

## بهینه‌یابی شبکه پایش سطح آب زیرزمینی در آبخوان ارومیه با استفاده از نقشه خود سازمانده

محمد نخعی<sup>۱\*</sup>، وهاب امیری<sup>۲</sup>، مرضیه امیری<sup>۳</sup>

تاریخ دریافت: ۱۳۹۸/۸/۱۹

تاریخ پذیرش: ۱۳۹۹/۴/۲۶

۱-استاد هیدروژئولوژی، دانشکده علوم زمین، دانشگاه خوارزمی

۲-استادیار هیدروژئولوژی، گروه زمین‌شناسی، دانشکده علوم، دانشگاه یزد

۳-دانش‌آموخته کارشناسی ارشد آب‌شناسی، دانشکده علوم زمین، دانشگاه خوارزمی

\*مسئول مکاتبات، پست الکترونیکی: nakhaei@khu.ac.ir

### چکیده

در این مطالعه از روش نوین نقشه‌های خودسازمانده (SOM) برای خوشه‌بندی پیژومترهای موجود در آبخوان ارومیه و شناسایی پیژومترهایی با رفتار مشابه و به عبارتی دیگر، تعیین شبکه پایش بهینه در این آبخوان استفاده شد. به این منظور، ابتدا به نرمال‌سازی داده‌های ماهیانه تراز آب زیرزمینی در یک بازه ۱۳ ساله (۱۳۸۱-۱۳۹۴) پرداخته شد و سپس الگوریتم خوشه‌بندی غیرسلسله مراتبی K-means برای تعیین تعداد خوشه‌های بهینه مورد استفاده قرار گرفت. در ادامه به کمک مدل SOM که از پرکاربردترین مدل‌های شبکه عصبی در خوشه‌بندی است؛ عملیات پیش‌پردازش مکانی جهت خوشه‌بندی مکانی پیژومترها انجام شد و نتایج خوشه‌بندی با استفاده از نقشه هم عمق سطح آب زیرزمینی مورد بررسی قرار گرفت. پیژومترهای نماینده هر خوشه از طریق به کارگیری فاصله اقلیدسی پیژومترها از یکدیگر تعیین شد. با توجه به نتایج خوشه‌بندی و سازگاری بسیار مناسب آن با تغییرات سطح آب زیرزمینی در محل هر پیژومتر، باید عنوان نمود که الگوریتم خوشه‌بندی SOM از قابلیت بالایی برای خوشه‌بندی برخوردار است. این نوع شبکه عصبی قادر به کشف الگوهای مناسبی از داده‌ها است که می‌تواند به شناسایی ویژگی‌های اعضای هر خوشه کمک کند. بنابراین، به نظر می‌رسد از الگوریتم خوشه‌بندی SOM می‌توان در تحلیل رفتار کمی آبخوان (تغییرات تراز آب زیرزمینی) و یافتن پیژومترهای نماینده معدود استفاده کرده و به این ترتیب امکان ارزیابی قابل قبول رفتار آبخوان در زمان کمی و با صرف کمترین هزینه فراهم آورد.

واژه‌های کلیدی: آب زیرزمینی، خوشه‌بندی، شبکه عصبی، نقشه‌های خودسازمانده، K-means

## Optimization of Groundwater Monitoring Network in Urmia Aquifer Using Self-Organizing Map

M Nakhaei<sup>1\*</sup>, V Amiri<sup>2</sup>, M Amiri<sup>3</sup>

Received: November 26, 2018 Accepted: July 5, 2020

<sup>1</sup>Prof. Faculty of Earth Sciences, Kharazmi University, Tehran, Iran

<sup>2</sup>Assist. Prof. Dept. of Geology, Faculty of Sciences, Yazd University, Yazd, Iran

<sup>3</sup>M.Sc. Graduate, Faculty of Earth Sciences, Kharazmi University, Tehran, Iran

