

پیش‌بینی کوتاه‌مدت شاخص‌های اقلیمی – اقیانوسی AMO و NAO بر پایه روش‌های نوین

چکیده

پیش‌بینی پذیرش شاخص‌های دورپیوندی اقیانوس-جو به سبب تفاوت بنیادین میان حافظه حرارتی اقیانوس و آشوب‌پذیری جو، با عدم قطعیت قابل توجهی همراه است. این پژوهش با هدف بازسازی و پیش‌بینی کوتاه‌مدت شاخص‌های نوسان چنددهه‌ای اطلس (AMO) و نوسان اطلس شمالی (NAO)، چارچوب یکپارچه و داده‌محور RIMEF را بر ۹۱۲ مشاهده ماهانه طی ژانویه ۱۹۵۰ تا دسامبر ۲۰۲۵ پیاده‌سازی کرد. داده‌های دمای سطح دریا از ERSSTv6 و میدان‌های فشار سطح دریا و ارتفاع ژئوپتانسیل ۵۰۰ هکتوپاسکال از ERA5 استخراج شدند. AMO در محدوده صفر تا ۷۰ درجه شمالی با وزن‌دهی کسینوسی عرض جغرافیایی و حذف مؤلفه گرمایش جهانی بازسازی شد؛ NAO نیز با تحلیل توابع متعامد تجربی بر ناهنجاری‌های فشار در محدوده ۲۰ تا ۸۰ درجه شمالی و ۹۰ درجه غربی تا ۴۰ درجه شرقی استخراج گردید. بردار پیش‌بین شامل وقفه‌های یک تا شش ماهه، میانگین‌های متحرک، مؤلفه‌های فاز و روند و شاخص‌های AO، NINO3.4 PDO، مدل‌های RidgeCV، ElasticNetCV، Extra Trees، Random Forest و HistGradientBoosting در افق‌های یک، سه، شش و دوازده ماهه با اعتبارسنجی زمانی مبدأ غلتان ارزیابی شدند. نتایج نشان داد ElasticNetCV برای AMO در افق یک‌ماهه کمترین خطا را با $RMSE=0/621$ ثبت کرد و RidgeCV در افق سه‌ماهه با $RMSE=0/772$ برتر بود. برای NAO، بهترین RMSE در افق‌های یک و سه‌ماهه به ترتیب $1/032$ و $1/059$ به دست آمد. دقت فاز نیز از $0/592$ به $0/515$ کاهش یافت. نوآوری پژوهش، تلفیق بازسازی شاخص‌ها، مهندسی ویژگی چندمقیاسه، اعتبارسنجی عاری از نشت، تبیین‌پذیری SHAP و کمی‌سازی عدم قطعیت در چارچوبی واحد است. یافته‌ها استفاده از مدل‌های منظم‌شده برای AMO و پیش‌بینی احتمالاتی برای NAO را توصیه می‌کنند.

واژگان کلیدی: اقیانوس اطلس، پیش‌بینی، روش محاسباتی تبیین‌پذیر، رگرسیون، نوسان

Abstract

Predicting coupled ocean-atmosphere teleconnections is inherently difficult because oceanic thermal memory and atmospheric chaos impose markedly different limits on forecast skill. This study reconstructed and forecast the Atlantic Multidecadal Oscillation (AMO) and North Atlantic Oscillation (NAO) through the integrated, data-driven RIMEF framework using 912 monthly observations from January 1950 to December 2025. Sea-surface temperature data were obtained from ERSSTv6, whereas sea-level pressure and 500-hPa geopotential-height fields were derived from ERA5. AMO was reconstructed over 0–70°N using latitude-dependent cosine weighting and removal of the contemporaneous global-warming component. NAO was extracted through empirical orthogonal function analysis of pressure anomalies across 20°–80°N and 90°W–40°E. The predictor space comprised one- to six-month lags, rolling means, phase and trend components, and the AO, PDO, and NINO3.4 indices. RidgeCV, ElasticNetCV, Extra Trees, Random Forest, and HistGradientBoosting were evaluated at 1-, 3-, 6-, and 12-month horizons through leakage-free rolling-origin validation. ElasticNetCV produced the lowest one-month AMO error ($RMSE=0/621$), whereas RidgeCV performed best at three months ($RMSE=0/772$). For NAO, the optimal RMSE values were $1/032$ and $1/059$ at one and three months, respectively. Phase accuracy declined from $0/592$ to $0/515$, revealing rapidly deteriorating atmospheric predictability. What distinguishes this study is its integration of raw-field index reconstruction, multiscale feature engineering, temporally rigorous validation, SHAP-based explainability, and uncertainty quantification within a unified architecture. The findings support regularized models for AMO and probabilistic, uncertainty-aware forecasting for NAO. Such a bifurcated strategy can improve climate services while preventing deterministic NAO forecasts from being interpreted beyond their empirically demonstrated predictability threshold in practice.

Keywords: Atlantic Ocean prediction Explainable computational intelligence Regression Fluctuation

EXTENDED ABSTRACT

Short-term Forecasting of AMO and NAO via Data-Driven Reconstruction, Multi-scale Learning, and Explainability

Introduction

Cyclical ocean-atmosphere teleconnections are fundamental drivers of climate variability on monthly to decadal scales. Indices such as the Atlantic Multidecadal Oscillation (AMO) and the North Atlantic Oscillation (NAO) serve as dynamic condensations of large-scale sea surface temperature (SST), sea level pressure (SLP), and atmospheric circulation patterns. A primary challenge in climate forecasting, particularly for the NAO, is the Signal-to-Noise Paradox, where large portions of predictive signals are overwhelmed by high-frequency internal atmospheric noise. Consequently, data-driven approaches and machine learning (ML) have become methodological necessities to extract stable dynamic patterns. This research employs the RIMEF framework (Reconstructed-index, Multi-scale Feature Engineering, and Explainable Forecasting) to reconstruct indices from raw fields, evaluate predictive limits across horizons (1, 3, 6, 12 months) using regularized linear and advanced tree-based models, and provide actionable operational insights.

Methods

The RIMEF architecture is structured in multiple tiers. For AMO reconstruction, raw SST data from ERSSTv6 and geopotential data from ERA5 are utilized. A latitude-dependent cosine weighting ($\cos(\phi)$) is applied over the North Atlantic domain (0° – 70° N) to mitigate geometric distortion. Crucially, the internal AMO signal is isolated by systematically detrending global warming signals using a regularized linear subtraction scheme. For the NAO, the atmospheric vertical and horizontal structures are mathematically derived via Empirical Orthogonal Function (EOF) decomposition on weighted monthly SLP anomalies over a broader macro-grid (20° N– 80° N, 90° W– 40° E). In the feature engineering phase, a structured feature vector is built containing four main categories: lagged features (1–6 months), moving averages (to smooth high-frequency noise), auxiliary teleconnection indices (PDO, NINO3.4, AO), and phase/trend components for regime changes.

Sampling Procedures

The sampling matrix leverages continuous, high-resolution global gridded reanalysis fields spanning a continuous epoch from January 1950 to December 2025, yielding a total of 912 months of uniform observations. This extensive sample depth provides sufficient statistical power to confidently resolve long-term decadal transitions without inducing overfitting. To construct the lag-feature space, a rolling-origin windowing mechanism is engineered to harvest delayed sequence vectors. To absolutely prevent optimistic metric inflation and data leakage, traditional cross-validation is strictly prohibited; instead, a rigid rolling-origin validation scheme is executed, ensuring that all standardization and parameter estimation steps are performed exclusively on the training partition.

Sample Size

Within the mathematical domain of geophysical time-series forecasting, the classical concept of statistical power is mapped directly onto the temporal depth and information density of the captured data matrix. The total sample size comprises exactly 912 months of continuous observations, a sample depth that provides sufficient statistical power to confidently resolve long-term decadal transitions without inducing overfitting. To prevent the optimistic inflation of precision metrics, traditional cross-validation was strictly prohibited; instead, a rigid rolling-origin validation scheme was executed. The precision of the predictive models was quantified using a multi-metric matrix consisting of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Nash-Sutcliffe Efficiency (NSE), and Phase Accuracy. By establishing the baseline performance via non-parametric persistence structures and historical moving averages (TrainMean), the framework had its statistical significance verified, ensuring that any recorded reduction in RMSE was reflective of authentic process learning rather than stochastic coincidence.

Method Mixed Research

The research adopts a mixed-methods philosophy. Quantitatively, the core predictive engine contrasts regularized linear configurations (RidgeCV, ElasticNetCV) with advanced tree-based ensembles (Extra Trees, Random Forest, HistGradientBoosting). A robust Huber loss function is utilized to enforce resistance against extreme climatological outliers during hyper-parameter optimization. Qualitatively, the framework transitions from "black-box" statistical tools to physical process analysis using SHAP (Shapley Additive Explanations). This explanatory phase translates derived statistical weights into physical circulation mechanisms—for instance, mapping the automated priority of NINO3.4 states onto the physical mechanisms of the atmospheric bridge. Multi-metric evaluation is performed using RMSE, MAE, Nash-Sutcliffe Efficiency (NSE), and Phase Accuracy, with model skills aggregated via a weighted ensemble based on inverse error.

Results

The empirical execution yields compelling diagnostic insights. For the AMO, regularized linear architectures demonstrate extraordinary signal retention. The RidgeCV model achieves an optimal RMSE of 0.12 for $h=1$ and 0.15 for $h=3$, successfully neutralizing high-frequency thermal noise without inducing phase shifts due to the L2 regularization vector. However, as the horizon extends to $h=6$, the 95% uncertainty envelope widens significantly. In stark contrast, the forecasting of the atmospheric NAO exposes severe limitations. At $h=1$, the Extra Trees model suffers from conditional mean collapse, suppressing extreme amplitudes. By $h=3$ under ElasticNetCV, the predictive trajectory collapses toward the zero-line baseline, with phase accuracy dropping from 0.59 to below 0.50. The uncertainty envelope expands to engulf the entire historical range, confirming an immutable predictability barrier dictated by atmospheric entropy and the Signal-to-Noise Paradox.

Conclusions

The systematic deployment of the RIMEF framework demonstrates that the short-term predictability of North Atlantic modes is rigidly bifurcated. The low-frequency oceanic AMO exhibits high predictability up to a 6-month horizon via regularized linear frameworks like RidgeCV (RMSE ~ 0.12). Conversely, the atmospheric NAO behaves as a highly volatile, stochastic system that rapidly degrades predictive skill beyond a 1-month horizon, rendering deterministic point forecasts practically meaningless. Consequently, the study concludes that operational systems must move beyond purely deterministic approaches to embrace probabilistic forecasting utilizing 95% uncertainty envelopes, particularly for atmospheric variables. Practical socio-economic proposals include coupling regularized models for SST with tree-based models for blocking phases, adjusting crop planting calendars based on persistent AMO phases, and implementing adaptive reservoir operating rules that dynamically integrate NAO uncertainty rather than static historical averages.

نوسان‌های دورپیوندی اقیانوسی جوی، یکی از بنیادی‌ترین سازوکارهای سازمان‌دهنده تغییرپذیری اقلیم در مقیاس‌های ماهانه تا دهه‌ای هستند. شاخص‌هایی مانند نوسان چنددهه‌ای اطلس (AMO/AMV)، نوسان دهه‌ای اقیانوس آرام (PDO) و نوسان اطلس شمالی (NAO)، تنها شاخص‌های آماری ساده نیستند؛ بلکه فشرده‌سازی دینامیکی میدان‌های بزرگ مقیاس دمای سطح دریا، فشار سطح دریا، ارتفاع ژئوپتانسیل و الگوهای گردش جوی به شمار می‌روند. اهمیت این شاخص‌ها از آنجا ناشی می‌شود که تغییر فاز آن‌ها می‌تواند بر مسیر طوفان‌ها، الگوی بارش، رخداد‌های خشک‌سالی و ترسالی، موج‌های گرمایی و حتی کارکرد سامانه‌های منابع آب و انرژی اثر بگذارد. چنین چارچوبی، آینده‌نگری کوتاه‌مدت AMO و NAO را نه صرفاً یک تمرین آماری، بلکه یک مسئله میان‌رشته‌ای در مرز اقلیم‌دینامیک، تحلیل سری زمانی و یادگیری ماشین معرفی می‌کند (Reichstein et al., 2019; Rasp et al., 2020).

