

TÍNH TOÁN HỆ SỐ VIRIAL BẬC HAI CỦA CÁC KHÍ Cl_2 , N_2 , CO VÀ Ar KẾT HỢP PHƯƠNG TRÌNH TRẠNG THÁI VIRIAL VÀ MÔ HÌNH ĐA BIẾN

Nguyễn Thành Được^{1,4}, Trần Dương², Phạm Văn Tất^{3,*}

¹Khoa Hóa học, Trường Đại học Khoa học, Đại học Huế

²Khoa Hóa học, Trường Đại học Sư phạm, Đại học Huế

³Bộ môn Kỹ thuật Môi trường, trường Đại học Hoa Sen, TP. Hồ Chí Minh

⁴Khoa Khoa học Tự nhiên, Đại học Thủ Dầu Một

*Email: vantat@gmail.com

Ngày nhận bài: 02/10/2018; ngày hoàn thành phản biện: 5/11/2018; ngày duyệt đăng: 10/12/2018

TÓM TẮT

Thành phần chất khí trong công nghiệp hóa học thường chứa các khí argon, nito, cacbon monoxit và clo. Sự phát thải trực tiếp các khí này vào khí quyển cần phải được hạn chế do tác động của chúng với môi trường. Để tách và lưu trữ chúng, chúng ta cần phải biết đầy đủ tính chất tương tác giữa các phân tử. Hệ số virial bậc hai đặc trưng cho tính chất tương tác phân tử. Bài báo sử dụng mô hình mạng thần kinh kết hợp phân tích thành phần chính ANN-PCA I(5)-HL(6)-O(3) với giá trị sai số MSE là 0,0069695 để dự đoán các hệ số a, b và c trong phương trình trạng thái virial dựa vào các tính chất tới hạn của các chất khí. Hệ số virial bậc hai được xác định một cách chính xác sử dụng các hệ số a, b và c được dự đoán. Hệ số virial tính toán cũng rất gần với các hệ số virial dự đoán từ của phương trình trạng thái Deiters và phù hợp với dữ liệu thực nghiệm.

Từ khóa: Mạng thần kinh ANN, hệ số virial, phương trình trạng thái virial, phân tích thành phần chính.

I. GIỚI THIỆU

Mô phỏng máy tính đã trở thành những công cụ không thể thiếu để nghiên cứu những chất lỏng và hỗn hợp chất lỏng. Kỹ thuật mô phỏng có thể tính toán các tính chất cấu trúc và nhiệt động học cũng như sự chuyển động của phân tử, hệ số virial là một tham số quan trọng về sự tương tác phân tử để đánh giá chất lượng tham số hàm thế tối ưu [[1]]. Kỹ thuật mô phỏng Monte Carlo cũng như động học phân tử, không thể thực hiện nếu không có thông tin đầu vào [[3]]. Thông thường có thể sử dụng một hàm thế đơn giản [[2]] như hàm thế cặp Lennard-Jones, có thể được sử dụng để khớp

với dữ liệu thực nghiệm tạo ra các tham số, và sau đó sử dụng để tiến hành mô phỏng [[4]]. Mô phỏng như vậy cho dự báo không lâu hơn, bởi vì quá trình mô phỏng cần thông tin đi vào cùng một tính chất giống nhau. Điều này cũng có thể là một hạn chế, cụ thể là nếu dữ liệu thực nghiệm không đủ [[5]].

Các kỹ thuật mô phỏng toàn cục đã được thực hiện cho các loại khí hiếm, mà hiện nay có thể được sử dụng để dự đoán cân bằng pha lỏng hơi mà không đòi hỏi dữ liệu thực nghiệm, cho kết quả với độ chính xác tương đương trong khoảng không chắc chắn của thực nghiệm [[8],[9],[11]]. Một trong những nỗ lực đầu tiên đạt được độ chính xác gần với thực nghiệm là các nghiên cứu của Deiters, Hloucha và Leonhard [[5],[6]] đối với khí neon. Sự tiến xa hơn trong các nỗ lực mô phỏng toàn cục của các chất khí hiếm đã công bố bởi nhóm Eggenberger và Huber [[7]-[10]] và Sandler [[11]]. Sử dụng một dạng hàm thế cho các thể khuếch tán của argon và krypton được đưa ra bởi Korona [[12]]; Nasrabad và Deiters thậm chí còn dự đoán cân bằng pha lỏng hơi ở áp suất cao của hỗn hợp khí hiếm [[13]]. Các hàm thế cặp của hỗn hợp các khí hiếm khác đã được López Cacheiro công bố [[14],[7]], nhưng vẫn chưa được sử dụng cho dự đoán cân bằng pha.

Phát triển các hàm thế cặp *ab initio* cho các phân tử phức tạp hơn nhiều do độ tự do của góc giữa các phân tử thay đổi. Đối với một số phân tử đơn giản những hàm thế này đã được xây dựng bởi Leonhard và Deiters [[5]] khi sử dụng hàm thế Morse với 5 vị trí, để biểu diễn tương tác cặp của các phân tử nitơ và đã dự đoán được áp suất và tỷ trọng hơi. Bock cũng đã đưa ra một thế cặp 5 vị trí cho phân tử CO_2 [[16]]; Ngoài ra các hàm thế tương tác bậc 2 có thể tiếp tục được hiệu chỉnh lượng tử bậc nhất cho các hệ số virial bậc 2 được Pack phát triển [[15]]. Naicker đã sử dụng lý thuyết nhiễu loạn đối xứng (SAPT) để phát triển một thế cặp tương tác 3 vị trí cho phân tử HCl [[11]], dựa trên cơ sở hàm thế Korona và một hàm thế Morse; nhóm của Naicker đã dự đoán thành công cân bằng pha lỏng hơi của HCl bằng mô phỏng Monte Carlo [[4]-[19],[20]].

