

## پیش‌بینی وقوع سیلاب با استفاده از نظریه مجموعه‌های زبر (مطالعه موردی: رودخانه هلیل رود)

امین حسین‌پور میل‌آغاردان<sup>۱</sup>، رحیم‌علی‌عباسپور<sup>۲\*</sup>، فاطمه شیدایی<sup>۳</sup>

۱. دانشجوی دکتری گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه برداری و اطلاعات مکانی،  
پردیس دانشکده‌های فنی دانشگاه تهران  
Amin\_hoseinpoor@ut.ac.ir

۲. استادیار گروه سیستم‌های اطلاعات مکانی، دانشکده مهندسی نقشه برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی  
دانشگاه تهران  
sheidaiefateme@gmail.com

۳. کارشناس ارشد بخش ترویج و آموزش کشاورزی دانشگاه شیراز

تاریخ پذیرش مقاله: ۹۴/۴/۲۴

تاریخ وصول مقاله: ۹۴/۲/۳

### چکیده

پژوهش حاضر روشی برای پیش‌بینی وقوع سیلاب به صورت روزانه، با استفاده از نظریه مجموعه‌های زبر ارائه کرده است تا علاوه بر مدیریت ریسک وقوع آن، عدم قطعیت مستخرج از داده‌های استفاده‌شده، آنالیز شود. به همین منظور پارامترهای میزان بارندگی، حداقل دما، مقدار تبخیر و دبی رودخانه به صورت روزانه، برای استخراج قوانین قوی به منظور پیش‌بینی وقوع سیلاب به کار گرفته شدند. با استفاده از این پارامترها و در نظر گرفتن تأثیر هم‌زمان آن‌ها در وقوع سیلاب، برای مدل‌سازی، قوانین محتمل برای وقوع سیلاب استخراج و با استفاده از روابط مطرح در نظریه مجموعه‌های زبر، بهترین قوانین به منظور پیش‌بینی روزانه سیلاب انتخاب شدند. داده‌های جمع‌آوری‌شده به مدت چهار سال از سد جیرفت با دقت روزانه برای آنالیز و استخراج قوانین استفاده شدند. در این مسیر ابتدا پیش‌پردازش داده‌ها صورت گرفت و بازه‌های زمانی وقوع سیلاب از هر سال جدا شدند. سپس، گسسته‌سازی داده‌ها، در پی آن تقریب و تقلیل داده‌ها انجام و با شناسایی هسته‌های مؤثر ویژگی‌ها، محتمل‌ترین قوانین استخراج شد. در نهایت داده‌های سال ۸۸ برای ارزیابی قدرت قوانین استخراج‌شده به کار گرفته شده و با استفاده از ماتریس آشفتگی مقدار ۰/۸۴ برای ضریب کاپا محاسبه شد. همچنین، قوانین قوی به‌دست‌آمده در نتایج با ۷۲ درصد از موارد وقوع سیلاب در سال ۱۳۸۸ مطابقت دارد.

### کلیدواژه

پیش‌بینی سیلاب، دبی رودخانه، نظریه مجموعه‌های زبر.

### ۱. سرآغاز

پارامترها در یک دهه گذشته محققان از روش‌های مختلفی برای پیش‌بینی سیلاب استفاده کرده‌اند. در این میان تمایل به استفاده از شبکه‌های عصبی مصنوعی در بین هیدرولوژیست‌ها به شدت افزایش یافته است (Kumar, et al., 2005; Kisi, 2008; Rezaeian Zadeh, et al., 2010; Mukerji, et al., 2009; Mutlu, et al., 2008; Sahay and Srivastava, 2014) که از جمله آن‌ها می‌توان به پژوهش چائو و همکارانش که به صورت ترکیبی از دو مدل شبکه عصبی مصنوعی برای پیش‌بینی سیلاب رودخانه در چین استفاده کردند (Chau, et al., 2005) و لیدین برای

سیل به مفهوم طغیان آب، خسارات بسیاری از لحاظ زیستی، اقتصادی و جانی در سراسر جهان به وجود آورده است. به طوری که بین سال‌های ۱۹۸۵ و ۲۰۰۴ بیش از ۱۵ میلیارد دلار خسارت مالی و نزدیک به ۷۵۰۰ مرگ و میر را همراه داشته است (Stromberg, 2007). بخش اصلی این تلفات و خسارات در کشورهای در حال توسعه رخ داده است (Rezaeianzadeh, et al., 2013). با توجه به عوامل متعدد تأثیرگذار در وقوع سیل و ناشناخته‌بودن روابط بین

یکی از معایب اشاره کرد. البته در برخی پژوهش‌ها برای حل مشکل عدم قطعیت وابسته به دانش کارشناسان، از شبکه عصبی مصنوعی در کنار سیستم استنتاج فازی و داده‌های پیشین برای ارائه قوانین بهینه استفاده شده است (Talei, et al., 2013; Yeh and Yang, 2014).

با این حال موضوعی که در این مطالعات کمتر به آن توجه شده است، عدم قطعیت در داده‌های استفاده‌شده برای استخراج الگوی وقوع سیلاب است، زیرا در نظر گرفتن عدم قطعیت، با توجه به حجم بالایی از داده‌ها که برای استخراج الگو به کار گرفته می‌شود، سبب بهبود نتایج خواهد بود (Rodríguez-Rincón, et al., 2014). این در حالی است که در این پژوهش، با انتخاب روش نظریه مجموعه‌های زبر<sup>۱</sup> علاوه بر در نظر گرفتن عدم قطعیت موجود در داده در پروسه پیش‌بینی، ارتباط وقوع سیلاب و پارامترهای مؤثر بررسی و با استخراج این ارتباطها، قوانین استخراج شده است که در عین سادگی، تأثیر همزمان مهم‌ترین پارامترهای مؤثر در وقوع سیلاب را ارائه می‌دهد. سپس، با استفاده از روابط موجود در این نظریه، ارتباط پارامترها به صورت چندتایی در وقوع سیلاب بررسی شده است و با حذف قوانین مشابه، کارآمدترین قوانین برای شناسایی محتمل‌ترین شرایط وقوع سیلاب به دست می‌آید. با استفاده از این روش علاوه بر اینکه عدم قطعیت موجود در داده در پروسه پیش‌بینی در نظر گرفته می‌شود، قوانین برای پیش‌بینی سیلاب استخراج می‌شوند که محتمل‌ترین شرایط برای وقوع سیلاب را ارائه می‌دهند. نظریه مجموعه‌های زبر را اولین بار پاولاک در ۱۹۸۲ ارائه کرد که در ارتباط با ابهام و عدم قطعیت در داده‌هاست و با در نظر گرفتن کران بالا و پایین از مجموعه مرجع، تقریبی از مجموعه زبر تحت بررسی را مشخص می‌کند (Pawlak, 1982). از مزیت‌های برجسته این روش نسبت به سایر روش‌ها، می‌توان به ویژگی استخراج دانش و قوانین محتمل برای مجموعه مورد نظر اشاره کرد.

