مجله تحقیقات مهندسی کشاورزی/جلد ۱۱/ شماره ۳/سال ۱۳۸۹/ص ۹۶-۸۷

ارزیابی مدل‌های رگرسیون در پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور

عباس روحانی، ایرج رنجبر، محمدحسین عباسپور فرد[*]،

یحیی عجب شیرجی و مصطفی ولی‌زاده[**]

[*] نگارنده مسئول، نشانی: مشهد، دانشگاه فردوسی مشهد، ص. پ: ۱۱۶۳، تلفن: ۸۷۹۵۶۲۰(۰۵۱۱)، پیام‌نگار: abaspour@ferdowsium.ac.ir

[**] به ترتیب دانشجوی دکتری؛ دانشیار گروه ماشین‌های کشاورزی دانشگاه تبریز؛ استادیار گروه ماشین‌های کشاورزی دانشگاه فردوسی مشهد؛ دانشیار گروه ماشین‌های کشاورزی دانشگاه تبریز؛ و استاد گروه زراعت دانشگاه تبریز

تاریخ دریافت: ۸۸/۱۱/۱۱؛ تاریخ پذیرش: ۸۹/۷/۱۰

چکیده

مدیریت جایگزینی ماشین (به ویژه تراکتور) یکی از فاکتورهای کلیدی در انجام دادن به موقع عملیات زراعی می‌باشد. بنابراین باید با دقت هزینه‌های تعمیر و نگهداری تراکتور پیش‌بینی شود. این تحقیق برای ارزیابی تکنیک رگرسیونی در پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور اجرا شد. در این مطالعه از داده‌های واقعی ۶۰ تراکتور دو چرخ محرک موجود در کشت و صنعت آستان قدس رضوی استفاده شده است. تجزیهٔ رگرسیونی انجام شد. متغیر پاسخ عبارت بود از مقدار هزینهٔ تعمیر و نگهداری و متغیر مستقل ساعات کارکرد تجمعی. مدل‌های چند جمله‌ای درجه دوم و درجه سوم و مدل‌های توانی و نمایی آزمایش شدند. نتایج حاصل نشان می‌دهد که مدل رگرسیونی درجه سوم بر اساس ضریب تبیین و معنی‌داری ضرایب رگرسیونی، بهترین مدل جهت پیش‌بینی هزینهٔ تعمیر، روغن، و سوخت است. همچنین، عملکرد مدل رگرسیونی درجه سوم در مرحلهٔ آزمون برای پیش‌بینی هزینهٔ تعمیرات با متوسط قدرمطلق خطا ۳/۷۲ بسیار بهتر از هزینه روغن و سوخت به ترتیب با متوسط قدرمطلق خطا ۵/۲۱ و ۹/۵۱ است.

واژه‌های کلیدی

تعمیر و نگهداری، رگرسیون، قابلیت پیش‌بینی

مقدمه

شاید یکی از مواردی که امروزه در کشور ما تصمیم‌گیری‌های کلان و خرد مدیریت کشاورزی و به ویژه مدیریت ماشین‌های کشاورزی را دچار مشکل کرده است، نبود درک متغیرهای تصمیم‌گیری صحیح مربوط به هر منطقه است. مهم‌ترین عامل اساسی در انجام دادن به موقع عملیات زراعی از قبل برنامه‌ریزی شده، در دسترس بودن ماشین است (Almassi *et al.*, 2008). اگر نتوان زمان بهینه جایگزینی ماشین را پیش‌بینی کرد، قابلیت دسترسی ماشین کاهش می‌یابد. پیش‌بینی یعنی تخمین پدیده در آینده بر اساس داده‌های تاریخی آن پدیده است (Oskounejad, 2003). تعیین ظرفیت و تعداد بهینه ماشین‌های مورد نیاز، برنامه‌ریزی کاری آنها در زمان‌های مشخص، فراهم کردن قطعات یدکی ضروری، و تعیین زمان بهینه جایگزینی به منظور حفظ قابلیت دسترسی ماشین، از عوامل ضرورت نیاز به پیش‌بینی است (Kim, 1989). پیش‌بینی دقیق هزینه‌های تعمیر و نگهداری، کلید موفقیت در تصمیم‌گیری جایگزینی است. به دلیل طبیعت تصادفی و غیر قابل انتظار بودن تعمیرات، مقدار و چگونگی هزینه‌های تعمیر و نگهداری به شدت وابسته به کیفیت مدیریت، مهارت کاربران، و شرایط آب و هوایی است (Mitchell, 1998). این عوامل، لزوم پیش‌بینی هزینه‌های تعمیر و نگهداری ماشین‌های کشاورزی و به ویژه تراکتور را به عنوان نیروی محرک سایر

