

Evaluation of the effect of solar and geomagnetic parameters in spatio-temporal modeling of ionosphere's total electron content using machine learning methods

Nezamzadeh, M. S.¹  | Voosoghi, B.²  | Ghaffari Razin, S. R.³  

1. Department of Geomatics Engineering, Faculty of Geodesy & Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: m.nezamzadeh97@gmail.com

2. Department of Geomatics Engineering, Faculty of Geodesy & Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: vosoghi@kntu.ac.ir

3. **Corresponding Author**, Department of Geoscience Engineering, Faculty of Geoscience Engineering, Arak University of Technology, Arak, Iran. E-mail: mr.ghafari@arakut.ac.ir

(Received: 26 Feb 2022, Revised: 4 Oct 2022, Accepted: 10 Jan 2023, Published online: 14 June 2023)

Summary

The ionosphere is the upper part of the Earth's atmosphere, which is considered to be approximately 70 to 1000 km above the Earth's surface. Ionosphere modeling has been one of the goals of spatial geodesy since 1970. In many ionosphere modeling using satellite measurements such as GPS, total electron content (TEC) are used as observational input data. In recent years, modeling and prediction of the TEC have been considered by researchers with methods that have high speed and accuracy. One of the branches that has been able to show good capabilities in the field of estimation and modeling is machine learning methods (ML). Machine learning includes fuzzy inference systems (FIS), artificial neural networks (ANNs), genetic algorithms (GAs), support vector machines (SVMs), and evolutionary communications (ECs). Since 1993, with the advancement of computer technology, many new and hybrid algorithms, such as the adaptive neuro-fuzzy inference system (ANFIS), have been developed in ML. Another new effective approach in ML is the support vector regression (SVR) method. The SVR is a kernel-based ML method for classification and regression in which the risk of incorrect classification is minimized. The structure of an SVR network has a lot in common with the ANN, and the main difference is practically in the way of the training algorithm. In general, this method is divided into linear and nonlinear modes.

In this paper, the TEC of the ionosphere is modeled and evaluated with ML models. Support vector regression (SVR) and artificial neural network (ANN) methods are used for local TEC modeling. In both models, the latitude and longitude of the GPS stations, day of the year (DOY), hours, AP, KP, DST, and F10.7 are considered an input vectors. Also, the value of VTEC is considered as the output of the models. The main innovation of this paper is in evaluating the effect of different physical parameters on the accuracy of ML models. Using observations of 15 GPS stations in the northwest of Iran from 193 to 228 in 2012, new models are evaluated. Also, the results of the new models are compared with the results of the global ionosphere map (GIM), the IRI2016, and NeQuick empirical models in two internal and one external control station. Statistical indices of root mean square error (RMSE), relative error, dVTEC, and correlation coefficient are used to evaluate the error of the models. Sensitivity analysis of SVR and ANN models to input parameters is performed and the importance of each physical parameter in spatio-temporal modeling of the ionosphere is investigated. The results obtained from this paper show that in both high and low geomagnetic and solar activities, the SVR model in internal control stations has a higher accuracy than other models. But at the external control station, the error of the SVR model is much higher than other models. Determining the parameters of the kernel function using observations at the territory of the studied network is the reason. Also, the sensitivity of SVR and ANN models is increased to the physical parameters F10.7, KP, DST, and AP, respectively. For precise local ionosphere modeling, the effect of these parameters must also be considered.

Keywords: Physical parameters, Ionosphere, Sensitivity analysis, GPS, SVR.

Cite this article: Nezamzadeh, M. S., Voosoghi, B., & Ghaffari Razin, S. R. (2023). Evaluation of the effect of solar and geomagnetic parameters in spatio-temporal modeling of ionosphere's total electron content using machine learning methods. *Journal of the Earth and Space Physics*, 49(1), 153-169. DOI: <http://doi.org/10.22059/jesphys.2023.339441.1007405>

ارزیابی تأثیر پارامترهای خورشیدی و ژئومغناطیسی در مدل‌سازی زمانی-مکانی محتوای الکترون کلی (TEC) یونسفر با استفاده از مدل‌های یادگیری ماشین

مهديه السادات نظام‌زاده^۱ | بهزاد وثوقی^۲ | سیدرضا غفاری رزین^۳ ✉

۱. گروه مهندسی ژئودزی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران. رایانامه: m.nezamzadeh97@gmail.com

۲. گروه مهندسی ژئودزی، دانشکده مهندسی نقشه‌برداری، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران. رایانامه: vosoghi@kntu.ac.ir

۳. نویسنده مسئول، گروه مهندسی نقشه‌برداری، دانشکده مهندسی علوم زمین، دانشگاه صنعتی اراک، ایران. رایانامه: mr.ghafari@arakut.ac.ir

(دریافت: ۱۴۰۰/۱۲/۷، بازنگری: ۱۴۰۱/۷/۱۲، پذیرش نهایی: ۱۴۰۱/۱۰/۲۰، انتشار آنلاین: ۱۴۰۲/۳/۲۴)

چکیده

در این مقاله مقدار محتوای الکترون کلی (TEC) یونسفر با مدل‌های یادگیری ماشین (ML)، به صورت مکانی-زمانی، مدل‌سازی شده و مورد ارزیابی و مقایسه قرار می‌گیرد. روش‌های رگرسیون بردار پشتیبان (SVR) و شبکه عصبی مصنوعی (ANN) جهت مدل‌سازی محلی TEC استفاده می‌شوند. نوآوری اصلی این مقاله در ارزیابی تأثیر پارامترهای فیزیکی مختلف (DST, AP, KP, F10.7) در دقت خروجی مدل‌های یادگیری ماشین است. نتایج به دست آمده از دو مدل جدید با نتایج مدل جهانی یونسفری (GIM)، مدل‌های تجربی IRI2016 و NeQuick در دو ایستگاه کنترل داخلی و یک ایستگاه کنترل خارجی مورد مقایسه قرار گرفته‌اند. شاخص‌های آماری جذر خطای مربعی میانگین (RMSE)، خطای نسبی، $dVTEC = |VTEC_{GPS} - VEC_{model}|$ و ضریب همبستگی برای ارزیابی خطای مدل‌ها، به کار گرفته شده است. ارزیابی تأثیر پارامترهای ژئومغناطیسی و خورشیدی در خروجی مدل‌های SVR و ANN نسبت به پارامترهای ورودی انجام و اهمیت هر کدام از پارامترهای فیزیکی در مدل‌سازی مکانی-زمانی یونسفر مورد بررسی قرار گرفته است. میانگین RMSE محاسبه شده در دو ایستگاه کنترل داخلی برای مدل‌های SVR، ANN، GIM، IRI2016 و NeQuick به ترتیب برابر با ۱/۰۴، ۳/۹۱، ۳/۰۲، ۶/۹۰ و ۷/۶۵ TECU شده است. همچنین میانگین ضریب همبستگی مدل‌ها در دو ایستگاه کنترل داخلی به ترتیب برابر با ۰/۹۷، ۰/۷۲، ۰/۸۴، ۰/۶۸ و ۰/۶۰ محاسبه شده است. نتایج به دست آمده از این مقاله نشان می‌دهد که در هر دو حالت فعالیت‌های ژئومغناطیسی و خورشیدی بالا و پایین، مدل SVR در ایستگاه‌های کنترل داخلی از دقت و صحت بالاتری نسبت به سایر مدل‌ها برخوردار است.

واژه‌های کلیدی: پارامترهای فیزیکی، یونسفر، GPS، SVR.

۱. مقدمه

عوامل یونیزاسیون است که انرژی باد خورشیدی را به صورت گرما به سمت یونسفر می‌کشد (سیر، ۲۰۰۳). این گرما باعث افزایش یونیزاسیون در لایه یونسفر شده و در نتیجه مقدار چگالی الکترونی را افزایش می‌دهد. لایه یونسفر به دلیل چگونگی تغییرات یونیزاسیون بر حسب ارتفاع به لایه‌های مختلفی تقسیم می‌شود. لایه یونسفر دارای تغییرات مکانی و زمانی است و این تغییرات می‌تواند ناشی از تغییرات روزانه، تغییرات فصلی، آنومالی‌های مختلف موجود و دوره فعالیت‌های

یونسفر قسمت بالایی جو زمین است که تقریباً در ارتفاعی بین ۷۰ تا ۱۰۰۰ کیلومتری سطح زمین در نظر گرفته می‌شود. با ورود پرتو ایکس و اشعه فرابنفش و پرتوها و بادهای خورشیدی که از فضای خارج به لایه‌های بالایی جو وارد می‌شوند، یونیزاسیون صورت می‌گیرد که باعث گسستگی پیوند مولکول‌ها و اتم‌های لایه یونسفر می‌شود. خورشید مهم‌ترین عامل یونیزه شدن است و میزان یونیزاسیون لایه یونسفر در درجه اول به فعالیت‌های خورشیدی بستگی دارد. میدان مغناطیسی زمین از دیگر

استناد: نظام‌زاده، مهديه السادات؛ وثوقی، بهزاد و غفاری رزین، سیدرضا (۱۴۰۲). ارزیابی تأثیر پارامترهای خورشیدی و ژئومغناطیسی در مدل‌سازی زمانی-مکانی محتوای الکترون کلی (TEC) یونسفر با استفاده از مدل‌های یادگیری ماشین. مجله فیزیک زمین و فضا، ۴۹(۱)، ۱۵۳-۱۶۹.

DOI: <http://doi.org/10.22059/jesphys.2023.339441.1007405>

از اندازه‌گیری‌های تک فرکانسه در بسیاری از کاربردهای ناوبری و تعیین موقعیت، مدلی دقیق و قابل‌اعتماد از یونسفر مورد نیاز است.

