





Analysis of drought characteristics using a copula-based method and a hybrid of soft computing algorithms (Case study: Qazvin Station)

Hamed Kiafar¹ , Hossein Babazadeh^{2*} , Hossein Sedghi³ , Ali Saremi⁴ 

¹ Ph.D. Graduated, Department of Water Science and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

² Professor, Department of Water Science and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

² Former Professor, Department of Water Science and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

⁴ Assistant Professor, Department of Water Science and Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

Extended Abstract

Introduction

Drought is a natural phenomenon that has significant negative impacts on various sectors, particularly due to changes in rainfall patterns—such as quantity, intensity, and distribution—which have led to annual droughts in certain regions. Climate change further complicates this issue by contributing to rising sea levels and altered weather conditions. The gradual onset of drought affects critical areas like water resources, agriculture, and the environment, yet the complex physical factors involved are not well understood, and reliable models are lacking. To better understand and predict drought, researchers often employ statistical methods in hydrological studies. Joint probability functions are effective for analyzing multiple variables while preserving their correlations. Additionally, the use of soft computing algorithms has become increasingly popular for prediction and estimation. Over the past two decades, advancements in data mining techniques have enabled researchers to extract valuable insights from large datasets and create predictive models. These methods are advantageous as they offer greater accuracy and efficiency compared to traditional numerical modeling without incurring laboratory costs. This study specifically examines drought occurrences in Qazvin Province by integrating joint probability functions and soft computing algorithms to analyze random data. The goal is to improve understanding and management of drought impacts in the region, thereby contributing to more effective responses to this pressing issue.

Materials and Methods

This article analyzes monthly rainfall data from January 1964 to December 2018 at the Qazvin Synoptic Station to identify drought characteristics in Qazvin Province. The average annual rainfall recorded at the station is 304 millimeters. The study employs a combination of copula functions and algorithms to model meteorological drought, utilizing randomly selected data for analysis. Specifically, the M5 algorithm is used to generate rainfall data, while various soft computing techniques—including Particle Swarm Optimization (PSO), genetic algorithms, CART, GEP, and GMDH algorithms—are applied. For drought analysis, the Standardized Precipitation Index (SPI) is utilized, as it is recognized as an effective measure for assessing drought conditions. The SPI time series derived from monthly rainfall data helps define and calculate key drought characteristics such as duration and severity. Drought is identified when the SPI falls below -1, with moist and dry conditions classified accordingly. Notably, some drought events may exhibit long durations without consistently reaching an SPI of -1 or lower; however, cumulative SPI values can still indicate drought conditions. A drought event is characterized by a continuous period of negative SPI values, with drought duration defined as the length of this negative period and severity measured by cumulative SPI during the drought. This research aims to enhance understanding of drought dynamics in Qazvin Province for better management strategies.

Results and Discussion

The amount of rainfall at the Qazvin Station has been estimated and predicted using the mentioned algorithms and copula forecasting and estimation methods. For the application of joint functions and analysis of two-variable drought,

the correlation between the variables should first be estimated. The comparison between observed drought variables (i.e., duration and intensity) and their corresponding fitted distributions indicates a significant correlation between the fitting distribution and the observed drought data. Modeling was performed using soft computing algorithms and rainfall data (70% for testing and 30% for validation) and the output of the joint functions to estimate and predict the correlation and error rates with M5-PSO, M5-GA, CART, M5-GEP, and M5-GMDH algorithms. Based on the modeling performance using soft computing and statistical indicators, the M5-GEP model is the best.

Conclusion

In this article, the analysis of meteorological drought characteristics (i.e., duration and intensity) for the Qazvin Station during the years 1964 to 2018 was examined using a combination of the piecewise function with soft computing algorithms. The M5 model was used for parameter classification. One of the important steps in using the combination of algorithms with the piecewise function for analyzing hydrological issues is the appropriate estimation of the dependent parameter of the piecewise function. Therefore, inaccurate estimation of the piecewise function parameter leads to underlying or overestimation in the modeling method, which was the main reason for our use of algorithms. The results showed that the M5-GEP model could be successfully used for drought modeling. The error rate and correlation coefficient of the model were RMSE=0.12 and CC=0.96, respectively, indicating that the M5-GEP model provided the best result among the models. Additionally, the CART model yielded an error rate and correlation coefficient of RMSE=0.46 and CC=0.87, respectively, for estimating drought duration and intensity, which was the worst result among the proposed models.

Keywords: Copula-based model, Drought analysis, Return period, Soft computing algorithms

Article Type: Research Article

Acknowledgment

The paper is based on a Ph.D. Dissertation. The authors are grateful for the support provided by the Science and Research Branch, Islamic Azad University.

Conflicts of interest

The authors declare no competing financial interests or personal relationships that might have influenced the work reported in this paper.

Data availability statement

Data is available upon request.

Authors' contribution

Hamed Kiafar: Experiment, writing; **Hossein Babazadeh, Hossein Sedghi, and Ali Saremi:** Supervision and advising.

*Corresponding Author, E-mail: h_babazadeh@iaua.ac.ir

Citation: Kiafar, H., Babazadeh, H., Sedghi, H., & Saremi, A. (2024). Analysis of drought characteristics using a copula-based method and a hybrid of soft computing algorithms (Case study: Qazvin Station). *Water and Soil Management and Modelling*, 4(4), 183-202.

DOI: 10.22098/mmws.2023.13351.1329

Received: 20 July 2023, Received in revised form: 08 September 2023, Accepted: 10 September 2023, Published online: 10 September 2023

Water and Soil Management and Modeling, Year 2024, Vol. 4, No. 4, pp. 183-202

Publisher: University of Mohaghegh Ardabili

© Author(s)





تحلیل ویژگی‌های خشکسالی با استفاده از روش مفصل و ترکیب الگوریتم‌های محاسبات نرم (مطالعه موردی: ایستگاه قزوین)

حامد کیافر^۱، حسین بابازاده^{۲*}، حسین صدقی^۳، علی صارمی^۴

^۱ دانش‌آموخته دکتری، گروه علوم و مهندسی آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۲ استاد، گروه علوم و مهندسی آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۳ استاد، گروه علوم و مهندسی آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۴ استادیار، گروه علوم و مهندسی آب، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

چکیده

خشکسالی یک پدیده طبیعی است که تأثیرات منفی قابل توجهی بر بخش‌های مختلف دارد. به‌ویژه به دلیل تغییرات در الگوهای بارندگی مانند مقدار، شدت و توزیع که منجر به خشکسالی‌های سالانه در برخی مناطق شده است. پیش‌بینی، پایش و مدیریت خشکسالی به‌عنوان پدیده‌ای طبیعی همواره از چالش‌های مهم جامعه علمی و تصمیم‌گیری بوده است. برای درک بهتر و پیش‌بینی خشکسالی، پژوهشگران معمولاً از روش‌های آماری در مطالعات هیدرولوژیکی استفاده می‌کنند. توابع احتمال مشترک برای تحلیل چندین متغیر به‌طور همزمان و حفظ همبستگی‌های آن‌ها مؤثر هستند. علاوه بر این، استفاده از الگوریتم‌های محاسبات نرم برای پیش‌بینی و برآورد به‌طور فزاینده‌ای محبوب شده است. بخش‌های مختلف کشور نیز با توجه به ویژگی‌های جغرافیایی با شدت و مدت متفاوتی از خشکسالی مواجه بوده است. در این مقاله به‌منظور پایش خشکسالی‌های هواشناسی ایستگاه قزوین در ایران، مدت و شدت خشکسالی با استفاده از داده‌های بارندگی ماهانه طی سال‌های ۱۳۴۳ تا ۱۳۹۷ با استفاده از الگوریتم‌های ترکیبی محاسبات نرم M5-GEP، M5-GMDH، GA، M5-GEP، M5-GMDH و CART بر مبنای توابع مفصل با نرم‌افزار کدنویسی R و GeneXproTools مورد بررسی قرار گرفته است. ویژگی‌های خشکسالی از شاخص استاندارد بارش ماهانه (SPI) محاسبه می‌شود. الگوریتم درخت M5 از نوع دسته‌بندی برای داده‌ها استفاده شد. داده‌های پیش‌بینی باران سنجی (از سال ۱۳۹۷-۱۴۰۱) به‌طور تصادفی که ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد داده‌ها برای آزمون برگزیده شد و سپس با خروجی پژوهش‌های دیگر (میزان خطا، همبستگی و صحت) مقایسه شدند. در بین الگوریتم M5-GEP بهترین جواب در مقایسه با سایر مدل‌ها نمایش داد. مقدار RMSE برای M5-GEP مبتنی بر مفصل ۰/۱۲ به‌دست آمد و برتری این مدل را در مقایسه با سایر الگوریتم‌های اشاره شده نشان می‌دهد. چنین ویژگی‌های احتمالی خشکسالی می‌تواند برای مدیریت و برنامه‌ریزی منابع آب استفاده شود.

واژه‌های کلیدی: الگوریتم‌های محاسبات نرم، تحلیل خشکسالی، توابع مفصلی، دوره بازگشت

نوع مقاله: پژوهشی

*مسئول مکاتبات، پست الکترونیکی: h_babazadeh@iau.ac.ir

استناد: کیافر، حامد، بابازاده، حسین، صدقی، حسین، صارمی، علی (۱۴۰۳). تجزیه و تحلیل ویژگی‌های خشکسالی با استفاده از روش مفصل و ترکیب الگوریتم‌های محاسبات نرم (مطالعه موردی: ایستگاه قزوین). *مدل سازی و مدیریت آب و خاک*، ۴(۴)، ۱۸۳-۲۰۲.
DOI:10.22098/mmws.2023.13351.1329

تاریخ دریافت: ۱۴۰۲/۰۲/۲۹، تاریخ بازنگری: ۱۴۰۲/۰۶/۱۷، تاریخ پذیرش: ۱۴۰۲/۰۶/۱۹، تاریخ انتشار: ۱۴۰۲/۰۶/۱۹

مدل سازی و مدیریت آب و خاک، سال ۱۴۰۳، دوره ۴، شماره ۴، صفحه ۱۸۳ تا ۲۰۲

ناشر: دانشگاه محقق اردبیلی © نویسندگان



۱- مقدمه

کاهش میزان بارندگی، اثر مراکز فشار و پدیده‌های جهانی همچون النینو، لاینیا و انسو پدیده خشکسالی همواره در طول تاریخ عواقب ناخوشایندی به دنبال داشته است. النینو باعث عدم توازن بزرگ آب و هوایی و تغییر اقلیم در بوم‌سازگان بسیاری از نقاط جهان می‌شود. بسیاری از پدیده‌های هیدرولوژیکی از قبیل بارندگی، سیلاب‌ها و خشکسالی‌ها به‌طور ذاتی طبیعی تصادفی دارند. خشکسالی به‌عنوان پدیده طبیعی، از دیرباز اثرات نامطلوبی به همراه داشته است. تغییر اقلیم بارندگی از لحاظ دفعات و میزان، باعث می‌شود خشکسالی در نقاط مختلف روی دهد و باعث خسارات زیادی در ابعاد مختلف شود (Caperaà et al., 1997). برای بهبود برنامه‌ریزی و مدیریت منابع آب، لازم است دوره‌های بازگشت مشروط و مشترک برای رویدادهای خشکسالی تخمین زده شود. بر اساس مجموعه‌ای از عوامل وضعیت خشکسالی و اثرات آن در استان‌های کشور در چهار دسته بحرانی، کمبود آبی، تنش آبی و وضعیت قابل تحمل طبقه‌بندی می‌شود. تمام پژوهش‌های انجام شده به این نکته اشاره دارد که آمادگی لازم برای پدیده خشکسالی که حداقل آسیب‌های محیط زیستی، اجتماعی و اقتصادی را همراه داشته‌اند (Dracup et al., 1980). عوامل و تقابل‌های فیزیکی پیچیده که بر این پدیده‌ها حاکم‌اند، شناخته شده نیستند و مدل‌های ریاضی و تجربی قابل اطمینانی وجود ندارد. بنابراین، به‌منظور توصیف مناسب چنین پدیده‌هایی، شیوه‌های آماری در بسیاری از مطالعات هیدرولوژیکی به کار گرفته شده است که یکی از دلایل استفاده از توابع مفصل که به‌طور هم‌زمان می‌توانند دو یا چند متغیر را تحلیل کنند و در عین حال همبستگی بین متغیرها را در محاسبات حفظ کنند. این توابع از توزیع‌های تک متغیره به‌عنوان توزیع حاشیه‌ای برای ساخت توزیع‌های چندمتغیره استفاده می‌کنند. پیش‌بینی و تخمین پدیده خشکسالی به‌دلیل پیچیدگی و پارامترهای تأثیرگذار بر آن دشوار است و مقدار دقیق آن قابل ارائه نیست به همین دلیل با میزان خطای بیان می‌شود که یکی از دلایل استفاده از الگوریتم‌های محاسبات نرم، پیش‌بینی و تخمین همین مورد است. محاسبات نرم از جمله روش‌های داده‌کاوی^۱، می‌تواند داده‌ها را به دانش قابل استفاده تبدیل کنند و برای تعریف مدل از آن‌ها استفاده کرد. پیش داده‌ها برای ورودی به مدل تعریف شده بسیار اهمیت دارد. استفاده از ترکیب این الگوریتم‌ها در علوم مهندسی بسیار رواج پیدا کرده است؛ زیرا هزینه مدل‌سازی آزمایشگاهی را ندارند و نسبت به مدل‌سازی‌های عددی صحت بسیار بالاتری در زمان کم‌تر ارائه می‌دهند.

