

توسعه مدل رگرسیون زنجیره مارکوف خاکستری و پیش‌بینی مصرف سالیانه گاز کشور^۱

شهران شفیعی

کارشناسی ارشد مهندسی صنایع، دانشگاه آزاد اسلامی واحد تهران شمال، ایران،
Shahran.shafiei@gmail.com

وحید برادران^۲

استادیار، گروه مهندسی صنایع، دانشگاه آزاد اسلامی واحد تهران شمال، ایران،
v_baradaran@iautnb.ac.ir

تاریخ دریافت: ۱۳۹۷/۱۲/۱۰ تاریخ پذیرش: ۱۳۹۸/۰۶/۱۰

چکیده

پیش‌بینی دقیق مصرف سالیانه گاز کل کشور نقش تعیین‌کننده‌ای در تدوین راهبردهای تأمین انرژی و سیاست‌گذاری این حوزه دارد. مدل رگرسیون زنجیره مارکوف خاکستری روشی مبتنی بر داده‌های سری‌های زمانی است که عملکرد مناسبی جهت تحلیل و پیش‌بینی داده‌های مصرف سالیانه گاز دارد. رگرسیون زنجیره مارکوف خاکستری، ترکیبی از مدل زنجیره مارکوف و مدل رگرسیون خاکستری و ابزار مناسبی برای پیش‌بینی داده‌های تعداد کم با نوسانات تصادفی است. در این مدل، خطاهای پیش‌بینی حاصل از مدل رگرسیون خاکستری به تعدادی بازه با فواصل یکسان تقسیم و از میانگین بازه خطاها و ماتریس احتمال انتقال وضعیت زنجیره مارکوف مربوط به وضعیت خطاها، مقدار اصلاح‌شده خطا محاسبه و به مقدار پیش‌بینی مدل رگرسیون خاکستری اضافه می‌شود. در این تحقیق پیشنهاد می‌شود که به‌جای استفاده از میانگین فواصل، خطای هر بازه به‌صورت عدد خاکستری بین حدود هر بازه به نحوی انتخاب شود که معیار دقت پیش‌بینی کمینه شود. به‌منظور تعیین نقطه بهینه فواصل خطا، از رویکرد طراحی آزمایش‌ها و طرح باکس-بنکن استفاده شده است. مطالعه کاربردی پیش‌بینی مصرف سالیانه گاز نشان می‌دهد که مدل رگرسیون زنجیره مارکوف خاکستری بهبودیافته، دقت بیشتری نسبت به مدل‌های خاکستری، رگرسیون خاکستری و رگرسیون زنجیره مارکوف خاکستری دارد.

طبقه‌بندی JEL: C15, C53, C90, L95

کلیدواژه‌ها: مدل خاکستری، رگرسیون خاکستری، زنجیره مارکوف، طراحی آزمایش‌ها

۱. این مطالعه در قالب طرح تحقیقاتی مصوب و با حمایت مالی شرکت ملی گاز ایران (شماره قرارداد ۹۷۴۰۱۵) انجام شده است. نگارندگان بر خود لازم می‌دانند مراتب تشکر صمیمانه خود را از مدیریت پژوهش و فناوری و مدیریت برنامه‌ریزی تلفیقی شرکت ملی گاز ایران که ما را در انجام این پژوهش یاری دادند، اعلام کنند.
۲. نویسنده مسئول

۱- مقدمه

ایران با داشتن حدود ۳۳/۶ تریلیون مترمکعب ذخایر اثبات شده (یعنی برابر ۱۸ درصد کل ذخایر گاز دنیا) به عنوان اولین کشور دارنده ذخایر گازی در جهان شناخته شده است (ناجی میدانی و رحیمی، ۱۳۹۴). این در حالی است که از نظر مصرف رتبه چهارم را به خود اختصاص داده است که این امر نشان از مصرف بالای گاز در کشور دارد. سهم بیش از ۷۰ درصدی گاز طبیعی در سبد انرژی کشور، رشد سریع مصرف گاز طبیعی در بازار جهانی در سال های اخیر و نقش مؤثر گاز به عنوان سوخت پاک در توسعه فعالیت های اقتصادی و افزایش رفاه اجتماعی کشور، اهمیت روزافزون این منبع انرژی را به عنوان محور توسعه نشان می دهد. مطالعه اعظمی و همکاران (۱۳۹۷) پیرامون رابطه مصرف گاز و رشد اقتصادی در یازده کشور صادرکننده گاز طبیعی نشان می دهد که میزان مصرف گاز طبیعی به طور مثبتی بر رشد اقتصادی این کشورها تأثیر می گذارد. از طرفی بررسی ها نشان می دهد که در بین انرژی های اولیه، مصرف گاز طبیعی در سال های آینده بیشترین رشد را خواهد داشت و این رشد برای کشورهای در حال توسعه بیشتر خواهد بود، به طوری که تا سال ۲۰۲۰ سهم گاز طبیعی از ترکیب انرژی جهان هر سال ۲ درصد افزایش می یابد (فطرس و همکاران، ۱۳۹۶). اما با توجه به محدود بودن منابع و به منظور پیشگیری از مواجه شدن با بحران در آینده، لزوم مدیریت صحیح تولید و مصرف گاز طبیعی از اهمیت بالایی برخوردار است. پیش بینی مصرف گاز طبیعی، نقش قابل ملاحظه ای در تعیین خط مشی انرژی و یکی از ابزارهای قوی تصمیم گیرندگان در جهت برنامه ریزی صحیح، هدایت مصرف و کنترل پارامترهای عرضه و تقاضای انرژی می باشد. به دلیل مؤلفه های مختلف تأثیرگذار بر روند مصرف انرژی و نوسانات بالای مربوط به آن، لزوم بهره گیری از یک مدل کارا و دقیق جهت پیش بینی تقاضای مصرف گاز در کشور بیش از پیش احساس می شود. بنابراین توسعه روش هایی برای پیش بینی دقیق مصرف سالیانه گاز در کشور که کاملاً منطبق با ویژگی های رفتار کمی و کیفی داده های مصرف باشد، ضرورت دارد. پیشنهاد این پژوهش برای پیش بینی مصرف سالیانه گاز و کاهش نوسانات داده های آن استفاده از مدل رگرسیون زنجیره مارکوف خاکستری^۱ است. در این مقاله، ضمن توسعه ای بر بهبود

مدل رگرسیون زنجیره مارکوف خاکستری، از آن برای پیش‌بینی دقیق‌تر میزان مصرف سالیانه گاز در کشور استفاده شده است.

مدل خاکستری^۱ در سال ۱۹۸۲ اولین بار توسط پرفسور دنگ^۲ بر مبنای تئوری سیستم خاکستری معرفی شد و از آن برای پیش‌بینی روند سیستم‌های نامطمئن با تعداد داده‌های مشاهده شده کم، کارا استفاده می‌شود (سان^۳ و همکاران، ۲۰۱۵). مدل خاکستری یک متغیره مرتبه یک یا به اختصار $GM(1,1)$ ، معروف‌ترین و محبوب‌ترین مدل مبتنی بر این تئوری است که می‌تواند در پیش‌بینی کوتاه‌مدت مبتنی بر مسائل نامعلوم با داده‌های گسسته و کم مورد استفاده قرار گیرد (هوسو^۴، ۲۰۰۳). این مدل، یک سری زمانی پرکاربرد در پیش‌بینی است که در معادله دیفرانسیل استفاده شده در آن از ضرایب متغیر زمانی بهره گرفته شده است. علیرغم مزیت‌های استفاده از مدل $GM(1,1)$ ، نظیر سادگی درک و فهم مدل و آسان بودن اجرای آن، استفاده از مدل فوق برای داده‌هایی که نوسانات^۵ بالایی دارند، تأثیرات ناخوشایندی از منظر منحنی برازش در پی خواهد داشت (چنگ سو^۶ و وو، ۱۹۹۷). یکی از توسعه‌های این مدل به‌منظور افزایش دقت پیش‌بینی، مدل رگرسیون زنجیره مارکوف خاکستری یک متغیره مرتبه یک یا به اختصار $MC-RGM(1,1)$ ^۷ است که از ترکیب دو مدل پیش‌بینی خاکستری یک متغیره مرتبه یک بهبودیافته و روش‌های رگرسیون و مارکوف حاصل شده است. به جهت رویکرد اصلاحی بر روی خطای باقیمانده‌ها، این روش می‌تواند تحلیل آماری بهتری برای داده‌های کم، بی‌ثبات و داده‌هایی که نوسانات معناداری دارند، حاصل کند (هوسو و همکاران، ۲۰۰۹؛ هوسو، ۲۰۰۳؛ کومار^۸ و جین، ۲۰۱۰؛ مائو ژان لی^۹ و همکاران، ۲۰۱۱؛ زی^{۱۰} و همکاران، ۲۰۱۵). به دلیل طبیعت غیرخطی و تعداد کم داده‌های مصرف گاز طبیعی، استفاده از مدل‌های ترکیبی مارکوف خاکستری نتایج

1. Grey Model (GM)
2. Deng
3. Xu Sun
4. Hsu
5. Fluctuation
6. Cheng Hsu
7. Markov Chain Regression Grey Model
8. Kumar
9. MAO Zhan-li
10. Xie