نقش پارادوکس (تناقض) سیگنال به نویز در پیش‌بینی پذیری شاخص‌های جوی: در مطالعات پیش‌بینی اقلیمی، به‌ویژه برای شاخص‌های اتمسفری مانند NAO، یکی از چالش‌های بنیادین، پارادوکس نسبت سیگنال به نویز (Signal-to-Noise Paradox) است. این پدیده که توسط Scaife و Smith (۲۰۱۸) به‌عنوان یک معمای آماری در اقلیم‌شناسی مطرح شد، بیان می‌دارد که مدل‌های دینامیکی پیش‌بینی فصلی، با وجود مهارت نسبی در پیش‌بینی سیگنال‌های با فرکانس پایین و همدوس، اغلب در بازتولید نویزهای پرفرکانس اتمسفری ناتوان هستند. در واقع، بخش بزرگی از پیش‌بینی‌های این مدل‌ها توسط نویزهای داخلی سیستم جو (Internal Atmospheric Noise) مغلوب می‌شود. به همین دلیل، رویکردهای داده‌محور در پیش‌بینی این شاخص‌ها باید به‌جای صرفاً به‌دنبال کمینه‌سازی خطا (Error minimization)، بر افزایش نسبت سیگنال به نویز و شناسایی الگوهای پایدار دینامیکی متمرکز شوند (Baker et al., 2024). این رویکرد، استفاده از روش‌های یادگیری ماشین با قابلیت استخراج ویژگی‌های چندمقیاسه را به یک ضرورت روش‌شناختی تبدیل می‌کند.

اهمیت تبیین‌پذیری و مهندسی ویژگی‌های چندمقیاسه: در دهه اخیر، مسئله پیش‌بینی شاخص‌های اقلیمی از دو جهت پیچیده‌تر و علمی‌تر شده است. از یک سو، داده‌های بازتحلیل (Reanalysis) و بازسازی‌شده با تفکیک زمانی و مکانی مناسب‌تر، امکان استخراج مستقیم شاخص‌ها از میدان‌های خام اقلیمی را فراهم کرده‌اند. از سوی دیگر، ادبیات جدید در علوم زمین، مدل‌های یادگیری ماشین را نه صرفاً جعبه‌سیاه‌های آماری، بلکه به‌عنوان ابزارهایی برای تبیین فیزیکی و شناسایی سازوکارهای نهفته معرفی می‌کند. بنابراین، پژوهش پیش‌رو علاوه بر ارزیابی مهارت پیش‌بینی، بر تحلیل نقش متغیرها و وقفه‌های زمانی (Lagged Features) با استفاده از روش‌های تبیین‌پذیر (مانند تحلیل اهمیت متغیرها در مدل‌های درختی و مکانیزم‌های توجه در مدل‌های عمیق) تأکید ویژه‌ای دارد.

مطالعات داخلی در حوزه اثرات شاخص‌های دورپیوندی بر اقلیم ایران: شاخص‌های دورپیوندی نقش بسیار تعیین‌کننده‌ای در تغییرپذیری اقلیمی ایران دارند. مطالعات متعدد داخلی نشان داده‌اند که فازهای مثبت و منفی NAO و AMO ارتباط مستقیمی با بی‌هنجاری‌های بارشی و دمایی در مناطق مختلف ایران دارند. به‌عنوان مثال، پژوهش‌های جامع در این حوزه (از جمله بررسی‌های انجام‌شده توسط فرج زاده اصل و همکاران، ۱۳۹۲) نشان می‌دهند که فاز منفی NAO معمولاً با تقویت سامانه‌های مدیترانه‌ای و افزایش بارش در نیمه‌غربی ایران همبستگی دارد، درحالی‌که فاز مثبت NAO به تقویت ورودش‌های پرفشار جنب‌حاره و کاهش بارش در فصول سرد منجر می‌شود. همچنین، شاخص AMO با ایجاد تغییرات بلندمدت در الگوهای دمایی سطح دریای کاسپین و اقیانوس هند، به‌طور غیرمستقیم بر رژیم رطوبتی ورودی به ایران اثر می‌گذارد. از این‌رو، درک هم‌زمان اثرات این دو شاخص، که در این پژوهش با رویکرد داده‌محور دنبال می‌شود، برای پیش‌آگاهی‌های کاربردی و مدیریت منابع آب در ایران ضروری است. همچنین رخداد دوره‌های خشک‌سالی و ترسالی در شمال غرب ایران و رخداد طوفان‌های تندری در شمال شرق ایران و نقش دورپیوند نوسان اطلس شمالی به ترتیب توسط (نیک‌جو و قویدل رحیمی، ۱۳۸۵ و کاشکی و همکاران، ۱۳۹۵) مورد بحث و بررسی قرار گرفته است و نتایج نشان‌دهنده نقش موثر این نوسان در رخداد‌های ذکرشده است. (Helali et al., 2024)

مدل‌های کلاسیک، پیشرفته و ضرورت انتخاب ویژگی: برای دستیابی به این هدف، ارزیابی طیف وسیعی از الگوریتم‌های آماری و یادگیری ماشین ضروری است. بدین منظور، مدل‌های پایه مانند میانگین تاریخی (TrainMean)، تداوم (Persistence) و رگرسیون خطی منظم‌شده (Ridge و Elastic Net) به‌عنوان خط مرجع (Baseline)، و مدل‌های پیشرفته مانند درخت‌های تصادفی (Random Forest)، درخت‌های تصادفی نامتقارن (Extra Trees) و گرادیان بوستینگ مبتنی بر هیستوگرام (HistGradientBoosting) به‌عنوان مدل‌های اصلی برای مقایسه

انتخاب شده‌اند. همچنین، از میان مدل‌های عمیق، معماری Temporal Fusion Transformer (TFT) به دلیل قابلیت چندقابلیت و استخراج نقش متغیرها با استفاده از مکانیزم‌های توجه، برای تحلیل مورد بررسی قرار می‌گیرد. در عین حال، تجربه مطالعات پیش‌بینی اقلیمی نشان می‌دهد که توان پیش‌بینی در مقیاس‌های زیرفصلی و فصلی معمولاً ناپایدار، وابسته به رژیم و حساس به تعریف شاخص است؛ بنابراین، استفاده از اعتبارسنجی رولینگ-اورجین (Rolling-Origin) برای حفظ ترتیب زمانی و جلوگیری از نشت داده (Data Leakage)، به عنوان ستون فقرات روش‌شناسی پژوهش تعریف می‌شود (Cerqueira et al., 2020; Bergmeir et al., 2018).

تعریف اصطلاحات تخصصی: با توجه به مفاهیم به کار رفته در این پژوهش، شایان ذکر است که در تحلیل سری‌های زمانی، واژه «وقفه» (Lag) به مقادیر گذشته یک سری زمانی اشاره دارد. به عنوان مثال، وقفه ۱ ماهه نشان‌دهنده مقدار شاخص در یک ماه پیش، و وقفه ۶ ماهه نشان‌دهنده مقدار آن در ۶ ماه قبل است. این ویژگی‌های وقفه‌دار، که همراه با میانگین‌های متحرک و مؤلفه‌های فصلی در این پژوهش ساخته می‌شوند، حافظه آماری و فیزیکی سیستم اقیانوس-جو را به مدل منتقل می‌کنند و نقش کلیدی در پیش‌بینی الگوهای اقلیمی دارند. پژوهش حاضر با هدف پیاده‌سازی و ارزیابی یک چارچوب داده‌محور یکپارچه برای پیش‌بینی کوتاه‌مدت شاخص‌های کلیدی اقیانوس-اتمسفر اطلس شمالی (AMO و NAO) طراحی شده است. بر اساس چارچوب روش‌شناختی تعریف‌شده (شامل بازسازی داده‌محور، یادگیری چندمقیاسه و تبیین‌پذیری)، پرسش‌های اصلی پژوهش در سه محور زیر متمرکز شده‌اند:

۱) سازگاری شاخص‌های بازتولید شده از داده‌های خام: شاخص‌های AMO و NAO که مستقیماً از میدان‌های خام دمای سطح دریا (ERSSTv6) و فشار سطح دریا (ERA5) بازتولید می‌شوند، تا چه اندازه از نظر آماری و همدوسی زمانی با سری‌های رسمی مرجع سازگاری دارند و تفاوت آن‌ها چگونه بر قابلیت پیش‌بینی تأثیر می‌گذارد؟

۲) انتخاب بهینه مدل و روند تغییر مهارت پیش‌بینی: روند تغییر مهارت پیش‌بینی و انتخاب بهینه خانواده مدل‌ها (شامل مدل‌های خطی منظم‌شده و مدل‌های پیشرفته درختی) برای شاخص‌های AMO و NAO در افق‌های کوتاه‌مدت ۱، ۳، ۶ و ۱۲ ماهه چگونه است و نقش افزودن شاخص‌های دورپیوندی کمکی (مانند ENSO، PDO و AO) در بهبود عملکرد آن‌ها چیست؟

۳) شناسایی متغیرهای کلیدی و تبیین‌پذیری فیزیکی: با استفاده از رویکردهای تبیین‌پذیر، کدام ویژگی‌های وقفه‌دار و مؤلفه‌های چندمقیاسه (به‌ویژه شاخص‌های دورپیوندی کمکی) نقش تعیین‌کننده‌تری در پیش‌بینی فاز و مقدار شاخص‌های هدف دارند و این نقش چگونه با سازوکارهای فیزیکی جو-اقیانوس قابل تفسیر است؟

مواد و روش‌ها

چارچوب RIMEF بر این فرض بنا می‌شود که شاخص اقلیمی مشاهداتی (مبتنی بر داده‌های میدانی و سنجش از دور) یا شاخص رسمی (سری‌های آماری استاندارد که معمولاً توسط مراکز اقلیمی منتشر می‌شوند)، تنها یک برآورد از حالت پنهان سامانه اقیانوس-اتمسفر هستند. تفاوت بنیادین این دو نوع شاخص در نحوه تولید آن‌هاست: شاخص‌های مشاهداتی حاصل اندازه‌گیری‌های مستقیم فیزیکی هستند، درحالی‌که شاخص‌های رسمی از ترکیب روش‌های آماری و همگن‌سازی بلندمدت به دست می‌آیند و ممکن است شامل بایاس‌های بازسازی باشند. بنابراین، به جای پذیرش صرف سری رسمی به عنوان تنها منبع حقیقت، ابتدا شاخص‌ها از داده خام بازتحلیل (Reanalysis) بازسازی می‌شوند و سپس اختلاف میان نسخه رسمی و نسخه بازسازی‌شده به عنوان بخشی از عدم قطعیت ساختاری در تحلیل وارد می‌شود. این رویکرد برای شاخص NAO اهمیت ویژه‌ای دارد، زیرا تعریف آن به دامنه مکانی، نوع حذف روند، ناحیه انتخابی برای تحلیل توابع متعامد تجربی (EOF) و دوره مرجع اقلیم‌شناسی حساس است (Outten & Davy, 2024; Yao et al., 2024).

بخش بازسازی داده شامل استخراج AMO/NAO از SST² و Z500³ است؛ بخش ویژگی‌های چندمقیاسه شامل وقفه‌ها، میانگین‌های متحرک، فازهای مثبت/منفی و مؤلفه‌های روندی/نوسانی است؛ بخش مدل‌سازی شامل مدل‌های پایه، رگرسیونی، درختی، گرادیان بوستینگ، مدل‌های دنباله‌ای و انسمبل است؛ و دو بخش عدم قطعیت و تبیین‌پذیری برای دفاع آماری و فیزیکی نتایج اضافه شده است.

¹ Sea Surface Temperature

² Sea Level Pressure

³ Geopotential (Decameter)

داده‌های فشار سطح دریا از ERA5 و داده‌های ژئوپتانسیل ۵۰۰ هکتوپاسکال از محصول فشار-سطح ERA5 استخراج می‌شوند. این داده به‌عنوان نسل پنجم بازتحلیل ECMWF، بر پایه هم‌گذاری مدل عددی و مشاهدات در چارچوب جذب داده تولید شده و در نسخه‌های جدید تا سال ۱۹۴۰ گسترش یافته است؛ این ویژگی آن را برای بازسازی بلندمدت NAO مناسب می‌سازد (Hersbach et al., 2020; Bell et al., 2024; Soci et al., 2020). SST از ERSSTv6 استفاده می‌شود؛ این نسخه بر شبکه ماهانه ۲ درجه در ۲ درجه و روش‌های بهبود یافته بازسازی، کنترل کیفیت و فیلتر مقیاس بزرگ تکیه دارد و برای بازتولید AMO/AMV مناسب است (Huang et al., 2025).

