

Research Article

Analysis of 70 Years Daily Precipitation in Tabriz Using Visual Data Mining Approach

M Taghi Sattari^{1*}, J Vahdat², S Javidan³

Received: October 18, 2021

Accepted: June 11, 2022

Received in Revised: May 29, 2022

Published online: September 23, 2023

1- Assoc. Prof., Dept. of Water Engineering, Faculty of Agriculture, University of Tabriz

2- M.Sc., Dept. of Statistics, Faculty of Mathematic, University of Tabriz

3- M.Sc. Student, Dept. of Water Engineering, Faculty of Agriculture, University of Tabriz

*Corresponding Author, Email: mtsattar@gmail.com

Abstract

Background and Objectives

Daily precipitation, which is completely stochastic, is one of the basic components of the water cycle and has an important role in the management of surface and ground water resources in terms of quantity and quality. Indispensable element of drought analysis and flood control research is precipitation. The efficient management of surface water resources directly depends on precipitation. The lack of long-term and reliable data has made it difficult to determine precipitation behavior. In this study, we tried to determine the behavior and pattern of precipitation through human-data interaction called visual data mining approach. One of the new approaches that focuses on the use of visualization and graphics in the analysis of complexities in data is visual data mining. Visual data mining can be thought of as a combination of two disciplines, visualization and data mining. Visual data mining is also closely related to computer graphics, multimedia systems, human-computer interaction, pattern recognition, and high-performance computing. The aim of this study is to analyze Tabriz daily precipitation data and discover the patterns in this data with the help of visual data mining approaches. Discovering these patterns and identifying rainfall behavior will help to manage floods on the one hand and droughts on the other.

Methodology

In this study, daily precipitation and temperature data of Tabriz Synoptic Station for the last seventy one years (1951-2021) were used to determine precipitation patterns. Tabriz has a generally cold climate as the center of the eastern Azerbaijan province and is surrounded by mountains. Recently, with the increase in data and software, data mining techniques have also started to attract attention. Data alone cannot mean anything. However, graphs consisting of data can give very meaningful information and messages. Visual data mining approach is an effort to bring data to life with different graphics. In this research, various softwares such as R, ArcGIS and Tableau were used for visualization. In addition, ExcelStat was used to check the accuracy of the data and for statistical tests. R statistical language was used to create the data mining structure and to display the data graphically. Then, different diagrams were drawn using the Tableau program. Finally, the drawn diagrams were evaluated and the final graphics were selected. ArcGIS software was used for spatial analysis and map drawing. Also, multiple linear regression method was used to predict precipitation amount and probability of occurrence.



Findings

According to the temperature histogram, the long-term average annual temperature in Tabriz varies between 12 and 13 degrees Celsius. Also, according to the precipitation histogram, precipitation over 5 mm in Tabriz varies between 10 and 20 days per year on average. The results obtained showed that there has been a change in precipitation patterns in the last 5 years (2017-2021). Although annual precipitation during these five years is above the 71-year average, it is still below the golden precipitation period (1961 to 1970). The results showed that the intensity of spring rains decreased significantly in Tabriz during the period 2006-2021. However, from 1971 to 1980 and from 2001 to 2010, it was observed that the spring rains were more in terms of both precipitation values and precipitation intensity. However, in the following periods from 2011 to 2021, both precipitation intensity and precipitation values decreased. The results showed that most of the spring precipitation in 1981 was 276.3 mm, making up 73.26% of the total precipitation for that year. According to the findings, precipitation has started to decrease in the spring season in recent periods. Decreased dry grain yield in the Azerbaijan region may be affected by decreased spring precipitation. As a result of this decrease, it is expected that the agricultural economy in the study area will be negatively affected.

Conclusion

In this study, as a first, daily precipitation in Tabriz was investigated with visualized data mining techniques. Thus, interesting findings were obtained with the help of different graphics. The results showed an increase in precipitation in the last 5-6 years. It has also been proven that the temperature behavior during this period shows an increase in average temperature, a confirmation of the increase in global temperature. Although the results of this research showed that visualized data mining is successful in precipitation analysis, it is recommended to conduct more comprehensive studies in this field in the future.

Keywords: Daily precipitation, Precipitation pattern, Precipitation trend, Regression, Visual data mining.

مقاله پژوهشی

تحلیل بارش‌های روزانه هفتاد ساله شهر تبریز با استفاده از داده‌کاوای بصری

محمدتقی ستاری^{۱*}، جواد وحدت^۲، سحر جاویدان^۳

تاریخ دریافت: ۱۴۰۰/۱۲/۱۵

تاریخ پذیرش: ۱۴۰۱/۰۴/۱۴

تاریخ دریافت ویرایش: ۱۴۰۱/۰۴/۰۵

تاریخ انتشار آنلاین: ۱۴۰۲/۷/۱

۱- دانشیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

۲- دانش‌آموخته کارشناسی ارشد، گروه آمار، دانشکده ریاضی، دانشگاه تبریز

۳- دانشجوی کارشناسی ارشد، گروه مهندسی آب، دانشکده کشاورزی، دانشگاه تبریز

*مسئول مکاتبات، پست الکترونیکی: mtsattar@gmail.com

چکیده

بارش روزانه با داشتن خصوصیتی کاملاً تصادفی یکی از مؤلفه‌های اساسی چرخه آب بوده و دارای نقش مهمی در مدیریت منابع آب‌های سطحی و زیرزمینی به لحاظ کمی و کیفی است. عدم وجود داده‌های طولانی مدت و قابل اتکای بارش، شناسایی رفتار آن را پیچیده نموده است. در این مقاله سعی شده از طریق تعامل انسان و داده که با عنوان داده‌کاوای بصری یاد می‌شود، اقدام به شناسایی رفتار و الگوی بارش نمود. در این پژوهش برای شناسایی الگوهای بارش از داده‌های روزانه بارش و دمای ایستگاه سینوپتیک تبریز در بازه زمانی هفتاد و یک ساله اخیر (۱۴۰۰-۱۳۳۰) استفاده شد. نتایج به دست آمده حاکی از تغییر الگوی بارش‌ها در ۵ سال اخیر (۱۴۰۰-۱۳۹۶) بود. علیرغم این که در این پنج سال میزان بارش سالانه بالای میانگین ۷۱ ساله بوده، ولی همچنان از دوره طلایی بارش در دهه ۴۰ پایین‌تر است. نتایج نشان داد که شدت بارش‌های بهار تبریز در دوره ۱۳۸۵ تا ۱۳۹۹ کاهش محسوسی داشته است. این در حالیست که در دهه‌های ۳۰ تا ۶۰ بارش‌های بهار هم به لحاظ میزان بارش و هم شدت بارش بیشتر بوده، اما در دوره‌های بعدی از دهه ۷۰ تا ۹۰ هم از شدت بارش و هم از میزان بارش کاسته شده است. یافته‌ها نشان داد بیشتر بارش بهار مربوط به سال ۱۳۶۰ به میزان ۲۷۶/۳ mm بوده که ۷۲/۲۶٪ از کل بارش آن سال را تشکیل می‌داد. همچنین رفتار دمایی در کل این مدت افزایش میانگین دما را نشان داد که تأییدی بر افزایش دمای کره زمین است.

واژه‌های کلیدی: الگوی بارش، بارش روزانه، داده‌کاوای بصری، رگرسیون، روند بارش.

مقدمه

بارش از عناصر مهم و پیچیده چرخه هیدرولوژیکی می‌باشد که برای توسعه اقتصادی-کشاورزی، مدیریت سیلاب-خشکسالی، برنامه‌ریزی و مدیریت منابع آب کشور بسیار حیاتی است. در یک منطقه و در طول سالیان دراز، بارش‌ها کم و بیش الگوهای رفتاری مشابهی از خود نشان می‌دهند. ولیکن کمبود داده‌های بارش، کشف و استخراج این الگوها را با مشکل مواجه می‌کند. طبیعتاً در صورتی که بتوان به داده‌های طولانی مدت بارش دسترسی پیدا کرد، زمینه برای تحلیل‌های واقع‌بینانه‌تر با ریسک و خطای کمتر مهیا خواهد شد. تغییرات زمانی- مکانی بارش و طبیعت تصادفی آن باعث توسعه روش‌های مختلف و دیدگاه‌های متفاوت در زمینه پیش‌بینی مقدار و شناسایی رفتار آن شده است. داده‌های بارش بصورت خام و به خودی خود توانایی کافی برای تحلیل پدیده بارش را ندارند. این در حالی است که تحلیل‌های آماری و نمودارها باعث جان گرفتن داده‌ها شده و می‌توانند پیام‌های گراندتری از ماهیت بارش ارائه کنند. حجم داده با گسترش فناوری و پیشرفت سخت‌افزارها و نرم‌افزارها روز به روز بیشتر می‌شود و می‌توان داده‌های بیشتری در زمینه‌های جدیدتری با دقت بالاتری را استخراج و ذخیره کرد. کاوش و تجزیه و تحلیل این ابرداده‌ها به‌طور فزاینده‌ای دشوار بوده و نیاز به سیستم‌هایی با قدرت پردازش بالا و فضایی ایمن برای ذخیره این حجم از داده‌ها است. به‌طور عادی نمی‌توان این حجم از داده‌ها را برای گزارشات مورد استفاده قرار داد، می‌بایست ابتدا تبدیل به اطلاعات شوند و در قالب‌های بصری مناسب که حاوی اطلاعات کافی از کل وضعیت داده‌ها باشند، مورد استفاده قرار گیرد. تجسم مناسب داده و داده‌کاوی بصری می‌تواند به مقابله با این ابرداده‌ها کمک کند. از مزیت کاوش داده‌های بصری می‌توان به این اشاره کرد که کاربر مستقیماً در فرآیند داده‌کاوی درگیر است و حتی در موارد زیادی اشتباهات موجود در مسیر