دقیق‌تری حاصل خواهد کرد. (مصطفایی^۱ و کردنوری، ۲۰۱۲؛ جانگ جو^۲ و ویگو وانگ، ۲۰۱۰). با توجه به تصادفی بودن داده‌ها، تعیین نقطه بهینه بازه خطاها موجب تعریف ضرایبی در این مدل‌ها می‌شود که تعیین تابع هدف معیار خطای ارزیابی، مستلزم محاسبات دستی و پیچیده‌ای می‌شوند. تاکنون مطالعات زیادی در جهت بهبود کارایی مدل زنجیره مارکوف خاکستری انجام گرفته است. بیشتر این تحقیقات در زمینه افزایش دقت و کارایی مدل مارکوف خاکستری، معطوف به استفاده از مدل‌های کمکی و ترکیبی نظیر مدل مارکوف خاکستری مبتنی بر تبدیل موجک، مدل مارکوف خاکستری بهبودیافته با تقریب تیلور، مارکوف فوریه خاکستری و ... بوده که با روش‌های ریاضی و محاسبات پیچیده آماری و بهره‌گیری از الگوریتم‌های ابتکاری یا فراابتکاری سعی در بهینه‌سازی پارامترهای مدل و افزایش دقت مدل داشته‌اند. کارایی مدل‌های پیش‌بینی به‌اندازه خطای حاصله از پیاده‌سازی مدل بستگی دارد. در محاسبات اصلاحی خطاهای مدل سنتی رگرسیون زنجیره مارکوف خاکستری، از مقدار میانگین بازه خطاها به‌عنوان نماینده فواصل بازه خطاها استفاده می‌گردد ولی چنانچه به‌جای استفاده از حد میانگین، بردار خطاها را به‌صورت عدد خاکستری تعریف نماییم آنگاه تعیین ضابطه تابع معیار خطای مدل و یافتن جواب بهینه، مشکل بوده، بالاخص به دلیل طبیعت تصادفی و تغییر داده‌های مصرف‌گاز بعضاً مستلزم محاسبات دستی و پیچیده و صرف زمان و هزینه بالا می‌باشد. در این تحقیق برای اولین بار بهره‌گیری از رویکرد طراحی آزمایش‌ها و روش رویه پاسخ (RSM)^۳، ضمن پرهیز از محاسبات پیچیده فوق‌الذکر، سعی در بالا بردن دقت مدل پیش‌بینی شده است. رویکرد فوق به‌عنوان تکنیکی سریع و با دقت قابل قبول نسبت به روش‌های کلاسیک بهینه‌سازی، به‌منظور تعیین مقدار بهینه ضرایب استفاده شده است، بنابراین ضمن ارائه رویکردی جدید برای تنظیم بهینه پارامترهای مدل‌های خاکستری، برای نخستین بار از این مدل، برای پیش‌بینی دقیق میزان مصرف گاز در کشور استفاده شده است.

1. Mostafaei

2. Gang Xu

3. Response Surface Methodology

۲- ادبیات موضوع

در این بخش ادبیات نظری تحقیق شامل مدل‌های خاکستری و رگرسیون زنجیره مارکوف خاکستری و پیشینه تحقیق به‌طور مختصر مرور می‌شوند.

مدل $GM(1,1)$ و رگرسیون خاکستری

مبنای توسعه سیستم‌ها و مدل‌های خاکستری، اعداد خاکستری‌اند (دنگ، ۱۹۸۲). عدد خاکستری $Xg \otimes$ عددی است که با حد بالا و پایین مشخص که به‌صورت $Xg = [X_d, X_u]$ بیان می‌شود. هر عدد خاکستری را می‌توان با ضریب سفیدسازی α مطابق رابطه (۱) به یک عدد قطعی (X_g) تبدیل کرد.

$$Xg = \alpha \times X_d + (1 - \alpha) \times X_u, \quad 0 \leq \alpha \leq 1 \quad (1)$$

سری زمانی بردار $x^{(0)}$ را شامل n مشاهده غیرمنفی به‌صورت $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ در نظر بگیرید که در آن $x^{(0)}(m)$ مقادیر مشاهده‌شده متغیر پیش‌بینی در دوره زمانی m است به‌طوری‌که $m=1, 2, \dots, n$ است. در مدل خاکستری یک متغیر مرتبه یک یا $GM(1,1)$ ، مراحل زیر بر داده‌های اولیه انجام می‌شود. ابتدا عملگر تولید تجمعی^۱ (AGO) مطابق رابطه (۲) بر روی داده‌های اولیه $x^{(0)}$ پیاده می‌شود تا سری زمانی $x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ حاصل شود. سری زمانی جدید $x^{(1)}$ کمی منظم‌تر از سری $x^{(0)}$ است و تصادفی بودن داده‌ها در آن کاهش یافته است.

$$x^{(1)}(k) = \sum_{m=1}^k x^{(0)}(m), \quad k = 1, 2, \dots, n \quad (2)$$

جهت هموارسازی بیشتر داده‌ها، سری میانگین جدید $z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n))$ تشکیل می‌شود که در آن $z^{(1)}(k)$ میانگین حسابی دو مقدار $x^{(1)}(k)$ و $x^{(1)}(k-1)$ مطابق رابطه (۳) است.

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k-1)), \quad k=1, 2, \dots, n \quad (3)$$

مدل $GM(1,1)$ بر اساس یک معادله دیفرانسیل خاکستری درجه اول مطابق رابطه (۴) تعیین می‌شود.

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (۴)$$

در رابطه (۴)، چنانچه واحد زمانی یک فرض شود (یک روز، یک ماه یا یک سال) $\frac{dx^{(1)}}{dt}$ برابر خواهد شد با $x^{(1)}(k) - x^{(1)}(k-1)$ که این مقدار همان $x^{(0)}(k)$ است و پس از جایگذاری و ساده سازی $x^{(1)}$ با متوسط وزنی $x^{(1)}(k)$ و $x^{(1)}(k-1)$ ، رابطه (۵) حاصل می شود (سان و همکاران، ۲۰۱۵).

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad k = 1, 2, \dots, n \quad (۵)$$

$z^1(k)$ مقدار سابقه یا زمینه^۱ مقدار $\frac{dx^{(1)}}{dt}$ و پارامترهای a و b به ترتیب ضریب توسعه^۲ و مقدار عمل خاکستری^۳ نامیده می شوند (سان و همکاران، ۲۰۱۵). از رابطه (۶) می توان پارامترهای a و b را به روش تخمین حداقل مربعات^۴ (LSE) تخمین زد.

$$[a \ b]^T = (B^T B)^{-1} B^T Y_n, \quad k = 1, 2, \dots, n \quad (۶)$$

که در آن ماتریس B به صورت (۷) و Y_n ترانهاده بردار داده های اولیه یا $Y_n = (x^{(0)}(1) \ x^{(0)}(2), \dots, x^{(0)}(n))^T$ است.

$$B = \begin{bmatrix} -z^1(1) & 1 \\ -z^1(2) & 1 \\ \dots & \dots \\ -z^1(n) & 1 \end{bmatrix} \quad (۷)$$

از طرفی با حل معادله دیفرانسیل (۴)، مقدار $x^{(1)}(t)$ برابر با رابطه (۸) می شود.

$$\hat{x}^{(1)}(t) = \left(x^{(0)}(0) - \frac{b}{a}\right)e^{-at} + \frac{b}{a}, \quad t = 1, 2, \dots, \quad (۸)$$

حال با استفاده از عملگر معکوس تولید تجمعی^۵ (IAGO) که در رابطه (۹) مشخص

شده، معادله نهایی مدل پیش بینی خاکستری (۱۰) به دست می آید:

$$x^{(0)}(t) = x^{(1)}(t) - x^{(1)}(t-1), \quad t = 1, 2, \dots, \quad (۹)$$

رابطه (۱۰) رابطه نهایی پیش بینی به روش مدل خاکستری است که در آن $x^{(0)}(0)$

اولین مقدار مشاهده در سری زمانی داده های اولیه است و ضرایب a و b از رابطه (۶)

1. background
2. development coefficient
3. grey action quantity
4. Least Square Estimation
5. inverse accumulated generating operation

برآورد شده‌اند. با جایگذاری مقدار t در رابطه ذیل، مقدار پیش‌بینی دوره مربوطه محاسبه می‌شود.

$$\hat{x}^{(0)}(t) = \left(x^{(0)}(0) - \frac{b}{a} \right) e^{-at} (1 - e^a), \quad t = 1, 2, \dots, \quad (10)$$

از آنجا که عموماً مدل‌های $GM(1,1)$ دارای خطا هستند، برای تصحیح مقدار خطای پیش‌بینی و کاهش فواصل مقادیر پیش‌بینی و مقادیر واقعی و با هدف افزایش کارایی مدل، از مدل‌های رگرسیونی استفاده می‌شود. در این مدل، پس از برآورد هر یک از مشاهدات به روش $GM(1,1)$ ، مدل رگرسیون خطی مطابق (۱۱) بین مقادیر پیش‌بینی به روش مدل خاکستری $(\hat{x}^{(0)}(t))$ و مقادیر واقعی مشاهدات $(x^{(0)}(t))$ به ازای مشاهدات آموزش تشکیل می‌شود. در مدل رگرسیون خطی (۱۱)، y_i و x_i به ترتیب مقادیر واقعی مشاهدات و مقادیر پیش‌بینی به روش مدل خاکستری می‌باشند. پارامترهای β_0 و β_1 با استفاده از روش حداقل مربع خطا در رگرسیون برآورد می‌شوند. پس از آنکه مقدار مشاهده جدیدی $(\hat{x}^{(0)}(t) | t = n + k; k \geq 1)$ با رابطه (۱۰) پیش‌بینی شد، مقدار اصلاح‌شده پیش‌بینی با رابطه (۱۱) به‌عنوان مقدار پیش‌بینی نهایی $(\hat{\hat{x}}^{(0)}(t))$ محاسبه می‌شود. این رویکرد به مدل رگرسیون خاکستری (RGM^1) معروف است.

$$y_i = \beta_0 + \beta_1 x_i \quad (11)$$

مدل رگرسیون زنجیره مارکوف خاکستری

مدل رگرسیون زنجیره مارکوف خاکستری یا $MC-RGM(1,1)$ ترکیبی از دو مدل رگرسیون خاکستری و مدل زنجیره مارکوف برای پیش‌بینی داده‌های کم با نوسانات و تغییرات تصادفی بالا توسعه یافته است. در این روش مقدار خطا (باقیمانده)‌های مدل $RGM(1,1)$ تبدیل به مقادیر گسسته (بازه‌هایی برابر) شده و در قالب یک زنجیره مارکوف مدل‌سازی می‌شوند. با استفاده از خاصیت مارکوفی مقادیر خطا و محاسبه احتمال رخداد خطاها به ازای یک مقدار پیش‌بینی، مقدار پیش‌بینی مدل $RGM(1,1)$ اصلاح می‌شود.

