

ارزیابی پیش‌بینی‌های کوتاه‌مدت بارش مدل عددی WRF در منطقه ایران

در دوره یک ماهه

فرحناز تقیوی^{۱*}، ابوالفضل نیستانی^۲ و سرمهد قادر^۳

^۱ استادیار، گروه فیزیک فضا، موسسه ژئوفیزیک دانشگاه تهران، ایران

^۲ دانش آموخته کارشناسی ارشد هواشناسی، موسسه ژئوفیزیک دانشگاه تهران، ایران

^۳ دانشیار، گروه فیزیک فضا، موسسه ژئوفیزیک دانشگاه تهران، ایران

(دریافت: ۹۰/۷/۲۶، پذیرش نهایی: ۹۲/۲/۲۴)

چکیده

پیش‌بینی عددی مقدار بارش مانند دیگر کمیت‌های هواشناختی در سطح زمین با خطا همراه است و همواره تلاش می‌شود با به کارگیری روش‌های مناسب، مقدار خطا کاهش یابد. بدین‌منظور نیاز به ارزیابی خروجی مدل‌های عددی بهمنظور سنجش میزان تطابق بین پیش‌بینی‌ها و مشاهدات برای انتخاب بهترین حالت وجود دارد. در این تحقیق، روش‌هایی برای ارزیابی پیش‌بینی‌های بارش کمی معروفی خواهد شد. شیوه ارزیابی رایج در پیش‌بینی بارش، روش‌های دسته‌ای است که اساساً بر پایه جدول‌های توافقی است. مشاهدات X و پیش‌بینی‌های Y به وقایع دو بخشی براساس اینکه آنها بیشتر از یک آستانه آهنگ بارش ثابت α باشند، تبدیل می‌شوند. سپس امتیازات ارزیابی دودویی از جدول‌های توافقی در بعضی آستانه‌های بارشی برآورد می‌شود. نتایج ارزیابی برای پیش‌بینی‌های بارش نشان داد که الگوی کلی نوارهای بارشی را مدل WRF بدروستی روی ناحیه برآورد می‌کند اما در دامنه بارش‌ها مشکلاتی وجود دارد. همچنین امتیازات مهارتی برای آستانه‌های متفاوت روی حوزه مورد تحقیق برای کل بازه یک‌ماهه و روزهای اوچ بارش، بیانگر عملکرد نسبتاً مطلوب مدل در رخداد بارش در بیشتر نواحی است. برای آستانه بارش $\alpha = 0.01$ میلی‌متر مدل در بیشتر ایستگاه‌های همدیدی منتخب تعداد دفعات بارش را به درستی برآورد می‌کند و امتیاز مهارتی مطلوبی دارد، اگرچه هنوز دامنه بارش‌های برآورد شده دارای مشکلاتی است. همچنین امتیازات ارزیابی مدل در برآورد پیش‌بینی‌های بارش ۲۴ ساعته نسبت به پیش‌بینی‌های بارش ۴۸ ساعته بهتر است.

واژه‌های کلیدی: ارزیابی پیش‌بینی، WRF، بارش، جدول توافقی، ایران

Short range precipitation forecasts evaluation of WRF model over IRAN

Taghavi, F.¹, Neyestani, A.² and Ghader, S.³

¹Assistant Professor, Space Physics Department, Institute of Geophysics, University of Tehran, Iran

²M.Sc. Student in Meteorology, Institute of Geophysics, University of Tehran, Iran

³Associate Professor, Space Physics Department, Institute of Geophysics, University of Tehran, Iran

(Received: 18 Oct 2011, Accepted: 14 May 2013)

Summary

Iran has a complex topography and it consists of rugged, mountainous rims surrounding high interior basins. Because of this condition, in some cases the NWP output has a significant error from mesoscale variations induced by the diverse topography.

Iran, covering an area of about 1,648,000 km², is located in the southwest of Asia approximately between 25° and 40° N and 44° and 64° E. This is predominately a semi-arid to arid region surrounded by Caspian Sea to the north and Persian Gulf to the south and crossed by the impressive Zagros and Alborz Mountains. However, for a next-

generation mesoscale forecast model (the Advanced Research Weather Research and Forecasting model, ARW) developed by NCAR (Skamarock et al. 2005), the performance of this model employed in the operational forecasts over Iran is not fully tested.

A few previous model studies (Evans and Smith 2001; Evans et al. 2004; Zaitchik et al. 2007; Marcella and Eltahir 2008; Xu et al. 2009) provided some interesting results for the basic weather simulation in SWA using a regional climate model [the second-generation National Center for Atmospheric Research (NCAR) Regional Climate Model (RegCM2)] or the fifth generation Pennsylvania State University–NCAR Mesoscale Model (MM5) model. They pointed out that the regional model has difficulty producing an accurate simulation of meteorological variables in certain sub regions this includes an accurate description of storm tracks, topographic interactions, and atmospheric stability.

On the other hand, verification is a critical component of the development and use of forecasting systems. Ideally, verification should play a role in monitoring the quality of forecasts, providing feedback to developers and forecasters to help improve forecasts, and provide meaningful information to forecast users to apply in their decision-making processes.

One of the purposes, in this study is to evaluate the performance of the WRF model in the complex terrain of Iran. This evaluation primarily concentrates on the precipitation forecasts. In this paper, Real-time gridded 24h and 48-h precipitation forecasts from NCAR models (the Advanced Research Weather Research and Forecasting model; WRF model) are verified over Iran from one to 28 February 2007. Network has 209 * 194 points, which have the center point 54E longitude and 32N latitude and the horizontal resolution of main domain is 21km. The observed precipitation data were taken from the FNL (Final Operational Global Analysis data) reanalysis data with horizontal resolution $1^\circ \times 1^\circ$.

All forecasts are mapped to a 21km latitude-longitude grid and have been verified against an operational precipitation analysis (more than 100 rain gauges), mapped to the same grid. In this study, first we will describe the forecasting errors of the WRFARW model for precipitation. Then, we will introduce some techniques for evaluating the forecasts. These techniques are particularly designed to examine the difficult features of the precipitation fields.

Forecasting weather variables by numerical models has a discrete nature. Assessment of discrete variables such as precipitation is examined by discrete evaluation techniques and evaluation of quantitative precipitation forecasts (QPF) is a challenging matter because of the noisy, discontinuous and non-normal nature of precipitation.

Common method for evaluation of quantitative precipitation forecast is a categorical procedure that is essentially based on contingency tables. X observations and Y forecasts convert to dichotomous events based on both of them more than a constant precipitation rate threshold ‘ u ’. Then, the behavior of binary categorical verification scores from contingency tables in some rainfall thresholds has been evaluated. Common measures of the binary events evaluation are hit rate, false alarm ratio, false alarm rate, skill scores, correct ratio etc.

Some measures from Signal Detection Theory (SDT) such as area under the ROC curve and discrimination distance is used. SDT offers two broad advantages. Firstly, it provide a means of assessing the performance of a forecasting system that distinguishes between the intrinsic discrimination capacity and the decision threshold of the system. The main analysis tool that accomplishes this is the relative operating characteristic (ROC) that is a graph of hit rate against false alarm rate as u varies, with false alarm rate plotted as the X-axis and hit rate as the Y –axis. Secondly, SDT provides a framework within which other methods of assessing binary forecasting performance can be analyzed

and evaluated.

Due to the complexity of the Iranian plateau and lack of knowledge in the estimation of the physical processes in this area, forecasters should have greater awareness of these limitations of the model when forecasting in this region.

Examining the calculated scores for QPF, show that the WRF model correctly estimates the general pattern of precipitation bands but there are problems in the actual rain value. Moreover, skill scores for different thresholds on total investigated area for one month period and in the intensity activity of synoptic systems days show good performance of the WRF model for estimating precipitation in most areas. In addition, using ROC curves gives a measure of performance in all thresholds. For 0.1mm rainfall threshold at selected synoptic stations, model estimates rainfall frequency properly and skill scores is desirable, although precipitation rate estimates still have problems. In addition, the verification scores of model in estimation of quantitative precipitation in 24 forecasts are better than 48h forecasts. The results suggest that improvements in initialization may be as important, or more so, than improvements in the physics for the land surface processes.

Keywords: Forecast evaluation, WRF, Precipitation, Contingency table, Iran

دارند که فقط وقتی آشکار می‌شوند که طرحواره‌های ارزیابی سطح بالا و پیچیده به کار رود. شناسایی چنین اُریبی‌هایی می‌تواند منجر به تحقیقاتی شود که هدف این تحقیقات ارتقا دانش درباره رخداد آنها است. این بهنوبه‌خود می‌تواند منجر به بهبود درک علمی از فرایندها برای بهبود مدل و سرانجام بهبود پیش‌بینی‌ها شود. کاربرد اجرایی ارزیابی پیش‌بینی قطعاً شامل بحث‌های اقتصادی می‌شود، اما سومین دلیل ارزیابی پیش‌بینی یعنی جنبه اقتصادی، معمولاً بیشتر در ارتباط با کاربران پیش‌بینی‌ها است. طرحواره‌های ارزیابی در این مورد باید به ساده‌ترین حد ممکن فقط در عباراتی از نتایج آنها برای کاربران بیان شوند. پیچیدگی بهدلیل اینکه کاربران متفاوت علائق متفاوتی دارند رخ می‌دهد. از این‌رو نیاز به طرحواره‌های ارزیابی مناسب برای هر کاربر وجود دارد. مدل‌های عددی وضع هوا ابزاری هستند که در گنار سایر امکانات، مانند تصاویر ماهواره‌ای، داده‌های راداری، در ارتقا کیفیت پیش‌بینی‌ها نقش بسزایی دارند. با این حال این مدل‌ها به دلایل متعددی دارای خطاهایی هستند. مقادیر این خطاهای بیویژه در پیش‌بینی کمیت‌های سطح زمین مانند دمای ۲ متری و بارش قابل توجه است. از

۱ مقدمه

ارزیابی پیش‌بینی بخشی ضروری از تحقیقات آب و هوایی و فعالیت‌های عملی پیش‌بینی است. با گسترش کاربرد مدل‌های عددی وضع هوا بر اهمیت ارزیابی پیش‌بینی میدان‌های خروجی مدل‌ها افزوده شده است. برای بیان دلیل ارزیابی پیش‌بینی سه نوع دیدگاه وجود دارد، که بنا به نظر برایر و آلن (1951) می‌توان انها را با عنوانین دیدگاه مدیریتی، علمی و اقتصادی توصیف کرد. به‌طور طبیعی هیچ کدام از دیدگاه‌ها کامل نیست و بین آنها همپوشانی وجود دارد. موضوع مهم این است که با توجه به هدف تحقیق، ابتدا نوع طرحواره‌های ارزیابی برای پاسخگویی به سوالات انتخاب شوند. از نظر اجرایی، داشتن بعضی از مقادیر آستانه برای ارزیابی عملکرد سامانه پیش‌بینی لازم است. برای این منظور، معمولاً یک یا چند شاخص مشخص، می‌تواند عملکرد سامانه پیش‌بینی را نشان دهد. در دیدگاه علمی، بیشترین توجه به درک سازوکار ارزیابی سامانه پیش‌بینی است. در کل ارزیابی جزئی از قدرت و ضعف یک مجموعه از پیش‌بینی‌ها است که معمولاً به بیشتر از یک یا دو امتیاز خلاصه شده نیاز دارد. بعضی از اوقات اُریبی‌های مشکوک در مدل‌های پیش‌بینی وجود

منطقه جنوب غرب آسیا برای کمیت‌های دمای ۲ متری، بارش و باد، در ماه مه ۲۰۰۶ ارزیابی کرده‌اند. آنها نتیجه گرفتند که پیش‌بینی‌های دمای ۲ متری با پیش‌بینی دیگر متغیرهای جوئی متفاوت است و خطاها در پیش‌بینی‌های دمای ۲ متری با تopoگرافی سطح شدیداً مرتبط است. همچنین با به کارگیری داده‌های ماهواره‌ای تابش و بپهود شرایط اولیه، خطاها مدل تا حدی روی بعضی از مناطق مورد بررسی کاهاش می‌یابد. صدوفی و همکاران (۲۰۱۰) به بررسی کارایی مدل ECMWF برای پیش‌بینی بارش به مدت یک سال روی ایران پرداختند. آنها نمره‌های ارزیابی شامل ریشه خطای میانگین مربعات، معیار برخورد و نسبت هشدارهای نادرست و امتیاز مهارتی آمار مهارت‌های درست را به منظور بررسی کیفیت پیش‌بینی‌های بارش مدل در آستانه‌های گوناگون مورد استفاده قرار دادند. نتایج تحقیق آنها روشن ساخت که مدل ECMWF فقط تا حدی موقعیت نوارهای بارشی را به درستی پیش‌بینی می‌کند، درحالی که موقعیت بارش‌های پیش‌بینی با مشاهدات سازگار است. همچنین در برآورد بارش‌های روزانه با استفاده از این مدل، اندکی برآورد اضافی وجود دارد.

در زمینه ارزیابی پیش‌بینی‌های بارش مدل‌های عددی، تحقیقاتی نیز در داخل کشور صورت گرفته است. برای نمونه، آزادی و همکاران (۱۳۸۸) در تحقیقی برونداد مدل WRF برای کمیت بارش روی ایران را برای ماه ژانویه در زمستان ۸، مورد راستی آزمایی قرار داده‌اند. به منظور راستی آزمایی برای دو وضعیت بارش و فقدان بارش و نیز در نظر گرفتن چهار آستانه برای بارش، جدول توافقی را تشکیل دادند و بعضی از امتیازات مهارتی را محاسبه کردند. مزرعه‌فرهانی و همکاران (۱۳۸۸) به بررسی اثر تغییر تفکیک افقی مدل میان‌مقیاس MM5 در شبیه‌سازی بارش ناشی از یک سامانه همدیدی در ایران پرداخته و نتیجه گرفتند که مدل قادر به شبیه‌سازی مناسب میدان

سال ۱۹۵۰ میلادی به بعد، پس از اجرای اولین نمونه از پیش‌بینی‌های وضع هوای عددی، روش‌های زیادی به منظور ارزیابی خروجی مدل‌های عددی معرفی و استفاده شده است. از جمله مورفی و وینکلر (۱۹۸۷)، در تحقیقی جامع، چارچوبی کلی برای ارزیابی پیش‌بینی‌ها براساس توزیع‌های مشترک احتمال بنا نهادند.

