Archive of SID

مهندسی صنایع و مدیریت شریف، (تابستان ۱۳۹۸)
دوره‌ی ۱-۳۵، شماره‌ی ۱/۲، ص. ۱۱۳-۱۲۱، (یادداشت فنی)

Research Note

پیش‌بینی کوتاه‌مدت تقاضای فصلی الکتریسیته با استفاده از مدل‌های ترکیبی هوشمند نرم

فاطمه چاهکوتاهی * **(دانشجوی کارشناسی ارشد)**
مهدی خاشعی (استادیار)
دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان

روش‌های پیش‌بینی از کارآمدترین ابزارهای موجود به‌منظور اتخاذ تصمیمات مدیریتی در حوزه‌های مختلف علوم هستند. دقت پیش‌بینی‌ها یکی از مهم‌ترین عامل‌های مؤثر بر کیفیت تصمیمات اتخاذی است که رابطه‌ی مستقیمی با کیفیت این تصمیمات دارند. پیش‌بینی تقاضای الکتریسته یکی از چالش‌برانگیزترین حوزه‌های پیش‌بینی است. مشخصه‌ی منحصربه فرد الکتریسته، که پیش‌بینی را در مقایسه با سایر کالاهای تولیدی دشوارتر می‌سازد، عدم امکان ذخیره‌سازی آن به‌منظور مصرف در آینده است. این موضوع سبب ایجاد سطح بالایی از ابهام در داده‌های مرتبط با این‌گونه از بازارها می‌شود. ابزارهای هوش محاسباتی و محاسبات نرم از جمله دقیق‌ترین روش‌های حال حاضر به‌منظور مدل‌سازی عدم قطعیت موجود در داده‌ها هستند. در این مقاله، با ترکیب روش‌های مذکور، یک روش هوشمند نرم به‌منظور پیش‌بینی الکتریسیته ارائه شده است. ایده‌ی اصلی مدل استفاده‌ی هم‌زمان از مزایای ابزارهای هوش محاسباتی و محاسبات نرم در مدل‌سازی سیستم‌های پیچیده است. نتایج نشان‌دهنده‌ی دقت بالاتر روش پیشنهادی نسبت به سایر مدل‌هاست.

واژگان کلیدی: ابزارهای هوش محاسباتی و محاسبات نرم، پیش‌بینی سری‌های زمانی، تقاضای فصلی الکتریسیته، پرسپترون‌های چندلایه (MLP)، خودرگرسیون میانگین متحرک انباشته‌ی فصلی (SARIMA).

f.chahkoutahi@in.iut.ac.ir
khashei@cc.iut.ac.ir

۱. مقدمه

پیش‌بینی تقاضای الکتریسیته نقش بسیار حیاتی در عملکرد عملیاتی - اقتصادی و همچنین امنیت سیستم‌های قدرتی ایفا می‌کند. از آن‌جا که بسیاری از تصمیم‌های مهم که عواقب اقتصادی نیز دارند بر اساس تقاضای الکتریسیته‌ی پیش‌بینی شده گرفته می‌شوند، دقت پیش‌بینی تقاضای الکتریسیته به‌طور فزاینده‌یی مورد توجه مدیران مالی و عملیاتی نیروگاه‌های تولید و توزیع الکتریسیته قرارگرفته است. بنابراین مطالعه‌ی پیش‌بینی بار الکتریکی از اهمیت حیاتی برخوردار است. محققان در دو دهه‌ی اخیر مدل‌های بسیار متنوعی را برای پیش‌بینی تقاضای الکتریسیته پیشنهاد کرده‌اند. چوی و همکارانش یک مدل تجزیه و تحلیل چندمتغیره با متغیرهای مستقل اقتصادی، جمعیتی و آب‌وهوایی ارائه کرده‌اند.[۱]

گویا و همکاران یک مدل پیش‌بینی اوج بار با استفاده از داده‌های تقاضای گرما در یک سیستم گرمایشی ساخته‌اند.[۲] نقطه‌ی ضعف این‌گونه از روش‌ها این است که تعیین متغیرهای توضیح‌دهنده‌ی مؤثر به‌منظور حصول نتایج مطلوب بسیار زمان‌بر و دشوار است. این متغیرها اساساً عوامل تعیین‌کننده‌ی دقت پیش‌بینی هستند؛ بنابراین، انتخاب آنها بسیار مهم و در عین حال دشوار است. در مقابل این‌گونه از مدل‌های سببی، مدل‌های بسیار متنوعی برای پیش‌بینی تقاضای الکتریسیته براساس روش‌های سری‌های زمانی ارائه شده است. برخی از مهم‌ترین این روش‌ها عبارت‌اند از: سری‌های زمانی فازی،[۳] مدل‌های خاکستری،[۴و۵] مدل‌های شبکه‌های عصبی مصنوعی،[۶و۷] رگرسیون بردار پشتیبان،[۸و۹] تبدیل موجک و مدل‌های تطبیقی،[۱۰] مدل‌های رگرسیون کلاسیک خطی[۱۱و۱۲] و روش‌های فصلی سه‌تایی.[۱۳]

برای پیش‌بینی سری‌های زمانی و به‌خصوص پیش‌بینی تقاضای الکتریسیته، طیف گسترده‌یی از مدل‌ها در دو دهه‌ی اخیر ارائه شده‌اند. هر یک از این مدل‌ها با این‌که مزایای بسیاری دارند، محدودیت‌هایی نیز دارند که از محبوبیت آن‌ها کاسته است. یکی از شناخته شده‌ترین روش‌ها مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی است. محدودیت اصلی این گونه از روش‌ها این است که مقدار آینده‌ی یک متغیر به‌صورت تابع خطی از چند مشاهده‌ی گذشته و خطاهای تصادفی فرض

* نویسنده مسئول
تاریخ: دریافت ۱۳۹۶/۲/۱۰، اصلاحیه ۱۳۹۶/۱۰/۲۶، پذیرش ۱۳۹۶/۱۱/۱۱.
DOI:10.24200/J65.2019.7089.1699

www.SID.ir

Archive of SID

می‌شود. همچنین این گونه از روش‌ها نیازمند داده‌های زیادی به‌منظور حصول نتایج مطلوب هستند. مدل‌های متعددی در پیشینه‌ی موضوع برای غلبه بر این محدودیت‌ها ارائه شده است. نتایج به دست آمده از شواهد علمی و عملی نشان‌گر آن است که ادغام و ترکیب مدل‌های مختلف می‌تواند یک روش مؤثر برای غلبه بر محدودیت‌های مدل‌های تکی و همچنین بهبود دقت پیش‌بینی‌ها باشد. اخیراً مدل‌های ترکیبی بسیاری در پیشینه‌ی موضوع روش‌های پیش‌بینی ارائه شده است.

این‌گونه از مدل‌های ترکیبی بارها برای پیش‌بینی سری‌های زمانی به‌کار رفته‌اند. خاشعی و بیجاری[۱۴] یک مدل که ترکیبی از مدل‌های شبکه‌های عصبی مصنوعی و خودرگرسیون میانگین متحرک انباشته است، برای پیش‌بینی سری‌های زمانی ارائه کرده‌اند. چن و وانگ[۱۵] به‌منظور پیش‌بینی سری‌های زمانی فصلی، یک مدل که ترکیبی از مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی و مدل ماشین‌های بردار پشتیبان است، پیشنهاد کرده‌اند. تی‌سنگ و تی‌ژنگ[۱۶] مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی - فازی را برای پیش‌بینی ارائه کرده‌اند که مزایای هر دو مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی و فازی را ترکیب کرده‌اند. همچنین تلاش‌های بسیاری توسط محققان به‌منظور بهبود فرایند پیش‌بینی تقاضای الکتریسیته صورت گرفته است. هووانگ و همکارانش،[۱۷] یک مدل پیش‌بینی به نام گری - مارکوف را برای پیش‌بینی تقاضای برق در چین ارائه کرده‌اند. پائو،[۱۸] یک مدل ترکیبی غیرخطی برای پیش‌بینی مصرف انرژی در تایوان پیشنهاد داده است.

