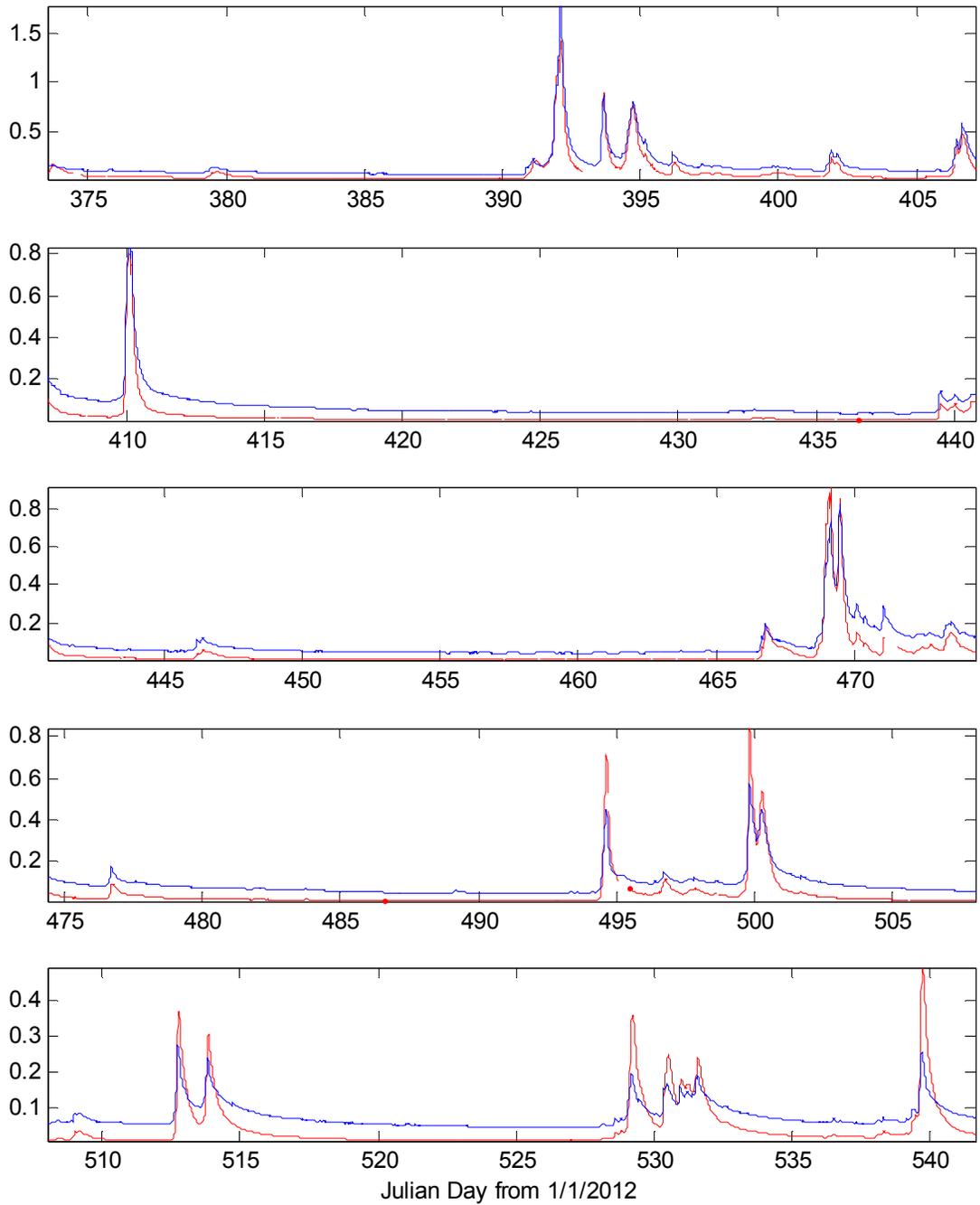
SI Figure S7. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside streamflow in ML/15min (scaled /300; blue line) at LI6



