فصلنامه اقتصاد کاربردی
سال چهارم/ شماره دوازدهم/ بهار ۱۳۹۲

ارائه الگوی ترکیبی FARIMA با به کارگیری روشهای ARIMA و رگرسیون فازی جهت پیش بینی قیمت جهانی نفت خام

قدرت الله امام وردی[۱] ـ مریم شهابی طبری[۲]

تاریخ دریافت: ۹۲/۵/۱۰ تاریخ پذیرش: ۹۲/۱۱/۲

چکیده

مدل ARIMA، مدل پیش بینی دقیقی برای بازه کوتاه مدت می باشد ولی محدودیت تعداد داده های گذشته را نیز دارد. در جامعه کنونی، با توجه به شرایط نااطمینانی و همین طور رشد سریع تکنولوژی، نیاز به پیش بینی در بازه کوتاه مدت احساس می شود. معمولا داده های در دسترس کمتر از آن تعدادی است که در مدل ARIMA باید به کار گرفته شود. در این میان مدلهای رگرسیون فازی قادرند با داده های اندک و در شرایط نااطمینانی به پیش بینی مقادیر بپردازند. اما نتایج حاصل از این مدلها نیز معمولا به دلیل گستردگی بازه پیش بینی نمی تواند چندان قابل اتکا باشد. بنابراین محققین سعی بر ارائه روشی دارند که بتواند علاوه بر استفاده از مزایای هر دو روش، معایب آنها را نیز بکاهد. لذا در این مقاله، مدل ترکیبی میانگین متحرک خودرگرسیون انباشته فازی (FARIMA) به منظور پیش بینی قیمت جهانی نفت خام پیشنهاد شده است. یافته ها بیانگر بهبود نتایج در روش پیشنهادی می باشند.

طبقه بندی JEL: C53, E37,Q30

واژگان کلیدی: مدل اتورگرسیو میانگین متحرک انباشته (ARIMA)؛ مدل اتورگرسیو میانگین متحرک انباشته فازی(FARIMA)؛ پیش بینی؛ قیمت جهانی نفت خام

[۱] عضو هیات علمی گروه اقتصاد دانشگاه آزاد اسلامی واحد تهران مرکزی *ghemamverdi@gmail.com*

[۲] کارشناس ارشد اقتصاد دانشگاه آزاد اسلامی واحد تهران مرکزی (مسئول مکاتبات) *m.shahaby@gmail.com*

۱- مقدمه

همواره عوامل متعددی موجب نوسانات قیمت در بازار بوده و پیش بینی های میان مدت و بلندمدت را با تردید مواجه کرده است. این در حالی است که با پیش بینی کوتاه مدت می توان فرآیند تصمیم گیری خرید و فروش در بازار را تسهیل و بهترین زمان برای انجام معاملات را نیز تعیین کرد. بنابراین طراحی الگویی که بتواند به بهترین شکل در جهان در حال تغییر کنونی موثر واقع شود، کمک بزرگی به اقتصاد جهانی خواهد داشت. مدلهای خودرگرسیون میانگین متحرک انباشته (ARIMA[۱])، از آن گونه مدلهایی است که از زمان پیشنهاد باکس-جنکینز[۲] در پیش بینی مسائل متعددی همچون مسائل اجتماعی، اقتصادی، مهندسی و مالی استفاده شده و نتایج مفید و موثری نیز در برداشته اند. اساس عملکرد این گونه مدلها بر این فرض اولیه استوار است که مقادیر آینده سری زمانی، رابطه تابعی مشخص و واضحی با مقادیر گذشته و فعلی سری زمانی و همچنین خطاهای خالص مدل دارند (۱). این گونه مدلها، برای پیش بینی های کوتاه مدت بسیار مفید بوده و پیش بینی های صحیحی نیز در صورت فراهم بودن شرایط مطلوب ایجاد خواهند کرد. اما برای مدلسازی با این روش نیاز به حداقل ۵۰ و ترجیحا ۱۰۰ مشاهده می باشد، علاوه بر این، این مدلها از مفهوم عبارت خطا (تفاوت بین مقادیر تخمین زده شده با مقادیر اصلی) استفاده می کنند که این مقادیر برآورد شده، مقادیری قطعی بوده و جزء خطا را شامل نمی شوند(۱۰).از آنجا که بسیاری از متغیرهای اقتصادی مانند قیمت نفت در بازار، رفتاری آشوبناک از خود به نمایش می گذارند، بنابراین نمی توان به نتایج حاصل از مدلهای خطی کلاسیک چون ARIMA چندان اتکا کرد و باید روش های غیرخطی را نیز برای حصول نتایج بهتر و قابل قبول تر مد نظر قرار داد. رگرسیون فازی نیز یک مدل فاصله ای غیر خطی جهت پیش بینی در شرایط تعداد داده اندک می باشد. اما عیب این مدلها در قابل اتکا نبودن نتایج پیش بینی است. چرا که بازه پیش بینی در این مدلها معمولا وسیع بوده و پراکندگی زیادی در نتایج پیش بینی به چشم می خورد (۷). در این میان، روش ترکیبی خودرگرسیون میانگین متحرک انباشته فازی (FARIMA)[۳] با ترکیب دو روش ARIMA (خطی) و رگرسیون فازی(غیرخطی) قادر است این معایب را تا حدی بهبود بخشد و همزمان از مزایای روش باکس-جنکینز (B.J) مدلهای ARIMA نیز در بررسی مدل خود استفاده کند(۲). همچنین روش ترکیبی FARIMA قادر خواهد بود با بهره گیری از مزیت رگرسیون فازی، بدون نیاز به تعداد داده های زیاد، به پیش بینی پرداخته و همزمان گستردگی بازه پیش بینی را نیز در مدل رگرسیون فازی کاهش دهد(۱). به منظور نشان دادن عملکرد مدل ترکیبی FARIMA ، داده های روزانه قیمت جهانی نفت خام OPEC[۴] جهت پیش بینی مورد بررسی قرار گرفته اند. بدین منظور در بخش بعدی به برخی مطالعات انجام شده با روشهای مذکور اشاره می‌شود. در بخش سوم و چهارم این مقاله به ترتیب مبانی نظری مدلهای ARIMA و رگرسیون فازی، و فرموله کردن مدل ترکیبی FARIMA ارائه می گردد. سپس در بخش پنجم، به تجزیه و تحلیل داده های قیمت نفت خام با مدل FARIMA پرداخته و در انتهای این بخش نیز با استفاده از معیارهای خطا مقایسه ای جهت بررسی عملکرد مدلهای FARIMA و ARIMA برای مشاهدات انجام می شود و سرانجام در بخش آخر این مطالعه نتیجه‌گیری کلی از بحث بیان خواهد شد.

