



روشی پیشنهادی برای مدلسازی عدم قطعیت در تخمینهای هیدروکلیماتولوژی با تاکید بر دوره‌های

خشکسالی

شهاب عراقی نژاد^{۱*}

تاریخ دریافت: ۱۳۹۲/۱/۵

تاریخ پذیرش: ۱۳۹۲/۶/۲۱

چکیده

بسیاری از تکنیک‌ها و روشهایی که در برنامه‌ریزی و مدیریت کشاورزی بر مبنای تغییرات و نوسانات اقلیمی وجود دارد، ارتباط مستقیم یا غیر مستقیم با تخمینهای هیدروکلیماتولوژیکی دارد. منظور از تخمین هیدروکلیماتولوژی تعیین و اعلام وقوع شرایطی در مکان و زمان خاص است که می‌تواند شامل پدیده‌هایی نظیر ترسالی یا خشکسالی شود. پیشرفتهای بیشتر در تخمینهای هیدروکلیماتولوژیکی منوط به شناخت بیشتر ارتباط پدیده‌های اقلیمی و رویدادهای هیدرولوژیکی، بهبود و توسعه مدلها یا پیشنهاد مدل‌های جدید می‌باشد. ارائه یک مدل تخمین هیدروکلیماتولوژیکی که دارای خصوصیات مدلسازی همبستگی غیر خطی بین سیگنالها و متغیرهای هیدرولوژیکی، مدلسازی سریهای غیر ایستا و ارائه پیش‌بینیهای احتمالاتی باشد، می‌تواند یک قدم به جلو در مطالعات انجام شده در این زمینه باشد. در این راستا مدلی بر اساس تحلیل‌های همبستگی و مدلسازی آماری در قالب این مقاله ارائه شده و قابلیت آن در مقایسه با یک مدل‌های مرسوم آمار ناپارامتری (رگرسیون نزدیک ترین همسایه) در تخمین نقطه‌ای و احتمالاتی سنجیده می‌شود. یکی از مهمترین چالشها در انجام تخمینهای هیدروکلیماتولوژیکی، انعکاس عدم قطعیتها در خروجیها است که در روش ارائه شده به آن پرداخته شده است. نوآوری این تحقیق در ارائه روشی ساده، اما کارآمد برای تخمینهای هیدروکلیماتولوژیکی با انعکاس عدم قطعیتها است. مطالعه موردی تحقیق تخمین آوردهای رودخانه زاینده رود در شرایط نرمال و خشکسالیها می‌باشد. نتایج نشان دهنده کارایی روش پیشنهادی در مقایسه با روشهای رگرسیون ناپارامتری است.

واژه‌های کلیدی: تخمینهای هیدروکلیماتولوژیکی، عدم قطعیت، تحلیل‌های زمین آمار، روش نزدیک ترین همسایه

مقدمه

می‌باشیم. دو سؤال مهم در این زمینه قابل طرح است: چه پارامترهایی برای توصیف اقلیم یک منطقه مورد نیاز است و بهترین تخمین از تغییرات اقلیمی در یک دوره زمانی چه می‌تواند باشد. بسیاری از تکنیک‌ها و روشهایی که در برنامه‌ریزی و مدیریت منابع آب وجود دارد، ارتباط مستقیم یا غیر مستقیم با تخمینهای هیدروکلیماتولوژیکی دارد. منظور از تخمین هیدروکلیماتولوژی تعیین و اعلام وقوع شرایطی در مکان و زمان خاص است که می‌تواند شامل پدیده‌هایی نظیر ترسالی یا خشکسالی شود. در سالهای اخیر، مدلها و ابزار تخمینهای هیدروکلیماتولوژیکی بر پایه روشهای تجربی و آماری در کنار مدل‌های عددی پیشنهاد و به کار گرفته

چرخه آب در طبیعت به خوبی تداخل و درگیری پارامترهای هواشناسی و هیدرولوژیکی را نشان می‌دهد. برای حل مسائل مرتبط در حیطه علم هیدروکلیماتولوژی، ناگزیر از در نظر گرفتن عوامل مؤثر اقلیمی بر سیستم هیدروکلیماتولوژیکی بوده و جهت تعریف الگوریتم چگونگی استفاده از این عوامل، ناگزیر به طرح سئوالات اساسی در زمینه اقلیم و تغییرات آن

^۱ استادیار گروه مهندسی آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران.

(araghinejad@ut.ac.ir)

(*) نویسنده مسئول.

داد. گرچه این مدلها روشهایی کلاسیک به حساب می آیند، اما در مواردی که مدلسازی ساختار داخلی فرایند هیدرولوژی مشکل باشد از این مدلها استفاده شد، ضمن اینکه به عنوان مبنای مقایسه ای در مقابل مدلهای دیگر به کار می روند.

مدل های ARIMA توسط باکس و جنکینز (۱۹۷۶) مد نظر قرار گرفتند و توسط محققین بسیار دیگری نظیر براکول و دیویس (۱۹۸۷) بسط داده شده و به شکل بسیار گسترده ای در مدلسازی سری های زمانی هیدرولوژیکی استفاده شدند.

در سال های اخیر شیوه های آماری نا پارامتری بسیار مورد توجه محققین در علم آمار و تحلیل سری های زمانی و همچنین مهندسیین قرار گرفته است. تفاوت اصلی مدل های نا پارامتری و مدل های پارامتری، همان طور که از نام این مدل ها بر می آید، در این است که مدل های نا پارامتری، دارای توزیع آماری پارامتری نمی باشند. اساس این مدل ها استفاده از تابعی به نام تابع کرنل است (لال و شارما، ۱۹۹۶). روشهای مبتنی بر هوش مصنوعی مانند شبکه های عصبی مصنوعی کاربرد زیادی در مدلسازی فرایندهای هیدرولوژیکی در دهه اخیر داشته اند. محققین مختلفی همچون شو و همکاران (۱۹۹۹)، و کولیالی و همکاران (۲۰۰۰) کاربردهای جالبی از این مدلها در مهندسی هیدرولوژی ارائه داده اند.

یکی از دغدغه های مهم در تخمین های هیدرولوژیکی، پیش بینی شرایط خشکسالی است. محققین مختلف با بکارگیری روشهای مختلف و البته استفاده از عوامل هیدرولیماتولوژیکی گوناگون به این مهم پرداخته اند. یکی از عواملی که تخمینهای هیدرولیماتولوژیکی در سالهای اخیر را رونق داده است مشاهده تأثیرات از راه دور اقلیم در الگوی بارشهای فصلی در نقاط مختلف جهان بوده است. از این رو استفاده از سیگنالهای بزرگ مقیاس اقلیمی به عنوان تخمین زننده بارش یا رواناب از چالشهای مهم مدلسازی بوده است.

مدرس پور (۱۳۷۷) به بررسی ناهنجاری های اقلیمی ایران با توجه به اطلاعات دما و بارش در سال های ال نینو در یک دوره ۳۰ ساله (۱۹۶۰ تا ۱۹۸۹) پرداخت. وی الگوهای دما و بارش را در ۱۵ ایستگاه سینوپتیک ایران، طی شش رخداد ال نینو در دوره مزبور مورد مطالعه قرار داد. بررسی انجام شده توسط وی در چهار فصل مانسون (می، ژوئن، ژوئیه و اوت) در سال وقوع ال نینو، پاییز (سپتامبر، اکتبر و نوامبر) در سال

شده اند. پیشرفتهای بیشتر در تخمینها، منوط به شناخت بیشتر ارتباط پدیده های اقلیمی و رویدادهای هیدرولوژیکی، بهبود و توسعه مدلها و پیشنهاد مدل های جدید می باشد.