با گسترش کاربرد مدل‌های عددی وضع هوای در دهه اخیر، تحقیقات در ارتباط با ارزیابی پیش‌بینی مدل‌ها وارد مرحله تازه‌ای شده است. ناتر و مانوبیانکو (۱۹۹۹) با ارزیابی مدل Eta در سه ایستگاه منتخب به بررسی کیفیت مدل در برآورد پیش‌بینی‌های متغیرهای سطحی و جوئی بالا در دو فصل سرد و گرم پرداختند. نتایج آنها نشان داد که مدل میان‌مقیاس Eta، تکامل میدان بزرگ‌مقیاس را به خوبی پیش‌بینی می‌کند، اما چالش اصلی در پیش‌بینی‌های زمان واقعی، تمایز بین خطای غیر سامان‌مند از خطای سامان‌مند است که باعث افزایش خطای کلی پیش‌بینی می‌شود. کنایویل و همکاران (۲۰۰۴) روشن ساختند که با کمک تحلیل هماهنگ (هارمونیک) می‌توان عملکرد مدل عددی پیش‌بینی وضع هوای را در تحلیل مدهای زمانی بارش شبیه‌سازی شده بررسی کرد. دیویس و همکاران (۲۰۰۶) به ارزیابی عینی پیش‌بینی‌های بارش مدل WRF روی ایالات متحده پرداخته‌اند. در این روش که به منظور تعریف مناطق بارانی برای اهداف ارزیابی بارش پیش‌بینی شده با مدل‌های عددی وضع هوای توسعه یافته، شیوه‌های بارش در هر دو مورد پیش‌بینی‌ها و مشاهدات بر پایه همامیخت (Convolution) و یک روش آستانه‌گذاری تعریف می‌شوند. نتایج آنها نشان داد که WRF روی هم‌رفته خطاها پیش‌بینی بارش مدل اُریبی‌های بزرگی در این منطقه ندارند، اما در ساعت‌بعد از ظهر اُریبی‌ها مثبت می‌شوند و به پیش‌بینی خود می‌رسند. شو و همکاران (۲۰۰۹) پیش‌بینی‌های مدل WRF را با گوارد داده ماهواره‌ای و بدون گوارد داده ماهواره‌ای در

شده C را به کار می‌برد با سامانه پیش‌پردازش مدل (WPS)، درون‌یابی انواع خاک جهانی، انواع زمین مورد استفاده، ارتفاع منطقه، میانگین سالانه دمای عمق خاک، کسر سبزی ماهیانه، سپیدایی (آلیدرو) برف بیشینه و دیگر داده‌های کاربری زمین و همچنین شرایط ابتدایی جوّ و شرایط مرزی جانی در شبکه مدل محاسبه می‌شود. سپس داده‌های خروجی از سامانه پیش‌پردازش در حکم ورودی به مدل اعمال می‌شود و خروجی در ساعات تنظیم شده در قالب دودویی تولید می‌شود که با نرم افزارهای پس‌پردازش قابل مشاهده و تحلیل است.

۳ داده‌ها و منطقه مورد بررسی

برای اجرای این تحقیق، شرایط ابتدایی جوّ و شرایط مرزی جانی شامل دما و رطوبت خاک از داده‌های تحلیل FNL (Final Operational Global Analysis) شده با فاصله‌های زمانی 6° ساعته و تفکیک افقی $1^{\circ} \times 1^{\circ}$ گرفته شده است. با استفاده از سامانه پیش‌پردازش مدل، انواع خاک جهانی، انواع زمین مورد استفاده، ارتفاع منطقه، میانگین سالانه دمای عمق خاک، کسر سبزی ماهیانه، سپیدایی برف بیشینه و مانند آن به شبکه‌های مدل در حوزه مورد بررسی درون‌یابی می‌شود. داده‌های جغرافیایی به کار رفته برای حوزه اصلی مدل داده‌های ۱۰ دقیقه‌ای و برای حوزه داخلی داده‌های ۵ دقیقه‌ای است. بسته‌های فیزیکی انتخاب شده شامل طرحواره میکروفیزیکی WSM3، طرحواره تابش طول موج بلند RRTM، طرحواره تابش طول موج کوتاه Dudhia، طرحواره لایه مرزی YSUPBL، طرحواره سطح زمین Noah و پارامترسازی کومولوس Kain-Fritsch هستند. حوزه شیوه‌سازی مدل برای ناحیه اصلی و ناحیه داخلی در شکل ۱ نشان داده شده است. فاصله نقاط شبکه برای ناحیه اصلی ۲۱ کیلومتر و برای ناحیه داخلی ۷ کیلومتر در نظر گرفته شده است. مرکز ناحیه در نظر گرفته شده در

بارش است و با کاهش فاصله شبکه‌ای میدان بارش دقت کیفی قابل قبولی پیدا می‌کند. همچنین هدایتی دزفولی و آزادی (۱۳۸۹) به بررسی راستی آزمایی پیش‌بینی بارش مدل منطقه‌ای MM5 در فصل سرد سال روی ایران پرداخته و با بدست آوردن پیش‌بینی‌های ۲۴، ۴۸ و ۷۲ ساعته برای بارندگی تجمعی ۲۴ ساعته، مقادیر شبیه‌سازی شده را با مقادیر دیده‌بانی شده مقایسه کردند.

در این تحقیق به ارزیابی پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش کمی مدل عددی WRF-ARW در مناطق گوناگون ایران در دوره یک‌ماهه (فوریه ۲۰۰۷) پرداخته می‌شود. شیوه معمول در ارزیابی پیش‌بینی بارش‌های کمی روش‌های دودویی است که براساس جدول‌های توافقی است. با محاسبه امتیازات مهارتی و بعضی کمیت‌ها از این جدول‌ها در آستانه‌های متفاوت بارشی، کیفیت پیش‌بینی‌های مدل از دیدگاه‌های متفاوت مورد بررسی قرار می‌گیرد. همچنین از نظریه آشکارسازی سیگنال (SDT) و منحنی مشخصه عملکرد نسبی (Relative Operator Characteristic Curve) که دارای مزایایی نسبت به روش‌های دودویی معمول هستند نیز در این تحقیق استفاده خواهد شد.

۲ مدل WRF-ARW

مدل مورد استفاده در این تحقیق، سامانه مدل‌سازی عددی تحقیقاتی پیشرفته WRF است که با هسته دینامیکی پیشرفته ARW استفاده می‌شود (WRF-ARW: میشالاک و همکاران، ۲۰۰۱؛ اسکاماروک و همکاران، ۲۰۰۵) و در چند سال گذشته توسعه یافته است. نسخه جاری این سامانه، نسخه ۳ است که از آوریل ۲۰۰۸ موجود است. ARW مناسب برای استفاده در دامنه وسیعی از کاربردها است که مقیاس آنها از چند متر تا هزاران کیلومتر را دربر می‌گیرد. این مدل کاملاً تراکم‌پذیر، و غیرهیدروروماتیک (با انتخاب هیدروروماتیک) است و شبکه آرکاوا جابه‌جا

نقاط ایستگاه‌های همدیدی باران‌سنجدی منتخب درون‌یابی خطی شده است. به علت زیاد بودن تفکیک افقی، اجرای مدل (۷ کیلومتر)، این روش درون‌یابی خطی قابل ملاحظه‌ای ایجاد نخواهد کرد.

۴ روش تحقیق

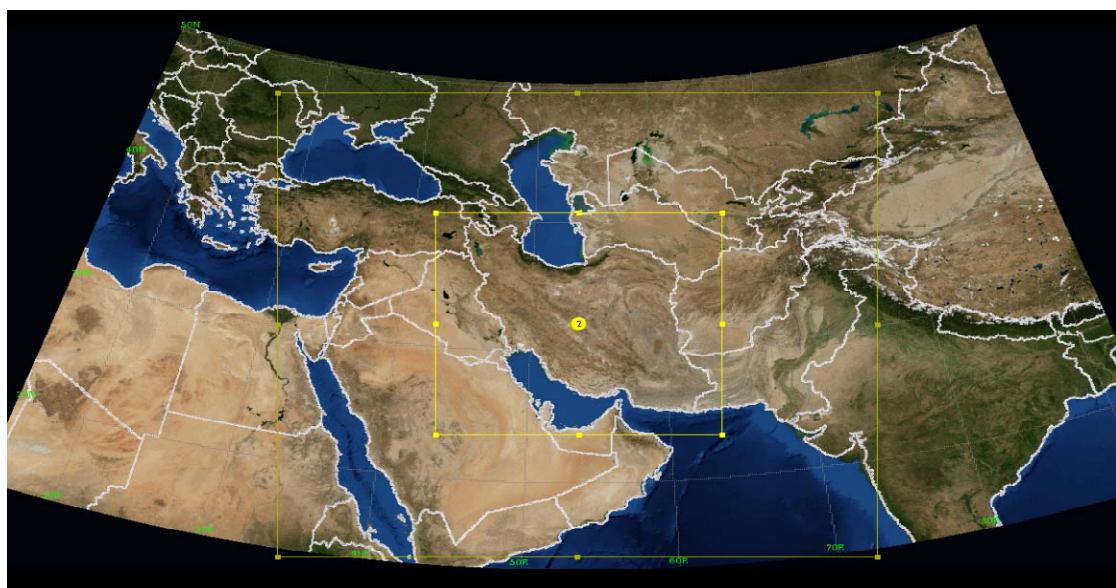
بارش‌های کمی به طور معمول با به کارگیری روش‌های ارزیابی پیوسته و دسته‌ای یا روش‌های تحقیقی، ارزیابی می‌شوند. روش‌های ارزیابی تحقیقی بر پایه مقایسه آماره‌های مشاهدات و پیش‌بینی‌ها است. ارزیابی میانگین‌های پیش‌بینی و مشاهده یکی از روش‌های ارزیابی تحقیقی است که به طور گسترده‌ای از آن استفاده می‌شود (ایری و هلم، ۱۹۹۵؛ مladck و همکاران، ۲۰۰۰).

روش‌های ارزیابی پیوسته بر پایه مقایسه نقطه‌به‌نقطه مقادیر پیش‌بینی و مشاهده است (ویلسون، ۲۰۰۰؛ ایری و هلم، ۱۹۹۵؛ اسپرن و هلم، ۱۹۹۸؛ مladck و همکاران، ۲۰۰۰).

طول جغرافیایی $53^{\circ}5'$ و عرض جغرافیایی 32° درجه قرار دارد. همچنین شبکه اصلی دارای 194×209 نقطه و شبکه داخلی دارای 276×300 نقطه است. با توجه به فاصله شبکه‌ای حوزه اصلی، گام زمانی انتگرال گیری 90 ثانیه در نظر گرفته شده است. اجرای مدل از ساعت ۰۰UTC تا ۰۰UTC ۴۸ ساعت بعد در یک دوره زمانی یک‌ماهه برای فوریه ۲۰۰۷ اجرا شده است. دلیل انتخاب این بازه یک‌ماهه فعالیت نسبتاً مناسب سامانه‌های همدیدی در مدت زمان مورد بررسی است.

داده‌های مشاهداتی بارش کمی نیز از اندازه‌گیری‌های روزانه 30 باران‌سنجد در مراکز استان‌ها و همچنین داده‌های ماهانه بیش از 100 ایستگاه باران‌سنجدی داخل منطقه ایران در بازه زمانی موردنظر از داده‌های کنترل کیفی شده سازمان هوواشناسی کشور به دست آمد. این باران‌سنجد به استثنای نواحی جنوب شرقی ایران در بقیه نواحی دارای حداقل توزیع مناسب در منطقه مورد بررسی‌اند.

داده‌های پیش‌بینی شده بارش کمی 24 و 48 ساعته خروجی مدل روی حوزه داخلی، به منظور ارزیابی روی



شکل ۱. نمایش حوزه اصلی و حوزه داخلی که شبیه‌سازی‌های مدل در این ناحیه انجام شده است. نسبت حوزه داخلی به حوزه والد ۱ به 3 است.

امتیازات دسته‌ای و امتیازات مهارتی در چارچوبی براساس احتمالات بیان می‌شود.

بسیاری از پدیده‌های آب‌وهوایی را می‌توان در حکم پیشامدهای دودویی ساده در نظر گرفت و پیش‌بینی و هشدار در مورد این پیشامدها اغلب به صورت گزارش‌های کامل که این پدیده‌ها رخ می‌دهند یا نمی‌دهند صادر می‌شود. بارش، سیل، توفان‌های شدید، مه و مانند آن یا رخ می‌دهند و یا رخ نمی‌دهند و پیش‌بینی‌ها یا اخطارهای مناسب یا صادر می‌شود یا صادر نمی‌شود. این نوع از پیش‌بینی‌ها گاهی اوقات با عنوان پیش‌بینی‌های بله/خبر شناخته می‌شوند. خروجی‌های امکان‌پذیر برای جدول (2×2) (پیشامدها) برای هر پیشامد در جدول ۲ نشان داده شده است. دو راه برای درست بودن پیش‌بینی وجود دارد: یا برخورد (Hit) و یا رد درست (Correct Rejection) و دو حالت هم برای نادرست بودن پیش‌بینی‌ها وجود دارد: یا هشدار نادرست (False Alarm) و یا به خط رفتن (Miss).

لازم به ذکر است که فقط یک مجموعه پیش‌بینی‌های دودویی مبنای مناسبی برای ارزیابی کیفیت سامانه پیش‌بینی فراهم نمی‌کنند، زیرا این حالت فقط کیفیت سامانه پیش‌بینی را در یک آستانه انتخابی معجزاً به ما می‌دهد. معیار برخورد و معیار هشدار نادرست را می‌توان برای بدست آوردن منحنی مشخصه عملکرد نسبی به کار برد (منحنی ROC؛ اسوتس، ۱۹۸۶). منحنی ROC از نظریه آشکارسازی سیگنال به این حوزه وارد شده است (میسن، ۱۹۸۲). این منحنی امکان یک ارزیابی امتیاز کلی را که وابسته به همه آستانه‌های تصمیم متفاوت است فراهم می‌کند.

