

مطالعه پدیده‌های طبیعی با استفاده از روش تجزیه و تحلیل

مؤلفه‌های اصلی^۱

سید کاظم علوی پناه^۲

چکیده

در این تحقیق از روش تجزیه و تحلیل مؤلفه‌های اصلی (PCA)^۳ برای کاهش تعداد باندهای طیفی و متغیرهای مکانی از طریق پیدا کردن ترکیب خطی متغیرهایی که بیشترین تغییرات را دارند استفاده شد. به منظور به کارگیری PCA، از داده‌های TM ماهواره لندست مربوط به ۵ منطقه مختلف مرکزی ایران و همچنین ۱۸ متغیر خاک استفاده گردید. نتایج PCA مربوط به باندهای TM، اهمیت مؤلفه اول (PC1) برای مطالعات خاک، PC2 و گاهی PC3 در مطالعات پوشش گیاهی را نشان داد. نتایج PCA نشان داد که ۷ باند TM و ۱۸ متغیر خاک عمدتاً در سه مؤلفه اصلی به ترتیب بیش از ۹۰ و ۵۵ درصد از اطلاعات را به خود اختصاص می‌دهند. براساس نتایج حاصل از این تحقیق، روش PCA برای اطلاعات طیفی و مکانی مختلف و کمک به طراحی نمونه‌برداری، به منظور صرفه‌جویی در وقت و هزینه می‌تواند به کار رود.

واژه‌های کلیدی: مؤلفه‌های اصلی، باندهای طیفی، متغیرهای خاک و داده‌های TM

۱- تاریخ دریافت: ۷۸/۱۲/۳، تاریخ تصویب نهایی: ۸۰/۲/۳۱

۲- استادیار مرکز تحقیقات مناطق کویری و بیابانی ایران، دانشگاه تهران

مقدمه

الف- کاربرد PCA در مطالعات مربوط به تصاویر

ماهواره‌ای

مهمترین فواید این روش، جمع‌آوری و مترکم ساختن اطلاعات پدیده‌های موجود در باندهای مختلف در تعدادی تصاویر یا مولفه‌های کمتر است، به عبارت دیگر، PCA برای حذف اطلاعات اضافی در داده‌های تصاویر ماهواره‌ای کاربرد فراوانی دارد. علاوه بر این، از مولفه‌های اصلی حاصل از تصاویر TM ماهواره لندست برای مطالعات دیگری مانند متمایز کردن خاک‌های گچی، آهکی و شنی استفاده شده است (۱).

ب- کاربرد PCA در مطالعات زمینی

این روش بر این اساس استوار است که یک یا چند متغیر از کل متغیرهایی که همبستگی قوی دارند، لحاظ می‌شود، به عبارت دیگر، در تجزیه و تحلیل مولفه‌های اصلی تعداد اندکی از مولفه‌های اصلی (PC) بیانگر بخش مهم و زیادی از کل اطلاعات می‌باشند. در روش PCA، به منظور اجتناب از اثر مقیاس‌های گوناگون متغیرها، می‌توان مقیاس‌ها را استاندارد کرد که در این صورت تمامی انحراف از معیارهای متغیرها برابر با یک خواهد شد.

روش PCA می‌تواند مبتنی بر ماتریس همبستگی (روش استاندارد) یا ماتریس واریانس - کو واریانس (غیراستاندارد) باشد. از روش PCA مبتنی بر ماتریس همبستگی، استفاده‌های متعددی برای تحلیل داده‌های خاک شده است (۱۴). در هنگام تجزیه و تحلیل مولفه‌های اصلی، متغیرهایی که دارای همبستگی خیلی بالایی باشند، بعد از انجام PC، محورها به گونه‌ای تغییر جهت داده می‌شود که حداکثر واریانس را به خود اختصاص دهند، برای مثال به نحوی که PC1 در جهت بیشترین واریانس و PC2 عمود بر PC1

جمع‌آوری اطلاعات و داده‌های آزمایشگاهی و صحرایی، معمولاً مستلزم صرف وقت و هزینه زیادی است، از طرفی تعدد داده‌های مربوط به تجزیه کامل نمونه‌هایی مانند نمونه‌های خاک، آب و گیاه، به پیچیدگی و صرف هزینه و وقت زیاد در امر تفسیر منجر می‌شود. اصولاً داده‌های ماهواره‌ای حاصل از باندهای طیفی گوناگون نظیر باندهای آبی، سبز، قرمز و مادون قرمز که حاوی اطلاعات مفیدی راجع به پدیده‌های زمینی هستند، از همبستگی بسیار بالایی برخوردارند. بنابراین مطالعاتی که بر داده‌های ماهواره‌ای و همچنین داده‌های زمینی نظیر خاک استوارند حذف برخی متغیرهایی که رابطه قوی دارند و لحاظ کردن یک یا چند متغیر از مجموع کل متغیرها به نحوی که بیانگر ویژگی‌های عمده و اصلی دیگر متغیرها باشد، می‌تواند صرفه‌جویی در وقت و هزینه را موجب شود (۱۰ و ۱۲). تجزیه و تحلیل مولفه‌های اصلی (PCA) روشی است که برای کاهش تعداد متغیرهای یک مجموعه داده از طریق پیدا کردن ترکیب خطی متغیرهایی که بیشترین تغییرات را دارند، استفاده می‌شود. نتایج برخی مطالعات نشان می‌دهد که PCA در طراحی نمونه‌برداری خاک کاربرد فراوانی دارد (۱۴). اگرچه دقت طبقه‌بندی داده‌های ماهواره‌ای و زمینی بستگی زیادی به نحوه نمونه‌برداری صحرایی دارد، ولی تاکنون سعی زیادی برای ارائه و طراحی نحوه نمونه‌برداری به عمل نیامده است (۲، ۷، ۸ و ۹). در این تحقیق، دو نوع کاربرد PCA در مناطق مختلف خشک و نیمه‌خشک ایران مورد بررسی قرار گرفت. دو نوع کاربرد مذکور به شرح زیر می‌باشند:

پلایای ابرکوه

این ناحیه از خاک‌های شور و غیرشور با بافت‌های متوسط تا سنگین و برخی گونه‌های خشکی‌پسند و شورپسند تشکیل شده است.

