

تاریخ پذیرش مقاله: ۳ اردیبهشت ۱۳۸۸

## Uncertainty Analysis of Artificial Neural Networks and Neuro-Fuzzy Models in River Flow Forecasting

A. Farokhnia<sup>1\*</sup> and S. Morid<sup>2</sup>

### Abstract

River flow forecasting in water resources management is of great importance. But, due to the high uncertainty in the factors affecting the rainfall-runoff process, the results are usually problematic. One of the procedures that can alleviate this problem is incorporating uncertainty analysis in forecasted results. Such an analysis has been traditionally used for statistical methods but less attention has been given to the Artificial Neural Networks (ANNs) and the Neuro-Fuzzy (ANFIS) models. These models have gained much popularity in recent years. This research has aimed to analyze the uncertainty of these techniques for 1 to 3 months forecasting of river flow. Sofy-Chay River at Tazekand gauging station in the northwest of Iran is selected as the study site to explore the methodology. The results show that ANFIS overall gave more accurate forecasts and less uncertainty. But, when it comes to high flows, the confidence interval for the two models increases quite obviously and this increases the risk for application of the results.

**Keywords:** River Flow Forecasting, Uncertainty, Artificial Neural Network, Neuro-Fuzzy.

## تحلیل عدم قطعیت مدل‌های شبکه عصبی و نروفازی در پیش‌بینی جریان رودخانه

اشکان فرخ‌نیا<sup>۱\*</sup> و سعید مرید<sup>۲</sup>

### چکیده

پیش‌بینی آورد رودخانه در مدیریت منابع آب از اهمیت فراوانی برخوردار است، اما به دلیل عدم قطعیت بالا در عواملی که فرآیند بارش - رواناب را سبب می‌گردند، همواره با مشکلاتی همراه بوده است. یکی از روش‌هایی که می‌تواند این مشکل را تا حدی کاهش دهد، تحلیل عدم قطعیت پیش‌بینی‌های انجام شده می‌باشد. این تحلیل‌ها در مدل‌های آماری سابقه طولانی دارند، ولی برای مدل‌های شبکه عصبی و نروفازی کمتر مورد استفاده قرار گرفته و این در شرایطی است که در سال‌های اخیر به مراتب توجه بیشتری به این تکنیک‌ها شده است. در مطالعه حاضر عدم قطعیت نتایج مدل‌های شبکه عصبی و نروفازی در پیش‌بینی ۱ تا ۳ ماه آینده جریان رودخانه مورد ارزیابی قرار گرفته و منطقه مطالعاتی رودخانه صوفی‌چای در محل ایستگاه تازه‌کند در نظر گرفته شده است. نتایج نشان می‌دهد که مدل نروفازی از دقت بالاتر در پیش‌بینی و عدم قطعیت کمتری برخوردار است، اما بطور مشخص برای پرابی‌ها در هر دو مدل عدم قطعیت بیشتر شده که خطرپذیری (ریسک) کاربرد نتایج را افزایش می‌دهد.

**کلمات کلیدی:** پیش‌بینی جریان رودخانه، عدم قطعیت، شبکه عصبی مصنوعی، نروفازی.

تاریخ دریافت مقاله: ۱۹ شهریور ۱۳۸۶

1- Ph.D. Student, Dep. of Water Structures Eng., Tarbiat Modares University, Tehran, Iran, Email: ashkan\_farokhnia@yahoo.com  
2- Associate Professor, Dep. of Water Structures Eng., Tarbiat Modares University, Tehran, Iran, Email: s\_morid@hotmail.com  
\*- Corresponding Author

۱- دانشجوی دکتری سازه‌های آبی، دانشگاه تربیت مدرس  
۲- دانشیار گروه سازه‌های آبی، دانشگاه تربیت مدرس  
\*- نویسنده مسئول

Coulibaly et al. (2000) میزان جریان روزانه ورودی به سدی در کانادا را با استفاده از شبکه‌های عصبی مختلفی پیش‌بینی و نتایج را با هم مقایسه کردند. در این تحقیق، آن‌ها برای جلوگیری از مشکل فوق‌برازشی<sup>۶</sup> از روش توقف زود هنگام آموزش<sup>۷</sup> استفاده نمودند. نتایج این تحقیق نشان دادند که شبکه‌های عصبی بازگشتی<sup>۸</sup> نسبت به سایر مدل‌های ANNs عملکرد بهتری داشته‌اند. Dibikey & Solomatine, 2001 مدل شبکه عصبی انتشار برگشتی و پایه شعاعی را با مدل مفهومی تانک<sup>۹</sup> به منظور پیش‌بینی مقدار جریان رودخانه‌ای در ونزوئلا مورد مقایسه قرار دادند که نتایج حاکی از عملکرد بهتر مدل‌های شبکه عصبی نسبت به مدل مفهومی مذکور می‌باشد و همچنین عملکرد شبکه عصبی انتشار برگشتی نسبت به شبکه عصبی پایه شعاعی مطلوب‌تر گزارش شده است. Kumar et al. (2004) دو مدل شبکه عصبی پیش‌خورد<sup>۱۰</sup> و شبکه عصبی برگشتی را در پیش‌بینی جریان ماهانه رودخانه‌ای در هند به صورت ۱ و ۱۲ ماه آینده مورد مقایسه قرار داده و نتایج مدل شبکه عصبی برگشتی را در هر دو مورد مناسب‌تر گزارش نمودند. Valenca et al. (2005) شبکه عصبی پرسپترون چند لایه و مدل PARMA<sup>۱۱</sup> را در پیش‌بینی جریان ماهانه رودخانه از ۱ تا ۱۲ ماه بعد مورد مقایسه قرار دادند که نتایج این مقایسه حاکی از حداقل ۵۰٪ کاهش خطا در پیش‌بینی‌های ANN نسبت به PARMA می‌باشد.

در مقایسه با شبکه عصبی، مدل نروفازی برای پیش‌بینی جریان رودخانه، کمتر مورد توجه محققان منابع آب قرار گرفته است. Nayak et al. (2004) برای پیش‌بینی جریان روزانه رودخانه‌ای در هند، از مدل نروفازی استفاده نمودند. آنها فقط از مقادیر آبدهی در روزهای قبل به عنوان ورودی مدل استفاده نموده و نتایج حاصله را با شبکه عصبی مورد مقایسه قرار دادند و عملکرد مدل نروفازی را کمی بهتر گزارش نمودند و نشان دادند که مدل ساخته شده در شرایطی که داده‌ها نرمال شده باشد، عملکرد بهتری نسبت به استفاده از داده‌های خام دارد. Aqil et al. (2007b) مدل‌های نروفازی و رگرسیون خطی را در پیش‌بینی جریان هفتگی رودخانه‌ای در اندونزی مورد مقایسه قرار دادند که عملکرد مدل نروفازی بسیار مناسب‌تر گزارش شد.

یکی از موارد مهمی که بر خلاف مدل‌های آماری در مدل‌های شبکه عصبی و نروفازی به آنها کمتر پرداخته شده است، تحلیل عدم قطعیت در هنگام ارائه نتایج می‌باشد. بدیهی است که پیش‌بینی‌ها قطعی نیستند و این تحلیل می‌تواند استفاده از نتایج مدل‌های

رشد جمعیت و افزایش تقاضای آب در بخش‌های مختلف شرب، کشاورزی و صنعت، نیاز به برنامه‌ریزی و مدیریت منابع آب را بیش از پیش ضروری نموده است. یکی از مهم‌ترین اقدامات در این راستا، پیش‌بینی مقادیر جریان رودخانه‌ها می‌باشد که این امر به علت پیچیدگی فرآیند و دخالت عوامل غیر قطعی در آن، مشکلات خاص خود را دارد. از این رو در دهه‌های اخیر، توجه زیادی به مدل‌هایی شده است که قابلیت ارتباط مستقیم بین داده‌های ورودی و خارجی بدون نیاز به درک کامل از فرآیند را فراهم آورند. این مدل‌ها که اصطلاحاً مدل‌های داده‌محور<sup>۱</sup> نامیده می‌شوند برای برآورد مقدار جریان زمانی که هدف اصلی پیش‌بینی با دقت بالا می‌باشد، بسیار سودمند هستند (Nayak et al, 2005).

دو مدل داده‌محوری که اخیراً توجه بسیاری را به خود معطوف نموده‌اند، شبکه‌های عصبی مصنوعی<sup>۲</sup> (ANNs) و سیستم نروفازی<sup>۳</sup> می‌باشند. این مدل‌ها مزایای خاص خود را دارند که از آن جمله می‌توان به توانایی شبیه‌سازی سیستم‌های غیرخطی، دقت بالا و زمان کمتر ساخت مدل و انجام محاسبات آن اشاره نمود (Aqil et al, 2007a; Nayak et al, 2004).

شبکه‌های عصبی مصنوعی بر مبنای ایجاد سامانه‌ای که توانایی شبیه‌سازی محاسبات مغز را دارا باشد، به وجود آمده‌اند. این ساختار شبکه شامل گره‌هایی است که به یکدیگر متصل شده‌اند. هر گره، یک واحد پردازش می‌باشد و اتصال بین گره‌ها بیانگر نحوه ارتباط آن‌ها می‌باشد (Jang et al, 1997). همچنین در روش نروفازی، از ترکیب منطق فازی (Zadeh, 1965) و شبکه‌های عصبی استفاده می‌شود که در آن از سامانه استنتاج فازی<sup>۴</sup> برای مدل کردن پدیده، و از قابلیت یادگیری شبکه عصبی به منظور بهینه‌سازی فراسنجه‌های آن استفاده می‌گردد (Jang and Gulley, 1995). مرسوم‌ترین دیدگاه توسعه مدل نروفازی، سامانه استنتاجی نروفازی تطبیقی<sup>۵</sup> (ANFIS) می‌باشد که نتایج خوبی را در مدل‌سازی فرآیندهای غیرخطی نشان داده است (Aqil et al, 2007b). ANFIS مشخصات سامانه را با توجه به داده‌های موجود آموزش دیده و فراسنجه‌های مربوطه را با توجه به معیار خطای مورد نظر تنظیم می‌کند.