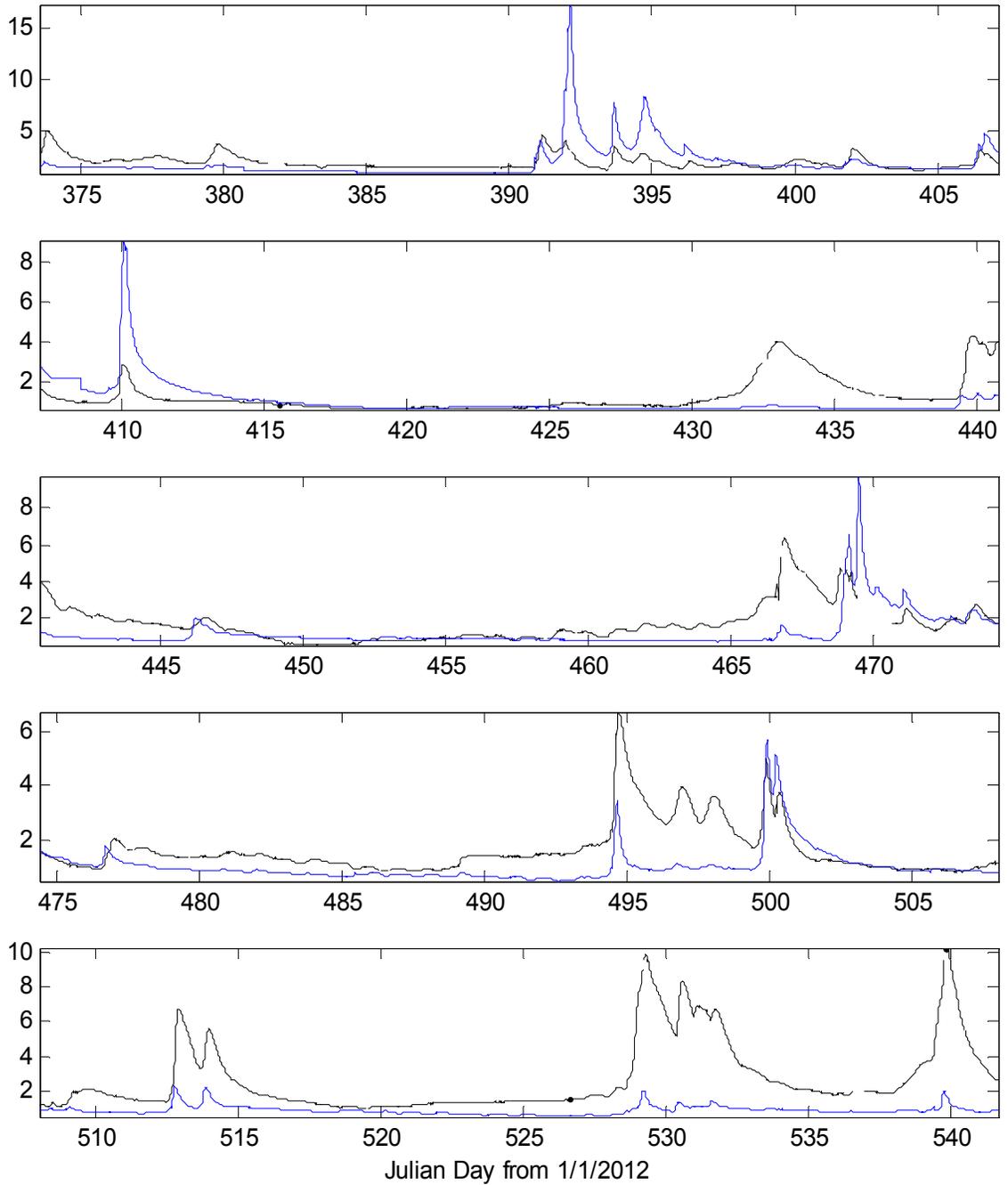
SI Figure S8. Full time series (8th January to 25th June 2013) of DOC concentration (mg/L per 15 min period; black line) alongside streamflow in ML/15min (scaled /30; blue line) at LI7



