

ارزیابی روش‌های درونیابی تأخیر وردسپهری حاصل از مشاهدات ایستگاه‌های پراکنده سامانه تعیین موقعیت جهانی

آیدا افشاری هرزویلی^۱ و یزدان عامریان^{۲*}

۱. دانشجوی کارشناسی ارشد، گروه ژئودزی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه‌نصیرالدین طوسی، تهران، ایران

۲. دانشیار، گروه ژئودزی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه‌نصیرالدین طوسی، تهران، ایران

چکیده

این مطالعه به بررسی روش‌های درونیابی مقدار بخار آب قابل بارش (Precipitable Water Vapor: PWV) با استفاده از داده‌های ایستگاه‌های (Global Positioning System: GPS) پراکنده در منطقه لس‌آنجلس می‌پردازد. منطقه مورد مطالعه به دلیل تنوع جغرافیایی و اقلیمی، شامل مناطق ساحلی، کوهستانی و دشت‌ها، و همچنین تغییرات فصلی، برای ارزیابی روش‌های مختلف انتخاب شده است. روش‌های مختلف درونیابی مورد بررسی، شامل عیارسنجی (Kriging)، ماشین بردار پشتیبان (Support Vector Machine: SVM)، جنگل تصادفی (Random Forest: RF)، همسایگی طبیعی (Natural Neighbor: NN) و شبکه عصبی مصنوعی (Artificial Neural Network: ANN)، بودند. ابتدا تأخیر تروپوسفری محاسبه و تأثیر پارامترهای هواشناسی مانند دمای سطح (Surface Temperature)، فشار سطح (Surface Pressure) و میانگین وزنی دما (Weighted Mean Temperature) بر PWV بررسی شد. نتایج نشان داد که مدل SVM به دلیل توانایی بالا در مدل‌سازی روابط غیرخطی، بهترین عملکرد را داشته و در مناطق کوهستانی دقت بیشتری ارائه داده است. همچنین، روش عیارسنجی نیز عملکرد مناسبی داشت، اما به دلیل فرض‌های ساده‌تر، ضعیف‌تر از SVM عمل کرد. جنگل تصادفی نیز به دلیل نیاز به داده‌های متراکم، نتایج مطلوبی ارائه نکرد. نتایج در تاریخ‌های ۲۴ ژوئیه ۲۰۲۱ و ۲۸ ژانویه ۲۰۲۲، با تحلیل‌های آماری تأیید شد. نقشه‌های متراکم از PWV نیز تهیه و تحلیل شدند که تغییرات زمانی و فضایی PWV را نشان دادند. این مطالعه به اهمیت انتخاب صحیح روش‌های درونیابی برای برآورد دقیق PWV و کاربرد آن‌ها در پیش‌بینی‌های جوی تأکید دارد.

واژه‌های کلیدی: بخار آب قابل بارش، جنگل تصادفی، سامانه تعیین موقعیت جهانی، شبکه عصبی مصنوعی، ماشین بردار پشتیبان.

۱. مقدمه

(Troposphere) و یونسپهر (Ionosphere)، دچار شکست می‌شوند و به جای طی کردن مسیر مستقیم، مسیر منحنی را طی می‌کنند. این پدیده در وردسپهر که در فناوری اطلاعات مکانی (Geomatics) به تأخیر تروپوسفری یا وردسپهری (Zenith Total Delay: ZTD) معروف است، یکی از مهم‌ترین چالش‌های GNSS به شمار می‌رود و منجر به کاهش دقت در تعیین موقعیت می‌شود (سیبر، ۲۰۰۳).

وردسپهر، پایین‌ترین لایه جو، حاوی حدود ۷۵ درصد از جرم کل جو و تقریباً تمام بخار آب موجود در هوا است (والاس و هابز، ۲۰۰۶). ضخامت این لایه بسته به فصل و موقعیت

سامانه‌های ناوبری جهانی ماهواره‌ای (Global Navigation Satellite Systems: GNSS)، یکی از پیشرفته‌ترین فناوری‌های راهبردی قرن بیست و یکم هستند که نقش محوری در گستره وسیعی از کاربردهای علمی، مهندسی و تجاری ایفا می‌کنند. این سامانه‌ها با دقت بالا امکان تعیین دقیق موقعیت مکانی، سرعت و زمان را در زمینه‌هایی چون زمین‌سنجی، نقشه‌برداری، ناوبری هوایی، کشاورزی دقیق و مدیریت بحران کاربرد گسترده‌ای یافته است (ولنهوف و همکاران، ۲۰۰۱). امواج GNSS هنگام عبور از جو زمین، به ویژه در لایه‌های وردسپهر

جغرافیایی بین ۷ تا ۱۸ کیلومتر متغیر است (کلیجر، ۲۰۰۴). تأخیر تروپوسفری به دلیل تأثیر دما، فشار و رطوبت در مسیر سیگنال‌های GNSS به وجود می‌آید و به دو مؤلفه اصلی تقسیم می‌شود: تأخیر آب‌ایستایی (Zenith Hydrostatic Delay: ZHD) و تأخیر تر (Zenith Wet Delay: ZWD). این اصطلاحات به طور خاص در مدل‌سازی و تصحیح سیگنال‌های GNSS مورد استفاده قرار می‌گیرند.

چالش اصلی در مدل‌سازی تأخیر تروپوسفری، پراکندگی و تراکم کم ایستگاه‌های GPS است. این پراکندگی باعث می‌شود که در بسیاری از مناطق داده‌های کافی برای تصحیح تأخیر وجود نداشته باشد. استفاده از روش‌های درون‌یابی برای برآورد این کمیت‌ها در مناطقی با ایستگاه‌های پراکنده ضروری است و این روش‌ها می‌توانند به دقت بیشتری در برآورد تأخیر تروپوسفری کمک کنند (لو و الموافی، ۲۰۱۱).

ساختار مقاله حاضر به شرح زیر است: در بخش دوم، به معرفی جزئیات مربوط به جمع‌آوری داده‌ها و روش‌های مورد استفاده برای درون‌یابی داده‌ها پرداخته می‌شود. بخش سوم به جزئیات ارزیابی نتایج اختصاص دارد. در نهایت، جمع‌بندی یافته‌ها و نتیجه‌گیری کلی از تحقیق در بخش چهارم ارائه شده است.

۲. روش پژوهش

در این بخش به بررسی و توضیح روش‌های مختلف درون‌یابی که برای ارزیابی تأخیر تروپوسفری با استفاده از داده‌های پراکنده ایستگاه‌های GPS و محاسبه مقدار PWV استفاده شده است، پرداخته می‌شود. از آنجا که تأخیر تروپوسفری به شدت به شرایط جوی و موقعیت مکانی وابسته است، استفاده از روش‌های پیشرفته درون‌یابی برای مدل‌سازی دقیق این تأخیر ضروری است. ابتدا مدل‌سازی و تصحیح تأخیر تروپوسفری و همچنین برآورد PWV بررسی می‌شود و سپس جزئیات مربوط به روش‌های درون‌یابی ارائه خواهد شد.

۲-۱. مدل‌سازی و تصحیح تأخیر تروپوسفری

ZHD به دلیل ثبات نسبی، حدود ۹۰ درصد از کل تأخیر تروپوسفری را شامل می‌شود. برای محاسبه ZHD از مدل ساستاموینن (Saastamoinen) استفاده شده است. این مدل با بهره‌گیری از کمیت‌های هواشناسی مانند فشار و ارتفاع ایستگاه، تأخیر خشک را به صورت دقیق برآورد می‌کند:

$$ZHD = \frac{0.002277P_S}{(1 - 0.0026\cos 2\varphi - 0.00000028H_S)} \quad (1)$$

که در آن؛ P_S فشار هوا در سطح زمین بر حسب هکتوپاسکال، φ عرض جغرافیایی ایستگاه و H_S ارتفاع ایستگاه بر حسب متر از سطح دریا است (ساستاموینن، ۱۹۷۲).

تأخیر کل تروپوسفری شامل دو مؤلفه ZHD و ZWD است. معادله ZTD به صورت زیر تعریف می‌شود (عبداله، ۲۰۱۶):

$$ZHD - ZTD = ZWD \quad (2)$$

بعد از محاسبه ZTD و ZHD، تأخیر تر تروپوسفری از طریق تفاضل ZHD و ZTD به دست می‌آید. ZWD به دلیل تأثیرات بخار آب در جو تغییرپذیر است و به صورت مستقیم بر برآورد دقیق کمیت PWV اثر می‌گذارد.

بخار آب قابل بارش به کل بخار آب موجود در ستونی از جو با سطح مقطعی به مساحت یک مترمربع گفته می‌شود که از سطح زمین تا بالاترین نقطه‌ای که بارش می‌تواند رخ دهد، امتداد دارد (کارلسون، ۱۹۹۳). برای محاسبه PWV از رابطه زیر استفاده شده است:

$$ZWD = ZTD - ZHD \quad (3)$$

که در آن Π ضریبی وابسته به شرایط جوی و عرض جغرافیایی است. این ضریب برای هر منطقه متفاوت است و بسته به دما و فشار هوا تغییر می‌کند (اگووا، ۲۰۱۵):

$$\Pi = \frac{10^6}{\rho_w \times R \times \left(\frac{k_3}{T_m} + k_2\right)} \quad (4)$$

در این رابطه مقادیر $k_2 = R = 0.4613 J.g^{-1}.K^{-1}$ و $k_3 = (3.739 \pm 0.12) \times 10^5 K^2.hpa^{-1}$ می‌باشند. ρ_w چگالی آب، R ثابت جهانی گاز، k_2 و k_3 ثابت‌های فیزیکی تجربی می‌باشند. از میان این کمیت‌ها، مهم‌ترین و تأثیرگذارترین مؤلفه، T_m میانگین وزنی دمای جو (Weighted Mean Temperature) بر حسب کلون است. این کمیت با استفاده از روش‌های مختلفی قابل محاسبه است، که در این مطالعه از رابطه ارائه شده توسط بیویس برای محاسبه میانگین وزنی دما استفاده شده است (بیویس و همکاران، ۱۹۹۲):

$$T_m = 85.63 + 0.668T_s \quad (5)$$

T_s دمای سطح (Surface Temperature) بر حسب کلون در مکان ایستگاه است و در مطالعه حاضر، از متغیر دمای ۲ متری (2m Temperature) موجود در داده‌های بازتحلیل شده ERA-5 برای آن استفاده شده است.

