

## کاربرد مدل داده‌های ترکیبی در پیش‌بینی سطح آب زیرزمینی

عزیزا... ایزدی<sup>۱</sup>، کامران داوری<sup>۲</sup>، امین علیزاده<sup>۲</sup> و بیژن قهرمان<sup>۲</sup>

### چکیده

منابع آب‌های زیرزمینی یکی از مهمترین و ارزانه‌ترین منابع آب به شمار می‌روند. شناخت صحیح و بهره‌برداری اصولی از آنها در مناطق خشک و نیمه خشک می‌تواند در توسعه پایدار فعالیت‌های اجتماعی و اقتصادی آن منطقه نقش به‌سزایی داشته باشد. برای آگاهی از وضعیت منابع آب زیرزمینی در چنین مناطقی لازم است پیش‌بینی دقیقی از نوسانات سطح آب زیرزمینی انجام شود. این پژوهش برای پیش‌بینی رفتار سطح آب زیرزمینی در دشت نیشابور با استفاده از مدل داده‌های ترکیبی انجام گردیده است. مدل داده‌های ترکیبی، مدل‌های رگرسیونی هستند که به دلیل در نظر گرفتن چندین پیرومتر در طول یک دوره زمانی، توانایی پیش‌بینی سطح آب زیرزمینی را در پیرومترهای مختلف به صورت توأم دارا می‌باشند. دشت فوق به دلیل داشتن بیش از ۵۰ پیرومتر که اکثراً دارای بیش از ۱۲ سال آمار هستند، برای این پژوهش انتخاب گردید. بدین منظور ابتدا پیرومترهای موجود در سطح دشت به روش "وارد" خوشه‌بندی شده و برای هر خوشه یک پیرومتر به عنوان نماینده انتخاب گردید. سپس، برای محدوده تحت پوشش پیرومترهای معرف هر خوشه، مقادیر متغیرهای مستقل شامل بارندگی، دمای حداکثر، حداقل و متوسط و سطح آب اولیه با استفاده از روش عکس فاصله محاسبه گردید. در نهایت عملکرد مدل‌های داده‌های ترکیبی مختلفی مانند داده‌های ترکیبی با اثرات مشترک، اثرات ثابت و اثرات تصادفی برای پیش‌بینی سطح آب زیرزمینی مورد بررسی قرار گرفت. نتایج نشان داد که مدل داده‌های ترکیبی با اثرات مشترک بهترین نتیجه را برای پیش‌بینی سطح آب زیرزمینی داشته است. معیارهای عملکرد  $R^2=0/99$  و  $RMSE=0/05$  در دوره آزمون نیز حاکی از کارایی این مدل می‌باشند. بعلاوه این نتایج با نتایج بدست آمده از کاربرد شبکه عصبی مصنوعی نیز مقایسه و برتری نسبی آن مورد تأیید قرار گرفت.

**واژه‌های کلیدی:** داده‌های ترکیبی، آب زیرزمینی، خوشه‌بندی وارد، دشت نیشابور

### مقدمه

منابع آب‌های زیرزمینی یکی از مهمترین و ارزانه‌ترین منابع آب به شمار می‌روند که شناخت صحیح و بهره‌برداری اصولی از آنها می‌تواند در توسعه پایدار فعالیت‌های اجتماعی و اقتصادی یک منطقه، به ویژه در مناطق خشک و نیمه خشک، نقش به‌سزایی داشته باشد. عدم شناخت صحیح و بهره‌برداری بی‌رویه از این منابع خسارت جبران ناپذیری مانند افت شدید و غیرقابل برگشت سطح آب زیرزمینی، کاهش دبی چاه‌ها و قنوات (باشقره، ۱۳۷۷)، تغییرات الگوی جریان آب زیرزمینی مانند پیشروی جبهه‌های آب شور و تداخل آب‌های شور و شیرین (Nayak et al., 2006) به دنبال خواهد داشت. بدین منظور برای آگاهی از وضعیت منابع آب زیرزمینی لازم است پیش‌بینی دقیقی از نوسانات سطح آب زیرزمینی انجام شود. با پیش‌بینی دقیق نوسانات سطح آب زیرزمینی می‌توان از آن در برنامه‌ریزی تأمین آب قابل اعتماد و نیز در مدیریت منابع آب استفاده نمود. برای پیش‌بینی نوسانات سطح آب زیرزمینی به علت پیچیده بودن ماهیت آن و عدم

قطعیت در مولفه‌های مربوطه نیاز به مدل‌های پیچیده‌ای می‌باشد (Nayak et al., 2006). بنابراین یکی از مهمترین رویکردها در برنامه‌ریزی مدیریت منابع آب به دست آوردن مدل مناسبی جهت پیش‌بینی رفتار منابع آب تحت متغیرهای موثر بر این پدیده می‌باشد. مهمترین مدل‌های رایج در این زمینه مدل‌های رگرسیونی با سری زمانی (Tankersley et al., 1993)؛ Van Geer and Zuur, 1997)؛ (Knotters and Van Walsum, 1997)، مدل‌های فیزیکی (Belmans et al., 1983 و Feddes et al, 1988) و مدل‌های جدید مانند شبکه عصبی مصنوعی (Maier and Dandy, 2000)؛ (Coulibaly et al., 2001 و Nayak et al., 2003)؛ (Allen et al., 2006 و Sami et al., 2002) و (2007) می‌باشند. از مدل‌های سری زمانی تجربی هر چند به طور گسترده‌ای برای مدل‌سازی سطح آب استفاده شده است، ولی هنگامی که رفتار دینامیک یک سیستم هیدرولوژیک با گذشت زمان تغییر می‌کند مدل‌های یاد شده برای پیش‌بینی مولفه‌های منابع آب توانایی کافی نداشته و مدل‌های مناسبی نیستند (Bierkens, 1998). از طرف دیگر، مدل‌های بر مبنای فیزیکی نیز در عمل به داده‌های زیادی برای پیش‌بینی سطح آب زیرزمینی نیاز دارند و از آنجا که روابط بین متغیرهای موثر بر سطح آب در یک سفره آب زیرزمینی

۱- دانشجوی دوره دکتری گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد،  
E-mail: az.izady@gmail.com

۲- به ترتیب استادیار، استاد و دانشیار گروه مهندسی آب، دانشکده کشاورزی، دانشگاه فردوسی مشهد

مدیریت منابع آب محدود بوده است. بنابراین هدف از این تحقیق نشان دادن پتانسیل کاربرد مدل فوق در مدیریت منابع آب و ارائه یک مدل رگرسیونی با داده‌های ترکیبی برای پیش‌بینی سطح آب زیرزمینی در دشت نیشابور می‌باشد. در واقع مدل فوق توانایی ارائه یک مدل رگرسیونی به صورت توأم برای چندین مقطع (پیزومتر) را دارا می‌باشد. در این مدل، ویژگی‌هایی که در روابط بین واحدهای مقطعی مختلف موثر می‌باشند به صورت خودکار در برآورد مدل در نظر گرفته می‌شوند.

## منطقه مورد مطالعه و داده‌ها

### معرفی منطقه

دشت نیشابور جزئی از حوضه آبریز کالشور نیشابور می‌باشد که در دامنه جنوبی ارتفاعات بینالود و در شمال شرق کویر مرکزی واقع شده است. این حوضه در طول جغرافیایی ۱۷° ۵۸' تا ۳۰° ۵۹' شرقی و عرض جغرافیایی ۴۰° ۳۵' تا ۳۹° ۳۶' شمالی واقع شده و از شمال به خط‌الرأس ارتفاعات بینالود، از شرق به بلندی‌های لیلاجوق و یال‌پلنگ، از جنوب به تپه ماهورهای نیزه‌بند، سیاه‌کوه و کوه نمک و از غرب به حوضه آبریز دشت سبزوآر محدود می‌شود (ولایتی و توسلی، ۱۳۷۰). شکل (۱) موقعیت منطقه مورد مطالعه را نشان می‌دهد.

