

پیش‌بینی کوتاه‌مدت حرکت قطبی با استفاده از شبکه‌های عصبی ژرف

محمدعلی شریفی^{۱*}، کورش شهریاری‌نیا^۲، شایان شیرافکن^۳، سید محسن خضرائی^۴، علیرضا امیری سیمکویی^۵

^۱ دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - دانشکده‌گان فنی - دانشگاه تهران
sharifi@ut.ac.ir

^۲ کارشناس ارشد هیدروگرافی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - دانشکده‌گان فنی - دانشگاه تهران
kouroshshahryari94@gmail.com

^۳ دانشجوی کارشناسی ارشد ژئودزی - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - دانشکده‌گان فنی - دانشگاه تهران
shayanshirafkan@ut.ac.ir

^۴ دکتری ژئودزی گروه مهندسی نقشه برداری - دانشکده عمران و حمل و نقل - دانشگاه اصفهان
S.mohsen.khazraei@gmail.com

^۵ استاد گروه مهندسی نقشه برداری - دانشکده عمران و حمل و نقل - دانشگاه اصفهان
ar.amirisimkooei@gmail.com

(تاریخ دریافت آذر ۱۴۰۰، تاریخ تصویب اسفند ۱۴۰۰)

چکیده

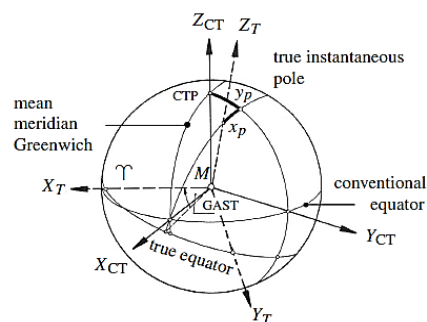
امروزه با توجه به نیاز روز افزون پیش‌بینی پارامترهای حرکت قطبی از روش‌های مختلفی استفاده می‌شود. وجود مجموعه داده‌های بلند مدت پارامترهای x و y حرکت قطبی سبب بوجود آمدن منبع ارزشمندی برای پیش‌بینی این پدیده‌ی ژئوفیزیکی است. در این پژوهش از داده‌های بلند مدت حرکت قطبی برای پیش‌بینی دوره‌های ۴۰ روزه به مدت ۵ سال استفاده شده است. برای انجام پیش‌بینی‌ها از شبکه‌های عصبی کانولوشنی، شبکه حافظه طولانی کوتاه مدت، شبکه پرسپترون چندلایه و روش آنالیز هارمونیک کمترین مربعات استفاده شده است. در ادامه نتایج شبکه‌های عصبی با یکدیگر و روش آنالیز هارمونیک کمترین مربعات مقایسه گردیده است. در پایان مشخص شد که شبکه‌های حافظه طولانی کوتاه مدت در بیشتر روزها از شبکه‌های کانولوشنی و پرسپترون چندلایه نتایج بهتری دارند. همچنین مشخص شد که این روش‌ها دقت پیش‌بینی بسیار بهتری نسبت به روش آنالیز هارمونیک کمترین مربعات دارند. برای شبکه‌های ژرف برای پارامتر x بهترین مقادیر RMSE برای روزهای اول و چهارم به ترتیب ۰/۶۸، ۹/۲۲ میلی ثانیه کمانی است و برای پارامتر y بهترین مقادیر RMSE برای روزهای اول و چهارم به ترتیب ۰/۹۵ و ۴۸/۴۰ میلی ثانیه کمانی است. در روش آنالیز هارمونیک کمترین مربعات مقدار RMSE برای پارامتر x برای روزهای اول و چهارم به ترتیب برابر با ۴۰/۹۵ و ۴۸/۴۰ میلی ثانیه کمانی است و مقدار RMSE پارامتر y برای روزهای اول و چهارم به ترتیب برابر با ۴۰/۸۶ و ۴۶/۵۳ میلی ثانیه کمانی است.

واژگان کلیدی: سری زمانی، حرکت قطبی، شبکه حافظه طولانی کوتاه مدت، شبکه عصبی کانولوشنی، شبکه پرسپترون چندلایه، آنالیز هارمونیک کمترین مربعات

* نویسنده رابط

۱- مقدمه

چرخش زمین نه تنها وضعیت حرکت کلی زمین را مشخص می‌کند، بلکه ارتباط بین زمین صلب و جو، اقیانوس‌ها، گوشته و هسته را در مقیاس‌های مکانی و زمانی منعکس می‌کند [۱] [۲]. پارامترهای توجیه زمین (EOP) شامل حرکت قطبی (PM) و طول روز (LOD) است. حرکت قطبی یک نمایش عینی از چرخش زمین است، بنابراین تغییرات آن نیز نشان‌دهنده تغییرات در عوامل مختلف ژئوفیزیکی است. لذا مطالعه‌ی تغییرات PM برای درک فرآیندهای ژئوفیزیکی، مانند حرکت جرم در پوسته (همرفت گوشته، توپوگرافی، جریانات الکترومغناطیسی و حرکت صفحه) و جریانات مواد روی سطح زمین از ارزش و اهمیت بالایی برخوردار است [۳]. با استفاده از مطالعه روی PM می‌توان به مطالعات جذابی همانند گردش اقیانوس و جزر و مد، حرکت اتمسفر، بازگشت پس از یخبندان، و ذوب یخچال‌های طبیعی و اثرات خارجی مانند تأثیر ماه، خورشید و سایر اجرام آسمانی بر روی زمین پرداخت [۴] [۵] [۶] [۷] [۸] [۹]. PM حرکت محور چرخشی زمین نسبت به پوسته است که معمولاً با استفاده از مختصات x و y بیان می‌شود و این محورهای مختصات به ترتیب در جهت‌های 0° درجه و 90° درجه غربی هستند. شکل ۱، این سیستم مختصات را نمایش می‌دهد.



شکل ۱- سیستم مختصات حرکت قطب به مبداء CIO [۱۰]

نوسانات بلندمدت، دوره تناوب‌های چندلر و سالیانه اجزای مهم PM هستند که در آنالیزهای مختلف سری‌های زمانی PM اهمیت زیادی دارند [۱۱]. دوره

تناوب‌های چندلر توسط چندلر در سال ۱۸۹۱ معرفی شدند و از آن زمان تاکنون، تحقیقات مداومی روی آن صورت گرفته است [۲] [۸]. با گذشت زمان، دوره تناوب چندلر از $1/13$ تا $1/20$ سال در نوسان است [۱۲]، و دامنه آن بین ۱۰۰ تا ۲۰۰ میلی ثانیه کمانی (mas) متغیر است [۸] [۱۳]. اگرچه مطالعات کنونی نشان داده است که دوره تناوب‌های چندلر به فعل و انفعالات بین اتمسفر، اقیانوس و دیگر سیالات زمین مرتبط است اما علل پیدایش آن‌ها همچنان مورد مطالعه بیشتر قرار دارد [۱] [۴] [۱۴]. این در حالی است که دوره تناوب و دامنه سیگنال سالیانه تقریباً ثابت می‌ماند و نوسانات بلندمدت تقریباً بین $3/3$ تا $3/5$ میلی ثانیه کمانی در سال است [۱۲] [۱۳]. برای استخراج بهتر سیگنال‌های PM و مطالعه ویژگی‌های متغیر زمانی آن‌ها، از روش‌های مختلف آنالیز سری‌های زمانی استفاده می‌شود. به عنوان مثال، پوپینسکی و همکاران در سال ۱۹۹۵ از روش تبدیل فوریه برای تجزیه و تحلیل تغییرات PM استفاده کردند و دوره تناوب‌های چندلر و سالیانه را از سری‌های زمانی IERS93C01 و IERS90C04 استخراج کردند. جدول ۱، دوره تناوب‌های غالب در سری زمانی حرکت قطبی را نمایش می‌دهد [۱۲] [۱۳].

جدول ۱- دوره تناوب‌های غالب در سری زمانی x و y [۱۵] [۱۶]

تعداد	دوره تناوب (روز)
۱	۴۳۳
۲	۳۶۵
۳	۳۰۰
۴	بین ۵۱۲ تا ۷۳۰

تبدیل تجزیه موجک برای تعیین، تجزیه و تحلیل دامنه‌های دوره تناوب‌های چندلر سالیانه اعمال شده است [۱۲] [۱۷]. از سال ۲۰۱۰، تجزیه و تحلیل طیف منفرد (SSA) برای تجزیه و تحلیل PM استفاده شده است. مالکین و میلر در سال ۲۰۱۰ از SSA برای استخراج دوره تناوب‌های چندلر و سالیانه PM استفاده کردند و نتایج نشان داد که این روش می‌تواند ساختار سری زمانی PM را با جزئیات بیشتری نشان دهد [۱۴]. وانگ و همکاران در سال

