

تأثیر دو نگرش متفاوت فاصله‌ی دو مجموعه بر روی ریزمقیاس‌نمایی فراکتالی درجه حرارت

در مشهد

شیرین تاج آبادی^{۱*} - بیژن قهرمان^۲ - علی نقی ضیائی^۳

تاریخ دریافت: ۱۳۹۴/۱۰/۱۴

تاریخ پذیرش: ۱۳۹۵/۱۰/۲۰

چکیده

با استفاده از اطلاعات ریزمقیاس‌شده می‌توان نیاز به وجود داده در مکان و زمان مختلف را رفع نمود. فراکتال اخیراً برای ریزمقیاس‌سازی داده‌های اندازه‌گیری شده، مورد استفاده قرار گرفته است. در این تحقیق از توابع درون‌یاب فرکتال برای تولید داده‌های ریزمقیاس‌شده روزانه و سه ساعته ایستگاه سینوپتیک مشهد استفاده شد. همچنین دو نگرش متفاوت در محاسبه فاصله هاسدرف در تعیین نقاط درون‌یابی (محاسبه فاصله هاسدرف نگرش اول: با داده‌های استانداردسازی نشده، نگرش دوم: با داده‌های استانداردسازی شده) استفاده شد و سه فاصله نقاط درون‌یابی متفاوت ۵، ۱۰، ۱۵ روز در نظر گرفته شد. نتایج مربوط به ریزمقیاس‌سازی با فاصله درون‌یابی ۵ و ۱۰ روز و نگرش اول از دیگر نتایج مناسب‌تر بودند، به دلیل خطای کم بین نتایج فاصله درون‌یابی ۵ و ۱۰ روز و با توجه به اهمیت زمان اجرای برنامه و استفاده از داده‌های کمتر، فاصله درون‌یابی ۱۰ روز بهترین نتیجه را حاصل کرد. آزمون‌های آماری مقادیر آماره R^2 را برای نگرش اول بین ۰/۷۴-۰/۹۸، و نگرش دوم ۰/۶۹-۰/۹۸، RMSE را برای نگرش اول بین ۵/۱۲-۱/۳۳ و نگرش دوم ۵/۹-۱/۴۴ درجه سانتیگراد و معیار اطلاعاتی آکائیک AICc را برای نگرش اول بین ۳/۱۹-۰/۵۵ و نگرش دوم ۳/۴۶-۲/۸۷ نشان دادند و همچنین عرض از مبدأها و شیب‌های خطوط مدل‌سازی در سطح ۵ درصد تفاوت معنی‌داری به ترتیب با صفر و یک ندارند. بر اساس نتایج بدست آمده، ریزمقیاس‌سازی زمانی روزانه و سه ساعته با دقت و کیفیت قابل قبول انجام شده است و در نهایت نگرش اول نتایج بهتری را نسبت به نگرش دوم ارائه کرده است.

واژه‌های کلیدی: استانداردسازی، ریزمقیاس‌سازی، فاصله هاسدرف

مقدمه

دامنه نوسانات پارامترهای هواشناسی مانند دمای هوا، در مقیاس‌های مختلف، متفاوت می‌باشد. هندسه فرکتالی‌شده‌ای از ریاضیات است که در میدان‌های گسسته و متناوب کاربردهای فراوان دارد (۸)، که از آن جمله می‌توان به مدل‌سازی سری‌های زمانی کمیت‌های مختلفی مانند الکتروکاردیوگرام، لوگ چاه و نوسانات صوت (۵)، آنالیز داده‌های تلاطم (۱۱)، پیش‌بینی فرسایش در سطحی از جنس آلیاژ برنج (۱۴) و مدل‌سازی سیگنال‌ها (۲) اشاره کرد. یکی دیگر از کاربردهای هندسه‌ی فرکتال تولید داده‌هایی است که مقیاس آنها با داده‌های اندازه‌گیری متفاوت است، که از آن به ریزمقیاس‌نمایی یاد می‌شود (۷ و ۸). کاربردی مشخص برای مدل‌سازی سری زمانی ساعتی بارندگی در ژاپن توسط پاتیرانا (۷) گزارش شده که در آن سری زمانی بارندگی با مقیاس زمانی ۵ دقیقه از روی بارندگی ۲ روزه تولید شده است. وی همچنین پس از مدل‌سازی داده‌های بارندگی که با مقیاس زمانی دو روز ثبت شده بودند توانست به

به منظور شناسایی بهتر رفتار حوضه آبریز، دستیابی به اطلاعات صحیح متغیرهای هیدرولوژیکی که در زمان و مکان تغییرپذیرند و به عنوان متغیرهای تصادفی و پویا شناخته می‌شوند، بسیار سودمند است. دستیابی صحیح به این اطلاعات نیازمند اندازه‌گیری مناسب در مکان و زمان مناسب می‌باشد، که عموماً در عمل امکان‌پذیر نمی‌باشد. بدین ترتیب برای دستیابی به اطلاعات ریزمقیاس در زمان و مکان و به منظور اعمال مدیریت بهینه در حوضه آبریز، ریزمقیاس‌نمایی زمانی و مکانی خصوصیات هیدرولوژیکی حوضه‌ی ضروری است (۷).

۱، ۲ و ۳- به ترتیب دانش‌آموخته کارشناسی ارشد منابع آب، استاد و استادیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد
(* نویسنده مسئول: Email: Tajabadi.sh.1991@gmail.com
DOI: 10.22067/jsw.v31i1.52049

عنوان دمای روزانه انتخاب شد. شهر مشهد مرکز استان خراسان رضوی، در شمال شرق ایران و در طول جغرافیایی ۵۹ درجه و ۱۵ دقیقه تا ۶۰ درجه و ۳۶ دقیقه و عرض جغرافیایی ۳۵ درجه و ۴۳ دقیقه تا ۳۷ درجه و ۸ دقیقه واقع است. از نظر دمایی، استان خراسان رضوی دارای تابستانهای گرم و زمستانهای نسبتاً سرد است و میانگین درجه حرارت سالانه از شمال به جنوب افزایش می‌یابد. میانگین سالانه دمای استان خراسان رضوی ۱۵/۶ درجه سانتی‌گراد می‌باشد (۱۲).

در فرآیند مدل‌سازی سری زمانی داده‌های دما روزانه از مدل فرکتالی خودمشابه برای تولید سیستم توابع تکرار استفاده می‌شود. در این مدل فرض می‌شود که تغییرات در هر زیر مجموعه‌ای از مجموعه داده‌های یک کمیت از تغییرات در کل مجموعه داده‌های کمیت مورد مطالعه پیروی می‌کنند. بنابراین برای تعیین نقاط بهینه درون‌یابی به منظور مدل‌سازی مجموعه داده‌های کمیت مورد مطالعه، از تمامی نقاط در مجموعه داده‌ها استفاده می‌گردد. بدین صورت که اولین نقطه در مجموعه داده‌ها (یعنی روز اول سال) اولین نقطه‌ی درون‌یابی و آخرین نقطه در مجموعه‌ی داده‌ها (یعنی روز آخر سال) آخرین نقطه‌ی درون‌یابی در نظر گرفته می‌شود. تعیین نقاط درون‌یابی دیگر صرفاً با روش جستجو صورت می‌گیرد. در این حال کنترل تمامی نقاط بین دو نقطه‌ی درون‌یابی اول و آخر از نظر محاسباتی بسیار سنگین می‌باشد. به همین دلیل معمولاً گام زمانی برای بررسی نقاط را بیشتر از یک انتخاب می‌کنند. نقاط درون‌یابی نیز با استفاده از الگوریتم ارائه شده توسط میزل و هیز تعیین می‌گردند (۵). در این جا ما از گام‌های متفاوت ۵، ۱۰، ۱۵ روز استفاده و تاثیر آن را بر روی مدل‌سازی فرکتالی و ریزمقیاس‌سازی زمانی درجه حرارت نیز بررسی کردیم. بدین ترتیب نقاط درون‌یابی تعیین شده و پس از آن توابع درون‌یاب فرکتال (نسبی برشی) محاسبه می‌شوند و در نتیجه سیستم توابع تکرار برای مدل‌سازی کمیت مورد مطالعه تولید می‌گردد. تابع درون‌یاب نسبی W در فضای دو بعدی به صورت رابطه (۱) تعریف می‌شود. در این معادله W تابع نسبی، $\begin{pmatrix} x \\ y \end{pmatrix}$ مختصات هر نقطه از فضای دو بعدی و ضرایب a, b, c, d, e, f پارامترهای تابع و اعداد حقیقی می‌باشند. پارامتر d را فاکتور مقیاس عمودی تابع می‌نامند. تابع نسبی برشی در واقع همان تابع نسبی است با این تفاوت که مقدار پارامتر b در این تابع معادل صفر در نظر گرفته می‌شود. در توابع نسبی برشی، پارامتر d (فاکتور مقیاس عمودی) به عنوان پارامتر آزاد برای تابع تعریف می‌شود. (۱).