وسعت کل حوضه ۷۳۵۰ کیلومتر مربع است که ۳۱۶۰ کیلومتر مربع آن را ارتفاعات و ۴۱۹۰ کیلومتر مربع را دشت تشکیل می‌دهد. بلندترین نقطه منطقه در ارتفاعات بینالود بوده که از سطح دریا ۳۳۰۰ متر ارتفاع دارد. پایین‌ترین نقطه در محل خروجی دشت (حسین‌آباد جنگل) قرار دارد که حدود ۱۰۵۰ متر از سطح دریا بلندتر است (ولایتی و توسلی، ۱۳۷۰). آب و هوای منطقه بری، نیمه خشک تا خشک است. میانگین دمای ماهانه در ایستگاه بار (معرف مناطق کوهستانی) ۱۳ درجه سانتی‌گراد و در ایستگاه محمدآباد - فدیشه (معرف مناطق دشتی) ۱۳/۸ درجه سانتی‌گراد است. متوسط بارندگی و تبخیر در کل حوضه به ترتیب ۲۳۴ و ۲۳۳۵ میلی‌متر در سال گزارش شده است (شرکت سهامی آب منطقه‌ای خراسان رضوی، ۱۳۷۷).

دشت نیشابور در سال‌های اخیر به علت برداشت زیاد از آب‌های زیرزمینی برای مصارف کشاورزی با مشکل افت شدید سطح آب زیرزمینی مواجه شده است. به‌طوریکه از سال ۱۳۶۵ به بعد وزارت نیرو دشت فوق را به عنوان دشت ممنوعه اعلام نموده است (حسینی و همکاران، ۱۳۸۴). اکثر چاه‌های موجود در سطح دشت در حال برداشت غیر مجاز (مغایر با پروانه یا بدون پروانه) می‌باشند؛ بعلاوه پمپاژ از آنها نیز بر مبنای مطالعات آب زیرزمینی صورت گرفته نیست. همچنین به دلیل غالب بودن کشت‌های سبزی برگ (با نیازهای آبی زیاد) در منطقه و سطح زیر کشت غیر متناسب با توان منابع آبی

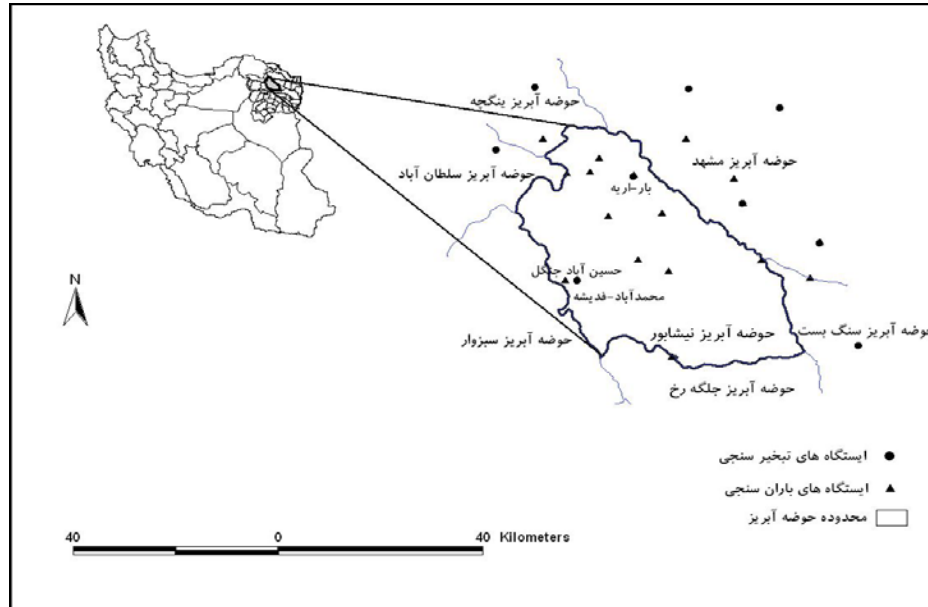
احتمالاً غیر خطی می‌باشند مدل‌های فوق در ارائه رابطه بین این متغیرها نیز نمی‌توانند به خوبی عمل کنند (Nayak et al., 2006). مدل‌های شبکه عصبی و منطق فازی هر چند برای مدل‌سازی رفتار دینامیک سیستم‌های غیر خطی مناسب می‌باشند ولی دارای این معایب هستند که هیچ‌گونه درکی از فیزیک مسئله نمی‌توانند داشته باشند و تنها برای زمانی مناسب‌اند که پیش‌بینی دقیق در حوزه کاری موردنظر مهمتر از درک فیزیک مسئله باشد (Daliakopoulos et al., 2005). در این میان، مدل‌های رگرسیونی با داده‌های ترکیبی<sup>۱</sup> به دلیل در نظر گرفتن واحدهای مقطعی در طی زمان توانایی زیادی در برآورد روابط بین متغیرهای موثر بر سطح آب زیرزمینی در چندین پیزومتر را به صورت توأم دارا می‌باشند. یعنی اینکه بر خلاف مدل‌های سری زمانی که قادر به پیش‌بینی سطح آب زیرزمینی در یک چاه را دارا هستند مدل‌های با داده‌های ترکیبی قادر به پیش‌بینی سطح آب زیرزمینی در چندین پیزومتر بصورت توأم در طول یک سری زمانی می‌باشند. به این صورت که علاوه بر زمان، مکان را نیز در پیش‌بینی در نظر می‌گیرد که در این صورت ویژگی‌هایی که در روابط بین پیزومترهای مختلف بر سطح آب تأثیرگذار می‌باشد در پیش‌بینی سطح آب زیرزمینی در نظر گرفته می‌شود. منظور از واحدهای مقطعی در این تحقیق پیزومترها می‌باشند که در این حالت چندین پیزومتر به صورت توأم در مدل داده‌های ترکیبی در نظر گرفته می‌شود. از مزایای مدل فوق می‌توان به افزایش حجم نمونه و افزایش کارایی اشاره کرد (Hsiao, 2003).

بررسی منابع نشان می‌دهد که کاربرد مدل فوق تا به امروز در مدیریت منابع آب محدود بوده و تحقیق جدی در این زمینه صورت نگرفته است. ولی در بحث اقتصاد کاربرد فراوانی داشته است. بر این اساس، طاهری (۱۳۸۴) نقش عوامل سطح زیر کشت، بارندگی و قیمت مناسب را در خودکفایی تولید گندم در ۲۸ استان کشور با استفاده از داده‌های ترکیبی مورد بررسی قرار داده است. بر اساس نتایج حاصل شده سطح زیر کشت بیشترین اهمیت را از آن خود ساخته و متغیرهای دیگر در رتبه بعدی قرار گرفته‌اند. مهرگان و همکاران (۱۳۸۵) رابطه تنگاتنگ بین نرخ تورم و نرخ بهره را در ۲۴ کشور طی دوره ۲۰۰۱ تا ۲۰۰۳ با استفاده از داده‌های ترکیبی مورد بررسی قرار داده است. نتایج بیانگر آن است که از لحاظ آماری رابطه علی یک طرفه از سوی نرخ بهره به سمت نرخ تورم وجود دارد. هادیان و همکاران (۱۳۸۵) برای تخمین تابع تقاضای بیمه درمان مکمل و (De Cian et al., 2007) برای بررسی اثر تغییرات دمایی بر تقاضای منابع انرژی (گاز، فرآورده‌های نفتی، زغال سنگ و الکتروسیته) از داده‌های ترکیبی استفاده کرده‌اند.

همانطور که در بالا ذکر شد کاربرد مدل فوق تا به امروز در

نیز در به وجود آمدن چنین شرایطی تأثیر فراوانی داشته است.

منطقه، برداشت از سفره آب زیرزمینی بیشترین تأثیر را در افت سطح آب داشته است. هر چند کاهش روند بارندگی و افزایش تبخیر-تعرق



شکل (۱) موقعیت منطقه مورد مطالعه

نزدیک بودن به کوهپایه‌ها و یا چاه‌های کشاورزی که روی نوسانات سطح آب داخل پیزومترها تأثیر گذار می‌باشند در نظر گرفته شده است. بدیهی است که نوسانات سطح آب در پیزومترهای نزدیک به رودخانه‌ها و یا چاه‌های کشاورزی زیاد بوده و به همین دلیل شاید معرف سطح آب واقعی نباشد که البته بیان فوق مورد تأیید کارشناسان امر نیز بوده و برای انتخاب پیزومترها از تجربیات آنها نیز استفاده شد.

