





دانشکده شیمی

گروه شیمی فیزیک

عنوان :

پیش بینی ارتباط کمی ساختار - خاصیت (QSPR) ضریب خودنفوذی آلکان‌ها

دانشجو: زهرا شهابی

استاد راهنما:

دکتر زهرا کلانتر

استاد مشاور:

دکتر حسین نیکوفرد

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

بهمن ماه ۱۳۹۲

حال که به یاری پروردگار، این دوره از تحصیلات خود را به پایان رساندم، بر خود لازم می‌دانم از زحمات بی‌شائبه‌ی استاد فرهیخته و

تعهد نامه

اینجانب **زهرا شهبابی** دانشجوی دوره کارشناسی ارشد رشته شیمی فیزیک دانشکده شیمی دانشگاه صنعتی شاهرود نویسنده پایان نامه: **پیش بینی ارتباط کمی ساختار - خاصیت (QSPR) ضریب خودنفوذی آلکان‌ها تحت راهنمایی سرکار خانم دکتر کلانتر متعهد می‌شوم.**

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است

تاریخ: ۹۲/۱۱/۲۹

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

پروردگار مهربانم

هر چه دارم از لطف بی کران تو ست

تقدیم به:

پدرم که وجود پر مهرش را کم دارم

مادر عزیزتر از جانم

و همسرم که بهترین است

مَشْکُر و قدردانی:

حال که به یاری پروردگار، این دوره از تحصیلات خود را به پایان رساندم، بر خود لازم می‌دانم از زحمات بی‌شائبه‌ی استاد فریخته و توانمندم سرکار خانم دکتر کلاشکر که راهنمایی‌ها و نظرات ارزنده‌ی ایشان نقش مهمی در به ثمر رساندن این پروژه داشت، صمیمانه‌ی مَشْکُر کنم. بی‌اغراق بدون لطف و مساعدت ایشان این پروژه ناتمام می‌ماند. از استاد مشاور بزرگوارم جناب آقای دکتر نیکوفرد که خالصانه‌ی یاری ام نمودند نیز نهایت قدردانی را دارم.

به‌چنین از جناب آقای دکتر بابایی و سرکار خانم دکتر سیربادی که سخاوتمندانه‌ی هر چه می‌دانستند به من آموختند سپاس‌گذاری می‌کنم.

و در پایان بایستی از پروفیسور لودمان و پروفیسور حریس قدردانی نمایم که بزرگوارانه‌ی اطلاعات مورد نیاز مرا در اختیارم قرار دادند.

سلامتی و موفقیت همیشگی این بزرگواران را از دگاه یزدان پاک خواستارم.

چکیده

در این تحقیق، مطالعات ارتباط کمی ساختار - خاصیت (QSPR) با استفاده از مدل سازی خطی و غیرخطی برای پیش بینی ضریب خودنفوذی آلکان های خطی به کار گرفته شده است. داده ی مربوط به ضریب خودنفوذی، با استفاده از روش تحلیل مؤلفه ی اصلی (PCA) به چهار دسته ی آموزش، ارزیابی، تست و تست بیرونی دسته بندی گردید. توصیف کننده های ساختار مولکولی محاسبه شد و پس از حذف توصیف کننده های اضافی و افزودن دو پارامتر دما و فشار، مدل سازی خطی آغاز گردید. از دو روش رگرسیون مؤلفه ی اصلی (PCR) و رگرسیون کمترین مربعات جزئی (PLS) جهت مدل سازی خطی بهره گرفته شد که در هر دو روش ابتدا مؤلفه های اصلی ساخته و تعداد بهینه ی مؤلفه ها محاسبه شد. سپس مدل سازی با مؤلفه های اصلی بهینه انجام شد و در نهایت مدل های نهایی برای پیش بینی داده های مربوط به سری تست بیرونی به کار گرفته شدند. مقدار میانگین مربعات خطای مربوط به سری تست بیرونی برای مدل های PCR و PLS به ترتیب $0/6129$ و $0/5327$ حاصل گردید.

در مدل سازی غیرخطی، جهت جداسازی توصیف کننده های بهینه از دو روش رگرسیون مرحله ای (SR) و الگوریتم ژنتیک (GA) استفاده شد. توصیف کننده های منتخب به شبکه ی عصبی وارد شدند و مدل سازی برای هر دو روش انتخاب توصیف کننده، صورت گرفت. شبکه های ساخته شده پس از آموزش و بهینه سازی جهت پیش بینی داده های مربوط به سری تست بیرونی مورد استفاده

قرار گرفتند. مقدار میانگین مربعات خطای مربوط به سری تست بیرونی برای مدل‌های SR-ANN و GNN به ترتیب ۰/۰۳۷۹ و ۰/۰۲۱۹ به دست آمد. این نتایج نشان‌دهنده برتری روش‌های غیر خطی در مرحله انتخاب توصیف‌کننده و در مرحله مدل‌سازی می‌باشند.

کلمات کلیدی: ارتباط کمی ساختار-خاصیت، رگرسیون مؤلفه‌ای اصلی، رگرسیون کمترین مربعات جزئی، شبکه‌ی عصبی مصنوعی، ضریب خودنفوذی، رگرسیون مرحله‌ای، الگوریتم ژنتیک

عنوان مقالات مستخرج از این پایان‌نامه که در همایش‌ها ارائه شده است:

Artificial neural network modeling for self-diffusion prediction of gas alkanes using theoretically derived descriptors

نهمین همایش ملی شیمی پیام‌نور، دانشگاه پیام‌نور بهشهر.

QSPR study of self-diffusion for n-alkanes using GA-ANN

۹

Artificial neural network modeling for self-diffusion prediction of n-alkanes over a wide range of temperatures and pressures using theoretically derived descriptors

نوزدهمین سمینار شیمی تجزیه ایران، دانشگاه فردوسی مشهد.

فهرست مطالب

فصل اول: ضریب خودنفوذی

۱-۱- مقدمه:	۲
۱-۲- خواص انتقالی:	۳
۱-۲-۱- قانون عمومی انتقال:	۴
۱-۲-۲- نفوذ:	۵
۱-۲-۲-۱- انواع نفوذ مولکولی:	۶
۱-۳- روش‌های تئوری محاسبه‌ی ضریب خودنفوذی در مایعات:	۸
۱-۳-۱- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل کره‌ی سخت:	۸
۱-۳-۲- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل سلول مکعبی:	۹
۱-۳-۳- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل حجم آزاد:	۱۰
۱-۴- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل ساختمان ویژه:	۱۱
۱-۵- محاسبه‌ی ضریب خودنفوذی با استفاده از معادله‌ی استوکس-انیشیتین:	۱۲
۱-۶- محاسبه‌ی ضریب خودنفوذی با استفاده از معادله‌ی بولتزمن:	۱۴
۱-۴- استفاده از روش‌های تجربی جهت اندازه‌گیری ضریب خودنفوذی:	۱۴
۱-۵- روش طیف سنجی NMR اسپین-اکو:	۱۵

۱-۵-۱- اندازه‌گیری ضریب خودنفوذی مایعات با استفاده از طیف سنجی NMR اسپین-اکو با میدان ثابت ۱۸

۱-۵-۲- اندازه‌گیری ضریب خودنفوذی مایعات با استفاده از طیف سنجی NMR اسپین-اکو با میدان متناوب ۱۹

فصل دوم: بخش تئوری

۲-۱- کمومتریکس..... ۲۲

۲-۲- رابطه‌ی ساختار- خاصیت یا ساختار- فعالیت..... ۲۲

۳-۲- مراحل انجام QSPR..... ۲۳

۲-۳-۱- انتخاب سری مولکولی..... ۲۳

۲-۳-۲- محاسبه و انتخاب توصیف‌کننده‌ها..... ۲۳

۱-۲-۳-۲- محاسبه‌ی توصیف‌کننده‌ها..... ۲۴

۲-۲-۳-۲- کاهش تعداد توصیف‌کننده‌ها..... ۲۴

۳-۳-۲- تقسیم بندی داده‌ها..... ۲۵

۱-۳-۳-۲- تقسیم بندی داده‌ها به روش تحلیل مؤلفه‌ی اصلی (PCA)..... ۲۵

۴-۳-۲- انتخاب بهترین توصیف‌کننده‌ها..... ۲۶

۱-۴-۳-۲- انتخاب بهترین توصیف‌کننده‌ها توسط رگرسیون مرحله‌ای (SR)..... ۲۶

۲-۴-۳-۲- انتخاب بهترین توصیف‌کننده‌ها توسط الگوریتم ژنتیک..... ۳۰

۲-۴-۳-۲- الف: الگوریتم ژنتیک..... ۳۱

۲-۴-۳-۲- ب: عملگرهای الگوریتم ژنتیک..... ۳۱

۲-۴-۳-۲- ج: نقاط قوت الگوریتم ژنتیک..... ۳۲

۵-۳-۲- مدل‌سازی..... ۳۳

۱-۵-۳-۲- رگرسیون خطی چندگانه (MLR)..... ۳۳

۲-۵-۳-۲- رگرسیون مؤلفه‌ی اصلی (PCR)..... ۳۴

۳۶	۳-۵-۳-۲- رگرسیون کمترین مربعات جزئی (PLS).....
۳۸	۳-۵-۴-۲- شبکه‌ی عصبی مصنوعی.....
۳۸	۳-۵-۴-۲- الف: ساختار شبکه‌ی عصبی مصنوعی.....
۴۰	۳-۵-۴-۲- ب: شبکه‌های عصبی با تغذیه‌ی پیشرو.....
۴۰	۳-۵-۴-۲- ج: آموزش شبکه‌ی عصبی مصنوعی.....
۴۳	۳-۶-۳-۲- ارزیابی اعتبار مدل‌ها.....

فصل سوم: بخش محاسبات

۴۶	۳-۱-۱- مدل‌سازی ضریب خودنفوذی آلکان‌های خطی.....
۴۶	۳-۱-۱- انتخاب سری داده‌ها.....
۴۶	۳-۱-۲- رسم و بهینه‌سازی ساختار مولکول‌ها.....
۴۸	۳-۱-۳- محاسبه‌ی توصیف‌کننده‌ها.....
۴۸	۳-۱-۴- کاهش تعداد توصیف‌کننده‌ها.....
۴۹	۳-۱-۵- خود مقیاس‌سازی داده‌ها.....
۵۰	۳-۱-۶- دسته‌بندی داده‌ها.....
۵۲	۳-۱-۷- انتخاب بهترین توصیف‌کننده‌ها.....
۵۲	۳-۱-۷-۱- انتخاب بهترین توصیف‌کننده‌ها به روش رگرسیون مرحله‌ای (SR).....
۵۵	۳-۱-۷-۲- انتخاب بهترین توصیف‌کننده‌ها با استفاده از الگوریتم ژنتیک.....
۵۹	۳-۱-۸- مدل‌سازی.....
۵۹	۳-۱-۸-۱- مدل‌سازی خطی با استفاده از روش PCR.....
۶۲	۳-۱-۸-۲- مدل‌سازی خطی با استفاده از روش PLS.....
۶۵	۳-۱-۸-۳- مدل‌سازی غیرخطی با استفاده از روش ANN.....

۱-۳-۸-۱-۳- بهینه‌سازی مدل شبکه عصبی مصنوعی با توصیف‌کننده‌های انتخابی به روش رگرسیون	۶۶
۲-۳-۸-۱-۳- بهینه‌سازی مدل شبکه‌ی عصبی مصنوعی با توصیف‌کننده‌های انتخابی به روش الگوریتم ژنتیک	۷۴
۹-۱-۳- ارزیابی مدل‌های خطی و غیر خطی بهینه شده	۸۳
۱-۹-۱-۳- ارزیابی مدل‌های خطی و غیر خطی با استفاده از پارامترهای آماری	۸۳
۲-۹-۱-۳- ارزیابی مدل‌های خطی و غیر خطی با استفاده از نمودار برگشتی	۸۴
۳-۹-۱-۳- ارزیابی مدل‌های خطی و غیر خطی با استفاده از نمودار خطای باقیمانده	۸۹
۴-۹-۱-۳- ارزیابی مدل‌های خطی و غیر خطی با استفاده از اعتبارسنجی بیرونی	۹۴
۵-۹-۱-۳- ارزیابی مدل‌های غیر خطی با استفاده از آزمون χ^2 تصادفی	۹۹
۱۰-۱-۳- بررسی ارتباط توصیف‌کننده‌های وارد شده در مدل با ضریب خودنفوذی	۱۰۰
۱-۱۰-۱-۳- توصیف‌کننده‌های هندسی	۱۰۰
۲-۱۰-۱-۳- Topological توصیف‌کننده‌های	۱۰۱
۳-۱۰-۱-۳- 3D- MORSE توصیف‌کننده‌های	۱۰۴
۱۱-۱-۳- بررسی میزان مشارکت توصیف‌کننده‌های منتخب شبکه‌ی عصبی	۱۰۵
۲-۳- نتیجه‌گیری	۱۰۷
۳-۳- آینده‌نگری	۱۰۸
پیوست	۱۰۹
فهرست منابع	۱۴۷

فهرست جداول

- جدول (۱-۳): محدوده‌ی دما و فشار و تعداد داده‌ی استفاده شده در مدل‌سازی برای هر آلکان به همراه مرجع مورد استفاده برای هریک از آن‌ها..... ۴۷
- جدول (۲-۳): مقایسه‌ی پارامترهای آماری مدل‌های ساخته شده توسط روش SR..... ۵۳
- جدول (۳-۳): نماد، گروه، نام کامل، و اثر متوسط توصیف‌کننده‌های انتخاب شده توسط روش SR..... ۵۴
- جدول (۴-۳): ماتریس همبستگی میان توصیف‌کننده‌های انتخاب شده توسط روش SR..... ۵۴
- جدول (۵-۳): مقادیر توصیف‌کننده‌های انتخاب شده توسط روش SR..... ۵۵
- جدول (۶-۳): توصیف‌کننده‌های منتخب برای ۱۲ بار اجرای الگوریتم ژنتیک..... ۵۶
- جدول (۷-۳): نماد، گروه، نام کامل و اثر متوسط توصیف‌کننده‌های انتخاب شده توسط روش GA..... ۵۷
- جدول (۸-۳): ماتریس همبستگی میان توصیف‌کننده‌های انتخاب شده توسط روش GA..... ۵۷
- جدول (۹-۳): مقادیر توصیف‌کننده‌های انتخاب شده توسط روش GA..... ۵۸
- جدول (۱۰-۳): مقدار میانگین مربعات خطای سری ارزیابی برای مدل‌های اولیه PCR..... ۶۱
- جدول (۱۱-۳): ضریب هر PC در مدل بهینه‌ی PCR و مقدار ویژه‌ی مربوط به آن..... ۶۲
- جدول (۱۲-۳): مقدار میانگین مربعات خطای سری ارزیابی برای مدل‌های اولیه PLS..... ۶۴
- جدول (۱۳-۳): ضریب هر LV در مدل بهینه‌ی PLS و واریانس مربوط به آن..... ۶۴
- جدول (۱۴-۳): مقادیر ضریب همبستگی توصیف‌کننده‌های انتخاب شده به روش SR با ضریب خودنفوذی..... ۶۶
- جدول (۱۵-۳): توابع و پارامترهای بهینه‌ی شبکه‌های طراحی شده با توصیف‌کننده‌های حاصل از روش SR..... ۷۲
- جدول (۱۶-۳): مقدار MSE سری ارزیابی برای تعداد دوره‌های آموزش متفاوت با توصیف‌کننده‌های منتخب روش SR..... ۷۲
- جدول (۱۷-۳): توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی با توصیف‌کننده‌های منتخب روش SR..... ۷۴
- جدول (۱۸-۳): مقادیر ضریب همبستگی توصیف‌کننده‌های منتخب روش GA با ضریب خودنفوذی..... ۷۴
- جدول (۱۹-۳): توابع و پارامترهای بهینه‌ی شبکه‌های طراحی شده با توصیف‌کننده‌های حاصل از الگوریتم ژنتیک..... ۸۰

جدول (۳-۲۰): مقدار MSE سری ارزیابی برای تعداد دورهای آموزش متفاوت با توصیف‌کننده‌های منتخب روش GA ۸۱

جدول (۳-۲۱): توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی با توصیف‌کننده‌های منتخب روش GA ۸۲

جدول (۳-۲۲): پارامترهای آماری مربوط به سری‌های ارزیابی و تست درونی ۸۴

جدول (۳-۲۳): مقادیر R^2 به دست آمده برای سری ارزیابی و تست از آزمون Y تصادفی در دو مدل SR-AAN و GA-

ANN ۹۹

جدول (۱-پ): سری داده‌های مربوط به ضریب خودنفوذی آلکان‌های خطی به همراه دما و فشار اندازه‌گیری ۱۰۹

جدول (۲-پ): مقایسه‌ی مقادیر تجربی و مقادیر پیش‌بینی شده‌ی ضریب خودنفوذی برای ترکیبات سری تست بیرونی ۱۴۲

فهرست شکل‌ها

- شکل (۱-۱): مراحل NMR اسپین-اکو ۱۸
- شکل (۱-۲): نمایی از یک شبکه‌ی عصبی مصنوعی ۳۹
- شکل (۲-۲): نمایی از شبکه‌ی عصبی مصنوعی با تغذیه‌ی پیشرو ۴۰
- شکل (۱-۳): پراکندگی داده‌ها بر روی نمودار PCA ۵۱
- شکل (۲-۳): نمودار تغییرات مقدار میانگین مربعات خطا بر حسب تعداد مؤلفه‌های اصلی برای سری ارزیابی ۶۰
- شکل (۳-۳): نمودار تغییرات مقدار میانگین مربعات خطا بر حسب تعداد متغیرهای پنهان مدل PLS ۶۳
- شکل (۴-۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف‌کننده‌های منتخب روش SR ۶۸
- شکل (۵-۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لومبرگ-مارکوارت با توصیف‌کننده‌های منتخب روش SR ۶۹
- شکل (۶-۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف‌کننده‌های منتخب روش SR ۷۰
- شکل (۷-۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد،

برای شبکه‌ی عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف‌کننده‌های منتخب روش SR..... ۷۱

شکل (۳-۸): نمودار مقدار MSE سری ارزیابی بر حسب تعداد دورهای آموزش با توصیف‌کننده‌های منتخب روش SR ۷۳
شکل (۳-۹): نمودار میانگین مربعات خطا حاصل از سری ارزیابی بر حسب پارامتر μ برای شبکه با توصیف‌کننده‌های منتخب روش SR..... ۷۳

شکل (۳-۱۰): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف‌کننده‌های منتخب روش GA..... ۷۶

شکل (۳-۱۱): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف‌کننده‌های منتخب روش GA..... ۷۷

شکل (۳-۱۲): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف‌کننده‌های منتخب روش GA..... ۷۸

شکل (۳-۱۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه‌ی عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف‌کننده‌های منتخب روش GA..... ۷۹

شکل (۳-۱۴): نمودار مقدار MSE سری ارزیابی بر حسب تعداد دورهای آموزش با توصیف‌کننده‌های منتخب روش GA ۸۱
شکل (۳-۱۵): نمودار میانگین مربعات خطای سری ارزیابی بر حسب پارامتر μ برای شبکه با توصیف‌کننده‌های منتخب روش GA..... ۸۲

شکل (۳-۱۶): نمودار برگشتی حاصل از روش PCR برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۸۵

شکل (۳-۱۷): نمودار برگشتی حاصل از روش PLS برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۸۶

- شکل (۳-۱۸): نمودار برگشتی حاصل از روش SR-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۸۷
- شکل (۳-۱۹): نمودار برگشتی حاصل از روش GA-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۸۸
- شکل (۳-۲۰): نمودار خطای باقیمانده حاصل از روش PCR برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۹۰
- شکل (۳-۲۱): نمودار خطای باقیمانده حاصل از روش PLS برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۹۱
- شکل (۳-۲۲): نمودار خطای باقیمانده حاصل از روش SR-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۹۲
- شکل (۳-۲۳): نمودار خطای باقیمانده حاصل از روش GA-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی..... ۹۳
- شکل (۳-۲۴): (الف) نمودار برگشتی و (ب) نمودار خطای باقیمانده حاصل از روش PCR برای سری تست بیرونی..... ۹۵
- شکل (۳-۲۵): (الف) نمودار برگشتی و (ب) نمودار خطای باقیمانده حاصل از روش PLS برای سری تست بیرونی..... ۹۶
- شکل (۳-۲۶): (الف) نمودار برگشتی و (ب) نمودار خطای باقیمانده حاصل از روش SR-ANN برای سری تست بیرونی..... ۹۷
- شکل (۳-۲۷): (الف) نمودار برگشتی و (ب) نمودار خطای باقیمانده حاصل از روش GA-ANN برای سری تست بیرونی..... ۹۸
- شکل (۳-۲۸): درصد مشارکت توصیف‌کننده‌ها در مدل SR-ANN..... ۱۰۶
- شکل (۳-۲۹): درصد مشارکت توصیف‌کننده‌ها در مدل GA-ANN..... ۱۰۶

فصل اول

ضریب خودنفوذی

۱-۱- مقدمه:

خواص انتقالی^۱ کمیت‌های بسیار تأثیرگذاری در اکثر زمینه‌های علوم مهندسی می‌باشند. دانش دقیق از فرآیندهای انتقالی سیالات برای تعیین توابع انرژی پتانسیل بین مولکولی و طراحی بهینه‌ی دستگاه‌های فرآیندهای شیمیایی بسیار حائز اهمیت است. به علاوه خواص انتقالی در تعیین سرعت پخش آلاینده‌ها در محیط نیز از اهمیت ویژه‌ای برخوردارند.

در بین شاخه‌های مختلف علوم مهندسی، محاسبات مهندسی شیمی، بیشترین ارتباط را با خواص انتقالی سیالات دارند. اما روش‌های موجود برای محاسبه‌ی خواص انتقالی سیالات (به غیر از گازهای تحت فشار کم تا متوسط) در مقایسه با روش‌های محاسبه برای دیگر خواص ترموفیزیکی ناتوان‌تر می‌نمایند و این مشکل در مورد مایعات، بزرگتر است. موانعی که سد راه این پیشرفت بوده‌اند دارای دو جنبه هستند: نخست مشکلات اساسی که در اندازه‌گیری دقیق این خواص وجود دارد و دوم پیچیدگی ساختاری که تجزیه و تحلیل تئوریک آن‌ها را با مشکل مواجه می‌سازد.

چندین دهه است که تئوری‌های مختلفی با فرضیات اولیه‌ی بسیار متنوع برای تشریح حالت مایع و خواص انتقالی آن ارائه گردیده است اما در مجموع می‌توان گفت به علت ناممکن بودن حل ریاضی مسأله، به جز با در نظر گرفتن فرضیات ساده کننده‌ی بسیار، در عمل تنها در موارد خاص نتایج حاصل از تئوری‌ها دقت نسبتاً کافی داشته‌اند. به عبارت دیگر تاکنون تئوری از رسیدن به جواب قابل قبولی برای این مسئله ناتوان بوده است. مسیری که به ناچار در زمینه‌ی یافتن خواص انتقالی مایعات برای کاربردهای مهندسی طی گردیده، استفاده از روابط تجربی یا نیمه تجربی بوده است. معمولاً اساس و ریشه‌ی این روابط معادلات نظری است که برای حالتی خاص صدق می‌کند. محققین سعی کرده‌اند با شاخ و برگ دادن به این روابط و افزودن پارامترهای تجربی به آن، به روابطی که در عمل برای مایعات مختلف قابل استفاده باشند برسند.

1- Transport properties

۱-۲- خواص انتقالی

هرگاه سیالی از لحاظ یکی از خواص ترکیب، دما و یا سرعت مولکول‌ها غیر یکنواخت باشد، پدیده‌های انتقالی برای یکنواخت شدن آن خاصیت در سیستم اتفاق می‌افتند. به طور کلی انتقال یک خاصیت معین مثل جرم، اندازه‌ی حرکت و یا انرژی از یک قسمت سیستم به قسمت دیگر را فرآیند انتقالی نامیده و میزان انتقال آن را با شار آن خاصیت بیان می‌کنند. شار مقداری از خاصیت است که از واحد سطح در واحد زمان می‌گذرد. مشاهدات تجربی خواص انتقالی نشان می‌دهد که شار یک خاصیت (J) با گرادیان^۱ خاصیت مربوطه (W) متناسب است:

$$J \propto \nabla W \quad (1-1)$$

به طور کلی پدیده‌های هدایت گرمایی^۲، ویسکوزیته^۳ و نفوذ^۴، خواص انتقالی نامیده می‌شوند چون در تمام این موارد یک خاصیت فیزیکی توسط سیال منتقل می‌شود. چنانچه گرادیان دمایی در درون سیستم وجود داشته باشد سیستم در حال تعادل گرمایی نیست و انرژی گرمایی در جهت گرادیان دما منتقل می‌شود. به این پدیده هدایت گرمایی می‌گویند.

اگر اختلاف غلظت مواد در نواحی مختلف یک محلول وجود داشته باشد، سیستم در حال تعادل مادی نیست و ماده آن‌قدر در جهت گرادیان غلظت منتقل می‌شود تا ترکیب یکنواختی حاصل شود، این پدیده نفوذ نام دارد.

پدیده‌ی ویسکوزیته که ناشی از انتقال اندازه حرکت خطی از میان یک سیال تحت گرادیان سرعت است نیز، یکی دیگر از خواص انتقالی سیالات محسوب می‌شود.

یکی دیگر از خواص انتقالی هدایت الکتریکی است. در پدیده‌ی هدایت الکتریکی، در اثر اعمال اختلاف پتانسیل الکتریکی به ذرات باردار الکتریکی (الکترون‌ها و یون‌ها) نیرو وارد می‌شود و این

1- Gradient
2- Thermal conductivity
3- Viscosity
4- Diffusion

ذرات باردار، تحت تأثیر گرادیان پتانسیل الکتریکی حرکت می‌کنند و به این ترتیب جریان الکتریکی ایجاد می‌شود. هریک از خواص انتقالی به متغیرهای ترمودینامیکی دما، فشار و چگالی وابسته‌اند [۱].

۱-۲-۱- قانون عمومی انتقال

همه قوانینی که بیانگر جریان سیال، رسانایی الکتریکی، هدایت گرمایی و نفوذ هستند از یک قانون کلی تابعیت می‌کنند که در آن سرعت انتقال متناسب با مشتق مکانی^۱ (گرادیان) یک کمیت فیزیکی است. به این قانون، قانون عمومی انتقال می‌گویند.

به سرعت انتقال کمیت فیزیکی W ، عمود بر جهت جریان سیال، شار یا جریان (J) اطلاق می‌شود که در آن:

$$J = \frac{1}{A} \left(\frac{\partial W}{\partial t} \right) \quad (۲-۱)$$

در این معادله، A سطح عمود بر جریان و t زمان است.

نیروی تعمیم یافته برای انتقال کمیت فیزیکی (X_W) مرتبط با منفی گرادیان کمیت فیزیکی ثانوی Y ، در جهت جریان (x) می‌باشد.

$$X_W = - \left(\frac{\partial Y}{\partial x} \right) \quad (۳-۱)$$

شار و نیروی تعمیم یافته از طریق قانون عمومی انتقال به یکدیگر مرتبط‌اند.

$$J_W = L_W X_W \quad (۴-۱)$$

که در آن L_W ثابت تناسب موسوم به ضریب پدیده شناختی است [۲].

1- Spatial derivative

۱-۲-۲- نفوذ

اگر دو نمونه که غلظت‌های متفاوتی از یک ذره دارند با یک صفحه‌ی متخلخل از هم جدا شوند، پدیده‌ی نفوذ در این سیستم مشاهده خواهد شد. تحت این شرایط انتقال ذره‌ی مورد نظر آن‌قدر از نمونه‌ی با غلظت بالاتر به نمونه‌ی با غلظت پایین‌تر ادامه می‌یابد تا غلظت ذره مورد نظر در دو نمونه یکسان شود [۳]. بنابراین، نفوذ شامل مخلوط شدن مولکول‌های دو نمونه از طریق حرکت‌های گرمایی تصادفی و برخورد مولکول‌ها تا رسیدن به یک ترکیب یکنواخت است. مولکول‌های هر نمونه به خاطر داشتن انرژی حرارتی دارای حرکات نامنظم می‌باشند. وجود این حرکات نامنظم باعث می‌شود مولکول‌های یک نمونه درون مولکول‌های نمونه‌ی دیگر نفوذ کنند. به علت برخوردهای متوالی، مقدار پیشروی مولکول در یک جهت معین نسبت به مجموع پیشروی که مولکول در جهت‌های مختلف دارد بسیار کوچکتر بوده و همین امر سبب می‌شود که نفوذ مولکولی پدیده‌ی بسیار کندی باشد. با این وجود با توجه به اصول نظریه‌ی جنبشی می‌توان با کم کردن فشار که منجر به کم شدن تعداد برخوردهای مولکولی می‌شود و افزایش درجه حرارت که منجر به ازدیاد سرعت متوسط مولکولی می‌گردد، سرعت این پدیده‌ی مولکولی را بالا برد. برخوردهای مولکولی متوالی مهم‌ترین دلیل کند بودن نفوذ مولکولی است [۴].

وقتی یک فاز گاز و یک فاز مایع که در تعادل ترمودینامیکی نیستند در تماس نزدیک با یکدیگر قرار داده شوند، ممکن است یک یا چند جزء به‌وسیله‌ی مکانیسم نفوذ مولکولی از فاز گاز به درون فاز مایع انتقال یابند یا برعکس. انتقال جرم ناشی از نفوذ مولکولی یک مکانیسم فیزیکی پایه‌ای است که در بسیاری از رشته‌های علمی مانند مهندسی خاک، مهندسی نفت، مهندسی شیمی، بیوتکنولوژی و مهندسی هسته‌ای کاربرد دارد. در سال‌های اخیر مسئله‌ی انتقال جرم به‌وسیله‌ی نفوذ مولکولی در مخازن نفتی، به خصوص در فرآیند تزریق گاز به داخل مخازن اهمیت زیادی پیدا کرده و مطالعات زیادی توسط شرکت‌های بزرگ در این زمینه صورت گرفته است [۵].

۱-۲-۲-۱- انواع نفوذ مولکولی

همان‌طور که گفته شد چنانچه غلظت اجزاء در نقاط مختلف یک محیط متفاوت باشد به علت نفوذ مولکولی این اختلاف به تدریج از بین می‌رود و در نتیجه محیطی با غلظت یکنواخت حاصل می‌گردد. شدت نفوذ مولکولی در یک محیط را می‌توان به وسیله‌ی مخلوط کردن، به هم زدن و ایجاد حرکت در آن محیط تسریع کرد. اگر به این صورت در محلول حرکت درهم ایجاد شود علاوه بر نفوذ مولکولی، انتقال جرم توسط چرخانه‌ها نیز صورت می‌گیرد که به این نوع انتقال جرم، نفوذ چرخانه‌ای^۱ یا نفوذ درهم^۲ گفته می‌شود. عامل اصلی نفوذ در سیالات ساکن یا سیالاتی که دارای حرکت آرام هستند، نفوذ مولکولی است. برعکس در حالتی که سیال دارای حرکات درهم است عامل اصلی انتقال جرم علی‌رغم وجود نفوذ مولکولی، نفوذ چرخانه‌ای است [۴].

سرعت نفوذ در جهت محور x با استفاده از قانون اول فیک به دست می‌آید [۶].

$$\frac{dN^*}{dt} = -DA \frac{\partial C}{\partial x} \quad (5-1)$$

علامت منفی در این رابطه، مشخص کننده‌ی نفوذ در جهتی است که اختلاف غلظت کم می‌شود. از معادله‌ی بالا می‌توان نتیجه گرفت که واحد نفوذ در سیستم cgs برابر cm^2/s و در سیستم SI برابر m^2/s می‌باشد [۴].

در یک سیستم دوتایی از اجزای A و B که در آن گرادیان غلظتی از جزء A در سیال وجود دارد، جزء A در جهت کاهش غلظت نفوذ می‌کند. در این حالت ضریب نفوذ جزء A در سیستم دوتایی A و B را ضریب نفوذ دوتایی^۳ نامیده و با D_{AB} نمایش می‌دهند که به وسیله‌ی قانون فیک چنین تعریف می‌شود:

1- Eddy diffusion
2- Turbulent diffusion
3- Mutual diffusion

$$J_{AB} = -D_{AB} \frac{\partial C_A}{\partial x} = -CD_{AB} \frac{\partial X_{WA}}{\partial x} \quad (6-1)$$

که در آن J_{AB} شار مولی جزء A در جهت x بوده و $\frac{dC_A}{dx}$ گرادیان غلظت در جهت x می‌باشد. C غلظت مولی مخلوط و X_{WA} جزء مولی ماده‌ی A در مخلوط می‌باشد. می‌توان نشان داد که در سیستم‌های دوتایی D_{AB} برابر D_{BA} است [۴].

در فرآیند نفوذ مولکولی، در حالت خاصی که محیط به صورت ساکن بوده و واکنش شیمیایی نیز وجود نداشته باشد، تغییرات غلظت با زمان و نسبت به فاصله، از قانون دوم فیک که به آن معادله‌ی نفوذ نیز گفته می‌شود، به دست می‌آید. قانون نفوذ مطابق با رابطه‌ی (۷-۱) می‌باشد [۴].

$$\frac{\partial C_i}{\partial t} = D \left[\frac{\partial^2 C_i}{\partial x^2} + \frac{\partial^2 C_i}{\partial y^2} + \frac{\partial^2 C_i}{\partial z^2} \right] \quad (7-1)$$

در یک ماده‌ی خالص تحرک ذرات مختلف با حرکات اتفاقی توأم است و احتمال حضور مولکول‌ها در هر نقطه از فضای اشغال شده توسط ماده برابر است. برای بررسی حرکت یک مولکول در ماده‌ی خالص باید مولکول نشان‌دار شود. از حرکت مولکول نشان‌دار شده در میان سایر مولکول‌ها نوعی از نفوذ به وجود می‌آید که به آن ضریب خودنفوذی^۱ (D_{AA}) گفته می‌شود. به عبارت دیگر ضریب خودنفوذی، نفوذ یک جزء درون مولکول‌های خودش است. ضریب خودنفوذی با جایگذاری غلظت مولی جزء علامت خورده به جای مقدار C_i در قانون دوم فیک، محاسبه می‌گردد [۷].

به همین ترتیب در یک سیستم چند جزئی کاملاً یکنواخت اگر بعضی از مولکول‌ها نشان‌دار شوند، گرادیان غلظتی بین اجزاء نشان‌دار و سایر اجزاء به وجود می‌آید. به این نوع از نفوذ، درون نفوذی^۲ (D_A^*) یا نفوذ ردیاب^۳ گفته می‌شود. هرگاه کسر مولی ماده‌ی A در محلول به سمت ۱ میل

1- Self-diffusion coefficient

2- Intra diffusion

3- Tracer diffusion

کند ($X_A \rightarrow 1$)، مقدار درون نفوذی ماده‌ی A به سمت ضریب خودنفوذی ماده‌ی A میل خواهد کرد
 $[7] (D_A^* \rightarrow D_{AA})$.

با توجه به گستردگی انواع نفوذ مولکولی و تئوری‌های مربوط به هریک از آن‌ها، در این بخش تنها به تئوری‌های مربوط به خاصیت مورد بررسی در این تحقیق یعنی ضریب خودنفوذی پرداخته شده است.

۱-۳- روش‌های تئوری محاسبه‌ی ضریب خودنفوذی در مایعات

محاسبه‌ی ضریب خودنفوذی در مایعات همانند دیگر خواص فیزیکی مایعات، دقت کمتری نسبت به گازها دارد. این امر عمدتاً به علت نبود یک تئوری کامل برای مایعات است. مدل‌های مختلفی برای بررسی ضریب خودنفوذی مایعات ارائه شده است که در بخش‌های عمده‌ی این مدل‌ها مایعات گاز فشرده و یا جامد بدون نظم در نظر گرفته شده‌اند. بیشتر مدل‌ها در سه گروه مدل کره‌ی سخت^۱، مدل سلول مکعبی^۲ و مدل حجم آزاد^۳ قرار دارند. سایر مدل‌ها نیز معمولاً ترکیبی از مدل‌های ذکر شده هستند. در ادامه به چند مدل و معادله‌ی مربوط به ضریب خودنفوذی پرداخته می‌شود.

۱-۳-۱- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل کره‌ی سخت

در مدل کره‌ی سخت مولکول‌ها ذراتی کروی، سخت و بدون جاذبه فرض می‌شوند که دارای حرکات اتفاقی هستند. لانگت هیگینز^۴ و پاپل^۵ در سال ۱۹۵۶ با استفاده از فرضیات کره‌ی سخت

1- Hard sphere model
 2- Cubic cell model
 3- Free volume model
 4- Longuet Higgins
 5- Pople

مدلی برای فرآیندهای انتقالی در سیالات فشرده ارائه کردند که قابلیت به دست آوردن معادلاتی برای ضریب خودنفوذی، ویسکوزیته و ضریب هدایت گرمایی را داشت. معادله‌ی (۸-۱) رابطه‌ی به دست آمده از مدل کره‌ی سخت برای ضریب خودنفوذی را نشان می‌دهد.

$$D_{AA} = \frac{\sigma}{2} \cdot \frac{\pi RT}{M} \left(\frac{PV}{RT} - 1 \right)^{-1} \quad (8-1)$$

که در آن σ قطر مولکولی، R ثابت گازها و M جرم مولکولی است.

از مقایسه‌ی مقادیر تجربی و مقادیر به دست آمده توسط این رابطه مشخص شده است که این رابطه وابستگی ضریب خودنفوذی با تغییرات دما را به درستی نشان نمی‌دهد [۷ و ۸].

۱-۳-۲- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل سلول مکعبی

در این مدل فرض بر آن است که مولکول اغلب در سلولی که توسط مولکول‌های همسایه (که در وضعیت تعادلیشان قرار دارند) به وجود آمده محبوس است. همچنین برای ساده سازی محاسبه‌ی فضای آزاد بین مولکولی، این فضای آزاد، کروی در نظر گرفته شده است. این مدل از تأثیر تابع پتانسیل لنارد-جونز^۱ بر مدل سلول ساده^۲ به وجود آمده و به مدل لنارد-جونز نیز معروف است.

هوگتون^۳ در سال ۱۹۶۴ با استفاده از مدل سلول، مدلی برای نفوذ و ویسکوزیته ارائه نمود که در آن نفوذ و ویسکوزیته به صورت یک فرآیند مداوم و اتفاقی فرض شده‌اند. در این مدل یک سلول مکعبی با حجم λ^3 در نظر گرفته می‌شود که مولکول مرکزی در آن محبوس است. سپس تأثیر متقابل نزدیک‌ترین همسایه‌ها با مولکول مرکزی بررسی می‌گردد. بر اساس اندازه‌گیری‌های اشعه‌ی ایکس،

1- Lennard Jones

2- Simple cell method

3-Houghton

مقدار اندازه‌ی سلول، λ ، به صورت رابطه‌ی (۹-۱) تخمین زده می‌شود.

$$\lambda = 2(V/N)^{1/3} \quad (9-1)$$

که در این رابطه V حجم مولی و N عدد آووگادرو می‌باشد.

معادله‌ی نهایی برای ضریب خودنفوذی به صورت رابطه‌ی (۱۰-۱) به دست آمده است.

$$D_{AA} = \frac{RT\rho}{6\eta M} \left(\frac{V}{N} \right)^{2/3} \quad (10-1)$$

که در این معادله R ثابت گازها، η ویسکوزیته، M جرم مولکولی و ρ دانسیته است. این معادله توسط هوگتون برای ۲۵ مایع مختلف در دمای ۵ درجه‌ی سانتیگراد مورد بررسی قرار گرفته و متوسط خطای ۹٪ توسط وی گزارش شده است [۷ و ۹].

۱-۳-۳- محاسبه‌ی ضریب خودنفوذی با استفاده از مدل حجم آزاد

یکی از نقص‌های مدل سلول، عدم توانایی آن در برآورد صحیحی از فضاهای خالی بین مولکولی است. در مدل فضای آزاد حجم آزاد بین مولکولی به تعدادی سلول تقسیم می‌شود که تعداد آن‌ها از تعداد مولکول‌ها بیشتر است و در نتیجه امکان در نظر گرفتن حفره‌هایی بین مولکولی وجود دارد. در نظر گرفتن سلول‌های خالی در شبکه‌ی مولکولی مایعات واقعیت‌های تجربی مانند نفوذ مولکولی را توجیه می‌کند. اولین بار نیز پدیده‌ی نفوذ مولکولی، آیرینگ^۱ را به سمت لزوم در نظر گرفتن حفره‌ها در مایعات راهنمایی کرد.

1- Eyring

کوهن^۱ و تورنبال^۲ در سال ۱۹۵۹، نقش اساسی در فرآیند نفوذ را به حجم آزاد مایع نسبت دادند و بیان کردند که مولکول در فاز مایع زمانی جابجایی نفوذی را انجام می‌دهد که یک فضای خالی بحرانی (حداقل به اندازه‌ی حجم خود مولکول) در اطرافش پدید آید. نتیجه‌ی نهایی این فرض به صورت معادله‌ی (۱۱-۱) بیان شده است.

$$D_{AA} = g\sigma\sqrt{\frac{3k_B T}{M}} \exp\left(-\frac{\gamma W^*}{V_f}\right) \quad (11-1)$$

که در آن γ یک فاکتور عددی تصحیح کننده، g فاکتور هندسی مولکول، σ قطر مولکولی، V^* حجم بحرانی، k_B ثابت بولتزمن و M جرم مولکولی است. کمیت V_f که حجم آزاد مولکول نامگذاری شده است از رابطه‌ی (۱۲-۱) محاسبه می‌شود.

$$V_f = \frac{V}{N} - V_0 \quad (12-1)$$

که در آن N عدد آووگادرو، V حجم مولی و V_0 حجم واندروالس مولکول است. با مقایسه‌ی نتایج حاصل از این معادله‌ی کوهن و تورنبال با مقادیر آزمایشگاهی مشخص شده است که این معادله، وابستگی ضریب خودنفوذی به دما را در دماهای بالا به درستی نشان نمی‌دهد [۷].

۱-۳-۴ - محاسبه‌ی ضریب خودنفوذی با استفاده از مدل ساختمان ویژه^۳

بر اساس تئوری ساختمان ویژه، در اثر افزایش دما فضاهای خالی در اطراف مولکول‌ها پدید می‌آید و این فضاهای خالی باعث افزایش ضریب خودنفوذی می‌گردد. این تئوری، خاصیت X را به صورت رابطه‌ی (۱۳-۱) محاسبه می‌نماید.

1- Cohen
2- Turnbull
3- Significant structure theory

$$X = X_s \frac{V_s}{V} + X_g \frac{V - V_s}{V} \quad (۱۳-۱)$$

که در آن X_s و X_g مقادیر خاصیت X در حالت جامد و گاز و V_s و V حجم مولی حالت جامد و مایع می‌باشند. بر اساس این مدل ضریب خودنفوذی می‌تواند به صورت رابطه‌ی (۱۴-۱) بیان شود:

$$D_{AA} = D_s \frac{V_s}{V} + D_g \frac{V - V_s}{V} \quad (۱۴-۱)$$

با توجه به اینکه تئوری‌های نفوذ برای حالت جامد و گاز به خوبی تکمیل شده‌اند استفاده از این رابطه به سادگی امکان پذیر است. این تئوری برای خواص تعادلی و سایر خواص انتقالی مانند ویسکوزیته و ضریب هدایت گرمایی نیز با موفقیت اعمال شده است [۷].

۱-۳-۵- محاسبه‌ی ضریب خودنفوذی با استفاده از معادله‌ی استوکس^۱-انیشیتین

از مهم‌ترین روابط ارائه شده برای محاسبه‌ی ضریب خودنفوذی معادله‌ی استوکس-انیشیتین است. در مجموع معادلات حاصل از مدل‌ها دارای پارامترهایی هستند که باید محاسبه شوند و اکثر آن‌ها به معادلاتی مشابه معادله‌ی استوکس-انیشیتین منتهی می‌گردند. در این معادله از رابطه‌ی انیشیتین استفاده شده است.

$$D = \frac{k_B T}{\xi} \quad (۱۵-۱)$$

استوکس با استفاده از هیدرودینامیک کلاسیک پارامتر ξ را برای حرکت یک کره با قطر σ در محیطی که ویسکوزیته‌ی η دارد، طبق رابطه‌ی (۱۶-۱) محاسبه نمود.

1- Stokes

$$\xi = 6\pi\eta \left(\frac{\sigma}{2} \right) \frac{1 + 2\eta/\beta(\sigma/2)}{1 + 3\eta/\beta(\sigma/2)} \quad (16-1)$$

که در آن β ثابت اصطکاک است. با فرض این که $\beta = \infty$ است یعنی لغزشی بین کره‌ی متحرک و سیال در تماس با آن وجود ندارد، رابطه‌ی (۱۶-۱) به شکل رابطه‌ی (۱۷-۱) ساده شده است.

$$\xi = 6\pi\eta \left(\sigma/2 \right) \quad (17-1)$$

با قرار دادن مقدار ξ در رابطه‌ی انیشتین، رابطه‌ی استوکس-انیشتین به صورت زیر ارائه شده است.

$$D_{AA} = k_B T / 6\pi\eta \left(\sigma/2 \right) \quad (18-1)$$

بر طبق اظهار نظر ساترلند^۱، هنگامی $\beta = \infty$ خواهد بود که حرکت یک ذره با قطر بزرگ در محلولی با جرم مولکولی نسبتاً کوچک بررسی شود. اما اگر $\beta = 0$ باشد؛ یعنی در حالتی که مولکول در بین مولکول‌هایی با اندازه‌ی قابل مقایسه نفوذ می‌نماید و لغزش بینابینی وجود دارد، آنگاه ξ از رابطه‌ی (۱۹-۱) به دست می‌آید.

$$\xi = 4\pi\eta \left(\sigma/2 \right) \quad (19-1)$$

به این ترتیب رابطه‌ی استوکس-انیشتین به صورت زیر در می‌آید.

$$D_{AA} = k_B T / 4\pi\eta \left(\sigma/2 \right) \quad (20-1)$$

هرچند این روابط برای حالت خاصی از ماده بیان گردیده‌اند اما در همه‌ی زمینه‌ها برای محاسبه‌ی ضریب نفوذ از این دو معادله استفاده می‌شود [۷].

1- Sutherland

۱-۳-۶- محاسبه‌ی ضریب خودنفوذی با استفاده از معادله‌ی بولتزمن

رابطه‌ی مهم دیگری که برای محاسبه‌ی ضریب خودنفوذی از آن استفاده می‌شود از حل معادله‌ی بولتزمن^۱ در حد دانسیته‌ی صفر به دست می‌آید.

$$D_{AA} = \frac{3}{8} \left(\frac{k_B^3 T^3}{m\pi} \right)^{1/2} \frac{1}{p\sigma^2 \Omega^{(1,1)*} T^*} \quad (۲۱-۱)$$

که در آن k_B ثابت بولتزمن، T دما در مقیاس کلوین، m جرم مولکولی و p فشار است. σ یکی از پارامترهای مهم پتانسیل است و نشان‌دهنده‌ی مکانی است که مقدار پتانسیل مولکول در آن جا برابر با صفر است (غیر از فاصله‌ی بی‌نهایت). $\Omega^{(1,1)*}$ انتگرال برخورد کاهش یافته است که توسط برنامه‌ی رایانه‌ای که بین انتگرال‌های برخورد و پتانسیل کاهش یافته ارتباط برقرار می‌کند، محاسبه می‌شود [۱۰].

۱-۴- استفاده از روش‌های تجربی جهت اندازه‌گیری ضریب خودنفوذی

برای تعمیم دادن روابط مربوط به حالت‌های جامد و گاز به مایعات، مایعات را جامد بدون نظم (شبه کریستالی) یا گاز فشرده فرض می‌کنند اما هیچیک از مدل‌های شبه کریستالی و گاز فشرده به طور کامل خواص حالت مایع را توجیه نمی‌کنند. تنها راه بررسی درستی معادلات حاصل از مدل‌ها مقایسه آن‌ها با نتایج آزمایشگاهی است. این آزمون‌ها با در اختیار داشتن تکنیک‌های طیف سنجی رزونانس مغناطیسی هسته^۲ ممکن می‌شود و امکان مطالعات بیشتری در وابستگی ضریب خودنفوذی به فشار و دما را ممکن می‌سازد.

1- Boltzmann

2- Nuclear magnetic resonance spectroscopy (NMR)

۱-۵- روش طیف سنجی NMR اسپین-اکو^۱

مناسب‌ترین راه برای سنجش ضریب خودنفوذی مایعات، استفاده از تکنیک NMR اسپین-اکو می‌باشد. در ادامه به شرح این روش پرداخته می‌شود.

به منظور بررسی خواص هسته‌های خاص، لازم است فرض شود که این هسته‌ها حول یک محور می‌چرخند و بنابراین خاصیت اسپین دارند. هسته‌های دارای اسپین، اندازه حرکت زاویه‌ای P دارند که حداکثر تعداد اجزای اسپین یا مقادیر P برای یک هسته‌ی خاص به عدد کوانتومی اسپین هسته (I) بستگی دارد. مشخص شده است که یک هسته $2I+1$ حالت مجزای اسپین دارد. هسته‌هایی با $I=0$ در طیف سنجی NMR قابل مطالعه نیستند زیرا هیچ میدان مغناطیسی تولید نمی‌کنند که با میدان خارجی برهم کنش ایجاد کند. دو هسته‌ی مورد توجه شیمی‌دانان ^1H و ^{13}C هستند که عدد کوانتومی اسپین آن‌ها $\frac{1}{2}$ است. هسته‌های ^1H و ^{13}C دو عدد کوانتومی مغناطیسی $+\frac{1}{2}$ و $-\frac{1}{2}$ دارند. در غیاب یک میدان خارجی حالت‌های مختلف اسپین انرژی‌های یکسان دارند اما هنگامیکه هسته‌ای با عدد کوانتومی اسپین $\frac{1}{2}$ در میدان مغناطیسی خارجی B_0 قرار می‌گیرد، نسبت به میدان جهت گیری می‌کند. این جهت گیری یا در جهت میدان مغناطیسی خارجی است و یا در خلاف جهت آن. انرژی پتانسیل هسته در هریک از این دو جهت یا دو حالت کوانتومی با رابطه‌ی (۲۲-۱) تعیین می‌شود.

$$E = -\frac{\gamma m h}{2\pi} B_0 \quad (22-1)$$

که در آن h ثابت پلانک^۲ است.

بنابراین، اختلاف انرژی بین دو حالت پایه ($I = +\frac{1}{2}$) و برانگیخته ($I = -\frac{1}{2}$) طبق رابطه‌ی

1- Espin-echo
2- Planck constant

(۲۳-۱) به دست می‌آید.

$$\Delta E = \frac{\gamma \hbar}{4\pi} B_0 - \left(-\frac{\gamma \hbar}{4\pi} B_0 \right) = \frac{\gamma \hbar}{2\pi} B_0 \quad (۲۳-۱)$$

این اختلاف انرژی با انرژی پرتوهای فرکانس رادیویی^۱ متناظر است. در نتیجه با تاباندن پرتوهای فرکانس رادیویی، RF، هسته‌ها از حالت پایه به حالت برانگیخته منتقل می‌شوند.

نسبت پروتون‌های موجود در دو حالت پایه و برانگیخته با استفاده از رابطه‌ی توزیع بولتزمن محاسبه می‌شود.

$$\frac{N_j}{N_0} = \exp\left(\frac{\Delta E}{k_B T}\right) \quad (۲۴-۱)$$

که در این رابطه N_j تعداد پروتون‌ها در حالت انرژی بالاتر، N_0 تعداد پروتون‌ها در حالت انرژی پایین‌تر، k_B ثابت بولتزمن و T دمای مطلق است.