پیش‌بینی جریان رودخانه در سودان اشاره کرد (Shamseldin, 2010). توکار و جانسون نیز به وسیله این روش و با استفاده از داده‌های روزانه بارندگی، تبخیر، تعرق و دما در حوزه رودخانه مرلند امریکا به پیش‌بینی و برآورد رواناب پرداختند (Tokar and Johnson, 1999). همچنین، در بسیاری از مطالعات، مدل‌های هیدرولوژیک شناخته‌شده برای پیش‌بینی هفتگی یا ماهانه سیلاب به کار گرفته شده‌اند (Bartholmes and Todini, 2005; Refsgaard, et al., 1988). رگرسیون (England and Hisdal, 2009; Eslamian, et al., 2010; Rezaeianzadeh, et al., 2013) و سیستم استنتاج فازی (Nayak, et al., 2005) استفاده کرده‌اند. به طور کلی، بررسی پژوهش‌ها در زمینه پیش‌بینی سیلاب بیانگر آن است که در اکثر این مطالعات از داده‌های اندازه‌گیری شده مانند دبی رودخانه یا دما و رطوبت استفاده شده است. این داده‌ها طی چند سال جمع‌آوری و در روش‌های مذکور به کار گرفته شده‌اند. در این بین وجه مشترک تمامی مطالعات، استفاده از الگوی داده‌ها برای آموزش مدل و در نهایت پیش‌بینی سیلاب است. به طوری که روشی همانند شبکه عصبی مصنوعی که مورد استفاده بسیاری از محققان بوده است، تنها با تکیه بر توانایی تشخیص الگو و برقراری ارتباط میان ورودی و خروجی داده‌ها نتایج مورد نظر را ارائه می‌دهد (Kisi, 2004; Nourani, et al., 2009; Tiwari, et al., 2013). به این ترتیب در روش رگرسیون، محاسبه ضرایب با استفاده از الگوی پیشین داده‌ها و بررسی میزان تأثیر هر پارامتر به طور مستقل در وقوع سیلاب، صورت می‌گیرد (Sahay, 2013). در برخی مطالعات نیز که از روش فازی استفاده کرده‌اند، تنها دانش کارشناسان برای تعریف توابع عضویت فازی و قوانین موتور استنتاج فازی به کار گرفته شده است (Mahabir, et al., 2007). در این میان می‌توان به وجود عدم قطعیت در نظرهای کارشناسان و اتکا این روش‌ها به نظرهای آنان به منزله

در منطقه مورد مطالعه، که دارای آب و هوایی گرمسیری با میزان بارش اندک در سال است، رخ نمی‌دهد. بنابراین، از میان داده‌های چهار سال ۱۳۸۲ تا ۱۳۸۶، داده‌های با دقت ماهانه برای آنالیز انتخاب شده‌اند و از دستگاه‌های هیدرومتری نصب‌شده در ورودی سد جیرفت در بازه زمانی ۱۳۸۲ تا ۱۳۸۸ استفاده شده است. این داده‌ها ابتدا پیش‌پردازش، سپس استفاده شده‌اند و هر یک از آنان حداقل یک مورد سیلاب با دبی بیشتر از ۱۰ متر مکعب بر ثانیه را دربر دارند. در نهایت داده‌های ثبت‌شده سال ۱۳۸۸ برای ارزیابی نتایج به کار گرفته شدند. جدول ۱ نمونه‌ای از داده‌های اندازه‌گیری شده در ورودی سد جیرفت را نشان می‌دهد.

#### ۴. روش پژوهش

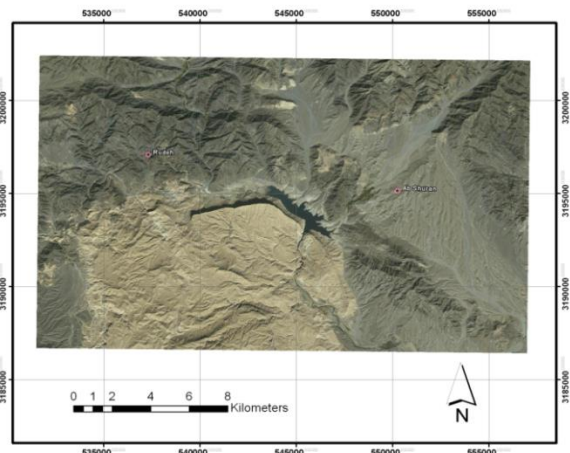
نظریه مجموعه‌های زبر با تحلیل جدول‌های داده در ارتباط است. این جداول می‌توانند از طریق اندازه‌گیری یا افراد متخصص و خبره به دست آمده باشند. هدف اصلی از تحلیل مجموعه زبر به دست آوردن مفاهیم تقریبی از طریق داده‌های اکتسابی است و روش‌هایی را برای زدودن و کاستن اطلاعات دانش غیردقیق یا مازاد بر نیاز از پایگاه‌های داده‌ها مهیا می‌کند. فرایند حذف داده‌های زائد، بر مبنای آموزش و بدون از دست دادن داده‌های اساسی پایگاه داده‌ها صورت می‌گیرد. در نتیجه تقلیل اطلاعات، مجموعه‌ای از قوانین تلخیص شده و پرمعنا حاصل می‌شوند که کار تصمیم‌گیری را بسیار ساده‌تر می‌کنند. در حقیقت می‌توان گفت که مجموعه زبر با کاهش فضای داده‌ها و برگزیدن ترم‌های مهم، یک نگاهت از فضای داده‌های خام و ترم‌ها به فضای سمانتیک (مفاهیم) انجام می‌دهد. لذا با توجه به رشد انفجاری حجم اطلاعات، مجموعه زبر می‌تواند نقش بسیار مؤثری در سیستم‌های پشتیبان تصمیم‌گیری داشته باشد (Pawlak, 1982).

