

برآورد دمای روزانه خاک با استفاده از شبکه عصبی مصنوعی

کیما امیر مرادی و امید بهمنی¹

دانشجوی سابق کارشناسی ارشد دانشگاه بوعلی سینا همدان؛ kimiaamirmoradi568@gmail.com

استادیار دانشگاه بوعلی سینا؛ Omid_bahmani@yahoo.com

دریافت: 92/2/25 و پذیرش: 93/5/21

چکیده

دمای خاک یکی از متغیرهای مهم در مطالعات هیدرولوژی، هواشناسی، کشاورزی و اقلیم‌شناسی است که اندازه‌گیری و برآورد آن ضروری است. با توجه به این‌که دمای خاک فقط در ایستگاه‌های سینوپتیک کشور اندازه‌گیری می‌شود، کمبود آن در نقاط فاقد ایستگاه از چالش‌های بزرگ در بسیاری از مطالعات مرتبط با کشاورزی است. در این پژوهش، با استفاده از پارامترهای هواشناسی ایستگاه سینوپتیک شیراز در یک دوره 9 ساله (2000-2008) به برآورد دمای خاک در سه عمق پنج، 10 و 30 سانتی‌متری با استفاده از شبکه عصبی مصنوعی پرسپترون چندلایه پرداخته شد. به علت تعداد زیاد متغیرهای مورد استفاده برای برآورد دمای خاک، شناخت متغیرهای مؤثر بر شبکه می‌تواند باعث بهبود نتایج شود. بدین منظور، با استفاده از تکنیک آماری چند متغیره آنالیز مولفه‌های اصلی (PCA) که باعث کاهش تعداد متغیرها و ورود متغیرهای مؤثر به شبکه می‌شود، اقدام به برآورد دمای خاک شد (PCA-ANN). ابتدا، از روش PCA برای کاهش متغیرهای ورودی استفاده شد و هشت متغیر هواشناسی به هشت مولفه اصلی تبدیل شد. چهار مولفه اصلی اول بیش از 99 درصد واریانس کل را به خود اختصاص داد. برای ارزیابی دو مدل ANN و PCA-ANN از شاخص‌های آماری ضریب همبستگی (r)، میانگین جذر مربعات خطا (RMSE) و میانگین خطای اریبی (MBE) استفاده شد. نتایج به دست آمده حاکی از نقش مؤثر پیش‌پردازش روی متغیرها با استفاده از PCA بود. دست‌آوردها نشان داد که شاخص‌های آماری r ، RMSE و MBE (در دوره صحت سنجی) به ترتیب برابر 0/98، 1/61 و 0/2 برای مدل PCA-ANN بر روی پارامتر دمای خاک در عمق پنج سانتی‌متری نتیجه بهتری را به دنبال داشت. روی هم رفته، سنجش دست آوردهای مدل PCA-ANN با دستاوردهای مدل ANN و داده‌های مشاهده شده نشانگر برتری مدل PCA-ANN نسبت به مدل ANN است. لذا، مدل PCA-ANN با ساختاری ساده‌تر، سرعت آموزش شبکه بیشتر و نتایجی دقیق‌تر می‌تواند جایگزین مدل ANN برای برآورد پارامتر دمای خاک باشد.

واژه‌های کلیدی: آنالیز مولفه‌های اصلی، ایستگاه سینوپتیک شیراز، مدل PCA-ANN

¹ نویسنده مسئول، آدرس: همدان، دانشگاه بوعلی سینا، دانشکده کشاورزی، گروه علوم و مهندسی آب

مقدمه

برآورد دقیق متغیرهای محیط خاک به منظور توسعه مدل‌های قابل اعتماد از جریان و انتقال آب و املاح آن در سراسر زنجیره خاک، گیاه و جو برای مدیریت کارآمد از منابع و حفظ کیفیت محیط زیست ضروری است. دمای خاک یکی از متغیرهای مهم در مطالعات هیدرولوژی، هواشناسی، کشاورزی و اقلیم-شناسی است (بهاایگل و همکاران، 2007) و نقش مهمی را در اکوسیستم‌های مختلف از بیابان تا جنگل ایفا می‌کند (کمپ و همکاران، 1992 و ین و همکاران، 1993). دمای خاک بر نفوذ آب از سطح خاک تأثیرگذار است (سینگ، 1992). دمای خاک بر فرآیندهای زیستی گیاهان، حشرات و سایر موجودات مؤثر است (مکسن و همکاران، 1991 و کلاندر و همکاران، 1993). علاوه بر این، سرعت تجزیه و جنبش مواد مغذی و مواد شیمیایی موجود در خاک با نوسانات درجه خاک در ارتباط است (گراندمن و همکاران، 1995 و ویگیل و کیسی، 1995) که یک عامل بسیار مهم در رشد انواع زیادی از محصولات به شمار می‌آید (برار و همکاران، 1992).

دمای خاک متأثر از عوامل متعددی از جمله توپوگرافی، تابش خورشید، دمای هوا، توزیع بارش، میزان رطوبت خاک، نوع و ویژگی گرمایی خاک، همانند ظرفیت گرمایی، ضریب رسانایی گرمایی و گرمای ویژه است (نجفی و همکاران، 1387). به همین دلیل، برآورد دمای خاک به ویژه در نزدیکی سطح زمین که نوسانات آن بیشتر است، تا حدودی مشکل است (می‌هالاکاکو، 2002). اندازه‌گیری دمای خاک در ایستگاه‌های سینوپتیک با انواع مختلف حسگر¹ و یا دماسنج‌های معمولی صورت می‌گیرد. اندازه‌گیری دمای خاک با حسگر هزینه‌بر بوده و به نیروی انسانی ماهر و پایش مداوم نیاز دارد. برای تعیین تغییرات مکانی دمای خاک در اعماق مختلف خاک، چندین حسگر و یا دماسنج مختلف مورد نیاز است (سبزی‌پرور و همکاران، 1389).

اهمیت دمای خاک در علوم کشاورزی و هیدرولوژی از یک سو و وجود مشکلات فراوان در ثبت این پارامتر حیاتی پژوهشگران را بر آن داشته است تا به دنبال یافتن رابطه‌ای بین دمای خاک و پارامترهای دیگر باشند تا بتوانند آن را با دقت مطلوب برآورد کنند. در این خصوص، استفاده از مدل‌های شبکه‌های عصبی مصنوعی (ANNs)² به طور گسترده‌ای طی سال‌های اخیر مورد

توجه واقع شده است (چنگو ایسلام، 2000 و طبری و همکاران، 2011). در مورد فرآیندهایی که تعریف دقیق و کاملی از فیزیک آن‌ها وجود ندارد، مدل‌های ANNs بسیار مؤثر عمل می‌کنند. خصوصیت دیگر این مدل‌ها که آن‌ها را نسبت به سایر روش‌ها و الگوریتم‌ها متمایز می‌سازد، حساسیت کمتر آن‌ها نسبت به وجود خطا در ورودی‌ها است که علت این امر، پردازش توزیعی اطلاعات³ است (تکار، 2000). یکی از مشکلات در طراحی شبکه‌های عصبی مواجه شدن با ساختارهای پیچیده است که منجر به شبکه‌ای با ساختار سنگین می‌شود. در این رابطه، کولیالی و همکاران (2000) اطلاعات مسئله را به سه دسته آموزش⁴، صحت‌یابی⁵ و تست⁶ تقسیم نمودند که این روش را الگوریتم توقف آموزش⁷ نام‌گذاری کردند. یکی از مراحل مهم در طراحی ANNs، انتخاب ورودی-های شبکه است. بسته به نوع فرآیند، ممکن است عوامل متعددی را بتوان یافت که این امر نیز می‌تواند سبب ایجاد شبکه‌هایی با ورودی زیاد و ساختار پیچیده شود.