یکی از رایج ترین مدلهای مورد استفاده برای تخمین های هیدرولیماتولوژیکی، مدلهای بر مبنای رگرسیون چند متغیره می باشد (دی، ۱۹۸۳). این مدلها بصورت روابط خطی و غیر خطی به کار گرفته می شوند. گارن (۱۹۹۲) روشهایی برای بهبود نتایج استفاده از این گونه مدلها ارائه چرخه آب در طبیعت به خوبی تداخل و درگیری پارامترهای هواشناسی و هیدرولوژیکی را نشان می دهد. برای حل مسائل مرتبط در حیطه علم هیدرولیماتولوژی، ناگزیر از در نظر گرفتن عوامل مؤثر اقلیمی بر سیستم هیدرولیماتولوژیکی بوده و جهت تعریف الگوریتم چگونگی استفاده از این عوامل، ناگزیر به طرح سئوالات اساسی در زمینه اقلیم و تغییرات آن می باشیم. دو سؤال مهم در این زمینه قابل طرح است: چه پارامترهایی برای توصیف اقلیم یک منطقه مورد نیاز است و بهترین تخمین از تغییرات اقلیمی در یک دوره زمانی چه می تواند باشد. بسیاری از تکنیک ها و روشهایی که در برنامه ریزی و مدیریت منابع آب وجود دارد، ارتباط مستقیم یا غیر مستقیم با تخمینهای هیدرولیماتولوژیکی دارد. منظور از تخمین هیدرولیماتولوژی تعیین و اعلام وقوع شرایطی در مکان و زمان خاص است که می تواند شامل پدیده هایی نظیر ترسالی یا خشکسالی شود. در سالهای اخیر، مدلها و ابزار تخمینهای هیدرولیماتولوژیکی بر پایه روشهای تجربی و آماری در کنار مدلهای عددی پیشنهاد و به کار گرفته شده اند. پیشرفتهای بیشتر در تخمینها، منوط به شناخت بیشتر ارتباط پدیده های اقلیمی و رویدادهای هیدرولوژیکی، بهبود و توسعه مدلها و پیشنهاد مدل های جدید می باشد.

یکی از رایج ترین مدلهای مورد استفاده برای تخمین های هیدرولیماتولوژیکی، مدلهای بر مبنای رگرسیون چند متغیره می باشد (دی، ۱۹۸۳). این مدلها بصورت روابط خطی و غیر خطی به کار گرفته می شوند. گارن (۱۹۹۲) روشهایی برای بهبود نتایج استفاده از این گونه مدلها ارائه داد. گرچه این مدلها روشهایی کلاسیک به حساب می آیند، اما در مواردی که مدلسازی ساختار داخلی فرایند هیدرولوژی

روش شناسی پیشنهادی

اگر Z متغیر وابسته و عوامل موثر بر آن $P = \{p_1, p_2, \dots, p_m\}$ باشد، یک مدل پیش بینی را می‌توان به صورت تابع f به شکل زیر نشان داد:

$$Z_t = f(P_t) + \varepsilon_t \quad t=1, 2, \dots, n \quad (1)$$

که n تعداد اطلاعات مشاهده شده و ε خطای نگاشت است. یک فضای مجازی در نظر بگیرید که محورهای مختصات این فضا توسط عوامل موثر تعریف می‌شود. بنابراین مختصات هر نقطه در این فضا توسط این عوامل مشخص می‌شود. متغیر Z در رابطه (۱) را به عنوان متغیری در این فضا در نظر بگیرید که موقعیت آن با متغیرهای موثر مشخص می‌شود. چنانچه n مشاهده تاریخی از Z در اختیار باشد، آنگاه n نقطه با مقدار معلوم در این فضای مجازی خواهیم داشت و هر متغیر پیش بینی، به عنوان نقطه‌ای با مقدار نامعلوم در این فضا در نظر گرفته می‌شود. چنین فرضی مسئله تخمین را تبدیل به یک مسئله شناخته شده در مطالعات صحرایی و تحلیل‌های جغرافیایی می‌کند. بسیاری از متغیرها در طبیعت تعداد اندازه‌گیری‌های محدودی در نقاط محدود دارای اندازه‌گیری دارند. مقادیر این متغیرها در مکان‌های دیگر بدون اندازه‌گیری باید از این اطلاعات محدود و در دسترس تخمین زده شود. ژئواستاتستیک (زمین‌آمار) ترکیبی از روش‌هایی است که تخمین یک متغیر را از نقاط اطراف ممکن می‌سازد.

عوامل موثر در تخمین یک متغیر هیدروکلیماتولوژیکی معمولاً با هم همبستگی دارند. به عنوان نمونه، همبستگی زیادی بین مقادیر بارش و رطوبت خاک در یک دوره زمانی می‌تواند وجود داشته باشد و در عین حال هر دو به عنوان عوامل موثر در تخمین رواناب مطرح می‌باشند. از طرفی محورهای مختصات دکارتی بر یکدیگر عمودند و عوامل موثر نمی‌توانند فضایی با محورهای عمود بر یکدیگر تشکیل دهند. بنابراین در روش پیشنهاد شده از مولفه‌های اصلی عوامل موثر برای تشکیل فضای مجازی تشریح شده استفاده می‌شود. این فضای مجازی و داده‌های اندازه‌گیری شده و

وقوع ال نینو و بهار (مارس، آوریل و می) سال بعد از وقوع ال نینو نشان داد که در سال وقوع ال نینو، در فصل پاییز مناطق جنوب غرب، جنوب شرق و بخشی از مرکز کشور دارای بارندگی زیر میانگین و مناطق شمال غرب، بخشی از غرب، شمال و تهران دارای بارندگی بالای میانگین بوده‌اند. خوش اخلاق (۱۳۷۷) به صورت آماری ارتباط بین فازهای سرد و گرم ENSO را با بارش ایران بررسی نمودند. نتایج وی که حاصل از همبستگی SOI و بارش می‌باشد، نشان می‌دهد که خشکسالی‌های ایران همزمان با لائینا و ترسالی‌ها همزمان با ال نینو بوده است.

کیانی‌پور (۱۳۷۹) شرایط سینوپتیکی حاصل از پدیده ال نینو در ایران و تغییرات ایجاد شده در نحوه فعالیت و آرایش سیستم‌ها را بررسی نموده است. ایشان در بررسی‌های خود، سال ۸۳-۱۹۸۲ را به عنوان یک سال نمونه برای حالت ال نینو و سال ۷۵-۱۹۷۴ را به عنوان یک سال نمونه برای حالت لائینا در نظر گرفته و مطالعات خود را صرفاً برای ماه‌های دسامبر، ژانویه و فوریه این دو سال انجام داده است. نتایج وی نشان می‌دهد که آرایش سیستم‌ها در سه ماه مزبور در سال لائینا شرایط مناسب‌تری برای ایجاد بارش نسبت به سال ال نینو داشته است.

