# A Temperature-Dependent Quantum Mechanical/Neural Net Model for Vapor Pressure

Andrew J. Chalk,[†] Bernd Beck,*[,‡] and Timothy Clark*[,†]

Computer-Chemie-Centrum, Friedrich-Alexander-Universität Erlangen-Nürnberg and Accelrys Inc., Computer-Chemie-Centrum, Nägelsbachstrasse 25, D-91052 Erlangen, Germany

Received February 7, 2001

We present a temperature-dependent model for vapor pressure based on a feed-forward neural net and descriptors calculated using AM1 semiempirical MO-theory. This model is based on a set of 7681 measurements at various temperatures performed on 2349 molecules. We employ a 10-fold cross-validation scheme that allows us to estimate errors for individual predictions. For the training set we find a standard deviation of the error $s = 0.322$ and a correlation coefficient ($R^2$) of 0.976. The corresponding values for the validation set are $s = 0.326$ and $R^2 = 0.976$. We thoroughly investigate the temperature-dependence of our predictions to ensure that our model behaves in a physically reasonable manner. As a further test of temperature-dependence, we also examine the accuracy of our vapor pressure model in predicting the related physical properties, the boiling point, and the enthalpy of vaporization.

## INTRODUCTION

Recently, many studies investigating the relationship between molecular properties, or descriptors, and physical properties have been published. Two reviews provide an overview of recent work in this field.[1,2] Ideally the descriptors used should be derivable from only the chemical structure and no other empirical parameters. It is obviously advantageous to be able to calculate physical properties without the need for possibly expensive and time-consuming experiments quickly and easily.[1,2]

Vapor pressure is an interesting property for several reasons; it is intimately related to other properties, such as boiling points, where intermolecular interactions are important. Knowledge of vapor pressures is also useful for chemicals that will be used in the environment. For example, the magnitude of the vapor pressure is important in determining the environmental impact of pesticides.[3]

There have been several studies published[4−7] over the past few years that have proposed models for vapor pressure using both linear regression and neural network models. All these studies have two things in common, however. They only attempt to predict vapor pressure at 298 K and use relatively small datasets. We therefore investigated creating a model that includes temperature-dependence and makes no assumptions about what functional form this dependence should take. To achieve this, we take data at all available temperatures and include the temperature as an additional descriptor. This allows us to include a much larger dataset than would otherwise be available, around five times as many molecules as used in previous models. We expect this to result in a model that can provide reliable predictions for a much larger range of molecules than those published previously.

This study aims to develop a robust, temperature-dependent vapor pressure model that includes error estimations for individual compounds and is applicable to compounds with a wide range of functionality. We employ a neural network approach with descriptors derived from quantum chemical calculations. These descriptors can be calculated for any molecule of interest (within the limits of the quantum chemical method) and most have simple physical interpretations. We also examine the predicted temperature-dependence in detail in order to ensure that it is physically correct. As a further test of temperature-dependence, we use our model for the prediction of boiling points and heats of vaporization. Comparisons between our work and a recently published model are given.

## I. METHODS

**A. Quantum Chemical and Descriptor Calculations.** Hydrogen suppressed 3D-molecular structures exported from the Beilstein database[8] were processed by CORINA[9] to add hydrogen atoms in the appropriate positions and to ensure a consistent set of conformations. Only a single conformation was produced for each molecule. The resulting structure was then fully optimized using the AM1[10] Hamiltonian within VAMP 7.5.[11] The standard parameters for AM1 for all elements were used throughout. Electrostatic potentials were calculated using the natural atomic orbital/point charge (NAO-PC)[12−14] model. All surface-based descriptors are calculated on the solvent-excluded surface (SES), which is determined by an algorithm based on GEPOL[15] and the Marsili marching cube algorithm[16] with a solvent radius of 1.4 Å. Polarizabilities were calculated using a parametrized atomic approach.[17] All descriptors were calculated with Propgen 1.0.[18]

**B. Descriptor Selection.** Descriptors were chosen by first creating a neural net model, with the methods discussed in Section C, using all 69 available descriptors as well as the absolute temperature. The "importance" of each descriptor

* Corresponding author phone: ++49 (0)9131 8522948; e-mail: clark@chemie.uni-erlangen.de (Clark) and phone ++49 (0)9131 879 622; e-mail: beck@chemie.uni-erlangen.de (Beck).

[†] Universität Erlangen-Nürnberg.

[‡] Oxford Molecular Ltd.