اختلاف انرژی بین دو حالت پایه و برانگیخته کم است و همین امر سبب می‌شود که تفاوت جمعیت اسپین‌ها در دو حالت پایه و برانگیخته نیز کم باشد. فرآیند جذب پرتوهای RF تنها زمانی اتفاق می‌افتد که $N_j < N_0$ باشد پس همیشه این خطر وجود دارد که فرآیند جذب سبب برابر شدن تعداد هسته‌ها در دو حالت پایه و برانگیخته شود و در نتیجه طیف جذبی کاهش یافته و به سمت صفر برود. این حالت را حالت اسپین سیر شده می‌نامند. به منظور اهتزاز از سیرشدگی لازم است که زمان آسایشی برای هسته‌ها در نظر گرفته شود تا هسته‌های برانگیخته با از دست دادن انرژی به حالت پایه برگردند. دو نوع فرآیند آسایش اسپین-شبکه و اسپین-اسپین برای آسایش هسته‌های برانگیخته به کار برده می‌شود.

1- Radio frequency (RF)

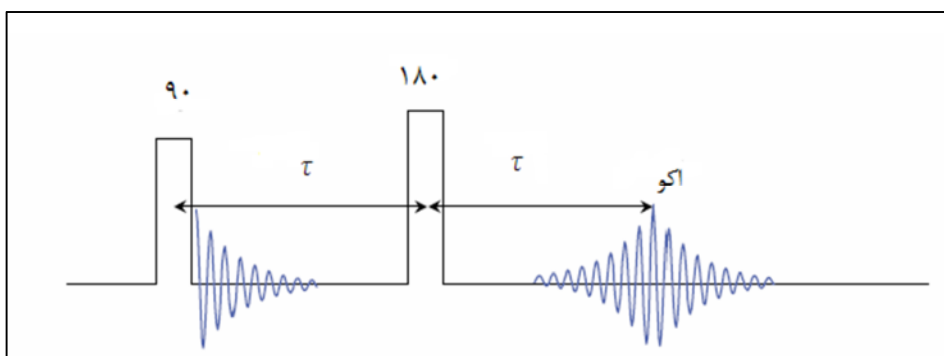
الف) آسایش اسپین-شبهه: هسته‌های جاذب در یک آزمایش NMR قسمتی از مجموعه‌ی بزرگ اتم‌ها هستند که نمونه را تشکیل می‌دهد، این مجموعه شبکه نامیده می‌شود. هسته‌های مختلف تشکیل دهنده‌ی شبکه، حرکت‌های ارتعاشی و چرخشی دارند. در این بین بعضی از شبکه‌ها با فرکانسی برابر با فرکانس مربوط به هسته‌ی برانگیخته ارتعاش می‌کنند. این امر اجازه می‌دهد که انرژی از اسپین برانگیخته به شبکه منتقل شود و اسپین برانگیخته به حالت پایه برگردد. آسایش اسپین-شبهه سیستم را با ثابت زمانی t_1 به حالت تعادل بر می‌گرداند.

ب) آسایش اسپین-اسپین: این آسایش به واسطه‌ی برهمکنش بین هسته‌های همسایه و هسته‌ی برانگیخته حاصل شده و به یک تبادل درونی حالت‌های کوانتومی بین دو هسته منجر می‌شود؛ یعنی هسته‌ی در حالت اسپین پایین‌تر برانگیخته شده و هسته‌ی برانگیخته به حالت انرژی پایه برمی‌گردد. هیچ تغییر خالصی در جمعیت نسبی هسته‌ها در دو حالت پایه و برانگیخته به وجود نمی‌آید؛ یعنی هیچ کاهشی در سیرشدگی اسپین روی نمی‌دهد، اما طول عمر متوسط یک هسته‌ی برانگیخته‌ی خاص، کوتاه می‌شود. آسایش اسپین-اسپین سیستم را با ثابت زمانی t_2 به حالت تعادل برمی‌گرداند. در اکثر موارد $t_2 < t_1$ به دست می‌آید [۱۱].

مهم‌ترین روش اندازه‌گیری t_2 ، روش NMR اسپین-اکو است. در این روش ابتدا یک پالس ۹۰ درجه نسبت به میدان مغناطیسی نمونه، توسط دستگاه اعمال می‌شود. این پالس سبب به وجود آمدن میدان مغناطیسی می‌شود که با میدان حاصل از اسپین هسته‌ها برهمکنش دارد و در نتیجه یک سیگنال خروجی حاصل می‌شود. اما در اثر توزیع اتفاقی میدان‌های مغناطیسی محلی که ناشی از غیر یکنواختی میدان مغناطیسی اعمال شده است، بردارهای اسپین به تدریج از هم باز می‌شوند تا جایی که مقدار سیگنال به صفر می‌رسد. بعد از گذشت زمان کوتاه τ ، با ایجاد یک پالس رادیو فرکانس شدید ۱۸۰ درجه، همه‌ی بردارهای اسپین باز شده، به هم می‌پیوندند و در نتیجه سیگنالی ایجاد می‌شود که پس از زمان 2τ به ماکزیمم می‌رسد و سپس مجدداً به سمت صفر کاهش می‌یابد. طیف حاصل

چون از شدت بالایی برخوردار است به آن اکو گفته می‌شود.

در شکل (۱-۱) مراحل NMR اسپین اکو به طور خلاصه نشان داده شده است. برای به دست آوردن t_2 چندین بار در بازه‌های زمانی τ پالس‌های 180° درجه تکرار می‌شوند تا یک سری اسپین-اکو حاصل شود. پیک‌های مربوط به سیگنال‌های ناشی از پالس‌های 180° درجه، به طور نمایی با ثابت زمانی مساوی با زمان آسایش اسپین-اسپین افت می‌کنند [۱۱].



شکل (۱-۱): مراحل NMR اسپین-اکو

۱-۵-۱- اندازه‌گیری ضریب خودنفوذی مایعات با استفاده از طیف سنجی NMR

اسپین-اکو با میدان ثابت

این روش در ابتدا توسط هان^۱ پیشنهاد گردید و سپس توسط کار^۲ و پورسل^۳ توسعه یافت. آن‌ها دریافتند که با به کارگیری گرادیانی برای یک میدان مغناطیسی ثابت و خطی و با بررسی مدت زمان t_2 ، می‌توان نفوذ مولکولی مولکول‌های نشان‌دار شده را با توجه به گشتاور مغناطیسی هسته‌ای آن‌ها محاسبه نمود. رابطه‌ای که کار و پورسل به آن دست یافتند عبارت است از:

1- Hahn
2- Carr
3- Purcell

$$A_{(2\tau)} = A_0 \exp\left[-2\tau/T_2\right] \exp\left[-\left(\frac{2}{3}\right)(\gamma g \tau)^2 D_{AA} \tau\right] \quad (25-1)$$

که در آن 2τ زمان به ماکزیمم رسیدن سیگنال ثانویه، A میزان تحلیل دامنه‌ی اکو، A_0 دامنه‌ی اکوی اول، γ نسبت ژيرومغناطیسی هسته‌ی مورد بررسی، τ زمان بین پالس‌های 90° و 180° درجه، D_{AA} ضریب خودنفوذی و $g = KI$ است که در آن I جریان سیم پیچ دستگاه و K یک عامل درجه بندی است که توسط آن مقدار g سنجیده می‌شود.

تعیین ضریب خودنفوذی با این رابطه نیازمند دانستن مقدار دقیق g است که از آن به عنوان شیب میدان مغناطیسی یاد می‌شود. این مهم معمولاً با استفاده از ماده‌ای که ضریب خودنفوذی آن به دقت محاسبه شده است انجام می‌شود و با استفاده از ماده‌ی دومی که t_2 و D_{AA} مشخص و کاملاً متفاوتی با ماده‌ی مورد سنجش دارد، کنترل می‌شود. این رابطه به علت استفاده از میدان مغناطیسی ثابت و خطی، معمولاً نتایجی با خطای بالا حاصل می‌کند [۱۲ و ۱۳].

۱-۵-۲- اندازه‌گیری ضریب خودنفوذی مایعات با استفاده از طیف سنجی NMR

اسپین-اکو با میدان متناوب

معادله‌ای که بیشترین کاربرد را در تعیین ضریب خودنفوذی مایعات دارد، با استفاده از تکنیک میدان مغناطیسی متناوب به دست می‌آید که نخستین بار توسط تانر^۱ مطرح گردید. تعیین ضریب خودنفوذی در این تکنیک، با تغییر جریان سیم پیچ (I) صورت می‌گیرد:

$$A_{(2\tau)} = A_0 \exp\left[-2\tau/T_2\right] \exp\left[-(\gamma \delta g)^2 \left(\Delta - \delta/3\right) D_{AA} \tau\right] \quad (26-1)$$

1- Tanner

در این رابطه τ زمان بین پالس‌های ۹۰ و ۱۸۰ درجه، δ مدت زمان یک ضربان شیب میدان مغناطیسی و Δ زمان بین دو ضربان شیب میدان مغناطیسی و $g = KI$ نیز شیب میدان مغناطیسی است [۱۲].

به منظور اندازه‌گیری ضریب خودنفوذی آلکان‌های خطی مایع، از سل‌هایی با قابلیت تحمل فشار بالا و از جنس آلیاژی قوی و غیرمغناطیس استفاده شده است. نمونه‌ی مایع در اتوکلاوهایی از جنس آلیاژهای تیتانیوم و یا برلیوم قرار داده شده و در یک دما و فشار ثابت مورد بررسی قرار گرفته است. یک دستگاه RF برای ایجاد پالس‌های ۹۰ و ۱۸۰ درجه مورد استفاده قرار گرفته و با استفاده از رابطه‌ی (۱-۳۰) ضریب خودنفوذی هر آلکان محاسبه شده است [۱۲ و ۱۳].

برای مدل‌سازی ضریب خودنفوذی در این پایان‌نامه، از داده‌های ضریب خودنفوذی آلکان‌های خطی مایعی که با استفاده از روش NMR اسپین-۱ کو با میدان متناوب اندازه‌گیری شده، استفاده گردیده است.

فصل دوم

بخش تئوری

۲-۱- کمومتریکس^۱

برای درک مکانیسم فرآیندهای مختلف شیمیایی، کشف و توسعه‌ی مواد جدید و زمینه‌های دیگر علم شیمی، هنوز توانایی حل مسائل به طور کامل وجود ندارد و معمولاً برای حل دقیق مسائل، نیاز به سیستم‌های بسیار پیچیده‌ای است که اجرای آن‌ها مستلزم صرف وقت و هزینه‌های بسیار است. در جهت حل این مشکل، روش‌های محاسباتی کمومتریکس می‌توانند مفید باشند. هدف کمومتریکس، بهبود بخشیدن فرآیندهای اندازه‌گیری و استخراج اطلاعات مفیدتر از داده‌های اندازه‌گیری شده‌ی فیزیکی و شیمیایی است. عبارت کمومتریکس در ابتدا توسط اسوانت ولد^۲ در سال ۱۹۷۴ ابداع گردید و پس از مدت کوتاهی، مجمع بین‌المللی کمومتریکس (ICS) توسط کووالسکی^۳ و ولد که به عنوان پیشگامان در این حوزه بودند تأسیس گردید [۱۴]. طبق تعریف مجمع بین‌المللی کمومتریکس، کمومتریکس علمی است که با به کارگیری آمار و ریاضیات، بین اندازه‌گیری‌های انجام شده بر روی یک سیستم و یا فرآیند شیمیایی با خواص ماده ارتباط ایجاد می‌کند. مطالعات کمومتریکس محدوده‌ی وسیعی از روش‌های مختلف را که می‌توانند در شیمی به کار گرفته شوند در بر می‌گیرد.

۲-۲- رابطه‌ی ساختار- خاصیت یا ساختار- فعالیت

یکی از زمینه‌های مهم کاربرد کمومتریکس در مطالعاتی است که خواص مولکول‌ها را به ویژگی‌های ساختاری آن‌ها نسبت می‌دهد. از نظر شیمی‌دانان فعالیت و خواص یک ترکیب ناشی از ویژگی‌های ساختاری آن است. این مطالعات به بررسی کیفی ارتباط ساختار- فعالیت^۴ (QSAR)، و همچنین بررسی کمی ارتباط ساختار- خاصیت^۵ (QSPR) معروف می‌باشد.

هدف از مطالعات QSPR پیدا کردن رابطه‌ای است که بین رفتار فیزیکوشیمیایی یک مولکول

1- Chemometrics

2- Svante Wold

3- B.Kowalski

4- Quantitative structure-activity relationship

5- Quantitative structure-property relationship

با پارامترهای ساختاری آن وجود دارد. نتایج این مطالعات علاوه بر شفاف سازی نحوه‌ی ارتباط خواص مولکول‌ها و ویژگی‌های ساختمانی آن‌ها، به پژوهشگران در پیش‌بینی رفتار مولکول‌های جدید بر اساس رفتار مولکول‌های مشابه کمک می‌کند. ولی در QSAR هدف پیدا کردن رابطه‌ای هماهنگ میان فعالیت بیولوژیکی یا خاصیت مواد با ویژگی‌های مولکول است، تا بتوان این رابطه را برای ارزیابی فعالیت یا خاصیت ترکیبات جدید به کار برد.

۲-۳- مراحل انجام QSPR

به‌طور کلی مطالعات QSPR/QSAR شامل چهار بخش اصلی است: انتخاب سری مولکولی، محاسبه و انتخاب توصیف‌کننده‌ها، مدل‌سازی و انتخاب بهترین مدل، ارزیابی اعتبار مدل‌های ساخته شده. در ادامه هر یک از این مراحل به اختصار شرح داده خواهد شد.

۲-۳-۱ انتخاب سری مولکولی

اولین مرحله در مطالعات QSPR/QSAR، جمع‌آوری و انتخاب تعدادی مولکول است که خاصیت یا فعالیت مورد نظر آن‌ها به طور تجربی در دسترس باشد. لازم به ذکر است که خاصیت یا فعالیت مورد نظر ترکیبات انتخاب شده باید در شرایط یکسان اندازه‌گیری شده باشد. به‌علاوه مجموعه ترکیبات مورد مطالعه باید تا حدودی تشابه ساختاری داشته باشند تا تعداد توصیف‌کننده‌های مورد نیاز برای ایجاد یک مدل مناسب، کم شود. بدیهی است که هرچه اطلاعات تجربی قابل دسترس برای طراحی مدل بیشتر باشد، مدل کارایی بهتری خواهد داشت.

۲-۳-۲ محاسبه و انتخاب توصیف‌کننده‌ها

توصیف‌کننده‌ها پارامترهای عددی هستند که ویژگی‌های متفاوتی از مولکول را بیان می‌کنند.

به عبارت دیگر، هریک از این توصیف‌کننده‌ها، اطلاعات خاصی از مولکول را در اختیار می‌گذارند. محاسبه و انتخاب توصیف‌کننده‌ها در مدل‌سازی از اهمیت به سزایی برخوردار است. چرا که اگر توصیف‌کننده‌های انتخاب شده مناسب نباشند، عملیات مدل‌سازی با مشکل مواجه شده و مدل مناسب به دست نخواهد آمد.

۲-۳-۲-۱ محاسبه‌ی توصیف‌کننده‌ها

توصیف‌کننده‌های مولکولی یا تئوری نقش اساسی در شیمی، علوم داروسازی، علوم زیست محیطی و مطالعات مسائل سلامتی و کنترل کیفیت دارند. این توصیف‌کننده‌ها امکان یافتن ارتباط بین پاسخ مورد نظر و ساختار ترکیبات را فراهم می‌آورند. توصیف‌کننده‌های مولکولی بیان ریاضی مولکول بوده و اطلاعات شیمیایی مولکولی را که با یک بیان نمادین نمایش داده شده‌اند، به اعداد قابل استفاده تبدیل می‌کنند. برای محاسبه‌ی بعضی از توصیف‌کننده‌ها به مختصات سه بعدی اتم‌ها نیاز است که برای تحقق این مهم باید ساختمان مولکولی ترکیبات بهینه شود؛ یعنی پایدارترین وضعیت ساختاری آن‌ها با حداقل انرژی تعیین گردد.

۲-۳-۲-۲ کاهش تعداد توصیف‌کننده‌ها

تا این مرحله، محاسبه‌ی توصیف‌کننده‌ها توسط نرم افزار و به صورت پیش‌فرض صورت گرفته است. بنابراین، احتمال این‌که برخی از توصیف‌کننده‌ها برای مدل‌سازی مناسب نباشند وجود دارد و باید آن‌ها را حذف کرد. توصیف‌کننده‌هایی که حذف می‌شوند معمولاً در یک یا چند دسته زیر قرار دارند:

- توصیف‌کننده‌هایی که کمتر از ۱۰٪ مقادیر غیر صفر دارند یا دارای بیش از ۹۰٪ مقادیر یکسان هستند.

▪ توصیف‌کننده‌هایی که با توصیف‌کننده‌های دیگر همبستگی بالایی دارند ($r > 0.9$).

▪ توصیف‌کننده‌هایی که با متغیر وابسته، همبستگی کمی دارند.

حذف توصیف‌کننده‌های نامناسب معادلات نهایی را ساده‌تر کرده و از پیچیدگی محاسبات جلوگیری می‌کند.

۲-۳-۳- تقسیم بندی داده‌ها

اصولاً سری ترکیبات یا داده‌ها به دو دسته‌ی سری مرجع^۱ (آموزشی) و سری تست^۲ (آزمایشی)، یا به سه دسته‌ی سری مرجع، سری ارزیابی^۳ و سری تست تقسیم می‌شوند. تقسیم بندی معمولاً به‌صورتی انجام می‌شود که ۷۰٪ داده‌ها در سری آموزش، ۱۵٪ در سری ارزیابی و ۱۵٪ در سری تست قرار گیرند. سری آموزش اکثریت داده‌ها را دربر می‌گیرد و گروهی است که برای ساخت مدل از آن استفاده می‌شود. سری تست شامل بقیه داده‌ها می‌باشد و از این سری به منظور ارزیابی قدرت پیش‌بینی مدل‌های ساخته شده استفاده می‌شود. در بعضی مواقع از روش‌هایی برای مدل‌سازی استفاده می‌شود که در آن‌ها سری تست به نحوی در مدل‌سازی درگیر می‌شود بنابراین، برای ارزیابی صحیح بهتر است از گروه سومی استفاده شود که هیچ دخالتی در مدل‌سازی نداشته باشد.

۲-۳-۳-۱- تقسیم بندی داده‌ها به روش تحلیل مؤلفه‌ی اصلی^۴ (PCA)

به منظور تقسیم‌بندی داده‌ها به سه سری آموزش، ارزیابی و تست می‌توان از روش تقسیم تصادفی استفاده نمود به این ترتیب که مولکول‌ها بر حسب مقدار عددی خاصیت مورد نظر، به‌صورت

1- Training set

2- Test set

3- Validation set

4- Principal component analysis

صعودی یا نزولی مرتب شوند، سپس از بین آن‌ها سه سری آموزش، ارزیابی و تست به نسبت ۷۰:۱۵:۱۵ طوری انتخاب شوند که دسته‌بندی داده‌ها از لحاظ آماری دارای توزیع نرمال باشد.

اما برای تقسیم بندی دقیق‌تر می‌توان از روش تحلیل مؤلفه‌ی اصلی، PCA، بهره گرفت. روش تحلیل مؤلفه‌ی اصلی، ابعاد کلیه مشاهدات را براساس شاخص ترکیب و دسته‌بندی مشاهدات مشابه کاهش می‌دهد. معمولاً دو مؤلفه‌ی اول و دوم مقدار قابل توجهی از پراکندگی داده‌ها را در نظر می‌گیرند و بنابراین، نمودار پراکندگی براساس این دو مؤلفه رسم می‌شود. روش تحلیل مؤلفه‌های اصلی چون هیچ تابع توزیعی برای مشاهدات در نظر نمی‌گیرد، روش آماری محسوب نمی‌شود بنابراین، دانستن این نکته که روش تحلیل مؤلفه‌ی اصلی تنها برای ارائه‌ی داده‌ها به شکلی ساده‌تر و با ابعاد کمتر به کار می‌رود حائز اهمیت است [۱۶ و ۱۷].

۲-۳-۴- انتخاب بهترین توصیف‌کننده‌ها

هرچه در مدل نهایی تعداد توصیف‌کننده‌ها کمتر باشد سبب سادگی و کارآمدی بیشتر مدل می‌شود. بنابراین، کاهش دادن تعداد توصیف‌کننده‌ها پیش از آغاز مرحله‌ی مدل‌سازی از اهمیت ویژه‌ای برخوردار است. برای این منظور می‌توان از روش‌های خطی و غیرخطی گوناگونی استفاده نمود. در ادامه روش خطی رگرسیون مرحله‌ای^۱ (SR) و روش غیر خطی الگوریتم ژنتیک^۲ (GA) که در این پایان‌نامه مورد استفاده قرار گرفته اند شرح داده خواهند شد.

۲-۳-۴-۱- انتخاب بهترین توصیف‌کننده‌ها توسط رگرسیون مرحله‌ای (SR)

در رگرسیون مرحله‌ای مدل‌سازی اولیه با استفاده از روش رگرسیون خطی چندگانه^۳ (MLR)

1- Stepwise regression

2- Genetic algorithm

3- Multiple linear regression

انجام می‌گیرد. در روش رگرسیون مرحله‌ای توصیف‌کننده‌ها به صورت مرحله‌ای و به ترتیب ضریب همبستگی که با خاصیت مورد نظر دارند به مدل وارد شده و مدل‌سازی به روش MLR انجام می‌شود. در هر مرحله برای سنجش معنادار بودن ضریب همبستگی متغیرهای وارد شده به مدل، پارامترهای آماری مدل مورد سنجش قرار می‌گیرند و چنانچه این پارامترها معنادار باشند، آن توصیف‌کننده در مدل باقی مانده و در غیر این صورت حذف خواهد شد. به این ترتیب کلیه‌ی مدل‌های ممکن ساخته شده و برای هر یک از آن‌ها پارامترهای آماری تعیین می‌شود. از مقایسه‌ی پارامترهای آماری بهترین مدل مشخص شده و توصیف‌کننده‌های بهترین مدل که توصیف‌کننده‌های منتخب به روش رگرسیون مرحله‌ای هستند، شناسایی می‌شوند.

از جمله مهم‌ترین پارامترهای آماری که برای انتخاب بهترین مدل مدنظر قرار می‌گیرند ضریب همبستگی رگرسیون^۱ (R)، ضریب تعیین^۲ (R^2)، آماره‌ی F ، پارامتر FIT و میانگین مربعات خطا^۳ (MSE) می‌باشند.

زمانیکه تنها یک متغیر مستقل در مدل وارد شده باشد، مقدار ضریب همبستگی رگرسیون برابر با ضریب همبستگی متغیر مستقل و متغیر وابسته است. اما در صورت وجود چندین متغیر مستقل در مدل، مقدار R ، ضریب همبستگی چند متغیره بین متغیر وابسته و مجموع متغیرهای مستقل معادله‌ی رگرسیون می‌باشد. هر چه R یک مدل به یک نزدیکتر باشد، آن مدل مناسب‌تر خواهد بود. ضریب همبستگی دو متغیر x و y براساس رابطه‌ی (۱-۲) تعریف می‌شود.

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (1-2)$$

پارامتر دیگری که برای بیان دقت خط رگرسیون برآورد شده و نشان دهنده‌ی نسبت تغییرات متغیر وابسته به متغیر مستقل است، ضریب تعیین می‌باشد که طبق رابطه‌ی (۲-۲) محاسبه می‌شود.

1- Correlation coefficient
2- Determination coefficient
3- Mean square error

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (2-2)$$

که SSR ^۱، بیانگر مجموع مربعات انحراف مقادیر پیش‌بینی شده‌ی متغیر وابسته از میانگین مقادیر آن است.

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (3-2)$$

و SST ^۲، نشانگر مجموع مربعات انحراف مقادیر واقعی متغیر وابسته از میانگین مقادیر آن است.

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (4-2)$$

SSE ^۳ نیز مبین مجموع انحرافات مقادیر واقعی متغیر وابسته از مقادیر پیش‌بینی شده است.

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5-2)$$

بنابراین می‌توان نوشت:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6-2)$$

طبق رابطه‌ی (۶-۲) اگر برای تمام نقاط $y_i = \hat{y}_i$ باشد، مقدار R^2 برابر یک می‌شود و کوچکترین انحراف از این حالت باعث می‌شود، ضریب تعیین از یک کوچکتر شود.

آماره‌ی F معیار دیگری برای ارزیابی مدل بوده و مقدار آن بیانگر میزان معنی‌دار بودن کلی مدل است. مقدار این پارامتر از نسبت میانگین مربعات رگرسیون (MSR) به میانگین مربعات باقیمانده‌ها (MSE) به دست می‌آید.

$$F = \frac{MSR}{MSE} = \frac{SSR/df_m}{SSE/df_{res}} \quad (7-2)$$

که در آن df_m درجه‌ی آزادی مدل و برابر با تعداد توصیف‌کننده‌ها است ($df_m = p$). df_{res} درجه‌ی

1- Sum square regression
2- Sum square total
3- Sum square error

آزادی باقیمانده‌ها است و به صورت $df_{res} = n - p - 1$ تعریف می‌شود.

پارامتر FIT نیز یک پارامتر آماری مهم برای مقایسه‌ی مدل‌های مختلف با تعداد متفاوت

توصیف‌کننده است و هر چه مقدار این پارامتر برای یک مدل بیشتر باشد، آن مدل مناسب‌تر است.

$$FIT = \frac{R^2(n-p-1)}{(n+p^2)(1-R^2)} \quad (۸-۲)$$

چند پارامتر دیگر برای ارزیابی مدل‌ها عبارتند از: خطای استاندارد پیش‌بینی^۱ (SEP)، میانگین

خطای مطلق^۲ (MAE)، خطای نسبی پیش‌بینی^۳ (REP)، میانگین مربعات خطا (MSE) و میانگین

خطای نسبی^۴ (MRE) که طبق روابط (۹-۲) تا (۱۳-۲) محاسبه می‌شوند [۱۸].

$$SEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (۹-۲)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (۱۰-۲)$$

$$REP(\%) = \frac{100}{\bar{y}} \times \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (۱۱-۲)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (۱۲-۲)$$

$$MRE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n} \times 100 \quad (۱۳-۲)$$

از پارامتر مجموع مربع باقیمانده‌ها^۵ ($PRESS$) نیز جهت ارزیابی مدل‌ها استفاده می‌شود. این

مقدار برابر مجموع مربعات تفاوت مقادیر واقعی و مقادیر پیش‌بینی شده است.

$$PRESS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (۱۴-۲)$$

1- Standard error of prediction

2- Mean absolute error

3- Relative error of prediction

4- Mean relative error

5- Predictive residual sum of squares

سه پارامتر مهم دیگر که برای ارزیابی مدل‌ها کاربرد دارند عبارتند از: انحراف میانگین^۱ ($bias$)، درصد انحراف میانگین مطلق^۲ (AAD) و انحراف ماکزیمم^۳ (D_{Max}). این مقادیر به ترتیب از روابط (۱۵-۲) تا (۱۷-۲) محاسبه می‌شوند.

$$bias = \frac{100}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} \quad (15-2)$$

$$AAD = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (16-2)$$

$$D_{Max} = \text{Max}(100 \left| \frac{y_i - \hat{y}_i}{y_i} \right|) \quad (17-2)$$

در تمامی روابط (۱-۲) تا (۱۷-۲) y_i و $\hat{y}_i$ به ترتیب مقدار واقعی و مقدار پیش‌بینی شده‌ی متغیر وابسته‌ی i ام، p تعداد توصیف‌کننده‌های مدل و n تعداد داده‌های وارد شده به مدل است [۱۹].

۲-۳-۴-۲- انتخاب بهترین توصیف‌کننده‌ها توسط الگوریتم ژنتیک

الگوریتم ژنتیک را شاید بتوان یکی از شناخته‌ترین الگوریتم‌های مدرن برای مسائل بهینه‌سازی دانست. در حالت کلی، الگوریتم ژنتیک بر سیستم‌های تکامل تدریجی و قواعد موروثی استوار است که در آن‌ها با استفاده از مجموعه‌ی راه‌حل‌های بالقوه‌ی موجود در جمعیت مورد مطالعه، نسل‌های آینده به وجود می‌آیند. این الگوریتم همان‌طور که از نام آن پیداست، از توارث صفات والدین به فرزندان به وسیله‌ی ترکیب کروموزوم‌های والدین، ایده می‌گیرد. الگوریتم ژنتیک بر روی یک سری جواب‌های مسأله، به امید به دست آوردن جواب‌های بهتر، قانون بقای بهترین را اعمال می‌کند. این فرآیند باعث می‌شود که نسل‌های جدید با شرایط مسأله سازگارتر باشند [۲۰].

1- Average deviation

2- Absolute average percent deviation

3- Maximum deviation

۲-۳-۴-۲-الف: الگوریتم ژنتیک

الگوریتم ژنتیک توسط گولبرگ^۱ در سال ۱۹۸۹ پیشنهاد گردید. این الگوریتم دارای سه عملگر اساسی انتخاب^۲، آمیزش یا ترکیب^۳ و جهش^۴ می‌باشد. الگوریتم‌های ژنتیک از اصول تکامل طبیعی داروین گرفته شده‌اند و یکی از شیوه‌های جستجو و بهینه‌سازی تصادفی هستند که بر مبنای اصول تکامل طبیعی پایه‌ریزی شده‌اند. الگوریتم ژنتیک با انبوهی از جواب‌ها سروکار دارد و به دلیل داشتن همین خصوصیت، امکان اجرای موازی الگوریتم و همچنین مهاجرت بین زیر جمعیت‌ها را که باعث تنوع ژنتیکی می‌شوند، فراهم می‌آورد [۲۰].

۲-۳-۴-۲-ب: عملگرهای الگوریتم ژنتیک

به طور خلاصه الگوریتم ژنتیک از عملگرهای زیر تشکیل شده‌است:

الف) کد گذاری^۵: این مرحله شاید مشکل‌ترین مرحله‌ی حل مسأله به روش الگوریتم ژنتیک باشد. الگوریتم ژنتیک به جای آن که بر روی پارامترها یا متغیرهای مسأله کار کند، با شکل کدگذاری شده‌ی آن‌ها سروکار دارد. یکی از روش‌های کد کردن، کد کردن دو دویی می‌باشد که در آن هدف تبدیل جواب مسأله به رشته‌ای از اعداد باینری (در مبنای ۲) است.

ب) ارزیابی^۶: تابع برازندگی را از اعمال تبدیل مناسب بر روی تابع هدف، یعنی تابعی که قرار است بهینه شود، به دست می‌آورند. این تابع هر رشته را با یک مقدار عددی ارزیابی می‌کند که کیفیت آن را مشخص می‌نماید. هرچه کیفیت رشته‌ی جواب بالاتر باشد مقدار برازندگی جواب بیشتر شده و احتمال مشارکت آن برای تولید نسل بعدی افزایش خواهد یافت.

1- Goldberg
2- Selection
3- Crossover
4- Mutation
5- Coding
6- Fitness

ج) ترکیب: مهمترین عملگر در الگوریتم ژنتیک، عملگر ترکیب است. ترکیب فرآیندی است که در آن نسل قدیمی کروموزوم‌ها با یکدیگر مخلوط و ترکیب می‌شوند تا نسل تازه‌ای از کروموزوم‌ها به وجود بیاید. جفت‌هایی که در قسمت انتخاب به عنوان والد در نظر گرفته شده‌اند، در این قسمت ژن‌هایشان را با هم مبادله می‌کنند و اعضای جدید به وجود می‌آورند. ترکیب در الگوریتم ژنتیک باعث از بین رفتن پراکندگی یا تنوع ژنتیکی جمعیت می‌شود زیرا اجازه می‌دهد ژن‌های خوب یکدیگر را بیابند.

د) جهش: جهش عملگری است که جواب‌های ممکن دیگری را تولید می‌کند. در الگوریتم ژنتیک بعد از این‌که یک عضو در جمعیت جدید به وجود آمد، هر ژن آن با احتمال جهش، جهش می‌یابد. در جهش ممکن است ژنی از مجموعه‌ی ژن‌های جمعیت حذف شود یا ژنی که تا به حال در جمعیت وجود نداشته به آن اضافه شود. اما جهش نباید زیاد اتفاق بیفتد زیرا در این صورت الگوریتم ژنتیک به جستجوی کاملاً تصادفی تبدیل خواهد شد.

ه) رمزگشایی^۱: رمزگشایی عکس عمل کدگذاری است. در این مرحله بعد از اینکه الگوریتم ژنتیک بهترین جواب را برای مسأله ارائه کرد لازم است عکس عمل کدگذاری یا همان رمزگشایی روی جواب‌ها اعمال شود تا بتوان نسخه‌ی واقعی جواب را به وضوح در دست داشت.

۲-۳-۴-۲-ج: نقاط قوت الگوریتم ژنتیک

مهمترین نقطه‌ی قوت الگوریتم ژنتیک این است که الگوریتم ژنتیک ذاتاً موازی است. این در حالی است که اکثر الگوریتم‌های دیگر موازی نیستند و فقط می‌توانند فضای مسأله مورد نظر را در هر لحظه در یک جهت معین جستجو کنند. اما چون الگوریتم ژنتیک چندین نقطه‌ی شروع دارد پس در یک لحظه می‌تواند فضای مسأله را از چند جهت مختلف جستجو کند.

1- Decoding

نقطه‌ی قوت دیگر الگوریتم ژنتیک آن است که این الگوریتم هیچ اطلاعاتی در مورد مسائلی که حل می‌کنند ندارد و اصطلاحاً «ساعت ساز نابینا»^۱ است. الگوریتم ژنتیک تغییراتی تصادفی در راه‌حل‌های اولیه به وجود می‌آورد و سپس از تابع برازش برای سنجش این که آیا تغییرات پیشرفتی ایجاد کرده‌اند یا نه، استفاده می‌کند. مزیت این تکنیک آن است که به الگوریتم ژنتیک اجازه می‌دهد با ذهنی باز شروع به حل مسائل کند.

۲-۳-۵- مدل‌سازی

مدل در واقع یک رابطه‌ی ریاضی است که بیان‌کننده‌ی ارتباط بین متغیر وابسته (خاصیت) و متغیرهای مستقل (ویژگی‌های مولکول یا توصیف‌کننده‌ها) می‌باشد. به کمک مدل می‌توان با داشتن مقادیر متغیرهای مستقل، مقدار متغیر وابسته را پیش‌بینی کرد. روش‌های مختلف خطی و غیرخطی برای این منظور به کار گرفته می‌شود. از جمله روش‌های مهم خطی، رگرسیون خطی چندگانه (MLR)، رگرسیون مؤلفه‌ی اصلی^۲ (PCR) و رگرسیون کمترین مربعات جزئی^۳ (PLS) هستند. از روش‌های مطرح غیرخطی نیز می‌توان به ماشین‌های بردار پشتیبان^۴ (SVM) و شبکه‌ی عصبی مصنوعی^۵ (ANN) اشاره کرد.

۲-۳-۵-۱- رگرسیون خطی چندگانه (MLR)

در میان روش‌های رگرسیون خطی چند متغیره برای مدل‌سازی، روش MLR به طور وسیعی مورد استفاده قرار می‌گیرد. این روش زمانی کاربرد دارد که در ورودی مدل بیش از یک متغیر مستقل

1- Blind watchmaker

2- Principle Component Regression

3- Partial Least Squares

4- Support vector machines

5- Artificial neural network

وجود داشته باشد. در این روش یک معادله‌ی خطی چند جمله‌ای بین متغیرهای مستقل $x_1, x_2, \dots, x_n$ و متغیر وابسته‌ی y به صورت زیر ایجاد می‌شود:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + e \quad (18-2)$$

که در این رابطه، از مقادیر $a_1, a_2, \dots, a_n$ با عنوان ضرایب رگرسیون یاد می‌شود. رگرسیون خطی چندگانه بسیار ساده و قابل تفسیر است، اما اگر تعداد متغیرهای مستقل زیاد و قابل مقایسه با تعداد مولکول‌ها باشد، کارایی لازم را نخواهد داشت. [۱۷].

۲-۳-۵-۲- رگرسیون مؤلفه‌ی اصلی (PCR)

اگر در هر مجموعه داده، n توصیف‌کننده وجود داشته باشد، هر توصیف‌کننده می‌تواند دارای چند بعد باشد. با توجه به این که اغلب درک و شهود فضای چند بعدی دشوار است، روش PCA ابعاد کلیه مشاهدات را براساس شاخص دسته‌بندی و ترکیب مشاهدات مشابه کاهش می‌دهد. ایده‌ی روش PCR ایجاد یک مدل خطی بین متغیر وابسته و مؤلفه‌های اصلی است. این روش شامل دو مرحله است: تجزیه‌ی ماتریس متغیرهای مستقل به مؤلفه‌های اصلی به کمک روش PCA و سپس برقراری یک رگرسیون خطی برای ساخت مدل. با این روش، متغیرهای اولیه به مؤلفه‌های جدید و مستقل با ضرایب همبستگی صفر برای هر دو مؤلفه تبدیل می‌شوند که مؤلفه‌های جدید، ترکیبی خطی از متغیرهای اولیه هستند. چون در تشکیل مؤلفه‌ها از تمام توصیف‌کننده‌ها استفاده می‌شود، در نتیجه اطلاعات توصیف‌کننده‌های اولیه با کمترین تلفات به وسیله مؤلفه‌های حاصل ارائه شده و باعث از دست رفتن جنبه‌های اطلاعاتی داده‌های اصلی نمی‌شود. [۲۱].

در روش MLR فرض بر این است که بهترین راه برای تخمین y از x به دست آوردن یک ترکیب خطی از متغیرهای موجود در x است به صورتی که خطای حاصل از پیش‌بینی y مینیمم شود. در این روش ارتباط بین x و y به صورت رابطه‌ی (۲-۱۹) بیان می‌شود.

$$y = xS \quad (۱۹-۲)$$

در این رابطه S ماتریسی است که از ضرایب رگرسیون تشکیل شده است. در روش رگرسیون مؤلفه-های اصلی ماتریس x با ماتریس x' جایگزین می‌شود. ماتریس x' دارای ستون‌هایی است که ترکیب خطی از ستون‌های ماتریس x هستند و تعداد سطرهای آن برابر با تعداد ردیف‌های ماتریس x است. به مؤلفه‌های به‌دست آمده از ماتریس x بردارهای ویژه^۱ نیز گفته می‌شود. بدین ترتیب معادله‌ی (۱۹-۲) به شکل زیر نوآرایی می‌گردد:

$$y = US \quad (۲۰-۲)$$

$$U = x \cdot V \cdot \lambda^{-1} \quad (۲۱-۲)$$

ماتریس U را ماتریس اسکور^۲ می‌نامند. مختصات هر داده نسبت به بردارهای ویژه، ماتریس اسکور را تشکیل می‌دهد. این ماتریس ابعادی برابر ماتریس x دارد و هر ستون آن یک مؤلفه‌ی اصلی، PC، است. در ماتریس V یا لودینگ^۳ تعداد سطرها برابر با تعداد ستون‌های ماتریس U است. این بردارها دارای ویژگی‌های خاصی هستند از جمله این که بردارهای اسکور دو به دو برهم عمودند و حاصل ضرب داخلی آن‌ها صفر است (اصطلاحاً این بردارها اورتوگونال^۴ هستند). بردارهای لودینگ طول واحد نیز دارند و به عبارتی اورتو نرمال^۵ هستند. سهم هر توصیف‌کننده در مؤلفه‌ی اصلی به کسینوس زاویه‌ی بین محور توصیف‌کننده و محور مؤلفه‌ی اصلی بستگی دارد. مقدار عددی کسینوس که گستره‌ای بین -۱ و +۱ اختیار می‌کند، ماتریس لودینگ را تشکیل می‌دهد. λ نیز یک ماتریس قطری تک مقدار است. در این ماتریس تمام عناصر قطری غیر صفر هستند. این مقادیر، مقادیر ویژه^۶ نامیده می‌شوند.

1- Eigen vector
2- Score
3- Loading
4- Orthogonal
5- Ortonormal
6- Eigen value

تفاوت دیگر مدل سازی MLR با PCR، وزن دار بودن مؤلفه های اصلی در مدل سازی PCR است. در مدل سازی MLR تمامی توصیف کننده ها وزن یکسان دارند اما در مدل سازی PCR هر مؤلفه ای اصلی با توجه به میزان اهمیت و اثرگذاری آن در مدل، وزن دار می شود. وزن هر مؤلفه ای اصلی در مدل سازی PCR با مقدار ویژه ی مؤلفه ای اصلی سنجیده می شود [۲۲]. پس از محاسبه ی ماتریس لودینگ ها، ماتریس مقادیر ویژه و ماتریس اسکورها و با جایگذاری مقادیر مربوطه در معادله ی (۲-۲۰)، ماتریس ضرایب مدل و از آنجا مدل حاصل از روش PCR به دست می آید.

۲-۳-۵-۳- رگرسیون کمترین مربعات جزئی (PLS)

روش PLS برای اولین بار در سال ۱۹۷۵ توسط ولد معرفی شد. PLS یک روش ساده اما مؤثر برای تخمین جواب با استفاده از الگوریتمی به نام نیپالس^۱ (کمترین مربعات جزئی متناوب غیر خطی) می باشد. [۵].

PLS روشی مشابه PCR است با این تفاوت که در آن تجزیه ماتریس داده ها به شکل متفاوتی انجام می شود، در روش PLS به جای شکستن جداگانه ی ماتریس اولیه به مجموعه ای از بردارهای ویژه و اسکورها ی مربوطه و سپس رگرسیون آن ها در مقابل متغیر وابسته، از اطلاعات متغیر وابسته به طور صحیحی در حین فرآیند شکستن ماتریس اولیه استفاده می شود. از این رو، بردارهای ویژه و اسکورها ی محاسبه شده در PLS و PCR متفاوت اند. در حقیقت مزیت روش PLS نسبت به روش PCR به کار بردن و ترکیب کردن اطلاعات بیشتر در ساخت مدل است.

$$X = T \cdot P \quad (۲-۲۲)$$

$$Y = U \cdot Q \quad (۲-۲۳)$$

1- Non linear iterative partial least squares (NIPALS)

در روابط فوق، T و U به ترتیب اسکورهای x و y هستند و P و Q نیز لودینگ‌ها می‌باشند اجزاء ماتریس T را متغیرهای پنهان^۱ (LV) می‌نامند.

مدل نهایی رگرسیون، از رابطه‌ی زیر به دست می‌آید:

$$y = xB \quad (24-2)$$

که در آن B ماتریس ضرایب رگرسیون PLS است که از رابطه‌ی زیر محاسبه می‌گردد [۲۱].

$$B = x^T U (T^T x x^T U)^{-1} T^T y \quad (25-2)$$

در مدل‌سازی PLS هر LV بر حسب ارتباط با ماتریس y وزن دار می‌گردد یعنی متغیرهایی که همبستگی بالایی با متغیرهای جواب نشان می‌دهند، وزن بیشتری پیدا می‌کنند زیرا آن‌ها در پیشگویی متغیر وابسته مؤثرتر خواهند بود [۲۳]. در رگرسیون PLS وزن هر متغیر پنهان در مدل با مقدار واریانس یا تغییراتی که هر متغیر پنهان شامل می‌شود، متناسب است. درصد واریانس یا تغییرات هر متغیر پنهان بر اساس میزان پراکندگی نقاط حول محور آن متغیر پنهان مشخص می‌شود. برای مثال چنانچه برای اولین متغیر پنهان درصد واریانس برابر ۹۸٪ محاسبه شود، بدین معنی است که خاصیت توصیف شده توسط اولین متغیر پنهان ۹۸٪ کل تغییرات مربوط به سیستم را به خود اختصاص داده و سایر متغیرهای پنهان مجموعاً ۲٪ تغییرات باقی‌مانده را در بر می‌گیرند.

روش PLS، ضرایب را به گونه‌ای تعیین می‌کند که مدل حاصل، بیشترین قدرت تفسیر و توضیح را دارا باشد، بدین معنا که مدل بتواند با بالاترین دقت و صحت، متغیر وابسته را پیش‌بینی نماید. معمولاً بین حالتی که در آن جواب شامل یک متغیر تنهاست، با حالتی که در آن جواب چند متغیره است فرق گذاشته می‌شود. حالت اول PLS1 و حالت دوم PLS2 نامیده می‌شود [۵ و ۲۴].

1- Latent variable

۲-۳-۴- شبکه‌ی عصبی مصنوعی

در بسیاری از کاربردهای شیمیایی، رابطه‌ی بین متغیرهای مستقل و پاسخ را نمی‌توان به صورت یک رابطه‌ی ساده‌ی خطی در نظر گرفت. در این حالت روش‌های رگرسیون خطی قادر به توصیف مناسب رابطه‌ی بین متغیرهای مستقل و وابسته نیستند. به همین دلیل روش‌های مدل‌سازی غیرخطی نظیر شبکه‌ی عصبی مصنوعی، به طور گسترده‌ای برای مدل‌سازی داده‌های شیمیایی به کار می‌روند.

۲-۳-۴-الف: ساختار شبکه‌ی عصبی مصنوعی

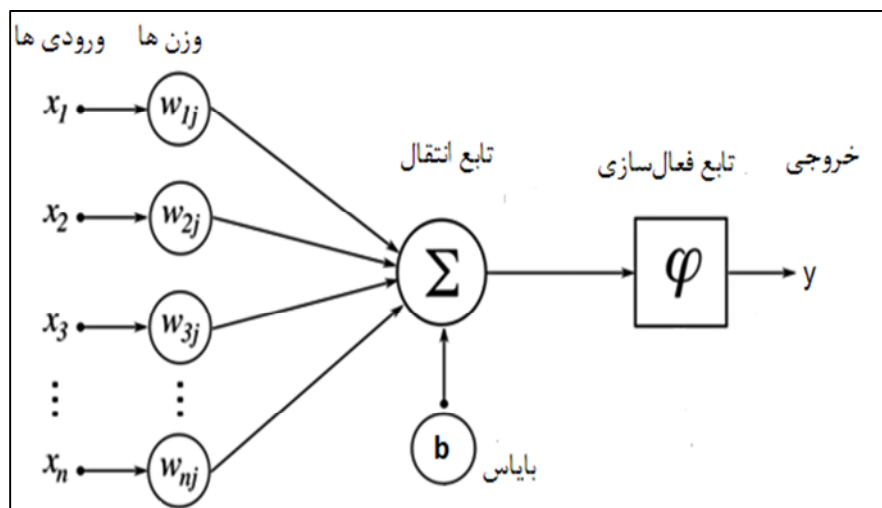
شبکه‌های عصبی مصنوعی با پردازش داده‌های تجربی، دانش یا قانون نهفته در ورای داده‌ها را به ساختار شبکه منتقل می‌کنند که به این عمل یادگیری گفته می‌شود. اصولاً توانایی یادگیری مهمترین ویژگی یک سیستم هوشمند است. شبکه‌ی عصبی مصنوعی، الگوریتمی برای پردازش اطلاعات می‌باشد که با تقلید از سیستم بیولوژیکی شبکه‌ی عصبی انسان ساخته شده است. عنصر کلیدی این شبکه ساختار سیستم پردازش اطلاعات آن است که از تعداد زیادی نرون^۱ با ارتباطات قوی داخلی که هماهنگ باهم برای حل مسائل کار می‌کنند، تشکیل شده است.

هر نرون مصنوعی همانند یک نرون زیستی شامل ورودی، جسم سلولی و یک خروجی است. هر ورودی قبل از ورود به هسته‌ی پردازشگر اصلی وزن دار می‌شود و سپس ورودی‌های وزن دار شده با هم جمع زده می‌شوند. وزن سیگنال بر انتشار آن در شبکه بسیار تأثیرگذار است. علاوه بر این نرون عصبی از یک بایاس^۲ هم برخوردار است. بایاس در واقع یک سیگنال ورودی به نرون است. سیگنال‌های وارد شونده به جسم سلولی که به هریک وزن ویژه‌ای اختصاص یافته است باهم جمع

1- Neuron

2- Bias

شده و پس از عبور از تابع انتقال^۱ به سیگنال‌های خروجی تبدیل می‌شوند. شکل (۱-۲) نمایی از یک شبکه‌ی عصبی مصنوعی نشان می‌دهد [۲۱].



شکل (۱-۲): نمایی از یک شبکه‌ی عصبی مصنوعی

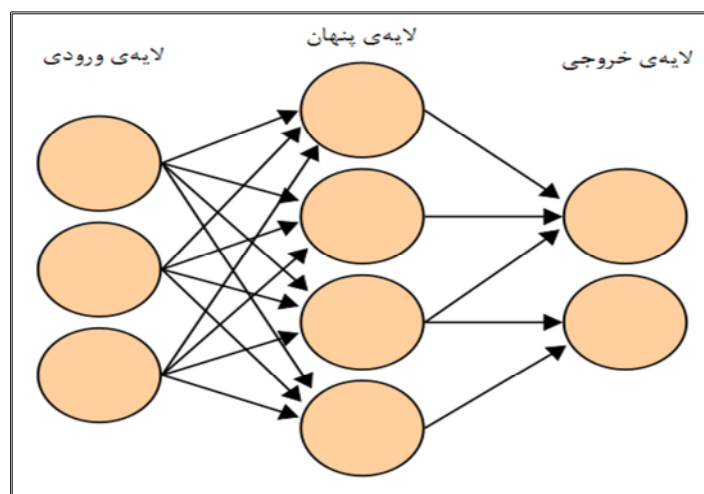
در شبکه‌های عصبی چند نوع اتصال و یا پیوند وزنی وجود دارد:

- پیشرو^۲: بیشترین پیوندهای شبکه‌های عصبی از این نوع است که در آن سیگنال‌ها تنها در یک جهت از ورودی به خروجی حرکت می‌کنند، هیچ بازخوردی بین نرون‌های یک لایه با نرون‌های لایه قبل وجود ندارد، خروجی هر لایه نیز بر همان لایه تأثیری نمی‌گذارد. رایج‌ترین شبکه‌ها، شبکه‌های چند لایه با تغذیه‌ی پیشرو هستند.
- پس‌رو^۳: داده‌ها از نرون‌های هر لایه به نرون‌های لایه‌های قبل از آن بازخورانده می‌شوند.
- جانبی: خروجی نرون‌های هر لایه به عنوان ورودی نرون‌های همان لایه استفاده می‌شوند.

1- Transfer function
2- Feed forward
3- Feed back

۲-۳-۵-۴-ب: شبکه‌های عصبی با تغذیه‌ی پیشرو

معمولاً سه لایه در این شبکه‌ها دیده می‌شود: لایه‌ی ورودی، لایه‌ی پنهان و لایه‌ی خروجی که نرون‌ها در این لایه‌ها چیده شده‌اند. تمام واحدهای لایه‌ی اول به تمام واحدهای لایه‌ی بعدی متصل شده‌اند. شبکه تمام سیگنال‌ها را از لایه‌ی ورودی دریافت می‌کند و پس از عبور از لایه‌ی پنهان به واحدهای لایه‌ی خروجی رسانده و خروجی شبکه را تولید می‌کند. ساختار عمومی این شبکه‌ها در شکل (۲-۲) نشان داده شده است.