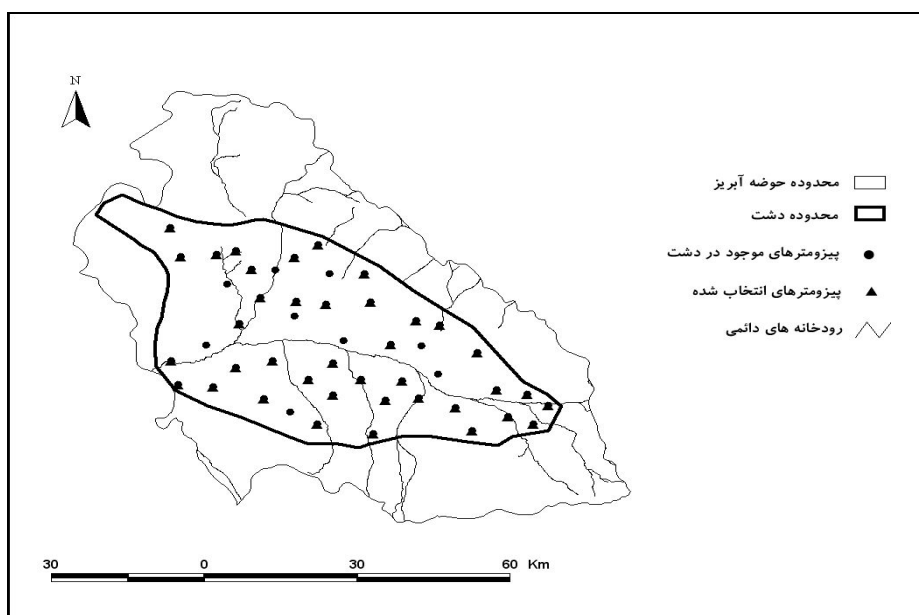
#### روش تحقیق

روش تحقیق در این مطالعه به این صورت است که ابتدا بوسیله تحلیل خوشه‌ای پیزومترهای انتخاب شده در سطح دشت را خوشه-بندی نموده تا بتوان برای هر یک از خوشه‌ها نماینده معرف آن خوشه را پیدا نمود. خوشه‌بندی پیزومترهای موجود به این خاطر می‌باشد که تحلیل تمامی پیزومترها مشکل بوده و به جای بررسی تک تک پیزومترها می‌توان رفتارهای یکسان را در قالب یک پیزومتر تجزیه و تحلیل نمود فرایندی که در تحلیل خوشه‌ای صورت می‌گیرد، جمع آوری اطلاعات و ترکیب آنها به گونه‌ای است که داده‌ها و مشاهدات به شکلی مناسب گروه‌بندی شوند (David, 1997). در این حالت باید گروه‌های مختلف حداکثر تفاوت ممکن را با هم داشته باشند و داده‌های موجود در یک گروه نیز بسیار شبیه به هم باشند. برای نیل به این هدف از سه متغیر میانگین ۱۲ ساله ماهانه تراز سطح آب زیرزمینی، مختصات جغرافیایی شامل طول، عرض و ارتفاع پیزومترهای موجود در سطح دشت به عنوان متغیرهای مناسب جهت

#### آمار و اطلاعات

با توجه به چنین وضعیتی، متغیرهای موردنظر برای پیش‌بینی سطح آب زیرزمینی شامل بارندگی، دمای حداکثر، حداقل و متوسط و سطح آب اولیه در مقیاس ماهانه می‌باشد. بارندگی شاخصی از تغذیه آب زیرزمینی بوده و دما نیز بیانگر تبخیر از سطح دشت در نظر گرفته می‌شود. قابل ذکر است که متغیرهای درنظر گرفته شده به طور گسترده‌ای در مقالات دیگر نیز به کار برده شده است (Coulbaly et al., 2005). با توجه به متغیرهای در نظر گرفته شده، اطلاعات خام اولیه مورد نیاز شامل سطح آب پیزومترهای موجود در سطح دشت، میزان بارندگی و دمای حداقل، حداکثر و متوسط در ایستگاه‌های موجود در سطح دشت و در مجاورت آن در مقیاس ماهانه در دوره سری زمانی ۱۲ ساله (از ۷۲-۱۳۷۱ تا ۸۳-۱۳۸۲) از شرکت سهامی آب منطقه‌ای خراسان رضوی تهیه گردید. ایستگاه‌های تبخیرسنجی و باران‌سنجی استفاده شده در شکل (۱) نشان داده شده است. اطلاعات تهیه شده با استفاده از روش میان‌یابی اسپلاین<sup>۱</sup> از نظر خلأهای آماری مورد بازسازی و ویرایش نهایی قرار گرفت. از بین پیزومترهای موجود در سطح دشت که بالغ بر ۵۰ عدد بودند، ۳۹ پیزومتر که دارای آمار با سری زمانی کامل و داده‌های دقیق می‌باشند انتخاب گردید (شکل ۲). برای انتخاب پیزومترها مواردی مانند کامل بودن داده‌ها، موقعیت مکانی پیزومترها از نظر قرارگیری در مجاورت رودخانه‌ها یا

1- Spline



شکل (۲) نمایش پیزومترهای موجود و انتخاب شده در سطح دشت

### تئوری مدل

#### مدل‌های رگرسیونی با داده‌های ترکیبی

اصولا داده‌های آماری به سه دسته داده‌های سری زمانی، مقطعی و ترکیبی تقسیم می‌شوند. در داده‌های ترکیبی، واحدهای مقطعی طی زمان مورد بررسی قرار می‌گیرند. استفاده از روش داده‌های ترکیبی نسبت به روش‌های مقطعی و سری‌های زمانی دو مزیت عمده دارد: ابتدا اینکه به محقق این امکان را می‌دهد تا ارتباط بین متغیرها و حتی واحدها را در طول زمان در نظر بگیرد و به بررسی آنها بپردازد. مزیت و توانایی دیگر این روش در بهره‌مندی آن از اثرات انفرادی مربوط به واحدها است که قابل مشاهده و اندازه‌گیری نیستند (Baltagi, 1995). روش تخمین در این مدل بیشتر بر عدم تجانس بین مقاطع تاکید می‌کند و هدف آنها بیشتر متمایز کردن مقاطع از همدیگر با استفاده از تخمین عرض از مبدأهای مختلف است. این تمایز به سه صورت اثرات مشترک<sup>۹</sup>، ثابت<sup>۱۰</sup> و تصادفی<sup>۱۱</sup> قابل انجام است (Yaffee, 2003).

#### روش اثرات مشترک

روش اثرات مشترک بر این فرض مبتنی است که عرض از مبدأ توابع ( $\alpha_i$ ) برای کلیه واحدها (مکان - سری زمانی یک پیزومتر) ثابت بوده و با هم برابرند ( $\alpha_i = \alpha$ ). این مدل با ترکیب داده‌های مقطعی و

انجام فرایند تحلیل خوشه‌ای به روش خوشه بندی پیوندی "وارد"<sup>۱</sup> که یکی از رایج ترین روش‌های خوشه بندی است استفاده گردید. نرم‌افزار به کار رفته برای این امر Minitab 2.3 می‌باشد. سپس با استفاده از روش عکس فاصله<sup>۲</sup> مقادیر متغیرهای مستقل شامل بارندگی، دمای حداکثر، حداقل و متوسط و سطح آب اولیه برای هر یک از نماینده‌های خوشه تعیین گردید. برای تعیین نماینده هر خوشه، ابتدا مجموع مربعات انحراف از میانگین هر خوشه را برای پیزومترهای موجود در داخل هر خوشه محاسبه و پیزومتری را که دارای کمترین مجموع انحرافات می باشد به عنوان نماینده آن خوشه انتخاب می‌گردد. بعد از آن با استفاده از آماره‌های چو<sup>۳</sup>، LM بروش<sup>۴</sup> - پاگان<sup>۵</sup> و هاسمن<sup>۵</sup> بهترین مدل داده‌های ترکیبی برای دوره آموزش (۷۱-۷۲ تا ۸۱-۸۲) انتخاب و در پایان نیز با استفاده از بهترین مدل انتخاب شده، تراز سطح آب زیرزمینی در دوره آزمون (مهر ۱۳۸۲ تا شهریور ۱۳۸۳) پیش‌بینی گردید. برای بررسی عملکرد مدل از معیارهای عملکرد متفاوتی مانند ضریب تبیین<sup>۶</sup>، میانگین مجذور مربعات خطا<sup>۷</sup> و حداکثر خطا<sup>۸</sup> استفاده به عمل آمد.

