

عملکرد دو روش ARIMA و شبکه عصبی GMDH در پیش‌بینی تقاضای گاز طبیعی در بخش‌های مختلف (ایران: ۱۳۸۰-۱۳۸۹)*

حمید ابریشمی^۱
فرخنده جبل‌عاملی^۲
معصومه ابوالحسنی^{۳*}
افشین جوان^۴

تاریخ پذیرش: ۱۳۹۳/۰۶/۱۸

تاریخ دریافت: ۱۳۹۲/۰۴/۱۲

چکیده

باتوجه به افزایش روزافزون مصرف گاز طبیعی، برنامه‌ریزی در بخش گاز طبیعی و بررسی و پیش‌بینی تقاضای گاز طبیعی جهت دستیابی به امنیت عرضه انرژی گاز طبیعی و به‌دنبال آن توسعه پایدار اهمیت فراوانی دارد. از این‌رو در این تحقیق تقاضای گاز طبیعی در بخش‌های خانگی-تجاری، صنعت و نیروگاه که جزء مصرف‌کنندگان عمده گاز طبیعی هستند مورد بررسی قرار گرفته و از دو روش ARIMA (Autoregressive Integrated Moving Average) و شبکه عصبی GMDH (Group Method of Data Handling) برای پیش‌بینی تقاضای گاز طبیعی و از معیارهای MSE (Mean Squared Error)، RMSE (Root Mean Squared Error)، درصد خطای پیش‌بینی و دقت پیش‌بینی برای مقایسه دو روش استفاده شده‌است. با توجه به نتایج، دقت پیش‌بینی به‌ترتیب در سه بخش خانگی-تجاری، صنعتی و نیروگاه در روش ARIMA، ۹۳/۸، ۹۸/۳ و ۸۷ درصد و در روش شبکه عصبی GMDH، ۹۶/۴، ۹۹ و ۹۸/۲ درصد به‌دست آمده است و معیارهای RMSE و MSE در هر سه بخش برای روش شبکه عصبی GMDH کوچک‌تر از روش ARIMA بوده است. از این‌رو می‌توان نتیجه گرفت که با توجه به مدلسازی صورت گرفته، روش شبکه عصبی GMDH عملکرد و دقت بالاتری نسبت به روش ARIMA در پیش‌بینی تقاضای گاز طبیعی دارد.

کلید واژه‌ها: پیش‌بینی تقاضای گاز طبیعی، شبکه عصبی GMDH، ARIMA

طبقه‌بندی JEL: C45, C53, Q47

Email: Abrishami@ut.ac.ir

۱. استاد دانشکده اقتصاد دانشگاه تهران

Email: Fameli@ut.ac.ir

۲. استادیار دانشکده اقتصاد دانشگاه تهران

Email: Abolhasani.masomeh@gmail.com

۳. کارشناس ارشد اقتصاد انرژی دانشگاه تهران (نویسنده مسئول)

Email: Af_javn@yahoo.co.uk

۴. عضو هیئت علمی - موسسه مطالعات بین‌المللی انرژی

* این تحقیق با حمایت شرکت ملی گاز ایران انجام شده‌است.

۱. مقدمه

انرژی نقش مهم و تعیین کننده در اقتصاد جهان ایفا می کند و از محورهای توسعه است. افزایش روزافزون نیاز جوامع صنعتی به انرژی و افزایش قیمت انرژی از یک سو و محدودیت منابع انرژی تجدیدناپذیر و پیامدهای استفاده از سوخت های فسیلی بر محیط زیست از سوی دیگر، کشورهای عمده مصرف کننده انرژی را به سمت محدودیت در مصرف انرژی به خصوص انرژی های فسیلی و توجه بیشتر به انرژی های نو و همچنین گاز طبیعی سوق داده است. گاز طبیعی انرژی برتر قرن ۲۱ شناخته شده است و به دلیل قیمت مناسب و آلاینده کمتری محیط زیست نسبت به سایر سوخت های فسیلی مورد توجه قرار گرفته است. طبق بررسی های انجام شده در بین انرژی های اولیه، مصرف گاز طبیعی در سال های آینده دارای بالاترین رشد خواهد بود و انتظار می رود بیشترین رشد مصرف گاز طبیعی متعلق به کشورهای در حال توسعه باشد. اهمیت انرژی در اقتصاد کشور ما کاملاً آشکار است و انرژی علاوه بر نقش زیربنایی، دارای نقش سیاسی، اجتماعی و اثر درآمدی است. به همین دلیل، مدیریت انرژی برنامه ریزی دقیق در زمینه عرضه و تقاضای انرژی و به خصوص گاز طبیعی در رسیدن به توسعه پایدار ضروری است. براساس گزارش ترازنامه انرژی، ایران دارای ۱۵٪ ذخایر گاز جهان و رتبه دوم در جهان است و به دلیل مزیت های فراوان گاز در کشورمان، مدتی است که در کانون توجه مسؤولان قرار گرفته است. از جمله مزیت ها می توان به وفور و پراکندگی منابع گاز طبیعی در ایران، هزینه کمتر استخراج نسبت به نفت و قیمت نسبی پایین اشاره کرد. با توجه به اهمیت روزافزون گاز طبیعی و ارتباط آن با سایر بخش ها می توان گفت گاز طبیعی نقش مهمی در فرآیند تصمیم گیری های اقتصادی ایفا می کند و سیاست گذاران برای برنامه ریزی در این بخش، نیازمند کسب اطلاعات از رفتار مصرف کنندگان گاز طبیعی هستند. بررسی و پیش بینی تقاضای گاز مسؤولان را در جهت اعمال مدیریت تقاضای گاز و سیاست گذاری های صادرات گاز در کشور یاری می کند (خالقی، ۸۶). از این رو هدف از این تحقیق مقایسه عملکرد دو روش ARIMA و شبکه عصبی GMDH و به دست آوردن مدل مطلوب برای پیش بینی تقاضای گاز طبیعی در بخش های خانگی-تجاری، صنعت و نیروگاه (که جزء مصرف کنندگان عمده گاز طبیعی در کشور قرار دارند) است.

این مقاله در پنج قسمت سامان یافته است. پس از مقدمه در بخش دوم، مبانی نظری مدل سازی در دو روش شبکه عصبی GMDH و باکس-جنکینز و استفاده از این دو روش در پیش بینی سری های زمانی ارائه شده است. در بخش سوم، مطالعات صورت گرفته در زمینه تقاضای گاز طبیعی بررسی شده و به مطالعاتی که در آن از روش های ARIMA و شبکه عصبی GMDH برای پیش بینی استفاده شده پرداخته شده است. بخش چهارم به مدل سازی و پیش بینی تقاضای گاز طبیعی با استفاده از دو

روش ARIMA و شبکه عصبی GMDH اختصاص دارد. بخش پنجم مقایسه عملکرد دو روش در سه بخش خانگی-تجاری، صنعت و نیروگاه و نتایج حاصله ارائه شده است.

۲. مبانی نظری پیش‌بینی تقاضای گاز طبیعی

مقدار یک متغیر در دوره‌های از زمان را سری زمانی می‌نامند. هر سری زمانی را می‌توان محصول تولید یک فرآیند تصادفی^۱ دانست و مجموعه پیوسته‌ای از داده‌ها، یک تحقق واقعی از فرآیند تصادفی اصلی است (گجراتی، ۱۳۸۳). در واقع می‌توان گفت فرآیند تصادفی مانند جامعه و تحقق واقعی آن مانند یک نمونه از جامعه است. فرآیند تصادفی ساکن یک نوع از فرآیند سری‌های زمانی است. با مدلسازی سری زمانی می‌توان از آن جهت پیش‌بینی استفاده کرد. در تحلیل سری‌های زمانی برخلاف مدل‌های ساختاری اقتصادسنجی، نیازی به تئوری اقتصادی نیست. در این روش با استفاده از رفتار گذشته یک متغیر، پیش‌بینی مقدار آینده آن صورت می‌گیرد. با توجه به این که در اقتصاد انرژی و به خصوص تقاضای گاز محدودیت‌های فراوانی در به‌دست آوردن داده‌ها در مورد متغیرهای مؤثر بر تقاضای گاز وجود دارد، روش تحلیل سری زمانی بسیار مفید است.

