

## پیش‌بینی احتمالاتی دماهای کمینه و بیشینه روزانه برای ایران با استفاده از سامانه همادی دو عضوی

مجید آزادی\*<sup>۱</sup>، سید عاطفه محمدی<sup>۲</sup>

۱- دانشیار، پژوهشگاه هواشناسی و علوم جو

۲- پژوهشگاه هواشناسی و علوم جو

(دریافت: ۱۳۹۷/۰۹/۲۶، بازنگری: ۱۳۹۷/۱۲/۱۳، پذیرش: ۱۳۹۸/۰۳/۱۸، انتشار آنلاین: ۱۳۹۸/۰۷/۰۱)

### چکیده

برای کمی کردن عدم قطعیت در پیش‌بینی وضع هوا و صدور پیش‌بینی احتمالاتی دماهای کمینه و بیشینه روزانه، در این تحقیق از دو روش آماری موسوم به آماره برونداد مدل (EMOS) و مدل میانگین‌گیری بیزی (BMA) برای برآورد تابع چگالی احتمال پیش‌بینی استفاده شده است. سامانه همادی در تحقیق حاضر، دو عضوی است که شامل پیش‌بینی‌های یک روزه تا پنج روزه دماهای کمینه و بیشینه روزانه برای ایستگاه‌های همدید مراکز استان‌های کشور در بازه زمانی از ۱۰ نوامبر ۲۰۱۷ تا ۳۱ مه ۲۰۱۸ با دو پیکربندی مختلف مدل WRF است. نتایج بدست آمده از این روش‌ها، خطای بسیار کمتری نسبت به خروجی منفرد خام مدل دارد؛ به طوری که متوسط میزان بهبود جذر میانگین مربعات خطای پیش‌بینی قطعی دماهای بیشینه و کمینه EMOS نسبت به میانگین پیش‌بینی خام دماهای بیشینه و کمینه به ترتیب ۳۷٪ و ۷٪ است. به طور کلی پیش‌بینی خام دمای کمینه، خطای کمتری نسبت به دمای بیشینه دارد اما بعد از پس‌پردازش نتایج بهبود چندانی ندارد. میانگین بدست آمده برای توابع چگالی احتمال در مقایسه با میانگین اعضای سامانه همادی، دارای و میانگین قدرمطلق خطای کمتری است.

**کلمات کلیدی:** پیش‌بینی عددی وضع هوا، دمای کمینه و بیشینه، سامانه همادی، پیش‌بینی احتمالاتی، تابع چگالی احتمال.

### ۱. مقدمه

توانایی مدل در لحاظ کردن همه‌ی فرآیندهای جو، کمبود داده‌های اولیه در برخی مناطق و آشوبی بودن سیستم دینامیکی جو نیز می‌توان اشاره کرد که باعث می‌شوند پیش‌بینی وضع هوا همراه با عدم قطعیت باشد (وارنر، ۲۰۱۱). برای تشخیص چگونگی وضع هوای آینده، علاوه بر پیش‌بینی قطعی رویکرد دیگری تحت عنوان پیش‌بینی احتمالاتی نیز وجود دارد. بحث پیش‌بینی احتمالاتی، با اعلام آشوبناکی سیستم دینامیکی جو توسط لورنز (۱۹۶۳) آغاز شد. با توجه به نظریه آشوب، سیستم‌های دینامیکی که کوچک‌ترین تغییر در شرایط اولیه آن‌ها منجر به تغییرات زیاد در نتیجه سیستم خواهد شد، آشوبناک نامیده می‌شوند. به

از سال ۱۹۵۰ با پیدایش رایانه‌ها با قدرت محاسباتی بالا، امکان اجرای مدل‌های پیش‌بینی عددی وضع هوا فراهم شده است. مدل‌های پیش‌بینی عددی وضع هوا شامل معادلات دیفرانسیلی هستند که قوانین فیزیکی و دینامیکی جو را توصیف می‌کنند. این معادلات، وابسته به مقادیر اولیه-مرزی هستند و با استفاده از روش‌های عددی، حل می‌شوند. به دلیل راه‌حل‌های تقریبی در روش‌های عددی، جواب به دست آمده از اجرای مدل‌های عددی، تقریبی از جواب واقعی است و همواره دارای خطا است. البته استفاده از روش‌های عددی در مدل‌ها، تنها علت بروز خطا در خروجی مدل نیست؛ بلکه به عواملی مانند وجود خطا در مقادیر اولیه-مرزی، عدم

می‌باشد نیز دارای خطاهای پراکندگی و نیز جابجایی میانگین توزیع خواهد بود (ویلکس، ۲۰۱۱). برای رفع این نواقص و کاهش خطا و دستیابی به تابع چگالی احتمال وضع هوای آینده، تعدادی روش‌های پس‌پردازش آماری مانند هیستوگرام رتبه‌ای (همیل و کلوجی<sup>۲</sup>، ۱۹۹۸)، برازش لجستیک (همیل و همکاران، ۲۰۰۴)، روش درسینگ (رولستون و اسمیت<sup>۳</sup>، ۲۰۰۳)، میانگین‌گیری به روش بیزی (BMA<sup>۴</sup>) (رافتری<sup>۵</sup> و همکاران، ۲۰۰۵) آماره برون‌داد مدل همدادی (EMOS<sup>۶</sup>) (نیتینگ<sup>۷</sup> و همکاران، ۲۰۰۵)، رگرسیون همدادی (آنگر<sup>۸</sup>، ۲۰۰۹)، تکنیک‌های تورم واریانس (جانسون و بولر<sup>۹</sup>، ۲۰۰۹)، کالیبره انتقال و کشش (اکل<sup>۱۰</sup> و همکاران، ۲۰۱۲) و روش EKDMO (گلان<sup>۱۱</sup>، ۲۰۰۹) ارائه شده است. در تحقیق حاضر، برای دستیابی به تابع چگالی احتمال پیش‌بینی دماهای بیشینه و کمینه از روش‌های EMOS و BMA استفاده شده است. در روش EMOS، میانگین و واریانس تابع چگالی احتمال نرمال، ترکیب خطی اعضای سامانه همدادی و واریانس آنها است. تخمین ضرایب با کمینه کردن امتیاز احتمالی رتبه‌ای پیوسته (CRPS<sup>۱۲</sup>) با استفاده از الگوریتم برویدن - فلچر - گلدفرب - شانو (BFGS<sup>۱۳</sup>) (پرس<sup>۱۴</sup> و همکاران، ۱۹۹۲) انجام می‌شود. در روش BMA، تابع چگالی احتمال نهایی از میانگین‌گیری وزنی هر یک از اعضای همدادی بدست می‌آید. در این روش ضرایب و وزن‌ها، با بیشینه کردن درستیابی<sup>۱۵</sup> تخمین زده می‌شوند. در بخش ۲، توضیحاتی در مورد روش‌های پس‌پردازش همدادی که در تحقیق حاضر به کار رفته‌اند (EMOS و BMA)، آورده شده است. معرفی داده‌ها و سامانه همدادی در بخش ۳،