در این بخش، بعضی از آماره‌های توصیفی اساسی و مقادیر کارایی در ارزیابی به روش دودویی براساس متغیرهای معرفی شده در جدول ۱ توصیف می‌شوند:

ضریب همبستگی و آماره‌های مشتق شده از آن (برای نمونه: ضرایب همبستگی بی‌亨جاری؛ پتس و همکاران، ۱۹۹۶) ارتباط خطی بین مشاهده و پیش‌بینی را اندازه می‌گیرد. امتیازات ارزیابی پیوسته، در مورد خاص MSE و RMSE، می‌تواند بیش از حد حساس به نوافه، شب‌های تند و مقادیر بزرگ داده‌های عددی دور از بقیه داده‌ها باشد. عامل ریشه میانگین مربعات (RMSE) که گلدنگ (۱۹۹۸) آن را معرفی کرد، بهتر با این ویژگی‌های میدان بارش مقابله می‌کند. امتیازات ارزیابی پیوسته، وابستگی فضایی در میدان‌های بارش که ناشی از حضور ساختارهای فضایی منسجم (ویژگی‌های بارش) است را شرح نمی‌دهند. علاوه بر این، آنها برخورد مطلوب با ناپیوستگی‌های فضایی تیز بین مناطق بارانی و غیر بارانی که میدان بارش را مشخص می‌کند، ندارند. آنها خطای بین مقادیر جفت شده در موقعیت یکسان را برآورد می‌کنند، اما این اندازه‌ها بیش از حد حساس به جابه‌جایی ویژگی‌های بارش هستند (کستی و همکاران، ۲۰۰۴).

روش‌های دسته‌ای به طور گسترده‌ای در ارزیابی مقادیر کمی بارش به کار می‌روند (السن و همکاران، ۱۹۹۵؛ مسینجر، ۱۹۹۶؛ جانسون و السن، ۱۹۹۸؛ ویلسن، ۲۰۰۰؛ بوگولت، ۲۰۰۳؛ صدوی و همکاران، ۲۰۱۰). ارزیابی دسته‌ای بارش‌های کمی، اساساً بر پایه جدول‌های توافقی و ساده ترین حالت آن ارزیابی دودویی است. مشاهدات X و پیش‌بینی‌های Y به وقایع دو بخشی (دودویی) براساس اینکه آنها بیشتر از یک آستانه آهنگ بارش ثابت α باشند، تبدیل می‌شوند. جدول‌های توافقی برای هر آستانه همان‌طور که در جدول ۱ نشان داده شده، ارزیابی می‌شود. سپس امتیازات دسته‌ای و امتیازات مهارتی از همه جدول‌های توافقی برآورد می‌شوند. توجه شود که جدول‌های توافقی دودویی به آسانی در عباراتی از توزیع‌های مشترک مشاهدات-پیش‌بینی‌ها تفسیر می‌شود (مورفی و وینکلر، ۱۹۸۷).

جدول ۱. پیشامدهای دودویی ممکن و تعریف آنها براساس توزیع‌های احتمال مشترک و حاشیه‌ای.

مشاهدات			پیش‌بینی
مجموع	خیر	بله	
a+b	b	a	بله
c+d	d	c	خیر
a+b+c+d=n	b+d	a+c	مجموع

مشاهدات			پیش‌بینی
مجموع	خیر	بله	
p(x̂=1)	p(x̂=1, x=0)	p(x̂=1, x=1)	بله
p(x̂=0)	p(x̂=0, x=0)	p(x̂=0, x=1)	خیر
1	P(x=0)	P(x=1)	مجموع

جدول ۲. خروجی‌های امکان‌پذیر برای جدول توافقی.

مشاهدات		پیش‌بینی
خیر	بله	
هشدار نادرست	برخورد	بله
رد درست	به خطأ رفت	خیر

احتمال پیشامد (آهنگ پایه)

اُریبی بسامدی
اُریبی بسامدی، نسبت تعداد پیش‌بینی‌های رخداده به تعداد رخدادهای مشاهدات واقعی است (دونالسون و همکاران، ۱۹۷۵):

$$B = \frac{a+b}{a+c} \quad (3)$$

معیار برخورد
این مقدار به صورت رابطه (۴) تعریف می‌شود:

$$H = POD = \frac{a}{a+c} = \hat{p}(\hat{x}=1|x=1) \quad (4)$$

معیار برخورد بخشی از رخدادها است که به درستی پیش‌بینی می‌شود. به عبارت دیگر، این معیار یک برآورد ساده از احتمال شرطی پیشامد پیش‌بینی شده است به شرط اینکه پیشامد مشاهده شود. همچنین نام دیگر این معیار، احتمال آشکارسازی (POD) است (دونالدسون و همکاران، ۱۹۷۵).

$$s = \frac{a+c}{n} = \hat{p}(x=1) \quad (1)$$

احتمال پیشامد، یک برآورد نمونه از احتمال حاشیه‌ای غیرشرطی از رخداد پیشامد مشاهده شده است که صرفاً یک ویژگی از مشاهدات است تا سامانه پیش‌بینی و به طور صریحی نباید ارتباط مستقیمی به ارزیابی مهارت پیش‌بینی داشته باشد، زیرا سامانه پیش‌بینی، کنترلی روی آهنگ رخداد پیشامدهای مشاهده شده ندارد.

احتمال رخداد یک پیش‌بینی

$$r = \frac{a+b}{n} = \hat{p}(\hat{x}=1) \quad (2)$$

بسامد حاشیه‌ای، یه برآورد نمونه از احتمال حاشیه‌ای مربوط به پیش‌بینی یک رخداد است.

به صورت رابطه (۸) تعریف می‌شود:

$$TSS = PSS = \frac{ad - bc}{(a + c)(b + d)} \quad (8)$$

در مقادیر بی‌مهارت، مقدار PSS برابر صفر است. این امتیاز مهارتی مستقل از رژیم‌های خشک و تر است (ایبرت و مک‌برايد، ۱۹۹۷).

شاخص موفقیت بحرانی (امتیاز خطر)

این اندازه را اولین بار گیلبرت در ۱۸۸۴ پیشنهاد کرد و پالمر و آلن (۱۹۴۹) آن را امتیاز خطر (TS) و دونالدسون و همکاران (۱۹۷۵) آن را شاخص موفقیت بحرانی (CSI) نامیدند. این شاخص به صورت رابطه (۹) تعریف می‌شود:

$$CSI = \frac{a}{a+b+c} \quad (9)$$

به طور گسترده‌ای برای ارزیابی پیش‌بینی وقایع نادر استفاده می‌شود. در مهارت کامل مقدار CSI برابر یک است و مقدار کمینه آن در پیش‌بینی‌های بی‌مهارت برابر صفر است.

نسبت شانس

نسبت شانس به صورت رابطه (۱۰) تعریف می‌شود:

$$\theta = \frac{ad}{bc} \quad (10)$$

برای مثال یک نسبت شانس ۲ به این معنی است که اگر رخداد پیش‌بینی مثبت باشد، احتمال رخداد مشاهده مثبت نسبت به مشاهده منفی دو برابر است.

مقادیر ارزیابی روی محورهای منحنی مشخصه عملکرد (ROC)

هر اندازه ارزیابی که در عبارت‌هایی از عضوهای جدول توافقی دودویی تعریف شده است، را می‌توان در عبارت‌هایی از عامل‌های معیار پایه احتمالی H (معیار برخورد)، F (معیار هشدار نادرست) و S (آهنگ پایه) بیان

معیار هشدار نادرست

معیار هشدار نادرست به صورت رابطه (۵) تعریف می‌شود:

$$F = POFD = \frac{b}{b+d} = \hat{p}(\hat{x} = 1 | x = 0) \quad (5)$$

معیار هشدار نادرست کسری از فقدان رخداد است که به طور نادرستی پیش‌بینی می‌شود. بعضی اوقات به این معیار، احتمال آشکارسازی اشتباه (POFD) می‌گویند.

نسبت درست

نسبت درست به صورت رابطه (۶) تعریف می‌شود:

$$PC = \frac{a+d}{n} = \hat{p}[(\hat{x} = 1, x = 1) \text{ or } (\hat{x} = 0, x = 0)] \quad (6)$$

که در واقع توزیع احتمال مشترک مشاهدات و پیش‌بینی‌ها است در صورتی که هر دو رخ دهنند یا هر دو رخ ندهند.

نسبت هشدارهای نادرست

نسبت هشدار نادرست به صورت رابطه (۷) تعریف می‌شود:

$$FAR = \frac{b}{a+b} = \hat{p}(x = 0 | \hat{x} = 1) \quad (7)$$

این نسبت برآورده نمونه از احتمالات شرطی پیش‌بینی‌های رخداد با توجه به اینکه رخداد پیش‌بینی شده، است. به عبارت دیگر بخشی از پیش‌بینی‌های رخداد است که با فقدان رخدادها دنبال می‌شود (پیش‌بینی شده اما رخ نداده) و باید بین این نسبت و معیار هشدار نادرست تفکیک قائل شد.

امتیاز مهارتی پیرس

این امتیاز برای جدول‌های توافقی چندطبقه را هنسن و کوپرس (۱۹۶۵) کشف کردند. همچنین این امتیاز را فلوک (۱۹۸۷) به طور مستقل دوباره کشف کرد که آمار مهارت درست (TSS) نامیده شد. امتیاز مهارتی پیرس

(2×2) با فقط یک نقطه مجزا روی منحنی ROC شامل فقط یک آستانه در نظر گرفته شده) این برآورد با $PSS(1+5/0)$ معادل است که PSS امتیاز ROCهای تجربی روی مهارتی پرس است. ROCهای تجربی روی محورهای احتمال به طور کلی محدب هستند (شکل ۲). بنابراین A_Z یک برآورد پارامتریک از عملکرد سامانه پیش‌بینی فراهم می‌کند.

کرد. از آنجاکه ROC یک نقشه H بر حسب F است، عامل‌بندی معیار پایه احتمالی نشان می‌دهد که اندازه‌های ارزیابی را می‌توان روی محورهای ROC رسم کرد و هم‌مقدارها را می‌توان با شکلی از ROCهای تجربی مقایسه کرد. این معیار از نظریه آشکارسازی سیگنال که در الکترونیک و روانشناسی کاربرد زیادی دارد به حوزه ارزیابی پیش‌بینی وضع هوا وارد شده است.

فاصله تمایز، d'

فاصله d' در ابتدا به متزله اندازه‌ای از نظریه آشکارسازی سیگنال برای سامانه‌هایی که واریانس‌های معادلی دارند توسط تائز و بردسل (۱۹۵۸) پیشنهاد شد. این فاصله با اختلاف بین انحراف‌های نرمال استاندارد مطابق با $F-1$ و $H-1$ برآورد می‌شود:

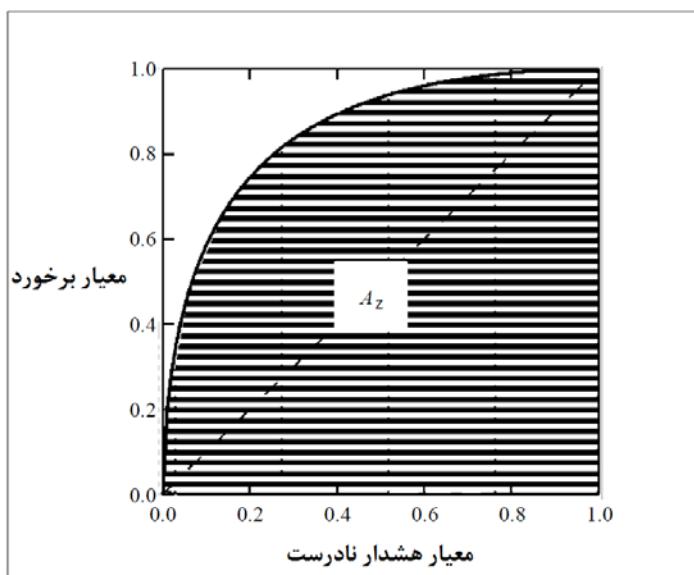
$$d' = \Phi^{-1}(1-F) - \Phi^{-1}(1-H) \quad (11)$$

مقادیر امکان‌پذیر برای d' ($-\infty$ ، $+00$) است، اما مقدار آن در پیش‌بینی آب‌وهوا به طور کلی از صفر تا چهار است. مهارت صفر با $d' = 0$ نشان داده می‌شود، در حالی که $d' > 3$ سطح خیلی بالایی از مهارت هست که اغلب در پیش‌بینی‌های وضع هوای عملی با این سطح مهارت کمتر برخورد می‌کنیم. فاصله d' مقدار بی‌ Mehارت صفر را برای حالت $H=F$ مطابق با پیش‌بینی‌های تصادفی به خود می‌گیرد، اما چون Φ^{-1} برای آرگومان‌های 0 یا 1 تعريف نشده است، وقتی که هردوی H یا F ، یا 1 هستند، d' تعريف نشده خواهد بود. شکل ۳، منحنی‌های ROC نظری تولید شده با مدل‌های آشکارسازی سیگنال با توزیع‌های گاوسی واریانس معادل و مقادیر متفاوت فاصله تمایز را نشان می‌دهد.

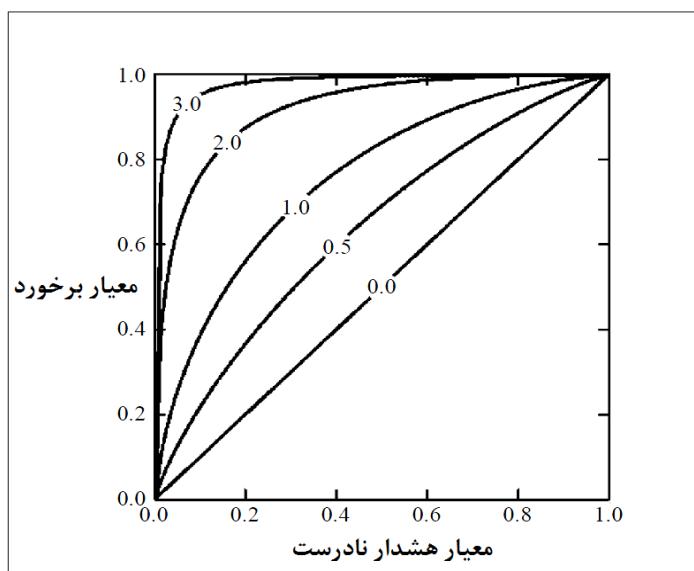
مساحت زیر منحنی ROC مدل شده، A_Z

اندازه مهارتی برپایه منحنی ROC که به طور خیلی وسیعی استفاده می‌شود، A_Z است که مساحت زیر منحنی ROC مدل شده است (اسوتس، ۱۹۸۸). شکل ۲ مساحت زیر منحنی ROC بر پایه مدل آشکارسازی سیگنال و توزیع‌های گاوسی را نشان می‌دهد. محدوده ممکن از A_Z [۰، ۱] است. مهارت صفر با $A_Z = 0/5$ نشان داده می‌شود، وقتی که ROC در امتداد قطر مثبت قرار می‌گیرد، این حالت نشان‌دهنده این است که پیش‌بینی کاملاً تصادفی است. همچنین مقدار A_Z در مهارت کامل برابر یک است، در این حالت منحنی ROC از (۰، ۰) تا (۱، ۰) بالا می‌آید و سپس به (۱، ۰) می‌رود. مقدار A_Z زیر قطر است و نشان‌دهنده سطح یکسانی از ظرفیت تمایز و مهارت است که فقط در امتداد قطر منعکس شده (منفی شده) است.