در این مطالعه، پارامترهای هواشناسی دمای سطحی و فشار سطحی به عنوان عوامل کلیدی در محاسبه PWV مورد استفاده قرار گرفته‌اند. دمای سطحی از طریق رابطه بیویس برای محاسبه T_m به کار رفته و به طور مستقیم بر مقدار Π تأثیر می‌گذارد. فشار سطحی نیز برای محاسبه ZHD به کار گرفته شده است، که یکی از اجزای اصلی در تعیین ZTD و در نتیجه PWV است.

۲-۲. جمع‌آوری داده‌ها

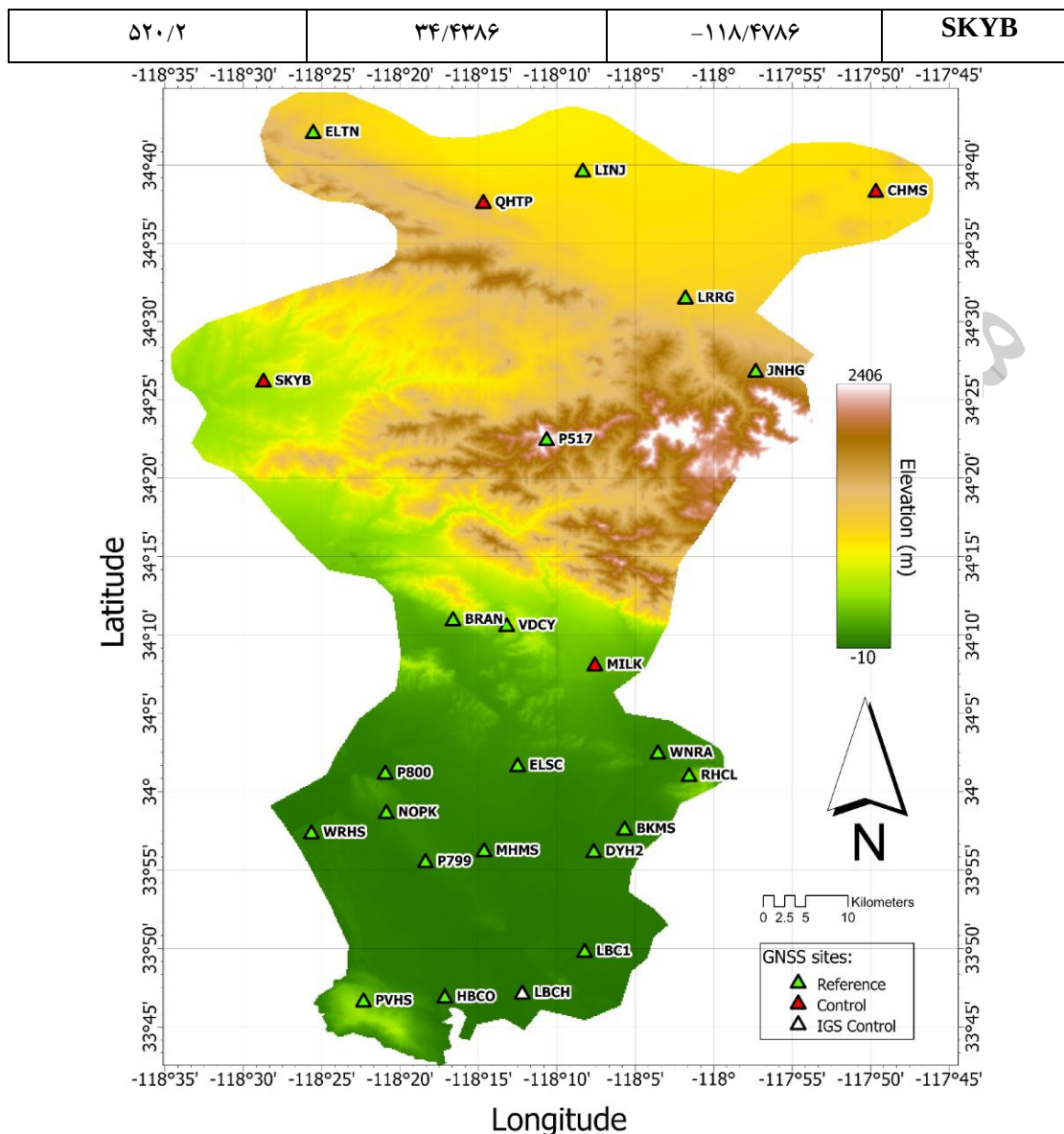
در این مطالعه، داده‌های GPS از ۲۵ ایستگاه پراکنده در محدوده جغرافیایی $33/75$ تا $34/75$ درجه شمالی و $117/5$ تا $118/5$ درجه غربی در منطقه لس‌آنجلس ایالات متحده آمریکا مورد استفاده قرار گرفتند. این منطقه به دلیل تنوع جغرافیایی و اقلیمی شامل مناطق ساحلی، کوهستانی و دشت‌ها انتخاب شد. این ویژگی‌ها امکان بررسی عملکرد مدل‌های درونیابی در شرایط مختلف فضایی و زمانی را فراهم می‌کند و چالشی مناسب برای ارزیابی دقت روش‌های مورد استفاده در پژوهش ایجاد می‌کند.

داده‌ها دو بازه زمانی ۲۴ تا ۳۰ ژوئیه ۲۰۲۱ (تابستان) و ۲۴ تا ۳۰ ژانویه ۲۰۲۲ (زمستان) جمع‌آوری شدند. این دو بازه زمانی به دلیل تفاوت‌های فصلی در مقادیر PWV و تأثیر تغییرات دما و رطوبت بر تأخیر تروپوسفری انتخاب شدند. در تابستان، دماهای بالاتر و رطوبت کمتر شرایطی متفاوت نسبت به زمستان با دماهای پایین‌تر و رطوبت بیشتر ایجاد می‌کنند. این شرایط به تحلیل عملکرد مدل‌ها در پیش‌بینی PWV در بازه‌های زمانی فضایی متنوع کمک می‌کند.

از ۲۵ ایستگاه موجود، ۲۰ ایستگاه به عنوان ایستگاه‌های مرجع شبکه UNAVCO برای پرآورد PWV استفاده شد و ۵ ایستگاه دیگر به عنوان ایستگاه‌های کنترل برای بررسی و مقایسه نتایج روش‌های درونیابی انتخاب شدند. جزئیات این ایستگاه‌ها در جدول ۱ ارائه شده است و موقعیت جغرافیایی به همراه توزیع ارتفاعی دقیقی منطقه لس‌آنجلس در شکل ۱ نشان داده شده است.

جدول ۱. جزئیات ایستگاه‌های کنترل.

نام ایستگاه	طول جغرافیایی (°)	عرض جغرافیایی (°)	ارتفاع ارتومتریک (m)
CHMS	-۱۱۸/۸۲۷۷	۳۴/۶۴۰۵	۷۶۰/۳
LBCH	-۱۱۸/۲۰۳۳	۳۳/۷۸۷۸	-۲۷/۶
MILK	-۱۱۸/۱۲۶۳	۳۴/۱۳۶۸	۲۴۷/۹
QHTP	-۱۱۸/۲۴۴۸	۳۴/۶۲۸۷	۸۶۳



شکل ۱. نقشه توزیع ارتفاعی منطقه لس آنجلس همراه با موقعیت ایستگاه‌های مرجع و کنترل GNSS.

۳-۲. منابع و پردازش داده‌ها

- منابع داده‌ها:

(۱) شبکه UNAVCO: داده‌های GPS شامل مختصات جغرافیایی (عرض، طول و ارتفاع)، ZTD، و اطلاعات زمانی از پایگاه داده شبکه UNAVCO استخراج شدند. این پایگاه داده یکی از معتبرترین منابع برای دسترسی به

داده‌های GNSS است. آدرس پایگاه داده به شرح زیر است:

<https://www.unavco.org/data/gps-gnss/data-access-methods/data-access-methods.html>

(۲) پایگاه داده ERA-5: داده‌های جوی شامل مقادیر دما و فشار سطحی از پایگاه داده بازتحلیل ERA-5 دریافت شدند. این داده‌ها در محاسبات مربوط به ZTD و PWV مورد

۲-۴. روش‌های اجرای درون‌یابی

در این مطالعه، پنج روش مختلف برای درون‌یابی PWV مورد ارزیابی قرار گرفت:

عیارسنجی: عیارسنجی یکی از پیشرفته‌ترین و پرکاربردترین روش‌های درون‌یابی فضایی است که برای برآورد مقادیر یک متغیر در نقاطی که فاقد داده استفاده می‌شود. این روش بر اساس همبستگی فضایی میان نقاط نمونه‌برداری شده عمل می‌کند (رحمان، ۲۰۱۸). در عیارسنجی، مقدار متغیر در نقطه‌ای ناشناخته ($Z(x)$) به عنوان ترکیبی خطی از مقادیر شناخته شده ($Z(x_i)$) در نقاط نمونه‌برداری شده (x_i) محاسبه می‌شود (رحمان، ۲۰۱۸):

$$Z(x) = \sum_{i=1}^n w_i Z(x_i) \quad (6)$$

که در آن؛ $Z(x_i)$ مقدار اندازه‌گیری شده در نقطه i ، w_i وزن عیارسنجی است که بر اساس فاصله و همبستگی فضایی تعیین می‌شود، n تعداد نقاط نمونه‌برداری شده است.