دلیل عدم انطباق شرایط اولیه در نظر گرفته شده برای مدل‌ها با واقعیت و با توجه به آشوبناکی سیستم دینامیکی جو، در نظر گرفتن چندین مقدار اولیه برای پیش‌بینی وضع هوا منطقی به نظر می‌رسد.

این در حالی است که در پیش‌بینی‌های قطعی یک مقدار اولیه نزدیک به شرایط واقعی جو به عنوان ورودی به مدل‌های عددی داده می‌شود و یک خروجی به عنوان محتمل‌ترین اتفاق در آینده بدست می‌آید. در این روش، وضعیت آینده جو بدون در نظر گرفتن عدم قطعیت ذاتی آن به طور قطعی پیش‌بینی می‌شود. اما با صدور پیش‌بینی احتمالاتی که شانس وقوع حالات ممکن آینده جو را محاسبه می‌کند، می‌توان عدم قطعیت را کمی ساخت و به اطلاعات بیشتر و دقیق‌تری دست یافت. به این طریق که به جای در نظر گرفتن فقط یک مقدار اولیه، می‌توان با نمونه‌گیری متناهی از تابع چگالی احتمال مربوط به عدم قطعیت شرایط اولیه و اجرای کامل مدل عددی با توجه به شرایط هر نمونه، یک سامانه همدادی با تعداد اعضای متناهی تولید کرد. با اعمال روش‌های پس‌پردازش آماری روی اعضای سامانه همدادی، یک تابع توزیع احتمال به دست خواهد آمد که عدم قطعیت وضعیت آینده جو را توصیف می‌کند و شامل اطلاعات کافی برای نیازهای کاربران مختلف است (ویلکس<sup>۱</sup>، ۲۰۱۱). از آنجا که ناهمواری‌ها و کاربری اراضی در مدل و واقعیت تفاوت زیادی با هم دارند و نیز به علت خطاهای نسبتاً زیاد در بخش فیزیک مدل، خروجی مدل‌های عددی پیش‌بینی وضع هوا همواره دارای خطای سامانمند می‌باشد. در نتیجه خروجی سامانه همدادی که مجموعه خروجی منفرد مدل‌های عددی

<sup>9</sup> Johnson & Bowler

<sup>10</sup> Eckel

<sup>11</sup> Glahn

<sup>12</sup> Continuous Ranked Probability Score

<sup>13</sup> Broyden-Fletcher-Goldfarb-Shanno

<sup>14</sup> Press

<sup>15</sup> Maximum Likelihood

<sup>1</sup> Wilks

<sup>2</sup> Hamill & Colucci

<sup>3</sup> Roulston & Smith

<sup>4</sup> Bayesian Model Averaging

<sup>5</sup> Raftery

<sup>6</sup> Ensemble Model Output Statistics

<sup>7</sup> Gneiting

<sup>8</sup> Unger

تخمین تعداد روزهای دوره آموزش در بخش ۴ و تحلیل نتایج حاصل در بخش ۵ ارائه شده است. بخش ۶ نیز یک جمع‌بندی کلی از نتایج مقاله است.

## ۲- پس‌پردازش سامانه همادی

همان‌طور که پیشتر گفته شد، روش‌های متنوعی برای پس‌پردازش آماری برون‌داد یک سامانه همادی معرفی شده است؛ اما روش‌های مورد توجه در تحقیق حاضر، روش EMOS (نیتینگ و همکاران، ۲۰۰۵) و روش BMA (رافتری و همکاران، ۲۰۰۵) برای پس‌پردازش سامانه همادی هستند. حاصل به کارگیری این روش‌ها، یک تابع چگالی احتمال نرمال برای پیش‌بینی دمای کمینه و بیشینه روزانه می‌باشد که در ادامه توضیحاتی راجع به تخمین آن آورده شده است.

### ۲-۱ روش EMOS

در روش EMOS، میانگین تابع چگالی احتمال نرمال از میانگین‌گیری وزنی اعضای سامانه همادی به دست می‌آید به طوری که ضرایب وزن‌ها از اطلاعات داده‌های دوره آموزش محاسبه می‌شوند. واریانس تابع چگالی احتمال در این روش، یک تابع خطی از واریانس اعضای همادی می‌باشد که در ادامه توضیحات بیشتر آورده شده است.