10.1021/ci0103222 CCC: $20.00 
Published on Web 06/26/2001

**Table 1.** Final Descriptor Set

| |
|---|
| 1. absolute temperature |
| Bulk Electrostatic Descriptors |
| 2. molecular electronic polarizability[17] |
| 3. dipolar density ($dipole^2/volume$) |
| Solvent Excluded Surface-based Electrostatic Descriptors[20,21] |
| 4. maximum electrostatic potential |
| 5. mean positive electrostatic potential |
| 6. mean negative electrostatic potential |
| 7. total mean electrostatic potential |
| 8. variance of the positive electrostatic potential |
| 9. total variance x balance |
| 10. local polarity |
| Weak Interactions |
| 11. sum of covalent hydrogen bond acidities[23] |
| 12. sum of covalent hydrogen bond basicities[23] |
| 13. sum of electrostatic hydrogen bond acidities[23] |
| 14. sum of electrostatic hydrogen bond basicities[23] |
| 15. number of hydrogen bonding acceptor groups (*nacc*) |
| 16. number of hydrogen bonding donor groups (*ndon*) |
| 17. cohesive index: $nacc \times \sqrt{ndon}/(surface\ area)$[33] |
| 18. number of aryl groups |
| Geometrical |
| 19. molecular weight |
| 20. globularity[25] |
| 21. number of rotatable bonds, excluding terminal groups |
| Electrotopological States (E-States) |
| 22. quantum mechanically based sum of E-states on N atoms[34] |
| 23. quantum mechanically based sum of E-states on O atoms[34] |
| 24. sum of E-states on oxygen atoms[24] |
| 25. sum of E-states on nitrogen atoms[24] |
| Electrostatic Field Based Sum of Atomic Charges on |
| 26. hydrogen |
| 27. oxygen |

was calculated by a simple weight-based method,[19] and the results were averaged over the 10 neural nets. We have examined the accuracy of predictions for models including various numbers of the most important descriptors and found that the best compromise between network size and accuracy was given by including the 30 most important descriptors. Three could be removed without significantly impacting on accuracy, resulting in the final set of 27 descriptors.

These descriptors have been grouped into several classes, as shown in Table 1. Intermolecular interactions play a pivotal role in determining many properties, including vapor pressure. We therefore now discuss the role each descriptor has to play in gauging the strength of these interactions.

The first class of descriptors are bulk electrostatic properties, such as polarizability and dipolar density. These properties are important in describing dipolar and induced dipolar interactions.

The electrostatic surface-based descriptors proposed by Politzer, Murray, and co-workers[20,21] describe the electrostatic potential on the solvent excluded surface, where interaction between molecules occurs. The utility of these descriptors in describing molecular interactions for the related property, boiling point, has already been demonstrated several times.[20−22] Hydrogen bonding plays a very important role in intermolecular interactions, hence it is no surprise that a number of descriptors specifically related to hydrogen bonds are found to be important. Included are the covalent and electrostatic hydrogen bond basicities and acidities[23] as well as a simple count of hydrogen bond donor and acceptor sites. The number of aryl groups gives information about possible $\pi-\pi$ interactions. Various E-state indices[24] on oxygen and nitrogen are also found to be important. Since hydrogen bonding predominantly occurs through these atoms or hydrogens attached to them, it seems likely that these descriptors provide additional information about the strength of hydrogen bonding.

Geometrical information is supplied by the globularity,[25] a measure of the deviation of the shape of the molecule from an ideal sphere and the number of rotatable bonds, which gives a measure of the flexibility of the molecule. Rotations of terminal groups (e.g. methyl) are not included in this count because they do not result in any new conformations.

**C. Feedforward Neural Nets.** Three-layer feedforward neural networks using sigmoid transfer functions ($1/(1+e^{-x})$) and trained with the back-propagation of errors algorithm[26−28] were employed throughout. For the sake of brevity, we refer to these simply as neural nets. We have examined neural nets with between 10 and 18 hidden nodes and have found that the best performance, as measured by the cross-validated RMS error, was a 27:15:1 network architecture. Such networks contain 420 weights or 0.055 weights per training sample.

Ten separate networks with random starting weights were trained with different training and testing sets, chosen such that each molecule appears in the test set for one and only one network. The cross-validated result for a given molecule is then the prediction of the net in which the molecule does not appear in the training set and should therefore give a worst case prediction for the given model. Training of the nets is halted when the root-mean-square (RMS) error of the cross-validated predictions for the entire dataset reaches a minimum. The prediction of the model is given by the mean of the results for the 10 nets. Using our method for prediction of individual errors discussed elsewhere,[29] we find that an estimate of the prediction error of an individual molecule can be given by 1.31 times the standard deviation of the results of the 10 neural nets for that species. This very low multiplication factor indicates that the data covers the range of chemical diversity in the training/test set evenly.

**D. Training/Test Data.** For this study we have extracted structures and experimental vapor pressures, including the measurement temperature, from the Beilstein[8] database. During initial neural net training runs, many large outliers were identified, and their experimental results were examined in detail. We have removed more than 200 measurements for which it was found that outliers had experimental values that were inconsistent with results at other temperatures or with the experimental boiling point. A number of measurements were also found where the temperature was reported in Celsius but appears to be in Kelvin. These results were also removed. The remaining data consists of 8542 measurements on 2349 compounds of diverse functionality, an average of around 3.6 different temperature measurements per compound. This dataset was divided into a training/test set of 7681 measurements and validation set of 861 measurements in such a manner that the validation set spans the full range of vapor pressures. No information regarding the accuracy of these data is available.