تحقیقات زیادی در زمینه استفاده از ANNs برای پیش‌بینی جریان رودخانه صورت گرفته که نتایج خوبی از آن‌ها گزارش شده‌اند.

$$net_i = b_i + \sum_{j=1}^n w_{ij} x_j \quad (1)$$

در ادامه مقادیر ورودی در تابع عضویت نرون قرار گرفته و خروجی به لایه بعد انتقال می‌یابد. معمول‌ترین تابع عضویت مورد استفاده در ANNs تابع سیگموئید می‌باشد که به شکل زیر تعریف می‌شود (Maier and Dandy, 2000):

$$f(net_i) = \frac{1}{1 + e^{-net_i}} \quad (2)$$

با توجه به حجم مقاله و وجود مراجع کافی برای ANNs (ASCE, 2000a,b; Maier and Dandy, 2000)، برای تعریف آن به همین اندازه بسنده می‌شود.

## ۲-۲- مدل نروفازی

یکی از روش‌هایی که اخیراً، در هیدرولوژی به آن توجه شده، مدل‌سازی بر مبنای قواعد فازی است. منطق فازی و تئوری مجموعه‌های فازی جهت توصیف تفکر و استدلال بشری در یک چارچوب ریاضی به کار می‌روند. مدل‌هایی که بر مبنای قواعد فازی هستند، نتایج را بطور کیفی شرح می‌دهند.

به مدل‌سازی فازی، سیستم استنتاج فازی (FIS) اطلاق می‌شود. ساختمان اولیه FIS شامل سه مولفه می‌باشد: الف) پایگاه قانون<sup>۱۶</sup> که شامل مجموعه‌ای از قوانین فازی است، ب) پایگاه داده<sup>۱۷</sup> که توابع عضویت<sup>۱۸</sup> (MFs) استفاده شده در قوانین فازی را تعریف می‌کند و ج) مکانیزم استدلال که براساس قوانین، الگوی ورودی را به خروجی مربوطه مرتبط می‌سازد.

FIS با استفاده از تعدادی قوانین اگر-آنگاه (IF-Then) یک رابطه جزء به جزء غیرخطی را از فضای ورودی به فضای خروجی شرح می‌دهد. ترکیبات مختلف توابع عضویت متغیرهای ورودی و خروجی قوانین را به وجود می‌آورند و این قوانین یک ناحیه فازی را از فضای ورودی تعریف کرده و در نهایت رابطه خروجی، خروجی مدل را تعیین می‌کند. کارایی FIS بستگی به برآورد فراسنجه‌های آن دارد که شامل فراسنجه‌های توابع عضویت و تابع خروجی هر قانون می‌باشد.

در مدل‌های نروفازی، برای حل مشکل شناسایی فراسنجه‌های FIS، از یک شبکه تطبیقی<sup>۱۹</sup> که حالت عمومی شبکه عصبی پیشرو چند لایه می‌باشد، استفاده می‌شود. یک شبکه تطبیقی، یک ساختار پیشرو چندلایه است که رفتار کلی خروجی آن بوسیله مقدار یک مجموعه

پیش‌بینی را کارآمدتر سازد. بررسی مراجع نشان داد که روش‌های محدودی برای تعیین عدم قطعیت در این مدل‌ها پیشنهاد شده که از آن جمله می‌توان به روش‌های Bootstrap و تخمین‌زننده ساندویچ<sup>۲۰</sup> (Tibshirani, 1995)، حداکثر درست‌نمایی<sup>۲۱</sup> و استنتاج بیزی<sup>۲۲</sup> (Dybowski, 1997) و روش مونت کارلو پیشنهاد شده توسط Marce et al. (2004) اشاره نمود. با بررسی این منابع، روش Marce et al. (2004) به دلیل جدید بودن و عملکرد مناسب گزارش شده از آن، در این تحقیق مورد توجه قرار گرفت. این روش بر مبنای قرار گیری مدل‌ها در قالب یک پروسه نمونه‌گیری تصادفی مونت کارلو می‌باشد که اولین بار برای برآورد میزان خروج سالانه مواد معدنی مغذی محلول در جریان رودخانه از حوضه، با استفاده از مدل نروفازی به کار گرفته شده است. Aqil et al. (2007b) نیز با استفاده از این روش، عدم قطعیت مقادیر خروجی مدل نروفازی را در پیش‌بینی جریان هفتگی رودخانه مورد بررسی قرار داده و آن را برای مدل نروفازی، مناسب گزارش نمودند.

## ۲- مواد و روش‌ها

### ۲-۱- شبکه‌های عصبی مصنوعی

مدل‌های شبکه عصبی قادرند رابطه میان ورودی‌ها و خروجی‌های یک سیستم را توسط شبکه‌ای از گره‌ها که همگی به هم متصل هستند، تعیین نمایند که در آن، میزان فعالیت هر یک از این اتصالات توسط اطلاعات تاریخی ثبت شده تنظیم می‌شود (فرآیند یادگیری) و در نهایت مدل قادر خواهد بود قوانین مرتبط میان ورودی‌ها و خروجی‌ها را کشف نماید، هرچند این قوانین غیرخطی و پیچیده باشند. در ANNs، نرون‌ها در دسته‌هایی که لایه نام دارند، مرتب می‌شوند که تعداد این لایه‌ها و تعداد نرون موجود در هر یک از لایه‌ها معماری شبکه را تعریف می‌کند. معماری معمول شبکه عصبی متشکل از سه لایه است: لایه ورودی، لایه پنهان و لایه خروجی. طراحی معماری ANNs عموماً بر مبنای روش سعی و خطا<sup>۲۳</sup> بنا شده است و طی آن با استفاده از ارقام مختلف لایه‌های پنهان و نرون‌های مربوط، می‌توان ابعاد مناسب شبکه را تعیین نمود. در ساختار ANNs اطلاعات رسیده از نرون‌های قبلی ( $x_1$  تا  $x_j$ ) به نرون  $i$ ام در وزن‌های مربوط به هر کدام ( $w_{ij}$  تا  $w_{il}$ ) ضرب و سپس جمع می‌شوند. برای جلوگیری از صفر شدن مقدار ورودی نرون، یک مقدار آستانه نیز به مقدار قبلی اضافه گردیده ( $b_i$ ) و بدین ترتیب مقدار ورودی نرون ( $net_i$ ) تعیین می‌گردد. وزن‌ها مبین شدت اتصال نرون‌ها به هم می‌باشند و طی فرآیند یادگیری اوزان و آستانه بهینه می‌گردند.

که در آن  $A_1, A_2, B_1, B_2$  به ترتیب توابع عضویت برای ورودی‌های  $x, y$  هستند.  $p_1, q_1, r_1, p_2, q_2, r_2$  نیز فراسنجه‌های توابع خروجی برای ۲ قانون تعریف شده می‌باشند. معماری مدل ANFIS مورد بحث در شکل (۲) ارائه شده است. جزئیات کامل الگوریتم‌های تولید و آموزش مدل ANFIS در مرجع (Jang (1993) آمده است.

### ۲-۳- برآورد عدم قطعیت در پیش‌بینی‌ها

همانطور که ذکر شد، یکی از اهداف مهم این تحقیق بررسی و تخمین عدم قطعیت در خروجی مدل‌های شبکه عصبی مصنوعی و نروفازی می‌باشد که تحقیق (Marce et al. (2004 مبنای قرار داده شده است. در این روش مدل‌های شبکه عصبی مصنوعی و نروفازی در قالب یک نمونه‌گیری مونت کارلو به کار گرفته می‌شوند. اساس کار شبیه‌سازی مونت کارلو بر پایه استفاده از اعداد تصادفی و حصول تابع توزیع احتمالاتی خروجی یک مدل می‌باشد. بدین صورت که متغیرها (ورودی به مدل) با استفاده از تابع توزیع احتمالاتی مربوط به طور تصادفی تولید شده، و سپس خروجی معادل آن‌ها از مدل مورد نظر به دست می‌آید و این عمل بارها تکرار می‌شود. سپس عدم قطعیت خروجی (تابع) از طریق محاسبه شاخص‌های آماری و یا تعیین تابع توزیع احتمالاتی برآورد می‌شود (Echhardt et al, 2003).

از فراسنجه‌های قابل اصلاح تعیین می‌گردد. با استفاده از این شبکه عصبی تطبیقی، مشکل اصلی استفاده از سیستم استنتاج فازی (که همان بدست آوردن قواعد "اگر- آنگاه" فازی و بهینه‌سازی فراسنجه‌های این مدل می‌باشد) برطرف می‌گردد.

در این تحقیق از سامانه استنتاجی فازی - عصبی تطبیقی<sup>۲۰</sup> (ANFIS) که از جمله مدل‌های فازی-عصبی است، استفاده شده است. معمول‌ترین نوع سامانه استنتاج فازی که قابلیت قرارگیری در یک شبکه تطبیقی را دارا می‌باشد سیستم فازی سوگنو (Sugeno and Yasukawa, 1993) می‌باشد که خروجی آن یک رابطه خطی است و فراسنجه‌های آن را می‌توان با ترکیب روش‌های حداقل مربعات خطی<sup>۲۱</sup> و انتشار خطا به عقب<sup>۲۲</sup> بر اساس کاهش گرادینان<sup>۲۳</sup> برآورد نمود.