فرض کنید  $X_1, \dots, X_m$  اعضای سامانه همادی و  $S^2$  واریانس اعضای همادی باشد. تابع چگالی احتمال توزیع نرمال سامانه همادی به این صورت تعریف می‌شود (نیتینگ و همکاران، ۲۰۰۵):

$$N(a + b_1 X_1 + \dots + b_m X_m, c + d S^2) \quad (1)$$

که در آن، میانگین  $a + b_1 X_1 + \dots + b_m X_m$  واریانس مجموعه همادی با اریبی تصحیح شده و واریانس  $c + d S^2$

یک تابع خطی از واریانس همادی است. ضرایب  $b_1, \dots, b_m$  نشان‌دهنده عملکرد اعضای همادی با توجه به داده‌های دوره آموزش می‌باشد. ضرایب واریانس  $c$  و  $d$  نشان‌دهنده پراکندگی همادی و عملکرد پیش‌بینی میانگین همادی است. امتیاز احتمال رتبه‌ای پیوسته (CRPS) روشی برای راستی-آزمایی پیش‌بینی احتمالاتی می‌باشد (ویلکس، ۲۰۱۱)، به طوری که اگر  $F$  تابع توزیع احتمال پیش‌بینی و  $y$  مشاهده دیدبانی باشد به صورت زیر تعریف می‌شود:

$$crps(F, y) = \int_{-\infty}^{+\infty} (F(t) - H(t - y))^2 dt \quad (2)$$

که  $H(t - y)$  تابع هویساید<sup>۱</sup> با تعریف زیر است:

$$H(t - y) = \begin{cases} 0 & t < y \\ 1 & \text{در غیر این صورت} \end{cases} \quad (3)$$

هرگاه  $F$  تابع توزیع احتمال نرمال با میانگین  $\mu$  و واریانس  $\sigma^2$  باشد، با انتگرال‌گیری جزء به جزء به رابطه زیر می‌رسیم:

$$crps(N(\mu, \sigma^2), y) \quad (4)$$

$$= \sigma \left\{ \frac{y - \mu}{\sigma} \left( 2\Phi \left( \frac{y - \mu}{\sigma} \right) \right) + 2\varphi \left( \frac{y - \mu}{\sigma} \right) - \frac{1}{\sqrt{\pi}} \right\}$$

که  $\varphi \left( \frac{y - \mu}{\sigma} \right)$  و  $\Phi \left( \frac{y - \mu}{\sigma} \right)$  به ترتیب تابع چگالی و تابع توزیع احتمال نرمال استاندارد است.

اگر  $F_i$ ها پیش‌بینی قطعی باشند، میانگین امتیاز  $CRPS = \frac{1}{n} \sum_{i=1}^n crps(F_i, y_i)$  به میانگین قدر مطلق خطا تبدیل می‌شود. در واقع میانگین امتیاز، مدل تعمیم‌یافته میانگین قدر مطلق خطا است که هرچه به صفر نزدیکتر باشد، پیش‌بینی تیزتر و کالیبره‌تر خواهد بود. بنابراین می‌توان با کمینه کردن امتیاز احتمال رتبه‌ای پیوسته (CRPS)، ضرایب  $a, b_1, \dots, b_m, c$  و  $d$  را تخمین زد.

### ۲-۲ روش BMA

<sup>1</sup> Heaviside

می‌خواهیم  $a_m$  و  $b_m$  و  $w_m$  برای  $m = 1, \dots, M$  و  $\sigma$  را از روی داده‌های دوره آموزش شامل پیش‌بینی‌های همادی و مشاهدات تخمین بزنیم. برای تخمین پارامترهای  $a_m$  و  $b_m$  از رگرسیون خطی ساده بین  $o_t$  و  $x_{tm}$  روی داده‌های دوره آموزش و برای تخمین  $w_m$  و  $\sigma$  از بیشینه درست‌نمایی به روی داده‌های دوره آموزش استفاده می‌شود. از آنجا که بیشینه کردن حالت لگاریتمی تابع درست‌نمایی ساده‌تر است، بنابراین تابع درست‌نمایی به صورت لگاریتمی نوشته می‌شود:

$$l(w_1, \dots, w_m, \sigma^2) = \sum_t \log \left( \sum_{m=1}^M w_m g_m(o_t | x_{tm}) \right) \quad (8)$$

بیشینه کردن این تابع، با استفاده از روش‌های عددی مانند روش نیوتن-رافسون بسیار پیچیده است. بنابراین رافتری و همکاران (۲۰۰۵) از الگوریتمی به نام EM<sup>۱</sup> برای بیشینه کردن تابع درست‌نمایی استفاده کردند. برای این منظور یک متغیر کمکی  $z_{tm}$  را معرفی نمودند به طوری که اگر  $m$ امین عضو مجموعه همادی بهترین پیش‌بینی برای زمان  $t$  باشد، آنگاه  $z_{tm} = 1$  و در غیر این صورت  $z_{tm} = 0$ . با این تعریف فقط یکی از  $[z_{t1}, \dots, z_{tM}]$  برابر با یک و بقیه برابر با صفر هستند. الگوریتم EM تکراری است و بین دو گام E و M تکرار می‌شود. این الگوریتم با یک حدس اولیه  $\theta^{(0)}$  برای بردار پارامتر  $\theta$  شروع می‌شود. در گام E،  $z_{tm}$  با حدس فعلی برای پارامترها تخمین زده می‌شود. مقدار بدست آمده  $z_{tm}$  لزوماً عدد صحیح نیست اگرچه مقدار واقعی آن صفر و یا یک است. در گام M، بردار پارامترها  $\theta$  با استفاده از  $z_{tm}$  تخمین زده شده در گام قبل، محاسبه می‌شود.