۲-پیشینه تحقیق

۱-۲- مطالعات خارجی:

رگرسیون خطی فازی اولین بار در سال ۱۹۸۲ توسط تاناکا و همکاران[۵] مطرح شد. آنها مدل های رگرسیون خطی با پارامترهای فازی، ورودی غیرفازی و خروجی فازی را به صورت مسائل برنامه ریزی خطی فرموله کردند. هدف آنها مینیمم کردن ابهام مدل رگرسیون خطی فازی بود به طوری که تکیه گاه[۶] مقادیر تخمین زده شده ،تکیه گاه مقادیر مشاهده شده را در یک سطح معین بپوشاند(۸). اما این مدل نیز معایبی داشت که از جمله مهمترین آنها می توان به وسیع شدن بیش از حد فاصله پیش بینی اشاره کرد. در ادامه ، چندین روش اصلاح شده رگرسیون فازی با استفاده از معیار مینیمم کردن ابهام مدل پیشنهاد شدند. تاناکا و ایشیبوچی(۱۹۹۱)[۷] توابع عضویت توان دوم را برای به دست آوردن ضرایب فازی معرفی کردند(۹). تاناکا و همکاران(۱۹۸۲)، رگرسیون امکان نمایی را پیشنهاد کردند. در این روشهای اصلاح شده از حداقل کردن ابهام مدل به عنوان معیار برازش استفاده شده است. در سال ۱۹۸۸، دیاموند[۸] روی مجموعه اعداد فازی یک متریک تعریف کرد و بر پایه آن روش کمترین مربعات فازی را برای مدلهای رگرسیون خطی فازی پیشنهاد کرد(۷). دیاموند حالتی را هم که ورودی وهم خروجی فازی هستند را نیز در نظر گرفت و پارامترهای فازی مدل را به قسمی تعیین کرد که مجموع فواصل بین مقادیر مشاهده شده و مقادیر تخمین زده شده حداقل شود . وی تنها مدلهای یک متغیره را مورد بحث قرار داد. ساکاوا و یانو(۱۹۸۱)[۹] مدلهای رگرسیون خطی فازی با خروجی- ورودی فازی را در نظر گرفتند و با

استفاده از سه شاخص برای تساوی بین اعداد فازی سه مسئله برنامه ریزی چندهدفه را برای تعیین پارامترهای فازی فرموله کردند. آنها از حداکثر(حداقل) کردن ابهام مدل و درجه تساوی بین مقادیر مشاهده شده و مقادیر تخمین زده شده به عنوان معیارهای برازش استفاده کردند.

۲-۲-مطالعات داخلی:

در سالهای اخیر مطالعه در زمینه سری های زمانی فازی و ترکیب آنها با سری های زمانی کلاسیک مورد توجه برخی محققین داخل کشور نیز قرار گرفته است.از جمله می توان به موارد زیر اشاره داشت:

خاشعی و بیجاری(۱۳۸۶) به پیش بینی نرخ ارز(دلار آمریکا در مقابل ریال) با ۴۲ داده روزانه این متغیر پرداختند. در انتهای مقاله نتایج مدل ترکیبی با مدل‌های سری زمانی فازی چن[۱۰] و مدل واتادا[۱۱] مقایسه شده است. نتایج این مقایسه نشان‌دهنده برتری مدل ترکیبی FARIMA بود .آنها همچنین در سال ۱۳۸۷، مقاله ای را تحت عنوان پیش بینی قیمت طلا با به کار گیری مدل ترکیبی FARIMA ارائه کردند و دریافتند که مدلهای سری زمانی فازی نسبت به مدلهای سری زمانی آماری کلاسیک نتایج بهتری را جهت پیش بینی با داده های کمتر ارائه می کنند (۱). زارع مهرجردی و نگارچی(۱۳۹۰)، در مقاله ای به مقایسه الگوهای ARIMA، رگرسیون فازی و FARIMA ، به منظور پیش بینی قیمت گوشت گوسفند پرداخته و نتیجه گرفتند مدلهای FARIMA از قابلیت بهتری جهت پیش بینی قیمتها برخوردارند چرا که می توانند بهترین و بدترین شرایط ممکن را نیز در کنار دقت بالای پیش بینی ارائه دهند (۲). ملکی (۱۳۹۰) در مطالعه‌ای به مقایسه روش‌های سری زمانی ARIMA و مدل‌های سری زمانی فازی در پیش‌بینی قیمت جهانی طلا در دو بازه زمانی متفاوت پرداخته است. وی همچنین از روش ترکیبی FARIMA به منظور پیش‌بینی قیمت جهانی طلا استفاده نموده و نتایج پژوهش حاکی از بهبود یافتن نتایج مدل ARIMA در بازه‌ی زمانی کوتاه‌مدت با استفاده از روش ترکیبیFARIMA بوده است (۴).