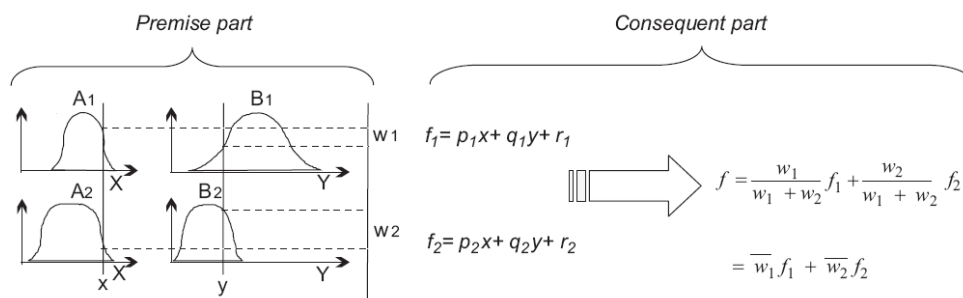
در شکل ۱ نمونه‌ای از سامانه استنتاج فازی سوگنو مرتبه اول که دارای دو ورودی  $x, y$  و خروجی  $z$  می‌باشد، نشان داده شده است. برای FIS نشان داده شده، یک نمونه پایگاه قانون فازی شامل ۲ قانون به شکل زیر قابل ارائه است:

قانون اول: اگر  $x$  مساوی  $A_1$  و  $y$  مساوی  $B_1$  باشد، آنگاه

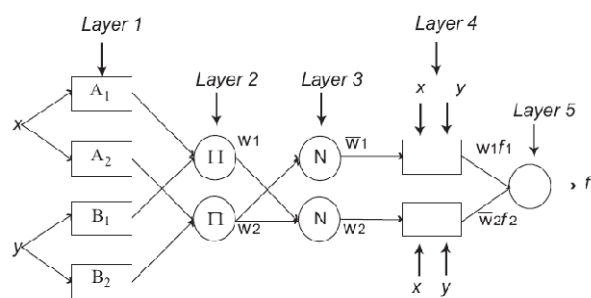
$$f_1 = p_1 x + q_1 y + r_1$$

قانون دوم: اگر  $x$  مساوی  $A_2$  و  $y$  مساوی  $B_2$  باشد، آنگاه

$$f_2 = p_2 x + q_2 y + r_2$$



شکل ۱- سامانه استنتاج فازی سوگنو



شکل ۲- معماری مدل ANFIS معادل سیستم استنتاجی مورد بحث

برای بررسی عدم قطعیت حدود اطمینان ۹۵٪ (۹۵PPU) برای پیش‌بینی‌ها مدنظر قرار گرفت که طبقه به دست آوردن آن تعیین حدود ۲/۵٪ ( $X_L$ ) و ۹۷/۵٪ ( $X_U$ ) توزیع تجمعی احتمال تجربی به دست آمده از ۱۰۰۰ پیش‌بینی انجام شده برای هر ماه می‌باشد. حدود اطمینان مناسب حدودی هستند که بالای ۸۰٪ از آبدهی مشاهده شده در آن حدود قرار گرفته باشند و دارای عرض متوسط قابل

قبولی باشند (Abbaspour et al, 2007). برای سنجش عرض متوسط باند اطمینان شاخص فاکتور عرض باند توسط (Abbaspour et al, 2007) مطابق زیر پیشنهاد شده‌است:

$$d.factor = \frac{\bar{d}_x}{\sigma_x} \quad (7)$$

که در آن  $\sigma_x$  انحراف استاندارد آبدهی مشاهده شده و  $\bar{d}_x$  عرض متوسط باند اطمینان می‌باشد و به صورت زیر تعریف می‌شود:

$$\bar{d}_x = \frac{1}{n} \sum_{i=1}^k (X_U - X_L)_i \quad (8)$$

که مقدار فاکتور عرض باند کوچکتر از ۱ مناسب به حساب می‌آید (Abbaspour et al, 2007). درصد داده‌های قرار گرفته در باند اطمینان ۹۵٪ نیز به صورت زیر محاسبه می‌شود:

$$Bracketed\ by\ 95PPU = \frac{1}{n} count(Q|X_L \leq Q \leq X_U) \times 100 \quad (9)$$

### ۳- ساخت و ارزیابی مدل‌ها

#### ۳-۱- منطقه مطالعاتی

رودخانه صوفی‌چای از رودخانه‌های بزرگ حوضه آبریز دریاچه ارومیه می‌باشد که ورودی اصلی به سد علویان را تشکیل می‌دهد و متوسط آبدهی سالیانه آن ۴/۶ متر مکعب در ثانیه است. این رودخانه از دامنه‌های جنوبی کوه‌های سه‌سهند سرچشمه گرفته و پس از عبور از قسمت غربی شهر مراغه و جنوب شهر بناب به دریاچه ارومیه می‌ریزد. اطلاعات بارندگی، تابش خورشیدی و دمای هوا (به صورت مقادیر حداقل، حداکثر و میانگین) در مقیاس ماهانه، به مدت ۱۸ سال (از اواسط ۱۳۶۲ تا اواخر ۱۳۸۳) از ایستگاه سینوپتیک مراغه و متوسط ماهانه آبدهی رودخانه صوفی‌چای نیز در محل ایستگاه هیدرومتری تازه‌کند تهیه گردید. با توجه به برف‌گیر بودن حوضه این رودخانه، استفاده از کلیه اطلاعات موجود مدنظر قرار گرفت. لذا در این مطالعه، اطلاعات ذکر شده هر کدام با سه تأخیر زمانی<sup>۲۹</sup> که مجموعاً ۱۸ متغیر را تشکیل می‌دهند، به عنوان گزینه‌هایی برای استفاده به عنوان ورودی مدل‌های پیش‌بینی مدنظر قرار گرفتند.

تفاوت روش مورد بحث در این تحقیق با تعریف بالا، استفاده از نمونه‌های تصادفی گرفته شده به جای اعداد تصادفی ساخته شده می‌باشد. این روش بسیار به روش نمونه‌گیری دودویی Bootstrap<sup>۲۴</sup> (Efron and Tibshirani, 1993) شباهت داشته و تنها فرق آن در عدم تکرار یک الگو در یک نمونه می‌باشد. در این روش، ابتدا از اطلاعاتی که برای واسنجی مدل در نظر گرفته شده‌اند، به دفعات زیاد (معمولاً ۱۰۰۰ بار) (Aqil et al, 2007b) بدون تکرار، نمونه‌گیری<sup>۲۵</sup> می‌شود؛ به نحوی که تعداد داده‌های آموزش و نظارت بر آموزش، در هر نمونه یکسان باشد و برای هر نمونه عملیات واسنجی مدل انجام شود. در نتیجه فراسنجه‌های مدل به دفعات زیاد و هر بار بر اساس نمونه‌ای از وقایع اتفاق افتاده در گذشته به دست می‌آیند. انجام این مراحل به تخمین عدم قطعیت در پارامترهای مدل و نتیجتاً خروجی مدل، عدم اریب<sup>۲۶</sup> بودن میانگین خروجی مدل و تعیین حدود اطمینان<sup>۲۷</sup> مقادیر پیش‌بینی شده می‌انجامد (Marce et al, 2004). در نتیجه با استفاده از این روش محاسباتی حجیم به جای پیش‌بینی یک عدد، مطلق مطابق با آنچه در تحقیقات مشابه انجام گرفته، محدوده‌ای از خروجی که نتیجه عدم قطعیت مدل ساخته شده می‌باشد، حاصل خواهد شد.

#### ۲-۴- معیارهای ارزیابی مدل‌ها

چهار شاخص آماری متداول ضریب همبستگی ( $R$ )، میانگین خطای مطلق ( $MAE$ )، ریشه میانگین مربعات خطا ( $RMSE$ ) و درصد میانگین خطای نسبی ( $MAPE$ ) برای سنجش خطای مدل‌ها بکار گرفته شد. این شاخص‌ها به صورت زیر محاسبه می‌شوند:

$$R = \frac{\sum_{i=1}^n (Q_i^o - \bar{Q}^o)(Q_i^e - \bar{Q}^e)}{\sqrt{\sum_{i=1}^n (Q_i^o - \bar{Q}^o)^2} \sqrt{\sum_{i=1}^n (Q_i^e - \bar{Q}^e)^2}} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Q_i^o - Q_i^e| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_i^o - Q_i^e)^2} \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Q_i^o - Q_i^e}{Q_i^o} \right| \times 100 \quad (6)$$

که در این روابط  $Q_i^o$  و  $Q_i^e$  به ترتیب مقادیر مشاهداتی و خروجی آبدهی در ماه  $i$ ام،  $\bar{Q}^o$  و  $\bar{Q}^e$  به ترتیب میانگین مقادیر مشاهداتی و خروجی آبدهی و  $n$  تعداد داده‌های مورد استفاده می‌باشد.

### ۳-۲- واسنجی مدل شبکه عصبی

اولین مرحله و در عین حال از مهمترین مراحل در ساخت مدل‌های ANNs، انتخاب ورودی‌های مؤثر از بین اطلاعات موجود می‌باشد (Zealand et al, 1999). انتخاب ورودی‌های مناسب به دلایل زیر مهم می‌باشد (Bowden et al, 2005): الف) با افزایش تعداد متغیرهای ورودی، پیچیدگی محاسباتی و حافظه مورد نیاز مدل به سرعت افزایش می‌یابد، ب) با ورود متغیرهای زیاد، آموزش مدل مشکل و ناکارآمد می‌شود و ج) ورود متغیرهای پرت سبب افزایش تعداد حداقل‌های موضعی<sup>۳۰</sup> در سطح خطا گردیده که این امر می‌تواند باعث عدم همگرایی<sup>۳۱</sup> و در نتیجه دقت پایین مدل گردد. برای انتخاب متغیرهای مؤثر تاکنون روش‌های زیادی ارائه گردیده است که می‌توان خلاصه‌ای از این روش‌ها را در کار Bowden et al. (2005) یافت نمود. در این تحقیق از متداول‌ترین روش، یعنی سعی و خطا استفاده شده است. بدین منظور ترکیبات مختلفی از متغیرهای مورد نظر به عنوان ورودی مدل در نظر گرفته می‌شود و ترکیبی از متغیرها که بهترین خروجی را داشته باشد، انتخاب می‌گردد. همچنین در انتخاب ورودی‌ها لازم است که توجیه فیزیکی ارتباط متغیرهای ورودی با متغیرهای خروجی، مد نظر باشد.