$$W \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = AX + t \quad (1)$$

هر تابع درون‌یاب به‌طور مجزا بر روی تمامی نقاط در مجموعه داده‌های کمیت مورد مطالعه اعمال شده و تمام نقاط این مجموعه را به فاصله‌ی میان دو نقطه درون‌یابی متناظر با هر تابع درون‌یاب

ریزمقیاس‌نمایی سری زمانی با مقیاس دو ساعت پردازد. ریزمقیاس‌نمایی با روش‌های مختلفی از جمله تابع رگرسیون یک متغیره، تابع رگرسیون چند متغیره، تابع اسپیلاین و تابع فراکتال انجام می‌شود (۴). ریزمقیاس‌نمایی با تابع فراکتال به این صورت انجام می‌شود که پس از بررسی داده‌های موجود برخی نقاط که پتانسیل بهتری برای مدل کردن کل نقاط دارند (نقاط ثابت) انتخاب شده و با کمک توابع درون‌یابی در هندسه فراکتالی (۱)، کل مجموعه را مدل می‌کنند. جهت تعیین بهترین نقاط درون‌یابی می‌باید برای تمامی نقاط پتانسیل، کمترین فاصله بین مجموعه نقاط اندازه‌گیری شده و مدل شده جستجو شود. این فاصله با روش‌های مختلفی، از جمله اقلیدوسی، برآورد می‌شود (۱). در ریزمقیاس‌نمایی زمانی، داده‌ها دو بعدی هستند به طوری که یکی از ابعاد زمان و بعد دیگر مربوط به واحدی است که پارامتر مورد نظر (برای مثال درجه حرارت) اندازه‌گیری می‌شود. در این صورت در فضای اقلیدوسی مشکل ابعادی به وجود می‌آید. در چنین شرایطی به‌طور معمول داده‌ها بدون بعد می‌شوند که از آن به استانداردسازی، نرمال‌سازی، مقیاس‌سازی یا بی‌بعدسازی هم یاد می‌شود. یک کاربرد مشخص در این زمینه، مربوط به خوشه‌بندی است که در آن مجموعه نقطه‌ها به خوشه‌هایی همگن افزاز می‌شوند. هر نقطه مشخصاتی متفاوت و با ابعاد متفاوت دارد. جهت سنجش و جلوگیری از غلبه کردن یکی از مشخصه‌ها بر روی سایر مشخصه‌ها، تمامی داده‌ها بدون بعد می‌شود (برای مثال شامکوئیان و همکاران (۱۰)، شاهدی و همکاران (۹)، قهرمان و داوری (۳)، به خوشه‌بندی هیدرولوژیکی حوضه‌های آبریز پرداخته‌اند). با این حال منابع موجود به‌طور مثال برنسلی (۱)، که فراکتال و کاربردهای آن را توضیح داده، و ولیدی و همکاران (۱۳) که با استفاده از توابع درون‌یاب فرکتال درجه حرارت روزانه را به ۳ ساعته ریزمقیاس کرده‌اند، به مبحث استانداردسازی داده‌ها اشاره‌ای نکرده‌اند.

بنابراین جدای از این که استاندارد نکردن داده‌ها موجب عدم توازن در روش‌های محاسبه‌ی دو مجموعه ایجاد می‌شود، در این مقاله به تاثیر استاندارد سازی داده‌ها بر روی تعداد نقاط درون‌یابی، زمان اجرای برنامه و نیز دقت در ریزمقیاس‌سازی دمای ایستگاه سینوپتیک مشهد پرداخته شده است.

مواد و روش‌ها

منطقه مورد مطالعه و داده‌ها

در این تحقیق از داده‌های ایستگاه سینوپتیک مشهد استفاده شد. داده‌های ایستگاه سینوپتیک مشهد به صورت روزانه و سه ساعته از سال ۲۰۰۷ تا ۲۰۰۹ میلادی تهیه و داده‌های دمای ساعت ۱۵ به

نیز به همان روشی که در قسمت فوق بیان شد عمل می‌گردد. چون در فضای متریک هاسدرف، تابع فاصله خاصیت جابجایی ندارد، مقادیر فاصله محاسبه شده میان مجموعه Z از W و W از Z متفاوت خواهد بود. با توجه به رابطه ۶ از میان دو عدد محاسبه شده، بیشترین مقدار به عنوان فاصله میان دو مجموعه بیان می‌گردد. در این رابطه، عملگر V بیانگر بیشترین مقدار می‌باشد (۱).

$$d(z, w) = \text{Max}\{d(x, w): x \in z\} \quad (۳)$$

$$d(x, w) = \text{Min}\{d(x, y): x \in z, y \in w\} \quad (۴)$$

$$d(w, z) = \text{Max}\{d(y, z): y \in w\} \quad (۵)$$

$$h(z, w) = d(z, w) \vee d(w, z) \quad (۶)$$

معیار ارزیابی صحت مدل

در بررسی‌های آماری از پارامتر ریشه دوم میانگین مربعات خطا، که معیاری برای بررسی میزان صحت مدل در ارائه نتایج می‌باشد، به عنوان شاخص واریانس بین مقادیر دمای اندازه‌گیری شده (T_{obs}) و دمای مدل سازی شده ($T_{\text{simulated}}$) (رابطه ۷) استفاده می‌شود. همچنین از مقایسه شیب و عرض از مبدأ خط رگرسیون بین مقادیر دمای اندازه‌گیری شده و دمای مدل سازی شده، به ترتیب با یک و صفر، به عنوان معیاری برای ارزیابی استفاده شده است.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (T_{\text{obs}} - T_{\text{simulated}})^2} \quad (۷)$$

همچنین از معیار اطلاعات آکائیک (AIC) (رابطه ۸) به عنوان یک پارامتر آماری برای مقایسه بین مدل‌ها استفاده شده است. AIC معیاری برای سنجش نیکویی برازش است که بر اساس مفهوم آنتروپی بنا شده و نشان می‌دهد که استفاده از مدل آماری به چه میزان باعث از دست رفتن اطلاعات می‌شود. به عبارت دیگر، این معیار تعادلی میان دقت مدل و پیچیدگی آن برقرار می‌کند. با توجه به داده‌ها، چند مدل متفاوت را با توجه به مقدار AIC رتبه بندی کرده و مدل دارای کمترین AIC را انتخاب می‌کنند. این معیار را می‌توان این‌گونه محاسبه کرد (۶):

$$AIC_c = \ln\left(\frac{RSS}{x}\right) + \frac{x + y}{x - y - 2} \quad (۸)$$

در این رابطه x : تعداد جفت داده‌ها، y : تعداد پارامترهای ورودی مدل و RSS : مجموع مربعات باقی‌مانده می‌باشند.