گام E با توجه به فرمول‌های BMA به قرار زیر است:

$$\hat{z}_{tm}^{(j)} = \frac{w_m^{(j-1)} g(o_t | x_{tm}, \sigma^{(j-1)})}{\sum_{i=1}^M w_i^{(j-1)} g(o_t | x_{ti}, \sigma^{(j-1)})} \quad (9)$$

این روش از جمله شیوه‌های مهم در پس‌پردازش پیش‌بینی‌های همادی برای برون‌داد اعضای سامانه همادی است که توسط رافتری و همکاران (۲۰۰۵) تکمیل و برای کمیت‌های دما و فشار جو در سطح دریا ارائه شده است. با استفاده از این روش ابتدا با برازش یک تابع چگالی احتمال (PDF<sup>۱</sup>) به پیش‌بینی هر عضو از مجموعه در بازه آموزش، وزن آن عضو با توجه به عملکردش در دوره آموزش مشخص می‌شود. سپس تابع چگالی احتمال نهایی با روش میانگین‌گیری بیزی استخراج می‌شود. به بیان دیگر، میزان مهارت هر عضو، سهم مشارکت آن را در صدور پیش‌بینی نهایی تعیین می‌کند. توضیحاتی که در ادامه درباره این روش آورده شده است، از پژوهش رافتری و همکاران (۲۰۰۵) اقتباس شده است. فرض کنید  $x_t = (x_{t1}, x_{t2}, \dots, x_{tM})$  یک سامانه همادی  $M$  عضوی در روز  $t$ ام،  $o_t$  مقدار مشاهده در روز  $t$ ام از دوره آموزش،  $N$  تعداد روزهای دوره آموزش و  $y_t$  پیش‌بینی احتمالاتی در روز  $t$ ام باشد. مدل پیش‌بینی BMA به صورت فرمول زیر است:

$$p(y_t | x_t) = \sum_{m=1}^M w_m g_m(o_t | x_{tm}) \quad (5)$$

که  $g$  تابع چگالی احتمال و  $w_m$  وزن  $m$ امین عضو همادی است و مقدار آن با توجه به عملکرد  $m$ امین عضو پیش‌بینی در دوره آموزش تعیین می‌شود.  $w_m$  ها مثبت بوده و جمع آنها برابر با یک است:

$$\sum_{m=1}^M w_m = 1 \quad (6)$$

اگر PDF مشروط، دارای توزیع نرمال باشد، مانند پیش‌بینی دما یا فشار سطح دریا، آنگاه فرمول زیر را داریم:

$$y_t | x_{tm} \sim N(a_m + b_m x_{tm}, \sigma^2) \quad (7)$$

برای متغیرهایی مانند سرعت باد و بارش، توزیع نرمال مناسب نیست و باید از توزیع‌های دیگر استفاده شود. در اینجا

<sup>۱</sup>Expectation Maximization

<sup>۱</sup> Probability Density Function

شمالی و ۴۲ تا ۶۵ درجه شرقی است. تفکیک مکانی برای دامنه بزرگ تر و کوچک تر به ترتیب ۲۷ و ۹ کیلومتر است. پیش‌بینی‌های سامانه همدادی در ساعت 12 UTC شروع شده‌اند و در بازه زمانی از ۱۰ نوامبر ۲۰۱۷ تا ۳۱ مه ۲۰۱۸ موجود هستند. خروجی‌های مدل (اعضای همدادی) با روش دوخطی<sup>۲</sup> به ایستگاه‌های دیدبانی درونیایی شده‌اند. علاوه بر داده‌های پیش‌بینی، داده‌های مشاهداتی شامل دماهای بیشینه و کمینه روزانه در ۳۱ ایستگاه هواشناسی سینوپتیک مراکز استان‌های ایران در بازه زمانی یاد شده است.

جدول ۱. پیکربندی‌های مختلف مدل WRF

عضو همدادی		
۲	۱	
Dudhia/rrtm	Goddard/rrtm	فیزیک تابش موج بلند/ موج کوتاه
Ferrier	Lin	خرد فیزیک
GD ensemble	KF	فیزیک همرفت
RUC	Unified Noah	سطح زمین
MM5 MO	MOJ	لایه سطحی
YSU	MYJ TKE	لایه مرزی

#### ۴- دوره آموزش

در دسترس بودن داده‌ها، سهولت محاسبات و هم‌چنین توانایی سازگاری سریع به تغییرات فصلی در الگوهای اریبی مدل و تغییرات در شرایط زیست‌محیطی، از مزایای دوره آموزش کوتاه مدت است. از طرفی دیگر، در دوره آموزش بلندمدت، تغییرپذیری آماری<sup>۳</sup> کاهش می‌یابد که منجر به تخمین بهتر پارامترها می‌شود. دوره آموزش در پژوهش حاضر، یک پنجره زمانی لغزان شامل N روز اخیر قبل از روز پیش‌بینی است.

که بالانویس z مربوط به زامین تکرار الگوریتم EM است و  $g(o_t|x_{tm}, \sigma^{(j-1)})$  چگالی نرمال با میانگین  $a_k + b_k x_{tm}$  و انحراف معیار  $\sigma^{(j-1)}$  در تکرار  $(j-1)$  الگوریتم است. گام M شامل تخمین  $w_m$  و  $\sigma$  با استفاده از فرمول‌های زیر قابل محاسبه می‌باشد:

$$w_m^{(j)} = \frac{1}{N} \sum_t \hat{z}_{tm}^{(j)}, \quad (10)$$

$$\sigma^{2(j)} = \frac{1}{N} \sum_t \sum_{m=1}^M \hat{z}_{tm}^{(j)} (o_t - x_{tm})^2. \quad (11)$$

پیش‌بینی احتمالاتی رویداد دودویی برای آستانه q به صورت زیر است:

$$\Pr(y_t \leq q) = \sum_{m=1}^M w_m G_m(q|\hat{x}_{tm}). \quad (12)$$

که G تابع توزیع تجمعی است.