موک استان فارس

این منطقه کوهستانی از گونه‌های جنگلی و مراتع وسیع و همچنین اراضی کشاورزی تشکیل شده است.

روش انجام تحقیق

به‌منظور بررسی کاربرد PCA در مطالعات مربوط به تصاویر ماهواره‌ای و مطالعات زمینی (خاک‌شناسی)، اقدامات زیر انجام شد:

۱- تجزیه و تحلیل داده‌های ماهواره‌ای

در این تحقیق، از تمامی ۷ باند مناطق پنج‌گانه مورد مطالعه به ابعاد 1200×1200 پیکسل استفاده شد. تاریخ برداشت داده‌های ماهواره‌ای مورد استفاده، مربوط به فصول تابستان و پاییز سال‌های ۱۹۹۰ یا ۱۹۹۱ است. در این مطالعه، ابتدا کیفیت باندها بررسی و داده‌های نادرست از محاسبات حذف شد. به‌منظور بررسی و مقایسه میزان اطلاعات موجود در تمامی هفت باند اصلی و سپس مولفه‌های اصلی (PC)، مراحل زیر انجام شد:

الف- PCهای هفت‌گانه مبتنی بر هفت باند

به روش ماتریس همبستگی ایجاد شده و سهم هر کدام از PCها از کل تغییرات تحت عنوان درصد واریانس (یا مقدار ویژه Eigenvalue) مورد مقایسه قرار گرفت.

ب- به‌دلیل نقش باند حرارتی (TM6) در

تفسیر خواص حرارتی مواد، ابتدا تمامی هفت باند برای استخراج ۷ مولفه اصلی استفاده شد و سپس با حذف TM6، میزان اطلاعات PCهای شش‌گانه مورد بررسی قرار گرفت.

باشد. در امتداد PCهای مراتب پایین‌تر، میزان واریانس (اطلاعات) کمتری وجود دارد، به‌عبارت دیگر، PC1، PC2 و... PCn به ترتیب بیشترین واریانس را دارا می‌باشند. از محاسن روش PCA، آن است که این روش محدود به جمعیت‌های با پراکنش نرمال نیست.

مواد و روش‌ها

در این تحقیق، به‌منظور ارائه کارایی روش تجزیه و تحلیل مولفه‌های اصلی در تفسیر پدیده‌های مناطق خشک و نیمه‌خشک ایران، ۵ منطقه (چهار منطقه بیابانی و یک منطقه غیربیابانی) واقع در استان‌های قم، یزد و فارس در نظر گرفته شدند. ویژگی‌های مهم مناطق مورد مطالعه به شرح زیر است:

پلایای قم (حوض سلطان)

این منطقه دارای پوسته‌های نمکی پوشیده از املاح به‌ویژه کلرورسدیم و دیگر رخساره‌ها نظیر رخساره‌های شور و مرطوب، رخساره‌های پف‌کرده با اشکال چندضلعی، رخساره‌های پوشیده از سنگریزه و رخساره‌های دارای پوشش گیاهی است که عمدتاً از گونه‌های مقاوم به شوری هستند.

پلایای سیاه کوه

این منطقه شامل دو بخش بالایی و پایینی است که حاشیه بالایی متعلق به چاه افضل با شرایط فوق‌العاده خشک است و رخساره‌های پف‌کرده، مرطوب و چسبناک، رخساره با سطح صاف و اشکال چندضلعی و همچنین سطوح نمکی از مهمترین پدیده‌های منطقه است. رخساره‌های حاشیه میانی علاوه بر رخساره‌های حاشیه بالایی، از رخساره‌های شدت فرسایش یافته و اراضی کشاورزی تشکیل شده‌اند.

۲- تجزیه و تحلیل داده‌های زمینی

در این تحقیق به منظور تجزیه و تحلیل تغییرات خاک و همچنین اثر سطح سفره آب زیرزمینی بر ویژگی‌های خاک، روش PCA برای دو شرایط متفاوت از نظر عمق سفره آب زیرزمینی به کار برده شد:

الف - عمق سفره آب زیرزمینی متغیر از ۲/۱ تا ۱۰ متر

ب- عمق سفره آب زیرزمینی کمتر از ۴ متر از مجموع ۵۵ پروفیل، ۳۷ پروفیل با سفره آب زیرزمینی کمتر از ۴ متر تشخیص داده شد. این داده‌ها برای مشخص کردن مهمترین متغیرها در خاک‌های منطقه به کار رفتند. مقایسه این دو حالت به منظور بررسی آثار سفره آب زیرزمینی در تغییرات پاره‌ای خصوصیات خاک سطح‌الارض صورت گرفت.

نتایج

در این تحقیق نتایج درصد واریانس PCها در جدول ۱ برای دو حالت مختلف زیر نشان داده شده است؛ حالت اول زمانی که از ۷ باند ورودی همان تعداد مولفه و حالت دوم هنگامی که از ۶ باند ورودی، شش مولفه اصلی استخراج گردید. همان‌گونه که جدول ۱ نشان می‌دهد، هرچه شماره مولفه‌های اصلی در مراتب پایین‌تر قرار می‌گیرد، میزان اطلاعات آن کمتر است و بدیهی است که PC1 (محور ۱) در جهت بیشترین واریانس و PC2 (محور ۲) عمود بر PC1 بوده و اطلاعات کمتری را داراست. بنابراین در تمامی موارد بیشترین درصد واریانس به PC1 و کمترین آن به PC7 مربوط است. مقایسه اشکال ۱ و ۲ نشان می‌دهد که ۳ تا PC اولیه بیشترین واریانس (معمولاً بیش از ۹۵ درصد) را دارد و میزان

به منظور تجزیه و تحلیل تغییرات خاک، از متغیرهای افق رویی خاک، مربوط به ۵۵ پروفیل خاک چاه افضل اردکان استفاده شد. در این مطالعه، علاوه بر متغیرهای افق رویی خاک ۵۵ پروفیل (موسسه خاک و آب یزد، ۲ و ۱۵)، از متغیرهای دیگری نظیر عمق سفره آب زیرزمینی (WTL)^۱، شوری آب زیرزمینی (ECwt)^۲ و ارتفاع (ELEV)^۳ که به ترتیب از نقشه‌های WTL، همتراز شوری سفره آب زیرزمینی (ECwt) و DEM استفاده شد. بدین نحو که پس از رقومی کردن نقشه‌های فوق با استفاده از نرم‌افزار ILWIS، عملیات رستر^۴ کردن، درون‌یابی^۵ و انطباق آنها انجام گرفت و سپس اطلاعات لازم از محل هر پروفیل به دست آمد. متغیرهای استفاده شده در روش PCA به شرح زیر است:

- ۱- هدایت الکتریکی (EC)
- ۲- کاتیون‌ها و آنیون‌های محلول خاک مانند Mg^{++} ، Ca^{++} ، Na^{+} ، HCO_3^{-} ، SO_4^{-} ، Cl^{-}
- ۳- بافت خاک (شن، سیلت و رس)
- ۴- ارتفاع از سطح دریای آزاد (ELEV)
- ۵- شوری آب زیرزمینی
- ۶- ضخامت افق سطحی (Al یا Cl)
- ۷- میزان رطوبت خاک (درصد وزنی) (MOIST)^۶
- ۸- درصد کربنات کلسیم ($CaCO_3$)
- ۹- گچ ($CaSO_4$)
- ۱۰- pH

۱- Water Table level

۲- Electrical Conductivity of Ground water

۳- Elevation

۴- Raster

۵- Interpolation

۶- Moisture

پوشش گیاهی است، بنابراین اختلاف ضرایب مثبت و منفی TM3 و TM4 نشان می‌دهد که در این مولفه اطلاعات زیادی روی پوشش گیاهی وجود دارد. قابل توجه است هنگامی که معادلات مربوط به تبدیل DNهای ۷ باند اصلی به DNهای مولفه‌های PCها به دست آمد، می‌توان به ارزیابی میزان و نوع اطلاعات موجود در هر PC پرداخت. برای مثال PC1 در تمامی مناطق مورد مطالعه همانند مجموع اطلاعات باندهای انعکاسی شش‌گانه است، اما PC مواردی نظیر سیاه‌کوه بیانگر تفاوت باندهای TM3 و TM4 و برخی اطلاعات از باند نرمال است. اطلاعات مربوط به پوشش گیاهی در مواردی هم در PC3 دیده می‌شود برای مثال، در منطقه فوق تفاوت بازتاب‌های باندهای TM3، TM4 و TM6 در PC3 نمایان‌تر است، زیرا PCهایی که بیشترین رابطه را با باند مادون قرمز نزدیک و همبستگی پایین‌تری با باندهای دیگر دارند، می‌توانند برای تخمین و استخراج اطلاعات پوشش گیاهی مورد استفاده قرارگیرند. دلیل این امر آن است که TM4 شاخصی از پوشش گیاهی است (۱۱)، علاوه بر این باند حرارتی TM6 نیز حاوی اطلاعات مفیدی درباره پوشش گیاهی است.

۲- نتایج حاصل از تجزیه مولفه‌های اصلی مربوط به متغیرهای خاک کل منطقه نشان می‌دهد که PC1 با واریانسی (اطلاعات) برابر ۳۲/۸ درصد بیشترین و PC18 با واریانسی نزدیک به صفر کمترین اطلاعات را به خود اختصاص داده است (جدول ۳)

اطلاعات سایر PCها بمراتب کمتر است. اگرچه PCهای مراتب پایین‌تر فاصله (PC4، PC5، PC6 و PC7) از لحاظ درصد اطلاعات، اهمیت زیادی ندارند، ولی ممکن است از لحاظ نوع اطلاعات حایز اهمیت باشند.

در شکل ۱، درصد واریانس و به عبارتی مقادیر ویژه (Eigenvalue) روی محور Y و PCها روی محور X نشان داده شده است که در محل PC3 مقادیر اطلاعات زیاد متوقف و اطلاعات کم شروع می‌شود. براساس همین نمودارهای ساده می‌توان ۳ مورد از PCهای اولیه را انتخاب و برای مراحل بعدی مورد استفاده قرار داد. هرچند PCهای پایین‌تر هم ممکن است دارای اطلاعات مفیدی باشند که PCهای اولیه فاقد آن می‌باشند. بنابراین این سوال مطرح می‌شود که با حذف PCهای پایین‌تر، چه اطلاعاتی از دست می‌رود؟ برای مثال برای منطقه ابرکوه، PC1 با واریانس ۸۱ درصد بیانگر این مهم است که تغییرات اصلی توسط این مولفه بیان می‌شود. برخی معتقدند همان تعداد PCهایی که تا ۹۰ درصد اطلاعات را بدهند، کافی است.

در جدول ۲، ماتریس بردار ویژه به کار رفته برای وزن باندهای موجود در مولفه‌های اصلی نشان داده شده است. این ضرایب می‌تواند بیانگر ارتباط برخی پدیده‌های زمینی با هر کدام از PCها باشد که البته این امر مستلزم شناخت کامل میزان انعکاس‌های پدیده‌ها در طیف‌های مختلف است. برای مثال در جدول ۲ منطقه سیاه‌کوه بالاترین ضریب (مثبت) به TM4 اختصاص یافته که دلالت بر وجود پوشش گیاهی دارد. از آنجایی که تفاوت TM3 و TM4 بیانگر

جدول ۱- درصد واریانس یا اطلاعات موجود در PCهای حاصل از مولفه‌های استخراج شده از مجموع ۷ (شامل باند

حرارتی) و ۶ باند به‌کاررفته (بدون باند حرارتی)