- 1- Ward Clustering Method
- 2- Distance Inverse Method
- 3- The Chow test
- 4- The Breusch – Pagan LM test
- 5- The Hausman test
- 6- Coefficient of Determination
- 7- Root Mean Square Error
- 8- Maximum Error

- 9- Common Effects
- 10- Fixed Effects
- 11- Random Effects

این فرضیه با توجه به آماره F مورد آزمون قرار می‌گیرد:

$$F(n-1, nT-n-k) = \frac{(R_f^2 - R_c^2)/(n-1)}{(1-R_f^2)/(nT-n-k)} \quad (۴)$$

که در آن n: تعداد مقاطع، T: تعداد مشاهدات سری زمانی، k: تعداد متغیرهای مستقل،  $R_f^2$  و  $R_c^2$  ضریب تبیین بترتیب در روش اثرات ثابت و مشترک می‌باشد. چنانچه آماره F محاسبه شده بیشتر از آماره مقدار بحرانی در سطح ۹۹٪ باشد، آنگاه فرضیه  $H_0$  رد می‌شود و این بدان معنی است که عرض از مبدأ پیرومترها متفاوت است. آزمون LM بروش - پاگان برای انتخاب بین دو روش اثرات مشترک و اثرات تصادفی با استفاده از باقیمانده‌های مدل اثرات مشترک قابل محاسبه می‌باشد:

$$LM = \frac{nT}{2(T-1)} \left[ \frac{T^2 \sum_{i=1}^n e_i^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} \right]^2 \quad (۵)$$

$e$ : باقیمانده مقادیر مشاهده شده و پیش‌بینی شده در مدل اثرات مشترک است. بقیه عوامل دقیقاً همان عوامل معادله (۴) می‌باشد. این آماره دارای توزیع کی دو ( $\chi^2$ ) با درجه آزادی یک می‌باشد. چنانچه آماره LM محاسبه شده بیشتر از آماره مقدار بحرانی در سطح ۹۹٪ باشد بدین مفهوم است که تفاوت بین مقاطع تصادفی است یعنی اینکه عرض از مبدأها متفاوت نبوده بلکه تفاوت‌ها در جمله اختلال ظاهر می‌شوند.

اگر الگوی موردنظر فقط با یک پارامتر تخمین زده شود آزمون هاسمن برای انتخاب بین دو روش اثرات ثابت و اثرات تصادفی به شکل ذیل خواهد بود:

$$h = \frac{(\beta_b - \beta_a)}{S_b^2 - S_a^2} \quad (۶)$$

که در آن  $(\beta_b - \beta_a)$  و  $S_b^2 - S_a^2$  به ترتیب پارامتر تخمین و خطای استاندارد مربوطه از مدل اثرات ثابت و مدل اثرات تصادفی می‌باشند. آزمون هاسمن برای مدل‌هایی با k پارامتر به شکل ماتریسی قابل محاسبه می‌باشد. آماره فوق دارای توزیع  $\chi^2$  با k (تعداد متغیرهای مستقل) درجه آزادی می‌باشد و چنانچه آماره h محاسبه شده کمتر از آماره مقدار بحرانی باشد بدین مفهوم است که عرض از مبدأ پیرومترهای مختلف متفاوت نبوده و تفاوت‌ها در جمله اختلال ظاهر می‌شوند.

## بحث و نتایج

### تحلیل خوشه‌ای

همانطور که در بالا اشاره شد برای انجام فرایند تحلیل خوشه‌ای از روش خوشه‌بندی پیوندی "وارد" که یکی از رایج‌ترین روش‌های خوشه بندی است استفاده گردید. نمودار درختی (دندروگرام) حاصل از

سری زمانی به وسیله حداقل مربعات معمولی<sup>۱</sup> تخمین زده می‌شود. به عبارت دیگر تخمین معادله ذیل با استفاده از n.T مشاهده انجام می‌شود:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}; i = 1, 2, \dots, n, t = 1, 2, \dots, T \quad (۱)$$

n: تعداد مقاطع (تعداد پیرومترها)، T: طول سری زمانی (طول دوره آماری)، Y: متغیر وابسته (تراز سطح آب زیرزمینی)، X: متغیرهای مستقل،  $\beta$ : ضرایب متغیرهای مستقل و  $\varepsilon$ : جمله اختلال.

### روش اثرات ثابت

روش اثرات ثابت بر این فرض استوار است که  $\alpha_i$  برای واحدهای (پیرومترهای) مختلف متفاوت می‌باشد. در این روش برای هر یک از مقاطع یک عرض از مبدأ ثابت در نظر گرفته می‌شود. در این حالت در هر معادله عرض از مبدأ یک عامل مجهول است که باید تخمین زده شود:

$$Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it} \quad (۲)$$

### روش اثرات تصادفی

در روش اثرات تصادفی فرض می‌شود که تفاوت میان واحدها می‌تواند در جمله اختلال ظاهر شود (Greene, 2003):

$$Y_{it} = \alpha + \beta X_{it} + u_i + \varepsilon_i \quad (۳)$$

$u_i$ : جزء تصادفی که در بر گیرنده عواملی است که توسط متغیرهای مستقل تبیین نمی‌شود و کاملاً تصادفی است.

در مدل‌های اثرات ثابت و تصادفی برای کارایی تخمین از روش برآورد حداقل مربعات تعمیم یافته<sup>۲</sup> و در مدل اثرات مشترک از روش برآورد حداقل مربعات معمولی استفاده می‌شود.

ابتدا الگوی موردنظر به سه روش بالا با استفاده از نرم‌افزار Eviews 3.0 (شیرین بخش و خونساری، ۱۳۸۴) برآورد می‌گردد. پس از آن باید این موضوع بررسی شود که نتایج حاصل از کدامیک از روش‌های فوق برای تجزیه و تحلیل نهایی ملاک عمل قرار گیرد. بدین منظور سه آزمون اساسی چو، LM بروش - پاگان و هاسمن برای انتخاب بین این سه روش انجام می‌پذیرد (Yaffee, 2003). در آزمون چو برای انتخاب بین دو روش اثرات ثابت و اثرات مشترک فرضیه ذیل مورد آزمون قرار می‌گیرد:

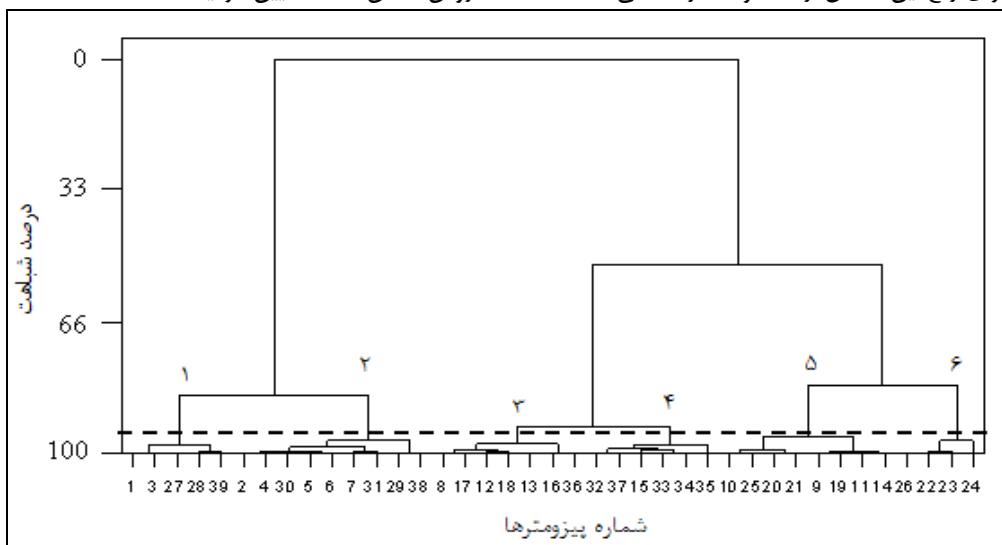
$H_0$ : کلیه واحدها دارای عرض از مبدأ مشترک می‌باشند (روش اثرات مشترک).