The distribution of vapor pressure (Vp) and temperature ($T$) values for the training dataset are shown in Figure 1. The vapor pressure values range from −8.63 to 5.47 log(Torr) units, while the temperatures range from 76 to 800

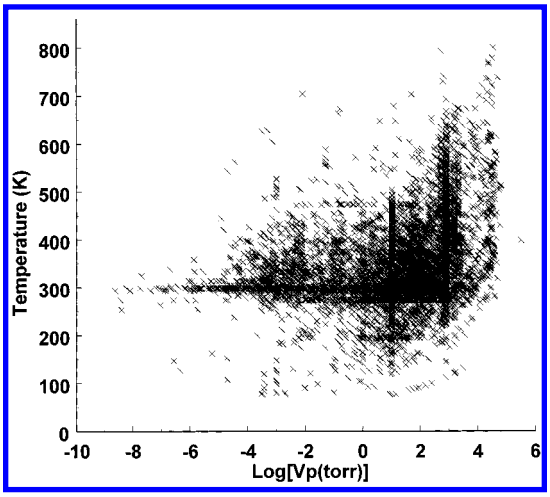



**Figure 1.** Vapor pressure [log(Vp)] and temperature distribution for training data.

K. Most of these data are concentrated however in the region $-4 < \log(\mathrm{Vp}) < 5$ and $200 < T < 600$ K. Several clusters of data are also visible in Figure 1, a horizontal band of 423 measurements corresponding to a temperature of 298 K. A vertical band of 457 measurements can also be seen at 2.88 log units, or atmospheric pressure, corresponding to the normal boiling point. Our dataset consists of molecules containing the following elements, with the number of molecules with this atom shown in parentheses; H, B (190), N (591), O (1,071), Si (112), P (119), S (190), F, Cl, Br, and I. Six hundred thirty-eight molecules in the training set contain halogen atoms, and all molecules contain carbon. Histograms showing the element and molecular weight distributions of the training molecules are shown in Figure 2.

## II. RESULTS AND DISCUSSION

**A. Performance of the Model.** Summaries of the results for the training, cross-validation, and validation sets, using 27:15:1 neural networks and the descriptors discussed above, are shown in Table 2. The results for these three sets are also shown graphically in Figures 3−5. We find that 63.8% and 65.2% of the results for the training and validation sets, respectively, are within their error bars. A robust model should have low maximum errors, and it can be seen in Figures 3 and 4 that some reasonably large deviations from experiment are observed. We have therefore examined the experimental values for the worst 20 predictions and found that in most of these cases, there is evidence that makes the accuracy of the experimental values doubtful. Several are

