فیزیک زمین و فضا، دوره ۴۱، شماره ۲، تابستان ۱۳۹۴، صفحهٔ ۲۸۱- ۲۹۹

مدل‌سازی تغییرات سطح لحظه‌ای آب دریای خزر با استفاده از مشاهدات ارتفاع‌سنجی ماهواره‌ای

عبدالرضا صفری[۱]، سیمین کلانتریون[۲*] و هادی امین[۳]

۱. دانشیار، گروه مهندسی نقشه‌برداری، پردیس دانشکده‌های فنی دانشگاه تهران، ایران
۲. دانشجوی کارشناسی‌ارشد، گروه مهندسی نقشه‌برداری، پردیس دانشکده‌های فنی دانشگاه تهران، ایران
۳. دانش‌آموختهٔ کارشناسی‌ارشد، گروه مهندسی نقشه‌برداری، پردیس دانشکده‌های فنی دانشگاه تهران، ایران

(دریافت: ۹۳/۲/۲۷، پذیرش نهایی: ۹۴/۳/۱۲)

چکیده

پیش‌بینی تراز آب دریا در صنایع کشتیرانی، تأسیسات نیرو در مناطق دریایی و بنادر، صنعت شیلات، شهرسازی بنادر، همچنین در علوم مختلف دریایی و اقیانوس‌شناسی اهمیت ویژه‌ای دارد. روش‌های متعددی از قبیل روش‌های تجربی، مدل‌های عددی و محاسباتی برای مطالعهٔ پیش‌بینی امواج دریا توسعه یافته است. یکی از این روش‌های مطالعاتی، روش سری‌های زمانی است. در این راستا مدل‌های مختلفی ارائه شده‌اند که از آن جمله می‌توان به مدل‌های سری زمانی ARMA، MA و AR اشاره کرد. در تحقیق حاضر، از مدل‌های مختلف سری زمانی همراه با روش‌های آنالیز طیفی فوریه برای یافتن فرکانس موجود در سری زمانی و روش کمترین مربعات با تکیه بر مفهوم ایستایی سری زمانی به‌منظور دستیابی به فرکانس دقیق‌تر استفاده و امکان افزایش دقت نسبی برآورد و پیش‌بینی تراز آب دریا در بازه زمانی بیشتر مطالعه شد. با توجه به نتایج به‌دست آمده از به‌کارگیری داده‌های ماهوارهٔ ارتفاع‌سنجی Jason-2 در بازهٔ زمانی ۲۰۰۸ تا ۲۰۱۲ در برآورد و تخمین مقادیر سری زمانی ارتفاع لحظه‌ای آب دریای خزر در بازه زمانی بلندمدت با استفاده از مدل‌های مختلف، همچنین کمترین مقدار محک آکائیک AIC، BIC و بیزی شوارتز SBC، مدل سری زمانی AR(6) کارآیی بهتری نسبت به مدل‌های MA(q) و ARMA(p,q) داشت و روند تغییرات سری زمانی را با خطای کمتری پیش‌بینی می‌کند. با توجه به تحلیل آماری نتایج حاصل از تخمین مقادیر ارتفاع لحظه‌ای آب دریا، مقدار میانگین مطلق خطا برابر ۳/۸ میلی‌متر و مجذور میانگین مربعات خطا برابر ۱/۴۳ سانتی‌متر در روز به‌دست آمد.

واژه‌های کلیدی: روش سری‌های زمانی، ماهوارهٔ ارتفاع‌سنجی Jason-2، محک آکائیک AIC، محک BIC، ملاک بیزی شوارتز SBC

۱. مقدمه

نوسانات تراز آب دریاها به‌واسطهٔ تأثیر بر صنایع مرتبط اهمیت ویژه‌ای دارد. کشتیرانی، حمل‌ونقل دریایی، شیلات و شهرسازی بنادر از جمله صنایعی است که تغییرات سطح آب دریاها به‌طور مستقیم روی آن تأثیرگذار است. از مهم‌ترین مواردی که در این راستا می‌توان به آن اشاره کرد، تأسیسات نیرو و انرژی در مناطق دریایی و بنادر است که نوسانات تراز آب دریا تأثیرات مخربی در این حوزه به‌دنبال دارد. بنابراین، با توجه به صدمات جبران‌ناپذیر زیست‌محیطی و اقتصادی، بررسی نوسانات آب دریاها و نحوهٔ تغییرات سطح آن طی زمان امری ضروری است. در سال‌های اخیر، محققان در علوم مختلف دریایی و اقیانوس‌شناسی مطالعات متعددی را در این زمینه انجام داده‌اند و الگوهای مختلفی برای پیش‌بینی کوتاه‌مدت و بلندمدت تراز آب دریاها ارائه کرده‌اند.

چنگ و کیو (۲۰۰۷) با استفاده از داده‌های ناهنجاری سطح دریا، روند تغییرات تراز آب دریای چین جنوبی را از سال ۱۹۹۳ تا ۲۰۰۵ بررسی کردند. مطالعات صورت گرفته نشان داد که تراز آب این دریا از سال ۱۹۹۳ تا سال ۲۰۰۱ روندی افزایشی به میزان ۱۱/۴ میلی‌متر در هر سال داشته و از سال ۲۰۰۱ تا ۲۰۰۵ سطح آب به میزان ۱۱/۸ میلی‌متر در هر سال پایین رفته است. همچنین، توزیع گرمایی و آثار آن بر تبادل آب و تغییرات سطح تراز دریا در این تحقیق بررسی شد. مارکوس و همکاران (۲۰۰۷) سری زمانی حاصل از داده‌های ارتفاع‌سنجی و تایدگیج‌ها را در خلیج بیسکای (Biscay) برای بررسی بالا آمدن

*نگارنده رابط: E-mail: simin.kalantarioun@yahoo.com

تراز آب دریا استفاده کردند. نتایج به‌دست آمده نشان‌دهندهٔ بالا آمدن سطح آب دریا از سال ۱۹۹۳ تا ۲۰۰۲ به میزان ۳/۰۹ میلی‌متر در هر سال است. در این تحقیق، بررسی نیروی اتمسفری نشان‌دهندهٔ تأثیر ۱۵ درصدی این نیرو بر نوسانات تراز آب منطقه بوده است.

یکی از مشکلات ملی در سال‌های اخیر، نوسانات تراز آب دریای خزر، تأثیرات مخرب آن بر تأسیسات نیروی دریایی، بندری و شیلات، صنایع کشتیرانی، حمل و نقل، نواحی مسکونی و نیز صدمات جبران‌ناپذیر زیست‌محیطی بوده است. نوسانات تراز آب دریای خزر آثار متفاوت، مخرب و بعضاً مفیدی بر کشورهای مجاور داشته است و چون اکثر شهرهای بزرگ در مجاور این دریا (شهرهای بندری) قرارگرفته‌اند، لذا بررسی و پیش‌بینی تغییرات سطح آب دریای خزر ضرورت است.

ریاضی (۱۳۷۸) با اندازه‌گیری مستقیم مشخصات امواج دریا به مدت طولانی و میانگین‌گیری از این مشخصات، پارامترهای مربوط به شکست امواج را پیش‌بینی کرد. وی با مطالعهٔ امواج فقط در جهت‌های خاص به بررسی و پیش‌بینی شکست امواج در آب‌های کم‌عمق ساحل خزر پرداخت.

فلاح و همکاران (۱۳۸۳) با اجرای مدل مایک ۲۱ و مقایسهٔ نتایج آن با داده‌های حاصل از مشاهدات و مدل عددی WAM، توانایی‌های این دو مدل را بررسی و دقت نتایج حاصل از کاربرد آن‌ها را در مدل‌سازی وضعیت امواج در دریای خزر ارزیابی کردند.

رجبی و همکاران (۱۳۹۲) با استفاده از مدل عددی SWAN (Simulating WAves Nearshore) که فقط برای آب‌های کم‌عمق و نیمه‌عمیق ساحلی توسعه داده شده است، به پیش‌بینی مشخصات امواج ناشی از باد در منطقهٔ انزلی پرداختند. در این تحقیق اطلاعات باد مرکز اروپایی پیش‌بینی میان‌مدت وضع هوا (ECMWF) در سال ۲۰۰۳ ورودی مدل استفاده شد. به دلیل نقص مدل SWAN و باد ECMWF، در به‌دست آوردن بلندترین ارتفاع موج شاخص یا حتی بلندترین پریود موج خطایی وجود خواهد داشت، لذا پیش‌بینی دقیقی در قله‌ها انجام نمی‌شود.

مدل‌های طیفی عددی بر اساس حل معادلهٔ تعادل انرژی طیف استوار است. در عین حال که این مدل‌ها ابزاری کارآمدند، به‌دلیل نیاز به استفاده از برنامه‌های رایانه‌ای پیچیده و زمان‌بر و حافظه و سرعت بالا برای رایانه‌های مورد استفاده، همچنین نیاز به اطلاعات دریایی و هواشناسی گسترده، بهره‌گیری از این روش‌ها پرهزینه و در بسیاری از موارد ناممکن است.

محجوبی و اردلان صمغی (۱۳۸۸) با استفاده از درخت‌های تصمیم رگرسیونی که یکی از ابزارهای داده‌کاوی است، پیش‌بینی مشخصات امواج ناشی از باد را در دریای خزر انجام دادند. بدین منظور از اطلاعات میدانی مربوط به دریای خزر در بندر امیرآباد و منطقهٔ نکا استفاده شد. برای منطقهٔ نکا در برخی زمان‌ها داده‌ای ثبت نشده بود و این خود یکی از عوامل ایجاد خطا در پیش‌بینی است. همچنین، برای منطقهٔ امیرآباد، آمار باد از ایستگاه هواشناسی بابلسر استخراج شده است که در فاصلهٔ ۶۶ کیلومتری از آن قراردارد و این موضوع نیز یکی از منابع ایجاد خطا در پیش‌بینی امواج در منطقهٔ امیرآباد محسوب می‌شود.

به موازات استفاده از روش‌های عددی در پیش‌بینی امواج، روش شبکه‌های عصبی، همچنین بهره‌گیری از روش‌های آماری متکی بر پیش‌بینی سری‌های زمانی نیز استفاده شد. پیش‌بینی پدیده‌های غیرخطی و پیچیده با استفاده از شبکه‌های عصبی مصنوعی نیز انجام می‌شود.

زمانی و عظیمیان (۱۳۸۳) با استفاده از داده‌های واقعی امواج دریای خزر که با بویهٔ موج‌نگار و سیستم ADCP اندازه‌گیری شد، در شبکهٔ مصنوعی عصبی، ارتفاع امواج ناشی از باد در دریای خزر پیش‌بینی شد. از مشکلات اساسی در این روش پیش‌بینی، می‌توان به دقت و سرعت روش تنها در دوره‌های زمانی کوتاه‌مدت اشاره کرد. در این تحقیق برای ۳، ۶، ۱۲ و ۲۴ ساعت بعد پیش‌بینی انجام شد که با افزایش زمان پیش‌بینی، دقت پیش‌بینی کاهش یافت. در پیش‌بینی برای ۲۴ ساعت آینده نتیجهٔ دقیقی حاصل نشد.

با ظهور ارتفاع‌سنجی ماهواره‌ای در سال ۱۹۷۳، امکان پایش سطح آب‌ها با دقت بالا فراهم شد (آنزنهوفر

و همکاران، ۱۹۹۹). ماهواره‌های ارتفاع‌سنجی اطلاعات ارتفاعی نقاط مختلف سطح کرهٔ زمین را در امتداد مدارهای تعیین‌شده جمع‌آوری می‌کنند. مأموریت اصلی این ماهواره‌ها اندازه‌گیری ارتفاع سطح آب دریاها و اقیانوس‌ها در زمان‌ها و مکان‌های مختلف است.

باربوسا و همکاران (۲۰۰۶) از داده‌های ماهوارهٔ ارتفاع‌سنجی Topex/Poseidon از سال ۱۹۹۲ تا ۲۰۰۵ برای بررسی نوسانات تراز آب در اقیانوس اطلس شمالی استفاده کردند. در این تحقیق تجزیه و تحلیل داده‌ها با استفاده از روش چند متغیرهٔ اتورگرسیو مرتبهٔ سوم انجام شد.