$H_1$ : عرض از مبدأ برای واحدهای مختلف متفاوت می‌باشد (روش اثرات ثابت).

1- Ordinary least square

2- Generalized least square

گرفته شد. متعاقباً نمایندگان دو خوشه تغییر یافتند. شکل (۴) پیژومترهای موجود در سطح دشت، خوشه‌های تعیین شده و نماینده هر یک از آنها را نشان می‌دهد. در نقاط کناری دشت به دلیل موجود نبودن پیژومتر و به تبع آن وجود نداشتن اطلاعات مربوط به سطح آب زیرزمینی در این نقاط، عمل خوشه‌بندی امکانپذیر نبوده و استفاده از اطلاعات چاه‌های کشاورزی موجود در نقاط فوق به دلیل دقیق نبودن داده‌های آنها مورد تأیید کارشناسان امور مربوطه نبود. بنابراین نقاط فوق در خوشه‌بندی پیژومترها در نظر گرفته نشد. بعد از انتخاب نماینده‌های معرف هر خوشه، مقادیر متغیرهای مستقل بارندگی و دمای حداکثر، حداقل و متوسط برای هر یک از خوشه‌ها با استفاده از روش عکس فاصله تعیین گردید.



شکل (۳) نمودار درختی حاصل از خوشه‌بندی پیژومترهای موجود در سطح دشت. پیژومترهای ۱- راه سلطان آباد، ۲- فیلخانه، ۳- امان آباد، ۴- اراضی چاه مهندس، ۵- امیرآباد و ۶- جنوب حسین آباد که به ترتیب نماینده هر یک از خوشه‌ها می‌باشند.

به عنوان یک محدوده قابل قبول شناخته می‌شود (Coulibaly et al., 2001). مطابق شکل فوق می‌توان گفت که مقدار خطا در اکثر پیژومترها در دوره آموزش در محدوده قابل قبول (کوچکتر از  $\pm 1$ ) قرار دارند. فقط در پیژومترهای فیلخانه و امیرآباد در یک سال خاص دارای خطای بیش از محدوده قابل قبول مشاهده می‌شود که با مراجعه به آمار سال مذکور در ماه‌های آن سال نوسانات غیرمعمول مشاهده گردید. که این نوسانات می‌تواند ناشی از خطای داده‌برداری در آن سال مذکور باشد. بعلاوه با توجه به موقعیت مکانی ایستگاه امیر آباد (نزدیکی آن به رودخانه) می‌توان فرض نمود که نوسانات غیر معمول پیژومتر در سال مذکور مرتبط با جریان‌های رودخانه‌ای باشند. در نهایت، مطابق جدول (۲) برای هر یک از پیژومترها، معادله جامع پیش‌بینی سطح آب زیرزمینی بر اساس متغیرهای مستقل دمای حداکثر، متوسط و حداقل، بارندگی و سطح آب اولیه ارائه گردید. همانطور که در بالا بیان شد از بین مدل‌های فوق، از ضرایب مربوط

خوشه‌بندی در شکل (۳) نشان داده شده است. با توجه به نمودار فوق برای تصمیم‌گیری در مورد تعداد خوشه‌ها، سطح شباهت ۹۰ درصد به دلیل در نظر گرفتن حداکثر سطح دشت در نظر گرفته شد. بعد از تعیین تعداد خوشه‌ها، مجموع مربعات انحراف از میانگین هر خوشه، برای پیژومترهای موجود در داخل هر خوشه محاسبه و پیژومتری که دارای کمترین مجموع انحرافات می‌باشد به عنوان نماینده آن خوشه انتخاب گردید.

به نظر می‌رسد که پیژومتر انتخابی در بعضی از نواحی انتخاب شده، نماینده شایسته‌ای از خوشه مورد نظر نمی‌باشد. یعنی اینکه پیژومتر انتخابی در مرکز ثقل خوشه نبوده و یا اینکه نزدیک به خوشه مجاور قرار دارد. برای رفع این مشکل از قضاوت کارشناسی کمک

### برآورد مدل و آزمون فرضیه‌ها

همان طور که در قسمت نظریه مدل بیان شد از آزمون‌های چو (آماره F)، LM، بروس - پاگان و هاسمن برای انتخاب بهترین مدل استفاده گردید. مطابق نتایج به دست آمده از دو آماره F و LM، مدل مناسب مبتنی بر الگوی اثرات مشترک بدست آمد بدین صورت که ابتدا با استفاده از آماره F از بین مدل‌های اثرات مشترک و اثرات ثابت، مدل اثرات مشترک انتخاب گردید. دوباره با استفاده از آماره LM از بین مدل‌های اثرات مشترک و اثرات تصادفی، مدل اثرات مشترک انتخاب شد. جدول (۱) مقادیر آماره محاسبه شده و بحرانی را نشان می‌دهد. مقدار خطای حاصل از برآورد مدل برای هر یک از پیژومترهای در نظر گرفته شده در مجموعه داده‌های آموزش (۷۲-۱۳۷۱ تا ۱۳۸۱) در شکل (۵) نشان داده شده است. یادآوری می‌گردد که در پیش‌بینی سطح آب زیرزمینی حداکثر خطای  $\pm 1$  متر

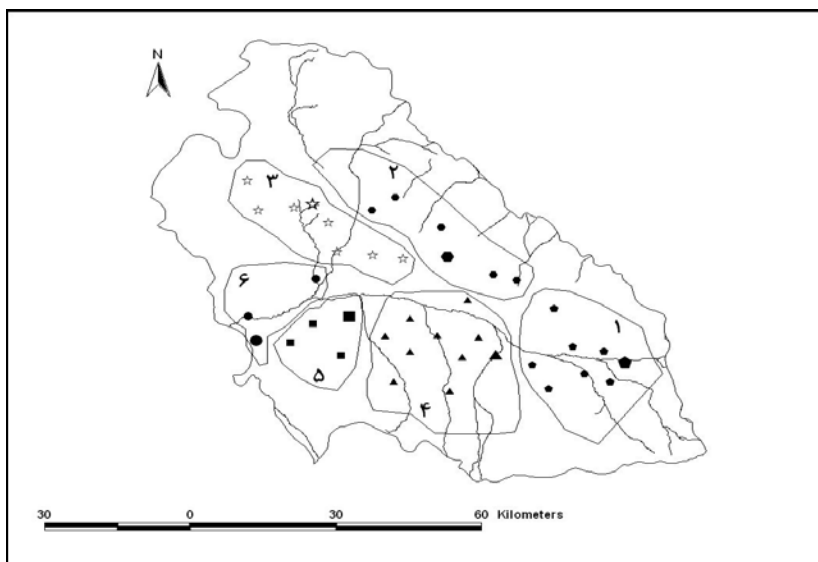
### صحت سنجی مدل

بعد از ارائه معادله جامع برای هر یک از پیژومترها، سطح آب زیرزمینی برای مجموعه آزمون (مهر ۱۳۸۲ تا شهریور ۱۳۸۳) با استفاده از معادلات فوق پیش‌بینی گردید. عملکرد مدل برای هر یک از پیژومترها در جدول (۳) ارائه شده است. مطابق جدول (۳) مدل در دوره آزمون نیز از عملکرد خیلی خوبی برخوردار می‌باشد به طوری که در همه پیژومترها خطا در محدوده قابل قبولی قرار داشته و مدل دارای حداکثر خطای برابر ۰/۴۶ متر برای پیژومتر فیلخانه می‌باشد که مقدار فوق نیز خطای نسبتاً کوچکی است. در شکل (۶) تراز سطح آب پیش‌بینی شده برای هر یک از پیژومترها (نماینده خوشه‌ها) در دوره آزمون نشان داده شده است که نتایج فوق منطبق با نتایج بدست آمده از جدول (۳) می‌باشد. یعنی اینکه پیژومتر امان آباد بهترین عملکرد را به خود اختصاص داده و پیژومتر امیرآباد در جایگاه بعدی قرار گرفته است. بدیهی است که هر یک از این پیژومترها نماینده یک خوشه بوده و سطح آب پیش‌بینی شده بیانگر سطح آب در آن خوشه می‌باشد.