یکی از مهمترین چالش‌ها در انجام تخمین‌های هیدروکلیماتولوژیکی انعکاس عدم قطعیتها در خروجیها است. اغلب روشهایی که در فوق مرور شد برای محاسبات و تخمین‌های نقطه‌ای (با به عبارتی قطعی) بکار می‌روند. ارائه بهره‌گیری از روشهایی که به نوعی عدم قطعیتها را در خروجی‌های تخمینی منعکس کنند در ایجاد اطمینان در بیان واقعی محاسبات هیدروکلیماتولوژیکی بسیار حائز اهمیت خواهد بود. نوآوری این تحقیق در ارائه روشی ساده اما کارآمد برای تخمین‌های هیدروکلیماتولوژیکی با انعکاس عدم قطعیتها است. ساختار مقاله ارائه شده بدین شکل است که پس از مقدمه روش شناسی پیشنهادی ارائه می‌شود. مطالعه موردی برای کاربرد این روش در بخش بعد ارائه می‌شود. نتایج و بحث آن در بخش بعد ارائه خواهد شد و مقاله با خلاصه و جمع‌بندی به انتها می‌رسد.

ایستای مرتبه دوم باشد، می‌توان فرض کرد که توزیع هر متغیر در هر مختصات یک توزیع نرمال است (وبستر و الیور، ۲۰۰۱).

یک فرایند، ایستای درجه دوم است اگر میانگین، واریانس و کوواریانس آن شرایط زیر را دارا باشد:

$$E(Z(x)) = m(x) = m \quad (۲)$$

$$\text{var}(Z(x)) = \sigma^2(x) = \sigma^2$$

$$\text{COV}(x_1, x_2) = \text{COV}(x_1 - x_2) = \text{COV}(h)$$

که در رابطه فوق، $h = x_1 - x_2$. فرایند تصادفی $Z(x)$ فرضیه ذاتی را ارضا می‌کند چنانچه تفاضل مرتبه اول آن دارای شرایط زیر باشد:

$$E(Z(x_1) - Z(x_2)) = m(h) \quad (۳)$$

$$\text{var}(Z(x_1) - Z(x_2)) = 2\gamma(h)$$

چنانچه فرایندی ایستای مرتبه دو باشد، خود به خود فرضیات ذاتی ارضا می‌شود اما عکس آن صادق نیست. به دلیل منطقه ای بودن اطلاعات در حقیقت از فضای نرمال استفاده می‌شود که فراتر از نرمال کردن اطلاعات است. از آنجایی که متغیر پیش بینی نرمال فرض شده است، تابع توزیع آن با معلوم بودن میانگین و واریانس به راحتی محاسبه می‌شود. چون این تابع توزیع برای هر مختصات منحصر به فرد است بنابراین این تابع، یک تابع احتمالاتی مشروط به هر مختصات در فضای در نظر گرفته شده (عوامل موثر) می‌باشد.

با توجه به مطالب ارائه شده روش پیشنهادی تخمین همراه با محاسبه عدم قطعیت در گام های زیر خلاصه می‌شود:

- ۱- تشخیص متغیر وابسته و عوامل موثر
- ۲- انتقال متغیر وابسته به یک محیط نرمال با ارضای شرط ایستایی مرتبه ۲
- ۳- با استفاده از محاسبه مولفه های اصلی عوامل موثر یک فضای مجازی ایجاد می‌شود که در آن متغیرهای وابسته

اندازه گیری نشده در شکل (۱) نشان داده شده است. همان طور که در این شکل مشاهده می‌گردد، مولفه های اصلی عوامل موثر (PC_i)، مختصات این فضا را می‌سازند. در این فضا، نقاط توپر به عنوان متغیرهای مشاهده شده و نقاط توخالی به عنوان متغیرهای تخمین نشان داده شده است. مختصات هر نقطه در این فضا بر اساس عوامل موثر تعیین می‌شود.

اولین گام برای تخمین اطلاعات مجهول در نقاط اندازه گیری شده، محاسبه همبستگی بین داده های موجود است. این همبستگی در روش های زمین آماری توسط تابعی به نام واریوگرام (تغییر نما) بیان می‌شود. واریوگرام تابعی از فاصله h است که در این حالت خاص، فاصله اقلیدسی بین متغیرهای مشاهداتی و به نوبه خود تابعی از فاصله بین عوامل موثر می‌باشد. با محاسبه تابع واریوگرام در فضای مجازی تعریف شده، مقدار متغیر تخمین در هر نقطه به وسیله حل معادلات سیستم کربجینگ (که یکی از روش های زمین آمار است) قابل انجام است. این تخمین به عنوان تخمین نقطه‌ای یا منفرد محسوب می‌شود.

تا این مرحله از تخمین، تابع f نشان داده شده در رابطه (۱) محاسبه شد. اما تخمین متغیر استوکستیک مجهول نیازمند تخمین تغییرات متغیر تصادفی نیز می‌باشد. بنابراین با حل سیستم معادلات کربجینگ (که در بخش بعد به آن پرداخته می‌شود) متوسط مربعات خطا در هر نقطه نیز محاسبه می‌شود. این مقدار در حقیقت واریانس متغیر تخمین می‌باشد. تخمین نقطه ای در هر مکان در حقیقت برابر با میانگین متغیر پیش بینی می‌باشد. با توجه به محاسبه واریانس این متغیر، چنانچه توزیع این متغیر نیز مشخص گردد، مقادیر نامعلوم در هر نقطه به صورت کلیه مقادیر محتمل به بدست می‌آیند. در مسائل عملی تخمین توزیع متغیر مورد تحلیل مشکل است چرا که اولاً تنها یک واقعه از آن متغیر وجود دارد، ثانیاً معمولاً توزیع های تجربی متفاوت از توزیع ثنوری می‌باشند. یک راه حل برای بر طرف کردن این مشکل انتقال توزیع تجربی داده های مشاهداتی به یک توزیع نرمال و انجام محاسبات در یک فضای نرمال است. چنانچه توزیع Z نرمال و فرایند، فرضیه ذاتی (intrinsic hypothesis) را ارضا کند و

و ماتریس‌های A و b را به صورت زیر تعریف کنیم:

$$b = \begin{bmatrix} -\gamma(\|x_1 - x_0\|) \\ -\gamma(\|x_2 - x_0\|) \\ \vdots \\ -\gamma(\|x_n - x_0\|) \\ 1 \end{bmatrix} \quad (5)$$

$$A = \begin{bmatrix} 0 & -\gamma(\|x_1 - x_2\|) & \cdots & -\gamma(\|x_1 - x_n\|) & 1 \\ -\gamma(\|x_2 - x_1\|) & 0 & \cdots & -\gamma(\|x_2 - x_n\|) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\gamma(\|x_n - x_1\|) & -\gamma(\|x_n - x_2\|) & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \quad (6)$$

آنگاه معادله ماتریسی زیر نشان دهنده معادلات خواهد بود.

$$AX = b \quad (7)$$

با حل مجموع معادلات فوق مقادیر مجهول λ_1 الی λ_n و v بدست می‌آیند.

همچنین معادله زیر جهت تخمین واریانس مقدار مجهول در

نقطه x_0 به کار می‌رود.