نتایج و بحث

شکل ۱ روند تغییرات دمای روزانه را برای سه سال مختلف ۲۰۰۷ تا ۲۰۰۹ نشان می‌دهند. واضح است که سری زمانی داده‌های

انتقال می‌دهد. بدین ترتیب مدل سازی مجموعه داده‌های کمیت مورد مطالعه انجام می‌پذیرد. در این تحقیق مدل سازی داده‌ها و پس از آن ریزمقیاس سازی، به دو روش انجام می‌شود، در روش اول از داده‌های اصلی (زمان و دما) استفاده شده، نقاط درون‌یابی انتخاب شده، مدل سازی اولیه انجام شده، فاصله‌ی بین دو مجموعه مدل سازی و اندازه‌گیری شده محاسبه شده، بهترین نقاط درون‌یابی انتخاب و مدل سازی نهایی و ریزمقیاس سازی صورت گرفته است. در روش دوم در ابتدا داده‌ها را به صورت استاندارد تبدیل کرده و پس از آن ادامه روند تعیین نقاط درون‌یابی و محاسبه فاصله دو مجموعه، مدل سازی و در نهایت ریزمقیاس سازی انجام شده است. برای استانداردسازی داده‌ها از رابطه (۲) استفاده شده است:

$$X(i)_{\text{new}} = \frac{X(i) - X(\min)}{X(\max) - X(\min)} \quad (۲)$$

که در آن $X(i)$ مقدار داده‌ی i -ام (به حسب مورد، زمان یا دما) و $X(\min)$ و $X(\max)$ به ترتیب کوچک‌ترین و بزرگ‌ترین داده در سری داده‌ها می‌باشد. در نتیجه بیشترین مقدار داده به یک و کمترین مقدار آن به صفر تبدیل شده و مقادیر دیگر بین صفر و یک قرار می‌گیرند.

محاسبه فاصله دو مجموعه از یکدیگر

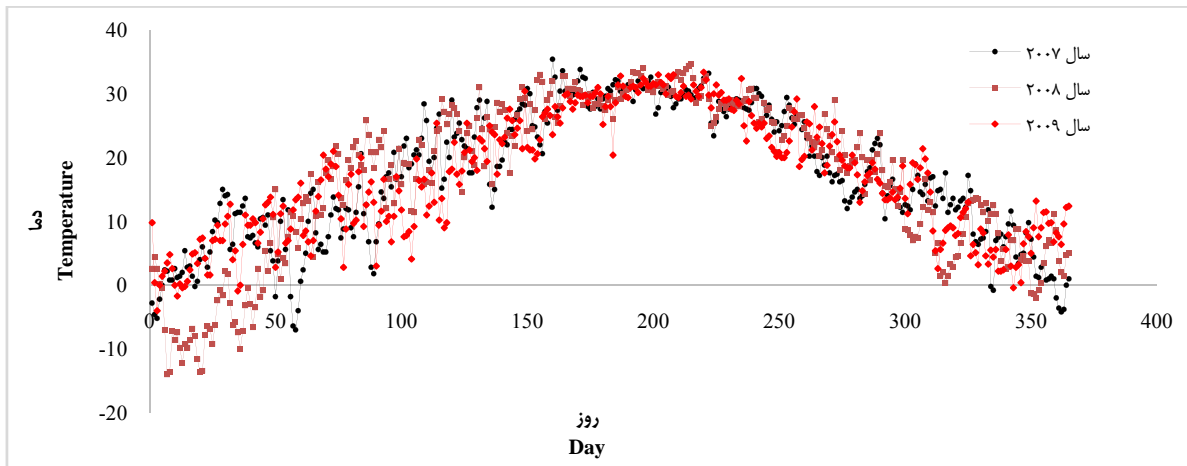
در مدل سازی داده‌ها با استفاده از نظریه‌ی فرکتال و با توجه به قضیه کولاج، سیستم توابع تکراری برای مدل سازی آن مجموعه از داده‌ها به عنوان بهترین سیستم توابع تکرار انتخاب می‌گردد که اختلاف میان مجموعه داده‌های اندازه‌گیری شده و مجموعه داده‌های حاصل از مدل سازی با آن سیستم توابع تکرار، کمترین مقدار باشد (۱). بدین ترتیب برای محاسبه فاصله دو مجموعه داده‌های اندازه‌گیری شده و داده‌های مدل سازی شده از روش فاصله هاسدرف استفاده شده است.

برای محاسبه فاصله هاسدرف میان دو مجموعه Z و W (به ترتیب مجموعه داده‌های اندازه‌گیری شده و مجموعه داده‌های مدل سازی شده)، ابتدا باید فاصله مجموعه Z را از مجموعه W ، حساب نمود (رابطه ۳). برای محاسبه فاصله مجموعه Z از W ، ابتدا باید فاصله هر عضو از مجموعه Z (مانند x_1, x_2, \dots, x_N) با تک تک اعضا در مجموعه W محاسبه گردد (با استفاده از متریک اقلیدسی یا من‌هتان). آنگاه از میان مقادیر به دست آمده، کمترین مقدار به عنوان فاصله نقطه مورد نظر (مانند x_1) از مجموعه Z تا مجموعه W تعریف می‌شود (رابطه ۴). این عمل برای سایر نقاط در مجموعه Z تکرار شده و سپس از میان مقادیر حداقل به دست آمده، بیشترین مقدار به عنوان فاصله مجموعه Z از W در نظر گرفته می‌شود.

سپس با استفاده از رابطه ۵، به محاسبه فاصله مجموعه W از مجموعه Z پرداخته می‌شود. برای محاسبه فاصله مجموعه W از Z

نقاط درون‌یابی منجر به تغییر در روند و تعداد نقاط درون‌یابی و همچنین تغییر آماره‌های خطا گردید.

دما به صورت سیکل سینوسی تغییر می‌کند. نمودار تغییرات دما با کمی اختلاف برای سال‌های مختلف تقریباً مشابه است. به طور کلی تغییر در شیوه‌ی استانداردسازی و تغییر در گام زمانی برای انتخاب



شکل ۱- روند تغییرات دمای روزانه مشهد برای سه سال متفاوت

Figure 1- Changing procedure daily temperature of Mashhad for three different years

نگرش بدون استانداردسازی نسبت به نگرش با استانداردسازی بیشتر بود. یعنی علاوه بر اینکه تعداد نقاط درون‌یابی در نگرش بدون استانداردسازی نسبت به نگرش با استانداردسازی بیشتر است، تغییرات فاصله بین نقاط درون‌یابی هم بیشتر و پراکندگی آن نیز نامنظم‌تر است. در کلیه‌ی حالات (سال و نیز گام فاصله‌ی درون‌یابی) استفاده از استانداردسازی منجر به کاهش تعداد نقاط درون‌یابی گردید. به نظر می‌رسد با انجام استانداردسازی داده‌ها به یکدیگر نزدیکتر شده‌اند و در بازه تغییر کمتری قرار می‌گیرند، در نتیجه در بین دو نقطه درون‌یابی متوالی تعداد نقاط بیشتری قرار می‌گیرد، پس می‌توان با نقاط درون‌یابی کمتری مدل‌سازی را انجام داد. جزئیات تحقیق و عملکرد دو نگرش، کاهش تعداد نقاط درون‌یابی در روند مثبت یا منفی نتایج ریزمقیاس‌سازی در دو بخش جداگانه ریزمقیاس‌سازی روزانه و سه ساعته بررسی شده است.

ریزمقیاس‌سازی برای داده‌های روزانه

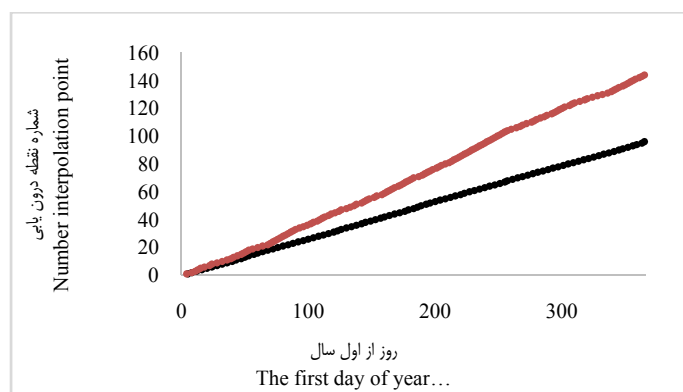
در این بخش به کمک داده‌های سال‌های میلادی ۲۰۰۷ تا ۲۰۰۹ به ریزمقیاس‌سازی زمانی روزانه دمای این سه سال پرداخته شد. در ابتدا از داده‌هایی با فواصل ۵ روز، دمای روزانه مدل‌سازی شد. شکل ۳ دمای شبیه‌سازی شده‌ی روزانه‌ی ایستگاه مشهد برای سال ۲۰۰۹ الف با فاصله درون‌یابی ۵ و بدون استانداردسازی و ب با فاصله درون‌یابی ۵ و با استانداردسازی را نشان می‌دهند.