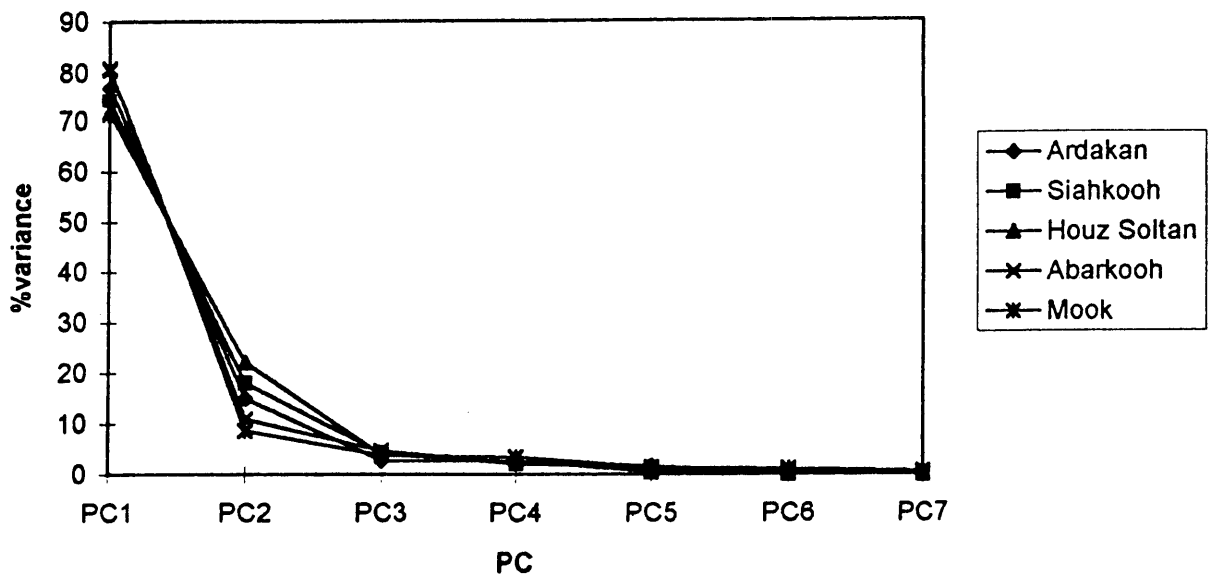
شماره PC	اردکان		چاه افضل		حوض سلطان		ابركوه		موک فارس	
	۷باند	۶باند	۷باند	۶باند	۷باند	۶باند	۷باند	۶باند	۷باند	۶باند
۱	۷۶/۹۷	۸۳/۲۰	۷۴/۱۰	۸۲/۶۴	۷۱/۷۸	۸۸/۶	۸۰/۵۶	۹۲/۰۴	۸۰/۹۱	۸۷/۵
۲	۱۵/۰	۱۰/۳۱	۱۸/۰۶	۱۱/۴۹	۲۲/۰۲	۵/۶۲	۱۰/۸۶	۳/۵۷	۸/۶۴	۴/۶۶
۳	۲/۷۸	۳/۱۹	۴/۲۹	۳/۳۳	۳/۸۹	۳/۵۶	۴/۶۹	۳/۲۳	۳/۸۲	۴/۰۱
۴	۲/۶۴	۲/۴۰	۲/۰۸	۲/۱۴	۲/۵۶	۱/۵۰	۲/۱۷	۰/۶۴	۳/۴۲	۱/۶۱
۵	۱/۷	۰/۵۱	۱/۲۹	۰/۲۶	۰/۴۸	۰/۴۹	۱/۱۴	۰/۳۸	۱/۳۸	۱/۴
۶	۰/۵۲	۰/۳۹	۰/۱۳	۰/۱۵	۰/۱۷	۰/۲۳	۰/۳۹	۰/۱۲	۱/۲	۰/۶۲
۷	۰/۴	-	۰/۰۵	-	۰/۰۵	-	۰/۱۸	-	۰/۶۲	-

جدول ۲- ماتریس بردار ویژه Eigenvector مربوط به وزن باندهای هفت‌گانه موجود در سه مولفه اولیه (Eigenvector)

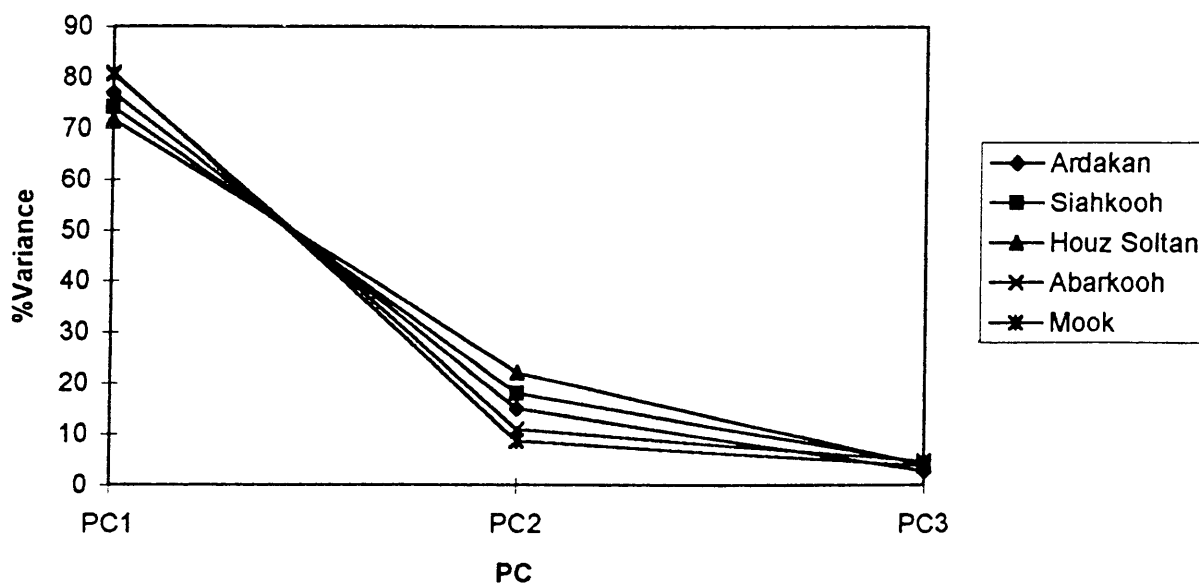
باندها	موک فارس			سیاه کوه			حوض سلطان			ابركوه		
	PC3	PC2	PC1	PC3	PC2	PC1	PC3	PC2	PC1	PC3	PC2	PC1
TM1	-۰/۰۸۸	-۰/۴۸۴	۰/۲۳۱	۰/۳۲۴	۰/۲۰۱	۰/۲۶۰	۰/۴۲۲	-۰/۴۲۵	۰/۴۲۹	۰/۳۵۵	۰/۴۱۰	۰/۴۱۱
TM2	-۰/۰۶۳	-۰/۳۲۷	۰/۲۳۷	۰/۱۹۹	۰/۳۰۱	۰/۲۷۰	۰/۱۳۱	-۰/۳۰۱	-۰/۱۵۳	۰/۲۲۱	۰/۱۰۹	۰/۱۰۹
TM3	-۰/۱۰۸	-۰/۳۶۶	۰/۳۸۰	۰/۲۷۳	-۰/۲۰۸	-۰/۳۱۵	۰/۲۸۷	۰/۲۸۵	۰/۳۱۶	-۰/۸۵۲	-۰/۰۵۶	۰/۰۶۷
TM4	-۰/۰۲۰	-۰/۲۸۱	۰/۳۱۰	۰/۱۸۸	۰/۱۷۵	۰/۳۱۸	-۰/۶۶۰	-۰/۰۰۴	-۰/۲۰۰	۰/۱۸۵	-۰/۲۷۴	-۰/۳۳۹
TM5	۰/۱۸۴	۰/۶۳۲	۰/۶۲۶	۰/۳۱۷	۰/۰۷۸	۰/۷۲۲	۰/۳۹۹	-۰/۲۹۶	-۰/۲۶۷	۰/۲۴۰	-۰/۵۳۰	-۰/۳۰۱
TM6	۰/۹۶۴	-۰/۰۱۵	۰/۱۶۴	-۰/۶۸۳	۰/۱۸۵	۰/۲۵۷	-۰/۳۲۶	۰/۲۱۳	۰/۴۲۸	-۰/۰۵۵	۰/۲۹۲	۰/۱۳۳
TM7	۰/۱۱۷	۰/۱۸۰	۰/۴۱۸	۰/۸۹۷	-۰/۲۶۴	۰/۱۱۷	۰/۵۳۵	-۰/۷۵۰	-۰/۶۳۰	۰/۰۳۸	-۰/۶۱۱	۰/۷۶۷

