

اصلاح پیش‌بینی فرایند توفان حاره‌ای با انتخاب ویژگی‌های موثر

فریده حبیبی^۱، مجید مزرعه فراهانی^{۲*}، بهزاد مشیری^۳، احمد نوحه‌گر^۴ و احمد کلهر^۵

^۱ دانشجوی دوره دکتری مشترک فیزیک جو، دانشگاه هرمزگان، ایران

^۲ استادیار، موسسه ژئوفیزیک دانشگاه تهران، ایران

^۳ استاد قطب علمی کنترل و پردازش هوشمند، دانشکده مهندسی برق و کامپیوتر دانشگاه تهران، ایران

^۴ دانشیار، گروه آبخیزداری دانشکده کشاورزی و منابع طبیعی دانشگاه هرمزگان، ایران

^۵ دانشجوی دوره دکتری برق-کنترل، دانشکده مهندسی برق و کامپیوتر دانشگاه تهران، ایران

(تاریخ دریافت: ۱۳۸۹/۵/۲۶، تاریخ پذیرش: ۱۳۹۰/۴/۷، دسترسی برخط: ۱۳۹۰/۱۲/۱)

چکیده

انتخاب ویژگی‌های مناسب برای بررسی و پیش‌بینی هر کمیت جوئی مرتبط با یک پدیده هواشناسی از قبیل وقوع توفان حاره‌ای یکی از چالش‌های اساسی شناسایی سامانه و مدل‌سازی آن است. در این تحقیق روش‌های گوناگون استخراج ویژگی از قبیل جست‌وجوی پی‌درپی پیشرو، پسرو، معیار همبستگی متقابل، تحلیل مولفه‌های اصلی و تحلیل عامل‌های اصلی برای استخراج کمیت‌های مرتبط با سرعت باد در ارتفاع ۱۰ متری از سطح زمین در زمان وقوع پدیده توفان حاره‌ای اعمال شده است. در این بررسی فقط داده‌های ۴۵ متغیر متفاوت در محدوده فعالیت توفان حاره‌ای گونو که در اوایل ماه ژوئن ۲۰۰۷ منطقه دریای عرب، دریای عمان و خلیج فارس را در نوردید و خسارت زیادی را به جنوب و جنوب‌شرقی ایران، شرق کشور عمان وارد ساخت پردازش شده است. از آنجا که برای استفاده از شبکه عصبی و منطق فازی نیاز به داده‌های آزمون و اعتبارسنجی است از داده‌های توفان حاره‌ای یمین که در اواخر ماه ژوئن همان سال در خلیج بنگال و خلیج فارس شکل گرفت درحکم داده آزمون و از داده‌های توفان حاره‌ای نرجس که در اوایل ماه مه ۲۰۰۸ در خلیج بنگال رخ داده بود به‌منزله داده اعتبارسنجی استفاده شده است. کمیت‌های انتخابی نهایی بعد از استخراج ویژگی درحکم ورودی به شبکه استنباط فازی-عصبی تطبیقی (ANFIS) که خروجی آن سرعت باد در ارتفاع ۱۰ متری از سطح زمین است داده شد و مقدار شاخص خطا در حالت‌های متفاوت آموزش شبکه تعیین و در داخل جدولی آورده شده است.

واژه‌های کلیدی: استخراج ویژگی، شبکه استنباط فازی-عصبی تطبیقی (ANFIS)، توفان حاره‌ای، جست‌وجوی پی‌درپی پیشرو (SFS)، معیار همبستگی متقابل

Improvement the tropical cyclone forecasting process with effective features selection

Farideh Habibi¹, Majied Mazraeh Frahani^{2*}, Behzad Moshiri³, Ahmad Nohegar⁴ and Ahmad Kalhor⁵

¹University of Hurmozgan, Iran

²Institute of Geophysics, University of Tehran, Iran

³University College of Engineering, University of Tehran, Iran

⁴Watershed Management Group of School of Agriculture and Natural Resources of University of Hurmozgan, Iran

(Received: 17 August 2010, accepted: 28 June 2011, available online: 20 February 2012)

Summary

The selection of adequate features for studying and forecasting of each atmospheric quantity regarding the meteorological phenomena (for example, a tropical cyclone) is one of the basic challenges in system recognition and modeling concepts.

This study uses various methods of feature extraction including sequential forward selection, sequential backward selection, mutual information, principal components analysis (PCA) and principal factor analysis (PFA) to extract quantities considering the wind speed at 10 meters above surface for a tropical cyclone activity period. This work studies 45 various quantities of a Guno tropical cyclone activity domain which occurred in early June 2007 over the Arabian Sea, the Gulf of Fars, and the Oman Sea regions which caused significant damage to the region due to storm surges. These quantities were initially selected from 286 quantities of the atmosphere which were determined in the global network of 1 by 1 degree under the standard conditions including standard pressure levels (1000 mb to 10 mb), surface land temperature (SLT), and surface sea temperature (SST). Due to the fact that the activity of the interesting phenomena (i.e. Guno tropical storm) concentrates at the low and medial troposphere, it is possible to exclude quantities of the upper atmosphere.

As a rule, the employment of neural network and fuzzy logic methods requires that the data be tested and validated. As a result, the Yemyin tropical cyclone data was considered as testing data (it happened exactly 8 days after the Guno storm in the west of India, and, after traveling over India, then entered into the southeast of Iran and, subsequently, into the northwest of Pakistan). The Nargis tropical cyclone data which occurred in May 2008 at the Bay of Bengal was considered as validation data because there is currently no valid and acceptable data regarding the history of the old Guno tropical storm which occurred in this region approximately 35 years ago.

The features extracted from the three first methods were tested using an adaptive network-based fuzzy inference system (ANFIS) approach. Of course, 44 features were considered as inputs and a speed of 10 meters of surface was considered as output. The Guno storm data was considered for training the network but the Yemyin storm data was used for testing the network and the Nargis storm data was used for the purpose of network validation.

Due to the fact that the KMO parameter for all three storms was greater than 0.80, the PCA method was used with high confidence for data mining. The results show that the first 18 new components of the storms Guno and Yemyin are greater than the others and they include more than 95% of existing information regarding 45 variables; consequently, the remaining components can be eliminated. However, this is true for the Nargis storm with the first 20 new components.

The results show that the computation time of the sequential backward selection method is very high in comparison with the sequential forward selection method. The

results show that the corresponding accuracy for sequential backward selection is also high.

The results of an ANFIS test for a different epoch show that the features extracted from the forward selection method have minimum errors within three epochs.

The study of coefficient components shows the basic part of the second component is related to wind speeds of 10 meters height over the surface and wind speeds in the layer 30 mb above the surface. Therefore, these components can be utilized as output for neural networks and fuzzy systems such as the ones used in the ANFIS method.

Key words: Features selection, Adaptive Network-based Fuzzy Inference System (ANFIS), Tropical Cyclone, Sequential Forward Selection (SFS), Mutual Information

۱ مقدمه

کالوگرو و همکاران (۱۹۹۷) از شبکه عصبی چندلایه‌ای پیش‌خور بارش ایستگاه را برآورد کردند، لیا سیدو و همکاران (۱۹۹۹) از شبکه عصبی کوهن برای شناسایی الگوهای سینوپتیکی (هم‌ارتفاع‌های 500 mb) روی نقشه‌های وضعیت جوئی استفاده کردند. مک کلوخ و همکاران (۱۹۹۵) با شبکه پس‌انتشار امکان بارش را از روی تصاویر ماهواره‌ای برآورد کردند. مرزبان و استامپف (۱۹۹۸) با شبکه پس‌انتشار توانستند تورنادو را پیش‌بینی کنند؛ البته پیش‌بینی آنها براساس بله و خیر است. به عبارتی دستگاه به این سؤال پاسخ می‌دهد که آیا چرخند میان‌مقیاس می‌تواند منجر به تشکیل تورنادو شود یا نه که این با استفاده از حالت فازی‌سازی ورودی و از فازی در آوردن خروجی امکان‌پذیر است. فلاح قاهری و همکاران نیز (۱۳۸۶) از سامانه استنباط فازی-عصبی تطبیقی و سامانه استنباط فازی ممدانی برای پیش‌بینی بارش سالیانه استفاده کردند. جانسون و لین (۱۹۹۵) با شبکه پس‌انتشاری که دارای توپولوژی $8 \rightarrow 1 \rightarrow 8$ بود توانستند توفند را ردگیری کنند. ورودی‌های این شبکه شامل سری‌های زمانی داده‌های مربوط به توفند از قبیل زمان و طول و عرض جغرافیایی موقعیت وقوع توفند، سرعت و جهت باد بیشینه، حداقل فشار، و اطلاع از مرحله زندگی چرخند بود. اشکال روش آنها این است که داده‌های آموزشی این شبکه فقط اطلاعات توفند یک فصل را در بر دارد.

