

I. Martínez*¹, J. Grifoll*¹, Francesc Giralt*¹, R. Rallo*², G. Espinosa*¹ and Y. Cohen³.

* Grup de Fenòmens de Transport, ¹Dept. d'Enginyeria Química, ²Dept. d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Av. dels Països Catalans 26, 43007 Tarragona, Catalunya, Spain.

³ Dept. of Chemical and Biomolecular Engineering, University of California, 5531 Boelter Hall, 405 Hilgard Avenue, Los Angeles, CA 90095, United States. Contact email: izacarjesus.martinez@urv.cat

BACKGROUND

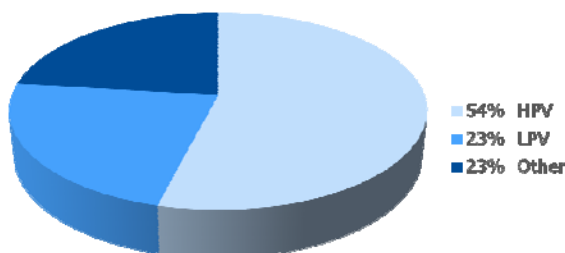
- ANN have been used in QSPR/QSAR models for predicting physicochemical properties and degradation rates related to multimedia chemical partition
- Screening of chemicals is usually based on their degradability in the environment.

GOALS

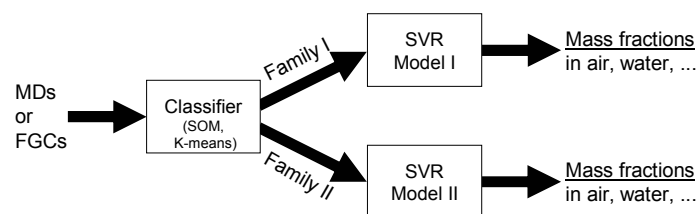
- Develop screening multimedia chemical partition (MCP) models from molecular information by using neural and kernel algorithms.
- Divide data sets into appropriate families to improve the prediction of chemical fate

DATA AND METHODOLOGY

- Experimental data of 383 chemicals (74 included in 2007 CERCLA priority list of hazardous substances)
- $31 \leq MW \leq 959$
- 58% Aromatics; 55% vP ($\tau_{1/2} \geq 60$ days in water)



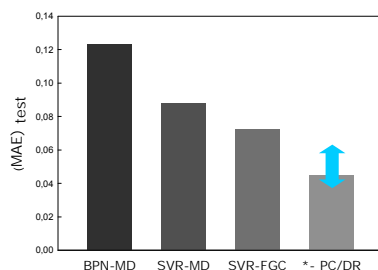
- MCP source code: SimpleBox 3
- Scenario: The Netherlands
- Reference emission: 1 ton/yr in air (other emissions have also been studied)
- Molecular Descriptors (MDs): MW, MR, IP, HOMO, LUMO, μ , PO, 2κ , SA, $2\chi^V$, ΔH_f and T_m
- Functional Group Counts (FGCs): 12 types of atoms, 3 bond types (single, double, triple), 18 different functional groups (Aldehyde, amide, etc.), 8 ring types
- Algorithms: BP NNs, support vector regression (SVR), self-organizing maps (SOM) and K-means



RESULTS AND DISCUSSION

Error trend: diagnostics and proposals

- SVR performs better than BP
- FGCs provide better chemical fingerprint than MDs
- Predictions with partition coefficients and degradation rates still unmatched



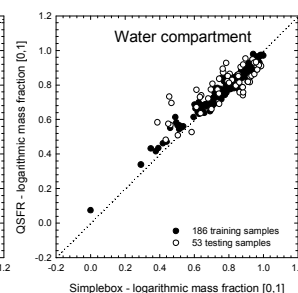
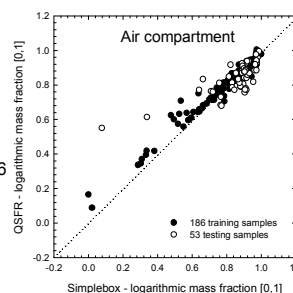
Criteria for chemical classification

- Classification according to degradation rates in air, water, sediments and soil
- Classes of low and high degradability have been obtained with SOM and K-means algorithms.

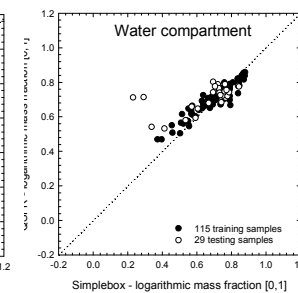
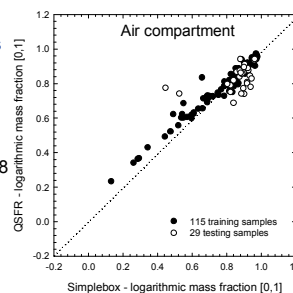
$$C_i = \frac{EMIS_i + IMP_i + \sum_{j,j \neq i} XCH_{ji} C_j}{V_i k_i + FLOW^S + DEP_i + \sum_{j,j \neq i} XCH_{ji}}$$

SVR models with MDs

Family I
SOM, K-means
(239; 186/53)
MAE_{all} = 0.077
q² = 0.7
MAE_{air} = 0.069
MAE_{water} = 0.076

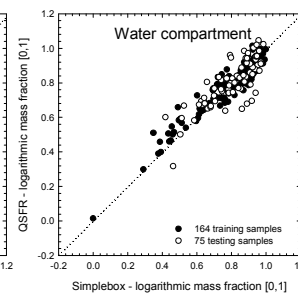
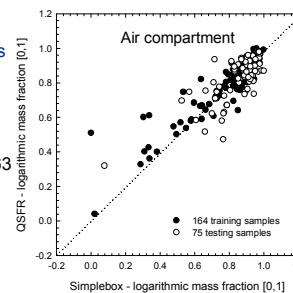


Family II
SOM, K-means
(144; 115/29)
MAE_{all} = 0.090
q² = 0.7
MAE_{air} = 0.076
MAE_{water} = 0.078

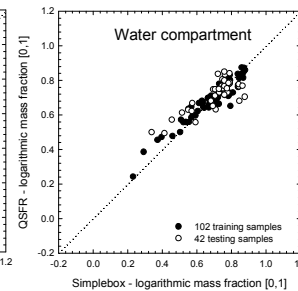
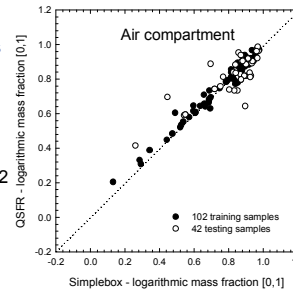


SVR models with FGCs

Family I
SOM, K-means
(239; 164/75)
MAE_{all} = 0.066
q² = 0.7
MAE_{air} = 0.070
MAE_{water} = 0.063



Family II
SOM, K-means
(144; 102/42)
MAE_{all} = 0.066
q² = 0.6
MAE_{air} = 0.062
MAE_{water} = 0.052



CONCLUSIONS

- Neural and kernel algorithms and FGCs can be used to screen the fate of chemicals
- Multimedia chemical fate is reasonably well estimated when chemicals are pre-classified according to their degradation

ACKNOWLEDGMENTS

This research was financially supported by the European Union (NoMiracle project, European Commission, FP6 Contract No. 003956) and by the Generalitat de Catalunya (2005SGR-00735). FG acknowledges the Distinguished Researcher grant received from the Catalan Government.