جدول ۳- مقایسه درصد کل واریانس مربوط به PCها

شماره	مولفه	WTL(2-10m)تجمعی	WTL(<4m)تجمعی
۱	PC1	۳۲/۸۱	۳۴/۹۱
۲	PC2	۴۵/۶۷	۴۸/۲۳
۳	PC3	۵۵/۴۱	۵۹/۰۳
۴	PC4	۶۳/۲۵	۶۷/۷۲
۵	PC5	۷۰/۱۵	۷۴/۱۴
۶	PC6	۷۶/۲۲	۷۹/۹۴
۷	PC7	۸۱/۶۵	۸۴/۴۱
۸	PC8	۸۵/۶۶	۸۸/۵۳
۹	PC9	۸۹/۲۳	۹۱/۳۵
۱۰	PC10	۹۲/۳۵	۹۳/۷۷
۱۱	PC11	۹۴/۸۳	۹۸/۹۵
۱۲	PC12	۹۶/۸۹	۹۷/۳۶
۱۳	PC13	۹۸/۲۲	۹۸/۴۲
۱۴	PC14	۹۹/۲۸	۹۹/۸۵
۱۵	PC15	۹۹/۸۶	۹۹/۸۵
۱۶	PC16	۹۹/۹۷	۹۹/۹۷
۱۷	PC17	۹۹/۹۹	۹۹/۹۹
۱۸	PC18	۱۰۰/۰۰	۱۰۰/۰۰



شکل ۱- درصد واریانس PCهای هفتگانه مناطق مورد مطالعه



شکل ۲- PCهای سه‌گانه اولیه مناطق مورد مطالعه

آنها در عملیات صحرایی و نمونه‌برداری بیشتر حایز اهمیت است. شکل ۴ PC1 و PC2 را با متغیرهای مورد مطالعه در منطقه الف (کل منطقه با تغییرات سفره آب زیرزمینی کمتر از ۴ متر) نشان می‌دهد، علاوه بر این، ملاحظه می‌شود که PC1 بیشترین رابطه را با متغیرهای EC، Na^+ ، Cl^- ، SO_4^{2-} ، رس و شن دارد. قابل توجه است که این متغیرها سهم زیادی در PC2 و PC3 نشان نمی‌دهند. برای مثال Na^+ با ضریب منفی (-۰/۳۹) بیشترین رابطه را با PC1 و همین متغیر کمترین ضرایب منفی (-۰/۰۲۵ و -۰/۰۸۹) را به ترتیب با PC2 و PC3 دارد (جدول ۴). زاویه‌های کوچک بین متغیرهای EC با Na^+ و Cl^- و SO_4^{2-} بیانگر رابطه قوی بین آنهاست. رابطه بالای متغیرهای شن و سیلت با PC1 نشان می‌دهد که آنها همبستگی زیادی باهم داشته‌اند. سهم بالای متغیرهای رس و شن با PC1 بیانگر آن است که شوری خاک بستگی به بافت خاک دارد. در شکل ۴، زاویه کوچک بین رس و EC نشان‌دهنده رابطه

جدول ۴ نشان می‌دهد که PC1 ۳۲/۸ درصد و PC2 حدود ۱۲/۸ درصد از کل اطلاعات را در خود جای داده است. بنابراین بیشترین اطلاعات در PC1 متراکم شده و کل اطلاعات مربوط به ۳ تا PC اول (حدود ۵۵/۴ درصد) و ۱۵ PC باقی‌مانده (حدود ۴۱ درصد) است. این نتایج نشان می‌دهد که بخش بسیار مهمی از مشخصات محیطی و فیزیکوشیمیایی خاک در ۳ تا PC اولیه متراکم شده است. در شکل ۱ درصد واریانس روی محور Y و PCها روی محور X نمایش داده شده، ملاحظه می‌شود که محل PC3، نقطه عطفی برای تغییرات اطلاعات است. براساس این معیار، تغییر وضعیت منحنی و انتخاب ۳ تا PC اول برای تجزیه و تحلیل مراحل بعدی مفیدترند.