جدول ۱. نمونه‌ای از داده‌های استفاده‌شده در پژوهش. منبع: سازمان آب منطقه‌ای استان کرمان، ۱۳۸۹

در ادامه، منطقه مورد مطالعه و داده‌های استفاده‌شده در بخش دوم و روش انجام پژوهش در بخش سوم ارائه شده و بخش چهارم شامل نتایج پژوهش و بحث روی نتایج است و در بخش انتهایی نتیجه‌گیری گنجانده شده است.

#### ۲. منطقه مورد مطالعه

منطقه مورد مطالعه رودخانه هلیل‌رود در جنوب‌شرقی کشور در استان کرمان قرار دارد. این منطقه در محدوده ۲۸ درجه و ۵۰ دقیقه و ۵۸ ثانیه تا ۲۸ درجه و ۵۴ دقیقه و ۵۳ ثانیه عرض جغرافیایی شمالی و ۵۷ درجه و ۲۳ دقیقه و ۶ ثانیه تا ۵۷ درجه و ۲۳ دقیقه و ۳ ثانیه طول جغرافیایی شرقی قرار دارد. شکل ۱ محدوده منطقه مورد مطالعه را نشان می‌دهد (برزگر و همکاران، ۱۳۸۹).



شکل ۱. منطقه مورد مطالعه (بخشی از رودخانه هلیل‌رود)

#### ۳. داده‌های استفاده‌شده

پارامترهای استفاده‌شده در این پژوهش شامل میزان بارندگی، حداقل روزانه دما، تبخیر و دبی ثبت شده است. کاربرد نظریه مجموعه‌های زبر، استخراج الگوی همزمان پارامترهای مؤثر در وقوع سیلاب، از جمله اهداف این پژوهش است. سیلاب در تمامی روزها و فصول، خصوصاً

| تاریخ    | بارندگی mm | حداقل دما C | حداکثر دما C | تبخیر mm | متوسط ورودی m <sup>3</sup> /s |
|----------|------------|-------------|--------------|----------|-------------------------------|
| ۸۴/۰۱/۰۱ | ۰/۰        | ۱۲/۰        | ۲۴/۰         | ۵/۰      | ۵۷/۹                          |
| ۸۴/۰۱/۰۲ | ۰/۰        | ۱۳/۰        | ۲۴/۵         | ۶/۰      | ۱/۸                           |
| ۸۴/۰۱/۰۳ | ۰/۰        | ۱۴/۰        | ۲۵/۰         | ۸/۰      | ۱/۸                           |
| ۸۴/۰۱/۰۴ | ۰/۰        | ۱۶/۰        | ۲۸/۰         | ۵/۰      | ۱/۸                           |
| ۸۴/۰۱/۰۵ | ۰/۰        | ۱۵/۰        | ۲۸/۰         | ۶/۰      | ۱/۸                           |
| ۸۴/۰۱/۰۶ | ۰/۰        | ۱۵/۰        | ۲۷/۰         | ۷/۰      | ۱/۸                           |
| ۸۴/۰۱/۰۷ | ۰/۰        | ۱۶/۰        | ۲۸/۰         | ۸/۰      | ۱/۹                           |
| ۸۴/۰۱/۰۸ | ۰/۰        | ۱۴/۰        | ۲۴/۰         | ۶/۰      | ۱/۹                           |
| ۸۴/۰۱/۰۹ | ۰/۰        | ۱۳/۰        | ۲۴/۰         | ۵/۰      | ۲/۲                           |
| ۸۴/۰۱/۱۰ | ۰/۰        | ۱۴/۰        | ۲۵/۵         | ۵/۰      | ۴/۷                           |
| ۸۴/۰۱/۱۱ | ۰/۰        | ۱۳/۵        | ۲۵/۰         | ۷/۰      | ۵/۱                           |
| ۸۴/۰۱/۱۲ | ۰/۰        | ۱۶/۵        | ۲۹/۰         | ۱۲/۰     | ۱۱/۵                          |
| ۸۴/۰۱/۱۳ | ۰/۰        | ۱۵/۰        | ۲۷/۰         | ۱۰/۰     | ۱۱/۵                          |
| ۸۴/۰۱/۱۴ | ۰/۰        | ۱۷/۰        | ۳۰/۰         | ۱۲/۰     | ۱۱/۵                          |
| ۸۴/۰۱/۱۵ | ۰/۰        | ۱۸/۰        | ۳۶/۵         | ۹/۰      | ۱۱/۵                          |
| ۸۴/۰۱/۱۶ | ۰/۰        | ۱۸/۰        | ۳۴/۰         | ۱۰/۰     | ۱۳/۶                          |
| ۸۴/۰۱/۱۷ | ۰/۰        | ۱۶/۰        | ۳۴/۰         | ۸/۰      | ۱۳/۰                          |
| ۸۴/۰۱/۱۸ | ۰/۰        | ۱۴/۰        | ۲۶/۰         | ۱۱/۰     | ۱۱/۳                          |
| ۸۴/۰۱/۱۹ | ۰/۰        | ۱۲/۰        | ۲۶/۰         | ۱۰/۰     | ۱۱/۱                          |

#### ۱.۴. سیستم اطلاعات<sup>۲</sup>

به طور کلی از این نظریه برای استخراج قوانین تصمیم‌گیری از یک سیستم تصمیم‌گیر، که حالت خاصی از یک سیستم اطلاعاتی است، استفاده می‌شود. یک سیستم اطلاعاتی به صورت  $S = (p, V, Q, U)$  است که در آن  $U$  بیانگر یک مجموعه متناهی و غیرتهی از اشیاست و با نام مجموعه مرجع خوانده می‌شود و  $Q$  بیانگر یک مجموعه متناهی و غیرتهی از ویژگی‌هاست. همچنین،  $V = U_{q \in Q} V_q$  است که در آن  $V_q$  دامنه ویژگی  $q$  است و  $p: U \times Q \rightarrow V$  تابعی کلی است که در آن به ازای هر  $X \in U$  و  $q \in Q$  یک تابع اطلاعات به صورت  $\rho(X, q) \in V_q$  تعریف می‌شود. همچنین، هر زوج  $q, v \in V_q$  و  $(v, q)$  به منزله یک توصیف‌کننده در  $S$  گفته می‌شود. سیستم اطلاعات معمولاً به صورت یک جدول ارائه می‌شود که در

گاهی مشاهداتی را در اختیار داریم که در آن تشخیص المان‌های مجموعه به دلیل وجود عدم قطعیت یا داده‌های غیردقیق ناممکن است و این دلیلی برای استفاده از نظریه مجموعه‌های زبر است. به عبارت دیگر، اطلاعات غیردقیق سبب تجزیه‌ناپذیری المان‌ها بر اساس داده‌های موجود می‌شوند. موضوع تجزیه‌ناپذیری المان‌ها از طبقه‌بندی دقیق آن‌ها جلوگیری می‌کند و از طرفی روابط تجزیه‌ناپذیری برای تعریف دو عملگر تقریب بالا و پایین روی داده‌ها استفاده می‌شوند. با استفاده از این دو عملگر، دقت تقریب بیان می‌شود که به صورت عددی در بازه  $[0, 1]$  است و کیفیت طبقه‌بندی را نشان می‌دهد. دیگر مفاهیم نظریه زبر شامل تقلیل ویژگی‌ها، طبقه‌بندی زبر و قوانین تصمیم‌گیری‌اند که در ادامه به آن‌ها اشاره می‌شود.

