Archive of SID

تعمیم روش برازش داده‌های ایستگاهی به شبکه منظم در هواشناسی

سیده عاطفه محمدی، دانشجوی دکتری، پژوهشکده توسعه تکنولوژی، mohamadi.atefeh@yahoo.com
مرتضی رحمانی*، عضو هیأت علمی، پژوهشکده توسعه تکنولوژی، rahmanimr@yahoo.com
مجید آزادی، عضو هیأت علمی، پژوهشکده هواشناسی و علوم جو، azadi۶۸@hotmail.com

چکیده: در فرایند صدور پیش‌بینی وضع هوای آینده از برازش داده های ایستگاهی به شبکه منظم استفاده می‌شود. در تحقیق حاضر ضمن اشاره به کاربرد آمار فضایی در روش‌های برازش داده های ایستگاهی به شبکه منظم، به ارائه روش خودکار در تعیین ایستگاه های موثر در شبکه منظم موضعی پرداخته خواهد شد.
کلمات کلیدی: پیش‌بینی وضع هوا، برازش داده ها، شبکه منظم، آمار فضایی

۱ مقدمه

جهت معرفی کاربرد ریاضیات در هواشناسی به دلیل گستردگی حوزه کاربرد، در این مجال تنها به حوزه برازش داده‌های ایستگاهی به شبکه منظم در مرحله پیش‌پردازش و پس‌پردازش پرداخته شده است. در بخش ۲ در باره پیش‌بینی قطعی و انواع روش‌های پس‌پردازش در نقاط ایستگاهی توضیح داده خواهد شد. در بخش ۳ درباره نحوه پیش‌بینی در شبکه منظم به طور اجمال مطالبی ذکر خواهد شد. در ادامه (بخش ۴) به توصیف روش‌های تعمیم داده‌های دیدبانی ایستگاهی به شبکه منظم در مرحله پیش‌پردازش پرداخته می‌شود.

*نویسنده مسئول

۲ پیش‌بینی قطعی و روش‌های پس‌پردازش آن

به دلیل راه حل‌های تقریبی در روش‌های عددی، جواب بدست آمده از اجرای مدل‌های عددی تقریبی است. البته استفاده از روش‌های عددی در مدل‌ها، تنها علت بروز خطا در خروجی مدل نمی‌باشد؛ بلکه این خطا به عواملی مانند وجود خطا در مقادیر اولیه- مرزی، کمبود داده‌های اولیه در برخی مناطق و آشوبی بودن سیستم دینامیکی جو نیز وابسته است. خطاهای ذکر شده در برونداد مدل‌های پیش‌بینی عددی وضع هوا به دو دسته سامانمند و تصادفی تقسیم می‌شوند. خطاهای سامانمند میزان میانگین انحراف برونداد مدل از واقعیت را نشان می‌دهند که با روش‌های

www.SID.ir

Archive of SID

آماری موسوم به روش های پس‌پردازش قابل حذف می‌باشد؛ اما خطای تصادفی از نظم خاصی پیروی نمی‌کند. برای کاهش و یا حذف خطای سامانمند پیش‌بینی در ایستگاه‌های دیدبانی، روش‌های متعددی مانند میانگین‌گیری لغزان، پیش‌یابی کامل (PPM) ، روش آماره برونداد مدل(MOS) ، پالایه کالمن، شبکه‌هاي عصبي و شبکه‌هاي فازي- عصبي ارائه شده است.

۳ تعمیم پیش بینی پس پردازش شده به شبکه منظم

در سال‌های اخیر نیاز و تحقیق در خصوص پیش‌بینی بدون اریبی در نقاط غیر ایستگاهی رو به افزایش است. از جمله تحقیقات در این زمینه می توان به موارد زیر اشاره کرد.

چارلز و همکاران [۱] برای کالیبره کردن پیش‌بینی بارش روی نقاط منظم شبکه منطقه جنوب شرقی استرالیا از سه روش تورم واریانس و احتمال توأم بیزی و رگرسیون بردار تکین استفاده کردند و نشان دادند که روش احتمال توأم بیزی نتایج بهتری در مقایسه با دو روش دیگر دارد. مس و همکاران [۲]روشی را برای کاهش اریبی سیستماتیک پیش‌بینی دمای دومتری، دمای نقطه شبنم و بارش روی نقاط شبکه شمال غربی اقیانوس آرام ارائه کردند. برای کاهش اریبی سیستماتیک پیش‌بینی روی نقاط شبکه، جل [۳] دو روشLOB وCART–ACE را پیشنهاد کرد و هکر و رایف [۴] از یک تحلیل آماری برای از بین بردن اریبی در شبکه میان مقیاس استفاده کردند. گلان و همکاران [۵] با استفاده از میانگین‌گیری وزنی و لحاظ کردن میزان تأثیر تغییرات ارتفاع بر کاهش دما، پس‌پردازش در نقاط شبکه‌ای را انجام دادند.