لبدف وکاستیانوی (۲۰۰۵) از داده‌های ارتفاع‌سنجی ماهواره‌ای، داده‌های ماهواره‌های Topex/Poseidon و Jason-1 در بازه زمانی سیزده ساله از سال ۱۹۹۲ تا ۲۰۰۴، برای بررسی تغییرات آنومالی فصلی و سالانه از سطح دریا، سرعت باد و ارتفاع موج در نقاط مختلف دریای خزر و سطح رودخانهٔ ولگا استفاده کردند. شجاعی (۱۳۸۳) مسئلهٔ ادغام اطلاعات سنجنده‌های ماهواره‌ای مختلف و اطلاعات ژئوفیزیکی به‌منظور دستیابی به اطلاعات سطح لحظه‌ای آب دریا در خلیج فارس و دریای عمان را با استفاده از مدل‌سازی سری‌های زمانی مطالعه کرد. جعفری (۱۳۸۵) به مدل‌سازی روند تغییرات سطح آب دریای خزر با استفاده از روش‌های آماری و مشاهدات ارتفاع‌سنجی ماهواره‌ای و تایدگیج‌های ساحلی پرداخت. در این مطالعه از داده‌های ارتفاع‌سنجی ماهواره‌های Topex/Poseidon و Jason-1 استفاده شد. داده‌های تایدگیج‌های ساحلی بندر انزلی، بندر صدرا و بندر نوشهر برای کنترل نتایج حاصل از ارتفاع‌سنجی ماهواره‌ای استفاده شد.

یکی از روش‌های مطالعاتی بررسی تغییرات تراز آب دریا و پیش‌بینی کوتاه و بلندمدت، روش سری‌های زمانی است. برای این منظور، روش‌ها و مدل‌های مختلفی ارائه شده است، از آن جمله می‌توان به مدل‌های سری زمانی ARMA، MA و AR اشاره کرد. در این تحقیق، با استفاده از فرایندهای مختلف سری زمانی، داده‌های ماهوارهٔ ارتفاع‌سنجی Jason-2 از سال ۲۰۰۸ تا سال ۲۰۱۲ تجزیه و تحلیل شد. برای بررسی میزان عملکرد این مدل‌ها در برآورد و تخمین مقادیر سری زمانی در بازه زمانی بلندمدت، از معیارهای میانگین مطلق خطا، مجذور میانگین مربعات خطا، محک آکائیک AIC، BIC و بیزی شوارتز SBC استفاده شد. هدف اصلی در این بررسی انتخاب مناسب‌ترین مدل پیش‌بینی تراز آب دریا در منطقهٔ مورد بررسی بر پایهٔ معیارهای خطای پیش‌بینی است. همچنین، روند نوسان‌های سطح آب دریا از سال ۲۰۰۸ تا سال ۲۰۱۲ و افزایش دقت نسبی برآورد و پیش‌بینی تراز آب دریا در بازه زمانی بلندمدت بررسی شد.

۲. مدل‌سازی سری‌های زمانی

تحلیل در قلمرو فرکانسی روشی است که نوسانات سری زمانی را برحسب رفتار سینوسی در فرکانس‌های مختلف بیان می‌کند. آنالیز طیفی و تبدیلات بعد فرکانس، روش ریاضی کارآمدی برای مدل‌سازی پدیده‌های تصادفی و غیرتصادفی است. در اغلب پدیده‌های طبیعی تناوب و تکرار به وضوح قابل درک است. سری زمانی داده‌های سطح لحظه‌ای آب متأثر از تغییرات فصلی، دوره‌ای و روند و تغییرات تصادفی ناشی از عوامل محیطی و جزآن است.

بررسی رفتار یک پدیده و به‌دنبال آن پیش‌بینی رفتار آن پدیده، مستلزم بیان آن رفتار در غالب مدل ریاضی درست و قابل اعتمادی است. مدل‌سازی یک پدیده، عموماً با نمونه‌برداری از رفتار پدیده در بازه زمانی و با نرخ مشخص آغاز می‌شود. سپس، مدل‌سازی مورد نظر با مطالعه و تجزیه و تحلیل نمونهٔ برداشته شده صورت می‌پذیرد. این روند در اصطلاح آنالیز طیفی نامیده می‌شود. برای مدل‌سازی رفتار یک پدیده که ماهیتی پریودیک داشته باشد، آنالیز فوریه ابزار ریاضی مناسب و کارآمدی است.

با در اختیار داشتن ارتفاع آب و لحظهٔ قرائت هر ارتفاع، یک سری زمانی از روند تغییرات به‌دست خواهد

آمد. فرکانس داده‌ها برداری به شکل رابطه (۱) را تشکیل می‌دهد.

$$\omega = \omega\{i\} \ , \ i = 1,2,...,n \qquad (۱)$$

مدل ارتفاع سطح آب به‌صورت ترکیب خطی از توابع مثلثاتی به‌صورت رابطه (۲) در نظر گرفته می‌شود.

$$y(t) = \sum_{i=0}^{n} a_i \cos \ \omega_i t \ + b_i \sin \ \omega_i t \qquad (۲)$$

رابطه (۲) سری فوریه دنبالهٔ $y(t)$ نامیده می‌شود که در آن a_i و b_i ضرایب فوریه است (نیرومند، ۱۳۷۱). اگر بتوان a و b را برای هر ω محاسبه کرد، می‌توان مقدار عددی پدیدهٔ پریودیک را در هر اپک t محاسبه کرد. اما مشکل اساسی در برآورد بردار ω و مجهول است. در واقع، مجهول در این معادلات، فرکانس‌های موجود در داده‌هاست. در صورت برآورد مقادیر ω_i، با استفاده از روش کمترین مربعات می‌توان مقادیر a_i و b_i را به‌دست آورد.

در روش آنالیز طیفی کمترین مربعات اگر مدل به‌صورت رابطه (۳) در نظر گرفته شود:

$$Y = AX \qquad (۳)$$

که در آن، Y بردار مشاهدات، A ماتریس ضرایب و X بردار مجهولات است که در واقع مقادیر a_i و b_i مربوط به هر فرکانس است. ماتریس‌های A و X را می‌توان به‌صورت رابطه (۴) نوشت:

$$A^T = \begin{bmatrix} \cos\omega_1 t_1 & \cos\omega_2 t_2 & ... & \cos\omega_n t_n \\ \sin\omega_1 t_1 & \sin\omega_2 t_2 & ... & \sin\omega_n t_n \end{bmatrix} \ , \ X = \begin{bmatrix} a_i \\ b_i \end{bmatrix} \qquad (۴)$$

و ضرایب با استفاده از رابطهٔ (۵) تعیین می‌شود.

$$X = (A^T A)^{-1} A^T Y \qquad (۵)$$

در نهایت، ضرایب و فرکانس‌ها به نحوی در فرایند سرشکنی شرکت داده می‌شود که بتوان به بهترین شکل ممکن موج مورد نظر را مدل‌سازی کرد. به‌عبارت دیگر V_x را مینیمم کرد.

دو نکتهٔ مهم در استفاده از این داده‌ها مطرح است. نخست، باید مجموعه اطلاعات اولیه‌ای از فرکانس داده‌ها در اختیار داشت. سپس، فرکانس را به تدریج تغییر داد و بهترین فرکانس را استخراج کرد. مشکلی که وجود دارد این است که اگر هیچ اطلاعی از فرکانس‌های موجود در داده‌ها در دسترس نباشد، یافتن بهترین فرکانس امری زمان‌بر است. دوم، مقدار نهایی فرکانس نیز مشخص نمی‌شود، بلکه تقریب خوبی از هر فرکانس اثرگذارنده در داده‌ها مشخص می‌شود.

۲. ۱. روش آنالیز فوریه

در این روش، داده‌ها از فضای زمان به فضای فرکانس تبدیل می‌شود. در صورتی که داده‌ها به‌صورت پیوسته در اختیار باشد، روابط فوریه به‌صورت انتگرالی و در صورت گسسته بودن مشاهدات، روابط با استفاده از مجموع سری بیان می‌شود. در این روش، بردار مشاهدات $x(t)$ از فضای زمان به بردار تبدیل‌یافته $X(f)$ در فضای فرکانس تبدیل می‌شود. بزرگ‌ترین فرکانس در دسترس از طریق رابطهٔ (۶) به‌دست می‌آید که طبق قضیهٔ نمونه‌برداری نایکوئیست–شانون بیان می‌شود. بنابر آن اگر فرکانس نمونه‌برداری بزرگ‌تر از دو برابر بالاترین مؤلفهٔ فرکانسی سیگنال باشد، می‌توان سیگنال را از روی سیگنال نمونه‌برداری شده به‌طور دقیق و بدون از دست دادن جزئیات بازسازی کرد (طوریان، ۱۳۸۵). فرکانس زیر به Nyquist Frequencyمشهور است.

$$\mathrm{F}_{\mathrm{nyquist}} = \frac{1}{2\Delta t} \qquad (۶)$$

حال هر مؤلفه از بردار $X(f)$ نشان‌دهندهٔ فرکانسی است که به‌صورت رابطه (۷) به‌دست می‌آید.

$$f \ i \ = \frac{i}{\frac{N}{2}} \mathrm{F}_{\mathrm{nyquist}} \qquad (۷)$$

در رابطهٔ (۷) i شمارهٔ هر مؤلفه و N تعداد کل داده‌هاست. باید توجه داشت که در نقطهٔ $i = 0$، اولین درایهٔ بردار X، فرکانس پدیده‌ای با پریود بی‌نهایت نشان داده شده است، بنابراین درایهٔ اول برای یافتن فرکانس موجود در داده‌ها قابل استفاده نیست.

پس از تبدیل داده‌ها، بزرگ‌ترین عدد موجود در بردار $X(f)$ تأثیرگذارترین مؤلفه روی بردار مشاهدات است. بنابراین با داشتن این درایه می‌توان مقدار فرکانس متناظر با آن را به‌دست آورد.

۲. ۱. ۱. استخراج کامل مؤلفه‌های پدیدهٔ پریودیک از بین داده‌ها

روش فوریه روشی برای یافتن فرکانس از روی اطلاعات است. اما برای اینکه بتوان اثر پدیدهٔ پریودیک را از روی سطح متوسط آب جدا کرد، باید سایر عوامل هر موج را نیز بررسی کرد. برای اینکه بتوان یک موج سینوسی استخراج کرد، باید سه المان هر موج یعنی فرکانس، دامنه و اختلاف فاز آن موج را به‌دست آورد. یک روش یافتن اختلاف فاز و دامنه، تقسیم منحنی کسینوسی به دو منحنی ساده است (طوریان، ۱۳۸۵).

$$h(t_i) = a\cos(\omega t_i) + b\sin(\omega t_i) \qquad (۸)$$

یا به‌شکل:

$$h(t_i) = A\cos(\omega t_i - \varphi) \qquad (۹)$$

که در آن، A و φ به ترتیب دامنه و اختلاف فازند. در رابطه (۸)، دو مجهول به‌صورت خطی با یکدیگر ترکیب شده‌اند، در نتیجه با یک مدل خطی می‌توان آن‌ها را در مدل پارامتریک کمترین مربعات استخراج کرد.

با یافتن a و b مجهولات دامنه و اختلاف فاز با دو رابطه (۱۰)، به‌راحتی به‌دست می‌آید.

$$\mathrm{A} = \sqrt{a^2 + b^2}\ ,\ \varphi = 2\tan^{-1}\left(\frac{b}{A+a}\right) \qquad (۱۰)$$

برای یافتن دامنه و فاز تمام مؤلفه‌های موجود در یک سری زمانی، روش‌های متعددی وجود دارد که در زیر به چند روش اشاره می‌شود (طوریان، ۱۳۸۵).

روش اول: استخراج تک به تک مؤلفه‌ها

در این حالت از روش پارامتریک کمترین مربعات برای یافتن a و b و در نتیجه مجهولات دامنه و اختلاف فاز هر موج استفاده می‌شود. در این روش، پس از اینکه یک موج به‌طور کامل از داده‌ها استخراج شد، فرمول ریاضی موج به‌صورت معادله بازسازی و اثر این موج از روی تمامی ارتفاعات قرائت شده برداشته می‌شود. سپس، تبدیل فوریه روی داده‌ها به‌کار می‌رود. در تبدیل فوریهٔ جدید یک فرکانس دیگر نیز از داده‌ها به‌دست آمد و اثر آن حذف شد. این روند تا کوچک‌تر شدن دامنهٔ مؤلفه‌های به‌دست آمده از دقت مورد نظر، ادامه می‌یابد.

روش دوم: یافتن دامنه و اختلاف فاز تمام مؤلفه‌ها به‌صورت هم‌زمان

در روش اول، هر موج به تنهایی از اطلاعات بیرون کشیده می‌شود، اما در این روش تمامی مؤلفه‌ها یکجا به‌دست می‌آیند. در این روش ابتدا از روی پریودوگرامی که در روش فوریه به‌دست می‌آید، فرکانس‌های دارای بزرگ‌ترین تأثیر به‌دست می‌آید. تعداد این فرکانس‌ها بسته به دقت کار، همچنین داشتن اطلاعات در مورد تابع مورد تقریب، انتخاب می‌شود. سپس، برای به به‌دست آوردن دامنه و اختلاف فاز، مدل مشاهدات خطی شد و مدل پارامتریک برای یافتن a و b تمام سیگنال‌ها تشکیل می‌شود. در این روش برای به‌دست آوردن تمامی n مؤلفهٔ اول، فقط یک تبدیل فوریه انجام می‌شود. پس از به‌دست آوردن a و bها، دامنه و فاز تمامی مؤلفه‌ها به‌دست می‌آید.