۱ Earth Orientation Parameters

۲ Polar Motion

۳ Length Of Day

۴ Conventional International Origin

۵ Milliarsecond

۶ Singular Spectral Analysis

بسیار استفاده شده‌اند برای مثال، چیمولا و همکاران در سال ۲۰۲۰ از این شبکه‌ها برای پیش‌بینی انتشار کووید ۱۹ در کشور کانادا تا ۱۴ روز آینده استفاده نمودند [۲۹]. البته باید گفت که شبکه‌های عصبی به شکل گسترده در پژوهش‌های مرتبط با ژئودزی و هیدروگرافی مورد استفاده قرار گرفته است. برای مثال، یانگ و همکاران سطح آب جزر و مدی ۱۷ بندر تابوان را به کمک روش‌های متفاوتی مانند ماشین بردار پشتیبانی (SVR^۱)، CNN، LSTM، ANN سنتی و آریمای (ARIMA^{۱۰}) پیش‌بینی کردند و در پایان مشخص شد که نتایج شبکه LSTM از نتایج سایر روش‌ها دقت بهتری دارند [۳۰] هم چنین تیرا از شبکه عصبی مصنوعی به عنوان ابزاری برای انجام تبدیلات بین سیستم‌های مرجع ژئودتیکی استفاده نمود. در این پژوهش از نقاط با مختصات مشخص در سیستم مرجع آمریکای جنوبی با نام (SAD69) و سیستم قدیمی تر Córrego Alegre استفاده شد [۳۱]. هر دوی این سیستم‌ها در آمریکای جنوبی استفاده می‌شوند. در همین راستا شبکه‌های عصبی برای پیش‌بینی پارامترهای توجیه زمین نیز استفاده شده است. برای مثال در پژوهش‌های مشابه، شو و همکاران از شبکه عصبی مصنوعی برای پیش‌بینی پارامترهای توجیه زمین استفاده نمودند. آن‌ها در ابتدا با استفاده از مدل‌هایی اثرات جزرومد زمین صلب و تغییرات فصلی اتمسفری را از روی سری زمانی LOD حذف نمودند. در ادامه با استفاده از شبیه‌ساز شبکه عصبی اشتوتگارت تلاش نمودند تا ساختارهای گوناگون شبکه عصبی را امتحان نموده و بهترین نتایج را استخراج کنند [۳۲]. لیاو با استفاده از شبکه عصبی مصنوعی این پارامترها را تا ۳۶۰ روز آینده پیش‌بینی نمود. برای این منظور از یک شبکه پیش‌خور استفاده نمودند و با استفاده از الگوریتم BP وزن‌ها را اصلاح کردند [۳۳]. پیش‌بینی‌ها شامل تمامی بخش‌های خطی و غیر خطی بودند. در پایان جواب‌ها با روش‌ها و کارهای مشابه مقایسه گردید. در این پژوهش، تلاش شده از شبکه‌های عصبی ژرف برای پیش‌بینی حرکت قطبی استفاده شود. همچنین برای سنجش دقت شبکه‌ها، نتایج به دست آمده با روش آنالیز هارمونیک کمترین مربعات (LSHA^{۱۱}) مقایسه شده است.

۲۰۱۶ یک روش ردیابی مبتنی بر فوریه را پیشنهاد کردند که می‌تواند به‌طور موثر اثر لبه را در استخراج دوره تناوب‌های چندلر و سالیانه از بین ببرد [۸]. از گذشته تاکنون روش‌های مختلف پیش‌بینی PM توسعه داده شده است، معمولاً، مدل‌های پیش‌بینی برون‌یابی براساس دوره‌های تناوب‌های مشخص PM با استفاده از روش برازش هارمونیک کمترین مربعات (LSHA^۱) با در نظر گرفتن سیگنال‌ها و ترند خطی به‌همراه یک مدل تصادفی برای پیش‌بینی باقیمانده‌های آن مورد استفاده قرار می‌گیرد [۱۸] [۱۹]. اگر چه در میان این مدل‌های موجود، مدل کمترین مربعات (LS^۲) و اتورگرسیو (AR^۳) در حال حاضر یکی از مؤثرترین مدل‌ها برای پیش‌بینی کوتاه‌مدت و بلندمدت PM در نظر گرفته می‌شود [۲۰] [۲۱] [۲۲] اما به‌دلیل وجود عوامل ژئوفیزیکی و پیچیدگی آن‌ها در ساختار سری زمانی PM پیش‌بینی آن کار ساده‌ای نیست [۴] و نتایج نشان داده است که مدل‌های خطی دقت کمتری نسبت به مدل‌های غیرخطی دارند [۲۳]. بنابراین بسیاری از محققین از مدل‌های غیرخطی مانند سیستم استنتاج فازی شبکه عصبی تطبیقی برای پیش‌بینی PM استفاده کرده‌اند [۲۴] [۲۵]. استفاده از شبکه‌های عصبی مصنوعی (ANN^۴) دارای پیشینه بسیار طولانی است. این شبکه‌ها با نام شبکه‌های عصبی پیش‌خور (FFNN^۵) هم شناخته می‌شوند و از انتهای دهه ۴۰ میلادی مطرح شدند. راملهارت و هینتون در سال ۱۹۸۶ با ارائه الگوریتم پس انتشار (BP^۶) توان پردازش و محاسبات شبکه‌های عصبی را افزایش دادند [۲۶]. لچان در سال ۱۹۸۹ شبکه‌ای طراحی نمود که توانایی تشخیص اعداد و نوشته‌های روی کاغذ را داشت [۲۷]. این شبکه‌ها با نام شبکه‌های کانولوشنی (CNN^۷) شناخته می‌شوند و در سال‌های بعد بسیار استفاده شدند. در انتهای دهه ۹۰ میلادی، شبکه‌های حافظه طولانی کوتاه - مدت (LSTM^۸) معرفی شدند که نوعی از شبکه‌های بازگشتی هستند [۲۸]. این شبکه‌ها دارای نوعی حافظه هستند و برای آنالیز سری‌های زمانی مختلف استفاده می‌شوند. شبکه‌های LSTM در حوزه‌های گوناگون علمی

^۱ Least Square Harmonic Analysis

^۲ Least-square

^۳ Autoregressive

^۴ Artificial Neural Network

^۵ Feed Forward Neural Network

^۶ Back-Propagation

^۷ Convolutional Neural Network

^۸ Long Short-Term Network

^۹ Support Vector Machines

^{۱۰} Autoregressive Integrated Moving Average

^{۱۱} Least Square Harmonic Estimation

۲- روش تحقیق

در این پژوهش برای پیش‌بینی هرچه دقیق‌تر حرکت قطبی از داده‌های بلند مدت استفاده شده است. این داده‌ها به شکل روزانه موجود هستند و توسط وب‌سایت (IERS^۱) ارائه می‌شوند. همانطور که در گذشته اشاره شد، واحد این داده‌ها میلی ثانیه کماتی است. داده‌های مذکور برای آموزش شبکه‌های عصبی استفاده می‌شوند. شبکه‌های عصبی استفاده شده در این پژوهش از انواع شبکه‌های عصبی ژرف پر استفاده در علوم مختلف هستند که در ادامه به شکل گسترده توضیح داده شده‌اند. برای بیان چگونگی انجام محاسبات در شبکه‌ها فلوچارت این شبکه‌ها رسم شده است. پس از پایان آموزش شبکه‌ها، پیش‌بینی‌های کوتاه مدت مربوط به دوره‌های ۴۰ روزه انجام شده است. همچنین به منظور سنجش کیفیت روش ارائه شده، همین پیش‌بینی‌ها به وسیله روش LSHA نیز انجام شده است و در ادامه نتایج شبکه‌های عصبی و روش LSHA در جدول‌هایی نشان داده شده و مقایسه گردیده است. در پایان توضیحاتی در مورد این نتایج بیان شده است.

۲-۱- شبکه‌های عصبی ژرف (عمیق)

مفاهیم ابتدایی شبکه عصبی در دهه ۴۰ میلادی بیان گردید و در طی سال‌های اخیر این شبکه‌ها با پیشرفت در سخت‌افزارهای کامپیوتری و به دلیل کارایی در زمینه‌های گوناگون علمی بسیار مورد توجه قرار گرفته است. در این شبکه‌ها از واحدهای محاسباتی بیشتری استفاده می‌شود و با توجه به پیشرفت‌های سخت‌افزاری رایانه‌ای امروزی بسیار استفاده می‌شوند. در این پژوهش تلاش شده که با توجه به اهمیت پیش‌بینی هرچه دقیق‌تر حرکت قطبی، از این شبکه‌ها برای این منظور استفاده گردد. در پایان نتایج شبکه‌ها با یکدیگر مقایسه می‌شوند و شبکه با دقت مناسب مشخص می‌شود.

۲-۱-۱- شبکه عصبی حافظه طولانی کوتاه مدت

در این شبکه‌ها از یک مکانیزم بازگشتی استفاده می‌شود، همچنین برای اصلاح وزن‌های شبکه نیز از الگوریتم BP استفاده می‌گردد [۲۸]. در الگوریتم

^۱ www.IERS.org

اصلاح وزن‌ها به وسیله بازگشت در طول شبکه و انجام مشتق زنجیره‌ای انجام می‌شود. اصلاح وزن‌ها به منظور کاهش مقدار تابع زیان در شبکه‌های با نظارت^۲ انجام می‌گیرد. برتری شبکه‌های LSTM نسبت به شبکه‌های بازگشتی سنتی، توانایی این شبکه‌ها در کار با دنباله‌های طولانی عددی و بدون از بین رفتن گرادیان^۳ است [۳۰]. این برتری به دلیل استفاده از دروازه فراموشی^۴ در شبکه‌های LSTM است. این دروازه، وظیفه تصمیم‌گیری در مورد دور انداختن اطلاعات ناخواسته را بر عهده دارد و این کار به وسیله یک تابع سیگموئید صورت می‌گیرد. فرض کنید $X=[x_1, x_2, \dots, x_t]$ ، بردار داده‌های ورودی به یک سلول شبکه LSTM است. روابط (۱) تا (۶) برای برزرسانی وضعیت سلول^۵ مورد استفاده قرار می‌گیرد که به همراه وضعیت پنهان^۶ برای محاسبات بعدی مورد استفاده قرار می‌گیرد. شکل ۲ ساختار یک سلول LSTM را نشان می‌دهد.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (1)$$

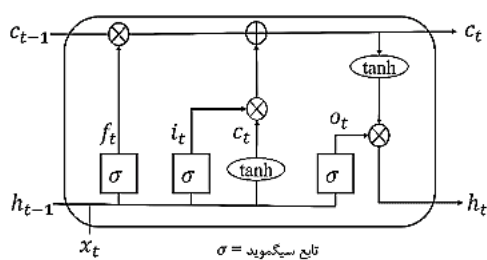
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \circ \tanh(C_t) \quad (6)$$



شکل ۲- ساختار داخلی حافظه طولانی کوتاه مدت

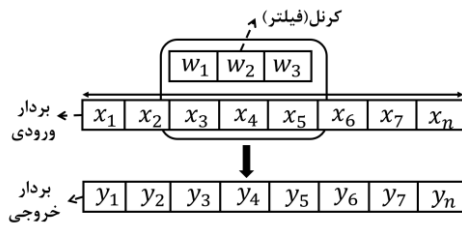
همان ضرب درایه‌ای است. در این شبکه، دروازه فراموشی وظیفه حفظ اطلاعات یا دور ریختن آن‌را برعهده