شبکه‌های عصبی تقریباً از ۱۹۹۱ در نوشتار هواشناسی وارد شدند و امروزه در بسیاری از مراکز پیش‌بینی، از آنها در پیش‌بینی دمای حداقل (اسپیزاس و همکاران، ۱۹۹۱)، مه (نارنجودیز و آلفونزو، ۱۹۹۵)، برآورد بارش از تصاویر ماهواره‌ای (مک کلوخ و همکاران، ۱۹۹۵)، پیش‌بینی ریزش برف و باران از تصاویر رادار هواشناسی (اوچپای و همکاران، ۱۹۹۵)، پیش‌بینی رعد و برق (فرانکل و همکاران، ۱۹۹۱)، پیش‌بینی تورنادو (مرزبان و استامپف، ۱۹۹۸)، تعقیب توفند (هاریکن) (جانسون و لین، ۱۹۹۵)، تشخیص ابر با حس‌گرهای زمینی (اویولت و همکاران، ۱۹۹۶)، طبقه‌بندی ابر با سنجش از دور (بنکرت و آها، ۱۹۹۶)، شناسایی پوشش یخ روی سطح دریا و تعقیب حرکت یخ با استفاده از سنجش از دور (رائو و همکاران، ۱۹۹۴)، پیش‌بینی ال‌نینو (در و اسلوتز، ۱۹۹۴) و مانند آن استفاده می‌شود. از ترکیب شبکه عصبی با منطق فازی نیز در تشخیص جابه‌جایی الگوها و تخریب الگوها (لیا سیدو و همکاران، ۱۹۹۹) و همچنین، پیش‌بینی وضعیت ازن استفاده می‌شود. برای مثال اسپیزاس و همکاران (۱۹۹۱) با استفاده از شبکه عصبی مصنوعی به روش پس‌انتشار توانستند دمای حداقل را به کمک داده‌های سری زمانی دما پیش‌بینی کنند.

آنها با استفاده از رابطه‌های همبستگی موجود در آمار که فقط روابط خطی بین متغیرها را آشکار می‌کند برای بررسی پدیده‌های دینامیکی جو که دارای روابط غیرخطی هستند، چندان مناسب نیست. البته در برخی موارد می‌توان از دانش قبلی فرد خبره در این خصوص کمک گرفت، ولی این دانش از خبره‌ای به خبره دیگر و از پدیده‌ای به پدیده دیگر متفاوت است. برای گریز از این معضل نیز می‌توان از روش‌های آماری و روش‌های هوش مصنوعی یعنی شبکه‌های عصبی مصنوعی و منطق فازی کمک گرفت.

در این تحقیق تلاش شده است که نه فقط این روش‌ها برای تعیین کمیت‌های مرتبط با هر کمیت خاص در یک پدیده توفان حاره‌ای اعمال شوند، بلکه کمیت‌های انتخابی در حکم ورودی به یک دستگاه فازی عصبی که خروجی آن کمیت مورد نظر است داده می‌شود و مقدار مدل برای صحت درستی انتخاب، سنجیده می‌شود.

۲ داده‌ها و روش تحقیق

هدف از این تحقیق، تعیین کمیت‌های مناسب هواشناسی برای بررسی پدیده‌های هواشناسی به کمک روش‌های انتخاب ورودی متداول در هوش مصنوعی است. در این تحقیق برای مثال سرعت باد در ارتفاع ۱۰ متری از سطح زمین در داخل یک توفان حاره‌ای انتخاب شده است (با استفاده از میزان سرعت باد به دست آمده و جدول سفیر-سیمپسون به راحتی می‌توان وقوع توفان حاره‌ای و شدت آن را پیش‌بینی کرد). برای این منظور، داده‌های ۳ توفان حاره‌ای گونو، یمین و نرجس از چند روز قبل از شکل‌گیری توفان تا چند روز بعد از اضمحلال آن از سایت NCEP گرفته شده است. برای اطلاع از محتویات درون فایل‌ها، ابتدا این داده‌ها با برنامه Wgrib از حالت دودویی (باینری) به اسکی تبدیل شدند سپس با استفاده از برنامه‌نویسی به زبان C متغیرهای مورد نیاز در محدوده

بررسی سری‌های زمانی داده‌های هواشناسی نشان می‌دهد که جو دارای ساختار دینامیکی است و نه تنها تغییرات فصلی و شبانه روزی بر آن تاثیر دارد بلکه عوامل ناشناخته‌ای نیز وجود دارند که مانع از عرضه یک مدل دقیق و ثابت برای پیش‌بینی می‌شوند. خاصیت اثر پروانه‌ای لورنز، نیز که مهم‌ترین شعار نظریه آشوب است بیان می‌کند که اگر پروانه‌ای در افریقا بال زند ممکن است گردبادی در امریکای جنوبی شکل گیرد. البته، عبارت «اثر پروانه‌ای» در اینجا در واقع اشاره به وابستگی حساس حل معادلات حاکم بر جو به شرایط اولیه و پارامترهای انتخابی دارد. پس فرایندهای جوئی از ماهیت آشوبی و تصادفی برخوردارند و وجود اصل پدیده آشوب در پژوهش‌های هواشناسی نیز بر غیرممکن بودن پیش‌بینی آب و هوا در دراز مدت تاکید دارند، البته روش‌های تطبیقی تا حدودی می‌توانند این مشکل را حل کنند اما برای آن نیاز به ابر رایانه‌هایی با حافظه بسیار زیاد برای ذخیره همه حالت‌های ممکن برای وقوع پدیده‌ها هست، که عملاً غیرممکن است.

پدیده‌هایی که در جو رخ می‌دهند نیز گرچه با یکدیگر در ارتباط هستند ولی شکل‌گیری آنها از نظر زمان و مکان و دوره تناوب متفاوت است. یک توفان حاره‌ای که روی اقیانوس شکل می‌گیرد و به سمت خشکی‌ها حرکت می‌کند، بعد از طی مسافتی روی خشکی از بین می‌رود. اما کمیت‌های جوئی که هم‌روزه در نقاط گوناگون کره زمین در ایستگاه‌های گوناگون اندازه‌گیری می‌شوند به حدی متنوع هستند که در نظر گرفتن همه آنها در مدل‌سازی‌های این گونه پدیده‌ها، مستلزم تحمیل بار محاسباتی بسیار زیادی است. به همین خاطر نیاز به انتخاب بعضی از این کمیت‌های موثر در بررسی آن پدیده است. البته برخی از این کمیت‌ها اطلاعات موجود در چندین کمیت دیگر را در بر دارند و می‌توان از همبستگی بین آنها کمک گرفت. اما حذف

ویژگی‌های زمانی و مکانی فعالیت توفان‌های مورد بررسی به شرح زیر هستند و مسیر حرکت آنها در شکل ۱ آورده شده است.

۱. توفان گونو، این توفان در اوایل ژوئن ۲۰۰۷ روی آب‌های ناحیه استوایی اقیانوس هند که دارای دمای بیشتر از ۲۶ درجه بودند شکل گرفت. گونو قوی‌ترین چرخند حاره‌ای است که تاکنون در دریای عرب و حتی در شمال اقیانوس هند رخ داده است. حداقل فشار در مرکز این توفان ۹۲۰ هکتوپاسکال، و بیشینه سرعت آن ۲۴۰ کیلومتر در ساعت بود، اما دلایلی که این توفان را بارز کرد سرعت و شدت تخریب آن نبود بلکه موارد زیر بود:

- تاثیر آن بر منطقه خلیج فارس که منجر به توقف چند ساعته صادرات نفت از این منطقه شد.
- وقوع هم‌زمان زلزله در منطقه که بر شایعه وقوع پدیده سونامی دامن زد.

- خاطره بدی که در اذهان مردم از توفند مخرب کاترینا (توفانی که در ۲۰۰۶ در امریکا رخ داد) وجود داشت.

- مسیر حرکت توفان‌های حاره‌ای که تا آن زمان در اقیانوس هند شکل گرفته بودند بیشتر به سمت خلیج بنگال بود اما این توفان با حرکت غیرعادی خود به سمت دریای عرب و سواحل کشور عمان و خلیج فارس، تغییر مسیر داد (مسیر شماره ۱ در شکل شماره ۱).