در روش اول، هر سیگنال به تنهایی استخراج می‌شود و اگر بخواهیم سیگنال دیگری علاوه بر سیگنال‌های قبلی به‌دست آوریم، نیازی به تشکیل مدل اضافی برای سیگنال‌های قبلی نیست و این مسئله، روند اجرا شدن عملیات را بسیار سریع می‌کند. در حالی که در روش دوم، حتماً باید تمام سیگنال‌ها هم‌زمان به‌دست آید.

در روش اول، سیگنال‌های مزاحم بیشتری وارد نتایج می‌شود، زیرا پس از حذف هر سیگنال از داده‌ها چون روش فوریه فرکانس موجود در داده را فقط تقریب می‌زند، در نتیجه مقدار کوچکی از سیگنال حذف شده روی داده‌ها تأثیرگذار است. در نتیجه در مراحل بعدی، این فرکانس‌ها تأثیر می‌گذارد و در واقع به‌صورت فرکانس غیرواقعی وارد می‌شود. البته، این سیگنال‌های اضافی را نیز می‌توان به نوعی سیگنال واقعی دانست با این تفاوت که منشأ ایجاد دو سیگنال واقعی و غیرواقعی، یکسان نیست. اما در روش دوم، چون همهٔ فرکانس‌ها با هم و در یک مدل سرشکن می‌شود، اثر فرکانس‌های غیرواقعی، در فرکانس‌های واقعی محو می‌شود و تعداد کمتری فرکانس غیرواقعی وجود خواهد داشت.

روش سوم: استخراج مؤلفه‌ها با مدل غیرخطی

در دو روش قبلی، هیچ فرکانسی در هنگام آنالیز و به‌کارگیری روش کمترین مربعات، تغییر نمی‌کند و نمی‌توان فرکانس‌های با مضارب غیرصحیح از فرکانس مینیمم را پیدا کرد. اما در روش غیرخطی، ابتدا به‌وسیلهٔ دو روش فوق فرکانس‌ها و سپس دامنه و اختلاف فازها تشخیص داده شد. سپس، با روش غیرخطی مجهولات به‌دست می‌آید. پس از حل این روش، فرکانس‌ها نیز تغییر کرد و به بهترین فرکانس‌های موجود در داده نزدیک شد. در این روش، مقادیر فرکانس‌ها تغییر زیادی نمی‌کند، به‌خصوص اگر فاصله‌های زمانی بین مشاهدات کوچک باشد، زیرا هرقدر که فاصلهٔ برداشت‌ها کوچک شود، فاصلهٔ بین فرکانس‌های انتخابی روش فوریه، که از بین آن‌ها بهترین فرکانس مشخص می‌شود نیز کمتر می‌شود. این مطلب باعث می‌شود که عملیات تقریب زدن بهترین فرکانس موجود، با دقت بیشتری انجام شود.

۲. ۲. ایستایی یک سری

نظریهٔ احتمال سری‌های زمانی بیشتر بر سری‌های زمانی ایستا متکی است. بنابراین در تجزیه و تحلیل سری‌های زمانی، برای استفاده از نظریه ایستایی لازم است سری‌های ناایستا به سری‌های ایستا تبدیل شود. هرگاه در یک سری زمانی تغییرات دوره‌ای حذف شده باشد و میانگین و واریانس تغییر منظمی را نشان دهد، آن سری را سری زمانی ایستا (Stationary time series) می‌نامند (نیرومند و بزرگ‌نیا، ۱۳۸۱). بنابراین بررسی ایستایی مورد نظر مستلزم بررسی میانگین و واریانس داده‌هاست.

پس از تجزیه و تحلیل روند و بررسی ایستایی، سعی خواهد شد با استفاده از روش باکس- جکینز یک مدل سری زمانی مناسب برای داده‌ها برازش داده شود. انواع مدل‌های سری زمانی عبارت‌اند از (نیرومند و بزرگ‌نیا، ۱۳۸۱):

الف) مدل اتورگرسیو AR(p) (Auto Regressive)

این مدل از مدل‌های متداول در آمار است و همان‌گونه که از نام آن مشخص است، روی جملات خود رگرسیون‌گیری را اعمال می‌کند. ساختار اصلی این مدل برای فرایند ایستای Z_t مطابق رابطهٔ (۱۱) یا (۱۲) است:

$$AR(p):\ Z_t = \varphi_1 Z_{t-1} + \ldots + \varphi_p Z_{t-p} + a_t \qquad (۱۱)$$

$$\varphi_p(B) Z_t = a_t \qquad (۱۲)$$

که در آن،

$$\varphi_p(B) = (1 - \varphi_1 B - \ldots + \varphi_p B^p) \qquad (۱۳)$$

$\varphi_1, \varphi_2, \ldots, \varphi_p$ ضرایب و پارامترهای مدل AR و a_t مقدار تصادفی و مستقل از زمان باقیمانده (noise) است که از توزیع نرمال با میانگین صفر تبعیت می‌کند. درصورتی که $\sum_{j=0}^{\infty} \varphi_j$ همگرا باشد، این مدل همواره وارون‌پذیر خواهد بود و برای ایستایی، باید ریشه‌های $\varphi_p(B) = 0$ خارج دایرهٔ واحد واقع شود.

ب) مدل میانگین متحرک MA(q) (moving Average)

در این مدل تغییر در زمان t از روی مقدار تصادفی همان لحظه به‌علاوهٔ q برابر مقدار تصادفی مربوط به زمان‌های قبل از t برآورد می‌شود. فرم عمومی این مدل با مرتبهٔ q به‌صورت رابطهٔ (۱۴) یا (۱۵) است.

$$MA(q):\ Z_t = a_t - \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q} \qquad (۱۴)$$

$$Z_t = \theta(B) a_t \qquad (۱۵)$$

که در آن،

$$\theta_p(B) = (1 - \theta_1 B - \ldots + \theta_p B^p) \qquad (۱۶)$$

در این مدل، $\theta_1, \theta_2, \ldots, \theta_p$ ضرایب و پارامترهای مدل است و چون $1 + \theta_1^2 + \ldots + \theta_q^2 < \infty$، لذا فرایند میا نگین متحرک، همواره ایستاست. اگر ریشه‌های $\theta(\mathrm{B}) = 0$ خارج دایرهٔ واحد باشد، فرایند میا نگین متحرک وارون‌پذیر خواهد بود.

ج) مدل خودهمبسته میانگین متحرک ARMA(p,q)

از ترکیب مدل Auto Regressive با مرتبهٔ p و مدل Moving Average با مرتبهٔ q، مدل ARMA با مرتبهٔ (p,q) حاصل می‌شود. ساختار کلی مدل مطابق رابطهٔ (۱۷) است.

$$X_t = \alpha_1 X_{t-1} + \ldots + \alpha_p X_{t-p} + Z_t - \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} \qquad (۱۷)$$

پارامترهای مربوط به این مدل نیز همانند مدل‌های MA و AR تعریف می‌شود.

به‌منظور تعیین پارامترهای p و q در فرایندهای اتورگرسیو و میانگین متحرک، نمودارهای توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) سری‌زمانی رسم و بررسی شد.

تابع خود همبستگی (Auto Correlation Function). برای فرایندی ایستا $[z_t]$، میانگین $E(z_t)=\mu$ و واریانس $Var(z_t)=E(z_t-\mu)^2=\sigma^2$ ثابت‌اند، و کواریانس‌های $\mathrm{cov}(z_t,z_s)$ توابعی فقط از اختلاف زمانی $|t-s|$ است. بنابراین، در این حالت کوواریانس بین z_t و z_{t+k} به‌صورت رابطه (۱۸)؛

$$\gamma_k = cov(z_t,z_{t+k}) = E(z_t-\mu)(z_{t+k}-\mu) \qquad (۱۸)$$

و همبستگی بین z_t و z_{t+k} به‌صورت رابطه (۱۹) تعریف می‌شود:

$$\rho_k = \frac{cov(z_t,z_{t+k})}{\sqrt{Var(z_t)}\sqrt{Var(z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \qquad (۱۹)$$

که در آن، $Var(z_t)=Var(z_{t+k})=\gamma_0$ است (نیرومند، ۱۳۷۱). در تحلیل سری‌های زمانی، γ_k که تابعی از k است را تابع اتوکواریانس و ρ_k را تابع خودهمبستگی با تأخیر k می‌نامند، زیرا کواریانس و همبستگی بین z_t و z_{t+k} فرایند را که فقط به‌وسیلهٔ k تأخیر زمانی از هم جدا شده نشان می‌دهد. معمولاً برای مقادیر k بزرگ‌تر از $N/4$ نباید ρ_k را محاسبه کرد (زیرا در این صورت بسیاری از اطلاعات سری زمانی از دست خواهد رفت).

برای تعبیر و تفسیر سری زمانی، نمودار تابع خود همبستگی، اغلب برای تأخیرهای نامنفی رسم می‌شود و این نمودار را همبستگی نگار می‌نامند.

تابع خودهمبستگی جزئی (Partial Auto Correlation Function). تابع مهم دیگر برای مطالعهٔ سری‌های زمانی تابع خودهمبستگی جزئی است. خودهمبستگی جزئی بین z_t و z_{t+k} برابر خودهمبستگی معمولی بین $(z_t-\hat{z}_t)$ و $(z_{t+k}-\hat{z}_{t+k})$ خواهد بود. لذا، اگر P_k خودهمبستگی جزئی بین z_t و z_{t+k} باشد (نیرومند، ۱۳۷۱):

$$P_k = \frac{\mathrm{Cov}[(z_t-\hat{z}_t),(z_{t+k}-\hat{z}_{t+k})]}{\sqrt{Var(z_t-\hat{z}_t)}\sqrt{Var(z_{t+k}-\hat{z}_{t+k})}} \qquad (۲۰)$$

برای شناسایی الگوی ARMA، باید حداقل تعداد N مشاهده وجود داشته باشد و تعداد ACF و PACF نمونه که باید محاسبه شود، در حدود $\frac{N}{4}$ است که بدین منظور کولوروگرام ρ_k در مقابل k ترسیم و خطوط محدودهٔ ۹۵ درصد احتمال از رابطهٔ $\pm\frac{2}{\sqrt{n}}$ روی آن مشخص می‌شود. با توجه به شکل کولوروگرام می‌توان به نوع سری زمانی پی برد (نیرومند و بزرگ‌نیا، ۱۳۸۱).

پس از شناسایی نوع و مرتبهٔ سری زمانی باید پارامترهای مربوط را برآورد کرد. برای برآورد پارامترهای یک الگوی سری زمانی می‌توان از روش‌های متداول برآورد نظیر روش گشتاورها، روش درست‌نمایی ماکزیمم و روش کمترین مربعات استفاده کرد. در روش گشتاور، گشتاورهای نمونه نظیر میانگین نمونه، واریانس نمونه و ACF نمونه جایگزین مقادیر موجود در فرمول می‌شود و معادلات حاصل حل می‌گردد. برای مثال در فرایند AR(P)، برای برآورد φ، ابتدا از $\rho_k=\varphi_1\rho_{k-1}+\varphi_2\rho_{k-2}+\ldots+\varphi_p\rho_{k-p}$ و $K>1$ برای محاسبهٔ دستگاه معادلات یول-والکر استفاده می‌شود.

$$Z_t = \varphi_1 Z_{t-1}+\ldots+\varphi_p Z_{t-p}+a_t \qquad (۲۱)$$

$$\begin{aligned}
\rho_1 &= \varphi_1 + \varphi_2\rho_1 + \varphi_3\rho_2 + \ldots + \varphi_p\rho_{p-1}\\
\rho_2 &= \varphi_1\rho_1 + \varphi_2 + \varphi_3\rho_1 + \ldots + \varphi_p\rho_{p-2}\\
&\vdots\\
\rho_p &= \varphi_1\rho_{p-1} + \varphi_2\rho_{p-2} + \varphi_3\rho_{p-3} + \ldots + \varphi_p
\end{aligned} \qquad (۲۲)$$

سپس اگر ρ_k جایگزین ρ_k و دستگاه معادلات خطی بالا حل شود، برآوردگرهای گشتاور $\varphi_1,\varphi_2,\ldots,\varphi_p$ به‌دست می‌آید. این برآوردگرها معمولاً برآوردگرهای یول-والکر نامیده می‌شود (نیرومند و بزرگ‌نیا، ۱۳۸۱).

