

بررسی احتمال وقوع روانگرایی و تخمین اهمیت نسبی پارامترهای مؤثر با استفاده از خوشه‌بندی فازی و برنامه‌ریزی ژنتیک

آرمین صاحب‌کرم علمداری^۱ و امیر نجفی^{۲*}

^۱ دانشجوی دکترای مهندسی عمران خاک، گروه مهندسی عمران، واحد تبریز، دانشگاه آزاد اسلامی تبریز

^۲ استادیار، گروه مهندسی صنایع، واحد زنجان، دانشگاه آزاد اسلامی زنجان

(دریافت: ۹۵/۱/۲۷، پذیرش: ۹۵/۷/۱۷، نشر آنلاین: ۹۵/۷/۱۸)

چکیده

روانگرایی خاک در اثر حوادث طبیعی یا انسان‌ساز به دلیل وارد آوردن خسارات مالی و گاهی جانی، جزء مباحث مهمی بوده است که در چند دهه اخیر توجه محققین امر را به خود جلب کرده است. در این بین بررسی احتمال وقوع روانگرایی خاک در اثر تحرکات لرزه‌ای زمین و البته شناخت پارامترهای مؤثر بر این امر بسیار حائز اهمیت می‌باشد. لذا در این مطالعه، ابتدا با استفاده از روش خوشه‌بندی فازی، داده‌های مطالعه موردی به گروه‌هایی از اعضای مشابه با خاصیت فیزیکی یکسان تقسیم گردیدند. سپس با توجه به نتایج خوشه‌بندی و بررسی داده‌های طبقه‌بندی شده در هر خوشه میزان حساسیت و تأثیر هر یک از پارامترهای مؤثر در روانگرایی خاک پرداخته شد. نتایج حاصل نشان داد گروه سه خوشه‌ای بهترین حالت را از نظر تفکیک داده‌ها با ویژگی‌های نزدیک به هم در هر خوشه ارائه کرده است. نتایج تحلیل حساسیت پارامترها نیز حاکی از تأثیر بالای پارامترهای عدد آزمایش نفوذ استاندارد (SPT) اصلاحی و حداکثر شتاب زلزله و تأثیر ناچیز عمق خاک در روانگرایی خاک بود. همچنین سعی شد با استفاده از داده‌های محلی آزمایش SPT جمع‌آوری شده از روانگرایی‌های به وقوع پیوسته در اثر زلزله در نقاط مختلف جهان (نیگاتای ژاپن و چی-چی تایوان)، مدلی به منظور بررسی احتمال وقوع روانگرایی با استفاده از روش برنامه‌ریزی ژنتیک توسعه داده شود. نتایج به دست آمده حاکی از دقت قابل قبول تخمین انجام شده به منظور احتمال وقوع روانگرایی بوده که مدلی با دقت معادل ۹۴/۷۷ درصد ارائه کرد.

کلیدواژه‌ها: احتمال وقوع، برنامه‌ریزی ژنتیک، خوشه‌بندی فازی، روانگرایی خاک.

۱- مقدمه

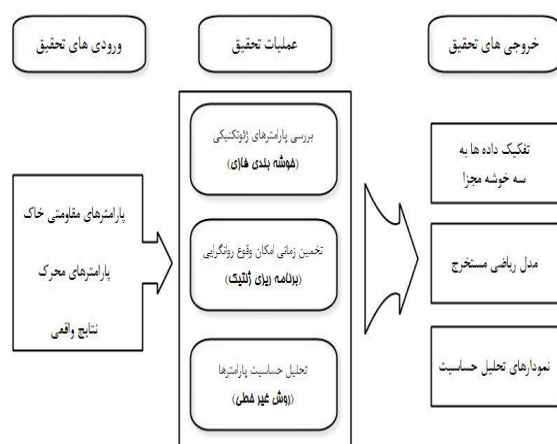
سیکلیک منحصربه‌فرد که تحریک زلزله را شبیه‌سازی می‌کرد نمود. Seed و همکاران (۱۹۸۳) رابطه روانگرایی را با استفاده از نتایج آزمون‌های درجای SPT و آزمایش نفوذ مخروط (CPT) و با بررسی پارامترهای تأثیرگذار در روانگرایی خاک، روابط جدیدی را برای ارزیابی پتانسیل روانگرایی توسعه دادند. Mourad Zeghal و Ahmet-W. Elgamel (۱۹۹۴) با استفاده از داده‌های زلزله سال ۱۹۸۷ برای تخمین رابطه مستقیم میان تنش- کرنش برشی وابسته به زلزله و تأثیر تاریخچه مسیر تنش اقدام نمودند. Olsen و Koester (۱۹۹۵) و Ishihara (۱۹۹۶) روابط خود با استفاده از داده‌های تجربی اضافه شده از روانگرایی‌های به وقوع پیوسته و داده‌های آزمایشگاهی تعریف نمودند. در سال‌های بعد

در طی چند دهه گذشته محققین متعددی مطالعات خود را در جهت توسعه رابطه بین مقاومت روانگرایی و پارامترهای لرزه-خیزی خاک‌های متعدد انجام داده‌اند. یکی از روش‌های ارزیابی روانگرایی که به صورت گسترده از آن استفاده شده است روش مبتنی بر تنش است، که برای اولین بار توسط Seed و Idriss (۱۹۷۱) و Whitman (۱۹۷۱) ارائه شده است. این روش تجربی بوده و بر پایه مشاهدات آزمایشگاهی و محلی استوار هستند. Castro (۱۹۸۲) با استفاده از آزمایش‌های فیزیکی در شرایط زهکشی نشده اقدام به محاسبه پتانسیل روانگرایی خاک تحت بار

* نویسنده مسئول؛ شماره تماس: ۰۲۴-۳۳۴۲۱۰۰۳

حالی که خاک‌ها ساختارهای کاملاً پیچیده‌ای دارند، که از آن جمله می‌توان به مشخصات فیزیکی مبهم و تغییرات نامحدود ایجاد شده توسط ساختار فیزیکی آنها اشاره کرد. بنابر این، مشخصات مکانیکی و دینامیکی خاک‌ها نشان دهنده رفتار غیر قابل اطمینان و متمایز در مقایسه با بیشتر مصالح مهندسی دیگر است. روش‌های بررسی ساده‌تری وجود دارند که اجازه توسعه نامحدود مدل برای تمام سیستم‌ها را می‌دهند، همانند برنامه‌ریزی ژنتیک (GP)، شبکه‌های عصبی مصنوعی (ANNS) و سیستم استنتاج عصبی- فازی (ANFIS) (Shahin و همکاران، ۲۰۰۳، Adel M. Hanna و همکاران، ۲۰۰۶، Tutmaz و Terzan، ۲۰۰۷، Kayadelen، ۲۰۰۸، Katadelen و همکاران، ۲۰۰۹).

در مطالعه حاضر سعی شده است تا در ابتدا با استفاده از خوشه- بندی فازی که در طی آن، داده‌ها به گروه‌هایی که اعضای آن مشابه و تقریباً هم خاصیت یکدیگر باشند تقسیم شوند. بنابر این می‌توان داده‌های با خصوصیات ژئوتکنیکی و لرزه‌ای نزدیک به هم را در یک خوشه قرار داد و همچنین امکان بررسی دقیق‌تر این داده‌ها به منظور شناخت هرچه بهتر عوامل تأثیرگذار بر روانگرایی خاک فراهم کرد. در ادامه مدلی برای تخمین زمانی احتمال وقوع روانگرایی خاک‌ها با روش برنامه‌ریزی ژنتیک ارائه می‌شود. در انتها با توجه به نتایج خوشه‌بندی انجام شده و بررسی داده‌های موجود در هر خوشه با استفاده از روش‌های غیر خطی به تحلیل حساسیت پارامترهای مؤثر در روانگرایی خاک و میزان اهمیت نسبی هر یک از آنها پرداخته می‌شود. شکل (۱) به صورت خلاصه بیانگر روند پیشران مطالعه حاضر می‌باشد.