محدوده مورد بررسی این توفان در این تحقیق ناحیه بین ۱۰ تا ۳۰ درجه عرض جغرافیایی و ۵۲ تا ۷۳ درجه طول جغرافیایی است و زمان مورد بررسی این توفان در این تحقیق از اول تا دهم ماه ژوئن انتخاب شده است.

۲. توفان یمین، این توفان در هفدهم ماه ژوئن به صورت اغتشاش کم‌فشار در ناوه مانسونی روی خلیج بنگال شکل گرفت. این چرخند حاره‌ای با سرعت باد ۵۵ کیلومتر در ساعت به سرعت در حال حرکت به سمت غرب-شمال‌غربی بود اما وجود پشته جنب حاره‌ای قوی

فعالیت هر توفان از فایل اصلی تفکیک شدند. هر فایل دودویی در کل شامل اطلاعات ۲۸۶ کمیت جو و دمای خاک است که در یک شبکه $1 \times 1^\circ$ و به فواصل زمانی ۶ ساعته برای کل کره زمین یعنی از عرض جغرافیایی $90^\circ S$ تا $90^\circ N$ و طول جغرافیایی $180^\circ W$ تا $180^\circ E$ تهیه شده است. البته ۱۹۸ کمیت از ۲۸۶ کمیت مربوط به دما، ارتفاع، مولفه مداری و نصف‌النهار باد گرادیان در سطوح فشاری استاندارد از ۱۰۰۰ تا ۱۰ هکتوپاسکال؛ و رطوبت نسبی، سرعت قائم فشار Pa/s، تاوایی، آب موجود در ابر در سطوح فشاری استاندارد تا ۱۰۰ هکتوپاسکال هستند؛ ۱۱ کمیت مربوط به دما و رطوبت خاک در اعماق ۰ تا ۲ متری از سطح زمین؛ ۲۱ کمیت مربوط به اندازه‌گیری‌های رطوبت نسبی، سرعت مولفه‌های افقی و قائم باد در لایه‌های متفاوت مختصات سیگما هستند. از آنجا که پدیده مورد بررسی در سطح آب‌های اقیانوس شکل می‌گیرد و در ورودی‌های پایین فعالیت می‌کند، فقط حدود ۵۱ کمیت جو در این پژوهش قابل استفاده بود. البته داده‌های سمت و سرعت باد، تاوایی و واگرایی در داده‌های خام موجود نیستند ولی آنها را می‌توان با استفاده از داده‌های مولفه‌های مداری و نصف‌النهار باد گرادیان محاسبه کرد. لازم به ذکر است که ۶ کمیت توفان گونو (سمت باد، سرعت باد و چینش سرعت قائم باد در 2000 (pv units) و (pv units) 34768) به دلیل نبود اندازه‌گیری صحیح آنها در بسیاری از نقاط شبکه محدوده مورد بررسی مرکز ملی پیش‌بینی محیطی (National Centers for Environmental Prediction) (که داده های gfs مدل جو از این مرکز گرفته شده است) به کل از حوزه عملیات بررسی حذف شده‌اند. چون تعداد نقاط مجهول به‌حدی زیاد بود که امکان استفاده از روش‌های گوناگون درون‌یابی میسر نبود. بنابراین فقط ۴۵ کمیت باقی‌مانده در بررسی شرکت داده شده‌اند که فهرست آنها در جدول ۱ آمده است.

سرعت باد آن در هنگام ظهر به ۴۸ کیلومتر در ساعت، فشار سطحی در مرکز سامانه به ۹۹۰ هکتوپاسکال (در ۹۰ کیلومتری کراچی) رسید و در صبح روز ۲۶ ژوئن در امتداد سواحل مکران و بلوچستان وارد خشکی و سپس مضمحل شد (مسیر شماره ۲ در شکل ۱).

محدوده مورد بررسی این توفان در این تحقیق ناحیه بین ۱۵ تا ۳۰ درجه عرض جغرافیایی و ۶۰ تا ۸۵ درجه طول جغرافیایی است. در این تحقیق زمان مورد بررسی این توفان از هجدهم تا بیست و هشتم ماه ژوئن انتخاب شده است.

در شمال خلیج بنگال سبب تضعیف چینش باد و فلج شدن سامانه شد. با این حال به تدریج با کمک همرفت شکل گرفته در منطقه و افزایش محتوای گرمایی اقیانوس تبدیل به چرخند حاره‌ای از نوع 3B شد و در اوایل روز ۲۲ ژوئن وارد خشکی شد که در برهمکنش با خشکی سیر تضعیف پیش گرفت ولی در روز ۲۴ ژوئن که وارد دریای عرب شد، تاوایی چرخندی آن در مواجهه با شارش مانسونی در تراز پایین جو؛ با چینش باد قائم ضعیف و سطح آب گرم افزایش یافت. این چرخند مجدد در روز ۲۵ ژوئن به سرعت فعال شد و با حرکت به سمت دریای عمان از منطقه جنوب شرقی ایران به سمت پاکستان روانه شد.

جدول ۱. کمیت‌های استفاده شده در مدل مورد استفاده برای بهبود پیش‌بینی.

ردیف	کمیت جوی	ردیف	کمیت جوی
۱	آب قابل بارش بر حسب $[kg/m^2]$	۲۳	جهت باد در ۱۰ متری از سطح متوسط دریا
۲	ارتفاع لایه مرزی سیاره‌ای بر حسب $[m]$	۲۴	ارتفاع سطح زمین
۳	دمای سطح زمین بر حسب درجه کلوین	۲۵	فشار در سطح زمین
۴	انرژی پتانسیلی موجود همرفتی در لایه $180 mb$	۲۶	فشار در سطح متوسط دریا
۵	از سطح زمین $[J/kg]$	۲۷	فشار در سطح تروپوپاز
۶	انرژی همرفتی بازدارنده $[J/kg]$ در لایه $180 mb$	۲۸	ارتفاع سطح تروپوپاز
۷	از سطح زمین $[J/kg]$	۲۹	دما در سطح تروپوپاز
۸	انرژی پتانسیلی همرفتی در دسترس در سطح زمین $[J/kg]$	۳۰	سرعت باد در سطح تروپوپاز
۹	انرژی همرفتی بازدارنده $[J/kg]$ در سطح زمین	۳۱	جهت باد در سطح تروپوپاز
۱۰	شاخص فرازش سطحی $[K]$	۳۲	فشار در سطح بیشینه باد
۱۱	بهترین شاخص فرازش (۴- لایه ای) $[K]$	۳۳	ارتفاع سطح بیشینه باد
۱۲	آنومالی ارتفاع ژئوپتانسیلی در $1000 mb$ $[gpm]$	۳۴	دما در سطح بیشینه باد
۱۳	آنومالی ارتفاع ژئوپتانسیلی در $500 mb$ $[gpm]$	۳۵	سرعت باد در سطح بیشینه باد
۱۴	آنومالی ارتفاع ژئوپتانسیلی موج $[gpm]$	۳۶	جهت باد در سطح بیشینه باد
۱۵	دما در 1829 متری از سطح متوسط دریا	۳۷	ارتفاع ترازوی که دما صفر است
۱۶	سرعت باد در 1829 متری از سطح متوسط دریا	۳۸	رطوبت نسبی در ترازوی که دما صفر است
۱۷	جهت باد در 1829 متری از سطح متوسط دریا	۳۹	دما در لایه $30 mb$ از سطح زمین
۱۸	دما در 2743 متری از سطح متوسط دریا	۴۰	سرعت باد در لایه $30 mb$ از سطح زمین
۱۹	سرعت باد در 2743 متری از سطح متوسط دریا	۴۱	جهت باد در لایه $30 mb$ از سطح زمین
۲۰	جهت باد در 2743 متری از سطح متوسط دریا	۴۲	رطوبت نسبی در لایه $30 mb$ از سطح زمین $[\%]$
۲۱	دما در 3658 متری از سطح متوسط دریا	۴۳	نم ویژه در لایه $30 mb$ از سطح زمین $[kg/kg]$
۲۲	سرعت باد در 3658 متری از سطح متوسط دریا	۴۴	دما در 2 متری از سطح زمین
	سرعت باد در 10 متری از سطح متوسط دریا	۴۵	رطوبت نسبی در 2 متری از سطح زمین

۳ روش‌های استخراج ویژگی‌ها

برای انتخاب پارامترهای مناسب برای تحقیق ابتدا باید ویژگی‌های مرتبط با پدیده موردنظر انتخاب شود. برای این منظور روش‌های زیادی وجود دارد (گایون و الیزسیف، ۲۰۰۳) که از بین آنها چند روش رایج به شرح زیر انتخاب شد:

۱. روش جست‌وجوی پی‌درپی پیشرو. ابتدا ویژگی مورد پیش‌بینی به‌داخل یک مجموعه تهی مانند S منتقل می‌شود سپس ویژگی‌های باقی‌مانده یک‌به‌یک در مدل ارزیابی می‌شوند تا ویژگی که کمترین $RMSE$ را برای مدل ایجاد کند به مجموعه S اضافه شود. این عملیات تا جایی ادامه می‌یابد که $RMSE$ مدل نسبت به حالت‌های قبلی افزایش یابد. در این مرحله ویژگی‌های موجود در داخل مجموعه S ویژگی‌هایی هستند که بهترین مدل با کمترین $RMSE$ مطابق روند انتخاب شده را فراهم می‌کنند و می‌توان از آنها در محاسبات استفاده کرد (وروریدیس و کاتروپولوس، ۲۰۰۵؛ گویترز-اوسانا، ۲۰۰۲).

