Wright State University
CORE Scholar

Browse all Theses and Dissertations | Theses and Dissertations

2011

Quantum Chemical Analysis of the pKa's of Alcohols and a Cellular Automata Model for the Distribution of Gases in the Earth's Atmosphere

Edur Basha BoyiniPalli
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all

 Part of the Chemistry Commons

Repository Citation

BoyiniPalli, Edur Basha, "Quantum Chemical Analysis of the pKa's of Alcohols and a Cellular Automata Model for the Distribution of Gases in the Earth's Atmosphere" (2011). *Browse all Theses and Dissertations*. 421.
https://corescholar.libraries.wright.edu/etd_all/421

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

QUANTUM CHEMICAL ANALYSIS OF THE pK_a's OF ALCOHOLS AND A CELLULAR AUTOMATA MODEL FOR THE DISTRIBUTION OF GASES IN THE EARTH'S ATMOSPHERE

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

By

EDUR BASHA BOYINI PALLI
B.Sc., Osmania University, 2004

2011
Wright State University

WRIGHT STATE UNIVERSITY

SCHOOL OF GRADUATE STUDIES

March 31, 2011

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Edur Basha Boyini Palli ENTITLED Quantum Chemical Analysis Of The pK_a's Of Alcohols And A Cellular Automata Model For The Distribution Of Gases In The Earth's Atmosphere BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Paul G. Seybold, Ph.D.
Thesis Director

Ken Turnbull, Ph.D., Chair
Department of Chemistry

Committee on
Final Examination

David A. Dolson, Ph.D.

Eric Fossum, Ph.D.

Paul G. Seybold, Ph.D.

Andrew Hsu, Ph.D.
Dean, School of Graduate Studies

ABSTRACT

Boyini Palli, Edur Basha. M.S., Department of Chemistry, Wright State University, 2011. Quantum Chemical Analysis of the pK_a's of Alcohols and a Cellular Automata Model for the Distribution of Gases in the Earth's Atmosphere.

Alcohols play important roles in many chemical and biological processes, and their acid/base behaviors are often important for these roles. In this study we ask the question, "Can the electronic properties of these compounds offer clues to the compounds' acid/base behaviors?" The study considers whether selected quantum chemical properties can be used to find correlations with the experimental pK_a's of the alcohols (aliphatic and aromatic). Calculations were carried out for these alcohols using the semi-empirical RM1 method and the more advanced density functional theory (DFT) B3LYP/6-31+G* method. Significant correlations were found for several quantum chemical descriptors. It was also found that conformer selection plays an important role in obtaining the lowest energy form of each alcohol for analysis.

Almost 200 years ago John Dalton proposed that the composition of the gases in the Earth's atmosphere should change with altitude, the heavier gases being relatively more abundant at lower elevations and the lighter ones relatively more abundant at higher altitudes. In 2006 this proposal was experimentally confirmed at low altitudes (0-4 meters) by careful measurements of the ratio of argon to nitrogen [Y. Adachi et al., *Science* **2006**, *311*, 142] at a desert location. In the present work a dynamic, isothermal cellular automata model for the distributions of nitrogen, oxygen, argon, and carbon dioxide in the atmosphere is presented and compared with the predictions of the barometric equation. The cellular automata model employs two rules: a gravitational rule based on the masses of the molecular constituents and a motional rule based on their

relative average speeds. The model captures the basic features of the gas distribution with altitude as well as the expected relative uncertainties caused by the diffusive motions of the gas molecules.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Firstly and most importantly, thank you to Dr. Paul G. Seybold for the guidance and inspiration throughout the two project works and during my graduate journey.

Thank you to the department of chemistry and the faculty members for the support in the past two years. Special thanks to Dr. Dolson for his advice when I was new to the program.

Thank you to my fellow graduate students for being very friendly and helpful. I would also like to thank the university's computer centre for helping out whenever our computers in the lab needed any updates and installations.

Finally, big thanks to dad and mom for the emotional support and encouragement in the past two years.

Chapter 1

Introduction

This thesis is combination of two projects. The first project (Chapter 2) is a quantum chemical analysis of the pK_a's of alcohols. Previous work in our research group on the pK_a's of substituted anilines[1], phenols,[2] benzoic acids,[3] organic phenols,[4] amines,[5] and anilines,[6] was the initial inspiration to start this work. The main goal was to see if it was possible to extend the earlier studies on the pK_a's to systems such as the alcohols. This study was begun by searching the literature for experimentally measured pK_a's for the alcohols. The second step involved selecting descriptors which were successful in previous studies and calculating these descriptors using two quantum chemical approaches, the more rapid semi-empirical RM1 method[7] and the more elaborate and time-consuming density functional theory (DFT) procedure.[8] Once the theoretical descriptors were determined, the third stage involved seeking correlation equations between the descriptors and the pK_a's of the alcohols. The selection of computational methods was based on previous studies where these methods were successful in finding good correlations with the pK_a's.

The RM1 method was applied initially to understand and verify whether this less demanding method could give reasonable results. For the more demanding *ab initio* DFT computations selection of a suitable basis set was very important. Here the 6-31+G* basis set was used to employ the polarization and diffuse functions for the non-hydrogen "heavy atoms". In particular inclusion of diffuse functions was judged advisable for

proper representation of the anions. In all cases the most stable structures of the molecules were sought through careful examinations of the possible conformers. To understand the influence of the solvent medium on these results two solvent models were examined: the continuum SM 5.4 aqueous solvent model[9] and the more advanced reaction-field aqueous solvent SM8 model[10].

The investigation in Chapter 3 represents a radical departure from that in Chapter 2. In Chapter 3 the distribution of gases in the earth's atmosphere with altitude was studied using a stochastic cellular automata (CA) technique. Stochastic CA methods, in contrast to the usual approach in physics, are based on simple probabilistic rules rather than on the "deterministic" results that come from solutions of differential equations. As a result of their construction and statistical underpinning, stochastic CA methods yield probabilistic results which include uncertainties, as distinguished from the "absolute" results obtained by solving the differential equations. The CA results obtained for the distributions of the gases in the atmosphere were compared with those predicted by the classic barometric equation,[11] which represents a deterministic solution for the atmospheric profile. The barometric equation doesn't provide any information on the fluctuations in the atmospheric concentrations of the gases that result from the normal diffusional motions of the gas molecules, whereas the CA approach, while coinciding in its average form with the barometric profile, in addition yields reasonable values for the uncertainties and density fluctuations that appear in the natural system. For statistical reasons the atmospheric gases were studied individually, starting with nitrogen and continuing to oxygen, argon, and carbon dioxide. The CA simulations were run multiple

times to obtain the variations in final patterns of these gases. Even though the CA simulations were repeated using the same rules for each particular gas, the results, while generally similar, were not absolutely identical due to the stochastic nature of the governing rules. The differing results thereby provided measures of the uncertainties present in the natural, dynamic system. The software program CASim[12] by Cheng, Kier, and Seybold was used for the CA simulations; it provides a dynamic picture of the system as it progresses from an initial stage of random placement of the agents on the grid to a later stage of dynamic equilibrium.

1.1 References

[1] Gross, K.C.; Seybold, P.G.; Hadad, C.M. Comparison of Different Atomic Charge Schemes for Predicting pK_a variations in Substituted Anilines and Phenols. *Int. J. Quantum Chem*. **2002**, *90*, 445.

[2] Gross, K.C.; Seybold, P.G. Substituent Effects on the Physical Properties and pK_a of Phenol. *Int. J. Quantum Chem*. **2001,** *85*, 569-579.

[3] Hollingsworth, C. A.; Seybold, P. G.; Hadad, C. M. Substituent Effects on the Electronic Structure and pK_a of Benzoic Acid. *Int. J. Quantum Chem*. **2002,** *90*, 1396-1403.

[4] Kreye, W. C.; Seybold, P. G. Correlations Between Quantum Chemical Indicies and the pK_a's of Diverse Set of Organic Phenols. *Int. J. Quantum Chem*. **2009,** *109*, 3679-3684.

[5] Seybold, P.G. Analysis of the pK_a's of Aliphatic Amines Using Quantum Chemical Descriptors. *Int. J. Quantum Chem*. **2008**, *108*, 2849-2855.

[6] Gross, K. C.; Seybold, P. G. Substituent Effects on the Structure and pK_a of Aniline,

Int. J. Quantum Chem. **2000**, *80*, 1107-1115.

[7] Rocha, G. B.; Freire, R.O.; Simas, A.M,; Stewart, J. J. P. RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J Comput Chem.* **2006**, 27, 1101.

[8] Robert G. Parr. Density Functional Theory. *Annu. Rev. Phys. Chem*. **1983**, *34*, 631-656.

[9] Chambers, C. C,; Hawkins, G. D,; Cramer, C. J,; Truhlar, D. G. A Model for Aqueous Solvation Based on Class IV Atomic Charges and First-Solvation-Shell Effects. *J Phys Chem*. **1996**, *100*, 16385.

[10] Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011

[11] R. A. Alberty, R. J. Silbey, *Physical Chemistry, 2nd edn* **1997**, (John Wiley: New York, NY) p.22.

[12] C.-K. Cheng, L .B. Kier, and P. G. Seybold, *CASim Version 1.0*, Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond ,VA 23284, USA, **2005.**

Chapter 2

Quantum Chemical Analysis of the pK_a's of Alcohols.

2.1 Introduction

Regression analyses correlating experimental pK_a's with various quantum chemical parameters have been performed in the past for a variety of chemical compound classes such as anilines[1], phenols,[2] benzoic acids,[3] organic phenols,[4] amines,[5] and anilines.[6] This chapter is a study of the aliphatic and aromatic alcohols, and seeks to determine whether analogous correlation equations between the pK_a's and quantum chemical parameters can be found for these compounds. As a first step, pK_a data were collected from the literature for various aliphatic and aromatic alcohols. Calculations were then carried out for these alcohols using the semi-empirical RM1 method[7] and the more advanced density functional theory (DFT) procedure.[8]

The following questions were asked before starting the research:

1. The previous studies on different compounds gave good correlations between the pKa's and selected descriptors using several methods. Is the RM1 semi-empirical method useful in getting good correlations for the separate studies of (a) aliphatic alcohols, (b) aromatic alcohols, and (c) aliphatic and aromatic alcohols combined together?

2. What effect does employment of the SM 5.4 aqueous solvent model[9] have on the RM1 results?

3. Is DFT method capable of yielding more accurate results than the RM1 method?

4. What are the effects of adding the SM 5.4 solvent model to the DFT results?

5. How well do DFT single-point calculations (using the gas-phase equilibrium geometries) perform when combined with the more advanced SM8 model?[10]

6. Is it necessary to optimize the geometries of the alcohols within the SM8 solvent model to get good results?

2.2 Methods

In order to find a correlation between the pK_a's and quantum chemical descriptors, experimental pK_a's of 95 alcohols were collected from various resources in preparation for the correlation studies. The main source for the experimental pK_a's was the *CRC Handbook of Chemistry and Physics*.[11] However, the CRC Handbook did not list a large number of pK_a values for aliphatic alcohols so other literature sources were also employed[12,13,14,15,16,17.] Table 2.1a summarizes the pK_a values for 29 aliphatic alcohols and Table 2.1b summarizes the pK_a values for 66 aromatic alcohols investigated in this study. The experimental pK_a values for the aromatic were collected from a recent work done by Zhang et al.[18] Zhang et al. studied different classes of compounds like alcohols, phenols, and carboxylic acids to predict the pK_a values[18]. The pK_a values reported in present work are all at temperature 25^0 C. The pK_a values fall in a range from 0.37 to 19.2, a very big range for a study of this sort. It is clear from the data collected that the aliphatic alcohols are weak acids, much weaker than the corresponding values found for the phenols.

The Spartan'10 (v.1.0.) [19] program was used for the quantum chemical calculations. The semi-empirical RM1 method was first employed for these calculations on the

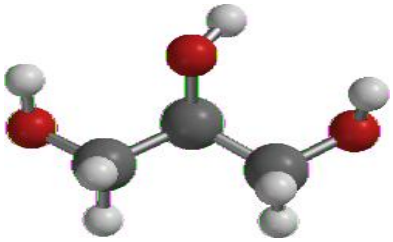
compounds in Tables 2.1a and 2.1b. Regression analyses were performed using Microsoft Office Excel 2010 on a MPC Client pro 365 workstation. After completing the RM1 calculations, further calculations were performed using the more advanced DFT method at the B3LYP/6-31+G* level. The basis set 6-31+G* was used to add polarization (*) and diffuse (+) functions to the basis set. Polarization functions were used for the heavy atoms (carbons and halogens) and diffuse functions were deemed especially desirable for proper description of the anionic forms. Two solvent models were examined - the SM5.4 model of Chambers et al. and the more recent SM8 model of Marenich et al.

The quantum chemical parameters selected were (1) the natural charge of the hydroxyl oxygen of the alcohol $Q_n(O)$, (2) the natural charge on the acidic proton of the alcohol $Q_n(H)$, (3) the total combined natural charge of the hydroxyl group of the alcohol $Q_n(OH)$, (4) the natural charge of the oxide oxygen of the anion $Q_n(O^-)$, (5) the difference in the gas phase energy between the alcohol and its anion, ΔE, and (6) the corresponding energy difference in the aqueous phase between the alcohol and its anion, ΔE_{aq} , this being the result found using the SM 5.4 solvent model. Since these descriptors gave successful regressions in the previous studies,[1-6] the aim was to find out whether they would be successful for this class of compounds.

In the regression studies the pK_a value was the dependent variable and the quantum chemical descriptors were the independent variables. The statistical results of special interest in the regressions were the number of compounds (n), the coefficient of determination (r^2) , the standard error (s), and the Fisher statistic (F). The errors associated with the parameter coefficients were also taken into account, since this reflects

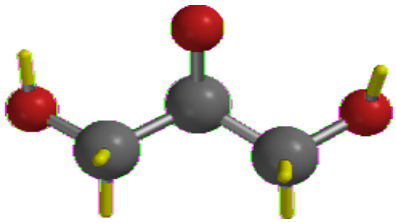
the significance of the parameter in the regression. When regressions were analyzed it was found some of the compounds were outliers in each category. The outliers were those compounds with standard residual values (the difference between experimental pK_a values and predicted pK_a values from regression analysis) of more than two standard deviations.

Also, while doing the calculations different conformers were studied for all of the compounds in order to find the lowest energy conformer. One example is glycerol:



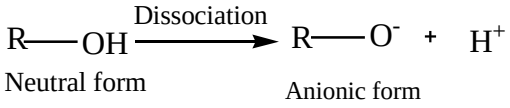
Scheme 1: Glycerol's most stable neutral form.

The RM1 calculations were employed to find out which structure of glycerol is most stable. This was done by changing the dihedral angles between the atoms. The most stable form gave energy, E of -612.01 kJ/mol (E_{aq} was -647.85 kJ/mol). In order to find which hydrogen dissociates first, calculations were done separately by successively removing a hydrogen atom from each of the three hydroxyl groups. The lowest energy anionic species was found to be the secondary anionic form shown in Scheme 2.



Scheme 2: Glycerol's most stable anionic form.

The dissociation scheme for the alcohols is:



In the more advanced DFT method, calculations were first performed on the alcohols in vacuum, which also yielded values for the SM5.4 solvent model. The purpose of introducing the solvent model was to get the difference in the aqueous phase energy between the alcohol and its anion. The parameters selected for the study were then subjected to more advanced SM8 solvent model. Finally, compounds were optimized fully in SM8 solvent model to see whether optimization improves the results or not.

2.3 Results and discussions

RM1 Calculations

The quantum chemical descriptors obtained by RM1 for aliphatic and aromatic alcohols are shown in Table 2.2 and 2.3 respectively. The summation of regression results obtained using the descriptors are shown in Tables 2.4-2.6. The first step was regression analysis of aliphatic alcohols and the best regression was obtained for the $Q_n(OH)$ ($r^2 = 0.658$) for a total of 29 aliphatic alcohols (Table2.4). For $Q_n(OH)$ the best regression was found by omitting 2 outliers (2-methyl-2-propanol and chloral hydrate). The r^2 value improved to 0.754 after omitting the outliers. This is shown in Equation 2.1 and plotted in Figure 2.1.

$$pK_a = -61.2\ (\pm 6.9)\ Q_n(OH) - 7.0\ (\pm 0.8) \qquad (2.1)$$

$$n = 27 \quad r^2 = 0.754 \quad s = 0.817 \quad F = 77$$

The SM 5.4 solvent model was not effective, since it gave $r^2 = 0.572$ (ΔE_{aq}) for the aliphatic alcohols.

The regression analysis of aromatic alcohols was performed separately as the second step. The difference of energies in gas phase ΔE, proved to be the best descriptor, with $r^2 = 0.838$ for a total of 66 aromatic alcohols (Table 2.5). The r^2 improved to 0.918 after excluding 6 outliers (2,6-dichlorophenol, 2,6-dinitrophenol, 4-methylsulfonylphenol, 2-hydroxypyridine, 2-hydroxypyrazine, and 5-amino-8-hydroxyquinoline). This is shown in Equation 2.2 and plotted in Figure 2.2.

$$pK_a = -0.040\ (\pm 0.001)\ \Delta E - 12.9\ (\pm 0.2) \qquad (2.2)$$

$n = 60 \quad r^2 = 0.918 \quad s = 0.530 \quad F = 649$

Addition of the SM 5.4 solvent model was not effective for the aromatic alcohols, giving $r^2 = 0.829$ (ΔE_{aq}) for the entire set of aromatics.

The third step was combined regression analysis of aliphatic and aromatic alcohols. The regression results for combined analysis proved to be more successful than the aliphatic and aromatic separately, which is good since the pKa's for 95 alcohols covers a wide range. The best regression for 95 alcohols was obtained for $Q_n(OH)$, $r^2 = 0.897$ (Table 2.6), and this improved to 0.944 after excluding 5 outliers (2-methyl-2-propanol, chloral hydrate, 2-hydroxypyridine, 4-amino-8-hydroxyquinoline, and 5-amino-8-hydroxyquinoline). This is shown in Equation 2.3 and plotted in Figure 2.3.

$$pK_a = -64.1\ (\pm 1.7)\ Q_n(OH) - 6.5\ (\pm 0.1) \qquad (2.3)$$

$n = 90 \quad r^2 = 0.944 \quad s = 0.753 \quad F = 1487$

The outliers for aliphatic alcohols (2-methyl-2-propanol, chloral hydrate) and the outliers for the aromatics (2-hydroxypyridine, 5-amino-8-hydroxyquinoline) appeared as

outliers in combined analysis also. The introduction of solvent using the SM 5.4 solvent model did improve the results for the combined analysis of alcohols from $r^2 = 0.479$ (ΔE) to $r^2 = 0.679$ (ΔE_{aq}). Overall for the aliphatic and aromatic compounds the solvent model SM 5.4 gave better regressions but individually, for the aliphatic and aromatic compounds separately, it didn't prove to be useful.

DFT calculations:

a) Vacuum and SM 5.4 solvent model calculations

After getting an idea which descriptors gave good correlations with experimental pKa's from the RM1 method, the set of aliphatic, aromatic and combined alcohols were treated with more advanced DFT method at the B3LYP/6-31+G* level. The results obtained for the descriptors are shown in Table 2.7 and 2.8 for aliphatic and aromatic alcohols respectively. The summation of regression results obtained using the descriptors are shown in Tables 2.9-2.11. For the aliphatic alcohols the best regression was obtained for the Q_n (OH) ($r^2 = 0.779$) for 29 aliphatic alcohols (Table 2.9). The r^2 improved to 0.841 after excluding 2 outliers (3, 3-diflouropropanol, and 2-methyl-2-propanol). This is shown in Equation 2.4 and plotted in Figure 2.4.

$$pK_a = -91.6\ (\pm 7.9)\ Q_n(OH) - 10.4\ (\pm 2.1) \qquad (2.4)$$

$$n = 27 \qquad r^2 = 0.841 \qquad s = 0.696 \qquad F = 131$$

The SM 5.4 solvent model was not at all effective, yielding $r^2 = 0.523$ for ΔE and $r^2 = 0.243$ for ΔE_{aq}.

For the aromatic alcohols the best regression was obtained for the Q_n (O) ($r^2 = 0.883$) for a total of 66 aromatic alcohols (Table 2.10). The r^2 improved to 0.949 after excluding

9 outliers (2, 6-dinitrophenol, 3-hydroxyquinoline, 2,4,5-trichlorophenol, 2,3,4-trichlorophenol, pentachlorophenol, 2-hydroxypyridine, 4-amino-8-hydroxyquinoline, 5-amino-8-hydroxyquinoline, and 8-hydroxy-4-methylquinoline). This is shown in Equation 2.5 and plotted in Figure 2.5.

$$pK_a = -122.3\ (\pm 3.8)\ Q_n\ (O) - 76.6\ (\pm 2.6) \qquad (2.5)$$

$$n = 57 \quad r^2 = 0.949 \quad s = 0.403 \quad F = 1044$$

The SM 5.4 solvent model was not effective with $r^2 = 0.832$ for ΔE decreasing to $r^2 = 0.779$ for ΔE_{aq}.

For the combined regression analysis the best regression was obtained for Q_n (OH) , $r^2 = 0.905$ (Table 2.11) and improved to 0.953 after excluding 7 outliers (3,3,3-triflouro-2-methylprop-2-nol, 2-methyl-2-propanol, 2-hydroxypyridine, 4-amino-8-hydroxyquinoline, 5-amino-8-hydroxyquinoline, 8-hydroxy-2-methylquinoline, and 8-hydroxy-4-methylquinoline). This is shown in Equation 2.6 and plotted in Figure 2.6

$$pK_a = -71.5\ (\pm 1.7)\ Q_n(OH) - 4.90\ (\pm 0.37) \qquad (2.6)$$

$$n = 88 \quad r^2 = 0.953 \quad s = 0.695 \quad F = 1758$$

The outliers for the aliphatic alcohols (2-methyl-2-propanol) and the outliers for the aromatics (2-hydroxypyridine, 4-amino-8-hydroxyquinoline, and 5-amino-8-hydroxyquinoline) were the common outliers in the combined analysis for the aliphatic and aromatic alcohols. The SM 5.4 solvent model was not effective with $r^2 = 0.886$ for ΔE decreasing to $r^2 = 0.775$ for ΔE_{aq}.

b) SM8 solvent model

Single point calculations using the vacuum DFT geometries were performed with the more advanced solvent model SM8 at the B3LYP/6-31+G* level. The results obtained for the descriptors are shown in Table 2.12 and 2.13 for aliphatic and aromatic alcohols respectively. The regression results obtained using the descriptors are shown in Tables 2.14-2.16. For the aliphatic alcohols, the best regression was obtained for the Q_n (OH) ($r^2 = 0.721$) for 29 aliphatic alcohols (Table 2.14). The r^2 improved to 0.808 after excluding 2 outliers (1,2,3,4-butanetetrol, and 2-methyl-2-propanol). This is shown in Equation 2.7 and plotted in Figure 2.7.

$$pK_a = -84.7\ (\pm 8.3)\ Q_n(OH) - 9.4\ (\pm 2.2) \qquad (2.7)$$

$$n = 27 \quad r^2 = 0.808 \quad s = 0.772 \quad F = 105$$

For the aromatic alcohols the best regression was obtained for the Q_n(OH) ($r^2 = 0.894$) for 66 aromatic alcohols (Table 2.15). The r^2 improved to 0.957 after excluding 10 outliers (2-aminophenol, 2-bromophenol, 2,3-dichlorophenol, 2,6-dichlorophenol, 4-nitrophenol, 6-hydroxyisoquinoline, 2,3,4-trichlorophenol, 2,3,4,6-tetrachlorophenol, pentachlorophenol, and 2-hydroxypyridine). This is shown in Equation 2.8 and plotted in Figure 2.8.

$$pK_a = -77.1\ (\pm 2.2)\ Q_n(OH) - 6.22\ (\pm 0.44) \qquad (2.8)$$

$$n = 56 \quad r^2 = 0.957 \quad s = 0.399 \quad F = 1192$$

For the combined regression analysis of aliphatic and aromatic alcohols, the best regression for 95 alcohols was obtained for Q_n(OH) , $r^2 = 0.936$ (Table 2.16) and improved to 0.965 after excluding 7 outliers (3,3-diflouropropanol, 3,3,3-triflouro-2-

methylprop-2-nol, 2-butanol, 1,2,3,4-butanetetrol, chloral hydrate, 2-methyl-2-propanol, and 2-hydroxypyridine). This is shown in Equation 2.9 and plotted in Figure 2.9

$$pK_a = -69.3\ (\pm 1.4)\ Q_n(OH) - 4.85\ (\pm 0.31) \qquad (2.9)$$

$$n = 88 \quad r^2 = 0.965 \quad s = 0.572 \quad F = 2406$$

The outliers for aliphatic alcohols (1,2,3,4-butanetetrol, and 2-methyl-2-propanol) and the outliers for the aromatics (2-hydroxypyridine) were the common outliers in the combined analysis for the aliphatic and aromatic alcohols. The ΔE_{aq} descriptor gave poor regressions in all three categories (Table 2.14-2.16).

Finally, the geometries of the alcohols were optimized within the SM8 solvent medium. The results obtained for the descriptors are shown in Table 2.17 and 2.18 for aliphatic and aromatic alcohols respectively. The regression results obtained using the descriptors are shown in Tables 2.19-2.21. For the aliphatic alcohols, the best regression was obtained for the $Q_n(OH)$ ($r^2 = 0.739$) for 29 aliphatic alcohols (Table 2.19). The r^2 improved to 0.883 after excluding 3 outliers (3,3-diflouropropanol, 1,2,3,4-butanetetrol, and 2-methyl-2-propanol). This is shown in Equation 2.10 and plotted in Figure 2.10.

$$pK_a = -81.2\ (\pm 6.1)\ Q_n(OH) - 8.4\ (\pm 1.7) \qquad (2.10)$$

$$n = 26 \quad r^2 = 0.883 \quad s = 0.608 \quad F = 181$$

The optimization improved the results for the aliphatic alcohols, improving $r^2 = 0.739$ for $Q_n(OH)$ from 0.721 for 29 aliphatic alcohols. The improvement was also seen in regressions for $Q_n(OH)$ without outliers as $r^2 = 0.808$ (27 alcohols) increased to 0.883(26 alcohols).

For the aromatic alcohols the best regression was obtained for the $Q_n(OH)$ ($r^2 = 0.895$) for 66 aromatic alcohols (Table 2.20). The r^2 value improved to 0.957 after excluding 10 outliers (2-aminophenol, 2-bromophenol, 2,3-dichlorophenol, 2,6-dichlorophenol, 4-nitrophenol, 6-hydroxyisoquinoline, 2,3,4-trichlorophenol, 2,3,4,6-tetrachlorophenol, pentachlorophenol, and 2-hydroxypyridine). This is shown in Equation 2.11 and plotted in Figure 2.11.

$$pK_a = -78.64\ (\pm 2.3)\ Q_n(OH) - 6.5\ (\pm 0.45) \qquad (2.11)$$

$$n = 56 \quad r^2 = 0.957 \quad s = 0.397 \quad F = 1206$$

There was no significant improvement in the results for $Q_n(OH)$ regressions after optimization for aromatic alcohols, changing $r^2 = 0.895$ from 0.894. The regressions results without outliers for $Q_n(OH)$ remained same as $r^2 = 0.957$(56 alcohols).

For the combined regression analysis, the best regression for the 95 alcohols was obtained for $Q_n(OH)$, $r^2 = 0.938$ (Table 2.21) and improved to 0.965 after excluding 7 outliers (3,3-diflouropropanol, 2-butanol, 1,2,3,4-butanetetrol, chloral hydrate, 2-methyl-2-propanol, pentachlorophenol, and 2-hydroxypyridine). This is shown in Equation 2.12 and plotted in Figure 2.12

$$pK_a = -68.5\ (\pm 1.4)\ Q_n(OH) - 4.71\ (\pm 0.31) \qquad (2.12)$$

$$n = 88 \quad r^2 = 0.965 \quad s = 0.562 \quad F = 2417$$

The common outliers in this case were 3,3-diflouropropanol, 1,2,3,4-butanetetrol, and 2-methyl-2-propanol from aliphatic and 2-hydroxypyridine from aromatic alcohols. There was no significant improvement in results for $Q_n(OH)$ after optimization, changing $r^2 = 0.938$ from 0.936. The regressions results without outliers remained same as $r^2 = 0.965$ (88 alcohols).

2.4 Conclusions:

Here are the possible answers to the questions asked earlier in this chapter. The first question was whether RM1 method could be useful in getting good correlations for the aliphatic alcohols*?* Yes, the RM1 method was successful in getting good correlations. In particular, Q_n(OH) was the best descriptor for the aliphatic and combined alcohols, whereas for the aromatic alcohols ΔE proved to be the best descriptor. Overall the best regression results were obtained for the combined analysis of the alcohols which consisted of 95 alcohols for Q_n(OH). This result is encouraging, since the analysis covers a wide range of pK_a's.

The second question was, what effect does employment of the SM 5.4 aqueous solvent model have on the RM1 results? Employment of the SM 5.4 aqueous solvent model improved the results in each category of study, but there was not a huge difference. For the aliphatic and aromatic alcohols there was very small improvement in the regressions, which cannot be considered as successful. But for the combined study, again the introduction of solvent model SM 5.4 improved the r^2 for ΔE_{aq} significantly, giving r^2 = 0.679 whereas r^2 for ΔE was 0.479. So it can be concluded that the solvent model SM 5.4 does play a reasonable role in getting better regressions.