تشخیص دهد و  $Bn_p(Y)$  شامل مجموعه‌اشیایی است که نمی‌تواند به درستی و با در نظر گرفتن مجموعه‌ویژگی  $P$  المان  $Y$  را احتمالاً طبقه‌بندی کند. همچنین، دقت تقریب المان  $Y$  با استفاده از مجموعه‌ویژگی  $P$  در  $S$  به صورت زیر بیان می‌شود.

$$\mu_p(Y) = \frac{Card(PY)}{Card(\overline{PY})}$$

در این رابطه  $card(PY)$  بیانگر تعداد اعضای مجموعه تقریب پایین و  $Card(\overline{PY})$  نیز بیانگر تعداد اعضای مجموعه تقریب بالاست (Pawlak, 1982).

#### ۴.۴. تقلیل ویژگی‌ها<sup>۸</sup>

با فرض دو مجموعه‌ویژگی  $R$  و  $P$  به طوری که  $R \subseteq Q$  و  $P \subseteq Q$ ، گفته می‌شود اگر  $\tilde{P} \in \tilde{R}$  باشد در این صورت مجموعه‌ویژگی  $R$  وابسته به مجموعه‌ویژگی  $P$  است. بدین ترتیب با شناسایی وابستگی بین ویژگی‌ها، تقلیل ویژگی‌ها ممکن خواهد بود. همچنین، مجموعه‌ویژگی  $P$  در  $Q$  مستقل خواهد بود اگر به ازای هر  $\tilde{P} \subseteq Q$  داشته باشیم  $\tilde{P} \supseteq P$  در غیر این صورت  $P \subseteq Q$  در  $S$  به منزله یک مجموعه‌ویژگی وابسته خواهد بود. در واقع با استفاده از تقلیل ویژگی‌ها به دنبال کاهش ویژگی‌هایی هستیم که در مجموعه  $S$  اضافی‌اند. در این صورت مجموعه‌ویژگی  $P \subseteq Q$  یک تقلیل از  $Q$  در  $S$  خواهد بود، اگر و تنها اگر  $P$  بزرگ‌ترین مجموعه‌ویژگی مستقل در  $Q$  باشد. یکی از راه‌های یافتن یک مجموعه‌ویژگی تقلیل‌یافته استفاده از دقت تقریب طبقه‌بندی است. در این صورت کوچک‌ترین مجموعه‌ویژگی، از میان مجموعه‌ویژگی‌هایی که دقت تقریب برابر دارند به منزله یک مجموعه‌ویژگی تقلیل در  $S$  شناخته می‌شود. همچنین، در صورتی که در یک مجموعه چندین مجموعه‌ویژگی با تعداد اعضای برابر داشته باشیم، به مجموعه‌ویژگی مشترک در این مجموعه‌های تقلیل، هسته گفته می‌شود (Pawlak, 1982).

آن سطرها نشان‌دهنده اشیا، ستون‌ها نشان‌دهنده ویژگی‌ها و ستون‌ها بیانگر ویژگی‌ها هستند و هر سلول حاوی مقادیر ویژگی‌ها برای هر شیء است (Pawlak, 1982).

#### ۲.۴. روابط تجزیه‌ناپذیر<sup>۳</sup>

در یک سیستم اطلاعات مانند  $S = (p, V, Q, U)$  و با فرض  $P \subseteq Q, X, Y \in U$ ، گفته می‌شود که  $X$  و  $Y$  با مجموعه‌ای از ویژگی‌های  $P$  در  $S$  تجزیه‌ناپذیرند (به صورت  $\tilde{P} Y X \tilde{P}$  بیان می‌شود) اگر به ازای هر  $q \in p$  داشته باشیم  $r(x, q) = r(y, q)$ ، همچنین، کلاس‌های یکسان مقادیر  $\tilde{P}$  را یک مجموعه  $P$ -elementary در  $S$  گویند. به همین ترتیب با در نظر گرفتن تمام ویژگی‌ها و کلاس‌های یکسان، یک مجموعه  $Q$ -elementary در  $S$  وجود خواهد داشت که به آن‌ها اتم گفته می‌شود (Pawlak, 1982).

#### ۳.۴. تقریب مجموعه‌زبر<sup>۴</sup>

به منظور ارزیابی کیفیت توصیف اشیا از طریق مجموعه  $\{Des_p(X), X \in P\}$ ، مفاهیم زیر از طریق پاولاک ارائه شد. با فرض  $P \subseteq Q$  و  $Y \subseteq U$  تقریب پایین المان<sup>۵</sup>  $Y$  (با در نظر گرفتن ویژگی‌های  $P$ ) به صورت  $\underline{PY}$  و تقریب بالای المان<sup>۶</sup>  $Y$  به صورت  $\overline{PY}$  نشان داده شده است و به صورت زیر تعریف می‌شود.

$$\underline{PY} = \bigcup X \{X \in P \text{ and } X \subseteq Y\}$$

$$\overline{PY} = \bigcup X \{X \in P \text{ and } X \cap Y \neq \emptyset\}$$

همچنین، مجموعه‌مرز با در نظر گرفتن ویژگی‌های  $P$  به صورت زیر تعریف شده است.

$$Bn_p(Y) = \overline{PY} - \underline{PY}$$

در توضیح این روابط باید گفت که تقریب پایین المان  $Y$  در واقع مجموعه‌ای از اشیا در مجموعه مرجع است که با در نظر گرفتن مجموعه‌ویژگی  $P$  می‌تواند المان  $Y$  را به درستی طبقه‌بندی و تشخیص دهد. تقریب بالای المان  $Y$  در مجموعه‌ای از اشیا در مجموعه مرجع است که با در نظر گرفتن مجموعه‌ویژگی  $P$  می‌تواند المان  $Y$  را احتمالاً