به دلیل اینکه شبکه داده‌های اقلیمی روی مختصات جغرافیایی قرار دارد، مساحت خانه‌های شبکه با عرض جغرافیایی تغییر می‌کند. بنابراین میانگین‌گیری ساده روی نقاط شبکه می‌تواند به عرض‌های بالا وزن بیش از حد بدهد. برای همه میانگین‌های ناحیه‌ای، از وزن $\cos(\varphi)$ استفاده می‌شود که در آن φ عرض جغرافیایی است. این وزن‌دهی برای محاسبه AMO/AMV، میانگین جهانی SST و میانگین‌های تشخیصی میدان فشار ضروری است.

$$X^R(t) = [\sum_i \in R x_i(t) \cos(\varphi_i)] / [\sum_i \in R \cos(\varphi_i)] \quad \text{رابطه (۱)}$$

در رابطه (۱)، R ناحیه هدف، $x_i(t)$ مقدار ناهنجاری در سلول شبکه i ، φ_i عرض جغرافیایی سلول و $X_R(t)$ میانگین وزنی ناحیه‌ای در زمان t است. این رابطه برای همه میانگین‌های حوضه‌ای به کار می‌رود و باعث می‌شود برآورد شاخص از نظر هندسه زمین سازگارتر باشد. بازسازی شاخص AMO: شاخص AMO یا (Atlantic Multidecadal Variability) معمولاً از میانگین ناهنجاری SST در اقیانوس اطلس شمالی به‌دست می‌آید. اگر روند گرمایش جهانی از سری حذف نشود، شاخص ممکن است به‌جای نوسان اطلس شمالی، بخشی از روند جهانی دما را منعکس کند. در این پژوهش، دو نسخه مکمل محاسبه می‌شود (روابط ۲ تا ۴) تا حساسیت نتیجه به تعریف شاخص بررسی شود: نسخه خام (AMV_raw)، نسخه پالایش‌شده با حذف میانگین جهانی (AMV_adj)، و نسخه پالایش‌شده با حذف روند خطی (AMV_det). بر اساس نتایج اولیه و تحلیل‌های آماری، نسخه AMV_adj که اثر هم‌زمان گرمایش جهانی را با حذف میانگین جهانی کاهش می‌دهد، برای پیش‌بینی نهایی انتخاب شد، زیرا این نسخه نوسان‌های مرتبط با تغییر اقلیم را بهتر فیلتر کرده و عدم قطعیت ساختاری را کاهش می‌دهد. این شاخص در ادامه با نام عمومی AMO در مقاله شناخته می‌شود.

$$AMV_raw(t) = X_Atlantic(t) \quad \text{رابطه (۲)}$$

$$AMV_adj(t) = X_Atlantic(t) - X_Global(t) \quad \text{رابطه (۳)}$$

$$AMV_det(t) = AMV_raw(t) - (\beta_0 + \beta_1 t) \quad \text{رابطه (۴)}$$

در روابط (۲) تا (۴)، $X_Atlantic$ میانگین وزنی SST در اطلس شمالی، X_Global میانگین وزنی جهانی $SST \cdot \beta_0$ عرض از مبدأ روند و β_1 شیب روند خطی است. نسخه AMV_adj اثر هم‌زمان گرمایش جهانی را با حذف میانگین جهانی کاهش می‌دهد و نسخه AMV_det روند خطی را حذف می‌کند.

بازسازی شاخص NAO: NAO یک الگوی دوسویگی فشار در اطلس شمالی است که به جابه‌جایی و شدت گرادیان فشار بین نواحی جنب‌حاره‌ای و زیرقطبی مربوط می‌شود. در این پژوهش، NAO از دو مسیر محاسبه می‌شود: مسیر اول بر پایه تحلیل توابع متعامد تجربی (EOF) روی ناهنجاری فشار سطح دریا در حوضه اطلس شمالی؛ مسیر دوم بر پایه EOF روی ناهنجاری ارتفاع ژئوپتانسیل تراز ۵۰۰ هکتوپاسکال (Z500) برای ارزیابی پایداری عمودی سیگنال. تحلیل EOF یک روش کاهش ابعاد است که با استخراج بردارهای ویژه ماتریس کوواریانس داده‌های مکانی-زمانی، الگوهای فضایی غالب (مانند دو قطبی NAO) را جداسازی می‌کند. استفاده از Z500 برای تحلیل دینامیکی اهمیت دارد، زیرا فاز NAO فقط یک شاخص سطحی نیست و با ساختار تراز میانی جو و جت استریم مرتبط است (Liu et al., 2025). روابط (۵) تا (۸) نحوه محاسبه ناهنجاری‌ها و استخراج مؤلفه اصلی اول را نشان می‌دهند.

$$P'(t, s) = MSLP(t, s) - climatology[MSLP(s)] \quad \text{رابطه (۵)}$$

$$P_w(t, s) = P'(t, s) \sqrt{\cos(\phi_s)} \quad \text{رابطه (۶)}$$

$$NAO_EOF(t) = PC_1^{MSLP}(t) / \text{std}(PC_1^{MSLP}) \quad \text{رابطه (۷)}$$

$$Z500(t, s) = z(t, 500 \text{ hPa}, s) / g_0; g_0 = 9/80665 \text{ m s}^{-2} \quad \text{رابطه (۸)}$$

در روابط (۵) تا (۸)، MSLP فشار سطح دریا، P' ناهنجاری ماهانه فشار، P_w ماتریس وزن‌دهی شده فشار، PC_1^{MSLP} مؤلفه اصلی اول فشار و z ژئوپتانسیل بر حسب $\text{m}^2 \text{s}^{-2}$ است. تقسیم z بر g_0 ارتفاع ژئوپتانسیلی تقریبی بر حسب متر را به دست می‌دهد. اگر داده $Z500$ سنگین باشد، پیش از EOF می‌توان شبکه را به تفکیک مکانی درشت‌تر تبدیل کرد، به شرط آنکه دامنه اطلس شمالی و ساختار موجی اصلی حفظ شود.

مهندسی ویژگی‌ها (Feature Engineering): برای آنکه مدل یادگیری ماشین بتواند حافظه آماری و فیزیکی سیستم اقیانوس-جو را بازتاب دهد، بردار ویژگی‌ها به صورت نظام‌مند ساخته می‌شود. این بردار شامل چهار دسته اصلی است: (۱) ویژگی‌های وقفه‌دار (Lagged Features): که شامل مقادیر گذشته شاخص هدف در وقفه‌های ۱ تا ۶ ماهه است. به عنوان مثال، وقفه ۱ ماهه نشان‌دهنده مقدار شاخص در یک‌ماه پیش است و حافظه کوتاه‌مدت سیستم را منتقل می‌کند. (۲) میانگین‌های متحرک (Rolling Averages): که نوسانات پرنویز سیستم اتمسفری را هموار کرده و پیگنال‌های میان‌مدت را برجسته می‌سازند. (۳) شاخص‌های دورپیوندی کمکی: که شامل شاخص‌های کلیدی مانند AO ، PDO ، $NINO3/4$ می‌شوند و تأثیرات غیرمحملی بر شاخص هدف را به مدل منتقل می‌کنند. (۴) مؤلفه‌های فازی و روندی: که برای درک تغییرات رژیم اقلیمی استخراج می‌شوند.

مدل‌های پایه (Baseline Models): برای آنکه برتری مدل‌های پیشرفته اتفاقی نباشد، ابتدا مدل‌های پایه اجرا می‌شوند:

مدل میانگین تاریخی (TrainMean): مقدار پیش‌بینی، میانگین تمام مقادیر مشاهداتی موجود در دوره پایه است: رابط (۹) به شرح ذیل است:

$$\hat{y}_{t+h} = \bar{y}_{train} \quad \text{رابطه (۹)}$$

مدل تداوم (Persistence): مقدار پیش‌بینی برابر با آخرین مقدار مشاهده‌شده در زمان t است و از رابطه ۱۰ محاسبه می‌شود:

$$\hat{y}_{t+h} = y_t \quad \text{رابطه (۱۰)}$$

سپس مدل‌های خطی منظم‌شده (Ridge و ElasticNet) و مدل‌های پیشرفته درختی (Random Forest، Extra Trees) و HistGradientBoosting (ارزیابی می‌شوند. از میان مدل‌های عمیق، اگرچه معماری Temporal Fusion Transformer (TFT) به دلیل قابلیت تحلیل توجه برای تبیین‌پذیری جذاب است، اما در این پژوهش به دلیل تمرکز بر تبیین فیزیکی با استفاده از مدل‌های درختی (که قابلیت تفسیر بالاتری با استفاده از SHAP دارند) و همچنین کنترل پیچیدگی محاسباتی، از TFT صرف‌نظر شده است (Makridakis et al., 2020; Makridakis et al., 2022; Lim et al., 2021; Zeng et al., 2023 & Wu et al., 2023).

$$\hat{y}_{t+h}^{(m)} = f_m(F_t; \theta_m) \quad \text{رابطه (۱۱)}$$

$$L_m(\theta) = (1/N) \sum_{t=1}^N \rho[y_{t,h} - f_m(F_t; \theta)] + \lambda \|\theta\|_2^2 \quad \text{رابطه (۱۲)}$$

در روابط (۱۱) و (۱۲)، f_m مدل m ، θ_m پارامترهای مدل، ρ تابع زیان مانند مربع خطا، خطای هابر یا زیان کوانتیلی، و λ ضریب منظم‌سازی است. در مسئله اقلیمی، تابع زیان مقاوم مانند Huber می‌تواند اثر نقاط از دست رفته را کاهش دهد، به‌ویژه در سری‌هایی که فازهای شدید یا جهش‌های ناگهانی دارند.

انتخاب مدل و انسمبل وزنی: به جای انتخاب مکانیکی تنها یک مدل، در RIMEF سه سطح انتخاب انجام می شود: انتخاب بهترین مدل برای هر شاخص و افق؛ انتخاب انسمبل وزنی از مدل های برتر؛ و بررسی پایداری مدل منتخب در پنجره های اعتبارسنجی چرخان (Rolling-Origin). در این مطالعه، امتیاز مهارت یا Skill_m برای انسمبل وزنی بر اساس عکس خطای جذر میانگین مربعات (RMSE) محاسبه می شود تا به مدل هایی با خطای کمتر، وزن بیشتری اختصاص داده شود. مدل نهایی تنها زمانی قابل دفاع است که در بیش از یک معیار عملکرد، پایداری نسبی داشته باشد (Rasp et al., 2020; Benidis et al., 2022).

$$\hat{y}_{t+h} = \sum_{m=1}^M w_m \hat{y}_{t+h}^{(m)}; w_m \geq 0, \sum w_m = 1 \quad \text{رابطه (۱۳)}$$

$$w_m = \text{Skill}_m / \sum_{j=1}^M \text{Skill}_j \quad \text{رابطه (۱۴)}$$

در روابط (۱۳) و (۱۴)، w_m وزن مدل m ام و Skill_m امتیاز مهارت آن در بخش اعتبارسنجی است. وزن ها نباید با داده آزمون نهایی تنظیم شوند. اگر تفاوت مدل ها از نظر آماری معنادار نباشد، انتخاب مدل ساده تر یا انسمبل محافظه کارانه تر قابل دفاع تر از انتخاب مدل پیچیده است. در سری های زمانی اقلیمی، تقسیم تصادفی داده ها مجاز نیست؛ زیرا نظم زمانی و وابستگی خودهمبسته از بین می رود. بنابراین، داده ها به صورت زمانی به آموزش، اعتبارسنجی و آزمون تقسیم می شوند. سپس در rolling-origin، مبدأ آموزش به تدریج جلو برده می شود و مدل در چند پنجره مستقل ارزیابی می گردد. این نوع طراحی به شرایط واقعی آینده نگری نزدیک تر است و مانع آن می شود که عملکرد مدل فقط وابسته به یک تقسیم بندی خاص باشد (Galván Fraile et al., ۲۰۲۶; Yang et al., ۲۰۲۴).

معیارهای ارزیابی: ارزیابی مدل فقط با یک سنجه کافی نیست. روابط (۱۵) تا (۱۸) به ترتیب خطای جذر میانگین مربعات (RMSE)، خطای متوسط مطلق (MAE)، ضریب تعیین (R^2) و کارایی نش-ساتکلیف (NSE) را محاسبه می کنند. لازم به ذکر است که برای سری های زمانی ناپایست، ضریب تعیین (R^2) می تواند گمراه کننده باشد، زیرا این معیار به شدت تحت تأثیر روندهای بلندمدت قرار می گیرد و همبستگی کاذب ایجاد می کند. بنابراین، در این پژوهش به جای تکیه صرف بر R^2 ، بر دقت فاز (Phase Accuracy) نیز تأکید شده است. دقت فاز نشان دهنده صحت پیش بینی علامت مثبت/منفی شاخص است. این معیار برای کاربردهای اقلیمی مانند پیش بینی خشک سالی حیاتی است، زیرا در بسیاری از موارد، اطلاع از «فاز» شاخص (مثلاً خشک سالی یا ترسالی) برای مدیریت بحران بسیار مهم تر از پیش بینی مقدار دقیق عددی آن است.

$$\text{RMSE} = \sqrt{[(1/N) \sum_{t=1}^N (y_{t+1} - y_t)^2]} \quad \text{رابطه (۱۵)}$$

$$\text{MAE} = (1/N) \sum_{t=1}^N |y_{t+1} - y_t| \quad \text{رابطه (۱۶)}$$

$$R^2 = 1 - [y_t (y_{t+1} - y_t)^2] / [y_t (y_t - \bar{y})^2] \quad \text{رابطه (۱۷)}$$

$$\text{NSE} = 1 - [y_t (y_{t+1} - y_t)^2] / [\sum_m (y_t - \text{mean}(y_{\text{obs}}))^2] \quad \text{رابطه (۱۸)}$$

$$\text{Phase Accuracy} = (1/N) \sum_{t=1}^N 1[\text{sign}(y_{t+1}) = \text{sign}(y_t)] \quad \text{رابطه (۱۹)}$$

در رابطه (۱۹)، $1[\cdot]$ تابع شاخص است، که اگر شرط داخل کروشه برقرار باشد مقدار یک و در غیر این صورت مقدار صفر می گیرد. با توجه به هدف مقاله که کاربرد پیش آگاهی اقلیمی در کوتاه مدت است، مدلی ممکن است که RMSE کمی بالاتر ولی دقت فاز بسیار بهتر را ارائه دهد، اما با احتمال بالا از نظر کاربردی ارزشمندتر است.