The Third question was, "Is DFT capable of yielding more accurate results than the RM1 method?" Yes, DFT gave good results compared to RM1 for the analysis of the aliphatic, aromatic and combined analysis at B3LYP/6-31+G* level, with Q_n(OH) again proving to be the best descriptor for aliphatic and combined analysis. But for the aromatics Q_n(O) also gave good regression results.

The fourth question was, "What are the effects of adding the SM 5.4 solvent model on to the DFT results?" Addition of the SM 5.4 solvent model to the DFT results did not improve the results. In fact, it often gave poorer results. In all three cases studies, the regression results for ΔE_{aq} dropped from those for ΔE. So, even though the SM 5.4 solvent model helped in the RM1 calculations, it didn't produce better results for the DFT method at the B3LYP/6-31+G* level.

The fifth question was, "What are the effects when single point calculations are performed with the more advanced solvent model SM8?" Introduction of the more advanced model did change the results, improving regressions for $Q_n(OH)$ for the aromatic and combined studies. But for the aliphatics there was no improvement. The SM8 regressions for ΔE_{aq} were unsuccessful while with the SM 5.4 model they were better. So, overall we can say the more advanced SM8 solvent model produced better results than the SM 5.4 solvent model for the $Q_n(OH)$ regressions, but failed for the parameter ΔE_{aq} .

The final question was, is it necessary to optimize the geometries of the aliphatic alcohols within the SM8 solvent medium to get good results? The best regressions were found for the $Q_n(OH)$ descriptor and the r^2 values did not improve much from that of the single point SM8 solvent model results. Therefore, we can conclude there is no need for optimization of alcohols in the SM8 solvent medium.

If we look at the regressions obtained for various descriptors selected for the alcohol study, we can conclude that for the RM1 method, $Q_n(OH)$ turned out to be the best descriptor for the aliphatic and combined alcohols. For the aromatics even though $Q_n(OH)$ gave good results, ΔE gave the best r^2 values. $Q_n(OH)$ continued to be the best

descriptor in the DFT method, both for the SM 5.4 and the SM8 solvent models, with the exception that $Q_n(O)$ was the best descriptor for the aromatic alcohols in SM5.4 medium. The possible reason why $Q_n(OH)$ gave good regressions is that it accounts for the total charge on the hydroxyl group which in turn is an integral part of the dissociating step. And the possible reasons for getting outliers in every case can be the wrong experimental pK_a value or perhaps the most stable conformer was not identified correctly. The equations obtained from the regression analysis in this study suggest that the pK_a value goes up as the $Q_n(OH)$ becomes more negative. The electron donating groups and withdrawing groups affects the charge on oxygen and hence the dissociation of hydrogen is also affected. The electron withdrawing groups decreases the negative charge on oxygen making the OH bond to break easily, whereas the electron donating groups increases the negative charge on oxygen making it difficult for the hydrogen to dissociate. In earlier studies, Zhang et al[18] discovered ΔE, as good descriptor (the only descriptor they examined) for all three cases aliphatic, aromatic, and combined alcohols. But then the solvent model used was COSMO[20] to calculate zero point energies using basis set 6-311+G**. In present study more descriptors were examined compared to Zhang et al[18] where only one descriptor was examined. And, it can also be concluded that selection of basis set and solvation model plays important role in getting good correlations.

Table 2.1a: Experimental pK_a's for the Aliphatic Alcohols

No.	Compound Name	Molecular formula	pK_a
1	Methanol	CH_3OH	15.51[11]
2	Ethanol	CH_3CH_2OH	15.5[11]
3	2-Bromoethanol	$BrCH_2CH_2OH$	14.38[17]
4	2-Ethoxyethanol	$C_2H_5OCH_2CH_2OH$	14.98[17]
5	2-Methoxyethanol	$CH_3OCH_2CH_2OH$	14.87[17]
6	2,2-Dichloroethanol	$CHCl_2CH_2OH$	12.89[14]
7	2,2-Dibromoethanol	Br_2CHCH_2OH	13.29[17]
8	2,2,2-trifloroethanol	F_3CCH_2OH	12.37[11],12.5[13]
9	2,2,2-Trichloroethanol	Cl_3CCH_2OH	12.24[11]
10	2,2,2-Tribromoethanol	Br_3CCH_2OH	12.7[17]
11	2-Cynoethanol	$NCCH_2CH_2OH$	14.03[17]
12	Ethylene glycol	$HOCH_2CH_2OH$	15.1[11],14.77[14]
13	Chloral Hydrate	$CCl_3CH(OH)_2$	11.0[12]
14	2,2,2-Triflouro-1-(p -tolyl)ethanol	$F_3CC(C_7H_7)OH$	12.04[16]
15	2,2,2-Triflouro-1-(4-methoxy phenyl)ethanol	$F_3CC(C_7H_7O)OH$	12.24[16]
16	1-Propanol	$CH_3CH_2CH_2OH$	16.1[15]
17	2-Propanol	$CH_3CHOHCH_3$	17.1[15]
18	2-Methyl-2-propanol	$CH_3CCH_3OHCH_3$	19.2[16]
19	2-Methyl-2-propen-1-ol	$CH_3CCH_2CH_2OH$	14.82[14]
20	3,3-Diflouropropanol	$CHF_2CH_2CH_2OH$	12.74[14]
21	2,2,3,3-Tetraflouropropanol	$F_2CF_2CCH_2OH$	12.74[16]
22	3,3,3-Triflouro-2-methylprop-2-nol	$CF_3C(CH_3)_2OH$	11.6[14]
23	Glycerol	$CH_2OHCHOHCH_2OH$	14.15[11]
24	Propargyl alcohol	$HCCCH_2OH$	13.6[11],13.55[14]
25	Allyl Alcohol	CH_2CHCH_2OH	15.5[11]
26	1-Butanol	$CH_3CH_2CH_2CH_2OH$	16.1[16]
27	2-Butanol	$CH_3CH_2CHOHCH_3$	17.6[16]
28	1,2,3,4-Butanetetrol	$HOCH_2CHOHCHOHCH_2OH$	13.9[11]
29	Phenyl methanol	$C_6H_5CH_2OH$	15.4[16]

Table 2.1b: Experimental pK_a's for Aromatic Alcohols.[a]

30	Phenol	C_6H_5OH	9.98
31	2-Aminophenol	C_6H_6NOH	9.96
32	3-Aminophenol	C_6H_6NOH	9.96
33	4-Aminophenol	C_6H_6NOH	10.46
34	2-Bromophenol	C_6H_4BrOH	8.45
35	3-Bromophenol	C_6H_4BrOH	8.87
36	4-Bromophenol	C_6H_4BrOH	9.35
37	2-tert-Butylphenol	$(CH_3)_3CC_6H_4OH$	10.62
38	3-tert-Butylphenol	$(CH_3)_3CC_6H_4OH$	10.12
39	4-tert-Butylphenol	$(CH_3)_3CC_6H_4OH$	10.23
40	2-Chlorophenol	C_6H_4ClOH	8.53
41	3-Chlorophenol	C_6H_4ClOH	8.88
42	4-Chlorophenol	C_6H_4ClOH	9.38
43	3-Cyanophenol	C_7H_4NOH	8.57
44	4-Cyanophenol	C_7H_4NOH	7.95
45	2,3-Dichlorophenol	$C_6H_3Cl_2OH$	7.71
46	2,4-Dichlorophenol	$C_6H_3Cl_2OH$	7.89
47	2,5-Dichlorophenol	$C_6H_3Cl_2OH$	7.51
48	2,6-Dichlorophenol	$C_6H_3Cl_2OH$	6.81
49	3,4-Dichlorophenol	$C_6H_3Cl_2OH$	8.62
50	3,5-Dichlorophenol	$C_6H_3Cl_2OH$	8.18
51	2,5-Dimethylphenol	C_8H_9OH	10.41
52	2,6-Dimethylphenol	C_8H_9OH	10.6
53	3,4-Dimethylphenol	C_8H_9OH	10.36
54	3,5-Dimethylphenol	C_8H_9OH	10.19
55	2,4-Dinitrophenol	$C_6H_4N_2O_5$	4.12
56	2,5-Dinitrophenol	$C_6H_4N_2O_5$	5.2
57	2,6-Dinitrophenol	$C_6H_4N_2O_5$	3.73
58	2-Flourophenol	C_6H_4FOH	8.7
59	3-Flourophenol	C_6H_4FOH	9.29
60	4-Flourophenol	C_6H_4FOH	9.81
61	2-Methylphenol	C_7H_7OH	10.32
62	3-Methylphenol	C_7H_7OH	10.09

63	4-Methylyphenol	C_7H_7OH	10.27
64	2-Methoxyphenol	$C_7H_8O_2$	9.98
65	3-Methoxyphenol	$C_7H_8O_2$	9.65
66	4-Methoxyphenol	$C_7H_8O_2$	10.24
67	3-Methylsulfonylphenol	$C_7H_8O_3S$	8.75
68	4-Methylsulfonylphenol	$C_7H_8O_3S$	7.83
69	2-Nitrophenol	$C_6H_5NO_3$	7.22
70	3-Nitrophenol	$C_6H_5NO_3$	8.36
71	4-Nitrophenol	$C_6H_5NO_3$	7.14
72	5-Hydroxyisoquinoline	C_9H_7NO	8.47
73	6-Hydroxyisoquinoline	C_9H_7NO	9.17
74	7-Hydroxyisoquinoline	C_9H_7NO	8.9
75	8-Hydroxyisoquinoline	C_9H_7NO	8.42
76	3-Hydroxyquinoline	C_9H_7NO	8.08
77	5-Hydroxyquinoline	C_9H_7NO	8.56
78	6-Hydroxyquinoline	C_9H_7NO	8.9
79	7-Hydroxyquinoline	C_9H_7NO	8.87
80	8-Hydroxyquinoline	C_9H_7NO	9.71
81	2,4,5-Trichlorophenol	$C_6H_2Cl_3OH$	7.07
82	2,3,4-Trichlorophenol	$C_6H_2Cl_3OH$	7.1
83	2,3,4,6-Tetrachlorophenol	C_6HCl_4OH	5.62
84	Pentachlorophenol	C_6Cl_5OH	4.9
85	2,4,6-Trimethylphenol	$C_9H_{11}OH$	10.89
86	2,4,6-Trinitrophenol	$C_6H_3N_3O_7$	0.37
87	2-Hydroxypyridine	C_5H_4NOH	11.99
88	2-Hydroxypyrazine	$C_4H_3N_2OH$	7.28
89	4-Amino-8-hydroxyquinoline	$C_9H_7N_2OH$	10.71
90	5-Amino-8-hydroxyquinoline	$C_9H_7N_2OH$	11.24
91	8-Hydroxy-2-methylquinoline	$C_{10}H_8NOH$	10.16
92	8-Hydroxy-4-methylquinoline	$C_{10}H_8NOH$	9.99
93	4-Chloro-3,5-dimethylphenol	C_8H_8ClOH	9.7
94	Dichloroxylenol	$C_8H_7Cl_2OH$	8.28
95	4-Hydroxybenzothiazole	C_7H_5NOS	8.85

[a] pK_a data collected from ref. 18.

Table 2.2: RM1 Calculation Data for Aliphatic Alcohols

				Neutral				Anion				
No.	Compound	Q(O)	Q(H)	Q(OH)	E	E aq	Q(O^-)	E	E aq	Del(E)	Del(E)aq	pKa
1	Methanol	-0.328	0.189	-0.139	-207.21	-227.07	-0.863	-82.34	-450.12	-124.87	-223.05	15.51
2	Ethanol	-0.33	0.19	-0.14	-241.49	-257.51	-0.785	-157.55	-469.24	83.94	-211.73	15.5
3	2-Bromoethanol	-0.319	0.193	-0.126	-225.62	-254.72	-0.757	-210.17	-474.99	15.45	-220.27	14.38
4	2-Ethoxyethanol	-0.331	0.205	-0.126	-411.5	-428.68	-0.773	-343.14	-637.44	68.36	-208.76	14.98
5	2,2,2-triflouro-1-(p-tolyl)ethanol	-0.302	0.217	-0.085	-830.98	-845.16	-0.558	-903.57	-1132.51	-72.59	-287.35	12.04
6	2,2,2-triflouro-1-(4-methoxy phenyl) ethanol	-0.302	0.215	-0.087	-948.07	-968.18	-0.556	-1029.32	-1261.44	-81.25	-293.26	12.24
7	2-methoxyethanol	-0.324	0.19	-0.134	-377.69	-400.91	-0.773	-317.42	-615.8	60.27	-214.89	14.87
8	2,2-Dibromoethanol	-0.316	0.209	-0.107	-204.1	-232.24	-0.709	-214.97	-451.27	-10.87	-219.03	13.29
9	2,2-Dichloroethanol	-0.315	0.214	-0.101	-299.17	-324.47	-0.711	-317.26	-561.5	-18.09	-237.03	12.89
10	2,2,2 tribromoethanol	-0.303	0.21	-0.093	-166.27	-191.4	-0.628	-197.44	-413.81	-31.17	-222.41	12.7
11	2,2,2-trichloroethanol	-0.298	0.215	-0.083	-302.71	-325.23	-0.377	-384.24	-596.88	-81.53	-271.65	12.24
12	2,2,2-trifloroethanol	-0.288	0.212	-0.076	-896.44	-915.14	-0.55	-955.56	-1210.2	-59.12	-295.06	12.37
13	Cynoethanol	-0.329	0.204	-0.125	-67.94	-99.99	-0.821	-25.99	-323.19	41.95	-223.2	14.03
14	ethylene glycol	-0.337	0.2	-0.137	-405.63	-433.3	-0.761	-356.77	-666.99	48.86	-233.69	15.1
15	1-Propanol	-0.33	0.189	-0.141	-262.55	-277.72	-0.781	-183.61	-483.04	78.94	-205.32	16.1
16	2-Propanol	-0.337	0.194	-0.143	-271.19	-283.73	-0.772	-198.54	-488.73	72.65	-205	17.1
17	2-methyl-2-propanol	-0.339	0.196	-0.143	-315.33	-324.91	-0.759	-246.28	-520.09	69.05	-195.18	19.2
18	2-methyl-2-propene-1-ol	-0.326	0.193	-0.133	-168.05	-177.78	-0.775	-116.16	-394.03	51.89	-216.25	14.82
19	2,2,3,3-tetraflouropropanol	-0.291	0.211	-0.08	-1047.41	-1066.89	-0.671	-1105	-1337.38	-57.59	-270.49	12.74
20	3,3-diflouropropanol	-0.328	0.2	-0.128	-651.87	-670.02	-0.754	-625.68	-896.49	26.19	-226.47	12.74
21	3,3,3-triflouro-2-methylprop-2-nol	-0.307	0.217	-0.09	-972.18	-984.59	-0.551	-1028.26	-1273.76	-56.08	-289.17	11.6
22	1-Butanol	-0.333	0.192	-0.141	-275.76	-289.36	-0.779	-206.33	-499.17	69.43	-209.81	16.1
23	2-Butanol	-0.335	0.193	-0.142	-290.86	-298.97	-0.766	-222.81	-498	68.05	-199.03	17.6
24	1,2,3,4-Butanetetrol	-0.328	0.209	-0.119	-803.67	-844.52	-0.599	-796.64	-1052.16	7.03	-207.64	13.9
25	Allyl Alcohol	-0.328	0.193	-0.135	-139.86	-152.46	-0.78	-73.14	-364.54	66.72	-212.08	15.5
26	Benzyl Alcohol	-0.329	0.196	-0.133	-107.95	-121.14	-0.761	-59.02	-324.06	48.93	-202.92	15.4
27	Chloral Hydrate	-0.342	0.229	-0.113	-504.4	-532.1	-0.404	-632.46	-849.11	-128.06	-317.01	11
28	Propargyl alcohol	-0.32	0.196	-0.124	13.38	-10.17	-0.755	49.5	-244.39	36.12	-234.22	13.6
29	Glycerol	-0.344	0.203	-0.141	-608.36	-639.65	-0.473	-628.71	-947.21	-20.35	-307.56	14.15

Table 2.3: RM1 Calculation Data for Aromatic Alcohols

	Aromatic alcohols			Neutral				Anion				
No.	Compound	Q(O)	Q(H)	Q(O+H)	E	E aq	Q(O⁻)	E	E aq	Δ(E)	Δ(E)aq	pKa
1	Phenol	-0.269	0.214	-0.055	-96.39	-111.66	-0.564	-166.94	-414.08	70.55	302.42	9.98
2	2-Aminophenol	-0.273	0.242	-0.031	-106.27	-126.81	-0.516	-157.56	-414.61	51.29	287.8	9.96
3	3-Aminophenol	-0.272	0.214	-0.058	-129.15	-160.14	-0.564	-192.57	-451.51	63.42	291.37	9.96
4	4-Aminophenol	-0.271	0.207	-0.064	-117.63	-148.78	-0.563	-173.4	-428.18	55.77	279.4	10.46
5	2-Bromophenol	-0.262	0.234	-0.028	-79.73	-92.41	-0.507	-177.93	-399.36	98.2	306.95	8.45
6	3-Bromophenol	-0.262	0.218	-0.044	-77.63	-97.95	-0.527	-194.2	-409.19	116.57	311.24	8.87
7	4-Bromophenol	-0.263	0.218	-0.045	-78.86	-99.44	-0.538	-189.82	-407.38	110.96	307.94	9.35
8	2-tert-Butylphenol	-0.288	0.219	-0.069	-191.4	-197.39	-0.559	-272.69	-487.1	81.29	289.71	10.62
9	3-tert-Butylphenol	-0.27	0.213	-0.057	-205.23	-217.31	-0.562	-276.33	-506.5	71.1	289.19	10.12
10	4-tert-Butylphenol	-0.269	0.213	-0.056	-204.96	-216.86	-0.556	-279.63	-505.64	74.67	288.78	10.23
11	2-Chlorophenol	-0.261	0.235	-0.026	-136.36	-148.17	-0.51	-232.57	-459.27	96.21	311.1	8.53
12	3-Chlorophenol	-0.262	0.219	-0.043	-132.94	-151.12	-0.53	-247.64	-467.92	114.7	316.8	8.88
13	4-Chlorophenol	-0.263	0.218	-0.045	-132.52	-151.1	-0.539	-243.06	-465.68	110.54	314.58	9.38
14	3-Cyanophenol	-0.261	0.219	-0.042	37.76	11.84	-0.538	-81.95	-297.53	119.71	309.37	8.57
15	4-Cyanophenol	-0.26	0.221	-0.039	33.87	6.45	-0.514	-107.7	-321.2	141.57	327.65	7.95
16	2,3-Dichlorophenol	-0.257	0.238	-0.019	-161.6	-176	-0.49	-289.65	-500.06	128.05	324.06	7.71
17	2,4-Dichlorophenol	-0.256	0.238	-0.018	-168.6	-182.46	-0.489	-301.11	-507.65	132.51	325.19	7.89
18	2,5-Dichlorophenol	-0.254	0.239	-0.015	-169.15	-182.86	-0.481	-305.94	-510.22	136.79	327.36	7.51
19	2,6-Dichlorophenol	-0.236	0.237	0.001	-159.04	-173.47	-0.46	-291.86	-502.16	132.82	328.69	6.81
20	3,4-Dichlorophenol	-0.258	0.222	-0.036	-157.34	-178.22	-0.514	-299.09	-505.87	141.75	327.65	8.62
21	3,5-Dichlorophenol	-0.255	0.223	-0.032	-164.31	-184.17	-0.503	-316.8	-515.72	152.49	331.55	8.18
22	2,5-Dimethylphenol	-0.273	0.217	-0.056	-171.59	-184.04	-0.556	-245.21	-478.09	73.62	294.05	10.41
23	2,6-Dimethylphenol	-0.271	0.217	-0.054	-168.51	-177.8	-0.55	-245.17	-473.93	76.66	296.13	10.6
24	3,4-Dimethylphenol	-0.269	0.213	-0.056	-172.81	-187.01	-0.558	-244.06	-478.47	71.25	291.46	10.36
25	3,5-Dimethylphenol	-0.27	0.214	-0.056	-175.62	-189.87	-0.561	-245.04	-482.17	69.42	292.3	10.19
26	2,4-Dinitrophenol	-0.266	0.291	0.025	-129.35	-120.16	-0.398	-374.97	-537.77	245.62	417.61	4.12
27	2,5-Dinitrophenol	-0.268	0.285	0.017	-117.33	-105.94	-0.413	-342.55	-505.77	225.22	399.83	5.2
28	2,6-Dinitrophenol	-0.238	0.291	0.053	-116.32	-104.93	-0.46	-291.86	-502.16	175.54	397.23	3.73
29	2-Flourophenol	-0.256	0.226	-0.03	-281.06	-291.97	-0.529	-376.04	-613.56	94.98	321.59	8.7
30	3-Flourophenol	-0.264	0.22	-0.044	-290.84	-303.65	-0.545	-393.02	-625.11	102.18	321.46	9.29
31	4-Flourophenol	-0.264	0.216	-0.048	-285.86	-299.25	-0.553	-379.8	-613.23	93.94	313.98	9.81
32	2-Methylphenol	-0.272	0.217	-0.055	-131.98	-145.07	-0.557	-206.14	-443.88	74.16	298.81	10.32
33	3-Methylphenol	-0.269	0.214	-0.055	-135.99	-150.75	-0.563	-205.98	-448.06	69.99	297.31	10.09
34	4-Methylyphenol	-0.269	0.213	-0.056	-134.81	-149.51	-0.559	-206.78	-445.58	71.97	296.07	10.27
35	2-Methoxyphenol	-0.266	0.231	-0.035	-250.84	-262.93	-0.526	-316.82	-557.01	65.98	294.08	9.98
36	3-Methoxyphenol	-0.268	0.217	-0.051	-259.94	-280.88	-0.55	-341.14	-585.97	81.2	305.09	9.65
37	4-Methoxyphenol	-0.268	0.212	-0.056	-252.73	-273.88	-0.562	-324.47	-568.79	71.74	294.91	10.24
38	3-Methylsulfonylphenol	-0.26	0.22	-0.04	-450.04	-507.02	-0.538	-581.79	-830.31	131.75	323.29	8.75
39	4-Methylsulfonylphenol	-0.253	0.229	-0.024	-470.06	-528.03	-0.471	-675.76	-912.38	205.7	384.35	7.83
40	2-Nitrophenol	-0.28	0.28	0	-120.74	-116.89	-0.445	-280.45	-486.91	159.71	370.02	7.22
41	3-Nitrophenol	-0.257	0.223	-0.034	-112	-125.51	-0.522	-255.34	-455.59	143.34	330.08	8.36
42	4-Nitrophenol	-0.254	0.225	-0.029	-119.36	-134.79	-0.481	-302.7	-503.15	183.34	368.36	7.14
43	5-Hydroxyisoquinoline	-0.263	0.219	-0.044	3.31	-30.72	-0.508	-130.99	-357.4	134.3	326.68	8.47
44	6-Hydroxyisoquinoline	-0.264	0.217	-0.047	-3.74	-40.73	-0.515	-131.41	-363.63	127.67	322.9	9.17
45	7-Hydroxyisoquinoline	-0.264	0.216	-0.048	-0.67	-36.44	-0.519	-122.93	-351.45	122.26	315.01	8.9
46	8-Hydroxyisoquinoline	-0.262	0.219	-0.043	2.08	-33.07	-0.508	-139.38	-366.85	141.46	333.78	8.42
47	3-Hydroxyquinoline	-0.253	0.216	-0.037	5.65	-27.17	-0.513	-121.35	-347.55	127	320.38	8.08
48	5-Hydroxyquinoline	-0.264	0.216	-0.048	5.54	-26.95	-0.513	-127.74	-359.2	133.28	332.25	8.56
49	6-Hydroxyquinoline	-0.266	0.216	-0.05	0.51	-32.52	-0.518	-116.14	-344.73	116.65	312.21	8.9
50	7-Hydroxyquinoline	-0.266	0.22	-0.046	-2.77	-36.63	-0.513	-114.94	-350.22	112.17	313.59	8.87
51	8-Hydroxyquinoline	-0.267	0.246	-0.021	-6.75	-28.04	-0.462	-105.16	-335.43	98.41	307.39	9.71
52	2,4,5-Trichlorophenol	-0.251	0.24	-0.011	-190.75	-205.65	-0.466	-352.55	-545.31	161.8	339.66	7.07
53	2,3,4-Trichlorophenol	-0.253	0.24	-0.013	-184.71	-200.53	-0.474	-339.26	-536.89	154.55	336.36	7.1
54	2,3,4,6-Tetrachlorophenol	-0.23	0.242	0.012	-203.11	-219.14	-0.428	-389.08	-574.02	185.97	354.88	5.62
55	Pentachlorophenol	-0.228	0.244	0.016	-217.3	-232.48	-0.415	-422.13	-598.74	204.83	366.26	4.9
56	2,4,6-Trimethylphenol	-0.271	0.216	-0.055	-206.94	-215.67	-0.546	-284.56	-506.08	77.62	290.41	10.89
57	2,4,6-Trinitrophenol	-0.225	0.301	0.076	-111.94	-90.54	-0.327	-422.28	-550.24	310.34	459.7	0.37
58	2-Hydroxypyridine	-0.261	0.24	-0.021	-76.41	-105.73	-0.517	-162.32	-435.89	85.91	330.16	11.99
59	2-Hydroxypyrazine	-0.246	0.244	-0.002	-29.83	-72.54	-0.495	-152.64	-430.34	122.81	357.8	7.28
60	4-Amino-8-hydroxyquinoline	-0.269	0.247	-0.022	-38.96	-75.42	-0.47	-119.61	-363.85	80.65	288.43	10.71
61	5-Amino-8-hydroxyquinoline	-0.271	0.239	-0.032	-22.2	-58.2	-0.456	-109.15	-345.9	86.95	287.7	11.24
62	8-Hydroxy-2-methylquinoline	-0.268	0.246	-0.022	-43.74	-61.7	-0.46	-139.82	-364.92	96.08	303.22	10.16
63	8-Hydroxy-4-methylquinoline	-0.268	0.246	-0.022	-44.42	-65.19	-0.461	-140.52	-368.06	96.1	302.87	9.99
64	4-Chloro-3,5-dimethylphenol	-0.265	0.217	-0.048	-213.84	-229.58	-0.539	-318.62	-535.8	104.78	306.22	9.7
65	Dichloroxylenol	-0.259	0.238	-0.021	-250.05	-260.79	-0.489	-374.48	-578.19	124.43	317.4	8.28
66	4-Hydroxybenzothiazole	-0.26	0.244	-0.016	14.71	-16.49	-0.475	-102.51	-342.1	117.22	325.61	8.85

Table 2.4: Summation of Regression results from RM1 method for the aliphatic alcohols

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	6.0	1.1	-68.3	9.7	0.658	1.161	49	29
$Q_n(O^-)$	7.2	1.4	-10.1	2.1	0.457	1.439	22	29
ΔE	14.1	0.2	0.02	0.01	0.536	1.331	31	29
ΔE_{aq}	23.6	1.5	0.03	0.01	0.572	1.278	36	29
$Q_n(O)$	-10.7	6.1	-77.7	18.9	0.385	1.532	16	29
$Q_n(H)$	43.3	4.1	-143.5	20.3	0.647	1.160	49	29

Table 2.5: Summation of Regression results from RM1 method for the aromatic alcohols

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	6.7	0.1	-62.1	4.55	0.744	0.992	186	66
$Q_n(O^-)$	-7.5	1.6	-32.1	3.17	0.614	1.217	102	66
ΔE	13.1	0.2	-0.03	0.002	0.838	0.787	333	66
ΔE_{aq}	25.6	0.9	-0.01	0.01	0.829	0.812	309	66
$Q_n(O)$	-25.3	3.7	-129.9	14.2	0.565	1.294	83	66
$Q_n(H)$	24.4	1.8	-68.2	8.1	0.520	1.359	69	66

Table 2.6: Summation of Regression results from RM1 method for the aliphatic and aromatic alcohols

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	6.6	0.1	-63.3	2.2	0.897	1.036	815	95
$Q_n(O^-)$	-2.5	0.8	-22.7	1.5	0.716	1.777	219	95
ΔE	13.1	0.3	-0.03	0.01	0.479	2.309	92	95
ΔE_{aq}	11.9	0.2	-0.01	0.01	0.679	1.851	195	95
$Q_n(O)$	-16.1	1.3	-94.8	4.7	0.812	1.403	402	95
$Q_n(H)$	36.5	1.9	-117.9	8.7	0.660	1.887	180	95

Table 2.7: Data for Aliphatic Alcohols using B3LYP/6-31+G* using SM 5.4 solvent model

