



معاونت پژوهش و فناوری



جمهوری اسلامی ایران

شبکه بندی برون داد پس پردازش مدل WRF برای کمیت دمای دو متری

روی ایران

کد طرح

(۵۵-۲۲۰۵)

پژوهشکده توسعه تکنولوژی

مرتضی رحمانی

گروه پژوهشی مهندسی صنایع

آبان ۱۳۹۴

Archive of SID

روزگار را با خود

شناسنامه گزارش

عنوان گزارش: گزارش پایانی طرح پژوهشی

شماره ویرایش:

عنوان فارسی طرح پژوهشی: شبکه بندی برون داد پس پردازش مدل WRF برای کمیت دمای دو متری روی ایران

عنوان انگلیسی طرح:

Gridding of post processed wrf output for 2-meter temperature over Iran

کد طرح: ۲۲۰۵-۵۵

تالیف کننده‌ها: مرتضی رحمانی، سیده عاطفه محمدی، مجید آزادی، مرتضی خاکزار، محمد رضا معمار جعفری

تاریخ طرحنامه: ۹۲/۷/۱۵

تاریخ تصویب در شورای بررسی نهایی طرح‌ها: ۹۳/۶/۱۲

تاریخ شروع اجرای طرح: ۹۳/۷/۱

سطح دسترسی به سند: بدون موافقت کتبی، نسخه برداری یا تکثیر ممنوع است

کارکرد در طرح به ساعت	عناوین فعالیت‌هایی که در طرح انجام داده‌اند	تخصص	محل استخدام	نام مسئول یا همکار
۸۵۰	توسعه و طراحی مدل - شبکه بندی - درونیایی	ریاضی کاربردی	پژوهشکده توسعه تکنولوژی	مرتضی رحمانی
۷۵۰	کدنویسی - مرور ادبیات - پیاده‌سازی مدل - مقایسه و توسعه مدل‌ها	آمار ریاضی کاربردی	پژوهشکده توسعه تکنولوژی	سیده عاطفه محمدی
۳۰۰	مطالعه مدل‌ها - روش‌های راست‌آزمایی و اعتبارسنجی	مهندسی صنایع	پژوهشکده توسعه تکنولوژی	مرتضی خاکزار
۳۰۰	تحلیل نتایج عددی - مطالعه روش‌های پس-پردازش - مشاوره علمی، نرم افزاری و کدنویسی	هواشناسی	پژوهشگاه هواشناسی	مجید آزادی
۸۰	آزمون‌های آماری - مطالعات مربوط به آمار فضایی - پردازش‌های آماری	آمار ریاضی کاربردی	پژوهشکده توسعه تکنولوژی	محمد رضا معمار جعفری

سازمان مجری: پژوهشکده توسعه تکنولوژی

گروه مجری: گروه پژوهشی مهندسی صنایع

نشانی: تهران، خیابان آزادی، خیابان شهید تیموری، کوچه شهید صادقی، روبروی درب شمال غربی دانشگاه صنعتی

شریف، سازمان جهاد دانشگاهی شریف، پژوهشکده توسعه تکنولوژی

تلفن محل کار تالیف کننده اصلی: ۰۲۱۶۶۰۷۵۲۳۶

نشانی پست الکترونیک تالیف کننده اصلی: Rahmanimr@jdsharif.ac.ir

این سند به همه همکارانی که برای انجام کارهای پژوهشی در کشور بی‌دریغ کوشش می‌نمایند، تقدیم می‌شود.

Archive of SID

چکیده

در مدل های پیش بینی عددی وضع هوا، قوانین حاکم بر جو به صورت معادلات دیفرانسیل با مقدار اولیه - مرزی - مدل سازی شده اند. برای حل این معادلات دیفرانسیل از روش های عددی استفاده می شود. بدیهی است که روش های عددی، تقریبی بوده و همواره دارای خطا می باشند. استفاده از روش های عددی در مدل ها، تنها علت بروز خطا در خروجی مدل نمی باشد؛ بلکه به عواملی مانند وجود خطا در مقادیر اولیه - مرزی، کمبود داده های اولیه در برخی مناطق و آشوبی بودن سیستم دینامیکی جو نیز می توان اشاره کرد. بنابراین، پیش بینی در یک چنین سیستمی با عدم قطعیت همراه خواهد بود و دارای خطا می باشد.

خطای ذکر شده در برون داد مدل های پیش بینی عددی وضع هوا به دو دسته سامانمند و تصادفی تقسیم می شود. خطای سامانمند میزان میانگین انحراف برون داد مدل از واقعیت را نشان می دهد که با روش های آماری موسوم به پس پردازش قابل حذف شدن می باشد؛ اما خطای تصادفی از نظم خاصی پیروی نمی کند. روش های مختلفی برای از بین بردن خطای سامانمند یا پس پردازش برون داد خام مدل های پیش بینی عددی وجود دارد که مقادیر پیش بینی شده روی ایستگاه ها را تصحیح می کنند. در طرح حاضر با استفاده از داده های دیدبانی در ایستگاه ها، برای کمی سازی عدم قطعیت در جو، پیش بینی احتمالی و همچنین پیش بینی پس پردازش شده دمای دو متری به صورت قطعی صادر شده است.

در سال های اخیر نیاز به پیش بینی بدون اریبی در نقاط غیر ایستگاهی یا اساساً نقشه پس پردازش شده زیاد شده است لذا روش هایی ابداع شده اند که می توانند پیش بینی پس پردازش شده برای هر نقطه دلخواه را بدست آورند. در این طرح با استفاده از روش های تصحیحات متوالی، روش GIDS و روش کریگیدن که یکی از روش های آمار فضایی است، پیش - بینی های قطعی پس پردازش شده روی یک شبکه منظم روی ایران محاسبه و مقایسه شده است.

کلمات کلیدی:

پیش بینی عددی وضع هوا، پس پردازش، شبکه بندی، روش کریگیدن

فهرست

فصل اول: پیش بینی وضع هوا و معرفی مدل های عددی.....	۱
(۱-۱) مقدمه.....	۲
(۲-۱) تاریخچه.....	۲
(۱-۲-۱) پیش بینی باستانی.....	۲
(۲-۲-۱) پیش بینی مدرن.....	۳
(۳-۲-۱) تاریخچه پیش بینی عددی.....	۴
(۳-۱) مدل WRF.....	۶
(۱-۳-۱) سامانه پیش پردازشگر WRF (WPS).....	۹
(۲-۳-۱) آغازگری سامانه مدل سازی WRF (WRF ARW).....	۱۱
(۴-۱) مدل MM5.....	۱۲
(۱-۴-۱) برنامه های سامانه مدل سازی MM5.....	۱۳
فصل دوم: منابع عدم قطعیت در پیش بینی عددی وضع هوا و روش های پس پردازش روی	
ایستگاه های دیدبانی.....	۱۶
(۱-۲) مقدمه.....	۱۷
(۲-۲) خطا در پیش بینی عددی وضع هوا.....	۱۷
(۳-۲) منابع عدم قطعیت.....	۱۸
(۴-۲) پیش بینی قطعی و احتمالی.....	۱۹
(۵-۲) پس پردازش آماری خروجی مدل.....	۲۲
(۶-۲) حذف خطای سامانمند پیش بینی روی ایستگاه ها.....	۲۲
(۱-۶-۲) روش پیش بینی کامل.....	۲۳

۲۴ (۲-۶-۲) آماره برون داد مدل (MOS)
۲۵ (۳-۶-۲) روش کالمن فیلتر
۲۵ (۷-۲) پس پردازش سامانه همادی
۲۶ (۱-۷-۲) تخمین کمینه CRPS
۲۸ (۲-۷-۲) میانگین گیری به روش بیزی
۳۲ فصل سوم: روش های پس پردازش آماری برای نقاط شبکه منظم
۳۳ (۱-۳) مقدمه
۳۳ (۲-۳) پس پردازش آماری روی نقاط شبکه منظم
۳۳ (۳-۳) حذف اریبی سامانند شبکه
۳۷ (۴-۳) شبکه بندی MOS
۳۹ (۵-۳) مقایسه سه روش برای حذف اریبی شبکه
۴۱ (۶-۳) روش GIDS
۴۳ (۷-۳) آمار فضایی
۴۵ (۸-۳) کریگیدن
 فصل چهارم: معرفی ایستگاه ها و داده های هواشناسی و وضعیت پیش بینی در سازمان
۴۹ هواشناسی کشور
۵۰ (۱-۴) مقدمه
۵۰ (۲-۴) پیش بینی در سازمان هواشناسی کشور
۵۲ (۴-۴) زمان و ساعت دیده بانی
۵۳ (۵-۴) مراکز پیش بینی و اطلاعات جوی
۵۴ (۶-۴) آمار تحقیقات هواشناسی
۵۴ (۷-۴) داده های رقومی ارتفاعی زمین SRTM
۵۷ فصل پنجم: بحث و نتیجه گیری

۵۸	۱-۵) مقدمه
۵۸	۲-۵) الگوریتم حرکت جمعی ذرات (PSO)
۵۹	۱-۲-۵) تعریف عملکرد PSO
۶۰	۲-۲-۵) فرمول سرعت
۶۱	۳-۲-۵) روش های بروز رسانی وزن ایستایی
۶۱	۴-۲-۵) انواع PSO
۶۳	۳-۵) پیش بینی احتمالی دمای سطح زمین روی ایستگاه ها
۶۴	۱-۳-۵) تعریف مسأله و نتایج
۶۴	۲-۳-۵) تابع هدف
۶۴	۳-۳-۵) محدودیت
۶۴	۴-۳-۵) داده های مسأله
۶۵	۵-۳-۵) پارامترهای الگوریتم PSO
۶۶	۶-۳-۵) نتایج بدست آمده
۷۰	۴-۵) نحوه راست آزمایی
۷۱	۵-۵) نتایج حاصل از پس پردازش پیش بینی در نقاط شبکه
۷۴	۶-۵) نتیجه گیری
۷۶	فهرست مراجع
۸۱	پیوست ۱: کد برنامه ها



فصل اول: پیش بینی وضع هوا و معرفی مدل های

عددی

Archive of SID

۱-۱) مقدمه

در این فصل به تاریخچه پیش‌بینی وضع هوا پرداخته می‌شود و تحول زمانی و پیشرفت در پیش‌بینی وضع هوا نشان داده می‌شود. سپس سراغ پیش‌بینی عددی وضع هوا که انقلابی در پیش‌بینی بود رفته و در انتها دو مدل عددی متداول در دنیا معرفی می‌شود.

۱-۲) تاریخچه

از هزاران سال پیش به طور غیر رسمی و از قرن نوزدهم به طور رسمی، همواره بشر سعی داشته است تا وضع هوا را پیش‌بینی کند. وضع هوا با جمع‌آوری اطلاعات کمی درباره وضعیت فعلی جو و با استفاده از یک درک علمی از فرایندهای جوی، پیش‌بینی می‌شود.

پیش‌بینی وضع هوا کاربران مختلفی دارد. هشدارهای آب و هوایی مهمترین پیش‌بینی‌ها هستند زیرا از حیات و دارایی انسان‌ها محافظت می‌کنند. پیش‌بینی دما و بارش برای کشاورزان، بازرگانان و مدیران منابع آب و غیره بسیار مهم است. همه مردم می‌خواهند از وضع هوا باخبر باشند تا بتوانند نسبت به انجام فعالیت‌های پیشرو تصمیم درستی اتخاذ کنند. فعالیت‌ها در هوای آزاد به شدت به باران، برف و باد بستگی دارد و پیش‌بینی می‌تواند به برنامه‌ریزی درست کمک کند.

۱-۲-۱) پیش‌بینی باستانی

بابلان^۱ ۶۵۰ سال پیش از میلاد مسیح، وضع هوا را از روی مدل ابرها پیش‌بینی می‌کردند. ارسطو^۲ ۳۴۰ سال قبل از میلاد مسیح، طرح‌های آب و هوایی در هواشناسی را توصیف کرد. بعد از آن، تئوفراستوس^۳ یک کتاب در پیش‌بینی وضع هوا به نام کتاب علائم^۴ گردآوری کرد. پیش‌بینی وضع هوای چینی به ۳۰۰ سال قبل از میلاد برمی‌گردد؛ در همان زمان که منجمان هندی، روش‌های پیش‌بینی وضع هوا را توسعه دادند.

^۱ Babylonians
^۲ Aristotle
^۳ Theophrastus
^۴ Book of Signs

ابن واشیا^۱ ۹۰۴ سال بعد از میلاد، درباره پیش بینی وضع هوا از روی علائم تغییرات جو و دگرگونی های ستاره ها و سیارات بحث کرد که از آن جمله، پیش بینی بارش از روی مشاهدات فازهای قمری^۲ و پیش بینی وضع هوا بر اساس حرکت بادها است.

پیش بینی وضع هوای باستانی مبتنی بر الگوهای مشاهداتی بود؛ به عنوان مثال اگر مشاهده می شد که غروب خورشید قرمز است روز بعد هوا خوب خواهد بود. در طول نسل ها، این تجربیات روی هم انباشته شد تا دانش آب و هوایی به وجود آید. معتبر بودن این قبیل پیش بینی ها هنوز اثبات نشده است.

۱-۲-۲) پیش بینی مدرن

بعد از اختراع تلگراف در سال ۱۸۳۵ میلادی، دوره جدیدی از پیش بینی وضع هوا آغاز شد. پیش از آن، سریعترین حالتی که یک گزارش آب و هوایی چه زمینی و چه دریایی به فاصله دور برده می شد ۱۶۰ کیلومتر در روز بود. تلگراف در اواخر دهه ۱۸۴۰، امکان انتقال فوری گزارش شرایط آب و هوایی از یک منطقه وسیع را برقرار کرد و بستر پیش بینی با استفاده از دانش شرایط آب و هوایی را به وجود آورد.

دو فردی که به تولد پیش بینی به عنوان یک علم نسبت داده شده اند فرانسیس بیوفرت^۳ افسر نیروی دریایی سلطنتی^۴ و زیردستش ربرت فیتزروی^۵ هستند. هر دوی آنها افرادی با نفوذ در دولت و نیروی دریایی بریتانیا بودند و پایه گذار دانش پیش بینی وضع هوا در حال حاضر می باشند؛ با این وجود که در مطبوعات زمان خود مورد تمسخر قرار گرفتند.

بیوفرت جداول معتبر جذر و مد در اطراف ساحل بریتانیا را توسعه داد و با دوستش ویلیام وول^۶ به ضبط و نگهداری سوابق آب و هوایی ۲۰۰ ایستگاه گارد ساحلی بریتانیا پرداخت. فیتزروی در سال ۱۸۵۴ به عنوان رئیس دپارتمان جدید در وزارت بازرگانی منصوب شد تا به جمع آوری داده های آب و هوایی در دریا به عنوان خدمتی به دریانوردان بپردازد. این اقدام طلایه دار اداره هواشناسی مدرن امروزه می باشد. یک طوفان وحشتناک در سال ۱۸۵۹ که باعث خسارات زیادی

^۱ Ibn Wahshiyya
^۲ lunar phases
^۳ Francis Beaufort
^۴ Royal Navy
^۵ Robert FitzRoy
^۶ William Whewell

شده بود، فیتزروی را برانگیخت که نقشه‌ها را برای پیش‌بینی وضع هوا توسعه دهد. بر اساس انتشار گزارشات روزانه او از وضع آب و هوا به وسیله تلگراف‌های جدید، ۱۵ ایستگاه زمینی ساخته شد. اولین سرویس هشدار وی برای کشتیرانی با استفاده از ارتباطات تلگرافی در فوریه ۱۸۶۱ آغاز شد.

۱-۲-۳) تاریخچه پیش‌بینی عددی

تا اواخر قرن نوزدهم، پیش‌بینی وضع هوا امری ذهنی و بر اساس قوانین تجربی همراه با درک بسیار محدودی از مکانیزم فیزیکی فرایندهای جوی بود. در سال ۱۹۰۱ کلیولند اب^۱ مؤسس اداره هواشناسی ایالات متحده عنوان کرد که جو به وسیله اصول و قواعد ترمودینامیک و هیدرودینامیکی که در قرن پیش مورد مطالعه واقع شده بود تبعیت می‌کند.

در سال ۱۹۰۴ ویلهلم برکنس^۲ یک روند دو مرحله‌ای برای پیش‌بینی مبتنی بر مدل وضع هوا ارائه کرد. ابتدا یک مرحله تشخیصی^۳ برای پردازش داده‌ها به منظور تولید شرایط اولیه استفاده شد و سپس در مرحله پیش‌بینی مسأله مقدار اولیه حل می‌شود. او همچنین هفت متغیر را که وضعیت جو در یک نقطه را تعیین می‌کنند شناسایی کرد: فشار، دما، چگالی، رطوبت و سه مؤلفه بردار سرعت. برکنس اشاره کرد که از معادلات مبتنی بر پیوستگی جرم^۴، پایستگی تکانه^۵، قوانین اول و دوم ترمودینامیک و قانون گاز ایده‌آل^۶ برای تخمین وضعیت جو در آینده از طریق روش‌های عددی استفاده خواهد شد. قوانین ذکر شده به غیر از قانون دوم ترمودینامیک، اساس و پایه معادلات به کار رفته در مدل‌های آب و هوایی امروزه را تشکیل می‌دهند.

در سال ۱۹۲۲ لوئیس فرای ریچاردسون^۷، اولین گام را در پیش‌بینی وضع هوا از طریق روش‌های عددی برداشت. با استفاده از تغییر هیدرواستاتیکی معادلات اولیه توسط برکنس، ریچاردسون یک پیش‌بینی شش ساعته در دو نقطه اروپای مرکزی با دست ارائه کرد که حداقل شش هفته به طول انجامید. اما به دلیل عدم تعادل یا هماهنگی بین فشار و سرعت باد در شرایط اولیه، خطایی بزرگ در نتایج بوجود آمد.

^۱ Cleveland Abbe

^۲ Vilhelm Bjerknes

^۳ diagnostic step

^۴ mass continuity

^۵ conservation of momentum

^۶ ideal gas law

^۷ Lewis Fry Richardson

اولین پیش‌بین عددی موفق به وسیله کامپیوتر دیجیتالی انیاک^۱ در سال ۱۹۵۰ توسط تیمی متشکل از هواشناسان آمریکایی چارنی^۲، تامسون^۳، گیتس^۴ و هواشناس نروژی به نام فرترف^۵ و ریاضیدانی به نام وون نیومن^۶ اجرا شد. آنها از فرم ساده شده دینامیک جو بر اساس حل معادله گردابی فشاروردی^۷ (در این معادله فرض شده است که سرعت و جهت باد، مستقل از ارتفاع است) در یک لایه جو با استفاده از محاسبه ارتفاع ژئوپتانسیل جو در لایه فشاری ۵۰۰ میلی بار استفاده کردند. این ساده سازی‌ها، میزان حافظه و مدت زمان محاسباتی کامپیوتر را کاهش داد تا در کامپیوترهای اولیه قابل اجرا باشد.

هنگامی که در سال ۱۹۵۰ خبر اولین پیش‌بینی با انیاک به ریچاردسون رسید وی اظهار داشت که این نتایج یک پیشرفت بزرگ علمی است. اولین محاسبات برای یک پیش‌بینی ۲۴ ساعته با انیاک تقریباً ۲۴ ساعت طول کشید. اما گروه چارنی ذکر کردند که بیشترین زمان صرف عملیات دستی شده است و اظهار امیدواری کردند که پیش‌بینی وضع هوا قبل از رخداد آن به زودی میسر می‌شود.

در سپتامبر ۱۹۵۴، گروه کارل گوستاو راسبی^۸ در مؤسسه هیدرولوژی و هواشناسی سوئیس اولین پیش‌بینی عملیاتی را صادر کردند. در سال ۱۹۵۵ پیش‌بینی وضع هوا به طور عملیاتی در ایالات متحده در یک پروژه مشترک بین نیروی هوایی، نیروی دریایی و اداره آب و هوایی زیر نظر واحد مشترک پیش‌بینی عددی وضع هوا^۹ انجام شد. آژانس هواشناسی ژاپن سومین سازمان پیش‌بینی وضع هوا در سال ۱۹۵۹ بود. از مهمترین مراکز هواشناسی در سراسر جهان می‌توان به مراکز ملی برای پیش‌بینی محیطی (NCEP)^{۱۰}، اداره ملی آب و هوا^{۱۱} (NWS)، اداره ملی جو و اقیانوس^{۱۲} (NOAA) و مرکز اروپایی برای پیش‌بینی میان مدت^۱ (ECMWF) اشاره کرد. با این اوصاف، پیش‌بینی عددی وضع هوا بزرگترین دستاورد قرن بیستم می‌باشد.

^۱ ENIAC

^۲ Jule Charney

^۳ Philip Thompson

^۴ Larry Gates

^۵ Ragnar Fjørtoft

^۶ John von Neumann

^۷ barotropic vorticity equation

^۸ Carl-Gustav Rossby

^۹ Joint Numerical Weather Prediction Unit

^{۱۰} National Centers for Environmental Prediction (NCEP)

^{۱۱} National Weather Service

^{۱۲} National Oceanic and Atmospheric Administration

مدل های عددی پیش بینی وضع هوا شامل معادلات دیفرانسیلی هستند که قوانین فیزیکی و دینامیکی جو را توصیف می کنند. این معادلات دیفرانسیل وابسته به مقادیر اولیه- مرزی هستند و به علت پیچیدگی بالای آنها با استفاده از روش های عددی حل می شوند. در حقیقت منطق پیش بینی وضع هوا به این صورت است که با داشتن اطلاعاتی از وضع فعلی جو و با توجه به قوانین فیزیکی و دینامیکی حاکم بر جو می توان وضع هوای آینده را پیش بینی نمود. از پر کاربردترین مدل ها می توان به مدل WRF^۲ و MM5^۳ اشاره نمود که در بخش های آینده درباره آنها توضیحاتی داده می شود.

۱-۳) مدل WRF

مدل WRF یک سامانه پیش بینی عددی وضع هوا (NWP) و شبیه سازی جو است که برای کاربردهای پژوهشی و عملیاتی طراحی شده است. توسعه و گسترش این مدل تبدیل به یک تلاش همگانی برای ایجاد یک مدل پیش بینی میان مقیاس و نیز ایجاد یک سامانه داده گزاری شده است تا فهم و پیش بینی وضعیت جوی را گسترش داده و نیز پیشرفت های تحقیقاتی را تسریع کند. طراحی مدل WRF نتیجه تلاش های مشترک نهادهایی در کشور آمریکا است.

مدل WRF سامانه ای با قابلیت های فراوان است که می توان وضعیت های مختلف جو را با استفاده از آن شبیه سازی نمود. مدل مذکور قابل نصب بر روی سامانه های مختلف و رایانه ای شامل خوشه رایانه می باشد. محدوده کاربرد آن بسیار وسیع است و گستره ای بین چندین متر تا هزاران کیلومتر را پوشش می دهد.

مدل WRF به صورت یک کد قابل انعطاف طراحی شده است به گونه ای که روی طیف گسترده ای از محیط های محاسبه ای از ابررایانه های موازی تا رایانه های رومیزی و خانگی قابل اجرا است. می توان کدهای تک منبعی و مطلق این مدل را به صورت منفرد برای کاربردهای پژوهشی و عملیاتی پیکربندی کرد. گزینه های فیزیکی و دینامیکی آن حاصل تجربه و تلاش گروه علمی گسترده ای است. در حال حاضر می توان از آن به منظور پژوهش و آموزش در زمینه های زیر استفاده کرد:

- شبیه سازی های یک حالت آرمانی (با در نظر گرفتن همرفت و امواج کژفشار)،

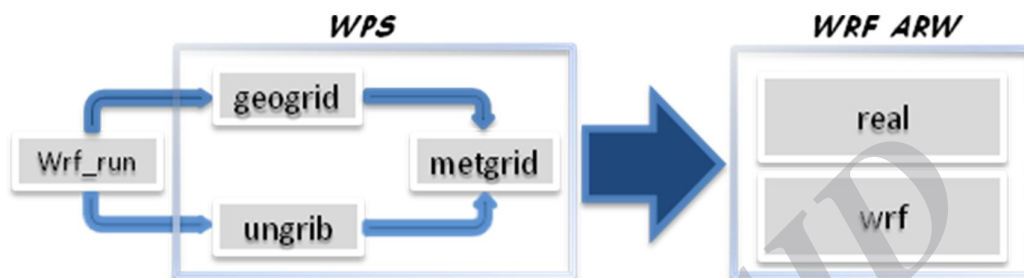
^۱ European center for medium weather forecasting

^۲ Weather Research and Forecasting

^۳ Fifth-Generation Penn State/NCAR Mesoscale Model

- مطالعه فیزیک مدل از طریق تغییر فراسنج سازی های فیزیکی،
 - پژوهش در زمینه داده گواری،
 - پژوهش در زمینه پیش بینی عددی وضع هوا،
 - پیش بینی عددی وضع هوا با استفاده از داده های زمان واقعی،
 - استفاده در مدل های جفت شده (ترکیب مدل های اقیانوسی و جوی)،
 - استفاده هم در محدوده وسیع جهانی و هم به صورت محدود در مقیاس منطقه ای،
 - استفاده برای کاربردهای آموزشی .
- بخشی از NCAR که در زمینه هواشناسی کوچک مقیاس و میان مقیاس فعال است، مسئولیت پشتیبانی مجموعه برنامه های منبع سامانه WRF را به شرح زیر برعهده دارد:
- چارچوب نرم افزاری WRF (WSF)،
 - حل کننده مسائل دینامیکی مدل پیشرفته تحقیقاتی WRF (ARW)، که شامل آشیانه سازی یک سویه و دوسویه و آشیانه سازی متحرک می شود؛
 - سامانه پیش پردازشگر WRF (WPS)،
 - سامانه داده گواری مدل (WRF-Var)،
 - توزیع بسته های فیزیکی متعدد توسط شرکای WRF و پژوهشگران،
 - چندین برنامه گرافیکی و برنامه تبدیل برای ابزارهای گرافیکی .
- مدل WRF یک مدل غیرهیدروستاتیک (با گزینه اختیاری هیدروستاتیک در حین اجرا) است که شبکه بندی افقی آن شبکه C-آراکوا می باشد. در این مدل از طرحواره های انتگرال گیری مرتبه دوم و سوم زمانی رانگ-کوتا و طرحواره های

مرتبہ دو تا شش برای فرارفت در دو راستای افقی و قائم استفاده می شود. این مدل از گام های زمانی کوچک برای مد- های اکوستیک و امواج گرانشی استفاده می کند. نمودار کلی اجرای مدل WRF در شکل ۱-۱ نشان داده شده است.



شکل ۱-۱: نمایی از اجرای مدل WRF

نسخه ی دوم مدل فوق دارای قابلیت های متعددی است که از آن جمله می توان به موارد زیر اشاره کرد:

- شبیه سازی ایده آل مدل و نیز شبیه سازی با داده های واقعی
- گزینه های مختلف برای مرزهای جانبی به منظور شبیه سازی مدل های ایده آل و نیز شبیه سازی با داده های واقعی
- انتخاب فیزیک مدل
- قابلیت آشیانه سازی یک و دو طرفه و نیز آشیانه ی متحرک
- دارای گزینه های هیدرواستاتیک و غیر هیدرواستاتیک (گزینه ی زمان اجرا)^۱
- قابلیت اجرا در گستره ی بین چندین متر تا هزاران کیلومتر

^۱ - runtime

۱-۳-۱) سامانه پیش پردازشگر WRF (WPS)

سامانه پیش پردازشگر WRF، مجموعه‌ای از سه برنامه است که نقش مشترکشان آماده‌سازی درون‌داد برنامه real برای شبیه‌سازی داده‌های واقعی می‌باشد. هر یک از این برنامه‌ها یک مرحله از آماده‌سازی را انجام می‌دهند:

○ برنامه geogrid، حوزه‌های مدل را تعریف کرده و داده‌های ثابت جغرافیایی (ناهمواری و کاربری اراضی) را وارد شبکه می‌کند.

