Journal of Agricultural Economics and Development
Vol. 25, No. 1, Spring 2011, p. 100-108

نشریه اقتصاد و توسعه کشاورزی (علوم و صنایع کشاورزی)
جلد ۲۵، شماره ۱، بهار ۱۳۹۰، ص. ۱۰۸-۱۰۰

مقایسه الگوهای میانگین متحرک خود رگرسیون انباشته، رگرسیون فازی و رگرسیون انباشته فازی به منظور پیش‌بینی قیمت (مطالعه موردی: قیمت گوشت گوسفند)

محمدرضا زارع مهرجردی [1*]– سمانه نگارچی [2]

تاریخ دریافت: ۸۹/۱۰/۳۰
تاریخ پذیرش: ۹۰/۲/۳۱

چکیده

امروزه به علت عدم قطعیت محیط و توسعه سریع تکنولوژی نوین معمولاً باید موقعیت‌های آینده را با استفاده از داده‌های کم و در بازه زمانی کوتاه‌مدت پیش‌بینی کرد. بنابراین به روش‌هایی برای پیش‌بینی نیاز است که به داده‌های کمتری احتیاج داشته باشد. مدل میانگین متحرک خودرگرسیون و روش شبکه عصبی مصنوعی برای دست‌یابی به نتایج دقیق نیاز به داده‌های زیادی دارند اما مدلهای رگرسیون فازی، مدل‌های مناسبی برای پیش‌بینی با استفاده از داده‌های کمتری نسبت به دیگر روشها می‌باشند. در این مقاله به منظور برطرف ساختن مشکل مذکور و حصول نتایج دقیق‌تر به بررسی سه روش میانگین متحرک خودرگرسیون انباشته، رگرسیون فازی و میانگین متحرک خودرگرسیون انباشته فازی که از ترکیب دو روش قبل بدست آمده، پرداخته شده است. مقایسه میزان صحت پیش‌بینی مدل‌های مذکور براساس دو معیار خطای ریشه متوسط مربعات(RMSE) و ضریب تعیین (R^2) حاکی از آن است که مدل میانگین متحرک خودرگرسیون انباشته فازی به عنوان الگوی برتر جهت پیش‌بینی مقادیر شاخص قیمت بوده است.

واژه های کلیدی: ARIMA، پیش‌بینی قیمت، رگرسیون انباشته فازی، رگرسیون فازی

مقدمه[1]

فرآیند پیش‌بینی معمولاً شامل اطلاعات تاریخی و تعمیم آن‌ها به آینده به کمک مدل‌های ریاضی است. از آن جا که پیش‌گویی وقایع آینده در فرآیند تصمیم‌گیری نقش عمده‌ای ایفا می‌کند، لذا پیش‌بینی برای بسیاری از سازمان‌ها و نهادها حائز اهمیت است و می‌توان پیش‌بینی را ابزاری مفید برای برنامه‌ریزی‌های کوتاه‌مدت و بلندمدت تلقی کرد (۸). تعدادی از اقتصاددانان با استفاده از روش‌های متنوع اقتصادسنجی با تأکید بر تقدم تئوری‌های اقتصادی، سعی در تبیین وضع موجود، ارائه سیاست‌های اقتصادی و همچنین پیش‌بینی مقادیر آتی نمودند. هرچند این مدل‌ها قادر به توضیح نسبی وضع موجود بوده و به عنوان ابزار مناسبی برای سیاست‌گزاری اقتصادی مورد استفاده قرار گرفتند، اما متأسفانه در زمینه پیش‌بینی چندان موفق نبودند (۱۰). وجود نقایص فوق و اهمیت روز افزون پیش‌بینی، اقتصاددانان را بر آن داشت که بدون توجه به تئوری‌های اقتصادی، پیش بینی متغیرهای اقتصادی را به عهده خودشان واگذار نموده و برای پیش‌بینی از روش‌هایی که به سری زمانی موسومند، بهره جویند. چراکه هر متغیر اقتصادی حاوی کلیه اطلاعات مربوط به خود بوده و لذا قویترین منبع برای توضیح تغییرات هر متغیر، خود متغیر می باشد (۴).

پیش‌بینی دقیق قیمت از طریق کاهش نوسانات قیمتی، می‌تواند موجب انتخاب نوع محصول و میزان تولید، تخصیص بهینه منابع، افزایش کارایی، افزایش مطلوبیت تولیدکننده و در نهایت افزایش درآمد گردد. همچنین پیش‌بینی قیمت نقش مؤثری در کاهش ناپایداری شدید قیمت‌ها و کاهش ریسک خواهد داشت. پیش‌بینی قیمت علاوه بر موارد یاد شده، نقش مؤثری در سیاست‌های دولت دارد. چرا که دولت سیاست‌های خود را نه صرفاً بر مبنای وضع موجود، بلکه بر مبنای پیش‌بینی‌های کوتاه‌مدت و بلندمدت از متغیرهای کلیدی اقتصادی از جمله قیمت محصولات مختلف، تدوین نموده و به اجرا می‌گذارد. لذا میزان دقت پیش‌بینی این متغیرها، صرف‌نظر از درستی و تناسب سیاست‌ها با شرایط موجود، از جمله

۱ و ۲– به ترتیب استادیار و دانشجوی کارشناسی ارشد (عضو انجمن پژوهشگران جوان) گروه اقتصاد کشاورزی، دانشگاه شهید با هنر کرمان
(*– نویسنده مسئول: Email: zare@mail.uk.ac.ir)

رموز موفقیت این سیاست‌ها به شمار می‌آیند.