همچنین سطح زیر منحنی ROC به سادگی می‌تواند با مرتب کردن نقاط داده و گوشش‌های (۰، ۰) و (۱، ۰) روی دیاگرام ROC با خطوط مستقیم و سپس جمع کردن ناحیه ذوزنقه‌های نتیجه‌گیری شده برآورد شود (اسوتس، ۱۹۸۶؛ ویلسون، ۲۰۰۰). در مورد جدول توافقی



شکل ۲. مساحت زیر منحنی ROC (A_z) بر پایه مدل آشکارسازی سیگنال بر پایه توزیع‌های گاوسی.



شکل ۳. ROC‌های نظری تولید شده با مدل‌های آشکارسازی سیگنال با توزیع‌های گاوسی واریانس معادل. برچسب‌ها مقادیر d' هستند.

شکل ۴ پریندهای مجموع بارش تجمعی مشاهده شده و شبیه‌سازی شده برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته برای ایستگاه‌های همدیدی منتخب روی منطقه ایران در فوریه ۲۰۰۷ را نشان می‌دهد. مناطق پُربارش و کم‌بارش به خوبی در این شکل ۴-الف مشخص است. مناطق غربی، رشته کوه زاگرس، رشته کوه البرز، سواحل دریای خزر و

۵ تجزیه و تحلیل نتایج در این بخش پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش کمی روزانه که با اجرای مدل عددی WRF-ARW روی ناحیه پیچیده ایران ایجاد می‌شود ارزیابی خواهد شد. نتایج به دست آمده مربوط به دوره یک‌ماهه از ۱ تا ۲۸ فوریه ۲۰۰۷ است.

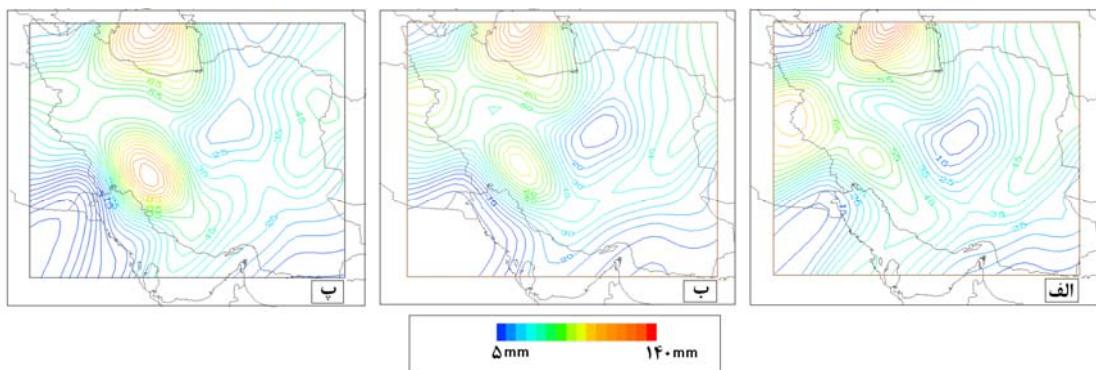
بارش روی منطقه به وقوع پیوسته و برآورده مدل از میزان بارش کمتر از مقدار مشاهده شده است در روزهایی با فعالیت بارشی ضعیفتر معمولاً برآورده مدل از مقدار بارش به خصوص برای پیش‌بینی‌های ۴۸ ساعته بیشتر بوده است.

با توجه به طبیعت گستره میدان بارش، استفاده از مقادیر کلی عملکرد مانند خطای میانگین مربعات که بیشتر برای متغیرهای پیوسته مانند دما به کار می‌رود، مطلوب نخواهد بود. در این موارد، بهتر است از روش‌های ارزیابی دودویی با استفاده از جدول‌های توافقی و شاخص‌های مربوط به آن برای آستانه‌های متفاوت استفاده کرد و یا روش‌های جدید که به تازگی در این حوزه معرفی شده‌اند را استفاده کرد. در این قسمت بارش را در حکم یک متغیر گستره در نظر می‌گیریم و از روش‌های دودویی به منظور تجزیه و تحلیل داده‌های بارش مشاهداتی و شبیه‌سازی شده مدل استفاده می‌شود. شاخص‌های این جدول‌های اعم از معیار برخورد (H)، معیار هشدار نادرست (F)، امتیازات مهارتی، مساحت زیر نمودارهای ROC و مانند آن برای آستانه‌های متفاوت و برای مناطق گوناگون به صورت زمانی و فضایی محاسبه و مقایسه‌های لازم صورت گرفته است.

نمودار پیشامد (جدول توافقی) برای کل ایستگاه‌های تحت بررسی برای آستانه‌های متفاوت بارش و پیش‌بینی-های ۲۴ و ۴۸ ساعته در شکل ۶ نشان داده شده است. با افزایش آستانه بارش، تعداد پیش‌بینی‌های فقدان بارش که با مشاهدات فقدان بارش منطبق می‌شوند افزایش می‌یابد (متغیر d ، در حالی که تعداد بارش‌هایی که پیش‌بینی شده‌اند و به وقوع هم پیوسته‌اند، کاهش می‌یابد (متغیر a)، زیرا آهنگ پایه اقلیمی (S) برای بارش‌های سنگین در حوزه مورد بررسی کاهش می‌یابد.

تا حدودی شمال شرق و شرق ایران از بارندگی مناسبی در این ماه برخوردار بوده‌اند و در نواحی مرکز، جنوب شرق، شمال غرب و جنوب غرب ایران مقدار بارش کم بوده است. شکل ۴-ب که پربندهای پیش‌بینی‌های ۲۴ ساعته را برای مقادیر بارش نشان می‌دهد، به خوبی مناطق پربارش و کم‌بارش را شبیه‌سازی کرده است. اما مقادیر بارش شبیه‌سازی شده گرچه برای بسیاری از نواحی به مقدار واقعی نزدیک است اما تفاوت‌هایی در این مقادیر وجود دارد. برای مثال گرادیان و مقدار بارش در نواحی جنوبی زاگرس را مدل بیشتر برآورده است و در شمال شرق و شرق و همچنین غرب ایران، مدل بارش را مقداری کمتر از مقدار مشاهده شده برآورده است. در شکل ۴-پ نیز که مربوط به پیش‌بینی ۴۸ ساعته مقدار بارش است، الگوی کلی به الگوی مشاهدات نزدیک است ولی مقادیر بارش نسبت به پیش‌بینی ۲۴ ساعته برای بسیاری از مناطق خطای بیشتری را نشان می‌دهد. در مرکز و شمال غرب ایران مقدار بارش زیادتر و در غرب کمتر برآورده شده و گرادیان و مقدار بارش را زاگرس جنوبی نسبت به پیش‌بینی ۲۴ ساعته افزایش پیدا کرده است.

به منظور مقایسه رفتار بارش شبیه‌سازی شده و مشاهده شده و مقدار بارش در حوزه مورد بررسی سری زمانی مجموع بارش روزانه مشاهده شده و شبیه‌سازی شده برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته در شکل ۵ نشان داده شده است. خروجی بارش تجمعی مدل برای پیش‌بینی ۲۴ و ۴۸ ساعته از لحظه رفتار تطابق خوبی با سری زمانی بارش مشاهده شده دارد، به طوری که مدل به خوبی روزهای فعالیت سامانه‌های جویی و روزهای فقدان فعالیت آنها را آشکار می‌کند. فقط برای روزهای ۷ و ۸ فوریه مقداری ناهمانگی وجود دارد. در ۴ و ۸ فوریه که بیشترین میزان



شکل ۴. (الف) پربندهای مجموع بارش تجمعی روی ایران در فوریه ۲۰۰۷. در رسم پربندها از اطلاعات بارش تجمعی بیش از ۱۰۰ ایستگاه همدیدی در سطح ایران استفاده شده است. (ب) پربندهای مجموع بارش شبیه‌سازی شده برای ایستگاه‌های همدیدی استفاده شده در قسمت (الف) برای پیش‌بینی‌های ۲۴ ساعته در همان دوره زمانی (ب). پربندهای مجموع بارش شبیه‌سازی شده برای ایستگاه‌های همدیدی استفاده شده در قسمت (الف) برای پیش‌بینی‌های ۴۸ ساعته در همان دوره زمانی.

آنگه پایه در پیش‌بینی‌های ۲۴ ساعته بارش نسبت به پیش‌بینی‌های ۴۸ ساعته بیشتر است زیرا تعداد رخدادهای بارش که پیش‌بینی نشده ولی اتفاق افتاده در پیش‌بینی ۲۴ ساعته برای همه آستانه‌ها بیشتر بوده است.

بسامد حاشیه‌ای، یک برآورد نمونه از احتمال حاشیه‌ای یک پیش‌بینی رخداد است. یعنی احتمال اینکه سامانه پیش‌بینی بارش را برای آن آستانه پیش‌بینی کند چه بارش مشاهده بشود و چه نشود. این عدد معروف یک ویژگی از پیش‌بینی‌های سامانه پیش‌بینی است. با توجه به جدول‌ها، احتمال رخداد یک پیش‌بینی بارش در پیش‌بینی‌های ۴۸ ساعته برای آستانه‌های بارش در نظر گرفته شده، اندکی بیشتر است.

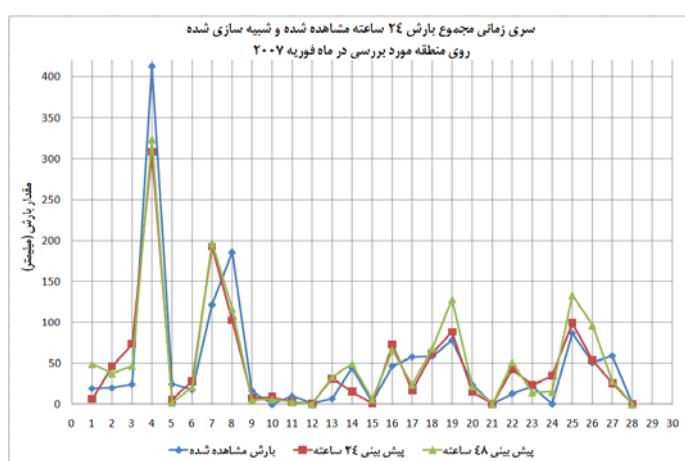
أُربی بسامدی که نسبت تعداد رخداد پیش‌بینی‌ها به تعداد رخدادهای مشاهدات واقعی است، در حالت ایدئال باید مقدار آن برابر ۱ باشد. در این حالت تعداد رخداد پیش‌بینی‌ها و مشاهدات برابر است. حال اگر این مقدار بیشتر از یک باشد، نشان دهنده آن است که مدل تعداد دفعات بارش را بیشتر از تعداد دفعات مشاهده شده بارش برآورد می‌کند و اگر کمتر از یک باشد، یعنی مدل تعداد دفعات بارش را کمتر از تعداد دفعات مشاهده شده بارش

متغیر b نیز با افزایش آستانه روندی کاهشی را خواهد داشت، در حالی که متغیر c در بعضی آستانه‌ها ممکن است افزایش یا کاهش یابد اما با افزایش آستانه بارش روند کلی کاهشی است. حالت ایدئال زمانی است که متغیرهای b و c صفر باشند، یعنی نه پیش‌بینی بارش و مشاهده فقدان بارش و نه پیش‌بینی فقدان بارش و مشاهده بارش در آن آستانه خاص را داشته باشیم و همچنین در این حالت a دارای بیشترین مقدار باشد. در پیش‌بینی بارش ۴۸ ساعته (۲۴ ساعت دوم) تعداد بارش‌هایی که پیش‌بینی شده ولی اتفاق نیفتاده (b)، نسبت به پیش‌بینی‌های بارش ۲۴ ساعته برای همه آستانه‌ها بیشتر است. بر عکس این حالت برای حالتی که بارش پیش‌بینی نشده ولی اتفاق افتاده (c)، برای همه آستانه‌ها تعداد دفعات رخداد در پیش‌بینی ۲۴ ساعته نسبت به پیش‌بینی ۴۸ ساعته بیشتر است.

پس از محاسبه مقادیر جدول پیش‌بینی دودویی برای آستانه‌های متفاوت، شاخص‌ها و امتیازات مهارتی برای ارزیابی عملکرد مدل روی کل حوزه فضایی مورد بررسی در دوره یک‌ماهه فوریه ۲۰۰۷ برای چند آستانه بارش منتخب محاسبه شده است. نتایج برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته به ترتیب در جداول‌های ۳ و ۴ خلاصه شده است.

آستانه ۱ میلی‌متر، تعداد دفعات پیش‌بینی رخداد بارش مدل با تعداد دفعات رخداد مشاهده بارش تقریباً برابر است ولی برای آستانه‌های بزرگ، مانند ۲۰ میلی‌متر، تعداد دفعات پیش‌بینی رخداد بارش از تعداد دفعات رخداد مشاهده بارش کمتر است. برای پیش‌بینی ۴۸ ساعته بارش این نظم دیده نمی‌شود و برای همه آستانه‌های بارش تعداد دفعات رخداد پیش‌بینی بارش حدود ۳۰٪ بیشتر از تعداد دفعات رخداد مشاهده بارش است.

