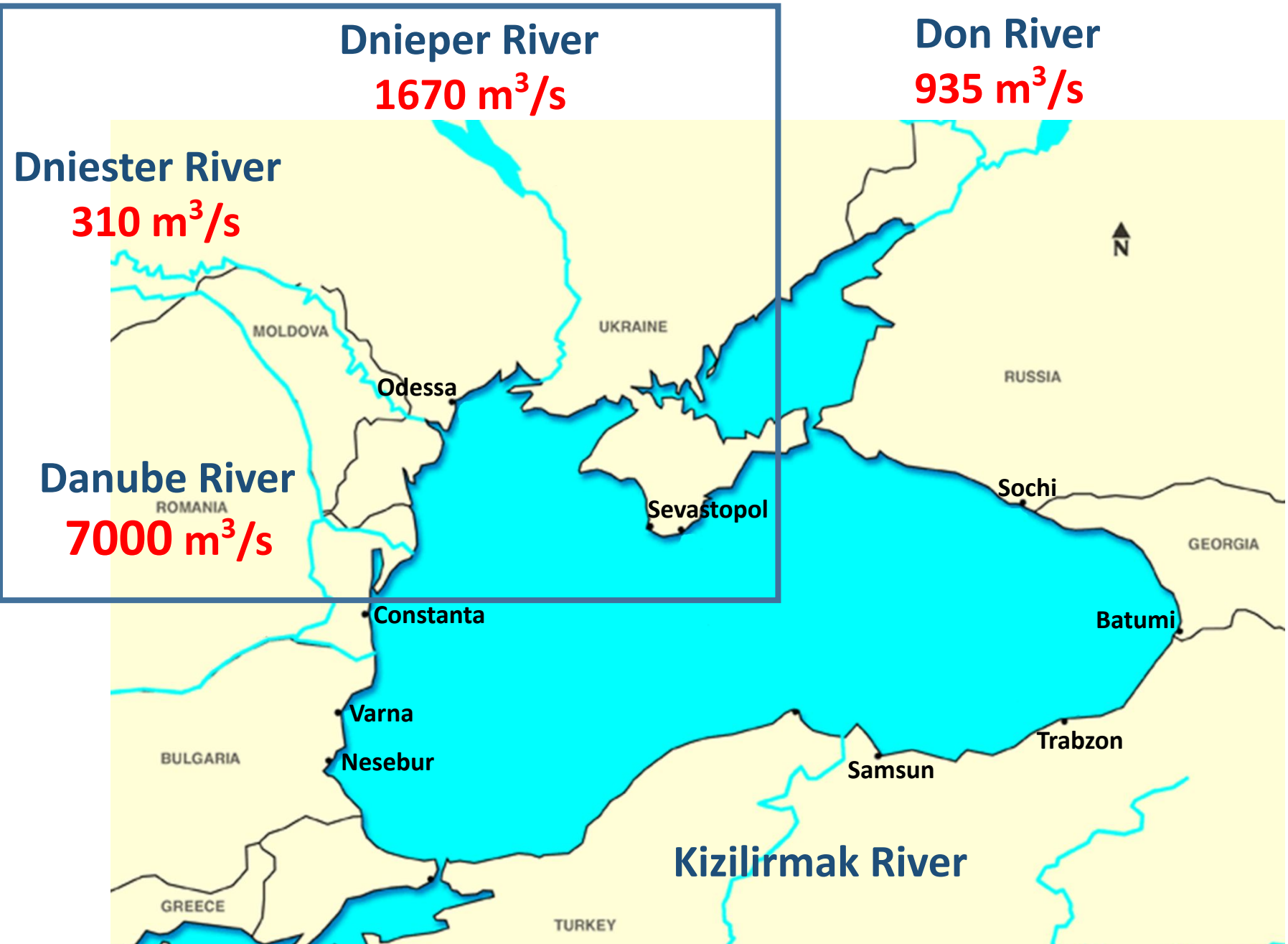