نظر به اینکه متغیر پیش‌بینی شونده میزان آبدهی می‌باشد، اثر میزان آبدهی ماه‌های گذشته در آن مشهود است و برای تصمیم‌گیری در مورد تعداد تأخیرهای زمانی مورد استفاده از تحلیل خودهمبستگی و خودهمبستگی جزئی استفاده شد. با توجه به شکل ۳، مشخص می‌شود که آبدهی با دو تأخیر زمانی، پارامتر ورودی مناسبی برای پیش‌بینی آبدهی ماه‌های آینده می‌باشد.

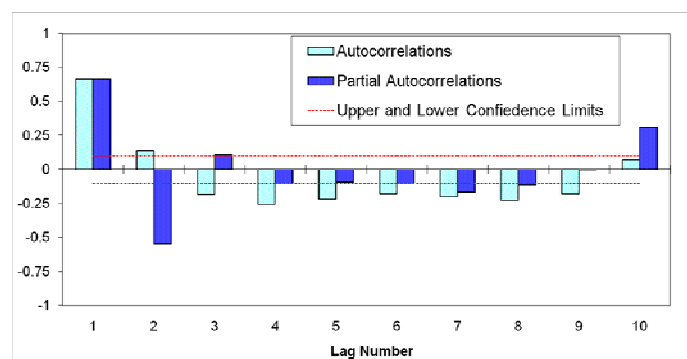
برای انتخاب بهترین فراسنجه ورودی به مدل، به همراه آبدهی با دو تأخیر زمانی، هر کدام از ۵ متغیر موجود با تأخیرهای زمانی مختلف به همراه دبی با دو تأخیر زمانی در مدل شبکه عصبی قرار داده شد

که خلاصه‌ای از نتایج به دست آمده در جدول ۱ آمده است. بر مبنای جدول ۱، مشاهده شد که دمای حداکثر با دو تأخیر زمانی تقریباً بهترین متغیری است که به همراه آبدهی می‌تواند برای ساخت مدل پیش‌بینی آبدهی یک تا سه ماه آینده به کار گرفته شود. این ترکیب منطقی بوده و می‌توان آن را ناشی از رژیم برقی رودخانه صوفی‌چای دانست. همچنین با توجه به اینکه ارتباط بین این متغیرها با آبدهی ماه‌های آینده توسط شبکه عصبی نشان داده شد، از همین ترکیب برای مدل نروفازی نیز استفاده شد. بدین ترتیب مدل پیش‌بینی مطابق رابطه (۱۰) پیشنهاد گردید:

$$Q(t+n) = f[Q(t), Q(t-1), T_{max}(t), T_{max}(t-1)] \quad (10)$$

که در آن  $n$  نشان دهنده زمان پیش‌بینی می‌باشد. به عنوان مثال، اگر  $n$  برابر یک باشد، مدل، پیش‌بینی ۱ ماه بعد را انجام می‌دهد.

برای واسنجی شبکه عصبی، ۴ سال اول داده‌ها برای نظارت بر آموزش، ۱۰ سال میانی برای آموزش و ۴ سال آخر برای صحت سنجی مورد استفاده قرار گرفت. به منظور اطمینان از توجه یکسان به متغیرهای مختلف در زمان آموزش شبکه عصبی، داده‌ها قبل از ورود به مدل طبق پیشنهاد Sanjikumar et al. (1999) در محدوده ۰/۱ تا ۰/۹ استاندارد شدند. Hornik et al. (1989) و Cybenko (1989) نشان دادند که شبکه عصبی با یک لایه پنهان (به شرط داشتن تعداد کافی نرون) توانایی تقریب هر تابع پیچیده و غیرخطی را داراست و افزایش تعداد لایه‌های پنهان تنها باعث پیچیدگی غیرضروری مدل و مشکلات ناشی از آن می‌شود. بنابراین در این تحقیق شبکه عصبی مصنوعی با ساختار یک لایه پنهان مدنظر قرار گرفت. برای تعیین تعداد نرون در لایه پنهان با در نظر داشتن توصیه‌های انجام شده توسط Tang & Fishwick (1993)، با توجه به وجود ۴ متغیر ورودی، تعداد نرون از ۲ تا ۸ مورد استفاده قرار گرفت که نتایج آن در جدول ۲ آمده است.



شکل ۳- نمودار خودهمبستگی و خودهمبستگی جزئی آبدهی رودخانه صوفی‌چای

Cybenko (1989) نشان دادند که شبکه عصبی با یک لایه پنهان (به شرط داشتن تعداد کافی نرون) توانایی تقریب هر تابع پیچیده و غیرخطی را داراست و افزایش تعداد لایه‌های پنهان تنها باعث پیچیدگی غیرضروری مدل و مشکلات ناشی از آن می‌شود. بنابراین در این تحقیق شبکه عصبی مصنوعی با ساختار یک لایه پنهان مدنظر قرار گرفت. برای تعیین تعداد نرون در لایه پنهان با در نظر داشتن توصیه‌های انجام شده توسط Tang & Fishwick (1993)، با توجه به وجود ۴ متغیر ورودی، تعداد نرون از ۲ تا ۸ مورد استفاده قرار گرفت که نتایج آن در جدول ۲ آمده است.

به این ترتیب ساختار مناسب ANN بر پایه عملکرد مدل در دوره صحت‌سنجی تعیین گردید که بر این اساس تعداد نرون در لایه پنهان برای مدل پیش‌بینی ۱ ماه آینده برابر ۴ و برای مدل پیش‌بینی ۲ و ۳ ماه بعد، برابر ۶ انتخاب شد.

برای واسنجی شبکه عصبی، ۴ سال اول داده‌ها برای نظارت بر آموزش، ۱۰ سال میانی برای آموزش و ۴ سال آخر برای صحت‌سنجی مورد استفاده قرار گرفت. به منظور اطمینان از توجه یکسان به متغیرهای مختلف در زمان آموزش شبکه عصبی، داده‌ها قبل از ورود به مدل، طبق پیشنهاد Sanjikumar et al. (1999) در محدوده ۰/۱ تا ۰/۹ استاندارد شدند. (Hornik et al. (1989)

جدول ۱- عملکرد شبکه عصبی به ازای ترکیب مختلف فراسنجه‌های ورودی به همراه دو تاخیر زمانی آبدهی برای پیش‌بینی ۱ تا ۳ ماه بعد آبدهی رودخانه صوفی چای

شماره تکرار	متغیرهای ورودی	پیش‌بینی ۱ ماه بعد		پیش‌بینی ۲ ماه بعد		پیش‌بینی ۳ ماه بعد	
		RMSE	R	RMSE	R	RMSE	R
1	$P(t-2), P(t-1), P(t)$	2.25	0.89	2.52	0.85	2.57	0.85
2	$P(t-1), P(t)$	2.31	0.88	2.61	0.84	2.59	0.85
3	$P(t)$	2.42	0.87	2.70	0.84	2.74	0.83
4	$Rad(t-2), Rad(t-1), Rad(t)$	1.98	0.93	2.21	0.90	2.28	0.89
5	$Rad(t-1), Rad(t)$	1.98	0.93	2.25	0.89	2.31	0.89
6	$Rad(t)$	1.96	0.93	2.34	0.89	2.37	0.88
7	$T_{min}(t-2), T_{min}(t-1), T_{min}(t)$	2.12	0.89	2.31	0.87	2.41	0.85
8	$T_{min}(t-1), T_{min}(t)$	2.19	0.88	2.30	0.87	2.47	0.84
9	$T_{min}(t)$	2.23	0.88	2.36	0.87	2.54	0.84
10	$T_{max}(t-2), T_{max}(t-1), T_{max}(t)$	1.89	0.94	2.21	0.90	2.29	0.88
11	$T_{max}(t-1), T_{max}(t)$	<u>1.89</u>	<u>0.94</u>	<u>2.20</u>	<u>0.90</u>	<u>2.27</u>	<u>0.89</u>
12	$T_{max}(t)$	1.90	0.94	2.25	0.89	2.31	0.88
13	$T_{mean}(t-2), T_{mean}(t-1), T_{mean}(t)$	1.94	0.93	2.20	0.89	2.29	0.89
14	$T_{mean}(t-1), T_{mean}(t)$	1.93	0.93	2.24	0.88	2.30	0.89
15	$T_{mean}(t)$	1.95	0.93	2.25	0.88	2.39	0.88

جدول ۲- شاخص‌های خطاسنجی شبکه عصبی با ساختارهای متفاوت برای مدل‌های پیش‌بینی ۱ تا ۳ ماه بعد

زمان پیش‌بینی	شاخص خطا	تعداد گره در لایه پنهان							
		2	3	4	5	6	7	8	
۱ ماه	R	0.83	0.85	<b>0.89</b>	0.89	0.88	0.89	0.89	
	MAE	1.26	1.18	<b>1.12</b>	1.13	1.12	1.13	1.15	
	RMSE	2.05	1.97	<b>1.89</b>	1.89	1.89	1.89	1.90	
	MAPE(%)	61.13	57.77	<b>54.24</b>	55.19	53.98	56.31	58.88	
۲ ماه	R	0.73	0.73	0.77	0.77	<b>0.80</b>	0.79	0.80	
	MAE	1.59	1.55	1.42	1.39	<b>1.30</b>	1.32	1.31	
	RMSE	2.64	2.61	2.41	2.34	<b>2.25</b>	2.24	2.25	
	MAPE(%)	77.87	73.23	68.66	67.89	<b>65.24</b>	64.53	65.88	
۳ ماه	R	0.71	0.74	0.75	0.76	<b>0.78</b>	0.77	0.77	
	MAE	1.76	1.68	1.54	1.44	<b>1.39</b>	1.40	1.42	
	RMSE	2.78	2.67	2.51	2.48	<b>2.31</b>	2.32	2.34	
	MAPE(%)	79.33	73.46	67.29	64.18	<b>58.68</b>	55.79	58.89	

$$D_i^{new} = D_i - D_{CI} \exp\left(-\frac{\|x_i - x_{CI}\|^2}{(r_b/2)^2}\right) \quad (12)$$