ازدیار تعداد متغیرهای ورودی می‌تواند سبب پیچیدگی ساختار شبکه و ناپایداری آن شد که یکی از دلایل آن افزایش تعداد وزن‌ها نسبت به تعداد ورودی-هاست (هایکین، 1999). از سویی، در مواردی متغیرها از همبستگی بالایی برخوردارند که استفاده از همگی آن‌ها نوعی ورود اطلاعات تکراری را به همراه دارد (ژانگ، 2007). روش تحلیل مولفه‌های اصلی⁸ یکی از راه‌هایی است که برای رفع مشکلات یادشده پیشنهاد شده و در کنار ANN در مطالعاتی مورد استفاده قرار گرفته است (ژانگ و همکاران، 2006 و ژانگ، 2007). هرچند که این روش در مطالعات هواشناسی کمتر استفاده شده، ولی در موضوعات مختلف زیست محیطی و منابع آب مورد استفاده قرار گرفته است.

از مطالعات انجام گرفته می‌توان به پژوهش نوری و همکاران (1387) در زمینه برآورد جریان رودخانه با استفاده از سه مدل PCA-ANN، ANNs و شبکه عصبی موجکی⁹ (WNN) اشاره کرد. نتایج حاکی از برتری مدل PCA-ANN نسبت به دو مدل دیگر بود. رامانجوم و همکاران (2011) با استفاده از الگوریتم PCA-ANN به بازیابی ساختار بارش در اتمسفر پرداخت. نتایج نشان داد

³ Distributed Processing

⁴ Train

⁵ Validation

⁶ Test

⁷ Stop Training Algorithm (STA)

⁸ Principal Component Analysis (PCA)

⁹ Wavelet Neural Network

¹ Sensor

² Artificial Neural Networks (ANNs)

(T_{max})، دمای خشک (T_d) همان دمای در حالت عادی است که با دماسنج اندازه‌گیری می‌شود، دمای تر (T_w) پایین‌ترین دمایی است که بسته هوا در فشار ثابت در اثر تبخیر آب بدون آن می‌تواند سرد شود، میانگین دمای هوا (T)، رطوبت نسبی هوا (RH)، ساعات آفتابی (n) و سرعت باد در ارتفاع دو متری (U) در مقیاس روزانه بودند. داده‌های دمای خاک نیز در سه عمق پنج، 10 و 30 سانتی‌متری به صورت روزانه در ایستگاه مورد نظر ثبت شده بود.

قبل از انجام تجزیه و تحلیل بر روی داده‌ها، آزمون کیفیت که شامل بررسی همگنی و پرت بودن داده‌ها، های مورد مطالعه بود، انجام شد. با استفاده از نرم افزار SPSS، همگنی و پرت بودن داده‌ها به ترتیب با آزمون‌های ران تست² و باکس پلات³ انجام شد. پس از انجام این مراحل و مرتب کردن داده‌ها، به اجرای مدل‌های ANN و PCA-ANN پرداخته شد. پارامترهای هواشناسی مورد استفاده در طی دوره آماری 9 ساله (2000-2008) است که مقادیر میانگین، بیشینه و کمینه پارامترهای مورد استفاده در طی این دوره برای ایستگاه شیراز در جدول (1) ارائه شده است. به منظور طراحی و آموزش یک شبکه عصبی مصنوعی باید داده‌ها را به دو مجموعه مختلف به نام نمونه آموزشی و نمونه آزمایشی تقسیم نمود (ژانگ و همکاران، 1998). گرانجر (1993)، پیشنهاد می‌کند در مدل‌های برآورد غیرخطی، باید حداقل 20 درصد داده‌ها را برای ارزیابی مدل به کار برد. از این روی، نزدیک 80 درصد داده‌ها (از اول ژانویه 2000 تا اواسط مارس 2007) برای آموزش و 20 درصد داده‌ها (از اواسط مارس 2007 تا آخر دسامبر 2008) برای آزمون کارایی شبکه بکار گرفته شد.

شبکه‌های عصبی مصنوعی

شبکه‌های عصبی شامل سه قسمت وزن (W)، بایاس (b) و تابع انتقال (f) است. این سه مولفه برای هر شبکه عصبی منحصر به فرد و مخصوص همان شبکه است. در رابطه‌های ارائه شده p و a به ترتیب ورودی خروجی شبکه عصبی است. پارامتر n که به نام ورودی شبکه⁴ شناخته می‌شود، ورودی تابع انتقال است و با توجه توجه به p و پارامترهای عصبی ساخته می‌شود. رابطه میان پارامترهای معرفی شده در شکل (1) به صورت شماتیکی نمایش داده شده است. شبکه عصبی مصنوعی از معادلات زیر پیروی می‌کند:

که شبکه عصبی قابلیت بالایی در بازیابی ساختار بارش دارد، ولی الگوریتم PCA-ANN با ورودی‌های کمتر و کاهش مقدار میانگین جذر مربعات خطا (RMSE) زمان محاسباتی را کاهش و دقت بازیابی را افزایش می‌دهد.

طبری و همکاران (2011) به مقایسه عملکرد شبکه عصبی مصنوعی از نوع پرسپترون چند لایه با روش رگرسیون خطی چند متغیره برای برآورد دمای روزانه خاک در عمق‌های 5، 10، 30 و 50 سانتی-متری در منطقه خشک ایران پرداختند. برای ورودی دو مدل شبکه عصبی و رگرسیون خطی چندگانه از میانگین روزانه پارامترهای هواشناسی که شامل دما، تابش خورشیدی و رطوبت نسبی بود، استفاده شد. نتایج مقایسه بین این دو مدل نشان داد که شبکه عصبی مصنوعی دقت بالاتری را در برآورد دمای روزانه خاک در منطقه مورد مطالعه داراست.

یانگ و همکاران (1997) به برآورد دمای خاک در اعماق مختلف با استفاده از شبکه عصبی مصنوعی پرداختند. شبکه عصبی را با ورودی‌های بارش روزانه، تبخیر تعرق پتانسیل و تعداد روزهای سال اجرا نمودند. از این بررسی، این نتیجه به دست آمد که شبکه عصبی بدون ورودی‌های هواشناسی می‌تواند نوسانات فصلی درجه حرارت خاک را به جز نزدیک به سطح خاک با دقت بالایی شبیه سازی کند. نتایج مطالعات انجام شده حاکی از دقت بالاتر، داشتن معماری ساده‌تر و سرعت آموزش سریعتر مدل PCA-ANN نسبت به مدل ANN بود.

هدف از انجام این پژوهش، مقایسه نقش پیش-پردازش داده‌های ورودی مدل شبکه عصبی با تکنیک PCA با نتایج خود شبکه عصبی مصنوعی از نوع پرسپترون چند لایه¹ (MLP) برای برآورد دمای خاک در اعماق پنج، 10 و 30 سانتی‌متری ایستگاه سینوپتیک شیراز در طی دوره آماری 1379-1387 است.

مواد و روش‌ها

داده‌ها و منطقه مطالعاتی

این مطالعه با استفاده از اطلاعات روزانه ثبت شده ایستگاه سینوپتیک هواشناسی شیراز انجام شد. طول جغرافیایی ایستگاه هواشناسی شیراز $36^{\circ} 52'$ شرقی و عرض جغرافیایی آن $29^{\circ} 32'$ شمالی است. انتخاب ایستگاه مورد مطالعه به علت کامل بودن داده‌های مورد نیاز و برآورد دقیقی آن‌ها نسبت به دیگر ایستگاه‌ها در جنوب کشور بود. داده‌های هواشناسی مورد استفاده در این مطالعه شامل دمای کمینه هوا (T_{min})، دمای بیشینه هوا

² Run- Test

³ Box Plot

⁴ Net input

¹ Multi-Layer Perceptron

این پژوهش مورد استفاده قرار گرفته است، تابع سیگموئید است که به صورت زیر تعریف می‌شود:

$$f(y) = \frac{1}{1+e^{-ky}} \quad (3)$$

که در آن: k پارامتر تعیین کننده‌ی واکنش $f(y)$ است. در این تحلیل‌ها، پارامتر k به خاطر عدم سازگاری آن با الگوریتم انتشار به عقب، مقدار ثابت یک در نظر گرفته می‌شود.