با توجه به اینکه هدف مقایسه دو نگرش است، در هر دو نگرش دماهای پیش‌بینی شده به خوبی رفتار دماهای اندازه‌گیری شده را

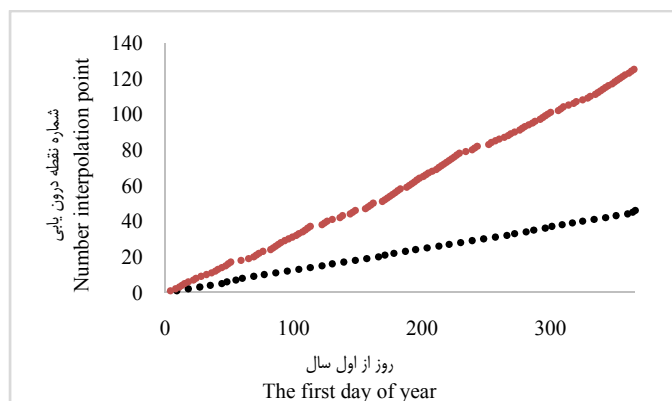
همانطور که بیان شد در این تحقیق با استفاده از توابع درون‌یاب فرانتال اقدام به مدل‌سازی سری زمانی دمای شهر مشهد برای سه سال میلادی ۲۰۰۷ تا ۲۰۰۹ و ریزمقیاس‌سازی زمانی روزانه و سه ساعته با دو نگرش استاندارد کردن و استاندارد نکردن داده‌ها بر فاصله هاسدرف دو مجموعه در تعیین نقاط درون‌یابی صورت گرفته است. همچنین همانطور که بیان شده فاصله بین نقاط درون‌یابی را به صورت متغیر ۵، ۱۰، ۱۵ در نظر گرفته شده است. در شکل ۲ تعداد نقاط درون‌یابی و روند تغییرات آنها در دو نگرش بدون استانداردسازی و با استانداردسازی برای سال ۲۰۰۹ و برای دو فاصله درون‌یابی مختلف ۵ و ۱۰ روز را نشان می‌دهد. کاهش تعداد نقاط درون‌یابی منجر به تعداد کمتری از توابع درون‌یابی شده که تاثیری مستقیم بر روی کاهش زمان اجرای برنامه‌ی مدل‌سازی فرکتال می‌شود. کاهش اجرای برنامه از موضوعاتی است که هنوز راه حل مشخصی برای آن وجود نداشته و تعداد نقاط درون‌یابی هنوز هم با سعی و خطا تعیین می‌شود. علاوه بر تعداد بازه‌های درون‌یابی، چگونگی توزیع نقاط درون‌یابی هم دارای اهمیت است. به طوری که هرچه توزیع مکانی نقاط درون‌یابی همگن‌تر باشد منجر به توزیع برابر دقت در ریزمقیاس‌سازی می‌گردد. ما در این جا این توزیع را با واریانس طول بازه‌های درون‌یابی تحلیل کردیم. در جدول ۱ تعداد نقاط درون‌یابی و واریانس طول بازه‌های درون‌یابی برای هر دو نگرش و فواصل درون‌یابی مختلف آورده شده است. واریانس طول بازه‌های درون‌یابی، تغییرات اختلاف فاصله نهایی نقاط درون‌یابی در یک سال را نشان می‌دهد. با توجه به جدول ۱، واریانس طول بازه‌های درون‌یابی در

آماره‌ی RMSE، R^2 و AICc، نگرش بدون استانداردسازی را با کمی اختلاف مناسب‌تر نشان داده‌اند ولی تفاوت چشمگیر نمی‌باشد.

تقلید کرده‌اند. پراکنش نقاط تقریباً یکسان به نظر می‌رسد. از این رو در جدول ۲ آماره‌های صحت و دقت مدل آورده شده‌اند. گرچه هر سه



الف



ب

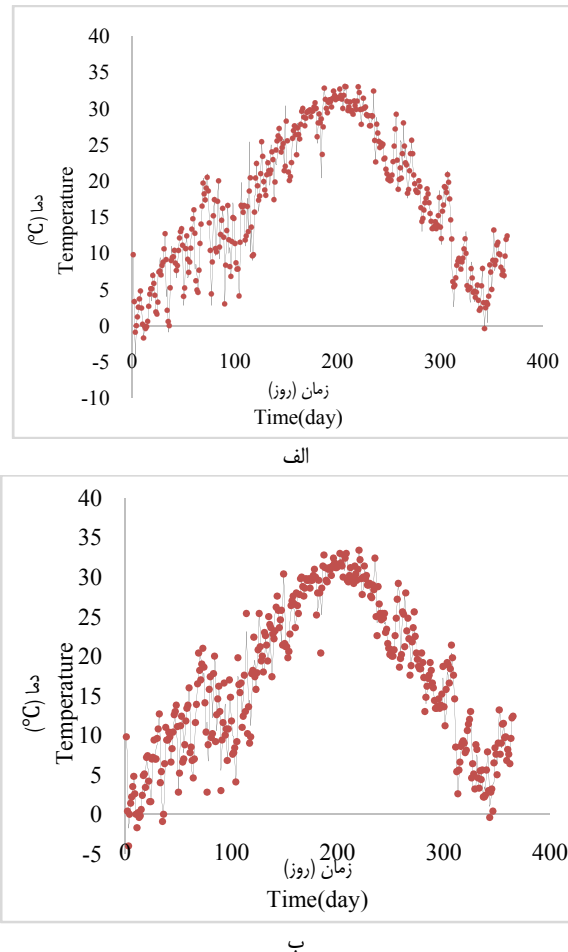
شکل ۲ - تعداد نقاط درون‌یابی و روند تغییرات آنها در دو نگرش مختلف استانداردسازی (رنگ آبی) و استاندارد نکردن (رنگ نارنجی) داده‌ها برای سال میلادی ۲۰۰۹ الف: فاصله درون‌یابی ۵ روز و ب: فاصله درون‌یابی ۱۰ روز

Figure 2- Number of interpolation points and trends in the two different approaches for Standardization (blue) and not standardization (orange) data for the year 2009 (a) interpolated 5 days and (b) interpolated 10 days

جدول ۱ - تعداد نقاط درون‌یابی و واریانس اختلاف فاصله نقاط درون‌یابی در نگرش‌های مختلف استانداردسازی

Table 1- The number of interpolation point and variance of differences between interpolation point in different approaches of Standardization

نگرش approach	فاصله interval سال year	فاصله درون‌یابی ۵ 5-interval			فاصله درون‌یابی ۱۰ 10-interval			فاصله درون‌یابی ۱۵ 15-interval		
		2007	2008	2009	2007	2008	2009	2007	2008	2009
بدون استانداردسازی Without standardization	تعداد نقاط درون‌یابی Number of interpolation point	119	128	144	80	107	125	43	54	72
	واریانس طول بازه‌های درون‌یابی Variance of interval interpolation	0.8	0.81	0.56	6.37	3.1	2.3	7.24	4.68	4.21
با استانداردسازی With standardization	تعداد نقاط درون‌یابی Number of interpolation point	95	96	94	46	48	48	35	38	37
	واریانس طول بازه‌های درون‌یابی Variance of interval interpolation	0.25	0.23	0.21	3.32	2.5	1.8	4.7	3.31	3.16