در این قسمت برخی از مطالعات انجام شده در حوزه پیش‌بینی قیمت‌های کشاورزی مرور شده‌است. کهزادی و همکاران علاوه بر پیش‌بینی قیمت سلف ذرت، مدل شبکه عصبی را با مدل میانگین متحرک خودرگرسیون انباشته (ARIMA)[1] مقایسه کردند. نتایج نشان داد خطای پیش‌بینی مدل شبکه عصبی کم‌تر از فرآیند ARIMA است (۱۶). گیلان‌پور و کهزادی (۷) قیمت فوب برنج تایلندی را با استفاده از فرآیند ARIMA ، پیش‌بینی کردند. مجاوریان و امجدی (۹) قیمت مرکبات را با استفاده از سه روش ARIMA بدون ملاحضات فصلی، فرآیند ARIMA با در نظر گرفتن اثرات فصلی و تابع مثلثاتی سینوسی پیش‌بینی و نشان دادند توابع مثلثاتی کارایی بیشتری در پیش‌بینی خارج از نمونه دارند. طرازکار (۴) با استفاده از روش میانگین ساده، میانگین متحرک، تعدیل نمایی یگانه و دوگانه، ARIMA، هارمونیک و ARCH و شبکه عصبی به پیش‌بینی قیمت محصولات کشاورزی در استان فارس پرداخت و نشان داد که برای افق زمانی یک و سه ماه روش شبکه عصبی مصنوعی و برای افق زمانی شش ماه روش تعدیل نمایی، پیش‌بینی بهتری ارایه می‌کند. این گونه از مدل‌ها برای پیش‌بینی‌های کوتاه‌مدت بسیار مفید بوده و احتیاج به حداقل پنجاه و ترجیحاً یکصد مشاهده یا بیشتر دارد. علاوه بر این، این‌گونه مدلها از مفهوم عبارت خطا (تفاوت بین مقادیر تخمین‌زده شده با مقادیر اصلی) استفاده می‌کنند. لذا به منظور به کارگیری این‌گونه از مدلها باید تمامی فرضهای مربوط به عبارت خطا را در مدلسازی لحاظ کرد. تاناکا به منظور جلوگیری از خطای مدلسازی، رگرسیون فازی که اساساً یک مدل پیش‌بینی فاصله‌ای است را پیشنهاد کرد (۱۷).

نظریه مجموعه‌های فازی در سال ۱۹۶۵ عرضه شد (۲۱). این نظریه یک قالب جدید ریاضی به منظورتجزیه و تحلیل مفاهیم و متغیرهای مبهم و سیستم‌های مبتنی بر روابط تقریبی ارائه می‌کند. در یک مدل رگرسیونی ممکن است خطای مدل ناشی از مبهم‌بودن و یا تقریبی‌بودن روابط بین متغیرهای مورد مطالعه باشد و ارتباطی به عدم اطمینان منسوب به خطای تصادفی نداشته باشد. هم چنین ممکن است در یک مدل رگرسیونی با مشاهدات مبهم روبه‌رو باشیم. در این موارد می توان از مدل های رگرسیون فازی به جای مدل‌های رگرسیون آماری استفاده کرد (۳).

رگرسیون فازی نخستین بار توسط تاناکا و همکاران معرفی شد. آنها مدل رگرسیون خطی با ضرایب فازی را مورد توجه قرار دادند (۱۳). پس از آن محققان بسیاری اقدام به مطالعه و بررسی جنبه‌های مختلف آن نمودند (۱۹). چی چن وانگ (۱۸) در مطالعه‌ای به مقایسه روش سری زمانی فازی و مدل ARIMA در راستای پیش‌بینی صادرات کشور تایوان پرداخته است. نتایج مقایسه نشان داد که روش سری زمانی فازی مطلوبتر از روش ARIMA است. آزاده و همکاران (۱۱) به پیش‌بینی میزان مصرفی نفت در کشورهای استرالیا، کانادا، آمریکا و ژاپن پرداخته و نشان داده است که رگرسیون فازی در راستای تخمین میزان مصرفی نفت، روش مناسبی است.

در این رهیافت، مسأله برازش مدل رگرسیون فازی عموماً معادل با یک مسأله برنامه‌ریزی خطی(و در بعضی حالت‌ها، برنامه‌ریزی غیرخطی) می‌شود. رهیافت دیگر، استفاده از روش کمترین مربعات برای بررسی و برازش مدل‌های رگرسیون فازی است که نخستین بار توسط کلمینس و دایاموند مورد توجه قرار گرفت. کیم و همکاران رگرسیون آماری (معمولی) را با رگرسیون فازی از جهات مختلف مانند فرض‌های اولیه، چگونگی برآورد پارامترها و زمینه‌های کاربردی مورد مقایسه قرار دادند (۲). هرچند مطالعات متعددی در زمینه پیش‌بینی قیمت محصولات کشاورزی در ایران انجام گرفته است، با این حال در هیچ‌کدام از این مطالعات از مدلهای فازی استفاده نشده است. لذا با توجه به توانایی بالای رگرسیون فازی در پیش‌بینی، در این مطالعه توانایی این روش را با روش ARIMA که روشی سنتی در پیش‌بینی قیمت محسوب می‌شود، مقایسه گردید. اهدافی که مقاله حاضر دنبال می کند به شرح زیر است:

- معرفی رگرسیون فازی و نشان‌دادن قابلیت‌های کاربردی آن در بسط و توسعه مدل‌های خطی (رگرسیون) به منظور پیش‌بینی قیمت.
- بررسی عملکرد سه مدل سری‌زمانی شامل مدل ARIMA ، رگرسیون فازی با تابع عضویت مثلثی، همچنین مدل میانگین متحرک خودرگرسیون انباشته فازی[2] که ترکیبی از دو روش ARIMA و رگرسیون فازی جهت پیش‌بینی شاخص قیمت.
- پیش بینی با استفاده از داده‌های کمتر نسبت به روش‌های معمول اقتصاد سنجی.
- تعیین بهترین و بدترین موقعیت‌های ممکن برای تصمیم‌گیری‌های صحیحتر و دقیقتر.