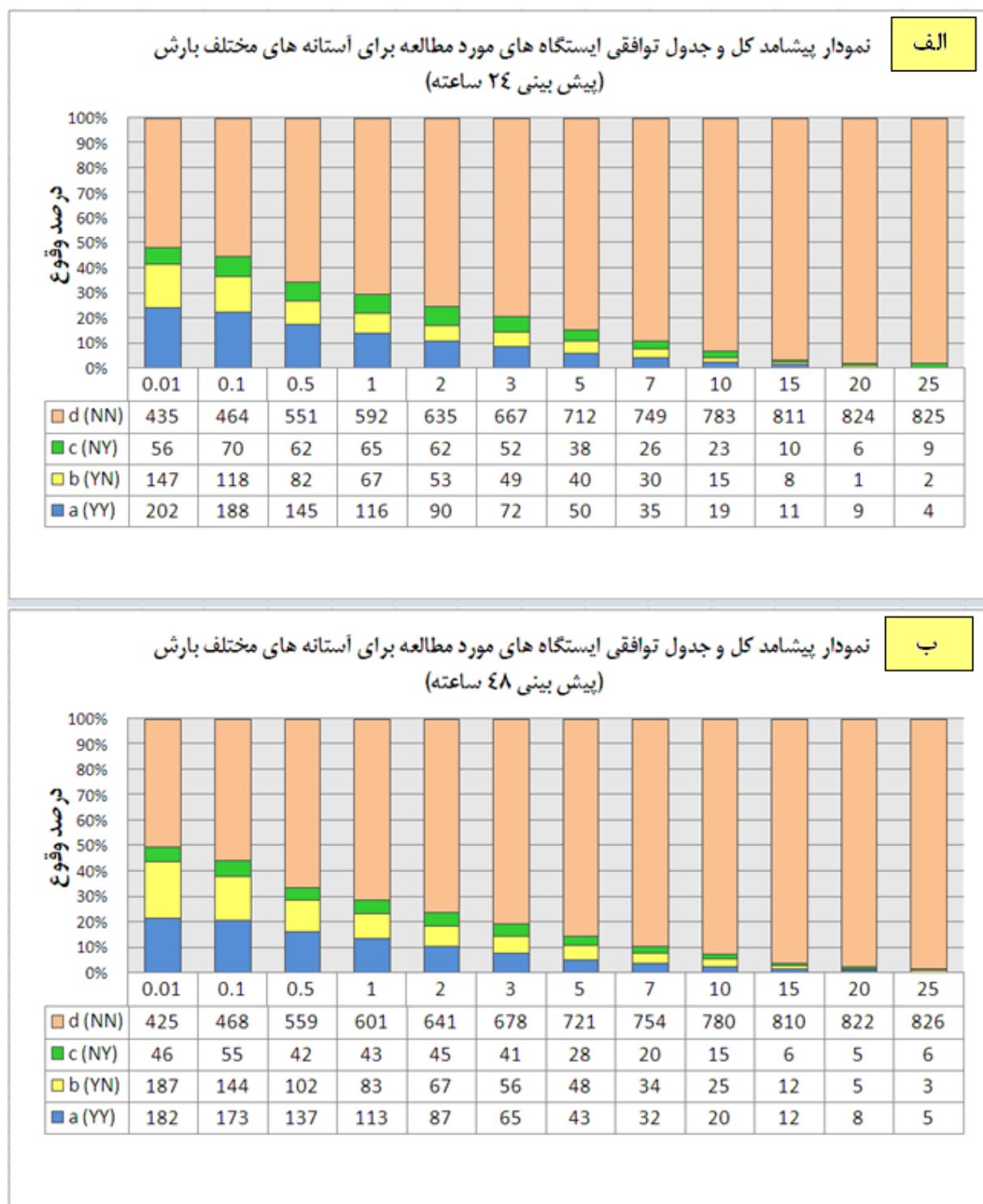
برآورد می‌کند. این عدد به تنها یکی از مهارت مدل نیست، زیرا ممکن است تعداد دفعات پیش‌بینی شده و مشاهده شده یکسان باشد ولی مقادیر بارش اختلاف زیادی داشته باشند. با توجه به جدول‌های ۴ و ۵ برای پیش‌بینی‌های ۲۴ ساعته بارش، اربی بسامدی با افزایش آستانه بارش، روندی کاهشی دارد، به گونه‌ای که مدل برای آستانه ۱/۰ میلی‌متر تعداد دفعات بارش را بیشتر از تعداد دفعات مشاهده شده بارش برآورد می‌کند و در



شکل ۵. سری‌های زمانی مجموع بارش روزانه مشاهده شده و شبیه‌سازی شده برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته در ۳۰ ایستگاه منتخب همدیدی در مراکز استان‌های ایران در ماه فوریه ۲۰۰۷.

جدول ۳. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در فوریه ۲۰۰۷ برای آستانه‌های بارش ۱/۰، ۰/۱، ۰/۵ و ۰/۸ میلی‌متر بر روز برای پیش‌بینی ۲۴ ساعته بارش کمتر.

آستانه بارش						
۲۰ mm	۱۰ mm	۵ mm	۱ mm	۰/۱ mm		
۰/۰۲	۰/۰۵	۰/۱	۰/۲۲	۰/۳۱	S	آهنج پایه
۰/۰۱	۰/۰۴	۰/۱۱	۰/۲۲	۰/۳۶	T	احتمال رخداد یک پیش‌بینی
۰/۶۷	۰/۸۱	۱/۰۲	۱/۰۱	۱/۱۹	Bias	اربی بسامدی
۰/۶	۰/۴۵	۰/۵۷	۰/۶۴	۰/۷۳	H	معیار بخورد
۰	۰/۰۲	۰/۰۵	۰/۱	۰/۲	F	معیار هشدار نادرست
۰/۹۹	۰/۹۵	۰/۹۱	۰/۸۴	۰/۷۸	PC	نسبت درست
۱۰۵۹	۳۹۱	۴۵۶	۵۲۰	۴۶۴		نسبت شناس
۰/۶	۰/۴۳	۰/۵۱	۰/۰۴	۰/۵۳	TSS - PSS	امتیاز مهارتی پیرسن (آمل مهارت‌های درست)
۰/۱	۰/۴۴	۰/۴۴	۰/۳۷	۰/۳۹	FAR	نسبت هشدارهای نادرست
۰/۵۶	۰/۲۳	۰/۳۹	۰/۴۷	۰/۵	CSI	شاخص موفقیت بحرانی
۲/۳۵	۱/۹۳	۱/۸۲	۱/۷۴	۱/۴۵	d'	فاصله تمایز
۰/۸	۰/۷۲	۰/۷۶	۰/۷۷	۰/۷۶	A _z	مساحت زیر منحنی ROC



شکل ۶. (الف) نمودار پیشامد و جدول توافقی برای آستانه های بارش $0/0.1, 0/0.5, 0/1, 0/2, 0/3, 0/5, 0/7, 0/10, 0/15, 0/20$ و $0/25$ میلی‌متر برای ۳۰ ایستگاه همدیدی منتخب (پیش‌بینی ۲۴ ساعته). متغیرهای a, b, c و d مربوط به جدول های توافقی دودویی هستند که از مقادیر خام بارش مشاهده شده و پیش‌بینی شده محاسبه شده‌اند. (ب) مشابه قسمت الف ولی برای پیش‌بینی ۴۸ ساعته.

جدول ۴. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در فوریه ۲۰۰۷ برای آستانه‌های بارش ۱۰/۱، ۵/۰، ۱۰ و ۲۰ میلی‌متر بر روز برای پیش‌بینی ۴۸ ساعته بارش کنی.

آستانه بارش	A_{sh}	S	F	PC	TSS - PSS	FAR	CSI	d'	A_z	ROC
آهنگ پایه	۰/۰۲	۰/۰۴	۰/۰۸	۰/۱۹	۰/۲۷	۱				
احتمال رخداد یک پیش‌بینی	۰/۰۳	۰/۰۵	۰/۱۱	۰/۲۳	۰/۳۸					
اریبی سامدی	۱/۲۳	۱/۲۹	۱/۲۸	۱/۲۶	۱/۳۹	Bias				
معیار برخورد	۰/۶۷	۰/۵۷	۰/۶۱	۰/۷۲	۰/۷۶	H				
معیار هشدار نادرست	۰/۰۱	۰/۰۳	۰/۰۶	۰/۱۲	۰/۲۴					
نسبت درست	۰/۹۸	۰/۹۵	۰/۹۱	۰/۸۵	۰/۷۶					
نسبت شانس	۵۴۰	۳۹۰	۴۰۸	۵۲۹	۴۰۷					
امتیاز مهارتی پیرس (آمار مهارت‌های درست)	۰/۶۵	۰/۵۴	۰/۵۴	۰/۶	۰/۵۲					
نسبت هشدارهای نادرست	۰/۵	۰/۵۶	۰/۵۳	۰/۴۲	۰/۴۵	FAR				
شاخص موفقیت بحرانی	۰/۴	۰/۳۳	۰/۲۶	۰/۴۷	۰/۴۷	CSI				
فاصله تمايز	۲/۷۷	۲/۰۶	۱/۱۸۴	۱/۶۶	۱/۴۲	d'				
مساحت زیر منحنی	۰/۸۳	۰/۷۷	۰/۷۷	۰/۸	۰/۷۶					

موقع پیش‌بینی یا فقدان پیش‌بینی بارش درست بوده است. با افزایش آستانه بارش این نسبت افزایش می‌یابد. برای پیش‌بینی‌های بارش ۲۴ و ۴۸ ساعته در هر آستانه انتخابی این نسبت نزدیک به هم است. برای نمونه به منظور پیش‌بینی بارش ۲۴ ساعته و آستانه بارش ۱/۰ میلی‌متر، نسبت درست برابر ۰/۷۸ است، یعنی در ۷۸ درصد موقع پیش‌بینی یا فقدان پیش‌بینی بارش درست بوده است. نسبت شانس نیز برای پیش‌بینی ۲۴ و ۴۸ ساعته در این جدول‌های در آستانه‌های بارش انتخابی محاسبه شده است. نسبت شانس برای همه آستانه‌ها در پیش‌بینی ۲۴ ساعته از پیش‌بینی ۴۸ ساعته بیشتر است. برای نمونه، نسبت شانس ۰/۴۶۴ در آستانه بارش ۱/۰ برای پیش‌بینی ۲۴ ساعته به این معنی است که اگر رخداد پیش‌بینی بارش مثبت باشد، احتمال رخداد مشاهده مثبت بارش نسبت به مشاهده منفی بارش در حوزه فضایی تحت بررسی و بازه زمانی یک ماهه، ۰/۴۶۴ برابر است.

امتیاز مهارتی پیرس (آمار مهارت‌های درست) برای آستانه‌های در نظر گرفته شده در پیش‌بینی‌های ۲۴ ساعته مقدار این امتیاز مهارتی بین ۰/۰۴۳ تا ۰/۰۶ و در پیش‌بینی -

معیار برخورد H بخشی از رخداد است که به درستی پیش‌بینی می‌شود. این میزان برای پیش‌بینی‌های ۴۸ ساعته نسبت به پیش‌بینی‌های ۲۴ ساعته برای همه آستانه‌های بارش در نظر گرفته شده، اندکی بیشتر است. برای پیش‌بینی ۲۴ ساعته این مقدار معیار برخورد بین ۰/۰۴۵ تا ۰/۰۷۳ و برای پیش‌بینی ۴۸ ساعته بین ۰/۰۵۷ تا ۰/۰۷۶ متغیر است. در هر دو پیش‌بینی‌های ۲۴ و ۴۸ ساعته معیار برخورد با افزایش آستانه بارش تا ۱۰ میلی‌متر کاهش می‌یابد ولی برای آستانه ۲۰ میلی‌متر مقدار معیار برخورد افزایش می‌یابد.

معیار هشدار نادرست که کسری از فقدان رخداد است که به طور نادرستی پیش‌بینی می‌شود، برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش در این جدول‌ها نشان داده شده است. معیار هشدار نادرست با افزایش آستانه بارش برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته کاهش می‌یابد و در پیش‌بینی‌های ۲۴ ساعته مقدار آن کمتر است. هرچه در یک آستانه خاص این مقدار کمتر باشد، پیش‌بینی عملکرد مطلوب تری دارد.

نسبت درست نشان‌دهنده این است که در چند درصد

فاصله تمایز برای پیش‌بینی‌های ۲۴ ساعته بارش بین ۱/۴۵ و ۳/۳۵ و برای پیش‌بینی‌های ۴۸ ساعته بین ۱/۴۲ تا ۲/۷۷ برای آستانه‌های منتخب تغییر می‌کند. مقدار فاصله تمایز وابسته به متغیرهای H و F است و از جدول‌های توزیع نرمال استاندارد به دست می‌آید. به طور کلی در این جدول‌ها برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته فاصله تمایز با افزایش آستانه بارش انتخابی، افزایش می‌یابد. برای آستانه‌های ۰/۱، ۱ و ۲۰ میلی‌متر فاصله تمایز بین رخداد یا فقدان رخداد پیش‌بینی بارش برای پیش‌بینی ۲۴ ساعته بیشتر است ولی برای آستانه ۵ و ۱۰ میلی‌متر پیش‌بینی ۴۸ ساعته فاصله تمایز را اندکی بهتر نشان می‌دهد. مساحت زیر منحنی ROC برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته محاسبه شده است. این مساحت در حکم یک امتیاز مهارتی برای هر آستانه در نظر گرفته می‌شود و ارتباط مستقیمی با امتیاز مهارتی پرسن دارد. اندازه مساحت برای پیش‌بینی ۲۴ و ۴۸ ساعته حاکی از این است که پیش‌بینی‌های بارش برای همه آستانه‌ها امتیاز قابل قبولی دارند (حدوداً بین ۰/۷ تا ۰/۸).