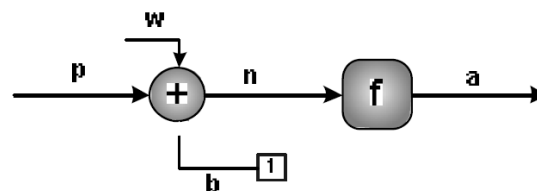
$$n = wp + b \quad (1)$$

$$a = f(n) = f(wp + b) \quad (2)$$

در فرآیند آموزش عصبی، w و b (پارامترهای عصبی) تا زمانی که بهترین تقریب برای خروجی را با توجه به پارامتر ورودی به دست آورند، ادامه دارد. طی فرآیند آموزش وزن‌ها (w) و مقادیر ثابتی که با آن‌ها جمع می‌شود، اصطلاحاً بایاس نامیده می‌شوند که به طور پی در پی تغییر می‌کند تا خطا به کمترین مقدار خود برسد. به طور کلی، توابع انتقالی به دو صورت خطی و غیر خطی وجود دارد. یک نمونه از توابع انتقال غیرخطی $f(y)$ که در

جدول 1- میانگین اقلیمی پارامترهای هواشناسی مورد استفاده در دوره 9 ساله (2000-2008) ایستگاه سینوپتیک شیراز

پارامتر	دمای هوا	رطوبت نسبی	ساعات آفتابی	سرعت باد در ارتفاع
	(°C)	(%)	(hour)	2متری (m.s ⁻¹)
نماد	T _w	RH	n	U ₂
میانگین	10/40	38/80	9/38	3/89
بیشینه	22/1	98/5	12/9	26/4
کمینه	-3	11/4	0/00	0/00



شکل 1- شماتیکی از عملکرد شبکه عصبی مصنوعی

است (سبزی پرور و علیانی، 1390 و کنانی و همکاران، 2008). شکل‌های (2) و (3) به ترتیب نشان‌دهنده‌ی تابع انتقال سیگموئید برای لایه پنهان و تابع انتقال خطی برای لایه ورودی و خروجی است. شبکه عصبی مورد استفاده در این مطالعه با 10000 دور تکرار در نظر گرفته شد و با استفاده از نرم‌افزار Matlab تخمین انجام گرفت. برای تعیین تعداد گره‌های لایه پنهان معمولاً از آزمون سعی و خطا استفاده می‌شود. MLP با یک لایه پنهان با تابع انتقال غیر خطی سیگموئید و لایه خروجی با تابع انتقال خطی به وسیله معادلات زیر می‌توان مدل‌سازی کرد:

$$a^1_j(t) = F\{\sum_{i=1}^R w^1_{i,j} p_i(t) + b^1_j\} \quad (4)$$

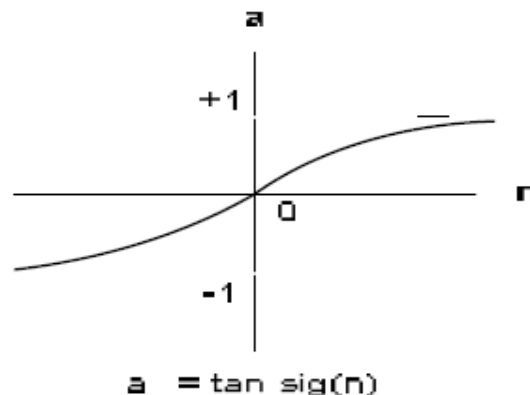
$$a^2_k(t) = G\{\sum_{j=1}^{S_1} w^2_{k,j} a^1_j(t) + b^2_k\} \quad (5)$$

که در آن‌ها، R تعداد اجزای بردار ورودی، S_1 تعداد نرون‌های لایه مخفی، P بردار ورودی، w^1 و w^2 به ترتیب ماتریس وزن در لایه‌های پنهان و خروجی، b^1 و b^2 به ترتیب بردارهای بایاس در لایه‌های پنهان و خروجی

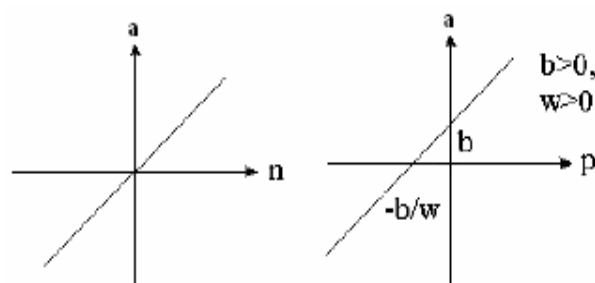
یکی از انواع شبکه‌های عصبی که در این پژوهش مورد استفاده قرار گرفته است، شبکه‌های عصبی پرسپترون چند لایه (MLP) است که از کاربرد روزافزونی بهره‌مند است. با توجه به تقریب جهانی، هر MLP با تابع انتقال سیگموئید برای لایه مخفی و یک لایه خروجی با تابع انتقال خطی قادر به برآورد هر پارامتر پیچیده است، به شرطی که تعداد گره‌های لایه مخفی به درستی انتخاب شود (تکاراند، 1999). بر اساس پژوهش‌های ASCE و برادوک و همکاران (1998) در خصوص مسائل آب شناختی و پژوهش‌های کولیپالی و همکاران (2000) در مسائل اقلیم شناختی، شبکه عصبی مصنوعی پرسپترون چند لایه با الگوریتم پس انتشار خطا برای انجام شبیه سازی پارامتر دمای خاک انتخاب شد.

با توجه به نظریه مطرح شده، شبکه عصبی مصنوعی مورد استفاده در این پژوهش شبکه عصبی پرسپترون چند لایه (MLP) با یک لایه پنهان با تابع انتقال سیگموئید و لایه ورودی و خروجی با تابع انتقال خطی

می‌باشند. F و G توابع انتقال در لایه‌های پنهان و خروجی است.



شکل 2- تابع انتقال سیگموئید لایه مخفی



شکل 3- تابع انتقال خطی برای لایه ورودی و خروجی

که در آن: X_i متغیر یا مولفه i ام، MIN_{X_i} حداقل مقدار متغیر یا مولفه، MAX_{X_i} حداکثر مقدار متغیر یا مولفه مورد نظر و N نیز مقادیر استاندارد شده است.

آنالیز مولفه‌های اصلی (PCA)

تجزیه مولفه اصلی از روش‌های آماری چند متغیره است که می‌توان از آن برای کاهش تعداد متغیرها و تفسیر بهتر اطلاعات استفاده کرد (کامدویرن و همکاران، 2005). هدف از تجزیه به مولفه‌های اصلی آن است که واریانس موجود در داده‌های چند متغیره را به مولفه‌هایی تجزیه کند که اولین مولفه تا آنجا که ممکن است علت بیشترین واریانس موجود در داده‌ها باشد. دومین مولفه علت بیشترین واریانس ممکن بعد از مولفه اول باشد و همچنین هیچ گونه همبستگی بین مولفه‌ها موجود نباشد. با اعمال این روش، متغیرهای ورودی اولیه به مولفه‌های جدید و بدون همبستگی تبدیل می‌شوند؛ به طوری که مولفه‌های ایجاد شده، ترکیبی خطی از متغیرهای ورودی-اند (لو و همکاران، 2003).

یک قضیه اصلی وجود دارد که می‌توان با استفاده از آن مولفه‌ها را به دست آورد. این قضیه بیان می‌دارد که ترکیباتی از P متغیر $X_1, X_2, \dots$ و X_p برای ایجاد P مولفه

مهمترین شروطی که می‌توان برای خاتمه الگوریتم شبکه به کار برد شامل توقف بعد از تکرار به دفعات معین و توقف وقتی که خطا از یک مقدار تعیین شده کمتر شود یا به عبارتی مربع خطای بین خروجی‌های شبکه و تابع هدف حداقل شود، است. اگر دفعات تکرار کم باشد، خطا وجود خواهد داشت و اگر زیاد باشد، مسئله Overfitting رخ خواهد داد. در این پژوهش، تعداد مناسب تکرار و حد خطای معیار برای توقف الگوریتم به ترتیب 10000 و 1×10^{-7} در نظر گرفته شد.