استخراج اطلاعات را به سبب ماهیت مسئله سریعاً متوجه می‌شود. در دهه گذشته و به سبب پیشرفت بسیار خوب منابع پردازشی و گرافیکی سیستم‌های کامپیوتری، تکنیک‌های تجسم اطلاعات برای پشتیبانی از اکتشاف مجموعه داده‌های بزرگ توسعه یافته‌اند. محققان دانشگاه برکلی تخمین می‌زنند که هر سال حدود یک اگزابایت (یک میلیون ترابایت) داده تولید می‌شود که البته در صورت وجود ابزار مناسب جهت رصد وقایع این عدد می‌تواند تا چندین برابر هم بیشتر شود (کیم ۲۰۰۲). هرچیزی که بتوان آن را ثبت کرد به عنوان داده حائز اهمیت هست. از دیدگاه افراد مختلف بسته به نیازشان اهمیت داده ثبت شده می‌تواند متفاوت باشد، ولی هر چیزی به عنوان داده و اطلاعات مفید بوده و ممکن است در آینده با گشوده شدن ابعاد تازه‌ای از علم محققین متوجه ارزش این داده‌های ثبت شده در گذشته شوند. برای اینکه داده‌کاوی مؤثر باشد، مهم است که انسان در فرآیند کاوش داده‌ها گنجانده شود و انعطاف‌پذیری، خلاقیت و دانش عمومی انسان با ظرفیت عظیم ذخیره‌سازی و قدرت محاسباتی رایانه‌های امروزی ترکیب شود. داده‌کاوی به‌عنوان علمی که به کاوش در مجموعه وسیعی از داده‌ها جهت کشف الگوهای پنهان آن می‌پردازد، در سال‌های اخیر مورد توجه دانشمندان هیدرولوژی و هواشناسی نیز قرار گرفته است. یکی از رویکردهای جدید که معطوف به بصری‌سازی و استفاده از نمودارها در تحلیل پیچیدگی‌های درون داده‌هاست، داده‌کاوی بصری می‌باشد. توسعه نرم‌افزاری و ایجاد پایگاه داده‌های بزرگ از یک سو و توسعه امکانات سخت‌افزاری ثبت داده‌ها باعث ایجاد تحولی عمیق در علم داده‌کاوی و مشخصاً داده‌کاوی بصری شده است. داده‌کاوی بصری به‌علت کمبود داده در بخش هیدرولوژی و هواشناسی کمتر امکان توسعه و کاربرد داشته است، اما پیش‌بینی می‌شود در سال‌های آتی با افزایش طول داده‌های اقلیمی، داده‌کاوی بصری بیش از گذشته مورد توجه دانشمندان منابع آب قرار گیرد. دنیل و همکاران (۲۰۰۳) در پژوهشی

رو، نتیجه گرفتند که هیچ تغییر آب و هوایی بر روی کویمباتور مشاهده نشده است. جهانبخش و همکاران (۲۰۱۵) به بررسی توزیع زمانی و مکانی بارش محدوده شهرستان تبریز با روش تحلیل خوشه‌ای براساس داده‌های ایستگاه‌های موجود در محدوده شهرستان تبریز پرداختند. نتایج محاسبات آن‌ها نشان داد که بیشترین میزان بارش محدوده مورد مطالعه مربوط به قسمت‌های جنوبی شهرستان می‌باشد که در این میان سهم ایستگاه‌های ليقوان و هروی چشمگیرتر بوده و قسمت‌های غربی شهرستان کمترین میزان بارش را دریافت کرده‌اند. جانا و همکاران (۲۰۱۶) روند بارش و روزهای بارانی آگرا در شمال هند را تحلیل نمودند. نتایج مطالعه آن‌ها نشان داد که در بارش سالانه، موسمی و زمستان و آگوست در آگرا روند کاهش قابل توجهی مشاهده می‌شود، در حالی که ماه مه افزایش بارش را نشان می‌دهد. دارند و ابراهیمی (۲۰۱۹) تغییرات مدت زمان انتظار رخداد بارش در استان کردستان را تجزیه و تحلیل نمودند. یافته‌ها نشان دادند که طی ماه‌های سرد سال مدت زمان انتظار رخداد بارش طولانی‌تر شده است، درحالی‌که در ماه‌های گرم سال فاصله زمانی بین دو رویداد بارشی پیاپی کوتاه‌تر شده است. نرخ کاهشی دوره انتظار رخداد بارش دست‌کم دو برابر نرخ افزایشی آن است. ایران‌پور و زهره‌وندی (۲۰۱۹) به منظور بررسی احتمال فراوانی و تداوم روزهای خشک و مرطوب، بارندگی‌های روزانه تبریز را تحلیل نمودند. آن‌ها نتیجه گرفتند که میانگین بارش در تبریز ۸۰ روز در سال و عمدتاً در فصول بهار و زمستان است. فردوسی و دین‌پژوه (۲۰۱۹) به بررسی توزیع زمانی بارش‌های روزانه با روش منحنی‌های بارش نرمال (NRC) در تبریز پرداختند. آن‌ها دریافتند با توجه به منحنی‌های NRC، در دوره آماری ۶۶ ساله، در ۲۵، ۵۰، ۷۵ و ۹۰٪ روزهای بارانی به ترتیب حدود ۲، ۱۰، ۳۰ و ۶۰٪ بارش در تبریز نازل شده است. گنتیلوچی و همکاران (۲۰۱۹) به تحلیل روند بارش شدید در سمت

مستقل به داده‌کاوای بصری مجموعه داده‌های مکانی بزرگ شامل پدیده‌های نقطه، خط و ناحیه پرداختند. آن‌ها دریافتند داده‌کاوای بصری ادراک انسانی را برای اکتشاف الگوهای درون مجموعه داده‌های بزرگ توسعه می‌دهد. ارائه داده‌ها به صورت تعاملی و گرافیکی اغلب بینش‌های جدیدی را تقویت کرده، شکل‌گیری و اعتبار فرضیه‌های جدید را برای پایان حل بهتر مسئله و کسب دانش عمیق‌تر مهیا می‌کند. وارد و همکاران (۲۰۰۴) تکنیک‌های تجسم، تعامل و مدیریت داده‌ها را برای برخورد مؤثر با مجموعه‌های داده‌ای که حاوی میلیون‌ها رکورد و یا صدها بعد هستند، توسعه دادند و روش‌هایی را برای اعمال رویکردهای مشابه برای گسترش سامانه برای مدیریت داده‌های اسمی و ترتیبی پیشنهاد کردند. وغ (۲۰۱۲) از داده‌کاوای بصری برای تعیین روابط خاک، اقلیم و کاربری زمین استفاده کرد. ارزیابی و مطالعه ایشان با استفاده از تکنیک‌های داده‌کاوای بصری برای شناسایی الگوهای انواع خاک انجام شد. نتایج حاکی از آن بود که کشاورزی در این مناطق با تنوع خاک بالا نیاز به مدیریتی متفاوت دارد. گارسیا و همکاران (۲۰۱۲) برای شناسایی الگوها و نقاط پرت در داده‌های ایستگاه‌های هواشناسی از داده‌کاوای بصری استفاده کردند. کنترل کیفیت داده‌های آب و هوایی به دست آمده از ایستگاه‌های هواشناسی ضروری است. یکی از راه‌های انجام این کنترل، مقایسه داده‌های دریافتی از یک ایستگاه با داده‌های ایستگاه‌های دیگر است که به نوعی انتظار می‌رود رفتار مشابهی از خود نشان دهند. هدف از این کار ارزیابی برخی تکنیک‌های داده‌کاوای بصری برای شناسایی گروه‌بندی (و نقاط پرت این گروه‌بندی‌ها) ایستگاه‌های هواشناسی با استفاده از داده‌های تاریخی بارش در یک بازه زمانی خاص بود. آرتی رانی و همکاران (۲۰۱۴) به تحلیل روند بارش و فراوانی روزهای بارانی در شهر کویمباتور کشور هند پرداختند. نتایج نشان داد که در بلندمدت، بارش ماهانه، فصلی، سالانه و فراوانی روزهای بارانی تغییری ایجاد نشده است. از این

برای گشودن پنجره‌ای کوچک به توسعه علم داده‌کاوی بصری در مهندسی آب است تا شاید مورد توجه پژوهشگران آب و آب-هواشناسی قرار گیرد. هدف پژوهش حاضر، تحلیل بارش‌های روزانه شهر تبریز در دوره آماری هفتاد و یک ساله و شناسایی الگوهای پنهان درون آن با استفاده از روش داده‌کاوی بصری، به کمک ابزارهای گرافیکی می‌باشد.

