Supporting Information

For *Environmental Science & Technology*

Quantification of Active Ingredient Losses from Formulating Pharmaceutical Industries and Contribution to Wastewater Treatment Plant Emissions

Sabine Anliker[a,b], Michael Patrick[a,b], Kathrin Fenner[a,b,c], Heinz Singer*[,a]

[a]Swiss Federal Institute of Aquatic Science and Technology, Eawag, CH-8600 Dübendorf, Switzerland
[b] Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, CH-8092 Zürich, Switzerland
[c] Department of Chemistry, University of Zürich, CH-8006 Zürich, Switzerland

* Corresponding Author
Email: heinz.singer@eawag.ch; Phone: +41 58 765 55 77

Number of Pages: 41
Number of Figures: 9
Number of Tables: 6

Contents

Materials and Methods

SI 1: Chemicals and Solutions

Organic solvents were supplied by Fisher Scientific (Switzerland), Sigma Aldrich (Switzerland), and Merck (Germany) at HPLC grade purity (≥ 98%). Formic acid (≥ 98%) was purchased from Merck (Germany). Ultrapure water was generated by a laboratory water purification system (Barnstead Nanopure, Thermo Scientific, USA). Reference standards of the target analytes were purchased from Sigma Aldrich (Switzerland), Dr. Ehrenstorfer (Germany), Toronto Research Chemicals (Canada) or ReseaChem (Switzerland) at analytical grade (> 94% purity). Isotope-labeled internal standards (ISTDs) at a purity ≥ 97% were obtained from the following suppliers: Sigma Aldrich (Switzerland), TRC Canada (Canada), Dr. Ehrenstorfer (Germany), TCI Europe (Belgium), Cambridge Isotope Laboratories (USA), HPC Standards (Germany), CDN Isotopes (Germany), Novartis (Switzerland), Cerilliant (USA), LGC (UK) and Lipomed (Switzerland).

Individual stock solutions of analytes and ISTDs were prepared at a concentration of 1 g/L in suitable solvents (preferably in ethanol or acetonitrile) depending on solubility and stability. Working standard solutions were produced by combining individual stock solutions into mixtures. Working standard solutions of target analytes were prepared at concentrations of 50, 5 and 0.5 µg/L in ethanol. The working standard solution of the ISTDs was prepared in ethanol at a concentration of 50 µg/L. All solutions were stored at -20°C. The list of 232 spiked ISTDs is in **SI 2 Table S1** and 162 target analytes are given in **SI 8 Table S5** in a separate EXCEL workbook.

SI 2: Table S1 List of Isotope-Labeled Internal Standards (ISTDs)

See separate file (Table_S1_ISTDs) in EXCEL workbook (es0c05178_si_002.xlsx).

SI 3: Additional Information on WWTPs

Table S2: Additional Information on WWTPs[a].

WWTP characterization	**WWTP_large[b]**	**WWTP_small[b]**
Connected inhabitants	130’000	15’000
Hospital beds in the catchment	650	0
Average wastewater from formulating pharmaceutical industries [%]	0.1	2.3
Sludge retention time [days]	10	14
Hydraulic retention time [h]	23 dry weather; 6 rainy weather	11; 5
Wastewater volume [m3/a]	19'000'000	1'900'000
Water quality parameters of effluent[c]		
TSS [mg/L]	1.5	9.0
DOC [mg/L]	5	9.1
COD [mg/L]	11.4	29.2
P-tot [mg/L]	0.1	0.4
BOD5 [mg/L]	2.7	11.1
NH_4-N [mg/L]	0.7	12.9
pH	7.6	7.4
Removal efficiency DOC [%]	93	90
Removal efficiency P-tot [%]	97	89

[a] Data was partially retrieved from the 2017 annual report of the WWTP (publicly available online) and partially received from cantonal authorities.

[b] Anonymity was granted to the WWTPs and the connected pharmaceutical company, therefore the exact locations of the sampling sites are not indicated.

[c] The water quality parameters of the effluent correspond to annual average values. The number of available measurements differed between WWTP and parameter; it ranged from 21 (biweekly measurement) to 344 (measurements every day).

BOD5 Biochemical oxygen demand (5 days)
COD Chemical oxygen demand
DOC Dissolved organic carbon
LOQ Limit of quantification
P-tot Total phosphorus
TSS Total suspended solids

SI 4: Analytical Instrumentation and Method

The liquid chromatography high-resolution mass spectrometry (LC-HRMS) measurements were performed with the following instrumentation: a PAL HTS-Xt autosampler (CTC Analytics), an Ultimate 3000 RS LC-pump (Thermo Scientific), an Atlantis T3 LC-column (3 µm particle size, 3.0 x 150 mm inner diameter, Waters) guarded with a precolumn of the same material and a precolumn filter, a column oven (Portmann Instruments) and a Q Exactive Plus mass spectrometer (Thermo Scientific) with a heated electrospray ion source (HESI II).

A sample volume of 100 µL was injected onto the LC column, which was kept at a temperature of 30 °C and the eluent flow rate was set to 300 µL/min. The LC mobile phase A consisted of ultrapure water and phase B of methanol, both acidified with 0.1% formic acid. The chromatographic run lasted 30.5 min. First 5% B was maintained for 1.5 min, then B linearly increased to 95% over 17 min, remained at 95% stayed for 7.5 min and finally decreased in 0.5 min to 5% B for a 4 min re-equilibration.

The MS method was chosen to allow for suspect and non-target analysis, in addition to target screening. It consisted of two full scans of different mass windows to extend the dynamic range of the measurement. The full scans were followed by six data independent (DIA) MS2 scans, resulting in a cycle time of 1.2 s and on average 10 data points per chromatographic peak of about 12 s width. An inclusion list was used containing the following *m/z* values as center masses for the DIA MS2 scans: 105.00000, 175.00000, 245.00000, 315.00000, 385.00000 and 710.00000. Data were acquired in positive and negative electrospray ionization (ESI) mode in separate runs. Details on the MS parameter settings are given in **Table S3.**

Table S3: Parameters Settings for Q Exactive Plus MS Measurement

	Positive mode	Negative mode
Ion source		
Spray voltage [kV]	4	-3
Capillary temperature [°C]	320	
Sheat gas flow (nitrogen) [arb]	40	
Auxiliary gas flow (nitrogen) [arb]	10	
Sweep gas flow [arb]	0	
S-Lense RF level [arb]	50	
MS		
1st full MS resolution at *m/z* 200	35'000	
1st full MS scan range [*m/z*]	50-105	
1st full MS AGC target [arb]	1E6	
1st full MS max injection time [ms]	100	
1st full MS spectrum data type	Profile	
2st full MS resolution at *m/z* 200	140'000	
2st full MS scan range [*m/z*]	100-1000	
2st full MS AGC target [arb]	1E6	
2st full MS max injection time [ms]	100	
2st full MS spectrum data type	Profile	
DIA-MS2 default charge state	1	
DIA-MS2 resolution at *m/z* 200	17'500	
DIA-MS2 AGC target [arb]	5E5	
DIA-MS2 max injection time [ms]	70	
DIA-MS2 isolation window [*m/z*]	80	
DIA-MS2 isolation offset [*m/z*]	0	
DIA-MS2 fixed first mass	-	
DIA-MS2 spectrum data type	Profile	
1st DIA-MS2 stepped NCE [arb]	110, 115, 135	
2nd DIA-MS2 stepped NCE [arb]	70, 90, 105	
3rd DIA-MS2 stepped NCE [arb]	45, 60, 75	
4th DIA-MS2 stepped NCE [arb]	15, 30, 45	
5th DIA-MS2 stepped NCE [arb]	15, 20, 30	
6th DIA-MS2 stepped NCE [arb]	15, 20, 30	

arb: arbitrary units

DIA: Data independent acquisition

ms: milliseconds

m/z: mass to charge ratio

SI 5: Quantification Method and Quality Control

Quantification

Quantification of the target analytes was performed with the software TraceFinder 4.1 EFS (Thermo Scientific) based on the area ratio of the reference standard of the analyte and the selected isotope-labeled internal standard (ISTD). For 91 analytes a structurally identical isotope-labeled analogue was available, while for 71 analytes the ISTD with the closest retention time (RT) was used for quantification. Target analyte peaks were extracted with a mass tolerance of 5 ppm and further identified by comparing the isotopic pattern and retention time to those of the matching reference standard. Only peaks with a signal-to-noise ratio larger than 10 and with at least five data points were considered for quantification. The calibration curves included the following 10 levels: 5, 10, 25, 50, 100, 500, 1000, 2'500, 5'000 and 10'000 ng/L, and were obtained by linear regression applying a 1/x weighting factor.

Limit of quantification

The limit of quantification (LOQ) was determined based on the lowest quantified calibration level (*i.e.,* signal-to-noise ratio larger than 10 and at least five data points). The lowest quantified calibration level was multiplied by the inverse of the absolute recovery (see below) to account for matrix effects. Additionally, if traces of the target analyte were found in the storage blanks (ultrapure water stored and processed in the same way as the samples), the LOQ was also calculated as follows:

$$LOQ = \ average\ conc.\ in\ blanks \quad + 10\ sd(conc.\ in\ blanks)$$

with *sd* being the standard deviation. Finally, if two LOQ were available (i.e., one derived from the lowest calibration level and one from the storage blanks), the higher value is reported.

Matrix effects

To assess matrix effects during instrumental analysis, the absolute recovery was determined for each analyte based on the chromatographic peak areas. No distinction was made between the two sampling sites, resulting in one absolute recovery per analyte. For target analytes with structurally identical ISTD, the absolute recovery was calculated as follows:

$$Absolute\ Recovery = \frac{average\ ISTD\ area\ in\ samples}{average\ ISTD\ area\ in\ calibration\ standards}$$

If no structurally identical ISTD was available, the absolute recovery was calculated based on the samples spiked with target analytes (*i.e*., for each sampling site two at a concentration level of 50 ng/L and two at 1000 ng/L), using the following equation:

$$Absolute\ Recovery = \frac{\sum \frac{area\ in\ spiked\ sample - area\ in\ unspiked\ sample}{area\ in\ corresponding\ calibration\ standard}}{number\ of\ spiked\ samples}$$

Accuracy

The accuracy of the measurement was assed with the relative recovery based on the concentrations found in the samples spiked with the target analytes (*i.e.*, for each sampling site two at a concentration level of 50 ng/L and two at 1000 ng/L). Relative recoveries were not calculated if the concentration in the unspiked samples was (i) higher than twice the spiked concentration or (ii) less than twice the LOQ. No distinction was made between the two sampling sites, resulting in one relative recovery per analyte. The following equation was used to calculate the relative recovery:

$$Relative\ Recovery\ [\%] = \frac{\sum \frac{conc.\ in\ spiked\ sample - conc.\ in\ unspiked\ sample}{conc.\ spiked}}{number\ of\ spiked\ samples} * 100$$

Reproducibility

Three samples from each WWTP were measured twice to assess the overall reproducibility (precision) of the analytical method. The reproducibility was computed as the average variation of the available replicates as follows:

$$Replicate\ variation\ = \frac{\sum \frac{conc.\ in\ sample\ i}{conc.\ in\ replicate\ i}}{number\ of\ replicates}$$

Carryover

Carryover between samples was estimated based on the concentration of target analyte found in the blank following the highest calibration standard (*i.e.*, 10'000 ng/L) in the measurement sequence using the following formula:

$$Carryover\ [\%] = \frac{conc.\ in\ blank}{conc.\ in\ highest\ calibration\ standard} * 100$$

SI 6: Processing Steps Included in enviMass v.4.2 Workflow

Data processing using enviMass (version 4.2)[1] included the following options:

- Mass recalibration based on isotope-labeled internal standards (ISTDs)
- Retention time alignment
- Detection of blank peaks
- Screening for compounds on the ISTD and the target compound list
- Intensity normalization based on ISTDs
- Peak shape correlation
- File-wise componentization including isotopologue and adduct grouping
- Profile componentization
- Time profile extraction
- Profile filtering including profile recalculation to omit peaks from blind files
- Profile blind detection

SI 7: Parameter Settings for enviMass v.4.2 Workflow

Table S4: Parameter Settings for enviMass v.4.2 Workflow

Instrument/Resolution	
	Q-Exactive, ExactivePlus_R140000@200
Peak Picking	
Data filtering	
Filter RT range?	No
Lower RT bound [minutes]	0
Upper RT bound [minutes]	25
Filter mass range?	No
Lower m/z bound	0
Upper m/z bound	2000
Parameter estimation	
Include estimation?	no
Extraction of ion chromatogram	
Max. retention time gap in EIC [s]	300
Max. m/z deviation of a centroid data point from its mean [ppm]	7
Peak picking	
Min. number of centroid data points per peak	3
…within a given RT window [s]	5
Max. RT gap width to be interpolated [s]	10
Max. RT gap width of a single peak from apex [s]	120
Min. log10 intensity threshold	-10

Minimum signal/noise	5
Minimum signal/base	2
Maximum possible number of peaks within a single EIC	3
Peak intensity: use peak area or peak intensoid?	Intensoid (max. int)
Advanced settings	
How often can peak detection fail to end the recursion?	1
Percentage of low-intensity data point do discard	0
Tolerances	
m/z tolerance	2
…given in	ppm
Max. RT deviation between peaks of the same analyte [s]	1.5
Intensity tolerance[%]	30
Mass recalibration	
Include mass recalibration for positive/negative ion mode files?	yes
Reference compounds	Internal standards
± m/z tolerance	1
Maximum allowable m/z correction	1
… given in	Absolute [mmu]
RT tolerance [s]	30
Advanced settings	
Only plot but do not apply recalibration results?	no
Alignment	
Include retention time alignment for positive/negative ion mode files?	Yes
Only plot but do not apply alignment results	Yes
+/- m/z tolerance for peak matches	3
… given in:	ppm
Advanced settings	
Reference peaks/masses	All peaks (recommended)
Maximum permissible RT shift correction [s]	30
Maximum number of most intense reference peaks to include	1000
Maximum number of iteration for match window adaption	4
Only include replicable peaks?	no
Blind	
Factor by which the sample peak intensity must exceed he blank not to be subtracted	100
+/- *m/z* tolerance	3
… given in:	ppm
RT tolerance [s]	30

Select blind/blank files for subtraction	Subtract with the next blank/blind file preceding a sample by its date & time
Screening	
IS	
RT tolerance of peaks relative to their expected RT [s]	30
Screen for MS2 fragments	FALSE
Restrict screening to latest files (covered during profiling) Include?	FALSE
Cutoff score [0,1]	0.8
Targets & Suspects	
RT tolerance of peaks relative to their expected RT [s]	30
Screen for MS2 fragments	FALSE
Restrict screening to latest files (covered during profiling) Include?	FALSE
Cutoff score [0,1]	0.8
Adducts	
Positive ions:	M+H, M+NH4, M+Na, M+K
Negative ions:	M-H, M-
Normalization	
Include normalization for positive/negative ion mode files?	Yes
Minimum of screened files covered by each IS profile [%]	95
Screening threshold	0.8
Minimum number of IS profile peaks	15
Show median deviation of blank/blind profiles	Yes
Use subsampling	Yes
Number of blank/blind profiles in subsample	250
Show median deviation of sample (i.e. non-blank) profiles?	Yes
Use subsampling	Yes
Number of profiles in subsample	250
Profiles	
Algorithm	Fine-tuned
Maximum number of newest files to be processed per ion mode	2300
Peak mass deviation within profiles: *m/z* tolerance	3
… given in	ppm
Peak deviation within profiles: RT tolerance [s]	60
Omit files with table entry profiled = FALSE?	TRUE
File-wise componentization	
Isotopoloque grouping	
Run atom bound estimation	FALSE
Adduct grouping	
Positive mode:	M+H, M+NH4, M+Na, M+K

Negative mode	M-H, M-, 2M-H
Peak shape correlation	
Min. number of MS1 scans over which peak pairs co-elute for their peak shape correlation:	10
Min. Spearman correlation [0,1] coefficient	0.95
Profile componentization	
Restrict profile componentization to a set of latest files only?	FALSE
Filter positive/negative mode components?	TRUE
Allow searching for additional adduct for peak shape correlated profiles?	TRUE
Restrict profile componentization to isotopologue and selected adduct relations only?	FALSE
Restrict profile componentization to top 100 most intense profiles only	FALSE