مدل‌سازی یونسفر از سال ۱۹۷۰ یکی از اهداف ژئودزی فضایی بوده است. در بسیاری از روش‌های مدل‌سازی یونسفر که با استفاده از اندازه‌گیری‌های ماهواره‌ای همانند سیستم GPS صورت می‌پذیرد، TEC به‌عنوان داده مشاهداتی در مدل‌سازی یونسفر بکار می‌رود. مدل‌های تجربی یونسفر به‌طور گسترده در زمینه‌های مطالعات آب‌وهوا، علوم رادیویی، ناوبری هوایی و دریایی مورد استفاده هستند. از جمله مدل‌های یونسفری می‌توان از مدل مرجع بین‌المللی یونسفر (IRI) نام برد که یک پروژه علمی دائمی مشترک کمیته تحقیقات فضایی COSPAR و اتحادیه بین‌المللی علوم رادیویی URSI است و از سال ۱۹۶۸ آغاز شده است (بلیتزا و راینیش، ۲۰۰۸). یکی دیگر از مدل‌های تجربی مدل NeQuick می‌باشد که توسط مرکز بین‌المللی فیزیک نظری عبدالسلام ایتالیا و با همکاری موسسه ژئوفیزیک، اخترفیزیک و هواشناسی دانشگاه گراتس اتریش توسعه یافته است. مدل NeQuick یک مدل سریع اجرایی برای چگالی الکترونی یونسفر است که برای کاربردهای انتشار گسترده یونسفری طراحی شده که امکان محاسبه چگالی الکترونی را در هر موقعیت یونسفر می‌دهد (ناوا و همکاران، ۲۰۰۸). این مدل چگالی الکترونی یونسفر را به شکل تابعی از ارتفاع، طول و عرض ژئوستریک و همچنین فعالیت‌های خورشیدی، فصلی و زمان (ساعت جهانی UT) در اختیار می‌گذارد.

با توجه به اینکه تعداد ایستگاه‌های گیرنده‌های GPS در داخل کشور ایران کافی نبوده، همچنین توزیع مکانی این گیرنده‌ها نامناسب می‌باشد و از طرفی ناپیوستگی مشاهداتی در حوزه زمان در مشاهدات این ایستگاه‌ها وجود دارد، محدودیت زمانی-مکانی در بررسی رفتار یونسفر امری اجتناب‌ناپذیر است و TEC به‌دست‌آمده از اندازه‌گیری‌های GPS دارای محدودیت زمانی-مکانی

خورشیدی باشد (اسچانک و ناچی، ۲۰۰۰). به‌طور کلی تغییرات مکانی یونسفر متأثر از میدان مغناطیسی زمین بوده و تغییرات زمانی یونسفر متأثر از فعالیت‌های خورشیدی می‌باشد. سیگنال عبوری از لایه یونسفر به دلیل خاصیت پاشندگی این لایه و تراکم الکترون‌های آزاد در آن تحت تأثیر قرار می‌گیرند که این تأثیر شکست یونسفری نامیده شده و به فرکانس سیگنال عبوری بستگی دارد (براوون و همکاران، ۱۹۹۵). تأثیر یونسفر بر امواج الکترومغناطیس حوزه‌های وسیع و گسترده‌ای را همانند سیستم‌های مشاهداتی فضا مبنای سیستم‌های ارتباطی و هواشناسی فضایی ایجاد کرده است.

یکی از مهم‌ترین سیستم‌های تعیین موقعیت و ناوبری ماهواره‌ای جهانی، سیستم تعیین موقعیت جهانی (GPS) است. امواج ارسال از ماهواره‌های GPS در مسیر انتشار از ماهواره به گیرنده تحت تأثیر خصوصیات فیزیکی لایه یونسفر قرار می‌گیرند (سیبر، ۲۰۰۳؛ هافمن و همکاران، ۲۰۰۷). اثر یونسفر بر سیگنال‌های GPS موجب تأخیر در کد و تقدم در فاز حامل می‌شود که برای محاسبه این تأخیر یونسفری و تغییرات مکانی و زمانی یونسفر نیازمند اطلاع از چگالی الکترونی و کمیت محتوی الکترون کلی (TEC) می‌باشیم. با استفاده از مشاهدات گیرنده‌های دوفرکانسه می‌توان مقدار TEC را در نقاط موسوم به نقاط نفوذ یونسفری محاسبه کرد و به کمک این مشاهدات می‌توان اثر لایه یونسفر در مشاهدات GPS را تعدیل کرد. کمیتی که از این روش محاسبه می‌شود محتوی الکترون کلی در راستای مایل نامیده می‌شود. این کمیت را می‌توان با استفاده از توابع نگاشت به محتوای الکترون کلی در راستای قائم تبدیل کرد. در حقیقت TEC تعداد کل الکترون‌های آزاد در مسیر مابین ماهواره تا گیرنده در ستونی به سطح مقطع یک مترمربع است. واحد اندازه‌گیری TEC، TECU بوده و یک TECU معادل 10^{16} (m²/ele) می‌باشد (سیبر، ۲۰۰۳). این کمیت را می‌توان از روی اندازه‌گیری‌های فاز و کد به‌دست‌آمده از سیستم تعیین موقعیت جهانی به دست آورد. برای استفاده

خواهند بود. بنابراین، لزوم ارائه مدل‌هایی با توانایی غلبه بر این محدودیت‌ها و برخورداری بودن دقت و صحت بالا کاملاً احساس می‌شود (غفاری رزین و وثوقی، ۲۰۱۷). جهت محاسبه و تعیین مقدار TEC در مناطقی که دارای مشاهده نبوده و یا توزیع ایستگاهی مناسب وجود ندارد، در این حالت مقدار TEC به دست آمده از اندازه‌گیری‌های GPS می‌بایستی با یک روش مناسب، درونیابی و یا برون‌یابی شوند. در نتیجه مقادیر TEC محاسبه شده، می‌توانند به عنوان مدلی جهت بررسی رفتار یونسفر در زمان و مکان مورد استفاده قرار گیرند.

تاکنون تلاش‌های بسیار زیادی جهت درونیابی و یا برون‌یابی TEC در مناطقی که مشاهده کافی و یا مناسب موجود نیست، صورت گرفته است. به طور مثال می‌توان به روش‌های کریجینگ، چندجمله‌ای، هارمونیک‌های کروی، هارمونیک‌های کلاه کروی، درونیابی به کمک توابع اسپیلاین، موجک، مدل‌سازی با توابع اسلپین، توابع اتو رگرسیون و اتو کوواریانس، درونیابی با روش المان‌های محدود^۱ C^۱ اشاره کرد (سایین و همکاران، ۲۰۰۸؛ ماتر و همکاران، ۲۰۰۵؛ عامریان و همکاران، ۲۰۱۳؛ شریفی و فرزانه، ۲۰۱۵؛ اعتمادفرد و حسینی، ۲۰۱۶؛ موتاروف و همکاران، ۲۰۰۲؛ نعمتی پور و همکاران، ۲۰۲۱). از این روش‌ها به کرات جهت تهیه و مدل‌سازی منطقه‌ای و جهانی مقدار TEC استفاده شده است. در سالان اخیر مدل‌سازی و پیش‌بینی کمیت TEC توسط محققان و با روش‌هایی که دارای سرعت عمل و دقت بالایی باشند مورد توجه قرار گرفته است. یکی از شاخه‌هایی که توانسته در زمینه تخمین و مدل‌سازی قابلیت‌های خوبی از خود نشان دهد روش‌های یادگیری ماشین (ML) هستند.

یادگیری ماشین شامل سیستم استنتاج فازی (FIS)، شبکه‌های عصبی مصنوعی (ANN)، الگوریتم ژنتیک (GA)، ماشین‌های بردار پشتیبان (SVMs) و ارتباطات تکاملی (EC) است (جانگ و توپال، ۲۰۱۴). از سال ۱۹۹۳، با پیشرفت فناوری رایانه، بسیاری از الگوریتم‌های

جدید و ترکیبی، مانند سیستم استنتاج عصبی فازی تطبیقی (ANFIS) در ML توسعه یافته‌اند (جنگ، ۱۹۹۳). در دهه‌های اخیر از روش‌های یادگیری ماشین در مدل‌سازی یونسفر و تروپوسفر استفاده شده است (کاندر، ۱۹۹۸؛ هابارولما و همکاران، ۲۰۱۱؛ هوانگ و همکاران، ۲۰۱۵؛ غفاری رزین و وثوقی، ۲۰۱۶a، ۲۰۱۷، ۲۰۱۸؛ اینیورت و سکرتهین، ۲۰۱۹؛ تبال و همکاران، ۲۰۱۹؛ فیضی و همکاران، ۲۰۲۰؛ ژیا و همکاران، ۲۰۲۱). یکی دیگر از رویکردهای جدید مؤثر در ML روش رگرسیون بردار پشتیبان (SVR) است (واپنیک، ۱۹۹۵). روش رگرسیون بردار پشتیبان یک روش یادگیری ماشین مبتنی بر هسته برای طبقه‌بندی و رگرسیون است که در آن ریسک عدم طبقه‌بندی صحیح را کمینه می‌کنند. ساختار یک شبکه SVR، اشتراکات زیادی با شبکه عصبی پرسپترون چند لایه دارد و تفاوت اصلی آن عملاً در شیوه یادگیری است. به طور کلی این روش به دو حالت خطی و غیرخطی تقسیم می‌شود (یگانه و همکاران، ۲۰۱۲).

هدف این مقاله استفاده از مدل‌های SVR و ANN برای مدل‌سازی تغییرات مکانی-زمانی محتوای الکترون کلی یونسفر به صورت به صورت محلی در ایران است. نوآوری اصلی این مقاله در آنالیز تأثیر پارامترهای فیزیکی مختلف در نتایج مدل‌های یادگیری ماشین است که از نقاط قوت مقاله محسوب می‌شود. ارزیابی هر دو مدل در شرایط مختلف مغناطیسی و خورشیدی انجام گرفته و نتایج در ایستگاه‌های کنترل با مدل‌های GIM، NeQuick و IRI2016 مقایسه می‌شوند. شاخص‌های آماری مختلفی نیز مورد بررسی و آنالیز قرار گرفته‌اند.