در این روش، ابتدا نیم‌پراش‌نما (Semivariogram) محاسبه می‌شود تا ساختار فضایی داده‌ها تحلیل و مدل پراش‌نگار (Variogram) مناسب برای آن تعیین شود. در این پژوهش، از مدل کروی (Spherical) به عنوان مدل پراش‌نگار استفاده شد که برای داده‌هایی با همبستگی قوی تا یک فاصله مشخص مناسب است (بلنج، ۲۰۰۴؛ وبستر و الیور، ۲۰۰۷).

پارامترهای مدل پراش‌نگار کروی به صورت خودکار بر اساس توزیع فضایی داده‌ها تعیین شدند تا دقت بالایی در برآورد فراهم شود. در این مطالعه، روش عیارسنجی معمولی سه‌بعدی (Ordinary Kriging Three Dimensional: OK 3D) به کار گرفته شد که از کمیت‌های عرض جغرافیایی، طول جغرافیایی و ارتفاع به عنوان ورودی مدل استفاده می‌کند. این کمیت‌ها امکان ایجاد پراش‌نگار فضایی دقیق و برآورد مقدار PWV در

استفاده قرار گرفتند. آدرس پایگاه داده به شرح زیر است:

<https://cds.climate.copernicus.eu/>

• پیش‌پردازش داده‌ها:

(۱) پردازش تأخیر تروپوسفری: داده‌های GPS با استفاده از نرم‌افزار برنیز (Bernese) پردازش و ZTD استخراج شدند. برای محاسبه ZHD، از مدل ساستاموین استفاده شد و مقادیر ZHD از ZTD تفاضل گرفته شد تا ZWD به دست آید. این مقادیر در محاسبات مرتبط با PWV به کار گرفته شدند.

(۲) داده‌های جوئی: مقادیر دما و فشار سطحی از داده‌های ERA-5 دریافت شدند و بر اساس رابطه ضریب PWV، Π به صورت دقیق محاسبه شد. این اطلاعات امکان تحلیل تأثیرات دما و فشار بر تغییرات PWV را فراهم کردند.

• پس‌پردازش داده‌ها:

(۱) نرمال‌سازی داده‌ها: برای اطمینان از مقایسه دقیق و درست داده‌ها، تمامی داده‌های ورودی مدل‌های یادگیری ماشین با استفاده از روش Min-Max نرمال‌سازی شدند. این روش داده‌ها را به یک باز استاندارد (۰ تا ۱) تبدیل کرد تا از اختلاف در مقیاس و دامنه داده‌ها جلوگیری شود. نرمال‌سازی باعث شد متغیرها برای تحلیل‌های آماری و یادگیری ماشین قابل مقایسه شوند و عملکرد مدل‌ها بهبود یابد.

(۲) ذخیره‌سازی نتایج: خروجی مدل‌ها و پیش‌بینی‌های انجام شده در قالب فایل‌های Excel ذخیره شدند تا امکان تحلیل دقیق و مقایسه نتایج فراهم شود. این فایل‌ها به صورت ساختاریافته برای ارزیابی عملکرد مدل‌های درون‌یابی و تحلیل آماری آماده شدند.

نقاطی که داده‌های مشاهده‌ای موجود نیست را فراهم می‌کنند (کمباردلا و همکاران، ۱۹۹۴).

- ماشین بردار پشتیبان: روش SVM یکی از روش‌های پرکاربرد در یادگیری ماشین (Machine Learning: ML) است که به دلیل قابلیت مدل‌سازی روابط غیرخطی و پیچیده، در این مطالعه به کار گرفته شد. این روش با استفاده از ترفند هسته (Kernel Trick)، داده‌ها را به فضای ویژگی با ابعاد بالاتر نگاشت می‌کند و یک ابرصفحه (Hyperplane) بهینه را برای جداسازی یا پیش‌بینی مقادیر تعیین می‌کند (کورتز و واپنیک، ۱۹۹۵؛ کریستیانینی و همکاران، ۲۰۰۱). در این مطالعه از نسخه وایزش این روش، یعنی SVR (Support Vector Regression)، برای پیش‌بینی مقادیر پیوسته PWV استفاده شد. رابطه SVR به صورت زیر است:

$$\min \frac{1}{2} \|w\|^2 \text{ subject to } \begin{cases} y_i - (wx_i + b) \leq \varepsilon \\ (wx_i + b) - y_i \leq \varepsilon \end{cases} \quad (V)$$

هسته تابع پایه شعاعی (Radial Basis Function Kernel: RBF Kernel) به دلیل توانایی در مدل‌سازی روابط غیرخطی میان داده‌ها انتخاب گردید (سو و همکاران، ۲۰۰۳؛ اسکولکوپف و اسمولا، ۲۰۱۸).

مقادیر پیش‌فرض برای پارامترهای کنترل نرمش (C) و پهنای باند کرنل (γ) در کتابخانه scikit-learn به کار گرفته شد. این مدل با استفاده از داده‌های نرمال‌سازی شده آموزش داده شد و به پیش‌بینی مقادیر PWV در ایستگاه‌های کنترل پرداخت.

به منظور حفظ شرایط یکسان میان روش‌ها و جلوگیری از سوگیری ناشی از تنظیمات متفاوت، پارامترهای پیش‌فرض مورد استفاده قرار گرفتند؛

همچنین ایستگاه‌های هدف از فرآیند داده‌های آموزشی حذف شدند تا احتمال بیش‌برازش کاهش یابد. شایان ذکر است که استفاده از مقادیر پیش‌فرض در مطالعات اولیه‌ی مقایسه‌ای روش‌ها، رویکردی رایج است، به‌ویژه زمانی که هدف، ارزیابی قابلیت نسبی مدل‌ها در شرایط یکسان و بدون اعمال تنظیمات پیچیده باشد.

- جنگل تصادفی: روش RF یکی از روش‌های قدرتمند یادگیری ماشین است که با ترکیب چندین درخت تصمیم (Decision Tree) و استفاده از نمونه‌برداری تصادفی، به پیش‌بینی دقیق و مقاوم در برابر نوفه (Noise) می‌پردازد (بریمن، ۲۰۰۱؛ کاتلر و همکاران، ۲۰۰۷؛ جیمز، ۲۰۱۳). در این مطالعه، جنگل تصادفی با ۵۰۰ درخت تصمیم‌گیری ($n_{estimators} = 500$) برای برآورد مقادیر PWV مورد استفاده قرار گرفت. تعداد ۵۰۰ درخت انتخاب شد تا تعادل مناسبی بین دقت و سرعت پردازش فراهم شود.

حداکثر عمق (max-depth) بدون محدودیت تنظیم گردید تا مدل بتواند پیچیدگی‌های داده‌ها را یاد بگیرد و روابط پیچیده میان متغیرهای ورودی، شامل ارتفاع، عرض و طول جغرافیایی، و مقادیر PWV را شناسایی کند. این روش توانایی بالایی در یادگیری الگوهای غیرخطی دارد و در بسیاری از موارد نتایج دقیقی ارائه می‌دهد (لیاو و وینر، ۲۰۰۲).

- همسایگی طبیعی: روش NN یکی از روش‌های پرکاربرد در درون‌یابی مکانی است که با استفاده از هندسه محاسباتی، به‌ویژه مثلث‌سازی دلونی (Delauny Triangulation) و سلول‌های ورونوی (Voronoi Cells)، مقادیر PWV را در نقاط فاقد داده برآورد می‌کند (آرونهامر، ۱۹۹۱؛ سوکومار و همکاران، ۱۹۹۸). این روش نیازی به تنظیم

فراکمیت‌ها (Hyperparameters) ندارد و به طور مستقیم با استفاده از هندسه محاسباتی و مثلث‌بندی دلونی، پیش‌بینی‌ها را انجام می‌دهد.

در این روش، وزن هر نقطه بر اساس مساحت اشتراکی سلول ورونوی همسایگان نزدیک تعیین می‌شود (سامریج و همکاران، ۱۹۹۵). این ویژگی باعث می‌شود که پیوستگی مکانی داده‌ها حفظ شود و نقشه‌های درون‌یابی شده بدون ناپیوستگی‌های غیرطبیعی باشند (سوتاهاران، ۲۰۱۶).

- شبکه عصبی مصنوعی: روش ANN از نوع پرسپترون چندلایه (Multilayer Perceptron Neural Network: MLPNN) در این مطالعه مورد استفاده قرار گرفت. این مدل شامل دو لایه پنهان با تعداد عصب (Neuron) های بهینه (۱۲ و ۶ عصب) بود که با استفاده از روش آزمون و خطا تعیین شدند (رهبر، ۲۰۱۶). مدل ANN با استفاده از راهکار پس‌انتشار خطا (Error Back Propagation Algorithm) و داده‌های نرمال‌شده آموزش داده شد و توانست الگوهای غیرخطی موجود در داده‌ها را یاد بگیرد (ژانگ و همکاران، ۲۰۲۲). راهکار گرادیان تصادفی (Stochastic Gradient Descent: SGD) با نرخ یادگیری ۰.۰۱ برای آموزی به کار گرفته شد. داده‌ها قبل از هر دوره (Epoch) به صورت تصادفی مرتب شدند و به دو مجموعه آموزشی (۸۰٪) و اعتبارسنجی (۲۰٪) تقسیم شدند. این فرآیند به بهینه‌سازی مدل و جلوگیری از بیش‌برازش (Overfitting) کمک می‌کند.