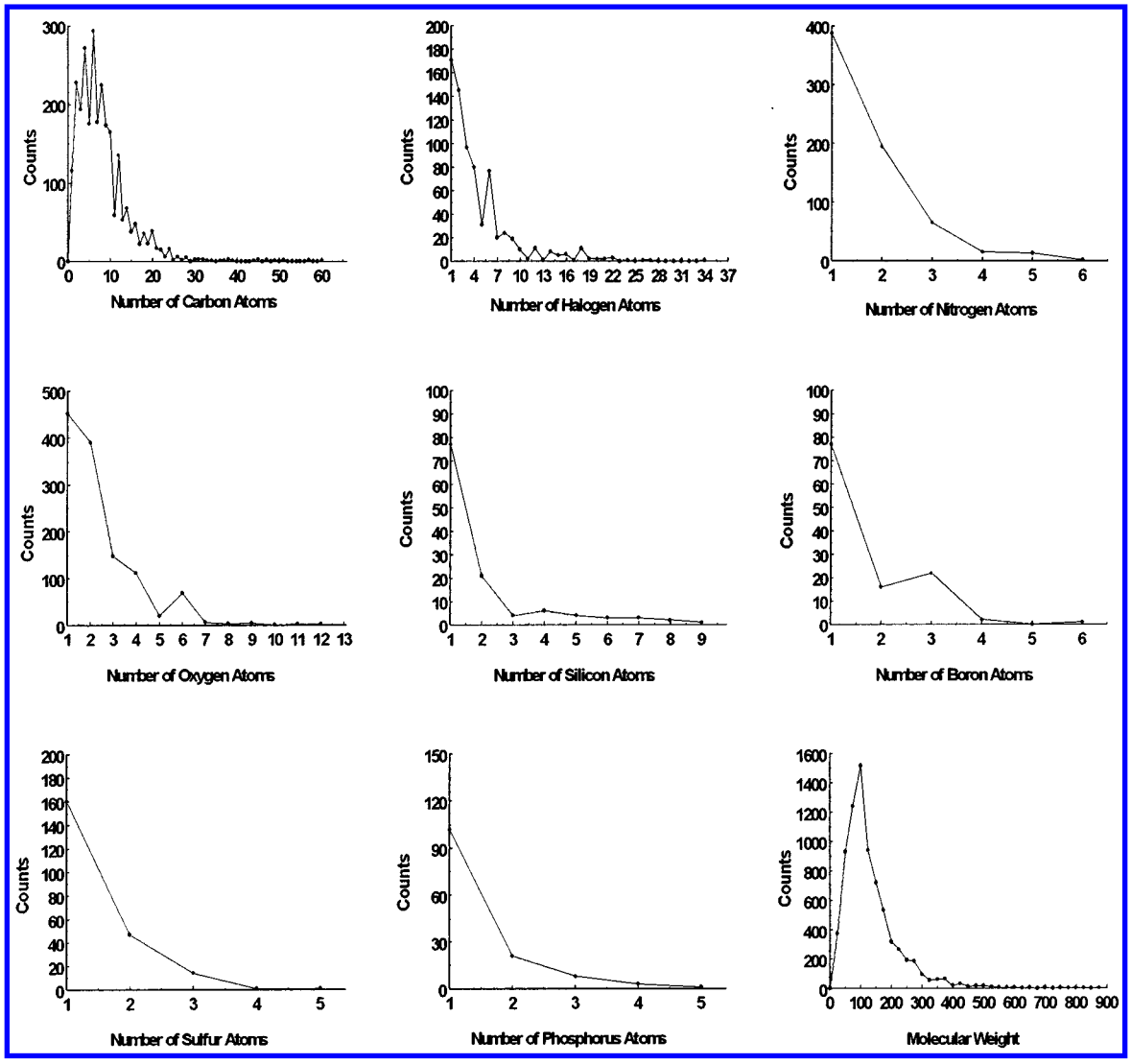



**Figure 2.** Distributions of elements and molecular weight for the training molecules.

**Table 2.** Comparisons of Predicted and Experimental Vapor Pressures (Log Vp)

| | correlation coeff ($R^2$) | SD | mean unsigned error | largest error |
|---|---|---|---|---|
| training[a] | 0.9762 | 0.322 | 0.211 | 2.02 |
| cross-validation[a] | 0.9516 | 0.459 | 0.291 | 4.15 |
| validation[b] | 0.9758 | 0.326 | 0.213 | 2.06 |

[a] $N = 7681$. [b] $N = 861$.

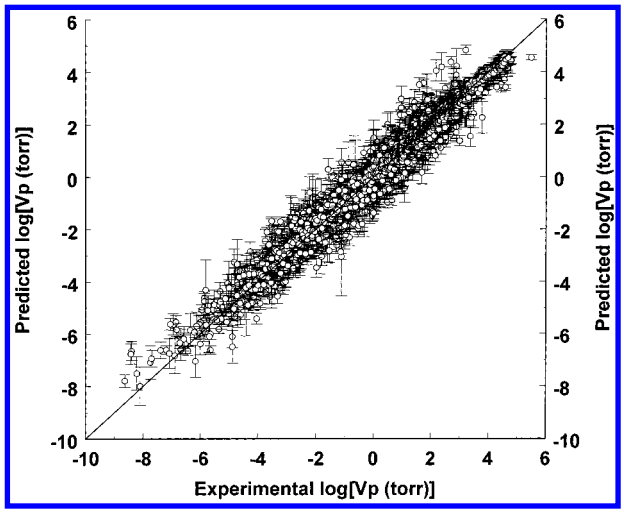



**Figure 3.** Predicted vs experimental vapor pressure [log(Vp)] for the train/test set.

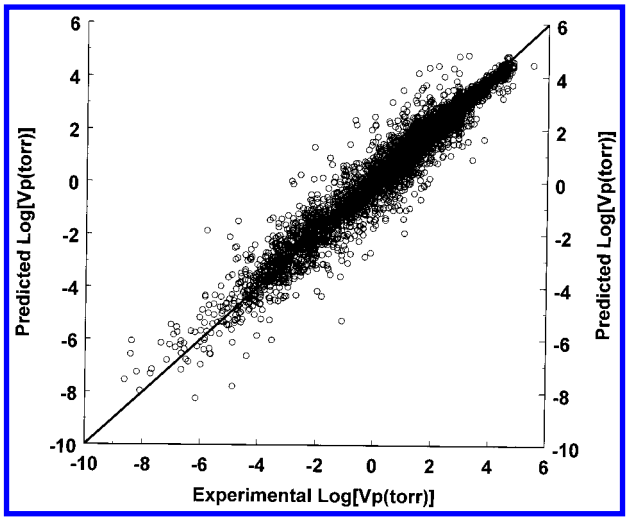



**Figure 4.** Cross-validated predicted vs experimental vapor pressure [log(Vp)] for the train/test set.

inconsistent with other experimental values or appear to be measured in Kelvin, rather than Celsius, as stated by the database. Many of the remaining outliers have only single measurements, so we cannot check the consistency of these results. The species with the largest error (1.94 log units) for which the experimental result appears consistent is difluoromethoxyborane, a molecule of a type not represented well in the training set. Due to the size of our training set, removal of a small number of possibly incorrect results will not affect the predictive power of the final model, hence these measurements were not removed from the training set.