\*Corresponding Author, Email: nakhaei@khu.ac.ir

### Abstract

In this study, the new method of self-organized maps (SOM) was used to cluster the piezometers in the Urmia aquifer and identify piezometers with similar behavior, thereby the optimal groundwater monitoring network was determined. To achieve this goal, the normalization of monthly groundwater level data over a period of 13 years (2002-2015) was done and the K-means non-hierarchical clustering algorithm was then used to determine the number of optimal clusters. Using the SOM model which is one of the most widely used neural network models in the clustering techniques, spatial preprocessing operations were performed for all piezometers and the clusters were detected using a groundwater depth map. Also, the representative piezometers of each cluster were determined using the piezometers Euclidean distance. According to the results of clustering and its good compatibility with changes in groundwater level at each piezometer, it should be noted that the SOM clustering algorithm has a high capability for clustering. Results show that this type of network is able to detect the appropriate data patterns that can help to identify the characteristics of the members of each cluster. Therefore, it seems that the SOM clustering algorithm can be used to analyze the quantitative behavior of the aquifer (groundwater level changes) and to find a few representative piezometers, thus providing an acceptable assessment of aquifer behavior in a short time and at the lowest cost.

**Keywords:** Cluster, Hydrogeology, K-means, Neural network, Self-organized maps

### مقدمه

تعیین الگوی رفتاری نوسانات تراز آب زیرزمینی به روش‌های مختلفی صورت می‌گیرد. در این بین، در سال‌های اخیر استفاده از روش‌هایی همچون مدل‌های شبکه عصبی مصنوعی به عنوان ابزاری قابل اعتماد در تحلیل هیدرولوژیکی آب زیرزمینی گسترش بیشتری پیدا کرده است. مدل‌های شبکه عصبی مصنوعی (ANN)<sup>۱</sup>، مدل‌های داده محوری هستند که با بررسی داده‌های بی‌ربط، غیرخطی و پویا، قابلیت حل مسائل هیدرولوژیکی پیچیده را دارند (نورانی و همکاران

آب زیرزمینی یکی از مهمترین و با ارزش‌ترین منابع آبی هستند. از آنجاکه برنامه‌ریزی‌های کوتاه‌مدت و بلندمدت تأمین آب در هر منطقه بر مبنای حجم منابع آب زیرزمینی و سطحی در دسترس صورت می‌گیرد، پایش مداوم و دقیق نوسانات سطح آب زیرزمینی که تأثیر مستقیمی بر میزان ذخیره آبخوان و متعاقباً تصمیمات بعدی در خصوص بهره‌برداری از منابع آب زیرزمینی دارد؛ بسیار ضروری است.

<sup>1</sup> Artificial neural network

آبخوان و ارائه تحلیل‌های کارشناسی دقیق از وضعیت هیدروژئولوژیکی یک منطقه دارد، شناسایی پیزومترهای نماینده<sup>۲</sup> می‌تواند الگوی نوسان تراز آب زیرزمینی را به بهترین نحو نشان دهد. به این ترتیب، با تعیین گروه‌های پیزومتری مشابه و شناسایی پیزومترهای نماینده هر گروه، هزینه، سرعت و در بسیاری از مواقع، کیفیت برداشت‌های پارامترهای کمی و کیفی در قالب شبکه بهینه شده بهبود پیدا می‌کند. در این مطالعه، آبخوان ارومیه برای بررسی روش انتخاب شده که در بالا به اختصار عنوان شد و در ادامه به تفصیل معرفی می‌شود؛ انتخاب شده است. انتخاب این محدوده به جهت اهمیت آن به عنوان مهمترین منبع آب شیرین در حاشیه دریاچه ارومیه و تأمین کننده بخش اعظمی از نیاز آبی منطقه است. تا کنون و پس از تشدید وضعیت دریاچه ارومیه و کاهش تراز آب آن، مطالعات متعددی در رابطه با برهمکنش منابع آب زیرزمینی و سطحی به انجام رسیده است که از آن جمله می‌توان به امیری و همکاران (۲۰۱۶ a,b,c) و امیری و همکاران (۲۰۱۷) اشاره کرد. با این وجود، سنجش روش‌های شناخته شده و دقیقی جهت ارزیابی هر چه بهتر رفتار کمی این آبخوان می‌تواند در تصمیم‌گیری مسئولان جهت مدیریت بهینه این منابع آبی ارزشمند سودمند واقع شود. بنابراین در این مطالعه سعی خواهد شد از روش نوین نقشه‌های خودسازمانده (SOM) برای خوشه‌بندی پیزومترهای موجود در آبخوان ارومیه و شناسایی پیزومترهایی با رفتار مشابه و به عبارتی دیگر، تعیین شبکه پایش بهینه در این آبخوان استفاده شود.

