رگرسیون خطی فازی با استفاده از برنامه‌ریزی آرمانی

حسن حسن‌پور [۱]، حمیدرضا ملکی [۲]، محمدعلی یعقوبی [۳]

hassanhassanpur@yahoo.com, maleki@sutech.ac.ir, yaghoobi@mail.uk.ac.ir

چکیده

با وجود این که زمان زیادی از پیدایش رگرسیون فازی نمی‌گذرد، این مبحث توسعه بسیار زیادی یافته است. در این مقاله سعی شده است برخی از روش‌های رگرسیون خطی فازی به اختصار مرور شود. سپس شیوه‌ای جدید برای محاسبه ضرایب مدل رگرسیون خطی فازی مبتنی بر رگرسیون کمترین قدرمطلق خطا و با استفاده از یک فاصله وزن‌دار بین اعداد فازی مثلثی ارائه شده است. در روش ارائه شده برای محاسبه ضرایب رگرسیون از برنامه‌ریزی آرمانی استفاده می‌شود. نتایج به‌دست آمده برتری این روش را نسبت به برخی از روش‌های موجود از نظر خطای برازش و حساسیت نسبت به داده‌های پرت نشان می‌دهد.

واژه‌های کلیدی: رگرسیون خطی فازی، برنامه‌ریزی آرمانی، برنامه‌ریزی خطی، فاصله وزن‌دار، عدد فازی.

۱ مقدمه

تحلیل رگرسیون، ابزاری آماری برای یافتن رابطه‌ای بین دو یا چند متغیر بر مبنای تعدادی مشاهده است. از این رابطه برای پیش‌بینی مقدار متغیر وابسته (خروجی یا پاسخ) به ازای مقادیر معینی از متغیر(های) مستقل (ورودی) استفاده می‌شود. مجموعه‌ای شامل n داده به‌صورت $(x_{i1}, x_{i2}, \ldots, x_{ip}, y_i)$ به ازای $i = 1, 2, \ldots, n$ را درنظر بگیرید که در آن x_{ij} مقدار متغیر مستقل x_j و y_i مقدار متغیر وابسته y در مشاهده i—ام است. هدف رگرسیون خطی، برازش مدلی خطی به‌صورت

$$Y = a_0 + a_1 x_1 + \cdots + a_p x_p = \sum_{j=0}^{p} a_j x_j, \qquad (1)$$

که در آن $x_0 = 1$، بر داده‌های فوق است به‌گونه‌ای که پاسخ‌های برآورد شده از مدل یعنی $Y_i = \sum_{j=0}^{p} a_j x_{ij}$ کمترین فاصله ممکن را تا پاسخ‌های داده شده y_i داشته باشند. در بسیاری از مسایل دنیای واقعی با داده‌های دقیق سر و کار نداریم. بسیاری اوقات از واژه‌هایی که یک مفهوم یا مقدار کاملا مشخصی را در ذهن تداعی می‌کنند استفاده نمی‌کنیم. برای مثال، از عباراتی مانند «تقریباً»، «کم و بیش»، «کوتاه»، «خیلی خوب»، و«نسبتاً خوب» استفاده می‌کنیم یا از کمیت‌های نادقیقی مانند «حدودا ۲» به جای عدد ۲، و «اعداد خیلی بزرگ» صحبت به میان می‌آوریم. در چنین مواردی، با داده‌های مبهم یا اصطلاحاً فازی سر و کار داریم. در مبحث رگرسیون نیز اگر بخشی یا تمام مشاهدات به جای اعداد حقیقی، اعداد فازی باشند، باید از رگرسیون فازی استفاده کرد. اگر ورودی‌ها اعداد حقیقی وپاسخ‌ها اعداد فازی باشند، هدف رگرسیون خطی فازی برازش مدلی به‌صورت زیر بر داده‌ها است

$$\tilde{Y} = \tilde{A}_0 \oplus \tilde{A}_1 x_1 \oplus \ldots \oplus \tilde{A}_p x_p \qquad (2)$$

که در آن ضرایب رگرسیون $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_p$ اعداد فازی

[۱] بخش ریاضی، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان
[۲] دانشکده علوم پایه، دانشگاه صنعتی شیراز
[۳] بخش آمار، دانشکده ریاضی و کامپیوتر، دانشگاه شهید باهنر کرمان

می‌باشند و از نماد $\oplus$ برای جمع اعداد فازی استفاده شده است. درصورتی که هم ورودی‌ها و هم پاسخ‌ها اعداد فازی باشند، هدف رگرسیون خطی فازی برازش مدلی به یکی از صورت‌های زیر بر داده‌ها است

$$\tilde{Y} = a_0 \oplus a_1\tilde{x}_1 \oplus \ldots \oplus a_p\tilde{x}_p, \qquad (۳)$$

$$\tilde{Y} = \tilde{A}_0 \oplus a_1\tilde{x}_1 \oplus \ldots \oplus a_p\tilde{x}_p, \qquad (۴)$$

$$\tilde{Y} = \tilde{A}_0 \oplus \tilde{A}_1\tilde{x}_1 \oplus \ldots \oplus \tilde{A}_p\tilde{x}_p. \qquad (۵)$$

در مدل (۳)، ضرایب رگرسیون همگی غیرفازی‌اند، در مدل (۴)، $\tilde{A}_0$ فازی ولی سایر ضرایب غیرفازی‌اند و در مدل (۵)، تمام ضرایب رگرسیون اعداد فازی‌اند. مقالات بسیار زیادی پیرامون رگرسیون خطی فازی توسط افراد مختلف نوشته شده است که در اکثر آنها خروجی‌ها و ضرایب رگرسیون اعداد فازی مثلثی متقارن و ورودی‌ها اعداد حقیقی فرض شده‌اند. در بخش بعد به برخی از آنها به اختصار اشاره خواهد شد. یک عدد فازی مثلثی مانند $\tilde{A}$، با یک سه‌تایی مرتب $\tilde{A} = (a, \alpha, \beta)$ نمایش داده می‌شود که در آن عدد حقیقی a مرکز و اعداد مثبت α و β به ترتیب پهناهای چپ و راست آن نام دارند. تابع عضویت $\tilde{A}$ عبارت است از

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-(a-\alpha)}{\alpha} & a-\alpha \le x \le a \\ \frac{(a+\beta)-x}{\beta} & a \le x \le a+\beta \\ 0 & \text{سایر جاها} \end{cases}$$