An obvious feature of Figure 5 is the extremely large error bars on one prediction. This corresponds to species **1** below. The large error bars for this prediction are most likely due

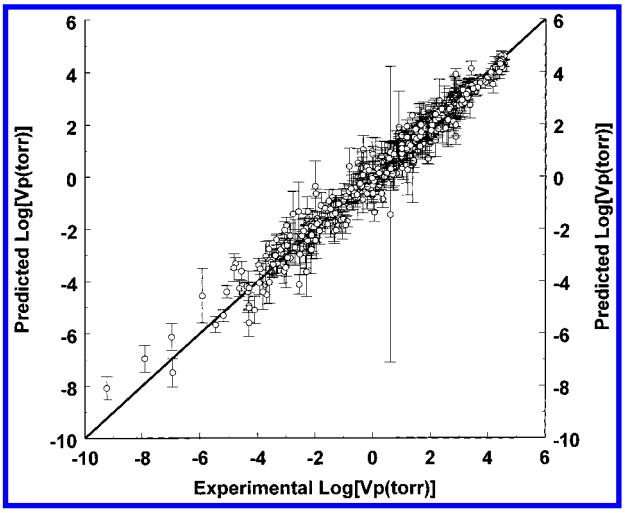



**Figure 5.** Predicted vs experimental log (vapor pressure) for the validation set.

**Table 3.** Comparisons of Predicted and Experimental Vapor Pressures (Log Vp) for McClelland and Jurs Dataset ($N = 420$)

| method | correlation coeff ($R^2$) | SD | mean unsigned error | weights per training input | largest error |
|---|---|---|---|---|---|
| present work | 0.971 | 0.228 | 0.158 | 0.054 | 1.24 |
| McClelland | 0.969 | 0.237 | 0.163 | 0.15 | 1.59 |

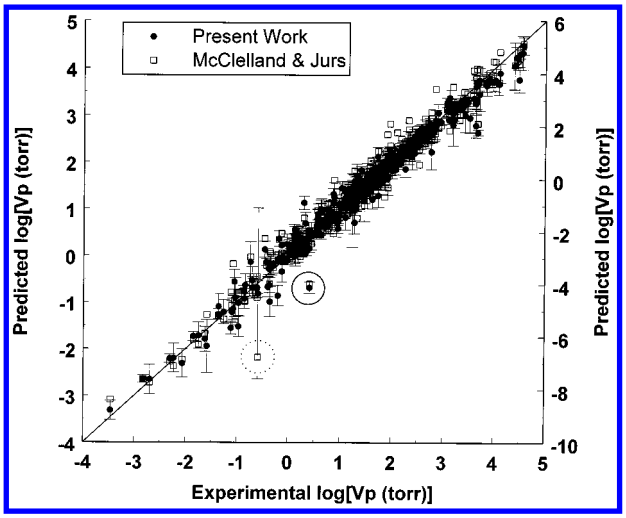
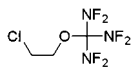



**Figure 6.** Comparison of vapor pressure prediction methods: predicted vs experimental vapor pressure [log(Vp)].

to the unusual nature of this molecule. There are no molecules in our training set containing N−F bonds.




**1**

To provide a basis for comparison for our model, we have investigated the 420 molecule dataset of McClelland and Jurs[5] which was used to create a neural network model for vapor pressure at 298 K. The results for these dataset using our model, and those of the 10-descriptor neural net model of McClelland and Jurs,[5] are summarized in Table 3 and shown graphically in Figure 6. The experimental values for this comparison were taken from McClelland and Jurs.[5] It can be seen that our model produces similar results to that

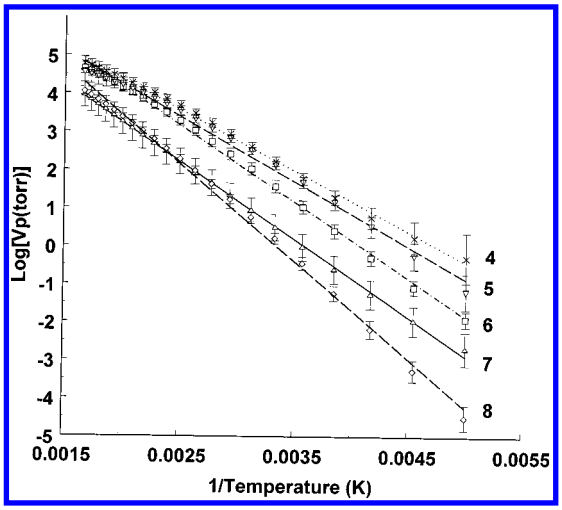



**Figure 7.** Sample temperature-dependence plots, log(Vp) vs temperature.

of McClelland and Jurs, despite the fact that we have less than half the number of weights per training input. Further improvement of this result is unlikely given that the accuracy of the predictions is already comparable with the expected experimental error.[5] Of these 420 molecules, measurements for only 72 appear in our database at a temperature of 298 K, with a further 286 having measurements at other temperatures.

