استخراج قواعد انجمنی جهت تحلیل رفتار آبخوان شهرستان قزوین توسط الگوریتم Apriori

سیدحسن میرهاشمی[۱]، پرویز حقیقت جو[۲]، فرهاد میرزایی[۳]، مهدی پناهی[۴]

چکیده:

ضرورت وجود الگوئی برای برنامه‌ریزی مناسب در راستای مدیریت منابع آب به ویژه آب‌های زیرزمینی و با توجه به شرایط کم آبی و افت آبخوان‌ها از اهمیت ویژه‌ای برخوردار است. در این تحقیق با استفاده از قوانین انجمنی، وابستگی بین تغییرات عمق آبخوان در محدوده شبکه آبیاری شهرستان قزوین و عوامل مختلف انسانی و طبیعی تأثیرگذار بر آن از سال ۱۳۸۰ تا ۱۳۹۴ توسط الگوریتم Apriori بررسی شد. جهت بررسی دقت قوانین از سه شاخص بهبود و درصد پشتیبان و اطمینان استفاده شد .نتایج نشان می‌دهند که، بیشترین مقدار همبستگی مثبت بین دو مجموعه مقادیر پیشین (شامل درصد زیاد رطوبت هوا، حجم کم آب در شبکه آبیاری و برداشتی از چاه کشاورزی) و پی‌آیند (بالاآمدگی عمق آبخوان بیشتر از یک متر) بوده است. با مقدار پشتیبان ۴/۸ درصد می‌توان بیان کرد که از مجموعه عوامل تأثیرگذار بر تغییرات عمق آبخوان، مقادیر حجم کم بارندگی، حجم زیاد آب در شبکه آبیاری و حجم زیاد نیاز آبی گیاهان با افت بین ۰/۳ تا ۱ متر به طور همزمان رخ داده است. در مواردی که مقادیر دمای هوا زیاد، درصد رطوبت هوا و حجم بارندگی کم بوده، مقدار افت بین ۰/۳ تا ۱ متر با مقدار اطمینان ۶۶ درصد شده است. همچنین با توجه به نتایج حاصل از استخراج قوانین انجمنی، با افزایش مقدار حجم آب برداشتی از چاه کشاورزی، دمای هوا، حجم آب در شبکه آبیاری و نیاز آبی گیاهان باعث افزایش افت عمق آبخوان و افزایش مقدار رطوبت هوا و بارندگی، باعث کاهش افت عمق آبخوان می‌شود.

واژه های کلیدی: بارندگی، رطوبت هوا، داده‌کاوی، شبکه آبیاری، مدیریت منابع آب.

[۱]-دانشجوی دکتری آبیاری و زهکشی، دانشگاه زابل، hassan.mirhashemi@yahoo.com
۲-استادیار گروه آب دانشکده آب و خاک، دانشگاه زابل ۰۹۱۵۵۴۲۵۱۳۶ (نویسنده مسئول)، phjou40@gmail.com
[۳]- دانشیار مهندسی گروه آبیاری و آبادانی پردیس کشاورزی و منابع طبیعی دانشگاه تهران ۰۹۱۲۸۶۱۳۸۲۵، fmirzaei@ut.ac.ir
[۴]- استادیار گروه مهندسی آب دانشکده کشاورزی، دانشگاه زنجان، ۰۹۱۲۶۰۷۰۳۲۳، panahi40@yahoo.com

مقدمه

با توسعه و پیشرفت کشاورزی و همچنین افزایش جمعیت که نیاز روز افزون به منابع آب را نیز به دنبال داشت، ناپایداری‌هایی را در مدیریت سنتی منابع آب ایجاد نمود. بخش عمده‌ای از عدم تعادل در منابع آب، ناشی از محدودیت طبیعی منابع آب و چرخه هیدرولوژی بوده و بخش دیگر تصمیمات و فعالیت‌های انسانی مؤثر بر روی منابع آب است که زمینه‌ساز مشکلات زیادی در امر بهره‌برداری از منابع آب زیرزمینی گشته است. همچنین مدیران در این بخش با روابط پیچیده و خصوصیات بسیار متنوع از حجم بسیار زیادی از داده‌های جمع‌آوری شده مواجه هستند که تجزیه و تحلیل و مدیریت آن‌ها به وسیله روش‌های تجربی و آماری، امری دشوار و در بسیاری از حوضه‌ها عملاً ناممکن می‌باشد.

نتایج تحقیقات نشان می‌دهد که برداشت بی‌رویه از سفره آب زیرزمینی دشت سیرجان، پیامدهایی همچون تغییر کیفیت آب زیرزمینی، افزایش مصرف انرژی برای استحصال آب زیرزمینی، افزایش آسیب‌پذیری دشت نسبت به خشکسالی، نشست زمین، از بین رفتن اکوسیستم منطقه و خشک شدن باغات و غیره را نیز به دنبال داشته است (عباس‌نژاد و شاهی‌دشت، ۱۳۹۲). سالانه بیش از ۱۹۹ میلیون متر مکعب اضافه برداشت از منابع آب زیرزمینی این دشت صورت می‌گیرد (باریکانی و همکاران،. ۱۳۹۰). با توجه به اینکه منابع آب‌های زیرزمینی دشت قزوین یکی از اصلی‌ترین منابع آبی مورد استفاده در بخش کشاورزی استان قزوین است، بنابراین برای نگهداری و برداشت بهینه از آب-های زیرزمینی مورد استفاده در بخش کشاورزی باید برنامه‌ریزی مناسب‌تری در بخش آب مصرفی کشاورزی انجام گیرد که برای تحقق این امر علاوه بر پیش‌بینی رفتار آبخوان با توجه تصمیمات عرضه و تقاضا، باید مقدار تأثیر عوامل مختلف را بر نوسانات عمق آبخوان بررسی کرد. ایران از جمله کشورهایی است که به دلیل کمبود منابع آب سطحی، بیشترین آب مصرفی در کشاورزی را از آب‌های زیرزمینی تامین می‌کند. بنابراین کمبود منابع آب های زیرزمینی و تراز رو به کاهش و منفی آن نیز یکی از بحران‌های حال حاضر کشور محسوب می‌گردد (عباس‌نژاد و شاهی‌دشت، ۱۳۹۲). در سال‌های گذشته نیز تحقیقات مختلفی در ارتباط با بررسی تغییرات مکانی و تخمین سطح آب-های زیرزمینی توسط محققین صورت گرفته است (jang et al.,2013).