۲. روش‌شناسی

با توجه به اینکه خروجی مدل‌های SVR و ANN مقدار VTEC حاصل از ایستگاه‌های GPS در نظر گرفته شده است، در نتیجه در این بخش، الگوریتم استخراج مشاهدات VTEC از اندازه‌گیری‌های ایستگاه‌های GPS

برای مشاهده کد و فاز میان گیرنده u و ماهواره m به ترتیب به صورت روابط (۱) و (۲) تشکیل می شود (کومجانی، ۱۹۹۷؛ لیک و همکاران، ۲۰۱۵):

$$P_{4,u}^m = A \left(\frac{f_1^2 - f_2^2}{f_1^2 f_2^2} \right) STEC_u^m - c(DCB^m + DCB_u) \quad (۱)$$

$$L_{4,u}^m = A \left(\frac{f_1^2 - f_2^2}{f_1^2 f_2^2} \right) STEC_u^m - c(DCB^m + DCB_u) + \Delta N^m \quad (۲)$$

در این روابط مقدار A برابر با $۴۰/۳$ ، f_1 و f_2 به ترتیب فرکانس موج حامل L_1 و L_2 سیستم تعیین موقعیت جهانی و $STEC_u^m$ مقدار محتوی کلی الکترونی مایل میان گیرنده i و ماهواره z و C سرعت نور در خلأ می باشد. DCB_u و DCB^m نیز به ترتیب ارباب تفاضلی کد ماهواره m و گیرنده زمینی u می باشند. ΔN^m اختلاف ابهام فاز موج حامل L_1 و L_2 است. برای یک زمان مشخص، محتوی الکترونی کلی مایل ($STEC$) با استفاده از داده های شبه فاصله و فاز حامل از هر ماهواره قابل محاسبه است. کمیت $STEC$ محاسبه شده از رابطه (۱) دارای اثر چند مسیری و نویز است؛ بنابراین خواهیم داشت:

$$STEC_u^m(n) = \frac{1}{A} \left(\frac{f_1^2 f_2^2}{f_1^2 - f_2^2} \right) [P_{4,u}^m(n) + c(DCB^m + DCB_u)] \quad (۳)$$

که در این رابطه n یک زمان مشخص در بازه $1 \leq n \leq N$ قرار دارد. برای محاسبه $STEC$ از رابطه (۲) ابهام فاز اولیه آن باید حل شود. در نتیجه خواهیم داشت:

$$B = \frac{1}{N_{me}} \sum_{n_{me}} (P_{4,u}^m(n_{me}) - L_{4,u}^m(n_{me})) \quad (۴)$$

در رابطه (۴) B مقداری است که با استفاده از اندازه گیری های فاز موج حامل در کمان های پیوسته محاسبه می شود. سپس می توان $STEC$ را با قرار دادن B در رابطه (۲) به صورت به صورت زیر محاسبه کرد:

$$STEC_u^m(n) = \frac{1}{A} \left(\frac{f_1^2 f_2^2}{f_1^2 - f_2^2} \right) [B + L_{4,u}^m(n) + c(DCB^m + DCB_u)] \quad (۵)$$

ارائه می شود. همچنین در ادامه، مدل های رگرسیون بردار پشتیبان و شبکه عصبی با جزئیات کامل توضیح داده می شوند.

۱-۲. محاسبه مشاهدات VTEC از اندازه گیری های GPS

شبه فاصله کد و فاز موج حامل مشاهدات پایه GPS می باشند. شبه فاصله موجود بین ماهواره و گیرنده از طریق مشاهدات کد P و C/A محاسبه می شود. برای همین منظور کدی مشابه کد تولید شده توسط ماهواره در گیرنده تولید شده و با تطبیق این دو کد با یکدیگر مدت زمان سیر سیگنال از ماهواره تا گیرنده محاسبه و با ضرب این زمان در سرعت نور شبه فاصله ماهواره تا گیرنده به دست می آید. از آنجاکه فاصله تعیین شده از این طریق آلوده به خطاهای مختلف بوده، به آن شبه فاصله اطلاق می شود و متفاوت از فاصله هندسی ماهواره تا گیرنده است. اندازه گیری فاز موج حامل نیز اساساً بر مبنای اندازه گیری اختلاف بین فاز موج حامل دریافت شده از ماهواره و فاز موج حامل تولید شده در نوسان ساز گیرنده استوار است. در حالت کلی طول موج حامل کوتاه تر از طول موج کد بوده و با دقت یک صدم دور قابل اندازه گیری است اما امکان استخراج مدت زمان سیر سیگنال از ماهواره تا گیرنده با استفاده از موج حامل وجود ندارد و گیرنده GPS فقط فاز موج حامل و تغییرات بعدی آن را اندازه گیری می کند. تعداد دورهای کامل طی شده توسط موج حامل ابهام فاز نام داشته که مجهول بوده و به روش های مختلف در پردازش اطلاعات GPS برآورد می شود. از طریق ضرب فاز موج حامل بر حسب دور در طول موج مربوطه، معادله مشاهده فاز موج حامل بر حسب واحد طول برای دو موج حامل L_1 و L_2 حاصل می شود. گیرنده های GPS سیگنال هایی را با دو فرکانس باند L ($f_1=1575.42$ MHz, $f_2=1227.6$ MHz) منتشر می کنند. با استفاده از مشاهدات کد و فاز GPS، ترکیب خطی مستقل از هندسه کد و ترکیب خطی مستقل از هندسه فاز

با هر جهش فاز رابطه (۵) محاسبات پایه دیگری را شروع می‌کند. پس از محاسبه STEC، VTEC را می‌توان با استفاده از تقریب پوسته نازک مدل یونسفر تک لایه به دست آورد:

$$VTEC_u^m(n) = \frac{STEC_u^m(n)}{M(\varepsilon_m(n))} \quad (6)$$

که در (۶) تابع نگاشت M به صورت زیر است:

$$M(\varepsilon_m(n)) = \left[1 - \left(\frac{R \cos \varepsilon_m(n)}{R+h} \right)^2 \right]^{\frac{1}{2}} \quad (7)$$

در روابط (۶) و (۷) M تابع تصویر، ε زاویه ارتفاعی ماهواره، R شعاع کره زمین و h ارتفاع مدل تک لایه یونسفری (۴۵۰ کیلومتر) است. بایاس داخل فرکانسی مربوط به ماهواره معمولاً از فایل IONEX محاسبه می‌شود که از محصولات IGS است. زاویه ارتفاعی نیز با استفاده از فاصله هندسی ماهواره و گیرنده و با داده‌های افریز ماهواره محاسبه می‌شود. افریزها اطلاعات مداری هستند که به صورت به صورت دقیق و غیردقیق برای ماهواره‌های سامانه‌های تعیین موقعیت ماهواره‌ای همانند GPS توسط شبکه جهانی IGS ارائه می‌شوند. معمولاً اطلاعات افریزهای دقیق هر دو ساعت یکبار برای هر ماهواره به روزرسانی می‌شوند در صورتی که در حالت غیردقیق، این به روزرسانی ممکن است ۲ تا ۱۴ روز به طول بیانجامد. در داخل فایل‌های افریز اطلاعات مداری ماهواره‌ها شامل المان‌های مداری کپلری، اطلاعات دریافت ساعت ماهواره و پارامترهای اغتشاشات مدار ماهواره ارائه می‌شود.

۲-۲. روش رگرسیون بردار پشتیبان

ماشین بردار پشتیبان (SVM) یک نوع سیستم یادگیری نظارت شده است که برای مسائل طبقه بندی بکار می‌رود به طوری که کمترین خطا در طبقه بندی داده‌ها رخ دهد (کورتیس و واپنیک، ۱۹۹۵). این روش بر مبنای تئوری یادگیری آماری استوار است که از اصل کمینه سازی خطای ساختاری بهره می‌جوید و موجب یک جواب بهینه

کلی می‌شود. این الگوریتم بعدها برای کار با مسائل رگرسیون یا تخمین داده‌ها توسعه یافت. الگوریتم جدید رگرسیون بردار پشتیبان (SVR) نامیده شد (اسمولا و اسچولکوپ، ۱۹۹۸). هدف رگرسیون بردار پشتیبان که مدل رگرسیونی SVM می‌باشد این است که تابع f را برای الگوی آموزشی x طوری تشخیص دهد که بیشینه حاشیه را از مقادیر آموزشی y داشته باشد. به عبارت دیگر مدل SVR مدلی است که منحنی با ضخامت ε را به داده‌ها برازش می‌دهد به نحوی که کمترین خطا در داده‌های آزمون صورت گیرد. با فرض آنکه l داده آموزشی وجود داشته باشد، اگر هر ورودی x دارای تعداد D ویژگی بوده و هر نقطه دارای مقدار y نظیر ویژه‌ای باشد، هدف پیدا کردن تابعی است که بین ورودی و خروجی رابطه زیر را برقرار کند:

$$f(x, w) = w^T x + b \quad (8)$$

برای به دست آوردن تابع f محاسبه مقادیر w و b ضروری است. برای محاسبه مقادیر w و b رابطه (۹) باید کمینه شود (کورتیس و واپنیک، ۱۹۹۵):

$$R(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{l} \sum_{i=1}^l L_\varepsilon(y_i, f_i(x, w)) \quad (9)$$

که در آن بیانگر فاکتور جریمه و ε سطح خطای مجاز برای تابع می‌باشند. هرچه C بزرگ تر باشد جریمه بیشتری به تابع اعمال می‌شود. تابع L_ε تابع وپنیک است که به صورت زیر تعریف می‌شود:

$$|y - f(x, w)|_\varepsilon = \begin{cases} 0 & \text{if } |y - f(x, w)| \leq \varepsilon \\ |y - f(x, w)| - \varepsilon & \text{otherwise} \end{cases} \quad (10)$$

مسئله بالا به صورت بیشینه شدن رابطه (۱۱) بازنویسی می‌شود:

$$L_p(\alpha_i, \alpha_i^*) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i \quad (11)$$

یک لایه خروجی با n پارامتر است. مدل مورد استفاده جهت محاسبه پارامترهای لایه خروجی به صورت رابطه زیر است (هایکین، ۱۹۹۴):

$$y = f\left(\sum_{j=1}^q W_j f\left(\sum_{l=1}^N w_{j,l} x_l + w_{j,0}\right) + W_0\right) \quad (15)$$

در رابطه (۱۵) وزن مابین لایه نوروں لایه پنهان و خروجی، $w_{j,l}$ وزن مابین لایه نوروں ورودی و لایه نوروں پنهان، x_l بیانگر لایه نوروں ورودی، $w_{j,0}$ بایاس مربوط به نوروں‌های مابین لایه‌های ورودی و پنهان و W_0 بایاس مربوط به نوروں‌های لایه‌های پنهان و خروجی هستند. جهت ارتباط میان لایه پنهان و خروجی به یک تابع فعال‌سازی مناسب نیاز است. یکی از مناسب‌ترین توابع فعال‌سازی تابع سیگموئید است که به صورت تابع ریاضی زیر تعریف می‌شود:

$$f(z) = \frac{1}{(1 + e^{-z})} \quad (16)$$

که در رابطه (۱۶) z اطلاعات ورودی نوروں و $f(z) \in [0,1]$ هستند. جهت آموزش شبکه عصبی، الگوریتم‌های آموزش متفاوتی وجود دارد. در این مقاله، از الگوریتم آموزش پس انتشار خطا (BP) استفاده می‌شود (مارس و همکاران، ۱۹۹۶). در این الگوریتم خروجی شبکه عصبی با خروجی مطلوب مقایسه شده و بر اساس تابع هدف، مقدار خطا محاسبه می‌شود. اگر مقدار خطا در حد دقت مطلوب نباشد، بر اساس مشتق تابع هدف، وزن‌های مربوط به نوروں‌ها تصحیح شده و مجدداً خروجی محاسبه می‌شود.