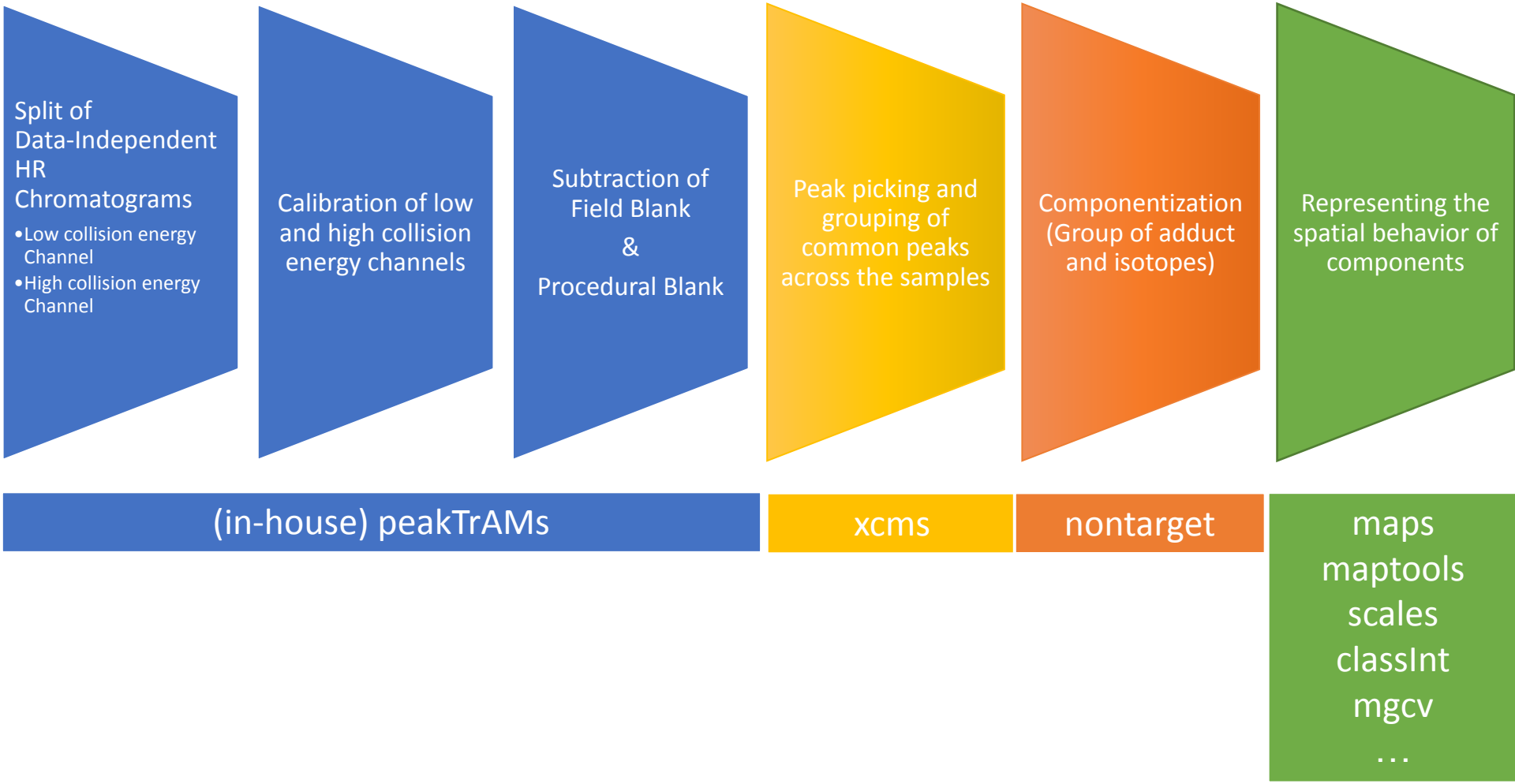
Anthropogenic Pressure on Black Sea



