پیش‌بینی و تحلیل نرخ خطای ناشی از رشد درختان در خطوط هوایی توزیع الکتریکی با استفاده از شبکه عصبی و تحلیل عاملی

مهدی صدقی[1]، دانشجوی دکتری، مسعود علی‌اکبر گلکار[1]، استاد، محمودرضا حقی‌فام[2]، استاد

۱- دانشکده مهندسی برق و کامپیوتر- دانشگاه صنعتی خواجه نصیرالدین طوسی- تهران- ایران
meh.sedghi@gmail.com, golkar@eetd.kntu.ac.ir
۲- دانشکده مهندسی برق- دانشگاه تربیت مدرس - تهران- ایران
haghifam@modares.ac.ir

چکیده: در مطالعه قابلیت اطمینان سیستم‌های قدرت، نرخ خطا پارامتری مهم است. نرخ خطا در مطالعات شبکه‌های توزیع معمولاً به طور تقریبی ثابت در نظر گرفته می‌شود ولی به طور دقیق‌تر و در عمل، پارامتری متغیر است که به عوامل داخلی و خارجی زیادی بستگی دارد. برای پیش‌بینی نرخ خطای متغیر، از وقایع قبلی و داده‌های آماری آنها استفاده می‌شود. در این مقاله، نرخ خطای متغیر در شبکه‌های توزیع فشار متوسط هوایی در اثر برخورد درختان مورد بررسی قرار گرفته است. از آنجا که جمع‌آوری و تعیین داده‌های آماری همواره با خطاهایی همراه است، برای کاوش در داده‌ها و حذف داده‌های پرت از روش تحلیل عاملی استفاده شده است. سپس برای پیش‌بینی نرخ خطا از یک شبکه عصبی چند لایه پرسپترون استفاده شده است. شبکه عصبی مذکور پس از آموزش، برای تحلیل ورودی‌ها نیز مورد استفاده قرار می‌گیرد. مطالعات عددی بر روی یک شبکه توزیع نمونه با ۳۲ فیدر نشان می‌دهد که روش‌های تحلیل عاملی و شبکه عصبی در تحلیل حساسیت، نتایج یکدیگر را تأیید می‌کنند. ضمناً روش مذکور می‌تواند هم برای مدلسازی ساده‌تر و کاهش پیچیدگی‌ها و هم برای حذف داده‌های پرت و افزایش اطمینان در پیش‌بینی نرخ خطا مورد استفاده قرار گیرد.

واژه های کلیدی: شبکه توزیع، قابلیت اطمینان، کیفیت توان، تحلیل عاملی، کاوش داده‌ها.

تاریخ ارسال مقاله : ۱۳۹۱/۶/۲۵

تاریخ پذیرش مقاله : ۱۳۹۱/۹/۷

نام نویسنده‌ی مسئول : مهدی صدقی

نشانی نویسنده‌ی مسئول : تهران – خیابان دکتر شریعتی – سید خندان – دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی خواجه نصیرالدین طوسی.

۱- مقدمه

قابلیت اطمینان در سیستم‌های قدرت نوین از اهمیت زیادی برخوردار است. یک سیستم مطمئن باید بتواند توان درخواستی مصرف کنندگان را بدون وقفه تأمین کند. معمولاً قابلیت اطمینان در بخش تولید و انتقال دارای اهمیت بیشتری است زیرا با معیوب شدن این بخش‌ها، مصرف کنندگان بیشتری دچار وقفه خواهند شد. با این حال، امروزه بررسی قابلیت اطمینان در سیستم توزیع برق نیز از اهمیت ویژه‌ای برخوردار است. بخصوص شبکه‌های توزیع هوشمند باید دارای قابلیت اطمینان بسیار بالایی باشند.

در این میان پیش‌بینی و تعیین نرخ خرابی سیستم توزیع برای بررسی و برنامه‌ریزی قابلیت اطمینان لازم است. معمولاً نرخ خرابی سیستم را به عنوان پارامتری ثابت در طول عمر مفید سیستم و بر حسب تعداد خطا در واحد طول- زمان (f/year-km) در نظر می‌گیرند. با اینحال نرخ خرابی در عمل پارامتری متغیر است و به عوامل زیادی بستگی دارد. با اینکه خرابی‌های شبکه توزیع به عوامل داخلی مانند نحوه ساخت و بهره‌برداری، عمر تجهیزات سیستم و اندازه آنها بستگی دارند و بسیاری از خطاها موقتی هستند، ولی نرخ خرابی شبکه توزیع به عوامل محیطی مانند باد، بارش، صاعقه، یخ زدگی، تماس درختان، پرندگان و حیوانات نیز وابسته است [۴-۱]. پیش‌بینی نرخ خطا در این حالت بر اساس اتفاقات گذشته و داده‌های آماری صورت می‌گیرد. عوامل تأثیرگذار مستقیم بر نرخ خطا و میزان تأثیرگذاری آنها (بر حسب فرکانس رخداد) برای یک شبکه توزیع نمونه در ایران در [۵] آورده شده است. اثرات حیوانات بر پیش‌بینی نرخ خطا در خطوط هوایی توزیع در [۷] ,[۶] بررسی شده است. در مرجع [۷] از مدل شبکه بیزین برای تحلیل احتمالاتی استفاده شده است و نتایج آن نشان می‌دهد که میزان تأثیر رفتار حیوانات بر نرخ خرابی با شرایط آب و هوایی مثل دما در رابطه است. همچنین تأثیر شرایط آب و هوایی بر نرخ خطا در خطوط هوایی شبکه توزیع در [۱۲-۸] در نظر گرفته شده است. در [۱۰] ,[۹] لزوم بررسی تأثیر هوای نامساعد بر نرخ خطا و مدلسازی آن نشان داده شده است. در مرجع [۱۱] عواملی مانند درجه حرارت، رطوبت، یخ زدگی، سرعت باد و شدت صاعقه مد نظر قرار گرفته‌اند و از دو مدل رگرسیون پواسن و شبکه بیزین[۱] برای تحلیل استفاده شده است. در [۱۲] نیز، با استفاده از روش رگرسیون، تأثیر صاعقه بر قابلیت اطمینان شبکه‌های توزیع، در مناطق مختلف نشان داده شده است.

تأثیر برخورد درختان در نرخ خطا برای خطوط هوایی توزیع نیز در [۱۵-۱۳] بررسی شده است. در [۱۴] تأثیر رشد و هرس درختان بر قابلیت اطمینان خطوط هوایی شبکه توزیع بررسی شده است و برای تعیین زمان‌بندی بهینه هرس درختان از الگوریتم ژنتیک استفاده شده است. در مرجع [۱۵] نیز از روش‌های خطی و غیرخطی ساده و همچنین شبکه عصبی برای پیش‌بینی نرخ خطا استفاده شده است ولی تحلیلی روی داده‌ها صورت نگرفته است. درحالیکه وجود داده‌های آماری نامربوط به صورت داده‌های پرت[۲] در میان اطلاعات خام می‌تواند روند پیش‌بینی را منحرف سازد. بنابراین تحلیل داده‌های آماری قبل از روند پیش‌بینی دارای اهمیت است. در این مقاله از داده‌های آماری مرجع [۱۵] استفاده شده است، ولی بر خلاف روش ارائه شده در [۱۵]، قبل از پیش‌بینی نرخ خطا، داده‌ها مورد تحلیل قرار گرفته‌اند.کاوش داده‌های[۳] آماری می‌تواند موجب حذف داده‌های نامربوط یا کم تأثیر و کاهش ابعاد داده‌ها و در نتیجه ساده‌تر شدن پیش‌بینی نرخ خطا گردد.