در این راستا، Sabziparvar and Kazemi (2010) در مقایسه هفت شاخص خشکسالی با آمار بارندگی ۳۵ ساله در ۲۲ ایستگاه در استان همدان پرداختند و نشان دادند شاخص‌هایی همچون DI، RAI و Z از نظر آماری به هم شبیه بوده و شاخص SPI^۲ برای ارزیابی در مناطق با آب و هوای سرد و نیمه خشک مناسب نیست. در ادامه، Amirataee et al., (2013) از توابع مفصل در استخراج رابطه شدت خشکسالی-منطقه-فراوانی (SAF) در حوضه دریاچه ارومیه استفاده کردند. مفصل فرانک به‌عنوان تابع مناسب برای استخراج منحنی SAF در منطقه مورد مطالعه انتخاب شد. همچنین، Mirabbasi Najafabadi et al., 2014 توزیع مشترک مدت و شدت خشکسالی را برای سنجش بارندگی در شرفخانه (ایران) با استفاده از توابع جفت دو متغیره مختلف مدل‌سازی کرد. توزیع گاما و نمایی به ترتیب برای مدت و شدت خشکسالی برازش داده شد. نتایج نشان داد که مفصل گالامبوس بهترین نتایج را برای رویدادهای خشکسالی ایستگاه شرفخانه ارائه می‌دهد.

از سایر پژوهش‌ها، Ekanayake and Perera (2014) ویژگی‌های خشکسالی منطقه Anuradhapura در سریلانکا را با استفاده از SPI سه ماهه بر اساس داده‌های بارندگی ماهانه طی سال‌های ۲۰۰۷-۱۹۵۱ تعیین کردند. نتایج نشان داد که توزیع گاما به‌عنوان بهترین توزیع حاشیه‌ای از نظر شدت و مدت خشکسالی انتخاب شد. مفصل فرانک نیز به‌عنوان بهترین مفصل برای تحلیل خشکسالی انتخاب شد. در پژوهشی، Tosunoglu and Can (2016) توزیع احتمال مشترک خشکسالی در ترکیه را با استفاده از توابع مفصل مدل‌سازی کردند. آن‌ها ویژگی‌های اصلی خشکسالی (مدت و شدت) را با استفاده از SPI محاسبه کردند. در پژوهش دیگری، Ayantobo et al. (2018) مدل‌سازی رویدادهای خشکسالی را با استفاده از توابع مفصل دو بعدی در چین بررسی کردند. آن‌ها شاخص استاندارد تبخیر-تعرق بارش (SPEI) را در نظر گرفتند و از تخمین حداکثر درستی‌مایی (MLE) و روش برازش منحنی (CFM) برای تخمین پارامترهای مفصل استفاده کردند. نتایج نشان داد که احتمال بیش‌تر خشکسالی‌های طولانی‌تر در مناطق فرعی مربوط به دوره‌های بازگشت کوتاه‌تر با شدت خشکسالی بیش‌تر است. همچنین، Grillakis (2019) روند خشکسالی رطوبت خاک برای اروپا را با استفاده از شاخص رطوبت خاک (SMI) برآورد شده روی شبیه‌سازی‌های شبیه‌ساز محیط زمین مشترک بریتانیا تحت دو مسیر تمرکز نماینده، سناریوهای RCP 2.6 و RCP 6.0 به‌منظور دسترسی به آب و تولید محصولات کشاورزی ارزیابی کردند. نتایج

² Standardized precipitation index

¹ Data mining

منطقه پرجمعیت‌ترین منطقه نیمه‌خشک در جهان است و به‌شدت مستعد خشکسالی است را ارزیابی کردند. هدف آن‌ها ارزیابی فراوانی خشکسالی‌ها در NRB و محاسبه دوره بازگشت هر رویداد خشکسالی با استفاده از تکنیک copula برای بارش استاندارد شده (SPI) در مقیاس زمانی سه ماهه اعمال کردند که مدت و شدت خشکسالی در NRB را در یک توزیع دو متغیره مشترک ادغام می‌کند و سه خانواده از توابع جفت ارشمیدس (-Gumbel Clayton، Frank و Hougaard) را آزمایش کردند. نتایج آنان نشان داد شدت و مدت، نوار شرقی NRB دارای طولانی‌ترین دوره بازگشت (بیش از ۶۰ ماه)، عمدتاً در بخش جنوب‌شرقی است که دوره بازگشت بالای ۹۰ ماه مشاهده شد. NRB شمالی کوتاه‌ترین دوره بازگشت (۴۵ ماه) را نشان می‌دهد. این بخش NRB با بیش‌ترین فراوانی خشکسالی‌های شدید است. این نتایج اطلاعات مفیدی را برای مدیریت خطر خشکسالی در NRB فراهم می‌کند.

با توجه به شرایط خشک و نیمه خشک اقلیمی ایران، تناوب خشکسالی با شدت‌های مختلف مشاهده شده است که رصد و پایش آن در برنامه‌ریزی برای مواجهه با آن با استفاده از ابزارهای نوین از اهمیت بالایی برخوردار است. در این راستا به‌منظور پایش خشکسالی‌های هواشناسی دشت قزوین، مدت و شدت خشکسالی با استفاده از داده‌های بارندگی ماهانه طی سال‌های ۱۳۴۳ تا ۱۳۹۷ با استفاده از الگوریتم‌های ترکیبی محاسبات نرم-M5، PSO، M5-GA، M5-GMDH، M5-GEP و CART بر مبنای توابع مفصل مورد بررسی قرار گرفت.

۲- مواد و روش

۲-۱- منطقه مورد مطالعه

استان قزوین در قسمت مرکزی ایران واقع شده است. مرکز استان قزوین، با ارتفاع ۱۲۷۸ متری از سطح دریا قرار دارد. مساحت آن برابر ۱۵۶۲۳ کیلومترمربع است. این استان از شمال به استان‌های گیلان و مازندران، از غرب به استان‌های زنجان و همدان، از جنوب به استان مرکزی و از سمت شرق به استان تهران ختم می‌شود. استان قزوین در دامنه جنوبی رشته‌کوه‌های البرز واقع شده است که به‌دلیل داشتن ارتفاعات متعدد و بارندگی‌های متوسط، از نقاط معتدل کشور به‌شمار می‌آید. مرتفع‌ترین کوه‌های استان شامل "سیاهلان، کی جکین، سفید کوه و سیاه کوه" هستند که بیش‌ترین ارتفاع و کوه‌های شمالی سیاهلان به ارتفاع ۴۱۷۵ متر از سطح دریا دست یافته‌اند. منطقه مرکزی و شرقی استان را دشت‌ها تشکیل می‌دهند که شیب آن‌ها از شمال غرب به جنوب شرق ممتد است و در پایین‌ترین نقطه ارتفاعی خود که در منطقه طارم سفلی و کناره‌های دریاچه سفیدرود واقع شده

آنان نشان داد که شرایط خشکسالی SMI در اروپا با تفاوت‌های اساسی در بین مناطق تشدید می‌شود. مناطق اروپای شرقی و مدیترانه بیش‌ترین آسیب را دیده‌اند.

(Nabaei et al. (2019 وضعیت خشکسالی هواشناسی را بر اساس SPI و توابع مفصل با استفاده از داده‌های بارندگی ماهانه ۱۰۲ ایستگاه در ایران ارزیابی کردند. آن‌ها سه ویژگی خشکسالی شامل شدت، مدت و اوج را برآورد کردند. نقشه‌های فضایی تولید شده اطلاعات مفیدی در مورد وضعیت خشکسالی برای مدیریت بهتر خشکسالی ارائه می‌دهد. در ادامه، Dodangeh et al. (2019 در تحقیقی عدم قطعیت‌ها را در مدل‌سازی دو متغیره فرآیند شدید بارش-رواناب با استفاده از کاپولا محاسبه کردند. آن‌ها از اوج‌های بیش از حد آستانه و سری حداکثر سالانه برای محاسبه عدم قطعیت‌ها استفاده کردند. نتایج نشان داد که عدم قطعیت مجموعه داده ورودی در تحلیل فرکانس دو متغیره مؤثرتر است. در برزیل، Cunha et al. (2019 از شاخص یکپارچه خشکسالی (IDI) که ترکیبی از شاخص خشکسالی مبتنی بر هواشناسی و شاخص مبتنی بر سنجش از دور است، برای ارزیابی رویدادهای خشکسالی از سال ۲۰۱۱ تا ۲۰۱۹ استفاده کردند. آن‌ها بیان نمودند که در این مدت، رویدادهای خشکسالی در سراسر کشور مشاهده شد که شدیدترین و گسترده‌ترین آن بین سال‌های ۲۰۱۱ تا ۲۰۱۹ بود.

در آسیای مرکزی Zhang et al. (2020 خطر خشکسالی را با استفاده از مفصل چندبعدی بر اساس شاخص استاندارد تبخیر-تعرق بارش (SPEI) بررسی نمودند. نتایج آنها نشان داد که خطر خشکسالی با استفاده از توزیع مشترک سه بعدی با توزیع دو بعدی یکسان است. در چین Wang et al. (2019 از مجموعه داده‌های SPEIbase v2.4 با وضوح فضایی ۰/۵ درجه برای استخراج شاخص تبخیر-تعرق بارش استاندارد (SPEI) استفاده کردند. همچنین، Mehr et al. (2020 تحلیل خشکسالی هواشناسی در سراسر آنکارا پایتخت ترکیه، را با استفاده از سه مدل آب و هوای جهانی (GCM) به‌صورت منطقه‌ای انجام دادند. برای این منظور، شاخص بارش استاندارد شده (SPI) و شاخص استاندارد تبخیر-تعرق بارش (SPEI) تحت سناریوهای RCP4.5 و RCP8.5 گازهای گلخانه‌ای پیش‌بینی شدند. به‌طور کلی، نتایج آنان نشان داد که آنکارا شش رویداد خشکسالی شدید و دو خشکسالی شدید را در طول دوره مرجع، ۲۰۰۰-۱۹۷۱ تجربه کرده است. با این حال، پیش‌بینی‌ها نشان می‌دهد که رویدادهای خشکسالی کم‌تری برای دوره آینده نزدیک ۲۰۴۰-۲۰۱۶ بدون رویدادهای خشکسالی شدید بالقوه است. در نهایت، Da Rocha Junior et al. (2020 دوره بازگشت خشکسالی دو متغیره را با استفاده از مفصل در منطقه شمال شرقی به‌منظور این‌که این