Methodology for tracing chemicals derive from Danube

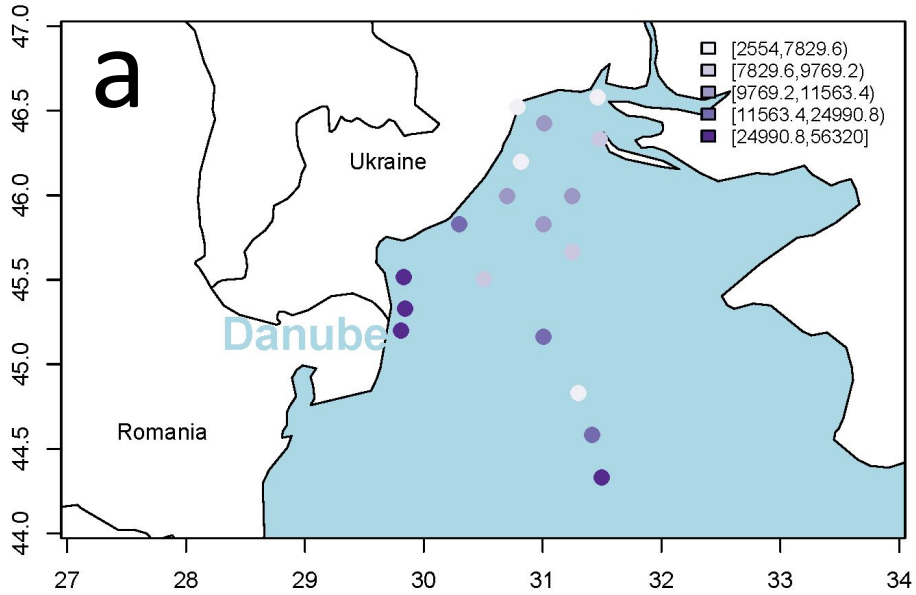


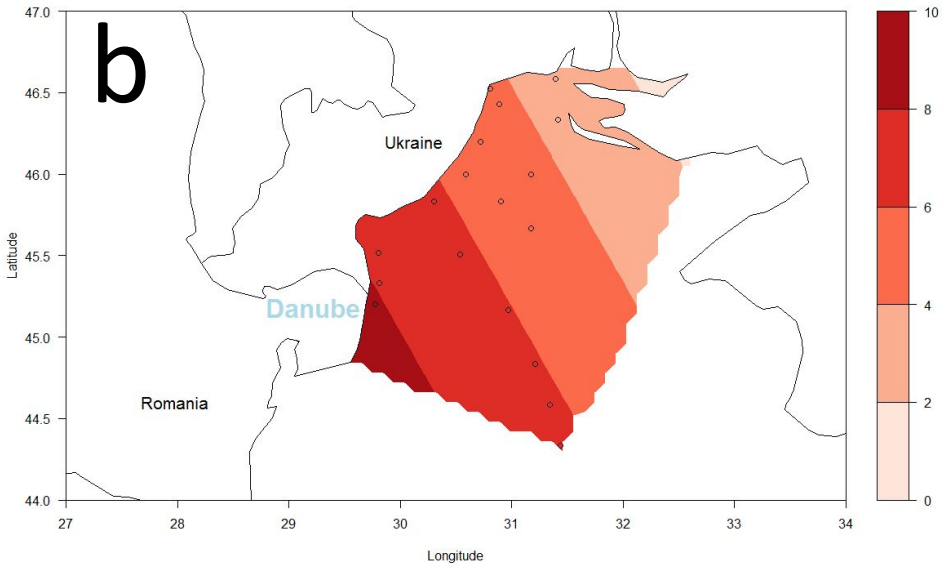
20440 (15,512 (+)-ESI-QTOF+4928 (-)-ESI-QTOF)

Sample	Latitudes	Longitudes	Positive ESI mz=201.955 rt=1.44	Positive ESI mz=240.1375 rt=8.64	Positive ESI mz=147.0308 rt=1.18	Positive ESI mz=416.1823 rt=7.81	Positive ESI mz=262.867 rt=1.28
GE01_B17831	41.55817	41.552167	6510	40895	18331	16757	5707
GE02_B17832	41.5675	41.536667	7372	43833	15453	15974	5447
GE03_B17833	41.665	41.591833	4364	19600	3425	6211	10986
GE04_B17834	41.67183	41.583667	9779	37235	10695	13211	14954
GE05_B17835	41.71	41.5395	11633	42172	18049	12947	6838
GE07_B17836	41.76133	41.712167	8393	34146	14567	8612	15175
GE08_B17837	41.9035	41.751833	6750	33089	13481	15896	4680
GE09_B17839	41.89917	41.674333	9627	21228	11026	7236	43130
GE10_B17840	41.88333	41.616667	9515	40292	3525	13542	8980
GE11_B17841	42.12367	41.619167	9335	28364	14220	5339	5939
GE12_B17842	42.1205	41.593333	9638	40789	3949	25223	4607
GE13_B17843	42.373	41.564	10501	29646	20197	8100	13875
GE14_B17844	42.373	41.558333	9256	42609	13450	13091	13120
GE15_B17845	42.36983	41.4975	6095	27534	14102	7961	14539
JOSS01_B17798	43.75378	31.710433	9611	40931	13863	12855	2813
JOSS02_B17799	43.5537	31.748983	11633	50338	4543	22050	3697
JOSS03_B17800	43.36748	31.833233	7181	47257	4791	11265	4715
JOSS04_B17801	43.30942	32.39169	14593	43001	3801	24368	3420
JOSS05_B17802	42.17707	40.113583	10522	41738	26921	14301	20886
JOSS06_B17803	42.09943	40.34415	10674	40596	28222	13290	5587
JOSS07_B17804	42.0172	40.57865	9235	35289	4223	11976	26551
JOSS08_B17805	41.93277	40.825267	21666	38829	21944	10322	29035
JOSS09_B17806	41.83253	41.02855	15046	46354	29663	20788	16525
JOSS10_B17807	41.78412	41.213567	9938	38646	24349	10133	26577
JOSS11_B17811	41.7041	41.40535	4675	40457	2789	14782	3968
JOSS12_B17812	42.2345	39.886017	8783	41379	28287	9557	7457