۳. بررسی صحت مدل‌های مورد بررسی

الگوسازی سری‌های زمانی روشی تکراری است که با شناخت الگو و برآورد پارامترها شروع می‌شود. پس از برآورد پارامترها، باید کارآیی الگو را با بررسی اینکه آیا فرض‌های الگو برقرارند ارزیابی کرد. فرض اساسی این

است که سری باقیمانده‌ها $\{a_t\}$ سری تصادفی محض است، بنابراین بررسی درستی تشخیص الگو با توجه به تحلیل دقیق سری باقیمانده‌ها انجام می‌شود (نیرومند و بزرگ‌نیا، ۱۳۸۱). در تحلیل سری‌های زمانی ممکن است چندین الگوی مناسب وجود داشته باشد که برای نشان دادن یک مجموعه دادهٔ معلوم استفاده می‌شوند.

برای تحلیل سری‌های زمانی یا حتی یک مجموعه داده معلوم، ممکن است چند الگوی قابل قبول وجود داشته باشد. در ادامه چند ملاک انتخاب الگوی مبتنی بر باقیمانده‌ها معرفی می‌شود.

الف) محک آکائیک AIC

برای ارزیابی کیفیت برازش الگو، آکائیک (۱۹۷۳ و ۱۹۷۴) ملاکی اطلاعاتی مطرح کرد. این ملاک را AIC (ملاک اطلاعاتی آکائیک) نامیده‌اند و به‌صورت رابطه (۲۳) تعریف می‌شود.

$$AIC(m) = -2\log L + 2m \qquad (۲۳)$$

که در آن، $m = (p+q)$ تعداد پارامترهای الگو و L تابع درست‌نمایی مدل است. بهترین مرتبهٔ الگو، به‌وسیلهٔ مقدار m که تابعی از p و q است، انتخاب می‌شود. در این صورت باید AIC(m) مینیمم باشد.

ب) محک BIC

شیباتا (۱۹۷۶) نشان داده است که ملاک AIC مرتبهٔ بالاتری از اتورگرسیو را بهترین الگو برآورد می‌کند. اخیراً آکائیک (۱۹۷۸ و ۱۹۷۹) تعمیم بیزی روش AIC مینیمم را به‌دست آورده که BIC نامیده می‌شود و به‌صورت رابطه (۲۴) است.

$$BIC(m) = -2log\hat{\sigma}_a^2 - (n-m)ln(1-m/n) + mlnn + mln[(\frac{\hat{\sigma}_z^2}{\hat{\sigma}_a^2} - 1)/m] \qquad (۲۴)$$

که در آن، $\hat{\sigma}_a^2$ برآورد درست‌نمایی ماکزیمم σ_a^2، m تعداد پارامترها و $\hat{\sigma}_z^2$ واریانس نمونهٔ سری است. با مطالعهٔ شبیه‌سازی آکائیک (۱۹۷۸) ادعا می‌کند که BIC احتمال اینکه مرتبهٔ بالاتری از اتورگرسیو را بهترین الگو برآورد کند، کم می‌شود. بهترین الگو با استفاده از این محک نیز در صورتی حاصل می‌شود که مقدار BIC(m) مینیمم باشد.

ج) ملاک بیزی شوارتز SBC

مشابه BIC آکائیک، شوارتز (۱۹۷۸) ملاک بیزی انتخاب الگویی (۲۵ رابطه) را پیشنهاد می‌کند که SBC (ملاک بیزی شوارتز) نامیده می‌شود.

$$SBC(m) = -2\log L + m\ln(n) \qquad (۲۵)$$

که در آن، n تعداد مشاهدات است. این محک نیز باید مقدارش مینیمم باشد تا بهینه‌ترین مدل انتخاب شود.

به‌منظور بررسی عملکرد مدل‌های توسعه‌یافته در پیش‌بینی، معیارهای میانگین مطلق خطا و مجذور میانگین مربعات خطا به ترتیب بر پایهٔ روابط (۲۶) و (۲۷) استفاده می‌شود.

$$RSME = \sqrt{\frac{1}{n}\sum_{1}^{n}(X_o - X_p)^2} \qquad (۲۶)$$

$$MAE = \frac{1}{n}\sum_{1}^{n}\left|X_o - X_p\right| \qquad (۲۷)$$

که در آن، X_o مقادیر مشاهدات، X_p مقادیرمحاسباتی و n تعداد مشاهدات است.

۴. پیش‌بینی بلندمدت سطح آب

از مهم‌ترین اهداف تحلیل سری زمانی می‌توان به پیش‌بینی مقادیر آیندهٔ آن اشاره کرد. حتی اگر هدف نهایی به الگو در آوردن سری‌های زمانی برای کنترل یک سیستم باشد، عملیات آن معمولاً مبتنی بر پیش‌بینی است.

پس از انتخاب مدل بهینه و برآورد پارامترهای مدل، پیش‌بینی داده‌ها با مدل مورد نظر صورت می‌پذیرد.

مدل عددی ARMA(1,1) بصورت رابطه (۲۸) است (نیرومند، ۱۳۷۱).

$$(1-\phi B)(z_t - \mu) = (1-\theta B)a_t \qquad (۲۸)$$

پیش‌بینی l مرحلهٔ بعد z_{n+l} یعنی $\hat{z}_n(l)$ به‌صورت رابطه (۲۹) است.

$$\hat{z}_n(l) = \mu + \phi^l(z_n - \mu) - \phi^{l-1}\theta a_n \ , \ l \geq 2 \qquad (۲۹)$$

همچنین رابطه (۲۹) برای پیش‌بینی AR و MA تعمیم‌پذیر است، به‌طوری که پیش‌بینی یک مرحله بعد برای AR(1) به‌صورت $\hat{z}_n(l) = \mu + \phi^l(z_n - \mu)$ و

پیش‌بینی یک مرحله بعد برای MA(1) به‌صورت $\hat{z}_n(l) = \mu - \theta a_n$ است (نیرومند، ۱۳۷۱).

پس از پیش‌بینی سری زمانی، برای رسیدن به ارتفاع لحظه‌ای سطح دریا، باید فرکانس‌هایی را که در ابتدا با روش فوریه از روی داده‌ها حذف شدند با تبدیل عکس فوریه به داده‌ها برگرداند.

۵. ارائهٔ نتایج در منطقهٔ مورد نظر

داده‌های مورد استفاده در این تحقیق، داده‌های ماهوارهٔ ارتفاع‌سنجی Jason-2 است. مسیر این ماهواره در منطقهٔ دریاچهٔ خزر در شکل ۱ نمایش داده شده است. این داده‌ها شامل ارتفاع سطح لحظه‌ای آب دریا از بیضوی مرجع است که برای آثار محیطی (آثار مرطوب تروپوسفری، خشک تروپوسفری و یونسفر) و ژئوفیزیکی (جزر و مد اقیانوس، جزر و مد زمین و جزر و مد قطب) تصحیح شده است. تغییرات سطح لحظه‌ای آب دریای خزر در نقطهٔ مورد مطالعه در بازه زمانی بین سال‌های ۲۰۰۸ تا ۲۰۱۲ به شکل سری زمانی در شکل ۲ نشان داده شده است. موقعیت نقطهٔ مورد نظر را می‌توان در شکل ۱ مشاهده کرد. در این مطالعه از مشاهدات سطح لحظه‌ای آب در فاصلهٔ زمانی بین سال‌های ۲۰۰۸ تا ۲۰۱۲ استفاده شده که شامل ۱۵۰ سیکل مشاهده است. ۱۰۰ سیکل از این تعداد برای مدل‌سازی به‌کارگرفته شده و از ۵۰ سیکل آخر برای اعتبارسنجی محاسبات استفاده شده است.

مدل‌سازی به‌وسیلهٔ روش‌های آماری سری زمانی مستلزم ایستا بودن داده‌های مورد نظر است. بنا به تعریف روند که سیر طبیعی سری زمانی در درازمدت است، در شکل ۲ نیز روند کاهش ارتفاع در طول زمان مشاهده می‌شود. برای ایستا بودن داده‌ها لازم است این روند از روی داده‌ها حذف شود. همچنین، با استفاده از تحلیل در قلمرو فرکانس می‌توان نوسانات سری زمانی را بر حسب رفتار سینوسی در فرکانس‌های مختلف بیان کرد.

بدین منظور، با استفاده از تبدیل فوریه، مشاهدات از فضای زمان به فضای فرکانس برده می‌شود و با استفاده از آنالیز فوریه فرکانس‌های موجود در مشاهدات به‌دست می‌آید. در شکل ۳ فرکانس مشاهدات به‌همراه بزرگی آن‌ها مشاهده می‌شود.

همان‌طور که ذکر شد درایهٔ اول برای یافتن فرکانس موجود در داده‌ها قابل استفاده نیست. بزرگ‌ترین عدد موجود در فرکانس‌ها نشان‌دهندهٔ تأثیرگذارترین مؤلفه روی بردار مشاهدات است. با استفاده از سرشکنی کمترین مربعات فرکانس‌ها را معلوم می‌گیریم و با توجه به رابطه (۳۰) ضرایب هر فرکانس به‌دست می‌آید.

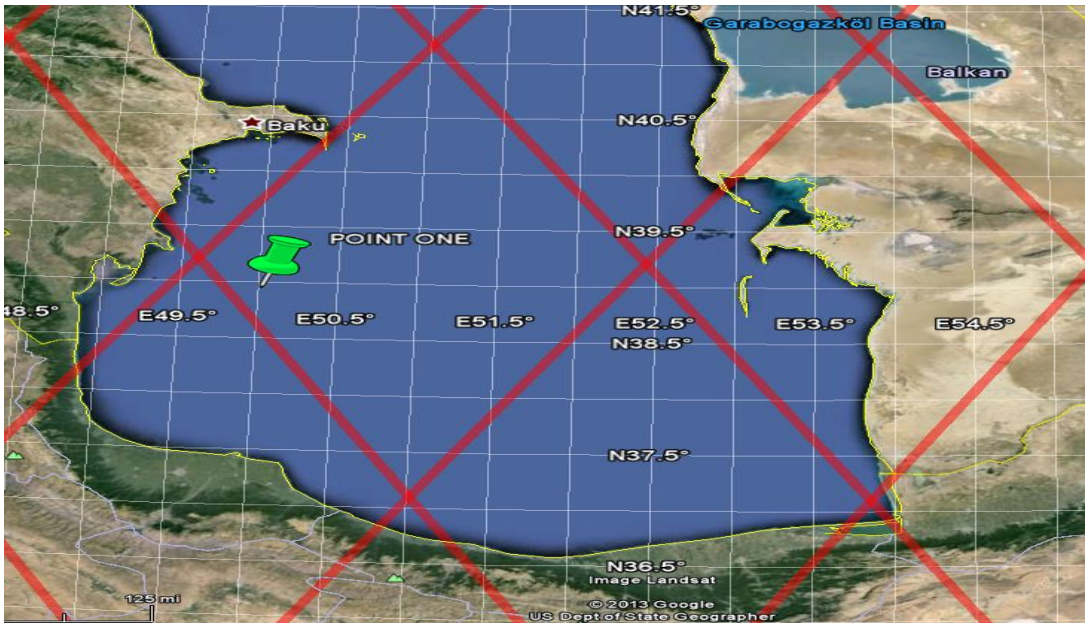
$$y\ t = a_0 + b_0 t + \sum_{i=o}^{n} a_i \cos\ \omega_i t\ + b_i \sin\ \omega_i t \qquad (۳۰)$$

که در آن، a_i و b_i ضرایب فوریه، a_0 بیانگر سطح متوسط دریا و b_0 نرخ کاهش سطح آب است. برای به‌دست آوردن فرکانس‌ها نیز از روش‌هایی که قبلاً ذکر شد استفاده می‌شود. طبق نظریهٔ نمونه‌برداری اگر فاصلهٔ زمانی داده‌ها ΔT باشد، آن‌گاه می‌توان فرکانس‌های با پریود بزرگ‌تر از $2\Delta T$ را از سری زمانی استخراج کرد. بنابراین با توجه به سری زمانی با پریود ۹/۹۱۵ روزه برای نقطهٔ برداشت شده در سطح دریا، می‌توان فرکانس‌های امواج با پریودهای بیشتر از ۱۹/۸۳ روز را مشخص کرد.