(۲۰۱۴). این مدل‌ها طیف گسترده‌ای از برنامه‌های کاربردی شامل تحلیل‌های هیدرولوژیکی، پیش‌بینی سیستم‌های غیرخطی و کیفیت آب را از سال ۱۹۹۰ به بعد پوشش می‌دهند (موسوی و همکاران ۲۰۱۳) و اخیراً به صورت گسترده‌ای توسط محققین مورد بررسی قرار گرفته‌اند (آداموسکی و چان ۲۰۱۱، تسای و همکاران ۲۰۱۴، حسینی و نخعی ۲۰۱۵، ودیعتی و همکاران ۲۰۱۶).

با این وجود، چنانچه انتخاب فصل، ویژگی‌های مشخص و کاهش نویز به دقت مورد توجه قرار نگیرد، مدل‌های داده محور ممکن است تخمینی گمراه کننده را ارائه دهند (نورانی و همکاران ۲۰۱۵). برای فائق آمدن بر چنین معضلاتی، روش‌های کمی مبتنی بر پیش‌پردازش یا پس‌پردازش ممکن است یک رویکرد موثر برای غلبه بر کاستی‌ها باشد (هان و همکاران ۲۰۱۶). تکنیک خوشه‌بندی ممکن است به عنوان یک روش پیش‌پردازش داده‌های فضایی جهت بهبود عملکرد مدل‌سازی هیدرولوژیکی مورد استفاده قرار گیرد. نقشه خود سازمانده<sup>۱</sup> (SOM) به عنوان یک تکنیک خوشه‌بندی غیرنظارتی مبتنی بر شبکه عصبی مصنوعی (نورانی و همکاران ۲۰۱۵) قادر به طرح‌ریزی ارتباط داده‌های ورودی بر روی نقشه براساس میزان تشابهات آن‌ها است. درک داده‌های چند بعدی عموماً دشوار است که به کمک این تکنیک می‌توان آن‌ها را بر روی نقشه‌های دو بعدی رسم کرده و تجسم بهتری از آن‌ها در اختیار محقق قرار داد (کاوامیورا و همکاران ۲۰۱۶). تکنیک SOM در بسیاری از زمینه‌های مطالعاتی منابع آب مورد استفاده برخی از محققین علم آب قرار گرفته است (چانگ و همکاران ۲۰۱۶، نورانی و همکاران ۲۰۱۶، هان و همکاران ۲۰۱۶، چن و همکاران ۲۰۱۸).

از آنجاکه رفتارسنجی آبخوان و شبکه پایش کمی، نقش قابل ملاحظه‌ای در ارزیابی‌های کمی و کیفی