Examples of signals that original from Danube

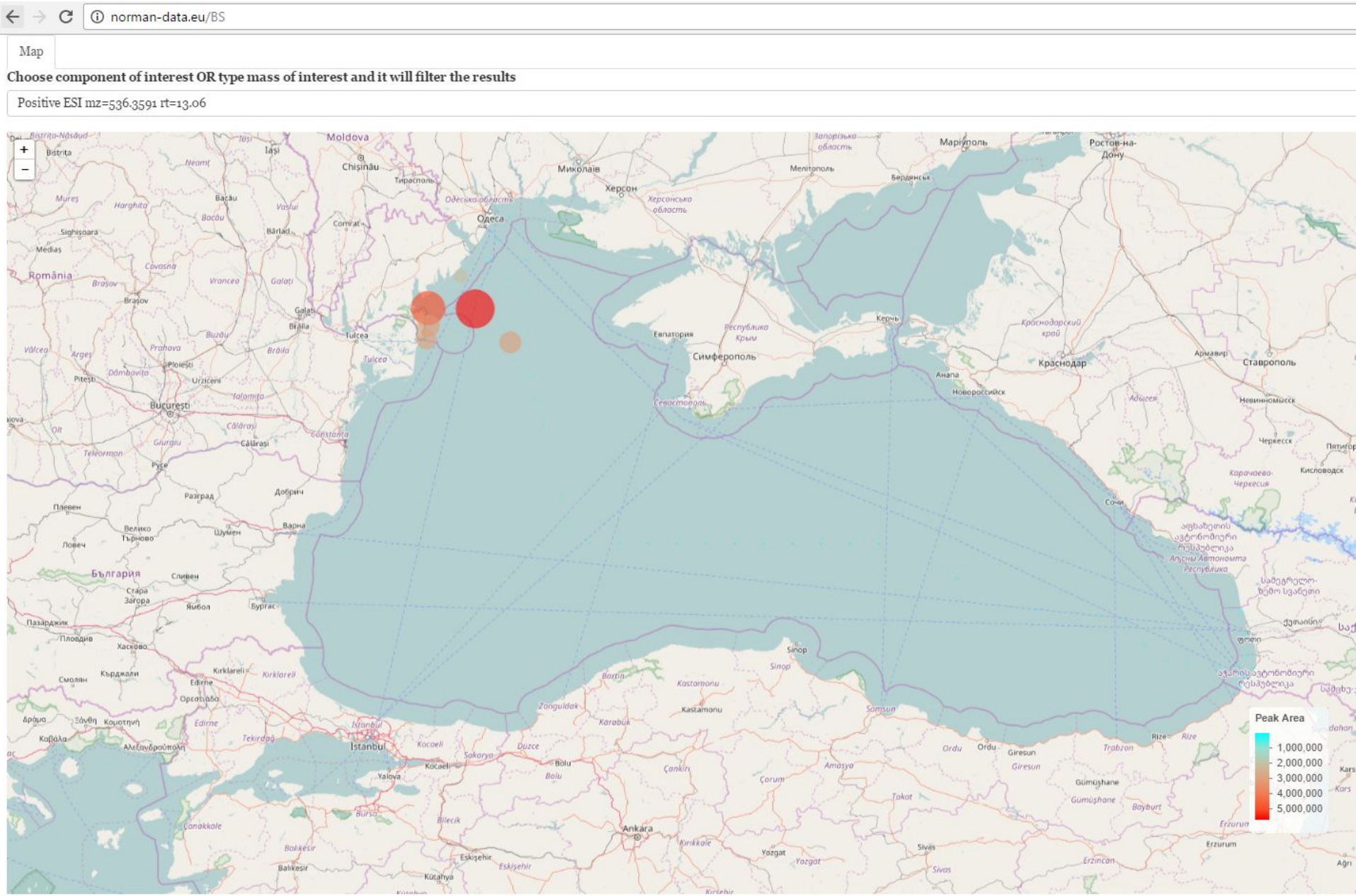
Valsartan





Name	[M+H]+	Molecular Formula	error [mDa]	error [ppm]	ΔRT	mSigma	Q1	Q2	Q3	Identification Level
Metolachlor	284.1412	C15H22ClNO2	-0.30	-1.06	0.17	25.4	252.1151	176.1427	212.0826	1
Metolachlor-ESA	330.137	C15H23N1O5S1	0.10	0.30	0.08	103*	298.1106	-	-	1
4-Acetamidoantipyrin	246.1237	C13H15N3O2	1.00	4.06	0.07	26.5	228.1127	204.1124	159.0905	1
4-Formylaminoantipyrine	232.1081	C12H13N3O2	0.50	2.15	0.06	17.2	214.0975	83.0604	187.0865	1
Benzotriazole (BTR)	120.0556	C6H5N3	-0.80	-6.66	0.01	35	92.0494	65.0384	-	1
Cotinine	177.1022	C10H12N2O1	-0.50	-2.82	0.13	38.5	80.0495	98.0600	70.06513	1
cotinine-Hydroxy	193.0972	C10H12N2O2	-0.40	-2.07	0.16	8.8	80.0495	134.0600	-	1
Metformin	130.1087	C4H11N5	-0.50	-3.84	0.13	6.1	71.0604	60.0556	85.0517	1
Carbamazepine	237.1022	C15H12N2O1	0.70	2.95	0.16	11.8	194.0964	192.0806	-	1
Dimethenamide	276.082	C12H18ClNO2S	0.00	0.00	0.13	19.3	244.0557	-	-	1
Atenolol acid (Metoprolol acid)	268.1543	C14H21N1O4	0.40	1.49	0.17	93.7*	191.0702	116.1070	-	1
Metoprolol	268.1907	C15H25N1O3	-2.40	-8.95	0.05	17.4	74.0607	116.1070	-	1

http://norman-data.eu/BS



Database for visualization of BS data

Deep learning workflow

- 652 figures were manually annotated as chemicals coming from the Danube, 559 figures were annotated as chemicals coming from the Dnieper river and 195 were labeled as substances with unknown source origin (http://dsfp.chem.uoa.gr/BlackSea/train/).
- We tried to develop our own neural network but we did not have enough data and we could not create an efficient and accurate model.