۲. روش جست‌وجوی پی‌درپی پسرو. ابتدا همه ویژگی در داخل مجموعه S قرار می‌گیرد و سپس در هر عملیات، فقط ویژگی حذف می‌شود که با حذف آن مقدار $RMSE$ مدل شروع به کاهش کند و با حذف ویژگی که $RMSE$ مدل شروع به افزایش کند، عملیات خاتمه می‌یابد. ویژگی‌های باقی‌مانده در مجموعه S در حکم ویژگی‌های مطلوب در نظر گرفته می‌شود (گایون و الیزسیف، ۲۰۰۳؛ گویترز-اوسانا، ۲۰۰۲).

۳. معیار اطلاعات متقابل (Mutual Information) کمی است که وابستگی متغیرها را به یکدیگر بیان می‌کند و در قیاس با معیار تحلیل همبستگی که فقط روابط خطی موجود بین متغیرهای تصادفی را آشکار می‌سازد، این معیار توانایی آشکارسازی روابط خطی و غیرخطی موجود بین متغیرها را دارد.

۳. توفان نرجس بعد از توفان نینا (۱۹۷۵) مهلک‌ترین چرخند در بستر اقیانوس هند شناخته شده است. این توفان ابتدا در ۲۷ آوریل ۲۰۰۸ به شکل یک وافشاری در مرکز خلیج بنگال ظاهر شد و چون پشته آن به سمت شمال قرار داشت لذا در ابتدا به کندهی به سمت شمال غرب حرکت می‌کرد. ولی بعد از مواجه شدن با شرایط مطلوب و کسب رطوبت از آب‌های گرم، دیواره‌های متحدالمرکز چشم توفان در آن تشخیص داده شد (مرکز اخطار تایفون (JTWC)، 2008a). در ابتدا، تصور می‌کردند که به سمت بنگلادش یا جنوب شرقی هند حرکت خواهد کرد ولی جریان همگرایی در مرکز آن به‌واسطه جریان‌های نشستی و خشکی هوا تضعیف شد و سامانه درهم ریخته و با تضعیف شدن به سمت شمال شرقی کشیده شد. این سامانه در اواخر ۲۹ آوریل علی‌رغم افزایش چینش باد مجدداً به‌واسطه همرفت هوا شکل گرفت. در اول ماه مه با رسیدن ناوه سطوح فوقانی و برقراری جریان‌های برون‌شارشی، تقویت آن سریع شد؛ به‌نحوی که در اوایل روز دوم، قطر چشم توفان به نوزده کیلومتر رسید بود. مرکز اخطار تایفون سرعت آن را در ساحل برمه ۲۱۵ کیلومتر در ساعت برآورد کرد که این سرعت به مدت یک دقیقه استمرار داشت و آن را به توفان نوع ۴ طبق جدول سفیر سیمسون تبدیل کرد. در دوم ماه مه به بیشینه باد ۱۶۵ کیلومتر در ساعت دست‌یافت که ۳ دقیقه استمرار داشت (مرکز اخطار تایفون، 2008b). حداقل فشار این چرخند ۹۶۲ هکتوپاسکال بود و با ورود به خشکی کشور برمه در سوم مه از بین رفت (مسیر شماره ۳ در شکل ۱). محدوده مورد بررسی این توفان در این تحقیق ناحیه بین ۷ تا ۲۱ درجه عرض جغرافیایی شمالی و ۸۲ تا ۱۰۰ درجه طول جغرافیایی شرقی است. در این تحقیق زمان مورد بررسی این توفان از بیست و پنجم آوریل تا چهارم ماه مه انتخاب شده است.

ممکن عضوهای دو مجموعه $((f_i \in F), (f_s \in S))$ یعنی

$$MI(f_i; f_s)$$

معیار $MI(f_i; y) - \frac{\beta}{M} \sum_{f_s \in S} MI(f_i; f_s)$ برای عضوهای

دو مجموعه $((f_i \in F), (f_s \in S))$ محاسبه شود M تا از

زیر مجموعه‌های دوتایی از ورودی‌هایی که این معیار را

بیشینه می‌کنند، به مجموعه S انتقال یابند. در این رابطه

β عددی درون بازه $[1, 1/5]$ است که برای وزن‌دهی

بین اطلاعات متقابل ورودی انتخاب شونده با خروجی

است (رضائی یوسفی، ۱۳۸۶). $|S|$ تعداد عضوهای

مجموعه S است و y خروجی شبکه است.

تحلیل مولفه‌های اصلی (Principal Components

Analysis (PCA)) که بیشتر در تحلیل داده اکتشافی و

همچنین در مدل‌های پیش‌بینی برای کاهش پیچیدگی

متغیرهای اولیه از آن استفاده می‌شود (اویانگ، ۲۰۰۵). در

این روش داده‌ها به مولفه‌های جدید و مستقل از هم تبدیل

می‌شوند که ترکیب خطی از متغیرهای اولیه هستند. اعتبار

استفاده از این روش با پارامتر

KMO (Kaiser - Meyer - Olkin) که مقدارش بین

صفر تا یک متغیر است، سنجیده می‌شود (نوری و

همکاران، ۱۳۸۶). اگر این پارامتر بزرگ‌تر از ۰/۵ باشد،

می‌توان از این روش برای تحلیل داده‌ها استفاده کرد

(سینگ و همکاران، ۲۰۰۴؛ کامدویرین و همکاران،

۲۰۰۵).

مراحل گوناگون تجزیه به عامل‌های اصلی در این

روش به‌قرار زیر است:

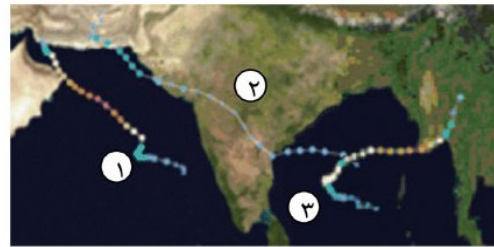
۱. محاسبه KMO ، پارامتر KMO به کمک ضرایب

همبستگی ساده و ضرایب همبستگی جزئی طبق رابطه

زیر محاسبه می‌شود.

KMO

$$= \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad i \neq j \quad (2)$$



شکل ۱. مسیر حرکت سه توفان حاره‌ای مورد بررسی. مسیر ۱ - توفان

حاره‌ای گونو؛ مسیر ۲ - توفان حاره‌ای یمین؛ مسیر ۳ - توفان حاره‌ای

نرجس.

بنابراین، برای دستگاه‌های غیرخطی با دینامیک

پیچیده چند بُعدی نظیر پدیده‌های جوئی مثل توفان، بسیار

مناسب است. این کمیت برای دو متغیر تصادفی گسسته

مانند X و Y به‌صورت زیر تعریف می‌شود:

$$MI(x, y) = \sum_{y \in Y} \sum_{x \in X} P(x, y) \log \left(\frac{P(x, y)}{P_1(x)P_2(y)} \right) \quad (1)$$

که در آن، $P(x, y)$ تابع توزیع احتمال مشترک X و Y و

$P_1(x)$ و $P_2(y)$ به‌ترتیب توابع توزیع احتمال حاشیه‌ای X

و Y است.