<sup>2</sup> Representative piezometers

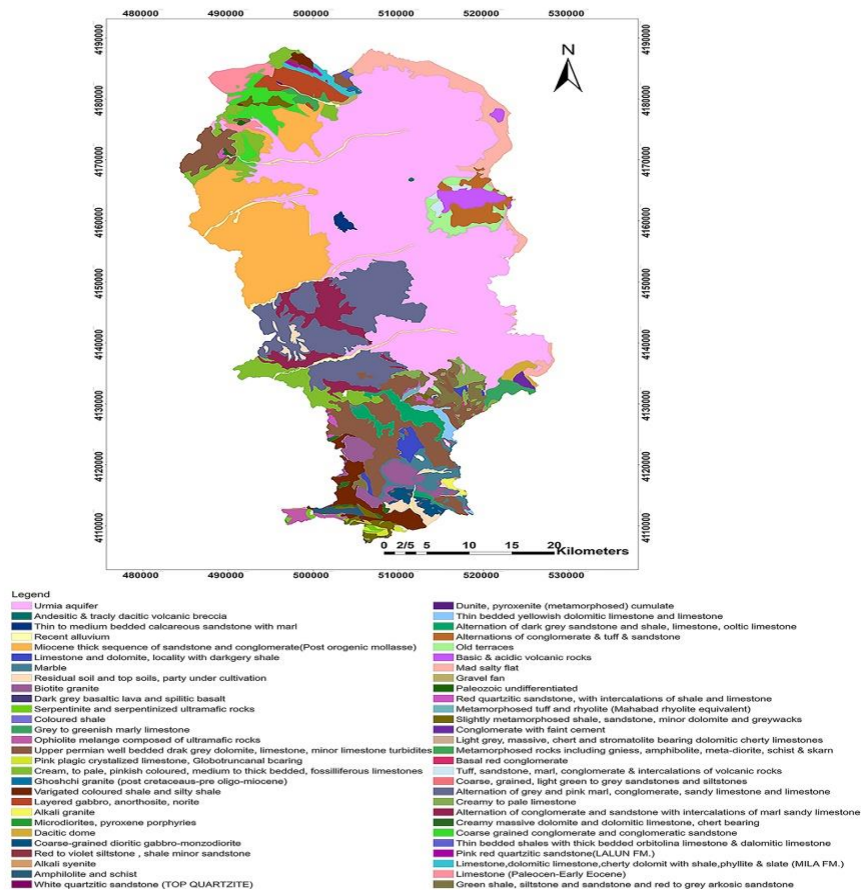
<sup>1</sup> Self organizing map

مواد و روش‌ها

منطقه مورد مطالعه

محدوده مطالعاتی ارومیه (شکل ۱) در استان آذربایجان غربی قرار گرفته و مساحت کلی آن ۲۱۶۶/۲ کیلومتر مربع است که حدود ۴/۱۸ درصد کل حوضه

آبریز دریاچه ارومیه را در بر می‌گیرد. از این میزان ۹۱۷ کیلومترمربع مربوط به وسعت دشت و ۱۲۴۹/۲ کیلومترمربع را نیز ارتفاعات محدوده مطالعاتی تشکیل می‌دهد (بی‌نام ۲۰۱۸).



شکل ۱- نقشه زمین‌شناسی محدوده مطالعاتی ارومیه.

محدوده توسط ایستگاه مرکزی ارومیه نشان می‌دهد که حجم بارش سالانه در حدود ۱۴۳۵۲ میلیون متر مکعب<sup>۱</sup> (MCM) است که از این مقدار، ۱۰۳۶ MCM به محیط زیرسطحی نفوذ کرده و موجب تغذیه منابع آب زیرزمینی می‌شود. علاوه بر این، در حدود ۹۷۷۴ MCM از حجم بارش سالانه نیز تبخیر شده و ۳۵۴۲ MCM باقیمانده به صورت رواناب در سطح این محدوده حرکت کرده و به دریاچه ارومیه ختم می‌شود و یا توسط بخش‌های

آبخوان ارومیه در فروافتادگی واقع در شرق ارتفاعات ارومیه تشکیل شده است. با توجه به رسوبگذاری ناشی از جریانات سیلابی در زمان‌های طولانی و تنوع سازندهای زمین‌شناسی در اطراف این دشت، رسوبات تشکیل‌دهنده آبخوان ارومیه دارای دامنه وسیعی از جنس‌ها (عمدتاً رسوبی و اندکی آذرین و دگرگونی) و اندازه ذرات (عمدتاً ماسه و تا حدودی گراول و رس) است (امیری ۲۰۱۶). بررسی بیلان بلندمدت این

<sup>1</sup> million cubic meters

خوشه‌های بهینه مورد استفاده قرار گرفت. در ادامه تکنیک SOM جهت خوشه‌بندی پیژومترها و فاصله اقلیدسی جهت تعیین پیژومتر نماینده هر خوشه به کار گرفته شد. در ادامه، مراحل ذکر شده به اختصار معرفی شده‌اند.

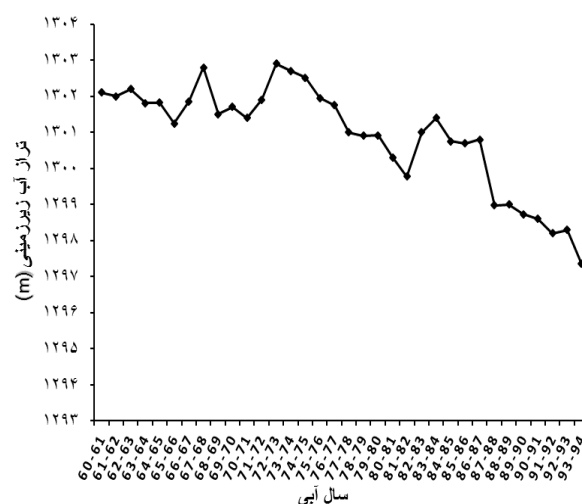
### پیش‌پردازش و نرمال‌سازی داده‌ها

در مرحله پیش‌پردازش داده‌ها، جهت بررسی داده‌های سطح آب زیرزمینی در یک بازه ۱۳ ساله (۱۳۸۱-۱۳۹۴)، از تعداد ۷۱ پیژومتر موجود در دشت، تعداد ۲۲ حلقه پیژومتر به علت نقص و کمبود داده‌ها و عدم پیوستگی، از روند بررسی خارج شد. در نهایت داده‌های ماهیانه مشاهدات سطح آب زیرزمینی مربوط به ۴۹ پیژومتر جهت خوشه‌بندی آماده گردید و ماتریس داده‌ها مرتب شد. از نرم افزار Excel و Notepad++ برای پیش‌پردازش و آماده‌سازی داده‌ها استفاده شد.