به مدل اثرات مشترک استفاده می‌شود که در روش فوق ضریب ثابت و ضرایب مربوط به متغیرهای مستقل برای همه پیژومترها یکسان می‌باشد. بدیهی است که برای پیش‌بینی سطح آب زیرزمینی در هر یک از پیژومترها باید مقادیر متغیرهای مستقل مربوط به همان پیژومتر در معادله جایگذاری شود. معادله کلی مدل اثرات مشترک عبارت است از:

$$(WTO)_i = a_0 - a_1 \times (T_{MAX})_i + a_2 \times (T_{MEAN})_i - a_3 \times (T_{MIN})_i - a_4 \times (P)_i + a_5 \times (WTi)_i \quad (7)$$

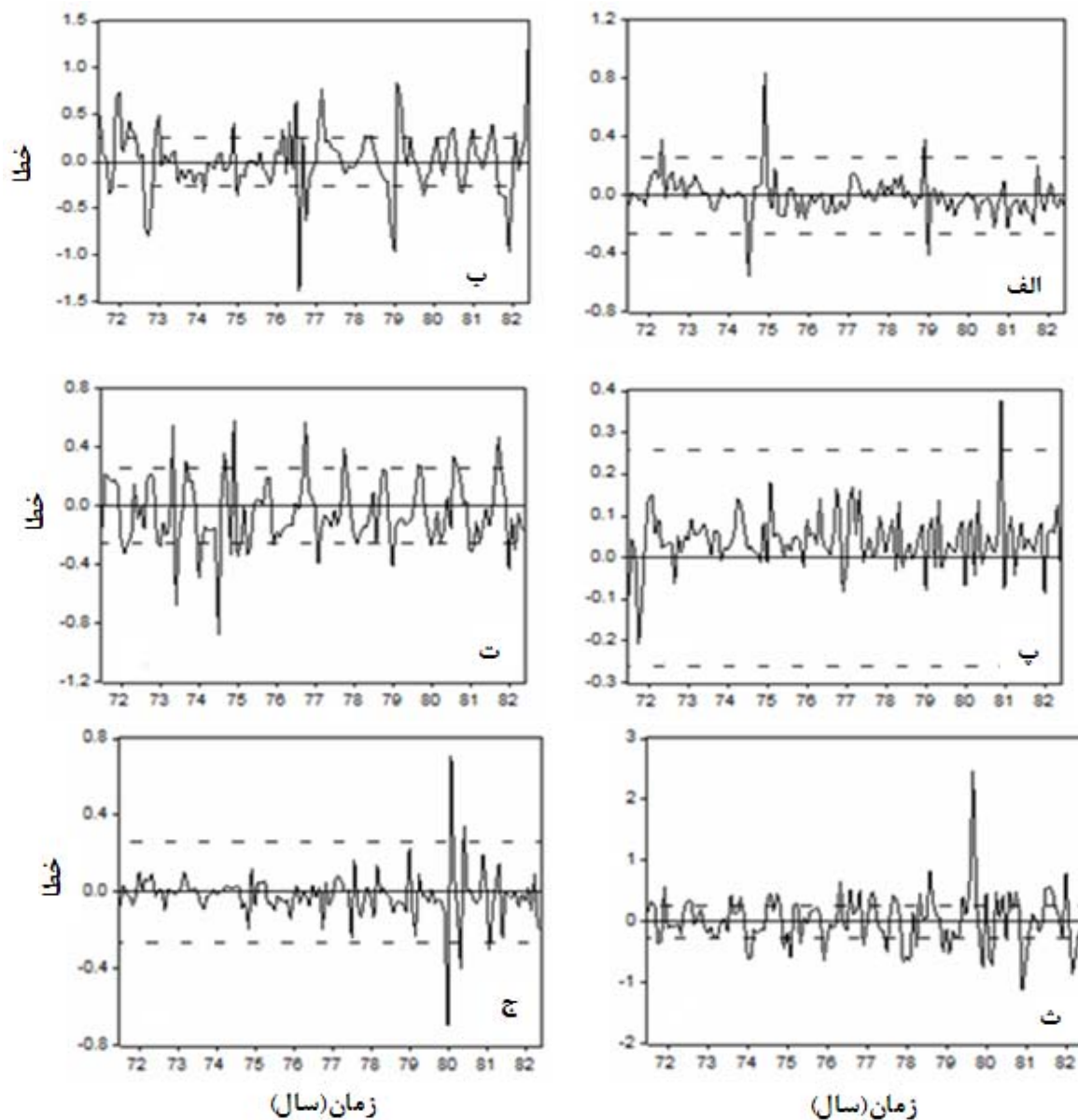
که WTO: تراز سطح آب زیرزمینی پیش‌بینی شده (متر)،  $T_{max}$ : دمای حداکثر (درجه سانتی‌گراد)،  $T_{min}$ : دمای حداقل (درجه سانتی‌گراد)،  $T_{mean}$ : دمای متوسط (درجه سانتی‌گراد)، P: بارندگی (میلی‌متر)، WTi: تراز سطح آب زیرزمینی اولیه (متر) و  $a_0$  تا  $a_5$  ضرایب معادله می‌باشند که می‌توان از جدول (۲) و با توجه به مدل اثرات مشترک تعیین کرد.



شکل (۴) پیژومترهای موجود در سطح دشت به همراه تعداد خوشه‌ها و نماینده خوشه‌ها به ترتیب ۱- راه سلطان آباد، ۲- فیلخانه، ۳- امان آباد، ۴- اراضی چاه مهندس، ۵- امیرآباد و ۶- جنوب حسین آباد

جدول (۱) مقادیر آماره‌های محاسبه شده و بحرانی

آماره	مقدار محاسباتی $\chi^2$ یا توزیع F مقدار بحرانی (توزیع
جو	۳/۰۶
بروش - پاگان LM	۱/۸۴



شکل (۵) مقدار خطای مدل در دوره آموزش برای پیزومترهای (الف) سلطان آباد، (ب) فیلخانه، (پ) امان آباد، (ت) چاه اراضی مهندس، (ث) امیرآباد و (ج) جنوب حسین آباد بترتیب مربوط به خوشه‌های ۱ الی ۶

### مقایسه مدل PD با مدل شبکه عصبی مصنوعی

نتایج مدل داده‌های ترکیبی با نتایج مدل شبکه عصبی مصنوعی مورد مقایسه قرار گرفت. نتایج مدل شبکه عصبی نیز در جدول (۳) و در شکل (۵) نشان داده شده است. مقایسه دو مدل نشان می‌دهد که مدل داده‌های ترکیبی دارای عملکرد نسبتاً بهتری نسبت به مدل شبکه عصبی می‌باشد. البته اختلاف عملکرد دو مدل یاد شده خیلی کم بوده و هر دو مدل عملکرد خیلی خوبی در پیش‌بینی سطح آب زیرزمینی داشته‌اند. قابل ذکر است که مدل داده‌های ترکیبی را به

دلیل سادگی کاربرد آن، پیچیده نبودن تئوری آن نسبت به مدل شبکه عصبی و همچنین در نظر گرفتن چندین پیزومتر به صورت توأم می‌توان برای پیش‌بینی متغیرهای منابع آب به کار برد. جدول (۳) مقادیر معیارهای عملکرد برای دوره آزمون در نظر گرفته شده است.

### نتیجه‌گیری

مدل داده‌های ترکیبی برای اولین مرتبه برای پیش‌بینی سطح آب زیرزمینی به کار رفت و در این تحقیق توانایی مدل‌های فوق به



بقیه پیژومترها نیز در محدوده قابل قبولی قرار داشتند. مقایسه مدل داده‌های ترکیبی با مدل شبکه عصبی مصنوعی نشان می‌دهد که مدل فوق دارای توانایی بالایی در پیش‌بینی سطح آب زیرزمینی می‌باشد. با این وجود با توجه به جدید بودن کاربرد مدل مذکور در مدیریت منابع آب توصیه می‌شود که این مدل فوق مورد آزمون‌های بیشتری قرار گرفته و کاربرد آن در سایر شاخه‌های مدیریت منابع آب نیز بررسی گردد.