امروزه داده‌کاوی در بسیاری از زمینه‌ها مانند بازاریابی، تشخیص تقلب، کشف علم و پزشکی استفاده می‌شود. قوانین انجمنی[1] یکی از تکنیکهای اصلی داده‌کاوی است و تقریباً مهمترین شکل از کشف و استخراج الگوها در سیستم‌های یادگیری غیر هدایت شده می-باشد. این روش تمام الگوهای جالب و تکرارپذیر در پایگاه داده‌ها را بازیابی می‌کند(Ramageri, 2011). هدف از کاوش قوانین انجمنی، یافتن ارتباطات بین اجزاء یک مجموعه می‌باشد. به این ترتیب جستجویِ یافتن وابستگی‌ها، همبستگی‌ها و ساختارهای علّی و معلولی موجود بین اعضاء و اشیای بانک‌های اطلاعاتی، تبادلات و معاملات و سایر انبارهای اطلاعاتی در این مقوله جای می‌گیرد(Tan et al., 2006). فتاحی و همکاران در سال ۱۳۹۱ با استفاده از قوانین انجمنی، ارتباط بارندگی پاییزه و زمستانه را بـا دمای سطح آب-های آزاد بدست آوردند. آن‌ها در تحقیق خود داده‌های سه ایستگاه سینوپتیک آبادان، اهواز و دزفول واقع در استان خوزستان را در طی سال‌های ۱۹۶۶-۲۰۱۰ مورد بررسی قرار دادند. در این تحقیق ابتدا شاخص دمای سطح آب (SST[2]) در خلیج فارس (SSTp) و دریای سرخ (SSTR) و شاخص بارندگی(SPI۳) به دست آوردند و سپس برای تولید قوانین انجمنی، شاخص‌ها بـه هفت دسته تقسیم‌بندی شدند. شاخص SPI به عنوان نماینده وقایع خشکسالی و ترسالی مورد نظر قرار گرفته شد. بعد از گسسته‌سازی داده‌ها، شاخصSPI به عنوان جملـه پـی‌آینـد[4] قوانین و SST خلیج فارس و دریای سرخ به عنوان سری پیشین[5] مورد استفاده قرار گرفتند.

[1] Association Rules
[2] Sea Surface Temperature
[3] Standardized Precipitation Index
[4] Consequent
[5] Antecedent

پارامترهای اقیانوسی، اقلیمی، و هواشناسی را برای کشف قواعدی که بتوانند رخداد خشکسالی را تشخیص دهند به کار گرفت و به این نتیجه رسید که بیشتر رخدادهای خشکسالی با مقادیر مثبت نمایهٔ نوسانات جنوبی، مقادیر منفی نمایهٔ ENSO چند متغیره و مقادیر منفی نوسانات دهکی آرام مرتبط هستند (Tadesse, 2002). از داده‌کاوی برای مدیریت خشکسالی استفاده کردند و در کاوش قواعد انجمنی متوالی با قیود تأخیر زمانی ارتباطاتی بین حوادث ضمنی اقلیمی در مقیاس جهانی و شرایط محلی خشکسالی یافتند (Harms and Deogun., 2004). برای ارتباط خشکسالی‌ها و سیل‌ها در هندوستان با داده‌های اقلیمی از قواعد انجمنی استفاده کردند و روابط مستحکمی بین نمایه‌های اقلیمی منتخب با رویدادهای بارش حدی به دست آوردند (Dhanya and Nagesh Kumar., 2009). قوانین انجمنی را در مطالعهٔ تطبیقی نمایه‌های هواشناسی SPIو SIAP به کار بردند و وقوع همزمانی طبقات خشکسالی این دو نمایه را بررسی کردند (قربانی و همکاران، ۱۳۸۸). نیکزاد و همکاران در سال ۱۳۹۲، از روش داده‌کاوی قوانین انجمنی برای آشکارسازی وابستگی‌های بین پارامترهای اقیانوسی-جوی و اقلیمی برای پایش خشکسالی استان خوزستان استفاده کردند و نشان دادند که وقایع خشکسالی بیشتر با وضعیت نرمال شاخص‌ها همراه می‌باشد. رحیمی‌خوب در سال ۱۳۸۹، با استفاده از قوانین انجمنی که یکی از تکنیک‌های داده‌کاوی می‌باشد، ارتباط بین دمای سطح گستره‌های آبی خلیج فارس و دریای سرخ، با حداکثر بارش ماهانه ایستگاه ایلام را بررسی کرد و نشان داد SST دو گستره آبی فوق قابل پیش‌بینی است. تحلیل وابستگی‌ها یک حالت غیر نظارتی داده‌کاوی می‌باشد که به جستجو برای یافتن ارتباط در مجموعه داده‌ها می‌پردازد. یکی از کاربردی‌ترین حالات تحلیل وابستگی (تحلیل سبد بازار) می‌باشد و در آن هدف یافتن کالاهایی است که معمولاً به طور همزمان خریداری می‌شوند (Tan. et al., 2006). از قواعد انجمنی برای رصد وقایع بارندگی و خشکسالی با استفاده از دمای سطح آب خلیج فارس و دریای سرخ استفاده کردند. نتایج نشان داد با اینکه الگوریتم تهیه شده قادر است شرایط خشکی متوسط تا رطوبت شدید را پیش‌بینی کند، شرایط خنثی نیز با احتمال بسیار زیاد و اطمینان بالا قابل پیش‌بینی هستند (فتاحی و همکاران، ۱۳۹۱).