شکل (۲-۲): نمایی از شبکه‌ی عصبی مصنوعی با تغذیه‌ی پیشرو

۲-۳-۵-۴-ج: آموزش شبکه عصبی مصنوعی

منظور از یادگیری یا آموزش، فرآیندی است که موجب می‌شود سیستم بر اثر تغییر در محرک‌های بیرونی، پاسخی متناسب با آن ارائه دهد. به عبارت دیگر، فرآیند پیدا کردن وزن‌ها و بایاس‌های مناسب برای شبکه‌ی عصبی مصنوعی، آموزش نامیده می‌شود. سه روش برای این منظور وجود دارد: یادگیری تحت نظارت^۱، یادگیری بدون نظارت^۲ یا خود سامانده.

1- Supervised
2- Unsupervised

الف) یادگیری تحت نظارت: این روش آموزش با تمرکز روی یک موضوع خاص و ارائه‌ی مثال‌های مختلفی از آن صورت می‌گیرد. شبکه، اطلاعات ورودی و مثال‌ها را تجزیه و تحلیل خواهد کرد به طوری که پس از مدتی قادر خواهد بود یک نوع جدید از آن دسته مثال‌ها را که قبلاً هرگز ندیده بود، شناسایی کند.

ب) یادگیری بدون نظارت یا یادگیری خود سامانده: در این یادگیری پارامترهای شبکه‌ی عصبی تنها توسط پاسخ سیستم اصلاح و تنظیم می‌شوند. به عبارتی تنها اطلاعات دریافتی شبکه را بردارهای ورودی تشکیل می‌دهند و به شبکه‌ی عصبی هیچ نمونه‌ای از تابعی که قرار است بیاموزد داده نمی‌شود.

آموزش شبکه‌های با تغذیه پیشرو معمولاً به صورت تحت نظارت انجام می‌شود و در طی دوره‌ی آموزش اوزان طبق قانون آموزش اصلاح می‌گردند. رایج‌ترین الگوریتم آموزش این شبکه‌ها، قانون آموزش انتشار معکوس است که در آن تنظیم کردن اوزان با توجه به تفاوت خروجی واقعی و خروجی مورد نظر صورت می‌گیرد. در این روش ابتدا یک مقدار تصادفی به اوزان اختصاص داده می‌شود سپس مقدار خروجی با استفاده از این اوزان به دست آمده و خطا به کمک تفاوت آن با خروجی واقعی به دست می‌آید.

از رایج‌ترین الگوریتم‌های آموزشی تحت نظارت، الگوریتم‌های لوبنرگ-مارکوات و تنظیم بایزین می‌باشد. الگوریتم آموزشی لوبنرگ-مارکوات در سال ۱۹۹۴ معرفی شد این الگوریتم روش مناسبی برای آموزش شبکه‌هایی است که شاخص عملکرد آن‌ها حداقل متوسط مربعات خطا می‌باشد. این روش تاکنون به عنوان سریع‌ترین و مؤثرترین روش آموزش شبکه‌های عصبی شناخته شده است. سرعت همگرایی این روش صدها برابر بیشتر از شیوه‌های استاندارد مشابه است [۲۱].

معادله‌ی مربوط به الگوریتم لوبنرگ-مارکوات برای بهینه‌سازی پارامترهای شبکه و کاهش تابع خطا در هر مرحله طبق رابطه‌ی (۲-۲۶) می‌باشد:

$$X_{k+1} = X_k - [J^T(X_k)J(X_k) + \mu_k I]^{-1} J^T(X_k)E(X_k) \quad (26-2)$$

که در آن X برداری است که مؤلفه‌هایش پارامترهای تنظیم پذیر شبکه (وزن‌ها و بایاس‌ها) هستند. k نشان‌دهنده‌ی مرحله‌ی آموزش، J معرف ماتریس ژاکوبین، I ماتریس یکانی و μ ضریب یادگیری مارکوات و عددی غیر منفی است. ماتریس ژاکوبین چنین تعریف می‌شود:

$$\begin{bmatrix} \frac{\partial e_1(x)}{\partial x_1} & \frac{\partial e_1(x)}{\partial x_2} & \dots & \frac{\partial e_1(x)}{\partial x_n} \\ \frac{\partial e_2(x)}{\partial x_1} & \frac{\partial e_2(x)}{\partial x_2} & \dots & \frac{\partial e_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_n(x)}{\partial x_1} & \frac{\partial e_n(x)}{\partial x_2} & \dots & \frac{\partial e_n(x)}{\partial x_n} \end{bmatrix} \quad (27-2)$$

یعنی ماتریس ژاکوبین شامل مشتقات مرتبه اول خطای شبکه نسبت به پارامترهای شبکه است.

برای جلوگیری از آموزش بیش از حد شبکه و ارتقاء عمومیت شبکه، الگوریتم آموزشی تنظیم بایزین پیشنهاد شده است که در این الگوریتم تابع عملکرد متداول (MSE) از طریق افزایش عبارات میانگین مربعات وزن‌ها و بایاس‌های شبکه اصلاح شده‌است.

$$F = \beta E_D + \alpha E_W \quad (28-2)$$

که در آن α و β پارامترهای تابع کارایی، E_D و E_W به ترتیب میانگین مربعات خطاها و میانگین مربعات وزن‌ها می‌باشند که به صورت زیر تعریف می‌شوند:

$$E_D = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2 \quad (29-2)$$

$$E_W = \frac{1}{N} \sum_{i=1}^N (w_j)^2 \quad (30-2)$$

که در این روابط e_i بیانگر خطای مطلق، t_i مقادیر تجربی، y_i مقادیر محاسبه شده برای پارامتر تجربی مدنظر و N نشان‌دهنده‌ی تعداد نمونه‌ها می‌باشد. تابع عملکرد، ترکیبی از میانگین مربعات خطا و میانگین مربع وزن‌هاست [۲۱].

۲-۳-۶- ارزیابی اعتبار مدل‌ها

به طور کلی می‌توان گفت مدلی مناسب‌تر است که ویژگی‌های زیر را دارا باشد:

- تعداد متغیرهای مستقل آن از سایر مدل‌ها کمتر باشد.
- ضریب همبستگی بین توصیف‌کننده‌هایی که در مدل شرکت دارند کم باشد.
- قدرت پیش‌بینی بالایی داشته باشد.
- پارامترهای آماری معتبری داشته باشد.

از جمله پارامترهای آماری که بایستی در مدل‌سازی خطی و غیر خطی مورد ارزیابی قرار بگیرد ضریب تعیین R^2 است که هرچه مقدار آن به یک نزدیک‌تر باشد، قدرت پیش‌بینی مدل بیشتر است. به علاوه می‌توان از پارامترهای MSE ، D_{max} ، AAD ، $bias$ ، $PRESS$ و MAE که در بخش (۲-۳-۴-۱) معرفی شدند نیز برای ارزیابی مدل استفاده نمود.

به منظور ارزیابی مستقل بودن مدل‌ها نسبت به سری داده‌ها، از روش ارزیابی تقاطعی^۱ (CV) استفاده می‌شود. در این روش، هر بار تعداد معینی از مولکول‌ها کنار گذاشته می‌شوند و با بقیه مولکول‌ها، مدل‌سازی انجام می‌شود. سپس مقدار متغیر وابسته برای مولکول‌های کنار گذاشته شده، پیش‌بینی می‌شود. این کار آن‌قدر تکرار می‌شود تا تمام مولکول‌ها، حداقل یکبار در دسته‌ای که در مدل‌سازی شرکت نمی‌کند، قرار گیرند. بسته به تعداد مولکول‌های کنار گذاشته شده، این روش به دو شیوه حذف یکتایی^۲ (LOO) و حذف چندتایی^۳ (LGO) قابل انجام است. در شیوه LOO هر بار یک مولکول از سری داده‌ها کنار گذاشته شده و مدل‌سازی با بقیه مولکول‌ها انجام می‌پذیرد. اما در شیوه LGO هر بار یک چندتایی از مولکول‌های سری داده‌ها کنار گذاشته می‌شود. بقیه مراحل LGO مانند LOO است. [۲۵ و ۲۶].

1- Cross validation

2- Leave one out

3- Leave group out

همچنین می‌توان برای ارزیابی قدرت پیش‌بینی مدل، از روش اعتبار سنجی بیرونی استفاده نمود. در این روش قسمتی از داده‌ها از ابتدا کنار گذاشته می‌شوند و در هیچیک از مراحل مدل‌سازی از اطلاعات این داده‌ها استفاده نمی‌شود. پس از ساخت و بهینه‌سازی مدل، داده‌های کنار گذاشته شده به مدل وارد می‌شوند و مقادیر متغیر وابسته‌ی مربوط به آن‌ها توسط مدل محاسبه می‌شود. با مقایسه‌ی مقادیر حقیقی و مقادیر محاسبه شده توسط مدل، می‌توان کارایی مدل را برای پیش‌بینی داده‌های جدید مورد ارزیابی قرار داد. این روش در مورد سنجش کارایی سایر روش‌های خطی و غیرخطی مدل‌سازی نیز کاربرد دارد.

آزمون تصادفی‌کردن داده‌ها^۱ یا به عبارتی آزمون Y- تصادفی نیز روشی برای ارزیابی اعتبار مدل به دست آمده از روش شبکه‌ی عصبی مصنوعی است که هدف این روش بررسی هر گونه ارتباط تصادفی بین داده‌هاست. در این روش به جای متغیر وابسته درست، داده‌ی تصادفی داده می‌شود. اگر تفاوت چشمگیری بین نتایج حاصل از داده‌های تصادفی و داده‌های تجربی وجود داشته باشد نشان‌دهنده‌ی صحت و اعتبار کار است [۲۷].

1- Y- randomization test

فصل سوم

بخش محاسبات

۳-۱- مدل سازی ضریب خودنفوذی آلکان های خطی

در این تحقیق برای مدل سازی ضریب خودنفوذی آلکان های خطی از دو روش خطی رگرسیون مؤلفه اصلی (PCR) و رگرسیون کمترین مربعات جزئی (PLS) و روش غیرخطی شبکه‌ی عصبی مصنوعی استفاده شده است. در ادامه به بررسی مراحل طراحی، بهینه سازی و ارزیابی مدل های ساخته شده پرداخته خواهد شد.

۳-۱-۱- انتخاب سری داده ها

با توجه به هدف تحقیق، داده های مربوط به ضریب خودنفوذی ۱۷ آلکان خطی مایع از متان تا n-هپتادکان در دامنهی وسیعی از دما و فشار جمع آوری گردید. نام ترکیبات به همراه محدوده ی دما و فشار و همچنین تعداد نقطه داده ی مورد استفاده از هر ترکیب به طور خلاصه در جدول (۳-۱) آورده شده است. همچنین سری داده های مربوط به این تحقیق به طور کامل در جدول (پ-۱) گزارش شده است.

۳-۱-۲- رسم و بهینه سازی ساختار مولکول ها

به منظور رسم و بهینه سازی مولکول ها از نرم افزار Hyperchem 7.0 استفاده شد. این نرم افزار امکان رسم سه بعدی مولکول ها را فراهم می آورد و اطلاعات فراوانی نظیر زوایای پیوندی، طول پیوندها، زوایای پیچشی و ... را در اختیار قرار می دهد. به علاوه این بسته ی نرم افزاری توانایی بهینه سازی ساختمان مولکول ها را نیز دارد. محاسبه ی توصیف کننده های ساختاری بر مبنای هندسه ی دقیق مولکولی استوار است و اگر ساختار مولکول ها به صورت صورتبندی با حداقل انرژی نباشد، مقادیر غیر صحیحی برای این توصیف کننده ها به دست می آید. بنابراین، باید ساختار هندسی مولکول ها

بهینه شود و مختصاتی از ساختار ترکیب به دست آید که در آن مختصات کارترین، انرژی مولکول حداقل باشد. در این تحقیق ساختار مولکول‌ها با استفاده از روش نیمه تجربی ¹AM1 موجود در این بسته نرم‌افزاری بهینه شد. در این روش فقط الکترون‌های لایه‌ی ظرفیت در محاسبات وارد می‌شوند.

جدول (۱-۳): محدوده‌ی دما و فشار ضریب خودنفوذی و تعداد داده‌ی استفاده شده در مدل‌سازی برای هر آلکان به همراه مرجع مورد استفاده برای هر یک از آن‌ها

مرجع	تعداد نقطه داده استفاده شده در مدل‌سازی	گستره‌ی فشاری (MPa)	گستره‌ی دمایی (K)	نام ترکیب
[۲۸ و ۲۹ و ۳۰ و ۳۱ و ۳۲ و ۳۳]	۳۰	۱/۵۳ - ۲۲۱/۵۸	۱۱۰ - ۱۶۰	متان
[۲۸ و ۲۹ و ۳۰]	۴۸	۲۵ - ۲۰۰	۱۳۶ - ۳۳۳/۱	اتان
[۲۸]	۷۴	۲۵ - ۲۰۰	۱۱۲ - ۳۵۳	پروپان
[۳۴]	۳۴	۴/۲ - ۲۰۰	۱۹۳ - ۴۱۶/۴۲	n-بوتان
[۳۵ و ۳۶]	۹۰	۰/۱ - ۲۴۵	۱۷۴ - ۴۹۷/۵	n-پنتان
[۲۹ و ۳۰ و ۳۷ و ۳۸ و ۳۹ و ۴۰ و ۴۱ و ۴۲]	۹۹	۰/۱ - ۳۹۳/۸	۲۱۴ - ۴۵۵	n-هگزان
[۳۵ و ۳۶ و ۳۷ و ۳۹ و ۴۲]	۵۸	۰/۱ - ۲۴۵	۱۹۵/۵ - ۳۵۳	n-هپتان
[۳۰ و ۳۷ و ۳۸ و ۴۳]	۵۳	۰/۱ - ۳۶۰/۸	۲۴۸/۱۱ - ۳۴۸/۹	n-اکتان
[۳۶ و ۳۷]	۵۵	۰/۱ - ۲۴۵	۲۹۳ - ۳۵۳	n-نونان
[۳۰ و ۳۴ و ۳۷]	۴۳	۵ - ۲۴۰	۲۶۴ - ۴۹۷/۵	n-دکان
[۳۶]	۵۵	۰/۱ - ۲۴۵	۲۹۳ - ۳۵۳	n-آن‌دکان
[۳۷ و ۳۸ و ۴۴ و ۴۵]	۱۰	۰/۱ - ۶۰	۲۹۸/۲ - ۴۴۳	n-دودکان
[۳۶]	۵۰	۰/۱ - ۲۴۵	۲۹۳ - ۳۵۳	n-تری‌دکان
[۳۷ و ۴۴ و ۴۶]	۲۷	۰/۱ - ۲۰۰	۲۹۳ - ۴۴۳/۱	n-تترادکان
[۳۶]	۵۲	۰/۱ - ۲۴۵	۲۹۳ - ۳۵۳	n-پنتاکان
[۳۷ و ۳۸ و ۴۲ و ۴۵ و ۴۷ و ۴۸]	۹۰	۰/۱ - ۶۰۰	۲۹۸/۱۵ - ۵۱۳	n-هگزادکان
[۳۶]	۴۴	۰/۱ - ۲۴۵	۳۰۳ - ۳۵۳	n-هپتادکان

1-Austian method

۳-۱-۳- محاسبه‌ی توصیف‌کننده‌ها

به منظور تبدیل ساختار سه بعدی مولکول به توصیف‌کننده‌های ساختاری از نرم‌افزار دراگون^۱ نسخه‌ی ۲/۱ استفاده شد. این نرم‌افزار توسط گروه کمومتریکس دانشگاه میلانو طراحی شده است و قابلیت محاسبه‌ی ۱۴۸۱ توصیف‌کننده برای هر مولکول را دارا می‌باشد. برای این منظور فایل خروجی نرم‌افزار Hyperchem به نرم‌افزار دراگون برده شد و ۱۴۸۱ توصیف‌کننده مولکولی در ۱۸ دسته بندی مختلف برای هر مولکول محاسبه گردید. به‌علاوه، ۲۱ توصیف‌کننده‌ی الکترونی نیز توسط نرم‌افزار Hyperchem به دست آمد. با توجه به وابستگی ضریب خودنفوذی به دما و فشار، دو متغیر تجربی دما و فشار نیز به توصیف‌کننده‌های ساختاری اضافه شدند و در مجموع برای هر مولکول ۱۵۰۴ توصیف‌کننده حاصل گردید.

۳-۱-۴- کاهش تعداد توصیف‌کننده‌ها

یکی از مشکلاتی که در هنگام ساخت مدل‌های QSPR/QSAR وجود دارد، تعداد زیاد متغیرهای مستقل است. معمولاً تعداد ترکیبات شیمیایی مورد بررسی، در مقایسه با تعداد توصیف‌کننده‌های مستقل کم بوده و این امر روش‌های مدل‌سازی را با محدودیت مواجه می‌کند. همچنین برخی از توصیف‌کننده‌ها دارای مقادیر ثابت هستند و برخی دیگر با سایر توصیف‌کننده‌ها همبستگی بالایی دارند، لذا بایستی این توصیف‌کننده‌ها شناسایی و حذف شوند. برای این منظور در مرحله‌ی اول توصیف‌کننده‌هایی که تمامی مقادیر آن‌ها صفر بود و همچنین توصیف‌کننده‌هایی که ۹۰ درصد مقادیر آن‌ها یکسان بود، حذف گردید.

در دومین مرحله، با توجه به این که برخی توصیف‌کننده‌ها با یکدیگر همبستگی بالایی دارند

1- Dragon

و تنها وجود یکی از آن‌ها در ساخت مدل کافی است و نیازی به حضور مابقی توصیف‌کننده‌ها نیست، توصیف‌کننده‌هایی که با یکدیگر همبستگی بالاتر از ۰/۹ داشتند توسط برنامه‌ای نوشته شده در نرم‌افزار Matlab مشخص و حذف شدند. در این مرحله بایستی توجه داشت که توصیف‌کننده‌هایی نگه داشته شدند که بیشترین همبستگی را با متغیر وابسته (در اینجا ضریب خودنفوذی) داشتند. در پایان این مرحله، ۲۹ توصیف‌کننده باقی ماند. این توصیف‌کننده‌ها عبارتند از:

Xt, CIC5, JGI2, JGI3, JGI5, Qmean, PJI3, Mor12u, Mor25u, Mor18v, Mor20v, Mor26e, Mor08p, Mor16p, Mor30p, E1u, E1p, P2s, G2s, HATS0m, HATS2m, HATS4m, R2u_A, R3m, R6m, R8m, DipolY, T, P.

۳-۱-۵- خود مقیاس‌سازی داده‌ها:

پس از حذف توصیف‌کننده‌های اضافی، بر روی ماتریس توصیف‌کننده‌ها فرآیند خود مقیاس‌سازی^۱ انجام گردید. در این مرحله مقادیر توصیف‌کننده‌ها بر اساس رابطه‌ی (۳-۱) به نحوی خود مقیاس شدند که دارای مقدار میانگین صفر و انحراف معیار یک باشند.

$$Z_i = \frac{X_i - \mu}{\sigma} \quad (۳-۱)$$

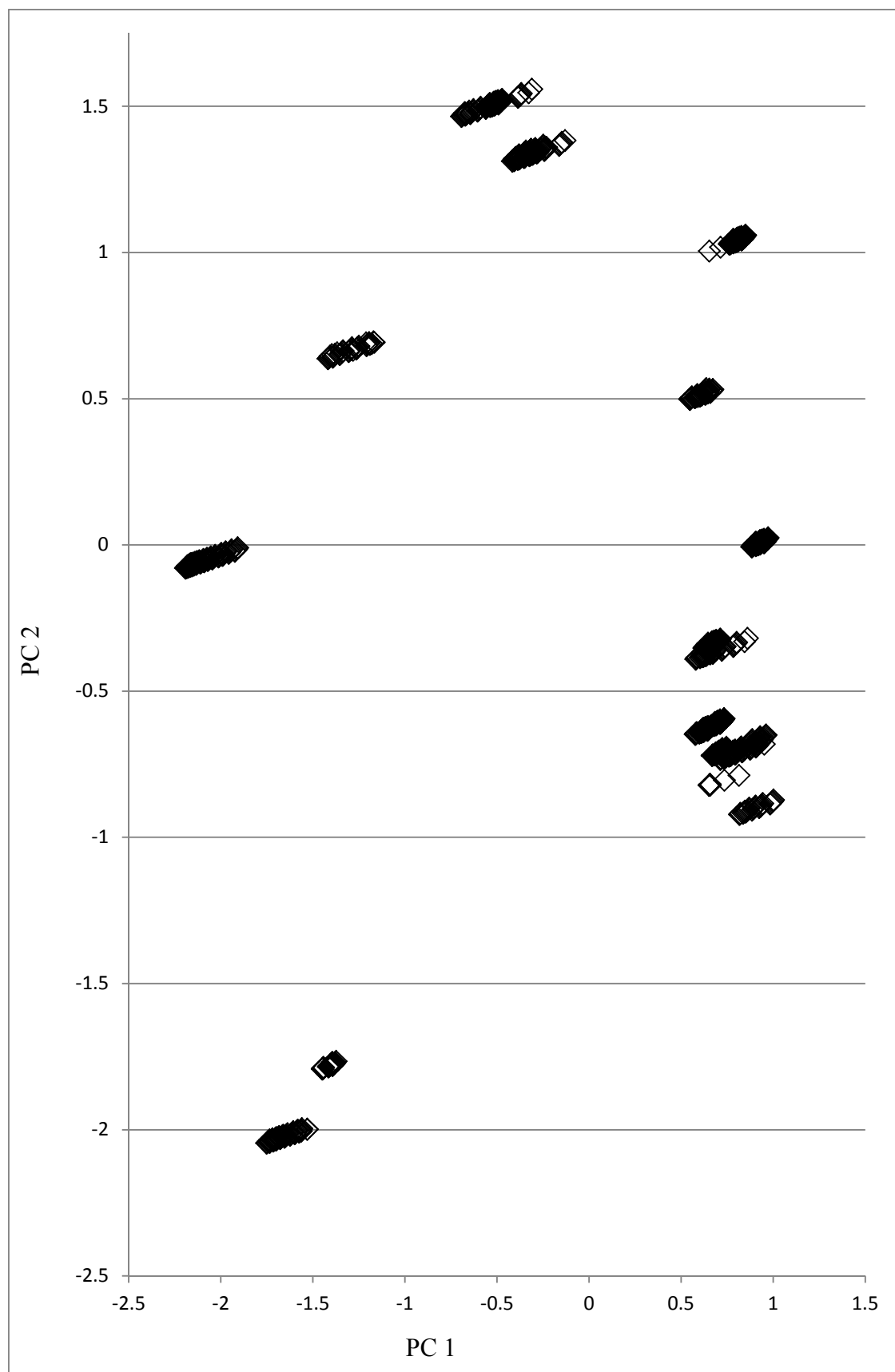
در این معادله، Z_i معادل مقدار خود مقیاس شده‌ی داده، X_i داده‌ی ورودی، μ میانگین هر توصیف‌کننده و σ نیز مقدار انحراف استاندارد برای هر توصیف‌کننده است. این عمل باعث می‌شود که تأثیر هر توصیف‌کننده از دامنه‌ی عددی آن جدا شده و دامنه‌ی تغییرات یک توصیف‌کننده، تأثیر کاذب بر روی میزان اهمیت آن ایجاد نکند [۴۹].

1-Autoscaling procedure

۳-۱-۶- دسته‌بندی داده‌ها

پیش از شروع مرحله‌ی انتخاب توصیف‌کننده‌های مناسب، داده‌ها باید به سه گروه آموزش، ارزیابی و تست تقسیم گردند. در این تحقیق برای دسته‌بندی داده‌ها از روش تحلیل مؤلفه‌های اصلی (PCA) که در فصل ۲ تشریح شد، استفاده گردید. برای این منظور در بسته‌ی نرم‌افزاری SPSS^۱، در سربرگ Analyze، گزینه‌ی Dimension Reduction و بعد گزینه‌ی Factor انتخاب شد تا مؤلفه‌های اصلی ساخته شوند. با توجه به این‌که معمولاً دو مؤلفه‌ی اول و دوم مقدار قابل توجهی از پراکندگی داده‌ها را در بر می‌گیرند، لذا نمودار پراکندگی براساس تغییرات مؤلفه‌ی اول برحسب مؤلفه‌ی دوم رسم شد. در شکل (۳-۱) پراکندگی داده‌ها بر روی نمودار PCA نشان داده شده است. بر اساس این نمودار نقطه‌های داده‌های محیطی در سری آموزش قرار گرفت و سری ارزیابی و تست از داده‌های داخلی به طور تصادفی انتخاب شدند. بدین ترتیب تعداد ۶۰۲ داده در سری آموزش، ۱۰۴ داده در سری ارزیابی و ۹۹ داده در سری تست قرار گرفتند. همچنین ۱۰۷ داده مربوط به دو مولکول n-پنتان و n-نونان به عنوان سری تست بیرونی انتخاب شدند.

1- Statistical package for the social science



شکل (۳-۱): پراکندگی داده‌ها بر روی نمودار PCA

۳-۱-۷- انتخاب بهترین توصیف‌کننده‌ها

پس از دسته‌بندی داده‌ها، بایستی بهترین توصیف‌کننده‌ها برای ساخت مدل انتخاب شوند. در این تحقیق از روش خطی رگرسیون مرحله‌ای و روش غیرخطی الگوریتم ژنتیک برای انتخاب توصیف‌کننده‌های مناسب استفاده شد که نتایج حاصل از هر روش در ادامه گزارش شده است.

۳-۱-۷-۱- انتخاب بهترین توصیف‌کننده‌ها به روش رگرسیون مرحله‌ای (SR)

در این مرحله داده‌های سری آموزش به نرم‌افزار SPSS داده شد سپس رگرسیون مرحله‌ای انجام شد و مدل‌های پیشنهادی ساخته شدند. بدین ترتیب تعداد و نوع توصیف‌کننده‌های پیشنهادی هر مدل مشخص شدند. مدلی که مقدار MSE کمتر، تعداد توصیف‌کننده‌ی کمتر، مقدار ضریب تعیین (R^2) بیشتر، مقدار FIT بزرگتر و بیشترین آماره‌ی F برای سری ارزیابی را دارا بود به عنوان بهترین مدل و توصیف‌کننده‌های آن به عنوان توصیف‌کننده‌های انتخاب شده توسط رگرسیون مرحله‌ای گزینش شدند. توصیف‌کننده‌های انتخاب شده با استفاده از روش رگرسیون مرحله‌ای و پارامترهای آماری مربوط به هر مدل، در جدول (۳-۲) گزارش شده است.

جدول (۳-۲): مقایسه‌ی پارامترهای آماری مدل‌های ساخته شده توسط روش SR

مدل	توصیف‌کننده‌ها	R ²	MSE	F	FIT
۱	Mor12u	۰/۳۵۲	۰/۴۱۰۵	۵۵/۳۶	۰/۵۴۱
۲	Mor12u , T	۰/۵۵۸	۰/۳۲۶۲	۱۸۲/۹۰	۱/۲۵۱
۳	Mor12u , T , P	۰/۶۲۹	۰/۲۹۹۸	۱۶۲/۸۷	۱/۶۶۸
۴	Mor12u , T , P , R2u-A	۰/۶۵۰	۰/۱۹۴۹	۳۳۷/۲۷	۱/۸۰۹
۵	Mor12u , T , P , R2u-A , CIC5	۰/۶۶۹	۰/۱۴۴۷	۲۷۷/۰۷	۱/۹۴۵
۶	Mor12u , T , P , R2u-A , CIC5 , HATS0m	۰/۶۸۰	۰/۱۱۲۲	۳۷۷/۵۵	۲/۰۱۶
۷	Mor12u , T , P , R2u-A , CIC5 , HATS0m , Mor26e	۰/۶۸۳	۰/۱۲۳۵	۳۴۰/۹۰	۲/۰۱۰
۸	Mor12u , T , P , R2u-A , CIC5 , HATS0m , Mor26e , P2s	۰/۶۸۷	۰/۱۲۸۲	۳۱۰/۸۷	۲/۰۱۰

با توجه به پارامترهای آماری به دست آمده مدل ۶، به عنوان مدل بهینه انتخاب گردید که نام کامل توصیف‌کننده‌های وارد شده در آن به همراه اثر متوسط^۱ هر یک از آن‌ها در جدول (۳-۳) گزارش شده است.

پارامتر اثر متوسط از جمله پارامترهایی است که می‌توان به وسیله‌ی آن مقدار و نحوه‌ی اثرگذاری هر توصیف‌کننده بر خاصیت مورد نظر را تخمین زد. این مقدار از معادله‌ی (۳-۲) محاسبه می‌شود.

$$t = \frac{\sum_{i=1}^n X_i B_i}{N} \quad (۳-۲)$$

که در آن t اثر متوسط توصیف‌کننده، X_i مقدار عددی توصیف‌کننده برای مولکول i ام، B_i ضریب توصیف‌کننده در رگرسیون مرحله‌ای و N تعداد کل مولکول‌ها است.

1- Mean effect

جدول (۳-۳): نماد، گروه، نام کامل و اثر متوسط توصیف‌کننده‌های انتخاب شده توسط روش SR

شماره	نماد	گروه	نام کامل	اثر متوسط
۱	Mor12u	3D MORSE	3D-MORSE-signal 12	۲۱/۸۲۰
۲	T	experimental	Temperature	۲۶/۱۶۹
۳	P	experimental	Pressure	-۱۲/۳۳۹
۴	R2u-A	GETAWAY	R autocorrelation of lag 2	۷/۵۵۴
۵	CIC5	Topologycal	Complementary information Content	۷/۲۹۴
۶	HATS0m	GETAWAY	Leverage- weighted utocorrelation of lag 0	۴/۵۵۵

همچنین در جدول (۴-۳) ماتریس همبستگی بین توصیف‌کننده‌های مدل ارائه شده است که عدم همبستگی بین آن‌ها به خوبی مشهود است.

جدول (۴-۳): ماتریس همبستگی میان توصیف‌کننده‌های انتخاب شده توسط روش SR

	Mor12u	T	P	R2u-A	CIC5	HATS0m
Mor12u	۱					
T	-۰/۶۵۵	۱				
P	-۰/۱۱۰	۰/۱۸۵	۱			
R2u-A	۰/۱۸۸	-۰/۳۱۲	-۰/۰۶۸	۱		
CIC5	-۰/۷۱۰	۰/۴۹۷	۰/۱۴۰	-۰/۴۷۸	۱	
HATS0m	۰/۶۰۹	-۰/۲۸۶	-۰/۰۳۶	-۰/۵۰۹	-۰/۴۲۵	۱

مقادیر توصیف‌کننده‌های به دست آمده از روش رگرسیون مرحله‌ای برای مولکول‌های موجود در سری داده‌ها در جدول (۵-۳) گزارش شده است.

جدول (۳-۵) مقادیر توصیف‌کننده‌های انتخاب شده توسط روش SR

ردیف	نام مولکول	نام توصیف‌کننده			
		Mor12u	R2u-A	CIC5	HATS0m
۱	متان	۰/۱۶۲	۲/۴۷۳	۰	۰/۰۱۶
۲	اتان	۰/۰۰۶	۲/۰۶۹	۲/۱۸۹	۰/۰۱۸
۳	پروپان	-۰/۲۴۷	۲/۰۵۱	۱/۷۷۴	۰/۰۲۶
۴	n-بوتان	-۰/۲۸۹	۲/۰۱۴	۱/۹۶۵	۰/۰۲۵
۵	n-پنتان	-۰/۳۹۷	۱/۹۷۷	۱/۷۳۶	۰/۰۲۷
۶	n-هگزان	-۰/۵۲۷	۲/۰۰۳	۱/۸۷۵	۰/۰۲۵
۷	n-هپتان	-۰/۵۷۴	۲/۰۱۳	۱/۷۱۸	۰/۰۲۴
۸	n-اکتان	-۰/۵۹۸	۲/۰۳۰	۱/۸۲۷	۰/۰۲۲
۹	n-نونان	-۰/۶۶۰	۲/۰۳۴	۱/۷۰۷	۰/۰۲۱
۱۰	n-دکان	-۰/۶۹۷	۲/۰۴۹	۱/۷۹۷	۰/۰۱۹
۱۱	n-آن‌دکان	-۰/۷۹۲	۲/۰۵۳	۱/۸۵۸	۰/۰۱۸
۱۲	n-دودکان	-۰/۸۷۳	۲/۰۶۴	۱/۹۸۷	۰/۰۱۷
۱۳	n-تری‌دکان	-۰/۹۰۰	۲/۰۶۷	۲/۱۳۴	۰/۰۱۶
۱۴	n-تترادکان	-۰/۹۵۱	۲/۰۷۶	۲/۲۸۵	۰/۰۱۵
۱۵	n-پنتادکان	-۱/۰۰۸	۲/۰۷۸	۲/۴۳۵	۰/۰۱۵
۱۶	n-هگزادکان	-۱/۰۵۲	۲/۰۸۵	۲/۵۸۰	۰/۰۱۴
۱۷	n-هپتادکان	-۱/۱۳۱	۲/۰۸۷	۲/۷۲۱	۰/۰۱۳

۳-۱-۷-۲- انتخاب بهترین توصیف‌کننده‌ها با استفاده از الگوریتم ژنتیک

به منظور انتخاب بهترین توصیف‌کننده‌ها توسط روش الگوریتم ژنتیک، از یک M-file

طراحی شده توسط ریکاردو لیاردی^۱ استفاده شد. هر الگوریتم ژنتیک برای ارزیابی عملکرد خود از یک

¹-Ricardo Leardi

روش رگرسیون درونی بهره می‌گیرد که در این برنامه روشی که الگوریتم ژنتیک برای بررسی عملکرد خود از آن استفاده می‌کند روش PLS می‌باشد. در ابتدا ماتریس داده‌های سری آموزش به همراه متغیر وابسته به عنوان ورودی به برنامه داده شد و سپس برنامه اجرا گردید. در این برنامه به صورت پیش فرض بهینه احتمال پیوند ۵۰ درصد، احتمال جهش ۱ درصد و تعداد کروموزوم‌ها ۳۰ در نظر گرفته شده است. در این تحقیق برنامه برای ۱۰۰ بار اجرا تنظیم شد و اجرای آن ۱۲ بار تکرار گردید تا احتمال انتخاب توصیف‌کننده‌های تصادفی به حداقل برسد. در جدول (۳-۶) توصیف‌کننده‌های منتخب برای ۱۲ بار اجرای الگوریتم ژنتیک گزارش شده است.

جدول (۳-۶): توصیف‌کننده‌های منتخب برای ۱۲ بار اجرای الگوریتم ژنتیک

توصیف‌کننده‌های منتخب توسط الگوریتم ژنتیک	شماره اجرای برنامه
T , P , PJI3 , E1p , Xt , G2s	۱
T , P , PJI3 , Xt , E1p , G2s	۲
T , P , E1p , PJI3 , G2s , Xt , Mor12u	۳
T , P , PJI3 , E1p , Mor12u , G2s	۴
T , P , PJI3 , Xt , Mor12u , G2s	۵
T , P , PJI3 , E1p , G2s , Xt	۶
T , P , PJI3 , E1p , G2s , Xt	۷
T , P , PJI3 , E1p , Xt , G2s	۸
T , P , E1p , Xt , PJI3 , G2s	۹
T , P , PJI3 , E1p , Xt , G2s , Hats0m , R2u-A	۱۰
T , P , PJI3 , Xt , E1p , G2s	۱۱
T , P , PJI3 , E1p , Xt , G2s	۱۲

از بین توصیف‌کننده‌های گزارش شده در جدول (۳-۶)، توصیف‌کننده‌هایی که بیش از ۶ بار تکرار شده‌اند، (بیش از ۵۰ درصد) انتخاب شدند. نام کامل توصیف‌کننده‌های انتخاب شده به روش

الگوریتم ژنتیک و اثر متوسط هر یک از آن‌ها در جدول (۷-۳) گزارش شده است.

جدول (۷-۳): نماد، گروه، نام کامل و اثر متوسط توصیف‌کننده‌های انتخاب شده توسط روش GA

شماره	نماد	گروه	نام کامل	اثر متوسط
۱	T	experimental	Temperature	۲۶/۵۵۸
۲	P	experimental	Pressure	-۱۲/۴۲۸
۳	PJI3	geometrical	3D Petijean shape index	-۱۸/۲۰۸
۴	Xt	Topologycal	Total Structure Connectivity index	۱۹/۴۱۰
۵	E1p	WHIM	1st component accessibility directional WHIM index	-۱۵/۸۹۱
۶	G2s	WHIM	3st component symmetry directional WHIM index	-۱۱/۹۹۹

در جدول (۸-۳) نیز ماتریس همبستگی بین توصیف‌کننده‌های انتخاب شده به روش

الگوریتم ژنتیک ارائه شده است. مقادیر موجود در این جدول به خوبی عدم همبستگی بین توصیف‌کننده‌ها را تأیید می‌کنند.

جدول (۸-۳): ماتریس همبستگی میان توصیف‌کننده‌های انتخاب شده توسط روش GA

	T	P	PJI3	Xt	E1p	G2s
T	۱					
P	۰/۱۸۵	۱				
PJI3	۰/۳۱۳	۰/۰۴۰	۱			
Xt	-۰/۴۸۶	-۰/۰۳۱	-۰/۴۶۴	۱		
E1p	۰/۳۲۴	۰/۰۷۲	-۰/۱۳۶	۰/۱۹۹	۱	
G2s	۰/۱۴۷	۰/۰۷۲	-۰/۶۴۳	۰/۱۷۳	۰/۳۵۴	۱

مقادیر توصیف‌کننده‌های انتخاب شده به روش الگوریتم ژنتیک برای مولکول‌های موجود در سری داده‌ها، در جدول (۹-۳) گزارش شده است.

جدول (۹-۳) مقادیر توصیف‌کننده‌های انتخاب شده توسط روش GA

ردیف	نام مولکول	نام توصیف‌کننده			
		PJI3	Xt	E1p	G2s
۱	متان	۰/۶۴۰	۰	۰/۲۱۵	۰
۲	اتان	۰/۴۳۵	۱/۰۰۰	۰/۶۷۵	۱/۰۰۰
۳	پروپان	۰/۹۸۶	۰/۷۰۷	۰/۴۲۸	۰/۳۸۷
۴	n-بوتان	۰/۶۰۸	۰/۵۰۰	۰/۴۵۶	۱/۰۰۰
۵	n-پنتان	۰/۹۶۳	۰/۳۵۴	۰/۴۷۲	۰/۳۰۱
۶	n-هگزان	۰/۷۳۴	۰/۲۵۰	۰/۴۸۱	۱/۰۰۰
۷	n-هپتان	۰/۹۹۸	۰/۱۷۷	۰/۴۹۲	۰/۲۶۳
۸	n-اکتان	۰/۷۷۹	۰/۱۲۵	۰/۴۹۸	۱/۰۰۰
۹	n-نونان	۰/۹۸۸	۰/۰۸۸	۰/۵۰۵	۰/۲۴۰
۱۰	n-دکان	۰/۸۲۵	۰/۰۶۳	۰/۵۰۸	۱/۰۰۰
۱۱	n-آن‌دکان	۰/۹۹۹	۰/۰۴۴	۰/۵۱۳	۰/۲۲۴
۱۲	n-دودکان	۰/۸۴۷	۰/۰۳۱	۰/۵۱۶	۱/۰۰۰
۱۳	n-تری‌دکان	۰/۹۹۴	۰/۰۲۲	۰/۵۱۹	۰/۲۱۳
۱۴	n-تترادکان	۰/۸۷۱	۰/۰۱۶	۰/۵۲۱	۱/۰۰۰
۱۵	n-پنتادکان	۰/۹۹۹	۰/۰۱۱	۰/۵۲۳	۰/۲۰۴
۱۶	n-هگزادکان	۰/۸۸۴	۰/۰۰۸	۰/۵۲۵	۱/۰۰۰
۱۷	n-هپتادکان	۰/۹۹۷	۰/۰۰۶	۰/۵۲۷	۰/۱۹۷

۳-۱-۸- مدل سازی

پس از انتخاب بهترین توصیف کننده ها، می توان با انواع روش های مدل سازی که در فصل ۲ ذکر شد، مدل سازی را آغاز کرد. در این تحقیق از روش های خطی رگرسیون مؤلفه ای اصلی (PCR)، رگرسیون کمترین مربعات جزئی (PLS) و روش غیرخطی شبکه ای عصبی مصنوعی (ANN) برای مدل سازی استفاده شده است.

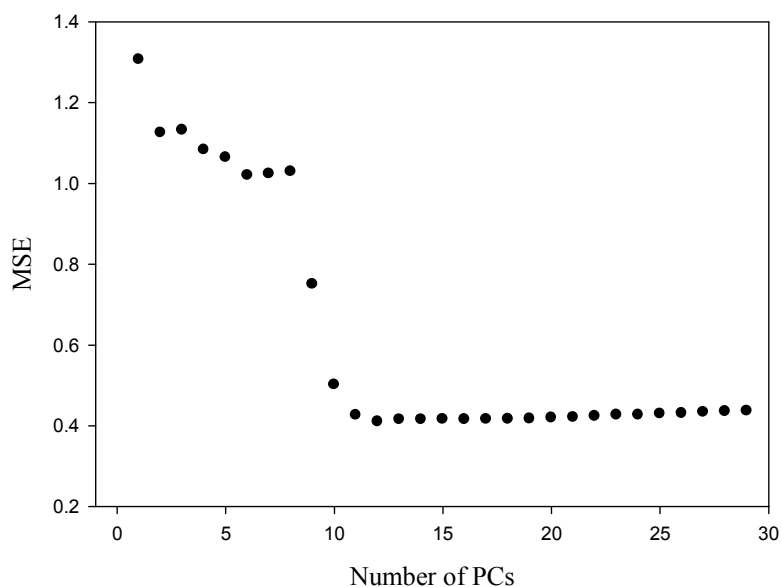
۳-۱-۸-۱- مدل سازی خطی با استفاده از روش PCR

در مدل سازی با استفاده از روش PCR در مرحله ای اول باید مؤلفه های اصلی از ترکیب خطی توصیف کننده ها ساخته شوند. جهت محاسبه ای مؤلفه های اصلی از توصیف کننده ها، چندین الگوریتم متفاوت وجود دارد از جمله الگوریتم نیپالس و الگوریتم تجزیه ای مقادیر منفرد^۱ SVD. به دلایل برتری های الگوریتم SVD نسبت به الگوریتم نیپالس، از جمله محاسبه ای دقیق تر مؤلفه های اصلی توسط الگوریتم SVD، در این تحقیق از الگوریتم SVD جهت محاسبه ای مؤلفه های اصلی استفاده شده است [۲۲].

لازم به تذکر است که در مدل سازی PCR و PLS کلیه توصیف کننده های باقیمانده از مرحله ای حذف توصیف کننده های اضافی، به مدل وارد می شوند و نیازی به اجرای روش های انتخاب بهترین توصیف کننده ها نیست. زیرا کاهش دادن تعداد توصیف کننده ها موجب از دست رفتن اطلاعات مفیدی می شود که در ساخت مؤلفه های اصلی تأثیرگذار هستند. به علاوه الگوریتم های PCR و PLS توانایی مدل سازی با اطلاعات ورودی زیاد را دارا هستند و به طور کلی این الگوریتم ها برای انجام محاسبات با ماتریس ورودی بزرگ طراحی شده اند.

1- Singular value decomposition

به منظور مدل‌سازی PCR، از برنامه‌ی نوشته شده توسط ریچارد برتون^۱ در محیط نرم افزار Matlab^۲ استفاده گردید. در ابتدا داده‌های مربوط به سری آموزش به برنامه داده شد و سپس به تعداد توصیف‌کننده‌ها، یعنی ۲۹ توصیف‌کننده، مؤلفه‌های اصلی ساخته شدند. برای دستیابی به تعداد مؤلفه‌های اصلی بهینه یک مدل‌سازی اولیه انجام پذیرفت. بدین ترتیب که ابتدا تنها با اولین مؤلفه‌ی اصلی مدل‌سازی انجام گرفت و مقادیر محاسباتی برای سری ارزیابی با استفاده از اولین مدل محاسبه شد. سپس مقادیر محاسبه شده با مقادیر حقیقی مقایسه شد و مقدار میانگین مربعات خطا محاسبه گردید. در مرحله‌ی بعد مدل‌سازی با مؤلفه‌های اصلی اول و دوم انجام شد و به همان روش قبل، مقدار میانگین مربعات خطای مدل دوم محاسبه شد. این عملیات تا ساخت مدلی با ۲۹ مؤلفه‌ی اصلی ادامه یافت و در نهایت نمودار مقدار میانگین مربعات خطا (MSE) بر حسب تعداد مؤلفه‌های اصلی برای سری ارزیابی رسم گردید. نقطه‌ای که در آن MSE کمترین مقدار را دارد نشان‌دهنده‌ی تعداد PC یا مؤلفه‌ی اصلی بهینه است. این نمودار در شکل (۳-۲) نشان داده شده‌است.



شکل (۳-۲): نمودار تغییرات مقدار میانگین مربعات خطا بر حسب تعداد مؤلفه‌های اصلی برای سری ارزیابی

1- Richard Breton

2- Matrix laboratory

همچنین مقدار میانگین مربعات خطای مربوط به هر مدل برای سری ارزیابی در جدول (۱۰-۳) به تفصیل گزارش شده است.

جدول (۱۰-۳): مقدار میانگین مربعات خطای سری ارزیابی برای مدل‌های اولیه PCR

مدل	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
MSE	۱/۳۰۶۵	۱/۱۲۵۳	۱/۱۳۱۵	۱/۰۸۲۷	۱/۰۶۳۸	۱/۰۱۹۹	۱/۰۲۳۵	۱/۰۲۸۸	۰/۷۵۰۱	۰/۵۰۱۳
مدل	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰
MSE	۰/۴۲۶۲	۰/۴۰۷۱	۰/۴۱۵۲	۰/۴۱۵۴	۰/۴۱۵۸	۰/۴۱۵۷	۰/۴۱۵۹	۰/۴۱۶۳	۰/۴۱۶۸	۰/۴۱۹۶
مدل	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹	
MSE	۰/۴۲۰۶	۰/۴۲۳۵	۰/۴۲۶۳	۰/۴۲۶۸	۰/۴۲۹۶	۰/۴۳۰۶	۰/۴۳۳۵	۰/۴۳۵۳	۰/۴۳۶۶	

با توجه به شکل (۲-۳) و جدول (۱۰-۳)، ۱۲ مؤلفه‌ی اصلی برای مدل‌سازی انتخاب گردید. در مرحله‌ی بعد برنامه با تعداد PC بهینه اجرا و مدل نهایی ساخته شد. ضریب هر مؤلفه‌ی اصلی در مدل نهایی طبق رابطه‌ی (۲۰-۲) و مقدار ویژه‌ی مربوط به هر مؤلفه‌ی اصلی نیز طبق رابطه‌ی (۲۱-۲) محاسبه گردید. همان‌طور که در فصل ۲ شرح داده شد مقدار ویژه‌ی یک مؤلفه‌ی اصلی با وزن آن مؤلفه‌ی اصلی در مدل PCR متناظر است. ضریب هر مؤلفه‌ی اصلی در مدل بهینه به همراه مقدار ویژه‌ی مربوط به آن در جدول (۱۱-۳) گزارش شده است.

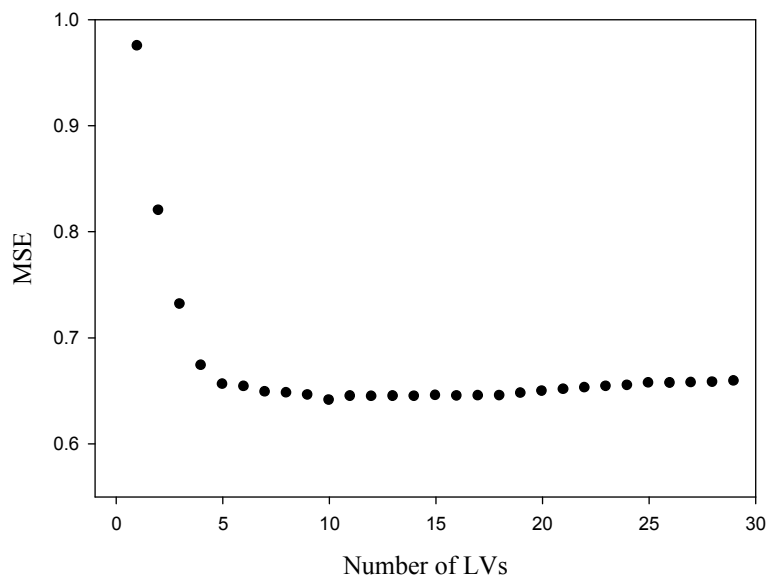
جدول (۳-۱۱): ضریب هر PC در مدل بهینه‌ی PCR و مقدار ویژه‌ی مربوط به آن

مقدار ویژه‌ی PC	ضریب PC در مدل	شماره‌ی PC
۴۶۵۱	۰/۰۶۲۱	۱
۴۰۰۹	۰/۱۴۶۲	۲
۳۱۵۲	۰/۰۲۵۹	۳
۲۲۷۶	-۰/۱۱۶۶	۴
۱۵۴۸	۰/۱۳۲۷	۵
۱۲۵۶	-۰/۱۷۹۸	۶
۸۷۹	-۰/۰۰۶۵	۷
۷۶۰	۰/۰۵۰۹	۸
۵۳۱	۰/۷۸۱۲	۹
۳۴۶	۰/۶۹۸۶	۱۰
۲۱۱	۰/۴۶۵۳	۱۱
۱۰۷	-۰/۰۵۱۸	۱۲

۳-۱-۸-۲- مدل سازی خطی با استفاده از روش PLS

جهت ساخت مدل به روش PLS، از برنامه‌ی نوشته شده توسط ریچارد برتون در محیط نرم افزار Matlab استفاده گردید. نخست داده‌های مربوط به سری آموزش به برنامه داده شد تا به تعداد توصیف‌کننده‌ها، متغیرهای پنهان (LV) ساخته شدند. جهت یافتن تعداد متغیرهای پنهان بهینه، یک مدل سازی اولیه به شرح زیر انجام پذیرفت. نخست تنها با اولین متغیر پنهان مدل سازی صورت گرفت و بعد مقادیر محاسباتی برای سری ارزیابی با استفاده از اولین مدل محاسبه شد. سپس مقادیر محاسبه شده با مقادیر حقیقی مقایسه شد و مقدار میانگین مربعات خطا به دست آمد. در مرحله‌ی بعد مدل سازی با متغیرهای پنهان اول و دوم انجام شد و به همان صورت، مقدار میانگین مربعات خطای مدل به دست آمده، محاسبه شد. عملیات مدل سازی تا ساخت مدلی با ۲۹ متغیر پنهان

ادامه یافت و در نهایت نمودار مقدار میانگین مربعات خطا (MSE) بر حسب تعداد متغیرهای پنهان مطابق با شکل (۳-۳) برای سری ارزیابی رسم گردید. نقطه‌ای که در آن MSE کمترین مقدار را دارد نشان‌دهنده تعداد متغیرهای پنهان بهینه است.



شکل (۳-۳) نمودار تغییرات مقدار میانگین مربعات خطا بر حسب تعداد متغیرهای پنهان مدل PLS

همچنین مقدار میانگین مربعات خطای مربوط به هر مدل برای سری ارزیابی در جدول (۱۲-۳) به تفصیل گزارش شده است. با توجه به شکل (۳-۳) و جدول (۱۲-۳) تعداد متغیرهای پنهان بهینه ۱۰ در نظر گرفته شد. در مرحله‌ی بعد برنامه با تعداد متغیرهای پنهان بهینه، یعنی ۱۰ متغیر پنهان، اجرا و مدل نهایی ساخته شد و ماتریس ضرایب طبق رابطه‌ی (۲-۲۵) به دست آمد. همان‌طور که در فصل ۲ تشریح شد در مدل‌سازی PLS، متغیرهای پنهان به وسیله‌ی مقدار واریانس‌شان وزن‌دار می‌شوند. ضریب هر متغیر پنهان در مدل بهینه به همراه درصد واریانس مربوط به آن در جدول (۱۳-۳) گزارش شده است.

