Reliability in Predictive Models under Alternative Treatments of Predictive Uncertainty

– QSPRs in Chemical Safety Assessments

Ullrika Sahlin[1,2] PhD
[1] CADASTER, Linneaus University, Sweden
[2] Centre of Environmental and Climate research, Lund University, Sweden

Background and aim

In the absence of experimentally tested physicochemical endpoints, European chemical regulation under REACH allows the use of non-testing strategies such as Quantitative Structure-Property Relationships (QSPR) to predict the required information using analogy models. A QSPR may be more or less reliable for predicting a property of a chemical compound. Quantitative measures of predictive reliability support the decision on whether a model is reliable enough to use for prediction. The integration of non-testing strategies into decision making calls for evaluated approaches to treat predictive uncertainty and reliability in QSPR predictions.

The aim was to illustrate the practical impact of the treatment of epistemic uncertainty in QSPR predictions on decision making, by showing how the judgment of compounds as being more or less inside the applicability domain of a predictive model may transfer to the treatment of predictive uncertainty.

An illustration of a practical impact of the treatment of epistemic uncertainty on decision making:

Fig 1. Predictive uncertainty assessment:

- Posterior Distribution by Bayesian Modeling*
  - mean (blue)
  - robust Bayes for two credibility levels (black)
- Estimates of Predictive Variance by Re-sampling and Gaussian (red)
- Interval Estimation by Sensitivity Analysis (green)

*Bayesian lasso

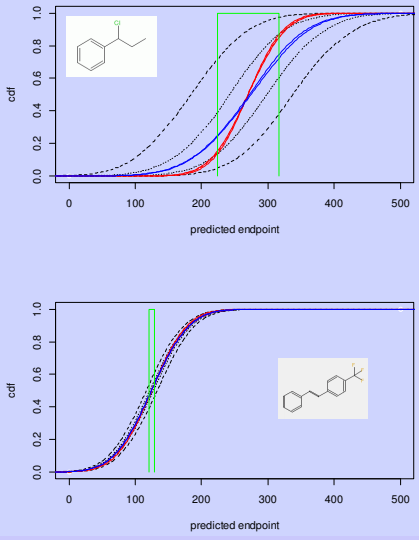


Fig 2. Predictive reliability and the extent of extrapolation (by leverage) in predictions of the external data set (red). Lines indicate predictive distributions for three items in the training set (solid black) and extrapolated items (solid red), to be compared with observed values (blue triangles).

Reflection:

Even though the width of the predictive distribution increases with leverage, reliability ought to be higher for compounds inside the applicability domain of the QSPR.

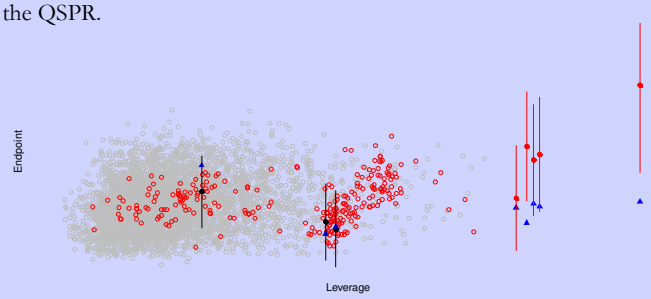
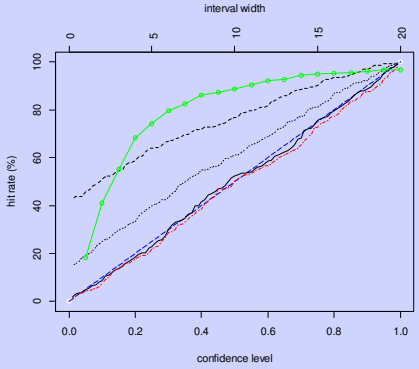


Fig 3. Global Predictive Reliability assessment:

A quantitative performance measure of predictive reliability is given by the empirical hit rate.

Result:

Similar degree of uncertainty aversion give similar reliability judgments



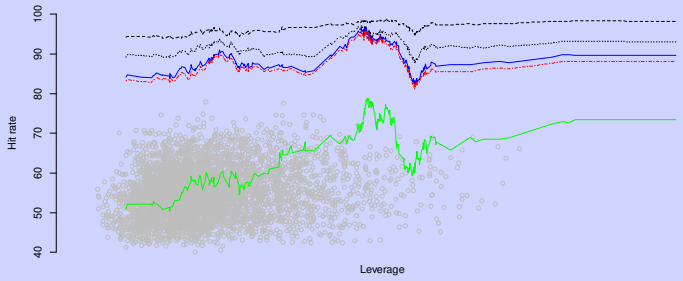


Fig 4. Local Reliability assessment: Local coverage by smoothed "kNN hit rates".

Result: Similar pattern between alternative assessments of predictive uncertainty. Higher aversion to uncertainty give higher, but more similar, reliability judgments.

Conclusions

Alternative approaches to assess predictive uncertainty detects the same pattern in uncertainty and reliability (Fig 1 & 4), but has an impact on the reliability in a model prediction and thereby the judgment on whether a compound is inside the applicability domain (Fig 2), since the treatment of predictive uncertainty

- forces us to balance predictive reliability to precision (Fig 3)
- reflect our preferences/aversion to uncertainty in the predictions (Fig 3 & 4)

Treatment of epistemic uncertainty depend on the kind of background knowledge and context, here the replacement of experimental data by non-testing information through analogy predictions in chemical regulation.

Acknowledgements

This study was funded by the European Seventh Framework Programme (FP7) through the project CADASTER (grant agreement 212668).

Linnæus University