در اینجا قابل ذکر است که اگرچه تعداد زیادی از خطاهای ایجاد شده توسط برخورد درختان با خطوط هوایی از نوع موقتی هستند و هزینه قابلیت اطمینان کمتری دارند، ولی موجب وقوع افت ولتاژ[۴] حتی در همه‌ی فیدرهای مجاور می‌شوند [۱۶]. این افت ولتاژها، از منظر کیفیت توان، می‌توانند هزینه‌های اضافی را به شبکه تحمیل کنند [۱۷].

در این مقاله پیش‌بینی و تحلیل نرخ خطا در خطوط هوایی شبکه توزیع، تحت تأثیر رشد و هرس درختان، دمای هوا و میزان بارش انجام شده است. بررسی داده‌های خام با استفاده از تحلیل عاملی[۵] صورت می‌گیرد و از شبکه عصبی، هم برای پیش‌بینی و هم برای آنالیز حساسیت استفاده می‌شود. جهت پیاده‌سازی همه مراحل نیز نرم‌افزار SPSS مورد استفاده قرار گرفته است. در بخش بعدی تأثیر رشد درختان بر نرخ خطای فیدرهای توزیع معرفی و مدلسازی می‌شود. سپس در بخش سوم روش تحلیل عاملی معرفی خواهد شد. مطالعات عددی بر روی یک شبکه نمونه در بخش چهارم نشان داده شده است و نتایج نیز در بخش پنجم آورده شده است.

۲- مدلسازی رشد درختان و نرخ خطا

تا کنون مدل‌های مختلفی برای رشد دینامیکی درختان ارائه شده است که مهمترین آنها بر دو نوع‌اند: مدل مبتنی بر پردازش و مدل مبتنی بر داده. یک مدل مبتنی بر داده به صورت زیر است:

$$y = a(1-e^{-bt})^c \qquad (۱)$$

که در آن y ارتفاع درخت، t عمر درخت و a، b و c پارامترهای رگرسیون هستند.

در عمل رشد درختان به پارامترهای دیگری نظیر دمای هوا و میزان بارش نیز بستگی دارد. به طوریکه می‌توان این پارامترها را نیز در مدل رشد درختان و در نتیجه نرخ خطا وارد کرد. معمولاً زمانی که فیدرهای هوایی ولتاژ متوسط توزیع از یک منطقه پردرخت عبور می‌کنند، برای افزایش قابلیت اطمینان، درختان به طور دوره‌ای هرس می‌شوند. این عملیات هرس کردن باید با برنامه‌ریزی و به موقع انجام شود که وابسته به پیش‌بینی نرخ خطا در شبکه است. زمانی که در یک منطقه درختان هرس می‌شوند، نرخ خطا در آن منطقه کاهش می‌یابد و بعد از مدتی با رشد دوباره درختان به مرور افزایش می‌یابد. بنابراین

نرخ خطای کل متغیر بوده و در صورت تجاوز از حد مجاز باید عملیات هرس درختان تکرار شود.

یک مدل پیشنهادی برای نرخ خطا در این حالت، که در مرجع [۱۵] پیشنهاد شده است، به صورت زیر می‌باشد:

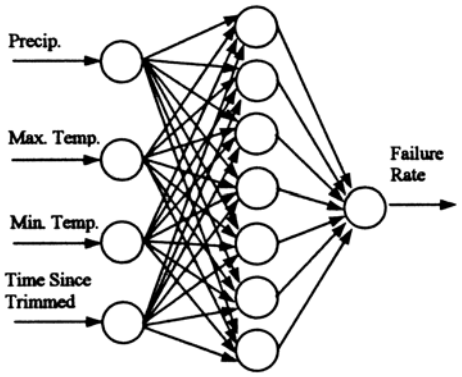
$$\lambda = b_0 + b_1 t + b_2 T_{\min} + b_3 T_{\max} + b_4 P \qquad (۲)$$

که در آن، λ نرخ خطا، t زمان بر حسب سال، $T_{\min}$ و $T_{\max}$ دمای مینیمم و ماکزیمم متوسط روزانه در طول یک سال و P میزان بارش متوسط روزانه در طول سال است. پارامترهای رگرسیون b_0، b_1، b_2، b_3 و b_4 با استفاده از نمونه‌های آماری و مینیمم سازی خطای مجموع مربعات وزن‌دار (SWSE) به صورت زیر به دست می‌آیند.

$$\text{Min. SWSE} = \sum_i w_i (\lambda_{i.est} - \lambda_i)^2 \qquad (۳)$$

که در آن، w_i طول فیدر i ام، λ_i نرخ خطای واقعی و $\lambda_{i,est}$ نرخ خطای تخمین زده شده از معادله (۲) است.

از آنجا که رابطه نرخ خطا با عوامل مؤثر در آن پیچیده است، مدل خطی فوق دارای دقت کافی نخواهد بود. به همین دلیل استفاده از شبکه عصبی به جای مدل خطی معادله (۲) دقت پیش‌بینی را افزایش می‌دهد. این مدل در شکل (۱) نشان داده شده است. در این مقاله از شبکه عصبی سه لایه پرسپترون استفاده می‌شود.



شکل (۱): مدل شبکه عصبی برای پیش‌بینی نرخ خطا

برای تحلیل میزان اثر ورودی‌ها بر نرخ خطای متغیر، می‌توان از آنالیز حساسیت در شبکه عصبی استفاده نمود. این تحلیل، زمانی مفید است که بخواهیم به کاوش داده‌ها بپردازیم و از پیچیدگی اطلاعات بکاهیم. روش‌های مختلفی برای آنالیز حساسیت در شبکه عصبی وجود دارد [۱۸]. روش به کار برده شده در این مقاله به صورت زیر است.

ابتدا برای هر ورودی pام و هر الگوی نمونه‌ی mام، مقدار d_{mp} به صورت زیر محاسبه می‌شود.

$$d_{mp} = \max_{x_{p1}, x_{p2} \in S_p} \left\| \hat{Y}_{p1}^{(m)} - \hat{Y}_{p2}^{(m)} \right\| \qquad (۴)$$

که در آن، $\hat{Y}$ بردار خروجی شبکه عصبی ناشی از بردار ورودی $(x_1^{(m)},...,x_{p-1}^{(m)},x_{pk}^{(m)},x_{p+1}^{(m)},...,x_N^{(m)})$ می‌باشد، به طوریکه مجموعه S_p به صورت $\{(1,0,...,0),(0,1,...,0),...,(0,0,...,1)\}$ تعریف می‌شود و N تعداد کل ورودی‌های شبکه عصبی است.