^۲ Supervised
^۳ Vanishing Gradient
^۴ Forget Gate
^۵ Cell State
^۶ Hidden State

دارد. این کار بوسیله تابع سیگموئید صورت می‌پذیرد. ورودی این قسمت x_t و h_{t-1} است که به ترتیب ورودی و وضعیت پنهان قبلی هستند و خروجی این بخش f_t است که با نام دروازه خروجی شناخته می‌شود رابطه (۱) عملیاتی که در قسمت دروازه خروجی انجام می‌شود را نشان می‌دهد. در ادامه از رابطه (۲) که دروازه ورودی^۱ (i_t) نام دارد برای بروزرسانی وضعیت سلول استفاده می‌کنیم. در این قسمت هم تابع سیگموئید وظیفه‌ای مشابه قبل دارد. در همین قسمت، وضعیت پنهان قبلی و ورودی کنونی به تابع تانژانت هیپربولیک وارد می‌شود (رابطه ۳). پس از آن خروجی تانژانت هیپربولیک و سیگموئید در یکدیگر ضرب شده و از آنجایی که سیگموئید خروجی به شکل ۰ و ۱ دارد، با ضرب در خروجی تابع هیپربولیک به نوعی تصمیم می‌گیرد که چه اطلاعاتی برای وضعیت سلول حفظ شود. در ادامه از رابطه (۴) برای بروزرسانی وضعیت سلول جدید (C_t) استفاده می‌کنیم و وضعیت سلول قبلی و مقادیر خروجی از دروازه ورودی با هم جمع می‌شوند و وضعیت سلول بروزرسانی می‌شود. در انتها دروازه خروجی قرار دارد (O_t). در این قسمت وضعیت پنهان بعدی محاسبه می‌شود (h_t). وضعیت پنهان به نوعی اطلاعاتی از ورودی‌های قبلی است و برای انجام پیش‌بینی‌های بعدی استفاده می‌شود. در این بخش، ابتدا وضعیت پنهان قبلی و ورودی‌های کنونی را وارد می‌کنیم (رابطه ۵). در ادامه وضعیت سلول بروزرسانی شده (خروجی قسمت قبل) را به عنوان ورودی تابع تانژانت هیپربولیک استفاده می‌کنیم و مقدار به دست آمده در خروجی دروازه خروجی ضرب می‌کنیم و وضعیت پنهان جدید به دست می‌آید (رابطه ۶). در روابط ۱ تا ۶، پارامترهای b بایاس و پارامترهای W وزن‌ها هستند.

$$z[n] = (x * g)(n) = \sum_m x[k] \cdot g[n - k] \quad (7)$$

در رابطه (۷)، علامت * عملیات ضرب کانولوشنی را نشان می‌دهد. همچنین x بردار ورودی تک‌بعدی، g کرنل (فیلتر) تک بعدی و $z[n]$ خروجی است که با نام نقشه ویژگی شناخته می‌شود. باید به این نکته اشاره کرد که شبکه‌های کانولوشنی زیرمجموعه‌ای از یادگیری با نظارت هستند و از تابع زیان و الگوریتم BP استفاده می‌کنند. شکل ۳ یک لایه ورودی و ضرب کانولوشنی را نشان می‌دهد. توجه نمایید که در این عملیات فیلتر بر روی داده در یک سمت حرکت می‌کند و پیوسته عملیات ضرب انجام می‌شود و بردار خروجی حاصل می‌شود. این بردارهای خروجی، ویژگی‌های شاخص و مهمی هستند که توسط فیلتر به وسیله ضرب انجام شده استخراج شده است.



شکل ۳- عملیات ضرب کانولوشنی در یک لایه CNN

۲-۱-۳- شبکه پرسپترون چند لایه^۲

MLP زیرمجموعه شبکه‌های عصبی پیشخور است و از سه بخش اصلی شامل لایه‌های ورودی، خروجی و پنهان تشکیل شده است. شبکه‌های MLP از الگوریتم BP برای اصلاح وزن‌ها استفاده می‌کند. در سال‌های اخیر، از بهینه‌ساز^۳ها و توابع فعال‌سازی جدید برای افزایش دقت و قدرت شبکه‌های MLP استفاده شده است. لایه ورودی MLP داده‌های ورودی را دریافت می‌کند در ادامه آن را به اولین لایه پنهان ارسال می‌کند. در این لایه محاسبات

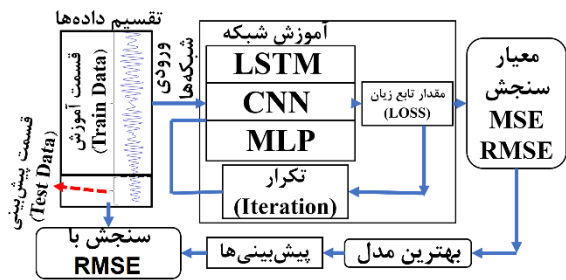
۲-۱-۲- شبکه عصبی کانولوشنی

شبکه CNN شناخته‌شده‌ترین نوع شبکه‌های ژرف است و در چند سال گذشته در زمینه‌های گوناگون و عمدتاً برای دسته‌بندی تصاویر و شناسایی اشیاء بسیار مورد استفاده قرار گرفته است [۳۴] [۳۵]. این شبکه‌ها توانایی کار با ساختارهای پیچیده تک بعدی و سری‌های زمانی را دارند و می‌توان در زمینه پیش‌بینی نیز از این شبکه‌ها

^۲ Multi-Layer Perceptron

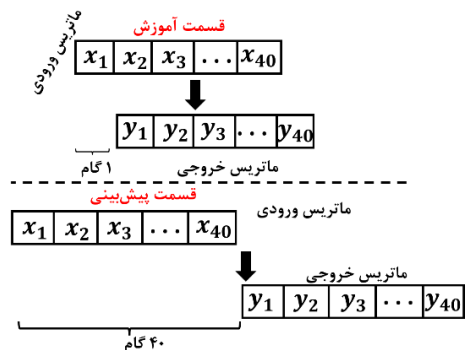
^۳ Optimizer

^۱ Input Gate



شکل ۵- فلوجارت چگونگی انجام محاسبات در شبکه‌ها

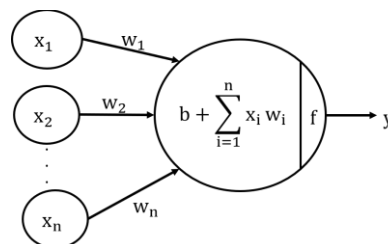
پیش‌بینی‌های انجام شده در این پژوهش برای بازه کوتاه مدت ۴۰ روزه است، بنابراین خروجی شبکه‌ها بردارهای ۴۰ روزه است. ورودی شبکه‌ها در قسمت آموزش هم بردارهای ۴۰ روزه است و در هر گام آموزش برای درک بهتر روابط و آموزش طولانی‌تر در قسمت آموزش تنها یک گام به جلو می‌رویم اما در قسمت پیش‌بینی پس از یک پیش‌بینی ۴۰ روزه، پیش‌بینی ۴۰ روز بعدی را انجام می‌دهیم، بنابراین بین پیش‌بینی‌ها ۴۰ روز فاصله است و داده‌های جدید وارد شبکه می‌شوند. شکل ۶ چگونگی انجام محاسبات را بیان می‌کند. پیش‌بینی‌های انجام شده در این پژوهش برای بازه کوتاه مدت ۴۰ روزه است. ورودی شبکه‌ها در قسمت آموزش هم بردارهای ۴۰ روزه است و در هر گام آموزش برای درک بهتر روابط و آموزش طولانی‌تر در قسمت آموزش تنها یک گام به جلو می‌رویم اما در قسمت پیش‌بینی پس از یک پیش‌بینی ۴۰ روزه، پیش‌بینی ۴۰ روز بعدی را انجام می‌دهیم، بنابراین بین پیش‌بینی‌ها ۴۰ روز فاصله است و داده‌های جدید وارد شبکه می‌شوند. شکل ۶ چگونگی انجام محاسبات را بیان می‌کند.



شکل ۶- نحوه وارد کردن داده‌ها به شبکه‌های عصبی ژرف

اصلی انجام می‌شود و خروجی‌های این لایه به لایه‌های بعدی ارسال می‌شوند تا به لایه خروجی نهایی می‌رسد. رابطه (۸) نشان دهنده چگونگی انجام محاسبات داخل یک این شبکه است. در اینجا w_i وزن، x_i ورودی و b بایاس است. f یک تابع فعال‌سازی غیرخطی است که برای یادگیری الگوهای پیچیده استفاده می‌شود. شکل ۴ ساختار یک نورون مصنوعی معمولی را نشان می‌دهد.

$$y = f \left(b + \sum_{i=1}^n x_i w_i \right) \quad (8)$$



شکل ۴- چگونگی انجام محاسبات در یک نورون شبکه عصبی

در این پژوهش از سه نوع شبکه گوناگون LSTM، CNN، MLP برای پیش‌بینی حرکت قطبی استفاده شده است و در پایان جواب‌ها مقایسه شده‌اند. شکل ۵ فلوجارت نحوه انجام محاسبات را نشان می‌دهد. همان طور که مشخص است، ابتدا داده‌ها به دو قسمت آموزش^۱ و قسمت پیش‌بینی یا آزمون^۲ تقسیم شده است. برای قسمت آموزش داده‌های روزانه پارامترهای x و y مربوط به بازه ۱ ژانویه ۱۹۸۰ تا ۳۱ دسامبر ۲۰۱۵ (۳۵ سال) استفاده شده است. داده‌های بازه پیش‌بینی هم مربوط به بازه ۱ ژانویه ۲۰۱۶ تا ۳۱ دسامبر ۲۰۲۰ (۵ سال) است. در ادامه، شبکه‌ها فرایند یادگیری را آغاز می‌کنند و پس از آموزش‌های پی در پی و در پایان هر مرحله، مقدار تابع زیان^۳ با معیارها مربوطه سنجیده می‌شود، در پایان آموزش هم بهترین مدل به دست آمده برای قسمت پیش‌بینی استفاده می‌شود. پیش‌بینی‌های به دست آمده با معیار خطای جذر میانگین مربعات (RMSE^۴) سنجیده می‌شوند که در ادامه بیان می‌شود.