۲-۱. فرآیند خودرگرسیون میانگین متحرک انباشته (ARIMA)^۲

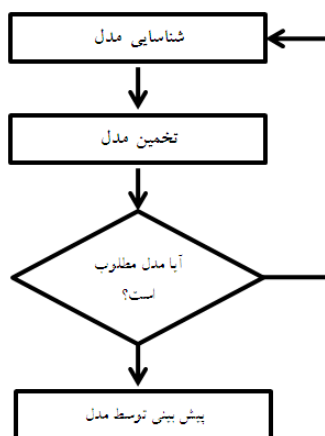
یکی از روش‌های معروف مدلسازی و پیش‌بینی سری زمانی، روش خودرگرسیون میانگین متحرک انباشته (ARIMA) است. به‌طور کلی فرآیندی را $ARMA(p,q)$ می‌نامند که شامل p مرتبه جمله خودرگرسیون و q مرتبه جمله میانگین متحرک باشد. یک فرآیند خودرگرسیون مرتبه p و میانگین متحرک مرتبه q به‌صورت زیر نشان داده می‌شود. ε_t در این رابطه جزء اختلال است.

$$Y_t = \mu + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \dots + \gamma_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

یکی از مسائل مهم در این فرآیند ساکن بودن سری زمانی است. یک سری زمانی در صورتی ساکن است که قدر مطلق ریشه‌های آن بزرگ‌تر از واحد باشند. در صورت غیرساکن بودن سری زمانی، برای مانا کردن آن نیاز به تفاضل‌گیری از سری زمانی است. بنابراین اگر یک سری زمانی پس از d بار تفاضل‌گیری مرتبه‌ی اول مانا شود، سری زمانی خودرگرسیونی میانگین متحرک انباشته^۳ یا $ARIMA(p,d,q)$ به‌دست می‌آید. روش مدلسازی ARIMA به متدولوژی باکس-جنکینز نیز معروف می‌باشد. باکس-جنکینز در سال ۱۹۷۶ ابزاری جهت پیش‌بینی ارائه کردند که از نظر تکنیکی به متدولوژی ARIMA شهرت یافته است. از آن‌جا که از تئوری اقتصادی در این روش استفاده

1. Stochastic
2. Autoregressive Moving Average
3. Autoregressive Integrated Moving Average

نمی‌شود، مدل غیرتئوریک نیز نامیده می‌شود. مسأله مهم در این مدل‌سازی تعیین نوع فرآیند و مرتبه آن است. باکس-جنکینز روشی را جهت تعیین نوع فرآیند و مرتبه آن ارائه می‌کند. مدل‌سازی از روش باکس-جنکینز شامل سه مرحله است که در شکل ۱ نشان داده شده‌است.



شکل ۱: مراحل مدل‌سازی در روش باکس-جنکینز (Suhartono 2011)

شناسایی^۱ اولین مرحله در مدل‌سازی به روش باکس-جنکینز است. منظور از شناسایی، تعیین مرتبه‌ی d, q, p تعیین است. برای تعیین مرتبه q, p از توابع خودهمبستگی و خودهمبستگی جزئی کمک می‌گیریم. در یک سری زمانی ساکن برای یک فرآیند $AR(p)$ تابع خودهمبستگی ACF ترکیبی از توابع نمایی و سینوسی با دامنه نوسانات کاهنده است و تابع خودهمبستگی یک فرآیند $MA(q)$ بعد از وقفه q برابر صفر می‌شود. تابع خودهمبستگی برای یک فرآیند $ARMA(p, q)$ پس از وقفه $q-p$ ترکیبی از توابع نمایی و سینوسی با دامنه نوسانات کاهنده است (ابریشمی، ۱۳۸۱). بعد از تشخیص مدل، مرحله تخمین و تعیین ضرایب مدل است. چندین روش برای تخمین پارامترهای مدل وجود دارد. بعضی از روش‌ها عبارتند از: روش گشتاورها، روش حداکثر درست نمایی و روش حداقل مربعات غیرخطی (ابریشمی، ۱۳۸۱). در مرحله بعد از تخمین ضرایب الگو به دنبال پیدا کردن بهترین مدل هستیم، به همین دلیل این مرحله را مرحله تأیید الگو نیز می‌نامند. در این مرحله براساس معیارها، الگو ارزیابی می‌شود و اگر الگو مطلوب نباشد دوباره به مرحله شناسایی بر می‌گردیم.

مسأله مهم در تحلیل سری زمانی، مانایی سری می‌باشد، لذا قبل از برآورد یک الگوی سری زمانی باید از مانا بودن آن اطمینان حاصل شود. بسیاری از سری‌های زمانی نامانا هستند. سری‌های زمانی

1. Identification

نامانا دارای میانگین و یا واریانس متغیر نسبت به زمان هستند. یک سری زمانی y_t را در صورتی که دارای شرایط زیر باشد مانا می‌نامیم (S. G. Hall D. Asterion 2007).

$E(y_t)$ (میانگین انتظاری y_t): در تمام دوره زمانی t ثابت باشد.

$Var(y_t)$ (واریانس y_t): در تمام دوره زمانی t ثابت باشد.

$Cov(y_t, y_{t+k})$ (کوواریانس y_t و y_{t+k}): برای تمام t و k ها ثابت باشد. به این معنی است که واریانس و کوواریانس سری زمانی در تمام دوره زمانی ثابت باشد.

در اقتصاد معمولاً سری‌های زمانی از طریق لگاریتم گیری مانا در واریانس می‌شوند. به همین دلیل تجزیه و تحلیل‌های سری زمانی عموماً بر روی لگاریتم متغیرها صورت می‌گیرد (حمید ابریشمی، ۱۳۸۱). جهت بررسی مانا بودن میانگین سری‌ها و تعیین تعداد دفعات تفاضل گیری برای مانا کردن سری از آزمون دیکلی فولر^۱ استفاده می‌شود. برای انجام آزمون دیکلی فولر رگرسیون به شکل زیر بیان می‌شود.

$$y_t = \mu + \beta t + \gamma y_{t-1} + \varepsilon_t \quad (2)$$

در آزمون دیکلی فولر فرضیه H_0 بدین گونه تعریف می‌شود.

$$H_0 : \gamma = 1$$

$$H_1 : |\gamma| < 1$$

آماره محاسباتی برای ضریب y_{t-1} (γ) با مقادیر بحرانی آماره t دیکلی فولر (τ) مقایسه می‌شود. چنانچه قدرمطلق محاسباتی بزرگ‌تر از مقادیر بحرانی باشد، فرضیه H_0 رد می‌شود. یعنی می‌توان فرضیه صفر مبنی بر نامانا بودن سری زمانی را رد کرد (William H. Greene, 2002). آماره t دیکلی فولر از رابطه زیر به دست می‌آید.

$$\tau = DF_t = \frac{\hat{\gamma} - 1}{Est.Std.Error(\hat{\gamma})} \quad (3)$$

۲-۲. شبکه‌های عصبی مصنوعی^۲

در سال‌های اخیر شبکه‌های عصبی به‌عنوان رقیبی در مقابل روش‌های سنتی اقتصاد سنجی مورد توجه واقع شده‌اند و به سرعت در حال پیشرفت و گسترش هستند. شبکه‌های عصبی مصنوعی یکسری

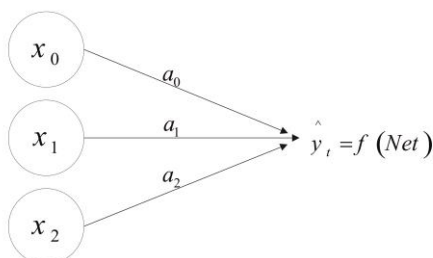
1. Dickey-Fuller
2. Artificial Neural Network

از مشاهدات رابه عنوان ورودی می‌گیرند و با استفاده از آن به قضاوت می‌پردازند و پس از آموزش به صورت عملیاتی و عمدتاً در پیش‌بینی سری‌های زمانی به کار گرفته می‌شوند. شبکه‌های عصبی مصنوعی معمولاً از چند لایه تشکیل شده‌اند که در هر لایه تعدادی عناصر پردازشگر که به نرون معروف است وجود دارد. هریک از نرون‌ها سیگنال‌هایی که حاوی اطلاعات کلی یا محرک‌های خارجی از دیگر نرون‌ها هستند را به عنوان ورودی دریافت کرده و آن‌ها را پردازش می‌کند و سیگنال خروجی راتحویل می‌دهد. در واقع نرون‌ها در شبکه‌های عصبی با استفاده از توابع فعال‌سازی پردازش اطلاعات را انجام می‌دهند و اطلاعات پس از پردازش به عنوان یک سیگنال خروجی به سایر نرون‌ها انتقال می‌یابد. تعداد نرون‌ها و لایه‌های مخفی و قرار گرفتن این نرون‌ها در لایه‌های مختلف و نحوه اتصال نرون‌ها به یکدیگر در هر لایه، معماری شبکه عصبی را تعیین می‌کند. شبکه عصبی در ساده‌ترین حالت که دارای تنها دو لایه ورودی و خروجی است را می‌توان به صورت $f: x \rightarrow y$ تعریف کرد. خروجی y و ورودی‌های x_0 ، x_1 و x_2 بردارهای $1 \times n$ هستند که در آن n تعداد مشاهدات است. مدل مورد بحث به علت حرکت اطلاعات از ورودی‌ها به خروجی‌ها به شبکه عصبی مصنوعی پیشخور^۱ معروف است. اهمیت‌سنجی هریک از ورودی‌ها و ارتباط آن با خروجی به وسیله وزن‌های آنها (a) مشخص می‌شود. ارزش نرون خروجی به صورت رابطه زیر به دست می‌آید:

$$Net_t = a_0x_{0t} + a_1x_{1t} + a_2x_{2t} = \sum_{i=0}^2 a_i x_{it} \quad (4)$$

با استفاده از تابع تبدیل یا فعال‌ساز که با $f(x)$ نشان داده می‌شود نرون های خروجی پردازش می‌شوند. تابع فعال‌ساز می‌تواند خطی و یا غیرخطی باشد. انتخاب نوع تابع فعال‌ساز بستگی به مسأله‌ای دارد که مورد بررسی شبکه عصبی قرار دارد. در ساده ترین شکل شبکه عصبی پیش‌خور، تابع فعال‌ساز خطی است. ارزش به دست آمده از رابطه‌ی ۴ و یک تابع فعال‌ساز خطی، خروجی نهایی شبکه را برای مشاهده t به ما می‌دهد. که در شکل ۲ و به صورت رابطه ۵ نشان داده شده است.