Việc tìm kiếm các phương pháp khác nhau để tính toán các hệ số virial là cần thiết cho việc xây dựng các hàm thế tương tác phân tử và các kỹ thuật mô phỏng. Trong công trình này, chúng tôi sử dụng kết hợp kỹ thuật đa biến với mạng thần kinh nhân tạo để dự đoán các tham số của phương trình trạng thái virial từ các tính chất tới hạn của các hợp chất. Các kiến trúc của mạng thần kinh nhân tạo được xây dựng từ phương pháp phân tích thành phần chính và sử dụng để dự đoán các tham số trong phương trình trạng thái virial. Những kết quả nhận được từ các mạng thần kinh nhân tạo so sánh với dữ liệu thực nghiệm.

II. PHƯƠNG PHÁP TÍNH TOÁN

1. Phương trình trạng thái virial

Trong khí quyển, các khí luôn kèm theo các tính chất nhiệt động học quan trọng liên quan đến trạng thái của chúng và hệ số virial là một tham số thể hiện khả năng tương tác của chúng. Các tính chất nhiệt động bao gồm áp suất tới hạn P_c , thể tích mol tới hạn $V_{c,m}$, và nhiệt độ tới hạn T_c , mà tại đó tỷ trọng của các pha lỏng và khí cùng tồn tại đồng nhất. Vì vậy, phương trình trạng thái của khí [[25],[26]] chính xác

$$pV_m / RT = 1 + B/V_m + C/V_m^2 + \dots \quad (1)$$

Trong đó R là hằng số khí $8,314\,510 \pm 0,000070 \text{ J K}^{-1} \text{ mol}^{-1}$; B và C là các hệ số virial bậc hai, bậc ba, Trong đó các hệ số virial bậc hai có thể được xác định bằng cách khớp các dữ liệu thực nghiệm ở các nhiệt độ khác nhau theo phương trình như sau:

$$B_2(T)/(\text{cm}^3 \text{mol}^{-1}) = a - b \exp \frac{c/K}{T} \quad (2)$$

Trong đó các tham số a , b , và c được cho ở vùng nhiệt độ $T = 80\text{K}$ đến 1000K cho khí argon và 90K đến 573K cho CO.

2. Phương trình trạng thái

Deiters đã xây dựng một phương trình trạng thái D1-EOS dựa trên lý thuyết chuỗi cầu cứng nhiễu loạn với mục đích sử dụng phương trình D1-EOS để tính nhiệt độ tới hạn, áp suất tới hạn và tỷ trọng của các thành phần tinh khiết [[22],[23]]. Deiters cũng đề xuất một phần mở rộng cho hỗn hợp bậc hai. Vì phương trình D1-EOS ban đầu được đề xuất bởi Deiters không thể lấy tích phân theo phép toán giải tích, Deiters đã thay một phần bằng một chuỗi đa thức. Ba tham số thành phần tinh khiết cần thiết cho phương trình trạng thái Deiters là: a , b và c . Trong nghiên cứu này, chúng tôi sử dụng phương trình trạng thái Deiters để tính các hệ số virial cho các khí tinh khiết Ar, N_2 , CO và Cl_2 , được biểu thị như sau:

$$z = 1 + cc_0 \frac{4\xi - 2\xi^2}{(1-\xi)^3} - \frac{ab\rho}{c^2T} (h_0\gamma^2\psi(\tilde{T}) + F_{i+1}) \quad \text{với} \quad F_t = \sum_{k=0}^3 \sum_{j=0}^6 \sum_{i=0}^{10} p_{ijk} (c-1)^k \tilde{T}^{-j} (b\rho)^i t \quad (3)$$

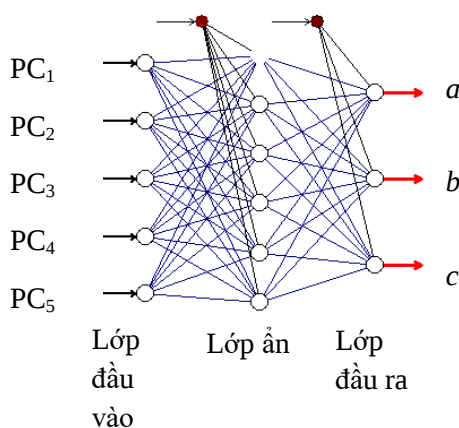
Trong đó p_{ijk} là hằng số thu được từ Deiters; $\psi(\tilde{T}) = \tilde{T}(\exp(\tilde{T}^{-1}) - 1)$; và $\gamma = 1 - 0.697816(c-1)^2$;

$$h_0 = 7.0794046, c_0 = 0.6887; \tilde{T} = \frac{ckT}{\varepsilon}; \xi = \frac{\pi\sqrt{2}}{6}b\rho$$

Các tham số trạng thái cho thành phần tinh khiết: $\frac{\varepsilon}{k} = a$

3. Xây dựng mạng thần kinh nhân tạo

Mạng thần kinh nhân tạo (ANN) [[2],[21]] được sử dụng trong nghiên cứu này là một hệ kết nối, chuyển tiếp đầy đủ và được luyện bằng một thuật toán lan truyền ngược nhanh. Nó liên quan đến hai pha: một pha chuyển tiếp trong đó thông tin đầu vào ở lớp ngoài tại các nút đầu vào được truyền tới để tính toán tín hiệu thông tin đầu ra ở lớp xuất; và một pha ngược lại trong đó cho phép sửa đổi các cường độ kết nối được thực hiện dựa trên sự khác biệt giữa tín hiệu thông tin được tính toán và quan sát tại các điểm đầu ra trên lớp xuất. Mạng lan truyền ngược sử dụng quá trình luyện có giám sát và so sánh kết quả đầu ra của nó với các kết quả mục tiêu [[2],[21]]. Các lỗi được truyền lại qua hệ để điều chỉnh các trọng số trong mỗi lớp. Trong quá trình luyện mạng, cùng một tập hợp dữ liệu được xử lý nhiều lần thì các trọng số kết nối được tinh chỉnh qua mỗi lần. Trong quá trình luyện, sai số giữa đầu ra của mô hình và kết quả mục tiêu giảm dần và mô hình luyện được tối ưu mối quan hệ giữa lớp đầu vào và lớp đầu ra. Các quy tắc học được lưu lại để quá trình luyện lặp đi lặp lại nhằm giảm thiểu sai số. Cấu trúc mạng thần kinh trong nghiên cứu này chứa một mạng gồm một lớp đầu vào, một lớp ẩn và một lớp đầu ra. Cấu trúc mạng được biểu thị trong Hình 1.