ماشین‌آلات کشاورزی روشن می‌سازد. تاکنون از روش‌های مختلف برای پیش‌بینی هزینه‌های تعمیر و نگهداری استفاده شده است. تکنیک‌های رگرسیونی به دلیل داشتن پشتوانهٔ ریاضی بسیار خوب، از سال ۱۹۷۰ در پیش‌بینی هزینه‌های تعمیر و نگهداری ماشین‌های کشاورزی استفاده شدند (Bowers & Hunt, 1970). تحلیل رگرسیونی روشی است که رابطهٔ بین یک متغیر وابسته نسبت به یک یا چند متغیر مستقل پیدا می‌کند (Lyman & Longnecker, 2001). در سالنامه انجمن مهندسان کشاورزی امریکا، برای اولین بار در سال ۱۹۹۱ مدل‌هایی جهت پیش‌بینی هزینه‌های تعمیر و نگهداری ماشین‌های کشاورزی ارائه شد ولی این مدل‌ها مشکلاتی مانند ارائهٔ دو مدل متفاوت برای هر ماشین داشتند که موجب سردرگمی کاربر در استفاده این مدل‌ها می‌شد که دلیلی بود بر غیر استاندارد بودن آنها. علاوه بر این، در بعضی از مدل‌ها از سطح پوشش عملیاتی تجمعی به عنوان سن ماشین استفاده شده بود که در این حالت اندازه ماشین تا حد زیادی موجب اریب یا خطا در براورد هزینه‌هایی می‌شد که مدل براورد می‌کرد. و سرانجام اینکه به دلیل کوچک بودن نمونهٔ آماری، مدل‌های به دست آمده قابل تعمیم برای سایر ماشین‌ها نبودند (Rotz, 1987; Morris, 1988; Nunnally, 1993).

در روش خط مستقیم برای پیش‌بینی هزینه‌های تعمیر و نگهداری، فرض می‌شود که هزینهٔ تعمیر و نگهداری در طول عمر ماشین ثابت است و این روش زمانی مفید خواهد بود که برای محاسبهٔ استهلاک از روش خط مستقیم استفاده شود و اغلب به دلیل روند تغییرات غیر خطی این هزینه‌ها، نمی‌توان از این روش استفاده کرد. مطالعهٔ داده‌های تعمیر و نگهداری ماشین‌های کشاورزی نشان می‌دهد که رابطهٔ بین هزینه‌های تعمیر و نگهداری ماشین و ساعات کارکرد تجمعی رابطه غیرخطی درجه دوم، درجه سوم، یا نمایی است (Nunnally, 1993; Fuls, 1999). در تعیین مدل هزینه‌های تعمیر و نگهداری باید کل ساعات کارکرد، عمر مفید، دما، شرایط کاری، کیفیت نگهداری، نوع استفاده، شانس، کیفیت ماشین، موقعیت کاربری ماشین، و سبک و کاربر را دخالت دارد. روش تخمین هزینه‌های تعمیر و نگهداری مشابه روش مجموع ارقام سال‌های عمر برای محاسبه استهلاک است (Mitchell, 1998). مطالعهٔ کامیون‌های ارتش امریکا نشان داده است که مدل سری‌های زمانی می‌تواند متوسط هزینه‌های تعمیر و نگهداری این کامیون‌ها را برازش کند ولی این روش تنها برای ماشین‌هایی با عمر طولانی عملی می‌باشد (Kim, 1989). برای براورد هزینه‌های تعمیر و نگهداری تراکتور، مدل توانی پیشنهاد شده است؛ این مدل می‌تواند روند تغییر هزینه‌ها را در طول عمر ماشین به طور واقعی و مطلوب نشان دهد. مشکل این مدل، استفاده از سطح پوشش عملیاتی به عنوان سن ماشین است که به تأثیر غیر واقعی اندازه، ماشین در براورد هزینه‌های ماشین می‌انجامد (Nunnally, 1993). در تحقیقی دیگر، مدلی بر پایهٔ مدل توانی پیشنهاد شده است با این تفاوت که پارامتری جدید برای لحاظ کردن تأثیر سیاست‌های مدیریتی و مهارت کاربر در تخمین هزینه‌های تعمیر و نگهداری تراکتور به مدل اضافه می‌شود (Fuls, 1999). در سال‌های اخیر، تحقیقات تقریباً زیادی در داخل کشور در این زمینه شده است که از آن جمله می‌توان به تعیین مدل‌های ریاضی برای پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتورهای موجود در کشت و صنعت‌های دشت ناز ساری و نیشکر کارون اشاره کرد (Almassi & Yaganeh, 2002; Ashtiani, 2005).

این مطالعه با هدف بررسی و شایستگی مدل‌های رگرسیونی به منظور پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور اجرا شد.