مواد و روش‌ها

ARIMA

مدل خودهمبسته میانگین متحرک[3] برای مدلسازی سری‌زمانی ایستا و مدل ARIMA برای مدلسازی سری‌های زمانی ناایستا کاربرد دارند. از جمله روش‌های تبدیل سری ناایستا به سری‌ایستا استفاده از روش تفاضل‌گیری ناایستا است. ساختار ریاضی مدل

1 -Auto-Regressive Integrated Moving Average

2-Fuzzy Auto Regressive Integrated Moving Average (FARIMA)

3 - Auto regressive moving average(ARMA)

ARIMA(p,d,q) بصورت رابطه زیر است (۶):

$$B(L)(1-L)^d(Y_t-\alpha)=\theta(L)u_t \qquad (۱)$$
$$B(L)=1-\beta_1 L-\beta_2 L^2-\ldots-\beta_P L^P$$
$$\theta(L)=1-\theta_1 L-\theta_2 L^2-\ldots-\theta_q L^q$$
$$t=1,2,\ldots,k$$

θ پارامترهای مدل میانگین متحرک، p مرتبه مربوط به مدل خودهمبسته، q مرتبه مربوط به مدل میانگین متحرک، d مرتبه مربوط به تفاضل و L عملگر پسرو، $\{Y_t\}$ بیانگر مقادیر مشاهده شده و α میانگین سری‌زمانی است.جمله خطای خالص u_t متغیری تصادفی با توزیع نرمال با میانگین صفر و واریانس ∂^2 فرض شده‌است. در حالت کلی مراحل مدلسازی در سری‌های زمانی براساس روند تکراری باکس و جنکینز شامل چهار مرحله شناسایی آزمایشی ساختار مدل، تخمین پارامترهای مجهول مدل، تشخیص دقت برازش مدل و پیش‌بینی با مدل انتخابی می‌باشد (۱۲). 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 مدل‌های منتخب، n تعداد پارامترهای(مجموع مرتبه‌های مدل خودهمبسته و میانگین متحرک)، N تعداد داده‌های مشاهداتی، $\sigma_{\varepsilon i}$ انحراف معیار مدل است. این تست براین مبنا استوار است که مرتبه‌ای که معیار آکائیک کمتری داشته باشد برازش بیشتری با سری مشاهداتی خواهد داشت.

رگرسیون فازی

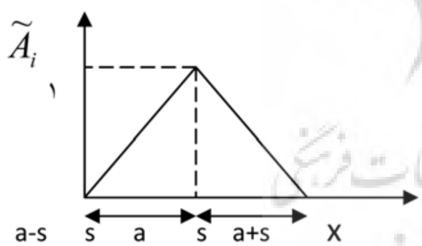
منطق فازی ابزاری توانمند جهت حل مسایل مربوط به سیستم های پیچیده‌ای که درک آنها مشکل و یا مسایلی که وابسته به استدلال، تصمیم‌گیری و استنباط بشری می‌باشد، بشمار می‌آید. به همین دلیل بهترین وسیله برای مدلسازی سیستم‌هایی است که دارای پیچیدگی زیاد بوده و داده‌های کافی از آنها موجود نیست و یا اطلاعات آنها مبهم و غیرصحیح می‌باشد. استفاده از این ابزار در رگرسیون فازی سبب ارتقاء این مدل از یک مدل آماری محض به سمت مدلی با قابلیت و کارایی بیشتری با تاکید بر در نظر گرفتن عدم قطعیت حاکم بر مسئله شده‌است (۵).

هدف اصلی در مدل رگرسیون پیدا کردن مدل ریاضی مناسب و تعیین ضرایب مدل با هدف بهترین برازش نتایج مدل رگرسیون با مقادیر مشاهداتی می‌باشد. اختلاف بین مقادیر مشاهداتی و تخمینی ناشی از خطای مشاهداتی است ولی این تفاوت در رگرسیون فازی ناشی از ابهام در ساختار سیستم می‌باشد (۲۰). یکی از روش‌های رگرسیون فازی، روش‌های معروف به رگرسیون امکانی است که ارتباط بین متغیرها، فازی فرض شده‌است. به عبارت دیگر ضرایب مدل رگرسیون فازی در نظر گرفته شده‌است. یافتن مدلهای رگرسیون بر پایه‌ی این روش‌ها غالباً مبتنی بر حل مسائل برنامه‌ریزی خطی یا غیرخطی است. در این مدل‌ها خطای پیش‌بینی متغیر وابسته تحلیل امکانی دارد. به عبارت دیگر هنگامی که بر پایه مدل رگرسیون و به ازای مقادیری از متغیرها ی مستقل مقدار متغیر وابسته را پیش‌بینی می‌کنیم، مقدار پیش بینی تحلیل امکانی دارد نه تعبیر احتمالی. صورت کلی مدل رگرسیون خطی با ضرایب فازی به صورت زیر است که در آن $\tilde{Y}$ متغیر وابسته فازی یا اصطلاحاً خروجی فازی می‌باشد.

$$\tilde{Y}=\tilde{A}_0+\tilde{A}_1x_1+\ldots+\tilde{A}_nx_n \qquad (۳)$$

هدف این است که بر اساس یک مجموعه از داده‌های مشاهداتی ضرایب فازی معادله بگونه‌ای بدست آورده می‌شود که معادله مزبور بهترین برازش بر روی مقادیر مشاهداتی را داشته باشد. ضرایب فازی با در نظر گرفتن توابع مثلثی متقارن به صورت $\tilde{A}_i=(a_i,s_i)_T \quad i=0,1,2,\ldots,n$ تعریف می‌شود که ضرایب همراه به ترتیب مرکز و پهنای تابع عضویت را نشان می‌دهد. تابع عضویت عدد فازی متقارن به صورت زیر است (۲و۱۵):

$$\tilde{A}(x)=\begin{cases}1-\dfrac{a-x}{s} & a-s\le x\le a\\[2mm] 1-\dfrac{x-a}{s} & a<x\le a+s\end{cases} \qquad (۴)$$



شکل۱- تابع عضویت مثلثی متقارن ضرایب فازی

اگر اعداد فازی متقارن و x_i ها اعداد حقیقی باشند آنگاه $\tilde{Y}$ یک عدد فازی مثلثی متقارن بصورت $\tilde{Y}=(f^c(\underline{x}),f^s(\underline{x}))_T$ خواهد بود که در آن:

(۵)

$$\begin{cases}f^c(\underline{x})=a_0+a_1x_1+\ldots+a_nx_n\\ f^s(\underline{x})=s_0+s_1|x_1|+\ldots+s_n|x_n|\end{cases}$$

به بیان دیگر

$$\mu(x)=\tilde{Y}(y)=\begin{cases}1-\dfrac{f^{c}(\underset{-}{x})-y}{f^{s}(\underset{-}{x})} & f^{c}(\underset{-}{x})-f^{s}(\underset{-}{x})\le y\le f^{c}(\underset{-}{x})\\ 1-\dfrac{y-f^{c}(\underset{-}{x})}{f^{s}(\underset{-}{x})} & f^{c}(\underset{-}{x})<y\le f^{c}(\underset{-}{x})+f^{s}(\underset{-}{x})\end{cases} \qquad (۶)$$

برای اینکه بتوانیم بر پایه مجموعه‌ای از مشاهدات پارامترهای مدل رگرسیون فازی یعنی ضرایب فازی مدل را برآورد کنیم باید ملاک‌ها و معیارهایی در نظر بگیریم. نخست اینکه مقدار عضویت هر y_i در $\tilde{Y}_j$(خروجی مدل به ازای $\underset{-}{x}_j$)عدد بزرگی باشد. به عبارت دیگر خروجی فازی $\tilde{Y}$ برای تمامی مقادیر $\tilde{Y}_j$ (j=1,2,…,m) دارای درجه عضویتی دست‌کم به بزرگی h باشد. یعنی:

$$\tilde{Y}_j(y_i)\ge h \qquad 0\le h\le 1 \qquad (۷)$$

دوم اینکه ابهام(که برابر با مجموع گسترش‌های تکی و مربوط به هریک از پارامترهای فازی مدل است،) در پیش‌بینی براساس مدل حداقل باشد. یعنی ضرایب فازی $\tilde{A}_i$ به گونه‌ای باشد که ابهام خروجی فازی مینیمم گردد. هرچه پهنای یک عدد فازی بیشتر باشد ابهام آن نیز بیشتر است. برای رسیدن به هدف فوق مجموع پهناهای خروجی‌های فازی $\tilde{Y}_j$ مربوط به کلیه مجموعه داده‌ها را حداقل می‌کنیم. با توجه به رابطه‌های بالا به یک برنامه‌ریزی خطی به صورت زیر دست پیدا می‌کنیم (۲و۱۸).

(۸)

$$\begin{aligned}&\min \qquad Z=\sum_{j=1}^{m}2\left[s_0+\sum_{i=1}^{n}s_i|x_{ji}|\right]\\ &s.t \quad 1-\frac{f^{c}(\underset{-}{x})-y}{f^{s}(\underset{-}{x})}\ge h\Rightarrow(1-h)f^{s}(\underset{-}{x}_j)-f^{c}(\underset{-}{x}_j)\ge -y_j\\ &\qquad 1-\frac{y-f^{c}(\underset{-}{x})}{f^{s}(\underset{-}{x})}\ge h\Rightarrow(1-h)f^{s}(\underset{-}{x}_j)+f^{c}(\underset{-}{x}_j)\ge y_j\\ &\qquad j=1,2,\dots,m \qquad s_i>0\end{aligned}$$

که در آن x_{ji} بیانگر مشاهده jام برای متغیرiام است. برای هر مشاهده دو محدودیت وجود دارد. با استفاده از روش‌های برنامه‌ریزی خطی تابع هدف z را با توجه به ۲m محدودیت تولید شده حداقل می‌کنیم. با حل مدل برنامه‌ریزی خطی فوق، ضرایب رگرسیون فازی بدست می‌آید. با قراردادن مقادیر بدست‌آمده برای ضرایب در معادله رگرسیون، مقدار متغیر خروجی به صورت فازی تعیین می‌شود. روش‌های مختلفی جهت غیرفازی سازی متغیر فازی و تبدیل آن به حالت قطعی وجود دارد که می‌توان به روش میانه حداکثر، روش مرکز مجموع‌ها و روش مرکز سطح اشاره کرد.

مدل رگرسیون انباشته فازی

مدل ARIMA مدل پیش‌بینی دقیقی برای دوره‌های کوتاه‌مدت است. اما دارای تعداد زیاد داده‌های گذشته (حداقل۵۰ و ترجیحاً ۱۰۰ یا بیشتر) می‌باشد. در صورتی که امروزه باید موقعیتهای آینده را با استفاده از داده‌های کم و در بازه زمانی کوتاه‌مدت پیش‌بینی کرد. مدل رگرسیون فازی یک مدل پیش‌بینی بازه‌ای مناسب در شرایط داده‌های قابل حصول‌کم است. در این قسمت با بهره‌گیری از مزیت های رگرسیون فازی و مدل ARIMA به پیش‌بینی شاخص قیمت با به کارگیری مدل رگرسیون انباشته فازی و برطرف‌کردن محدودیت‌های موجود در روش‌های رگرسیون فازی و ARIMA می‌پردازیم. پارامترهای مدل ARIMA قطعی‌اند، در صورتی که در این روش به جای پارامترهای قطعی، پارامترهای فازی که به شکل اعداد فازی مثلثی است به کار گرفته شده‌است. با استفاده از پارامترهای فازی نیاز به داده‌های گذشته کاهش می‌یابد (۱).

$$\tilde{B}_p(L)Y_t^{*}=\tilde{\theta}_q(L)u_t \qquad (۹)$$
$$Y_t^{*}=(1-L)^{d}(Y_t-\alpha)$$
$$\tilde{Y}_t^{*}=\tilde{\beta}_1Y_{t-1}^{*}+\tilde{\beta}_2Y_{t-2}^{*}+\dots+\tilde{\beta}_pY_{t-p}^{*}+u_t-\tilde{\theta}_{p+1}u_{t-1}-\tilde{\theta}_{p+2}u_{t-2}-\dots-\tilde{\theta}_{p+q}u_{t-q}$$