پیش پردازش: به علت تأثیر متفاوت داده‌های ورودی بر روی نرون‌ها، در ابتدا لازم است داده‌های ورودی به شبکه استاندارد شوند، یعنی بین یک دامنه عددی، که معمولاً (0 و 1) می‌باشد، قرار گیرند. با توجه به این‌که در لایه مخفی از تابع سیگموئید استفاده شد، از رابطه (6) که باعث می‌شود، داده‌ها بین 0/1 و 0/9 قرار گیرند، برای استاندارد کردن استفاده شد (داسون، 1998):

$$N = 0.8 \frac{(X_i - MIN_{X_i})}{(MAX_{X_i} - MIN_{X_i})} + 0.1 \quad (6)$$

متغیرهای هواشناسی مورد استفاده و 3288 تعداد روزها در طی دوره آماری (2000-2008) به عنوان ورودی مدل PCA وارد نرم افزار SPSS شد و سپس مقدار بردار ویژه و مقادیر ویژه و نمرات³ آن محاسبه شد.

معیار ارزیابی مدل

برای ارزیابی مدل ها از جذر میانگین مربعات خطا⁴ (RMSE)، ضریب همبستگی⁵ (r) و میانگین خطای اریبی⁶ (MBE) استفاده شد (چن و همکاران، 2003):

$$r = \frac{N \sum_{i=1}^N O_i y_{p,i} - (\sum_{i=1}^N O_i)(\sum_{i=1}^N y_{p,i})}{\sqrt{[N \sum_{i=1}^N O_i^2 - (\sum_{i=1}^N O_i)^2] \times [N \sum_{i=1}^N y_{p,i}^2 - (\sum_{i=1}^N y_{p,i})^2]}}$$

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(y_{p,i} - O_i)^2}{N}} \quad (11)$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (y_{p,i} - O_i) \quad (12)$$

که در آن ها، $y_{p,i}$ و O_i به ترتیب مقادیر محاسبه شده توسط مدل و مقادیر اندازه گیری شده ی متغیرهاست و N بیانگر تعداد مشاهدات است. کم بودن میزان خطا (RMSE و MBE) و بالا بودن ضریب همبستگی (r) بیانگر دقت قابل قبول مدل و ملاک برتری آن خواهد بود.

نتایج و بحث

آنالیز مولفه های اصلی

به ازا هشت متغیر ورودی که شامل دمای کمینه هوا (Tmin)، دمای بیشینه هوا (Tmax)، دمای خشک (Td)، دمای تر (Tw)، میانگین دمای هوا (T)، رطوبت نسبی هوا (RH)، ساعات آفتابی (n) و سرعت باد در ارتفاع دو متری (U) هشت مولفه اصلی تشکیل شد. مشخصات این مولفه ها در جدول (2) آمده است. در جدول (2) مقادیر ویژه و درصد واریانس متغیرهای اولیه که به وسیله هر مولفه بیان می شود، نشان داده شده است. این جدول نشان می دهد اولین مولفه تا آنجا که ممکن است علت بیشترین واریانس موجود در داده ها است. دومین مولفه علت بیشترین واریانس ممکن بعد از مولفه اول و الی آخر است. هنگام تجزیه به مولفه اصلی اغلب این امید وجود دارد که واریانس بسیاری از مولفه ها آنقدر کم باشد که قابل صرف نظر کردن باشند. در این حالت کارایی تجزیه بالا می رود، زیرا تغییرات در P متغیر اولیه X به وسیله تعداد کمی متغیر PC بیان می شود. همچنین، در این جدول ملاحظه شد که چهار مولفه اول 99/27

مستقل $PC_1, PC_2, \dots, PC_p$ برقرار می شود. در این روش اطلاعات پارامترهای اصلی با کمترین تلفات در مؤلفه های حاصل آورده می شود (جانسون و ویچرن، 1982). هر مؤلفه اصلی می تواند با دنباله زیر مشخص شود:

$$PC_i = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p \quad (7)$$

که در آن: (7)، PC_i معرف مولفه مورد نظر، a_{ij} بردار ویژه¹ مربوطه و X_i نیز متغیرهای اصلی است. این اطلاعات از حل معادله زیر به دست می آید (جانسون و ویچرن، 1982).

$$|A - \lambda I| = 0 \quad (8)$$

که در آن: I ماتریس واحد، A ماتریس دترمینان متغیرهای ورودی و λ نیز مقادیر ویژه این ماتریس است. از این مقادیر ویژه، بردارهای ویژه به دست می آیند. این بردارهای ویژه همان ضرایب مولفه های اصلی بوده و واریانس مولفه های اصلی همان مقادیر ویژه مربوطه است. روشی که به وسیله آن محقق قادر به تشخیص مناسب بودن داده ها برای تحلیل داده ها می باشد، آزمون KMO² است. مقدار این فاکتور بین صفر و یک متغیر است. این فاکتور به کمک ضرایب همبستگی ساده و ضرایب همبستگی جزئی طبق فرمول (9) محاسبه می شود:

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad i \neq j \quad (9)$$

که در آن: r_{ij} ضریب همبستگی ساده بین متغیرهای i و j و a_{ij} نیز ضریب همبستگی جزئی متغیرهای i و j به شرط ثابت بودن سایر متغیرهاست. با توجه به فرمول (9) کمیت فاکتور KMO با ضرایب همبستگی ساده رابطه مستقیم و با ضرایب همبستگی جزئی رابطه معکوس دارد. در صورتی که این فاکتور بزرگتر از 0/5 به دست آید، این امر نشان دهنده امکان اجرای تجزیه به مؤلفه اصلی روی متغیرهای ورودی است (سینگ و همکاران، 2004 و شرستا و کازاما، 2007).

مقدار فاکتور KMO برابر 0/763 برای داده های هواشناسی مورد استفاده حاصل شد. در این پژوهش، با توجه به مقدار فاکتور KMO می توان آنالیز مولفه های اصلی را بر روی داده های مورد مطالعه اجرا کرد. برای انجام آنالیز تجزیه به مولفه اصلی در این پژوهش یک ماتریس همبستگی با آرایش 3288×8 (هشت تعداد

³. score

⁴. Root Mean Square Error

⁵. Mean Bias Error

⁶. Correlation Coefficient

¹. Eigenvectors

². Kaiser- Meyer- Olkin

درصد کل پراکندگی و اطلاعات متغیرهای اصلی را بیان می‌کنند. بنابراین، می‌توان چهار مولفه اول اصلی اول را به عنوان مولفه اصلی قلمداد نمود. بر این اساس مولفه‌های اول تا چهارم برای مطالعات بعدی انتخاب شدند.

ضرایب متغیرها در هر کدام از مولفه‌های اصلی نیز در جدول (3) ارائه شده است. برای مشخص نمودن مؤثرترین متغیرها در شکل‌گیری چهار مولفه اصلی اول چرخش واریمکس بر روی ضرایب متغیرها انجام شد و نتایج در جدول (4) ارائه شده است. همان‌طور که از جدول (4) مشاهده می‌شود در تشکیل مولفه اول که به تنهایی بیش از 73 درصد پراکندگی کل را به خود اختصاص می‌دهد، متغیرهای مربوط به دما که در این پژوهش شامل دمای کمینه هوا ($T_{\min}$)، دمای بیشینه هوا ($T_{\max}$)، دمای خشک (T_d)، دمای تر (T_w)، میانگین دمای

هوا (T) است، مقدار بردار ویژه آن‌ها بزرگتر از 0/9 ظاهر شد (بردار ویژه نشان‌دهنده همبستگی بین متغیرها و PC می‌باشد) و همبستگی متغیرهای سرعت باد (U) و ساعات آفتابی (n) نسبت به دیگر متغیرها کمتر است.

بنابراین، این نتایج نشان می‌دهد که متغیرهای دما بیش‌ترین تأثیر را در شکل‌گیری این مولفه دارا بوده‌اند. در مولفه دوم بیش‌ترین عامل بارگذاری مربوط به متغیرهای ساعات آفتابی (n) و رطوبت نسبی (RH) است و بیش‌ترین مقدار را متغیر ساعات آفتابی به خود اختصاص داده است. نتایج حاصل از عامل‌های بارگذاری برای مولفه سوم نشان می‌دهد که بیش‌ترین مقدار همبستگی را این مولفه با متغیر سرعت باد (U) دارا است و مولفه چهارم بیش‌ترین مقدار همبستگی را با متغیر رطوبت نسبی داشته است.