#### ۳- سامانه همدادی و داده‌ها

سامانه همدادی مورد استفاده در تحقیق حاضر، شامل دو عضو است که هر عضو از اجرای مدل WRF با طرح‌واره‌های فیزیکی مختلف (جدول ۱) به دست آمده است (محمدی و همکاران، ۲۰۱۶). برای شرایط اولیه در اجرای مدل، از داده‌های GFS<sup>۱</sup> با تفکیک افقی ۰/۵ درجه استفاده شده است. مدل WRF با دو حوزه اجرا شده است که دامنه بزرگ‌تر قسمت جنوب‌غربی خاورمیانه (از ۱۰ تا ۵۱ درجه شمالی و ۲۰ تا ۸۰ درجه شرقی) را می‌پوشاند و دامنه داخلی (کوچک‌تر) ایران را پوشش می‌دهد که از ۲۳ تا ۴۱ درجه

<sup>3</sup> Statistical variability

<sup>1</sup> Global Forecast System

<sup>2</sup> Bilinear

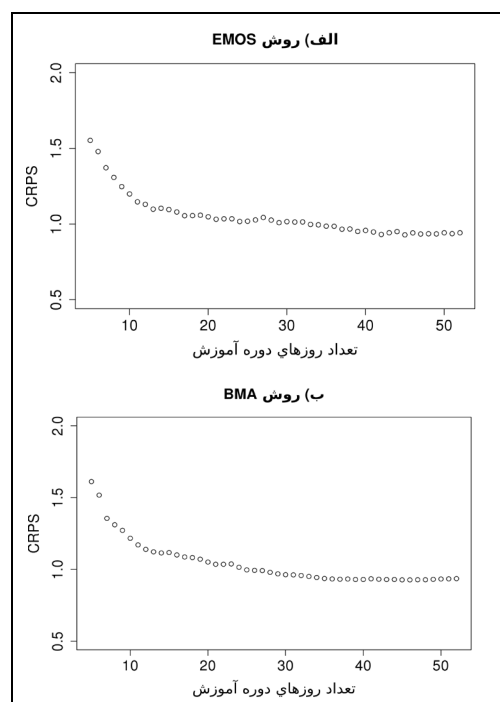
## ۵- نتایج حاصل از اعمال روش‌های پس‌پردازش

برای ارزیابی عملکرد روش‌های پس‌پردازش همادی به کار رفته (روش‌های EMOS و BMA)، میانگین تابع چگالی احتمال به دست آمده از اعمال این روش‌ها به عنوان یک پیش‌بینی قطعی در نظر گرفته شده و مقدار میانگین قدر مطلق خطا (MAE)<sup>۱</sup> و جذر میانگین مربعات خطای (RMSE)<sup>۲</sup> آنها محاسبه شده است. از آنجا که هر عضو همادی یک پیش‌بینی قطعی است، اربیی هر عضو از دوره آزمون با روش میانگین‌گیری لغزان (استنشورد<sup>۳</sup> و همکاران، ۱۹۹۶) با دوره آموزش لغزان ۱۵ روزه حذف شده است. بنابراین نتایج حاصل از پس‌پردازش همادی با پیش‌بینی‌های خام و پیش-بینی‌هایی که اربیی آنها به روش میانگین‌گیری لغزان حذف شده، مقایسه شده است.

مقدار خطای پیش‌گفته در پیش‌بینی‌های یک روزه تا پنج روزه دماهای کمینه و بیشینه به روش‌های مختلف روی ۴۶۸۱ مورد (31 × 151) در غالب یک نمودار میله‌ای در شکل ۲ نشان داده شده است. در شکل ۲، منظور از f1 پیش‌بینی خام اول (عضو اول همادی)، f2 پیش‌بینی خام دوم (عضو دوم همادی)، mean\_raw میانگین ساده پیش‌بینی‌های خام (اعضای اول و دوم همادی)، f1\_post پیش‌بینی اول که اربیی آن حذف شده، f2\_post پیش‌بینی دوم که اربیی آن حذف شده، mean\_post میانگین ساده پیش‌بینی‌هایی که اربیی آنها حذف شده، mean\_BMA میانگین تابع چگالی احتمال بدست آمده به روش BMA و mean\_EMOS میانگین تابع چگالی احتمال بدست آمده به روش EMOS است.

نتایج شکل ۲ نشان می‌دهد که اگرچه در اکثر مواقع پیش-بینی EMOS خطای کمتری نسبت به پیش‌بینی BMA

برای تعیین تعداد روزهای مناسب دوره آموزش، مقدار CRPS پیش‌بینی احتمالاتی برای تعداد مختلف روزهای دوره آموزش از ۵ تا ۵۲ روز با استفاده از هر دو روش EMOS و BMA محاسبه شده است. باید تذکر داد که از یک دوره آزمون یکسان برای ارزیابی همه دوره‌های آموزش استفاده شده است؛ به این صورت که از همه ۲۰۲ روز داده موجود، ۵۲ روز اول به عنوان دوره آموزش و ۱۵۱ روز باقیمانده (از ۱ ژانویه ۲۰۱۸ تا ۳۱ مه ۲۰۱۸) به عنوان دوره آزمون در نظر گرفته شده است. همان طور که در شکل ۱ دیده می‌شود، در هر دو روش مقدار CRPS برای دوره آموزش تا ۲۵ روز روند کاهشی دارد و از آنجا به بعد تقریباً ثابت مانده است. از آنجا که هر چه مقدار CRPS کمتر باشد پیش‌بینی به واقعیت نزدیک‌تر است، بنابراین دوره آموزش ۲۵ روزه انتخاب شده است.



شکل ۱. مقایسه تعداد روزهای مختلف دوره آموزش؛ الف) به روش EMOS و ب) به روش BMA.