شکل ۴ وابستگی‌های PC2 و PC1 را با ۱۸ متغیر موردنظر نشان می‌دهد. در این نمودار متغیرهایی که بزرگترین بردار ویژه مثبت یا منفی (قدر مطلق) را با PC1 یا PC2 نشان می‌دهند، مهمترین متغیرهایی هستند که مدنظر قرار دادن

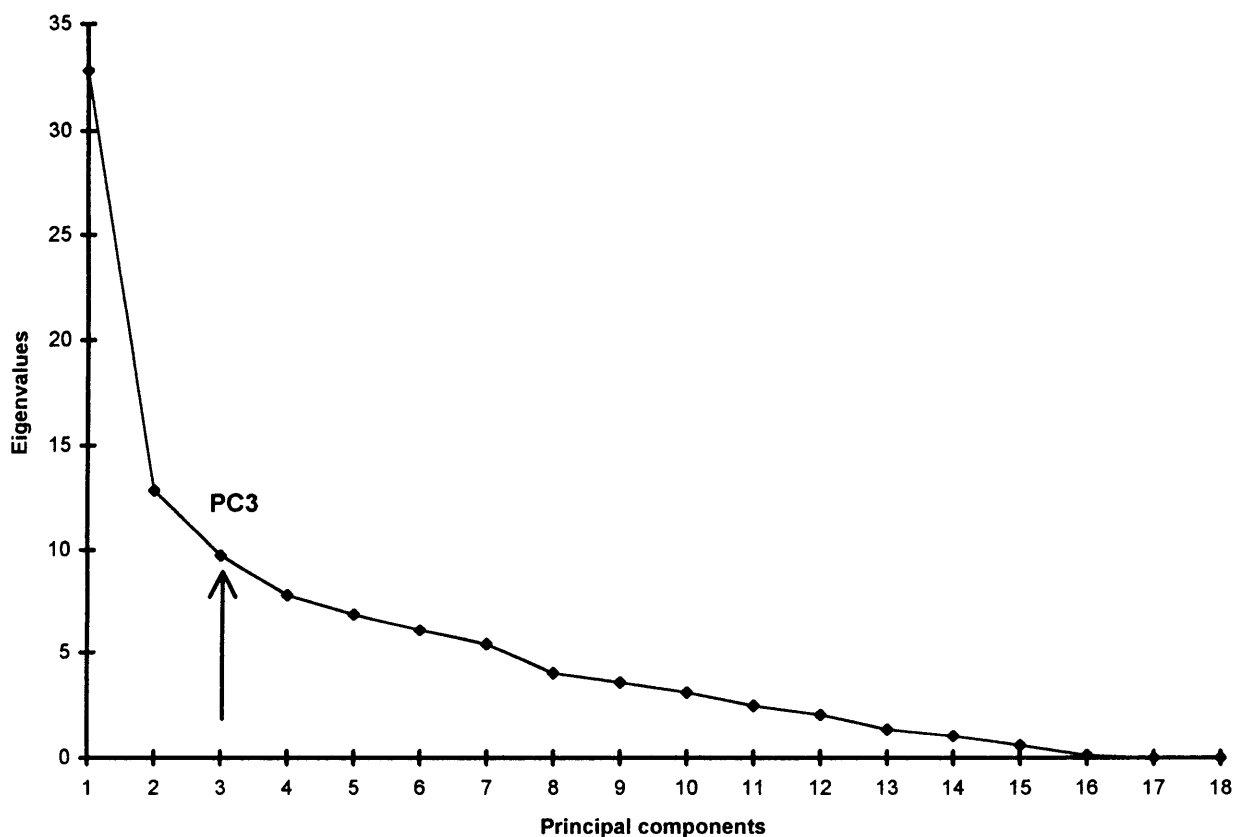
ضرایب PCها در دو حالت مورد مطالعه نشان می‌دهد که نتایج بسیار به هم نزدیک است و تنها اختلاف اصلی بین کل منطقه الف و منطقه ب این است که در حالت اخیر زاویه بین خط WTL و متغیرهای شوری (EC ، Na^+ و Cl^-) کاهش یافته است (شکل ۴ب) که دلیل این امر افزایش حرکت رو به بالای آب و در نتیجه تجمع بیشتر نمک در سطح خاک می‌باشد.

نتایج این تحقیق اهمیت PC1 را به عنوان بخش اصلی اطلاعات لازم برای تفسیر بسیاری از پدیده‌ها از جمله شوری خاک، آشکار می‌سازد.

قوی بین این دو متغیر است. از سوی دیگر، یک رابطه منفی برای متغیرهای شن و EC وجود دارد که با افزایش رس، شوری نیز افزایش می‌یابد. براساس این نتایج، PC1 را می‌توان به عنوان مولفه شوری خاک نامگذاری کرد، زیرا PC1 حاوی اطلاعات زیادی است که منعکس کننده پدیده شوری خاک سطح الارضی است. بنابراین شوری از مهمترین متغیرهایی است که در نمونه برداری و آنالیز خاک باید مدنظر قرار گیرد. PC1 در هر دو حالت (منطقه با $WTL < 4m$ و کل منطقه) سهم زیادی از $ELEV$ و WTL می‌برد که نشان می‌دهد آثار پستی و بلندی در PC2 جلوه گر شده است.

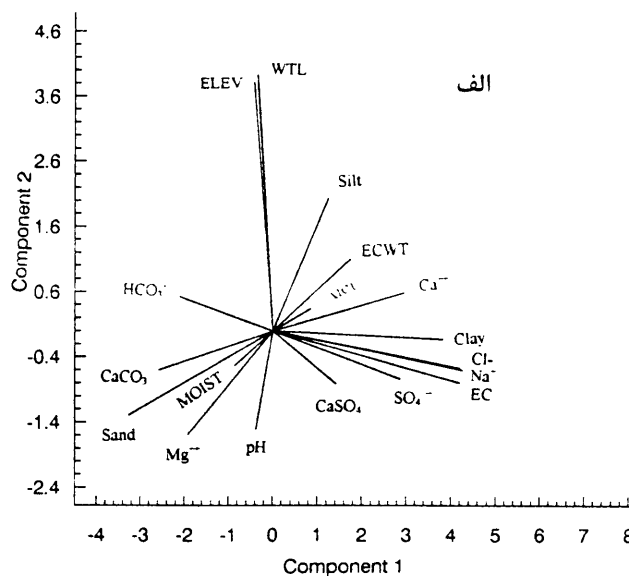
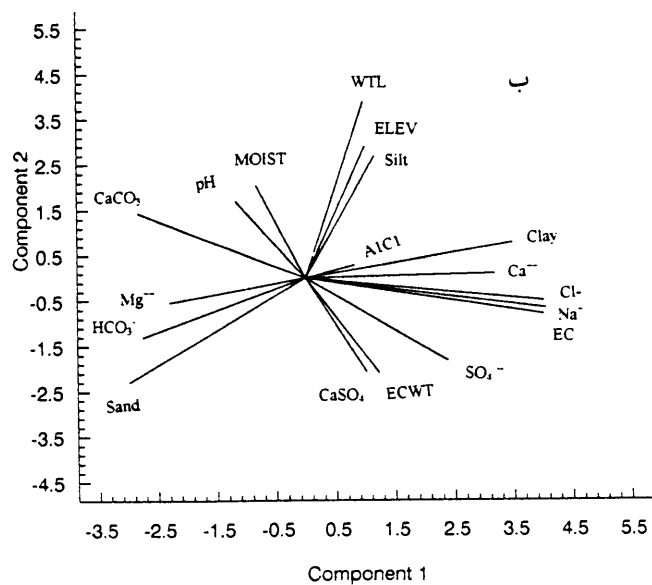
جدول ۴- ضرایب بردار ویژه (Eigenvector) سه مولفه اول براساس ماتریس همبستگی