	Aliphatic alcohols			Neutral				Anion				
No.	Compound	Q(O)	Q(H)	Q(OH)	E	E aq	Q(O-)	E	E aq	Δ(E)	Δ(E)aq	pKa
1	Methanol	-0.777	0.49	-0.287	-115.725	-115.732825	-1.044	-115.104	-115.244	-0.62139	-0.48905	15.51
2	Ethanol	-0.768	0.484	-0.284	-155.044	-155.050085	-1.002	-154.43	-154.559	-0.61317	-0.49065	15.5
3	2-Bromoethanol	-0.763	0.493	-0.27	-2728.33	-2728.33548	-0.853	-2727.75	-2727.85	-0.57751	-0.48512	14.38
4	2-Ethoxyethanol	-0.773	0.496	-0.277	-308.88	-308.887043	-0.993	-308.263	-308.387	-0.61727	-0.50036	14.98
5	2-methoxyethanol	-0.767	0.486	-0.281	-269.56	-269.56866	-1.016	-268.951	-269.075	-0.60875	-0.49409	14.87
6	2,2-Dibromoethanol	-0.759	0.501	-0.258	-5301.6	-5301.61146	-0.849	-5301.03	-5301.13	-0.57098	-0.48618	13.29
7	2,2-Dichloroethanol	-0.758	0.499	-0.259	-1074.23	-1074.23749	-0.951	-1073.64	-1073.75	-0.58572	-0.48675	12.89
8	2,2,2 tribromoethanol	-0.744	0.498	-0.246	-7874.86	-7874.87032	-0.886	-7874.29	-7874.39	-0.57145	-0.48091	12.7
9	2,2,2-trichloroethanol	-0.75	0.503	-0.247	-1533.81	-1533.82302	-0.866	-1533.25	-1533.34	-0.56714	-0.48566	12.24
10	2,2,2-trifloroethanol	-0.754	0.504	-0.25	-452.786	-452.791739	-0.945	-452.212	-452.314	-0.57428	-0.47768	12.37
11	Cynoethanol	-0.761	0.492	-0.269	-247.288	-247.299118	-0.985	-246.699	-246.812	-0.58896	-0.48754	14.03
12	ethylene glycol	-0.78	0.503	-0.277	-230.26	-230.271426	-0.961	-229.673	-229.787	-0.5867	-0.48426	15.1
13	1-Propanol	-0.774	0.487	-0.287	-194.36	-194.36673	-0.966	-193.753	-193.87	-0.60746	-0.49721	16.1
14	i-PrOH	-0.771	0.488	-0.283	-194.363	-194.368314	-0.992	-193.753	-193.874	-0.61005	-0.49479	17.1
15	2,2,3,3-tetraflouropropanol	-0.755	0.507	-0.248	-591.331	-591.336664	-0.961	-590.765	-590.86	-0.56617	-0.47624	12.74
16	3,3-diflouropropanol	-0.767	0.494	-0.273	-392.844	-392.850174	-1.008	-392.245	-392.36	-0.59932	-0.49048	12.74
17	3,3,3-triflouro-2-methylprop-2-nol	-0.76	0.502	-0.258	-531.428	-531.431411	-0.928	-530.85	-530.945	-0.57826	-0.48627	11.6
18	1-Butanol	-0.77	0.488	-0.282	-233.675	-233.68097	-1.015	-233.057	-233.179	-0.61754	-0.50244	16.1
19	2-Butanol	-0.783	0.491	-0.292	-233.68	-233.684463	-0.954	-233.077	-233.185	-0.60233	-0.49901	17.6
20	1,2,3,4-Butanetetrol	-0.775	0.508	-0.267	-459.326	-459.341821	-0.447	-458.755	-458.864	-0.57154	-0.47822	13.9
21	Allyl Alcohol	-0.768	0.49	-0.278	-193.125	-193.130728	-0.951	-192.523	-192.638	-0.60223	-0.49296	15.5
22	Benzyl Alcohol	-0.769	0.491	-0.278	-346.788	-346.79505	-0.94	-346.195	-346.299	-0.59288	-0.49622	15.4
23	Chloral Hydrate	-0.739	0.51	-0.229	-1609.04	-1609.04692	-0.802	-1608.48	-1608.57	-0.55331	-0.47499	11
24	Propargyl alcohol	-0.76	0.494	-0.266	-191.865	-191.875415	-0.945	-191.273	-191.387	-0.5917	-0.48841	13.6
25	2,2,2-triflouro-1-(p-tolyl)ethanol	-0.745	0.501	-0.244	-723.166	-723.170983	-0.911	-722.575	-722.662	-0.59088	-0.50934	12.04
26	2,2,2-triflouro-1-(4-methoxy phenyl) ethanol	-0.741	0.495	-0.246	-798.326	-798.335881	-0.949	-797.756	-797.854	-0.56984	-0.48184	12.24
27	Glycerol	-0.797	0.513	-0.284	-344.794	-344.807866	-0.943	-344.219	-344.347	-0.57522	-0.46114	14.15
28	2-methyl-2-propanol	-0.781	0.488	-0.293	-233.684	-233.687948	-0.961	-233.081	-233.188	-0.60322	-0.49978	19.2
29	2-methyl-2-propen-1-nol	-0.78	0.494	-0.286	-232.443	-232.447911	-0.959	-231.845	-231.954	-0.59861	-0.49348	14.82

Table 2.8: Data for Aromatic Alcohols using B3LYP/6-31+G* using SM 5.4 solvent model

	Aromatic alcohols			Neutral				Anion				
No.	Compound	Q(O)	Q(H)	Q(OH)	E	E aq	Q(O-)	E	E aq	Δ(E)	Δ(E)aq	pKa
1	Phenol	-0.706	0.502	-0.204	-307.48	-307.488185	-0.788	-306.921	-307.017	-0.5593	-0.47118	9.98
2	2-Aminophenol	-0.713	0.522	-0.191	-362.838	-362.846914	-0.791	-362.269	-362.366	-0.56884	-0.48071	9.96
3	3-Aminophenol	-0.706	0.499	-0.207	-362.836	-362.850096	-0.789	-362.271	-362.372	-0.56472	-0.47798	9.96
4	4-Aminophenol	-0.716	0.499	-0.217	-362.832	-362.846152	-0.791	-362.269	-362.366	-0.56309	-0.47995	10.46
5	2-Bromophenol	-0.699	0.509	-0.19	-2880.76	-2880.76816	-0.749	-2880.21	-2880.3	-0.54877	-0.46735	8.45
6	3-Bromophenol	-0.7	0.505	-0.195	-2880.76	-2880.76878	-0.763	-2880.22	-2880.3	-0.54456	-0.46737	8.87
7	4-Bromophenol	-0.702	0.504	-0.198	-2880.76	-2880.76865	-0.774	-2880.21	-2880.3	-0.54656	-0.46925	9.35
8	2-tert-Butylphenol	-0.709	0.502	-0.207	-464.732	-464.736098	-0.794	-464.181	-464.264	-0.55121	-0.4718	10.62
9	3-tert-Butylphenol	-0.707	0.501	-0.206	-464.738	-464.744543	-0.786	-464.179	-464.269	-0.55924	-0.47602	10.12
10	4-tert-Butylphenol	-0.708	0.501	-0.207	-464.738	-464.744087	-0.786	-464.178	-464.267	-0.55941	-0.47752	10.23
11	2-Chlorophenol	-0.699	0.512	-0.187	-767.077	-767.083101	-0.754	-766.527	-766.616	-0.55035	-0.46758	8.53
12	3-Chlorophenol	-0.7	0.505	-0.195	-767.076	-767.084107	-0.767	-766.53	-766.617	-0.54602	-0.46712	8.88
13	4-Chlorophenol	-0.703	0.504	-0.199	-767.075	-767.083579	-0.778	-766.527	-766.614	-0.54824	-0.46934	9.38
14	3-Cyanophenol	-0.697	0.508	-0.189	-399.726	-399.735828	-0.764	-399.19	-399.273	-0.53656	-0.46314	8.57
15	4-Cyanophenol	-0.692	0.507	-0.185	-399.728	-399.73783	-0.737	-399.199	-399.281	-0.52901	-0.45718	7.95
16	2,3-Dichlorophenol	-0.695	0.513	-0.182	-1226.67	-1226.67413	-0.739	-1226.13	-1226.21	-0.54115	-0.46446	7.71
17	2,4-Dichlorophenol	-0.696	0.513	-0.183	-1226.67	-1226.67678	-0.746	-1226.13	-1226.21	-0.54016	-0.46493	7.89
18	2,5-Dichlorophenol	-0.693	0.514	-0.179	-1226.67	-1226.67747	-0.736	-1226.13	-1226.21	-0.53793	-0.46282	7.51
19	2,6-Dichlorophenol	-0.682	0.514	-0.168	-1226.67	-1226.67425	-0.722	-1226.13	-1226.21	-0.53773	-0.46122	6.81
20	3,4-Dichlorophenol	-0.698	0.506	-0.192	-1226.67	-1226.67447	-0.761	-1226.13	-1226.21	-0.53801	-0.46518	8.62
21	3,5-Dichlorophenol	-0.694	0.507	-0.187	-1226.67	-1226.67806	-0.75	-1226.14	-1226.22	-0.53423	-0.46276	8.18
22	2,5-Dimethylphenol	-0.709	0.506	-0.203	-386.117	-386.123915	-0.788	-385.558	-385.649	-0.55923	-0.47499	10.41
23	2,6-Dimethylphenol	-0.713	0.507	-0.206	-386.117	-386.122873	-0.793	-385.559	-385.649	-0.55798	-0.47362	10.6
24	3,4-Dimethylphenol	-0.709	0.501	-0.208	-386.115	-386.12334	-0.79	-385.554	-385.645	-0.56175	-0.47834	10.36
25	3,5-Dimethylphenol	-0.708	0.501	-0.207	-386.117	-386.124662	-0.785	-385.556	-385.648	-0.56128	-0.47617	10.19
26	2,4-Dinitrophenol	-0.665	0.535	-0.13	-716.505	-716.5055961	-0.625	-715.995	-716.06	-0.51025	-0.44587	4.12
27	2,5-Dinitrophenol	-0.67	0.534	-0.136	-716.502	-716.501656	-0.647	-715.983	-716.048	-0.51874	-0.45394	5.2
28	2,6-Dinitrophenol	-0.637	0.532	-0.105	-716.492	-716.489769	-0.59	-715.977	-716.043	-0.51509	-0.44642	3.73
29	2-Flourophenol	-0.7	0.518	-0.182	-406.721	-406.726984	-0.769	-406.167	-406.259	-0.55453	-0.46802	8.7
30	3-Flourophenol	-0.7	0.505	-0.195	-406.722	-406.72875	-0.774	-406.173	-406.263	-0.54916	-0.46549	9.29
31	4-Flourophenol	-0.706	0.503	-0.203	-406.72	-406.727055	-0.793	-406.16	-406.25	-0.56094	-0.47693	9.81
32	2-Methylphenol	-0.709	0.506	-0.203	-346.798	-346.805694	-0.79	-346.24	-346.333	-0.5582	-0.47247	10.32
33	3-Methylphenol	-0.707	0.501	-0.206	-346.799	-346.806419	-0.786	-346.238	-346.333	-0.56036	-0.4738	10.09
34	4-Methylphenol	-0.708	0.501	-0.207	-346.798	-346.805608	-0.791	-346.237	-346.33	-0.56131	-0.47588	10.27
35	2-Methoxyphenol	-0.709	0.518	-0.191	-422.009	-422.014573	-0.769	-421.44	-421.534	-0.56835	-0.48025	9.98
36	3-Methoxyphenol	-0.705	0.503	-0.202	-422.007	-422.017377	-0.783	-421.447	-421.543	-0.55988	-0.47467	9.65
37	4-Methoxyphenol	-0.711	0.501	-0.21	-422.004	-422.014505	-0.8	-421.441	-421.537	-0.56279	-0.47773	10.24
38	3-Methylsulfonylphenol	-0.698	0.508	-0.19	-895.376	-895.399089	-0.758	-894.84	-894.934	-0.53627	-0.46499	8.75
39	4-Methylsulfonylphenol	-0.694	0.507	-0.187	-895.377	-895.400375	-0.735	-894.848	-894.939	-0.52916	-0.46108	7.83
40	2-Nitrophenol	-0.682	0.531	-0.151	-511.999	-512.000311	-0.669	-511.457	-511.537	-0.54219	-0.46286	7.22
41	3-Nitrophenol	-0.697	0.509	-0.188	-511.991	-511.999656	-0.762	-511.457	-511.535	-0.53433	-0.46439	8.36
42	4-Nitrophenol	-0.687	0.509	-0.178	-511.993	-512.002567	-0.703	-511.475	-511.553	-0.51839	-0.4491	7.14
43	5-Hydroxyisoquinoline	-0.698	0.506	-0.192	-477.163	-477.176111	-0.75	-476.626	-476.715	-0.53666	-0.46161	8.47
44	6-Hydroxyisoquinoline	-0.698	0.505	-0.193	-477.166	-477.179735	-0.744	-476.629	-476.718	-0.53699	-0.46205	9.17
45	7-Hydroxyisoquinoline	-0.701	0.504	-0.197	-477.164	-477.178153	-0.757	-476.624	-476.713	-0.54035	-0.46552	8.9
46	8-Hydroxyisoquinoline	-0.696	0.507	-0.189	-477.163	-477.17629	-0.738	-476.628	-476.715	-0.53529	-0.46081	8.42
47	3-Hydroxyquinoline	-0.7	0.506	-0.194	-477.166	-477.178576	-0.757	-476.626	-476.714	-0.54011	-0.46469	8.08
48	5-Hydroxyquinoline	-0.699	0.503	-0.196	-477.164	-477.176178	-0.751	-476.626	-476.715	-0.53751	-0.46116	8.56
49	6-Hydroxyquinoline	-0.702	0.504	-0.198	-477.167	-477.179716	-0.756	-476.623	-476.712	-0.54365	-0.46807	8.9
50	7-Hydroxyquinoline	-0.701	0.507	-0.194	-477.169	-477.181665	-0.748	-476.623	-476.713	-0.54547	-0.46901	8.87
51	8-Hydroxyquinoline	-0.7	0.525	-0.175	-477.176	-477.18387	-0.714	-476.613	-476.701	-0.56239	-0.48283	9.71
52	2,4,5-Trichlorophenol	-0.691	0.515	-0.176	-1686.26	-1686.26623	-0.73	-1685.73	-1685.81	-0.53052	-0.46026	7.07
53	2,3,4-Trichlorophenol	-0.693	0.514	-0.179	-1686.26	-1686.26289	-0.733	-1685.72	-1685.8	-0.53347	-0.46195	7.1
54	2,3,4,6-Tetrachlorophenol	-0.678	0.516	-0.162	-2145.85	-2145.85156	-0.704	-2145.32	-2145.4	-0.52205	-0.45473	5.62
55	Pentachlorophenol	-0.675	0.516	-0.159	-2605.43	-2605.43466	-0.694	-2604.91	-2604.98	-0.51621	-0.45136	4.9
56	2,4,6-Trimethylphenol	-0.715	0.507	-0.208	-425.435	-425.440298	-0.796	-424.875	-424.962	-0.55988	-0.47782	10.89
57	2,4,6-Trinitrophenol	-0.623	0.535	-0.088	-920.993	-920.984762	-0.559	-920.505	-920.557	-0.48795	-0.42791	0.37
58	2-Hydroxypyridine	-0.704	0.513	-0.191	-323.532	-323.542326	-0.753	-322.972	-323.075	-0.55957	-0.46753	11.99
59	2-Hydroxypyrazine	-0.696	0.515	-0.181	-339.563	-339.577433	-0.743	-339.018	-339.122	-0.54552	-0.45543	7.28
60	4-Amino-8-hydroxyquinoline	-0.703	0.527	-0.176	-532.537	-532.551039	-0.712	-531.97	-532.063	-0.56656	-0.48816	10.71
61	5-Amino-8-hydroxyquinoline	-0.708	0.524	-0.184	-532.528	-532.542117	-0.715	-531.96	-532.05	-0.56798	-0.49162	11.24
62	8-Hydroxy-2-methylquinoline	-0.703	0.524	-0.179	-516.497	-516.504512	-0.71	-515.928	-516.014	-0.5688	-0.49036	10.16
63	8-Hydroxy-4-methylquinoline	-0.701	0.526	-0.175	-516.495	-516.50338	-0.71	-515.931	-516.018	-0.56408	-0.48507	9.99
64	4-Chloro-3,5-dimethylphenol	-0.704	0.503	-0.201	-845.713	-845.720486	-0.778	-845.161	-845.246	-0.55185	-0.47412	9.7
65	Dichloroxylenol	-0.699	0.512	-0.187	-1305.31	-1305.31188	-0.747	-1304.76	-1304.84	-0.54539	-0.47024	8.28
66	4-Hydroxybenzothiazole	-0.696	0.522	-0.174	-797.939	-797.950859	-0.729	-797.384	-797.476	-0.55445	-0.47506	8.85

Table 2.9: Summation of regression results using B3LYP/6-31+G* for the aliphatic alcohols in vacuum and in SM5.4 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-12.3	2.7	-98.7	10.1	0.779	0.918	95	29
$Q_n(O^-)$	9.5	3.0	-5.1	3.2	0.082	1.872	2	29
ΔE	-30.4	8.2	-75.7	13.9	0.523	1.349	29	29
ΔE_{aq}	-32.5	15.9	-95.9	32.5	0.243	1.701	8	29
$Q_n(O)$	-64.0	14.1	-102.3	18.4	0.532	1.336	30	29
$Q_n(H)$	95.7	17.6	-164.2	35.5	0.441	1.460	21	29

Table 2.10: Summation of regression results using B3LYP/6-31+G* for the aromatic alcohols in vacuum and in SM5.4 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-5.0	0.9	-73.4	5.2	0.753	0.975	194	66
$Q_n(O^-)$	-16.7	2.1	-34.0	2.8	0.683	1.104	138	66
ΔE	-50.2	3.3	-107.8	6.0	0.832	0.803	317	66
ΔE_{aq}	-64.3	4.8	-156.0	10.3	0.779	0.922	225	66
$Q_n(O)$	-73.1	3.7	-117.3	5.3	0.883	0.670	484	66
$Q_n(H)$	65.6	10.7	-111.4	21.1	0.303	1.638	27	66

Table 2.11: Summation of regression results using B3LYP/6-31+G* for aliphatic and aromatic alcohols in vacuum and in SM5.4 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-4.4	0.5	-70.1	2.3	0.905	0.993	895	95
$Q_n(O^-)$	-8.6	1.4	-23.6	1.7	0.660	1.886	181	95
ΔE	-53.5	2.3	-114.2	4.2	0.886	1.092	724	95
ΔE_{aq}	-83.7	5.2	-198.5	11.0	0.775	1.536	320	95
$Q_n(O)$	-52.8	2.1	-88.0	2.9	0.902	1.009	864	95
$Q_n(H)$	119.5	9.7	-215.6	19.2	0.573	2.114	125	95

Table 2.12: Data for Aliphatic Alcohols using B3LYP/6-31+G* using SM8 solvent model

	Aliphatic alcohols		Neutral			Anion			
No.	Compound Name	Q(O)	Q(H)	Q(O+H)	E aq(SM8)	Q(O^-)	E aq	ΔEaq	pKa
1	Methanol	-0.806	0.506	-0.3	-115.7330905	-1.035	-115.250959	-0.4821315	15.51
2	Ethanol	-0.802	0.501	-0.301	-155.052894	-1.013	-154.566398	-0.486496	15.5
3	2-Bromoethanol	-0.783	0.509	-0.274	-2728.33547	-0.982	-2727.85222	-0.48325	14.38
4	2-Ethoxyethanol	-0.792	0.507	-0.285	-308.890981	-1.026	-308.396288	-0.494693	14.98
5	2-methoxyethanol	-0.792	0.507	-0.285	-269.571114	-1.024	-269.079214	-0.4919	14.87
6	2,2-Dibromoethanol	-0.774	0.509	-0.265	-5301.61103	-0.955	-5301.12154	-0.48949	13.29
7	2,2-Dichloroethanol	-0.778	0.508	-0.27	-1074.23917	-0.964	-1073.75489	-0.48428	12.89
8	2,2,2 tribromoethanol	-0.761	0.508	-0.253	-7874.87581	-0.912	-7874.39463	-0.48118	12.7
9	2,2,2-trichloroethanol	-0.765	0.513	-0.252	-1533.82163	-0.989	-1533.35114	-0.47049	12.24
10	2,2,2-trifloroethanol	-0.78	0.524	-0.256	-452.793109	-0.887	-452.284267	-0.508842	12.37
11	Cynoethanol	-0.789	0.508	-0.281	-247.298766	-0.982	-246.812188	-0.486578	14.03
12	ethylene glycol	-0.802	0.51	-0.292	-230.273546	-0.973	-229.784859	-0.488687	15.1
13	1-Propanol	-0.801	0.502	-0.299	-194.366433	-1.012	-193.875663	-0.49077	16.1
14	2-Propanol	-0.804	0.503	-0.301	-194.371928	-1.012	-193.882149	-0.489779	17.1
15	2,2,3,3-tetraflouropropanol	-0.781	0.522	-0.259	-591.336961	-0.91	-590.819296	-0.517665	12.74
16	3,3-diflouropropanol	-0.799	0.514	-0.285	-392.857893	-1.028	-392.361743	-0.49615	12.74
17	3,3,3-triflouro-2-methylprop-2-nol	-0.774	0.511	-0.263	-531.432046	-0.881	-530.924524	-0.507522	11.6
18	1-Butanol	-0.798	0.502	-0.296	-233.681426	-1.035	-233.189792	-0.491634	16.1
19	2-Butanol	-0.803	0.503	-0.3	-233.685255	-1.006	-233.191746	-0.493509	17.6
20	1,2,3,4-Butanetetrol	-0.823	0.521	-0.302	-459.351091	-0.94	-458.855864	-0.495227	13.9
21	Allyl Alcohol	-0.801	0.509	-0.292	-193.130537	-1.03	-192.644311	-0.486226	15.5
22	Benzyl Alcohol	-0.789	0.503	-0.286	-346.796963	-1.014	-346.311726	-0.485237	15.4
23	Chloral Hydrate	-0.777	0.524	-0.253	-1609.04682	-0.961	-1608.57621	-0.47061	11
24	Propargyl alcohol	-0.778	0.51	-0.268	-191.875337	-1.021	-191.401479	-0.473858	13.6
25	2,2,2-triflouro-1-(p-tolyl)ethanol	-0.762	0.512	-0.25	-723.17218	-0.887	-722.671286	-0.500894	12.04
26	2,2,2-triflouro-1-(4-methoxy phenyl) ethanol	-0.767	0.519	-0.248	-798.382687	-0.888	-797.875777	-0.50691	12.24
27	Glycerol	-0.812	0.521	-0.291	-344.811979	-0.947	-344.322356	-0.489623	14.15
28	2-methyl-2-propanol	-0.801	0.498	-0.303	-233.688801	-1.018	-233.195029	-0.493772	19.2
29	2-methyl-2-propen-1-nol	-0.802	0.511	-0.291	-232.449679	-1.024	-231.959084	-0.490595	14.82

Table 2.13: Data for Aromatic Alcohols using B3LYP/6-31+G* using SM8 solvent model

	Aromatic alcohols		Neutral			Anion			
No.	Compound Name	Q(O)	Q(H)	Q(O+H)	E_{aq}(SM8)	$Q(O^-)$	E aq	ΔEaq	pKa
1	Phenol	-0.732	0.52	-0.212	-307.489386	-0.859	-307.027294	-0.462092	9.98
2	2-Aminophenol	-0.748	0.526	-0.222	-362.850584	-0.903	-362.38227	-0.468314	9.96
3	3-Aminophenol	-0.731	0.517	-0.214	-362.850865	-0.92	-362.392597	-0.458268	9.96
4	4-Aminophenol	-0.738	0.518	-0.22	-362.848811	-0.939	-362.395698	-0.453113	10.46
5	2-Bromophenol	-0.722	0.515	-0.207	-2880.76987	-0.853	-2880.31597	-0.4539	8.45
6	3-Bromophenol	-0.723	0.524	-0.199	-2880.76999	-0.894	-2880.32011	-0.44988	8.87
7	4-Bromophenol	-0.714	0.521	-0.193	-2880.76829	-0.878	-2880.31383	-0.45446	9.35
8	2-tert-Butylphenol	-0.733	0.508	-0.225	-464.73701	-0.879	-464.278258	-0.458752	10.62
9	3-tert-Butylphenol	-0.732	0.518	-0.214	-464.744457	-0.92	-464.284858	-0.459599	10.12
10	4-tert-Butylphenol	-0.732	0.52	-0.212	-464.744088	-0.872	-464.274136	-0.469952	10.23
11	2-Chlorophenol	-0.716	0.516	-0.2	-767.082993	-0.872	-766.633179	-0.449814	8.53
12	3-Chlorophenol	-0.721	0.524	-0.197	-767.084621	-0.893	-766.634191	-0.45043	8.88
13	4-Chlorophenol	-0.725	0.523	-0.202	-767.084258	-0.891	-766.608876	-0.475382	9.38
14	3-Cyanophenol	-0.717	0.526	-0.191	-399.736787	-0.883	-399.284218	-0.452569	8.57
15	4-Cyanophenol	-0.706	0.529	-0.177	-399.739829	-0.794	-399.284342	-0.455487	7.95
16	2,3-Dichlorophenol	-0.716	0.522	-0.194	-1226.67482	-0.858	-1226.22781	-0.44701	7.71
17	2,4-Dichlorophenol	-0.717	0.524	-0.193	-1226.67756	-0.864	-1226.2295	-0.44806	7.89
18	2,5-Dichlorophenol	-0.713	0.525	-0.188	-1226.67829	-0.857	-1226.23324	-0.44505	7.51
19	2,6-Dichlorophenol	-0.706	0.523	-0.183	-1226.67535	-0.833	-1226.2312	-0.44415	6.81
20	3,4-Dichlorophenol	-0.715	0.526	-0.189	-1226.67456	-0.883	-1226.22555	-0.44901	8.62
21	3,5-Dichlorophenol	-0.713	0.528	-0.185	-1226.67812	-0.877	-1226.23363	-0.44449	8.18
22	2,5-Dimethylphenol	-0.733	0.521	-0.212	-386.125814	-0.909	-385.668891	-0.456923	10.41
23	2,6-Dimethylphenol	-0.734	0.521	-0.213	-386.125099	-0.895	-385.666355	-0.458744	10.6
24	3,4-Dimethylphenol	-0.709	0.501	-0.208	-386.115826	-0.926	-385.66518	-0.450646	10.36
25	3,5-Dimethylphenol	-0.708	0.501	-0.207	-386.117219	-0.918	-385.669645	-0.447574	10.19
26	2,4-Dinitrophenol	-0.669	0.539	-0.13	-716.515984	-0.68	-716.076248	-0.439736	4.12
27	2,5-Dinitrophenol	-0.686	0.536	-0.15	-716.511103	-0.714	-716.066528	-0.444575	5.2
28	2,6-Dinitrophenol	-0.658	0.534	-0.124	-716.50593	-0.66	-716.067311	-0.438619	3.73
29	2-Flourophenol	-0.716	0.522	-0.194	-406.726117	-0.889	-406.276699	-0.449418	8.7
30	3-Flourophenol	-0.721	0.524	-0.197	-406.730743	-0.886	-406.275947	-0.454796	9.29
31	4-Flourophenol	-0.727	0.523	-0.204	-406.729586	-0.861	-406.262667	-0.466919	9.81
32	2-Methylphenol	-0.733	0.521	-0.212	-346.807214	-0.904	-346.350383	-0.456831	10.32
33	3-Methylphenol	-0.731	0.519	-0.212	-346.807599	-0.889	-346.3471	-0.460499	10.09
34	4-Methylphenol	-0.733	0.52	-0.213	-346.806789	-0.786	-346.232173	-0.574616	10.27
35	2-Methoxyphenol	-0.739	0.525	-0.214	-422.018577	-0.905	-421.556884	-0.461693	9.98
36	3-Methoxyphenol	-0.728	0.521	-0.207	-422.018338	-0.913	-421.561353	-0.456985	9.65
37	4-Methoxyphenol	-0.734	0.52	-0.214	-422.015552	-0.875	-421.54344	-0.472112	10.24
38	3-Methylsulfonylphenol	-0.717	0.526	-0.191	-895.400424	-0.88	-894.947377	-0.453047	8.75
39	4-Methylsulfonylphenol	-0.709	0.528	-0.181	-895.401945	-0.776	-894.901873	-0.500072	7.83
40	2-Nitrophenol	-0.703	0.531	-0.172	-512.008498	-0.859	-511.544381	-0.464117	7.22
41	3-Nitrophenol	-0.714	0.528	-0.186	-512.004408	-0.858	-511.548307	-0.456101	8.36
42	4-Nitrophenol	-0.69	0.534	-0.156	-512.008802	-0.862	-511.550096	-0.458706	7.14
43	5-Hydroxyisoquinoline	-0.72	0.52	-0.2	-477.176681	-0.813	-476.721269	-0.455412	8.47
44	6-Hydroxyisoquinoline	-0.714	0.526	-0.188	-477.181559	-0.809	-476.722728	-0.458831	9.17
45	7-Hydroxyisoquinoline	-0.718	0.525	-0.193	-477.180138	-0.828	-476.71895	-0.461188	8.9
46	8-Hydroxyisoquinoline	-0.717	0.52	-0.197	-477.176599	-0.803	-476.722848	-0.453751	8.42
47	3-Hydroxyquinoline	-0.717	0.528	-0.189	-477.180575	-0.823	-476.723264	-0.457311	8.08
48	5-Hydroxyquinoline	-0.718	0.519	-0.199	-477.178084	-0.809	-476.724617	-0.453467	8.56
49	6-Hydroxyquinoline	-0.72	0.525	-0.195	-477.18219	-0.825	-476.720307	-0.461883	8.9
50	7-Hydroxyquinoline	-0.719	0.526	-0.193	-477.183763	-0.82	-476.723151	-0.460612	8.87
51	8-Hydroxyquinoline	-0.734	0.528	-0.206	-477.188086	-0.807	-476.719632	-0.468454	9.71
52	2,4,5-Trichlorophenol	-0.708	0.527	-0.181	-1686.26672	-0.846	-1685.82329	-0.44343	7.07
53	2,3,4-Trichlorophenol	-0.711	0.524	-0.187	-1686.26299	-0.849	-1685.81836	-0.44463	7.1
54	2,3,4,6-Tetrachlorophenol	-0.695	0.526	-0.169	-2145.85147	-0.814	-2145.41554	-0.43593	5.62
55	Pentachlorophenol	-0.69	0.528	-0.162	-2605.43414	-0.802	-2605.00189	-0.43225	4.9
56	2,4,6-Trimethylphenol	-0.735	0.521	-0.214	-425.442841	-0.906	-424.98283	-0.460011	10.89
57	2,4,6-Trinitrophenol	-0.63	0.542	-0.088	-921.004173	-0.614	-920.576824	-0.427349	0.37
58	2-Hydroxypyridine	-0.73	0.525	-0.205	-323.544903	-0.87	-323.091069	-0.453834	11.99
59	2-Hydroxypyrazine	-0.711	0.532	-0.179	-339.578293	-0.832	-339.131496	-0.446797	7.28
60	4-Amino-8-hydroxyquinoline	-0.74	0.528	-0.212	-532.556213	-0.85	-532.090881	-0.465332	10.71
61	5-Amino-8-hydroxyquinoline	-0.743	0.526	-0.217	-532.546566	-0.809	-532.067711	-0.478855	11.24
62	8-Hydroxy-2-methylquinoline	-0.736	0.527	-0.209	-516.509528	-0.814	-516.039441	-0.470087	10.16
63	8-Hydroxy-4-methylquinoline	-0.735	0.528	-0.207	-516.507696	-0.801	-516.036589	-0.471107	9.99
64	4-Chloro-3,5-dimethylphenol	-0.727	0.522	-0.205	-845.721941	-0.91	-845.267032	-0.454909	9.7
65	Dichloroxylenol	-0.721	0.522	-0.199	-1305.31395	-0.873	-1304.86399	-0.44996	8.28
66	4-Hydroxybenzothiazole	-0.726	0.527	-0.199	-797.951934	-0.816	-797.491285	-0.460649	8.85