Two interesting features are visible in Figure 6. The result for our largest outlier (shown with a solid circle, species **2** below) is very close to that calculated by McClelland and Jurs, a strong indication that the experimental value here is incorrect. The Beilstein database contains no entries for this species at 298 K. It can also be seen that in the case where the McClelland and Jurs model has its largest failure (dashed circle, molecule **3**), the prediction of our model is significantly better, although the error bars are relatively large. This indicates that the molecule is of a type not well represented in the two training sets.


**B. Temperature-Dependence.** One of the major aims of this study is to produce a temperature-dependent model. Since we make no assumptions about the form of this dependence, it is important to establish that our model behaves in a physically reasonable manner. In an ideal system, the relationship between temperature and vapor pressure (Vp) is given by the Clasius-Clapeyron equation, which can be written as follows[30]

$$\ln \mathrm{Vp} = \frac{-\Delta H_{\mathrm{vap}}}{RT} + C \qquad (1)$$

where $\Delta H_{\mathrm{vap}}$ is the enthalpy of vaporization and $C$ is a constant. According to this equation, there is a linear relationship between ln(Vp) (or log(Vp)) and $1/T$ with a slope of $-\Delta H_{\mathrm{vap}}/R$. In Figure 7 we have plotted log(Vp) vs $1/T$ for five molecules (Figure 8, species **4**−**8**) with temperatures ranging from 200 to 600 K, the range over which most of

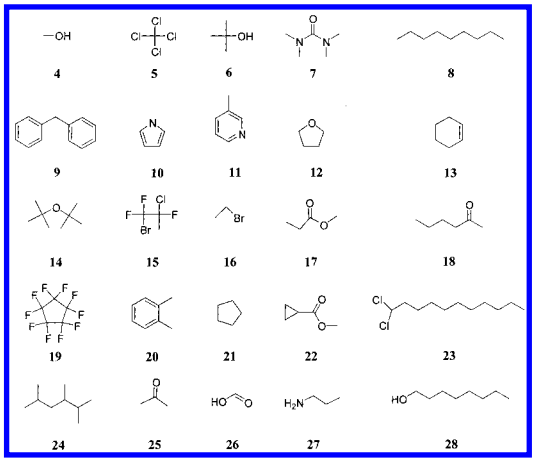



**Figure 8.** Molecule set for heat of vaporization predictions.

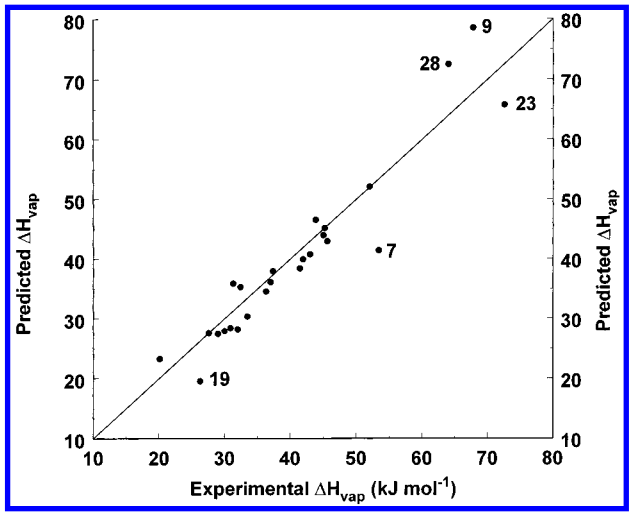



**Figure 9.** Predicted vs experimental enthalpy of vaporization ($\Delta H_{\mathrm{vap}}$).

the training data is concentrated. All these plots show little deviation from linearity over the temperature range studied. However, in some cases we have found that at very low vapor pressures (below −6 log units) linear behavior is no longer observed. Most likely this occurs because there are very few training data in this region, so that the interpolation procedure becomes less reliable.

From eq 1, it can be seen that the slope of the linear fits shown in Figure 7 is related to the enthalpy of vaporization. We have calculated ln(Vp) at 10 temperatures between 288 and 308 °C and used the slope of the linear regression fit to compute the enthalpy of vaporization for 25 molecules taken from the Beilstein database and the NIST Chemistry WebBook.[31] In all cases, the correlation coefficient ($R^2$) for the linear fit was found to be greater than 0.99. A comparison between the experimental (at 298 K) and calculated enthalpies of vaporization is shown in Figure 9. The standard deviation of the error for these predictions is 4.7 kJ mol$^{-1}$, and the maximum error is 12.0 kJ mol$^{-1}$. The five molecules producing the largest errors are labeled, and the structures can be found in Figure 8.