مواد و روش‌ها

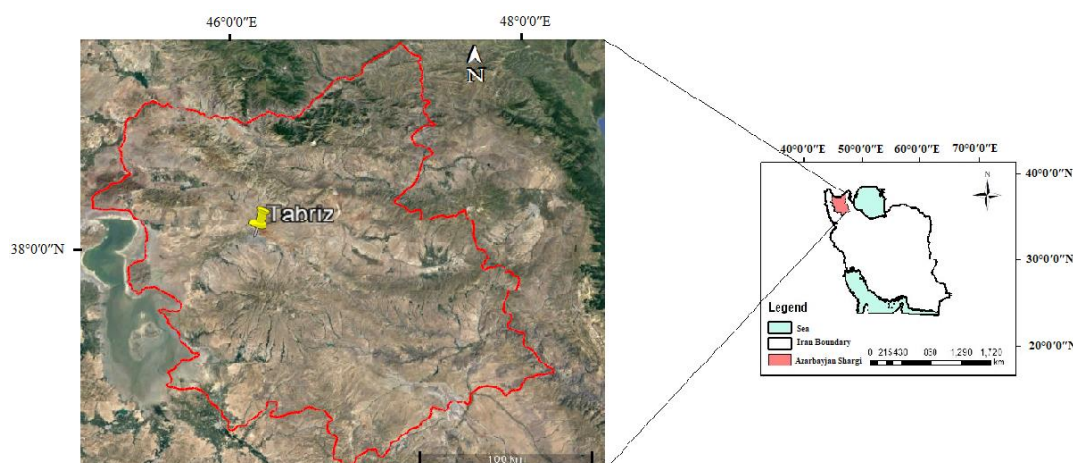
معرفی منطقه مورد مطالعه و داده‌های مورد استفاده

تبریز مرکز استان آذربایجان شرقی بوده و دارای مساحتی حدود 1200 km^2 می‌باشد. تبریز از شمال به کوه عینالی، از جنوب به دامنه کوه سهند و از غرب به جلگه تبریز و دریاچه ارومیه محدود است (عساکره و اکبرزاده ۲۰۱۷). در پژوهش حاضر از داده‌های مربوط به بارش روزانه ایستگاه سینوپتیک تبریز از سال ۱۳۳۰ تا ۱۴۰۰ استفاده شده است. داده‌های استفاده شده از اداره هواشناسی تبریز اخذ شده است. ایستگاه سینوپتیک تبریز در 46° درجه و $17'$ دقیقه طول شرقی و 38° درجه و $5'$ دقیقه عرض شمالی واقع شده است. ارتفاع آن از سطح دریا 1366 m است. مجموع بارش سالانه در ۷۱ سال اخیر به اندازه $283/6 \text{ mm}$ و متوسط دمای سالانه $12/4^\circ \text{C}$ بوده و بر اساس طبقه‌بندی دومارتن دارای اقلیم نیمه‌خشک می‌باشد. موقعیت مکانی ایستگاه مورد مطالعه در شکل ۱ آورده شده است:

آدریاتیک میانی، منطقه مارکه پرداختند. آن‌ها دریافتند که بارش شدید تحت تأثیر پارامترهای ژئوتوپوگرافی نیستند، بلکه بیشتر از همه تحت تأثیر دینامیک آب و هوا هستند. پاوار و همکاران (۲۰۲۰) به منظور تجزیه و تحلیل روند بلندمدت سطح منطقه‌ای از بارش، روزهای بارانی و خشکسالی در حوضه رودخانه سند مطالعه‌ای انجام دادند. بر اساس مطالعه آن‌ها تحلیل روند بارش و روزهای بارانی، روند کاهشی قابل توجهی را برای قسمت شمال شرقی حوضه نشان داد. نتایج تحلیل روند شاخص بارش استاندارد شده^۱ نیز نشان‌دهنده روند افزایشی خشکسالی در بخش بزرگی از حوضه بود. یاکوب و تایفور (۲۰۲۰) تغییرات خشکسالی هواشناسی و روند بارش در کل موریتانی را به صورت مکانی و زمانی تحلیل کردند. نتایج تجزیه و تحلیل روند نشان داد که هیچ کاهشی در بارش در منطقه شمالی مشاهده نشده است. قربانی و همکاران (۲۰۲۱) به تحلیل مکانی-زمانی بارش و میانگین دمای هوا در اردبیل پرداختند. نتایج مطالعه آن‌ها نشان داد که توزیع زمانی بارش یکنواخت نیست و به طور کلی عمده‌ترین مقادیر بارش در فصل بهار، اوایل پاییز و اواخر زمستان متمرکز شده است.

بررسی منابع نشان می‌دهد اگرچه بارش‌های تبریز با استفاده از روش‌های مختلفی طی مطالعات گوناگون مورد توجه پژوهشگران قرار گرفته است، اما تاکنون از روش داده‌کاوی بصری برای تحلیل بارش‌ها استفاده نشده است. پژوهش حاضر تلاشی هر چند اندک

^۱ Standardize Precipitation Index



شکل ۱- موقعیت مکانی ایستگاه سینوپتیک تبریز.

داده‌کاوی بصری^۲

داده‌کاوی بصری به‌عنوان مجموعه‌ای از تکنیک‌ها و رویکردهای مورد استفاده برای استخراج و درک اطلاعات رمزگذاری شده در مجموعه داده‌ها با استفاده از سیستم ادراک بصری انسان به‌عنوان بخشی از وظیفه پردازش داده است (سیموف و همکاران ۲۰۰۸). داده‌کاوی بصری ممکن است به کشف یا برجسته کردن ویژگی‌های داده، آسان‌تر و سریع‌تر کردن درک ویژگی‌ها و برای تأیید متقابل نتیجه‌گیری‌های به‌دست‌آمده از روش‌های دیگر کمک کند (کوپاناکیس و همکاران ۲۰۰۵). داده‌کاوی بصری دانش ضمنی و مفید را از مجموعه داده‌های بزرگ با استفاده از داده‌ها و یا تکنیک‌های تجسم دانش کشف می‌کند. سیستم بینایی انسان توسط چشم‌ها و مغز کنترل می‌شود که می‌توان آن‌ها را به‌عنوان یک موتور پردازش و استدلال قدرتمند، موازی و دارای پایگاه دانش بزرگ در نظر گرفت. داده‌کاوی بصری اساساً قدرت این مؤلفه‌ها را ترکیب می‌کند و آن را به ابزاری بسیار جذاب و مؤثر برای درک توزیع داده‌ها، الگوها، خوشه‌ها و نقاط پرت در داده‌ها تبدیل می‌کند. داده‌کاوی بصری را می‌توان به‌عنوان ادغام دو رشته تجسم داده و داده‌کاوی در نظر گرفت. همچنین داده‌کاوی بصری ارتباط نزدیکی با گرافیک کامپیوتری،

سیستم‌های چند رسانه‌ای، تعامل انسان و کامپیوتر، تشخیص الگو و محاسبات با کارایی بالا دارد (هان و پی ۲۰۱۲). ارائه داده‌ها به صورت تعاملی و گرافیکی اغلب بینش‌های جدیدی را تقویت می‌کند، شکل‌گیری و اعتبار فرضیه‌های جدید را برای پایان حل بهتر مسئله و کسب دانش عمیق‌تر حوزه تسهیل می‌کند. داده‌کاوی بصری معمولاً امکان کاوش سریع‌تر داده را فراهم می‌کند و اغلب نتایج بهتری را ارائه می‌دهد، به‌ویژه در مواردی که الگوریتم‌های داده‌کاوی خودکار با شکست مواجه می‌شوند. علاوه بر این، تکنیک‌های کاوش داده‌های بصری درجه بسیار بالاتری از رضایت و اطمینان کاربر را در یافته‌های کاوش فراهم می‌کند. این واقعیت منجر به تقاضای بالایی برای تکنیک‌های کاوش بصری می‌شود و آن‌ها را در ارتباط با تکنیک‌های کاوش خودکار ضروری می‌سازد (کیم و آنکرست ۲۰۰۵). در این پژوهش از نرم‌افزارهای مختلفی چون اکسل، R، ArcGIS و Tableau جهت بصری‌سازی استفاده گردیده است. ترکیب زبان‌های مختلف برنامه‌نویسی و استفاده از ویژگی‌های منحصر به فرد هر کدام باعث می‌شود تا در عین جذابیت نمودارها به حداکثر اطلاعات دریافتی در یک تصویر رسید. در این پژوهش از نرم‌افزار اکسل و افزونه اکسل استت^۳ برای مرتب کردن، آزمون‌های اولیه آماری برای

^۳ Xlstat

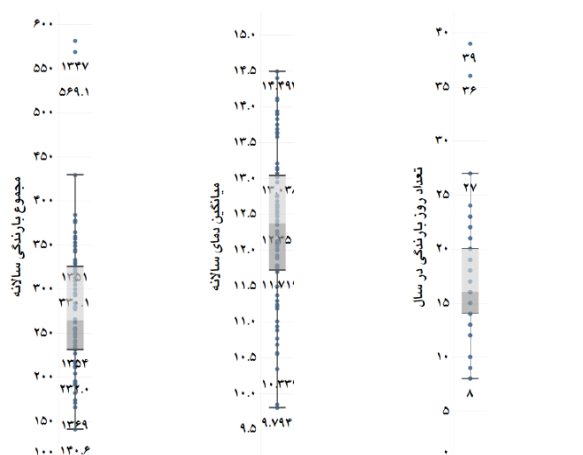
^۲ Visual Data Mining

در این پژوهش با آزمون‌های آماری، صحت و دقت داده‌ها مورد بررسی قرار گرفت. در جدول ۱ ویژگی‌های آماری پارامترهای میانگین دمای سالانه، مجموع بارش سالانه و روزهای بارانی در هر سال ارائه گردیده است.

جدول ۱- آماره‌های توصیفی داده‌های میانگین دمای سالانه، مجموع بارش سالانه و روزهای بارانی در هر سال.

تعداد روزهای بارانی در هر سال	میانگین دمای سالانه (C°)	مجموع بارش سالانه (mm)
۹	۶/۹	۱۴۰/۶
۳۹	۱۴/۶	۵۸۳
۱۳۶۷	۱۳۶۱	۱۳۹۶
۱۳۴۲	۱۳۹۶	۱۳۴۲
۱۴	۱۱/۸	۲۳۲/۲
۱۶	۱۲/۵	۲۷۲/۴
۱۷	۱۲/۴	۲۸۳/۶
۲۰	۱۳/۱	۳۳۴/۹

در شکل ۲ نمودار جعبه‌ای مربوط به تعداد دفعات بارش، مجموع بارش و میانگین دمای سالانه ارائه گردیده است.



شکل ۲- نمودار جعبه‌ای برای تعداد روزهای بارش، میانگین دمای سالانه، مجموع بارش سالانه.

بررسی صحت و دقت داده‌ها و فرمت خروجی داده‌ها استفاده گردید. از زبان آماری R برای ایجاد ساختار داده‌کاوی، تبدیل داده‌ها به فرمت دلخواه و مورد نیاز جهت نمایش گرافیکی و بیان ویژگی‌های توصیفی استفاده گردید. سپس با استفاده از برنامه مصورسازی Tableau نمودارهای مختلفی ترسیم و مورد ارزیابی قرار گرفت و نهایتاً نمودارهای نهایی انتخاب گردید. از نرم‌افزار ArcGIS هم به منظور ارزیابی مکانی و صحت داده‌های مکانی و نقشه‌های دقیق جغرافیایی استفاده شد.