نتایج حاصل از روش‌های مختلف درون‌یابی و مقایسه دقت آن‌ها در برآورد تأخیر تروپسفری و PWV در ایستگاه‌های کنترل در ۲ روز خاص مربوط به ۲۴ ژوئیه ۲۰۲۱ و ۲۸ ژانویه ۲۰۲۲ که به عنوان

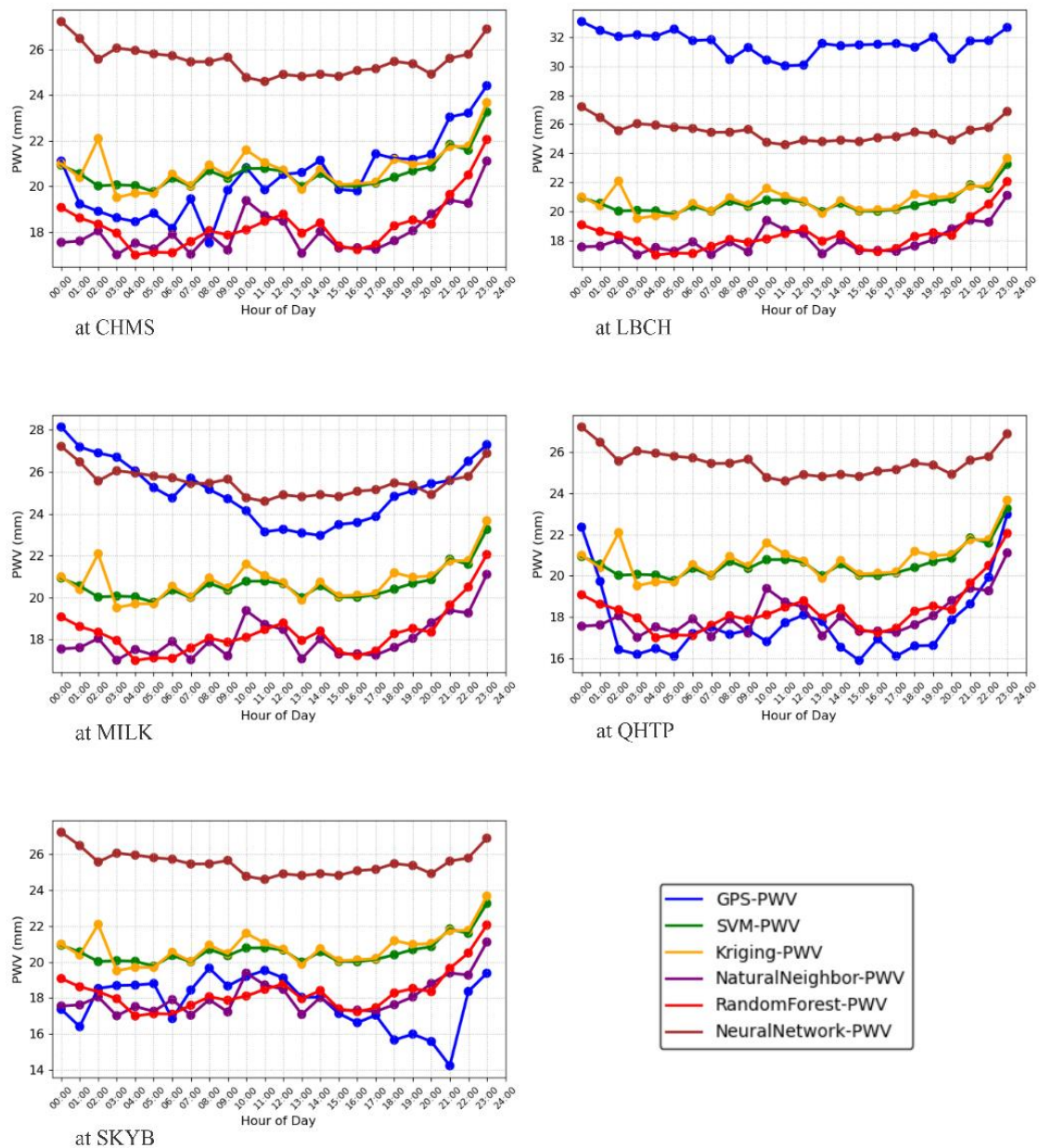
نماینده شرایط فصلی متفاوت در نظر گرفته شده‌اند، در بخش بعد مورد تحلیل قرار می‌گیرد.

۳. بحث

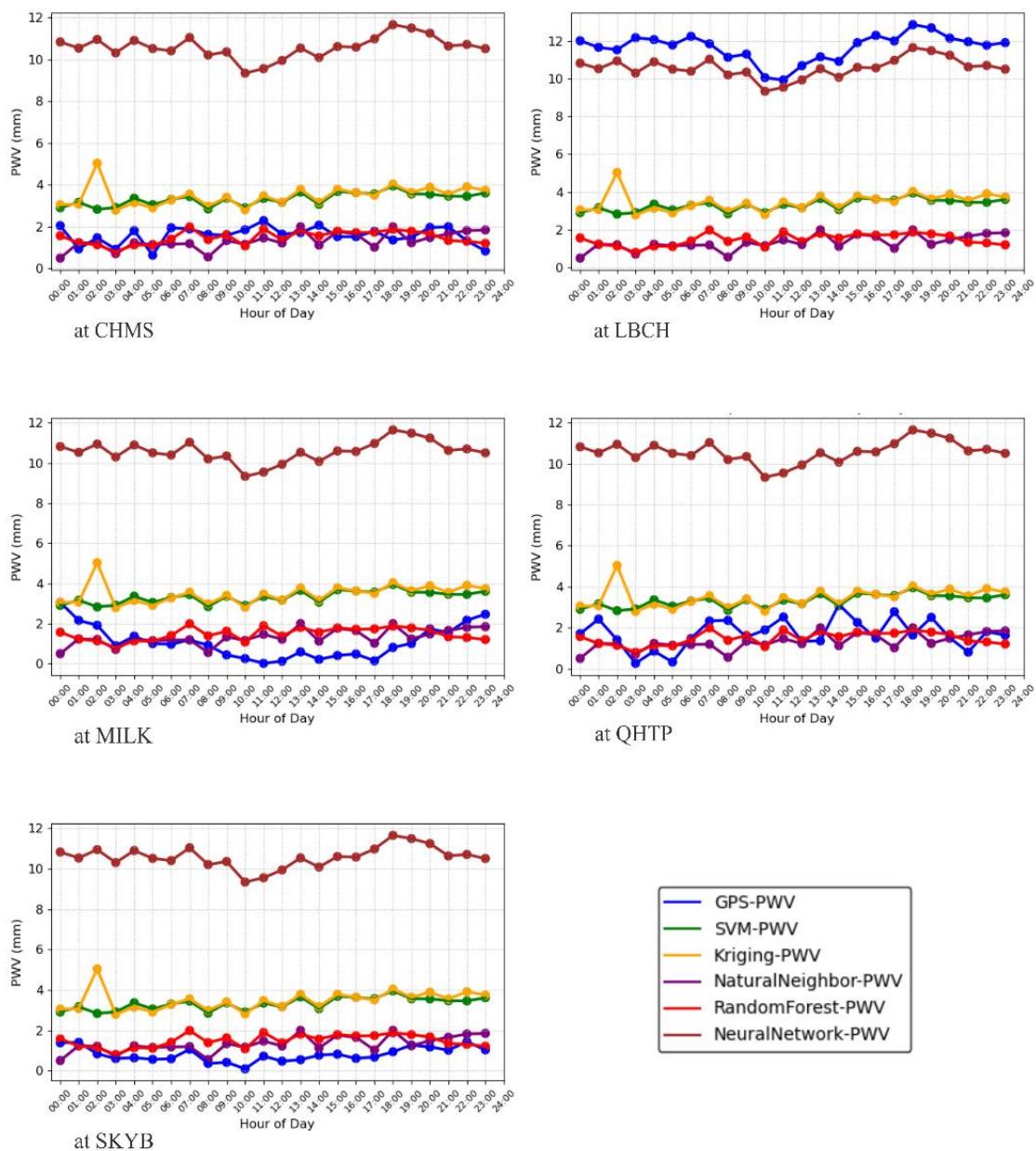
در این بخش، برای تحلیل عملکرد مدل‌های مختلف درون‌یابی PWV، در ۵ ایستگاه کنترل (CHMS, LBCH, MILK, QHTP, SKYB) دو تاریخ ۲۴ ژوئیه ۲۰۲۱ (تابستان) و ۲۸ ژانویه ۲۰۲۲ (زمستان) در نظر گرفته شدند. لازم به ذکر است که مدل‌های درون‌یابی برای هر یک از این دو روز به صورت جداگانه آموزش داده شده‌اند؛ به این ترتیب که برای هر تاریخ، داده‌های ایستگاه‌های GPS در همان روز به عنوان ورودی آموزشی برای برآورد مقادیر PWV در ایستگاه‌های هدف مورد استفاده قرار گرفتند. موقعیت این ایستگاه‌ها در شکل ۱ نشان داده شده و مقادیر پیش‌بینی شده توسط مدل‌ها در مقایسه با مقادیر مشاهده شده، در نمودارهای شکل‌های ۲ و ۳ با تفکیک زمانی ۱ ساعت ارائه شده‌اند.

شکل ۲ نشان‌دهنده عملکرد مدل‌ها در تابستان است، جایی که تغییرات سریع دما و رطوبت موجب نوسانات بیشتر در مقادیر PWV شده و دقت پیش‌بینی مدل‌ها را کاهش داده است. در این بازه، به‌ویژه در ساعات اوج تابش خورشید، خطای مدل‌ها افزایش چشمگیری داشته است. در مقابل، شکل ۳ عملکرد مدل‌ها در زمستان را نشان می‌دهد، که به دلیل پایداری بیشتر شرایط جوّی، دقت بیشتری داشته و نوسانات خطا کمتر بوده است.

مقایسه نتایج نشان می‌دهد که مدل SVM به طور کلی بهترین عملکرد را داشته است، به‌ویژه در شرایطی که نوسانات داده‌ها زیاد بوده است. با این حال، در زمستان که شرایط جوّی پایدارتر بوده، مدل‌های RF و NN توانسته‌اند عملکردی قابل توجه ارائه دهند. این یافته‌ها نشان می‌دهند که شرایط فصلی و جوّی تأثیر قابل توجهی بر دقت مدل‌های درون‌یابی دارند و باید در طراحی و بهینه‌سازی این روش‌ها در نظر گرفته شوند.



شکل ۲. نتایج مربوط به روش‌های درون‌یابی برای ۵ ایستگاه کنترل با تفکیک زمانی ۱ ساعت در تاریخ ۲۴ ژوئیه ۲۰۲۱.