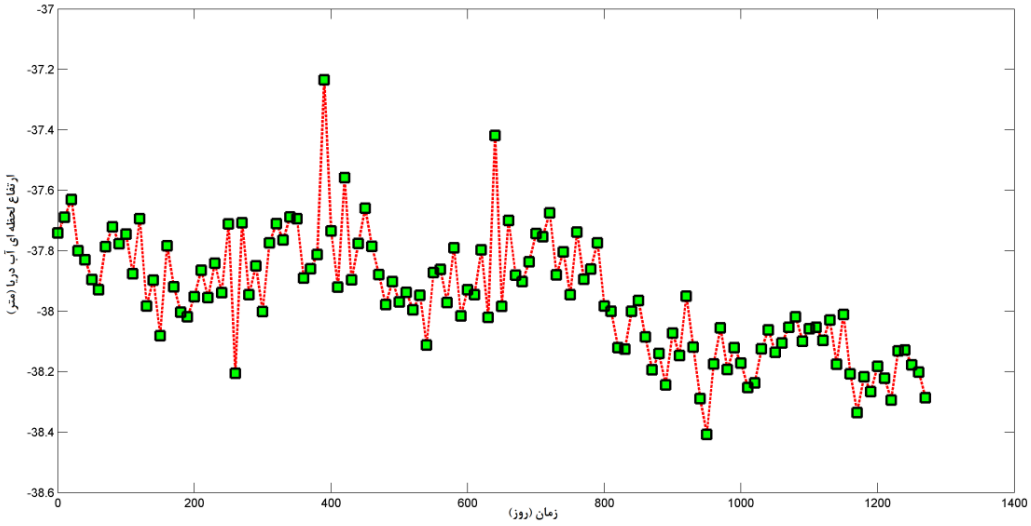
سپس، اثر امواجی که دامنهٔ آن‌ها از دقت مشاهدات بالاتر است حذف می‌شود. در شکل ۴ سری زمانی ارتفاعات و سیگنال ناشی از حذف اثر امواج مورد نظر نمایش داده شده است. نتایج تفاضل ارتفاعات از سیگنال ناشی از حذف اثر امواج مورد نظر در شکل ۵ نمایش داده شده است.

جدول ۱. ویژگی‌های نظری نمودارهای ACF و PACF در سری‌های‌زمانی مانا (نیرومند و بزرگ‌نیا، ۱۳۸۱).

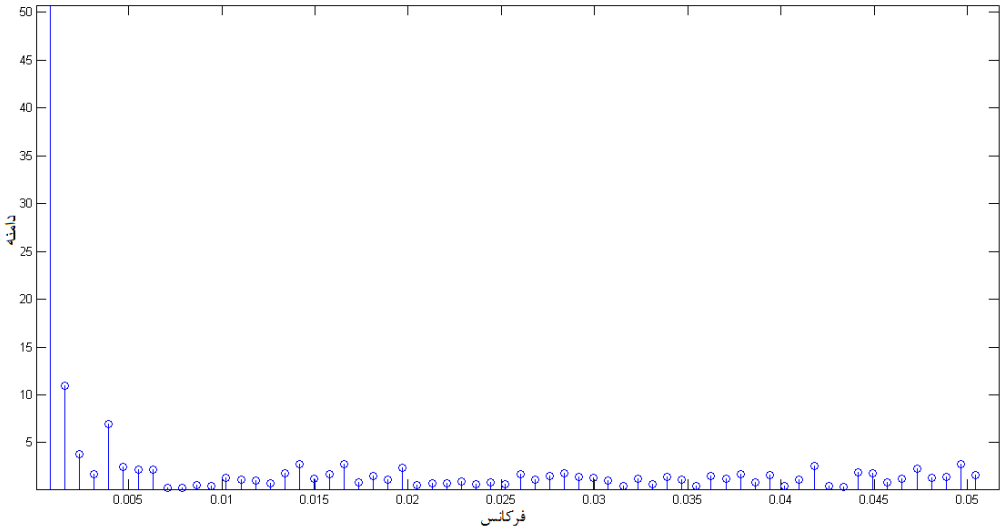
فرایند	ACF	PACF
AR (p)	به‌صورت نمایی یا موج سینوسی میرا به سمت صفر میل می‌کند.	بعد از تأخیر p قطع می‌شود.
MA (q)	بعد از تأخیر q قطع می‌شود.	به‌صورت نمایی یا موج سینوسی میرا به سمت صفر میل می‌کند.
ARMA (p,q)	بعد از تأخیر q-p به سمت صفر میل می‌کند.	بعد از تأخیر p-q به سمت صفر میل می‌کند.



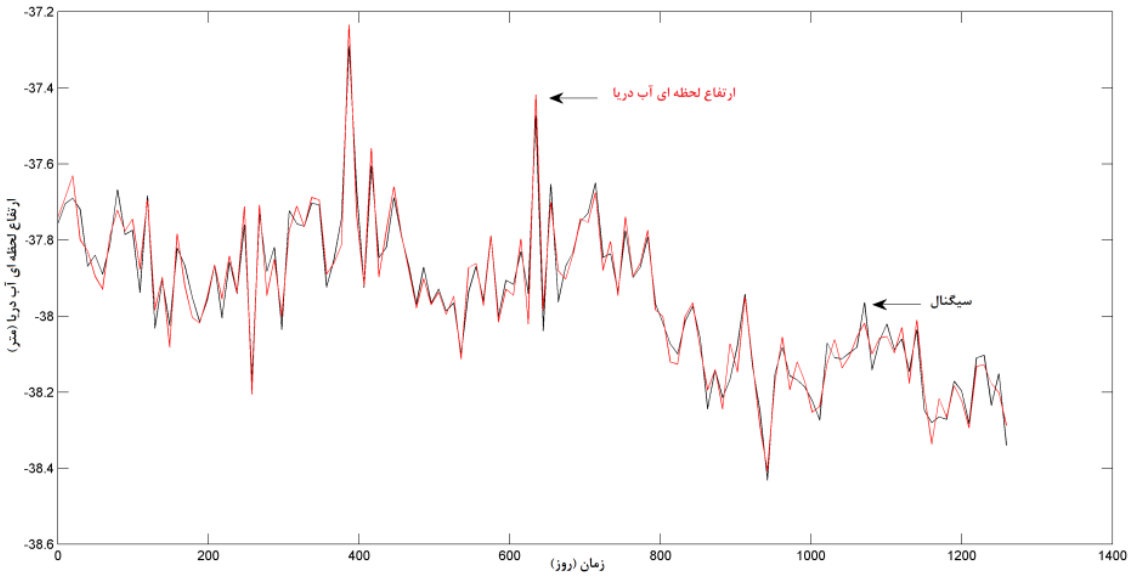
شکل ۱. نمایش مسیر حرکت ماهوارهٔ ارتفاع‌سنجی Jason-2 در دریای خزر و محل نقطهٔ مورد بررسی.



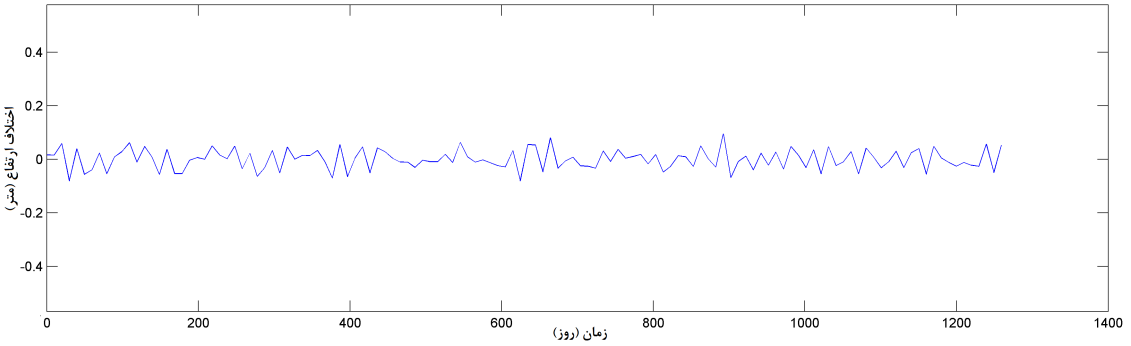
شکل ۲. تغییرات سطح لحظه‌ای آب دریا در بازه زمانی بین سال‌های ۲۰۰۸ تا ۲۰۱۲ برای نقطهٔ مورد نظر.



شکل ۳. پریودوگرام سری زمانی.



شکل ۴. نمایش سری زمانی ارتفاعات و سیگنال ناشی از حذف اثر امواج مورد نظر.



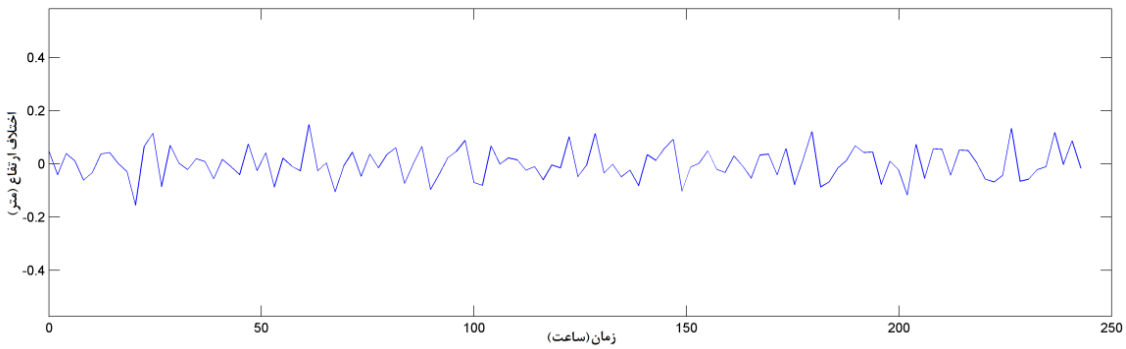
شکل ۵. نتایج تفاضل سری زمانی ارتفاعات از سیگنال ناشی از حذف فرکانس‌های اصلی برای پریودهای بیش از ۱۹/۸۳ روز.

اگر مشاهداتی که ۹/۹۱۵ روز با هم فاصله دارند با توجه به ساعت انجام مشاهده مرتب شود، در می‌یابیم که مشاهدات طول شبانه‌روز را پوشش می‌دهد. در واقع، می‌توان مشاهدات را در طی یک شبانه‌روز ۲۴ ساعته مرتب کرد. بنابراین نمونه‌برداری کامل از موج خواهیم داشت، به‌قسمی که فاصلهٔ زمانی این نمونه‌برداری برابر ۲/۰۴ ساعت است. طبق نظریهٔ نمونه‌برداری با چنین فاصلهٔ زمانی نمونه‌برداری قادر خواهیم بود امواج با پریود چهار ساعت و بیشتر را مدل‌سازی کنیم. پس از انجام مجدد آنالیز فوریه و سرشکنی کمترین مربعات برای داده‌های ۲/۰۴ ساعته نتایج به‌صورت شکل ۶ است.

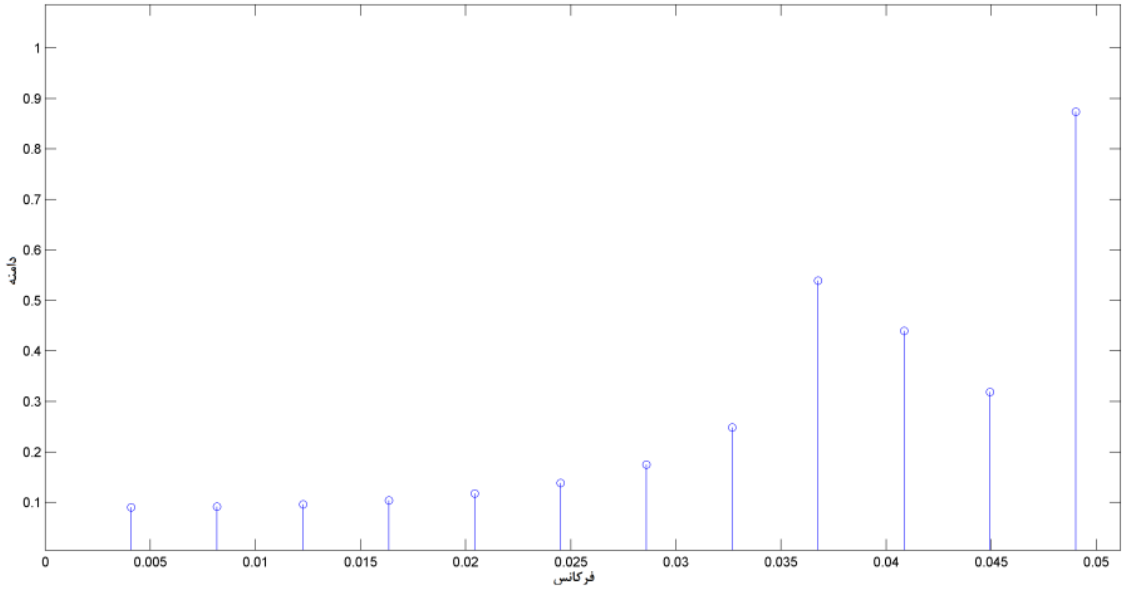
با استفاده از آنالیز فوریه فرکانس‌های امواج با پریودهای بیش از ۴/۰۸ ساعت در سری زمانی اخیر مشخص می‌شود. پریودوگرام مربوط به این سری زمانی در شکل ۷ نمایش داده شده است. سپس، اثر امواجی که دامنهٔ آن‌ها از دقت مشاهدات بالاتر است حذف می‌شود. در شکل ۸ سری زمانی ارتفاعات و سیگنال ناشی از حذف اثر امواج مورد نظر نمایش داده شده است.