بازه‌های $[a-(1-h)\alpha, a+(1-h)\beta]$ که $0 < h \le 1$ و $[a-\alpha, a+\beta]$ به ترتیب $h-$ برش و تکیه‌گاه $\tilde{A}$ نام دارند. اگر $\alpha = \beta$، $\tilde{A}$ یک عدد فازی مثلثی متقارن نام دارد و با $\tilde{A} = (a, \alpha)$ نمایش داده می‌شود. با استفاده از اصل گسترش اعداد فازی [۲۱]، برای دو عدد فازی $\tilde{A}_1 = (a_1, \alpha_1, \beta_1)$ و $\tilde{A}_2 = (a_2, \alpha_2, \beta_2)$ و عدد حقیقی r داریم

$$\tilde{A}_1 \oplus \tilde{A}_2 = (a_1 + a_2, \alpha_1 + \alpha_2, \beta_1 + \beta_2), \qquad (۶)$$

$$r\tilde{A}_1 = \begin{cases} (ra_1, r\alpha_1, r\beta_1) & r \ge 0, \\ (ra_1, |r|\beta_1, |r|\alpha_1) & r < 0. \end{cases} \qquad (۷)$$

در بخش بعد برخی از روش‌های موجود برای برآورد ضرایب مدل رگرسیون خطی به اختصار مرور شده و به بعضی از اشکالات آنها اشاره شده است. در نهایت شیوه‌ای جدید برای محاسبه ضرایب رگرسیون ارائه شده و خواص آن مورد بررسی قرار گرفته است.

۲ مروری بر برخی از روش‌های رگرسیون خطی فازی

دو دیدگاه کلی برای مواجهه با مساله رگرسیون خطی فازی وجود دارد: دیدگاه کمترین مربعات و دیدگاه برنامه‌ریزی ریاضی. استفاده از روش کمترین مربعات در رگرسیون خطی فازی اولین بار توسط کلمینس [۴] [۲] در سال ۱۹۸۷ ارائه شد و توسط دیگران توسعه یافت (برای مثال به [۱، ۲، ۴، ۱۰، ۱۱، ۱۲] مراجعه نمایید). استفاده از برنامه‌ریزی خطی در رگرسیون خطی فازی اولین بار توسط تاناکا[۵] و همکاران [۲۰] در سال ۱۹۸۲ ارائه شد. آنها یک مدل برنامه‌ریزی خطی برای محاسبه ضرایب مدل رگرسیون (۲) با این فرض که خروجی و ضرایب مدل اعداد فازی مثلثی متقارن باشند ارائه کردند که تابع هدف آن کمینه کردن مجموع پهناهای ضرایب بود و محدودیت‌های آن باعث می‌شد که به ازای مقدار مشخصی از h، $h-$برش پاسخ‌های برآورد شده از مدل شامل $h-$برش پاسخ‌های داده شده‌ی متناظرشان باشند. از جمله نواقص این مدل این است که جواب به مقیاس متغیرهای ورودی x_j وابسته است [۷]، بسیاری از پهناهای ضرایب مدل صفر به‌دست می‌آیند [۷]، و دیگر این که مدل نسبت به داده‌های پرت بسیار حساس است [۱۵].

[۵] Tanaka

تاناکا [۳] مدل قبلی را به این صورت اصلاح کرد که به جای مجموع پهناهای ضرایب، مجموع پهناهای پاسخ‌های برآورد شده را در نظر گرفت و دو مدل برنامه‌ریزی خطی ارائه کرد. یکی از این دو مدل مجموع مذکور را تحت همان شرایط مدل تاناکا و همکارانش کمینه می‌سازد. مدل دوم مجموع مذکور را با این محدودیت‌ها که به ازای مقدار مشخصی از h، h—برش پاسخ‌های برآورد شده از مدل مشمول h—برش پاسخ‌های داده شده‌ی متناظرشان باشند، بیشینه می‌کند. مدل کمینه‌سازی ارائه شده همواره جواب شدنی دارد ولی مدل بیشینه‌سازی ارائه شده گاهی اوقات جواب شدنی ندارد [۳]. علاوه بر این، هر دو مدل نسبت به داده‌های پرت بسیار حساس می‌باشند.

ایراد دیگری که ساویک و پدریچ [۶] [۱۸] بر مدل تاناکا و همکاران گرفتند این بود که مدل مذکور مراکز داده‌ها را در محاسبه ضرایب رگرسیون به حساب نمی‌آورد. لازم به ذکر است که اکثر روش‌های موجود این ضعف را دارند. حال آنکه این مراکز به دلیل داشتن بیشترین درجه عضویت، بیشترین اهمیت را دارند. ساویک وپدریچ برای برطرف کردن این نقیصه یک روش دو مرحله‌ای برای محاسبه ضرایب رگرسیون ارائه کردند. در مرحله اول، مراکز ضرایب به روش کمترین مربعات محاسبه می‌شوند و در مرحله دوم، برای محاسبه پهناهای ضرایب، یک مدل برنامه‌ریزی خطی مشابه مدل تاناکا و همکاران با استفاده از مراکز به‌دست آمده در مرحله اول تشکیل و حل می‌شود.

کیم و بیشو[۷] [۱۱] نقیصه دیگری از مدل تاناکا و همکارانش و همچنین روش دو مرحله‌ای ساویک و پدریسز را مطرح کرده‌اند و آن این که هر دو روش برای محاسبه ضرایب رگرسیون از مجموع پهناهای ضرایب رگرسیون به‌عنوان درجه ابهام مدل رگرسیون استفاده می‌کنند. در نتیجه هرچه پهنای ضرایب به‌دست آمده بیشتر باشد ابهام خروجی‌های برآورد شده توسط مدل بیشتر خواهد بود. به‌عبارت دیگر، خروجی‌های برآورد شده کارایی کمتری خواهند داشت. کیم و بیشو برای رفع این نقیصه روش دیگری ارائه نمودند که در آن برای محاسبه نقاط ابتدا، مراکز و نقاط انتهای ضرایب فازی مثلثی، سه مدل کمترین مربعات معمولی حل می‌شود. در [۶] نشان داده شده است که روش کیم و بیشو در بسیاری از موارد پهناهای منفی برای ضرایب رگرسیون به‌دست می‌دهد، و این نقیصه با اعمال شرایط اضافی بر مدل‌های کمترین مربعات کیم و بیشو برطرف شده‌است [۶].