مواد و روش‌ها

داده‌های هزینه‌های تعمیر و نگهداری مزرعه‌ای

امکان دستیابی به داده‌های آزمایشگاهی مناسب برای این مطالعه باشد وجود ندارد ولی داده‌های مزرعه‌ای تصویری واقعی از چگونگی تغییرات هزینه‌های تعمیر و نگهداری در طول زمان را نشان می‌دهند. برای اجرای این تحقیق از داده‌های هزینه‌های تعمیر و نگهداری ماهانه در ۱۸ سال مربوط به چهار نوع تراکتور دو چرخ محرک فعال در مزرعهٔ نمونهٔ آستان قدس رضوی در استان خراسان رضوی استفاده شد که عبارت‌اند از جاندیر ۳۱۴۰، جاندیر ۴۴۵۰، فرگوسن ۲۸۵ و فیات ۴۴۵ به ترتیب ۲۸، ۵، ۱۷ و ۱۰ و مجموعاً ۶۰ تراکتور. این مزرعه حدود ۳۰۰۰ هکتار زمین زراعی دارد. این داده‌ها شامل هزینهٔ تعمیرات (قطعات یدکی و دستمزد تعمیرات)، هزینهٔ روغن (روانسازها و فیلترهای روغن)، هزینهٔ سوخت (گازوییل و فیلترهای سوخت)، هزینهٔ تعمیر و نگهداری، سال خرید، و ساخت هر تراکتور است.

استانداردسازی داده‌ها

در ابتدا و قبل از محاسبهٔ هزینهٔ تجمعی، باید اثر تورم بر هزینه‌ها را تعدیل کرد (Mitchell, 1998; Nunnally, 1993). به دلیل متفاوت بودن هزینه‌های تعمیر و نگهداری و نیز قیمت خرید اولیه انواع تراکتورهای مورد مطالعه و نیز متفاوت بودن طبیعت داده‌های مربوط به هر نوع تراکتور به دلیل به کارگیری آنها در عملیات زراعی مختلف، تمام این هزینه‌ها با شاخص زیر استانداردسازی شد (Mitchell,1998):

$$CCI_t = \frac{\sum_{0}^{t} C_t}{PP_0} \times 100 \qquad (1)$$

در آن، CCI_t= شاخص هزینه تجمعی در زمان؛ t= برای هر یک از هزینه های سوخت، روغن و تعمیرات؛ C_t= هزینهٔ تعمیر و نگهداری در زمان t برای هر یک از هزینه‌های سوخت، روغن، و تعمیرات؛ و PP_0= قیمت خرید اولیه تراکتور است. این شاخص در طول عمر تقویمی تراکتور همواره روند افزایشی دارد یا ثابت است. بعد از استانداردکردن هزینه‌ها، شاخص‌های تجمعی به صورت شاخص هزینهٔ تعمیراتی تجمعی (CCI_{repair})، شاخص هزینهٔ روغن تجمعی (CCI_{oil})، و شاخص هزینهٔ سوخت تجمعی (CCI_{fuel}) محاسبه شد. از این شاخص‌ها به عنوان متغیرهای وابسته در هر مدل رگرسیونی استفاده شد. در تصمیم‌گیری اقتصادی مدیریت ماشین برای تعیین عمر مفید، اغلب از شاخص هزینه تعمیر و نگهداری استفاده می‌شود. که برابر است با مجموع سه شاخص هزینهٔ تجمعی سوخت، روغن، و تعمیرات.

طول عمر

عموماً، متغیر مستقل در مدل رگرسیونی هزینهٔ تجمعی، عمر تراکتور است. برای طول عمر سه تعریف وجود دارد عمر تقویمی، عمر برحسب واحدهای تولید محصول، و عمر بر حسب ساعات کارکرد تجمعی. از بین این سه تعریف، ساعات کارکرد تجمعی (CHU[1])، مناسب‌ترین تعریفی است که می‌توان از عمر تراکتور داشت (Mitchell, 1998). CHU تعداد ساعاتی را نشان می‌دهد که تراکتور به طور فیزیکی کار کرده است و نیز تغییرات بسیار زیاد هزینه‌های تعمیر و نگهداری را به خوبی در طول زمان تعدیل می‌کند. تعداد ساعات کارکرد هر تراکتور بر اساس تعداد تعویض ماهیانه روغن موتور محاسبه شد. برای دستیابی به عملکرد بهتر مدل‌بندی این هزینه‌ها توسط روش رگرسیونی، CHU بر حسب صد ساعت محاسبه شد.