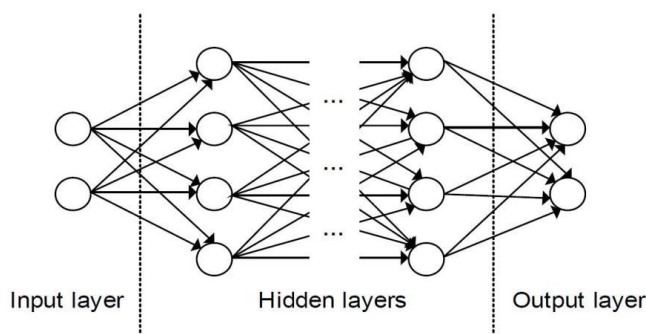
1. Feed forward artificial neural network



شکل ۲: مدل شبکه عصبی ساده پیش خور

$$\hat{y}_t = f(Net_t = a_0x_{0t} + a_1x_{1t} + a_2x_{2t}) = a_0x_{0t} + a_1x_{1t} + a_2x_{2t} \quad (5)$$

همان طور که مشاهده می شود یک شبکه عصبی پیش خور با دو لایه و تابع فعال ساز خطی مشابه مدل رگرسیون خطی چند متغیره است. نرون های ورودی همان متغیره های مستقل هستند و نرون خروجی همان برآورد متغیر وابسته است. وزن های مختلف شبکه نیز، مشابه پارامترهای مدل رگرسیون است (سلیمانی کیا، ۱۳۸۶). در مثال بالا تابع فعال ساز به صورت خطی فرض شده است، اما در واقع برای رسیدن به نتیجه مطلوب و عملکرد بهتر باید از تابع فعال ساز غیرخطی استفاده کرد. رایج ترین تابع فعال ساز غیرخطی در شبکه عصبی تابع توزیع لجستیک است. در شبکه عصبی معمولاً از یک لایه پنهان^۱ و یا بیشتر جهت توضیح پیچیدگی های رفتار متغیر استفاده می کنند و در واقع شبکه های عصبی با لایه پنهان عملکرد بهتری نسبت به شبکه عصبی با دو لایه ورودی و خروجی دارند. در شکل ۳ شبکه عصبی با دو لایه پنهان نمایش داده شده است.



شکل ۳: شبکه عصبی با ۲ لایه پنهان (G. Kayakutlu D. Çınar, 2007)

1. Hidden Layer

لایه پنهان نیز مانند لایه خروجی شامل مجموع وزنی از متغیرهای ورودی است که توسط تابع فعالساز که در بیشتر مواقع تابع لجستیک است، پردازش در آن صورت می‌گیرد. یکی از مراحل مدلسازی در شبکه عصبی مصنوعی آموزش^۱ است که به معنی تعیین وزن‌های شبکه و در واقع معادل تخمین ضرایب در رگرسیون‌های سنتی است. آموزش و تعیین وزن‌های شبکه باید براساس هدف الگو صورت گیرد. هدف الگو کاهش خطای پیش‌بینی نسبت به خروجی‌های الگو است. در واقع در آموزش، هر دو متغیر ورودی (X) و ارزش متغیر هدف (y) به شبکه داده می‌شود و هدف کاهش خطای اختلاف بین خروجی شبکه (f(x)) و ارزش متغیر هدف (y) است. در هر بار تکرار پردازش اطلاعات، وزن‌های شبکه تعدیل می‌شود به گونه‌ای که خطای پیش‌بینی کاهش یابد. پس از آموزش مدل، باید قدرت پیش‌بینی را ارزیابی کرد به همین منظور داده‌ها را دو قسمت می‌کنند. بخش اول را که به مجموعه‌ی آموزش معروف است جهت آموزش شبکه و بخش دوم را که به مجموعه‌ی آزمون معروف است را جهت بررسی قدرت پیش‌بینی مدل به کار می‌گیرند. با استفاده از معیار میانگین مربع خطا^۲ و ریشه میانگین مربع خطا^۳ و درصد خطای پیش‌بینی^۴ می‌توان عملکرد یک مدل را در پیش‌بینی ارزیابی نمود.

همان‌طور که گفته شد شبکه‌های عصبی دارای مزایای فراوانی نسبت به روش‌های سنتی اقتصادسنجی هستند، اما در بعضی موارد شبکه‌های عصبی دارای محدودیت‌هایی از قبیل نیاز به داده‌های ورودی زیاد جهت آموزش شبکه و عدم ارائه اطلاعات کافی در مورد روابط بین متغیرهای ورودی و خروجی هستند. ضعف مدل‌های آماری و الگوریتم‌های شبکه عصبی در پیش‌بینی و ارائه مدل مطلوب منجر به ایجاد روش سازماندهی گروهی داده‌ها^۵ (GMDH) شد که روشی جهت آموزش آماری شبکه است. در طراحی شبکه‌های عصبی GMDH هدف جلوگیری از رشد و واگرایی شبکه و نیز مرتبط کردن شکل ساختار شبکه به یک یا چند پارامتر عددی است به گونه‌ای که با تغییر این پارامترها ساختار شبکه نیز تغییر کند (فخرایی، ۸۶). الگوریتم GMDH توسط ایواخنکو (۱۹۶۶) به صورت یک جمله‌ای بسیط با مراتب بالا ایجاد شده است و این روش برای به دست آوردن روابط بین متغیرهای ورودی و خروجی با مسأله مهم در تحلیل‌های رگرسیونی، پیش‌داوری در مورد ساختار

1. Training

2. Mean Squared Error
$$MSE = \frac{\sum (\bar{y}_t - y_t)^2}{n}$$

3. Root Mean Squared Error
$$RMSE = \sqrt{\frac{\sum (\bar{y}_t - y_t)^2}{n}}$$

4. Error Percentage
$$EP = \frac{\sqrt{MSE}}{Meany} * 100 = \frac{RMSE}{Meany} * 100$$

5. Group method of data handling

مدل و متغیرهای مرتبط است، که اما در الگوریتم GMDH مدل فقط بر پایه داده‌ها طراحی می‌شود و هیچ‌گونه پیش‌زمینه نظری درباره‌ی نحوه عملکرد داده‌ها از سوی محقق در آن وجود ندارد. الگوریتم GMDH ایواخنکو یک مدل خود تنظیم‌کننده می‌سازد که امکان پیش‌بینی و تشخیص را دارد. مبنای الگوریتم عبارت است از فرآیندی جهت ساختن یک چند جمله‌ای با مراتب بالا است که به صورت رابطه ۶ است:

$$\hat{y} = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \dots \quad (6)$$

در واقع هدف این الگوریتم به دست آوردن ضرایب مجهول a_i می‌باشد. همان‌طور که در شکل ۴ نشان داده شده است، متغیرهای ورودی شامل x_1 و x_2 و x_3 و ... و x_m و y متغیر خروجی است که m تعداد متغیرهای ورودی و هر یک از متغیرهای ورودی به شبکه، شامل n مشاهده و nt تعداد مشاهدات در مجموعه آموزش است.

Y		X			
	y_1	x_{11}	x_{12}	...	x_{1m}
مجموعه	y_2	x_{21}	x_{22}	...	x_{2m}
آموزش	\vdots	\vdots	\vdots	...	\vdots
	y_{nt}	$x_{nt,1}$	$x_{nt,2}$...	$x_{nt,m}$
	\vdots	\vdots	\vdots	...	\vdots
مجموعه	\vdots	\vdots	\vdots	...	\vdots
کنترل	\vdots	\vdots	\vdots	...	\vdots
	y_n	x_{n1}	x_{n2}	...	x_{nm}
	y	x_1	x_2	...	x_m

شکل ۴: متغیرهای ورودی الگوریتم GMDH (بوروبور ۱۳۹۰)

الگوریتم دارای لایه‌های مختلفی می‌باشد. در مرحله‌ی اول وقتی داده‌های ورودی به شبکه GMDH داده می‌شود، ترکیبی از داده‌های اولیه تولید می‌شود (Z_i) و به لایه اول فرستاده می‌شود. ورودی‌های این لایه سپس به صورت ترکیبی از داده‌ها طبقه‌بندی می‌شود و به عنوان ورودی به لایه بعدی فرستاده می‌شود. این روند تا زمانی ادامه دارد که نتایج لایه $(n+1)$ بهتر از لایه (n) باشد. زمانی که توسط آزمون همگرایی به این نتیجه برسد نتایج به دست آمده از لایه $(n+1)$ بهتر از لایه

