ReadMe: Additional Physicochemical Models for StarDrop

Version 1.0

These predictive models of physicochemical properties are provided free of charge to all StarDrop users. They have been built using the StarDrop's Auto-Modeller module based on the data sets provided with the US Environmental Protection Agency's Toxicity Estimation Software Tool (TEST) (http://www.epa.gov/nrmrl/std/cppb/qsar/#TEST). Full details of the model generation and validation can be found in the accompanying document "Additional Physicochemical Models for StarDrop".

Extracting the Archive

Extract the contents of the .zip file into a convenient directory by right clicking on the file and selecting "Extract All".

Directory Contents

The extracted directory contains seven StarDrop model files (.aim) and a PDF associated with each. The PDF for each model contains details such as validation results, descriptors used and (where applicable) variable importance information.

An additional PDF, "Additional Physicochemical Models for StarDrop" provides a full description of the data and methods used to build the models along with their validation results.

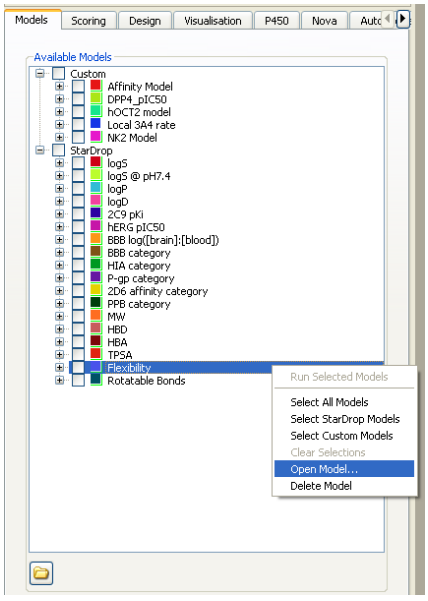
A subdirectory, "Alternative Models," contains two additional models and associated PDFs. For details of these, please see below.

Installation

The models can be opened in StarDrop in two ways:

If you wish to load an individual model:

- Right click on the "Models" Tab in StarDrop
- Select "Open Model…" from the menu that appears
- Navigate to the directory into which you extracted the model files
- Choose the model you wish to load and click "Open"
- The model will appear in the list of "Custom" models in the "Models" tab.
- You will need to reopen the model in the same way, each time you restart StarDrop.



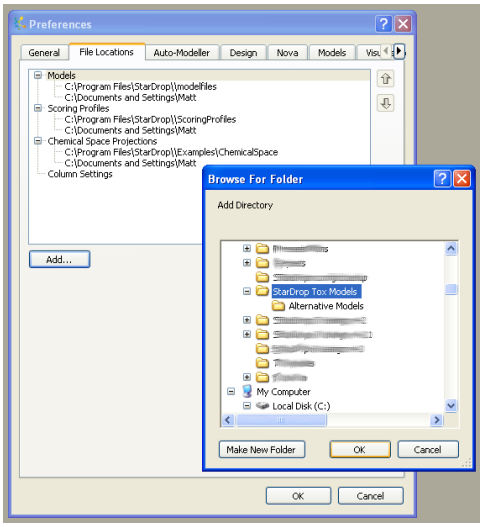
7226 IQ Cambridge
Beach Drive, Cambridge
CB25 9TL, UK

Tel: +44 1223 815900
Fax: +44 1223 815907

Email: info@optibrium.com
Website: www.optibrium.com

Alternatively, if you would like the models to be available in StarDrop every time you start the application, you can configure StarDrop to load all of the models automatically from a directory, as follows:

- Select the File->Preferences menu option in StarDrop
- Choose the "File Locations" Tab
- Select the "Models" line and click "Add"
- Navigate to the directory in which you extracted the model files
- Select the directory and click "OK"
- Click "OK" in the "Preferences" dialogue box
- All of the models will appear in the "Models" tab under "Custom"
- The models will be present every time you start StarDrop
- If you want to remove the models at a later date, return to the "Preferences" dialogue box and delete the directory containing the models under the "Models" line in the "File Locations" tab.



Alternative Models

In the main directory containing the models, you will find a subdirectory "Alternative Models." These contain alternative models for Surface Tension and Viscosity. These are Gaussian Processes (GP) models with marginally lower performance on the independent test set than the best StarDrop models for these properties, which are based on the RBF method. However, GP models have some desirable properties over RBF models; for example, GP models produce individual estimates of the uncertainties in the predicted values for compounds and they tend to perform better when extrapolating beyond the chemical space on which the model was trained.

Therefore, we have included these alternative models for you to try on your compounds if you wish. The validation data for these models are shown in the table below:

Property	Alternative StarDrop Model						
	Method	R^2	RMSE	$R^2_{Pearson}$	$\frac{R^2_{Pearson} - R^2_0}{R^2_{Pearson}}$	k	MAE
Surface Tension	GPFixed	0.82	2.65	0.83	0.01	1.00	1.30
Viscosity	GP2DSearch	0.90	0.16	0.91	0.00	0.87	0.12