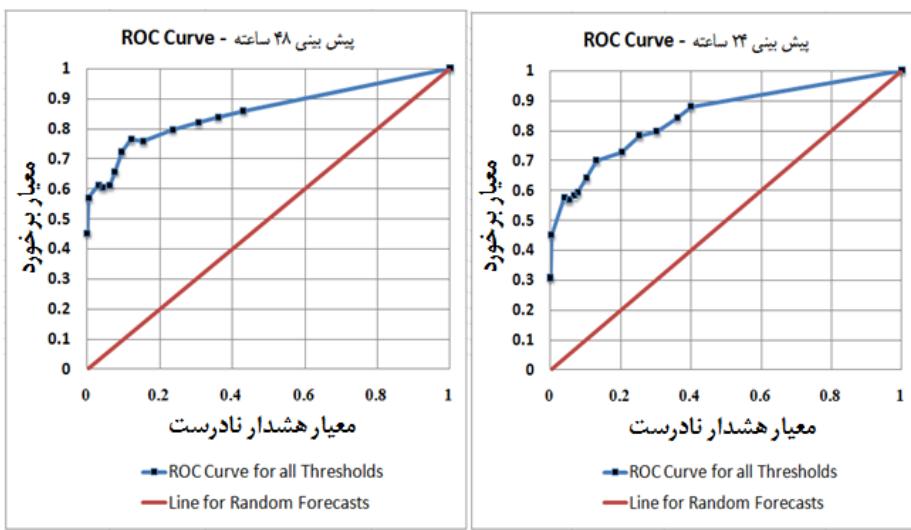
منحنی ROC برای مقادیر بارش مشاهده شده و پیش‌بینی شده در حوزه تحت بررسی در شکل ۷ برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته و آستانه‌های متفاوت ترسیم شده است. خط راست مورب قرار گرفته روی منحنی مربوط به پیش‌بینی‌های تصادفی است. نقاطی که روی نمودار در سطح بالاتری قرار گرفته‌اند مربوط به آستانه‌های بارش پایین‌تر هستند. هر دو منحنی شباهت زیادی به هم دارند، تنها در آستانه‌های خیلی پایین، ROC مربوط به پیش‌بینی ۲۴ ساعته مقداری معیار برخورد بیشتر و معیار هشدار نادرست کمتری را نشان می‌دهد اما در آستانه‌های خیلی بالا، معیار برخورد پیش‌بینی‌های ۴۸ ساعته بیشتر است.

های ۴۸ ساعته بین ۰/۵۲ تا ۰/۶۵ متغیر است. برای یک پیش‌بینی ایدئال، مقدار این امتیاز مهارتی ۱ است.

کمیت نسبت هشدارهای نادرست نشان می‌دهد که چند درصد موارد پیش‌بینی شده نادرست بوده است. این نسبت برای پیش‌بینی‌های ۲۴ ساعته بارش در همه آستانه‌ها نسبت به پیش‌بینی‌های ۴۸ ساعته کمتر است. پیرای مثال در آستانه ۰/۱ میلی‌متر، ۳۹ درصد موارد در پیش‌بینی ۲۴ ساعته و ۴۵ درصد موارد در پیش‌بینی ۴۸ ساعته، موارد پیش‌بینی شده بارش نادرست بوده است.

شاخص موفقیت بحرانی یا امتیاز خطر به طور گسترده‌ای در حکم اندازه‌ای از اجرا برای پیش‌بینی وقایع نادر استفاده می‌شود. در مهارت کامل مقدار آن برابر یک است و مقدار کمینه آن برابر صفر است. مقدار برای مهارت صفر بسته به نسبت پیش‌بینی‌های رخداد به فقدان رخداد می‌تواند بین صفر و آهنگ پایه اقلیمی (S) باشد. این امتیاز وابسته به آهنگ پایه اقلیمی است. برای پیش‌بینی‌های ۲۴ ساعته بارش این مقدار بین ۰/۳۳ تا ۰/۵۶ و برای پیش‌بینی‌های ۴۸ ساعته بین ۰/۳۳ تا ۰/۴۷ برای آستانه‌های منتخب است.

نمودار مشخصه عملکرد نسبی (ROC) و اندازه‌گیری‌های مرتبط با آن، توانایی ذاتی سامانه پیش‌بینی برای تمایز رخداد یا فقدان رخداد پیشامد و همچنین برآوردهای کمی از احتمالات خروجی‌های پیش‌بینی برای هر آستانه تصمیم‌گیری را نشان می‌دهد که دو مورد از اندازه‌های مرتبط با آن یعنی مساحت زیر منحنی ROC (A_Z) و فاصله تمایز (d') برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته روی کل حوزه تحت بررسی در بازه زمانی یک‌ماهه در جدول ۳ و ۴ محاسبه شده است.



شکل ۷. نمودار مشخصه عملکرد نسبی (ROC) مربوط به پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش کمی مدل در حوزه فضایی تحت بررسی برای بازه یکماهه فوریه ۲۰۰۷.

بین ۰/۱۲ تا ۰/۰۹ متر برای پیش‌بینی‌های ۴۸ ساعته در آستانه‌های در نظر گرفته شده تغییر می‌کند. البته امتیاز پیرسون برای بیشتر آستانه‌ها و در روزهای مورد بررسی برای پیش‌بینی ۲۴ ساعته بیشتر است.

جدول ۹ بعضی از امتیازات ارزیابی پیوسته را برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته در حوزه فضایی تحت بررسی برای روزهای اوج فعالیت بارش در ماه فوریه ۲۰۰۷ نشان می‌دهد. به طور کلی میانگین‌های بارش در پیش‌بینی‌های ۴۸ ساعته نسبت به پیش‌بینی‌های ۲۴ ساعته در همه روزهای ذکر شده برآورد اضافی دارد. به نظر می‌رسد که این حالت برای همه روزهای اوج فعالیت بارشی قابل توجه روی ایران در بازه زمانی مورد بررسی وجود داشته باشد. خطای اریبی در هر دوی پیش‌بینی‌های ۲۴ و ۴۸ ساعته برای ۴ و ۸ فوریه منفی و برای ۷ و ۱۹ فوریه مثبت است؛ یعنی در روزهایی است که مقدار بارش بیشتری در ایستگاه‌های تحت بررسی رخ داده، مدل برآورد کمتری از میزان بارش داشته و در نتیجه در این روزها میانگین پیش‌بینی‌های ۴۸ ساعته به مشاهدات نزدیک‌تر است.

علاوه بر امتیازات و شاخص‌های ارزیابی برای کل دوره یکماهه، امتیازات و شاخص‌های ارزیابی برای روزهایی که بارش قابل ملاحظه‌ای روی کشور اتفاق افتاده برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته محاسبه شده است. نتایج در جدول‌های ۵، ۶، ۷ و ۸ خلاصه شده است. آستانه‌های در نظر گرفته شده برای بارش، ۰/۱، ۱ و ۵ میلی‌متر است. میزان آهنگ پایه و احتمال رخداد یک پُربارش نسبت به میانگین یکماهه خیلی بیشتر است. اریبی بسامدی بیانگر برآورد اضافی یا کم تعداد بارش‌هایی است که در حوزه فضایی اتفاق افتاده است. در روزها و آستانه‌هایی که اریبی بسامدی بزرگ‌تر از ۱ است مدل تعداد بارش‌ها را روی منطقه زیاد برآورد کرده است. برای مثال در روز ۷ فوریه برای هر سه آستانه بارش در نظر گرفته شده این حالت برقرار است. در پیش‌بینی‌های ۴۸ ساعته برای روزهای تحت بررسی این حالت بیشتر اتفاق افتاده است.

امتیاز مهارتی پیرس در روزهای پُربارش ماه فوریه ۲۰۰۷ بین ۰/۳۴ تا ۰/۷۶ برای پیش‌بینی‌های ۲۴ ساعته و

جدول ۵. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در روزهای اوج فعالیت سامانه‌های جوی (۴ و ۷ فوریه ۲۰۰۷) و برای آستانه‌های بارش ۱، ۱ و ۵ میلی‌متر بر روز برای پیش‌بینی ۲۴ ساعته بارش کمی.

۷ فوریه							۴ فوریه		آستانه بارش
۵ mm	۱ mm	۰/۱mm	۵ mm	۱ mm	۰/۱mm				
۰/۲۷	۰/۵	۰/۶۷	۰/۶	۰/۷۳	۰/۷۷	s			آهنگ پایه
۰/۳۷	۰/۵۳	۰/۷۳	۰/۵۳	۰/۷	۰/۸۳	r			احتمال رخداد یک پیش‌بینی
۱/۳۸	۱/۰۷	۱/۱	۰/۸۹	۰/۹۵	۱/۰۹	Bias			اریبی بسامدی
۰/۶۳	۰/۸۷	۰/۹	۰/۷۸	۰/۸۶	۰/۹۱	H			معیار برخورد
۰/۲۷	۰/۲	۰/۴	۰/۱۷	۰/۲۵	۰/۵۷	F			معیار هشدار نادرست
۰/۷	۰/۸۳	۰/۸	۰/۸	۰/۸۳	۰/۸	PC			نسبت درست
۹	۳۱	۱۸	۲۳	۲۳	۱۱				نسبت شناس
۰/۳۵	۰/۶۷	۰/۵	۰/۶۱	۰/۶۱	۰/۳۴	TSS			امتیاز مهارتی پیرس (آمار مهارت‌های درست)
۰/۵۵	۰/۱۹	۰/۱۸	۰/۱۳	۰/۱	۰/۱۶	FAR			نسبت هشدارهای نادرست
۰/۳۶	۰/۷۲	۰/۷۵	۰/۷	۰/۷۹	۰/۷۸	CSI			شاخص موفقیت بحرانی
۰/۹۴	۱/۸۷	۱/۵۴	۱/۷۲	۱/۶۶	۱/۱۶	d'			فاصله تمایز
۰/۶۷	۰/۸۳	۰/۷۵	۰/۸	۰/۸	۰/۶۷	A _z			مساحت زیر منحنی ROC

جدول ۶. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در روزهای اوج فعالیت سامانه‌های جوی (۸ و ۱۹ فوریه ۲۰۰۷) و برای آستانه‌های بارش ۱، ۱ و ۵ میلی‌متر بر روز برای پیش‌بینی ۲۴ ساعته بارش کمی.

۱۹ فوریه							۸ فوریه		آستانه بارش
۵ mm	۱ mm	۰/۱mm	۵ mm	۱ mm	۰/۱mm				
۰/۱۷	۰/۴۳	۰/۶	۰/۴۳	۰/۵۷	۰/۷۳	s			آهنگ پایه
۰/۱۷	۰/۳۷	۰/۵۷	۰/۲۳	۰/۴	۰/۵۷	r			احتمال رخداد یک پیش‌بینی
۱	۰/۸۵	۰/۹۴	۰/۵۴	۰/۷۱	۰/۷۷	Bias			اریبی بسامدی
۰/۸	۰/۶۲	۰/۷۸	۰/۵۴	۰/۷۱	۰/۶۸	H			معیار برخورد
۰/۰۴	۰/۱۸	۰/۲۵	۰	۰	۰/۲۵	F			معیار هشدار نادرست
۰/۹۳	۰/۷۳	۰/۷۷	۰/۸	۰/۸۳	۰/۷	PC			نسبت درست
۴۸	۱۴	۱۸	۲۰	۳۱	۱۰				نسبت شناس
۰/۷۶	۰/۴۴	۰/۵۳	۰/۵۴	۰/۷۱	۰/۴۳	TSS			امتیاز مهارتی پیرس (آمار مهارت‌های درست)
۰/۲	۰/۲۷	۰/۱۸	۰	۰	۰/۱۲	FAR			نسبت هشدارهای نادرست
۰/۶۷	۰/۵	۰/۶۷	۰/۵۴	۰/۷۱	۰/۶۳	CSI			شاخص موفقیت بحرانی
۲/۵۹	۱/۲۳	۱/۴۵	۲/۱۹	۲/۶۴	۱/۱۵	d'			فاصله تمایز
۰/۸۸	۰/۷۲	۰/۷۶	۰/۷۷	۰/۸۵	۰/۷۱	A _z			مساحت زیر منحنی ROC

جدول ۷. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در روزهای اوج فعالیت سامانه‌های جوی (۴ و ۷ فوریه ۲۰۰۷) و برای آستانه‌های بارش ۰/۱ و ۵ میلی‌متر بر روز برای پیش‌بینی ۴۸ ساعته بارش کهنس.

۷ فوریه				۴ فوریه				آسانه بارش
۵ mm	۱ mm	۰/۱mm	۵ mm	۱ mm	۰/۱mm			
۰/۲۳	۰/۱۴	۰/۶	۰/۵	۰/۶۳	۰/۶۷	s	آنگ پایه	
۰/۳۷	۰/۷	۰/۸۳	۰/۵۷	۰/۷۳	۰/۸۳	r	احتمال رخداد یک پیش‌بینی	
۱/۵۷	۱/۶۲	۱/۳۹	۱/۱۳	۱/۱۶	۱/۲۵	Bias	اریبی بسامدی	
۰/۷۱	۰/۷۷	۰/۹۴	۰/۸	۰/۸۴	۰/۹۵	H	معیار برخورد	
۰/۲۶	۰/۶۵	۰/۶۷	۰/۲۳	۰/۵۵	۰/۶	F	معیار هشدار نادرست	
۰/۷۳	۰/۵۳	۰/۷	۰/۷۳	۰/۷	۰/۷۷	PC	نسبت درست	
۱۱	۴	۸	۱۵	۹	۱۱		نسبت شناس	
۰/۴۵	۰/۱۲	۰/۲۸	۰/۴۷	۰/۳	۰/۳۵	TSS	امتیاز مهارتی پیرس (آمار مهارت‌های درست)	
۰/۵۵	۰/۵۲	۰/۲۲	۰/۲۹	۰/۲۷	۰/۲۴	FAR	نسبت هشدارهای نادرست	
۰/۳۸	۰/۴۲	۰/۶۵	۰/۶	۰/۶۴	۰/۷۳	CSI	شاخص موفقیت بحرانی	
۱/۲	۰/۳۵	۱/۱۲	۱/۵۸	۰/۸۸	۱/۳۹	d'	فاصله تمایز	
۰/۷۳	۰/۵۶	۰/۶۴	۰/۷۳	۰/۶۵	۰/۶۸	A _z	مساحت زیر منحنی ROC	

مشاهداتی خطای RMSE و خطای انحراف معیار برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته افزایش می‌یابد. ضریب همبستگی بین مقادیر بارش مشاهداتی و پیش‌بینی شده برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته در ایستگاه‌های تحت بررسی نیز موید آن است که به طور کلی پیش‌بینی‌های ۲۴ ساعته همبستگی بیشتری دارند.

برعکس روزهای با میانگین بارش مشاهداتی کمتر (مانند ۷ و ۱۹ فوریه) مدل برآورده اضافی از بارش دارد و میزان این برآورده برای پیش‌بینی ۲۴ ساعته کمتر است و بنابراین میانگین پیش‌بینی ۲۴ ساعته به میانگین مشاهدات نزدیک‌تر است. در روزهای با میانگین بارش خیلی کم، نظم خاصی وجود ندارد. برای هر چهار روز مورد بررسی، با افزایش میانگین بارش

جدول ۸. شاخص‌ها و امتیازات مهارتی پیش‌بینی‌های دودویی بارش برای کل حوزه فضایی مورد بررسی در روزهای اوج فعالیت سامانه‌های جوی (۸ و ۱۹ فوریه ۲۰۰۷) و برای آستانه‌های بارش ۰/۱ و ۵ میلی‌متر بر روز برای پیش‌بینی ۴۸ ساعته بارش کهنس.