حال معادله بالا بصورت زیر تبدیل می شود:

$$\tilde{Y}_t^{*}=\tilde{A}_1Y_{t-1}^{*}+\tilde{A}_2Y_{t-2}^{*}+\dots+\tilde{A}_pY_{t-p}^{*}+u_t-\tilde{A}_{p+1}u_{t-1}-\tilde{A}_{p+2}u_{t-2}-\dots-\tilde{A}_{p+q}u_{t-q} \qquad (۱۰)$$

با استفاده از پارامتر های فازی $\tilde{A}$ به صورت اعداد فازی مثلثی و تابع عضویت Y^{*} مطابق رابطه (۱۱) خواهد بود:

که h سطح آستانه‌ای برای میزان توابع عضویت تمامی مشاهدات است ($\tilde{Y}_{(Y^*)}\ge h$) و S و محدودیت‌ها مطابق (۱۲) محاسبه می‌شود.

در نهایت مدل رگرسیون انباشته فازی بصورت رابطه (۱۳) به دست می‌آید:

علاوه براین، در این مطالعه متدولوژی ارائه شده توسط ایشیبوچی و تاناکا برای شرایطی که دامنه پیش‌بینی وسیع می‌شود به کارگرفته شده‌است (۱۴). در این مدل وقتی که دامنه مدل رگرسیون انباشته فازی وسیع می‌شود، داده‌های حد بالا و پایین مدل حذف خواهند شد. به منظور ساختن مدلی شامل همه شرایط ممکن رگرسیون انباشته فازی، اگر مجموعه داده‌ها شامل تفاوت‌های مشخص یا موارد خارج از محدوده‌اند، c_jها بسیار گسترده خواهند شد. طبق نظرات ایشیبوچی داده‌های اطراف مرزهای بالا و پائین مدل حذف می‌شود و سپس مدل

مجدداً فرمولبندی می‌شود (۱۷).

(۱۱)

$$\tilde{Y}_{(Y^*)}=\begin{cases}1-\dfrac{\left|Y_t^*-\sum_{i=1}^{p}a_iY_{t-i}^*-u_t+\sum_{i=p+1}^{p+q}a_iu_{t+p-i}\right|}{\sum_{i=1}^{p}s_i\left|Y_{t-i}^*\right|-u_t+\sum_{i=p+1}^{p+q}s_i\left|u_{t+p-i}\right|} & for\quad Y_t^*\neq 0,u_t\neq 0\\ 0 & otherwise\end{cases}$$

(۱۲)

$$\begin{aligned}Min\quad & S=\sum_{i=1}^{p}\sum_{t=1}^{k}s_i\left|\beta_i\right|\left|Y_{t-i}^*\right|+\sum_{i=p+1}^{p+q}\sum_{t=1}^{k}s_i\left|\theta_{i-p}\right|\left|u_{t+p-i}\right|\\ S.t\quad & (1-h)(\sum_{i=1}^{p}s_i\left|Y_{t-i}^*\right|+\sum_{i=p+1}^{p+q}s_i\left|u_{t+p-i}\right|)+(\sum_{i=1}^{p}a_iY_{t-i}^*+u_t-\sum_{i=p+1}^{p+q}a_iu_{t+p-i})\geq Y_t^*\\ & (1-h)(\sum_{i=1}^{p}s_i\left|Y_{t-i}^*\right|+\sum_{i=p+1}^{p+q}s_i\left|u_{t+p-i}\right|)-(\sum_{i=1}^{p}a_iY_{t-i}^*+u_t-\sum_{i=p+1}^{p+q}a_iu_{t+p-i})\geq -Y_t^*\\ & s_i\geq 0\qquad t=1,2,\ldots,k\qquad for\quad i=1,2,\ldots,p+q\end{aligned}$$

(۱۳)

$$\tilde{Y}_t^*=(a_1,s_1)Y_{t-1}^*+\ldots+(a_p,s_p)Y_{t-p}^*+u_t-(a_{p+1},s_{p+1})u_{t-1}-\ldots-(a_{p+q},s_{p+q})u_{t-q}$$

ارزیابی مدل ها

به منظور ارزیابی عملکرد مدل می‌توان از معیارهای مربوطه استفاده کرد. در این تحقیق از معیارهای خطای ریشه متوسط مربعات(RMSE) و ضریب تعیین(R^2)جهت این امر استفاده شده‌است. این آماره‌ها با روابط زیر محاسبه می‌شود:

(۱۴)

$$RMSE=\sqrt{\sum_{k=1}^{N}(X_K-Y_K)^2/N}$$

(۱۵)

$$R^2=\left[\frac{\sum_{K=1}^{N}(X_K-\bar{X})(Y_K-\bar{Y})}{\sum_{K=1}^{N}(X_K-\bar{X})^2\sum_{K=1}^{N}(Y_K-\bar{Y})^2}\right]^2$$

که در آن N = تعداد مجموعه داده ها، X= مقدار خروجی محاسبه شده و Y= مقدار خروجی اندازه‌گیری شده می‌باشد.

انتخاب مدل بهینه

مدلی که دارای کمترین مقدارRMSE باشد مدل برتری نسبت به دیگر مدل‌هاست. همچنین مدلی که دارای بیشترین مقدار برای آمارهR^2 باشد برتر از دیگر مدل‌هاست. حال برای استفاده از این معیارها، روش مرکز سطح برای تبدیل یک عدد فازی به یک عدد کلاسیک بکار برده شده که به صورت رابطه (۱۶) بیان می‌شود: که A یک عدد کلاسیک است.