Results and Discussion

SI 8: Table S5 Quantification Quality Control

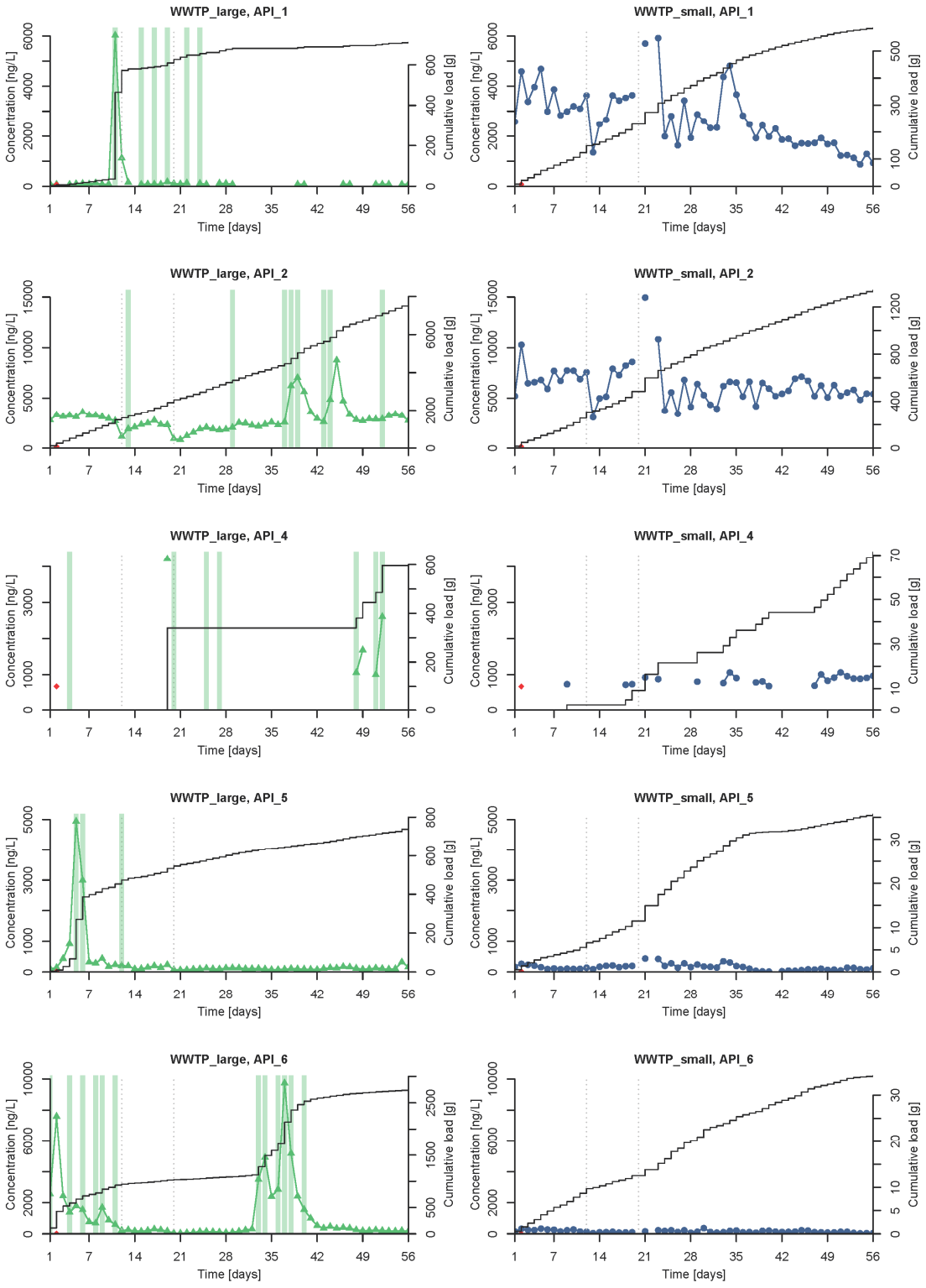
See separate file Table_S5_Quality_control in EXCEL workbook (es0c05178_si_002.xlsx).

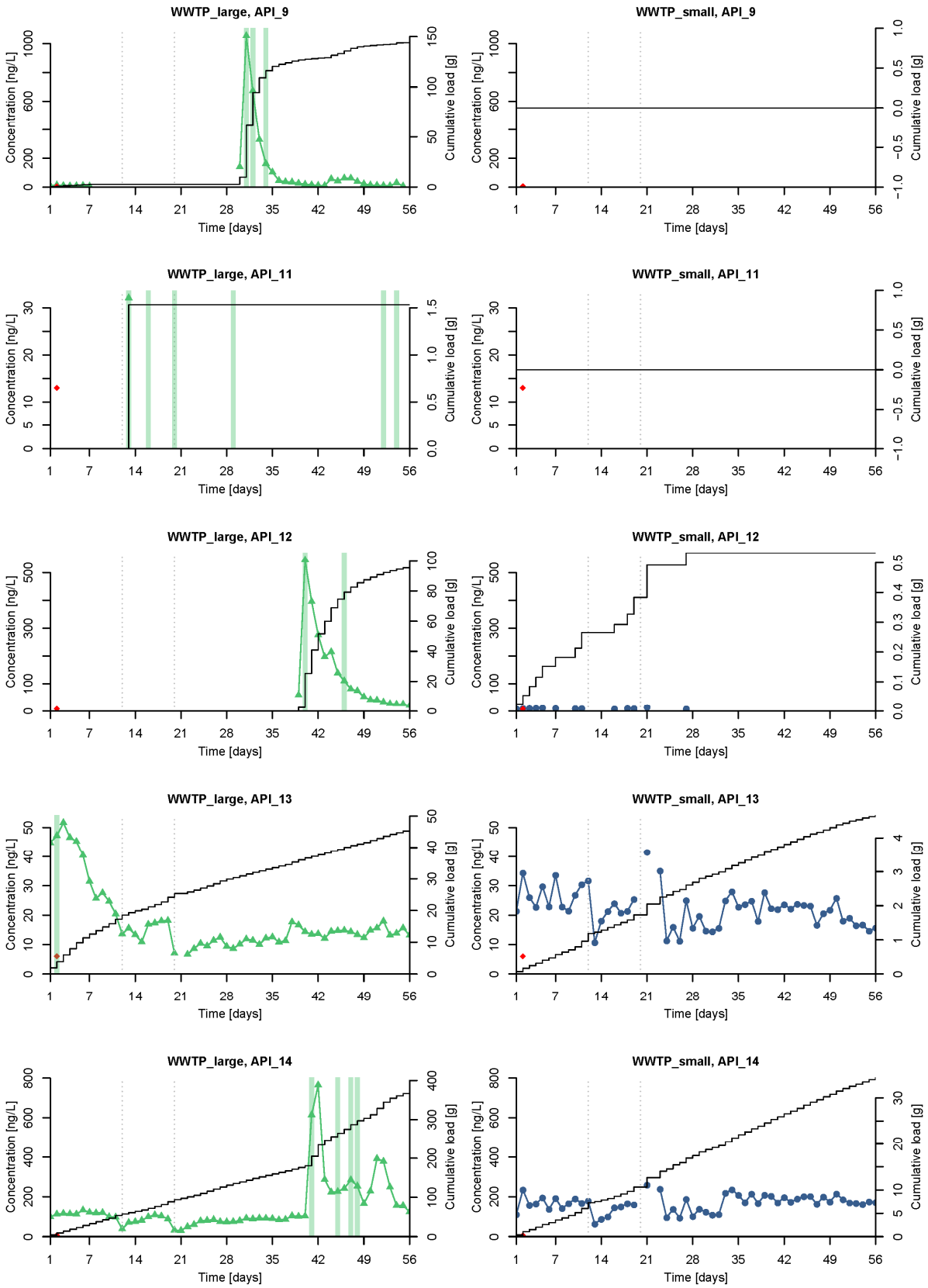
SI 9: Table S6 Quantification Results

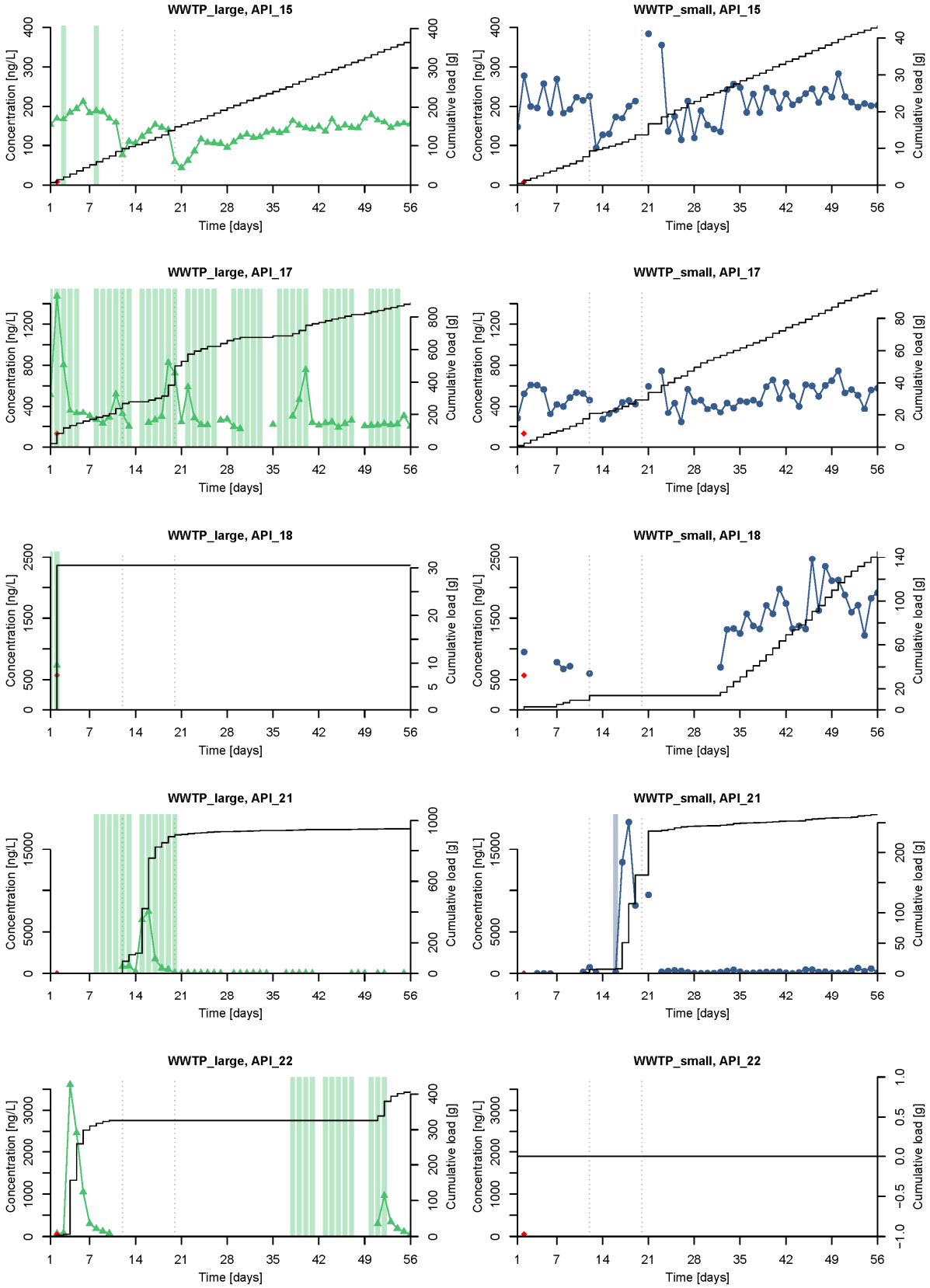
See separate file Table_S6_Quant_results in EXCEL workbook (es0c05178_si_002.xlsx).

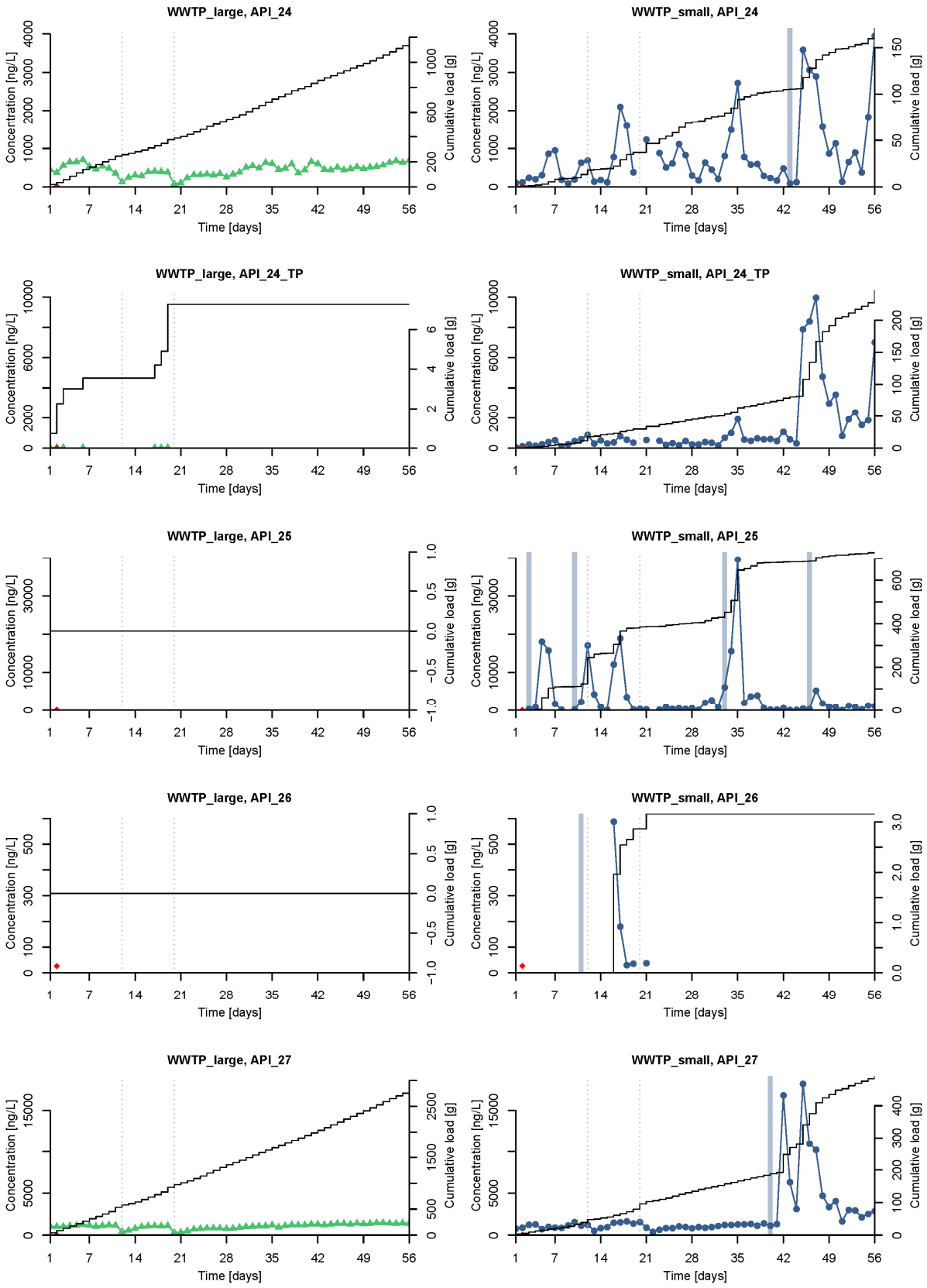
SI 10: Time Series of Quantified Compounds