شکل ۱- شماتیک پیشران تحقیق

این روش بر اساس مطالعات جدید و اطلاعات به دست آمده از روانگرایی‌های به وقوع پیوسته توسط محققین دیگری اصلاح شده است (NCEER، ۱۹۹۷ و Youd و همکاران، ۲۰۰۱). Green و همکاران (۲۰۰۰) مدل تولید فشار منفذی برپایه انرژی (GMP) که یک مدل تجربی برای ایجاد رابطه بین پراکندگی انرژی برشی و نسبت افزایش فشار منفذی پس‌ماند r_u می‌باشد ارائه نمودند. Cetin و همکاران (۲۰۰۴) به بررسی روابط جدیدی برای ارزیابی احتمال آغاز روانگرایی نمودند، که منجر به حذف و کاهش مقدار زیادی از عدم قطعیت‌ها و مغایرت‌ها در روابط قبلی شده است. Muhsung Chang و همکاران (۲۰۱۱) به بررسی روش‌های تجربی موجود برای برآورد پتانسیل روانگرایی با استفاده از داده‌های مربوط به زلزله تایوان با هدف یافتن روشی مناسب برای پیش‌بینی روانگرایی در کشور تایوان پرداختند و در نتیجه برای عمق‌های مختلف خاک میزان درستی هر یک از روش‌ها بررسی شده است. Hadi Shahir و همکاران (۲۰۱۲) به بررسی تغییر در نفوذپذیری خاک در طول روانگرایی و اثرات آن بر پاسخ لرزه‌ای خاک با استفاده از تجزیه و تحلیل پویا را به طور کامل مورد مطالعه قرار گرفته است. نتایج شبیه‌سازی عددی و اندازه‌گیری با استفاده از سانتریفیوژ نشان می‌دهد که در تمام مراحل اتلاف و روانگرایی رابطه مستقیمی ما بین ضریب نفوذپذیری و نسبت فشار آب حفره‌ای وجود دارد. Hariharan و Pijush Samui (۲۰۱۴) با یک مدل طبقه‌بندی متحد به نام ماشین احتمال مینیماکس (MPM) برای مدل‌سازی پتانسیل روانگرایی لرزه‌ای خاک مبتنی بر CPT اقدام کرده‌اند. KumarMuduli و SaratKumarDas (۲۰۱۵) با مدل عدم قطعیت روش مبتنی بر SPT برای ارزیابی پتانسیل لرزه- ای روانگرایی خاک با استفاده از برنامه‌ریزی ژنتیک چند ژن اقدام کرده‌اند.

Guoxing Chen و همکاران (۲۰۱۵) اقدام به واسنجی یک مدل CRR بر پایه یک پایگاه داده مبتنی بر SPT گسترش یافته برای ارزیابی پتانسیل روانگرایی خاک کرده‌اند. Chen و همکاران (۲۰۱۵) یک روش عددی برای آنالیز پایداری از راه حل‌های معادلات حاکم بر پایداری روانگرایی خاک با استفاده از یک روش المان محدود بر پایه طرح گسسته جز ترکیبی را ارائه کرده‌اند. Zhang و همکاران (۲۰۱۶) با استفاده از تغییران درون منطقه‌ای روش روبرتسون و راید برای آنالیز خطر روانگرایی خاک پرداخته‌اند.

به هر حال روابط توسعه داده شده با استفاده از روش‌های سنتی برپایه داده‌های تجربی محدود شده و پیش‌بینی درست و پایداری را فراهم نمی‌کنند و همچنین ضعف عمده روابط ریاضی، ساختار کامل و فضای ناهمگن از خاک و روانگرایی را نشان نمی‌دهند، در

۲- روش انجام تحقیق

۲-۱- خوشه‌بندی

خوشه‌بندی یکی از شاخه‌های یادگیری بدون نظارت و فرآیند خودکاری می‌باشد که در طی آن، داده‌ها به گروه‌هایی که اعضای آن مشابه و تقریباً هم خاصیت یکدیگر باشند تقسیم می‌شوند. به هر یک از این دسته یا گروه‌های مشخص خوشه گفته می‌شود. بنابراین خوشه مجموعه‌ای از داده‌ها و اطلاعات می‌باشد که در آن اعضای مجموعه دارای خصوصیتی مشابه و یکسان بوده و با نمونه‌های موجود در خوشه‌های دیگر غیر همسان و نامشابه می‌باشند. به منظور پی بردن به شباهت هر یک از داده‌ها و تشخیص چگونگی آن می‌توان معیارهای مختلفی را در نظر گرفت (Bezdek, ۱۹۸۱). طرز کار اغلب الگوریتم‌های اجرایی در خوشه‌بندی بدین صورت است که، در ابتدا یک نماینده برای هر گروه به عنوان داده یا نمونه-ای که میانگین خصوصیات را دارا باشد تعیین می‌شود، و سپس مقایسه دیگر نمونه‌ها از آن جهت که به چه میزان با نمونه نماینده شباهت داشته‌اند، صورت می‌پذیرد. نسبت به شباهت و یکسان بودن نمونه‌ها با نمونه نماینده و برگزیده تکلیف قرارگیری هر کدام در دسته مربوطه مشخص می‌گردد. بعد از این مرحله نماینده‌های جدید از روی نمونه‌های متعلق به خوشه محاسبه شده و دوباره نمونه‌ها با این نماینده‌ها مقایسه می‌شوند تا مشخص شود که به کدام خوشه تعلق دارند و این کار آنقدر تکرار می‌شود تا زمانی که نماینده‌های مشخص شده برای هر خوشه تغییری نکنند (Bezdek, ۱۹۸۱).

هدف از خوشه‌بندی یافتن دسته‌هایی مشابه از نمونه‌ها در بین داده‌های ورودی می‌باشد. برای به دست آوردن یک معیار خاص برای پی بردن به چگونگی کیفیت خوشه‌بندی و این که با چه تعداد خوشه باید به کار ادامه دهیم می‌توان گفت، در واقع هیچ معیار مطلقی برای بهترین خوشه‌بندی وجود ندارد بلکه این بستگی به نظریه‌های تخصصی در بحث مربوطه و مقایسه نمونه‌های قرار گرفته در یک خوشه دارد که آیا نمونه‌ها به درستی در یک خوشه قرار داده شده‌اند یا خیر (Bezdek, ۱۹۸۱).

به طور کلی می‌توان گفت تعداد خوشه‌های انتخاب شده برای هر خوشه‌بندی، تراکم نمونه‌ها در هر خوشه را نشان می‌دهد. همچنین قرارگیری نمونه‌هایی با خصوصیات مختلف که می‌توان از مقایسه با حالت طبیعی دریافت شود، معیار مناسبی برای خوشه-بندی صحیح می‌باشد (Bezdek, ۱۹۸۱).

مهم‌ترین ویژگی خوشه‌بندی فازی در آن است که ورودی‌ها اجازه حضور در خلال خوشه‌های متعدد با درجه عضویت متعدد را خواهند داشت و این در حالی است که در خوشه‌بندی‌های مرسوم

(Hard clustering)، هر ورودی تنها اجازه حضور در بین یک خوشه را دارد (Bezdek, ۱۹۸۱).

در سال‌های اخیر الگوریتم‌های متعددی از انواع خوشه‌بندی توسعه یافته است که از آن جمله می‌توان به موارد زیر اشاره نمود: الگوریتم خوشه‌بندی فازی (Javadi و همکاران، ۲۰۰۶)، الگوریتم خوشه‌بندی قله‌ای (Filev و Yager, ۱۹۹۴)، خوشه‌بندی کاهشی (Chiu, ۱۹۹۴).

۲-۱-۱- الگوریتم خوشه‌بندی فازی C-Mean

روش C-Mean یکی از مرسوم‌ترین روش‌های الگوریتم خوشه-بندی فازی و کاوش اطلاعات می‌باشد که در ابتدا توسط دون (۱۹۷۳) ارائه شده و سپس با استفاده از نظریات و به حالت کنونی خود تبدیل گشت (Dunn, ۱۹۷۳). مفهوم اصلی این خوشه توسط روسپینی (۱۹۶۹) در روابط زیر تعریف شده است (Bezdek, ۱۹۸۱).

در حالی که $X = \{X_1, X_2, \dots, X_n\}$ مجموعه داده‌ها باشد و در آن هر داده به صورت X_k ($k = 1, \dots, n$) برداری از فضای $\mathbb{R}^p$ محسوب گردد، U_{cn} مجموعه‌ای از ماتریس حقیقی $c \times n$ و C یک عدد صحیح به صورت $2 \leq c < n$ خواهد بود. فضای در نظر گرفته شده برای X به صورت زیر می‌باشد:

$$M_{f cn} = \{U \in U_{cn} : u_{ik} \in [0, 1] \sum_{i=1}^c u_{ik} = 1, 0 < \sum_{k=1}^n u_{ik} < n\} \quad (۱)$$

u_{ik} عضوی به میزان X_k در خوشه i ($i = 1, \dots, c$) می‌باشد. پیدا کردن بهینه‌ترین گونه خوشه‌بندی فازی یکی از اهداف مهم برای روش C-Mean می‌باشد. به طوری که تابع هدف این روش اختصاراً به نحو زیر تشریح می‌گردد:

$$J_m(U, V; X) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|x_k - v_i\|^2 \quad (۲)$$

در رابطه فوق، $V = (v_1, v_2, \dots, v_c)$ ماتریسی برای مرکز ناشناخته خوشه‌بندی، $v_i \in \mathbb{R}^p$ ، $\| \cdot \|$ نرم اقلیدس، و در نهایت توان وزن m در بازه $[1, \infty)$ ، مقداری پیوسته در بین مقادیر دیگر اعضای می‌باشد. به منظور ساده‌تر کردن رابطه J_m ، در شرایط تعریف شده رابطه (۱)، بزداک (۱۹۸۱) الگوریتم اجرایی زیر را به عنوان متغیرهای ساده شده الگوریتم خوشه‌بندی فازی معرفی نمود (Ruspini, ۱۹۶۹). با انتخاب m, c ، و هنگامی که $\varepsilon > 0$ به عنوان یک مقدار ثابت بسیار کوچک، تولید نسل تفکیکی فازی به صورت U^0 انجام می‌گیرد. با انتخاب میزان تکرار به اندازه $t = 0$ و بر

شود (Koza و Poli، ۲۰۰۵). به هر حال، تنظیم شاخه ممکن است شامل تعدادی از اعداد ثابت و ورودی‌های بیرونی برنامه باشد.