در مرحله نخست، از اختلاف مقدار هر مشاهده آموزش $(x^{(0)}(t))$ و مقدار پیش‌بینی با مدل $RGM(1,1)$ ، مقدار خطای مربوطه طبق رابطه زیر محاسبه می‌شود.

$$e(t) = x^{(0)}(t) - \hat{x}^{(0)}(t), \quad t = 1, 2, \dots, n \quad (12)$$

بازه حداقل و حداکثر خطاها (R) به r بازه مساوی تقسیم می‌شود که $R = [\min e(t), \max e(t)]$. تعداد بازه‌ها (r) برابر بخش عدد صحیح $\frac{\ln(n)}{\ln 2}$ در نظر گرفته می‌شود. هر یک از بازه‌ها را می‌توان یک فاصله خاکستری فرض کرد که توسط عدد خاکستری $(\otimes s_j)$ تعریف می‌گردد که در رابطه (۱۳) نشان داده شده است.

$$\otimes s_j = [L_j, U_j], \quad j = 1, 2, \dots, r \quad (13)$$

حد بالا و پایین هر وضعیت j مطابق معادلات (۱۴) به‌دست خواهد آمد.

$$\begin{cases} L_j = \min_t e(t) + \frac{j-1}{r} (\max_t e(t) - \min_t e(t)) \\ U_j = \min_t e(t) + \frac{j}{r} (\max_t e(t) - \min_t e(t)) \end{cases} \quad (14)$$

بر اساس وضعیت خطاها، ماتریس مربعی گذار مرحله m مطابق (۱۶) تشکیل می‌شود که درآیه‌های آن احتمال گذار از یک وضعیت خطا به وضعیت دیگر در پیش‌بینی m دوره بعدی است.

$$P^{(m)} = \begin{bmatrix} P_{11}^m & \dots & P_{1r}^m \\ \vdots & \ddots & \vdots \\ P_{r1}^m & \dots & P_{rr}^m \end{bmatrix} \quad (15)$$

تابع احتمال درآیه‌های ماتریس گذار، P_{ij}^m ، یعنی احتمال انتقال از وضعیت i به j پس از m مرحله یا دوره پیش‌بینی را به‌صورت رابطه (۱۶) تعریف می‌شود.

$$P_{ij}^m = \frac{M_{ij}^m}{M_i}, \quad j = 1, 2, \dots, r \quad (16)$$

M_{ij}^m تعداد دفعات انتقال اتفاق افتاده از وضعیت i به وضعیت j تحت m مرحله و M_i تعداد داده‌هایی است که خطاهای مربوطه در وضعیت i قرار گرفته است. بردار افقی سطر i ام در ماتریس (۱۵) یا $a_{j(i)}^m = (p_{i1}^m, p_{i2}^m, \dots, p_{ir}^m)$ را در نظر بگیرید. هر یک از بردار افقی این ماتریس (p_{ij}^m) بیانگر این مطلب است که اگر خطای دوره‌ای در بازه s_i قرار گیرد، خطایی به اندازه بازه s_j با احتمال p_{ij}^m به پیش‌بینی m دوره بعد آن اضافه یا کم می‌شود. از آنجا که مقدار خطای s_j به‌صورت بازه است، به‌طور قراردادی میانگین حد بالا و پایین بازه مربوطه به‌عنوان مقدار اصلاحی در نظر گرفته می‌شود. به‌منظور انجام

این رویه، بردار وضعیت خطاها را به صورت $V = (v(1), v(2), \dots, v(r))$ تعریف می‌شود که به ازای هر وضعیت j مقدار $v(j)$ از رابطه (۱۷) محاسبه می‌شود.

$$v(j) = \alpha_j L_j + (1 - \alpha_j) U_j \quad j = 1, 2, \dots, r \quad (17)$$

مقدار خطایی که باید به مقدار پیش‌بینی m دوره بعدی روش $RGM(1,1)$ یا $\hat{x}^{(0)}(i+m)$ اضافه شود از حاصل ضرب بردار افقی ماتریس توان m احتمال مربوط به آن وضعیت در ترانهاده بردار وضعیت خطاها و از رابطه زیر محاسبه می‌شود.

$$\hat{e}_r(t+m) = (a_r)^m (V)^T, \quad t = 1, 2, \dots, n \quad (18)$$

در صورتی که فرض کنیم خطای پیش‌بینی دوره t تنها بر یک دوره بعدی اثرگذار باشد و مقدار اصلاحی تنها تحت تأثیر مقدار خطای یک دوره قبلی باشد یا به عبارتی $m=1$ آنگاه مقدار پیش‌بینی حاصل از مدل تصحیح یافته مارکوف $(\tilde{x}^{(0)}(t))$ به صورت ذیل تعریف می‌شود.

$$\tilde{x}^{(0)}(t+1) = \hat{e}_r(t+1) + \hat{x}^{(0)}(t+1), \quad t = 1, 2, \dots, n \quad (19)$$

که $\hat{e}_r(t+1)$ در معادله (۱۹) در واقع مقدار اصلاحی مارکوف و $\hat{x}^{(0)}(t+1)$ و $\tilde{x}^{(0)}(t+1)$ به ترتیب مقادیر پیش‌بینی حاصل از مدل رگرسیون خاکستری و رگرسیون زنجیره مارکوف خاکستری است.

پیشینه تحقیق

در گذشته مسأله پیش‌بینی مصرف انرژی و به تبع آن گاز طبیعی مورد توجه محققین بسیاری بوده است. سبالج و همکاران^۱ (۲۰۱۷) به بررسی اجمالی و تجزیه و تحلیل سیستماتیک آخرین مقالات و تحقیقات صورت گرفته در زمینه پیش‌بینی مصرف گاز طبیعی از سال ۲۰۰۴ تا ۲۰۱۷ پرداخته‌اند. روش‌های این تحقیقات از نظر افق زمانی به سه دسته کوتاه‌مدت (ساعتی، روزانه)، میان‌مدت (هفتگی یا ماهانه) و بلندمدت (سالانه) و از منظر حوزه پیش‌بینی نیز به چهار حوزه کشوری، منطقه‌ای، شهری و خانگی تقسیم شدند. در جدول ۱ به طور خلاصه، مهم‌ترین تحقیقات این حوزه به تفکیک روش به کار رفته، افق زمانی و حوزه پیش‌بینی از سال ۲۰۰۴ تا ۲۰۱۷ ارائه شده است. بر اساس این جدول، مدل‌های شبکه عصبی متداول‌ترین و بیشترین روش

1. Sebalj et al.

مورد استفاده محققان در پیش بینی مصرف گاز بوده همچنین افق پیش بینی روزانه و سالیانه و همچنین حوزه پیش بینی کشوری به ترتیب بیشترین تمرکز محققان را در برداشته است.

جدول ۱۱. مروری بر پیشینه تحقیق حوزه پیش بینی مصرف گاز

نوع مدل					حوزه پیش بینی					افق پیش بینی			
کشور	منطقه	شهر	خانه	سال	فصل	ماه	هفته	روز	ساعت				
۲	۲	۴	۱	۱			۱	۳	۲				
۲							۱	۱					
۴				۴									
۲		۲	۱	۱				۳					
۱		۲		۱	۱	۲							
۴	۲	۲	۱	۳	۱		۱	۵	۱				
۱	۲	۱				۱	۱	۲					
۲						۱		۱					
		۱											

منبع: سبالج و همکاران (۲۰۱۷)

علاوه بر تحقیقات ذکر شده در جدول ۱۱، در سال های اخیر پژوهش هایی جهت توسعه مدل زنجیره مارکوف خاکستری و یا استفاده از آن برای پیش بینی تقاضای انرژی انجام شده است. لی^۴ و همکاران (۲۰۱۳) با ترکیب روش مارکوف خاکستری، مدل رگرسیون و روش تقریب تیلور، مدلی با عنوان "نسخه بهبود یافته روش مارکوف خاکستری" ارائه کردند و بر اساس داده های تولید برق حرارتی ۴۰ سال اخیر در ژاپن بهبود مدل را اعتبارسنجی کردند. سان و همکاران (۲۰۱۷) مدل مارکوف خاکستری مبتنی بر تبدیل موجک^۱ را جهت حذف نویزهای مربوط به داده های اصلی ارائه نموده و

1. Adaptiv Network-based Fuzzy Interference System
2. Support Vector Regression
3. Support Vector Machine
4. Guo-Dong Li
1. Wavelet Transform

با استفاده از آن پیش‌بینی تولید و مصرف انرژی در چین را برای سال‌های ۲۰۱۵ تا ۲۰۲۰ انجام دادند. کاظمی^۱ و همکاران (۲۰۱۲) پیش‌بینی تقاضای انرژی بخش حمل‌ونقل در ایران را با استفاده از مدل زنجیره مارکوف خاکستری انجام دادند. آنها نتایج پیش‌بینی را با مدل خاکستری و مدل رگرسیونی مقایسه کردند و همچنین تقاضای انرژی بخش حمل و نقل در ایران را تا سال ۱۴۰۰ با استفاده از مدل زنجیره مارکوف خاکستری پیش‌بینی و نشان دادند که مدل زنجیره مارکوف خاکستری دقت بالاتری از دو مدل خاکستری و مدل رگرسیونی دارد.