Hình 1. Cấu trúc mạng thần kinh ANN- PCA I(5)-HL(6)-O(3); kí hiệu: ○, các nơ ron ở lớp đầu vào, ẩn và đầu ra; ●: Các nút sai lệch chéo cho nơ ron lớp ẩn và lớp đầu ra.

Các tham số tối hạn quan trọng P_c , V_c , T_c , T_L và T_u được lựa chọn và các hệ số thực nghiệm trong phương trình 2 được dùng cho xây dựng mạng thần kinh nhân tạo, được thể hiện trong Bảng 1. Các tham số này cũng phụ thuộc nhiều vào đặc tính của nhau [[26],[28]]. Do đó, chúng tôi thực hiện chuyển các tham số tối hạn sang dạng các thành phần chính PC_n ($n = 1 - 5$) bằng cách sử dụng kỹ thuật phân tích thành phần chính tương ứng. Ngoài ra chúng tôi cũng chuyển các hệ số của phương trình virial

sang dạng $\log(x)$ (với $x = a, b$, hoặc c). Những tham số này được sử dụng để xây dựng mạng thần kinh nhân tạo.

Các dữ liệu trong Bảng 1, được chia thành dữ liệu luyện mạng gồm dữ liệu của các chất không bao gồm argon, nito, cacbon monoxit và clo. Nhóm kiểm tra có thể sử dụng các chất trong nhóm ngoại gồm argon, nito, cacbon monoxit và clo. Đây cũng là các chất nghiên cứu thuộc công trình này.

Bảng 1. Dữ liệu ban đầu của các tính chất tới hạn P_c , V_c , T_c , T_L và T_U , và hệ số a , b và c [[26],[28]].

Các hợp chất	P_c	V_c	T_c	T_L	T_U	a	b	c
NH ₃	11,35	72,50	405,50	273,00	573,00	44,30	23,60	766,60
CS ₂	7,90	173,00	552,00	280,00	430,00	211,00	167,10	538,70
N ₂ O	7,25	97,40	309,60	200,00	423,00	180,70	114,80	305,40
F ₂	5,22	66,00	144,30	80,00	300,00	71,40	48,00	165,00
He	0,23	57,20	5,19	7,00	150,00	114,10	98,70	3,25
H ₂	1,30	65,00	33,20	14,00	400,00	315,00	289,70	9,47
HCl	8,31	81,00	324,70	190,00	480,00	57,70	37,80	495,90
H ₂ S	8,94	98,50	373,20	278,00	493,00	47,70	30,30	632,90
Kr	5,50	91,20	209,40	110,00	700,00	189,60	148,00	145,30
Ne	2,76	41,70	44,40	44,00	973,00	81,00	63,60	30,70
NO	6,48	57,70	180,00	122,00	311,00	15,90	11,00	372,30
O ₂	5,04	73,40	154,60	90,00	400,00	152,80	117,00	108,80
SO ₂	7,88	122,00	430,80	265,00	473,00	134,40	72,50	606,50
SF ₆	3,77	199,00	318,70	200,00	525,00	422,10	281,30	273,50
UF ₆	4,66	250,00	505,80	321,00	469,00	540,50	380,90	445,00
H ₂ O	22,06	56,00	647,10	293,00	1248,00	33,00	15,20	1300,70
Xe	5,84	118,00	289,70	160,00	650,00	245,60	190,90	200,20
N ₂	3,39	89,50	126,20	75,00	700,00	185,40	141,80	88,70
Ar	4,90	74,60	150,90	80,00	1024,00	154,20	119,30	105,10
CO	3,50	93,10	132,90	90,00	573,00	202,60	154,20	94,20
Cl ₂	7,98	124,00	416,90	360,00	700,00	201,90	131,80	409,90
CO ₂	7,38	94,00	304,10	220,00	1100,00	137,60	87,70	325,70

III. KẾT QUẢ VÀ THẢO LUẬN

1. Phân tích thành phần chính

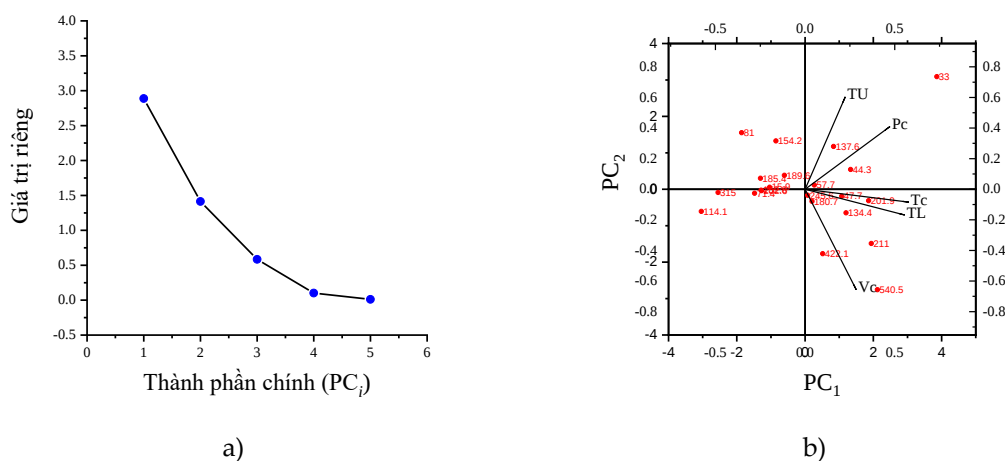
Trong ma trận tương quan ở Bảng 2, dựa vào hệ số tương quan giữa các biến số để đánh giá chiều hướng và mức độ ảnh hưởng tương quan giữa các biến số. Có nhiều giá trị tương quan lớn hơn 0,3. Phân tích thành phần chính là một công cụ lựa chọn thích hợp để loại bỏ tính cộng tính tuyến tính của các biến số. Cũng từ Bảng 2, hệ số tương quan có thể sử dụng để xây dựng vectơ riêng cho các thành phần chính.