۳- مبانی نظری مدلهای خودهمبسته میانگین متحرک انباشته ARIMA و رگرسیون فازی

۱-۳-مدل ARIMA

ساختار ریاضی مدل ARIMA(p,d,q) به صورت رابطه زیر است:

$$B(L)(1-L)^d(Y_t-\alpha) = \theta(L)u_t \qquad (۱)$$

$$B(L)=1-\beta_1 L-\beta_2 L^2 - \cdots - \beta_P L^P$$
$$\theta(L)=\theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q$$
$$t=1,2,\ldots,k$$

θ پارامترهای مدل میانگین متحرک، p مرتبه مربوط به مدل خودهمبسته، q مرتبه مدل میانگین متحرک ، d مربوط به تفاضل و L عملگر پسرو، $\{Y_t\}$ بیانگر مقادیر مشاهده شده و α میانگین سری زمانی است. جمله خطای خالص u_t متغیری تصادفی با توزیع نرمال با میانگین صفر و واریانس δ^2 فرض شده است.

در حالت کلی مراحل مدلسازی در سری های زمانی بر اساس روند تکراری باکس و جنکینز شامل چهار مرحله شناسایی آزمایشی ساختار مدل، تخمین پارامترهای مجهول مدل، تشخیص دقت برازش مدل و پیش بینی با مدل انتخابی می باشد(۶). معیار مورد استفاده در این تحقیق تست معیار آکائیک(AIC) است.

$$AIC(i) = N\ln(\sigma_{\varepsilon i}{}^2)+2n_i \qquad (۲)$$

i دلالت بر تعداد مدل های منتخب، n تعداد پارامترها (مجموع مرتبه های مدل خودهمبسته و میانگین متحرک)، N تعداد داده های مشاهداتی، $\sigma_{\varepsilon i}$ انحراف معیار مدل است. این آزمون بر این مبنا استوار است که مرتبه ای که معیار آکائیک کمتری داشته باشد برازش بیشتری با سری مشاهداتی خواهد داشت.

۲-۳-رگرسیون فازی

رگرسیون فازی تعمیمی از رگرسیون معمولی است که برای محاسبه رابطه تابعی بین متغیرهای وابسته و متغیر مستقل در یک محیط فازی استفاده می شود. در روشهای رایج رگرسیون، تفاوت بین مقادیر مشاهده شده و مقادیر پیش بینی شده توسط مدل، خطای پیش بینی در مقایسه با مشاهدات بوده و یک متغیر تصادفی فرض می شود. حد بالا و حد پایین مقدار پیش بینی محاسبه شده و احتمال قرار گرفتن مقدار پیش بینی در بین این دو حد بیانگر اطمینان تخمین است. به عبارت دیگر مدلهای رایج رگرسیون ساختار احتمالی دارند. در حالی که در رگرسیون فازی، اختلاف بین مقادیر پیش بینی و مشاهده شده ناشی از ابهام ذاتی در سیستم است. مقدار خروجی برای ورودی های مشخص، دامنه مقادیر ممکن فرض می شود. بنابراین در رگرسیون فازی، امکان مقدار خروجی با توجه به مقادیر متغیرهای ورودی بررسی می شود. ضمن اینکه این مدل، مدل پیش بینی بازه ای مناسب در شرایط داده های قابل حصول کم است (۱). در مدل رگرسیون فازی عمومی، ضرایب متغیرها

اعداد فازی فرض می شوند. یک مدل عمومی رگرسیون خطی فازی به شرح ذیل است:

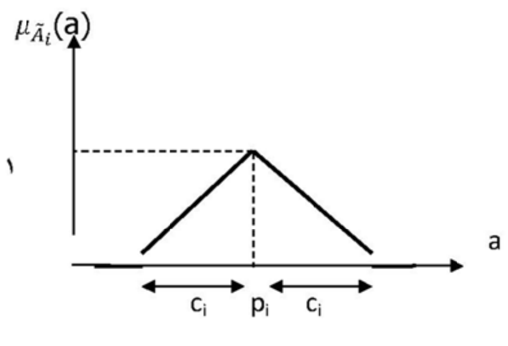
$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \ldots + \tilde{A}_n x_n \qquad (۳)$$

که در آن $\tilde{Y}$ متغیر وابسته فازی یا اصطلاحاً خروجی فازی می باشد و A_i، عدد فازی i امین ضریب است..

هنگامی که پارامترهای مدل رگرسیون خطی به جای توزیع احتمالی با توابع عضویت فازی بیان شوند مدل، رگرسیون خطی فازی نامیده می شود. مدل های رگرسیون خطی فازی در بسیاری از مسائل از جمله پیش بینی و مهندسی به کار گرفته می شوند(۲). هدف از تحلیل رگرسیون خطی فازی پیدا کردن یک مدل رگرسیون است که با معیار مشخصی از برازش، همه داده ها را برازش کند. وابسته به معیار برازش، مدلهای رگرسیون خطی فازی متفاوتی می توانند به دست آیند. ضرایب فازی با در نظر گرفتن توابع عضویت مثلثی متقارن به صورت $\tilde{A}=(p_i , c_i)$ $i=0,1,2,\ldots,n$ تعریف می شود (البته اعداد فازی دیگری نیز می توانند منظور شوند.) که به ترتیب p مرکز و c پهنای تابع عضویت(گستره عدد فازی) بوده که بیانگر میزان فازی بودن عدد را نشان می دهد . تابع عضویت عدد فازی مثلثی متقارن به صورت زیر است:

(۴)

$$\mu_{\tilde{A}_i}(a_i)=\begin{cases} 1-\frac{|a_i-p_i|}{c_i} & p_i-c_i \le a_i \le p_i+c_i \\ 0 & a_i < p_i-c_i \ , a_i > p_i+c_i \end{cases}$$



شکل(۱)- تابع عضویت عدد فازی مثلثی

در نتیجه تابع عضویت متغیر خروجی رگرسیون به صورت ذیل به دست می آید:

(۵)

$$\mu_{\tilde{Y}}(y)=\begin{cases} \max(\min\{\mu_{\tilde{A}_i}(a_i)\}) & , \{a|y=f(x,a)\}=\emptyset \\ 0 & , \quad otherwise \end{cases}$$