۲-۴. شاخص‌های آماری مورد استفاده

ارزیابی خطای مدل‌های SVR و ANN با استفاده از شاخص‌های آماری جذر خطای مربعی میانگین (RMSE)، خطای نسبی، $dVTEC = |VTEC_{GPS} - VTEC_{model}|$ ، ضریب همبستگی و هیستوگرام باقیمانده‌ها انجام می‌گیرد. این شاخص‌ها به صورت به صورت زیر تعریف می‌شوند:

که شرایط آن به صورت رابطه (۱۲) است:

$$\begin{cases} \sum_{i=1}^I (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C \\ 0 \leq \alpha_i^* \leq C \end{cases} \quad (12)$$

با حل معادله بالا می‌توان تابع SVR یعنی f را با استفاده از تابع کرنل به صورت زیر محاسبه کرد:

$$f(x, w) = w_0^T x + b = \sum_{i=1}^I (\alpha_i - \alpha_i^*) x_i^T \quad (13)$$

برای ساخت بهترین مدل SVR باید پارامترهای تأثیرگذار آن بهینه شود. به منظور بهینه‌سازی پارامترهای مدل می‌توان از روش‌های مختلفی استفاده کرد که در این مقاله از تجزیه کننده الگوریتم حداقل بهینه‌سازی متوالی با تابع کرنل شعاعی (گوسین) استفاده شد. این تابع کرنل برای دو متغیر a و b مطابق رابطه (۱۴) تعریف می‌شود:

$$K(a, b) = \exp\left(-\gamma \|a - b\|^2\right) = \exp\left(-\frac{\|a - b\|^2}{2\sigma^2}\right) \quad (14)$$

که در این رابطه پارامتر γ با عکس مجذور عرض (پهنای) کرنل گوسی (σ) متناسب است. توابع کرنل دیگر نیز به صورت سعی و خطا آزمایش شدند؛ ولی بهترین نتیجه، با استفاده از تابع پایه شعاعی به دست آمد.

۲-۳. روش شبکه عصبی مصنوعی

شبکه‌های عصبی مصنوعی (ANNs) و یا به عبارت دیگر سیستم‌های اتصال‌گر، ابزارهای محاسباتی هستند که منشأ الهام آن‌ها سیستم‌های بیولوژیکی و سیستم‌های طبیعی است. شبکه عصبی مصنوعی با تشبیه به شبکه‌های عصبی زیستی، مبتنی بر واحدهای متصل و غیرخطی به نام نوروں هستند که کوچک‌ترین واحد پردازشگر اطلاعات بوده و این اجزا تعیین‌کننده رفتار شبکه هستند (سیمپسون، ۱۹۹۰). شبکه عصبی پرسپترون سه لایه مورد استفاده دارای یک لایه ورودی با N مشاهده، یک لایه پنهان و

کنترل خارج از محدوده شبکه مورد مطالعه انتخاب شده‌اند. مشاهدات ایستگاه‌های کنترل در مرحله آموزش مورد استفاده قرار نگرفته‌اند. با استفاده از الگوریتم اشاره شده در بخش دوم مقدار VTEC در بازه‌های زمانی دو ساعته (۰، ۲، ۴، ...، ۲۲) و در موقعیت ایستگاه‌های آموزش (۱۳ ایستگاه) محاسبه شده و به عنوان خروجی، به مدل‌های SVR و ANN معرفی می‌شوند. بردار ورودی مدل‌های SVR و ANN شامل پارامترهای طول و عرض جغرافیایی ایستگاه GPS، روز از سال (DOY)، ساعت، شاخص‌های ژئومغناطیسی AP، KP و DST، شاخص فعالیت‌های خورشیدی F10.7 می‌باشد. به عبارت دیگر عبارت دیگر بردار ورودی شامل ۸ پارامتر و خروجی شامل VTEC می‌باشد. رابطه ریاضی زیر را می‌توان برای مدل‌های SVR و ANN در نظر گرفت:

$$VTEC_{SVR \text{ ANN}} = f(Lat., Lon., DOY, time, KP, AP, DST, F10.7) \quad (21)$$

پارامتر F10.7 با واحد sfu نشان‌دهنده شار خورشیدی است که در حال حاضر یکی از بهترین شاخص‌های فعالیت خورشیدی بوده و از ۵۰ سال گذشته در حال اندازه‌گیری می‌باشد. پارامترهای Kp و AP بیانگر شاخص فعالیت‌های ژئومغناطیسی هستند. پارامتر Kp با رزولوشن زمانی ۳ ساعت، جهت پیش‌بینی فعالیت‌های میدان مغناطیسی زمین در دوره سه ساعته مورد استفاده قرار می‌گیرد. این پارامتر می‌تواند هر مقداری بین ۰ تا ۹ باشد. مقدار کمتر از ۴ بیانگر فعالیت ژئومغناطیسی آرام و مقدار ۹ نشان‌دهنده فعالیت ژئومغناطیسی بالا است. ضریب AP هم دارای بازه ۰ تا ۴۰۰ می‌باشد که میانگین روزانه فعالیت‌های ژئومغناطیسی را نشان می‌دهد. پارامتر Dst نیز با رزولوشن زمانی ۱ ساعت داده‌های زمان اختلال طوفان است. روزهای دارای فعالیت ژئومغناطیسی و خورشیدی بالا، روزهایی است که در آن، پارامترهای $F10.7 \geq 100$, $AP \geq +20$, $KP \geq 4$, $Dst \leq -20$ باشند. شکل (۲) چگونگی تغییرات پارامترهای KP،

$$RMSE = \sqrt{\frac{1}{Q} \sum_{q=1}^Q (VTEC_{model}^q - VTEC_{GPS}^q)^2} \quad (17)$$

$$Re.(%) = \frac{|VTEC_{model} - VTEC_{GPS}|}{VTEC_{GPS}} \times 100 \quad (18)$$

$$dVTEC = |VTEC_{GPS} - VTEC_{model}| \quad (19)$$

$$R = \frac{\sum_{i=1}^N (VTEC_{model}^i - \overline{VTEC}_i) (VTEC_{GPS}^i - \overline{VTEC}_{GPS}^i)}{\sqrt{\sum_{i=1}^N (VTEC_{model}^i - \overline{VTEC}_i)^2} \sqrt{\sum_{i=1}^N (VTEC_{GPS}^i - \overline{VTEC}_{GPS}^i)^2}} \quad (20)$$

در روابط بالا، Q تعداد نمونه‌های انتخاب شده برای تست و به دست آوردن خطاها، $VTEC_{GPS}$ بیانگر مقدار حاصل از مشاهدات GPS و $VTEC_{model}$ نشان‌دهنده مقدار حاصل از مدل‌های مختلف می‌باشند.

۳. نتایج عددی

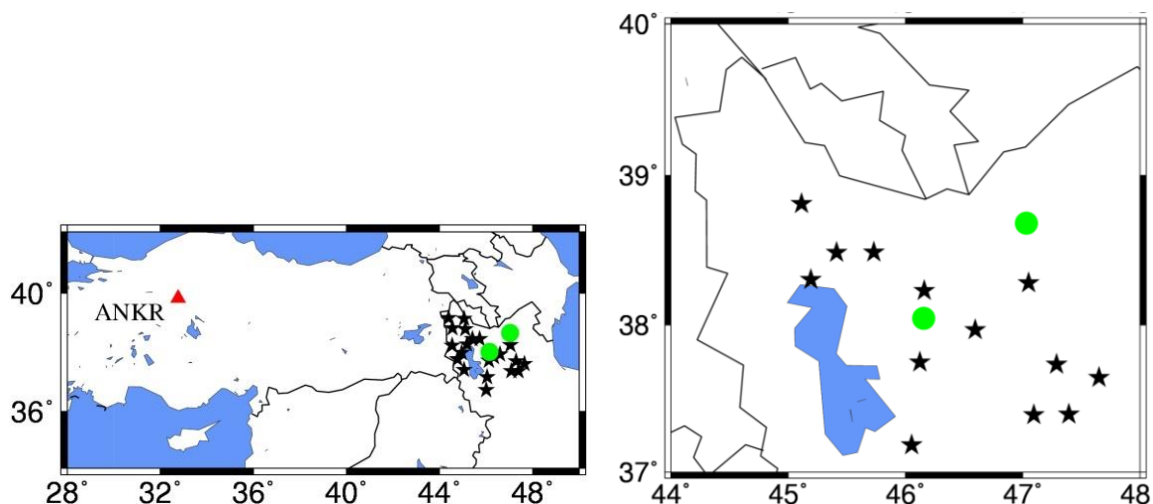
در این بخش، مشاهدات مورد استفاده، نتایج عددی حاصل از مقاله و آنالیزهای آماری انجام گرفته جهت ارزیابی دقت و صحت مدل‌ها، ارائه می‌شوند.