Bảng 2. Ma trận tương quan giữa các tham số tới hạn P_c , V_c , T_c , T_L , T_U của các hợp chất

	P_c	V_c	T_c	T_L	T_U
P_c	1	-0,087	0,778	0,650	0,476
V_c	-0,087	1	0,525	0,560	-0,162
T_c	0,778	0,525	1	0,934	0,247
T_L	0,650	0,560	0,934	1	0,182
T_U	0,476	-0,162	0,247	0,182	1

Phương pháp phân tích thành phần chính làm giảm tính phức tạp của dữ liệu và có thể giải thích quy luật của tập dữ liệu đa biến lớn mang lại các cấu trúc tuyến tính cơ bản, và có thể phát hiện các mối liên quan bất thường giữa các dữ liệu.

Để xác định số lượng các thành phần chính được giữ lại, trước hết chúng ta phải chạy phân tích thành phần chính và sau đó tiến hành phân tích đánh giá dựa trên kết quả của nó.



Hình 2: Đồ thị sườn dốc lựa chọn (a) và đồ thị kép (b) có lợi để xác định số lượng thành phần chính thích hợp trong không gian con

Kết quả phân tích đánh giá cho thấy rằng sau khi có các giá trị riêng của ma trận tương quan, bốn thành phần chính đầu tiên giải thích 86% phương sai và các thành phần còn lại đều đóng góp 5% hoặc ít hơn. Vì vậy, bốn thành phần chính có thể được sử dụng để xem xét mối tương quan giữa các thành phần.

Bảng 3. Dữ liệu thành phần chính tương ứng với các tính chất tới hạn P_c , V_c , T_c , T_L và T_U , và các hệ số a , b và c chuyển đổi thành logarit [[26],[28],[29]].

Các chất	PC_1	PC_2	PC_3	PC_4	PC_5	$\log a$	$\log b$	$\log c$
Nhóm luyện								
NH_3	1,3347	0,5422	-0,9545	-0,2961	0,0328	1,6464	1,3729	2,8846
CS_2	1,9397	-1,4894	-0,1667	0,3041	-0,3271	2,3243	2,2230	2,7313
N_2O	0,2043	-0,3331	-0,6204	-0,0672	0,0137	2,2570	2,0599	2,4849
F_2	-1,4785	-0,1106	-0,8596	0,1898	0,0813	1,8537	1,6812	2,2175
He	-3,0407	-0,6086	-0,7406	0,0196	-0,0732	2,0573	1,9943	0,5119
H_2	-2,5470	-0,0880	-0,1006	0,1090	-0,0435	2,4983	2,4619	0,9763

HCl	0,2701	0,1141	-0,7225	-0,0041	-0,0626	1,7612	1,5775	2,6954
H ₂ S	1,0744	-0,1913	-0,6825	-0,4012	0,0816	1,6785	1,4814	2,8013
Kr	-0,6050	0,3816	0,3889	0,1762	-0,0101	2,2778	2,1703	2,1623
Ne	-1,8598	1,5529	1,0834	-0,2686	-0,1016	1,9085	1,8035	1,4871
NO	-1,0392	0,0557	-1,0860	-0,0135	0,0989	1,2014	1,0414	2,5709
O ₂	-1,2877	-0,0265	-0,5117	0,1432	0,0672	2,1841	2,0682	2,0366
SO ₂	1,1999	-0,6486	-0,4426	-0,1473	-0,1655	2,1284	1,8603	2,7828
SF ₆	0,5162	-1,7685	0,9599	0,3376	0,2104	2,6254	2,4492	2,4370
UF ₆	2,1190	-2,7576	0,9847	0,1819	0,0695	2,7328	2,5808	2,6484
H ₂ O	3,8622	3,0915	-0,4272	0,6556	0,0646	1,5185	1,1818	3,1142
Xe	0,0750	-0,1647	0,3813	0,1626	-0,0432	2,3902	2,2808	2,3015
Nhóm kiểm tra ngoại								
N ₂	-1,3057	0,3018	0,6544	0,1003	0,0273	2,2681	2,1517	1,9479
Ar	-0,8545	1,3268	1,2413	0,0557	-0,0399	2,1881	2,0766	2,0216
CO	-1,2749	-0,0398	0,3125	0,0667	0,0829	2,3066	2,1881	1,9741
Cl ₂	1,8615	-0,3145	0,1232	-0,9036	0,0701	2,3051	2,1199	2,6127
CO ₂	0,8360	1,1744	1,1853	-0,4010	-0,0336	2,1386	1,9430	2,5128

Hình 2 là đồ thị sườn dốc trực quan cho phép xác định hiệu quả số lượng thành phần chính thích hợp, biểu diễn sự thay đổi độ dốc theo số thành phần như trên Hình 2a. Ngoài ra số thành phần phụ thuộc vào điểm gấp “khuyết tay” mà tại đó dựa vào các giá trị riêng còn lại tương đối nhỏ và tất cả có cùng kích thước để lựa chọn thành phần chính. Điểm này không thể hiện rõ trong đồ thị sườn dốc Hình 2a, nhưng chúng ta vẫn có thể kết luận điểm thứ tư là điểm gấp “khuyết tay”. Trong Hình 2b đồ thị mô tả sự thay đổi các thành phần liên quan PC_1 và PC_2 trong không gian bao gồm P_c , T_c , T_L , V_c and T_U . Bằng kỹ thuật phân tích thành phần chính đã chứng minh được các thành phần chính được chọn trong trường hợp này phù hợp với các tính chất nhiệt động của các hệ chất, năm thành phần chính đã được chọn một cách thích hợp, như được chỉ ra trong Hình 2.