با جایگزینی رابطه (۴) در رابطه(۵) تابع عضویت متغیر خروجی به صورت ذیل به دست می آید:

(۶)

$$\mu_{\tilde{Y}}(y)=\begin{cases} 1-\frac{|y-\sum_{i=1}^{n} p_i x_i|}{\sum_{i=1}^{n} c_i |x_i|} & , \ x_i \ne 0 \\ 1 & , x_i = 0, \quad y = 0 \\ 0 & , \ x_i = 0, \ y \ne 0 \end{cases}$$

محققین روشهای مختلفی برای تعیین مدلهای رگرسیون خطی فازی پیشنهاد کرده اند. که از آن جمله می توان به روشهای "تحلیل رگرسیون با داده های غیر فازی" ، "تحلیل رگرسیون با داده های فازی" ، "مدل رگرسیون فازی حداقل مربعات"، اشاره کرد(۳). در این پژوهش از روش تحلیل رگرسیون فازی با داده های غیر-فازی(واقعی) استفاده شده است.

۴- فرموله کردن مدل ترکیبی اتورگرسیو میانگین متحرک انباشته فازی(FARIMA)

ARIMA یک مدل پیش بینی دقیق برای دوره های کوتاه مدت است. اما دارای تعداد زیاد داده های گذشته(حداقل ۵۰ و ترجیحا ۱۰۰ یا بیشتر) می باشد. در صورتی که امروزه باید موقعیتهای آینده را با استفاده از داده های کم و در بازه زمانی کوتاه مدت پیش بینی کرد. مدل های رگرسیون فازی یک مدل پیش بینی بازه ای مناسب در شرایط داده های قابل حصول کم است. اما این مدلها نیز معمولا مشکل وسیع شدن بازه پیش بینی در برخی از شرایط خاص را دارند. بازه پیش بینی می تواند به دلیل وجود داده های پراکنده مخصوصا در محیط های ناآرام اقتصادی وسیع شود.و این امر کاهش دقت پیش بینی را در پی خواهد داشت(۱۰). مدل رگرسیون انباشته فازی(FARIMA) از مزیتهای رگرسیون فازی و مدل ARIMA بهره گرفته و می کوشد تا محدودیتهای هر یک از دو مدل فوق را برطرف کند(۷). ضمن اینکه قابلیت انتخاب بهترین و بدترین حالت ممکن را برای تصمیم گیرندگان دارد. پارامترهای مدل ARIMA قطعی اند، درصورتی که در این روش به جای پارامترهای قطعی ، پارامترهای فازی به صورت اعداد فازی مثلثی به کار گرفته شده است(۱). بنابراین می توان نیاز به داده های گذشته را تا حدی کاهش داد. همچنین طبق نظر ایشیبوچی و تاناکا می توان تا حدی گستردگی بازه پیش بینی را کاهش داد و نتایج بهتری را ارائه کرد، که این مسئله در ادامه بحث مطرح می گردد.

یک مدل آریمای فازی (Fuzzy ARIMA) با توابع و پارامترهای فازی بدین صورت است:

$$\widetilde{\Phi}_p(B)W_t = \widetilde{\theta}_q(b)u_t \tag{۷}$$

$$W_t = (1-B)^d(Z_t - \propto) \tag{۸}$$

$$\widetilde{W}_t = \widetilde{\varphi}_1 W_{t-1} + \widetilde{\varphi}_2 W_{t-2} + \ldots + \widetilde{\varphi}_p W_{t-p} + u_t - \widetilde{\theta}_{p+1}u_{t-1} - \widetilde{\theta}_{p+2}u_{t-2} - \ldots - \widetilde{\theta}_{p+q}u_{t-q} \tag{۹}$$

که $\{z_t\}$ مشاهدات، $\widetilde{\varphi}_1, \widetilde{\varphi}_2, \ldots, \widetilde{\varphi}_p$ و $\widetilde{\theta}_1, \widetilde{\theta}_2, \ldots, \widetilde{\theta}_q$ اعداد فازی هستند.

حال معادله (۹) به صورت زیر تبدیل می‌گردد.

$$\widetilde{W}_t = \widetilde{\beta}_1 W_{t-1} + \widetilde{\beta}_2 W_{t-2} + \ldots + \widetilde{\beta}_p W_{t-p} + u_t - \widetilde{\beta}_{p+1}u_{t-1} - \widetilde{\beta}_{p+2}u_{t-2} - \ldots - \widetilde{\beta}_{p+q}u_{t-q} \tag{۱۰}$$

پارامترهای فازی در این معادله بصورت اعداد فازی مثلثی مطابق زیر در نظرگرفته شده‌اند.

$$\propto_{\widetilde{\beta}_i}(\beta_i) = \begin{cases} 1 - \dfrac{|\alpha_i - \beta_i|}{c_i} & \alpha_i - c_i \le \beta_i \le \alpha_i + c_i \\ 0 & otherwise \end{cases} \tag{۱۱}$$

بطوری‌که $\propto_{\widetilde{\beta}}(\beta_i)$ تابع عضویت مجموعه فازی‌ایست که با پارامترهای β_i و α_i مشخص می‌گردند. حال با استفاده از پارامترهای فازی β_i به صورت اعداد فازی مثلثی و همچنین اصل گسترش، تابع عضویت W مطابق ذیل خواهد بود.

$$\propto_{\widetilde{W}}(W_t) = \begin{cases} 1 - \dfrac{\left|W_t - \sum_{i=1}^{p}\alpha_i W_{t-i} - u_t + \sum_{i=p+1}^{p+q}\alpha_i u_{t+p-i}\right|}{\sum_{i=1}^{p} c_i|W_{t-i}| + \sum_{i=p+1}^{p+q} c_i|u_{t+p-i}|} & for \quad W_t \ne 0, \quad u_t \ne 0 \\ 0 & otherwise \end{cases} \tag{۱۲}$$

