

مطالعه مولکولی ثابت هنری گاز کربن دی اکسید در حلال‌های خالص

زهرا اسحاقی گرجی، علی ابراهیم پور گرجی، سیاوش ریاحی*⁺

تهران، دانشگاه تهران، پردیس دانشکده‌های فنی، دانشکده مهندسی شیمی، انیستیتو مهندسی نفت

چکیده: حذف گازهای اسیدی از جمله کربن دی اکسید از جریان‌های گازی، یکی از معضله‌های اساسی در صنعت است؛ چون وجود این گاز مشکل‌های زیادی را در فرایندهای دیگر ایجاد می‌کند. هدف از این کار، یافتن حلال‌هایی است که بهترین عملکرد را در جذب CO_2 داشته باشند. مقدارهای ثابت هنری برای کربن دی اکسید در گروهی از حلال‌ها با استفاده از مدلی که از روش رابطه کمی - ساختار ویژگی (QSPR) ایجاد شده است، محاسبه شده است. در ابتدا مولکول‌های حلال براساس یکی از روش‌های نظریه تابعیت چگالی (DFT) به نام B3LYP و با مجموعه تابع‌های ++ G (d, p) -۳۱، در نرم‌افزار گوسین بهینه‌سازی شده و سپس توصیف‌کننده‌های مناسب با استفاده از نرم‌افزار دراگون محاسبه شدند. توانایی مدل با دو توصیف‌کننده ساده مولکولی شامل تعداد گروه عاملی هیدروکسیل (n-ROH) و اندیس محوری بالابان (BAC)، برای پیش‌بینی ثابت هنری به روش اعتبارسنجی متقاطع حذف یکتایی سنجیده شده است و مقدار ضریب تعیین برای مدل اصلی ارائه شده به تقریب برابر با ۰٫۹۵ می‌باشد. میانگین انحراف نسبی این مدل نیز کمتر از ۷٪ بود. نتیجه‌های به دست آمده نشان از دقت بالای مدل و توافق بسیار خوب آن با مقدارهای اندازه‌گیری شده در منابع علمی گوناگون دارد.

واژه‌های کلیدی: مدل؛ ثابت هنری؛ کربن دی اکسید؛ گروه عاملی هیدروکسیل؛ اندیس محوری بالابان.

KEYWORDS: Model; Henry's law constant; Carbon dioxide; Hydroxyl functional group; Balaban centric index.

مقدمه

در صورت تشکیل در خطوط لوله انتقال گاز، منجر به بسته شدن آن می‌شوند. خوردگی لوله‌ها و مخازن و کاهش ارزش گرمایی گاز طبیعی از دیگر دلایل‌های حذف کربن دی اکسید است.

با توجه به مقدار دی اکسید کربن در گاز طبیعی، از روش‌های گوناگونی برای حذف آن استفاده می‌کنند که مهم‌ترین آن‌ها شامل جذب فیزیکی^(۱)، جذب شیمیایی^(۲) و استفاده از غشای^(۳) است. فرایند جذب با حلال‌های فیزیکی زمانی استفاده می‌شود که فشار جزئی CO_2 بالا بوده و مقدار بالایی از جریان گاز

شیرین‌سازی گاز یا حذف گازهای اسیدی (به‌عنوان نمونه گاز کربن دی اکسید CO_2) و هیدروژن سولفید (H_2S) در بسیاری از صنایع استفاده می‌شود [۱] و به تقریب همه گازهای مخزن که در سرتاسر جهان وجود دارد، دارای این ترکیب‌ها هستند. همراه بودن CO_2 با آب منجر به تشکیل هیدرات گازی می‌شود. هیدرات‌ها ترکیب‌های بلوری شبیه به یخ هستند که در شرایط دمایی و فشاری خاصی تشکیل شده و در فرایند تشکیل آن، مولکول‌های گاز در داخل شبکه‌ای از مولکول‌های آب به دام می‌افتند و

*عهده دار مکاتبات

+E-mail: riahi@ut.ac.ir

(۱) Physical absorption

(۲) Chemical absorption

(۳) Membrane

یکی از روش‌های ترمودینامیکی پیچیده با محاسبه‌های طولانی که در به‌دست آوردن ثابت هنری مورد استفاده قرار گرفته، روش سهم گروه^(۱) است. در این روش هر ترکیب به گروه‌های عاملی سازنده خود تقسیم شده و خاصیت آن ماده از جمع ویژگی‌های گروه‌ها به دست می‌آید [۷، ۸]. گروه دیگری از پژوهشگران از انواع معادله‌های حالت مانند معادله پنگ-رابینسون^(۲) [۹، ۱۰]، پی سی سفت^(۳) [۱۱]، معادله مکعبی به اضافه تجمعی^(۴) [۱۲]، مدل کنت-هایزنبرگ^(۵) [۱۳]، ضریب‌های فعالیت ویلسون، ون لار و ان. آر. تی. ال^(۶) [۱۴] برای محاسبه مقدار حلالیت کربن دی اکسید بهره برده‌اند. دسته‌ی دیگری از روش‌های نظری که عملکرد بسیار خوبی در پیش‌بینی ویژگی‌های مولکولی دارد، فناوری مطالعه‌ی کمی ساختار - ویژگی یا به اختصار QSPR^(۷) است [۱۵]. این روش می‌تواند هر خاصیتی که مرتبط با مولکول است را به صورت کمی به ساختار آن مولکول مرتبط سازد و این کار با استفاده از توصیف‌کننده^(۸)ها انجام می‌گیرد. توصیف‌کننده‌ها، مقدارهای عددی هستند که به ساختار و شکل مولکول وابسته‌اند [۱۶]. بنابراین در روش QSPR، دانستن تنها ساختار مولکول کافی است؛ بدون این که نیازی به انجام محاسبه‌های پیچیده و شرایط آزمایشگاهی خاص باشد. به عنوان مثال، مؤمنی و ریاحی با استفاده از این فناوری به مطالعه ظرفیت جذب آمین در برابر کربن دی اکسید در فرایند شیرین‌سازی گاز، به روش مطالعه کمی ساختار-ویژگی پرداخته‌اند. مدل خطی براساس ۲۳ حلال آمینی و با استفاده از سه توصیف‌کننده ساده ارایه شده است. قدرت تخمین مدل با انواع آزمون‌های آماری مورد ارزیابی قرار گرفته و نتیجه‌های خوبی به همراه داشته است. آن‌ها نشان دادند که بهترین حلال آمینی برای جذب CO₂، حلالیست با یک ساختار خطی که تعداد زیادی از آمین‌های اولیه و ثانویه به آن متصل است [۱۷].

با توجه به مطالعه‌های انجام شده، این اولین بار است که اثر ساختار مولکولی بر ثابت هنری کربن دی اکسید در حلال‌های الکلی، اتری، کتونی و گلایکولی، مورد بررسی قرار گرفته است. هدف از این کار، ارایه مدل QSPR با استفاده از روش برازش خطی چندگانه^(۹) است تا ثابت هنری کربن دی اکسید را در انواع حلال

وجود داشته باشد. اگرچه حلال‌های شیمیایی مانند آلکانوآمین‌ها از جمله پرکاربردترین حلال‌های مورد استفاده در صنعت هستند، اما حلال‌های فیزیکی به خاطر غیرخورنده و غیرسمی بودن و احیاسازی تنها با کاهش فشار و مصرف بسیار پایین انرژی، توجه زیادی را به خود جلب کرده است [۲].