شکل ۳. نتایج مربوط به روش‌های درون‌یابی برای ۵ ایستگاه کنترل با تفکیک زمانی ۱ ساعت در تاریخ ۲۸ ژانویه ۲۰۲۲.

۳-۱. مقایسه روش‌ها

- در نمودارهای جعبه‌ای مربوط به روزهای ۲۴ ژوئیه ۲۰۲۱ و ۲۸ ژانویه ۲۰۲۲ در شکل ۴، تفاوت عملکرد روش‌های مختلف درون‌یابی بررسی شده است. این نمودارها نشان می‌دهند که روش‌های عیارسنجی و

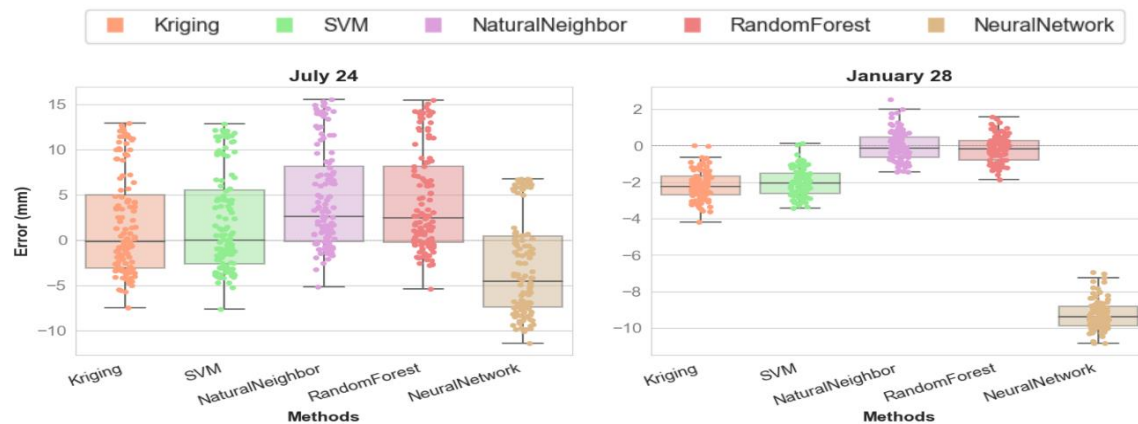
SVM در هر دو بازه زمانی بهترین عملکرد را از نظر دقت برآورد داشته‌اند. خطاهای این مدل‌ها کوچک بوده و میانگین خطا نزدیک به صفر است، که نشان‌دهنده دقت بالای این روش‌ها است. در نمودارهای پراکندگی برای دو روز ۲۴ ژوئیه ۲۰۲۱ و ۲۸ ژانویه ۲۰۲۲ در شکل‌های ۵ و ۶، رابطه بین

مقادیر پیش‌بینی شده توسط هر مدل و مقادیر واقعی PWV بررسی شده است. مدل‌های عیارسنجی و SVM در هر دو روز پراکندگی کمتری را نشان می‌دهند، که به معنای تطابق بهتر بین مقادیر پیش‌بینی شده و مقادیر مرجع است.

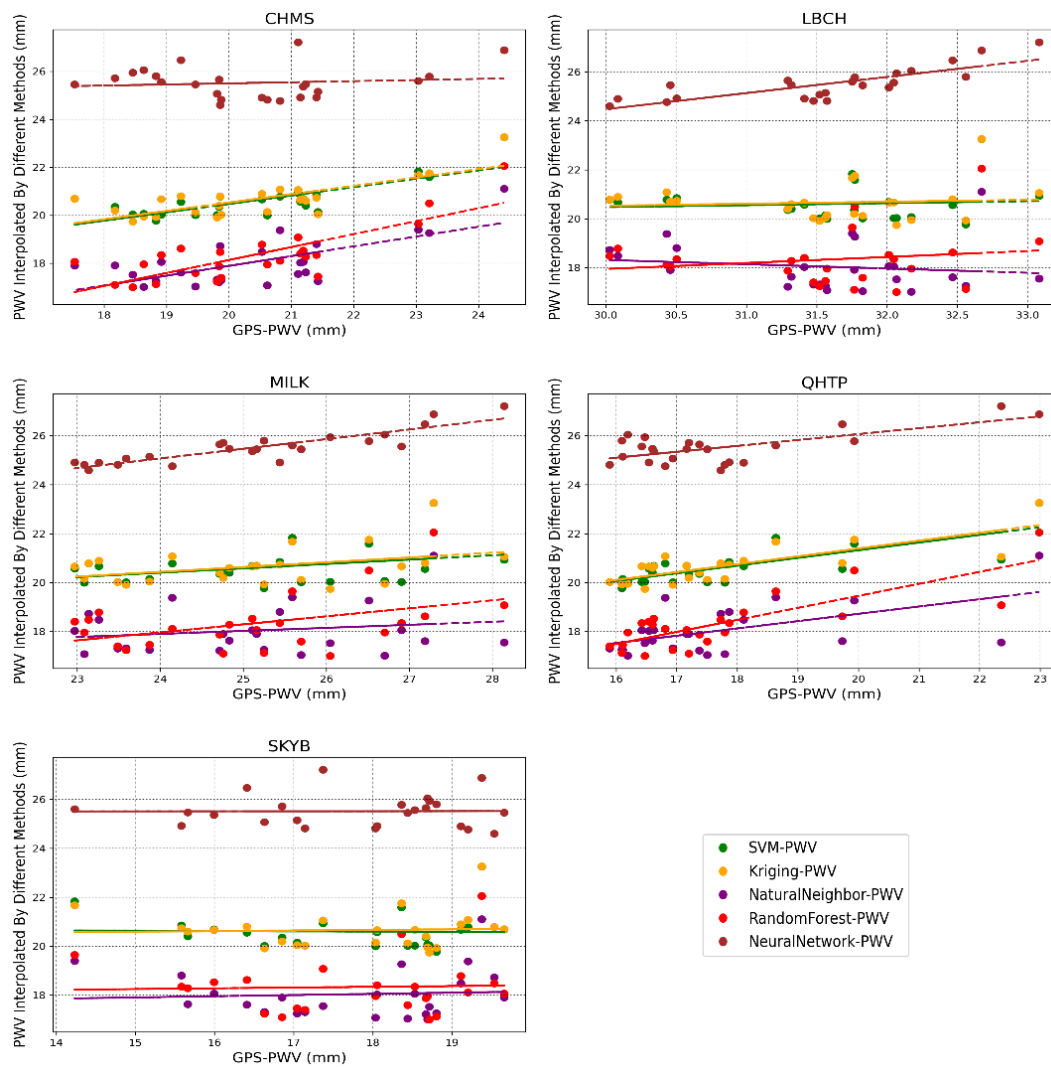
بزرگتر بودن خطای پیش‌بینی مدل‌ها در تابستان نسبت به زمستان، ناشی از شرایط جوئی ناپایدارتر و نوسانات بزرگتر در مقادیر PWV در تابستان است. در تابستان، تغییرات سریع دما و رطوبت، نوسانات بیشتری در مقادیر PWV ایجاد می‌کند که پیش‌بینی را دشوارتر می‌سازد. همچنین، بررسی نمودارهای جعبه‌ای نشان می‌دهد که مدل‌های RF و NN در

زمستان عملکرد بهتری نسبت به SVM دارند. این موضوع به پایداری بیشتر شرایط جوئی در زمستان و توانایی ذاتی RF و NN در کاهش نوفه و تطبیق با داده‌های با نوسانات کمتر مرتبط است.

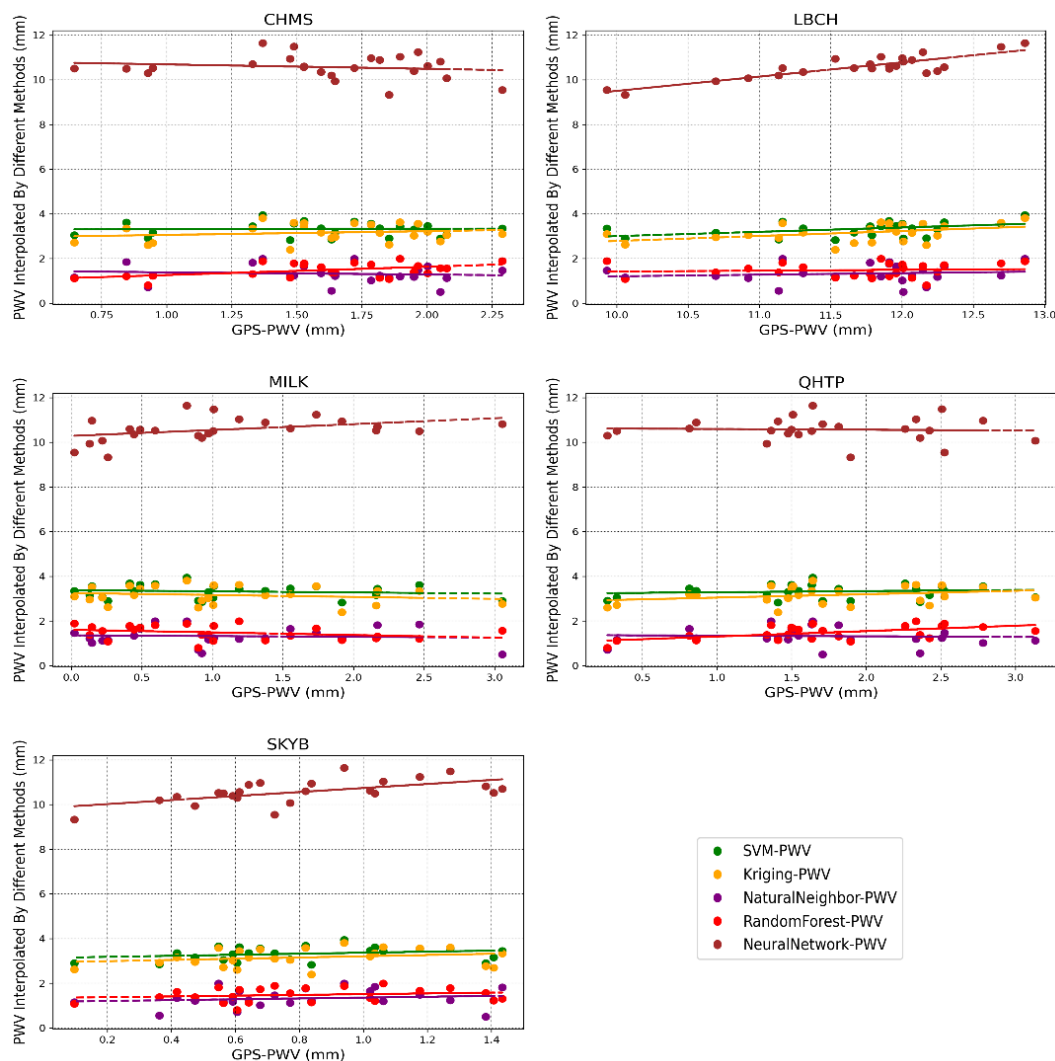
در مقابل، SVM به دلیل تنظیمات پیش‌فرض و حساسیت به نوسانات بزرگ‌تر، در تابستان دقت کمتری نشان داده است. این یافته‌ها نشان می‌دهند که انتخاب متغیرهای اختصاصی برای هر فصل یا تنظیم کمیت‌های مدل‌ها بر اساس شرایط فصلی می‌تواند در بهبود دقت مدل‌ها مؤثر باشد و تفاوت‌های فصلی در خطای پیش‌بینی را کاهش دهد.