خان و آبرهام،[۱۹] یک مدل که ترکیبی از شبکه‌ی هوش مصنوعی و منطق فازی است، برای پیش‌بینی بار در جمهوری چک پیشنهاد کرده‌اند. آن‌ها دریافتند که مدل ترکیبی فازی - شبکه‌ی عصبی بهترین گزینه برای تجزیه و تحلیل بار در جمهوری چک هستند. لیائو و تیسائو،[۲۰] یک مدل که ترکیبی از روش فازی - شبکه‌ی عصبی و شبیه‌سازی تبریدی است، به‌منظور کاهش خطای بار الکتریکی ارائه کرده‌اند. مدل پیشنهادی آن‌ها با استفاده از داده‌هایی با دوره‌های یک‌ساله، یک‌ماهه و یک‌روزه آزمایش شد. نتیجه نشان‌دهنده‌ی دقت بالاتر مدل پیشنهادی بوده است. مطالعه‌ی دیگری توسط کودوگیانیس و آناگنوستاکیس،[۲۱] به‌منظور پیش‌بینی بار الکتریکی کوتاه‌مدت به روش مبتنی‌بر شبکه عصبی برای سیستم برق جزیره‌ی کرت در یونان صورت گرفته است. نتایج بیان‌گر آن بوده است که مدل‌های پیش‌بینی بار توسعه یافته نسبت به سایر روش‌های معمول پیش‌بینی از دقت بالاتری برخوردار هستند. شرف و تی‌جینگ،[۲۲] یک روش پیش‌بینی تقاضای الکتریسیته‌ی کوتاه‌مدت فازی - عصبی را، که بر اساس شبکه‌های عصبی و منطق فازی عمل می‌نماید، پیشنهاد کرده‌اند. سان و کیم[۲۳] یک روش ترکیبی برای پیش‌بینی تقاضای الکتریسیته‌ی کوتاه‌مدت بر اساس مدل‌های رگرسیون بردار پشتیبان، فازی و الگوریتم بهینه‌سازی ازدحام ذرات ارائه کرده‌اند. در مدل پیشنهادی آنها جنبه‌ی فصلی بودن داده‌های الکتریسیته بررسی نشده است. ژانگ و همکاران[۲۴] یک روش ترکیبی بر اساس مدل‌های ماشین بردار پشتیبان، تجزیه و تحلیل طیف منحصر به فرد و الگوریتم جستجوی فاخته برای پیش‌بینی بار الکتریکی ارائه کرده‌اند. روش پیشنهادی آنها ابهام و عدم‌قطعیت موجود و همچنین الگوهای فصلی موجود در داده‌ها را نادیده گرفته است. رازا و همکاران[۲۵] از یک روش ترکیبی مبتنی‌بر شبکه‌های عصبی مصنوعی با لایه‌های پیش‌رو برای شبکه‌های انرژی و پیش‌بینی بار الکتریکی کوتاه‌مدت استفاده کرده‌اند. روش پیشنهادی آنها الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها را در نظر نگرفته است. سادهیر و ساسیلاتها[۲۶] یک روش که ترکیبی از مدل‌های انتقال موجک، هالت - وینترز و نزدیک‌ترین همسایگی وزن‌دهی‌شده است، به‌منظور پیش‌بینی بار الکتریکی ارائه کرده‌اند. کاووسی‌فرد و همکاران[۲۷] از یک روش ترکیبی بر پایه‌ی مدل‌های رگرسیون بردار پشتیبان و الگوریتم کرم شب‌تاب اصلاح شده برای پیش‌بینی کوتاه‌مدت بار الکتریکی استفاده کرده‌اند.

همان‌گونه که گفته شد، هریک از مدل‌های ذکر شده فقط جنبه‌ی خاصی از داده‌های بار الکتریسیته را مدل‌سازی می‌کنند و به‌طور هم‌زمان تمام الگوها و ساختارها را بررسی نمی‌کنند. برخی از پژوهش‌ها جنبه‌ی غیرخطی و فصلی بودن داده‌های بار الکتریسیته را در نظر می‌گیرند و مدل‌سازی را انجام می‌دهند و ابهام و عدم‌قطعیت موجود در داده‌ها را نادیده می‌گیرند. برخی فصلی بودن داده‌ها را در مدل‌سازی اعمال نمی‌کنند. سایر روش‌های بررسی شده در سال‌های اخیر و معایب آنها نسبت به مدل پیشنهادی در جدول ۱ خلاصه شده است.

بر این اساس، در این مقاله به‌منظور حصول نتایج دقیق‌تر در مدل‌سازی‌های الگوهای فصلی، یک مدل ترکیبی برای پیش‌بینی کوتاه‌مدت تقاضای الکتریسیته پیشنهاد شده است. مدل پیشنهادی از ترکیب مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی، پرسپترون‌های چندلایه و مدل فازی ایجاد شده است. در این مدل، پرسپترون‌های چندلایه و مدل‌های فازی به‌منظور غلبه بر محدودیت خطی و محدودیت عدم قطعیت در مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی به‌منظور حصول نتایج دقیق‌تر مورد استفاده قرار گرفته‌اند. در این روش هر یک از مدل‌های تشکیل‌دهنده‌ی مدل پیشنهادی، جنبه‌یی خاص از داده‌ها را در نظر می‌گیرند و پیش‌بینی را انجام می‌دهد. بدین‌منظور مدل اول (خودرگرسیون میانگین متحرک انباشته‌ی فصلی) به‌منظور بررسی الگوهای فصلی و همچنین ساختارهای خطی موجود در داده‌ها به‌کارگرفته شده است. مدل دوم (پرسپترون‌های چندلایه)، ساختارهای غیرخطی موجود در نتایج پیش‌بینی با مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی را مدل‌سازی می‌کند و در آخر مدل فازی به‌منظور رفع ابهام و عدم‌قطعیت به مدل افزوده می‌شود. نتایج حاصل از به‌کارگیری مدل ترکیبی پیشنهادی در بازارهای بین‌المللی الکتریسیته بیان‌گر برتری روش پیشنهادی نسبت به مدل‌های کلاسیک همچون خودرگرسیون میانگین متحرک انباشته‌ی فصلی (SARIMA) و رگرسیون خطی چندگانه‌ی کلاسیک (MLR) و همچنین شبکه‌های عصبی مصنوعی (ANN) همچون پرسپترون‌های چندلایه (MLP) و ماشین‌های بردار پشتیبان (SVM) بوده است.

سایر قسمت‌های این مقاله بدین صورت است: در قسمت بعد مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی به‌طور خلاصه ارائه شده است. در قسمت سوم، فرمول‌سازی مدل ترکیبی پیشنهادی ارائه شده است. در قسمت چهارم، مدل پیشنهادی برای پیش‌بینی کوتاه‌مدت تقاضای الکتریسیته مورد استفاده قرارگرفته و نتایج با سایر روش‌ها مقایسه شده است. در نهایت نیز نتایج مدل‌ها بررسی و تجزیه و تحلیل شده‌اند.

۲. مدل خودرگرسیون میانگین متحرک انباشته فصلی

بیش از نیم قرن است که مدل‌های خودرگرسیون میانگین متحرک انباشته به‌منظور تجزیه و تحلیل سری‌های زمانی، توسط بوکس - جنکینز[۳۳] پیشنهاد شده‌اند. هرچند این‌گونه از مدل‌ها در پیش‌بینی سری‌های زمانی غیرفصلی بسیار استفاده می‌شوند، اما گاهی اوقات ممکن است یک جزء فصلی نیز در سری زمانی مورد مطالعه وجود داشته باشد. بنابراین، بوکس - جنکینز یک مدل گسترش یافته از این‌گونه از مدل‌ها را که توانایی پیش‌بینی داده‌های فصلی را نیز داشته باشد، ارائه کردند.[۳۳] این مدل اصطلاحاً مدل خودرگرسیون میانگین متحرک انباشته فصلی نامیده می‌شود.

www.SID.ir

Archive of SID

جدول ۱. ادبیات موضوع بررسی شده در سال‌های اخیر.