۱۹ فوریه				۸ فوریه				آستانه بارش
۵ mm	۱ mm	۰/۱mm	۵ mm	۱ mm	۰/۱mm			
۰/۱۷	۰/۴۳	۰/۵۳	۰/۳۳	۰/۴۷	۰/۶۳	s	آنگ پایه	
۰/۲۳	۰/۴	۰/۶۳	۰/۱۳	۰/۳۳	۰/۵	r	احتمال رخداد یک پیش‌بینی	
۱/۴	۰/۹۲	۱/۱۹	۰/۴	۰/۷۱	۰/۷۹	Bias	اریبی بسامدی	
۱	۰/۷۷	۰/۸۸	۰/۴	۰/۵۷	۰/۶۳	H	معیار برخورد	
۰/۰۸	۰/۱۲	۰/۳۶	۰	۰/۱۳	۰/۲۷	F	معیار هشدار نادرست	
۰/۹۳	۰/۸۳	۰/۷۷	۰/۸	۰/۷۳	۰/۶۷	PC	نسبت درست	
۵۸	۳۰	۱۸	۱۳	۱۴	۱۰		نسبت شناس	
۰/۹۲	۰/۶۵	۰/۵۲	۰/۶	۰/۴۵	۰/۳۶	TSS	امتیاز مهارتی پیرس (آمار مهارت‌های درست)	
۰/۲۹	۰/۱۷	۰/۲۶	۰	۰/۲	۰/۲	FAR	نسبت هشدارهای نادرست	
۰/۷۱	۰/۶۷	۰/۶۷	۰/۴	۰/۵	۰/۵۵	CSI	شاخص موفقیت بحرانی	
۲/۸۲	۱/۹۲	۱/۵۶	۲/۲۴	۱/۳۱	۰/۹۴	d'	فاصله تمایز	
۰/۹۶	۰/۸۳	۰/۷۶	۰/۷	۰/۷۲	۰/۶۸	A _z	مساحت زیر منحنی ROC	

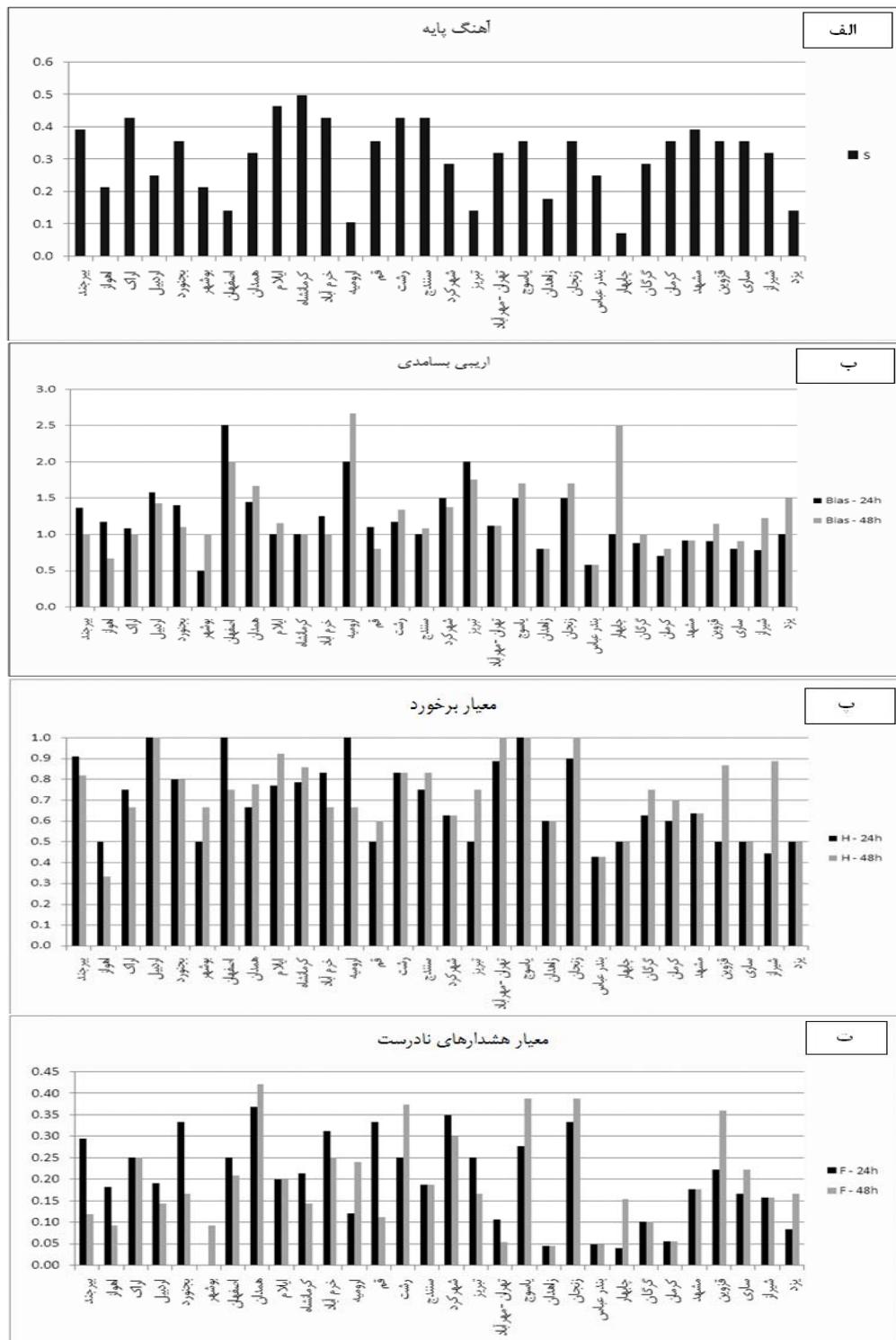
جدول ۹. بعضی از امتیازات ارزیابی پیوسته برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته در حوزه تحت بررسی برای روزهای اوج فعالیت سامانه‌های جوی در ماه فوریه ۲۰۰۷

روزهای اوج بارش	۴فوریه	۷فوریه	۱۰فوریه	۱۳فوریه	۱۶فوریه
میانگین بارش مشاهده شده (میلی‌متر)					۶/۲
میانگین بارش برای پیش‌بینی ۲۴ ساعته (میلی‌متر)	۳	۳/۴	۶/۴	۱۰/۳	
میانگین بارش برای پیش‌بینی ۴۸ ساعته (میلی‌متر)	۴/۳	۳/۹	۶/۶	۱۰/۸	
خطای اربیی برای پیش‌بینی ۲۴ ساعته	۰/۳	-۲/۸	۲/۴	-۲/۵	
خطای اربیی برای پیش‌بینی ۴۸ ساعته	۱/۷	-۲/۳	۲/۵	-۳	
خطای RMSE برای پیش‌بینی ۲۴ ساعته	۳/۸	۶/۵	۶/۲	۱۳/۷	
خطای RMSE برای پیش‌بینی ۴۸ ساعته	۳/۸	۸/۳	۵/۷	۱۶/۴	
خطای انحراف معیار برای پیش‌بینی ۲۴ ساعته	۳/۸	۵/۸	۵/۸	۱۳/۳	
خطای انحراف معیار برای پیش‌بینی ۴۸ ساعته	۳/۴	۸	۵/۲	۱۶/۲	
همبستگی برای پیش‌بینی ۲۴ ساعته	۰/۷۶	۰/۸۱	۰/۹۱	۰/۶۳	
همبستگی برای پیش‌بینی ۴۸ ساعته	۰/۹۶	۰/۶۴	۰/۸۹	۰/۴۹	

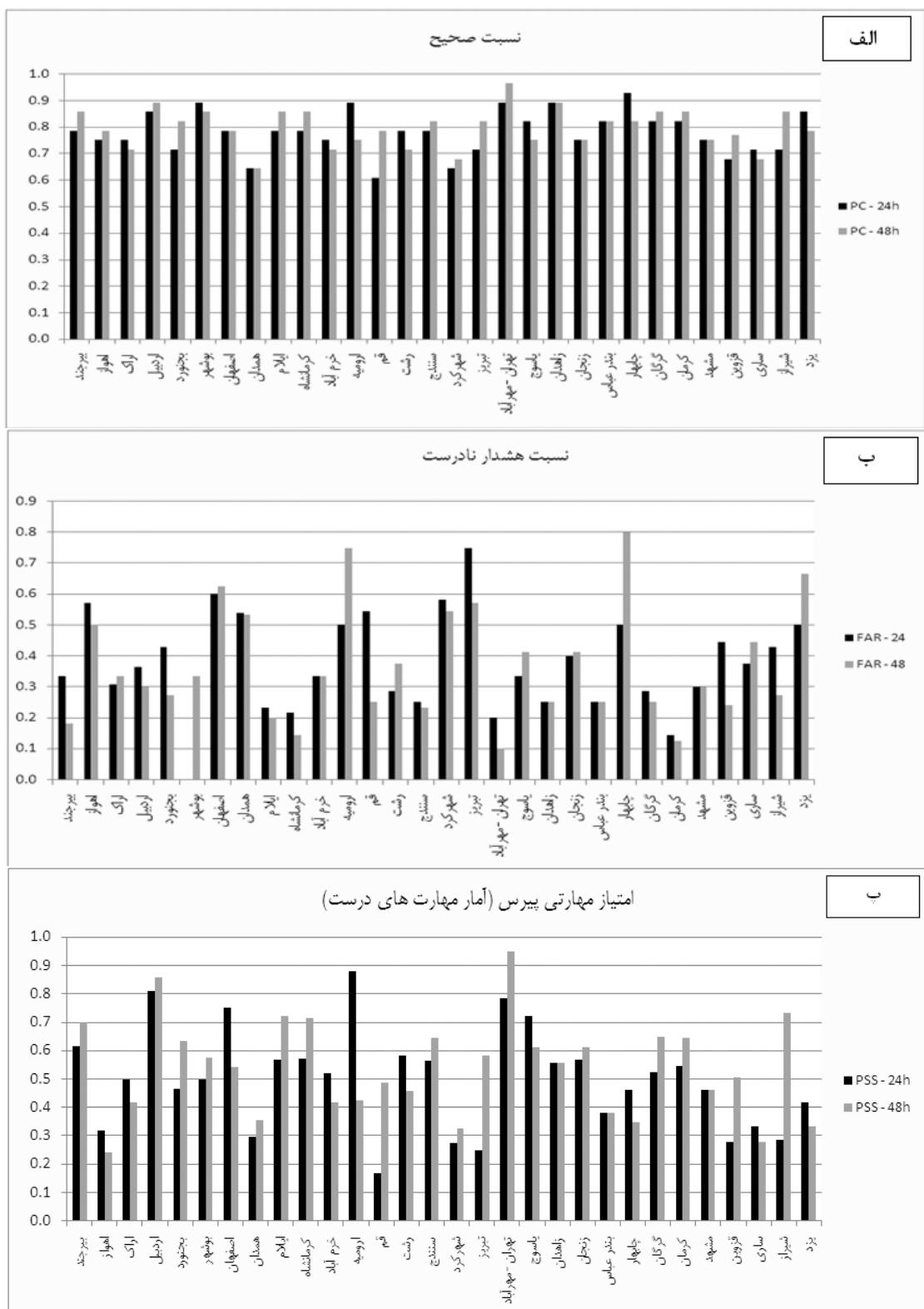
تعداد دفعات کمتری نسبت به تعداد دفعات بارش مشاهده شده دارد.

در ایستگاه‌هایی مانند اراک، کرمانشاه و ایلام، مقدار اُریبی بسامدی نزدیک ۱ است و مدل بهخوبی تعداد دفعات بارش را پیش‌بینی کرده است (شکل ۸-ب). میزان معیار برخورد برای ایستگاه‌های یاسوج و اردبیل برای هر دوی پیش‌بینی‌های ۲۴ و ۴۸ ساعته بسیار مناسب است (برابر ۱)، اما در ایستگاه‌هایی مانند اهواز و بندرعباس، معیار برخورد نسبت به دیگر ایستگاه‌ها برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته بارش کمتر است (شکل ۸-پ). در ایستگاه‌هایی همدان، زنجان، شهرکرد و یاسوج معیار هشدار نادرست نسبت به ایستگاه‌های دیگر بیشتر و در ایستگاه‌هایی مانند زاهدان، کرمان و بندرعباس برای هر دو پیش‌بینی‌های ۲۴ و ۴۸ ساعته معیار هشدار نادرست کم است (شکل ۸-ت).

به منظور بررسی میزان مهارت پیش‌بینی بارش مدل در هر یک از ایستگاه‌های همدیدی منتخب در ماه فوریه ۲۰۰۷، بعضی از شاخص‌ها و امتیازات مهارتی ۳۰ ایستگاه همدیدی مورد بررسی برای آستانه بارش ۱/۰ میلی‌متر در شکل ۸ و ۹ نشان داده شده است. در میان ایستگاه‌های همدیدی مورد بررسی، ایستگاه‌های کرمانشاه، ایلام، خرم‌آباد، ستندج و رشت بیشترین آهنگ پایه و ایستگاه‌های چابهار و ارومیه کمترین آهنگ پایه در ماه فوریه ۲۰۰۷ را داشته‌اند (شکل ۸-الف). برای پیش‌بینی‌های ۲۴ ساعته، اصفهان، ارومیه و تبریز و برای پیش‌بینی‌های ۴۸ ساعته، ارومیه، چابهار و اصفهان بیشترین اُریبی بسامدی برای آستانه بارش ۱/۰ میلی‌متر را داشته‌اند. این بدین معنی است که مدل برای تعداد دفعات بارش در ایستگاه‌های ذکر شده برآورد اضافی زیادی دارد. بر عکس، در ایستگاه‌هایی مثل بوشهر و بندرعباس برای پیش‌بینی ۲۴ ساعته مدل برآورد



شکل ۸ (الف) آهنگ پایه اقلیمی برای بارش ۳۰ ایستگاه همدیدی کشور در ماه فوریه ۲۰۰۷، (ب) مشابه قسمت الف اما برای اربی بسامدی پیش‌بینی‌های ۲۴ و ۴۸ ساعته، (پ) مشابه قسمت ب اما برای آهنگ برخورد و (ت) مشابه قسمت ب اما برای آهنگ هشدارهای نادرست.