○ برنامه ungrib، میدان‌های جوی حاصل از پیش‌بینی‌های یک مدل تمام کره‌ای را از فایل‌هایی با قالب GRIB کدگشایی و استخراج می‌کند.

○ وظیفه برنامه metgrid، درون‌یابی افقی میدان‌های استخراج شده در مرحله قبل بوسیله برنامه ungrib، روی شبکه‌های تعریف شده توسط geogrid است، درون‌یابی قائم این میدان‌ها روی ترازهای WRFeta، توسط برنامه real انجام می‌شود.

۱-۳-۱-۱) برنامه Geogrid

هدف برنامه geogrid معین کردن حوزه‌های شبیه‌سازی شده، و وارد کردن مجموعه داده‌های زمینی گوناگون روی شبکه‌های مدل است. حوزه‌های شبیه‌سازی شده با استفاده از اطلاعات معین شده توسط کاربر در بخش "geogrid" فایل namelist.wps تعریف می‌شوند. همچنین برای محاسبه‌ی طول، عرض و عامل‌های مقیاس نقشه در هر نقطه شبکه، برنامه geogrid، انواع خاک، نوع کاربری زمین، ارتفاع منطقه، میانگین سالیانه دمای خاک در عمق‌های متفاوت، درصد پوشش گیاهی ماهانه، بیشترین مقدار آلودگی برف، و نوع شیب را برای شبکه‌های مدل به صورت پیش فرض وارد خواهد کرد. مجموعه داده‌های جدید از طریق فایل GEOGRID.TBL به حوزه شبیه‌سازی وارد می‌شود؛ که این فایل همه میدان‌هایی که توسط geogrid تولید خواهند شد را تعیین می‌کند. همچنین این فایل روش‌های درون‌یابی کردن را که برای یک میدان استفاده می‌شود، مشخص می‌کند.

برون‌داد این برنامه، فایل‌هایی با نام geo_em.d0*.nc می‌باشد. هر یک از این فایل‌ها متعلق به یک حوزه است، که * شماره حوزه فرض می‌شود. برون‌دادها به شکل NetCDF I/O نوشته می‌شوند.

۱-۳-۱) برنامه Ungrib

برنامه ungrib فایل‌های GRIB را خوانده و در یک قالب استاندارد که قالب میانه^۱ نام دارد، می‌نویسد. فایل‌های GRIB میدان‌های هواسنجی را که با زمان تغییر می‌کنند، شامل می‌شوند و اغلب برون‌داد دیگر مدل‌های ناحیه‌ای یا جهانی مانند GFS یا NCEP هستند. برنامه ungrib می‌تواند GRIB Edition 1 و اگر با گزینه "GRIB2" کامپایل شود، GRIB Edition 2 را بخواند.

پس از اجرای این برنامه فایل‌هایی با نام "FILE:yyyy-mm-dd_hh" تولید می‌گردند که yyyy-mm-dd_hh معرف تاریخ و ساعت است، همان طور که از نام فایل پیداست هر یک از این فایل‌ها به تاریخ مشخصی ختم می‌شوند که ساختار و توالی این تاریخ‌ها را می‌توان در namelist.wps تنظیم کرد.

۱-۳-۱) برنامه Metgrid

برنامه metgrid داده‌های هواسنجی قالب میانی را که توسط برنامه ungrib استخراج می‌شوند، به روی حوزه‌های شبیه‌سازی تعریف شده توسط برنامه geogrid به صورت افقی درون‌یابی می‌کند. حوزه داده‌هایی که توسط metgrid درون‌یابی می‌شوند در بخشی به نام "share" در namelist.wps تعیین شده است، و حوزه داده‌ها باید به طور جداگانه در این فایل برای هر حوزه شبیه‌سازی شده معین شود. از آن‌جا که کار این برنامه (به مانند برنامه ungrib) وابسته به زمان است، هر زمان که یک شبیه‌سازی جدید آغاز می‌شود، اجرا می‌گردد.

بطور کلی برنامه metgrid برون‌داد دو برنامه دیگر را به عنوان ورودی دریافت کرده و با درون‌یابی افقی از میدان‌ها و انتقال کمیت‌ها بر روی شبکه شطرنجی^۲، فایل‌هایی به نام met_em.d0*_yyyy-mm-dd_hh.nc تولید می‌کند که * معرف شماره حوزه و yyyy-mm-dd_hh معرف تاریخ و ساعت است. این فایل‌ها برای هر حوزه و هر ساعت، جداگانه تولید می‌شوند. برون‌دادها به شکل NetCDF I/O نوشته می‌شوند.

^۱-Intermediate Format
^۲- Staggered

۱-۳-۲) آغازگری سامانه مدل سازی WRF (WRF ARW)

مدل WRF دو کلاس بزرگ شبیه سازی دارد که می تواند هم شبیه سازی هایی با داده های ایده آل و هم شبیه سازی -هایی با داده های واقعی را تولید کند. شبیه سازی های ایده آل معمولاً یک فایل با شرایط اولیه برای مدل WRF تولید می کنند که ناشی از گمانه زنی یک بعدی یا دو بعدی است و در آنها توپوگرافی به صورت ساده فرض می شود. در مورد به کارگیری داده های واقعی به یک پیش پردازش از بسته WPS نیاز داریم، که همه میدان های جوی و ثابت را به نحو مقتضی ایجاد می کند.

برنامه آغازگری در موارد ایده آل، ideal.exe و در موارد داده واقعی real.exe نام دارد. این دو برنامه هرگز باهم استفاده نمی شوند و هر دو برنامه پیش از اجرای مدل WRF پردازش می شوند. مدل WRF برای هر دو حالت یکسان عمل می کند، یعنی مدل برای اجرای هر یک از این دو برنامه، مجموعه فعالیت های مشابهی را دنبال می کند. مدل پیشرفته WRF شامل برنامه مقدماتی real.exe و برنامه انتگرال گیری عددی wrf.exe می باشد.

۱-۳-۲-۱) برنامه Real.exe

برنامه "real.exe" فایل های met_em.d0*.yyyy-mm-dd_hh.nc را که برون داد برنامه "metgrid.exe" در WPS هستند، به عنوان داده های ورودی دریافت می کند. با اجرای این برنامه داده ها از مختصات فشاری به مختصات سیگما تغییر می یابد. علاوه بر این در زمان های بعد نقاط مرز را نیز از داده ها جدا می کند. برون داد این برنامه در فایل هایی به نام wrfinput_d0* ذخیره می شود که * معرف شماره حوزه است.

۱-۳-۲-۲) برنامه WRF.exe

برنامه "wrf.exe" فایل های wrfinput_d0* را که برون داد "real.exe" است به عنوان داده های ورودی دریافت می کند. با اجرای این برنامه پیش بینی برای روزهای آینده (تعداد روز و تاریخ در namelist در ابتدا مشخص می شود) تولید می شود. برون داد این برنامه در فایل هایی به نام wrfout_d0*_ yyyy-mm-dd_hh * معرف شماره حوزه و yyyy-mm-dd_hh معرف تاریخ و ساعت مشخص است.

برای بدست آوردن اطلاعات بیشتر در رابطه با مدل WRF می توان به سایت اینترنتی

http://www.mmm.ucar.edu/wrf/users/docs/user_guide_V3/contents.html

و http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf مراجعه کرد.

۱-۴) مدل MM5

سامانه مدل سازی MM5 شامل چند برنامه کمکی و یک مدل منطقه محدود میان مقیاس است که قبل از دهه ۱۹۷۰ بوسیله آنتس^۱ در دانشگاه ایالتی پنسلوانیا^۲ توسعه داده شد. بعدها در سال ۱۹۸۷ بوسیله آنتس و وارنر^۳ کامل تر شد و از آن زمان تاکنون تغییرات زیادی در برنامه های تشکیل دهنده آن داده شده است که در نتیجه کاربردهای مدل نیز تفاوت زیادی پیدا کرده است. از جمله این موارد می توان به توانایی داشتن چند حوزه، استفاده از دینامیک غیر ایستایی (هیدرواستاتیک) و توانایی داده گواری چهار بعدی داده ها اشاره کرد. علاوه بر این گزینه های مربوط به بخش فیزیک مدل افزایش یافته است. همچنین امروزه گستره وسیعی از سامانه های رایانه ای شامل رایانه های شخصی با سامانه عامل لینوکس قادر به اجرای این مدل هستند.

این مدل، داده ها را از یک شبکه طول و عرض جغرافیایی بعنوان ورودی دریافت و بر روی یکی از نگاشت های قطبی، مخروطی لامبرت و یا مرکاتور درون یابی می کند (برنامه های TERRAIN و REGRID). این داده ها را می توان با داده های دیدبانی از ایستگاه های همدیدی یا جو بالا بهبود بخشید (برنامه های RAWINS/little-r). سپس درون یابی قائم از مختصات فشاری به مختصات سیگمای مدل انجام می گیرد (برنامه INTERPF).

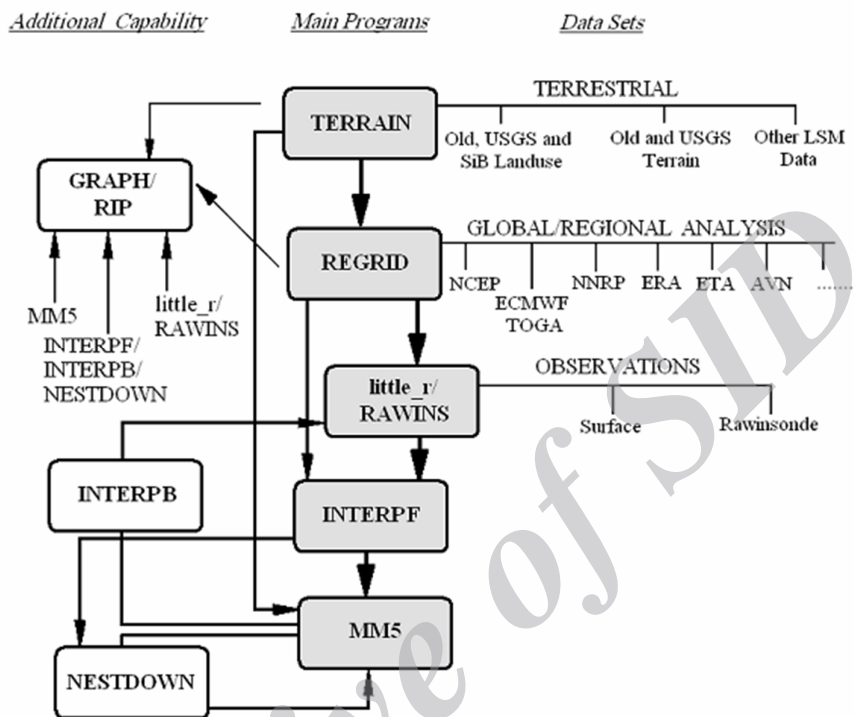
ویژگی سطح سیگما آن است که در نزدیکی سطح زمین تقریباً از عوارض زمینی و در سطوح بالاتر از سطوح فشاری پیروی می کند. چون MM5 یک مدل منطقه محدود است، پس جهت اجرا به شرایط اولیه و شرایط مرزی جانبی نیاز دارد. برای شرایط مرزی جانبی، به داده های هواشناسی بر روی نقاط جانبی شبکه منظم در طی بازه زمانی اجرا نیاز

1- Anthes

2- Pensilvania

3- Warner

است. به منظور روشن شدن موضوع در ذیل به ساختار (شکل ۱-۲) و وظیفه هر یک از برنامه های سامانه مدل سازی MM5 اشاره می شود.



شکل ۱-۲: نمایی از اجرای مدل MM5

۱-۴-۱) برنامه های سامانه مدل سازی MM5

۱-۱-۴-۱) برنامه Terrain

نخستین برنامه ای که جهت انجام هر پیش بینی یا شبیه سازی کامل در سامانه مدل سازی MM5 به اجرا در می آید، برنامه Terrain است. این برنامه طول و عرض جغرافیایی، ارتفاع عوارض زمینی و رشد و نمو گیاهان (کاربری زمین) را برای حوزه های منتخب میان مقیاس درون یابی می کند.

وظایف برنامه Terrain شامل چهار بخش زیر است:

۱- تعیین منبع داده های ورودی

۲- درون یابی داده های منبع از طول و عرض جغرافیایی به شبکه میان مقیاس

۳- تنظیم اثر متقابل حوزه و بازخورد

۴- فراهم کردن برون داد ارتفاع عوارض زمینی.

۱-۱-۲) برنامه Regrid

برنامه Regrid تحلیل های هواشناختی بایگانی شده را از ترازهای فشاری می خواند و آنها را از شبکه محلی و تصویر نقشه به شبکه تصویر نقشه تعریف شده توسط برنامه پیش پردازنده Terrain در سامانه مدل سازی MM5 درون یابی می کند. در روی این ترازها درون یابی دو بعدی انجام می گیرد. در نمودار کلی سامانه مدل سازی MM5، برنامه Regrid دومین مرحله است، این برنامه ورودی خود را از برنامه Terrain می گیرد. برون داد برنامه فایل هایی است که به عنوان ورودی برنامه Interpf به کار می رود. این فایل ها عموماً به عنوان حدس اولیه برنامه انجام تحلیل عینی یا به عنوان تحلیل هایی که باید به طور مستقیم به ترازهای مدل MM5 جهت شرایط مرزی یا جانبی درون یابی شوند، مورد استفاده قرار می گیرند.

این برنامه شامل دو بخش Pregrid و Regridder می باشد:

الف) برنامه Pregrid

برنامه pregrid می تواند بسیاری از داده های با فرمت GRIB و مخلوط GRIB و غیر GRIB را که به طور معمول در دسترس کاربران مدل MM5 قرار دارد، بخواند. بیشتر برنامه مستقل pregrid، به ویژه آن دسته که با داده های GRIB مرتبط هستند، به جدول هایی نیاز دارند که مشخص می کند برنامه pregrid از طریق فایل های ورودی به کدام میدان ها دست یابی دارد.

ب) برنامه Regridder

برنامه regridder فایل های میدان های دما، مولفه های باد افقی، رطوبت نسبی، ارتفاع سطوح فشاری، فشار سطح دریا، دمای سطح دریا و داده های پوشش برف را به عنوان ورودی از برون داد pregrid دریافت و جهت استفاده بهینه سامانه، درون یابی می کند.

۱-۴-۱) برنامه Interpf

این برنامه وظیفه ارسال داده های مورد نیاز را از برنامه تحلیل به مدل میان مقیاس به عهده دارد که شامل درون یابی قائم، محاسبات فرایابی و قالب بندی مجدد داده هاست. Interpf ورودی خود را از برون داد برنامه Regrid می گیرد و سپس آغازگری مدل، شرایط مرزی جانبی و شرایط مرزی پایین را تولید می کند.

۱-۴-۱) برنامه MM5

این برنامه بخش اصلی سامانه و همان مدل میان مقیاس و منطقه محدود MM5 است. این بخش، انتگرال گیری زمانی مدل را انجام می دهد. این مدل میان مقیاس می تواند برای طیف گسترده ای از مطالعات نظری و زمانی واقعی شامل شبیه سازی پیش بینی و داده گواری چهار بعدی برای مونسون، هاریکن و چرخند به کار رود. همچنین برای مطالعاتی نظیر سامانه های همرفتی میان مقیاس، جبهه و سیستم دریا، چرخش های کوهستانی-دره ای و جزایر گرمایی شهری نیز به کار می رود.

۱-۴-۱) برنامه Nestdown

تولید شبکه ریز ورودی مدل از شبکه بزرگ تر برون داد مدل، توسط این برنامه انجام می شود.

برای بدست آوردن اطلاعات بیشتر در رابطه با مدل MM5 می توان به سایت اینترنتی

<http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes-pdf/mm5.pdf> مراجعه کرد.

فصل دوم: منابع عدم قطعیت در پیش‌بینی عددی
وضع هوا و روش‌های پس‌پردازش روی ایستگاه‌های

دیدبانی

Archive of SID

۲-۱) مقدمه

در فصل حاضر، علل وجود خطا و همچنین منابع عدم قطعیت در پیش بینی عددی وضع هوا بیان می شود. روش های آماری برای حذف آریبی که موسوم به روش های پس پردازش می باشند، توضیح داده می شوند. این روش ها به منظور از بین بردن خطای پیش بینی در ایستگاه های دیدبانی از داده های مشاهداتی استفاده می کنند. در این فصل به هر دو روش های پس پردازش قطعی و احتمالی روی ایستگاه های دیدبانی پرداخته شده است.

۲-۲) خطا در پیش بینی عددی وضع هوا

در پیش بینی وضع هوا با استفاده از مدل های عددی، با خطای خروجی مواجه هستیم که از دلایل این امر می توان به موارد زیر اشاره نمود:

- معادلات مدل نمی توانند همه واقعیت های جو را در نظر بگیرند؛ زیرا برخی رفتارها غیر خطی هستند که بیان این رفتارها در غالب معادلات غیر خطی و حل آنها حتی با استفاده از ابررایانه های سریع، در یک زمان منطقی امکان پذیر نخواهد بود.
- مدل نمی تواند فرایندهای جوی را که کمتر از یک فاصله مشخص باشند، در نظر بگیرد؛ لذا بعضی پدیده های اقلیمی محدوده مورد پیش بینی نادیده گرفته می شود.
- مدل های پیش بینی از نوعی هستند که باید مقادیر اولیه - مرزی داشته باشند تا حل و اجرا شوند. از آنجایی که مقادیر اولیه دارای کیفیت و دقت کافی نیستند، خروجی مدل نیز مقداری دقیق نخواهد بود.
- کمبود داده های اولیه در بعضی مناطق مانند دریا و نقاط صعب العبور باعث بروز خطا در پیش بینی می گردد.

خطای ذکر شده در برون داد مدل های پیش بینی عددی وضع هوا به دو دسته سامانمند و تصادفی تقسیم می شود. خطای سامانمند میزان میانگین انحراف برون داد مدل از واقعیت را نشان می دهد که به آن اریبی^۱ گفته می شود و با روش های آماری موسوم به پس پردازش قابل حذف شدن می باشد؛ اما خطای تصادفی از نظم خاصی پیروی نمی کند و تصادفی است. روش های مختلفی برای از بین بردن خطای سامانمند یا پس پردازش برون داد خام مدل های پیش بینی عددی وجود دارد اما همه این روش ها مقادیر پیش بینی شده روی ایستگاه ها را تصحیح می کنند. در سال های اخیر نیاز به پیش بینی بدون اریبی در نقاط غیر ایستگاهی یا اساساً نقشه پس پردازش شده زیاد شده است پس روش هایی مانند روش تصحیحات متوالی^۲ (Sheets, 2009 & Glahn, Gilbert, Cosgrove, Ruth) ابداع شده اند که می توانند پیش بینی پس پردازش شده برای هر نقطه دلخواه را بدست آورند.

۲-۳) منابع عدم قطعیت

در علم جو از تنها موردی که اطمینان کامل داریم، عدم قطعیت در پیش بینی های وضع هوا و اقلیمی است. به دو دلیل نمی توان از مدل های عددی پیش بینی وضع هوا، انتظار یک پیش بینی همراه با قطعیت را داشت:

۱- اگر چه مدل های عددی موجود تقریب خوبی از رفتار جو هستند، اما فیزیک جو را به درستی توصیف نمی کنند. دلیل مهم و اجتناب ناپذیر این مشکل، عدم توانایی مدل در لحاظ کردن برخی فرایندهای فیزیکی در مقیاس واقعی آنها می باشد. برخی فرایندهای فیزیکی جو روی مقیاس های کوچک عمل می کنند که به طور واضح روی این مدل ها قابل نشان دادن نیستند و فقط در مقیاس های بزرگتر تقریب زده می شوند. از طرفی دیگر اگر فرایندهای جوی، پایدار و یا متناوب می بودند، توصیف آنها به صورت ریاضی ساده می بود. اما با توجه به اینکه جو دارای نوسانات غیرمنظم می باشد، مدلسازی ریاضی هرگز نمی تواند نمایانگر رفتار واقعی آن باشد و فقط به طور تقریبی می تواند آنچه که واقعا اتفاق می افتد را توصیف کند.

۲- حتی اگر بتوان مشکل ۱ را به طریقی بر طرف کرد که فیزیک جو به درستی در مدل های عددی بیان شود، اما نمی توان عدم قطعیت تولید شده از آشوب دینامیکی جو را نادیده گرفت. محققى به نام لورنز^۳ که پایه گذار نظریه آشوب می باشد، برای اولین بار به آشوبناک بودن جو پی برد. حساسیت به مقادیر اولیه در برخی

^۱ Bias

^۲ Successive Correction Method (SCM)

^۳ Lorenz

سیستم‌های دینامیکی، تعریفی ساده و کلی از سیستم‌های آشوبی می‌باشد. در یک چنین سیستم‌های آشوبی که جو نیز یک از آنهاست، ممکن است به ازای مقادیر اولیه مختلف که تفاوت ناچیزی با هم دارند خروجی‌هایی با تفاوت چشمگیر حاصل شود.

از آنجا که مقدار پارامترهای جوی به دلیل خطاهای اندازه‌گیری به طور کامل قابل مشاهده نیستند، مقادیر اولیه‌ای که به عنوان ورودی به مدل‌ها داده می‌شوند با شرایط واقعی جو تفاوت دارند. بنابراین فرض کنید که مدل‌های عددی بتوانند فیزیک جو را به طور کامل مدلسازی کنند، اما به دلیل آشوبی بودن سیستم دینامیکی جو و مقادیر اولیه متفاوت با شرایط واقعی، پیش‌بینی دقیق و قطعی وضع هوا غیر ممکن خواهد بود. به طوری که زنگ^۱ در این باره می‌گوید: "همان گونه که نظریه نسبیت تصور نیوتن درباره زمان-مکان مطلق و یا نظریه کوانتوم رژیای انیشتین و نیوتن را درباره فرایند اندازه‌گیری قابل کنترل از بین برد، نظریه آشوب نیز تصور پیش‌بینی قطعی طولانی مدت وضع هوا را از بین برد."

واقعیت این است که حتی با وجود سیستم‌های داده‌گویی داده‌ها^۲ (کاهش خطای مقدار داده‌های اولیه)، دسترسی به شرایط اولیه کامل محال است. بنابراین در نظر گرفتن صرفاً یک مقدار اولیه به عنوان ورودی مدل‌های عددی پیش‌بینی وضع هوا و طبیعتاً به دست آوردن یک خروجی به عنوان مقدار پیش‌بینی، به معنی چشم‌پوشی از عدم قطعیت در سیستم جو می‌باشد. بنابراین هم دینامیک آشوبی و هم خطای اجتناب-ناپذیر در مدلسازی ریاضی جو، بر این نکته تأکید می‌کند که وضع هوا را نمی‌توان به طور قطعی پیش‌بینی نمود.

۲-۴) پیش‌بینی قطعی^۳ و احتمالی^۴

برای تشخیص چگونگی وضع هوای آینده، علاوه بر پیش‌بینی قطعی رویکرد دیگری تحت عنوان پیش‌بینی احتمالی نیز وجود دارد. ایده پیش‌بینی احتمالی از عدم قطعیت موجود در جو نشأت گرفته است.

^۱ Zeng
^۲ Data assimilation
^۳ Deterministic
^۴ Probabilistic

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، در مواجهه با شرایط اولیه یکسان ممکن است وضعیت هوای آینده متفاوت باشد.

پیش‌بینی احتمالی سعی دارد تا با بیان احتمالی از شانس وقایع ممکن آینده، عدم قطعیت را کمی سازد. بدین طریق که به جای در نظر گرفتن فقط یک مقدار اولیه، از یک توزیع احتمال که عدم قطعیت وضعیت اولیه را توصیف می‌کند نمونه‌گیری می‌شود که نتیجتاً پیش‌بینی یک توزیع احتمال خواهد بود که عدم قطعیت وضعیت آینده جو را توصیف می‌کند. پیش‌بینی احتمالی 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.

بحث پیش‌بینی احتمالی با اثبات آشوبناکی سیستم دینامیکی جو توسط لورنز (Lorenz, 1963) 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 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 احتمالی 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 بجای در نظر گرفتن فقط یک مقدار اولیه می‌توان با نمونه‌گیری متناهی از تابع چگالی احتمال عدم قطعیت شرایط اولیه و اجرای کامل مدل عددی با توجه به شرایط هر نمونه، یک سامانه همادی با تعداد اعضای متناهی تولید کرد. با اعمال روش‌های آماری روی اعضای سامانه همادی یک تابع توزیع احتمال بدست خواهد آمد که عدم قطعیت وضعیت آینده جو را توصیف می‌کند و شامل اطلاعات کافی برای نیازهای کاربران مختلف

می‌باشد. از دلایل برتری پیش بینی احتمالی به پیش بینی قطعی می‌توان به موارد زیر اشاره نمود (Estephenson, 2007):

- با پیش بینی قطعی نمی‌توان اطلاعات کافی از وضعیت آینده جو بدست آورد اما در پیش بینی احتمالی که وضعیت آینده جو با یک تابع توزیع احتمال مشخص می‌شود می‌توان اطلاعات کاملتری کسب نمود.
- میانگین یک پیش بینی همادی، بسیار دقیق تر از یک پیش بینی منفرد می‌باشد.
- پیش بینی احتمالی، به کاربران اجازه اخذ تصمیم بهینه را می‌دهد؛ در حالی که پیش بینی قطعی تصمیمی می‌باشد که قبلاً توسط پیش بین گرفته شده است.
- پیش بینی احتمالی برای ارزیابی کمی ریسک لازم است.
- واریانس یا پراکندگی اعضای همادی نشان دهنده میزان عدم قطعیت در پیش بینی می‌باشد.
- تابع چگالی احتمال وضعیت هوای آینده، اطلاعات خوبی درباره وقایع فرین^۱ جهت اعلام هشدارهای لازم می‌دهد.

با وجود این پیش بینی قطعی نیز در برخی موارد به پیش بینی احتمالی برتری‌هایی دارد:

- هزینه صدور پیش بینی قطعی بسیار کمتر از پیش بینی احتمالی است.
- اینگونه نیست که همه کاربران خواهان تصمیم‌گیری بهینه باشند؛ اغلب برخی ترجیح می‌دهند که پیش بین به طور دقیق درباره چگونگی وضع هوای آینده گزارش دهد (با وجود این واقعیت که این امر ناممکن است).
- برای بیان پیش بینی احتمالی نیاز به مبادله حجم بیشتری از اطلاعات است و به سختی می‌توان این اطلاعات را به طور خلاصه انتقال داد (مثلاً در یک برنامه کوتاه تلویزیونی) اما گزارش پیش بینی قطعی ساده‌تر و برای کاربران قابل فهم‌تر می‌باشد.
- برای پیش بینی قطعی امکان رسم نقشه وجود دارد.

^۱ Extreme event

بنابراین با تشکیل یک سامانه همادی^۱ و با استفاده از روش های رایج پس پردازش همادی، وضع هوای آینده به صورت یک تابع توزیع احتمال پیش بینی می شود.