نویسنده/نویسندگان	سال	حوزه	مدل‌ها	معایب روش نسبت به روش پیشنهادی
ژانگ و همکاران[۲۴]	۲۰۱۷	پیش‌بینی بار الکتریکی	ماشین بردار پشتیبان، تجزیه و تحلیل طیف منحصربه‌فرد، و الگوریتم جستجوی فاخته	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
رازا و همکاران[۲۵]	۲۰۱۷	پیش‌بینی ساعتی تقاضای الکتریسیته	شبکه‌های عصبی مصنوعی لایه‌های پیشرو و الگوریتم بهینه‌سازی ازدحام ذرات	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
کائو و وو[۲۸]	۲۰۱۶	پیش‌بینی مصرف الکتریسیته	رگرسیون بردار پشتیبان و الگوریتم بهینه‌سازی مگس میوه	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
سادهیر و ساسیلاتها [۲۶]	۲۰۱۵	پیش‌بینی بار الکتریکی کوتاه‌مدت	انتقال موجک، هموارسازی نمایی سه‌تایی و مدل نزدیک‌ترین همسایگی وزن‌دار	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
ژیائو و همکاران[۲۹]	۲۰۱۵	پیش‌بینی بار الکتریکی	شبکه‌های عصبی پسخور، شبکه‌های عصبی المان و الگوریتم ژنتیک و شبکه عصبی رگرسیون کلی	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
کاووسی‌فرد و همکاران [۲۷]	۲۰۱۴	پیش‌بینی بار الکتریکی کوتاه‌مدت	رگرسیون بردار پشتیبان و الگوریتم کرم شب‌تاب اصلاح شده	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها و همچنین ساختارهای خطی موجود در داده‌ها
ژانگ و همکاران[۳۰]	۲۰۱۲	پیش‌بینی قیمت الکتریسیته	انتقال موجک، خودرگرسیون میانگین متحرک انباشته، ماشین بردار پشتیبان حداقل مربعات و الگوریتم بهینه‌سازی ازدحام ذرات	عدم بررسی الگوهای فصلی، ابهام و عدم قطعیت موجود در داده‌ها
چه و همکاران[۳۱]	۲۰۱۲	پیش‌بینی بار الکتریکی	نقشه‌های خودسازمان‌دهنده، رگرسیون بردار پشتیبان و مدل فازی	عدم بررسی الگوهای فصلی و همچنین ساختارهای خطی موجود در داده‌ها
نی و همکاران[۳۲]	۲۰۱۲	پیش‌بینی بار الکتریکی کوتاه‌مدت	خودرگرسیون میانگین متحرک انباشته و ماشین بردار پشتیبان	عدم بررسی الگوهای فصلی، ابهام و عدم‌قطعیت موجود در داده‌ها

در حالت کلی، در سری‌های زمانی فصلی دو نوع تناوب وجود دارد. نوع اول بین مشاهدات متوالی و نوع دوم بین مشاهدات نظیر به نظیری است که متعلق به فصل‌های متوالی‌اند. اساساً مدل خودرگرسیون میانگین متحرک انباشته برای توصیف رابطه بین مقادیر مشاهدات متوالی استفاده می‌شود. در حالی‌که مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی به‌منظور مدل‌سازی رابطه بین مشاهدات نظیر به نظیر فصل‌های متوالی استفاده می‌شود. اصطلاحاً یک سری زمانی $\{y_t|t=1,2,\dots,k\}$ از مدل سری زمانی بوکس - جنکینز توسط فرایند $SARIMA(p,d,q)(PD,Q)s$ با میانگین μ تولید می‌شود، اگر:

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D(y_t-\mu)=\Theta_Q(B^s)\theta_q(B)a_t \qquad (1)$$

به قسمی که y_t و a_t مقادیر مشاهده شده و خطای تصادفی در دوره‌ی زمانی $(t=1,2,\dots,k)$ هستند. $\phi_p(B)=1-\sum_{i=1}^{p}\phi_iB^i$ و $\theta_q(B)=1-\sum_{i=1}^{q}\theta_iB^i$ به ترتیب عملگرهای خودرگرسیون غیرفصلی (AR) و میانگین متحرک غیرفصلی (MA) هستند. $\Phi_P(B^s)=1-\sum_{i=1}^{P}\Phi_{is}B^{is}$ و $\Theta_Q(B^s)=1-\sum_{i=1}^{Q}\Theta_{is}B^{is}$ عملگرهای خودرگرسیون و میانگین متحرک فصلی هستند. عملگر B عملگر انتقال پسرو است. $\nabla^d=(1-B)^d$ عملگر تفاضل‌گیری غیرفصلی و $\nabla_s^D=(1-B^s)^D$ عملگر تفاضل‌گیری فصلی است. s تعداد دوره‌ها، p مرتبه‌ی خودرگرسیون غیرفصلی، P مرتبه‌ی خودرگرسیون فصلی، q مرتبه‌ی میانگین متحرک غیرفصلی و Q مرتبه‌ی میانگین متحرک فصلی است. a_t متغیری تصادفی با توزیع نرمال با میانگین صفر و واریانس σ^2 و مستقل از مشاهدات است. در حالت کلی فرایند فرمول‌بندی مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی شامل چهار مرحله‌ی تکراری مطابق زیر است:

www.SID.ir

Archive of SID

۱. تعیین ساختار $SARIMA(p,d,q)(P,D,Q)$

۲. تخمین پارامترهای مجهول.

۳. آزمون نیکویی برازش بر روی باقی‌مانده‌ها.

۴. پیش‌بینی نتایج بر اساس داده‌های معلوم.

مرحله‌ی اول، تعیین صحیح مرتبه‌ی مدل خودرگرسیون میانگین متحرک انباشته است. یک راه حل اساسی به‌منظور شناسایی مدل، استفاده از توابع خودهمبستگی (ACF) و خودهمبستگی جزئی ($PACF$) است. در این مرحله با رسم نمودارهای خودهمبستگی و خودهمبستگی جزئی، ایستایی در میانگین ارزیابی می‌شود و مراتب p، P، q و Q را می‌توان تعیین کرد. البته روش‌های دیگری نیز در پیشینه‌ی موضوع برای شناسایی مرتبه‌ی مدل پیشنهاد شده است. معیار اطلاعاتی آکائیک یکی از شناخته شده‌ترین این روش‌هاست. به‌علاوه، در سال‌های اخیر روش‌های مختلفی بر اساس روش‌های هوشمند مثل شبکه‌های عصبی، الگوریتم ژنتیک،[۳۴] یا سیستم فازی[۳۵] به‌منظور شناسایی دقیق‌تر مرتبه‌ی مدل‌های خودرگرسیون میانگین متحرک انباشته پیشنهاد شده است. پس از این‌که یک مدل آزمایشی تشخیص داده شد، تخمین پارامترهای مدل کار ساده‌یی است و از طریق کمینه‌سازی جمله‌ی خطا به دست می‌آیند. در این فرایند، مدلی که دارای مجموع مربعات خطا یا میانگین مربعات خطای کوچکتری باشد، انتخاب می‌شود. آخرین گام مدل‌سازی، تشخیص مناسب بودن مدل است و اساساً به‌منظور بررسی این‌که فرضیات مدل در مورد خطاها صدق می‌کند یا نه استفاده می‌شود. اگر مدل انتخابی مناسب نباشد یک مدل آزمایشی جدید باید تشخیص داده شود و موارد فوق دوباره تکرار شوند.

۳. مدل ترکیبی پیشنهادی

پیش‌بینی سری‌های زمانی فصلی یکی از مهم‌ترین زمینه‌های تحقیقاتی در سال‌های اخیر بوده است. مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی یکی از شناخته شده‌ترین و پرکاربردترین روش‌های آماری خطی برای پیش‌بینی سری‌های زمانی فصلی‌اند. اگرچه مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی، هم در مسائل علمی و هم در مسائل عملی موفقیت‌های چشمگیری در طول سه دهه‌ی اخیر کسب کرده‌اند، این‌گونه از مدل‌ها دو محدودیت اساسی دارد که از محبوبیت آنها برای پیش‌بینی سری‌های زمانی فصلی کاسته است؛ در مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی، مقدار آینده‌ی متغیر وابسته به‌صورت تابعی خطی از چند مشاهده‌ی گذشته و خطای تصادفی در نظر گرفته می‌شود. به‌دلیل این محدودیت، اگر سازوکار تولید داده‌ها غیرخطی باشد، ممکن است تقریب مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی کاملاً نامناسب باشد. در حالی‌که بر اساس پیشینه‌ی موضوع، سیستم‌های دنیای واقع اغلب غیرخطی‌اند.[۳۶] بر این اساس، این فرض که یک سری زمانی توسط یک فرایند خطی تولید شده است، ممکن است کاملاً غیرمنطقی به نظر برسد.