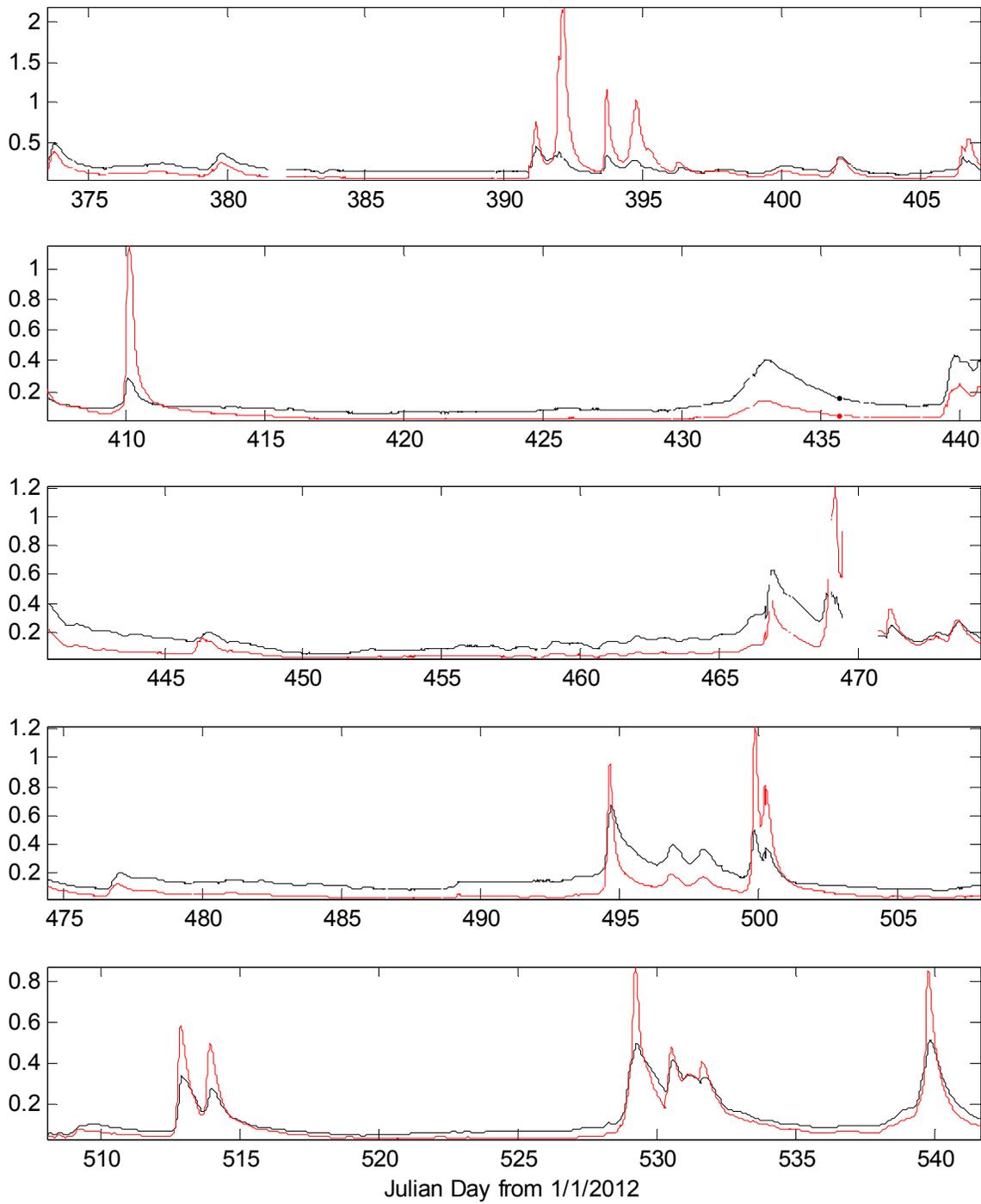
SI Figure S9. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside DOC concentration in mg/L (scaled /10; black line) at LI7



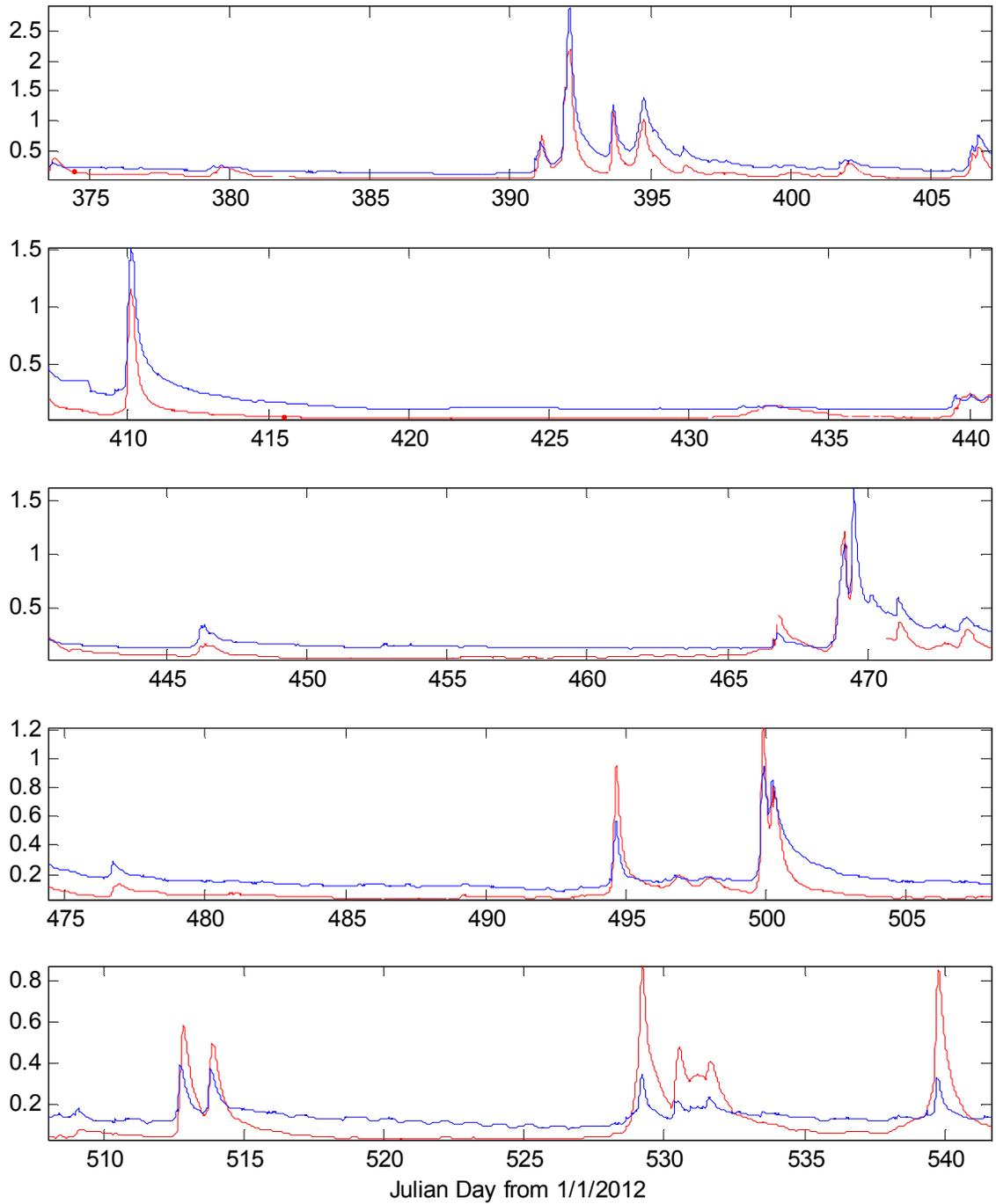
SI Figure S10. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside streamflow in ML/15min (scaled /300; blue line) at LI7