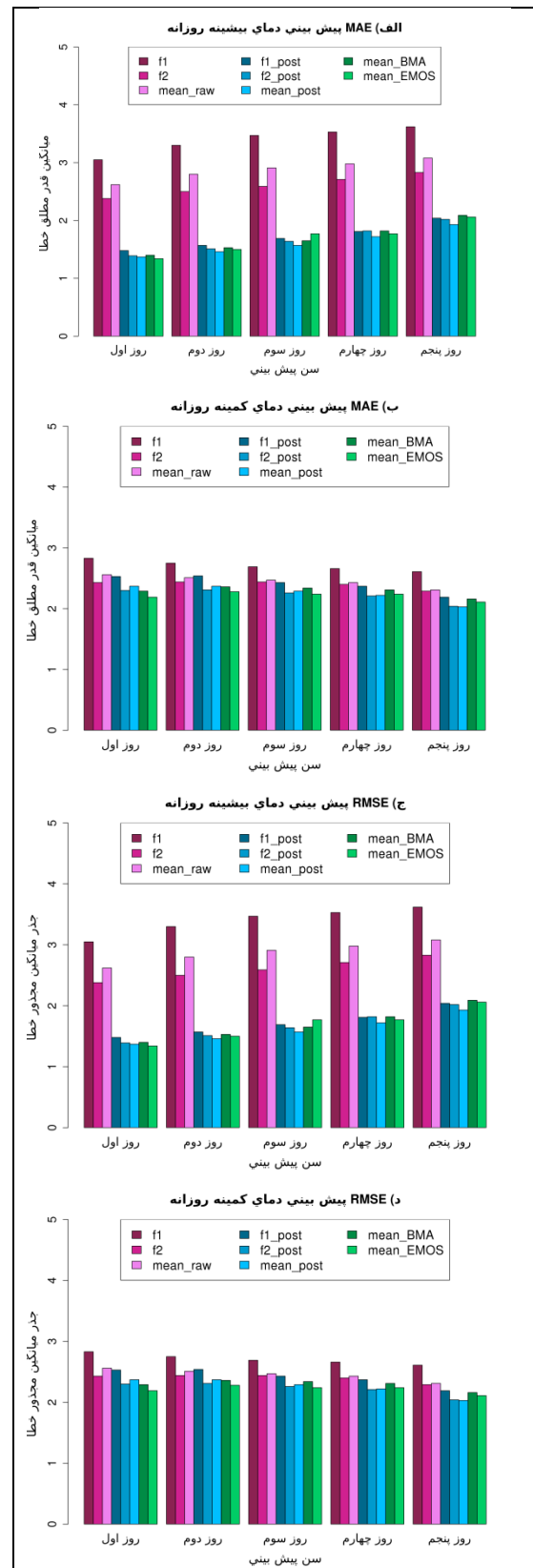
<sup>3</sup> Stensrud

<sup>1</sup> Mean Absolute Error

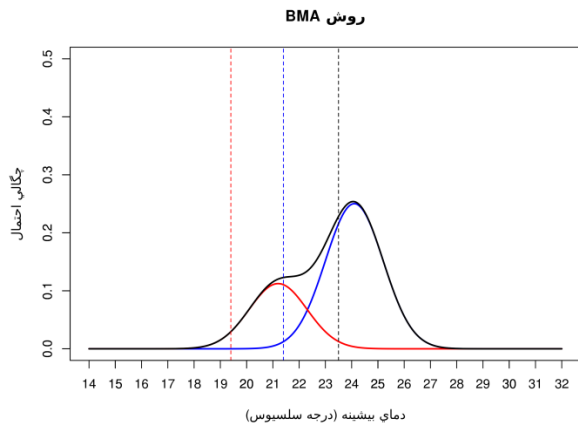
<sup>2</sup> Root Mean Absolute Error

درد اما تفاوتشان معنادار نیست و می‌توان عملکردشان را یکسان فرض نمود. همچنین هر دو پیش‌بینی‌های EMOS و BMA نسبت به پیش‌بینی‌های خام خطای کمتری دارند و نسبت به پیش‌بینی‌های با اریبی حذف شده به روش میانگین-گیری لغزان خطای نسبتاً مشابهی دارند که علت این امر، تعداد کم اعضای همادی (دو عضو) است. با ارجاع به تحقیق صورت گرفته توسط محمدی و همکاران (۲۰۱۶) برای پیش‌بینی ۴۸ ساعته دمای دومتری در مراکز استان‌های ایران با استفاده از یک سامانه همادی ۹ عضوی به روش EMOS، مقدار MAE و RMSE پیش‌بینی همادی به طور معناداری بسیار کمتر از پیش‌بینی پس‌پردازش شده است. بنابراین با افزایش تعداد اعضای سامانه همادی می‌توان به پیش‌بینی‌های قطعی‌تری دست یافت.

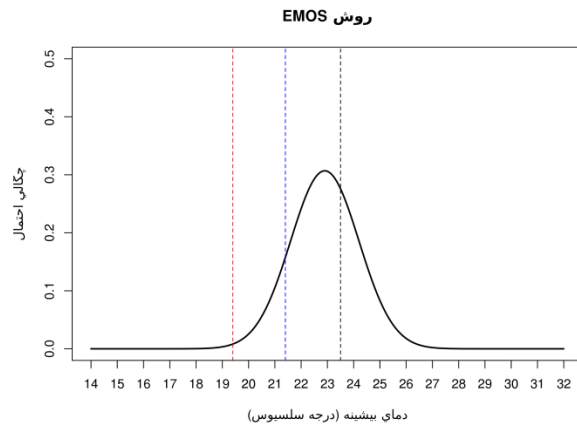
با پس‌پردازش سامانه همادی، علاوه بر دستیابی به یک پیش-بینی قطعی با دقت بالا، تابع چگالی احتمال پیش‌بینی نیز بدست می‌آید که اصلی‌ترین هدف از صدور پیش‌بینی‌های همادی است. با دسترسی به تابع چگالی احتمال پیش‌بینی، کاربر می‌تواند اطلاعات زیادی از چگونگی وضع هوای آینده استخراج کند. شکل ۳ نمونه‌ای از تابع چگالی احتمال پیش‌بینی دو روزه دمای بیشینه برای ایستگاه مهرآباد در تاریخ دهم مه ۲۰۱۸ را نشان می‌دهد که با استفاده از روش EMOS بدست آمده است. تابع چگالی احتمال مذکور دارای توزیع نرمال با میانگین ۲۲/۹ و انحراف معیار ۱/۳ است که با استفاده از آن می‌توان احتمال وقوع هر آستانه‌ای را تخمین زد. همان-طور که در شکل ۳ دیده می‌شود، میانگین تابع چگالی احتمال نرمال نسبت به هر یک از اعضای همادی، به مقدار مشاهداتی نزدیکتر است و واریانس توزیع نرمال میزان عدم قطعیت موجود در پیش‌بینی را نشان می‌دهد.