Table 2.14: Summation of regression results using single point B3LYP/6-31+G* for the aliphatic alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-9.9	2.8	-86.6	10.3	0.721	1.031	69	29
$Q_n(O^-)$	-10.4	5.1	-25.2	5.2	0.456	1.440	22	29
ΔE_{aq}	21.0	16.7	13.8	34.2	0.006	1.948	0.16	29
$Q_n(O)$	-49.2	13.6	-80.4	17.2	0.446	1.453	21	29
$Q_n(H)$	115.9	17.5	-199.2	34.4	0.553	1.305	33	29

Table 2.15: Summation of regression results using single point B3LYP/6-31+G* for the aromatic alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-6.3	0.6	-77.6	3.3	0.894	0.638	541	66
$Q_n(O^-)$	-11.1	2.2	-23.4	2.6	0.546	1.321	77	66
ΔE_{aq}	-13.3	5.3	-48.2	11.6	0.210	1.743	17	66
$Q_n(O)$	-57.6	3.2	-92.4	4.5	0.863	0.723	406	66
$Q_n(H)$	101.3	14.6	-176.9	28.0	0.383	1.540	39	66

Table 2.16: Summation of regression results using single point B3LYP/6-31+G* for the aliphatic and aromatic alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-4.7	0.4	-68.7	1.8	0.936	0.819	1358	95
$Q_n(O^-)$	-19.5	1.7	-33.7	1.93	0.765	1.569	303	95
ΔE_{aq}	-36.2	4.9	-99.8	10.6	0.485	2.324	87	95
$Q_n(O)$	-49.2	2.0	-80.6	2.7	0.902	1.011	863	95
$Q_n(H)$	159.4	10.3	-286.8	19.8	0.690	1.801	207	95

Table 2.17: Optimization Data for Aliphatic Alcohols using B3LYP/6-31+G* using SM8 solvent model

	Aliphatic alcohols		Neutral			Anion			
No.	Compound	Q(O)	Q(H)	Q(O+H)	E aq	$Q(O^-)$	E aq	ΔEaq	pKa
1	Methanol	-0.808	0.507	-0.301	-115.73323	-1.082	-115.257152	-0.476077	15.51
2	Ethanol	-0.804	0.502	-0.302	-155.05301	-1.071	-154.574424	-0.478588	15.5
3	2-Bromoethanol	-0.784	0.51	-0.274	-2728.3356	-1.045	-2727.8612	-0.47439	14.38
4	2-Ethoxyethanol	-0.793	0.507	-0.286	-308.89114	-1.074	-308.412925	-0.478216	14.98
5	2-methoxyethanol	-0.793	0.507	-0.286	-269.57126	-1.073	-269.092482	-0.478778	14.87
6	2,2-Dibromoethanol	-0.775	0.509	-0.266	-5301.6111	-1.014	-5301.13658	-0.47454	13.29
7	2,2-Dichloroethanol	-0.778	0.51	-0.268	-1074.2395	-1.027	-1073.76492	-0.47461	12.89
8	2,2,2 tribromoethanol	-0.762	0.508	-0.254	-7874.876	-0.972	-7874.40566	-0.47033	12.7
9	2,2,2-trichloroethanol	-0.75	0.503	-0.247	-1533.823	-1.005	-1533.35415	-0.46888	12.24
10	2,2,2-trifloroethanol	-0.78	0.525	-0.255	-452.79339	-1.037	-452.325271	-0.46812	12.37
11	Cynoethanol	-0.791	0.51	-0.281	-247.29914	-1.037	-246.822465	-0.476675	14.03
12	ethylene glycol	-0.803	0.51	-0.293	-230.27355	-1.033	-229.796527	-0.477021	15.1
13	1-Propanol	-0.804	0.502	-0.302	-194.36655	-1.073	-193.885189	-0.481359	16.1
14	2-Propanol	-0.806	0.503	-0.303	-194.37206	-1.071	-193.892065	-0.479992	17.1
15	2,2,3,3-tetraflouropropanol	-0.781	0.523	-0.258	-591.33718	-1.048	-590.865014	-0.47217	12.74
16	3,3-diflouropropanol	-0.8	0.514	-0.286	-392.85809	-1.079	-392.375204	-0.482882	12.74
17	3,3,3-triflouro-2-methylprop-2-nol	-0.776	0.52	-0.256	-531.43228	-1.023	-530.957686	-0.474596	11.6
18	1-Butanol	-0.8	0.502	-0.298	-233.68152	-1.088	-233.201541	-0.479979	16.1
19	2-Butanol	-0.805	0.503	-0.302	-233.68538	-1.067	-233.20271	-0.482667	17.6
20	1,2,3,4-Butanetetrol	-0.823	0.519	-0.304	-459.35176	-0.988	-458.874598	-0.477163	13.9
21	Allyl Alcohol	-0.803	0.51	-0.293	-193.13065	-1.065	-192.654596	-0.476056	15.5
22	Benzyl Alcohol	-0.792	0.503	-0.289	-346.79727	-1.062	-346.318365	-0.478901	15.4
23	Chloral Hydrate	-0.776	0.525	-0.251	-1609.0474	-0.929	-1608.59002	-0.45741	11
24	Propargyl alcohol	-0.779	0.51	-0.269	-191.87549	-1.055	-191.405694	-0.469798	13.6
25	2,2,2-triflouro-1-(p-tolyl) ethanol	-0.767	0.516	-0.251	-723.17256	-1.013	-722.704155	-0.468408	12.04
26	2,2,2-triflouro-1-(4-methoxy phenyl) ethanol	-0.768	0.522	-0.246	-798.38322	-1.018	-797.914211	-0.469011	12.24
27	Glycerol	-0.812	0.52	-0.292	-344.81232	-1.008	-344.33918	-0.473139	14.15
28	2-methyl-2-propanol	-0.803	0.498	-0.305	-233.68892	-1.069	-233.204573	-0.48435	19.2
29	2-methyl-2-propen-1-nol	-0.803	0.511	-0.292	-232.44985	-1.067	-231.972171	-0.477678	14.82

Table 2.18: Optimization Data for Aromatic Alcohols using B3LYP/6-31+G* using SM8 solvent model

	Aromatic alcohols		Neutral			Anion			
No.	Compound	Q(O)	Q(H)	Q(O+H)	E aq	$Q(O^-)$	E aq	ΔEaq	pKa
1	Phenol	-0.732	0.52	-0.212	-307.489386	-0.91	-307.033007	-0.456379	9.98
2	2-Aminophenol	-0.748	0.526	-0.222	-362.850584	-0.935	-362.397673	-0.452911	9.96
3	3-Aminophenol	-0.73	0.518	-0.212	-362.852076	-0.912	-362.393457	-0.458619	9.96
4	4-Aminophenol	-0.739	0.518	-0.221	-362.848825	-0.941	-362.385698	-0.463127	10.46
5	2-Bromophenol	-0.722	0.515	-0.207	-2880.76987	-0.853	-2880.31607	-0.4538	8.45
6	3-Bromophenol	-0.723	0.524	-0.199	-2880.76998	-0.893	-2880.31997	-0.45001	8.87
7	4-Bromophenol	-0.724	0.523	-0.201	-2880.76949	-0.9	-2880.31728	-0.45221	9.35
8	2-tert-Butylphenol	-0.733	0.508	-0.225	-464.736839	-0.88	-464.278474	-0.458365	10.62
9	3-tert-Butylphenol	-0.732	0.518	-0.214	-464.74447	-0.92	-464.284671	-0.459799	10.12
10	4-tert-Butylphenol	-0.732	0.52	-0.212	-464.744091	-0.929	-464.284164	-0.459927	10.23
11	2-Chlorophenol	-0.723	0.521	-0.202	-767.08449	-0.872	-766.633168	-0.451322	8.53
12	3-Chlorophenol	-0.721	0.524	-0.197	-767.084608	-0.892	-766.634058	-0.45055	8.88
13	4-Chlorophenol	-0.725	0.523	-0.202	-767.084246	-0.9	-766.630728	-0.453518	9.38
14	3-Cyanophenol	-0.717	0.526	-0.191	-399.736802	-0.882	-399.28409	-0.452712	8.57
15	4-Cyanophenol	-0.706	0.529	-0.177	-399.739773	-0.832	-399.288707	-0.451066	7.95
16	2,3-Dichlorophenol	-0.716	0.522	-0.194	-1226.67484	-0.858	-1226.22763	-0.44721	7.71
17	2,4-Dichlorophenol	-0.717	0.524	-0.193	-1226.67756	-0.863	-1226.22951	-0.44805	7.89
18	2,5-Dichlorophenol	-0.713	0.525	-0.188	-1226.67829	-0.857	-1226.23331	-0.44498	7.51
19	2,6-Dichlorophenol	-0.706	0.523	-0.183	-1226.67537	-0.833	-1226.23118	-0.44419	6.81
20	3,4-Dichlorophenol	-0.715	0.526	-0.189	-1226.67456	-0.881	-1226.22542	-0.44914	8.62
21	3,5-Dichlorophenol	-0.713	0.528	-0.185	-1226.67812	-0.876	-1226.23336	-0.44476	8.18
22	2,5-Dimethylphenol	-0.733	0.521	-0.212	-386.125814	-0.909	-385.668085	-0.457729	10.41
23	2,6-Dimethylphenol	-0.734	0.521	-0.213	-386.125097	-0.9	-385.668057	-0.45704	10.6
24	3,4-Dimethylphenol	-0.709	0.501	-0.208	-386.115826	-0.925	-385.665041	-0.450785	10.36
25	3,5-Dimethylphenol	-0.708	0.501	-0.207	-386.117214	-0.918	-385.668819	-0.448395	10.19
26	2,4-Dinitrophenol	-0.669	0.539	-0.13	-716.515984	-0.742	-716.072264	-0.44372	4.12
27	2,5-Dinitrophenol	-0.686	0.536	-0.15	-716.511104	-0.828	-716.061986	-0.449118	5.2
28	2,6-Dinitrophenol	-0.669	0.536	-0.133	-716.506822	-0.83	-716.060316	-0.446506	3.73
29	2-Flourophenol	-0.725	0.527	-0.198	-406.728349	-0.898	-406.276445	-0.451904	8.7
30	3-Flourophenol	-0.721	0.524	-0.197	-406.730749	-0.886	-406.275908	-0.454841	9.29
31	4-Flourophenol	-0.727	0.523	-0.204	-406.729585	-0.91	-406.270163	-0.459422	9.81
32	2-Methylphenol	-0.733	0.521	-0.212	-346.807219	-0.905	-346.350616	-0.456603	10.32
33	3-Methylphenol	-0.731	0.519	-0.212	-346.807599	-0.917	-346.351479	-0.45612	10.09
34	4-Methylphenol	-0.733	0.52	-0.213	-346.806796	-0.791	-346.236597	-0.570199	10.27
35	2-Methoxyphenol	-0.739	0.525	-0.214	-422.018618	-0.898	-421.549154	-0.469464	9.98
36	3-Methoxyphenol	-0.728	0.521	-0.207	-422.018346	-0.912	-421.561175	-0.457171	9.65
37	4-Methoxyphenol	-0.734	0.52	-0.214	-422.015546	-0.93	-421.552783	-0.462763	10.24
38	3-Methylsulfonylphenol	-0.717	0.526	-0.191	-895.400144	-0.877	-894.946669	-0.453475	8.75
39	4-Methylsulfonylphenol	-0.709	0.528	-0.181	-895.401362	-0.848	-894.94982	-0.451542	7.83
40	2-Nitrophenol	-0.703	0.531	-0.172	-512.008355	-0.869	-511.549715	-0.45864	7.22
41	3-Nitrophenol	-0.714	0.528	-0.186	-512.004396	-0.874	-511.552095	-0.452301	8.36
42	4-Nitrophenol	-0.69	0.534	-0.156	-512.008784	-0.756	-511.565072	-0.443712	7.14
43	5-Hydroxyisoquinoline	-0.72	0.52	-0.2	-477.176679	-0.864	-476.727352	-0.449327	8.47
44	6-Hydroxyisoquinoline	-0.714	0.526	-0.188	-477.181528	-0.809	-476.722728	-0.4588	9.17
45	7-Hydroxyisoquinoline	-0.718	0.525	-0.193	-477.180106	-0.88	-476.725503	-0.454603	8.9
46	8-Hydroxyisoquinoline	-0.717	0.52	-0.197	-477.176566	-0.852	-476.729175	-0.447391	8.42
47	3-Hydroxyquinoline	-0.717	0.528	-0.189	-477.18058	-0.873	-476.729289	-0.451291	8.08
48	5-Hydroxyquinoline	-0.719	0.519	-0.2	-477.178049	-0.859	-476.730225	-0.447824	8.56
49	6-Hydroxyquinoline	-0.72	0.525	-0.195	-477.182165	-0.877	-476.726695	-0.45547	8.9
50	7-Hydroxyquinoline	-0.719	0.526	-0.193	-477.183763	-0.873	-476.729453	-0.45431	8.87
51	8-Hydroxyquinoline	-0.734	0.528	-0.206	-477.188064	-0.866	-476.726427	-0.461637	9.71
52	2,4,5-Trichlorophenol	-0.708	0.527	-0.181	-1686.26672	-0.846	-1685.82331	-0.44341	7.07
53	2,3,4-Trichlorophenol	-0.711	0.524	-0.187	-1686.26299	-0.849	-1685.81834	-0.44465	7.1
54	2,3,4,6-Tetrachlorophenol	-0.695	0.526	-0.169	-2145.85147	-0.815	-2145.41551	-0.43596	5.62
55	Pentachlorophenol	-0.69	0.528	-0.162	-2605.434412	-0.801	-2605.00193	-0.432482	4.9
56	2,4,6-Trimethylphenol	-0.735	0.521	-0.214	-425.442841	-0.905	-424.982617	-0.460224	10.89
57	2,4,6-Trinitrophenol	-0.63	0.542	-0.088	-921.004174	-0.724	-920.573705	-0.430469	0.37
58	2-Hydroxypyridine	-0.73	0.525	-0.205	-323.544923	-0.871	-323.0913	-0.453623	11.99
59	2-Hydroxypyrazine	-0.711	0.532	-0.179	-339.578347	-0.833	-339.131688	-0.446659	7.28
60	4-Amino-8-hydroxyquinoline	-0.739	0.528	-0.211	-532.556336	-0.866	-532.09119	-0.465146	10.71
61	5-Amino-8-hydroxyquinoline	-0.743	0.526	-0.217	-532.546526	-0.886	-532.076842	-0.469684	11.24
62	8-Hydroxy-2-methylquinoline	-0.736	0.527	-0.209	-516.509528	-0.876	-516.048858	-0.46067	10.16
63	8-Hydroxy-4-methylquinoline	-0.734	0.528	-0.206	-516.507591	-0.857	-516.043438	-0.464153	9.99
64	4-Chloro-3,5-dimethylphenol	-0.727	0.522	-0.205	-845.721897	-0.91	-845.266836	-0.455061	9.7
65	Dichloroxylenol	-0.721	0.522	-0.199	-1305.31395	-0.873	-1304.86378	-0.45017	8.28
66	4-Hydroxybenzothiazole	-0.726	0.527	-0.199	-797.951927	-0.873	-797.498308	-0.453619	8.85

Table 2.19: Summation of regression results using B3LYP/6-31+G* for the optimized aliphatic alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-8.4	2.6	-81.2	9.2	0.739	0.998	76	29
$Q_n(O^-)$	-21.9	7.6	-34.7	7.3	0.455	1.442	22	29
ΔE_{aq}	-111.5	19.1	-264.7	40.2	0.615	1.212	43	29
$Q_n(O)$	-46.4	12.4	-76.8	15.7	0.467	1.425	23	29
$Q_n(H)$	110.4	16.1	-188.3	31.6	0.567	1.285	35	29

Table 2.20: Summation of regression results using B3LYP/6-31+G* for the optimized aromatic alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-6.6	0.6	-79.0	3.3	0.895	0.634	548	66
$Q_n(O^-)$	-20.8	3.1	-33.8	3.6	0.577	1.275	87	66
ΔE_{aq}	-0.40	0.04	0.03	0.01	0.006	0.016	0.45	66
$Q_n(O)$	-59.4	3.3	-94.8	4.6	0.866	0.717	415	66
$Q_n(H)$	103.5	14.5	-181.0	27.8	0.398	1.521	42	66

Table 2.21: Summation of regression results using B3LYP/6-31+G* for the optimized alcohols in SM8 medium

	Intercept	Intercept Error	X-variable	X-variable error	r^2	Standard error	Fisher	Observations
$Q_n(OH)$	-4.7	0.4	-68.8	1.8	0.938	0.802	1422	95
$Q_n(O^-)$	-0.50	0.01	1.80	0.06	0.892	0.029	769	95
ΔE_{aq}	-49.8	6.6	-130.6	14.4	0.467	2.362	81	95
$Q_n(O)$	-49.2	2.0	-80.4	2.7	0.903	1.003	875	95
$Q_n(H)$	158.9	10.3	-285.6	19.9	0.688	1.806	205	95

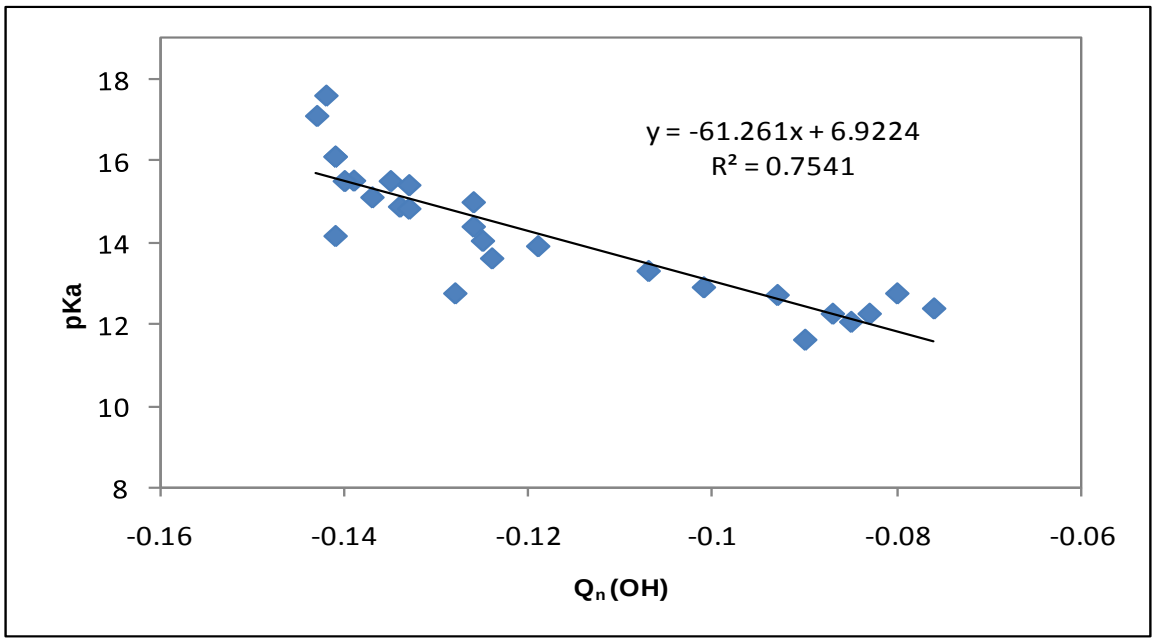


Figure 2.1: Best regression plot using RM1 for aliphatic alcohols (Eq. 2.1)

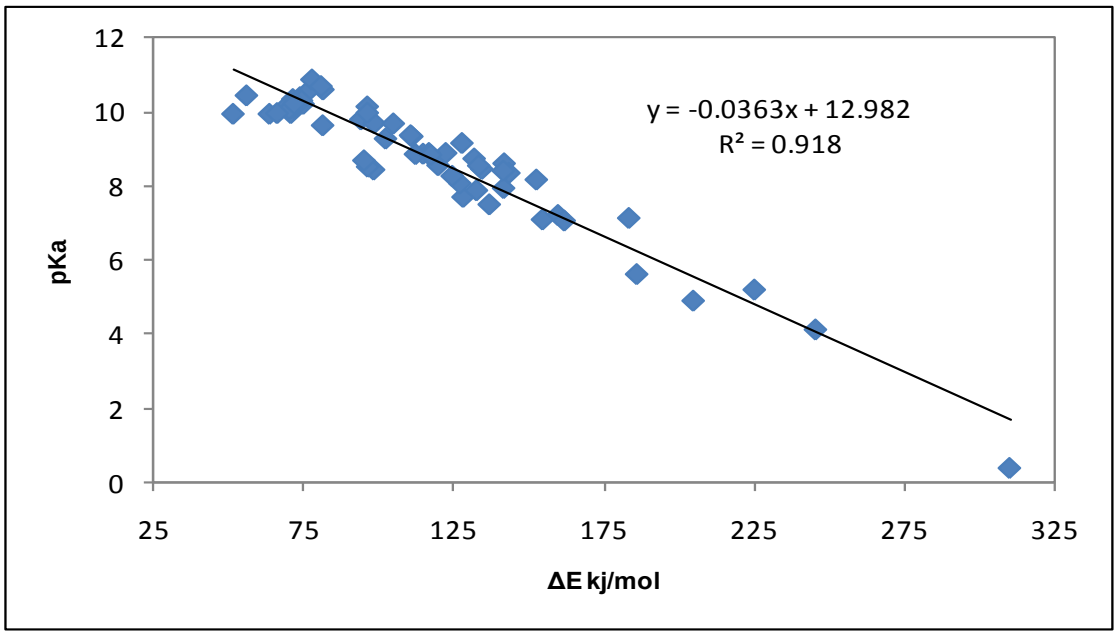


Figure 2.2: Best regression plot using RM1 for aromatic alcohols (Eq.2.2)

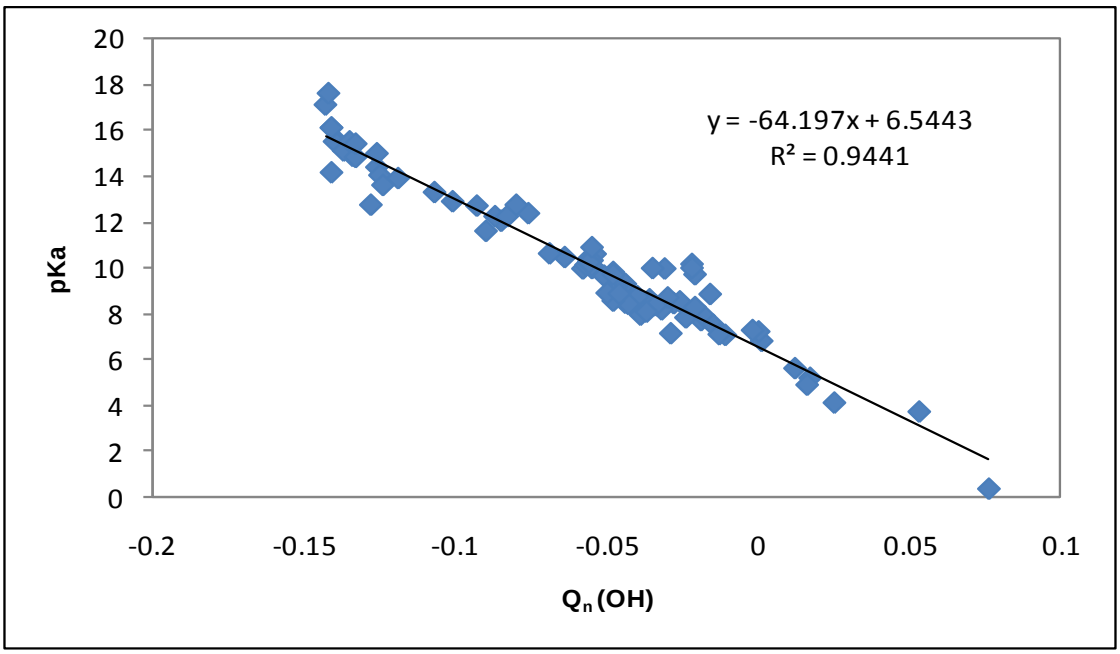


Figure 2.3: Best regression plot using RM1 for aliphatic and aromatic alcohols (Eq. 2.3)

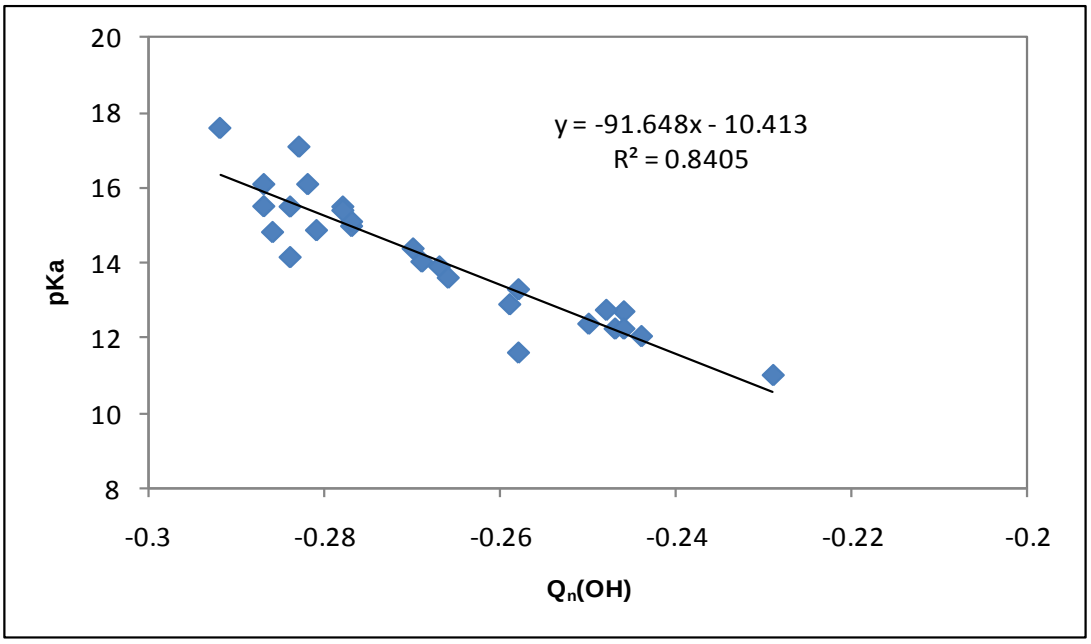


Figure 2.4: Best regression plot using B3LYP/6-31+G* for aliphatic alcohols in SM5.4 medium (Eq. 2.4)

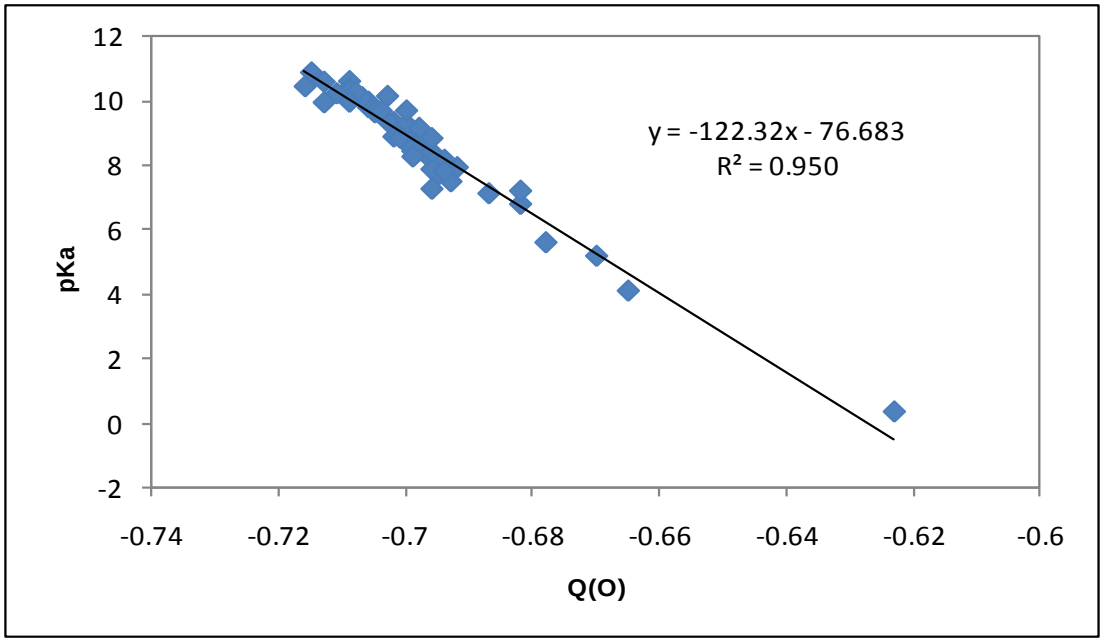


Figure 2.5: Best regression plot using B3LYP/6-31+G* for aromatic alcohols in SM5.4 medium. (Eq. 2.5)