۳-۱. منطقه مورد مطالعه، مشاهدات مورد استفاده

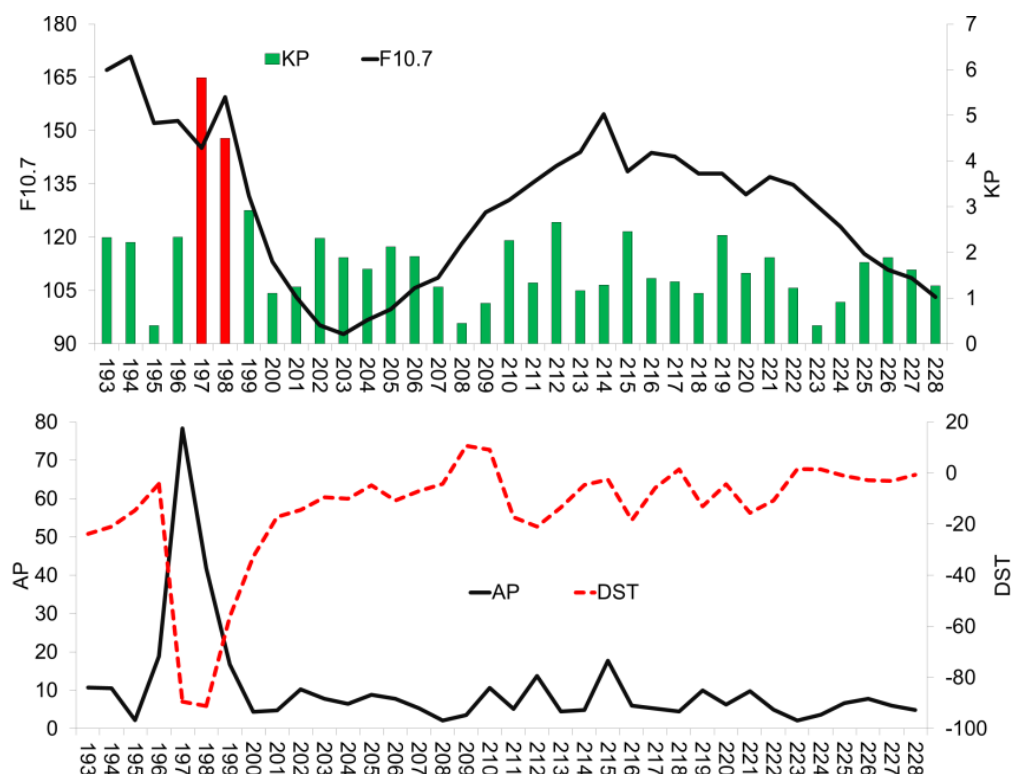
جهت ارزیابی مدل‌های رگرسیون بردار پشتیبان (SVR) و شبکه عصبی مصنوعی (ANN) در مدل‌سازی و پیش‌بینی مکانی-زمانی محتوای الکترون کلی یونسفر و همچنین ارزیابی تأثیر پارامترهای فیزیکی مختلف در مدل‌سازی TEC، از مشاهدات ۱۵ ایستگاه GPS شبکه محلی شمال غرب ایران در روزهای ۱۹۳ الی ۲۲۸ از سال ۲۰۱۲ استفاده می‌شود. کلیه فایل‌های راینکس مربوط به ایستگاه‌های GPS از سازمان نقشه‌برداری کشور (NCC) تهیه شده است. شکل (۱) چگونگی پراکندگی ایستگاه‌های مورد استفاده در این تحقیق را نمایش می‌دهد. بایستی اشاره شود که از ۱۵ ایستگاه مورد استفاده در این تحقیق، دو ایستگاه KLBR (۳۸/۶۸، ۴۷/۰۳) و TABZ (۴۶/۱۵، ۳۸/۰۴) به عنوان به عنوان ایستگاه‌های کنترل داخلی و ایستگاه ANKR (۳۲/۶۴، ۴۰/۱۵) به عنوان ایستگاه

از ۴ باشند (روزهای ۱۹۷ و ۱۹۸). همچنین بر اساس شکل (۲) به جز در روزهای ۲۰۲، ۲۰۳، ۲۰۴ و ۲۰۵، شاخص فعالیت خورشیدی بالای ۱۰۰ است.

DST و AP را در بازه زمانی روزهای مورد استفاده در این تحقیق را نشان می دهد. ستون های قرمز رنگ روزهایی را نمایش می دهد که مقدار KP بزرگ تر



شکل ۱. چگونگی پراکندگی ایستگاه های GPS مورد استفاده در این تحقیق، ستاره های سیاه ایستگاه های مورد استفاده در آموزش مدل های SVR و ANN، دایره های سبز رنگ ایستگاه های کنترل داخلی و مثلث قرمز رنگ ایستگاه کنترل خارج از محدوده شبکه است.



شکل ۲. چگونگی تغییرات پارامترهای KP (باکس های سبز و قرمز)، F10.7، DST و AP در روزهای مورد مطالعه، داده های مربوط به پارامترهای ژئومغناطیسی و خورشیدی از وبسایت <https://omniweb.gsfc.nasa.gov/form/dx1> اخذ شده است.

نتایج حاصل از مرحله آموزش را برای دو مدل SVR و ANN نمایش می‌دهد. مقایسه نتایج هر دو مدل در مرحله آموزش با VTEC حاصل از GPS است. همچنین کلیه پردازش‌ها برای دو مدل در یک سیستم کامپیوتری با سخت‌افزار و نرم‌افزار یکسان انجام گرفته است.

بر اساس نتایج حاصل از جدول (۱)، مقدار RMSE مرحله آموزش برای مدل SVR کمتر از مدل ANN است. به عبارت دیگر مدل SVR با دقت بالاتری، مدل‌سازی مکانی-زمانی VTEC را در مرحله آموزش انجام داده است. همچنین ضریب همبستگی مدل SVR از مدل ANN بالاتر است؛ اما زمان همگرایی به جواب بهینه در مدل ANN کمتر از مدل SVR است؛ یعنی مدل ANN سریع‌تر از مدل SVR به جواب بهینه همگرا شده است. دلیل این امر را می‌توان پیچیدگی محاسباتی در مرحله انتخاب کرنل بهینه و محاسبات آن برای مدل SVR دانست. در مدل ANN از تابع فعال‌سازی سیگموئید در لایه پنهان و تابع خطی در لایه خروجی استفاده شده است. همین امر سرعت محاسبات را در مدل ANN افزایش داده است. پس از مرحله آموزش دو مدل SVR و ANN، با استفاده از مدل‌های آموزش‌دیده مقدار VTEC در ایستگاه‌های کنترل داخلی و ایستگاه کنترل خارجی برآورد شده و با مقدار VTEC حاصل از GPS و مدل‌های GIM، IRI2016 و NeQuick مقایسه و ارزیابی می‌شوند. در مرحله آزمون خطای مدل‌ها نیز، پارامترهای آماری مختلفی مقایسه و ارزیابی می‌شوند. جدول (۲) مقایسه پارامترهای آماری مختلف خطای مدل‌ها را نمایش می‌دهد. بایستی اشاره شود که میانگین نتایج حاصل از ۳۶ روز در جدول (۲) ارائه شده است. همچنین VTEC حاصل از GPS به عنوان مشاهده مرجع در محاسبه پارامترهای آماری در نظر گرفته شده است.

پایگاه داده مورد استفاده جهت آموزش مدل‌های SVR و ANN شامل $12 \times 13 \times 36$ مقدار است (۳۶ بیانگر تعداد روزهای مورد استفاده، ۱۳ تعداد ایستگاه‌های آموزش و ۱۲ تعداد مقادیر VTEC در هر روز). همچنین جهت مقایسه نتایج حاصل از مدل‌های SVR و ANN، مقادیر پیش‌بینی شده VTEC در ایستگاه‌های کنترل داخلی و ایستگاه کنترل خارجی با مقادیر VTEC حاصل از GPS، VTEC حاصل از شبکه جهانی IGS (GIM)، مدل مرجع بین‌المللی یونسفر ۲۰۱۶ (IRI2016) و مدل تجربی NeQuick مقایسه می‌شوند.

۲-۳. آنالیز خطای مدل‌های SVR، ANN، GIM،

IRI2016 و NeQuick

در مرحله آموزش با استفاده از پایگاه داده مشکل از ۸ پارامتر ورودی و خروجی متناظر VTEC، مدل‌های SVR و ANN آموزش داده شده و مدل بهینه حاصل می‌شود. بایستی اشاره شود که در مرحله آموزش پایگاه داده یکسانی برای هر دو مدل استفاده شده است. این پایگاه داده دارای ۵۶۱۶ سطر با ۹ ستون است که هشت ستون آن مربوط به پارامترهای ورودی و ستون آخر مربوط به خروجی متناظر VTEC است. معیار انتخاب مدل بهینه، کمترین مقدار RMSE برای مرحله آموزش است. البته در مرحله آموزش، زمان همگرایی به جواب بهینه و ضریب همبستگی نیز مورد ارزیابی و مقایسه قرار گرفته است. ضریب همبستگی، نشان‌دهنده وجود ارتباط مابین دو متغیر مورد بررسی (در اینجا VTEC حاصل از مدل و VTEC حاصل از GPS به عنوان مبنا) است. هر چقدر مقدار این کمیت به یک نزدیک‌تر باشد، ارتباط مابین دو متغیر بالا است. به عبارت دیگر به عبارت دیگر در تحقیق حاضر، ضریب همبستگی بالا بیانگر این موضوع است که VTEC حاصل از مدل، تغییرات موجود در VTEC حاصل از GPS را با دقت بالایی شناسایی کرده است. جدول (۱)

جدول ۱. مقایسه پارامترهای زمان همگرایی به جواب بهینه (ثانیه)، RMSE (TECU) و ضریب همبستگی برای مرحله آموزش مدل های SVR و ANN.

ضریب همبستگی (R)	RMSE (TECU)	زمان همگرایی به جواب بهینه (ثانیه)	
۰/۹۶	۰/۸۹	۱۱۲	SVR
۰/۹۵	۱/۰۳	۹۶	ANN

جدول ۲. مقایسه پارامترهای آماری ضریب همبستگی (R)، RMSE (TECU)، خطای نسبی (%) در ایستگاه های کنترل داخلی و خارجی برای مدل های SVR، ANN، GIM، IRI2016 و NeQuick.