There is a causal relationship between vapor pressure and boiling points. The latter simply being the temperature at which the vapor pressure equals atmospheric pressure (760 Torr or 2.88 log units). Therefore, as a further test of our model's temperature-dependence, we have calculated the

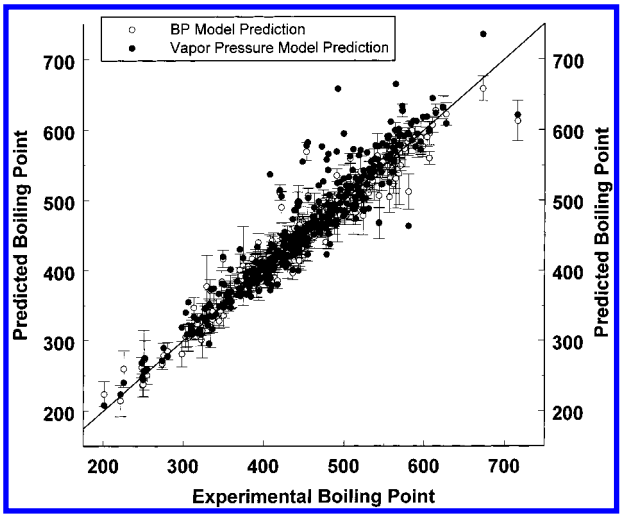



**Figure 10.** Comparison of predicted vs experimental boiling point predictions from boiling point and vapor pressure models.

**Table 4.** Comparison of Boiling Point Predictions with Boiling Point and Vapor Pressure Models ($N = 392$)

| method | correlation coeff ($R^2$) | SD (K) | mean unsigned error (K) | largest error (K) |
|---|---|---|---|---|
| present work | 0.889 | 28.59 | 18.68 | 166.1 |
| Boiling Point | 0.928 | 21.86 | 13.55 | 140.0 |

temperature corresponding to a log(Vp) of 2.88, i.e., that boiling point. We have achieved this using the bisection algorithm[32] for finding the root of the function VP($T$)-2.88 numerically, where VP represents the neural net output as a function of T.

Using the above procedure we have calculated the boiling points for a subset of 392 molecules from a previous study on boiling points.[22] Only three of these compounds appear in the vapor pressure model training set. The predicted and experimental points are shown graphically in Figure 10. For comparison, we have also plotted the predictions of our boiling point model.[22] Statistics for both methods are shown in Table 4. As expected, the boiling point model gives superior results. However, apart from a small number of outliers, the accuracy of the predictions from the vapor pressure model is still quite good. In fact from Figure 8, it can be seen that for a number of outliers the two models produce similar results. This suggests problems with the experimental values in these cases. We have made no allowance for the uncertainty of the predicted vapor pressure in this calculation, which would certainly result in improved results. Fifty-five percent of the predicted boiling points from the vapor pressure model lie within the error-limits of the boiling point model's prediction.

## III. CONCLUSIONS

We have presented a robust, temperature-dependent model for vapor pressure, based on a large and diverse training set. This training set is much larger than would be possible by restricting our model to measurements at 298 K. The large dataset also allows us to include much more diverse molecules than would otherwise be possible, allowing our model to be more generally applicable than models trained on smaller datasets. We have examined the temperature-dependence of our model and find that it is physically reasonable and accurate enough to make prediction of boiling point and enthalpy of vaporization possible. This model is found to perform comparably to a very recent model trained on a much smaller dataset.[5] Inclusion of error estimates allow the reliability of individual predictions to be gauged.

As with our previous study on boiling points,[22] we find that the large number of training samples used allows the model to be robust enough to identify erroneous experimental results. Given the large size of the set, it seems unlikely that we have identified all bad results. The large size of the dataset, however, also means that the remaining incorrect training datapoints will have little effect on the accuracy of the final model.

## ACKNOWLEDGMENT

We thank Oxford Molecular Ltd for financial support.