به همین علت، در این مقاله از شبکه‌ی عصبی مصنوعی به‌منظور غلبه‌بر محدودیت خطی و از مدل فازی به‌منظور غلبه بر ابهام موجود در داده‌ها در مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی استفاده شده است. بر اساس پیشینه‌ی موضوع، شبکه‌های عصبی مصنوعی چارچوب‌های محاسباتی غیرخطی انعطاف‌پذیری هستند که می‌توانند در بسیاری از مسائل پیش‌بینی استفاده شوند. بنابراین با ترکیب شبکه‌های عصبی مصنوعی، مدل فازی و مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی، می‌توان ساختارهای خطی و غیرخطی موجود در داده‌ها را به‌صورت هم‌زمان مدل‌سازی کرد. در مدل پیشنهادی ساختارهای خطی و غیرخطی موجود در داده‌ها به ترتیب توسط مدل‌های خودرگرسیون میانگین متحرک انباشته‌ی فصلی و پرسپترون‌های چندلایه مدل‌سازی شده‌اند.

مطابق با فرایند مدل پیشنهادی، در ابتدا ساختار خطی (L_t) موجود در داده‌های سری زمانی مورد مطالعه $\{y_t\}$ توسط یک مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی برازش داده می‌شود. نتایج این مرحله، برآوردهای مقادیر واقعی سری‌زمانی مورد مطالعه ($\hat{L}_t$) و پارامترهای خطی مدل خواهند بود. محاسبات مربوط به این مرحله‌در روابط زیر به‌صورت خلاصه بیان شده است:

$$\begin{aligned} z_t = & \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{j=1}^{P} + \Phi_{js} z_{t-js} + \\ & \sum_{i=1}^{p}\sum_{j=1}^{P} \phi_i \Phi_{js} z_{t-i-js} + a_t + \\ & + \sum_{i=1}^{q} \theta_i a_{t-i} + \cdots + \sum_{j=1}^{Q} \Theta_{js} \cdot a_{t-js} + \\ & + \sum_{i=1}^{q}\sum_{j=1}^{Q} \theta_i \Theta_{js} a_{t-i-js}, \end{aligned} \quad (۲)$$

که در آن، $\phi_p(p = 1, 2, \ldots, p)$ و $\theta_q(q = 1, 2, \ldots, q)$ به ترتیب پارامترهای غیرفصلی و $\Phi_p(P = 1, 2, \ldots, P)$ و $\Theta_Q(Q = 1, 2, \ldots, Q)$ پارامترهای فصلی مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی، $z_t = \nabla^d \nabla_s^D (y_t - \mu)$ و $B(\mu) = (1 - B^s)^D (1 - B)^d (y_t - \mu)$ عملگر انتقال پسرو، D و d نیز اعداد صحیحی هستند که اصطلاحاً مرتبه‌ی تفاضلی فصلی و غیرفصلی نامیده می‌شوند.

حال با در نظر گرفتن یک سری زمانی به‌صورت ترکیب خطی یک جزء خطی و یک جزء غیرخطی $y_t = N_t + L_t$ داریم و با توجه به مطالب ذکر شده می‌توان استنباط کرد که مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی نمی‌تواند الگوهای غیرخطی (N_t) موجود در داده‌ها را مدل‌سازی کند. بر این اساس، باقی‌مانده‌های حاصل از مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی، فقط شامل الگوهای فصلی غیرخطی ($e_t = y_t - \hat{L}_t$) خواهند بود. بنابراین در مرحله‌ی دوم، پرسپترون‌های چندلایه، برای مدل‌سازی باقی‌مانده‌های مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی استفاده خواهند شد. با مدل‌سازی باقی‌مانده‌های مرحله‌ی اول و استفاده از شبکه‌ی عصبی مصنوعی، روابط فصلی غیرخطی موجود در داده‌ها را نیز می‌توان مدل‌سازی کرد. حال با در نظرگرفتن M نرون ورودی، نتایج مدل پرسپترون چندلایه مطابق رابطه‌ی ۳ خواهد بود. باقی‌مانده‌ها به‌صورت زیر خواهد بود:

$$\begin{aligned} e_t = & f(e_{t-1}, \ldots, e_{t-M}) + \varepsilon_t = w_0 + \\ & + \sum_{j=1}^{N} w_j \; g(w_{0j} + \sum_{i=1}^{M} w_{i,j} e_{t-i}) + \varepsilon_t = \sum_{j=0}^{N} w_j u_{t,j} + \varepsilon_t \end{aligned} \quad (۳)$$

که f تابعی غیرخطی است که توسط شبکه‌ی عصبی تعیین شده است. $u_{t,j} = g(w_{0j} + \sum_{i=1}^{M} w_{i,j} e_{t-i})$ و ε_t خطای تصادفی است. البته باید توجه داشت که اگر مدل f مناسب نباشد، عبارت خطا لزوماً تصادفی نخواهد بود. بنابراین، شناسایی صحیح مدل، گامی بسیار اساسی در فرایند طراحی مدل پیشنهادی خواهد بود. در انتها نیز با استفاده از مقادیر پیش‌بینی شده با استفاده از رابطه‌ی ۳، پیش‌بینی ترکیبی

www.SID.ir

Archive of SID

$$\mu_{\tilde{z}_t}(z_t)=\begin{cases} I-\dfrac{\left|z_t-\left[\sum_{i=0}^{p}\sum_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js}+\sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1s}+\sum_{i=l_2}^{l_2+N}\alpha_i\cdot u_{t,i_2}\right]\right|}{\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot|z_{t-i-js}|+\sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}c_{ij}\cdot|a_{t-i_1-j_1s}|+\sum_{i=l_2}^{l_2+N}c_i\cdot|u_{t,i_2}|}, & z_t\neq 0, a_t\neq 0, u_{t,i}\neq 0,\\ 0 & Otherwise \end{cases} \quad (8)$$

به‌صورت زیر به دست خواهد آمد:

$$\begin{aligned}\hat{z}_t=\hat{L}_t+\hat{N}_t=&\Big(\sum_{i=1}^{p}\phi_i z_{t-i}+\sum_{j=1}^{P}\Phi_{js}\cdot z_{t-js}\\&+\sum_{i=1}^{p}\sum_{j=1}^{P}\phi_i\Phi_{js}z_{t-i-js}+\sum_{i=1}^{q}\theta_i\cdot a_{t-i}\\&+\sum_{j=1}^{Q}\Theta_{js}\cdot a_{t-js}+\sum_{i=1}^{q}\sum_{j=1}^{Q}\theta_i\Theta_{js}z_{t-i-js}\Big)\\&+\Big(w_0+\sum_{j=1}^{N}w_j\cdot u_{t,j}\Big).\end{aligned} \quad (4)$$

در قسمت بعد، پارامترهای دو مدل ذکر شده، یعنی $\varphi_p(p=1,2,\dots,p)$، $\theta_q(q=1,2,\dots,q)$، $\phi_p(P=1,2,\dots,P)$، $\Theta_Q(Q=1,2,\dots,Q)$ و $w_j(j=0,1,2,\dots,N)$ را به صورت اعداد فازی در نظر خواهیم گرفت. به‌عبارت دیگر، این پارامترها به صورت: $\tilde{\varphi}_p(p=1,2,\dots,p)$، $\tilde{\theta}_q(q=1,2,\dots,q)$، $\tilde{\phi}_p(P=1,2,\dots,P)$، $\tilde{\Theta}_Q(Q=1,2,\dots,Q)$ و $\tilde{w}_j(j=0,1,2,\dots,N)$ تبدیل خواهند شد. سپس رگرسیون فازی برای محاسبه‌ی پارامترهای مدل فازی مورد استفاده قرار خواهد گرفت. روش پیشنهادی همچنین روش ایشیبوچی و تاناکا[۳۷] را برای شرایطی که فواصل پیش‌بینی شده بسیار عریض می‌شوند، به‌کار برده است. بر این اساس، مدل ترکیبی پیشنهادی به‌صورت رابطه‌ی ۵ تبدیل خواهد شد:

$$\begin{aligned}\tilde{z}_t=&\sum_{i=1}^{p}\tilde{\phi}_i z_{t-i}+\sum_{j=1}^{P}\tilde{\Phi}_{js}\cdot z_{t-js}+\\&+\sum_{i=1}^{p}\sum_{j=1}^{P}\tilde{\phi}_i\tilde{\Phi}_{js}z_{t-i-js}+\dots\\&\cdots+\sum_{i=1}^{q}\tilde{\theta}_i\cdot a_{t-i}+\sum_{j=1}^{Q}\tilde{\Theta}_{js}\cdot a_{t-js}+\\&\sum_{i=1}^{q}\sum_{j=1}^{Q}\tilde{\theta}_i\tilde{\Theta}_{js}z_{t-i-js}+\sum_{j=0}^{N}\tilde{w}_j\cdot u_{t,j}\end{aligned} \quad (5)$$