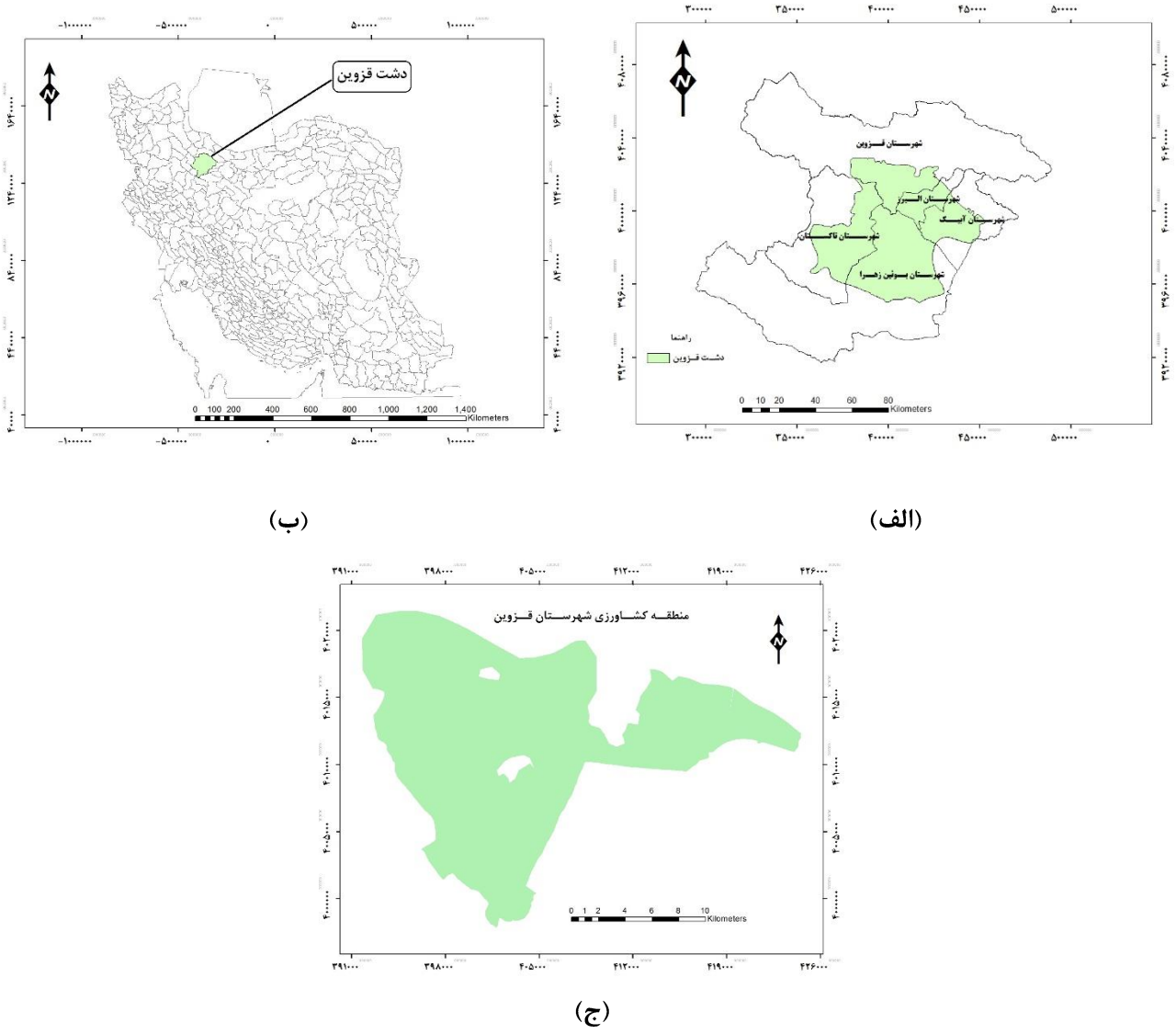
با اینکه داده‌کاوی در مطالعات مختلف در زمینه‌های هواشناسی و خشکسالی به کار رفته، در مطالعات اندکی جهت استخراج قوانین انجمنی بین روابط تغییرات عمق آبخوان و عوامل مختلف تأثیرگذار بر آن استفاده شده است. بر این اساس، هدف این تحقیق به استخراج قوانین انجمنی بین روابط تغییرات عمق آبخوان و عوامل مختلف تأثیرگذار بر آن در محدوده شبکه آبیاری منطقه کشاورزی دشت شهرستان قزوین می‌باشد.

مواد و روش ها

منطقه مطالعاتی

دشت قزوین با مساحتی حدود ۴۵۰۰۰۰ هکتار در محدوده‌ی طول جغرافیایی ۴۹ درجه و ۲۵ دقیقه تا ۵۰ درجه و ۳۵ دقیقه شرقی و عرض‌های جغرافیایی ۳۵ درجه و ۲۵ دقیقه تا ۳۶ درجه و ۲۵ دقیقه شمالی قرار دارد. این دشت از جلگه آبرفتی پهناوری تشکیل شده که از رسوبات جریان‌های سطحی کوه‌های اطراف تشکیل شده است (محمدی و همکاران، ۱۳۹۰). مجموع تغذیه آبخوان دشت قزوین ۱۲۵۹/۴۶ میلیون مترمکعب و مجموع عوامل تخلیه آبخوان دشت قزوین ۱۴۵۸/۶۶ میلیون مترمکعب است. بر این اساس سهم تخلیه بخش کشاورزی حدود ۱۳۵۲/۹۲ میلیون مترمکعب می‌باشد که حدود ۸۵۷/۳ میلیون هکتار آن سهم بخش زراعی است. با توجه به محدودیت منابع آب سطحی و فصلی بودن این منابع بخش عمده آب آبیاری از منابع زیرزمینی استحصال می‌شود. در وضعیت موجود، برداشت باعث افت سالیانه ۱/۵ متر سطح سفره‌های زیرزمینی و فرو نشست زمین در این منطقه تا حدود ۲۵ سانتیمتر در سال شده است. با توجه به اهمیت دشت قزوین به عنوان یک منطقه مستعد کشاورزی از یک سو و وجود معضل افت شدید عمق

سطح ایستابی در این منطقه از سوی دیگر، توجه به پایداری منابع آب زیرزمینی در تولید محصولات کشاورزی و انتخاب الگوی کشت این منطقه ضروری به نظر می‌رسد(باریکانی و همکاران، ۱۳۹۰). استان قزوین دارای شش شهرستان آبیک، آوج، البرز، بوئین‌زهرا، تاکستان و قزوین می‌باشد. از این شش شهرستان بخش‌هایی از شهرستان‌های آبیک، البرز، بوئین‌زهرا، تاکستان و قزوین داخل دشت قزوین هستند و با توجه به متفاوت بودن رفتار آبخوان دشت قزوین، در بخش-های مختلف دشت، در این مطالعه بخشی از آبخوان دشت قزوین که در محدوده منطقه کشاورزی شهرستان قزوین می‌باشد، مورد مطالعه قرار گرفت (شکل۱).



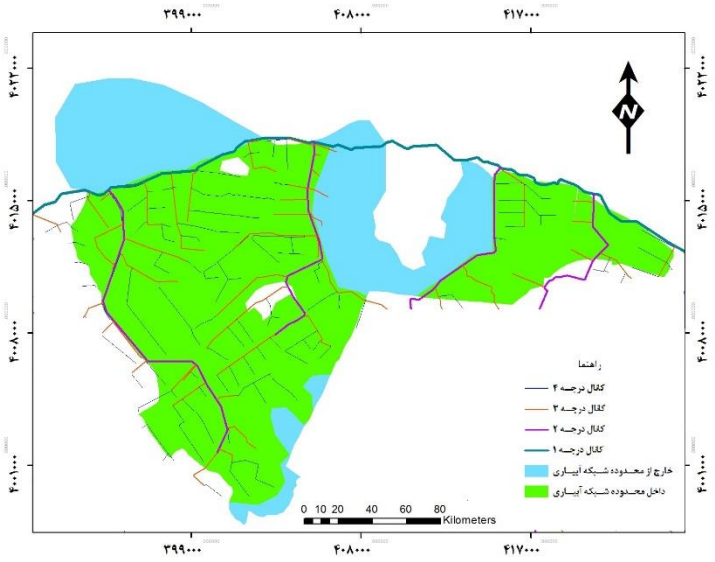
(ب) (الف)

(ج)

شکل ۱- الف-موقعیت دشت قزوین در بین دشت‌های ایران ب-موقعیت شهرستان‌های قزوین در دشت قزوین ج-منطقه کشاورزی شهرستان قزوین

شبکه آبیاری دشت قزوین دارای یک کانال به طول ۹۴ کیلومتر بوده که روی خط تراز ۱۳۰۰ متر احداث شده است و تقریبا به موازات اتوبان تهران-زنجان قرار دارد. ظرفیت کانال اصلی از ۱ تا ۳۰ مترکعب در ثانیه متغییر است. (ستوده‌نیا و ستوده‌نیا، ۱۳۹۵).