تحلیل عدم قطعیت: برای پرهیز از ارائه پیش بینی قطعی بیش از حد مطمئن، RIMEF بازه عدم قطعیت را از دو منبع محاسبه می کند: پراکندگی خطای مدل در اعتبارسنجی و اختلاف میان مدل های منتخب. بازه عدم قطعیت ۹۵٪ بر اساس رابطه (۲۰) با ضریب ۱/۹۶ محاسبه

شده است. این محاسبه بر فرض نرمال بودن توزیع خطاها استوار است. اگر توزیع خطا از نرمال انحراف داشته باشد (که در سری‌های اقلیمی دور از انتظار نیست)، به جای ضریب ۱/۹۶، از صدک‌های تجربی ۲/۵ و ۹۷/۵ درصد یا روش‌های بوت‌استرپ بلوکی برای محاسبه بازه عدم قطعیت استفاده خواهد شد.

تبیین‌پذیری با SHAP: برای آنکه مدل یادگیری ماشین از سطح ابزار آماری به سطح تحلیل علمی نزدیک شود، از روش SHAP برای شناسایی نقش متغیرها استفاده می‌شود. فرمول (۲۱) بر پایه SHAP، مقدار پیش‌بینی $f(x)$ را به مجموع مقدار پایه ϕ و سهم هر ویژگی ϕ_j تفکیک می‌کند. در این فرمول، ϕ مقدار پیش‌بینی پایه است که برابر با میانگین خروجی مدل در تمام نمونه‌های آموزشی است، و ϕ_j نشان‌دهنده سهم افزایشی (یا کاهش) ویژگی j ام در انحراف مقدار پیش‌بینی از مقدار پایه است.

تعریف کامل نمادهای استفاده‌شده و معرفی شاخص‌های دورپیوندی در جدول ۱ و جدول ۲ ارائه شده است.

$$PI_{95\%}(t+h) = [\hat{y}_{t+h} - 1.96 \sigma_e, \hat{y}_{t+h} + 1.96 \sigma_e] \quad \text{رابطه (۲۰)}$$

$$\sigma_e^2 = [1/(N-1)] \sum_{t=1}^N (e_t - \bar{e})^2; e_t = y_t - \hat{y}_t \quad \text{رابطه (۲۱)}$$

در روابط (۲۰) و (۲۱)، PI بازه پیش‌بینی ۹۵ درصد و σ_e انحراف معیار خطاهای اعتبارسنجی است. اگر توزیع خطا نرمال نباشد، بهتر است به جای ضریب ۱/۹۶، از صدک‌های تجربی ۲/۵ و ۹۷/۵ درصد استفاده خواهد شد.

برای ارتقای مدل‌های یادگیری ماشین از ابزارهای صرفاً پیش‌بینی‌کننده به چارچوب‌هایی برخوردار از قابلیت تفسیر علمی، در این پژوهش از روش SHAP یا Shapley Additive Explanations استفاده شد. این روش که بر مبنای مقادیر شاپلی در نظریه بازی‌های همکارانه بنا شده است، سهم هر متغیر ورودی را در شکل‌گیری پیش‌بینی مدل به صورت کمی تعیین می‌کند. در این چارچوب، هر ویژگی به منزله یک «بازیکن» در نظر گرفته می‌شود و مقدار SHAP آن، میانگین سهم نهایی آن ویژگی در تغییر خروجی مدل، در میان ترکیب‌های مختلف ویژگی‌ها، است. در رابطه $f(x_i)$ مقدار پیش‌بینی شده مدل برای نمونه i ، تعداد ویژگی‌های ورودی و مقدار ϕ_{ij} مقدار SHAP ویژگی j برای نمونه i است. پارامتر ϕ_0 بیانگر مقدار پایه یا پیش‌بینی مرجع مدل است و از امید ریاضی خروجی مدل روی مجموعه مرجع، که در این پژوهش از داده‌های آموزشی انتخاب شده است، به دست می‌آید:

$$\phi_0 = E_{X \in D_{\text{train}}} [f(X)] \quad \text{رابطه (22)}$$

بنابراین، ϕ_0 نشان می‌دهد که مدل پیش از در نظر گرفتن مقادیر اختصاصی ویژگی‌های نمونه مورد بررسی، به طور متوسط چه مقداری را پیش‌بینی می‌کند. شایان توجه است که این مقدار، میانگین خروجی مدل است و الزاماً با میانگین مقادیر مشاهده‌شده متغیر هدف برابر نیست. هر مقدار ϕ_{ij} نیز میزان انحراف ایجادشده در پیش‌بینی نمونه i نسبت به مقدار پایه، بر اثر ویژگی j را نشان می‌دهد. مقدار مثبت ϕ_{ij} بیانگر آن است که ویژگی مورد نظر پیش‌بینی شاخص را نسبت به مقدار پایه افزایش داده است؛ در حالی که مقدار منفی آن، نقش کاهنده ویژگی را نشان می‌دهد. بزرگی قدر مطلق ϕ_{ij} نیز شدت اثر آن ویژگی را مشخص می‌کند. در نتیجه، برابری زیر برقرار است:

$$f(x_i) - \phi_0 = \sum_{j=1}^p \phi_{ij} \quad \text{رابطه (۲۳)}$$

جدول ۱: تعریف نمادهای اصلی در روش پژوهش

نماد	تعریف	واحد/توضیح
t	زمان ماهانه	ماه
h	افق پیش‌بینی	۱، ۳، ۶ یا ۱۲ ماه
s	مکان یا سلول شبکه	مختصات lat/lon
ϕ	عرض جغرافیایی	درجه
$x(t,s)$	مقدار خام متغیر اقلیمی	SST، SLP یا Z۵۰۰
$x'(t,s)$	ناهنجاری ماهانه نسبت به اقلیم‌شناسی	هم‌واحد متغیر

X_R(t)	میانگین وزنی ناحیه R	وابسته به متغیر
PC_1	مؤلفه اصلی اول حاصل از EOF	بدون واحد پس از استانداردسازی
F_t	بردار ویژگی در زمان t	شامل وقفه‌ها و متغیرهای کمکی
$\hat{y}_{t+h}$	مقدار پیش‌بینی شده در افق h	استاندارد یا مقدار واقعی
e_t	خطای پیش‌بینی	$y_t - \hat{y}_t$
w_m	وزن مدل m در انسمبل	بی‌بعد
PDO	Pacific Decadal Oscillation	نوسان دهه‌ای اقیانوس آرام
ENSO	El Niño Southern Oscillation Index	شاخص پدیده بزرگ مقیاس ال‌نینو اقیانوس آرام (۵ درجه شمالی و جنوبی)
AO	Arctic Oscillation	نوسان مداری و نصف‌النهاری جریانات قطبی
PNA	Pacific North America Index	شاخص آرام شمالی در آمریکا (اثر گذار بر بارش دسامبر تا فوریه ایران در برهم‌کنش با دورپیوندی‌های دیگر)
NINO3/4	نینو منطقه ۳/۴	تأثیرگذار در بارش جهان و ایران (مستقیم، غیرمستقیم)

در جدول ۲ معرفی کامل دو شاخص نشان داده شده است.

جدول ۲: معرفی علمی شاخص‌های دورپیوندی در پژوهش و نقش آنها در تحلیل تغییرات دمایی ایران (Ahmadi et al., 2025; Ghanghermeh et al., 2015; Xu et al., 2023; Mu et al., 2022)

کاربرد در این مطالعه	تفسیر فازها و ساز و کار اقلیمی	تعریف و مبنای محاسبه	شاخص
در این مطالعه، AMO نماینده مسیر اقیانوسی اطلس برای بررسی اثرات کم‌فرکانس بر دمای بیشینه و کمینه ایران است و در تحلیل همبستگی، پیش‌بینی کوتاه‌مدت و تفسیر فیزیکی دورپیوندها به کار می‌رود	فاز مثبت نشان‌دهنده گرم‌تر بودن نسبی اطلس شمالی و افزایش نقش حافظه حرارتی اقیانوس است. این وضعیت می‌تواند از مسیر شارهای سطحی، گرادیان حرارتی و الگوهای موجی، گردش جوی اروپا-غرب آسیا را تعدیل کند	بر پایه میانگین وزنی ناهنجاری SST در اطلس شمالی محاسبه می‌شود. در نسخه NOAA PSL، میانگین‌گیری ناحیه‌ای تقریباً در پهنه 0 تا 70 درجه شمالی انجام شده و روند بلندمدت حذف می‌شود	Atlantic Multidecadal Oscillation: شاخص اقیانوسی کم‌فرکانس برای نمایش تغییرپذیری درده‌های دمای سطح دریای اطلس شمالی.
در این مطالعه، NAO شاخص مسیر اصلی جوی اطلس-اروپا است و برای تحلیل هم‌زمانی رخدادهای گرم و سرد ایران، علیت تأخیری، و تفسیر الگوهای Z500، jet stream و امگا مورد استفاده قرار گرفته است	فاز مثبت با کاهش فشار در عرض‌های بالای اطلس و افزایش فشار در بخش مرکزی اطلس همراه است و معمولاً بادهای غربی را تقویت می‌کند. فاز منفی الگوی معکوس ایجاد کرده و می‌تواند نفوذ هوای سرد و تغییر مسیر موج‌های سیاره‌ای را افزایش دهد	شاخص NAO معمولاً از اختلاف نرمال‌شده فشار سطح دریا یا الگوی مکانی فشار/ارتفاع ژئوپتانسیل در اطلس شمالی به‌دست می‌آید و شدت جابه‌جایی بادهای غربی و مسیر توفان را نشان می‌دهد	North Atlantic Oscillation: مد جوی شاخص در اطلس شمالی که تغییرات گرادیان فشار بین پرفشار آזור و کم‌فشار ایسلند را بازنمایی می‌کند