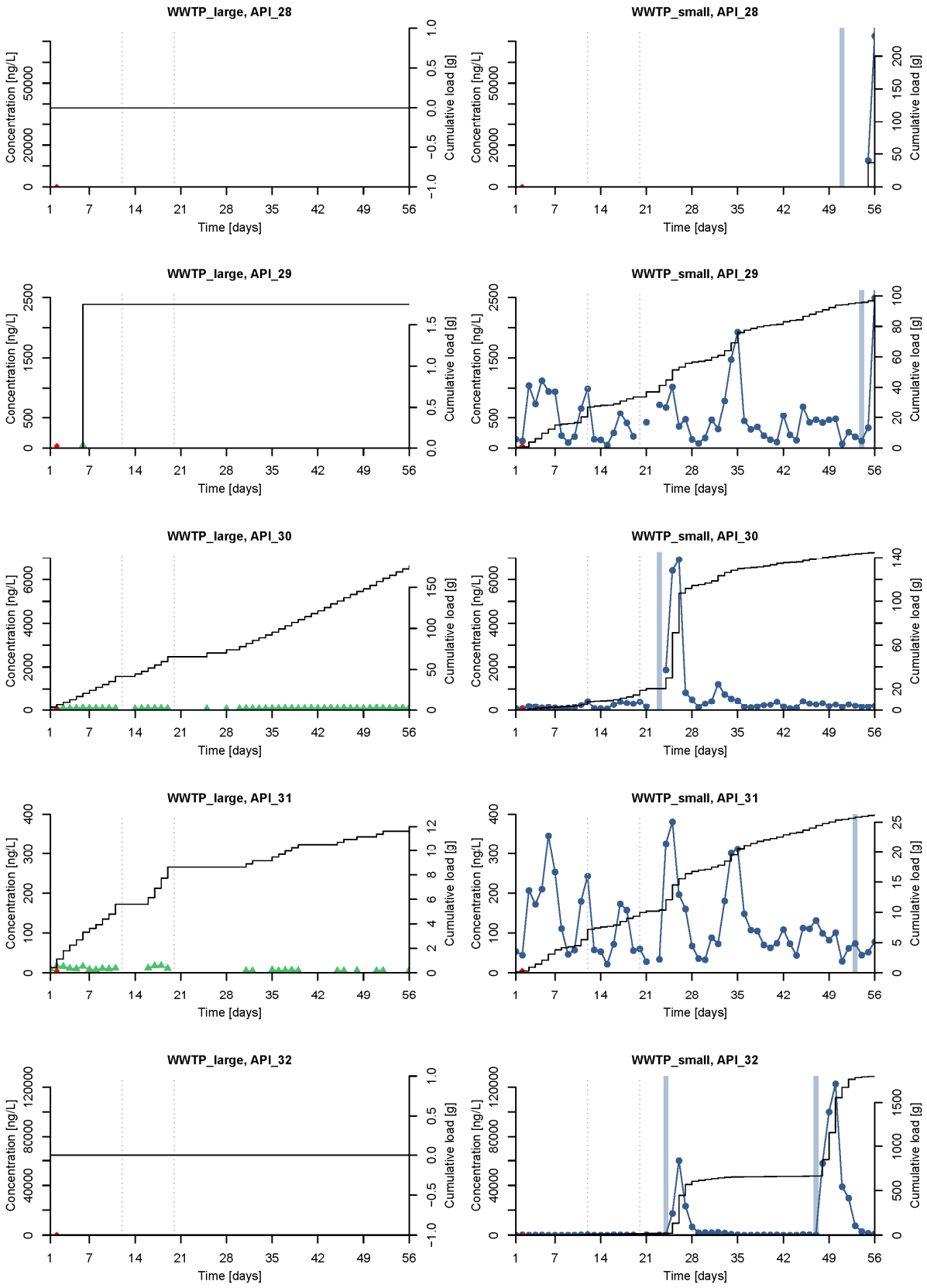
Active Ingredients Processed by Formulating Pharmaceutical Industries




WWTP_large, API_9
WWTP_small, API_9
WWTP_large, API_11
WWTP_small, API_11
WWTP_large, API_12
WWTP_small, API_12
WWTP_large, API_13
WWTP_small, API_13
WWTP_large, API_14
WWTP_small, API_14
Concentration [ng/L]
Cumulative load [g]
Time [days]


WWTP_large, API_15
WWTP_small, API_15
WWTP_large, API_17
WWTP_small, API_17
WWTP_large, API_18
WWTP_small, API_18
WWTP_large, API_21
WWTP_small, API_21
WWTP_large, API_22
WWTP_small, API_22
Concentration [ng/L]
Cumulative load [g]
Time [days]


WWTP_large, API_24
WWTP_small, API_24
WWTP_large, API_24_TP
WWTP_small, API_24_TP
WWTP_large, API_25
WWTP_small, API_25
WWTP_large, API_26
WWTP_small, API_26
WWTP_large, API_27
WWTP_small, API_27
Concentration [ng/L]
Cumulative load [g]
Time [days]


WWTP_large, API_28
WWTP_small, API_28
WWTP_large, API_29
WWTP_small, API_29
WWTP_large, API_30
WWTP_small, API_30
WWTP_large, API_31
WWTP_small, API_31
WWTP_large, API_32
WWTP_small, API_32
Concentration [ng/L]
Cumulative load [g]
Time [days]

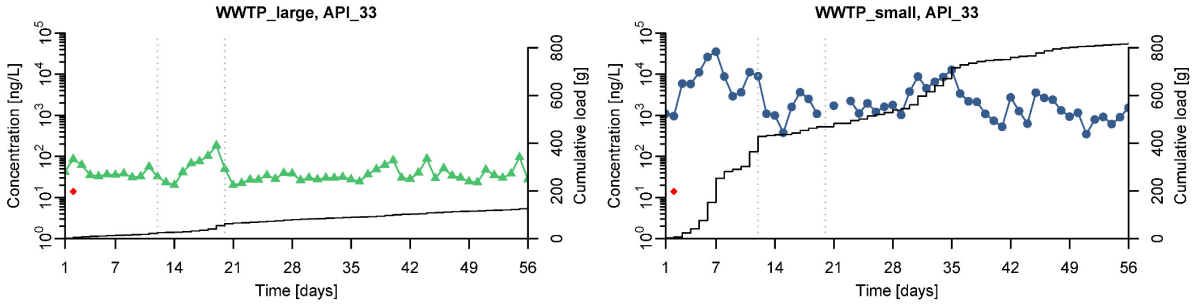
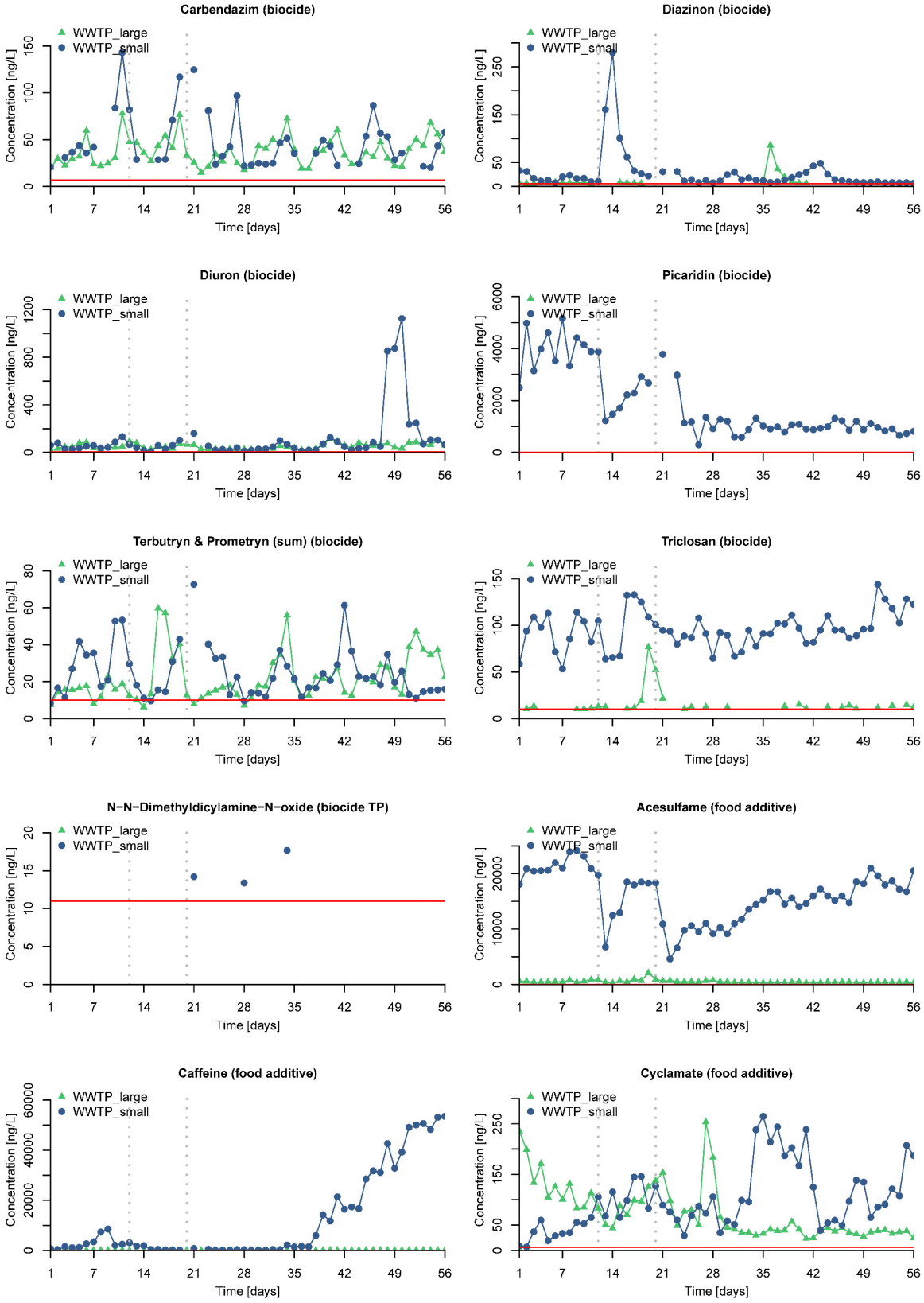
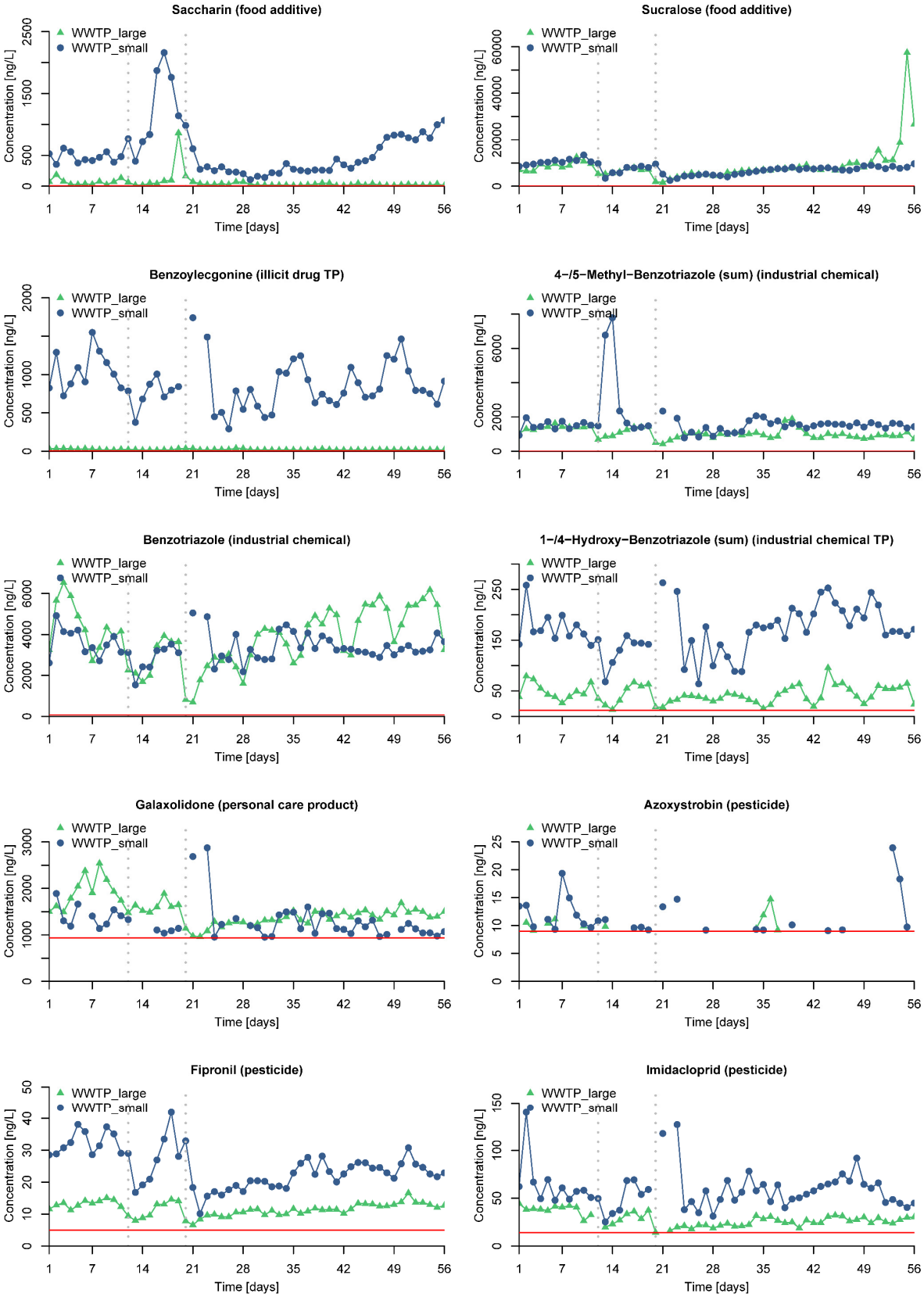
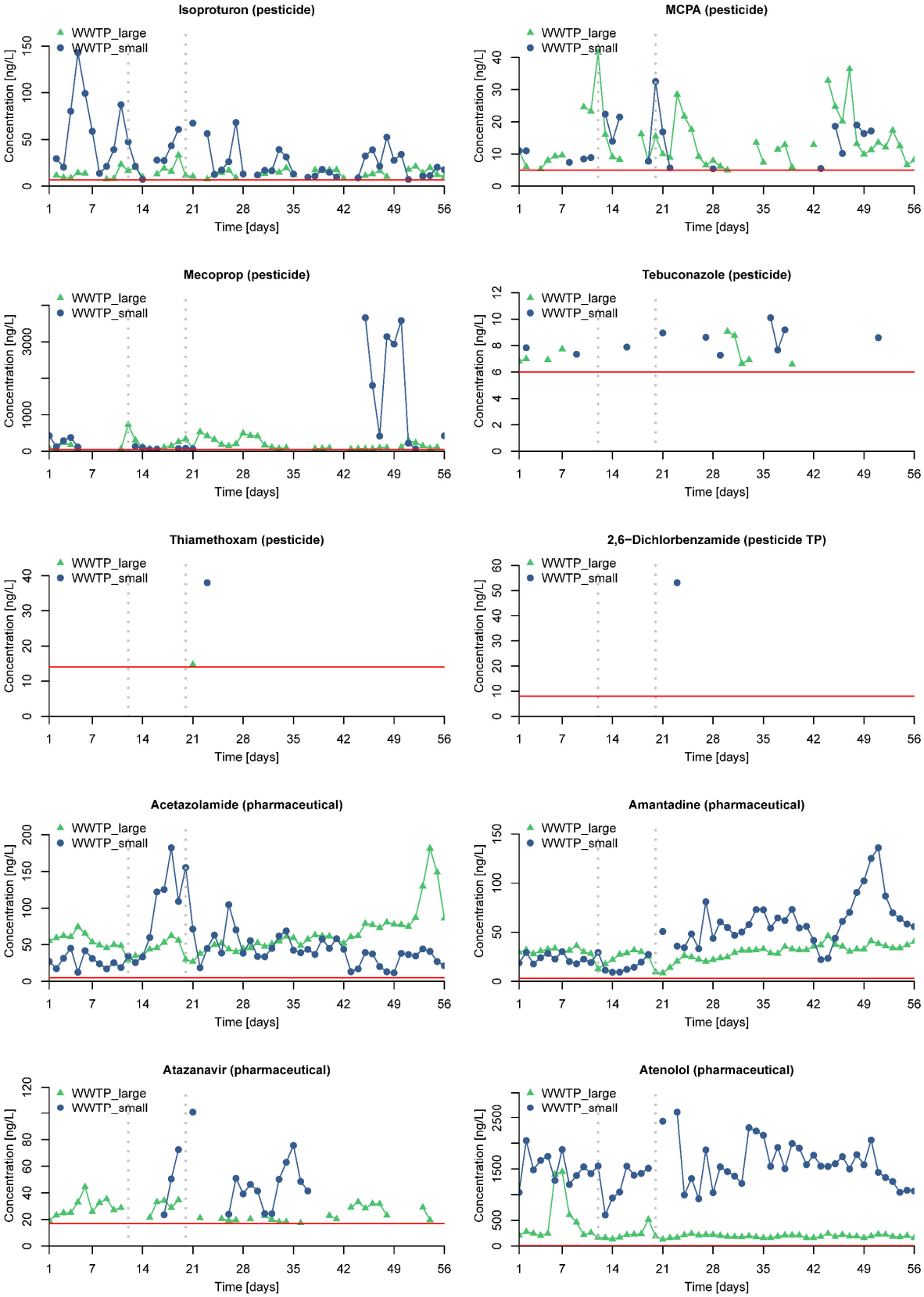