Các biến thành phần chính được xác định từ sự kết hợp tuyến tính của các biến ban đầu. Các thành phần chính được trích xuất từ các vectơ riêng cung cấp cho các hệ số của phương trình. Đồ thị kép được minh họa trong Hình 2b cho thấy cả trọng số và điểm số cho hai thành phần được lựa chọn song song. Nó có thể cho biết phép chiếu của một quan sát trên không gian con với các điểm số tương ứng. Nó cũng có thể tìm thấy tỷ lệ quan sát và các biến trong không gian con của hai thành phần ban đầu. Điều này cũng cho phép có thể kiểm tra được các thành phần khác nhau trên không gian con.

Từ Bảng 1, dữ liệu hóa lý tương ứng với các hợp chất, chúng tôi tiến hành xác định thành phần chính tương ứng với tính chất hóa lý P_c , V_c , T_c , T_L và T_U đối với nhóm luyện và nhóm kiểm tra. Các kết quả phân tích thành phần chính được đưa ra ở Bảng 3, được sử dụng để luyện mạng thần kinh nhân tạo và đánh giá khả năng dự đoán của mạng thần kinh dựa vào các chất N₂, Ar, CO, Cl₂ và CO₂ trong nhóm ngoại.

2. Xây dựng mạng thần kinh nhân tạo

Kiến trúc mạng thần kinh nhân tạo I(5)-HL(6)-O(3) được thành lập với năm tham số đầu vào PC₁, PC₂, PC₃, PC₄ và PC₅ của lớp đầu vào; với sáu nút của lớp ẩn HL(6) và ba tham số đầu ra log_a, log_b và log_c của lớp đầu ra [[2],[20],[21]]. Hàm truyền trên các nút hình chữ S với thuật toán luyện mạng Levenberg-Marquardt và các tham số luyện mạng: momen 0,7; tỷ lệ học 0,7 được sử dụng để luyện mạng thần kinh ANN-PCA. Các giá trị MSE = 0,001702; 0,0053802 và 0,0058694 cho các kết quả tương ứng với nút đầu ra log_a, log_b và log_c, tương ứng thu được từ quá trình luyện mạng sau 5000 vòng luyện.

Bảng 4. So sánh kết quả hệ số tính toán (cal.) với giá trị gốc từ thực nghiệm (exp.) [[26],[28],[29]]. Các hệ số tính toán từ mạng I(5)-HL(6)-O(3) được phục hồi trở lại (recover) từ các giá trị logarit.

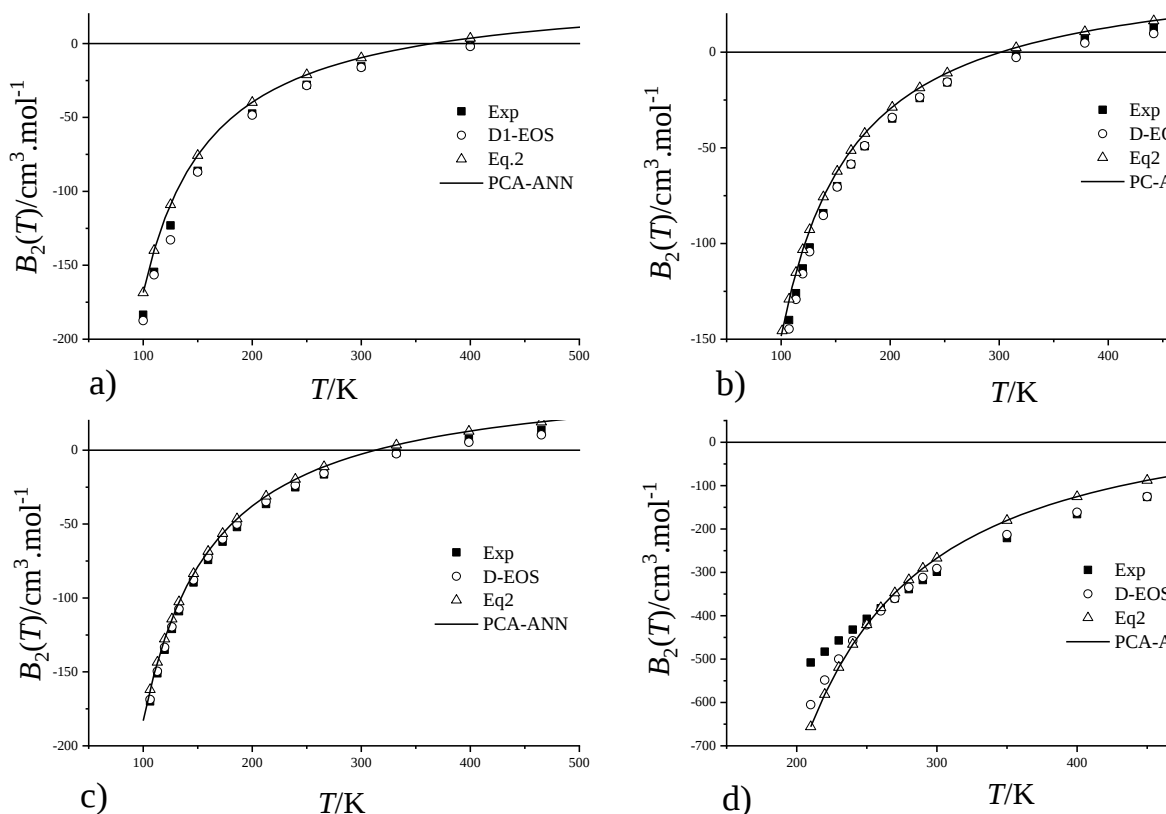
Các khí	Các hệ số chuyển đổi								
	log _a			log _b			log _c		
	exp.	cal.	ARE%	exp.	cal.	ARE%	exp.	cal.	ARE%
N ₂	2,268	2,269	0,044	2,152	2,142	0,437	1,948	1,944	0,2
Ar	2,188	2,18	0,375	2,077	2,058	0,92	2,022	2,013	0,435
CO	2,307	2,311	0,169	2,188	2,18	0,366	1,974	1,973	0,056
Cl ₂	2,305	2,35	1,944	2,12	2,099	0,972	2,613	2,611	0,054
CO ₂	2,139	2,17	1,454	1,943	1,987	2,244	2,513	2,51	0,127
MARE, %			0,633			0,673			0,186