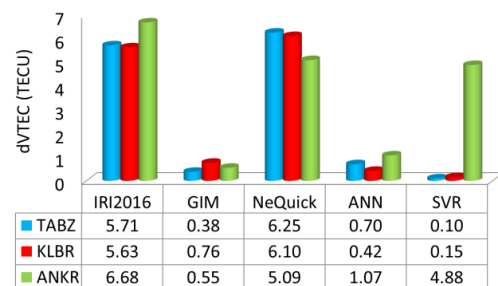
NeQuick	IRI2016	GIM	SVR	ANN		
۰/۶۰	۰/۶۹	۰/۸۴	۰/۹۷	۰/۷۲	ضریب همبستگی (R)	ایستگاه کنترل داخلی TABZ
۷/۷۳	۶/۹۲	۲/۹۲	۱/۰۶	۳/۹۱	RMSE (TECU)	
۳۱/۳۹	۲۶/۴۹	۹/۰۵	۳/۶۴	۱۵/۷۱	خطای نسبی (%)	
۰/۶۰	۰/۶۹	۰/۸۴	۰/۹۷	۰/۷۲	ضریب همبستگی (R)	ایستگاه کنترل داخلی KLBR
۷/۵۷	۶/۸۸	۳/۱۲	۱/۰۲	۳/۹۲	RMSE (TECU)	
۳۱/۵۲	۲۶/۶۳	۹/۹۱	۳/۵۷	۱۶/۲۵	خطای نسبی (%)	
۰/۶۳	۰/۶۶	۰/۸۲	۰/۵۵	۰/۶۳	ضریب همبستگی (R)	ایستگاه کنترل خارجی ANKR
۷/۲۹	۶/۳۱	۳/۰۶	۶/۰۱	۴/۱۱	RMSE (TECU)	
۴۴/۶۰	۲۹/۶۰	۹/۵۵	۲۶/۲۷	۱۷/۶۴	خطای نسبی (%)	

ایستگاه های GPS استفاده می کند، در نتیجه دقت TEC حاصل از این مدل بالاتر از مدل های تجربی است؛ اما در ایستگاه کنترل خارجی و بر اساس نتایج حاصل از جدول (۲)، دقت و صحت مدل GIM نسبت به سایر مدل ها بالاتر است. در ایستگاه کنترل خارجی ANKR، مدل SVR نسبت به مدل ANN خطای بیشتری داشته است. به عبارت دیگر این مدل در محدوده شبکه مورد مطالعه دارای دقت بالایی است، اما خارج از محدوده، دقت این مدل به شدت کاهش پیدا می کند. دلیل این امر را می توان در تابع کرنل مورد استفاده و همچنین عدم استفاده از مشاهدات ایستگاه کنترل خارجی در مرحله آموزش مدل دانست. پارامترهای تابع کرنل بر اساس مشاهدات ورودی به مدل SVR تعیین می شوند. در نتیجه این پارامترها برای مناطقی که خارج از محدوده مورد مطالعه است دارای

مطابق با نتایج حاصل از آنالیز انجام گرفته در جدول (۲)، در هر سه پارامتر ضریب همبستگی، RMSE و خطای نسبی، مدل SVR در دو ایستگاه کنترل داخلی TABZ و KLBR از دقت و صحت بالاتری نسبت به سایر مدل های مورد بررسی برخوردار است. مدل های تجربی IRI2016 و NeQuick در دو ایستگاه کنترل داخلی دارای بیشترین خطا هستند. عمده ترین منابع مشاهداتی برای این دو مدل، از مشاهدات ایستگاه های اندازه گیری مستقیم یونسفر است. با توجه به اینکه در منطقه ایران تنها یک ایستگاه غیرفعال وجود دارد، در نتیجه خطای این دو مدل برای یونسفر در منطقه ایران افزایش می یابد. مدل GIM نسبت به مدل های تجربی IRI2016 و NeQuick در منطقه مورد مطالعه از دقت و صحت بالاتری برخوردار است. با توجه به اینکه این مدل مستقیماً از اندازه گیری های

اعتبار نخواهند بود.

نوع دیگر ارزیابی انجام گرفته برای مدل‌های SVR و ANN، استفاده از شاخص dVTEC مطابق با رابطه (۲۰) در هر سه ایستگاه کنترل است. مقدار این شاخص به‌صورت به‌صورت ساعتی و برای تمامی روزهای مورد مطالعه محاسبه شده و میانگین آن در سه ایستگاه کنترل محاسبه شده است. شکل (۳) نتایج این ارزیابی را نمایش می‌دهد.



شکل ۳. مقادیر شاخص dVTEC مدل‌های ANN، SVR، GIM، IRI2016 و NeQuick محاسبه شده برای سه ایستگاه کنترل TABZ، KLBR و ANKR به‌صورت به‌صورت میانگین برای تمامی روزهای مورد مطالعه.

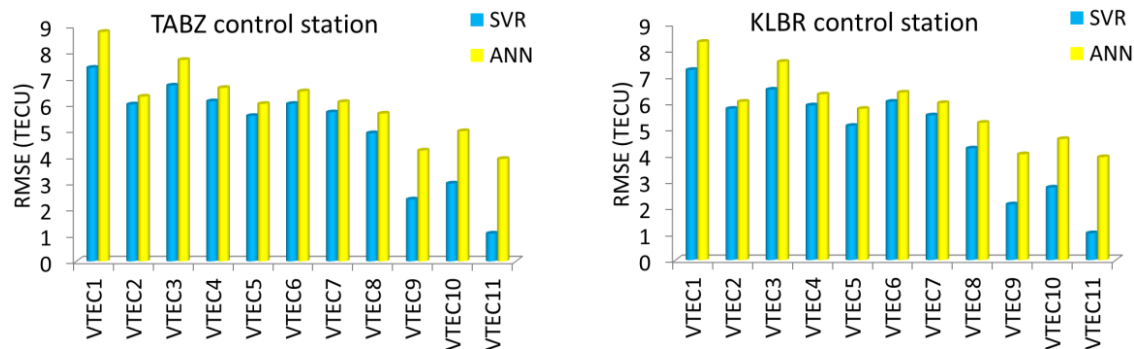
بر اساس نتایج حاصل از آنالیز شاخص dVTEC، مقدار خطای مدل‌های SVR و ANN در ایستگاه‌های کنترل داخلی TABZ و KLBR کمتر از سایر مدل‌های مورد بررسی است؛ اما در ایستگاه کنترل خارجی ANKR مقدار خطای مدل SVR افزایش قابل توجهی داشته است. این نتیجه مطابق با نتایج به‌دست آمده در جدول (۲) است. به عبارت دیگر مدل SVR خارج از محدوده شبکه مورد مطالعه، دقت پایینی دارد.

۳-۳. ارزیابی تأثیر پارامترهای خورشیدی و ژئومغناطیسی KP، F10.7، AP و DST در خروجی مدل‌های SVR و ANN با توجه به این نکته که مقدار محتوای الکترون کلی به

فعالیت‌های خورشیدی و ژئومغناطیسی وابسته است، لذا آموزش مدل‌های ANN و SVR با استفاده از چهار پارامتر فیزیکی KP، AP، F10.7 و DST انجام گرفت. به عبارت دیگر چهار پارامتر طول و عرض جغرافیایی، روز از سال و زمان به عنوان پارامترهای اصلی در مدل‌سازی بوده و در هر مرحله پارامترهای فیزیکی به ورودی دو مدل اضافه شدند. در نتیجه سناریوهای زیر مورد بررسی قرار گرفت:

$$\begin{aligned}
 VTEC_{(SVR\ ANN)_1} &= f(Lat., Lon., DOY, time) \\
 VTEC_{(SVR\ ANN)_2} &= f(Lat., Lon., DOY, time, KP) \\
 VTEC_{(SVR\ ANN)_3} &= f(Lat., Lon., DOY, time, AP) \\
 VTEC_{(SVR\ ANN)_4} &= f(Lat., Lon., DOY, time, DST) \\
 VTEC_{(SVR\ ANN)_5} &= f(Lat., Lon., DOY, time, F10.7) \\
 VTEC_{(SVR\ ANN)_6} &= f(Lat., Lon., DOY, time, KP, AP) \\
 VTEC_{(SVR\ ANN)_7} &= f(Lat., Lon., DOY, time, KP, DST) \\
 VTEC_{(SVR\ ANN)_8} &= f(Lat., Lon., DOY, time, DST, F10.7) \\
 VTEC_{(SVR\ ANN)_9} &= f(Lat., Lon., DOY, time, DST, F10.7, KP) \\
 VTEC_{(SVR\ ANN)_{10}} &= f(Lat., Lon., DOY, time, DST, F10.7, AP) \\
 VTEC_{(SVR\ ANN)_{11}} &= f(Lat., Lon., DOY, time, KP, AP, DST, F10.7)
 \end{aligned}
 \quad (22)$$

جهت آموزش دو مدل SVR و ANN، ابتدا پارامترهای فیزیکی به‌صورت جداگانه به مدل معرفی شده و میزان تأثیر آن پارامتر استخراج می‌شود، سپس در هر مرحله سایر پارامترهای فیزیکی وارد مدل شده و در نهایت به منظور تکمیل تر شدن مدل، هر چهار پارامتر فیزیکی در نظر گرفته می‌شوند. با توجه به اینکه در جدول (۲)، خطای هر دو مدل در ایستگاه کنترل خارجی ANKR بیشتر از ایستگاه‌های کنترل داخلی بود، در مرحله ارزیابی تأثیر پارامترهای خورشیدی و ژئومغناطیسی، تنها نتایج مربوط به ایستگاه‌های TABZ و KLBR ارائه شده است. شکل (۴) نتایج ارزیابی دو مدل SVR و ANN را نسبت به ۴ پارامتر فیزیکی KP، F10.7، DST و AP در دو ایستگاه کنترل داخلی نمایش می‌دهد.



