



Estimation of precipitation using the combined method of Support Vector Machine- Simulated Annealing algorithm (case study: Gorgan synoptic station)

Saeed Azadi ¹ , Maryam Vafae ² , Hamed Nozari ³

¹ Ph. D graduated of Water Resources Engineering, Department of Water science engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: s_azadi_ir@yahoo.com

² Ph. D graduated of Water Resources Engineering, Department of Water science engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: ma.vafae2014@gmail.com

³ Corresponding Author, Associate Professor, Department of Water science engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: h.nozari@basu.ac.ir

ABSTRACT

Introduction

Precipitation is one of the essential components of the water cycle and is considered one of the most critical input components of the hydrological cycle. Artificial intelligence methods have a comprehensive ability to model and predict complex problems. These methods are based on dynamic systems. They can create a relationship between dependent and independent variables by analyzing input data and the results without considering the physical process governing the system. The SVM intelligent hybrid models and the SA optimization algorithm are intelligence methods algorithms that have been used separately or in combination in water science and engineering. In the current research, the accuracy of the Simulated Annealing algorithm based on a Support Vector Machine (SVM-SA) was evaluated in the simulation of precipitation changes.

Methodology

To verify the results, we used the precipitation data of the Gorgan synoptic station during the 40 years from 1971 to 2010. Non-precipitation meteorological data have been used to estimate rainfall, including cloud cover, maximum and minimum temperature, water vapor pressure, wind speed, maximum and minimum relative humidity, and dew point. The SVM-SA algorithm can avoid the time-consuming trial and error phase to obtain the results and minimize the estimation error. In this method, the input parameters are first determined in the SVM algorithm; then, the initial parameters σ and γ are specified. The SVM finds the hyperplane that gives the most considerable minimum distance to the training data. SA algorithm avoids wasting time trying to find the optimum values of σ and γ parameters.

In this research, the skewness and elongation of the data in SPSS software indicated the normal distribution of the studied parameter. Therefore, due to the non-normality of the data, the normalization of non-rainfall meteorological data was done using the minimum-maximum conversion method (Eq. 1) based on the maximum and minimum values.

$$Y_t = \frac{x_t - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where, Y_t is the normalized value, x_t is the value of each data, x_{min} is the lowest value of the data, and x_{max} is the highest value of the data. So, the scale of both features is 0 to 1, and different classification algorithms on the data can now be implemented fairly.

The root mean square error (RMSE), standard error (SE), and correlation coefficient (R2) were used to evaluate the model's validity (Eq. 2 to 4). After normalizing the data, the effective meteorological parameters on rainfall using SPSS software were determined. Then, the value of the Pearson correlation coefficient between non-precipitation meteorological parameters is at a significant level of 99% for the selection of input parameters to the algorithm was calculated.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_m - Y_p)^2}{n}} \quad (2)$$

$$SE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_m - Y_p)^2}}{\bar{Y}_m} \quad (3)$$

$$R^2 = \frac{[\sum_{i=1}^n (Y_m - \bar{Y}_m)(Y_p - \bar{Y}_p)]^2}{\sum_{i=1}^n (Y_m - \bar{Y}_m)^2 \cdot \sum_{i=1}^n (Y_p - \bar{Y}_p)^2} \quad (4)$$

Where n represents the number of days in the study period, Y_m is the measured values for each day, Y_p is the predicted values by model, \bar{Y}_m is the average of measured data, and \bar{Y}_p is the average of the predicted data.

Results and discussion

The results of Pearson's correlation coefficient showed that out of the eight considered parameters, cloudiness parameters, average maximum temperature, water vapor pressure, maximum relative humidity, and dew point have the most significant correlation at the 99% level with rainfall, so parameters The inputs were considered to predict the rainfall parameter in the SVM-SA algorithm. Based on the results, using five non-precipitation meteorological parameters, including cloud cover, maximum temperature, water vapor pressure, maximum relative humidity, and dew point, the RMSE, SE, and R^2 in the training section are equal to 6.02 mm, 0.01, and 0.999. In the testing section, 18.72 mm, 0.03, and 0.925 mm, were calculated, respectively.

Conclusions

The results showed that the SVM-SA could be highly accurate in simulating precipitation changes in the study area, and this model can be introduced as a new hybrid algorithm in engineering sciences.

Keywords: Artificial intelligence; Precipitation; Simulated Annealing; Simulation; Support Vector Machine

Article Type: Research Article

Article history: Received: 01 September 2022 Revised: 01 November 2022 Accepted: 09 November 2022 ePublished: 22 November 2022

Cite this article: Azadi, S., Vafae, M., & Nozari, H. (2022). Estimation of precipitation using the combined method of Support Vector Machine- Simulated Annealing algorithm (case study: Gorgan synoptic station), *Advanced Technologies in Water Efficiency*, 2(3), 14-34, DOI: 10.22126/ATWE.2022.8393.1028

Publisher: Razi University

© The Author(s).





برآورد بارندگی با استفاده از روش ترکیبی ماشین بردار پشتیبان - الگوریتم تبرید (مطالعه موردی: ایستگاه سینوپتیک گرگان)

سعید آزادی^۱ ، مریم وفايي^۲ ، حامد نوذری^۳ ✉

^۱ دانش آموخته دکتری مهندسی منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران، رایانامه: s_azadi_ir@yahoo.com
^۲ دانش آموخته دکتری مهندسی منابع آب، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران، رایانامه: ma.vafae2014@gmail.com
^۳ نویسنده مسئول، دانشیار گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران، رایانامه: h.nozari@basu.ac.ir

چکیده

بارندگی یکی از اجزای اساسی چرخه آب است و به عنوان یکی از مهم ترین مؤلفه های ورودی چرخه هیدرولوژیکی به شمار می رود. در تحقیق حاضر دقت الگوریتم هیبریدی شبیه سازی تبرید بر پایه ماشین بردار پشتیبان (SVM-SA) در شبیه سازی تغییرات بارندگی مورد ارزیابی قرار گرفت. به منظور واسنجی و صحت سنجی نتایج از آمار و اطلاعات بارندگی ایستگاه سینوپتیک گرگان طی دوره ۴۰ ساله ۱۹۷۱ تا ۲۰۱۰ استفاده شد. در روند تحقیق ابتدا پارامترهای هواشناسی مؤثر بر مقدار بارندگی تعیین شد. سپس به منظور انتخاب پارامترهای ورودی به الگوریتم مورد نظر، با استفاده از نرم افزار SPSS مقدار ضریب همبستگی پیرسون بین پارامترهای هواشناسی در سطح معنی داری ۹۹ اعمال گردید. نتایج ضریب همبستگی پیرسون نشان داد که از هشت پارامتر مورد نظر، پارامترهای ابرناکی، میانگین دمای حداکثر، فشار بخار آب، رطوبت نسبی حداکثر و نقطه شبنم بیشترین همبستگی معنی دار در سطح ۹۹ درصد را با بارندگی دارند. بنابراین به عنوان پارامترهای ورودی برای پیش بینی پارامتر بارندگی در الگوریتم SVM-SA در نظر گرفته شد. بر اساس نتایج، با استفاده از ۵ پارامتر هواشناسی مذکور، مقدار شاخص های SE ، $RMSE$ و R^2 در بخش آموزش به ترتیب برابر با ۶/۰۲ میلی متر، ۰/۰۱ و ۰/۹۹۹ و در بخش آزمون مقدار این شاخص ها به ترتیب ۱۸/۷۲ میلی متر، ۰/۰۳ و ۰/۹۲۵ محاسبه گردید. نتایج نشان داد که روش ترکیبی SVM-SA می تواند در شبیه سازی تغییرات بارندگی در منطقه مورد مطالعه از دقت بالایی برخوردار باشد و این مدل را به عنوان الگوریتم هیبریدی جدید در حوزه علوم مهندسی می توان معرفی کرد.

واژه های کلیدی: بارندگی، شبیه سازی، ماشین بردار پشتیبان، هوش مصنوعی، هیبریدی شبیه سازی تبرید

نوع مقاله: مقاله پژوهشی

سابقه مقاله: دریافت: ۱۰ شهریور ۱۴۰۱ اصلاح: ۱۰ آبان ۱۴۰۱ پذیرش: ۱۸ آبان ۱۴۰۱ چاپ الکترونیکی: ۰۱ آذر ۱۴۰۱

استناد: آزادی، س، وفايي، م، و نوذری، ح. (۱۴۰۱). برآورد بارندگی با استفاده از روش ترکیبی ماشین بردار پشتیبان - الگوریتم تبرید (مطالعه موردی: ایستگاه سینوپتیک گرگان)، فناوری های پیشرفته در بهره وری آب، ۳(۳)، ۱۴-۳۴، شناسه دیجیتال: 10.22126/ATWE.2022.8393.1028



© نویسندگان.

ناشر: دانشگاه رازی

مقدمه

محدودیت منابع آبی جهت تأمین آب موردنیاز کشاورزی و غیر کشاورزی موجب بروز مشکلات عمده‌ای شده است. لذا پیش‌بینی بارندگی به‌عنوان یکی از مهم‌ترین منابع آبی موجود (وانگ و همکاران^۱، ۲۰۱۳) برای اهداف مختلفی نظیر برآورد سیلاب، خشک‌سالی، مدیریت حوضه آبریز، کشاورزی و غیره دارای اهمیت است و پیش‌بینی بارندگی در مقیاس‌های مختلف یکی از اقدامات مفید و مؤثر جهت برنامه‌ریزی‌های خرد و کلان اقتصادی و اجتماعی است (جباری^۲، ۲۰۰۵).

روش‌های هوش مصنوعی قابلیت وسیعی در مدل‌سازی و پیش‌بینی مسائل پیچیده دارند. این روش‌ها مبتنی بر سیستم‌های دینامیکی بوده و با تجزیه و تحلیل داده‌های ورودی و نتایج نظیر آن‌ها و بدون در نظر گرفتن فرآیند فیزیکی حاکم بر سیستم، قادر به ایجاد رابطه بین متغیر وابسته و مستقل هستند (پالمر^۳، ۲۰۲۰). از جمله روش‌های هوش مصنوعی، الگوریتم SVM است. این روش به‌نوبه خود به‌ویژه زمانی که روابط فیزیکی کاملاً شناخته‌شده نیست به کار گرفته می‌شود. این روش نیازی به دانستن فرآیند پدیده ندارد و زیرمجموعه روش‌های جعبه سیاه است.