(۸)

$$\sigma_0^2 = E[(\hat{z}_0 - z(x_0))^2] = -v + \sum_{i=1}^n \lambda_i \gamma(\|x_i - x_0\|)$$

مقادیر پیش بینی احتمالاتی توسط این روش بین دو حد تغییر می‌کند. یک حد آن پیش بینی کامل که برابر با مقدار واقعی بوده و واریانس آن صفر است. این حالت هنگامی حاصل می‌شود که همبستگی کامل بین داده‌ها وجود داشته باشد، به عبارتی مقدار واریوگرام برابر با صفر باشد. حالت حدی دیگر هنگامی رخ می‌دهد که همبستگی بین اطلاعات وجود نداشته باشد و واریوگرام به ازای همه مقادیر h برابر با ۱ باشد. در این حالت از معادله (۷) مقدار v برابر با صفر به دست می‌آید. بر اساس رابطه (۸) واریانس متغیر پیش بینی برابر با واریانس داده‌ها ($=1$) و متوسط آن برابر با میانگین کل داده‌ها می‌باشد. در این حالت در حقیقت توزیع شرطی پیش بینی به توزیع تجربی داده‌ها تبدیل می‌شود، بدین

مشاهده شده به عنوان اندازه‌گیری‌ها در آن فضا فرض می‌شود. این فضای مجازی در شکل (۱) نشان داده شده است.

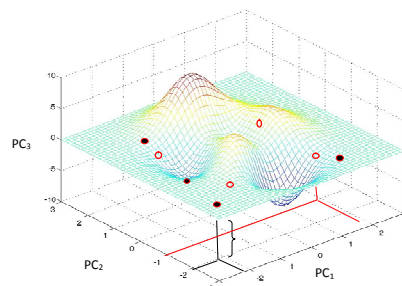
۴- واریوگرافی نقاط اندازه‌گیری شده و برازش واریوگرام تئوری بر روی آن

۵- در نظر گرفتن عوامل موثر مشاهده شده در زمان حال به عنوان مختصات مقدار مجهول مورد نظر

۶- حل معادلات سیستم کریجینگ با استفاده از واریوگرام گام چهارم و محاسبه مقدار میانگین (برابر با پیش بینی منفرد) و واریانس مقدار مجهول در این نقطه

۷- تشکیل یک تابع توزیع نرمال با میانگین و واریانس بدست آمده از گام ششم

۸- تبدیل تابع توزیع گام هفتم به توزیع تجربی اولیه



شکل ۱ - فضای مجازی در نظر گرفته شده برای تخمین متغیر مجهول

سیستم ماتریسی حل معادلات

بخش عمده‌ای از محاسبات بر اساس تحلیل ژئواستاتستیک انجام میشود، در این بخش معادلات لازم برای انجام این کار ارائه می‌گردد. فرضیه کریجینگ برای بار اول بر اساس کار تجربی کریج (۱۹۹۶) ارائه شد. معمولاً سیستم معادلات کریجینگ به صورت ماتریسی نوشته می‌شود؛ اگر X را به صورت ماتریس مقادیر مجهول به صورت زیر در نظر بگیریم:

$$X = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ v \end{bmatrix} \quad (4)$$

در حالی که لال و شارما (۱۹۹۶) نشان دادند که با تعمیم رابطه (۱۱) تابع کرنل زیر به دست می آید که می تواند جواب های بهتری در مدل سازی فرایندهای مارکوفی ایجاد کند:

$$K(j(t)) = \frac{1/j}{\sum_{j=1}^k 1/j} \quad (12)$$

در نهایت مقدار تخمین از رابطه زیر محاسبه می شود:

$$D_r = \sum_{j=1}^k K(j(t)) * D_t(j) \quad (13)$$

که $D_t(j)$ مقدار متغیر وابسته تاریخی دارای رتبه j می باشد.

توزیع احتمالاتی شرطی در این روش با استفاده از تعیین احتمال وقوع همسایه ها و تابع کرنل مورد استفاده به دست می آید.

مدل $K-NN$ در هر بار پیش بینی، K الگوی شبیه به شرایط فعلی را شناسایی کرده و به آن ها وزن می دهد. در این روش پارامتر K (تعداد همسایه ها) ثابت می باشد. روش ارائه شده نیز در حقیقت برای هر پیش بینی نزدیک ترین حالت های به وقوع پیوسته را انتخاب کرده و در محاسبه متغیر پیش بینی در نظر می گیرد. تعداد همسایه ها در این روش ثابت نمی باشد و از طریق پارامتر $range$ در واریوگرام برای هر نقطه محاسبه می گردد. با توجه به این پارامتر، تعداد همسایه ها و الگوهای شبیه به شرایط پیش بینی، بسته به توزیع داده ها و موقعیت متغیر پیش بینی مقداری متغیر می باشد. این کار انعطاف بیشتری به نداشت رفتار غیر خطی به مدل می دهد و از نرم کردن نتایج پیش بینی جلوگیری می کند.

مدل $K-NN$ با وزن دهی به هر یک از پیش بینی کننده ها سهم آن ها را در تعیین متغیر پیش بینی محاسبه می کند. در مدل پیشنهاد شده با فرض غیر ایزوتروپیک بودن فضای پیش بینی کننده ها با محاسبه واریوگرام های مختلف در جهات مختلف، سهم هر یک از پیش بینی کننده ها در تعیین متغیر پیش بینی با وزن های مختلف در نظر گرفته می شود.

معنی که استفاده از پیش بینی کننده ها هیچ اطلاعات اضافی برای تخمین متغیر پیش بینی در اختیار قرار نمی دهد.

مقایسه روش ارائه شده با روش K - نزدیکترین همسایه

روش پیشنهاد شده شباهت مفهومی زیادی با مدل $K-NN$ نزدیک ترین همسایه دارد و در جهت کاهش خطاهای آن توسعه یافته است. الگوریتم انجام پیش بینی با استفاده از روش $K-NN$ (K-nearest neighborhood) به صورت زیر است:

بردار d بعدی متغیرهای مستقل در مسئله مورد نظر پیش بینی تعیین می شود. به طوریکه:

$$D_t = f(x_{1t}, x_{2t}, \dots, x_{dt}) \quad (9)$$

متغیر پیش بینی را برابر با Dr فرض کرده و با استفاده از فاصله وزنی اقلیدسی زیر k نزدیک ترین همسایه آن از میان مقادیر Dt انتخاب می شود.

$$Dist_{rt} = \left(\sum_{j=1}^d w_j (x_{rj} - x_{tj})^2 \right)^{1/2} \quad (10)$$

که در آن مقادیر w_j به ترتیب وزن هایی است که برای پیش بینی کننده ها در نظر گرفته می شود. نزدیک ترین همسایه ها بر اساس فاصله محاسبه شده فوق از ۱ تا k رتبه بندی می شوند به نحوی که همسایه با کمترین فاصله کمترین رتبه (j) را دارد و بر عکس. چنانچه دو یا چند همسایه دارای یک فاصله باشند، رتبه همسایه ای که در سری تاریخی فاصله زمانی کمتری با زمان پیش بینی دارد، کمتر در نظر گرفته می شود.