ساکاوا و یانو[۸] [۱۷] به کمک سه شاخص دبوا و پراد[۹] [۵] برای تساوی اعداد فازی، سه مدل برنامه‌ریزی خطی مختلف برای محاسبه ضرایب هریک از مدل‌های (۲) و (۵) ارائه کرده‌اند. علاوه بر این، سه مدل برنامه‌ریزی خطی دو هدفی نیز توسط این نویسندگان به کمک سه شاخص مذکور ارائه شده است.

حجتی[۱۰] و همکاران [۷] دو مدل برنامه‌ریزی خطی با استفاده از برنامه‌ریزی آرمانی برای محاسبه ضرایب مدل‌های (۲) و (۵) ارائه کرده‌اند. این دو مدل نیز همانند اکثر مدل‌های ارائه شده مراکز اعداد مثلثی را در محاسبه ضرایب رگرسیون به حساب نمی‌آورند.

دو مدل برنامه‌ریزی خطی نیز توسط پیترز[۱۱] [۱۴] و ازلکان و داکستین[۱۲] [۱۳] برای محاسبه ضرایب مدل (۲) ارائه شده

[۶] Savic and Pedrycz
[۷] Kim and Bishu
[۸] Sakawa and Yano
[۹] Dubois and Prade
[۱۰] Hojati
[۱۱] Peters
[۱۲] Özelkan and Ducksetein

است. قیود هر دو مدل شامل اعداد ثابتی است که قبل از حل مدل‌ها باید تعیین و در مدل‌ها گنجانده شوند. اما روشی عملی برای انتخاب مناسب این ثابت‌ها ارائه نشده است. این امر استفاده از این روش‌ها را مشکل می‌کند.

چن[۱۳] [۳] برای کاهش تاثیر داده‌های پرت، تعدادی محدودیت به محدودیت‌های مدل تاناکا و همکاران اضافه کرده است. لازم به ذکر است که همان اشکال ذکر شده در بند قبل در مدل چن نیز وجود دارد. زیرا محدودیت‌های اضافه‌شده شامل ثابتی است که قبل از حل مدل باید تعیین و در مدل گنجانده شود.

یک روش حذفی نیز برای تعیین داده‌های پرت توسط هونگ و یانگ[۱۴] [۸] ارائه شده است. در این روش برای تعیین داده‌های پرت میزان تغییر در تابع هدف مدل در ازای حذف هریک از داده‌ها مورد بررسی قرار می‌گیرد.

ایشیبوچی و نی[۱۵] [۹] برخی از محدودیت‌های استفاده از مدل رگرسیون با ضرایب فازی مثلثی متقارن را ذکر کرده‌اند. به‌دنبال آن برای مدل رگرسیون خطی فازی (۲) با ورودی‌های حقیقی و خروجی و ضرایب فازی مثلثی، ضرایب مثلثی متقارن را به اعداد مثلثی و ذوزنقه‌ای نامتقارن تعمیم داده‌اند.

۳ محاسبه ضرایب رگرسیون با استفاده از برنامه‌ریزی آرمانی

برخی از اشکالات روش‌های مرور شده در بخش ۲ عبارت‌اند از

- اغلب روش‌ها روی داده‌های متقارن متمرکز شده‌اند.
- برخی از روش‌ها نسبت به داده‌های پرت بسیار حساس می‌باشند.
- اکثر روش‌ها مراکز داده‌ها را در محاسبه ضرایب رگرسیون به‌حساب نمی‌آورند.

در این مقاله برای رفع اشکالات فوق، روشی مبتنی بر رگرسیون کمترین قدرمطلق خطا و با استفاده از یک تابع فاصله جدید روی اعداد فازی مثلثی ارائه شده است. روش را برای محاسبه ضرایب مدل (۲) با خروجی و ضرایب فازی مثلثی شرح می‌دهیم. این روش برای محاسبه ضرایب مدل‌های (۲)، (۳) و (۴) با خروجی و ضرایب فازی ذوزنقه‌ای و همچنین برای محاسبه ضرایب مدل (۵) با داده‌های مثلثی قابل کاربرد است.

همانگونه که قبلا اشاره شد، هدف رگرسیون معمولی برازش یک مدل بر تعدادی داده است به‌گونه‌ای که پاسخ‌های برآورد شده از مدل کمترین فاصله ممکن را با پاسخ‌های داده شده داشته باشند. برای این منظور در رگرسیون معمولی تابعی از تفاضلات بین پاسخ‌های برآورد شده و پاسخ‌های داده شده را کمینه می‌کنند. به‌طور مشابه، در رگرسیون فازی هدف برازش یک مدل فازی بر مجموعه‌ای از داده‌های فازی است به‌گونه‌ای که پاسخ‌های فازی برآورد شده از مدل تا حد ممکن به پاسخ‌های فازی داده شده نزدیک باشند. به‌وضوح طبیعی است که انتظار داشته باشیم که نقاط با بیشترین درجه عضویت در پاسخ‌های برآورد شده به نقاط با بیشترین درجه عضویت در پاسخ‌های داده شده، وهمچنین نقاط با درجات عضویت دیگر درپاسخ‌های برآورد شده به نقاط متناظر درپاسخ‌های داده شده نزدیک باشند. بنابراین سعی می‌کنیم توابع عضویت پاسخ‌های داده شده و برآورد شده را تا حد ممکن به هم نزدیک کنیم. چون در این مقاله پاسخ‌های مشاهده شده اعداد فازی مثلثی فرض شده‌اند، سعی برآن است که این کار با نزدیک کردن

[۱۳] Chen
[۱۴] Hung and Yang
[۱۵] Ishibuchi and Nii

مراکز و پهناهای پاسخ‌های برآورد شده به مقادیر متناظر در پاسخ‌های داده شده انجام شود. از این رو فاصله وزن دار دو عدد فازی مثلثی را به‌صورت زیر تعریف می‌کنیم.