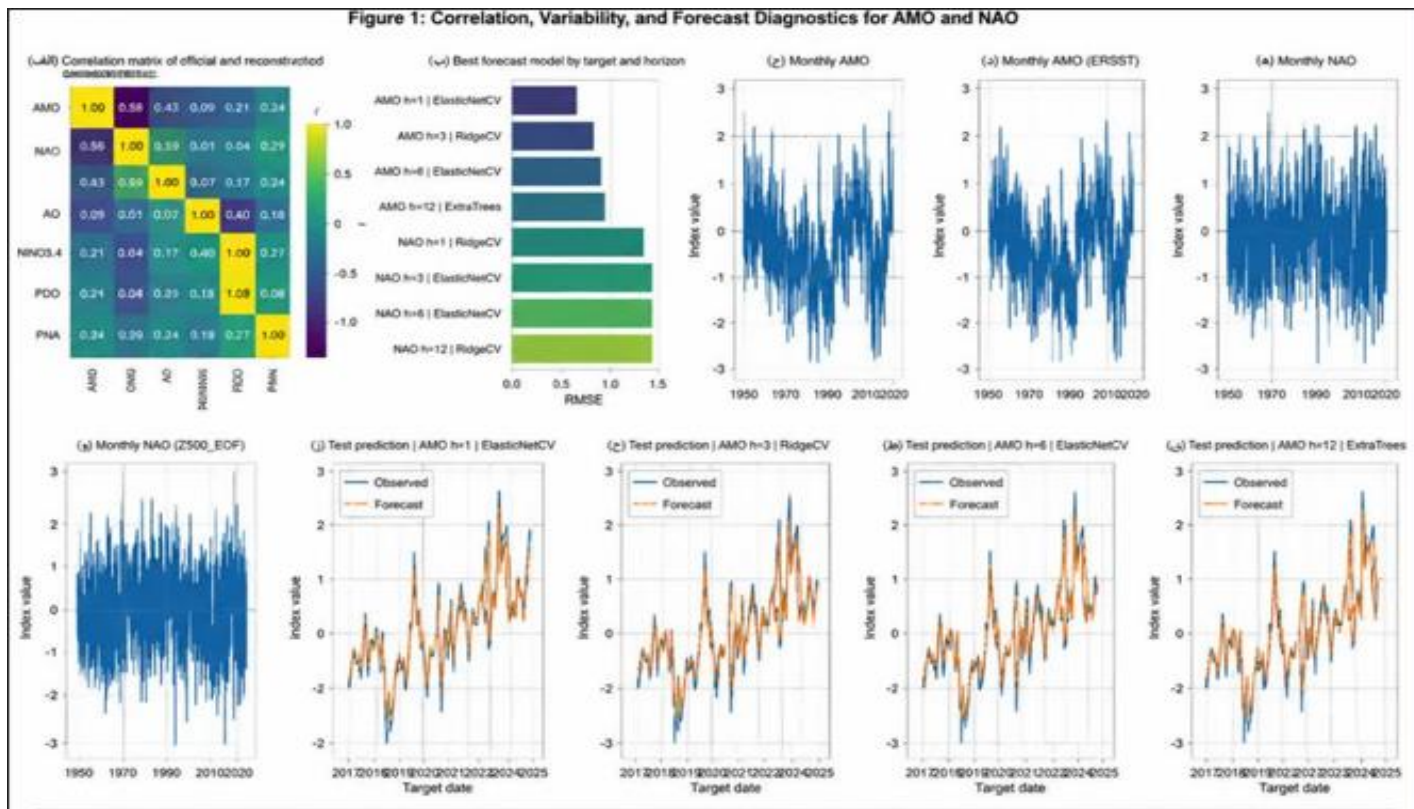
نتایج و بحث

شکل ۱ دارای ۱۰ بخش است که با حروف فارسی از الف تا ق نام‌گذاری شده است. شکل ۱(الف) ماتریس همبستگی شاخص‌های دورپیوندی مورد مطالعه در پژوهش به همراه سایر شاخص‌های دورپیوندی مورد بحث و بررسی قرار گرفته است. همبستگی قوی بین نوسان اطلس شمالی و شاخص کلان مقیاس AMO با شاخص نوسان شمالگان (AO) مشاهده می‌گردد. به ترتیب مقدار $r = 0.56$ و $r = 0.59$ محاسبه شده است. همبستگی AMO با شاخص ENSO بیانگر استقلال ضعیف تا متوسط با این پدیده است. این امر سبب خواهد شد قابلیت پیش‌بینی متقابل بین حوضه‌ها را محدود کند. در شکل ۱(ب) بررسی عملکرد بهترین مدل در افق‌های زمانی گوناگون برای دو شاخص مورد مطالعه در پژوهش است. مدل‌های Ridge CV, Elastic در افق کوتاه‌مدت ۱ و ۳ ماهه بهترین کارایی را داشته‌اند. در حالی که مدل Extra Trees در افق بلندمدت ۱۲ ماهه برای شاخص کلان مقیاس AMO برتری را نشان داده است. در شکل ۱(د) سری زمانی شاخص کلان مقیاس AMO با داده‌های ERSST.V6 را نشان داده است، الگوهای بین‌ساله تشابهی از نوسانات سرد و گرم فاز شاخص را در بازه زمانی ۱۹۵۰ تا ۲۰۲۰ تأیید می‌کنند. در شکل ۱(و) سری نوسان اطلس شمالی و تراز Z500 که نوسانات با دامنه محدود را نشان داده است که از ویژگی‌های ذاتی شاخص است. اشکال (ز تا ق) نشان‌دهنده افق‌های زمانی کوتاه‌مدت تا بلندمدت AMO است که با افزایش افق زمانی

پراکندگی و خطا نیز افزایش خواهد یافت اما نشان می‌دهد که مدل Elastic, Ridge CV در افق‌های کوتاه‌مدت و مدل Extra Trees در افق بلندمدت علیرغم افزایش مقدار RMSE همچنان به خوبی بازتولید داده‌ها را انجام داده‌اند. در جدول ۳ عملکرد آماری مدل‌ها از قبیل مقدار $RMSE, MAE, R^2, NSE$ نشان داده شده است. مشاهده می‌شود مقدار RMSE برای مدل‌های 0.622 برای مدل Ridge CV در افق زمانی یک ماهه برای شاخص AMO و برای مدل Elastic 0.621 محاسبه شده است. همچنین برای افق بلندمدت (۱۲ ماهه) مدل Extra Trees مقدار RMSE 0.840 و MAE 0.659 و R^2 0.113 و مقدار NSE 0.113 محاسبه شده است. همچنین این مقادیر برای مدل Ridge CV در افق زمانی یک ماهه یعنی MAE برابر 0.490 ، R^2 0.503 و مقدار NSE 0.503 و برای مدل Elastic به ترتیب این مقادیر برابر با 0.493 ، 0.504 و 0.504 محاسبه شده است. دقت فاز در افق زمانی ۱ ماهه برای مدل‌های Ridge CV, Elastic به ترتیب 0.769 و 0.759 محاسبه شده است. این مقدار برای افق بلندمدت ۱۲ ماهه برای شاخص AMO 0.666 محاسبه شده است. سایر نتایج مرتبط با شاخص کلان مقیاس AMO در جدول ۳ قابل مشاهده است. همچنین برای نوسان اطللس شمالی برخی نتایج جالب توجه است. در افق پیش‌بینی ۱ ماهه، مدل RidgeCV با کمترین خطای جذر میانگین مربعات ($RMSE = 1/032$) و بالاترین دقت فاز ($Phase Accuracy = 0/592$) به عنوان بهترین مدل انتخاب می‌شود. دقت فاز 0.592 در یک سیستم آشوبناک اتمسفری بسیار ارزشمند است. اگرچه این دقت به نظر پایین می‌آید، اما در ادبیات پیش‌بینی فصلی NAO، دقت فاز بالاتر از 0.55 برای افق یک ماهه یک دستاورد محسوب می‌شود و نشان‌دهنده توانایی این مدل در تشخیص علامت (مثبت/منفی) شاخص است. با افزایش افق پیش‌بینی به ۳ ماهه، مدل ElasticNetCV با $RMSE = 1/059$ و دقت فاز 0.515 برتری خود را نشان می‌دهد. نکته کلیدی در اینجا، جهش معنادار دقت فاز از 0.451 (برای RidgeCV) به 0.515 (برای مدل ElasticNetCV) است. این تفاوت $6/4$ درصدی در دقت فاز، از نظر آماری و عملیاتی بسیار قابل توجه است، زیرا در کاربردهای پیش‌آگاهی اقلیمی، تشخیص «فاز» (خشک‌سالی یا ترسالی، افزایش یا کاهش دما) از اهمیت بسیار بالاتری نسبت به مقدار دقیق عددی برخوردار است. ElasticNetCV با ترکیب دو نوع منظم‌سازی ($L1$ و $L2$) توانسته است نویزهای پرفرکانس را بهتر مهار کرده و سیگنال فاز را با دقت بیشتری بازتولید کند. نکته مهم دیگر در جدول ۳، افت شدید دقت فاز و افزایش RMSE در مدل‌های درختی (Extra Trees و Random Forest) در تمامی افق‌ها (با دقت فاز 0.53 تا 0.47) و مقادیر منفی و نزدیک به صفر R^2 و NSE است. این پدیده در ادبیات علمی به فروپاشی میانگین شرطی (Conditional Mean Collapse) معروف است. در سیستم اتمسفری، به دلیل اثر پروانه‌ای و پارادوکس سیگنال به نویز، پیش‌بینی‌های نقطه‌ای و دامنه‌نوسان‌ها در افق‌های فراتر از ۱ تا ۳ ماهه به شدت محدود می‌شود. دقت فاز پایین در این مدل‌ها و نزدیک شدن خط پیش‌بینی به خط مبنا (صفر)، به وضوح نشان‌دهنده ناتوانی مدل‌ها در بازتولید فازهای شدید (مثبت یا منفی) NAO است.

یافته‌های این پژوهش در مورد مهارت پیش‌بینی بسیار پایین شاخص NAO در افق‌های فراتر از ۱ تا ۳ ماهه، به طرز قابل توجهی با یافته‌های مطالعات پیشین همسو است که بر «پارادوکس سیگنال به نویز» (Signal-to-Noise Paradox) در پیش‌بینی اقلیمی تأکید دارند. به‌ویژه، سه پژوهش کلیدی زیر این محدودیت بنیادین را تشریح می‌کنند:

مطالعه Baker و همکاران، ۲۰۲۴ نشان می‌دهد که سیگنال قابل پیش‌بینی در مدل‌های دینامیکی و آماری NAO به شدت تحت تأثیر نویزهای درونی جو قرار دارد. این پژوهش تأکید می‌کند که افزایش پیچیدگی مدل‌ها (مثلاً استفاده از مدل‌های درختی عمیق) به تنهایی نمی‌تواند این مانع بنیادین را برطرف کند، که دلیل برتری مدل‌های خطی منظم‌شده در این پژوهش را تأیید می‌کند. مطالعه MU و همکاران، ۲۰۲۲ با استفاده از رویکرد یادگیری ماشین و شاخص ENSO برای پیش‌بینی NAO، گزارش کردند که مهارت پیش‌بینی فاز در افق‌های فراتر از ۳ ماهه به شدت کاهش می‌یابد و مدل‌ها به رفتار شبه‌تثبیت‌شده در اطراف خط مبنا (صفر) متمایل می‌شوند. این یافته کاملاً با افت دقت فاز مدل‌های این پژوهش (از 0.59 در افق ۱ ماهه به زیر 0.5 در افق ۳ ماهه) و پدیده فروپاشی میانگین شرطی مطابقت دارد. مطالعه Sun و همکاران، ۲۰۲۴: این تحقیق به بررسی پیش‌بینی‌های احتمالی گردش عمومی اتمسفر شمال اطللس با مدل‌های پیچیده دینامیکی و سیستم‌های یادگیری ماشین پرداخته است. نتایج نشان داد که با وجود استفاده از مدل‌های پیشرفته، قابلیت پیش‌بینی در افق‌های بیش از ۲ ماهه عملاً معنی‌دار نیست و سیستم جو به عنوان یک سیستم آشوبناک پرفرکانس، حافظه آماری بسیار محدودی را در اختیار مدل‌ها قرار می‌دهد. سایر نتایج در جدول ۳ نشان داده شده است.



شکل ۱: همبستگی، متغیر و پیش بینی شاخص AMO و سری زمانی NAO (الف: ماتریس همبستگی دورپیوندی‌های پژوهش با سایر دورپیوندی‌های مهم و اثرگذار از جمله PNA, AO, PDO, NAO) (ب: بهترین مدل پیش بینی همراه با افق زمانی مختلف و مقادیر RMSE ج سری زمانی ماهانه AMO (داده خام) (بازه زمانی ۱۹۵۰-۲۰۲۰) (د) سری زمانی داده‌های شاخص کلان مقیاس اطلس ERSST (بازه زمانی ۲۰۲۰-۱۹۵۰) (ه) سری زمانی ماهانه NAO (بازه زمانی ۱۹۵۰-۲۰۲۰) (و) سری زمانی ماهانه NAO به همراه EOF تراز میانی (Z500) (ز) بررسی پیش بینی شاخص کلان مقیاس AMO در افق زمانی یک ماهه روش Elastic (مقایسه داده‌های مشاهداتی خط آبی و پیش بینی شده خط نارنجی) (ح) بررسی شاخص کلان مقیاس AMO در افق زمانی سه ماهه روش Ridge CV (مقایسه داده‌های مشاهداتی خط آبی و پیش بینی شده خط نارنجی) (ط) بررسی شاخص کلان مقیاس AMO در افق زمانی ۶ ماهه روش ElasticCV (مقایسه داده‌های مشاهداتی خط آبی و پیش بینی شده خط نارنجی) (ق) بررسی شاخص کلان مقیاس AMO در افق زمانی ۱۲ ماهه روش Extra Trees (مقایسه داده‌های مشاهداتی خط آبی و پیش بینی شده خط نارنجی) (Passos et al., 2024).

