QSPR in the solubility challenge

- Hewitt et al tried four QSPR models [5]
  - Multiple linear regression (426 descriptors, genetic algorithm, no more than 5 used at once)
  - Artificial neural network
  - Category-specific models based on H bond ability
  - Various commercial QSPR models
  - Consensus, mean of four models
- The best was a three-descriptor linear regression
  - based on log P, boiling point, and R maximal autocorrelation of lag 2 (related to size and connectivity).
  - $R^2_{train} = 0.74, R^2_{test} = 0.51$
- "none of the other modeling approaches used in this study was able to improve upon the predictions made by the MLR model"

16

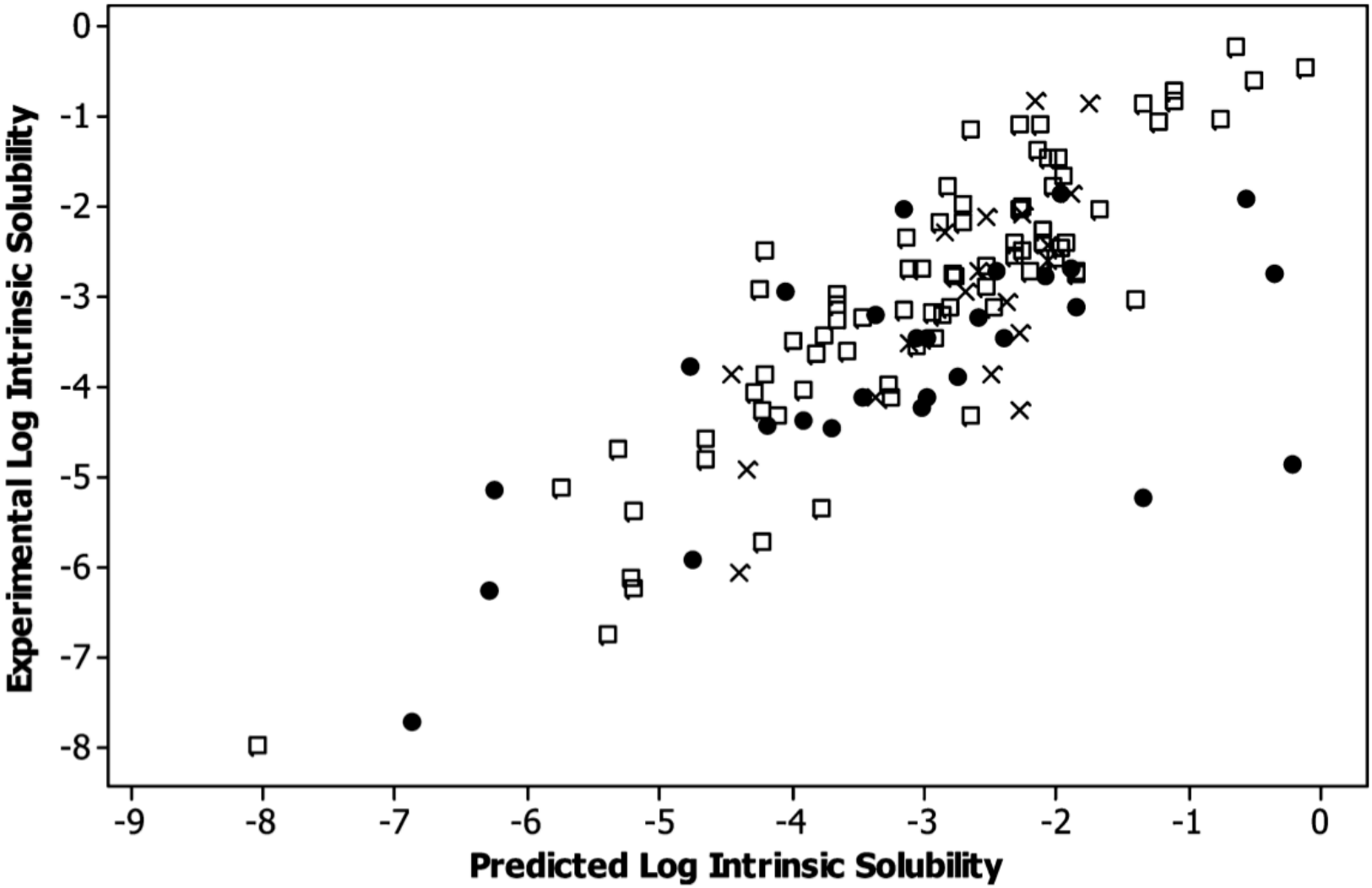

Figure 1. Plot of experimental versus predicted log(intrinsic solubility) using the MLR model (eq 2) for (□) the training set, (×) the validation set, and (●) the test set.