جدول 2- مقادیر ویژه و درصدی از واریانس متغیرهای اولیه که به‌وسیله هر مولفه ارائه می‌شود

مولفه	مقادیر ویژه	درصد واریانس	درصد واریانس تجمعی
1	5/866	73/324	73/324
2	0/964	12/044	85/368
3	0/909	11/359	96/727
4	0/204	2/550	99/277
5	0/044	0/556	99/833
6	0/010	0/125	99/958
7	0/003	0/042	100/00
8	$2/216 \times 10^{-6}$	$2/77 \times 10^{-5}$	100/00

جدول 3- ضرایب متغیرها در هر کدام از مولفه‌های اصلی

متغیر	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
U_2	0/332	0/434	0/837	0/041	0/018	0/003	0/00	$-2/979 \times 10^{-8}$
T_d	0/990	0/067	-0/107	-0/007	0/023	-0/002	0/052	$2/653 \times 10^{-7}$
T_w	0/933	0/243	-0/214	0/123	0/073	-0/069	-0/016	$-5/258 \times 10^{-7}$
$T_{\min}$	0/946	0/254	-0/116	0/023	-0/165	0/003	-0/007	0/001
$T_{\max}$	0/988	-0/32	-0/099	-0/019	0/101	0/057	-0/014	0/001
T	0/988	0/099	-0/109	$-4/048 \times 10^{-5}$	-0/019	0/033	-0/011	-0/001
RH	-0/829	0/381	-0/221	0/343	0/008	0/030	0/007	$1/784 \times 10^{-7}$
n	0/609	-0/701	0/260	0/262	-0/024	-0/003	0/00	$-4/729 \times 10^{-8}$

جدول 4- ضرایب متغیرهای 4 مولفه اصلی اول به وسیله چرخش واریماکس

متغیر	PC1	PC2	PC3	PC4
U_2	0/146	0/047	0/986	0/060
T_d	0/923	0/279	0/107	0/236
T_w	0/981	0/137	0/090	0/039
T_{min}	0/951	0/124	0/176	0/146
T_{max}	0/884	0/353	0/069	0/274
T	0/933	0/255	0/120	0/221
RH	-0/488	-0/541	-0/146	-0/669
n	0/239	0/959	0/038	0/148

شبکه عصبی مصنوعی

در این پژوهش، به منظور مدل‌سازی دمای خاک در سه عمق پنج، 10 و 30 سانتی‌متری ایستگاه سینوپتیک شیراز از داده‌های هواشناسی اشاره شده در طی دوره‌ی آماری 2000-2008 استفاده شد. همچنین، داده‌های مربوط به شش سال (2000-2006) برای آموزش شبکه‌ها و دو سال (2007-2008) برای صحت سنجی شبکه‌ها مورد استفاده قرار گرفت. بعد از استاندارد کردن هشت متغیر اولیه و تعیین ساختار و ساختار شبکه به برآورد دمای خاک در اعماق یاد شده پرداخته شد. برای تعیین تعداد گره‌های لایه مخفی معمولاً از آزمون سعی و خطا استفاده می‌شود. یکی از این روابط محدوده‌ی مقادیر گره‌های لایه پنهان را بین $2n+1$ و $2n^{1/2}+m$ تخمین می‌زند (کوانگ - مینگ، 2005). در این روابط m تعداد گره‌های خروجی و n تعداد گره‌های ورودی است. در این پژوهش لایه پنهان با 12 گره بهترین جواب را نشان داد.

شبکه عصبی و آنالیز مولفه‌های اصلی

در این بخش نیز با انجام تکنیک PCA، هشت متغیر اولیه به هشت مولفه تبدیل شدند. چهار مولفه اصلی اول بیش از 99 درصد کل واریانس متغیرهای اولیه را دارا بودند. برای انتخاب تعداد مولفه‌های لازم برای شبکه PCA-ANN از آزمون سعی و خطا استفاده شد. برای انجام این کار تعداد مولفه‌های مختلف مثلاً سه تا مولفه اول، چهار تا و ... مورد آزمون قرار گرفت. نتایج به دست آمده نشان داد که چهار مولفه اول بهترین ورودی را نسبت به دیگر تعداد مولفه‌ها برای شبکه عصبی مصنوعی به خود اختصاص دادند. در این قسمت نیز برای داده‌های مربوط به شش سال (2000-2006) برای آموزش شبکه‌ها و دو سال (2007-2008) برای صحت سنجی شبکه‌ها مورد استفاده قرار گرفت. برای این مدل نیز تعداد گره‌های لایه مخفی با استفاده از دامنه $2n^{1/2}+m$ و $2n+1$ تخمین زده شده است. نتایج نشان داد که لایه پنهان با هشت گره بهترین جواب را به خود اختصاص داده است. توابع انتقالی مورد استفاده نیز مانند قسمت قبلی که مربوط به

خود شبکه عصبی بود، انتخاب شد. مدل 4PCs-ANN نتایج بهتری را نسبت به دیگر تعداد مولفه‌ها ارائه داد.

ارزیابی مدل‌های ANN و 4PCs-ANN

به منظور ارزیابی عملکرد مدل‌های ANN و 4PCs-ANN در شبیه‌سازی دمای خاک در سه عمق پنج، 10 و 30 سانتی‌متری از سه شاخص آماری r ، RMSE و MBE استفاده شد و مقدار این شاخص‌ها برای دو مدل ANN و 4PCs-ANN محاسبه و نتایج حاصله از این دو مدل جدول (5) ارائه شده است. در این جدول مشاهده می‌شود که مدل 4PCs-ANN نتایج بهتری را نسبت به مدل ANN ارائه داده است. به طوری که تمامی شاخص‌های آماری مقادیر مطلوب‌تری را در مدل 4PCs-ANN داراست. در این مدل، مقدار خطا (RMSE و MBE) کمتر از مدل ANN و مقدار ضریب همبستگی (r) آن بیشتر از مدل ANN در هر سه عمق داراست که این مسئله بیانگر بالا بودن دقت مدل 4PCs-ANN نسبت به مدل ANN است. در واقع مدل 4PCs-ANN با ساختاری ساده‌تر نتایج قابل قبول‌تری را به خود اختصاص داده است.

شکل‌های (4) تا (6) به ترتیب نمودار بین مقادیر مشاهده شده و برآورد شده برای دو مدل ANN و 4PCs-ANN به منظور برآورد دمای خاک در اعماق پنج، 10 و 30 سانتی‌متری را نشان می‌دهد. از نمودارها روشن است که مدل 4PCs-ANN مقدار ضریب تعیین (R^2) بالاتری نسبت به مدل ANN برای برآورد دمای خاک در هر سه عمق اشاره شده، دارد که نشان‌دهنده نزدیک‌تر بودن مقادیر تخمین زده شده مدل 4PCs-ANN به مقادیر اندازه‌گیری شده است.

معادله برازش خطی که برای هر نمودار به صورت معادله $y=a_0x+a_1$ تعریف شده است. ضریب a_0 هر چه - قدر به سمت یک میل کند و فاصله کمتری با یک داشته باشد نشان‌دهنده عملکرد بهتر مدل مربوطه است. نمودارهای ارائه شده در شکل (7) نشان می‌دهند که مدل 4PCs-ANN نسبت به مدل ANN عملکرد بهتری را داشته است. مقدار شاخص MBE در عمق 30 سانتی‌متری

سینوپتیک شیراز با استفاده از داده‌های هواشناسی مورد ارزیابی قرار گرفت. در مرحله اول با استفاده از مدل ANN و با هشت متغیر ورودی کمینه هوا (T_{min})، دمای بیشینه هوا (T_{max})، دمای خشک (T_d)، دمای تر (T_w)، میانگین دمای هوا (T)، رطوبت نسبی هوا (RH)، ساعات آفتابی (n) و سرعت باد در ارتفاع دو متری (U) برآورد روزانه دمای خاک در سه عمق پنج، 10 و 30 سانتی متری صورت گرفت. در مرحله دوم، پیش‌پردازش متغیرهای ورودی با استفاده از تکنیک PCA انجام و به اِزاء هشت متغیر ورودی هشت مولفه اصلی تشکیل شد. دست‌آوردها نشان داد که چهار مولفه اول 99/27 درصد کل پراکندگی و اطلاعات متغیرهای اصلی را بیان می‌کنند. بنابراین می‌توان چهار مولفه اول اصلی اول را به عنوان مولفه اصلی قلمداد نمود. برای مشخص نمودن مؤثرترین متغیرها در شکل‌گیری چهار مولفه اصلی اول چرخش واریمکس بر روی ضرایب متغیرها انجام شد. نتایج حاصل نشان داد که متغیرهای دما بیش‌ترین تأثیر را در شکل‌گیری مولفه اصلی اول دارا بوده‌اند در مولفه دوم بیش‌ترین عامل بارگذاری مربوط به متغیر ساعات آفتابی، مولفه سوم بیش‌ترین مقدار همبستگی با متغیر سرعت باد (U) و مولفه‌ی چهارم بیش‌ترین مقدار همبستگی را با متغیر رطوبت نسبی داشت.