عنوان یک ابزار سودمند و البته نسبتاً جدید در مدیریت منابع آب برای پیش‌بینی سطح آب زیرزمینی در دشت نیشابور مورد ارزیابی قرار گرفت. با توجه به انواع مدل‌های در نظر گرفته شده، مدل داده‌های ترکیبی با اثرات مشترک به عنوان بهترین مدل برای پیش‌بینی سطح آب زیرزمینی با توجه به آماره‌های چو و بروش - پاگان انتخاب گردید. مدل فوق دارای عملکرد خیلی خوبی هم در دوره آموزش و هم در دوره آزمون بوده به طوری که در دوره آزمون پیژومتر امان آباد بهترین و پیژومتر فیلخانه بدترین عملکرد را به خود اختصاص دادند و

جدول (۲) ضرایب معادلات مربوط به هر یک از مدل‌های داده‌های ترکیبی

متغیر	روش اثرات مشترک	روش اثرات ثابت	روش اثرات تصادفی
ضریب ثابت کلی معادله	-	-	۰/۰۸۴۷۵
ضریب ثابت پیژومتر فیلخانه	۰/۰۸۶۴۸	۱/۴۰۳۴۹۵	۰/۰۰۹۹۰۹
ضریب ثابت پیژومتر امان آباد	۰/۰۸۶۴۸	۱/۳۸۳۴۳۳	۰/۰۳۲۵۷۹
ضریب ثابت پیژومتر امیر آباد	۰/۰۸۶۴۸	۱/۳۳۱۱۰۴	۰/۰۱۶۹۱۲
ضریب ثابت پیژومتر حسین آباد	۰/۰۸۶۴۸	۱/۲۳۶۶۹۰	-۰/۰۱۴۶۷۹
ضریب ثابت پیژومتر سلطان آباد	۰/۰۸۶۴۸	۱/۴۱۲۶۵۸	-۰/۰۰۴۹۷۴
ضریب ثابت پیژومتر چاه مهندس	۰/۰۸۶۴۸	۱/۳۰۸۱۶۴	-۰/۰۳۹۷۴۷
دمای حداکثر	-۰/۰۲۳۹۸۶	-۰/۰۲۵۳۶۷	-۰/۰۰۴۱۷۳
دمای متوسط	۰/۰۳۸۵۹۱	۰/۰۳۹۸۸۶	- <sup>۱</sup>
دمای حداقل	-۰/۰۱۶۵۵۳	-۰/۰۱۶۳۲۸	۰/۰۰۲۲۶۴
بارندگی	-۵/۸۷×۱۰ <sup>-۵</sup>	-۸/۶۱×۱۰ <sup>-۵</sup>	-۳/۳۶×۱۰ <sup>-۵</sup>
سطح آب اولیه	۰/۹۹۹۹۵	۰/۹۹۸۸۵۴	۰/۹۹۹۹۴۱

<sup>۱</sup> در روش اثرات تصادفی تعداد مقاطع (پیژومترها) باید بیشتر از تعداد متغیرها (مستقل و وابسته) باشد (شیرین بخش و خونساری، ۱۳۸۴)

## سپاسگزاری

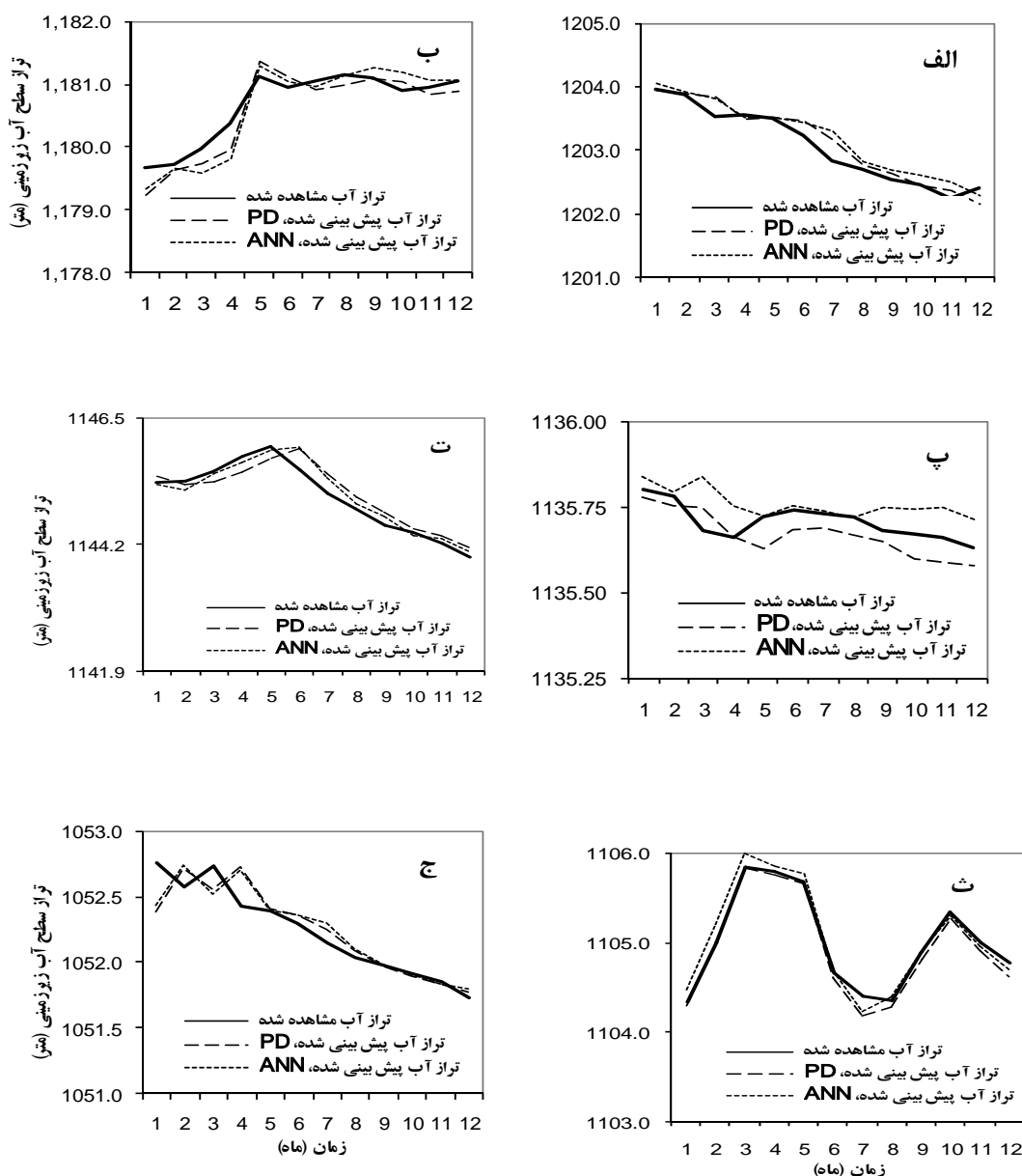
همکاری آقای مهندس حمید طاهرپور دانشجوی کارشناسی ارشد اقتصاد دانشگاه فردوسی مشهد قدردانی می‌نمایند.

بدینوسیله از حمایت‌های دانشگاه فردوسی مشهد (طرح شماره ۳۲۳ پ دانشکده کشاورزی) سپاسگزاری می‌شود همچنین مولفین از

جدول (۳) مقادیر معیارهای عملکرد در نظر گرفته شده برای دوره آزمون

شماره خوشه	نام پیژومتر (نماینده)	RMSE		R <sup>2</sup>		Max Error	
		ANN <sup>1</sup>	PD <sup>1</sup>	ANN	PD	ANN	PD
۱	راه سلطان آباد نمک	۰/۱۷	۰/۱۷	۰/۹۳	۰/۹۴	-۰/۳۳	-۰/۳۳
۲	فیلخانه	۰/۲۳	۰/۲۳	۰/۹۴	۰/۹۵	۰/۴۶	۰/۴۶
۳	امان آباد	۰/۰۵	۰/۰۵	۰/۲۳	۰/۵۹	۰/۰۹	۰/۰۹
۴	اراضی چاه مهندس	۰/۲۲	۰/۲۲	۰/۹۳	۰/۸۸	-۰/۳۸	-۰/۳۸
۵	امیرآباد	۰/۰۹	۰/۰۹	۰/۹۷	۰/۹۹	۰/۲۳	۰/۲۳
۶	جنوب حسین آباد	۰/۱۶	۰/۱۶	۰/۷۹	۰/۷۸	۰/۳۷	۰/۳۷

<sup>۱</sup> ANN و PD بترتیب بیانگر مدل داده‌های ترکیبی و شبکه عصبی می‌باشند.