تعریف ۱ فرض کنید $\tilde{y} = (y, \alpha_y, \beta_y)$ و $\tilde{Y} = (Y, \alpha_Y, \beta_Y)$ دو عدد فازی مثلثی باشند و $w_1, w_2 \geq 0$ که حداقل یکی از آنها غیر صفر است. فاصله وزن دار $\tilde{y}$ و $\tilde{Y}$ را به‌صورت زیر تعریف می کنیم

$$d^w(\tilde{y}, \tilde{Y}) = w_1|y - Y| + w_2(|\alpha_y - \alpha_Y| + |\beta_y - \beta_Y|).$$

به‌سادگی می‌توان نشان داد که d^w یک تابع فاصله روی $T(\mathbb{R}) \times T(\mathbb{R})$ است که $T(\mathbb{R})$ مجموعه اعداد فازی مثلثی و $\mathbb{R}$ مجموعه اعداد حقیقی است. همچنین در انتهای این بخش به چگونگی انتخاب وزن های w_1 و w_2 اشاره خواهد شد.

تعداد n داده به‌صورت $(x_{i1}, x_{i2}, \ldots, x_{ip}, \tilde{y}_i)$، به ازای $i = 1, 2, \ldots, n$، که در آن ورودی‌های x_{ij} اعدادی حقیقی و پاسخ‌های $\tilde{y}_i$ اعداد فازی مثلثی به‌صورت $\tilde{y}_i = (y_i, \alpha_{y_i}, \beta_{y_i})$ می‌باشند مفروض است. مدل (۲) که در آن به ازای $j = 0, 1, \ldots, p$، $\tilde{A}_j = (a_j, \alpha_j, \beta_j)$ را درنظر بگیرید. بدون آن که از کلیت موضوع کاسته شود می‌توان فرض کرد که به ازای هر i و j، $x_{ij} > 0$. در این صورت با توجه به روابط (۶) و(۷)، i—امین پاسخ برآورد شده عبارت است از

$$\tilde{Y}_i = (\sum_{j=0}^{p} a_j x_{ij}, \sum_{j=0}^{p} \alpha_j x_{ij}, \sum_{j=0}^{p} \beta_j x_{ij}). \qquad (۸)$$

اکنون برای محاسبه ضرایب مدل با توجه به توضیحات قبل مدل برنامه‌ریزی غیرخطی زیر را ارائه می کنیم

$$\begin{array}{l} minimize \quad \sum_{i=1}^{n} d^w(\tilde{y}_i, \tilde{Y}_i) \\ s.t. \quad a_j \in \mathbb{R}, \alpha_j, \beta_j \geq 0, j = 0, 1, \ldots, p. \end{array} \qquad (۹)$$

که $\tilde{Y}_i$ در رابطه (۸) و d^w در تعریف ۱ آمده است. با توجه به تعریف d^w، مدل فوق مبتنی بر رگرسیون کمترین قدر مطلق خطا می‌باشد. ویژگی استفاده از تعریف ۱ این است که با انتخاب متغیرهای انحرافی مناسب [۱۶]، مدل غیرخطی (۹) به‌راحتی به مدل برنامه‌ریزی آرمانی وزن‌دار زیر تبدیل می‌شود

$$\begin{array}{l} GP1(w_1, w_2): \\ minimize \quad \sum_{i=1}^{n}(w_1(n_{ic} + p_{ic}) + \\ w_2(n_{il} + p_{il} + n_{ir} + p_{ir})) \end{array} \qquad (۱۰)$$

$$s.t.$$

$$\sum_{j=0}^{p} a_j x_{ij} + n_{ic} - p_{ic} = y_i, i = 1, \ldots, n \qquad (۱۱)$$

$$\sum_{j=0}^{p} \alpha_j x_{ij} + n_{il} - p_{il} = \alpha_{y_i}, i = 1, \ldots, n \qquad (۱۲)$$

$$\sum_{j=0}^{p} \beta_j x_{ij} + n_{ir} - p_{ir} = \beta_{y_i}, i = 1, \ldots, n \qquad (۱۳)$$

$$n_{ik} \cdot p_{ik} = 0, i = 1, \ldots, n, \ k = l, c, r \qquad (۱۴)$$

$$n_{ik}, \ p_{ik} \geq 0, i = 1, \ldots, n, \ k = l, c, r \qquad (۱۵)$$

$$a_j \in \mathbb{R}, \alpha_j, \beta_j \geq 0, j = 0, 1, \ldots, p. \qquad (۱۶)$$

در این مدل، n_{ic} و p_{ic} به ترتیب انحرافات منفی و مثبت بین مراکز پاسخ‌های برآورد شده و پاسخ‌های داده شده می‌باشند. به همین ترتیب n_{il} و p_{il} به‌ترتیب انحرافات منفی و مثبت بین پهناهای چپ پاسخ‌های برآورد شده و پاسخ‌های داده شده، و در نهایت n_{ir} و p_{ir} به‌ترتیب انحرافات منفی و مثبت بین پهناهای راست پاسخ‌های برآورد شده و پاسخ‌های داده شده می‌باشند. در واقع اگر $\sum_{j=0}^{p} a_j x_{ij} < y_i$، داریم $n_{ic} > 0$ و $p_{ic} = 0$ و چنانچه $\sum_{j=0}^{p} a_j x_{ij} > y_i$، داریم $n_{ic} = 0$ و $p_{ic} > 0$. سایر متغیرهای انحرافی نیز ویژگی مشابهی دارند.

اکنون توجه خود را به بررسی خواص مدل فوق معطوف می‌کنیم. اگر پاسخ‌های داده شده متقارن باشند به ازای هر i داریم $\alpha_{y_i} = \beta_{y_i}$. برای برآورد پاسخ‌های متقارن کافی است