Các khí	Các hệ số hồi phục								
	a			b			c		
	exp.	recover	ARE%	exp.	recover	ARE%	exp.	recover	ARE%
N ₂	185,4	185,8	0,23	141,8	138,8	2,127	88,7	87,89	0,91
Ar	154,2	151,3	1,869	119,3	114,2	4,306	105,1	103,0	2,014
CO	202,6	204,4	0,902	154,2	151,4	1,818	94,2	93,97	0,248
Cl ₂	201,9	223,8	10,87	131,8	125,7	4,63	409,9	408,6	0,312
CO ₂	137,6	147,8	7,412	87,7	96,95	10,55	325,7	323,3	0,743
MARE, %			3,467			3,22			0,871

Các sai số được truyền ngược lại qua hệ thống để điều chỉnh các tham số trong mỗi lớp. Trong quá trình luyện mạng trên cùng một tập dữ liệu được xử lý nhiều lần vì trọng số kết nối đã được tinh chỉnh qua mỗi lần. Trong quá trình học, sai số lệch giữa đầu ra của mô hình mạng và đầu ra mong muốn giảm xuống dần theo quá trình luyện và được mô hình được kết nối tối ưu giữa đầu vào và đầu ra. Các quy tắc luyện được lưu lại để quá trình luyện lặp đi lặp lại trong quá trình luyện được giảm tối thiểu các sai số lỗi. Cấu trúc mạng thần kinh trong nghiên cứu này được biểu thị bằng một mạng lưới ba lớp, như đưa ra trong Hình 1.

Mô hình mạng thần kinh ANN-PCA I(5)-HL(6)-O(3) được sử dụng để dự đoán các hệ số log_a, log_b và log_c. Các giá trị dự đoán log_a, log_b và log_c được phục hồi và so sánh với các tham số ban đầu, như được nêu trong Bảng 4. Các tham số log_a, log_b và log_c nhận được từ mô hình ANN-PCA I(5)-HL(6)-O(3) được so sánh với các giá trị

được chuyển đổi từ tham số thực nghiệm, như được đưa ra trong Bảng 4. So sánh này không chỉ cho biết độ lớn của các giá trị sai số MARE, % đối với các giá trị được chuyển đổi và dự đoán $\log a$, $\log b$ và $\log c$; mà còn được sử dụng cho cả dữ liệu gốc thực nghiệm và giá trị phục hồi. Đây là những giá trị đo lường mức độ tin cậy và độ chính xác của các phép tính, đã khẳng định các kết quả tính toán phù hợp với thực nghiệm và nằm trong khoảng không chắc chắn của thực nghiệm.



Hình 3: So sánh hệ số virial bậc hai của các chất khí được tính bằng mạng ANN-PCA I(5)-HL(6)-O(3), phương trình trạng thái virial (2) [25], phương trình trạng thái Deiters D1-EOS [24], và thực nghiệm, Exp. được lấy từ [25,26]. a) argon; b) khí N₂; c) khí CO; d) khí Cl₂.

Các giá trị sai số MARE, % cho các tham số được tính bởi công thức sau:

$$MARE, \% = \left(\sum_{i=1}^n \frac{|x_{\text{exp.}} - x_{\text{cal}}|}{x_{\text{exp.}}} \right) \frac{100}{n} \quad \text{với } n \text{ từ 1 đến 4} \quad (4)$$

Trong đó các giá trị $x_{\text{exp.}}$ và $x_{\text{cal.}}$ là các giá trị chuyển đổi logarit $\log a$, $\log b$ và $\log c$ từ thực nghiệm và các giá trị tính toán từ mạng I(5)-HL(6)-O(3). Điều này cũng được áp dụng tương tự cho dữ liệu thực nghiệm ban đầu và dữ liệu phục hồi.

Các dữ liệu phục hồi được sử dụng để tính toán các hệ số virial bậc hai cho khí argon, N₂, CO và Cl₂ được thể hiện trong Hình 3. Trong quá trình luyện mạng thần kinh nhân tạo ANN-PCA I(5)-HL(6)-O(3) đã được đánh giá chéo bằng cách sử dụng kỹ

thuật loại bỏ từng trường hợp. Mô hình ANN-PCA I(5)-HL(6)-O(3) này thể hiện khả năng dự đoán tốt cho các tham số chuyển đổi $\log a$, $\log b$ và $\log c$.

Các tham số chuyển đổi $\log a$, $\log b$ và $\log c$ được sử dụng để phục hồi lại hệ số dạng ban đầu trong phương trình virial (2), như được đưa ra trong Bảng 4. Phương pháp phân tích phương sai một yếu tố [[27]] cũng được sử dụng để xác nhận các giá trị dự đoán của $\log a$ ($F = 0,089 < F_{0,05} = 5,318$), đối với $\log b$ ($F = 0,00024 < F_{0,05} = 5,987$), đối với $\log c$ ($F = 0,00024 < F_{0,05} = 5,987$); mà còn phục hồi dữ liệu cho các tham số a ($F = 0,094 < F_{0,05} = 5,987$), đối với tham số b ($F = 0,0044 < F_{0,05} = 5,987$) và đối với tham số c ($F = 0,0002 < F_{0,05} = 5,987$). Điều này cho thấy sự khác biệt giữa khả năng dự báo của mô hình ANN-PCA I(5)-HL(6)-O(3) và các giá trị thực tế là không đáng kể. Vì vậy, mô hình mạng thần kinh ANN-PCA I(5)-HL(6)-O(3) có thể được sử dụng để đánh giá các tham số a , b , và c của phương trình trạng thái virial (2).