شکل ۴. نمودارهای جعبه‌ای تفاضل مقادیر مدل و مقادیر کنترل در روز ۲۴ ژوئیه ۲۰۲۱ (سمت چپ) و روز ۲۸ ژانویه ۲۰۲۲ (سمت راست).



شکل ۵. نمودار پراکندگی روش‌های مختلف درونیابی مربوط به ۵ ایستگاه کنترل در تاریخ ۲۴ ژوئیه ۲۰۲۱.



شکل ۶. نمودار پراکندگی روش‌های مختلف درون‌یابی مربوط به ۵ ایستگاه کنترل در تاریخ ۲۸ ژانویه ۲۰۲۲.

۲۰۲۱ به ترتیب از دقت بالا به پایین قرار دارند. در حالی که در روزهای ۲۴ تا ۳۰ ژانویه ۲۰۲۲ این ترتیب کمی تغییر می‌کند؛ روش SVM همچنان دقیق‌ترین روش است، اما سایر روش‌ها به ترتیب دقت از بالا به پایین به صورت RF، عیارسنجی، ANN و NN قرار می‌گیرند.

جدول ۲. مقایسه آماری شاخص RMSE با توجه به نتایج به دست آمده در تاریخ ۲۴ تا ۳۰ ژوئیه ۲۰۲۱.

- جدول‌های ۲ و ۳، نشان‌دهنده مقایسه آماری شاخص ریشه میانگین مربع خطاها (Root Mean Square Error: RMSE) با توجه به نتایج به دست آمده در تاریخ‌های ۲۴ تا ۳۰ ژوئیه ۲۰۲۱ و ۲۴ تا ۳۰ ژانویه ۲۰۲۲ برای روش‌های مختلف درون‌یابی در هر ایستگاه کنترل می‌باشد. روش‌های SVM، RF، کریجینگ، NN و ANN در روزهای ۲۴ تا ۳۰ ژوئیه

روش	ایستگاه	CHMS RMSE(mm)	LBCH RMSE(mm)	MILK RMSE(mm)	QHTP RMSE(mm)	SKYB RMSE(mm)
کریجینگ		۲/۳	۲	۱/۷	۲/۱	۲
ماشین بردار پشتیبان		۲/۲	۱/۲	۱/۱	۱	۱/۹
جنگل تصادفی		۱/۴	۲/۷	۲/۳	۲	۲
همسایگی طبیعی		۲	۲/۴	۲/۱	۲/۲	۲/۳
شبکه عصبی مصنوعی		۴/۱	۲	۱/۶	۳/۷	۲/۷

جدول ۳. مقایسه آماری شاخص RMSE با توجه به نتایج به دست آمده در تاریخ ۲۴ تا ۳۰ ژانویه ۲۰۲۲.

روش	ایستگاه	CHMS RMSE(mm)	LBCH RMSE(mm)	MILK RMSE(mm)	QHTP RMSE(mm)	SKYB RMSE(mm)
کریجینگ		۰/۷	۰/۹	۱	۰/۸	۰/۸
ماشین بردار پشتیبان		۰/۶	۰/۶	۰/۷	۰/۷	۰/۷
جنگل تصادفی		۰/۵	۰/۹	۰/۹	۰/۷	۰/۷
همسایگی طبیعی		۰/۷	۱	۱/۱	۰/۸	۰/۹
شبکه عصبی مصنوعی		۱/۱	۰/۸	۰/۹	۱/۱	۰/۹

- جدول‌های ۴ و ۵، نشان‌دهنده مقایسه آماری ریشه نسبی میانگین مربع خطاها (Relative Root Mean Square Error: RRMSE) با توجه به نتایج به دست آمده در تاریخ‌های ۲۴ تا ۳۰ ژوئیه ۲۰۲۱ و ۲۴ تا ۳۰ ژانویه ۲۰۲۲ برای روش‌های مختلف درون‌یابی در هر ایستگاه کنترل می‌باشد. روش‌های SVM، RF، عیارسنجی، NN و ANN در روزهای ۲۴ تا ۳۰ ژوئیه ۲۰۲۱ به ترتیب دقت از بالا به پایین قرار دارند. در

حالی‌که در روزهای ۲۴ تا ۳۰ ژانویه ۲۰۲۲ این ترتیب کمی تغییر می‌کند؛ روش SVM همچنان دقیق‌ترین روش است، اما سایر روش‌ها به ترتیب دقت از بالا به پایین به صورت RF، عیارسنجی، ANN و NN قرار می‌گیرند. این نتایج نشان‌دهنده تغییرات مقادیر PWV در روزها و ساعات مختلف شبانه‌روز است. نکته مهم این است که در اکثر موارد، روش SVM عملکرد بهتری داشته و دقت بیشتری را نشان می‌دهد.

جدول ۴. مقایسه آماری شاخص RRMSE با توجه به نتایج به دست آمده در تاریخ ۲۴ تا ۳۰ ژوئیه ۲۰۲۱.

روش	ایستگاه	CHMS RRMSE(%)	LBCH RRMSE(%)	MILK RRMSE(%)	QHTP RRMSE(%)	SKYB RRMSE(%)
کریجینگ		۹/۵	۸	۷/۷	۸/۵	۸

ماشین بردار پشتیبان	۹/۲	۴/۳	۴/۸	۴/۵	۷/۹
جنگل تصادفی	۶/۲	۱۱/۶	۱۰	۸/۱	۸/۵
همسایگی طبیعی	۹/۵	۱۱/۳	۱۰	۱۰/۴	۱۰/۹
شبکه عصبی مصنوعی	۱۴/۵	۸	۶/۸	۱۲/۹	۹/۳

جدول ۵. مقایسه آماری شاخص RRMSE با توجه به نتایج به دست آمده در تاریخ ۲۴ تا ۳۰ ژانویه ۲۰۲۲.

روش	ایستگاه	CHMS RRMSE(%)	LBCH RRMSE(%)	MILK RRMSE(%)	QHTP RRMSE(%)	SKYB RRMSE(%)
کریجینگ		۱۳/۷	۱۸/۷	۲۰/۹	۱۷/۳	۱۶/۴
ماشین بردار پشتیبان		۱۲/۳	۴/۹	۷/۲	۹/۴	۱۵
جنگل تصادفی		۱۶/۴	۲۰/۶	۲۳/۴	۲۳/۱	۲۴
همسایگی طبیعی		۲۶/۷	۲۹/۸	۳۲	۳۲/۶	۳۳
شبکه عصبی مصنوعی		۸/۹	۱۷/۴	۱۰/۱	۱۵/۶	۸/۲

۳-۱-۱. تهیه نقشه‌های متراکم از PWV

با توجه به دقت قابل توجه روش SVM نسبت به سایر روش‌ها در قسمت قبل، در ادامه از این روش به منظور تحلیل دقیق توزیع فضایی PWV در منطقه لس‌آنجلس استفاده شد. نقشه‌های PWV برای دو تاریخ ۲۴ ژوئیه ۲۰۲۱ و ۲۸ ژانویه ۲۰۲۲ در بازه‌های زمانی ۴ ساعت تهیه شدند. این نقشه‌ها توزیع PWV را در سطح منطقه نشان می‌دهند و می‌توانند در تحلیل‌های جوی و پیش‌بینی‌های هواشناسی استفاده شوند. ابتدا داده‌های ارتفاعی با تفکیک مکانی ۳۰ متر از مدل ارتفاعی رقومی ناسا (National Aeronautics and Space Administration Digital Elevation Model: NASA DEM) استخراج شد. سپس این داده‌ها برای بهینه‌سازی محاسبات به تفکیک مکانی ۹۰ متر نمونه‌برداری شدند. مختصات جغرافیایی و ارتفاع مرکز هر نقطه شبکه به عنوان ورودی‌های اصلی مدل SVM استفاده شدند. مدل با استفاده از داده‌های مرجع آموزش داده شد و PWV را در ساعات مختلف شبانه‌روز پیش‌بینی کرد. نقشه‌های نهایی با تفکیک مکانی ۹۰ متر

برای دو تاریخ ۲۴ ژوئیه ۲۰۲۱ (تابستان) و ۲۸ ژانویه ۲۰۲۲ (زمستان) تهیه شدند.