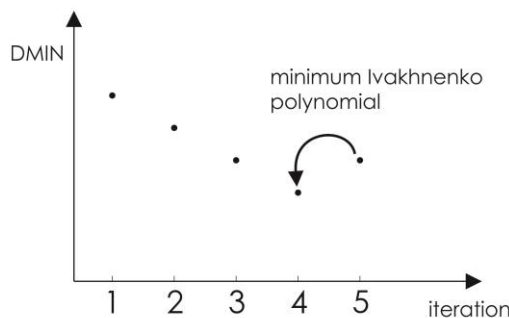
(n) نمی‌باشد روند متوقف می‌شود. متغیرهای Z از برآورد ترکیب $\begin{bmatrix} m \\ 2 \end{bmatrix}$ حداقل مربعات چند درجه‌ای رابطه‌ی ۷ به دست می‌آید. سپس مقدار عددی Z_i به دست آمده از ترکیب زوج‌های مرتب را جایگزین متغیرهای ورودی اولیه (xi) می‌کنیم. بنابراین متغیرهای $\left(z_1, z_2, \dots, z_{\begin{bmatrix} m \\ 2 \end{bmatrix}} \right)$ را به عنوان متغیرهای جدید در نظر می‌گیریم:

$$z(x_i, x_j) = v_0 + v_1 x_i + v_2 x_j + v_3 x_i^2 + v_4 x_j^2 + v_5 x_i x_j \quad (7)$$

در مرحله دوم، در هر لایه غربال‌سازی صورت می‌گیرد. در این مرحله باید Z هایی که بهترین برآوردکننده هستند را انتخاب کرده و مابقی Z ها را حذف می‌کنیم. جهت پیدا کردن بهترین برآوردگر Z_i از معیار حداقل مربع خطای d_j استفاده می‌کنیم. d_j به صورت رابطه‌ی ۸ به دست می‌آید.

$$d_j^2 = \sum_{i=nt+1}^n (y_i - z_{ij})^2 \quad j = 1, 2, k, \begin{bmatrix} m \\ 2 \end{bmatrix} \quad (8)$$

ستون‌های Z را براساس افزایش حداقل مربع خطاهای d_j مرتب می‌کنیم و ستون‌هایی که در آنها d_j از مقدار از پیش تعیین شده M کوچک‌تر باشد را نگه داشته و جایگزین ستون‌های X می‌کنیم. تعداد ستون‌های متغیر جدید ممکن است کوچک‌تر و یا بزرگ‌تر از متغیرهای قبلی (که در اصل m تا بوده) باشد (رضایی، ۱۳۸۷). در مرحله‌ی سوم آزمون همگرایی صورت می‌گیرد. کوچک‌ترین d_j های به دست آمده را DMIN می‌نامیم. در هر لایه اگر مقدار DMIN به دست آمده کمتر از لایه قبل باشد دوباره مراحل ۱ و ۲ برگشته و آن را تکرار می‌کنیم ولی اگر مقدار DMIN به دست آمده در لایه جدید بزرگ‌تر از DMIN قبلی باشد فرآیند را متوقف می‌کنیم و نتایج به دست آمده را مورد استفاده قرار می‌دهیم. برای مثال در شکل ۵ پس از پنج مرتبه تکرار فرآیند متوقف می‌شود.



شکل ۵: فرآیند آزمون همگرایی پس از پنج مرتبه تکرار (رضایی، ۱۳۸۷)

یکی از مسائل مهم در طراحی شبکه GMDH جهت به دست آوردن رابطه‌ی مطلوب بین متغیرهای ورودی و خروجی تعیین مقدار لایه‌ها و نرون‌های هر لایه و ارتباط بین نرون‌ها می‌باشد (سلیمانی کیا، ۱۳۸۶).

۳. پیشینه تحقیق

اولین مطالعات مربوط به تقاضای گاز حدوداً از دهه ۱۹۶۰ آغاز شد و با افزایش اهمیت گاز طبیعی به عنوان منبع انرژی ارزان تر و تمیزتر نسبت به دیگر حامل‌های انرژی و با پیشرفت تکنولوژی‌های جدید تبدیل گاز و به دنبال آن افزایش تقاضا برای گاز طبیعی، مطالعات در این زمینه نیز گسترش یافت. برخی از مطالعات داخلی و خارجی صورت گرفته در زمینه تقاضای گاز طبیعی به صورت زیر هستند.

لیو و کابودان^۱ (۲۰۰۳) تقاضای کوتاه مدت گاز طبیعی کشور آمریکا را با روش برنامه ریزی ژنتیک (GP) سیستم معادلات چندگرسیون پیش بینی و مصرف گاز طبیعی را برای بخش‌های خانگی - تجاری و صنعتی و برق برآورد کرده‌اند. آنها تقاضای گاز هر بخش را تابعی از قیمت گاز، قیمت انرژی‌های جایگزین، شرایط اقتصادی و متغیر دما در نظر گرفته‌اند. نتایج نشان می‌دهد استفاده از برنامه ریزی ژنتیک موجب کاهش خطای پیش بینی شده است.

ادیگر و آکار^۲ (۲۰۰۷) در مطالعه خود از مدل خودرگرسیون میانگین متحرک انباشته (ARIMA) خودرگرسیون میانگین متحرک انباشته فصلی (ARIMA) جهت پیش بینی تقاضای حامل‌های انرژی

1. Kaboudan, M. A. and Liu

2. Ediger, V. and Akar. S

و تقاضای کل انرژی ترکیه در سال‌های ۲۰۰۵ تا ۲۰۲۰ استفاده کردند. نتایج به‌دست آمده نشان می‌دهد که تقاضای همه حامل‌های انرژی مورد بررسی به جز چوب و بیوماس افزایش خواهد یافت و سهم سوخت‌های فسیلی در کل سبد انرژی مصرفی کشور از ۸۷/۶٪ در سال ۲۰۰۵ به ۹۱/۶٪ در سال ۲۰۲۰ خواهد رسید و در آینده گاز طبیعی عامل کلیدی در سیستم انرژی آینده ترکیه خواهد بود. اردوگر^۱ (۲۰۱۰) به بررسی تابع تقاضای گاز طبیعی در بخش‌های صنعتی، نیروگاه، خانگی کشور ترکیه پرداخته است و با استفاده از مدل ARIMA تقاضای گاز را برای سال ۲۰۰۸ تا ۲۰۳۰ پیش‌بینی کرده است. نتایج حاصل از بررسی کشش‌های قیمتی و درآمدی بیانگر آن است که کشش قیمتی تقاضا در بخش خانگی بیشتر از بخش صنعت و نیروگاهی است. در بخش نیروگاه در آمد اثر بیشتری بر تقاضای گاز طبیعی نسبت به قیمت دارد و در نهایت نتایج به‌دست آمده از پیش‌بینی تقاضای گاز طبیعی از رشد متوسط سالانه ۴٪ طی دو دهه آینده حکایت می‌کند.

باغجری (۱۳۸۴) به بررسی ساختاری تقاضای گاز طبیعی ایران در دو بخش خانگی-تجاری و صنعت که بیشترین سهم مصرف گاز طبیعی در کشور دارند پرداخته است. او تقاضای گاز طبیعی در هر دو بخش را با استفاده از فرم تابعی لگاریتم خطی خودرگرسیون مدل‌سازی کرده و از روش پانل و روش تعدیل جریان جزئی به برآورد کشش‌های قیمتی (درآمدی) تقاضای گاز طبیعی پرداخته است.

آذربایجانی و همکاران (۱۳۸۶) عوامل مؤثر بر تقاضای گاز طبیعی در بخش صنعت را مورد بررسی قرار داده‌اند و سپس از طریق تحلیل سری زمانی و با استفاده از الگوی پویای خود توضیح‌دهندگی با وقفه‌های توزیعی (ARDL) و الگوی تصحیح خطا (ECM) تابع تقاضای گاز طبیعی در بخش صنعت را برآورد و تجزیه و تحلیل نموده‌اند. نتایج بررسی نشان می‌دهد قیمت واقعی گاز طبیعی و برق در بخش صنعت تأثیری بر مصرف گاز طبیعی در بخش صنعت ندارد.

شبکه‌های عصبی از دهه ۵۰ شناخته شده بودند اما در اواسط دهه ۸۰ بود که الگوریتم‌ها و روش‌های مربوط به شبکه‌های عصبی مصنوعی به درجه‌ای از پیشرفت رسیدند که در حل مسائل واقعی از آن‌ها استفاده شد (صادقی و ذوالفقاری، ۸۸). ابتدا شبکه‌های عصبی در مطالعات مربوط به بازارهای مالی کاربرد داشت اما بعد از آن به دلیل استفاده از الگوی غیر خطی و خطای کمتر در مطالعات متعددی از این روش برای پیش‌بینی متغیرهای مختلفی مورد استفاده قرار گرفته است. بعضی از مطالعات صورت گرفته در سال‌های اخیر در ذیل آمده است.

آشو و همکاران^۲ (۲۰۰۱) تقاضای کوتاه‌مدت آب در کانپور هند را مورد بررسی قرار داده‌اند. در این مطالعه از سه روش تجزیه تحلیل رگرسیون، تحلیل سری‌های زمانی و شبکه عصبی جهت پیش‌بینی

1. Erdogdu, E.

2. Ashu, J. et al

تقاضای آب استفاده شده است. نتایج به دست آمده بیانگر آن است که الگوهای پیچیده شبکه عصبی دارای عملکرد بهتری نسبت به الگوهای ساده شبکه عصبی است و مدل‌های غیرخطی نسبت به مدل‌های خطی دقت بالاتری دارند.

مورات و سیلان^۱ (۲۰۰۶) با استفاده از روش شبکه عصبی، تقاضای انرژی در بخش حمل و نقل ترکیه را پیش‌بینی کرده‌اند. تولید ناخالص ملی، جمعیت و کیلومتر وسایل نقلیه به عنوان مؤثر بر تقاضای انرژی حمل و نقل در نظر گرفته شده‌اند. در این مطالعه، نتایج حاصل از پیش‌بینی توسط روش شبکه عصبی با پیش‌بینی‌های ارائه شده توسط وزارت انرژی و منابع ترکیه تحت سناریوهای مختلف، مورد بررسی قرار گرفته است.