که در آن  $r_b$  یک عدد مثبت ثابت است که به طور معمول  $1/5$  برابر  $r_a$  در نظر گرفته می‌شود. ورودی این مدل نیز مانند قبل، آبدی و درجه حرارت با دو تأخیر زمانی در نظر گرفته شد، از توابع عضویت گوسین استفاده شد و مقدار مناسب  $r_a$  به روش سعی و خطا تعیین گردید. به این منظور اعداد  $0/1$  تا  $0/9$  با افزایش  $0/05$  مورد استفاده قرار گرفتند و در پایان مقدار مناسب  $r_a$  برای هر مدل پیش‌بینی (۱) تا ۳ ماه آینده) تعیین گردید. مقدار مناسب  $r_a$  برای مدل پیش‌بینی ۱ ماه بعد، برابر  $0/3$  (۷ خوشه)، در مدل پیش‌بینی ۲ ماه بعد، برابر  $0/35$  (۶ خوشه) و در مدل پیش‌بینی ۳ ماه بعد، برابر  $0/25$  (۷ خوشه) به دست آمد. همچنین مرکز خوشه‌های ساخته شده و شعاع اثر هر یک نیز مشخص گردید. سپس با استفاده از آن‌ها توابع عضویت هر متغیر ورودی مشخص می‌شود و از ترکیب توابع عضویت متناظر متغیرهای ورودی و ارتباط آنها با متغیر خروجی توسط یک رابطه خطی، قوانین تعریف می‌شوند. پس از آن، پارامترهای مدل (ضرائب رابطه خطی خروجی هر قانون و پارامترهای توابع عضویت) با قرار گرفتن سیستم استنتاج در بطن یک شبکه تطبیقی<sup>۳۴</sup> تصحیح می‌شوند. تعداد تکرار الگوریتم ترکیبی برای تصحیح پارامترهای مدل، برابر  $30$  و خطای هدف برابر  $10^{-4}$  قرار داده شد. در جدول ۳، شاخص‌های خطای مدل نروفازی در دوره صحت‌سنجی به ازای تعداد مختلف قوانین آورده شده‌اند و برای نمونه، مرکز خوشه‌ها و قوانین ساخته شده برای مدل نروفازی پیش-بینی ۱ ماه بعد (پیش از بهینه‌سازی فراسنجه‌های مدل) در جدول ۴ آورده شده است.

#### ۳-۴- تعیین عدم قطعیت مدل‌ها

در این بخش ۱۴ سال اول آماری برای واسنجی مدل و ۴ سال آخر برای صحت‌سنجی مورد استفاده قرار گرفت. در هر مرحله (از ۱۰۰۰ سال تکرار)، از ۱۴ سال داده انتخاب شده برای واسنجی مدل‌ها، ۱۰ سال به صورت تصادفی برای آموزش و ۴ سال باقیمانده برای نظارت بر آموزش استفاده شد. نظارت بر آموزش در مدل شبکه عصبی به صورت توقف آموزش پس از افزایش خطای اطلاعات نظارتی و در مدل نروفازی به صورت در نظر گرفتن فراسنجه‌های مدل در هنگام حداقل بودن خطای اطلاعات نظارت بر آموزش انجام گردید. حداکثر تعداد تکرار الگوریتم آموزش برای تصحیح فراسنجه‌ها در شبکه عصبی  $300$  و در مدل نروفازی  $30$  انتخاب شد و خطای هدف برای هر دو مدل  $10^{-4}$  قرار داده شد. ساختار بهینه هر یک از مدل‌ها که در قسمت قبل تعیین شده بود در این بخش مورد استفاده قرار گرفت.

به این ترتیب ساختار مناسب ANN بر پایه عملکرد مدل در دوره صحت‌سنجی تعیین گردید که بر این اساس تعداد نرون در لایه پنهان برای مدل پیش‌بینی ۱ ماه آینده برابر ۴ و برای مدل پیش‌بینی ۲ و ۳ ماه بعد، برابر ۶ انتخاب شد.

#### ۳-۳- واسنجی مدل نروفازی

در مدل ANFIS در لایه اول، مقادیر هر متغیر ورودی باید به چند کلاس برای ساختن قوانین مربوط، دسته‌بندی شود که قوانین فازی از ترکیب ۲ یا تعداد بیشتری توابع عضویت در لایه دوم ساخته می‌شوند. برای طبقه‌بندی داده‌های ورودی و ساخت قوانین، روش‌های متعددی پیشنهاد گردیده‌اند که رایج‌ترین آنها عبارتند از: افراز شبکه‌ای<sup>۳۳</sup> (Jang and Sun, 1995) و خوشه‌بندی فازی کاشی<sup>۳۳</sup> (Chiu, 1994). زمانی که تعداد متغیرهای مورد استفاده کم است، افراز شبکه‌ای، یک روش مناسب برای طبقه‌بندی داده‌ها می‌باشد ولی در این تحقیق به علت وجود ۴ متغیر ورودی و نیاز به وجود تعداد قابل توجهی تابع عضویت برای هر کدام نمی‌توان از این روش استفاده کرد. به عنوان مثال اگر ۴ متغیر ورودی و برای هر متغیر، ۴ تابع عضویت داشته باشیم، تعداد قوانین برابر  $4^4$  خواهد بود که برآورد فراسنجه‌های این مدل با توجه به تعداد داده‌های موجود و مشکلات محاسباتی میسر نمی‌باشد. لذا در این تحقیق از روش خوشه‌بندی فازی کاشی برای ساخت FIS استفاده شد. این روش بر مبنای اندازه‌گیری تراکم نقاط موجود در فضای متغیرهای مورد استفاده انجام می‌شود. به این منظور، فضای متغیرهای موجود استانداردسازی می‌شود، به نحوی که ابعاد تمامی متغیرها به بازه  $[0-1]$  منتقل شود. در ابتدا هر یک از نقاط به عنوان نقطه‌ای که پتانسیل مرکزیت یک خوشه را دارد در نظر گرفته می‌شود، سپس شاخص تراکم ( $D_i$ ) نقاط موجود در اطراف نقطه  $x_i$  طبق رابطه زیر محاسبه می‌شود:

$$D_i = \sum_{j=1}^n \exp\left(-\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right) \quad (11)$$

که در آن  $r_a$  یک عدد مثبت ثابت است. اگر یک نقطه دارای تعداد زیادی نقاط همسایگی باشد، شاخص تراکم بالایی خواهد داشت. پس از محاسبه شاخص تراکم برای هر نقطه، نقطه‌ای که دارای بالاترین شاخص تراکم می‌باشد، به عنوان مرکز اولین خوشه انتخاب می‌شود. برای تعیین خوشه‌های بعدی، اگر نقطه  $x_{CI}$  با شاخص تراکم  $D_{CI}$  به عنوان مرکز خوشه انتخاب شده باشد، شاخص تراکم هر یک از نقاط باقیمانده برای انتخاب مرکز خوشه بعدی طبق رابطه زیر تصحیح می‌شود (Yager and Filev, 1994):



جدول ۳- شاخص‌های خطاسنجی مدل نروفازی با ساختارهای متفاوت برای مدل‌های پیش‌بینی ۱ تا ۳ ماه بعد

زمان پیش‌بینی	شاخص خطا	تعداد قوانین						
		2	3	4	5	6	7	8
۱ ماه بعد	R	0.67	0.72	0.79	0.85	0.84	<b>0.88</b>	0.81
	MAE	1.42	1.33	1.23	1.14	1.13	<b>1.11</b>	1.17
	RMSE	2.16	2.04	1.99	1.95	1.93	<b>1.93</b>	1.99
	MAPE	58.34	54.12	46.25	45.85	41.57	<b>38.22</b>	41.74
۲ ماه بعد	R	0.69	0.74	0.78	0.80	<b>0.82</b>	0.78	0.70
	MAE	1.57	1.46	1.32	1.24	<b>1.17</b>	1.27	1.44
	RMSE	2.61	2.48	2.25	2.11	<b>2.01</b>	2.14	2.37
	MAPE	77.87	73.23	68.66	67.89	<b>44.57</b>	64.53	65.88
۳ ماه بعد	R	0.64	0.69	0.72	0.74	0.75	<b>0.79</b>	0.78
	MAE	1.81	1.67	1.52	1.43	1.40	<b>1.37</b>	1.39
	RMSE	2.87	2.61	2.42	2.37	2.30	<b>2.24</b>	2.28
	MAPE	81.92	67.27	62.48	55.21	53.22	<b>49.71</b>	48.54