میزان حلالیت CO₂ در حلال‌هایی که به صورت فیزیکی این گاز را جذب می‌کنند از قانون هنری تبعیت می‌کند، زیرا ظرفیت جذب این حلال‌ها متناسب با فشار جزئی گاز است [۳]. قانون هنری بیان می‌کند که در یک دمای ثابت، مقدار گاز حل شده در یک مایع، به طور مستقیم متناسب با فشار جزئی آن گاز در حالت تعادل با مایع است. ثابت این تناسب را ثابت هنری نامیده و با H_i نشان می‌دهند و به صورت نسبت فوگاسیته جزء i در فاز گاز (f_i) به جزء مولی آن جزء در فاز مایع (x_i)، زمانیکه جزء مولی به سمت صفر میل می‌کند، تعریف می‌شود [۴]. شایان ذکر است که این پارامتر با حلالیت رابطه عکس دارد. بیان ریاضی این قانون در معادله (۱) آورده شده است:

$$H_i = \lim_{x_i \rightarrow 0} f_i / x_i \quad (1)$$

تعیین ثابت هنری کربن دی اکسید در حلال‌های گوناگون به صورت تجربی انجام شده است که به عنوان نمونه، پژوهشگران در سال ۲۰۱۱ میلادی، به بررسی میزان حلالیت CO₂ در گروه‌های گوناگونی از حلال شامل الکل‌ها، گلایکول‌ها، اترها و کتون‌ها در فشارهای بالا پرداخته‌اند. آن‌ها نتیجه گرفتند که کتون‌ها بیش‌ترین ظرفیت جذب را داشته و گروه عاملی کربونیل و باند اتر منجر به بهبود جذب فیزیکی دی اکسید کربن می‌شود و گروه عاملی هیدروکسیل از جذب آن توسط حلال جلوگیری خواهد کرد [۵]. اما در حالت کلی، اندازه‌گیری مقدار ثابت هنری کار آسانی نیست؛ به خاطر هزینه‌ی بالای دستگاه‌ها و تجهیزات، زمان‌بر بودن آزمایش‌ها، در دسترس نبودن برخی از مواد شیمیایی موردنیاز و به ویژه مشکل‌هایی که در رابطه با ترکیب‌ها با حلالیت کم وجود دارد [۶]. به همین دلیل ممکن است مقدارهای گوناگونی از ثابت هنری در منابع علمی گزارش شده باشد. از این رو، بسیاری از پژوهشگران به استفاده از روش‌های تئوری تمایل پیدا کرده‌اند.

(۱) Group contribution

(۲) Peng-robinson

(۳) PC-SAFT

(۴) Cubic plus association

(۵) Kent-eisenberg

(۶) NRTL

(۷) Quantitative structure-property relationship

(۸) Descriptor

(۹) Multiple linear regression

قادر به ایجاد ۴۸۸۵ توصیف‌کننده از یک ترکیب است، استفاده شد. توجه شود که وجود این تعداد توصیف‌کننده، ماتریسی با اندازه‌های بسیار بزرگ ایجاد کرده و بار انجام محاسبات را بالا می‌برد. به منظور کاهش این تعداد، تنها پنج نوع از توصیف‌کننده‌های ساده شامل: توصیف‌کننده‌های اساسی^(۱۰)، توپولوژیکی^(۱۱)، گروه عاملی^(۱۲)، ویژگی‌های مولکولی^(۱۳) و اجزای میان اتمی^(۱۴) که از همه ساده‌تر هستند، انتخاب شده‌اند. با این کار، تعداد توصیف‌کننده‌های محاسبه شده به ۷۰ کاهش خواهد یافت.

ایجاد مدل و اعتبارسنجی

انتخاب توصیف‌کننده به روش الگوریتم ژنتیک [۲۲] انجام شد که منجر به تولید چندین معادله خطی می‌شود. تأثیرگذارترین توصیف‌کننده‌های مولکولی بر مقدار ثابت هنری کربن دی اکسید در حلال‌هایی که در ده مدل برتر ظاهر شده، در جدول ۱ گزارش شده است. عملکرد و قدرت تخمین مدل‌های مورد نظر در این کار، با استفاده از فناوری‌های گوناگون اعتبارسنجی داخلی و خارجی ارزیابی شد که می‌توان به (۱) اعتبارسنجی متقاطع حذف یکتایی (LOO-CV)^(۱۵)، (۲) دسته‌بندی داده‌ها به دو دسته آموزشی^(۱۶) و ارزیابی^(۱۷) و (۳) ارزیابی توانایی مدل در تخمین ثابت هنری با مجموعه داده‌های خارجی، اشاره کرد. یکی از پرکاربردترین روش‌ها برای ارزیابی داخلی مدل، استفاده از فناوری اول یا همان اعتبارسنجی متقاطع حذف یکتایی است. در این روش، در هر مرحله یک داده از مجموعه داده‌ها کنار گذاشته می‌شود و مدل‌سازی با ترکیب‌های باقی‌مانده انجام می‌گیرد و بعد از ساخت مدل، خاصیت ترکیب کنار گذاشته شده پیش‌بینی می‌شود. این روند آن‌قدر تکرار می‌شود که همه ترکیب‌ها یک بار کنار گذاشته شوند. در فناوری دوم، بخشی از داده‌ها برای ایجاد مدل به کار گرفته می‌شوند (دسته آموزشی) و بخش دیگر، توانایی مدل را می‌سنجند (دسته ارزیابی). این تقسیم‌بندی داده‌ها به روش تجزیه به مولفه‌های اصلی (PCA)^(۱۸)

پیش‌بینی کند. مدل موردنظر، توصیف‌کننده‌های ساده و قابل تفسیری دارد که به راحتی قابل محاسبه‌اند و نتیجه‌های دلخواه به دست آمده از انواع روش‌های اعتبارسنجی، نشان از توانایی بالای مدل برای تخمین ثابت هنری دارد.