۳- روش تحقیق

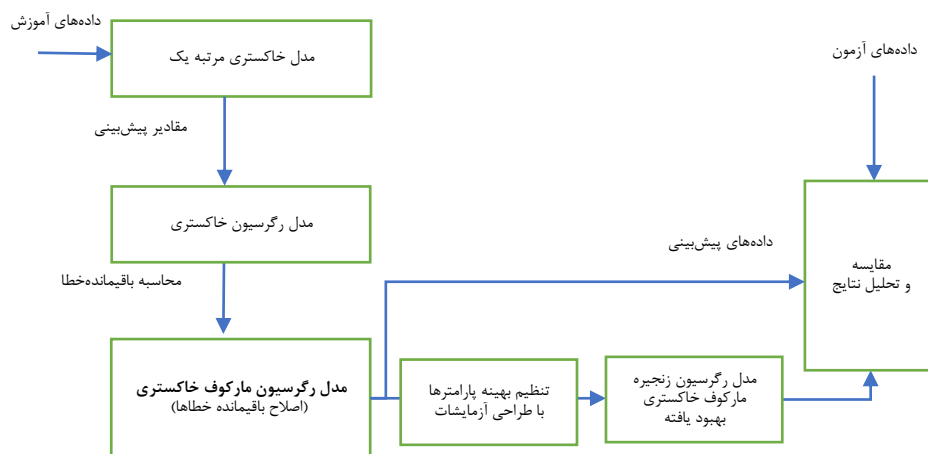
جهت توسعه روش‌هایی برای پیش‌بینی مصرف سالانه گاز در کشور، داده‌های مصرف کل گاز ایران (مجموع مصارف خانگی، تجاری، صنایع عمده و نیروگاهی) طی سال‌های ۱۳۸۰ الی ۱۳۹۷ (۱۸ سال) از مجموعه گزارش‌های سالیانه عملکرد شرکت ملی گاز ایران مربوط به مدیریت برنامه‌ریزی تلفیقی شرکت ملی گاز گردآوری شد. به‌منظور ساخت مدل‌های پیش‌بینی و ارزیابی دقت آنها، داده‌ها به دو دسته داده‌های آموزش (داده‌های سال‌های ۱۳۸۰ الی ۱۳۹۴) و داده‌های آزمون (داده‌های ۱۳۹۵ الی ۱۳۹۷) تقسیم شده‌اند. داده‌های اولیه آموزش $(x^{(0)}(t) | t=0, 1, \dots, 15)$ در جدول ۲ و کل مشاهدات اولیه در شکل ۱ نشان داده شده است. فرایند انجام پژوهش در شکل ۱ نشان داده شده است.

همان‌طور که در شکل ۱ نشان داده شده، برای پیش‌بینی مصرف گاز از چهار روش مدل $GM(1,1)$ ، $RGM(1,1)$ ، $MC-RGM(1,1)$ و مدل رگرسیون زنجیره مارکوف خاکستری بهبودیافته (مدل بهینه‌شده تنظیم پارامترهای مدل رگرسیون زنجیره مارکوف خاکستری) استفاده و در نهایت دقت مدل‌ها مذکور با داده‌های آموزش و آزمون بررسی شده‌اند.

پیش‌بینی مدل خاکستری $GM(1,1)$ و رگرسیون خاکستری $RGM(1,1)$

داده‌های ستون $x^{(0)}(t)$ در جدول را به‌عنوان داده‌های بردار $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ به ازای $n=15$ در مدل خاکستری $GM(1,1)$ در نظر بگیرید. به کمک روابط (۲) و (۳) و داده‌های بردار $x^{(0)}$ به ترتیب درآیه‌های بردارهای $x^{(1)}$ و $z^{(1)}$ محاسبه می‌شوند. به کمک رابطه (۶)، مقدار ضرایب a و b به ترتیب برابر -0.0531248 و $84100/956$ در مدل $GM(1,1)$ برآورد شده‌اند. با استفاده از ضرایب برآورد شده و داده‌های آموزش $x^{(0)}(t)$ و استفاده از رابطه (۱۰)، مقادیر پیش‌بینی هر دوره با مدل خاکستری $GM(1,1)$ محاسبه و مقادیر مربوطه $\hat{x}^{(0)}(t)$ در ستون سوم جدول ۲ درج شده است.

جهت کاهش مقدار خطا در مدل خاکستری، مدل رگرسیون خطی درجه اول بین داده‌های دو ستون $x^{(0)}(t)$ و $\hat{x}^{(0)}(t)$ بر اساس رویکرد مدل رگرسیون خاکستری رابطه (۱۱) تشکیل شده است. رابطه مذکور با بهره‌گیری از نرم‌افزار متلب و ضریب تعیین (R^2) برابر $94/2$ درصد به شکل $x^{(0)}(t) = 1014 + 9935/0 \cdot \hat{x}^{(0)}(t)$ تعیین شده است. پس از جایگذاری مقادیر پیش‌بینی مدل خاکستری $\hat{x}^{(0)}(t)$ در رابطه رگرسیونی، مقادیر پیش‌بینی به روش رگرسیون خاکستری محاسبه شده که در ستون $\hat{\hat{x}}^{(0)}(t)$ در جدول ۲ نشان داده شده است.



شکل ۱. فرایند انجام تحقیق

منبع: یافته‌های تحقیق

پیش‌بینی مدل رگرسیون زنجیره مارکوف خاکستری MC-RGM(1,1)

به‌طور مشابه، مراحل مدل رگرسیون زنجیره مارکوف خاکستری که قبلاً تشریح شد را بر داده‌های آموزش در جدول ۲ اجرا شده است. در اولین مرحله مقادیر خطاها $(e(t))$ با رابطه (۱۲) و از اختلاف $\hat{x}^{(0)}(t)$ و $x^{(0)}(t)$ حاصل می‌شود که نتایج محاسبات آن در جدول ۲ نشان داده شده است. با توجه به تعداد مشاهدات $(n=15)$ ، فاصله بیشینه و کمینه خطاها به سه بازه مساوی $(r=3)$ و بر اساس رابطه (۱۴) به شرح اعداد خاکستری زیر تقسیم می‌شوند. هر عدد خاکستری بیانگر یک وضعیت در زنجیره مارکوف است. وضعیت هر یک از خطاها در جدول ۲ مشخص شده است.

$$\otimes s_1 = [-10518.35578 \quad -3324.071507]$$

$$\otimes s_2 = [-3324.071507 \quad 3870.212766]$$

$$\otimes s_3 = [3870.212766 \quad 11064.49704]$$

اگر هر یک از اعداد خاکستری فوق بیانگر یک وضعیت خطا باشد، ماتریس احتمال گذار زنجیره مارکوف مربوطه در رابطه (۱۵)، به شرح زیر خواهد بود. درآیه‌های این ماتریس بر اساس رابطه (۱۶) و مقدار خطاها $(e(t))$ محاسبه شده است.

$$P = \begin{bmatrix} 0.666666667 & 0.333333333 & 0 \\ 0 & 0.333333333 & 0.666666667 \\ 0.4 & 0 & 0.6 \end{bmatrix}$$

در روش رگرسیون زنجیره مارکوف خاکستری، در رابطه (۱۷) عموماً از ضریب ۰/۵ (میانگین حد پایین و بالای اعداد خاکستری خطاها) برای سفیدسازی تمام اعداد خاکستری استفاده می‌شود. بردار مرکزی خطاها با استفاده از ضریب ۰/۵ برابر است با:

$$V = (0.571/6885-, 6513/228, 3597/73421)$$

مقدار تعدیل و پیش‌بینی یک دوره بعد از هر دوره جاری با روابط (۱۸) و (۱۹) محاسبه شده و در جدول ۲ مقادیر پیش‌بینی مدل رگرسیون زنجیره مارکوف خاکستری برای همه مشاهدات آزمایش، نشان داده شده است.

جدول ۲. نتایج پیش‌بینی مدل‌های خاکستری، رگرسیون خاکستری و رگرسیون مارکوف

خاکستری

سال	t	مقادیر واقعی $(x^{(0)}(t))$	مدل خاکستری $(\hat{x}^{(0)}(t))$	مدل رگرسیون خاکستری $(\hat{\hat{x}}^{(0)}(t))$	باقیمانده خطاها	وضعیت خطا	رگرسیون زنجیره مارکوف خاکستری $\tilde{x}^{(0)}(t)$ به ازای $\alpha_j = 0.5$	رگرسیون زنجیره مارکوف خاکستری بهینه
۱۳۸۰	۰	۹۰۳۲۰/۵	۸۶۵۷۹/۱۰	۸۷۰۲۷/۳۶۶	۳۲۹۳/۱۳	۲	۸۷۳۰۰/۴۳۶۷۳	۹۰۸۹۷/۵۷
۱۳۸۱	۱	۱۰۱۶۰۴/۲	۹۱۳۰۲/۹۷	۹۱۷۲۰/۳۵۶	۹۸۸۳/۸۴	۳	۹۶۷۸۹/۶۱۷۰۳	۱۰۰۳۸۶/۷
۱۳۸۲	۲	۸۸۵۱۶/۷۲	۹۶۲۸۴/۵۷	۹۶۶۶۹/۴۰۳	-	۱	۹۸۳۸۱/۳۳۰۶۹	۹۹۱۰۰/۷۵
۱۳۸۳	۳	۹۱۳۷۰/۱۲	۱۰۱۵۳۷/۹	۱۰۱۸۸۸/۴۷	-	۱	۹۷۳۶۵/۳۵۶۸۹	۹۶۱۶۶/۳۰
۱۳۸۴	۴	۹۷۷۱۱	۱۰۷۰۷۸/۰	۱۰۷۳۹۲/۳۰	-	۱	۱۰۲۸۶۹/۱۸۸۷	۱۰۱۶۷۰/۱
۱۳۸۵	۵	۱۰۹۱۰۶/۷	۱۱۲۹۲۰/۳	۱۱۳۱۹۶/۴۳	-	۱	۱۰۸۶۷۳/۳۱۶۶	۱۰۷۴۷۴/۳
۱۳۸۶	۶	/۸۲	۱۱۹۰۸۱/۴	۱۱۹۳۱۷/۲۴	۲۳۵۷/۵۷	۲	۱۱۴۷۹۴/۱۲۵	۱۱۳۵۹۵/۰
۱۳۸۷	۷	۱۳۲۶۵۰/۴	۱۲۵۵۷۸/۶	۱۲۵۷۷۲/۰۱	۶۸۷۸/۳۸	۳	۱۳۰۸۴۱/۲۷۱۶	۱۳۴۴۳۸/۴
۱۳۸۸	۸	/۰۱	۱۳۲۴۳۰/۳	۱۳۲۵۷۸/۹۵	۸۶۵۶/۰۵	۳	۱۳۴۲۹۰/۸۸۶۸	۱۳۵۰۱۰/۳
۱۳۸۹	۹	۱۵۰۸۲۱/۸	۱۳۹۶۵۵/۹	۱۳۹۷۵۷/۳۰	۱۱۰۶۴/۴	۳	۱۴۱۴۶۹/۲۳۰۴	۱۴۲۱۸۸/۶
۱۳۹۰	۱۰	/۱۵	۱۴۷۲۷۵/۷	۱۴۷۳۲۷/۳۰	۵۴۰۵/۸۴	۳	۱۴۹۰۳۹/۲۳۳۶	۱۴۹۷۵۸/۶
۱۳۹۱	۱۱	/۴۴	۱۵۵۳۱۱/۳	۱۵۵۳۱۰/۳۳	-	۱	۱۵۷۰۲۲/۲۶۵۷	۱۵۷۷۴۱/۶
۱۳۹۲	۱۲	/۶۱	۱۶۳۷۸۵/۳	۱۶۳۷۲۸/۹۳	-	۱	۱۵۹۲۰۵/۸۱۵۹	۱۵۸۰۰۶/۷
۱۳۹۳	۱۳	/۰۱	۱۷۲۷۲۱/۶	۱۷۲۶۰۶/۸۶	-	۲	۱۶۸۰۸۳/۷۴۱۸	۱۶۶۸۸۴/۶
۱۳۹۴	۱۴	۱۸۰۷۷۰	۱۸۲۱۴۵/۵	۱۸۱۹۶۹/۱۷	-	۲	۱۸۷۰۳۸/۴۳۷۶	۱۹۰۶۳۵/۵