1- Cumulative Hours of Usage (CHU)

فرضیات مدل‌سازی

در ابتدای تحلیل هزینه‌های تعمیر و نگهداری، فرض زیر را باید در نظر گرفت:

۱- اطمینان به داده‌های موجود، برای تصدیق مقدار هزینه‌ها، امکان برگشت به گذشته وجود ندارد و از این‌رو باید به مقدار هزینه‌های ثبت شده از طرف شرکت اطمینان کرد.

۲- هزینهٔ تعمیر و نگهداری تراکتور در ابتدای عمر آن صفر است. این فرض کاملاً قابل قبول و پذیرش آن ضروری است زیرا هزینهٔ تعمیرات احتمالی قبل از به کارگیری تراکتور را شرکت سازنده پرداخت می‌کند؛ براساس این فرضیه، ضریب عرض از مبدا مدل‌های رگرسیونی (β_0) صفر خواهد شد و همچنین نباید هزینه‌هایی که مدل تخمین می‌زند کمتر از صفر باشد.

۳- ساعات کارکرد تجمعی تنها متغیر مستقل برای مدل‌های رگرسیونی است.

اگرچه ممکن است متغیرهای زیادی در تخمین هزینه‌های تعمیر و نگهداری تراکتور تاثیرگذار باشند ولی به دلیل وجود شرایط یکسان به کارگیری آنها از جمله سطح مدیریتی، شرایط آب و هوایی، و تا حدودی یکسان بودن سطح مهارت کاربران، می‌توان سایر متغیرها را ثابت در نظر گرفت.

تقسیم‌بندی داده‌ها

در ابتدا داده‌های موجود به طور تصادفی به دو دسته، یکی مجموعهٔ داده انتخاب مدل با ۱۳۰ عضو (دو سوم کل داده‌ها) برای انتخاب مناسب‌ترین مدل رگرسیونی و دیگری مجموعهٔ آزمون با ۸۶ عضو (یک سوم کل داده‌ها) برای ارزیابی عملکرد مدل‌ها تقسیم‌بندی شدند. چنانچه این تقسیم‌بندی منجر به نتایج مطلوب نشود، می‌توان این مرحله را مجدداً تکرار کرد.

معیارهای ارزیابی عملکرد مدل

برای ارزیابی قابلیت روش رگرسیونی در پیش‌بینی هزینه‌های تعمیر و نگهداری تراکتور، از معیارهای متوسط قدرمطلق درصد خطا (MAPE[1]) ، ریشه متوسط مربعات خطا (RMSE[2])، و ضریب تبیین معادلهٔ خطی بین مقادیر پیش‌بینی شده توسط مدل رگرسیونی و مقادیر واقعی آنها (R^2) استفاده شد:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(d_i - p_i)^2}{m}} \qquad (۷)$$

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{d_i - p_i}{d_i}\right| \times 100 \qquad (۸)$$

$$R^2 = \frac{\left(\sum_{i=1}^{m}(d_i - \bar{d})(p_i - \bar{p})\right)^2}{\sum_{i=1}^{m}(d_i - \bar{d})^2 \quad . \quad \sum_{i=1}^{m}(p_i - \bar{p})^2} \qquad (۹)$$

که در آن روابط، d_i= مؤلفه iام از خروجی مدل؛ p_i= مؤلفه iام از خروجی پیش‌بینی شده توسط مدل؛ $\bar{d}$ و $\bar{p}$ = به ترتیب متوسط خروجی‌های واقعی و پیش‌بینی شده توسط مدل؛ و m= تعداد داده‌های استفاده شده است.

تحلیل رگرسیونی

برای انتخاب بهترین مدل رگرسیونی از بین مدل‌های توصیه شده درجه دوم، درجه سوم، نمایی، و توانی از مجموعه داده‌های آموزش و برای ارزیابی اعتبار مدل‌های انتخاب شده برای هر شاخص هزینه تعمیر و نگهداری از مجموعه داده‌های آزمون استفاده شد. مدل‌های رگرسیونی بر اساس معنی‌داری ضرایب رگرسیونی و پارامتر F و بزرگی ضریب تبیین در محیط نرم افزاری SPSS انتخاب شدند. مدل‌های رگرسیونی انتخاب شده بر

1- Mean Absolute Percentage Error 2- Root Mean Squared Error

اساس معیارهای MAPE، RMSE و R^2 در محیط نرم افزاری MATLAB ارزیابی شدند.

بحث و نتایج

در این مرحله به منظور توجیه کردن روابط $CCI_{repair} = f(CHU)$، $CCI_{oil} = f(CHU)$ و $CCI_{fuel} = f(CHU)$ بهترین مدل‌های رگرسیونی انتخاب خواهد شد.