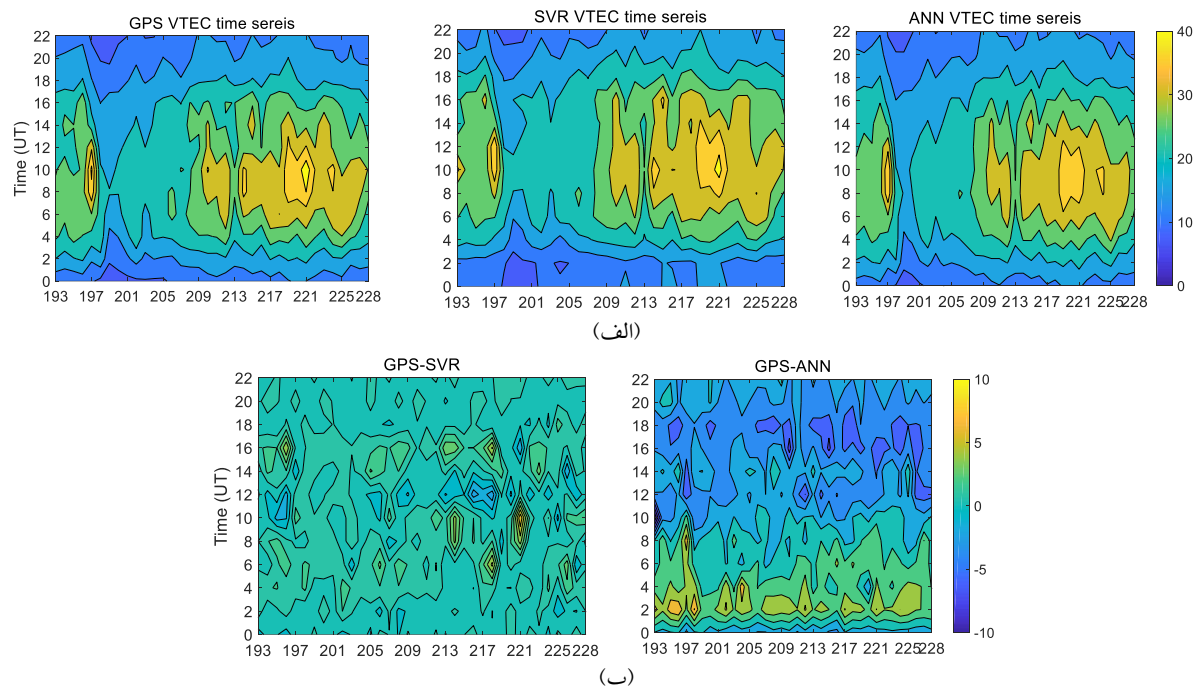
شکل ۴. مقادیر RMSE محاسبه شده بر اساس ۱۱ سناریو ارائه شده در رابطه (۲۲) و برای ایستگاه‌های کنترل داخلی TABZ و KLBR.

تأثیر زیادی در نتایج دو مدل SVR و ANN نداشته است؛ بنابراین، مطابق با آنالیز انجام گرفته در شکل (۴)، پارامترهای $F_{10.7}$ ، KP، DST و AP به ترتیب در اولویت‌های اول تا چهارم مدل‌سازی با SVR و ANN قرار دارند.

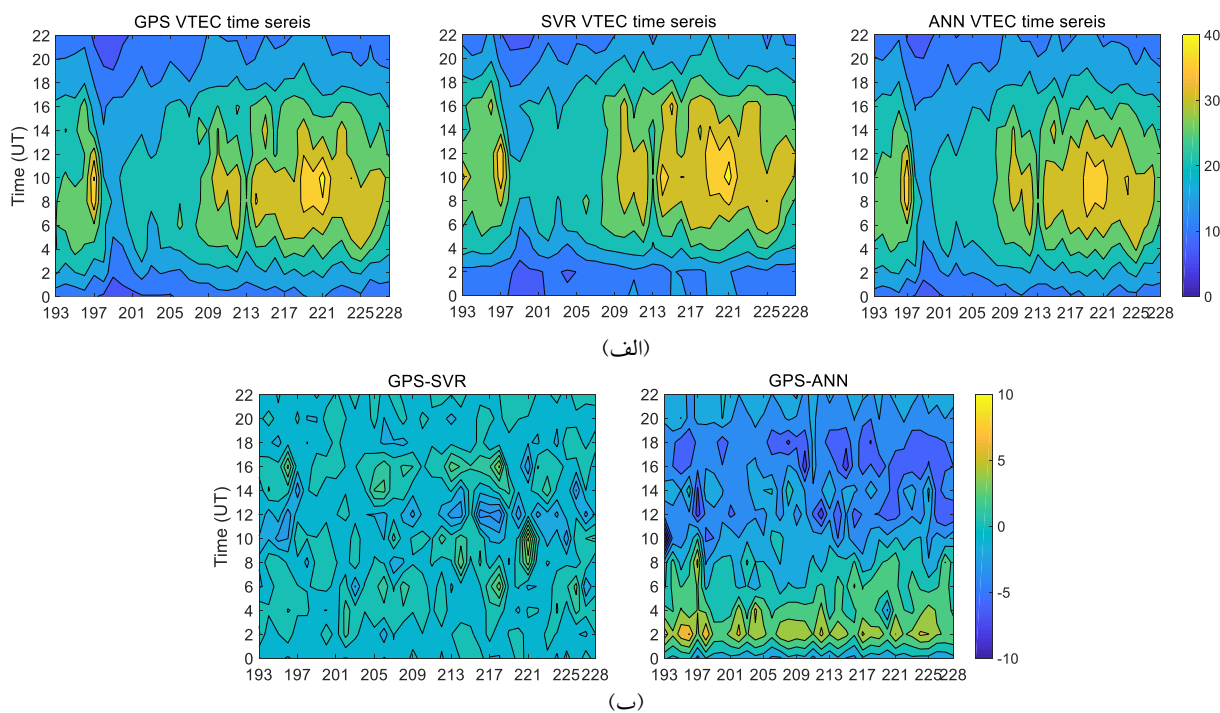
۳-۴. مدل‌سازی مکانی-زمانی TEC

پس از مرحله آموزش و تست دو مدل SVR و ANN و همچنین ارزیابی تأثیر پارامترهای فیزیکی در خروجی دو مدل، حال می‌توان با استفاده از مدل‌های آموزش دیده، تغییرات زمانی-مکانی TEC را در محدوده شبکه مورد مطالعه برآورد کرده و مورد ارزیابی قرار داد. با توجه به اینکه در مرحله ارزیابی دقت و صحت دو مدل SVR و ANN در ایستگاه‌های کنترل، مشخص شد که خطای این دو مدل در ایستگاه‌های کنترل داخلی کمتر از مدل‌های IRI2016، NeQuick و GIM است در نتیجه در این بخش تنها نتایج حاصل از دو مدل یادگیری ماشین ارائه می‌شوند. اشکال (۵) و (۶) سری زمانی VTEC برآورده شده توسط مدل‌های SVR و ANN را در مقایسه با VTEC حاصل از GPS در دو ایستگاه کنترل TABZ و KLBR نشان می‌دهند.

مطابق با آنالیز انجام گرفته در شکل (۴) و در هر دو ایستگاه کنترل TABZ و KLBR، در صورت استفاده از تنها چهار پارامتر اصلی طول و عرض جغرافیایی ایستگاه، روز از سال و زمان، مقدار RMSE دو مدل SVR و ANN بالای ۷ TECU شده است. در صورت مقایسه با نتایج جدول (۲)، در سناریو یک، مقدار خطای دو مدل یادگیری ماشین تقریباً برابر با مدل‌های تجربی IRI2016 و NeQuick است؛ اما با اضافه کردن یک پارامتر فیزیکی مانند KP در سناریو دو به پارامترهای ورودی دو مدل SVR و ANN، RMSE در حدود ۲ TECU کاهش داشته است. به عبارت دیگر، نتایج هر دو مدل به پارامتر فیزیکی KP حساسیت بالایی دارند. در مورد پارامتر فیزیکی AP و DST نیز بهبودی در حدود ۱/۵۵ TECU در نتایج هر دو مدل دیده می‌شود. در سناریو ۵ و با در نظر گرفتن پارامتر فیزیکی $F_{10.7}$ به همراه چهار پارامتر اصلی، مقدار RMSE هر دو مدل ۲/۲۵ TECU کاهش داشته است. به عبارت دیگر حساسیت دو مدل SVR و ANN نسبت به پارامتر فیزیکی $F_{10.7}$ کمی بیشتر از پارامتر KP است. مقایسه سناریوهای ۹ و ۱۰ نیز نشان می‌دهد که اثر پارامتر فیزیکی AP در نتایج دو مدل مورد بررسی کمتر از پارامتر فیزیکی KP می‌باشد. همچنین مقایسه سناریوهای ۹ و ۱۱ نشان می‌دهد که اضافه شدن پارامتر فیزیکی AP



شکل ۵. سری زمانی VTEC حاصل از مدل‌های SVR و ANN در مقایسه با سری زمانی حاصل از GPS (الف) و تفاضل سری‌های زمانی ایجاد شده با دو مدل از GPS (ب) در کلیه روزهای مورد مطالعه و موقعیت ایستگاه کنترل TABZ. محور افقی بیانگر روز از سال و محور عمودی نشان‌دهنده زمان به وقت جهانی است.



شکل ۶. سری زمانی VTEC حاصل از مدل‌های SVR و ANN در مقایسه با سری زمانی حاصل از GPS (الف) و تفاضل سری‌های زمانی ایجاد شده با دو مدل از GPS (ب) در کلیه روزهای مورد مطالعه و موقعیت ایستگاه کنترل KLBR. محور افقی بیانگر روز از سال و محور عمودی نشان‌دهنده زمان به وقت جهانی است.

یادگیری ماشین ارائه شده در این مقاله به سری زمانی حاصل از GPS بسیار شباهت دارد. تغییرات زمانی موجود

مطابق با نتایج حاصل از اشکال (۵) و (۶)، به‌وضوح قابل مشاهده است که سری زمانی VTEC حاصل از مدل‌های

کنترل TABZ، مقدار RMSE مدل‌های ANN، SVR، GIM، IRI2016 و NeQuick به ترتیب برابر با ۱/۰۶، ۳/۹۱، ۲/۹۲، ۶/۹۲ و ۷/۷۳ TECU حاصل شدند. در ایستگاه کنترل KLBR مقدار RMSE به ترتیب برابر با ۱/۰۲، ۳/۹۲، ۳/۱۲، ۶/۸۸ و ۷/۵۷ TECU به دست آمد؛ اما در ایستگاه کنترل خارجی از محدوده شبکه مورد مطالعه، مقدار RMSE مدل GIM از بقیه مدل‌ها کمتر بود. آنالیز خطای انجام گرفته در ایستگاه‌های کنترل نشان داد که مدل‌های یادگیری ماشین در محدوده شبکه مورد مطالعه از دقت و صحت بالایی برخوردار هستند. خارج از محدوده شبکه، دقت مدل SVR به شدت کاهش پیدا می‌کند. دلیل این امر نیز وابستگی پارامترهای تابع کرنل به مشاهدات شبکه مورد مطالعه است. در مرحله ارزیابی تأثیر پارامترهای ژئومغناطیسی و خورشیدی در مدل‌های ANN و SVR نیز مشخص شد که پارامترهای فیزیکی KP و F10.7 تأثیر بیشتری در نتایج مدل‌سازی دارند. پارامترهای DST و AP در اولویت‌های بعدی می‌باشند. به عبارت دیگر برای بررسی رفتار یونسفر با دقت بالا، می‌بایستی تأثیر پارامترهای KP و F10.7 را نیز در نظر گرفت. نتایج این مقاله نشان داد که مدل‌های یادگیری ماشین علی‌الخصوص مدل SVR دارای دقت و صحت بسیار بالایی در مدل‌سازی محلی یونسفر هستند و در صورت استفاده از پارامترهای فیزیکی KP و F10.7 در ورودی‌های این مدل، نتایج حاصل از آن به شدت قابل اعتماد و تفسیر خواهد بود.