۲-۵) پس پردازش آماری^۲ خروجی مدل

پس پردازش آماری یا کالیبراسیون خروجی مدل پیش بینی عددی وضع هوا (NWP^۳) متداول می باشد زیرا مهارت کسب شده از آن با سال ها بهبود مدل پایه برابری می کند. حتی در مقایسه با سایر رویکردهای مرسوم بهبود مهارت مانند افزایش تفکیک مدل، دارای مهارت بیشتری است. از لحاظ تاریخی، روش های پس پردازش آماری برای تشخیص متغیرهایی که به طور مستقیم از مدل های عددی وضع هوای نسل اول با تفکیک پایین قابل پیش بینی نبودند، مورد استفاده واقع می شدند. متغیرهای مستقل مدل با سایر متغیرهای غیر قابل پیش بینی مانند بارش یخ زده^۴، مه و پوشش ابر به طور آماری مرتبط می شوند. اما مدل های نسل فعلی با تفکیک بالا قادر به پیش بینی آن متغیرها هستند و روش های آماری برای کاهش خطای سامانمند^۵ به کار می روند.

به طرق مختلف می توان روش های پس پردازش آماری را دسته بندی کرد. ممکن است آنها از نظر تکنیک های آماری به کار رفته و همچنین نوع داده های پیش بینی کننده برای توسعه روابط آماری طبقه بندی شوند. بین روش ایستا و پویا تفاوت به وجود می آید. در روش های ایستا، الگوریتم های آماری برای از بین بردن خطای سامانمند، از یک دوره آموزش طولانی مدت بر اساس نسخه یکسان از مدل توسعه داده می شود و بدون تغییر برای یک دوره قابل توجهی از زمان به کار می رود. به دلیل هزینه محاسباتی مرتبط با محاسبات روابط آماری، مدل اغلب نمی تواند ارتقا یابد زیرا این کار نیازمند تجدید محاسبات روابط خواهد بود.

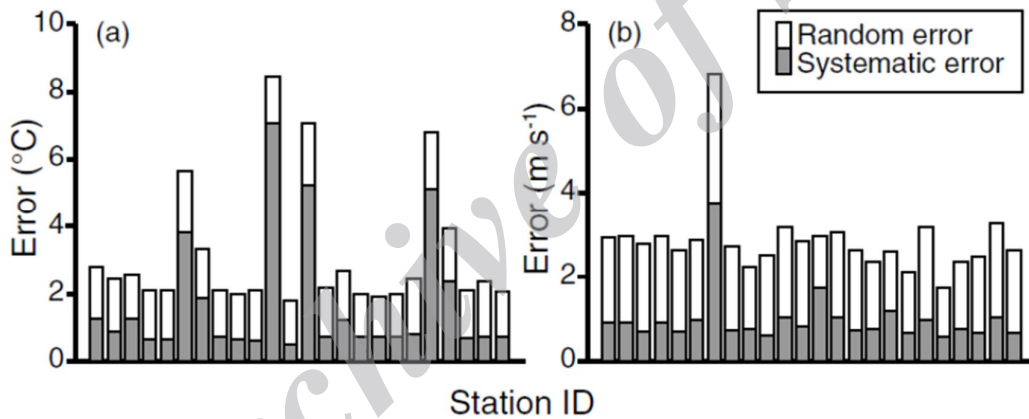
۲-۶) حذف خطای سامانمند پیش بینی روی ایستگاهها

بخش حاضر به معرفی روش های مختلف اصلاح پیش بینی های مدل NWP برای کاهش خطای سامانمند می پردازد. روش های ایستا نیازمند به یک دوره طولانی مدت از داده های بازپیش بینی^۶ مدل برای ایجاد یک رابطه بین خروجی و

^۱ Ensemble system
^۲ Statistical post-processing
^۳ Numerical Weather Prediction
^۴ freezing rain
^۵ systematic error
^۶ reforecast

مشاهدات گذشته می باشد. این روابط آماری غالباً برورسانی نمی شوند. در مقابل، روش های پویا تصحیحات را بر اساس دوره آموزش کوتاه تر انجام می دهند. هدف در هر دو حالت، کاهش خطای پیش بینی فعلی با استفاده از برآورد خطای گذشته است.

باید توجه داشت که با استفاده از این روش ها، فقط خطای سامانمند کاهش می یابد و خطای تصادفی^۱ که به طور آماری حذف شدنی نیست باقی می ماند. با این وجود، خطای سامانمند می تواند بخش قابل توجهی از خطای کل به ویژه در سطح زمین باشد بنابراین حذف آن از طریق پس پردازش های آماری بسیار مفید است. به عنوان مثال شکل ۱-۲ خطای تصادفی و سامانمند پیش بینی دمای سطح زمین و سرعت باد با استفاده از مدل منطقه ای میان مقیاس^۲ برای شمال غربی ایالت متحده را نشان می دهد. کاملاً مشخص است که خطای سامانمند در برخی ایستگاه های دیدبانی بیشتر از خطای تصادفی می باشد.



شکل ۱-۲: مقایسه خطای تصادفی و سامانمند دمای سطح زمین و سرعت باد

۱-۶-۲ روش پیش بینی کامل

اولین رویکرد در پس پردازش آماری پیش بینی کامل^۳ می باشد (Klein, Lewis, & Enger, 1959). در اینجا، مشاهدات کمیت هایی که به وسیله مدل پیش بینی شده اند (پیش بینی کننده) به مشاهدات یک پیش بینی شونده که لزوماً توسط مدل پیش بینی نشده است به طور آماری مرتبط می شود. سپس پیش بینی مدل NWP برای پیش بینی

^۱ Random error

^۲ regional mesoscale model

^۳ perfect-prognosis

کننده‌ها در روابط رگرسیون به کار می‌رود تا متغیر پیش‌بینی شونده پیش‌بینی شود. به دلیل آنکه روابط آماری از پیش-بینی مدل به دست نیامده است، خطای مدل را کاهش نمی‌دهد. این روش به سادگی متغیرهای پیش‌بینی شده را به متغیرهای غیر قابل پیش‌بینی یا ضعیف پیش‌بینی شده مرتبط می‌کند. در واقع، فرض بر این است که پیش‌بینی مدل کامل است. زیرا همانطور که قبلاً اشاره شد، مدل‌های فعلی در پیش‌بینی بسیاری از متغیرهایی که قبلاً باید با روش پیش‌بینی کامل پیش‌بینی می‌شدند توانمند هستند. این روش به صورت عملیاتی کمتر استفاده می‌شود. مزیت روش پیش‌بینی کامل در این است که از مدل مستقل است و روابط آماری با تغییر مدل دیگر نیاز به محاسبه مجدد ندارند.

۲-۶-۲) آماره برون داد مدل (MOS)

روش MOS^۱ توسط گلان و لوری در سال ۱۹۷۲ پیشنهاد داده شد (Glahn & Lowry, 1972). در این روش، پیش-بینی‌های قبلی یک متغیر و مشاهدات مربوط به آن متغیر برای تعیین خطای سیستماتیک (اریبی) در هر نقطه دیدبانی به طور آماری به یکدیگر مرتبط می‌شوند. اریبی محاسبه شده در هر نقطه دیدبانی برای تصحیح پیش‌بینی‌های آینده مورد استفاده قرار می‌گیرند. از آنجا که روش MOS نیازمند محاسبات جداگانه آماره‌ها بر اساس زوج (پیش‌بینی، مشاهدات) برای هر زمان، برای هر مکان دیدبانی و برای هر متغیر است تعداد معادلات زیادی درگیر هستند.

جکس و همکاران در سال ۱۹۹۰ (Jacks, Bower, & Dagostaro, 1990) خلاصه‌ای از سیستم MOS که شامل پیش‌بینی کننده‌های دما، فرارفت دما^۲، ضخامت^۳، میزان بارش، آب بارش شو^۴، رطوبت نسبی، سرعت قائم، مؤلفه‌های باد افقی، سرعت باد، فرا رفت تاوایی^۵، تاوایی نسبی^۶، پایداری^۷ و همگرایی رطوبت^۸ می‌باشد را ارائه کرد. سیستمی شامل چندین هزار معادله که از نظر محاسباتی بسیار هزینه‌بر بود.

ویلسون و والی (Wilson & Vallée, 2002) در روش UMOS^۹، با تغییر مدل NWP معادلات حاصل از روش MOS را بروز می‌کنند. این کار توسط وزن‌هایی که تحت کنترل کاربر است انجام می‌شود؛ به این صورت که پس از تغییر مدل، به خصوصیات آماری مدل جدید و مدل قدیم وزن داده شده و به رابطه آماری اضافه می‌شود. به این معنی

^۱ model output statistics

^۲ temperature advection

^۳ thickness

^۴ precipitable water

^۵ vorticity advection

^۶ relative vorticity

^۷ stability

^۸ moisture convergence

^۹ Updatable MOS

که، در روش UMOS به جای استفاده از یک رابطه آماری جدید با دوره آموزش طولانی مدت، با وزندهی به مدل جدید رابطه قبلی بروز می شود.

۲-۶-۳) روش کالمن فیلتر^۱

فیلتر کالمن یک روش دیگر پس پردازش می باشد که از پیش بینی ها و مشاهدات گذشته برای تخمین اریبی مدل به کار می رود. در این روش از معادله زیر برای تخمین اریبی استفاده می شود:

$$B_{t+\Delta t} = B_t + \beta_t (y_t - B_t) \quad (1)$$

به طوری که $B_{t+\Delta t}$ برآوردی از اریبی پیش بینی یک متغیر بعد از زمان Δt ، B_t برآورد اریبی در انتهای پیش بینی گذشته، y_t خطای پیش بینی مشاهده شده (خطای سامانمند و تصادفی) در انتهای چرخه گذشته و β_t عامل وزندهی می باشد. برای پیش بینی در حال حاضر، اریبی B_t از قبل تخمین زده شده و به عنوان حدس اولیه برای تخمین اریبی در Δt زمان بعد یعنی $B_{t+\Delta t}$ مورد استفاده قرار می گیرد. به دلیل آنکه پیش بینی در زمان t به وقوع پیوسته است، y_t (تفاضل پیش بینی از مشاهده) قابل محاسبه می باشد. بنابراین اریبی آینده، با برآوردی از اخیرترین اریبی که با یک تفاضل وزنی بین اریبی برآورد شده و خطای کل مشاهده شده جمع شده است، تخمین زده می شود.

۲-۷) پس پردازش سامانه همادی

از آنجا که ناهمواری و کاربری اراضی در مدل و واقعیت تفاوت زیادی با هم دارند و نیز به علت خطاهای نسبتاً زیاد در بخش فیزیک مدل، خروجی مدل های عددی پیش بینی وضع هوا همواره دارای خطای سامانمند می باشد. در نتیجه خروجی سامانه همادی که مجموعه خروجی منفرد مدل های عددی می باشد نیز دارای خطای سامانمند خواهد بود.

با ایجاد سامانه همادی، می توان به تابع چگالی احتمال وضع هوای آینده دست پیدا کرد؛ به طوری که اعضای همادی نمونه هایی از تابع چگالی احتمال عدم قطعیت وضعیت آینده جو می باشند. برای رفع این نواقص و کاهش خطا و همچنین دستیابی به تابع چگالی احتمال وضع هوای آینده، تعدادی روش های پس پردازش آماری مانند هیستوگرام رتبه-

^۱ Kalman filter

ای (Hamill & Colucci, 1998)، برازش لجستیک (Hamill et al., 2004)، روش درسینگ (Roulston & Smith, 2003)، میانگین گیری به روش بیزی (Raftery et al., 2005)، آماره برون داد مدل همادی (EMOS)^۱ (Gneiting et al., 2005)، رگرسیون همادی (Unger, 2009)، تکنیک های تورم واریانس (Johnson & Bowler, 2009)، کالیبره انتقال و کشش (Eckel et al., 2012) و روش EKDMO (Glahn, 2009) (Veenhuis, 2013) ارائه شده است. در داخل کشور نیز پیش بینی احتمالی بارش با استفاده از خروجی مدل MM5^۲ (Azadi & Zakeri, 2010) و مدل WRF-MM5 (Azadi et al., 2009) صادر و برتری آن نسبت به پیش بینی قطعی بررسی شده است.

در ادامه چند روش به تفصیل توضیح داده می شود.

۲-۷-۱) تخمین کمینه CRPS^۳

نیتینگ^۴ و همکاران برای تخمین تابع چگالی احتمال سامانه همادی از کمینه کردن CRPS استفاده کردند (Goldman, 2005 & Gneiting, Raftery, Westveld). فرض کنید X_1, \dots, X_m اعضای مجموعه همادی و y مشاهده و S^2 واریانس همادی باشد. تابع چگالی احتمال توزیع نرمال سامانه همادی به این صورت تعریف می شود:

$$N(a + b_1X_1 + \dots + b_mX_m, c + dS^2) \quad (2)$$

$a + b_1X_1 + \dots + b_mX_m$ میانگین وزنی مجموعه همادی با اربیی تصحیح شده می باشد. واریانس $c + dS^2$ یک تابع خطی از واریانس همادی است. ضرایب b_1, \dots, b_m نشان دهنده عملکرد اعضای همادی با توجه به داده های دوره آموزش می باشد. ضرایب واریانس c و d نشان دهنده پراکندگی همادی و عملکرد پیش بینی میانگین همادی است.

یک تکنیک کلاسیک برای تخمین ضرایب a, b_1, \dots, b_m, c و d برآورد بیشینه درستنمایی می باشد. از آنجا که ماکزیمم کردن لگاریتمی تابع درستنمایی ساده تر است، تابع لگاریتمی درستنمایی به صورت زیر نوشته می شود:

^۱ Ensemble Model Output Statistics

^۲ Mesoscale model version 5

^۳ Continuous ranked probability score

^۴ Gneiting

$$\begin{aligned}
 l(a; b_1, \dots, b_m; c; d) & \quad (3) \\
 &= -\frac{1}{2} \left\{ k \log(2\pi) \right. \\
 &+ \sum_{i=1}^k \frac{[Y_i - (a + b_1 X_{i1} + \dots + b_m X_{im})]^2}{c + d S_i^2} \\
 &+ \left. \sum_{i=1}^k \log(c + d S_i^2) \right\}
 \end{aligned}$$

که k طول دوره آموزش و X_{i1}, \dots, X_{im} اعضای آامین مجموعه همدادی و S_i^2 و Y_i به ترتیب واریانس و مشاهده آامین مجموعه همدادی در دوره آموزش می باشد. با استفاده از الگوریتم EM^1 می توان تابع درستنمایی را ماکزیمم کرد. اما رویکردی دیگر مینم کردن امتیاز CRPS است. در این روش ضرایب a, b_1, \dots, b_m, c و d با مینم کردن مقدار CRPS با استفاده از داده های دوره آموزش تخمین زده می شوند. بنابراین فرمول زیر را داریم:

$$\begin{aligned}
 \Gamma(a; b_1, \dots, b_m; c; d) & \quad (4) \\
 &= \frac{1}{k} \sum_{i=1}^k (c + d S_i^2)^{1/2} \left\{ Z_i [2\Phi(Z_i) - 1] \right. \\
 &+ \left. 2\Phi(Z_i) - \frac{1}{\sqrt{\pi}} \right\}
 \end{aligned}$$

به طوری که

$$Z_i = \frac{Y_i - (a + b_1 X_{i1} + \dots + b_m X_{im})}{(c + d S_i^2)^{1/2}} \quad (5)$$

فرمول (4) را می توان با استفاده از الگوریتم Broyden-Fletcher-Goldfarb-Shanno کمینه کرد.

¹ Expectation maximization

۲-۷-۲) میانگین گیری به روش بیزی

آنالیز آماری استاندارد مانند آنالیز رگرسیون، روی یک مدل آماری مفروض به صورت مشروط اعمال می شود. معمولاً این مدل از بین چندین مدل ممکن برای داده ها انتخاب شده است به طوری که تحلیلگر داده ها مطمئن نیست که مدل مفروض بهترین باشد. از طرفی دیگر، سایر مدل های قابل قبول دارای جواب های متفاوت می باشند. بنابراین یک چنین رویکردی که فقط یک مدل را به عنوان بهترین مدل مدنظر قرار می دهد باعث ایجاد عدم قطعیت در نتیجه خواهد شد.

میانگین گیری از مدل بیزی^۱ (BMA) با در نظر گرفتن همه مدل های آماری، این مشکل را برطرف می کند (Raftery, Polakowski, 2005 & Gneiting, Balabdaoui). فرض می کنیم کمیت y بر اساس داده های دوره آموزش که با y^T نشان می دهیم با استفاده از تعداد K مدل آماری (M_1, \dots, M_K) پیش بینی می شود. بر اساس قانون احتمال کل، PDF پیش بینی یعنی $p(y)$ از فرمول زیر بدست می آید:

$$p(y) = \sum_{k=1}^K p(y|M_k)p(M_k|y^T) \quad (6)$$

که در آن $p(y|M_k)$ تابع چگالی احتمال پیش بینی بر مبنای مدل M_k و $p(M_k|y^T)$ احتمال پسین مدل M_k بر اساس داده های دوره آموزش است.

مبنای ایده اخیر این است که برای هر مجموعه همادی، یک بهترین عضو وجود دارد که ما نمی دانیم کدامیک است. این عدم قطعیت درباره بهترین عضو به وسیله BMA کمی می شود. فرض کنید y مقداری است که پیش بینی می شود و f_k ، k امین عضو مجموعه همادی است. هر کدام از f_k ها دارای یک PDF مشروط $g_k(y|f_k)$ می باشد به طوری که احتمال وقوع y به شرط f_k را مشخص می کند. در صورتی که f_k بهترین عضو همادی باشد. مدل پیش بینی BMA به صورت فرمول زیر است:

$$p(y|f_1, \dots, f_K) = \sum_{k=1}^K w_k g_k(y|f_k) \quad (7)$$

^۱ Bayesian Model Averaging

که w_k احتمال پسین k امین پیش بینی است و مقدار آن با توجه به عملکرد k امین پیش بینی در دوره آموزش تعیین می شود. به دلیل آنکه w_k ها احتمال هستند بنابراین مثبت بوده و جمع آنها برابر با یک می باشد:

$$\sum_{k=1}^K w_k = 1 \quad (8)$$

برای پیش بینی دما و فشار سطح دریا منطقی به نظر می رسد که PDF مشروط با توزیع نرمال با میانگین $a_k + b_k f_k$ و انحراف معیار σ تقریب زده شود:

$$y|f_k \sim N(a_k + b_k f_k, \sigma) \quad (9)$$

برای راحتی، شرایطی را در نظر می گیریم که PDF مشروط دارای توزیع نرمال باشد. این فرض برای برخی متغیرها مانند دما و فشار سطح دریا منطقی است اما برای متغیرهایی مانند سرعت باد و بارش، توزیع نرمال مناسب نمی باشد و باید از توزیع های دیگر استفاده شود. در اینجا می خواهیم a_k و b_k و w_k برای $k = 1, \dots, K$ و σ را از روی یک مجموعه داده ای آموزش شامل پیش بینی های همادی و مشاهدات تخمین بزنیم. مکان و زمان با زیر نویس های s و t مشخص شده اند به طوری که f_{kst} نشان دهنده k امین پیش بینی در مجموعه همادی برای مکان s و زمان t و y_{st} نشان دهنده مشاهده مربوط به آن زمان و مکان است.

برای تخمین پارامترهای a_k و b_k از رگرسیون خطی ساده بین f_{kst} و y_{st} روی داده های دوره آموزش و برای تخمین w_k و σ از بیشینه درست نمایی بروی داده های دوره آموزش استفاده می شود. از آنجا که ماکزیم کردن حالت لگاریتمی تابع درست نمایی ساده تر است، بنابراین تابع درست نمایی به صورت لگاریتمی نوشته می شود:

$$l(w_1, \dots, w_k, \sigma^2) = \sum_{s,t} \log \left(\sum_{k=1}^K w_k g_k(y_{st} | f_{kst}) \right) \quad (10)$$

بیشینه کردن این تابع با استفاده از روش های عددی مانند روش نیوتن-رافسون بسیار پیچیده است. بنابراین از الگوریتمی به نام الگوریتم EM^1 برای بیشینه کردن تابع درست نمایی استفاده می کنیم. برای این منظور یک متغیر

¹ Expectation Maximization

کمکی Z_{kst} را معرفی می کنیم به طوری که اگر k امین عضو مجموعه همادی بهترین پیش بینی برای مکان s و زمان t باشد آنگاه $Z_{kst} = 1$ و در غیر این صورت $Z_{kst} = 0$. با این تعریف برای هر (s, t) فقط یکی از $[Z_{1st}, \dots, Z_{Kst}]$ برابر با یک و بقیه برابر با صفر هستند.

الگوریتم EM تکراری است و بین دو گام E و M تکرار می شود. این الگوریتم با یک حدس اولیه $\theta^{(0)}$ برای بردار پارامتر θ شروع می شود. در گام E، Z_{kst} با حدس فعلی برای پارامترها تخمین زده می شود. مقدار بدست آمده Z_{kst} لزوماً عدد صحیح نمی باشد اگرچه مقدار واقعی آن صفر و یا یک می باشد. در گام M، بردار پارامترها θ با استفاده از Z_{kst} تخمین زده شده در گام قبل، محاسبه می شود.

گام E با توجه به فرمول های BMA به قرار زیر است:

$$\hat{z}_{kst}^{(j)} = \frac{w_k^{(j-1)} g(y_{st}|f_{kst}, \sigma^{(j-1)})}{\sum_{i=1}^K w_i^{(j-1)} g(y_{st}|f_{ist}, \sigma^{(j-1)})} \quad (11)$$

که بالانویس z مربوط به زامین تکرار الگوریتم EM است و $g(y_{st}|f_{kst}, \sigma^{(j-1)})$ چگالی نرمال با میانگین $a_k + b_k f_{kst}$ و انحراف معیار $\sigma^{(j-1)}$ می باشد.

گام M شامل تخمین w_k و σ با استفاده از فرمول های زیر می باشد:

$$w_k^{(j)} = \frac{1}{n} \sum_{s,t} \hat{z}_{kst}^{(j)} \quad (12)$$

$$\sigma^{2(j)} = \frac{1}{n} \sum_{s,t} \sum_{k=1}^K \hat{z}_{kst}^{(j)} (y_{st} - f_{kst})^2 \quad (13)$$

۲-۷-۳) رگرسیون همادی

آنگر و همکاران (Unger, Dool, O, Lenic, & Collins, 2009) یک مدل رگرسیونی را برای صدور پیش‌بینی احتمالی توسعه دادند. آنها در این تحقیق اثبات کردند که ضرایب مدل رگرسیون خطی بین مشاهدات و بهترین عضو سامانه همادی با ضرایب مدل رگرسیونی بین مشاهدات و میانگین اعضای سامانه همادی برابر است. ضرایب مدل رگرسیونی را با روش حداقل مربعات و با داده‌های دوره آموزش تخمین زدند و با استفاده از آن ضرایب و مدل بدست آمده برای هر عضو سامانه همادی مقداری را به عنوان وضع هوای آینده پیش‌بینی کردند. برای پیش‌بینی تابع توزیع احتمال از روش تابع چگالی کرنل گاوسی^۱ استفاده کرده‌اند که مقادیر پیش‌بینی شده از هر عضو سامانه همادی به عنوان مرکز هر کرنل و خطای رگرسیون به عنوان پهناهای هر کرنل در نظر گرفته شده است.

$$Y = a_0 + a_1 f_m \quad (14)$$

$$a_1 = R_m \frac{S_Y}{S_m}, \quad a_0 = \langle Y \rangle - a_1 \langle f_m \rangle \quad (15)$$

به طوری که Y مشاهده، f_m میانگین اعضای سامانه همادی، a_0 و a_1 ضرایب رگرسیون، R_m ضریب همبستگی بین مشاهدات و میانگین‌های همادی، S_Y و S_m انحراف معیار مشاهدات و میانگین‌های همادی در دوره آموزش می‌باشند.

^۱Gaussian Kernel Density Distribution

فصل سوم: روش‌های پس‌پردازش آماری برای نقاط

شبکه منظم

Archive of SID

۳-۱) مقدمه

با توجه به نیاز روز افزون کاربران به پیش بینی وضع هوای بدون خطا در هر نقطه دلخواه، در این فصل به روش های آماری پس پردازش در هر نقطه شبکه منظم پرداخته شده است. در این روش ها به دلیل عدم وجود داده مشاهده ای در نقاط شبکه، از مقدار پس پردازش شده روی ایستگاهها برای این منظور استفاده می شود. در واقع پیش بینی پس پردازش شده در نقاط ایستگاهی به نقاط شبکه درونیابی می شوند.

۳-۲) پس پردازش آماری روی نقاط شبکه منظم

روش های مختلفی برای از بین بردن خطای سامانمند یا پس پردازش برون داد خام مدل های پیش بینی عددی وجود دارد اما روش های ذکر شده، مقادیر پیش بینی شده روی ایستگاهها را تصحیح می کنند. در سال های اخیر نیاز به پیش بینی بدون اریبی در نقاط غیر ایستگاهی یا اساساً نقشه پس پردازش شده زیاد شده است و روش هایی ابداع شده اند که می توانند پیش بینی پس پردازش شده برای هر نقطه دلخواه را بدست آورند که در ادامه درباره برخی از آنها توضیحاتی داده خواهد شد.

۳-۳) حذف اریبی سامانمند شبکه

اکل^۱ و مس^۲ در سال ۲۰۰۵ برای سیستم همادی پیش بینی شبکه بندی شده مدل MM5^۳ قبل از محاسبه میانگین همادی و صدور تابع چگال احتمال، اریبی را حذف کردند (MASS, BAARS, WEDAM, GRIMIT, & STEED, 2007). به این صورت که با میانگین گیری اریبی دو هفته پیش و با استفاده از آنالیز شبکه بندی، اریبی را تصحیح کردند. یوسف و استنسرود در سال ۲۰۰۶ (Yussouf & Stensrud, 2006) اریبی یک دوره ۱۲ روزه در

^۱ Eckel

^۲ Mass

^۳ fifth-generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model (MM5)

ایستگاه‌های دیدبانی را با استفاده از روش کرسمن (Cressman, 1959) به روی نقاط شبکه درون‌یابی کردند که نتایج بدست آمده نشان دهنده بهبود در پیش‌بینی متغیرهای سطح زمین در منطقه اکلاهما^۱ بود. سازمان آب و هوایی ملی^۲ اقدام به ایجاد یک سیستم شبکه‌بندی MOS شبیه به MOS متعارف برای کاهش آریبی سامانند کرده است (Dallavalle & Glahn, 2005). در این سیستم ابتدا مقادیر MOS در ایستگاه‌های دیدبانی محاسبه شده و سپس با روش تغییر یافته کرسمن به نقاط شبکه درون‌یابی می‌شوند. تفاوت این روش با روش کرسمن در نظر گرفتن ارتفاع نقاط دیدبانی و نقاط شبکه در درون‌یابی می‌باشد. به علاوه، نوع سطح نیز در نظر گرفته شده است؛ به این صورت که داده‌های زمینی برای نقاط شبکه زمینی و داده‌های آبی برای نقاط شبکه آبی به کار رفته‌اند.