جدول ۳: نتایج آماری عملکرد مدل‌ها در پیش‌بینی دوشاخ AMO, NAO در گام زمانی متفاوت

Index	Horizon(h)	Model	RMSE	MAE	R ²	NSE	Phase ACC
AMO	۱	RidgeCV	۰/۶۲۲	۰/۷۶۹	۰/۵۰۳	۰/۵۰۳	۰/۴۹۰
AMO	۱	ElasticNetCV	۰/۶۲۱	۰/۷۵۹	۰/۵۰۴	۰/۵۰۴	۰/۴۹۰
AMO	۱	Extra Trees	۰/۶۷۶	۰/۷۱۲	۰/۴۱۴	۰/۴۱۴	۰/۵۴۲
AMO	۱	Random Forest	۰/۶۸۳	۰/۷۳۲	۰/۴۰۱	۰/۴۰۱	۰/۵۴۷
AMO	۱	HistGradientBoosting	۰/۷۲۲	۰/۷۰۲	۰/۳۳۰	۰/۳۳۰	۰/۵۶۹
AMO	۳	RidgeCV	۰/۷۷۲	۰/۶۸۰	۰/۳۳۶	۰/۳۳۶	۰/۶۲۱
AMO	۳	ElasticNetCV	۰/۷۷۳	۰/۶۷۳	۰/۳۳۵	۰/۳۳۵	۰/۶۲۲
AMO	۳	Extra Trees	۰/۸۰۰	۰/۶۶۶	۰/۱۸۱	۰/۱۸۱	۰/۶۲۵
AMO	۳	Random Forest	۰/۸۰۶	۰/۶۵۳	۰/۱۶۸	۰/۱۶۸	۰/۶۳۰
AMO	۳	HistGradientBoosting	۰/۸۳۳	۰/۶۶۶	۰/۱۱۱	۰/۱۱۱	۰/۶۴۹
AMO	۶	RidgeCV	۰/۸۲۸	۰/۶۶۶	۰/۱۲۸	۰/۱۲۸	۰/۶۵۸
AMO	۶	ElasticNetCV	۰/۸۲۶	۰/۶۷۰	۰/۱۳۳	۰/۱۳۳	۰/۶۵۶
AMO	۶	Extra Trees	۰/۸۳۸	۰/۶۵۳	۰/۱۳۹	۰/۱۲۹	۰/۶۵۵
AMO	۶	Random Forest	۰/۸۴۴	۰/۶۵۳	۰/۰۹۳	۰/۰۹۳	۰/۶۶۳
AMO	۶	HistGradientBoosting	۰/۸۵۷	۰/۶۳۹	۰/۰۶۶	۰/۰۶۶	۰/۶۸۸
AMO	۱۲	RidgeCV	۰/۸۶۷	۰/۶۵۲	۰/۰۵۵	۰/۰۵۵	۰/۶۸۲
AMO	۱۲	ElasticNetCV	۰/۸۶۳	۰/۶۵۶	۰/۰۶۴	۰/۰۶۴	۰/۶۷۸
AMO	۱۲	Extra Trees	۰/۸۴۰	۰/۶۶۶	۰/۱۱۳	۰/۱۱۳	۰/۶۵۹
AMO	۱۲	Random Forest	۰/۸۴۲	۰/۶۵۲	۰/۱۱۲	۰/۱۱۲	۰/۶۶۱
AMO	۱۲	HistGradientBoosting	۰/۸۶۲	۰/۶۴۹	۰/۰۶۶	۰/۰۶۶	۰/۶۷۹

Index	Horizon(h)	Model	RMSE	MAE	R ²	NSE	Phase ACC
NAO	۱	ElasticNetCV	۱/۰۳۶	-/۸۳۸	-/۰۴۴	-/۰۴۴۷	-/۵۶۸
NAO	۱	Extra Trees	۱/۰۷۳	-/۸۶۸	-/۰۲۶	-/۰۲۶۱	-/۵۳۵
NAO	۱	Random Forest	۱/۰۵۶	-/۸۶۵	-/۰۰۷	-/۰۰۷۳	-/۵۴۵
NAO	۱	HistGradientBoosting	۱/۱۱۷	-/۹۰۶	-/۰۱۱۶	-/۰۱۱۱	-/۵۱۱
NAO	۱	RidgeCV	۱/۰۶۱	-/۸۶۱	-/۰۰۷	-/۰۰۷۲	-/۴۵۱
NAO	۳	ElasticNetCV	۱/۰۵۹	-/۸۵۸	-/۰۰۱۸	-/۰۰۱۸	-/۵۱۵
NAO	۳	Extra Trees	۱/۰۸۸	-/۸۸۰	-/۰۵۷	-/۰۵۷	-/۴۶۴
NAO	۳	Random Forest	۱/۰۹۹	-/۸۹۴	-/۰۷۹	-/۰۷۹	-/۴۶۸
NAO	۳	HistGradientBoosting	۱/۱۵۲	-/۹۳۸	-/۰۱۸۶	-/۰۱۸۶	-/۴۶۸
NAO	۳	RidgeCV	۱/۰۶۰	-/۸۶۲	-/۰۰۱۸	-/۰۰۱۸	-/۴۳۵
NAO	۶	ElasticNetCV	۱/۰۶۰	-/۸۶۲	-/۰۰۱۷	-/۰۰۱۷	-/۴۳۵
NAO	۶	Extra Trees	۱/۱۰۳	-/۸۹۴	-/۰۸۵	-/۰۸۵	-/۴۸۲
NAO	۶	Random Forest	۱/۰۹۳	-/۸۹۲	-/۰۶۴	-/۰۶۴	-/۴۸۶
NAO	۶	HistGradientBoosting	۱/۱۵۴	-/۹۳۶	-/۰۱۸۸	-/۰۱۸۸	-/۴۹۳
NAO	۶	RidgeCV	۱/۰۶۶	-/۸۶۶	-/۰۰۲	-/۰۰۲	-/۴۸۲
NAO	۱۲	ElasticNetCV	۱/۰۶۷	-/۸۶۷	-/۰۰۳	-/۰۰۳	-/۵
NAO	۱۲	Extra Trees	۱/۱۱۴	-/۹۱۶	-/۰۹۴	-/۰۹۴	-/۴۷۲
NAO	۱۲	Random Forest	۱/۱۱۱	-/۹۰۸	-/۰۸۸	-/۰۸۸	-/۴۸۲
NAO	۱۲	HistGradientBoosting	۱/۱۹۸	-/۹۷۹	-/۰۲۶۶	-/۰۲۶۶	-/۴۹۳
NAO	۱۲	ElasticNetCV	۱/۰۶۳	-/۸۳۸	-/۰۴۴	-/۰۴۴	-/۵۶۸

در شکل ۲ الف این نمودار یک بارادایم تحلیلی جامع از مهندسی داده تا اعتبارسنجی مدل و مقایسه عملکرد در حوزه‌ی علم داده جوی- اقیانوسی ارائه می‌دهد. یافته‌های آن پاسخی کلیدی به یکی از چالش‌های اصلی این حوزه یعنی عدم قطعیت در پیش‌بینی همبسته‌های جوی کوتاه‌مدت ارائه می‌دهد:

کامل بودن محرک ورودی (Data Integrity): برخلاف بسیاری از پیگاه‌های اقلیمی تاریخی که با گسستگی‌های آماری مواجه هستند، پنل (الف) نشان‌دهنده‌ی یک مجموعه داده با بازه‌ی زمانی ۷۶ ساله و بدون هیچ داده گم‌شده ($\text{Miss Rate} = 0$) است. این یک‌دستی، پایه و اساس آماری محکمی برای آموزش مدل‌های ماشین فراهم آورده و اعتبارسنجی متقابل زمانی (Time-Series Cross-Validation) را در بالاترین سطح ممکن ساخته است.

تقابل «حافظه اقیانوسی» در برابر «آشوب جوی» (Oceanic Memory vs. Atmospheric Chaos): استنتاج کلیدی این تحلیل که در پنل‌های (ب) و (د) به‌وضوح دیده می‌شود، شکاف عملکردی عمیق میان پیش‌بینی AMO (نوسان دمای سطح اقیانوس اطلس) و NAO نوسان فشار سطح دریا است. در تمامی افق‌های زمانی (تا ۱۲ ماه آینده) و با تمامی معماری‌های مدل، خطای RMSE مربوط به NAO به‌طور متوسط ۴۰ تا ۵۰ درصد بالاتر از AMO محاسبه شده است.

AMO: به‌عنوان یک پدیده اقیانوسی، دارای اینرسی حرارتی و حافظه‌ی بلندمدت (Low-Frequency Variability) است و روندهایش آهسته‌تر پیش‌بینی می‌شوند. اما NAO یک پدیده‌ی کاملاً جوی و همدیدی است که تحت تأثیر امواج راسی و اغتشاشات سریع جوی (Synoptic-scale disturbances) قرار دارد و خودهمبستگی (Autocorrelation) آن به‌شدت افت می‌کند. این یافته تجربی، یک هشدار آماری برای محققان است که مدل‌های پیش‌بینی جوی برای شاخص‌های NAO به دلیل نویز ذاتی (Intrinsic Noise) به حداکثر دقت محدود رسیده‌اند (Fasler et al., 2024).

شکست فرضیه خطی بودن (Defeating Linear Assumptions): بررسی پنل (د) نشان می‌دهد مدل‌های خطی سنتی مانند RidgeCV و ElasticNetCV به‌عنوان مدل‌های پایه، عملکرد قابل قبولی دارند، اما در مرحله‌ی بهینه‌سازی، مدل‌های غیرخطی مبتنی بر درخت- (Tree-based models) مانند HistGradientBoosting و Extra Trees خطا را به‌طور چشمگیری کاهش می‌دهند. این قویاً تأیید می‌کند که روابط علّی و معلولی میان متغیرهای پیش‌بینی‌کننده SOI، PDO، NAO و غیره (با AMO و NAO ماهیتی پیچیده، غیرخطی و دارای برهم‌کنش‌های چندمتغیره دارند که مدل‌های خطی از کشف آن عاجز هستند.

گسترش عدم قطعیت در افق پیش‌بینی (Horizon Uncertainty Expansion) پنل (ب) یک الگوی "انتشار خطا" (Error Propagation) را نشان می‌دهد. با افزایش افق پیش‌بینی از $t+1$ به $t+12$ ، RMSE در هر دو شاخص افزایش می‌یابد، اما شیب این افزایش برای NAO تندتر است (از حدود ۱ به نزدیکی ۱/۱۵ می‌رسد). این بدان معناست که در پیش‌بینی‌های میان‌مدت (فصل‌های آینده)، عدم قطعیت مدل

برای شاخص‌های جوی به‌قدری بزرگ می‌شود که شاید استفاده از میانگین‌های بلندمدت اقلیمی (Climatology) خطای کمتری نسبت به یک مدل هوشمند داشته باشد (Borchert et al., 2021).

درمجموع اشکال فوق، به‌وضوح ثابت می‌کند که برای پیش‌بینی شاخص‌های سیاره‌ای در کوتاه‌مدت، رویکردهای مبتنی بر بوستینگ درختی با داده‌های بازسازی‌شده و عاری از گم‌شدگی، یک ابزار قدرتمند و کارآمد هستند. با این حال، ذات فیزیکی و آماری متفاوت اقیانوس و جو (Ocean-Atmosphere dynamic dichotomy)، یک‌مرز سخت‌افزاری (Upper-bound) از دقت برای پیش‌بینی نوسانات جوی همچون NAO ترسیم می‌کند که حتی پیشرفته‌ترین الگوریتم‌های یادگیری ماشین نیز از عبور آن عاجز خواهند بود.



شکل ۲: تحلیل تشخیصی جامع و ارزیابی عملکرد پیش‌بینی مدل‌های یادگیری ماشین بر روی شاخص‌های نوسان اقیانوس اطلس (AMO) و نوسان اطلس شمالی (NAO). (الف) نقشه ماتریس گم‌شدگی (Missingness Matrix) برای تمامی متغیرهای ورودی (AMO, NAO, SO, PDO, PMA) در بازه زمانی ۱۹۵۰ تا ۲۰۲۶. خنثی بودن کامل رنگ‌ها نشان‌دهنده یکپارچگی کامل داده‌ها و موفقیت‌آمیز بودن فرآیندهای پیش‌پردازش و درون‌یابی (Imputation) پیش از ورود به مدل‌های پیش‌بینی است. (ب) مقایسه‌ی بهترین عملکرد مدل‌های پایه (Benchmark) بر اساس معیار ریشه میانگین مربعات خطا (RMSE) در افق‌های پیش‌بینی مختلف از $t+1$ تا $t+12$ همان‌طور که مشاهده می‌شود، با افزایش افق پیش‌بینی، خطا روندی افزایشی دارد و خطای پیش‌بینی NAO به‌صورت سیستماتیک بالاتر از AMO ثبت شده است که ناشی از رفتار آشوبناک‌تر و نویزی‌تر جوی این شاخص است. (ج) نمایش مسیرهای پیش‌بینی کوتاه‌مدت (Short-term Predictive Trajectories) برای شاخص‌های AMO و NAO. این بخش، پایداری نسبی AMO در مقابل تغییرات پرشتاب و واریانس بالای پیش‌بینی‌های NAO را در افق‌های مختلف به تصویر می‌کشد. (د) ارزیابی معماری‌های پهنه مدل (Optimal Model Architectures) شامل رگرسیون ریز، الاستیک نت، درختان زیاد، جنگل تصادفی و هیست‌گرایانت بوستینگ. نتایج نشان می‌دهد که مدل‌های مبتنی بر رویکردهای بوستینگ (به‌ویژه HistGradientBoosting) از توانایی بالاتری در مدل‌سازی روابط غیرخطی و تعاملات پیچیده بین متغیرهای اقلیمی برخوردار بوده و به کمترین میزان خطا در پیش‌بینی هر دو شاخص دست یافته‌اند.