3. Dự đoán hệ số virial

Các hệ số virial của khí argon, N_2 , CO và Cl_2 thu được từ quá trình tính toán bằng phương trình trạng thái virial (2) [[25]] sử dụng các hệ số a , b và c dự đoán trong Bảng 4. Các hệ số virial này cũng được tính bằng cách sử dụng phương trình trạng thái Deiters D1-EOS [[24]]. Các kết quả nhận được, so sánh với nhau và với dữ liệu thực nghiệm [[25],[26]], nhận thấy các hệ số virial bậc hai của các chất khí tính toán từ các phương pháp trong công trình này rất gần với dữ liệu thực nghiệm [[25],[26]] và với kết quả tính toán từ phương trình trạng thái Deiters, đã được mô tả ở Hình 3. Sự khác biệt giữa các kết quả tính toán và dữ liệu thực nghiệm là không đáng kể. Các hệ số virial tương tác bậc hai được tạo ra nằm trong vùng không chắc chắn của các phép đo thực nghiệm.

IV. KẾT LUẬN

Nghiên cứu này bước đầu mô tả thành công mối quan hệ giữa các tham số nhiệt động học tới hạn của các khí bằng phương pháp phân tích thành phần chính kết hợp sử dụng mạng thần kinh ANN-PCA I(5)-HL(6)-O(3). Kiểu kiến trúc mạng được xây dựng bằng cách sử dụng kỹ thuật phân tích thành phần chính nhằm nâng cao chất lượng mô hình và chất lượng dự đoán các hệ số $\log a$, $\log b$ và $\log c$ trong phương trình trạng thái. Sau đó các dữ liệu được tái tạo lại dữ liệu định dạng ở dạng hệ số trong phương trình đã đạt chất lượng dự đoán cao với sai số MARE% rất nhỏ. Công trình này cũng đã sử dụng thành công thuật toán di truyền để tìm kiếm cấu trúc mạng thần kinh tối ưu I(5)-HL(6)-O(3) đáp ứng khả năng ứng dụng thực tế. Kỹ thuật này có thể cho phép hứa hẹn trong tương lai sử dụng để thành các mô hình mới theo cách này và đánh giá dữ liệu nhiệt động học của các hệ khác nhau.

Các hệ số virial thu được từ mô hình ANN-PCA I(5)-HL(6)-O(3) khớp rất tốt với các hệ số từ phương trình trạng thái virial và phương trình trạng thái Deiters. Trong tương lai chúng có thể được sử dụng hỗ trợ dự đoán các tham số sigma và epsilon trong hàm thế Lennard-Jones của các hệ chất lỏng. Điều này có thể được sử dụng để dự đoán dữ liệu nhiệt động học của hệ cân bằng lỏng hơi bằng sử dụng cho mô phỏng Monte Carlo toàn cục.

TÀI LIỆU THAM KHẢO

- [1]. H. R. Maier, G. C. Dandy, *J. Mathematical and Computer Modelling*, Vol. 33, P. 669-682, (2001).
- [2]. R. Babuška, H. Verbruggen, Neuro-fuzzy methods for nonlinear system identification, *Annual Reviews in Control*, 27, 73-85, (2003).
- [3]. A.J. Stone, The theory of intermolecular forces, published in the United State by oxford university Press Inc., New York, (1996).
- [4]. A. Z. Panagiotopoulos, *Mol. Phys.*, 61, 813-826, (1987).
- [5]. K. Leonhard and U. K. Deiters, *Mol. Phys.*, 100, 2571-2585, (2002).
- [6]. T. M. Haarmann and W. Koschel, *Proc. Appl. Math. Mech.*, 2, 360-361, (2003).
- [7]. M. Abbaspour, E. K. Goharshadi, Determination of potential energy functions of CO-CO, CO₂-CO₂, and N₂O-N₂O and calculation of transport properties, *J. Chemical Physics*, 330, 313-325 (2006).
- [8]. A. E. Nasrabad, R. Laghaei, Computational studies on thermodynamic properties, effective diameters, and free volume of argon using an *ab initio* potential, *J. Chem. Phys.* 125, 084510, (2006).
- [9]. A. E. Nasrabad, R. Laghaei, U. K. Deiters, Prediction of the thermophysical properties of pure neon, pure argon, and the binary mixtures neon-argon and argon-krypton by Monte Carlo simulation using *ab initio* potentials, *J. Chem. Phys.* 121, 6423, (2004).
- [10]. J. Stoll, J. Vrabec, H. Hasse, A set of molecular models for carbon monoxide and halogenated hydrocarbons, *J. Chem. Phys.* 119, 11396, (2003).
- [11]. P. K. Naicker, A. K. Sum, and S. I. Sandler, *J. Chem. Phys.*, 118, 4086-4093, (2003).
- [12]. A. E. Nasrabad and U. K. Deiters, *J. Chem. Phys.*, 119, 947-952, (2003).
- [13]. A. E. Nasrabad, R. Laghaei, and U. K. Deiters, *J. Chem. Phys.*, 121, 6423-6434, (2004).
- [14]. J. López Cacheiro, B. Fernandez, D. Marchesan, S. Coriani, C. Hattig, and A. Rizzo, *J. Mol. Phys.*, 102, 101-110, (2004).
- [15]. R. T. Pack, *J. Chem. Phys.*, 78, 7217-7222, 1983.
- [16]. S. Bock, E. Bich, and E. Vogel, *Chem. Phys.*, 257, 147-156, (2002).
- [17]. P. Diep and J. K. Johnson, *J. Chem. Phys.*, 112, 4465-4473, (2000).
- [18]. P. Diep and J. K. Johnson, *J. Chem. Phys.*, 113, 3480-3481, (2000).