جدول (۳-۱۲): مقدار میانگین مربعات خطای سری ارزیابی برای مدل‌های اولیه PLS

مدل	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
MSE	۰/۹۷۵۰	۰/۸۱۹۹	۰/۷۳۱۶	۰/۶۷۳۶	۰/۶۵۵۹	۰/۶۵۳۷	۰/۶۴۸۶	۰/۶۴۷۷	۰/۶۴۵۷	۰/۶۳۸۹
مدل	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰
MSE	۰/۶۴۴۷	۰/۶۴۴۴	۰/۶۴۴۷	۰/۶۴۴۶	۰/۶۴۵۳	۰/۶۴۴۸	۰/۶۴۴۹	۰/۶۴۵۲	۰/۶۴۷۴	۰/۶۴۹۴
مدل	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹	
MSE	۰/۶۵۱۱	۰/۶۵۲۶	۰/۶۵۴۰	۰/۶۵۴۸	۰/۶۵۷۱	۰/۶۵۷۰	۰/۶۵۷۴	۰/۶۵۷۸	۰/۶۵۸۹	

جدول (۳-۱۳): ضریب هر متغیر پنهان در مدل بهینه‌ی PLS و درصد واریانس مربوط به آن

شماره‌ی LV	ضریب LV در مدل	درصد واریانس
۱	۰/۲۹۷۲	۲۲/۱۹
۲	۰/۳۱۶۸	۸/۴۱
۳	۰/۱۴۹۱	۱۴/۶۸
۴	۰/۱۵۰۴	۱۵/۴۷
۵	۰/۰۹۸۸	۱۰/۴۱
۶	۰/۰۴۱۴	۱۱/۷۲
۷	۰/۰۵۳۸	۷/۴۳
۸	۰/۰۴۷۱	۱/۳۹
۹	۰/۱۲۳۶	۵/۱۷
۱۰	۰/۰۵۷۸	۳/۱۲

۳-۱-۸-۳- مدل سازی غیرخطی با استفاده از روش ANN

برای ساخت مدل غیرخطی با استفاده از روش شبکه‌ی عصبی مصنوعی، ابتدا توصیف‌کننده‌های انتخاب شده به روش‌های رگرسیون مرحله‌ای و الگوریتم ژنتیک، به عنوان ورودی به شبکه‌ی عصبی مصنوعی داده شدند. با توجه به این که شبکه، توصیف‌کننده‌ها را به صورت مرحله‌ای به مدل وارد می‌کند، لذا ترتیب چیدمان توصیف‌کننده‌ها در ماتریس ورودی شبکه، از اهمیت ویژه‌ای برخوردار است. در حقیقت توصیف‌کننده‌ها بایستی برحسب میزان اهمیت، در ماتریس ورودی مرتب شوند. یک روش مرتب سازی توصیف‌کننده‌ها، مرتب سازی برحسب میزان همبستگی با متغیر وابسته است که به آن رتبه‌بندی همبستگی^۱ (CR) گفته می‌شود. برای این منظور در نرم افزار SPSS از منوی Analyze و زیر منوی Correlate، گزینه‌ی Bivariate انتخاب شد. همه‌ی متغیرها اعم از مستقل و وابسته انتخاب شدند و میزان همبستگی هر توصیف‌کننده با خاصیت مورد نظر یا همان متغیر وابسته سنجیده شد و توصیف‌کننده‌ای که بیشترین مقدار همبستگی با متغیر وابسته را داشت، در ستون اول ماتریس ورودی شبکه قرار گرفت. سایر توصیف‌کننده‌ها نیز به همین ترتیب در ماتریس ورودی شبکه مرتب شدند [۴۹ و ۱۶].

پس از مرتب سازی توصیف‌کننده‌ها به روش رتبه‌بندی همبستگی، توصیف‌کننده‌ها به مدل وارد و عملیات مدل‌سازی آغاز گردید. در شبکه‌ی عصبی به ازای هر توصیف‌کننده یک نرون ورودی وجود دارد و لایه‌ی خروجی نیز شامل یک نرون است که نشان دهنده‌ی ضریب خودنفوذی متناظر با هر ورودی می‌باشد. اما تعداد نرون‌های لایه‌ی پنهان در ابتدا مشخص نیست و باید بهینه شود. همچنین تعداد ورودی‌ها، تعداد دوره‌های آموزش، نوع تابع انتقال و نوع تابع آموزش نیز جزء پارامترهای مؤثر بر آموزش شبکه‌ی عصبی هستند که باید بهینه گردند. معیار برای پایان دادن آموزش شبکه و بهینه شدن پارامترهای شبکه شامل مقادیر اوزان ارتباطی و بایاس‌ها، زمانی است که طبق

1- Correlation ranking

تابع کارایی تعیین شده، مقدار خطای شبکه کمترین مقدار شود. در این تحقیق تابع کارایی، میانگین مربعات خطا (MSE) می باشد.

در مدل سازی از دو تابع آموزشی لونیگ-مارکوارت (trainlm) و بایزین (trainbr) و دو تابع انتقال لگاریتم سیگموئیدی (logsig) و تانژانت سیگموئیدی (tansig) استفاده گردید. بنابراین در هر قسمت از ترکیب دو تابع آموزشی و دو تابع انتقال، چهار شبکه ی عصبی مصنوعی مختلف طراحی و هر کدام به طور جداگانه بهینه سازی شدند. در ادامه نتایج مدل سازی غیرخطی شبکه ی عصبی مصنوعی را به طور جداگانه با استفاده از توصیف کننده های منتخب به روش های رگرسیون مرحله ای، و الگوریتم ژنتیک شرح داده می شود.

۳-۱-۸-۳-۱- بهینه سازی مدل شبکه عصبی مصنوعی با توصیف کننده های

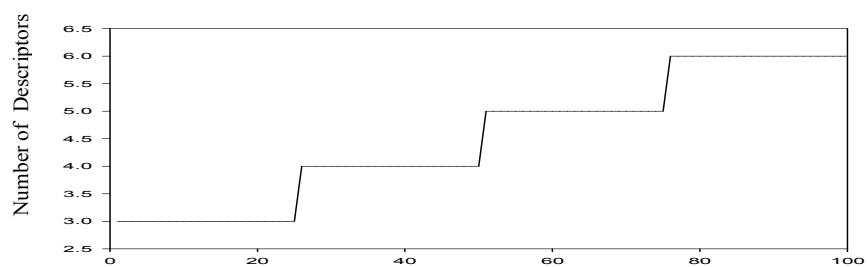
انتخابی به روش رگرسیون مرحله ای

توصیف کننده های انتخاب شده توسط روش رگرسیون مرحله ای که در جدول (۳-۳) نمایش داده شده اند، به روش رتبه بندی همبستگی، CR، به شبکه عصبی وارد شدند. در جدول (۳-۱۴) مقادیر همبستگی توصیف کننده های انتخاب شده توسط روش رگرسیون مرحله ای با متغیر وابسته گزارش شده است.

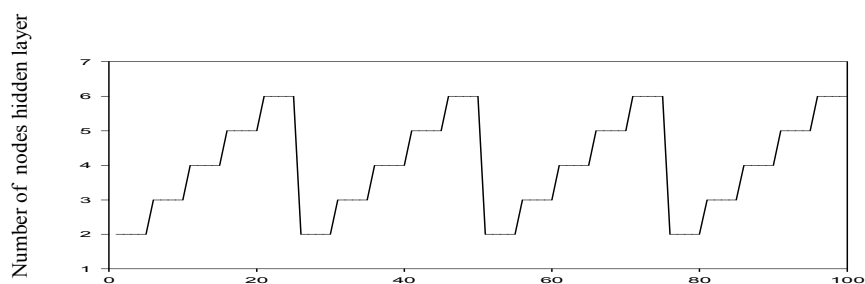
جدول (۳-۱۴): مقادیر ضریب همبستگی توصیف کننده های انتخاب شده به روش SR با ضریب خود نفوذی

	Mor12u	T	CIC5	P	HATS0m	R2u-A
D	۰/۵۳۶	۰/۳۸۰	-۰/۳۳۴	-۰/۲۶۵	۰/۲۵۶	۰/۱۴۱

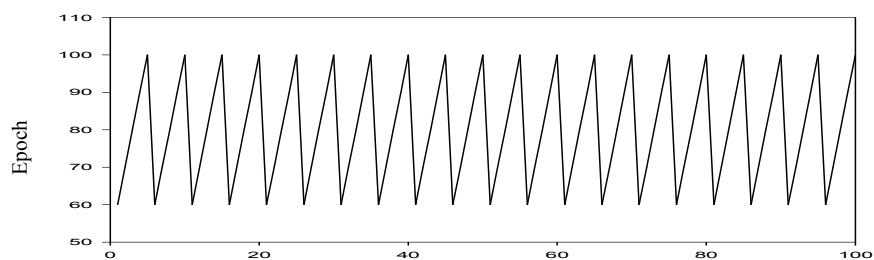
سپس به منظور بهینه سازی تعداد ورودی‌های شبکه، نوع تابع آموزش، نوع تابع انتقال و تعداد گره‌های لایه‌ی پنهان، از روش بهینه‌سازی همزمان استفاده گردید به این صورت که مقدار ممنتوم (μ) به صورت ثابت ۰/۰۰۱ در نظر گرفته شد. تعداد ورودی‌های شبکه از ۲ تا ۶، تعداد گره‌های لایه‌ی پنهان از ۲ تا ۱۰ گره و تعداد دور آموزشی شبکه از ۱۰ تا ۱۰۰ دور و با گام ۱۰ انتخاب شد و به ازای همه‌ی ترکیب‌های ممکن این سه پارامتر، چهار شبکه‌ی عصبی طراحی شده و بهینه گردیدند. با توجه به تأثیر تعداد دورهای آموزش بر بهبود آموزش شبکه، این پارامتر نیز همزمان با سایر پارامترها مورد بهینه سازی قرار گرفت. لازم به تذکر است که تابع انتقال خطی Pureline به عنوان تابع انتقال لایه‌ی خروجی انتخاب شد. در روند بهینه‌سازی در هر یک از شبکه‌های طراحی شده، به حداقل رساندن میانگین مربعات خطا مربوط به سری ارزیابی، به عنوان معیار انتخاب متغیرهای بهینه مورد استفاده قرار گرفت. قسمتی از روند تغییرات پارامترهای شبکه در حین بهینه سازی بر حسب یک بردار مرجع در شکل‌های (۳-۴) تا (۳-۷) نشان داده شده است. لازم به ذکر است که بردار مرجع، شمارنده‌ای است که تعداد ترکیب‌های ممکن از سه پارامتر تعداد توصیف‌کننده، تعداد نرون لایه‌ی پنهان و تعداد دور آموزش را نشان می‌دهد. مقادیر بهینه‌ی پارامترهای مربوط به هر شبکه نیز در جدول (۳-۱۵) خلاصه شده است.



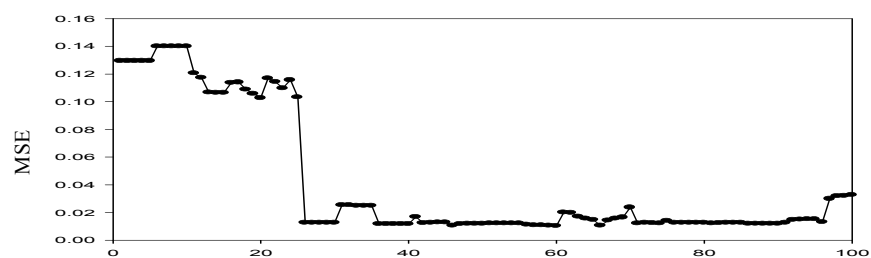
(الف)



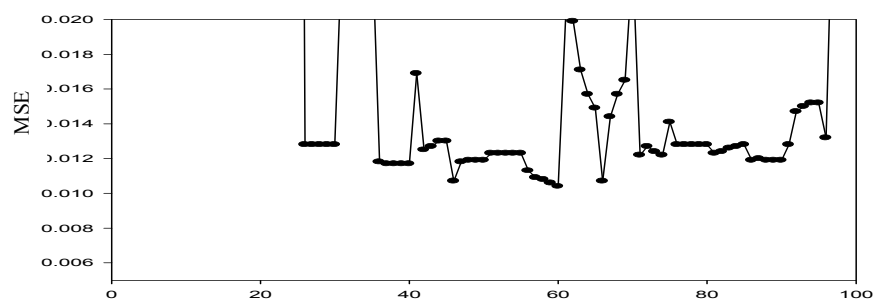
(ب)



(ج)



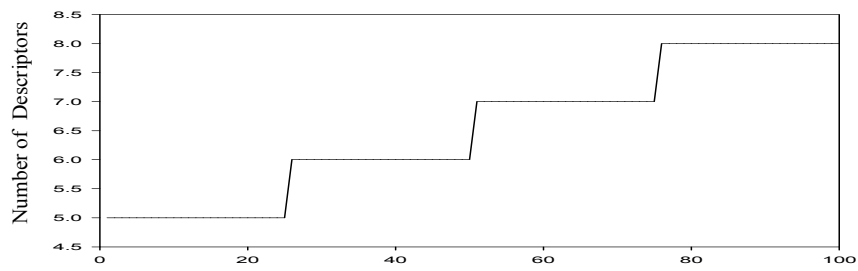
(د)



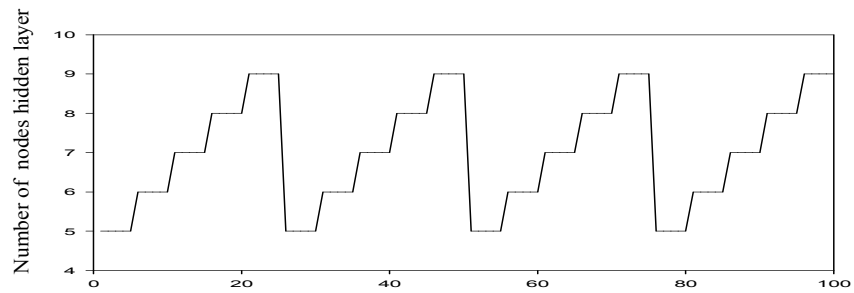
(ه)

بردار مرجع (شمارنده)

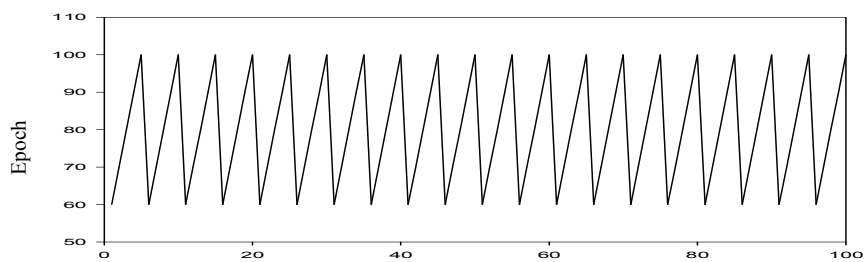
شکل (۳-۴): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می دهد، برای شبکه عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف کننده های منتخب روش SR



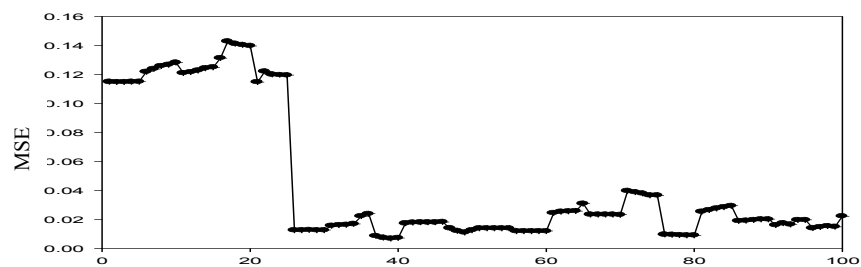
(الف)



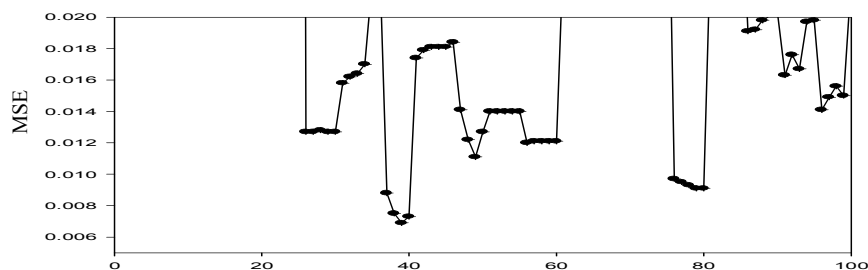
(ب)



(ج)



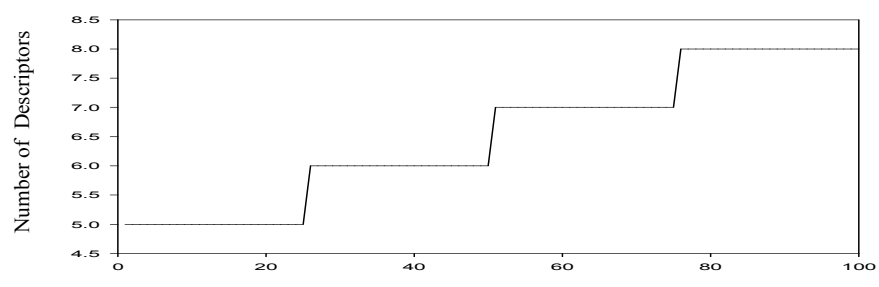
(د)



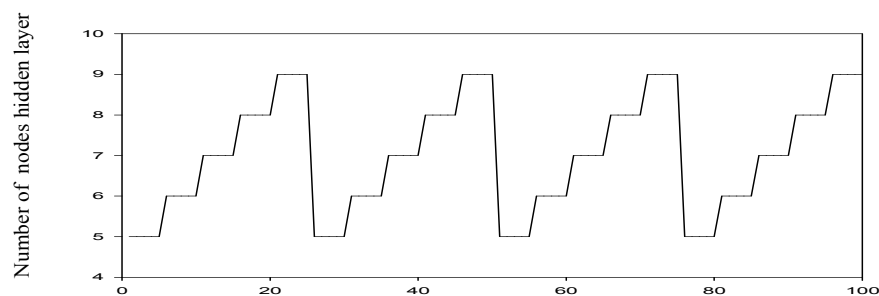
(ه)

بردار مرجع (شمارنده)

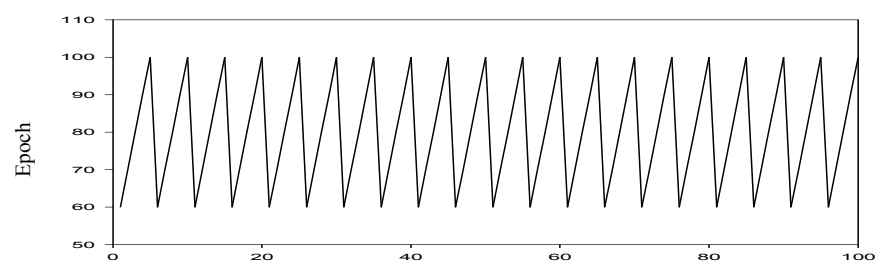
شکل (۳-۵): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می دهد، برای شبکه عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لوببرگ-مارکوارت با توصیف کننده های منتخب روش SR



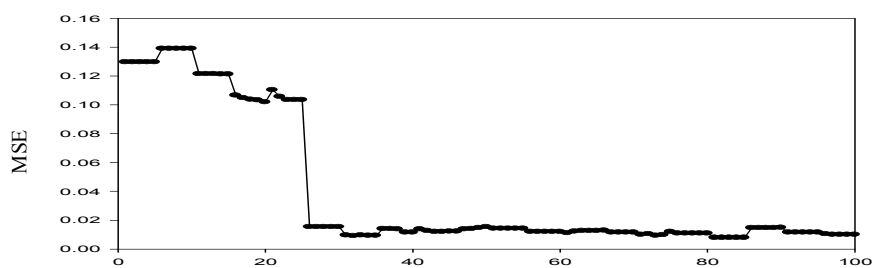
(الف)



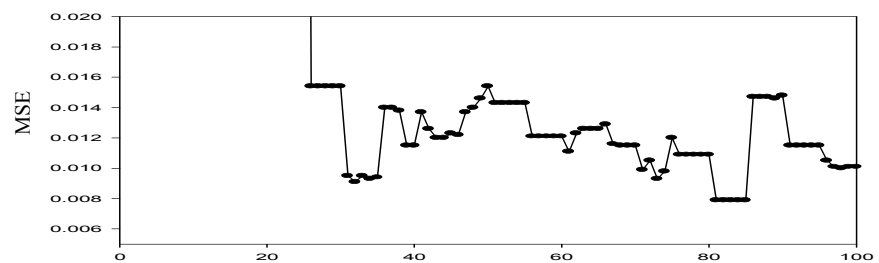
(ب)



(ج)



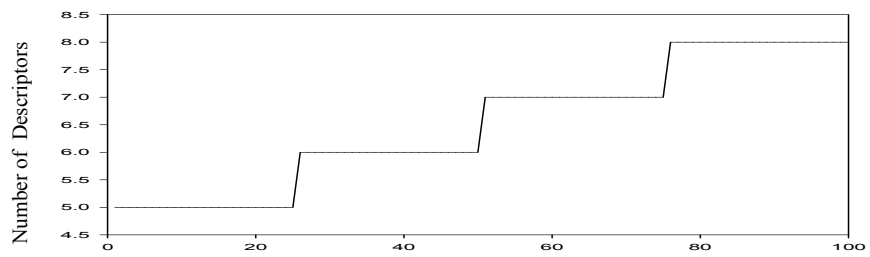
(د)



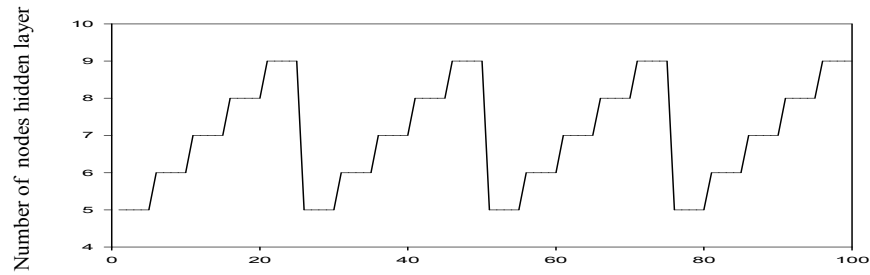
(ه)

بردار مرجع (شمارنده)

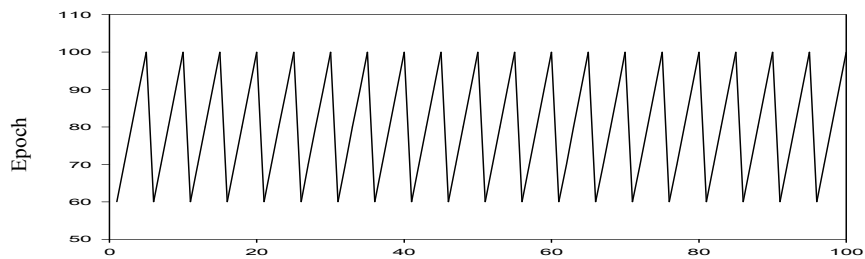
شکل (۳-۶): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف کننده‌های منتخب روش SR



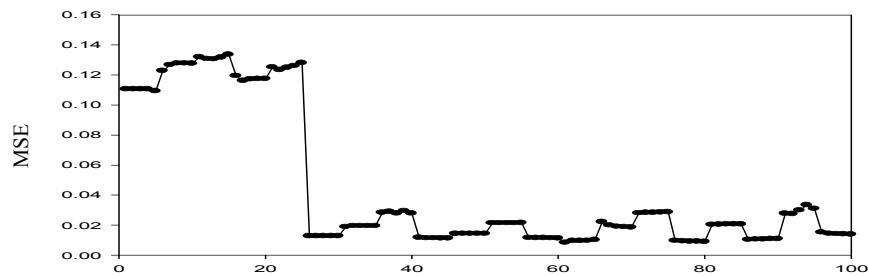
(الف)



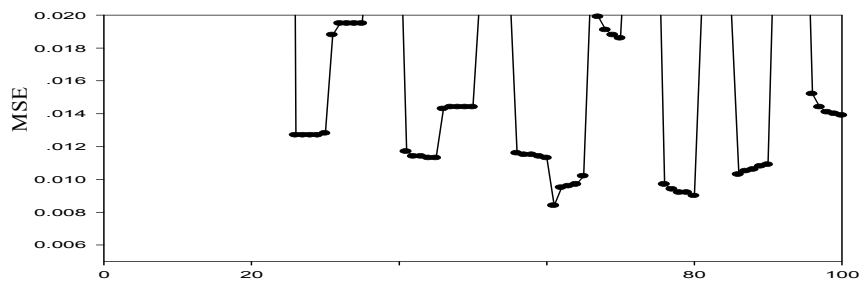
(ب)



(ج)



(د)



(ه)

بردار مرجع (شمارنده)

شکل (۷-۳): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف کننده‌های منتخب روش SR

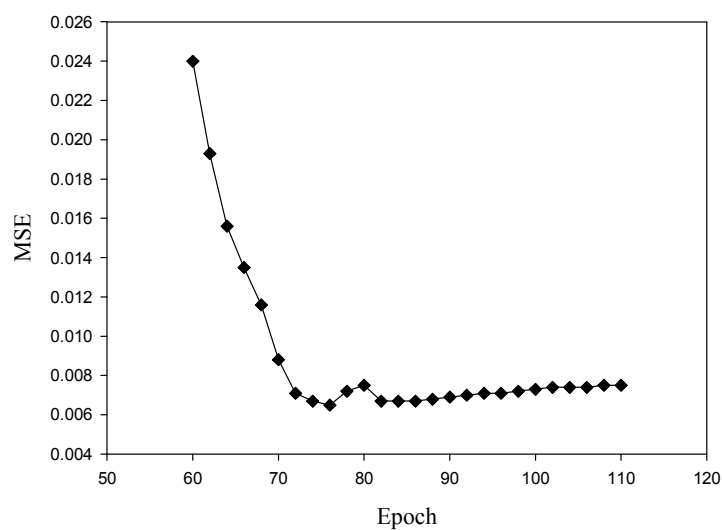
جدول (۳-۱۵): توابع و پارامترهای بهینه‌ی شبکه‌های طراحی شده با توصیف‌کننده‌های حاصل از روش SR

تابع انتقال	تابع آموزش	تعداد توصیف کننده	تعداد نرون لایه‌ی پنهان	تعداد دور آموزش	MSE
لگاریتم سیگموئید	تنظیم بایزین	۵	۷	۴۰	۰/۰۰۸۲
لگاریتم سیگموئید	لونبرگ-مارکوارت	۴	۴	۹۰	۰/۰۰۶۹
تانژانت سیگموئید	تنظیم بایزین	۶	۳	۵۰	۰/۰۰۷۹
تانژانت سیگموئید	لونبرگ-مارکوارت	۵	۴	۵۰	۰/۰۰۸۰

با توجه به نتایج، شبکه‌ی حاصل از الگوریتم آموزشی لونبرگ-مارکوارت و تابع انتقال لگاریتم سیگموئید با ۴ توصیف‌کننده در لایه‌ی ورودی، ۴ گره در لایه‌ی پنهان و ۹۰ دور آموزشی کمترین مقدار MSE برای سری ارزیابی را داراست. این شبکه به عنوان شبکه‌ی بهینه حاصل از توصیف‌کننده‌های انتخاب شده توسط رگرسیون مرحله‌ای انتخاب گردید. جهت محاسبه‌ی دقیق‌تر دوره‌های آموزش بهینه، شبکه‌ای با مقادیر بهینه‌ی تعداد نرون لایه‌ی ورودی، تعداد نرون لایه‌ی پنهان، تابع آموزش و تابع انتقال اما با تعداد دوره‌های آموزش از ۶۰ تا ۱۱۰ با گام ۲ مورد بررسی قرار گرفت. مقادیر میانگین مربعات خطا برای تعداد دوره‌های آموزش متفاوت در جدول (۳-۱۶) گزارش شده است. نمودار تغییرات مقادیر MSE برحسب تعداد دوره‌های آموزش نیز در شکل (۳-۸) نشان داده شده است. با توجه به این مقادیر دور آموزشی ۷۶ با کمترین مقدار MSE برای سری ارزیابی، به عنوان تعداد دور آموزش بهینه انتخاب گردید.

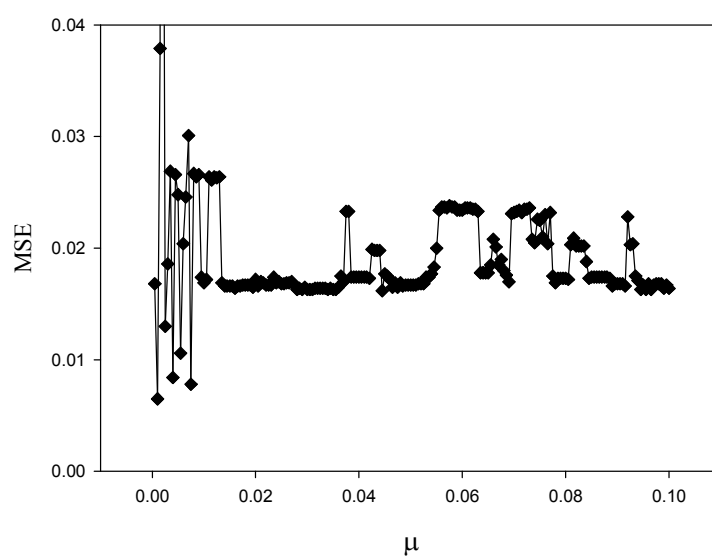
جدول (۳-۱۶): مقدار MSE سری ارزیابی برای تعداد دوره‌های آموزش متفاوت با توصیف‌کننده‌های منتخب روش SR

دور آموزش	۶۰	۶۲	۶۴	۶۶	۶۸	۷۰	۷۲	۷۴	۷۶
MSE	۰/۰۲۴۰	۰/۰۱۹۳	۰/۰۱۵۶	۰/۰۱۳۵	۰/۰۱۱۶	۰/۰۰۸۸	۰/۰۰۷۱	۰/۰۰۶۷	۰/۰۰۶۵
دور آموزش	۷۸	۸۰	۸۲	۸۴	۸۶	۸۸	۹۰	۹۲	۹۴
MSE	۰/۰۰۷۲	۰/۰۰۷۵	۰/۰۰۶۷	۰/۰۰۶۷	۰/۰۰۶۷	۰/۰۰۶۸	۰/۰۰۶۹	۰/۰۰۷۰	۰/۰۰۷۱



شکل (۳-۸): نمودار مقدار MSE سری ارزیابی بر حسب تعداد دوره‌های آموزش با توصیف‌کننده‌های منتخب روش SR

جهت بهینه‌سازی مقدار پارامتر ممتم در الگوریتم آموزشی لونبرگ-مارکوارت، شبکه‌ی بهینه با مقدار μ از ۰/۰۰۰۵ تا ۰/۱ با گام ۰/۰۰۰۵ اجرا گردید. با توجه به شکل (۳-۹) که میانگین مربعات خطا برای سری ارزیابی بر حسب تغییرات ممتم را نشان می‌دهد، مقدار ۰/۰۰۱ به عنوان مقدار بهینه‌ی ممتم انتخاب شد.



شکل (۳-۹): نمودار میانگین مربعات خطا حاصل از سری ارزیابی بر حسب پارامتر μ برای شبکه با توصیف‌کننده‌های منتخب روش SR

مقادیر پارامترهای بهینه‌ی شبکه‌ی عصبی ساخته شده با توصیف‌کننده‌های منتخب روش رگرسیون مرحله‌ای در جدول (۳-۱۷) گزارش شده است.

جدول (۳-۱۷) توابع و مقادیر پارامترهای بهینه‌ی شبکه‌ی عصبی با توصیف‌کننده‌های منتخب روش SR

تابع آموزشی	Train lm
تابع انتقال لایه‌ی پنهان	Logsig
تابع انتقال لایه‌ی خروجی	Pure line
تعداد نرون لایه‌ی پنهان	۴
تعداد متغیرهای ورودی	۴
تعداد دوره‌های آموزش	۷۶
پارامتر ممنت	۰/۰۰۱
مقدار MSE سری ارزیابی	۰/۰۰۶۵

۳-۱-۸-۳-۲- بهینه‌سازی مدل شبکه‌ی عصبی مصنوعی با توصیف‌کننده‌های

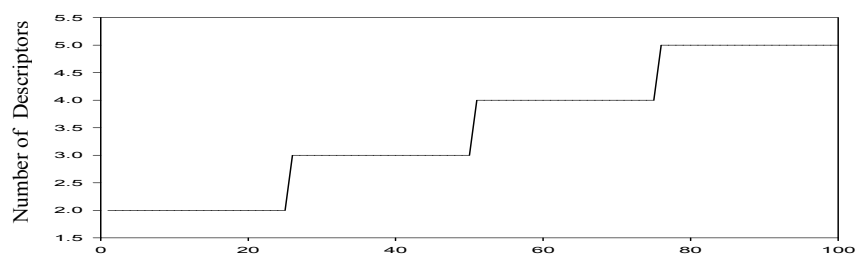
انتخابی به روش الگوریتم ژنتیک

برای ساخت مدل شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده‌های انتخاب شده توسط روش الگوریتم ژنتیک، همبستگی توصیف‌کننده‌ها با متغیر وابسته به ترتیبی که در قسمت ۳-۱-۸-۳ شرح داده شد، محاسبه گردید و توصیف‌کننده‌های معرفی شده در جدول (۳-۷)، به روش رتبه‌بندی همبستگی، CR، به شبکه‌ی عصبی وارد شدند. در جدول (۳-۱۸) مقادیر همبستگی توصیف‌کننده‌های انتخاب شده توسط روش الگوریتم ژنتیک با متغیر وابسته گزارش شده است.

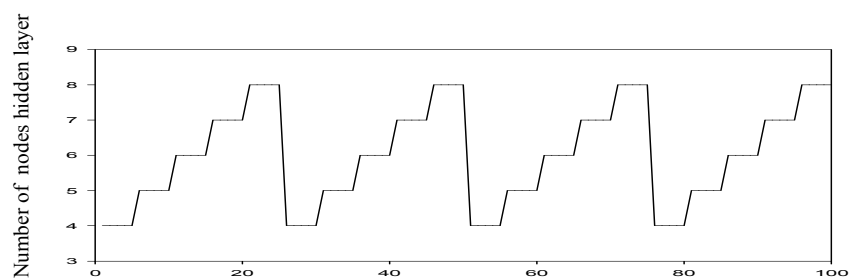
جدول (۳-۱۸): مقادیر ضریب همبستگی توصیف‌کننده‌های منتخب روش GA با ضریب خودنفوذی

	PJl3	Xt	T	P	E1p	G2s
D	-۰/۴۰۸	۰/۴۰۰	۰/۳۸۰	-۰/۲۵۶	-۰/۱۵۸	۰/۰۵۷

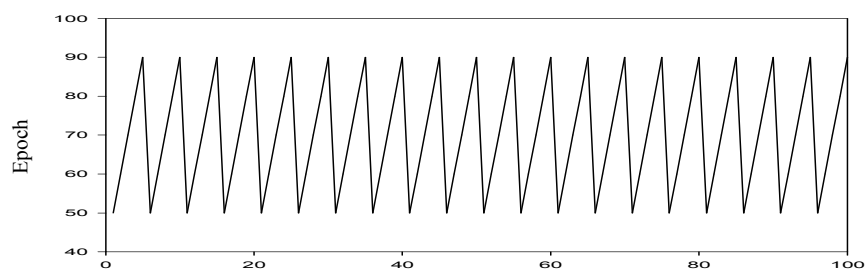
در این بخش نیز از شبکه‌ی عصبی سه لایه، شامل لایه‌ی ورودی، لایه‌ی میانی و لایه‌ی خروجی استفاده شد. برای طراحی یک شبکه‌ی بهینه، مانند قبل باید پارامترهای شبکه بهینه شوند. بر این اساس شبکه‌ای با مقدار μ ثابت ۰/۰۰۱ و تعداد متغیر ورودی از ۲ تا ۶ توصیف‌کننده، تعداد متفاوت گره در لایه‌ی پنهان از ۲ تا ۱۰ و تعداد دور آموزشی از ۱۰ تا ۱۰۰ با گام ابتدایی ۱۰ توسط دو الگوریتم آموزشی لونبرگ-مارکوارت و تنظیم بایزین و همچنین توابع انتقال لگاریتم سیگموئیدی و تانژانت سیگموئیدی در لایه‌ی میانی، آموزش داده شد. در تمامی مراحل بهینه‌سازی شبکه‌ها، حداقل رساندن تابع عملکرد MSE در سری ارزیابی مورد نظر بوده است. شبکه‌های طراحی شده به ازای همه ترکیب‌های ممکن از این سه پارامتر بهینه‌سازی گردید که نتایج این بررسی‌ها در شکل‌های (۳-۱۰) تا (۳-۱۳) نشان داده شده است.



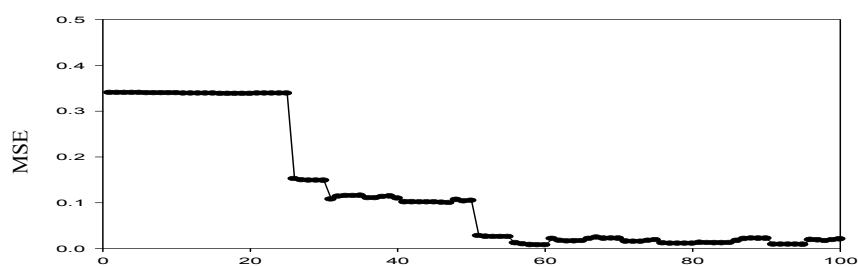
(الف)



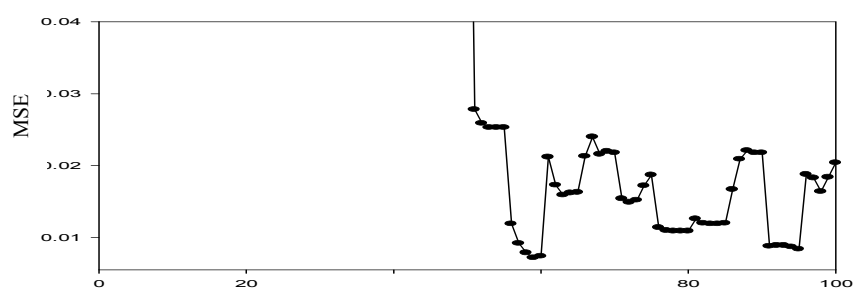
(ب)



(ج)



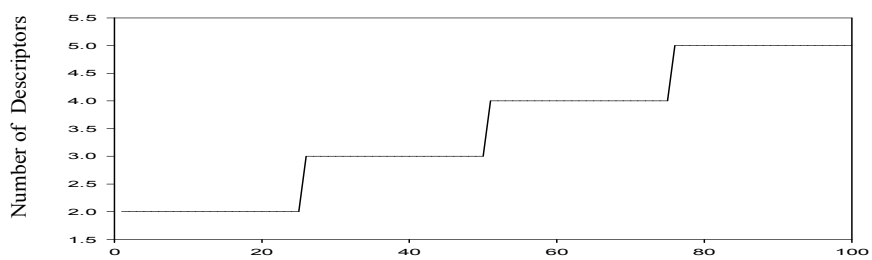
(د)



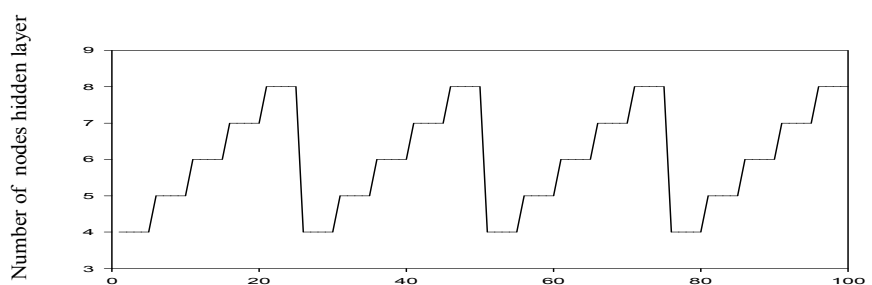
(ه)

بردار مرجع (شمارنده)

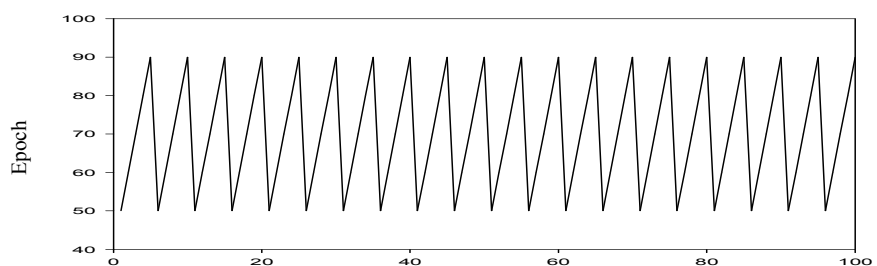
شکل (۳-۱۰): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف کننده‌های منتخب روش GA



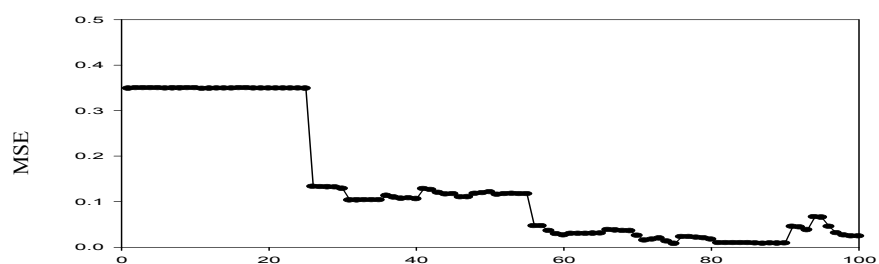
(الف)



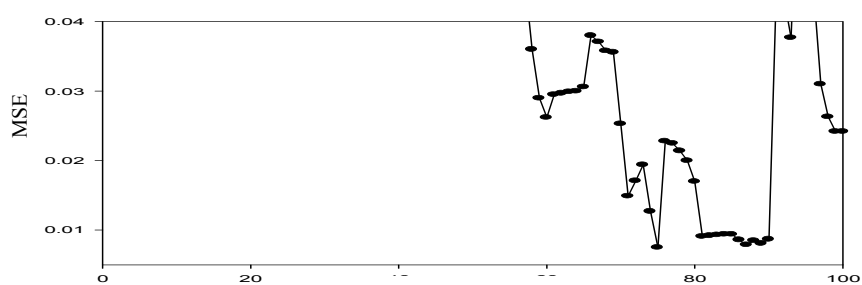
(ب)



(ج)



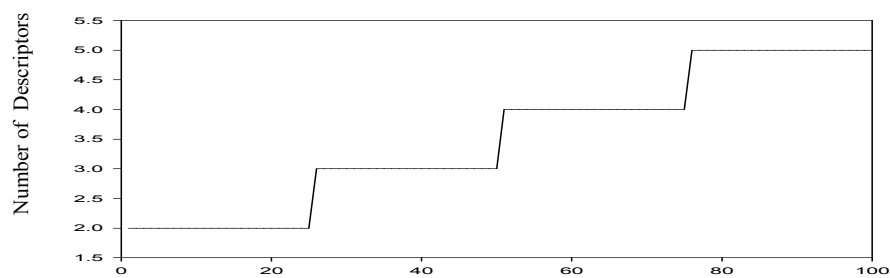
(د)



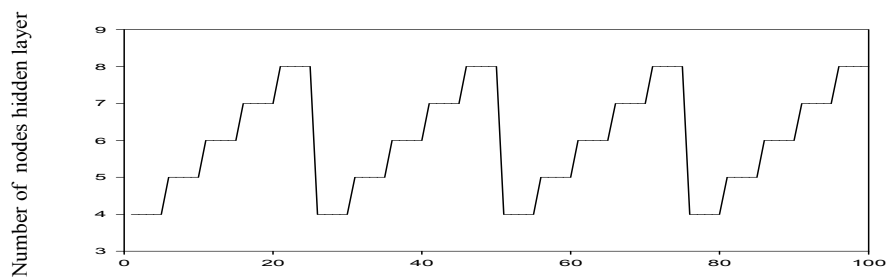
(ه)

بردار مرجع (شمارنده)

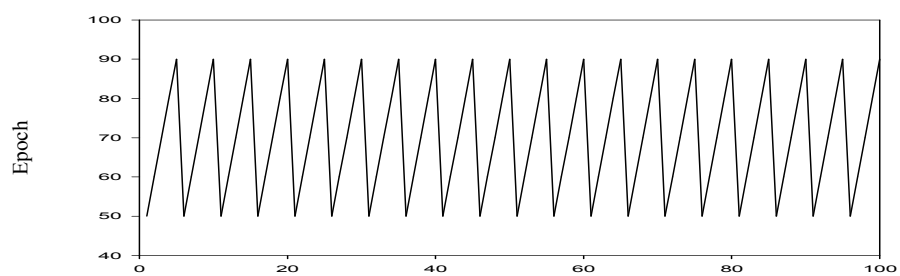
شکل (۳-۱۱): نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می دهد، برای شبکه عصبی با تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف کننده های منتخب روش GA



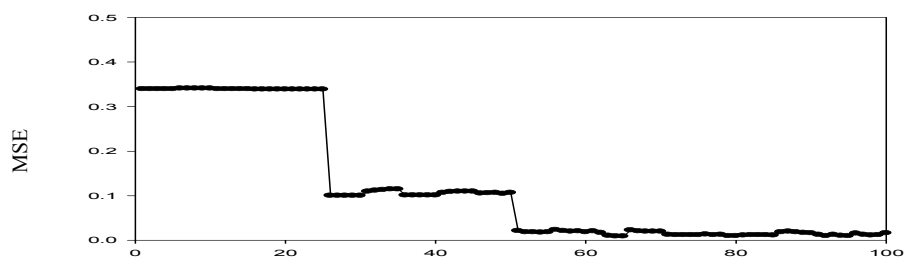
(الف)



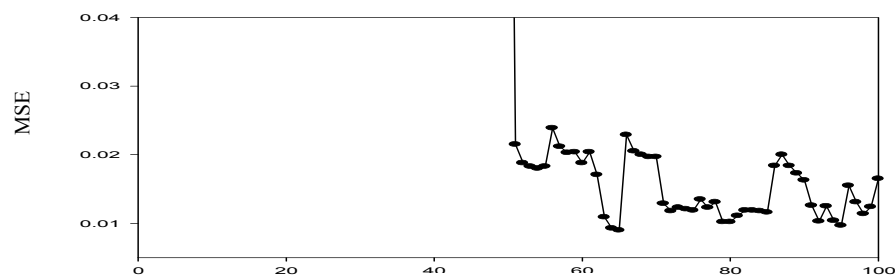
(ب)



(ج)



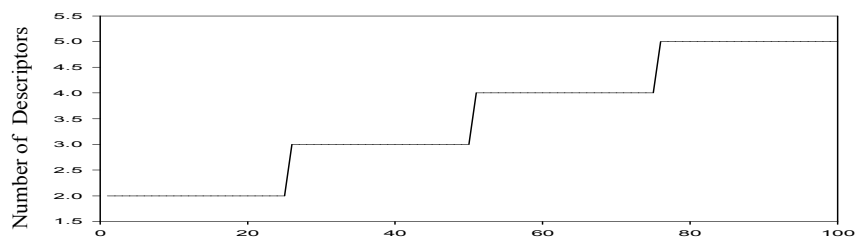
(د)



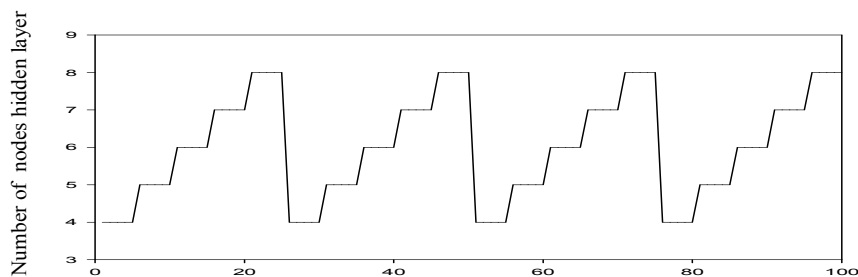
(ه)

بردار مرجع (شمارنده)

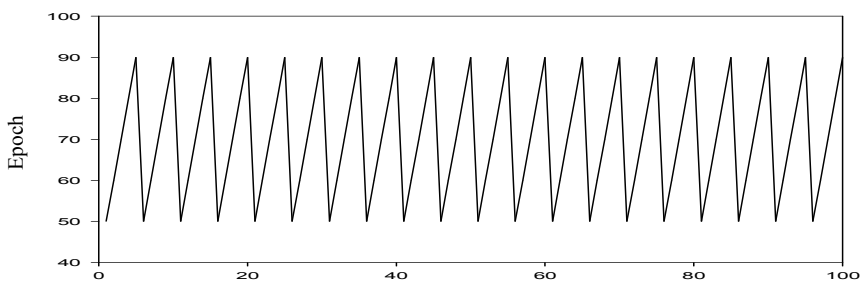
شکل (۳-۱۲): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین با توصیف‌کننده‌های منتخب روش GA



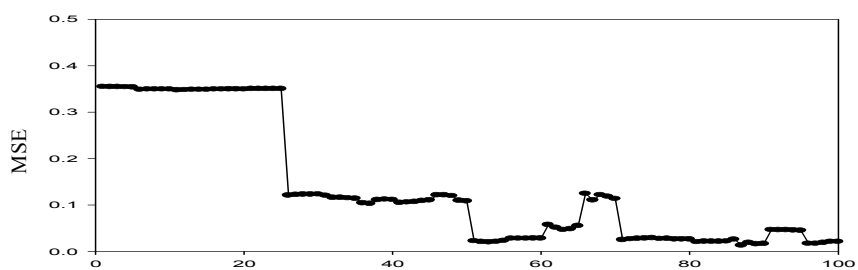
(الف)



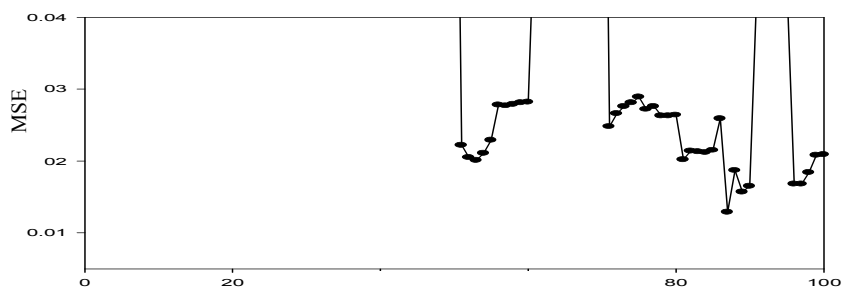
(ب)



(ج)



(د)



(ه)

بردار مرجع (شمارنده)

شکل (۳-۱۳): نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی پنهان، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بر حسب بردار مرجع بهتر نشان می‌دهد، برای شبکه عصبی با تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونبرگ-مارکوارت با توصیف‌کننده‌های منتخب روش GA

در جدول (۳-۱۹) توابع و پارامترهای بهینه‌ی شبکه‌های عصبی GA-ANN آورده شده است. نتایج موجود در این جدول نشان می‌دهد که شبکه‌ای با الگوریتم آموزشی تنظیم بایزین، تابع انتقال لگاریتم سیگموئیدی با ۴ توصیف‌کننده در لایه‌ی ورودی، ۵ گره در لایه‌ی پنهان و دور آموزشی بهینه ۸۰ کمترین مقدار MSE برابر با ۰/۰۰۷۲ را برای سری ارزیابی به دست می‌دهد، در نتیجه این شبکه به عنوان شبکه‌ی بهینه انتخاب گردید.