^۱ Train
^۲ Test
^۳ Loss Function
^۴ Root mean square error

۲-۲- روش آنالیز هارمونیک کمترین مربعات (LSHA)

روش LSHA توسط امیری سیمکویی در سال ۲۰۰۷ معرفی شد. این روش ابتدا برای تشخیص فرکانس‌های موجود در سری‌های زمانی GPS ارائه شد. نتایج نشان داد که این روش می‌تواند با دقت بالایی فرکانس‌های موجود در این سری زمانی را استخراج نماید. در مطالعات بعدی، این روش بر روی سری‌های زمانی پدیده‌های مختلف ژئوفیزیکی پیاده شد [۳۷] [۳۸]. اساس این روش تکمیل مدل تابعی بر اساس آزمون‌های آماری است. مدل تابعی در سری‌های زمانی مطابق رابطه (۹) تعریف می‌شود. در مدل تابعی، ماتریس طرح اولیه A ، از دو ستون مربوط به ترم آفست و ترندخطی تشکیل شده است که در رابطه (۱۰) معرفی شده‌اند. همچنین در مدل تابعی ماتریس A_k که مرتبط با فرکانس ω_k است به صورت رابطه (۱۱) تعریف می‌شود.

$$y = Ax + \sum_{k=1}^q A_k x_k \quad (9)$$

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \quad (10)$$

$$A_k = \begin{bmatrix} \cos(\omega_k t_1) & \sin(\omega_k t_1) \\ \cos(\omega_k t_2) & \sin(\omega_k t_2) \\ \vdots & \vdots \\ \cos(\omega_k t_m) & \sin(\omega_k t_m) \end{bmatrix} \quad (11)$$

$$x = \begin{bmatrix} y_0 \\ v \end{bmatrix}, \quad x_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix} \quad (12)$$

در این روش، آزمون آماری برای فرکانس‌های مختلف بررسی می‌شود و فرکانسی معنادار است که در این آزمون آماری پذیرفته شود.

$$H_0 = Ax + \sum_{k=1}^{i-1} A_k x_k \quad \text{فرض صفر} \quad (13)$$

$$H_a = Ax + \sum_{k=1}^i A_k x_k \quad \text{فرض مقابل} \quad (14)$$

در آزمون آماری که توسط روابط (۱۳) و (۱۴) تعریف می‌شود، رد شدن فرض صفر نشان دهنده وجود فرکانس ω_k در سری زمانی مورد بررسی است. با پذیرفته شدن هر ω_k ، مولفه‌های سینوسی و کسینوسی مربوط به آن

فرکانس به فرض صفر اضافه شده و آزمون برای فرکانس‌های بعدی تکرار خواهد شد. قابل ذکر است که در هر مرحله، فرکانس ω_k (با کمترین نرم بردار باقیمانده‌ها تحت فرض مقابل) به مدل تابعی افزوده می‌شود. این فرکانس از مقدار آرگومان بیشینه طیف توان رابطه (۱۵) برای مقادیر مختلف ω_j ، حاصل می‌شود (رابطه ۱۶).

$$P(\omega_j) = \hat{\epsilon}_0^T Q_y^{-1} A_j (A_j^T Q_y^{-1} p_A^T A_j)^{-1} A_j^T Q_y^{-1} \hat{\epsilon}_0 \quad (15)$$

$$\omega_k = \arg \max P(\omega_j) \quad (16)$$

در روابط فوق

$$\hat{\epsilon}_0 = p_A^T y \quad (17)$$

$$p_A^T = I - A(A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} \quad (18)$$

هستند و A_j نیز مطابق با رابطه (۱۱) تشکیل می‌شود. همچنین ماتریس Q_y در این تحقیق یک ماتریس یکه در نظر گرفته شده است [۳۷] [۳۸].

۳- نتایج عددی

در شبکه‌های عصبی از پارامترهای گوناگونی استفاده می‌شود که روی عملکرد و دقت شبکه‌ها تاثیر مستقیم دارد. در هر ۳ نوع شبکه استفاده شده در این پژوهش، ما از تابع فعالسازی ReLU ($y = \max(0, x)$) برای درک روابط غیرخطی استفاده شد، این تابع فعالسازی سرعت پردازش را افزایش می‌دهد و از محوشدگی گرادینان جلوگیری می‌کند [۳۹]. نرخ یادگیری برابر با ۰/۰۰۱ تا ۰/۰۰۱ است. به‌منظور دستیابی به میزان دقت بالاتر، باید از انتخاب مقادیر بزرگ برای نرخ یادگیری اجتناب کنیم، همچنین استفاده از مقادیر بسیار کوچک برای نرخ یادگیری سرعت شبکه را بسیار کم می‌کند و زمان زیادی هدر می‌رود [۴۰]. برای تابع بهینه‌ساز هم از ADAM استفاده نمودیم. این تابع برای مجموعه داده‌های بزرگ و نویزی مناسب است زیرا سرعت شبکه را افزایش می‌دهد و از لحاظ محاسباتی بهینه است. همینطور استفاده از این بهبود دهنده نیاز به صرف زمان در تنظیم پارامترهای دیگر را کاهش می‌دهد [۴۱]. تعداد اپک‌های آموزش هم برابر با ۱۰۰۰ است. در ادامه نتایج شبکه‌های ژرف استفاده شده درون جداول ۳ و ۴ نشان داده شده است. باید توجه داشت که در این جداول

جدول ۲- معماری‌های مختلف شبکه‌های استفاده شده

تعداد	شبکه‌ها
۱	LSTM (۸۰)
۲	LSTM (۸۰) - LSTM (۶۰)
۳	CNN (۸۰ (f۲))
۴	CNN (۸۰ (f۳))
۵	CNN (۸۰ (f۵))
۶	CNN (۸۰ (f۲)) - CNN (۶۰ (f۳))
۷	MLP (۸۰)
۸	MLP (۸۰) - MLP (۶۰)

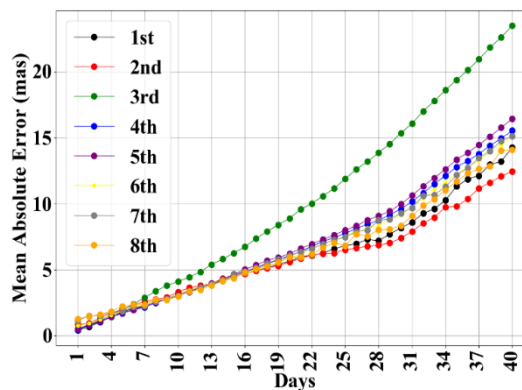
به طور کلی از جدول‌های ۳ و ۴ مشخص است که با طولانی شدن مدت پیش‌بینی طبیعتاً دقت‌ها کاهش یافته است، این مورد در تمامی روش‌ها مشهود است. برای مثال در شکل‌های ۹ و ۱۰ دوره ۱۸ ام پیش‌بینی‌ها نسبت به مقادیر واقعی نشان داده شده است و در این شکل مشخص است که با گذشت زمان مقادیر پیش‌بینی از مقدار واقعی فاصله گرفته است. همچنین طبق جدول‌های ۳ و ۴ مشخص است که به ازای تعداد لایه‌ها و پارامترهای گوناگون شبکه‌ها، دقت‌ها و جواب‌ها متفاوت است و حتی در برخی مواقع تفاوت دقت‌ها بسیار بزرگ است. برای مثال در شبکه سوم CNN با فیلتر ۲، دقت در روزهای پایانی بسیار بدتر از روزهای دیگر است، در حالی‌که در روزهای ابتدایی این شبکه دارای دقت بسیار خوبی است. این مورد در شکل ۷ به طور کامل قابل مشاهده است. در این شکل فاصله زیاد این شبکه از تمامی شبکه‌ها در روزهای نهایی واضح است (خط سبز رنگ). از جدول‌های ۳ و ۴ همچنین مشخص است که اگر شبکه‌ها ژرف‌تر از شبکه‌های دیگر باشند، لزوماً به این معنا نیست که دقت‌ها بهبود می‌یابند بلکه ما نیاز داریم تا معماری‌های متفاوتی از شبکه عصبی را امتحان کنیم و جواب‌ها را با یکدیگر مقایسه کنیم. برای مثال از شکل‌های ۷ و ۸ مشخص است که شبکه آخر که یک شبکه دو لایه پرسپترون است، در بسیاری از روزها دقت بدتری نسبت به شبکه MLP تک لایه دارد، همچنین شبکه ششم که یک شبکه CNN دو لایه است، نسبت به شبکه‌های CNN یک لایه دقت بدتری داشته است، بنابراین لزوماً شبکه ژرف‌تر دقت بهتری ندارد.

برای پارامتر x از جدول ۳ و از خط قرمز رنگ در شکل‌های ۷ و ۸ مشخص است که شبکه LSTM دو لایه یعنی شبکه دوم تقریباً دقت بهتری نسبت به شبکه‌های LSTM تک لایه و شبکه‌های دیگر دارد. در جدول ۲ ساختار

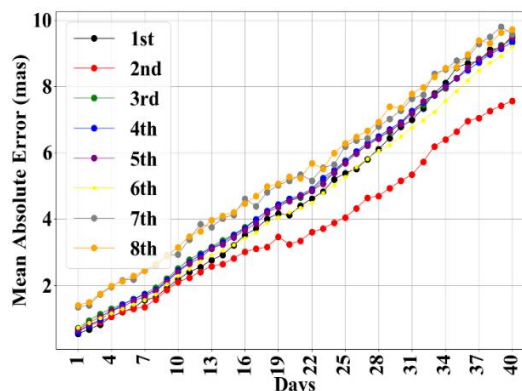
نتایج بر اساس مقادیر RMSE دوره‌های ۴۰ روزه گوناگون بیان شده است. هم چنین برای درک بهتر تغییرات، مقادیر خطای میانگین قدرمطلق (Mean Absolute Error) هم در اشکال ۷ و ۸ برای هر کدام از پارامترها آورده شده است. در ادامه، روابط (۱۹) و (۲۰) چگونگی محاسبه MAE و RMSE را نشان می‌دهد. در این رابطه‌ها n نشان دهنده تعداد دوره‌ها برای بازه p روزه است و اولین روز هر دوره پیش‌بینی با یکدیگر متفاوت است. همچنین y_i مقدار واقعی و \tilde{y}_i مقدار پیش‌بینی شده است.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i| \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (20)$$