مبانی نظری و پیشینه پژوهش

در سال‌های اخیر مطالعات زیادی از موارد کاربرد روش SVM و الگوریتم بهینه‌سازی SA^۴ به‌صورت جداگانه یا ترکیبی در علوم و مهندسی آب صورت گرفته است. به‌عنوان مثال دیبیک و همکاران^۵ (۲۰۰۱) در تحقیقی از مدل SVM در مدل‌سازی بارندگی- رواناب استفاده کردند. در تحقیقی دیگر، آسفا و همکاران^۶ (۲۰۰۶) از SVM برای پیش‌بینی دبی جریان استفاده نمودند که نتایج مطلوب به‌دست‌آمده، نشان از توانایی این تکنیک دارد.

تریپاتی و همکاران^۷ (۲۰۰۶) جهت برآورد ماهانه بارندگی در هند از مدل SVM استفاده کردند. در این مطالعه محققان پیشنهاد کردند که SVM به‌عنوان گزینه برتر نسبت به سایر روش‌های معمول برای پیش‌بینی بارندگی مورد استفاده قرار گیرد. یو و لینگ^۸ (۲۰۰۶) از مدل SVM به‌منظور پیش‌بینی سری زمانی متغیرهای هیدرولوژیکی استفاده نمودند. چن و همکاران^۹ (۲۰۱۰)، با مدل SVM ریز مقایس نمودن بارندگی روزانه را بررسی و با روش آنالیز چند متغیره مقایسه نمودند. نتایج نشان داد پیش‌بینی‌های حاصل از مدل SVM دقیق‌تر از روش آنالیز چند متغیره است. حمیدی و همکاران^{۱۰} (۲۰۱۴) به پیش‌بینی بارندگی ماهانه ایستگاه‌های سینوپتیک فرودگاه و نوزه در استان همدان با دو مدل SVM و شبکه عصبی مصنوعی پرداختند. نتایج نشان داد مدل SVM نسبت به شبکه عصبی از کارایی بالاتری برخوردار است. ضابط پیشخانی و همکاران (۱۳۹۵) به مقایسه الگوسازی بارندگی ماهانه با مدل‌های SVM و ANFIS^{۱۱} در شهر گنبد کاووس پرداختند. نتایج پیش‌بینی بارندگی با بارندگی ایستگاه‌های منطقه برای هر دو روش قابل‌قبول بود. نتایج نشان داد که SVM عملکرد بهتری نسبت به ANFIS در پیش‌بینی بارندگی داشته است. عیسی زاده و همکاران (۱۳۹۶) به ارزیابی تأثیر نرمال‌سازی توزیع احتمالاتی رواناب ماهانه بر عملکرد مدل‌های SVM و ANN^{۱۲} در شبیه‌سازی جریان ماهانه رودخانه‌ها در حوزه زربینه‌رود پرداختند. نتایج نشان داد که مدل SVM از دقت بالاتری برخوردار است. نوذری و توکلی (۱۳۹۸) به ارزیابی کارایی مدل‌های مختلف خطی و غیرخطی در پیش‌بینی بارندگی ماهانه در تغییرات اقلیم استان همدان پرداختند. در این مطالعه، به‌منظور پیش‌بینی مقادیر ماهانه بارندگی از مدل‌های SVM، SVM تلفیق شده با تبدیل

¹ Wang et al

² Jabbari

³ Palmer

⁴ Support Vector Machine

⁵ Simulated Annealing

⁶ Dibike et al

⁷ Asefa et al

⁸ Tripathi et al

⁹ Yu & Liang

¹⁰ Chen et al

¹¹ Hamidi et al

¹² Adaptive neuro fuzzy inference system

¹³ Artificial neural network

موجک (W-SVM)، ARIMA^۲ و ARMAX استفاده گردید. در نهایت نتایج نشان داد، مدل SVM دارای پارامترهای قابل تنظیم کمتری نسبت به مدل های دیگر است. لذا این مدل با سهولت بیشتر و در زمان کمتری قادر به پیش بینی بارندگی بوده و از این نظر نسبت به سایر روش ها ارجحیت دارد.

شرفی و همکاران (۱۳۹۹) به پیش بینی بارندگی ماهانه شهر اردبیل با استفاده از مدل های برنامه ریزی ژنتیک و SVM پرداختند. به طور کلی نتایج نشان داد که مدل SVM برای مدل سازی و پیش بینی بارندگی ماهانه استان اردبیل مناسب تر بوده است و این مدل به عنوان یک ابزار بسیار مؤثر برای مدل سازی سیستم برای پیش بینی بارندگی ماهانه است.

اما در مورد استفاده و تلفیق مدل SVM و الگوریتم بهینه سازی SA مطالعاتی توسط محققین انجام شده است که از آن جمله می توان به تحقیق پی و هونگ^۳ (۲۰۰۶) اشاره کرد. آن ها از الگوریتم هیبریدی شبیه سازی تبرید بر پایه ماشین بردار پشتیبان (SVM-SA) برای حل مسائل رگرسیون غیرخطی و سری زمانی استفاده کردند. نتایج نشان داد که این روش به پیش بینی های بهتری نسبت به SVM به صورت مجزا منجر می شود.

لین و همکاران^۵ (۲۰۰۸) از روش الگوریتم بهینه سازی SA برای تعیین پارامتری الگوریتم SVM، به نام SVM-SA استفاده کردند. روش پیشنهادی با جستجوی شبکه ای که یک روش مرسوم برای اجرای تنظیمات پارامتر در سیستم است مقایسه شد. نتایج نشان داد که طبقه بندی روش SVM-SA از جستجوی شبکه و سایر رویکردها دقیق تر است.

جیا و همکاران^۶ (۲۰۱۱) ویژگی های شیرهای هیدرولیک را با روش الگوریتم SVM-SA بررسی کردند. نتایج نشان داد که SVM-SA در مقایسه با ANFIS و شبکه عصبی مصنوعی از دقت بالاتری برخوردار است. جیا و همکاران^۷ (۲۰۲۰) به بررسی و تخمین تبخیر و تعرق با استفاده از داده های هواشناسی برای سواحل غربی شبه جزیره مالزی پرداختند. در این تحقیق آن ها از مدل های SVM-Mak^۸، SVM-HS^۹، SVM-Rom^{۱۰} و SVM-SA استفاده نمودند. نتایج نشان داد تلفیق مدل SVM با مدل های دیگر می تواند از دقت بیشتر و خطای کمتری جهت تخمین تبخیر و تعرق برخوردار باشد. همچنین چن^{۱۱} (۲۰۱۶) با استفاده از SVM-SA به استخراج داده های هیدرولوژیکی و انتقال آن ها در حوادث سیل در حوضه رودخانه یانگ تایوان را مورد بررسی قرارداد و نشان داد که این الگوریتم در حل مسائل مختلف توانمند است.

مهربانی بشار و همکاران^{۱۲} (۲۰۲۲) به ارزیابی عوامل مؤثر بر مصرف آب شرب روستایی استان همدان با استفاده از مدل های هیبریدی هوشمند پرداختند. در این مطالعه، اثر عوامل بر تقاضای آب شرب روستایی با استفاده از روش ANFIS و مدل های ترکیبی مانند ANFIS-GA، ANFIS-PSO و SVM-SA مورد بررسی قرار گرفت. به منظور ارزیابی مصرف آب آشامیدنی پنج عامل بارندگی، رطوبت نسبی، دما، تعداد مشترکین و قیمت آب انتخاب شدند. نتایج نشان داد مدل ترکیبی SVM-SA عملکرد و دقت بالاتری نسبت به مدل های دیگر در بررسی عوامل مؤثر بر مصرف آب شرب روستایی داشته است.

محمدی و حیدر پور (۱۳۹۸) به بررسی تأثیر سنجه های اقلیمی بر روی تخمین دمای خاک توسط الگوریتم SVM-SA پرداختند. نتایج نشان داد که الگوریتم SVM-SA می تواند به عنوان ابزاری کارآمد در مدل سازی دمای عمیق خاک استفاده شود.

¹ Wavelet- Support Vector Machines

² Autoregressive Integrated Moving Average

³ Pai & Hong

⁴ Support Vector Machines - Simulated Annealing Algorithm

⁵ Lin et al

⁶ Jia et al

⁷ Chia et al

⁸ Support Vector Machines- Makkink

⁹ Support Vector Machines- Hargreaves-Samani

¹⁰ Support Vector Machines- Romanenko

¹¹ Chen

¹² MehrabaniBashar et al

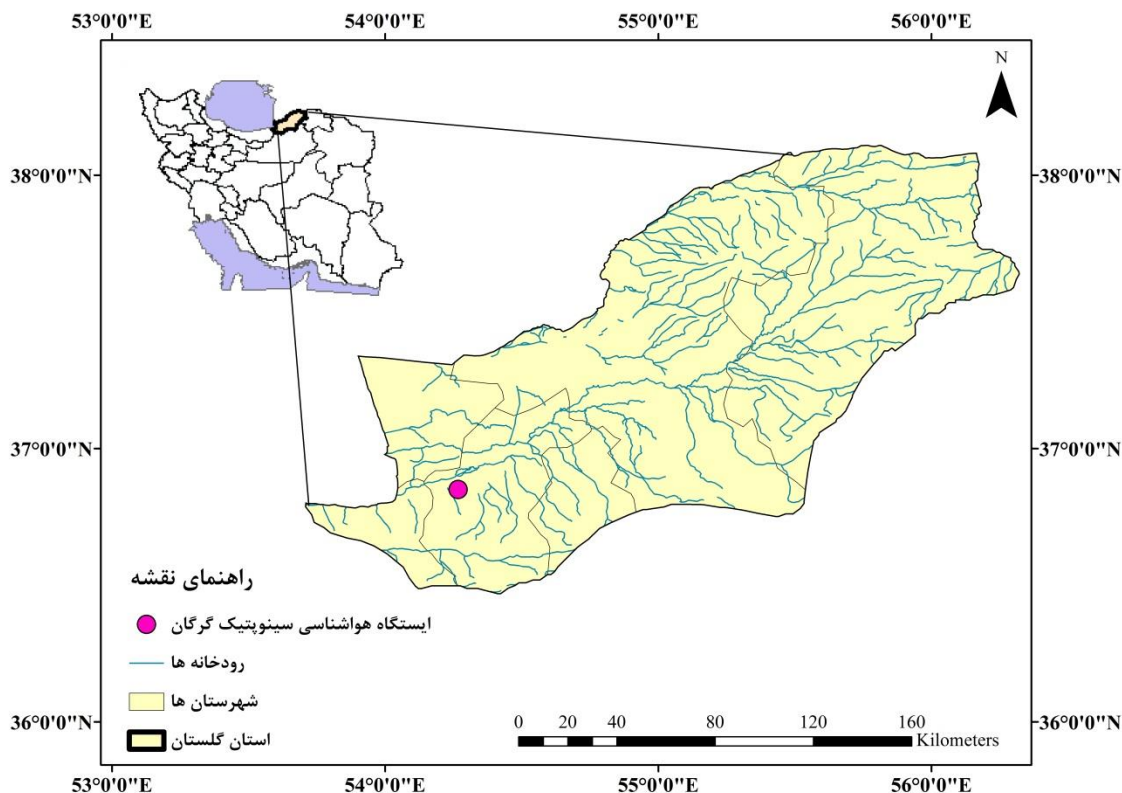
محمدی و مؤذنزاده (۱۳۹۸) نیز به پیش‌بینی جریان با استفاده از مدل‌های ترکیبی هوشمند در مقیاس ماهانه در منطقه رودخانه زرین رود پرداختند. در این مطالعه از داده‌های هواشناسی و سری زمانی ماهانه دبی رودخانه زرین رود (ایستگاه صفوانکه) آذربایجان شرقی استفاده شد. نتایج نشان داد که شبیه‌سازی الگوریتم ترکیبی SVM-SA نسبت به سایر مدل‌های هیبریدی هوشمند عملکرد بهتری برای شبیه‌سازی دبی جریان دارد.