نرمال‌سازی برای الگوریتم‌های دسته‌بندی همچون شبکه‌های عصبی یا اندازه‌گیری فاصله همچون دسته‌بندی از طریق نزدیک‌ترین همسایه و خوشه‌بندی مفید است. در این الگوریتم‌ها نرمال‌سازی باعث می‌شود که وقتی داده‌ها برای اندازه‌گیری فاصله به کار می‌روند، داده‌های با مقیاس بزرگ نتیجه را به سمت خود منحرف نکنند (غضنفری و همکاران ۲۰۰۸) و در سازماندهی نقشه خروجی SOM ارزش یکسانی را به خود اختصاص دهند. نرمال‌سازی داده‌ها از طریق تابع  $Som\_normalize$  که در جعبه ابزار<sup>۲</sup> SOM وجود دارد؛ انجام گرفت. این تابع در درون خود از تابع  $Som\_norm\_variable$  جهت مدیریت عملیات نرمال‌سازی استفاده می‌کند. متغیر استفاده شده جهت نرمال‌سازی داده‌ها به طوری که واریانس داده‌ها برابر ۱ باشد،  $variance\ normalization$  است. روش محاسبه فاصله در این متغیر استفاده از

مختلف (مانند کشاورزی و صنعت) مورد استفاده قرار می‌گیرد (بی‌نام ۲۰۱۸).

تعداد منابع آب زیرزمینی واقع در آبخوان آبرفتی دشت ارومیه شامل ۱۷۴۷۴ حلقه چاه مجاز و ۱۵۲۱۹ حلقه چاه غیر مجاز (بعد از سال ۱۳۸۵) می‌باشد. براساس بیلان آب زیرزمینی آبخوان آبرفتی دشت ارومیه، مجموع آب ورودی به آبخوان آبرفتی دشت سالانه ۳۸۱/۵۳ میلیون مترمکعب و مجموع آب خروجی آن سالانه ۳۸۵/۶۳ میلیون مترمکعب است. بنابراین آبخوان آبرفتی دشت ارومیه دارای تغییرات ذخیره ۴/۱ - میلیون مترمکعب در سال است. براساس آبنمود تهیه شده (شکل ۲)، افت متوسط سالیانه درازمدت سطح آب زیرزمینی ۰/۱۵- متر و افت جمعی سطح آب زیرزمینی از سال آبی ۶۱-۶۰ تا ۹۴-۹۳ برابر ۵/۳۶- متر است (بی‌نام ۲۰۱۸).



شکل ۲- آبنمود معرف دشت ارومیه.

در این تحقیق از روش SOM برای خوشه‌بندی و تعیین پیژومترهای نماینده آبخوان ارومیه استفاده شده است. محیط پیاده‌سازی، نرم افزار Matlab R2015a می‌باشد. ابتدا داده‌ها پیش‌پردازش و ماتریس داده‌ها مرتب و نرمال‌سازی شد. سپس الگوریتم خوشه‌بندی غیر سلسله‌مراتبی K-میانگین<sup>۱</sup> جهت تعیین تعداد

<sup>2</sup>Toolbox

<sup>3</sup> Mahalanobis distance

<sup>1</sup> Non-hierarchical k-means clustering

مبنای اصل کمترین شباهت بین خوشه‌ای و بیشترین شباهت درون خوشه‌ای استوار است.