شکوری و همکاران (۲۰۱۱) به بررسی عوامل مؤثر بر تقاضای گاز طبیعی در بخش خانگی-تجاری ایران و پیش‌بینی آن پرداخته‌اند. تولید ناخالص داخلی، تعداد خانوار، قیمت گاز، قیمت گازوئیل و سرمایه گذاری در ساخت و ساز به عنوان متغیر ورودی در نظر گرفته شده‌اند. در این مطالعه با استفاده از روش شبکه عصبی و داده‌های سال‌های ۱۹۶۷ تا ۲۰۰۷ ابتدا تک‌تک متغیرها و سپس تقاضای گاز بخش خانگی-تجاری را برای سال‌های ۲۰۰۸ تا ۲۰۲۰ پیش‌بینی شده‌اند.

رضانی (۱۳۸۳) به پیش‌بینی تقاضای ماهیانه برق ایران با استفاده از دو روش شبکه عصبی و ARIMA پرداخته است. جهت مدل‌سازی از داده‌های سال‌های ۱۳۷۰ تا ۱۳۷۹ استفاده کرده و برای دو سال ۱۳۸۰ و ۱۳۸۱ نتایج حاصل از پیش‌بینی با استفاده از دو روش را باهم مقایسه کرده است. براساس معیارهای RMSE و MAPE الگوهای شبکه عصبی در پیش‌بینی دارای عملکرد بهتری هستند.

شرزه‌ای و همکاران (۱۳۸۷) از الگوهای ساختاری، سری زمانی و شبکه عصبی نوع GMDH مبتنی بر الگوریتم ژنتیک، جهت پیش‌بینی تقاضای سرانه آب در شهر تهران استفاده کرده‌اند. نتایج حاصل از این مطالعه بیانگر آن است که روش شبکه عصبی نوع GMDH در مقایسه با الگوهای ساختاری و سری زمانی، در پیش‌بینی تقاضای آب با توجه به معیارهای MAE, RMSE و MAPE دارای خطای کمتر و دقت بالاتری است. در این مطالعه علت دقت کمتر پیش‌بینی توسط روش ARIMA نسبت به شبکه عصبی را، عدم استفاده الگوی ARIMA از داده‌های سایر متغیرها مانند قیمت، درآمد و دما و همچنین کافی نبودن تعداد مشاهدات بیان کرده‌اند.

1. Murat, Y. and Ceylan, H.

۴. طراحی مدل و برآورد آن

۴-۱. مدلسازی و پیش‌بینی تقاضای گاز با استفاده از الگوی سری زمانی ARIMA

در این بخش به مدلسازی و پیش‌بینی تقاضای گاز طبیعی در سه بخش خانگی-تجاری، صنعت و نیروگاه به روش ARIMA پرداخته شده است. جهت برآورد الگو از نرم‌افزار Eviews5 و از داده‌های ماهانه مصرف گاز طبیعی سال‌های ۱۳۸۰ تا ۱۳۸۹ استفاده شده است بدین‌صورت که داده‌های ماهانه فروردین ۱۳۸۰ تا بهمن ۱۳۷۸ جهت مدلسازی و از داده‌های ماهانه اسفند ۱۳۸۷ تا اسفند ۱۳۸۹ جهت ارزیابی عملکرد مدل در پیش‌بینی به‌کار گرفته شده‌اند. در ابتدا برای کاهش نوسانات سری زمانی مصرف گاز در هر سه بخش خانگی-تجاری، صنعت و نیروگاه و پایا کردن آن در واریانس، از آن‌ها لگاریتم گرفته شده است. جهت بررسی مانایی سری‌های زمانی مصرف گاز در هر سه بخش، از توابع خودهمبستگی و آزمون دیکی‌فولر استفاده شده است. نتایج بیانگر آن است که سری فوق در سطح ناماناست و با یک‌بار تفاضل‌گیری مانا می‌شود. به همین دلیل در مدلسازی از تفاضل مرتبه اول سری استفاده شده است.

۴-۱-۱. بخش خانگی-تجاری

در این روش پس از اطمینان از مانایی سری، باید مراتب AR و MA تعیین شود. با توجه به نمودار خودهمبستگی و خودهمبستگی جزئی مراتب تعیین و مدل‌های فراوانی برآورد شدند و با اعمال قواعد مدلسازی با روش ARIMA و با توجه به معیار آکاییک و شوآرتز مدل مطلوب که دارای متغیرهای $AR(1)$, $AR(3)$, $AR(12)$, $MA(2)$ می‌باشد انتخاب شده است. در این مدل تفاضل مرتبه اول لگاریتم مصرف گاز طبیعی در بخش خانگی-تجاری (dLRgc) به‌عنوان متغیر وابسته وارد شده است. نتایج حاصل از برآورد مدل در دوره فروردین ۱۳۸۰ تا بهمن ۱۳۸۷ در جدول ۱ ارائه شده است.

جدول ۱: نتایج برآورد مدل در بخش خانگی-تجاری

نام متغیر	ضریب	خطای معیار
AR(1)	۰/۱۸۰۶	۰/۰۶۱۹
AR(3)	-۰/۱۲۱۴	۰/۰۳۴۴
AR(12)	۰/۸۴۹	۰/۰۵۵۲
MA(2)	-۰/۵۷۳۳	۰/۱۰۹

منبع: نتایج تحقیق

در مدل مطلوب ارائه شده خودهمبستگی و ناهمسانی واریانس وجود ندارد. پس از انتخاب بهترین مدل، از آن جهت پیش‌بینی در دوره اسفند ۱۳۸۷ تا اسفند ۱۳۸۹ استفاده شده است. نتایج حاصل از پیش‌بینی با توجه به معیارهای مورد نظر در جدول ۲ ارائه شده است.

جدول ۲: خطای پیش‌بینی تقاضای گاز خانگی - تجاری

معیار	RMSE	MSE	درصد خطا
خطا	۰/۵۴۳۴	۰/۲۹۵۲	۶/۲

منبع: نتایج تحقیق

۲-۱-۴. بخش صنعت

پس از اطمینان از مانایی سری، جهت تعیین مراتب مدل از نمودار خودهمبستگی و خودهمبستگی جزئی کمک گرفته شده است. مدل‌های فراوانی مورد برآورد قرار گرفت اما بعضی از مدل‌ها به دلیل نامناسب بودن حذف شدند. با توجه به مدل‌های برآورد شده، مدلی که دارای متغیرهای $AR(1)$ $MA(1)$ می‌باشد، دارای آکاییک و شوارتز کوچک‌تر و مدل مطلوب می‌باشد. در این مدل تفاضل مرتبه‌ی اول لگاریتم مصرف گاز طبیعی در بخش صنعت (dLlgc) به‌عنوان متغیر وابسته در نظر گرفته شده است. نتایج حاصل از برآورد مدل در دوره فروردین ۱۳۸۰ تا بهمن ۱۳۸۷ در جدول ۳ ارائه شده است.

جدول ۳: نتایج برآورد مدل در بخش صنعت

متغیر	ضریب	انحراف معیار
AR(1)	۰/۴۹۹۷	۰/۰۹۸۲
AR(12)	-۰/۲۳۸۶	۰/۰۸۸۸
MA(1)	-۰/۹۷۹۸	۰/۰۱۸۶

منبع: نتایج تحقیق

پس از انتخاب مدل مطلوب، از آن جهت پیش‌بینی استفاده شده است. با توجه به معیارهای موردنظر، نتایج حاصل از پیش‌بینی با استفاده از مدل مطلوب در دوره اسفند ۱۳۸۷ تا اسفند ۱۳۸۹ در جدول ۴ ارائه شده است.

جدول ۴: خطای پیش‌بینی تقاضای گاز صنعت

معیار	RMSE	MSE	درصد خطا
خطا	۰/۱۵۰۳	۰/۰۲۲۶	۱/۲۴

منبع: نتایج تحقیق

۳-۱-۴. بخش نیروگاه

با توجه به نمودار خودهمبستگی و خودهمبستگی جزئی تفاضل مرتبه‌ی اول سری زمانی مراتب AR و MA که باید در مدل وارد شود بررسی و تعیین شد. مدل‌های فراوان و مختلفی برآورد شد و برخی از مدل‌ها به دلیل نامناسب بودن حذف شدند. با توجه به معیارهای موردنظر مدلی که دارای متغیرهای $AR(1)$, $AR(2)$, $AR(3)$, $AR(12)$, $AR(13)$ $MA(12)$ است به‌عنوان مدل مطلوب انتخاب

شد که در این مدل تفاضل مرتبه اول لگاریتم مصرف گاز طبیعی در بخش نیروگاه متغیر وابسته است. نتایج بدست آمده از برآورد مدل در دوره فروردین ۱۳۸۰ تا بهمن ۱۳۸۷ در جدول ۵ ارائه شده است.