از یک تابع کرنل گسسته برای وزن دهی به همسایه ها استفاده می شود. کارلسون و یاکوویتز (۱۹۸۷) تابع کرنل زیر را پیشنهاد دادند:

$$K(Dist(t)) = \frac{1/Dist_{rt}}{\sum_{t=1}^k 1/Dist_{rt}} \quad (11)$$

فاصله بین پیش بینی و مقدار مشاهداتی را با استفاده از آماره زیر تحت عنوان مقدار LEPS اندازه گیری کردند:

$$S = 3 \left(1 - |P_f - P_v| + P_f^2 - P_f + P_v^2 - P_v \right) - 1 \quad (14)$$

که در آن P_f احتمال تجمعی متغیر پیش بینی می باشد که از منحنی توزیع تجمعی (تجربی) تاریخی متغیر به دست آمده باشد و P_v احتمال متناظر با مقدار مشاهداتی به دست آمده از توزیع تجمعی تاریخی متغیر جریان می باشد. نحوه تعیین این مقادیر در شکل (۲) نشان داده شده است. توزیع تجربی ورودی ها تنها یک واقعه می باشد که بر اساس شرایط اقلیمی و هیدرولوژیکی خاص یک منطقه ثبت شده است. بدیهی است که هرچه اطلاعات ما شامل عوامل موثر بر میزان حجم رواناب بیشتر باشد، توزیع پیش بینی متفاوت و دارای گستره کمتری نسبت به توزیع تجربی خواهد بود. پیچوتا و همکاران (۱۹۹۶) الگوریتم محاسبه آماره LEPScore را برای سنجش یک توزیع پیوسته تعمیم دادند. در این پایان نامه نیز از این روش تعمیم داده شده برای ارزیابی نتایج مدل پیشنهادی بهره گرفته می شود.

بر اساس این الگوریتم، از روی توزیع متغیر پیش بینی، مقادیر مختلفی انتخاب می شود (که در این جا تعداد ۱۰۰ پیش بینی در نظر گرفته شده است). مقدار احتمال متناظر با هر یک از این مقادیر از روی توزیع احتمالاتی تاریخی جریان خوانده می شود و به ازای هر یک از این مقادیر، مقدار S از رابطه (۱۴) محاسبه می گردد. مجموع مقادیر S آماره S'' را به شرح زیر برای هر پیش بینی به دست می دهد.

$$S''_j = \sum_{k=0}^1 S_k \quad (15)$$

در رابطه فوق k از صفر تا یک در بازه های یک صدمی تغییر می کند. و متوسط آماره LEPScore، (SK) برای تمام زوج های پیش بینی کننده ها و متغیر مستقل از رابطه زیر بدست می آید:

$$SK = \frac{\sum_{j=1}^n 100S''_j}{\sum_{j=1}^n S''_{mj}} \quad (16)$$

در روش $K-NN$ سهم هر یک از همسایه ها در تعیین مقدار پیش بینی از بهینه (مینیمم) سازی معیار حداقل بایاس به دست می آید. در حالی که در روش پیشنهادی (در حقیقت کریجینگ) سهم هر یک از همسایه ها نه تنها از معیار حداقل کردن بایاس بلکه از حداقل کردن واریانس تخمین حاصل می شود. چنین رویکرد به خودی خود پیش بینی هایی با قطعیت بیشتر ایجاد می کند.

توزیع احتمالات شرطی متغیر پیش بینی در روش $K-NN$ به طور ضمنی و با محاسبه تابع کرنل گسسته محاسبه می شود. در حالی که تابع توزیع احتمالاتی پیش بینی در روش ارائه شده به طور صریح و پیوسته به دست می آید.

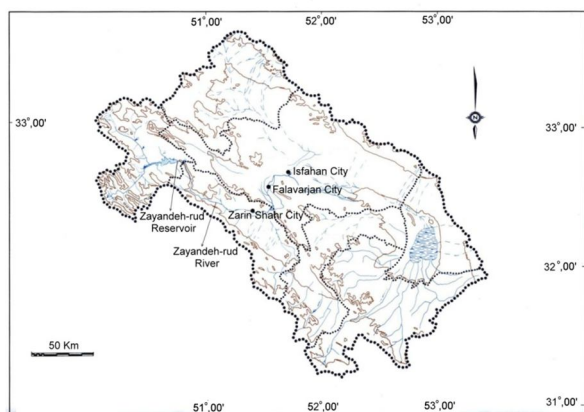
اندازه گیری دقت و قطعیت تخمین

یک مدل تخمین احتمالاتی مناسب باید بتواند توزیعی از پیش بینی ارائه دهد که هم دقیق و هم قطعی باشد. توزیع پیش بینی هنگامی دقیق است که مقدار مورد انتظار پیش بینی به مقدار واقعی نزدیک باشد. قطعیت یک توزیع پیش بینی به وسیله گستره آن اندازه گیری می شود. بدترین حالت پیش بینی یک توزیع نادقیق است زیرا حس اطمینانی نسبت به یک پیش بینی به دست می دهد که عملاً اشتباه است و با واقعیت فاصله زیادی دارد. شکل (۲) پیش بینی های احتمالاتی مناسب و نامناسب را نیز نشان می دهد. چنانچه انحراف معیار را به عنوان شاخص گستره یک توزیع در نظر بگیریم، ترکیبی از این شاخص و مقدار مورد انتظار می تواند معیاری برای بررسی توام دقت و قطعیت یک پیش بینی ارائه دهد.

یکی از آماره های اندازه گیری دقت یک پیش بینی اندازه گیری خطای خطی در فضای احتمالاتی (Linear error in probability space score=LEPS) برای اولین بار پاتس و همکاران (۱۹۹۶) این آماره را برای یک پیش بینی نقطه ای و مقایسه آن با توزیع تاریخی متغیرهای هواشناسی به کار بردند. ایشان نشان دادند که این آماره بهتر از آماره های دیگر نظیر ریشه متوسط مربعات خطا، پیش بینی متغیرهای اقلیمی را ارزیابی و به خصوص در پیش بینی مقادیر حدی دقیقتر عمل می کند. پاتس و همکاران (۱۹۹۶)

سطحی می‌باشد. وجود سد زاینده‌رود اهمیت مطالعه آوردها و شرایط هیدرولوژیکی این حوضه را برای کل منطقه دو چندان می‌کند. برای شبیه سازی سیستم واقعی، علاوه بر ورودی طبیعی حوضه سد زاینده رود، آوردهای حاصل از انتقال بین حوضه‌ای شامل آورد تونل های اول و دوم کوه‌رنگ، در مطالعات و مدل سازی ها در نظر گرفته شده است. شکل (۳) نقشه حوضه گاوخونی (زاینده رود) و موقعیت سد زاینده رود و حوزه آن را نشان می‌دهد.

در تحلیل ها و تصمیم‌گیری های مربوط به بخش های مختلف اعم از مدل های پیش بینی آورد رودخانه زاینده‌رود و بهره‌برداری از سد، از اطلاعات هیدرولوژیکی حوزه، و اطلاعات مربوط به سیستم مدیریت منابع آب منطقه استفاده شده است. اطلاعات تاریخی بارش در حوزه سد، حجم رواناب ورودی به سد زاینده رود، اطلاعات بودجه برفی به علاوه مشخصات فیزیکی سد و نیازهای آبی منطقه، از جمله اطلاعات مورد استفاده می‌باشد.