شکل ۲. پیش‌بینی دماهای کمینه و بیشینه روزانه در روش‌های مختلف؛ الف) MAE پیش‌بینی دمای بیشینه، ب) MAE پیش‌بینی دمای کمینه، ج) RMSE پیش‌بینی دمای بیشینه و د) RMSE پیش‌بینی دمای کمینه.



شکل ۴. توزیع نرمال برازش شده برای پیش‌بینی دو روزه دمای بیشینه برای ایستگاه مهرآباد در تاریخ دهم مه ۲۰۱۸ به روش BMA؛ خط‌چین‌های آبی، قرمز و مشکی به ترتیب مربوط به مقدار پیش‌بینی عضو اول، عضو دوم و مقدار مشاهداتی هستند. توزیع نرمال‌های آبی و قرمز به ترتیب مربوط به عضو اول و دوم و توزیع نهایی به رنگ مشکی است.



شکل ۳. توزیع نرمال برازش شده برای پیش‌بینی دو روزه دمای بیشینه برای ایستگاه مهرآباد در تاریخ دهم مه ۲۰۱۸ به روش EMOS؛ خط‌چین‌های آبی، قرمز و مشکی به ترتیب مربوط به مقدار پیش‌بینی عضو اول، عضو دوم و مقدار مشاهداتی هستند.

## ۶. جمع‌بندی

در پیش‌بینی‌های قطعی، یک مقدار اولیه نزدیک به شرایط واقعی جو به عنوان ورودی به مدل‌های عددی داده می‌شود و یک خروجی به عنوان محتمل‌ترین اتفاق در آینده بدست می‌آید. در این روش، به عدم قطعیت حاکم بر جو هیچ گونه توجهی نمی‌شود؛ به طوری که اگر با مقادیر اولیه یکسان، مدل عددی را چندین بار اجرا کنیم نتیجه تغییری نخواهد کرد. در حالی که به علت آشوبی بودن سیستم دینامیکی جو، با شرایط اولیه یکسان ممکن است وضعیت هوای آینده متفاوت باشد.

پیش‌بینی احتمالاتی سعی دارد تا با بیانی احتمالی از فضای پیشامد وقایع ممکن آینده، عدم قطعیت را کمی سازد. بدین طریق که به جای در نظر گرفتن تنها یک مقدار اولیه، از یک توزیع احتمال که عدم قطعیت وضعیت اولیه را توصیف می‌کند، نمونه‌گیری می‌شود که در نتیجه آن، پیش‌بینی یک توزیع احتمال خواهد بود که عدم قطعیت وضعیت آینده جو را توصیف می‌کند. پیش‌بینی احتمالاتی

شکل ۴ تابع چگالی احتمال پیش‌بینی دو روزه دمای بیشینه برای ایستگاه مهرآباد در تاریخ دهم مه ۲۰۱۸ را نشان می‌دهد که با استفاده از روش BMA بدست آمده است. تابع چگالی احتمال گفته شده از میانگین وزنی تابع چگالی احتمال یک از اعضای همادی حاصل شده است. تابع چگالی احتمال برازش شده برای هر یک از اعضای همادی دارای توزیع نرمال با میانگین  $24/1$  و  $21/2$  به ترتیب برای عضو اول و دوم و انحراف معیار یکسان  $1/1$  برای هر دو عضو است. وزن برآورد شده برای عضو همادی اول و دوم به ترتیب برابر با  $0/69$  و  $0/31$  است. مقدار دمای بیشینه پیش‌بینی شده عضو اول  $21/4$  است که با برازش تابع چگالی احتمال برای آن، مقدار میانگین  $(24/1)$  به مقدار مشاهداتی  $(23/5)$  نزدیکتر شده است. همچنین مقدار دمای بیشینه پیش‌بینی شده عضو دوم  $19/4$  است که با برازش تابع چگالی احتمال برای آن، مقدار میانگین  $(21/2)$  به مقدار مشاهداتی  $(23/5)$  نزدیکتر شده است. همان طور که در شکل ۴ دیده می‌شود، اگرچه تابع چگالی احتمال هر عضو دارای توزیع نرمال است، اما تابع چگالی احتمال نهایی لزوماً دارای توزیع نرمال نیست.

استفاده می‌توان آن را برای کمیت‌های دیگر مانند بارش یا باد تعمیم داد.