در شکل(۲) محل قرارگیری شبکه آبیاری در منطقه کشاورزی شهرستان قزوین را نشان می‌دهد. مناطق آبی رنگ داخل شکل (۲) خارج از محدوده شبکه آبیاری که در حدود ۹۶۰۴/۲۵ هکتار و مناطقی که با رنگ سبز مشخص شده داخل محدوده شبکه آبیاری که در حدود ۲۲۰۷۹/۷۰ هکتار می‌باشد را مشخص می‌کند.



شکل ۲- موقعیت شبکه آبیاری در منطقه کشاورزی شهرستان قزوین

نرم‌افزار داده‌کاوی IMB SPSS Modeler

IBM SPSS نرم‌افزاری است که برای تحلیل‌های آماری به کار می‌رود. IBM SPSS Modeler نرم افزاری از شرکت SPSS است که در ابتدا با نام کلمنتاین (Clementine 12) ارائه می شد که در نسخه جدید به SPSS Modeler تغییر نام پیدا کرد. این نرم افزار یکی از بهترین ابزارهای داده‌کاوی است و برنامه‌ای حرفه ای برای انجام محاسبات پیچیده و آنالیزهای آماری به صورت سرور و کلاینت می باشد. نرم افزار SPSS IBM Modeler که از بهترین ابزارهای داده کاوی است دارای مزیتهای نسبت به نرم افزار Clementine 12 است: بهینه کردن تکنیک ها برای مجموعه داده‌های بزرگ، افزایش مصورسازی الگورتیم های کلیدی مانند شبکه های عصبی و درخت تصمیم، محیط مصورسازی جدید برای الگوریتم‌های کلیدی و مدل های ترکیبی در ساخت نتایج برای فهم بهتر، بروزرسانی الگوریتم‌های شبکه عصبی و خطی، توانایی در خواندن و نوشتن داده از xml و بهینه کردن پردازش پایگاه داده بزرگ برای الگوریتم‌های کلیدی را اشاره کرد.

الگوریتم قوانین انجمنی

داده‌کاوی انجمنی و استخراج قوانین انجمنی از مجموعه داده‌ها از جنس قوانین احتمالی هستند که اولین بار توسط Agrawal و همکاران در سال ۱۹۹۴ ارائه شد. در واقع برای مشخص شدن وابستگی‌های مهم و پنهان رکوردهای یک پایگاه داده از قوانین انجمنی استفاده می‌شود. این قوانین وابستگی اتفاق و وقوع یک شی را بر اساس وقوع سایر اشیا پیش‌بینی می‌کند. اغلب مشاهده می‌شود که وابستگی نزدیک بین مجموعه‌ای از داده‌های معین وجود دارد. بنابراین الگوریتم‌های یافتن قوانین وابستگی، تمامی قواعد

انجمنی ممکن را درون پایگاه داده پیدا می‌کنند. خروجی مهم در این روش، عبارت است از مجموعه‌ای از قوانین اگر آن‌گاه که بیانگر ارتباطات میان رخداد توامان مجموعه‌ای از اشیا با یکدیگر هستند. این الگوریتم مجموعه قوانین به دست آمده از داده‌ها را

معرفی اصطلاحات پایه مربوط به قوانین انجمنی

ضریب پشتیبان[1] : احتمال شمول مقدم (پیشین) و تالی (پی‌آیند) در یک قانون می‌باشد. به عبارت دیگر میزان پشتیبانی نسبت وقایع هم زمان A و B به کل وقایع موجود در مجموعه را نشان می دهد که از تقسیم تعداد وقایعی که شامل A و B می‌باشند بر تعداد کل وقایع به دست می آید به عبارت دیگر این پارامتر احتمال وجود همزمان A و B را در قانون A→B بیان می‌کند.

ضریب اطمینان[2] : میزان صحت یک قاعده را مشخص می‌کند. این ضریب نشان می‌دهد اگر تبادل، شرایط مقدم را اختیار کند آنگاه با چه احتمالی شرایط تالی را نیز کسب می‌کند. به عبارت دیگر نسبت تعداد وقایع شامل A و B به تعداد وقایع شامل A می‌باشد. این پارامتر احتمال شرطی است برای آنکه تراکنش A دارای B نیز باشد.

بهبود[3]: معیاری برای ارزیابی مقدار همبستگی دو مجموعه داده (مقادیر پیشین و پی‌آیند) می‌باشد. بطوریکه اگر مقدار بهبود کمتر از یک باشد دو مجموعه دارای همبستگی منفی و اگر بیشتر از یک باشد دو مجموعه داده دارای همبستگی مثبت و اگر برابر یک باشند دو مجموعه بدون همبستگی و مستقل هستند.

سه پارامتر اخیر پشتیبان، اطمینان و بهبود جهت ارزیابی قانون‌های تولید شده استفاده می‌شود. در رابطه A→Bمقدار پشتیبان، اطمینان و بهبود به ترتیب از سه رابطه زیر محاسبه می‌گردند:

$$\text{support}=\frac{n(A\cup B)}{N} \qquad (۱)$$

$$Confidence(A\rightarrow B)=\frac{\text{Support}(A\cup B)}{\text{Support}(A)}=\frac{n(A\cup B)}{n(A)} \qquad (۲)$$

$$lift=\frac{p(\frac{y}{x})}{p(y)}=\frac{p(x\cap y)}{p(x).p(y)}\ lift=\frac{Confidence(X\Rightarrow Y)}{Support(Y)} \qquad (۳)$$

n بیانگر تعداد رکوردهای دارای شرایط و N تعداد کل رکوردها می‌باشند.