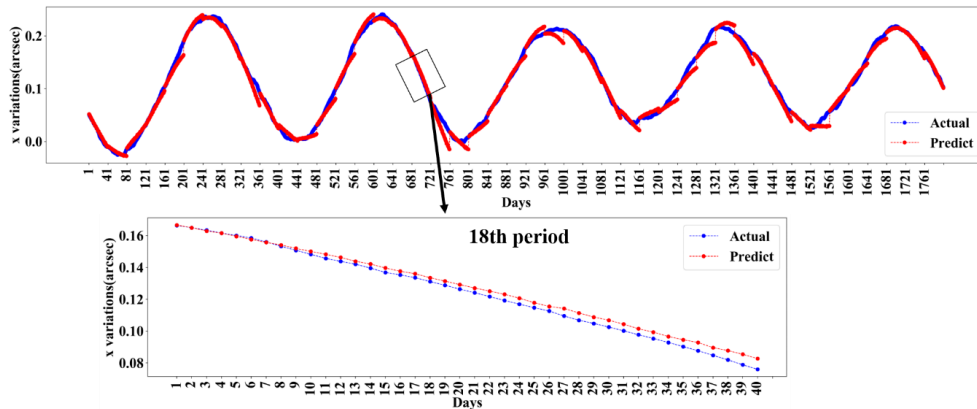
شکل ۷- مقادیر MAE پارامتر x برای ۴۰ روز



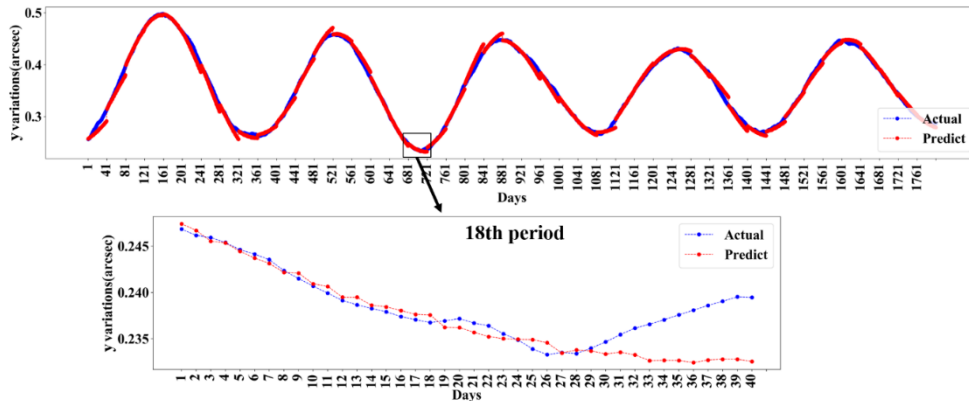
شکل ۸- مقادیر MAE پارامتر y برای ۴۰ روز

در جدول ۲ ساختار هر شبکه به همراه تعداد واحدهای محاسباتی (یا سلول) و لایه‌ها بیان شده است. برای مثال در شبکه ششم، ۲ لایه CNN با فیلترهای ۲ و ۳ و به ترتیب ۸۰ و ۶۰ واحد محاسباتی و یا شبکه دوم از ۲ لایه LSTM با به ترتیب ۸۰ و ۶۰ واحد محاسباتی ساخته شده است.

می‌شود (خط قرمز). برای این شبکه مقدار کمیته RMSE مربوط به روز اول و برابر با 0.076 mas است و مقدار بیشینه RMSE برای روز ۴۰ ام و برابر با 0.229 mas است که این مقدار بسیار کمتر از شبکه‌های دیگر است. برای روز اول کمترین مقادیر خطا مربوط به شبکه اول یا LSTM تک لایه است. مقدار کمیته RMSE برای روز اول در این شبکه برابر با 0.068 mas است که از شبکه LSTM دو لایه کمتر است. در مورد شبکه‌های CNN هم شبکه‌های چهارم و پنجم که به ترتیب فیلترهای ۳ و ۵ دارد، دقت‌های مشابهی دارند. در این مورد هم شبکه‌های MLP دقت بسیار بدتری نسبت به شبکه‌های دیگر دارد. یکی از برتری‌های شبکه‌های CNN مدت زمان کمتر آموزش این شبکه‌ها در هر اپک است. در رایانه‌ای که محاسبات این پژوهش صورت گرفت، مدت زمان آموزش شبکه CNN در هر اپک برابر با ۲ ثانیه و در شبکه‌های LSTM در هر اپک حدوداً ۶ ثانیه است. بنابراین آموزش شبکه‌های LSTM حدوداً یک ساعت و چهل دقیقه طول کشید. در ادامه برای انجام مقایسه و سنجش روش ارائه شده، الگوریتم LSHA بر روی سری‌های زمانی x و y پیاده شد. جدول ۵ دوره تناوب‌های کشف شده از الگوریتم LSHA برای سری‌های زمانی x و y را برحسب روز نشان می‌دهد.



شکل ۹- پیش‌بینی‌ها نسبت به مقادیر واقعی پارامتر x . در پایین برای نمونه دوره ۱۸ام پیش‌بینی آورده شده است



شکل ۱۰- پیش‌بینی‌ها نسبت به مقادیر واقعی پارامتر y . در پایین برای نمونه دوره ۱۸ام پیش‌بینی آورده شده است

این شبکه مشخص است، این شبکه دارای دو لایه LSTM با به ترتیب ۸۰ و ۶۰ واحد محاسباتی است و از روز ۱۸ام به بعد دقت بسیار بهتری نسبت به شبکه‌های دیگر، یعنی شبکه CNN و شبکه‌های دیگر دارد، البته برای روزهای ابتدایی هم دقت این شبکه خوب است. این مورد در شکل ۷ ام به خوبی قابل مشاهده است (خط قرمز رنگ). در این شبکه برای پارامتر x ، مقدار کمیته RMSE برای روز اول برابر با 0.076 mas است و مقدار بیشینه RMSE در روز ۴۰ ام برابر با 0.229 mas است. ذکر این نکته ضروری است که در روز ۴۰ام، این شبکه بسیار بهتر از شبکه‌های دیگر عمل کرده است. در مقابل در روزهای ابتدایی شبکه‌های LSTM تک لایه دقت بهتری دارند، برای مثال در این شبکه در روز اول، مقدار RMSE برابر با 0.049 mas است. دقت این شبکه تا روز ۱۸ام بهتر از شبکه LSTM دو لایه است. در مورد شبکه‌های CNN هم مشخص است که شبکه CNN با فیلتر ۳ بهتر از شبکه‌های دیگر CNN با فیلترهای ۲ یا ۵ است و برای روزهای ابتدایی این مقادیر خطا حتی کمتر از شبکه دوم یا LSTM دو لایه است. در رابطه با پارامتر y هم شبکه دوم یا همان LSTM دو لایه از روز ۴ام تا انتها دقت بهتری نسبت به شبکه‌های دیگر دارد این مورد در شکل ۸ به خوبی دیده

جدول ۳- مقادیر RMSE پیش‌بینی حرکت قطب توسط شبکه‌های عصبی برای پارامتر x بر حسب mas