حال رابطه‌ی ۵ را می‌توان به صورت زیر بازنویسی کرد:

$$\begin{aligned}\tilde{z}_t=&\sum_{i=0}^{p}\sum_{j=0}^{P}\tilde{\beta}_{ij}\cdot z_{t-i-js}\\&+\sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1}\tilde{\beta}_{ij}\cdot a_{t-i_1}-j_{1s}+\sum_{i=l_2}^{l_2+N}\tilde{\beta}_i\cdot u_{t,i_2}\end{aligned} \quad (6)$$

که $\tilde{\beta}_{0,j}=\tilde{\Phi}_{js}$، $\tilde{\beta}_{i,0}=\tilde{\phi}_i$، $\tilde{\beta}_{0/0}=0$ برای: $(i=0,1,2,\dots,p)$ و $(j=0,1,2,\dots,P)$. $\tilde{\beta}_{l_1,l_1}=\tilde{\theta}_1$، $\tilde{\beta}_{i_1,l_1}=0$، $\tilde{\beta}_{il_1,j}=\tilde{\theta}_i\tilde{\Theta}_{js}$، $\tilde{\beta}_{0,jt}=\tilde{\Theta}_{js}$، برای $(i_1=l_1,l_1+1,\dots,l_1+q)$، $(j_1=l_1,l_1+1,\dots,l_1+Q)$ و $l_1=P\times p$، $i_1=i-l_1$. برای $\tilde{\beta}_i=\tilde{w}_i$، $(i_2=l_2,l_2+1,\dots,l_2+N)$، $l_2=(P\times p)+(q\times Q)$ و $i_2=i-l_2$.

سپس پارامترهای فازی به‌صورت اعداد فازی مثلثی مطابق زیر در نظرگرفته می‌شوند:

$$\mu_{\tilde{\beta}_i}=(\beta_i)=\begin{cases} I-\dfrac{|\alpha_i-\beta_i|}{c_i} & \alpha_i-c_i\le\beta_i\le\alpha_i+c_i,\\ 0 & Otherwise, \end{cases} \quad (7)$$

به قسمی که $\mu_{\hat{\beta}}(\beta_i)$ تابع عضویت مجموعه‌ی فازی است که توسط پارامترهای β_i بیان می‌شود، α_i مرکز عدد فازی و c_i گسترش عدد فازی است. حال با استفاده از پارامترهای فازی β_i به صورت اعداد فازی مثلثی و استفاده از اصل تعمیم، تابع عضویت z_t در رابطه‌ی ۶ به‌صورت رابطه‌ی ۸ تبدیل خواهد شد که z_t بیان‌گر tامین مشاهده و سطح h مقدار آستانه‌یی است که باید توسط تمام مشاهدات $(z_1,z_2,\dots,z_k)$ برآورده شود.

$$\mu_{\tilde{z}}(z_t)\ge h \qquad for\ \ t=1,2,\dots,k \quad (9)$$

که اندیس t به تعداد داده‌ی غیرفازی استفاده شده برای ساخت مدل اشاره دارد. به عبارت دیگر، s بیان‌گر کل ابهام مترتب به پارامترهای فازی مدل پیشنهادی است که توسط رابطه‌ی ۱۰ محاسبه می‌شود:

$$\begin{aligned}S=&\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot|z_{t-i-js}|\\&+\sum_{t=1}^{k}\sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}c_{ij}\cdot|a_{t-i_1}-j_{1s}|\\&+\sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N}c_i\cdot|u_{t,i_2}|,\end{aligned} \quad (10)$$

به قسمی که φ_i ضریب خودرگرسیون غیرفصلی وقفه‌ی زمانی iام و ϕ_{is} ضریب خودرگرسیون فصلی وقفه‌ی زمانی isام است. همچنین θ_i ضریب میانگین متحرک غیرفصلی وقفه‌ی زمانی iام و Θ_{is} ضریب میانگین متحرک فصلی وقفه‌ی زمانی isام است. w_i وزن بین نرون خروجی و iامین نرون پنهان است. مدل پیشنهادی را می‌توان در حالت کلی مطابق رابطه‌ی ۱۱ فرمول‌بندی کرد. مراحل کلی روش ترکیبی پیشنهادی را می‌توان در مراحل زیر خلاصه کرد:

مرحله‌ی اول (فاز اول): مدل‌سازی خطی: یک مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی با استفاده از اطلاعات موجود در داده‌های مورد مطالعه برازش می‌شود. نتایج حاصل از این فاز، جواب بهینه‌ی پارامترهای $\alpha^*=(\alpha_1^*,\alpha_2^*,\dots,\alpha_{l_2}^*)$ و باقی‌مانده‌های α_t هستند که به عنوان مجموعه‌ی داده ورودی در مرحله‌ی بعد استفاده می‌شوند.

مرحله‌ی اول (فاز دوم): مدل‌سازی غیرخطی: یک شبکه‌ی عصبی پرسپترون چندلایه با استفاده از باقی‌مانده‌های مربوط به مدل خودرگرسیون میانگین متحرک انباشته‌ی فصلی آموزش داده می‌شود. نتایج حاصل از این فاز، جواب بهینه‌ی پارامترهای $\alpha^*=(\alpha_{l_2+1}^*,\alpha_{l_2+2}^*,\dots,\alpha_{l_2+N}^*)$ و مقادیر خروجی نرون‌های پنهان هستند که به عنوان مجموعه داده‌های ورودی در مرحله‌ی بعد استفاده می‌شوند.

مرحله‌ی اول (فاز سوم): ترکیب نتایج: در این فاز، نتایج فاز اول و دوم که به مدل‌سازی

www.SID.ir

Archive of SID

$$Minimize\, S = \sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|z_{t-i-js}| + \sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|a_{t-i_1-j_1 s}| + \sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N} c_i\cdot|u_{t,i_2}|$$

$$subject.to \begin{cases} \sum_{i=0}^{p}\sum_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js} + \sum_{i-l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1 s} + \sum_{i=1_2}^{1_2+N}\alpha_i\cdot u_{t,i_2} + (1-h)\cdot(\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|z_{t-i-js}| + \dots \\ \quad \dots + \sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|a_{t-i_1-j_1 s}| + \sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N} c_i\cdot|u_{t,i_2}|) \geq z_t \qquad t=1,2,\dots,k, \\ \sum_{i=0}^{p}\sum_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js} + \sum_{i-l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1 s} + \sum_{i=l_2}^{l_2+N}\alpha_i\cdot u_{t,i_2} - (1-h)\cdot(\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|z_{t-i-js}| = + \dots \\ \quad \dots + \sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P} c_{ij}\cdot|a_{t-i_1-j_1 s}| + \sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N} c_i\cdot|u_{t,i_2}|) \geq z_t \qquad t=1,2,\dots,k \\ c_i \geq 0 \qquad for\ i=1,2,\dots,l_2+N. \end{cases}$$

(۱۱)

تمام روابط تشکیل‌دهنده در داده‌ها منجر می‌شود، با یکدیگر ترکیب خواهند شد. نتایج این مرحله برآورد نهایی داده‌های واقعی مورد مطالعه خواهد بود.