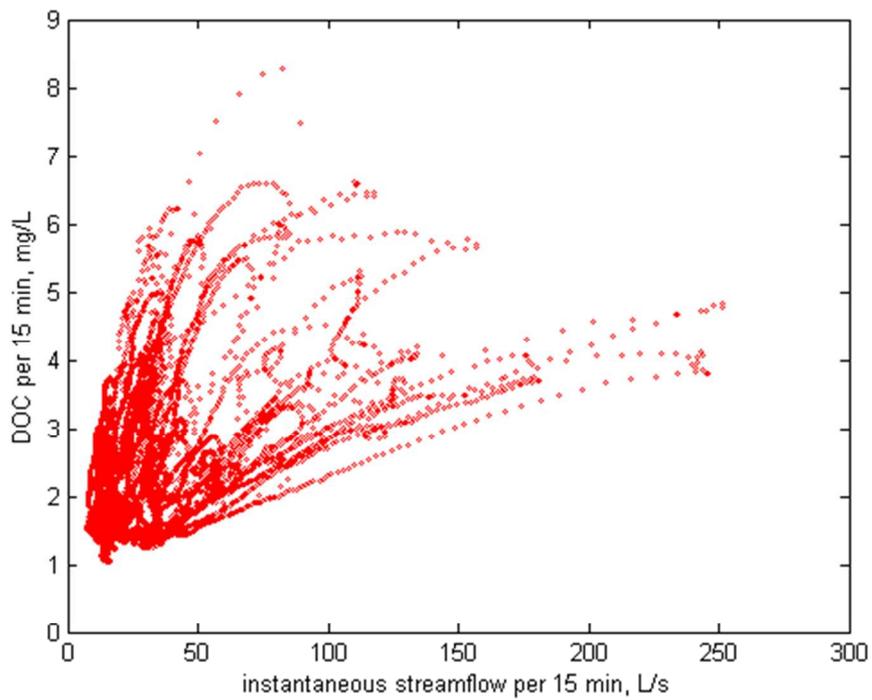
SI Figure S11. Full time series (8th January to 25th June 2013) of DOC concentration (mg/L per 15 min period; black line) alongside streamflow in ML/15min (scaled /50; blue line) at LI8



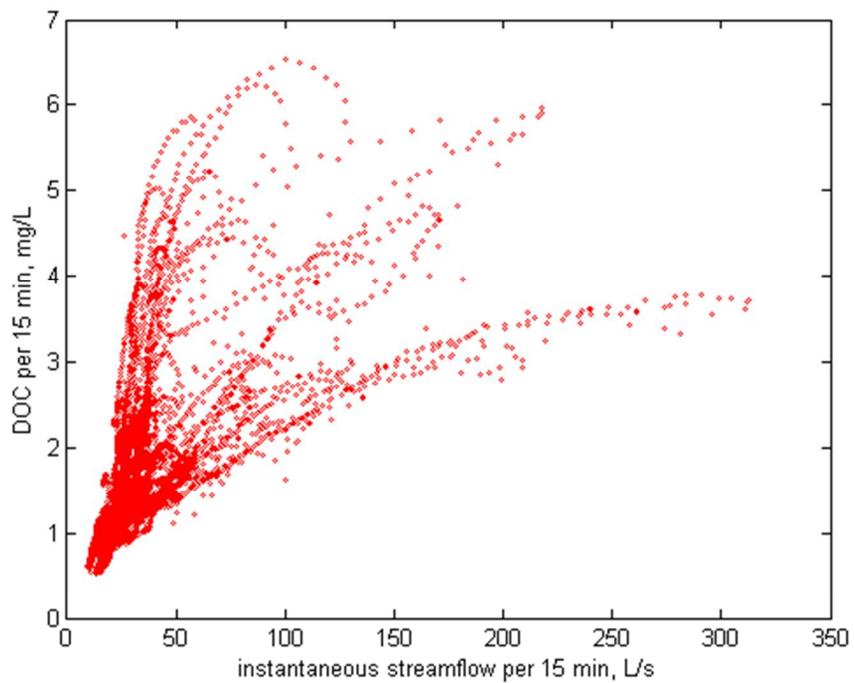
SI Figure S12. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside DOC concentration in mg/L (scaled /10; black line) at LI8