WTL(<4m)			WTL(2-10m)			متغیر	شماره
PC3	PC2	PC1	PC3	PC2	PC1		
-۰/۰۸۵	-۰/۱۰۲	۰/۳۷۴	۰/۰۰	۰/۱۱۸	۰/۳۸۴	EC	۱
-۰/۳۷۲	۰/۳۷۰	۰/۰۹۳	۰/۲۱۵	۰/۵۶۴	-۰/۰۳۹	ELEV	۲
-۰/۱۵۷	۰/۲۶۸	۰/۱۱۵	-۰/۱۶۲	۰/۱۳۶	۰/۱۵۸	ECwt	۳
-۰/۲۸۳	۰/۴۹۵	۰/۰۹۱	۰/۱۹۷	۰/۵۸۱	-۰/۰۳۲	WTL	۴
-۰/۰۵۶	-۰/۰۸۵	۰/۳۷۸	-۰/۰۲۵	-۰/۰۸۹	-۰/۳۹۰	Na^+	۵
۰/۰۸۴	۰/۰۱۲	۰/۲۹۶	۰/۳۱۲	۰/۰۸۵	۰/۲۶۹	Ca^{++}	۶
-۰/۳۰۹	-۰/۰۶۸	-۰/۲۱۴	-۰/۱۰۸	-۰/۲۳۴	-۰/۱۷۷	Mg^{++}	۷
۰/۴۵۸	۰/۲۱۸	-۰/۱۰۹	۰/۳۱۳	-۰/۲۲۱	-۰/۰۳۵	PH	۸
-۰/۰۱۰	-۰/۱۶۶	-۰/۲۵۶	۰/۴۰۰	۰/۰۷۷	-۰/۱۹۵	HCO_3^-	۹
۰/۱۲۸	-۰/۲۳۳	۰/۲۲۲	۰/۱۹۹	-۰/۱۰۸	۰/۲۶۱	SO_4^{--}	۱۰
-۰/۰۹۲	-۰/۰۶۴	۰/۳۷۸	-۰/۰۰۳	-۰/۰۸۵	۰/۳۸۵	Cl	۱۱
۰/۱۱۲	-۰/۲۶۵	۰/۰۹۵	۰/۴۲۳	-۰/۱۱۹	۰/۱۲۹	$CaSO_4$	۱۲
۰/۱۳۳	/۱۸۳	-۰/۲۶۳	-۰/۲۵۷	-۰/۰۸۸	-۰/۲۳۷	$CaCO_3$	۱۳
-۰/۲۶۵	-۰/۲۹۴	-۰/۲۷۸	۰/۲۸۱	-۰/۱۹۰	-۰/۳۰۰	Sand	۱۴
۰/۳۳۳	-۰/۳۴۳	۰/۱۰۸	-۰/۳۷۶	۰/۳۰۰	۰/۱۱۳	Silt	۱۵
۰/۰۶۴	۰/۰۹۹	۰/۳۲۴	-۰/۰۴۱	-۰/۰۱۸	۰/۳۴۹	Clay	۱۶
-۰/۲۱۹	۰/۲۶۱	-۰/۰۷۶	۰/۰۰۲	-۰/۰۷۸	-۰/۰۷۹	MOIST	۱۷
-۰/۳۷۷	۰/۰۳۶	-۰/۰۷۶	-۰/۱۰۷	-۰/۰۵۸	۰/۰۷۶	AlCl	۱۸



شکل ۳- درصد واریانس مولفه‌ها روی محور Y و مولفه‌های اصلی روی محور X نشان داده شده است

Biplot for First Two Principal Component



شکل ۴- الف) رابطه متغیرهای هجده‌گانه با PC1 و PC2 در اراضی با سفره آب زیرزمینی ۲ - ۱۰ متر

و ب) اراضی با سفره آب زیرزمینی کمتر از ۴ متر

بحث و نتیجه‌گیری

براساس نتایج حاصل از تحقیق، ۳ PC طیفی اولیه، بخش اصلی و مهم اطلاعات ۷ باند TM را در خود جمع نموده که PC1 برای تفسیر خاک، PC2 یا گاهی PC3 برای تفسیر پوشش گیاهی مناسب‌اند. براساس نتایج PCهای حاصل از داده‌های زمینی، PC1 ناشی از ۱۸ متغیر خاک، اهمیت تغییرات شوری و بافت خاک را آشکار می‌سازد. براساس نتایج اولیه، ضرورت نمونه‌برداری شوری و بافت خاک ممکن است کمک زیادی به تفسیر آنها با استفاده از PC1 حاصل از تصاویر هفت‌گانه ماهواره‌ای بنماید. در این تحقیق با استفاده از نتایج حاصل از PCA و اطلاعات جنبی و عملیات نمونه‌برداری، متغیرهای مهم مانند EC و بافت تفسیر شد و معلوم گردید که شوری خاک عامل بسیار مهمی است که در تفسیر تصاویر ماهواره‌ای باید مورد توجه قرار گیرد (۲، ۳ و ۴) و لازم است عواملی که بر شوری خاک موثرند، به نحو مطلوبی شناسایی شوند (۱۳). براساس نتایج به‌دست آمده از مولفه‌های اصلی حاصل از باندهای ماهواره لندست و همچنین متغیرهای خاک، بیشترین درصد اطلاعات در سه مولفه اصلی متراکم می‌شود، ولی میزان اطلاعات مولفه‌های داده‌های طیفی بمراتب بیشتر از داده‌های مکانی است. براساس این مطالعات درمی‌یابیم که با استفاده از مولفه‌های اصلی مربوط به هر دو نوع اطلاعات طیفی و مکانی، می‌توان متغیرهای کم‌اهمیت برای نمونه‌برداری را شناسایی و سپس حذف کرد. این امر صرفه‌جویی در هزینه و وقت را به‌دنبال خواهد