## REFERENCES AND NOTES

(1) Katritzky, A. R.; Maran, U.; Lobanov, V. S.; Karelson, M. Structurally Diverse Quantitative Structure−Property Relationship Correlations of Technologically Relevant Physical Properties. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1−8.
(2) Livingstone, D. J. The Characterization of Chemical Structures Using Molecular Properties. A Survey. *J. Chem. Inf. Comput. Sci.* **2000**, *40*.
(3) Hornsby, A. G.; Wauchope, R. D.; Herner, A. E. *Pesticide Properties in the Environment*; Springer-Verlag: New York, 1995.
(4) Liang, C.; Gallagher, D. A. QSPR Prediction of Vapor Pressure from Solely Theoretically-Derived Descriptors. *J. Chem. Inf. Comput. Sci.* **1998**, *1998*, 321−324.
(5) McClelland, H. E.; Jurs, P. C. Quantitative Structure−Property Relationships for the Prediction of Vapor Pressures of Organic Compounds from Molecular Stucture. *J. Chem. Inf. Comput. Sci.* **2000**, *50*, 967−975.
(6) Goll, E. S.; Jurs, P. C. Prediction of Vapor Pressures of Hydrocarbons and Halohydrocarbons from Molecular Structure with a Computational Neural Network Model. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1081−1089.
(7) Basak, S. C.; Gute, B. D.; Grunwald, G. D. Use of Topostructural, Topochemical, and Geometric Parameters in the Prediction of Vapor Pressure: A Hierarchial QSAR Approach. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 651−655.
(8) Beilstein; Beilstein Informationssysteme GmbH.: Frankfurt am Main, 2000.
(9) Sadowski, J.; Gasteiger, J. Corina 1.8; Oxford Molecular: Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, UK.
(10) Dewer, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. P. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(11) Clark, T.; Alex, A.; Beck, B.; Chandrasekhar, J.; Gedeck, P.; Horn, A.; Hutter, M.; Rauhut, G.; Sauer, W.; Steinke, T. VAMP 7.0; Oxford Molecular Ltd.: Medawar Centre, Oxford Science Park, Standford-on-Thames, Oxford, OX4 4GA, U.K., 1998.
(12) Beck, B.; Clark, T. Some Biological Applications of Semiemprical MO Theory. *Perspectives Drug Discovery Design* **1998**, *9/10/11*, 131−159.
(13) Rauhut, G.; Clark, T. Multicenter Point Charge Model for High Quality Molecular Electrostatic Potentials from AM1 Calculations. *J. Comput. Chem.* **1993**, *14*, 503−509.
(14) Beck, B.; Rauhut, G.; Clark, T. The Natural Atomic Orbital Point Charge Model for PM3: Multiple Moments and Molecular Electrostatic Potentials. *J. Comput. Chem.* **1995**, *15*, 1064−1073.
(15) Pascual-Ahuir, J. L.; Silla, E.; Tuñon, I. GEPOL: An improved Description of Molecular Surfaces III. A New Algorithm for the Computation of a Solvent-Excluded Surface. *J. Comput. Chem.* **1994**, *15*, 1127−1138.
(16) Marsili, M. In *Physical Property Prediction in Organic Chemistry*; Jochum, C., Hicks, M. G., Sunkel, J., Eds.; Springer: Berlin-Heidelberg, 1988; p 249.
(17) Martin, B.; Gedeck, P.; Clark, T. Additive NDDO−Based Atomic Polarizabilty Model. *Int. J. Quantum Chem.* **2000**, *77*, 473−497.
(18) Beck, B. Propgen 1.0; Oxford Molecular Ltd.: Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, UK, 2000.
(19) Tetko, I. G.; Villa, A. E. P.; Livingstone, D. J. Neural Network Studies. 2. Variable Selection. *J. Chem. Inf. Comput. Sci.* **1996**, *26*, 794−803.

(20) Murray, J. S.; Politzer, P. Statistical Analysis of the Molecular Surface Electrostatic Potential: An Approach to Describing Noncovalent Interaction in Condensed Phases. *J. Mol. Struct. (THEOCHEM)* **1998**, *425*, 107−114.
(21) Murray, J. S.; Lane, P.; Brinck, T.; Paulsen, K.; Grice, M. E.; Politzer, P. Relationships of Critical Constants and Boiling Points to Computed Molecular Surface Properties. *J. Phys. Chem.* **1993**, *97*, 9369−9373.
(22) Chalk, A. J.; Beck, B.; Clark, T. A Quantum Mechanical/Neural Net Model for Boiling Points With Error Estimation. *J. Chem. Inf. Comput. Sci.* **2000**, Accepted for publication.
(23) Cronce, D. T.; Famini, G. R.; DeSoto, J. A.; Wilson, L. Y. Using Theoretical Descriptors in Quantitative Structure−Property Relationships: Some Distribution Equilibria. *J. Chem. Soc., Perkin Trans. 2* **1998**, 1293−1301.
(24) Kier, L. B.; Hall, L. H. *Molecular Structure Description*; Academic Press: New York, 1999.
(25) Meyer, A. Y. The Size of Molecules. *Chem. Soc. Rev.* **1986**, *15*, 449−475.
(26) Müller, B.; Reinhardt, J.; Strickland, M. T. *Neural Networks − An introduction*, 2nd ed.; Springer-Verlag: Berlin, Heidelberg, 1995.
(27) Pao, Y.-H. *Adaptive Pattern Recognition and Neural Networks*; Addison-Wesley Publishing Company: Reading, 1989.
(28) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*; VCH Verlag: Weinheim, 1993.
(29) Beck, B.; Breindl, A.; Clark, T. QM/NN QSPR Models with Error Estimation: Vapor Pressure and logP. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1046−1051.
(30) Wedler, G. *Lehrbuch der Physikalischen Chemie*, 3rd ed.; VCH Verlagsgesellschaft: Weinheim, 1987; p 279.
(31) NIST Chemistry Web-book; http://webbook.nist.gov/chemistry/.
(32) Press: W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran*, 2nd ed.; Cambrige University Press: Cambridge, 1992; p 346.
(33) Jorgenson, W. L.; Duffy, E. M. Prediction of Drug Solubility from Monte Carlo Simulations. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1155−1158.
(34) Beck, B. 2000, Unpublished results.

CI0103222