در GP، ابتدا بلوک‌های موجود که شامل متغیرهای ورودی و هدف و نیز تابع ارتباط دهنده آنها می‌باشند، تعریف گردیده و سپس ساختار مناسب الگو و ضرایب آن تعیین می‌شود. این روش شامل یک معادله ارتباط دهنده بین متغیرهای ورودی و خروجی بوده، لذا قادر به انتخاب خودکار متغیرهای مناسب الگو و حذف متغیرهای غیر مرتبط است، که این امر سبب کاهش ابعاد متغیرهای ورودی خواهد شد. انتخاب ورودی‌های مناسب، یکی از مهم‌ترین مواردی است که بایستی در این روش مورد توجه قرار گیرد. این امر در شرایطی که از داده‌های ورودی مختلفی بهره‌برده می‌شود، از اهمیت مضاعفی برخوردار خواهد شد. چرا که ارائه داده-های ورودی غیر مرتبط، سبب کاهش دقت الگو و ایجاد الگوهای پیچیده‌تری می‌شود که تفسیر آنها با دشواری‌های بیشتری مواجه است. فرآیند گام به گام برنامه ریزی ژنتیک به صورت مراحل زیر است:

- ۱) یک جمعیت اولیه از توابع مرکب نشان دهنده الگوهای پیش-بینی، به صورت تصادفی در نظر گرفته می‌شود (ایجاد کروموزوم‌ها)
- ۲) معرفی جمعیت اولیه (کروموزوم‌ها) به رایانه و ارزیابی هر یک از افراد (ژن) جمعیت مذکور با استفاده از توابع برازش (شناسایی مؤثرترین افراد در ماهیت پدیده)
- ۳) انتخاب ژن‌های مؤثر به منظور تکثیر، جهش، جفت‌گیری و تولید مثل افراد جدید با صفات اصلاح شده (فرزندان)
- ۴) اعمال فرآیند توسعه‌ای تکراری بر روی فرزندان در هر تولید گام چهارم به تعداد معین و یا تا حصول بهترین پاسخ تکرار خواهد شد (Koza، ۲۰۰۸؛ Liang و همکاران، ۲۰۰۲).

شکل (۲) نمایش دهنده روندی عمومی GP است، که حاوی روش‌هایی از برنامه‌ریزی ژنتیک می‌باشد. جزئیات بیشتر در مورد برنامه‌ریزی ژنتیک و عملگرهای آن می‌تواند در Koza (۱۹۹۲) یافت شود.

۲-۳- تحلیل حساسیت با لینگو

یکی از مواردی که در مسائل بهینه‌سازی مطرح است تحلیل حساسیت سنجی تابع هدف و یا خروجی در نقطه بهینه نسبت به پارامترهای ورودی است. با توجه به این که پارامترهای ورودی در شرایط یکسان در نظر گرفته شده‌اند و بدون هیچ پیش فرضی نسبت به یک پارامتر مدل‌سازی شده‌اند لذا می‌بایستی تحلیل حساسیت نسبت به همه پارامترها صورت گیرد که این کار را نرم‌افزار لینگو انجام می‌دهد (در قسمت رنج یا دامنه ایجاد شده است که به راحتی می‌توان انجام داد).

اساس مقدار عضویت مشخص $u_{ik}^{(t)}$ ، مرکز هر خوشه $v_i^{(t)}$ ($i = 1, \dots, c$) به صورت زیر محاسبه می‌گردد:

$$v_i^{(t)} = \frac{\sum_{k=1}^n (u_{ik}^{(t)})^m x_k}{\sum_{k=1}^n (u_{ik}^{(t)})^m} \quad (3)$$

مرکز هر خوشه که در دوره‌های تکراری تغییر می‌یابد میزان قبلی عضویت یعنی $u_{ik}^{(t)}$ را به صورت زیر به روز خواهد کرد:

$$u_{ik}^{(t+1)} = \left[\sum_{j=1}^c \left(\frac{\|x_k - v_i^{(t)}\|^2}{\|x_k - v_j^{(t)}\|^2} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (4)$$

این روند در صورتی که به میزان مشخصی از تکرار یا $|U^{(t+1)} - U^{(t)}| \leq \varepsilon$ رسیده باشد، متوقف خواهد شد.

۲-۲- برنامه‌ریزی ژنتیک

برنامه‌ریزی ژنتیک (GP) یک ساختار آموزشی الهام گرفته از علم زیست‌شناسی برای آنالیز داده‌های آماری و الگوی تشخیص می‌باشد. GP با موفقیت بسیار بالایی برای حل انواع مختلفی از مسائل پیچیده مهندسی ژئوتکنیک استفاده شده است. از آن جمله می‌توان به مواردی اشاره کرد همانند: پیش‌بینی رونگرایی که منجر به جابه‌جایی‌های افقی می‌شود، نشست پی‌های عمیق و نسبت V_{max}/a_{max} برای حرکت‌های شدید زمین (Javadi و همکاران، ۲۰۰۶؛ Rezania و Javadi، ۲۰۰۷؛ Jafarian و همکاران، ۲۰۱۰). Koza (۱۹۹۲) برای اولین بار این روش مفید را معرفی کرد، که شاخه‌ای از ژنتیک الگوریتم (GA) متداول توسعه داده شده توسط Holland (۱۹۷۵) می‌باشد.

تفاوت اصلی بین GP و GA نحوه ارائه روش حل می‌باشد. حلقه‌ای از اعداد برای ارائه راه حل GA ایجاد می‌شود، در حالی که مراحل GP در برنامه‌های کامپیوتری در حالت عادی به صورت ساختار درختی ارائه شده است. GAها به صورت عمومی در بهینه‌سازی پارامترها برای در نظر گرفتن بهترین مقدار برای تنظیمات داده شده برای پارامترهای مدل استفاده می‌شوند، در حالی که GP از سوی دیگر، یک ساختار پایه از مدل تقریبی با استفاده از مقادیر پارامترها را نیز می‌دهد (Torres و همکاران، ۲۰۰۹).

در GP، مسائل در ساختارهای درختی ارائه می‌شود و تا اندازه‌ای جمعیت بزرگ اتفاقی از اجزا جدا از هم بر پایه درخت با میزان اختلاف بالا ایجاد شده است. هر عضو جمعیت متشکل از عملگرها و سرشاخه‌ها هستند؛ در حالتی که عملگرها می‌تواند برپایه عملکردهای ریاضی (جمع و ضرب و غیره) و همچنین عامل منطق برلین (And و OR و غیره) و یا هر نمایه تعریف شده دیگر انتخاب

شده است (Chen و همکاران، ۲۰۰۵). همچنین در این مطالعه از نتایج ۸۲ آزمایش که در ۴۳ منطقه مختلف ثبت شده مورد استفاده قرار گرفته است. این نتایج شامل ۵۹ داده روانگرا شده و ۲۳ داده روانگرا نشده می‌باشد. این داده‌ها مربوط به زلزله به وقوع پیوسته در سال ۱۹۹۹ در منطقه چی-چی تایوان با بزرگای $M_w=7/3$ می‌باشد، که منجر به تلفات بسیار و روانگرایی خاک در شهرهای تایچونگ، تانتوا، ژانگهوا، یانلین و جیایی در مرکز تایوان شده، استخراج شده است (Bolton, ۱۹۸۵). در مجموع داده‌های مورد استفاده ۱۷۴ مورد می‌باشد.