برای به دست آوردن قوانین انجمنی، الگوریتم Aprioriروی دو مجموعه داده پیشین (مقدم) و پی-آیند (تالی) پیاده سازی خواهند شد. هدف یافتن روابط بین متغیرهای ورودی (پیشین) و متغیر هدف (پی‌آیند) خواهد بود. متغییرهای ورودی (مقادیر پیشین) شامل مقادیر حجم آب برداشتی از چاه‌های کشاورزی (میلیون مترمکعب)، حجم آب در شبکه آبیاری (میلیون مترمکعب)، حجم نیاز آبی گیاهان (میلیون مترمکعب)، مقادیر حجم بارندگی (میلیون مترمکعب)، دما (سانتی-گراد) و رطوبت هوا (درصد) انتخاب شدند. همچنین از داده‌های مربوط به مقدار تغییرات عمق آبخوان (متر) به عنوان متغییر هدف (مقدار پی‌آیند) در چندین حالت مختلف بر اساس دسته‌های مختلف تقسیم‌بندی شدند. مجموعه داده‌های مورد استفاده در این تحقیق از اطلاعات مربوط به منطقه کشاورزی دشت شهرستان قزوین طی دوره ۱۵ ساله بین سال‌های ۱۳۸۰ تا ۱۳۹۴ به صورت ماهانه در محدوده شبکه آبیاری می‌باشد.

شهرستان قزوین دارای ۲۴ پیزومتر اندازه‌گیری عمق آبخوان در محدوده دشت قزوین می‌باشد. با توجه به پراکندگی پیزومترهای اندازه‌گیری عمق آب زیرزمینی در قسمت‌های مختلف آبخوان و همچنین گسترش می‌دهد و قوانین با بالاترین سطح اطلاعات را بیرون می‌آورد. یکی از الگوریتم های تکامل یافته کشف قوانین انجمنی الگوریتم Apriori می‌باشد که در این تحقیق استفاده شده است. این الگوریتم در سال ۱۹۹۶ توسط Cheung و همکارانش ابداع شد.

[1] Support
[2] Confidence
[3] Lift

برای تشخیص تقریبی میزان بهره‌مندی و تأثیر هر پیزومتر از پارامترهای مختلف همچون بارندگی، عرضه و مصرف آب از روش تیسن برای تعیین محدوده مساحت تأثیر هر پیزومتر محاسبه گردید. برای تحقق این امر از نرم افزار ArcGIS 10 استفاده گردید. در این خصوص با استفاده از نرم‌افزار GIS ابتدا لایه‌های اطلاعاتی پیزومتر و مرزهای مناطق کشاورزی شهرستان قزوین در محدوده دشت قزوین تهیه گردید.

از ۱۷۴حلقه چاه مجاز مورد استفاده در منطقه کشاورزی شهرستان قزوین، ۱۳۸ حلقه چاه از نوع بهره‌برداری کشاورزی، تعداد ۲۳ حلقه چاه از نوع بهره-برداری تلفیقی و تعداد ۱۳ حلقه چاه از نوع بهره‌بردای چندمنظوره هستند. از ۱۷۴حلقه چاه موجود، تعداد ۴۰ حلقه چاه در خارج از محدوده شبکه آبیاری و تعداد ۱۳۴ حلقه چاه در محدوده شبکه آبیاری قرار دارند. میانگین عمق چاه‌ها ۱۲۳ متر و بیشترین عمق چاه‌ها ۲۰۰ متر و همچنین متوسط آبدهی چاه‌ها ۳۶ لیتر در ثانیه و حداکثر آبدهی چاه‌ها ۹۰ لیتر در ثانیه می‌باشد.

برای تعیین حجم نیاز آبی گیاهان، مقدار نیاز خالص آبی هر محصول در مقدار سطح زیر کشت همان محصول ضرب شد. مقادیر بارندگی بر حسب میلی‌متر از شرکت آب منطقه‌ای استان قزوین، به صورت ماهانه تهیه گردید. با توجه به اینکه میزان اثر باران برای هر پیزومتر متفاوت است، لذا میزان مقدار حجم باران مربوط به وزن تأثیر هر پیزومتر، بر حسب میلیون متر مکعب محاسبه شد.

با توجه به اینکه مقادیر ورودی به الگوریتم Apriori باید به صورت دسته‌بندی باشد به همین دلیل مقادیر بارندگی، نیاز آبی گیاهان، حجم آب برداشتی از چاه کشاورزی، حجم آب در شبکه آبیاری، دمای هوا، درصد رطوبت هوا و همچنین مقدار نوسانات عمق آبخوان به صورت جدول‌های ۱ تا ۴ در ماه‌های زراعی اسفند تا مهر دسته‌بندی شدند. با توجه به اینکه مقادیر داده‌های ورودی در ماه‌های مختلف با هم متفاوت می-باشند، بدین منظور برای هر هشت ماه زراعی، داده‌های مختلف در سه دسته کم، متوسط و زیاد تقسیم‌بندی شدند.

جدول ۱- دسته‌بندی مقادیر حجم بارندگی و حجم آب در شبکه آبیاری

نام ماه	حجم بارندگی (میلیون متر مکعب)			حجم آب در شبکه آبیاری(میلیون مترمکعب)		
	کم	متوسط	زیاد	کم	متوسط	زیاد
اسفند	<۰/۱۲	۰/۱۲-۰/۴	>۰/۴	-	-	-
فروردین	<۰/۳	۰/۳-۰/۸	>۰/۸	<۰/۲	۰/۲-۰/۴	>۰/۴
اردیبهشت	<۰/۱۶	۰/۱۶-۰/۵۲	>۰/۵۲	<۱	۱-۱/۲	>۱/۲
خرداد	<۰/۰۲۵	۰/۰۲۵-۰/۱۳	>۰/۱۳	<۰/۸	۰/۸-۱/۲	>۱/۲
تیر	<۰/۰۰۶	۰/۰۰۶-۰/۰۵۱	>۰/۰۵۴	<۰/۵	۰/۵-۰/۸	>۰/۸
مرداد	<۰/۰۰۲	۰/۰۰۰۲-۰/۰۱۴	>۰/۰۱۴	<۰/۶	۰/۶-۱	>۱
شهریور	<۰/۰۰۲	۰/۰۰۲-۰/۰۳۵	>۰/۰۳۵	<۰/۵	۰/۵-۰/۸	>۰/۸
مهر	<۰/۰۱	۰/۰۱-۰/۰۷	>۰/۰۷	<۰/۱	۰/۱-۰/۲	>۰/۲

جدول ۲- دسته‌بندی حجم آب خروجی از چاه کشاورزی و حجم نیاز آبی گیاهان

نام ماه	حجم آب برداشتی از چاه کشاورزی(میلیون مترمکعب)			حجم نیاز آبی گیاهان (میلیون مترمکعب)		
	کم	متوسط	زیاد	کم	متوسط	زیاد
اسفند	<۰/۰۲۷	۰/۰۲۷-۰/۱	>۰/۱	<۰/۰۶	۰/۰۶-۰/۱۴	>۰/۱۴
فروردین	<۰/۱	۰/۱-۰/۳۶	>۰/۳۶	<۰/۲۱	۰/۲۱-۰/۵	>۰/۵
اردیبهشت	<۰/۱۵	۰/۱۵-۰/۵۸	>۰/۵۸	<۰/۳۴	۰/۳۴-۰/۷۷	>۰/۷۷
خرداد	<۰/۲۶	۰/۲۶-۰/۹۷	>۰/۹۷	<۰/۵۵	۰/۵۵-۱/۲۵	>۱/۲۵
تیر	<۰/۲۶	۰/۲۶-۰/۹۷	>۰/۹۷	<۰/۵۸	۰/۵۸-۱/۲۶	>۱/۲۶
مرداد	<۰/۲	۰/۲-۰/۷۶	>۰/۷۶	<۰/۴۷	۰/۴۷-۱	>۱
شهریور	<۰/۰۷۵	۰/۰۷۵-۰/۳	>۰/۳	<۰/۱۸	۰/۱۸-۰/۴	>۰/۴
مهر	<۰/۰۳	۰/۰۳-۰/۱۳	>۰/۱۳	<۰/۰۷	۰/۰۷-۰/۱۶	>۰/۱۶