۵.۴. جدول تصمیم‌گیری<sup>۹</sup>

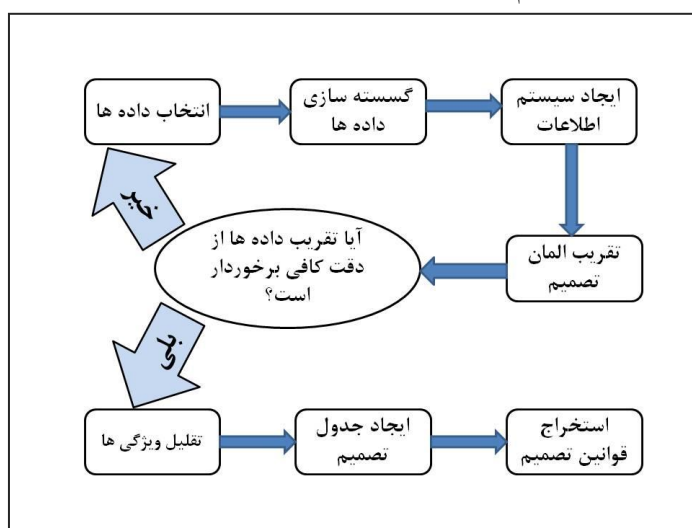
آخرین قسمت از مراحل نظریه مجموعه زبر، ایجاد جدولی است که شامل چند مجموعه ویژگی A (در این جدول، شرایط تصمیم گفته می‌شود) و چند المان تصمیم‌گیری D است. البته پس از ایجاد این جدول بایستی المان‌هایی را که با یک یا چند ویژگی مشترک یک تصمیم را نتیجه می‌دهند کاهش داد تا جدول بهینه شود. از اطلاعات این جدول می‌توان برای طبقه‌بندی اشیا و استخراج قوانین تصمیم‌گیری<sup>۱۰</sup> استفاده کرد. قوانین تصمیم‌گیری مستخرج از این جدول به صورت  $a_k \rightarrow d_j$  می‌باشند که در آن یکسری مجموعه ویژگی به منزله شرایط تصمیم‌گیری ( $a_k$ ) و یک المان تصمیم ( $d_j$ ) وجود دارد (Pawlak, 1982).

## ۵. پیاده‌سازی و نتایج

با توجه به مراحل ارائه شده برای نظریه مجموعه‌های زبر، سیر کلی انجام این پژوهش در شکل ۲ ملاحظه می‌شود و اولین قدم در آن انتخاب داده‌هاست. برای این منظور داده‌های روزانه چهار سال ۱۳۸۲ تا ۱۳۸۶ به صورت چهار پارامتر میزان بارندگی، حداقل دما، تبخیر و دبی رودخانه، به منزله ورودی تحقیق انتخاب شدند و پیش‌پردازش داده‌ها به منظور حذف داده‌های پرت روی آن‌ها اعمال شد. در مرحله بعدی، گسسته‌سازی داده‌ها برای هر فصل، به صورت جداگانه انجام گرفت و چهار سیستم اطلاعات

ایجاد شد. سپس، تقریب بالا و پایین برای المان تصمیم در هر یک از سیستم‌های اطلاعات محاسبه شده است و در صورت قابل قبول بودن مقادیر مربوطه، تقلیل ویژگی روی سیستم اطلاعات انجام می‌شود. در نتیجه تقلیل ویژگی، جدول تصمیم تهیه می‌شود و قوانین تصمیم از این جدول استخراج خواهند شد. در انتها برای ارزیابی نتایج، داده‌های سال ۱۳۸۸ استفاده شد. در این تحقیق برای پردازش داده‌ها، نرم‌افزار تخصصی ROSE2 به کار گرفته شد که در خصوص نظریه مجموعه‌های زبر ارائه شده است.

در اولین مرحله پژوهش کلیه داده‌های سال‌های ۱۳۸۲ تا ۱۳۸۶ به صورت سالانه و کامل برای پردازش انتخاب شدند. بدین ترتیب که پارامترهای میزان بارندگی، حداقل دما و تبخیر روز  $t$  به منزله مجموعه ویژگی‌ها و میزان دبی رودخانه در روز  $(t+1)$  برای المان تصمیم انتخاب شد. در مرحله بعد برای گسسته‌سازی داده‌ها از آنجا که میانگین تبخیر و حداقل دما در فصول مختلف سال متفاوت است امکان قراردادن تمامی داده‌ها در یک جدول سیستم اطلاعاتی میسر نیست، به همین دلیل داده‌ها بر حسب فصول مختلف جدا و در نهایت به صورت چهار جدول سیستم اطلاعاتی تهیه شدند. به طوری که هر جدول شامل داده‌های فصلی برای چهار سال است. بعد از انتخاب داده‌ها، گسسته‌سازی داده با در نظر گرفتن میانگین آن‌ها مطابق جدول ۲ صورت گرفت.



شکل ۲. مراحل انجام پژوهش

جدول ۲. مقادیر انتخاب‌شده برای گسسته‌سازی

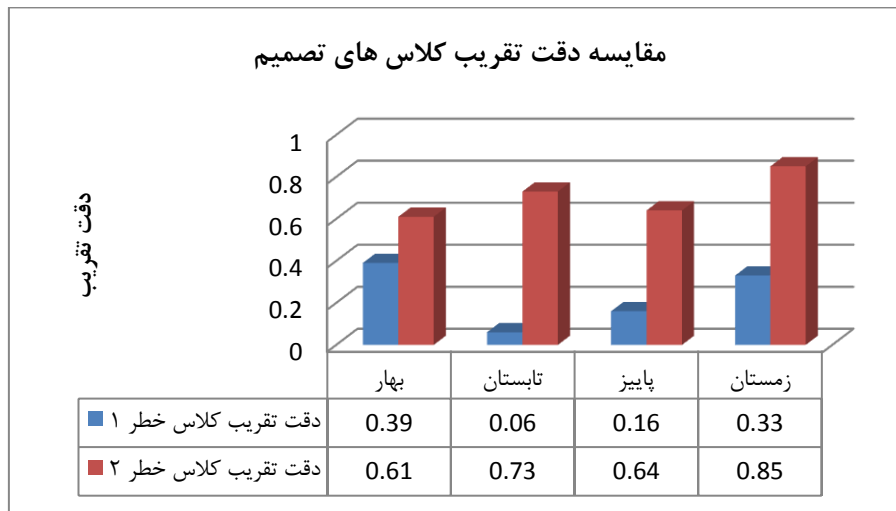
| ردیف | زمان    | میزان بارندگی | حداقل دما | تبخیر | دبی رودخانه |
|------|---------|---------------|-----------|-------|-------------|
| ۱    | بهار    | ۰             | ۱۵        | ۶     | ۱۰          |
| ۲    | تابستان | ۰             | ۲۵        | ۱۲    | ۱۰          |
| ۳    | پاییز   | ۰             | ۱۳        | ۷     | ۱۰          |
| ۴    | زمستان  | ۰             | ۸         | ۳     | ۱۰          |