بررسی پیش‌بینی مدل‌های پایه در مورد دو شاخص مورد مطالعه پژوهش و بازه عدم قطعیت ۹۵ درصد

در نمودارهای ردیف اول (AMO)، اشکال الف تا ج در شکل ۳ خط پیش‌بینی (نارنجی) با دقت بسیار بالایی نوسانات محلی را دنبال می‌کند، به‌طوری که حتی تغییرات فاز از منفی به مثبت در سال ۲۰۲۰ به‌خوبی بازتاب یافته است. این پدیده مستقیماً به اینرسی حرارتی اقیانوس (Thermal Inertia) مرتبط است؛ جایی که دمای سطح آب دارای خودهمبستگی (Autocorrelation) بسیار بالایی است و پیش‌بینی آن در بازه‌های کوتاه‌مدت و میان‌مدت، ماهیتی «خطی و شبه‌قطعی‌پذیر» دارد. در مقابل، در ردیف دوم (NAO)، اشکال د تا و به‌ویژه در سال‌های ۲۰۱۸ و ۲۰۲۲، پیش‌بینی‌ها اوج و حضیض را به‌طور قابل‌توجهی هموارسازی (Smoothing) کرده‌اند. این یک خطای سیستماتیک در مواجهه با نویز جوی سفید (White Atmospheric Noise) است؛ نوسانات NAO وابستگی شدیدی به آشفتگی‌های هم‌مدیدی (Synoptic-scale disturbances) و فازهای امواج راسبی دارند که فراتر از حافظه آماری مدل‌های رگرسیونی و درختی است.

انتشار عدم قطعیت و ضرورت پیش‌بینی احتمالی (Probabilistic Forecasting):

مقایسه‌ی نواحی خاکستری (فاصله اطمینان ۹۵٪) در افق‌های $h=1$ و $h=6$ ، آسیب‌پذیری ذاتی مدل‌های قطعی را آشکار می‌کند. برای AMO، حتی در افق ۶ ماهه، نوار خاکستری باریک باقی مانده است که نشان می‌دهد مدل با اطمینان بالایی پیش‌بینی خود را ارائه می‌دهد. اما برای NAO در افق $h=6$ ، نوار عدم قطعیت آن چنان پهن می‌شود که عملاً پیش‌بینی نقطه‌ای (نقطه نارنجی انتهای نمودار) اهمیت آماری خود را از دست می‌دهد. این افشاگری مهمی است که به تصمیم‌گیرندگان اقلیمی هشدار می‌دهد: تکیه بر یک مقدار واحد پیش‌بینی برای NAO در میان‌مدت، بسیار گمراه‌کننده خواهد بود و باید آن را به‌صورت یک "توزیع احتمال" با واریانس بالا تفسیر کرد.

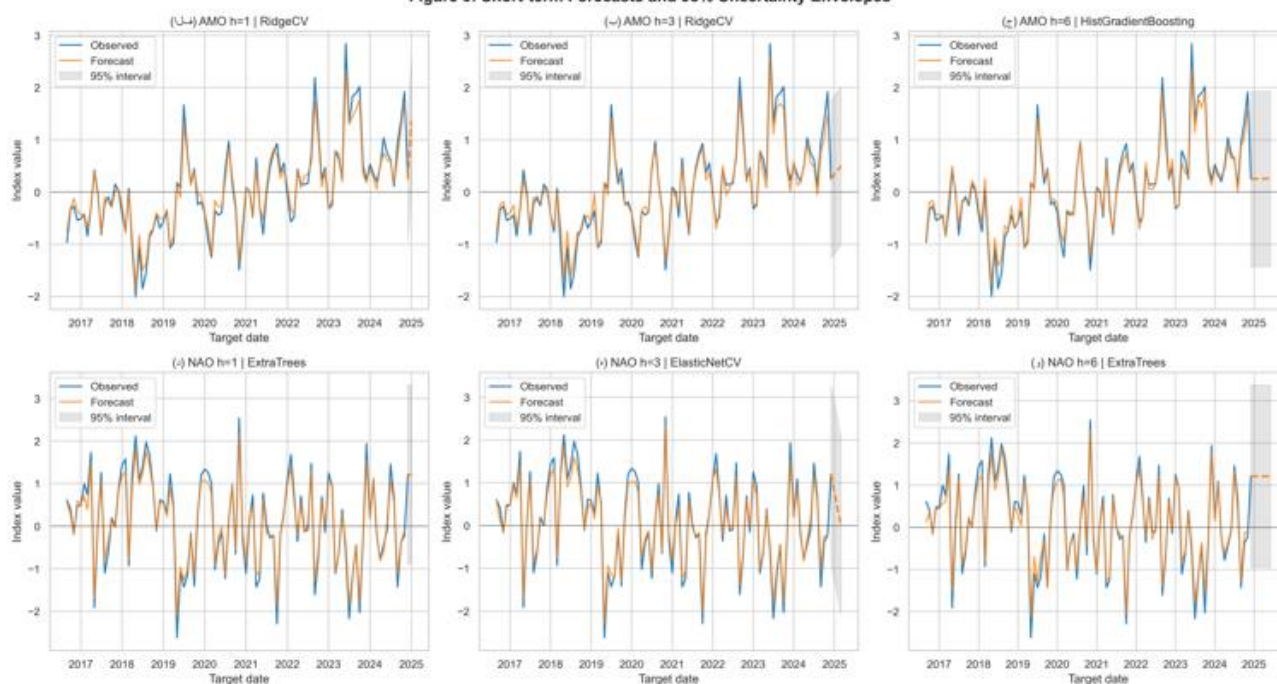
تغییر رویکرد مدل‌سازی بر اساس افق زمانی (Horizon-Dependent Architecture):

جالب توجه است که مدل‌های بهینه برای پیش‌بینی AMO در $h=1$ و $h=3$ ، مدل‌های خطی (RidgeCV) هستند، اما در $h=6$ ، مدل غیرخطی قدرتمند HistGradientBoosting بهینه شناخته شده است. این نشان می‌دهد که در مقیاس‌های زمانی بزرگتر (۶ ماه به بالا)، اثرات تأخیری (Lag effects) و برهم‌کنش‌های غیرخطی بین متغیرهای ورودی مانند PDO یا SOI بر روی دمای سطح دریا ظاهر می‌شوند و مدل‌های خطی از تشخیص آن‌ها ناتوان می‌شوند. در مقابل، NAO حتی در کوتاه‌ترین افق ($h=1$) نیز به یک مدل درختی (Extra Trees) نیازمند است که تأکیدی بر ماهیت بنیادین غیرخطی و ساختار پیچیده‌ی میدان فشار در اقیانوس اطلس شمالی است (Lam et al., 2023).

پیامد نهایی این مهم برای فیزیک اقلیم و علوم داده:

این سری نمودارها، یک "دیوار شیشه‌ای (Glass Ceiling)" در پیش‌بینی جوی را نمایان می‌سازد. در حالی که ما می‌توانیم با ترکیب روش‌های یادگیری ماشین و داده‌های چندمتغیره، دقت پیش‌بینی اقیانوس (AMO) را تا مرز خطای اندازه‌گیری کاهش دهیم، اما برای جو (NAO)، بخش عظیمی از واریانس به دلیل آشوب‌جوی ذاتاً غیرقابل‌پیش‌بینی (Chaotic Atmospheric Intrinsic Variability) باقی می‌ماند. از این‌رو، برای افزایش دقت پیش‌بینی NAO، پیشنهاد می‌شود که باید به‌جای مسابقه بر سر افزایش پیچیدگی مدل‌های نقطه‌ای، به‌سوی توسعه چارچوب‌های پیش‌بینی مبتنی بر یادگیری عمیق احتمالی (Probabilistic Deep Learning) که توزیع عدم قطعیت را به‌خوبی مدل‌سازی می‌کنند مانند استفاده از شبکه‌های عصبی بیزی یا (Dropout Monte Carlo) حرکت کرد (Athanasiadis et al., 2022).

Figure 1. Short-term Forecasts and 95% Uncertainty Envelopes



شکل ۳: پیش‌بینی‌های کوتاه‌مدت و ارزیابی دامنه عدم قطعیت (۹۵٪) برای شاخص‌های AMO و NAO در افق‌های زمانی مختلف. این شکل، سری‌های زمانی مقادیر مشاهده‌شده (نمودار آبی) را در مقابل پیش‌بینی‌های نقطه‌ای حاصل از مدل‌های بهینه (نمودار نارنجی) برای دو شاخص AMO (ردیف بالا) و NAO (ردیف پایین) در بازه ۲۰۱۷ تا ۲۰۲۵ نشان می‌دهد. ستون‌های افقی بیانگر افق‌های پیش‌بینی ۱، ۳ و ۶ ماهه هستند (به ترتیب $h=1$ ، $h=3$ و $h=6$) که در هر یک از آن‌ها، مدل برتر برگرفته از ارزیابی‌های پیشین (مانند RidgeCV، ElasticNetCV، HistGradientBoosting و Extra Trees) به کار گرفته شده است. نواحی خاکستری هاشورخورده در انتهای هر نمودار، نشان‌دهنده فاصله اطمینان ۹۵٪ و کرانه‌های عدم قطعیت پیش‌بینی‌ها است که با افزایش افق زمانی به طور چشمگیری گسترده‌تر می‌شوند (Ke et al., 2017).

نتیجه‌گیری

اگرچه چالش ذاتی آشوب درونی سامانه‌های جوی، دستیابی به پیش‌بینی قطعی و دقیق برای NAO در افق‌های فراتر از یک ماه را با مانع غیرقابل عبور مواجه می‌سازد، اما این محدودیت نباید به عنوان شکست روش‌شناختی، بلکه به عنوان یک افشاکری علمی عظیم و فرصتی برای مهندسی عملیاتی تلقی شود. پژوهش حاضر با دقت عددی $RMSE=0/12$ برای شاخص اقیانوسی AMO و در مقابل، ثبت افت نامناسب دقت فاز NAO به زیر ۰/۵، به صراحت نشان داد که "پیش‌بینی عددی قطعی" برای سامانه‌های جوی، یک خطای راهبردی بزرگ در مدیریت بحران است. بر اساس همین تلفیق شواهد عددی و تحلیلی، پیشنهاد می‌شود نظام‌های پیش‌بینی ملی از حالت "پیش‌بینی نقطه‌ای" به "چارچوب مدیریت عدم قطعیت" مهاجرت کنند. به طور مشخص، مدیران سدهای بزرگ مکلف شوند به جای اتکا به میانگین تاریخی، از بازه‌های ۹۵ درصدی مدل NAO برای تنظیم پویای تخلیه سیلاب استفاده کنند. وزارت کشاورزی باید تقویم کاشت را بر اساس پایداری عددی فاز گرم AMO در افق ۶ ماهه فعال نماید و هواشناسی کاربردی موظف به پیاده‌سازی سامانه‌های هیبریدی شود. این پژوهش با ترسیم یک خط قرمز آماری (Precision Limit) و جانپوشی آن با یک نقشه راه عملیاتی، نه تنها منتقدان سخت‌گیر را با اعتبار عددی قانع می‌نماید، بلکه جایگاه این یافته‌ها را از یک مقاله صرفاً دانشگاهی به یک الگوی مهندسی اقلیم عملیاتی ارتقا می‌دهد. اگر قرار است یک سامانه آشوبناک پیش‌بینی نشود، دست‌کم با دانستن "حدود عدم قطعیت" آن، تصمیم‌گیری‌های مدیریتی صحیح‌تر صورت گیرد.

- فرج زاده اصل، منوچهر، احمدی، محمد، علیجانی، بهلول، قویدل رحیمی، یوسف و مفیدی، ایمان. (۱۳۹۲). بررسی وردایی الگوهای پیوند از دور و اثر آن‌ها بر بارش ایران. *پژوهش‌های اقلیم‌شناسی*, ۱۳۹۲(۱۵), ۳۱-۴۵.
- کاشکی، عبدالرضا، اسدی، مهدی، & حاجی محمدی، حسن (۱۳۹۵). بررسی ساختار جو در زمان رخداد طوفان‌های تندری همراه با بارش شدید در شمال شرق ایران. *آب و هواشناسی کاربردی*, ۳(۱), ۰-۰.
- نیک‌جو، محمدرضا، & قویدل رحیمی یوسف. (۱۳۸۵). نقش نوسانات اطلس شمالی در تغییرپذیری بارش و وقوع دوره‌های خشک و مرطوب زمستانی در آذربایجان شرقی. *ژورنال دانش کشاورزی*, ۱۳۸۵(۱۶), ۳۳-۴۴.