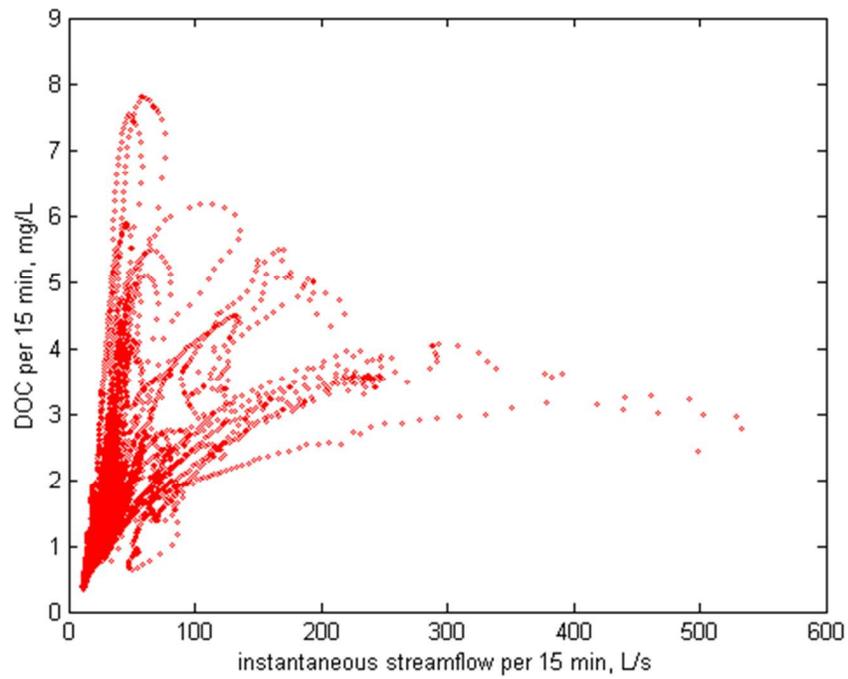
SI Figure S13. Full time series (8th January to 25th June 2013) of DOC load in kg/15min (red line) alongside streamflow in ML/15min (scaled /300; blue line) at LI8



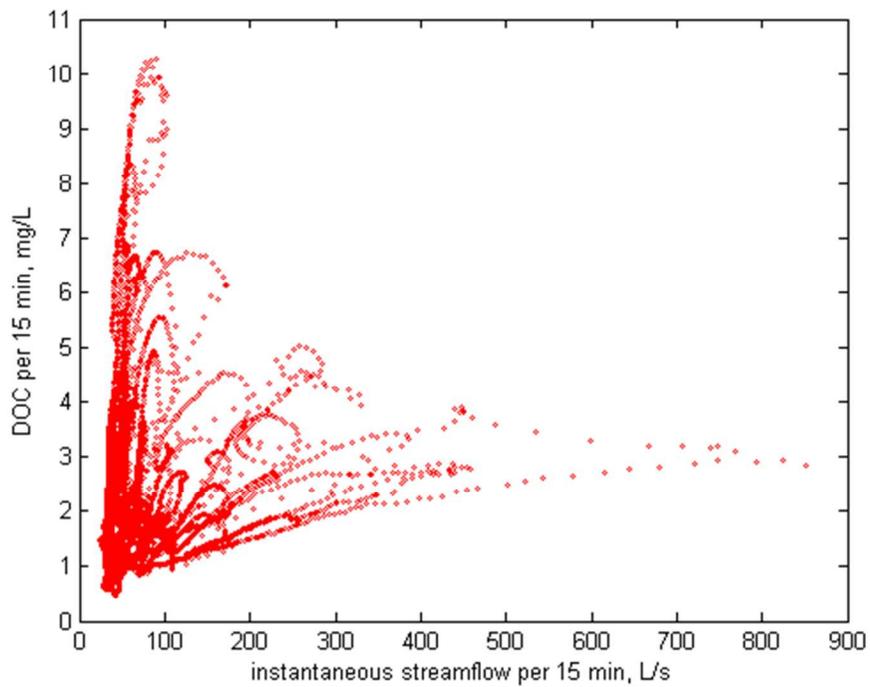
SI Figure S14. Relationship between instantaneous streamflow recorded every 15 minutes (L/s) and DOC concentration recorded every 15 minutes (mg/L) within the LI3 stream near Llyn Brianne, UK