جدول ۵: نتایج برآورد مدل در بخش نیروگاه

متغیر	ضریب	انحراف معیار
AR(1)	-۰/۱۵۴۱	-۰/۱۰۷۷
AR(2)	-۰/۱۲۶۵	-۰/۰۵۵۹
AR(3)	-۰/۱۰۳۶	-۰/۰۴۷۸
AR(12)	۰/۹۱۹	-۰/۰۴۱۶
AR(13)	۰/۴۴۸۹	-۰/۱۲۱۴
MA(12)	-۰/۸۷۴۲	-۰/۰۴۹۲

منبع: نتایج تحقیق

جهت ارزیابی عملکرد مدل مطلوب، از آن برای پیش‌بینی استفاده شده است. نتایج حاصل از پیش‌بینی با توجه به معیارهای موردنظر در جدول ۶ ارائه شده است.

جدول ۶: خطای پیش‌بینی تقاضای گاز نیروگاه

معیار	RMSE	MSE	درصد خطا
خطا	۱۸۳۳/۱	۴۰۰۳/۱	۸/۱۳

منبع: نتایج تحقیق

۲-۴. مدل‌سازی و پیش‌بینی تقاضای گاز طبیعی با استفاده شبکه‌های عصبی GMDH

در این بخش تقاضای گاز طبیعی برای سه بخش خانگی-تجاری، صنعت و نیروگاه با استفاده از روش شبکه عصبی GMDH مدل‌سازی و پیش‌بینی شده است. شبکه مورد استفاده دارای یک لایه ورودی و دو لایه پنهان می‌باشد. لازم به ذکر است که تعداد متغیرهای ورودی در تعیین تعداد لایه‌ها تأثیر دارد و افزایش تعداد لایه‌های پنهان سبب پیچیده‌تر شدن محاسبات و کاهش خطای مدل‌سازی می‌شود اما با مشکلاتی چون افزایش حجم و زمان محاسبات و پیچیدگی فرآیند پیش‌بینی همراه خواهد بود.

۱-۲-۴. متغیرها و داده‌های مدل

برای مدل‌سازی و پیش‌بینی با استفاده از روش شبکه عصبی GMDH از نرم‌افزار Matlab 7.1 و داده‌های ماهانه سال ۱۳۸۰ تا ۱۳۸۹ استفاده شده است. ۸۰ درصد داده‌ها (فروردین ۱۳۸۰ تا بهمن ۱۳۸۷) جهت آموزش شبکه و ۲۰ درصد داده‌ها (اسفند ۱۳۸۷ تا اسفند ۱۳۸۹) به منظور پیش‌بینی در نظر گرفته شده است. در جدول ۷ متغیرهایی که در مدل‌سازی تقاضای گاز در هر سه بخش استفاده شده، ارائه شده است. انتخاب متغیرهایی که بیشترین تأثیر را بر روی خروجی شبکه دارد توسط شبکه

صورت می‌گیرد. با توجه به توانایی الگوریتم GMDH در شناسایی متغیرهای زائد و انتخاب متغیرهای مهم در فرآیند مدل‌سازی، نگرانی بابت استفاده از متغیرهای مختلف در مدل وجود ندارد. برای کاهش نوسانات داده‌ها و قابل مقایسه بودن دو روش ARIMA و شبکه عصبی GMDH، داده‌های متغیرهای قیمت و مصرف گاز در هر سه بخش به صورت لگاریتم وارد مدل شده‌اند. لازم به ذکر است که مأخذ داده‌های قیمت^۱ و مصرف گاز طبیعی، شرکت ملی گاز ایران، واحد مدیریت گاز رسانی و مأخذ داده‌های دمای هوا، سازمان هواشناسی ایران است.

جدول ۷: متغیرهای استفاده شده در مدل‌سازی تقاضای گاز

متغیرها	توضیح
Rgc	مصرف گاز طبیعی در بخش خانگی-تجاری
Rgc(-1)	مصرف یک دوره قبل گاز طبیعی در بخش خانگی-تجاری
Rgc(-2)	مصرف دو دوره قبل گاز طبیعی در بخش خانگی-تجاری
Rgc(-3)	مصرف سه دوره قبل گاز طبیعی در بخش خانگی-تجاری
Pgc	مصرف گاز طبیعی در بخش نیروگاه
Pgc(-1)	مصرف یک دوره قبل گاز طبیعی در بخش نیروگاه
Pgc(-2)	مصرف دو دوره قبل گاز طبیعی در بخش نیروگاه
Igc	مصرف گاز طبیعی در بخش صنعت
Igc(-1)	مصرف یک دوره قبل گاز طبیعی در بخش صنعت
Igc(-2)	مصرف دو دوره قبل گاز طبیعی در بخش صنعت
Pr	متوسط قیمت واقعی گاز طبیعی در بخش خانگی-تجاری
Pr(-1)	متوسط قیمت واقعی یک دوره قبل گاز طبیعی در بخش خانگی-تجاری
Pr(-2)	متوسط قیمت واقعی دو دوره قبل گاز طبیعی در بخش خانگی-تجاری
Pp	قیمت واقعی گاز طبیعی در بخش نیروگاه
Pp(-1)	قیمت واقعی یک دوره قبل گاز طبیعی در بخش نیروگاه
Pp(-2)	قیمت واقعی دو دوره قبل گاز طبیعی در بخش نیروگاه
Pi	قیمت واقعی گاز طبیعی در بخش صنعت
Pi(-1)	قیمت واقعی یک دوره قبل گاز طبیعی در بخش صنعت
Pi(-2)	قیمت واقعی دو دوره قبل گاز طبیعی در بخش صنعت
T	متوسط وزنی حداقل دمای کشور
T(-1)	متوسط وزنی حداقل دمای یک دوره قبل کشور
T(-2)	متوسط وزنی حداقل دمای دو دوره قبل کشور

منبع: نتایج تحقیق

۱. متغیر قیمت در هر سه بخش برپایه شاخص قیمت ماهانه سال ۱۳۸۳ واقعی شده‌اند.

اثر دمای هوا بر مصرف گاز طبیعی به خصوص در بخش خانگی-تجاری بسیار زیاد است، به همین دلیل از متغیر دمای هوا در مدل سازی استفاده شده است. اما باید توجه کرد که در بعضی از مناطق کشور ممکن است دمای هوا در فصول سرد کاهش یابد اما به دلیل عدم دسترسی به گاز طبیعی و یا تعداد کم مصرف کنندگان گاز در این منطقه، کاهش دما تاثیر چندانی بر مصرف گاز نخواهد داشت. به عبارت دیگر علاوه بر دمای هوای هر استان، تعداد مصرف کنندگان و یا مشترکین گاز طبیعی در آن استان نیز در مصرف گاز طبیعی تأثیر دارد. به همین دلیل در این مطالعه از متوسط حداقل دمای هوای کشور به صورت میانگین وزنی دمای هوای استان های کشور (که در آن وزن ها، تعداد واحد مصرف کننده گاز در هر استان است) استفاده شده است. دلیل استفاده از حداقل دمای هوا نیز این است که مصرف گاز در فصول سرد با کاهش دما با هدف گرمایشی افزایش می یابد. اما می توان گفت که حساسیت مصرف کنندگان گاز طبیعی به خصوص در بخش خانگی-تجاری به حداقل دمای هوا بیشتر از متوسط دمای هوا می باشد. در واقع کاهش دما در صورتی باعث روشن شدن دستگاه های گرمایشی در منازل می شود که دما از یک درجه مشخصی کمتر شود.

۲-۲-۴. بخش خانگی-تجاری

برای مدل سازی تقاضای گاز در این بخش پنج مدل مورد بررسی قرار گرفته و در جدول ۸ مدل های برآورد شده ارائه شده است. در جدول ۸، ستونی با عنوان مؤثرترین متغیر آورده شده است، این بدین معنی که در مدل سازی بیشترین اثر را بر خروجی شبکه دارد. در واقع الگوریتم GMDH خود متغیرهای مؤثر و زائد در هر مدل را شناسایی می کند و متغیر زائد را حذف می کند.

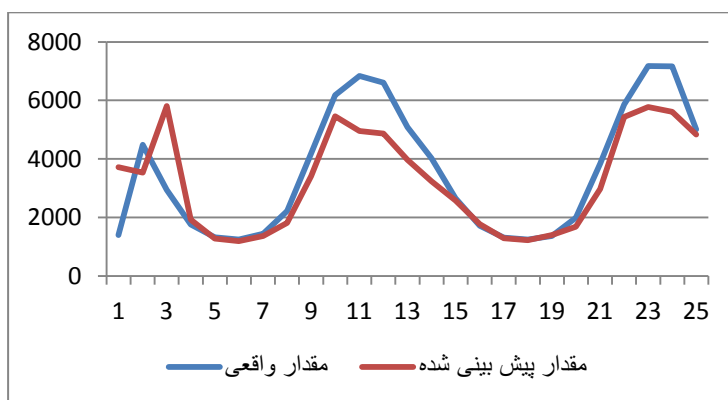
جدول ۸: نتایج حاصل از پیش بینی تقاضای گاز خانگی-تجاری