ضرایب مدل رگرسیون را متقارن درنظر بگیریم. یعنی به ازای هر j قرار دهیم $\alpha_j = \beta_j$. در این صورت قیود (۱۲) و (۱۳) معادل خواهند شد و می‌توان یکی از آنها را حذف کرد، که در این صورت تعداد n قید از قیود مدل کاسته می‌شود. علاوه بر این، با توجه به ویژگی روش سیمپلکس [۱۶]، قیود غیرخطی (۱۴) را می‌توان حذف کرد و مدل را به یک مدل برنامه‌ریزی خطی تبدیل نمود که آن را $LP۱(w_۱, w_۲)$ می‌نامیم. وجود وزن های $w_۱$ و $w_۲$ باعث می‌شود که بتوان با تخصیص مقادیر مختلف به آنها بسته به اهمیت نسبی مراکز و پهناهای پاسخ‌ها، جواب‌های مختلفی به‌دست آورد. به این ترتیب که هرچه اهمیت مراکز پاسخ‌ها نسبت به پهناهای آنها بیشتر باشد، به $w_۱$ مقدار بیشتری (در مقایسه با $w_۲$) اختصاص می‌دهیم. برای مثال، اگر اهمیت مراکز دو برابر پهناها تشخیص داده شود می‌توان قرار داد $w_۱ = ۲$ و $w_۲ = ۱$. به‌علاوه، با روش سعی و خطا ممکن است مقادیر وزن‌های مناسب‌تر و نتایج مطلوب‌تری به‌دست آید. فرض کنید مراکز و پهناهای پاسخ‌ها از اهمیت یکسانی برخوردار باشند. قرار می‌دهیم $w_۱ = w_۲ = ۱$ و مدل به‌دست آمده را برای سادگی $GP۱$ و مدل برنامه‌ریزی خطی حاصل از آن را $LP۱$ می‌نامیم. ویژگی مهم این دو مدل این است که چون قیود (۱۱)، (۱۲) و (۱۳) مستقل از هم می‌باشند، هریک از این دو مدل را می‌توان به سه مدل مجزا تجزیه کرد که تعداد قیود هریک به مراتب از تعداد قیود مسایل اصلی کمتر است. این ویژگی به خصوص زمانی که تعداد مشاهدات زیاد باشد از اهمیت ویژه‌ای برخوردار است زیرا تعداد قیود به میزان قابل توجهی کاهش می‌یابد. مدل‌های برنامه‌ریزی خطی حاصل از تجزیه $LP۱$ عبارت‌اند از

$(LP۲)$:

$$minimize \quad \sum_{i=۱}^{n}(n_{ic} + p_{ic})$$

$$s.t. \quad \sum_{j=۰}^{p} a_j x_{ij} + n_{ic} - p_{ic} = y_i, i = ۱, \ldots, n$$
$$a_j \in \mathbb{R} \qquad j = ۰, \ldots, p$$
$$n_{ic}, p_{ic} \geq ۰ \qquad i = ۱, \ldots, n$$

$(LP۳)$:

$$minimize \quad \sum_{i=۱}^{n}(n_{il} + p_{il})$$

$$s.t. \quad \sum_{j=۰}^{p} \alpha_j x_{ij} + n_{il} - p_{il} = \alpha_{y_i}, i = ۱, \ldots, n$$
$$\alpha_j, n_{il}, p_{il} \geq ۰, j = ۰, \ldots, p, i = ۱, \ldots, n$$

$(LP۴)$:

$$minimize \quad \sum_{i=۱}^{n}(n_{ir} + p_{ir})$$

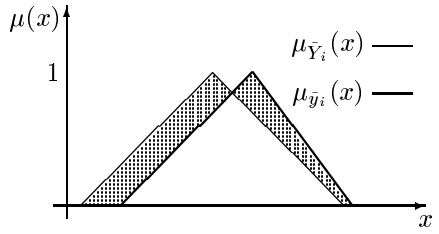
$$s.t. \quad \sum_{j=۰}^{p} \beta_j x_{ij} + n_{ir} - p_{ir} = \beta_{y_i}, i = ۱, \ldots, n$$
$$\beta_j, n_{ir}, p_{ir} \geq ۰, j = ۰, \ldots, p, i = ۱, \ldots, n$$

۴ نتایج عددی

در این بخش دو معیار برای ارزیابی مدل‌های رگرسیون ارائه می‌شود، و سپس با حل مثال‌هایی روش ارائه شده با برخی از روش‌های قبل مقایسه می‌شود. کیم و بیشو [۱۱] برای ارزیابی کارایی مدل رگرسیون خطی فازی با داده‌های مثلثی، از مساحت ناحیه غیر مشترک زیر نمودار توابع عضویت دو عدد فازی مثلثی (ناحیه سایه‌دار در شکل ۱) به‌عنوان فاصله آن دو استفاده کرده‌اند

$$E_i = \int_{S_{\tilde{y}_i} \cup S_{\tilde{Y}_i}} |\mu_{\tilde{y}_i}(x) - \mu_{\tilde{Y}_i}(x)| dx \qquad (۱۷)$$

که در آن $S_{\tilde{y}_i}$ و $S_{\tilde{Y}_i}$ به‌ترتیب تکیه‌گاه‌های اعداد فازی مثلثی $\tilde{y}_i$ و $\tilde{Y}_i$ می‌باشند. ما نیز E_i را به‌عنوان خطای i—امین پاسخ برآورد شده می‌پذیریم. به‌وضوح هرچه E_i کوچکتر باشد توابع عضویت پاسخ‌های برآورد شده و پاسخ‌های داده شده به هم نزدیکترند.



شکل ۱. فاصله بین دو تابع عضویت

مراکز اعداد فازی مثلثی به‌دلیل داشتن بیشترین درجه عضویت از اهمیت ویژه‌ای برخوردارند. برای ارزیابی مدل رگرسیون از نظر میزان نزدیکی مراکز پاسخ‌های برآورد شده و داده شده، پارامتر جدیدی را معرفی می‌کنیم که عبارت است از مجموع فواصل مراکز پاسخ‌های برآورد شده و داده شده

$$\lambda = \sum_{i=1}^{n} |y_i - \sum_{j=0}^{p} a_j x_{ij}| \qquad (۱۸)$$

واضح است که مقدار کوچکتر λ نشان دهنده برازش بهتر مراکز پاسخ‌های برآورد شده بر مراکز پاسخ‌های داده شده است. در ادامه سه مثال عددی ارائه می‌کنیم و نتایج حاصل از چند روش را با هم مقایسه می‌کنیم.

مثال ۱ داده‌های جدول شماره ۱ شامل پنج مقدار یک متغیر مستقل و مقادیر متغیر وابسته نظیرشان را که اعداد فازی مثلثی متقارن می‌باشند درنظر بگیرید. این مثال توسط تاناکا و همکاران [۲۰] ارائه شده است.