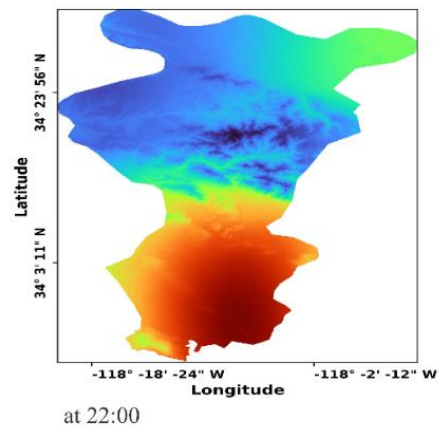
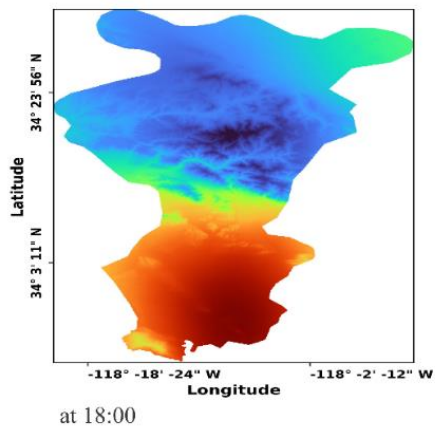
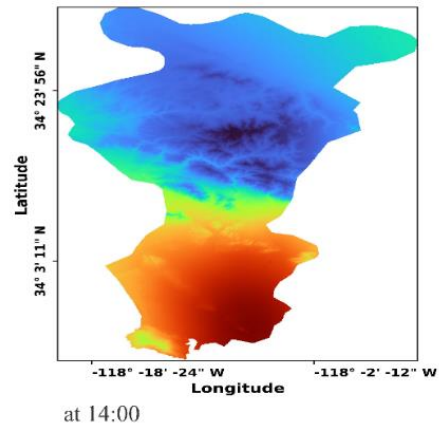
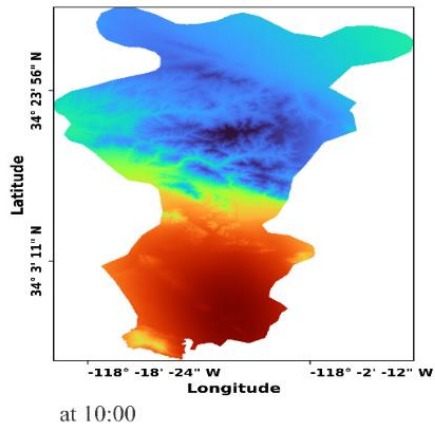
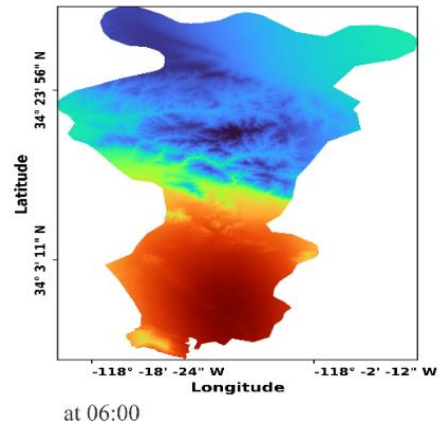
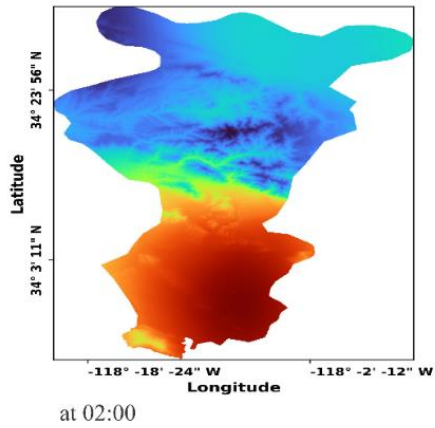
۳-۱-۲. تحلیل نقشه‌های متراکم از PWV

نقشه‌های PWV در شکل‌های ۸ و ۹ نشان می‌دهند که مقادیر PWV در طول روز و در بازه‌های زمانی متفاوت دچار تغییرات قابل توجهی می‌شوند.

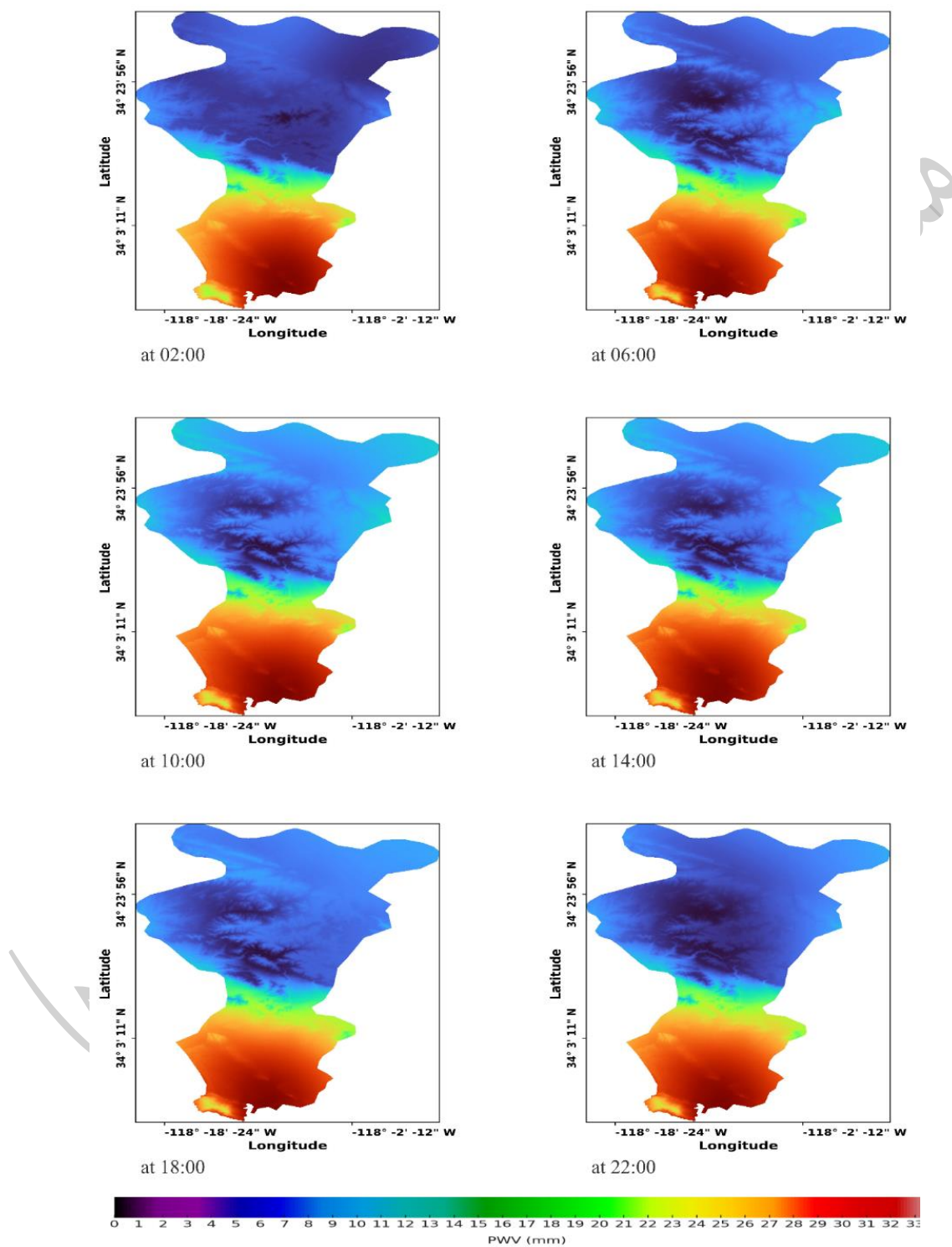
- در تاریخ ۲۴ ژوئیه ۲۰۲۱ (تابستان)، مقادیر PWV در ساعات اولیه صبح در مناطق پایین‌تر و نزدیک به دریا به دلیل رطوبت بالا بیشتر بوده است. با پیشروی روز و افزایش دما، PWV نیز افزایش یافته و در ساعات بعد از ظهر بیشترین مقدار در مناطق ساحلی مشاهده شده است. در شب، PWV مجدداً کاهش یافته و توزیع به حالت اولیه بازمی‌گردد.
- در تاریخ ۲۸ ژانویه ۲۰۲۲ (زمستان)، الگوی مشابهی با تابستان مشاهده شد؛ با این تفاوت که تغییرات PWV در زمستان محدودتر و بیشتر به مناطق ساحلی منحصر بود. در ساعات بعد از ظهر مقادیر PWV به

اوج خود رسید، اما به دلیل سرمای هوا این افزایش نسبت به تابستان محدودتر بود. این شرایط نشان‌دهنده

تأثیر مستقیم شرایط جوی تابستانی و زمستانی بر توزیع فضایی PWV در منطقه لس‌آنجلس است.



شکل ۷. نقشه‌های متراکم از مقدار بخار آب قابل بارش در تاریخ ۲۴ ژوئیه ۲۰۲۱.



شکل ۸. نقشه‌های متراکم از مقدار بخار آب قابل بارش در تاریخ ۲۸ ژانویه ۲۰۲۲.

۴. نتیجه گیری

این مطالعه با هدف ارزیابی روش های مختلف درون یابی بخار آب قابل بارش بر پایه داده های ایستگاه های پراکنده GPS در منطقه لس آنجلس انجام شد. تنوع جغرافیایی و اقلیمی این منطقه، شامل مناطق ساحلی، دشت ها و کوهستانی، شرایطی ایده آل برای بررسی عملکرد مدل ها در شرایط مختلف فضایی و فصلی فراهم کرد. علاوه بر این، دو بازه زمانی مشخص در تابستان (۲۴ ژوئیه ۲۰۲۱) و زمستان (۲۸ ژانویه ۲۰۲۲) انتخاب شدند تا تأثیر شرایط جوی و فصلی بر دقت مدل ها تحلیل شود.