رگرسیون چندجمله‌ای^۴

گاهی می‌توان یک مدل غیرخطی بین متغیرهای توصیفی و پاسخ به صورت یک چندجمله‌ای درجه k تعیین کرد. مدل رگرسیونی چندجمله‌ای در چنین مواقعی می‌تواند رابطه بین متغیرهای پیش‌گو و وابسته را به خوبی بیان کند (استرتاگوا ۲۰۱۲). معادله یک خط راست توسط دو نقطه از آن خط، قابل تعیین است. هر چه تعداد نقاط بیشتر باشد، درجه منحنی یا چندجمله‌ای که از آن نقاط بگذرد، بیشتر خواهد شد. یک رابطه یا تابع چندجمله‌ای مرتبه k به صورت زیر نوشته می‌شود.

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_kx^k \quad [1]$$

رگرسیون چندجمله‌ای سعی دارد با استفاده از تعداد متغیرهای بسیار کمتر از تعداد نقطه‌ها، پارامترهای چندجمله‌ای یا رگرسیون را برآورد کند. در نتیجه پیچیدگی مدل کاهش یافته و خطا به وجود می‌آید. به این ترتیب توازن بین بیش‌برازش^۵ و کم‌برازش^۶ در مدل برقرار خواهد شد. در این پژوهش از رگرسیون چندجمله‌ای برای پیش‌بینی مقدار بارش و احتمال وقوع آن استفاده گردیده است.

نتایج و بحث

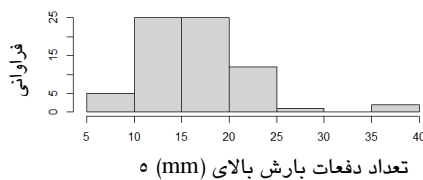
^۱ Underfitting

^۴ Polynomial Regression

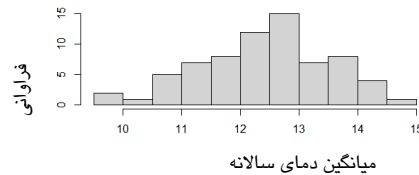
^۵ Overfitting

۷۱ سال اخیر خواهند داشت. اما داده‌های مجموع بارش سالانه از جمع بارش در هر سال به دست آمده و واریانس بیشتری دارد. شکل ۳ بافت‌نگار توزیع فراوانی میانگین دماهای سالانه و شکل ۴ توزیع فراوانی دفعات بارش بالای ۵ mm در هر سال در ایستگاه تبریز را برای دوره مطالعاتی هفتاد و یک ساله نشان می‌دهد.

در شکل ۲ نمودار داده‌های پرت به صورت نقاطی دور از آخرین قسمت نمودار جعبه‌ای به نمایش گذاشته شده است. در نمودار جعبه‌ای از پایین به بالا کمترین داده‌ها، چارک اول، میانگین، چارک سوم و بیشترین داده‌ها مشاهده می‌شود. همانطور که دیده می‌شود، میانگین دمای سالانه داده‌های منسجم‌تری دارد که علت آن استفاده از تابع میانگین برای دماهای کل طول سال است. معیار میانگین خاصیت تمایل به مرکز دارد و لذا این داده‌ها پس از میانگین‌گیری، واریانس کمتری را در



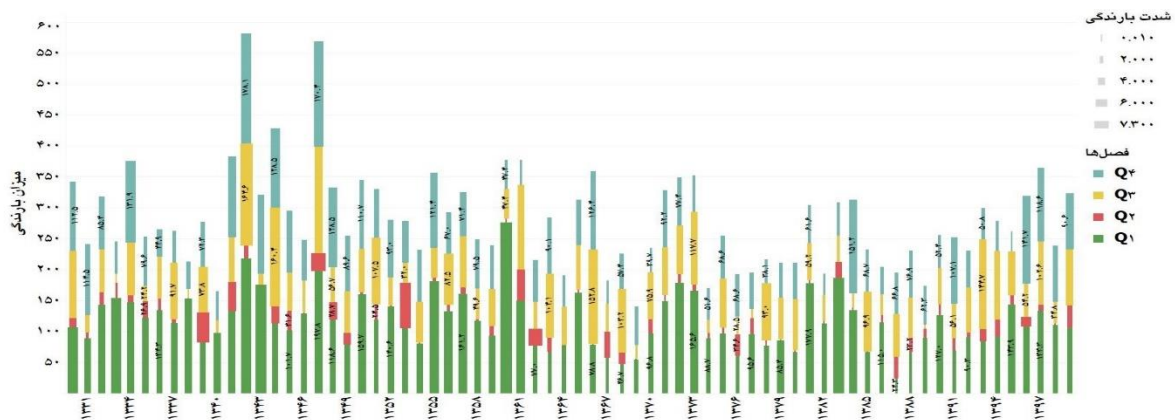
شکل ۴- نمودار هیستوگرام توزیع فراوانی دفعات بارش بالای ۵ میلی‌متر در هر سال.



شکل ۳- نمودار هیستوگرام توزیع فراوانی میانگین دماهای سالانه.

بارش بالای ۵ mm اتفاق افتاده است. در حالی که سال‌های بسیار کمی بوده‌اند که بارش در آن سال‌ها بیشتر از ۲۵ روز اتفاق افتاده است. شکل ۵ نمودار بارش سالانه به تفکیک هر فصل همراه با شدت بارش در دوره مطالعاتی را نشان می‌دهد.

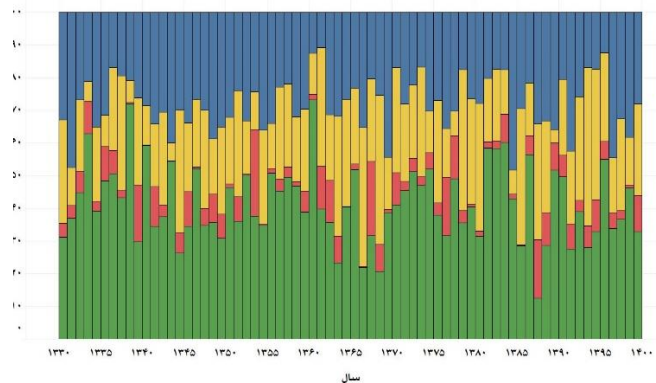
با توجه به شکل ۳ می‌توان مشاهده کرد که در اغلب سال‌ها میانگین دمای سالانه بین 12°C تا 13°C بوده است و به نوعی نمودار زنگوله‌ای شکل (توزیع نرمال) را تشکیل داده است. میانگین این نمودار همانطور که در جدول ۱ آمده است برابر $12/41^{\circ}\text{C}$ می‌باشد. با توجه به شکل ۴ در اکثر سال‌ها حدوداً ۱۰ تا ۲۰ روز



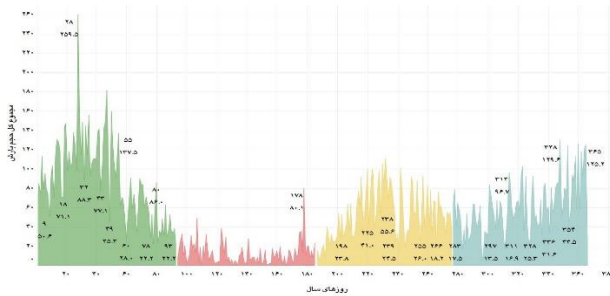
شکل ۵- نمودار بارش سالانه به تفکیک هر فصل همراه با شدت بارش.

در شکل ۶ هر ستون نشان‌دهنده یک سال شمسی بوده که به چهار بخش مختلف تقسیم شده و نمایانگر فصول بهار، تابستان، پاییز و زمستان که به ترتیب با رنگ‌های سبز، قرمز، نارنجی و آبی نشان داده شده است. در سال ۱۳۶۰ بیشترین درصد بارش بهار رخ داده به طوری که $73/26\%$ از کل بارش آن سال در این فصل باریده است. همانطور که ملاحظه می‌شود در دهه اخیر از سال ۱۳۹۱ تا ۱۳۹۵ درصد بارش‌های پاییزه افزایش یافته و بیشتر بارش‌های فصول سرد در پاییز رخ داده است. اما در سال ۱۳۹۶ این الگو عوض شده و در آن سال $44/46\%$ از کل بارش سالانه در زمستان رخ داد. کمترین درصد بارش فصلی همان‌طور که انتظار می‌رفت مربوط به تابستان می‌باشد. بطوریکه در سال ۱۳۴۳ فقط $0/03\%$ از کل بارش آن سال در فصل تابستان باریده است. بیشترین سهم بارش تابستانه مربوط به سال ۱۳۵۳ بوده بطوریکه $26/61\%$ بارش آن سال در تابستان رخ داده است. کمترین درصد بارش بهار مربوط به سال ۱۳۸۷ و به میزان $12/42\%$ از کل بارش همان سال است. کمترین و بیشترین درصد بارش پاییزه مربوط به سال‌های ۱۳۸۹ و ۱۳۹۳ که به ترتیب $4/18$ و $48/32\%$ از کل بارش‌های آن سال‌ها را تشکیل داده است. این آمار برای فصل زمستان برای کمترین و بیشترین درصد بارش زمستانه مربوط به سال‌های ۱۳۶۱ و ۱۳۸۴ به ترتیب $10/86$ و $48/31\%$ بوده است. شدت بارش‌ها به تفکیک هر فصل در طول دوره مطالعاتی هفتاد و یک ساله در شکل ۷ دیده می‌شود.