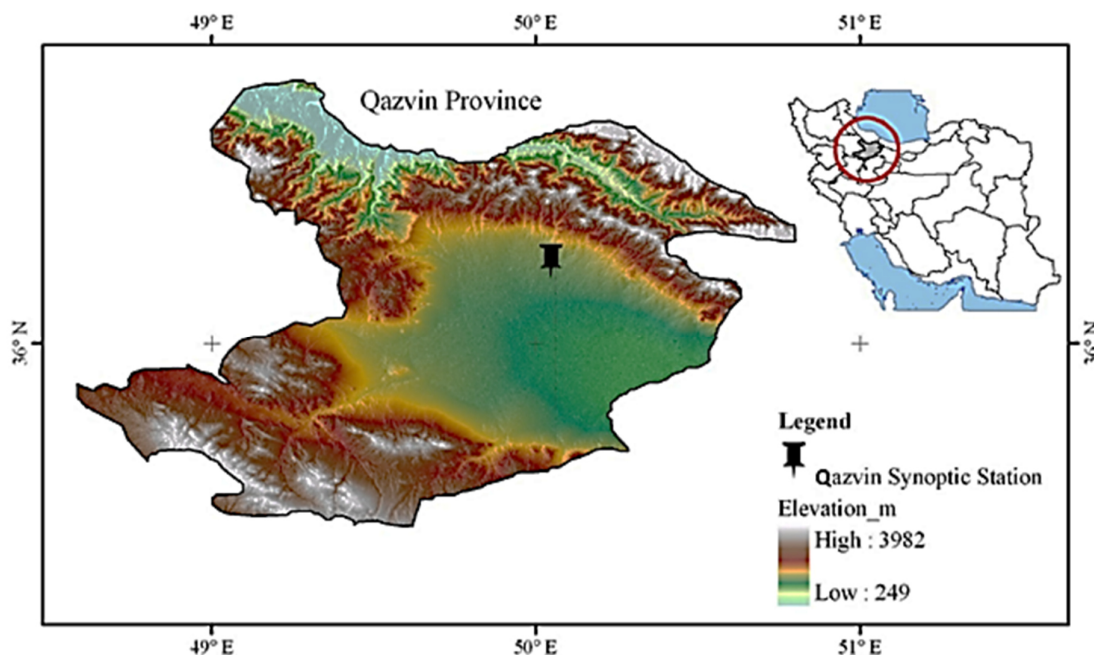
۳-۲- شرایط آب و هوای استان قزوین

شرایط هواشناسی استان قزوین تحت تأثیر دو سامانه هوایی اصلی قرار دارد. این دو سامانه شامل سامانه غربی و مدیترانه‌ای که مهم‌ترین منبع بارش برای استان است و سامانه پرفشار شمال که در تعیین سرمای و رطوبت منطقه دخالت دارد. همچنین، در استان دو جریان بادی وجود دارد. جریان اول به نام "باد مه" از سمت شمال غرب به سمت جنوب شرق حرکت می‌کند و بیش‌ترین بارش‌ها را به همراه دارد. جریان دوم به نام "باد شره" از سمت جنوب شرق به سمت شمال غربی جاری است. آب‌های سطحی استان بیش‌تر در دو حوزه آبخیز جاری است. حوزه آبخیز شمالی (سفیدرود) و حوزه آبخیز جنوبی (شور) است. شکل ۱ موقعیت ایستگاه باران‌سنجی قزوین را نشان می‌دهد. در مقاله حاضر، داده‌های بارندگی ماهانه از ژانویه ۱۹۶۴ (دی‌ماه ۱۳۴۲) تا دسامبر ۲۰۱۸ (۱۳۹۷) برای ایستگاه سینوپتیک قزوین جهت شناسایی ویژگی‌های خشکسالی آن استفاده شد. میانگین بارندگی در ایستگاه قزوین ۳۰۴ میلی‌متر است.

است، ارتفاع آن ۱۱۳۰ متر از سطح دریا است. حداقل ارتفاع آن نیز در شمال غربی و در بخش طارم سفلی و نزدیکی دریاچه سفیدرود با ارتفاعی حدود ۳۰۰ متر از سطح دریا ثبت شده است. بر اساس آخرین تقسیمات کشوری تا پایان سال ۱۳۹۰، استان قزوین شامل شش شهرستان، ۲۹ شهر، ۱۹ بخش، ۴۶ دهستان، و ۱۱۴۷ آبادی است.

۲-۲- شرایط اقلیمی استان قزوین

استان قزوین دارای شرایط اقلیمی متنوعی است و این تنوع اقلیمی باعث فراهم آوردن زمینه مناسبی برای کشت انواع محصولات گرمسیری در مناطقی مانند طارم سفلی و رودبار الموت و کشت محصولات سردسیری در سایر نقاط استان شده است. جنگل‌های این استان در ارتفاعاتی میان دو تا سه هزار متری در بخش‌های الموت و طارم سفلی واقع شده و به‌صورت درختان پراکنده و بی‌نظم رویده‌اند. این جنگل‌ها به‌عنوان ذخیره‌گاهی از تنوع ژنتیکی درختان عمل می‌کنند.



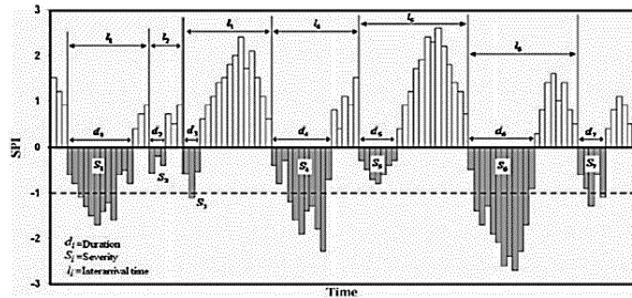
شکل ۱- موقعیت ایستگاه باران‌سنجی قزوین
Figure 1- Position of Qazvin rain gauge station

(1999)، سری زمانی SPI بر اساس بارندگی ماهانه برای تعریف و محاسبه ویژگی‌های خشکسالی مانند مدت و شدت در این پژوهش استفاده شد. نقش شدت و مدت زمان برای مدیریت خشکسالی مهم است (شکل ۲) (McKee et al., 1993).

۴-۲- شاخص استاندارد بارش و ویژگی‌های خشکسالی

بر اساس روابط تجربی

با توجه به این که SPI (شاخص استاندارد بارش) به‌عنوان شاخص مناسب برای تحلیل خشکسالی شناخته می‌شود (Hayes et al.,)



شکل ۲- توصیف ویژگی‌های خشکسالی (مدت و شدت)
Figure 2 - Description of drought characteristics (i.e., duration and intensity)

در الگوریتم‌های داده‌مبنا، داده‌ها جهت آموزش و آزمون به‌صورت کاملاً تصادفی با سعی و خطا انتخاب شدند که نماینده از همه داده‌ها باشد. مدت زمان خشکسالی یک دوره SPI منفی مداوم است که شدت خشکسالی به‌عنوان مقادیر تجمعی SPI در طول مدت خشکسالی محاسبه می‌شود که در رابطه (۲) توضیح داده شده است.

$$S = - \sum_{i=1}^D SPI_i \quad (2)$$

در رابطه فوق، S شدت خشکسالی و D طول دوره خشکسالی (بر حسب ماه) است (McKee et al., 1993).

۲-۵- توابع مفصل تجربی

توابع مفصل توزیع چندمتغیره هستند که احتمال مشترک توزیع‌ها را به توزیع‌های حاشیه‌ای مربوطه مرتبط می‌کنند (Nelsen, 2006). برای نمونه با اندازه n، مفصل تجربی d بعدی C_n برابر است با:

$$C_n \left(\frac{k_1}{n}, \frac{k_2}{n}, \dots, \frac{k_d}{n} \right) = \frac{a}{n} \quad (3)$$

در رابطه (۳)، a تعداد نمونه‌ها {X₁, ..., X_d} با 1 ≤ k₁, ..., k_d ≤ n آماره‌های ترتیبی نمونه هستند. داده‌های بارندگی ما به‌طور معمول به‌صورت خطی تفکیک نمی‌شوند. بنابراین، باید از تکنیک‌ها و نگاشت‌های مختلفی استفاده کرد تا مسأله را به یک فضای جدید منتقل کرد که داده‌ها در آن به‌صورت خطی تفکیک شوند. به‌عبارت دیگر، باید یک تابع نگاشت با نام "φ" تعریف کنیم که داده‌های موجود در فضای X را به یک فضای جدید به نام Z منتقل می‌کند. در این فضای جدید، داده‌ها به‌صورت خطی قابل تفکیک خواهند بود و از آن‌جا به سادگی می‌توان به تحلیل آن‌ها پرداخت.

$$\varphi: X \quad (4)$$

خشکسالی را می‌توان به‌عنوان دوره‌هایی تعریف کرد که SPI به زیر منفی و یا کمتر می‌رسد. شرایط مرطوب و خشک با توجه به مقادیر SPI طبقه‌بندی شده‌اند و در جدول ۱ ارائه شده است. لازم به ذکر است که برخی از رخداد‌های خشکسالی ممکن است مدت خشکسالی طولانی داشته‌باشند اما SPI متوسط برای هر دوره به منفی یک یا کمتر نمی‌رسد (Mirabbasi Najafabadi et al., 2012). SPI تجمعی برای چنین رخداد‌هایی علی‌رغم این‌که احتمال بیش‌تر از رخداد‌هایی با مدت زمان کوتاه است ولی SPI کم‌تر از منفی یک دارند. یک رخداد خشکسالی را می‌توان به‌عنوان یک دوره پیوسته تعریف کرد که SPI زیر صفر است (Shiau, 2006). مدت زمان خشکسالی یک دوره SPI منفی مداوم است که شدت خشکسالی به‌عنوان مقادیر تجمعی SPI در طول مدت خشکسالی محاسبه می‌شود که در رابطه (۱) توضیح داده شده است.

$$S = - \sum_{i=1}^D SPI_i \quad (1)$$

در رابطه بالا، S شدت خشکسالی و D طول دوره خشکسالی (بر حسب ماه) است. مقدار SPI بر اساس جدول زیر دسته‌بندی می‌شود (McKee et al., 1993).

جدول ۱- طبقه‌بندی آب و هوا بر اساس مقادیر SPI (McKee et al., 1993)

SPI	دسته‌بندی
≤ -2.00	فوق‌العاده خشک
-1.50 - -1.99	خشکسالی شدید
-1 - -1.49	نسبتاً خشک
+0.99 - +1.00	نزدیک به نرمال
+1.49 - +1.00	نسبتاً مرطوب
+1.99 - +1.5	خیلی مرطوب
+ 2.00 ≤	فوق‌العاده مرطوب

یک مدل رگرسیون خطی ساخته می شود که به عنوان یک "مختص" برای آن زیر فضا عمل می کند. در مدل درختی M5، ابتدا تقریبی از آستانه های مورد نیاز برای دودویی سازی ویژگی ها با تحلیل کل داده های مربوط به هر ویژگی به دست می آید. الگوریتم از پیش از هر مرحله ای از مدل سازی، داده ها را به صورت استاندارد می کند. این فرآیند استانداردسازی داده ها به دلیل مهم بودن آن در مدل سازی انجام می شود. برای این منظور، ابتدا فرض می شود X_1, X_2, \dots, X_n که داده های n تایی دارای میانگین و انحراف استاندارد s هستند. سپس با انجام تبدیل های مربوطه، داده ها به شکل استاندارد با میانگین صفر و انحراف استاندارد یک تبدیل می شوند:

$$z_i = \frac{x_i - \bar{x}}{s}, \quad i = 1, 2, \dots, n \quad (12)$$

داده های جدید Z_1, Z_2, \dots, Z_i که به نام داده های استاندارد معروف هستند، به سادگی با میانگین صفر و انحراف معیار یک ایجاد می شوند. این ویژگی ها باعث می شود که اصطلاح "استاندارد" برای توصیف آن ها مناسب باشد. شکل ۳ نحوه دسته بندی داده های بارش آب های سطحی شهر قزوین را با الگوریتم M5 نشان می دهد.

۲-۶-۲- الگوریتم PSO

الگوریتم PSO با ذرات تولید شده به طور تصادفی در جست و جوی یک راه حل بهینه شده مقدماتی است. تخمین پارامتر مفصل بر اساس الگوریتم PSO هر ذره دارای موقعیت و سرعت است. رابطه (۱۳) سرعت ذرات در فرآیند بهینه سازی را نشان می دهد. با استفاده از دو مقدار Ps و G، نسل ها از راه یک فرآیند تکراری به دست می آیند. Ps نشان دهنده بهترین راه حل برای هر ذره ای با کمترین خطا یا نزدیکترین به هدف تعریف شده است که تاکنون به دست آورده است و G بهترین مقداری ثبت شده ای است که تاکنون توسط هر ذره از جمعیت به دست آمده است. شکل ۴ روند تشکیل چارچوب کلی الگوریتم PSO را نشان می دهد.

$$z = \varphi(x) \quad (5)$$

در نتیجه رابطه تفکیک کننده ما باید به صورت زیر تبدیل شود:

$$w^T z + b = 0 \quad (6)$$

$$w^A T \varphi(x) + b = 0 \quad (7)$$

در این صورت مسأله ما تغییراتی خواهد داشت که در رابطه زیر قابل مشاهده است.

$$\text{Min} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \varphi(x_i)^T \varphi(x_j) - \sum_{i=1}^n \alpha_i \quad (8)$$

$$\sum_{i=1}^n \alpha_i y_i = 0; \quad 0 \leq \alpha_i \leq C \quad (9)$$

به جای عبارت $\varphi(x_i)^T \varphi(x_j)$ می توانیم یک حالت جامع تر به صورت $K(x_i, x_j)$ تعریف کرد که به وسیله آن بتوان به فضاهای پیچیده تر راه یافت. فرآیند مذکور ترند کرنل (Kernel) نامیده می شود. عبارت $K(x_i, x_j)$ در حالت کلی یک تابع غیرخطی است و می تواند شکل های مختلفی داشته باشد. در نتیجه مسأله به صورت زیر قابل اصلاح است:

$$y = \text{Sign} \left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \quad (10)$$

$$b = \frac{1}{|S|} \sum_{i=1}^n [y_i - \sum_{j=1}^n \alpha_j y_j K(x_j, x)] \quad (11)$$

۲-۶-۲- معرفی و روش الگوریتم های محاسبات نرم استفاده شده در مدل

در این مطالعه خشکسالی هواشناسی استان قزوین با استفاده از نرم افزار کد نویسی R و GeneXproTools 4 از ترکیب توابع مفصل و الگوریتم ها محاسبات نرم خشکسالی ها الگوریتم ها ترکیب و داده های به صورتی تصادفی در مدل ها انتخاب شده است.