نتایج و بحث

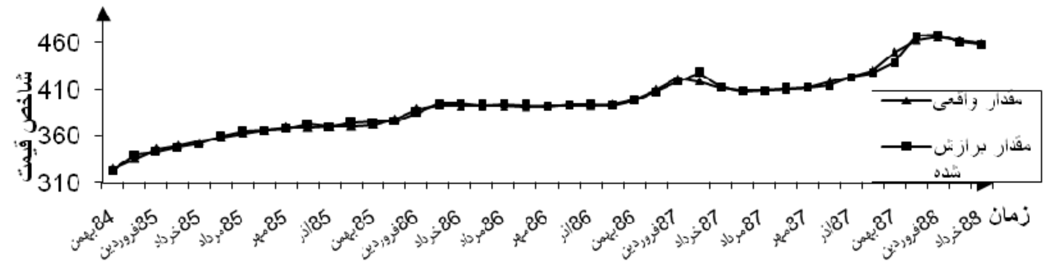
در مدلهای ARIMA حداقل پنجاه و ترجیحاً یکصد یا بیشتر مشاهده به کار گرفته می‌شود. در این مقاله با بکارگیری نرم‌افزار Microfit از۲۰۴ داده ماهانه شاخص قیمت گوشت گوسفند (از فروردین سال ۱۳۷۲ تا اسفند ۱۳۸۸) استفاده شده‌است که از این تعداد ۱۹۵ داده برای تخمین مدل و ۹ داده آخر برای ارزیابی مدل مورد استفاده قرار گرفته‌است. مدل برازش شده توسط این نرم‌افزار(ARIMA(5,1,0 در رابطه (۱۷) مطرح شده‌است:

(۱۶)

$$A=\frac{\int\mu(x).x.dx}{\int\mu(x).dx}$$

(۱۷)

$$Y_t^*=0.6373+0.7894Y_{t-1}^*-0.2513Y_{t-2}^*+0.2115Y_{t-11}^*+0.3002Y_{t-12}^*-0.2867Y_{t-13}^*+u_t$$



شکل ۲- مقدار واقعی و مقدار برازش شده مدل میانگین متحرک خود رگرسیون انباشته

که Y_t^* تفاضل مرتبه اول Y_t می‌باشد. شکل(۲) مقادیر واقعی و مقادیر برازش شده مدل میانگین متحرک خودرگرسیون انباشته را نشان می‌دهد.

در این مقاله از مدل حاصله به منظور پیش‌بینی دامنه تغییرات شاخص قیمت (قیمت گوشت گوسفند) استفاده شده‌است. از ۵۰ داده مورد نظر (بهمن ماه ۱۳۸۴ تا اسفند ماه ۱۳۸۸)، ۴۱ مشاهده برای تخمین مدل رگرسیون فازی در نظر گرفته شده‌است و ۹ مشاهده برای ارزیابی و آزمون مدل بکار رفته است. با توجه به مدل بدست آمده در روش ARIMA مدل رگرسیون فازی را بصورت رابطه (۱۸) در نظر گرفته‌ایم.

$$\widetilde{Y} = \widetilde{A}_0 + \widetilde{A}_1 Y_{t-1} + \widetilde{A}_2 Y_{t-2} + \widetilde{A}_3 Y_{t-3} + \widetilde{A}_4 Y_{t-11} + \widetilde{A}_5 Y_{t-12} + \widetilde{A}_6 Y_{t-13} + \widetilde{A}_7 Y_{t-14} \qquad (۱۸)$$

در این روش با توجه به تعداد مشاهدات(m=۴۱)؛ ۲m=۸۲ محدودیت داریم و تابع هدف بصورت زیر است(h=0).با استفاده از نرم‌افزار WinQSB برنامه به دست آمده را حل نموده و پارامترهای a_i و s_i به دست آمدند که معادله رگرسیون فازی (۱۹) در نهایت حاصل شد.

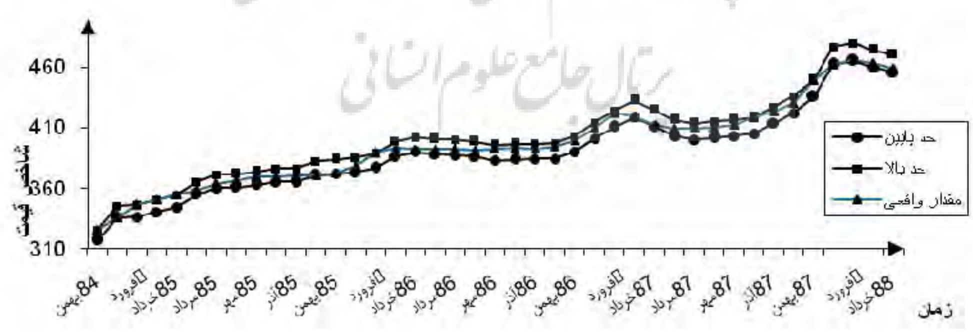
همانطور که در شکل (۳) نشان داده شده است، مقادیر واقعی در فواصل فازی قرار گرفته است.

حال با استفاده از نرم‌افزار برنامه‌ریزی خطی پارامترهای رگرسیون انباشته فازی را به دست آورده (h=0) و نتیجه آن معادله (۲۰) شده‌است:

در شکل (۴) مقادیر واقعی و فواصل فازی نمایش داده شده‌است. طول فواصل فازی بسیار وسیع شده و مدل رگرسیون انباشته فازی فواصل مناسبی را به دست نمی‌دهد.

با حذف محدودیت‌های مربوط به مشاهداتی که خارج از فواصل فازی قرار گرفته‌اند و سپس تکرار مراحل بالا به معادله (۲۱) دست پیدا کرده‌ایم:

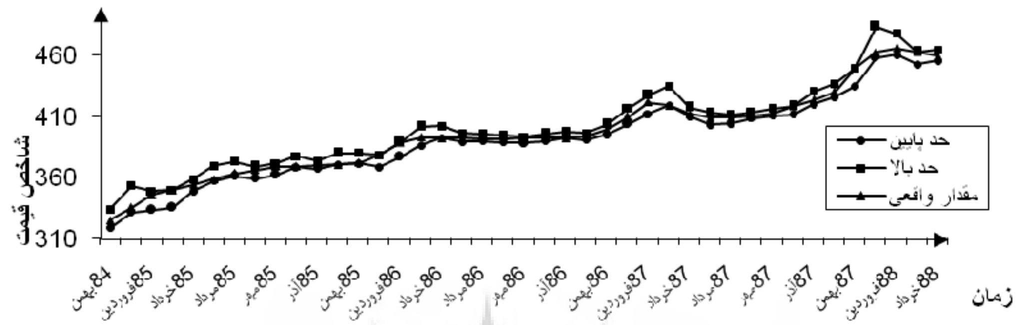
$$\widetilde{Y} = -49.2558 + (1.3560\,,\,0.0049)\,Y_{t-1} - 0.2431\,Y_{t-2} + 0.0485\,Y_{t-3} + (0.6957\,,\,0.0120)\,Y_{t-11} - 0.3514\,Y_{t-12} - 0.3588\,Y_{t-13} - 0.0319\,Y_{t-14} \qquad (۱۹)$$



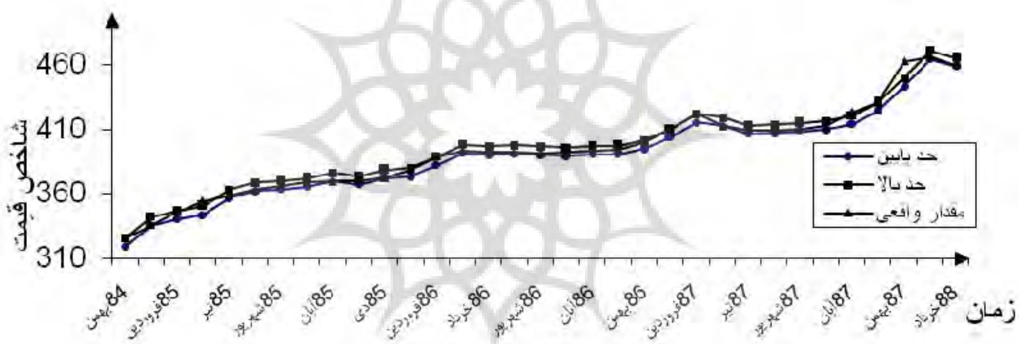
شکل۳- مقادیر واقعی و حد بالا و پایین مقادیر بدست آمده از رگرسیون فازی

$$\tilde{Y}^{*} = (1.5176\,,2.3629) + (0.8782\,,0.4445)\,Y^{*}_{t-1} + (-0.0659\,,0.003)\,Y^{*}_{t-2} + (0.3587\,,0.1727)\,Y^{*}_{t-11} + 0.3167\,Y^{*}_{t-12} - 0.7370\,Y^{*}_{t-13} \qquad (۲۰)$$

$$\tilde{Y}^{*} = (1.2383\,,3.2836) + 0.7339\,Y^{*}_{t-1} - 0.34339\,Y^{*}_{t-2} + 0.1365\,Y^{*}_{t-11} + 0.2933\,Y^{*}_{t-12} - 0.2029\,Y^{*}_{t-13} \qquad (۲۱)$$



شکل٤- مقادیر واقعی و حد بالا و پایین حاصل از رگرسیون انباشته فازی



شکل۵- مقادیر واقعی و حد بالا و پایین حاصل از رگرسیون انباشته فازی بعد از حذف

نتایج مدل نهایی در شکل (۵) آورده شده‌است.

مقایسه شکل‌های رسم شده برای روش‌های قبل نشان می‌دهد که عرض بازه حاصله از روش ترکیبی نسبت به دیگر روش‌ها کمتر بوده و نتایج مطلوبتری را ارائه کرده است.

همچنین با استفاده از مدل‌های ارائه شده، مقادیر آینده متغیر وابسته پیش‌بینی شده‌است. حال از دو آماره جذر میانگین مربعات خطا RMSE و ضریب تعیین R^2 برای بررسی این مقادیر بدست آمده از روش‌های مختلف پرداخته‌ایم. جدول (۱) مقادیر مربوط به دو آماره را برای سه مدل مطرح شده نشان می‌دهد.

برای انتخاب بهترین روش از دو آماره ضریب تعیین و جذر میانگین مربعات خطا استفاده شده است مقایسه مقادیر جدول (۱) نشان می‌دهد که مدل رگرسیون انباشته بعد از حذف دارای کمترین مقدار برای آماره RMSE می‌باشد. همچنین مقدار R^2 در این روش نسبت به دیگر روشها بیشتر است. بنابراین بهترین روش برای پیش‌بینی شاخص قیمت روش رگرسیون انباشته فازی می‌باشد.

نتیجه گیری

در حال حاضر در بسیاری از مراکز تحقیقات اقتصادی داخل کشور و مراکز مطالعاتی سازمانها و همچنین دفاتر مطالعات استراتژیکی که ملزم به پیش‌بینی می‌باشند از مدل‌های ساختاری برای پیش‌بینی استفاده می‌گردد. هر چند که شاید بسیار دقیق بودن میزان پیش‌بینی در برخی موارد از اهمیت چندانی برخوردار نباشد ولی مسلماً پیش‌بینی‌های کوتاه‌مدت برای بسیاری از متغیرها نزد سیاست‌گزاران از اهمیت بالایی برخوردار است.

جدول۱ – مقایسه روش‌های ARIMA، رگرسیون فازی و رگرسیون انباشته فازی (قبل و بعد از حذف)

معیار ها	ARIMA	رگرسیون فازی	رگرسیون انباشته فازی(قبل از حذف)	رگرسیون انباشته فازی(بعد از حذف)
RMSE	۵/۳۸	۳/۹۸	۳/۴۲	۲/۹۴
R^2	۰/۰۰۸	۰/۰۰۵	۰/۰۱۰	۰/۰۳۲

این تحقیق منظور مقایسه عملکرد مدل ARIMA و رگرسیون فازی و رگرسیون انباشته فازی برای پیش‌بینی ماهانه قیمت انجام گرفته‌است. نتایج حاصله بیانگر آن است که مدل میانگین متحرک خودرگرسیون انباشته فازی توانایی انجام یک پیش‌بینی مناسب را داشته و در نتیجه می‌توان از مدل رگرسیون انباشته فازی به عنوان ابزاری دقیق‌تر برای پیش‌بینی متغیرهای اقتصادی در کنار دیگر روشها بهره جست. همچنین مدل پیشنهادی، برای تصمیم‌گیرندگان بهترین و بدترین حالت ممکن را نیز فراهم می‌سازد و با فازی در نظر گرفتن خروجی‌ها اولاً از فروض مربوط به جمله خطا رهایی می‌یابد. ثانیاً به داده‌های کمتری نسبت به مدل ARIMA معمولی نیاز داشته و مدل مناسبی به منظور پیش‌بینی در شرایط داده‌های قابل حصول کم از لحاظ کمی و کیفی خواهد بود.