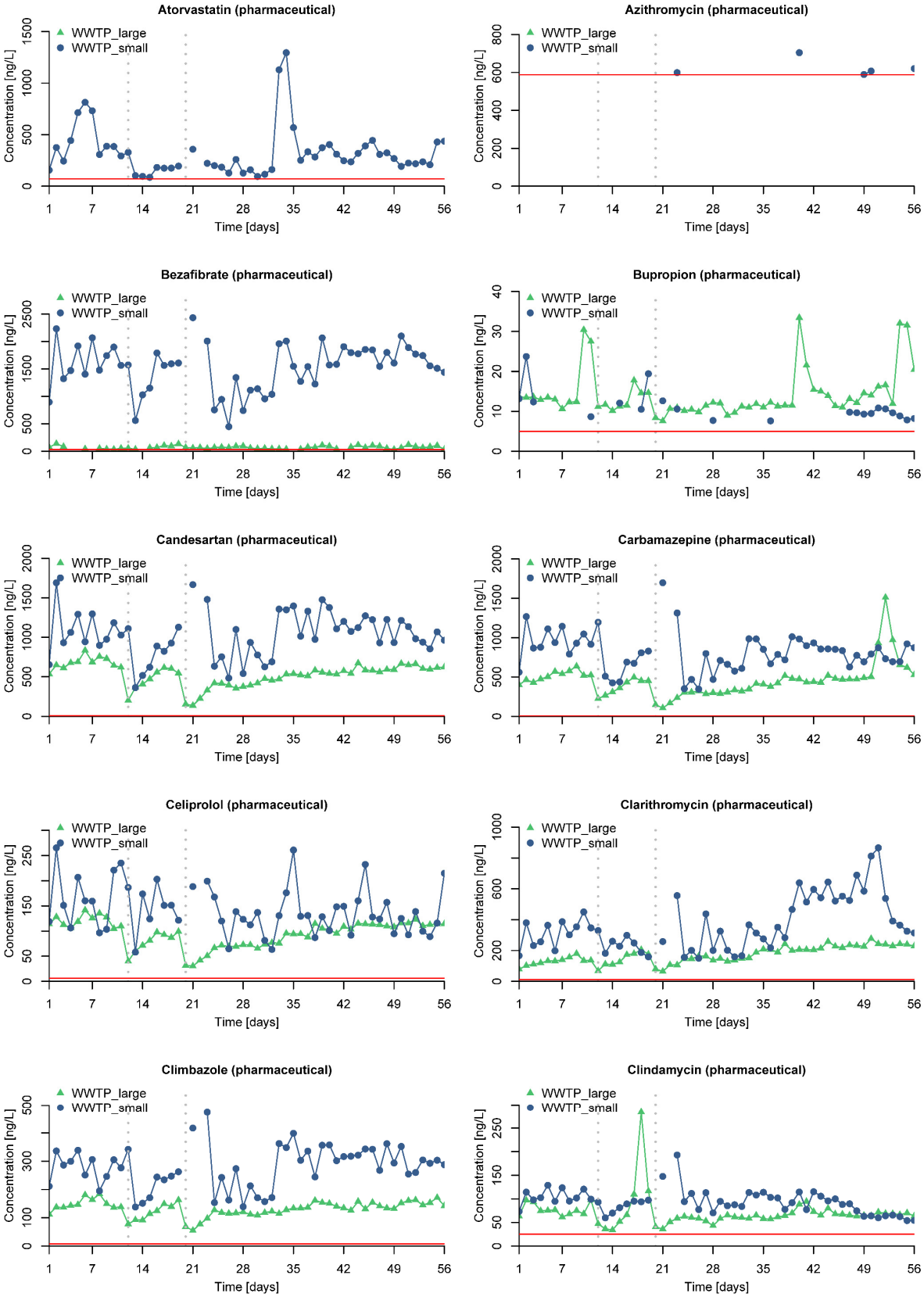
Figure S1: Concentration time profiles of active pharmaceutical ingredients (APIs) processed by the formulating pharmaceutical industries (FPIs) during the sampling campaign. For each compound, the time profile detected in the effluents of WWTP_large is shown on the right and the one detected in WWTP_small on the left. Colored lines with symbols indicate concentrations, colored vertical bars (green for WWTP_large and blue for WWTP_small) mark days at which the respective API was processed by a FPI, the black stair plots display the cumulative loads, the red diamonds on the left mark the limit of quantification and the vertical dotted grey lines mark days of maximal discharge caused by two rain events. Missing data points are due to concentration values below the LOQ. At WWTP_small data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

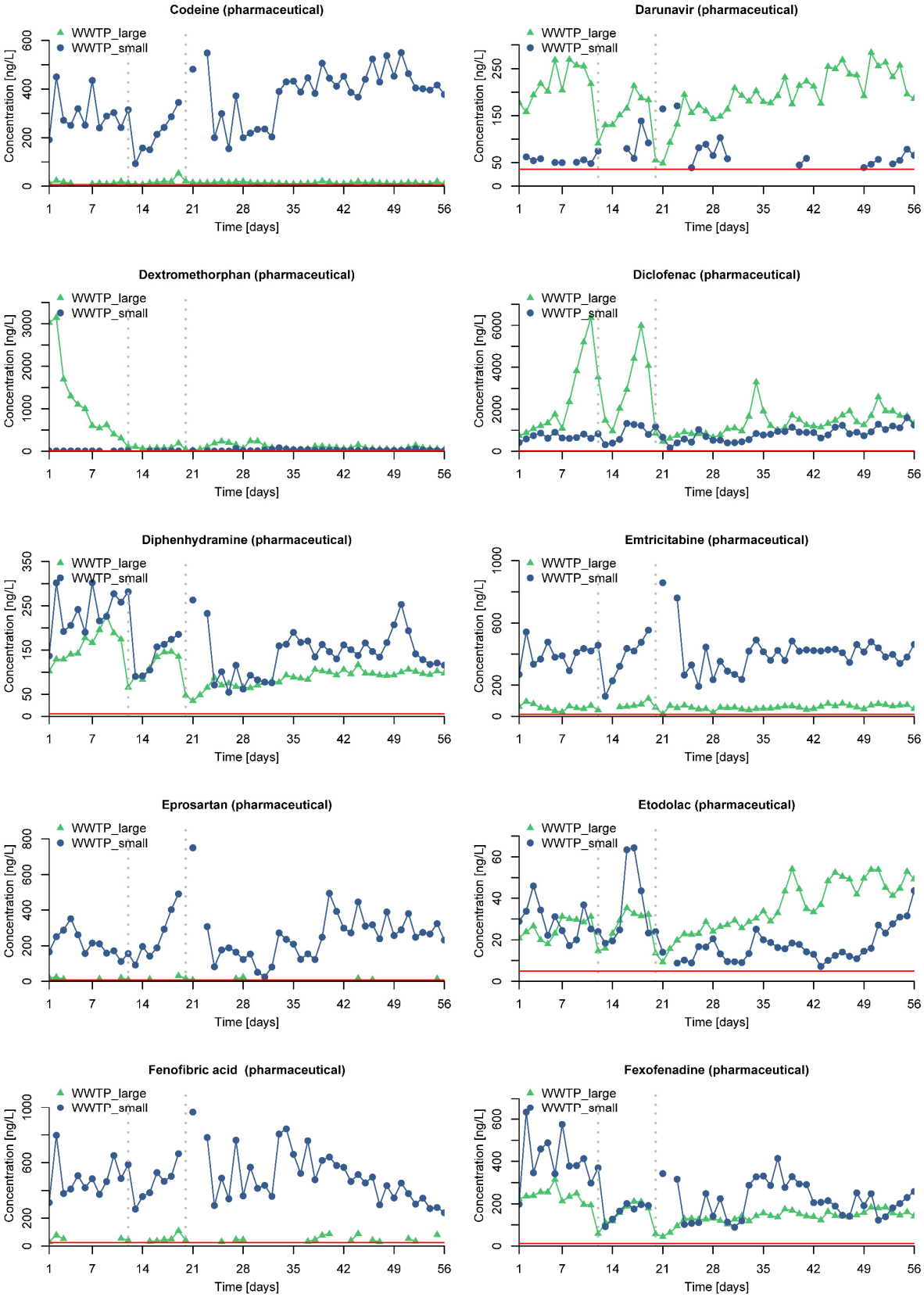
Compounds Commonly Detected in Domestic WWTP Effluent

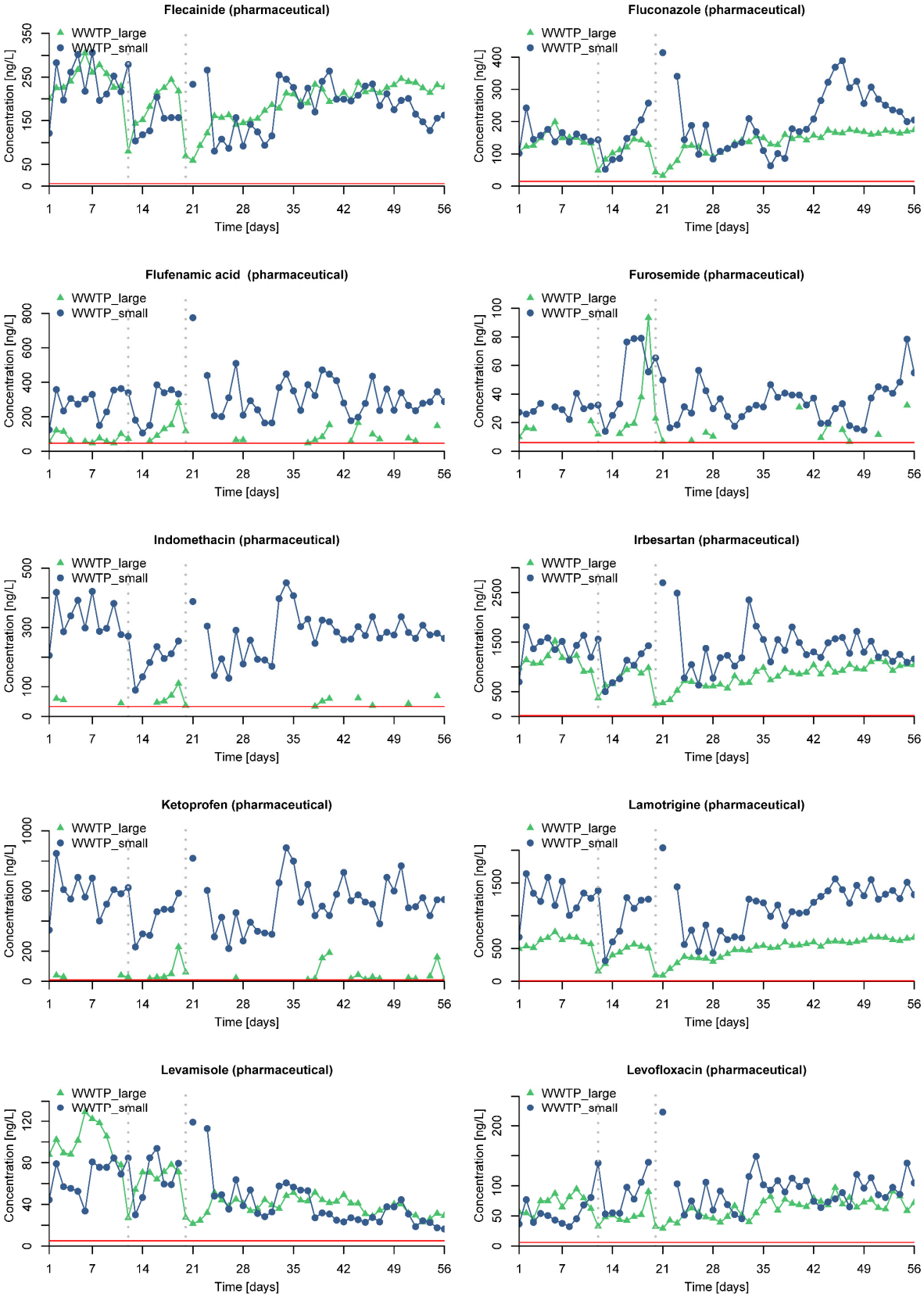


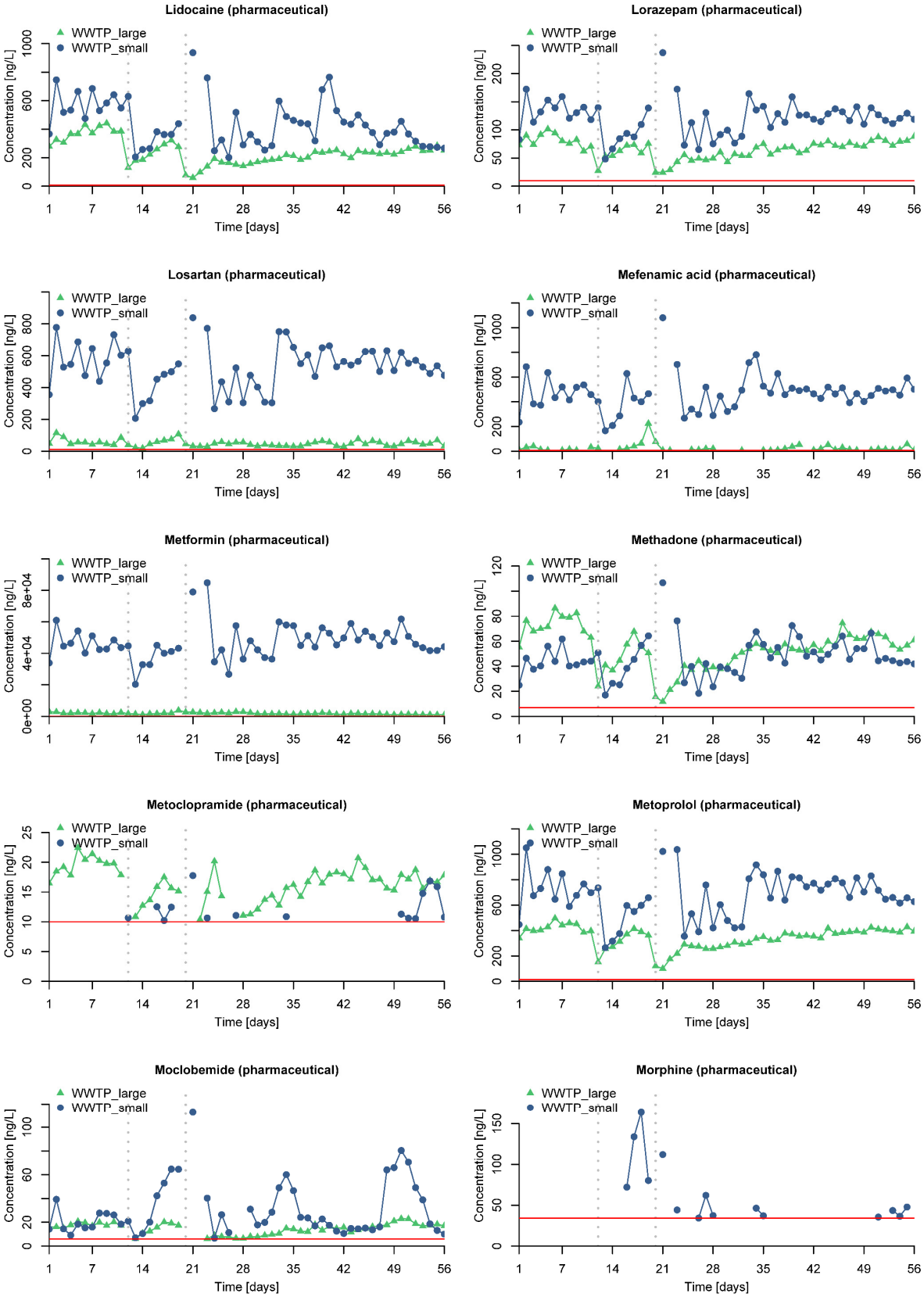

Saccharin (food additive)
Sucralose (food additive)
Benzoylecgonine (illicit drug TP)
4−/5−Methyl−Benzotriazole (sum) (industrial chemical)
Benzotriazole (industrial chemical)
1−/4−Hydroxy−Benzotriazole (sum) (industrial chemical TP)
Galaxolidone (personal care product)
Azoxystrobin (pesticide)
Fipronil (pesticide)
Imidacloprid (pesticide)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