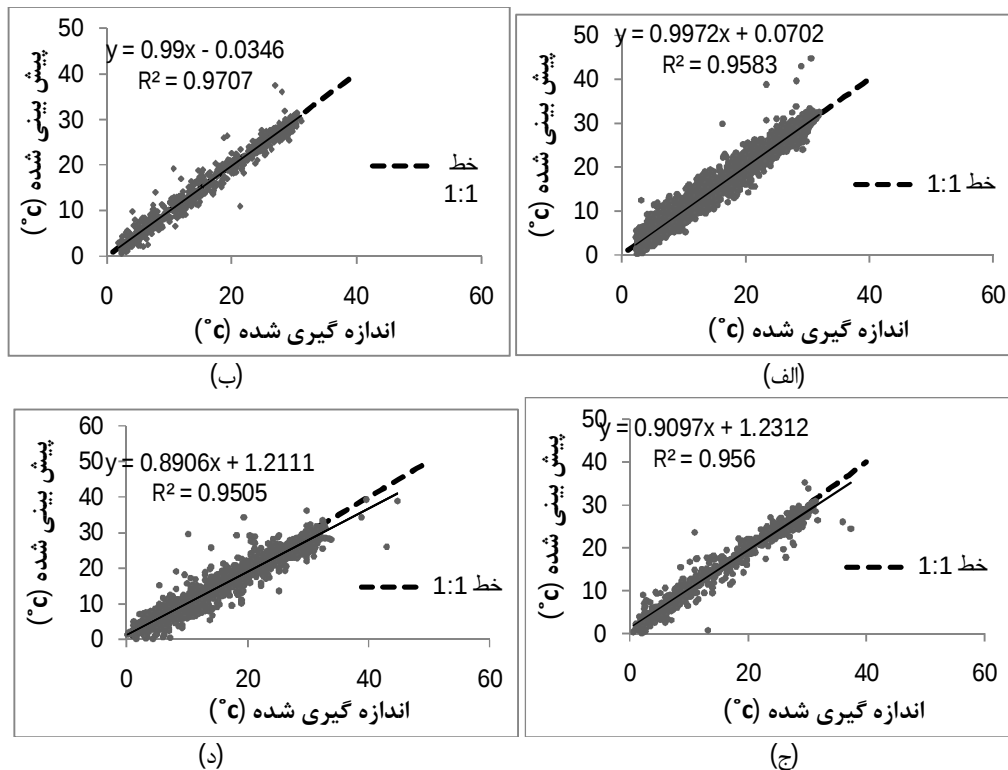
در مرحله صحت‌سنجی مدل ANN نسبت به مدل 4PCs-ANN کمتر مشاهده می‌شود. اختلاف این شاخص در عمق ارائه شده در حد دهم است و بر روی نتیجه حاصل برای دقیق‌تر بودن مدل 4PCs-ANN تأثیر گذار نخواهد بود. اگرچه از روش PCA-ANN در مطالعات مربوط به برآورد دمای خاک استفاده نشده، ولی استفاده از این روش در دیگر زمینه‌های علوم با بهبود عملکرد شبکه عصبی همراه بوده است (چویی، 2001، لو، 2003 و ژانگ، 2006). نوری و همکاران (1387) به برآورد جریان رودخانه با استفاده از سه مدل ANNs، PCA-ANN و شبکه عصبی موجکی (WNN) پرداختند که نتایج حاکی از برتری مدل PCA-ANN نسبت به دو مدل دیگر بود. برای مقایسه آماری نتایج به دست آمده از دو مدل 4PCs-ANN و ANN نمودار مقایسه‌ای خط 1:1 دو مدل برای عمق‌های 5، 10 و 30 سانتی متری در شکل (7) ارائه شده است. توزیع داده‌های حاصل از دو مدل مربوطه در اطراف خط 1:1 ترسیم شد. همانطور که در شکل (9) مشاهده می‌شود تابع رگرسیونی حاصل برای هر سه عمق پنج، 10 و 30 سانتی متری دارای شبیهی نزدیک به یک می‌باشد که نشان‌دهنده مقایسه مناسب هر دو مدل در اطراف خط 1:1 است.

نتیجه‌گیری

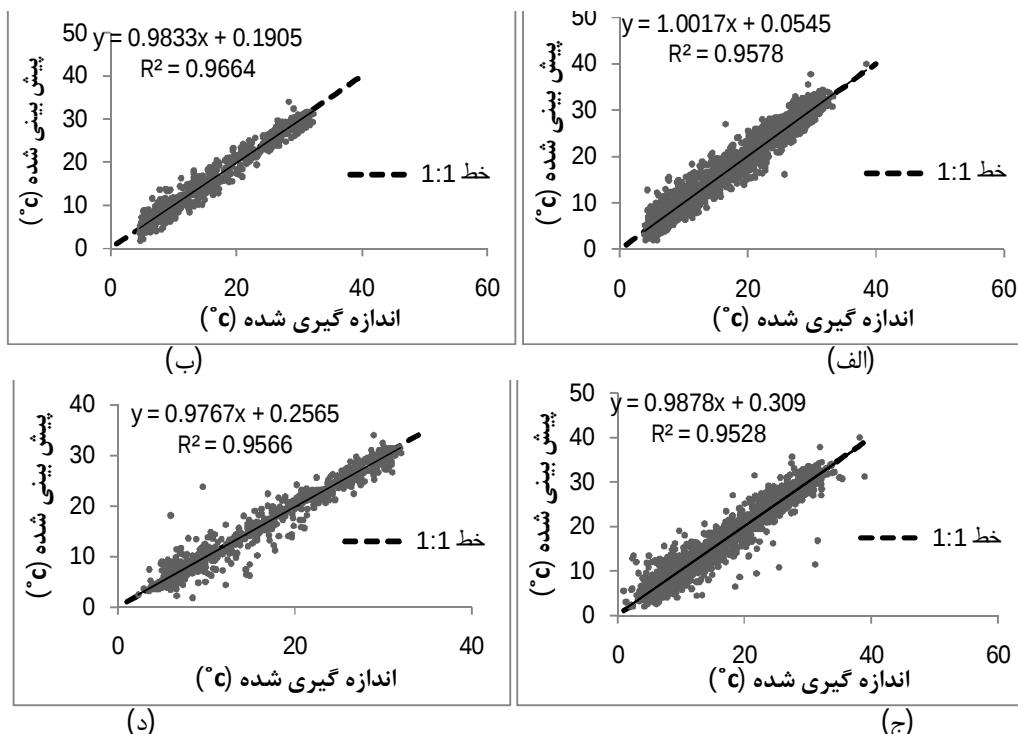
در پژوهش حاضر عملکرد دو مدل ANN و PCA-ANN برای برآورد دمای روزانه خاک ایستگاه

جدول 5- نتایج مراحل آموزش و صحت‌سنجی مدل‌های ANN و 4PCs-ANN در برآورد دمای خاک در عمق‌های 5، 10 و 30 سانتی متری