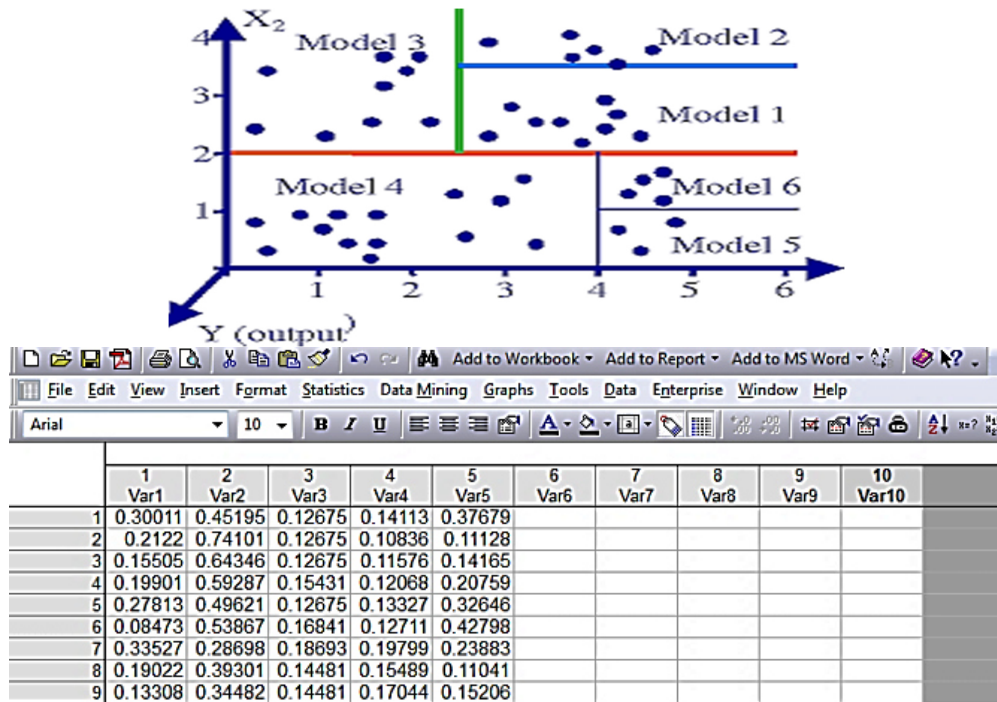
۲-۶-۱- الگوریتم M5

مدل سازی پیچیده با استفاده از فضای پارامتر چندبعدی به زیر فضاهای مختلف تقسیم می شود و در هر یک از این زیر فضاهای

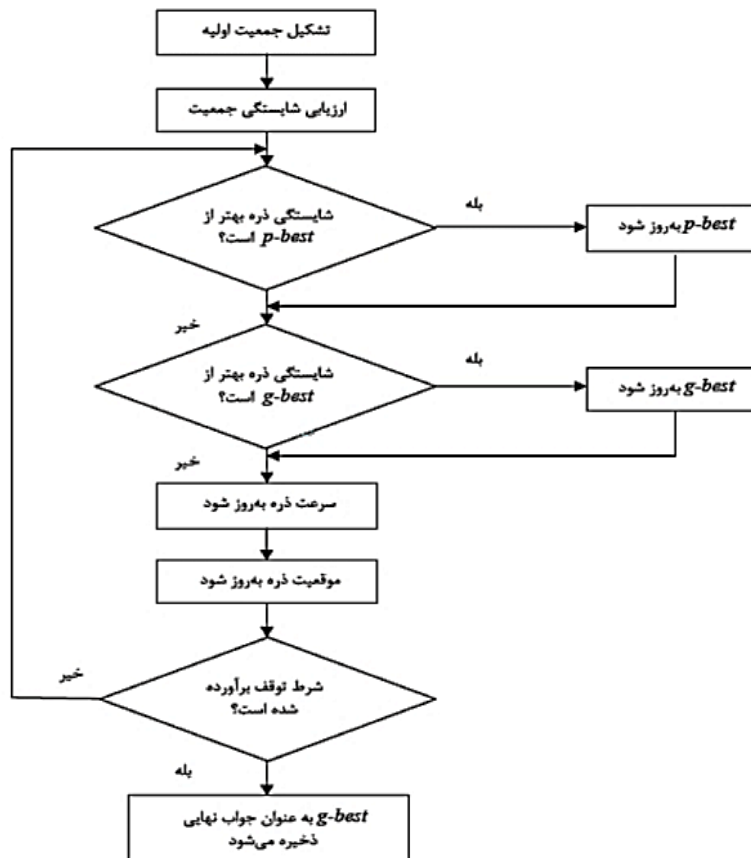
$$\text{velocity of particle } i \text{ at time } k+1 \rightarrow v_{k+1}^i = w v_k^i + c_1 \text{rand} \frac{(p^i - x_k^i)}{\Delta t} + c_2 \text{rand} \frac{(p_k^g - x_k^i)}{\Delta t} \quad (13)$$

current motion
particle memory influence
swarm influence

inertia factor range: 0.4 to 1.4
self confidence range: 1.5 to 2
swarm confidence range: 2 to 2.5



شکل ۳- نحوه دسته‌بندی الگوریتم M5
Figure 3- Classification method using M5 algorithm

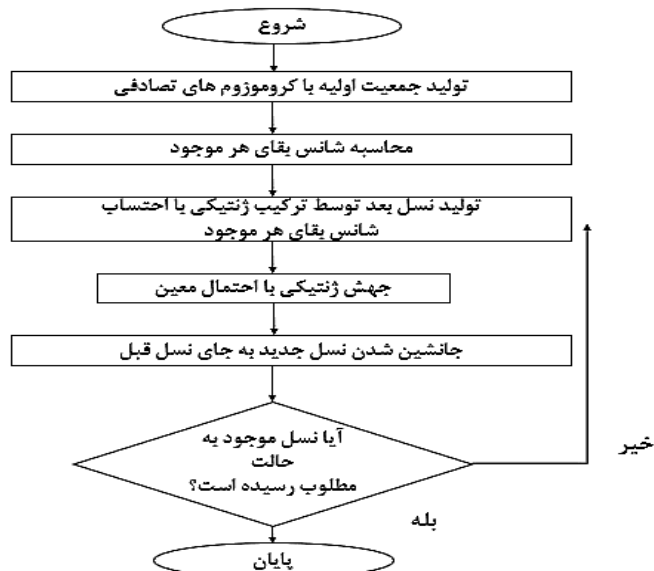


شکل ۴- نمودار جریان‌یابی تشکیل الگوریتم PSO
Figure 4- PSO algorithm formation flowchart

۳-۶-۲- الگوریتم ژنتیک (GA)

است؛ زیرا فضای تولید با تولید پتانسیل تصادفی جست و جو می شود. راه حل ها جمعیتی از افراد به عنوان راه حل های نامزد، به تدریج تغییر می کنند تا به راه حل مطلوب برسند که در شکل ۵ نشان داده شده است.

الگوریتم ژنتیک (GA) یک روش بهینه سازی است که از تکامل بیولوژیکی در طبیعت تقلید می کند. این بهینه سازی از نوع تصادفی



شکل ۵- ساختار الگوریتم ژنتیک

Figure 5- Genetic algorithm structure

تصمیم طبقه بندی استفاده شد و برای درخت های رگرسیون از معیاری به نام انحراف حداقل مربعات^۱ (LSD) استفاده شد. این معیار به صورت زیر تعریف می شود:

$$SS(t) = \sum_{i=1}^{N_t} (y_i(t) - \bar{y}(t))^2 \quad (16)$$

در رابطه (۱۶)، N_t تعداد رکوردها (داده ها) در گره برگ $y_i(t)$ مقدار خروجی (متغیر هدف) در گره برگ t و $\bar{y}(t)$ میانگین مقادیر متغیر هدف در گره برگ t است. به عنوان نمونه در گره ریشه (اولین گره): تعداد کل داده ها N_t ، میانگین مقادیر متغیر هدف برای همه داده ها $\bar{y}(t)$. حال متغیر ورودی S ، زمانی بهترین متغیر برای ایجاد شاخه در گره t است که مقدار $Q(S, t)$ را بیشینه کردیم.

$$Q(S, t) = SS(t) - SS(t_R) - SS(t_L) \quad (17)$$

در رابطه بالا، $SS(t_R)$ و $SS(t_L)$ به ترتیب بیان گر میزان $SS(t)$ در شاخه سمت راست و چپ گره t هستند.

۵-۶-۲- الگوریتم GEP

الگوریتم GEP با غلبه بر محدودیت نقش دوگانه کروموزوم ها در الگوریتم های پیش از خود امکان اعمال عملگرهای متعدد ژنتیک را با ضمانت سلامت همیشگی کروموزوم های فرزند فراهم می سازد

۴-۶-۲- الگوریتم CART

الگوریتم طبقه بندی و رگرسیون درختان CART یک روش آماری برای تحلیل و طبقه بندی است که قادر به ارائه قوانین تصمیم گیری در صورت تصمیم گیری برای پیش بینی متغیر وابسته بر اساس طیف متغیرهای مستقل است. در این الگوریتم، برای انتخاب بهترین متغیر ورودی برای ایجاد شاخه از شاخص Gini استفاده می شود.

$$gini(D) = 1 - \sum_{j=1}^n p_j^2 \quad (14)$$

در این رابطه، در آن p_j فراوانی نسبی از کلاس j در مجموعه D و n تعداد کلاس های این مجموعه است. حال بهترین انتخاب برای تقسیم مجموعه D به دو مجموعه D_1 و D_2 از معیار زیر تبعیت می کند:

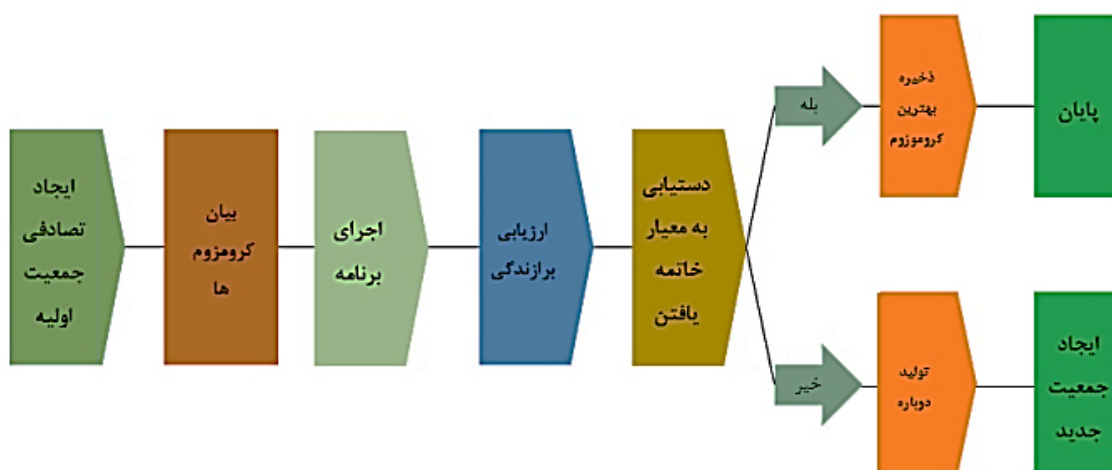
$$gini_{split}(D) = \frac{|D_1|}{|D|} gini(D_1) + \frac{|D_2|}{|D|} gini(D_2) \quad (15)$$

در این رابطه، علامت $| \cdot |$ به معنای تعداد اعضای هر مجموعه است. این معیار را برای همه متغیرها در گره امتحان کرد و آن متغیری که مقدار (رابطه ۱۵) برای آن از همه کوچکتر باشد، برای ایجاد شاخه انتخاب می شود. دقت شود شاخص Gini در ساخت درخت های

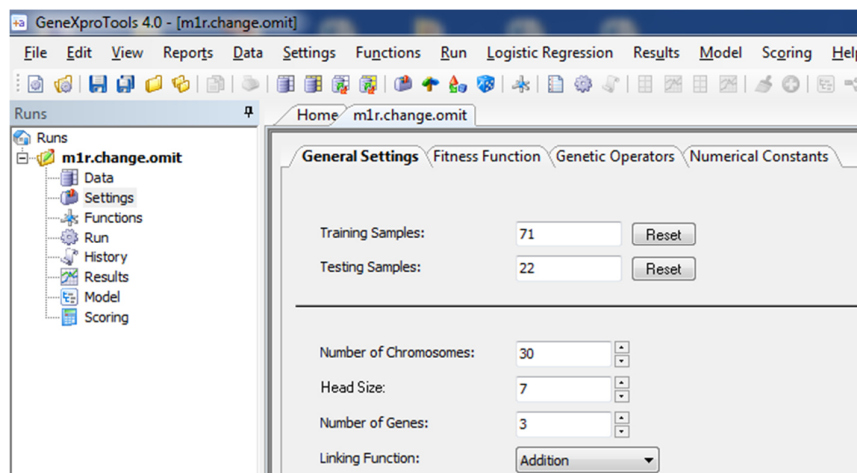
¹ Least square deviation

شکل ۶ فرایند الگوریتم GEP و در شکل ۷ تصویری از صفحه مدل مورد استفاده ارائه شده است.