تشکر و قدردانی

نویسندگان مقاله از داوران محترم به دلیل ارائه نکات بسیار ارزنده و علمی تشکر و قدردانی می‌کنند. همچنین، از سازمان نقشه‌برداری کشور (NCC) به دلیل در اختیار قرار دادن مشاهدات ایستگاه‌های شبکه محلی آذربایجان در روزهای مورد درخواست، تشکر و قدردانی می‌شود.

منابع

در VTEC در ساعات مختلف روز و در روزهای مختلف، قابل مشاهده است. همچنین، تفاضل سری‌های زمانی حاصل از مدل‌های ANN و SVR از سری زمانی GPS در هر دو ایستگاه کنترل نشان می‌دهد که مقدار VTEC حاصل از SVR به GPS بسیار نزدیک است. نکته دیگر این است که بیشینه VTEC در ساعات ۸ الی ۱۱ بوقت جهانی به وقوع پیوسته است. نتایج بیانگر این موضوع است که مدل‌های یادگیری ماشین قابلیت بسیار بالایی برای مدل‌سازی و پیش‌بینی تغییرات زمان-مکانی VTEC یونسفر را دارند و از این روش‌ها می‌توان با دقت بالایی تغییرات یونسفر را مورد بررسی و آنالیز قرار داد.

۴. نتیجه‌گیری و بحث

هدف اصلی این مقاله ارزیابی تأثیر پارامترهای فیزیکی مختلف در نتایج دو مدل یادگیری ماشین یعنی مدل رگرسیون بردار پشتیبان (SVR) و شبکه عصبی مصنوعی (ANN) بود. برای این منظور از مشاهدات ۱۵ ایستگاه GPS در شمال غرب ایران در روزهای ۱۹۳ الی ۲۲۸ از سال ۲۰۱۲ استفاده شد (دلیل انتخاب این منطقه در دسترس بودن اطلاعات کاملی از اندازه‌گیری‌های ایستگاه‌های GPS بود). پس استخراج مقادیر VTEC از مشاهدات ایستگاه‌های GPS، با استفاده از ۸ پارامتر ورودی (طول و عرض جغرافیایی ایستگاه GPS، روز از سال (DOY)، ساعت، شاخص‌های ژئومغناطیسی AP، KP و DST، شاخص فعالیت‌های خورشیدی F10.7) مدل‌ها آموزش داده شدند. ارزیابی خطای مدل‌ها در دو ایستگاه کنترل داخلی (TABZ و KLBR) و یک ایستگاه کنترل خارجی (ANKR) و با استفاده از پارامترهای آماری مختلف انجام گرفت. همچنین، نتایج حاصل از دو مدل ANN و SVR با VTEC حاصل از GPS و مدل‌های تجربی IRI2016، NeQuick و مدل GIM مقایسه و ارزیابی شدند.

بر اساس نتایج حاصل از آنالیزهای انجام گرفته در ایستگاه

- Amerian, Y., Voosoghi, B., & Mashhadi Hossainali, M. (2013). Regional Ionosphere Modeling in Support of IRI and Wavelet Using GPS Observations. *Acta Geophysica*, 61(5), 1246-1261, DOI: 10.2478/s11600-013-0121-5.
- Bilitza, D., & Reinisch, B. W. (2008). International reference ionosphere 2007: Improvements and new parameters. *Advances in space research*, 42(4), 599-609.
- Browne, S., Hargreaves, J., & Honary, B. (1995). An imaging riometer for ionospheric studies. *Electronics & communication engineering journal*, 7(5), 209-217.
- Cander, R. (1998). Artificial neural network applications in ionospheric studies. *Annali di Geofisica*, 5-6(1998), 757-766.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- Etemadfar, H., & Mashhadi Hossainali, M. (2016). Application of Slepian theory for improving the accuracy of SH-based global ionosphere models in the Arctic region. *Journal of Geophysical Research: Space Physics*, 121(3), 2583-2594.
- Feizi R, Voosoghi B, & Ghaffari Razin M. R. (2020). Regional modeling of the ionosphere using adaptive neuro-fuzzy inference system in Iran. *Advances in Space Research*, 65, 2515-2528.
- Ghaffari Razin, M. R., & Voosoghi, B. (2017). Ionosphere tomography using wavelet neural network and particle swarm optimization training algorithm in Iranian case study. *GPS Solutions*, 21(3), 1301-1314.
- Ghaffari-Razin, M. R., & Voosoghi, B. (2018). Application of Wavelet Neural Networks for Improving of Ionospheric Tomography Reconstruction over Iran. *Journal of the Earth and Space Physics*, 44(4), 99-114.
- Ghaffari Razin, M. R., & Voosoghi, B. (2016a). Modeling of ionosphere time series using wavelet neural networks (case study: NW of Iran). *Advances in Space Research*, 58(1), 74-83.
- Ghaffari Razin, M. R., & Voosoghi, B. (2016b). Wavelet neural networks using particle swarm optimization training in modeling regional ionospheric total electron content. *Journal of Atmospheric and Solar-Terrestrial Physics*, 149, 21-30.
- Habarulema, J. B., McKinnell, L. A., & Opperman, B. D. (2011). Regional GPS TEC modeling; Attempted spatial and temporal extrapolation of TEC using neural networks. *Journal of Geophysical Research: Space Physics*, 116(A4), 1-14.
- Huang, Z., Li, Q., & Yuan, H. (2015). Forecasting of ionospheric vertical TEC 1-h ahead using a genetic algorithm and neural network. *Adv. Space Res.* 55, 1775-1783.
- Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*, MacMillan College Publishing Co. New York.
- Hofmann-Wellenhof, B., Lichtenegger, H., & Wasle, E. (2007). GNSS—global navigation satellite systems: GPS, GLONASS, Galileo, and more: Springer Science & Business Media.
- Inyurt, S., & Sekertekin, A. (2019). Modeling and predicting seasonal ionospheric variations in Turkey using artificial neural network (ANN). *Astrophysics and Space Science*, 364(4), 1-8.
- Jang, H., & Topal, E. (2014). A review of soft computing technology applications in several mining problems. *Applied Soft Computing*, 22, 638-651.
- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.
- Komjathy, A. (1997). Global ionospheric total electron content mapping using the Global Positioning System. University of New Brunswick Fredericton.
- Leick, A., Rapoport, L., & Tatarnikov, D. (2015). GPS satellite surveying: John Wiley & Sons.
- Mars, P., Chen, J., Nambiar, R., & Fidler, J. (1996). *Learning Algorithms: Theory and Applications in Signal Processing*: CRC Press, Inc.
- Muhtarov, P., Kutiev, I., & Cander, L., (2002). Geomagnetically correlated autoregression model for short-term prediction of ionospheric parameters. *Inverse Problems*. 18(1), 49.
- Mautz, R., Ping, J., Heki, K., Schaffrin, B., Shum, C., & Potts, L. (2005). Efficient spatial and temporal representations of global ionosphere maps over Japan using B-spline wavelets. *Journal of Geodesy*, 78(11), 662-667.
- Nava, B., Coisson, P., & Radicella, S. (2008). A new version of the NeQuick ionosphere electron density model. *Journal of Atmospheric and Solar-Terrestrial Physics*, 70(15), 1856-1862.
- Nematipour, P., Raoofian-Naeeni, M., & Ghaffari Razin, M. R. (2021). Regional application of C1 finite element interpolation method in modeling of ionosphere total electron content over Europe. *Advances in Space Research*, 69(3), 1351-1365.
- Sayin, I., Arikan, F., Arikan, O. (2008). Regional TEC mapping with random field priors and kriging. *Radio Science*, 43(5), 1-14.
- Schunk, R.W., & Nagy, A.F. (2000). *Ionospheres: Physics, Plasma Physics, and Chemistry*, Cambridge University Press, 554.
- Seeber, G. (2003). *satellite geodesy: foundations*.

- Methods and applications*, Walter de Gruyter, Berlin and New York, 53.
- Sharifi, M. A., & Farzaneh, S. (2015). Regional TEC dynamic modeling based on Slepian functions. *Advances in Space Research*, 56(5), 907-915.
- Simpson, P. (1990). Artificial neural system-foundation, paradigm, application and implementation Pergamon Press New York.
- Smola, A. J., & Schölkopf, B. (1998). On a kernel-based method for pattern recognition, regression, approximation, and operator inversion. *Algorithmica*, 22(1), 211-231.
- Tebabal, A., Radicella, S., Damtie, B., Migoya-Orue, Y., Nigussie, M., & Nava, B. (2019). Feed forward neural network based ionospheric model for the East African region. *Journal of Atmospheric and Solar-Terrestrial Physics*, 191, 105052.
- Vapnik, V. (1995). Support-vector Networks. *Machine Learning*, 20, 273-297.
- Xia, G., Liu, Y., Wei, T., Wang, Z., Huang, W., Du, Z., Zhang, Z., Wang, X., & Zhou, C. (2021). Ionospheric TEC forecast model based on support vector machine with GPU acceleration in the China region. *Advances in Space Research*, 68(3), 1377-1389.
- Yeganeh, B., Motlagh, M. S. P., Rashidi, Y., & Kamalan, H. (2012). Prediction of CO concentrations based on a hybrid Partial Least Square and Support Vector Machine model. *Atmospheric Environment*, 55, 357-365.