بخش نظری

مجموعه داده‌ها^(۱) و توصیف‌کننده^(۲)ها

مقدارهای تجربی از کار دو گروه از پژوهشگران، ژریو و همکاران [۵] و هنی^(۳) و همکاران [۱۸]، جمع‌آوری شده است. از بین ۱۴ حلال مورد بررسی در مرجع [۱۸]، سه مورد حذف شد؛ زیرا یکی از حلال‌ها جزء حلال‌های پلیمری بوده و مورد دوم، تنها حلالی است که در ساختار مولکولی خود اتم نیتروژن دارد. شایان گفتن است که حلال‌های مورد استفاده در این مطالعه، در ساختار خود تنها اتم‌های کربن، اکسیژن و هیدروژن دارند. مقدار ثابت هنری CO₂ در اتیلن گلیکول مونو متیل اتر، در هر دو مرجع یاد شده گزارش شده است که برای جلوگیری از تکرار، از مرجع [۱۸] حذف شده است. همه مقدارهای تجربی ثابت هنری به صورت لگاریتم طبیعی درآورده شده که مقدار عددی آن بین ۱ و ۴ است. نزدیک بودن ساختار مولکول به ساختار طبیعی خود، می‌تواند بر پیش‌بینی خاصیت موردنظر اثر گذار باشد؛ بنابراین ساختار مولکولی ترکیب‌ها باید در پایدارترین حالت خود قرار داشته باشند. برای این کار، ساختار مولکولی در نرم‌افزار هایپرکم^(۴) [۱۹] رسم شده و با استفاده از روش نیمه تجربی RMI و به کار بردن الگوریتم پلاک-رایبر^(۵) بهینه‌سازی شده‌اند. بهینه‌سازی بیشتر در نرم‌افزار گوسین^(۶) [۲۰] که براساس یکی از روش‌های نظریه تابعیت چگالی (DFT)^(۷) به نام B3LYP^(۸) و با مجموعه توابع G (d, p)++-۳۱-۶ است، انجام می‌گیرد. محاسبه توصیف‌کننده‌ها با نرم‌افزارهای گوناگونی ممکن است انجام گیرد. در این مطالعه، از نرم‌افزار دراگون^(۹) [۲۱] که

(۱) Data set

(۲) Descriptor

(۳) Henni

(۴) Hyperchem

(۵) Polak-Ribiere

(۶) Gaussian

(۷) Density Functional Theory

(۸) Becke, 3-parameter, Lee-Yang-Parr

(۹) Dragon 6

(۱۰) Constitutional Descriptors

(۱۱) Topological Descriptors

(۱۲) Functional group counts

(۱۳) Molecular properties

(۱۴) Atom-Centered Fragments

(۱۵) Leave-One-Out cross validation

(۱۶) Training set

(۱۷) Test set

(۱۸) Principal Component Analysis

جدول ۱- توصیف‌کننده‌های ظاهر شده در ده مدل برتر QSPR.

تعریف	نوع	توصیف‌کننده
نمایه زاگرب اول با درجات راس واحد ^(۱)	اندیس توپولوژیک	ZM1V
نمایه زاگرب دوم با درجات راس واحد ^(۲)	اندیس توپولوژیک	ZM2V
اندیس محوری بالابان	اندیس توپولوژیک	BAC
دگرگونی الکترو توپولوژیک مولکولی ^(۳)	اندیس توپولوژیک	DELS
تعداد گروه‌های عاملی هیدروکسیل	گروه عاملی	nROH
تعداد اترها ^(۴)	گروه عاملی	nROR
جرم مولکولی	اندیس اساسی	MW
تعداد اتم‌های اکسیژن ^(۵)	اندیس اساسی	nO
تعداد هترواتم‌ها ^(۶)	اندیس اساسی	nHet
مجموع مرتبه پیوندهای قراردادی (بدون توجه به پیوند هیدروژن) ^(۷)	اندیس اساسی	SCBO
سطح قطبی توپولوژیک با استفاده از سهم قطبیت نیتروژن و اکسیژن ^(۸)	خواص مولکولی	TPSA (NO)

(۱) First Zagreb index by valence vertex degrees, (۲) Second Zagreb index by valence vertex degrees

(۳) Molecular electrotopological variation, (۴) Number of ethers

(۵) Number of oxygen atoms, (۶) Number of hetroatoms

(۷) Sum of conventional bond orders (H-depleted), (۸) Topological polar surface area using N, O polar contribution

اگر ساختار مولکولی بدون اتم هیدروژن به‌عنوان یک گراف در نظر گرفته شود، اندیس مرکزی بالابان براساس هرس کردن گراف^(۲) تعریف می‌شود که شامل یک روش گام به گام برای حذف همه اتم‌هایی است که در راس‌های پایانی واقع شده و تنها یک اتم در همسایگی خود دارند. n_g تعداد اتم‌هایی است که در مرحله g ام حذف می‌شود و R تعداد کل مراحل است که در آن همه رئوس حذف می‌شوند [۲۵]. این توصیف‌کننده نشان‌دهنده مقدار شاخه‌های مولکولی موجود در یک ترکیب است و هرچه مقدار آن بزرگتر باشد، شاخه‌های مولکولی بیش‌تر خواهد بود.

در فناوری سوم، افزون بر اینکه قدرت تخمین مدل برای داده‌های خارجی موجود در منابع علمی دیگر که در مدل‌سازی استفاده نشده‌اند، بررسی می‌شود؛ بلکه به نوعی به صحت‌سنجی آن داده تجربی هم کمک می‌شود. در کنار روش‌های گوناگونی که برای اعتبارسنجی مدل استفاده شده، از پارامترهای آماری گوناگون شامل؛ ضریب تعیین^(۳) (R^2)، ضریب تعیین شده (R^2_{adj})، ضریب تعیین در حذف یکتایی (Q^2)، آماره فیشر (F) و

انجام شده است [۲۳] و حدود ۸۰٪ از مجموعه داده‌ها (۱۷ حلال) برای دسته آموزشی و ۲۰٪ آن‌ها (پنج حلال) برای دسته ارزیابی انتخاب شدند. بعد از تقسیم‌بندی ۲۲ حلال به دو دسته آموزشی و ارزیابی، مدل‌هایی به روش برازش خطی چندگانه با یک و دو توصیف‌کننده و با استفاده از ۱۷ حلالی که در دسته آموزشی قرار دارند، ارایه شده است که به ترتیب در معادله‌های (۲) و (۳) قابل دیدن است.

$$\ln(H) = 0.543 nROH^2 + 1.65069 \quad (2)$$

$$\ln(H) = 0.47448 nROH^2 - 0.3582 BAC + 2.25664 \quad (3)$$

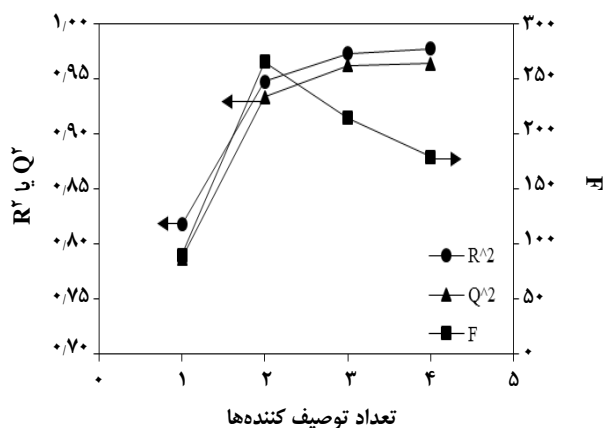
$nROH$ تعداد گروه عاملی هیدروکسیل در هر ترکیب را نشان می‌دهد و BAC ^(۱) یا اندیس محوری بالابان، یکی از توصیف‌کننده‌های توپولوژیکی است که به‌صورت معادله (۴) تعریف می‌شود [۲۴]:

$$BAC = \sum_{g=1}^R n_g^{\gamma} \quad (4)$$

(۱) Index

(۲) Pruning of the graph

(۳) Coefficient of Determination



شکل ۱- اثر تعداد توصیف کننده در کیفیت مدل و تعیین تعداد بهینه توصیف کننده با نمودار شکست.