مدل	متغیر های مدل	مؤثر ترین متغیر	MSE	RMSE	درصد خطا پیش بینی	درصد دقت پیش بینی
۱	T-Pr	Pr	۱/۳۸۶۳	۱/۱۷۷۴۱	۱۴	۸۶
۲	T, T(-1), Pr, Pr(-1), Rgc(-1)	Rgc(-1), T	-۱۰۸۷۵۲	-۰/۲۹۵۸۴	۳/۷	۹۶/۳
۳	T, T(-1), Pr, Pr(-1), Rgc(-1), Rgc(-2)	Rgc(-1), Rgc(-2), T	۳/۷۲	۱/۹۲۸۸	۲۴	۷۶
۴	T-T(-1), T(-2), Pr, Pr(-1), Rgc(-1)	Rgc(-1)	-۰/۶۸	-۰/۸۲۴۶	۱۰	۹۰
۵	T, T(-1), T(-2), Pr, Pr(-1), Rgc(-1), Rgc(-2)	Rgc(-1), T(-1), T	-۱۰۸۳۹۶	-۰/۲۸۹۷	۳/۶	۹۶/۴

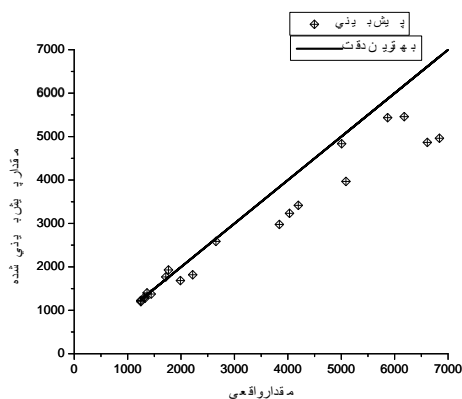
منبع: نتایج تحقیق

با توجه به نتایج به دست آمده، مدل پنجم که دارای میانگین مجذور خطا و درصد خطای کمتر و دقت پیش بینی بیشتری می باشد، مدل مطلوب است. دقت پیش بینی این مدل ۹۶/۴ درصد می باشد و

در این مدل متغیرهای مصرف دوره قبل، حداقل دمای هوا و حداقل دمای هوا دوره قبل دارای بیشترین اثر بر متغیر مصرف گاز طبیعی در بخش خانگی-تجاری است. در نمودار ۱ مقادیر واقعی و پیش‌بینی شده و در نمودار ۲ دقت پیش‌بینی توسط بهترین مدل شبکه عصبی GMDH نشان داده شده است.



نمودار ۱: مقادیر واقعی و پیش‌بینی شده تقاضای گاز در بخش خانگی-تجاری



نمودار ۲: دقت پیش‌بینی در بخش خانگی-تجاری

۳-۲-۴. بخش صنعت

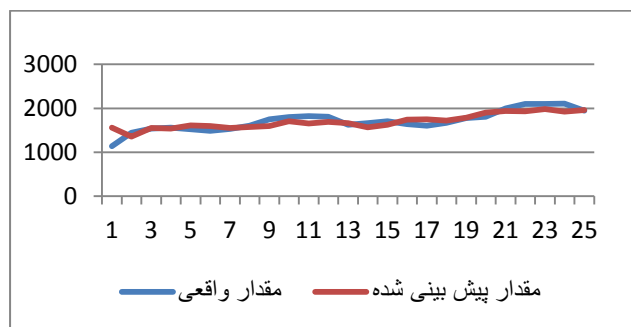
برای مدل‌سازی و پیش‌بینی تقاضای گاز در بخش صنعت، سه مدل با متغیرهای ورودی مختلف مورد بررسی قرار گرفتند. در جدول ۹ نتایج حاصل از پیش‌بینی توسط مدل‌ها ارائه شده است.

جدول ۹: نتایج حاصل از پیش‌بینی تقاضای گاز صنعت

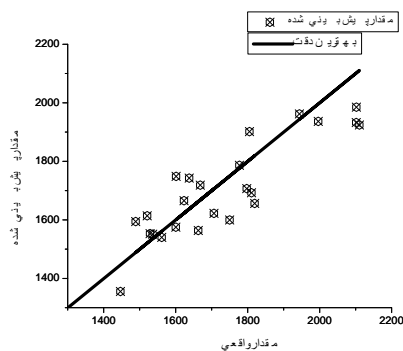
درصد دقت پیش‌بینی	درصد خطا پیش‌بینی	RMSE	MSE	مؤثرترین متغیر	متغیرهای مدل	مدل
۹۵	۵	۰/۴۱۳۲	۰/۱۷۰۸	Pi(-1)	T,T(-1),Pi,Pi(-1),Igc(-1)	۱
۹۹	۱	۰/۰۸۴۶۷	۰/۰۰۷۱۷	Igc(-1)	T,T(-1),Pi,Pi(-1),Igc(-1),Igc(-2)	۲
۹۹	۱	۰/۰۸۴۶۷	۰/۰۰۷۱۷	Igc(-1)	T,T(-1),Pi,Pi(-1),Pi(-2),Igc(-1),Igc(-2)	۳

منبع: نتایج تحقیق

با توجه به نتایج به‌دست آمده مدل دوم و سوم در پیش‌بینی تقاضای گاز دارای عملکرد یکسانی هستند. هر دو مدل دارای دقت پیش‌بینی ۹۹ درصد می‌باشند که نتیجه مطلوبی است. در این مدل‌ها متغیر مصرف دوره قبل گاز طبیعی، متغیر مؤثر است. در نمودار ۳ مقادیر واقعی و پیش‌بینی و در نمودار ۴ دقت پیش‌بینی توسط بهترین مدل شبکه عصبی GMDH نشان داده شده‌است.



نمودار ۳: مقادیر واقعی و پیش‌بینی شده تقاضای گاز در بخش صنعت



نمودار ۴: دقت پیش‌بینی در بخش صنعت

۴-۲-۴. بخش نیروگاه

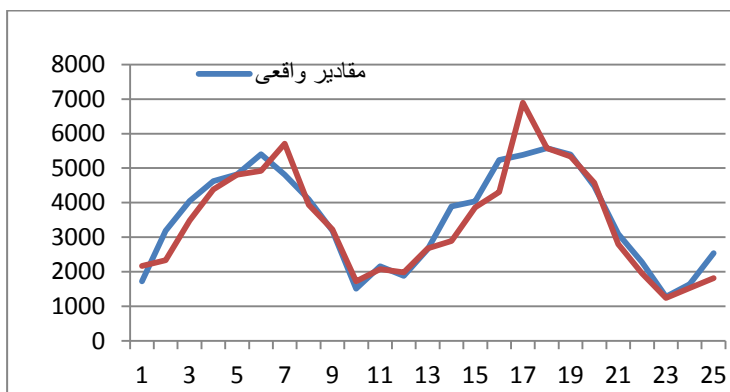
مدل‌های مختلفی جهت پیش‌بینی تقاضای گاز در بخش نیروگاه مورد بررسی قرار گرفت. می‌توان گفت اثر دمای هوا بر مصرف گاز نیروگاه به‌طور غیرمستقیم است، از این‌رو به‌عنوان متغیر ورودی در مدل‌ها وارد شده‌است. در جدول ۱۰ نتایج حاصل از پیش‌بینی توسط مدل‌ها ارائه شده است.

جدول ۱۰: نتایج حاصل از پیش‌بینی تقاضای گاز نیروگاه

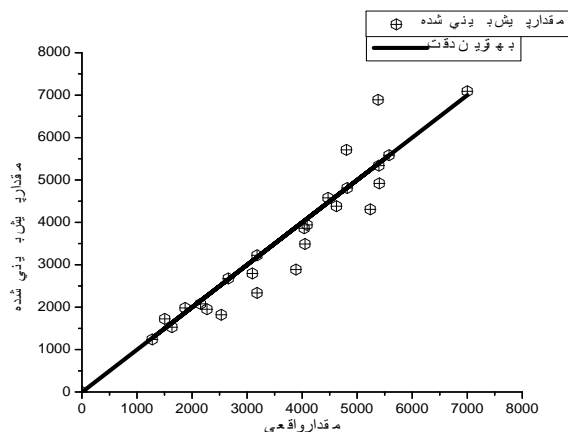
مدل	متغیرهای مدل	مؤثرترین متغیر	MSE	RMSE	درصد خطا پیش‌بینی	درصد دقت پیش‌بینی
۱	T, T(-1), Pp, Pp(-1), Pgc(-1)	T(-1)	-/۳۸۱۷۴	-/۱۹۵۳۸	۲/۴	۹۷/۶
۲	T, T(-1), T(-2), Pp, Pp(-1), Pgc(-1), Pgc(-2)	Pgc(-1)	-/۰۲۳۱۹	-/۱۵۳۳۱	۱/۸	۹۸/۲
۳	Pp, Pp(-1), Pgc(-1), Pgc(-2)	Pgc(-1) Pgc(-2)	-/۲۹۳۶	-/۵۴۰۹	۶/۶	۹۳/۴

منبع: نتایج تحقیق

با توجه به نتایج به‌دست آمده، مدل دوم که دارای میانگین مجذور خطا و درصد خطای کمتر است، بهترین مدل است. دقت پیش‌بینی در مدل دوم ۹۸/۲ درصد می‌باشد که نتیجه مطلوبی است. مؤثرترین متغیر در این مدل مصرف دوره قبل گاز طبیعی است. در نمودار ۵ مقادیر واقعی و پیش‌بینی شده و در نمودار ۶ دقت پیش‌بینی توسط مدل دوم ارائه شده است.