در مرحله بعد به منظور ارزیابی داده‌ها، تقریب بالا و پایین هر یک از سیستم‌های اطلاعات به صورت مجزا محاسبه شد که نتایج آن در شکل ۴ ملاحظه می‌شود. در این شکل برای هر سیستم اطلاعات، تقریب پایین و بالا برای هر یک از کلاس‌های تصمیم‌گیری همراه دقت تقریب ارائه شده است. از آنجا که هدف این پژوهش استخراج قوانین وقوع سیلاب است، تنها سیستم‌های اطلاعاتی را - که بیشترین دقت تقریب را برای کلاس ۱ تصمیم‌گیری داشته باشند - می‌توان برای مرحله بعد انتخاب کرد. همان‌طور که در این شکل مشاهده می‌شود کمترین دقت تقریب برای تابستان و پاییز و بیشترین آن برای بهار و زمستان است.

در این جدول مقادیر گسسته‌سازی داده‌ها ارائه شده است. بدین ترتیب که میزان بارندگی صفر بیانگر این است که مقادیر بیشتر از صفر به منزله شرایط مساعد برای سیلاب و با عدد یک و مقادیر کمتر از آن با عدد صفر در جدول سیستم اطلاعات نشان داده خواهد شد. مقادیر دبی بیشتر از ۱۰ با عدد یک و کمتر از آن با عدد ۲ در جدول سیستم اطلاعاتی وارد می‌شوند. در خصوص ویژگی تبخیر و حداقل دما مرز گسسته‌سازی برای هر فصل متفاوت است. بدین ترتیب که برای ویژگی حداقل دما و تبخیر مقادیر کمتر از عدد ارائه‌شده در جدول ۲، با عدد ۱ در جدول سیستم اطلاعاتی ارائه می‌شوند. نمونه‌ای از سیستم اطلاعاتی ایجادشده در شکل ۳ ملاحظه می‌شود.

|    | precipitation | Min_Temperature | evaporation | Decision |
|----|---------------|-----------------|-------------|----------|
| 1  | 0             | 0               | 0           | 2        |
| 2  | 1             | 1               | 1           | 1        |
| 3  | 1             | 1               | 1           | 1        |
| 4  | 0             | 0               | 1           | 2        |
| 5  | 0             | 0               | 1           | 2        |
| 6  | 0             | 0               | 1           | 2        |
| 7  | 0             | 0               | 0           | 2        |
| 8  | 0             | 1               | 1           | 2        |
| 9  | 0             | 0               | 0           | 2        |
| 10 | 0             | 0               | 0           | 2        |
| 11 | 0             | 0               | 1           | 2        |
| 12 | 0             | 0               | 0           | 2        |
| 13 | 0             | 0               | 0           | 2        |

شکل ۳. سیستم اطلاعاتی ایجادشده برای پیش‌بینی سیلاب



شکل ۴. نمودار مقایسه دقت تقریب کلاس‌های تصمیم

مقادیر A1، A2 و A3 به ترتیب بیانگر ویژگی‌های بارش، حداقل دما و تبخیرند. همان‌طور که ملاحظه می‌شود تنها تقلیل ویژگی از سیستم‌های اطلاعات خرداد و فروردین نتیجه شده است با این تفاوت که ویژگی‌های تقلیل آن‌ها معادل نیستند. در این پژوهش پس از تعیین ویژگی‌های تقلیل، جدول تصمیم و استخراج قوانین تصمیم صورت گرفت که نتایج آن در جدول ۴ ارائه شده است. نتایج این جدول شامل دو نوع قوانین قوی و تقریبی‌اند. قوانین قوی شامل یک کلاس تصمیم و قوانین تقریبی هر دو کلاس تصمیم را شامل می‌شوند. برخی از این قوانین مانند ۱ و ۲ و ۵ با هم همپوشانی دارند که می‌توان در مرحله استفاده از قوانین، قانون جزئی را حذف کرد. در این پژوهش نیز قانون ۵ حذف شد.

شایان یادآوری است که تفاوت نسبی دقت تقریب کلاس‌ها به دلیل وجود دبی ثبت شده کمتر از ۱۰ متر مکعب بر ثانیه، در بیشتر روزهای سال است. با توجه به این نمودار برای رسیدن به نتایج بهتر سیستم اطلاعاتی بهار و زمستان برای تقلیل ویژگی انتخاب شد. برای تقلیل ویژگی‌ها در این پژوهش ابتدا سیستم‌های اطلاعاتی بهار و زمستان ادغام شدند و یک سیستم اطلاعات ایجاد شد. سپس، با استفاده از روابط مطرح در نظریه مجموعه زبر تقلیل ویژگی برای این سیستم اطلاعات جدید انجام شد که در پی آن هیچ‌یک از ویژگی‌ها حذف نشدند. دلیل این امر را می‌توان در تنوع شرایط و حجم بالای داده‌ها در این سیستم جستجو کرد. در ادامه برای رسیدن به قوانین ساده تصمیم، تقلیل ویژگی با تقسیم سیستم اطلاعات جدید به فصل‌ها و تشکیل سیستم‌های اطلاعات مجزا برای هر ماه از فصل بهار تکرار شد که نتایج آن در جدول ۳ ملاحظه می‌شود.