در ابتدا برای اینکه مشخص شود چه توصیف کننده‌ای اهمیت بالاتری در تعیین مقدار ثابت هنری CO₂ دارد، مدلی با یک توصیف کننده ارائه شد. با توجه به معادله (۲)، دیده می‌شود که از بین ۷۰ توصیف کننده، مجذور تعداد گروه عاملی هیدروکسیل در مدل ظاهر شده است. همان گونه که در این معادله دیده می‌شود، مقدارهای ثابت هنری در مقیاس لگاریتمی با مجذور تعداد گروه عاملی هیدروکسیل رابطه مستقیم دارد و با افزایش این توصیف کننده مقدار ثابت هنری به طور چشمگیری افزایش می‌یابد. با توجه به شکل ۲، ۲۲ حلال بر اساس این توصیف کننده به سه دسته مجزا تقسیم می‌شوند که هر دسته شامل حلال‌هایی است که تعداد گروه عاملی هیدروکسیل برابر دارند. ۷ حلال در دسته اول قرار گرفته‌اند که فاقد گروه عاملی هیدروکسیل هستند. ۱۳ حلال در دسته دوم با یک گروه عاملی هیدروکسیل و دو حلال در دسته سوم با دو گروه عاملی هیدروکسیل جای گرفته‌اند. همچنین این شکل بیان می‌کند که بازه‌ی تغییرهای ln H برای هر دسته، یک بازه‌ی مشخصی بوده و با تغییر مقدار این توصیف کننده، بازه‌ی جدیدی از ln H اشغال می‌شود. بازه تغییرهای در دسته‌های اول، دوم و سوم به ترتیب ۱/۸۲-۱/۰۹، ۲/۶۱-۱/۶۶ و ۳/۸۶-۳/۶۴ است. مقدارهای تجربی ثابت هنری کربن دی اکسید در حلال‌هایی که در یک دسته واقع شده‌اند، نزدیک به هم است و این موضوع بیان کننده اثر بسیار بالای این توصیف کننده در تعیین ثابت هنری CO₂ در این حلال‌هاست.

مقدار خطای استاندارد (S) نیز استفاده شده است که معادله‌های مربوط به هر کدام در ادامه آورده شده است (معادله‌های (۶)-(۱۱)). بهترین مدل براساس نزدیک‌تر بودن مقدارهای R²، R²_{adj} و Q² به یک، نزدیک‌تر بودن مقدار S به صفر و بالاترین مقدار F انتخاب می‌شود. مقدار انحراف نسبی (RD%) که اختلاف مقدار محاسبه شده و مقدار تجربی نسبت به مقدار تجربی را نشان می‌دهد نیز می‌تواند قدرت پیش‌بینی مدل را ارزیابی کند.

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (6)$$

$$R^2_{adj} = 1 - \left(1 - R^2\right) \left(\frac{n-1}{n-p-1}\right) \quad (7)$$

$$Q^2_{Loo} = 1 - \frac{PRESS}{TSS} = 1 - \frac{\sum (y_i - \hat{y}_{iii})^2}{\sum (y_i - \bar{y})^2} \quad (8)$$

$$F = \frac{MSS/df_m}{RSS/df_e} = \frac{\sum (\hat{y}_i - \bar{y})^2 / p}{\sum (y_i - \hat{y}_i)^2 / (n-p-1)} \quad (9)$$

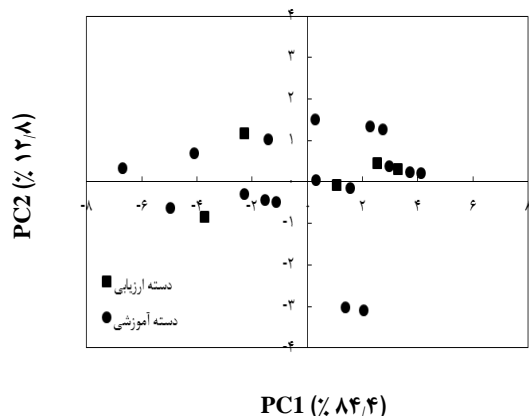
$$s = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n-p-1}} \quad (10)$$

$$RD(\%) = \frac{\left| \left(x_{exp} \right) - \left(x_{pre} \right) \right|}{x_{exp}} \times 100 \quad (11)$$

نتیجه‌ها و بحث

برای تعیین تعداد بهینه توصیف کننده‌ها در مدل QSPR، از نموداری موسوم به نمودار شکست^(۱) استفاده می‌شود. در این کار مدل‌های گوناگونی با تعداد یک، دو، سه و چهار توصیف کننده ساخته شده و برخی از پارامترهای آماری آن از جمله R²، Q² و F محاسبه شدند. شکل ۱ نمودار شکست را نشان می‌دهد که در آن پارامترهای آماری بر حسب تعداد توصیف کننده‌ها رسم شده است. دیده می‌شود که استفاده از سه توصیف کننده در مدل‌سازی نسبت به دو توصیف کننده، نه تنها اثر چندانی بر پارامترهای آماری R² و Q² ندارد بلکه باعث کاهش چشمگیر مقدار پارامتر F می‌شود. بنابراین از دو توصیف کننده در مدل نهایی استفاده شده است.

(۱) Breaking point plot



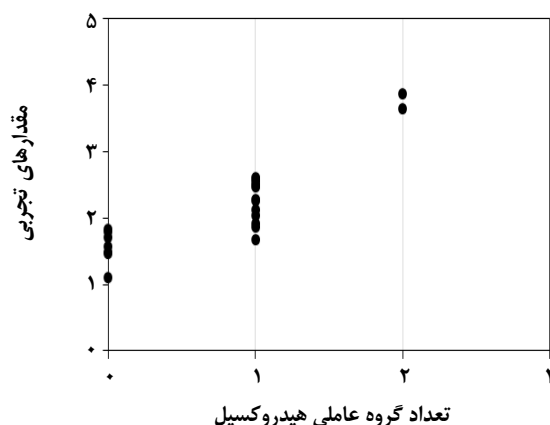
شکل ۳- آنالیز مؤلفه اصلی برای مولکول‌ها در دو مجموعه آموزشی و ارزیابی.

از مجموعه کل داده‌ها دور هستند، نباید در مجموعه ارزیابی واقع شوند چون خط همبستگی را به شدت تحت تأثیر قرار می‌دهند. با بررسی شکل ۳ می‌توان نتیجه گرفت که دسته ارزیابی به خوبی انتخاب شده است.