منبع: یافته‌های تحقیق

۴- بهبود مدل رگرسیون زنجیره مارکوف خاکستری

در رویکرد رگرسیون زنجیره مارکوف خاکستری برای اصلاح مقادیر پیش‌بینی حاصل از روش رگرسیون خاکستری با احتمالی که از ماتریس گذار مارکوف محاسبه می‌شود، میانگین خطا در هر وضعیت (با ضرایب α برابر با ۰/۵) به مقدار پیش‌بینی اضافه می‌شود. به نظر می‌رسد، انتخاب میانگین فواصل در اعداد خاکستری خطاها معیار مناسبی برای اصلاح مقادیر پیش‌بینی در رابطه (۱۸) نیست. استفاده صرف از مرکز وضعیت خطا موجب می‌شود سایر اطلاعات توزیع خطاها را نادیده گرفته شود و به

همین دلیل مرکز بردار خطاها نمی‌تواند نماینده مناسبی برای وضعیت خطاها باشد (سان و همکاران، ۲۰۱۵). الگوی منطقی برای انتخاب میانگین بازه‌ها وجود ندارد و مقدار ۰/۵ برای سفیدسازی بازه‌ها به صورت قراردادی است. در واقع مسئله این پژوهش تعیین مقدار بهینه متغیرهای α_j است به‌طوری‌که مدل رگرسیون زنجیره مارکوف خاکستری برازش بهتری برای داده‌های آموزش انجام دهد. پیشنهاد این مقاله، تعیین مقادیر بهینه این پارامترها استفاده از رویکرد طراحی آزمایش‌ها^۱ و رویه پاسخ است. طراحی آزمایشات یک رویکرد آماری برای بررسی اثر متغیرهایی مانند α_j ها بر رفتار متغیرهای پاسخ یا وابسته^۲ است. در مسئله مذکور چنانچه r متغیر α_j در مدل رگرسیون زنجیره مارکوف خاکستری، متغیر مؤثر بر میزان خطای پیش‌بینی (به‌عنوان متغیر پاسخ) در نظر گرفته شوند، رویکرد طراحی آزمایشات قادر خواهد بود اثر سطوح مختلف متغیرهای اثرگذار را بر متغیر پاسخ بررسی و مقدار بهینه‌ای برای آنها پیشنهاد نماید. در این روش، ابتدا هر یک از متغیرهای اثرگذار به تعدادی سطح (عموماً دو سطح حداقلی و حداکثری) تقسیم می‌شوند. هر یک از ترکیبات سطوح مختلف متغیرها یک آزمایش نامیده می‌شود. می‌توان پیش‌بینی مصرف سالیانه گاز با مدل رگرسیون زنجیره مارکوف خاکستری را به ازای سطوح مختلف آزمایش (منظور مقادیر مختلف آزمایش به‌ازای α_j) و اعمال ضرایب در رابطه (۱۷) انجام داده و مقدار خطای کل پیش‌بینی را (مانند میانگین انحراف مطلق یا میانگین درصد انحراف مطلق^۳) به‌عنوان متغیر پاسخ برای داده‌های آموزش محاسبه کرد. در واقع فضای جواب نامحدود برای تعیین ضرایب سفیدسازی با سطح‌بندی آنها و انجام آزمایش به ازای هر آزمایش، به فضای جواب محدودتر و گسسته محدود می‌شود. روش رویه پاسخ به دنبال برازش تابع متغیر پاسخ بر حسب متغیرهای اثرگذار و تعیین مقدار بهینه آنها برای کمینه یا بیشینه کردن متغیر پاسخ است.

هر یک از سه متغیر α_1 ، α_2 و α_3 در مطالعه موردی این پژوهش به سه سطح یک (بالاترین مقدار ممکن)، ۰/۵ (مقدار متوسط به‌عنوان مقدار میانی) و صفر (کمترین مقدار ممکن) تقسیم شده‌اند. به‌منظور طراحی آزمایش‌ها (ترکیب سطوح مختلف

1. Design of Exprimtent
2. Response or Dependent
3. Mean Absolute Percentage Error

متغیرها) از طرح باکس- بنکن^۱ سه سطحی از جمله روش‌های تعیین آزمایش‌ها استفاده شده است. در صورت عدم استفاده از طرح باکس-بنکن با توجه به این که هر سه متغیر سه سطح دارند، در مجموع $3^3=27$ آزمایش قابل طرح است. اما استفاده از این طرح تعداد آزمایش‌های را به ۱۵ آزمایش کاهش می‌دهد.

به‌عنوان مثال یکی از ۱۵ آزمایش پیشنهادی، ترکیب $(\alpha_1 = 1, \alpha_2 = 0.5, \alpha_3 = 0)$ بوده است. به ازای هر یک از آزمایش‌های پیشنهادی طرح باکس-بنکن، مدل پیش‌بینی رگرسیون زنجیره مارکوف خاکستری بر داده‌های آموزشی اجرا شده و مقدار درصد انحراف مطلق (MAPE) مطابق (۲۰) به ازای هر آزمایش محاسبه شده است.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{x^{(0)}(t) - \hat{x}^{(0)}(t)}{x^{(0)}(t)} \right| \times 100 \% \quad (20)$$

از داده‌های اجرای ۱۵ آزمایش شامل سطوح سه متغیر اثرگذار و درصد انحراف مطلق پیش‌بینی به‌عنوان مقدار پاسخ هر آزمایش برای تشکیل رویه پاسخ یا معادله رگرسیونی MAPE و α_j استفاده می‌شود. در مدل باکس-بنکن امکان تشکیل یک مدل درجه دو کامل^۲ به شکل رابطه (۲۱) وجود دارد که در آن علاوه بر اثرات متغیرهای اصلی (α_j) ، مربع متغیرهای اصلی (α_j^2) اثرات برهمکنش‌های^۳ دوتایی $(\alpha_i \times \alpha_j)$ متغیرهای اصلی نیز سنجش می‌شود.

$$MAPE = \beta_0 + \beta_1 \times \alpha_1 + \beta_2 \times \alpha_2 + \beta_3 \times \alpha_3 + \beta_{11} \times \alpha_1^2 + \beta_{22} \times \alpha_2^2 + \beta_{33} \times \alpha_3^2 + \beta_{12} \times \alpha_1 \times \alpha_2 + \beta_{13} \times \alpha_1 \times \alpha_3 + \beta_{23} \times \alpha_2 \times \alpha_3 \quad (21)$$

در رابطه فوق، β_k ها ضرایب مدل رگرسیونی یا رویه پاسخ است که بیانگر ضرایب تأثیر هر یک متغیرها یا اثر برهمکنش دوتایی آنها است.

در جدول ۳، بخشی از نتایج برازش مدل رگرسیونی (۲۱) شامل برآورد ضرایب و آزمون معنی‌داری آنها نشان داده شده است. در این جدول تنها ضرایب معنی‌دار و تأثیرگذار در مدل رویه پاسخ نشان داده شده است. مقدار ضریب تعیین مدل رگرسیون ۹۹/۸۷ درصد برآورد شده و نتایج تحلیل واریانس مدل (سطح معنی‌داری مدل برابر ۰/۰۰۰۰) نیز بیانگر برازش مناسب مدل بر داده‌های جمع‌آوری شده است. مقدار سطح معنی‌داری (P-value) تمام ضرایب ذکر شده در جدول ۳، کوچک‌تر از ۰/۰۵ است که

1. Box-Behnken
2. Full quadratic
3. Interactions

نتیجه آن تأثیر تمام متغیرهای α_j در رابطه (۲۱) بر مقدار MAPE است. به عبارتی دیگر می‌توان نتیجه گرفت مقدار هر سه متغیر α_1 ، α_2 و α_3 به دلیل معنادار بودن باید در معادله رویه سطح قرار داشته باشند. همچنین جملات درجه دوم α_1^2 ، α_2^2 و α_3^2 و جملات برهمکنش هر دو متغیر (α_1 و α_2) و (α_2 و α_3) نیز بر مقدار خطای پیش‌بینی مؤثر است.