بررسی منابع نشان از کارایی SVM و SA به‌صورت جداگانه یا ترکیبی در زمینه آب را دارد. باید خاطر نشان کرد، با توجه به مطالعات صورت گرفته تحقیقی جهت پیش‌بینی تغییرات بارندگی با روش ترکیبی SVM-SA تاکنون صورت نگرفته است. بنابراین هدف از تحقیق حاضر، ارزیابی قابلیت مدل ترکیبی SVM-SA در شبیه‌سازی پارامترهای هواشناسی غیر بارندگی در برآورد بارندگی است. نتایج این مطالعه می‌تواند در تعیین نوع و تعداد پارامترهای مؤثر در تخمین بارندگی ایستگاه سینوپتیک مفید واقع گردد.

روش پژوهش

منطقه مورد مطالعه

شهر گرگان از لحاظ مختصات جغرافیایی در محدوده ۴۷ درجه و ۴۸ دقیقه تا ۴۸ درجه و ۳۹ دقیقه طول جغرافیایی و ۳۷ درجه و ۵۶ دقیقه تا ۳۸ درجه و ۳۳ دقیقه عرض جغرافیایی قرار دارد. این شهر مرکز استان گلستان محسوب می‌شود و با مساحت حدودی ۴۰ کیلومتر مربع در جنوب استان واقع شده است. گرگان از شمال با شهرستان‌های آق‌قلا و بندر ترکمن، از جنوب با استان سمنان، از شرق با شهرستان علی‌آباد و از غرب با شهرستان کردکوی همسایه است. از نظر اقلیم طبیعی، شهر گرگان به سه دسته کوهستانی، جلگه‌ای و کوهپایه‌ای تقسیم می‌شود و دارای هوای معتدل مدیترانه‌ای است. بیشترین میزان رطوبت هوا در ماه‌های فروردین، آبان و بهمن است. بارش باران در ماه‌های فروردین و دی به بیشترین حد خود می‌رسد. ایستگاه هواشناسی این شهر از نوع سینوپتیک بوده و موقعیت مکانی آن در ۳۶ درجه و ۵۱ دقیقه عرض شمالی و ۵۴ درجه و ۱۶ دقیقه طول غربی قرار گرفته است. شکل ۱ موقعیت جغرافیایی منطقه مورد مطالعه را نشان می‌دهد.



شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه

در این تحقیق از داده‌های سالانه بارندگی ایستگاه سینوپتیک گرگان طی دوره آماری ۴۰ ساله ۱۹۷۱-۲۰۱۰ استفاده شده است. جهت تخمین بارندگی از داده‌های هواشناسی غیر بارندگی شامل: ابرناکی، میانگین دمای حداکثر، میانگین دمای حداقل، فشار بخار آب، سرعت باد، رطوبت نسبی حداکثر، رطوبت نسبی حداقل و نقطه شبنم استفاده شده است.

الگوریتم ماشین بردار پشتیبان (SVM)

SVM یک الگوریتم یادگیری کارآمد بر مبنای نظریه بهینه‌سازی مقید است که از اصل استقرای کمینه‌سازی خطای ساختاری استفاده می‌کند و منجر به یک جواب بهینه کلی می‌شود (نوری، ۲۰۱۱). این روش، یک تکنیک یادگیری برای کلاسه‌بندی داده‌های نهان به طور صحیح است. برای این کار، SVM یک ابر صفحه ایجاد می‌کند که داده‌ها را به دسته‌های مختلف جدا می‌نماید و مجموعه داده‌ها ممکن است به طور خطی از هم جدا باشند (پولت و گیونس، ۲۰۰۷).

در مدل رگرسیونی SVM، تابعی مرتبط با متغیر وابسته y که خود تابعی از چند متغیر مستقل x است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی، فرض می‌شود که رابطه میان متغیرهای مستقل و وابسته با تابع جبری (رابطه ۱) به علاوه مقداری اغتشاش به عنوان خطای مجاز (ϵ) به نام $noise$ (رابطه ۲) تعریف شده است (نوری، ۲۰۱۱).

¹ Nouri et al

² Polat & Gunes

$$f(x) = w^T \times \phi(x) + b \quad (۱)$$

$$y = f(x) + \text{noise} \quad (۲)$$

در این روابط w بردار ضرایب، b ثابت مشخصه‌های تابع رگرسیونی و $\phi(x)$ تابع کرنل و هدف پیدا کردن تابع $f(x)$ است. این هدف با آموزش مدل SVM توسط مجموعه‌ای از نمونه‌ها (مجموعه آموزش) محقق می‌شود. برای محاسبه w و b می‌توان نوشت:

$$w^T \times \phi(x_i) + b - y_i \leq \varepsilon + \varepsilon_i^* \quad (۳)$$

$$\frac{1}{2} \times w \times w + C \sum_{i=1}^n \varepsilon_i + C \sum_{i=1}^n \varepsilon_i^* \times y_i - w^T \times \phi(x_i) - b \leq \varepsilon + \varepsilon_i^* \quad i = 1, 2, \dots, n \text{ and } \varepsilon_i, \varepsilon_i^* \geq 0 \quad (۴)$$

در روابط فوق C ، عددی صحیح و مثبت است که عامل تعیین جریمه در هنگام رخ دادن خطای آموزش مدل است، n تعداد نمونه‌ها، ε_i و ε_i^* متغیرهای کمبود هستند که حد بالا و پایین خطای آموزش مرتبط با مقدار خطای مجاز ε را مشخص می‌کنند. در نهایت تابع SVM را می‌توان به صورت زیر بازنویسی کرد:

$$f(x) = \sum_{i=1}^n \bar{\alpha}_i \times \phi(x_i)^T \times \phi(x) + b \quad (۵)$$

$\bar{\alpha}_i$ میانگین ضرایب لاگرانژ است که فقط مجموعه داده‌هایی که ضرایب $\bar{\alpha}_i$ آن‌ها غیر صفر است در معادله رگرسیون نهایی وارد می‌شوند. این مجموعه داده‌ها به عنوان بردارهای پشتیبان شناخته می‌شوند (فلچر، ۱۹۸۷). همچنین محاسبه $\phi(x)$ در فضای تابع هدف بسیار پیچیده است و برای حل این مشکل روند معمول در مدل SVM انتخاب یک تابع کرنل است. متداول‌ترین تابع کرنل برای SVM، تابع RBF_kernel (رابطه ۶)، تابع RBF_kernel به همراه γ با یک مشخصه هدف (رابطه ۷) و تابع Polynomial_kernel (رابطه ۸) می‌باشند:

$$\phi(x, x_i, \sigma) = e^{-\frac{|x-x_i|^2}{\sigma^2}} \quad (۶)$$

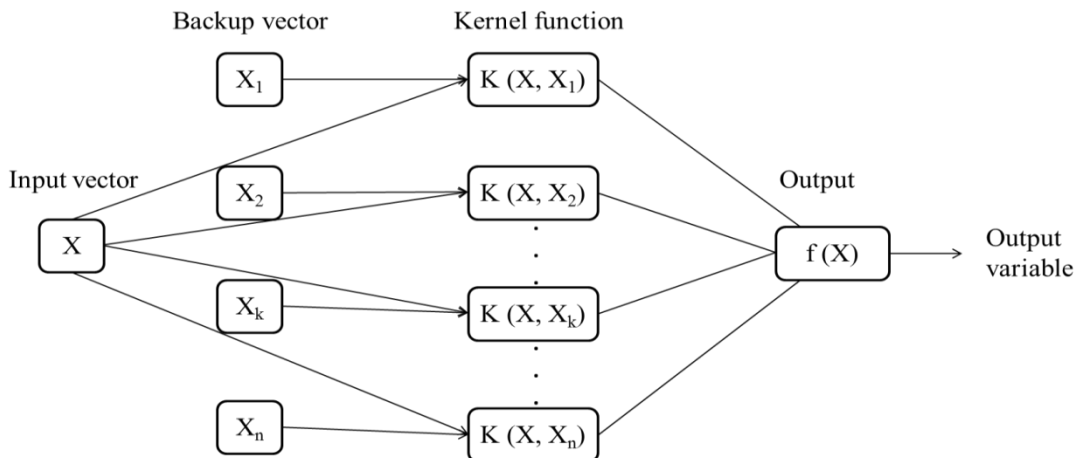
$$\phi(x, x_i, \gamma) = e^{-\gamma |x_i - x|^2} \quad (۷)$$

$$\phi(x, x_i, p) = (1 + x_i \cdot x_i)^p \quad (۸)$$

در روابط فوق σ و γ پارامترهای قابل تنظیم و کنترل‌کننده نتایج و تعمیم دهنده مدل SVM برای هسته‌های RBF_kernel می‌باشند. p پارامتر قابل تنظیم برای هسته Polynomial_kernel است. در این تحقیق، از تابع هسته RBF_kernel برای SVM استفاده شده است، زیرا تابع هسته RBF می‌تواند داده را با ابعاد بالا آنالیز کند و در این تابع باید مشخصه‌های σ و γ بهینه شوند (پولت و گیونس، ۲۰۰۷). در شکل ۲ ساختار کلی الگوریتم SVM نشان داده شده است. لازم به ذکر است فرآیند محاسبات الگوریتم SVM، به صورت کد نویسی در نرم‌افزار MATLAB انجام و پارامترهای تابع کرنل از طریق سعی و خطا بهینه می‌شود.