سپس مقدار d_p به صورت زیر محاسبه می‌شود:

$$d_p = \frac{1}{M} \sum_{m=1}^{M} d_{pm} \qquad (۵)$$

که در آن، M تعداد نمونه‌ها برای آموزش شبکه عصبی است. این تحلیل پس از آموزش شبکه عصبی اجرا می‌شود و مقدار d_p به دست آمده از معادله (۵) میزان اهمیت ورودی pام را نشان می‌دهد. مقادیر d_p را می‌توان نرمالیزه کرد به طوریکه معادله (۶) برقرار شود.

$$\sum_{p=1}^{N} d_p = 1 \qquad (۶)$$

به این ترتیب می‌توان میزان اهمیت هر ورودی را با استفاده از مقادیر نرمال شده‌ی d_p به دست آورد.

۳- تحلیل عاملی

یکی از روش های آماری برای تجزیهٔ اطلاعات موجود در مجموعه داده ها روش تجزیه عامل‌ها یا تحلیل عاملی است. این روش توسط کارل پیرسون و چارلز اسپیرمن برای اولین بار هنگام اندازه‌گیری هوش مطرح شد و برای تعیین تأثیرگذارترین متغیرها در زمانیکه تعداد متغیرهای مورد بررسی زیاد و روابط بین آنها ناشناخته باشد، استفاده می‌شود [۱۹].

روش تجزیهٔ عاملی در واقع گسترش روش تحلیل مؤلفه‌های اصلی (PCA[۶]) است. در هر دو روش تلاش بر آن است که ماتریس کوواریانس تقریب زده شود، اما این تقریب در روش تحلیل عاملی از دقت و ظرافت بیشتری برخوردار است. در این روش متغیرها در عامل‌هایی قرار می‌گیرند، به طوریکه از عامل اول به عامل‌های بعدی درصد واریانس کاهش می یابد، از اینرو متغیرهایی که در عامل‌های اولی قرار می‌گیرند، تأثیرگذارترین هستند. تحلیل عاملی بر دو نوع اکتشافی و تأییدی است. در تحلیل عاملی تأییدی معمولاً یک سری پیشفرض برای رابطه داده‌ها بر اساس تجربه وجود دارد. در حالیکه در نوع اکتشافی، اطلاع قبلی از رابطه عوامل وجود ندارد و باید کشف شود. در این مقاله از تحلیل عاملی نوع اکتشافی استفاده شده است. به طور کلی، هدف از تحلیل عاملی کاهش ابعاد مسأله است، به طوریکه به جای تحلیل تعداد زیادی متغیر، تعداد کمتری عامل[۷] استخراج شده و تحلیل شوند. این عامل‌ها معمولاً مستقل از هم هستند و تحلیل آنها ساده‌تر است.

در تحلیل عاملی، نمونه‌های آماری مربوط به یک گروه از متغیرها را می‌توان به صورت ترکیب خطی از یک سری عوامل و بارها نوشت. این ترکیب خطی برای m نمونه از n متغیر به صورت زیر است:

$$x_{ji} = \mu_j + a_{j1}F_{1i} + \ldots + a_{jt}F_{ti} + \ldots + a_{jk}F_{ki} + d_j U_{ji} \;, i = 1,\ldots,m,\; j = 1,\ldots,n,\; k < n \qquad (۷)$$

که در آن از نظر آماری، x_{ji} نمره‌ی[۸] نمونه i ام در متغیر jام، F_{ti} نمره‌ی نمونه‌ی i ام در t امین عامل مشترک و ضرایب a_{jt} بارهای عاملی هستند. U_{ji} نمره‌ی نمونه‌ی i ام در عامل اختصاصی و d_j بار مربوط به آن است. عامل اختصاصی، عاملی است که تنها در یک متغیر واحد موجود است.

در معادله‌ی (۷)، μ_j میانگین متغیر jام است و نمره‌های x_{ji} و F_{ti} دارای میانگین صفر و واریانس واحد می‌باشند. با مجذور کردن دو طرف معادله‌ی (۷) و سپس جمع آنها بر m مورد و تقسیم آن بر m و این فرض که نمره‌های عاملی ناهمبسته هستند، می‌توان نوشت:

$$S_j^2 = 1 = a_{j1}^2 + a_{j2}^2 + \ldots + a_{jk}^2 + d_j^2 \qquad (۸)$$

این معادله نشان می‌دهد که واریانس کل را می‌توان به دو بخش جمع‌پذیر واریانس مشترک و واریانس منفرد تقسیم کرد.

میزان اشتراک یک متغیر که اغلب با علامت h_j^2 نشان داده می‌شود، برابر است با مجموع مجذورات بارهای عاملی مشترک:

$$h_j^2 = 1 = a_{j1}^2 + a_{j2}^2 + \ldots + a_{jk}^2 \qquad (۹)$$

میزان اشتراک بخشی از واریانس است که می‌توان آن را به عوامل مشترک نسبت داد. بخشی از واریانس که باقی می‌ماند و نمی‌توان آن را به عوامل مشترک نسبت داد، واریانس منفرد نامیده شده و با علامت d_j^2 نشان داده می‌شود. معادلات (۹)-(۷) معادلات پایه در تحلیل عاملی هستند، به طوریکه با اعمال این روابط به داده‌ها، پارامترهای موردنظر تخمین زده می‌شوند. برای جزییات بیشتر به مرجع [۲۰] مراجعه شود.

پیش از آنکه بخواهیم تحلیل عاملی را به داده‌های آماری اعمال کنیم، لازم است از مناسب بودن این داده‌ها برای تحلیل عاملی مطمئن شویم، زیرا هر نوع داده‌ای برای تحلیل عاملی مناسب نیست. در این مقاله از معیار KMO[۹] برای این تحلیل استفاده می‌شود. ضریب KMO بین صفر و یک تغییر می‌کند و از رابطه زیر به دست می‌آید.

$$\mathrm{KMO} = \frac{\sum_i \sum_{j \neq i} r_{ij}^2}{\sum_i \sum_{j \neq i} r_{ij}^2 + \sum_i \sum_{j \neq i} a_{ij}^2} \qquad (۱۰)$$

که در آن r_{ij} ضریب همبستگی ساده‌ی بین متغیرهای i و j، و a_{ij} ضریب همبستگی جزئی بین آنهاست. اگر مجموع ضرایب همبستگی جزئی بین همه زوج متغیرها در مقایسه با مجموع مجذورات ضرایب همبستگی کوچک باشد، اندازهٔ KMO نزدیک به یک خواهد بود. مقادیر کوچک KMO بیانگر آن است که همبستگی بین زوج متغیرها نمی‌تواند توسط دیگر متغیرها تبیین شود، بنابراین کاربرد تحلیل عاملی متغیرها ممکن است قابل توجیه نباشد. معمولاً متغیرهایی با ضریب KMO > 0.5 برای تحلیل عاملی مناسب‌اند [۱۹].

برای انجام تحلیل عاملی ابتدا باید روش استخراج عامل‌ها و معیار تعیین آنها مشخص شود. برای استخراج عامل‌ها روش‌های مختلفی وجود دارد که برحسب مقدار و نوع واریانسی که توسط متغیرهای هر عامل در مدل توجیه می‌شود، متفاوتند. در این مقاله از روش تجزیه مؤلفه‌های اصلی استفاده شده است. در روش تجزیه‌ی مؤلفه های اصلی، عامل‌ها همهٔ واریانس هر متغیر از جمله واریانس مشترک با سایر متغیرهای مجموعه و نیز واریانس خاص خود متغیر را توجیه می‌کنند. پس تعداد عامل‌ها در این روش از نظر تئوری باید با تعداد متغیرها برابر باشد، زیرا همهٔ واریانس هر متغیر باید توسط عامل‌ها تبیین شود. به عبارت دیگر در تجزیهٔ مؤلفه‌های اصلی به تعداد متغیرها، مؤلفه وجود دارد، ولی فقط عامل‌هایی استخراج می شوند که بیشترین مقدار واریانس را تبیین کنند.