جدول ۱- دامنه تغییرات پارامترهای ورودی

min	max	
۰/۹۱	۱۸	Depth
۱۵/۸	۳۵۳/۴۵	σ
۱۵/۸	۳۱۵/۸۸	σ'
۰	۹۹	FC (%)
۲	۳۰	N1 (60)
۰/۰۷	۰/۳۲	a_{max}
۵/۳	۸	M_v
۱	۲۹۰	R

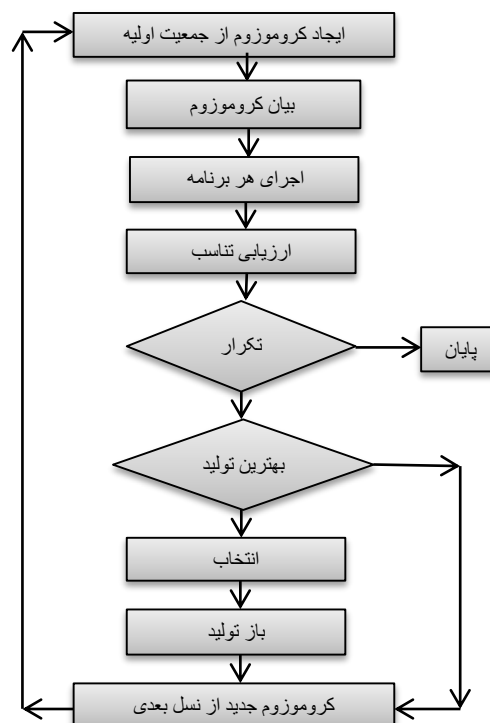
۳- نتایج و بحث

۳-۱- تحلیل خوشه‌بندی فازی

تحلیل خوشه‌بندی فازی بر روی تعدادی از داده‌های روانگرایی-های به وقوع پیوسته در اثر زلزله:

اطلاعات بدست آمده از روانگرایی‌های به وقوع پیوسته، با احتساب پارامترهای تأثیرگذار بر روانگرایی و روابط بین این پارامترها در مدل خوشه‌بندی فازی مورد تحلیل قرار گرفت.

در این بین سه گروه: سه، پنج و هفت خوشه‌ای مورد بررسی قرار گرفته است که نتایج به دست آمده نشان می‌دهد که، گروه‌های پنج و هفت خوشه‌ای به دلیل تفکیک موارد لغزش به خوشه‌های متعدد، موجب قرارگیری موارد مشابه و یکسان در خوشه‌های مجزا گردیده است و بنابراین تفکیک مناسبی برای پارامترها با بازه‌های مختلف ارائه نکرده است و باعث شده برخی از خوشه‌ها دارای تقسیم‌بندی مشابهی باشند و نتوان تفاوت‌چندانی بین برخی خوشه‌های مذکور اشاره کرد. در جدول (۲) می‌توان جای‌گیری موارد روانگرایی در گروه‌های سه خوشه‌ای را مشاهده نمود. بنابر این تنها گروه سه‌تایی دارای دقت کافی و مناسب در خوشه‌بندی می‌باشد. به طوری که در جدول (۲) نشان داده شده است، در گروه مذکور داده‌های روانگرایی با تفکیک مناسب در سه خوشه مورد نظر جای گرفته است. که در این سه خوشه داده‌ها به سه بازه مختلف به صورت تقریباً مجزا از هر بازه قرار گرفته‌اند.



شکل ۲- فلوچارت برنامه‌ریزی ژنتیک

در رویکرد بهینه‌سازی ما برنامه‌ریزی خطی و غیر خطی داریم که تابع هدف ما غیر خطی می‌باشد و در برنامه‌ریزی غیر خطی ما توابع محدود و غیر محدود داریم که تابع ما از جنس غیر خطی دارای محدودیت است. نحوه کار بدین صورت است که ما یک تابع هدف داریم که در این مدل برنامه‌ریزی غیر خطی مقید است و همچنین این تابع مبتنی بر یک سری محدودیت‌های خطی حل خواهند شد که برای حل آن از نرم‌افزار لینگو استفاده شده است. در پیوست ۱ کدنویسی صورت گرفته برای تحلیل حساسیت در نرم‌افزار لینگو نوشته شده است.

۳-۲- منطقه و داده‌های مطالعه موردی

در این مطالعه از نتایج ۹۲ آزمایش که در مناطق مختلف جهان ثبت شده مورد استفاده قرار گرفته است. این داده‌ها شامل ۴۷ داده روانگرا شده و ۴۵ داده روانگرا نشده می‌باشند. از جمله زلزله‌های به وقوع پیوسته که از داده‌های آنها استفاده شده است می‌توان به زلزله به وقوع پیوسته در سال ۱۹۶۴ در شهر نیگاتای ژاپن با بزرگای $M_w=7/5$ اشاره کرد، که به دلیل روانگرایی گسترده خاک در نواحی کم عمق باعث بروز خرابی‌های عمده در شهر نیگاتا شده است. مورد دیگر زلزله به وقوع پیوسته در جنوب کالیفرنیا در منطقه سن فرناندو که در سال ۱۹۷۱ با بزرگای $M_w=6/4$ اتفاق افتاده است و در اثر این رویداد انواع مختلفی از گسیختگی‌های زمین را منجر

تأثیرگذاری و حساسیت بالای این پارامتر در روانگرایی و یا عدم روانگرایی داده‌ها را نشان می‌دهد.

بیش از ۷۰٪ داده‌های موجود در دسته سوم روانگرا نشده‌اند. با بررسی داده‌های موجود در این خوشه پارامتری که بار دیگر تأثیرگذاری خود بر روی روانگرایی و یا روانگرا نشدن داده‌ها را نشان می‌دهد پارامتر عدد نفوذ استاندارد اصلاحی می‌باشد که همچون دسته اول تأثیر به سزایی بر روی این گروه از داده‌ها داشته است، ولی این پارامتر تنها پارامتر تأثیرگذار این دسته نمی‌باشد.

با بررسی دقیق‌تر و مقایسه داده‌ها، تأثیرگذاری عامل محرک به خوبی نمایان می‌شود، یعنی داده‌ها با مشخصات ژئوتکنیکی کاملاً مشابه ولی به دلیل اختلاف بزرگای زلزله، شتاب حداکثر و یا حتی فاصله تا کانون زلزله نیز که تا به اکنون زیاد مورد توجه نبوده است، نتایج روانگرایی متفاوتی ارائه کرده‌اند.

پارامتر دیگری که تأثیر مستقیمی بر روی داده‌ها را به خوبی نشان می‌دهد درصد ریزدانه موجود در خاک می‌باشد که تفاوت در میزان ریزدانه خاک عاملی برای روانگرایی و یا روانگرا نشدن خاک شده است.

با توجه به این تفاسیر اشاره شده بررسی دقیق‌تر تأثیرگذاری تمامی پارامترها که در روانگرایی خاک نقش دارند مهم و اساسی جلوه می‌کند، که سعی خواهد شد در بخش بعدی به این مسئله با دقت و ظرافت بیشتری رسیدگی شود.

۳-۲- مدل‌سازی تخمین زمانی روانگرایی با استفاده از برنامه‌ریزی ژنتیک

با استفاده از داده‌های جمع‌آوری شده از زلزله‌های به وقوع پیوسته در مناطق مختلف جهان که دارای تمامی پارامترهای مدنظر در تخمین زمانی روانگرایی خاک می‌باشد و با علم به روانگرایی و عدم روانگرایی خاک با خصوصیات مورد نظر؛ مدلی برای تخمین زمانی روانگرایی خاک توسط برنامه‌ریزی ژنتیک نموده‌ایم. منظور از زمان در مطالعه حاضر مدنظرگیری یک بازه زمانی نمی‌باشد و در اینجا صرفاً برای تعریف و نسبی کردن وقایع مختلف به صورت هم-زمان یا یک لحظه خاص استفاده شده است. به عنوان مثال منظور از تخمین روانگرایی در زمان (n) یعنی در نظرگیری امکان وقوع روانگرایی در لحظه n که چند شرایط خاص به صورت هم‌زمان در همان لحظه n مد نظر قرار گرفته‌اند. مدل GP توسعه داده شده در اینجا اساساً با هدف ایجاد عملکردهای ریاضی برای تخمین زمانی پتانسیل روانگرایی صورت گرفته می‌شود. برای توسعه این مدل از ۸ پارامتر ورودی عمق بحرانی، تنش، تنش مؤثر، در صد ریزدانه خاک، عدد SPT اصلاحی، شتاب حداکثر زلزله، بزرگای زلزله و

جدول ۲- ترکیب‌بندی خوشه‌ها

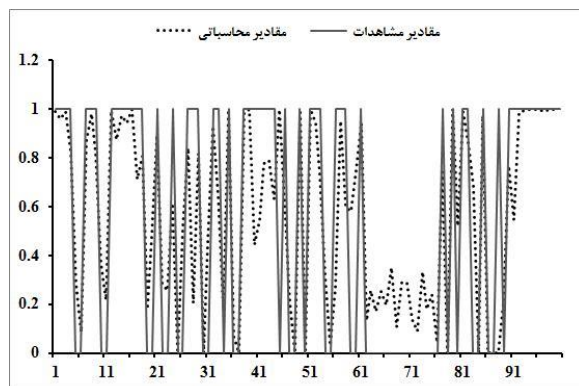
شماره خوشه‌ها	خوشه اول	خوشه دوم	خوشه سوم
داده‌های قرار گرفته در سه خوشه	D83,D77,D67,D62,D66,D69,D87,D86,...	D168,D98,D105,D112,D131,D38,D109,D89,D45,D164,D51,D149,D19,D49,D138,D162,D104,D13,D36,D113,...	D117,D19,D128,D15,D37,D129,D150,D24,D120,D52,D153,...