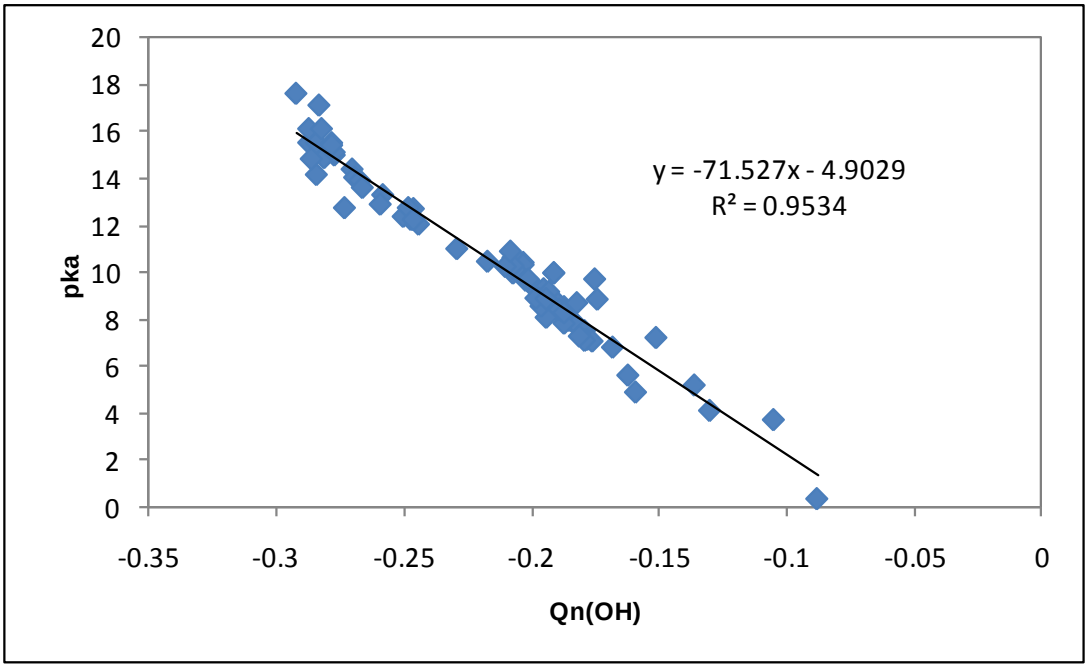


Figure 2.6: Best regression plot using B3LYP/6-31+G* for aliphatic and aromatic alcohols in SM5.4 medium. (Eq. 2.6)

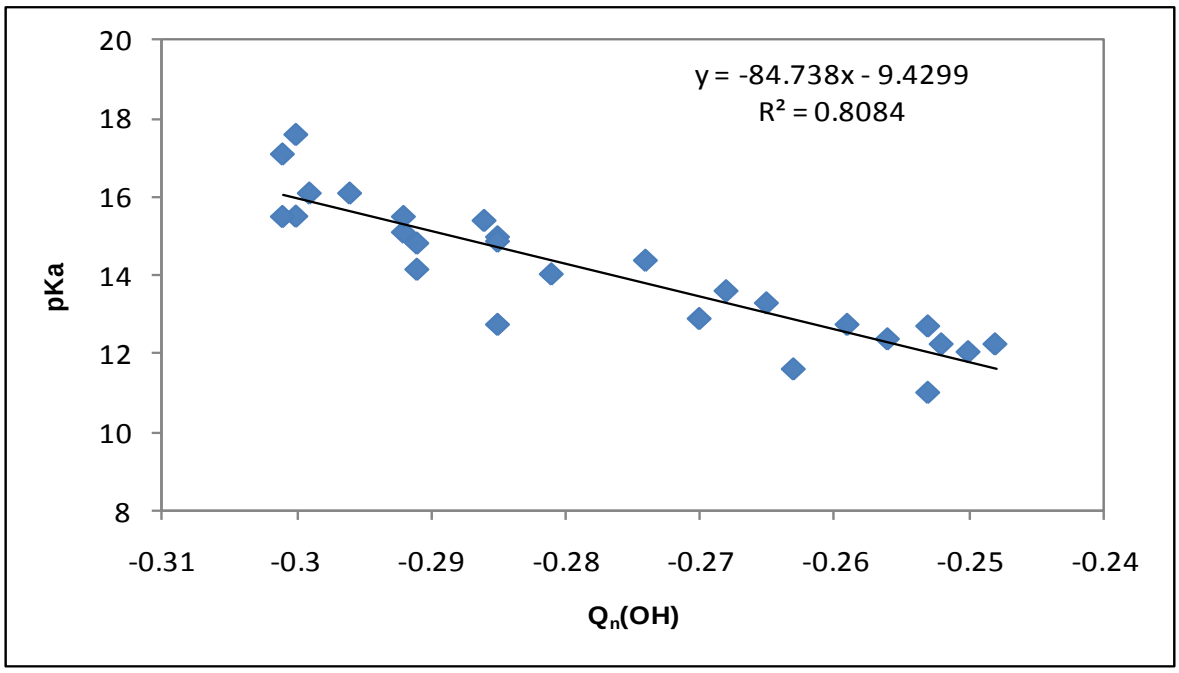


Figure 2.7: Best regression plot using B3LYP/6-31+G* for aliphatic alcohols using SM8 solvent model. (Eq. 2.7)

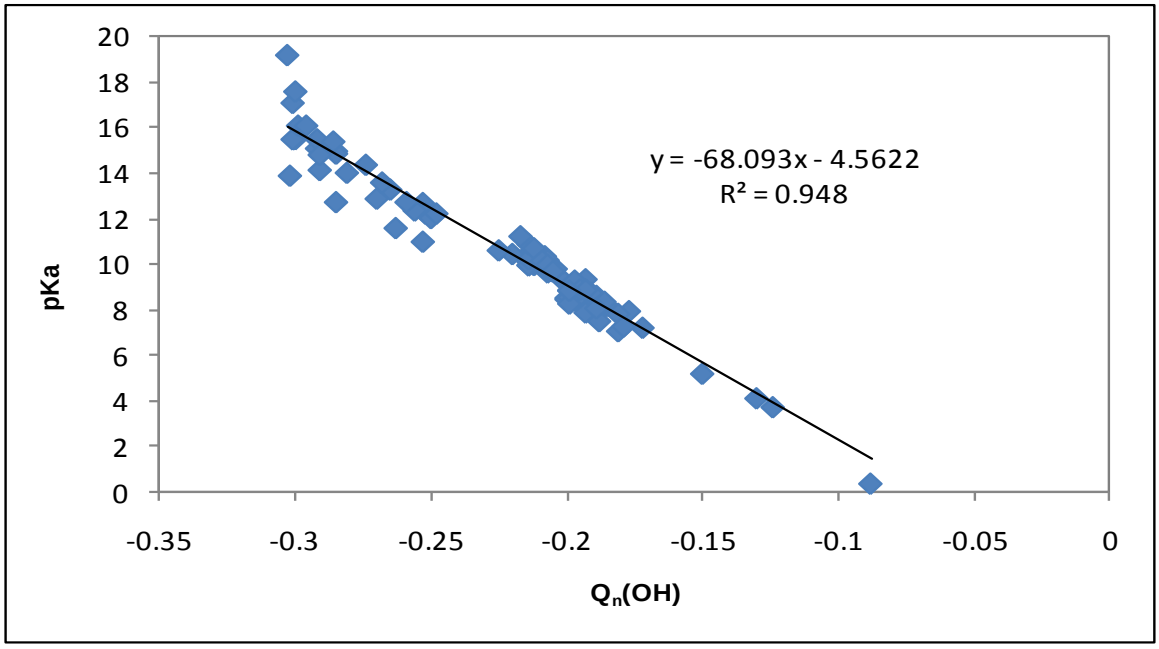


Figure 2.8: Best regression plot using B3LYP/6-31+G* for aromatic alcohols using SM8 solvent model. (Eq. 2.8)

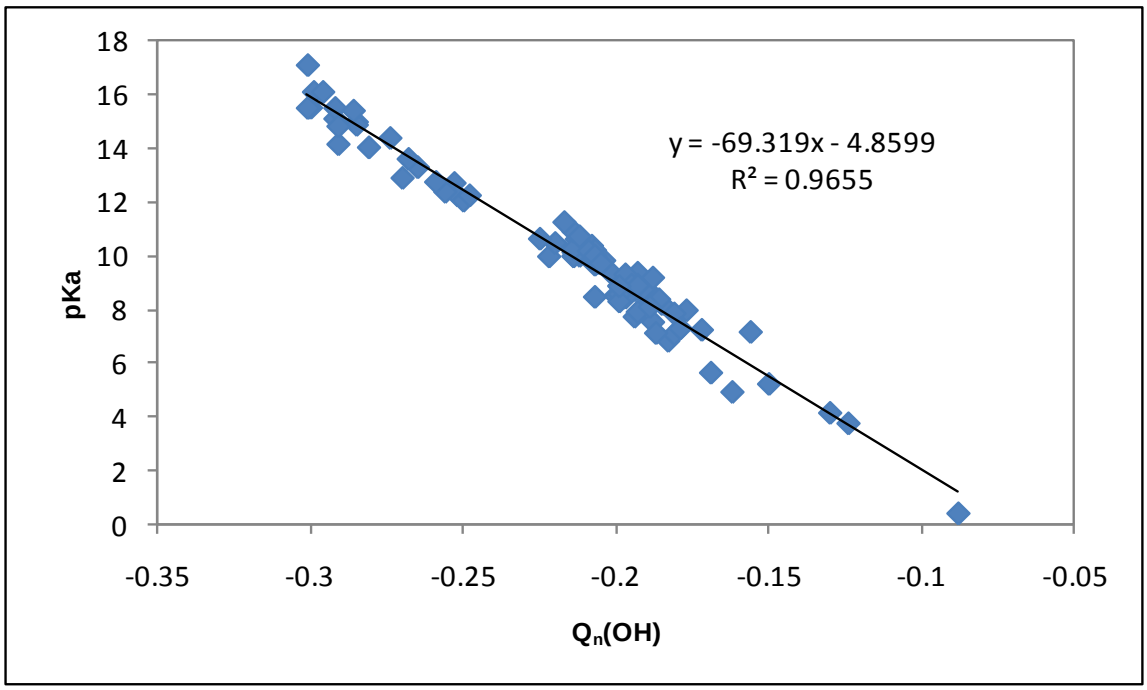


Figure 2.9: Best regression plot using B3LYP/6-31+G* for aliphatic and aromatic alcohols using SM8 solvent model. (Eq. 2.9)

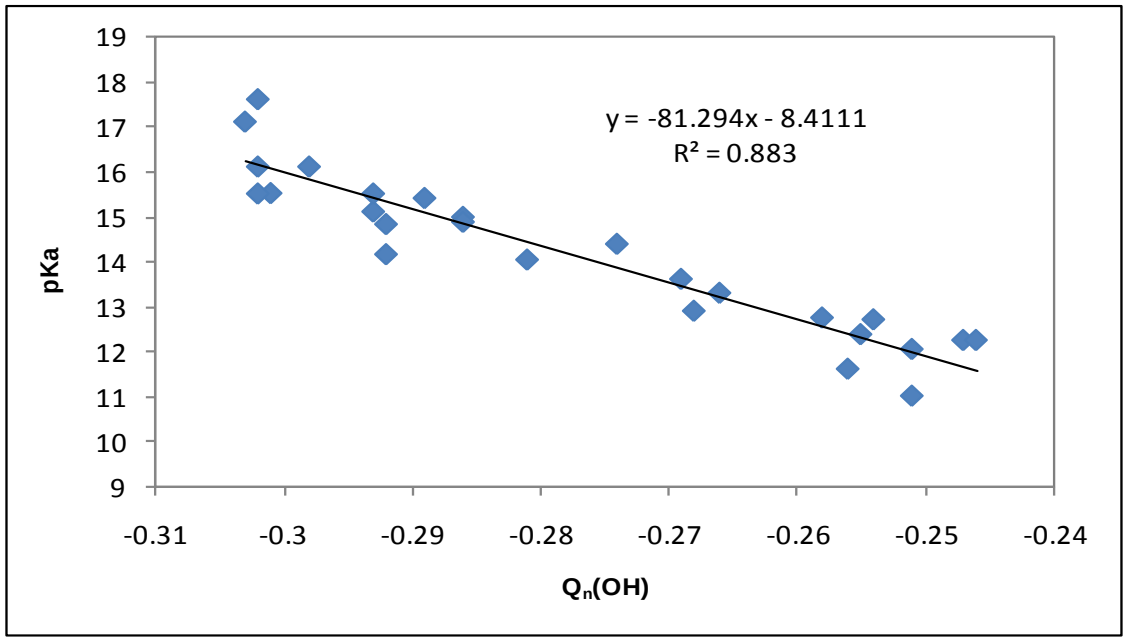


Figure 2.10: Best regression plot for optimization using B3LYP/6-31+G* for aliphatic alcohols using SM8 solvent model. (Eq. 2.10)

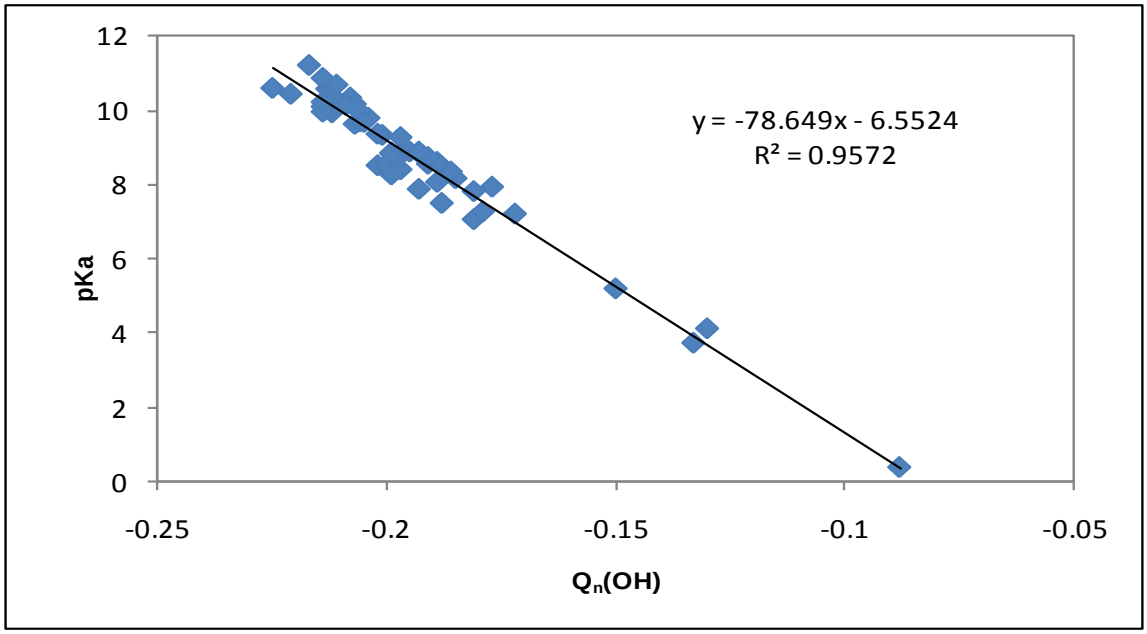


Figure 2.11: Best regression plot for optimization using B3LYP/6-31+G* for aromatic alcohols using SM8 solvent model. (Eq. 2.11)

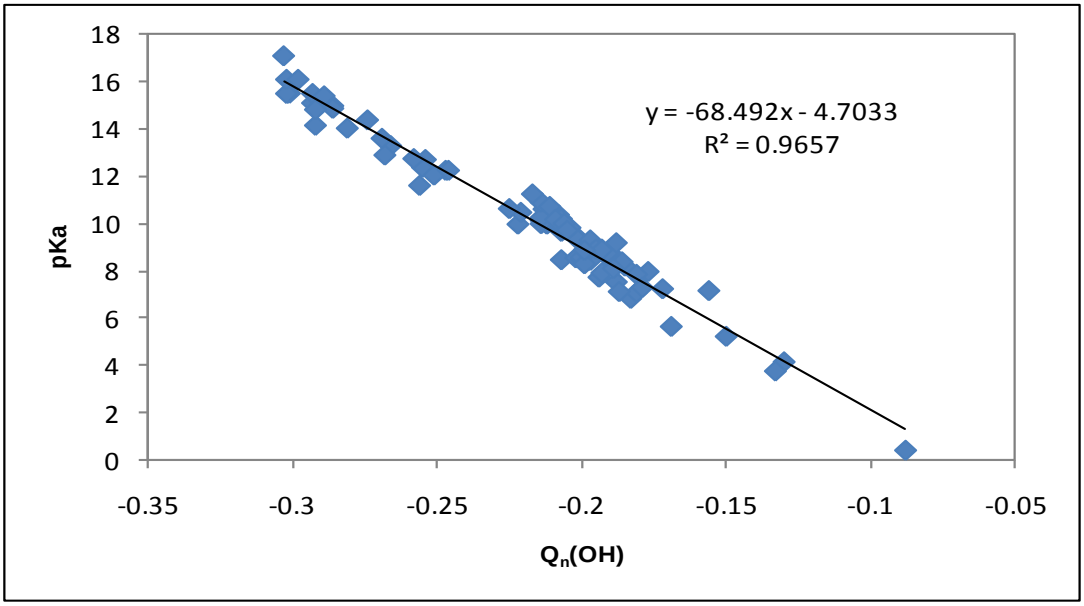


Figure 2.12: Best regression plot for optimization using B3LYP/6-31+G* for aliphatic and aromatic alcohols using SM8 solvent model. (Eq. 2.12)

2.5 References

[1] Gross, K.C.; Seybold, P.G.; Hadad, C.M. Comparison of Different Atomic Charge Schemes for Predicting pKa variations in Substituted Anilines and Phenols. *Int. J. Quantum Chem*. **2002**, *90*, 445.

[2] Gross, K.C.; Seybold, P.G. Substituent Effects on the Physical Properties and pK_a of Phenol. *Int. J. Quantum Chem*. **2001,** *85*, 569-579.

[3] Hollingsworth, C. A.; Seybold, P. G.; Hadad, C. M. Substituent Effects on the Electronic Structure and pK_a of Benzoic Acid. *Int. J. Quantum Chem*. **2002,** *90*, 1396-1403.

[4] Kreye, W. C.; Seybold, P. G. Correlations Between Quantum Chemical Indicies and the pK_as of a Diverse Set of Organic Phenols. *Int. J. Quantum Chem*. **2009,** *109*, 3679-3684.

[5] Seybold, P.G. Analysis of the pKas of Aliphatic Amines Using Quantum Chemical Descriptors. *Int. J. Quantum Chem*. **2008**, *108*, 2849-2855.

[6] Gross, K. C.; Seybold, P. G. Substituent Effects on the Structure and pK_a of Aniline, *Int. J. Quantum Chem*. **2000**, *80*, 1107-1115.

[7] Rocha, G. B.; Freire, R.O.; Simas, A.M,; Stewart, J. J. P. RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J Comput Chem.* **2006**, *27*, 1101.

[8] Robert G. Parr. Density Functional Theory. *Annu. Rev. Phys. Chem*. **1983**, *34*, 631-656.

[9] Chambers, C. C,; Hawkins, G. D,; Cramer, C. J,; Truhlar, D. G. A Model for Aqueous Solvation Based on Class IV Atomic Charges and First-Solvation-Shell Effects. *J Phys*

Chem. **1996**, *100*, 16385.

[10] Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges. *J. Chem. Theory Comput.* **2007**, *3*, 2011

[11] Lide, D.R., ed.; CRC Handbook of Chemistry and Physics, 88th ed. 2007; pp. 42-51.

[12] Gawron, O.; Draus, F. *J. Am. Chem. Soc*. **1958**, *80*, 5392

[13] Ballinger, P.; Long, F. A. *J. Am. Chem. Soc*. **1959**, *81*, 1050

[14] Ballinger, P.; Long, F. A. *J. Am. Chem. Soc*. **1960**, *82*, 795

[15] Stewart, R. The Proton: *Applications to Organic Chemistry*; Wasserman, H.H., Ed.; Vol 46 of Organic Chemistry, A series of Monographs. Academic Press: New York, 1985

[16] Dixon, S. L.; Jurs, P. C. Estimation of pKa for Organic Oxyacids Using Calculated Atomic Charges. *J. Comput. Chem.* **1993**, 14(*12*), 1460-1467.

[17] Takahashi, S,; Cohen, L. A,; Miller, H. K,; Peake, E. G. *J .Org. Chem*. **1971**, 36(9), 1207.

[18] Zhang, S,; Baker, J,; Pulay, P. A Reliable and Efficient First principles-Based Method for Predicting pKa Values. 2. Organic Acids. *J. Phys. Chem*. A. **2010**, *114*, 432-442.

[19] Spartan' 08, Wave function. Inc., Irvine, CA 92612, U.S.A.

[20] Klamt, A,; Schuurmann, G. *J. Chem.. Soc*., Perkin Trans. 2 **1993**, 799

Chapter 3

A Cellular Automata Model for the Distribution of Gases in the Earth's Atmosphere

3.1 Introduction

Almost 200 years ago John Dalton proposed that the composition of the gases in the earth's atmosphere should change with altitude,[1] the heavier gases being relatively more abundant at lower elevations and the lighter ones relatively more abundant at higher altitudes. In 2006 this proposal was experimentally confirmed at low altitudes (0-4 meters) by careful measurements of the ratio of argon to nitrogen at a desert location.[2] Earth's atmosphere is composed of different gases of which nitrogen (N_2 =78.084%), oxygen (O_2 = 20.9476%), argon (Ar = 0.934%), carbon dioxide (CO_2 = 0.0314%) are the major components.[3] In an attempt to further study the distribution of these gases in the atmosphere with respect to altitude, the present study was carried out using a cellular automaton model. The results obtained were compared with the classic barometric equation[4] to understand the atmospheric profile better.

In the past, scientists have studied the diffusive separation of gases in the lower atmosphere using diffusion equations.[2] In 2006, a cellular automata (CA) model for vapor-liquid equilibria was introduced by Seybold et al.[5] It was apparent at the time that this model could be extended to features of the atmosphere and this possibility encouraged us to carry out the present project. The cellular automata model was a stochastic (i.e., probabilistic) CA model, which implemented just two rules: a rule governing attractions between the agents and a gravitational (downward moving) preference for motions of the agents. Application of the attraction rule alone gave a

dynamic mist-like pattern, and application of the gravity rule by itself yielded an isothermal atmospheric concentration profile. Application of both rules led to formation of a condensed phase with a vapor phase above it, i.e. a pattern of liquid-vapor equilibrium. The previous study and its ideas gave us an opportunity to study the variation of the density of nitrogen, oxygen, argon and carbon dioxide gases with altitude. The results of the resulting CA model could then be compared with the barometric equation, which is the customary standard for analyzing gas/altitude distributions.

The barometric equation is[4]

$$P = P_o e^{-gMh/RT} \qquad (3.1)$$

or $$\ln(P/P_0) = -gMh/RT$$

where P = the pressure at height h, P_0 = the pressure at the base, g = the acceleration due to gravity, M = the molar mass of the gas in kg/mol, h = the height in meters, R = the gas constant, and T = the absolute temperature in kelvins.

As will be shown, the CA simulations yield ingredient distributions with altitude as row populations. They also yield fluctuations in these populations, which cannot be obtained from the barometric equation. In general, the behavior of one ingredient, such as an individual nitrogen molecule, is unpredictable, but the collective outcome for numerous ingredients over several runs yields consistent patterns, in the same way that laboratory experiments, when repeated, yield similar but not identical outcomes. Accordingly, the individual CA simulations yield similar, but not identical patterns, which can be analyzed to determine statistical properties.[5] The promise, then, was that a CA study based on relatively simple rules and using modest computing facilities could be

used to obtain an informative model for the distributions of gases in the earth's atmosphere.

3.2 Cellular automata (CA) models

Cellular automata concepts were first proposed by the mathematical physicist John von Neumann[6] and the mathematician Stanislaw Ulam[7,8] more than half century ago. Initially, Von Neumann's interest was in the construction of "self-reproducing automata." His idea was to construct a mechanical device that would gather and assemble the parts necessary to reproduce itself. A suggestion by Ulam led him to consider more abstract systems consisting of grids with moving agents, operating under sets of rules. One modern version of a CA model is the so-called "Game of Life," invented by John Horton Conway.[9] The game of life is a deterministic CA model based on very simple rules, which nonetheless can yield quite very complex patterns. Here "deterministic" means that the governing rules of the model are absolute and fixed, as distinguished from the alternative "stochastic" possibility, where the rules are expressed as probabilities of actions. The game starts with the application of simple rules on a (usually simple) pattern of ingredients placed on the grid. The pattern on the grid then evolves with time as the rules are applied to successive generations (or "iterations"), sometimes leading to complex patterns. For the most part the latter patterns are virtually impossible to predict without actually running the program. In the same way, other CA models implementing different sets of rules often yield complex patterns, which in selected cases can be interpreted as representing or simulating the behaviors of a variety of natural systems.[10]

Cellular automata models are discrete, agent-based models that can be used for the simulation of complex systems.[5] They are composed of the following components:

1. *A Grid*: This is the frame containing the cells and the agents. The grid might be one, two, or three-dimensional. Three general types of two-dimensional grids are distinguished by their boundaries: box, cylinder, and torus. In a box moving agents encounter boundaries on all four sides. In a cylinder, there are no restrictions on horizontal motions of the agents, but there are boundaries at the top and bottom of the grid. Thus agents moving off the grid to the right appear on the left, and agents moving off the grid to the left appear on the right side. In a torus, no boundaries restrict the agents' horizontal or vertical movements. The grid type used is based on what type of system one wants to simulate.

2. *Ingredients*: The agents are called ingredients and they occupy specific cells on the grid.

3. *Initial conditions*: In a CA model the starting distribution of the ingredients on the grid can be chosen. This choice includes the numbers of the different ingredients and the placement of these ingredients on the grid. The number of iterations in the simulation is also normally fixed before starting the simulation.

4. *The Rules*: The behaviors of the agents during the CA simulation are governed by a set of local rules. The rules are further explained in the next section.

3.3 Rules

Rules control the actions taken by the agents in the simulations. They may take the form of probability of an action. The rules applied during the present study are briefly explained as:

Motion rules

1. *The Free moving probability*: P_m, defines the probability that an agent A in a cell will move to one of the four adjacent cells if that space is unoccupied. If $P_m = 1$, movement of an agent in one of the directions always happens.

2. *The Joining probability*: The interaction rule J(AB) defines the movement of an agent A toward or away from a second agent B when the two are separated by a vacant cell.

If J(AB) < 1, A will tend to move away from B.

If J(AB) > 1, A will tend to move toward B.

If J(AB) = 1, there is no preference in either direction.

3. *The Breaking probability*: P_B, This parameter in effect assigns a persistence to an encounter between two agents that are in contact, i.e., touching each other on the grid.

If $P_B = 0$, the agents will not separate from each other.

If $P_B = 1$, they have no tendency to adhere to one another.

If $0 < P_B < 1$, there is an intermediate persistence.

4. *The Absolute gravity rule*, $G_A(A)$: This determines the relative probability of moving downward on the grid. Motions in all the other directions are equally probable, so that any value of $G_A(A)$ greater than 0 signifies some, usually slight, tendency for the agents to move downward. The $G_A(A)$ values for the gases in the present study were modified based on their molecular weights. Initially $G_A(A)$ was set as 0.100 for nitrogen gas(N_2) and this value was modified accordingly for the other gases studied in this chapter.

3.4 Methods

In order to run CA simulations to study the distribution of the gases in the earth's atmosphere, the CASim[11] software program was used. A thousand ingredients of each

gas (starting with nitrogen) on a 100 x 100 grid were subjected to 11,000 iterations in each run. The ingredients were initially placed randomly on the 100 x 100 grid (this excludes the top and bottom rows, which are occupied by the cylinder barriers). The first 10,000 iterations were used to establish equilibrium, and then altitude (row population) data were taken at 100 iteration intervals starting at 10,100 iterations. CASim displays the population pattern of the system as it varies with time (iterations). Initial and later stage snapshots of the dynamic process are shown in Figures 3.1 and 3.2, respectively. These steps were performed individually for all the gases studied in this chapter: nitrogen, oxygen, argon, and carbon dioxide.

To summarize, the input for carrying out the simulations was:

1. Grid: 100 rows by 100 columns (100 x 100) plus rows occupied by the top and bottom barriers.

2. Number of ingredients of the gas: 1000

3. Grid type: Cylinder (The molecules are confined in a cylinder with restraining top and bottom barriers. Molecules that leave from the right side appear on the left side and vice versa.)