بعد از تقسیم‌بندی داده‌ها به دو دسته ارزیابی و آموزشی، مقدارهای پیش‌بینی شده ثابت هنری CO_2 از معادله‌های (۲) و (۳) به همراه مقدارهای تجربی در مقیاس لگاریتمی در جدول ۲ گزارش شده است. با توجه به نتیجه‌های معادله (۲) دیده می‌شود که مقدار ثابت هنری برخی از حلال‌ها برابر است، درحالی که مقدارهای تجربی متفاوتی دارند. نتیجه‌های معادله (۳) بسیار دلخواه بوده و تطابق خوبی بین مقدارهای پیش‌بینی شده و تجربی وجود دارد.

در شکل ۴، مقدارهای تجربی بر حسب مقدارهای پیش‌بینی شده مربوط به معادله (۳)، ترسیم شده است تا درک بهتری را از توانایی مدل برای تخمین ثابت هنری ارایه کند. همان‌گونه که دیده می‌شود، کلیه داده‌ها در حول خط ۴۵ درجه واقع شده‌اند و این مسئله کیفیت بالای مدل ارایه شده را نشان می‌دهد. همچنین از شکل ۴ مشخص است که دو مولکول که تأثیر بیشتری نسبت به سایر مولکول‌ها در ایجاد مدل دارند، از سایر مولکول‌ها که در ناحیه‌ای متراکم شده‌اند، فاصله بیشتری داشته و این دو نباید در دسته ارزیابی قرار گیرند. در این کار نیز دو مولکول یاد شده در دسته آموزشی قرار گرفته‌اند.

یکی از پرکاربردترین روش‌ها برای تعیین پایداری و ارزیابی داخلی مدل، استفاده از روش اعتبارسنجی متقاطع حذف یکتایی است.



شکل ۲- نمودار مقدارهای تجربی ثابت هنری برحسب تعداد گروه عاملی هیدروکسیل.

از آنجایی که حلال‌های واقع شده در هر دسته مقدار nROH^2 برابری دارند، انتظار می‌رود که با استفاده از مدل تک متغیره، مقدار ثابت هنری کربن دی‌اکسید در آن‌ها یکسان باشد، درحالی که در واقعیت چنین نتیجه‌ای درست نیست. همچنین با توجه به شکل ۱ نیز مشخص شد که برای داشتن مدلی معتبر، استفاده از دو توصیف‌کننده ضروری است. به این ترتیب معادله (۳) با دو توصیف‌کننده شامل nROH^2 و BAC که به راحتی قابل محاسبه و قابل تفسیر هستند، به عنوان بهترین مدل انتخاب شد. ضریب همبستگی خطی بین این دو توصیف‌کننده، برابر با ۰/۱۰۷ است. این موضوع نشان می‌دهد که توصیف‌کننده‌های موردنظر مستقل از هم بوده و می‌توانند برای ساخت یک مدل QSPR مورد استفاده قرار گیرند.

همان‌گونه که پیش از این گفته شد یکی از راه‌های اعتبارسنجی داخلی مدل، استفاده از روش PCA برای تقسیم‌بندی داده‌ها به دو دسته آموزشی و ارزیابی است که در شکل ۳ نشان داده شده است. مقدارهای PC با استفاده از توصیف‌کننده‌های گزارش شده در جدول ۱ محاسبه می‌شود. نتیجه‌ها نشان داده که ۹۷/۲ درصد از کل اطلاعات داده‌ها در PC1 و PC2 قرار دارند که ۸۴/۴٪ آن در PC1 و ۱۲/۸٪ در PC2 است و به همین دلیل این دو PC، در شکل ۳ نسبت به هم رسم شده‌اند. توجه شود که در انتخاب داده‌های دسته ارزیابی باید نکاتی رعایت شود از جمله: (۱) حتی الامکان از هر کدام از ناحیه‌های چهارگانه، مولکولی انتخاب شود، (۲) تعداد مولکول‌های انتخابی از هر ناحیه متناسب با کل مولکول‌ها در آن ناحیه باشد، (۳) نقاطی که

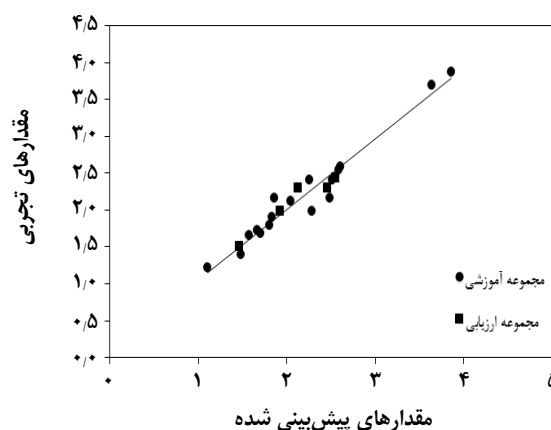
جدول ۲- مقدارهای ثابت هنری تجربی و پیش‌بینی شده برای دی اکسید کربن در ۲۲ حلال در دمای ۲۹۸ کلین.

نام ماده	ln H _{exp} (MPa)	معادله (۲)	نام ماده	ln H _{exp} (MPa)	معادله (۳)	نام ماده	ln H _{exp} (MPa)
		ln H _{pre} (MPa)			ln H _{pre} (MPa)		
متانول	۲٫۶۱۲	۲٫۱۵	دی اتیلن گلایکول متیل‌اتر	۲٫۵۹	۲٫۱۶	متانول	۲٫۶۱۲
اتانول	۲٫۵۸۷	۲٫۱۵	تری اتیلن گلایکول متیل‌اتر*	۲٫۵۵	۱٫۹۸	اتانول	۲٫۵۸۷
ان - پروپانول*	۲٫۵۵۳	۲٫۱۵	اتیلن گلایکول دی متیل‌اتر	۲٫۴۴	۱٫۶۸	ان - پروپانول*	۲٫۵۵۳
ان - بوتانول	۲٫۵۱	۲٫۱۵	دی اتیلن گلایکول دی متیل‌اتر	۲٫۴۱	۱٫۶۵	ان - بوتانول	۲٫۵۱
ان - پنتانول*	۲٫۴۶۴	۲٫۱۸	تری اتیلن گلایکول دی متیل‌اتر	۲٫۳	۱٫۴	ان - پنتانول*	۲٫۴۶۴
اتیلن گلایکول	۳٫۸۶۴	۳٫۷۸	تترا اتیلن گلایکول دی متیل‌اتر	۳٫۸۷	۱٫۲۲	اتیلن گلایکول	۳٫۸۶۴
پروپیلن گلایکول	۳٫۶۴۱	۳٫۹۶	دی اتیلن گلایکول اتیل‌اتر	۳٫۶۹	۲٫۱۲	پروپیلن گلایکول	۳٫۶۴۱
استون	۱٫۸۲۶	۱٫۶۱	دی اتیلن گلایکول دی اتیل‌اتر*	۱٫۹	۱٫۵	استون	۱٫۸۲۶
۲-بوتانول	۱٫۸۰۱	۱٫۶۱	اتیلن گلایکول بوتیل‌اتر	۱٫۷۹	۲٫۱۶	۲-بوتانول	۱٫۸۰۱
اتیلن گلایکول متیل‌اتر	۲٫۲۵۸	۲٫۱۸	دی اتیلن گلایکول بوتیل‌اتر	۲٫۴۱	۱٫۹۸	اتیلن گلایکول متیل‌اتر	۲٫۲۵۸
اتیلن گلایکول اتیل‌اتر*	۲٫۱۲۵	۲٫۱۸	تری اتیلن گلایکول بوتیل‌اتر	۲٫۳	۱٫۷۳	اتیلن گلایکول اتیل‌اتر*	۲٫۱۲۵

با مدل است. همچنین میانگین انحراف نسبی معادله (۳) هم در حالت کلی ۵٫۲۸ درصد بوده که قدرت بالای مدل را در تخمین ثابت هنری نشان می‌دهد.