جدول ۳- دسته بندی مقدار دما و رطوبت نسبی در ماه‌های زراعی

نام ماه	دما (سانتی‌گرد)			رطوبت نسبی هوا (درصد)		
	کم	متوسط	زیاد	کم	متوسط	زیاد
اسفند	<۸/۵	۸/۵-۹/۸	>۱۴/۵	<۵۵	۵۵-۶۱	>۶۱
فروردین	<۱۳	۱۳-۱۴/۴	>۱۴/۴	<۵۰	۵۰-۵۷	>۵۷
اردیبهشت	<۱۸	۱۸-۱۹	>۱۹	<۴۵	۴۵-۵۷	>۵۷
خرداد	<۲۳/۴	۲۳/۴-۲۴/۴	>۲۴/۴	<۳۹	۳۹-۴۳	>۴۳
تیر	<۲۶/۲	۲۶/۲-۲۷	>۲۷	<۴۲	۴۲-۴۴	>۴۴
مرداد	<۲۵/۶	۲۵/۶-۲۶/۳	>۲۶/۳	<۳۷	۳۷-۴۱	>۴۱
شهریور	<۲۱/۴	۲۱/۴-۲۲/۴	>۲۲/۴	<۳۹	۳۹-۴۳	>۴۳
مهر	<۱۵/۶	۱۵/۶-۱۶/۲	>۱۶/۲	<۴۴	۴۴-۵۵	>۵۵

برای افزایش قابلیت استفاده از نتایج و اجرای الگوریتم، مقادیر تغییرات عمق آب‌های زیرزمینی در دسته‌های مختلف تقسیم‌بندی شدند. مقادیر منفی در جدول (۴) نشان دهنده مقدار افت می‌باشد.

جدول ۴- تعیین محدوده داده‌های تغییرات عمق آبخوان

محدوده داده‌ها	میانگین داده‌ها	تغییرات عمق آبخوان
-۱≥...	-۱/۸۱	افت زیاد
-۰/۳≥...>-۱	-۰/۵۲	افت متوسط
۰ >...>-۰/۳	-۰/۱۶	افت کم
۰/۳≥...>۰	۰/۱۵	بالاآمدگی کم
۱≥...>۰/۳	۰/۵۰	بالاآمدگی متوسط
...> ۱	۱/۷۷	بالاآمدگی زیاد

جدول (۴) شامل خلاصه‌ای از مشخصات آماری به صورت ماهانه مربوط به داده‌های آبخوان منطقه کشاورزی شهرستان قزوین در محدوده شبکه آبیاری برای دوره ۱۵ ساله از سال ۱۳۸۰ تا ۱۳۹۴ می‌باشد.

جدول ۵- مشخصات آماری مربوط به عمق آبخوان در داخل محدوده مورد مطالعه

محدوده مطالعاتی	میانگین کل(m)	میانگین مقادیر افت(m)	میانگین مقادیر بالا آمدگی(m)	حداکثر افت(m)	حداکثر بالاآمدگی(m)
محدوده شبکه آبیاری	-۰/۳۳	-۰/۵۹	۰/۴۰	-۴/۶۱	۳/۳
تعداد داده‌ها	۱۵۱۲	۱۱۱۲	۴۰۰		

نتایج

اثرات ناشی از افت آب زیرزمینی باعث بروز مشکلات روز افزون بر جوامع انسانی و از جمله کشاورزی می‌شود. برای بررسی عوامل تأثیرگذار بر نوسانات عمق آبخوان از الگوریتم Apriori در این تحقیق استفاده شد. از جمله عوامل تاثیر گذار بر تغییرات عمق آبخوان که در این تحقیق استفاده شد شامل مقادیر حجم آب برداشتی از چاه کشاورزی (میلیون مترمکعب)، حجم نیاز آبی گیاهان (میلیون - مترمکعب)، حجم بارندگی (میلیون مترمکعب)، حجم آب در شبکه آبیاری (میلیون مترمکعب)، دمای هوا و درصد رطوبت هوا به عنوان مقادیر پیشین و همچنین نوسانات عمق آبخوان (متر) به عنوان مقدار پی‌آیند مورد استفاده قرار گرفته شدند. با توجه به اینکه داده-های ورودی به الگوریتم Apriori باید گسسته باشد با استفاده از اطلاعات جدول‌های ۱، ۲، ۳ و ۴ برای اجرای الگوریتم استفاده شد. سپس با استفاده از سه شاخص مقدار پشتیبان، اطمینان و بهبود مورد ارزیابی قرار گرفتند. جدول (۶) اطلاعات آماری قوانین به دست آمده از اجرای الگوریتم Apriori روی مجموعه داده-های تأثیرگذار بر تغییرات عمق آبخوان را نشان می-دهد.