نتایج تفاضل ارتفاعات از سیگنال ناشی از حذف اثر امواج مورد نظر با پریودهای بیش از ۴/۰۸ ساعت در شکل ۹ مشاهده می‌شود.

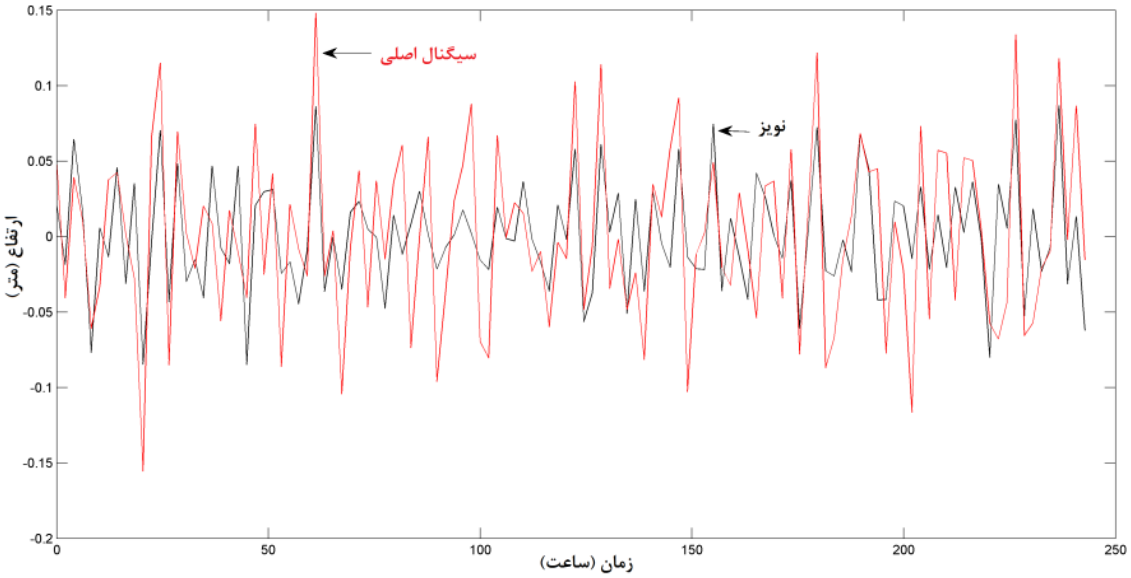
بررسی فرکانس‌های پریودیک نشان می‌دهد که این فرکانس‌ها با مؤلفه‌های اصلی جزرومدی شامل SSA، M6، S4، S2، M2، MF، MM و SA همخوانی دارد.



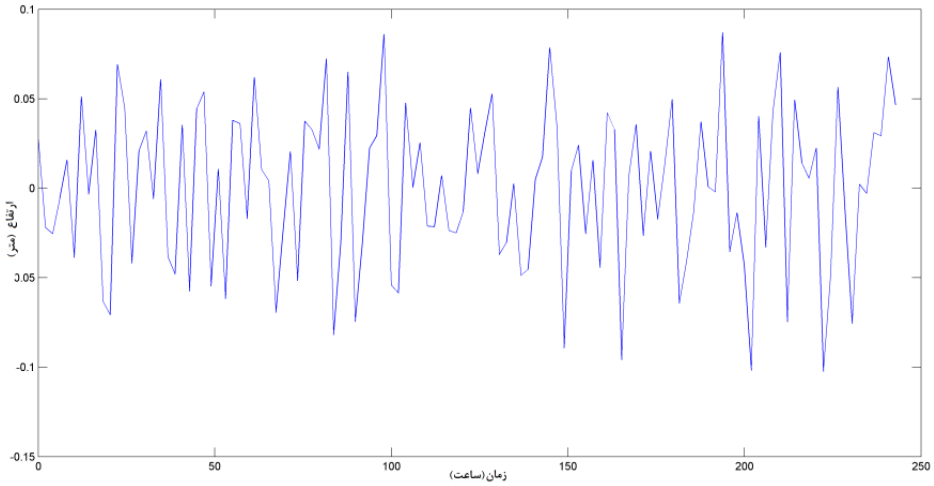
شکل ۶. سری زمانی ۲/۰۴ ساعته ارتفاع نسبت به زمان بر حسب ثانیه.



شکل ۷. فرکانس امواج موجود در سری زمانی ۲/۰۴ ساعته.



شکل ۸. نمایش سری زمانی ۲/۰۴ ساعت و سیگنال ناشی از حذف اثر امواج مورد نظر برای پریودهای بیش از ۴/۸ ساعت.



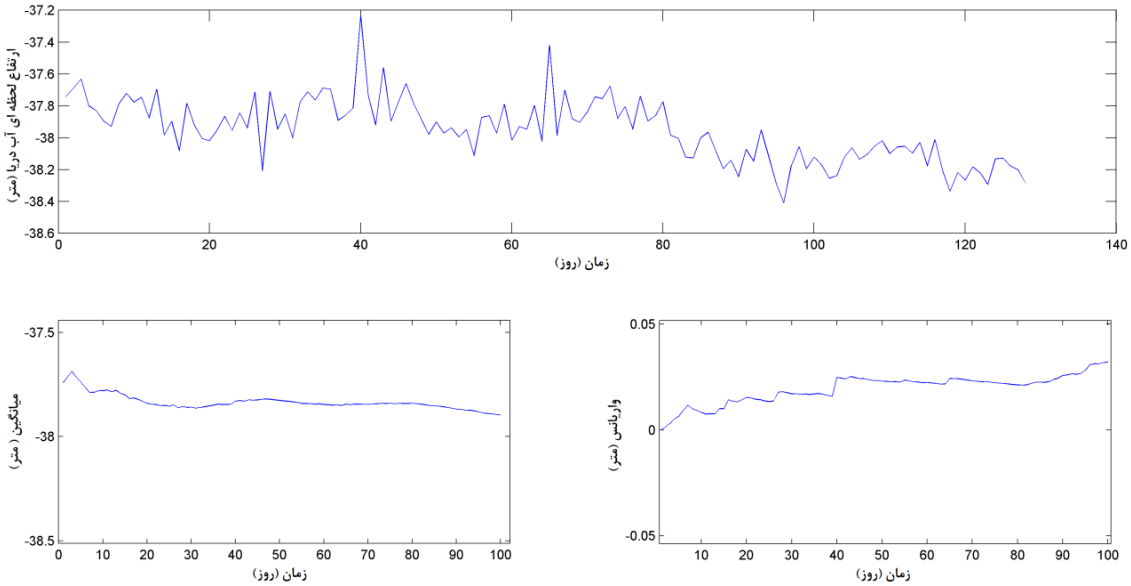
شکل ۹. نتایج تفاضل سری زمانی ۲/۰۴ ساعت و سیگنال ناشی از حذف اثر امواج مورد نظر در پریودهای بیش از ۴/۰۸ ساعت.

همان‌گونه که گفتیم ، مدل‌سازی سری زمانی به‌دست آمده با استفاده از تجزیه و تحلیل به روش سری زمانی مستلزم ایستا بودن سری مورد نظر است. با توجه به تعریف، میانگین و واریانس برای سری زمانی اولیهٔ ارتفاع از سطح دریا نسبت به زمان محاسبه شده است که به علت نداشتن میانگین و واریانس ثابت، داده‌های مورد نظر داده‌های ایستایی نیست. در شکل ۱۰ می‌توان میانگین و واریانس سری زمانی اولیهٔ ارتفاع از سطح دریا نسبت به زمان را مشاهده کرد.

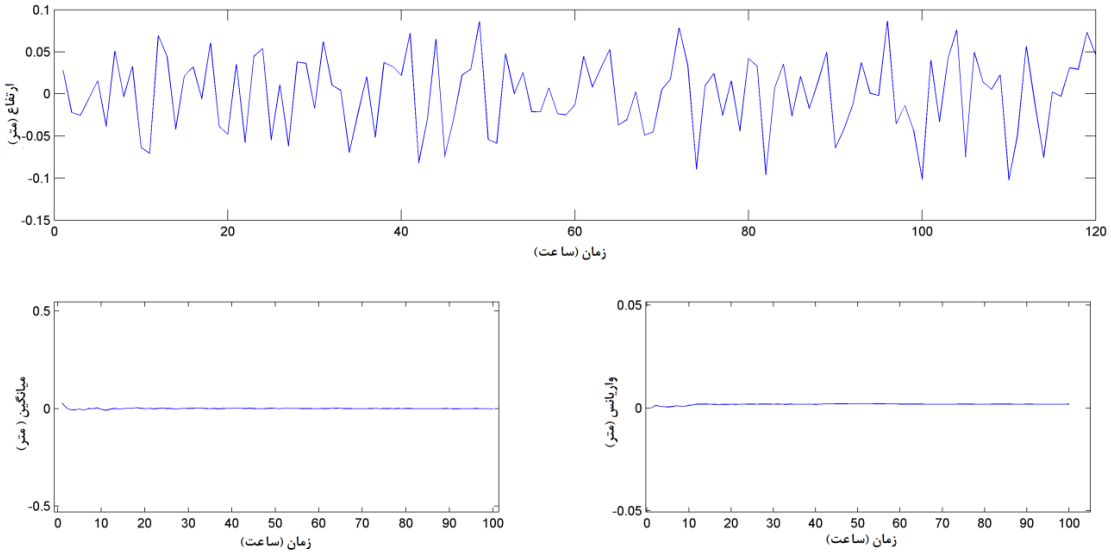
محاسبهٔ میانگین و واریانس داده‌های نهایی حذف شده از طریق آنالیز فوریه، نشان‌دهندهٔ ایستا بودن این داده‌هاست. این کمیت‌ها را می‌توان در شکل ۱۱ مشاهده کرد.

در مرحلهٔ شناسایی الگو، نمودارهای خودهمبستگی ACF و خودهمبستگی جزئی PACF داده‌ها بررسی شد که میزان همبستگی زمانی داده‌ها را نشان می‌دهد. شکل ۱۲ و ۱۳ به‌ترتیب نمودارهای ACF و PACF مربوط به سری زمانی را نشان می‌دهد.

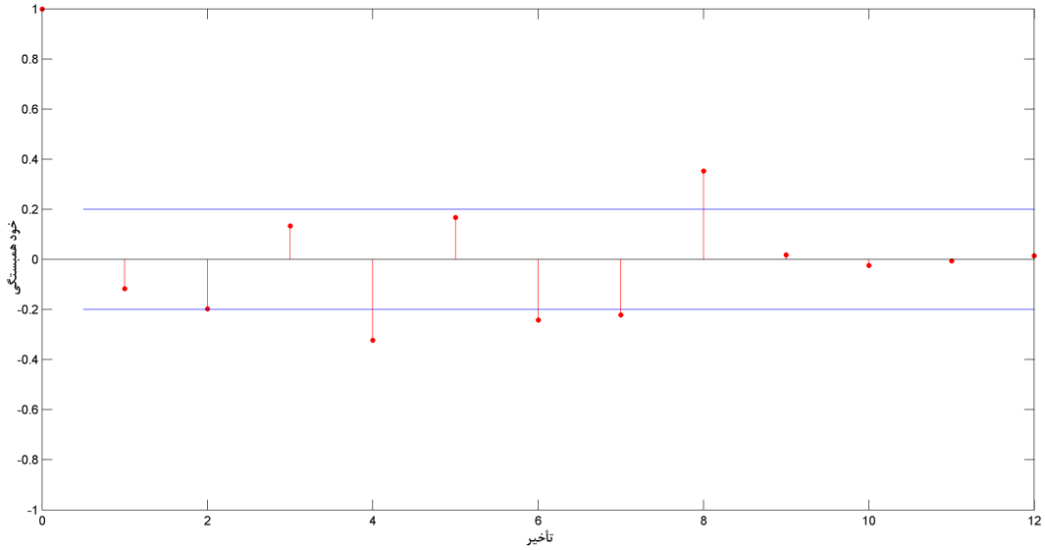
با توجه به مطالبی که در جدول ۱ ذکر شد، برای نمودار ACFسه حالت ممکن است رخ دهد: ۱. میرا شدن به‌صورت نمایی و یا موج سینوسی، ۲. قطع شدن پس از تأخیر p، ۳. به سمت صفر میل کردن بعد از تأخیر $p-q$.



شکل ۱۰. میانگین و واریانس سری زمانی اولیهٔ ارتفاع از سطح دریا نسبت به زمان.