شکل ۹. (الف) نسبت صحیح برای پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش ۳۰ ایستگاه همدیدی کشور در ماه فوریه ۲۰۰۷، (ب) مشابه قسمت (الف) اما برای نسبت هشدارهای نادرست و (پ) مشابه قسمت (الف) اما برای امتیاز مهارتی پیرس.

متفاوت مقایسه شد. همچنین از مفاهیم نظریه آشکارسازی سیگنال و منحنی‌های ROC برای ارزیابی پیش‌بینی‌های بارش در آستانه‌های متفاوت استفاده شد. نتایج نشان داد که پیش‌بینی‌های مدل WRF برای بسیاری از مناطق موقعیت نوارهای بارش را به خوبی نشان می‌دهد (همان‌طور که برای میانگین ماهانه پیش‌بینی‌های ۲۴ و ۴۸ ساعته بارش روزانه مشاهده شد)، اما هنوز مسائل زیادی در خصوص برآورد دامنه بارش در هر منطقه وجود دارد. فرض دیگر در استفاده از روش‌های دودویی این است که مدل زمان بارش را دقیقاً به درستی پیش‌بینی کرده است و هیچ‌گونه اختلاف فازی بین داده‌های روزانه مشاهده شده و پیش‌بینی شده در کل منطقه وجود ندارد؛ که تقریباً فرض درستی بود، زیرا رفتار بارش پیش‌بینی شده و مشاهده شده در سری‌های زمانی مقدار بارش روزانه تا حدود زیادی مشابه بود. آمار مهارت‌های درست مدل در کل حوزه تحت بررسی و بازه یک‌ماهه فوریه ۲۰۰۷ برای همه آستانه‌های بارش در نظر گرفته شده برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته مقدار بارش روزانه مقدار مناسبی داشت، اگرچه در روزهای اوج فعالیت سامانه‌های جویی، پیش‌بینی‌های ۲۴ ساعته دارای امتیاز بیشتری بودند.

منبع اصلی خطا در پیش‌بینی بارش، بیشتر ناشی از خطاهای تصادفی است که بخشی از آن ناشی از عدم قطعیت در شرایط اولیه و انتخاب مقیاس‌های غیرمنطقی برای اجرای مدل است. این خطاهای تصادفی مانع پیش‌بینی دقیق مقادیر بارش خواهند شد.

نتایج نشان داده شده در این تحقیق فقط مربوط یک ماه است، شناسایی دقت در عملکرد پیش‌بینی ها، نیاز به ارزیابی و تحقیق پیش‌بینی‌های زمان واقعی بیشتری دارد. آمارهای و نتایج بدست آمده در این تحقیق و تحقیقات مشابه، اطلاعات اضافی برای کاربران پیش‌بینی و توسعه دهنده‌گان مدل فراهم خواهد کرد، به‌گونه‌ای که منجر به شناسایی مناطقی خواهد شد که نیاز به توجه و تحقیق

نسبت صحیح برای همه ایستگاه‌ها و پیش‌بینی‌های ۲۴ و ۴۸ ساعته بین ۰/۹ تا ۰/۶ برای ایستگاه‌های متفاوت، متغیر است (شکل ۹-الف). نسبت هشدارهای نادرست در ایستگاه‌های اهواز، اصفهان، همدان، ارومیه، قم، شهرکرد، تبریز، چابهار و یزد دارای مقدار زیادی است (بیش از ۰/۵) و در ایستگاه‌هایی مانند کرمان و تهران-مهرآباد مقدار آن کم است (شکل ۹-ب). در شکل ۹-پ امتیاز مهارتی پیرس که مستقل از آهنگ پایه اقلیمی است برای ایستگاه‌های همدیدی منتخب برای آستانه بارش ۰/۱ میلی‌متر مقایسه شده است. ایستگاه تهران-مهرآباد و اردبیل برای هر دو پیش‌بینی ۲۴ و ۴۸ ساعته امتیاز مهارتی بالایی دارند. ایستگاه قم کمترین امتیاز مهارتی را برای پیش‌بینی ۲۴ ساعته و ایستگاه اهواز دارای کمترین امتیاز مهارتی برای پیش‌بینی ۴۸ ساعته بارش است.

۶ نتیجه‌گیری

در این تحقیق پیش‌بینی‌های کوتاه‌مدت بارش کمی روزانه مدل میان‌مقیاس مرکز NCAR (WRF-AEW) در یک دوره یک‌ماهه در زمستان ۲۰۰۷ روی منطقه ایران مورد ارزیابی قرار گرفت. به‌واسطه پیچیدگی منطقه و نبود اطلاعات کافی در برآورد فرایندهای فیزیکی در این ناحیه، پیش‌بینی‌ها باستی محدودیت‌های مدل را در نظر بگیرند.

به‌منظور بررسی عملکرد مدل در پیش‌بینی بارش‌های روزانه، بارش به‌منزله متغیر گستره در نظر گرفته و از روش‌های ارزیابی دودویی استفاده شد. البته به‌منظور مقایسه، خطاهای پیوسته نیز در شرایط خاص جویی محاسبه شد. با بررسی یک دوره یک‌ماهه در زمستان ۲۰۰۷، دقت مدل WRF در برآورد مقادیر بارش روزانه، ماهانه و روزهای با بارش سنگین روی منطقه ایران آزمایش شد. امتیازات مهارتی و شاخص‌های ارزیابی پیش‌بینی برای هر مورد خاص و برای آستانه‌های بارش

- verification of spatial precipitation forecasts, *Meteorological Applications*, **11**, 141-154.
- Davis, C., Brown, B., and Bullock, R., 2006, Object-based verification of precipitation forecast. Part I: Methodology and application to mesoscale rain areas, *Monthly Weather Review*, **134**, 1772-1784.
- Donaldson, R. J., Dyer, R. M. and Kraus, M. J., 1975, An objective evaluator of techniques for predicting severe weather events, In: Preprints, Ninth Conference on Severe Local Storms, Norman, Oklahoma, American Meteorological Society, 321-326.
- Ebert, E. E. and McBride, J. L., 1997, Methods for verifying quantitative precipitation forecasts: application to the BMRC LAPS model 24-hour precipitation forecast, Technique development. Report No. 2, BMRC, Melbourne, Australia.
- Evans, J. P., and R. Smith, 2001, Modelling the climate of south west Asia. Proc. Int. Congress on Modelling and Simulation (MODSIM01), Canberra, ACT, Australia, Modelling and Simulation Society of Australia and New Zealand Inc. [Available online at <http://www.mssanz.org.au/MODSIM01/Vol%202/Evans.pdf>.]
- Evans, J. P., and R. Smith, and R. J. Oglesby, 2004, Middle East climate simulation and dominant precipitation processes, *Int. J. Climatol.*, **24**, 1671-1694. Flueck, J. A., 1987, A study of some measures of forecast verification, In: P reprint s, 10th Conference on Probability and Statistics in Atmospheric Science. Edmonton, AB, Canada, American Meteorological Society, 69-73.
- Golding, B. W., 1998, NIMROD: a system for generating automated very short range forecast, *Meteorological Applications*, **5**, 1-16.
- Hanssen, A. W. and Kuipers,W. J. A., 1965, On the relationship between the frequency of rain and various meteorological parameters, *Mededeelingen en Verhandelingen*, Royal Netherlands Meteorological Institute, 81, 24-29.
- Johnson, L. E. and Olsen, B. G., 1998, Assessment of quantitative precipitation forecasts, *Weather and Forecasting*, **13**, 75-83.
- Knievel, J. C., Ahijevych, D. A. and Manning, K. W., 2004, Using temporal modes of rainfall to evaluate the performance of a numerical weather prediction model, *Monthly Weather Review*, **132**, 2995-3009.
- Marcella, M. P. and E. A. Eltahir, 2008, Modeling the hydroclimatology of Kuwait: The role of

بیشتری دارند. در این مناطق می‌توان با بررسی دقیق عمل فیزیکی رخداد خطاهای سامان‌مند، میزان این خطاهای تغییر طرحواره‌های مدل عددی و پارامترسازی فیزیکی کاهش داد. همچنین توجه به این نکته ضروری است که خطاهای تصادفی خیلی پیچیده هستند. گوارد داده‌های جوی تا اندازه‌ای می‌تواند شرایط اولیه نزدیک‌تر به جو واقعی را فراهم کند و آن بخش از این خطاهای را که مربوط به عدم قطعیت در شرایط اولیه است، بهبود بخشد.

مراجع

- آزادی، م.، قاضی‌سعید، م. و جعفری، س.، ۱۳۸۸، ارزیابی عملکرد مدل WRF برای پیش‌بینی بارش روی ایران به مدت یک ماه، دوازدهمین کنفرانس دینامیک شاره ها، دانشگاه صنعتی نوشیروانی بابل.
- مزرعه‌فرهانی، م.، وظیفه، ا. و آزادی، م.، ۱۳۸۸، بررسی اثر تغییر تفکیک افقی مدل میانمقیاس MM5 در شبیه‌سازی بارش حاصل از سامانه همدیدی اکتبر ۲۰۰۴ روی ایران، م. فیزیک زمین و فضا، **۳۵**(۴)، ۱۳۱-۱۴۶.

هدایتی دزفولی، ا. و آزادی، م.، ۱۳۸۹، راست‌آزمایی پیش‌بینی بارش مدل منطقه‌ای MM5 روی ایران، م. فیزیک زمین و فضا، **۳۶**(۳)، ۱۲۹-۱۱۵.

- Airey, M. and Hulme, M, 1995, Evaluating climate model simulations of precipitation: methods, problems and performances, *Progress in Physical Geography*, **19**, 427-448.
- Bougeault, P., 2003, The WGNE survey of verification methods for numerical prediction of weather elements and severe weather events, Technical report, WMO, Available at, <http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/Bougeault/> Bougeault Verication-methods.htm.
- Brier, G. W. and Allen, R. A., 1951, Verification of weather forecasts, In: *Compendium of Meteorology* (ed. T.F. Malone), Boston, American Meteorological Society, 841-848.
- Casati, B., Ross, G. and Stephenson, D. B., 2004, A new intensity-scale approach for the

- Palmer, W. C. and Allen, R. A., 1949, Note on the accuracy of forecasts concerning the rain problem, U.S. Weather Bureau manuscript, Washington, D.C.
- Potts, J. M., Folland, C. K., Jolliffe, I. T. and Sexton, D., 1996, Revised LEPS score for assessing climate model simulations and long range forecasts, *Journal of Climate*, **9**, 34-53.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W. and Powers, J. G., 2005, A description of the advanced research WRF, Version 2, NCAR Tech, Note NCAR/TN-4681STR, 94 pp.
- Sodoodi, S., Noorian, A. and Reimer, M. G., 2010, Daily precipitation forecast of ECMWF verified over Iran, *Theor Appl Climatol*, **99**, 39-51.
- Swets, J. A., 1986, Indices of discriminations or diagnostic accuracy, their ROCs and implied models, *Psychol. Bull.*, **99**, 100-117.
- Swets, J. A., 1988, Measuring the accuracy of diagnostic systems, *Science*, **240**, 1285-1293.
- Tanner, W. P. JR. and Birdsall, T. G., 1958, Definitions of d' and η as psychological measures, *J. Acoust. Soc. Amer.*, **30**, 922-928.
- Wilson, C., 2000, Review of current methods and tools for verification of numerical forecasts of precipitation, Technical Report WP2.1, UK Met Office.
- Xu, J., Rugg, S., Byerle, L. and Liu, Z., 2009, Weather forecasts by the WRF-ARW model with the GSI data assimilation system in the complex terrain areas of Southwest Asia, *Weather and Forecasting*, **24**, 987-1008.
- Xu, J., Rugg, S., Horner, M. and Byerle, L., 2009, Application of ATOVS radiance with ARW WRF/GSI data assimilation system in the prediction of Hurricane Katrina, *Open Atmos. Sci. J.*, **3**, 13-28.
- Zaitchik, B. F., Evans, J. P., Geerken, R. A. and Smith, R. B., 2007, Climate and vegetation in the Middle East: Interannual variability and drought feedbacks. *J. Climate*, **20**, 3924-3941.
- subcloud evaporation in semiarid climates, *J. Climate*, **21**, 2976-2989.
- Mason, I. B., 1982, A model for assessment of weather forecasts, *Austral. Met. Mag.*, **30**, 291-303.
- Mesinger, F., 1996, Improvements in quantitative precipitation forecasts with the ETA regional model at the national centers for environmental prediction, the 48-Km Upgrade, *Bulletin of American Meteorological Society*, **77**, 2637-2649.
- Michalakes, J., Chen, S., Dudhia, J., Hart, L., Klemp, J., Middlecoff, J. and Skamarock, W., 2001, Development of a next generation regional weather research and forecast model. Development in Teracomputing, *Proceedings of the Ninth ECMWF Workshop on the Use of High Performance Computing in Meteorology*, W. Zwiefelhofer and N. Kreitz, Eds., World Scientific, 269-276.
- Mladek, R., Barckicke, J., Binder, P., Bougeault, P., Brzovic, N., Frei, C., Geleyn, J. F., Hoffman, J., Ott, W., Paccagnella, T., Paturno, P., Pottier, P. and Rossa, A., 2000, Intercomparison and evaluation of precipitation forecasts for MAP seasons 1995 and 1996, *Meteorology and Atmospheric Physics*, **72**, 111-129.
- Murphy, A. H. and Winkler, R. L., 1987, A general framework for forecast verification, *Monthly Weather Review*, **115**, 1330-1338.
- Nutter, P. A. and Manobianco, J., 1999, Evaluation of the 29-km Eta model, Part I: objective verification at three selected stations, *Wea. Forecasting*, **14**, 5-17.
- Olson, D. A., Junker, N. W. and Korty, B., 1995, Evaluation of 33 years of quantitative precipitation forecasting at NMC, *Weather and Forecasting*, **10**, 498-511.
- Osborn, T. J. and Hulme, M., 1998, Evaluation of the European daily precipitation characteristics from the Atmosphere Model Intercomparison Project, *International Journal of Climatology*, **18**, 505-522.