جدول ۱.داده‌های مثال ۱

i	x_{i1}	$\tilde{y}_i$
۱	۱	(۸, ۱/۸)
۲	۲	(۶/۴, ۲/۲)
۳	۳	(۹/۵, ۲/۶)
۴	۴	(۱۳/۵, ۲/۶)
۵	۵	(۱۳, ۲/۴)

مقادیر خطای E_i، خطای کل، $TE = \sum_{i=1}^{n} E_i$، و λ برای مدل تاناکا و همکارانش [۲۰] (T)، مدل کمینه‌سازی تاناکا [۳] $(T-min)$، مدل بیشینه‌سازی تاناکا [۳] $(T-max)$، مدل ساکاوا و یانو [۱۷] (SY)، مدل کیم و بیشو [۱۱] (KB)، مدل حجتی و همکاران [۷] (HBS)، مدل عربپور و تاتا [۱] (AT)، و مدل ارائه شده در این مقاله $(GP1)$ در جدول ۲ آمده است. مقایسه مقادیر دو سطر آخر جدول ۲ نشان می‌دهد که هم TE و هم λ در $GP1$ در این مثال از تمامی روش‌های مورد مقایسه (به استثنای HBS که TE در آن کمتر از TE در $GP1$ و λ در آن مساوی λ در $GP1$ است) بهتر می‌باشد. مدل‌های رگرسیون به‌دست آمده از روش‌های مذکور به‌قرار زیراند

$$\tilde{Y}_T = (۷/۷۰۰۰, ۰) \oplus (۰/۱۷۵۰, ۱/۹۲۵۰)x$$
$$\tilde{Y}_{T-min} = (۳/۸۵۰۰, ۳/۸۵۰۰) \oplus (۲/۱۰۰۰, ۰)x$$
$$\tilde{Y}_{T-max} = (۴/۶۳۳۳, ۰) \oplus (۱/۷۷۵۰, ۰/۲۰۸۳)x$$
$$\tilde{Y}_{SY} = (۴/۶۳۳۳, ۰) \oplus (۱/۵۶۶۷, ۰)x$$
$$\tilde{Y}_{KB} = (۴/۹۵۰۰, ۱/۸۴۰۰) \oplus (۱/۷۱۰۰, ۰/۱۶۰۰)x$$
$$\tilde{Y}_{HBS} = (۶/۷۵۰۰, ۱/۶۵۰۰) \oplus (۱/۲۵۰۰, ۰/۱۵۰۰)x$$
$$\tilde{Y}_{AT} = (۴/۹۵۰۰, ۱/۸۴۰۰) \oplus (۱/۷۱۰۰, ۰/۱۶۰۰)x$$
$$\tilde{Y}_{GP1} = (۶/۷۵۰۰, ۱/۸۰۰۰) \oplus (۱/۲۵۰۰, ۰/۲۰۰۰)x$$

مثال ۲ در این مثال، برای مقایسه دقیق‌تر روش ارائه شده با روش‌های مذکور در مثال ۱، یک مطالعه شبیه‌سازی به‌صورت زیر انجام شده است.

تعداد پنج مجموعه از داده‌ها هر یک شامل $n = ۲۰$ مقدار یک متغیر حقیقی و متغیر وابسته فازی مثلثی تولید شده است. ۲۰ مقدار متغیر مستقل x از توزیع یکنواخت روی بازه $(۰, u)$ به ازای ۵ مقدار ۱۰، ۲۰، ۳۰، ۴۰ و ۵۰ برای u استخراج شده‌اند. همچنین برای تولید ۲۰ مقدار متغیر وابسته $\tilde{y}$، هر بار سه عدد y_1 و y_2 و y_3 از توزیع یکنواخت روی بازه‌ی $(۰, u)$ به ازای مقادیر فوق برای u استخراج شده و برای تشکیل یک عدد فازی مثلثی، به‌ترتیب صعودی مرتب شده‌اند. خطای کل TE و λ به ازای انتخاب‌های مختلف u و نیز میانگین آنها، $\overline{TE}$ و $\overline{\lambda}$ به‌ترتیب در جداول ۳ و ۴ آمده‌اند. مقایسه $\overline{TE}$ و $\overline{\lambda}$ در جداول مذکور برتری $GP1$ بر سایر روش‌ها را هم از نظر TE و هم از نظر λ نشان می‌دهد.

مثال ۳ در مثال قبل پنج نمونه و هر یک شامل $n = ۲۰$ داده روی بازه‌های $(۰, u)$ به ازای $u = ۱۰, ۲۰, ..., ۵۰$ درنظر گرفته شد. در این مثال نیز یک مطالعه شبیه‌سازی، مانند مثال قبل انجام شده است، با این تفاوت که تعداد ۱۰۰ نمونه، و هر یک شامل $n = ۲۰$ داده تولید شده است. مقادیر متغیرهای مستقل و وابسته در نمونه‌ی k—ام از توزیع یکنواخت روی بازه‌ی $(۰, k)$ به ازای $k = ۱, ۲, ..., ۱۰۰$ استخراج شده‌اند. برای رعایت اختصار، فقط مقادیر $\overline{TE}$ و $\overline{\lambda}$ در جدول ۵ ارائه شده است. مجدداً مقایسه این مقادیر برتری $GP۱$ بر سایر روش‌ها را هم از نظر TE و هم از نظر λ نشان می‌دهد.

مثال ۴ اکنون داده‌های مثال ۱ را درنظر بگیرید که در آن $\tilde{y}_۱$ را با داده پرت $(۲۰, ۱/۸)$ عوض کرده‌ایم. مجدداً مقادیر E_i، TE، و λ برای مدل‌های مذکور در مثال ۱ محاسبه شده و در جدول ٦ ارائه شده است. همچنین اختلاف مقادیر دو سطر آخر جداول ۲ و ٦ در جدول ۷ آمده است.

این جدول نشان می‌دهد که از نظر TE، کمترین حساسیت نسبت به داده پرت را مدل $GP۱$ دارد، و از نظر λ، حساسیت مدل‌های HBS و $GP۱$ نسبت به داده پرت از سایر مدل‌های مورد مقایسه کمتر است. ضمنا مدل $T - max$ پس از اضافه شدن داده پرت جواب ندارد، که بیانگر بیشترین حساسیت نسبت به داده پرت است.