¹ Fletcher

² Radial Basis Function



شکل ۲. ساختار الگوریتم SVM

الگوریتم بهینه‌سازی باز پخت (SA)

روش الگوریتم SA یک روش بهینه‌سازی بر مبنای روش مونت کارلو است (گونزالو و همکاران؛^۱ ۲۰۰۱). این الگوریتم یک روش جستجوی محلی است که سعی در به دست آوردن جواب بهینه سراسری دارد. اساس الگوریتم SA بر مبنای رابطه بین ساختار اتمی، آنتروپی و دما در طول تبرید یک ماده و شباهت آن با پدیده آنیلینگ است (تران و همکاران؛^۲ ۲۰۰۳).

مفهوم اصلی الگوریتم SA از فرآیند فیزیکی تبرید فلزات مذاب سرچشمه می‌گیرد. در فرآیند تبرید، یک فلز مذاب با دمای بسیار بالا به تدریج خنک می‌شود. در دماهای بالا، اتم‌های سازنده فلز به صورت تصادفی قرار گرفته‌اند و لذا می‌توانند به آسانی با یکدیگر جابه‌جا شوند. با کاهش تدریجی دما، حرکت اتم‌ها محدود می‌شود به طوری که اتم‌ها شروع به مرتب شدن نموده و تشکیل کریستال می‌دهند. سطح انرژی کریستال تشکیل شده، به سرعت سرد کردن فلز بستگی دارد. اگر کاهش دما خیلی سریع صورت بگیرد، ممکن است ساختار کریستالی، تشکیل نشده و به جای آن ساختاری غیر کریستالی با سطح انرژی بالا تشکیل شود. بنابراین برای رسیدن به کمترین میزان سطح انرژی، فرآیند سرد کردن باید به آرامی صورت پذیرد (فابیان؛^۳ ۱۹۹۷). تابع کاهش دما و حرکت به سمت سرد شدن سیستم، نیازمند ضابطه‌ای است که در رابطه ۹ ارائه شده است.

$$T_i = \alpha \times (T_{i-1}) \quad , \quad \forall i=1, \dots, n \quad (9)$$

در رابطه فوق α ضریب تبرید یا ضریب کاهش دما است که مقدار آن کمتر از یک بوده و معمولاً بین ۰/۵ تا ۰/۹۹ در نظر گرفته می‌شود (کریک‌پاتریک و همکاران؛^۴ ۱۹۸۳). T نشان‌دهنده دمای سیستم و n تعداد دفعات کاهش سطح دمای سیستم است. روش الگوریتم SA از توزیع احتمالی بولتزمن، که در رابطه زیر نشان داده شده است استفاده می‌کند.

$$P(E) = \exp\left(-\frac{E}{k_b \times T}\right) \quad (10)$$

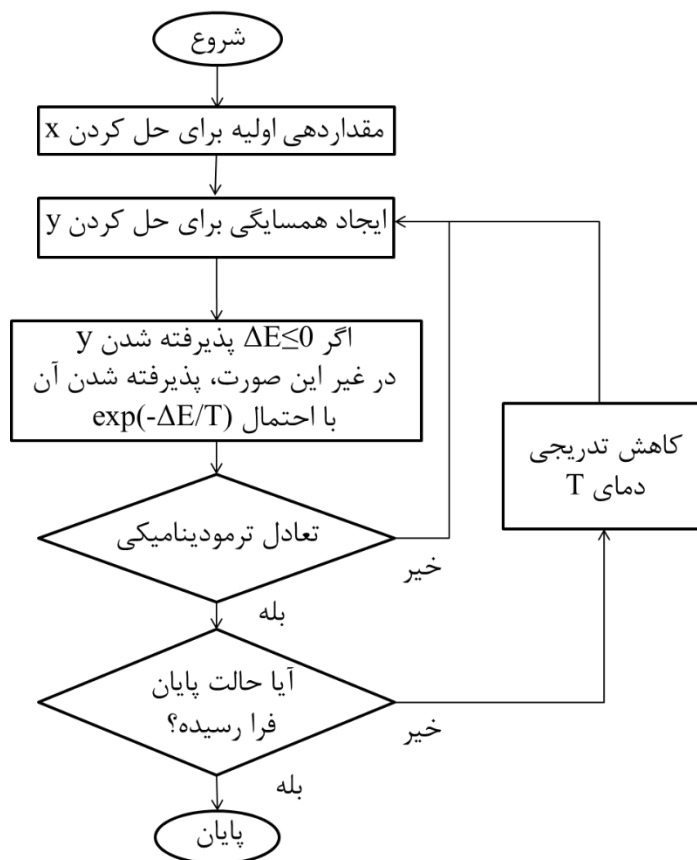
¹ Gonzalez et al

² Tran et al

³ Fabian

⁴ Kirkpatrick et al

E در رابطه فوق نشان‌دهنده انرژی سیستم و k_B نشان‌دهنده ثابت بولتزمن است. این توزیع بر این نکته تأکید می‌کند که وقتی یک سیستم در دمای T در تعادل گرمایی قرار دارد، یک توزیع انرژی وجود دارد که در بین کل حالات مختلف انرژی توزیع داده شده است. همیشه این امکان وجود دارد که حتی در یک دمای پایین، حالت انرژی سیستم بالا باشد. بنابراین شانس اینکه سیستم از یک مقدار انرژی مینیمم موضعی بیرون آمده و به یک مقدار مطلق همگرا شود وجود دارد (دوران و ویت، ۲۰۰۰). شکل ۳ فلوجارت الگوریتم SA را نشان می‌دهد. مطابق با شکل ۳ الگوریتم SA ابتدا از یک جواب اولیه (x_0) شروع می‌کند و سپس در یک حلقه تکرار t به سوی جواب‌های همسایه $N(x_t)$ حرکت می‌کند. اگر جواب همسایه بهتر از جواب فعلی باشد، الگوریتم آن را به‌عنوان جواب فعلی قرار می‌دهد. در غیر این صورت الگوریتم آن جواب را با احتمال $\exp(-\frac{\Delta E}{T})$ به‌عنوان جواب فعلی می‌پذیرد. در این رابطه ΔE تفاوت بین تابع هدف جواب فعلی و جواب همسایه است. در هر دما T ، چندین تکرار اجرا و سپس دما به آرامی کاهش داده می‌شود. در گام‌های اولیه دما خیلی بالا قرار داده می‌شود تا احتمال بیش‌تری برای پذیرش جواب‌های بدتر وجود داشته باشد. با کاهش تدریجی دما، در گام‌های پایانی احتمال کمتری برای پذیرش جواب‌های بدتر وجود دارد، بنابراین الگوریتم به سمت یک جواب خوب همگرا می‌شود (کریک‌پاتریک و همکاران، ۱۹۸۳). با توجه به توضیحات فوق درباره الگوریتم SA، برخلاف روش‌های بهینه‌سازی موضعی که فقط می‌توانند یک مقدار مینیمم نزدیک به حدس اولیه را پیدا کنند، این روش مقدار مینیمم مطلق را پیدا می‌کند (تران و همکاران، ۲۰۰۳).



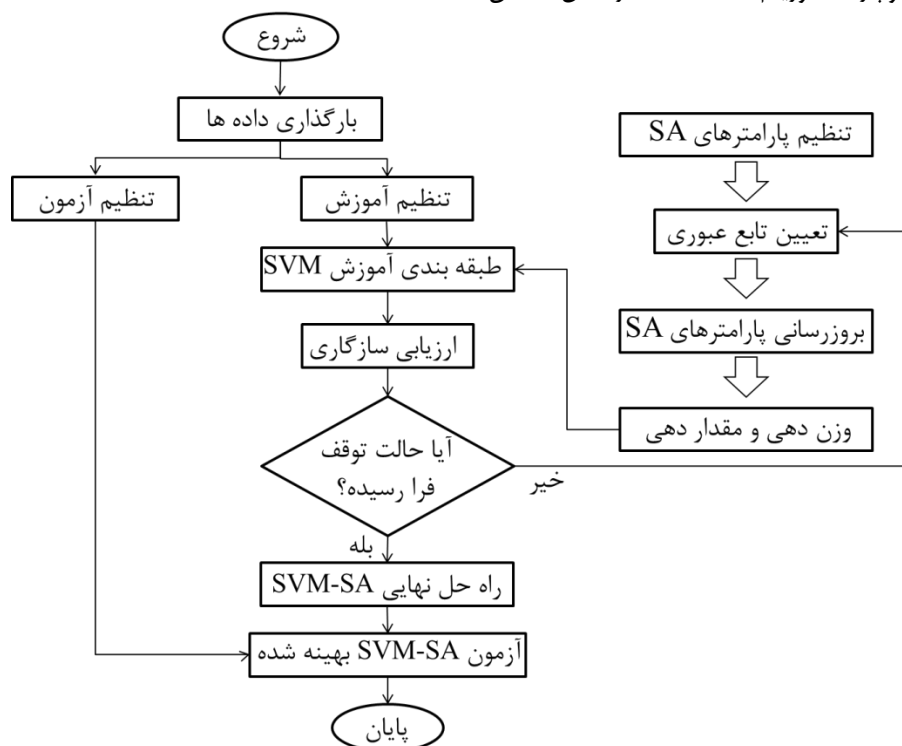
شکل ۳. فلوجارت الگوریتم SA

¹ Durand & White

الگوریتم هیبریدی شبیه‌سازی تبرید بر پایه ماشین بردار پشتیبان (SVM-SA)

الگوریتم SVM-SA، با تلفیق SVM و الگوریتم SA می‌تواند از زمان بر بودن مرحله سعی و خطا برای دریافت نتایج و کمینه شدن خطای تخمین جلوگیری نماید (تران و همکاران، ۲۰۰۳). در این روش در گام اول پارامترهای ورودی SVM تعیین و مقدار پارامترهای اولیه σ و γ مشخص می‌شوند. لذا در این روش در گام اول در الگوریتم SVM پارامترهای ورودی تعیین می‌شود و در گام دوم مقدار پارامترهای اولیه σ و γ مشخص می‌شوند. الگوریتم SVM به این صورت است که در فاز آموزش سعی می‌کند که مرز تصمیم‌گیری به گونه‌ای انتخاب گردد که حداقل فاصله آن با هر یک از طبقه‌بندی داده‌های موردنظر ماکزیمم شود به همین دلیل برای رسیدن به جواب بهینه به تکرارهای دیگر جهت یافتن بهترین پارامترهای اولیه برای شبیه‌سازی نیاز است.