شکل (۶) مقایسه مقادیر پیش‌بینی شده و مشاهده شده در دوره آزمون برای بیژومتر الف) راه سلطان آباد نمک، ب) فیلیخانه، پ) امان آباد، ت) اراضی چاه مهندس، ث) امیرآباد، ج) جنوب حسین آباد جنگل بترتیب مربوط به خوشه‌های ۱ الی ۶

## مراجع

شرکت سهامی آب منطقه‌ای خراسان رضوی، (۱۳۷۷). گزارش محاسبه بیلان منابع آب حوضه آبریز دشت نیشابور. شیرین بخش، ش. و خونساری، ح.، (۱۳۸۴). "کاربرد Eviews در اقتصاد سنجی". پژوهشکده امور اقتصادی، تهران. طاهری، ص.، (۱۳۸۴). "عوامل موثر بر تولید گندم در ایران به ترتیب ضریب اهمیت (۱۳۷۷-۱۳۸۱)". اقتصاد کشاورزی و توسعه، سال سیزدهم، شماره ۵۰، صفحه ۸۱ تا ۹۴. مهرگان، ن.، عزتی، م. و اصغرپور، ح.، (۱۳۸۵). "بررسی رابطه علی

باشقره، ع. ا.، (۱۳۷۷). "ارزیابی پتانسیل قابل بهره‌برداری از آبهای زیرزمینی دشت مهران به کمک GIS و RS". پایان نامه کارشناسی ارشد، دانشکده علوم انسانی، دانشگاه تربیت مدرس. حسینی، آ.، فرج زاده، م.، ولایتی، س. (۱۳۸۴). "تحلیل بحران آب در نیشابور با رویکرد برنامه ریزی محیطی". کمیته تحقیقات شرکت سهامی آب منطقه‌ای خراسان.

- De Cian, E., Lanzi, E. and Roson, R., (2007). "The Impact of Temperature Change on Energy Demand: A Dynamic Panel Analysis".
- Feddes, R. A., Kabat, P., Van Bakel, P. J. T., Bronswijk, J. J. B. and Halbertsma, J., (1988). "Modeling soil water dynamics in the unsaturated zone-State of art". *Journal of Hydrology*, 100, 69-111.
- Greene, William H., (2003). *Econometric Analysis*; 3<sup>rd</sup> Edition, Prentice Hall, New Jersey.
- Hsiao, C., (2003). *Analysis of Panel Data*; 2<sup>nd</sup> Edition, Cambridge University Press.
- Knotters, M. and Van Walsum, P. E., (1997). "Estimating fluctuation quantities from time series of water table depths using models with a stochastic component". *Journal of Hydrology*, 197, 25-46.
- Maier, H.R. and Dandy, G.C., (2000). "Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications". *Environmental Modeling & Software*, 15, 101-124.
- Nayak, P., Satyaji Rao, Y. R., and Sudheer, K. P., (2006). "Groundwater level forecasting in a shallow aquifer using artificial neural network approach". *Water Resources Management*, 2(1)77-99.
- Sami S. M, Nasr, I. M. and Galal, H. G. (2002), "Evaluation of Groundwater Resources by Using Fuzzy Logic in Wadi Dara Area, Gulf of Suez, Egypt", 2<sup>nd</sup> International Conference on Earth Observation and Environmental Information, Cairo, Egypt.
- Tankersley, C. D., Graham, W. D. and Haltfield, K., (1993). "Comparison of univariate and transfer function models of groundwater fluctuations, *Water Resources Research*, 29(10)3517-3533.
- Van Geer, F. C. and Zuur, A. F., (1997). "An extension of Box-Jenkins transfer/noise from time series of groundwater head series". *Journal of Hydrology*, 192, 65-80.
- Yaffee, R., (2003). "A Primer for Panel Data Analysis". Social Science, Statistics & Mapping Group of ITS' Academic Computing Services, New York University.
- بین نرخ بهره و تورم: با استفاده از داده‌های تابلویی". فصلنامه پژوهش های اقتصادی، سال ششم، شماره سوم، صفحه ۹۱ تا ۱۰۵.
- ولایتی، س. و توسلی، س.، (۱۳۷۰). "منابع و مسائل آب خراسان. موسسه چاپ و انتشارات آستان قدس رضوی، مشهد.
- هادیان، م.، قادری، ح. و معینی، م.، (۱۳۸۵). "تخمین تابع تقاضای بیمه درمان مکمل، مطالعه موردی: شرکت سهامی بیمه ایران". فصلنامه پژوهش های اقتصادی، سال ششم، شماره چهارم، صفحه ۶۷ تا ۸۲.
- Allen, D. M., Schuurman, N. and Zhang, Q., (2007). "Using fuzzy logic for modeling aquifer architecture". *Journal of Geograph. System*, 9, 289-310.
- Baltagi, B., (1995). *Econometric Analysis of Panel Data*, 1<sup>st</sup> Edition, John Wiley & Sons Press, New York, U.S.A.
- Bierkens, M. F. P., (1998). "Modeling water table fluctuations by means of a stochastic differential equation". *Water Resour. Res.*, 34(10)2485-2499.
- Blemans, C., Wesseling, J. G. and Feddes, R. A., (1983). "Simulation of water balance of a cropped soil: SWATRE". *Journal of Hydrology*, 63271-286.
- Coppola, J. M., Szidarovszky, F., Poulton, M. and Charles, E., (2003). "Artificial neural network approach for predicting transient water levels in a multi layered groundwater system under variable state, pumping, and climate conditions". *Journal of Hydrologic Engineering*. 8(6)348-360.
- Coulibaly, P., Anctil, F., Aravena, R. and Bernard B., (2001). "Artificial neural network modeling of water table depth fluctuations". *Journal of Hydrology*, 307(4)92-111.
- Daliakopoulos, I. N., Coulibaly, P. and Tsanis, I. K., (2004). "Groundwater level forecasting using artificial neural networks". *Journal of Hydrology*. 309(4)229-240.
- David, W.S., (1997). *Cluster Analysis: Multivariate Statistics: Concepts, Models, and Applications.*, <http://www.psychstat.missouristate.edu/multibook/mt04.htm> (2008/05/20).

## Application of Panel Data Model in Predicting Groundwater Level

A. Izadi, K. Davari, A. Alizadeh and B. Ghahreman<sup>1</sup>

### Abstract

Groundwater resources are considered as one of the significant and economical water resources. Comprehensive recognition and proper utilization of this valuable resource especially in arid and semiarid areas has an important effect on sustainable development of social and economic activities. It is necessary to predict groundwater level fluctuations for a better understanding of the aquifer behavior in these areas. This research was aimed at prediction of groundwater level in Neyshabour plain using "*Panel Data*" model. The plain was selected due to presence of over 50 observation wells, mostly with more than 12 years of record. Inasmuch as "*Panel Data*" model considers historic data for several observation wells, it was able to predict groundwater levels in different observation wells simultaneously. At the first step, the available observation wells in the plain were clustered with "Ward" method which ended-up in six areal zones. Then, for each cluster, an observation well was selected as its representative, and for each zone, values of independent variables (precipitation, temperature, and initial water table) were estimated by "Distance Inverse method". Finally, the performances of different *Panel Data* regression models such as *Panel Data* with *Common effects*, *Fixed effects* and *Random effects*, for groundwater level prediction were investigated. The results shows *Common effects* model has the best results for prediction of groundwater level. The performance indicators ( $R^2=0.99$  and  $RMSE=0.05$ ) reveals the effectiveness of this method. In addition, these results were compared with the results of an ANN model, which showed relevant superiority of "*Panel Data*" over ANN.

**Key words:** *Panel Data*, groundwater, Ward clustering, Neyshabour plain

---

1 - Contribution from water Engineering Department, College of Agriculture, Ferdowsi University of Mashhad , Iran