از دیگر نتایج جدول ۳، تأثیر منفی متغیرهای α_1 و تأثیر مثبت متغیر α_2 و α_3 بر متغیر پاسخ است. جهت کمینه‌کردن متغیر پاسخ (درصد انحراف مطلق) نیاز است که متغیر α_1 افزایش یافته و کاهش متغیرهای α_2 و α_3 موجب بهترشدن تابع هدف می‌شود.

جدول ۳. نتایج مدل رگرسیونی رویه پاسخ

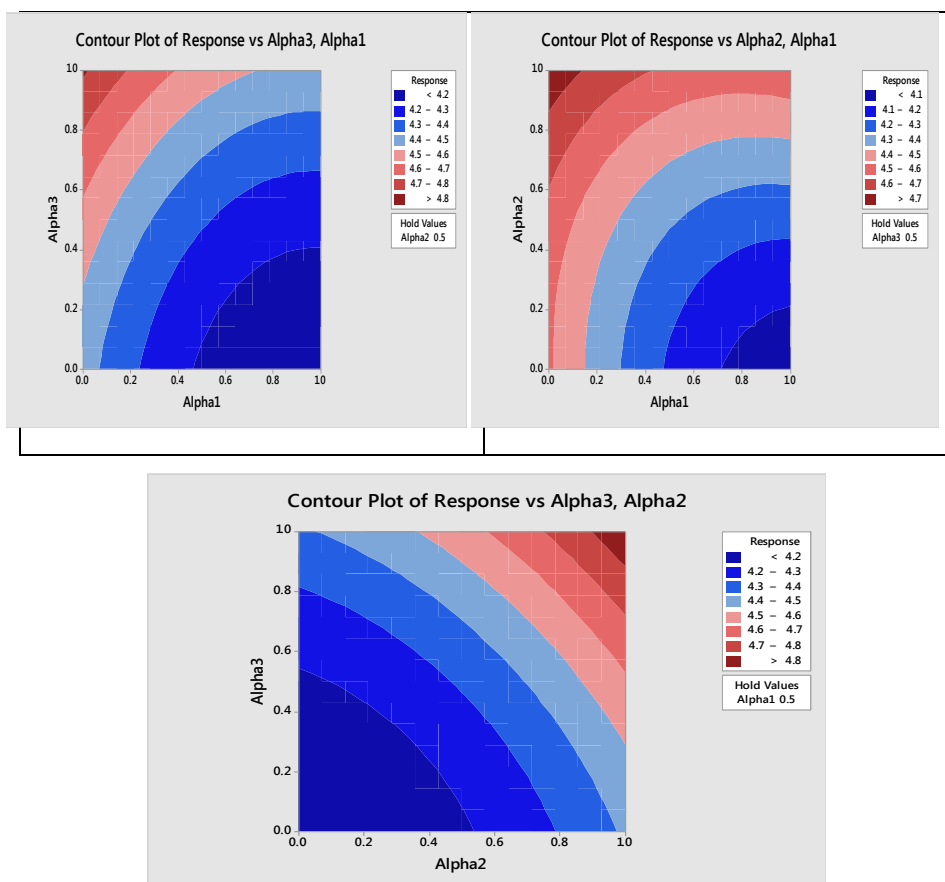
عبارت	ضریب	مقدار تأثیر	مقدار برآورد	انحراف معیار	T-Value	P-Value
ضریب ثابت	β_0		۴/۳۱۳۱	۰/۰۰۸۰۰	۵۳۹/۴۴	۰/۰۰۰
ضریب α_1	β_1	-۰/۳۳۷۷۶	-۰/۱۶۸۸۸	۰/۰۰۴۹۰	-۳۴/۴۹	۰/۰۰۰
ضریب α_2	β_2	۰/۳۹۸۷۶	۰/۱۹۹۳۸	۰/۰۰۴۹۰	۴۰/۷۲	۰/۰۰۰
ضریب α_3	β_3	۰/۳۷۲۴۰	۰/۱۸۶۲۰	۰/۰۰۴۹۰	۳۸/۰۳	۰/۰۰۰
ضریب α_1^2	β_{11}	۰/۱۷۷۹۸	۰/۰۸۸۹۹	۰/۰۰۷۲۱	۱۲/۳۵	۰/۰۰۰
ضریب α_2^2	β_{22}	۰/۱۴۷۳۰	۰/۰۷۳۶۵	۰/۰۰۷۲۱	۱۰/۲۲	۰/۰۰۰
ضریب α_3^2	β_{33}	۰/۱۲۱۳۱	۰/۰۶۰۶۶	۰/۰۰۷۲۱	۸/۴۲	۰/۰۰۰
برهمکنش α_2 و α_1	β_{12}	۰/۱۴۶۵۳	۰/۰۷۳۲۶	۰/۰۰۶۹۲	۱۰/۵۸	۰/۰۰۰
برهمکنش α_3 و α_2	β_{13}	۰/۰۸۹۸۶	۰/۰۴۴۹۳	۰/۰۰۶۹۲	۶/۴۹	۰/۰۰۱

منبع: یافته‌های تحقیق

جهت درک بهتر تأثیر تغییرات متغیرها بر متغیر پاسخ، نمودارهای کانتور^۱ تغییرات متغیرها بر متغیر پاسخ در شکل ۲ رسم شده است. همان‌طور که از در نمودارهای کانتور و جدول راهنمای سمت راست آنها مشخص است، هرچه از منطقه قرمز رنگ به سمت منطقه آبی حرکت شود مقدار پاسخ یا تابع هدف کاهش خواهد یافت. بر این اساس با افزایش متغیر α_1 و کاهش متغیر α_2 مقدار تابع هدف کاهش می‌یابد و به همین ترتیب

1. Contour plot

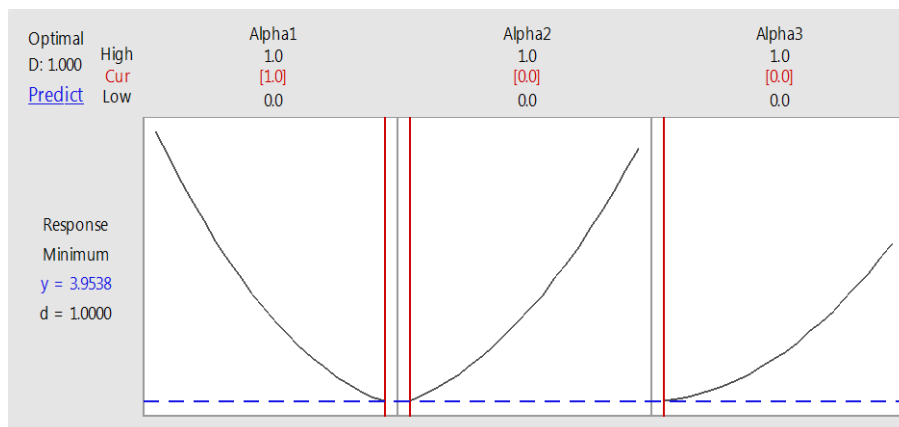
با افزایش فاکتور α_1 و کاهش α_3 مقدار تابع هدف کاهش می‌یابد و در نهایت با کاهش α_2 و α_3 مقدار تابع هدف کاهش می‌یابد.



شکل ۲. خطوط کانتور تغییرات متغیرها بر متغیر پاسخ

منبع: یافته‌های تحقیق

هدف تعیین مقدار بهینه متغیرهای α_j ترسیم شده است. آنچه در این نمودار مشخص است، کمترین مقدار MAPE برابر ۳/۹۵۳۸ تعیین شده است. این مقدار بهینه تابع هدف به ازای α_1 برابر یک (حد بالای تغییرات مجاز آن)، α_2 برابر صفر (حد پایین تغییرات مجاز آن) و α_3 برابر صفر (حد پایین تغییرات مجاز آن) حاصل می‌شود.



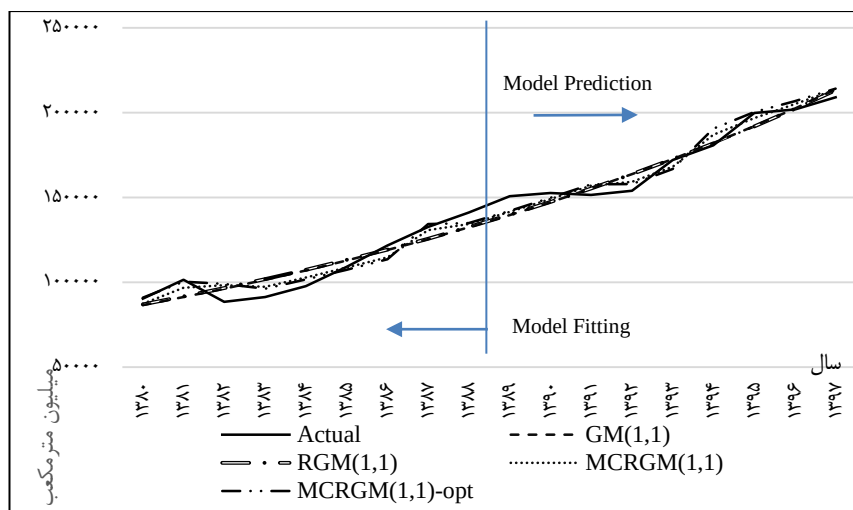
شکل ۳. نمودار تعیین مقدار بهینه متغیرها مبتنی بر داده‌های آموزش

منبع: یافته‌های تحقیق

اکنون که مقدار بهینه ضرایب سفیدسازی در رابطه (۱۷) تعیین شده، مجدداً مدل رگرسیون زنجیره مارکوف خاکستری با ضرایب بهینه شده (به جای ضرایب یکسان ۰/۵ در مدل سنتی این روش) اجرا و مقادیر پیش‌بینی مجدد در ستون آخر جدول ۲ ارائه شده است.