مرحله‌ی دوم: تعیین حداقل ابهام: در این مرحله، با استفاده از رابطه‌ی ۱۱ و $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_{l_2+N}^*)$، حداقل ابهام مترتب به پارامترهای مدل محاسبه می‌شود. در نتیجه مدل پیشنهادی به‌صورت زیر به دست خواهد آمد:

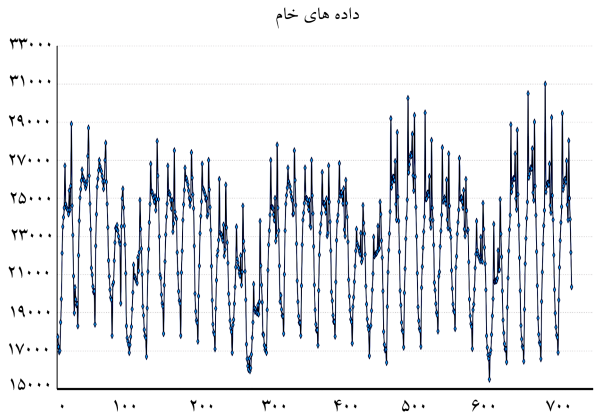
$$\tilde{z}_t = \sum_{i=0}^{p}\sum_{j=0}^{P}\langle\alpha_{ij}, c_{ij}\rangle\cdot z_{t-i-js} + \sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\langle\alpha_{ij}, c_{ij}\rangle\cdot a_{t-i_1-j_1 s} + \sum_{i=l_2}^{l_2+N}\langle\alpha_i, c_i\rangle_i\cdot u_{t,i_2} \qquad (۱۲)$$

که z_t برابر با $\nabla^d\nabla_s^D(y_t-\mu) = (1-B^s)^D(1-B)^d(y_t-\mu)$ است، به طوری که B عملگر انتقال پسرو، D و d اعدادی صحیح هستند که اصطلاحاً مرتبه‌ی تفاضلی فصلی و غیرفصلی نامیده می‌شوند، μ میانگین مشاهدات، α_i مرکز اعداد فازی و c_i گسترش مترتب به اعداد فازی است.

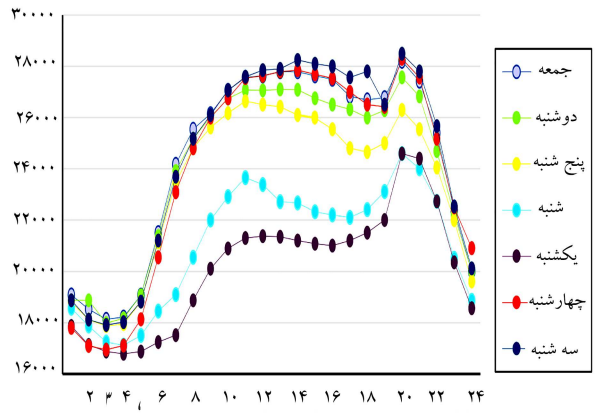
مرحله‌ی سوم: حذف داده‌های پرت: در حالت کلی، زمانی که مجموعه داده‌های مورد استفاده در مدل پیشنهادی شامل داده‌های پرت یا تفاوت‌های معنی‌دار باشد، آنگاه کران‌های بالا یا پایین حاصل از مدل پیشنهادی، بنابر نظرات ایشیبوچی و تاناکا[۳۷] حذف خواهند شد. به عبارت دیگر، به‌منظور ساختن مدلی با تمام این شرایط اگر مجموعه داده‌ها شامل تفاوت‌های مشخص یا موارد خارج از محدوده باشند، آنگاه c_jها بسیار گسترده خواهند شد. ایشیبوچی و تاناکا پیشنهاد داده‌اند که داده‌های اطراف کران بالا و پایین مدل حذف شوند و سپس مدل رگرسیون فازی مجدداً فرمول‌بندی شود. پس در روش پیشنهادی نیز در صورت وجود داده‌های پرت یا نمونه‌های خارج از محدوده، از روش ایشیبوچی و تاناکا استفاده خواهد شد.

۴. بررسی عملکرد مدل پیشنهادی در پیش‌بینی تقاضای الکتریسیته

در این قسمت، عملکرد مدل پیشنهادی در مقابل مدل‌های تشکیل‌دهنده‌ی آن و سایر مدل‌های موجود همچون مدل نی و همکاران[۳۲] در پیش‌بینی کوتاه‌مدت تقاضای الکتریسیته بررسی می‌شود. به‌منظور مقایسه‌ی مدل‌های مختلف با یکدیگر، داده‌های مربوط به تقاضای الکتریسیته استفاده شده‌اند. این داده‌ها از جمله داده‌های محک تقاضای الکتریسیته هستند که به‌صورت ساعتی از بازار کالیفرنیا گرفته شده‌اند.[۳۸] این داده‌ها مربوط به بازه‌ی زمانی یکم تا سی‌ام آوریل سال ۱۹۹۸ هستند. نمودار



شکل ۱. داده‌ها مربوط به تقاضای الکتریسیته (یکم تا سی‌ام آوریل سال ۱۹۹۸).



شکل ۲. میانگین چرخه‌های هفتگی تقاضای الکتریسیته (یکم تا سی‌ام آوریل سال ۱۹۹۸).

سری زمانی داده‌ها در شکل ۱ نمایش داده شده است. از ویژگی‌های مهم این سری زمانی وجود دو چرخه‌ی فصلی: هفتگی و روزانه است. چرخه‌ی فصلی روزانه بیانگر وجود رابطه‌ی معنی‌دار بین تقاضا در هر ساعت خاص نسبت به روز گذشته است. همچنین در چرخه‌ی فصلی هفتگی این روند به روزهای مشابه در هفته بستگی دارد. میانگین چرخه‌های مربوط به تقاضای الکتریسیته‌ی روزانه، برای هر روز از هفته، در شکل ۲ نشان داده شده است. در این مقاله، مطابق با مقالات قبلی، داده‌های مربوط به یکم تا بیست‌ونهم آوریل به‌منظور آموزش و روز آخر ماه به‌منظور

www.SID.ir

Archive of SID

جدول ۲. میانگین و خطای استاندارد مدل پیشنهادی و سایر مدل‌ها.

مدل	داده‌های آموزش		فاصله‌ی اطمینان ۹۰ درصدی میانگین	
	میانگین	خطای استاندارد	کران پایین	کران بالا
رگرسیون چندگانه	۶۰۰٫۵	۲۲٫۱	۵۵۷٫۲	۶۴۳٫۸
خودرگرسیون میانگین متحرک انباشته‌ی فصلی	۶۲۵٫۸	۲۰٫۹	۵۸۴٫۸	۶۶۶٫۸
ماشین‌های بردار پشتیبان	۶۴۱٫۶	۲۶٫۳	۵۹۰٫۵	۶۹۳٫۱
پرسپترون‌های چندلایه	۷۷۸٫۴	۲۲٫۷	۷۳۳٫۹	۸۲۲٫۹
مدل ترکیبی نی و همکاران[۳۲]	۵۰۲٫۷۲	۱۹٫۴	۴۶۴٫۷	۵۴۰٫۷
مدل ترکیبی خطی ـ غیرخطی غیرفازی	۴۹۹٫۵	۱۴٫۲	۴۷۱٫۶	۵۲۷٫۳
مدل ترکیبی پیشنهادی (خطی ـ غیرخطی فازی)	۴۲۴٫۵	۱۲٫۱	۴۰۰٫۸	۴۴۸٫۲
مدل	داده‌های آزمون		فاصله‌ی اطمینان ۹۰ درصدی میانگین	
	میانگین	خطای استاندارد	کران پایین	کران بالا
رگرسیون چندگانه	۶۹۳٫۰	۱۱۲٫۳	۴۷۲٫۹	۹۱۳٫۱
خودرگرسیون میانگین متحرک انباشته‌ی فصلی	۶۵۹٫۸	۸۱٫۵	۵۰۰٫۱	۸۱۹٫۵
ماشین‌های بردار پشتیبان	۶۹۷٫۱	۱۴۷٫۰	۴۰۹٫۰	۹۸۵٫۲
پرسپترون‌های چندلایه	۷۶۳٫۰	۱۲۵٫۲	۵۱۷٫۶	۱۰۰۸٫۴
مدل ترکیبی نی و همکاران[۳۲]	۶۳۱٫۵	۹۵٫۱	۴۴۵٫۱	۸۱۷٫۹
مدل ترکیبی خطی ـ غیرخطی غیرفازی	۴۶۰٫۱	۶۸٫۴	۳۲۶٫۰	۵۹۴٫۲
مدل ترکیبی پیشنهادی (خطی ـ غیرخطی فازی)	۳۹۱٫۱	۵۸٫۱	۲۷۷٫۲	۵۰۴٫۹