برای تعیین عامل‌ها نیز معیارهای مختلفی نظیر معیار مقدار ویژه، معیار پیشین، معیار تست بریدگی و معیار درصد واریانس تجمعی وجود دارند [۱۹]. در این مقاله از معیار مقدار ویژه استفاده شده است. برای به دست آوردن این معیار، مقدار ویژه و بردار ویژه‌ی ماتریس همبستگی متغیرها (X^TX) محاسبه می‌شود [۲۱]. بنابراین تعداد مقادیر ویژه برابر تعداد کل عامل‌هاست. هر چه مقدار ویژه‌ی یک عامل بیشتر باشد، آن عامل واریانس بیشتری را تبیین می‌کند. بر این اساس تعداد عامل‌ها با توجه به مقدار ویژهٔ هر عامل مشخص می‌شود و عامل‌هایی که مقدار ویژه آنها بیشتر از یک باشد، به عنوان عامل های معنی‌دار در نظر گرفته می‌شوند و از مابقی صرفنظر می‌شود. بنابراین ماتریس بارگیری عامل mام را می‌توان از رابطه زیر استخراج کرد:

$$\Lambda_m = \Omega_m \Gamma_m^{1/2} \qquad (۱۱)$$

که در آن، $\Omega_m = (\omega_1, \omega_2, \ldots, \omega_m)$ ، $\Gamma_m = \mathrm{diag}(|\gamma_1|, |\gamma_2|, \ldots, |\gamma_m|)$ و $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_m$ مقادیر ویژه ماتریس همبستگی و $\omega_1, \omega_2, \ldots, \omega_m$ بردارهای ویژه ماتریس همبستگی هستند. در این حالت، مقدار اشتراک متغیر i ام به صورت زیر است:

$$h_i = \sum_{j=1}^{m} |\gamma_j| \omega_{ij}^2 \qquad (۱۲)$$

با توجه به مقادیر ویژه و بردارهای ویژه، عواملی که دارای مقدار ویژه کمتر از یک هستند، در ماتریس بارگیری Λ حذف می‌شوند، بنابراین ماتریس بارگیری Λ معمولاً یک ماتریس ستونی خواهد بود.

با اینحال، تفسیر این خروجی تحلیل عاملی ممکن است پیچیده باشد. به طور مثال زمانی که پس از اجرای تحلیل عاملی، یک متغیر به بیش از یک عامل مربوط شده باشد. در این حالت برای ساده کردن تفسیر تحلیل عاملی از تبدیلاتی نظیر دوران[۱۰] استفاده می‌شود. دوران را می‌توان با روش‌های مختلفی اجرا نمود. در این مقاله از روش واریماکس[۱۱] برای دوران استفاده شده است. واریماکس از جمله

روش‌های متعامد است که استقلال میان عامل‌های استخراجی را حفظ می‌کند. این روش متغیرهای دارای بار عاملی بزرگتر را به کمترین تعداد تقلیل می‌دهد، به طوریکه جمع واریانس بارها در ماتریس عاملی را بیشترین مقدار می‌کند و در این حالت تفسیر عامل‌ها ساده می‌شود [۱۹]. بنابراین، برای اجرای دوران، ماتریس بارگیری Λ باید طوری دوران یابد که معیار واریماکس زیر ماکزیمم شود:

$$\text{Max.}\quad V=\sum(q_{j,l}^2-\bar{q}_{j,l}^2)^2 \qquad (۱۳)$$

که در آن، $q_{j,l}^2$ مجذور مقدار بارگیری متغیر jام در عامل l ام و $\bar{q}_{j,l}^2$ میانگین مجذور آن است.

عملیات اجرای دوران با توجه به مسأله بهینه سازی فوق، به صورت تکراری انجام می‌شود. به این صورت که اگر ماتریس بارگیری Λ یک ماتریس $d\times m$ (که معمولاً $d>m$) باشد، آنگاه برای هر جفت از ستون‌های ماتریس بارگیری (مثلاً $[a_k\ \ a_l]$)، ماتریس دوران $\mathrm{R}_{2\times 2}$ محاسبه شده و در ماتریس $[a_k\ \ a_l]$ ضرب می‌شود. ماتریس R متعامد بوده و از رابطه زیر به دست می‌آید:

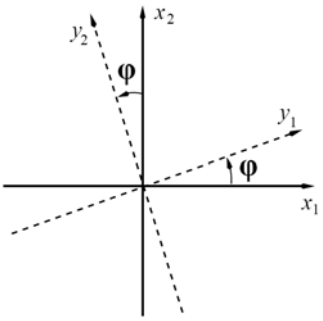
$$\mathrm{R}=\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix} \qquad (۱۴)$$

که در آن، $\varphi=0.25\times\arctan(t/b)$ است و t و b تابعی از عناصر ماتریس بارگیری به صورت زیر هستند [۲۲]:

$$t=2\times[d\times\sum_{j=1}^{d}(a_{jk}^2-a_{jl}^2)\times(2a_{jk}a_{jl})-\sum_{j=1}^{d}(a_{jk}^2-a_{jl}^2)\times(2a_{jk}a_{jl})] \qquad (۱۵)$$

$$b=d\times\sum_{j=1}^{d}[(a_{jk}^2-a_{jl}^2)^2-(2a_{jk}a_{jl})^2]-[\sum_{j=1}^{d}(a_{jk}^2-a_{jl}^2)]^2-(\sum_{j=1}^{d}2a_{jk}a_{jl})^2 \qquad (۱۶)$$

با توجه به معادله‌ی (۱۴)، رابطه‌ی $\mathrm{R}^{\mathrm{T}}\mathrm{R}=\mathrm{I}$ برقرار بوده و لذا ماتریس R متعامد است. عملیات دوران توسط ماتریس R در شکل (۲) نشان داده شده است که در آن y_1 و y_2 محورهای دوران یافته هستند.



شکل (۲): نحوه اجرای دوران توسط ماتریس متعامد *R*

به طور خلاصه عملیات دوران واریماکس را می‌توان به صورت الگوریتم زیر بیان کرد [۲۲]:

گام ۱) ماتریس بارگیری اولیه $\Lambda_{d\times m}$ دریافت شود.

گام ۲) قرار بده $k\leftarrow 1$ و $l\leftarrow 2$.

گام ۳) قرار بده $\mathrm{A}_{d\times 2}=[a_k\ \ a_l]$.

گام ۴) ماتریس R را از روابط (۱۶)-(۱۴) محاسبه کن.

گام ۵) بروزرسانی ماتریس A توسط R: $\mathrm{A}\leftarrow\mathrm{A}\times\mathrm{R}$.