روش بهینه برای حذف آریبی یک شبکه پیش‌بینی باید دارای ویژگی‌های زیر باشد:

- پایدار و قابل انطباق به هر منطقه‌ای باشد.
- برای هر میزان فاصله بین نقاط شبکه (۱ تا ۱۰ کیلومتر) خروجی مدل‌های میان مقیاس^۳ قابل استفاده باشد.
- برای مناطق با تراکم کم، کارایی داشته و از اطلاعات داده‌های موجود در مناطق با تراکم بیشتر بهره گیرد.
- در زمان‌هایی که سوابق اقلیمی دراز مدت یا آرشیوی از پیش‌بینی‌های مدل در طولانی مدت موجود نیست مورد استفاده قرار گیرد.
- قابلیت انطباق با تغییرات رژیم که ممکن است باعث تغییر ناگهانی آریبی مدل شود، را دارا باشد.

مراحل تصحیح آریبی پیش‌بینی دمای دو متری بر اساس داده‌های مشاهداتی به صورت زیر می‌باشد:

(۱) ابتدا خطای پیش‌بینی در ایستگاه‌های دیدبانی برآورد می‌شود. خطای پیش‌بینی در نقاط دیدبانی از تفریق داده مشاهداتی از پیش‌بینی که با روش درون‌یابی دوخطی^۴ چهار نقطه شبکه‌ای اطراف دیدبانی بدست آمده، محاسبه

^۱ Oklahoma

^۲ National Weather Service

^۳ Mesoscale model

^۴ Bilinear

می شود. آهنگ افت دمای جوی استاندارد^۱ ($-6.5 \text{ } ^\circ\text{C km}^{-1}$) برای اختلاف ارتفاع زمین لحاظ شده در مدل و ارتفاع واقعی ایستگاه دیدبانی استفاده می شود.

۲) در تصحیح اریبی فقط از داده های دیدبانی نزدیک و هم ارتفاع با نقاط شبکه مورد نظر استفاده می شود. در ادامه توضیح داده خواهد شد که از میانگین گیری وزنی بر مبنای فاصله ایستگاه از نقطه شبکه استفاده نمی شود تا تأثیر ایستگاه نزدیک که دارای خطای بازنمودی^۲ می باشد را کاهش دهد.

در تصحیح اریبی، از ایستگاه های دیدبانی که دارای کاربری اراضی^۳ مشابه با نقطه شبکه ای مورد نظر است استفاده می شود. زیرا به طور تجربی مشاهده شده است که کاربری اراضی تأثیر زیادی روی ماهیت بسیاری از اریبی های سطحی داشته است. به عنوان مثال مناطق پوشیده شده با آب، اریبی متفاوتی از مناطق خشک دارند و مناطق کویری دارای اریبی متفاوت با جنگل ها و مناطق حاصلخیز هستند. برای این روابط، کاربری اراضی به ۲۴ گروه که در مدل MM5 استفاده می شود، دسته بندی می شوند (جدول ۳-۱) و آنهایی که دارای ویژگی های مشابه هستند به ۹ گروه که در جدول ۳-۲) نشان داده شده دسته بندی می شوند.

۳) این روش به صورتی است که اثرات تغییر رژیم که یک مشکل عمده در اغلب روش های تصحیح اریبی است کاهش می دهد. وقتی یک تغییر رژیم بزرگ رخ دهد، میانگین گرفتن از داده های دوره قبل ممکن است باعث تخمین نادرست اریبی شود. زیرا طبیعت اریبی با تغییر رژیم عوض می شود و باعث تصحیح نادرست پیش بینی می شود. اثرات تغییر رژیم با دو روش کاهش می یابد. اولاً فقط خطای پیش بینی هایی که مقدار مشابه دارند در تخمین اریبی در یک نقطه شبکه مورد استفاده قرار می گیرند. مثلاً اگر پیش بینی دمای دومتری در یک نقطه $70 \text{ } ^\circ\text{F}$ باشد، فقط خطای پیش بینی های مشابه (بین $65 \text{ } ^\circ\text{F}$ تا $75 \text{ } ^\circ\text{F}$) در برآورد اریبی آن نقطه استفاده می شوند. ثانیاً خطاهای روزهای اخیر با تعداد کافی در برآورد اریبی استفاده می شوند.

^۱ Standard atmospheric lapse rate

^۲ representational error

^۳ Land-use

جدول ۳-۱: ۲۴ گروه کاربری اراضی در مدل MM5

Vegetation integer identification	Vegetation description	Albedo (%)		Moisture available (%)		Emissivity (% at 9 mm)		Roughness length (cm)		Thermal inertia (cal cm ⁻² K ⁻¹ s ^{-1/2})	
		Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter
		1	Urban	18	18	10	10	88	88	50	50
2	Dryland/cropland/pasture	17	23	30	60	92	92	15	5	0.04	0.04
3	Irrigated cropland/pasture	18	23	50	50	92	92	15	5	0.04	0.04
4	Mixed dry-irrigated cropland/pasture	18	23	25	50	92	92	15	5	0.04	0.04
5	Cropland-grassland mosaic	18	23	25	40	92	92	14	5	0.04	0.04
6	Cropland-wood mosaic	18	20	35	60	93	93	20	20	0.04	0.04
7	Grassland	19	23	15	30	92	92	0.12	0.1	0.03	0.04
8	Shrubland	22	25	10	20	88	88	10	10	0.03	0.04
9	Mixed shrubs-grassland	20	24	15	25	90	90	11	10	0.03	0.04
10	Savanna	20	20	15	15	92	92	15	15	0.03	0.03
11	Deciduous broadleaf	16	17	30	60	93	93	50	50	0.04	0.05
12	Deciduous needleleaf	14	15	30	60	94	93	50	50	0.04	0.05
13	Evergreen broadleaf	12	12	50	50	95	95	50	50	0.05	0.05
14	Evergreen needleleaf	12	12	30	60	95	95	50	50	0.04	0.05
15	Mixed forest	13	14	30	60	94	94	50	50	0.04	0.06
16	Water bodies	8	8	100	100	98	98	0.01	0.01	0.06	0.06
17	Herbaceous wetland	14	14	60	75	95	95	20	20	0.06	0.06
18	Wooded wetland	14	14	35	70	95	95	40	40	0.05	0.05
19	Barren/sparse vegetation	25	25	2	50	85	85	10	10	0.02	0.02
20	Herbaceous tundra	15	60	50	90	92	92	10	10	0.05	0.05
21	Wooded tundra	15	50	50	90	93	93	30	30	0.05	0.05
22	Mixed tundra	15	55	50	90	92	92	15	15	0.05	0.05
23	Bare ground tundra	25	70	2	95	85	95	0.1	5	0.02	0.05
24	Snow or ice	55	70	95	95	95	95	5	5	0.5	0.5

جدول ۳-۲: ۹ گروه کاربری اراضی

Combined land-use category	Component MM5 land-use categories
1, urban	1
2, cropland	2, 3, 4, 5, 6
3, grassland	7, 8, 9, 10
4, forest	11, 12, 13, 14, 15
5, water	16
6, wetland	17, 18
7, barren tundra	19, 23
8, wooded tundra	20, 21, 22
9, snow-ice	24

۴) اریبی برای هر ساعت از روز باید محاسبه شود زیرا اریبی در ساعات مختلف روز متفاوت است. همچنین در یک زمان یکسان از روز، ویژگی‌های اریبی با سن پیش‌بینی^۱ تغییر می‌کند.

۵) با توجه به شرایط ذکر شده در بندهای قبل، اریبی در هر نقطه شبکه با میانگین‌گیری ساده اریبی محاسبه شده در نقاط دیدبانی اطراف آن نقطه برآورد می‌شود. در واقع ۵ ایستگاه دیدبانی نزدیکتر به نقطه شبکه‌ای که دارای کاربری اراضی، ارتفاع و مقدار پیش‌بینی مشابه هستند انتخاب می‌شوند. از میانگین‌گیری ساده بدون وزن‌دهی بر مبنای فاصله استفاده می‌شود تا از انتشار خطایی که در مقیاس کوچکتر از شبکه در یک ایستگاه دیدبانی رخ می‌دهد بر سایر نقاط اطراف شبکه جلوگیری کند. با میانگین‌گیری از اریبی نقاط مختلف دیدبانی برای محاسبه اریبی نقاط شبکه، تأثیر ایستگاه‌های دیدبانی دارای مشکل کاهش می‌یابد.

۳-۴) شبکه بندی MOS

گلان و همکاران در مقاله خود در سال ۲۰۰۹ روش MOS را به نقاط شبکه تعمیم داده‌اند (Glahn, Gilbert, Ruth, Cosgrove, & Sheets, 2009). روشی که توسط برگترسن و دوس (Bergthorsson & Doos, 1955) پیشنهاد و توسط کرسمن (Cressman, 1959) بکار گرفته شده، موسوم به روش BCD است. در این روش با یک الگوریتم تکراری و مقادیر دیدبانی موجود در نقاط ایستگاهی، مقدار پیش‌بینی شده در گره‌های شبکه منظم بهبود پیدا می‌کنند. در این روش برای هر گره شبکه یک دایره با مرکز گره و شعاع نفوذ R در نظر گرفته می‌شود. شعاع نفوذ باید به اندازه‌ای باشد که حداقل یک ایستگاه داخل دایره مفروض وجود داشته باشد. برای هر ایستگاه تفاضل مقدار درونیابی شده و مقدار دیدبانی بدست می‌آید. این تفاضل برای نقاط شبکه‌ای که اطراف آن ایستگاه هستند بکار می‌رود و معمولاً با استفاده از فاصله ایستگاه تا نقطه شبکه وزن‌دهی می‌شوند. سپس با توجه به تعداد ایستگاه‌ها در داخل دایره میانگین وزنی تفاضلات محاسبه و با مقدار قبلی نقطه شبکه جمع می‌شود. این محاسبات به صورت تکراری هستند که معمولاً با ۴ بار تکرار انجام می‌شوند.

روش BCDG توسعه یافته روش BCD می‌باشد، که توسط گلان و همکاران (Glahn, Gilbert, Cosgrove, Ruth, & Sheets, 2009) برای پس‌پردازش پیش‌بینی دمای ماکزیمم روی شبکه منظم معرفی شد. به علت

^۱ forecast projection

وابستگی دما به ارتفاع، در این روش علاوه بر تفاضل مقدار دیدبانی از مقدار درون یابی، تغییرات دما نسبت به تغییرات ارتفاع نیز در نظر گرفته شده است.

مقدار پیش بینی برای هر گره با توجه به دایره در نظر گرفته شده و ایستگاه های داخل دایره از فرمول تکراری زیر بدست می آید:

$$c^{t+1} = c^t + \frac{\frac{1}{n} \sum_{i=1}^n W_i D_i}{\frac{1}{n} \sum_{i=1}^n W_i} \quad (16)$$

که در آن

$$W_i = \frac{R^2 - d_i^2}{R^2 + d_i^2} \quad (17)$$

و c^t مقدار گره در مرحله t ام، n تعداد ایستگاهها در دایره مفروض، d_i فاصله بین ایستگاه و گره، R شعاع نفوذ و D_i ضابطه ای است که با تفاضل مقدار پس پردازش شده و مقدار پیش بینی شده با در نظر گرفتن تغییرات دما نسبت به ارتفاع رابطه مستقیم دارد:

$$D = (S - BB) + \left(VCE - \frac{(G - BB)}{ELEDIF} \right) * ELEDIF \quad (18)$$

$$= S - G + VCE * ELEDIF$$

که در آن S مقدار پس پردازش شده در ایستگاه، BB مقدار درون یابی داده گره به ایستگاه، G مقدار داده گره، $ELEDIF$ تفاضل ارتفاع گره از ارتفاع ایستگاه می باشد. VCE ^۱ میزان تغییرات دما نسبت به تغییرات ارتفاع است. برای محاسبه VCE برای ایستگاه A فرض می کنیم TA و EA به ترتیب میزان دما و ارتفاع در ایستگاه A و TB_i و EB_i

^۱ Vertical Change with Elevation

میزان دما و ارتفاع در ایستگاه B_i ام باشد. B_i ها ایستگاههای نزدیک به ایستگاه A هستند به طوری که مثلا حداکثر ۳۴۰ کیلومتر فاصله و حداقل ۱۳۰ متر اختلاف ارتفاع داشته باشند:

$$VCE(A) = \sum (TB_i - TA_i) / \sum (EB_i - EA_i) \quad (19)$$

۳-۵) مقایسه سه روش برای حذف اریبی شبکه

جل در سال ۲۰۰۷ (Gel Y. , 2007) یک رابطه آماری بین اریبی پیش‌بینی فعلی در یک نقطه شبکه و اریبی پیشین در ایستگاه‌های دیدبانی را با سه روش برآورد کرد و نتایج را مقایسه نمود. رویکرد این مقاله توسعه روش‌هایی است که مس در سال ۲۰۰۳ (Mass, 2003)، و دام و همکاران در سال ۲۰۰۵ (Wedam, Mass, & Steed, 2005)، جلو همکاران در سال ۲۰۰۴ (Gel, Raftery, & T. Gneiting, 2004) و تبالدی در سال ۲۰۰۲ (Tebaldi, 2002) به کار برده‌اند. سه روش برای تصحیح اریبی داده‌های نقاط شبکه با هم مقایسه شده است.

روش اول، روش مبتنی بر مشاهدات محلی^۱ (LOB) می‌باشد که نیازی به آرشیو پیشینه داده مشاهداتی در نقطه دلخواه ندارد. در غیاب داده دیدبانی در نقطه شبکه، اطلاعات هواشناسی از ایستگاه‌های دیدبانی در همسایگی آن نقطه بر اساس نزدیکی، اطلاعات کاربری اراضی، ارتفاع ناهمواری^۲ و مشابهت در ویژگی‌های اقلیمی استخراج می‌شود. اریبی در یک نقطه را می‌توان به صورت تابعی از اطلاعات مشاهده شده در اریبی ایستگاه‌های دیدبانی همسایگی در طول یک دوره زمانی مشخص در نظر گرفت. برای این منظور، اریبی وزن‌دار در ایستگاه‌های دیدبانی در یک دوره زمانی T روزه محاسبه شده و از پیش‌بینی فعلی در نقطه شبکه تفریق می‌شود.

در روش LOB به آرشیو داده‌های مشاهداتی با گستردگی کمتری نسبت به روش MOS نیاز است. در حقیقت، یک نمایش فضایی غنی‌تر که می‌تواند تعداد بیشتری از همسایگی با مدت زمان کمتر، اطلاعات کافی برای تخمین مدل

^۱ Local Observation Based
^۲ terrain height

آماري پایدار را فراهم کند. علاوه بر این، روش LOB با ارائه روشی ساده برای تصحیح اریبی در نقاط شبکه، به طور بالقوه از مشکلات مرتبط با تأثیرات فصلی و روزانه و ناهمگنی فضایی^۱ داده‌های پیش‌بینی اجتناب می‌کند.

روش دوم، توسعه یافته روش MOS می‌باشد. به این صورت که بر خلاف روش MOS، بین اریبی در نقطه شبکه به عنوان پیش‌بینی شونده و طول، عرض، ارتفاع، اطلاعات کاربری اراضی و زمان به عنوان پیش‌بینی کننده یک رابطه غیر خطی با استفاده از روش‌های درخت رگرسیون و دسته‌بندی (CART)^۲ (Breiman, Olshen, & Stone, 1984) (Fielding, 1999) (Burrows, et al., 1995) (Spark & Connor, 2004) و امید ریاضی شرطی (ACE)^۳ (Hastie & Tibshirani, 1990) (Friedman, Estimating optimal transformations for multiple regression and correlations (with (Buja & Kass, 1985)(discussion), 1985) برقرار می‌شود.

ایده روش CART-ACE ایجاد یک مدل فضایی زمانی^۴ از اریبی است که قابل اعمال در هر مکانی، مستقل از وجود آرشیو مشاهداتی در آن مکان باشد. فرایند مدل‌سازی به سه مرحله تقسیم می‌شود: آنالیز تغییرات^۵ فضایی اریبی، آنالیز تغییرات زمانی اریبی و در نهایت ادغام دو مرحله قبل برای ایجاد یک مدل فضایی زمانی از اریبی. در این روش ابتدا، روی تغییرپذیری فضایی متمرکز شده و مدل رگرسیون خطی را تشکیل می‌دهیم:

$$Y = \alpha + \sum_{j=1}^p \alpha_j X_j + \varepsilon \quad (20)$$

که Y اریبی و X_j ها متغیرهای فضایی مانند طول، عرض، ارتفاع، اطلاعات کاربری اراضی و یا ترکیبی از آنهاست که با دو روش رگرسیونی جدید CART و ACE انجام می‌شود.

^۱ Spatial Inhomogeneity

^۲ Classification And Regression Tree (CART)

^۳ Alternative Conditional Expectation (ACE)

^۴ spatiotemporal

^۵ variation

روش CART با توجه به طول، عرض و ارتفاع، کل دامنه را به زیر دامنه‌هایی تقسیم می‌کند که مشاهدات در هر زیر دامنه دارای رفتار مشابه می‌باشند. در روش ACE از تابع اسپلاین^۱ طول، عرض و ارتفاع به عنوان پیش‌بینی کننده‌های مدل استفاده می‌شود. هر دو روش CART و ACE روش‌های رگرسیونی ناپارامتری جدید هستند که اریبی را به طور غیر خطی در نظر می‌گیرند در حالی که این موضوع با رگرسیون خطی کلاسیک امکان پذیر نیست.

از ملاک اطلاعات آکائیکه^۲ (AIC) برای تعیین لیست نهایی پیش‌بینی کننده‌های آماری معنادار (Weisberg, 1985) استفاده شده است. در مقایسه با روش MOS، مزیت اصلی روش CART-ACE توانایی آن در تصحیح اریبی هر نقطه دلخواه بدون نیاز به سوابق پیشینه داده‌های مشاهداتی است.

روش سوم، ترکیبی از روش‌های LOB و CART-ACE است که اطلاعات بدست آمده از روش LOB به عنوان یک پیش‌بینی کننده دیگر در مدل رگرسیون استفاده شده است.

۳-۶ روش GIDS^۳

روش GIDS^۴ در سال ۱۹۹۸ توسط نالدر و وین (Wein, 1998 & Nalder) برای درونیابی داده‌های اقلیمی معرفی شد. استال و همکاران در سال ۲۰۰۶ [۱۴] روش‌های مختلفی را برای درونیابی داده‌های مشاهداتی روی شبکه منظم با هم مقایسه کردند که روش GIDS نسبت به سایر روش‌ها برتری داشت. در تحقیق حاضر، از روش GIDS برای درونیابی پیش‌بینی پس پردازش شده روی شبکه منظم استفاده شده است که توضیحات در بخش بعد آورده شده است. روش‌های مختلف برای درونیابی فضایی دمای هوا از سه جنبه زیر با یکدیگر تفاوت دارند:

۱- مدل مورد استفاده برای توصیف تنوع فضایی درجه حرارت هوا

۲- نحوه در نظر گرفتن تأثیر ارتفاع

۳- چگونگی انتخاب نقاط ایستگاهی

^۱ spline

^۲ Akaike Information Criterion (AIC)

^۳ Gradient-plus-Inverse-Distance-Squared

^۴ Gradient plus Inverse Distance Squared

در بعضی از روش های درونیابی، رویه ای مانند رگرسیون چندگانه و یا اسپلاین برای نقطه مورد نظر برازش می شود. روش دیگر میانگین گیری وزنی است که به صورت فرمول زیر می باشد:

$$T_p = \frac{\sum_{i=1}^n w_i (T_i + f(h_p - h_i))}{\sum_{i=1}^n w_i} \quad (21)$$

که T_p پیش بینی در نقطه مورد نظر، T_i میزان دما در نقطه ایستگاهی i ام، w_i مربوط به نقطه ایستگاهی i ام و $f(h_p - h_i)$ تابعی (معمولاً خطی) از تفاوت ارتفاع (h) در نقطه ایستگاهی و نقطه مورد نظر برای درونیابی می باشد. وزن ها به طرق مختلف محاسبه می شوند؛ مانند معکوس فاصله^۱، فیلتر گاوسی کوتاه شده^۲ و روش های زمین آماری مبتنی بر گریگیدن.

تأثیر ارتفاع بر دما که به آن آهنگ افت دما^۳ گفته می شود، معمولاً به میزان ۶ درجه سانتیگراد در هر کیلومتر است. البته آهنگ افت دما در شرایط هواشناسی مختلف و فصول مختلف مقداری متفاوت دارد. برای این منظور، آهنگ افت دما با استفاده از داده های موجود در ایستگاه های دیدبانی در هر بازه زمانی را می توان با روش رگرسیون ساده و در نظر گرفتن ارتفاع به عنوان پیش بینی کننده و یا رگرسیون چندگانه با در نظر گرفتن طول جغرافیایی، عرض جغرافیایی و ارتفاع به عنوان پیش بینی کننده تخمین زد.

برای انتخاب نقاط ایستگاهی، می توان شعاعی مشخص را از قبل تعیین کرده و ایستگاه های داخل دایره با آن شعاع را در نظر گرفت. همچنین برای تعیین شعاع می توان از نمودار وریوگرام^۴ استفاده نمود. رویکرد دیگر به این صورت است که تعدادی مشخص از نزدیکترین ایستگاه ها را در نظر گرفت.

در ادامه، روش کار انجام شده در تحقیق حاضر توضیح داده خواهد شد. برای تخمین مقدار دما در یک نقطه از شبکه ابتدا چند ایستگاه اطراف نقطه مورد نظر که فاصله آنها تا آن نقطه کمتر از یک مقدار مفروض است انتخاب می شود.

^۱ Inverse distance

^۲ Truncated Gaussian filter

^۳ Laps Rate

^۴ variogram

سپس با توجه به طول و عرض جغرافیایی و ارتفاع و مقدار دمای مشاهداتی ایستگاههای انتخابی یک مدل رگرسیون چندگانه به صورت زیر برآزش کرده و ضرایب رگرسیون b_1 ، b_2 و b_3 را بدست می آوریم:

$$T = b_0 + b_1 long + b_2 lat + b_3 h \quad (22)$$

که T دما و h ، lat ، $long$ به ترتیب طول جغرافیایی، عرض جغرافیایی و ارتفاع می باشد. با استفاده از ضرایب رگرسیونی بدست آمده و مختصات جغرافیایی ایستگاهها و نقطه مورد نظر مقدار دمای تخمین زده شده به صورت زیر بدست می آید:

که d_i فاصله ایستگاه نام از نقطه، T_i و T_p به ترتیب دمای دیدبانی در ایستگاه نام و دمای تخمینی در نقطه مورد نظر می باشد. در تحقیق حاضر از مدل رگرسیون درجه دوم نیز استفاده شده است که به دلیل عدم بهبود در خطای پیش بینی از ذکر آن خودداری می شود.

$$T_p = \left[\sum_{i=1}^n \left(\frac{1}{d_i} \right)^2 \right]^{-1} \times \sum_{i=1}^n \left\{ [T_i + b_1(long_p - long_i) + b_2(lat_p - lat_i) + b_3(h_p - h_i)] \left(\frac{1}{d_i} \right)^2 \right\} \quad (23)$$

۳-۷) آمار فضایی

اغلب روش های معمولی آمار مبتنی بر استقلال مشاهدات نمونه مستخرج از جامعه بنا شده اند. اما در عمل موارد زیادی وجود دارند که مشاهدات مستقل نبوده و بر حسب موقعیت قرار گرفتن آنها در فضای مورد مطالعه به یکدیگر وابسته هستند. اگر این وابستگی تابعی از فاصله بین موقعیت مشاهدات باشد به گونه ای که مشاهدات نزدیک به هم وابسته تر و

مشاهدات دورتر از هم وابستگی کمتری داشته باشند، این گونه مشاهدات داده‌های فضایی^۱ نامیده می‌شوند. بدیهی است که تحلیل آماری داده‌های فضایی با روش‌های آماری معمول مقدور نیست؛ زیرا شرط اساسی استقلال داده‌ها محقق نمی‌شود. لذا شاخه آمار فضایی برای تحلیل اینگونه مشاهدات شکل گرفته و در حال توسعه و فراهم آمدن فن‌های مختلف می‌باشد.

اولین بار توجه به داده‌های فضایی در سال ۱۶۸۶ توسط هالی (Halley, 1686) جغرافیدان و ستاره‌شناس انگلیسی صورت پذیرفته است که برای بررسی وزش بادهای موسمی نقشه زمین‌های کشاورزی در نزدیکی نواحی گرمسیر استوایی، مسیر حرکت بادهای موسمی را اضافه کرد و تلاش نمود علل طبیعی آنها را تعیین کند. در دهه‌های اخیر افراد زیادی به بحث و بررسی پیرامون داده‌های فضایی پرداخته‌اند. استیودنت (Student, 1907) در بررسی توزیع ذرات معلق در مایعات، پیشنهاد کرد که داده‌ها در سطوح مختلف به طور جداگانه به دست آورده شوند. فیشر (Fisher, 1935) با توجه به اینکه کرت‌های نزدیک به هم از لحاظ محصول شباهت بیشتری نسبت به کرت‌های دور از هم دارند به وجود همبستگی فضایی در طرح آزمایش‌های کشاورزی اشاره کرد؛ اما مدل مناسبی برای توصیف چنین پدیده‌هایی ارائه نشد. ماترون (Matheron, 1962) با انتشار مقاله‌ای در سال ۱۹۶۲ پایه‌های زمین‌آمار^۲ را بنا کرد. زمین‌آمار شاخه‌ای از علم آمار است که بر اساس نظریه متغیرهای ناحیه‌ای به تجزیه و تحلیل داده‌های همبسته فضایی می‌پردازد. در واقع بین مقادیر مختلف یک متغیر و فاصله و جهت قرارگیری آنها ارتباطی را برقرار می‌کند که ساختار فضایی نامیده می‌شود و بررسی و مطالعه آن یکی از مسائل مهم آمار فضایی است. دیوید (David, 1977) در مورد پیشگویی فضایی ذخائر معدنی مطالب بسیاری ارائه نمود. یورنل و هیگبرتس (Journel & Huijbregts, 1978) به طور مفصل روش‌های زمین‌آمار را با اصطلاحات خاص شرح دادند. ریپلی (Ripley, 1981) در کتاب آمار فضایی پیشگویی فرایندهای تصادفی را که ارتباط نزدیک با سری‌های زمانی دارند مطرح نمود. کرسی (Cressie, 1993) برای بیان ضرورت لحاظ نمودن همبستگی فضایی در تحلیل داده‌ها به ذکر مسئله برآورد میانگین جامعه‌ای با واریانس معلوم σ^2 پرداخته است.

یکی از روش‌هایی که پیوسته در تجزیه و تحلیل داده‌های فضایی به کار گرفته می‌شود درونیابی آنها براساس مشاهدات است که عموماً به منظور تخمین مقدار متغیر پاسخ در موقعیت‌های جدید و تهیه نقشه‌های جغرافیایی از متغیر بررسی شده به کار می‌رود. در آمار فضایی روش‌های مختلفی برای درونیابی به کار گرفته می‌شود. اولین بار برایان و وایز

^۱ Spatial Data
^۲ Geostatistics

روش‌های عکس مجذور فاصله و عکس توان p فاصله را برای درونیابی داده‌های فضایی مورد استفاده قرار دادند. سپس واهبا (Wahba, 1990) روش اسپلاین را برای درونیابی داده‌ها به کار گرفت. پس از آن کرسی کریگیدن را که اولین بار ماترون (Matheron, 1962) معرفی کرد به عنوان بهترین تخمین‌گر خطی نااریب مورد استفاده قرار داد.