که h سطح آستانه‌ای برای میزان توابع عضویت تمامی مشاهدات است.

$$Z_z(Z_t) \ge h \qquad for\, i = 1,2,\ldots,k \tag{۱۳}$$

به عبارت دیگر s مطابق زیر تعریف می‌گردد.

$$S = \sum_{i=1}^{p}\sum_{t=1}^{k} c_i|\varphi_{ii}||W_{t-i}| + \sum_{i=p+1}^{p+q}\sum_{t=1}^{k} c_i|\rho_{i-p}||u_{t+p-i}| \tag{۱۴}$$

به قسمی ρ_{i-p} که ضریب خود همبستگی در وقفه زمانی $i-p$ و φ_{ii} ضریب خودهمبستگی جزئی در وقفه زمانی i ام می‌باشد. مراحل روش اریما فازی مطابق زیر می‌باشد:

مرحله اول: با برازش مدل ARIMA با استفاده از اطلاعات موجود در مشاهدات، ضرایب به صورت غیرفازی به دست می آیند. نتیجه فاز یک، جواب بهینه پارامترها $\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_{p+q}^*)$ خطای خالص می‌باشند که به عنوان یکی از مجموعه داده‌های ورودی در فاز دوم مورد استفاده قرار می‌گیرد.

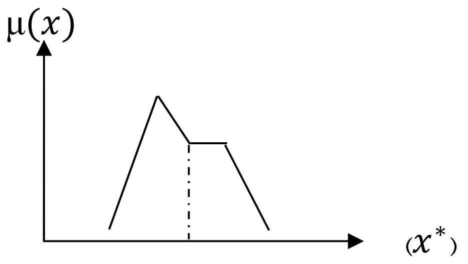
$$Minimize \quad S = \sum_{i=1}^{p}\sum_{t=1}^{k} c_i|\varphi_{ii}||W_{t-i}| + \sum_{i=p+1}^{p+q}\sum_{t=1}^{k} c_i|\rho_{i-p}||u_{t+p-i}| \tag{۱۵}$$

$$subject.to \quad \sum_{i=1}^{p}\alpha_i W_{t-i} + u_t - \sum_{i=p+1}^{p+q}\alpha_i u_{t+p-i} + (1+h)\left(\sum_{i=1}^{p} c_i|W_{t-i}| + \sum_{i=p+1}^{p+q} c_i|u_{t+p-i}|\right) \ge W_t \qquad t = 1,2,\ldots k$$

$$\sum_{i=1}^{p}\alpha_i W_{t-i} + u_t - \sum_{i=p+1}^{p+q}\alpha_i u_{t+p-i} + (1-h)\left(\sum_{i=1}^{p} c_i|W_{t-i}| + \sum_{i=p+1}^{p+q} c_i|u_{t+p-i}|\right) \le W_t \qquad t = 1,2,\ldots k$$

$$c_i \ge 0 \qquad for\, i = 1,2,\ldots,p+q$$

مرحله‌دوم: حال با استفاده از برنامه ریزی خطی تابع هدف S را (با توجه به محدودیتهایی که تعداد آنها دو برابر تعداد مشاهدات است) حداقل می کنیم. ابهام مدل با استفاده از معیارهایی همانند معادله (۱۵) و $\alpha^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_{p+q}^*)$ حداقل می گردد. بدین معنی که ضرایب فازی باید به گونه ای باشد که ابهام خروجی فازی حداقل شود. چرا که هر چه پهنای یک عدد فازی بیشتر باشد ابهام آن نیز بیشتر است. در انتها با حل رابطه (۱۵) مدل خودرگرسیو میانگین متحرک انباشته فازی به صورت تابعی معادله (۱۶) به دست می آید.



شکل (۲)- روش مرکز سطح جهت غیرفازی کردن داده ها

با به کارگیری رابطه (۱۷)، مقادیر فازی به مقادیر قطعی (x^*) تبدیل می شوند.(۳)

$$\widetilde{W}_t = \langle \alpha_1, c_1 \rangle W_{t-1} + ... + \langle \alpha_p, c_p \rangle W_{t-p} + u_t - \langle \alpha_{p+1}, c_{p+1} \rangle u_{t-1} - ... - \langle \alpha_{p+q}, c_{p+q} \rangle u_{t-q} \qquad (۱۶)$$

مرحله‌سوم: با توجه به نظرات ایشیبوچی و تاناکا داده‌های حد بالا و پایین وقتی که دامنه مدل وسیع می‌گردد، حذف خواهد شد، به منظور ساختن مدلی مناسب، اگر مجموعه داده ها شامل تفاوت های مشخص یا موارد خارج از محدوده باشند، C_i ها بسیار گسترده خواهند شد. طبق نظر ایشیبوچی داده های اطراف مرزهای بالا و پایین مدل در این حالت حذف می شود و مدل مجددا فرمول بندی می شود (Ishibuchi& Tanaka,1998,312).

غیر فازی کردن[۱۲] :

نتایج به دست آمده از مراحل بالا به صورت فازی می باشند ولی جهت مقایسه مقادیر به دست آمده با مقادیر واقعی لازم است که این مقادیر فازی طی فرآیندی به مقادیر قطعی و واقعی تبدیل شوند که به این فرآیند "غیر فازی کردن" گفته می شود. جهت این کار روشهای متعددی بسته به نوع کار و متغیر مورد نظروجود دارد. روش مرکز سطح[۱۳] یکی از پرکاربردترین این روشها می باشد که در این پژوهش نیز از این روش استفاده شده است. این روش که مرکز ناحیه یا مرکز ثقل نیز نام دارد به صورت زیر محاسبه می شود:

$$x^* = \frac{\int \mu(x) . x \, dx}{\int \mu(x) \, dx} \qquad (۱۷)$$

نماد $\int$ انتگرال جبری و $\mu(x)$ تابع عضویت مشاهده x را نشان می دهد. شکل (۲) مفهوم مرکز سطح را به تصویر کشیده است:

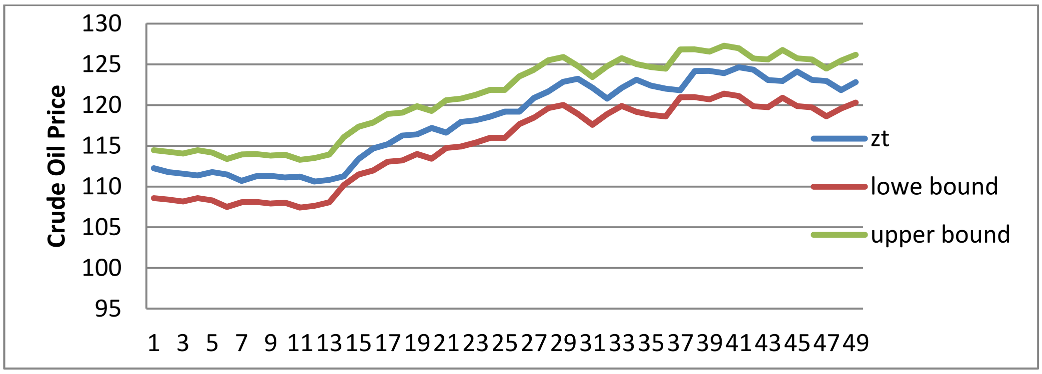
۵-تجزیه و تحلیل داده های قیمت جهانی نفت خام

از آنجا که مدل ARIMA دارای محدودیت تعداد داده های بالا (حداقل ۵۰ و ترجیحا ۱۰۰ مشاهده یا بیشتر) می باشد و به نظر می رسد که نتواند برای مشاهدات اندک نتایج قابل اعتمادی را ارائه کند . روش ترکیبی FARIMA با ترکیب دو روش ARIMA (خطی) و رگرسیون فازی(غیرخطی) قادر است این نتایج را تا حدی بهبود بخشد و همزمان از مزایای روش باکس-جنکینز (B.J) نیز در بررسی مدل خود استفاده کند. همچنین روش ترکیبی FARIMA قادر خواهد بود با بهره گیری از رگرسیون فازی خطاهای بالای مدل ARIMA را حداقل کند.

حال با توجه به مباحثی که توضیح داده شد به مدلسازی قیمت نفت خام به روش میانگین متحرک اتورگرسیو انباشته فازی می پردازیم. لازم به توضیح است این مدل سازی بر روی ۵۰ مشاهده پایانی قیمت روزانه نفت خام انجام گرفته است و این درحالی ست که مدل ARIMA با استفاده از ۵۲۲ مشاهده قیمت روزانه نفت خام OPEC(از ۲۰۱۰/۳/۲۶ تا ۲۰۱۲/۳/۲۶) برآورد شده است. بدین ترتیب می توان توانایی پیش بینی مدل FARIMA را در کوتاه مدت بررسی و نتایج را با پیش بینی در بازه کوتاه مدت مدل ARIMA مقایسه نمود ، مراحل روش اریمای فازی به شرح ذیل می باشد:

فاز ۱: برازش مدل ARIMA با به کارگیری نرم افزار EVIEWS انجام گرفته است و بهترین مدل ARIMA(1,1,0) تخمین زده شد. معادله مدل حاصله به همراه باقیمانده های مدل(خطای خالص) در رابطه(۱۸) آورده شده است:

Z_t = ۰/۱۵۲۷۲۷ +۰/۲۳۱۶۵۸ Z_{t-1}+ U_t (۱۸)



شکل (۳)- مقادیر واقعی قیمت نفت به همراه حدود بالا و پایین آنها

ماخذ: یافته های تحقیق

فاز ۲ : (تعیین حداقل ابهام): با قرار دادن α_0 و α_1 در معادله به عنوان مجهول و با استفاده از معادله (۱۵) و (h=0) و حل این برنامه ریزی خطی در نرم افزار Lingo، معادله با پارامترهای فازی به صورت زیر به دست می آید[۱۴]:

$$\tilde{Z}_t = (۰, ۲/۹۳۵۱۲۷) + (۰/۹۹۷۶۸۴۰, ۰)\tilde{Z}_{t-1} \quad (۱۹)$$

شکل(۳) مقادیر واقعی به همراه حدود پایین و بالای آنها را نشان می دهد. همان طور که مشاهده می شود، مقادیر واقعی در فواصل فازی (حد بالا و پایین) قرار گرفته‌اند.

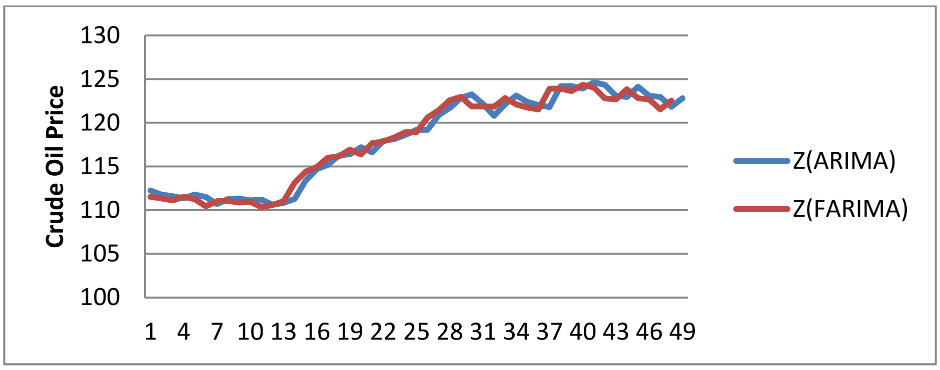
پس از محاسبه واریانس سری z_t (مقادیر واقعی)، حدود بالا و پایین به ترتیب به صورت z_t+2var و z_t-2var محاسبه می شوند. (varبه معنی واریانس سری می باشد.)