QSPR in the solubility challenge

- Hewitt et al tried four QSPR models [5]
  - Multiple linear regression (426 descriptors, genetic algorithm, no more than 5 used at once)
  - Artificial neural network
  - Category-specific models based on H bond ability
  - Various commercial QSPR models
  - Consensus, mean of four models
- The best was a three-descriptor linear regression
  - based on log P, boiling point, and R maximal autocorrelation of lag 2 (related to size and connectivity).
  - $R^2_{train} = 0.74, R^2_{test} = 0.51$
- "none of the other modeling approaches used in this study was able to improve upon the predictions made by the MLR model"

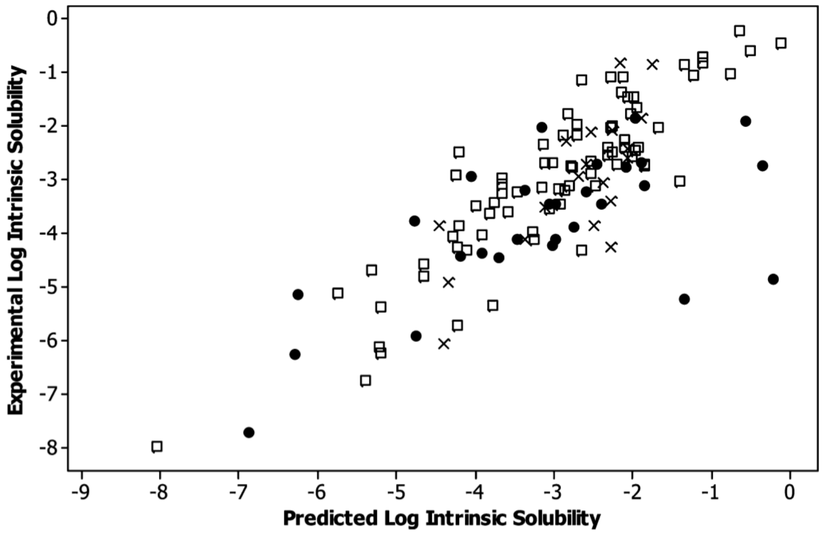

Figure 1. Plot of experimental versus predicted log(intrinsic solubility) using the MLR model (eq 2) for (□) the training set, (×) the validation set, and (●) the test set.

Can deep learning do better? [6]

Table 5

Prediction performance and standard deviations using 10- fold cross validation on the Solubility Challenge Dataset (125 molecules)

Models	R^2	std R^2	RMSE	std RMSE	AAE	std AAE
UG-RNN	0.32	0.03	1.41	0.12	1.08	0.10
UG-RNN-LogP	0.45	0.04	1.27	0.13	1.03	0.11
UG-RNN-CR-LogP	0.44	0.09	1.28	0.18	1.03	0.16
UG-RNN-Huusk	0.43	0.02	1.16	0.03	0.93	0.03
UG-RNN-Huusk-Sub	0.48	0.02	1.11	0.03	0.84	0.01
UG-RNN-LogP-Huusk	0.54	0.02	1.00	0.03	0.82	0.03
UG-RNN-LogP-Huusk-Sub	0.60	0.02	0.94	0.02	0.71	0.02
UG-RNN-CR-LogP-Huusk	0.62	0.03	0.96	0.06	0.83	0.06
UG-RNN-CR-LogP-Huusk-Sub	0.67	0.03	0.90	0.06	0.74	0.05
NN-Sol-Chal[11]	0.40	-	1.51	-	-	-
MLR-Sol-Chal[11]	0.51	-	0.95	-	0.77	-
New *in silico* consesus[11]	0.60	-	0.90	-	0.68	-

Better than simple neural network, but worse than MLR

Huusk results are based on more data.
Sub results use different solubility values.

In summary, deep learning can do better if there is more data.