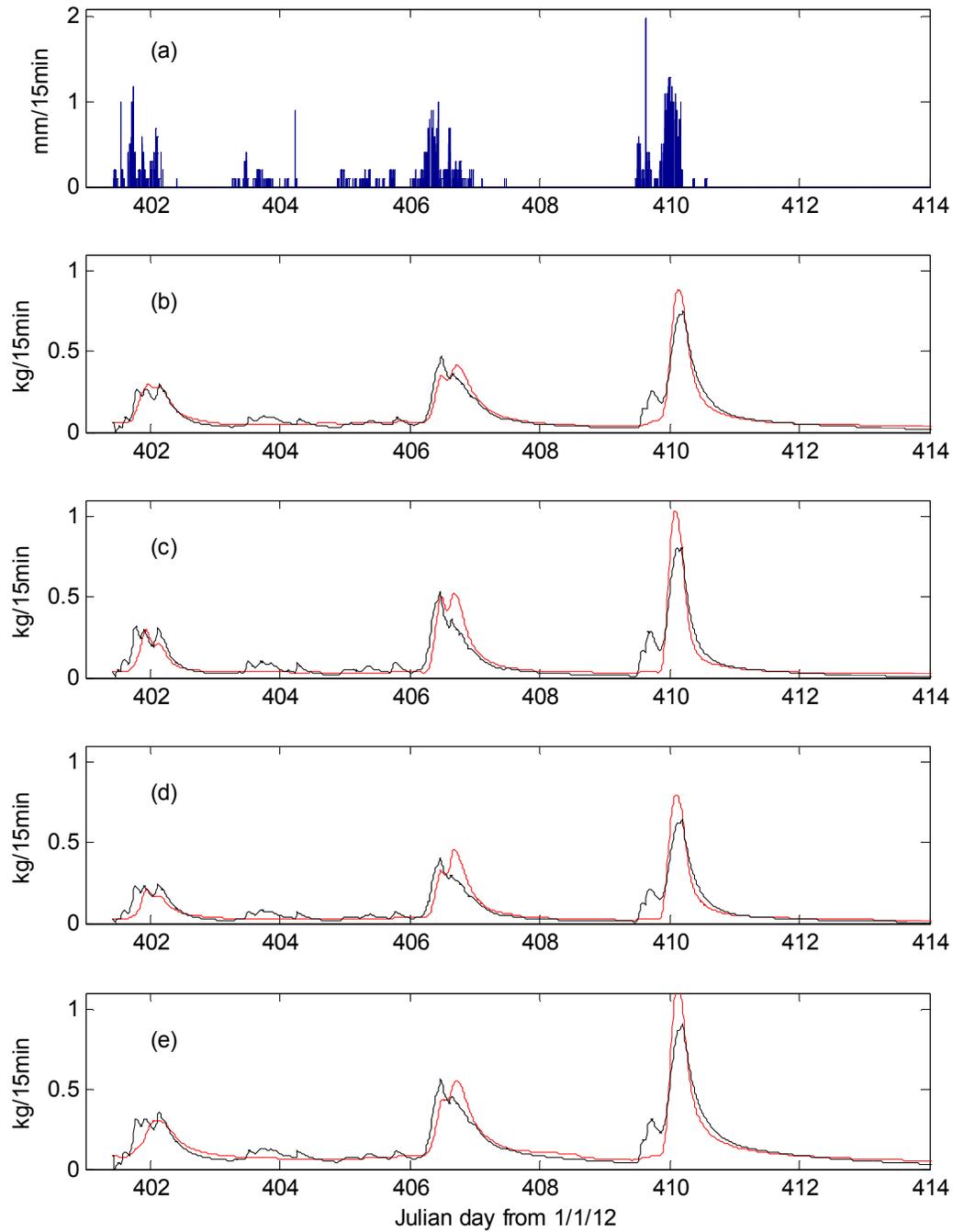
SI Figure S15. Relationship between instantaneous streamflow recorded every 15 minutes (L/s) and DOC concentration recorded every 15 minutes (mg/L) within the LI6 stream near Llyn Brianne, UK