برای انتخاب ویژگی در این روش چندین الگوریتم

وجود دارد که یکی از آنها الگوریتم همبستگی متقابل

نزدیک به بهینه براساس انتخاب ویژگی (NOMIFS)

Near Optimal Mutual Information based Feature (

Selection) به شرح زیر است:

■ قرار دادن کل ورودی‌ها در مجموعه‌ای مانند F و

در نظر گرفتن مجموعه S به صورت تهی

■ محاسبه اطلاعات متقابل بین تک‌تک ورودی‌ها با

خروجی

■ انتخاب M ورودی که بیشترین اطلاعات متقابل را

با خروجی دارد عدد M باید کوچکتر از تعداد کل

ورودی‌ها باشد برای مثال نصف آنها و انتقال آن از

مجموعه F به S محاسبه اطلاعات متقابل بین تمام زوج‌های

که جملات خطا در روش PFA همگی دارای واریانس یکسان‌اند، در آن صورت این دو روش یکسان می‌شوند.

۴ دستگاه استنباط فازی-عصبی تطبیقی (ANFIS)

ترکیبی از دو مدل فازی و شبکه عصبی است که در بخش فازی رابطه بین ورودی و خروجی به کمک دانش تجربی عرضه می‌شود و در بخش دیگر پارامترهای مربوط به توابع عضویت بخش فازی که از داده‌های عددی در شناسایی پارامترها استفاده می‌کند.

این دستگاه شامل ۵ لایه است (طاهر و ناصح، ۲۰۰۹؛ نورانی و صالحی، ۱۳۸۷) که به شرح زیر هستند:

۱. لایه اول یا لایه فازی‌سازی ورودی‌ها: که در آن ورودی‌های (X_1, X_2, \dots, X_n) به متغیرهای فازی (A_i) با توابع عضویت (membership functions) $\mu_{A_i}(x)$ متناسب با گره‌های مربوطه تبدیل می‌شوند. این توابع می‌توانند سیگموئید، گاوسی متقارن، اسپلاین و مانند آن باشند.

۲. لایه دوم: در این لایه قوانین (عملگرهای «و»؛ «یا») اعمال می‌شود.

۳. لایه سوم: در این لایه عمل استنتاج از مقدمه نتیجه به شکل نرمال‌سازی وزن قوانین صورت می‌گیرد.

۴. لایه چهارم: در این لایه نتایج قوانین گره‌ها با یکدیگر تطبیق داده می‌شود.

۵. لایه پنجم: در این لایه عمل غیرفازی کردن خروجی صورت می‌گیرد.

۵ پیاده‌سازی مدل

بعد از تفکیک داده‌های ۴۵ کمیت انتخابی توفان‌های گونو، یمین و نرجس از بین داده‌های جهانی و قرار دادن آنها در سه ماتریس $N \times 45$ مجزا که N تعداد مشاهدات هر کمیت است، داده‌ها از نظر توزیع نرمال بودن با

که در آن r_{ij} ضریب همبستگی ساده بین متغیرهای i و j ؛ و a_{ij} ضرایب همبستگی جزئی متغیرهای i و j به شرط ثابت بودن سایر متغیرها است.

۲. استاندارد کردن داده‌ها با استفاده از روش‌های موجود

۳. محاسبه ماتریس کواریانس داده‌ها

۴. محاسبه بردارهای ویژه و مقادیر ویژه ماتریس کواریانس و واریانس مولفه‌های استخراجی

۵. بازچینی بردارهای ویژه برای کاهش ابعاد داده

۶. انتخاب مؤلفه‌ها و قرار دادن آنها در ماتریس بردار ویژگی

۷. استخراج داده‌های جدید از طریق ضرب ترانهاده ماتریس بردار ویژگی در ترانهاده داده‌های استاندارد شده

تحلیل مولفه‌های عاملی (principal Factor Analysis) روشی است که با تحقیق درباره ماهیت روابط متقابل بین متغیرهای یک مجموعه معین، مشخص می‌کند که آیا این مجموعه متغیر را می‌توان برحسب تعداد کمتری از ابعاد (عامل‌ها) در قیاس با کل تعداد متغیرها توصیف کرد یا نه؟ و هر یک از عامل‌ها معرف چه صفت یا ویژگی‌اند (ویلی، ۲۰۰۱). در این روش متغیرهای مشاهده شده به صورت ترکیب‌های خطی تعداد کمتری از متغیرهای تصادفی غیر قابل مشاهده بیان می‌شوند.

در اینجا ذکر این نکته ضروری است که اغلب PCA را با PFA یکسان فرض می‌کنند اما باید توجه داشت که این دو با هم در ارتباط هستند اما لزوماً یکسان نیستند چون در PCA واریانس کل داده‌ها بررسی می‌شود و با چرخش مولفه سعی می‌شود که با بیشینه کردن فضای متغیرها، همه تغییرپذیری در متغیرها بررسی شود. اما برعکس در PFA، واریانس مشترک بررسی می‌شود تا میزان تغییرپذیری به‌واسطه عامل‌های مشترک تعیین شود. چنانچه فرض شود

معیارهای کشیدگی، چاولگی و آزمون کولموگوروف - اسمیرنوف بررسی شدند. نتایج آزمایش کولموگوروف - اسمیرنوف برای متغیرهای سه توفان مورد بررسی در شکل ۲ آورده شده است. در این بررسی در مورد کمیت‌هایی که از توزیع نرمال برخوردار نبودند در نرمال‌سازی از کمیت میانه به جای میانگین استفاده شده است.

استانداردسازی داده‌های کمیت‌ها بعد از اطمینان از توزیع داده‌های آنها به کمک رابطه $X_{ji}' = \frac{X_{ji} - \bar{X}_j}{\sigma}$ صورت گرفت که در آن X_{ji} مشاهده i ام از کمیت j ام، \bar{X}_j میانگین کمیت j ام و σ انحراف معیار کمیت j ام است.

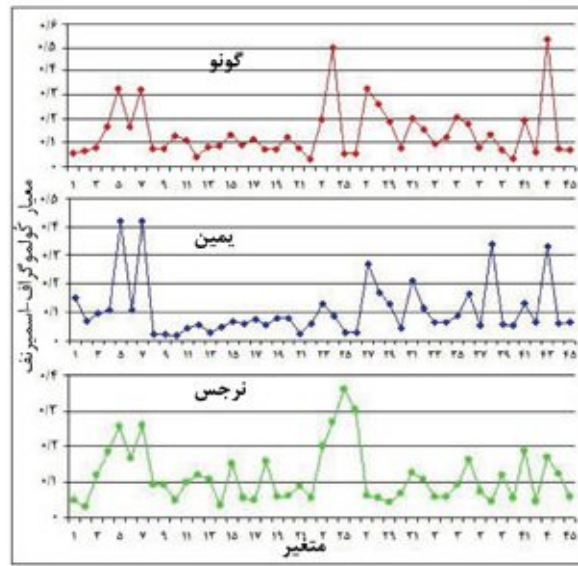
برای تحلیل هر پدیده‌ای با استفاده از داده‌های جمع‌آوری شده، ابتدا باید بُعد مشاهدات را طوری کاهش دهیم که با از دست دادن حداقل اطلاعات همراه باشد. به عبارتی بررسی خود را به متغیرهایی اختصاص دهیم که از یک مشاهده به مشاهده دیگر زیاد تغییر نکنند تا بتوانیم آنها را به منزله ثابت‌های مسئله در نظر بگیریم. بدین منظور متغیرهای با واریانس کم را حذف اما متغیرهای با واریانس زیاد را حفظ می‌کنیم. در این بخش چند روش از بین روش‌های رایج انتخاب ویژگی مانند روش جست‌وجوی پی‌درپی پیشرو، پسرو، معیار اطلاعات متقابل، تحلیل مولفه‌های اصلی و تحلیل عاملی انتخاب و روی داده‌ها اعمال شدند که نتایج سه روش نخست در شکل ۳ آورده شده است. البته در پایین این شکل جدولی عرضه شده که در آن ویژگی‌های استخراج شده مناسب برای هر روش با رنگ کرم مشخص شده است. در این محاسبات در روش‌های جست‌وجوی پی‌درپی پیشرو، پسرو برای انتخاب متغیرهای مناسب از داده‌های توفان گونو برای آموزش و از داده‌های توفان‌های یمین و نرجس در حکم داده‌های آزمون و اعتبارسنجی شبکه استفاده شده است. از آنجا که پارامتر KMO برای سه توفان گونو، یمین و نرجس به ترتیب ۰/۸۴۴، ۰/۸۵۳ و ۰/۸۰ به دست آمد،

در نتیجه با اطمینان می‌توانیم از روش PCA برای تقلیل متغیرها استفاده کنیم. نتایج حاصل از بررسی تغییرات واریانس که در شکل ۴ آورده شده است نشان می‌دهد که مقادیر ویژه مربوط به ۱۸ مولفه نخست توفان‌های گونو و یمین بزرگ‌تر از بقیه بوده‌اند و بیش از ۹۵٪ اطلاعات موجود در ۴۵ متغیر را به دست می‌دهند. در نتیجه می‌توان به راحتی از بقیه مولفه‌ها صرف‌نظر کرد. اما در توفان نرجس، این مسئله برای ۲۰ مولفه نخست صادق است. در اینجا فقط ضرایب ترکیب خطی این مولفه برای داده‌های توفان گونو در جدول ۲ و شکل ۵ آورده شده است. بررسی مقادیر ضرایب مولفه‌ها نشان می‌دهد که بخش عمده دومین مولفه مربوط به کمیت سرعت باد در ارتفاع ۱۰ متری از سطح زمین و سرعت باد در لایه ۳۰ هکتوپاسکالی از سطح زمین است. به همین خاطر در استفاده از شبکه ANFIS از این مولفه می‌توان در حکم خروجی شبکه استفاده کرد.