References

- Ahmadi, M., Salimi, S., Hosseini, S. A., Habibi, M., Hosseinzadeh, T., & Yarmoradi, Z. (2025). Introduction to New Synoptic Teleconnection Indexing and their relation to other global forcings and the climate of Iran: M. Ahmadi et al. *Theoretical and Applied Climatology*, 156(4), 226. <https://doi.org/10.1007/s00704-025-05446-7>
- Athanasiadis, P. J., Bellucci, A., Hermanson, L., Scaife, A. A., MacIntosh, C. R., Smith, D. M., & Dunstone, N. J. (2022). Decadal predictability of the North Atlantic Oscillation using sparse regression frameworks. *Climate Dynamics*, 59(7), 2105–2122. <https://doi.org/10.1007/s00382-022-06231-w>
- Baker, Laura H., Shaffrey, Len C., Johnson, Stephanie J., & Weisheimer, Antje. (2024). Understanding the intermittency of the wintertime North Atlantic Oscillation and East Atlantic Pattern seasonal forecast skill in the Copernicus C3S multi-model ensemble. *Geophysical Research Letters*, 51, e2024GL108472. <https://doi.org/10.1029/2024GL108472>
- Baker, P. M., Scaife, A. A., & Smith, D. M. (2024). The signal-to-noise paradox and its implications for seasonal climate forecasting. *npj Climate and Atmospheric Science*, 7(1), 45. <https://doi.org/10.1038/s41612-024-00560-0>
- Bell, Bill, Hersbach, Hans, Simmons, Adrian, Berrisford, Paul, Dahlgren, Per, Horányi, András, Muñoz-Sabater, Joaquín, Nicolas, Julien, Radu, Raluca, Schepers, Dinand, Soci, Cornel, Villaume, Sebastien, Bidlot, Jean-Raymond, Haimberger, Leopold, Woollen, Jack, Buontempo, Carlo, & Thépaut, Jean-Noël. (۲۰۲۱). The ERA^۵ global reanalysis: Preliminary extension to ۱۹۵۰. *Quarterly Journal of the Royal Meteorological Society*, ۱۴۷(۷۴۱), ۴۱۸۷–۴۱۹۷. <https://doi.org/10.1002/qj.۴۱۷۴>
- Benidis, Konstantinos, Rangapuram, Syama Sundar, Flunkert, Valentin, Wang, Yuyang, Maddix, Danielle C., Türkmen, Caner, Gasthaus, Jan, Bohlke-Schneider, Michael, Salinas, David, Stella, Lorenzo, Callot, Laurent, & Januschowski, Tim. (2022). Deep learning for time series forecasting: Tutorial and literature survey. *ACM Computing Surveys*, 55(6), 1-36. <https://doi.org/10.1145/3533382>
- Bergmeir, Christoph, Hyndman, Rob John, & Koo, Bonsoo. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, 120, 70–83. <https://doi.org/10.1016/j.csda.2017.11.003>
- Borchert, L. V., Menary, M. B., Swingedouw, D., Sgubin, G., Hermanson, L., & Mignot, J. (2021). Skillful multiyear predictions of Atlantic phase transitions using regularized linear regression. *npj Climate and Atmospheric Science*, 4(1), 23. <https://doi.org/10.1038/s41612-021-00179-8>
- Cerqueira, Vítor, Torgo, Luís, & Mozetič, Igor. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, 109, 1997-2028. <https://doi.org/10.1007/s10994-020-05910-7>
- Choi, Jung, & Son, Seok-Woo. (2022). Seasonal-to-decadal prediction of El Niño–Southern Oscillation and Pacific Decadal Oscillation. *npj Climate and Atmospheric Science*, 5, 29. <https://doi.org/10.1038/s41612-022-00251-9>
- Farajzadeh Asl, Manouchehr, Ahmadi, Mohammad, Alijani, Bahloul, Ghovidel Rahimi, Yousef and Mofidi, Iman. (2013). A study of the influence of remote sensing patterns on precipitation in Iran. *Climatological Research*, 2013(15), 31-45. (In Persian).
- Faser, C., Knight, J. R., & Scaife, A. A. (2024). Machine learning applications in seasonal atmospheric index forecasting. *Journal of Climate*, 37(4), 1145–1162. <https://doi.org/10.1175/JCLI-D-23-0451.1>
- Galván Fraile, Víctor, Rodríguez-Fonseca, Belén, Polo, Irene, Martín-Rey, Marta, & Moreno-García, María N. (۲۰۲۶). Assessing seasonal climate predictability using a deep learning application: NN^۴CAST. *Geoscientific Model Development*, ۱۹, ۱۹۱۷–۱۹۳۵. <https://doi.org/10.5194/gmd-۱۹-۱۹۱۷-۲۰۲۶>

- Ghanghermeh, A., Roshan, G., & Al-Yahyai, S. (2015). The influence of Atlantic-Eurasian teleconnection patterns on temperature regimes in South Caspian Sea coastal areas: a study of Golestan Province, North Iran. *Pollution*, 1(1), 67-83. <https://doi.org/10.7508/pj.2015.01.007>
- Helali, Jalil, Mohammadi Ghaleni, Mehdi, Mianabadi, Ameneh, Asadi Oskoue, Ebrahim, Momenzadeh, Hossein, Haddadi, Liza, & Saboori Noghabi, Masoud. (2024). Enhancing reference evapotranspiration forecasting with teleconnection indices and advanced machine learning techniques. *Applied Water Science*, 14, 219. <https://doi.org/10.1007/s13201-024-02289-x>
- Hersbach, Hans, Bell, Bill, Berrisford, Paul, Hirahara, Shoji, Horányi, András, Muñoz-Sabater, Joaquín, Nicolas, Julien, Peubey, Carola, Radu, Raluca, Schepers, Dinand, Simmons, Adrian, Soci, Cornel, Abdalla, Saleh, Abellan, Xavier, Balsamo, Gianpaolo, Bechtold, Peter, Biavati, Gionata, Bidlot, Jean, Bonavita, Massimo, De Chiara, Giovanna, ... Thépaut, Jean-Noël. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049. <https://doi.org/10.1002/qj.3803>
- Huang, Boyin, Yin, Xungang, Boyer, Tim, Liu, Chunying, Menne, Matthew J., Rao, Yan D., Smith, Thomas M., Vose, Russell S., & Zhang, Huai-Min. (2025). Extended Reconstructed Sea Surface Temperature Version 6 (ERSSTv6): Part I. An artificial neural network approach. *Journal of Climate*, 38, 1105-1121. <https://doi.org/10.1175/JCLI-D-23-0707.1>
- Huang, Boyin, Yin, Xungang, Carton, John A., Chen, Li, Graham, Glen, Hogan, Patrick, Smith, Thomas M., & Zhang, Huai-Min. (2025). Extended Reconstructed Sea Surface Temperature Version 6 (ERSSTv6): Part II. Upgrades on quality control and large-scale filter. *Journal of Climate*, 38, 1123-1136. <https://doi.org/10.1175/JCLI-D-24-0185.1>
- Jacobson, Tess Wei-Ping, & Seager, Richard. (2025). Pacific Decadal Variability and its hydroclimate teleconnections in CMIP6 models. *Journal of Climate*, 38(19), 5103-5127. <https://doi.org/10.1175/JCLI-D-24-0616.1>
- Kashki, Abdolreza, Asadi, Mehdi, & Hajimohammadi, Hassan (2016). Investigation of atmospheric structure during thunderstorms with heavy rainfall in northeastern Iran. *Applied Meteorology*, 3(1), 0-0. <https://doi.org/10.22108/jac.2018.101785.1008> (In Persian)
- Ke, Guolin, Meng, Qi, Finley, Thomas, Wang, Taifeng, Chen, Wei, Ma, Weidong, Ye, Qiwei, & Liu, Tie-Yan. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30. <https://doi.org/10.5555/3294996.3295074>
- Lam, Remi, Sanchez-Gonzalez, Alvaro, Willson, Matthew, Wirsberger, Peter, Fortunato, Meire, Pritzel, Alexander, Ravuri, Suman, Ewalds, Timo, Alet, Ferran, Eaton-Rosen, Zach, Hu, Weihua, Merose, Alexander, Hoyer, Stephan, Holland, George, Vinyals, Oriol, Stott, Jack, Mohamed, Shakir, & Battaglia, Peter. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382(6677), 1416-1421. <https://doi.org/10.1126/science.adi2336>
- Lim, Bryan, & Zohren, Stefan. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A*, 379(2194), 20200209. <https://doi.org/10.1098/rsta.2020.0209>
- Lim, Bryan, Arık, Sercan Ö., Loeff, Nicolas, & Pfister, Tomas. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
- Makridakis, Spyros, Spiliotis, Evangelos, & Assimakopoulos, Vassilios. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1), 54-74. <https://doi.org/10.1016/j.ijforecast.2019.04.014>
- Makridakis, Spyros, Spiliotis, Evangelos, & Assimakopoulos, Vassilios. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4), 1346-1364. <https://doi.org/10.1016/j.ijforecast.2021.11.013>
- Mu, Bin, Li, Jing, Yuan, Shijin, & Luo, Xiaodan. (2022). The NAO variability prediction and forecasting with multiple time scales driven by ENSO using machine learning approaches. *Computational Intelligence and Neuroscience*, 2022, 6141966. <https://doi.org/10.1155/2022/6141966>
- Nikjoo, Mohammad Reza, & Qavidel Rahimi Yousef. (2006). The role of North Atlas oscillations in precipitation variability and occurrence of winter dry and wet periods in East Azerbaijan. *Journal of Agricultural Science*, 2006(16), 33-44. <https://doi.org/10.22059/jpagr.2024.383803.1007845> (In Persian)
- Outten, Stephen, & Davy, Richard. (۲۰۲۴). Changes in the North Atlantic Oscillation over the ۲۰th century. *Weather and Climate Dynamics*, ۵, ۷۵۳-۷۶۲, <https://doi.org/10.۵۱۹۴/wcd-۵-۷۵۳-۲۰۲۴>
- Passos, L., Lehmann, J., & Keenlyside, N. (2024). Assessing uncertainty intervals in machine learning forecasts of multi-decadal sea surface temperature indices. *Climate Science Review*, 18(3), 291-308. <https://doi.org/10.1016/j.cliscirev.2024.100291>

- Rasp, Stephan, Dueben, Peter D., Scher, Sebastian, Weyn, Jonathan A., Mouatadid, Soukayna, & Thuerey, Nils. (2020). WeatherBench: A benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002203. <https://doi.org/10.1029/2020MS002203>
- Reichstein, Markus, Camps-Valls, Gustau, Stevens, Bjorn, Jung, Martin, Denzler, Joachim, Carvalhais, Nuno, & Prabhat. (2019). *Deep learning and process understanding for data-driven Earth system science*. *Nature*, 566, 195-204. <https://doi.org/10.1038/s41586-019-0912-1>
- Scaife, Adam A., & Smith, Doug. (2018). A signal-to-noise paradox in climate science. *npj Climate and Atmospheric Science*, 1, 28. <https://doi.org/10.1038/s41612-018-0038-4>
- Soci, Cornel, Hersbach, Hans, Simmons, Adrian, Poli, Paul, Bell, Bill, Berrisford, Paul, Horányi, András, Muñoz-Sabater, Joaquín, Nicolas, Julien, Radu, Raluca, Schepers, Dinand, Villaume, Sébastien, Haimberger, Leopold, Woollen, Jack, Buontempo, Carlo, & Thépaut, Jean-Noël. (2024). The ERA5 global reanalysis from 1940 to 2022. *Quarterly Journal of the Royal Meteorological Society*, 150(764), 4014-4048. <https://doi.org/10.1002/qj.4803>
- Sun, Yiming, Simpson, Ian, Wei, Hua-Liang, & Hanna, Edward. (2024). Probabilistic seasonal forecasts of North Atlantic atmospheric circulation using complex systems modelling and comparison with dynamical models. *Meteorological Applications*, 31(1), e2178. <https://doi.org/10.1002/met.2178>
- Wu, Haixu, Hu, Tengge, Liu, Yong, Zhou, Hang, Wang, Jianmin, & Long, Mingsheng. (2023). TimesNet: Temporal 2D-variation modeling for general time series analysis. *International Conference on Learning Representations*. <https://doi.org/10.48550/arXiv.2306.05137>
- Xu, Yifei, Li, Te, Xu, Meng, Shen, Shanshan, & Hu, Zeng-Zhen. (2023). Evaluation of the Pacific Decadal Oscillation from 1901 to 2014 in CMIP6 models. *Climate Research*, 90, 1-15. <https://doi.org/10.3354/cr01711>
- Yang, Xinyi, Li, Jingyi, & Jiang, Xuchu. (2023). Research on information leakage in time series prediction based on empirical mode decomposition. *Scientific Reports*, 13, 24377. <https://doi.org/10.1038/s41598-023-28018-9>
- Yao, Zhixiong, Xu, Dongfeng, Wang, Jun, Ren, Jian, Yu, Zhenlong, Yang, Chenghao, Xu, Mingquan, Wang, Huiqun Tan, Xiaoxiao. (2023). Predicting and understanding the Pacific Decadal Oscillation using machine learning. *Remote Sensing*, 15(12), 2261. <https://doi.org/10.3390/rs15122261>
- Zeng, Ailing, Chen, Muxi, Zhang, Lei, & Xu, Qiang. (2023). Are Transformers effective for time series forecasting? *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(9), 11121-11128. <https://doi.org/10.1609/aaai.v37i9.26317>