روش BCDG توسط گلان و همکاران [۵] برای پس‌پردازش دمای ماکزیمم روی یک شبکه منظم معرفی شد. 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 پس‌پردازش شده و مقدار دیدبانی بدست می‌آید. این تفاضل برای نقاط شبکه‌ای که اطراف آن ایستگاه هستند به کار می‌رود و معمولاً با استفاده از فاصله ایستگاه تا نقطه شبکه وزن دهی 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 معمولا با ۴ بار تکرار انجام می‌شوند. به علت وابستگی دما به ارتفاع، در این روش علاوه بر تفاضل مقدار دیدبانی از مقدار درونیابی، تغییرات دما نسبت به تغییرات ارتفاع نیز در نظر گرفته شده است. مقدار بهبود یافته برای هر گره با توجه به دایره در نظر گرفته شده و ایستگاه‌های داخل دایره از فرمول تکراری زیر بدست می‌آید:

$$c_{t+1} = c_t + \frac{\frac{1}{n}\sum_{i=1}^{n} w_i D_i}{\frac{1}{n}\sum_{i=1}^{n} w_i} \qquad (1)$$

در معادله (۱) عبارت c_t مقدار گره در تکرار tام، n تعداد ایستگاه‌ها در دایره مفروض، w_i وزن هر ایستگاه که با فاصله‌اش تا گره نسبت عکس دارد و D_i ضابطه‌ای است که با تفاضل مقدار دیدبانی و مقدار پیش‌بینی شده با در نظر گرفتن تغییرات دما نسبت به ارتفاع رابطه مستقیم دارد:

$$D = S - G + VCE \times ELEDIF \qquad (2)$$

در معادله (۲) عبارت S مقدار دیدبانی در ایستگاه، D مقدار درونیابی داده گره به ایستگاه، G مقدار داده گره، $ELEDIF$ تفاضل ارتفاع گره از ارتفاع ایستگاه و VCE میزان تغییرات دما نسبت به تغییرات ارتفاع است. برای سنجش میزان بهبودی مقدار بدست آمده کافی است چند ایستگاه را به طور تصادفی انتخاب کرده

www.SID.ir

Archive of SID

و مقدار بدست آمده برای گره‌ها را به ایستگاه‌های مزبور درونیابی کنیم و با مقدار دیدبانی مقایسه کنیم.

۴ روش‌های تعمیم داده‌های دیدبانی ایستگاهی به شبکه منظم

می توان روش‌های درونیابی را به سه دسته قطعی، احتمالی و ترکیبی از قطعی و احتمالی طبقه‌بندی کرد [۶]. روش‌های قطعی با استفاده از مشخصات هندسی نقاط دیدبانی، یک سطح پیوسته ایجاد می‌کنند. روش‌های احتمالی با استفاده از نظریه احتمال و مفهوم تصادفی بودن و با در نظر گرفتن واریانس خطا درونیابی را انجام می‌دهند. دسته سوم نیز شامل ادغام روش‌های دو دسته قبل می‌باشد. از جمله روش‌های بکار رفته می توان به مدل‌های رگرسیونی، اسپلاین‌ها، دسته‌بندی، درخت رگرسیون، سری‌های فوریه و آنالیز سطح روند اشاره نمود.

روش GIDS [۷] یکی از روش‌های درونیابی قطعی است که نتیجه بهتری نسبت به بقیه روش‌ها برای کمیت دما داشته است. 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 می‌شود. سپس با توجه به طول و عرض جغرافیایی و ارتفاع و مقدار دمای مشاهداتی ایستگاه‌های انتخابی، یک مدل رگرسیون چندگانه و مختصات جغرافیایی ایستگاه‌ها و نقطه مورد نظر، مقدار دمای تخمین زده شده به صورت زیر بدست میـآید:

$$T_p = (\sum_{i=۱}^{n}(\frac{۱}{d_i})^۲)^{-۱} \times \sum_{i=۱}^{n}(T_i + b_۱(long_p - long_i) + b_۲(lat_p - lat_i) + b_۳(h_p - h_i))(\frac{۱}{d_i})^۲ \quad (۳)$$

که در معادله (۳) عبارت d_i فاصله ایستگاه i ام از نقطه، T، $long$، lat و h به ترتیب دما، طول، عرض و ارتفاع و ضرایب b ضرایب رگرسیونی و اندیس‌های p و i مربوط به نقطه و ایستگاه‌های اطراف نقطه می‌باشد.