داشت. قابلیت نامگذاری و تفسیر مولفه‌های اصلی مربوط به داده‌های طیفی و مکانی نشان می‌دهد که با استفاده از این روش علاوه بر متراکم ساختن اطلاعات پراکنده، می‌توان آنها را براساس ماهیت و نوع اطلاعات طبقه‌بندی کرد. نظر به اینکه تعیین رابطه دقیق بین نوع اطلاعات مولفه‌های مکانی نیاز به اطلاعات کاملی از داده‌های زمینی دارد، بنابراین پیشنهاد می‌گردد که در تحقیقات بعدی رابطه دقیق‌تر قابلیت تفسیر مولفه‌های طیفی براساس مولفه‌های مکانی مدنظر قرار گیرد. برای پی بردن به کاربرد این روش در منابع طبیعی، پیشنهاد می‌شود این مطالعات در مناطق دارای پوشش گیاهی انبوه نیز صورت گیرد تا بدین طریق نقش متغیرهای گیاهی و بوم‌شناختی در رابطه بین داده‌های مکانی و طیفی روشن شود. علاوه بر این، پیشنهاد می‌گردد که مطالعات در مناطق دیگر کویری با به‌کارگیری متغیرهای موثر و گسترده‌تر نظیر نوع نمک و ساختمان خاک و کانی‌ها جهت طراحی و نحوه بهتر نمونه‌برداری انجام شود.

قدردانی

بدین‌وسیله از حمایت‌های معاونت پژوهشی دانشگاه تهران در فراهم ساختن امکانات مالی و همکاری اعضای محترم مرکز تحقیقات مناطق کویری و بیابانی ایران و نیز بخش آب و خاک سازمان جهاد کشاورزی یزد و مرکز سنجش از دور ایران به خاطر در اختیار قراردادن داده‌های ماهواره تقدیر و تشکر می‌شود.

منابع

- 1-Adb El-Hady, 1992. Discrimination of gypsiferous calcareous and sandy soil surfaces using the Principal Components (PC) Transformation of Thematic Mapper, Egypt. J. Soil .Sci. Vol 32 (1):71-80.
- 2-Alavi Panah, S.K. 1997. Study of soil salinity in the Ardakan area (Iran) based upon field observations, remote sensing and a GIS P.292. Gent: University of Gent (Ph.D.Thesis).
- 3-Alavi Panah, S.K., R. Goossens & M. De Dapper, 1999. Study of soil Salinity in the Ardakan Area, Iran, based upon field observations and remote sensing, G.J.A. Nieuwenhuis, R.A Vaughan & M. Molenaar (ed.), Operational Remote Sensing for Sustainable Development, Proceedings of 18th EARSeI Symposium, Enschede, Netherlands, 11-14May, 1998:419-425.
- 4- Alavi Panah, S.K., M.De Dapper, R.Goossens & M.Massoudi, 2001. The use of TM thermal band for land cover/land use mapping in two different conditions of Iran. Sci Tech. Vol 3,331-340.
- 5-Asrar, G. 1989. Theory and applications of practical remote sensing, John Wiley, NewYork.
- 6-Byrne, G.F., P.K. Crapper and K.K. Mayo, 1980. Monitoring land cover change by component analysis of multi-temporal Landsat data. Remote Sensing of Environment. Vol 1(10): 175-184.
- 7-Campbell, J.B. & J.O. Browder, 1995. Field data collection for remote sensing analysis: SPOT data, Rondonia, Brazil. Intern. Journal of Remote Sensing, Vol 16(2): 333-350.
- 8-Cole, J. & C.L. King, 1968. Quantitative geography, John Wiley and Sons Ltd., 692 p.
- 9-Daeles, L. & M. Antrop, 1977. The extraction of soil information from remote sensing documents. Pedologic, Vol: 27:123-190.
- 10- James, R.C. & R.Matanawi, 1999. Correspondance analysis for principal components transformation of multispectral and hyperspectral digital images. Photogrametric Engineering and remote sensing, vol 65 (8) : 909-914.
- 11-Lillesand, & Kiefer, 1994. Remote sensing and image interpretation, John wiley and Sons, newYork.
- 12-Montgomery, O.L. 1976. An investigation of the relationship between the spectral reflectance and the chemical, physical and genetic characteristics of soil, Ph.D. Dissertation, Purdue Univ., West Lafayette, IN. Univ . Microfilm. Ann Arbor, MI (Diss. Abstr. 37/08/-8:3707).
- 13-Mulder, M. 1987. Remote Sensing in soil science. Development in Soil science, Elsevier. Amsterdam, The Nethrlands, 379 p.
- 14-Severson, R.C. 1981. Evaluating chemical character of soil material for

suitability in mined land in the Sanjuan Basin, New Mexico. Soil Sci. Soc. Am.J.Vol:45: 396-404.

15-Yazd Soil and Water Institute, 1992. Soil and salinity map and data, Yazd, Iran.

Study of Natural Features Based On Principal Component Analysis

S.K. Alavi Panah¹

Abstract

In this study, PC technique was used to reduce the number of spectral bands or spatial variables in a data set by finding their linear combinations. To apply the PCA for spectral/spatial data set, Landsat TM data recorded from 5 different areas in Central Iran, and 18 soil variables were used. The result of PCA transformation for TM bands revealed the importance of PC1 in soil studies, PC2 and or PC3 for vegetation investigations. The results of PCA for Landsat TM and soil data showed that the TM data of 7 bands and 18 soil variables were mainly compressed to just three PCs that describe more than 90% and 55% of spectral and spatial information respectively. Based on the obtained results we may conclude that PCA can be applied to different sources of spectral/spatial data, for a better establishment of sampling plan and save in money and time.

Keywords: Soil variables, Spectral band, Spatial information, Principal component

¹ - Asst. Prof., Iran Desert Research Center, University of Tehran