۳-۸) کریگیدن^۱

فرض کنید مشاهدات $Z = (Z(t_1), Z(t_2), \dots, Z(t_n))$ در موقعیت‌های t_1, t_2, \dots, t_n برای میدان تصادفی $\{Z(t); t \in D\}$ در اختیار باشند. فرض کنید به ازای هر t میانگین میدان تصادفی مقدار ثابت و نامعلوم μ و $\sigma(t)$ نیز یک میدان تصادفی ایستای ذاتی با میانگین صفر و نیم‌واریانس معلوم $2\gamma(h)$ باشد. بهترین تخمین‌گر نااریب

$$\hat{Z}(t_0) = \sum_{i=1}^n \lambda_i Z(t_i)$$

بر اساس مشاهدات Z کریگیدن معمولی نامیده می‌شود که در آن ضرایب $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ به گونه‌ای انتخاب می‌شوند که میانگین مربعات خطای

$$\sigma_e^2 = E \left(\hat{Z}(t_0) - Z(t_0) \right)^2 \quad (24)$$

با شرط $\sum_{i=1}^n \lambda_i = 1$ به عنوان تضمین کننده نااریبی آن کمینه گردد. برای این منظور عبارت

$$E \left[Z(t_0) - \sum_{i=1}^n \lambda_i Z(t_i) \right]^2 - 2m \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (25)$$

بر حسب ضرایب $\lambda_1, \lambda_2, \dots, \lambda_n$ و m که در آن m ضریب لاگرانژ است کمینه می‌شود و بردار ضرایب $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$ به صورت $\lambda = \left(\gamma + \mathbf{1} \frac{(1 - \mathbf{1}^T \Gamma^{-1} \gamma)}{\mathbf{1}^T \Gamma^{-1} \mathbf{1}} \right)^T \Gamma^{-1}$ حاصل می‌گردد، که در آن $\mathbf{1}$ یک بردار $n \times 1$ با عناصر واحد، $\gamma = (\gamma(t_0 - t_1), \dots, \gamma(t_0 - t_n))^T$ و Γ یک ماتریس $n \times n$ با (i, j) امین عنصر $\gamma(t_i - t_j)$ است.

^۱ Kriging

اگر میانگین $\mu(t)$ بر حسب t ثابت نباشد، با فرض آنکه یک ترکیب خطی نامعلوم مانند $\mu(t) = \sum_{j=1}^{p+1} \phi_{j-1}(t)\beta_j$ از توابع معلوم $\{\phi_0(t), \dots, \phi_p(t)\}$ باشد، تخمین مقدار $Z(t_0)$ بر حسب مشاهدات، کریگیدن عمومی نامیده می شود. برای محاسبه آن، ضرایب رابطه خطی $\hat{Z}(t_0) = \sum_{i=1}^n \lambda_i Z(t_i)$ به گونه ای تعیین می شوند که عبارت (۲۵) کمینه شود. برای این منظور باید رابطه زیر بر حسب ضرایب کمینه گردد:

$$E \left[Z(t_0) - \sum_{i=1}^n \lambda_i Z(t_i) \right]^2 - 2 \sum_{j=1}^{p+1} m_{j-1} \left\{ \sum \lambda_i \phi_{j-1}(t_i) - \phi_{j-1}(t_0) \right\} \quad (26)$$

در این صورت ضرایب بهینه به صورت $\lambda_u = \Gamma_u^{-1} \gamma_u$ به دست می آیند که در آن یک ماتریس متقارن $(n+p+1) \times (n+p+1)$ است:

$$\lambda_u = (\lambda_1, \dots, \lambda_n, m_0, \dots, m_p)^T \quad (27)$$

$$\gamma_u = (\gamma(t_0 - t_1), \dots, \gamma(t_0 - t_n), 1, \phi_1(t_0), \dots, \phi_p(t_0))^T \quad (28)$$

تعاریف زیر را داریم:

u : بردار مختصات فضایی

$Z(u)$: مقدار متغیر در مکان فضایی

h : بردار تأخیر^۱

^۱ Lag

$z(u+h)$: مقدار متغیر در بردار مکان فضایی با تأخیر

$N(h)$: تعداد زوج‌هایی از مکانها که با تأخیر h از هم فاصله دارند.

کوواریانس^۱:

$$c(h) = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} z(u_{\alpha}) \cdot z(u_{\alpha} + h) - m_0 m_h \quad (29)$$

همبستگی^۲:

$$\rho(h) = \frac{c(h)}{\sqrt{\sigma_0 \cdot \sigma_h}} \quad (30)$$

نیمه واریانس :

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_{\alpha} + h) - z(u_{\alpha})]^2 \quad (31)$$

که m_0 ، m_h ، σ_0 و σ_h به ترتیب میانگین متغیر در مکان فعلی، میانگین متغیر در فاصله h و انحراف معیار متناظر با آنها می‌باشد و به صورت زیر تعریف می‌شوند:

$$m_0 = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} z(u_{\alpha}) \quad (32)$$

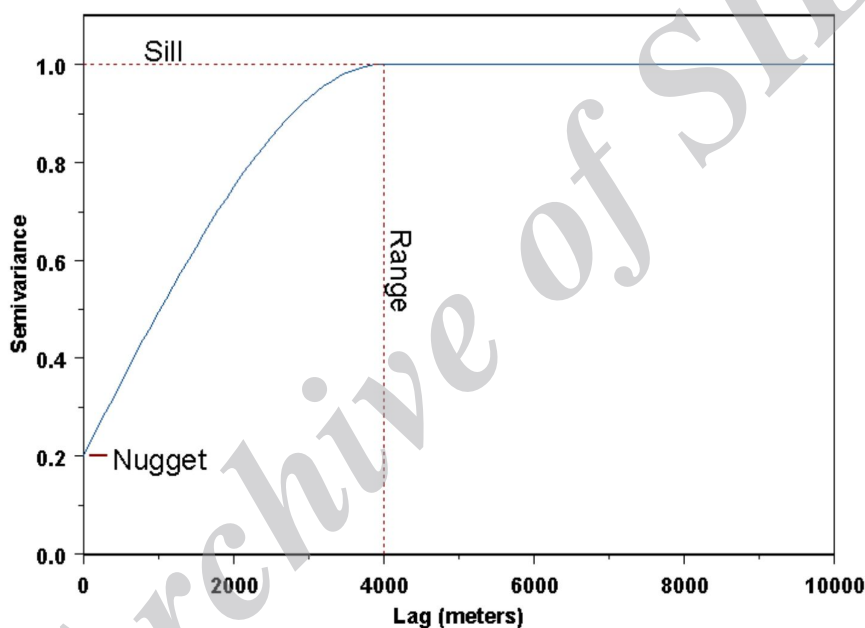
$$m_h = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} z(u_{\alpha} + h) \quad (33)$$

^۱ Covariance
^۲ Correlation

$$\sigma_0 = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - m_0]^2 \quad (34)$$

$$\sigma_h = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha + h) - m_h]^2 \quad (35)$$

برای رسم وریوگرام^۱ باید نیم‌واریانس در هر h محاسبه شود. برای نمونه، شکل ۳-۱ را در نظر بگیرید:



شکل ۳-۱: نمونه یک وریوگرام

Sill ماکزیمم مقدار نیم‌واریانس است؛ range فاصله‌ای است که نیم‌واریانس به اندازه sill می‌رسد و از آنجا به بعد همبستگی بین نقاط صفر می‌شود. برای تعیین شعاع همسایگی نقطه شبکه‌ای می‌توان از range استفاده کرد زیرا همبستگی بین نقاط از آن فاصله به بعد صفر است.

^۱ Variogram

فصل چهارم: معرفی ایستگاه‌ها و داده‌های
هواشناسی و وضعیت پیش‌بینی در سازمان
هواشناسی کشور

Archive of SID

۴-۱) مقدمه

در فصل حاضر، ابتدا به نحوه پیش‌بینی در سازمان هواشناسی کشور می‌پردازیم و سپس به دلیل استفاده از داده‌های موجود در ایستگاه‌ها برای شبکه‌بندی پیش‌بینی پس‌پردازش شده، انواع ایستگاه‌های موجود را معرفی می‌کنیم. ما در این تحقیق از اطلاعات ایستگاه‌های سینوپتیک استفاده کرده‌ایم. در انتها به علت نیاز به ارتفاع هر نقطه دلخواه در شبکه به معرفی داده‌های SRTM که داده‌های مربوط به ارتفاع می‌باشند می‌پردازیم.

۴-۲) پیش‌بینی در سازمان هواشناسی کشور

سازمان هواشناسی کشور عضو سازمان جهانی هواشناسی (WMO)^۱ می‌باشد. نرم افزاری به نام synergy از کشور فرانسه خریداری شده است که قابلیت دسترسی به چندین مدل پیش‌بینی را دارد. این نرم افزار قادر است در مدت زمان کوتاهی به مدل‌ها GFS و RPEG دسترسی داشته و خروجی را به صورت نقشه برای پیش‌بین‌ها نمایش دهد. پیش‌بین‌ها در جلسه discussion که روزانه برگزار می‌شود نتایج حاصل از نقشه‌ها و تجربه محلی خود را ترکیب کرده و پیش‌بینی نهایی را صادر می‌کنند. متأسفانه هیچ گونه فرایند پیش‌پردازش و پس‌پردازش در سازمان هواشناسی کشور انجام نمی‌شود. نمایی از مراحل صدور پیش‌بینی در شکل ۴-۱ آمده است.

۴-۳) وضعیت شبکه ایستگاه‌های هواشناسی کشور

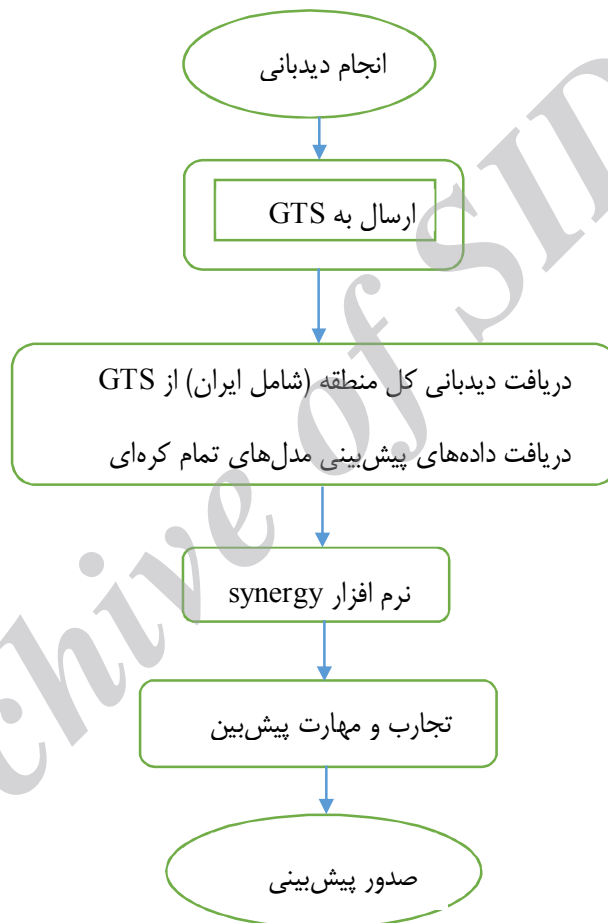
• ایستگاه‌های هواشناسی سینوپتیک:

ساعت به ساعت و یا هر ساعت یک بار حدود بیست پارامتر جوی را دیده‌بانی می‌کنند و از طریق شبکه مخابراتی هواشناسی کشور به مرکز پیش‌بینی تهران ارسال می‌دارند، از این نوع ایستگاه‌ها در زمان انقلاب ۵۰ ایستگاه ایجاد شده و در ابتدای سال ۱۳۶۴ به ۹۰ ایستگاه، ابتدای سال ۱۳۶۵ به ۹۱ ایستگاه و در پایان سال ۱۳۶۵ به ۱۱۴ و تا سال ۱۳۶۹ به ۱۲۲ ایستگاه در سال ۱۳۷۷ بالغ بر ۱۶۰ ایستگاه و در حال حاضر به ۳۹۵ ایستگاه رسیده است.

^۱ World Meteorological Organization

• ایستگاه‌های کاوش جو (جو بالا):

در هر شبانه روز دو مرتبه توسط بالن‌هایی که با گاز هیدروژن پر شده‌اند دستگاه‌هایی به نام رادیو سوند را به سطوح فوقانی جو حمل می‌کنند و مشخصات لایه‌های جو را دقیقه به دقیقه تا پایان عمل کاوش (سوندینگ) گزارش می‌کند. از این نوع ایستگاه‌ها قبل از پیروزی انقلاب اسلامی ۷ ایستگاه و سال ۱۳۶۴ تعداد ۱۰ ایستگاه و امروز ۱۱ ایستگاه در کشور مشغول کار هستند.



شکل ۴-۱: مراحل صدور پیش‌بینی در سازمان هواشناسی کشور

• ایستگاه‌های تحقیقاتی کشاورزی:

علاوه بر دیده‌بانی‌های سینوپتیک یک سری دیده‌بانی‌های فنولوژی و اندازه‌گیری‌های بیومتری روی ۱۷ محصول مهم کشور انجام می‌دهند و هم زمان یک سری مطالعه و تحقیق بر روی ارتباط عوامل جوئی با رشد

محصولات مختلف در ایستگاه به عمل می‌آورند. از این نوع ایستگاه‌ها قبل از انقلاب اسلامی ۲ ایستگاه وجود داشته که امروزه تعداد آنها به ۱۸ ایستگاه رسیده است.

- **ایستگاه‌های هواشناسی دریائی سینوپتیک:**

محلّی که پس از انقلاب به مجموعاً ۱۸ ایستگاه بر روی سکوها، جزایر و بنادر جنوب کشور تاسیس شده است. این ایستگاه‌ها در سواحل دریاها، اقیانوس‌ها و جزایر و کشتی‌ها بصورت ثابت و متحرک قرار دارند. علاوه بر اطلاعات جوی، اطلاعاتی در مورد دمای آب دریا، طول امواج و سرعت امواج و غیره را اندازه‌گیری می‌کنند.

- **ایستگاه‌های اقلیم شناسی:**

فقط ۳ نوبت در روز پارامترهای مهم جوی را دیده‌بانی و در دفاتر مخصوص ثبت می‌کنند و از این نوع در حال حاضر ۲۵۰ ایستگاه در کشور مشغول به کار هستند.

- **ایستگاه‌های باران سنجی:**

آمار مورد نیاز بیشتر مؤسسات و سازمان‌های تحقیقاتی کشور را در زمینه کشاورزی و آبیاری فراهم می‌کند پس از انقلاب به علت نیاز، رشد بیشتری داشته‌اند و به طوری که تعداد کل ایستگاه‌های باران سنجی کشور (مربوط به سازمان هواشناسی کشور) تا کنون به حدود ۱۹۷۵ رسیده است.

- **ایستگاه‌های تشعشع سنجی:**

که میزان کل تشعشع خورشیدی را بصورت تشعشع کلی و تشعشع پراکنده و بازتاب زمینی اندازه‌گیری و ثبت می‌کنند. تعداد این ایستگاه‌ها در سال ۱۳۶۴، ۲۱ ایستگاه بوده که امروز به ۲۵ ایستگاه رسیده است.

۴-۴) زمان و ساعت دیده‌بانی

برای اینکه بتوان عوامل جوی دیده‌بانی شده را از ایستگاه‌های متعدد هواشناسی دریافت و روی نقشه‌های سینوپتیک ترسیم و آنالیز نمود، در مرحله اول لازم است کلیه گزارشات در رأس ساعت معینی انجام شود و همزمان باشند و جهت رسیدن به این هدف بستگی به این دارد که تمام ایستگاه‌های دیده‌بانی از ساعت‌های محلی (وقت محلی) استفاده ننموده و مبدأ زمان معینی را در نظر بگیرند. برای این منظور از وقت و زمان متوسط نصف النهار گرینویچ UTC (زمان

هماهنگ شده بین المللی) استفاده می‌شود و کلیه ایستگاه‌های هواشناسی موظفند گزارش خود را رأس ساعت معین UTC تهیه و ارسال نمایند. اختلاف افق تهران با ساعت گرینویچ سه ساعت و ۳۱ دقیقه است ولی برای محاسبه وقت UTC در ایستگاه‌های هواشناسی ایران از ساعت رسمی کشور سه ساعت و ۳۰ دقیقه کسر می‌کنیم یعنی ساعت ۱۲ ظهر بوقت تهران، ساعت ۸:۳۰ صبح بوقت گرینویچ می‌باشد.

۴-۵) مراکز پیش بینی و اطلاعات جوی

مراکز مخابرات هواشناسی کشورهای مختلف همانگونه که اطلاعات کشور خود را به نقاط دیگر جهان ارسال می‌کنند اطلاعات کشورهای دیگر را نیز دریافت می‌کنند. این اطلاعات به مراکز پیش‌بینی فرستاده می‌شود. در مراکز پیش‌بینی هواشناسی دی‌کُد می‌شود یعنی از حالت رمز خارج می‌شود و روی نقشه‌های جغرافیایی مخصوصی که برای همین منظور تهیه شده پلات می‌شود. یعنی اطلاعات هر ایستگاهی در محل همان ایستگاه یادداشت می‌شود. سپس نقشه‌های تهیه شده آنالیز می‌شود. خطوط هم مقدار روی آن رسم می‌گردد. مناطق کم فشار و پر فشار، توده‌های مختلف هوا و جبهه‌های همراه با آنها و همچنین سمت و سرعت حرکت و تقویت و تضعیف سیستم‌ها شناسایی می‌شود. پس از آن با استفاده از این نقشه‌ها و تصاویر ماهواره‌ای و مدل‌های پیش‌بینی عددی که با استفاده از همین اطلاعات توسط کامپیوتر بدست می‌آید وضعیت جوی روزهای آینده پیش‌بینی می‌شود.

بعلت عدم شناسایی بعضی از عوامل که روی تقویت و تضعیف و حرکت سیستم‌های جوی تاثیر می‌گذارد هنوز هیچ یک از کشورهای جهان حتی کشورهای توسعه یافته با همه تجهیزات پیشرفته نتوانسته‌اند صحت پیش‌بینی هوا را به صد در صد برسانند.

یکی از عوامل مهمی که روی حرکت و فعالیت سیستم‌های جوی اثر می‌گذارد ناهمواری‌های زمین و رشته کوه‌ها می‌باشد. وجود سلسله جبال زاگرس و البرز در ایران تاثیر زیادی روی حرکت و فعالیت سیستم‌های جوی دارد. به عنوان مثال پیش‌بینی وضع هوا برای کشوری مثل انگلستان که رشته کوه ندارد به مراتب آسانتر از ایران است.

پیش‌بینی‌هایی که به این نحو صادر می‌شود از طریق رسانه‌های گروهی و صدا و سیما در اختیار مردم قرار می‌گیرد. پیش‌بینی‌های مخصوص برای فرودگاه‌ها و مسیر پرواز هواپیماها و بنادر کشتیرانی صادر و در اختیار آنها قرار داده می‌شود. پیش‌بینی‌ها نیز تا سه روز آینده برای امور کشاورزی صادر می‌شود از این پیش‌بینی‌ها برای آبیاری، سمپاشی، برداشت محصول و غیره استفاده می‌شود.

۴-۶) آمار تحقیقات هواشناسی

اطلاعات هواشناسی در موقع دیده بانی در دفاتر مخصوص ثبت می‌گردد. همزمان با دیده بانی بیشتر عوامل هواشناسی توسط دستگاه‌های ثبت بطور مداوم اندازه‌گیری و ثبت می‌گردد. بعنوان مثال دما، فشار، رطوبت هوا بیش از چهل سال قبل تا کنون در آرشیو هواشناسی موجود است و برای هر لحظه که نیاز باشد می‌توان به آرشیو مراجعه و آن را بدست آورد. ذکر این نکته ممکن است جالب باشد که حتی در زمان جنگ تحمیلی ایستگاه هواشناسی آبادان تعطیل نشد و در کار تهیه آمار ایستگاه گسستگی و خللی به وجود نیامد.

به هر حال دفاتر ثبت اطلاعات و گراف‌های دستگاه‌های ثبت در پایان هر ماه به مرکز کامپیوتر سازمان هواشناسی ارسال می‌گردد. در مرکز کامپیوتر اطلاعات وارد حافظه کامپیوتر گشته و پردازش می‌شود. اطلاعات پردازش شده به اشکال مختلف بر حسب نوع نیاز در اختیار استفاده کنندگان قرار داده می‌شود. این اطلاعات برای شناخت رفتار اتمسفر و علل تغییرات آن و شناسایی اقلیم نقاط مختلف، طرح‌های توسعه‌ای مانند سد سازی، شهرسازی، کشاورزی، دامداری، شیلات و احداث فرودگاه‌ها و غیره بکار برده می‌شود.

هم اکنون کمیته‌های هشت‌گانه تحقیقاتی کشور مانند کمیته کشاورزی، آبشناسی، هواشناسی دریایی، اقیانوس شناسی فیزیکی و آلودگی جو با همکاری کارشناسان سازمان هواشناسی و اساتید دانشگاه‌ها و محققین کشور در این زمینه مشغول مطالعه و تحقیق هستند.

۴-۷) داده‌های رقومی ارتفاعی زمین SRTM

حروف اختصاری DEM و DTM مجموع ای از داده‌های رقومی در مدل سازی سطح توپوگرافی (سطحی که داده‌های ارتفاعی را نشان می‌دهد) بکار می‌رود. به منظور مدل سازی دقیق یک سطح، باید تعداد زیادی از مشاهدات را ثبت کرد. از آنجا که این امر میسر نیست، مدل سطح بطور تقریبی به یک سطح پیوسته همانند است که با استفاده از تعداد محدودی از مشاهدات فراهم می‌آید. بنابراین، تعداد مناسبی از مشاهدات باید به همراه موقعیت جغرافیایی آنها انتخاب شوند.

مفهوم اولیه مدل ارتفاع رقومی در اواخر دهه ۱۹۵۰ میلادی از ایالات متحده آمریکا آغاز شد. از آن زمان تاکنون پیشرف‌های قابل توجهی بویژه در مورد روش‌های گردآوری و پردازش اطلاعات زمینی، حاصل گردید. اصطلاح مدل رقومی زمین عبارتست از بهره‌گیری از اطلاعات ارتفاعی نیمرخ (مقطع) به منظور تشریح شکل زمین. این اطلاعات

ارتفاعی می تواند هم منظم (رئوس شبکه قائم الزاویه) و هم غیر منظم باشد. واژه های دیگری هم معمولاً به عنوان معادل مدل رقومی زمین مورد استفاده قرار می گیرند که تعدادی از آنها عبارتند از: مدلسازی رقومی منطقه^۱، مدلسازی رقومی ارتفاع^۲، مدلسازی رقومی ارتفاع^۳، مدلسازی رقومی زمین^۴، مدلسازی رقومی ارتفاع منطقه^۵. مدل ارتفاع رقومی و دیگر زیرشاخه های منشعب از آن (شاخص های شیب، حوضه، شبکه، منحنی، توپوگرافیکی و...) پارامترهای مهمی برای استخراج اطلاعات یا ارزیابی هرگونه پردازش مورد نیاز تحلیل زمین هستند. از این زیرشاخه ها برای موارد متعدد استفاده می شود؛ مدل جریان آب، ارزیابی میزان رواناب، شبیه سازی سیل و مدیریت، حرکت توده ها، تحلیل کاربری زمین، تهیه نقشه ناهمواری ها، خطرات آتشفشانی، مجازی سازی زمین و نقشه کشی، مطالعات اقلیمی و آب و هوا شناختی. برون داد این مدل ها بستگی به صحت مدلسازی رقومی ارتفاع دارد.

پروژه Shuttle Radar Topographic Mission توسط ناسا جهت تهیه مدل رقومی ارتفاعی ۸۰ درصد از جهان کلید خورد. این داده ها به صورت رایگان با قدرت تفکیک تقریبی ۹۰ متر انتشار یافته است. این داده ها مربوط به توپوگرافی زمین است که بصورت مخفف SRTM خوانده می شود و نشان از یک تلاش تحقیقاتی جهانی برای رسیدن به مدل ارتفاعی رقومی از سطح زمین دارد. این داده ها تقریباً دارای یک پوشش جهانی از ۵۶ درجه جنوبی تا ۶۰ درجه شمالی می باشند.

مدل ارتفاعی بدست آمده بصورت قطعاتی که هر کدام دارای طول یک درجه و عرض یک درجه می باشند ارائه شده اند. دقت هر سلول یک آرک ثانیه (حدود ۳۰ متر) است ولی چنین دقتی فقط برای ایالات متحده بیرون داده شده است و برای بقیه جهان بصورت دو آرک ثانیه (حدود ۹۰ متر) ارائه شده است. مدل ارتفاعی بدست آمده از داده های SRTM در سیستم های مختلف اطلاعات جغرافیائی می تواند استفاده شود.

مزایای داده های SRTM چیست؟

داده های SRTM چند مزیت مهم دارند. اول آنکه این گونه داده های DEM در مناطق وسیع که از یک منبع به دست آمده باشند مثل (SRTM)، بسیار مورد نیاز هستند؛ زیرا پایدار بوده و در مناطق بزرگ قابل استفاده هستند. در حالی که

^۱ Digital Terrain Modelling

^۲ Digital Elevation Modelling

^۳ Digital Height Modelling

^۴ Digital Ground Modelling

^۵ Digital Terrain Elevation Modelling

^۶ National Aeronautics and Space Administration (NASA)

دیگر DEM های با قدرت تفکیک بالا از منابع متغیری مثل زوج تصاویر ماهواره‌ای به دست می‌آیند. همچنین از آنجا که اشعه رادار از ابر عبور می‌کند، سنجنده‌های راداری محدودیتی از نظر پوشش ابر ندارند. برای تهیه داده‌های سه بعدی از روشی برای تداخل سنجی راداری استفاده می‌کنند. در این روش دو تصویر راداری از دو نقطه با اختلاف مکانی کم گرفته می‌شوند. از اختلاف میان این تصاویر، ارتفاع نقاط زمین یا تغییرات آن قابل محاسبه است. تداخل سنجی، مطالعه الگوهای تداخلی است که از ترکیب دو مجموعه سیگنال راداری حاصل می‌شوند. اگر تا کنون یک گودال آب با لایه نازکی از روغن روی آن دیده باشید، احتمالاً نوارهای رنگی روی سطح آن را ملاحظه نموده‌اید. این نوارهای رنگی به وسیله شعاع‌های نور منعکس شده از سطح صاف روغن و آب زیر آن که الگوهای تداخلی را ایجاد کرده‌اند، به وجود می‌آیند. برای اخذ دو تصویر راداری از دو منظر متفاوت، سیستم SRTM، شامل یک آنتن رادار در داخل دستگاه و یک آنتن رادار ثانویه نصب شده در انتهای یک دکل به طول ۶۰ متر (۱۹۵ فوت) در خارج دستگاه است. SRTM از روش تداخل سنجی با بازده ثابت استفاده می‌کند. یعنی دو مجموعه داده راداری در یک زمان اخذ می‌شوند و آنتن‌هایی که این داده‌ها را جمع‌آوری می‌کنند در فاصله ثابتی از هم قرار گرفته‌اند.