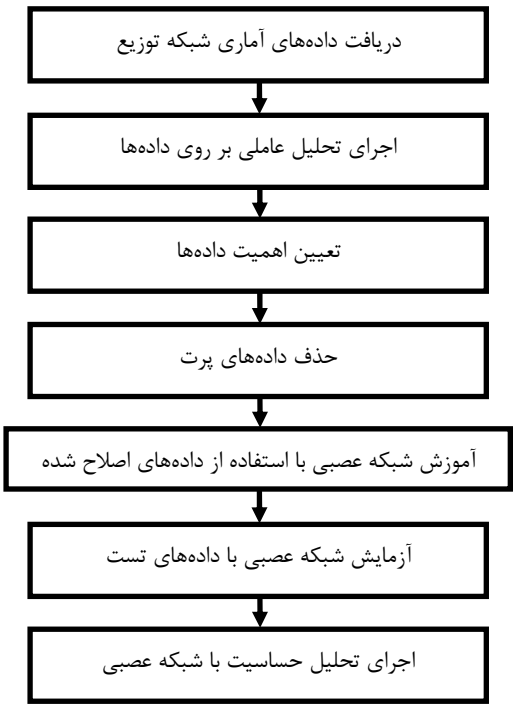
گام ۶) قرار بده $\Lambda(j,k)\leftarrow\mathrm{A}(j,1)$ و $\Lambda(j,l)\leftarrow\mathrm{A}(j,2)$، که $j=1,\ldots,d$.

گام ۷) اگر $l<m$ است، آنگاه قرار بده: $l\leftarrow l+1$ و برو به گام ۳.

گام ۸) اگر $k<m-1$ است، آنگاه قرار بده $k\leftarrow k+1$ و سپس $l\leftarrow k+1$ و برو به گام ۳، وگرنه پایان.

الگوریتم فوق پس از اجرای کامل تجزیه مؤلفه‌های اصلی در تحلیل عاملی انجام می‌شود تا تفسیر خروجی آن راحت‌تر شده و وابستگی بین مؤلفه‌ها به حداقل برسد.

روش تحلیل عاملی معمولاً در علوم آماری به کار برده می‌شود ولی در این مقاله از روش تحلیل عاملی برای کاوش داده‌های خام آماری استفاده می‌شود به طوریکه با کشف عامل‌های مهم در میان داده‌های زیاد ناشی از وجود تعداد زیاد فیدرها در شبکه توزیع، اطلاعاتی که پرت یا نامربوط می‌باشند، شناخته شده و حذف گردند. این اطلاعات نامربوط به دلایل مختلفی مانند وجود نویز در داده‌ها، اتفاقات تصادفی در شبکه، اشتباه در جمع‌آوری اطلاعات و ... ایجاد



شکل (۳): فلوچارت روش پیشنهادی برای تحلیل و پیش‌بینی نرخ خطای تغییرپذیر در شبکه توزیع

می‌شوند. همچنین با این تحلیل می‌توان اهمیت داده‌ها را با یکدیگر مقایسه کرد. روند روش پیشنهادی در شکل (۳) نشان داده شده است. این روند که موجب بهبود دقت پیش‌بینی و تحلیل داده‌ها می‌شود، در سایر مراجع مورد استفاده قرار نگرفته است.

۴- مطالعات عددی

برای بررسی روش‌های فوق در تحلیل و پیش‌بینی نرخ خطای متغیر شبکه توزیع، مطالعات عددی بر روی اطلاعات یک شبکه توزیع با ۳۲ فیدر صورت می‌گیرد که مشخصات آن در مرجع [۱۵] آورده شده است. در این شبکه، نرخ خطای مورد نظر تحت تأثیر خرابی‌های ناشی از برخورد گیاهان و درختان اطراف فیدرهای هوایی است. بنابراین درختان به صورت دوره‌های ۷ ساله اصلاح می‌شوند به طوریکه اصلاحات برای هر فیدر مستقل بوده و پس از سال شماره یک برای هر فیدر، تا هفت سال اقدام به اصلاحات نمی‌شود. اطلاعات آماری مربوط به این شبکه طی چهار سال در جدول (۱) آورده شده است که شامل طول هر فیدر، شماره مربوط به سال اصلاحات برای هر فیدر و نرخ خطای ناشی از برخورد درختان در هر سال است. مشخصات آماری مربوط به شرایط آب و هوایی و محیطی نیز در جدول (۲) آورده شده است.

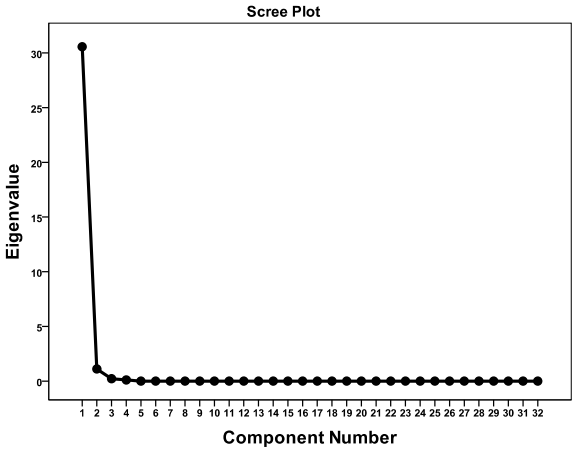
در مرحله اول، پس از محاسبه ضریب KMO، تحلیل عاملی را بر روی داده‌های جدول (۱) اعمال می‌کنیم. در این تحلیل از معیار مقدار ویژه و دوران متعامد واریماکس استفاده می‌شود. نمودار مقادیر ویژه در شکل (۴) نشان داده شده است. این شکل نشان می‌دهد که طبق معیار مقدار ویژه، فقط دو عامل دارای مقدار ویژه بزرگتر از یک و معنی دار هستند.

جدول (۱): مشخصات آماری مربوط به شبکه توزیع تحت مطالعه طی چهار سال [۱۵]

Feeder ID	Length (miles)	Trim				Failure Rates (f/mile-year)			
		1993	1994	1995	1996	1993	1994	1995	1996
1	5.075	5	6	7	1	0.0000	0.1971	0.1971	0.3941
2	3.615	5	6	7	1	0.0000	0.0000	0.0000	0.2638
3	5.894	5	6	7	1	0.3393	0.6786	0.6786	0.0000
4	12.285	5	6	7	1	0.0000	0.0326	0.0814	0.3256
5	7.739	5	6	7	1	0.1657	0.0829	0.1657	0.0000
6	7.575	1	2	3	4	0.2640	0.2640	0.3960	0.0000
7	21.089	7	1	2	3	0.0946	0.2839	0.2366	0.4258
8	24.218	6	7	1	2	0.1228	0.4737	0.2281	0.1930
9	27.322	7	1	2	3	0.1098	0.2196	0.2562	0.6222
10	9.665	5	6	7	1	0.5173	0.8277	0.2069	0.2609
11	6.576	1	2	3	4	0.0000	0.1521	0.0000	0.0000
12	20.223	5	6	7	1	0.4987	0.4080	0.3174	0.4987
13	24.472	5	6	7	1	0.6130	0.8581	0.6538	0.5721
14	28.894	5	6	7	1	0.4845	0.5537	0.4499	0.3807
15	23.579	7	1	2	3	0.7331	0.6284	0.6284	0.5237
16	3.883	1	2	3	4	0.0000	0.2575	0.0000	0.0000
17	30.650	6	7	1	2	0.2936	0.5220	0.4568	0.6199
18	9.106	1	2	3	4	0.1038	0.2076	0.2076	0.2076
19	4.509	5	6	7	1	0.0000	0.4435	0.0000	0.6653
20	28.059	1	2	3	4	0.2851	0.1782	0.3564	0.2138
21	18.408	6	7	1	2	0.1724	0.1034	0.1897	0.1724
22	5.349	2	3	4	5	0.1870	0.3739	0.1870	0.5609
23	8.979	1	2	3	4	0.1112	0.1112	0.1112	0.0000
24	3.289	1	2	3	4	0.0000	0.0000	0.1429	0.0000
25	2.751	2	3	4	5	0.0000	0.0000	0.7271	0.0000
26	3.736	5	6	7	1	0.5353	0.4087	0.0000	0.5353
27	28.974	4	5	6	7	0.8283	0.5867	0.6213	0.4832
28	8.735	1	2	3	4	0.6867	0.5723	0.2289	1.0301
29	22.697	2	3	4	5	0.3525	0.3084	0.2644	0.1322
30	5.884	5	6	7	1	0.0000	1.0309	1.0309	0.3436
31	11.296	2	3	4	5	0.1699	0.5098	0.1699	0.0000
32	2.910	2	3	4	5	0.1771	0.3541	0.0885	0.1771