در گام سوم برای جلوگیری از اتلاف زمان در تکرارهای متعدد با استفاده از الگوریتم SA به‌طور مستقیم با روش سیمپلکس مقداردهی جدید پارامترهای σ و γ به صورت بهینه (نقطه بهینه مطلق) تعیین می‌شود. در نهایت با جایگزین کردن پارامترهای به دست آمده از الگوریتم SA در الگوریتم SVM، بهترین حالتی که پایین‌ترین درصد خطا با تابع کرنل پایه‌ای شعاعی رادار است به دست آورده و با ترکیب دو الگوریتم بهینه-سازی برای بخش آموزش و آزمون با کمترین اختلاف خطا و بهترین مقادیر خطا با توجه به تعداد ورودی‌ها، مدل بهینه را می‌توان با دقت بالاتری تخمین زد. فلوجارت الگوریتم SVM-SA در شکل ۴ نشان داده شده است.



شکل ۴. فلوجارت الگوریتم SVM-SA

روش نرمال‌سازی داده‌ها

توزیع نرمال یکی از متداول‌ترین توزیع‌های آماری است که اولین بار توسط گوس^۱ ارائه و در تئوری خطای اندازه‌گیری به کار گرفته شد (بهرامی و همکاران، ۱۳۹۶). عملیات نرمال‌سازی قبل از بسیاری از الگوریتم‌های داده‌کاوی بایستی انجام گیرد تا ابعاد مختلف به صورت عادلانه توسط

^۱ Guos

الگوریتم بررسی شوند و تأثیر یکی بیشتر از بقیه نباشد. در این تحقیق از روش نرمال سازی تبدیل مینیمم- ماکزیمم که روشی رایج در هیدرولوژی است استفاده شده است.

این روش، تبدیل خطی را روی داده اصلی انجام می دهد و ارتباط بین مقادیر اولیه را حفظ می کند. اگر یک مقدار ورودی در خارج از محدوده اولیه قرار گیرد، خطای خارج محدوده رخ خواهد داد (ناظری تهرودی و همکاران، ۱۳۹۳). در این روش، داده ها به مجموعه جدیدی تبدیل می شوند که در آن، همه مقادیر بین صفر و یک هستند. فرمول کلی روش نرمال سازی تبدیل مینیمم - ماکزیمم برای تبدیل داده ها به بازه موردنظر به صورت رابطه زیر است (اکسوی و هارالیک، ۲۰۰۱):

$$Y_t = \frac{x_t - x_{min}}{x_{max} - x_{min}} \quad (11)$$

در رابطه فوق Y_t مقدار نرمال شده، x_t مقدار هر یک از داده ها، x_{min} کمترین مقدار داده ها و x_{max} بیشترین مقدار داده ها است. یعنی مقیاس هر دو ویژگی در بازه ۰ تا ۱ قرار گرفته و حالا می توان الگوریتم های مختلف خوشه بندی یا طبقه بندی را بر روی آن ها به صورت منصفانه اجرا کرد.

شاخص های آماری

به منظور ارزیابی و اعتبار یابی نتایج مدل، از شاخص های آماری ریشه میانگین مربعات خطا (RMSE)، خطای استاندارد (SE) و ضریب تبیین (R^2) استفاده شد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_m - Y_p)^2}{n}} \quad (12)$$

$$SE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_m - Y_p)^2} \quad (13)$$

$$R^2 = \frac{[\sum_{i=1}^n (Y_m - \bar{Y}_m)(Y_p - \bar{Y}_p)]^2}{\sum_{i=1}^n (Y_m - \bar{Y}_m)^2 \cdot \sum_{i=1}^n (Y_p - \bar{Y}_p)^2} \quad (14)$$

در این روابط، n تعداد روزهای دوره مورد مطالعه Y_m مقادیر اندازه گیری شده، Y_p مقدار پیش بینی شده با استفاده از مدل، \bar{Y}_m میانگین داده های اندازه گیری شده و \bar{Y}_p میانگین داده های پیش بینی شده است.

یافته ها

همان گونه که پیشتر اشاره شد، در این تحقیق با استفاده از الگوریتم SVM-SA به پیش بینی بارندگی پرداخته شد. جهت تخمین بارندگی در الگوریتم SVM-SA از داده های غیر بارندگی ایستگاه سینوپتیک گرگان استفاده شده است. روند کار شامل سه مرحله است: در مرحله اول به منظور تعیین پارامترهای هواشناسی مؤثر بر مقدار بارندگی از داده های هواشناسی غیر بارندگی شامل ابرناکی، میانگین دمای حداکثر، میانگین دمای حداقل، فشار بخار آب، سرعت باد، رطوبت نسبی حداکثر، رطوبت نسبی حداقل، نقطه شبنم استفاده شد. جدول ۱ خلاصه آماری پارامترهای مورد استفاده شامل میانگین، حداقل، حداکثر، چولگی و کشیدگی پارامترهای هواشناسی مورد استفاده را نشان می دهد.

¹ Aksoy & Haralick

جدول ۱. نتایج مشخصات آماری پارامترهای هواشناسی از مدل SPSS

پارامتر	واحد	میانگین	حداقل	حداکثر	چولگی	کشیدگی
ابرناسی	ساعت	۱۴/۱	۹۱/۰	۱۴۹/۰	۰/۳۳	-۰/۷۴۸
میانگین دمای حداکثر	درجه سانتی گراد	۲۲/۸	۲۱/۵	۲۵/۱	۰/۱۷	۰/۰۸۳
میانگین دمای حداقل	درجه سانتی گراد	۱۲/۷	۱۱/۲	۱۴/۵	۰/۱۶	-۰/۲۵۲
فشار بخار آب	هکتو پاسکال	۱۴/۹	۱۳/۶	۱۶/۱	-۰/۲۸	-۰/۷۴۰
سرعت باد	متر بر ثانیه	۲/۲	۰/۵	۵/۷	۱/۰	-۰/۳۹۲
رطوبت نسبی حداکثر	درصد	۸۶/۰	۷۵/۰	۹۳/۰	-۰/۸	-۰/۳۸۵
رطوبت نسبی حداقل	درصد	۵۳/۳	۴۸/۰	۵۸/۰	-۰/۱۰	-۰/۵۸۹
نقطه شبنم	درجه سانتی گراد	۱۱/۵	۹/۹	۱۲/۸	-۰/۴۶	-۰/۶۲۲
بارندگی	میلی متر	۵۵۸	۳۱۴/۶	۷۹۲/۶	-۰/۰۰۹	-۰/۳۵۳

نتایج جدول ۱ نشان می‌دهد، میانگین بارندگی در منطقه مورد مطالعه ۵۵۸ میلی‌متر است. حداکثر دمای هوای ثبت شده طی دوره مورد مطالعه در سال ۱۳۸۷ برابر ۲۵/۱ درجه سانتی‌گراد مشاهده شد. میانگین کمترین دمای هوای ثبت شده طی ۴۰ سال معادل ۱۱/۲ درجه سانتی‌گراد است. مقدار چولگی و کشیدگی داده‌ها بیان‌گر توزیع نرمال پارامترهای مورد مطالعه می‌باشند.

در مرحله دوم، قبل از استفاده داده‌ها باید از نرمال بودن آن‌ها مطمئن شد. به دلیل نرمال نبودن داده‌ها، نرمال‌سازی داده‌های هواشناسی غیر بارندگی با استفاده از روش تبدیل مینیمم-ماکزیمم (رابطه ۱۱) انجام شد. در نظر گرفتن بازه ۰ تا ۱ جهت نرمال‌سازی، باعث می‌شود شاخص‌های آماری با خطای کمتر و دقت بالاتری برآورد شوند. نتایج نرمال‌سازی نشان داد این روش از دقت بالایی جهت نرمال نمودن و آماده‌سازی داده‌ها برخوردار است.

همچنین در تحقیقی بهرامی و همکاران (۱۳۹۶) به تعیین اثر پیش پردازش داده بر عملکرد شبکه عصبی مصنوعی به منظور پیش‌بینی بارندگی ماهانه در شهرستان آبادیه پرداختند. ایشان برای نرمال‌سازی داده‌ها از سه روش تبدیل مینیمم-ماکزیمم، نرمال‌سازی رتبه‌ای و تبدیل نرمال استاندارد استفاده کردند. نتایج نشان داد روش مینیمم-ماکزیمم در مقایسه با روش‌های دیگر به‌عنوان بهترین روش در نرمال‌سازی داده‌ها جهت شبیه‌سازی در شبکه عصبی انتخاب شد.