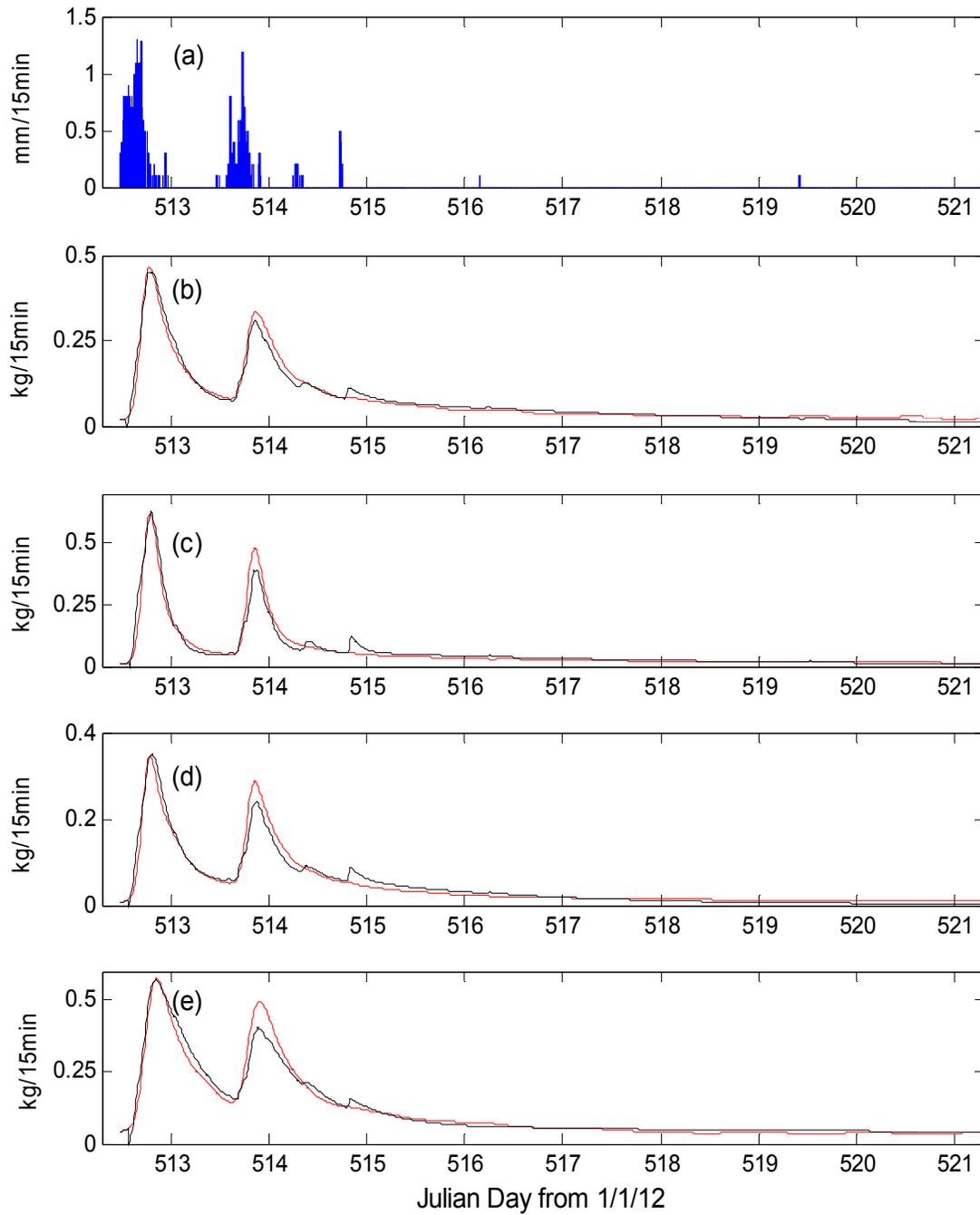
SI Figure S16. Relationship between instantaneous streamflow recorded every 15 minutes (L/s) and DOC concentration recorded every 15 minutes (mg/L) within the LI7 stream near Llyn Brianne, UK



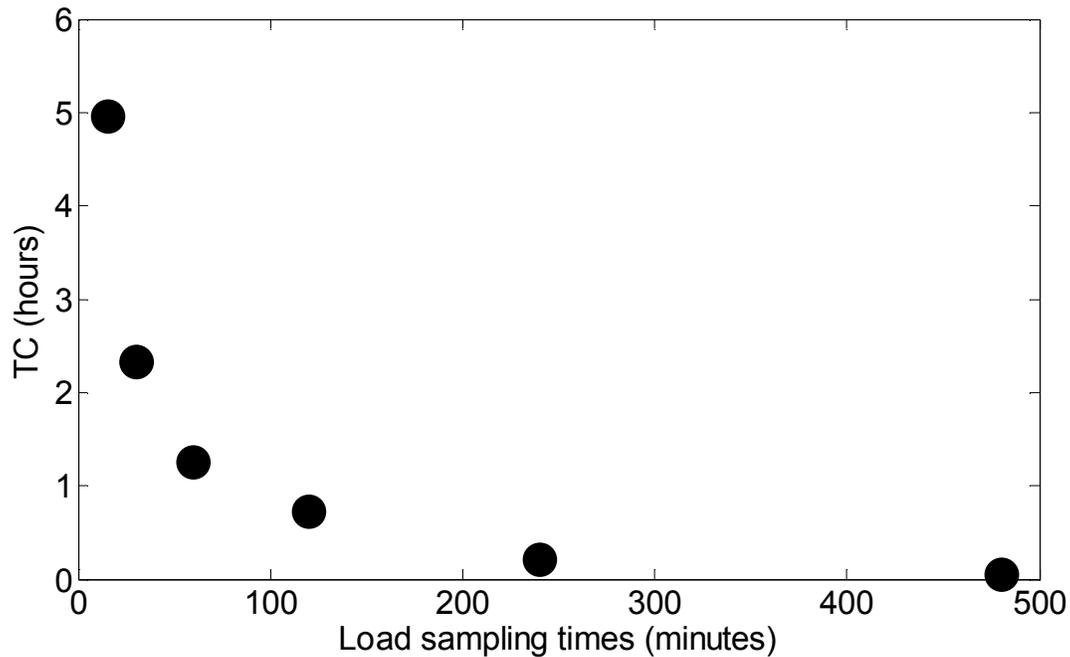
SI Figure S17. Relationship between instantaneous streamflow recorded every 15 minutes (L/s) and DOC concentration recorded every 15 minutes (mg/L) within the LI8 stream near Llyn Brianne, UK