آزمون مدل‌ها استفـاده شده‌اند. در این مقاله سه معیار سنجش: میانگین قدرمطلق خطا، جذر میانگین مربعات خطا و میانگین درصدی قدرمطلق خطا مورد استفاده قرارگرفته‌اند. این معیارها مطابق روابط زیر محاسبه می‌شود:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - A_i| \qquad (۱۳)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2} \qquad (۱۴)$$

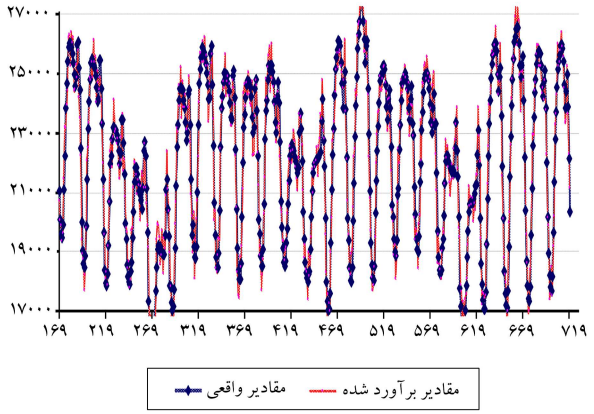
$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_i - A_i}{A_i}\right| \times 100 \qquad (۱۵)$$

که p_i و A_i به ترتیب iامین مقدار پیش‌بینی شده و واقعی و n تعداد کل داده‌هاست.

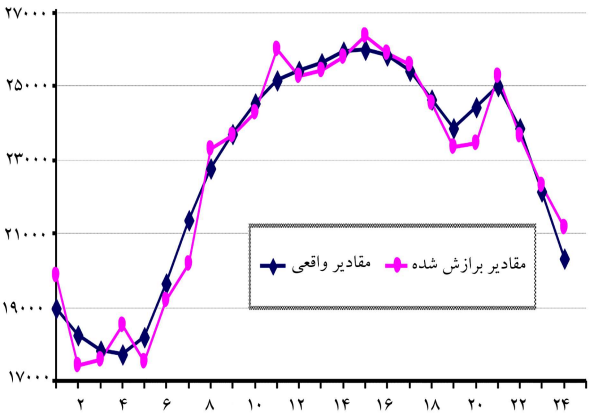
۱.۴. نتایج

در جدول ۲ نتایج مربوط به میانگین و خطای استاندارد مدل‌های مختلف نشان داده شده است. این نتایج بیانگر دقت بالاتر روش پیشنهادی در مقایسه با سایر روش‌هاست. خطای استاندارد روش پیشنهادی نیزکمتر از سایر روش‌هاست. پس نتایج دقیق‌تر مدل پیشنهادی، قابلیت اطمینان بالاتری نیز دارند. مقادیر میانگین قدرمطلق خطا، جذر میانگین مربعات خطا و میانگین درصدی قدرمطلق خطا برای مدل پیشنهادی به ترتیب در داده‌های آموزش ۵۰۲٫۷، ۵۷۳٫۴۸ و ۳٫۳٪ و داده‌های آزمون ۶۳۱٫۵، ۶۹۷٫۵ و ۳٫۰٪ است.

مقادیر واقعی و پیش‌بینی شده‌ی مربوط به مدل پیشنهادی برای کل داده‌ها و داده‌های آزمون به ترتیب در شکل ۳ و ۴ آمده است. همچنین در جدول ۳ شاخص‌های ارزیابی عملکرد مربوط به نتایج پیش‌بینی نشان داده شده است. این نتایج نیز نشان می‌دهند که مدل پیشنهادی چه در داده‌های آزمون و چه در داده‌های آموزش عملکرد بهتری نسبت به سایر مدل‌ها داشته است.



شکل ۳. مقادیر واقعی و پیش‌بینی شده‌ی تقاضای الکتریسیته کل داده‌ها (یکم تا سی‌ام آوریل سال ۱۹۹۸).



شکل ۴. مقادیر واقعی و پیش‌بینی شده‌ی تقاضای الکتریسیته برای داده‌های آزمون (روز سی‌ام آوریل سال ۱۹۹۸).

www.SID.ir

Archive of SID

جدول ۳. معیارهای عملکردی مدل پیشنهادی و سایر مدل‌ها.

مدل	داده‌های آموزش			داده‌های آزمون		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE
رگرسیون چندگانه	۶۰۰/۵	۸۳۶/۸	٪۳/۵	۶۹۳/۰	۸۷۷/۸	٪۳/۱
خودرگرسیون میانگین متحرک انباشته‌ی فصلی	۶۲۵/۸	۷۸۸/۸	٪۳/۹	۶۵۹/۸	۷۶۶/۸	٪۳/۰
ماشین‌های بردار پشتیبان	۶۴۱/۶	۹۶۵/۸	٪۳/۹	۶۹۷/۱	۹۳۲/۲	٪۳/۱
پرسپترون‌های چندلایه	۷۷۸/۴	۹۸۱/۹	٪۳/۵	۷۶۳/۰	۹۷۰/۸	٪۳/۳
مدل ترکیبی نی و همکاران[۳۲]	۵۰۲/۷	۵۷۳/۴۸	٪۳/۳	۶۳۱/۵	۶۹۷/۵	٪۳/۰
مدل ترکیبی خطی ـ غیرخطی غیرفازی	۴۹۹/۵	۵۹۶/۷	٪۳/۳	۴۶۰/۱	۵۶۴/۹	٪۳/۲
مدل ترکیبی پیشنهادی (خطی ـ غیرخطی فازی)	۴۲۴/۵	۵۰۷/۲	٪۳/۰	۳۹۱/۱	۴۸۰/۲	٪۱/۹

۵. نتیجه‌گیری

دقت پیش‌بینی‌ها یکی از مهم‌ترین عامل‌های مؤثر بر کیفیت تصمیمات اتخاذی توسط مدیران مالی و عملیاتی شرکت‌هاست. این مهم‌ترین عاملی است که باعث شده تحقیقات برای ارائه و توسعه‌ی مدل‌های دقیق‌تر پیش‌بینی در پژوهش‌ها هرگز متوقف نشود. این‌گونه از مدل‌ها براساس معیارهای متفاوتی تقسیم‌بندی می‌شوند. مثلاً این‌گونه از مدل‌ها را می‌توان به مدل‌های خطی/غیرخطی، مدل‌های قطعی/غیرقطعی، مدل‌های فصلی/غیرفصلی و ... تقسیم‌بندی کرد. مدل‌های پیش‌بینی خطی، همان‌طورکه از نامشان پیداست، فقط قادر به مدل‌سازی الگوها و ساختارهای خطی موجود در داده‌ها هستند. به‌همین ترتیب، مدل‌های پیش‌بینی قطعی، فقط توانایی مدل‌سازی الگوهای غیرفازی و قطعی موجود در داده‌ها را دارند. این موضوع در مورد مدل‌های فصلی/غیرفصلی نیز صادق است و ساختارهای خود همبسته‌ی فصلی موجود در داده‌ها در مدل‌سازی‌های غیرفصلی لحاظ نمی‌شوند. بر این اساس، هر یک از این مدل‌ها فقط می‌توانند در حوزه‌ی تعریف شده‌ی خود به کار بروند و به کارگیری آنها در سایر موارد منطقی یا کارآمد نخواهد بود. در این مقاله، یک روش ترکیبی به‌منظور پیش‌بینی هم‌زمان الگوهای خطی و غیرخطی ارائه شده است که توانایی مدل‌سازی الگوهای فصلی و غیرقطعی موجود در داده‌ها را نیز دارد. ایده‌ی روش پیشنهادی ترکیب مدل‌های خطی خودرگرسیون میانگین متحرک انباشته‌ی فصلی با مدل‌های غیرخطی پرسپترون‌های چندلایه‌ی غیرفصلی به‌منظور بهره‌گیری از مزایای منحصربه‌فرد هر یک از این مدل‌ها و همچنین استفاده از منطق فازی در مدل‌سازی عدم قطعیت داده‌هاست. بر این اساس با استفاده از مدل پیشنهادی می‌توان الگوهای خطی و غیرخطی، الگوهای فصلی و غیرفصلی و همچنین الگوهای قطعی و غیرقطعی را به‌صورت هم‌زمان مدل‌سازی کرد. به‌منظور مقایسه‌ی عملکرد مدل پیشنهادی با سایر مدل‌ها، داده‌های محک تقاضای ساعتی الکتریسیته از یکم تا سی‌ام آوریل سال ۱۹۹۸ در نظر گرفته شده است. نتایج نشان‌گر آن است که دقت پیش‌بینی‌های مدل پیشنهادی در تمام معیارهای ارزیابی از سایر مدل‌های بررسی‌شده بهتر بوده است. در حالی که خطای استاندارد مدل پیشنهادی نیز از سایر مدل‌ها کمتر بوده است. از این‌رو، روش ترکیبی پیشنهادی می‌تواند راهکاری مناسب به‌منظور جایگزینی با سایر روش‌های تکی باشد.