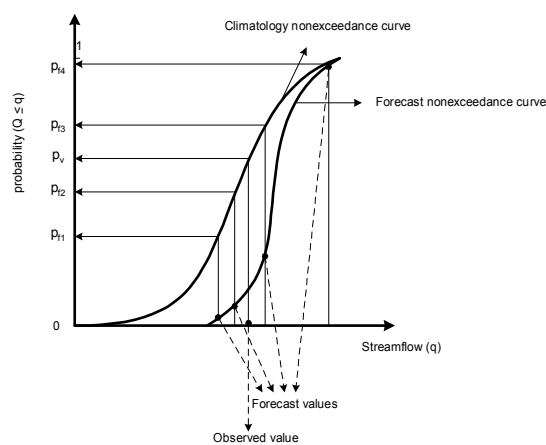


شکل ۳ - موقعیت حوضه سد زاینده رود و حوزه گاوخونی

بهره برداری از سد زاینده رود به منظور تامین نیازهای آبی شرب، صنعت و کشاورزی استان اصفهان که در حال حاضر مجموعاً سالانه ۱۵۰۰ میلیون متر مکعب تخمین زده می‌شود، انجام می‌گیرد. ۷۰ درصد از این نیاز متعلق به بخش کشاورزی است. فصل زراعی در استان شامل دو فصل پاییز-زمستان و بهار-تابستان است که فصل اول شامل کشت هایی است که در خلال ماه های آبان الی اسفند انجام می‌گیرد. فصل دوم شامل کشت هایی است که در خلال ماه های فروردین الی مهر انجام شده و نیاز آبی آن حدوداً دو و نیم

در رابطه فوق S_m'' همانند S'' به دست می‌آید با این تفاوت که هنگامی که S'' مثبت باشد، S_m'' با فرض بهترین پیش بینی $p_f = p_v$ به دست می‌آید و هنگامی که S'' منفی باشد، S_m'' با فرض بدترین پیش بینی به دست می‌آید. بدترین پیش بینی به این صورت به دست می‌آید که هنگامی که p_v کمتر از ۰/۵ است، $p_f = 1$ و $p_v = 0$ فرض می‌شود و هنگامی که p_v بیش از ۱ است $p_f = 0$ و $p_v = 1$ در نظر گرفته می‌شود.

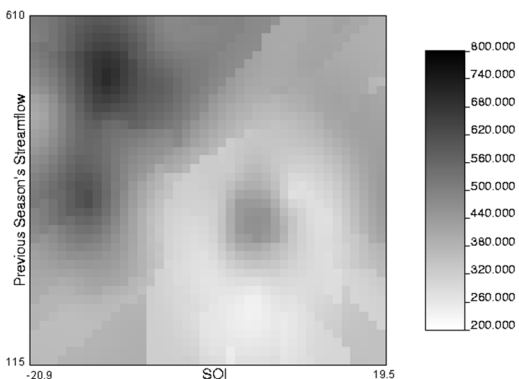
مقدار LEPSscore بین ۱۰۰- تا ۱۰۰+ برای بدترین بهترین پیش بینی تغییر می‌کند. بر اساس پیچوتا و همکاران (۱۹۹۶) پیش بینیهایی که LEPSscore بیش از ۱۰ دارند به عنوان پیش بینی های خوب شناخته می‌شوند.



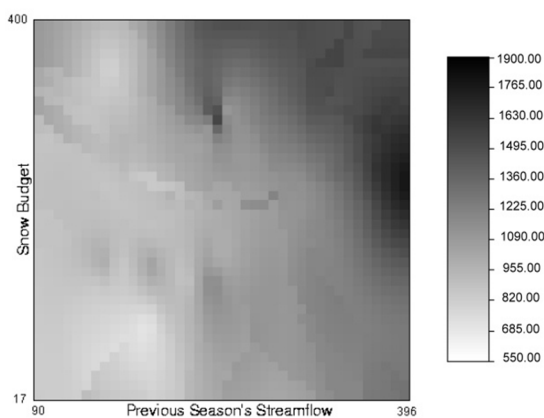
شکل ۲- نحوه تعیین p_f و p_v در محاسبه LEPS score

مطالعه موردی

منطقه مورد مطالعه، حوضه آبریز سد زاینده‌رود از سرشاخه‌های آن در ارتفاعات کوه‌رنگ تا سد زاینده رود می‌باشد. این حوضه یکی از ۵ زیر حوضه اصلی گاوخونی است که در حد فاصل طول‌های جغرافیایی ۵۰°-۰۲' و ۴۵°-۵۲' و عرض‌های جغرافیایی ۳۱°-۱۵' و ۳۳°-۴۵' در منطقه مرکزی ایران به وسعت ۳۲۰۰۰ کیلومتر مربع قرار دارد. حوضه آبریز سد زاینده‌رود به وسعت ۴۲۰۰ کیلومتر مربع مهم ترین حوضه آبریز این منطقه به لحاظ آوردها و مدیریت منابع



شکل ۴- تغییرات رواناب پاییز-اسفند با تغییرات SOI و رواناب تابستان



شکل ۵- تغییرات رواناب بهار تابستان با تغییرات NAO و رواناب زمستان

در این الگوریتم هر یک از زوج اطلاعات پیش‌بینی‌کننده‌ها و متغیرهای مستقل یکبار حذف شده و با استفاده از کالیبراسیون بقیه زوج اطلاعات، پیش‌بینی می‌شوند. این فرایند تا در نظر گرفتن همه زوج اطلاعات ادامه پیدا می‌کند. مزیت این الگوریتم استفاده از کلیه اطلاعات موجود در فرایند کالیبراسیون و صحت‌سنجی در مواردی است که اطلاعات بلند مدت در دسترس نمی‌باشد.

نتایج آماره‌های مختلف برای پیش‌بینی فصل‌های اول و دوم با استفاده از روش پیشنهادی و روش رایج‌ترین نزدیکترین همسایه به ترتیب در جداول (۱) و (۲) ارائه شده است. کلیه نتایج ارائه شده در این جداول نیز با استفاده از الگوریتم صحت‌سنجی متقاطع (cross validation) ۳۲ ساله ارائه شده است. برای تحلیل نتایج پیش‌بینی در شرایط اقلیمی مختلف، نتایج این آماره‌ها به تفکیک سال‌های خشک و تر در جدول (۳) ارائه شده است. سال‌های خشک و تر بر اساس در

برابر بیش از نیاز آبی فصل اول کشت است. در ابتدای هر فصل زراعی، سیاست تامین نیازهای کشاورزی تعیین شده و بر اساس این سیاست مدیریت سد خود را متعهد به تامین بخشی از یا تمامی نیازهای آبی آن فصل کشت می‌سازد. از این پس ماه‌های آبان الی اسفند به عنوان فصل اول و ماه‌های فروردین الی شهریور به عنوان فصل دوم نامیده می‌شوند که به ترتیب نماینده فصول اول و دوم زراعی در حوضه زاینده رود بر اساس مشاهده سیگنال‌های اقلیمی نوسانات جنوبی، نوسانات آتلانتیک شمالی و رواناب مشاهداتی می‌باشد. متوسط و انحراف معیار حجم جریان ورودی به سد زاینده رود در نیمه اول سال به ترتیب برابر با ۴۰۸ و ۱۵۰ میلیون متر مکعب می‌باشد. متوسط و انحراف معیار حجم ورودی به سد زاینده رود در نیمه دوم سال به ترتیب برابر با ۱۱۶۴ و ۳۷۳ میلیون متر مکعب می‌باشد.