با توجه به مطالب ذکر شده فوق و تصمیم بر انتخاب گروه سه خوشه‌ای به منظور تفکیک اصولی داده‌های روانگرایی در دسته‌های متعدد، بررسی دقیق‌تر هر یک از خوشه‌های به دست آمده صورت می‌پذیرد.

جدول ۳- مشخصات دو داده از هر خوشه

	خوشه اول	خوشه دوم	خوشه سوم	خوشه اول	خوشه دوم	خوشه سوم
Depth	۶/۴	۶/۱	۴/۴۵	۱۲	۱۵	ناروانگر
σ	۱۲۰/۶	۱۳۲/۱	۸۶/۵۲	۲۳۳/۸	۱۷۹/۱	ناروانگر
σ'	۶۷/۰۳	۸۴/۲۶	۵۶/۶	۱۳۴/۹	۳۰۸/۶	ناروانگر
R	۱۴۰	۳۰	۳۰/۸۲	۳۴/۲۷	۳۳/۱۲	ناروانگر
M_v	۶/۷	۷/۴	۷/۹	۷/۶	۷/۶	ناروانگر
FC	۰	۰	۵	۲۲	۱۲	ناروانگر
a_{max}	۰/۱	۰/۲	۰/۳۲	۰/۷۹	۰/۱۹	ناروانگر
N1(60)	۱۲/۴	۱۱/۸	۱۱	۱۱	۲۰	ناروانگر

خوشه دوم از گروه سه خوشه‌ای دارای بیشترین تعداد داده روانگرایی می‌باشد. بیش از ۸۰٪ داده‌های موجود در این خوشه روانگرا شده‌اند. از مشخصات این خوشه می‌توان به عمق کم (۰/۹۱-۸/۲۳) داده‌های موجود در این خوشه و همین طور دارای تنش (۱۵۸-۱۵/۸) و تنش مؤثر (۱۰۸-۱۵/۸) کم می‌باشند. همین‌طور شتاب حداکثر این خوشه بازه بالایی داشته که حتی در برخی موارد که روانگرا شده‌اند تا ۱ نیز می‌رسد. فاصله محل‌های روانگرایی در این خوشه با کانون زلزله (۵۶-۱) کم می‌باشد (نزدیک به منبع انرژی لرزه‌ای قرار دارند). بنابر این در این خوشه اکثر پارامترهای تأثیرگذار نقش به سزایی در روانگرایی داده‌های این خوشه داشته‌اند. بیش از ۶۰٪ داده‌های موجود در خوشه اول روانگرا نشده‌اند. مهم‌ترین پارامتر این خوشه که تأثیر به سزایی در روانگرا نشدن یا روانگرایی داده‌های این خوشه مشخص است پارامتر $N_1(60)$ یعنی پارامتر عدد نفوذ استاندارد می‌باشد که به وضوح در مقایسه داده‌های روانگرا و غیر روانگرا تأثیر خود را نشان می‌دهد، به طوری که داده‌های با مشخصات کاملاً یکسان به دلیل داشتن $N_1(60)$ متفاوت روانگرایی و یا روانگرا نشدنشان مشخص شده است. همین مسئله



شکل ۳- نمودار مقایسه نتایج تخمین روانگرایی مدل برنامه ریزی ارائه شده GP با مشاهدات محلی

برای مقایسه هرچه بهتر و نشان دادن میزان مناسب درصد پیش‌بینی از مدل ارائه شده توسط GP، داده‌های آزمایش شده قبلی با استفاده از روش تجربی و مقبولی که اکثراً مورد توجه مهندسی و محققین می‌باشد (روش سید و ادیس)، تخمین زمانی روانگرایی انجام گرفته می‌شود. جواب‌های به دست آمده از این روش در مقایسه با واقعیت تنها برای ۸۰٪ داده‌ها پیش‌بینی درستی از خود نشان داده است. نتایج به دست آمده از مدل و روش تجربی با نتایج مدل Hanna و همکاران که در دسترس بود با استفاده از Fscore مقایسه صورت گرفته است. نتایج به دست آمده از فرمول زیر را در جدول (۵) ملاحظه می‌فرمایید.

$$F_{\beta} = \frac{(1 + \beta^2) \cdot tp}{(1 + \beta^2) \cdot tp + \beta^2 \cdot fn + fp}$$

tp: true positive
fn: false negative
fp: false positive

جدول ۵- مقایسه نتایج مدل‌ها توسط Fscore

	Fscore
نتایج به دست آمده از روش تجربی سید و ادیس	۰/۸۸
نتایج به دست آمده از مدل Hanna و همکاران	۰/۹۷۶
نتایج به دست آمده از مدل ارائه شده توسط GP	۰/۹۸۲

نتایج حاصل از تخمین زمانی انجام شده توسط مدل نشان می‌دهد که در بین ۱۷۴ داده موجود که توسط مدل به دست آمده مورد تخمین قرار گرفته است در مقایسه با داده‌های واقعی موجود تنها در ۴ مورد از داده‌های روانگرا شده و در ۷ مورد از داده‌های روانگرا نشده تخمین درستی توسط مدل صورت نگرفته است. در مقایسه، تخمین انجام شده توسط روش تجربی در ۱۰ مورد از داده‌های روانگرا شده و در ۲۵ مورد از داده‌های روانگرا نشده تخمین

فاصله کانونی محل تا زلزله استفاده می‌شود. به منظور رسیدن به مدل مناسب، برنامه به تعداد متعددی و با تنظیمات اولیه مختلفی اجرا می‌شود و عملکرد هر یک از مدل‌های توسعه یافته مورد تجزیه و تحلیل قرار گرفته و براساس معیارهای آماری و مقایسه تخمین‌های انجام شده با واقعیت بهترین مدل انتخاب می‌شود. پارامترهای لازم جهت تجزیه و تحلیل در این تحقیق که جزء گام‌های اولیه برنامه‌ریزی ژنتیک می‌باشند به صورتی که در جدول (۴) نشان داده شده است انتخاب می‌شود.

جدول ۴- مشخصات مدل ارائه شده

اندازه جمعیت (Population size)	۱۵۰۰
تولید (Generation)	۱۵۰۰
ارتباط عملکرد (Linking function)	+
مجموعه تابع (Function set)	+, -, ×, √, exp, log, ln
مقاطع (Crossover)	۰/۹
جهش (Mutation)	۰/۰۵
تکثیر (Reproduction)	۰/۲
جایگشت (Permutation)	۰/۰۰۵
ثابت‌های تصادفی (Random constants)	۴

در این رابطه Depth: عمق، σ : تنش، σ' : تنش مؤثر، FC(%): درصد ریزدانه خاک، $N_{1,60}$: عدد SPT اصلاحی، a_{max} : شتاب حداکثر سطح زمین، M_v : بزرگای زلزله، R(km): فاصله کانونی زلزله تا محل می‌باشند.

تابع تخمین زمانی (y) ایجاد شده به صورت زیر تولید شده است:

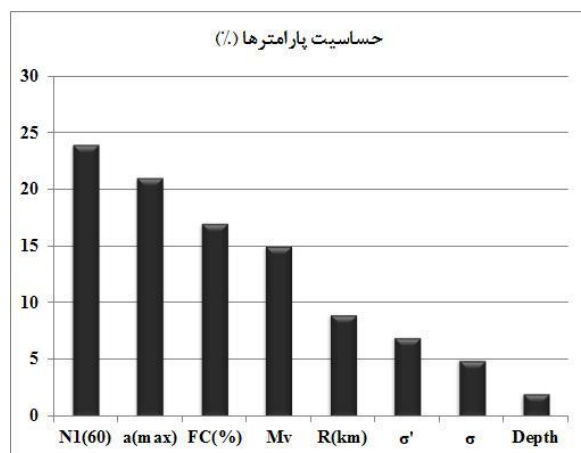
$$y = \left(\left(\left(8.2371 \times (N_1(60)) + (\sqrt{R \times M}) \right) \right) + ((3.4495 \times (a_{max})) \times (\log(\text{Critical Depth}))) + ((2.9473 \times (\ln(FC(\%))) \times (5.3279 \times (\ln \sigma \times \ln \sigma'))) + ((N_1(60) \times FC(\%)) + (R^2)) + ((3.4495 \times (a_{max})) \times (M)) \right) \quad (5)$$

جواب‌های به دست آمده که اعداد بین صفر و یک می‌باشد (که صفر برای داده‌هایی که روانگرا نشده‌اند و یک برای داده‌هایی است که روانگرایی در آنها صورت پذیرفته است) با استفاده از همین مدل ارائه شده به دست می‌آید، با نتیجه داده‌های واقعی از وقوع یا عدم وقوع روانگرایی در دست است مقایسه می‌شوند. ۹۴/۷۷٪ تخمین انجام شده توسط مدل درست بوده و جواب‌ها با واقعیت هم‌خوانی دارد، که می‌توان گفت درصد مناسبی را برای تخمین ارائه نموده است. نمودار ارائه شده در شکل (۳) (محور قائم آن جواب‌های به دست آمده از مدل ارائه شده ما بین ۰ و ۱ و محور افقی شماره داده‌ها می‌باشد) مقایسه تخمین بلند مدت نتایج حاصل از مدل برنامه‌ریزی GP با مشاهدات محلی را به وضوح نشان می‌دهد.