شکل ۵ نمودار مجموع کل بارش‌های سالانه شهر تبریز را از ۱۳۳۰ تا ۱۴۰۰ به تفکیک فصول نشان می‌دهد. در این نمودار رنگ سبز نشان‌دهنده بارش در فصل بهار، رنگ قرمز بارش در تابستان، رنگ زرد بارش در پاییز و رنگ آبی میزان بارش در زمستان می‌باشد. پهنای هر میله در هر فصل نشان‌دهنده شدت بارش می‌باشد که از تقسیم میزان بارش به مدت زمان بارش به دست آمده است. همچنان‌که در شکل ۵ ملاحظه می‌شود در طول ۷۱ سال دوره مطالعاتی، بیشترین بارش‌ها در دهه ۴۰ شمسی اتفاق افتاده است. بیشترین بارش‌های سالانه به ترتیب مربوط به سال ۱۳۴۲ به میزان 583 mm و سال ۱۳۴۷ به میزان 577 mm و کمترین بارش در سال ۱۳۶۹ به میزان 140 mm اتفاق افتاده است. در سال ۱۳۴۱ شدت بارش در فصل پاییز و در سال ۱۳۴۷ در تابستان بیشتر بوده است. با توجه به شکل ۵ می‌توان مشاهده کرد که شدت بارش‌های بهار تبریز در دوره ۱۳۸۵ تا ۱۳۹۹ کاهش محسوسی یافته است. این در حالیست که در دهه‌های ۳۰ تا ۶۰ بارش‌های بهار هم به لحاظ میزان بارش و هم شدت بارش بیشتر بوده است. اما در دوره‌های بعدی از دهه ۷۰ تا ۹۰ هم از شدت بارش و هم از میزان بارش کاسته شده است. بیشتر بارش بهار مربوط به سال ۱۳۶۰ به میزان $276/3$ mm که $73/26\%$ از کل بارش آن سال را شامل شده است.

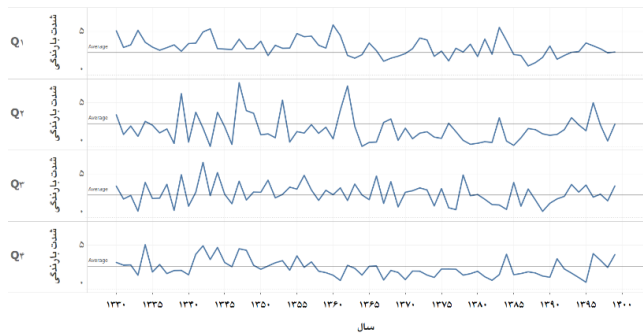


شکل ۶- نمودار درصدی بارش در هر فصل از سال.



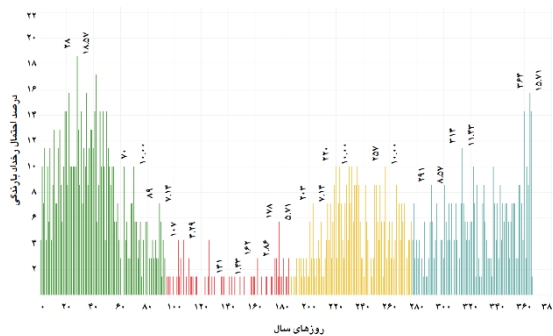
شکل ۸- نمودار مجموع بارش‌های روزانه تبریز در طول دوره هفتاد و یک ساله (۱۳۳۰-۱۴۰۰) به تفکیک هر فصل.

در شکل ۸ مجموع بارش‌های روزانه در طول ۳۶۵ روز از سال مشاهده می‌شود. نمودار نشان می‌دهد که در هر روز از سال در طول ۷۱ سال گذشته در مجموع چند میلی‌متر باران باریده است. به دلیل وجود روزهایی بدون باران با مقدار بارش صفر، از معیار میانگین استفاده نگردید. معیار میانگین یک معیار تمایل به مرکز هست و باعث می‌شود رخدادهایی که بارش زیاد باریده با زمانی که باران نباریده در اثر این معیار (میانگین گرفتن) از بین بروند. از این‌روی از عملکرد مجموع استفاده گردید تا اطلاعات بسندگی کافی را داشته باشند. رنگ‌های سبز، قرمز، زرد و آبی به ترتیب بیانگر فصل‌های بهار، تابستان، پاییز و زمستان می‌باشد. همانطور که انتظار می‌رفت، بیشترین بارش‌ها معمولاً از دهه سوم اولین ماه بهار (فروردین) شروع و تا اواخر دومین ماه بهار (اردیبهشت) ادامه پیدا کرده است. در فرهنگ محلی به این بارش‌ها اصطلاحاً بارش‌های نیسان (لیسان) گفته می‌شود. همچنان‌که از شکل ۸ ملاحظه می‌شود کمترین میزان بارش در طول این مدت به فصل تابستان تعلق داشته است. در فصل تابستان در بازه‌ای از روزهای ماه دوم تابستان (مرداد) در طول این ۷۱ سال تقریباً هیچ وقت بارشی اتفاق نیافتاده است. همچنین شکل ۸ نشان می‌دهد در روز ۲۸ سال که در واقع روز ۲۸ فروردین ماه می‌باشد، در طول دوره مطالعاتی هفتاد و یک ساله جمعاً ۲۵۹ mm و به تعداد ۳۸ بار بارش اتفاق افتاده است. در ارتباط با نمودار شکل ۸ لازم به توضیح هست که عدد بالایی شماره روز از سال و عدد پایینی میزان کل بارش



شکل ۷- نمودار شدت بارش‌ها به تفکیک هر فصل از بالا به پایین بهار، تابستان، پاییز و زمستان.

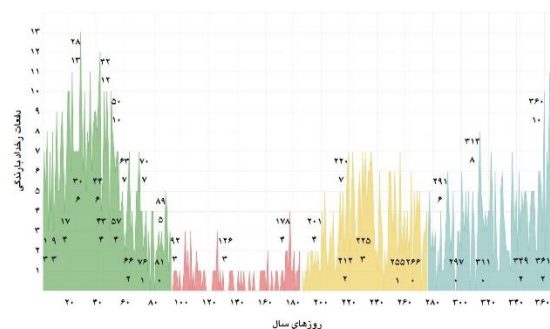
در شکل ۷، ردیف اول از بالا شدت بارش‌ها در فصل بهار را نشان می‌دهد. خط میانگین بارش که برای هر چهار ردیف به ترتیب از بالا بهار، تابستان، پاییز و زمستان کشیده شده است، بیانگر میانگین شدت بارش در کل دوره مطالعاتی می‌باشد که برابر با ۲/۶ mm است. در فصل بهار تا سال ۱۳۶۲ شدت بارش‌ها کاملاً بالای میانگین می‌باشد. از سال ۱۳۶۲ تا سال ۱۳۸۵ تقریباً حول میانگین می‌باشد و در دهه ۱۳۸۵ تا ۱۳۹۵ شدت بارش‌های بهار زیر میانگین قرار دارد که متأسفانه در این بازه میزان بارش هم افت قابل مشاهده‌ای داشته است. در سال‌های اخیر این شاخص بهبود یافته و در بالای میانگین قرار گرفته است. همانطور که انتظار می‌رود این شدت بارش در فصل تابستان در اکثر سال‌ها زیر خط میانگین قرار دارد. در فصل پاییز نیز شدت بارش‌ها عموماً حول میانگین قرار داشته و در سال‌های ۱۳۸۰ تا ۱۳۹۲ این شاخص افت قابل ملاحظه‌ای داشته است. در فصل زمستان تا سال ۱۳۵۷ شدت بارش‌ها بالای خط میانگین قرار دارد که شواهد تاریخی نیز مؤید این حجم و شدت بارش‌ها در زمستان می‌باشد. در سال‌های بعد از آن یعنی ۱۳۵۸ تا ۱۳۹۵ شدت بارش‌ها عموماً در زیر خط میانگین قرار داشته اما سال‌های اخیر در بالای این میانگین قرار گرفته است. در هر یک از این چهار فصل در چندسال اخیر شاهد افزایش شاخص شدت بارش هستیم.



شکل ۱۰- نمودار ستونی احتمال رخداد بارش در هر روز از سال در دوره هفتاد و یک ساله (۱۴۰۰-۱۳۳۰) به تفکیک فصل.

شکل ۱۰ نمودار ستونی احتمال وقوع رخداد بارش بالای ۵ mm در هر روز از سال به تفکیک فصل در طول دوره مطالعاتی را نشان می‌دهد. احتمال براساس داده‌های بارش روزانه و به شکل تقسیم تعداد رخداد بارش در یک روز خاص در طی این دوره هفتاد و یک ساله به عدد ۷۱ بدست آمده است. همانطور که در شکل ۱۰ مشاهده می‌گردد، بیشترین احتمال بارش در روز خاصی از سال مربوط به روزهای فصل بهار و زمستان می‌باشد. در این نمودار، اعداد روی شکل بدین صورت است که عدد اول بیانگر روز از سال و عدد دوم بیانگر احتمال رخداد بارش در آن روز می‌باشد شکل ۱۱ نمودار حبابی بارش ایستگاه تبریز در هر روز از سال در دوره مطالعاتی هفتاد و یک ساله را به تفکیک فصول نشان می‌دهد. شکل ۱۱ روزهای سال را به ترتیب از مرکز دایره به بیرون نشان می‌دهد و بزرگی و کوچکی شعاع دایره‌ها، میزان بارش (بالای ۵ mm) در آن روز از سال را براساس داده‌های ۷۱ ساله به تفکیک فصول نشان می‌دهد. همچنان‌که از شکل ۱۱ استنباط می‌شود، در روزهای ۲۸-ام و ۴۲-ام سال، مربوط به فصل بهار بیشترین بارش‌ها اتفاق افتاده است. در بعضی از روزهای تابستان اصلاً بارش وجود نداشته یا خیلی کم بوده است، بنابراین آن روزها در نمودار نمایش داده نشده است.