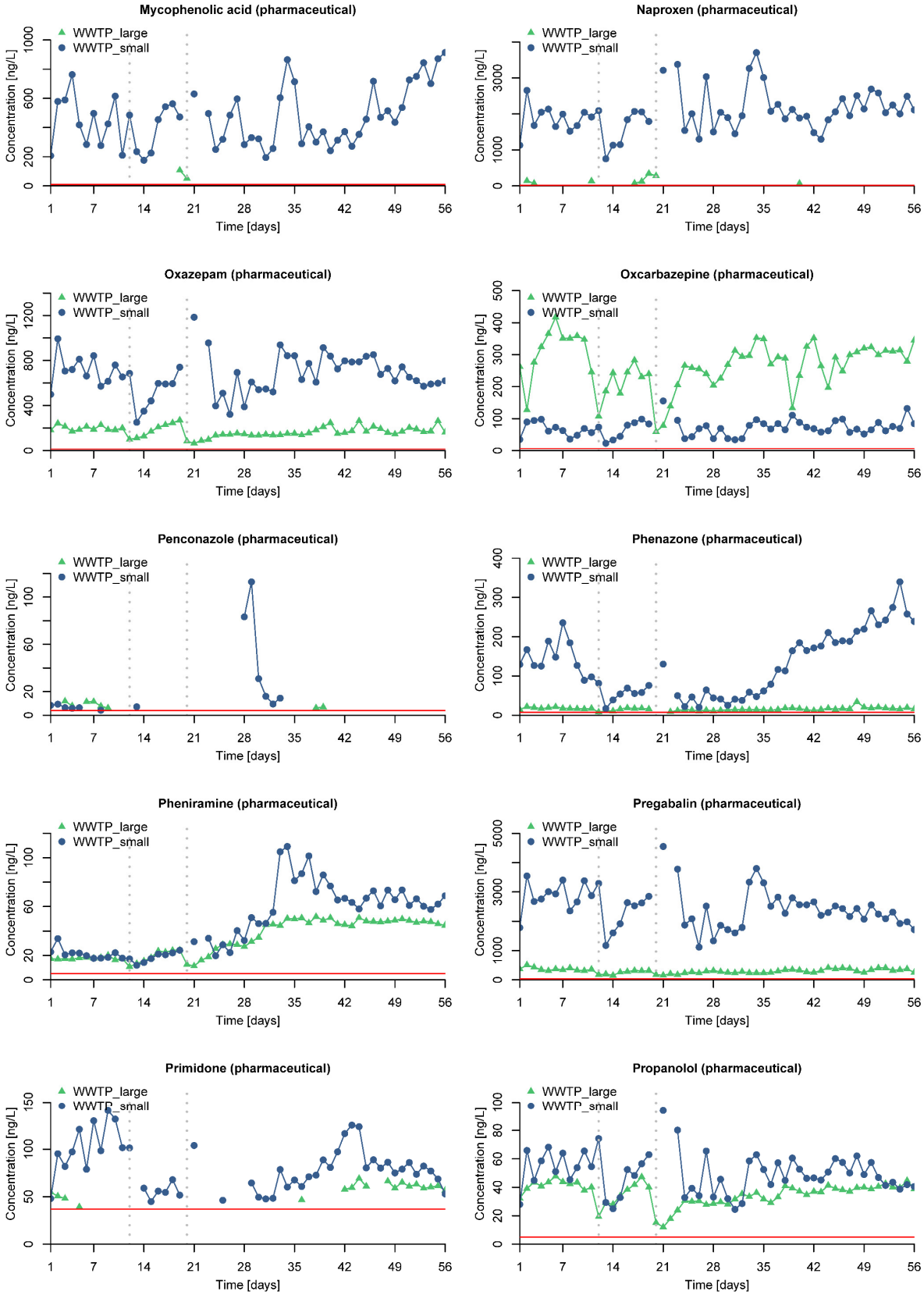

Isoproturon (pesticide)
MCPA (pesticide)
Mecoprop (pesticide)
Tebuconazole (pesticide)
Thiamethoxam (pesticide)
2,6−Dichlorbenzamide (pesticide TP)
Acetazolamide (pharmaceutical)
Amantadine (pharmaceutical)
Atazanavir (pharmaceutical)
Atenolol (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

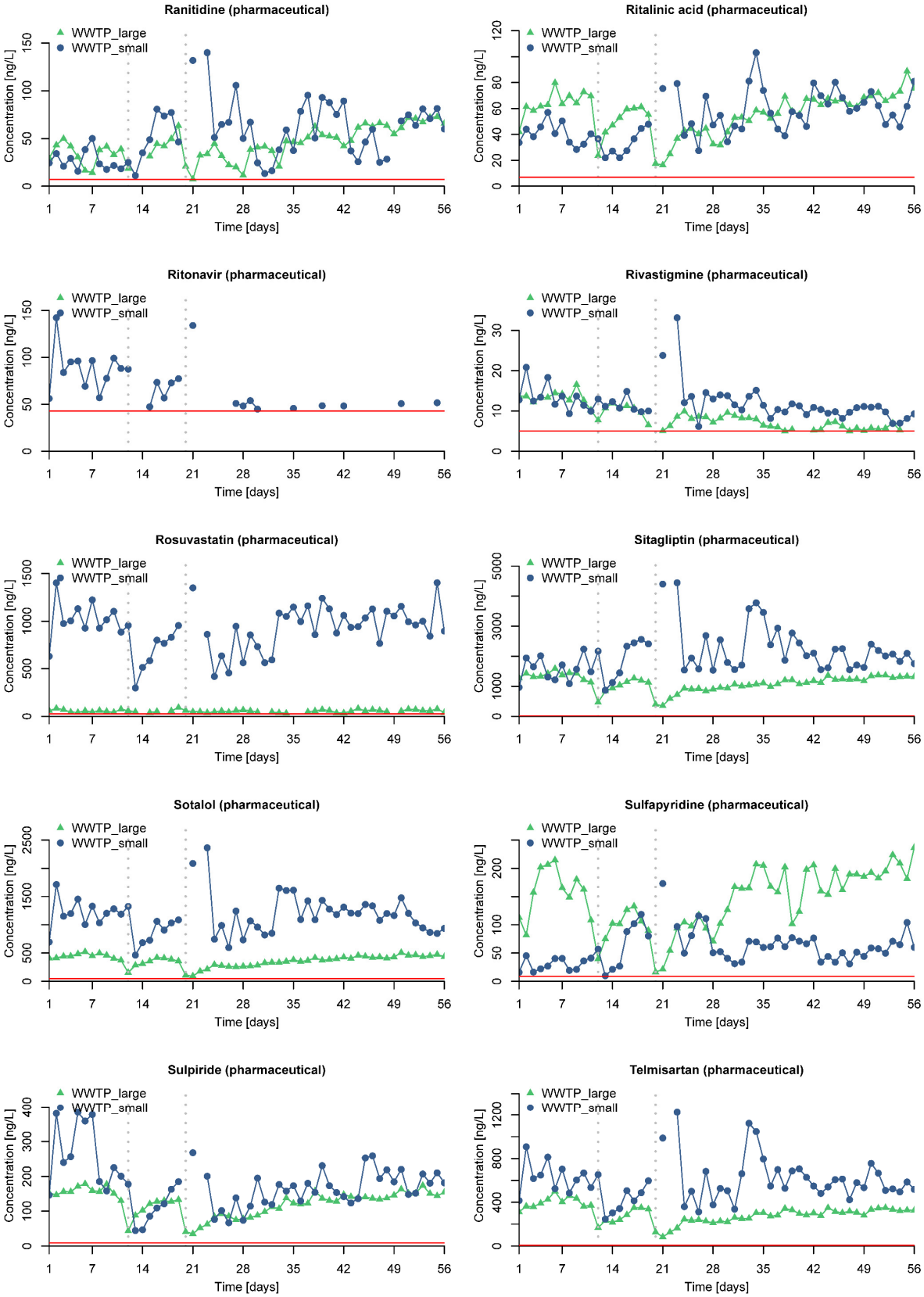

Atorvastatin (pharmaceutical)
Azithromycin (pharmaceutical)
Bezafibrate (pharmaceutical)
Bupropion (pharmaceutical)
Candesartan (pharmaceutical)
Carbamazepine (pharmaceutical)
Celiprolol (pharmaceutical)
Clarithromycin (pharmaceutical)
Climbazole (pharmaceutical)
Clindamycin (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

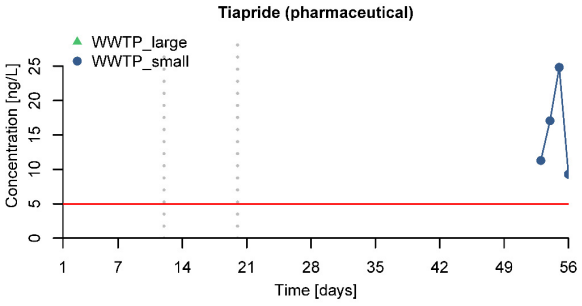

Codeine (pharmaceutical)
Darunavir (pharmaceutical)
Dextromethorphan (pharmaceutical)
Diclofenac (pharmaceutical)
Diphenhydramine (pharmaceutical)
Emtricitabine (pharmaceutical)
Eprosartan (pharmaceutical)
Etodolac (pharmaceutical)
Fenofibric acid (pharmaceutical)
Fexofenadine (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

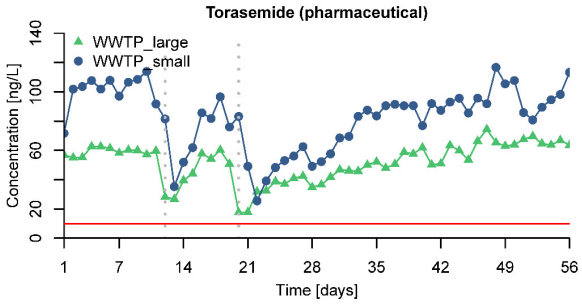

Flecainide (pharmaceutical)
Fluconazole (pharmaceutical)
Flufenamic acid (pharmaceutical)
Furosemide (pharmaceutical)
Indomethacin (pharmaceutical)
Irbesartan (pharmaceutical)
Ketoprofen (pharmaceutical)
Lamotrigine (pharmaceutical)
Levamisole (pharmaceutical)
Levofloxacin (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

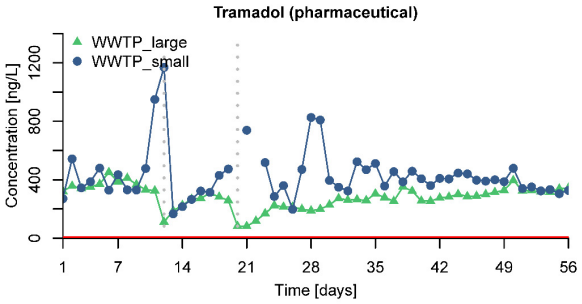

Lidocaine (pharmaceutical)
Lorazepam (pharmaceutical)
Losartan (pharmaceutical)
Mefenamic acid (pharmaceutical)
Metformin (pharmaceutical)
Methadone (pharmaceutical)
Metoclopramide (pharmaceutical)
Metoprolol (pharmaceutical)
Moclobemide (pharmaceutical)
Morphine (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

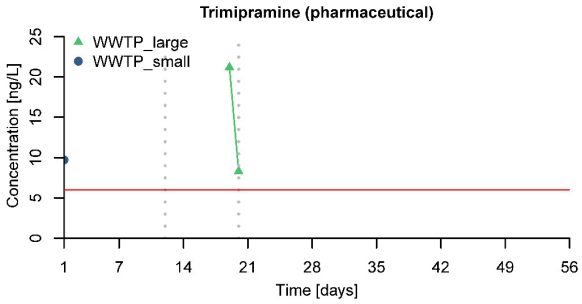

Mycophenolic acid (pharmaceutical)
Naproxen (pharmaceutical)
Oxazepam (pharmaceutical)
Oxcarbazepine (pharmaceutical)
Penconazole (pharmaceutical)
Phenazone (pharmaceutical)
Pheniramine (pharmaceutical)
Pregabalin (pharmaceutical)
Primidone (pharmaceutical)
Propanolol (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

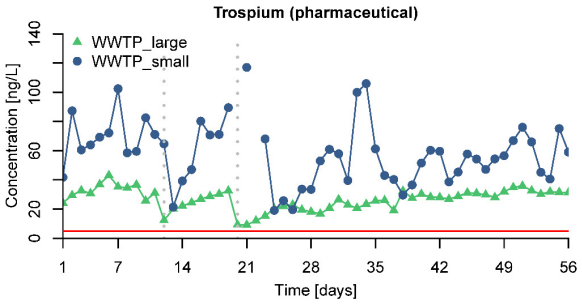

Ranitidine (pharmaceutical)
Ritalinic acid (pharmaceutical)
Ritonavir (pharmaceutical)
Rivastigmine (pharmaceutical)
Rosuvastatin (pharmaceutical)
Sitagliptin (pharmaceutical)
Sotalol (pharmaceutical)
Sulfapyridine (pharmaceutical)
Sulpiride (pharmaceutical)
Telmisartan (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
Time [days]

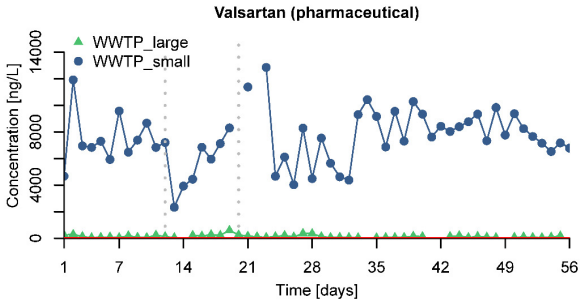

Tiapride (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 5 10 15 20 25
1 7 14 21 28 35 42 49 56
Time [days]

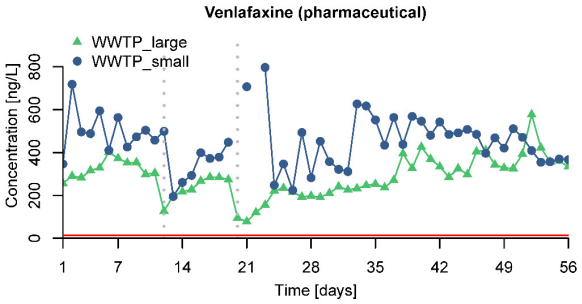

Torasemide (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 20 60 100 140
1 7 14 21 28 35 42 49 56
Time [days]