و با سرعتی بیش از GP به دلیل تنوع ساختاری بالاتر از GA، فضای پاسخ‌های ممکن را به صورت کامل‌تری جست و جو می‌کند. در



شکل ۶- نمودار مراحل تحلیل GEP
Figure 6- Flowchart of GEP analysis stages



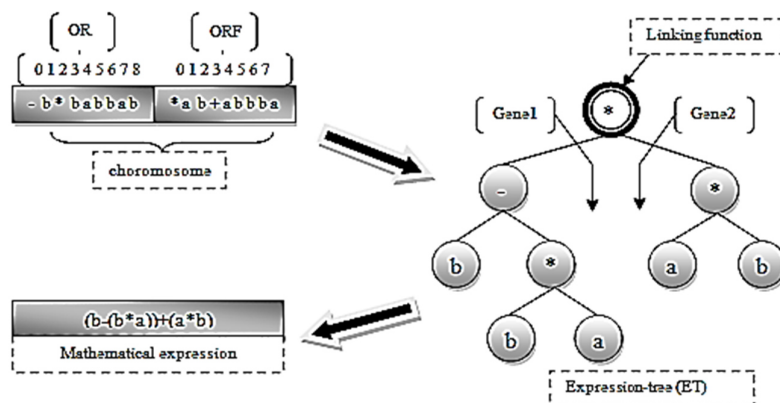
شکل ۷- آزمایش داده‌ها
Figure 7- Data experiment

در ابداع کروموزوم‌هایی با قابلیت نمایش به صورت درخت تحلیل، بینش اساسی از (GEP) شامل می‌شود. این کروموزوم‌ها با استفاده از زبان کاروا که برای خواندن و تفسیر اطلاعات کدگذاری شده در کروموزوم‌ها طراحی شده است، کدگذاری و بیان می‌شوند. به علاوه، ساختار کروموزوم‌ها به گونه‌ای طراحی شده است که اجازه ایجاد چندین ژن را می‌دهد، هر یک از آن‌ها برای برنامه‌های کوچک‌تر یا کدگذاری درختی متفاوت مسأله استفاده می‌شوند. یکی از جوانب حیاتی در GEP، انتخاب تابع برازش است و هدف آن یافتن یک راه‌حل که برای تمام موارد برازش با صحت یکسان عمل کند. شکل ۸ نشان می‌دهد چگونه

با توجه به ماهیت این الگوریتم فضای پاسخ‌های ممکن به صورت کامل‌تری جست و جو شد. در این مدل ابتدا یکبار داده‌ها برای آموزش و سپس داده‌ها برای آزمون برازش داده شد و در انتها به صورت یک‌جا ارائه شدند.

در فرآیند برازش توسط الگوریتم تعدادی از مسائل هدف بررسی می‌شوند. اگر یک راه‌حل قابل قبول با کیفیت مناسب پیدا شود یا نسل‌ها به تعداد معینی برسند، تکامل متوقف می‌شود و بهترین راه‌حل تاکنون به عنوان نتیجه گزارش داده می‌شود. در غیر این صورت، بهترین راه‌حل‌ها از نسل حاضر به عنوان نخبه‌ها انتخاب می‌شوند و بقیه راه‌حل‌ها به فرآیند گزینشی واگذار می‌شوند.

یک کروموزوم با دو ژن به صورت یک رشته خطی کدگذاری شده و چه طور به صورت ET بیان شده است.



شکل ۸- کدگذاری رشته خطی و کدگذاری کروموزومی در ET برای یک کروموزوم با دو ژن در GEP
Figure 8- Linear String Encoding and Chromosome Encoding in ET Form for a Chromosome with Two Genes in GEP

متناظر با درصد خطای نسبی است که برای صحت تعریف می شود. چنانچه $P_{ij}=T_j$ خطا صفر می شود. پس برای یک برازش کامل که در آن عبارت داخل قدرمطلق صفر و $f_i=f_{max}=nR$ می شود.

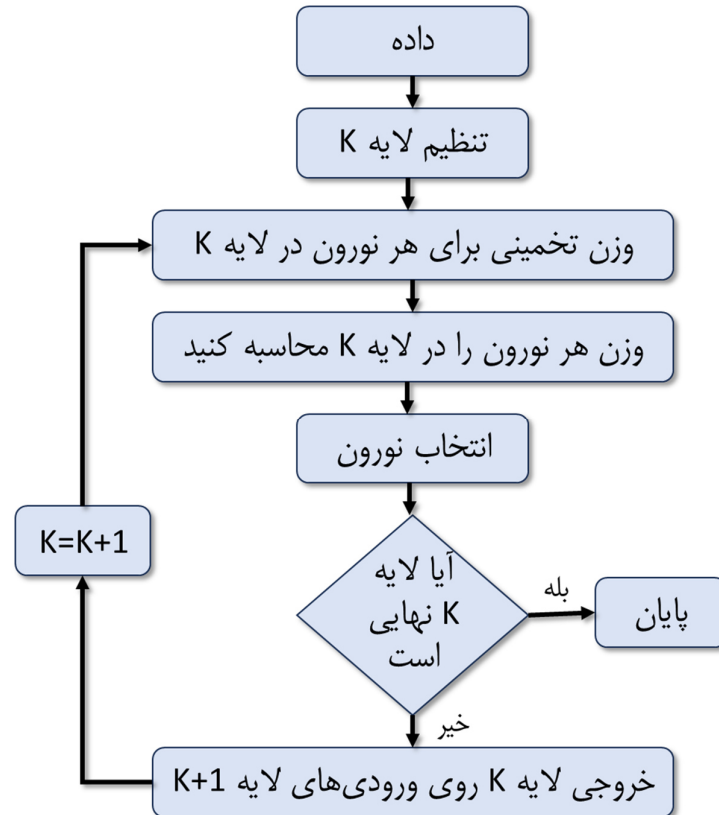
۶-۶-۲- الگوریتم GMDH

شبکه های تشکیل شده توسط GMDH بر داده ورودی های تعریف شده به صورت شکل ۹ خروجی مدل را پیش بینی می کند.

در GEP از لحاظ ریاضی، برازش f_i از یک برنامه انفرادی i به صورت رابطه (۱۸) بیان می شود.

$$f_i = \sum_{j=1}^n \left(R - \left| 100 * \frac{P_{ij} - T_j}{T_j} \right| \right) \quad (18)$$

در محدوده انتخاب شده، مقدار P_{ij} پیش بینی شده توسط برنامه انفرادی i برای مورد برازش j (از میان n مورد برازش) و T_j مقدار هدف برای مورد برازش j است. رابطه داخل قدرمطلق



شکل ۹- نمودار جریان‌ی الگوریتم GMDH
Figure 9- GMDH algorithm flowchart

۲-۸- ضریب همبستگی

ضریب همبستگی یک مؤثر مهم برای تعیین نوع و قدرت ارتباط بین دو متغیر، به‌ویژه شدت و مدت زمان خشکسالی است. این ضریب معمولاً مقادیری بین یک و منفی یک اختیار می‌کند. در واقعیت، اگر مقدار آن برابر با یک باشد، نشان‌دهنده یک ارتباط کاملاً مستقیم و کامل بین دو متغیر است. به‌عبارت دیگر، هنگامی که ضریب همبستگی برابر با یک است، افزایش یکی از متغیرها به تنهایی منجر به افزایش دقیقاً معادلی در متغیر دیگر می‌شود. اگر ضریب همبستگی برابر با منفی یک باشد، نشان‌دهنده یک ارتباط کاملاً معکوس و کامل بین متغیرها است. در این حالت، افزایش یکی از متغیرها به تنهایی به کاهش دقیقاً معادلی در متغیر دیگر منجر می‌شود. اگر مقدار ضریب همبستگی برابر با صفر باشد، نشان‌دهنده عدم وجود هیچ ارتباط خطی معناداری بین دو متغیر است. به‌عبارت دیگر، ضریب همبستگی صفر نشان می‌دهد که تغییرات در یک متغیر هیچ تأثیر معناداری بر تغییرات در متغیر دیگر ندارد و بین آن‌ها هیچ ارتباط خطی

۲-۷- محاسبه خطا و همبستگی مدل‌ها

برای ارزیابی میزان صحت هر یک از رابطه‌ها از اندازه‌گیری‌های مختلفی استفاده شده است. این اندازه‌گیری‌ها عبارت‌اند از: پارامترهای شاخص انحراف اریبی (Bias)، نسبت اختلاف (DR^1)، پراکندگی (SI^2)، ضریب همبستگی (CC^3)، جذر متوسط مربع خطا ($RMSE^4$)، شاخص توافق (Ia^5)، میانگین خطای مطلق (MAE^6) استفاده شد. هر یک از این پارامترها برای ارزیابی عملکرد مدل‌ها به‌صورت انحراف (اریبی) که اختلاف میانگین داده‌های اندازه‌گیری و محاسباتی است، نسبت اختلاف و شاخص پراکندگی به‌ترتیب به صورت رابطه‌های (۱۹)، (۲۰) و (۲۱) معرفی می‌شوند:

$$Bias = \bar{y} - \bar{x} \quad (19)$$

$$DR = \frac{1}{N} \sum \frac{y_i}{x_i} \quad (20)$$

$$SI = 1 - \frac{\sqrt{\frac{1}{N} \sum (XI - Yi)^2}}{\bar{x}} \quad (21)$$

⁴ Root mean square error

⁵ Index of agreement

⁶ Mean absolute error

¹ Discrepancy ratio

² Scatter index

³ Correlation coefficient

(۲۹)

$$D \geq d \text{ and } S \geq s \Rightarrow T_{DS} = \frac{E(L)}{1 - F_D(d) - F_S(s) + C(F_D(d), F_S(s))}$$

(۳۰)

$$D \geq d \text{ or } S \geq s \Rightarrow T'_{DS} = \frac{E(L)}{1 - C(F_D(d), F_S(s))}$$

در آن $E(L)$ زمان بازه خشکسالی مورد انتظار است، T_{DS} دوره بازگشت مشترک را برای شرایطی نشان می‌دهد که $D \geq d$ و $S \geq s$ و T'_{DS} دوره بازگشت مفصل برای $D \geq d$ یا $S \geq s$ است. برای موقعیت‌های مشروط، دوره‌های بازگشت شدت و مدت خشکسالی را می‌توان به‌عنوان رابطه‌های (۳۱) و (۳۲) توصیف کرد. مشابه موارد قبلی، برای این وضعیت دو مورد می‌توان در نظر گرفت که شامل دوره بازگشت مدت خشکسالی با توجه به شدت خشکسالی بیش از مقدار آستانه و دوره بازگشت شدت خشکسالی با توجه به مدت زمان خشکسالی بیش از آستانه ثابت است (Shiau, 2006).

$$T_{D|S \geq s} = \frac{E(L)}{[1 - F_S(s)][1 - F_D(d) - F_S(s) + C(F_D(d), F_S(s))]} \quad (۳۱)$$

$$T_{S|D \geq d} = \frac{E(L)}{[1 - F_D(d)][1 - F_D(d) - F_S(s) + C(F_D(d), F_S(s))]} \quad (۳۲)$$

در روابط فوق $T_{S|D \geq d}$ دوره بازگشت مشروط برای S را نشان می‌دهد وقتی $D \geq d$ و $T_{D|S \geq s}$ نشان‌دهنده دوره بازگشت مشروط برای D زمانی است که $S \geq s$. چنین احتمالی را نمی‌توان با استفاده از ارزیابی جداگانه شدت و مدت خشکسالی محاسبه کرد.