نتایج و بحث

با استفاده از روش پیشنهادی این تحقیق، شکل‌های (۴) و (۵) به عنوان مبنای تخمین‌ها ترسیم شده‌اند. شکل (۵) تغییرات بزرگی جریان فصل پاییز و زمستان بر اساس تغییرات شاخص نوسانات جنوبی (SOI) و جریان رودخانه در تابستان قبل را نشان می‌دهد. همچنین شکل (۶) تغییرات جریان رودخانه در فصل بهار را با توجه به تغییرات نوسانات آتلانتیک شمالی (NAO)، جریان زمستان و بودجه برفی مشاهده شده در زمستان نشان می‌دهد. این شکل‌ها برای تخمین و بررسی مقدار جریان و بررسی تغییرات آنها با عوامل موثر بسیار مفیدند. نتایج تخمین‌های نقطه‌ای (منفرد) انجام شده به وسیله سه معیار کمترین متوسط خطا (RMSE)، درصد خطا (%VE)، همبستگی مقادیر پیش‌بینی شده و واقعی (CORR) و در حالت پیش‌بینی احتمالاتی توسط LEPS Score سنجیده می‌شود.

نتایج حاصل از پیش‌بینی برای هر دو فصل در دوره ۳۲ ساله و بر اساس صحت‌سنجی متقاطع (Cross validation) و مقایسه آن با مقادیر واقعی ارائه شده است.

نظر گرفتن متوسط آوردها به عنوان آستانه خشکسالی به دست آمده اند. تعداد سال های خشک در فصل اول و دوم به ترتیب ۱۹ و ۲۰ می باشد. بقیه سال ها به عنوان سال های تر در نظر گرفته شده اند.

جدول ۱- نتایج صحت سنجی متقاطع ۳۲ ساله برای پیش بینی فصل اول

K-NN	روش پیشنهادی		
۰/۵۰	۰/۵۶	CORR	SOI
۲۳/۲	۲۴/۷	RMSE	
۲۸/۰	۲۷/۰	%VE	
۹/۶	۲۵/۱	LEPS	
-۰/۱۷	۰/۵۱	CORR	جریان تابستان
۲۹/۱	۲۸/۲	RMSE	
۳۶/۱	۳۲/۱	%VE	
۲/۳	۱۸/۱	LEPS	
۰/۵۸	۰/۶۲	CORR	جریان تابستان SOI+
۲۱/۳	۲۱/۸	RMSE	
۲۳/۷	۲۳/۰	%VE	
۱۲/۲	۳۱/۲	LEPS	

جدول ۲- نتایج صحت سنجی متقاطع ۳۲ ساله برای پیش بینی فصل دوم

K-NN	روش پیشنهادی		
۰/۱۴	۰/۲۵	CORR	NAO
۵۶/۶	۵۱/۱	RMSE	
۲۵/۲	۲۶/۲	%VE	
۸/۲	۱۹/۱	LEPS	
۰/۶۸	۰/۸۲	CORR	جریان زمستان + بودجه برفی
۴۳/۳	۳۴/۹	RMSE	
۱۸/۶	۱۲/۵	%VE	
۱۱/۴	۲۸/۱	LEPS	
۰/۷۳	۰/۹۲	CORR	جریان زمستان + بودجه برفی NAO +
۴۳/۰	۴۸/۴	RMSE	
۱۷/۲	۱۲/۳	%VE	
۱۳/۲	۳۰/۲	LEPS	

جدول ۳- نتایج صحت سنجی متقاطع به تفکیک سال‌های خشک و تر

K-NN		روش پیشنهادی			آماره	فصل
سال‌های تر	سال‌های خشک	سال‌های تر	سال‌های خشک	سال‌های خشک		
۰/۳۲	۰/۶۹	۰/۵۹	۰/۵۴	CORR	فصل اول	
۳۹/۰	۲۲/۳	۳۷/۰	۲۶/۴	RMSE		
۲۲/۱	۲۳/۳	۲۲/۲	۲۴/۱	%VE		
۲/۳	۹/۶	۱۴/۶	۳۹/۶	LEPS		
۰/۰	۰/۳۲	۰/۸۹	۰/۵۱	CORR	فصل دوم	
۷۲/۱	۳۹/۲	۳۲/۶	۳۲/۰	RMSE		
۱۲/۳	۱۹/۷	۶/۲	۱۵/۱	%VE		
۲۵/۲	۸/۱	۴۱/۳	۱۶/۱	LEPS		

جریان فصل دوم در منطقه مورد مطالعه قویا وابسته به شرایط هیدرولوژیکی منطقه است. استفاده از بودجه برفی و جریان مشاهده شده در زمستان نتایج پیش بینی خوبی را نتیجه داده است. به طوری که متوسط خطا در این حالت برابر با ۱۲/۵ درصد و آماره LEPSscore برابر با ۲۸/۱ بوده است. در عین حال در نظر گرفتن NAO به عنوان یک پیش بینی کننده دیگر نتایج پیش بینی را بهبود بخشیده و بهترین آماره های آزمون پیش بینی در همین حالت به دست آمده است. با افزودن NAO بهبود قابل توجهی در دقت پیش بینی حاصل نشده است اما قطعیت پیش بینی ۱۰ درصد افزایش یافته است. حاصل بیشترین متوسط آماره LEPSscore در این حالت نیز با به کارگیری مدل پیشنهادی و در نظر گرفتن هر سه پیش‌بینی کننده و برابر با ۳۰/۲ بدست آمده است.

نتایج ارائه شده در جدول (۳) نشان می دهد که در پیش بینی فصل اول نتایج به دست آمده در پیش بینی دوره‌های خشک دقیقتر و با قطعیت بیشتری نسبت به پیش بینی در دوره های تر بوده است به طوری که متوسط آماره LEPSscore برای دوره های خشک برابر با ۳۹/۶ بوده است که این مقدار برای دوره تر ۱۴/۶ بوده است. بر خلاف آن مدل های پیش بینی در فصل دوم دوره‌های تر را با دقت بیشتری پیش‌بینی کرده‌اند زیرا LEPSscore پیش‌بینی

نتایج ارائه شده توسط مدل ها بر اساس دقت و قطعیت تخمین با یکدیگر مقایسه می‌شوند. دقت پیش بینی بر اساس سه معیار RMSE، CORR، %VE و قطعیت پیش بینی بر اساس معیار LEPSscore مورد مقایسه قرار گرفته است.

با مراجعه به جدول ۱ ملاحظه می شود که در نظر گرفتن توام SOI و همبستگی بین اطلاعات جریان بهترین تخمین در مقایسه با سایر موارد (در نظر گرفتن پیش بینی کننده های منفرد) را نتیجه می دهد. کمترین RMSE و %VE، همچنین بیشترین ضریب همبستگی در این حالت به دست آمده است. از نتایج ارائه شده در این بخش بدست می آید که SOI هم دقت و هم قطعیت پیش بینی را در مقایسه با در نظر گرفتن تنها همبستگی بین جریان برای پیش بینی، افزایش می دهد و همانطور که در نتایج دیده می‌شود در نظر گرفتن SOI تا ۳۸ درصد آماره LEPSscore را در مقایسه با پیش بینی تنها بر اساس همبستگی بین اطلاعات افزایش داده است. به عبارتی قطعیت و دقت پیش بینی نسبت به استفاده تنها از همبستگی جریان، تا ۳۸ درصد افزایش پیدا کرده است. بیشترین متوسط LEPSscore در حالت استفاده از مدل پیشنهادی و برابر با ۳۱/۲ بدست آمده است که در زمره یک پیش بینی خوب دسته بندی می شود. در این حالت متوسط خطای پیش بینی ۳۲ ساله برابر با ۲۳ درصد بوده است.