ماهواره‌های دارای سیستم راداری

کلمه رادار از عبارت "Radio Detection & Ranging" گرفته شده و اولین بار بطور آزمایشی در سال‌های ۱۹۲۵ و ۱۹۲۶ در کشورهای انگلستان و ایالات متحده آمریکا از این سیستم استفاده شد. سپس تا ۱۹۶۰ در اهداف نظامی بکار گرفته شد، پس از آن به منظور مطالعات زمینی در هواپیما تعبیه گردید. در حال حاضر نیز به عنوان یک سنجنده فعال در سکوی فضایی مورد استفاده قرار گرفته است. از سری ماهواره‌هایی که در آن از سیستم راداری استفاده شده است می‌توان ماهواره‌های Radar sat (آژانس فضایی کانادا ۱۹۹۵)، Sea Sat (آمریکا ۱۹۸۷)، فضایی شاتل Shuttle (آمریکا ۱۹۸۶ و ۱۹۸۴)، ERS (سازمان فضایی اروپا ۱۹۹۰) و Jers (ژاپن ۱۹۹۳) را نام برد.

فصل پنجم: بحث و نتیجه‌گیری

Archive of SID

۵-۱) مقدمه

در این فصل، ابتدا به نتایج حاصل از پس پردازش و صدور پیش بینی احتمالی در ایستگاه‌های دیدبانی پرداخته می‌شود. پس پردازش با استفاده از روش کمینه کردن CRPS انجام شده است. برای کمینه کردن این تابع از الگوریتم فراابتکاری^۱ PSO استفاده شده است. بنابراین الگوریتم PSO توضیح داده می‌شود. در ادامه نتایج حاصل از مقایسه سه روش BCDG, GIDS و kriging به منظور پس پردازش بر نقاط یک شبکه منظم نشان داده خواهد شد.

۵-۲) الگوریتم حرکت جمعی ذرات (PSO)

در طبیعت، برخی از جانوران دارای زندگی اجتماعی می‌باشند که از آن جمله می‌توان به زندگی گروهی پرندگان، ماهی-ها، زنبورها، مورچگان و ... اشاره کرد. زندگی اجتماعی اینگونه از جانوران، الهام بخش پیدایش تعدادی الگوریتم‌های بهینه‌سازی فراابتکاری مانند الگوریتم‌های ژنتیک، کلونی مورچگان، زنبور عسل^۲ و حرکت جمعی ذرات شده است. شاید ظاهراً این طور به نظر برسد که نام الگوریتم حرکت جمعی ذرات، که الگوریتم انتخابی این تحقیق می‌باشد، ربطی به زندگی گروهی جانوران نداشته باشد؛ ولی اساس کار این الگوریتم، تقلیدی از حرکت دسته جمعی پرندگان، وقتی در جستجوی غذا هستند، می‌باشد.

در سال ۱۹۸۳، ریوس (Reeves, 1983) برای مدل‌های پویا که نمی‌توان آنها را به صورت چند ضلعی یا مسطح در نظر گرفت، سیستم ذرات^۳ را پیشنهاد کرد. در این سیستم، ذرات تشکیل دهنده آتش، دود، آب و ابر از هم مستقل هستند و حرکتشان به وسیله یک سری قوانین می‌باشد. رینولدز (Reynolds, 1987) برای شبیه‌سازی رفتار جمعی پرندگان، از سیستم ذرات استفاده کرد و دو سال بعد از آن، در یک نوع مشابه، هپنر و جرناندر (F. Heppner, 1990)، جستجوی لانه‌ای مناسب برای پرندگان را شبیه‌سازی کردند. هر دو این مدل‌ها باعث ایجاد قوانینی شدند که چگونگی حرکت ذرات را در یک فضای جستجو، مشخص می‌کردند. بعدها، این قوانین به عنوان شالوده اصول اساسی PSO در

^۱ PARTICLE Swarm Optimization

^۲ Bee Algorithm

^۳ Particle system

نظر گرفته شدند (Dorigo, 2008). بنابراین PSO نامش را از سیستم ذرات استفاده شده در دو مدل مذکور گرفته است؛ ولی روش کار آن الهام گرفته از حرکت دسته جمعی پرندگان می باشد. در حرکت جمعی پرندگان (swarm)، یک رابطه پیچیده در رفتار حرکتی آنها وجود دارد. پرندگان در حرکتشان به سمت غذا از تصمیم دیگر پرندگان نیز استفاده می کنند تا نتیجه بهتری کسب کنند. در واقع این خصوصیت یک زندگی اجتماعی است. ما انسان ها نیز برای تصمیم گیری در امور مهم زندگی، برای رسیدن به یک نتیجه خوب و قابل قبول، سعی می کنیم تا علاوه بر در نظر گرفتن عقاید شخصی خود و تجربه هایی که تاکنون کسب کرده ایم، از تجربه های دیگران نیز استفاده کنیم. در واقع در یک زندگی اجتماعی، به علت تعامل متقابل، افراد با به اشتراک گذاشتن عقاید و نگرش های خود می توانند دستیابی به بهترین نتایج را آسانتر کنند.

۵-۲-۱) تعریف عملکرد PSO

PSO، الگوریتم جستجوی فراابتکاری مبتنی بر جمعیتی از پرندگان است که اولین بار توسط ابرهارت و کندی (J. Kennedy R. E., 1995)، برای بهینه سازی توابع غیرخطی پیوسته ارائه شده است. در این الگوریتم گروهی از پرندگان به قصد پیدا کردن غذایی که از محل آن بی خبر هستند، در یک فضای مورد نظر پرواز می کنند. پرندگان در ابتدا از یک مکان تصادفی، پروازشان را آغاز کرده و در مراحل بعد، سعی می کنند از پرندگانی که به غذا نزدیکترند پیروی کنند. به عبارتی دیگر آنها برای حرکت در فضای جستجو، در هر مرحله با بروزرسانی مکان و سرعت خود، به غذای مورد نظر نزدیکتر می شوند. برای تعیین مکان جدید هر پرنده، باید سرعتش را که نشان دهنده جهت حرکت است، محاسبه نمود. سرعت هر پرنده نیز با در نظر گرفتن فاصله خودش و دیگر پرندگان تا غذا و انتخاب بهترین مکانی که در آن کمترین فاصله تا غذا مشاهده شده است، بروزرسانی می شود. در واقع میزان فاصله پرندگان تا غذا، ملاک تصمیم گیری آنهاست. فاصله آنها تا غذا، نشان دهنده میزان شایستگی آنها می باشد که هر چه فاصله کمتر، شایستگی پرنده برای یافتن غذا بیشتر است.

برای تطبیق هر یک از این مفاهیم با مسأله برنامه ریزی مورد نظر، باید توجه داشت که در مدل PSO، هر ذره که نماینده یک راه حل از مسأله می باشد، دارای حرکتی معادل با حرکت یک پرنده در حرکت جمعی پرندگان است. جواب مسأله که یکتاست، همان غذایی است که پرندگان در جستجوی آن هستند و در نهایت، شایستگی هر پرنده را تابع هدف مسأله مشخص می کند. بنابراین به طور کلی اساس کار PSO بدین ترتیب است که در ابتدا به تعداد ذرات که

دلخواه می باشد، جواب های اولیه کاملاً تصادفی در نظر گرفته می شود. هریک از این جواب ها دارای مکان هایی در فضای قابل قبول مسأله هستند؛ که این مکان ها در هر مرحله به وسیله متغیری تحت عنوان سرعت، بروزرسانی می شوند. در واقع در هر مرحله جواب های جدید دیگری تولید و بررسی می شوند و بهترین جواب در حافظه ذرات باقی می ماند این فرایند طی تعداد مراحل دلخواه تکرار می شود و بهترین جوابی که تاکنون در حافظه ذرات ضبط شده است به عنوان جواب بهینه اعلام می گردد.

۵-۲-۲) فرمول سرعت

تعیین سرعت ذرات در PSO و به تبع آن، تعیین جهت حرکت و مکان جدید ذرات، باید به طور هوشمندانه انجام شود تا در هر مرحله جواب هایی تولید شوند که نهایتاً از بین آنها بهترین جواب نزدیک به بهینگی انتخاب گردد. فرمول سرعت و مکان در PSO به این صورت است:

$$v(t+1) = w(t)v(t) + c_p u(x(t) - x_{pbest}) + c_g u(x(t) - x_{gbest}) \quad (36)$$

$$x(t+1) = x(t) + v(t+1) \quad (37)$$

t : اندیس تکرار؛ $t=1, \dots, T$ که T تعداد تکرارهاست و مقدارش دلخواه می باشد.

u : عدد تصادفی از توزیع یکنواخت^۱ در بازه $[0, 1]$

c_p و c_g : ضرایب یادگیری

$w(t)$: وزن لختی^۲ در تکرار t ام

$v(t)$: سرعت ذره در تکرار t ام

$x(t)$: مکان ذره در تکرار t ام

x_{pbest} : بهترین مکانی که تا به حال توسط خود ذره شناسایی شده است که به آن بهترین تجربه شخصی گفته می شود.

x_{gbest} : بهترین مکانی که تا به حال توسط همه ذرات شناسایی شده است که به آن بهترین تجربه گروه گفته می شود.

^۱ Uniform distribution

^۲ Inertia weight

۵-۲-۳) روش‌های بروزرسانی وزن ایستایی

وزن لختی، ضریبی برای در نظر گرفتن مقداری از سرعت قبلی ذره می‌باشد؛ که تعیین مقداری مناسب برای این ضریب، بر عملکرد الگوریتم PSO تأثیرگذار است. بنابراین تعدادی از محققان، سعی در ارائه روش‌هایی برای تعیین مقادیری مناسب برای وزن لختی داشته‌اند. از مهمترین روش‌ها برای بدست آوردن مقادیر مناسب وزن لختی، یک تابع نزولی خطی است که در تحقیق ابره‌ارت و شی^۱ (Y.H. Shi, 1998)، در سال ۱۹۹۸ ارائه شده است. رن و گائو^۲ (Y. Gao, 2007) در تحقیقشان در سال ۲۰۰۷، یک تابع نزولی غیر خطی را برای وزن لختی پیشنهاد کردند. در سال ۲۰۰۸، آروم‌گام و رائو^۳ (M.S. Arumugam, 2008) برای بروزرسانی مقدار وزن ایستایی، از مقادیر pbest و gbest در یک تکرار خاص، استفاده نمودند. یک پارامتر که خودش را با شرایط مسأله مطابقت دهد، مناسب به نظر می‌رسد. این ایده آی و وراتاس^۴ (T.J. Ai) در سال ۲۰۰۸ بود که باعث شد تا آنها روشی ارائه دهند که پارامترها بتوانند خودشان را بصورت اتوماتیک با شرایط سازگار کنند؛ که جواب بدست آمده از این روش، بهتر از روش‌های قبل بوده است اما زمان محاسبات بیشتری را می‌طلبد.

۵-۲-۴) انواع PSO

محققان بسیاری، در جهت بهبود عملکرد الگوریتم PSO، روش‌های مختلفی برای بروزرسانی سرعت و مکان ذرات ارائه داده‌اند؛ که در این بخش، به بعضی از آنها اشاره می‌شود.

- **PSO دودوئی گسسته^۵:** در ابتدا، PSO برای جستجو در فضای پیوسته ارائه شد؛ اما تعدادی مسائل هستند که باید در فضای گسسته بررسی شوند. این دلیلی شد تا ابره‌ارت و کندی (J. Kennedy R. E., 1997)، نوع دیگری از این الگوریتم را با نام PSO دودوئی گسسته ارائه دهند. در این مدل، مکان ذره گسسته است اما سرعت آن مقداری پیوسته می‌باشد. مکان ذره تحت شرایط خاصی مقادیر ۰ و ۱ را اختیار می‌کند و سرعت نیز مانند روش PSO استاندارد محاسبه می‌شود.

^۱ Shi

^۲ Ren & Gao

^۳ Arumugam & Rao

^۴ Ai & Voratas

^۵ Discrete binary

- PSO با وزن نزولی خطی (LDWPSO)^۱: در سال ۱۹۹۸، ابرهات و شی (Y. Shi, 1998) از ضریب ایستایی نزولی خطی، برای محاسبه سرعت در رابطه (۱-۲) استفاده کرده‌اند.
- PSO با ضرایب انقباضی^۲: این نوع PSO، توسط کلرک و کندی (M. Clerc, 2002) بیان شد. ضرایب انقباضی به عنوان خروجی یک آنالیز نظری از پویاشناسی حرکت جمعی (swarm) معرفی شده‌اند. در این روش برای بروزرسانی سرعت، از یک ماتریس قطری استفاده می‌شود که عناصر روی قطر اصلی آن با یک فرمول خاص محاسبه می‌گردند.
- PSO بی استخوان^۳: در سال ۲۰۰۳، کندی (Kennedy, 2003) برای بروزرسانی متغیرهای سرعت و مکان ذره، از یک توزیع نرمال که میانگین و واریانس آن از pbest و gbest بدست می‌آید، استفاده کرد.
- GLNPSO: ویراماچانی و همکاران (K. Veeramachaneni, 2003) روشی ارائه دادند که در آن علاوه بر pbest و gbest، از بهترین تجربه محلی (lbest)^۴ و بهترین تجربه همسایگان (nbest)^۵ برای بروزرسانی سرعت استفاده شده است؛ که توضیح این روش در فصل بعد خواهد آمد.
- PSO کاملاً مطلع (FIPSO)^۶: در PSO استاندارد، یک ذره به سمت بهترین همسایه‌اش جذب می‌شود؛ ولی در سال ۲۰۰۴، مندرس و همکاران (R. Mendes, 2004) در تحقیقشان پیشنهاد کردند که از اطلاعات همه همسایه‌ها برای بروز رسانی سرعت استفاده شود تا بدین طریق، حرکت ذرات کاملاً از روی آگاهی باشد.
- PSO جدید (NPSO)^۷: در این روش هرذره، مکانش را بر اساس بدترین تجربه قبلی خودش و گروه تنظیم می‌کند. استراتژی این نوع PSO که توسط یانگ و سیمون (Yang & Simon, 2005) ارائه شده است، دور شدن از بدترین جواب‌های قبلی می‌باشد و مقادیر پارامترها مانند PSO استاندارد تعیین می‌گردند.

^۱ Linear Decreasing Weight PSO

^۲ Constriction coefficient

^۳ Bare Bones PSO

^۴ Local best

^۵ Neighbour best

^۶ Fully Informed PSO

^۷ New PSO

- PSO **قبیله‌ای**^۱: این روش توسط چن و همکاران (K. Chen, 2006)، مطرح گردید؛ به طوری که آنها ذرات را به ۲ لایه و رویه بهینه‌سازی را به ۳ فاز، طی روش خاصی تقسیم‌بندی می‌کنند. این روش نسبت به PSO استاندارد، به پارامترها و محاسبات بیشتری نیاز دارد.
- PSO **مرکزی**^۲: لیو و همکاران (Liu et. al., 2007) از یک ذره اضافی به نام ذره مرکزی استفاده کردند. مکان این ذره با میانگین گرفتن از مکان همه ذرات بدست می‌آید. این ذره اضافی سرعت ندارد؛ اما به عنوان یک ذره معمولی در هر مرحله تلقی می‌شود. بروزرسانی سرعت بقیه ذرات مانند روش LDWPSO محاسبه می‌گردد.
- PSO **اصلاح شده**^۳: در سال ۲۰۰۷، قمشه و همکاران (Ghomsheh, Shoorehdeli, & Teshnehlab, 2007) با ترکیب کردن عامل GA با الگوریتم PSO، رویکرد جدیدی را در این زمینه ایجاد کردند. در تحقیق آنها هدف از بین بردن بدترین ذره و جایگذاری آن با یک ذره جدید می‌باشد؛ که چگونگی برآورد بدترین ذره و تولید ذره جدید، موضوع مورد بحث تحقیقشان است.
- PSO **پویا** (DPSO)^۴: در این روش، جیائو و همکاران (B. Jiao, 2008) یک ضریب وزن لختی پویا را جهت بروزرسانی سرعت ذرات ارائه داده‌اند.

۳-۵) پیش‌بینی احتمالی دمای سطح زمین روی ایستگاه‌ها

یک تابع چگالی احتمال نرمال برای پیش‌بینی احتمالی دمای دومتری برآورد شده است. میانگین تابع چگالی احتمال نرمال، میانگین وزنی اعضای سامانه همادی است که اریبی خطای آنها حذف شده و واریانس آن یک ترکیب خطی از واریانس اعضای همادی می‌باشد. با کمینه کردن امتیاز احتمالاتی رتبه‌ای پیوسته (CRPS) که سنج‌های برای راستی-آزمایی پیش‌بینی احتمالی است، میانگین و واریانس تابع توزیع پیش‌بینی برآورد شده‌اند. از الگوریتم حرکت جمعی ذرات (PSO) برای کمینه کردن CRPS در دوره آموزش استفاده شده است. داده‌های مسأله، پیش‌بینی ۲۴ ساعته دمای

^۱ Tribe PSO
^۲ Central PSO
^۳ Modified PSO
^۴ Dynamic

دومتري روی ايستگاه‌هاي همديد کشور در فصل‌هاي پاييز و زمستان در سال‌هاي ۲۰۱۱ و ۲۰۱۲ با پيکربندي‌هاي مختلف مدل WRF است. نتايج بدست آمده، خطاي بسيار کمتری نسبت به خروجی منفرد خام مدل دارد به طوری که میانگین بدست آمده در تابع چگالی احتمال در مقایسه با میانگین اعضای سامانه همادی، دارای میانگین مجذور خطا و میانگین قدرمطلق خطای کمتری است؛ همچنین مقدار CRPS آن حدود ۱/۲ و تابع چگالی احتمال بدست آمده نسبت به سامانه همادی کالیبره تر است

۵-۳-۱) تعريف مسأله و نتايج

با توجه به اینکه در اجرای الگوریتم‌های فراابتکاری، تابع هدف و محدودیت‌ها به طور مفهومی اعمال می‌شوند لزومی به ارائه مدل ریاضی مسأله وجود ندارد و تابع هدف و محدودیت‌ها در ذیل تعریف می‌شوند.

۵-۳-۲) تابع هدف

همانطور که پیشتر گفته شد تابع امتیاز احتمال رتبه‌ای پیوسته مدلی از تابع میانگین مجذور خطا است. بنابراین هدف مسأله، تخمین ضرایب میانگین و واریانس تابع چگالی احتمال با کمینه کردن تابع CRPS می‌باشد.

۵-۳-۳) محدودیت

با توجه به اینکه ضرایب b_1, \dots, b_m نشان دهنده میزان عملکرد اعضای سامانه همادی می‌باشند تنها محدودیت مسأله، تخمین این ضرایب در بازه $[0, 1]$ است.

۵-۳-۴) داده‌های مسأله

در این تحقیق، سامانه همادی شامل ۹ عضو است که هر عضو آن یک پیش‌بینی ۴۸ ساعته‌ی دمای دومتري با پيکربندي‌هاي فیزیکی مختلف می‌باشد که در جدول ۵-۱ مشخصات پيکربندي‌ها آورده شده است. برای پیش‌بینی‌ها از مدل WRF^۱ (Skamarock, et al., 2008) (یکی از متداولترین مدل‌های عددی پیش‌بینی وضع هوا) با مقادیر اولیه داده‌های GFS^۲ استفاده شده است. داده‌های دیدبانی از سازمان هواشناسی کشور در دسترس می‌باشد. بازه زمانی از روز چهارم ماه سپتامبر سال ۲۰۱۱ تا روز دوازدهم ماه فوریه ۲۰۱۲ میلادی برای ۳۱ مرکز استان‌ها در نظر گرفته شده است که جمعاً ۱۶۰ روز می‌باشد.

^۱ Weather Research and Forecasting

^۲ Global forecast System

جدول ۵-۱: پیکربندی های فیزیکی سامانه همادی

عضو همادی	لایه مرزی	لایه سطحی	سطح زمین	فیزیک همرفت	خرد فیزیک	فیزیک تابش موج بلند/ موج کوتاه
۱	MYJ TKE	MOJ	Unified Noah	KF	Lin	Goddard/ rrtm
۲	MYJ TKE	MOJ	Unified Noah	KF	Kessler	Dudhia/rrtm
۳	MYJ TKE	MOJ	Unified Noah	GD ensemble	Kessler	Dudhia/rrtm
۴	YSU	MM5 MO	Unified Noah	KF	WSM 5- class	Dudhia/rrtm
۵	MYJ TKE	MOJ	Unified Noah	BMJ	WSM 5- class	Dudhia/rrtm
۶	YSU	MM5 MO	Unified Noah	KF	Ferrier	Dudhia/rrtm
۶	MYJ TKE	MOJ	Unified Noah	BMJ	Ferrier	Dudhia/rrtm
۸	YSU	MM5 MO	RUC	GD ensemble	Ferrier	Dudhia/rrtm
۹	MYJ TKE	MOJ	Unified Noah	GD ensemble	Ferrier	Dudhia/rrtm

۵-۳-۵) پارامترهای الگوریتم PSO

انتخاب مجموعه ای خوب از پارامترها، در جواب بدست آمده بی تاثیر نمی باشد. این مقادیر با چندین بار آزمایش مسأله مورد نظر و انتخاب مقادیر مناسب به طور تجربی بدست می آیند که تنها برای همان مسأله مناسب هستند و نمی توان

آن را به مسائل دیگر تعمیم داد و از درستی جواب‌های بدست آمده مطمئن بود. در این تحقیق، مقدار پارامترهای استفاده شده به قرار ذیل است:

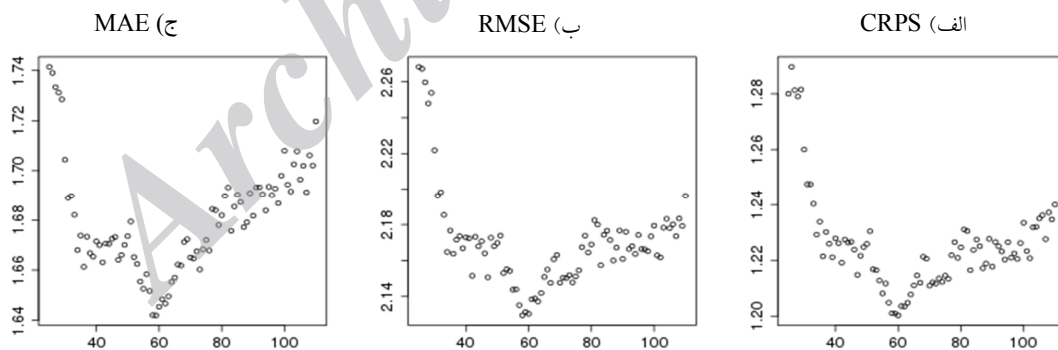
تعداد ذرات (I): ۵۰، تعداد تکرارها (T): ۱۰۰، وزن ایستایی (w): ۰/۵.

ضریب شتاب بهترین تجربه شخصی (C_p): ۲

ضریب شتاب بهترین تجربه گروهی (C_g): ۲

۵-۳-۶) نتایج بدست آمده

وقتی از یک دوره آموزش مثلاً ۶۰ روزه صحبت می‌شود، منظور یک دوره آموزش لغزان می‌باشد که بلافاصله ۶۰ روز قبل از روز پیش‌بینی را شامل می‌شود. اما طول دوره آموزش برای تخمین ضرایب رگرسیون چقدر باشد؟ برای پاسخ دادن به این سؤال، دوره‌های آموزش ۲۵، ۲۶، ...، ۱۱۰ در نظر گرفته و با مقایسه سنج‌های مختلف خطا، دوره آموزش مناسب تعیین شد. به این صورت که برای در اختیار داشتن یک دوره آزمون یکسان برای هر دوره آموزش، ۱۱۰ روز اول را کنار گذاشته و ۵۰ روز انتهایی از مجموعه داده‌ها را به عنوان دوره آزمون در نظر گرفتیم که بازه زمانی از روز بیست و پنجم ماه دسامبر سال ۲۰۱۱ تا روز دوازدهم ماه فوریه ۲۰۱۲ میلادی است. سپس CRPS، میانگین قدرمطلق خطا (MAE)^۱ و ریشه میانگین مربع خطای (RMSE)^۲ داده‌های آزمون را با طول‌های ۲۵ تا ۱۱۰ روزه برای تعیین بهترین طول دوره آموزش محاسبه و در شکل ۵-۱ نشان داده شده‌اند.



شکل ۵-۱: مقایسه طول دوره آموزش: الف) امتیاز احتمال پیوسته (CRPS) ب) RMSE پیش‌بینی قطعی وابسته

به تابع چگالی احتمال ج) MAE پیش‌بینی قطعی وابسته به تابع چگالی احتمال

^۱ Mean Absolute Error
^۲ Root Mean Squared Error

در شکل ۵-۱، محور افقی طول دوره آموزش و محور عمودی میزان خطا می باشد. در این قسمت (الف) مقدار میانگین CRPS در دوره آزمون، در (ب) RMSE و در (ج) مقدار MAE با در نظر گرفتن میانگین تابع چگالی نرمال در هر روز به عنوان یک پیش بینی قطعی، نشان داده شده است. همانطور که از شکل ۵-۱ دیده می شود مقدار خطا با دوره آموزش ۵۰ تا ۷۰ روزه کمترین مقدار خود را دارد. بنابراین در تحقیق حاضر طول دوره آموزش ۶۰ روز در نظر گرفته شده است. از طرف دیگر طول دوره آزمون ۵۰ روز می باشد. به گونه ای که در کل ۱۶۰ روز با داده های موجود، ۵۰ روز آخر به عنوان دوره آزمون و برای هر روز ۶۰ روز بلافاصله قبل از آن به عنوان داده های دوره آموزش لحاظ شده است. به بیان دیگر دوره آموزش یک پنجره ۶۰ روزه و لغزان است و CRPS را به صورت لغزان برای دوره آزمون محاسبه نموده و میانگین آنها را به عنوان امتیاز بهینه در نظر می گیریم.

MAE و RMSE برای میانگین ساده اعضای همادی در هر دو حالت خام و پس پردازش شده با روش میانگین لغزان^۱ و دوره آموزش ۱۵ روزه و همچنین برای اعضای همادی ۲ و ۸ (به عنوان مثال) و میانگین تابع چگالی احتمال محاسبه شده و در جدول ۵-۲ با یکدیگر مورد مقایسه قرار گرفته اند. همانطور که ملاحظه می شود خطا به میزان قابل توجهی کاهش یافته است.