انتخاب مدل: در جدول ۱، نتایج حاصل از تحلیل رگرسیونی، آزمون معنی‌داری ضرایب رگرسیونی مدل‌های درجه دوم، درجه سوم، نمایی، و توانی، پارامتر F متناظر هر مدل، و ضرایب تبیین و متوسط قدرمطلق درصد خطای هر مدل برای هر شاخص هزینهٔ تعمیر و نگهداری تراکتور دو چرخ محرک نشان داده شده است. نتایج حاصل از آزمون معنی‌داری ضرایب رگرسیونی نشان می‌دهد که تمام ضرایب رگرسیونی کلیهٔ مدل‌ها برای چهار شاخص هزینه تعمیر و نگهداری در سطح احتمال یک درصد معنی‌دار هستند. نتایج مؤید این موضوع هستند که چهار مدل درجه دوم، درجه سوم، نمایی، و توانی قابلیت توجیه تغییرات رفتار شاخص‌های هزینه تعمیر و نگهداری را در طول عمر تراکتور دارند. معنی‌دار شدن پارامتر F در سطح احتمال یک درصد در تمام مدل‌ها نشان می‌دهد که بین ساعات کارکرد تجمعی و شاخص‌های هزینه تعمیر، روغن، و سوخت رابطه‌ای معنی‌دار وجود دارد. شایستگی هر مدل بعد از معنی‌دار شدن ضرایب و پارامتر F، اغلب بر اساس بزرگی ضریب تبیین R^2 ارزیابی می‌شود، اما، همان‌طور که می‌بینیم، ضرایب تبیین همهٔ مدل‌ها نزدیک به هم و بالا هستند به ویژه ضریب تبیین دو مدل درجه دوم و درجه سوم بسیار نزدیک به هم هستند، بنابراین، برای انتخاب دقیق‌تر بهترین مدل برای هر شاخص هزینه، از پارامتر MAPE استفاده شد. با توجه به این معیار، مدلی بیشترین شایستگی را دارد که دارای کمترین مقدار MAPE باشد. با در نظر گرفتن این موضوع، مدل درجه سوم می‌تواند مناسب‌ترین مدل جهت پیش‌بینی هر سه شاخص هزینهٔ تعمیر، روغن، و سوخت باشد زیرا مدل‌های درجه دوم و درجه سوم هزینه دارای ضرایب تبیین تقریباً مساوی و بیشتر از مدل‌های توانی و نمایی در هر سه شاخص هزینه تعمیر و نگهداری هستند ولی مدل درجه سوم در مقایسه با مدل درجه دوم منجر به خطای کمتری در حدود ۹۰، ۳۰، و ۴۵ درصد به ترتیب برای شاخص‌های تعمیر، سوخت، و روغن شده است.

جدول ۱- نتایج تحلیل رگرسیونی داده‌های مجموعه انتخاب مدل

شاخص هزینه	نوع مدل	ضرایب رگرسیونی			F	R^2	MAPE (درصد)
		β_1	β_2	β_3			
CCI_{repair}	درجه دوم	۰/۳۹۸**	۰/۰۰۱**	-	۲۶۸۸۷**	۰/۹۹۸	۲۷/۳۷
	درجه سوم	۰/۰۴۶ **	۰/۰۰۵**	$-۱/۱\times۱۰^{-۵}$**	۸۴۹۱۳۰**	۰/۹۹۹	۲/۹۴
	نمایی	۳/۹۹۳**	۰/۰۱۸**	-	۶۴۶**	۰/۹۰۶	۶۲
	توانی	۰/۰۳۳**	۱/۵۴۱**	-	۳۲۳۳۲**	۰/۹۹۸	۶/۶۴
CCI_{oil}	درجه دوم	۰/۰۵۴**	$۲/۸\times۱۰^{-۴}$**	-	۱۰۵۰۴**	۰/۹۹۹	۴/۰۲
	درجه سوم	۰/۰۴۷**	$۲/۸\times۱۰^{-۴}$**	$-۳/۰\times۱۰^{-۷}$**	۱۲۶۹۴۸**	۰/۹۹۹	۲/۸۹
	نمایی	۱/۱۶۹**	۰/۰۱۴**	-	۷۹۳**	۰/۸۴۸	۳۹/۷۱
	توانی	۰/۰۲۷**	۱/۲۱۵**	-	۴۶۳۹۱**	۰/۹۹۷	۴/۷۲
CCI_{fuel}	درجه دوم	۰/۰۲۹*	۰/۰۰۰۳۰**	-	۹۶۴۵۳**	۰/۹۹۹	۸/۹۵
	درجه سوم	۰/۰۴۱**	۰/۰۰۰۱۱**	$۵/۱۵\times۱۰^{-۷}$**	۱۰۴۹۲۷**	۰/۹۹۹	۴/۸۹
	نمایی	۱/۰۰۹**	۰/۰۱۵**	-	۱۳۰۷**	۰/۹۰۲	۲۸
	توانی	۰/۰۲۸**	۱/۱۸۲**		۸۳۸۳**	۰/۹۸۳	۱۱/۹۸