امروزه می‌توان با استفاده از نتایج حاصل از مطالعات پیش‌بینی قیمت، از قبل نسبت به اقدامات لازم برای ایجاد تعادل در بازار محصولات از طریق واردات یا اعطای مجوز صادرات، برنامه‌ریزی‌های لازم را انجام دهیم. با توجه به توانایی رگرسیون انباشته فازی در پیش‌بینی قیمت محصولات کشاورزی، استفاده از این روش به وسیله دستگاه‌های مسئول در برنامه‌ریزی و سیاست‌گزاری توصیه می‌شود.

منابع

۱- خاشعی م. بیجاری م. ۱۳۸۶. به کارگیری مدل میانگین متحرک خود رگرسیون انباشته فازی به منظور پیش‌بینی نرخ ارز. استقلال(ویژه نامه: روش‌های عددی در مهندسی) سال ۲۶ شماره ۲: ۷۵-۶۷.

۲- طاهری م. ماشین چی م. ۱۳۸۷. مقدمه‌ای بر احتمال و آمار فازی. انتشارات دانشگاه شهید باهنر کرمان.

۳- طاهری م. محمدی ح. ۱۳۸۴. برازش توابع انتقالی خاک با استفاده از رگرسیون فازی. مجله علوم و فنون کشاورزی و منابع طبیعی. سال نهم. شماره ۲: ۶۰-۵۱.

۴- طرازکار م. ۱۳۸۶. پیش بینی قیمت برخی محصولات زراعی در استان فارس: کاربرد شبکه عصبی مصنوعی. مجله علوم وفنون کشاورزی و منابع طبیعی سال ۱۱. شماره اول(ب):۵۱۱-۵۰۱.

۵- کوره پزان ا. ۱۳۸۴. اصول تئوری مجموعه‌های فازی و کاربردهای آن در مدلسازی مسایل مهندسی آب. انتشارات جهاد دانشگاهی واحد صنعتی امیرکبیر.

۶- گجراتی د. ۱۳۷۷. مبانی اقتصادسنجی. دانشگاه تهران. موسسه انتشارات و چاپ تهران. ویرایش دوم.

۷- گیلان‌پور الف. و ن کهزادی. ۱۳۷۶. پیش‌بینی قیمت برنج در بازارهای بین‌المللی با استفاده از الگوی خود رگرسیون میانگین متحرک. فصلنامه اقتصاد کشاورزی و توسعه۸: ۲۰۰-۱۸۹.

۸- یا علی جهرمی م. محمدی ح و ز فرجزاده. ۱۳۸۸. پیش‌بینی قیمت چغندر در ایران. مجله چغندرقند جلد ۲۵. شماره۱: ۱۱۱-۹۷.

۹- مجاوریان م. و الف. امجدی. ۱۳۷۸. مقایسه روش‌های معمول با تابع مثلثاتی در قدرت پیش‌بینی سری زمانی قیمت محصولات کشاورزی همراه با اثرات فصلی: مطالعه مورد مرکبات. فصلنامه اقتصاد کشاورزی و توسعه ۲۵: ۴۳-۶۲.

۱۰- مشیری س. ۱۳۸۰. پیش‌بینی تورم ایران با استفاده از مدل ساختاری. سری زمانی و شبکه عصبی مصنوعی. مجله تحقیقات اقتصادی۵۸: ۱۸۴-۱۴۷.

11- Azadeh A., Khakestani M., and Saberi M., 2009. A flexible fuzzy regression algorithm for forecasting oil consumption estimation. Energy Policy 37: 5567–5579.

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

13- Hui-Kuang Yu T., and Huarng K. H., 2010. A neural network-based fuzzy time series model to improve forecasting . Expert Systems with Applications 37: 3366–3372.

14- Ishibuchi H., and Tanaka H., 1988. "Interval regression Analysis Based on Mixed ۰-۱ Integer programming problem," J. Japan Soc., Ind. Eng, vol.40, No.5, pp. 312-319.

15- Khashei M., Hejazi S. R. and Bijari M., 2008. A new hybrid artificial neural networks and fuzzy regression model for time series forecasting. Fuzzy Sets and Systems 159: 769 – 786.

16- Kohzadi N., Boyd M. S., Kaastra J., Kermanshahi B. S. and Scuse D., 1995. Neural network for rorecasting: an introduction. Can. J. Agric. Econ. 43: 463-474.

17- Tanaka H., and Ishibuchi H., 1992. "Possibility Regression Analysis Based on Linear programming," in: J.Kacprzyk, M.Fedrizzi(Eds), Fuzzy Regression Analaysis, Omnitech press. Warsaw and Physica-Verlag, Heidelberg, pp: 47-60.

18- Wang C. C.,2011. A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. Expert Systems with Applications 38: 9296–9304.

19- Watada, 1992."Fuzzy Time Series Analysis and Forecasting of scales Volume," in : Kacprzyk J., Fedrizzi M.(Eds.), Fuzzy Regression Analysis, Omintech Press. Warsaw and physica- Verlag, Heidelberg. pp:211-227.

20- Yen K.K, Choshary S., and Raig G. 1999. A linear model using traanular Fuzzy number coefficients. Fuzzy sets and systems 106:167-177.

21- Zadeh L.A., 1965. Fuzzy sets and Information control 8: 338-353.