در شکل ۵ مقدار خطای به دست آمده از تخمین ثابت هنری با معادله (۳) بر حسب مقدارهای تجربی نشان داده شده است. هرچه نقطه‌ها به خط افقی نزدیک‌تر باشند، مقدار خطای مدل کم‌تر خواهد بود. همان‌گونه که دیده می‌شود، برخی از نقاط بالاتر از خط افقی و برخی پایین‌تر هستند و این نشان‌دهنده نبود خطای منظم یا سیستمی در مدل می‌باشد.

دامنه کاربرد پذیری^(۱) مدل QSPR به دست آمده، می‌تواند با استفاده از نمودار ویلیام^(۲) که در شکل ۶ نشان داده شده، به دست آید [۲۷]. این نمودار با رسم مقدارهای خطای استاندارد مدل بر حسب مقدارهای لوریج^(۳) به دست می‌آید. نمودار ویلیام هم می‌تواند به تشخیص داده‌های پرت کمک کند و هم بررسی می‌کند که آیا حلال جدید در بازه‌ی ساختاری حلال‌های مدل شده، قرار دارد یا خیر. هر چه مقدار لوریج برای ترکیبی بیش‌تر باشد، آن ترکیب تاثیر بیشتری در ایجاد مدل نهایی دارد. مقدار لوریج با استفاده از معادله (۱۲) قابل محاسبه است [۲۸].



شکل ۴- نمودار مقدارهای ثابت هنری پیش‌بینی شده بر حسب مقدارهای تجربی در دو دسته آموزشی و ارزیابی.

پارامتر Q^2 مربوط به معادله (۳) در جدول ۳ گزارش شده است. مقایسه این مقدار با معیارهایی که توسط پژوهشگران بیان شده [۲۶]، پایداری مدل را تایید می‌کند. دیگر پارامترهای آماری نیز در جدول ۳ گزارش شده است از جمله مقدار ضریب تعیین که برابر با ۰٫۹۵۱۵ به دست آمد و نشان‌دهنده برازش خوب داده‌ها

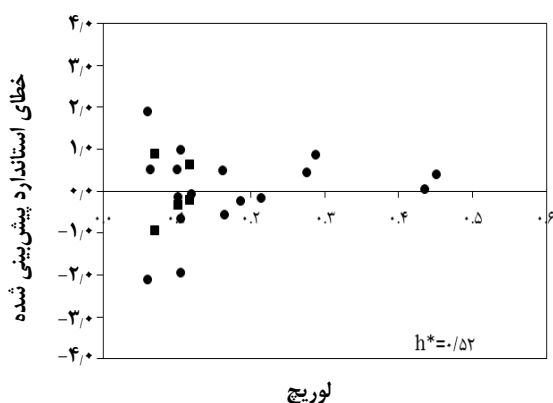
(۱) Applicability domain

(۲) William's plot

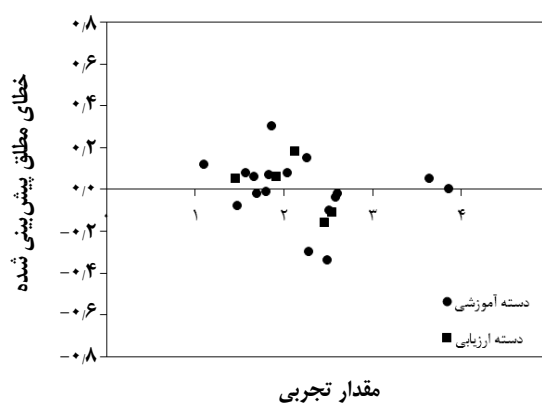
(۳) Leverage

جدول ۳- پارامترهای آماری دو دسته‌ی آموزشی و ارزیابی برای معادله (۳).

ARD	S	F	R^2_{adj}	Q^2	R^2	N	
۵/۳۸	۰/۱۶۴	۱۵۰/۳۳	۰/۹۵	۰/۹۴۵	۰/۹۵۵۵	۱۷	دسته آموزشی
۴/۹۳	۰/۱۸۹	۷/۹۳	۰/۸۴	-	۰/۹۱۷	۵	دسته ارزیابی
۵/۲۸	۰/۱۵۴	۱۸۶/۴۹	۰/۹۴۶۴	۰/۹۴۳۱	۰/۹۵۱۵	۲۲	حالت کلی



شکل ۶ - نمودار ویلیام.



شکل ۵ - نمودار خطای مقدار پیش‌بینی شده برحسب مقدار تجربی در مقیاس لگاریتمی.

از بررسی شکل ۶ می‌توان نتیجه گرفت که (۱) هیچ حلالی مقدار لوریج بالاتر از لوریج بحرانی ندارد. بنابراین مدل ارائه شده توسط هیچ یک از حلال‌ها به‌طور جدی تأثیر پذیر نیست که این بدین مفهوم است که همه داده‌ها سهم به تقریب یکسانی در تعیین مدل داشته‌اند، (۲) با توجه به اینکه همه داده‌ها در بازه‌ی کاربرد پذیری واقع شده‌اند، داده پرتی وجود ندارد، (۳) مولکول‌های سری ارزیابی به درستی انتخاب شده‌اند، چون پراکندگی مناسبی نسبت به مقدارهای خطا و مجموعه داده‌ها دیده می‌شود.

اعتبار سنجی خارجی نیز روی مدل انجام شده است. در ابتدا داده‌های تجربی ثابت هنری CO_2 در حلال‌هایی غیر از حلال‌های مورد استفاده در مدلسازی، از منابع گوناگون جمع‌آوری شده [۲۹، ۳] و سپس با استفاده از معادله (۳) مقدارهای ثابت هنری تخمین زده می‌شود. جدول ۴ مقدارهای به دست آمده را با مقدارهای تجربی مقایسه می‌کند و نشان می‌دهد که خطاهای نسبی گزارش شده در حد قابل پذیرش است و پیش‌بینی مدل حاضر، می‌تواند برای گونه‌های گوناگون حلال قابل اطمینان باشد.