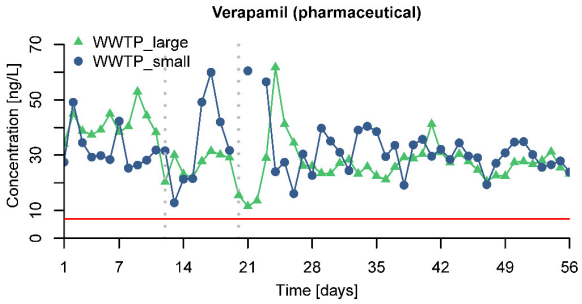

Tramadol (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 400 800 1200
1 7 14 21 28 35 42 49 56
Time [days]

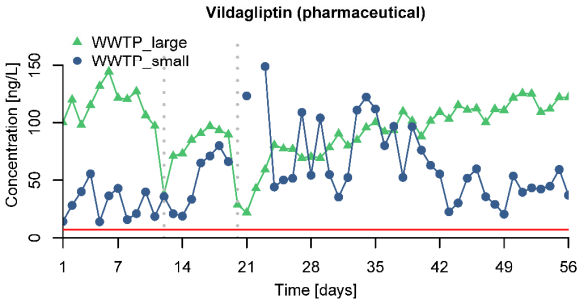

Trimipramine (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 5 10 15 20 25
1 7 14 21 28 35 42 49 56
Time [days]

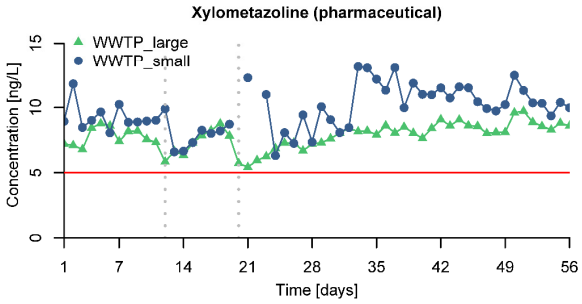

Trospium (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 20 60 100 140
1 7 14 21 28 35 42 49 56
Time [days]

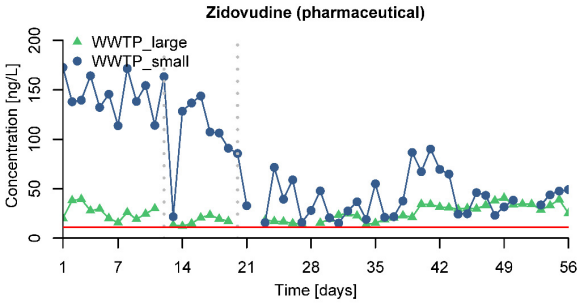

Valsartan (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 4000 8000 14000
1 7 14 21 28 35 42 49 56
Time [days]

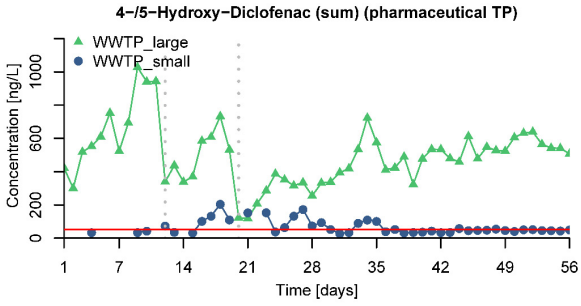

Venlafaxine (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 200 400 600 800
1 7 14 21 28 35 42 49 56
Time [days]

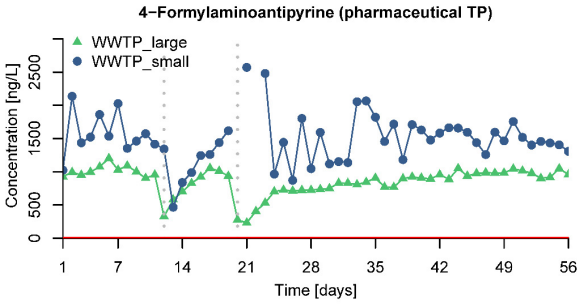

Verapamil (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 10 30 50 70
1 7 14 21 28 35 42 49 56
Time [days]

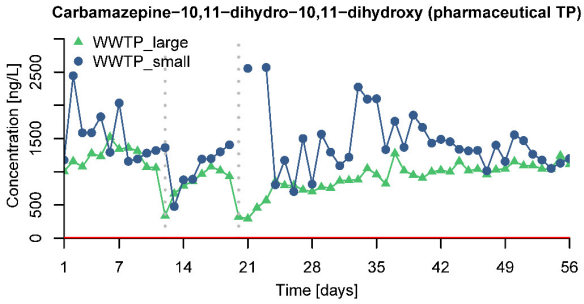

Vildagliptin (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 50 100 150
1 7 14 21 28 35 42 49 56
Time [days]

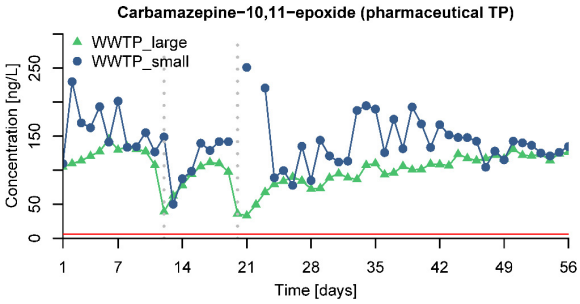

Xylometazoline (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0 5 10 15
1 7 14 21 28 35 42 49 56
Time [days]

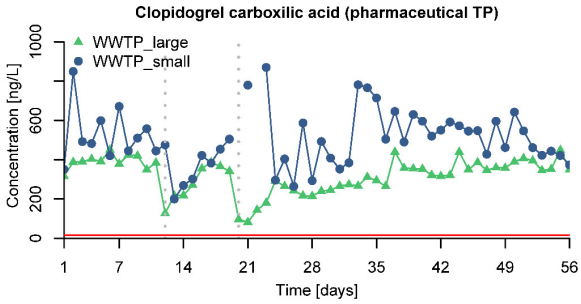

Zidovudine (pharmaceutical)
WWTP_large
WWTP_small
Concentration [ng/L]
0
50
100
150
200
1
7
14
21
28
35
42
49
56
Time [days]

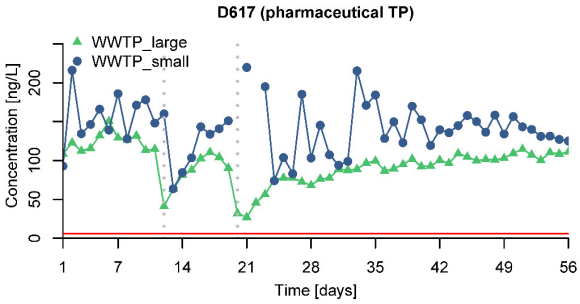

4−/5−Hydroxy−Diclofenac (sum) (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
200
600
1000
1
7
14
21
28
35
42
49
56
Time [days]

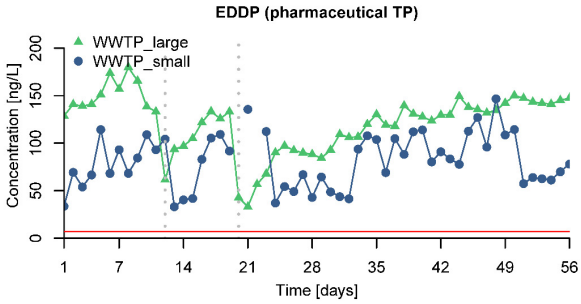

4−Formylaminoantipyrine (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
500
1500
2500
1
7
14
21
28
35
42
49
56
Time [days]

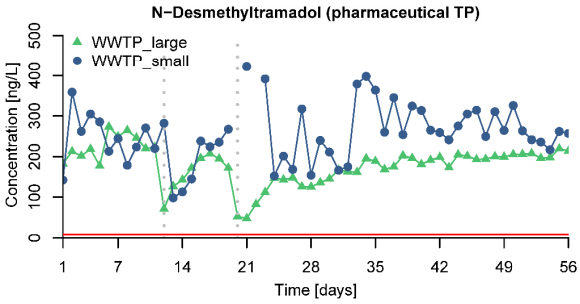

Carbamazepine−10,11−dihydro−10,11−dihydroxy (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
500
1500
2500
1
7
14
21
28
35
42
49
56
Time [days]

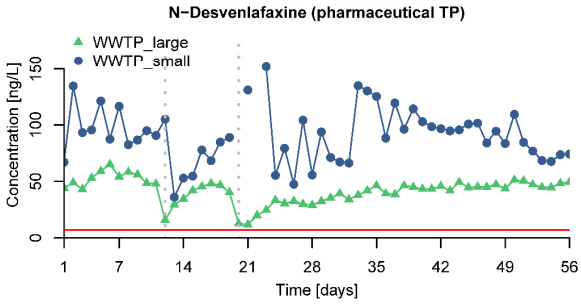

Carbamazepine−10,11−epoxide (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
50
150
250
1
7
14
21
28
35
42
49
56
Time [days]


Clopidogrel carboxilic acid (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
200
600
1000
1
7
14
21
28
35
42
49
56
Time [days]


D617 (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
50
100
200
1
7
14
21
28
35
42
49
56
Time [days]


EDDP (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
50
100
150
200
1
7
14
21
28
35
42
49
56
Time [days]


N−Desmethyltramadol (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
100
200
300
400
500
1
7
14
21
28
35
42
49
56
Time [days]


N−Desvenlafaxine (pharmaceutical TP)
WWTP_large
WWTP_small
Concentration [ng/L]
0
50
100
150
1
7
14
21
28
35
42
49
56
Time [days]

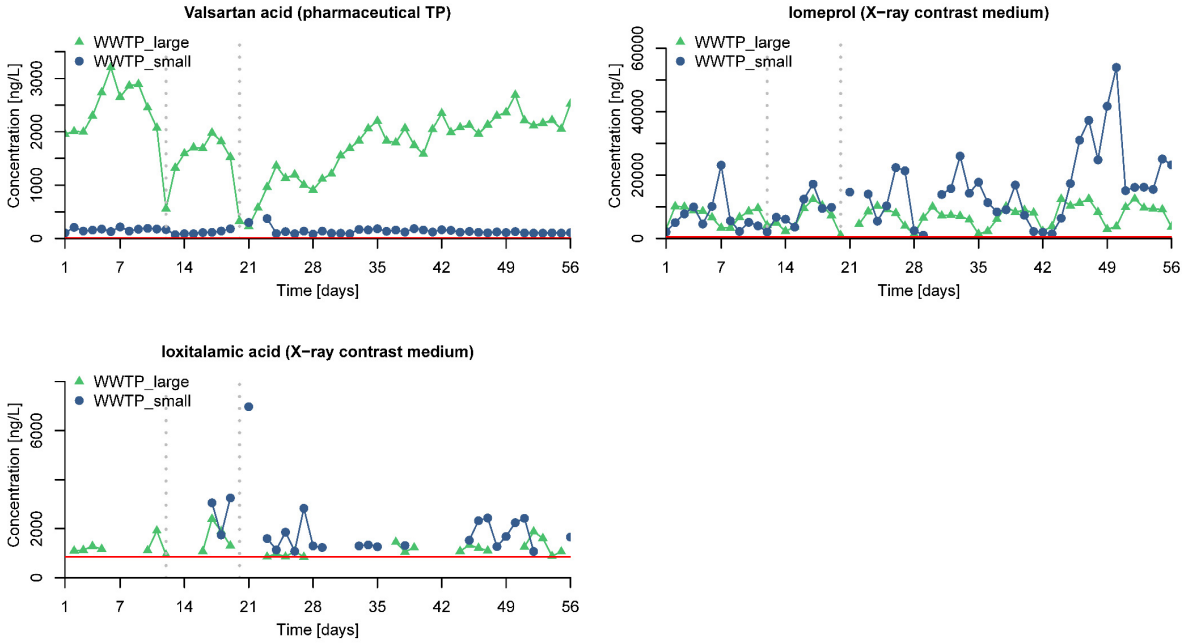


Figure S2: Concentration time profiles of compounds commonly detected in WWTP effluent sorted by compound class and alphabetical order, in green, the data of WWTP_large are shown and in blue, the data of WWTP_small. The horizontal red lines indicate the limit of quantification and the vertical dotted grey lines mark days of maximal discharge caused by two rain events. Missing data points are due to concentration values below the LOQ. At WWTP_small data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

SI 11: Time Series of APIs Formulated by Industry Detected in Suspect Screening

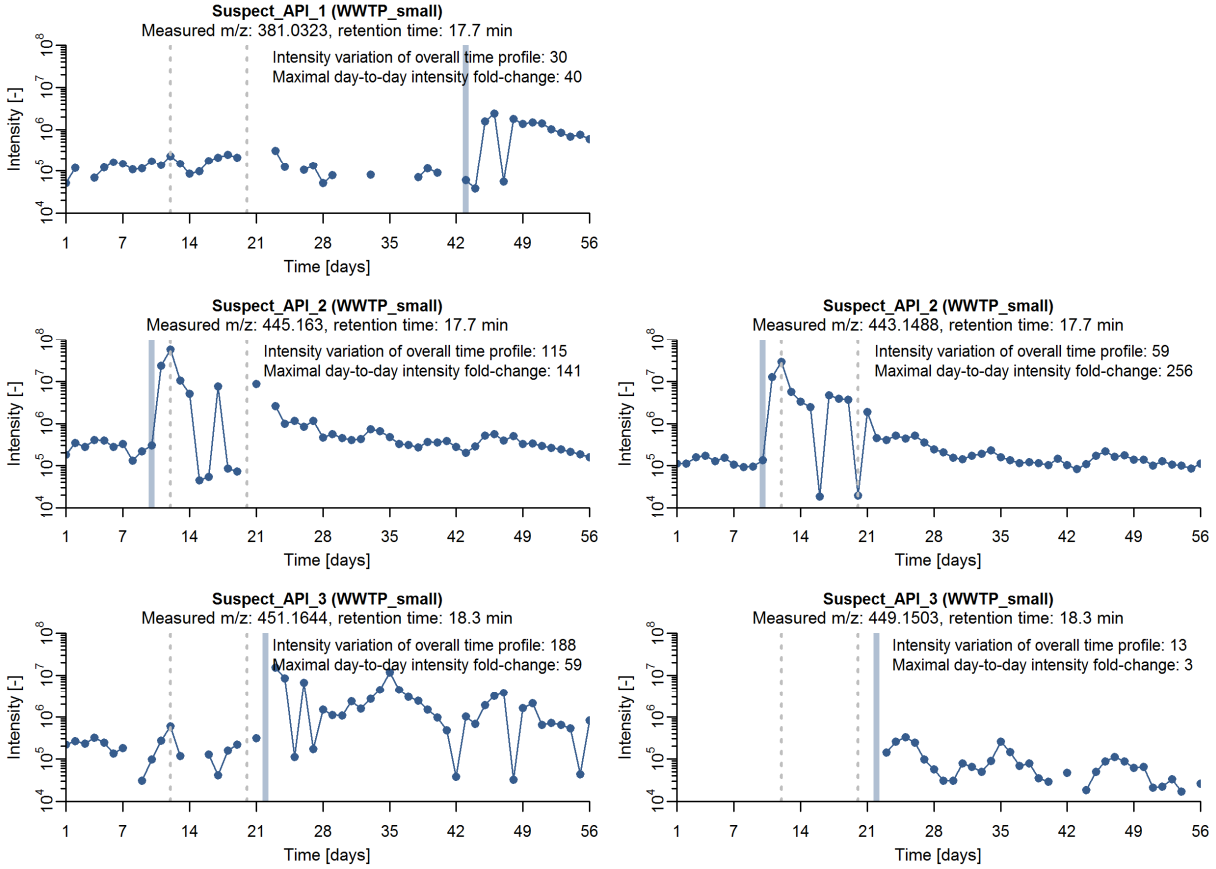


Figure S3: Intensity time profiles of APIs, for which industrial discharges were detected in the suspect screening at WWTP_small. At WWTP_large none of the suspect APIs displayed an industrial emission pattern. Blue bars mark days on which the respective compound was processed by a formulating pharmaceutical industry (FPI) in the catchment and the vertical dotted grey lines indicate days of maximal WWTP discharge caused by two rain events. Please note the logarithmic scale of the intensity (y-axis). The intensity variation of the overall profile and the maximal day-to-day intensity fold change, indicated in the legend, were used to allocate effluent emission time patterns to FPI wastewater (see SI 12). API_2 and API_3 were detected in both positive and negative ESI mode; therefore, two profiles are show for these compounds. API_1 was confirmed with an authentic reference standard (*i.e.*, level 1 according Schymanski, et al. [2]); however, the applied analytical method was not suitable for quantification. For API_2 and API_3 no reference standards were available, for API_2 matching library spectra were found (*i.e.,* identification level 2a), whereas API_3 is a pharmaceutical still under development for which no library spectra were available (*i.e.*, identification level 2b). Missing data points are due to non-detects. Data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

SI 12: Source Allocation

Identifying the origin of compounds detected in the WWTP effluents was necessary, because several of the pharmaceuticals processed by the FPIs are also commonly present in household wastewater. Therefore, we analysed the time pattern of the processed active pharmaceutical ingredients (APIs) to assess if industrial emissions were present. As proposed in our previous publication[3], compounds displaying concentration variation in the overall time series > 10-fold were classified as having industrial origins. This classification is based on the assumption that industrial emissions cause concentration fluctuations much larger than those in WWTP discharge, which was 2.5 and 2 for WWTP_small and WWTP_large, respectively. Applying the time series variation criterion, 54% (*i.e.*, 14 of 26) of the detected APIs processed by industry were classified as industrial emissions. However, in-depth inspection of the time profiles revealed that several industrial discharges were missed. These misclassifications were related to the following reasons; (i) some of processed APIs were only detected in single samples after manufacturing and at low concentrations and (ii) some showed sharp concentration increases within a single day and returned to the background level almost as quickly. In case of such very short-term emissions, the true fluctuations in the time series were greatly underestimated by the intensity variation criterion, which was not calculated based on maximal measured values, but on the ratio of the 0.95-quantile to the 0.05-quantile. This definition was reasonable, not to give too much weight to potential outlier values, in the context of non-target time pattern analysis, for which the intensity spread criterion had been originally developed. However, in the case of quantitative data, the inclusion of outlier values is much less critical, because each value was checked manually. Hence, here having quantitative data and production data available allowed for two additional criteria to be used to also capture low concentrated industrial contributions and very short-term industrial emissions. These criteria were: day-to-day concentration changes > 5-fold and detection only after reported manufacturing. Applying these refined source allocation criteria led to the detection of 8 additional API time series displaying an industrial emission pattern, *i.e.*, finally for 85% (22 of 26) of the quantified APIs detected in the effluents and processed by the FPIs an industrial contribution could be identified.