جدول۶- قوانین به دست آمده از الگوریتم Apriori

شماره قانون	پیشین	پی‌آیند	درصد پشتیبان	درصد اطمینان	بهبود
۱	حجم آب در شبکه آبیاری= کم حجم بارندگی= کم حجم آب برداشتی از چاه کشاورزی= کم دمای هوا = زیاد	افت زیاد	۲/۲	۳۹/۲۸	۲/۶۶
۲	حجم نیاز آبی گیاهان =زیاد حجم آب برداشتی از چاه کشاورزی = زیاد درصد رطوبت هوا= کم	افت زیاد	۳/۱۵	۴۱/۳۷	۲/۸۱
۳	دمای هوا = زیاد درصد رطوبت هوا =کم حجم بارندگی=کم	افت متوسط	۳/۲	۶۶/۶۷	۱/۹۳
۴	حجم آب در شبکه آبیاری= زیاد حجم بارندگی=کم حجم نیاز آبی گیاهان=زیاد	افت متوسط	۴/۸۲	۶۲/۶۵	۲/۲
۵	حجم نیاز آبی گیاهان=متوسط حجم آب برداشتی از چاه کشاورزی= متوسط حجم بارندگی=متوسط دمای هوا =زیاد	افت کم	۳/۱۵	۵۸/۶۲	۲/۲۶
۶	حجم آب در شبکه آبیاری= متوسط درصد رطوبت هوا =کم حجم نیاز آبی گیاهان =کم حجم بارندگی=متوسط	افت کم	۴/۳	۵۴/۵۵	۲/۱
۷	حجم بارندگی= کم دمای هوا = متوسط حجم آب برداشتی از چاه کشاورزی= متوسط	بالاآمدگی کم	۳/۴۶	۳۷/۸۳	۳/۱
۸	حجم بارندگی= کم دمای هوا = کم درصد رطوبت هوا =متوسط	بالاآمدگی کم	۴/۲۳	۳۵/۴۸	۲/۹۱
۹	حجم بارندگی= متوسط دمای هوا = کم درصد رطوبت هوا =متوسط	بالاآمدگی متوسط	۳/۶	۳۵/۷۱	۳/۹۸
۱۰	حجم آب در شبکه آبیاری= متوسط حجم نیاز آبی گیاهان=کم دمای هوا = کم حجم بارندگی= متوسط	بالاآمدگی متوسط	۱/۱۵	۶/۸۹	۰/۷۶
۱۱	درصد رطوبت هوا = زیاد حجم آب در شبکه آبیاری=کم حجم آب برداشتی از چاه کشاورزی= کم	بالاآمدگی زیاد	۴/۷۸	۱۷/۷۸	۵/۶
۱۲	حجم نیاز آبی گیاهان =کم حجم آب برداشتی از چاه کشاورزی= کم درصد رطوبت هوا = زیاد دمای هوا = کم	بالاآمدگی زیاد	۳/۷۸	۱۱/۳۶	۳/۵۷

با توجه به جدول (۶) بیشترین مقدار اطمینان مربوط به قانون سوم و با مقدار اطمینان ۶۶/۶۷ درصد می‌باشد لذا می‌توان بیان کرد که ۶۶ درصد مواردی که مقادیر دمای هوا زیاد، درصد رطوبت هوا و حجم بارندگی کم بوده، مقدار افت متوسط (افت بین ۰/۳ تا ۱ متر) شده است. بیشترین مقدار پشتیبان مربوط به قانون چهارم و با مقدار پشتیبان ۴/۸۲ درصد می‌باشد لذا می‌توان بیان کرد که با مقدار پشتیبان ۴ درصد مقادیر حجم زیاد آب در شبکه آبیاری، حجم کم بارندگی و حجم زیاد نیاز آبی گیاهان با مقدار افت متوسط همزمان رخ داده است و همچنین بیشترین مقدار بهبود در قانون ۱۱ بین دو مجموعه مقادیر پیشین (درصد زیاد رطوبت هوا، حجم کم آب در شبکه آبیاری و حجم کم آب برداشتی از چاه کشاورزی) و پی‌آیند (بالاآمدگی بیشتر از یک متر) و دارای بیشترین همبستگی مثبت بوده است.

با توجه به جدول (۶) بیشترین مقدار افت در قانون اول و دوم می‌باشد. در قانون اول با وجود مقادیر حجم کم آب در شبکه آبیاری و حجم کم آب برداشتی از چاه کشاورزی ولی به دلیل دمای زیاد هوا و بارندگی کم باعث افت زیاد آبخوان با مقدار بهبود، درصد اطمینان و پشتیبان به ترتیب برابر ۲/۶۶، ۳۹/۲۸ و ۲/۲ شده است و همچنین در قانون دوم هر سه عامل پیشین حجم زیاد نیاز آبی گیاهان، حجم زیاد آب برداشتی از چاه کشاورزی و درصد کم رطوبت هوا باعث افت زیاد آبخوان با مقدار بهبود، درصد اطمینان و پشتیبان به ترتیب برابر ۲/۸۱، ۴۱/۳۷ و ۳/۱۵ شده است.

مقدار بهبود بیشتر از یک نشان دهنده کارامدی قانون و همبستگی مثبت بین دو مجوعه مقادیر پیشین و پی‌آیند می‌باشد. در تمامی قوانین جدول ۶ به جز قانون شماره ۱۰ مقدار بهبود بیشتر از یک می‌باشد. در قانون شماره ۱۰ مقدار بهبود برابر مقدار ۰/۷۶ می‌باشد. لذا قانون شماره ۱۰ قانون ناکارامدی می‌باشد و همبستگی منفی بین دو مجموعه مقادیر پیشین (حجم متوسط آب در شبکه آبیاری، حجم کم نیاز آبی گیاهان، دمای کم هوا و حجم متوسط بارندگی) و پی-آیند (بالاآمدگی بین ۰/۳ تا ۱ متر) وجود دارد.

بیشترین مقدار بالاآمدگی در قانون یازدهم و دوازدهم و در مقادیر پیشین هر دو قانون مقادیر درصد رطوبت هوا زیاد و حجم آب برداشتی از چاه کشاورزی کم می‌باشد. در قانون یازدهم در ۱۷ درصد مواردی که مقدار درصد رطوبت هوا زیاد، حجم آب در شبکه آبیاری و حجم آب برداشتی از چاه کشاورزی کم بوده، مقدار بالاآمدگی بیشتر از یک متر شده است. همچنین با مقدار پشتیبان ۴ درصد احتمال رخ دادن مقادیر پیشین و پی‌آیند قانون یازدهم همزمان بوده است. مقدار بهبود نیز برابر ۵/۶ می‌باشد که نشان دهنده کارامدی این قانون و همبستگی مثبت دو مجموعه مقادیر پیشین و پی‌آیند قانون یازدهم می‌باشد. در قانون دوازدهم در ۱۱ درصد مواردی که حجم نیاز آبی گیاهان، حجم آب برداشتی از چاه کشاورزی و دمای هوا کم و درصد رطوبت هوا زیاد بوده، مقدار بالاآمدگی بیشتر از یک متر شده است و با مقدار پشتیبان ۳ درصد مقادیر پیشین و پی‌آیند همزمان رخ داده است. مقدار بهبود نیز برابر ۳/۵۷ که از مقدار یک بیشتر می-باشد، لذا همبستگی مثبت بین دو مجموعه مقادیر پیشین و پی‌آیند قانون دوازدهم وجود دارد.

در مقادیر پیشین ۷۵، ۶۶ و ۵۸ درصد کل قوانین جدول (۶)، به ترتیب شامل مقدار حجم بارندگی، دمای هوا و درصد رطوبت هوا بوده است. لذا می‌توان بیان کرد مقدار بارندگی بیشترین تأثیر و بعد از آن به ترتیب مقادیر دما و رطوبت هوا بیشترین تأثیر در تغییرات عمق آبخوان داشته‌اند.