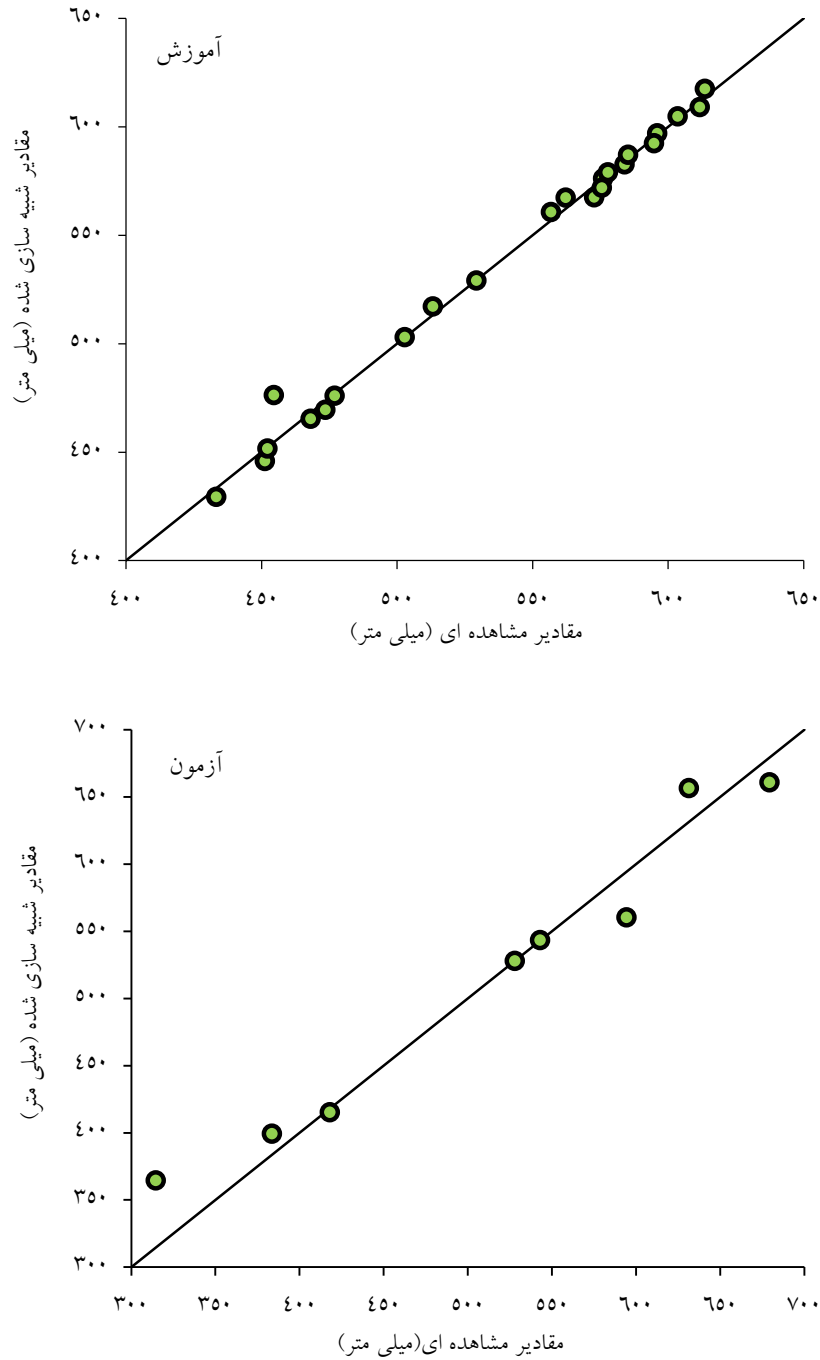
در نهایت، در مرحله سوم پس از نرمال‌سازی داده‌ها جهت تعیین پارامترهای مؤثر با استفاده از نرم‌افزار SPSS، مقدار ضریب همبستگی پیرسون بین پارامترهای هواشناسی و بارندگی در سطح معنی‌داری ۹۹ درصد بررسی شد. جدول ۲ نتایج ضریب همبستگی پارامترهای هواشناسی، به‌منظور انتخاب پارامترهای ورودی را نشان می‌دهد.

جدول ۲. ماتریس همبستگی پارامترهای هواشناسی در نرم افزار SPSS

ابرناکی	میانگین دمای حداکثر	میانگین دمای حداقل	فشار بخار آب	سرعت باد	رطوبت نسبی حداکثر	رطوبت نسبی حداقل	نقطه شبنم	بارندگی
۱								
میانگین دمای حداکثر	۱							
میانگین دمای حداقل	۰/۵۲۸**	۱						
فشار بخار آب	۰/۵۱۴**	۰/۰۷۰	۱					
سرعت باد	۰/۵۶۵**	-۰/۰۳۷	۰/۴۶۵**	۱				
رطوبت نسبی حداکثر	۰/۱۳۲	-۰/۴۴۱**	۰/۷۸۱**	۰/۳۶۴**	۱			
رطوبت نسبی حداقل	-۰/۴۷۲**	-۰/۴۷۴**	-۰/۴۴۱**	-۰/۰۹۷	-۰/۶۷۷**	۱		
نقطه شبنم	-۰/۵۳۷**	۰/۰۸۲	-۰/۹۷۸**	-۰/۴۹۲**	۰/۸۰۰**	-۰/۴۴۸**	۱	
بارندگی	-۰/۵۰۲**	۰/۱۹۳	۰/۴۷۱**	-۰/۴۰۱*	-۰/۴۸۷**	۰/۰۵۱	-۰/۴۴۳**	۱

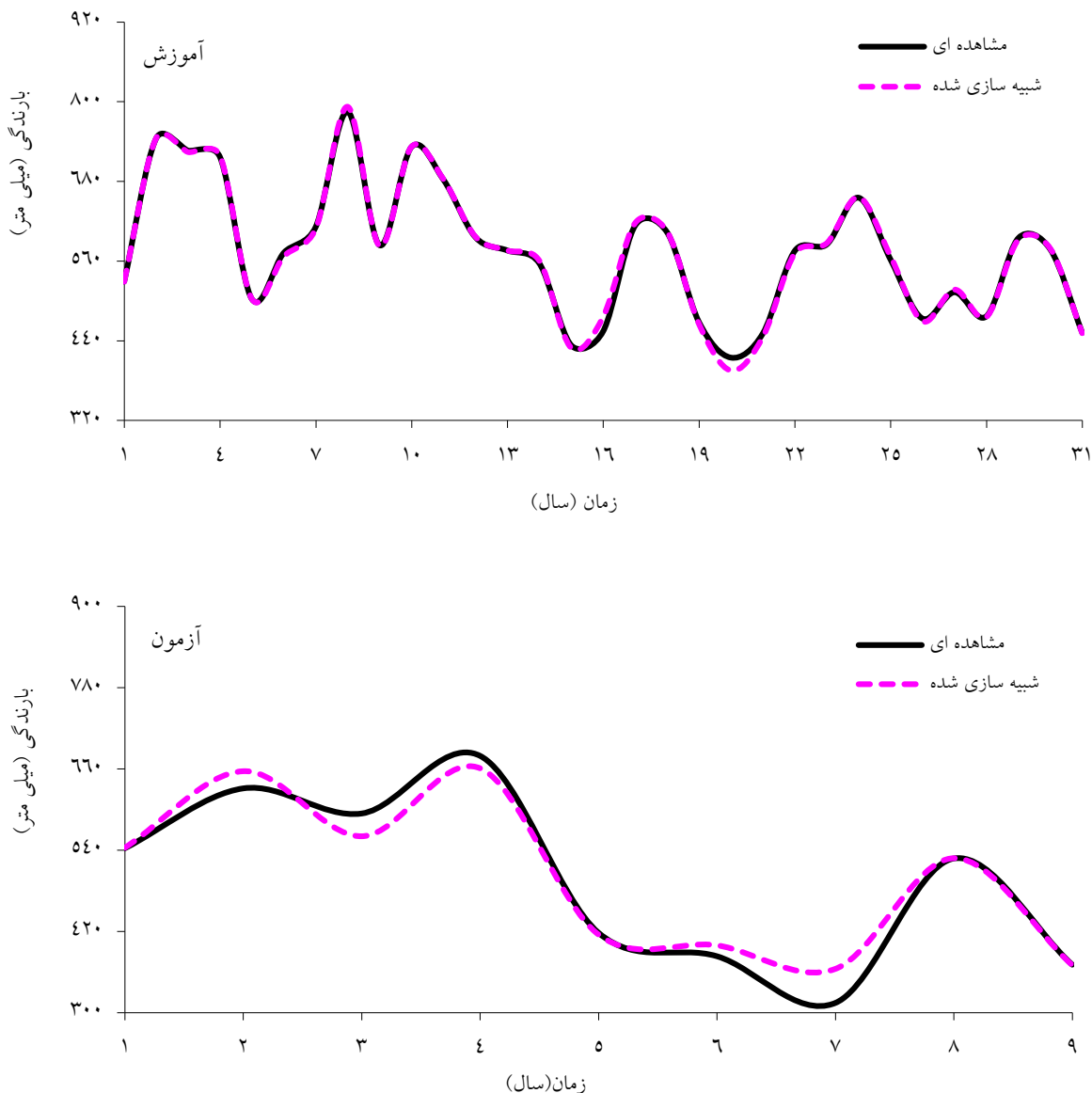
** پارامترهای پرننگ در سطح ۹۹ درصد معنی دار می باشند و به عنوان ورودی نرم افزار SPSS، برای پارامتر مؤثر بر بارندگی مورداستفاده قرار می گیرند.

همان گونه که جدول ۲ نشان می دهد، پارامترهای ابرناکی، میانگین دمای حداکثر، فشار بخار آب، رطوبت نسبی حداکثر و نقطه شبنم بیشترین همبستگی معنی دار در سطح ۹۹ درصد را با بارندگی دارند که در جدول به صورت پرننگ و دو ستاره نشان داده شده اند. بنابراین از ۸ پارامتر، ۵ پارامتر فوق به عنوان پارامترهای ورودی برای پیش بینی پارامتر بارندگی در الگوریتم SVM-SA در نظر گرفته شد. پس از انتخاب پارامترهای ورودی مناسب، در الگوریتم SVM-SA به تعیین مقادیر σ و γ که پارامترهای اولیه مدل بهینه SVM می باشند، پرداخته شد. در این راستا، ۷۵ درصد داده ها در بخش آموزش و ۲۵ درصد داده ها در بخش آزمون مورداستفاده قرار گرفتند. در ابتدا مقادیر σ و γ به ترتیب برابر با ۳/۰۷ و ۴۹۸/۹۶ فرض شد و در نهایت پس از اجرای مدل، مقادیر بهینه محاسبه شده توسط الگوریتم SA به ترتیب برابر با ۲/۵۴ و ۴۲۳/۲۸ محاسبه گردید. با جایگزین کردن مقادیر بهینه این دو پارامتر در الگوریتم SVM، شبیه سازی انجام شد. شکل ۵ پیش بینی پارامتر بارندگی با مدل الگوریتم SVM-SA را در بخش آموزش و آزمون در مقابل داده های مشاهده ای نشان می دهد.



شکل ۵. مقادیر شبیه سازی شده تغییرات بارندگی با مدل SVM-SA در مرحله آموزش و آزمون

همچنین شکل ۶، سری زمانی تغییرات مقادیر بارندگی شبیه سازی شده توسط مدل الگوریتم SVM-SA در بخش آموزش و آزمون در برابر داده های مشاهده ای را نشان می دهد.