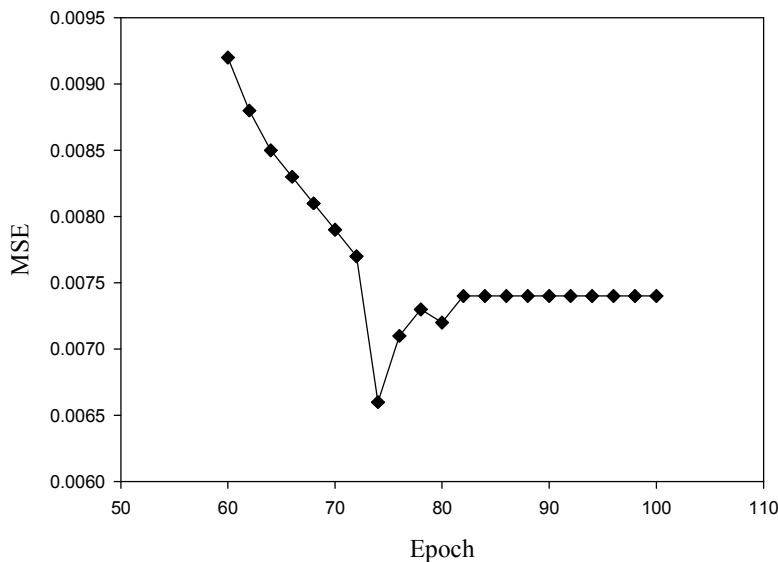
جدول (۳-۱۹): توابع و پارامترهای بهینه‌ی شبکه‌های طراحی شده با توصیف‌کننده‌های حاصل از الگوریتم ژنتیک

MSE	تعداد دور آموزش	تعداد نرون لایه‌ی پنهان	تعداد توصیف کننده	تابع آموزش	تابع انتقال
۰/۰۰۷۲	۸۰	۵	۴	تنظیم بایزین	لگاریتم سیگموئید
۰/۰۰۷۹	۶۰	۶	۵	لونبرگ-مارکوارت	لگاریتم سیگموئید
۰/۰۰۸۰	۱۰۰	۶	۴	تنظیم بایزین	تانژانت سیگموئید
۰/۰۰۸۸	۷۰	۹	۵	لونبرگ-مارکوارت	تانژانت سیگموئید

برای تعیین دقیق‌تر دور آموزش، شبکه‌ای با همان مقادیر بهینه‌ی تعداد نرون لایه‌ی ورودی، تعداد لایه‌ی پنهان، تابع آموزش و تابع انتقال اما با تعداد دورهای ۶۰ تا ۱۰۰ با گام ۲ مورد بررسی قرار گرفت. مقادیر میانگین مربعات خطا برای تعداد دورهای آموزش متفاوت در جدول (۳-۲۰) گزارش شده است. نمودار تغییرات مقادیر MSE برحسب تعداد دورهای آموزش نیز در شکل (۳-۱۴) نشان داده شده است. طبق این مقادیر دور آموزشی ۷۴ با مقدار MSE، ۰/۰۰۶۶ به عنوان بهترین مقدار برای تعداد دور آموزش در شبکه‌ی بهینه انتخاب گردید.

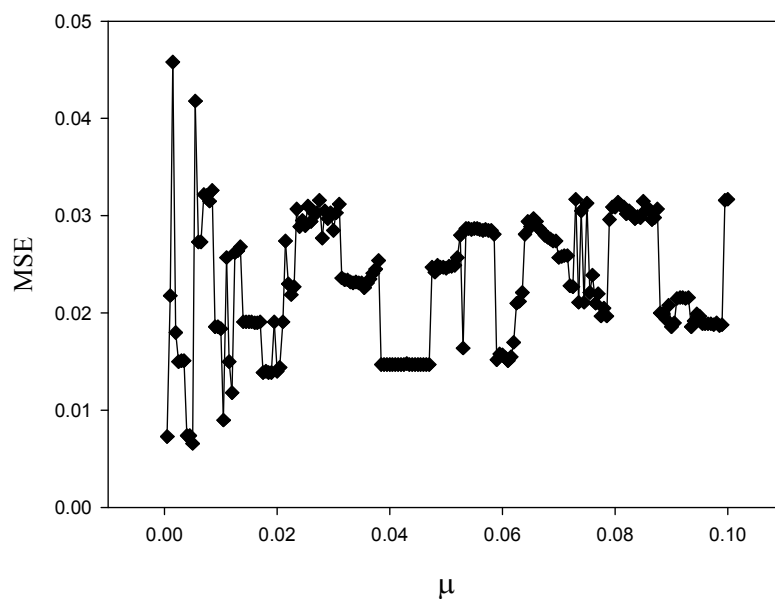
جدول (۳-۲۰): مقدار MSE سری ارزیابی برای تعداد دوره‌های آموزش متفاوت با توصیف‌کننده‌های منتخب روش GA

دور آموزش	۶۰	۶۲	۶۴	۶۶	۶۸	۷۰	۷۲	۷۴	۷۶
MSE	۰/۰۰۹۲	۰/۰۰۸۸	۰/۰۰۸۵	۰/۰۰۸۳	۰/۰۰۸۱	۰/۰۰۷۹	۰/۰۰۷۷	۰/۰۰۶۶	۰/۰۰۷۱
دور آموزش	۷۸	۸۰	۸۲	۸۴	۸۶	۸۸	۹۰	۹۲	۹۴
MSE	۰/۰۰۷۳	۰/۰۰۷۲	۰/۰۰۷۴	۰/۰۰۷۴	۰/۰۰۷۴	۰/۰۰۷۴	۰/۰۰۷۴	۰/۰۰۷۴	۰/۰۰۷۴



شکل (۳-۱۴): نمودار مقدار MSE سری ارزیابی بر حسب تعداد دوره‌های آموزش با توصیف‌کننده‌های منتخب روش GA

در ادامه‌ی بهینه‌سازی شبکه‌ی GA-ANN، مقدار پارامتر μ نیز بهینه‌شد؛ برای این منظور در ساختار شبکه‌ی بهینه با ۴ متغیر ورودی، ۵ گره در لایه‌ی پنهان، الگوریتم آموزشی تنظیم بایزن، تابع انتقال لگاریتم سیگموئیدی و دور آموزشی ۷۴، مقدار μ از ۰/۰۰۰۵ تا ۰/۱ با گام‌های ۰/۰۰۰۵ تغییر داده شد و آنگاه برای هر مورد، مقدار میانگین مربعات خطا برای سری ارزیابی محاسبه گردید که مقدار ۰/۰۰۵ برای پارامتر μ با مقدار MSE برابر با ۰/۰۰۶۶ به عنوان مقدار بهینه‌ی μ انتخاب شد. نمودار مقدار میانگین مربعات خطا بر حسب μ حول نقطه‌ی بهینه در شکل (۳-۱۵) رسم شده است.



شکل (۳-۱۵): نمودار میانگین مربعات خطای سری ارزیابی بر حسب پارامتر μ برای شبکه با توصیف‌کننده‌های منتخب روش GA

توابع و مقدار پارامترهای بهینه شده‌ی شبکه‌ی GA-ANN برای پیش‌بینی ضریب خودنفوذی آلکان‌های خطی، در جدول (۳-۲۱) ارائه شده است.

جدول (۳-۲۱): توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی با توصیف‌کننده‌های منتخب روش GA

تابع آموزشی	Train br
تابع انتقال لایه‌ی پنهان	Logsig
تابع انتقال لایه‌ی خروجی	Pure line
تعداد نرون لایه‌ی پنهان	۵
تعداد متغیرهای ورودی	۴
تعداد دوره‌های آموزش	۷۴
پارامتر ممتنم	۰/۰۰۵
مقدار MSE سری ارزیابی	۰/۰۰۶۶

۳-۱-۹- ارزیابی مدل‌های خطی و غیر خطی بهینه شده

در این مرحله مدل‌های بهینه شده حاصل از روش‌های خطی PCR و PLS و مدل‌های بهینه شده حاصل از روش غیرخطی شبکه‌ی عصبی مصنوعی مورد ارزیابی قرار گرفتند. برای این منظور از پارامترهای آماری، نمودار برگشتی، نمودار خطای باقیمانده و اعتبارسنجی بیرونی برای ارزیابی مدل‌های خطی و از پارامترهای آماری، نمودار برگشتی، نمودار خطای باقیمانده، اعتبارسنجی بیرونی و آزمون Y-تصافی برای ارزیابی مدل‌های غیرخطی استفاده گردید که در ادامه در ارتباط با هر یک توضیحاتی ارائه خواهد شد.

۳-۱-۹-۱- ارزیابی مدل‌های خطی و غیر خطی با استفاده از پارامترهای آماری

همان‌طور که گفته شد، پارامترهای آماری مقادیر تعیین کننده‌ای جهت ارزیابی مدل‌ها هستند. لذا در مرحله‌ی اول از این مقادیر برای ارزیابی مدل‌های ساخته شده بهره گرفته شد. شایان ذکر است مقادیر به دست آمده توسط مدل‌ها که به صورت خودمقیاس محاسبه شده بود، به مقادیر واقعی تبدیل شده و از مقادیر واقعی در جداول و رسم نمودارهای مربوطه استفاده شده است.

در جدول (۳-۲۲) پارامترهای آماری مربوط به سری‌های ارزیابی و تست درونی برای ۴ مدل بهینه شده به طور مختصر آورده شده است. با مقایسه مقادیر R^2 می‌توان نتیجه گرفت که مقادیر پیش‌بینی شده‌ی ضریب خودنفوذی حاصل از روش غیرخطی شبکه‌ی عصبی مصنوعی در مقایسه با مقادیر حاصل از روش‌های خطی PCR و PLS به مقادیر آزمایشگاهی ضریب خودنفوذی نزدیک‌ترند. همچنین با توجه به پارامترهای آماری مشاهده می‌شود که مقادیر ضریب خودنفوذی محاسبه شده با استفاده از روش‌های مدل‌سازی خطی در مقایسه با مقادیر محاسبه شده با روش‌های مدل‌سازی غیر خطی با خطای بیشتری همراه هستند. به علاوه می‌توان دریافت که از میان دو روش خطی مدل‌سازی،

روش PLS با خطای کمتری همراه است. همچنین از مقایسه دو روش انتخاب توصیف‌کننده‌ی رگرسیون مرحله‌ای و الگوریتم ژنتیک، توصیف‌کننده‌های انتخاب شده توسط روش غیرخطی الگوریتم ژنتیک برای پیش‌بینی ضریب خودنفوذی بهتر عمل نموده‌اند.

جدول (۳-۲۲): پارامترهای آماری مربوط به سری‌های ارزیابی و تست درونی

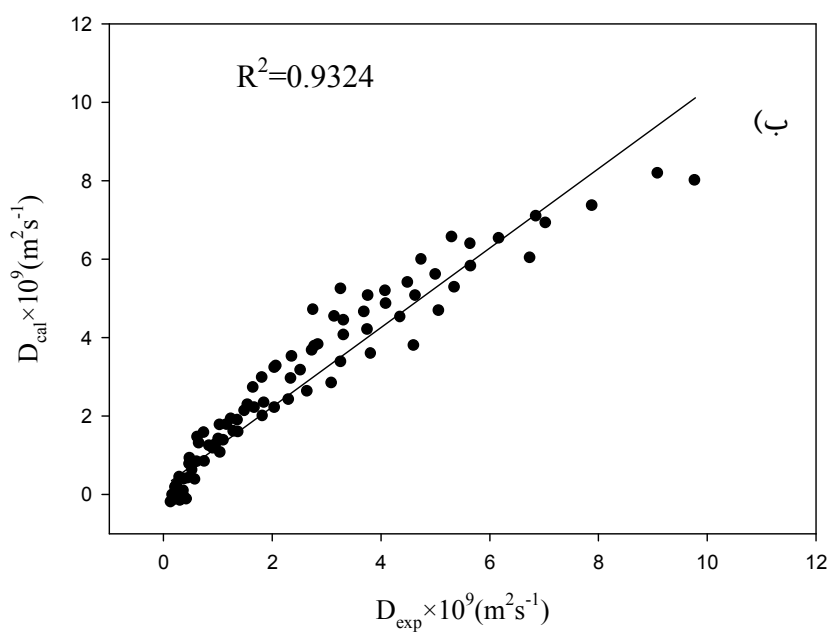
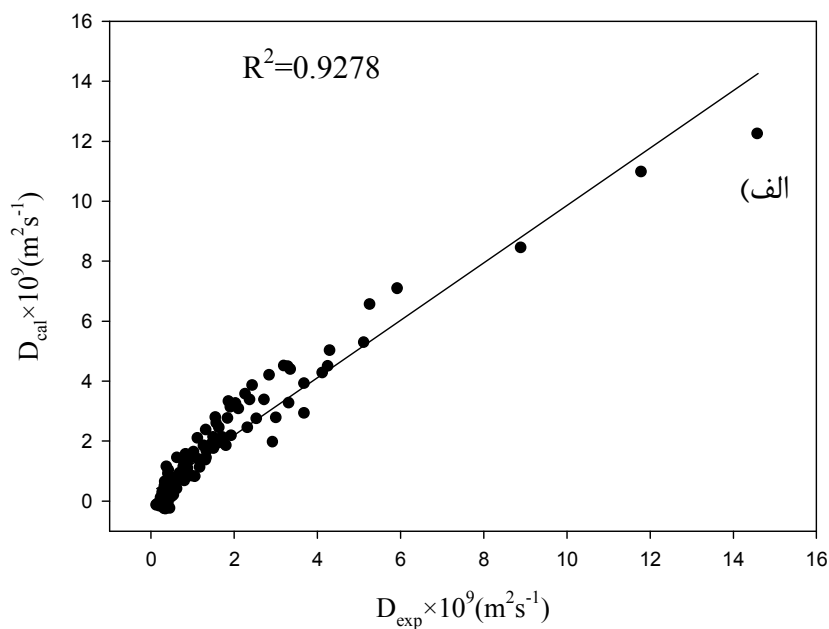
PCR		PLS		SR-ANN		GA-ANN		
ارزیابی	تست	ارزیابی	تست	ارزیابی	تست	ارزیابی	تست	
۰/۴۰۰۴	۰/۴۵۳۱	۰/۳۹۳۷	۰/۳۸۹۱	۰/۰۲۶۸	۰/۰۲۶۲	۰/۰۰۹۴	۰/۰۱۱۸	MSE
۰/۹۲۷۸	۰/۹۳۲۴	۰/۹۴۰۱	۰/۹۳۶۱	۰/۹۹۶۴	۰/۹۹۶۰	۰/۹۹۸۰	۰/۹۹۷۹	R ²
۲/۰۸۱۷	۲/۳۸۳۱	۲/۰۰۱۲	۲/۱۲۰۶	۱/۱۸۰۷	۱/۴۱۸۶	۰/۶۶۶۰	۰/۸۶۳۹	D _{Max}
۴/۷۱۲۹	۵/۱۱۰۱	۴/۵۱۱۲	۴/۷۸۴۹	۱/۷۷۱۱	۲/۱۹۳۶	۰/۹۷۴۹	۱/۲۲۷۲	AAD
۱/۲۱۹۱	۱/۲۷۶۹	۱/۱۰۲۵	۱/۲۱۲۹	-۱/۵۹۲۱	-۱/۷۶۵۹	-۰/۰۸۶۱	۰/۸۸۴۹	bias
۴۲/۸۳۱۲	۴۱/۹۸۰۶	۳۹/۶۸۲۵	۳۹/۰۶۱۸	۲/۷۸۷۱	۲/۶۰۰۴	۰/۹۷۸۲	۱/۱۷۰۲	PRESS
۰/۴۸۵	۰/۵۰۱	۰/۴۱۷	۰/۴۷۶	۰/۱۲۰۴	۰/۱۳۲۴	۰/۰۷۴۰	۰/۰۸۳۴	MAE

۳-۱-۹-۲- ارزیابی مدل‌های خطی و غیر خطی با استفاده از نمودار برگشتی

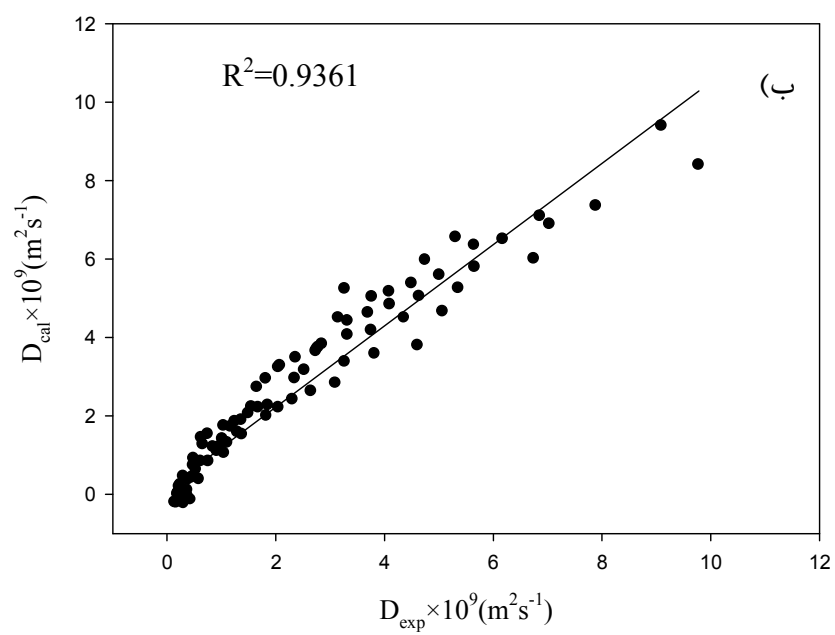
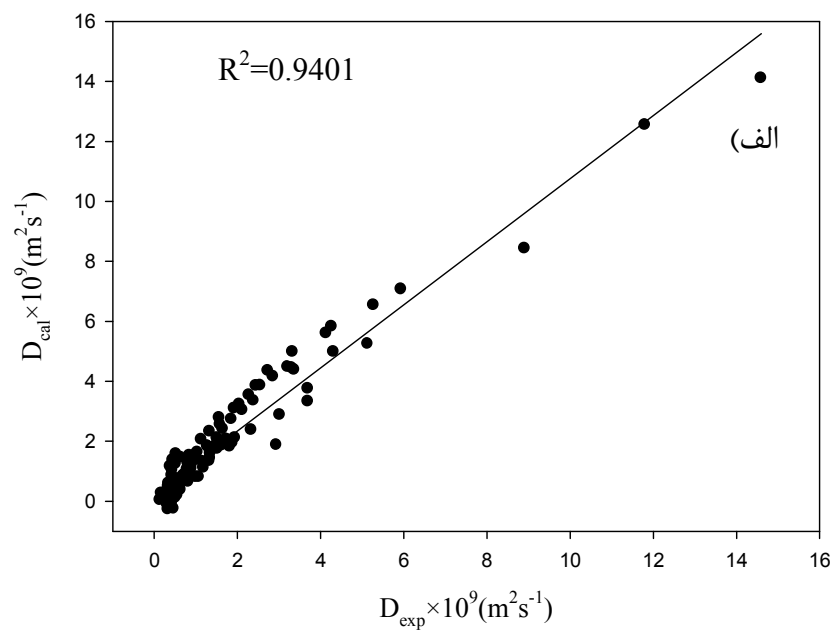
در نمودار برگشتی مقادیر پیش‌بینی شده بر حسب مقادیر تجربی رسم می‌گردد و با توجه به مقدار ضریب تعیین (R^2) به دست آمده از نمودار، پراکندگی نقاط در اطراف خط برگشت تعیین می‌شود. تغییرات مقدار R^2 بین ۱ و ۰ است. هر چه مقدار R^2 به دست آمده به یک نزدیک‌تر باشد به این معناست که همبستگی بیشتری بین داده‌ها وجود داشته و نتایج پیش‌بینی شده به واقعیت نزدیک‌تر می‌باشند. نتایج حاصل از نمودارهای برگشتی مربوط به چهار روش PCR، PLS، SR-ANN و GA-ANN به همراه مقادیر R^2 هریک از آن‌ها برای سری ارزیابی و سری تست در شکل‌های

(۱۶-۳)، (۱۷-۳)، (۱۸-۳) و (۱۹-۳) نشان داده شده است. مقدار R^2 گزارش شده‌ی نمودارهای

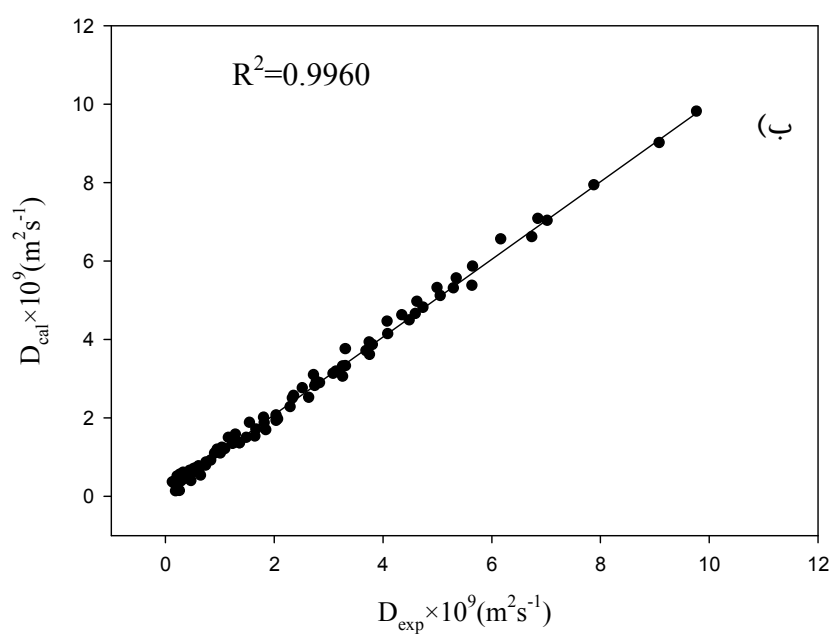
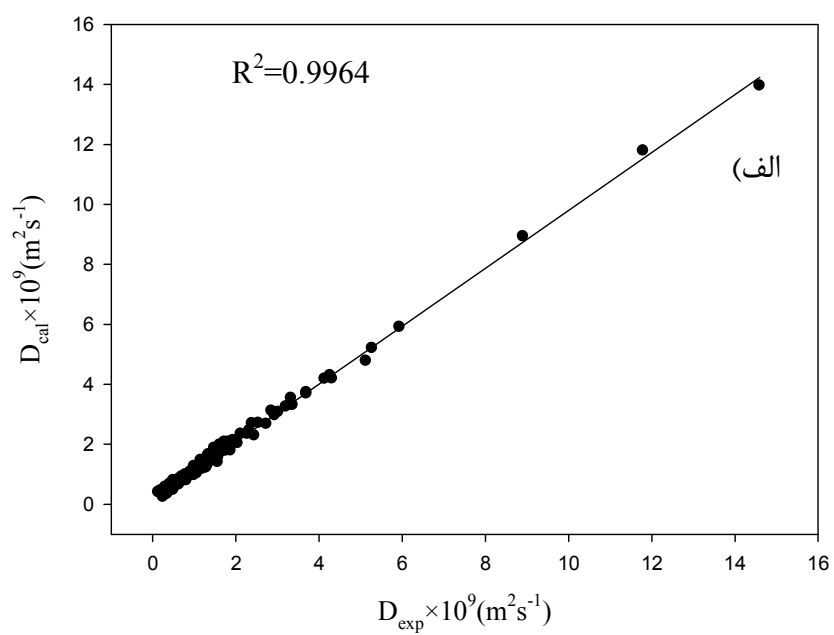
برگشتی، معیاری برای مقایسه‌ی میزان توانایی مدل‌ها در پیش‌بینی مقادیر ضریب خودنفوذی است.



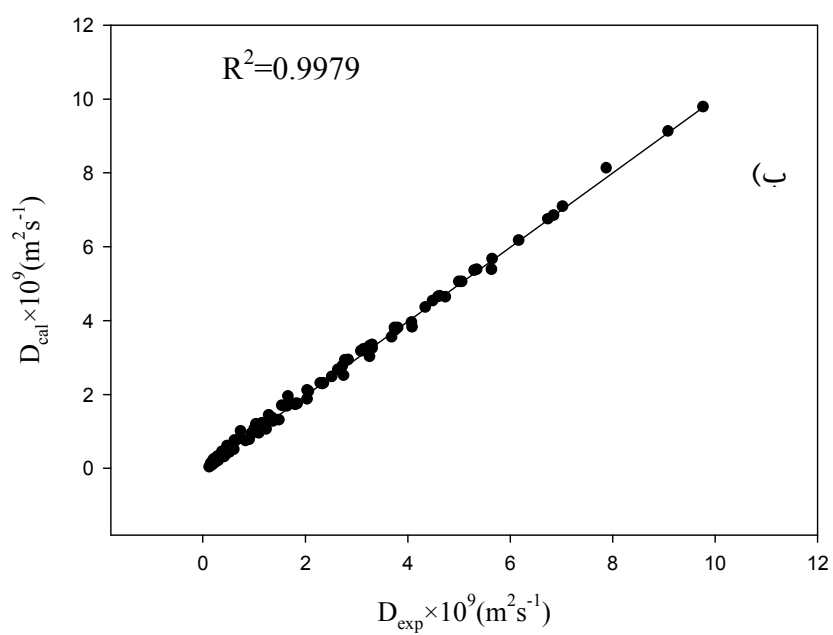
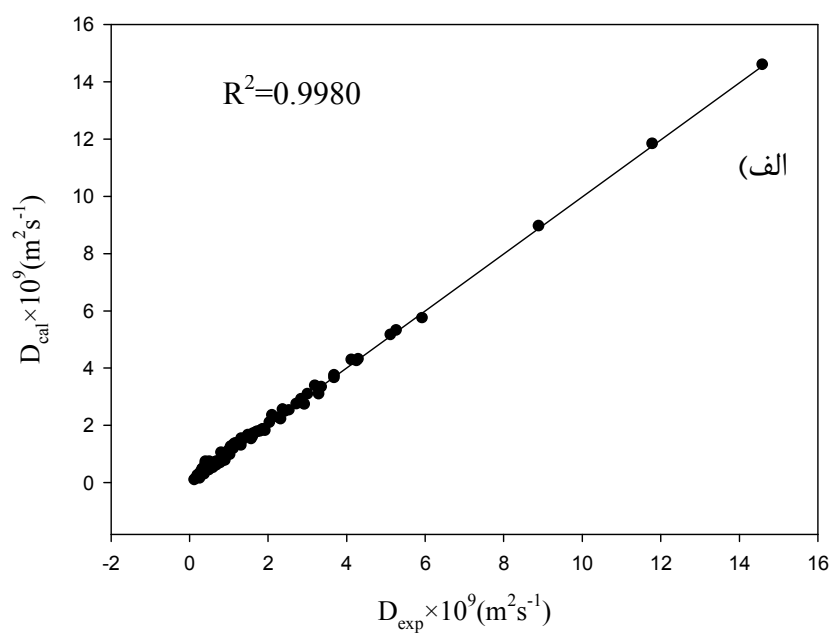
شکل (۱۶-۳): نمودار برگشتی حاصل از روش PCR برای (الف) سری ارزیابی و (ب) سری تست درونی



شکل (۳-۱۷): نمودار برگشتی حاصل از روش PLS برای (الف) سری ارزیابی و (ب) سری تست درونی



شکل (۳-۱۸): نمودار برگشتی حاصل از روش SR-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی



شکل (۳-۱۹): نمودار برگشتی حاصل از روش GA-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی

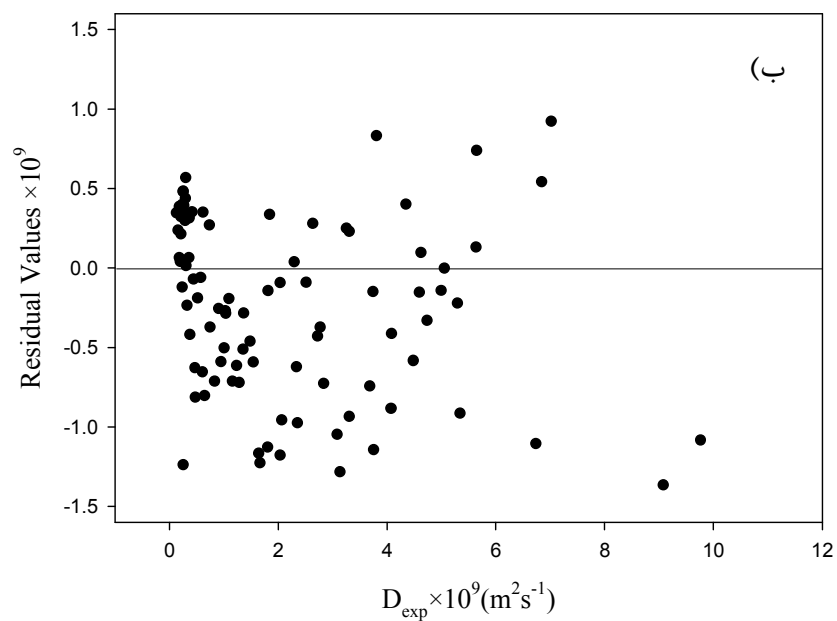
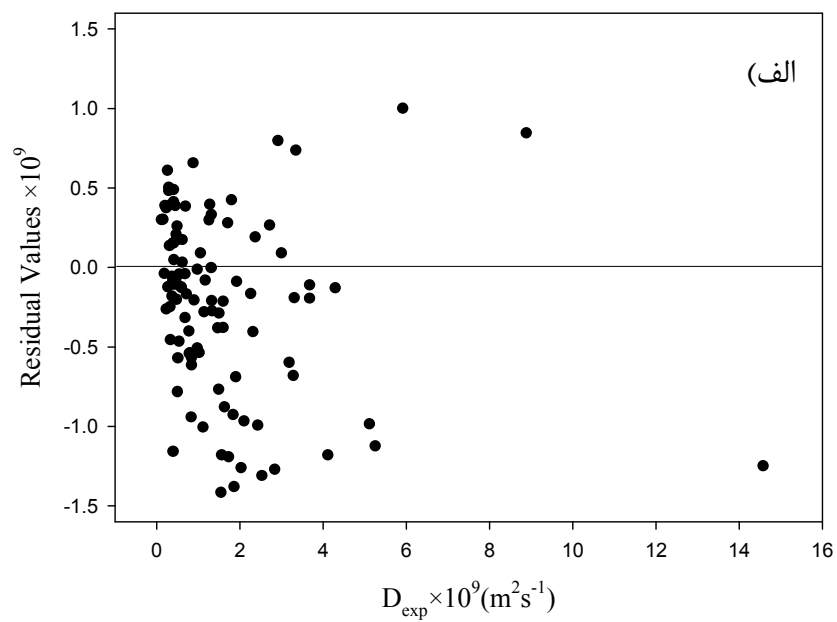
۳-۹-۱-۳- ارزیابی مدل‌های خطی و غیرخطی با استفاده از نمودار خطای

باقیمانده

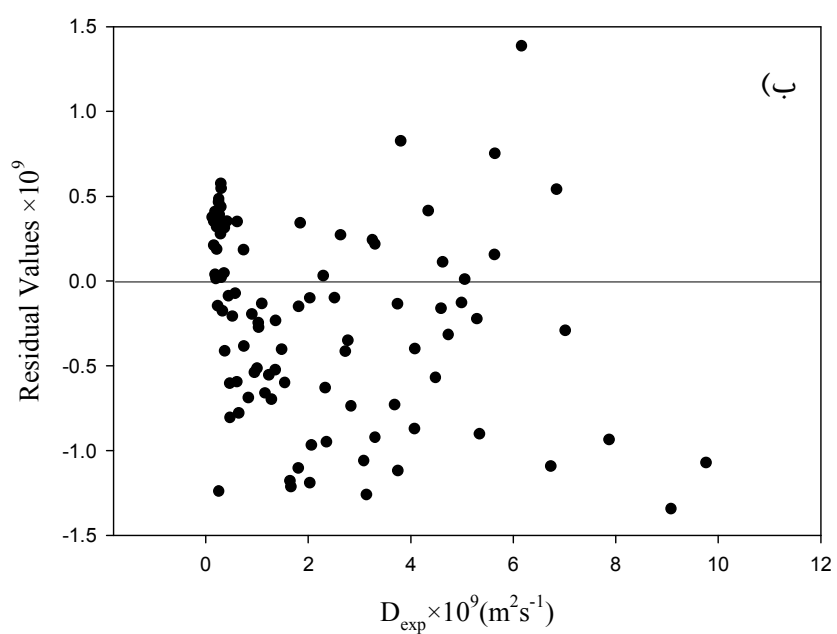
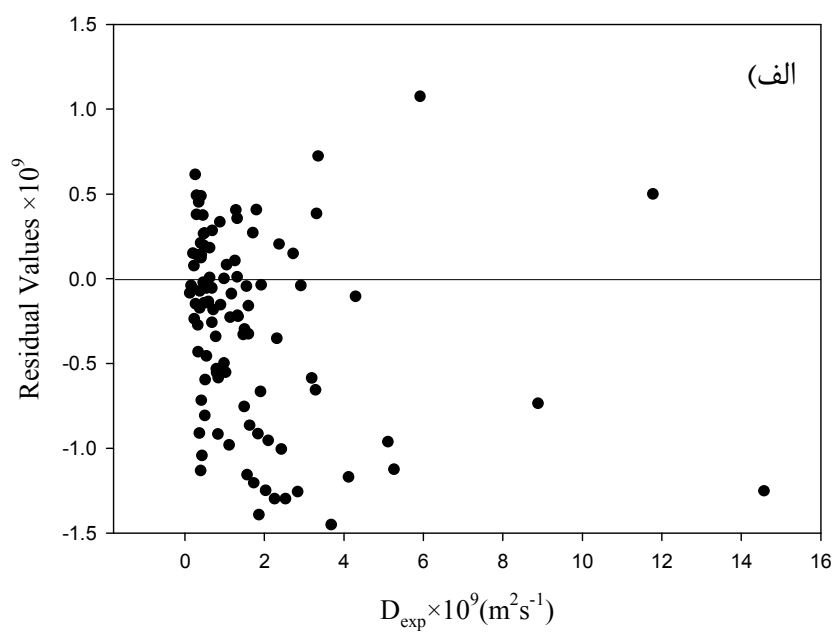
خطای باقیمانده به تفاوت مقادیر پیش‌بینی شده و مقادیر تجربی گفته می‌شود و از رسم نموداری از خطاهای باقیمانده برحسب مقادیر تجربی نمودار باقیمانده حاصل می‌شود. اگر پراکندگی مقادیر در حول محور خطای صفر این نمودار یکنواخت باشد، نشان‌دهنده‌ی مناسب بودن مدل و تصادفی بودن خطاها است.

شکل‌های (۳-۲۰) تا (۳-۲۳) نمودار خطای باقیمانده‌های مربوط به چهار روش PCR، PLS، SR-ANN و GA-ANN برای سری ارزیابی و سری تست را نشان می‌دهند. همان طور که گفته شد پراکندگی اتفاقی نقاط در حول محور خطای صفر نمودار، نشان دهنده‌ی عدم وجود خطای سیستماتیک در مدل‌سازی است. همچنین در نمودارهایی که تمرکز نقاط به سمت محور خطای صفر مشاهده می‌شود، نشان از نزدیک بودن مقادیر پیش‌بینی شده به مقادیر تجربی ضریب خودنفوذی دارد.

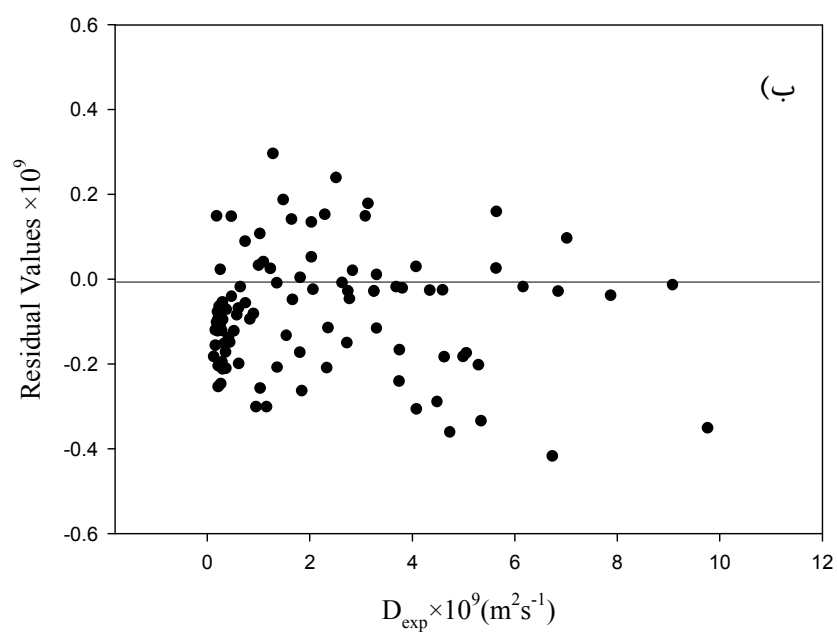
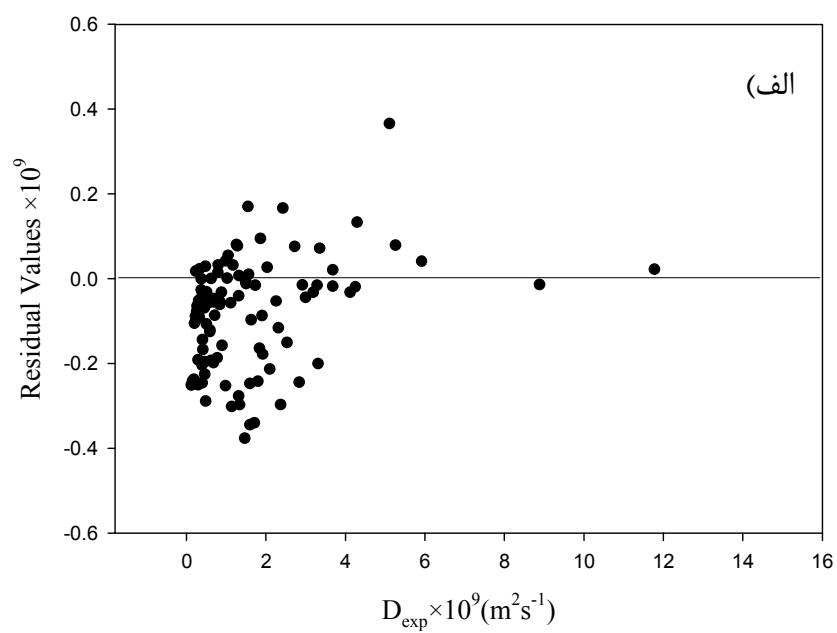
شایان ذکر است که در نمودارهای خطای باقیمانده مربوط به روش‌های PCR و PLS، بازه‌ی تغییرات خطای باقیمانده از $1/5$ تا $1/5$ و در نمودارهای باقیمانده‌ی مربوط به دو روش SR-ANN و GA-ANN این بازه بین $0/6$ تا $0/6$ در نظر گرفته شده است.



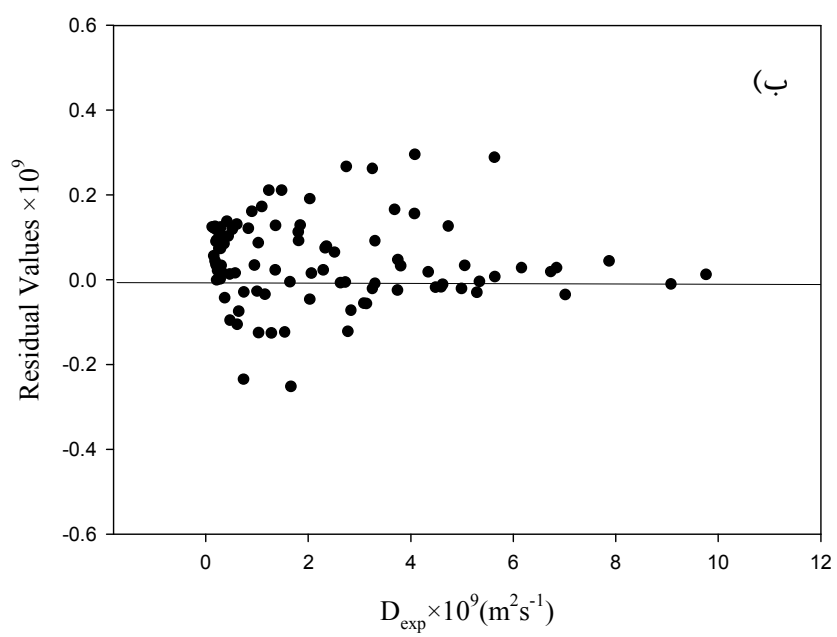
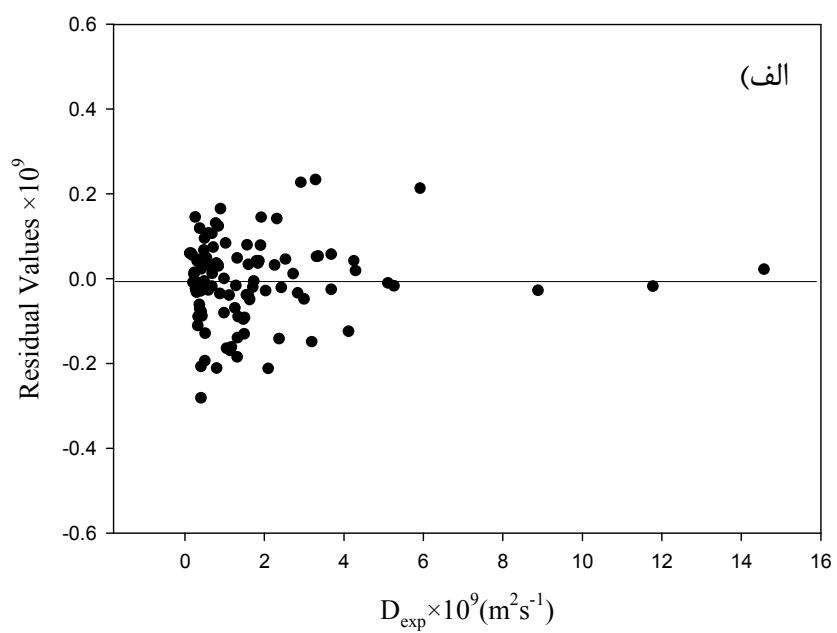
شکل (۳-۲۰): نمودار خطای باقیمانده حاصل از PCR برای (الف) سری ارزیابی و (ب) سری تست درونی



شکل (۳-۲۱): نمودار خطای باقیمانده حاصل از PLS برای (الف) سری ارزیابی و (ب) سری تست درونی



شکل (۳-۲۲): نمودار خطای باقیمانده حاصل از SR-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی

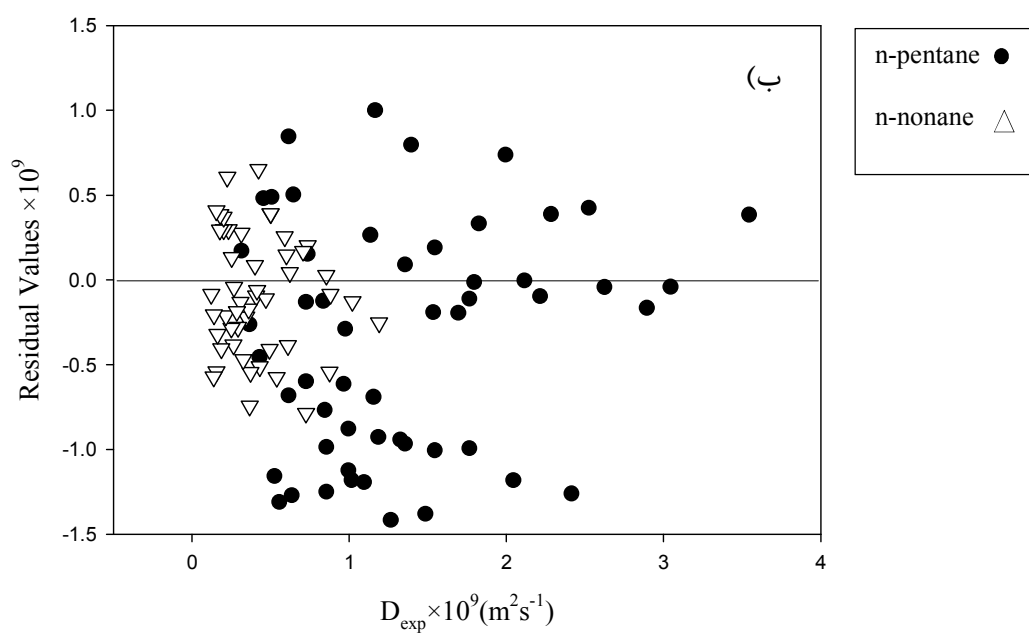
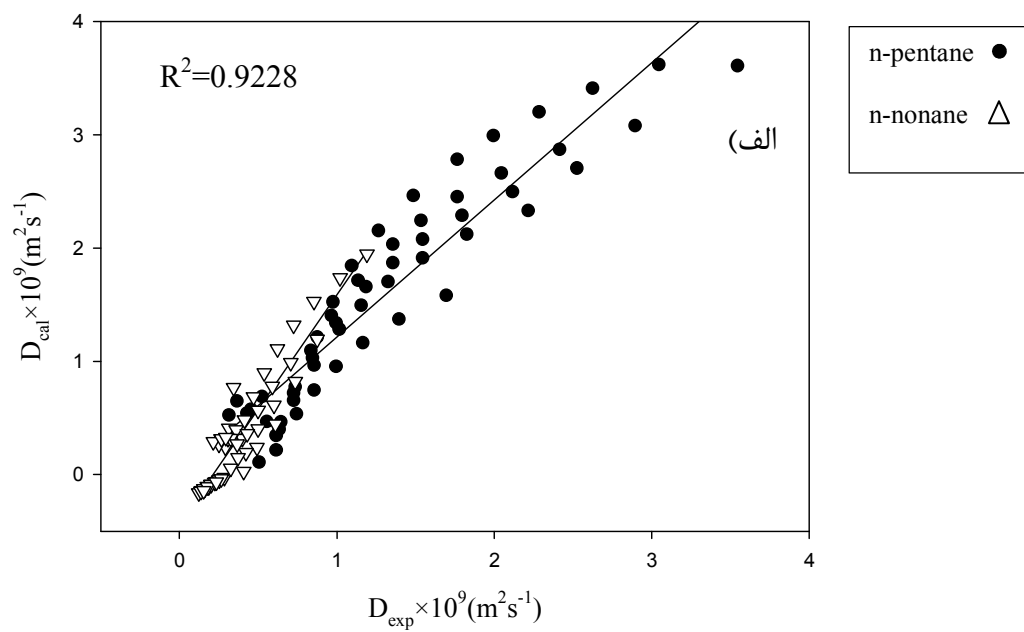


شکل (۳-۲۳): نمودار خطای باقیمانده حاصل از GA-ANN برای (الف) سری ارزیابی و (ب) سری تست درونی

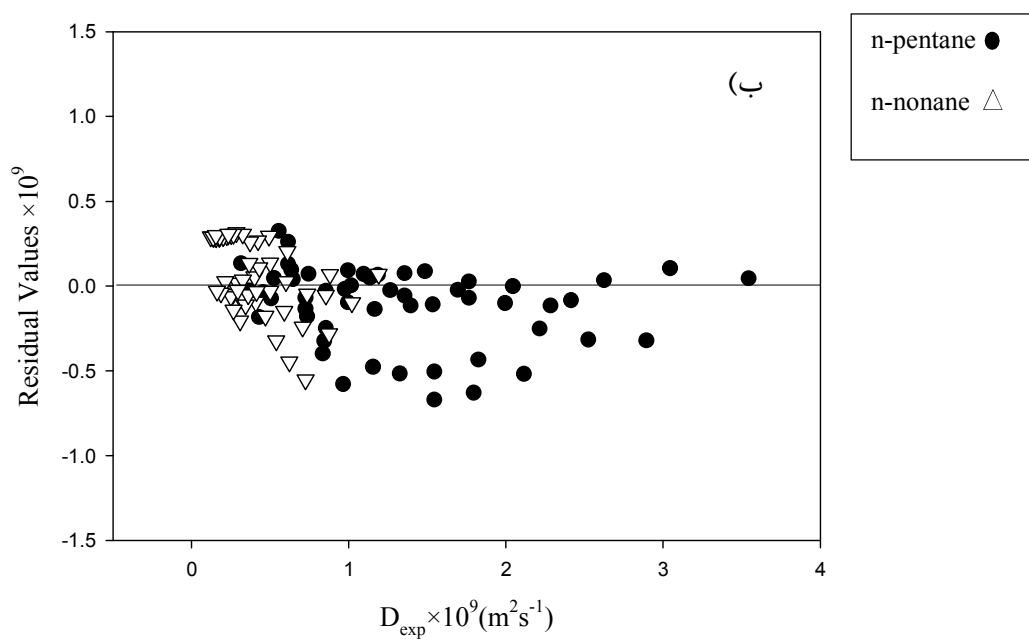
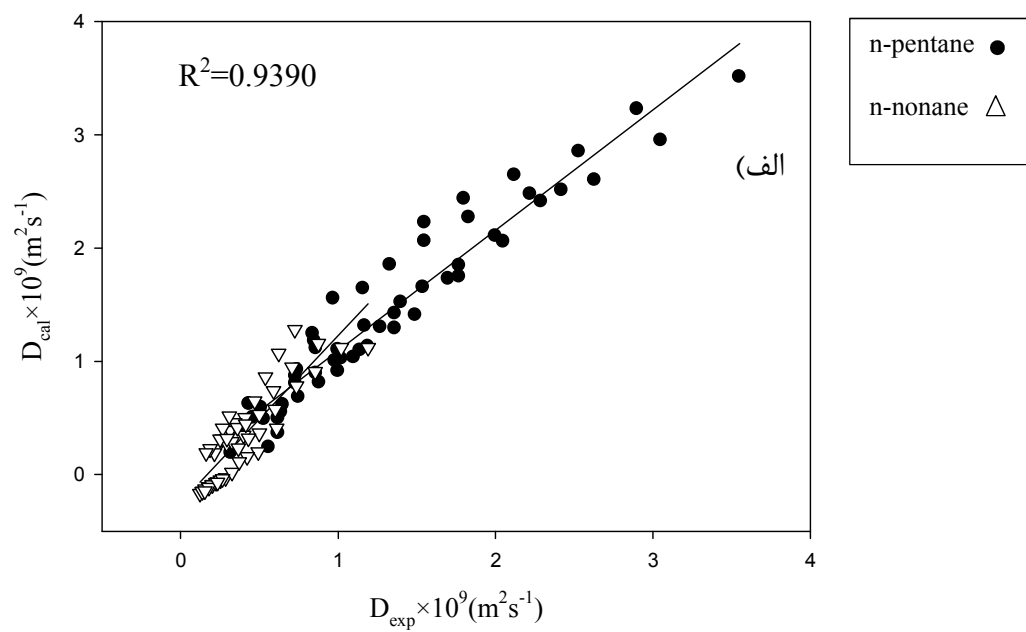
۳-۱-۹-۴- ارزیابی مدل‌های خطی و غیر خطی با استفاده از اعتبارسنجی

بیرونی

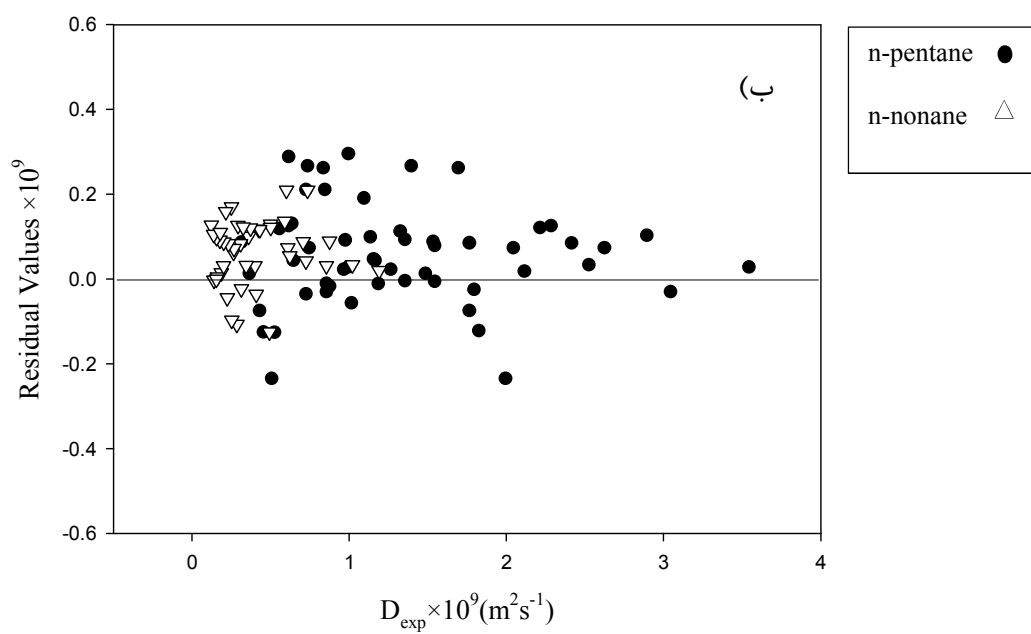
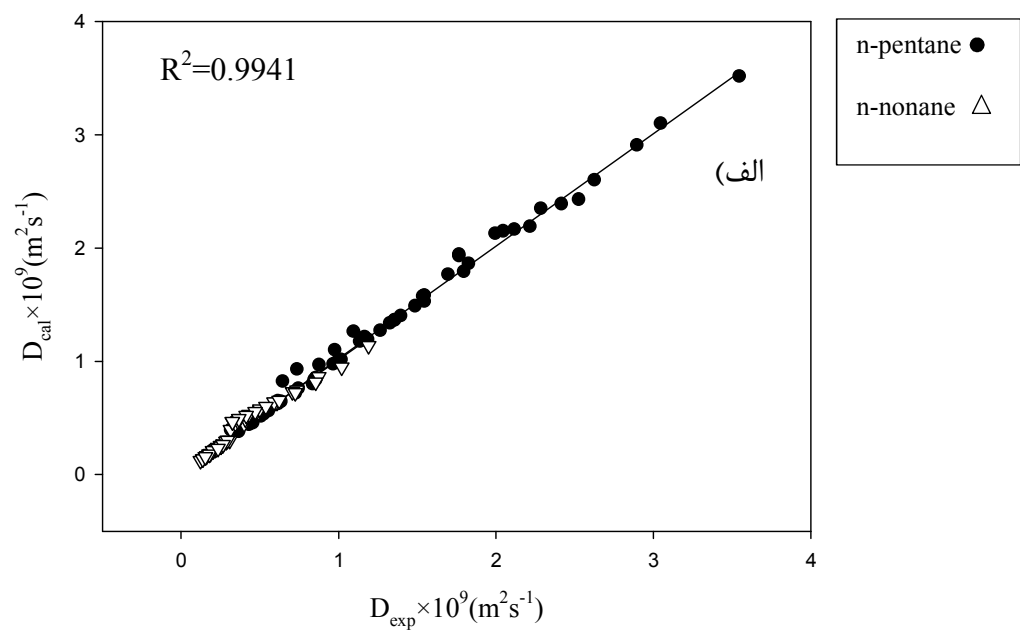
همان‌طور که گفته شد در تقسیم‌بندی داده‌ها، داده‌های مربوط به دو مولکول n -پنتان و n -نونان از سری داده‌های مدل‌سازی جدا شده و این داده‌ها در هیچیک از مراحل انتخاب توصیف‌کننده و مدل‌سازی نقشی نداشتند. مقادیر پیش‌بینی شده برای این دو مولکول توسط مدل‌های مختلف، معیار بسیار تعیین‌کننده‌ای جهت سنجش کارآمدی و اعتبار مدل ساخته شده محسوب می‌شود. نمودارهای برگشتی و نمودارهای خطای باقیمانده مربوط روش‌های PCR، PLS، SR-ANN و GA-ANN برای سری تست بیرونی به ترتیب در شکل‌های (۳-۲۴)، (۳-۲۵)، (۳-۲۶) و (۳-۲۷) آورده شده است. همچنین مقادیر محاسبه شده‌ی ضریب خودنفوذی توسط روش‌های PCR، PLS، SR-ANN و GA-ANN برای سری تست بیرونی یعنی دو مولکول n -پنتان و n -نونان در جدول (پ-۲) با مقادیر آزمایشگاهی مقایسه شده‌اند.