منابع

باریکانی، الف.، م. احمدیان و ص. خلیلیان. ۱۳۹۰. بهره‌برداری بهینه پایدار از منابع آب زیرزمینی در بخش کشاورزی: مطالعه موردی زیربخش زراعت دشت قزوین. نشریه اقتصادی و توسعه کشاورزی (علوم و صنایع کشاورزی)، ۲۵: ۲۶۲-۲۵۳.

رحیمی‌خوب، ع. ۱۳۸۹. پیش‌بینی حداکثر بارندگی ماهانه ایستگاه ایلام از روی دمای سطح آب خلیج فارس و دریای سرخ با استفاده از روش داده‌کاوی، جله تحقیقات آب و خاک ایران، (۱) ۴۲، ۱-۷.

ستوده‌نیا، ع.، و س. ستوده‌نیا. ۱۳۹۵. برسی نقش احداث سد مخزنی طالقان بر رسوبگذاری شبکه آبیاری دشت قزوین، پانزدهمین کنفرانس هیدرولیک ایران، دانشکده فنی و مهندسی ، دانشگاه بین المللی امام خمینی، قزوین ۲۴ و ۲۵ آذر ۱۳۹۵

شهرابی، ج.، و ر. ص. نیستانی. ۱۳۸۷. "تجزیه و تحلیل سبد خرید مشتریان با استفاده از قوانین وابستگی در فروشگاه-های زنجیرهای شهروند "، دومین کنفرانس داده‌کاوی، دانشگاه امیرکبیر

عباس‌نژاد، ا.، و ع. شاهی‌دشت. ۱۳۹۲. بررسی آسیب‌پذیری دشت سیرجان با توجه به برداشت بی‌رویه از سفره آب زیرزمینی منطقه. جغرافیا و آمایش شهری- منطقه‌ای، (۷): ۹۶ -۸۵.

فتاحی، م. ه.، الف. بامداد و ع. رحیمی خوب. ۱۳۹۱. کاربرد قواعد انجمنی در رصد کردن وقایع بارندگی و خشکسالی با

استفاده از دمای سطح آبهای آزاد، مجله مهندسی منابع آب، سال پنجم، ص .۹۶-۸۹

قربانی، خ.، ع، خلیلی .، س ک، علوی‌پناه و غ. ر. نخعی‌زاده. ۱۳۸۸. مطالعه تطبیقی نمایه‌های هواشناسی خشکسالی SIAP, SPIبه روش داده‌کاوی (مطالعه موردی استان کرمانشاه). نشریه آب و خاک. ۲۴(۳): ۴۱۷-۴۲۶).

محمدی، م.، م. محمدی‌قلعه‌نی و ک. ابراهیمی. ۱۳۹۰. تغییرات زمانی و مکانی کیفیت آب زیرزمینی دشت قزوین. مجله پژوهش آب ایران، ۸ : ۵۲-۴۱.

نیکزاد، م.، م. ر. بهبهانی و ع. رحیمی‌خوب. ۱۳۹۲. آشکارسازی وابستگی‌های بین پارامترهای اقیانوسی-جوی و اقلیمی برای پایش خشکسالی با روش داده‌کاوی مطالعه موردی: استان خوزستان، مجله پژوهش آب ایران، (۱۳)۷، ۱۷۵-۱۸۳.

Agrawal, R., T. Imielinski, and A. Swami. 1993. Mining association rules between sets of items in large databases. conference on Management of data. June 1, 1993, Washington DC, USA. PP: 207–216.

Cheung, D. W., J. Han, V, Ng and Wong, C. Y. 1996. Maintenance of discovered association rules in large databases: An incremental updating technique. In Proc. 1996 Int. Conf.Data Engineering (ICDE'96), pages 106–114.

Dhanya, C. T. and D. Nagesh-Kumar. 2009. Data mining for evolution of association rules for droughts and floods in India using climate inputs. J. Geophys. Res., 114, D02102, doi:10.1029/2008JD010485.

Harms, S. K. and J. S. Deogun. 2004. Sequential association rule mining time lags. Journal of Intelligent Information Systems. 22(1): 7-22..

Jang, C.S., S.K. Chen. and Y.M. Ku. 2013. Applying indicator-based geostatistical approaches to determine potential zones of groundwater recharge based on borehole data. Catena, 101, 178–187.

Khan, S., H. F. Gabriel and T. Rana. 2008. Standard precipitation index to track drought and assess impact of rainfall on watertables in irrigation areas. Irrig Drainage Syst, 22, 159–177.

Ramageri, M. 2011. Data mining techniques and applications. Indian Journal of Computer Science and Engineering.;4(1): 301-5.

Tadesse, T. 2002. Identifying Drought and its association with climatic and Oceanic Parameters Using Data Mining Techniques. Nebraska, Graduate college University of Nebraska

Tan, P, M. Steinbach and V Kumar. 2006. Introduction to data mining. 2th ed. Boston: Addison-Wesley

The mining association rules for water table fluctuations analysis of Ghazvin city by Apriori algorithm

Seyyed Hassan Mirhashemi[1], Parviz Haghighat jou [2], Farhad Mirzaei[3], and Mehdi Panahi[4]

Abstract

The necessity of existence of a model is very important for suitable planning along with management of ground water resources and with regard to water scarcity and decline aquifers. Because of various factors in the amount of changes the depth of the aquifer can be effective. in this article by using association rules, dependency between aquifer depth changes in irrigation network of Qazvin city and human and natural factors affecting on it, was investigated. for the years 2001 to 2015 using Apriori algorithm. to evaluate accuracy of the rules used three indicator such as lift, support percent and confidence percent.The results show that the maximum positive correlation has been between two sets of the Antecedent value, consisting high percent of relative humidity and low volume of water in irrigation network and withdraw of agricultural water, and Consequent, rising of aquifer depth more than one meter. With 4.8 percent of support it can be stated that effective factors on aquifer fluctuations are low precipitation, high volume of water in irrigation network and high crop water requirements occurred simultaneously with decline of 0.3 to one meter. In cases that the temperature was high, air humidity and precipitation were low, the decline has been 0.3 to one meter with 66 percent of confidence. Furthermore, due to the derived association rules by increasing withdrawal of water from the well, air temperature, volume of water in irrigation network and crop water requirements, causes aquifer drops and On the other hand, increasing air humidity and precipitation causes a slump in aquifer decline.

Keywords: Air humidity, Association algorithm, Data mining, Irrigation network, Rainfall depth.

[1] Ph.D. student, Dept. of Water Engineering, Faculty of Water and Soil, University of Zabol.
[2] Assistant Professor, Dept. of Water Engineering, Faculty of Water and Soil, University of Zabol.
[3] Associate Professor, Dept. of Irrigation and Drainage, Faculty of Agriculture and Natural Resources, University of Tehran.
[4] Assistant Professor, Dept. of Water Engineering, Faculty of Agriculture, University of Zanjan.