** اختلاف معنی‌دار در سطح احتمال ۱ درصد، و * اختلاف معنی‌دار در سطح احتمال ۵ درصد

مدل درجه دوم: $y=\beta_1 x+\beta_2 x^2$؛ مدل درجه سوم: $y=\beta_1 x+\beta_2 x^2+\beta_3 x^3$؛ مدل نمایی: $y=\beta_1 e^{\beta_2 x}$؛ مدل توانی: $y=\beta_1 x^{\beta_2}$

ارزیابی قابلیت پیش‌بینی مدل: در اینجا، نتایج حاصل از ارزیابی مدل‌های رگرسیونی انتخاب شده ارائه می‌شود زیرا معتبر بودن این نتایج دال بر موفق و معتبر بودن مدل‌های رگرسیونی دارد. هدف از این مرحله، ارزیابی تعمیم‌پذیری و شایستگی مدل‌های رگرسیونی انتخاب شده است. بنابراین، مدل‌ها با استفاده از مجموعه داده‌های غیر از مجموعه داده‌های انتخاب مدل (مجموعه داده‌های آزمون) ارزیابی می‌شوند. در جدول ۲ برخی ویژگی‌های آماری داده‌های استفاده شده در مرحله آزمون همراه با مقادیر پیش‌بینی شده آنها توسط مدل رگرسیونی نشان داده شده است. بین مقادیر پیش‌بینی شده شاخص هزینه تعمیر و مقادیر واقعی آنها برای هر سه مدل رگرسیونی تفاوت زیادی دیده نمی‌شود. ولی این اختلاف در مورد شاخص هزینهٔ روغن و سوخت، در مقایسه با شاخص هزینه تعمیرات، بسیار چشمگیر است که دلیل آن ممکن است نوع مدل‌های رگرسیونی انتخاب شده برای این هزینه‌ها باشد.

جدول ۲- ویژگی‌های آماری مقادیر داده‌های واقعی و پیش‌بینی شده از مجموعه آزمون

شاخص هزینه		ویژگی‌های آماری							
		میانگین	واریانس	انحراف معیار	کمینه	بیشینه	کشیدگی	چولگی	مجموع
CCI_{repair}	dv	۶۲/۳۱	۲۲۹۴/۷	۴۷/۹	۰/۱۳۲	۱۴۳/۱۳	۱/۶	۰/۱۷۳	۴۴۸۶
	pv	۶۲/۳۷	۲۲۹۶/۴	۴۷/۹	۰/۰۶۷	۱۴۲/۳۶	۱/۵۹	۰/۱۷۵	۴۴۹۰
CCI_{oil}	dv	۱۰/۱	۵۲/۳۷	۷/۲۴	۰/۰۶۱	۲۳/۲۲	۱/۷۴	۰/۲۴	۷۲۵
	pv	۹/۷	۴۶/۳۸	۶/۸۱	۰/۰۶	۲۱/۷۸	۱/۷۳	۰/۲۰۱	۶۹۶
CCI_{fuel}	dv	۹/۰۷	۴۹/۷۹	۷/۰۵	۰/۰۷۲	۲۳/۰۴	۱/۹۶	۰/۴۸۹	۶۵۳
	pv	۸/۱	۳۶/۹۵	۶/۱	۰/۰۵۳	۲۰/۱۵	۱/۹۳	۰/۴۰	۵۸۴

dv: مقادیر واقعی داده‌ها؛ pv: مقادیر پیش‌بینی شده توسط مدل رگرسیونی درجه سوم

میانگین، واریانس، و توزیع آماری مقادیر واقعی و مقادیر پیش‌بینی شدهٔ هزینه، از مجموعه داده‌های آزمون از نظر آماری مقایسه می‌شوند تا قابلیت پیش‌بینی روش رگرسیونی بیشتر بررسی شود. در اینجا فرضیهٔ صفر بر تساوی میانگین، واریانس، و توزیع آماری هر دو سری داده دلالت دارد. هر فرضیه در سطح احتمال ۹۵ درصد به کمک پارامتر p آزمون شد. بنابراین، اگر p محاسبه شده برای هر مقایسه بیش از ۰/۰۵ باشد، فرضیهٔ صفر را نمی‌توان رد کرد. برای مقایسهٔ میانگین، واریانس، و توزیع آماری به ترتیب از آزمون t، F، و کولموگرف-اسمیرنو استفاده شد. مقادیر p محاسبه شده برای هر سه شاخص هزینه در جدول ۳ نشان داده شده است. نتایج نشان می‌دهد که میانگین، واریانس، و توزیع آماری مقادیر واقعی و پیش‌بینی شده، شاخص هزینهٔ تعمیر، روغن، و سوخت توسط مدل رگرسیونی درجه سوم اختلاف معنی‌داری با هم ندارند و لذا مدل رگرسیونی درجه سوم در پیش‌بینی شاخص‌های هزینه تعمیر و نگهداری قابلیت بالایی دارد. اما دوباره می‌توان مقادیر p مربوط به هر مقایسه را به ترتیب نزولی از شاخص هزینه تعمیر، روغن و سوخت مرتب کرد.