۳- نتایج و بحث

با توجه به پیچیدگی پدیده خشکسالی، استفاده از تحلیل چندمتغیره روش دقیق‌تری از تحلیل یک متغیره خشکسالی دارد. با استفاده از الگوریتم‌های معرفی شده و مفصل میزان بارندگی ایستگاه قزوین پیش‌بینی و تخمین و سپس با مدل تجربی تابع مفصل صحت‌سنجی شده است (Salvadori and Michele, 2006, Shiau, 2006). مفصل‌های نظری (پارامتری) ذکر شده شامل فرانک، کلیتون، گالامبوس، پلاکت و گامبل-هوگارد برای تحلیل خشکسالی دومتغیره مورد استفاده قرار گرفته است. مفصل‌های نظری (پارامتری) ذکر شده با مفصل‌های تجربی مقایسه می‌شوند. مفصل‌های تجربی توزیع‌های احتمال مشترک و بر اساس رتبه‌بندی تجربی هستند. پدیده خشکسالی با استفاده از ترکیب الگوریتم‌ها رایج است چون پارامترهای مؤثر آن دارای رنج‌های مختلفی می‌باشند یک پارامتر شامل بازه بزرگ و دیگری دارای بازه کوچک باشد و اهمیت و تأثیر هر پارامتر از روی کمیت

وجود ندارد. به‌طور کلی، ضریب همبستگی به ما کمک می‌کند تا ارتباط بین متغیرها را درک کرده و تفسیر کنیم. در ادامه، رابطه‌های ریشه میانگین مربعات خطا (رابطه ۲۳)، میانگین خطای مطلق (رابطه ۲۴) ارائه شده است

$$C = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (۲۲)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - T_i)^2}{N}} \quad (۲۳)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}| \quad (۲۴)$$

توابع مفصل شامل فرانک، کلیتون، گالامبوس، پلاکت و گامبل-هوگارد برای تحلیل خشکسالی دومتغیره مورد استفاده قرار گرفته است. مفصل‌های نظری (پارامتری) ذکر شده با مفصل‌های تجربی مقایسه و صحت‌سنجی می‌شوند. مفصل‌های تجربی توزیع‌های احتمال مشترک و بر اساس رتبه‌بندی تجربی هستند صحت‌سنجی کرد:

$$C_e(u_i, v_i) = \frac{1}{N} \sum_{i=1}^N I\left(\frac{D_i}{N+1} \leq u_i, \frac{S_i}{N+1} \leq v_i\right) \quad (۲۵)$$

پس از ساخت توزیع دومتغیره و انتخاب بهترین تابع جفت، احتمال فراتر رفتن (Nelsen, 2006). مفصل تجربی دومتغیره داده‌های مشاهده شده (u_i, v_i) را می‌توان با استفاده از (رابطه ۲۵) تخمین

و

هم‌زمان شدت و مدت خشکسالی از یک آستانه معین را می‌توان با استفاده از روابط ۲۶ تا ۲۸ تخمین زد (Shiau, 2006).

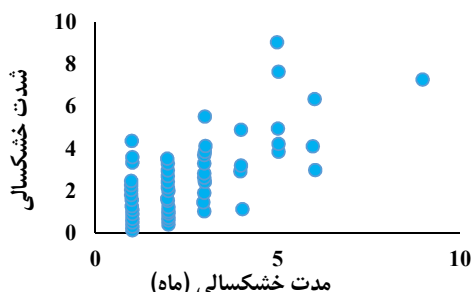
$$P(D \geq d, S \geq s) = 1 - FD^{(d)} - FS^{(s)} + C(FD^{(d)}, FS^{(s)}) \quad (۲۶)$$

$$P(S \leq s | D \geq d) = \frac{FS(s) - C(FD^{(d)}, FS(s))}{1 - FD^{(d)}} \quad (۲۷)$$

$$P(D \leq d | S \geq s) = \frac{FD(d) - C(FD(d), FS(s))}{1 - FS(s)} \quad (۲۸)$$

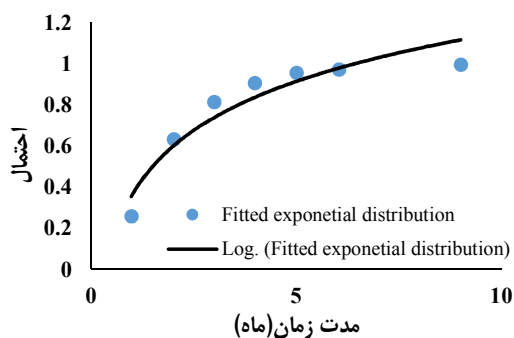
توزیع مشروط شدت خشکسالی با توجه به این که مدت خشکسالی از مقدار معینی فراتر می‌رود (یعنی d') را می‌توان به راحتی با استفاده از رابطه (۲۵) محاسبه کرد. برعکس از رابطه (۲۶) با توجه به این که شدت خشکسالی از مقدار معینی (یعنی s') فراتر رفته است، می‌توان برای ارزیابی توزیع مدت خشکسالی استفاده کرد. مشابه احتمالاتی که قبلاً ذکر شد، دوره‌های بازگشت خشکسالی را می‌توان تخمین زد. در پژوهشی، Shiau (2006) دوره‌های بازگشت خشکسالی مشترک و دوره بازگشت شدت را برای دو شرایط تعریف کرد. اولی (رابطه ۲۹) دوره بازگشت برای $D \geq d$ و $S \geq s$ و دومی (رابطه ۳۰) برای دوره بازگشت $D \geq d$ یا $S \geq s$ است.

برای کاربرد توابع مفصل و تحلیل خشکسالی دومتغیره، ابتدا همبستگی بین متغیرها برآورد شد که نشان‌دهنده پراکندگی مدت خشکسالی و شدت متناظر آن است که از ۱۳۱ خشکسالی مشاهده شده در طول دوره مورد مطالعه استخراج شده است (شکل ۱۱).



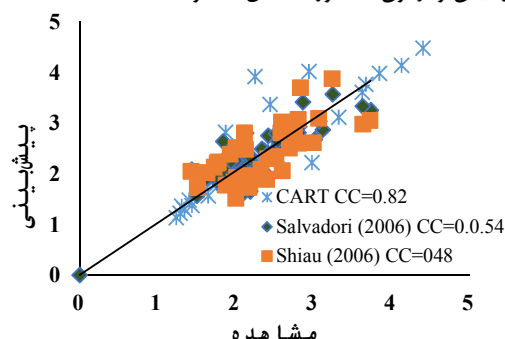
شکل ۱۱- نمودار پراکندگی شدت خشکسالی مشاهده شده و مدت خشکسالی برای ایستگاه قزوین
Figure 11- Scatter plot of observed drought intensity and duration for Qazvin Station

با توجه به شکل ۱۱، همبستگی بالایی بین شدت خشکسالی و مدت خشکسالی بر اساس رابطه تجربی SPI مشاهده می‌شود. مقایسه بین متغیرهای خشکسالی مشاهده شده (یعنی مدت و شدت) و توزیع‌های برازش مربوطه در شکل‌های ۱۲ و ۱۳ ارائه شده است. همان‌طور که در این نمودارها مشاهده می‌شود، همبستگی مناسبی بین توزیع برازش و داده‌های خشکسالی مشاهده شده وجود دارد. همان‌طور که در شکل ۱۳ مشخص است، همبستگی بالایی بین شدت خشکسالی و مدت خشکسالی بر اساس رابطه تجربی SPI مشاهده می‌شود.



شکل ۱۲- مدت زمان خشکسالی مشاهده شده و توزیع نمایی متناسب مربوطه
Figure 12- Observed drought duration and corresponding exponential distribution

آن مشخص نشود که این مورد را با وزن‌دهی به داده‌های مورد نظر برطرف می‌نماییم. برای صحت مدل پیشنهاد شده ابتدا الگوریتم پیشنهادی روی داده‌های مقاله نام برده شده پیاده‌سازی شد. برای این کار ابتدا با تک الگوریتم CART این کار را انجام دادیم به این صورت که روابط تجربی در مقالات نام برده خشکسالی را بر حسب پارامترها به صورت دوره بازگشت خشکسالی ارائه می‌دهند در صورتی که الگوریتم‌ها به صورت تخمین و پیش‌بینی ارائه می‌دهند که دقت آن نسبت به رابطه تجربی خشکسالی SPI مقایسه شده و میزان همبستگی نتایج آزمایش و آزمون به صورت شکل ۱۰ بود:

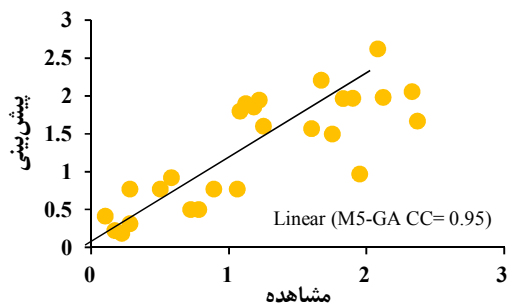


شکل ۱۰- مقایسه مدل ارائه شده الگوریتم CART با مدل Shiau (2006) و Salvadori and Michele (2006)
Figure 10- Comparison of the presented model of the CART algorithm with the model of Salvadori and Michele (2006) and Shiau (2006)

با استفاده از ترکیب الگوریتم CART میزان بارش در مقایسه با روش SPI (Salvadori and Michele, 2006) پیش‌بینی شد که نتایج خطا و همبستگی، $I_a=0.94$ ، $CC=0.82$ ، $DR (\text{mean} \pm \text{sdt}) = 1.03 \pm 0.83$ و $SI (\%) = 28.3$ را نشان داد.

مفصل‌های نظری (پارامتری) ذکر شده با مفصل‌های تجربی مقایسه می‌شوند. مفصل‌های تجربی توزیع‌های احتمال مشترک و بر اساس رتبه‌بندی تجربی هستند که الگوریتم‌های ترکیبی (Hybrid algorithms) را آن‌ها مقایسه و صحت‌سنجی شد. تکاملی به این شکل، از مکانیسم‌ها و عملیات اولیه برای حل مسأله بهره بردیم و در طول یک دوره تکراری اقدام به یافتن یک راه‌حل مناسب برای مسأله کردیم. این الگوریتم‌ها با شروع از یک جمعیتی که حاوی راه‌حل‌های تصادفی است، شروع به کار کردند و در هر مرحله تکراری سعی در بهبود مجموعه راه‌حل‌ها داشتند. در ابتدا، تعدادی از داده‌ها به صورت تصادفی حدس زده می‌شدند و سپس تابع هدف برای هر یک از این افراد محاسبه می‌شد و نسل اول ایجاد می‌شد. اگر هیچ‌یک از معیارهای خاتمه بهینه‌سازی مشاهده نشود، یک نسل جدید ایجاد می‌شد

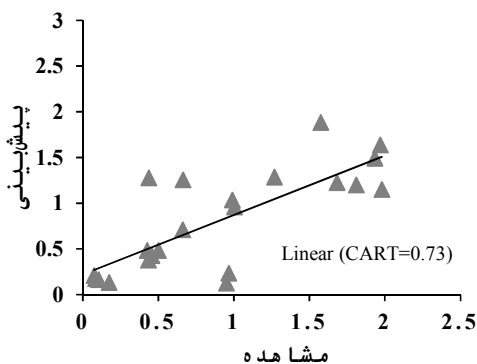
۳-۲- تخمین پارامتر مفصل بر اساس الگوریتم مدل M5-GA با استفاده از الگوریتم M5 که دسته بندی داده ها انجام شد و با استفاده از GA که الگوریتمی برای بهینه سازی است میزان بارش پیش بینی شد که نتایج عبارتند از: $I_a=0.94$, $CC=0.92$, DR (mean \pm sdt)= 25.8 و $RMSE=0.34$, SI (%)= 25.8 و 1.01 ± 0.14 را نشان داد.



شکل ۱۵- مقایسه بین مقادیر اندازه گیری و پیش بینی شده از مدل M5-GA

Figure 15- Comparison between measured and predicted values from the M5-GA model

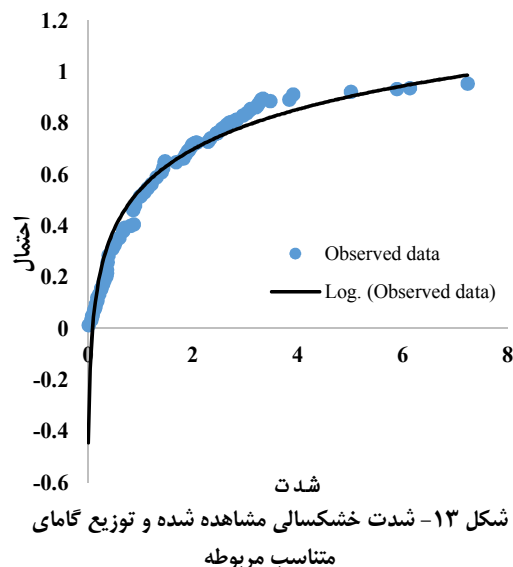
۳-۳- تخمین پارامتر مفصل بر اساس الگوریتم مدل CART در این قسمت داده های ثبت شده که دسته بندی آن ها بر اساس متغیرهای وابسته معلوم است انتخاب شده اند برای تشکیل ساختار یک درخت که بتواند متغیرهای جدید با رشدش و شاخه های تخمین و پیش بینی کند.