جدول ۳. نتایج تقلیل ویژگی

| ویژگی‌های تقلیل یافته | بهار       | زمستان     | فروردین | اردیبهشت   | خرداد   |
|-----------------------|------------|------------|---------|------------|---------|
|                       | A3، A2، A1 | A3، A2، A1 | A3 و A1 | A3، A2، A1 | A3 و A2 |



جدول ۴. قوانین استخراج‌شده

| قانون   | ویژگی A1 | ویژگی A2 | ویژگی A3 | کلاس تصمیم | تعداد موارد مطابقت با سال ۱۳۸۸ |
|---------|----------|----------|----------|------------|--------------------------------|
| قانون ۱ | ۱        | *        | *        | ۱          | ۱۵                             |
| قانون ۲ | ۱        | ۱        | ۰        | ۱          | -                              |
| قانون ۳ | ۰        | ۱        | ۱        | ۱          | ۴                              |
| قانون ۴ | ۱        | *        | ۰        | ۲          | -                              |
| قانون ۵ | ۱        | ۱        | ۱        | ۱          | -                              |
| قانون ۶ | ۰        | ۱        | ۰        | ۱          | ۲                              |
| قانون ۷ | *        | ۱        | ۱        | ۱ یا ۲     | -                              |
| قانون ۸ | ۰        | ۱        | *        | ۱ یا ۲     | -                              |
| قانون ۹ | ۰        | ۱        | ۰        | ۱ یا ۲     | -                              |

۷۲ درصد تطابق قوانین استخراج‌شده با موارد وقوع سیلاب است که بیانگر توانایی این قوانین برای شناسایی موارد محتمل وقوع سیلاب است. البته شایان یادآوری است که قوانین قوی استخراج‌شده از طریق روش نظریه زبر بر اساس تقریب پایین کلاس ۱ است. بدین ترتیب که المان‌های تشکیل‌دهنده مرز یا عدم قطعیت در این قوانین استفاده نشده‌اند و تنها برای قوانین تقریبی از المان‌های مرزی و تقریب بالای کلاس ۱ استفاده شده است. بدین ترتیب عدم قطعیت موجود در داده‌ها شناسایی و از پروسه استخراج قوانین حذف شده است. این امر یکی از اهداف پژوهش است که با استفاده از نظریه مجموعه زبر محقق شده است. همچنین، در این تحقیق به منظور بررسی دقیق‌تر نتایج از ماتریس آشفتگی<sup>۱۱</sup> استفاده شده است که در جدول ۵ ملاحظه می‌شود.

با استفاده از مقادیر به‌دست‌آمده برای ماتریس آشفتگی مقدار ضریب کاپا ۰/۸۴ محاسبه شد. همچنین، دقت کلی ۹۵ درصد برای نتایج تحقیق به دست آمده است که البته این پارامتر در برگرفته پیش‌بینی عدم وقوع سیلاب نیز می‌شود و با توجه به فراوانی کمتر موارد عدم وقوع سیلاب نسبت به وقوع سیلاب، دقت کلی افزایش یافته است.

در نهایت برای ارزیابی قوانین به‌دست‌آمده از داده‌های هیدرولوژی سال ۱۳۸۸ استفاده شد. بررسی این داده‌ها بیانگر وقوع ۲۹ مورد دبی بیش از ۱۰ متر مکعب بر ثانیه است. نتایج گسسته‌سازی داده‌های وقوع سیلاب سال ۱۳۸۸ بر اساس مجموعه ویژگی‌های A1، A2 و A3 و مقایسه قوانین به‌دست‌آمده با سیستم اطلاعات ۱۳۸۸، بیانگر وقوع ۲۱ مورد از ۲۹ مورد سیلاب بر اساس قوانین استخراج‌شده برای کلاس ۱ است. نتایج این مقایسه در جدول ۴ ارائه شده است. در این میان ۱۵ مورد بر اساس قانون ۱ است که در واقع قانون کلی و دربرگیرنده قوانین ۴ و ۵ است. ۴ مورد مطابق قانون ۳ بوده است، بدین معنا که ویژگی بارندگی معادل صفر و دو ویژگی دیگر مساعد وقوع سیلاب‌اند. دلیل این موارد را می‌توان در بارندگی بالادست اشاره کرد که از طریق دستگاه‌های مستقر در سد اندازه‌گیری نشده‌اند. ۲ مورد نیز مطابق قانون ۶ بوده است و دلیل آن را می‌توان در مقادیر بارندگی و تبخیر نزدیک به مرز گسسته‌سازی جستجو کرد. البته برای حل این مشکل می‌توان از سیستم استنتاج فازی استفاده کرد که با توجه به کم‌بودن وقوع این المان در سیستم اطلاعاتی از انجام آن صرف‌نظر شده است. همان‌طور که گفته شد نتایج بیانگر

## جدول ۵. ماتریس آشفستگی نتایج پژوهش

| مقادیر واقعی |                |            |                |
|--------------|----------------|------------|----------------|
| جمع ردیف     | عدم وقوع سیلاب | وقوع سیلاب |                |
| ۳۱           | ۱۰             | ۲۱         | وقوع سیلاب     |
| ۳۳۴          | ۳۲۶            | ۸          | عدم وقوع سیلاب |
| ۳۶۵          | ۳۳۶            | ۲۹         | جمع ستون       |
|              |                | دقت کلی    | %۹۵            |

مقادیر پیش‌بینی شده

## ۶. نتیجه‌گیری و پیشنهادها

استخراج‌شده با موارد وقوع یا عدم وقوع سیلاب را ارائه می‌دهد. از دیگر نتایج این ماتریس می‌توان به خطای پیش‌بینی وقوع سیلاب اشاره کرد که برابر با ۹ درصد از کل موارد پیش‌بینی وقوع سیلاب است که نسبت به دقت بالای پیش‌بینی قابل چشم‌پوشی است. این موضوع، نشان‌دهنده توانایی نظریه مجموعه‌زبر در استخراج قوانین پیش‌بینی وقوع پدیده‌های طبیعی است که پیچیدگی‌های زیادی دارند. از نتایج این پژوهش می‌توان در برنامه‌ریزی مدیریت بحران و مهار سوانح طبیعی بهره جست.