در روابط زیر،  $V$  مرکز خوشه،  $DB$  مقدار نهایی شاخص دیویس-بولدین،  $d$  نشان‌دهنده فاصله خوشه‌ها از یکدیگر،  $S$  نشان‌دهنده پراکندگی داخل خوشه،  $q$  و  $t$  یک عدد صحیح،  $A_i$  مجموعه‌ای از رکوردهایی است که در هر مرحله در خوشه  $i$  قرار می‌گیرد،  $|A_i|$  تعداد عناصر مجموعه  $A_i$  و  $c$  نشان‌دهنده تعداد خوشه‌ها در هر مرحله از محاسبه شاخص است (آقایی ۲۰۱۲). بنابراین داریم:

$$S_{i,q} = \left( \frac{1}{|A_i|} \sum_{x \in A_i} \|x - V_i\|_2^q \right)^{\frac{1}{q}} \quad [2]$$

$$d_{ij,s} = \left\{ \sum_{s=1}^{\rho} |v_{si} - v_{sj}|^t \right\}^{\frac{1}{t}} = \|v_i - v_j\|_t \quad [3]$$

$$R_{i,qt} = \max_{j \in c, j=1} \left\{ \frac{S_{i,q} + S_{j,q}}{d_{ij,s}} \right\} \quad [4]$$

و رابطه نهایی شاخص دیویس-بولدین برای تعداد  $c$  خوشه به صورت زیر است که طبق این رابطه تعداد خوشه بهینه برابر کمترین مقدار  $DB(c)$  است.

$$DB(c) = \frac{1}{c} \sum_{i=1}^c R_{i,q} \quad [5]$$

### نقشه‌های خودسازمانده

نقشه خودسازمانده که گاهی نقشه مشخصه‌های خودسازمانده<sup>۲</sup> و یا نقشه کوهونن<sup>۳</sup> نامیده می‌شود، توسط تئوو کوهونن<sup>۴</sup> ابداع شده است (غضنفری و همکاران ۲۰۰۸). SOM یکی از پرکاربردترین مدل‌های شبکه عصبی رقابتی بدون ناظر است که در خوشه‌بندی و پیش‌پردازش اطلاعات کاربرد دارد (موسوی ۲۰۱۵) و قادر است روابط آماری پیچیده و غیرخطی بین داده‌های با ابعاد بزرگ را به روابط هندسی ساده تبدیل کند و با تولید یک نقشه یک یا دو بعدی باعث تجسم بهتر داده‌های با ابعاد بزرگ شود (کوهونن ۱۹۹۸). از آنجاکه نقشه

فاصله ماهالانوبیس<sup>۱</sup> است. این فاصله معیاری از فاصله هر یک از مشاهدات در فضای چند بعدی از مرکز میانگین تمام مشاهدات است. برتری عمده فاصله ماهالانوبیس نسبت به سایر فاصله‌ها، در نظر گرفته شدن ماتریس کواریانس در آن است (فیلزموسر و همکاران ۲۰۰۵) که موجب می‌شود شکل و اندازه داده‌ها تعیین شود (نوحه گر و همکاران ۲۰۱۸). رابطه محاسبه فاصله ماهالانوبیس هر نمونه به صورت زیر بیان می‌شود (قنادپور و همکاران ۲۰۱۶):

$$D^2 = ([X] - [\bar{X}]][S]^{-1}([X] - [\bar{X}])^T \quad [1]$$

که در آن  $[X]$  بردار داده‌ها برای نمونه مورد نظر،  $[\bar{X}]$  بردار میانگین کل داده‌ها و  $[S]^{-1}$  ماتریس کواریانس است. همان‌طور که مشاهده می‌شود فاصله مورد نظر معادل کسر بردار میانگین داده‌ها از بردار هر یک از داده‌ها تقسیم بر ماتریس  $[S]^{-1}$  است.

### تعیین تعداد خوشه بهینه

الگوریتم خوشه‌بندی Non-hierarchical K-means

مهمترین و پرکاربردترین الگوریتم خوشه‌بندی است. یکی از مهمترین معایب الگوریتم K-means، عدم مشخص بودن K است. ابتدا باید این الگوریتم را برای Kهای مختلف محاسبه کرده و سپس با استفاده از یک شاخص معین، K مناسب جهت خوشه‌بندی بهینه تعیین شود (هان و همکاران ۲۰۱۱). در این تحقیق شاخص دیویس-بولدین به دلیل گستردگی استفاده به کار گرفته شده است. این شاخص در واقع میانگین شباهت بین هر خوشه با شبیه‌ترین خوشه مربوط به آن را محاسبه می‌کند و هر چه مقدار این شاخص کمتر باشد، خوشه‌های بهتری تولید شده است (دهقان و همکاران ۲۰۱۲). انتخاب K بهینه در شاخص دیویس-بولدین بر