جدول (۲): مشخصات دما و بارش طی چهار سال برای شبکه مورد مطالعه [۱۵]

Year	Precipitation (in)	T_{max} (°F)	T_{min} (°F)
1993	0.079	59.825	41.416
1994	0.100	60.686	42.788
1995	0.105	61.565	42.987
1996	0.100	62.553	43.938



شکل (۴): نمودار مقادیر ویژه در تحلیل عاملی

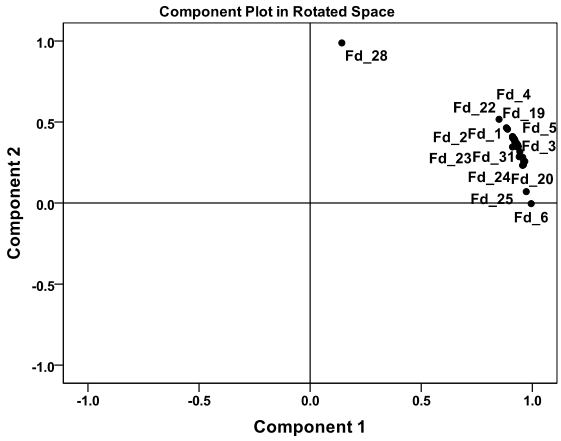
جدول (۳) ماتریس بارگیری نمونه‌ها را بر روی این دو عامل نشان می‌دهد. موقعیت قرارگیری نمونه‌ها در فضای عامل‌ها پس از چرخش نیز در شکل (۵) نشان داده شده است.

جدول (۳): ماتریس بارگیری شماره فیدرها بر روی عامل‌ها

Rotated Component Matrix

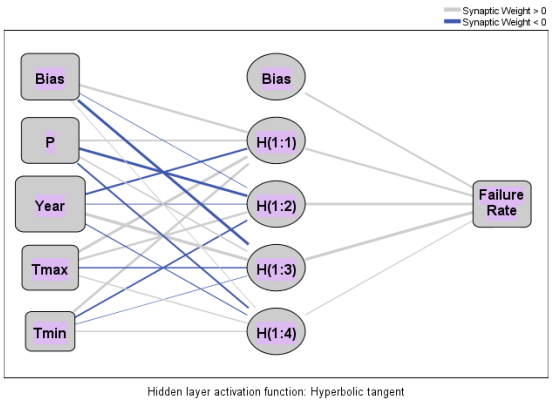
	Component	
	1	2
Fd_6	.995	
Fd_25	.973	
Fd_3	.966	
Fd_24	.961	
Fd_30	.960	
Fd_23	.957	
Fd_20	.956	
Fd_29	.944	.316
Fd_31	.941	
Fd_5	.937	.347
Fd_13	.935	.354
Fd_27	.933	.345
Fd_14	.933	.359
Fd_15	.932	.360
Fd_8	.932	.363
Fd_21	.928	.370
Fd_18	.927	.360
Fd_11	.925	.359
Fd_10	.922	.375
Fd_7	.920	.390
Fd_17	.920	.390
Fd_1	.916	.398
Fd_12	.915	.401
Fd_32	.912	.400
Fd_4	.912	.406
Fd_2	.911	.409
Fd_9	.911	.409
Fd_16	.910	.346
Fd_26	.888	.456
Fd_19	.883	.464
Fd_22	.850	.517
Fd_28		.988

شکل (۵) نشان می‌دهد که فیدر شماره ۲۸ دارای بارگیری متفاوتی نسبت به سایر فیدرهاست که آن را به عنوان یک نمونه پرت معرفی می‌کند. با رجوع به جدول (۱) مشخص می‌شود که علیرغم اجرای اصلاحات در سال اول (۱۹۹۳) برای فیدر ۲۸، نرخ خطا در این سال برای فیدر ۲۸ بسیار بالا بوده و حتی از همه‌ی فیدرهای دیگر نیز بیشتر است. این در حالی است که دمای هوا و میزان بارش برای همه فیدرها در این سال یکسان بوده است. بنابراین تفاوت فاحش اطلاعات برای فیدر ۲۸ نسبت به سایر فیدرها ناشی از دلایل دیگری است. یکی از این دلایل می‌تواند نحوه ثبت اطلاعات باشد، به طوریکه اصلاح درختان در اواخر سال اول انجام شده ولی نرخ خطا برای کل طول سال اول ثبت شده است. اگرچه این موضوع در مرجع [۱۵] بررسی نشده است ولی نتایج این تحلیل نشان می‌دهد که پیش از پردازش داده‌ها (با شبکه عصبی) داده‌های مربوط به فیدر ۲۸ (به عنوان نمونه‌ای ناهمخوان) باید حذف شوند.



شکل (۵): نحوه قرارگیری نمونه‌ها در فضای عامل‌ها

در مرحله‌ی دوم، به پیش‌بینی نرخ خطا با استفاده از شبکه عصبی چند لایه پرسپترون (MLP) می‌پردازیم. ساختار این شبکه عصبی در شکل (۶) نشان داده شده است.



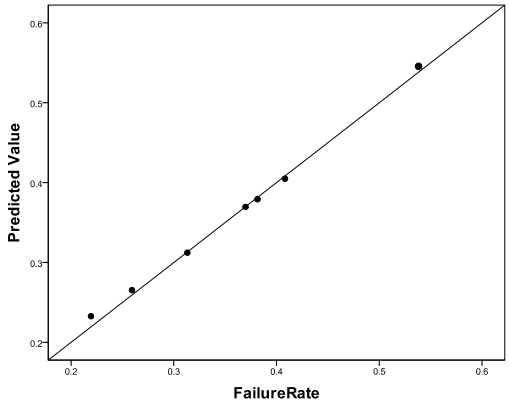
شکل (۶): ساختار شبکه عصبی پرسپترون برای پیش‌بینی نرخ خطا

پس از آموزش شبکه عصبی با استفاده از داده‌های جداول (۱) و (۲)، وزن‌های آن به دست می‌آید که در جدول (۴) نشان داده شده است.