جدول ۴- مرکز خوشه‌های و قوانین ساخته شده برای مدل نروفازی پیش‌بینی ۱ ماه بعد

شماره قانون	IF-Part (Cluster)								Then-Part (Output Linear Equation)
	wf1 [ $T_{max}(t-1)$ ]		wf2 [ $T_{max}(t)$ ]		wf3 [ $Q(t-1)$ ]		wf4 [ $Q(t)$ ]		
	Center	Radius	Center	Radius	Center	Radius	Center	Radius	
1	0.165	0.091	0.134	0.091	0.828	0.091	0.803	0.091	output=-0.395wf1+0.147wf2+0.431wf3-2.016wf4+0.126
2	0.154	0.091	0.162	0.091	0.254	0.091	0.228	0.091	output= -0.201wf1+0.357wf2+0.046wf3-0.078wf4+0.131
3	0.144	0.091	0.146	0.091	0.577	0.091	0.388	0.091	output= 1.024wf1+0.395wf2+0.133wf3+0.093wf4-0.163
4	0.116	0.091	0.132	0.091	0.751	0.091	0.581	0.091	output= -0.017wf1-1.107wf2-0.195wf3-0.027wf4 0.458
5	0.16	0.091	0.212	0.091	0.212	0.091	0.357	0.091	output= -1.192wf1+1.099wf2+0.755wf3-3.096wf4+1.762
6	0.473	0.091	0.195	0.091	0.781	0.091	0.841	0.091	output= -0.370wf1+0.568wf2+0.734wf3-1.06wf4+0.557
7	0.144	0.091	0.261	0.091	0.359	0.091	0.555	0.091	output= 0.642wf1-0.158wf2+1.205wf3-2.045wf4+0.165

منظور ارائه متوسط پیش‌بینی جریان در هر ماه و مقایسه بیشتر دو مدل، از ۱۰۰۰ مورد پیش‌بینی انجام شده توسط مدل‌ها برای هر ماه میانگین گرفته شد. شاخص‌های خطاسنجی آبدهی پیش‌بینی شده و مشاهداتی برای میانگین خروجی هر مدل و در هر زمان پیش‌بینی محاسبه شد که در جدول ۶ نشان داده شده‌اند و سری زمانی آبدهی مشاهداتی و میانگین پیش‌بینی شده دو مدل برای یک ماه بعد در شکل‌های ۵ و ۶ آورده شده‌اند.

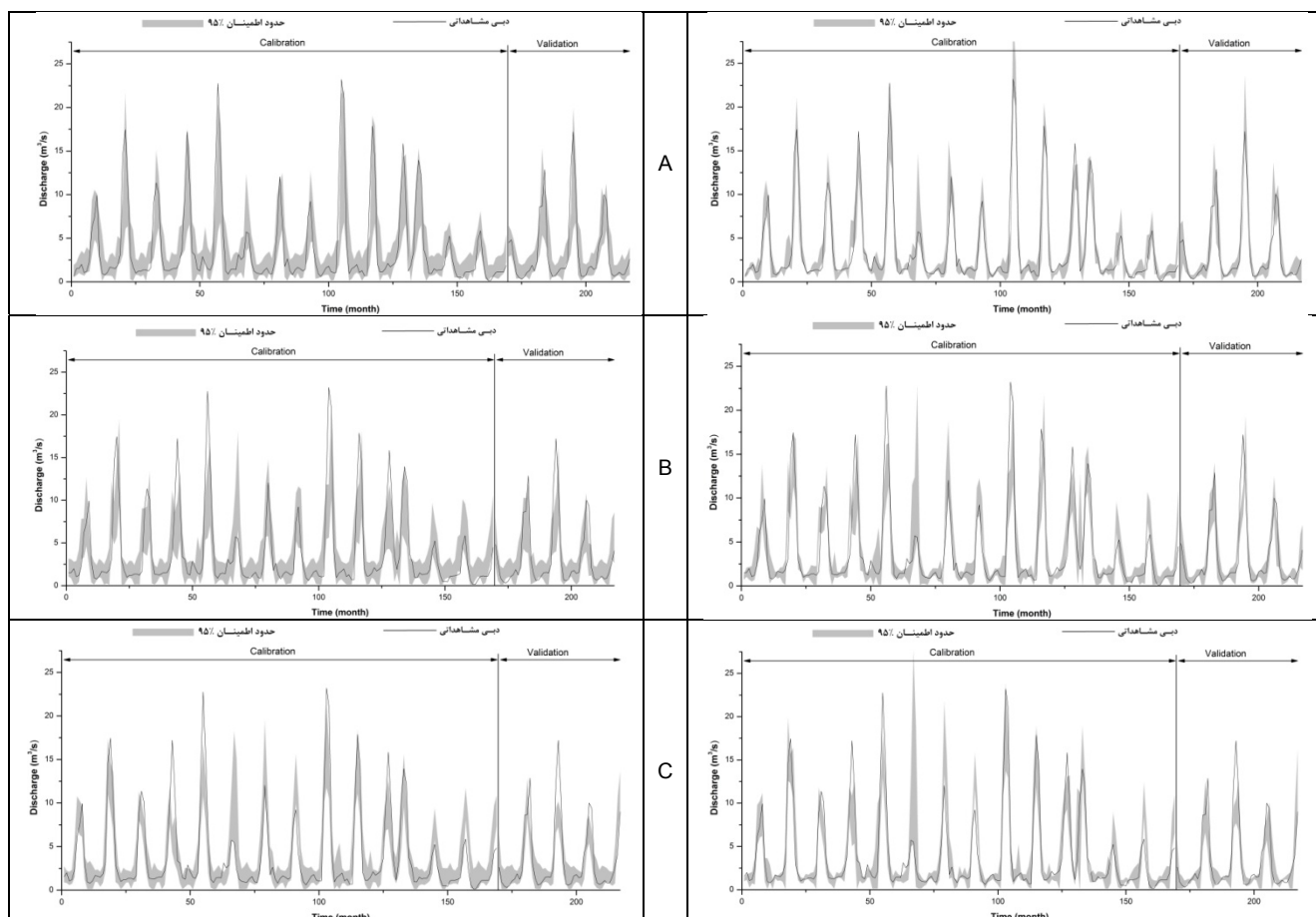
پس از آموزش مدل‌ها و به دست آوردن ۱۰۰۰ مقدار برای پیش‌بینی هر ماه، حدود اطمینان ۹۵٪ برای پیش‌بینی ۱ تا ۳ ماه آینده توسط مدل‌های شبکه عصبی مصنوعی و نروفازی از طریق تعیین مقادیر ۲/۵٪ و ۹۷/۵٪ توزیع احتمالاتی تجربی به دست آمده برای هر ماه مشخص گردید. حدود اطمینان ۹۵٪ مدل‌های پیش‌بینی شبکه عصبی و نروفازی برای ۱ تا ۳ ماه آینده به ترتیب در شکل‌های ۴ و ۵ و شاخص‌های عدم قطعیت در جدول ۵ ارائه شده‌اند. همچنین به

جدول ۵- شاخص‌های عدم قطعیت مدل‌های شبکه عصبی و نروفازی برای پیش‌بینی‌های ۱ تا ۳ ماه بعد

زمان پیش‌بینی	شاخص خطا	ANN		ANFIS	
		واسنجی	صحت‌سنجی	واسنجی	صحت‌سنجی
۱ ماه بعد	d-factor	0.86	0.93	0.48	0.63
	(%) 95PPU Bracketed by	81.06	79.16	73.96	77.08
۲ ماه بعد	d-factor	0.85	0.92	0.59	0.68
	(%) 95PPU Bracketed by	73.96	72.91	70.41	70.83
۳ ماه بعد	d-factor	0.69	0.71	0.57	0.62
	(%) 95PPU Bracketed by	76.33	75.00	71.60	62.50

جدول ۶- شاخص های خطای میانگین ۱۰۰۰ مورد پیش بینی انجام شده و پیش بینی های منفرد توسط مدل های ANN و ANFIS  
 ۱ تا ۳ ماه بعد

زمان پیش بینی	شاخص خطا	نتیجه حاصل از میانگین خروجی ۱۰۰۰ مورد پیش بینی				نتیجه حاصل از مدل منفرد			
		ANN		ANFIS		ANN		ANFIS	
		واستجی	صحت سنجی	واستجی	صحت سنجی	واستجی	صحت سنجی	واستجی	صحت سنجی
۳ ماه بعد	R	0.93	0.9	0.95	0.9	0.91	0.89	0.9	0.88
	MAE	1.02	1.03	0.89	1.03	1.1	1.12	0.99	1.11
	RMSE	1.71	1.68	1.44	1.78	1.77	1.89	1.66	1.93
	MAPE	46.07	49.77	34.46	36.02	49.85	54.24	32.95	38.22
۲ ماه بعد	R	0.82	0.84	0.87	0.88	0.81	0.8	0.87	0.82
	MAE	1.56	1.25	1.35	1.01	1.55	1.3	1.44	1.17
	RMSE	2.67	2.13	2.34	1.86	2.61	2.25	2.51	2.01
	MAPE	63.21	60.98	45.75	39.18	63.76	65.24	50.32	44.57
۱ ماه بعد	R	0.88	0.84	0.91	0.87	0.85	0.78	0.83	0.79
	MAE	1.29	1.3	1.1	1.2	1.34	1.39	1.19	1.37
	RMSE	2.23	2.18	1.9	2.01	2.32	2.31	2.04	2.24
	MAPE	47.73	53.21	33.71	45.59	55.76	58.68	40.42	49.71



شکل ۴- حدود اطمینان ۹۵٪ ANN  
 (A) یک ماه بعد، (B) دو ماه بعد و (C) سه ماه بعد

شکل ۵- حدود اطمینان ۹۵٪ ANFIS  
 (A) یک ماه بعد، (B) دو ماه بعد و (C) سه ماه بعد

## ۴-۱- تحلیل و مقایسه عدم قطعیت مدل‌ها

همانطور که در شکل‌های ۴ و ۵ مشاهده می‌شود، در تمام زمان‌های پیش‌بینی شبکه عصبی دارای عدم قطعیت بیشتری در مقادیر خروجی بوده و این امر درحالی بوده که با وجود افزایش قابل توجه عرض محدوده اطمینان ۹۵٪ آن نسبت به نروفازی، تعداد داده‌های مشاهداتی که در این باند قرار گرفته‌اند، افزایش چشمگیری نیز نسبت به نروفازی نشان نمی‌دهد. با توجه به شاخص عرض متوسط باند اطمینان به دست آمده (جدول ۵) مشاهده می‌شود که شبکه عصبی حساسیت بیشتری به داده‌های دسته آموزش دارد، چنانچه حتی در آبدی‌های کم، که توقع وجود عدم قطعیت کمی در پیش‌بینی آنها می‌رود، شبکه عصبی باند اطمینان بزرگ و نامناسبی را برآورد نموده است. مدل نروفازی، محدوده اطمینان کوچکتری را ایجاد نموده است که به طبع تعداد کمتری از داده‌های مشاهداتی را در محدوده اطمینان خود جای داده است، ولی عدم قطعیت این مدل به خصوص در جریان‌های کم، واقع بینانه‌تر است و کاهش درصد تعداد داده‌های مشاهده شده در محدوده اطمینان آن نسبت به شبکه عصبی چندان قابل توجه نمی‌باشد. نکته قابل توجه دیگری که در جدول ۵ مشاهده می‌شود، کاهش عرض محدوده اطمینان ۹۵٪ شبکه عصبی با افزایش افق پیش‌بینی می‌باشد که این امر برخلاف اصول احتمالاتی بوده و نشان می‌دهد که شبکه عصبی در پیش‌بینی‌های بلندمدت دارای ضعف و نوعی بی‌تفاوتی به داده‌ها می‌باشد در حالی که این نقیصه در مدل نروفازی کمتر به چشم می‌آید و این مدل روند منطقی‌تری را در عدم قطعیت پیش‌بینی‌های خود نشان می‌دهد.