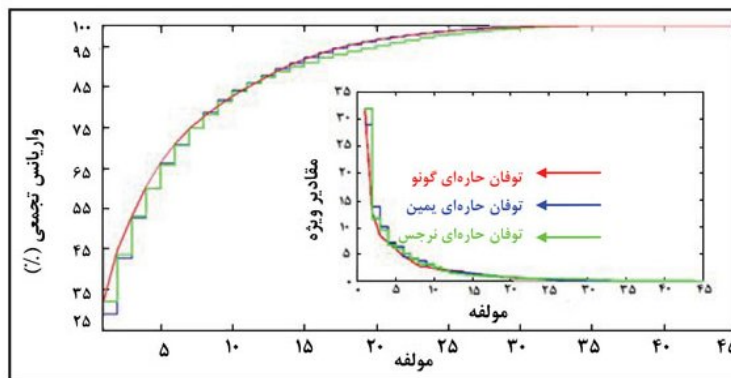
داده‌های کمیت‌های مورد بررسی با استفاده از روش PFA نیز بررسی شدند که نتایج بررسی تغییرات واریانس آنها در شکل ۶ آمده است.

ویژگی‌های استخراج شده از ۳ روش نخست با شبکه عصبی ANFIS مورد آزمون قرار گرفت. شبکه عصبی ANFIS مورد استفاده در این تحقیق دارای ویژگی‌های زیر است.

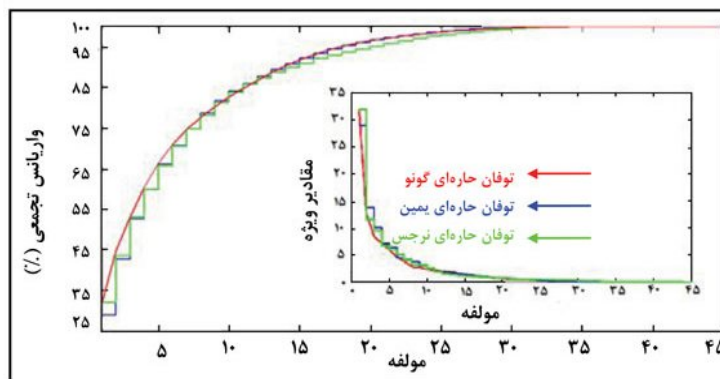
۱. برای ویژگی‌های استخراج شده از روش جست‌وجوی پی‌درپی پیشرو، شبکه هوشمند از ۱۳ ورودی داده شده در جدول متصل به شکل ۳ (در جدول به رنگ کرم مشخص شده‌اند) و یک خروجی استفاده می‌کند که شرح آنها در جدول ۱ آمده است. خروجی این شبکه سرعت باد در ارتفاع ۱۰ متری از سطح زمین است. طرح‌واره‌ای از این شبکه هوشمند در شکل ۷ نشان داده شده است.



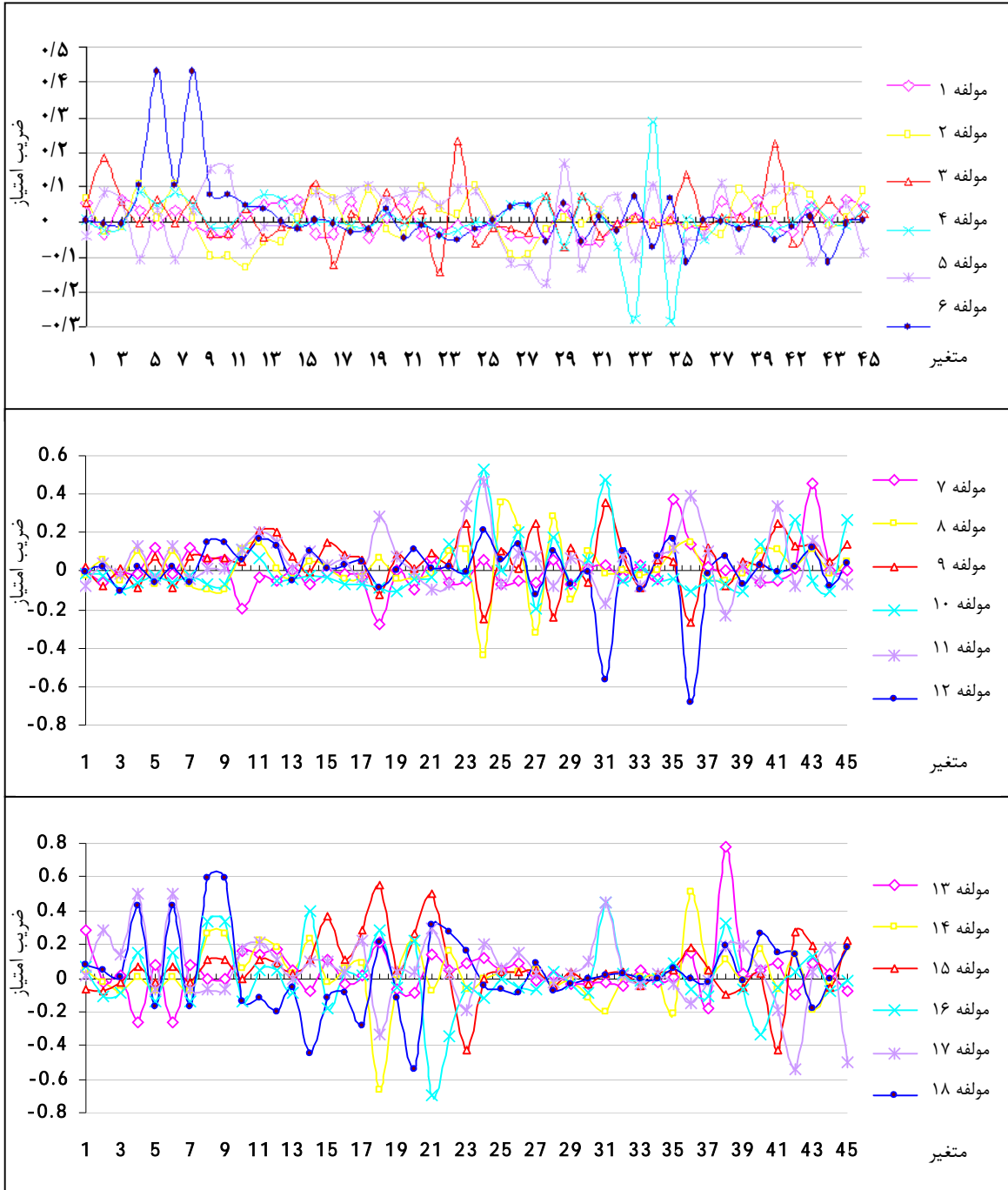
شکل ۲. آزمون کولموگوروف-اسمیرنوف برای متغیرهای سه توفان، قبل از استانداردسازی.



شکل ۳. ویژگی‌های استخراجی برای سه روش پی‌درپی پیشرو، پسرو و همبستگی متقابل همراه با میزان RMSE، در جدول زیر شکل ترتیب استخراج ویژگی آورده شده است که در آن ویژگی‌های مناسب به رنگ کرم مشخص شده‌اند.



شکل ۴. واریانس تجمعی و مقادیر ویژه متغیرهای توفان‌های مورد بررسی در تحلیل مولفه‌های اصلی.



شکل ۵. ضرب امتیاز هر متغیر در ترکیب مولفه‌های اصلی برای داده‌های توفان حاره‌ای گونو.