شکل ۱۶- مقایسه بین مقادیر اندازه گیری و پیش بینی از مدل CART

Figure 16- Comparison between measured and predicted values from the CART model

با استفاده از الگوریتم M5 که دسته بندی داده ها انجام شده و با استفاده از CART که الگوریتمی از نوع درخت و یک روش آماری برای تحلیل و میزان بارش پیش بینی شد که نتایج عبارتند از: $I_a=0.95$, $CC=0.87$, $RMSE=0.46$, SI (%)= 24 و DR (mean \pm sdt)= 1.12 \pm 0.38 را نشان داد.

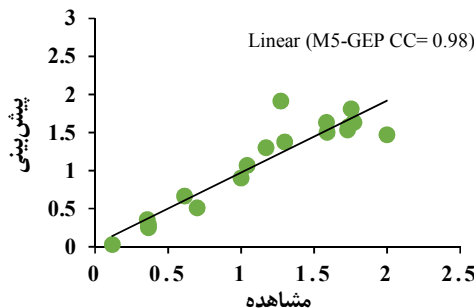


شکل ۱۳- شدت خشکسالی مشاهده شده و توزیع گامای متناسب مربوطه

Figure 13- Observed drought intensity and corresponding gamma distribution

۳-۱- مدل و نتایج با الگوریتم های محاسبات نرم

در این قسمت با استفاده از داده های بارش (۷۰ درصد آزمون و ۳۰ درصد آزمایش) و خروجی توابع مفصل، تخمین و پیش بینی و میزان همبستگی و خطا با الگوریتم های M5-PSO, M5-GA, M5-GMDH و CART انجام شد. سپس نمودار نتایج توابع مفصل و الگوریتم های آن به صورت مجزا ارائه و در انتها تمامی الگوریتم ها با هم مقایسه شده اند و بهینه ترین آن انتخاب شد (شکل های ۱۴ تا ۱۹).



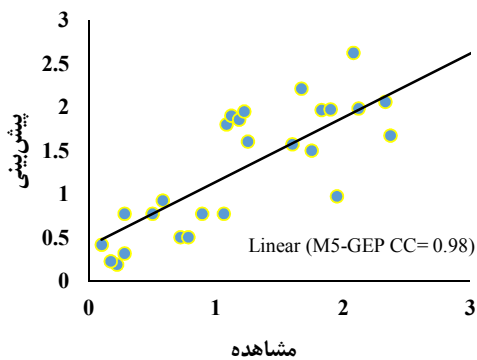
شکل ۱۴- مقایسه بین مقادیر اندازه گیری و پیش بینی شده از مدل M5-PSO

14- Comparison between measured and predicted values from the M5-PSO model

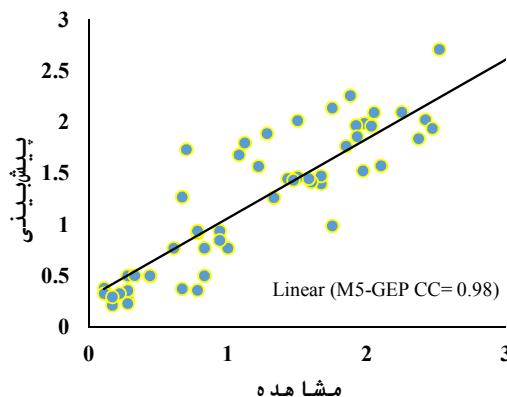
با استفاده از ترکیب الگوریتم M5 و Pso میزان بارش پیش بینی شد که نتایج خطا و همبستگی $I_a=0.94$, $CC=0.95$, DR (mean \pm sdt) = 28.3 و $RMSE=0.32$, SI (%)= 28.3 و 1.03 ± 0.83 را نشان داد.

۳-۴- ORF و بیان درختی ET ژن‌ها

کروموزوم‌ها به صورت بیان درختی (ETS) نشان داده، که این مرحله هم مطابق با یک تابع برازش ارزیابی می‌شود تا میزان مناسب بودن یک راه حل در حیطه مسأله تعیین شد.

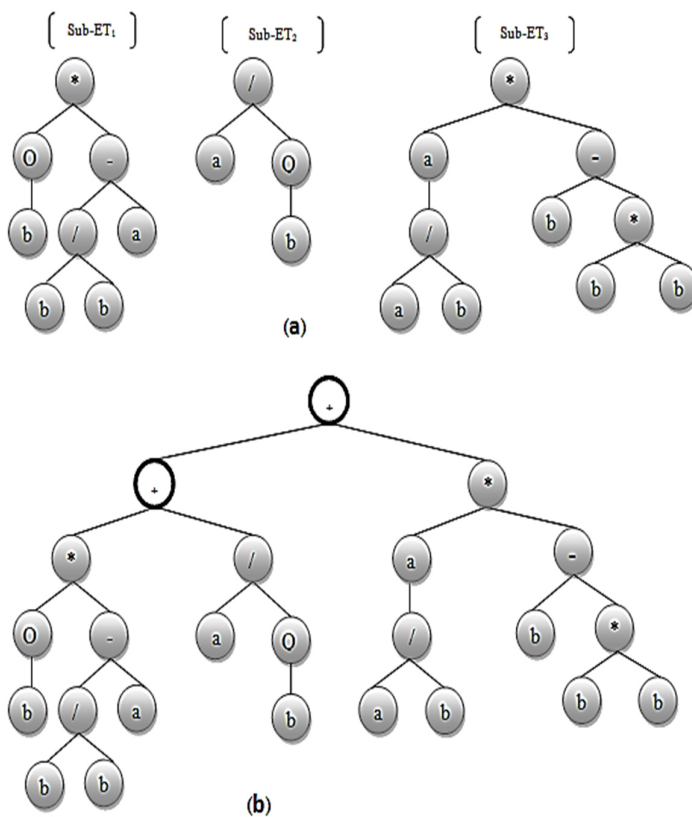


شکل ۱۸- نمایش بین مقدار اندازه‌گیری با مقدار پیش‌بینی با مدل الگوریتم M5-GEP (داده‌ها برای آزمون)
Figure 18- Comparison of measured and predicted values using the M5-GEP algorithm model (test data)

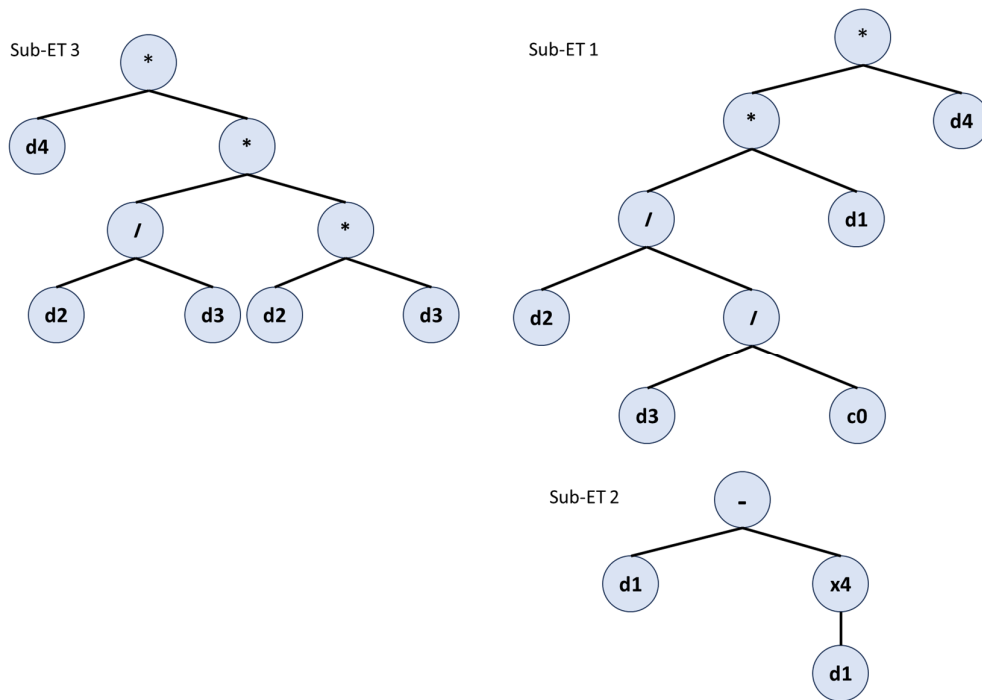


شکل ۱۷- نمایش بین مقدار اندازه‌گیری با مقدار پیش‌بینی با مدل الگوریتم M5-GEP (داده‌ها برای آموزش)
Figure 17- Comparison of measured and predicted values using the M5-GEP algorithm model (training data)

هر ژن به صورت یک زیردرخت (sub-ET) کدگذاری شده بود. این زیردرخت‌های فرعی (sub-ETs) از طریق تأثیر متقابل بر یکدیگر، مجموعه‌ای از واحدهای زیری ET را ایجاد کردند و خروجی مورد نیاز را تولید می‌کردند. رابطه پیش‌بینی خشکسالی بر اساس درخت تصمیم در روش برنامه‌ریزی بیان ژن به صورت زیر استخراج شد (شکل ۱۹). شکل ۲۰ خروجی حاصل از ساختار این ژن تکامل را نشان می‌دهد.

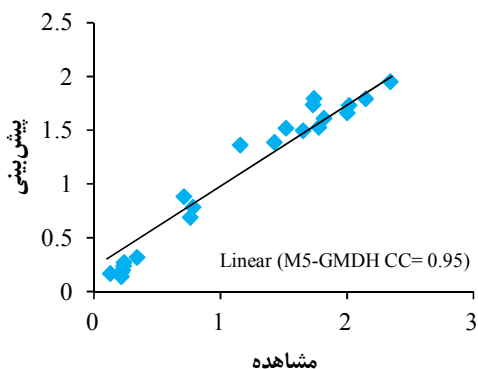


شکل ۱۹- تولید رابطه به دست آمده
Figure 19- Derived relationship output



شکل ۲۰- خروجی مدل
Figure 20- Model output

جدول ۲ و همه مدل های ترکیب شده با مفصل در شکل ۲۳ ارائه داده شده اند.

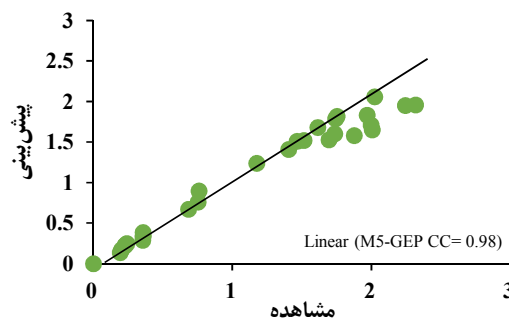


شکل ۲۲- مقایسه بین مقادیر اندازه گیری شده و پیش بینی شده با الگوریتم مدل شده M5-GMDH

Figure 22- Comparison between measured and predicted values using modeled M5-GMDH algorithm

نمودار و جدول مقایسه الگوریتم ها به ترتیب M5-GEP، M5-GMDH، M5-GAM5- PSO، CART از بهترین مدل تا بدترین را نشان دادند. یکی از گام های مهم برای استفاده از ترکیب الگوریتم ها با مفصل برای تحلیل مسائل هیدرولوژیکی، تخمین مناسب پارامتر وابستگی تابع مفصل است. بنابراین، تخمین نادقیق پارامتر مفصل

۳-۴- تخمین پارامتر مفصل بر اساس الگوریتم مدل M5-GEP نتایج شکل ۲۱ و $SI, CC=0.98, I_a=0.97, RMSE=0.17$ و $DR (mean \pm sdt) = 0.77 \pm 0.84$ و $(\%) = 31.6$ را نشان داد.