جدول (۴): پارامترهای به دست آمده از آموزش شبکه عصبی

Parameter Estimates

Predictor		Predicted				
		Hidden Layer 1				Output Layer
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	Failure Rate
Input Layer	(Bias)	.710	-.064	-1.030	.119	
	P	.336	-.992	.307	-.399	
	Year	-.481	-.117	1.869	-.120	
	Tmax	1.358	.644	-.275	.230	
	Tmin	.755	-.389	-.027	.221	
Hidden Layer 1	(Bias)					.426
	H(1:1)					.505
	H(1:2)					.897
	H(1:3)					1.530
	H(1:4)					.191

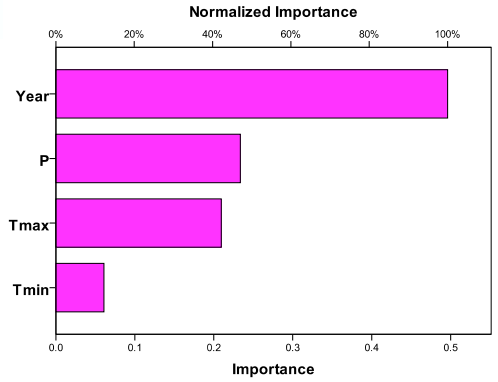


شکل (۷): نرخ خطا (λ) پیش‌بینی شده توسط شبکه عصبی بر حسب مقادیر واقعی آن

شکل (۷) مقادیر پیش‌بینی شده نرخ خطا در شبکه توزیع نمونه را بر حسب مقادیر واقعی آن طی ۷ سال نشان می‌دهد که با همدیگر منطبق‌اند. در این شکل خطای مجموع مربعات برابر ۰/۰۱۱ به دست می‌آید. همچنین این شکل نشان می‌دهد که دقت پیش‌بینی در این مقاله از روش مرجع [۱۵] (که در آن خطای مجموع مربعات برابر ۰/۲۱ می‌باشد) بهتر است.

برای اطمینان بیشتر از صحت پیش‌بینی، شبکه عصبی فوق را با داده‌های ۶ سال آموزش می‌دهیم و سپس نرخ خطا را در سال هفتم (به عنوان داده‌ی تست) پیش‌بینی می‌کنیم. با مقایسه‌ی نتیجه‌ی پیش‌بینی شده با مقدار واقعی نرخ خطا در سال هفتم، مقدار خطای مطلق پیش‌بینی برابر $۲/۹۱۳ \times ۱۰^{-۵}$ می‌شود که نشان دهنده‌ی صحت پیش‌بینی است. این در حالی است که در مرجع [۱۵] این آزمایش انجام نشده است.

پس از پیش‌بینی نرخ خطا با شبکه عصبی، به تحلیل نتایج می‌پردازیم. با استفاده از معادلات (۶)-(۴) مقادیر اهمیت متغیرهای مؤثر در تغییر نرخ خطا را به دست می‌آوریم که نتیجه‌ی آن در شکل (۸) نشان داده شده است.

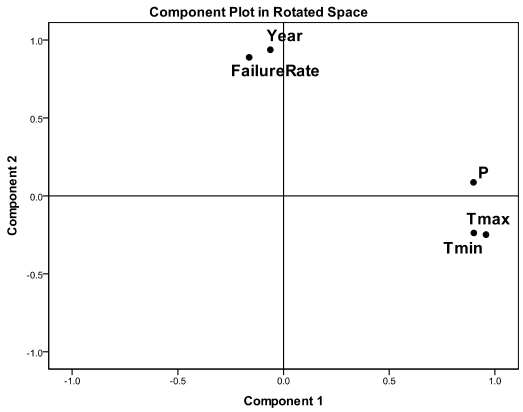


شکل (۸): میزان اهمیت متغیرهای مؤثر در پیش‌بینی نرخ خطا

شکل (۸) نشان می‌دهد که زمان اجرای اصلاح درختان، بیشترین تأثیر را بر مقدار نرخ خطا دارد در حالیکه شرایط آب و هوایی تأثیر کمتری دارند. در میان شرایط آب و هوایی نیز، نرخ بارش سالیانه (در مقایسه با دما) تأثیر بیشتری بر پیش‌بینی نرخ خطا دارد. در واقع افزایش خطای اندازه‌گیری در آن دسته از پارامترهای ورودی که مؤثرتر هستند (یعنی Year و P)، تأثیر بیشتری در افزایش خطای پیش‌بینی خروجی دارد. بنابراین، علیرغم هزینه‌ی بیشتر، لازمست که این ورودی‌ها با دقت بیشتری اندازه‌گیری یا پیش‌بینی شوند.

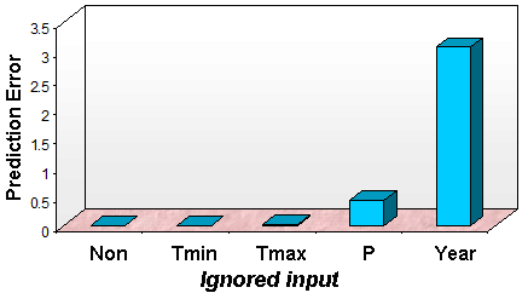
در مرحله سوم، برای اطمینان از صحت نتایج به دست آمده از آنالیز حساسیت در شبکه عصبی، به تحلیل اهمیت عوامل مؤثر در نرخ خطا، با استفاده از تحلیل عاملی می‌پردازیم. پس از محاسبه‌ی ضریب KMO، تحلیل عاملی را اجرا می‌کنیم. در این حالت ضریب KMO برابر ۰/۵۰۴ به دست می‌آید که نشان می‌دهد بکارگیری تحلیل عاملی برای این داده‌ها قابل قبول است. معیار مقدار ویژه، استخراج دو عامل را نتیجه می‌دهد که بارگیری متغیرها بر روی این دو عامل پس از اجرای دوران واریماکس در شکل (۹) نشان داده شده است.

شکل (۹) نشان می‌دهد که رابطه‌ی نرخ خطا با زمان اجرای اصلاحات درختان بیشتر از دما و بارش سالیانه است. این نتیجه با نتایج حاصل از



شکل (۹): خروجی تحلیل عاملی پس از دوران برای متغیرهای مؤثر در نرخ خطا

تحلیل شبکه عصبی مطابقت دارد. بنابراین، با وجود حذف اطلاعات آب و هوایی برای کاهش پیچیدگی مدلسازی، همچنان می‌توان بـه دقـت پیش‌بینی نرخ خطا، امیدوار بود. بـه عنـوان آزمـایش، اگـر هـر یـک از ورودی‌های شبکه عصبی را به صورت جداگانـه حـذف کنـیم و شـبکه عصبی را با داده‌های باقیمانده آموزش داده و تست کنیم، مقدار خطای پیش‌بینی در هر مرحله به صورت شکل (۱۰) به دست می‌آید.



شکل (۱۰): تأثیر حذف هر ورودی بر خطای پیش‌بینی

در شکل (۱۰) خطای پیش‌بینی در هر مورد، با استفاده از مفهـوم خطای مجموع مربعات محاسبه شده است. این شـکل نشـان مـی‌دهـد زمانیکـه هیچیـک از ورودی‌هـا حـذف نشـوند (Non)، مقـدار خطـای پیش‌بینی به کمترین مقـدار خـود مـی‌رسـد. اگـر ورودی‌هـای دمـای مینیمم یـا مـاکزیمم در ورودی شـبکه عصـبی حـذف شـوند، خطـای پیش‌بینی همچنان قابل قبول است، ولـی در صـورت حـذف داده‌هـای مربوط به میزان بارش یـا سـال اصـلاحات، صـحت پـیش‌بینـی معتبـر نخواهد بود.