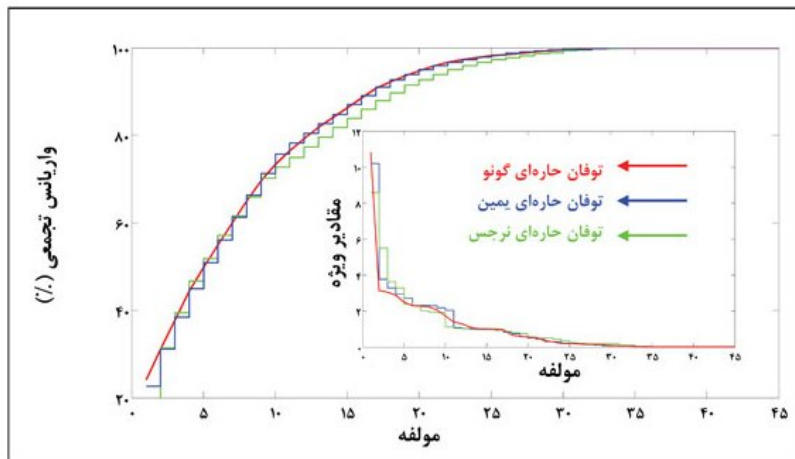
و ... x_N متعلق به خوشه X_N است آن‌گاه y_1 متعلق به خوشه Y_1 تعیین شده‌اند.

این شبکه‌ها برای تعداد زوج‌های آموزشی متفاوت اجرا شدند که نتایج به‌دست آمده در شکل ۸ عرضه شده است. در این آزمایش داده‌های گونو در حکم داده‌های آموزش و یمین در حکم داده‌های آزمون و نرجس در حکم داده‌های اعتبارسنجی شرکت داده شدند. البته این ویژگی‌ها به‌مثابه ورودی و سرعت باد در ۱۰ متری از سطح زمین در حکم خروجی شبکه در نظر گرفته شدند.

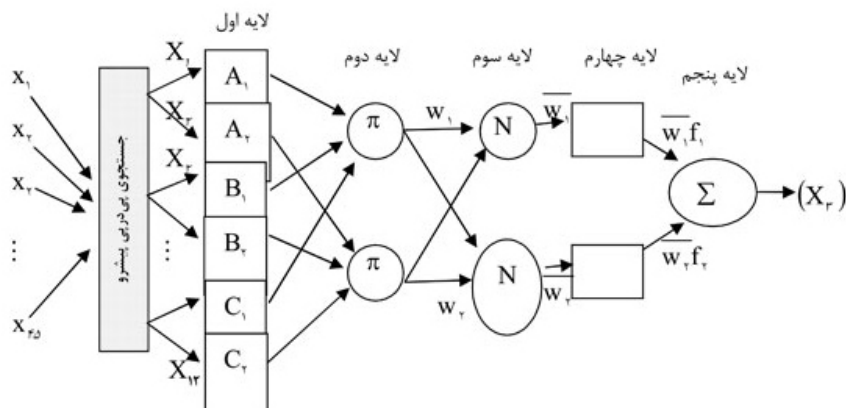
۲. شبکه هوشمند برای ویژگی‌های استخراج شده از روش جست‌وجوی پی‌درپی پسرو از نوزده ورودی داده شده در جدول متصل به شکل ۳ و یک خروجی استفاده می‌کند.

۳. شبکه هوشمند برای ویژگی‌های استخراج شده از روش همبستگی متقابل نیز از پانزده ورودی داده شده در جدول متصل به شکل ۳ و یک خروجی استفاده می‌کند.

۴. توابع عضویت ورودی‌ها در این بررسی به‌شکل گاوسی، تابع عضویت خروجی به شکل خطی و قوانین نیز با استفاده از و یا به‌صورت « اگر x_1 متعلق به خوشه X_1

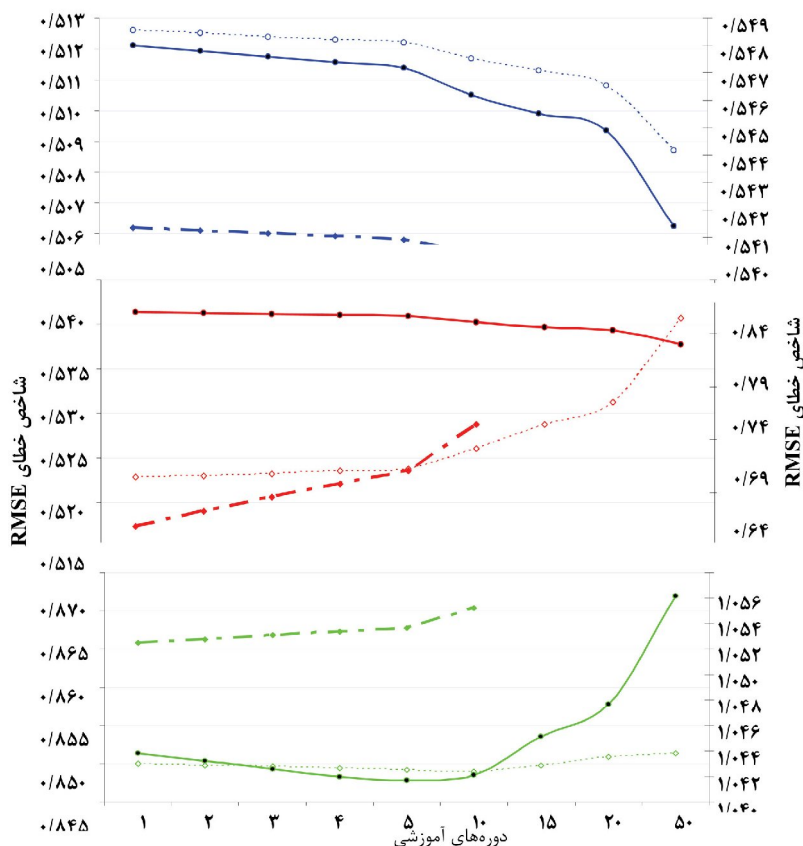


شکل ۶. واریانس تجمعی و مقادیر ویژه متغیرهای توفان‌های مورد بررسی در تحلیل مؤلفه‌های عاملی.



شکل ۷. طرح‌واره‌ای از شبکه هوشمند (ترکیبی از شبکه ANFIS و سامانه جست‌وجوی پی‌درپی پیشرو) برای پیش‌بینی سرعت باد در ارتفاع ۱۰ متری از سطح

زمین.



شکل ۸ نتایج پیش‌بینی سرعت باد در ارتفاع ۱۰ متری از زمین با استفاده از شبکه هوشمند (ترکیب ANFIS با یکی از روش‌های جست‌وجو): منحنی خط چین ضخیم برای ترکیب ANFIS با روش جست‌وجوی پیشرو، منحنی خط چین نازک برای ترکیب ANFIS با روش همبستگی متقابل و منحنی توپر برای ترکیب ANFIS با روش جست‌وجوی پسرو است. آموزش شبکه به روش ترکیبی که ترکیبی از روش کمترین مجذورات^۱ با روش گردایان نزولی پس‌انتشار خطا است. در این پیش‌بینی داده‌های توفان گونو برای آموزش، یمین برای آزمون و نرجس برای اعتبارسنجی به کار رفته است. مقادیر عرضه شده در سمت راست محور قائم مربوط به شبکه هوشمندی است که از ترکیب ANFIS با روش جست‌وجوی پسرو ساخته شده است.

۶ نتیجه‌گیری

زمان محاسبات در روش جست‌وجوی پی‌درپی پسرو طولانی‌تر از روش پیشرو است و از دقت زیادی نیز برخوردار است. چون شروع محاسبات با تمام ویژگی‌ها است و در هر نوبت یکی از ویژگی‌های نامناسب حذف می‌شود و این تا جایی ادامه می‌یابد که شاخص انتخاب ویژگی‌ها بهینه شود و لزوماً تا آخرین ویژگی ادامه نمی‌یابد. برای مثال، با روش پسرو از مجموع ۴۵ ویژگی به ۲۰ ویژگی می‌رسیم و حجم محاسبات کل نیز متناسب است با :

$$\text{حجم محاسبات (۴۵)} + \text{حجم محاسبات (۴۴)} + \dots + \text{حجم محاسبات (۲۱)} + \text{حجم محاسبات (۲۰)}.$$

حال آنکه ویژگی‌های مناسب در روش پیشرو یک‌به‌یک افزوده می‌شود. در نتیجه حجم محاسبات در شروع پردازش کمتر است و در عین حال لازم نیست انتخاب ویژگی‌ها تا انتها ادامه یابد. در حقیقت این انتخاب ویژگی‌ها تا مرحله‌ای که شاخص انتخاب ویژگی بهینه شود ادامه خواهد یافت. برای مثال از مجموع ۴۵ ویژگی با کمک این روش از ۱ به ۱۳ ویژگی می‌رسیم و حجم محاسبات کل متناسب است با :

مطالعه موردی رودخانه کارون: مجله علمی-پژوهشی آب و فاضلاب، ۶۳، ۶۹-۶۰.