جدول ۳- مقایسهٔ آماری مقادیر واقعی و پیش‌بینی شده از مجموعه داده‌های آزمون

شاخص هزینه	نوع تحلیل آماری		
	مقایسه میانگین	مقایسه واریانس	مقایسه توزیع
CCI_{repair}	۰/۹۹۴	۰/۹۹۸	۱/۰۰۰
CCI_{oil}	۰/۷۳۱	۰/۶۱۰	۰/۹۹۳
CCI_{fuel}	۰/۳۸۹	۰/۲۱۱	۰/۷۴۱

ارزیابی بر اساس معیارهای عملکردی: در اینجا عملکرد مدل رگرسیونی درجه سوم انتخاب شده، از نظر معیارهای متوسط قدرمطلق درصد خطا (MAPE) و نیز ریشه متوسط مربعات خطا (RMSE) در دو مرحله انتخاب مدل و آزمون با هم مقایسه می‌شود (جدول ۴). با توجه به نتایج درج شده در این جدول، MAPE و RMSE در هر سه شاخص هزینه برای مرحلهٔ آزمون بیشتر از مرحله انتخاب است زیرا داده‌های به کار گرفته شده در مرحلهٔ آزمون برای مدل رگرسیونی کاملاً تازگی دارند. مقایسه مدل‌های رگرسیونی در هر دو مرحلهٔ انتخاب مدل و آزمون بر اساس این دو معیار نشان از برتری مطلق مدل رگرسیونی شاخص هزینهٔ تعمیر نسبت به شاخص هزینه روغن و سوخت دارد زیرا MAPE و RMSE برای شاخص هزینه تعمیر به ترتیب ۳۰ و ۳ درصد کمتر از شاخص هزینه روغن و ۷۰ و ۶۰ درصد کمتر از شاخص هزینه سوخت است.

جدول ٤- مقایسه عملکرد مدل رگرسیونی در پیش‌بینی هزینهٔ تعمیر و نگهداری تراکتور

شاخص هزینه	معیار عملکردی مدل			
	MAPE		RMSE	
	مرحله انتخاب مدل	مرحله آزمون	مرحله انتخاب مدل	مرحله آزمون
CCI_{repair}	۲/۹۴	۳/۷۲	۰/۵۷۷	۰/۶۲۱
CCI_{oil}	۲/۸۹	۵/۲۱	۰/۵۹۲	۰/۶۳۸
CCI_{fuel}	۴/۸۹	۹/۵۱	۱/۳۴	۱/۳۹

نتیجه‌گیری

نتیجهٔ تحلیل رگرسیونی داده‌های مربوط به هزینه‌های تعمیر و نگهداری تراکتور از مجموعه داده‌های انتخاب مدل نشان می‌دهد که هر چهار مدل درجه دوم، درجه سوم، نمایی، و توانی اگرچه قابلیت برازش هزینه‌های تعمیر، روغن، و سوخت را در طول ساعات کارکرد تجمعی دارند ولی تنها مدل درجهٔ سوم، در مقایسه با سایر مدل‌ها به دو دلیل داشتن خطای کمتر در حدود ۹۰، ۳۰ و ۴۵ درصد به ترتیب برای هزینهٔ تعمیر، روغن، و سوخت، بیشترین قابلیت را برای برازش داده‌های مورد نظر دارد. مطابق با برخی یافته‌های قبلی، تغییرات هزینهٔ تعمیر و نگهداری در طول دوره مالکیت کاملاً غیر خطی است (Kim, 1989; Nunnally, 1993; Mitchell, 1998). این موضوع را می‌توان به طبیعت تصادفی تعمیرات نسبت داد. همچنین، بر اساس یافته‌های این مقاله، بهتر آن است که برای انتخاب مستدل‌تر بهترین مدل جهت پیش‌بینی هزینه علاوه بر معیارهایی مانند معنی‌داری ضرایب مدل‌ها، ضریب تبیین، و پارامتر F هر مدل، از معیار MAPE نیز استفاده شود (جدول ۱). اگرچه در چند منبع موجود بر مدل توانی به عنوان بهترین مدل جهت توجیه تغییرات هزینه‌های تعمیر و نگهداری در طول عمر تراکتور تاکید شده است، اما نتیجهٔ تحلیل رگرسیونی موجود در جدول ۱ نشان می‌دهد که این مدل اگرچه از نظر معنی‌دار بودن ضرایب مدل و نیز ضریب تبیین کاملاً قابل قبول است ولی مقدار خطای آن در سه هزینهٔ تعمیر، روغن، و سوخت تقریباً ۳ برابر مدل درجه سوم است، و از این رو در شرایط مطالعهٔ موجود، مدل توانی نمی‌تواند بهترین مدل باشد (Kim, 1989). مقایسهٔ عملکرد مدل درجه سوم برای سه نوع هزینهٔ تعمیر، روغن، و سوخت نشان می‌دهد که این مدل در پیش‌بینی هزینهٔ تعمیر، نسبت به هزینهٔ روغن و سوخت، دارای عملکرد بیشتری است. بنابراین اگرچه روش رگرسیون می‌تواند روشی نسبتاً مناسب برای پیش‌بینی هزینه‌های تعمیر و نگهداری باشد ولی مدل‌های رگرسیونی انتخاب شده تنها برای جمعیت تراکتوری مورد مطالعه اعتبار دارند.