We are aware that these allocation criteria have some restrictions. With regard to false negatives, it should be noted that the detectability of industrial emissions not only depended on the actual emission in the WWTP effluents but also on the background level of the respective API in domestic wastewater and on the method LOQ. Moreover, industrial contributions will fail in to be recognized if they are continuous or if there is a large background load of domestic emissions for the specific compound. Regarding false positives, it should be considered that industrial emissions are not the only reason for large overall fluctuations in time series and large changes in daily loads, as discussed in detail in a previous publication[3]. Here, the rate of false positive identifications of industrial emissions was assessed based on the APIs detected in the effluents and not reported to be processed by the FPIs (*i.e.*, >80 at each WWTP) and was calculated to be < 5% (3 APIs at WWTP_large and 5 at WWTP_small). Eventually, the three APIs displaying an industrial emission pattern at WWTP_large were found on the product list of a FPI not involved in our study but located in the catchment of WWTP_large, resulting in an even lower false positive rate.

SI 13: Differences in Degradation Efficiency between the two WWTPs

Concentrations of several pharmaceuticals emitted in domestic wastewater (*e.g.*, bezafibrate, codeine, metformin and valsartan) were about one order of magnitude higher in the effluent of WWTP_small compared to WWTP_large. This is doubtfully due to differences in consumption, since some of these compounds are widely used, but likely a consequence of the better removal efficiency of WWTP_large compared to WWTP_small. Evidence for this comes from the concentrations of valsartan and its transformation product (TP) valsartan acid. That is, in WWTP_small concentrations of the parent compound valsartan were much higher compared to those of TP valsartan acid. Exactly the opposite was observed in WWTP_large, where the concentrations of valsartan acid were much higher than those of valsartan. This strongly suggests that valsartan is readily degraded in WWTP_large but not in WWTP_small, because valsartan acid (2'-(2H-tetrazol-5-yl)-[1,1'-biphenyl]-4-carboxylic acid in the cited literature) is a TP known to be formed during biological treatment in WWTPs[4, 5]. Different degradation efficiencies likely arise from the additional removal steps at WWTP_large, *i.e.*, a nitrification-denitrification step, compared to WWTP_small. In summary, these findings demonstrate the well-known influence of the WWTP degradation efficiency on the effluent contaminant concentrations.

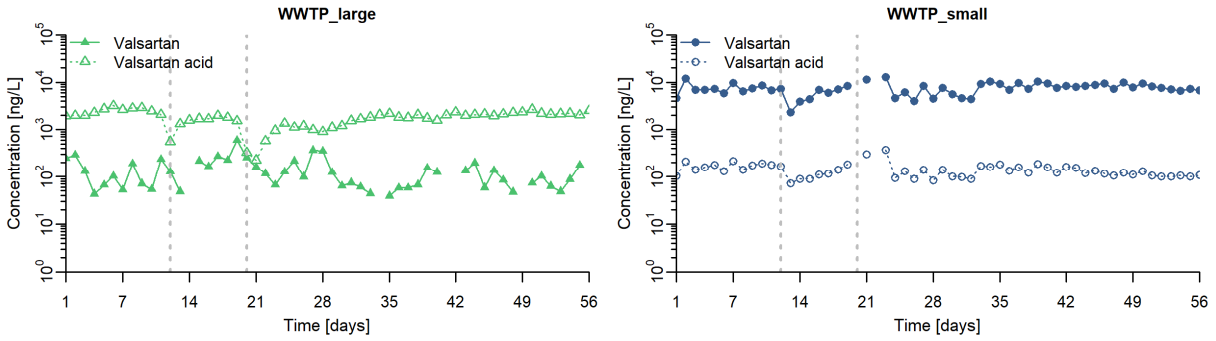


Figure S4: Concentration time series of valsartan and its transformation product valsartan acid at WWTP_large (left) and WWTP_small (right). The vertical dotted grey lines indicate days of maximal WWTP discharge. Please note the logarithmic scale of the y-axis. Missing data points are due to concentration values below the LOQ. At WWTP_small data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

SI 14: Estimation of Loss Factors

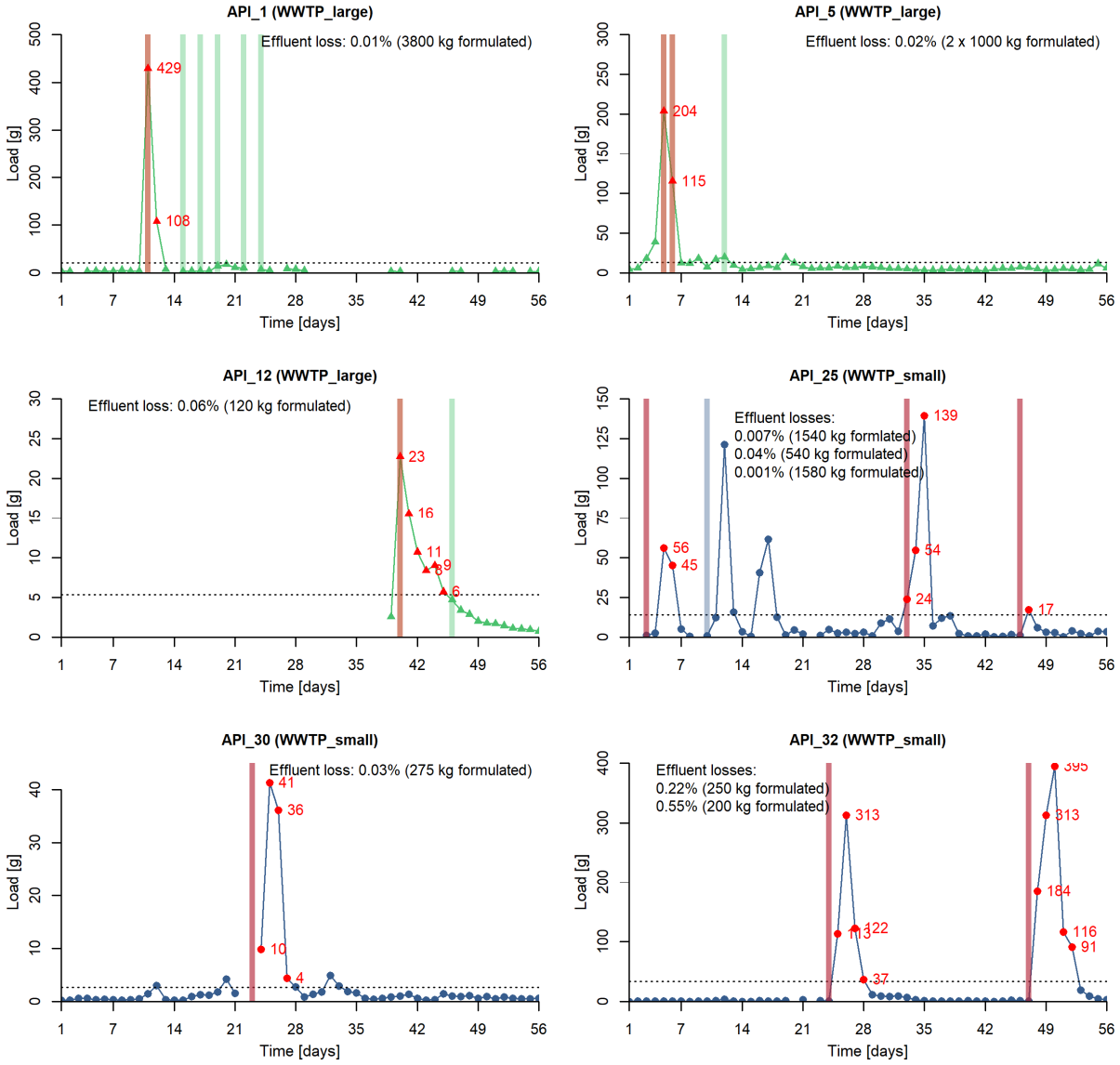


Figure S5: Load time series of the six formulated APIs for which effluent loss factors were estimated. API time series measured in the effluent of WWTP_large and WWTP_small are plotted in green and blue, respectively. Red vertical bars mark formulation events for which loss factors were estimated, whereas green and blue vertical bars indicate days on which the respective API was formulated, but no loss factors were estimated. Loss factors were only estimated for peak emissions that were clearly relatable to a formulation event. Only load values marked with red symbols were considered in the calculations, *i.e.*, values subsequent to the respective formulation event and above the mean load (indicated by the horizontal black dotted line) of the time series. The processed API quantity per formulation event is given in the legend in chronological order. Missing data points are due to concentration values below the LOQ. At WWTP_small data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

SI 15: Times Series of Unexpected Discharges

Suspect Screening at WWTP_large

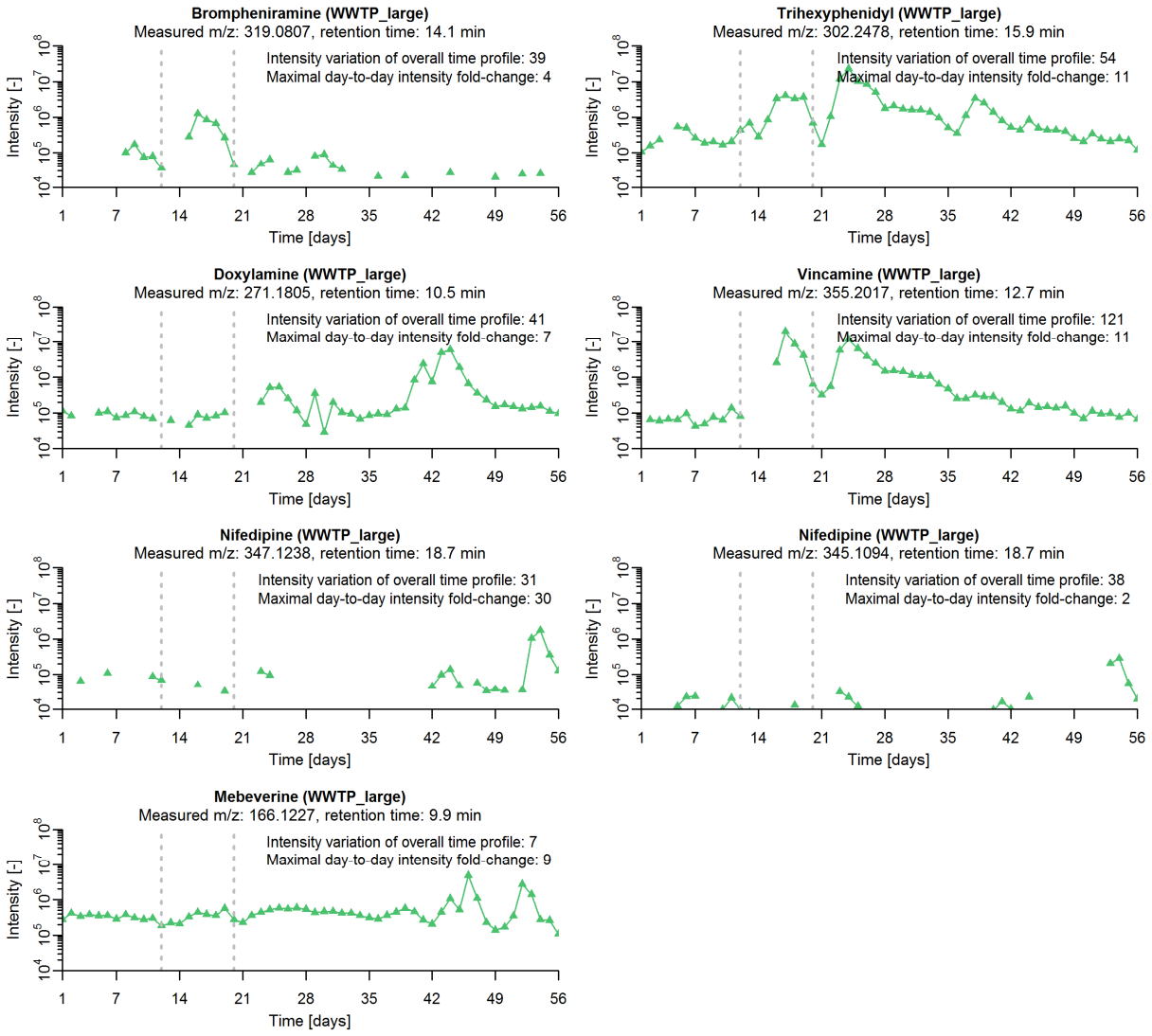
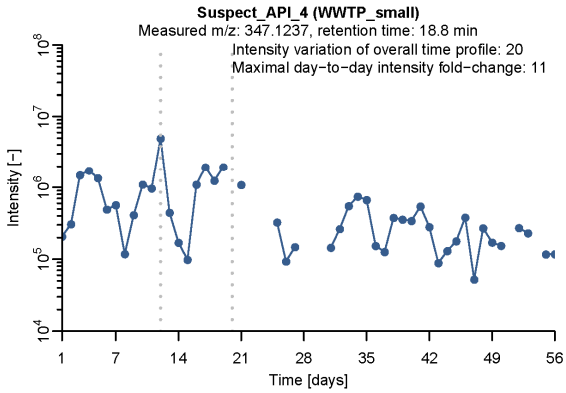
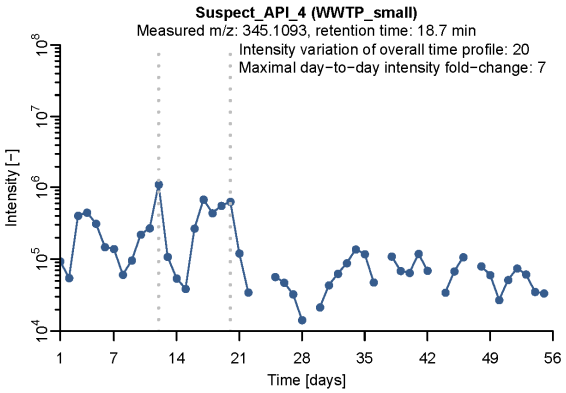
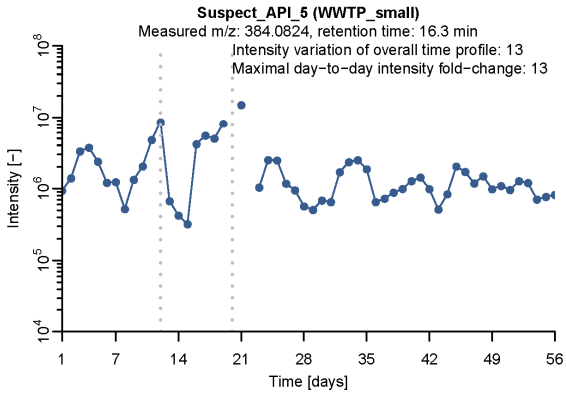