شکل ۶. تغییرات سری زمانی مقادیر بارندگی مشاهده‌ای در برابر مقادیر شبیه‌سازی‌شده با مدل SVM-SA در مرحله آموزش و آزمون

مطابق با شکل‌های ۵ و ۶ ملاحظه می‌شود که تغییرات بارندگی شبیه‌سازی‌شده، تطابق بسیار خوبی با مقادیر مشاهده‌ای دارد. همچنین برای ارزیابی دقت عملکرد مدل SVM-SA، از شاخص‌های آماری RMSE، SE و R^2 استفاده شد. در جدول ۳ مقدار شاخص‌های آماری در بخش‌های آموزش و آزمون نشان داده شده است. مقادیر SE بین صفر و یک متغیر بوده و هرچه قدر مقدار این پارامتر به صفر نزدیک‌تر باشد، نشان‌دهنده تطابق خوب بین مقادیر مشاهده‌ای و شبیه‌سازی‌شده است. لذا با توجه به جدول ۳ بر اساس نتایج شاخص‌های آماری می‌توان انطباق بین نتایج شبیه‌سازی‌شده و داده‌های مشاهده‌ای در تحقیق حاضر را با روش ترکیبی SVM-SA بسیار خوب ارزیابی کرد که نشان از ارتباط قوی بین مقادیر برآوردی و مشاهده‌ای است.

جدول ۳. مقادیر شاخص‌های آماری تغییرات پارامتر بارندگی برای مدل SVM-SA

آزمون			آموزش			مرحله
R ²	SE	RMSE	R ²	SE	RMSE	شاخص
۰/۹۲۵	۰/۰۳	۱۸/۷۲	۰/۹۹۹	۰/۰۱	۶/۰۲	مقدار

بحث

بارندگی یکی از مهم‌ترین داده‌های ورودی به سیستم‌های هیدرولوژیکی است. با توجه به اینکه کشور ایران جزء کشورهای خشک به شمار می‌رود، برآورد صحیح میزان بارندگی اهمیت بسزایی دارد. این تحقیق باهدف برآورد بارندگی و استفاده از متغیرهای ورودی پارامترهای هواشناسی غیر بارندگی طی دوره ۴۰ ساله ۱۹۷۱ تا ۲۰۱۰ با استفاده از SVM-SA در منطقه گرگان صورت پذیرفت. پس از نرمال‌سازی داده‌ها، جهت تعیین پارامترهای مؤثر با استفاده از نرم‌افزار SPSS، مقدار ضریب همبستگی پیرسون بین پارامترهای هواشناسی و بارندگی در سطح معنی‌داری ۹۹ درصد بررسی گردید. پس از شبیه‌سازی، میزان برازش میان مقادیر مشاهده‌ای و شبیه‌سازی‌شده، با شاخص‌های آماری RMSE، SE و R² محاسبه شد.

نتیجه‌گیری

بارش یکی از مهم‌ترین داده‌های ورودی به سیستم‌های هیدرولوژیکی است. با توجه به اینکه کشور ایران جزء کشورهای خشک به شمار می‌رود، برآورد صحیح میزان بارش اهمیت بسزایی دارد. این تحقیق باهدف برآورد بارش و استفاده از متغیرهای ورودی پارامترهای هواشناسی غیربارشی طی دوره ۴۰ ساله ۱۹۷۱ تا ۲۰۱۰ با استفاده از SVM-SA در منطقه گرگان صورت پذیرفت. پس از نرمال‌سازی داده‌ها، جهت تعیین پارامترهای مؤثر با استفاده از نرم‌افزار SPSS، مقدار ضریب همبستگی پیرسون بین پارامترهای هواشناسی و بارندگی در سطح معنی‌داری ۹۹ درصد بررسی گردید.

نتایج مقدار شاخص‌های آماری RMSE، SE و R² در پیش‌بینی تغییرات بارندگی با استفاده از داده‌های هواشناسی در بخش آموزش به ترتیب برابر با ۶/۰۲ میلی‌متر، ۰/۰۱ و ۰/۹۹۹ است. همچنین مقادیر این شاخص‌ها در بخش آزمون برابر با ۱۸/۷۲ میلی‌متر، ۰/۰۳ و ۰/۹۲۵ به دست آمد. حلییان و همکاران (۱۳۹۱) با استفاده از داده‌های هواشناسی در پیش‌بینی بارندگی ماهانه اصفهان با شبکه عصبی مصنوعی میزان شاخص آماری R² در شبیه‌سازی بارندگی را ۰/۸۷ گزارش نمودند. بنابراین با توجه به نتایج به‌دست‌آمده، می‌توان انطباق بین نتایج شبیه‌سازی‌شده و داده‌های مشاهده‌ای در تحقیق حاضر را بسیار خوب ارزیابی کرد.

محمدی و حیدرپور (۱۳۹۸) اذعان داشتند که الگوریتم SVM-SA می‌تواند به‌عنوان ابزاری کارآمد در شبیه‌سازی نسبت به شبیه‌سازی مدل‌های دیگر باشد که با نتایج حاصل از تحقیق حاضر تطابق دارد. همچنین تحقیق محمدی و مؤذن‌زاده (۱۳۹۸) نیز نشان داد که در بین مدل‌های مختلف با ساختار ورودی یکسان، شبیه‌سازی الگوریتم ترکیبی SVM-SA نسبت به سایر مدل‌های هیبریدی هوشمند عملکرد بهتری برای در شبیه‌سازی دارد. لذا با توجه به این تحقیقات و نتایج به‌دست‌آمده از تحقیق حاضر می‌توان گفت شبیه‌سازی در روش ترکیبی SVM-SA از دقت بالایی برخوردار است.

نتایج نشان داد که با یافتن روابط بین پارامترهای هواشناسی می‌توان دقت برآورد الگوریتم را افزایش داد. با توجه به قابلیت و دقت بالای مدل ترکیبی در تخمین بارندگی می‌توان گفت نتیجه‌گیری این تحقیق با نتیجه‌گیری سلیمی‌سرتاخطی و همکاران (۲۰۱۱)، یوتاو و لین (۲۰۲۱) و مارک و همکاران (۲۰۲۱) تطابق دارد و این مدل را به‌عنوان الگوریتم هیبریدی جدید در حوزه علوم مهندسی می‌توان معرفی کرد.

¹ Salimisartakhti et al

² Youtao & Lin

³ Mahareek et al

از قابلیت مدل ترکیبی تهیه شده در این تحقیق، علاوه بر شناسایی روابط بین متغیرهای ورودی و خروجی، می توان به سرعت بالا در محاسبات موازی، مقاومت در برابر تغییر پارامترها و قابلیت تعمیم اشاره کرد. لذا می توان استفاده از این مدل را در مدیریت منابع آب و برنامه ریزی به منظور استفاده بهینه از منابع آب منطقه توصیه نمود.

تقدیر و تشکر

در پایان از سازمان هواشناسی و شرکت آب منطقه ای استان گلستان که اطلاعات مورد نیاز این پژوهش را در اختیار قرار دادند، سپاسگزاری می گردد.

منابع

- بهرامی، مهدی، امیری، محمدجواد، رضایی مہارلویی، فاطمه، و غفاری، کرامت اله. (۱۳۹۶). تعیین اثر پیش پردازش داده بر عملکرد شبکه عصبی مصنوعی به منظور پیش بینی بارندگی ماهانه در شهرستان آباء. نشریه اکوهیدرولوژی، ۴(۱)، ۳۷-۲۹. <https://dx.doi.org/10.22059/IJE.2017.60880.29-37>
- حلبیان، امیر، حسین، و دارند، محمد. (۱۳۹۱). پیش بینی بارندگی اصفهان با استفاده از شبکه های عصبی مصنوعی. نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۲(۲۶). <https://www.sid.ir/paper/102413/fa>
- شرفی، میلاد، صمدیان فرد، سعید، و هاشمی، سجاد. (۱۳۹۹). پیش بینی بارندگی ماهانه با استفاده از مدل های برنامه ریزی ژنتیک و ماشین بردار پشتیبان. نشریه سامانه های نوین سطوح آبگیر باران، ۸(۴)، ۷۱-۶۳. <https://www.sid.ir/paper/958629/fa>
- ضابط پیشخانی، نسرين، سیدیان، سیدمرتضی، حشمت پور، علی، و روحانی، حامد. (۱۳۹۵). مقایسه الگوسازی بارندگی ماهانه با مدل های SVM و ANFIS (مطالعه موردی: شهر گنبد کاووس). نشریه آب و خاک (علوم و صنایع کشاورزی)، ۳۰(۱)، ۲۴۶-۲۳۶. <https://civilica.com/doc/666817.236-246>
- عیسی زاده، محمد، احمدزاده، حجت، و قربانی، محمدعلی. (۱۳۹۶). ارزیابی تأثیر نرمال سازی توزیع احتمالاتی رواناب ماهانه بر عملکرد مدل های SVM و ANN در شبیه سازی جریان ماهانه رودخانه ها (مطالعه موردی: حوزه زربینه رود). پژوهشنامه مدیریت حوزه آبخیز، ۸(۱۶)، ۳۳-۲۲. https://jsw.um.ac.ir/article_38332.html
- محمدی، بابک، و مودن زاده، روزبه. (۱۳۹۸). پیش بینی جریان جریان با استفاده از مدل های ترکیبی هوشمند در مقیاس ماهانه (مطالعه موردی: رودخانه زربینه رود). فصلنامه علوم و تکنولوژی محیط زیست، ۲۱(۹)، ۸۱-۷۱. <https://dx.doi.org/10.22034/JEST.2020.24315.3331>
- محمدی، بابک، و حیدرپور، باقر. (۱۳۹۸). بررسی تأثیر سنجه های اقلیمی بر روی تخمین دمای خاک توسط الگوریتم هیبریدی شبیه سازی تبرید. مجله علمی و ترویجی نیوار، ۱۰۵-۱۰۴. <https://doi.org/10.30467/nivar.2019.83519.104-105>
- ناظری تهرودی، محمد، خلیلی، کیوان، عباس زاده افشار، مرضیه، و ناظری تهرودی، زهرا. (۱۳۹۳). مقایسه تبدیل های نرمال ساز جهت نرمال کردن داده های بارندگی ماهانه مناطق مختلف ایران. نشریه آب و خاک (علوم و صنایع کشاورزی)، ۲۸(۲)، ۳۷۲-۳۶۵. <https://www.sid.ir/paper/141631/fa>
- نوذری، حامد، و توکلی، فاطمه. (۱۳۹۸). ارزیابی کارایی مدل های مختلف خطی و غیرخطی در پیش بینی بارندگی ماهانه در تغییرات اقلیم استان همدان. پژوهشنامه مدیریت حوزه آبخیز، ۱۰(۲۰)، ۱۲-۱۱. <https://doi.org/10.29252/jwmr.10.20.19.1-12>