۵- نتیجه

در این مقاله، پیش‌بینـی و تحلیـل نـرخ خطـا در خطـوط هـوایی توزیع برق در اثر رشد و برخورد درختان مورد بررسی قـرار گرفـت. در این حالت، بر خلاف مدل‌های ثابت معمول، نرخ خطا در شبکه توزیـع، به صورت متغیر با زمان مدلسازی شده اسـت. تـأثیر رشـد درختـان در نرخ خطا، خود تابعی از شرایط آب و هـوایی نظیـر دمـا و بـارش بـاران سالیانه و همچنـین هـرس و اصـلاحات دوره‌ای درختـان اسـت. بـرای تحلیل داده‌ها و همچنین حذف داده‌های پرت، استفاده از تحلیل عاملی پیشنهاد گردید. ضمناً شبکه عصبی چندلایه پرسپترون، نه تنهـا بـرای پیش‌بینی نرخ خطا، بلکه برای تحلیل حساسیت نیز مورد استفاده قـرار گرفت. مطالعات عددی بر روی یک شـبکه توزیـع نمونـه بـا ۳۲ فیـدر، قابلیت روش‌های فوق را نشان می‌دهد. تحلیل انجام شده بـر روی ایـن شبکه نشان می‌دهد که تأثیر زمان انجام اصلاح درختـان بـر روی نـرخ خطا بیشـتر از تـأثیر شـرایط آب و هـوایی اسـت. بـا اسـتفاده از روش پیشنهادی و حذف داده‌های پرت، نرخ خطا با اطمینان و دقت بیشتری پیش‌بینی می‌شود و می‌توان از پیچیدگی داده‌ها نیز کاست. مزیت ایـن عمل در این است که نرخ خطای پیش‌بینـی شـده، بـرای برنامـه‌ریـزی قابلیت اطمینان شبکه و اقدامات به موقع، می‌تواند مورد اسـتفاده قـرار گیرد.

مراجع

[1] R.E. Brown, "*Electric power distribution reliability*", Marcel Dekker, Inc., New York, 2002.

[2] S. Sahai, A. Pahwa, "*Failures of overhead distribution system lines due to animals*", Proceedings of the 36th Annual North American Power Symposium, University of Idaho, Moscow, Idaho, Aug. 2004.

[3] A. Pahwa, "*Effect of environmental factors on failure rate of overhead distribution feeders*", IEEE Power Engineering Society General Meeting, pp. 691-692, Jun. 2004.

[4] M.-R Haghifam, E. Akhavan-Rezai, A. Fereidunian, "*Failure rate modeling: a non-parametric data mining approach to MV network field data*", IEEE Electrical Power and Energy Conf., pp. 1-6, 2009.

[5] E. Akhavan-Rezai, M.R. Haghifam, A. Fereidunian, "*Data-driven reliability modeling, based on data mining in distribution network fault statistics*", IEEE Power Tech. Conf., Bucharest, June 28 - July 2, 2009.

[6] M.-Y. Chow, S. O. Yee, L.S. Taylor, "*Recognizing animal-caused faults in power distribution systems using artificial neural networks*", IEEE Trans. on Power Delivery, Vol . 8, No. 3, pp. 1268-1274, 1993.

[7] S. Sahai, A. Pahwa, "*A probabilistic approach for animal-caused outages in overhead distribution systems*", 9th Int. Conf. on Probabilistic Methods Applied to Power Syst., Stockholm, Jun. 11-15, 2006.

[8] Y. Zhou, A. Pahwa, S. Das, "*Prediction of weather-related failures of overhead distribution feeders*", Proc. Probability Methods Applications Power Systems Conf., Ames, IA, Sep. 2004.

[9] M. H. J. Bollen, "*Effects of adverse weather and aging on power system reliability*", IEEE Trans. on Ind. Appl., Vol. 37, No. 2, pp. 452-457, 2001.

[10] P. Wang, R. Billinton, "*Reliability cost/worth assessment of distribution systems incorporating time-varying weather conditions and restoration resources*", IEEE Trans. on Power Delivery., Vol. 17, No. 1, pp. 260-265, 2002.

[11] Y. Zhou, A. Pahwa, S.Yang, "*Modeling weather-related failures of overhead distribution lines*", IEEE Trans. on Power Syst., Vol. 21, No. 4, pp. 1683-1690, 2006.

[12] J. McDaniel, C. Williams, A. Vestal, "*Lightning and distribution reliability – a comparison of three utilities*", Proceedings of IEEE PES T&D Conference and Expo, Dallas, TX, Sept. 2003.

[13] L. Xu, M.-Y. Chow, L. S. Taylor, "*Analysis of tree-caused faults in power distribution systems*", Proceedings of the 35th North American Power Symposium, University of Missouri – Rolla, Rolla, MO, Oct. 2003.

[14] P. A. Kuntz, R. D. Christie, S. S. Venkata, "*Optimal vegetation maintenance scheduling of overhead electric power distribution systems*", IEEE Trans. on Power Delivery, Vol. 17, No. 4, pp. 1164-1169, 2002.

[15] D.T. Radmer, P.A. Kuntz, R.D. Christie, S.S. Venkata, R.H. Fletcher, "*Predicting vegetation-related failure rates for overhead distribution feeders*", IEEE Trans. on Power Delivery, Vol. 17, No. 4, pp. 1170-1175, 2002.

[16] R. Mohammadi, M. Sedghi, M.-R. Haghifam, M. Aliakbar-Golkar, M. Akhbari, "*Optimal distribution*

network planning considering DVR", 25th Int. Power Syst. Conf. (PSC), No. 10-F-PDS-1463, Tehran, Iran, 2010.

[17] J.V. Milanovic, Yan Zhang, "*Global minimization of financial losses due to voltage sags with FACTS based devices*", IEEE Trans. on Power Delivery., Vol. 25, No. 1, pp.298-306, 2010.

[18] C. Vasilakos, K. Kalabokidis, J. Hatzopoulos, I. Matsinos, "*Identifying wildland fire ignition factors through sensitivity analysis of a neural network*", Nat Hazards, Vol. 50, No. 1, pp. 125-143, 2009.

[19] http://utcan.ut.ac.ir/member/syllabus/mazare/multivariate.pdf

[20] H.H. Harman, "*Modern Factor Analysis*", Chicago: University of Chicago Press., 3rd Ed., 1976.

[21] I.T. Jolliffe, "*Principal Component Analysis*". Springer-Verlag, 1986.

[22] G.H. Golub and C.F. Van Loan, "*Matrix Computations*", Baltimore, MD: Johns Hopkins University Press., 3rd Ed., 1996.

زیرنویس‌ها

[1] Bayesian
[2] Outlier
[3] Data mining
[4] Voltage dip
[5] Factor analysis
[6] Principal Component Analysis
[7] Factor
[8] Scale
[9] Kaiser Meyer Olkin
[10] Rotation
[11] Varimax