۵- مقایسه عملکرد مدل‌های پیش‌بینی

با استفاده از روابط چهار مدل این مقاله شامل مدل خاکستری، رگرسیون خاکستری، رگرسیون زنجیره مارکوف خاکستری و رگرسیون زنجیره مارکوف خاکستری بهبودیافته (مدل پیشنهادی این مقاله)، مقدار مصرف گاز طی سه سال بعد از آخرین دوره داده‌های آموزش یعنی سال‌های ۱۳۹۵ تا ۱۳۹۷ به‌عنوان داده آزمون (چون مقادیر واقعی این سال‌ها مشخص و از داده‌های این دوره‌ها برای ساخت مدل‌ها استفاده نشده است) پیش‌بینی شده است. در شکل ۴ مقادیر واقعی و پیش‌بینی مصرف گاز در کل دوره‌های مورد مطالعه (به تفکیک داده‌های آموزش و داده‌های آزمون) با هر یک از مدل‌های مذکور نشان داده شده است.



شکل ۴. منحنی روند تغییر مقادیر واقعی و پیش‌بینی مصرف سالیانه گاز

منبع: یافته‌های تحقیق

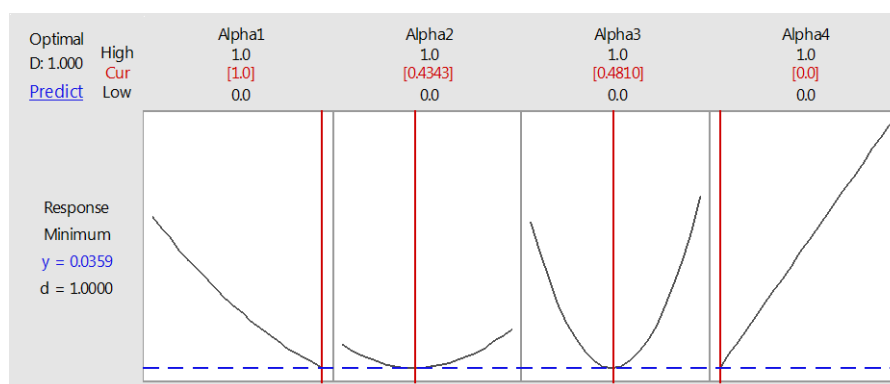
به منظور اندازه‌گیری و نمایش میزان خطای پیش‌بینی از شاخص میانگین درصد انحراف مطلق (MAPE) استفاده شده است. در جدول ۴، مقدار شاخص دقت پیش‌بینی مدل‌های مختلف به تفکیک داده‌های آموزش و آزمون ارائه شده است. همان‌طور که از نتایج این جدول مشخص است، خطای مدل رگرسیون خاکستری نسبت به مدل خاکستری کمتر و در نتیجه روش رگرسیون و مشخصاً تکنیک کمینه‌سازی مربعات خطای باقیمانده، موجب عملکرد بهتر مدل شد. همچنین بهره‌گیری از خواص مارکوفی مربوط به باقیمانده خطاها و محاسبات ماتریس احتمال انتقال، موجب تصحیح میزان خطاها شد که نشان‌دهنده عملکرد بهتر مدل رگرسیون زنجیره مارکوف خاکستری نسبت به مدل رگرسیون خاکستری می‌باشد و در نهایت بهره‌گیری از رویکرد طراحی آزمایشات و تکنیک رویه پاسخ جهت تعیین دقیق مقدار بهینه ضرائب سفیدسازی مدل، موجب گردید که مدل رگرسیون زنجیره مارکوف خاکستری بهبودیافته، نسبت به سایر مدل‌ها از خطای کمتر و از دقت بالاتری برخوردار شود.

جدول ۴. مقایسه دقت پیش‌بینی مدل‌های پیش‌بینی

مدل‌ها	میانگین درصد انحراف مطلق داده‌های آموزش	میانگین درصد انحراف مطلق داده‌های آزمون
خاکستری	۵/۴۵۵۱۳۶۰۵۱	۲/۱۵۵۸۷۸۹۷
رگرسیون خاکستری	۵/۴۳۷۷۱۷۵۶۴	۲/۰۸۴۰۲۴۸۳۵
رگرسیون زنجیره مارکوف خاکستری	۴/۳۱۳۳۰۷۶۶	۱/۷۷۳۵۹۱۶۱۸
رگرسیون زنجیره مارکوف خاکستری بهینه شده پیشنهادی	۳/۹۸۱۴۵۸۲۷۶	۱/۷۳۸۱۳۰۰۷۱

منبع: یافته‌های تحقیق

پس از اعتبارسنجی مدل پیشنهادی، در جدول ۶ مقادیر پیش‌بینی مصرف سالیانه کشور بر اساس داده‌های مصرف ۱۸ سال و به کمک مدل خاکستری، مدل رگرسیون خاکستری و رگرسیون زنجیره مارکوف خاکستری پیشنهادی با α بهینه محاسبه شده است. از آنجا که تعداد داده‌های مصرف اولیه ۱۸ می‌باشد، تعداد اعداد خاکستری خطاها چهار ($t=4$) خواهد بود و مجدداً از طرح باکس-بنکن استفاده شده تا مقدار بهینه α ها محاسبه شود. در شکل ۵ تغییرات میانگین انحراف مطلق خطاها بر اساس تغییرات ضرایب سفیدسازی رسم شده است. با تحلیل این نمودار مقدار بهینه α_1 برابر ۱، α_2 برابر ۰/۴۳۴۳۴۳، α_3 برابر ۰/۴۸۰۹۷۹ و α_4 برابر صفر و مقدار بهینه خطا نیز برابر ۰/۰۳۵۹ خواهد شد.



شکل ۵. نمودار بهینه‌سازی مبتنی بر داده‌های ۱۸ سال

منبع: یافته‌های تحقیق

در نهایت از روش پیشنهادی این مقاله و با ضرایب بهینه تعیین شده، پیش‌بینی مصرف سالیانه گاز سه سال آتی (۱۳۹۸ الی ۱۴۰۰) انجام شده است. نتایج برآوردهای این روش نیز در جدول ۶ ارائه شده است. درصد انحراف مطلق خطاها در مدل خاکستری، رگرسیون خاکستری، رگرسیون زنجیره مارکوف خاکستری و مدل رگرسیون زنجیره مارکوف خاکستری بهبود داده شده برای داده‌های ۱۸ سال به ترتیب ۰/۰۵۱۱۹۰۸۰۱، ۰/۰۴۹۱۳۷۵۶۷، ۰/۰۳۷۵۴۵۸۸۶ و ۰/۰۳۵۸۱۶۲۰۹ می‌باشد.

جدول ۵. پیش‌بینی سه ساله مصرف گاز کشور بر اساس مدل رگرسیون زنجیره مارکوف خاکستری بهینه و داده‌های ۱۸ سال (۱۳۸۰-۱۳۹۷)

t	$x^{(0)}(t)$ (میلیون مترمکعب)	مدل خاکستری	مدل رگرسیون خاکستری	رگرسیون زنجیره مارکوف خاکستری بهینه
۱۳۸۰	۹۰۳۲۰۵	۸۵۷۴۶۹۱۵	۸۷۲۸۱۴۷۰۸	۹۰۲۷۸۶۸
۱۳۸۱	۱۰۱۶۰۴۲	۹۰۵۵۴۳۰۹۱	۹۱۹۲۷۱۱۴۹	۹۸۵۵۰۱۰
۱۳۸۲	۸۸۵۱۶۷۲	۹۵۶۳۱۲۲۹۴	۹۶۸۳۳۲۱۶۶	۹۹۲۷۷۶۰
۱۳۸۳	۹۱۳۷۰۱۲	۱۰۰۹۹۲۷۸۷	۱۰۲۰۱۴۳۷۸	۹۵۶۰۹۴۴
۱۳۸۴	۹۷۷۱۱	۱۰۶۶۵۴۹۳۹	۱۰۷۴۸۶۰۲۲	۱۰۱۰۸۱۰۰۸
۱۳۸۵	۱۰۹۱۰۶۷	۱۱۲۶۳۴۵۴	۱۱۳۲۶۴۴۳۲	۱۰۶۸۵۹۴۹
۱۳۸۶	۱۲۱۶۷۴۸۲	۱۱۸۹۴۹۳۸۷	۱۱۹۳۶۶۸۰۹	۱۱۹۱۷۴۶۱
۱۳۸۷	۱۳۲۶۵۰۴	۱۲۵۶۱۸۲۷۵	۱۲۵۸۱۱۳۱۵	۱۳۳۴۳۴۳۰
۱۳۸۸	۱۴۱۲۳۵۰۱	۱۳۲۶۶۱۰۵۴	۱۳۲۶۱۷۱۳۲	۱۳۵۰۶۱۵۱
۱۳۸۹	۱۵۰۸۲۱۸	۱۴۰۰۹۸۶۸۵	۱۳۹۸۰۴۵۱۶	۱۴۲۲۴۸۸۹
۱۳۹۰	۱۵۲۷۳۳۱۵	۱۴۷۹۵۳۳۰۷	۱۴۷۳۹۴۸۶۱	۱۴۹۸۳۹۲۴
۱۳۹۱	۱۵۱۶۱۸۴۴	۱۵۶۲۴۸۲۹۸	۱۵۵۴۱۰۷۵۷	۱۶۲۰۳۳۷۴
۱۳۹۲	۱۵۴۰۶۶۶۱	۱۶۵۰۰۸۳۴۶	۱۶۳۸۷۶۰۶۴	۱۶۳۶۸۳۸۶
۱۳۹۳	۱۷۲۰۶۳۰۱	۱۷۴۲۵۹۵۲۶	۱۷۲۸۱۵۹۷۸	۱۶۶۴۱۱۰۴
۱۳۹۴	۱۸۰۷۷۰	۱۸۴۰۲۹۳۷۳	۱۸۲۲۵۷۱۰۸	۱۸۲۰۶۴۹۰
۱۳۹۵	۱۹۹۹۰۰	۱۹۴۳۴۶۹۶۵	۱۹۲۲۲۷۵۵۳	۱۹۲۰۳۵۳۵
۱۳۹۶	۲۰۱۸۴۸	۲۰۵۲۴۳۰۱۳	۲۰۲۷۵۶۹۹۱	۲۰۵۲۰۱۳۷
۱۳۹۷	۲۰۹۰۱۲	۲۱۶۷۵۰	۲۱۳۸۷۶۷۶۱	۲۱۳۶۸۴۵۶
۱۳۹۸		۲۲۵۶۱۹۹۶	۲۲۸۹۰۲۰۱۵	۲۲۵۴۲۷۷۶
۱۳۹۹		۲۳۸۰۲۱۵۴۱	۲۴۱۷۳۵۳۸۹	۲۳۸۴۷۷۱۴
۱۴۰۰		۲۵۱۱۱۸۴۱۵	۲۵۵۲۸۸۲۶۴	۲۵۱۱۵۰۶۳