نمودار ۵: مقادیر واقعی و پیش‌بینی شده تقاضای گاز در بخش نیروگاه



نمودار ۶: دقت پیش‌بینی در بخش نیروگاه

۵. مقایسه قدرت پیش‌بینی الگوهای ARIMA و GMDH و نتایج حاصل از

پیش‌بینی تقاضای گاز طبیعی

روش شبکه عصبی GMDH توانایی مدل‌سازی و توضیح روابط غیرخطی بین متغیرها را دارد در صورتی که روش ARIMA یک الگوی خطی است و در نتیجه نمی‌تواند روابط غیرخطی و پیچیده متغیرها را توضیح دهد. اما اگر روابط بین متغیرها از یک فرآیند خطی تبعیت کند، شبکه عصبی ممکن است نسبت به مدل خطی ARIMA نتایج بهتری ارائه نکند. اگرچه الگوهای شبکه عصبی الگوهای غیرتئوریک هستند اما از مزایای شبکه‌های عصبی GMDH این است که می‌توان متغیرهای ورودی مدل را براساس پایه‌های تئوریک تعیین کرد و شبکه عصبی در میان متغیرهای ورودی، متغیرهای مهم و زائد را شناسایی کرده و متغیرهای زائد را از مدل حذف می‌کند. اما در الگوهای سری زمانی ARIMA امکان وارد نمودن متغیرهای دیگر و بررسی اثر سایر عوامل در متغیر موردنظر وجود ندارد، لذا این الگوها برای سیاست‌گذاری اهمیت کمتری دارند. البته برای استفاده از روش شبکه عصبی GMDH برای پیش‌بینی تقاضای گاز، باید مقادیر آینده متغیرهای ورودی در شبکه وارد شود.

در این مطالعه از دو روش ARIMA فصلی و الگوریتم GMDH جهت پیش‌بینی تقاضای گاز طبیعی بخش‌های خانگی-تجاری، صنعت و نیروگاه استفاده شده است. مدل‌های مختلفی با استفاده از هردو روش مورد برآورد قرار گرفتند و بهترین مدل‌ها با توجه به معیارهای موردنظر انتخاب شدند. نتایج حاصل از پیش‌بینی تقاضای گاز در بخش خانگی-تجاری با استفاده از بهترین مدل‌های هردو روش در جدول ۱۱ ارائه شده است.

جدول ۱۱: خطای پیش‌بینی تقاضای گاز در بخش خانگی-تجاری

مدل	MSE	RMSE	درصد خطای پیش‌بینی	درصد دقت پیش‌بینی
ARIMA	۰/۲۹۵۲	۰/۵۴۳۴	۶/۲	۹۳/۸
GMDH	۰/۰۸۳۹	۰/۲۸۹۷	۳/۶	۹۶/۴

منبع: نتایج تحقیق

با توجه به نتایج ارائه شده در جدول و مقایسه دو الگو می‌توان گفت که روش شبکه عصبی GMDH عملکرد بهتری نسبت به روش ARIMA در پیش‌بینی تقاضای گاز بخش خانگی-تجاری دارد. دقت پیش‌بینی در روش GMDH بیشتر از روش SARIMA و حدود ۹۶/۴ درصد است. در جدول ۱۲ نتایج پیش‌بینی تقاضای گاز طبیعی در بخش صنعت با استفاده از هر دو روش ارائه شده است.

جدول ۱۲: خطای پیش‌بینی تقاضای گاز طبیعی در بخش صنعت

مدل	MSE	RMSE	درصد خطای پیش‌بینی	درصد دقت پیش‌بینی
ARIMA	۰/۰۲۲۶	۰/۱۵۰۳	۱/۷	۹۸/۳
GMDH	۰/۰۰۷۱	۰/۰۸۴۶	۱	۹۹

منبع: نتایج تحقیق

همان‌طور که در جدول مشاهده می‌شود، روش شبکه عصبی GMDH در پیش‌بینی تقاضای گاز بخش صنعت نسبت به روش تحلیل سری‌زمانی ARIMA عملکرد بهتری دارد. میانگین مجذور خطا در روش GMDH کمتر می‌باشد. دقت پیش‌بینی توسط هر دو روش تا حدی به هم نزدیک می‌باشد و اختلاف زیادی ندارد. خطای پیش‌بینی در روش ARIMA حدود ۱/۷ درصد و خطای پیش‌بینی در روش GMDH حدود ۱ درصد است. دقت پیش‌بینی حاصل از روش GMDH حدود ۹۹ درصد بدست آمده است که نتیجه مطلوبی بوده و می‌توان گفت این مدل قادر به توضیح تغییرات تقاضای گاز در بخش صنعت است. در جدول ۱۳ نتایج حاصل از پیش‌بینی تقاضای گاز طبیعی در بخش نیروگاه با استفاده از دوروش شبکه عصبی GMDH و تحلیل سری زمانی ARIMA ارائه شده است.

جدول ۱۳: خطای پیش‌بینی تقاضای گاز طبیعی در بخش نیروگاه

مدل	MSE	RMSE	درصد خطای پیش‌بینی	درصد دقت پیش‌بینی
ARIMA	۱/۴۰۰۳	۱/۱۸۳۳	۱۳/۸	۸۶/۲
GMDH	۰/۰۲۳۱	۰/۱۵۲۳	۱/۸	۹۸/۲

منبع: نتایج تحقیق

همان‌طور که در جدول مشاهده می‌شود، اختلاف زیادی بین نتایج دو روش وجود دارد. با توجه به دقت پیش‌بینی ۹۸/۲ درصد حاصل از روش GMDH می‌توان گفت که این مدل برای پیش‌بینی تقاضای گاز در بخش نیروگاه بسیار مناسب می‌باشد.

نتایج به‌دست آمده حاکی از آن است که دقت پیش‌بینی تقاضای گاز حاصل از روش شبکه عصبی GMDH در هر سه بخش خانگی-تجاری، بالاتر از دقت پیش‌بینی در روش ARIMA می‌باشد. در نتیجه می‌توان گفت که مدل به‌دست آمده از الگوریتم GMDH توانایی پیش‌بینی و توضیح تغییرات تقاضای گاز طبیعی با دقت بالا دارد. پیشنهاد می‌شود سیاست‌گذاران از این روش برای پیش‌بینی تقاضای گاز طبیعی در سال‌های آینده استفاده کنند و با توجه به خطای بسیار کم مدل می‌توانند از نتایج حاصله در برنامه‌ریزی‌ها و سیاست‌گذاری‌ها برای مدیریت تقاضای گاز طبیعی و صادرات آن استفاده کنند.

منابع

- ابریشمی، حمید (۱۳۸۱): *اقتصادسنجی کاربردی (رویکردهای نوین)*، انتشارات دانشگاه تهران، تهران.
- باغجری، محمود (۱۳۸۴): تخمین تابع تقاضای گاز، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- بوربور، فاطمه (۱۳۹۰): بررسی سیر تطور قیمت‌گذاری گاز طبیعی و پیش‌بینی قیمت گاز در آمریکا، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- خالقی، شهلا (۱۳۸۶): "مدیریت و سیاست‌گذاری در بخش گاز"، فصلنامه مطالعات اقتصاد انرژی، شماره ۱۵: ۱۶۴-۱۸۶.
- رضایی، منیژه (۱۳۸۷): پیش‌بینی قیمت گازوئیل در بازار خلیج فارس با استفاده از شبکه عصبی GMDH، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- رمضانی، رضا (۱۳۸۳): مقایسه عملکرد پیش‌بینی در مدل‌های شبکه عصبی و ARIMA، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- فخرایی، حسن (۱۳۸۶): مقایسه پیش‌بینی تقاضای آب با استفاده از الگوهای ساختاری، سری زمانی و شبکه‌های عصبی، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- سلیمانی‌کیا، فاطمه (۱۳۸۶): مدلسازی و پیش‌بینی قیمت بنزین در بورس سنگاپور با استفاده از شبکه عصبی نوع GMDH براساس الگوریتم ژنتیک، پایان‌نامه‌ی کارشناسی‌ارشد، دانشکده اقتصاد، دانشگاه تهران.
- گجراتی، دامودار (۱۳۸۳): *مبانی اقتصادسنجی*، ترجمه حمید ابریشمی، انتشارات دانشگاه تهران، تهران.
- Akar, S. and Ediger, V. (2007); "ARIMA Forecasting of Primary Energy Demand by Fuel Turkey", *Energy Policy*, Vol. 35: 1701-1708.
- Erdogdu, E. (2010); "Natural Gas Demand in Turkey", *Applied Energy*, Vol. 87: 211-219.
- Greene, W. H. (2002); *Econometric Analysis*, Pearson Education press, New Jersey.
- Hall, S. G. and Asterion, D. (2007); *Applied Econometrics A Modern Approach*, Palgrave Macmillan press, New York.
- Kayakutlu, G. and Çınar, D. (2007); "Forecasting Production of Renewable Energy Using Cognitive Mapping and Artificial Neural Networks", 19th International Conference on Production Research, Istanbul.
- Liu, Q.W. and Kaboudan, M. A. (2003); "Forecasting Quarterly US Demand for Natural Gas", *ITEM (Information Technology for Economics and Management)*, Vol. 2.
- Li, X.; J Dong, J.; Shanguan, J. and Hook, M. (2011); "Forecasting The Growth of China's Natural Gas Consumption", *Energy*, Vol. 26: 1380-1385.
- Suhartono (2011); "Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average", *Journal of Mathematics and Statistics*, Vol. 7: 20-27.