## ۷. مراجع

- 1- Eckel, F., Allen M., Sittel, M., 2012, Estimation of ambiguity in ensemble forecasts, *Weather and Forecasting*, vol. 27, pp. 50-69.
- 2- Glahn, B., Peroutka, M., Wiedenfeld, J., Wagner, J., Zylstra, G., Schuknecht, B., Jackson, B., 2009, MOS uncertainty estimates in an ensemble framework, *Monthly Weather Review*, vol. 137, pp. 246-268.
- 3- Gneiting, T., Raftery, A. E., Westveld, A. H., Goldman, T., 2005, Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation, *Monthly Weather Review*, vol. 133, pp. 1098-1118.
- 4- Hamill, T. M., Colucci, S. J., 1997, Verification of Eta-RSM short-range forecasts, *Monthly Weather Review*, vol. 125, pp. 1312-1327.
- 5- Hamill, T. M., Whitaker, J. S., Wei, X., 2004, Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts, *Monthly Weather Review*, vol. 132, pp. 1434-1447.
- 6- Johnson, C., Bowler, N., 2009, On the reliability and calibration of ensemble forecasts, *Monthly Weather Review*, vol. 137, pp. 1717-1720.
- 7- Lorenz, E. N., 1963, *The essence of chaos*, University of Washington Press.
- 8- Mohamadi, S. A., Rahmani, M., Azadi, M., 2016, Meta-heuristic CRPS minimization for the calibration of short-range probabilistic forecasts, *Meteorology and Atmospheric Physics*, vol. 128, pp. 429-440.
- 9- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, A. B., 1992, *Numerical Recipes in FORTRAN: The Art of Scientific*, Cambridge University Press.
- 10- Raftery, A., Gneiting, T., Balabdaoui, F., Polakowski, A. M., 2005, Using Bayesian model averaging to calibrate forecast ensemble, *Monthly Weather Review*, vol. 133, pp. 1155-1174.
- 11- Roulston, M. S., Smith, L. A., 2003, Combining dynamical and statistical ensembles, *Tellus A*, vol. 55, pp. 16-30.
- 12- Stensrud, D., Skindlov, J., 1996, Gridpoint predictions of high temperature from a mesoscale model, *Weather and Forecasting*, vol. 11, pp. 103
- 13- E., Collins, D., 2009, Ensemble regression, *Monthly Weather Review*, vol. 137, pp. 2365-2379.

به کاربر این امکان را می‌دهد که با آگاهی از میزان عدم قطعیت پیش‌بینی، تصمیم‌گیری بهینه داشته باشد. برای کاربر لازم است تا با در اختیار داشتن تخمینی از احتمالات همه خروجی‌های ممکن، میزان ریسک را بسنجد. بنابراین با ترکیبات مختلفی از شرایط اولیه متفاوت، استفاده از چندین مدل پیش‌بینی عددی و اجرای یک مدل با چندین پیکربندی فیزیکی مختلف، اعضای همادی بدست می‌آیند. با ایجاد سامانه همادی، می‌توان به تابع چگالی احتمال وضع هوای آینده دست یافت؛ به طوری که اعضای همادی نمونه‌هایی از تابع چگالی احتمال عدم قطعیت وضعیت آینده جو می‌باشند. به علت محدودیت در امکانات موجود، یک سامانه همادی دو عضوی با دو پیکربندی فیزیکی مختلف در بازه زمانی از ۱۰ نوامبر ۲۰۱۷ تا ۳۱ مه ۲۰۱۸ تشکیل شد.

در پژوهش حاضر برای برآورد تابع چگالی احتمال پیش-بینی دماهای کمینه و بیشینه روزانه از دو روش EMOS و BMA استفاده شد. هر دو روش از پر اهمیت‌ترین روش‌های موجود می‌باشند. اگرچه مهم‌ترین هدف از بکارگیری روش‌های مذکور، تخمین تابع چگالی احتمال و صدور پیش‌بینی احتمالاتی است، اما میانگین آنها را نیز می‌توان به عنوان یک پیش‌بینی قطعی در نظر گرفت. پس از راستی-آزمایی نتایج به دست آمده، نشان داده شد که هر دو روش مذکور دارای عملکرد یکسانی هستند و نسبت به پیش‌بینی خام دارای خطای بسیار کمتری هستند. همچنین یک نمونه از تابع چگالی احتمال پیش‌بینی دو روزه دمای بیشینه برای ایستگاه مهرآباد در تاریخ دهم مه ۲۰۱۸ با هر دو روش EMOS و BMA تخمین زده شد که به خوبی نحوه عملکرد این روش‌ها را نشان می‌دهد. یک مزیت روش BMA آن است که به سادگی و با تغییر توابع کرنل مورد

- 14- Warner, T. T., 2011, Numerical Weather and Climate Prediction, Cambridge University Press.
- 15- Wilks, D., 2011, Statistical Methods in the Atmospheric Sciences, Academic Press, New York.

## Probabilistic forecasting of diurnal maximum and minimum temperature using a 2-member ensemble system over Iran

Majid Azadi<sup>\*1</sup>, Seyede Atefeh Mohamadi<sup>2</sup>

1- Associate Professor, Atmospheric Science and Meteorological Research Center (ASMERC)

2- Atmospheric Science and Meteorological Research Center (ASMERC)

(Received: 17 Dec 2018, Revised: 04 Mar 2019, Accepted: 08 Jun 2019, Published online: 23 Sep 2019)

### Abstract

To quantify the uncertainty exists in the weather forecasting, and producing the probabilistic forecasting of diurnal maximum and minimum temperature, two important methods named Ensemble Model Output Statistics (EMOS) and Bayesian Model Averaging (BMA) are used to estimate the forecast probabilistic density function. In this study, the ensemble system has two members consisting of 1 to 5 ahead forecasting of diurnal maximum and minimum temperature over synoptic stations of Iran provinces from 10 November 2017 to 31 May 2018. These members are the outputs of the WRF model with two different physical configurations. Results show that the deterministic post-processed forecasts have improved the root mean squared error (RMSE) of deterministic raw forecast, 37% and 7% for maximum and minimum temperature, respectively. Generally, raw forecasts of minimum temperature have less error than maximum temperature, but they are not be improved considerably after post-processing.

**Keywords:** Ensemble system, probabilistic forecasting, Probabilistic density function.

---

\* Corresponding author:

Email Address: azadi68@hotmail.com