منبع: یافته‌های تحقیق

۶- نتیجه‌گیری و پیشنهادها

در این تحقیق روشی مبتنی بر رویکرد طراحی آزمایشات جهت بهبود مدل پیش‌بینی رگرسیون زنجیره مارکوف خاکستری پیشنهاد شد. به این منظور در ابتدا با ترکیب مدل خاکستری با تکنیک رگرسیون، فاصله داده‌های واقعی را با مقادیر پیش‌بینی خاکستری کمینه کرده و جهت غلبه بر تأثیر خواص تصادفی و نوسانی داده‌ها، از مدل باقیمانده خطای اصلاح شده زنجیره مارکوف استفاده شد و با ترکیب آن با مدل رگرسیون خاکستری، مدل رگرسیون زنجیره مارکوف خاکستری پیاده سازی گردید. در ادامه، مقادیر بهینه ضرائب سفیدسازی مدل فوق با رویکرد طراحی آزمایشات و براساس روش باکس بنکن، با هدف کاهش معیار پاسخ "میزان درصد متوسط خطا" حاصل شد.

در تحقیق حاضر جهت اعتبارسنجی مدل پیشنهادی، داده‌های مصرف سالیانه گاز ایران مربوط به ۱۵ سال (۱۳۸۰ الی ۱۳۹۴) را به‌عنوان داده‌های مدل فرض کرده و پیش‌بینی ۳ سال (۱۳۹۵-۱۳۹۷) را براساس سه مدل خاکستری، رگرسیون خاکستری و رگرسیون زنجیره مارکوف خاکستری انجام و نتایج حاصله با داده‌های واقعی مربوط به این سه سال مقایسه شد.

مطالعه کاربردی ما در زمینه پیش‌بینی مصرف سالیانه گاز کشور به خوبی نشان می‌دهد که مدل پیشنهادی رگرسیون زنجیره مارکوف خاکستری (OP-MCRGM) دقت بالاتری نسبت به مدل‌های خاکستری GM، مدل RGM و مدل سنتی MCRGM دارد.

در این تحقیق از داده‌های مصرف سالیانه گاز که مجموع سه عدد، مصرف (خانگی، تجاری و صنایع غیرعمده)، مصرف صنایع عمده و مصرف نیروگاه‌ها می‌باشد، به‌عنوان داده‌های کاربردی استفاده شده است، لذا می‌توان مدل بهینه پیشنهادی این تحقیق را به‌صورت مجزا به ترتیب برای داده‌های سه بخش مصرف (خانگی، تجاری و صنایع غیرعمده)، مصرف صنایع عمده و مصرف نیروگاه‌ها انجام داده و با نتایج پیش‌بینی حاصل از مجموع مصرف مقایسه کرد. ثانیاً با توجه به منطق کلی بهینه‌سازی استفاده شده در این مقاله و نیاز به حداقل داده‌ها (۴ داده) جهت مدل‌سازی به این روش،

می‌توان حوزه کاربردی مورد استفاده در مقاله را به سایر حوزه‌های انرژی منجمله نفت و آب و برق نیز تعمیم داد.

پیشنهاد تحقیق آتی، مربوط به انتخاب رویکردهای بهینه‌سازی کلاسیک برنامه‌ریزی خطی (غیرخطی) جهت بهینه‌سازی ضرائب سفیدسازی می‌باشد. علیرغم اینکه تعیین تابع هدف بهینه‌سازی در این روش مستلزم محاسبات دستی طولانی می‌باشد، ولیکن می‌تواند با استفاده از تحلیل نتایج مربوطه، گام تأثیرگذاری در کمک به تصمیم‌گیرندگان حوزه انرژی کشور برداشت.

منابع

اعظمی، سمیه، الماسی، مجتبی، علی نیایی، زهرا (۱۳۹۷)، ارتباط میان مصرف گاز طبیعی و رشد اقتصادی در کشورهای صادرکننده گاز طبیعی با در نظر گرفتن شکست‌های ساختاری: کاربردی از آزمون علیت دومیتروش-هرلین، فصلنامه مطالعات اقتصاد انرژی، شماره ۵۷، صفحات ۳۹-۶۴.

فطرس، محمدحسن، مصطفی، امیدعلی، گلوانی، امیرمحمد (۱۳۹۶)، پیش‌بینی تراز داخلی گاز طبیعی: با استفاده از مدل ترکیبی ARDL و میانگین کتحرک خودهمبسته یکپارچه (ARIMA)، پژوهشنامه اقتصاد انرژی ایران، شماره ۲۵، صفحات ۹۵-۱۲۴.

ناجی‌میدانی، علی اکبر، رحیمی، غلامعلی، پاییز (۱۳۹۴)، تعیین اولویت بازارهای صادراتی گاز طبیعی ایران، فصلنامه مطالعات اقتصاد انرژی، شماره ۴۶، صفحات ۱۸۱-۲۱۹.

Cheng, C. S., Hsu, Y. T., & Wu, C. C. (1997). Grey modeling using neural network. *The Journal of Grey System*, 3, 219e232.

Deng JL. Control problem of grey systems. *Syst Control Lett* 1982;1(5):288-94.

Dario Sebalj, Josip Mesarić, Davor Dujak. (2017), Predicting Natural Gas Consumption – A Literature Review. *28th CECIIS*, 27-29, Varazdin, Croatia

Gang Xu., & Weiguo Wang. (2010). Forecasting China's natural gas consumption based on a combination model. *Journal of Natural Gas Chemistry*, 19(2010)493-496.

- Hsu, L.C. (2003). Applying the grey prediction model to the global integrated circuit industry. *Technological Forecasting and Social Change*, 70(6), 563-574.
- Hsu, Y. T., Liu, M. C., Yeh, J., & Hung, H. F. (2009). Forecasting the turning time of stock market based on Markov Fourier grey model. *Expert Systems with Applications*, 36(4), 8597e8603.
- Kazemi, A. Modarres, M. Mehregan, M. (2012), Energy Demand Forecast of Iran's Industrial sector Using Markov Chain Grey Model. *Part A: Intl. J. Humanities*, (2013) Vol. 20(1): (1-12).
- Kumar, U., & Jain, V. K. (2010). Time series models (Grey Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India. *Energy*, 35(4), 1709e1716.
- Li, G. D., Masuda, S., & Nagai, M. (2013). The prediction model for electrical power system using an improved hybrid optimization model. *Electrical Power and Energy Systems*, 44, 981e987
- MAO Zhan-li, SUN Jin-hua (2011). Application of Grey-Markov Model in Forecasting Fire Accidents. *Procedia Engineering* 11(2011) 314-318
- Mostafaei, H., Kordnoori, SH. (2012). Hybrid Grey Forecasting Model for Iran's Energy Consumption and Supply. *International Journal of Energy Economics and Policy* Vol. 2, No. 3, 2012, pp. 97-102.
- Sun Wei., Xu Yanfeng,. (2017). Research on China's energy supply and demand using an improved Grey-Markov chain model based on wavelet transform. *Energy* 118 (2017) 969-984.
- Xie, N.-ming, Yuan, C.-qing, & Yang, Y.-jie (2015). Forecasting China's energy demand and self-sufficiency rate by grey forecasting model and Markov model. *Electrical Power and Energy Systems*, 66(2015), 1e8.
- Xu Sun., Wangshu S., Jianzhou Wang., Yixin GAO (2015). Using a Grey Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China. *Tourism Management* 52 (2016) 369e379.

Development of a Markov Chain Grey Regression Model to Forecast Annual Natural Gas Consumption in Iran

Shahran Shafiei

Master of Industrial Engineering, Islamic Azad University-North Branch,
Tehran, Iran Shahran/shafiei@gmail.com

Vahid Baradaran¹

PhD. Industrial Engineering Department, Faculty of Engineering Islamic Azad
University-North Branch, Tehran, Iran, v_baradaran@iau-tnb.ac.ir

Received: 2018/02/06

Accepted: 2019/12/17

Abstract

Accurate forecasting of annual gas consumption of the country plays an important role in energy supply strategies and policy making in this area. Markov chain grey regression model is considered to be a superior model for analyzing and forecasting annual gas consumption. This model Markov is a combination of the Markov chain and grey regression models. According to this model, the residual errors generated from the grey regression model are divided into a number of equal portions. We will add the calculated error terms to the values obtained through the grey regression in order to increase its accuracy. We use the box-bench design to calculate the optimal value of the error term which produces the most accurate forecasts when using the Markov chain grey regression model. The experimental study of the forecasting of natural gas consumption indicates that the proposed Markov chain grey regression model is more accurate than they grey or conventional Markov chain regression models.

JEL Classification: C15, C53, C90, L95

Keywords: Grey Model, Regression Technique, Markov Chain, Design of Experiments

1. Corresponding Author