فاز ۳: طبق نظرات ایشیبوچی[۱۵]، در صورتی که مشاهدات ما در یکی از دو مرز بالا و پایین قرار بگیرند باید محدودیت مربوط به آن مشاهده حذف و مدل مجددا فرموله شود.(h=0). نتایج حاصله از شکل(۳) نشان می دهد که هیچ یک از مشاهدات بر روی دو حد بالا و پایین قرار نگرفته اند و مدل آریمای فازی نتایج خوبی را ارائه می کند. پس نیازی به انجام این مرحله نخواهد بود.

غیر فازی کردن :

از انجا که مقادیر مشاهدات واقعی قیمت نفت، قطعی هستند جهت مقایسه مقادیر پیش بینی شده از مراحل بالا که به صورت مقادیر فازی برآورد شده اند با مشاهدات واقعی قیمت نفت نیاز به تبدیل مقادیر فازی به مقادیر قطعی می باشد که به این فرآیند "غیر فازی کردن" گفته می شود[۱۶]. جهت این کار روش مرکز سطح[۱۷] که یکی از پرکاربردترین این روشها می باشد در این پژوهش استفاده شده است. نهایتا رابطه (۱۷) را برای تک تک مشاهدات به کار گرفته و بدین صورت مقادیر فازی به مقادیر قطعی (x^*) تبدیل می شوند

در شکل(۴) می توان مقادیر پیش بینی قیمت نفت خام به روش مذکور را با مشاهدات واقعی مقایسه کرد. همان طور که در شکل مشاهده می شود، مقادیر پیش بینی شده در این روش تا حد زیادی بر مقادیر واقعی قیمت نفت منطبق هستند.



شکل (۴)- مقادیر واقعی و پیش بینی قیمت نفت به روش FARIMA

ماخذ: یافته های تحقیق

جدول(۱)- مقادیر خطا در مدل های سری زمانی فازی پیشنهادی و FARIMA

معیار / مدل	ریشه میانگین مجذور خطا RMSE	میانگین درصد خطا MPE	میانگین مطلق خطا MAE	میانگین مطلق درصد خطا MAPE
ARIMA	۰،۸۹۰۶۶۸	۰/۰۰۰۶۵۹	۰/۷۱۰۲۹۴	۰/۰۰۵۹۷۸
FARIMA	۰/۸۳۹۲۲۶	-۰/۰۰۰۱۶	۰/۶۷۳۷۱۱	۰/۰۰۵۸۰۷

حال با استفاده از معیارهای خطا به بررسی توان پیش بینی هر مدل پرداخته می شود.در این تحقیق از معیارهای خطای ریشه میانگین مجذور خطا[۱۸](RMSE) ، میانگین درصد خطا[۱۹](MPE)، میانگین مطلق خطا[۲۰](MAE)، میانگین مطلق درصد خطا[۲۱] (MAPE) استفاده شده است.

همانطور که در جدول(۱) مشاهده می شود، مدل FARIMA توانسته مقادیر کمتری از معیارهای خطا را نسبت به مدل ARIMA به خود اختصاص دهد. و بنابراین می توان نتیجه گرفت که مدل ترکیبی FARIMA از توان پیش بینی بالاتری برخوردار بوده و قادر است نتایج بهتری را برای سری زمانی قیمت نفت خام OPEC ارائه دهد.

۶- نتیجه گیری

با پیشرفت و توسعه مداوم، روشهای هوش مصنوعی، کاربردهای روزافزونی در بحث پیش بینی پیدا کرده اند. در این میان روش آریمای فازی (Fuzzy ARIMA)نیز از آنجا که قادر به بیان بدترین و بهترین حالات ممکن می باشد روش مناسبی به خصوص برای تصمیم گیرندگان و سیاست گذاران بوده است و امروزه کاربردهای فراوانی جهت پیش بینی در بازارهای مالی داشته است. در این مطالعه، مقادیر معیارهای خطا حاکی از این مدل نشان داد که روش FARIMA قادر است با داده هایی کمتر از مدل ARIMA ، نتایج مناسبتری را در پیش بینی قیمت جهانی نفت خام ارائه دهد. همچنین مدل پیشنهادی با فازی در نظر گرفتن خروجی ها، از فروض مربوط به جمله خطا رهایی می یابد . بنابراین روش پیشنهادی می تواند به عنوان یک مدل پیش بینی فاصله ای در مدل سازی ها مورد استفاده قرار گیرد.

منابع:

۱- خاشعی،مهدی؛ بیجاری،مهدی؛۱۳۸۷؛پیش بینی قیمت طلا با به کارگیری مدل ترکیبی مدلهای خودرگرسیون میانگین متحرک انباشته کلاسیک با منطق فازی؛ جلد ۳۱؛ شماره ۳.

۲- زارع مهرجردی،محمدرضا،نگارچی،سمانه،۱۳۹۰،مقایسه الگوهای میانگین متحرک خودرگرسیون انباشته ، رگرسیون فازی و رگرسیون انباشته فازی به منظور پیش بینی قیمت(مطالعه موردی قیمت گوشت گوسفند)، نشریه اقتصاد و توسعه کشاورزی،جلد ۲۵،شماره :۱۰۸-۱۰۰.

۳- شوندی، حسن؛ ۱۳۸۹؛ نظریه مجموعه های فازی و کاربرد آن در مهندسی صنایع و مدیریت؛ نشر گسترش علوم پایه؛ تهران.

۴- ملکی،گلناز،۱۳۹۰، پیش بینی قیمت جهانی طلا با استفاده از روش های سری زمانی و منطق فازی،پایان نامه کارشناسی ارشد،دانشکده اقتصاد دانشگاه آزاد تهران مرکزی.