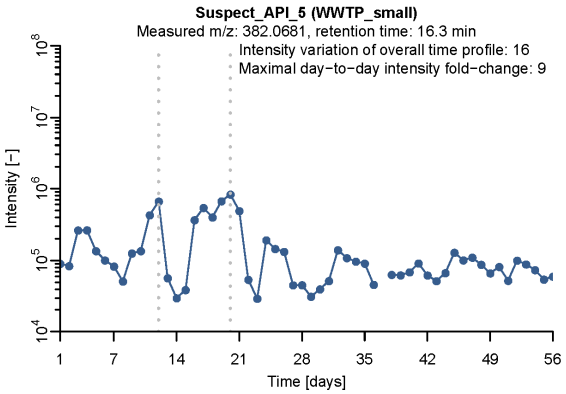
Figure S6: Intensity time profiles of APIs, for which industrial discharges were detected in the suspect screening in the effluents of WWTP-large. These APIs were not reported to be processed during the sampling period by the formulating pharmaceutical industries (FPIs) involved the study, but they were found on the product list on the website of an additional FPI in the catchment of WWTP_large. Please note the logarithmic scale of the intensity (y-axis). The vertical dotted grey lines in the plot indicate days of maximal WWTP discharge caused by two rain events. The intensity variation of the overall profile and the maximal day-to-day intensity fold change, indicated in the legend, were used to allocate effluent emission time patterns to FPI wastewater (see SI 12). Nifedipine was detected in positive and negative electrospray ionization mode; therefore two profiles are shown for this compound. Doxylamine was confirmed with an authentic reference standard (*i.e.*, level 1 according Schymanski, et al. [2]), whereas for the other suspect APIs matching characteristic fragments were found in MassBank or mzCould library spectra (*i.e.*, identification level 2a). Missing data points are due to non-detects.

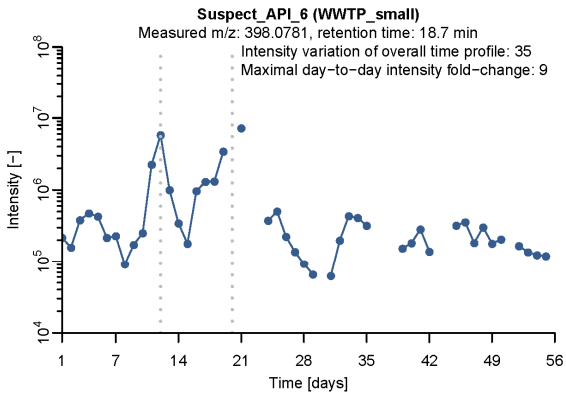
Suspect Screening at WWTP_small

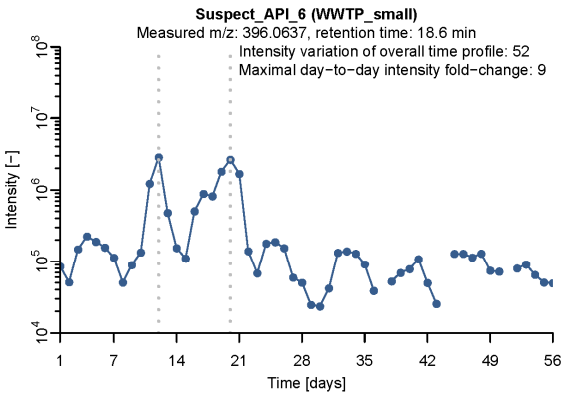


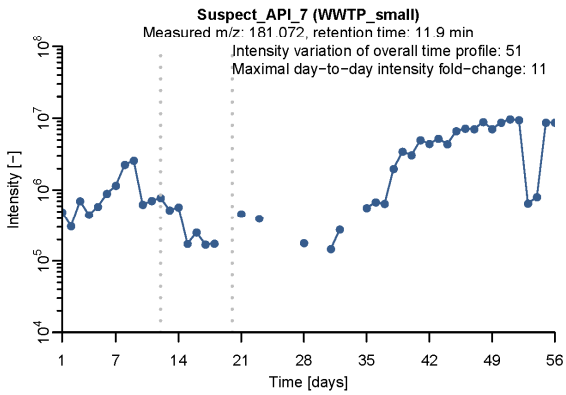


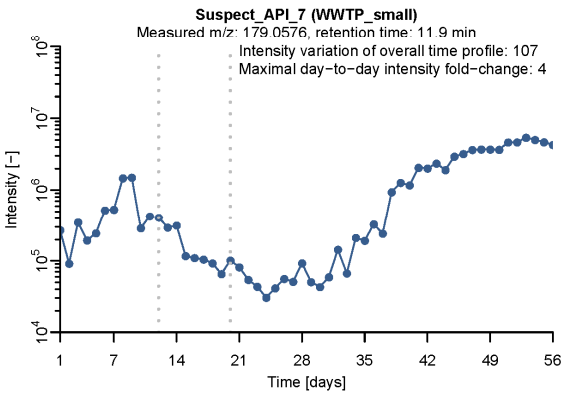












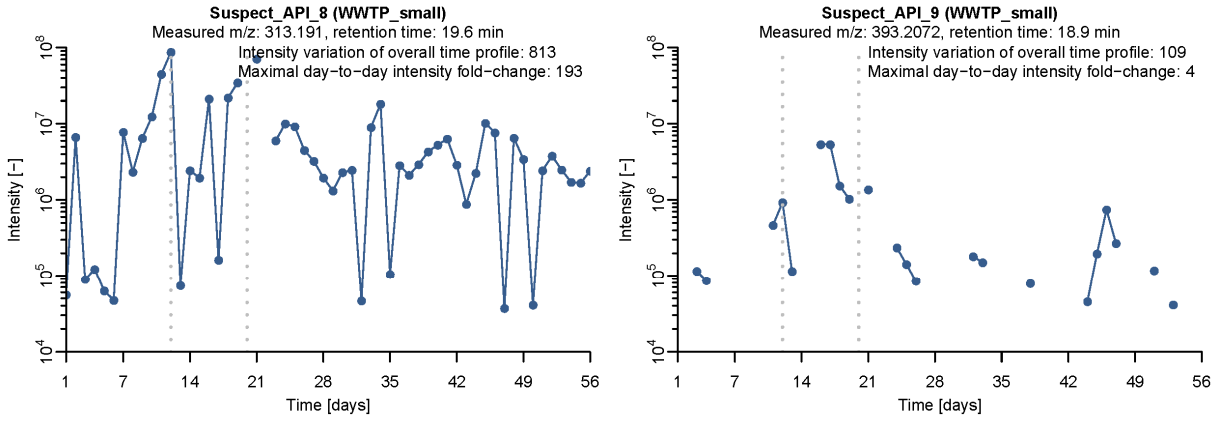


Figure S7: Intensity time profiles of APIs, for which industrial discharges were detected in the suspect screening. These APIs were not reported to be formulated during the sampling period of this study, but from the pre-study it was known that these compounds have been processed in the past by the formulating pharmaceutical industry (FPI_4) in the catchment of WWTP_small. Please note the logarithmic scale of the intensity (y-axis). The vertical dotted grey lines in the plot indicate days of maximal WWTP discharge caused by two rain events. The intensity variation of the overall profile and the maximal day-to-day intensity fold change, indicated in the legend, were used to allocate effluent emission time patterns to FPI wastewater (see SI 12). Suspect_API_4, -5, -6 and -7 were detected in positive and negative electrospray ionization mode; therefore, two profiles are show for these compounds. API_9 was confirmed with an authentic reference standard (*i.e.*, level 1 according Schymanski, et al. [2]), whereas for the other suspect APIs matching characteristic fragments were found in MassBank or mzCould library spectra (i.e., identification level 2a). Missing data points are due to non-detects. Data is generally missing for days 20 and 22 for compounds acquired in ESI positive mode because of corrupt measurement files.

SI 16: Fraction of APIs Emitted from FPIs on Total Contaminant Effluent Concentration

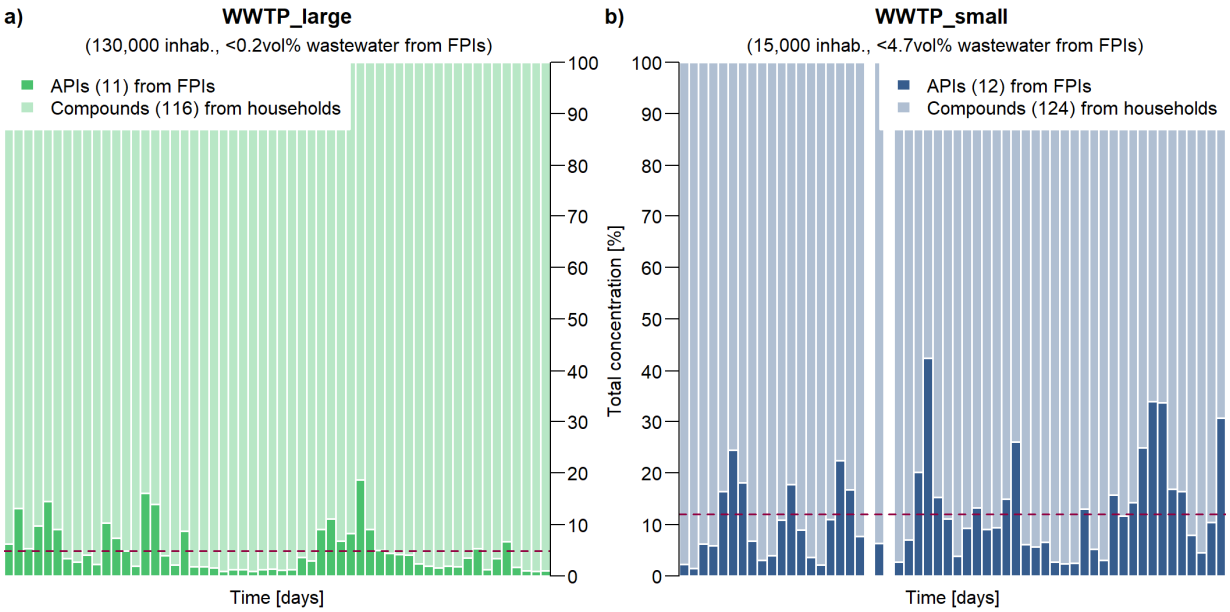


Figure S8: Column chart showing the concentration fraction of active pharmaceutical ingredients (APIs) emitted by formulating pharmaceutical industries (FPIs) compared to all quantified compounds, *i.e.*, not only APIs, from households detected in the effluents of **(a)** WWTP_large and **(b)** at WWTP_small. The number of detected compounds is given in brackets in the legends. The red dashed lines indicate the average of the total contaminant concentration of APIs emitted from FPIs during the sampling period of 8 weeks, *i.e.*, 5% at WWTP_large and 12% at WWTP_small. For WWTP_small total concentrations are not shown for two days (white bars) because only ESI negative mode data was available.

SI 17: Fraction of Most Concentrated Household Compounds on Total Concentration

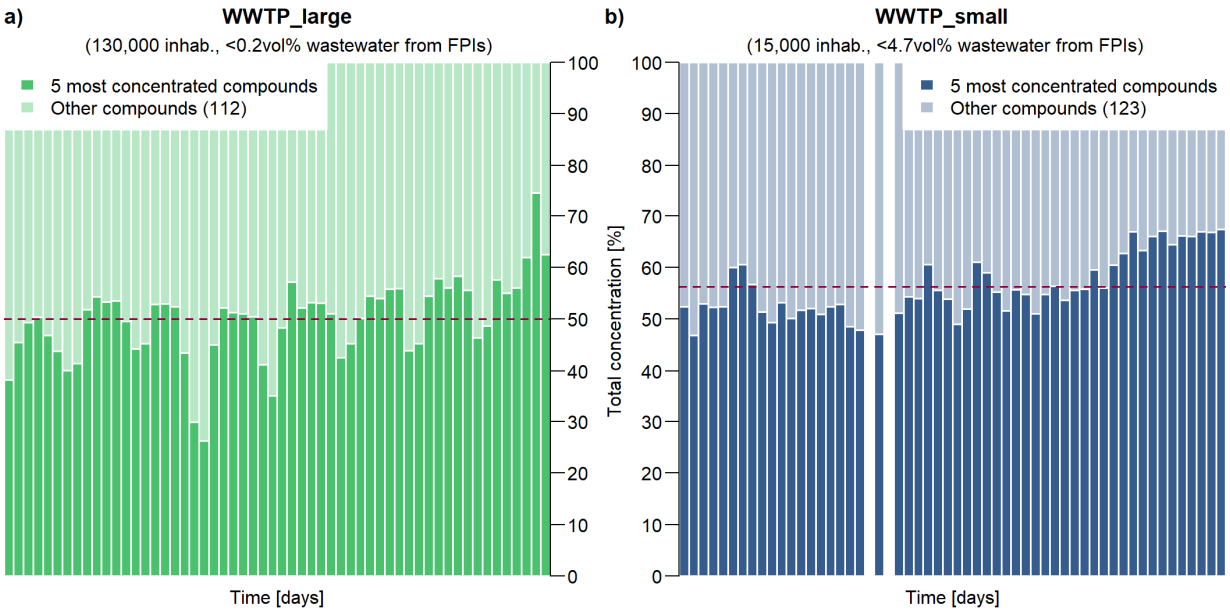


Figure S9: Column chart showing the fraction of the five most concentrated compounds emitted from households on the total compound concentration detected in the effluents of **(a)** WWTP_large and **(b)** WWTP_small. The red dashed lines indicate the average of the most concentrated compounds on the total concentration during the sampling period of 8 weeks, i.e., 50% at WWTP_large and 60% at WWTP_small. The number of detected compounds is given in brackets in the legends. For WWTP_small total concentrations are not shown for two days (white bars), because only ESI negative mode data was available. At WWTP_large the five most concentrated compounds of domestic origin were: Sucralose, iomeprol, API_2, benzotriazole and diclofenac. The weekly pattern observed in the data of the most concentrated compounds at WWTP_large was caused by the concentration profiles of benzotriazole and iomeprol. At WWTP_small the five most concentrated compounds of domestic origin were: metformin, iomeprol, caffeine, acesulfam and API_2.

References

(1) Loos, M., enviMass version 4.2, **2020**. https://www.envibee.ch/eng/enviMass/overview.htm.

(2) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environmental Science & Technology* **2014,** *48*, (4), 2097-2098.

(3) Anliker, S.; Loos, M.; Comte, R.; Ruff, M.; Fenner, K.; Singer, H., Assessing Emissions from Pharmaceutical Manufacturing Based on Temporal High-Resoluion Mass Spectrometry Data. *Environmental Science & Technology* **2020,** *54*, (7), 4110-4120.

(4) Helbling, D. E.; Hollender, J.; Kohler, H. P. E.; Singer, H.; Fenner, K., High-Throughput Identification of Microbial Transformation Products of Organic Micropollutants. *Environmental Science & Technology* **2010,** *44*, (17), 6621-6627.

(5) Kern, S.; Baumgartner, R.; Helbling, D. E.; Hollender, J.; Singer, H.; Loos, M. J.; Schwarzenbach, R. P.; Fenner, K., A tiered procedure for assessing the formation of biotransformation products of pharmaceuticals and biocides during activated sludge treatment. *Journal of Environmental Monitoring* **2010,** *12*, (11), 2100-2111.