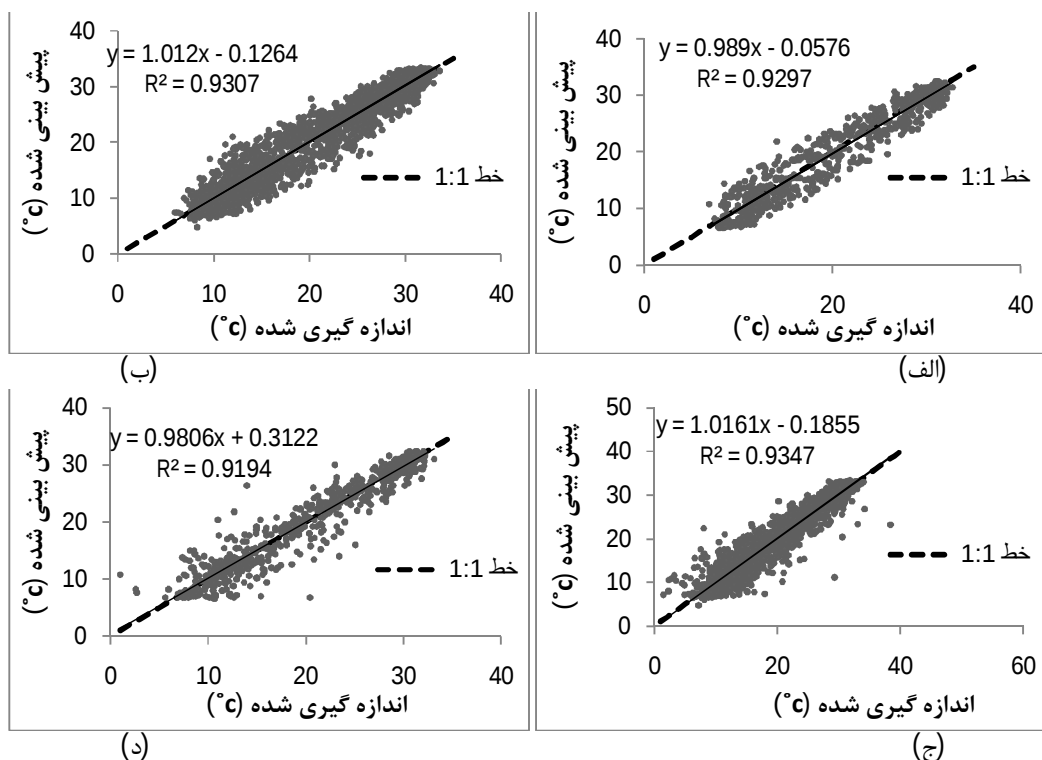
مدل	عمق (سانتی‌متر)	r		RMSE (°C)		MBE (°C)	
		آموزش	صحت- سنجی	آموزش	صحت- سنجی	آموزش	صحت- سنجی
ANN	5	0/974	0/977	2/22	2/02	-0/66	-0/28
	10	0/975	0/977	1/98	1/89	-0/08	0/18
	30	0/964	0/960	2/14	2/33	-0/15	0/09
4PCs-ANN	5	0/980	0/984	1/87	1/61	-0/02	0/2
	10	0/978	0/982	1/87	1/66	-0/08	0/12
	30	0/964	0/964	2/2	2/19	0/12	0/28



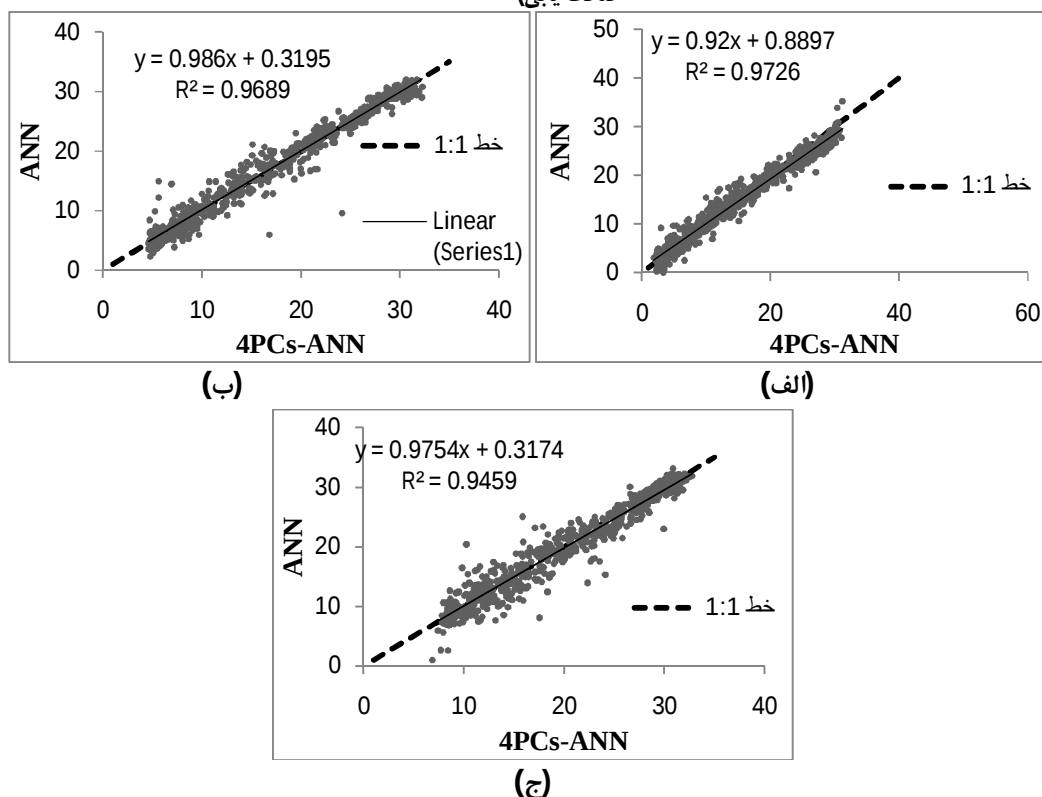
شکل 4- مقایسه نتایج محاسباتی و مشاهداتی پارامتر دمای روزانه خاک ایستگاه سینوپتیک شیراز در عمق 5 سانتی متری توسط مدل (الف) 4PCs-ANN (دورهی آموزش)، (ب) مدل 4PCs-ANN (دورهی صحت یابی)، (ج) مدل ANN (دورهی آموزش) و (د) مدل ANN (دورهی صحت یابی)



شکل 5- مقایسه نتایج محاسباتی و مشاهداتی پارامتر دمای روزانه خاک ایستگاه سینوپتیک شیراز در عمق 10 سانتی متری توسط مدل (الف) 4PCs-ANN (دورهی آموزش)، (ب) مدل 4PCs-ANN (دورهی صحت یابی)، (ج) مدل ANN (دورهی آموزش) و (د) مدل ANN (دورهی صحت یابی)



شکل 6- مقایسه نتایج محاسباتی و مشاهداتی پارامتر دمای روزانه خاک ایستگاه سینوپتیک شیراز در عمق 30 سانتی متری توسط الف) مدل 4PCs-ANN (دورهی آموزش)، ب) مدل 4PCs-ANN (دورهی صحت یابی)، ج) مدل ANN (دورهی آموزش) و د) مدل ANN (دورهی صحت یابی)



شکل 7- مقایسه‌ی نتایج حاصل شده از دو مدل 4PCs-ANN و ANN برای عمق‌های: الف) 5، ب) 10 و ج) 30 سانتی متری

مربعات خطا (RMSE) و میانگین خطای اریبی (MBE) استفاده شد. مقایسه عملکرد دو مدل بکار رفته نشان داد که مدل PCA-ANN با داشتن ساختاری ساده‌تر و سرعت آموزش سریع‌تر و همچنین با ارائه دادن نتایجی بهتر و با دقتی بیشتر نسبت به شبکه عصبی در برآورد دمای روزانه خاک مدلی مناسب است. در حالت کلی پیش‌پردازش متغیرهای ورودی اثر مثبتی در عملکرد شبکه عصبی داشته است.