شکل ۳- دمای روزانه شبیه‌سازی شده‌ی سال ۲۰۰۹ با فاصله درون‌یابی ۵ روز و بدون استانداردسازی (الف) و با استانداردسازی (ب). دمای مدل‌سازی شده با خط ممتد و دمای اندازه‌گیری شده با نقطه نشان داده شده است

Figure 3- Simulated temperature of Mashhad year 2009 with 5-interval interpolation (a) without standardization (b) with standardization. Modeled data with stretch and observed data with points

نشان می‌دهد. در این شکل‌ها نقاط تقریباً بر روی نیمساز ربع اول واقع شده که نشان دهنده نزدیک بودن نقاط شبیه‌سازی به نقاط اندازه‌گیری می‌باشد. عرض از مبدأ و شیب این خطوط در دو مدل‌سازی در سطح ۵ درصد تفاوت معنی‌داری به ترتیب با صفر و یک نداشتند که نشان‌دهنده‌ی مدل‌سازی مناسب می‌باشد. با این حال در دو شکل ۵-ب و ۶-ب که به ترتیب مربوط به استانداردسازی با فاصله نقاط درون‌یابی ۵ و ۱۰ می‌باشد، نسبت به شکل‌های ۵-الف و ۶-الف که به ترتیب مربوط به استانداردسازی نشده با فاصله نقاط درون‌یابی ۵ و ۱۰ می‌باشد، پراکندگی نقاط نسبت به نیمساز اول کمی بیشتر می‌باشد. به طور کل از جدول ۱ و ۲ و شکل‌های ۵ و ۶ نتیجه می‌شود که به رغم کاهش تعداد نقاط درون‌یابی و زمان اجرای برنامه در نگرش با استانداردسازی، معیارهای صحت و دقت مدل در حالت نگرش بدون استانداردسازی، با اختلافی اندک نتایج بهتری را نشان داده‌اند. همچنین در تعیین فاصله نقاط درون‌یابی بهینه، فاصله

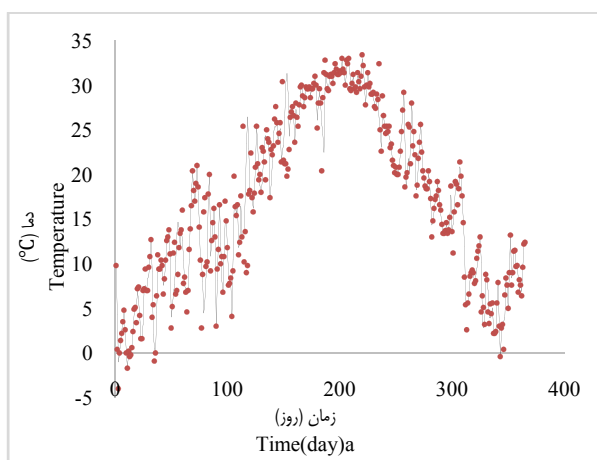
فاصله درون‌یابی را به ۱۰ روز تغییر داده و از داده‌هایی با فواصل ده روز، دمای روزانه را مدل‌سازی نمودیم و نتایج دمای شبیه‌سازی شده‌ی روزانه‌ی ایستگاه مشهد برای سال ۲۰۰۹ توسط فراکتال را با هر دو نگرش بدون استانداردسازی و با استانداردسازی در شکل ۴ارایه کردیم. تقریباً همان رفتار در شکل ۳ (مربوط به فاصله‌ی درون‌یابی ۵ روزه) در این جا نیز وجود دارد و این که تفاوت بین دو نگرش محسوس نمی‌باشد. مقایسه‌ی دقیق‌تر آماری در جدول ۲ انجام شده است. در جدول شماره ۲ آماره‌های صحت و دقت مدل بیان شده‌اند. روند تغییر آماره‌ها برای فاصله درون‌یابی ۱۰ روز مشابه فاصله درون‌یابی ۵ روز می‌باشد. همچنین شبیه‌سازی با نگرش بدون استانداردسازی برای هر دو فاصله درون‌یابی نتایج بهتری را نسبت به نگرش دوم با استانداردسازی نشان می‌دهد. شکل‌های ۵ و ۶ مقایسه‌ی تطبیقی بین دماهای روزانه‌ی شبیه‌سازی شده و اندازه‌گیری شده برای ایستگاه مشهد را

ریزمقیاس‌سازی سه ساعته هم انجام شده است. شکل ۷ دمای شبیه‌سازی شده‌ی روزانه‌ی ایستگاه مشهد برای سال ۲۰۰۹ و با فاصله درون‌یابی ۵ الف بدون استانداردسازی و ب با استانداردسازی را نشان می‌دهند. همچنین این مدل‌سازی برای فاصله درون‌یابی ۱۰ و ۱۵ روز هم انجام شده است. روند مشابه درون‌یابی ۵ روزه در فواصل دیگر درون‌یابی هم تکرار شده است. با توجه به مناسب بودن هر دو مدل‌سازی و هر دو نگرش، بازهم نگرش اول بدون استانداردسازی با کمی اختلاف نتایج بهتری را نسبت به نگرش دوم با استانداردسازی نشان می‌دهد.

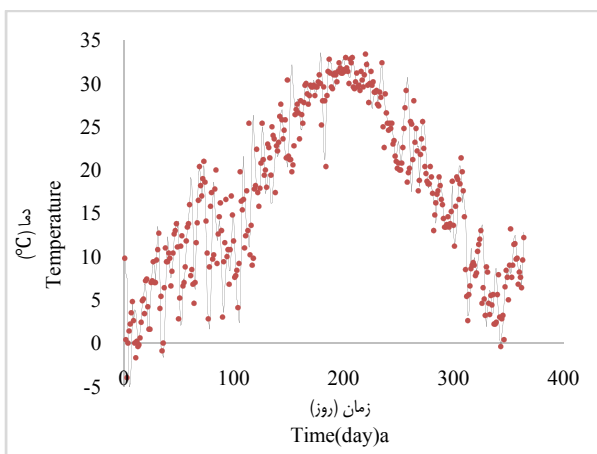
درون‌یابی پنج روزه با اندکی اختلاف نسبت به ده روز نتایج بهتری را نشان می‌دهد، با این حال با توجه به اینکه فاصله بیشتر بین نقاط درون‌یابی منجر به کاهش تعداد نقاط درون‌یابی و کاهش زمان اجرای برنامه می‌شود، می‌توان فاصله درون‌یابی ده روزه را به عنوان فاصله درون‌یابی بهینه قلمداد کرد.

ریزمقیاس‌سازی سه ساعته

در این بخش به کمک داده‌های سال‌های میلادی ۲۰۰۷ تا ۲۰۰۹ به ریزمقیاس‌سازی زمانی سه ساعته دمای این سه سال پرداخته شده است. در ابتدا از داده‌هایی با فاصله پنج روز، دمای سه ساعته را مدل‌سازی نمودیم. مانند ریزمقیاس‌سازی روزانه، مدل‌سازی و