4. Initial condition: Random placement of the ingredients on the grid.

5. Total number of iterations in each run = 11,000

6. Number of runs for each gas = 12

7. $G_A(N_2)$, Gravity value for nitrogen = 0.100

8. Breaking Probability, P_B =1; this assigns zero stickiness to the interaction between two ingredients.

9. Joining Probability J(AB) =1; this defines no tendency for A to move toward or

away from another ingredient B when these two are separated by an empty cell.

10. Free moving probability, P_m (N_2) = 1; this rule means that movement in one of the directions always happen.

In each run, the output obtained was grouped into 10 snapshots starting from 10,100, 10,200, 10,300….up to 11,000 iterations, each focusing on the populations of the bottom 30 rows of the grid (These lower rows had enough ingredients for statistical significance). At first, the bottom 30 rows were studied, but then it was found that the bottom barrier causes an anomalous effect in which the ingredients tend be too sparse in the lowest rows. In order to avoid this artifact, the lowest 3 row counts were ignored and all the gases including nitrogen were studied by taking the bottom 30 rows minus the lowest three rows. Each gas was subjected to 12 runs and the average of total counts for each row was calculated. Then the natural log was taken of this average value. This obtained value, which is natural log of average counts of each row was plotted against the row number to quantify the molecules behavior with respect to height (in this case, row). A standard deviation was also calculated for each average. The output thus obtained was in two forms, visual and numerical. The output was analyzed and compared with the barometric equation.

Also in the nitrogen gas simulations, a binning experiment was carried out. In this binning experiment, grouping of the row counts was done for one run using bins of 10 rows each to get statistically valid data. This binning was performed because there were very low counts in the upper rows. Later on, this was helpful in comparing with the data from the bottom 30 rows for nitrogen.

Finally, CA simulations for all four gases were carried out individually by the changing the $G_A(A)$ values, starting with $G_A(N_2) = 0.200$ to check whether increasing the gravity yields better results or not.

3.5 Results and discussion

After setting the rules and the initial conditions for each gas, the ingredients, 1000 for each gas, were subjected to 12 runs of 11000 iterations each. 1000 ingredients were selected to get sufficient data for statistical purposes. The snapshots shown in Figures 3.1 and 3.2 shows that in the early stage of the simulation for nitrogen gas the ingredients appear essentially randomly on the grid. Figure 3.2, the snapshot of the simulation after 10,000 iterations, shows that the ingredients reach a dynamic steady state after 10,000 iterations.

a) Approach 1:

The data obtained after running the nitrogen gas CA simulation for 11000 iterations for 12 runs are summarized in Table 3.1 which shows the total number of counts (ingredients) in each run for the bottom 30 rows of the grid. Only the bottom 30 rows were considered and the natural log of the number of counts was taken in order to get good statistics. In Table 3.1 we can see that the higher rows tend to have fewer counts compared to the lower rows. This tells us that the gas gets less dense as we go up in the atmosphere. In order to get an exponential fall off profile for the nitrogen gas molecules, the average row counts N(r) were plotted against the row number 'r' for the lowest 30 rows. This is shown in Figure 3.3. Further regression analysis was performed between the natural log of the average counts and row number and the equation obtained is shown below and plotted in Figure 3.4.

$$\ln (N) = -0.038 (\pm 0.001)\, r + 6.069(\pm 0.014) \qquad (3.2)$$

$$n = 30, \quad R^2 = 0.977, \quad s = 0.053, \quad F = 1824$$

where n is the number of observations from the bottom 30 rows, R^2 is the coefficient of determination, s is the standard error of the estimate for the regression, and F is the Fisher statistic.

Table 3.2 is the data for the binning experiment with ten bins, each bin summing the counts of molecules for 10 rows. The plot between the bin number and natural log of the bin population is plotted in Figure 3.5, and was similar to the plot in Figure 3.4. The binning was done so as to have enough counts so that the data can be properly analyzed. The data obtained from the binning experiment can be compared with the Figure 3.4 results (which is for the lowest 30 rows and 12 runs). The results obtained for one run (Table 3.2 and Figure 3.5) are similar to the results obtained for 12 runs for nitrogen gas (Table 3.1 and Figure 3.4), which suggests that just as when laboratory experiments are repeated and yield similar results, CA simulations also give similar results when repeated, with slight variations in the number of counts. The resultant equation for the binning experiment is:

$$\ln(N) = -0.538(\pm 0.022)\, r + 8.954(\pm 0.136) \qquad (3.3)$$

$$n = 10, \quad R^2 = 0.987, \quad s = 0.199, \quad F = 607$$

where N is the sum of each bin and r is the bin number which contains 10 rows.

b) Approach 2:

From the data obtained for 12 runs for the nitrogen gas (Table 3.1), it was found that the zeroth row had low counts compared to the rows above it, which was puzzling since the bottom row should give the highest number of counts. After analyzing the data it was

discovered that the CASim software is designed in such a way that the lower barrier, which prevents downward motion, tends to favor motion of ingredients in an upward direction. This artificially depletes the population in the very lowest rows and slightly exaggerates the populations in the nearest rows. This artifact due to the lower barrier is clearly visible in Figures 3.3 and 3.4, where the zero row data point is not fitting into the remaining data. In order to avoid this programming artifact the data were collected for the 27 rows excluding the bottom three rows, as summarized in Table 3.3. We can call this "Approach 2". Figure 3.6 shows the exponential fall off profile for the nitrogen gas molecules, the average row counts N(r) was plotted against the row number 'r' for the lowest 27 rows. The regression statistics for the nitrogen gas for this approach (taking 27 of the bottom 30 rows) was performed between the natural log of the average counts and row number.

The equation obtained is shown below and is plotted in Figure 3.7

$$\ln (N) = -0.034(\pm 0.001)\, r + 5.930(\pm 0.014) \qquad (3.4)$$

$$n = 27, \quad R^2 = 0.984, \quad s = 0.035, \quad F = 731$$

Since this approach for the nitrogen gas gave better regression statistics, with R^2 = 0.984 and s = 0.035, the same approach was adopted for the rest of the gases: oxygen, argon and carbon dioxide. The rules applied for the rest of the gases were the same, except that the $G_A(A)$ and P_m values for each of these gases were changed based on the molecular weight and free moving probability with respect to the nitrogen gas. The initial conditions were kept the same. To obtain a correspondence between the CA results above and the barometric equation the following calculation was done for nitrogen gas.

c) Nitrogen gas

From the CA plot, Figure 3.7, ln (N) = -0.034(± 0.001) r + 5.930(± 0.014)

and from the Barometric equation, ln (P/Po) = - (gM/RT) h (3.5)

Substituting numerical values into equation (3.5) [g = 9.8 ms^{-2}, M = 28 $x10^{-3}$ kg mol^{-1}, R = 8.314 $m^2kgs^{-2}K^{-1}$ mol^{-1}, and T = 298.15 K (25+ 273.15)] we get,

- (gM/RT) = - (9.8 x28 x 10^{-3})/(8.314 x 298.15) (ms^{-1}. kg mol^{-1})/($m^2kgs^{-2}K^{-1}$ mol^{-1}. K)

- (gM/RT) = -0.111 x $10^{-3}m^{-1}$

ln(P/Po) = -0.111 x $10^{-3}m^{-1}$ (3.6)

now, comparing the slopes in equations (3.4) and (3.6),

-0.034 (± 0.001) r = -0.111 x $10^{-3}m^{-1}$

therefore, 1 row = -0.034 (± 0.001)/-0.111 x $10^{-3}m^{-1}$ = 306 ± 9 meters. Thus each row in the CA simulation for N_2 corresponds to about 306 m in altitude.

d) Oxygen gas

Since the molecular mass of the oxygen gas(32 g/mol) is greater than that of nitrogen(28 g/mol), the parameters $G_A(O_2)$ and $P_m(O_2)$ were modified. The $G_A(O_2)$ value was calculated by comparing the ratio of molecular weights of two gases to their G_A values. The calculations as how $G_A(O_2)$ was calculated is shown below:

$G_A(O_2)/ G_A(N_2)$ = 32/28, (32 and 28 are molecular weights of O_2 and N_2)

$G_A(N_2)$ = 0.1 (set initially for nitrogen gas)

therefore, $G_A(O_2)$ = (32/28)*0.1 = 0.114. This calculation reflects the greater gravitational attraction of heavier gas.

Now according to the kinetic molecular theory of gases, at a given temperature the molecules of all species of gas have the same average kinetic energy (given by 1/2 mv^2).

So based on this assumption the calculation for the free moving probability $P_m(O_2)$ is shown below:

$1/2\ m(O_2)v^2(O_2) = 1/2\ m(N_2)v^2(N_2)$

$v^2(O_2)/\ v^2(N_2) = m(N_2)/m(O_2),\ \ v(O_2) = [\sqrt{}\ m(N_2)/m(O_2)].\ v(N_2)$

Since we can assume that $P_m\ (O_2)$ is proportional to $v(O_2)$,
$P_m\ (O_2)/\ P_m(N_2) = \sqrt{}28/32$, and since $P_m(N_2) = 1.0$ (set initially for nitrogen gas)
$P_m\ (O_2) = 1.0*\sqrt{}(28/32) = 0.935$. This calculation answers that heavier gases move more slowly than the lighter gas.

The CA simulation results for oxygen gas for 12 runs are tabulated in Table 3.4. It was found as expected, that the counts in the lower rows for oxygen were greater compared to those for nitrogen under the same set of conditions. This is mainly because of the increase in $G_A(O_2)$.

From the CA plot in Figure 3.9 we get,

$$\ln (N) = -0.041\ (\pm\ 0.001)\ r + 6.079(\pm\ 0.016)\ , \qquad (3.7)$$

$$n = 27,\quad R^2 = 0.985,\quad s = 0.040,\quad F = 937$$

and, $\ln (P/Po) = -\ (gM/RT)\ h$ (from the barometric equation). (3.8)

With $M = 32 \times 10^{-3}$ kg mol^{-1} in the barometric equation for oxygen one obtains

$$\ln(P/Po) = -0.127 \times 10^{-3} m^{-1} \qquad (3.9)$$

Now, comparing the slopes in equations (3.7) and (3.9),

$-0.041\ (\pm\ 0.001)\ r = -0.127 \times 10^{-3} m^{-1}$

Therefore, 1 row = $-0.041(\pm\ 0.001)/\ -0.127 \times 10^{-3} m^{-1} = 323 \pm 8$ meters for O_2. Thus each row in the CA simulation corresponds to about 323 m in altitude. The change in the parameters $G_A(A)$ and P_m does change the output, showing that the counts for the oxygen gas are in the lower rows and fewer in the upper rows compared to the nitrogen gas.

The exponential falloff is shown in Figure 3.8 and the regression analysis plot between the natural log of the average counts and row number and the equation obtained is shown in Figure 3.9 which also gives us $R^2 = 0.985$, which indicates that the CA simulations gave reasonable results for oxygen gas also. The more negative slope for oxygen gas (Eq. 3.6) suggests that oxygen, being heavier than nitrogen, falls more sharply in concentration with altitude under the applied rules.

e) Argon gas

Approach 2 was also followed in the case of argon gas. The reasons for changing the G_A and P_m for argon gas are the same as for oxygen. The calculations for $G_A(Ar)$ and $P_m(Ar)$ are shown below:

$G_A(Ar)/ G_A(N_2) = 39.95/28$, (39.95 and 28 are the molecular weights of Ar and N_2, respectively) so that, $G_A(Ar) = (39.95/28)*0.1 = 0.143$.

For $P_m (Ar)$, $[P_m (Ar)/ P_m(N_2)]^2 = 28/39.95$, so that $P_m (Ar) = 1.0* \sqrt{(28/39.95)} = 0.837$

The input for carrying out the simulations for argon gas were same as for N_2 and O_2 except that $G_A(Ar) = 0.143$ and $P_m(Ar) = 0.837$. The CA simulation results for argon gas for 12 runs are tabulated in Table 3.5.

Now from the CA plot (Figure 3.11) we get,

$$\ln (N) = -0.053 (\pm 0.001)\, r + 6.270(\pm 0.023) \qquad (3.10)$$

$$n = 27,\ R^2 = 0.982,\ s = 0.058,\ F = 819$$

From the barometric equation, $\ln (P/Po) = - (gM/RT)h$ (3.11)

With $M = 39.95 \times 10^{-3}$ kg mol^{-1} for argon one obtains

$$\ln(P/Po) = -0.158 \times 10^{-3} m^{-1} \qquad (3.12)$$

Now, comparing the slopes in equations (3.10) and (3.12),

$-0.053(\pm 0.001)$ r = $-0.158\ X10^{-3}m^{-1}$. Therefore, 1 row = $-0.053(\pm 0.001)/ -0.158\ X10^{-3}m^{-1}$ = 335 ± 6 meters and each row in the CA simulation corresponds to about 335 m in altitude. The exponential fall off is shown in Figure 3.10. The regression analysis plot between the natural log of the average counts and row number is shown Figure 3.11, which also gives us $R^2 = 0.982$. This suggests that the CA simulations gave reasonable results for argon gas also.

f) Carbon dioxide gas

The greenhouse gas carbon-dioxide gas was also included in this study of atmospheric gases. The calculations for $G_A(CO_2)$ and $P_m(CO_2)$ are shown below:

$G_A(CO_2)/ G_A(N_2) = 44/28$, (44 and 28 are the molecular weights of CO_2 and N_2), and since $G_A(N_2) = 0.1$, therefore, $G_A(CO_2) = (44/28)*0.1 = 0.157$. For $P_m(CO_2)$, $[P_m(CO_2)/ P_m(N_2)]^2 = 28/44$, and therefore, $P_m(CO_2) = 1.0 *\sqrt{(28/44)} = 0.798$

The CA simulation routine was the same as in the previous studies except for the parameters $G_A(CO_2) = 0.157$ and $P_m(CO_2) = 0.798$. The CA simulation results for 12 runs are tabulated in Table 3.6.

Now from the CA plot (Figure 3.13) we get,

$$\ln (N) = -0.058(\pm 0.002)\, r + 6.346(\pm 0.281)\,, \qquad (3.13)$$

$n = 27,\ R^2 = 0.977,\ s = 0.072,\ F = 1389$

and, $\ln (P/Po) = -(gM/RT)h$ (from Barometric equation) (3.14)

With $M = 44 \times 10^{-3}$ kg mol^{-1} for carbon dioxide one obtains

$$\ln(P/Po) = -0.174\ X10^{-3}m^{-1} \qquad (3.15)$$

Now, comparing equations (3.13) and (3.15),

$-0.058(\pm 0.002)$ r = $-0.174\ X10^{-3}m^{-1}$. Therefore, 1 row = $-0.058(\pm 0.002)/-0.174X10^{-3}m^{-1}$ = 333 ± 11 meters, so that each row in the CA simulation corresponds to about 333 m in

altitude. The exponential falloff is shown in Figure 3.12 and the regression analysis plot between natural log of the average counts and row number and the equation obtained is shown in Figure 3.13, which gives us $R^2 = 0.977$. This suggests that the CA simulation gives a reasonable model for carbon dioxide gas also. Table 3.7 summarizes the regression statistics for all four gases for 27 of the bottom 30 rows. The plot showing the variation of all four gases with respect to height is shown in Figure 3.14, where we can see how the gases fall under the fixed rules. The somewhat unclear nature of the curves at the bottom might be due to the bottom barrier effect, which alters the ingredient distribution.

g) Approach 3 - Effect of increase in gravity rule

In an attempt to find out the effect of increasing the gravity parameter, $G_A(A)$, CA simulations were run for all of the gases starting with nitrogen as $G_A(N_2) = 0.200$. The $G_A(A)$ values for the remaining three gases were changed based on the nitrogen value. The P_m values were kept constant. The results are shown in Table 3.8 and the plot showing the variation of the four gases is shown in Figure 3.15. As we can see in Table 3.8, the R^2 values did not improve much and the fall of for all the gases was curved (Figure 3.15) rather than fitting a trend line. A possible reason for this result could be the intense crowding of ingredients in the bottom rows, which could block the ingredients above them from moving down under the increased gravity. As noted above the bottom barrier also causes an anomaly making the ingredients bump back from the lowest rows. Hence this approach of attempting the CA simulation at high gravity didn't prove to be successful. For a real gas the average volume occupied by a molecule is 0.00405×10^7 cubic angstroms at $25^{O}C$. At STP, the volume occupied by a mole of gas is 22.4 liters.

At 298K ($25^{O}C$) the volume occupied by one mole of gas is 24.4 liters. 24.4 liters are equal to 0.0244 cubic meters. Now in one mole we have 6.023×10^{23} molecules, so it means we have 6.023×10^{23} molecules in 0.0244 cubic meters. Converting the volume to cubic angstroms we get on average 4.05×10^{4} cubic angstroms per molecule at $25^{0}C$. This corresponds to 34.34 angstroms of distance of separation between two molecules. In CA approach the crowding and the distance between two ingredients changes based on gravity parameter.

h) Comparison of CA results with the Barometric equation:

In order to compare the CA results with the barometric equation, the experimental results obtained from cellular automata simulations of nitrogen and argon gases were compared with the barometric equation calculations for the Ar/N_2 ratio at different heights up to 4m.

Calculation of Ar/N_2 ratio using the barometric equation:

The barometric equation is $\ln (P/P_0) = -gMh/RT$, or $P/P_0 = e^{-gMh/RT}$ (3.16)

For nitrogen, at height 0 m the slope in eq. 3.16 becomes,

$e^{-gMh/RT} = e^{(-9.8 \times 28 \times 10\text{-}3/8.314 \times 298.15) \times 0} = 1$

As the percentage of nitrogen gas in the atmosphere is 78.084% so we get,

$e^{-gMh/RT} = e^{(-9.8 \times 28 \times 10\text{-}3/8.314 \times 298.15) \times 0} = 1 \times 0.78084 = 0.780840000$ (% of N_2 at 0 m)

Similarly, at 1m, $e^{(-9.8 \times 28 \times 10\text{-}3/8.314 \times 298.15) \times 1} = 0.998893642 \times 0.78084 = 0.780753568$,

at 2m = 0.780667145, 3m = 0.780580733 and at 4m = 0.78049433

For Argon, at the base, 0 m, the slope in eq.3.15 becomes,

$e^{-gMh/RT} = e^{(-9.8 \times 39.95 \times 10\text{-}3/8.314 \times 298.15) \times 0} = 1$

The percentage of argon gas in the atmosphere is 0.9304% so we get,

$e^{-gMh/RT} = e^{(-9.8 \times 39.95 \times 10\text{-}3/8.314 \times 298.15) \times 0} = 1 \times 0.009340000 = 0.009340000$ (% of Ar at 0m)

Similarly, at 1 m = 0.009338524, 2 m = 0.00933705, 3 m = 0.009335575, and at 4 m = 0.009334101.

To calculate the Ar/N_2 ratios at different heights the formula used was

$$Ar/N_2 = \{[(Ar/N_2)_S - (Ar/N_2)_R] / (Ar/N_2)_R\} \quad (3.17)$$

where, S is for the sample and R is the reference (i.e. the ratio at 0 m).

The $(Ar/N_2)_R$ is calculated by taking the ratio at 0 m (0.00934/0.78084 = 0.011961477.

$$\text{Therefore } Ar/N_2 \text{ at 1m} = \frac{(0.009338524/0.780753568) - 0.011961477}{0.011961477} = -0.000047318$$

As the obtained ratio is very small, there is a need to convert the resulting ratio to "meg" (parts per million). Converting to meg gives $-0.000047318 \times 10^6 = -47.3$ meg

at 2m = -94.5 meg

at 3m = -141.7 meg

at 4m = -188.9 meg

Calculation of Ar/N_2 ratio from the CA results:

For nitrogen from the CA plot we got,

$$\ln (N) = -0.034(\pm 0.001)\, r + 5.930(\pm 0.014) \quad (3.18)$$

At the base, 0 m, ln(N) = 5.930

N = 376.1545

As the CA results are in rows so we need to convert rows to meters,

1 row = 306 ± 9 meters from CA calculations

N = 376.1545/306 = 1.229263117 x 0.78084 = 0.959857812

at 1m, ln (N) = [-0.034(± 0.001) x 1]/306

$N = e^{[-0.034(\pm 0.001) \times 1]/306} = 0.99988895 \times 0.78084 = 0.780753244$ (% of N_2 at 1m)

at 2m = 0.780666499, at 3m = 0.780579763 and at 4m = 0.780493037

For Argon from the CA plot we got,

ln (N) = -0.053 (± 0.001) r + 6.270(± 0.023) (3.19)

at 0m, ln(N) = 6.270

N = 528.477

1 row = 335 ± 6 meters from CA calculations

N = 528.477/335 = 1.577544412 x 0.009340000 = 0.014734264

at 1m, ln (N) = [-0.053(± 0.001) x 1]/335

$N = e^{[-0.053(\pm 0.001) \times 1]/335} = 0.999841803 \times 0.009340000 = 0.009338522$

at 2m = 0.009337045, at 3m = 0.009335568 and at 4m = 0.009334091

To calculate the Ar/N_2 ratios at different heights the formula in eq.3.16 was used

$(Ar/N_2)_{1m}$ = [(0.009338522/0.780753244) - 0.011961477] / 0.011961477 = -47.151 meg

$(Ar/N_2)_{2m}$ = -94.2 meg

$(Ar/N_2)_{3m}$ = -141.3 meg

$(Ar/N_2)_{4m}$ = -188.4 meg

The calculations for the barometric equation and CA results are summarized in Table 3.9 and plotted in Figures 3.16 and 3.17. The two plots show that the CA results were very close to the barometric equation results which support the credibility of the CA simulations of the atmospheric gases. The negative numbers are the reduced parts per million starting from height = 0 as the starting point, where the ratio of argon to nitrogen at the ground level is is used as a reference.

3.6 Conclusions

The results obtained for different gases and the R^2 values suggest that CA methods can be used to simulate the disposition of atmospheric gases with altitude. Figures 3.1 and 3.2 show how the simulation of the gas changes as the system evolves from a random placement. Approach 1 shows that the counts in the very lowest rows are less than expected, apparently because the barrier at the bottom of the cylinder stops the molecules from going down and exaggerates their upward movements. With this exception, the CA model (Approach 2) accurately simulates the atmospheric system, showing that when one goes higher in the atmosphere the density of the gases decreases exponentially. The logarithmic equations were used to calculate the row/height equivalent using the barometric formula. An attempt to simulate the gas behavior by increasing the gravity (Approach 3), proved to be unsuccessful as the fall offs of the gases were not linear, but curved presumably due to excessive crowding of the ingredients.

It was also concluded that when a simulation is repeated with the same initial conditions, the outcomes vary slightly, which reflects the expected results in a natural fluctuating system. We can see this variation in Table 1, where the final number of molecules in each row varies for the 12 runs. Thus the model captures the complexity of this natural system. However, a disadvantage is that the model is an isothermal model, which is not true practically in the Earth's atmosphere. Another disadvantage is that all the gases cannot be put together in one CA simulation to analyze the data. As the counts of carbon dioxide and argon are very small compared to those of nitrogen and oxygen, it is impractical to combine all the gases in a single simulation.

The comparisons of CA results with the barometric equation were successful. The CA results for the Argon/Nitrogen ratio were close to the theoretical barometric equation predictions (Table 3.9 and Figures 3.16-3.17). The CA simulation output gives us the errors associated with the system that is not possible from the barometric equation. This can be considered an advantage of the CA model over the barometric equation model.

Overall, CA simulation results are visual and dynamic. The simulation results can be associated with the statistical fluctuations present in the natural processes, and which are not revealed by the barometric equation model. Cellular automata simulations can be carried out using modest computing facilities, employ rather simple rules, and most importantly, yield realistic representations of the systems under study.

Table 3.1: The No. of molecules in bottom 30 rows of nitrogen gas for 12 runs.

						RUNS								
Row No.	1	2	3	4	5	6	7	8	9	10	11	12	Avg 12 Runs	Std Dev	ln Avg
29	137	121	126	133	121	123	154	137	145	134	152	140	135	11	4.905
28	117	143	145	146	149	152	163	134	150	168	136	154	146	13	4.983
27	147	158	146	153	149	133	157	134	139	153	154	161	149	9	5.003
26	149	173	163	152	155	131	149	178	144	155	150	152	154	13	5.036
25	157	157	164	162	151	160	163	154	170	146	167	184	161	10	5.081
24	168	187	158	174	188	178	186	184	176	183	167	172	177	9	5.176
23	172	159	185	167	159	207	174	176	169	191	173	203	178	16	5.181
22	188	187	192	176	182	196	190	171	202	172	202	195	188	11	5.236
21	179	230	212	182	197	220	181	207	203	199	198	187	200	16	5.298
20	219	213	206	226	225	210	214	202	206	217	194	205	211	9	5.351
19	225	230	227	199	210	242	217	215	208	222	215	235	220	12	5.39
18	224	237	209	217	218	225	215	213	204	217	239	241	222	12	5.402
17	259	229	237	244	233	256	254	229	235	235	238	244	241	10	5.484
16	254	229	273	228	233	270	250	242	241	247	272	226	247	17	5.509
15	246	279	243	257	270	257	253	257	257	265	225	268	256	14	5.545
14	250	289	263	284	250	266	273	265	265	247	272	266	266	13	5.583
13	293	283	285	283	297	282	276	276	282	296	269	269	283	9	5.645
12	297	262	319	295	293	287	296	288	269	273	294	281	288	15	5.662
11	315	310	299	294	249	295	324	311	304	308	316	326	304	20	5.717
10	297	295	311	304	307	321	290	294	304	278	323	282	301	14	5.707
9	306	318	325	303	328	312	331	287	305	290	302	306	309	14	5.733
8	326	317	359	331	342	337	320	347	355	329	355	322	337	15	5.82
7	346	354	332	354	359	356	345	365	353	331	313	339	346	15	5.846
6	359	300	357	363	315	334	345	343	341	347	337	359	342	19	5.834
5	362	395	371	359	373	370	343	319	345	350	358	347	358	19	5.88
4	382	387	362	371	375	362	373	359	365	374	360	368	370	9	5.913
3	378	386	395	385	409	355	379	393	371	396	387	328	380	21	5.94
2	416	385	417	373	418	365	400	389	384	387	383	370	391	18	5.968
1	379	426	400	403	412	399	393	425	378	393	401	396	400	15	5.991
0	365	340	384	377	359	371	363	358	366	344	360	346	361	13	5.888

Table 3.2: Data for binning experiment

Bin No.	Bin Sum	ln Bin Sum
10	29	3.367
9	66	4.189
8	104	4.644
7	152	5.023
6	319	5.765
5	662	6.495
4	1162	7.057
3	1736	7.459
2	2483	7.817
1	3287	8.097

.

Table 3.3: The No. of molecules in bottom 27 rows of nitrogen gas for 12 runs.

						Runs									
Row No.	1	2	3	4	5	6	7	8	9	10	11	12	Avg of 12 runs	Std Dev	ln(avg)
27	138	144	131	144	155	157	143	144	142	137	157	143	145	8	4.974
26	180	167	129	132	154	162	138	150	131	136	147	172	150	17	5.010
25	152	143	144	157	148	146	161	150	144	144	145	135	147	7	4.993
24	132	165	167	172	179	162	175	171	161	151	172	167	165	13	5.103
23	174	155	147	166	177	172	146	175	167	176	156	159	164	11	5.101
22	163	192	156	187	189	172	176	178	179	179	182	166	177	11	5.174
21	177	201	165	171	181	203	197	162	166	200	187	192	184	15	5.212
20	177	186	192	202	184	185	209	175	174	167	181	193	185	12	5.223
19	216	190	215	198	198	203	203	217	196	218	210	221	207	10	5.333
18	203	201	199	214	226	217	205	207	221	211	200	212	210	9	5.346
17	204	206	209	234	229	214	225	201	201	199	215	228	214	12	5.365
16	229	240	210	244	205	228	224	228	233	234	213	223	226	12	5.420
15	238	240	235	232	224	222	250	218	221	231	221	235	231	10	5.441
14	262	232	257	239	246	243	240	235	246	241	235	232	242	9	5.490
13	244	240	229	241	257	260	229	217	242	276	242	249	244	16	5.496
12	263	259	239	269	250	262	271	259	268	267	252	272	261	10	5.564
11	263	255	278	290	275	267	257	269	261	257	266	264	267	10	5.587
10	282	256	257	272	270	262	299	281	279	287	280	249	273	15	5.609
9	308	289	281	274	279	264	278	287	265	274	287	288	281	12	5.639
8	314	286	294	277	301	305	294	250	304	281	297	275	290	17	5.669
7	307	316	286	288	312	303	304	302	309	293	281	321	302	12	5.710
6	304	277	308	318	316	295	314	305	306	316	304	303	306	11	5.722
5	321	298	306	311	305	312	318	321	301	289	324	317	310	11	5.737
4	336	319	317	314	343	331	305	337	327	321	317	313	323	11	5.779
3	347	317	338	319	317	364	309	337	331	311	327	345	330	17	5.800
2	349	337	328	330	330	338	350	335	338	330	329	346	337	8	5.819
1	347	333	345	327	346	344	321	343	352	301	360	352	339	16	5.827

Table 3.4: The No. of molecules in bottom 27 rows of Oxygen gas for 12 runs.