در آن روز می‌باشد. به عنوان مثال در روز ۲۹۶ که در فصل زمستان واقع هست مجموعاً ۵/۶ mm در طول دوره مطالعاتی بارش اتفاق افتاده است. به دلیل استفاده از این نوع سبک نمایش داده که حداکثر اطلاعات را نشان می‌دهد و به سبب تعداد روزهای سال و محدودیت عرض نمایش بعضی روزها با اطلاعات روزهای حول خود تجمع شده که نمایش بهتر شود و در جاهایی که اطلاعات آن روز خاص متفاوت از بقیه بوده سعی شده به صورت برچسب، روی نمودار عدد روز و میزان بارش آن آورده شود. شکل ۹ نمودار تعداد رخداد بارش بالای ۵ mm ایستگاه سینوپتیک تبریز در هر روز از سال در دوره هفتاد و یک ساله به تفکیک فصول را نشان می‌دهد. با توجه به شکل ۹ می‌توان مشاهده کرد که بیشترین تعداد رخداد بارش مربوط به فصل بهار بوده (۲۸ فروردین ماه دقیقاً ۱۳ بار در طول ۷۱ سال بارش بالای ۵ mm اتفاق افتاده است) و بعد از آن به ترتیب زمستان، پاییز و تابستان بیشترین رخداد بارش را داشته‌اند.

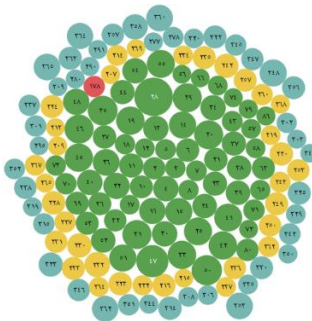


شکل ۹- نمودار تعداد رخداد بارش بالای ۵ میلی‌متر در هر روز از سال در دوره هفتاد و یک ساله (۱۴۰۰-۱۳۳۰) به تفکیک فصل.

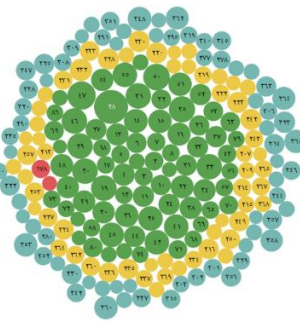
لازم به ذکر هست که وقوع بارش در یک روز برای بارش بالای ۵ mm در نظر گرفته شده است و کمتر از این مقدار به عنوان بارش لحاظ نگردیده است. با این فرض در طول فصل تابستان روزهایی از ماه و یا بازه‌هایی از چندین روز بودند که حتی یک بار هم در این ۷۱ سال بارش اتفاق نیفتاده است.

در این نمودار سعی شده تا فضای خالی ایجاد نشود، به همین سبب بعضی روزها که به لحاظ میزان بارش با روز بعد خود متفاوت بودند، کمی جابه‌جا شدند تا نمودار فضای خالی کمتری را در بین حباب‌ها داشته باشد و این کار زیبایی این نمودار را دوچندان می‌کند.

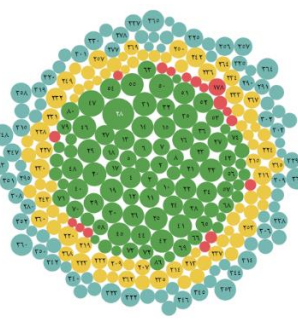
شکل ۱۲ نمودار حبابی بارش‌های بالای ۱۰ mm، ۲۰ mm، ۴۰ میلی‌متر و ۶۰ mm در هر روز از سال در دوره هفتاد و یک ساله تبریز را به تفکیک فصل نشان می‌دهد.



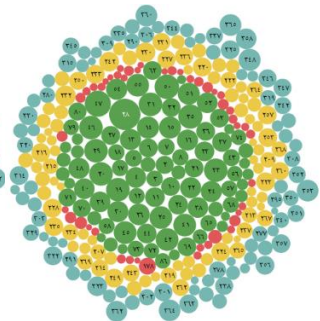
د



ج



ب



الف

شکل ۱۲- نمودار حبابی بارش‌های بالای الف) ۱۰ میلی‌متر (ب) ۲۰ میلی‌متر (ج) ۴۰ میلی‌متر (د) ۶۰ میلی‌متر در هر روز از سال در دوره هفتاد و یک ساله (۱۳۳۰-۱۴۰۰) به تفکیک فصل.

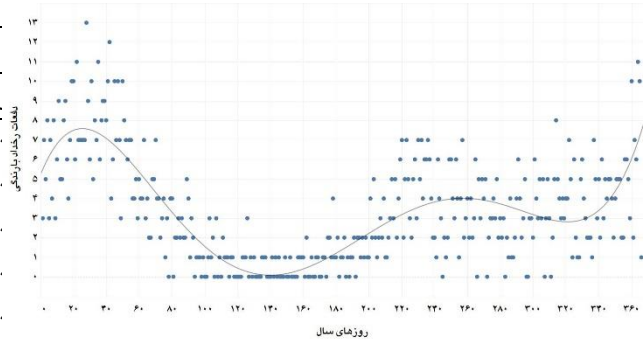
بطوریکه با هربار افزایش آستانه نمایش بارش این رنگ‌ها (حباب‌های قرمز) کمتر شده و در نهایت یک روز باقی مانده است. همچنین می‌توان دریافت که عمده بارش‌ها در فصل بهار بوده است. چراکه با هر بار تغییر آستانه نمایش مجموع بارش روزانه، رنگ سبز در نمودار که نمایانگر روزهای فصل بهار است، تغییر آنچنانی رخ نداده و عمده بارش‌ها را به خود اختصاص داده‌اند. شکل ۱۳ نمودار نقطه‌ای تعداد روزهای بارانی در هر روز از سال در طول دوره آماری هفتاد و یک ساله را نشان

در شکل ۱۲ به ترتیب آستانه نمایش مجموع بارش در روزهای سال در طول دوره مطالعاتی افزایش پیدا کرده تا بتوان روزهای با بارش بیشتر را بهتر مشاهده نمود. با گذر از هر کدام از این آستانه‌ها روزهای حذف شده (بارش‌های کمتر از ۵ mm) را می‌توان مشخص نمود. در شکل ۱۲-الف چون آستانه بارش mm ۱۰ در نظر گرفته شده، بنابراین تعداد روزهای زیادی در نمودار دیده می‌شود که با گذر از این تصاویر و در قسمت د مشاهده می‌شود که فقط یک روز بارانی در تابستان وجود داشته (روز ۱۷۸) که مجموع بارش در کل طول دوره مطالعاتی بالای ۶۰ mm بوده است. لذا تصاویر این چینی به وضوح، میزان کم شدن روزهایی با میزان بارش بالا را نمایش می‌دهند و می‌توان فهمید که در اکثر روزهای تابستان مجموع بارش روزانه در طول دوره مطالعاتی در بازه کمتر از ۶۰ mm بوده است.

می‌دهد.

جدول ۲- ضرایب مدل غیرخطی رابطه ۲.

مقدار	ضرایب		
	مقدار	خطای استاندارد	t آماره
a_1	$1/414e-10$	$1/176e-11$	$12/02$
a_2	$-1/315e-07$	$1/084e-08$	$-12/13$
a_3	$4/286e-05$	$3/611e-06$	$11/86$
a_4	$-0/00556$	$0/00052$	$-10/61$
a_5	$0/20784$	$0/03111$	$6/68$
a_0	$5/22805$	$0/56768$	$9/22$



شکل ۱۳- نمودار نقطه‌ای تعداد روزهای بارانی در هر روز از سال در طول دوره آماری هفتاد و یک ساله (۱۴۰۰-۱۳۳۰).

در جدول ۲ همانطور که دیده می‌شود ضرایب a_1 تا a_3 نزدیک به صفر می‌باشند و قابل چشم‌پوشی هستند. اگر مدل به مرتبه ۲ کاهش داده شود، با تحمل کمی خطا در پیش‌بینی به مدل ساده‌تری می‌توان رسید، ولی از آنجایی که منطق مدل بارش‌ها هیچ همخوانی با نمودار سهمی درجه ۲ ندارد، مقدار ضریب تبیین به شدت کاهش پیدا می‌کند و از این‌روی از این کار خودداری شد. در جدول ۳ حالت‌های مختلف مراتب مدل چندجمله‌ای آورده شده است.

جدول ۳- حالت‌های مختلف مراتب مدل چندجمله‌ای.

مجموع خطای استاندارد ضرایب	مقدار ضریب تبیین	مرتبه رگرسیون چندجمله‌ای
$0/3526$	$0/3588$	۲
$0/4523$	$0/4265$	۳
$0/5761$	$0/4295$	۴
$0/5993$	$0/5930$	۵
$0/7061$	$0/7092$	۶
$0/8352$	$0/7092$	۷

همانطور که در جدول ۳ مشاهده می‌شود در مدل مرتبه ۶ و ۷ مقدار ضریب تبیین افزایش پیدا نمی‌کند، ولی خطای استاندارد متغیرها افزایش پیدا می‌کند که قابل قبول نیست. هدف از افزایش مرتبه مدل رگرسیون چندجمله‌ای رسیدن به ضریب تبیین بالاتر با کنترل خطای استاندارد می‌باشد. از این‌روی در این مدل، مرتبه ۵ مناسب‌تر تشخیص داده می‌شود.

در شکل ۱۳ محور افقی روزهای سال و محور عمودی تعداد دفعات بارش (بالای ۵ mm) در آن روز از سال در طی دوره مطالعاتی هفتاد و یک ساله را نشان می‌دهد. اگر بتوان منحنی مناسبی با کمترین خطا از بین نقاط برازش داد، می‌توان تعداد احتمالی وقوع بارش در هر روز در طول دوره آماری را پیش‌بینی نمود. در این تحقیق بدین منظور مدل رگرسیون چندجمله‌ای با مرتبه ۵ و با پی-مقدار کمتر از $0/0001$ و ضریب تبیین $0/593$ مناسب تشخیص داده شد (رابطه ۲). در این مدل شماره روز در هر سال به عنوان متغیر مستقل و تعداد دفعات بارش به عنوان متغیر پاسخ وابسته در نظر گرفته شده است. بطور واضح مشخص است که پس از پیش‌بینی و بدست آوردن مقدار دفعات بارش در یک روز خاص از سال می‌توان آن را به تعداد کل دوره مطالعه شده (۷۱ سال) تقسیم و احتمال رخداد بارش در آن روز را بدست آورد. رابطه مدل برازش داده شده به شکل زیر می‌باشد:

$$[2] \quad = a_0 + a_1 D^5 + a_2 D^4 + a_3 D^3 + a_4 D^2 + a_5 D$$

در رابطه فوق، a ضریب مدل غیرخطی و D

شماره روز در هر سال است.