توصیف‌کننده‌هایی که در یک مدل QSPR ظاهر می‌شوند، افزون بر این که باید مدلی با کیفیت و عملکرد بالا ایجاد کنند،

$$h = X(X^T X)^{-1} X^T \quad (12)$$

در فرمول بالا، X ماتریس توصیف‌کننده‌هاست و X^T هم ماتریس ترانپوز است. پارامتر دیگری به نام لوریج بحرانی وجود دارد که در معادله (۱۳) تعریف شده است.

$$h^* = 3(p+1)/n \quad (13)$$

p تعداد توصیف‌کننده و n تعداد داده‌ها را نشان می‌دهد. اگر مقدار لوریج برای یک حلال تازه بیش از حد لوریج بحرانی باشد، مدل نهایی به‌دست آمده در این کار برای آن حلال کاربرد پذیر نمی‌باشد. چگونگی تشخیص دادن پرت با استفاده از نمودار ویلیام به این شکل است که اگر مقدار لوریج برای ترکیبی بزرگتر از h^* بوده و خطای استاندارد آن بزرگتر از ۳ و کمتر از ۳- باشد، داده موردنظر داده پرت خواهد بود. مقدار لوریج بحرانی برای این مجموعه داده‌ها برابر با ۰/۵۲ بوده که به‌صورت خط چین عمودی نشان داده شده است. به ناحیه مستطیلی شده در شکل ۶ که بین محور عمودی نمودار، دو خط چین افقی در مختصات ۳ و ۳- و خط عمودی لوریج به‌وجود آمده است، دامنه کاربردپذیری می‌گویند.

جدول ۴ - اعتبار سنجی خارجی معادله ۳ در پیش‌بینی ثابت هنری گاز دی اکسید کربن در دمای ۲۹۸ کلوین.

مرجع	RD	AD	$\ln H_{pre}$ (MPa)	$\ln H_{exp}$ (MPa)	حلال
[۳]	۴/۴۵۵	۰/۰۸۲	۱/۷۵۵	۱/۸۲۷	دی متیل کربنات
[۲۹]	۱/۹۵۹	۰/۰۴۲	۲/۱۸۵	۲/۱۴۳	پروپیلن کربنات
[۲۹]	۲۴/۷۳	۰/۵۴۱	۱/۶۴۷	۲/۱۸۹	متیل سیانو استات
[۲۹]	۱۱/۱۴	۰/۲۱۹	۲/۱۸۵	۱/۹۶۶	ان - فرمیل مورفولین
[۲۹]	۲۳/۹۷	۰/۳۰۵	۰/۹۶۷	۱/۲۷۲	سلکسول

از اتم‌های واقع شده در رأس مولکول، شکسته می‌شود و این موضوع تعداد حفره‌های لازم برای مولکول‌های گاز را افزایش می‌دهد. در نتیجه با افزایش BAC، مقدار حلالیت گاز افزایش و ثابت هنری کاهش می‌یابد و علامت منفی این توصیف‌کننده در معادله هم تایید کننده این موضوع است.

نتیجه‌گیری

در این مطالعه، مدل‌های QSPR در دمای ۲۹۸ کلوین با استفاده از روش برازش خطی چندگانه برای بررسی مقدار ثابت هنری کربن دی اکسید در گروهی از حلال‌های ارایه شده است. تعداد گروه عاملی هیدروکسیل و اندیس محوری بالابان به‌عنوان توصیف‌کننده‌های مولکولی در مدل ظاهر شده‌اند که هم قابل تفسیر بوده و هم به سادگی قابل محاسبه هستند. میزان دقت و عملکرد مدل با استفاده از روش‌های گوناگون اعتبارسنجی داخلی و خارجی سنجیده شده است و همه نتیجه‌ها نشان داده‌اند که این مدل از قدرت بالایی برای تخمین ثابت هنری برخوردار است. به گونه‌ای که مقدار میانگین انحراف نسبی برابر با ۵/۳۸ است. همچنین توانایی تخمین مدل برای حلال‌هایی که از نظر ساختاری به طور کامل متفاوت از حلال‌های مورد استفاده در مدل‌سازی بوده‌اند، بررسی شد و در همه موارد، مقدار خطای مدل ناچیز بوده است. در نتیجه مدل موردنظر می‌تواند با اطمینان بالایی برای گروه‌های گوناگون حلال استفاده شود.

با توجه به نتیجه‌های به‌دست آمده، بهترین حلال برای جذب CO_2 حلالی است که در ساختار خود فاقد گروه عاملی هیدروکسیل باشد و زنجیره اتمی طولانی‌تری داشته باشد. این یافته‌ها می‌تواند به پژوهشگران کمک کند که با استفاده از حلال‌هایی با این ویژگی، فرایند حذف کربن دی اکسید از گاز طبیعی و محیط زیست با بازده بالاتری انجام گیرد.

باید ساده و قابل تفسیر هم باشند. در این کار، معادله (۳) نه تنها ساده بوده و دارای پارامترهای آماری قابل قبولی است؛ بلکه دارای دو متغیری است که این متغیرها در کار پژوهشگران دیگر به عنوان عوامل مؤثر در حلالیت گاز کربن دی اکسید در حلال‌ها به شمار می‌آیند. در این قسمت به تفسیر دو توصیف‌کننده مولکولی پرداخته می‌شود. پیش از این، اهمیت گروه عاملی هیدروکسیل در میزان جذب کربن دی اکسید در حلال‌های شیمیایی آمین، توسط پژوهشگران اثبات شده بود. سینگ و همکاران در بررسی ساختار حلال‌های آمینی بر میزان ظرفیت (حلالیت) این گاز به این نتیجه رسیدند که هر چه طول زنجیره بین آمین و گروه عاملی هیدروکسیل بیشتر شود، ظرفیت جذب گاز کربن دی اکسید در این حلال‌های آمینی بیشتر می‌شود [۳۰]. همچنین آن‌ها در کار دیگر خود دریافته‌اند که حتی مکان گروه عاملی هیدروکسیل در ساختار حلال‌های آمینی نیز تأثیر به‌سزایی در میزان ظرفیت این نوع حلال‌ها دارد [۳۱]. جذب گاز در حلال‌های فیزیکی در دو مرحله رخ می‌دهد. در مرحله اول، مولکول‌های حلال - حلال شکسته شده و حفره‌هایی تولید خواهد شد و در مرحله بعد، این حفره‌ها توسط مولکول‌های گاز پر می‌شود. زمانی که تعداد گروه عاملی هیدروکسیل افزایش می‌یابد، پیوندهای هیدروژنی بیشتری در حلال تشکیل می‌شود و این موضوع از ایجاد حفره‌های لازم جلوگیری می‌کند. در نتیجه منجر به کاهش حلالیت گاز و افزایش مقدار ثابت هنری خواهد شد؛ بنابراین انتظار می‌رود که این توصیف‌کننده با علامت مثبت در معادله حاضر شود. BAC توصیف‌کننده دیگری است که در مدل ظاهر شده و افزایش مقدار آن به شاخه‌دارتر شدن مولکول نسبت داده می‌شود. ولی از آنجایی که همه ترکیب‌ها به کار رفته در این مطالعه ساختار خطی دارند، افزایش این اندیس به معنای طولانی‌تر شدن زنجیره اتم‌هاست. بنابراین واضح است که هرچه مقدار این توصیف‌کننده افزایش یابد، یعنی تعداد بیشتری

قدردانی

نویسندگان از دانشکده مهندسی شیمی دانشگاه تهران برای همکاری با این کار سپاسگزارند.