روش‌های احتمالی بر پایه مفهوم تصادفی بودن استوار بوده و شامل روش‌هایی مانند مدل‌های رگرسیون خطی، درونیابی بهینه و زمین‌آماری می‌باشند. مکدونل [۸] نشان داد که وقتی داده‌ها دارای پراکندگی تنک هستند کریگیدن(kriging) بهترین روش درونیابی موجود می‌باشد. درونیابی کریگیدن بر این فرضیه است که تغییرپذیری فضایی داده‌ها نا منظم‌تر از آن است که با یک تابع ساده مدلسازی شود. در این روش مقدار تخمین زده شده به صورت زیر تعریف می‌شود:

$$z(x) = \sum_{i=۱}^{n} w_i z(x_i)$$

که $z(x)$ مقدار داده در نقاط پراکنده و w_i وزن‌هایی هستند که با استفاده از مینیمم کردن سمی واریانس که به صورت زیر می‌باشد بدست می‌آید:

$$\hat{\gamma} = \frac{۱}{۲n}\sum_{i=۱}^{n}(z(x_i) - z(x_i + h))^۲$$

که n تعداد زوج‌هایی از نقاط مشاهداتی است که دارای فاصله h از یکدیگر می‌باشند.

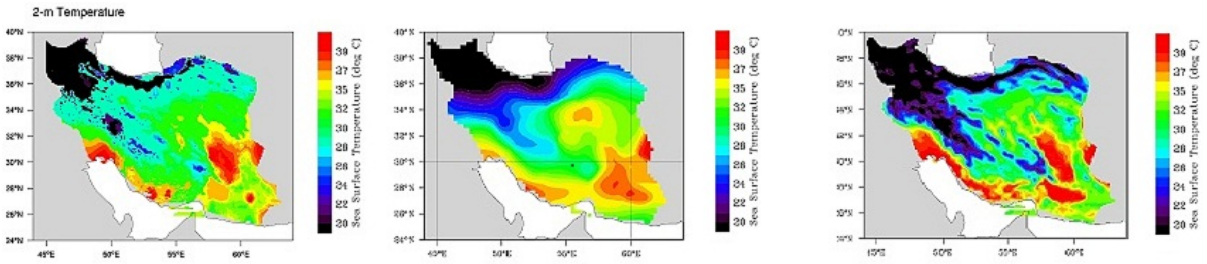
۵ نتیجه‌گیری

در این تحقیق با استفاده از دمای پیش‌بینی پس‌پردازش شده در ایستگاه‌ها به روش میانگین‌گیری لغزان، با استفاده از روش GIDS روی یک شبکه منظم دمای پیش‌بینی پس‌پردازش شده بدست آمده است. در روش‌های متداول تعداد ایستگاه‌ها به طور تجربی انتخاب می‌گردند. اما در تحقیق حاضر ضمن استفاده از آمار فضایی و اعمال مختصات بومی در محاسبات مربوط به تعیین دما روی شبکه از روش وریوگرام برای تعیین ایستگاه‌های تاثیرگذار بر داده‌های موضعی شبکه استفاده می شود.

نتایج حاصل از بکارگیری روش GIDS (شکل ۱سمت چپ) روی مدل WRF (شکل۱ سمت راست

www.SID.ir

Archive of SID



شکل ۱: دمای ایستگاهی و پردازش شده روی شبکه منظم

) و مشاهدات ایستگاهی (شکل ۱وسط) در تصویر نشان داده شده است. مقدار MAE برابر ۱٫۹۴ و مقدار RMSE برابر ۲٫۳ می باشد.

مراجع

[1] A. Charles, H. Hendon, Q. Wang, D. Robertson, a. E. Lim, Comparison of techniques for the calibration of coupled model forecasts of Murray Darling Basin seasonal mean rainfall. CAWCR Technical Report. 2011 ,vol. 40.

[2] C. Mass, J. Baars, G. Wedam, E. Grimit, A. R. Steed. Removal of Systematic Model Bias on a Model Grid. Weather And Forecasting. 2007, vol. 23, pp. 438-459.

[3] Y. Gel, 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. 2007, vol. 22, pp. 1243-1256.

[4] J. Hacker, A. D. Rife, A Practical Approach to Sequential Estimation of Systematic Error on Near-Surface Mesoscale Grids. Weather And Forecasting. 2007 , vol. 22, pp. 1257-1273.

[5] B. Glahn, K. Gilbert, R. Cosgrove, D. P. Ruth, K. Sheets, The gridding of MOS. Weather and Forecasting 2009, vol. 24, pp. 520-529.

[6] R. Sluiter, Interpolation methods for climate data. KNMI, R and D Information and Observation Technology. 2009.

[7] K. Stahl, R.D. Moore, J.A. Floyer, M.G. Asplin, I.G. McKendry, Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agricultural and Forest Meteorology. 2006, vol. 139, pp. 224–236.

[8] P.A. Burrough, R.A. McDonnell, Principles of Geographical Information Systems. Oxford University Press. 1998

www.SID.ir