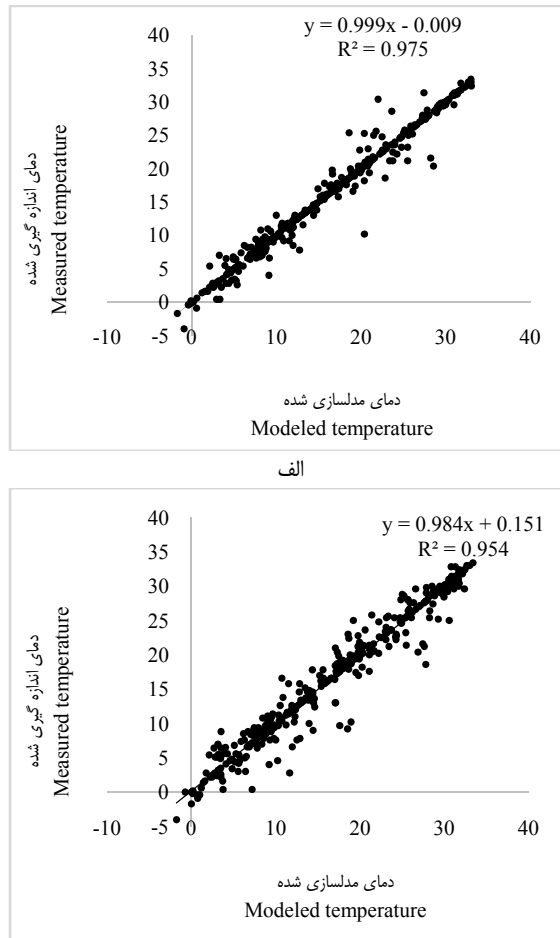
الف



ب

شکل ۴- دمای روزانه شبیه‌سازی شده‌ی سال ۲۰۰۹ با فاصله درون‌یابی ۱۰ روز و بدون استانداردسازی (الف) و با استانداردسازی (ب). دمای مدل‌سازی شده با خط ممتد و دمای اندازه‌گیری شده با نقطه نشان داده شده است

Figure 4- Simulated temperature of Mashhad year 2009 with 10-interval interpolation (a) without standardization (b) with standardization. Modeled data with stretch and observed data with points



ب

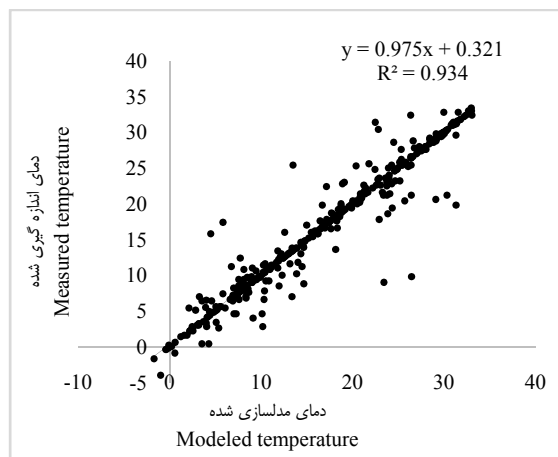
شکل ۵- مقایسه مقادیر دمای اندازه‌گیری شده با مقادیر شبیه‌سازی شده در مقیاس زمانی روزانه در سال ۲۰۰۹ با فاصله درونیابی ۵ روز و بدون استانداردسازی (الف) و با استانداردسازی (ب)

Figure 5- Comparison of measured and simulated daily temperature in year 2009 with 5-interval interpolation (a) without standardization (b) with standardization

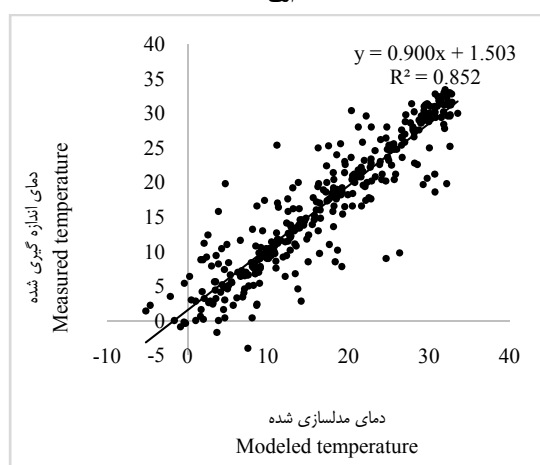
جدول ۲- معیارهای ارزیابی صحت و دقت در مدل فراکتال برای مدل سازی داده های دمای روزانه در ایستگاه مشهد

Table 2- Evaluation of fractal model for daily temperature modeling in Mashhad

نگرش Approach	فاصله interval سال Year	فاصله درونیابی ۵ 5-interval			فاصله درونیابی ۱۰ 10-interval			فاصله درونیابی ۱۵ 15-interval		
		2007	2008	2009	2007	2008	2009	2007	2008	2009
بدون استانداردسازی Without standardization	<i>F</i>	1.44	1.49	1.48	2.25	1.76	2.45	2.36	1.89	2.51
		0.98	0.98	0.97	0.95	0.97	0.93	0.91	0.93	0.89
	AICc	0.55	0.8	0.76	1.61	1.11	1.76	1.94	2.02	2.16
با استانداردسازی With standardization	<i>F</i>	1.44	1.76	2.05	2.51	3.27	3.8	2.6	3.32	3.95
		0.98	0.97	0.95	0.94	0.89	0.85	0.88	0.83	0.86
	AICc	0.7	1.14	1.41	1.82	2.35	2.65	2.13	2.41	2.81



الف



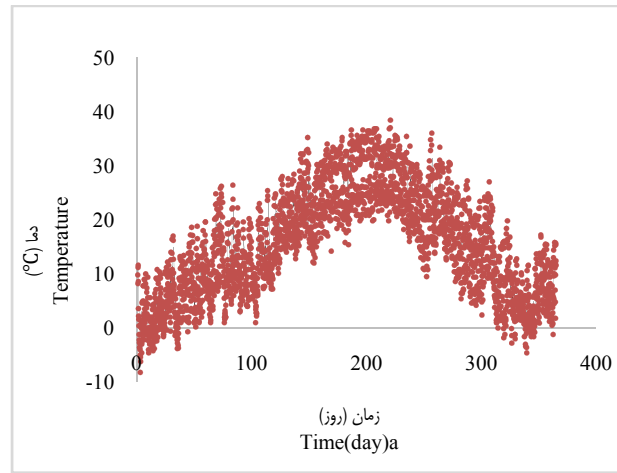
ب

شکل ۶- مقایسه مقادیر دمای اندازه‌گیری شده با مقادیر شبیه‌سازی شده به روش فراکتال در مقیاس زمانی روزانه در سال ۲۰۰۹ با فاصله درون‌یابی ۱۰ روز و بدون استانداردسازی (الف) و با استانداردسازی (ب)

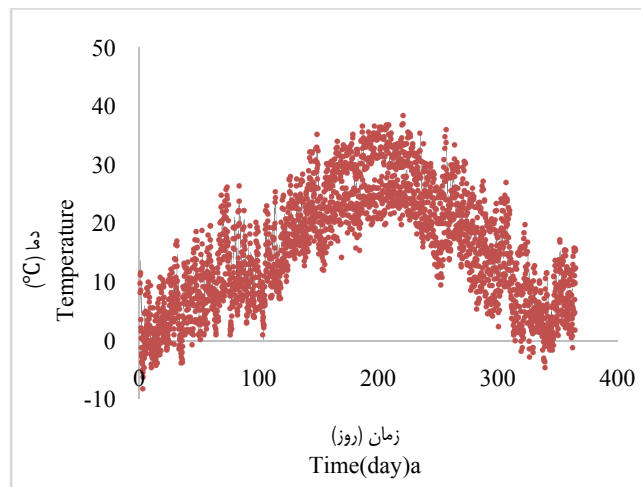
Figure 6- Comparison of measured and simulated daily temperature in year 2009 with 10-interval interpolation (a) without standardization (b) with standardization

در شکل ۸ مقایسه‌ی تطبیقی بین دماهای روزانه‌ی شبیه‌سازی شده و اندازه‌گیری شده سال ۲۰۰۹ برای ایستگاه مشهد انجام شده است. در این شکل نقاط تقریباً بر روی نیمساز ربع اول واقع شده، عرض از مبدأ و شیب این خطوط در دو مدل‌سازی در سطح ۵ درصد تفاوت معنی‌داری به ترتیب با صفر و یک ندارند. اختلاف فاصله بین نقاط اندازه‌گیری شده و مدل‌سازی شده در فاصله درون‌یابی ۱۰ و ۱۵ روز هم مشابه روند در فاصله درون‌یابی ۵ روز می‌باشد. مانند ریزمقیاس‌سازی روزانه نقاط در نگرش دوم نسبت به نگرش اول پراکندگی بیشتری دارند. در جدول ۳ هم مانند جدول ۲ هر سه آماره‌ی $RMSE$ ، R^2 و $AICc$ ، نگرش بدون استانداردسازی را با کمی اختلاف مناسب‌تر نشان داده‌اند. به نظر می‌رسد که استانداردسازی و به عبارتی قرارگیری داده‌ها در بازه تغییر کمتر تاثیر مثبتی بر روی خاصیت فرکتال ندارد و خاصیت طبیعی فرکتال را کمی برهم می‌زند. به طور کلی واضح است که نتایج حاصل از ریزمقیاس‌سازی سه ساعته عیناً نتایج ریزمقیاس‌سازی روزانه را نشان می‌دهند و این رویداد را می‌توان چنین تحلیل نمود که در ریزمقیاس‌سازی سه ساعته عیناً رخداد ریزمقیاس‌سازی روزانه در مقیاس کوچکتر (خاصیت فرکتال) رخ می‌دهد، پس طبیعتاً نتایج باید در مقیاس کوچکتر هم تکرار شوند و روند یکسانی را نشان دهند، گرچه خطا مقداری افزایش می‌یابد.