- [19]. P. L. Fast, M. L. Sanchez, and D. G. Truhlar, *J. Chem. Phys.*, 111, 2921-2926, (1999).
- [20]. Jean-Loup Faulon, Andreas Bender., *Handbook of chemoinformatics algorithms.*, Taylor and Francis Group., (2010).
- [21]. J. Zupan., J. Gasteiger., *neural networks in Chemistry and Drug Design.*, Wiley-VCH, (1992).
- [22]. Deiters, U. A new semiempirical equation of state for fluids – I. Derivation. *Chem. Eng. Sci.*, 36, 1139, (1981).
- [23]. Deiters, U. A new semiempirical equation of state for fluids – II. Application to pure substances. *Chem. Eng. Sci.*, 36, 1147, (1981).
- [24]. U. K. Deiters, ThermoC project homepage: <http://thermoc.uni-koeln.de/index.html>.
- [25]. J. H. Dymond and E. B. Smith, *The Virial Coefficients of Pure Gases and Mixtures*. Clarendon Press, Oxford, (1980).
- [26]. D. R. Lide, *Handbook of Chemistry and Physics*. CRC Press, Raton, 85th edition, (2000).
- [27]. Issa Bass., *SixSigma Statistics with Excel and Minita*, McGraw-Hill, New York, (2007).
- [28]. R. D. Etters, R. Danilowicz, and W. England, *Phys. Reviews A*, 12, 2199-2212, (1975).
- [29]. B. E. Poling, J. M. Prausnitz, J. P. O'Connell., *The Properties of Gases and Liquids*, McGraw-Hill, New York, 3rd edn, p. 53 (2000).

CALCULATION OF SECOND VIRIAL COEFFICIENTS OF GASES Cl₂, N₂, CO AND Ar COMBINING VIRIAL STATE EQUATION AND MULTIVARIATE MODEL

Nguyen Thanh Duoc^{1,4}, Tran Duong², Pham Van Tat^{3,*}

¹ Faculty of Chemistry, University of Sciences, Hue University

² Faculty of Chemistry, University of Education, Hue University

³ Hoa Sen University

⁴ Thu Dau Mot University

*Email: vantat@gmail.com

ABSTRACT

The gaseous components in the chemical industry usually contain of mainly gases such as argon, nitrogen, carbon monoxide and chlorine. The direct emission of these gases into the atmosphere should be limited by their impact on the environment. For its storage and separation, we should sufficiently know the intermolecular interaction and the second virial coefficients that characterize the interaction of molecules. The paper uses the neural network model combining the

principal component analysis ANN-PCA I(5)-HL(6)-O(3) with the MSE error value of 0.0069695 to predict the coefficients a , b and c in the virial state equation based on the critical properties of gases. Second virial coefficients are correctly determined using the predicted coefficients a , b , and c . The computed virial coefficients are very close to the predicted virial coefficients of the Deiter equation and appropriate to the experimental data.

Keywords: ANN neural network, virial coefficients, virial state equation; principal component analysis.



Nguyễn Thành Được sinh ngày 10/10/1976 tại An Giang. Năm 2000, ông tốt nghiệp cử nhân Sư phạm ngành Hóa học tại Trường Đại học Cần Thơ. Năm 2008, ông tốt nghiệp thạc sĩ chuyên ngành Hóa lý thuyết và Hóa lý tại Đại học Cần Thơ. Từ năm 2011 đến nay, ông giảng dạy tại Trường Đại học Thủ Dầu Một, Tỉnh Bình Dương. Từ năm 2012 đến nay, ông là nghiên cứu sinh chuyên ngành Hóa lý thuyết và Hóa lý tại Trường Đại học Khoa học, Đại học Huế.



Trần Dương sinh ngày 01/01/1960 tại Quảng Ngãi. Năm 1981 ông tốt nghiệp ngành Hóa học tại trường Đại học Sư phạm Huế; tốt nghiệp thạc sĩ chuyên ngành Hóa vô cơ năm 1989 tại trường Đại học Sư phạm Hà Nội. Năm 2005 ông bảo vệ luận án tiến sĩ chuyên ngành Hóa vô cơ tại trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội. Năm 2016 ông được công nhận chức danh phó giáo sư chuyên ngành hóa vô cơ. Từ năm 1982 đến nay ông là giảng viên tại Khoa Hóa học, Trường Đại học Sư phạm, Đại học Huế.

Lĩnh vực nghiên cứu: vật liệu vô cơ, hóa tính toán.



Phạm Văn Tất sinh ngày 30/11/1966 tại Nam Định. Năm 1989 ông tốt nghiệp ngành Hóa học tại trường đại học Tổng hợp Hà Nội; tốt nghiệp Thạc sĩ ngành Hóa học Phân tích năm 2001 tại trường đại học Đà Lạt. Năm 2006 ông bảo vệ thành công luận án Tiến sĩ tại Viện Hóa lý và Hóa lý thuyết, trường đại học Cologne, Cộng hòa Liên bang Đức. Năm 2010 ông được nhà nước công nhận chức danh phó Giáo sư chuyên ngành Hóa lý và Hóa lý thuyết. Hiện nay là trưởng Bộ môn ngành Công nghệ Kỹ thuật Môi trường, trường đại học Hoa Sen.

Lĩnh vực nghiên cứu: Hóa học tính toán lượng tử và Mô phỏng Monte Carlo các trạng thái cân bằng, Hóa học Phân tích, Hóa học Môi trường và Quản lý và Đánh giá chất lượng Môi trường bằng GIS, Viễn thám và Hệ thống Trí tuệ Nhân tạo.