جدول ۵-۲: مقایسه خطای پیش بینی های قطعی

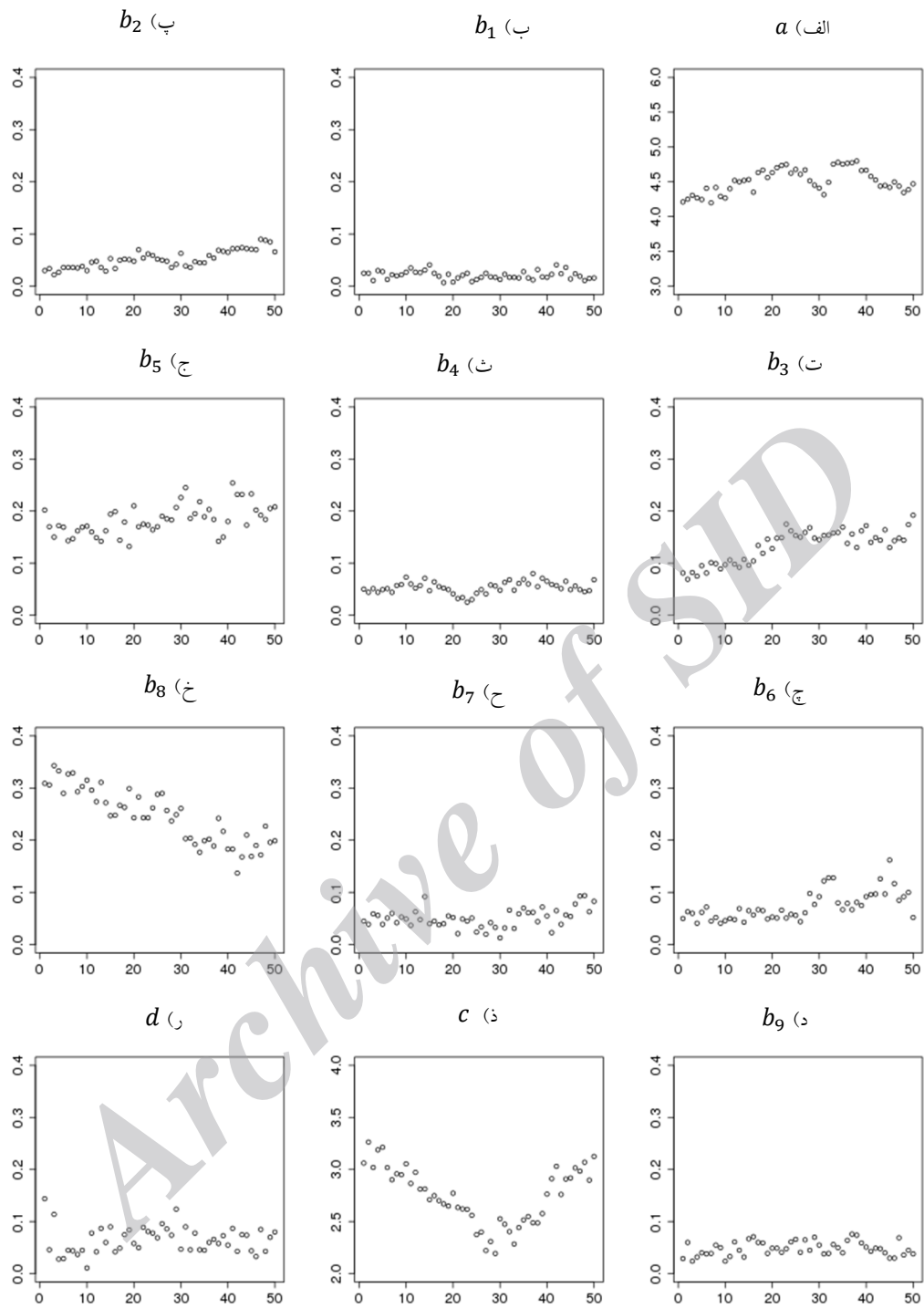
RMSE	MAE	
۷/۵۱	۶/۳۸	عضو ۲
۵/۰۴	۴/۰۷	عضو ۲ با اریبی تصحیح شده
۴/۴۲	۳/۶۱	عضو ۸
۳/۰۴	۲/۴۵	عضو ۸ با اریبی تصحیح شده
۴/۳۸	۳/۹۷	میانگین اعضای همادی خام
۲/۹۷	۲/۵۴	میانگین اعضای همادی با اریبی تصحیح شده
۲/۱۴	۱/۶۴	میانگین تابع چگالی احتمال نرمال

^۱ Moving Average

ضرایب تخمین زده شده در شکل ۵-۲ نشان داده شده‌اند به طوری که محور افقی، روزهای دوره آزمون و محور عمودی مقدار ضرایب می‌باشد. همانطور که مشاهده می‌شود عضوهای ۳، ۵ و ۸ دارای وزن‌های بیشتری هستند که نشان دهنده اهمیت بیشتر آنها نسبت به سایر اعضا است.

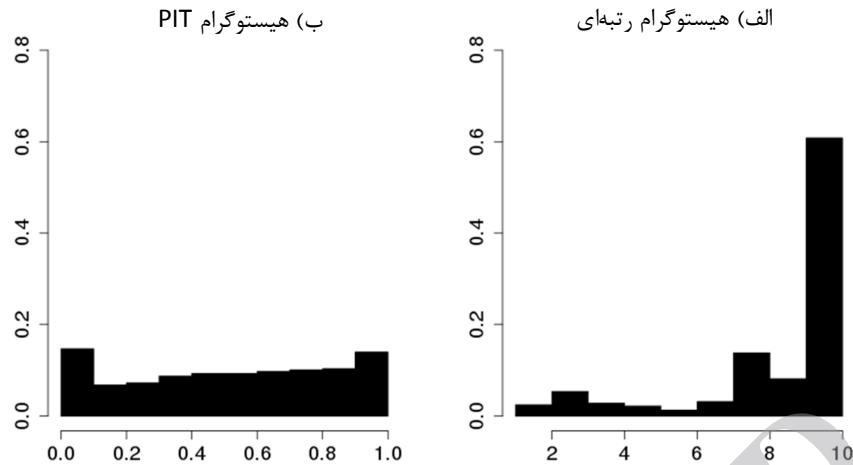
هیستوگرام رتبه‌ای^۱ (Hamill, 2001) روشی برای ارزیابی کالیبره بودن سامانه همادی می‌باشد؛ به این صورت که رتبه مشاهدات در میان اعضای سامانه همادی تعیین و در طول دوره آزمون میزان فراوانی نسبی رتبه‌ها محاسبه و هیستوگرام آنها رسم می‌شود. هیستوگرام^۲ PIT (Dawid, 1984) روشی برای ارزیابی کالیبره بودن توزیع احتمال می‌باشد. برای رسم هیستوگرام PIT، فراوانی نسبی مقدار توزیع تجمعی احتمال به ازای مقدار مشاهدات محاسبه و هیستوگرام آن رسم می‌شود. در شکل ۵-۳ هیستوگرام رتبه‌ای سامانه همادی و هیستوگرام PIT تابع چگالی احتمال بدست آمده ارائه شده است. در هر دو حالت هیستوگرام رتبه‌ای و هیستوگرام PIT هر چه یکنواخت‌تر باشند نشان دهنده کالیبره بودن آن است. همانطور که در شکل ۵-۳ می‌بینید هیستوگرام تابع چگالی یکنواختتر از هیستوگرام اعضای خام همادی می‌باشد.

^۱ Rank histogram
^۲ Probability Integral Transform



شکل ۵-۳: ضرایب تخمین زده شده در یک دوره آزمون ۵۰ روزه. الف) عرض از مبدأ. ب) تا د) وزن هر کدام از

اعضای همادی. ذ) و ر) ضرایب واریانس



شکل ۵-۳: مقایسه هیستوگرام (الف) هیستوگرام رتبه‌ای سامانه همادی. (ب) هیستوگرام PIT تابع چگالی احتمال.

۵-۴) نحوه راست‌آزمایی

منظور از فرآیند راست‌آزمایی همان، ارزیابی کیفیت پیش‌بینی‌های هواشناسی است که در آن نتایج فرآیند پیش‌بینی با دیدبانی متناظر مقایسه می‌شود. ممکن است این پرسش مطرح شود که اصلاً چه نیازی به راست‌آزمایی داریم. هر چند پاسخ این پرسش تا حدودی روشن به نظر می‌رسد، بریر^۱ و آلن^۲ در ۱۹۵۱ سه دلیل برای آن بر شمرند:

۱- از نظر مدیریتی، سامانه پیش‌بینی باید کنترل شود؛ یعنی بررسی شود که آیا روند کار و سامانه پیش‌بینی با گذشت زمان بهبود می‌یابد یا نه.

۲- از نظر علمی، باید موارد قوت و ضعف سامانه پیش‌بینی شناسایی شود تا از بازخورد آن در بهبود سامانه پیش‌بینی بهره‌گیری شود.

۳- از دیدگاه اقتصادی، کاربران و سرمایه‌گذاران همواره انتظار دارند نسبت به درستی پیش‌بینی‌هایی که دریافت کرده‌اند، توجیه شوند. یا به بیان دیگر بتوانند عملکرد سامانه‌های مختلف پیش‌بینی را با هم مقایسه کنند. برای نمونه اگر کاربران در ازای دریافت خدمات پیش‌بینی وضع هوا، مبلغی را می‌پردازند، ضروری است که با انجام فرآیند راست‌آزمایی آنها را از درصد درستی آنچه دریافت کرده‌اند، مطمئن ساخت.

1- Brier
2- Allan

از سنج‌های متداول آماری برای راست‌آزمایی می‌توان به میانگین قدر مطلق خطا^۱ (MAE) و ریشه میانگین مربع خطا^۲ (RMSE) اشاره نمود که فرمول آنها به قرار زیر می‌باشد بطوریکه f_i پیش‌بینی و y_i مشاهده در نقطه آم است:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - f_i| \quad (27)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2} \quad (28)$$

برای راست‌آزمایی پیش‌بینی نقاط شبکه‌ای، مشکلی که پیش می‌آید عدم وجود مشاهدات در آن نقاط است که برای این منظور می‌توان دو رویکرد را در نظر گرفت:

۱) برای راست‌آزمایی نتایج بدست آمده، از اعتبار سنجی متقابل^۳ استفاده می‌کنیم به این ترتیب که به اندازه تعداد

ایستگاه‌های دیدبانی پس‌پردازش را انجام می‌دهیم به طوری که در هر تکرار یک ایستگاه را حذف کرده و

میانگین قدر مطلق خطا و جذر میانگین مجذور خطا را با توجه به مقدار مشاهدات در ایستگاه حذف شده و

مقدار پیش‌بینی که از چهار نقطه اطراف آن درونیابی شده است محاسبه می‌کنیم.

در ابتدا چند ایستگاه دیدبانی را به طور تصادفی حذف می‌کنیم و پس‌پردازش را با توجه به اطلاعات ایستگاه‌های

باقیمانده انجام می‌دهیم. سپس پیش‌بینی پس‌پردازش شده را به ایستگاه‌های حذف شده درونیابی کرده و میانگین قدر

مطلق خطا و ریشه میانگین مربع خطا را محاسبه می‌کنیم.

۵-۵) نتایج حاصل از پس‌پردازش پیش‌بینی در نقاط شبکه

در این تحقیق، داده‌های مسئله پیش‌بینی‌های ۴۸ ساعته‌ی دمای دومتری با طول جغرافیایی از 44.1° تا 65.4° و

عرض جغرافیایی از 24.8° تا 39.7° و فاصله بین نقاط شبکه 0.1° (حدوداً ۱۱ کیلومتر) می‌باشد. برای پیش‌بینی‌ها از

مدل WRF با مقادیر اولیه داده‌های GFS، استفاده شده است. داده‌های دیدبانی از سازمان هواشناسی کشور برای ۳۹۴

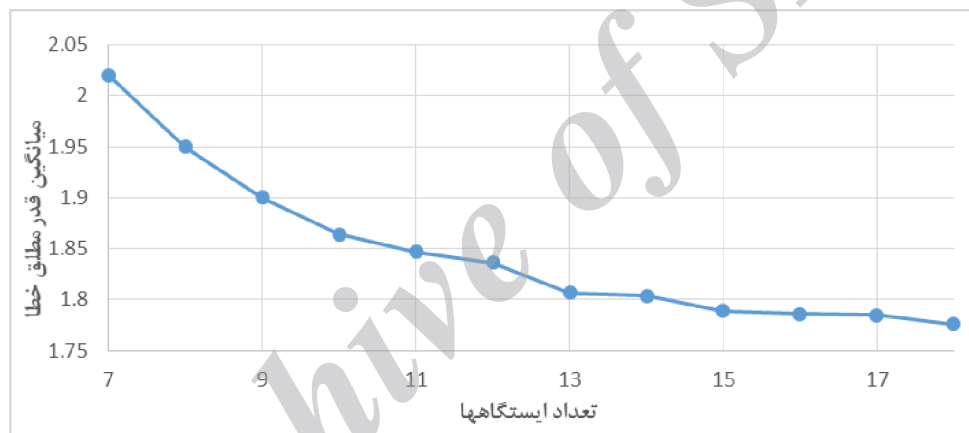
ایستگاه سینوپتیک در دسترس می‌باشد. بازه زمانی، کل ۳۱ روز ماه دسامبر سال ۲۰۱۱ است.

^۱ Mean absolute error

^۲ Root mean squared error

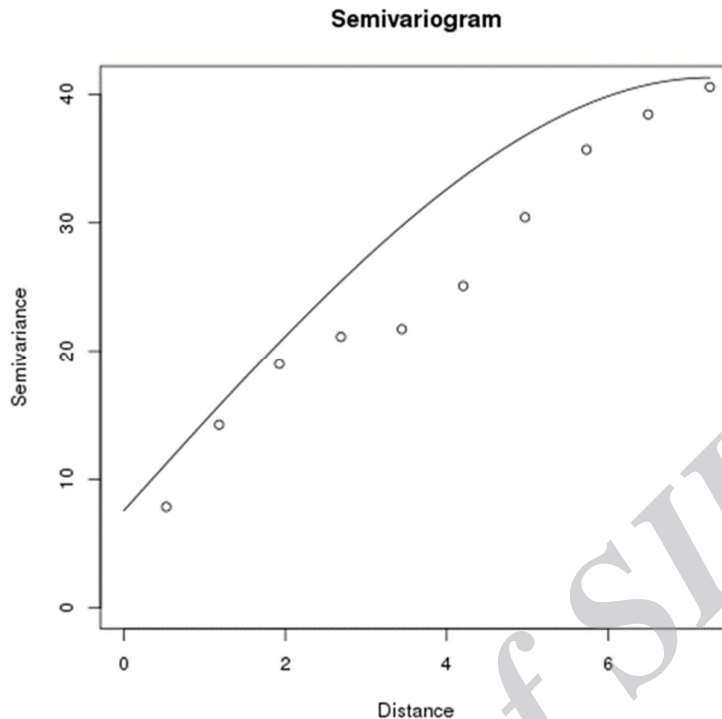
^۳ Cross validation

برای راستی آزمایی پیش بینی نقاط شبکه ای، مشکلی که پیش می آید عدم وجود مشاهدات در آن نقاط می باشد. برای این منظور، به تعداد ایستگاه های دیدبانی پس پردازش را انجام می دهیم به طوری که در هر تکرار یک ایستگاه را حذف کرده و میانگین قدر مطلق خطا (MAE) و ریشه میانگین مربع خطا (RMSE) را با توجه به مقدار مشاهده در ایستگاه حذف شده و مقدار پیش بینی که از چهار نقطه اطراف آن درونیابی شده است محاسبه می کنیم. برای تخمین تعداد مناسب ایستگاه های نزدیک، میانگین مربع خطا را برای تعداد ایستگاه های مختلف محاسبه کرده که در نمودار (۵-۱) نشان داده شده است. همان طور که در نمودار ملاحظه می شود تا تعداد ایستگاه ۱۵، میانگین قدر مطلق خطا در حال کاهش است و بعد از آن تقریباً ثابت می باشد. بنابراین ۱۵ ایستگاه به عنوان مناسبترین تعداد انتخاب شده است.



نمودار ۵-۱: تعیین تعداد ایستگاهها بر حسب میانگین قدر مطلق خطا

نمودار و ریوگرام حاصل از داده های موجود در ایستگاهها در شکل (۵-۴) نشان داده شده است. مقدار range ، $7/26$ ، مقدار sill ، $41/29$ و مقدار nugget ، $7/55$ به دست آمده است. فاصله محاسبه شده در محور افقی بر حسب درجه است که اگر هر درجه 110 کیلومتر در نظر گرفته شده باشد مقدار range برابر با 798 کیلومتر خواهد بود. یعنی ایستگاهها در فاصله 798 کیلومتری از هم دارای همبستگی می باشند.



شکل ۵-۴: نمودار وریوگرام

جدول (۳-۵)، نتایج راستی آزمایی بدست آمده برای میانگین قدر مطلق خطا و ریشه میانگین مربع خطا را نشان داده شده است. همانطور که در جدول مشهود است، روش GIDS دارای خطای کمتری نسبت به روش های BCDG و kriging می باشد. روش BCDG حتی از پیش بینی خام نیز بدتر شده است.

جدول (۳-۵): میانگین قدر مطلق خطا، ریشه میانگین مربع خطا، درصد تغییرات نسبت به پیش بینی خام و نسبت بهبود به عدم بهبود پیش بینی خام و پیش بینی پس پردازش شده

RMSE	MAE	
۳/۶۱	۲/۸۹	پیش بینی خام
۵/۱۱	۳/۵	پیش بینی پس پردازش شده به روش BCDG
۲/۳۳	۱/۷۹	پیش بینی پس پردازش شده به روش GIDS
۲/۷۶	۲/۱۴	پیش بینی پس پردازش شده به روش kriging

برای محاسبه نسبت بهبود به عدم بهبود^۱، صورت کسر تعداد ایستگاه‌هایی است که در آنها وضعیت بهبود یافته یا به عبارتی تفاضل آریبی خام از آریبی پس پردازش شده بیشتر از ۲ درجه کلونین و مخرج کسر تعداد ایستگاه‌هایی است که وضعیت بدتر شده یا به عبارتی تفاضل آریبی خام از آریبی پس پردازش شده کمتر از ۲- درجه کلونین است. اگر مقدار این نسبت بزرگتر از ۱ باشد یعنی تعداد ایستگاه‌های بهبود یافته بیشتر است. این مقدار برای روش GIDS برابر با ۳/۹۸ شده است که دلیل دیگری بر کارایی بالای این روش است.

۵-۶ نتیجه گیری

در گزارش حاضر در فصل اول مروری بر ادبیات پیش‌بینی قطعی و احتمال و پیش‌بینی روی شبکه منظم آورده شد. در فصل دوم توضیحاتی راجع به مدل‌های متداول پیش‌بینی عددی و چگونگی پیش‌بینی در کشور ارائه شد. در فصل سوم درباره الگوریتم بهینه سازی حرکت جمعی ذرات که در فصل چهارم از آن برای کمینه سازی تابع CRPS استفاده شده صحبت شده است. نتایج حاصل از فصل چهارم به قرار زیر است.

یک تابع چگالی احتمال نرمال با میانگین و واریانس حاصل از ترکیب خطی از اعضا و واریانس اعضای سامانه همدادی، برای پیش‌بینی احتمالی دمای سطح زمین تولید شده است. در این تحقیق، بر اساس کار نیتینک و همکاران (Goldman, 2005 & Gneiting, Raftery, Westveld) ضرایب ترکیب خطی با کمینه کردن CRPS که یک معیار سنجش خطا برای پیش‌بینی احتمالی می‌باشد تخمین زده شده است. برای کمینه کردن CRPS از الگوریتم فراابتکاری PSO استفاده شده است. نتایج بدست آمده نشان‌دهنده آن است که مقدار CRPS، ۱/۲ بوده و میانگین قدرمطلق خطا و ریشه میانگین مربع خطای میانگین تابع چگالی نرمال نسبت به میانگین اعضای همدادی با آریبی تصحیح شده به ترتیب ۳۵٪ و ۲۸٪ کمتر و نسبت به میانگین اعضای خام همدادی به ترتیب ۵۹٪ و ۵۱٪ کمتر و پیش-بینی احتمالاتی نیز کالیبره می‌باشد. به این معنی که احتمال پیش‌بینی شده برای آستانه‌های مختلف با فراوانی نسبی متناظر با دیدبانی تقریباً برابر است.

در فصل پنجم، نحوه پس پردازش توضیح داده شد. به این صورت که ابتدا توضیحاتی راجع به وضعیت ایستگاه‌های موجود داده شد. در ادامه درباره نحوه پیش‌بینی و استخراج داده‌ها از فایل‌های SRTM و استخراج پیش‌بینی واقع بر

^۱ Improve to hurt ratio

ایستگاه‌ها از خروجی مدل WRF و همچنین کدهای نوشته شده به زبان C، فرترن و R به طور مفصل توضیح داده شد. سپس نتایج حاصل از روش‌های BCDG، GIDS و kriging مقایسه شدند.

نتایج نشان داد که از بین روش‌های مذکور، روش GIDS دارای کمترین خطا است. اگر چه این روش نسبت به سایر روش‌ها دارای فرمول ریاضی ساده‌ای بود اما نتیجه بهتری داشت. البته فرمول ریاضی ساده مزیت این روش به حساب می‌آید که می‌توان آن را به راحتی عملیاتی کرد. روش‌های بکار رفته در این تحقیق به جز روش BCDG، برای بار اول است که در دنیا برای پس‌پردازش نقاط شبکه منظم مورد استفاده قرار می‌گیرند. روش‌های GIDS و kriging برای شبکه‌بندی داده‌های دیدبانی به کار رفته‌اند و در آن تحقیقات نیز روش GIDS دارای برتری زیادی نسبت به سایر روش‌ها داشته است.

فهرست مراجع

- Ai T.J., Kachitvichyanukul V. "A study on adaptive particle swarm optimization for solving vehicle routing problems. The 9th Asia Pacific Industrial Engineering and Management Systems Conference, 2262- 2268, 2008.
- Azadi M., Zakeri Z. "Probabilistic Precipitation Forecasting Using a Deterministic Model Output over Ira", Research Journal of Environment Sciences, 138-148, 2010.
- Bergthorsson P., Doos A.B. "Numerical weather map analysis", Tellus, 7, 329-340, 1955.
- Breiman L., Friedman J. H. "Estimating optimal transformations for multiple regression and correlations (with discussion)", J. Amer. Stat. Assoc, 80, 580-619, 1985.
- Breiman L., Olshen R.A., Stone C.J. "Classification and Regression Trees", Wadsworth and Brooks/Cole, 385, 1984.
- Buja A., Kass R.E. "Estimating optimal transformations for multiple regression and correlation", J. Amer. Stat. Assoc, 80, 602-607, 1985.
- Burrows W.R., Benjamin M., Beauchamp S., Lord E.R., Mc-Collor D., Thompson B. "CART decision-tree statistical analysis and prediction of summer season maximum surface ozone for Vancouver, Montreal, and Atlantic regions of Canada", J. Appl. Meteor, 34, 1848-1862, 1995.
- Chen K., Li T., Cao L. "Tribe-PSO: A novel global optimization algorithm and its application in molecular docking", Chemometrics and Intelligent Laboratory Systems, 82, 248-259, 2006.
- Cressie N. "Statistics for Spatial Data", John Wiley and Sons, 1993.
- Cressman G.P. "An operational objective analysis system", Monthly Weather Review, 87, 367-374, 1959.
- Dallavalle J., Glahn H.R. "Toward a gridded MOS system", 21st Conf on Weather Analysis and Forecasting and 17th Conf. on Numerical Weather Prediction, Washington, DC, Amer. Meteor. Soc., 2005.
- David M. "Geostatistical Ore Reserve Estimation", Elsevier, Amsterdam, 1977.

- Dorigo M. "Particle swarm optimization. Scholarpedia", 3(11), 1486, 2008.
- Estephenson D. "an introduction to probability forecasting", Seasonal Climate: Forecasting and Managing Risk, Netherlands: Springer, 241-264, 2007.
- Fielding A.H. "Ecological Applications of Machine Learning Methods", Kluwer Academic, 280, 1999.
- Fisher R.A. "The Design of Experiments", Oliver and Boyd, Edinburgh, 1935.
- Gao Y., Ren Z. "Adaptive particle swarm optimization algorithm with genetic mutation operation", Proceedings of the Third International Conference on Natural Computation, 2007.
- Gel Y. "Comparative Analysis of the Local Observation-Based (LOB) Method and the Nonparametric Regression-Based Method for Gridded Bias Correction in Mesoscale Weather Forecasting", Weather And Forecasting, 22, 1243-1256, 2007.
- Gel Y., Raftery A.E., Gneiting T. "Combining global and local grid-based bias correction for mesoscale numerical weather prediction models", 17th Conf. on Probability and Statistics in the Atmospheric Sciences, Seattle, 2004.
- Ghomsheh V., Shoorehdeli M.A., Teshnehlab M. "Training ANFIS Structure with Modified PSO Algorithm", Proceedings of the 15th Mediterranean Conference on Control & Automation, Greece, 1-6, 2007.
- Glahn B., Gilbert K., Cosgrove R., Ruth D.P., Sheets K. "The gridding of MOS", Weather and Forecasting, 24, 520-529, 2009.
- Glahn H., Lowry R. "The use of model output statistics (MOS) in objective weather forecasting", J. Appl. Meteor, 11, 1203-1211, 1972.
- Gneiting T., Raftery A., Westveld A., Goldman A.T. "Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation", Monthly Weather Review, 133, 1098-1118, 2005.
- Halley E. "An Historical of the Trade Winds, and Monsoons, Observable in the Seas between and near the Tropicks; With an attempt to assign the physical cause of Said Winds", Philosophical Thransection, 183, 153-168, 1686.
- Hastie T.J., Tibshirani R.J. "Generalized Additive Models", CRC Press, 352, 1990.

- Heppner F., Grenander G. "A stochastic nonlinear model for coordinated bird flocks", *The Ubiquity of Chaos*, Washington, DC: AAAS Publications, 233-238, 1990.
- Jacks E., Bower J.B., Dagostaro V.J. "New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind", *Wea. Forecasting*, 5, 128-138, 1990.
- Jiao B., Liana Z. "A dynamic inertia weight particle swarm next term optimization algorithm", *Chaos, Solitons and Fractals* 37, 698-705, 2008.
- Journel A.G., Huijbregts C.J. "Mining Geostatistics", London: Academic Press, 1978.
- Kennedy J., Eberhat R. "Particle Swarm Optimization", in *Proceedings of IEEE International Conference on Neural Networks*, 5, 1942-1948, 1995.
- Kennedy J., Eberhat R. "A discrete binary version of the particle swarm algorithm", In *Proceedings of the IEEE International Conference on Systems*, 4104-4108, 1997.
- Kennedy J. "Bare bones particle swarms", In *Proceedings of the IEEE Swarm Intelligence Symposium*, 80-87, 2003.
- Klein W.H., Lewis B.M., Enger I. "Objective prediction of five-day mean temperatures during winter", *J. Atmos. Sci*, 16, 672-682, 1959.
- Lorenz E.N. "The essence of chaos", University of Washington Press, 1963.
- Clerc M., Kennedy J. "The particle swarm-explosion, stability and convergence in a multidimensional complex space", *IEEE Transactions on Evolutionary Computation*, 6(1), 58-73, 2002.
- Arumugam M.S., Rao M.V.C. "On the improved performances of the particle swarm optimization algorithms with adaptive parameters, cross-over operators and root mean square (RMS) variants for computing optimal control of a class of hybrid systems", *Applied soft computing*, 8, 324-336, 2008.
- Mass C.F. "IFPS and the future of the National Weather Service", *weather and forecasting*, 18, 75-79, 2003.
- Mass C.F., Baars J., Wedam G., Gritmit E., Steed R. "Removal of Systematic Model Bias on a Model Grid", *weather and forecasting*, 23, 438-459, 2007.