References

- Aksoy, S., & Haralick, M. (2001). Feature Normalization and Likelihood-based Similarity Measures for Image Retrieval. *Journal Pattern Recognition Letters*, 22(5), 563-582. [https://doi.org/10.1016/S0167-8655\(00\)00112-4](https://doi.org/10.1016/S0167-8655(00)00112-4).
- Asefa, T., Kemblowski, M., McKee, M., & Khalil, A. (2006). Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology*, 318, 7-16. <https://doi.org/10.1016/j.jhydrol.2005.06.001>

- Bahrami, M., Amiri, M., J., RezaeiMaharlouei, F., & Ghafari, K.A. (2017). Determining the effect of data preprocessing on the performance of artificial neural network in order to predict monthly rainfall in Abadeh city. *Journal of Ecohydrology*, 4(1), 29-37. <https://dx.doi.org/10.22059/IJE.2017.60880> [in Persian]
- Chen, S., T., Yu, P., Sh., & Tang., H.,Y. (2010). Statistical downscaling of daily Precipitation using support vector machines and multivariate analysis. *Journal of Hydrology*, 385, 13-23. <https://doi.org/10.1016/j.jhydrol.2010.01.021>
- Chen, sh. (2016). Application Hydrologic Data Mining Using Artificial Neural Network. *Entropy*. 12, 83-98. <https://www.earticle.net/Article/A241921>
- Chia, M., Y., Huang, Y., F., & Koo, C., H. (2020). Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. *Journal Computers and Electronics in Agriculture*, 175, 105577. <https://doi.org/10.1016/j.compag.2020.105577>
- Dibike, Y., B., Velickov, S., Solomatine, D., & Abbott, M., B. (2001). Model induction with support vector machines: introduction and applications. *Journal of Computing in Civil Engineering*, 15, 208-216. [https://doi.org/10.1061/\(ASCE\)0887-3801\(2001\)15:3\(208\)](https://doi.org/10.1061/(ASCE)0887-3801(2001)15:3(208))
- Durand, M., D., & White, S., R. (2000). Trading accuracy for speed in parallel simulated annealing with simultaneous moves. *Elsevier Parallel Computing*, 26, 135-150. [https://doi.org/10.1016/S0167-8191\(99\)00099-X](https://doi.org/10.1016/S0167-8191(99)00099-X)
- Eysazadeh, M., Ahmadzadeh, H., & Ghorbani, M., A. (2017). Evaluation of the effect of normalizing the distribution of monthly runoff probabilities on the performance of SVM and ANN models in simulating the monthly flow of rivers (Case study: Zarineh Rood basin). *Watershed Management Journal*, 8(16), 22-33. https://jsw.um.ac.ir/article_38332.html [in Persian]
- Fabian, V. (1997) Simulated Annealing Simulated. *Computers & Mathematics with Applications*, 33, 81-94. [https://doi.org/10.1016/S0898-1221\(96\)00221-0](https://doi.org/10.1016/S0898-1221(96)00221-0)
- Fletcher, R. (1987). *Practical Methods of Optimization*. 2nd Edition, John Wiley and Sons, Chichester. [https://www.scirp.org/\(S\(351jmbntvnsjt1aadkozje\)\)](https://www.scirp.org/(S(351jmbntvnsjt1aadkozje)))
- Gonzalez, R., C., & Perez, V., S. (2001). Two procedures for stochastic simulation of vuggy formations. (2001) SPE 69663, Latin American and Caribbean Petroleum Engineering Conference, Buenos Aires, Argentina, pp. 25-28 March. <https://doi.org/10.2118/69663-MS>
- Guo, H., C., Liu, L., Huang, G., H., Fuller, G., A., Zou, R., & Yin, Y., Y. (2001). A system dynamics approach for regional environmental planning and management: A study for the Lake Erhai Basin. *Journal of Environmental Management*, 61, 93-111. <https://doi.org/10.1006/jema.2000.0400>
- Halabian, A., H., & Darend, M. (2012). Rainfall forecasting in Isfahan using artificial neural networks. *Applied Research Journal of Geographical Sciences*, 12(26). <https://www.sid.ir/paper/102413/fa> [in Persian]
- Hamidi, O., J., Poorolajal, M., Sadeghifar, H., Abbasi, Z., Maryanaji, H., Faridi, R., & Tapak, L. (2014) . A comparative study of support vector machines and artificial neural networks for predicting precipitation in Iran. *Theoretical and Applied Climatology*, 119, 723-731. <https://link.springer.com/article/10.1007/s00704-014-1141-z>
- Harris, R., & Jarvis, C. (2011). *Statistics for Geography and Environmental Science*. ed: 1st Edition, London ,Routledge, 280. <https://doi.org/10.4324/9781315847610>
- Jia, Zh, y., Ma, J, w., Wang, F, J., & Liu, W. (2011). Hybrid of simulated annealing and SVM for hydraulic valve characteristics prediction. *Journal Expert Systems with Applications*, 38, 8030-8036. <https://doi.org/10.1016/j.eswa.2010.12.132>

- Kirkpatrick, S., Gelatt, C., D., & Vecchi, M., P. (1983). Optimization by simulated annealing. *Science*, 220, 671–680. <https://doi.org/10.1126/science.220.4598.671>
- Lin, S., W., Lee, Z., J., Chen, S., C., & Tseng, T., Y. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied Soft Computing*, 8(4), 1505-1512. <https://doi.org/10.1016/j.asoc.2007.10.012>
- Mahareek, E., A., Desuky, A., S., & AbdullahEl-Zhni, H. (2021). Simulated annealing for SVM parameters optimization in student performance prediction. *Bulletin of Electrical Engineering and Informatics*, 10(3), 1211-1219. <https://doi.org/10.11591/eei.v10i3.2855>
- Mehrabani Bashar, A., Nozari, H., Marofi, S., Mohamadi, M., & Ahadiiman, A. (2022). Investigation of factors affecting rural drinking water consumption using intelligent hybrid models. *Journal Water Science and Engineering*. <https://www.sciencedirect.com/journal/water-science-and-engineering>
- Mohammadi, B., & Heydarpour, B. (2019). Investigating the influence of climate parameters on the estimation of soil temperature by hybrid refrigeration simulation algorithm. *Nivar scientific and promotional magazine*, 105-104. <https://doi.org/10.30467/nivar.2019.83519> [in Persian]
- Mohammadi, B., & Mouzenzadeh, R. (2019). Stream flow forecasting using intelligent hybrid models on a monthly scale (case study: Zarin Rood River). *Environmental Science and Technology Quarterly*, article, 21 (9), 71-81. <https://dx.doi.org/10.22034/JEST.2020.24315.3331> [in Persian]
- Nazeri Tahrudi, M., Khalili, K., AbbaszadeAfshar, M., & NazeriTahrudi, Z. (2014). Compared to the normal mechanism becomes the normal monthly rainfall data from different regions of Iran. *Journal of Water and soil*, 28 (2), 365- 372. <https://www.sid.ir/paper/141631/fa> [in Persian]
- Nouri, R., Karbassi, A., R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M., H., Farokhnia, A., & GhafariGousheh, M. (2011). Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology*, 401, 177-189. <https://doi.org/10.1016/j.jhydrol.2011.02.021>
- Nozari., H., & Tavakoli., F. (2019). Evaluation of the Efficiency of Linear and Nonlinear Models in Predicting Monthly Rainfall (Case Study: Hamedan Province). *jwmr*, 10(20), 1-12. <https://doi.org/10.29252/jwmr.10.20.19> [in Persian]
- Pai, P., F., & Hong,W., CH. (2006).Software reliability forecasting by support vector machines with simulated annealing algorithms. *Journal of Systems and Software*, 79, 747–755. <https://doi.org/10.1016/j.jss.2005.02.025>
- Palmer, T. (2020). A vision for numerical weather prediction in 2030.arXiv preprint arXiv:2007.04830.
- Polat, K., & Gunes., S. (2007). Classification of epileptic form EEG using a hybrid system based on decision tree classifier and fast Fourier transform. *Applied Mathematics and Computation*, 187(2), 1017-1026. <https://doi.org/10.1016/j.amc.2006.09.022>
- Salimisartakhti, J., Zangooei, M., H., & Mozafari, K. (2012) Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA). *Computer methods and programs in biomedicine*, 108, 570–579. <https://doi.org/10.1016/j.cmpb.2011.08.003>
- Sharfi, M., SamadianFard, S., & Hashemi, S. (2020). Forecasting monthly rainfall using genetic programming models and support vector machine. *Rain catchment surface systems*, 8(4), 71-63. <https://www.sid.ir/paper/958629/fa> [in Persian]
- Tran, N., H., Chen, Z., & Rahman, S., S. (2003). Characterizing and Modeling of Naturally Fractured Reservoirs with the use of Object-Based Global Optimization. *Petroleum Society's Canadian International Petroleum Conference*, Calgary, Alberta, Canada, 10 – 12. <https://doi.org/10.2118/2003-179>

- Tripathi, Sh., Srinivas, V., & Nanjundiah, R., S. (2006). Downscaling of precipitation for climate change scenarios: A support vector machine approach. *Journal of Hydrology*, 330, 62- 640. 21- Vapnik V.N. 1998. *Statistical learning theory*. 1st Ed., Wiley, New York. <https://doi.org/10.1016/j.jhydro.2006.04.030>
- Wang, S., Feng, J., Liu, G, J., M., & Modelling, C. (2013). Application of seasonal time series model in the precipitation forecast. 58(3-4), 677-683. <https://doi.org/10.1016/j.mcm.2011.10.034>
- Youtao, X., & Lin, C. (2021). Forecast of natural gas market yield based on data analysis. *Resources & Industries*, 23(1), 79-86. <https://doi.org/10.13776/j.cnki.resourcesindustries.20201211.001>
- Yu, X., & Liong, S.,Y. (2006). Forecasting of hydrologic time series with ridge regression in feature space. *Journal of Hydrology*, 332, 290-302. <https://doi.org/10.1016/j.jhydro.2006.07.003>
- ZabetPishkhani, N., Seydian, S., M., Heshmatpour, A., & Rouhani, H. (2016). Comparison of monthly rainfall patterning with SVM and ANFIS models (case study: City Gonbadkavos). *Water and soil magazine*. 30(1), 236-246. <https://civilica.com/doc/666817> [in Persian]