شکل ۸ عرض محدوده اطمینان ۹۵٪ دو مدل را نسبت به جریان مشاهده شده ۱ ماه بعد نشان می‌دهد. همانطور که در این شکل نیز مشاهده می‌شود، محدوده اطمینان مدل نروفازی در اکثر مقادیر پیش‌بینی شده ۱ ماه بعد باریک‌تر می‌باشد که نشان‌دهنده قابلیت بیشتر مدل نروفازی و حساسیت کمتر آن نسبت به تغییرات در داده‌های آموزشی در قیاس با شبکه عصبی می‌باشد. محدوده اطمینان هر دو مدل در ماه‌های پرآب عریض‌تر می‌شود که این افزایش به دلیل پیچیدگی بیشتر پدیده در این حالت و تأثیر عوامل دیگری که به علت کمبود اطلاعات یا عدم آگاهی در ساخت مدل‌ها مد نظر قرار نگرفته‌اند، می‌باشد. علاوه بر این و با وجود عدم قطعیت بالای ماه‌هایی که جریان زیاد دارند، همانطور که در اشکال ۴ و ۵ دیده می‌شود، تعداد قابل توجهی از آبدی‌های مشاهداتی این

ماه‌ها در محدوده اطمینان ۹۵٪ قرار نگرفته‌اند که دلیل این امر، کم بودن اطلاعات برای واسنجی مدل در این حالت نسبت به ماه‌های کم جریان می‌باشد که نشان دهنده مشکلات بیشتر مدل برای شناسایی رفتار سامانه در این شرایط می‌باشد.

## ۴-۲- تحلیل و مقایسه نتایج پیش‌بینی مدل‌های منفرد و میانگین خروجی از ۱۰۰۰ مورد شبیه‌سازی

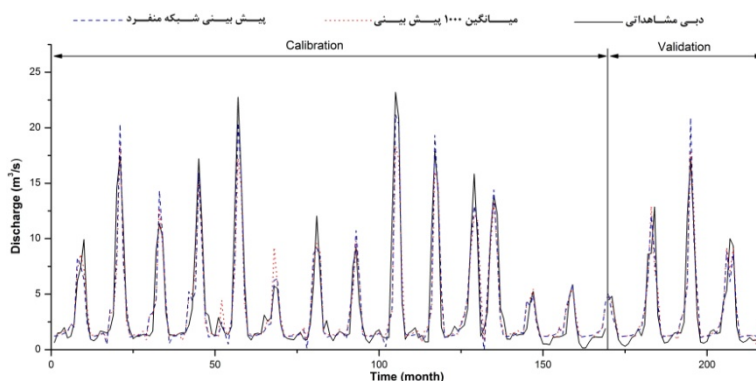
همانطور که در اشکال ۶ و ۷ مشاهده می‌شود، پیش‌بینی‌های انجام شده توسط مدل‌های اولیه و میانگین ۱۰۰۰ تکرار همان مدل‌ها نزدیک به هم می‌باشد که این امر نشان دهنده پایداری هر دو مدل می‌باشد، هر چند با توجه به جدول ۶ می‌توان گفت که پیش‌بینی‌های حاصل از میانگین شبیه‌سازی‌ها نسبت به نتایج مدل‌های منفرد کمی بهتر بوده‌اند که علت این امر می‌تواند کاهش اریب<sup>۳۵</sup> در نتایج خروجی مدل‌ها باشد. بر این اساس در مورد هر یک از مدل‌ها میانگین هر ماه به عنوان پیش‌بینی مناسب انتخاب شد و مقایسه بین نتایج دو مدل در هر افق پیش‌بینی در زیر ارائه می‌گردد:

۱- در پیش‌بینی ۱ ماه بعد، مدل نروفازی در دوره آموزش جریان‌های کم و زیاد را با دقت خوبی پیش‌بینی نموده است که این امر برای جریان‌های کم دوره صحت‌سنجی نیز مشاهده می‌شود، ولی در پیش‌بینی جریان‌های زیاد این دوره چندان خوب عمل نکرده است و آن‌ها را کمی بیش از مقدار واقعی برآورد نموده است. شبکه عصبی نیز کم و بیش عملکرد مشابهی برای این دوره پیش‌بینی نموده و کلاً در پیش‌بینی ۱ ماه بعد نسبت به نروفازی در جریان‌های زیاد کمی بهتر و برای جریان‌های کم، ضعیف‌تر بوده است.

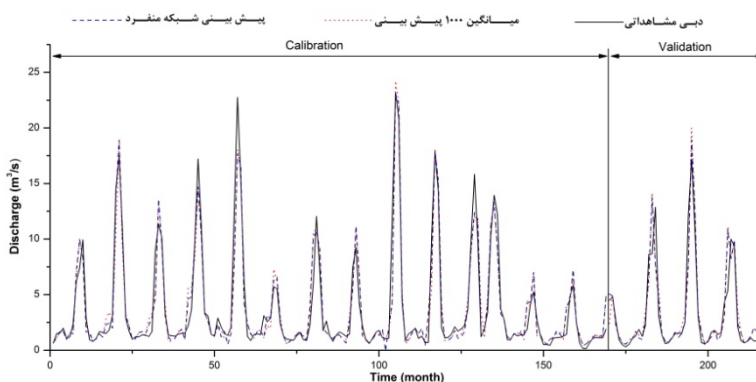
۲- در پیش‌بینی ۲ ماه بعد، مدل نروفازی در دوره آموزش جریان‌های کم را به خوبی پیش‌بینی نموده ولی جریان‌های زیاد را کمتر از مقدار واقعی برآورد نموده و در دوره صحت‌سنجی، هر دو نوع جریان را با دقت خوبی پیش‌بینی نموده است. شبکه عصبی، هم در دوره آموزش و هم در صحت‌سنجی، عملکرد ضعیفی داشته است و جریان‌های کم را بیش از مقدار واقعی و جریان‌های زیاد را کمتر از مقدار واقعی برآورد نموده است. مدل نروفازی مقدار شاخص‌های  $MAE$ ،  $RMSE$  و  $MAPE$  را در دوره صحت‌سنجی به ترتیب ۴/۸٪، ۱۹/۲٪، ۱۲/۷٪ و ۳۵/۷٪ نسبت به شبکه عصبی بهبود بخشیده است.

جریان‌ها را بسیار بهتر و نزدیک‌تر به مقادیر واقعی پیش‌بینی نموده است. در دوره صحت‌سنجی این افق پیش‌بینی، مدل نروفازی در مقایسه با شبکه عصبی در شاخص‌های  $R$ ،  $MAE$ ،  $RMSE$  و  $MAPE$  به ترتیب  $3/5\%$ ،  $7/7\%$ ،  $7/8\%$  و  $14/3\%$  بهبود نشان می‌دهد.

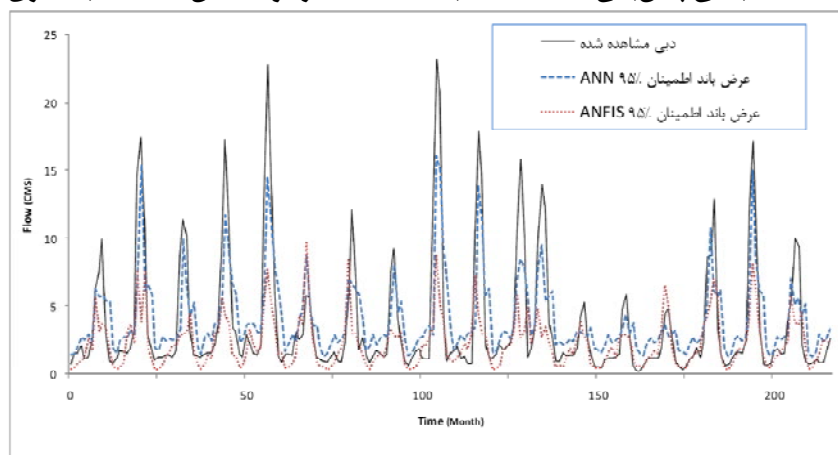
۳- هر دو مدل در پیش‌بینی‌های ۳ ماه بعد، در برآورد جریان‌های زیاد در هر دو دوره آموزش و صحت‌سنجی عملکرد مناسبی نداشته‌اند و این جریان‌ها را به شکل فاحشی کمتر از مقدار واقعی برآورد نموده‌اند که تفاوت برآورد این جریان‌ها در مدل شبکه عصبی بیشتر می‌باشد. همچنین شبکه عصبی جریان‌های کم را نیز بیشتر از مقدار مشاهداتی برآورد نموده است؛ در حالی که مدل نروفازی این



شکل ۶- آبدهی پیش‌بینی شده یک ماه بعد ANN منفرد و میانگین ۱۰۰۰۰ شبیه‌سازی



شکل ۷- آبدهی پیش‌بینی شده یک ماه بعد ANFIS منفرد و میانگین ۱۰۰۰۰ شبیه‌سازی



شکل ۸- عرض محدوده اطمینان ۹۵٪ ANN و ANFIS و مقدار واقعی جریان برای پیش‌بینی ۱ ماه بعد آبدهی رودخانه صوفی چای

## ۵- نتیجه گیری

در این تحقیق دو روش شبکه عصبی مصنوعی و نروفازی برای پیش‌بینی یک تا سه ماه آینده جریان رودخانه صوفی‌چای همراه با لحاظ نمودن عدم قطعیت در محاسبات مربوط، مورد مقایسه قرار گرفت که نتایج زیر برای آن قابل ارائه می‌باشد:

۱- مدل نروفازی در مقایسه با شبکه عصبی دارای عدم قطعیت کمتر در مقادیر خروجی می‌باشد؛ به نحوی که عرض محدوده اطمینان آن در عین داشتن پهنای قابل قبول، در اکثر پیش‌بینی‌ها به طور بهتری عمل کرده است.

