

OCHEM.eu – Public QSAR framework for modeling PK/PD parameters

Ahmed Abdelaziz, Robert Körner, Sergii Novotarskyi, Anil Kumar Pandey, Iurii Sushko, Matthias Rupp, Igor V. Tetko
Chemoinformatics group, Institute of Bioinformatics & Systems Biology, 85764 Neuherberg, Germany

Chemical property database – structured compound and property information as the basis for modeling

Searching: compounds, properties, conditions

Example output of a search for endothelial permeability

Online chemical database and modeling environment that is easy-to-use, user-driven, and public.

- All data supported by published evidence; easy verification.
- Among the wealth of data available related to PK/modeling is plasma protein binding (2691 data points), bio-distribution in different organs, oral absorption percentage, Pgp induction.
- Data include experimental conditions.
- Batch upload of many compounds in one operation.
- Wiki-style documentation, including an introduction.

Experimental data: conditions

Measurements with experimental conditions

Modeling environment - ADME/T prediction for drug optimization & environmental risk estimation (REACH)

Regression Model

Quantitative property prediction

Descriptor selection

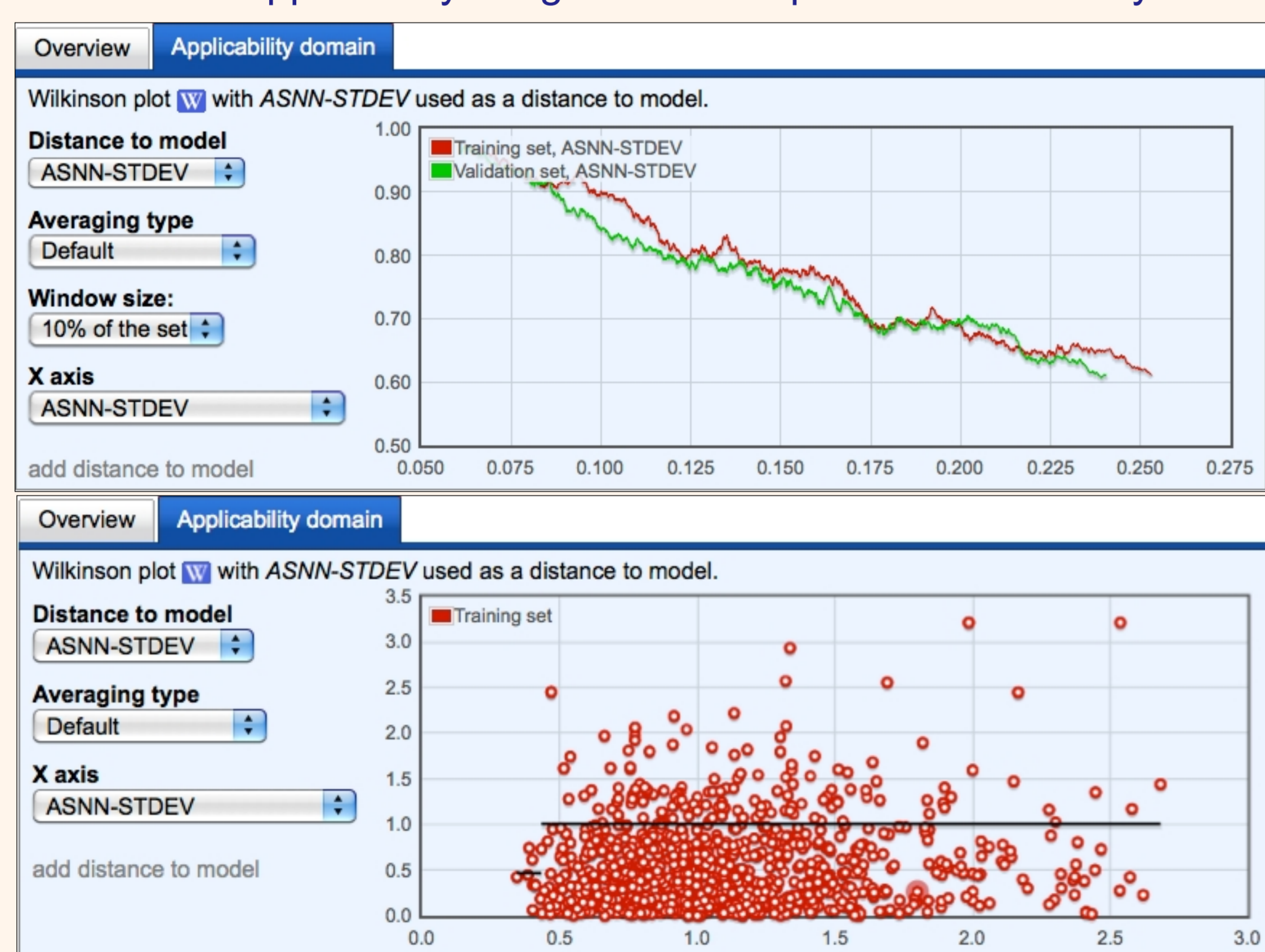
Classification Model

Qualitative property prediction

Model quality – Applicability Domain to estimate model quality

Accuracy

Domain of applicability usage increases prediction accuracy



- Development of new models
- Published model repository
- Refinement of existing models
- Many descriptors (E-state indices, Dragon descriptors, graph kernels, quantum mechanics descriptors)
- Automatic domain of applicability estimation
- Multiple prediction methods (regression, nearest neighbors, neural networks, kernel methods, support vector machines and others)

Summary:

- The database currently contains
 - 250K experimental measurements
 - 500 models built with this data
 - 375 property
 - 1 million compound
 - 4800 references to scientific publications
 - 4000 registered users via VCCLAB
- In-silico models created using OCHEM can provide reliable and accurate predictions of physico-chemical and ADME/T properties relevant to drug development and environmental risk estimation. Such predictions can reduce animal tests and costs of experimental measurements.

Future work:

- Models for Ames test, CYP 450 1A2, BBB
- Prediction of pKa, logD, solubility

We thank Prof. R. Todeschini & Prof. J. Gasteiger for their DRAGON & Corina programs, Dr. P. Ertl for JME Editor. We also thank CDK and InChi communities for the open source software, Prof. A. Varnek for fragmental descriptors software as well as ChemAxon for visualisation software.