						Runs									
Row No.	1	2	3	4	5	6	7	8	9	10	11	12	Avg of 12 runs	Std Dev	ln(avg)
27	129	119	127	145	119	119	146	127	131	139	123	135	130	10	4.867
26	130	147	146	143	147	147	138	130	115	154	143	131	139	11	4.936
25	139	149	138	174	161	161	166	141	148	171	138	154	153	13	5.033
24	163	155	170	153	170	170	182	153	158	177	161	152	164	10	5.098
23	153	161	162	157	161	161	150	169	166	171	170	164	162	7	5.088
22	173	178	173	174	166	166	174	183	182	178	168	177	174	6	5.161
21	177	197	199	178	171	171	170	183	169	203	179	196	183	13	5.208
20	190	191	187	207	199	199	214	186	225	207	186	201	199	12	5.295
19	214	206	213	192	199	199	199	219	202	219	198	230	208	11	5.335
18	220	195	209	229	210	210	213	205	221	211	214	215	213	9	5.360
17	209	232	210	206	232	232	228	220	227	220	204	212	219	11	5.391
16	230	226	224	236	231	231	230	233	249	238	241	263	236	11	5.464
15	263	247	235	242	234	234	252	240	251	215	236	246	241	12	5.486
14	265	254	265	246	260	260	236	244	264	258	274	248	256	11	5.546
13	246	266	278	281	251	251	243	262	284	260	279	239	262	16	5.567
12	290	276	278	270	285	285	264	288	271	286	294	262	279	11	5.632
11	290	271	284	300	282	282	259	311	293	281	299	296	287	14	5.661
10	284	302	304	317	308	308	288	294	307	271	305	285	298	13	5.696
9	324	301	330	308	292	292	287	300	331	310	339	295	309	18	5.734
8	304	320	327	313	299	299	290	324	338	307	350	335	317	18	5.759
7	334	331	348	314	337	337	313	340	311	298	342	315	327	16	5.789
6	347	351	362	350	310	310	316	326	341	339	346	360	338	18	5.824
5	335	328	368	349	357	357	332	365	363	350	348	345	350	13	5.857
4	377	349	380	374	334	334	314	326	339	363	400	368	355	26	5.872
3	369	366	387	379	402	402	312	348	377	373	365	389	372	25	5.920
2	402	367	369	382	374	374	363	383	420	368	376	389	381	16	5.942
1	400	433	399	405	388	388	331	408	416	390	408	390	396	24	5.982

Table 3.5: The No. Of molecules in bottom 27 rows of Argon gas for 12 runs.

						Runs									
Row No.	1	2	3	4	5	6	7	8	9	10	11	12	Avg 12 runs	Std Dev	Ln Avg
27	111	121	127	101	120	112	106	100	119	99	130	120	114	11	4.735
26	123	109	131	116	125	119	119	109	141	136	144	130	125	12	4.830
25	134	118	136	141	129	142	140	118	140	142	135	132	134	9	4.897
24	132	149	129	124	137	149	161	123	144	155	142	153	142	13	4.952
23	160	141	149	150	143	156	171	142	161	169	156	156	155	10	5.040
22	168	154	186	169	171	190	170	151	168	154	175	167	169	12	5.127
21	198	169	181	166	147	169	186	140	153	180	178	170	170	17	5.134
20	174	196	197	188	189	175	207	162	170	190	216	221	190	18	5.249
19	210	192	205	197	206	182	204	197	199	199	204	201	200	7	5.297
18	212	207	221	237	201	205	210	198	236	208	223	229	216	13	5.373
17	231	227	206	245	231	235	231	247	227	244	240	223	232	11	5.448
16	252	247	255	243	236	217	231	232	240	235	270	242	242	14	5.488
15	256	227	287	265	263	236	263	262	248	279	244	266	258	17	5.553
14	264	287	249	252	271	277	246	284	284	265	273	273	269	14	5.594
13	275	278	300	293	295	289	269	277	258	304	276	282	283	14	5.645
12	305	289	299	301	312	313	307	287	316	286	282	303	300	12	5.704
11	319	301	313	302	313	327	290	312	285	328	307	298	308	13	5.730
10	324	300	329	329	329	348	336	334	331	335	312	303	326	14	5.786
9	331	335	337	350	331	352	335	369	347	338	345	347	343	11	5.838
8	344	362	348	353	377	308	360	364	360	350	333	339	350	18	5.857
7	349	363	376	372	367	368	385	371	372	401	369	358	371	13	5.916
6	351	373	361	394	383	403	352	407	389	363	374	385	378	19	5.935
5	399	406	378	370	420	391	390	410	388	390	401	414	396	15	5.982
4	423	432	390	427	411	408	418	410	413	383	388	402	409	16	6.013
3	433	451	431	429	464	400	439	422	411	437	417	433	431	17	6.065
2	441	450	436	433	451	450	427	461	436	417	435	436	439	12	6.085
1	429	497	433	460	472	445	442	445	443	453	462	439	452	19	6.113

Table 3.6: The No. of molecules in bottom 27 rows of Carbon dioxide for 12 runs.

						Runs									
Row No.	1	2	3	4	5	6	7	8	9	10	11	12	Avg 12 runs	Std dev	ln avg
27	104	118	116	108	129	106	121	112	86	76	118	96	108	15	4.677
26	102	107	125	118	104	103	109	102	101	101	131	104	109	10	4.691
25	130	119	142	116	126	119	124	135	129	127	115	126	126	8	4.834
24	153	147	144	128	141	139	128	130	122	133	141	130	136	9	4.915
23	143	142	158	123	153	139	146	145	142	162	152	139	145	10	4.979
22	170	150	152	162	149	161	160	154	129	135	151	152	152	11	5.024
21	177	182	176	152	183	197	156	163	177	168	186	174	174	13	5.160
20	180	173	196	195	174	181	184	179	175	192	183	187	183	8	5.211
19	212	182	212	168	200	203	191	184	195	197	196	208	196	13	5.276
18	201	203	224	217	209	216	202	193	206	199	201	210	207	9	5.332
17	211	220	232	263	222	229	238	226	237	229	221	225	229	13	5.436
16	248	256	244	231	273	253	228	231	243	246	229	240	244	13	5.495
15	250	249	263	245	242	252	273	256	253	267	229	269	254	13	5.537
14	275	271	254	252	258	294	297	254	273	290	276	272	272	16	5.606
13	295	302	279	271	295	298	278	289	286	305	301	292	291	11	5.673
12	296	311	309	329	294	316	302	328	302	306	305	315	309	11	5.735
11	324	321	309	368	321	328	324	304	308	323	332	336	325	17	5.783
10	334	358	354	338	339	348	339	348	332	359	340	344	344	9	5.842
9	344	334	366	355	332	349	340	365	344	357	323	341	346	13	5.846
8	381	374	359	378	383	373	361	368	376	368	379	374	373	8	5.921
7	382	395	383	366	388	383	414	381	386	362	344	388	381	18	5.943
6	374	365	391	380	401	400	398	428	395	408	388	374	392	17	5.971
5	407	408	404	432	407	420	400	425	425	426	399	417	414	11	6.026
4	438	435	406	460	411	433	423	425	444	449	441	426	433	15	6.070
3	449	424	455	459	481	434	443	440	458	468	433	440	449	16	6.106
2	456	468	438	477	445	447	451	457	453	452	445	457	454	11	6.118
1	456	476	472	461	465	474	459	482	504	486	483	467	474	14	6.161

Regression statistics

Gas	M.W g/mol	$G_A(A)$	P_m	Slope	Slope error(±)	Intercept	Intercept error(±)	R^2	Std. error	Fisher	Row Eqv (meter)
N_2	28	0.100	1.000	-0.034	0.001	5.930	0.014	0.984	0.035	1559	306±9
O_2	32	0.114	0.935	-0.041	0.001	6.079	0.016	0.985	0.040	1726	323±8
Ar	39.95	0.143	0.837	-0.053	0.001	6.270	0.023	0.982	0.058	1356	335±6
CO_2	44	0.157	0.798	-0.058	0.002	6.346	0.028	0.977	0.072	1074	333±11

Table 3.7: Regression statistics for four gases for the bottom 27 rows with 12 runs at $G_A(N_2) = 0.100$, Pm = 1.000, $G_A(O_2) = 0.114$, Pm = 0.935, $G_A(Ar) = 0.143$, Pm = 0.837, $G_A(CO_2) = 0.157$, Pm = 0.798

Gas	M.W g/mol	$G_A(A)$	P_m	Slope	Slope error(±)	Intercept	Intercept error(±)	R^2	Std. error	Fisher	Row Eqv (meter)
N_2	28	0.200	1.000	-0.055	0.002	6.329	0.032	0.967	0.082	732	492±18
O_2	32	0.229	0.935	-0.077	0.003	6.583	0.051	0.959	0.128	587	606±24
Ar	39.95	0.285	0.837	-0.097	0.004	6.788	0.069	0.952	0.176	497	614±25
CO_2	44	0.314	0.798	-0.111	0.005	6.917	0.084	0.946	0.213	441	638±28

Table 3.8: Regression statistics for the four gases for the bottom 27 rows with 12 runs at $G_A(N_2) = 0.200$, Pm = 1.000, $G_A(O_2) = 0.229$, Pm = 0.935, $G_A(Ar) = 0.285$, Pm = 0.837 $G_A(CO_2) = 0.314$, Pm = 0.798

height ⟶	1m	2m	3m	4m
(Ar/N_2) ratio Barometric Eq.	-47.3 meg	-94.5 meg	-141.7 meg	-188.9 meg
(Ar/N_2) ratio CA results	-47.2 meg	-94.2meg	-141.3 meg	-188.4 meg

Table 3.9: Summary of calculations for barometric equation and CA results

Figures

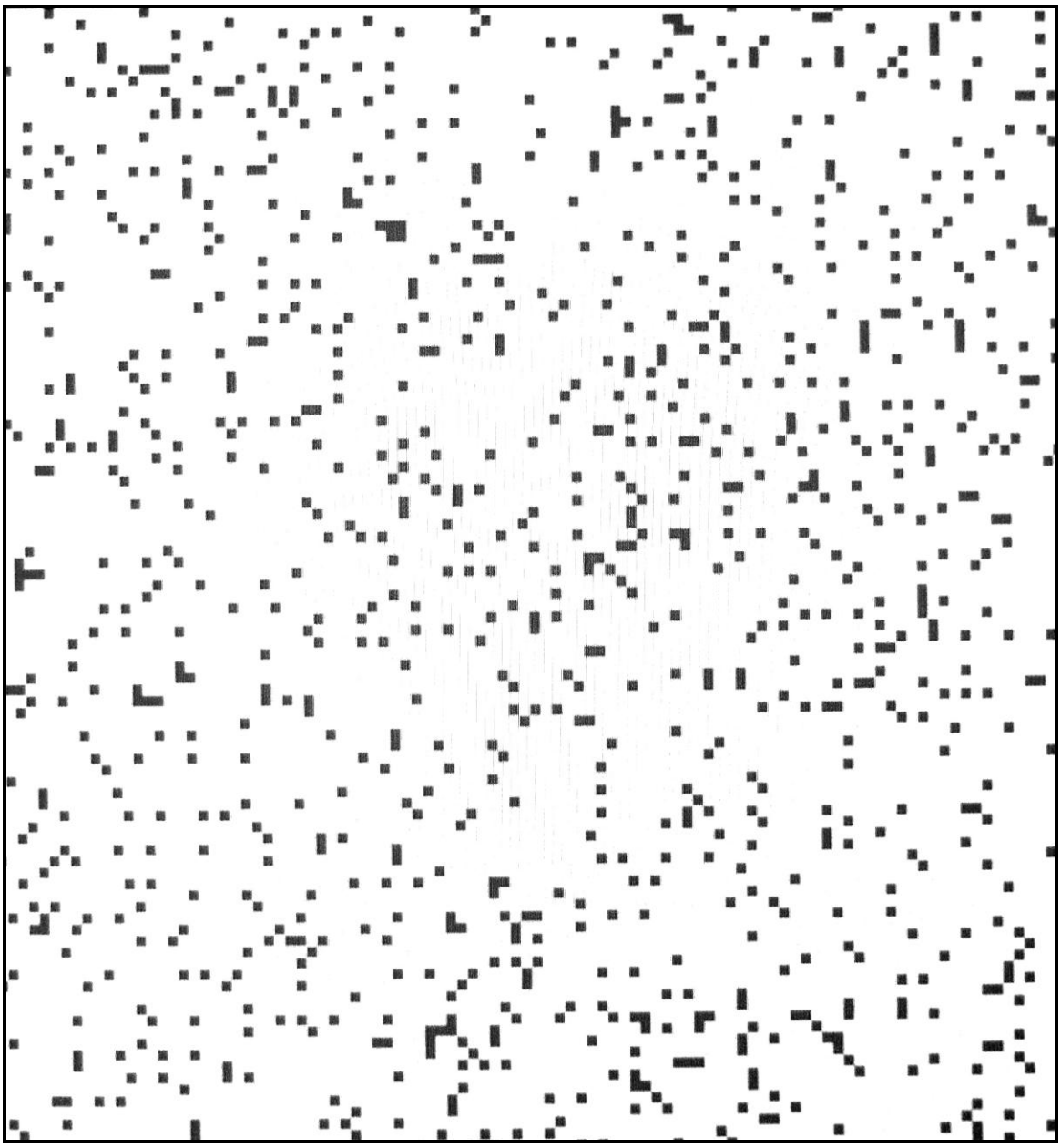


Figure 3.1: A snapshot at a very early stage of the simulation showing random placement of ingredients on the 100 x 100 grid.

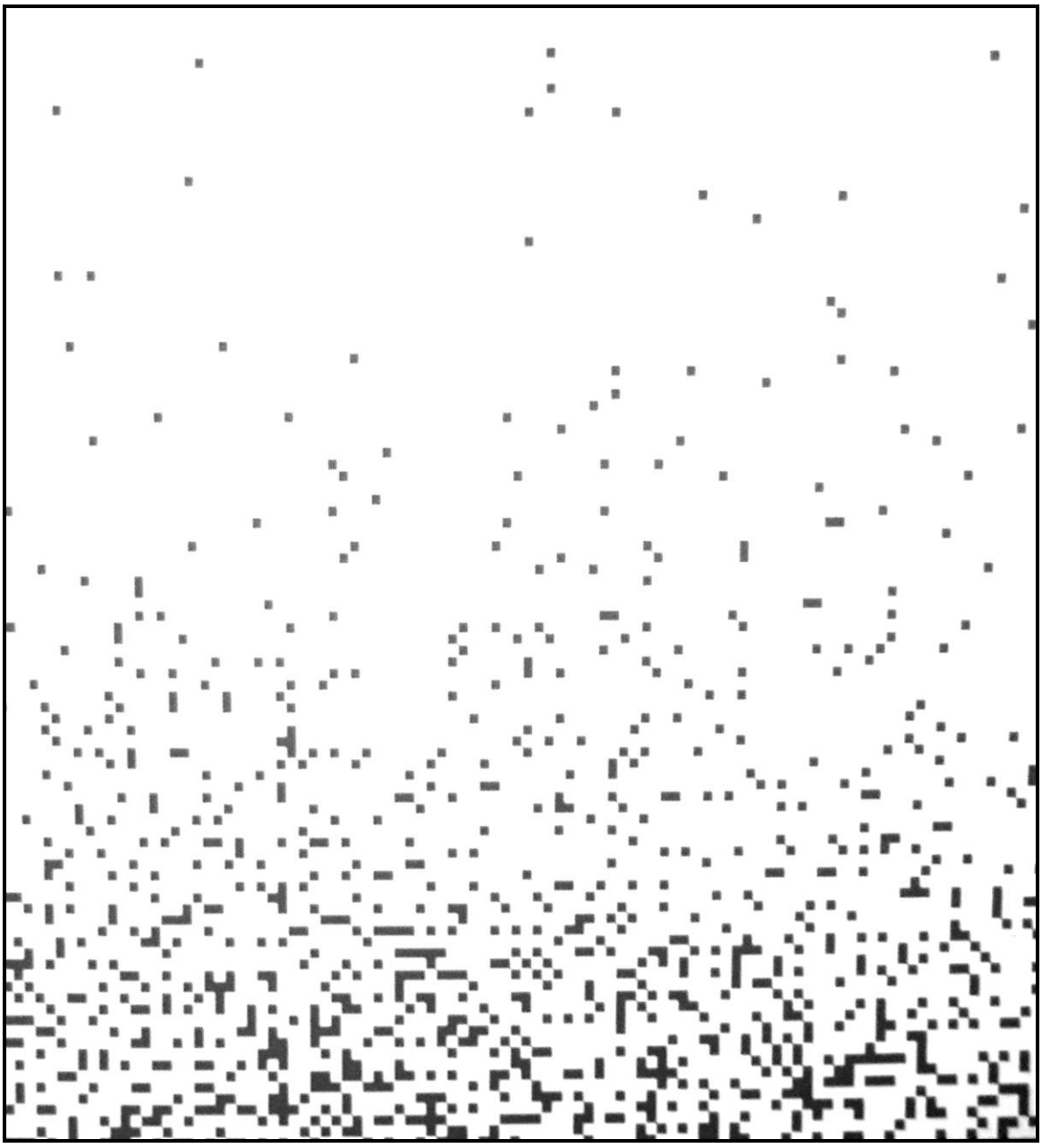


Figure 3.2: A snapshot at a later stage, showing the tendency of the ingredients to cluster near the bottom of the 100 x 100 grid

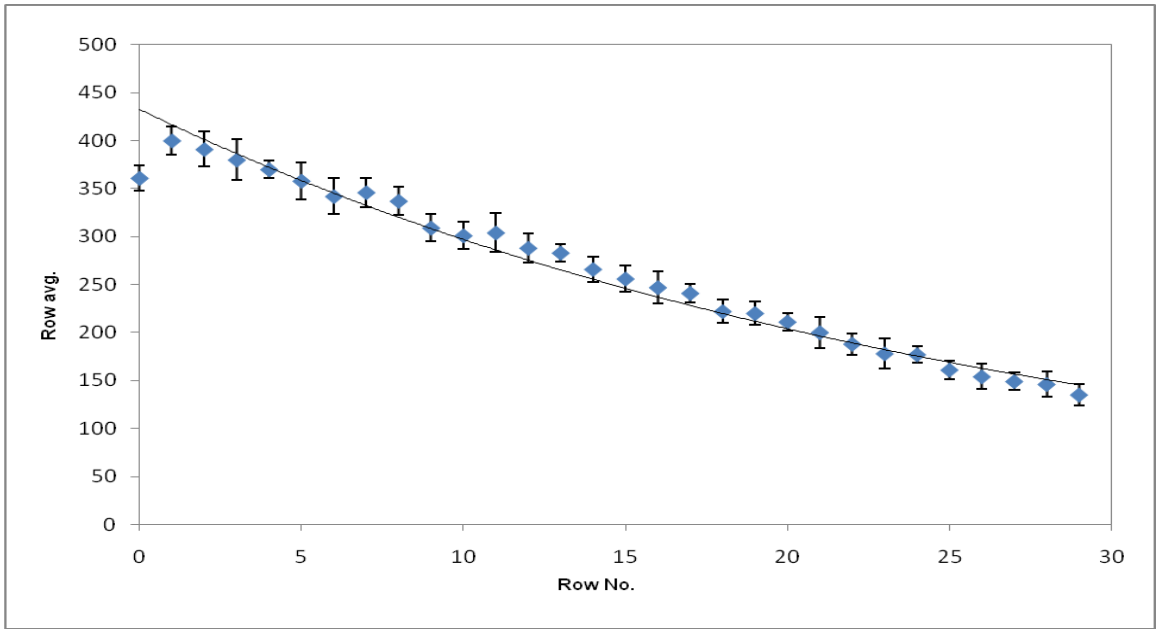


Figure 3.3: Plot of row average counts number vs. row number showing the exponential fall with respect to height and the errors associated with data points for nitrogen gas for the bottom 30 rows.

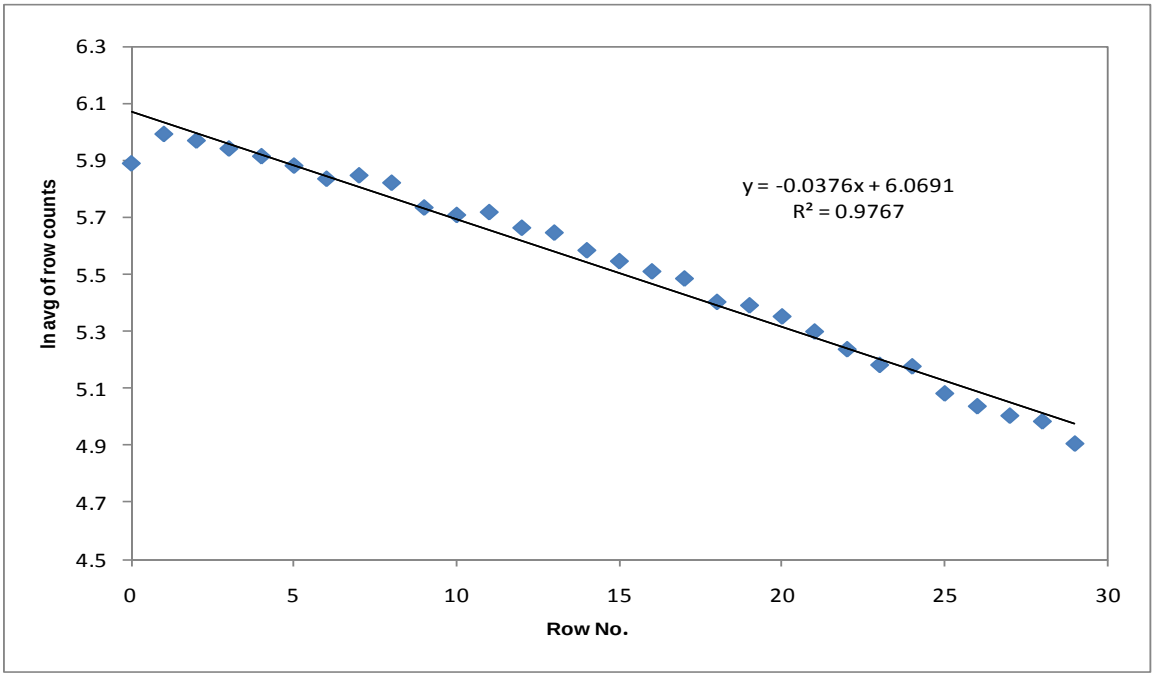


Figure 3.4: Plot of natural log of the average row counts vs. row No. for 12 runs for nitrogen gas for the bottom 30 rows.

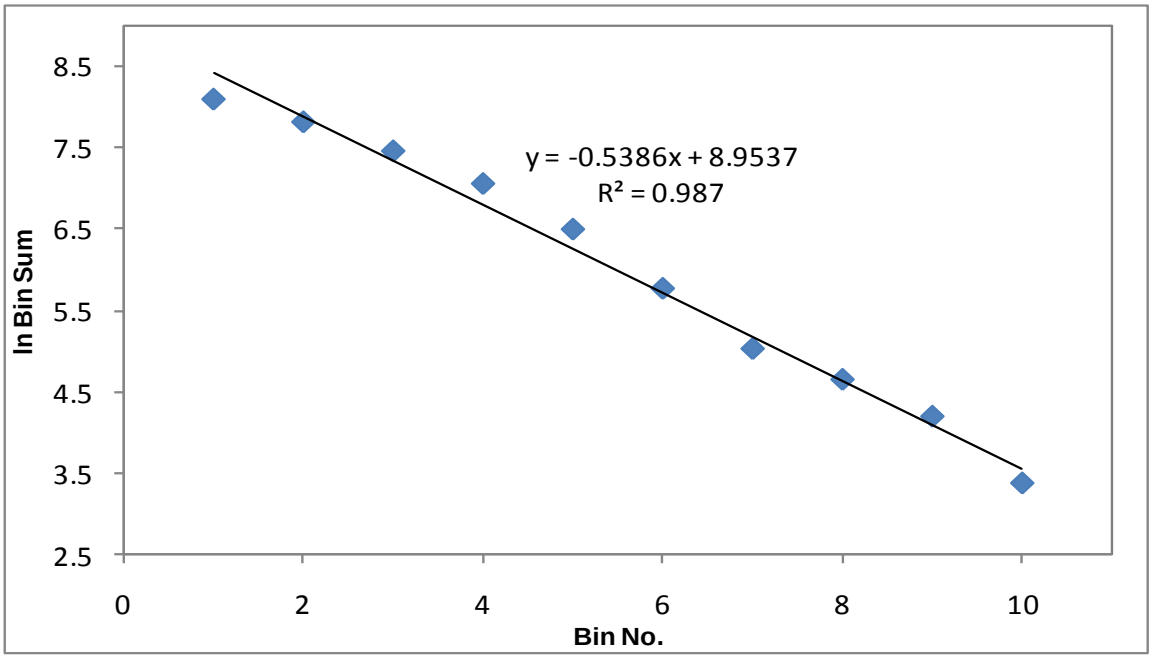


Figure 3.5: Binned Results for nitrogen gas

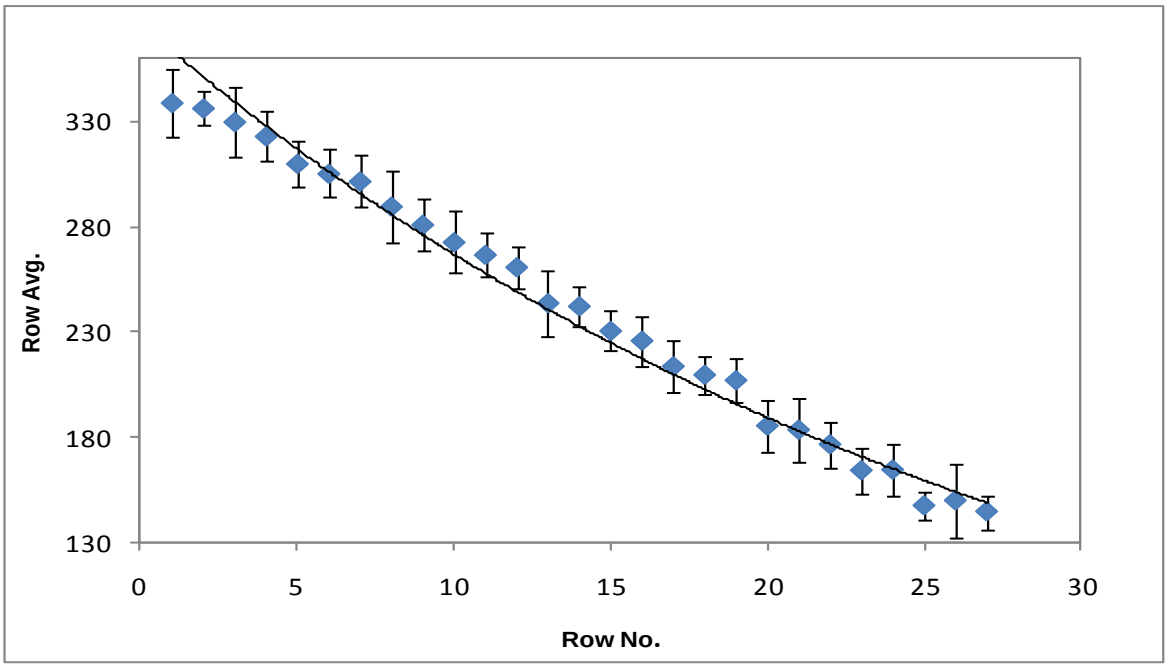


Figure 3.6: Plot of row average vs. row number showing the exponential fall with respect to height and the errors associated with data points for nitrogen gas for bottom 27 rows for nitrogen gas.

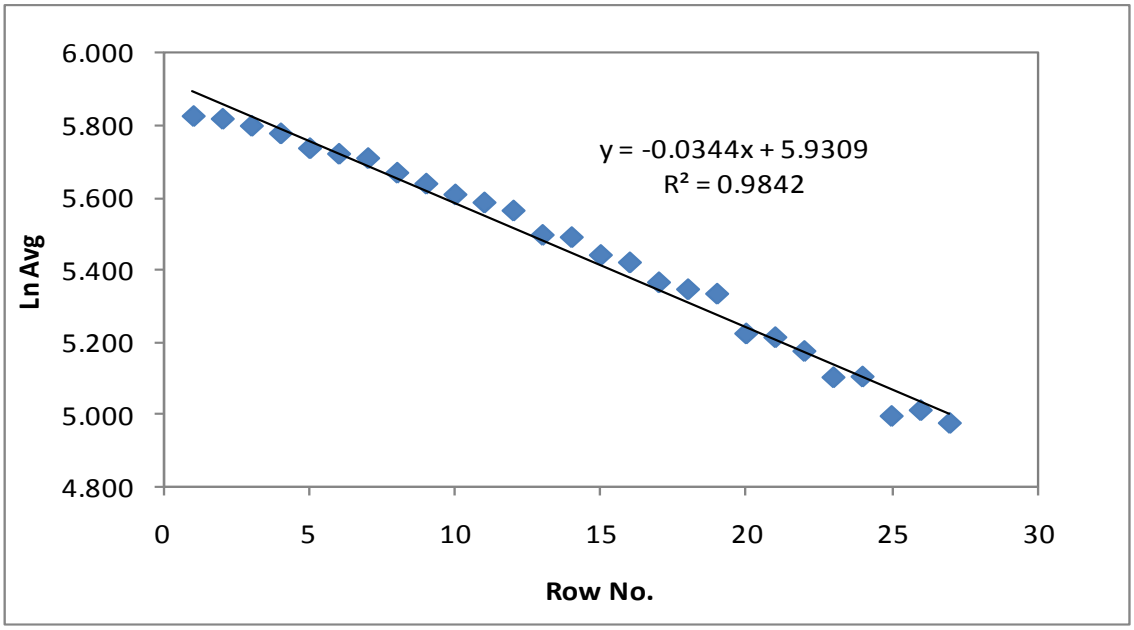


Figure 3.7: Plot of natural log of the average row counts vs. row No. for 12 runs for nitrogen gas for the bottom 27 rows

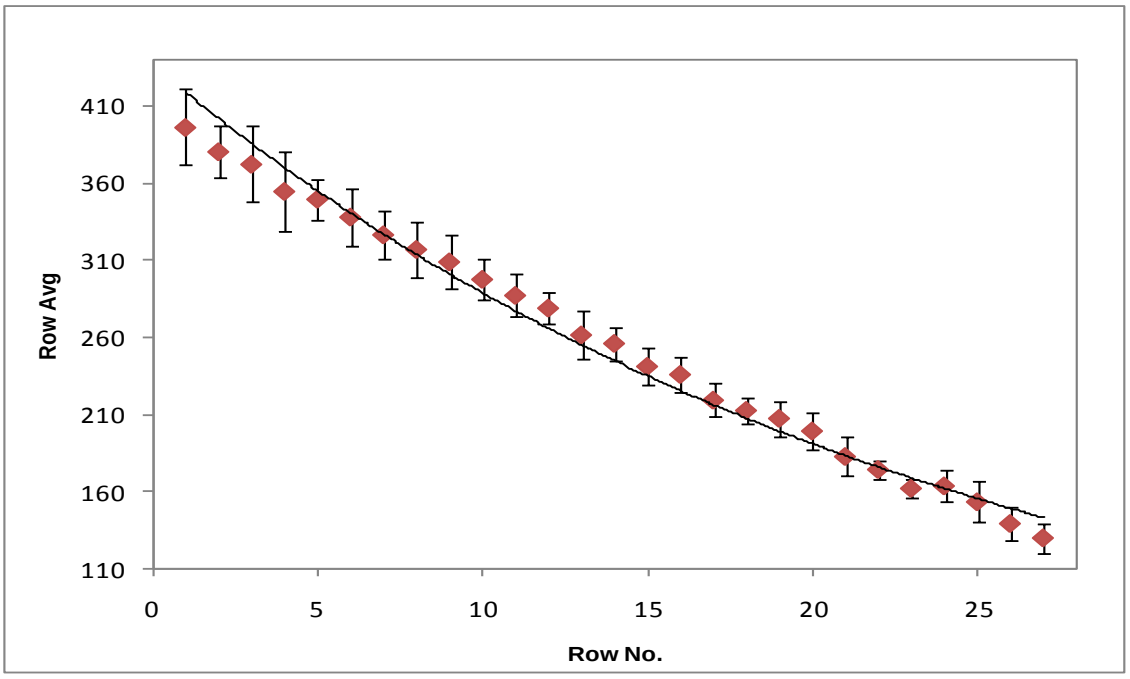


Figure 3.8: Plot of row average counts number vs. row number showing the exponential fall with respect to height and the errors associated with them for oxygen gas.