هدف ثانویه مطالعات ارائه شده در این فصل این بود که آیا سیگنال های در نظر گرفته شده پیش بینی کننده های خوبی برای جریانهای فصلی رودخانه زاینده رود در محل ورودی به سد می باشند؟ نتایج نشان می دهد که SOI می تواند شاخصی برای تغییرات اقلیمی منطقه به ترتیب در نیمه اول سال آبی بوده و برای پیش بینی های بلند مدت مورد استفاده قرار بگیرد اما بهره گیری از NAO تاثیر چندانی در بهبود نتایج نداشته است. این اطلاعات علاوه بر افزایش اطمینان پذیری پیش بینی ها در این شرایط، می تواند راهنمایی برای یافتن پیش بینی کننده های جدید باشد. در نهایت از مزایای روش پیشنهادی نسبت به سایر رویکردهای مشابه برای تولید تابع توزیع پیش بینی موارد زیر است:

امکان در نظر گرفتن همزمان چند پیش بینی کننده و تولید تابع توزیع احتمالاتی پیوسته است.

پیش بینی سری های غیرایستا

تخمین متغیرها با انعکاس عدم قطعیت آن تخمین

از دیدگاه مدیریت کشاورزی این رویکرد تخمین می تواند به عنوان ابزاری برای افزایش کارایی بهره برداری و بهره برداری بر اساس ریسک مورد استفاده قرار گیرد.

دوره های تر فصل دوم برابر با $41/3$ در مقابل $16/1$ برای دوره های خشک بوده است. نتایج آماره های پیش بینی نقطه ای نیز این نتایج را تایید کرده است. این نتیجه گیری می تواند در تعیین پیش بینی کننده های مناسب دیگر برای این دو فصل مورد استفاده قرار بگیرد.

نتیجه گیری

در مطالعه حاضر رویکردی برای تخمین هیدروکلیماتولوژیک با استفاده از سیگنال های بزرگ مقیاس اقلیمی و شرایط هیدرولوژیکی یک حوضه ارائه شد. مدل ذکر شده توانایی بهره گیری از چند عامل ورودی و انجام پیش بینی به صورت توزیع احتمالاتی پیوسته شرطی به صورت صریح را دارد. در موارد کوتاه مدت یا میان مدت بودن اطلاعات، خطای استفاده از این مدل کمتر از خطای استفاده از مدل ناپارامتری $K-NN$ می باشد، که دارای الگوریتمی مشابه با این رویکرد است. نتایج پیش بینی هر دو فصل زمستان و بهار نشان می دهد که دقت و اطمینان مدل تخمین ارائه شده بهتر از مدل ناپارامتری نزدیکترین همسایه می باشد. این نتیجه در استفاده از تصمیم گیری های بر اساس ریسک در مدیریت کشاورزی بسیار حائز اهمیت است.

منابع

- خوش اخلاق، ف. ۱۳۷۷. پدیده انسو و تاثیر آن بر رژیم بارش ایران. فصلنامه تحقیقات جغرافیایی، شماره پیاپی ۵۱، سال ۱۳ شماره ۴.
- کیانی پور، م. ۱۳۷۹. بررسی سینوپتیکی پدیده النینو و ارتباط آن با ناهنجاری بارش های جنوب و جنوب غرب ایران. پایان نامه کارشناسی ارشد رشته جغرافیای طبیعی (هیدرواقلم)، دانشگاه تربیت مدرس، تهران.
- مدرس پور، آ. ۱۳۷۷. ناهنجاریهای اقلیمی ایران ENSO. پایان نامه کارشناسی ارشد فیزیک دریا، دانشگاه آزاد اسلامی واحد شمال، تهران.
- Box, G. E. P., Jenkins, F. M. 1976. Time Series Analysis: Forecasting and Control. Oakland, CA: Holden-Day, 2nd ed.
- Brockwell, P. J., Davis, R. A. 1987. Time series theory and Methods", Springer-Verlag, New York.
- Coulbali, P., Anctil, F., Bobée, B. 2000. Daily streamflow forecasting using neural networks with stopped training approach. J. Hydrol., 230: 244-257.
- Day, G. N. 1983. Extended Streamflow Forecasting Using NWSRFS. J. Water Res. Plan. Manag., 111(2):157-170.

- Garen, D. C. 1992. Improved techniques in regression-based streamflow volume forecasting. *J. Water Res. Plan. Manag.*, 118 (6): 654-670.
- Hsu, K. L., Gupta, H. V., Sorooshian, S. 1995. Artificial neural networks modeling of rainfall-runoff process, *Water Resour. Res.*, 31(10): 2517-2530.
- Karlsson, M., Yakowitz, S. 1987. Nearest-neighbor methods for nonparametric rainfall-runoff forecasting, *Water Resour. Res.*, 23(7):1300-1308.
- Krieg, D. G. 1966. Two-dimensional wighted moving average trend surfaces for ore-evaluation. *J. South Afr. Inst. Min. Metall.*, 66: 13-38.
- Lall, U., Sharma, A. 1996. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour. Res.*, 32(3): 679-694.
- Potts, J. M., Folland, C. K., Jolliffe, I. T., Sexton, D. 1996. Revised LEPS scores for assessing climate model simulations and long-range forecast. *J. Climatol.*, 9: 34-53.
- Webster, R., Oliver, M. A. 2001. *Geostatistics for Environmental Scientists*. John Wiley & Sons, Ltd., New York, NY.



A proposed method for uncertainty modeling in hydroclimatological estimation with emphasis on drought events

S. Araghinejad*¹

Received: 25-3-2013

Accepted: 14-9-2013

Abstract

Most techniques and methods in the “agricultural planning and management based on the climate variation and climate change” are implicitly or explicitly dependent to the hydroclimatological estimations. Hydroclimatological estimation is actually determination and warning of specific situations in a certain time and place such as drought and floods. Further improvements in the modeling for those estimations are up to the growing knowledge on the interaction between climatic and hydrological events as well as developing new models and improvement of existing ones. Presenting a hydroclimatological model to be able to map the nonlinear relationship between climate signals and hydrological models; to model the non stationary time series; and to provide probabilistic estimation could be a significant contribution. A model is presented in this paper based on the correlation analysis and statistical modeling which is compared with a conventional nonparametric model (nearest neighborhood) in point and interval estimation. One of the important challenges in hydroclimatological estimation is to calculate estimation uncertainty in which the proposed method deals with it. The contribution of this paper is to provide a simple but efficient method for hydroclimatological estimation in an uncertain environment. The case study of this research is forecasting of streamflow of Zayandeh-rud River in drought and normal situations. The results demonstrate the supremacy of the proposed model in comparison with the conventional models.

Keywords: Hydroclimatological estimation, Uncertainty, Geostatistical analysis, K-nearest neighborhood method

¹ Assistant Professor, Department of Irrigation and Reclamation Engineering, University of Tehran, Karaj
(*Corresponding Email Address: araghinejad@ut.ac.ir)