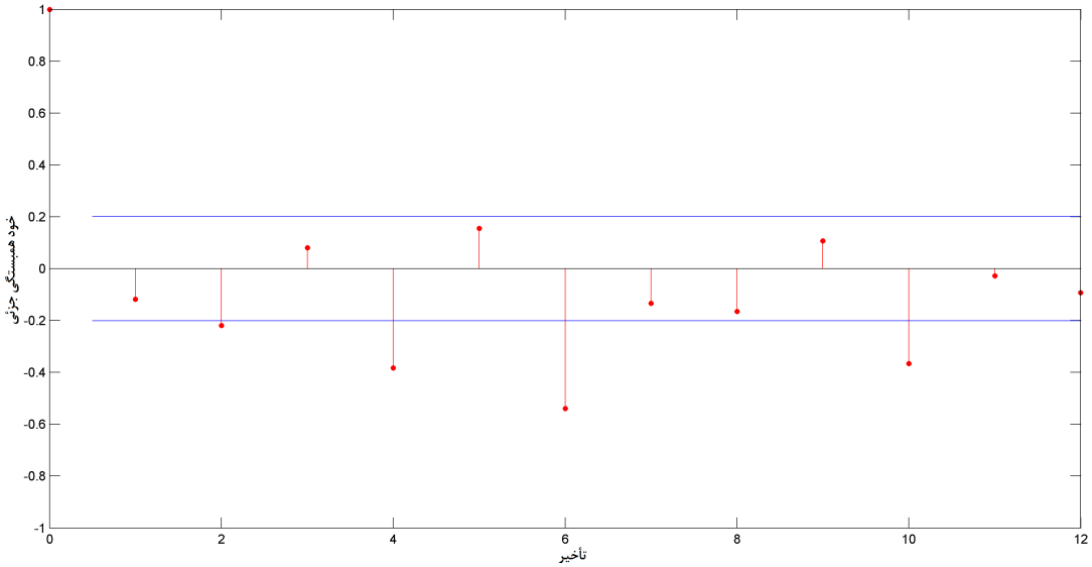
شکل ۱۱. میانگین و واریانس سری زمانی نهایی.



شکل ۱۲. نمودار تابع خودهمبستگی ACF.

اما با توجه به شکل ACF سری مورد نظر، می‌توان دریافت که این نمودار به‌صورت ترکیبی از موج نمایی و سینوسی میرا می‌شود و به سمت صفر میل می‌کند و قطع نمی‌شود که مدل AR(p) را پیشنهاد می‌کند. اما از سه موردی که برای نمودار PACF در جدول ۱ ذکر شد، رفتار این نمودار به‌صورت میراشدن به‌صورت نمایی یا موج سینوسی نیست. با بررسی دقیق متوجه می‌شویم که این نمودار دارای چهار مقدار معنادار در تأخیرهای ۲، ۴، ۶ و ۱۰ است. مقادیر همبستگی جزئی در این تأخیرها از خط اطمینان یا همان خط محدودهٔ ۹۵ درصد احتمال بیشتر شده است که در تأخیر ۲ مقدار ناچیزی است. بنابراین سه مورد دیگر بررسی شد. سه تأخیر مورد نظر امکان یک الگوی AR(4) یا AR(6) یا حتی AR(10) را تأیید می‌کند.

برای انتخاب بهترین مدل از بین این سه مدل اتورگرسیو، محک‌های آکائیک AIC، BIC و بیزی شوارتز SBC آن‌ها بررسی شد که مقادیر آن در جدول ۲ آمده است. برای انتخاب بهترین مدل در برازش بیشتر با سری مورد نظر، باید مقدار هر سه محک آن مینیمم باشد که همان‌طور که در جدول ۲ مشاهده می‌شود، مدل AR(6) بهترین الگوی سری زمانی برای سری ایستا تعیین شد.



شکل ۱۳. نمودار تابع خودهمبستگی جزئی PACF.

جدول ۲. محک آکائیک AIC، BIC و ملاک بیزی شوارتز SBC، میانگین مطلق خطا MAE و مجذور میانگین مربعات خطا RMSE.

Model	RMSE (m)	MAE (m)	AIC	BIC	SBC
AR(1)	۰/۰۱۵۴۰۷	۰/۰۰۴۶۷۶	-۶۱۹/۸۸۸	-۶۱۶/۲۸۸	-۶۱۷/۲۸۳
AR(2)	۰/۰۱۵۰۰	۰/۰۰۴۶۹۹	-۶۲۲/۷۲۷	-۶۱۶/۹۲۳	-۶۱۷/۵۱۶
AR(3)	۰/۰۱۵۴۹۸	۰/۰۰۴۷۰۲	-۶۲۱/۳۷۱	-۶۱۳/۸۹۷	-۶۱۳/۵۵۶
AR(4)	۰/۰۱۵۷۰۵	۰/۰۰۴۸۰۹	-۶۳۳/۲۵۱	-۶۲۴/۴۵۶	-۶۲۲/۸۳۰
AR(5)	۰/۰۱۵۳۰۳	۰/۰۰۴۶۱۶	-۶۳۳/۳۲۶	-۶۲۳/۴۷۵	-۶۲۰/۳۰۰
AR(6)	۰/۰۱۴۱۵۸۰	۰/۰۰۳۷۳	-۶۵۸/۶۰۸	-۶۴۷/۹۱۱	-۶۴۲/۹۷۷
AR(7)	۰/۰۱۴۷۳۴	۰/۰۰۳۹۹۴	-۶۵۸/۱۳۱	-۶۴۶/۷۶۷	-۶۳۹/۸۹۵
AR(8)	۰/۰۱۴۶۸۵۰	۰/۰۰۳۸۹۱	-۶۵۶/۵۸۶	-۶۴۴/۷۰۹	-۶۳۵/۷۴۵
AR(9)	۰/۰۱۴۵۲۱	۰/۰۰۳۹۳۵	-۶۵۶/۸۴۵	-۶۴۴/۵۹۱	-۶۳۳/۳۹۹
AR(10)	۰/۰۱۴۳۲۱	۰/۰۰۳۸۸۲	-۶۶۰/۸۴۷	-۶۴۸/۳۹۹	-۶۳۴/۷۹۵
AR(11)	۰/۰۱۴۲۹۸	۰/۰۰۳۸۶۵	-۶۵۸/۴۵۷	-۶۴۶/۲۰۵	-۶۳۰/۲۰۰
AR(12)	۰/۰۱۴۲۸۵	۰/۰۰۳۸۶۵	-۶۵۶/۸۶۰	-۶۴۴/۱۶۷	-۶۲۵/۵۹۸
MA(1)	۰/۰۱۵۳۲۸	۰/۰۰۴۵۸۶	-۶۱۹/۹۱۳	-۶۱۶/۳۱۳	-۶۱۷/۳۰۸
MA(2)	۰/۰۱۵۲۷۴	۰/۰۰۴۵۷۲	-۶۲۱/۷۰۳	-۶۱۵/۸۹۹	-۶۱۶/۴۹۳

در بررسی الگوهای مختلف سری‌های زمانی برای سری ایستا، مقادیر $\varphi_1,\varphi_2,\ldots,\varphi_p$ و $\theta_1,\theta_2,\ldots,\theta_p$ در الگوی ARMA(p,q) بزرگ‌تر از ۱ به‌دست آمد که با فرض ایستایی سری در تناقض است، بنابراین الگوی ARMA(p,q) مدل مناسبی نیست.

با توجه به نمودارهای همبستگی و خودهمبستگی جزئی، همچنین مقادیر محک‌های مختلف برای مدل MA(q)، این مدل نیز مدل ایده‌آلی مطرح نیست.

در مرحلهٔ برازش الگو و برآورد پارامترهای $\varphi_1,\varphi_2,\ldots,\varphi_6$ از روش گشتاورها استفاده می‌شود که نیاز به مقادیر توابع خودهمبستگی ACF و خودهمبستگی جزئی PACF است.

با توجه به کمترین مقدار میانگین مطلق خطا و مجذور میانگین مربعات خطا برای مدل‌های مختلف AR(p)، در این بررسی نیز AR(6) بهترین الگو برای برازش روی سری زمانی ایستاست. رابطهٔ (۳۱) برای الگوی AR(6) ارائه شده است:

$$Z_t = 0.02702Z_{t-1} - 0.4497Z_{t-2} + 0.1120Z_{t-3} - 0.4801Z_{t-4} + 0.1354Z_{t-5} - 0.4845Z_{t-6} + a_t \quad (۳۱)$$

که در آن،

$$\varphi_1 = 0.02702,\ \varphi_2 = -0.4497,\ \varphi_3 = 0.1120,\ \varphi_4 = -0.4801,\ \varphi_5 = 0.1354,\ \varphi_6 = -0.4845 \quad (۳۲)$$

ضرایب تمام مدل‌ها، همچنین بهترین مدل، AR(6)، در جدول ۳ آورده شده است.

با توجه به اینکه فرایند AR ایستاست، اگر و فقط اگر:

$$\varphi_2 - \varphi_1 < 1 \quad ،\ ،\quad \varphi_2 + \varphi_1 < 1 \quad |\varphi_2| < 1 \quad (۳۳)$$

این شرط برای ضرایب به‌دست آمدهٔ مدل AR برقرار است. در نتیجه این فرایند ایستاست.

پیش‌بینی با الگوی AR(6) مستلزم وجود سری تصادفی $\{a_t\}$ است که دارای میانگین صفر و واریانس σ_a^2 است و:

$$\sigma_a^2 = \sigma_z^2\ 1 - \varphi_2\rho_1 - \ldots - \varphi_6\rho_6 \quad (۳۴)$$

و مقدار واریانس برابر $\sigma_a^2 = 0.0002$ محاسبه شد. با توجه به کوچک بودن مقدار $\{a_t\}$ که دارای میانگین صفر و واریانس بسیار کوچک $\sigma_a^2 = 0.0002$ است، جایگذاری میانگین به‌جای اجزای سری $\{a_t\}$ اختلاف فاحشی را در محاسبات ایجاد نمی‌کند. در نتیجه رابطهٔ پیش‌بینی سری به‌صورت رابطه (۳۵) است.

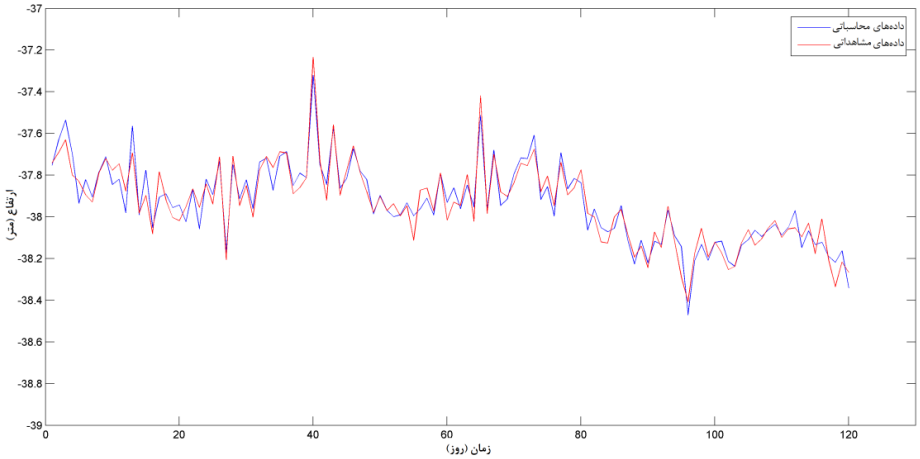
$$Z_t = 0.02702Z_{t-1} - 0.4497Z_{t-2} + 0.1120Z_{t-3} - 0.4801Z_{t-4} + 0.1354Z_{t-5} - 0.4845Z_{t-6} \quad (۳۵)$$

حال با استفاده از عکس تبدیل فوریه فرکانس‌های حذف شده به سری اضافه می‌شود تا بتوان ارتفاع سطح لحظه‌ای آب دریا را پیش‌بینی کرد.

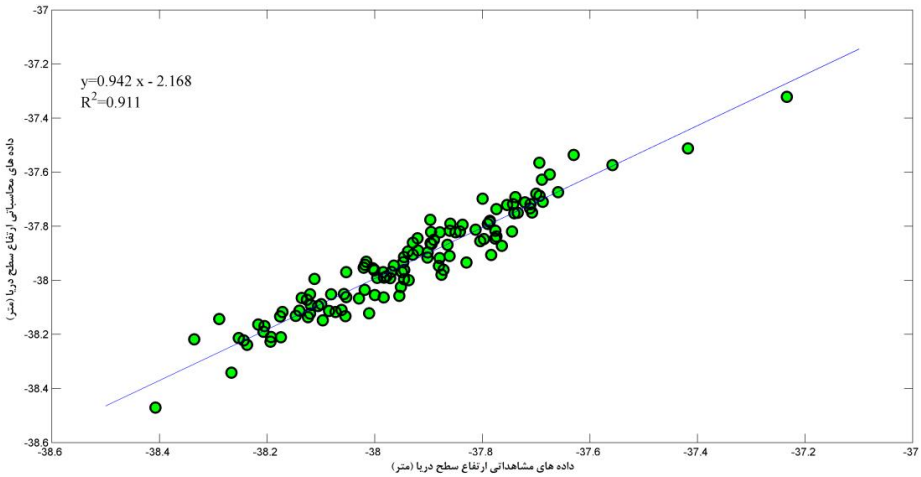
جدول ۳. مقادیر پارامترهای مدل‌های سری زمانی.