جدول ۲. خطای روش‌های مختلف در مثال ۱

E_i								i
$GP۱$	AT	HBS	KB	SY	$T-max$	$T-min$	T	
۰/۲۰۰	۲/۲۰۷	۰/۰۰۰	۲/۲۰۷	۱/۸۰۰	۱/۹۲۲	۳/۳۵٦	۰/۲۴٦	۱
۳/۸۵۴	۳/۰۵۰	۳/۷۴۳	۳/۰۵۰	۲/۲۰۰	۲/۳۵۱	۲/۸۵۰	۲/۸۵۰	۲
۱/۸۰۰	۱/۰۹۲	۱/۷۸۷	۱/۰۹۲	۲/٦۰۰	۲/۰۱٦	۱/۵۲۲	۳/۴۹۳	۳
۲/۹۱۱	۲/۸۴۴	۲/۸٦۹	۲/۸۴۴	۲/٦۰۰	۲/٦۲۴	۲/۲۵۸	۷/٦۷۵	۴
۰/۴۰۰	۰/۹۵۰	۰/۰۰۰	۰/۹۵۰	۲/۴۰۰	۱/۴۷۴	۲/۴۱۵	۸/۳۰۷	۵
۹/۱٦۵	۱۰/۱۴۴	۸/۳۹۹	۱۰/۱۴۴	۱۱/٦۰۰	۱۰/۳۸۷	۱۲/۴۰۲	۲۲/۵۷۰	TE
۵/٦۰۰	٦/۱۰۰	۵/٦۰۰	٦/۱۰۰	٦/۴٦٦	٦/۱۰۸	٦/۹۵۰	۱۲/۵۷۵	λ

جدول ۳. مقادیر TE در مثال ۲

خطای کل (TE)								u
$GP۱$	AT	HBS	KB	SY	$T-max$	$T-min$	T	
۳۵/٦۹۴	۳٦/۲۸۴	۴۰/۵۵۰	۳٦/۲۸۴	۵۴/۲۹٦	۴۴/۱۴۷	٦۵/۳۰۳	۱۱۲/۲٦۱	۱۰
۸۳/۷۵۷	۹۰/۵۱۵	۸٦/۵۸۴	۹۰/۵۱۵	۱۰۳/۵۳۳	۱۰۷/۴۵۵	۱۳٦/۸۹۱	۱۳۷۹/۷۲۹	۲۰
۲۰۷/۱۵۱	۲۱۳/۰۰٦	۱۵۹/۵۲۰	۲۱۳/۰۰٦	۱۴٦/٦۸۸	۱۵۲/۹۳۳	۲۲٦/۲۲۰	۳۵۳/۳۴۱	۳۰
۱۹۳/۱۴۹	۲۱۸/۹۰٦	۲۳۰/٦۷۵	۲۱۸/۹۰٦	۲۴۹/۷۰۰	۱٦۰/۵۳۵	۳۲۸/٦۱۹	۵۳۷/۹۹۳	۴۰
۲۱۴/٦۳۷	۲۱٦/۳۰۳	۲۳۱/۱۸۸	۲۱٦/۳۰۳	۲۷۹/۰۱۹	۲۷٦/۱۰۷	۳۵۷/۵۴۸	۵۴۲۴/۰۸۵	۵۰
۱۴٦/۸۷۸	۱۵۵/۰۰۳	۱۴۹/۷۰۳	۱۵۵/۰۰۳	۱٦٦/٦۴۷	۱۴۸/۲۳۵	۲۲۲/۹۱٦	۱۵٦۱/۴۸۲	$\overline{TE}$

جدول ۴. مقادیر λ در مثال ۲

λ								u
$GP۱$	AT	HBS	KB	SY	$T-max$	$T-min$	T	
۲۹/۰۲۵	۲۹/۲۲۸	۴۷/۴۵٦	۲۹/۲۲۸	۳٦/۵۱۵	۳٦/۱۰۷	۸۰/۷۰۳	۱۴٦/۵۲٦	۱۰
۷۳/۹۵۷	۷۹/۸۱۲	۸۵/۷۴۱	۷۹/۸۱۲	۸۳/۵۲۹	۷۷/۵۸۱	۱٦۷/۲۸۹	۲۴۴۱/٦۷۸	۲۰
۸۵/۵۹۷	۸۸/۱۳۸	۲۰۳/۵۳۵	۸۸/۱۳۸	۱۰۳/۴۳۴	۹۸/۸۹۳	۲۹۴/۰۳۹	۴٦۸/۲۵۵	۳۰
۲۰۰/۵۲۵	۲۰۳/۲۳۹	۲۴۵/۸۲۲	۲۰۳/۲۳۹	۲۲۱/۳۲۱	۲۵۳/۲٦۹	۴۱۱/۱۴۵	۷۳۰/۰۴۰	۴۰
۱۴۹/۴٦۲	۱۵۱/۷۱۳	۲۳۵/۱۱۱	۱۵۱/۷۱۳	۲۳۷/۸۸۴	۱٦۲/۱۴۸	۴٦۱/٦۸۸	۹۷۳۰/۵۵۵	۵۰
۱۰۷/۷۱۳	۱۱۰/۴۲٦	۱٦۳/۵۳۳	۱۱۰/۴۲٦	۱۳٦/۵۳۷	۱۲۵/٦۰۰	۲۸۲/۹۳۷	۲۷۰۳/۴۱۱	$\overline{\lambda}$

جدول ۵. مقادیر $\overline{TE}$ و $\overline{\lambda}$ در مثال ۳

روش	T	$T-min$	$T-max$	SY	KB	HBS	AT	$GP۱$
$\overline{TE}$	۲۵۸۴/۹۹۰	۳۷۲/۵۵۸	۲۵۰/۹۵۷	۲۶۴/۶۵۸	۲۶۰/۰۵۱	۲۵۱/۱۰۱	۲۶۰/۰۵۱	۲۳۴/۵۲۵
$\overline{\lambda}$	۲۰۸۹/۵۲۴	۴۷۷/۴۷۹	۲۲۹/۳۵۹	۲۲۷/۱۲۳	۱۷۸/۸۴۶	۲۷۶/۹۶۷	۱۷۸/۸۴۶	۱۷۳/۷۴۲

جدول ۶. خطای روش‌های مختلف در مثال ۴

i	E_i							
	T	$T-min$	$T-max$	SY	KB	HBS	AT	$GP۱$
۱	۳/۵۶۶	۸/۴۷۸	–	۵/۶۵۰	۳/۸۰۰	۴/۲۰۰	۳/۸۰۰	۳/۸۰۰
۲	۸/۳۰۲	۸/۳۰۲	–	۶/۰۵۰	۴/۳۶۰	۴/۶۰۰	۴/۳۶۰	۴/۴۰۰
۳	۹/۷۱۹	۵/۵۱۰	–	۱/۹۶۱	۴/۱۵۵	۴/۸۳۸	۴/۱۵۵	۴/۹۵۰
۴	۱۶/۶۶۰	۷/۲۴۰	–	۶/۰۵۳	۲/۸۴۴	۰/۳۹۲	۲/۸۴۴	۰/۰۰۰
۵	۲۱/۳۷۱	۸/۱۸۵	–	۶/۲۵۰	۳/۰۸۴	۰/۰۰۰	۳/۰۸۴	۰/۴۰۰
TE	۵۹/۶۱۷	۳۷/۷۱۴	–	۲۵/۹۶۴	۱۸/۲۴۳	۱۴/۰۳۰	۱۸/۲۴۳	۱۳/۵۵۰
λ	۴۱/۰۰۰	۲۳/۲۰۰	–	۲۳/۸۵۰	۱۹/۵۰۰	۱۷/۶۰۰	۱۹/۵۰۰	۱۷/۶۰۰