روز	شبکه اول	شبکه دوم	شبکه سوم	شبکه چهارم	شبکه پنجم	شبکه ششم	شبکه هفتم	شبکه هشتم
۱	۰/۴۹	۱/۰۲	۰/۶۹	۰/۵۷	۰/۶۲	۰/۹۵	۱/۳۹	۱/۶۵
۲	۰/۸۸	۱/۱۲	۱/۰۸	۰/۹۷	۱/۰۳	۱/۱۷	۱/۸۰	۱/۹۲
۳	۱/۳۴	۱/۵۸	۱/۵۴	۱/۳۹	۱/۴۳	۱/۵۰	۱/۹۴	۲/۲۱
۴	۱/۷۷	۲/۰۴	۲/۰۴	۱/۸۱	۱/۸۵	۱/۸۹	۲/۳۱	۲/۳۲
۵	۲/۱۷	۲/۴۱	۲/۵۶	۲/۱۴	۲/۲۱	۲/۲۲	۲/۲۹	۲/۷۵
۶	۲/۵۳	۲/۷۲	۲/۹۶	۲/۳۹	۲/۴۴	۲/۵۰	۲/۶۹	۲/۸۱
۷	۲/۷۹	۳/۰۵	۳/۴۸	۲/۶۸	۲/۷۴	۲/۸۳	۲/۸۰	۲/۸۳
۸	۳/۲۹	۳/۴۶	۴/۰۰	۳/۰۷	۳/۱۱	۳/۱۲	۳/۲۱	۳/۳۴
۹	۳/۵۴	۳/۷۴	۴/۵۱	۳/۳۹	۳/۴۶	۳/۴۳	۳/۳۷	۳/۵۱
۱۰	۳/۸۶	۴/۰۵	۴/۹۰	۳/۷۴	۳/۷۹	۳/۷۶	۳/۶۹	۳/۸۳
۱۱	۴/۱۹	۴/۴۵	۵/۳۶	۴/۰۷	۴/۱۵	۴/۱۵	۴/۱۶	۴/۲۵
۱۲	۴/۵۳	۴/۶۴	۵/۸۰	۴/۴۳	۴/۴۹	۴/۴۸	۴/۴۶	۴/۴۴
۱۳	۴/۸۲	۴/۹۳	۶/۴۰	۴/۷۶	۴/۸۶	۴/۸۱	۴/۹۲	۴/۷۷
۱۴	۵/۱۷	۵/۳۳	۶/۸۸	۵/۱۲	۵/۲۰	۵/۱۳	۵/۱۹	۵/۱۶
۱۵	۵/۴۴	۵/۴۸	۷/۴۳	۵/۴۹	۵/۵۷	۵/۴۵	۵/۵۹	۵/۴۸
۱۶	۵/۷۲	۵/۷۶	۸/۰۰	۵/۸۵	۵/۹۵	۵/۷۳	۵/۹۳	۵/۷۸
۱۷	۶/۰۲	۵/۹۴	۸/۶۵	۶/۲۳	۶/۳۵	۶/۱۷	۶/۲۷	۶/۱۹
۱۸	۶/۳۰	۶/۲۰	۹/۲۴	۶/۶۲	۶/۷۹	۶/۴۸	۶/۶۹	۶/۵۲
۱۹	۶/۶۵	۶/۴۶	۹/۷۸	۶/۹۸	۷/۱۰	۶/۸۲	۷/۰۴	۶/۸۸
۲۰	۶/۹۲	۶/۶۹	۱۰/۳۶	۷/۳۶	۷/۴۸	۷/۳۱	۷/۳۲	۷/۳۵
۲۱	۷/۳۲	۶/۹۲	۱۱/۱۴	۷/۸۶	۷/۹۵	۷/۶۱	۷/۶۹	۷/۳۱
۲۲	۷/۴۵	۷/۳۱	۱۱/۷۳	۸/۲۸	۸/۳۹	۸/۰۷	۸/۰۲	۷/۶۱
۲۳	۷/۸۲	۷/۴۲	۱۲/۳۳	۸/۶۹	۸/۷۸	۸/۳۷	۸/۴۸	۸/۳۶
۲۴	۸/۲۲	۷/۶۶	۱۲/۹۴	۹/۰۷	۹/۲۲	۸/۷۷	۸/۹۱	۸/۷۸
۲۵	۸/۶۴	۸/۲۲	۱۳/۶۸	۹/۵۲	۹/۶۶	۹/۲۶	۹/۰۹	۸/۵۸
۲۶	۸/۹۵	۸/۴۰	۱۴/۴۳	۱۰/۰۱	۱۰/۱۵	۹/۷۲	۹/۶۷	۹/۶۶
۲۷	۹/۴۹	۸/۵۶	۱۵/۰۶	۱۰/۴۳	۱۰/۶۵	۱۰/۰۴	۹/۸۲	۹/۶۰
۲۸	۹/۴۹	۸/۸۷	۱۵/۷۴	۱۰/۸۰	۱۱/۰۳	۱۰/۳۷	۱۰/۶۷	۱۰/۱۸
۲۹	۱۰/۰۸	۹/۱۹	۱۶/۴۰	۱۱/۲۵	۱۱/۴۴	۱۰/۸۸	۱۰/۸۴	۱۰/۲۴
۳۰	۱۰/۵۵	۹/۶۳	۱۷/۲۰	۱۱/۷۱	۱۱/۹۷	۱۱/۲۲	۱۱/۳۷	۱۰/۶۱
۳۱	۱۰/۹۹	۱۰/۱۳	۱۷/۹۰	۱۲/۲۶	۱۲/۶۱	۱۱/۷۹	۱۱/۸۲	۱۱/۲۸
۳۲	۱۱/۶۸	۱۰/۸۶	۱۸/۹۰	۱۲/۹۳	۱۳/۳۳	۱۲/۵۰	۱۲/۸۳	۱۲/۰۶
۳۳	۱۲/۱۴	۱۱/۳۴	۱۹/۷۷	۱۳/۶۸	۱۴/۰۵	۱۳/۲۹	۱۲/۹۸	۱۲/۷۲
۳۴	۱۲/۹۰	۱۲/۴۳	۲۰/۷۲	۱۴/۴۳	۱۴/۸۰	۱۳/۹۷	۱۳/۷۲	۱۳/۴۲
۳۵	۱۳/۹۷	۱۲/۳۷	۲۱/۶۰	۱۵/۲۰	۱۵/۶۸	۱۴/۷۴	۱۴/۶۹	۱۴/۱۰
۳۶	۱۴/۴۵	۱۲/۹۹	۲۲/۴۲	۱۵/۷۳	۱۶/۲۶	۱۵/۳۰	۱۵/۲۶	۱۴/۷۲
۳۷	۱۴/۷۰	۱۳/۷۰	۲۳/۲۷	۱۶/۳۱	۱۶/۹۹	۱۵/۸۱	۱۵/۹۴	۱۵/۲۳
۳۸	۱۵/۵۶	۱۴/۱۳	۲۴/۱۹	۱۶/۹۴	۱۷/۵۶	۱۶/۴۴	۱۶/۴۶	۱۵/۳۶
۳۹	۱۵/۸۰	۱۴/۵۸	۲۴/۹۸	۱۷/۴۶	۱۸/۲۳	۱۷/۰۱	۱۷/۱۹	۱۶/۳۹
۴۰	۱۶/۹۳	۱۵/۰۵	۲۵/۸۹	۱۸/۰۶	۱۸/۹۳	۱۷/۶۱	۱۷/۶۴	۱۶/۶۰

جدول ۴- مقادیر RMSE پیش‌بینی حرکت قطب توسط شبکه‌های عصبی برای پارامتر γ بر حسب mas

روز	شبکه اول	شبکه دوم	شبکه سوم	شبکه چهارم	شبکه پنجم	شبکه ششم	شبکه هفتم	شبکه هشتم
۱	۰/۶۸	۰/۷۶	۰/۸۹	۰/۷۵	۰/۸۱	۰/۸۹	۱/۶۸	۱/۸۶
۲	۰/۸۶	۱/۰۷	۱/۱۵	۱/۰۱	۱/۰۷	۱/۱۲	۱/۷۶	۱/۸۸
۳	۱/۰۲	۱/۱۹	۱/۲۵	۱/۲۲	۱/۲۸	۱/۲۷	۲/۱۳	۲/۱۱
۴	۱/۳۱	۱/۳۷	۱/۵۳	۱/۴۵	۱/۵۰	۱/۴۴	۲/۲۸	۲/۳۳
۵	۱/۴۶	۱/۵۶	۱/۷۱	۱/۶۷	۱/۶۶	۱/۵۸	۲/۶۱	۲/۵۰
۶	۱/۶۰	۱/۶۲	۱/۸۹	۱/۸۵	۱/۸۵	۱/۷۲	۲/۵۹	۲/۷۲
۷	۱/۸۹	۱/۷۴	۲/۱۲	۲/۰۹	۲/۰۷	۱/۹۶	۲/۹۳	۲/۹۴
۸	۲/۰۷	۱/۹۵	۲/۴۳	۲/۳۶	۲/۳۶	۲/۲۱	۳/۱۵	۳/۲۲
۹	۲/۴۸	۲/۳۲	۲/۸۲	۲/۷۸	۲/۷۴	۲/۶۳	۳/۶۰	۳/۵۸
۱۰	۲/۸۳	۲/۶۰	۳/۱۸	۳/۱۳	۳/۰۹	۲/۹۵	۳/۶۷	۳/۹۳
۱۱	۳/۰۷	۲/۷۸	۳/۴۸	۳/۴۳	۳/۳۸	۳/۱۷	۴/۱۸	۴/۳۱
۱۲	۳/۳۵	۳/۰۹	۳/۷۲	۳/۷۵	۳/۶۵	۳/۴۸	۴/۷۵	۴/۵۲
۱۳	۳/۶۴	۳/۳۳	۴/۰۳	۴/۰۵	۴/۰۰	۳/۷۶	۴/۷۱	۵/۰۳
۱۴	۳/۸۷	۳/۴۶	۴/۳۰	۴/۲۹	۴/۲۱	۳/۹۸	۵/۰۶	۵/۲۶
۱۵	۴/۱۰	۳/۷۶	۴/۵۳	۴/۵۴	۴/۴۷	۴/۲۰	۵/۲۶	۵/۲۹
۱۶	۴/۴۷	۴/۰۱	۴/۸۱	۴/۸۳	۴/۷۷	۴/۵۲	۵/۷۶	۵/۷۱
۱۷	۴/۷۴	۴/۱۹	۵/۱۲	۵/۱۸	۵/۰۶	۴/۸۰	۵/۶۱	۵/۹۸
۱۸	۵/۱۲	۴/۳۸	۵/۴۲	۵/۴۶	۵/۴۱	۵/۱۲	۶/۱۲	۶/۳۱
۱۹	۵/۳۱	۴/۷۸	۵/۶۴	۵/۶۸	۵/۶۰	۵/۲۹	۶/۳۲	۶/۴۲
۲۰	۵/۳۱	۴/۷۲	۵/۷۹	۵/۸۵	۵/۷۸	۵/۴۳	۶/۵۰	۶/۶۶
۲۱	۵/۶۲	۴/۸۸	۵/۹۱	۵/۹۷	۵/۹۱	۵/۵۶	۶/۶۹	۶/۶۵
۲۲	۵/۸۶	۵/۱۳	۶/۱۲	۶/۱۵	۶/۰۷	۵/۷۵	۶/۴۷	۷/۰۹
۲۳	۵/۹۹	۵/۲۸	۶/۴۱	۶/۴۴	۶/۲۷	۶/۰۱	۶/۹۳	۶/۹۵
۲۴	۶/۳۴	۵/۴۲	۶/۵۴	۶/۶۰	۶/۵۵	۶/۱۹	۶/۹۰	۷/۳۴
۲۵	۶/۵۲	۵/۵۷	۶/۸۸	۶/۸۳	۶/۷۹	۶/۳۶	۷/۴۳	۷/۵۵
۲۶	۶/۶۵	۵/۶۸	۷/۱۰	۷/۱۲	۷/۰۶	۶/۶۹	۷/۵۸	۷/۷۵
۲۷	۶/۹۸	۶/۰۲	۷/۳۳	۷/۳۶	۷/۳۱	۶/۹۴	۷/۶۵	۷/۹۳
۲۸	۷/۲۵	۶/۱۳	۷/۶۱	۷/۵۹	۷/۵۲	۷/۱۶	۸/۰۲	۸/۲۳
۲۹	۷/۵۲	۶/۲۸	۷/۸۰	۷/۸۰	۷/۷۶	۷/۳۵	۸/۲۳	۸/۶۹
۳۰	۷/۹۱	۶/۵۰	۸/۱۰	۸/۰۵	۸/۰۲	۷/۶۸	۸/۴۸	۸/۶۶
۳۱	۸/۱۸	۶/۷۱	۸/۴۱	۸/۴۰	۸/۴۱	۸/۰۱	۸/۹۱	۹/۰۹
۳۲	۸/۵۲	۷/۱۰	۸/۷۱	۸/۷۴	۸/۸۴	۸/۳۲	۹/۰۶	۹/۳۲
۳۳	۹/۰۸	۷/۶۰	۹/۰۹	۹/۰۲	۹/۱۱	۸/۶۴	۹/۷۳	۹/۶۹
۳۴	۹/۳۸	۷/۷۹	۹/۳۵	۹/۳۳	۹/۳۷	۸/۹۷	۹/۹۴	۹/۹۸
۳۵	۹/۹۲	۸/۰۶	۹/۷۰	۹/۶۷	۹/۷۳	۹/۳۲	۱۰/۲۱	۱۰/۰۹
۳۶	۱۰/۱۱	۸/۴۰	۱۰/۱۰	۱۰/۰۰	۱۰/۱۷	۹/۷۴	۱۰/۴۱	۱۰/۵۴
۳۷	۱۰/۲۵	۸/۵۴	۱۰/۴۲	۱۰/۳۰	۱۰/۵۰	۱۰/۱۰	۱۰/۹۴	۱۰/۹۸
۳۸	۱۰/۷۰	۸/۸۲	۱۰/۶۴	۱۰/۶۱	۱۰/۷۳	۱۰/۳۸	۱۱/۱۵	۱۰/۹۵
۳۹	۱۰/۸۰	۹/۰۵	۱۰/۹۶	۱۰/۸۳	۱۰/۹۶	۱۰/۶۱	۱۱/۴۸	۱۱/۳۱
۴۰	۱۱/۳۱	۹/۲۲	۱۱/۲۴	۱۱/۱۰	۱۱/۳۴	۱۰/۹۶	۱۱/۲۸	۱۱/۵۲