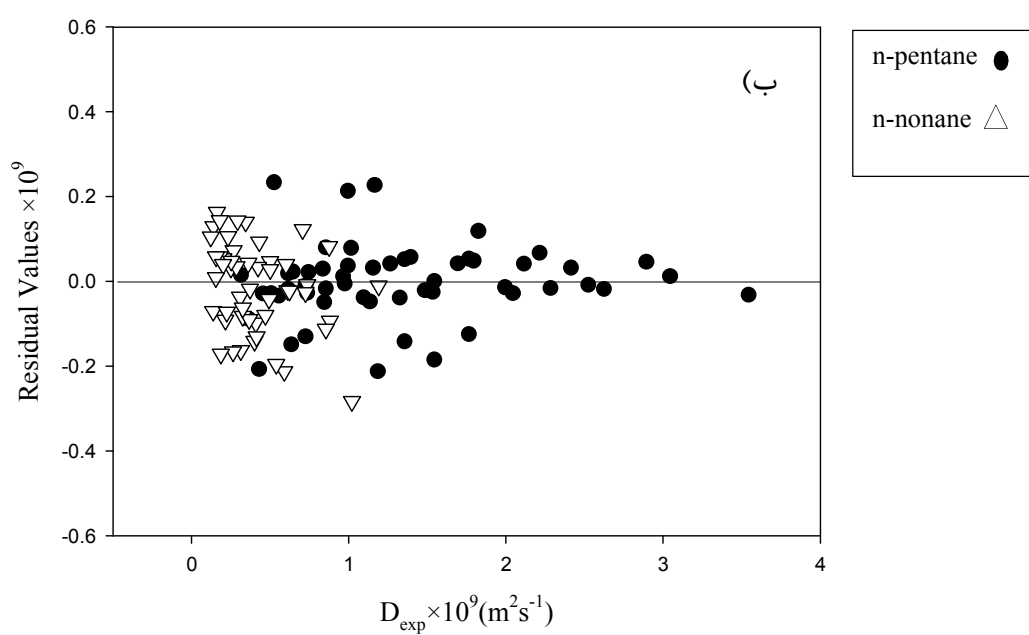
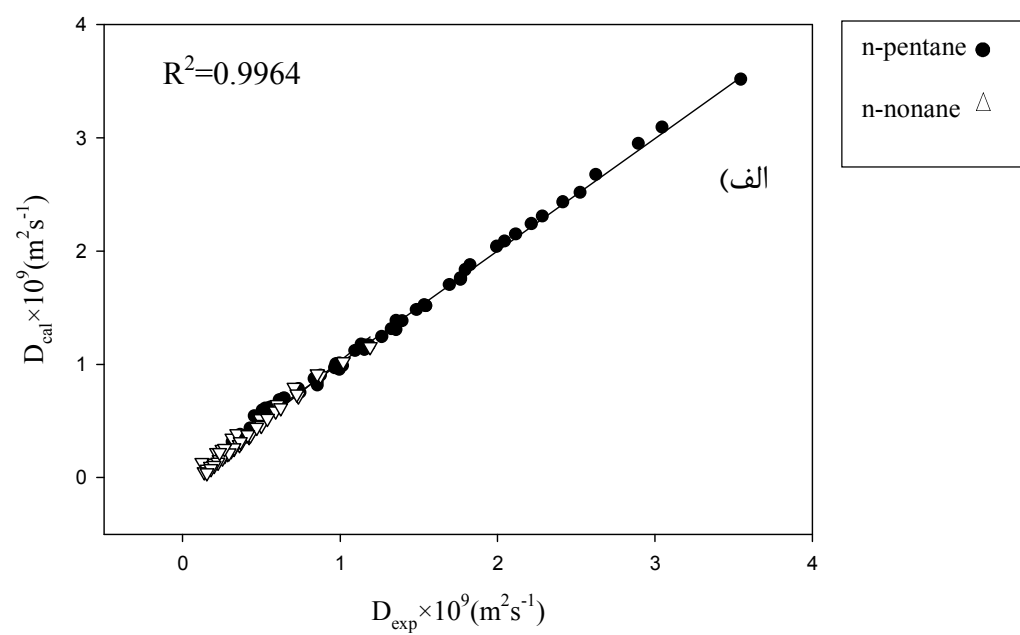
شکل (۳-۲۴): الف) نمودار برگشتی و ب) نمودار خطای باقیمانده حاصل از روش PCR برای سری تست بیرونی



شکل (۳-۲۵): الف) نمودار برگشتی و ب) نمودار خطای باقیمانده حاصل از روش PLS برای سری تست بیرونی



شکل (۳-۲۶): الف) نمودار برگشتی و ب) نمودار خطای باقیمانده حاصل از روش SR-ANN برای سری تست بیرونی



شکل (۳-۲۷): الف) نمودار برگشتی و ب) نمودار خطای باقیمانده حاصل از روش GA-ANN برای سری تست بیرونی

۳-۱-۹-۵- ارزیابی مدل‌های غیر خطی با استفاده از آزمون Y تصادفی

در این روش مقادیر متغیر وابسته یا همان خاصیت مورد اندازه‌گیری، در گستره‌ی داده‌های تجربی به صورت تصادفی تغییر داده شدند و شبکه‌های عصبی بهینه شده با استفاده از متغیرهای مستقل و مقادیر تصادفی، اجرا شدند. سپس همبستگی متغیرهای مستقل با متغیر وابسته توسط ضریب تعیین (R^2) برای سری‌های ارزیابی و تست محاسبه گردید. مقادیر پایین ضریب تعیین نشان‌دهنده‌ی عدم وابستگی در مدل توسعه یافته توسط شبکه‌ی عصبی می‌باشد. در جدول (۳-۲۳) نتایج حاصل از تکرار این آزمون برای دو سری ارزیابی و تست نشان داده شده است.

جدول (۳-۲۳): مقادیر R^2 به‌دست آمده برای سری ارزیابی و تست از آزمون Y تصادفی در دو مدل SR-ANN و

GA-ANN

شماره‌ی تکرار آزمون	SR-ANN		GA-ANN	
	R^2 Validation	R^2 Test	R^2 Validation	R^2 Test
۱	۰/۳۳۲۸	۰/۳۵۵۳	۰/۱۱۴۲	۰/۱۴۶۴
۲	۰/۱۹۲۵	۰/۱۷۰۲	۰/۳۰۰۲	۰/۲۹۹۴
۳	۰/۳۱۵۹	۰/۳۵۷۱	۰/۱۹۵۱	۰/۱۵۸۴
۴	۰/۱۹۹۸	۰/۱۷۴۲	۰/۱۶۳۱	۰/۱۴۷۲
۵	۰/۱۴۰۶	۰/۱۹۷۱	۰/۰۶۸۱	۰/۰۳۶۰
۶	۰/۱۵۵۱	۰/۱۰۵۲	۰/۰۹۷۲	۰/۱۰۰۴
۷	۰/۰۵۸۲	۰/۰۲۰۶	۰/۰۸۱۳	۰/۰۲۹۵
۸	۰/۰۶۶۰	۰/۰۴۱۱	۰/۲۰۹۳	۰/۱۰۸۱
۹	۰/۲۲۶۴	۰/۳۱۶۰	۰/۱۸۱۷	۰/۱۰۰۲
۱۰	۰/۳۳۱۳	۰/۲۹۸۳	۰/۰۲۶۳	۰/۰۴۷۸

۳-۱-۱۰- بررسی ارتباط توصیف‌کننده‌های وارد شده در مدل با ضریب

خودنفوذی

در این بخش توصیف‌کننده‌های تئوری منتخب شبکه‌های SR-ANN و GA-ANN معرفی شده و ارتباط آن‌ها با ضریب خودنفوذی مورد بررسی قرار می‌گیرد. این توصیف‌کننده‌ها عبارتند از: X_t ، $PJ13$ ، $CIC5$ ، $Mor12u$.

۳-۱-۱۰-۱- توصیف‌کننده‌های هندسی

این توصیف‌کننده‌ها در دسته‌ی توصیف‌کننده‌های سه بعدی قرار می‌گیرند زیرا برای محاسبه‌ی آن‌ها نیاز به ساختار سه بعدی مولکول است. این توصیف‌کننده‌ها اطلاعات جامع‌تری نسبت به توصیف‌کننده‌های توپولوژیکی در اختیار قرار می‌دهند زیرا ساختار سه بعدی مولکول به خوبی موقعیت اتم‌ها در فضا، نوع پیوندهای اتمی، زوایای پیوندی و ... را نشان می‌دهد. توصیف‌کننده‌ی $PJ13$ در دسته‌ی توصیف‌کننده‌های سه بعدی قرار دارد. این توصیف‌کننده زیر مجموعه‌ای از توصیف‌کننده‌های شکل^۱ مولکولی محسوب می‌شود. شکل مولکولی به چندین عامل فیزیکی-شیمیایی وابسته است و درجه‌ی انحراف مولکول از حالت کروی تحت تأثیر این عوامل را آنیزومتري^۲ می‌نامند. شاخص‌های متفاوتی برای تعیین مقدار آنیزومتري وجود دارد از جمله شاخص شکل پتیت‌جین^۳ که عبارت شکل گراف-تئوری (I_2) نیز نامیده می‌شود.

$$I_2 = \frac{D-R}{R} \quad 0 \leq I_2 \leq 1 \quad (۱۰-۳)$$

که در آن D و R به ترتیب شعاع توپولوژیکی و قطر توپولوژیکی هستند. که توسط تئوری گراف

1-Shape
2- Anisometry
3- Petitjean

مولکولی محاسبه می‌شوند.

عبارت شکل هندسی I_3 از رابطه‌ی مشابهی به دست می‌آید.

$$I_3 = \frac{{}^G D - {}^G R}{{}^G R} \quad (۱۱-۳)$$

که در این رابطه ${}^G R$ و ${}^G D$ شعاع و قطر هندسی هستند که به وسیله‌ی هندسه‌ی فضایی محاسبه می‌شوند [۵۰].

طبق اطلاعات گزارش شده در جدول (۳-۹) با افزایش اندازه‌ی مولکول‌ها مقدار عددی این توصیف‌کننده کم و بیش افزایش می‌یابد. از طرف دیگر با افزایش اندازه‌ی مولکول‌ها مقدار ضریب خودنفوذی کاهش می‌یابد. بنابراین ارتباط این توصیف‌کننده با کمیت D_{AA} ، ارتباط معکوس است. این امر بدیهی به نظر می‌رسد زیرا با افزایش اندازه و قطر هندسی مولکول، نفوذ مولکول‌ها در یکدیگر با ممانعت فضایی روبرو می‌گردد.

۳-۱-۱۰-۲- توصیف‌کننده‌های Topological

توصیف‌کننده‌های توپولوژیکی یا توصیف‌کننده‌های جانمایی، توصیف‌کننده‌های عمومی هستند و به طور مستقیم از ساختار مولکول محاسبه می‌شوند و مستقل از صورتبندی مولکول‌ها هستند. این توصیف‌کننده‌ها از گراف مولکولی به دست می‌آیند و بیانی از شاخه‌ای بودن مولکول و اندازه‌ی آن هستند. توصیف‌کننده‌های $CIC5$ و Xt در دسته‌ی توصیف‌کننده‌های توپولوژیکی قرار دارند.

توصیف‌کننده‌ی Xt از اندیس‌های ارتباط مولکولی محسوب می‌شود. اندیس‌های ارتباط مولکولی جزء مطرح‌ترین توصیف‌کننده‌های توپولوژیکی هستند که برای محاسبه‌ی این توصیف‌کننده‌ها باید ساختار مولکول با استفاده از تئوری گراف شیمیایی رسم شود. این اندیس، اندیس شاخه‌ای یا اندیس

اتصال نیز نامیده می‌شود.

رندیک^۱ برای نخستین بار رابطه‌ی (۳-۳) را برای محاسبه‌ی این اندیس پیشنهاد نمود.

$$X_R = {}^1X = \sum_b (\delta_i \cdot \delta_j)_b^{-1/2} \quad (3-3)$$

که در این معادله δ_i و δ_j درجات رأس^۲ اتم‌های دو سر پیوند در یک مولکول، هنگام برخوردی بین مولکولی را نشان می‌دهد و این مقدار در زنجیره‌ی مولکولی محاسبه می‌شود. هر عبارت $(\delta_i \cdot \delta_j)^{-1/2}$ به تنهایی، لبه‌ی^۳ اتصال نامیده می‌شود و می‌تواند برای مشخص کردن طول پیوندهای قبلی مولکول به کار برده شود. در محاسبات مربوط به اندیس ارتباط مولکولی برای سادگی محاسبات از اتم‌های هیدروژن صرف نظر شده است.

موقعیت مکانی اتم‌های یک مولکول در هنگام برخوردهای مولکولی، در توانایی برقراری پیوند آن مولکول با سایر مولکول‌ها در محیطی بدون واسطه تأثیرگذار است. اندیس 1X عبارتی برای بیان سهم یک مولکول در انجام واکنش‌های احتمالی در هنگام برخوردهای بین مولکولی است.

$${}^1X = \sqrt{\sum_{b=1}^B \sum_{b'=1}^B (\delta_i \cdot \delta_j)_b^{-1/2} \cdot (\delta_k \cdot \delta_l)_{b'}^{-1/2}} \quad (4-3)$$

که دو سیگمای معادله بیانگر حضور تمامی پیوندهای مولکولی و کلیه‌ی درجات رئوس هستند.

Xt یا اندیس ارتباطی کل، از محاسبه‌ی هم‌زمان تمامی اندیس‌های ارتباط برای تمامی اتم‌های یک مولکول حاصل می‌شود [۵۰].

$$Xt = \left(\prod_{i=1}^A \delta_i \right)^{-1/2} \quad (5-3)$$

1- Randic
2- Vertex degrees
3- Edge

برای آلکان‌های خطی مقدار X_t در جدول (۳-۹) گزارش شده است. طبق مقادیر این جدول هرچه اندازه‌ی مولکول بزرگتر می‌شود اندیس ارتباط مولکولی کل (X_t) کاهش می‌یابد. از طرف دیگر با افزایش اندازه‌ی مولکول‌ها ضریب خودنفوذی نیز کاهش می‌یابد. به این ترتیب این توصیف‌کننده با ضریب خودنفوذی ارتباط مستقیم دارد. این موضوع با اثر متوسط مثبت گزارش شده برای این توصیف‌کننده در جدول (۳-۷) مطابقت دارد. می‌توان گفت با افزایش اندیس ارتباط کل یک مولکول تمایل مولکول‌ها برای نفوذ در یکدیگر افزایش پیدا می‌کند و این امر موجب افزایش ضریب خودنفوذی می‌گردد. شایان ذکر است که مقدار اندیس ارتباط مولکولی برای مولکول متان برابر صفر است زیرا متان تنها یک اتم غیر هیدروژنی دارد.

توصیف‌کننده‌ی CIC5 نیز از توصیف‌کننده‌های توپولوژیکی است که به تقارن همسایه‌های اتم مرکزی مربوط می‌شود. این توصیف‌کننده با تعداد اتم‌های کربن مشابهی که از شکستن پیوندها به وجود می‌آیند، ارتباط دارد و عدد توصیف‌کننده نیز مرتبه‌ی شکست پیوندها را مشخص می‌سازد [۵۰]. روند تغییرات CIC5 برای سری مولکولی این تحقیق در جدول (۳-۵) نشان داده شده است. در مولکول‌های با تعداد کربن زوج مقدار عددی CIC5 بیشتر از مولکول‌های با تعداد کربن فرد است چون در مولکول‌های با تعداد اتم زوج اگر مولکول از وسط شکسته شود دو شاخه‌ی یکسان تشکیل می‌شود اما مولکول‌های با تعداد اتم کربن فرد چنین ویژگی ندارند. همچنین مقدار این توصیف‌کننده برای مولکول متان برابر صفر محاسبه شده است زیرا در مولکول متان پیوند C-C وجود ندارد. رابطه‌ی این توصیف‌کننده نیز با ضریب خودنفوذی معکوس است زیرا با افزایش مقدار عددی این توصیف‌کننده ضریب خودنفوذی کاهش می‌یابد.

۳-۱۰-۱-۳- توصیف‌کننده‌های 3D- MORSE^۱

توصیفگرهای 3D- MORSE جزء توصیف‌کننده‌های نمایش سه بعدی ساختارهای مولکولی براساس پراش الکترون طبقه‌بندی می‌شوند. هدف توصیف‌کننده‌های 3D-MORSE این است که ساختار سه بعدی مولکول را به یک کد ثابت تبدیل کنند. این عمل توسط معادله‌ی (۳-۶) که به منظور فراهم آوردن منحنی‌های نظری تفرق در مطالعات پراش الکترونی به کار رفته، انجام می‌گیرد.

$$Mor(S, A) = I(s, A) = \sum_{i=2}^N \sum_{j=i}^{i-1} A_i A_j \frac{\sin(sr_{ij})}{sr_{ij}} \quad (۳-۶)$$

که در این رابطه I شدت الکترون پراکنده شده، A_i و A_j خاصیت اتمی اتم‌های i و j ، r_{ij} فاصله‌ی بین اتم‌های i و j ، N تعداد اتم‌ها و پارامتر S بیانگر پراکندگی در جهات مختلف توسط مجموعه‌ای از اتم‌ها که در موقعیت r_{ij} قرار گرفته‌اند، می‌باشد. زاویه پراکندگی از رابطه‌ی (۳-۷) به دست می‌آید.

$$S = 4\pi \cdot \sin(v/2) / \lambda \quad (۳-۷)$$

که در این رابطه v زاویه‌ی پراکندگی و λ طول موج پرتو الکترون است. تمام روابط با فرض اینکه مولکول‌ها صلب باشند، ارائه شده‌اند. با توجه به اینکه توصیف‌کننده‌های 3D-MORSE آرایش سه بعدی در مولکول‌ها را بدون آنکه به اندازه‌ی مولکول ارتباط داشته باشند بیان می‌کنند، برای بسیاری از مولکول‌ها با ساختار اتمی متفاوت مورد استفاده قرار می‌گیرند. توصیف‌کننده‌ی Mor12u، بر خلاف سایر توصیف‌کننده‌های این گروه وزن‌دار نمی‌شود و به‌صورت بدون وزن محاسبه می‌گردد [۵۰].

همان‌طور که در جدول (۳-۴) گزارش شده است، مقدار این توصیف‌کننده با کاهش تقارن

1- 3D- Molecular representation of structure based on electron diffraction descriptors

کروی مولکول‌ها به شدت کاهش می‌یابد به صورتی که مقدار عددی این توصیف‌کننده برای دو مولکول متان و اتان مثبت و برای سایر آلکان‌های خطی منفی است. این روند با توجه به کاهش ضریب‌خودنفوذی در برابر کاهش تقارن کروی مولکول‌ها، کاملاً مطابقت دارد. ضریب تأثیر مثبت گزارش شده در جدول (۳-۳) نیز این امر را تأیید می‌نماید.

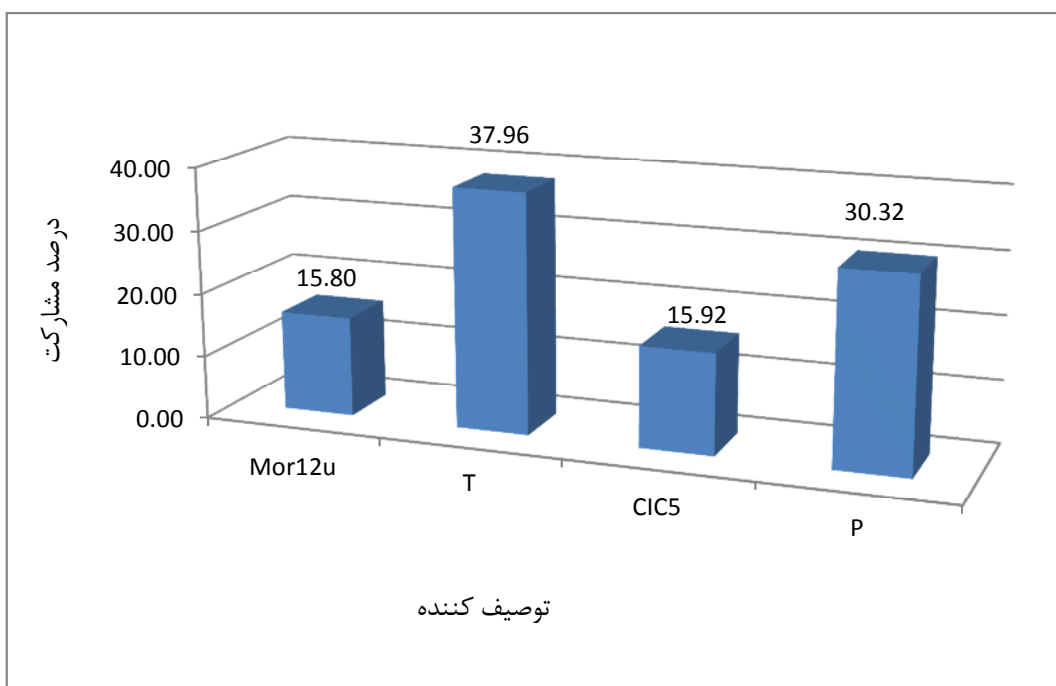
۳-۱-۱۱- بررسی میزان مشارکت توصیف‌کننده‌های منتخب شبکه‌ی عصبی

به منظور محاسبه‌ی میزان مشارکت توصیف‌کننده‌های منتخب، ابتدا توصیف‌کننده‌ی اول به همراه وزن مربوط به آن از شبکه‌ی بهینه حذف گردید. سپس مقدار متغیر وابسته یعنی ضریب خودنفوذی، با استفاده از سایر توصیف‌کننده‌ها برای هر یک از ترکیبات سری ارزیابی محاسبه شده و میانگین خطای مطلق یا MAE برای سری ارزیابی محاسباتی اندازه‌گیری شد. این محاسبات برای هریک از توصیف‌کننده‌های مربوط به شبکه‌ی عصبی بهینه انجام گرفت. در نهایت درصد مشارکت توصیف‌کننده‌ها از رابطه‌ی (۳-۸) محاسبه گردید:

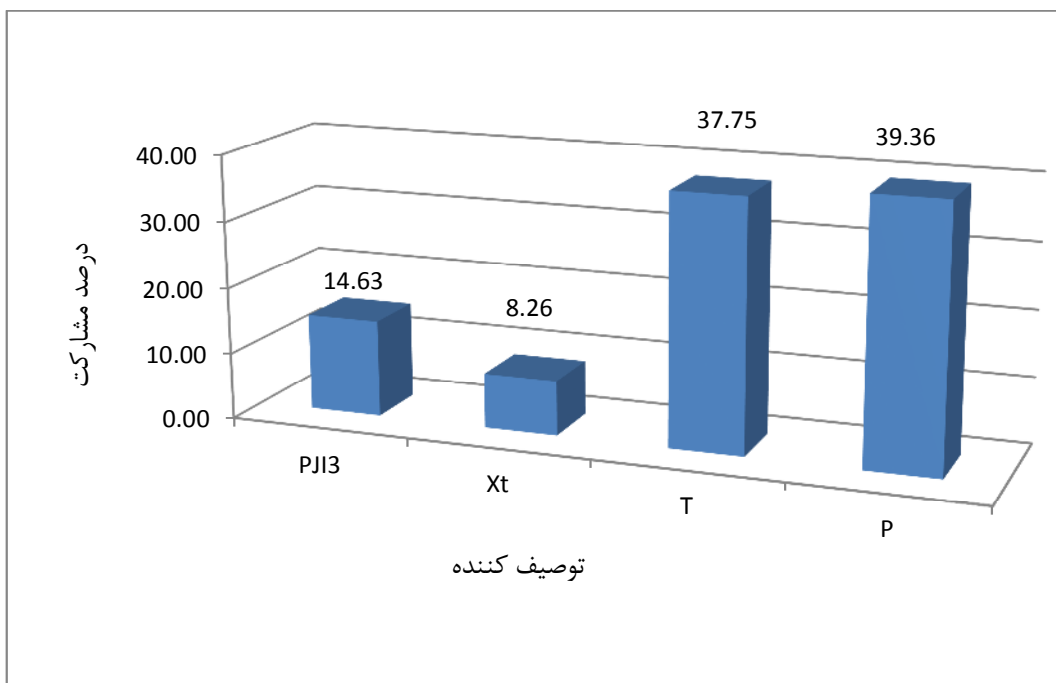
$$C_i = 100 \frac{\Delta m_i}{\sum_{i=1}^N \Delta m_i} \quad (۳-۸)$$

که در این رابطه C_i درصد مشارکت توصیف‌کننده‌ی حذف شده‌ی i ام، Δm_i میانگین خطای مطلق حاصل از سری ارزیابی در غیاب توصیف‌کننده‌ی i ام و N تعداد توصیف‌کننده‌ها در مدل است [۱۸]. مراحل فوق برای محاسبه‌ی درصد مشارکت سایر توصیف‌کننده‌ها تکرار شد. نتایج این محاسبات در شکل‌های (۳-۲۸) و (۳-۲۹) نمایش داده شده است. بر اساس نتایج حاصل از سهم مشارکت توصیف‌کننده‌ها، می‌توان دریافت که دما و فشار در هر دو شبکه بیشترین اثر مشارکت را دارا می‌باشند. وابستگی بالای ضریب خودنفوذی به دو پارامتر دما و فشار، در داده‌های تجربی نیز قابل

مشاهده است.



شکل (۳-۲۸): درصد مشارکت توصیف کننده‌ها در مدل SR-ANN



شکل (۳-۲۹): درصد مشارکت توصیف کننده‌ها در مدل GA-ANN

۳-۲- نتیجه‌گیری

در این تحقیق از روش‌های خطی و غیرخطی جهت پیش‌بینی ضریب خودنفوذی آلکان‌های خطی بهره گرفته شد. نتایج نشان می‌دهند روش‌های خطی PCR و PLS با وجود برتری‌هایی که نسبت به روش‌های خطی مشابه مانند MLR دارند اما در پیش‌بینی ضریب خودنفوذی سری داده‌های این تحقیق، کارآمد نیستند چراکه مقادیر پیش‌بینی شده توسط این دو مدل در مقایسه با مقادیر آزمایشگاهی با خطای بالایی همراه هستند. اما به طور کلی نتایج حاصل از روش PLS در مقایسه با روش PCR، با خطای کمتری همراه است.

روش غیرخطی به کار برده شده در این تحقیق، روش شبکه‌ی عصبی مصنوعی است که توصیف‌کننده‌های وارد شده به آن، یک بار توسط روش خطی رگرسیون مرحله‌ای و بار دیگر توسط روش غیرخطی الگوریتم ژنتیک انتخاب شدند. گرچه در هر دو شبکه‌ی SR-ANN و GA-ANN تعداد توصیف‌کننده‌های انتخاب شده توسط شبکه‌ها یکسان هستند اما نتایج نشان می‌دهد که مقادیر پیش‌بینی شده توسط شبکه‌ی GA-ANN به مقادیر حقیقی نزدیک‌ترند. پس در نهایت می‌توان نتیجه گرفت مدل‌سازی غیرخطی که توصیف‌کننده‌های وارد شده به آن نیز با روش غیرخطی انتخاب شده‌اند، بهترین نتایج را حاصل نموده‌اند. این واقعیت از مقایسه‌ی مقادیر مربوط به ضریب تعیین هر مدل که در جدول (۳-۲۲) گردآوری شده‌است نتیجه می‌شود. همچنین مقایسه‌ی پارامترهای آماری نشان می‌دهد که مقادیر پیش‌بینی شده با استفاده از روش‌های خطی PCR و PLS در مقایسه با روش غیرخطی شبکه‌ی عصبی مصنوعی با خطای بیشتری همراه هستند.

در فصل یک شرح داده شد که ضریب خودنفوذی وابستگی بالایی به دما و فشار دارد. شبکه‌ی GA-ANN که به عنوان کارآمدترین روش پیش‌بینی ضریب خودنفوذی آلکان‌های خطی در این تحقیق شناخته شد، اهمیت دما و فشار را بر ضریب خودنفوذی به درستی منعکس می‌سازد.

۳-۳- آینده‌نگری

- می‌توان از روش شبکه‌ی عصبی مصنوعی برای پیش‌بینی ضریب خودنفوذی سایر سیالات آلی نیز استفاده نمود.
- برای انتخاب توصیف‌کننده‌های مناسب، می‌توان از سایر روش‌ها نظیر روش‌های تخمین تابع ژنتیک (GFA)^۱، الگوریتم جستجوی ممنوع^۲، الگوریتم اجتماع مورچگان^۳، الگوریتم بهینه‌سازی جمعیت ذره‌ها^۴ و اتحاد جداسازی داده و انتخاب خاصیت (CDFS)^۵ استفاده و نتایج را با نتایج حاصل از این تحقیق مقایسه کرد.
- برای مدل‌سازی می‌توان از روش‌های ترکیبی PC-SR-ANN و PC-GA-ANN استفاده نمود. برای این منظور باید با استفاده از روش‌های انتخاب توصیف‌کننده مانند رگرسیون مرحله‌ای یا الگوریتم ژنتیک، بهترین مؤلفه‌های اصلی ساخته شده توسط PCA را شناسایی نموده و سپس با وارد نمودن مؤلفه‌های اصلی منتخب به شبکه‌ی عصبی مصنوعی، به طراحی و بهینه‌سازی مدل پرداخت. همچنین می‌توان از روش‌های تابع پایه‌ی شعاعی (RBF)^۶، ماشین‌های بردار پشتیبان (SVM)، حداقل مربعات ماشین‌های بردار پشتیبان (LS-SVM)^۷ و سیستم استنتاج فازی عصبی سازگار (ANFIS)^۸ برای مدل‌سازی استفاده کرد.

1- Genetic function approximation

2- Tabu search

3- Ant colony optimization

4- Honey-bee matikg optimization algorithm

5- Combined data splitting-feature selection

6- Radial basis function

7- Least squares support vector machines

8- Adaptive neuron-fuzzy inference system

پیوست

جدول (۱-پ): سری داده‌های مربوط به ضریب خودنفوذی آلکان‌های خطی به همراه دما و فشار اندازه‌گیری

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۱	متان	۱۱۰	۲/۱	۵/۱۹۰
۲	متان	۱۱۰	۲/۸	۵/۱۲۰
۳	متان	۱۱۰	۹/۱	۴/۷۸۰
۴	متان	۱۱۰	۹/۹	۴/۶۳۰
۵	متان	۱۱۰	۲۰/۸	۴/۱۸۰
۶	متان	۱۱۰	۳۳/۲	۳/۸۲۰
۷	متان	۱۱۰	۴۶/۷	۳/۴۸۰
۸	متان	۱۱۰	۶۱/۴	۲/۹۵۰
۹	متان	۱۱۰	۶۵/۶	۲/۹۱۰
۱۰	متان	۱۱۰	۸۰/۱	۲/۶۲۰
۱۱	متان	۱۴۰	۱/۵	۱۱/۱۰۰
۱۲	متان	۱۴۰	۵/۴	۱۰/۴۰۰
۱۳	متان	۱۴۰	۱۸/۱	۸/۸۶۰
۱۴	متان	۱۴۰	۳۴/۸	۷/۶۸۰
۱۵	متان	۱۴۰	۵۶/۳	۶/۴۸۰
۱۶	متان	۱۴۰	۸۴/۰	۵/۴۶۰
۱۷	متان	۱۴۰	۸۴/۸	۵/۴۰۰
۱۸	متان	۱۴۰	۱۱۹/۹	۴/۴۱۰
۱۹	متان	۱۴۰	۱۶۴/۴	۳/۵۲۰
۲۰	متان	۱۴۰	۲۲۱/۵	۲/۸۵۰
۲۱	متان	۱۶۰	۱/۶	۱۷/۵۰۰
۲۲	متان	۱۶۰	۷/۲	۱۵/۸۰۰
۲۳	متان	۱۶۰	۱۵/۴	۱۳/۹۰۰
۲۴	متان	۱۶۰	۲۶/۸	۱۲/۵۰۰
۲۵	متان	۱۶۰	۴۲/۱	۱۰/۷۰۰
۲۶	متان	۱۶۰	۶۱/۸	۹/۰۳۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۲۷	متان	۱۶۰	۸۶/۸	۷/۹۱۰
۲۸	متان	۱۶۰	۱۱۸/۱	۶/۵۹۰
۲۹	متان	۱۶۰	۱۵۶/۷	۵/۶۲۰
۳۰	متان	۱۶۰	۲۰۴/۸	۴/۵۸۰
۳۱	اتان	۱۳۶	۲۵/۰	۱/۶۶۰
۳۲	اتان	۱۳۶	۵۰/۰	۱/۴۲۰
۳۳	اتان	۱۳۶	۱۰۰/۰	۱/۰۹۰
۳۴	اتان	۱۳۶	۱۵۰/۰	۰/۸۷۰
۳۵	اتان	۱۳۶	۲۰۰/۰	۰/۷۰۰
۳۶	اتان	۱۵۰	۲۵/۰	۲/۴۸۰
۳۷	اتان	۱۵۰	۵۰/۰	۲/۰۶۰
۳۸	اتان	۱۵۰	۱۰۰/۰	۱/۶۰۰
۳۹	اتان	۱۵۰	۱۵۰/۰	۱/۲۸۰
۴۰	اتان	۱۵۰	۲۰۰/۰	۱/۰۳۰
۴۱	اتان	۱۶۸	۲۵/۰	۳/۴۳۰
۴۲	اتان	۱۶۸	۵۰/۰	۲/۹۷۰
۴۳	اتان	۱۶۸	۱۰۰/۰	۲/۲۹۰
۴۴	اتان	۱۶۸	۱۵۰/۰	۱/۸۷۰
۴۵	اتان	۱۶۸	۲۰۰/۰	۱/۵۱۰
۴۶	اتان	۱۸۲	۲۵/۰	۴/۲۷۰
۴۷	اتان	۱۸۲	۵۰/۰	۳/۷۰۰
۴۸	اتان	۱۸۲	۱۰۰/۰	۲/۹۴۰
۴۹	اتان	۱۸۲	۱۵۰/۰	۲/۴۰۰
۵۰	اتان	۱۸۲	۲۰۰/۰	۲/۰۲۰
۵۱	اتان	۲۰۲	۲۵/۰	۶/۰۴۰
۵۲	اتان	۲۰۲	۵۰/۰	۵/۰۵۰
۵۳	اتان	۲۰۲	۱۰۰/۰	۳/۸۰۰
۵۴	اتان	۲۰۲	۱۵۰/۰	۳/۲۷۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۵۵	اتان	۲۰۲	۲۰۰/۰	۲/۷۶۰
۵۶	اتان	۲۲۳	۲۵/۰	۷/۸۹۰
۵۷	اتان	۲۲۳	۵۰/۰	۶/۸۶۰
۵۸	اتان	۲۲۳	۱۰۰/۰	۵/۳۱۰
۵۹	اتان	۲۲۳	۱۵۰/۰	۴/۳۴۰
۶۰	اتان	۲۲۳	۲۰۰/۰	۳/۷۵۰
۶۱	اتان	۲۵۱	۲۵/۰	۱۱/۳۰۰
۶۲	اتان	۲۵۱	۵۰/۰	۹/۴۴۰
۶۳	اتان	۲۵۱	۱۰۰/۰	۷/۱۸۰
۶۴	اتان	۲۵۱	۱۵۰/۰	۵/۹۴۰
۶۵	اتان	۲۵۱	۲۰۰/۰	۵/۲۸۰
۶۶	اتان	۲۷۳	۲۵/۰	۱۴/۶۰۰
۶۷	اتان	۲۷۳	۵۰/۰	۱۱/۸۰۰
۶۸	اتان	۲۷۳	۱۰۰/۰	۸/۹۱۰
۶۹	اتان	۲۷۳	۱۵۰/۰	۷/۲۴۰
۷۰	اتان	۲۷۳	۲۰۰/۰	۶/۲۷۰
۷۱	اتان	۲۹۴	۲۵/۰	۱۸/۷۰۰
۷۲	اتان	۲۹۴	۵۰/۰	۱۴/۰۰۰
۷۳	اتان	۲۹۴	۱۰۰/۰	۱۰/۳۰۰
۷۴	اتان	۲۹۴	۱۵۰/۰	۸/۵۷۰
۷۵	اتان	۲۹۴	۲۰۰/۰	۷/۳۰۰
۷۶	اتان	۳۰۳/۲	۵۰/۰	۱۵/۲۰۰
۷۷	اتان	۳۳۳/۱	۳۰/۰	۲۵/۰۰۰
۷۸	اتان	۳۳۳/۱	۴۰/۰	۲۱/۰۰۰
۷۹	پروپان	۱۱۲	۲۵/۰	۰/۲۴۹
۸۰	پروپان	۱۱۲	۵۰/۰	۰/۱۹۹
۸۱	پروپان	۱۱۲	۱۰۰/۰	۰/۱۴۴
۸۲	پروپان	۱۱۲	۱۵۰/۰	۰/۱۰۲

ادامہی جدول (۱-پ)

ضریب خودنفوذی ($\times 10^{-9}$)	فشار (Mpa)	دما ($^{\circ}\text{C}$)	نام مولکول	شماره
۰/۳۴۵	۲۵/۰	۱۱۹	پروپان	۸۳
۰/۲۶۵	۵۰/۰	۱۱۹	پروپان	۸۴
۰/۱۹۳	۱۰۰/۰	۱۱۹	پروپان	۸۵
۰/۱۳۷	۱۵۰/۰	۱۱۹	پروپان	۸۶
۰/۱۰۱	۲۰۰/۰	۱۱۹	پروپان	۸۷
۰/۴۲۵	۲۵/۰	۱۲۶	پروپان	۸۸
۰/۳۵۲	۵۰/۰	۱۲۶	پروپان	۸۹
۰/۲۵۲	۱۰۰/۰	۱۲۶	پروپان	۹۰
۰/۱۷۸	۱۵۰/۰	۱۲۶	پروپان	۹۱
۰/۱۳۴	۲۰۰/۰	۱۲۶	پروپان	۹۲
۰/۵۹۸	۲۵/۰	۱۳۵	پروپان	۹۳
۰/۵۰۱	۵۰/۰	۱۳۵	پروپان	۹۴
۰/۳۶۴	۱۰۰/۰	۱۳۵	پروپان	۹۵
۰/۲۷۰	۱۵۰/۰	۱۳۵	پروپان	۹۶
۰/۲۰۰	۲۰۰/۰	۱۳۵	پروپان	۹۷
۰/۷۵۲	۲۵/۰	۱۴۴	پروپان	۹۸
۰/۶۶۰	۵۰/۰	۱۴۴	پروپان	۹۹
۰/۴۸۴	۱۰۰/۰	۱۴۴	پروپان	۱۰۰
۰/۳۶۷	۱۵۰/۰	۱۴۴	پروپان	۱۰۱
۰/۲۷۵	۲۰۰/۰	۱۴۴	پروپان	۱۰۲
۱/۰۲۰	۲۵/۰	۱۵۴	پروپان	۱۰۳
۰/۸۷۸	۵۰/۰	۱۵۴	پروپان	۱۰۴
۰/۶۶۰	۱۰۰/۰	۱۵۴	پروپان	۱۰۵
۰/۴۹۹	۱۵۰/۰	۱۵۴	پروپان	۱۰۶
۰/۳۹۴	۲۰۰/۰	۱۵۴	پروپان	۱۰۷
۱/۳۳۶	۲۵/۰	۱۶۵	پروپان	۱۰۸
۱/۱۳۳	۵۰/۰	۱۶۵	پروپان	۱۰۹
۰/۸۵۱	۱۰۰/۰	۱۶۵	پروپان	۱۱۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۱۱۱	پروپان	۱۶۵	۱۵۰/۰	۰/۶۶۳
۱۱۲	پروپان	۱۶۵	۲۰۰/۰	۰/۵۱۷
۱۱۳	پروپان	۱۸۵	۲۵/۰	۲/۰۰۰
۱۱۴	پروپان	۱۸۵	۵۰/۰	۱/۷۱۹
۱۱۵	پروپان	۱۸۵	۱۰۰/۰	۱/۳۰۴
۱۱۶	پروپان	۱۸۵	۱۵۰/۰	۱/۰۴۴
۱۱۷	پروپان	۱۸۵	۲۰۰/۰	۰/۸۴۹
۱۱۸	پروپان	۲۰۳	۲۵/۰	۲/۷۹۰
۱۱۹	پروپان	۲۰۳	۵۰/۰	۲/۳۷۰
۱۲۰	پروپان	۲۰۳	۱۰۰/۰	۱/۸۲۱
۱۲۱	پروپان	۲۰۳	۱۵۰/۰	۱/۴۵۲
۱۲۲	پروپان	۲۰۳	۲۰۰/۰	۱/۱۸۶
۱۲۳	پروپان	۲۲۱	۲۵/۰	۳/۶۷۷
۱۲۴	پروپان	۲۲۱	۵۰/۰	۳/۱۰۳
۱۲۵	پروپان	۲۲۱	۱۰۰/۰	۲/۳۸۹
۱۲۶	پروپان	۲۲۱	۱۵۰/۰	۱/۹۲۴
۱۲۷	پروپان	۲۲۱	۲۰۰/۰	۱/۵۸۵
۱۲۸	پروپان	۲۴۳	۲۵/۰	۵/۱۳۳
۱۲۹	پروپان	۲۴۳	۵۰/۰	۴/۳۱۵
۱۳۰	پروپان	۲۴۳	۱۰۰/۰	۳/۳۰۵
۱۳۱	پروپان	۲۴۳	۱۵۰/۰	۲/۶۸۷
۱۳۲	پروپان	۲۴۳	۲۰۰/۰	۲/۲۳۲
۱۳۳	پروپان	۲۷۳	۲۵/۰	۷/۳۲۹
۱۳۴	پروپان	۲۷۳	۵۰/۰	۶/۱۰۶
۱۳۵	پروپان	۲۷۳	۱۰۰/۰	۴/۶۴۴
۱۳۶	پروپان	۲۷۳	۱۵۰/۰	۳/۷۶۹
۱۳۷	پروپان	۲۷۳	۲۰۰/۰	۳/۱۵۲
۱۳۸	پروپان	۲۹۴	۲۵/۰	۹/۰۹۵

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۱۳۹	پروپان	۲۹۴	۵۰/۰	۷/۰۳۶
۱۴۰	پروپان	۲۹۴	۱۰۰/۰	۵/۶۵۰
۱۴۱	پروپان	۲۹۴	۱۵۰/۰	۴/۵۷۵
۱۴۲	پروپان	۲۹۴	۲۰۰/۰	۳/۸۳۲
۱۴۳	پروپان	۳۲۳	۲۵/۰	۱۲/۳۱۰
۱۴۴	پروپان	۳۲۳	۵۰/۰	۹/۹۹۲
۱۴۵	پروپان	۳۲۳	۱۰۰/۰	۷/۴۱۱
۱۴۶	پروپان	۳۲۳	۱۵۰/۰	۶/۰۷۸
۱۴۷	پروپان	۳۲۳	۲۰۰/۰	۵/۱۲۱
۱۴۸	پروپان	۳۵۳	۲۵/۰	۱۵/۷۱۰
۱۴۹	پروپان	۳۵۳	۵۰/۰	۱۲/۲۹۰
۱۵۰	پروپان	۳۵۳	۱۰۰/۰	۹/۰۶۹
۱۵۱	پروپان	۳۵۳	۱۵۰/۰	۷/۲۷۳
۱۵۲	پروپان	۳۵۳	۲۰۰/۰	۶/۱۰۷
۱۵۳	n-بوتان	۱۹۳	۵/۰	۱/۸۶۰
۱۵۴	n-بوتان	۱۹۳	۵۰/۰	۱/۵۰۰
۱۵۵	n-بوتان	۱۹۳	۱۰۰/۰	۱/۱۴۰
۱۵۶	n-بوتان	۱۹۳	۱۵۰/۰	۰/۹۹۰
۱۵۷	n-بوتان	۱۹۳	۲۰۰/۰	۰/۸۱۷
۱۵۸	n-بوتان	۲۱۷	۵/۰	۳/۱۱۰
۱۵۹	n-بوتان	۲۱۷	۵۰/۰	۲/۲۶۰
۱۶۰	n-بوتان	۲۱۷	۱۰۰/۰	۱/۸۱۰
۱۶۱	n-بوتان	۲۱۷	۱۵۰/۰	۱/۵۷۰
۱۶۲	n-بوتان	۲۲۲	۲۰۰/۰	۱/۳۶۰
۱۶۳	n-بوتان	۲۵۱	۵/۰	۴/۷۵۰
۱۶۴	n-بوتان	۲۵۱	۵۰/۰	۳/۵۶۰
۱۶۵	n-بوتان	۲۵۱	۱۰۰/۰	۲/۹۳۰
۱۶۶	n-بوتان	۲۵۱	۱۵۰/۰	۲/۴۶۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۱۶۷	n-بوتان	۲۵۱	۲۰۰/۰	۲/۰۱۰
۱۶۸	n-بوتان	۲۹۵	۵/۰	۷/۶۱۰
۱۶۹	n-بوتان	۲۹۵	۵۰/۰	۵/۸۴۰
۱۷۰	n-بوتان	۲۹۵	۱۰۰/۰	۴/۷۰۰
۱۷۱	n-بوتان	۲۹۵	۱۵۰/۰	۴/۰۰۰
۱۷۲	n-بوتان	۲۹۵	۲۰۰/۰	۳/۴۵۰
۱۷۳	n-بوتان	۳۱۴	۸/۹	۷/۶۰۰
۱۷۴	n-بوتان	۳۳۲	۵/۰	۱۰/۸۰۰
۱۷۵	n-بوتان	۳۳۲	۵۰/۰	۸/۰۸۰
۱۷۶	n-بوتان	۳۳۲	۱۰۰/۰	۶/۵۵۰
۱۷۷	n-بوتان	۳۳۲	۱۵۰/۰	۵/۳۵۰
۱۷۸	n-بوتان	۳۶۳	۲۰۰/۰	۵/۳۳۰
۱۷۹	n-بوتان	۳۸۲	۵/۰	۲۰/۰۰۰
۱۸۰	n-بوتان	۳۸۲	۵۰/۰	۱۲/۰۰۰
۱۸۱	n-بوتان	۳۸۲	۱۰۰/۰	۸/۸۶۰
۱۸۲	n-بوتان	۳۸۲	۱۵۰/۰	۷/۵۸۰
۱۸۳	n-بوتان	۳۹۸	۴/۲	۲۵/۰۰۰
۱۸۴	n-بوتان	۴۰۶	۱۵/۴	۱۵/۲۰۰
۱۸۵	n-بوتان	۴۰۷	۱۰۱/۹	۷/۸۰۰
۱۸۶	n-بوتان	۴۱۶	۳۹/۶	۱۲/۶۰۰
۱۸۷	n-پنتان	۱۷۴	۵/۰	۰/۶۸۲
۱۸۸	n-پنتان	۱۷۴	۵۰/۰	۰/۴۶۷
۱۸۹	n-پنتان	۱۷۴	۱۰۰/۰	۰/۳۲۶
۱۹۰	n-پنتان	۱۷۴	۱۵۰/۰	۰/۲۲۳
۱۹۱	n-پنتان	۱۷۴	۲۰۰/۰	۰/۱۶۲
۱۹۲	n-پنتان	۱۹۳	۵/۰	۱/۱۲۰
۱۹۳	n-پنتان	۱۹۳	۵۰/۰	۰/۸۰۶
۱۹۴	n-پنتان	۱۹۳	۱۰۰/۰	۰/۵۶۷