SI Figure S18. Rainfall (a) and simulated DOC load (black line) and observed DOC load (red line) for optimal second-order continuous time transfer function models of: (b) LI3 stream, (c) LI6 stream, (d) LI7 stream, and (e) LI8 stream for the contiguous storms over the 5th to 18th February 2013 period



SI Figure S19. Rainfall (a) and simulated DOC load (black line) and observed DOC load (red line) for optimal second-order continuous time transfer function models of: (b) LI3 stream, (c) LI6 stream, (d) LI7 stream, and (e) LI8 stream for the contiguous storms over the 26th May to 5th June 2013 period



SI Figure S20. The effect on TC_{fast} of sampling DOC_{LOAD} at 15-mins (original rate see Table 1 research article) and sub-sampling DOC_{LOAD} from the 15-min time-series at 30-mins, 1 h, 2 h, 4 h and 8 h for the example of the LI3 stream over 5th – 18th Feb 2013. The rainfall used as input to the CT-TF models was integrated over the respective sampling periods.

References

- (1) Young, P.C. The refined instrumental variable method. *J. Eur. Syst. Automat.* **2008**, *42*, 149–179.
- (2) Jones, T.D.; Chappell, N.A. Streamflow and hydrogen ion interrelationships identified using Data-Based Mechanistic modelling of high frequency observations through contiguous storms. *Hydrol. Res.* **2014**, *in press*, doi: 10.2166/nh.2014.155.
- (3) Jakeman, A.J.; Littlewood, I.G.; Whitehead, P.G. An assessment of the dynamic response characteristics of streamflow in the Balquhiddy catchments. *J. Hydrol.* **1993**, *145*, 337-355.
- (4) Young, P.C. The estimation of continuous-time rainfall-flow models for flood risk management. In *Role of Hydrology in Managing Consequences of a Changing Global Environment*; Walsh, C., Ed.; British Hydrological Society: Newcastle upon Tyne, 2010; 303–310.

- (5) Rier, S.T.; Stevenson, R.J.; LaLiberté, G.D. Photo-acclimation response of benthic stream algae across experimentally manipulated light gradients: a comparison of growth rates and net primary productivity. *J. Phycol.* **2006**, *42*, 560–567.
- (6) Rowland, A.P.; Neal, C.; Reynolds, B.; Neal, M.; Lawlor, A.J.; Sleep, D. Manganese in the upper Severn mid-Wales. *Journal of Environmental Monitoring* **2012**, *14*, 155-164.
- (7) Grayson, R.; Holden, J. Continuous measurement of spectrophotometric absorbance in peatland streamwater in northern England: implications for understanding fluvial carbon fluxes. *Hydrol. Process.* **2012**, *26*, 27-39.
- (8) Young, P.C. Recursive estimation, forecasting and adaptive control. In *Control and Dynamic Systems: Advances in Theory and Applications*; Leondes, C.T., Ed.; Academic Press: San Diego, 1990; 119–166.
- (9) Ockenden, M.C.; Chappell, N.A. Identification of the dominant runoff pathways from the data-based mechanistic modelling of nested catchments in temperate UK. *J. Hydrol.* **2011**, *402*, 71–79.
- (10) Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*, 4th edition; Wiley: Hoboken, 2008.
- (11) Weatherley, N.S.; Ormerod, S.J. The impact of acidification on macroinvertebrate assemblages in Welsh streams: towards an empirical model. *Environ. Pollut.* **1987**, *46*, 223–240.
- (12) Littlewood, L.G. *The Dynamics of Acid Runoff from Moorland and Conifer Afforested Catchments Draining into Llyn Brianne, Wales*. PhD Thesis, University of Wales, 1989.
- (13) Reynolds, B.; Norris, D.A. *Llyn Brianne Acid Waters Project – Summary of Catchment Characteristics*. Institute of Terrestrial Ecology, 1990.
- (14) FAO-Unesco. *Soil Map of the World, Revised Legend*. FAO, Rome, 1990..