فهرست نمادها

انحراف استاندارد

انحراف نسبی

اندیس محوری بالابان

ترانهاده ماتریس توصیف کننده ها

تعداد گروه عاملی هیدروکسیل

تعداد اتم های با یک همسایگی در مرحله گام

تعداد توصیف کننده ظاهر شده در مدل

تعداد داده های استفاده شده در ایجاد مدل

تعداد کل مراحل

ثابت هنری جزء i ام

جزء مولی جزء i در فاز مایع

مراجع

- [1] Bohloul, M.R., Vatani, A., Peyghambarzadeh, S.M., [Experimental and Theoretical Study of CO₂ Solubility in N-Methyl-2-Pyrrolidone \(NMP\)](#), *Fluid Phase Equilib*, **365**: 106-111 (2014).
- [2] Intergovernmental Panel on Climate Change, "Carbon Dioxide Capture and Storage", United Nations, 1st Edition, (2006).
- [3] Gui X., Tang Z.G., Fei W., [CO₂ Capture with Physical Solvent Dimethyl Carbonate at High Pressures](#), *J. Chem. Eng. Data*, **55**: 3736-3741 (2010).
- [4] Sandler S. I., "Chemical and Engineering Thermodynamics", (3rd ed.), New York: Wiley, (1999).
- [5] Gui X., Tang Z.G., Fei W., [Solubility of CO₂ in Alcohols, Glycols, Ethers, and Ketones at High Pressures from 288.15 to 318.15 K](#), *J. Chem. Eng. Data*, **56**: 2420-2429 (2011).
- [6] Altschuh J., Bruggemann R., Santl H., Eichinger G., Piringer O.G., [Henry's Law Constant for a Diverse Set of Organic Chemicals: Experimental Determination and Comparison of Estimation Methods](#), *Chemosphere*, **39**(11): 1871-1887 (1999).
- [7] Lin Sh.T., Sandler S.I., [Henry's Law Constant of Organic Compounds in Water from a Group Contribution Model with Multipole Corrections](#), *Chem. Eng. Sci.*, **57**: 2727-2733 (2002).
- [8] Majer V., Sedlbauer J., Bergin G., [Henry's Law Constant and Related Coefficients for Aqueous Hydrocarbons, CO₂ and H₂S Over a Wide Range of Temperature and Pressure](#), *Fluid Phase Equilib*, **272**: 65-74 (2008).

- [9] Jou F.Y., Desmukh R.D., Otto F.D., Mather A.E., **Vapor-Liquid Equilibria for Acid Gases and Lower Alkanes in Triethylene Glycol**, *Fluid Phase Equilib*, **36**: 121-140 (1987).
- [10] Lotfollahi, M.N., Baseri, H., Haghighi Asl A., **High Pressure Phase Equilibrium of (Solvent + Salt + CO₂) Systems by the Extended Peng-Robinson Equation of State**, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **27**(4): 97-105 (2008).
- [11] Uyan M., Sieder G., Ingram Th., Held Ch., **Predicting CO₂ Solubility in Aqueous N-methyldiethanolamine Solutions with ePC-SAFT**, *Fluid Phase Equilib*, **393**: 91-100 (2015).
- [۱۲] سید حمید حسینی، امیرعباس ایزدپناه، حسین رهیده، مدل سازی ترمودینامیکی حلالیت کربن دی اکسید در محلول آبی متیل دی اتانول آمین با استفاده از معادله حالت مکعبی به اضافه تجمعی (CPA)، *نشریه شیمی و مهندسی شیمی ایران*، (۴)۳۴: ۴۵ تا ۵۷ (۱۳۹۴).
- [13] Goharrokhi M., Taghikhani V., Ghotbi S., Safekordi A.A., Najibi H., **Correlation and Prediction of Solubility of CO₂ in Amine Aqueous Solutions**, *Iran. J. Chem. Chem. Eng. (IJCCE)*, **29**(1): 111-124 (2010).
- [14] Rayer A.V., Henni A., Tontiwachwuthikul P., **High Pressure Physical Solubility of Carbon Dioxide (CO₂) in Mixed Polyethylene Glycol Dimethyl Ethers (Genosorb 1753)**, *Can. J. Chem. Eng.*, **90**(3): 576-583 (2012).
- [15] Hansch C., Leo A., Exploring QSAR. "Fundamentals and Applications in Chemistry and Biology", ACS, Washington, D.C., (1995).
- [16] Todeschini R., Consonni V., "Molecular Descriptors for Chemoinformatics", 2nd ed., Wiley-VCH, Weinheim, (2009).
- [17] Momeni M., Riahi S., **Prediction of Amines Capacity for Carbon Dioxide Absorption in Gas Sweetening Processes**, *J. Nat. Gas. Sci. and Eng.*, **21**: 442-450 (2014).
- [18] Henni A., Tontiwachwuthikul P., Chakma A., **Solubilities of Carbon Dioxide in Polyethylene Glycol Ethers**, *Can. j. Chem. Eng.*, **83**: 358-361 (2005).
- [19] <http://www.hyper.com>
- [20] Frisch, Michael J., Nielsen, Alice B., Frisch, Aeleen (Eds.), "Gaussian 98", Gaussian Incorporated, 1998.
- [21] <http://www.taletе.mi.it>
- [22] Michalewicz, Z., "Genetic Algorithms + Data Structures = Evolution Programs", 3rd ed., Berlin: Springer-Verlag, (1998).
- [23] Jolliffe I.T., "Principal Component Analysis", Springer-Verlag, New York, (1986).
- [24] Todeschini R., Consonni V., "Handbook of Molecular Descriptors", Wiley VCH Verlag GmbH, 11, (2000).
- [25] Balaban, A.T., **XXXIV. Five New Topological Indices for the Branching of Tree-Like Graphs**, *Theor. Chim. Acta*, **53**: 355-375 (1979).
- [26] A. Golbraikh A., Tropsha A., **Beware of q²!**, *J. Mol. Graph. Model.*, **20**(4): 269-276 (2002).

- [27] Netzeva T.I., Worth A.P., Aldenberg T., Benigni R., Cronin M.T., Gramatica P., Yang C., [Current Status of Methods for Defining the Applicability Domain of \(Quantitative\) Structure – Activity Relationships](#), *ATLA*, **33**: 155-173 (2005).
- [28] Gramatica P., [Principles of QSAR Models Validation: Internal and External](#), *QSAR Comb. Sci.*, **26**(5): 694–701 (2007).
- [29] Xu Y., Schutte R.P., Hepler L.G., [Solubilities of Carbon Dioxide, Hydrogen Sulfide and Sulfur Dioxide in Physical Solvents](#), *Can. J. Chem. Eng.*, **70** (1992).
- [30] Singh P., Niederer J. PM., Versteeg G.F., [Structure and Activity Relationships for Amine Based CO₂ absorbents—I](#), *Int. J. Greenh. Gas. Con.*, **5-10** (2007).
- [31] Singh P., Niederer J. PM, Versteeg G.F., [Structure and activity Relationships for Amine Based CO₂ Absorbents-II](#), *Chem. Eng. Res. Des.*, **87**: 135-144 (2009).