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۱۹۵	n-پنتان	۱۹۳	۱۵۰/۰	۰/۴۱۵
۱۹۶	n-پنتان	۱۹۵	۲۰۰/۰	۰/۳۱۲
۱۹۷	n-پنتان	۲۱۷	۵/۰	۱/۸۲۰
۱۹۸	n-پنتان	۲۱۷	۵۰/۰	۱/۳۳۰
۱۹۹	n-پنتان	۲۱۷	۱۰۰/۰	۱/۰۰۰
۲۰۰	n-پنتان	۲۱۷	۱۵۰/۰	۰/۷۷۶
۲۰۱	n-پنتان	۲۱۷	۲۰۰/۰	۰/۵۳۹
۲۰۲	n-پنتان	۲۵۲	۵/۰	۲/۹۰۰
۲۰۳	n-پنتان	۲۵۲	۵۰/۰	۲/۳۵۰
۲۰۴	n-پنتان	۲۵۲	۱۰۰/۰	۱/۷۵۰
۲۰۵	n-پنتان	۲۵۲	۱۵۰/۰	۱/۳۰۰
۲۰۶	n-پنتان	۲۵۲	۲۰۰/۰	۱/۰۵۰
۲۰۷	n-پنتان	۲۹۳	۵/۰	۵/۰۷۰
۲۰۸	n-پنتان	۲۹۳	۲۰/۰	۴/۳۶۰
۲۰۹	n-پنتان	۲۹۳	۵۰/۰	۳/۷۶۰
۲۱۰	n-پنتان	۲۹۳	۱۰۰/۰	۲/۷۴۰
۲۱۱	n-پنتان	۲۹۳	۱۵۰/۰	۲/۰۰۰
۲۱۲	n-پنتان	۲۹۳	۲۰۰/۰	۱/۷۳۰
۲۱۳	n-پنتان	۲۹۳	۰/۱	۵/۱۶۰
۲۱۴	n-پنتان	۲۹۳	۱۹/۶	۴/۴۵۰
۲۱۵	n-پنتان	۲۹۳	۳۹/۲	۳/۸۲۰
۲۱۶	n-پنتان	۲۹۳	۵۸/۸	۳/۳۷۰
۲۱۷	n-پنتان	۲۹۳	۷۸/۴	۳/۰۱۰
۲۱۸	n-پنتان	۲۹۳	۹۸/۱	۲/۷۴۰
۲۱۹	n-پنتان	۲۹۳	۱۲۸/۰	۲/۳۹۰
۲۲۰	n-پنتان	۲۹۳	۱۵۷/۰	۲/۱۲۰
۲۲۱	n-پنتان	۲۹۳	۱۸۶/۰	۱/۸۶۰
۲۲۲	n-پنتان	۲۹۳	۲۱۶/۰	۱/۶۵۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۲۲۳	n-پنتان	۲۹۳	۲۴۵/۰	۱/۵۱۰
۲۲۴	n-پنتان	۳۱۳	۱۹/۶	۵/۳۶۰
۲۲۵	n-پنتان	۳۱۳	۳۹/۲	۴/۶۴۰
۲۲۶	n-پنتان	۳۱۳	۵۸/۸	۴/۱۰۰
۲۲۷	n-پنتان	۳۱۳	۷۸/۴	۳/۷۰۰
۲۲۸	n-پنتان	۳۱۳	۹۸/۱	۳/۳۲۰
۲۲۹	n-پنتان	۳۱۳	۱۲۸/۰	۲/۹۱۰
۲۳۰	n-پنتان	۳۱۳	۱۵۷/۰	۲/۵۸۰
۲۳۱	n-پنتان	۳۱۳	۱۸۶/۰	۲/۲۸۰
۲۳۲	n-پنتان	۳۱۳	۲۱۶/۰	۲/۰۴۰
۲۳۳	n-پنتان	۳۱۳	۲۴۵/۰	۱/۸۳۰
۲۳۴	n-پنتان	۳۲۳	۱۹/۶	۶/۰۳۰
۲۳۵	n-پنتان	۳۲۳	۳۹/۲	۵/۱۲۰
۲۳۶	n-پنتان	۳۲۳	۵۸/۸	۴/۵۳۰
۲۳۷	n-پنتان	۳۲۳	۷۸/۴	۴/۱۰۰
۲۳۸	n-پنتان	۳۲۳	۹۸/۱	۳/۷۰۰
۲۳۹	n-پنتان	۳۲۳	۱۲۸/۰	۳/۲۱۰
۲۴۰	n-پنتان	۳۲۳	۱۵۷/۰	۲/۸۶۰
۲۴۱	n-پنتان	۳۲۳	۱۸۶/۰	۲/۵۵۰
۲۴۲	n-پنتان	۳۲۳	۲۱۶/۰	۲/۲۸۰
۲۴۳	n-پنتان	۳۲۳	۲۴۵/۰	۲/۰۵۰
۲۴۴	n-پنتان	۳۳۳	۱۹/۶	۶/۷۵۰
۲۴۵	n-پنتان	۳۳۳	۳۹/۲	۵/۶۶۰
۲۴۶	n-پنتان	۳۳۳	۵۸/۸	۵/۰۱۰
۲۴۷	n-پنتان	۳۳۳	۷۸/۴	۴/۵۰۰
۲۴۸	n-پنتان	۳۳۳	۹۸/۱	۴/۰۹۰
۲۴۹	n-پنتان	۳۳۳	۱۲۸/۰	۳/۵۸۰
۲۵۰	n-پنتان	۳۳۳	۱۵۷/۰	۳/۱۸۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۲۵۱	n-پنتان	۳۳۳	۱۸۶/۰	۲/۸۱۰
۲۵۲	n-پنتان	۳۳۳	۲۱۶/۰	۲/۵۲۰
۲۵۳	n-پنتان	۳۳۳	۲۴۵/۰	۲/۲۷۰
۲۵۴	n-پنتان	۳۵۳	۱۹/۶	۸/۳۰۰
۲۵۵	n-پنتان	۳۵۳	۳۹/۲	۶/۷۰۰
۲۵۶	n-پنتان	۳۵۳	۵۸/۸	۵/۸۰۰
۲۵۷	n-پنتان	۳۵۳	۷۸/۴	۵/۲۰۰
۲۵۸	n-پنتان	۳۵۳	۹۸/۱	۴/۷۲۰
۲۵۹	n-پنتان	۳۵۳	۱۲۸/۰	۴/۱۴۰
۲۶۰	n-پنتان	۳۵۳	۱۵۷/۰	۳/۷۰۰
۲۶۱	n-پنتان	۳۵۳	۱۸۶/۰	۳/۳۳۰
۲۶۲	n-پنتان	۳۵۳	۲۱۶/۰	۳/۰۲۰
۲۶۳	n-پنتان	۳۵۳	۲۴۵/۰	۲/۷۴۰
۲۶۴	n-پنتان	۳۵۵	۵/۰	۹/۷۸۰
۲۶۵	n-پنتان	۳۵۵	۵۰/۰	۶/۱۸۰
۲۶۶	n-پنتان	۳۵۵	۱۰۰/۰	۴/۷۵۰
۲۶۷	n-پنتان	۳۵۵	۱۵۰/۰	۴/۰۳۰
۲۶۸	n-پنتان	۳۵۵	۲۰۰/۰	۳/۲۶۰
۲۶۹	n-پنتان	۴۵۰	۵/۰	۲۸/۸۰۰
۲۷۰	n-پنتان	۴۵۰	۲۰/۰	۱۶/۲۰۰
۲۷۱	n-پنتان	۴۵۰	۵۰/۰	۱۲/۲۰۰
۲۷۲	n-پنتان	۴۵۰	۱۰۰/۰	۸/۴۰۰
۲۷۳	n-پنتان	۴۵۰	۱۵۰/۰	۶/۹۵۰
۲۷۴	n-پنتان	۴۵۰	۲۰۰/۰	۵/۷۷۰
۲۷۵	n-پنتان	۴۹۷/۵	۶۳/۴	۱۳/۵۰۰
۲۷۶	n-پنتان	۴۹۷/۵	۲۴۰/۰	۶/۱۰۰
۲۷۷	n-هگزان	۲۱۴	۵/۰	۱/۱۵۰
۲۷۸	n-هگزان	۲۱۴	۵۰/۰	۰/۷۴۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۲۷۹	n-هگزان	۲۱۴	۱۰۰/۰	۰/۴۹۹
۲۸۰	n-هگزان	۲۱۴	۱۵۰/۰	۰/۳۶۸
۲۸۱	n-هگزان	۲۱۴	۲۰۰/۰	۰/۲۷۲
۲۸۲	n-هگزان	۲۲۳/۱۵	۰/۱	۱/۳۲۰
۲۸۳	n-هگزان	۲۲۳/۱۵	۴۸/۷	۰/۸۸۰
۲۸۴	n-هگزان	۲۲۳/۱۵	۵۱/۴	۰/۸۶۰
۲۸۵	n-هگزان	۲۲۳/۱۵	۱۰۰/۷	۰/۶۱۰
۲۸۶	n-هگزان	۲۲۳/۱۵	۱۵۰/۴	۰/۴۳۰
۲۸۷	n-هگزان	۲۲۳/۱۵	۲۰۰/۵	۰/۳۱۰
۲۸۸	n-هگزان	۲۲۳/۱۵	۲۵۲/۹	۰/۲۲۰
۲۸۹	n-هگزان	۲۲۳/۱۵	۲۹۲/۳	۰/۱۷۰
۲۹۰	n-هگزان	۲۴۰	۵/۰	۱/۸۰۰
۲۹۱	n-هگزان	۲۴۰	۵۰/۰	۱/۲۵۰
۲۹۲	n-هگزان	۲۴۰	۱۰۰/۰	۰/۸۷۰
۲۹۳	n-هگزان	۲۴۰	۱۵۰/۰	۰/۶۶۶
۲۹۴	n-هگزان	۲۴۳	۲۰/۰	۱/۹۶۰
۲۹۵	n-هگزان	۲۴۳	۲۰۰/۰	۰/۵۰۹
۲۹۶	n-هگزان	۲۴۸/۱۵	۰/۱	۲/۰۸۰
۲۹۷	n-هگزان	۲۴۸/۱۵	۲۵/۸	۱/۶۸۰
۲۹۸	n-هگزان	۲۴۸/۱۵	۵۵/۷	۱/۳۷۰
۲۹۹	n-هگزان	۲۴۸/۱۵	۱۰۰/۸	۱/۰۲۰
۳۰۰	n-هگزان	۲۴۸/۱۵	۱۵۴/۴	۰/۷۶۰
۳۰۱	n-هگزان	۲۴۸/۱۵	۱۹۷/۵	۰/۵۹۰
۳۰۲	n-هگزان	۲۴۸/۱۵	۲۵۰/۳	۰/۴۴۰
۳۰۳	n-هگزان	۲۴۸/۱۵	۲۹۲/۹	۰/۳۵۰
۳۰۴	n-هگزان	۲۴۸/۱۵	۳۰۵/۰	۰/۳۴۰
۳۰۵	n-هگزان	۲۴۸/۱۵	۳۰۸/۰	۰/۳۳۰
۳۰۶	n-هگزان	۲۴۸/۱۵	۳۵۰/۶	۰/۲۶۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۳۰۷	n-هگزان	۲۴۸/۱۵	۳۹۱/۲	۰/۲۱۶
۳۰۸	n-هگزان	۲۷۱	۵۰/۰	۲/۷۶۰
۳۰۹	n-هگزان	۲۷۱	۲۰/۰	۲/۴۳۰
۳۱۰	n-هگزان	۲۷۱	۵۰/۰	۱/۹۹۰
۳۱۱	n-هگزان	۲۷۱	۱۰۰/۰	۱/۵۷۰
۳۱۲	n-هگزان	۲۷۱	۱۵۰/۰	۱/۱۹۰
۳۱۳	n-هگزان	۲۷۳/۱۵	۰/۱	۲/۹۶۰
۳۱۴	n-هگزان	۲۷۳/۱۵	۲۵/۰	۲/۴۷۰
۳۱۵	n-هگزان	۲۷۳/۱۵	۲۵/۲	۲/۴۹۰
۳۱۶	n-هگزان	۲۷۳/۱۵	۹۹/۸	۱/۵۱۰
۳۱۷	n-هگزان	۲۷۳/۱۵	۱۵۲/۳	۱/۱۴۰
۳۱۸	n-هگزان	۲۷۳/۱۵	۱۹۸/۷	۰/۹۱۰
۳۱۹	n-هگزان	۲۷۳/۱۵	۲۴۸/۶	۰/۷۱۰
۳۲۰	n-هگزان	۲۷۳/۱۵	۲۴۹/۵	۰/۷۰۰
۳۲۱	n-هگزان	۲۷۳/۱۵	۳۰۲/۸	۰/۵۶۰
۳۲۲	n-هگزان	۲۷۳/۱۵	۳۴۴/۰	۰/۴۷۰
۳۲۳	n-هگزان	۲۷۳/۱۵	۳۹۳/۸	۰/۳۸۰
۳۲۴	n-هگزان	۲۹۸/۱۵	۰/۱	۴/۲۱۰
۳۲۵	n-هگزان	۲۹۸/۱۵	۲۷/۶	۳/۳۲۰
۳۲۶	n-هگزان	۲۹۸/۱۵	۵۰/۰	۲/۸۵۰
۳۲۷	n-هگزان	۲۹۸/۱۵	۱۰۱/۴	۲/۰۸۰
۳۲۸	n-هگزان	۲۹۸/۱۵	۱۰۵/۳	۲/۰۵۰
۳۲۹	n-هگزان	۲۹۸/۱۵	۱۵۳/۰	۱/۶۶۰
۳۳۰	n-هگزان	۲۹۸/۱۵	۲۰۰/۱	۱/۳۲۰
۳۳۱	n-هگزان	۲۹۸/۱۵	۲۴۹/۵	۱/۰۷۰
۳۳۲	n-هگزان	۲۹۸/۱۵	۲۹۰/۸	۹/۸۰۰
۳۳۳	n-هگزان	۲۹۸/۱۵	۲۹۵/۸	۸/۹۰۰
۳۳۴	n-هگزان	۲۹۸/۱۵	۲۹۹/۱	۸/۶۰۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۳۳۵	n-هگزان	۲۹۸/۱۵	۳۵۳/۵	۷/۰۰۰
۳۳۶	n-هگزان	۲۹۸/۲	۰/۱	۴/۲۳۰
۳۳۷	n-هگزان	۲۹۸/۲	۱۰/۰	۳/۸۷۰
۳۳۸	n-هگزان	۲۹۸/۲	۲۰/۰	۳/۵۲۰
۳۳۹	n-هگزان	۲۹۸/۲	۳۰/۰	۳/۲۶۰
۳۴۰	n-هگزان	۲۹۸/۲	۴۰/۰	۳/۰۱۰
۳۴۱	n-هگزان	۲۹۸/۲	۵۰/۰	۲/۸۰۰
۳۴۲	n-هگزان	۲۹۸/۲	۶۰/۰	۲/۵۹۰
۳۴۳	n-هگزان	۳۰۳/۲	۳۰/۰	۳/۴۰۰
۳۴۴	n-هگزان	۳۰۳/۲	۴۰/۰	۳/۲۰۰
۳۴۵	n-هگزان	۳۰۳/۲	۵۰/۰	۲/۸۰۰
۳۴۶	n-هگزان	۳۱۳/۰	۵/۰	۴/۸۳۰
۳۴۷	n-هگزان	۳۱۳/۰	۵۰/۰	۳/۳۷۰
۳۴۸	n-هگزان	۳۱۳/۰	۱۰۰/۰	۲/۴۵۰
۳۴۹	n-هگزان	۳۱۳/۰	۱۵۰/۰	۱/۸۸۰
۳۵۰	n-هگزان	۳۱۳/۰	۲۰۰/۰	۱/۵۷۰
۳۵۱	n-هگزان	۳۳۳/۰	۳۰/۰	۴/۵۰۰
۳۵۲	n-هگزان	۳۳۳/۰	۴۰/۰	۴/۲۰۰
۳۵۳	n-هگزان	۳۳۳/۰	۵۰/۰	۴/۱۰۰
۳۵۴	n-هگزان	۳۳۳/۱۵	۰/۱	۵/۹۷۰
۳۵۵	n-هگزان	۳۳۳/۱۵	۲۱/۹	۴/۸۵۰
۳۵۶	n-هگزان	۳۳۳/۱۵	۴۲/۰	۴/۱۹۰
۳۵۷	n-هگزان	۳۳۳/۱۵	۶۹/۸	۳/۵۶۰
۳۵۸	n-هگزان	۳۳۳/۱۵	۹۹/۱	۳/۰۹۰
۳۵۹	n-هگزان	۳۳۳/۱۵	۹۹/۳	۳/۰۲۰
۳۶۰	n-هگزان	۳۳۳/۱۵	۱۴۸/۳	۲/۴۳۰
۳۶۱	n-هگزان	۳۳۳/۱۵	۲۰۳/۹	۱/۹۱۰
۳۶۲	n-هگزان	۳۳۳/۱۵	۲۴۹/۰	۱/۶۲۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۳۶۳	n-هگزان	۳۳۳/۱۵	۲۹۶/۲	۱/۳۵۰
۳۶۴	n-هگزان	۳۳۳/۱۵	۳۴۷/۶	۱/۱۴۰
۳۶۵	n-هگزان	۳۳۳/۱۵	۳۹۲/۴	۰/۹۶۰
۳۶۶	n-هگزان	۳۵۹/۰	۲۰۰/۰	۲/۴۸۰
۳۶۷	n-هگزان	۳۷۰/۰	۵/۰	۸/۲۳۰
۳۶۸	n-هگزان	۳۷۰/۰	۵۰/۰	۵/۴۳۰
۳۶۹	n-هگزان	۳۷۰/۰	۱۰۰/۰	۴/۵۹۰
۳۷۰	n-هگزان	۳۷۰/۰	۱۵۰/۰	۳/۳۶۰
۳۷۱	n-هگزان	۴۴۳/۰	۵/۰	۱۴/۶۰۰
۳۷۲	n-هگزان	۴۴۳/۰	۵۰/۰	۸/۶۹۰
۳۷۳	n-هگزان	۴۴۳/۰	۱۰۰/۰	۶/۹۵۰
۳۷۴	n-هگزان	۴۴۳/۰	۱۵۰/۰	۵/۳۳۰
۳۷۵	n-هگزان	۴۵۵/۰	۲۰۰/۰	۴/۷۵۰
۳۷۶	n-هپتان	۱۹۵/۵	۰/۱	۰/۳۹۷
۳۷۷	n-هپتان	۲۵۰/۰	۰/۱	۱/۴۶۰
۳۷۸	n-هپتان	۲۹۹/۷	۰/۱	۳/۰۴۰
۳۷۹	n-هپتان	۲۹۳/۰	۰/۱	۲/۸۵۰
۳۸۰	n-هپتان	۲۹۳/۰	۱۹/۶	۲/۴۵۰
۳۸۱	n-هپتان	۲۹۳/۰	۳۹/۲	۲/۱۱۰
۳۸۲	n-هپتان	۲۹۳/۰	۵۸/۸	۱/۸۴۰
۳۸۳	n-هپتان	۲۹۳/۰	۷۸/۴	۱/۶۲۰
۳۸۴	n-هپتان	۲۹۳/۰	۹۸/۱	۱/۴۳۰
۳۸۵	n-هپتان	۲۹۳/۰	۱۲۸/۰	۱/۲۰۰
۳۸۶	n-هپتان	۲۹۳/۰	۱۵۷/۰	۱/۰۳۰
۳۸۷	n-هپتان	۲۹۳/۰	۱۸۶/۰	۰/۸۸۰
۳۸۸	n-هپتان	۲۹۳/۰	۲۱۶/۰	۰/۷۷۰
۳۸۹	n-هپتان	۲۹۳/۰	۲۴۵/۰	۰/۶۷۰
۳۹۰	n-هپتان	۳۱۳/۰	۰/۱	۳/۶۵۰

ادامه ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۳۹۱	n-هپتان	۳۱۳/۰	۱۹/۶	۳/۱۰۰
۳۹۲	n-هپتان	۳۱۳/۰	۳۹/۲	۲/۶۵۰
۳۹۳	n-هپتان	۳۱۳/۰	۵۸/۸	۲/۳۱۰
۳۹۴	n-هپتان	۳۱۳/۰	۷۸/۴	۲/۰۵۰
۳۹۵	n-هپتان	۳۱۳/۰	۹۸/۱	۱/۸۳۰
۳۹۶	n-هپتان	۳۱۳/۰	۱۲۸/۰	۱/۵۵۰
۳۹۷	n-هپتان	۳۱۳/۰	۱۵۷/۰	۱/۳۴۰
۳۹۸	n-هپتان	۳۱۳/۰	۱۸۶/۰	۱/۱۷۰
۳۹۹	n-هپتان	۳۱۳/۰	۲۱۶/۰	۱/۰۲۰
۴۰۰	n-هپتان	۳۱۳/۰	۲۴۵/۰	۰/۸۹۰
۴۰۱	n-هپتان	۳۲۳/۰	۰/۱	۴/۱۰۰
۴۰۲	n-هپتان	۳۲۳/۰	۱۹/۶	۳/۴۷۰
۴۰۳	n-هپتان	۳۲۳/۰	۳۹/۲	۲/۹۸۰
۴۰۴	n-هپتان	۳۲۳/۰	۵۸/۸	۲/۶۰۰
۴۰۵	n-هپتان	۳۲۳/۰	۷۸/۴	۲/۳۰۰
۴۰۶	n-هپتان	۳۲۳/۰	۹۸/۱	۲/۰۴۰
۴۰۷	n-هپتان	۳۲۳/۰	۱۲۸/۰	۱/۷۵۰
۴۰۸	n-هپتان	۳۲۳/۰	۱۵۷/۰	۱/۵۲۰
۴۰۹	n-هپتان	۳۲۳/۰	۱۸۶/۰	۱/۳۴۰
۴۱۰	n-هپتان	۳۲۳/۰	۲۱۶/۰	۱/۱۹۰
۴۱۱	n-هپتان	۳۲۳/۰	۲۴۵/۰	۱/۰۷۰
۴۱۲	n-هپتان	۳۳۳/۰	۰/۱	۴/۶۱۰
۴۱۳	n-هپتان	۳۳۳/۰	۱۹/۶	۳/۸۲۰
۴۱۴	n-هپتان	۳۳۳/۰	۳۹/۲	۳/۲۷۰
۴۱۵	n-هپتان	۳۳۳/۰	۵۸/۸	۲/۵۳۰
۴۱۶	n-هپتان	۳۳۳/۰	۷۸/۴	۲/۳۵۰
۴۱۷	n-هپتان	۳۳۳/۰	۹۸/۱	۲/۲۶۰
۴۱۸	n-هپتان	۳۳۳/۰	۱۲۸/۰	۱/۹۴۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۴۱۹	n-هپتان	۳۳۳/۰	۱۵۷/۰	۱/۷۰۰
۴۲۰	n-هپتان	۳۳۳/۰	۱۸۶/۰	۱/۵۰۰
۴۲۱	n-هپتان	۳۳۳/۰	۲۱۶/۰	۱/۳۳۰
۴۲۲	n-هپتان	۳۳۳/۰	۲۴۵/۰	۱/۱۸۰
۴۲۳	n-هپتان	۳۵۳/۰	۰/۱	۵/۶۰۰
۴۲۴	n-هپتان	۳۵۳/۰	۱۹/۶	۴/۶۲۰
۴۲۵	n-هپتان	۳۵۳/۰	۳۹/۲	۳/۹۵۰
۴۲۶	n-هپتان	۳۵۳/۰	۵۸/۸	۳/۴۵۰
۴۲۷	n-هپتان	۳۵۳/۰	۷۸/۴	۳/۰۵۰
۴۲۸	n-هپتان	۳۵۳/۰	۹۸/۱	۲/۷۵۰
۴۲۹	n-هپتان	۳۵۳/۰	۱۲۸/۰	۲/۳۸۰
۴۳۰	n-هپتان	۳۵۳/۰	۱۵۷/۰	۲/۰۸۰
۴۳۱	n-هپتان	۳۵۳/۰	۱۸۶/۰	۱/۸۵۰
۴۳۲	n-هپتان	۳۵۳/۰	۲۱۶/۰	۱/۶۶۰
۴۳۳	n-هپتان	۳۵۳/۰	۲۴۵/۰	۱/۵۱۰
۴۳۴	n-اکتان	۲۹۸/۲	۰/۱	۲/۳۶۰
۴۳۵	n-اکتان	۲۹۸/۲	۱۰/۰	۲/۰۸۰
۴۳۶	n-اکتان	۲۹۸/۲	۲۰/۰	۱/۸۹۰
۴۳۷	n-اکتان	۲۹۸/۲	۳۰/۰	۱/۷۳۰
۴۳۸	n-اکتان	۲۹۸/۲	۴۰/۰	۱/۶۲۰
۴۳۹	n-اکتان	۲۹۸/۲	۵۰/۰	۱/۴۹۰
۴۴۰	n-اکتان	۲۹۸/۲	۶۰/۰	۱/۳۸۰
۴۴۱	n-اکتان	۳۰۳/۳	۳۰/۰	۱/۹۰۰
۴۴۲	n-اکتان	۳۰۳/۳	۴۰/۰	۱/۷۵۰
۴۴۳	n-اکتان	۳۰۳/۳	۵۰/۰	۱/۶۷۰
۴۴۴	n-اکتان	۳۳۳/۰	۳۰/۰	۲/۸۰۰
۴۴۵	n-اکتان	۳۳۳/۰	۴۰/۰	۲/۶۰۰
۴۴۶	n-اکتان	۳۳۳/۰	۵۰/۰	۲/۴۰۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۴۴۷	n-اکتان	۲۴۸/۲۵	۰/۱	۰/۹۸۵
۴۴۸	n-اکتان	۲۴۸/۱۱	۰/۱	۰/۹۷۸
۴۴۹	n-اکتان	۲۴۸/۱۵	۲۷/۸	۰/۷۵۸
۴۵۰	n-اکتان	۲۴۸/۱۴	۵۰/۲	۰/۶۴۴
۴۵۱	n-اکتان	۲۴۸/۱۳	۷۶/۴	۰/۵۰۹
۴۵۲	n-اکتان	۲۴۸/۲۳	۹۹/۷	۰/۴۲۴
۴۵۳	n-اکتان	۲۴۸/۱۰	۱۲۴/۴	۰/۳۲۷
۴۵۴	n-اکتان	۲۷۳/۴۳	۰/۱	۱/۶۸۰
۴۵۵	n-اکتان	۲۷۲/۹۸	۰/۱	۱/۶۸۰
۴۵۶	n-اکتان	۲۷۳/۹	۲۵/۰	۱/۳۳۰
۴۵۷	n-اکتان	۲۷۳/۸	۵۰/۱	۱/۰۸۸
۴۵۸	n-اکتان	۲۷۳/۷	۷۵/۳	۰/۸۹۶
۴۵۹	n-اکتان	۲۷۳/۷	۹۹/۸	۰/۷۴۸
۴۶۰	n-اکتان	۲۷۳/۹	۱۴۷/۵	۰/۵۳۷
۴۶۱	n-اکتان	۲۷۲/۹۴	۱۵۰/۵	۰/۵۱۹
۴۶۲	n-اکتان	۲۸۷/۶۷	۰/۱	۲/۰۸۶
۴۶۳	n-اکتان	۲۹۸/۶۷	۰/۱	۲/۳۴۷
۴۶۴	n-اکتان	۲۹۸/۱۵	۰/۱	۲/۳۳۶
۴۶۵	n-اکتان	۲۹۸/۱۵	۲۵/۴	۱/۹۴۲
۴۶۶	n-اکتان	۲۹۸/۱۵	۴۹/۷	۱/۶۲۱
۴۶۷	n-اکتان	۲۹۸/۱۵	۷۴/۲	۱/۳۵۵
۴۶۸	n-اکتان	۲۹۸/۱۵	۹۹/۴	۱/۱۵۹
۴۶۹	n-اکتان	۲۹۸/۱۵	۱۴۸/۵	۰/۸۶۷
۴۷۰	n-اکتان	۲۹۸/۱۵	۱۹۷/۶	۰/۶۶۱
۴۷۱	n-اکتان	۲۹۸/۱۵	۲۴۶/۰	۰/۵۱۱
۴۷۲	n-اکتان	۲۹۸/۱۵	۲۹۳/۱	۰/۴۰۸
۴۷۳	n-اکتان	۲۹۸/۱۵	۳۶۰/۸	۰/۲۸۶
۴۷۴	n-اکتان	۳۲۵/۷	۰/۱	۳/۳۱۵

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۴۷۵	n-اکتان	۳۲۳/۱۵	۰/۱	۳/۲۱۰
۴۷۶	n-اکتان	۳۲۳/۱۳	۲۴/۶	۲/۵۲۰
۴۷۷	n-اکتان	۳۲۳/۱۹	۵۲/۶	۲/۰۸۰
۴۷۸	n-اکتان	۳۲۳/۱۳	۷۶/۳	۱/۸۱۰
۴۷۹	n-اکتان	۳۲۳/۱۳	۱۰۲/۲	۱/۵۵۹
۴۸۰	n-اکتان	۳۲۳/۱۷	۱۵۰/۰	۱/۱۷۲
۴۸۱	n-اکتان	۳۲۳/۱۶	۱۹۸/۰	۰/۹۶۵
۴۸۲	n-اکتان	۳۲۳/۰۶	۲۴۴/۸	۰/۷۴۱
۴۸۳	n-اکتان	۳۴۸/۱۴	۰/۱	۴/۲۶۰
۴۸۴	n-اکتان	۳۴۸/۱۵	۷۶/۰۷	۲/۳۰۰
۴۸۵	n-اکتان	۳۴۸/۹	۱۰۳/۱۰	۱/۹۸۰
۴۸۶	n-اکتان	۳۴۸/۹	۱۴۴/۷	۱/۵۶۸
۴۸۷	n-نونان	۲۹۳/۰	۰/۱	۱/۷۰۰
۴۸۸	n-نونان	۲۹۳/۰	۱۹/۶	۱/۴۰۰
۴۸۹	n-نونان	۲۹۳/۰	۳۹/۲	۱/۱۷۰
۴۹۰	n-نونان	۲۹۳/۰	۵۸/۸	۱/۰۰۰
۴۹۱	n-نونان	۲۹۳/۰	۷۸/۴	۰/۸۶۰
۴۹۲	n-نونان	۲۹۳/۰	۹۸/۱	۰/۷۵۰
۴۹۳	n-نونان	۲۹۳/۰	۱۲۸/۰	۰/۶۲۰
۴۹۴	n-نونان	۲۹۳/۰	۱۵۷/۰	۰/۵۱۲
۴۹۵	n-نونان	۲۹۳/۰	۱۸۶/۰	۰/۴۳۵
۴۹۶	n-نونان	۲۹۳/۰	۲۱۶/۰	۰/۳۷۰
۴۹۷	n-نونان	۲۹۳/۰	۲۴۵/۰	۰/۳۲۰
۴۹۸	n-نونان	۳۱۳/۰	۰/۱	۲/۲۲۰
۴۹۹	n-نونان	۳۱۳/۰	۱۹/۶	۱/۸۳۰
۵۰۰	n-نونان	۳۱۳/۰	۳۹/۲	۱/۵۵۰
۵۰۱	n-نونان	۳۱۳/۰	۵۸/۸	۱/۳۳۰
۵۰۲	n-نونان	۳۱۳/۰	۷۸/۴	۱/۱۶۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۵۰۳	n-نونان	۳۱۳/۰	۹۸/۱	۱/۰۲۰
۵۰۴	n-نونان	۳۱۳/۰	۱۲۸/۰	۰/۸۶۰
۵۰۵	n-نونان	۳۱۳/۰	۱۵۷/۰	۰/۷۳۰
۵۰۶	n-نونان	۳۱۳/۰	۱۸۶/۰	۰/۶۲۰
۵۰۷	n-نونان	۳۱۳/۰	۲۱۶/۰	۰/۵۳۰
۵۰۸	n-نونان	۳۱۳/۰	۲۴۵/۰	۰/۴۶۰
۵۰۹	n-نونان	۳۲۳/۰	۰/۱	۲/۵۳۰
۵۱۰	n-نونان	۳۲۳/۰	۱۹/۶	۲/۱۲۰
۵۱۱	n-نونان	۳۲۳/۰	۳۹/۲	۱/۸۰۰
۵۱۲	n-نونان	۳۲۳/۰	۵۸/۸	۱/۵۵۰
۵۱۳	n-نونان	۳۲۳/۰	۷۸/۴	۱/۳۶۰
۵۱۴	n-نونان	۳۲۳/۰	۹۸/۱	۱/۱۹۰
۵۱۵	n-نونان	۳۲۳/۰	۱۲۸/۰	۱/۰۰۰
۵۱۶	n-نونان	۳۲۳/۰	۱۵۷/۰	۰/۸۵۰
۵۱۷	n-نونان	۳۲۳/۰	۱۸۶/۰	۰/۷۳۰
۵۱۸	n-نونان	۳۲۳/۰	۲۱۶/۰	۰/۶۴۰
۵۱۹	n-نونان	۳۲۳/۰	۲۴۵/۰	۰/۵۶۰
۵۲۰	n-نونان	۳۳۳/۰	۰/۱	۲/۹۰۰
۵۲۱	n-نونان	۳۳۳/۰	۱۹/۶	۲/۴۲۰
۵۲۲	n-نونان	۳۳۳/۰	۳۹/۲	۲/۰۵۰
۵۲۳	n-نونان	۳۳۳/۰	۵۸/۸	۱/۷۷۰
۵۲۴	n-نونان	۳۳۳/۰	۷۸/۴	۱/۵۴۰
۵۲۵	n-نونان	۳۳۳/۰	۹۸/۱	۱/۳۶۰
۵۲۶	n-نونان	۳۳۳/۰	۱۲۸/۰	۱/۱۴۰
۵۲۷	n-نونان	۳۳۳/۰	۱۵۷/۰	۰/۹۷۰
۵۲۸	n-نونان	۳۳۳/۰	۱۸۶/۰	۰/۸۴۰
۵۲۹	n-نونان	۳۳۳/۰	۲۱۶/۰	۰/۷۴۰
۵۳۰	n-نونان	۳۳۳/۰	۲۴۵/۰	۰/۶۵۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۵۳۱	n-نونان	۳۵۳/۰	۰/۱	۳/۵۵۰
۵۳۲	n-نونان	۳۵۳/۰	۱۹/۶	۳/۰۵۰
۵۳۳	n-نونان	۳۵۳/۰	۳۹/۲	۲/۶۳۰
۵۳۴	n-نونان	۳۵۳/۰	۵۸/۸	۲/۲۹۰
۵۳۵	n-نونان	۳۵۳/۰	۷۸/۴	۲/۰۰۰
۵۳۶	n-نونان	۳۵۳/۰	۹۸/۱	۱/۷۷۰
۵۳۷	n-نونان	۳۵۳/۰	۱۲۸/۰	۱/۴۹۰
۵۳۸	n-نونان	۳۵۳/۰	۱۵۷/۰	۱/۲۷۰
۵۳۹	n-نونان	۳۵۳/۰	۱۸۶/۰	۱/۱۰۰
۵۴۰	n-نونان	۳۵۳/۰	۲۱۶/۰	۰/۹۸۰
۵۴۱	n-نونان	۳۵۳/۰	۲۴۵/۰	۰/۸۸۰
۵۴۲	n-دکان	۲۶۴/۰	۵/۰	۰/۶۸۶
۵۴۳	n-دکان	۲۶۴/۰	۵۰/۰	۰/۴۳۶
۵۴۴	n-دکان	۲۶۴/۰	۷۵/۰	۰/۳۴۵
۵۴۵	n-دکان	۲۶۴/۰	۱۰۰/۰	۰/۲۷۳
۵۴۶	n-دکان	۲۸۶/۰	۵/۰	۱/۰۳۰
۵۴۷	n-دکان	۲۸۶/۰	۵۰/۰	۰/۶۷۷
۵۴۸	n-دکان	۲۸۶/۰	۱۰۰/۰	۰/۴۵۶
۵۴۹	n-دکان	۲۸۶/۰	۱۵۰/۰	۰/۳۰۳
۵۵۰	n-دکان	۲۸۶/۰	۲۰۰/۰	۰/۲۱۹
۵۵۱	n-دکان	۲۹۸/۲	۱۰/۰	۱/۲۳۵
۵۵۲	n-دکان	۲۹۸/۲	۲۰/۰	۱/۱۳۰
۵۵۳	n-دکان	۲۹۸/۲	۳۰/۰	۱/۰۲۰
۵۵۴	n-دکان	۲۹۸/۲	۴۰/۰	۰/۹۴۹
۵۵۵	n-دکان	۲۹۸/۲	۵۰/۰	۰/۸۶۱
۵۵۶	n-دکان	۲۹۸/۲	۶۰/۰	۰/۸۱۲
۵۵۷	n-دکان	۳۰۳/۳	۳۰/۰	۱/۱۰۰
۵۵۸	n-دکان	۳۰۳/۳	۴۰/۰	۱/۰۳۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۵۵۹	n-دکان	۳۰۳/۳	۵۰/۰	۰/۹۵۰
۵۶۰	n-دکان	۳۱۴/۰	۵/۰	۱/۶۰۰
۵۶۱	n-دکان	۳۱۴/۰	۵۰/۰	۱/۱۳۰
۵۶۲	n-دکان	۳۱۴/۰	۱۰۰/۰	۰/۷۷۹
۵۶۳	n-دکان	۳۱۴/۰	۱۵۰/۰	۰/۵۹۶
۵۶۴	n-دکان	۳۱۴/۰	۲۰۰/۰	۰/۴۰۶
۵۶۵	n-دکان	۳۳۳/۰	۳۰/۰	۱/۷۲۰
۵۶۶	n-دکان	۳۳۳/۰	۴۰/۰	۱/۵۷۰
۵۶۷	n-دکان	۳۳۳/۰	۵۰/۰	۱/۴۵۰
۵۶۸	n-دکان	۳۴۸/۰	۵/۰	۲/۵۰۰
۵۶۹	n-دکان	۳۴۸/۰	۵۰/۰	۱/۸۰۰
۵۷۰	n-دکان	۳۴۸/۰	۱۰۰/۰	۱/۲۵۰
۵۷۱	n-دکان	۳۴۸/۰	۱۵۰/۰	۰/۹۶۰
۵۷۲	n-دکان	۳۴۸/۰	۲۰۰/۰	۰/۷۸۵
۵۷۳	n-دکان	۳۹۲/۰	۵/۰	۳/۹۴۰
۵۷۴	n-دکان	۳۹۲/۰	۵۰/۰	۲/۷۸۰
۵۷۵	n-دکان	۳۹۲/۰	۱۰۰/۰	۲/۰۷۰
۵۷۶	n-دکان	۳۹۲/۰	۱۵۰/۰	۱/۵۶۰
۵۷۷	n-دکان	۳۹۲/۰	۲۰۰/۰	۱/۲۲۰
۵۷۸	n-دکان	۴۴۸/۰	۵/۰	۶/۲۴۰
۵۷۹	n-دکان	۴۴۸/۰	۵۰/۰	۴/۲۰۰
۵۸۰	n-دکان	۴۴۸/۰	۱۰۰/۰	۳/۱۴۰
۵۸۱	n-دکان	۴۴۸/۰	۱۵۰/۰	۲/۴۶۰
۵۸۲	n-دکان	۴۴۸/۰	۲۰۰/۰	۲/۰۹۰
۵۸۳	n-دکان	۴۹۷/۵	۶۳/۴	۵/۱۰۰
۵۸۴	n-دکان	۴۹۷/۵	۲۴۰/۰	۲/۶۰۰
۵۸۵	n-آن دکان	۲۹۳/۰	۰/۱	۰/۹۸۰
۵۸۶	n-آن دکان	۲۹۳/۰	۱۹/۶	۰/۷۸۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۵۸۷	n-آن دکان	۲۹۳/۰	۳۹/۲	۰/۶۵۰
۵۸۸	n-آن دکان	۲۹۳/۰	۵۸/۸	۰/۵۵۲
۵۸۹	n-آن دکان	۲۹۳/۰	۷۸/۴	۰/۴۶۵
۵۹۰	n-آن دکان	۲۹۳/۰	۹۸/۱	۰/۴۲۰
۵۹۱	n-آن دکان	۲۹۳/۰	۱۲۸/۰	۰/۳۳۳
۵۹۲	n-آن دکان	۲۹۳/۰	۱۵۷/۰	۰/۲۷۰
۵۹۳	n-آن دکان	۲۹۳/۰	۱۸۶/۰	۰/۲۱۸
۵۹۴	n-آن دکان	۲۹۳/۰	۲۱۶/۰	۰/۱۸۲
۵۹۵	n-آن دکان	۲۹۳/۰	۲۴۵/۰	۰/۱۵۲
۵۹۶	n-آن دکان	۳۱۳/۰	۰/۱	۱/۳۸۰
۵۹۷	n-آن دکان	۳۱۳/۰	۱۹/۶	۱/۱۱۰
۵۹۸	n-آن دکان	۳۱۳/۰	۳۹/۲	۰/۹۲۰
۵۹۹	n-آن دکان	۳۱۳/۰	۵۸/۸	۰/۷۹۰
۶۰۰	n-آن دکان	۳۱۳/۰	۷۸/۴	۰/۶۸۰
۶۰۱	n-آن دکان	۳۱۳/۰	۹۸/۱	۰/۶۰۰
۶۰۲	n-آن دکان	۳۱۳/۰	۱۲۸/۰	۰/۴۸۵
۶۰۳	n-آن دکان	۳۱۳/۰	۱۵۷/۰	۰/۴۰۰
۶۰۴	n-آن دکان	۳۱۳/۰	۱۸۶/۰	۰/۳۴۰
۶۰۵	n-آن دکان	۳۱۳/۰	۲۱۶/۰	۰/۲۸۰
۶۰۶	n-آن دکان	۳۱۳/۰	۲۴۵/۰	۰/۲۲۸
۶۰۷	n-آن دکان	۳۲۳/۰	۰/۱	۱/۶۰۰
۶۰۸	n-آن دکان	۳۲۳/۰	۱۹/۶	۱/۲۷۰
۶۰۹	n-آن دکان	۳۲۳/۰	۳۹/۲	۱/۰۸۰
۶۱۰	n-آن دکان	۳۲۳/۰	۵۸/۸	۰/۹۲۰
۶۱۱	n-آن دکان	۳۲۳/۰	۷۸/۴	۰/۷۹۵
۶۱۲	n-آن دکان	۳۲۳/۰	۹۸/۱	۰/۷۰۰
۶۱۳	n-آن دکان	۳۲۳/۰	۱۲۸/۰	۰/۵۰۲
۶۱۴	n-آن دکان	۳۲۳/۰	۱۵۷/۰	۰/۴۸۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۶۱۵	n-آن دکان	۳۲۳/۰	۱۸۶/۰	۰/۴۰۰
۶۱۶	n-آن دکان	۳۲۳/۰	۲۱۶/۰	۰/۳۴۰
۶۱۷	n-آن دکان	۳۲۳/۰	۲۴۵/۰	۰/۲۹۰
۶۱۸	n-آن دکان	۳۳۳/۰	۰/۱	۱/۸۶۰
۶۱۹	n-آن دکان	۳۳۳/۰	۱۹/۶	۱/۵۰۰
۶۲۰	n-آن دکان	۳۳۳/۰	۳۹/۲	۱/۲۵۰
۶۲۱	n-آن دکان	۳۳۳/۰	۵۸/۸	۱/۰۸۰
۶۲۲	n-آن دکان	۳۳۳/۰	۷۸/۴	۰/۹۳۵
۶۲۳	n-آن دکان	۳۳۳/۰	۹۸/۱	۰/۸۲۰
۶۲۴	n-آن دکان	۳۳۳/۰	۱۲۸/۰	۰/۶۷۵
۶۲۵	n-آن دکان	۳۳۳/۰	۱۵۷/۰	۰/۵۷۰
۶۲۶	n-آن دکان	۳۳۳/۰	۱۸۶/۰	۰/۴۹۰
۶۲۷	n-آن دکان	۳۳۳/۰	۲۱۶/۰	۰/۴۱۲
۶۲۸	n-آن دکان	۳۳۳/۰	۲۴۵/۰	۰/۳۶۵
۶۲۹	n-آن دکان	۳۵۳/۰	۰/۱	۲/۴۲۰
۶۳۰	n-آن دکان	۳۵۳/۰	۱۹/۶	۱/۹۵۰
۶۳۱	n-آن دکان	۳۵۳/۰	۳۹/۲	۱/۶۵۰
۶۳۲	n-آن دکان	۳۵۳/۰	۵۸/۸	۱/۴۱۰
۶۳۳	n-آن دکان	۳۵۳/۰	۷۸/۴	۱/۲۳۰
۶۳۴	n-آن دکان	۳۵۳/۰	۹۸/۱	۱/۰۸۰
۶۳۵	n-آن دکان	۳۵۳/۰	۱۲۸/۰	۰/۸۸۵
۶۳۶	n-آن دکان	۳۵۳/۰	۱۵۷/۰	۰/۷۷۰
۶۳۷	n-آن دکان	۳۵۳/۰	۱۸۶/۰	۰/۶۶۰
۶۳۸	n-آن دکان	۳۵۳/۰	۲۱۶/۰	۰/۵۷۵
۶۳۹	n-آن دکان	۳۵۳/۰	۲۴۵/۰	۰/۵۰۵
۶۴۰	n-دو دکان	۲۹۸/۲	۰/۱	۰/۸۴۵
۶۴۱	n-دو دکان	۲۹۸/۲	۱۰/۰	۰/۷۵۴
۶۴۲	n-دو دکان	۲۹۸/۲	۲۰/۰	۰/۶۶۸

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۶۴۳	n-دو دکان	۲۹۸/۲	۳۰/۰	۰/۶۲۴
۶۴۴	n-دو دکان	۲۹۸/۲	۴۰/۰	۰/۵۵۳
۶۴۵	n-دو دکان	۲۹۸/۲	۵۰/۰	۰/۵۱۰
۶۴۶	n-دو دکان	۲۹۸/۲	۶۰/۰	۰/۴۷۴
۶۴۷	n-دو دکان	۳۰۴/۰	۱/۴	۰/۹۶۰
۶۴۸	n-دو دکان	۳۷۳/۰	۱/۴	۲/۶۱۰
۶۴۹	n-دو دکان	۴۴۳/۰	۱/۴	۵/۲۳۰
۶۵۰	n-تری دکان	۲۹۳/۰	۰/۱	۰/۶۱۰
۶۵۱	n-تری دکان	۲۹۳/۰	۱۹/۶	۰/۵۰۰
۶۵۲	n-تری دکان	۲۹۳/۰	۳۹/۲	۰/۴۱۲
۶۵۳	n-تری دکان	۲۹۳/۰	۵۸/۸	۰/۳۴۲
۶۵۴	n-تری دکان	۲۹۳/۰	۷۸/۴	۰/۲۸۵
۶۵۵	n-تری دکان	۲۹۳/۰	۹۸/۱	۰/۲۳۵
۶۵۶	n-تری دکان	۳۱۳/۰	۰/۱	۰/۹۰۰
۶۵۷	n-تری دکان	۳۱۳/۰	۱۹/۶	۰/۷۳۲
۶۵۸	n-تری دکان	۳۱۳/۰	۳۹/۲	۰/۶۱۰
۶۵۹	n-تری دکان	۳۱۳/۰	۵۸/۸	۰/۵۱۵
۶۶۰	n-تری دکان	۳۱۳/۰	۷۸/۴	۰/۴۴۰
۶۶۱	n-تری دکان	۳۱۳/۰	۹۸/۱	۰/۳۷۵
۶۶۲	n-تری دکان	۳۱۳/۰	۱۲۸/۰	۰/۳۰۴
۶۶۳	n-تری دکان	۳۱۳/۰	۱۵۷/۰	۰/۲۴۸
۶۶۴	n-تری دکان	۳۱۳/۰	۱۸۶/۰	۰/۲۰۵
۶۶۵	n-تری دکان	۳۱۳/۰	۲۱۶/۰	۰/۱۷۲
۶۶۶	n-تری دکان	۳۱۳/۰	۲۴۵/۰	۰/۱۴۰
۶۶۷	n-تری دکان	۳۲۳/۰	۰/۱	۱/۰۷۰
۶۶۸	n-تری دکان	۳۲۳/۰	۱۹/۶	۰/۸۷۶
۶۶۹	n-تری دکان	۳۲۳/۰	۳۹/۲	۰/۷۳۵
۶۷۰	n-تری دکان	۳۲۳/۰	۵۸/۸	۰/۶۲۴

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۶۷۱	n-تری دکان	۳۲۳/۰	۷۸/۴	۰/۵۳۵
۶۷۲	n-تری دکان	۳۲۳/۰	۹۸/۱	۰/۴۶۰
۶۷۳	n-تری دکان	۳۲۳/۰	۱۲۸/۰	۰/۳۷۲
۶۷۴	n-تری دکان	۳۲۳/۰	۱۵۷/۰	۰/۳۰۶
۶۷۵	n-تری دکان	۳۲۳/۰	۱۸۶/۰	۰/۲۵۵
۶۷۶	n-تری دکان	۳۲۳/۰	۲۱۶/۰	۰/۲۱۵
۶۷۷	n-تری دکان	۳۲۳/۰	۲۴۵/۰	۰/۱۸۲
۶۷۸	n-تری دکان	۳۳۳/۰	۰/۱	۱/۲۸۰
۶۷۹	n-تری دکان	۳۳۳/۰	۱۹/۶	۱/۰۴۰
۶۸۰	n-تری دکان	۳۳۳/۰	۳۹/۲	۰/۸۶۰
۶۸۱	n-تری دکان	۳۳۳/۰	۵۸/۸	۰/۷۳۲
۶۸۲	n-تری دکان	۳۳۳/۰	۷۸/۴	۰/۶۳۴
۶۸۳	n-تری دکان	۳۳۳/۰	۹۸/۱	۰/۵۵۲
۶۸۴	n-تری دکان	۳۳۳/۰	۱۲۸/۰	۰/۴۵۲
۶۸۵	n-تری دکان	۳۳۳/۰	۱۵۷/۰	۰/۳۷۵
۶۸۶	n-تری دکان	۳۳۳/۰	۱۸۶/۰	۰/۳۱۵
۶۸۷	n-تری دکان	۳۳۳/۰	۲۱۶/۰	۰/۲۶۸
۶۸۸	n-تری دکان	۳۳۳/۰	۲۴۵/۰	۰/۲۳۰
۶۸۹	n-تری دکان	۳۵۳/۰	۰/۱	۱/۷۰۰
۶۹۰	n-تری دکان	۳۵۳/۰	۱۹/۶	۱/۴۰۰
۶۹۱	n-تری دکان	۳۵۳/۰	۳۹/۲	۱/۱۹۰
۶۹۲	n-تری دکان	۳۵۳/۰	۵۸/۸	۰/۹۹۰
۶۹۳	n-تری دکان	۳۵۳/۰	۷۸/۴	۰/۸۵۶
۶۹۴	n-تری دکان	۳۵۳/۰	۹۸/۱	۰/۷۶۰
۶۹۵	n-تری دکان	۳۵۳/۰	۱۲۸/۰	۰/۶۲۸
۶۹۶	n-تری دکان	۳۵۳/۰	۱۵۷/۰	۰/۵۲۴
۶۹۷	n-تری دکان	۳۵۳/۰	۱۸۶/۰	۰/۴۴۴
۶۹۸	n-تری دکان	۳۵۳/۰	۲۱۶/۰	۰/۳۸۲