- We used models such as ResNet50, MobileNetV2, InceptionV3 with pretrained weights.
- Four new layers were added, one global spatial average pooling layer, one fully connected intermediate layer, using the relu activation function, a dropout layer and the final logistic layer which outputs the 3 classes: ‘Danube’, ‘Dnieper’ and ‘UnknownOrigin’. The final layer utilized the softmax function.
- Holdout evaluation was used. The data was split into 80% (train set) and 20% (test set)
- Finally a set of images was kept for evaluation purposes.
- The networks performed really well in terms of accuracy:
  - ResNet50 - 86% evaluation and 97% validation accuracy
  - InceptionV3 - 99% evaluation and 96% validation accuracy
  - MobileNetV2 - 93% evaluation and 98% validation accuracy

Summary

- The objective of this study was to investigate the sources of chemicals in the northern west shelf of the Black Sea (compounds coming from the Dnieper river, from Danube and compounds with unknown source origin).
- For this purpose, 1406 figures were manually labelled and we compared the performance of three pre-trained models ResNet50, MobileNetV2, InceptionV3 by adding four new layers.
- In order to evaluate the models, we used both holdout evaluation and keeping a separate evaluation set
- InceptionV3 proved to have the best performance (99% accuracy in evaluation set and 96% accuracy in validation set)
- **We found out that 2091 unknown chemical compound originate from Danube, 3045 originate from Dnieper and 4666 come from an unknown origin**