در مرحله سوم، با استفاده از مدل PCA-ANN برآورد روزانه دمای خاک در سه عمق پنج، 10 و 30 سانتی‌متری انجام شد. برای انتخاب تعداد مولفه‌های لازم برای شبکه PCA-ANN از آزمون سعی و خطا استفاده شد. نتایج به دست آمده نشان داد که چهار مولفه اول بهترین ورودی را نسبت به دیگر تعداد مولفه‌ها برای شبکه عصبی مصنوعی به خود اختصاص دادند. برای ارزیابی نتایج حاصل از دو مدل ANN و PCA-ANN از سه شاخص آماری ضریب همبستگی (r)، میانگین جذر

فهرست منابع:

1. سبزی پرور، ا. و طبری، ح. 1389. برآورد میانگین روزانه دمای خاک در چند نمونه اقلیمی ایران با استفاده از داده‌های هواشناسی. مجله علمی - پژوهشی علوم آب و خاک (علوم و فنون کشاورزی و منابع طبیعی)، سال چهاردهم، شماره پنجاه و دوم، صفحه: 125-137.
2. سبزی پرور، ا. و علیانی، ا. 1390. ارزیابی عملکرد شبکه عصبی مصنوعی در برآورد تابش خورشیدی کل روزانه و مقایسه آن با نتایج مدل انگستروم (مطالعه موردی: ایستگاه همدیدی تبریز). مجله ژئوفیزیک ایران، جلد 5، شماره 3، صفحه: 41-30.
3. نجفی مود، م.ح، علیزاده، ا، محمدیان، ا. و موسوی، ج. 1387. بررسی رابطه دمای هوا و دمای اعماق مختلف خاک و برآورد عمق یخبندان (مطالعه موردی استان خراسان رضوی). مجله علمی - پژوهشی آب و خاک (علوم و صنایع کشاورزی)، جلد 22، شماره 2، صفحه: 456-466.
4. نوری، ر، فرخ‌نیا، ا، مرید، س. و ریاحی مدوار، ح. 1388. تأثیر پیش‌پردازش‌های متغیرهای ورودی به شبکه عصبی برای برآورد جریان ماهانه با آنالیز مولفه‌های اصلی و موجک. مجله آب و فاضلاب، شماره 1، صفحه: 22-13.
5. ASCE Task Committee. 2000. Artificial neural network in hydrology. Journal of Hydrologic Engineering, 5: 124-144.
6. Béhaegel, M., Sailhac, P., and Marquis G. 2007. On the use of surface and ground temperature data to recover soil water content information. J. Appl. Geophys., 62: 234-243.
7. Braddock, R.D., Kremmer, M.L and Sanzogni, L. 1998. Feed forward artificial neural network model for forecasting rainfall-runoff. Environmental Sciences, 9: 419-432.
8. Brar, G. S., Steiner, J. L., Unger, P. W., and Prihar, S. S. 1992. Modeling sorghum seedling establishment from soil wetness and temperature of drying seed zones. Agronomy Journal, 84: 905-910.
9. Camdevyren, H., Demyr, N., Kanik, A., and Keskin, S. 2005. Use of principal component scores in multiple linear regression models for prediction of Chlorophyll- a in reservoirs. Ecological Modeling, 181: 581-589.
10. Chang, D.H., and Islam, S. 2000. Estimation of Soil Physical Properties Using Remote Sensing and Artificial neural network. Remote Sensing of Environment, 74: 534-544.
11. Chen, J.C., Chang, N.B., and Shieh, W.K. 2003. Assessing wastewater reclamation potential by neural network model. Journal of Engineering Application of Artificial Intelligence, 16: 149-157.
12. Choi, D.J., and Park, H. 2001. A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process. Water Res. 35: 3959-3967.

13. Coulibaly, P., Anctil, F. and Bobe B. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, 230:244-257.
14. Dawson, C.W. and Wilby, R. 1998. An artificial neural network approach to rainfall-runoff modeling. *Journal of Hydrological Sciences*, 43: 14-66.
15. Granger, C. W. J. 1993. Strategies for modeling nonlinear time series relationships. *The Economic Record*. 69: 233-238.
16. Grundmann, G. L., Renault, P., Rosso, L. and Bardin, R. 1995. Differential effects of soil water content and temperature on nitrification and aeration. *Soil Science Society of America Journal*, 59: 1342-1349.
17. Guang-ming, Z., Hong-wei, L.U., Xiang-can, JIN. and XU Min. 2005. Assessment of the Water Quality and Nutrition of the Dongting Lake with Wavelet Neural Network. *Journal of Hunan University*, 2005-01: 3291-94.
18. Haykin, S. 1999. *Neural networks: a comprehensive foundation*. 2nd Ed, Prentice Hall, New Jersey, USA.
19. Johnson, R. A. and Wichern, D. W. 1982. *Applied multivariate statistical analysis*. Prentice-Hall Inc, Englewood Cliffs, SA, 590 pp.
20. Kanani, S., Asadollahfardi, G. and Ghanbari A. 2008. Application of Artificial Neural Network to Predict Total Dissolved Solid in Achechay River Basin. *World Applied Sciences Journal*, 4 (5): 646-654.
21. Kemp, P. R., Cornelius, J. M., and Reynolds, J. F. 1992. A simple model for predicting soil temperatures in desert ecosystems. *Journal of Soil Science*, 153: 280-287.
22. Kluender, R. A., Thompson, L. C., and Steigerwald, D. M. 1993. A conceptual model for predicting soil temperatures. *Soil Science*, 156:10-19.
23. Lu, W.Z., Wang, W.J., Wang, X.K., Xu, Z.B. and Leung, A.Y.T. 2003. Using improved neural network to analyze RSP, NOX and NO2 levels in urban air in Mong Kok, Hong Kong. *Environmental Monitoring and Assessment*, 87: 235-254.
24. McCann, I.R., McFarland, M.J. and Witz, J.A. 1991. Near surface bare soil temperature model for biophysical models. *Trans. Am. Soc. Agric. Eng.* 34: 748-755.
25. Mihalakakou, G. 2002. On estimating soil surface temperature profiles. *Energy and Buil.*, 34: 251-259.
26. Ramanujam, S., Chandrasekar, R. and Chakravarthy, B. 2011. A new PCA-ANN algorithm for retrieval of rainfall structure in a precipitating atmosphere. *International Journal of Numerical Methods for Heat & Fluid Flow*, 21:1002-1025.
27. Sezin Tokar, A. and Momcilo, M. 2000. Precipitation. Runoff modeling using artificial neural networks and conceptual models. *Journal of Hydrologic Engineering*, 5: 156-161.
28. Shrestha, S. and Kazama, F. 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fujiriver basin, Japan. *Environ. Model & Softw*, 22: 464-475.
29. Singh, K. P., Malik, A., Mohan, D. and Sinha, S. 2004. Multivariate statistical techniques of the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study. *Water Res.*, 38: 3980-3992.
30. Singh, V. P. 1992. *Elementary hydrology*. Prentice-Hall, Englewood Cliffs, NJ. 973 pp.
31. Tabari, H., Sabziparvar, A., and Ahmadi, M. 2011. Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region. *Journal of Meteorology and Atmospheric Physics*, 110: 135-142.
32. Tokar, A.S. and Markus, M. 2000. Precipitation runoff modeling using artificial neural network and conceptual models. *J. Hydrol. Eng., ASCE*. 5: 156-161.

33. Vigil, M.F. and Kissel, D.E. 1995. Rate of nitrogen mineralized from incorporated crop residues as influenced by temperature. *Soil Science Society of America Journal*, 59: 1636–1644.
34. Yang, C.C., Prasher, S.O. and Mehuys, G.R. 1997. An artificial neural network to estimate soil temperature. *Can. J. Soil Sci.*, 77: 421–429.
35. Yin, X. and Arp, P.A. 1993. Predicting forest soil temperatures from monthly air temperature and precipitation records. *Can. J. For. Res.*, 23: 2521–2536.
36. Zhang, G., Patuwo, B.E. and Hu, M.Y. 1998. Forecasting with Artificial Neural Network: The state of art. *International Journal of Forecasting*, 14: 35-62.
37. Zhang, Y., Li, H., Hou, A. and Havel, J. 2006. Artificial neural networks based on principal component analysis input selection for quantification in overlapped capillary electro phoresis peaks. *Chemometrics and Intelligent Laboratory Systems*, 82 (1-2): 165-175.
38. Zhang, Y.X. 2007. Artificial neural networks based on principal component analysis input selection for clinical pattern recognition analysis. *Talanta*, 73: 68-75.