مراجع

Almassi, M. and Yaganeh, H. R. 2002. Determining a suitable mathematical model to predict the repair and maintenance costs of farm tractors in Karoon agro-industry Co. Iranian J.Agric.Sci. 33(4): 707-716. (in Farsi)

Almassi, M. Kiani, S. H. and Loveimi, N. 2008. Principles of Agricultural Mechanization. Jangal Press. (in Farsi)

Ashtiani, A. 2005. Determining the optimum mathematical model for prediction of tractors repair and maintenance costs in Sari Dasht-e-Naz Farm Company. Agric. Sci. Tabriz Uni. 15(4):102-112. (in Farsi)

Bowers, W. and Hunt, D. R. 1970. Application of mathematical formula to repair cost data. Trans. ASAE. 13, 806-809.

Fuls, J. 1999. The Correlation of repair and maintenance costs of agricultural machinery with operating hours management policy and operator skills for South Africa. http://www.arc.agric.za

Kim, Y. H. 1989. A forecasting methodology for maintenance cost of long-life equipment. PhD Thesis. University of Alabama.

Lyman, O. and Longnecker, M. 2001. An introduction to statistical methods and data analysis. R R. Donnelley & Sons, Inc./Willard. United States of America. 532-825.

Mitchell, Z. W. 1998. A Statistical Analysis of Construction Equipment Repair Costs Using Field Data & the Cumulative Cost Model. PhD Thesis. Faculty of the Virginia Polytechnic Institute and State University.

Morris, J. 1988. Estimation of Tractor Repair and Maintenance Costs. J. Agric. Eng. Res. 41, 191-200.

Nunnally, S. W. 1993. Construction means and methods. Prentice-Hall. Englewood Cliffs. NJ.

Oskounejad, M. M. 2003. Engineering economic evolution of industrial project. Amirkabir Industrial Univ. Press. (in Farsi)

Rotz, C. A. 1987. A Standard Model for Repair Costs of Agricultural Machinery. Appl. Eng Agric. (1): 3-9.

Journal of Agricultural Engineering Research/Vol.11/No.3/2010/P: 87-96

Evaluation Regression Techniques in Prediction of Tractor Repair and Maintenance Costs

A. Rohani, I. Ranjbar, M. H. Abbaspour-Fard*, Y. Ajabshir and M. Valizadeh

* Corresponding Author: Assistant Professor, Mashhad University, P. O. Box: 1163. Mashhad, Iran. E-mail: abaspour@ferdowsi.um.ac.ir

The management of machine replacement (specifically tractors) is one of the most critical factors for performing field operations on time. Therefore, tractor repair and maintenance costs must be accurately predicted. This research was carried out to evaluate the use of the regression technique in predicting tractor repair and maintenance costs. The study was conducted using empirical data on 60 two-wheel drive tractors from Astan Ghods-e Razavi agro-industry. Regression analyses were carried out. The response variable was repair and maintenance cost value and the independent variable was cumulative hours of use. Different second and third-order polynomial models and exponential and power models of age were examined. The results showed that the cubic regression model was the best model for predicting repair, oil and fuel costs based on the coefficient of determination and significance of regression coefficients. The performance cubic regression model performed much better in the test phase for the prediction of repair costs with a mean absolute percentage error of 3.72 in comparison to oil and fuel costs with a mean absolute percentage errors of 5.21 and 9.51, respectively.

Key Words: Prediction Capability, Regression, Repair and Maintenance