جدول ۴- ضرایب مدل غیرخطی رابطه ۳.

	ضرایب			پی-مقدار
	مقدار	خطای استاندارد	مقدار آماره t	
a_1	$-۸/۳۰۱e-۱۲$	$۱/۶۴۱e-۱۲$	$-۵/۰۵۹۷۷$	$< ./.۰۰۰۱$
a_2	$۱/۱۲۸e-۰۸$	$۱/۸۱۳e-۰۹$	$۶/۲۷۸۶۸$	$< ./.۰۰۰۱$
a_3	$-۵/۹۰۷e-۰۶$	$۷/۶۶۴e-۰۷$	$-۷/۷۰۶۶۸$	$< ./.۰۰۰۱$
a_4	$-۰/۰۰۱۴۳۴۹$	$۰/۰۰۰۱۵۴۸$	$۹/۲۶۷۴$	$< ./.۰۰۰۱$
a_5	$-۰/۱۵۹۲۸$	$۰/۰۱۵۱۶۶۸$	$-۱۰/۵۰۱۹$	$< ./.۰۰۰۱$
a_6	$۶/۱۰۸۹۴$	$۰/۶۴۲۰۴۲$	$۹/۵۱۴۸۷$	$< ./.۰۰۰۱$
a_0	$۵۲/۴۴۵۷$	$۸/۵۶۲۵۶$	$۶/۱۲۵$	$< ./.۰۰۰۱$

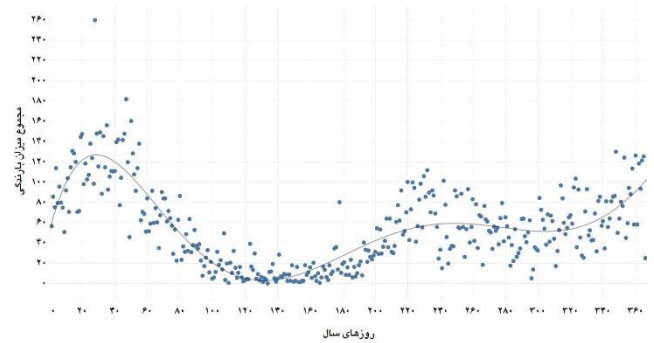
مشابه با جدول ۲ در جدول ۴ هم ضرایب a_1 تا a_4 تأثیر بسیار کمی در مدل دارند و این خاصیت مدل رگرسیون چندجمله‌ای هست که وزن بیشتری به متغیرها با توان کمتر می‌دهد.

جدول ۵- مقایسه مدل رگرسیون چندجمله‌ای با مراتب

مختلف.

مرتبه رگرسیون چندجمله‌ای	مقدار ضریب تبیین	مجموع خطای استاندارد ضرایب
۲	$۰/۳۹۸۵$	$۴/۹۶۷۹$
۳	$۰/۴۶۲۳$	$۶/۳۷۰۶$
۴	$۰/۴۶۸۴$	$۸/۰۸۹۸$
۵	$۰/۶۶۲۶$	$۷/۹۳۶۵$
۶	$۰/۶۸۵۱$	$۹/۲۱۹۹$
۷	$۰/۶۸۶۹$	$۱۰/۸۷۴۶$

در جدول ۵ به مقایسه مدل رگرسیون چندجمله‌ای با مراتب مختلف پرداخته شد و مشاهده گردید که با لحاظ کردن بیشترین میزان ضریب تبیین و کنترل خطای استاندارد مدل مرتبه ۶ برای این مدل نسبتاً مناسب‌تر هست.



شکل ۱۴- نمودار نقطه‌ای مقادیر بارش در هر روز از سال در طول دوره آماری هفتاد و یک ساله (۱۳۳۰-۱۴۰۰).

در نمودار شکل ۱۴ محور افقی روزهای سال و محور عمودی میزان بارش در آن روز مشخص از سال را به طور مجموع بارش ۷۱ ساله به صورت نقطه‌ای نشان می‌دهد. اگر بتوان منحنی مناسبی با کمترین خطا از بین نقاط برازش داد، می‌توان مقدار بارش در هر روز از سال در طول دوره آماری را پیش‌بینی نمود. در این تحقیق بدین منظور مدل رگرسیون چندجمله‌ای با مرتبه ۶ و با پی-مقدار کمتر از $۰/۰۰۰۱$ و ضریب تبیین $۰/۶۸$ مناسب تشخیص داده شد (رابطه ۳). در این مدل شماره روز در هر سال به عنوان متغیر مستقل و مقدار بارش به عنوان متغیر پاسخ وابسته در نظر گرفته شده است.

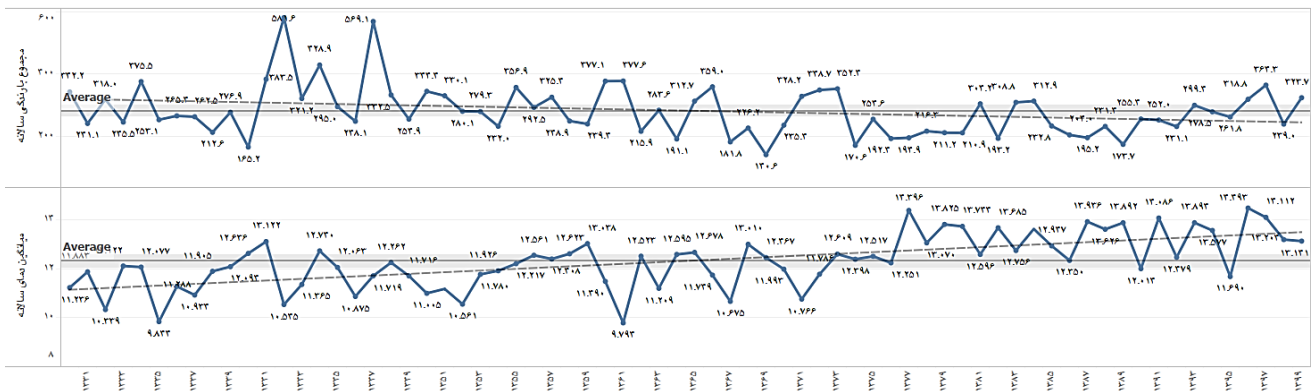
$$[۳] \quad \text{میزان بارش در روز} = a_0 + a_1 D^6 + a_2 D^5 + a_3 D^4 + a_4 D^3 + a_5 D^2 + a_6 D$$

در رابطه فوق، a ضریب مدل غیرخطی و D

شماره روز در هر سال است.

جدول ۴ ضرایب مربوط به مدل غیرخطی رابطه ۳

را نشان می‌دهد.



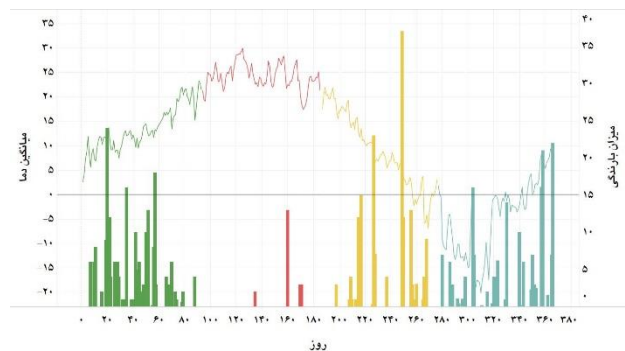
شکل ۱۵- نمودار مجموع بارش سالانه و میانگین دمای سالانه به همراه خط روند.

بارش‌هاست. زیرا در سالی که میانگین دمای سالانه پایین بوده میزان بارش در آن سال نسبت به سال‌های پیشین خودش زیاد بوده است. نمودار دمای سالانه با شیب خیلی ملایم رو به افزایش می‌باشد که بیانگر افزایش نسبی دما می‌باشد. این موضوع می‌تواند دلیل کاهش بارش‌های چند دهه اخیر در ایستگاه تبریز باشد. اما همانطور که از نمودار مشخص است از سال ۱۳۹۶ شیب نمودار میانگین دمای سالانه روبه پایین است و روند بارش‌های سالانه از این سال به جز سال ۱۳۹۸ افزایشی است. جهت بررسی تغییرات توأم بارش و دما شکل‌های ۱۶ و ۱۷ به عنوان نمونه برای سال‌های ۱۳۹۹ و ۱۳۹۲ ترسیم گردید.

در شکل ۱۵ در ردیف اول از بالا نمودار مجموع بارش سالانه در طول دوره مطالعاتی و در ردیف دوم میانگین دمای سالانه در طول دوره رسم شده است. خط‌چین رسم شده روند داده‌ها را نشان می‌دهد و خط مستقیم به همراه حاشیه کناری میانگین کل داده‌ها به همراه فاصله اطمینان ۹۵٪ را نشان می‌دهد. هرچا اگر خط روند بالای این میانگین قرار گیرد، به منزله روند مثبت و صعودی و اگر خط روند زیر این میانگین قرار گیرد، به منزله روند منفی و نزولی است. در نواحی فاصله‌های اطمینان حضور این دو خط میانگین و خط روند به معنای خنثی بودن روند (عدم وجود روند) می‌باشد. همانطور که در شکل ۱۵ دیده می‌شود، میانگین دمای سالانه و میزان بارش سالانه قرینه همدیگر هستند که سازگار با طبیعت



شکل ۱۷- تغییرات توأم دما و بارش در سال ۱۳۹۹.