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۶۹۹	n-تری دکان	۳۵۳/۰	۲۴۵/۰	۰/۳۳۶
۷۰۰	n-تترا دکان	۲۹۳/۰	۰/۱	۰/۴۷۱
۷۰۱	n-تترا دکان	۲۹۳/۰	۵۰/۰	۰/۲۷۸
۷۰۲	n-تترا دکان	۲۹۳/۰	۵۰/۰	۰/۲۷۰
۷۰۳	n-تترا دکان	۳۱۱/۰	۰/۱	۰/۷۲۹
۷۰۴	n-تترا دکان	۳۱۱/۱	۵۰/۰	۰/۴۸۸
۷۰۵	n-تترا دکان	۳۱۱/۱	۱۰۰/۰	۰/۲۹۰
۷۰۶	n-تترا دکان	۳۲۴/۰	۰/۱	۰/۹۰۹
۷۰۷	n-تترا دکان	۳۲۴/۱	۵۰/۰	۰/۵۴۸
۷۰۸	n-تترا دکان	۳۲۴/۱	۱۰۰/۰	۰/۳۶۹
۷۰۹	n-تترا دکان	۳۲۴/۱	۱۵۰/۰	۰/۲۶۴
۷۱۰	n-تترا دکان	۳۲۴/۰	۲۰۰/۰	۰/۲۳۱
۷۱۱	n-تترا دکان	۳۵۵/۰	۰/۱	۱/۴۶۱
۷۱۲	n-تترا دکان	۳۵۵/۱	۵۰/۰	۰/۹۶۰
۷۱۳	n-تترا دکان	۳۵۵/۱	۱۰۰/۰	۰/۶۶۹
۷۱۴	n-تترا دکان	۳۵۵/۱	۱۵۰/۰	۰/۴۸۴
۷۱۵	n-تترا دکان	۳۵۵/۱	۲۰۰/۰	۰/۳۶۰
۷۱۶	n-تترا دکان	۳۹۰/۰	۰/۱	۲/۲۶۰
۷۱۷	n-تترا دکان	۳۹۰/۰	۵۰/۰	۱/۴۰۱
۷۱۸	n-تترا دکان	۳۹۰/۱	۵۰/۰	۱/۴۰۰
۷۱۹	n-تترا دکان	۳۹۰/۱	۱۰۰/۰	۱/۰۰۴
۷۲۰	n-تترا دکان	۳۹۰/۱	۱۵۰/۰	۰/۷۷۶
۷۲۱	n-تترا دکان	۳۹۰/۱	۲۰۰/۰	۰/۵۷۰
۷۲۲	n-تترا دکان	۴۴۳/۰	۰/۱	۳/۳۹۴
۷۲۳	n-تترا دکان	۴۴۳/۱	۵۰/۰	۲/۱۷۴
۷۲۴	n-تترا دکان	۴۴۳/۱	۱۰۰/۰	۱/۵۵۲
۷۲۵	n-تترا دکان	۴۴۳/۱	۱۵۰/۰	۱/۲۳۰
۷۲۶	n-تترا دکان	۴۴۳/۱	۲۰۰/۰	۱/۰۱۲

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۷۲۷	n-پنتا دکان	۲۹۳/۰	۰/۱	۰/۴۰۰
۷۲۸	n-پنتا دکان	۲۹۳/۰	۱۹/۶	۰/۳۱۳
۷۲۹	n-پنتا دکان	۲۹۳/۰	۳۹/۲	۰/۲۶۴
۷۳۰	n-پنتا دکان	۳۱۳/۰	۰/۱	۰/۶۱۰
۷۳۱	n-پنتا دکان	۳۱۳/۰	۱۹/۶	۰/۴۹۳
۷۳۲	n-پنتا دکان	۳۱۳/۰	۳۹/۲	۰/۴۰۸
۷۳۳	n-پنتا دکان	۳۱۳/۰	۵۸/۸	۰/۳۴۴
۷۳۴	n-پنتا دکان	۳۱۳/۰	۷۸/۴	۰/۲۹۲
۷۳۵	n-پنتا دکان	۳۱۳/۰	۹۸/۱	۰/۲۵۰
۷۳۶	n-پنتا دکان	۳۱۳/۰	۱۱۸/۰	۰/۲۱۵
۷۳۷	n-پنتا دکان	۳۱۳/۰	۱۳۷/۰	۰/۱۸۶
۷۳۸	n-پنتا دکان	۳۱۳/۰	۱۵۷/۰	۰/۱۶۲
۷۳۹	n-پنتا دکان	۳۱۳/۰	۱۷۷/۰	۰/۱۴۰
۷۴۰	n-پنتا دکان	۳۱۳/۰	۱۹۶/۰	۰/۱۲۲
۷۴۱	n-پنتا دکان	۳۲۳/۰	۰/۱	۰/۷۳۵
۷۴۲	n-پنتا دکان	۳۲۳/۰	۱۹/۶	۰/۶۰۰
۷۴۳	n-پنتا دکان	۳۲۳/۰	۳۹/۲	۰/۵۰۰
۷۴۴	n-پنتا دکان	۳۲۳/۰	۵۸/۸	۰/۴۲۲
۷۴۵	n-پنتا دکان	۳۲۳/۰	۷۸/۴	۰/۳۶۰
۷۴۶	n-پنتا دکان	۳۲۳/۰	۹۸/۱	۰/۳۰۹
۷۴۷	n-پنتا دکان	۳۲۳/۰	۱۱۸/۰	۰/۲۶۷
۷۴۸	n-پنتا دکان	۳۲۳/۰	۱۳۷/۰	۰/۲۳۲
۷۴۹	n-پنتا دکان	۳۲۳/۰	۱۵۷/۰	۰/۲۰۲
۷۵۰	n-پنتا دکان	۳۲۳/۰	۱۷۷/۰	۰/۱۷۷
۷۵۱	n-پنتا دکان	۳۲۳/۰	۱۹۶/۰	۰/۱۵۵
۷۵۲	n-پنتا دکان	۳۲۳/۰	۲۱۶/۰	۰/۱۳۷
۷۵۳	n-پنتا دکان	۳۳۳/۰	۰/۱	۰/۸۷۵
۷۵۴	n-پنتا دکان	۳۳۳/۰	۱۹/۶	۰/۷۰۷

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۷۵۵	n-پنتا دکان	۳۳۳/۰	۳۹/۲	۰/۵۹۰
۷۵۶	n-پنتا دکان	۳۳۳/۰	۵۸/۸	۰/۵۰۰
۷۵۷	n-پنتا دکان	۳۳۳/۰	۷۸/۴	۰/۴۳۰
۷۵۸	n-پنتا دکان	۳۳۳/۰	۹۸/۱	۰/۳۷۲
۷۵۹	n-پنتا دکان	۳۳۳/۰	۱۱۸/۰	۰/۳۲۵
۷۶۰	n-پنتا دکان	۳۳۳/۰	۱۳۷/۰	۰/۲۸۵
۷۶۱	n-پنتا دکان	۳۳۳/۰	۱۵۷/۰	۰/۲۵۲
۷۶۲	n-پنتا دکان	۳۳۳/۰	۱۷۷/۰	۰/۲۲۴
۷۶۳	n-پنتا دکان	۳۳۳/۰	۱۹۶/۰	۰/۲۰۰
۷۶۴	n-پنتا دکان	۳۳۳/۰	۲۱۶/۰	۰/۱۷۹
۷۶۵	n-پنتا دکان	۳۳۳/۰	۲۴۵/۰	۰/۱۵۴
۷۶۶	n-پنتا دکان	۳۵۳/۰	۰/۱	۱/۱۹۰
۷۶۷	n-پنتا دکان	۳۵۳/۰	۱۹/۶	۱/۰۲۰
۷۶۸	n-پنتا دکان	۳۵۳/۰	۳۹/۲	۰/۸۵۵
۷۶۹	n-پنتا دکان	۳۵۳/۰	۵۸/۸	۰/۷۲۵
۷۷۰	n-پنتا دکان	۳۵۳/۰	۷۸/۴	۰/۶۲۲
۷۷۱	n-پنتا دکان	۳۵۳/۰	۹۸/۱	۰/۵۳۸
۷۷۲	n-پنتا دکان	۳۵۳/۰	۱۱۸/۰	۰/۴۷۰
۷۷۳	n-پنتا دکان	۳۵۳/۰	۱۳۷/۰	۰/۴۱۳
۷۷۴	n-پنتا دکان	۳۵۳/۰	۱۵۷/۰	۰/۳۶۶
۷۷۵	n-پنتا دکان	۳۵۳/۰	۱۷۷/۰	۰/۳۲۵
۷۷۶	n-پنتا دکان	۳۵۳/۰	۱۹۶/۰	۰/۲۹۴
۷۷۷	n-پنتا دکان	۳۵۳/۰	۲۱۶/۰	۰/۲۶۶
۷۷۸	n-پنتا دکان	۳۵۳/۰	۲۴۵/۰	۰/۲۳۴
۷۷۹	n-هگزا دکان	۲۹۸/۱۵	۰/۱	۰/۳۷۸
۷۸۰	n-هگزا دکان	۲۹۸/۳۰	۰/۱	۰/۳۸۰
۷۸۱	n-هگزا دکان	۲۹۸/۳۲	۹/۹	۰/۳۳۱
۷۸۲	n-هگزا دکان	۲۹۸/۳۳	۱۸/۸	۰/۳۰۶

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۷۸۳	n-هگزا دکان	۲۹۸/۳۳	۲۷/۳	۰/۲۷۵
۷۸۴	n-هگزا دکان	۲۹۸/۲	۰/۱	۰/۳۸۸
۷۸۵	n-هگزا دکان	۳۱۳/۵	۰/۱	۰/۵۰۰
۷۸۶	n-هگزا دکان	۳۱۳/۵	۲۵/۰	۰/۳۸۰
۷۸۷	n-هگزا دکان	۳۱۳/۵	۵۰/۰	۰/۳۰۰
۷۸۸	n-هگزا دکان	۳۱۳/۵	۱۰۰/۰	۰/۲۰۰
۷۸۹	n-هگزا دکان	۳۲۳/۰	۱/۴	۰/۶۸۰
۷۹۰	n-هگزا دکان	۳۲۳/۱۵	۰/۱	۰/۶۳۲
۷۹۱	n-هگزا دکان	۳۲۳/۱۵	۲۳/۴	۰/۵۰۰
۷۹۲	n-هگزا دکان	۳۲۳/۱۵	۴۸/۷	۰/۳۹۲
۷۹۳	n-هگزا دکان	۳۲۳/۱۵	۵۰/۹	۰/۳۸۲
۷۹۴	n-هگزا دکان	۳۲۳/۱۵	۷۷/۲	۰/۳۱۱
۷۹۵	n-هگزا دکان	۳۲۳/۱۵	۹۹/۶	۰/۲۵۲
۷۹۶	n-هگزا دکان	۳۲۳/۱۵	۱۱۹/۳	۰/۲۲۱
۷۹۷	n-هگزا دکان	۳۲۲/۹۱	۱۴۱/۷	۰/۱۸۹
۷۹۸	n-هگزا دکان	۳۴۰/۵	۰/۱	۰/۸۲۰
۷۹۹	n-هگزا دکان	۳۴۰/۵	۲۵/۰	۰/۶۴۰
۸۰۰	n-هگزا دکان	۳۴۰/۵	۵۰/۰	۰/۵۱۰
۸۰۱	n-هگزا دکان	۳۴۰/۵	۱۰۰/۰	۰/۳۴۰
۸۰۲	n-هگزا دکان	۳۴۰/۵	۱۵۰/۰	۰/۲۵۰
۸۰۳	n-هگزا دکان	۳۴۰/۵	۲۰۰/۰	۰/۱۹۰
۸۰۴	n-هگزا دکان	۳۴۰/۵	۲۵۰/۰	۰/۱۵۰
۸۰۵	n-هگزا دکان	۳۴۸/۱۵	۰/۱	۱/۰۰۰
۸۰۶	n-هگزا دکان	۳۴۸/۱۵	۲۵/۲	۰/۷۸۷
۸۰۷	n-هگزا دکان	۳۴۸/۱۵	۵۲/۷	۰/۶۱۰
۸۰۸	n-هگزا دکان	۳۴۸/۱۵	۷۴/۹	۰/۵۱۲
۸۰۹	n-هگزا دکان	۳۴۸/۱۵	۹۹/۷	۰/۴۲۲
۸۱۰	n-هگزا دکان	۳۴۸/۱۵	۱۰۰/۴	۰/۴۳۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۸۱۱	n-هگزا دکان	۳۴۸/۱۵	۱۴۸/۳	۰/۳۱۳
۸۱۲	n-هگزا دکان	۳۴۸/۱۵	۱۹۹/۰	۰/۲۲۷
۸۱۳	n-هگزا دکان	۳۴۸/۱۵	۲۴۵/۸	۰/۱۷۳
۸۱۴	n-هگزا دکان	۳۴۸/۱۵	۲۷۹/۲	۰/۱۴۲
۸۱۵	n-هگزا دکان	۳۷۱/۰	۱/۴	۱/۴۳۰
۸۱۶	n-هگزا دکان	۳۷۱/۵	۰/۱	۱/۳۰۰
۸۱۷	n-هگزا دکان	۳۷۱/۵	۲۵/۰	۱/۰۰۰
۸۱۸	n-هگزا دکان	۳۷۱/۵	۵۰/۰	۰/۸۱۰
۸۱۹	n-هگزا دکان	۳۷۱/۵	۱۰۰/۰	۰/۵۶۰
۸۲۰	n-هگزا دکان	۳۷۱/۵	۱۵۰/۰	۰/۴۱۰
۸۲۱	n-هگزا دکان	۳۷۱/۵	۲۰۰/۰	۰/۳۱۰
۸۲۲	n-هگزا دکان	۳۷۱/۵	۲۵۰/۰	۰/۲۵۰
۸۲۳	n-هگزا دکان	۳۷۱/۵	۳۰۰/۰	۰/۲۰۰
۸۲۴	n-هگزا دکان	۳۷۱/۵	۳۵۰/۰	۰/۱۷۰
۸۲۵	n-هگزا دکان	۴۰۴/۵	۰/۱	۱/۹۵۰
۸۲۶	n-هگزا دکان	۴۰۴/۵	۲۵/۰	۱/۵۲۰
۸۲۷	n-هگزا دکان	۴۰۴/۵	۵۰/۰	۱/۲۳۰
۸۲۸	n-هگزا دکان	۴۰۴/۵	۱۰۰/۰	۰/۸۵۰
۸۲۹	n-هگزا دکان	۴۰۴/۵	۱۵۰/۰	۰/۶۳۰
۸۳۰	n-هگزا دکان	۴۰۴/۵	۲۰۰/۰	۰/۴۹۰
۸۳۱	n-هگزا دکان	۴۰۴/۵	۲۵۰/۰	۰/۳۹۰
۸۳۲	n-هگزا دکان	۴۰۴/۵	۳۰۰/۰	۰/۳۲۰
۸۳۳	n-هگزا دکان	۴۰۴/۵	۳۵۰/۰	۰/۲۷۰
۸۳۴	n-هگزا دکان	۴۰۴/۵	۴۰۰/۰	۰/۲۲۰
۸۳۵	n-هگزا دکان	۴۰۴/۵	۴۵۰/۰	۰/۱۹۰
۸۳۶	n-هگزا دکان	۴۰۴/۵	۵۰۰/۰	۰/۱۶۰
۸۳۷	n-هگزا دکان	۴۰۴/۵	۵۵۰/۰	۰/۱۴۰
۸۳۸	n-هگزا دکان	۴۰۴/۵	۶۰۰/۰	۰/۱۳۰

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۸۳۹	n-هگزا دکان	۴۴۳/۰	۱/۴	۳/۲۳۰
۸۴۰	n-هگزا دکان	۴۴۴/۰	۰/۱	۳/۰۰۰
۸۴۱	n-هگزا دکان	۴۴۴/۰	۲۵/۰	۲/۳۵۰
۸۴۲	n-هگزا دکان	۴۴۴/۰	۵۰/۰	۱/۹۵۰
۸۴۳	n-هگزا دکان	۴۴۴/۰	۱۰۰/۰	۱/۴۲۰
۸۴۴	n-هگزا دکان	۴۴۴/۰	۱۵۰/۰	۱/۰۷۰
۸۴۵	n-هگزا دکان	۴۴۴/۰	۲۰۰/۰	۰/۸۴۰
۸۴۶	n-هگزا دکان	۴۴۴/۰	۲۵۰/۰	۰/۶۷۰
۸۴۷	n-هگزا دکان	۴۴۴/۰	۳۰۰/۰	۰/۵۵۰
۸۴۸	n-هگزا دکان	۴۴۴/۰	۳۵۰/۰	۰/۴۶۰
۸۴۹	n-هگزا دکان	۴۴۴/۰	۴۰۰/۰	۰/۳۹۰
۸۵۰	n-هگزا دکان	۴۴۴/۰	۴۵۰/۰	۰/۳۳۰
۸۵۱	n-هگزا دکان	۴۴۴/۰	۵۰۰/۰	۰/۲۸۰
۸۵۲	n-هگزا دکان	۴۴۴/۰	۵۵۰/۰	۰/۲۵۰
۸۵۳	n-هگزا دکان	۴۴۴/۰	۶۰۰/۰	۰/۲۲۰
۸۵۴	n-هگزا دکان	۴۷۲/۵	۰/۱	۳/۹۰۰
۸۵۵	n-هگزا دکان	۴۷۲/۵	۲۵/۰	۳/۱۰۰
۸۵۶	n-هگزا دکان	۴۷۲/۵	۵۰/۰	۲/۵۵۰
۸۵۷	n-هگزا دکان	۴۷۲/۵	۱۰۰/۰	۱/۸۵۰
۸۵۸	n-هگزا دکان	۴۷۲/۵	۱۵۰/۰	۱/۴۰۰
۸۵۹	n-هگزا دکان	۴۷۲/۵	۲۰۰/۰	۱/۱۰۰
۸۶۰	n-هگزا دکان	۴۷۲/۵	۲۵۰/۰	۰/۹۰۰
۸۶۱	n-هگزا دکان	۴۷۲/۵	۳۰۰/۰	۰/۷۴۰
۸۶۲	n-هگزا دکان	۴۷۲/۵	۳۵۰/۰	۰/۶۲۰
۸۶۳	n-هگزا دکان	۴۷۲/۵	۴۰۰/۰	۰/۵۳۰
۸۶۴	n-هگزا دکان	۴۷۲/۵	۴۵۰/۰	۰/۴۶۰
۸۶۵	n-هگزا دکان	۴۷۲/۵	۵۰۰/۰	۰/۴۰۰
۸۶۶	n-هگزا دکان	۴۷۲/۵	۵۵۰/۰	۰/۳۵۰

ادامہی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۸۶۷	n-هگزا دکان	۴۷۲/۵	۶۰۰/۰	۰/۳۱۰
۸۶۸	n-هگزا دکان	۵۱۳/۰	۱/۴	۵/۷۶۰
۸۶۹	n-هپتا دکان	۳۰۳/۰	۰/۱	۰/۳۶۰
۸۷۰	n-هپتا دکان	۳۰۳/۰	۱۹/۶	۰/۲۷۹
۸۷۱	n-هپتا دکان	۳۰۳/۰	۳۹/۲	۰/۲۲۸
۸۷۲	n-هپتا دکان	۳۱۳/۰	۰/۱	۰/۴۲۵
۸۷۳	n-هپتا دکان	۳۱۳/۰	۱۹/۶	۰/۳۴۲
۸۷۴	n-هپتا دکان	۳۱۳/۰	۳۹/۲	۰/۲۸۸
۸۷۵	n-هپتا دکان	۳۱۳/۰	۵۸/۸	۰/۲۴۱
۸۷۶	n-هپتا دکان	۳۱۳/۰	۷۸/۴	۰/۲۰۵
۸۷۷	n-هپتا دکان	۳۱۳/۰	۹۸/۱	۰/۱۷۷
۸۷۸	n-هپتا دکان	۳۲۳/۰	۰/۱	۰/۵۲۲
۸۷۹	n-هپتا دکان	۳۲۳/۰	۱۹/۶	۰/۴۳۵
۸۸۰	n-هپتا دکان	۳۲۳/۰	۳۹/۲	۰/۳۵۹
۸۸۱	n-هپتا دکان	۳۲۳/۰	۵۸/۸	۰/۳۰۰
۸۸۲	n-هپتا دکان	۳۲۳/۰	۷۸/۴	۰/۲۵۲
۸۸۳	n-هپتا دکان	۳۲۳/۰	۹۸/۱	۰/۲۱۵
۸۸۴	n-هپتا دکان	۳۲۳/۰	۱۱۸/۰	۰/۱۸۵
۸۸۵	n-هپتا دکان	۳۲۳/۰	۱۳۷/۰	۰/۱۶۰
۸۸۶	n-هپتا دکان	۳۲۳/۰	۱۵۷/۰	۰/۱۴۲
۸۸۷	n-هپتا دکان	۳۳۳/۰	۰/۱	۰/۶۴۰
۸۸۸	n-هپتا دکان	۳۳۳/۰	۱۹/۶	۰/۵۲۲
۸۸۹	n-هپتا دکان	۳۳۳/۰	۳۹/۲	۰/۴۳۵
۸۹۰	n-هپتا دکان	۳۳۳/۰	۵۸/۸	۰/۳۶۷
۸۹۱	n-هپتا دکان	۳۳۳/۰	۷۸/۴	۰/۳۱۱
۸۹۲	n-هپتا دکان	۳۳۳/۰	۹۸/۱	۰/۲۶۶
۸۹۳	n-هپتا دکان	۳۳۳/۰	۱۱۸/۰	۰/۲۲۸
۸۹۴	n-هپتا دکان	۳۳۳/۰	۱۳۷/۰	۰/۱۹۸

ادامه‌ی جدول (۱-پ)

شماره	نام مولکول	دما (C°)	فشار (Mpa)	ضریب خودنفوذی ($\times 10^{-9}$)
۸۹۵	n-هپتا دکان	۳۳۳/۰	۱۵۷/۰	۰/۱۷۳
۸۹۶	n-هپتا دکان	۳۳۳/۰	۱۷۷/۰	۰/۱۵۲
۸۹۷	n-هپتا دکان	۳۳۳/۰	۱۹۶/۰	۰/۱۳۵
۸۹۸	n-هپتا دکان	۳۳۳/۰	۲۱۶/۰	۰/۱۲۲
۸۹۹	n-هپتا دکان	۳۳۳/۰	۲۴۵/۰	۰/۱۰۶
۹۰۰	n-هپتا دکان	۳۵۳/۰	۰/۱	۰/۹۲۰
۹۰۱	n-هپتا دکان	۳۵۳/۰	۱۹/۶	۰/۷۶۶
۹۰۲	n-هپتا دکان	۳۵۳/۰	۳۹/۲	۰/۶۲۸
۹۰۳	n-هپتا دکان	۳۵۳/۰	۵۸/۸	۰/۵۲۷
۹۰۴	n-هپتا دکان	۳۵۳/۰	۷۸/۴	۰/۴۵۰
۹۰۵	n-هپتا دکان	۳۵۳/۰	۹۸/۱	۰/۳۸۷
۹۰۶	n-هپتا دکان	۳۵۳/۰	۱۱۸/۰	۰/۳۳۵
۹۰۷	n-هپتا دکان	۳۵۳/۰	۱۳۷/۰	۰/۲۹۳
۹۰۸	n-هپتا دکان	۳۵۳/۰	۱۵۷/۰	۰/۲۵۸
۹۰۹	n-هپتا دکان	۳۵۳/۰	۱۷۷/۰	۰/۲۲۸
۹۱۰	n-هپتا دکان	۳۵۳/۰	۱۹۶/۰	۰/۲۰۳
۹۱۱	n-هپتا دکان	۳۵۳/۰	۲۱۶/۰	۰/۱۸۳
۹۱۲	n-هپتا دکان	۳۵۳/۰	۲۴۵/۰	۰/۱۵۷

جدول (۲-پ): مقایسه‌ی مقادیر تجربی و مقادیر پیش‌بینی شده‌ی ضریب خودنفوذی برای ترکیبات سری تست بیرونی

n-پنتان	مقادیر تجربی	PCR	PLS	SR-ANN	GA-ANN
۱	۱/۷۰۰	۱/۵۷۴	۱/۷۲۸	۱/۷۶۱	۱/۶۹۵
۲	۱/۴۰۰	۱/۳۶۶	۱/۵۲۰	۱/۳۹۵	۱/۳۷۶
۳	۱/۱۷۰	۱/۱۵۶	۱/۳۱۱	۱/۲۱۰	۱/۱۶۱
۴	۱/۰۰۰	۰/۹۴۷	۱/۱۰۲	۰/۹۹۱	۱/۰۰۲
۵	۰/۸۶۰	۰/۷۳۸	۰/۸۹۳	۰/۸۴۸	۰/۸۰۸
۶	۰/۷۵۰	۰/۵۲۸	۰/۶۸۳	۰/۷۵۴	۰/۷۴۴
۷	۰/۶۲۰	۰/۲۰۸	۰/۳۶۴	۰/۶۴۴	۰/۶۶۰
۸	۰/۵۱۲	۰/۱۰۱	۰/۵۹۰	۰/۵۰۹	۰/۵۸۸
۹	۰/۴۳۵	۰/۵۳۵	۰/۶۲۳	۰/۴۳۵	۰/۴۳۱
۱۰	۰/۳۷۰	۰/۶۴۱	۰/۴۰۱	۰/۳۷۵	۰/۳۷۲
۱۱	۰/۳۲۰	۰/۵۱۸	۰/۱۹۰	۰/۳۸۹	۰/۳۰۹
۱۲	۲/۲۲۰	۲/۳۲۳	۲/۴۷۷	۲/۱۸۴	۲/۲۳۴
۱۳	۱/۸۳۰	۲/۱۱۵	۲/۲۶۹	۱/۸۵۷	۱/۸۷۱
۱۴	۱/۵۵۰	۱/۹۰۵	۲/۰۶۰	۱/۵۷۶	۱/۵۰۹
۱۵	۱/۳۳۰	۱/۶۹۶	۱/۸۵۱	۱/۳۳۱	۱/۳۰۶
۱۶	۱/۱۶۰	۱/۴۸۷	۱/۶۴۲	۱/۱۸۳	۱/۱۲۳
۱۷	۱/۰۲۰	۱/۲۷۷	۱/۰۲۲	۱/۰۰۹	۰/۹۸۲
۱۸	۰/۸۶۰	۰/۹۵۷	۱/۱۱۳	۰/۸۳۲	۰/۸۴۲
۱۹	۰/۷۳۰	۰/۶۴۸	۰/۸۰۴	۰/۷۱۱	۰/۷۵۶
۲۰	۰/۶۲۰	۰/۳۳۸	۰/۴۹۵	۰/۶۱۸	۰/۶۷۸
۲۱	۰/۵۳۰	۰/۶۸۰	۰/۴۸۹	۰/۵۲۸	۰/۶۰۴
۲۲	۰/۴۶۰	۰/۵۶۹	۰/۵۰۱	۰/۴۵۱	۰/۵۳۶
۲۳	۲/۵۳۰	۲/۶۹۷	۲/۸۵۱	۲/۴۲۳	۲/۵۰۹
۲۴	۲/۱۲۰	۲/۴۸۹	۲/۶۴۳	۲/۱۵۹	۲/۱۴۲
۲۵	۱/۸۰۰	۲/۲۸۰	۲/۴۳۴	۱/۷۸۷	۱/۸۲۸
۲۶	۱/۵۵۰	۲/۰۷۰	۲/۲۲۵	۱/۵۲۳	۱/۵۰۹
۲۷	۱/۳۶۰	۱/۸۶۱	۱/۴۲۱	۱/۳۵۹	۱/۳۷۹
۲۸	۱/۱۹۰	۱/۶۵۱	۱/۱۲۹	۱/۱۸۶	۱/۱۶۲

ادامه‌ی جدول (۲-پ)

n-پنتان	مقادیر تجربی	PCR	PLS	SR-ANN	GA-ANN
۲۹	۱/۰۰۰	۱/۳۳۲	۰/۹۱۲	۰/۹۹۰	۰/۹۴۹
۳۰	۰/۸۵۰	۱/۰۲۲	۱/۱۷۸	۰/۸۳۱	۰/۸۵۳
۳۱	۰/۷۳۰	۰/۷۱۳	۰/۸۶۹	۰/۷۲۹	۰/۷۶۷
۳۲	۰/۶۴۰	۰/۳۹۲	۰/۵۴۹	۰/۶۳۹	۰/۶۸۶
۳۳	۰/۵۶۰	۰/۴۶۰	۰/۲۴۰	۰/۵۵۷	۰/۶۱۲
۳۴	۲/۹۰۰	۳/۰۷۲	۳/۲۲۶	۲/۹۰۲	۲/۹۴۲
۳۵	۲/۴۲۰	۲/۸۶۴	۲/۵۰۹	۲/۳۸۴	۲/۴۲۶
۳۶	۲/۰۵۰	۲/۶۵۴	۲/۰۵۶	۲/۱۴۴	۲/۰۸۰
۳۷	۱/۷۷۰	۲/۴۴۵	۱/۸۴۵	۱/۹۳۸	۱/۷۵۴
۳۸	۱/۵۴۰	۲/۲۳۶	۱/۶۵۴	۱/۵۶۹	۱/۵۱۶
۳۹	۱/۳۶۰	۲/۰۲۶	۱/۲۹۰	۱/۳۵۷	۱/۳۰۰
۴۰	۱/۱۴۰	۱/۷۰۶	۱/۰۹۶	۱/۱۷۰	۱/۱۷۰
۴۱	۰/۹۷۰	۱/۳۹۷	۱/۵۵۳	۰/۹۶۹	۰/۹۶۳
۴۲	۰/۸۴۰	۱/۰۸۷	۱/۲۴۴	۰/۷۹۴	۰/۸۶۷
۴۳	۰/۷۴۰	۰/۷۶۷	۰/۹۲۴	۰/۹۲۴	۰/۷۷۶
۴۴	۰/۶۵۰	۰/۴۵۸	۰/۶۱۵	۰/۸۱۸	۰/۶۹۵
۴۵	۳/۵۵۰	۳/۶۰۲	۳/۵۱۰	۳/۵۱۱	۳/۵۰۹
۴۶	۳/۰۵۰	۳/۶۱۳	۲/۹۵۰	۳/۰۹۴	۳/۰۸۶
۴۷	۲/۶۳۰	۳/۴۰۳	۲/۶۰۰	۲/۵۹۵	۲/۶۶۷
۴۸	۲/۲۹۰	۳/۱۹۴	۲/۴۱۰	۲/۳۴۴	۲/۳۰۰
۴۹	۲/۰۰۰	۲/۹۸۵	۲/۱۰۶	۲/۱۲۳	۲/۰۳۳
۵۰	۱/۷۷۰	۲/۷۷۵	۱/۷۴۸	۱/۹۲۵	۱/۷۴۳
۵۱	۱/۴۹۰	۲/۴۵۶	۱/۴۰۸	۱/۴۸۳	۱/۴۷۵
۵۲	۱/۲۷۰	۲/۱۴۶	۱/۳۰۰	۱/۲۶۶	۱/۲۳۷
۵۳	۱/۱۰۰	۱/۸۳۶	۱/۰۳۴	۱/۲۵۸	۱/۱۱۴
۵۴	۰/۹۸۰	۱/۵۱۶	۱/۰۰۳	۱/۰۹۵	۰/۹۹۷
۵۵	۰/۸۸۰	۱/۲۰۷	۰/۸۱۱	۰/۹۶۳	۰/۸۹۵

ادامہی جدول (۲-پ)

n-نونا	مقادیر تجربی	PCR	PLS	SR-ANN	GA-ANN
۱	۰/۴۰۰	۰/۳۱۷	۰/۳۴۶	۰/۴۴۸	۰/۳۸۷
۲	۰/۳۱۳	۰/۴۰۸	۰/۳۸۹	۰/۳۹۶	۰/۳۴۸
۳	۰/۲۶۴	۰/۳۱۴	۰/۳۱۲	۰/۲۶۲	۰/۲۴۷
۴	۰/۶۱۰	۰/۴۴۶	۰/۴۰۷	۰/۶۳۷	۰/۶۴۵
۵	۰/۴۹۳	۰/۲۳۸	۰/۱۹۹	۰/۵۵۰	۰/۵۲۰
۶	۰/۴۰۸	۰/۰۲۸	۰/۵۰۲	۰/۴۹۲	۰/۳۷۵
۷	۰/۳۴۴	۰/۷۶۷	۰/۴۵۶	۰/۴۵۱	۰/۳۸۷
۸	۰/۲۹۲	۰/۲۴۳	۰/۲۷۶	۰/۳۰۱	۰/۲۲۸
۹	۰/۲۵۰	۰/۲۶۸	۰/۳۱۲	۰/۲۵۷	۰/۲۴۸
۱۰	۰/۲۱۵	۰/۲۸۹	۰/۱۸۹	۰/۲۱۸	۰/۲۲۱
۱۱	۰/۱۸۶	-۰/۱۰۹	۰/۲۲۷	۰/۱۸۴	۰/۱۲۰
۱۲	۰/۱۶۲	-۰/۱۲۹	۰/۱۹۲	۰/۱۶۳	۰/۰۹۳
۱۳	۰/۱۴۰	-۰/۱۴۴	-۰/۱۴۷	۰/۱۳۹	۰/۱۱۱
۱۴	۰/۱۲۲	-۰/۱۶۴	-۰/۱۶۸	۰/۱۲۳	۰/۱۳۳
۱۵	۰/۷۳۵	۰/۸۲۰	۰/۷۸۲	۰/۷۴۴	۰/۷۱۹
۱۶	۰/۶۰۰	۰/۶۱۲	۰/۵۷۴	۰/۶۳۶	۰/۶۴۳
۱۷	۰/۵۰۰	۰/۴۰۳	۰/۳۶۵	۰/۵۶۳	۰/۴۶۸
۱۸	۰/۴۲۲	۰/۱۹۳	۰/۱۵۶	۰/۵۱۲	۰/۳۶۲
۱۹	۰/۳۶۰	۰/۳۹۵	۰/۴۱۲	۰/۴۷۴	۰/۲۹۱
۲۰	۰/۳۰۹	۰/۳۱۸	۰/۵۱۶	۰/۲۹۹	۰/۲۳۷
۲۱	۰/۲۶۷	۰/۳۱۵	۰/۴۰۹	۰/۲۷۰	۰/۱۹۴
۲۲	۰/۲۳۲	-۰/۰۶۴	-۰/۰۶۸	۰/۲۳۹	۰/۱۵۹
۲۳	۰/۲۰۲	-۰/۰۸۵	-۰/۰۸۹	۰/۲۰۷	۰/۱۲۶
۲۴	۰/۱۷۷	-۰/۱۰۶	-۰/۱۱۰	۰/۱۷۹	۰/۰۹۸
۲۵	۰/۱۵۵	-۰/۱۲۷	-۰/۱۳۰	۰/۱۵۶	۰/۰۷۳
۲۶	۰/۱۳۷	-۰/۱۴۸	-۰/۱۵۲	۰/۱۳۷	۰/۰۵۰
۲۷	۰/۸۷۵	۱/۱۹۵	۱/۱۵۶	۰/۸۶۲	۰/۸۹۹
۲۸	۰/۷۰۷	۰/۹۸۷	۰/۹۴۸	۰/۷۳۰	۰/۷۹۶

ادامه‌ی جدول (۲-پ)

n-نونا	مقادیر تجربی	PCR	PLS	SR-ANN	GA-ANN
۲۹	۰/۵۹۰	۰/۷۷۷	۰/۷۳۹	۰/۶۴۰	۰/۵۸۷
۳۰	۰/۵۰۰	۰/۵۶۸	۰/۵۳۰	۰/۵۷۷	۰/۴۶۰
۳۱	۰/۴۳۰	۰/۳۵۹	۰/۳۲۱	۰/۵۳۰	۰/۳۷۴
۳۲	۰/۳۷۲	۰/۱۴۹	۰/۱۱۱	۰/۴۹۳	۰/۳۰۸
۳۳	۰/۳۲۵	۰/۳۱۷	۰/۲۸۹	۰/۴۶۳	۰/۲۵۵
۳۴	۰/۲۸۵	-۰/۰۲۷	-۰/۰۳۰	۰/۲۹۰	۰/۲۱۱
۳۵	۰/۲۵۲	-۰/۰۴۸	-۰/۰۵۲	۰/۲۴۸	۰/۱۷۱
۳۶	۰/۲۲۴	-۰/۰۶۹	-۰/۰۷۳	۰/۲۲۷	۰/۱۳۶
۳۷	۰/۲۰۰	-۰/۰۹۰	-۰/۰۹۳	۰/۲۰۹	۰/۱۰۶
۳۸	۰/۱۷۹	-۰/۱۱۰	-۰/۱۱۴	۰/۱۸۱	۰/۰۷۸
۳۹	۰/۱۵۴	-۰/۱۴۲	-۰/۱۴۵	۰/۱۵۷	۰/۰۴۲
۴۰	۱/۱۹۰	۱/۹۴۴	۱/۱۲۱	۱/۱۳۸	۱/۱۵۴
۴۱	۱/۰۲۰	۱/۷۳۶	۱/۱۱۹	۰/۹۴۷	۱/۰۲۴
۴۲	۰/۸۵۵	۱/۵۲۶	۰/۹۰۹	۰/۸۱۵	۰/۹۱۴
۴۳	۰/۷۲۵	۱/۳۱۷	۱/۲۷۹	۰/۷۲۲	۰/۷۳۷
۴۴	۰/۶۲۲	۱/۱۰۸	۱/۰۷۰	۰/۶۵۴	۰/۶۱۵
۴۵	۰/۵۳۸	۰/۸۹۸	۰/۸۶۰	۰/۶۰۰	۰/۵۲۲
۴۶	۰/۴۷۰	۰/۶۸۵	۰/۶۴۸	۰/۵۵۷	۰/۴۴۳
۴۷	۰/۴۱۳	۰/۴۸۲	۰/۴۴۶	۰/۵۲۲	۰/۳۷۹
۴۸	۰/۳۶۶	۰/۲۶۹	۰/۲۳۲	۰/۴۹۲	۰/۳۱۸
۴۹	۰/۳۲۵	۰/۰۵۵	۰/۰۱۹	۰/۴۶۶	۰/۲۶۳
۵۰	۰/۲۹۴	۰/۳۲۵	۰/۳۱۹	۰/۳۰۱	۰/۲۱۶
۵۱	۰/۲۶۶	-۰/۰۳۹	-۰/۰۴۰	۰/۲۶۳	۰/۲۵۶
۵۲	۰/۲۳۴	-۰/۰۶۷	-۰/۰۷۱	۰/۲۳۰	۰/۲۲۱

فهرست منابع

- [۱]. کلانتر ز، (۱۳۸۵)، رساله‌ی دکتری، "پیش‌بینی خواص ترمودینامیکی و انتقالی سیالات آلی چگال با استفاده از روش سهم گروه‌ها و شبکه‌ی عصبی موجک"، دانشکده شیمی، دانشگاه صنعتی اصفهان.
- [۲]. متر ک، (۱۳۸۴)، "نظریه و مسائل شیمی فیزیک"، سالاری ع، منجمی م، انتشارات امیر کبیر، تهران.
- [3]. Silbey R. J, Alberty R. A, Bawendi M. G, (2008), "**Physical Chemistry**", Fourth Edition, John Wiley & Sons, New York.
- [4]. Treybal R. E, (1980), "**Mass Transfer Operation**", Third Editon, Mc Graw-Hill, New York.
- [۵]. میرجردوی ن، (۱۳۹۰)، پایان نامه ارشد، "بررسی آزمایشگاهی ضریب نفوذ مولکولی گاز-نفت برای سیالات میادین نفتی ایران"، دانشکده مهندسی شیمی و نفت، دانشگاه صنعتی شریف.
- [۶]. لواین ای، (۱۳۸۰)، "شیمی فیزیک"، اسلامپور غ، پارسافر غ، مقاری ع، نجفی ب، جلد ۲، چاپ دوم، انتشارات فاطمی، تهران.
- [۷]. پورشاه نظری ک، (۱۳۶۹)، پایان نامه ارشد، "پیش بینی خواص انتقالی مایعات"، دانشکده مهندسی شیمی، دانشگاه صنعتی شریف.
- [8]. Longuet Higgins H. C, Pople J. A, (1956), "Transport Properties of a Dense Fluid of Hard Spheres", *J. Chem. Phys.*, **25**, pp 884.
- [9]. Houghton G, (1964), "Cubic Cell Model for Self-Diffusion in Liquids", *J. Chem. Phys.*, **40**, pp 1628.
- [10]. Maitland G. C, Rigby M, Smith E. B. and Wakeham W. A, (1987), "**Intermolecular Forces: Their Origin and Determination**", Oxford university Press, London.
- [۱۱]. قلی زاده ا، (۱۳۸۶)، پایان نامه ارشد، "استاندارد کردن روش متابونومیک بر پایه طیف سنجی NMR"، دانشکده علوم شیمی، دانشگاه صنعتی شریف.
- [12]. Lang E. W and Ludemann H. -D, (1993), "Density Dependence of Rotational and Translational Molecular Dynamics in Liquids Studied by high Pressure NMR", *J. Progress in NMR Spectroscopy*, **25**, pp 507.
- [13]. Vardag T, Bachl F, Wappmann S and Ludemann H. -D, (1990), "Pressure Dependence of Self Diffusion in Some Neat Alkanes and Binary Mixtures", *J. Ber.*

Bunsenges. Phys. Chem., **94**, pp 336.

[14]. Meyers R. A, (2000), "**Encyclopedia of Analytical Chemistry**", John Wiley & Sons, New York.

[15]. Tamm K, (2006), PhD.thesis, "QSPR Modeling of Some Properties of Organic Compounds", Chemistry department, Tartu University.

[16]. Hemmateenejad B, Safarpour M. A, Miri R and Nesari N, (2005), "Toward on Optimal Procedure for PC-ANN Model Building: Prediction of the Carcinogenic Activity of a Larg Set of Drugs", *J. Chem. Inf. Model.*, **45**, pp 190.

[17]. Martens H, Naes T, (1991), "**Multivariate calibration**", John Wiley & Sons, New York.

[۱۸]. اشرفی م، (۱۳۸۹)، پایان نامه ارشد، "مطالعه ارتباط کمی ساختار -فعالیت مشتقات تیوکربامات ها به عنوان دسته ی جدیدی از بازدارنده های غیر نوکلئوزیدی HIV"، دانشکده علوم، دانشگاه صنعتی شاهرود.

[19]. Zamora-Lopez H, Galicia L. A, Elizalde-Solis O, Hernandez-Rosales I. P, Mendez-Lango E, (2012), "Derived thermodynamic properties for the (ethanol+decane) and (carbon dioxide+ethanol+decane) systems at high pressures", *J. Chem. Thermodynamics*, **55**, pp 130.

[20]. Goldberg D. E, (2000), "**Genetic Algorithm in Search, Optimization and Machine Learning**", Addison-Wesley Longman, Boston.

[۲۱]. کیانی آ، (۱۳۸۶)، رساله ی دکتری، "توسعه ی شبکه های عصبی مصنوعی و سیستم های فازی و کاربرد آن ها در مطالعات پروتئومیکس و طراحی دارو"، دانشکده علوم شیمی، دانشگاه صنعتی شریف.

[22]. Brereton R. G, (2003), "**Chemometrics Data Analysis fr the Laboratory and Chemical Plant**", John Wiley & Sons, England.

[23]. Geladi P, Kowalski B. R, (1986), "Partial Least-Squares Regression: A Tutorial", *J. Anal. chim. Acta*, **185**, pp 1.

[24]. Lorber A, Wagnen L. E, Kowalski B. R, (1987), "A Theoretical foundation for the PLS algorithm", *J. Chemometr.*, **7**, pp 19.

[25]. Aptula A. O, Jeliaskova N. G, Schultz T. W and Cronin M. T. D, (2005), "QSAR", *J. Comb. Sci.*, **24**, pp 385.

[26]. Martins J. P. A, Barbosa E. G, Pasqualoto K. F. M, Ferreira, M. M. C, (2009), "LQTA-QSAR: A new \$D\$-QSAR Methodology", *J. Chem. Inf. model.*, **49**, pp 1428.

[27]. Tropsha A, Gramatica P, Gombar V. K, (2003), "QSAR", *J. Comb. Sci.*, **22**, pp 69.

[28]. Greiner Schmid A, Wappmann S, Has M and Ludemann H. -D, (1991), "Self

diffusion in the compressed fluid lower alkanes: methane, ethane, and propane”, *J. Chem. Phys.*, **94**, pp 5643.

[29]. Helbaek M, Hafskjold B, Dysthe D. K and Sorland G. H, (1996), ”Self diffusion coefficients of methane or ethane mixtures with hydrocarbons at high pressure by NMR”, *J. Chem. Eng. Data.*, **41**, pp 598.

[30]. Helbaek M, Hafskjold B, Dysthe D. K and Sorland G. H, (1996), ”Self diffusion coefficients of methane or ethane mixtures with hydrocarbons at high pressure by NMR”, *J. Chem. Eng. Data.*, **41**, pp 603.

[31]. Harris K. R, (1978), ”The density dependence of the self diffusion coefficient of methane at -50, 25 and 50 C”, *J. Physica*, **94A**, pp 448.

[32]. Harris K. R. and Trappeniers N. J, (1980), ”The density dependence of the self diffusion coefficient of liquid methane”, *J. Physica*, **104A**, pp 262.

[33]. Dymond J. H, (1974), ”Corrected enskog theory and transport coefficients of liquids”, *J. Chem. Phys.*, **60**, pp 969.

[34]. Bachl F, Ludemann H. -D, (1986), ”Pressure and temperature dependence of self diffusion in liquid linear hydrocarbons”, *J. Z. Naturforsch*, **41A**, pp 963.

[35]. Fishman E, (1955), ”Self diffusion in liquid n-pentane and n-heptane”, *J. Phys. Chem.*, **59**, pp 469.

[36]. Arkhipov V.P, (1982), ”Experimental study of the diffusion coefficient in liquid n-paraffins at high pressures”, *J. Izvestiia vysshikh uchebnykh zavedenii. Neft' i gaz*, pp 34.

[37]. Marbach W, Hertz H. G, (1996), ”Self and mutual diffusion coefficients of some n-alkanes at elevated temperatures and pressures”, *J. Z. Phys. Chem.*, **193**, pp 19.

[38]. Marbach W, Hertz H. G, Weingaertner H, (1995), ”Self and mutual diffusion coefficients of some binary liquid n-alkane mixtures. A velocity correlation study”, *J. Phys. Chem.*, **189**, pp 63.

[39]. Li S. F. Y, Yue L. S, (1990), ”Composition dependence of binary diffusion coefficients in alkane mixtures”, *J. Thermophys*, **11**, pp 537.

[40]. Eastale A. J, Wolf L. A, (1984), ”Solute-solvent interaction effects on tracer diffusion coefficients”, *J. Chem. Soc. Faraday Trans I*, **80**, pp 1287.

[41]. Albright J. G, Edge A. V. J, Mills R, (1983), ”Relations between intra diffusion and viscosity coefficients”, *J. Chem. Soc. Faraday Trans I*, **79**, pp 1327.

[42]. Harris K. R, (1982), ”Temperature and density dependence of the self diffusion coefficient of n-hexane from 223 to 333 K and up to 400 MPa”, *J. Chem. Soc. Faraday Trans I*, **78**, pp 2265.

[43]. Harris K. R, Alexander J. J, Goscinska T, (1993), ”Temperature and density dependence of the self diffusion coefficient of liquid n-octane and toluene”, *J.*

Molecular Physics, **78**, pp 235.

[44]. Holtz M, Heil S. R, Sacco A, (2000), "Temperature dependence self diffusion coefficient of water and six selected molecular liquids for calibration in accurate H NMR PFG measurements", *J. Phys. Chem. Chem. Phys.*, **2**, pp 4740.

[45]. Erkey C, Rodden J. B, Matthews M. A, Akgerman A, (1989), "Application of rough hard-sphere theory to diffusion in n-alkanes", *J. Thermophys*, **10**, pp 953.

[46]. Tarassov I. N, Wappmann S, Ludemann H. -D, (1997), "p,T-dependence of self diffusion in the fluid one phase region of the system tetradecane/methane", *Pol. J. Chem.*, **71**, pp 1763.

[47]. Vardag T, Karger N, Ludemann H. -D, (1991), "Temperature and pressure dependence of self diffusion in long liquid n-alkanes", *J. Ber. Bunsen-Ges.*, **95**, pp 859.

[48]. Dymondt J. H, Harris K. R, (1992), "The temperature and density dependence of the self diffusion coefficient of n-hexadecane", *J. Molecular Physics*, **75**, pp 461.

[49]. Hemmateenejad and Shamsipur M, (2004), "Quantitative Structure-Electromistry Relationship Study of Some Organic Compounds Using PC-ANN and PCR", *J. Molecular Design*, **3**, pp 316.

[50]. Todeschini R and Consonni V, (2000), "**Handbook of Molecular Descriptors**", Wiley-VCH, Germany.

Abstract

In this research, studies on Quantitative Structure-Property Relationship (QSPR) has been applied by using linear and nonlinear modeling in order to predict the self-diffusion coefficient for linear alkanes. The self-diffusion coefficient of cases were classified by the use of principal component analysis (PCA) into four categories; training, validation, test and external test. Descriptors of molecular structure was calculated. The linear modeling started after removing the additional descriptors and adding two parameters: temperature and pressure. Two methods, principal component regression (PCR) and partial least squares regression (PLS), were used for linear modeling which in both principal components has been made first and then optimum number of principal components has been calculated. Afterward modeling started by optimized principal components and the final models were applied in order to case prediction related to external test set. The Mean squared error (MSE) related to external test set for PCR and PLS models were obtained respectively 0.6129 and 0.5327.

Two methods, stepwise regression (SR) and genetic algorithm (GA), were used for separating optimum descriptors in nonlinear modeling. Optimum descriptors entered the neural network and modeling was done for both methods in those descriptors have been chosen. After training and optimization, built networks were used for predicting related cases to external testing series. The MSE for the external test set in SR-ANN and GNN models was 0.0379 and 0.0219, respectively. These results indicate the superiority of nonlinear methods in describing selection phase and modeling stage.

Keywords: Quantitative structure –property relationship(QSPR), Principal Component Regression(PCR), Partial Least Squares, artificial neural network(ANN), Self-diffusion(D), Stepwise regression (SR), Genetic algorithm (GA)



Shahrood University of Technology

Faculty of Chemistry

M.Sc. Thesis in Physical Chemistry

**Quantitative Structure- Property Relationship study of Self-diffusion for
n-alkanes using linear and nonlinear methods**

Zahra Shahabi

Supervisors:

Dr. Z. Kalantar

Advisor:

Dr. H. Nikofard

February 2014