۳-۳- تحلیل حساسیت پارامترها

همان گونه که در تحقیقات گذشته نیز اثبات شده است، تأثیرگذاری بسیار بالای برخی پارامترها از جمله عدد SPT اصلاحی و حداکثر شتاب زلزله در روانگرا شدن خاک مشخص می باشد. با این وجود سعی بر آن شده است تا با استفاده از نتایج به دست آمده از روش های غیر خطی توسط لینگو که یک نرم افزار جامع طراحی شده برای ایجاد و حل کردن آسان تر و با کیفیت تر مدل های بهینه سازی ریاضی می باشد، تحلیل حساسیت بر روی مدل ارائه شده توسط برنامه ریزی ژنتیک، درصد حساسیت داخلی پارامترها در نمودار میله ای نشان داده شود. نتایج به دست آمده از تحلیل حساسیت انجام شده به صورت زیر می باشد:

همان گونه که در شکل (۴) ملاحظه می شود دو پارامتر عدد SPT اصلاحی و حداکثر شتاب زلزله بیشترین تأثیر و پارامتر عمق خاک کمترین تأثیر را در روانگرایی خاک در بین پارامترهای موجود داشته اند. این بدان معنی است که تأثیر مقدار پارامترهایی که حساسیت بیشتری بر روی مدل داشته اند، در روانگرا شدن و یا عدم روانگرایی خاک ها نیز به همین منوال است. می توان گفت تأثیری که کم یا زیاد شدن به عنوان مثال پارامتر عدد SPT اصلاحی بر روی روانگرا شدن خاک می گذارد، پارامتر عمق خاک مورد بررسی و یا تنش خاک نمی تواند تأثیرگذار باشد. می توان به این مسئله نیز اشاره کرد که در ارائه روش های پیش گیری و مقابله با روانگرایی بیشتر باید بر روی پارامترهای عدد SPT اصلاحی، حداکثر شتاب زلزله، درصد ریزدانه خاک، بزرگای زلزله و فاصله از کانون زلزله توجه و تمرکز نمود و برای جلوگیری از پدیده روانگرایی، اصلاحات و تدابیر لازم بر روی این پارامترهای مؤثر انجام داد.



شکل ۴- مطالعه حساسیت پارامترها

درستی انجام نشده است. با توجه به مقایسه صورت گرفته در مورد داده های به دست آمده از هر دو روش به نظر می رسد مدل ارائه شده در تحقیق حاضر عملکرد بهتری نسبت به نتایج تخمین زمانی روش تجربی سید و ادريس دارد. البته لازم به ذکر است شرایط فیزیکی هر دو تحقیق متفاوت بوده که نتایج بر حسب آن گزارش شده است.

در ادامه به آنالیز حساسیت پارامترهای مدل پرداختیم که نتایج به صورتی که در زیر آمده است در مورد هر یک از پارامترهای دخیل ارائه می شود.

یکی از پارامترهایی که تأثیرگذاری بالای آن اثبات شد عدد SPT اصلاحی بود. در ابتدا بر روی تأثیر این پارامتر بر روی مدل ارائه شده پرداختیم. با ایجاد تغییر این پارامتر باز هم تأثیرگذاری آن بر روی مدل کاملاً مشخص می شود. داده های روانگرا شده به دلیل کم بودن این پارامتر با تغییر اندک و افزایش مقدار این پارامتر از حالت روانگرایی در آمده و ناروانگرا می شوند و بالعکس داده های روانگرا نشده به دلیل بالا بودن این پارامتر با جایگزینی مقدار کمتری روانگرا می شوند. این مسئله نشان دهنده این مطلب هم می باشد که برای جلوگیری از روانگرایی خاک های دارای پتانسیل بالا می توان با ایجاد تغییر همین خصوصیت ژئوتکنیکی خاک از پتانسیل روانگرایی آن کاسته شود.

پارامتر دیگر بزرگای زلزله می باشد که باز هم تأثیر بالایی بر روی روانگرایی خاک دارد. در داده های روانگرا شده، کم کردن بزرگای زلزله باعث عدم روانگرایی خاک می شود و همین طور برای روانگرا نشدن خاک ها کاهش این پارامتر، تأثیر خود را کاملاً نشان می دهد. دیگر پارامتر عامل محرک حداکثر شتاب زلزله می باشد که تأثیرگذاری همچون پارامتر بزرگای زلزله را نشان می دهد.

پارامتر دیگر که با تغییر آن تأثیرگذاری خود را بر نتایج مدل نشان می دهد درصد ریزدانه خاک می باشد. نتایج مطالعات پیشین نشان می دهد که میزان ریزدانه خاک هایی که دارای پتانسیل روانگرایی بالایی می باشند از یک حد نصاب خاصی تبعیت می کنند ولی نتایج بررسی های ما گویای این مطلب است که این پارامتر تأثیرگذاری ثابتی بر روی نتایج مدل نشان نمی دهد، یعنی ممکن است افزایش این پارامتر باعث روانگرایی خاک شود و در شرایطی دیگر از روانگرایی خاک جلوگیری کند.

دیگر پارامترهای دخیل در روانگرایی تا یک بازه خاصی افزایش و یا کاهش آنها در نتایج مدل تأثیرگذار نبوده اند ولی در صورتی که مقدار این پارامترها از یک میزان خاصی بالاتر یا پایین تر می رود تأثیرگذاری خود را در نتایج نشان می دهند.

۴- نتیجه‌گیری و پیشنهادات

با توجه به داده‌های در دسترس و پارامترهای دخیل انتخاب شده تقسیم‌بندی داده‌ها به سه خوشه مجزا بهترین تقسیم‌بندی را با توجه به جداسازی داده‌ها و عدم قرار گرفتن داده‌ها در خوشه‌های تکراری و با ویژگی‌های نزدیک به هم را به ما ارائه کرد. که این تقسیم‌بندی امکان بررسی بهتر را به ما داد. در این تقسیم‌بندی داده‌ها، اندازه تمامی پارامترها به خوبی به سه بازه جداگانه تقسیم‌بندی شده است. با توجه به جمع‌بندی و تفسیر هر یک از سه خوشه ارائه شده در بین پارامترهای ژئوتکنیکی، عدد SPT اصلاحی ($N_1(60)$) تأثیرگذاری به مراتب بیشتری نسبت به دیگر پارامترها را نشان می‌دهد. دیگر پارامتر تأثیرگذار عامل محرک یعنی بزرگای زلزله، شتاب حداکثر زلزله و پارامتر فاصله تا کانون زلزله می‌باشد. که در اینجا اثبات تأثیرگذاری فاصله تا کانون و نقش آن در تعیین روانگرایی خاک می‌تواند عامل مهم و قابل ذکر باشد.

نتایج به دست آمده از تحلیل حساسیت داده‌ها نشان می‌دهد دو پارامتر عدد SPT اصلاحی و حداکثر شتاب زلزله بیشترین تأثیر و پارامتر عمق بحرانی خاک کمترین تأثیر را در روانگرایی خاک در بین پارامترهای موجود را دارند.

در مجموع مطالعه انجام شده، روش GP را به عنوان یک روش دقیق و عملی برای مدل‌سازی رابطه تخمین زمانی روانگرایی خاک-ها معرفی می‌نماید. دقت بالای نتایج حاصل از این روش از یک سو (۹۴/۷۷٪) تخمین انجام شده توسط مدل درست بوده و جواب‌ها با واقعیت هم‌خوانی دارد) و سادگی استفاده از این روش و عدم نیاز به دانش بالای مبانی ریاضی الگوریتم و برنامه‌ریزی ژنتیک از سوی دیگر مؤید قابلیت این روش به عنوان یک ابزار مفید و سریع در پیش‌بینی روانگرایی خاک‌ها می‌باشد. اگرچه با علم بر مبانی ریاضی الگوریتم ژنتیک و روش‌های مدل برنامه‌ریزی ریاضی غیر خطی پیشرفته (برنامه‌ریزی محدب، الگوریتم روزن و...) می‌توان معادله صریح حاکم بر پیش‌بینی روانگرایی خاک‌ها را نیز ارائه کرد. استخراج این رابطه و مقایسه دقت آن با نتایج حاصل از روش‌های بهینه‌سازی پیشرفته‌تر می‌تواند به عنوان موضوع تحقیق دیگری مطرح شود.

در ادامه تحلیل حساسیت انجام شده بر روی مدل تأثیرگذاری بالای پارامترهای عدد SPT اصلاحی و حداکثر شتاب زلزله را نشان می‌دهد. که این کار واند در ادامه به ارائه روش‌های پیشگیری از وقوع روانگرایی خاک با توجه هرچه بیشتر به این دو پارامتر مهم کمک کند.