- Matheron G. "Le Krigeage Universal Cahirs du Centre de Morphologie Mathematique, No. 1", Fontainebleau, France, *Jornal of Econometrics*, 37, 87-114, 1962.
- Mendes R., Kennedy J., Neves J. "The fully informed particle swarm: simpler, maybe better", *IEEE Transactions on Evolutionary Computation*, 8(3), 204-210, 2004.
- Nalder I., Wein R. "Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest", *Forest Meteorol*, 92, 211-225, 1998.
- Qin Z., Shi Z., Lu J., Liu Y. "Center particle swarm optimization", *Neurocomputing*, 70, 627-679, 2007.
- Raftery A., Gneiting T., Balabdaoui F., Polakowski A. M. "Using Bayesian model averaging to calibrate forecast ensemble", *Monthly Weather Review*, 133, 1155-1174, 2005.
- Reeves W.T. "Particle systems-A technique for modeling a class of fuzzy objects", *ACM Transactions on Graphics*, 2, 91-108, 1983.
- Reynolds C.W. "Flocks, herds, and schools: A distributed behavioral model", *ACM Computer Graphics*, 21(4), 25-34, 1987.
- Ripley B.D. "Spatial Statistics", New York: Wiley, 1981.
- Spark E., Connor G.J. "Wind forecasting for the sailing events at the Sydney 2000 Olympic and Paralympic Games", *weather and forecasting*, 19, 181-199, 2004.
- Shi Y.H., Eberhat R. "A Modified Particle Swarm Optimizer", *IEEE International Conference on Evolutionary Computation*, Anchorage, Alaska, 1998.
- Shi Y., Eberhat R. "Parameter selection in particle swarm optimization", *Evolutionary Programming VII, Lecture Notes in Computer Science*, 591-600, 1998.
- Student. "On the Error of Counting with a Haemacytometer", *Biometrika*, 5, 351-360, 1907.
- Tebaldi C. "Looking far back vs. looking around enough: Operational weather forecasting by spatial composition of recent observations", *16th Conf. on Probability and Statistics in the Atmospheric Sciences*, Orlando, 2002.
- Unger D.A., Dool H.V., O'Lenic E., Collins A.D. "Ensemble regression", *Monthly Weather Review*, 7(137), 2365-2379, 2009.

Veeramachaneni K., Peram T., Mohan C., Osadciw L.A. "Optimization using particle swarms with near neighbor interaction", Proceedings of Genetic and Evolutionary Computation Conference, 2003.

Wahba G. "Spline models for Observational Data", siam, Philadelphia, 1990.

Wedam G, Mass C. F., Steed R. "Grid-based bias removal of surface parameters", Pacific Northwest Weather Workshop, Seattle, WA, University of Washington and NWS, 2005. [Available online at http://www.wrh.noaa.gov/sew/WorkShop_05/session_3/Wedam.pdf].

Wilson L.J., Vallée M. "The Canadian Updatable Model Output Statistics (UMOS) system: Design and development tests", Wea. Forecasting, 17, 206–222, 2002.

Yang C., Simon D. "A New Particle Swarm Optimization Technique", Proceedings of the 18th International Conference on Systems Engineering, 2005.

Yussouf N., & Stensrud D. J. "Prediction of near-surface variables at independent locations from a bias-corrected ensemble forecasting system", monthly weather review, 134, 3415–3424, 2006.

Archive of SID

پیوست ۱: کد برنامه‌ها

استخراج پیش‌بینی از خروجی مدل WRF

خروجی‌های wrfout حاصل از اجرای مدل WRF با اجرای برنامه UPP، به فرمت اسکی تبدیل می‌شوند. با اجرای UPP، برای هر روز و هر ساعت فایلی به نام wrfprs_dnn.hhh تولید می‌شود که nn شماره دامنه و hhh ساعت پیش‌بینی می‌باشد. برای مثال wrfprs_d02.24 مربوط به دامنه دوم و پیش‌بینی ۲۴ ساعته می‌باشد. خروجی‌ها روی یک شبکه منظم که شامل کل ایران است، صادر شده‌اند. طول جغرافیایی از 44.1° تا 65.4° و عرض جغرافیایی از 24.8° تا 39.7° است و فاصله بین نقاط شبکه 0.1° (حدوداً ۱۱ کیلومتر) می‌باشد. هر یک از این فایل‌ها پیش‌بینی‌های پارامترهای مختلف مانند دما، رطوبت، فشار، ارتفاع ژئوپتانسیل، سرعت باد و را در ارتفاعات مختلف جو در بر دارد.

برای استخراج پیش‌بینی‌های مربوط به پارامتر مورد نظر ابتدا باید شماره رکورد مربوط به آن پارامتر را پیدا نمود. برای این منظور با استفاده از دستور زیر، فایلی ایجاد می‌شود که شامل اطلاعات همه پارامترها می‌باشد:

```
$ wgrib -V wrfprs_d02.24 > data
```

با اجرای دستور فوق اطلاعات مربوط به پیش‌بینی ۲۴ ساعته برای یک روز دلخواه در فایل data ریخته می‌شود. یک نمونه از این فایل در شکل ۱ نشان داده شده است. همانطور که در شکل دیده می‌شود رکورد ۲۲۸ مربوط به پیش‌بینی ۲۴ ساعته دمای دومتری برای تاریخ شروع در ساعت ۱۲، روز ۱۰، ماه ۱ و سال ۲۰۱۲ (۲۰۱۲۰۱۱۰۱۲)(YYYYMMDDHH) است. رکورد ۲۲۹ مربوط به پارامتر رطوبت، رکورد ۲۳۰ برای دمای نقطه شبنم و الی آخر می‌باشد.

```
rec 228:11613846:date 2012011012 TMP kpds5=11 kpds6=105 kpds7=2 levels=(0,2) grid=255 2 m above gnd 24hr fcst: bitmap: 2029 undef
TMP=Temp. [K]
timerange 0 P1 24 P2 0 TimeU 1 nx 214 ny 150 GDS grid 0 num_in_ave 0 missing 0
center 7 subcenter 0 process 89 Table 2 scan: WE:NS winds(grid)
latlon: lat 24.800000 to 39.700000 by 0.100000 nxny 32100
long 44.100000 to 65.400000 by 0.100000, (214 x 150) scan 0 mode 136 bdsgrid 1
min/max data 264.699 302.038 num bits 16 BDS_Ref 264699 DecScale 3 BinScale 0

rec 229:11678092:date 2012011012 SPFH kpds5=51 kpds6=105 kpds7=2 levels=(0,2) grid=255 2 m above gnd 24hr fcst: bitmap: 2029 undef
SPFH=Specific humidity [kg/kg]
timerange 0 P1 24 P2 0 TimeU 1 nx 214 ny 150 GDS grid 0 num_in_ave 0 missing 0
center 7 subcenter 0 process 89 Table 2 scan: WE:NS winds(grid)
latlon: lat 24.800000 to 39.700000 by 0.100000 nxny 32100
long 44.100000 to 65.400000 by 0.100000, (214 x 150) scan 0 mode 136 bdsgrid 1
min/max data 0.0015 0.016 num bits 8 BDS_Ref 15 DecScale 4 BinScale 0

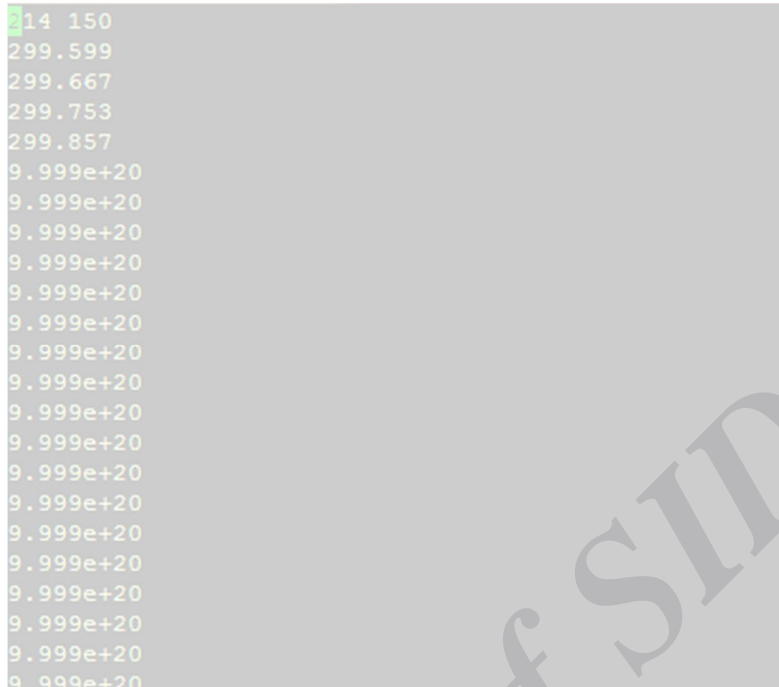
rec 230:11712266:date 2012011012 DPT kpds5=17 kpds6=105 kpds7=2 levels=(0,2) grid=255 2 m above gnd 24hr fcst: bitmap: 2029 undef
DPT=Dew point temp. [K]
timerange 0 P1 24 P2 0 TimeU 1 nx 214 ny 150 GDS grid 0 num_in_ave 0 missing 0
center 7 subcenter 0 process 89 Table 2 scan: WE:NS winds(grid)
latlon: lat 24.800000 to 39.700000 by 0.100000 nxny 32100
```

شکل ۱: خروجی حاصل از اجرای wgrib

بعد از مشخص شدن شماره رکورد، حال می توان پیش بینی پارامتر مورد نظر را روی شبکه منظم با دستور زیر روی فایل به نام ۲۰۱۲۰۱۱۰۱۲ قرار داد:

```
$ wgrib -d 228 -o 2012011012 -text wrfprs_d02.24
```

به طوری که -d برای دیکد کردن شماره رکورد، -o برای نامگذاری فایل خروجی و -text برای نوع خروجی می باشد. با اجرای دستور فوق فایل 2012011012 ایجاد می شود که چند خط اول آن در شکل ۲ آورده شده است. در خط اول عدد ۲۱۴ تعداد نقاط شبکه در طول و ۱۵۰ تعداد نقاط شبکه در عرض می باشد که این فایل حدود ۳۲۱۰۰ خط می باشد. خط های بعدی مقدار پیش بینی روی نقاط شبکه به ترتیب از گوشه سمت چپ پایین تا گوشه سمت راست بالا را نشان می دهد. برای مثال، خط دوم مقدار پیش بینی دمای دو متری به واحد کلوین در نقطه شبکه ای به مختصات $(44.1^\circ, 24.8^\circ)$ و خط سوم مقدار پیش بینی در مختصات $(44.2^\circ, 24.8^\circ)$ و است. عدد $9.999e+20$ داده missing می باشد.



شکل ۲: خروجی اسکی دمای سطح زمین به کلومین

برای تعیین مقدار پیش‌بینی در نقاط ایستگاهی که طول و عرض آنها مشخص است، طول و عرض را طوری گرد کرده که هر کدام از نقاط ایستگاهی روی یک نقطه شبکه واقع شود. سپس شماره خط مورد نظر در فایل فوق را محاسبه کرده و به عنوان پیش‌بینی در آن ایستگاه در نظر می‌گیریم. برنامه نوشته شده به زبان C در شکل ۳ قابل مشاهده است. فرمول مورد استفاده برای یافتن شماره خط، داخل کادر آبی رنگ نشان داده شده است. a و b به ترتیب عرض و طول ایستگاه می‌باشد که در ۱۰ ضرب شده است. عدد ۲۱۴ تعداد نقاط در طول، ۲۴۸ کوچکترین عرض ضربدر ۱۰ و ۴۴۱ کوچکترین طول ضربدر ۱۰ است.

```
ss[19]=((int)s[19])-48;  
ss[20]=((int)s[20])-48;  
  
a=100*ss[6]+10*ss[7]+ss[9];  
if(s[9]==' ')  
a=100*ss[6]+10*ss[7];  
  
b=100*ss[16]+10*ss[17]+ss[19];  
  
if(ss[10]>=5)  
a++;  
if(ss[20]>=5)  
b++;  
  
k=214*(b-248)+(a-441+1)+1;  
  
for(i=1;i<=k;i++)  
fgets(q,300,e);  
  
fclose(e);  
  
id=strtod(s,&ptr);  
long int id=(int)id;
```

شکل ۳: نمونه برنامه برای استخراج دمای ایستگاهها

در این برنامه از فایلی که مشخصات ایستگاهها شامل طول و عرض جغرافیایی و ارتفاع در آن موجود می باشد، خوانده می شود و مقدار پیش بینی پارامتر مورد نظر برای ایستگاهها بدست می آید. شکل ۴ فایل مشخصات ایستگاهها است که ستون اول اسم استان، ستون دوم اسم شهر، ستون سوم شماره شناسه ایستگاه، ستون چهارم طول جغرافیایی، ستون پنجم عرض جغرافیایی و ستون آخر ارتفاع ایستگاه می باشد.

Ardebil	Parsabad	40700	47.779	39.604	72.600
Azərbayjan-e-Gharbi	Maku	40701	44.392	39.379	1411.200
Azərbayjan-e-Sharghi	Jolfa	40702	45.667	38.750	736.200
Azərbayjan-e-Gharbi	Khoy	40703	44.995	38.558	1103.400
Azərbayjan-e-Sharghi	Ahar	40704	47.067	38.433	1391.000
Ardebil	Meshkinshahr	40705	47.678	38.375	1560.600
Azərbayjan-e-Sharghi	Tabriz	40706	46.243	38.122	1361.000
Azərbayjan-e-Sharghi	Sahand	40707	46.117	37.933	1641.000
Ardebil	Ardebil	40708	48.329	38.218	1335.200
Gilan	Astara	40709	48.855	38.365	-21.100
Azərbayjan-e-Sharghi	Sarab	40710	47.533	37.933	1682.000
Azərbayjan-e-Sharghi	Kalibar	40711	47.017	38.867	1180.000
Azərbayjan-e-Gharbi	Orumiyeh	40712	45.055	37.659	1328.000
Azərbayjan-e-Sharghi	Maragheh	40713	46.146	37.347	1344.000
Ardebil	Germi	40714	48.057	39.050	749.000
Zanjan	Mahneshan	40715	47.684	36.740	1284.500
Azərbayjan-e-Sharghi	Mianeh	40716	47.700	37.450	1110.000
Ardebil	Khalkhal	40717	48.538	37.609	1797.400
Gilan	Bandar-e-Anzali	40718	49.458	37.480	-23.600
Gilan	Rasht	40719	49.624	37.322	-8.600
Gilan	Manjil	40720	49.410	36.729	338.300
Golestan	Maravehtappeh	40721	55.943	37.302	460.000
Azərbayjan-e-Gharbi	Salmas	40722	44.848	38.216	1339.300
North Khorasan	Bojnurd	40723	57.303	37.487	1065.000

شکل ۴: فایل مشخصات ایستگاهها

استخراج ارتفاع هر نقطه شبکه روی ایران

به منظور دانلود و استفاده راحتتر از داده‌های SRTM با تفکیک ۹۰ متر، آنها را در غالب موزائیک‌هایی با ابعاد ۵ درجه در ۵ درجه تهیه کرده‌اند. برای بدست آوردن ارتفاع در ایران، از ۱۲ عدد از آن موزائیک‌ها استفاده شده است. اولین موزائیک در گوشه جنوبی غربی، شامل طول از ۴۴/۹ تا ۴۹/۹ و عرض از ۲۴/۹ تا ۲۹/۹ است. یک نمونه از فایل آن موزائیک‌ها در شکل ۵ نشان داده شده است:

```
cols      6001
nrows     6001
xllcorner 44.999583696539
yllcorner 24.999583914469
cellsize  0.000833333333333333
NODATA_value -9999
```

شکل ۵: یک نمونه فایل SRTM

که در آن خط اول و دوم تعداد ستون‌ها و سطرها، خط سوم و چهارم مختصات گوشه جنوب غربی، خط پنجم فاصله دو نقطه به درجه و خط ششم داده غیر معتبر را نشان می‌دهد. خطوط بعدی که با علامت @ شروع شده‌اند اطلاعات یک سطر را شامل می‌شوند. به طوری که اگر نشانگر را روی آن خط ببریم تمامی ارتفاعات نقاط واقع در آن سطر را نشان می‌دهد که یک نمونه از آن در شکل ۶ آورده شده است.

```
650 652 652 646 643 644 644 639 636 640 637 636 635 628 627 625 622 623
618 618 618 616 616 618 616 615 615 614 618 618 619 618 618 614 615 61
3 602 600 598 600 602 602 602 602 601 602 602 602 604 600 599 600 598 5
91 591 590 592 591 598 603 605 604 604 605 608 611 607 601 602 601 599
596 594 595 594 595 600 599 597 595 594 595 595 596 597 595 592 593 594
588 585 582 584 585 588 587 584 587 591 591 598 601 603 601 597 601 60
0 568 569 568 566 566 569 566 565 571 571 569 566 567 566 567 565 562 5
62 561 562 563 562 560 559 560 561 562 560 557 560 557 558 558 559 557
559 558 558 559 564 568 571 570 568 569 571 574 575 572 570 570 573 570
575 579 584 574 570 583 587 574 572 589 585 569 571 576 578 581 592 60
2 606 596 591 589 596 609 608 601 599 594 579 582 584 581 577 577 577 5
07 606 606 602 599 605 602 599 598 595 602 603 606 609 609 613 614 618
633 630 627 627 628 640 649 647 649 655 660 656 661 646 659 641 657 665
654 655 662 672 684 690 692 692 692 692 694 694 694 693 691 691 689 69
2 692 692 693 691 692 694 692 693 691 690 692 694 694 694 693 693 693 6
91 692 692 693 693 693 693 693 693 694 694 692 693 693 693 692 692 693
693 694 693 692 692 694 695 694 694 694 692 694 694 693 695 695 693 693
```



```
!NOW READ THE TOPOGRAPHY DATA
do j=my,1,-1
write(*,*)j
read(1,*)(topo(i,j),i=1,my)
enddo
do j=1,my
lat(j)=y11+(j-1)*d
write(*,*)j
do i=1,mx
lon(i)=x11+(i-1)*d
!write(4,10)lon(i),lat(j),topo(i,j)
10 format(3f12.4)
enddo
enddo
!PUT YOUR OWN LAT LON HERE
do jj=0,50
mylat=35+.1*jj
write(*,40)'Doing for lat:',mylat
do ii=0,50
mylon=40+.1*ii
!write(*,20)mylon,mylat
ichap=1
jchap=1
dchap=99999.

do j=1,my
dis=(mylat-lat(j))**2
if(dchap.gt.dis)then
jchap=j
dchap=dis
endif
enddo
dchap=99999.
```

شکل ۷: برنامه استخراج ارتفاع از فایل SRTM

45.0	25.0	790.000
45.1	25.0	781.000
45.2	25.0	746.000
45.3	25.0	750.000
45.4	25.0	722.000
45.5	25.0	750.000
45.6	25.0	721.000
45.7	25.0	713.000
45.8	25.0	731.000
45.9	25.0	762.000
46.0	25.0	871.000
46.1	25.0	845.000
46.2	25.0	797.000
46.3	25.0	835.000
46.4	25.0	724.000
46.5	25.0	670.000
46.6	25.0	659.000
46.7	25.0	619.000
46.8	25.0	597.000
46.9	25.0	664.000
47.0	25.0	635.000
47.1	25.0	677.000

شکل ۸: فایل ارتفاعها

روش محاسبه فاصله دو نقطه به روش فضایی

در تمام برنامه‌های نوشته شده، فاصله بین دو نقطه مورد نظر به صورتی که در ادامه به آن اشاره خواهد شد محاسبه شده است. از آنجا که مختصات نقاط به صورت کروی و بر حسب درجه می‌باشد ابتدا مختصات آنها را به دستگاه دکارتی تبدیل کرده‌ایم. به این صورت که اگر طول و عرض جغرافیایی را به ترتیب با lon و lat نشان دهیم، آنگاه طول و عرض و ارتفاع در مختصات دکارتی به صورت زیر محاسبه می‌شود:

$$x = \cos(lon) \times \cos(lat)$$

$$y = \cos(lon) \times \sin(lat)$$

$$z = \sin(lon)$$

سپس اگر کره زمین به صورت یک دایره در نظر گرفته شود، برای محاسبه فاصله دو نقطه که زاویه بین آنها از مبدأ برابر با θ است از رابطه شعاع کره زمین ضربدر θ استفاده می‌کنیم. برای بدست آوردن θ از ضرب داخلی دو بردار استفاده می‌کنیم. بنابراین داریم:

$$\theta = \cos^{-1}(x_1x_2 + x_1x_2 + x_1x_2)$$

لازم به ذکر است که شعاع کره زمین در محاسبات تقریباً برابر با ۶۳۷۰ کیلومتر فرض شده است.

درونیابی بر روی شبکه منظم به روش GIDS

ابتدا فایل های ورودی به برنامه معرفی می شوند. برای هر روز یک فایل ۷ ستونه وجود دارد که ستون های اول تا هفتم به ترتیب، کد ایستگاه ها، سال، ماه، روز، مقدار دیدبانی، پیش بینی خام و پیش بینی پس پردازش شده در ایستگاه را شامل می شود. روش پس پردازش روی ایستگاه ها به روش میانگین گیری وزنی لغزان با دوره آموزش ۱۵ روزه و وزن نمایی می باشد. روش میانگین گیری لغزان به این صورت است که به مقدار خطای هر روز، میانگین وزنی خطای ۱۵ روز گذشته اضافه می شود. فایل بعدی فایل ارتفاعها (elavation) که قبلاً معرفی شد می باشد.

در این برنامه برای هر روز، یک ماتریس ۶ ستونه که سطرها ی آن به تعداد ایستگاه ها می باشد به نام st ساخته می شود. ستون های اول تا ششم به ترتیب شامل کد ایستگاه، طول، عرض، ارتفاع، پیش بینی خام و پیش بینی پس پردازش شده است. مشاهدات دیدبانی در آرایه ای به نام obs ذخیره شده اند. در این برنامه ابتدا برای هر نقطه شبکه نزدیکترین ایستگاه ها به تعدادی مشخص انتخاب می شوند تا از پیش بینی پس پردازش شده آن ایستگاه ها استفاده شود.

اطلاعات ایستگاه ها به اضافه گره مورد نظر در یک فایل ۴ ستونه که به ترتیب شامل طول، عرض، ارتفاع و پیش بینی پس پردازش شده ذخیره می شود تا مدل رگرسیون که طول و عرض و ارتفاع، پیش بینی کننده های آن و پیش بینی پس پردازش شده، پیش بینی شونده در نظر گرفته شده است تخمین زده شود. برای تخمین مدل رگرسیون از نرم افزار R استفاده شده است که در برنامه فرترن، مجبور به رفت و برگشت به R در حین اجرای برنامه هستیم. برای این منظور با استفاده از دستور زیر، برنامه ای با عنوان reg.R در نرم افزار R را فراخوانی می کنیم:

```
call system("R CMD BATCH reg.R")
```

برنامه reg.R فایل حاوی اطلاعات نزدیکترین ایستگاه ها را به عنوان ورودی می گیرد و مدل رگرسیون را با استفاده از تابع lm() تخمین می زند و با استفاده از تابع predict() مقدار پیش بینی با استفاده از مدل رگرسیون در ایستگاه ها و گره مورد نظر را در فایل routput.txt می ریزد. برنامه reg.R و توابع مذکور را می توانید در شکل ۹ ببینید.

```
file1=read.table("/home/mohammadi/rtest/rinput.txt")
#te=FALSE,col.name=FALSE,row.name=FALSE)
datafile<-file1[file1$V1<100&file1$V2<100&file1$V3<100&file1$V4>-9000, ]
colnames(datafile)=c("temp","long","lat","h")
attach(datafile)
n=length(long)
print(n)
wtemp=temp[2:n]
wlong=long[2:n]
wlat=lat[2:n]
wh=h[2:n]
lmfit=lm(wtemp~wlong+wlat+wh)
print(lmfit)
yy=c(1,1)
for(i in 1:n)
{
y=predict(lmfit,newdata=data.frame(wlong=long[i],wlat=lat[i],wh=h[i]))
yy[i]=y
}
print(yy)
write.table(yy,file="routput.txt",quote=FALSE,col.name=FALSE,row.name=FALSE)
detach(datafile)
-
```

شکل ۹: برنامه reg.R

برای محاسبه خطا و ارزیابی میزان درستی پس پردازش از روش cross validation استفاده شده است. به این صورت که در هر تکرار یک ایستگاه از لیست ایستگاه‌های دیدبانی حذف می‌شود و مکان ایستگاه حذف شده به عنوان یک گره در نظر گرفته می‌شود و درونیایی در آن گره اعمال می‌شود و مقدار خطا محاسبه می‌شود. بعد از تکرار این کار بر روی همه ایستگاه‌ها میانگین قدر مطلق خطا و ریشه میانگین مربع خطا محاسبه می‌شود. قسمتی از برنامه درونیایی در شکل ۱۰ نشان داده شده است.

```
do t=1,(num_st+1) !do not consider the station itself
write(*,*) dist(t),ind(t)
if(dist(t).ne.0) then
write(20,*) st(ind(t),5),st(ind(t),2),st(ind(t),3),st(ind(t),4)
write(20,*) obs(ind(t)),st(ind(t),2),st(ind(t),3),st(ind(t),4)
endif
enddo
close(20)
call system("R CMD BATCH reg.R")
open(4,file="routput.txt")
read(4,*) pg !predicted grid by regression
soorat=0
makhraj=0
do t=1,(num_st+1)
if(ind(t).ne.kk) then
read(4,*) pst !predicted station by regression
! read(4,*) pg !predicted grid by regression
!if(ind(t).ne.kk)
soorat=soorat+(st(ind(t),5)+(pg-pst))*((1/dist(t))**2)
makhraj=makhraj+((1/dist(t))**2)
! write(*,*) (dist(t),ind(t)
! write(*,*) x y z dist(t) st(ind(t)) pg ind(t)
```

شکل ۱: برنامه GIDS

درونیابی به روش کریگیدن

برای درونیابی به این روش از پکیج آماری kriging.R که در محیط نرم افزار R اجرا می شود استفاده شده است. برای این منظور ابتدا با دستور زیر، پکیج را دانلود و نصب کرده و سپس آن را در محیط مورد نظر لود می کنیم:

```
>Install.packages("kriging.R")
```

```
>library("kriging.R")
```

با استفاده از تابع kriging می توان درونیابی را انجام داد و همچنین نمودار و ریوگرام آن را رسم کرد. تابع kriging مختصات نقاط و مقدار پیش بینی در آن نقاط را به عنوان آرگومان های ورودی می گیرد. تعداد lag و تعداد پیکسل ها مواردی هستند که توسط کاربر اختیار می شود. کد نوشته شده در شکل ۱۱ نشان داده شده است.

```
datafile=read.table("coord")
colnames(datafile)=c("id","long","lat","temp_obs","fore")
attach(datafile)
kriged=kriging(long,lat,fore,pixels=100,lags=10)
image(kriged)
points(long,lat)
print(kriged$model)
print(kriged$nugget)
print(kriged$range)
print(kriged$sill)
print(kriged$map)
plot(kriged)
detach(datafile)
```

شکل ۱۱: برنامه kriging

Archive of SID

Abstract

In general, NWP models include several highly nonlinear partial differential equations that describe the governing physical and dynamical laws of the atmosphere. These differential equations along with proper initial and boundary conditions have to be solved by the numerical methods. Because of the approximate solutions in the numerical methods, the output of the numerical models is an approximation of the real answer and always involves error. The use of numerical methods in the models is not the only source of error. There are also some other sources of error. Among them the presence of errors in the initial/boundary values, the lack of data in some areas and chaotic nature of the dynamical system of the atmosphere are the most important.

The mentioned bias in NWP models includes both random and systematic errors. The systematic error is the average of deviation of prediction to observation and can be removed by statistical post-processing. But random error does not have any rule and cannot be removed. There are several methods to remove the systematic errors that they use the observation of the stations. In this project, both the probability forecasting and post-processed deterministic forecasting is done to quantify the uncertainty of the atmosphere.

In the recent years, the requirement to the post-processed weather prediction in the non-station points has been increased. Therefore in this project, deterministic forecast is post-processed, by three methods naming BCDG, GIDS and kriging for grid points over Iran.

Keywords: numerical weather prediction, post-processing, gridding, kriging



ACECR



Islamic Republic of Iran

Gridding of post processed wrf output for 2-meter temperature over Iran

Technology development institute

(ACECR), Sharif branch

Morteza Rahmani

November , 2015