۵- نصر آبادی، ابراهیم، ۱۳۸۲؛ تحلیل مدل های رگرسیون خطی فازی: روشهای برنامه ریزی ریاضی؛ پایان نامه کارشناسی ارشد؛ داشکده مهندسی صنایع دانشگاه صنعتی شریف.

6- Box P, and Jenkins G.M. 1976. Time Series Analysis, Forecasting and control, Holden-day Inc, San Francisco, CA.

7-Diamond.P, 1988, Fuzzy least squares, Inform. Sci. 46 , 141–157.

8- Tanaka, H, 1987, "Fuzzy data analysis by possibility linear models", Fuzzy Sets and Systems, 24(3),363- 375.

9-Tanaka, H. Ishibuchi, 1991, Identi0cation of possibilistic linear systems by quadratic membership functions of fuzzy parameters, Fuzzy Sets and Systems 41 , 145–160.

10-Wang, Ming- Jyh and et al,2009, A Fuzzy ARIMA Model by using Quadratic Programming Approuch for Time Series Data, Vol 5,No1,41-51.

پیوست: (نتایج خروجی نرم افزار LINGO)

Global optimal solution found.
Objective value: 2.935127
Infeasibilities: 0.000000
Total solver iterations: 3

Variable	Value	Reduced Cost
C0	2.935127	0.000000
C1	0.000000	1135.262
P0	0.000000	0.2594810E-01
P1	0.9976840	0.000000

Row	Slack or Surplus	Dual Price
1	2.935127	-1.000000
2	2.638167	0.000000
3	1.786557	0.000000
4	2.256417	0.000000
5	2.792917	0.000000
6	2.730537	0.000000
7	1.386418	0.000000
8	2.397281	0.000000
9	2.727335	0.000000
10	2.280719	0.000000
11	2.448637	0.000000
12	2.107942	0.000000
13	2.147960	0.000000
14	3.349099	0.000000
15	5.076523	0.000000
16	5.870255	0.000000
17	3.847594	0.000000
18	3.829685	0.000000
19	3.662348	0.000000
20	3.226178	0.000000
21	2.732919	0.000000
22	3.113941	0.000000
23	4.173923	0.000000
24	2.889655	0.000000
25	3.557724	0.000000
26	3.035177	0.000000
27	4.078855	0.000000
28	5.035949	0.000000
29	4.134054	0.000000
30	3.892400	0.000000
31	1.543863	0.000000
32	0.000000	-0.4870259
33	2.769780	0.000000
34	5.175376	0.000000
35	2.496222	0.000000
36	1.203531	0.000000
37	2.124778	0.000000
38	4.765922	0.000000
39	4.958808	0.000000
40	1.734071	0.000000
41	2.974483	0.000000
42	3.012359	0.000000
43	0.8059364	0.000000
44	1.222504	0.000000
45	3.862211	0.000000
46	2.690505	0.000000
47	1.094307	0.000000
48	1.514231	0.000000
49	2.425792	0.000000
50	4.491145	0.000000
51	3.232088	0.000000
52	4.083698	0.000000
53	3.613838	0.000000
54	3.077338	0.000000
55	3.139718	0.000000
56	4.483836	0.000000
57	3.472973	0.000000
58	2.517256	0.000000
59	3.589536	0.000000
60	3.421618	0.000000
61	3.762313	0.000000
62	3.722294	0.000000
63	2.521155	0.000000
64	0.7937314	0.000000
65	0.000000	0.5129741
66	2.022661	0.000000
67	2.040569	0.000000
68	2.207907	0.000000
69	2.644077	0.000000
70	3.137335	0.000000
71	2.756314	0.000000
72	1.696332	0.000000
73	2.980599	0.000000
74	2.312531	0.000000
75	2.835078	0.000000
76	1.791400	0.000000
77	0.8343060	0.000000
78	1.736201	0.000000
79	1.977855	0.000000
80	4.326391	0.000000
81	5.870255	0.000000
82	3.100475	0.000000
83	0.6948789	0.000000
84	3.374032	0.000000
85	4.666723	0.000000
86	3.745477	0.000000
87	1.104333	0.000000
88	0.9114470	0.000000
89	4.136184	0.000000
90	2.895771	0.000000
91	2.857895	0.000000
92	2.129191	0.000000
93	4.647750	0.000000
94	2.008044	0.000000
95	3.179749	0.000000
96	4.775947	0.000000
97	4.356023	0.000000
98	3.444463	0.000000
99	1.379109	0.000000
100	2.935127	0.000000

یادداشت‌ها

[1] *Auto Regressive Integrated Moving Average*
[2] *Box-Jenkins*
[۳] به این مدل *Fuzzy ARIMA* نیز اطلاق می شود.
[4] *Organization of Petroleum Exporting Countries*
[5] *Tanaka et al*
[6] *Supportive*
[7] *Tanaka, Ishibuchi*
[8] *Diamond*
[9] *Sakava, yano*
[10] *Chen*
[11] *Watada*
[12] *Defuzzification*
[13] *Centroid Method*
[14] نتایج نرم افزار در پیوست آورده شده است.
[15] *Ishibuchi*
[۱۶] مراحل محاسباتی این فرآیند در نرم افزار *MATLAB* انجام گردیده است
[17] *Centroid Method*
[18] *RMSE*= $\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Z_t - \hat{Z}_t)^2}$
[19] *MPE*= $\frac{1}{T}\frac{\sum_{t=1}^{T}(Z_t-\hat{Z}_t)^2}{\sum_{t=1}^{T}Z_t}$
[20] *MAE*= $\frac{1}{T}\sum_{t=1}^{T}\left|(Z_t - \hat{Z}_t\right|$
[21] *MAPE*= $\frac{1}{T}\frac{\sum_{t=1}^{T}|(Z_t-\hat{Z}_t|}{\sum_{t=1}^{T}Z_t}$

Z_t مقدار داده واقعی و $\hat{Z}_t$ مقدار پیش بینی شده در زمان t و T تعداد مشاهدات را نشان می دهد.