شکل ۲۱- مقایسه بین مقادیر اندازه گیری شده و پیش بینی با الگوریتم مدل M5-GEP

Figure 21- Comparison between measured and predicted values using M5-GEP model algorithm

۳-۵- تخمین پارامتر مفصل بر اساس الگوریتم مدل M5-GMDH با استفاده از الگوریتم M5 دسته بندی داده ها انجام شد و با استفاده از GMDH بر اساس زوج داده ورودی-خروجی، عملکرد نامشخص یک مدل پیچیده پیش بینی شد. نتایج $CC=0.95, I_a=0.95$ و $DR (mean \pm sdt) = 0.76 \pm 0.95$ و $(\%) = 29.4$ را نشان داد (شکل ۲۲). در انتها کلیه نتایج مدل های ترکیبی جهت پیش بینی و تخمین میزان بارش منطقه مطالعاتی در

الگوریتم ژنتیک چندژنی است. در واقعیت، در سامانه‌های ژنوتیپ/فنوتیپ با عملکرد صحیح، نقش یک فرزند ایجاد افراد پیچیده‌تری که از چندین ژن تشکیل شده بود را داشت و تطبیق زیاد سامانه GEP را به‌خوبی نمایش می‌دهد.

۴- نتیجه‌گیری

در این مقاله، تحلیل ویژگی‌های خشکسالی هواشناسی (یعنی مدت و شدت) برای ایستگاه قزوین طی سال‌های ۱۳۳۳ تا ۱۴۰۱ با استفاده از ترکیب تابع مفصل با الگوریتم‌های محاسبات نرم مورد بررسی قرار گرفت. نتایج نشان داد که مدل M5-GEP را می‌توان با موفقیت برای مدل‌سازی خشکسالی نمونه موردی قزوین استفاده کرد. میزان خطا و ضریب همبستگی مدل به‌ترتیب برابر $RMSE=0.17$ و $CC=0.98$ شد که مدل M5-GEP بهترین نتیجه را از بین مدل‌ها داد. همچنین، مدل CART برای تخمین و شدت خشکسالی میزان خطا و ضریب همبستگی به‌ترتیب برابر با $RMSE=0.52$ و $CC=0.73$ نتیجه حاصل شد که در بین مدل‌های پیشنهاد شده بدترین نتیجه را داده است که در مقایسه مدل تجربی با $RMSE=0.19$ و $CC=0.54$ با $RMSE=0.25$ و $CC=0.4$ بهتر است. دقت بالا و خطای کم‌تر این الگوریتم‌ها نسبت به روش‌های تجربی به‌دست آمد. پس می‌توان از این الگوریتم‌ها برای تخمین پیش‌بینی خشکسالی استفاده نمود.

سپاسگزاری

این مقاله برگرفته از رساله دکتری دانشگاه آزاد اسلامی است که بدین‌وسیله از حمایت‌های انجام شده قدردانی می‌شود.

تضاد منافع نویسندگان

نویسندگان این مقاله اعلام می‌دارند که هیچ‌گونه تضاد منافی در خصوص نگارش و انتشار مطالب و نتایج این تحقیق وجود ندارد.

دسترسی به داده‌ها

داده‌ها و نتایج استفاده شده در این تحقیق از طریق مکاتبه با نویسنده مسئول در اختیار قرار خواهد گرفت.

مشارکت نویسندگان

حامد کیافر: انجام کلیه مراحل مدل‌سازی و نگارش مقاله؛ حسین بابازاده، حسین صدقی، و علی صارمی: راهنمایی و مشاوره.

منابع

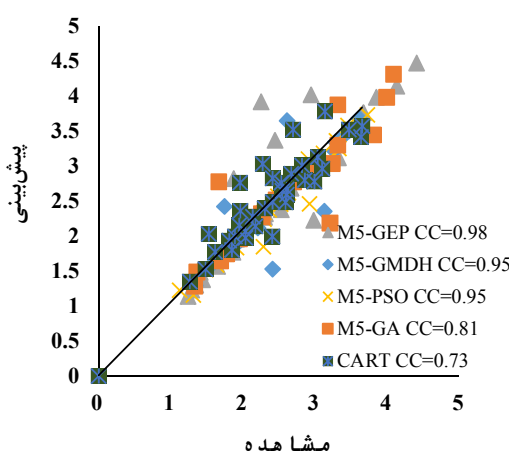
امیر عطایی، بابک، منتصری، مجید، و یاسی، مهدی (۱۳۹۲). مقایسه عملکرد ذاتی هفت شاخص خشکسالی در پایش خشکسالی با

منجر به دست کم گرفتن یا برآورد بیش از حد در روش مدل‌سازی می‌شود که دلیل اصلی استفاده ما از الگوریتم‌ها بود.

جدول ۲- مشخصه آماری روش‌های مختلف جهت تخمین بارش سال‌های ۱۴۰۱-۱۳۴۳

Table 2- Statistical Characteristics of different methods for estimating annual rainfall for the years 1964-2018 "

RMSE	SI (درصد)	I _a	CC	DR (mean ± sdt)	پارامترهای خطا و الگوریتم‌های ترکیب شده
0.52	22	0.89	0.73	1.14±0.33	CART
0.41	24.6	0.91	0.81	1.04±0.34	M5-GA
0.32	28.3	0.94	0.95	1.03±0.83	M5- PSO
0.30	29.4	0.95	0.95	0.95±0.76	M5- GMDH
0.17	31.6	0.97	0.98	0.77±0.84	M5-GEP



شکل ۲۳- مقایسه بین مقادیر اندازه‌گیری شده و پیش‌بینی شده با تمام الگوریتم مدل شده

Figure 23- Comparison between measured and predicted values using all modeled algorithms

ساختار کروموزوم‌های GEP به گونه‌ای است که هرگونه تغییری که در ژنوم ایجاد می‌شود، همواره به ایجاد برنامه‌های صحیح و درست منجر می‌شود. این الگوریتم از ترکیب تک نقطه‌ای و دونقطه‌ای برای تکامل استفاده می‌کند. نوع ترکیب دونقطه‌ای تا حدودی جالب‌تر است چرا که می‌تواند نواحی کدگذاری نشده در کروموزوم‌ها را به مراتب بیش‌تری تغییر دهد. همچنین، با توجه به داده‌های بارش، GEP طبیعت چندژنی بی‌همتای آن را نشان داد و اجازه تکامل برنامه‌های پیچیده‌تری که از چندین زیربرنامه تشکیل شده را داد. در GEP از یک روش نخبه‌گزینی ساده استفاده شد و بهترین افراد یک نسل را برای نسل بعدی حفظ شد. این الگوریتم برخلاف دیگر الگوریتم‌ها چندین عملگر ژنتیکی برای تکثیر افراد با اصلاحات داشت که آن را از سایر الگوریتم‌ها متمایز کرد. به‌علاوه، ساختار کروموزوم‌ها به گونه‌ای طراحی شده که امکان ایجاد چندین ژن که هر کدام برای برنامه‌های کوچک‌تر یا کدگذاری زیردرختی وجود داشت را فراهم کرد. این نکته قابل توجه است که GEP تنها

- میرعباسی نجفآبادی، رسول، فاخری فرد، احمد، دین پڑوه، یعقوب و اسلامیان، سعید (۱۳۹۲). پایش بلندمدت خشکسالی های ارومیه با استفاده از شاخص کمبود توأم. *دانش آب و خاک*، ۳۳(۴)، ۸۷-۱۰۳.
- استفاده از شبیه سازی مونت کارلو. مهندسی عمران و محیط زیست، ۴۳(۱)، ۲۵-۳۹.
- سبزی پرور، علی اکبر، و کاظمی، آزاده (۱۳۸۹). ارزیابی تطبیقی هفت نمایه خشکسالی هواشناسی با استفاده از تحلیل خوشه ای. *علوم و تکنولوژی محیط زیست*، ۹۷-۱۱۲.
- to time scales. In: 8th conference on applied climatology. American Meteorological Society, Boston. Atmospheric and Climate Sciences, 17 - 22 January 1993, Anaheim, California
- Mehr, A.D., Sorman, A.U., Kahya, E., & Afshar, M.H. (2020). Climate change impacts on meteorological drought using SPI and SPEI: Case study of Ankara, Turkey. *Hydrological Sciences Journal* 65, 254–268. doi:10.1080/02626667.2019.1691218
- Mirabbasi Najafabadi, R., Fakheri-Fard, A., & Dinpashoh, Y. (2012). Bivariate drought frequency analysis using the copula method. *Theoretical and Applied Climatology*, 108(2-1), 191-206.
- Mirabbasi Najafabadi, R., Fakheri-Fard, A., Dinpashoh, Y., & Eslamian, S. (2014). Longterm drought monitoring of Urmia using joint deficit index (JDI). *Water and Soil Science*, 23(4), 87-103. (In Persian)
- Nabaei, S., Sharafati, A., Yaseen, Z.M., & Shahid, S. (2019). Copula based assessment of meteorological drought characteristics: regional investigation of Iran. *Agricultural and Forest Meteorology*, 76, 107611. doi:10.1016/j.agrformet.2019.06.010
- Nelsen, R.B. (2006) An Introduction to Copulas. Springer. <https://doi.org/10.1007/0-387-28678-0>.
- Sabziparvar, A., Kazemi, A. (2010). Study of solid waste management of hotels placed in District No 6 of the City of Tehran. *Journal of Environmental Science and Technology*, 1, 97-112. (In Persian)
- Salvadori, G., De Michele, C. (2006). Statistical characterization of temporal structure of storms. *Advances in Water Resources*, 29(6), 827–842. doi:10.1016/j.advwatres.2005.07.013
- Shiau, J. (2006). Fitting drought duration and severity with twodimensional copulas. *Water Resources Management*, 20(5), 795–815. doi:10.1007/s11269-005-9008-9
- Tosunoglu, F., & Can, I. (2016). Application of copulas for regional bivariate frequency analysis of meteorological droughts in Turkey. *Journal of the International Society for the Prevention and Mitigation of Natural Hazards*, 82, 1457–1477. doi:10.1007/s11069-016-2253-9
- Wang, Y., Liu, G., & Guo, E., (2019). Spatial distribution and temporal variation of drought in Inner Mongolia during 1901–2014 using standardized precipitation evapotranspiration index. *Science of the Total Environment*, 654, 850–862. doi:10.1016/j.scitotenv.2018.10.425
- Zhang, L., Wang, Y., Chen, Y., Bai, Y., & Zhang, Q. (2020). Drought risk assessment in central Asia using a probabilistic copula function approach. *Water*, 12(2), 421. doi:10.3390/w12020421
- Amirataee, B., Montaseri, M., & Yasi, M. (2013). Comparison of Inherent Performance of Seven Drought Indices in Drought Mitigation Using a Monte Carlo Simulation Approach. *Journal of Civil and Environmental Engineering*, 43(1), 25-39. (In Persian)
- Ayantobo, O.O., Li, Y., Song, S., Javed, T., Yao, N., (2018). Probabilistic modelling of drought events in China via 2-dimensional joint copula. *Journal of Hydrology*, 559, 373–391. doi:10.1016/j.jhydrol.2018.02.022
- Capéraà, P., Fougères, A.L., & Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84(3), 567-577.
- Cunha, A. P. M. A., Zeri, M., Deusdará Leal, K., Costa, L., Cuartas, L. A., Marengo, J. A., Tomasella, J., Vieira, R. M., Barbosa, A. A., Cunningham, C., Cal Garcia, J. V., Broedel, E., Alvalá, R., & Ribeiro-Neto, G. (2019). Extreme Drought Events over Brazil from 2011 to 2019. *Atmosphere*, 10(11), 642. <https://doi.org/10.3390/atmos10110642>
- Da Rocha Júnior, R.L., dos Santos Silva, F.D., Costa, R.L., Gomes, H.B., Pinto, D.D.C., Herdies, D.L. (2020). Bivariate assessment of drought return periods and frequency in Brazilian northeast using joint distribution by copula method. *Geosciences*, 10(4), 135. doi:10.3390/geosciences10040135.
- Dodangeh, E., Shahedi, K., Solaimani, K., Shiau, J.T., & Abraham, J. (2019). Databased bivariate uncertainty assessment of extreme rainfall-runoff using copulas: comparison between annual maximum series (AMS) and peaks over threshold (POT). *Environmental Monitoring and Assessment*, 191(2), 67. doi:10.1007/s10661-019-7202-0.
- Dracup, J.A., Lee, K.S., & Paulson, E.G.R. (1980). on the statistical characteristics of drought events. *Water Resources Research*, 16(2), 289-96. doi:10.1029/WR016i002p00289.
- Grillakis, M.G. (2019). Increase in severe and extreme soil moisture droughts for Europe under climate change. *Science of the Total Environment*, 60, 1245–1255. doi:10.1016/j.scitotenv.2019.01.001
- Hayes, M.J., Svoboda, M.D., Wilhite, D.A., Vanyarkho, O.V. (1999). Monitoring the 1996 drought using the standardized precipitation index. *Bulletin of American Meteorological Society*, 80(3), 429–438 doi:10.1175/1520-0477(1999)080<0429:MTDUTS>2.0.CO;2
- McKee, T.B., Doesken, N.J., & Kleist, J. (1993). The relationship of drought frequency and duration