پارامترهای مدل	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6	φ_7	θ_1	θ_2
AR(1)	−۰/۱۱۶۸								
AR(2)	−۰/۱۴۱۸	−۰/۲۱۴۳							
AR(3)	−۰/۱۲۳۵	−۰/۲۰۲۲	۰/۰۸۵۲						
AR(4)	−۰/۰۹۲۸	−۰/۲۷۵۲	۰/۰۴۰۶	−۰/۳۶۰۵					
AR(5)	−۰/۰۳۰۹	−۰/۲۸۲۱	۰/۰۸۷۹	−۰/۳۴۴۵	۰/۱۷۱۷				
AR(6)	۰/۰۵۳۵	−۰/۴۵۱۵	۰/۱۳۱۱	−۰/۴۸۳۲	۰/۱۵۶۵	−۰/۴۹۱۶			
AR(7)	۰/۰۱۰۰	−۰/۴۳۷۷	۰/۰۸۸۴	−۰/۴۷۱۶	۰/۱۱۶۶	−۰/۴۸۶۹	−۰/۰۸۸۴		
MA(1)								۰/۶۵۳۰	
MA(2)								۰/۴۷۰۹	۰/۲۴۷۳

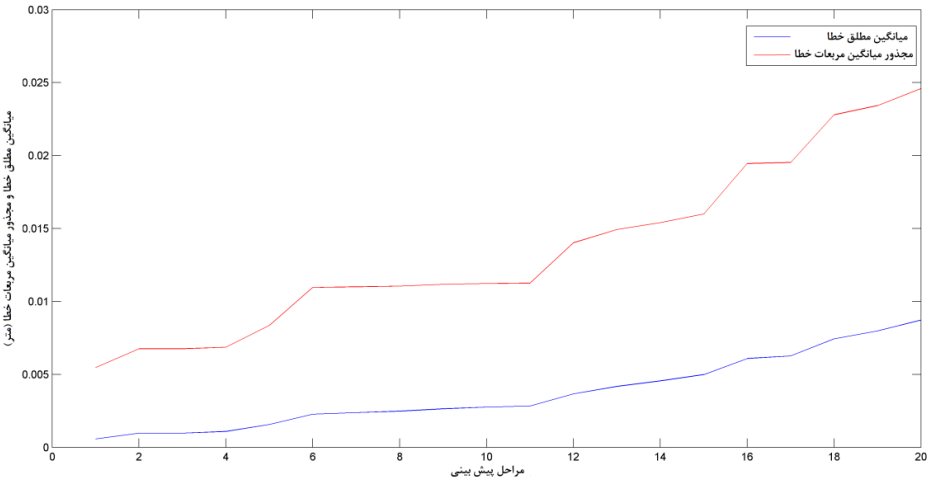
همان‌طور که در شکل ۱۴ مشاهده می‌شود، سری محاسباتی برازش مناسبی با داده‌های مشاهداتی دارد. شکل ۱۵ نیز همبستگی بین مقادیر محاسباتی و مشاهداتی را نشان می‌دهد. مقدار ضریب همبستگی $R = 0.911$ دقت بالای مدل AR(6) برازش شده به داده‌های ارتفاع لحظه‌ای آب دریای خزر را به خوبی نمایش می‌دهد. شکل ۱۶ و جدول ۴ مقادیر RMSE و MAE پیش‌بینی پس از گذشت هر مرحله از پیش‌بینی را نشان می‌دهد.



شکل ۱۴. نمودار مقایسهٔ داده‌های محاسباتی و مشاهداتی ارتفاع سطح دریا.



شکل ۱۵. نمودار تعیین همبستگی داده‌های محاسباتی و مشاهداتی ارتفاع سطح دریا.



شکل ۱۶. نمودار تغییرات RMSE و MAE پس از هر مرحلهٔ پیش‌بینی.

جدول ۴. تغییرات RMSE و MAE پس از هر مرحلهٔ پیش‌بینی.

مرحلهٔ پیش‌بینی	RMSE(m)	MAE(m)
۱	۰/۰۰۵۴۷	۰/۰۰۰۵۷۱
۲	۰/۰۰۶۷۵۵	۰/۰۰۰۹۶۸
۳	۰/۰۰۶۷۵۵	۰/۰۰۰۹۷۱
۴	۰/۰۰۶۸۵۹	۰/۰۰۱۰۸۹
۵	۰/۰۰۸۳۶۱	۰/۰۰۱۵۶۷
۶	۰/۰۱۰۹۵۵	۰/۰۰۲۲۷۵
۷	۰/۰۱۱۰۰۸	۰/۰۰۲۳۸۴
۸	۰/۰۱۱۰۴۷	۰/۰۰۲۴۷۶
۹	۰/۰۱۱۱۷۳	۰/۰۰۲۶۴۴
۱۰	۰/۰۱۱۲۳۱	۰/۰۰۲۷۵۷
۱۱	۰/۰۱۱۲۵۳	۰/۰۰۲۸۲۸
۱۲	۰/۰۱۴۰۲۱	۰/۰۰۳۶۶۵
۱۳	۰/۰۱۴۹۳۷	۰/۰۰۴۱۷۹
۱۴	۰/۰۱۵۴۰۶	۰/۰۰۴۵۵۷
۱۵	۰/۰۱۵۹۹۳	۰/۰۰۴۹۸۶
۱۶	۰/۰۱۹۴۵۶	۰/۰۰۶۰۹۴
۱۷	۰/۰۱۹۵۳۱	۰/۰۰۶۲۶۵
۱۸	۰/۰۲۲۷۹۹	۰/۰۰۷۴۴۱
۱۹	۰/۰۲۳۴۲۹	۰/۰۰۷۹۸
۲۰	۰/۰۲۴۶۰۵	۰/۰۰۸۷۳۲

۶. نتیجه‌گیری

در این مقاله به مدل‌سازی تغییرات سطح لحظه‌ای آب دریای خزر با استفاده از پنج سال مشاهدات ماهوارهٔ ارتفاع‌سنجی Jason-2 پرداخته‌ایم. با در اختیار داشتن ارتفاع آب و لحظهٔ قرائت هر ارتفاع، یک سری زمانی با فاصلهٔ نمونه‌برداری دو ساعته تشکیل شد. تخمین فرکانس‌ها یکی از مراحل مهم مدل‌سازی است که برای این منظور از روش آنالیز طیفی فوریه برای یافتن فرکانس موجود در سری زمانی و روش کمترین مربعات با تکیه بر مفهوم ایستایی سری زمانی به منظور دستیابی به فرکانس دقیق‌تر استفاده شد. نتایج نشان داد پس از حذف روند و فرکانس‌های مهم در دو مرحلهٔ با پریود بیش از نوزده روز و با پریود بیش از چهار ساعت داده‌ها به‌صورت کامل ایستا و آماده برای مدل‌سازی از طریق سری‌های زمانی شد.

در این بررسی عملکرد مدل‌های سری زمانی اتورگرسیو، میانگین متحرک و مدل خودهمبستهٔ میانگین متحرک در پیش‌بینی ارتفاع لحظه‌ای سطح دریای خزر بر پایهٔ معیارهای خطا و محک‌های آماری مقایسه شد.

مدل اتورگرسیو مرتبهٔ ششم بهترین مدل شناخته شد، به‌طوری که هم روند تغییرات سری زمانی را بهتر شبیه‌سازی کرد و هم به خطای کمتری منجر گردیشد. مقایسه با معیارهای مقدار میانگین مطلق خطا و مجذور میانگین مربعات خطا، همچنین تعیین همبستگی مقادیر محاسبه شده با استفاده از رابطهٔ سری زمانی با داده‌های مشاهداتی، دقت مدل ارائه شده را تأیید کرد.

مراجع

جعفری، ع.، ۱۳۸۵، بررسی روند تغییرات سطح آب دریای خزر با استفاده از مشاهدات ارتفاع‌سنجی ماهواره‌ای و تایدگیج‌های ساحلی، رساله کارشناسی ارشد، دانشکده فنی دانشگاه تهران.

رجبی، س.، چگینی، و.، لاری، ک.، کامران زاد، ب. و حسینی، ط.، ۱۳۹۲، مدل‌سازی عددی امواج ناشی از باد در منطقه انزلی با استفاده از مدل عددی SWAN، نشریه علمی-پژوهشی اقیانوس‌شناسی.

زمانی، ا.، عظیمیان، ا.، ۱۳۸۳، پیش‌بینی ارتفاع امواج دریای خزر با استفاده از شبکه مصنوعی عصبی، دانشگاه صنعتی اصفهان، نهمین کنفرانس دینامیک شاره‌ها.

شجاعی کیاده، ح.، ۱۳۸۳، ادغام اطلاعات سنجنده‌های ماهواره‌ای مختلف و اطلاعات ژئوفیزیکی به‌منظور دستیابی به سطح لحظه‌ای آب دریا، رساله کارشناسی ارشد، دانشکده فنی دانشگاه تهران.

فلاح، ع. ع.، کرمی خانیکی، ع. و آزرمسا، س. ع.، ۱۳۸۳، ارزیابی مدل‌های ریاضی WAM و MIKE ۲۱ در شبیه‌سازی امواج دریای خزر، م. آبخیز، ۳(دی)، ۲۴-۳۰.

طوریان، م، ۱۳۸۵، مدل‌سازی جزرومد دریای عمان و خلیج فارس با استفاده از مشاهدات ارتفاع‌سنجی ماهواره‌ای و مشاهدات تایدگیج‌های ساحلی، رساله کارشناسی ارشد، دانشکده فنی دانشگاه تهران.

محجوبی، ج. و اردلان صمغی، ح.، ۱۳۸۸، پیش‌بینی پارامترهای امواج ناشی از باد در دریای خزر با استفاده از روش درختان تصمیم رگرسیونی و شبکه‌های عصبی مصنوعی، نشریه مهندسی دریا، **۵**(۹)، ۶۵-۷۱.

نیرومند، ح.، ۱۳۷۱، تجزیه و تحلیل سری‌های زمانی، تألیف، جاناتان دی کرایر، مشهد؛ دانشگاه فردوسی مشهد.

نیرومند، ح.، بزرگ نیا، ا.، ۱۳۸۱، مقدمه‌ای بر تحلیل سری‌های زمانی، انتشارات دانشگاه فردوسی مشهد.

Akaike, H., 1973, Information theory and an extension of the maximum likelihood principle, proc., 2nd International Symposium on Information Theory, (Eds. B. N. petrov and F. Csaki), 267-281, Akademiai kiado, Budapest.

Akaike, H., 1974, A new look at the statistical model identification, IEEE Transactions on Automatic Control, AC-19, 716-723.

Akaike, H., 1978, A Bayesian analysis of the minimum AIC procedure, Ann.Inst. Statist. Math., 30A, 9-14.

Akaike, H., 1979, A Bayesian analysis of the minimum AIC procedure of autoregressive model fitting, Biometrika, 66, 237-242.

Anzenhofer, M., Shum, C. K. and Rentsh, M., 1999, Coastal altimetry and applications, Department of Civil and Environmental Engineering and Geodetic Science of the Ohio StateUniversity, Columbus Ohio, Report No. 464.

Barbosa, S, M., Silva, M, E. and Fernandes, M, J., 2006, Multivariate autoregressive modeling of sea level time series from TOPEX/Poseidon satellite altimetry, Nonlin, Processes Geophysics, 13, 177-184.

Cheng, X. and Qi, Y., 2007, Trends of sea level variations in the South China Sea from merged altimetry data, Global and Planetary Change., 57, 371-382.

Lebedev, S. A. and Kostianoy, A. G., 2005, Satellite altimetry of the Caspian Sea, Moscow: Sea, 366 pp., (in Russian).

Marcos, M., Woppelmann, G., Bosch, W. and Savcenko, R., 2007, Decadal sea level trends in the Bay of Biscay from tide gauges, GPS and TOPEX, Journal of Marine, Systems, MARSYS-01435.

Schwartz, G., 1978, Estimating the dimension of a model, ann. Statist., 6, 461-464.

Shibata, R., 1976, Selection of the order of an autoregressive model by Akaike's information criterion, Biometrika, 63, 117-126.