۵- مراجع

- Bezdek J, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York, 1981.
- Bezdek J, Hathaway R, "Recent convergence results for the fuzzy c-means clustering algorithms", Journal of Classification, 1988, 5 (2), 237-247.
- Castro G et al, "Liquefaction induced by cyclic loading", Winchester, Mass: Geotechnical Engineers, 1982.
- Cetin KO, et al, "Standard penetration test-based probabilistic and deterministic assessment of seismic soil liquefaction potential", Geotechnical and Geoenvironmental Engineering, Vol, 2004, 130, No 12.
- Chang Muhsung et al, "Comparison of SPT-N-based analysis methods in evaluation of liquefaction potential during the 1999 Chi-chi earthquake in Taiwan", Computers and Geotechnics, 2011, 38, 393-406.
- Chen YR, SC Hsieh, Chen JW, Shih C, "Energy-based probabilistic evaluation of soil liquefaction", Soil Dynamics and Earthquake Engineering, 2005, 25 55-68.
- Chen, O-tani H, Hori M, "Stability analysis of soil liquefaction using a finite element method based on particle discretization scheme", Computers and Geotechnics, 2015, 67 (2015) 64-72
- Chiu S, "Fuzzy model identification based on cluster estimation", J. Intel. Fuzzy System, 1994, 2, 267-278.
- Dunn J, "A fuzzy relative of the Isodata process and its use in detecting compact, well-separated clusters", Journal of Cybernetics, 1973, 3 (3), pp. 32-57.
- Gandomi AH, Alavi AH, "Multi-stage genetic programming: a new strategy to nonlinear system modeling", Information Sciences, 2011, 181 (23), 5227-5239.
- Green R, et al, "An Energy-based excess pore pressure generation model for cohesionless soils". Proceedings of the John Booker Memorial Symposium Sydney, New South Wales, Australia, November, 2000, 16-17.
- Guoxing Chen, Lingyu Xu, Mengyun Kong, Xiaojun Li, "Calibration of a CRRmodel based on an expanded SPT-based database for assessing soil liquefaction potential", Engineering Geology, 2015 196, 305-312.
- Hadi Shahir, Ali Pak, Mahdi Taiebat, Boris Jeremic, "Evaluation of variation of permeability in liquefiable soil under earthquake loading", Computers and Geotechnics, 2012, 40, 74-88.
- Hanna Adel M, et al, "Neural Network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data", Soil Dynamics and Earthquake Engineering, 2006, 27, 521-540.
- Holland JH, "Adaption in Natural and Artificial Systems" University of Michigan Press, Ann Arbor, 1975, 228pp.
- Ishihara K, "Soil behavior in earthquake geotechnics". New York: Oxford University Press, 1996.
- Jafarian Y, Kermani E, Baziar MH, "Empirical predictive model for the V_{max}/a_{max} ratio of strong ground

- Seed HB, Idriss IM, "Simplified Procedure for Evaluating Soil Liquefaction Potential", Journal of the Soil Mechanics and Foundations Division, 1971, 97 (SM9), pp1249-1273.
- Seed HB, Idriss IM, Arrango, "Evaluation of liquefaction potential using field data", Journal of Geotechnical Engineering, ASCE, 1983, 109, 458-484.
- Shahin MA, Maier HR, Jaksa MB, "Settlement prediction of shallow foundations on granular soils using B-spline neuro fuzzy models". Computers and Geotechnics, 2003, 30, 637-647.
- Torres RS, Falcao, Gonc alves AX, Papa MA, Zhang JP, Fan W, "A genetic programming framework for content-based image retrieval", Pattern Recognition, 2009, 42 (2), 283e292.
- Tütmez B, Tercan AE, "Spatial estimation of some mechanical properties of rocks by fuzzy modeling". Computers and Geotechnics, 2007, 34, 10-18.
- Whitman RV, "Resistance of Soil to Liquefaction and Settlement", Soils and Foundations, 1971, 11 (4), pp59-68.
- Yager R, Filev D, "Generation of fuzzy rules by mountain clustering", J. Intel. Fuzzy Systems, 1994, 2, 209-219.
- Youd TL, Idriss IM, Andrus RD, Arango I, Castro G, Christian J T, Dobry R, Finn W D L, Harder L F, Hynes M E, Ishihara K, Koester J P, Liao S C, Marcuson W F, Martin GR, Mitchell JK, Moriwaki Y, Power MS, Robertson PK, Seed RB, Stokoe KH, "Liquefaction Resistance of Soils: Summary Report from the 1996 NCEER and 1998 NCEER/NSF Workshops on Evaluation of Liquefaction Resistance of Soils", Journal of Geotechnical and Geoenvironmental Engineering, 2011, 127 (4), pp297-313.
- Zeghal Mourad, Elgamal Ahmet-W, "Analysis of site liquefaction using earthquake records", Geotechnical Engineering, Vol, 1994, 120, No.6046.
- Zhang CH, Juang JR, Martin HW, Huang, "Inter-region variability of Robertson and Wride method for liquefaction hazard analysis", Engineering Geology, 2016, 203 (2016) 191-203.
- motions using genetic programming", Computers and Geosciences, 2011a, 36 (12), 1523-1531.
- Javadi A, Rezanian M, Mousavinezhad M, "Evaluation of liquefaction induced lateral displacements using genetic programming", Computers and Geotechnics, 2016, 33, 222-233.
- J. Geotechnical Engineering 1985, 10.1061/ (ASCE), 733-9410 (1985) 111:12 (1425), 1425-1445.
- Kayadelen C, "Estimation of effective stress parameter of unsaturated soils by using artificial neural networks". International Journal for Numerical and Analytical Methods in Geomechanics, 2008, 32, 1087-1106.
- Kayadelen C, Günaydin O, Fener M, Demir A, Özvan A, "Modeling of the angle of shearing resistance of soils using soft computing systems". Expert Systems with Applications, 2009, 36, 11814-11826.
- Koza JR, "Genetic Programming", www.genetic-programming.com, The home page of John R. Koza at Genetic Programming, 2008.
- Koza JR, Poli R, "Genetic programming. In: Burke E, Kendall G. (Eds.), Search Methodologies; Introductory Tutorialsin Optimization and Decision Support Techniques", Springer Science +Business Media, pp. 2005, 127-164.
- Koza JR, "Genetic Programming: On the Programming of Computers by Means of Natural Selection". MIT Press, Cambridge, MA, 1992, 813pp.
- Liong SY, Gautam T R, Khu ST, Babovic V, Keijzer M, Muttill N, "Genetic programming. A new paradigm in rainfall runoff modeling", J. Am. Water Response Association, 2003, 38 (3), 705-718.
- Muduli Pradyut Kumar, Das Sarat Kumar, "Model uncertainty of SPT-based method for evaluation of seismic soil liquefaction potential using multi-gene genetic programming", Soils and Foundations, 2015, 2015;55 (2): 258-275.
- NCEER, "Proceeding of the NCEER Workshop on Evaluation of Liquefaction Resistance of Soils, (T.L. Youd and Idriss I M, eds.), Technical Report NCEER-97-0022, National Center for Earthquake Engineering Research", State University of New York at Buffalo, 1997, 276pp.
- Olsen RS, Koester JP, "Prediction of liquefaction resistance using the CPT", In International symposium on cone penetration testing, CPT 95 Linkoping, Sweden, 1995, pp. 251-256.
- Rezanian M, Javadi A, "An ewgenetic programming model for predicting settlements of shallow foundations", Canadian Geotechnical Journal, 2017, 44(12), 1462-1473.
- Ruspini E, "A new approach to clustering", Information and Control, 1969, 15, pp. 22-32.
- Samui Pijush, R Hariharan, "A unified classification model for modeling of seismic liquefaction potential of soil based on CPT", Journal of Advanced Research, 2004.
- Seed B, Tokimatsu H, Harder KL, Chung R, "Influence of SPT Procedures in Soil Liquefaction Resistance Evaluations".

EXTENDED ABSTRACT

The Study of the Liquefaction Probability and Estimation of the Relative Importance of Effective Parameters Using Fuzzy Clustering and Genetic Programming

Armin Sahebkar Alamdari ^a, Amir Najafi ^{b,*}

^a Department of Civil Engineering, Islamic Azad University, Tabriz branch, Tabriz

^b Department of Industrial Engineering, Islamic Azad University, Zanzan branch, Zanzan

Received: 16 May 2016; Accepted: 08 October 2016

Keywords:

Probability, genetic programming, fuzzy clustering, soil liquefaction

1. Introduction

During several past decades, many researchers have conducted studies to develop the relationship between liquefaction resistance and seismic parameters of various soils. One of the methods to assess liquefaction which is widely used is the stress-based method, which was presented for the first time by Seed and Idriss (1971) and Whitman (1971). That is an experimental method and is based on experimental and local observations.