۲- نتایج نشان داد که هر دو مدل در ماه‌های پر آب دارای عدم قطعیت بیشتری می‌باشند به نحوی که بعضاً در این ماه‌ها عرض محدوده اطمینان بسیار زیاد بوده و با مقدار مشاهداتی آبدهی برابری می‌کند. این امر می‌تواند به علت پیچیدگی بیشتر پدیده در این ماه‌ها و تاثیر عواملی باشد که به علت کمبود اطلاعات یا عدم آگاهی، در ساخت مدل‌های پیش‌بینی مدنظر قرار نگرفته‌اند.

۳- مقایر پیش‌بینی منفرد مدل‌ها با میانگین ۱۰۰۰ مورد پیش‌بینی آنها مقایسه شد که نشان دهنده برآورد غیراریب‌تر میانگین می‌باشد، اما نتایج تفاوت معنی‌داری را از آنها بیان نداشت.

۴- روش‌شناسی انجام این تحقیق به نحو مطلوبی بحث عدم قطعیت را در مدل‌های ANN و ANFIS وارد نمود و با توجه به نتایج مناسب و منطقی به دست آمده، استفاده از این روش در تحقیقات مشابه توصیه می‌شود.

۵- مطالعات انجام شده در این مقاله می‌تواند مکمل طرح‌های تحقیقاتی پیش‌بینی در صنعت آب کشور باشد. به‌طوریکه برنامه‌ریزان به‌جای مواجهه به یک رقم به عنوان پیش‌بینی، ترکیب این رقم و محدوده اطمینان را در اختیار داشته و می‌توانند تصمیمات واقع‌بینانه‌تری اتخاذ نمایند.

## پی‌نوشت‌ها

- 1- Data-Driven Models
- 2- Artificial Neural Networks
- 3- Neuro-Fuzzy Inference System
- 4- Fuzzy Inference System
- 5- Adaptive Neuro-Fuzzy Inference System
- 6- Overfitting
- 7- Early Stop Training Approach
- 8- Recurrent Neural Network
- 9- Tank Conceptual Model
- 10- Feed-Forward Neural network
- 11- Periodic Auto Regressive Moving Average
- 12- Sandwich Estimator

- 13- Maximum Likelihood
- 14- Bayesian Inference
- 15- Trial and Error
- 16- Rulebase
- 17- Database
- 18- Membership Functions
- 19- Adaptive network
- 20- Adaptive Neuro-Fuzzy Inference System
- 21- Least Square Error
- 22- Backpropagation
- 23- Gradient Descent
- 24- Bootstrap Pairs Resampling
- 25- Resampling Without Replacement
- 26- Unbiased
- 27- Confidence Intervals
- 28- 95 Percent Prediction Uncertainty
- 29- Lag
- 30- Local Minima
- 31- Misconvergence
- 32- Grid Partition
- 33- Subtractive Fuzzy Clustering
- 34- Adaptive Network
- 35- Bias

## ۶- مراجع

- Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. and Srinivasan, R. (2007), "Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT." *Journal of Hydrology*, 333, pp. 413-430.
- Aqil, M., Kita, I., Yano, A. and Nishiyama, S. (2007a), "A comparative study of artificial neural networks and neuro-fuzzy in continuous modeling of the daily and hourly behaviour of runoff." *Journal of Hydrology*, 337, pp. 22-34.
- Aqil, M., Kita, I., Yano, A. and Nishiyama, S. (2007b), "Analysis and prediction of flow from local source in a river basin using a Neuro-fuzzy modeling tool." *Journal of Environmental Management*, 85(1), pp. 215-223.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000a), "Artificial neural networks in hydrology. I: Preliminary concepts." *Journal of Hydrologic Engineering*, ASCE, 5(2), pp. 115-123.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. (2000b), "Artificial neural networks in hydrology. II: Hydrologic applications." *Journal of Hydrologic Engineering*, ASCE, 5(2), pp. 124-137.
- Bowden, G. J., Dandy, G. C. and Maier, H. R. (2005), "Input determination for neural network models in

- Maier, H. R. and Dandy, G. C. (2000), "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications." *Journal of Hydrology*, 15, pp. 101-124.
- Marce, R., Comerma, M., García, J. C. and Armengol, J. (2004), "A neuro-fuzzy modeling tool to estimate fluvial nutrient loads in watersheds under time-varying human impact." *Limnology and Oceanography: Methods*, 2, pp. 342-355.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M. and Ramasastri, K. S. (2004), "A neuro-fuzzy computing technique for modeling hydrological time series." *Journal of Hydrology*, 291, pp. 52-66.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S. (2005), "Short-term flood forecasting with a neurofuzzy model." *Water Resources Research*, 41, pp. 2517-2530.
- Sajikumar, N. and Thandaveswara, B. S. (1999), "A non-linear rainfall-runoff model using an artificial neural network." *Journal of Hydrology*, 216, pp. 32-55.
- Sugeno, M. and Yasukawa, T. (1993), "A fuzzy-logic based approach to qualitative modelling." *IEEE Transactions on Fuzzy Systems*, 1(1), pp. 7-31.
- Tang, Z. and Fishwick, P. A. (1993), "Feedforward neural nets as models for time series forecasting." *ORSA Journal on Computing*, 5(4), pp. 374-385.
- Tibshirani, R. (1994), "A Comparison of Some Error Estimates for Neural Network Models." *Technical Working Paper No. 94-10*, Department of Statistics, University of Toronto.
- Valenca, M. J. S., Ludermir, T. B. and Valenca, A. C. B. (2005), "River Flow Forecasting for Reservoir Management Through Neural Networks." In: The Fifth International conference on Hybrid Intelligent Systems, Rio de Janeiro. *Proceedings of HIS05*. Los Alamitos: IEEE Computer Society, p. pp. 545-547.
- Yager, R. and Filev, D. (1994), "Generation of fuzzy rules by mountain clustering." *Journal of Intelligent and Fuzzy Systems*, 2(3), pp. 209-219.
- Zadeh, L. A. (1965), "Fuzzy sets." *Information Control*, 8(3), pp. 338-353.
- Zealand, C. M., Burn, D. H. and Simonovic, S. P. (1999), "Short term streamflow forecasting using artificial neural network." *Journal of Hydrology*, 214, pp. 32-48.
- water resources applications. Part 1—background and methodology." *Journal of Hydrology*, 301, pp. 75-92.
- Chiu, S. L. (1994), "Fuzzy Model Identification Based on Cluster Estimation." *Journal of Intelligent and Fuzzy Systems*, 2(3), pp. 267-278.
- Coulibaly, P., Anctil, F. and Bobee, B. (2000), "Daily reservoir inflow forecasting using artificial neural networks with stopped training approach." *Journal of Hydrology*, 230, pp. 244-257.
- Cybenko, G. (1989), "Approximation by superpositions of a sigmoidal function." *Mathematics of Control, Signals and System*, 2(4), pp. 303-314.
- Dibike, Y. B. and Solomatine, D. P. (2001), "River Flow Forecasting Using Artificial Neural Networks." *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(1), pp. 1-8.
- Dybowski, R. (1997), "Assigning confidence intervals to neural network predictions." *Technical report*, Division of Infection (St Thomas' Hospital), King's College London.
- Eckhardt, K., Breuer, L. and Frede, H. G. (2003), "Parameter uncertainty and the significance of simulated land use change effects." *Journal of Hydrology*, 273, pp. 164-176.
- Efron, B. and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, Chapman & Hall, New York.
- Hornik, K., Stinchcombe, M. and White, H. (1989), "Multilayer feedforward networks are universal approximators." *Neural networks*, 2, pp. 359-366.
- Jang, J. S. R. and Gulley, N. (1995), *The Fuzzy Logic Toolbox for Use with MATLAB*, The Mathworks Inc, Natick, MA.
- Jang, J. S. R. and Sun, C. T. (1995), "Neuro-Fuzzy Modeling and Control." *Proceedings of the IEEE*, 83, pp. 378-406.
- Jang, J. S. R. (1993), "ANFIS: Adaptive-Network-Based Fuzzy Inference System." *IEEE Transactions on Systems, Man, AND Cybernetics*, 23(3), pp. 665-685.
- Jang, J. S. R., Sun, C. T. and Mizutani, E. (1997), *Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice-Hall, New Jersey.
- Kumar, D. N., Raju, K. S. and Sathish, K. (2004), "River Flow Forecasting Using Recurrent Neural Network." *Water Resources Management*, 18, pp. 143-161.