اختلاف مناسب‌تر نشان داده‌اند. به نظر می‌رسد که استانداردسازی و به عبارتی قرارگیری داده‌ها در بازه تغییر کمتر تاثیر مثبتی بر روی خاصیت فرکتال ندارد و خاصیت طبیعی فرکتال را کمی برهم می‌زند. به طور کلی واضح است که نتایج حاصل از ریزمقیاس‌سازی سه ساعته عیناً نتایج ریزمقیاس‌سازی روزانه را نشان می‌دهند و این رویداد را می‌توان چنین تحلیل نمود که در ریزمقیاس‌سازی سه ساعته عیناً رخداد ریزمقیاس‌سازی روزانه در مقیاس کوچکتر (خاصیت فرکتال) رخ می‌دهد، پس طبیعتاً نتایج باید در مقیاس کوچکتر هم تکرار شوند و روند یکسانی را نشان دهند، گرچه خطا مقداری افزایش می‌یابد.



الف



ب

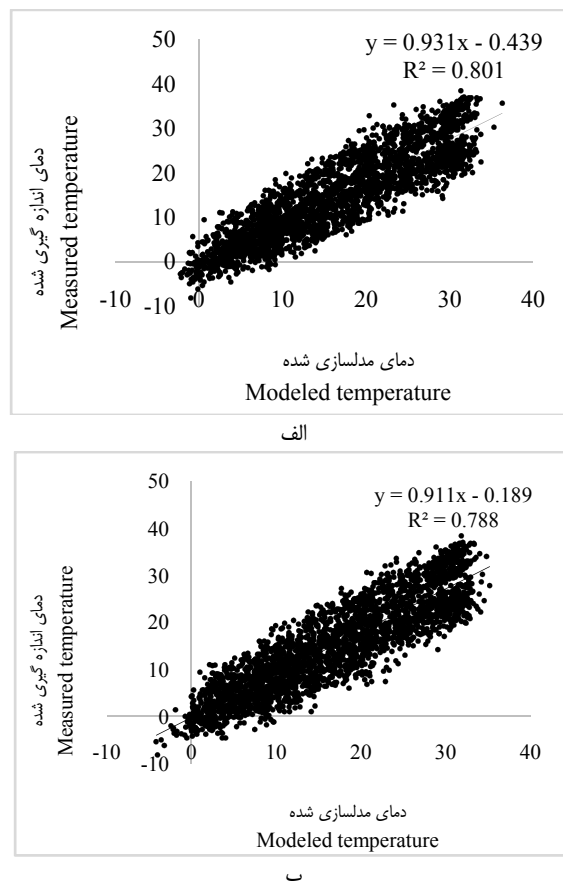
شکل ۷- دمای سه ساعته شبیه‌سازی شده‌ی سال ۲۰۰۹ با فاصله درونیابی ۵ روز و الف: بدون استانداردسازی و ب با استانداردسازی. دمای مدل‌سازی شده با خط ممتد و دمای اندازه‌گیری شده با نقطه نشان داده شده است

Figure 7- Simulated 3-hour temperature of year 2009 with 5-interval interpolation (a) without standardization (b) with standardization. Modeled data with stretch and observed data with points

جدول ۳- معیارهای ارزیابی صحت و دقت در مدل فراکتال برای مدل‌سازی داده‌های سه ساعته در ایستگاه مشهد

Table 3- Evaluation of fractal model for downscaling 3-hours temperature modeling in Mashhad

نگرش Approach	فاصله interval سال Year	فاصله درونیابی ۵ 5-interval			فاصله درونیابی ۱۰ 10-interval			فاصله درونیابی ۱۵ 15-interval		
		2007	2008	2009	2007	2008	2009	2007	2008	2009
بدون استانداردسازی Without standardization	<i>F</i>	4.69	4.86	4.69	4.92	4.91	5.06	5.1	5.07	5.12
	AICc	0.83	0.87	0.8	0.81	0.86	0.76	0.79	0.84	0.74
		2.86	2.96	2.86	3.01	2.98	3.02	3.14	3.04	3.19
با استانداردسازی With standardization	<i>F</i>	4.8	4.9	4.88	5.08	5.5	5.82	5.3	5.7	5.9
	AICc	0.82	0.86	0.78	0.79	0.83	0.7	0.71	0.74	0.69
		2.9	3	2.95	3.08	3.22	3.35	3.16	3.3	3.46



شکل ۸- مقایسه مقادیر دمای اندازه‌گیری شده با مقادیر شبیه‌سازی شده در مقیاس زمانی سه ساعته در سال ۲۰۰۹ با فاصله درون‌یابی ۵ روز
الف: بدون استانداردسازی و ب: با استانداردسازی
Figure 8- Comparison of measured and simulated 3-hour temperature in year 2009 with 5-interval interpolation (a) without standardization (b) with standardization

نتیجه‌گیری کلی

ها در محدوده تغییر کوچکتری قرار می‌گیرند، پس مجموعه فرکتالی به بخش‌های بزرگتری تقسیم می‌شود که در هر بخش داده‌های بیشتری قرار می‌گیرد، پس تعداد بخش‌ها و در نتیجه تعداد نقاط درون‌یابی برای مدل‌سازی کاهش یافته است. همچنین نتایج مربوط به ریزمقیاس‌سازی با فاصله درون‌یابی ۵ و ۱۰ روز و نگرش اول از دیگر نتایج با کمی اختلاف نتایج بهتری را نشان می‌دهد، به دلیل خطای کم بین نتایج فاصله درون‌یابی ۵ و ۱۰ روز و با توجه به اهمیت زمان اجرای برنامه و استفاده از داده‌های کمتر در نتیجه فاصله درون‌یابی ۱۰ بهترین نتیجه را حاصل می‌کند. آزمون‌های آماری مقادیر آماره R^2 را برای نگرش اول بین ۰/۷۴-۰/۹۸ و برای نگرش دوم بین ۰/۶۹-۰/۹۸، RMSE را برای نگرش اول بین ۱/۳۳-۵/۱۲ و برای نگرش دوم بین ۱/۴۴-۵/۹ درجه سانتیگراد و AIC را برای نگرش اول بین ۳/۱۹-۰/۵۵ و برای نگرش دوم بین ۲/۸۷-۳/۴۶ نشان دادند و همچنین عرض از مبدأها و شیب‌های خطوط در مدل‌سازی‌ها در سطح ۵ درصد تفاوت معنی‌داری به ترتیب با صفر و