Anyway, the relationships developed by the use of the traditional method based on experimental data are limited and don't provide a proper and stable prediction, they also don't show major weaknesses in mathematical relations, complete structure and heterogeneous space of soil and liquefaction, while soils have quite complicated structures, there is simpler methods of examination that allow unlimited development of a model for all systems, such as genetic programming (GP), artificial neural networks (ANNS) and Neuro-fuzzy inference system (ANFIS) (Shahin et al., 2003). In this study, using fuzzy clustering, it has been initially tried to divide data to groups whose members are similar and have relatively same properties. So data with similar geotechnical and seismic properties can be placed in a cluster and hence, a closer study of the data to better understand the factors affecting soil liquefaction can be possible. Next, a model was presented for time estimation of the soil liquefaction probability using the method of genetic programming. Finally, considering the results of the conducted clustering and the study of data in each cluster, the sensitivity of parameters affecting soil liquefaction and their relative importance were analyzed using nonlinear methods.

2. Methodology

2.1. Clustering

Clustering is one of the branches of learning without supervision and it is an automated process during which the data is divided into groups whose members are similar and have the same properties. Each of these specific categories or groups is called a cluster. Then, a cluster is a collection of data and information where the members of the group have similar properties and are dissimilar with the samples in other clusters.

2.1.1. Fuzzy C-Mean clustering algorithm

The method of C-Mean is one of the most popular methods of fuzzy clustering algorithm and information exploration which was represented initially by Duun (1973) and then using the theories, it was turned into its

* Corresponding Author

E-mail addresses: armin.sahebkar@gmail.com (Armin Saheb karam Alamdari), asdnjf@gmail.com (Amir Najafi).

present state (Dunn, 1973, Bezdek, 1981. Bezdek and Hathaway, 1988). The main concept of the cluster has been defined by Ruspini (1969) in some relations (Bezdek, 1981).

2. 2. Genetic Programming

Genetic programming (GP) is an educational structure inspired by biology to analyze statistical data and recognition pattern. GP has been used very successfully to solve the various kinds of complicated issues of Geotechnical engineering. Some of them include: prediction of liquefaction which results in horizontal displacement, settlement of deep foundations and the V_{max} / a_{max} ratio for strong ground motions (Javadi et al., 2006, Rezaia and Javadi, 2007. Jafarian et al., 2010). Koza (1992) introduced this method as an effective method for the first time, which is a branch of typical genetic algorithms (GA), developed by Holland (1975) (Holland, 1975. Gandomi and Alavi, 2011).

2. 3. Sensitivity analysis with Lingo

One of the issues that are considered in optimization problems is the analysis of sensitivity of target function or output at the optimal point with regard to the input parameters. Since the input parameters are considered in the same condition and are modeled without any presumption concerning a parameter, then the sensitivity analysis must be done with respect to all parameters, which is conducted by Lingo software (it has been developed in the section of the range which can be done easily).

3. Results and Discussion

3. 1. Fuzzy clustering analysis

Information obtained from the occurred liquefaction was analyzed considering the parameters which affect the liquefaction and the relationships between these parameters in the fuzzy clustering model. Among them, three groups with three, five and seven clusters have been studied, which the obtained results show that the groups with five and seven clusters have led to the placement of similar and same cases in separate clusters due to the separation of slipping cases into multiple clusters, and therefore they have not provided an appropriate separation for parameters with different intervals and have caused that some clusters have similar classifications.

So, only the group with three clusters involved enough accuracy and was suitable for clustering. As shown in table (2), in this group, liquefaction data has been placed with appropriate separation in the three given clusters. In these three clusters, the data has almost separately been placed in three different intervals.

3. 2. Modeling of the time estimation of liquefaction using genetic programming

Using data obtained from earthquakes that have occurred in different regions of the world, which involve all parameters needed for the time estimation of soil liquefaction and by knowing soil liquefaction and non-liquefaction with given properties, we have presented a model for the time estimation of soil liquefaction by the genetic programming.

This time estimation function (y) has been developed as follows:

$$y = (((8.2371 \times (N_{-1} (60)) + (\sqrt{(R \times M)}))) + ((3.4495 \times (a_{max})) \times (\log(\text{Critical Depth}))) + ((2.9473 \times ((\ln(FC(\%)))) \times (5.3279 \times (\ln \sigma \times \ln \sigma^{' }))) + (((N1 (60) \times FC(\%)) + (R^2)) + ((3.4495 \times (a_{max})) \times (M))))$$

The answers that are the numbers between zero and one (zero for the data in which liquefaction has not occurred, 1 for the data in which liquefaction has occurred) are obtained by the presented model and are compared with the real data which are available about the occurrence or non-occurrence of liquefaction. 94.77% of the estimation conducted by the model were correct and the answers were consistent with the reality, which can be said that it has provided an appropriate percentage for estimation.

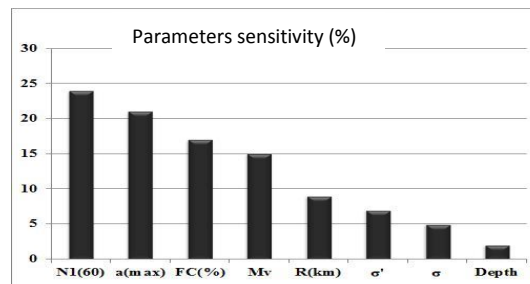
Time estimation of liquefaction has been done to compare better and to show the appropriate amount of the prediction percentage of the model provided by GP and the previous data tested by experimental method which is accepted by all engineers and researchers (the method of Seed and Idriss, 1971). The answers obtained from this method in comparison with reality have shown proper prediction for only 80% of the data. Using Fscore, the results of the experimental method have been compared with the model of Hana et al. (2006) which was available.

Table 1. Comparison of the models' results using Fscore

	Fscore
The results obtained from the experimental method of Seed and Idriss	0.88
The results obtained from the model of Hana et al	0.976
The results obtained from the model provided by GP	0.982

3.3. Analysis of the Parameters Sensitivity

Using the results obtained from the nonlinear method by Lingo, which is a comprehensive software designed for easier and high quality development and solving the mathematical optimization models, it has been tried to analyze sensitivity on the model presented by genetic programming and the percentage of internal sensitivity of parameters has been displayed in the following bar graph.

**Fig 1.** The study of parameters' sensitivity

4. Conclusion and Recommendations

According to available data and selected, involved parameters, dividing the data into three separate clusters provided the best dividing with respect to the data separation and not placing data in repeated clusters with similar properties. Totally, the present study introduces the GP method as an accurate and practical method for modeling the time estimation of soil liquefaction. High accuracy of the results obtained from this method on the one hand (94.77% of the estimation conducted by the method has been proper and the answers are consistent with reality) and the simplicity of the use of the method and no need to high knowledge of the basic mathematical principles of algorithms and genetic programming on the other hand approves the ability of the method as an effective and quick instrument for predicting soil liquefaction.

Next, sensitivity analysis conducted on the model shows the high effectiveness of the parameters of the corrected SPT number and maximum acceleration of earthquake, which can help to offer methods for prevention of soil liquefaction by more attention to these two important parameters.

5. References

- Dunn J, "A fuzzy relative of the Isodata process and its use in detecting compact, well-separated clusters", *Journal of Cybernetics*, 1973, 3 (3), pp. 32-57.
- Gandomi AH, Alavi AH, "Multi-stage genetic programming: a new strategy to nonlinear system modeling", *Information Sciences*, 2011, 181 (23), 5227-5239.
- Hanna Adel M et al, "Neural Network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data", *Soil Dynamics and Earthquake Engineering*, 2006, 27, 521-540.
- Holland JH, "Adaption in Natural and Artificial Systems", University of Michigan Press, Ann Arbor, 1975, 228pp.
- Jafarian Y, Kermani E, Baziar MH, "Empirical predictive model for the V_{max}/a_{max} ratio of strong ground motions using genetic programming", *Computers and Geosciences*, 2011a, 36 (12), 1523-1531.
- Koza JR, "Genetic Programming: On the Programming of Computers by Means of Natural Selection". MIT Press, Cambridge, MA, 1992, 813pp.
- Rezania M, Javadi A, "An ewgenetic programming model for predicting settlements of shallow foundations", *Canadian Geotechnical Journal*, 2017, 44 (12), 1462-1473.
- Seed HB, Idriss IM, "Simplified Procedure for Evaluating Soil Liquefaction Potential", *Journal of the Soil Mechanics and Foundations Division*, 1971, 97(SM9), pp1249-1273.
- Shahin MA, Maier HR, Jaksa MB, "Settlement prediction of shallow foundations on granular soils using B-spline neuro fuzzy models", *Computers and Geotechnics*, 2003, 30, 637-647.
- Tütmez B, Tercan AE, "Spatial estimation of some mechanical properties of rocks by fuzzy modeling". *Computers and Geotechnics*, 2007, 34, 10-18.
- Whitman RV, "Resistance of Soil to Liquefaction and Settlement", *Soils and Foundations*, 1971, 11 (4), pp 59-68.