## یادداشت‌ها

1. Rough sets Theory
2. Information System
3. Indiscernibility Relation
4. Approximation Of Sets
5. Lower Approximation
6. Upper Approximation
7. P-Boundary
8. Reduction of Attributes
9. Decision Table
10. Decision Rules
11. Confusion Matrix

وقوع سیلاب یکی از بحران‌های طبیعی به شمار می‌رود که تهدیدی بسیار جدی برای زیرساخت‌های جوامع است و جبران خسارات ناشی از آن به راحتی ممکن نیست. در بیشتر مطالعات انجام‌شده در این زمینه، با در نظر گرفتن داده‌های پیشین، از الگوی وقوع سیلاب برای پیش‌بینی استفاده شده است. نکته درخور توجه در این پژوهش‌ها لحاظ‌نکردن عدم قطعیت موجود در حجم بالای داده‌های استفاده‌شده برای شناسایی الگوی وقوع سیلاب است. این در حالی است که هدف اصلی این پژوهش در نظر گرفتن عدم قطعیت داده‌ها و حذف آن‌ها از پروسه استخراج قوانین و بررسی ارتباط پارامترها با وقوع سیلاب به طور همزمان با استفاده از نظریه مجموعه‌های زبر است. در این تحقیق برای ارزیابی نتایج از داده‌های سال ۱۳۸۸ استفاده شد و نتایج پژوهش بیانگر تطابق ۷۲ درصد از سیلاب‌های ثبت‌شده در سال ۱۳۸۸ با قوانین به دست آمده است. همچنین، نتایج ماتریس آشفستگی مقدار ۰/۸۴ برای ضریب کاپا و دقت کلی ۹۵ درصد در تطبیق قوانین قوی

## منابع

برزگر، ف.، صدیقیان، ا.، و نجم‌وی، م. ۱۳۸۹. «یادداشتی بر زمین‌شناسی محل سد هلیل‌رود (جیرفت)»، مرکز اطلاعات علمی مجازی: دوره ۱، شماره ۱، صص ۳۶-۵۵.

شرکت آب منطقه‌ای استان کرمان. ۱۳۸۹. «گزارش سالیانه آماری شرکت سهامی آب منطقه‌ای کرمان»، کرمان: شرکت سهامی آب منطقه‌ای، تعداد صفحات: ۵۵.

- Bartholmes, J., and Todini, E. 2005. Coupling meteorological and hydrological models for flood forecasting. *Journal of Hydrology and Earth System Sciences*. 9: pp. 333-346.
- Chau, K., Wu, C., and Li, Y. 2005. Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering*. 10: pp. 485-491.
- Engeland, K., and Hisdal, H. 2009. A comparison of low flow estimates in ungauged catchments using regional regression and the HBV model. *Journal of Water Resource Manage*. 23: pp. 2567-2586.
- Eslamian, S., Ghasemzadeh, M., Biabanaki, M., and Talebizadeh, M. 2010. A principal component regression method for estimating low flow index. *Journal of Water Resource Manage*. 24: pp. 2553-2566.
- Kumar, A.P.S., Sudheer, K. P., Jain, S.K., and Agarwal., P. K. 2005. Rainfall runoff modeling using artificial neural networks comparison of network types. *Journal of Hydrological Processes*. 19: pp. 1277-1291.
- Kisi, O. 2008. River flow forecasting and estimation using different neural network techniques. *Journal of Hydrological Processes*. 39: pp. 27-40.
- Kisi, O. 2004. River flow modeling using artificial neural network. *Journal of Hydrologic Engineering*. 9: pp. 60-63.
- Mukerji, A., Chatterjee, C., and Raghuwanshi, N. 2009. Flood forecasting using ANN, neuro-fuzzy, and neuro-GA Models. *Journal of Hydrologic Engineering*. 14: pp. 647-652.
- Mahabir, C., Hicks, F.E., and Robinson Fayek, A. 2007. Transferability of a neuro-fuzzy river ice jam flood forecasting model. *Journal of Hydrological Processes*. 48: pp. 188-201.
- Mutlu, E., Chaubey, I., Hexmoor, H., and Bajwa, S.G. 2008. Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. *Journal of Hydrological Processes*. 22: pp. 5097-5106.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S. 2005. Short-term flood forecasting with a neurofuzzy model. *Journal of Water Resources Research*. 41: 2005.
- Nourani, V., Komasi, M., and Mano, A. 2009. A multivariate ANN-wavelet approach for rainfall-runoff modeling. *Journal of Water Resour Manage*. 23: pp. 2877-2894.
- Pawlak, Z. 1982. Rough sets. *International Journal of Computer and Information Sciences*. 11: pp. 341-356.
- Rezaeianzadeh, M., Tabari, H., Arabi Yazdi, A., Isik, S., and Kalin., L. 2013. Flood flow forecasting using ANN, ANFIS and regression models. *Journal of Neural Computing and Applications*. 25: pp. 25-37.
- Rezaeian Zadeh, M., Amin, S., Khalili D., and Singh, V. P. 2010. Daily outflow prediction by multi-layer perceptron with logistic sigmoid and tangent sigmoid activation functions. *Journal of Water Resour Manag*. 24: pp. 2673-2688.
- Refsgaard, J.C., Havno, K., Ammentorp, H.C., and Verwey, A. 1988. Application of hydrological models for flood forecasting and flood control in India and Bangladesh. *Journal of Advances in Water Resources*. 11: pp. 101-105.
- Rodríguez-Rincón, J. P., Breña-Naranjo, J. A., and Pedrozo-Acuña, A. 2014. Uncertainty propagation in a hydro-meteorological approach: from the cloud to the flood map. conference: hic 2014-11th. international conference on hydroinformatics, at new york, usa august. 11: pp. 7977-8011.
- Shamseldin, A.Y. 2010. Artificial neural network model for river flow forecasting in a developing country. *Journal of Hydroinformatics*. 12: pp. 22-35.
- Stromberg, D. 2007. Natural disasters, economic development, and humanitarian aid. *Journal of Economic Perspectives*. 21: pp. 199-222.
- Sahay, R. R., and Sehgal., V. 2013. Wavelet regression models for predicting flood stages in rivers: a case study in Eastern India. *Journal of Flood Risk Manag*. 6: pp. 146-155.
- Sahay, R. R., and Srivastava, A. 2014. Predicting monsoon floods in rivers embedding wavelet transform, genetic algorithm and neural network. *Journal of Water Resour Manage*. 28: pp. 301-317.

Tokar, A. S., and Johnson, A. 1999. Rainfall-runoff modeling using artificial neural networks. *Journal of Hydrologic Engineering*. 4: pp. 232–239.

Tiwari, M. K., Song, K.Y., Chatterjee, C., and Gupta, M. M. 2013. Improving reliability of river flow forecasting using neural networks, wavelets and self-organizing maps *Journal of Hydrologic Engineering*. 15: pp. 486-502.

Talei, A., Chua, L., Quek, C., and Jansson, P. 2013. Runoff forecasting using a Takagi–Sugeno neuro-fuzzy model with online learning. *Journal of Hydrology*. 488: pp. 17-32.

Yeh, J., and Yang, R. 2014. Application of the Adaptive Neuro-Fuzzy Inference System for Optimal Design of Reinforced Concrete Beams. *Journal of Intelligent Learning Systems and Applications*. 6: pp. 162-175.