نورانی، و. و صالحی، ک.، ۱۳۸۷، مدل‌سازی بارش-رواناب با استفاده از روش شبکه عصبی فازی تطبیقی و مقایسه آن با روش‌های شبکه عصبی و استنتاج فازی: چهارمین کنگره ملی مهندسی عمران، دانشگاه تهران.

فلاح قاهری، غ. و وحیدیان کامیاد، ع.، موسوی بایگی، م.، و حبیبی نوخندان، م.، ۱۳۸۶، کاربرد سیستم استنباط فازی-عصبی تطبیقی (ANFIS) و سیستم استنباط فازی ممدانی در پیش‌بینی بارش سالیانه: هشتمین کنفرانس سیستم‌های هوشمند، دانشگاه فردوسی مشهد.

Aviolat, F., Cattani, D., and Cornu, T., 1996, Recognition of meteorological situations with neural networks: Proceedings of the Third Biennial Joint Conference on Engineering Systems Design & Analysis, ESDA, Montpellier, July.

Bankert, R. L., and Aha, D. W., 1996, Improvement to a neural network cloud classifier: Journal of Applied Meteorology, 35, 2036-2039.

Camdevyren, H., Demyr, N., Kanik, A., and S. Keskin, 2005, Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs: Ecol. Model., 181, 581-589.

Derr, V. E., and Slutz, R. J., 1994, Prediction of el nino events in the pacific by means of neural networks: AI Applications, 8(2), 51-63.

Frankel, D., Schiller, I., Draper, J. S., and Barnes, A. A. Jr., 1991, Use of neural networks to predict lightning at Kennedy space center: International Joint Conference on Neural Networks, (in Seattle, Wa.), IEEE, New York, 319-324.

Gutierrez-Osuna, R., 2002, LECTURE 11: Sequential feature selection from introduction to pattern analysis: Texas A & M University (http://research.cs.tamu.edu/prism/lectures/pr/pr_111.pdf).

Guyon, I., and Elisseeff, A., 2003, An introduction to variable and feature selection: Journal of Machine Learning Research, 3, 1157-1182.

حجم محاسبات (۱) + حجم محاسبات (۲) + ... + حجم محاسبات (۱۲) + حجم محاسبات (۱۳).

تحلیل مولفه‌های اصلی چون از ترکیب همه کمیت‌ها استفاده می‌کند مناسب‌تر است ولی نسبت به بقیه پیچیدگی بیشتری دارد.

همان‌گونه که در شکل ۸ ملاحظه می‌کنید، شبکه هوشمند با ویژگی‌های استخراج شده از روش جست‌وجوی پی‌درپی پسرو نسبت به دو روش دیگر بهتر آموزش دیده است. نتایج این بررسی برای مثال عرضه شده نشان داد که می‌توان با استفاده از شبکه هوشمند کمیت‌های مرتبط با هر کمیت را بدون نیاز به دانش فرد خبره از بین کمیت‌های موجود انتخاب کرد. امروزه در مراکز بررسی توفان فرایند پیش‌بینی براساس پیش‌بینی جمعی صورت می‌گیرد که همه کمیت‌های اندازه‌گیری شده هواشناسی در آن بررسی می‌شود. این تحقیق روشن ساخت که نه تنها می‌توان برخی از کمیت‌ها را از بررسی حذف کرد بلکه نیازی به عملیات زمان‌بر در مراکز موردنظر نیست.

منابع

رضائی یوسفی، م. م.، ۱۳۸۶، انتخاب متغیرهای ورودی در شناسایی سیستم‌ها و کاربرد آن در پیش‌بینی سری‌های زمانی: پایان‌نامه (کارشناسی ارشد در مهندسی برق-گرایش کنترل)، پردیس فنی دانشگاه تهران.

طاهر، ع. و ناصح، م.، ۲۰۰۹، پیش‌بینی کوتاه‌مدت بار با استفاده از شبکه‌های عصبی RBF/BP و سیستم استنتاج وفقی فازی-عصبی ANFIS در محیط قیمت گذاری بلادرنگ: بیست و چهارمین کنفرانس بین‌المللی برق، تهران.

نوری، ر.، کراچیان، ر. و خدادادی دربان، ا.، ۱۳۸۶، ارزیابی اهمیت ایستگاه‌های پایش کیفی رودخانه با استفاده از آنالیز مولفه‌های اصلی و آنالیز فاکتور،

- Ochiai, K., Suzuki, H., Shinozawa, K., Fujii, M., and Sonehara, N., 1995, Snowfall and rainfall reforecasting from weather radar images with artificial neural networks: Proceedings of IEE International Conference on Neural Networks, Nov. 27-Dec. 1, (in Perth, Australia), IEEE, New York, New York, U. S. A., 1182-1187.
- Ouyang, Y., 2005, Application of principal component and factor analysis to evaluate surface water quality monitoring network: *Water Res.*, **39**, 2621-2635.
- Rau, Y. C., Cosimo, J. C., and Lure, F. Y. M., 1994, Application of neural networks for ice coverage and movements from satellite imagery: Proceedings of the International Geoscience and Remote Sensing Symposium. Surface and Atmospheric Sensing: Technologies, Data Analysis and Interpretation (August 8-12, in Pasadena, Ca., USA), IEEE, New York, U. S.A., 1407-1409.
- Schizas, C. N., Michaelides, S. C., Pattichis, C. S., and Livesay, R. R., 1991, Artificial neural networks in forecasting minimum temperature: *Institution of Electrical Engineers, Publ.*, **349**, 112-114.
- Singh, K. P., Malik, A., Mohan, D., and Sinha, S., 2004, Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti river (India) a case study: *Water Res.*, **38**, 3980-3992.
- Ververidis, D., and Kotropoulos, C., 2005, Sequential forward feature selection with low computational cost, (<http://www.eurasip.org/Proceedings/Eusipco/Eusipco2005/defevent/papers/cr1411.pdf>).
- Wyly, E., 2001, Quantitative geographical analysis principal components and factor analysis: Background paper for March 19, seminar discussion. http://www.geog.ubc.ca/~ewyly/teaching/606_pca.pdf.
- Johnson, G. P., and Lin, F. C., 1995, Hurricane tracking by backpropagation neural network: Proceedings of the IEE International Conference on Neural Networks, Nov. 27-Dec. 1, (in Perth, Australia), IEEE, New York, U. S. A., 1103-1106.
- Joint Typhoon Warning Center, 2008a, Cyclone nargis warning NR 006: <ftp://ftp.met.fsu.edu/pub/weather/tropical/GuamStuff/2008042821-WTIO.PGTW>.
- Joint Typhoon Warning Center, 2008b, Cyclone nargis warning NR 020: <ftp://ftp.met.fsu.edu/pub/weather/tropical/GuamStuff/2008043009-WTIO.PGTW>.
- Kalogirou, S., Neocleous, C., Michaelides, S. C., and Schizas, C. N., 1997, A time series reconstruction of precipitation records using artificial neural networks: 5th European Congress on Intelligent Techniques and Soft Computing, EUFIT 97, Aachen, Germany, September, 8-11.
- Liassidou, F., Michaelides, S. C., Neocleous, C. C., and Schizas, C. N., 1999, Identification of synoptic patterns on weather charts by artificial neural networks: 5th International Conference on Engineering Applications of Neural Networks, EANN99, Warsaw, Poland, September.
- Marzban, C., and Stumpf, G. C., 1998, A neural network for tornado prediction based on doppler radar-derived attributes: *Journal of Applied Meteorology*, **35**, 17-626.
- McCullagh, J., Bluff, K., and Ebert, E., 1995, A neural network for rainfall estimation: Conference Paper, Proceedings of the Second New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, IEEE Computing Society Press, Los Alamitos, Ca. U.S.A., 389-392.
- Naranjo-Diaz, L. R., and Alfonso, A. P., 1995, Fog forecasting in Cuba, neural networks versus discriminant analysis: *Meteorological Applications*, **1**, 31-33.