جدول ۷. میزان تغییر در TE و λ در روش‌های مختلف پس از اضافه شدن داده پرت در مثال ۴

روش	T	$T-min$	$T-max$	SY	KB	HBS	AT	$GP۱$
تغییر در TE	۳۷/۰۴۷	۲۵/۳۱۲	–	۱۴/۳۶۴	۸/۰۹۹	۵/۶۳۱	۸/۰۹۹	۴/۳۸۵
تغییر در λ	۲۸/۴۲۵	۱۶/۲۵۰	–	۱۷/۳۸۴	۱۳/۴۰۰	۱۲/۰۰۰	۱۳/۴۰۰	۱۲/۰۰۰

۵ نتیجه

در این مقاله، روش جدیدی برای محاسبه ضرایب مدل رگرسیون خطی فازی با ورودی‌های غیرفازی و خروجی فازی ارائه شد. در روش ارائه شده، ابتدا با استفاده از یک تابع فاصله وزن‌دار جدید روی اعداد فازی مثلثی، یک مدل بهینه‌سازی غیرخطی برای محاسبه ضرایب رگرسیون ارائه گردید که به سادگی به یک مدل برنامه‌ریزی آرمانی و در نهایت به یک مدل برنامه‌ریزی خطی تبدیل می‌شود. مزیت این تبدیل این است که روش سیمپلکس جواب دقیق مدل برنامه‌ریزی خطی را به‌دست می‌دهد، حال آن که الگوریتم‌های موجود برای حل مدل‌های غیرخطی اغلب جواب تقریبی به‌دست می‌دهند.

ویژگی مهم مدل برنامه‌ریزی خطی ارائه شده تجزیه‌پذیر بودن آن است، که باعث می‌شود تعداد قیود مدل به میزان قابل توجهی کاهش یابد. نتایج عددی نشان می‌دهد که روش ارائه شده از نظر نزدیکی توابع عضویت پاسخ‌های برآورد شده و داده شده و به‌خصوص از نظر نزدیکی مراکز آنها بر بسیاری از روش‌های قبلی مورد مقایسه برتری دارد (توجه داریم که مراکز پاسخ‌های فازی مثلثی به‌دلیل داشتن بیشترین درجه عضویت، اهمیت ویژه‌ای دارند). ضمناً روش ارائه شده برای داده‌های ذوزنقه‌ای و مثلثی متقارن و نامتقارن قابل استفاده است.

مراجع

[1] Arabpour A. R. and Tata, M., Estimating the parameters of a fuzzy linear regression model, *Iranian Journal of Fuzzy Systems*, To appear.

[2] Celmins, A. (1987), Least squares model fitting to fuzzy vector data, *Fuzzy Sets and Systems*, 22, 245-269.

[3] Chen, Y.S. (2001), Outliers detection and confidence interval modification in fuzzy regression, *Fuzzy Sets and Systems*, 119, 259-272.

[4] Diamond, P. (1988), Fuzzy Least squares, *Inform. Sci.*, 46, 141-157.

[5] Dubois, D. and Prade, H. (1983), Ranking fuzzy numbers in setting of possibility theory, *Inform. Sci.*, 30, 183-224.

[6] Hassanpur, H. and Maleki, H.R. (2007), A simple method for evaluation of fuzzy linear regression models by ordinary least-squares method, *Extended Abstracts of the 38th Annual IMC, 3-6 Sept 2007, University of Zanjan, Zanjan, Iran.*

[7] Hojati, M., Bector, C.R. and Smimou, K. (2005), A simple method for computation of fuzzy linear regression, *EJOR*, 166, 172-184.

[8] Hung, W.L. and Yang, M.S. (2006), An omission approach for detecting outliers in fuzzy regression models, *Fuzzy Sets and Systems*, 157, 3109-3122.

[9] Ishibuchi, H. and Nii, M. (2001), Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks, *Fuzzy Sets and Systems*, 119, 273-290.

[10] Kao, C. and Chyu, C.L. (2003), Least-squares estimates in fuzzy regression analysis, *EJOR*, 148, 426-435.

[11] Kim, B. and Bishu, R.R. (1998), Evaluation of fuzzy linear regression models by comparing membership functions, *Fuzzy Sets and Systems*, 100, 343-352.

[12] Modarres, M., Nasrabadi, E., and Nasrabadi, M.M. (2005), Fuzzy linear regression models with least square errors, *Applied Mathematics and Computation*, 163, 977-989.

[13] Özelkan, E.C. and Duckstein, L. (2000), Multi-objective fuzzy regression: a general framework, *Computers & Operations Research*, 27, 635-652.

[14] Peters, G. (1994), Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems*, 63, 45-55.

[15] Redden, D. T. and Woodall, W.H. (1994), Properties of certain fuzzy linear regression models, *Fuzzy Sets and Systems*, 64, 361-375.

[16] Sakawa, M. (1993). *Fuzzy Sets and Interactive Multiobjective Optimization*, Plenum Press, New York.

[17] Sakawa, M. and Yano, H. (1992), Multiobjective fuzzy linear regression analysis for fuzzy input-output data, *Fuzzy Sets and Systems*, 47, 173-181.

[18] Savic D.A. and Pedrycz, W. (1991), Evaluation of fuzzy linear regression models, *Fuzzy Sets and Systems*, 39, 51-63.

[19] Tanaka, H. (1987), Fuzzy data analysis by possibilistic linear models, *Fuzzy Sets and Systems*, 24, 363-375.

[20] Tanaka, H., Uejima, S., and Asai, K. (1982), Linear regression analysis with fuzzy model, *IEEE Transactions on Systems, Man, and Cybernetics*, 12, 903-907.

[21] Zadeh, L.A. (1978), Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems*, 1, 3-28.