جدول ۵ - دوره تناوب‌های کشف شده از الگوریتم LSHA برای سری‌های زمانی x و y برحسب روز

فرکانس‌های x	فرکانس‌های y
۳۰/۳۳۰	۳۰/۳۳۰
۴۵/۱۲۰	۴۵/۱۲۰
۵۹/۹۴۰	۵۹/۹۴۰
۹۰/۴۴۰	۹۰/۴۴۰
۱۲۲/۳۷۰	۱۲۲/۳۷۰
۱۸۲/۶۳۰	۱۸۲/۶۳۰
۲۳۳/۸۹۰	۲۳۳/۸۹۰
۲۹۶/۱۰۰۰	۲۹۶/۱۰۰۰
۳۱۴/۲۰۰	۳۱۴/۲۰۰
۳۴۷/۲۳۰	۳۴۷/۲۳۰
۳۶۵/۲۵۰	۳۶۵/۲۵۰
۴۰۶/۴۵۰	۴۰۶/۴۵۰
۴۳۳/۰۶۰	۴۳۳/۰۶۰
۴۸۴/۴۰۰	۴۸۴/۴۰۰
۷۳۱/۷۰۰	۷۳۱/۷۰۰
۱۳۰۹/۰۰۰	۱۳۰۹/۰۰۰

جدول ۶- مقادیر RMSE پیش‌بینی حرکت قطب توسط روش LSHA برای پارامتر x و y بر حسب mas

روز	x	y	روز	x	y
۱	۴۰/۹۵	۴۰/۸۶	۲۱	۴۵/۴۱	۴۴/۴۰
۲	۴۱/۲۱	۴۱/۱۳	۲۲	۴۵/۵۲	۴۴/۵۵
۳	۴۱/۴۶	۴۱/۳۶	۲۳	۴۵/۵۳	۴۴/۶۵
۴	۴۱/۷۸	۴۱/۹۲	۲۴	۴۵/۵۵	۴۴/۷۷
۵	۴۲/۱۷	۴۲/۱۴	۲۵	۴۵/۶۳	۴۴/۹۳
۶	۴۲/۵۳	۴۲/۲۶	۲۶	۴۵/۷۵	۴۵/۰۹
۷	۴۳/۰۲	۴۲/۴۲	۲۷	۴۵/۹۴	۴۵/۲۶
۸	۴۳/۱۵	۴۲/۶۲	۲۸	۴۶/۱۷	۴۵/۳۷
۹	۴۳/۳۳	۴۲/۷۹	۲۹	۴۶/۴۱	۴۵/۴۷
۱۰	۴۳/۵۸	۴۲/۹۹	۳۰	۴۶/۶۱	۴۵/۵۰
۱۱	۴۳/۸۶	۴۳/۱۶	۳۱	۴۶/۸۴	۴۵/۶۵
۱۲	۴۴/۱۴	۴۳/۲۹	۳۲	۴۶/۹۹	۴۵/۷۳
۱۳	۴۴/۴۱	۴۳/۴۰	۳۳	۴۷/۰۸	۴۵/۸۱
۱۴	۴۴/۴۷	۴۳/۴۶	۳۴	۴۷/۲۰	۴۵/۹۳
۱۵	۴۴/۶۳	۴۳/۵۵	۳۵	۴۷/۳۴	۴۶/۱۲
۱۶	۴۴/۸۰	۴۳/۶۱	۳۶	۴۷/۵۱	۴۶/۲۴
۱۷	۴۴/۹۴	۴۳/۷۰	۳۷	۴۷/۷۰	۴۶/۳۴
۱۸	۴۵/۰۷	۴۳/۸۰	۳۸	۴۷/۹۶	۴۶/۳۸
۱۹	۴۵/۱۸	۴۳/۹۷	۳۹	۴۸/۲۳	۴۶/۴۷
۲۰	۴۵/۲۸	۴۴/۲۱	۴۰	۴۸/۴۰	۴۶/۵۳

همانطور که از جدول ۶ مشخص است، در روش آنالیز هارمونیک کمترین مربعات مقدار RMSE برای پارامتر x برای روزهای اول و چهارم به ترتیب برابر با ۴۰/۹۵ و ۴۸/۴۰ میلی ثانیه‌ی کمانی است. همچنین مقدار RMSE پارامتر y برای روزهای اول و چهارم به ترتیب برابر با ۴۰/۸۶ و ۴۶/۵۳ میلی ثانیه‌ی کمانی است. همانطور که در تمامی روش‌های ارائه شده می‌توان دید، پیش‌بینی‌های پارامتر y از پارامتر x دقیق‌تر است. به طور کلی باید گفت دلیل اصلی برتری LSTM نسبت به شبکه‌های دیگر مانند CNN این است که شبکه‌های بازگشتی نوعی از شبکه‌های عصبی مصنوعی هستند که مجهز به نوعی حافظه کوتاه مدت هستند و این ویژگی این شبکه‌ها را برای کار با دنباله‌ها و سری‌های زمانی کارآمد ساخته است. در این شبکه‌ها در هر مرحله ورودی‌ها شامل نوعی حافظه هستند که وابستگی‌های کوتاه مدت بین دنباله‌ها را در ورودی‌ها شامل می‌شوند. در شبکه‌های LSTM تا حدی مشکل ناپدید شدن گرادینان حل شده است زیرا واحدهای LSTM به جریان گرادینان این امکان را می‌دهند که بدون تغییر بماند [۲۸] [۴۲]. شبکه CNN ویژگی‌های مهم داده‌های ورودی را با ضرب کانولوشنی بدست می‌آورد. هر لایه پیچیدگی‌های داده‌های ورودی را درک می‌کند و نقشه‌های ویژگی^۱ را به عنوان خروجی تولید می‌کند. این نقشه‌های ویژگی حاوی الگوهای شناخته شده ای است که به عنوان ورودی لایه بعدی استفاده می‌شود. در نتیجه، شبکه‌های CNN که عمدتاً برای برنامه‌های بینایی کامپیوتری طراحی شده‌اند، بیشتر و بیشتر در پیش‌بینی مبتنی بر سری زمانی استفاده می‌شوند [۳۶]. باید اضافه کرد که روش LSHA یک روش پارامتریک است که در آن دامنه و فاز فرکانس‌ها، ثابت در نظر گرفته می‌شوند. علت افت دقت در این روش نسبت به روش یادگیری عمیق این است که در روش یادگیری عمیق دامنه و فاز فرکانس‌ها ثابت نیستند و این روش می‌تواند رفتار غیر پارامتریک سری‌های زمانی x و y را به خوبی مدل‌سازی نماید همچنین روش آنالیز هارمونیک کمترین مربعات فقط رفتار تناوبی موجود در سری زمانی تحت بررسی را ارزیابی و پیش‌بینی می‌کند بنابراین رفتار غیر تناوبی در این روش بررسی نمی‌شود در حالی که در روش‌های غیر پارامتریک، رفتار غیر تناوبی نیز ارزیابی و پیش‌بینی می‌شود. [۳۶] [۳۷]

^۱ Feature Map

هارمونیک کمترین مربعات برای پیش‌بینی هر چه دقیق‌تر حرکت قطبی استفاده شده است. در این پژوهش از داده‌های ۳۵ سال برای آموزش شبکه‌های عصبی استفاده شد و پیش‌بینی‌های ۴۰ روزه از ابتدای ژانویه ۲۰۱۶ انجام گرفت و به دلیل وجود پارمترها و لایه‌های گوناگون، تلاش شد تا از معماری‌های متفاوت شبکه‌های ژرف استفاده شود تا مناسب‌ترین و دقیق‌ترین جواب‌ها نسبت به مقادیر واقعی پیش‌بینی شود. با تکرارهای پی در پی در اپک‌های متوالی تلاش شد تا مقدار تابع زیان به حداقل برسد. همانطور که در این پژوهش مشاهده شد، شبکه LSTM دو لایه دارای بهترین دقت در طولانی مدت است و در ۴۰ امین روز اختلاف فاحشی بین پیش‌بینی‌های این شبکه و شبکه‌های دیگر وجود دارد. شبکه‌های CNN دقت کمتری نسبت به شبکه LSTM داشتند اما در مجموع پیش‌بینی‌ها، اختلاف کمی با مقادیر واقعی داشتند. در پایان می‌توان گفت شبکه‌های ژرف برای پیش‌بینی دقیق حرکت قطبی مناسب هستند. همچنین نتایج نشان داد دقت روش آنالیز هارمونیک کمترین مربعات نسبت به روش‌های ژرف کمتر می‌باشد.

همچنین از دیگر عوامل موثر در افت دقت روش LSHA، می‌توان به عدم یکسان بودن دوره تناوب‌های غالب سری‌های زمانی x و y اشاره نمود که روش‌های پارامتریکی مانند LSHA نمی‌توانند این چالش را حل کنند [۴۳].