نتایج نشان داد که روش ماشین بردار پشتیبان به دلیل قابلیت بالای خود در مدل سازی روابط غیرخطی و پیچیده، بهترین عملکرد را در برآورد PWV داشت. این روش توانست در هر دو بازه زمانی تابستان و زمستان، دقت بالایی را ارائه دهد، به ویژه در مناطق با کوهساری پیچیده. همچنین، نمودارهای جعبه ای و پراکندگی نشان دادند که مقادیر پیش بینی شده توسط SVM تطابق بالایی با مقادیر واقعی دارند و پراکندگی کمتری در داده ها مشاهده شد.

روش عیارسنجی به عنوان دومین روش دقیق شناسایی شد. این روش، هرچند به دلیل مدل سازی ساده تر، دقت کمتری نسبت به SVM داشت، اما در مناطقی با تراکم بالای ایستگاه ها عملکرد مناسبی ارائه داد. جنگل تصادفی و همسایگی طبیعی نیز در زمستان به دلیل شرایط جوی پایدارتر، نتایج بهتری نسبت به تابستان داشتند. روش شبکه عصبی مصنوعی به دلیل تنظیمات بهینه نشده و تعداد محدود لایه های پنهان، عملکرد ضعیف تری نسبت به SVM نشان داد. با این حال، بهینه سازی تنظیمات ANN می تواند به بهبود دقت آن منجر شود.

تحلیل نقشه های توزیع فضایی PWV نشان داد که شرایط جوی تابستانی باعث افزایش نوسانات و پیچیدگی توزیع PWV می شود، در حالی که در زمستان توزیع یکنواخت تر و با نوسانات

کمتری است. این تحلیل همچنین نقش مهم عوامل جوی مانند دما، فشار و رطوبت را در تغییرات PWV برجسته کرد.

همچنین، برای بهبود دقت پیش بینی مدل های درون یابی، پیشنهاد می شود از روش های تنظیم پارامتر مانند GridSearchCV و RandomizedSearchCV برای تعیین بهینه ترین مقادیر پارامترها در مدل های یادگیری ماشین، به ویژه در روش های حساس مانند SVM و RF استفاده شود. این کار می تواند خطر بیش برآزش را کاهش داده و دقت پیش بینی را در شرایط مختلف جوی افزایش دهد.

علاوه بر این، ترکیب مدل های مختلف مانند ترکیب SVM با Kriging یا استفاده از رویکردهای (Ensemble) می تواند باعث افزایش پایداری و دقت نتایج شود. همچنین، تنظیم ساختار مدل ها و کمیت های آن ها با توجه به ویژگی های فصلی و محلی داده ها، از جمله پراکندگی مکانی ایستگاه ها و شدت تغییرات PWV، می تواند به بهبود عملکرد مدل ها کمک کند. استفاده از داده های بیشتر در بازه های زمانی مختلف و شرایط جوی متنوع نیز موجب افزایش تعمیم پذیری مدل ها خواهد شد. یافته های این پژوهش می توانند مبنایی برای طراحی سیستم های هوشمند پیش بینی پارامترهای جوی و بهبود تحلیل های هواشناسی در مناطق مختلف با ویژگی های مشابه جغرافیایی و اقلیمی قرار گیرند.

مراجع

- Abdallah, A. T. M. (2016). Bernese GNSS Software Handout University of Stuttgart, Germany.
- Aurenhammer, F. (1991). Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)* **23**, 345-405.
- Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H. (1992). GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system. *Journal of Geophysical Research: Atmospheres* **97**, 15787-15801.

- Rahbar, S. S., M. (2016). Comparison of different methods of determining the geodetic height correction level: A case study of Tehran city. *Iran Geophysics Journal* **10**, 40-52.
- Rahman, H. (2018). Evaluation of Synthetic CPT and Soil Boring Data by Various Spatial Interpolation Techniques.
- Saastamoinen, J. (1972). Contributions to the theory of atmospheric refraction. *Bulletin Géodésique (1946-1975)* **105**, 279-298.
- Sambridge, M., Braun, J., and McQueen, H. (1995). Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophysical Journal International* **122**, 837-857.
- Scholkopf, B., and Smola, A. J. (2018). "Learning with kernels: support vector machines, regularization, optimization, and beyond", MIT press.
- Seeber, G. (2003). "Satellite Geodesy: foundations, methods and application," 2nd Edition/Ed.
- Sukumar, N., Moran, B., and Belytschko, T. (1998). The natural element method in solid mechanics. *International journal for numerical methods in engineering* **43**, 839-887.
- Suthaharan, S. (2016). Support Vector Machine. In "Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning" (S. Suthaharan, ed.), pp. 207-235. Springer US, Boston, MA.
- Wallace, J. M., and Hobbs, P. V. (2006). Preface to the Second Edition. In "Atmospheric Science (Second Edition)" (J. M. Wallace and P. V. Hobbs, eds.), pp. xi-xiii. Academic Press, San Diego.
- Webster, R., and Oliver, M. A. (2007). "Geostatistics for environmental scientists," John Wiley & Sons.
- Zhang, R., Shen, Y., Tang, Z., Li, W., and Zhang, D. (2022). A review of numerical research on the pressure swing adsorption process. *Processes* **10**, 812.
- Blanch, J. (2004). Using Kriging to Bound Satellite Ranging Errors Due to the Ionosphere, Stanford University, California.
- Breiman, L. (2001). Random forests. *Machine learning* **45**, 5-32.
- Cambardella, C. A., Moorman, T. B., Novak, J., Parkin, T. B., Karlen, D., Turco, R., and Konopka, A. E. (1994). Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sci Soc Am J* **58**.
- Carlson, T. (1993). Mid-Latitude Weather Systems. *Transactions of the Institute of British Geographers* **18**.
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning* **20**, 273-297.
- Cristianini, Nello, and Shawe-Taylor, J. (2001). An introduction to support vector machines and other kernel-based learning methods. Repr. *Introduction to Support Vector Machines and other Kernel-Based Learning Methods* **22**.
- Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology* **88**, 2783-2792.
- Egova, E. S. (2015). Integrated water vapour comparison from GNSS and WRF model for Bulgaria in 2013., Sofia University Unpublished master's thesis. Bulgaria.
- Hofmann-Wellenhof, B., Lichtenegger, H., and Collins, J. (2001). "Global Positioning System. Theory and practice".
- Hsu, C.-W., Chang, C.-C., and Lin, C.-J. (2003). A practical guide to support vector classification. Taipei, Taiwan.
- James, G. (2013). An Introduction to Statistical Learning. Springer.
- Kleijer, F. (2004). Troposphere Modeling and Filtering for Precise GPS Leveling.
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R news* **2**, 18-22.
- Lo, J., and El-Mowafy, A. (2011). Interpolation of the GNSS wet troposphere delay. In "Surveying and Spatial Sciences Conference, 2011", pp. 425-438. New Zealand Institute of Surveyors and the Surveying and Spatial Sciences....

Evaluation of Tropospheric Delay Interpolation Methods from Scattered GPS Station Observations

Summary

This study comprehensively evaluates the effectiveness of five interpolation methods—Kriging, Support Vector Machine (SVM), Random Forest (RF), Natural Neighbor (NN), and Artificial Neural Network (ANN)—in estimating Precipitable Water Vapor (PWV) based on GPS data collected from 25 strategically located stations across the diverse geographic region of Los Angeles. The predictors utilized in this study included critical factors such as latitude, longitude, elevation, and tropospheric delay components derived from high-precision GPS observations. The analysis primarily focuses on two representative dates, July 24, 2021 (summer), and January 28, 2022 (winter), specifically chosen for their contrasting meteorological conditions. These dates enable a detailed evaluation of seasonal variability in PWV distribution and provide an opportunity to test the robustness of the selected methods under varying atmospheric conditions.

Tropospheric delay, a key parameter in GNSS-based atmospheric studies, was computed by separating it into its hydrostatic (Zenith Hydrostatic Delay: ZHD) and wet (Zenith Wet Delay: ZWD) components. ZHD was accurately calculated using the well-established Saastamoinen model, which relies on meteorological variables such as surface pressure and station altitude. ZWD was subsequently derived as the difference between ZHD and the Zenith Total Delay (ZTD). The final PWV values were estimated by applying a region-specific coefficient that depends on the weighted mean temperature (T_m). This critical parameter, T_m , was determined using ERA-5 reanalysis data to ensure precise calculations.

The results demonstrate that SVM emerged as the most effective interpolation method, achieving the lowest Root Mean Square Error (RMSE) of 0.6 mm in winter and exhibiting remarkable robustness across diverse spatial and temporal conditions. Kriging, another reliable method, provided accurate results in regions with dense station coverage but encountered difficulties in sparsely populated areas. RF and NN exhibited better performance in winter conditions, benefiting from the reduced atmospheric noise and more stable meteorological conditions during this season. Conversely, ANN, while theoretically capable of modeling complex relationships, was limited in this study by suboptimal network configurations and sensitivity to sparse data distribution. This underscores the importance of careful architectural design and parameter tuning to unlock its full potential.

Seasonal differences in PWV distribution were clearly depicted in the high-resolution maps generated for the selected dates. During summer, PWV values exhibited significant diurnal fluctuations, with peaks in coastal regions during the afternoon due to elevated temperatures and humidity levels. In contrast, the winter maps displayed more stable distributions with lower peak values, reflecting cooler temperatures and reduced atmospheric moisture. These observations highlight the challenges posed by the dynamic summer conditions while emphasizing the critical role of meteorological parameters such as temperature, pressure, and humidity in influencing PWV estimation accuracy.

This study underscores the necessity of selecting appropriate interpolation methods tailored to specific conditions for accurate PWV estimation. SVM demonstrated exceptional capability in handling nonlinear relationships and scattered datasets, making it the most reliable method in this study. Furthermore, while ANN showed room for improvement, its performance could be significantly enhanced with better configurations and deeper architectures specifically tailored for atmospheric complexities. These findings provide valuable insights into GNSS-based atmospheric research and contribute to the advancement of meteorological modeling, weather forecasting, and climate science.

Keywords: Artificial Neural Network, Kriging, Precipitable Water Vapor, Random Forest, Support Vector Machine