شکل ۱۶- تغییرات توأم دما و بارش در سال ۱۳۹۲.

داده شده است. محور عمودی سمت چپ میانگین دما را در هر روز به صورت نمودار خطی نشان می‌دهد. یک خط افقی از نقطه صفر جهت مشخص نمودن نقطه صفر

در اشکال ۱۶ و ۱۷ محور افقی روزهای سال و نمودار میله‌ای میزان بارش در آن روز سال را نشان می‌دهد که در محور عمودی سمت راست این اعداد نشان

در حالیست که در دهه‌های ۳۰ تا ۶۰ بارش‌های بهاره هم به لحاظ میزان بارش و هم شدت بارش بیشتر بوده است. اما در دوره‌های بعدی از دهه ۷۰ تا ۹۰ هم از شدت بارش و هم از میزان بارش کاسته شده است. بیشتر بارش بهاره مربوط به سال ۱۳۶۰ به میزان ۲۷۶ mm که ۷۳/۲۶٪ از کل بارش آن سال را شامل شده است. کاهش عملکرد غلات دیم در منطقه آذربایجان می‌تواند تحت تأثیر کاهش بارش‌های بهاره باشد. در نیمه اول دهه اخیر از سال ۱۳۹۱ تا ۱۳۹۵ درصد بارش‌های پاییزه افزایش یافته است، یعنی بیشتر بارش‌های فصول سرد در پاییز رخ داده است. اما در سال ۱۳۹۶ این الگو عوض شده و در آن سال ۴۴/۴۶٪ از کل بارش سالانه در زمستان رخ داده است. در روز ۲۸ سال که در واقع روز ۲۸ فروردین ماه می‌باشد، در طول دوره مطالعاتی هفتاد و یک ساله جمعاً ۲۵۹ mm و به تعداد ۳۸ بار بارش اتفاق افتاده است. به عبارتی دیگر شانس وقوع بارش در ۲۸ فروردین در ایستگاه تبریز حدود ۵۵ درصد می‌باشد. با توجه به نتایج به‌دست آمده مبنی بر پایدار بودن کم و بیش الگوی بارش ایستگاه تبریز در هفتاد و یک سال اخیر، به نظر می‌رسد کمبود منابع آب و مشکلات ناشی از آن بیشتر معطوف به مسائل مدیریتی و انسانی باشد تا مسائل طبیعی. در چنین شرایطی موارد متعددی به‌عنوان راه حل پیشنهاد می‌گردد. (۱) عدم توسعه اراضی کشاورزی، (۲) استفاده از سامانه‌های آبیاری کم‌فشار و یا تحت فشار، (۳) تغییر الگوی کشت به سمت محصولات کم‌آبر، (۴) گذر کامل از کشاورزی سنتی به سمت کشاورزی مدرن و صنعتی، (۵) توسعه صنعتی عمدتاً مرتبط با کشاورزی، (۶) استفاده از پتانسیل‌های گردشگری به‌عنوان درآمد جایگزین در بخش کشاورزی.

دما ترسیم شده تا بازخوانی نمودار بهتر شود. در همه حالت‌ها از چهار رنگ‌بندی جهت تفکیک فصول سال استفاده شده است که رنگ سبز، قرمز، زرد و آبی برای فصول سال به ترتیب بهار، تابستان، پاییز و زمستان مورد استفاده قرار گرفته است.

نتیجه‌گیری کلی

در این پژوهش سعی گردید با رویکردی نوین مبنی بر داده‌کاوی بصری به توصیف بارش‌های هفتاد و یک ساله اخیر تبریز پرداخته شود و الگوهای پنهان درون داده‌ها شناسایی گردد. در این راستا یافته‌های زیر قابل توجه است. نتایج حاکی از تغییر الگوی بارش از سال ۱۳۹۵ تا ۱۳۹۹ بود به طوری که در این سال‌ها روند بارش‌ها صعودی شده است. در مدت ۵ سال اخیر روند بارش‌ها به بیش از میانگین ۷۱ ساله افزایش پیدا کرده است ولی همچنان از دهه ۴۰ پایین‌تر است. از سویی دیگر تغییرات رفتار دمایی نشان از افزایش میانگین دمای سالانه در سال‌های اخیر دارد که اطلاعات دمای جهانی نیز این افزایش دمای نسبی را تأیید می‌کند. با توجه به بافت‌نگار میانگین دمای سالانه، در بیشتر سال‌ها میانگین دمای سالانه در ایستگاه بین ۱۲ تا ۱۳ درجه سانتی‌گراد بوده است. با توجه به بافت‌نگار دفعات بارش سالانه ایستگاه تبریز در اغلب سال‌ها بین ۱۰ الی ۲۰ رخداد بارش بالای ۵ mm اتفاق افتاده است. همچنین نتایج نشان داد که بیشترین بارش‌های سالانه به ترتیب مربوط به سال ۱۳۴۲ به میزان ۵۸۳ mm و سال ۱۳۴۷ به میزان ۱۴۰ mm و کمترین بارش در سال ۱۳۶۹ به میزان ۱۴۰ mm اتفاق افتاده است. در سال ۱۳۴۱ شدت بارش در فصل پاییز و در سال ۱۳۴۷ در تابستان بیشتر بوده است. براساس یافته‌های پژوهش، شدت بارش‌های بهاره تبریز در دوره ۱۳۸۵ تا ۱۳۹۹ کاهش محسوسی یافته است. این

منابع مورد استفاده

- Arthi Rani B, Manikandan N and Maragatham N, 2014. Trend analysis of rainfall and frequency of rainy days over Coimbatore. MAUSAM Journal 65 (3): 379-384.
- Asakereh H and Akbarzadeh Y, 2017. Simulation of temperature and precipitation changes of Tabriz Synoptic Station using statistical downscaling and canesm2 climate change model output. Journal of Geography and Environmental Hazards 21: 153-174. (In Persian with English abstract)
- Daniel A, Christian P and Mike S, 2003. Visual data mining of large spatial data sets. Lecture Notes Computer Science 2822: 201-215.
- Darand M and Ebrahimi B, 2019. Temporal-spatial analysis of changes in waiting time for rainfall in Kurdistan province. Journal of Water Resources Engineering 11: 17-30. (In Persian with English abstract)
- Ferdosi F and Dinpashoh Y, 2019. Investigation of temporal distribution of daily rainfall using the normalized rainfall curves (NRC) (case study: Tabriz Station). Journal of Water and Soil Science 2(29): 1-14. (In Persian with English abstract)
- Garcia JRM, Monteiro AMV and Santos RDC, 2012. Visual data mining for identification of patterns and outliers in weather stations' data. Pp. 245-252. International Conference on Intelligent Data Engineering and Automated Learning. 29-31 August, Natal, Brazil.
- Gentilucci M, Barbieri M, Lee L and Zardi D, 2019. Analysis of rainfall trends and extreme precipitation in the middle adriatic side, marche region (Central Italy). Water 11 (9). doi:10.3390/w11091948.
- Ghorbani MA, Mahmoud Alilou S, Javidan S and Raghavendra S, 2021. Assessment of spatio-temporal variability of rainfall and mean air temperature over Ardabil province, Iran. SN Applied Sciences 3 (728). <https://doi.org/10.1007/s42452-021-04698-y>.
- Han J and Pei J, 2012. Data Mining Trends and Research Frontiers. Data Mining (3rd Edition) Pp. 585-631. <https://doi.org/10.1016/B978-0-12-381479-1.00013-7>.
- Iranpour F and Zohreh Wandi H, 2019. Analysis of daily rainfall in Tabriz to study the probability of frequencies and the persistence of dry and wet days. Journal of Climate Research 36: 91-105. (In Persian with English abstract)
- Jahanbakhsh S, Abtahi V, Ghorbani MA, Tadayoni M and Valayi A, 2015. Temporal and spatial distribution of rainfall in Tabriz county using hierarchical cluster analysis. Quarterly Journal of Geographical Space 50 (15): 59-81. (In Persian with English abstract)
- Jana C, Sharma C, Alam N, Mishra PK, Dubey S and Kumar R, 2016. Trend analysis of rainfall and rainy days of Agra in Northern India. International Journal of Agricultural and Statistical Sciences 12 (1): 263-270.
- Keim DA, 2002. Information visualization and visual data mining. IEEE Transactions on Visualization and Computer Graphics 8 (1): 1-8. doi: 10.1109/2945.981847.
- Keim D and Ankerst M, 2005. Visual data mining techniques. Visualization Handbook. Pp. 831-843. <https://doi.org/10.1016/B978-012387582-2/50045-9>.
- Kopanakis I, Pelekis N, Karanikas H and Mavroudkis T, 2005. Visual techniques for the interpretation of data mining outcomes. Pp. 25-35. Panhellenic Conference on Informatics. 11-13 November, Volas, Greece.
- Ostertagova E, 2012. Modelling using polynomial regression. Procedia Engineering 48: 500-506.
- Pawar P, Rawat U, Yadav A, Rajput A, Vasht D and Nema S, 2020. Long term trend analysis of rainfall, rainy days and drought for Sindh River Basin, Madhya Pradesh, India. International Journal of Current Microbiology and Applied Sciences 9 (12): 2738-2749.
- Simoff SJ, Bohlen MH and Mazeika A, 2008. Visual Data Mining: An Introduction and Overview. Visual Data Mining. Lecture Notes in Computer Science, vol 4404. Springer, Berlin, Heidelberg. 1-12.
- Vagh Y, 2012. The application of a visual data mining framework to determine soil, climate and land use relationships. Procedia Engineering 32: 299-306.
- Ward M, Peng W and Wang X, 2004. Hierarchical visual data mining for large-scale data. Computational Statistics 19: 147-158.
- Yacoub E and Tayfur G, 2020. Spatial and temporal of variation of meteorological drought and precipitation trend analysis over whole Mauritania. Journal of African Earth Sciences 163 (1): 1-12. <https://doi.org/10.1016/j.jafrearsci.2020.103761>.