در این تحقیق به ریزمقیاس‌سازی روزانه و سه ساعته سری زمانی دمای ایستگاه سینوپتیک مشهد به کمک فراکتال با دو نگرش بر فاصله هاسدرف برای تعیین نقاط درون‌یابی (نگرش اول استفاده از داده‌بدون استانداردسازی و نگرش دوم استفاده از داده‌ها با استانداردسازی) پرداخته شده است، همچنین فاصله نقاط درون‌یابی به صورت متغیر ۵، ۱۰، ۱۵ در نظر گرفته شده است. از آماره‌های ریشه دوم میانگین مربعات خطا، ضریب همبستگی و شاخص اطلاعاتی آکائیک برای بررسی میزان دقت و صحت مدل استفاده شده است. به طور کلی در هر دو نگرش ریزمقیاس‌سازی زمانی روزانه و سه ساعته با دقت و کیفیت بسیار خوب انجام شده است. در نگرش دوم با توجه به اینکه استانداردسازی انجام شده است تعداد نقاط درون‌یابی در نگرش دوم نسبت به نگرش اول به حدود دو سوم کاهش یافته است و به دنبال آن زمان اجرای برنامه کاهش یافته است، به نظر می‌رسد که در نگرش دوم با انجام استانداردسازی پراش داده‌ها کاهش یافته و داده-

منابع

- 1- Barnsley M.F. 1993. Fractals Everywhere. 2nd ed. New York, Academic Press.
- 2- Chuanzhen L., Xiangdong, G. and Shuping Sh. 2000. A speedup method for fractal encoding of digital signals. Signal Processing, 5th International Conference on WCCC-ICSP, 2: 1115-1118.
- 3- Ghahraman B. and Davary K. 2014. Adopting Hierarchical Cluster Analysis to Improve the Performance of K-mean Algorithm. Journal of Water and Soil, Vol. 28, 3, p: 471-480. (in Persian with English abstract).
- 4- Li Z.F. and Li X.F. 2008. An explicit fractal interpolation algorithm for reconstruction of seismic data. Chinese Physics Letters, 3: 1157-1159.
- 5- Mazel D.S. and Hayes M.H. 1992. Using integrated function systems to model discrete sequences. IEEE Transactions on Signal Processing, 40 (7) :1724-1734.
- 6- McQuarrie A. D. and Tsai C. L. 1998. Regression and time series model selection, World Scientific Publishing Co. Pte. Ltd.
- 7- Pathirana A. 2001. Fractal modeling of rainfall: Downscaling in time and space for hydrological applications. PhD thesis, University of Tokyo, Japan.
- 8- Puente C.E. 1995. Geometric modeling of rainfall fields. Water Resources Center Technical Completion Report W-804. University of California, Davis.
- 9- Shahedi M., Sanaiinejad S.H. and Ghahraman B. 2012. Regional Frequency Analysis of Annual Maximum 1-day and 2-day Rainfalls Using Clustering and L-moments, Case study: Khorasan Razavi Province. . Journal of Water and Soil, Vol. 27(1): 80-89. (in Persian with English abstract).
- 10- Shamkoueyan H., Ghahraman B., Davary K. and Sarmad M. 2009. Flood frequency analysis using Linear moment and flood index method in Khorasan provinces. Journal of Water and Soil, Vol. 23(1): 31-43. (in Persian with English abstract).
- 11- Strahle W.C. 1991. Turbulent combustion data analysis using fractals. AIAA, J, 3: 409-417.
- 12- Tajabadi Sh., Ghahraman B. and Ziaei A.N. 2016. Fractal analysis of temperature time series. Msc thesis, Ferdowsi University of Mashhad, Iran.
- 13- Validi N., Ziaei A.N., Ghahraman B. and Ansari H. 2014. Using Fractal Interpolation Functions for Temporal Downscaling of Temperature Data. Journal of Water and Soil, Vol. 27(6): 1123-1132.
- 14- Zhou G.Y. and Leu M.C. 1993. Fractal geometry model for wear prediction. Wear, 170: 1-14.

The Effect of Different Approaches of Distance Between Two Sets on Fractal Downscaling of Temperature in Mashhad

Sh. Tajabadi^{1*} - B. Ghahraman²- A.N. Ziaei³

Received: 04-01-2016

Accepted: 09-01-2017

Introduction: The range of meteorological parameters, such as temperature, are different at different scales. Fractal geometry is a branch of mathematics that has many applications in the field of discrete and continuous domains. Downscaling may be done by different methods, including univariate, multivariate regression functions, splined function and fractal function. Finding the best model for fractal downscaling, is needed to implement the distance between measured and modeled data sets. This distance may be estimated by different methods, including Euclidian. For temporal downscaling, the data are two-dimensional, i.e. time and that of principal variable (e.g. temperatures). In such a case, the dimensionality problem arises in Euclidean space. In these cases, data are usually changed to non-dimensional forms which are referred to standardization, normalization, scaling, or non-dimensionalizing. So, in addition to imbalance of data calculating distance between two sets, we are also considering the impact of standardized data on the number of interpolation points, run time, and accuracy of downscaling the temperature of Mashhad synoptic station.

Materials and Methods: In this paper, fractal model was used for modeling and downscaling temperature datasets for the period of 2007- 2009 at Mashhad Synoptic stations with two approaches of Hasdurf distance to determine the interpolation points (first approach: in first approach original data was used. Second approach: in second approach the data were standardized). We adopted some criteria, such as root mean squared error, correlation, and Akaike information criteria to assess the accuracy of fractal downscaling.

Mashhad is the second most populous city in Iran and capital of Razavi Khorasan Province. It is located in the northeast of the country, close to the borders of Turkmenistan and Afghanistan. It is built-up (or metro) area was home to 2,782,976 inhabitants including Mashhad Taman and Torqabeh cities. It was a major oasis along the ancient Silk Road connecting with Merv in the East. The city is located at 36.20° North latitude and 59.35° East longitude. Mashhad features a steppe climate with hot summers and cool winters. The city only receives about 250 mm of precipitation per year, summers are typically hot and dry, with high temperatures sometimes exceeds 35 °C (95 °F). Winters are typically cool to cold and somewhat humid, with overnight lows routinely dropping below freezing.

At first, fractal method was used to produce daily temperature from daily datasets with two attitude and different interval interpolation (5, 10, 15days). Then the same process was applied to produce 3-hours temperature.

Results and Discussion: Downscaling for daily temperature: In this part, we considered that which standardizing approach and which interval interpolation, will carry the best accuracy for the fractal modeling. Although RMSE, R2, AIC, show that standardized approach is not better, but the difference is not substantial.

Results from fractal modeling from 5-day interval interpolation and 10-day interval interpolation with daily measured temperature in Mashhad compared based on 1:1 line of perfect agreement, and showed acceptable ($\alpha=5\%$) behavior. In both approaches and two interval interpolation with both 5 and 10 days, predicted temperatures imitate the behavior of the measured temperatures. However, simulation with no standardization approach show better results for both distance interpolation compared to the second approach with standardization.

2. Downscaling daily temperature to 3-hour interval: We compared downscaled 3-hour temperature from two standardizing approaches and two times interpolation based on daily temperature with 3-hour measured temperature and compared the results with respect to 1:1 line of perfect agreement. It is clear that the results of the three-hour downscaling show the same results with daily downscaling, because temperature shows the fractal behavior. Although both approaches perform well but un-standardizing is better, yet the difference is not pronounced.

Conclusion: Overall, in both approaches, three-hour and daily downscaling is done precisely and with high

1, 2 and 3- M.Sc. Student, Professor and Assistant Professor, Department of Water Engineering, College of Agriculture, Ferdowsi University of Mashhad

(* - Corresponding Author Email: Tajabadi.sh.1991@gmail.com)

quality. The number of interpolation points was reduced by 30% under the second standardizing approach, which followed by considerable computer runtime. However, the result shows that the first approach had better modeling.

The comparison results of the modeling with 5 intervals interpolation and with 10, the 10 intervals interpolation were more acceptable, such that correlation coefficient was between (first approach: 0.98 and 0.7, second approach: 0.98 and 0.65) while RMSE was between (first approach: 1.33 and 3.27 ° C and second approach: 1.44 and 6.02 ° C), and AICc was between (first approach: 0.55-3.27 and second approach: 2.87-3.51). The intercepts and slopes of regression lines between measured and predicted temperatures were not statistically (5% level of significant) different from 0 and 1, respectively.

Keywords: Downscaling, Housdorf distance, Standardization