منابع (References)

1. Chui, F., Elkamel, A., Surit, R. and et al. "Long-term electricity demand forecasting for power system planning using economic, demographic and climatic variables", *European Journal of Industrial Engineering*, **3**, pp. 277-304 (2009).

2. Goia, A., May, C. and Fusai, G. "Functional clustering and linear regression for peak load forecasting", *International Journal of Forecasting*, **26**(4), pp. 700-711 (2010).

3. Kucukali, S. and Baris, K. "Turkeys short-term gross annual electricity demand forecast by fuzzy logic approach", *Energy Policy*, **38** (5), pp. 2438-2445 (2010).

4. Zhou, P., Ang, B.W. and Poh, K.L."A trigonometric grey prediction approach to forecasting electricity demand", *Energy*, **31**(14), pp. 2839-2847 (2006).

5. Akay, D. and Atak, M. "Grey prediction with rolling mechanism for electricity demand forecasting of Turkey", *Energy*, **32**(9), pp. 1670-1675 (2007).

6. Azadeh, A., Ghaderi, S.F. and Sohrabkhani, S."A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran", *Energy Policy*, **36**(7), pp. 2637-2644 (2008).

7. Chang, P.C., Fan, C.Y. and Lin, J.J. "Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach", *International Journal of Electrical Power Energy Systems*, **33**(1), pp. 17-27 (2011).

8. Kavaklioglu, K. "Modeling and prediction of turkeys electricity consumption using support vector regression", *Applied Energy*, **88**(1), pp. 368-375 (2011).

9. Wang, J.Z., Zhu, W.J., Zhang, W.Y. and et al."A trend fixed on firstly and seasonal adjustment model combined

www.SID.ir

Archive of SID

with the e-SVR for short-term forecasting of electricity demand", *Energy Policy*, **37**(11), pp. 4901-4909 (2009).

10. Nguyen, H.T. and Nabney, I.T. "Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models", *Energy*, **35**(9), pp. 3674-3685 (2010).
11. Bianco, V., Manca, O. and Nardini, S. "Electricity consumption forecasting in Italy using linear regression models", *Energy*, **34**(9), pp. 1413-1421 (2009).
12. Goia, A., May, C. and Goia, GF. "Functional clustering and linear regression for peak load forecasting", *International Journal of Forecasting*, **26**(4), pp. 700-711 (2010).
13. Taylor, J.W. "Triple seasonal methods for short-term electricity demand Forecasting", *European Journal of Operational Research*, **204**(1), pp. 139-152 (2010).
14. Khashei, M. and Bijari, M. "An artificial neural network (p, d, q) model for time series forecasting", *Expert Systems With Applications*, **37**, pp. 479-489 (2010).
15. Chen, K. and Wang, C. "A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in taiwan", *Expert Systems With Applications*, **32**, pp. 254-264 (2007).
16. Tseng, F.M. and Tzeng, G.H. "A fuzzy seasonal ARIMA model for forecasting", *Fuzzy Sets and Systems*, **126**, pp. 367-376 (2002).
17. Huang, M., He, Y. and Cen, H.Y. "Predictive analysis on electric-power supply and demand in China", *Renewable Energy*, **32**(7), pp. 1165-1174 (2007).
18. Pao, H.T. "Forecasting energy consumption in taiwan using hybrid nonlinear models", *Energy*, **34**(10), pp. 1138-1146 (2009).
19. Khan, M.R. and Abraham, A. "Short term load forecasting models in Czech Republic using soft computing paradigms", *International Journal of Knowledge-Based Intelligent Engineering Systems*, **7**(4), pp. 172-179 (2003).
20. Liao, G.C. and Tsao, T.P. "Application of fuzzy neural networks and artificial intelligence for load forecasting", *Electric Power Systems Research*, **70**(3), pp. 237-244 (2004).
21. Kodogiannis, V.S. and Anagnostakis, E.M. "A study of advanced learning algorithms for short-term load forecasting", *Engineering Applications of Artificial Intelligence Journal*, **44**(9), pp. 159-173 (1999).
22. Sharaf, A.M. and Tjing, T. "A novel neuro-fuzzy based self-correcting online electric load forecasting model", *Electric Power Systems Research*, **34**(2), pp. 121-125 (1995).
23. Son, H. and Kim, C. "Short-term forecasting of electricity demand for the residential sector using weather and social variables", *Resources, Conservation and Recycling*, **123**, pp.200-207 (2017).
24. Zhang, X., Wang, J. and Zhang, K. "Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm", *Electric Power Systems Research*, **146**, pp. 270-285 (2017).
25. Raza, M.Q., Nadarajah, M., Hung, D.Q. and et al. "An intelligent hybrid short term load forecast model for seasonal prediction of smart power grid", Sustainable Cities and Society, (2016).
26. Sudheer, G. and Suseelatha, A. "Short term load forecasting using wavelet transform combined with holt-winters and weighted nearest neighbor models", *International Journal of Electrical Power & Energy Systems*, **64**, pp. 340-346 (2015).
27. Kavousi-Fard, A., Samet, H. and Marzbani, F. "A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting", *Expert Systems With Applications*, **41**(13), pp. 6047-6056 (2014).
28. Cao, G. and Wu, L. "Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting", *Energy*, **115**, pp. 734-745 (2016).
29. Xiao, L., Wang, J., Yang, X. and et al. "A hybrid model based on data preprocessing for electrical power forecasting", *International Journal of Electrical Power & Energy Systems*, **64**, pp. 311-327 (2015).
30. Zhang, J., Tan, Z. and Yang, S. "Day-ahead electricity price forecasting by a new hybrid method", *Computers & Industrial Engineering*, **63**(3), pp. 695-701 (2012).
31. Che, J., Wang, J. and Wang, G. "An adaptive fuzzy combination model based on self-organizing map and support vector regression for electric load forecasting", *Energy*, **37**(1), pp. 657-664 (2012).
32. Nie, H., Liu, G., Liu, X. and et al. "Hybrid of airma and SVMs for short-term load forecasting", *Energy Procedia*, **16**, pp. 1455-1460 (2012).
33. Box, P. and Jenkins, G.M., *Time Series Analysis: Forecasting and Control*, Holden-day Inc, San Francisco (1976).
34. Minerva, T. and Poli, I., *Building ARMA Models With Genetic Algorithms*, Lecture Notes in Computer Science, 2037, pp. 335-342 (2001).
35. Haseyama, M. and Kitajima, H. "An ARMA order selection method with fuzzy reasoning", *Signal Process*, **81**, pp. 1331-1335 (2001).
36. Zhang, G., Patuwo, B.E. and Hu, M.Y. "Forecasting with artificial neural networks: the state of the art", *International Journal of Forecasting*, **14**, pp. 35-62 (1998).
37. Ishibuchi, H. and Tanaka, H. "Interval regression analysis based on mixed 0-1 integer programming problem", *J Japan Soc. Industrial Engng*, **40**(5), pp. 312-319 (1988).
38. http://www.ucei. berkeley. edu/ datamine/px_sup_dmd_bid. htm

www.SID.ir