۴- نتیجه گیری

با استفاده از روش‌های مدرن فضایی مانند (Satellite Laser Ranging) SLR، VLBI (Very Long Base Line Interferometry)، GPS و روش‌های دیگر، تعیین EOP از جمله PM به دقت و وضوح بالایی رسیده است. از این پارامترها در علوم گوناگونی مانند نجوم، زمین‌شناسی، اقیانوس‌شناسی استفاده می‌شود. با این حال باید اشاره کرد که به دلیل پردازش‌های بسیار استفاده از این پارامترها به شکل در لحظه ممکن نیست و در عین حال برای کاربردهای مختلفی همچون تعیین مدار ماهواره به مقادیر پارامترها برای آن روز نیاز است و باتوجه به اینکه این پارامترها به صورت روزانه در دسترس نمی‌باشند بایستی این پارامترها با دقت بالایی پیش‌بینی و برآورد شوند. در این تحقیق از چندین روش یادگیری ژرف و روش آنالیز

مراجع

- [1] Nastula J, Ponte RM (1999) Further evidence for oceanic excitation of polar motion. *Geophys J Int* 139(1):123–130. <https://doi.org/10.1046/j.1365-246X.1999.00930.x>. (n.d.).
- [2] Zotov LV, Bizouard C (2015) Regional atmospheric influence on the Chandler wobble. *Adv Space Res* 55(5):1300–1306. <https://doi.org/10.1016/j.asr.2014.12.013>. (n.d.)
- [3] Jin, X., Liu, X., Guo, J., & Shen, Y. (2021). Analysis and prediction of polar motion using MSSA method. *Earth, Planets and Space*, 73(1), 147. <https://doi.org/10.1186/s40623-021-01477-2>.
- [4] Chen JL, Wilson CR (2005) Hydrological excitations of polar motion, 1993–2002. *Geophys J Int* 160(3):833–839. <https://doi.org/10.1111/j.1365-246X.2005.02522.x>. (n.d.).
- [5] Lambeck K (2005) *The Earth's variable rotation: Geophysical causes and consequences*. Cambridge University Press. (n.d.)
- [6] Dill R, Dobslaw H (2010) Short-term polar motion forecasts from earth system modeling data. *J Geodesy* 84(9):529–536. <https://doi.org/10.1007/s00190-010-0391-5>. (n.d.).
- [7] Jin S, van Dam T, Wdowinski S (2013) Observing and understanding the Earth system variations from space geodesy. *J Geodyn* 72:1–10. <https://doi.org/10.1016/j.jog.2013.08.001>. (n.d.).
- [8] Wang G, Liu L, Su X, Liang X, Yan H, Tu Y, Li Z, Li W (2016) Variable chandler and annual wobbles in Earth's polar motion during 1900–2015. *Surv Geophys* 37(6):1075–1093. <https://doi.org/10.1007/s10712-016-9384-0>. (n.d.).
- [9] Adhikari S, Ivins ER (2016) Climate-driven polar motion: 2003–2015. *Sci Adv* 2(4): e1501693. <https://doi.org/10.1126/sciadv.1501693>. (n.d.)
- [10] G. Seeber, *Satellite geodesy, 2nd completely rev. and Extended ed.* Berlin ; New York: Walter de Gruyter, 2003

- [11] Nastula J, Pańnicka M, Kołaczek B (2011) Comparison of the geophysical excitations of polar motion from the period: 1980.0–2009.0. *Acta Geophys* 59(3):561–577. [https:// doi. Org/ 10. 2478/ s11600- 011- 0008-2](https://doi.org/10.2478/s11600-011-0008-2). (n.d.).
- [12] Schuh H, Nagel S, Seitz T (2001) Linear drift and periodic variations observed in long time series of polar motion. *J Geodesy* 74(10):701–710. [https:// doi. Org/ 10. 1007/ s00190 00000 133](https://doi.org/10.1007/s001900000133). (n.d.).
- [13] Gross RS (2000) The excitation of the Chandler wobble. *Geophys Res Lett* 27(15):2329–2332. [https:// doi. Org/ 10. 1029/ 2000G L01114 50](https://doi.org/10.1029/2000GL011145). (n.d.).
- [14] Malkin Z, Miller N (2010) Chandler wobble: Two more large phase jumps revealed. *Earth Planets Space* 62(12):943–947. [https:// doi. Org/ 10. 5047/ eps. 2010. 11. 002](https://doi.org/10.5047/eps.2010.11.002). (n.d.).
- [15] H. B. İz, 'Polar motion modeling, analysis, and prediction with time dependent harmonic coefficients', *J. Geod.*, vol. 82, no. 12, pp. 871–881, Dec. 2008, doi: 10.1007/s00190-008-0215-z.
- [16] Höpfner J (2004) Low-frequency variations, chandler and annual wob- bles of polar motion as observed over one century. *Surveys in geophysics*. Kluwer, New York, vol. 25, pp 1–54 'Iz HB, Chen YQ (1999) VLBI rates with first order autoregressive disturbances. *J Geodynam* 28(2–3):131–145.
- [17] Liu L, Hsu H, Grafarend EW (2007) Normal Morlet wavelet transform and its application to the Earth's polar motion. *J Geophys Res Solid Earth* 112: B08401. [https:// doi. Org/ 10. 1029/ 2006J B0048 95](https://doi.org/10.1029/2006JB004895). (n.d.).
- [18] Kosek W, McCarthy DD, Luzum BJ (1998) Possible improvement of Earth orientation forecast using autocovariance prediction proce- dures. *J Geod* 72(4):189–199. Doi:10.1007/s001900050160. (n.d.).
- [19] Yao YB, Yue S, Chen P (2013) A new LS + AR model with additional error correction for polar motion forecast. *Sci China Earth Sci* 56(5):818–828. Doi:10.1007/s11430-012-4572-3. (n.d.).
- [20] Kosek W, Kalarus M, Niedzielski T (2007) Forecasting of the Earth orientation parameters comparison of different algorithms. In: Capitaine N (ed) *Proceedings of the "Journées systèmes de deréférence spatio-temporels (2007) Observatoire de Paris, 17–19 Sept 2007, Paris, France*. Pp 155–158. (n.d.).
- [21] Kalarus M, Schuh H, Kosek W, Akyilmaz O, Bizouard Ch, Gambis D, Gross R, Jovanovic B, Kumakshev S, Kutterer H, Ma L, Mendes Cerveira PJ, Pasynok S, Zotov L (2010) Achievements of the Earth orientation parameters prediction comparison campaign. *J Geod* 84(10):587–596. Doi:10.1007/s00190-010-0387-1. (n.d.).
- [22] Xu X, Zhou Y (2015) EOP prediction using least square fitting and autoregressive filter over optimized data intervals. *Adv Space Res* 56(10):2248–2253. Doi: 10.1016/j.asr.2015.08.007. (n.d.).
- [23] Shen, Y., Guo, J., Liu, X., Kong, Q., Guo, L., & Li, W. (2018). Long-term prediction of polar motion using a combined SSA and ARMA model. *Journal of Geodesy*, 92(3), 333–343. <https://doi.org/10.1007/s00190-017-1065-3>
- [24] Akyilmaz O, Kutterer H (2004) Prediction of Earth rotation parameters by fuzzy inference systems. *J Geod* 78(1):82–93. Doi:10.1007/s00190-004-0374-5. (n.d.).
- [25] Akyilmaz O, Kutterer H, Shum CK, Ayan T (2011) Fuzzy-wavelet based prediction of Earth rotation parameters. *Appl Soft Comput* 11(1):837–841. Doi: 10.1016/j.asoc.2010.01.003. (n.d.).
- [26] Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. Learning representations by back-propagating errors. *nature*, 323(6088), pp.533-536.
- [27] LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W. and Jackel, L., 1989. Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, 2.
- [28] Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, 9(8), pp.1735-1780.
- [29] Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, 135, 109864.
- [30] Yang, C.H., Wu, C.H. and Hsieh, C.M., 2020. Long short-term memory recurrent neural network for tidal level forecasting. *IEEE Access*, 8, pp.159389-159401.
- [31] Tierra, A., Dalazoana, R. and De Freitas, S., 2008. Using an artificial neural network to improve the transformation of coordinates between classical geodetic reference frames. *Computers & Geosciences*, 34(3), pp.181-189.

- [32] Schuh, H., Ulrich, M., Egger, D., Müller, J., & Schwegmann, W. (2002). Prediction of Earth orientation parameters by artificial neural networks. *Journal of Geodesy*, 76(5), 247-258.
- [33] Liao, D. C., Wang, Q. J., Zhou, Y. H., Liao, X. H., & Huang, C. L. (2012). Long-term prediction of the earth orientation parameters by the artificial neural network technique. *Journal of Geodynamics*, 62, 87-92.
- [34] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, pp.1097-1105.
- [35] Erhan, D., Szegedy, C., Toshev, A. and Anguelov, D., 2014. Scalable object detection using deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2147-2154).
- [36] Harbola, S. and Coors, V., 2019. One dimensional convolutional neural network architecture for wind prediction. *Energy Conversion and Management*, 195, pp.7
- [37] A. Amiri-Simkooei, *Least-squares variance component estimation: theory and GPS applications*. Delft: NCG, 2007.
- [38] A. R. Amiri-Simkooei and J. Asgari, 'Harmonic analysis of total electron contents time series: methodology and results', *GPS Solut.*, vol. 16, no. 1, pp. 77-88, Jan. 2012, doi: 10.1007/s10291-011-0208-x.
- [39] Agarap, A.F., 2018. Deep learning using rectified linear units (ReLU). *arXiv preprint arXiv:1803.08375*.
- [40] Wilson, D.R. and Martinez, T.R., 2001, July. The need for small learning rates on large problems. In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No. 01CH37222) (Vol. 1, pp. 115-119)*. IEEE.
- [41] Kingma, D.P. and Ba, J., 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [42] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471
- [43] G. Beutler, A. Villiger, R. Dach, A. Verdun, and A. Jäggi, 'Long polar motion series: Facts and insights', *Adv. Space Res.*, vol. 66, no. 11, pp. 2487-2515, Dec. 2020, doi: 10.1016/j.asr.2020.08.033.