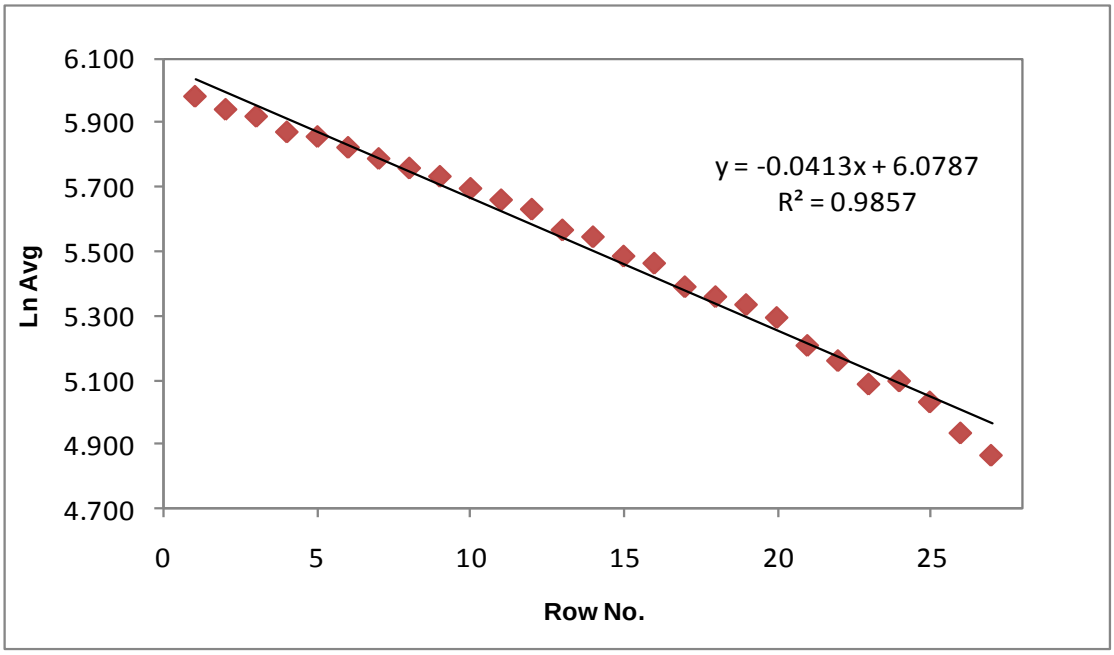


Figure 3.9: Oxygen gas behavior with respect to height for 12 runs.

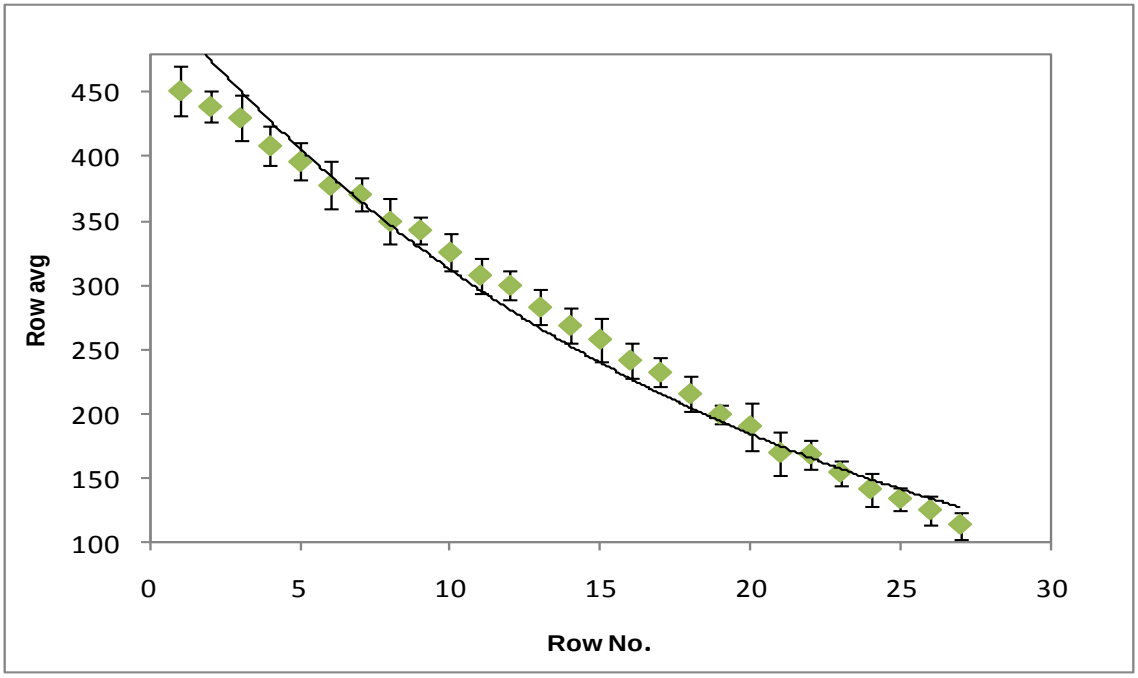


Figure 3.10: Plot of row average counts number vs. row number showing the exponential fall with respect to height and the errors associated with them for argon gas.

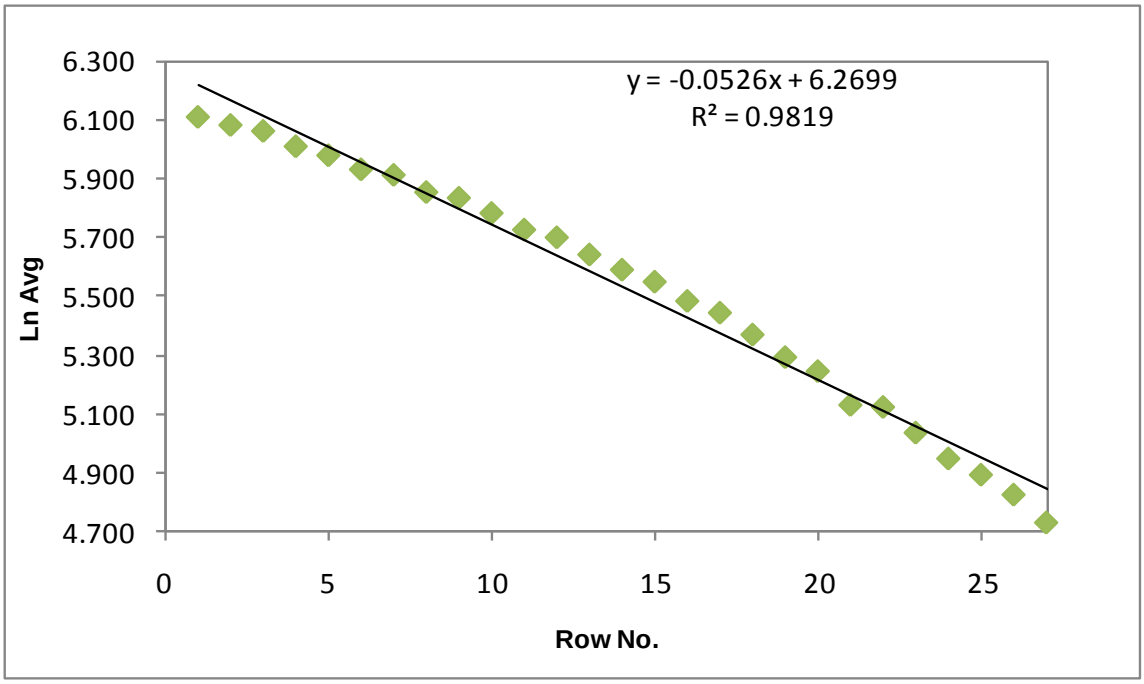


Figure 3.11: Argon Gas Behavior with respect to height for 12 runs.

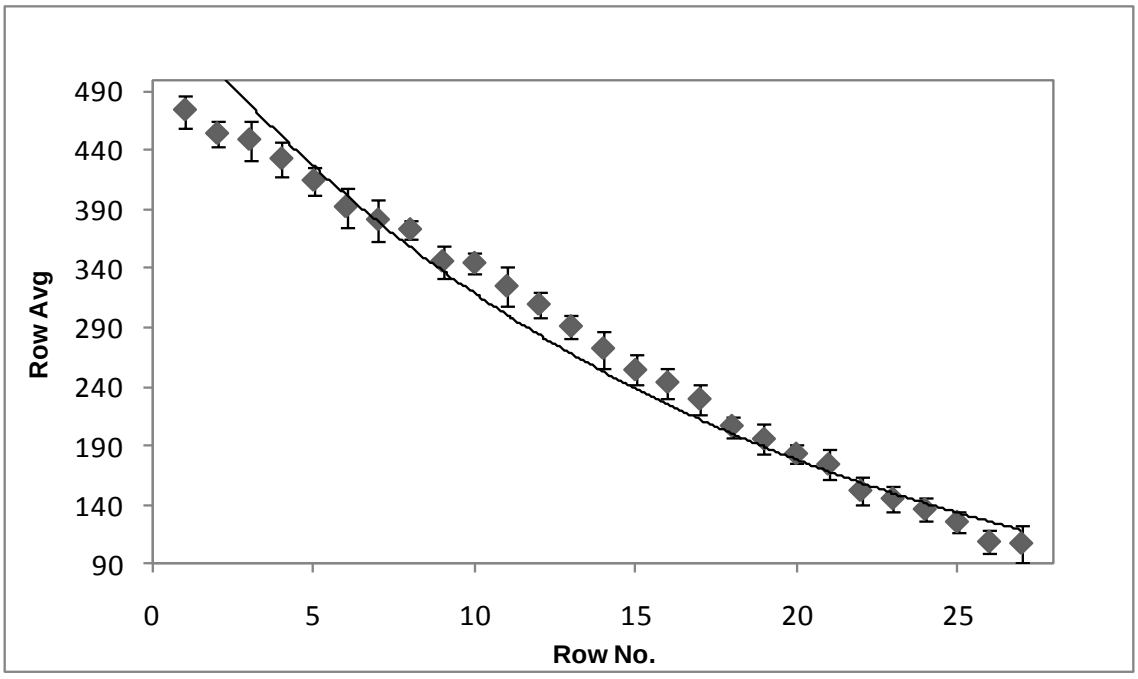


Figure 3.12: Plot of row average counts number vs. row number showing the exponential fall with respect to height and the errors associated with them for CO_2

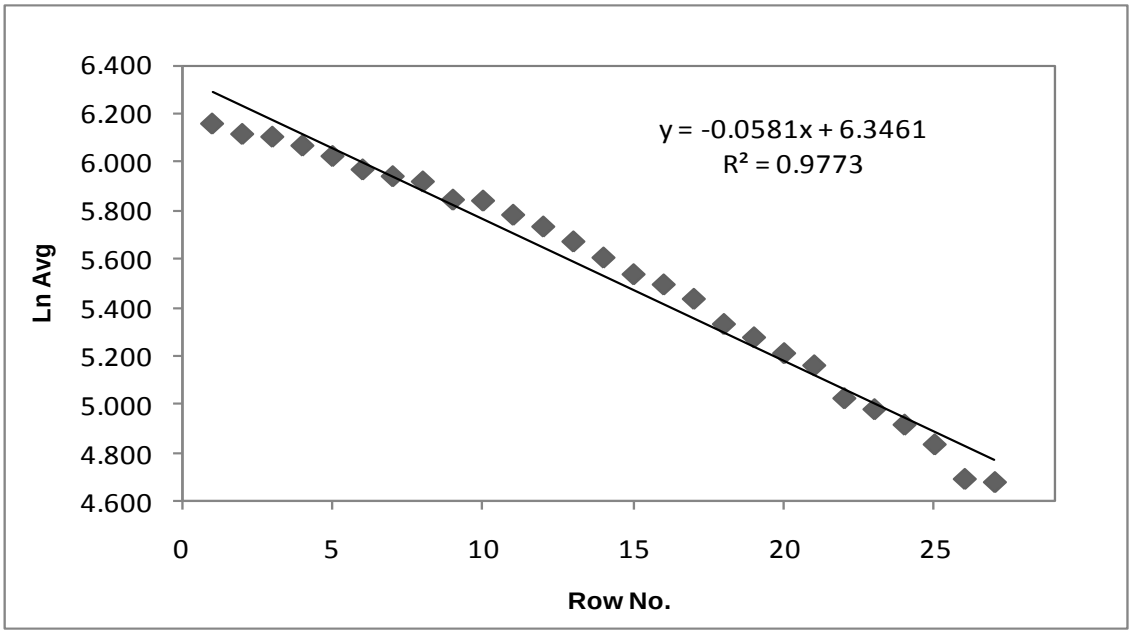


Figure 3.13: CO_2 Gas Behavior with respect to height for 12 runs.

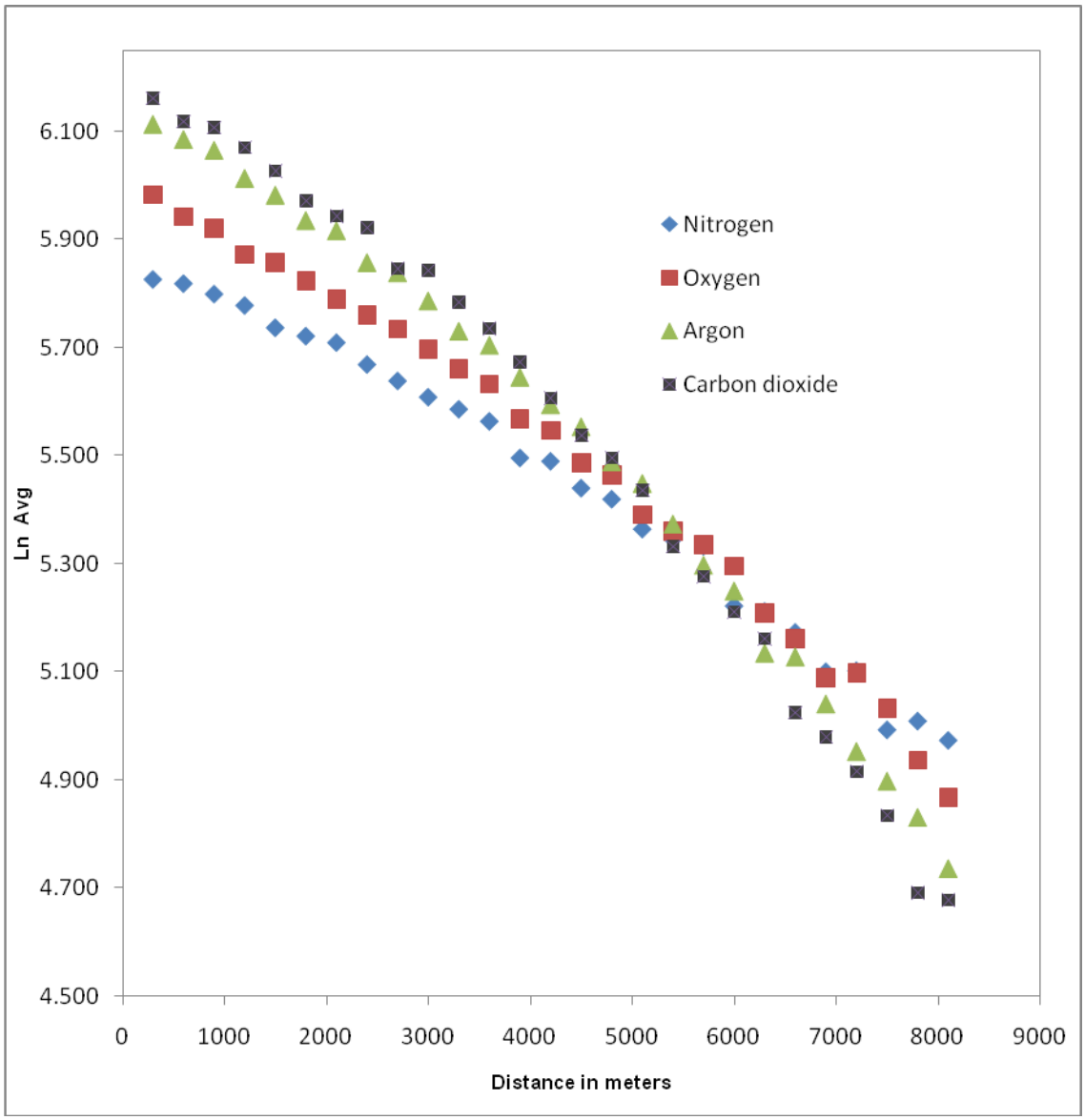


Figure 3.14: Plot showing the distribution of all four gases when the simulation was ran at $G_A(N_2) = 0.100$, Pm = 1.000, $G_A(O_2) = 0.114$, Pm = 0.935

$G_A(Ar) = 0.143$, Pm = 0.837

$G_A(CO_2) = 0.157$, Pm = 0.798

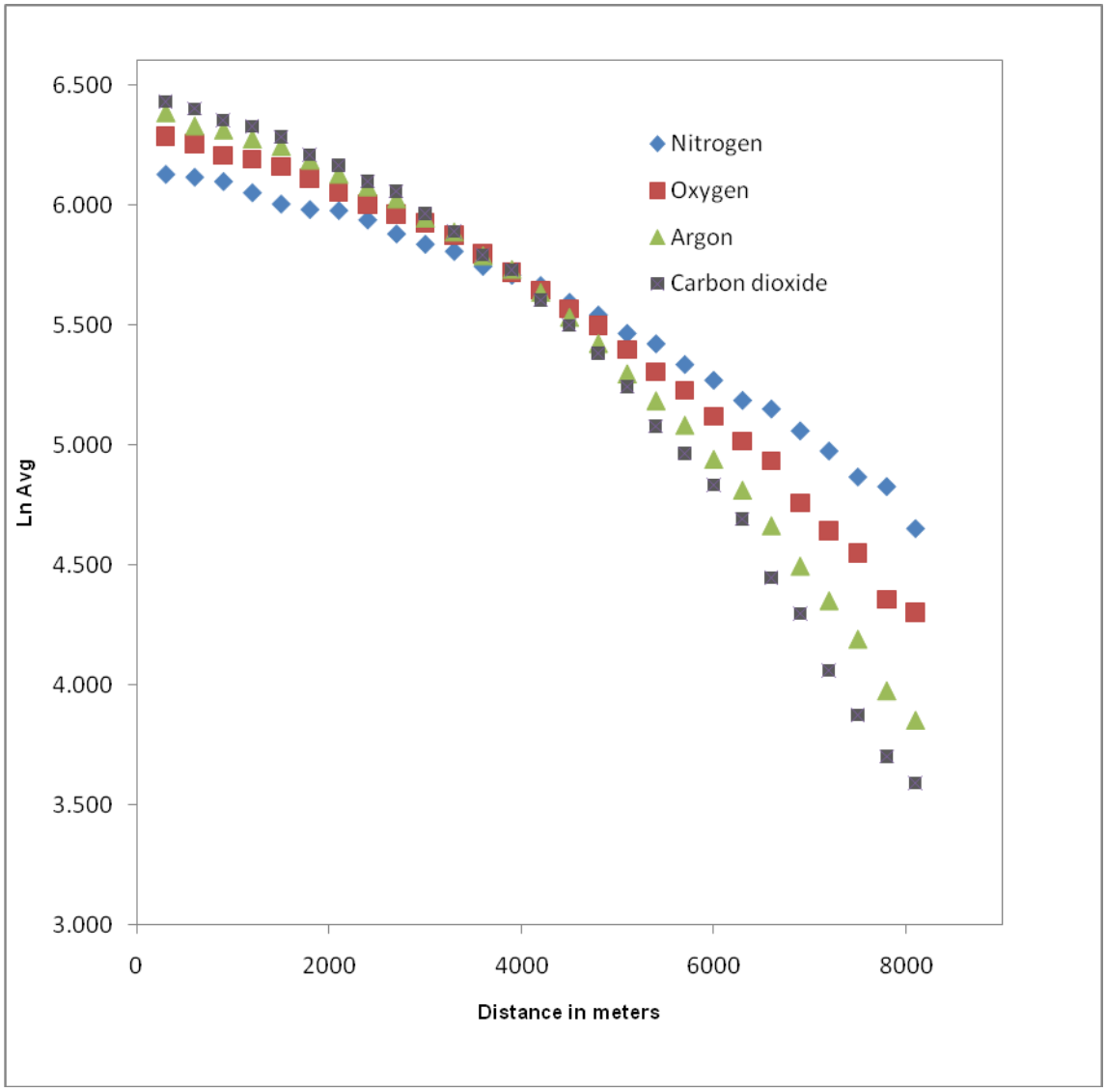


Figure 3.15: Plot showing the distribution of all four gases when the simulation was ran at

$G_A(N_2) = 0.200$, Pm = 1.000

$G_A(O_2) = 0.229$, Pm = 0.935

$G_A(Ar) = 0.285$, Pm = 0.837

$G_A(CO_2) = 0.314$, Pm = 0.798

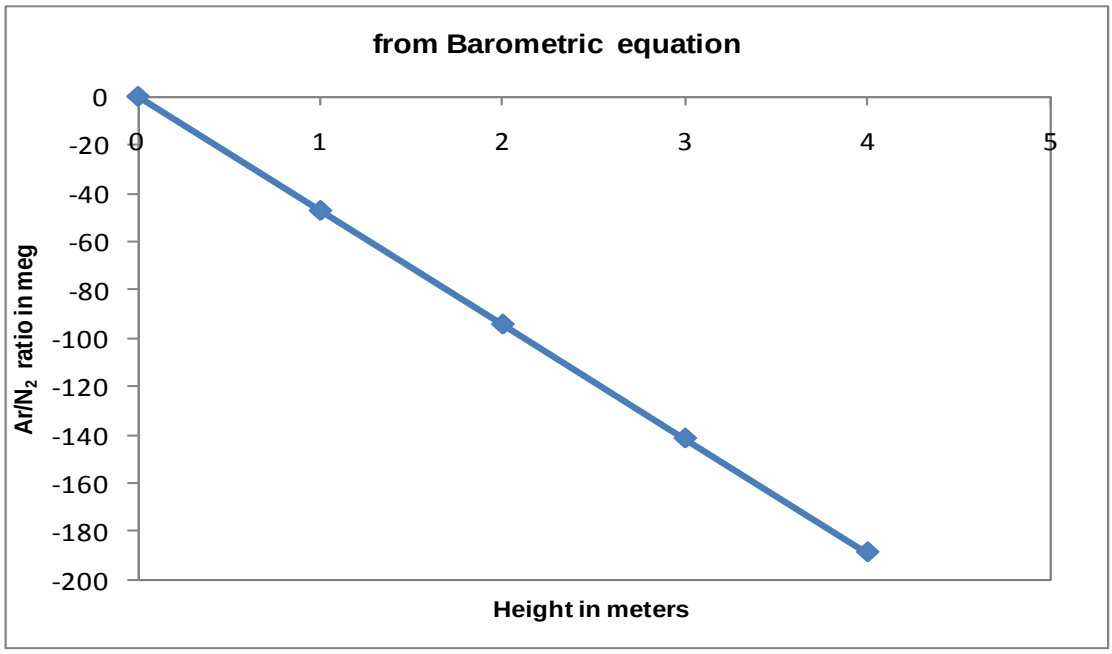


Figure 3.16: Plot showing the Ar/N_2 ratios at different heights using the barometric equation

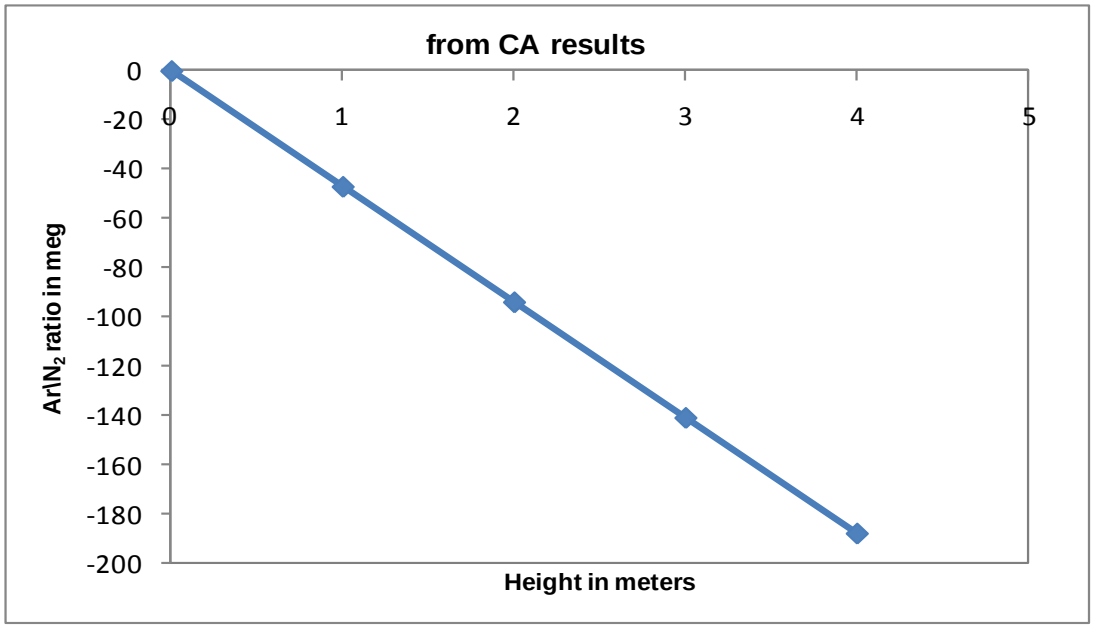


Figure 3.17: Plot showing the Ar/N_2 ratios at different heights using the CA results

3.7 References:

[1] Dalton, J. *Philos. Trans. R. Soc. London.* **1826**, *116*, 174.

[2] Adachi, Y.; Kawamura, K.; Armi, L.; Keeling, R. F. Diffusive Separation of the Lower Atmosphere. *Science.* **2006**, *311*, 1429.

[3] Lide, D.R., ed.; CRC Handbook of Chemistry and Physics, 89th ed. 2008-9; pp. 14-19.

[4] Alberty, R. A.; Silbey, R. J. *Physical Chemistry*, 2nd ed.; John Wiley: New York, 1997; pp 22.

[5] Seybold, P. G.; O'Malley, M. J,; Kier, L. B,; Cheng, C.-K. Cellular Automata Simulations of Vapor-Liquid Equilibria. *Australian J. Chem*. **2006,** *59*, 865-868

[6] J. von Neumann, in *Theory of Self-Replicating Automata*, A. Burks, ed. University of Illinois Press, Urbana, 1966.

[7] S. M. Ulam, Random processes and transformations. *Proc. Int. Congr. Math*. 1952 (held in 1950), 2, 264

[8] S. M. Ulam, *Adventures of a Mathematician*. Charles Scribner's Sons, New York, 1976.

[9] Gardner, M. MATHEMATICAL GAMES The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*. **1970**, *223*, 120-123.

[10] Kier, L. B,; Seybold, P. G. Cellular Automata Modelling of Complex Biochemical Systems, Springer: Heidelberg, 2009; pp 848-865.

[11] C.K. Cheng, L .B. Kier, and P. G. Seybold, *CASim Version 1.0*, Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond ,VA 23284,USA, **2005**