<sup>1</sup> Self-organizing feature maps: (SOFM)

<sup>2</sup> Kohonen maps

<sup>3</sup> Teuvo Kohonen

فعلی است. فاصله اقلیدسی با رابطه ۶ داده می شود که در آن  $X$  بردار ورودی فعلی و  $W$  بردار وزن های نورون هاست (غضنفری و همکاران ۲۰۰۸):

$$Dist = \sqrt{\sum_{i=0}^n (X_i - W_i)^2} \quad [6]$$

قدم بعدی پس از تعیین BMU، یافتن همسایگان BMU است. وزن BMU و نورون های همسایه در قدم بعدی تغییر می یابد. برای این کار ابتدا باید شعاع همسایگی محاسبه و سپس با روش فیثاغورث، وجود هر نورون در داخل شعاع تعیین شود. ویژگی منحصر به فرد الگوریتم یادگیری کوهونن، کوچک شدن همسایگی در طی زمان است. این کار با کم کردن شعاع در طول زمان انجام می شود. برای این کار از تابع کاهش نمایی استفاده می شود:

$$\sigma(t) = \sigma_0 e^{-\frac{t}{\lambda}} \quad t = 1, 2, 3... \quad [7]$$

که در آن  $\sigma_0$  نشان دهنده عرض شبکه در زمان  $t_0$  و  $\lambda$  یک ثابت زمانی است.  $t$  گام زمانی فعلی (دور حلقه) می باشد. مقدار  $\lambda$  وابسته به  $\sigma$  و تعداد دور انتخاب شده برای اجرای الگوریتم است. در طی زمان، همسایگی به کوچکی یک نورون یعنی همان BMU می شود. وقتی نورونی در همسایگی پیدا می شود آنگاه بردار وزن های آن از طریق رابطه زیر اصلاح می شود:

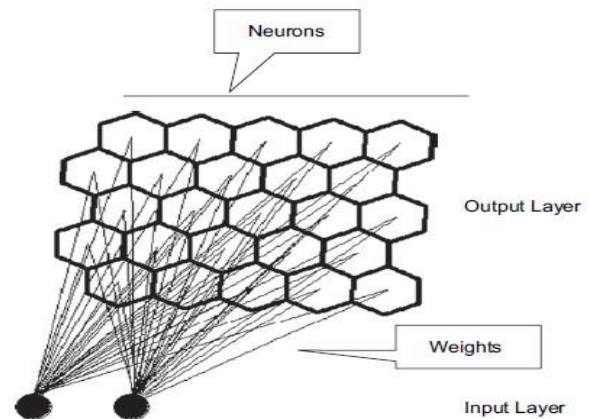
$$W(t+1) = W(t) + L(t)(X(t) - W(t)) \quad [8]$$

که در آن  $t$  نشانگر گام زمانی و  $L$  متغیر کوچکی به نام نرخ یادگیری است که در طول زمان کم می شود. در واقع این رابطه بیان می کند که وزن اصلاح شده جدید برابر وزن قدیمی ( $W$ ) به اضافه بخشی ( $L$ ) از تفاوت بین وزن قدیمی و بردار ورودی ( $X$ ) است. کاهش نرخ یادگیری در هر دور از طریق رابطه ۹ انجام می شود:

$$L(t) = L_0 e^{-\frac{t}{\lambda}} \quad t = 1, 2, 3... \quad [9]$$

در ابتدا نرخ یادگیری مقدار ثابتی مثل ۰/۱ است و به تدریج در طول زمان به صفر میل می کند. نه تنها نرخ

تولید شده توسط SOM معمولاً قادر به نشان دادن شباهت های بین داده هاست، بنابراین مجموعه داده های مشابه می توانند در یک گروه قرار گیرند تا به این طریق طبقه بندی های مختلف با الگوهای ورودی مشابه حاصل شود (کوهونن ۱۹۹۷). شبکه از نورون هایی ایجاد می شود که هر یک به طور کامل به لایه ورودی وصل شده اند (شکل ۳). هر نورون دارای موقعیت مکانی مشخصی بوده و دارای برداری از وزن ها با همان ابعاد بردارهای ورودی می باشد. اگر داده های آموزشی دارای بردارهای  $X$  با  $n$  بعد باشند:  $X_1, X_2, X_3, \dots, X_n$  آنگاه هر نورون دارای بردارهای وزن  $W$  با  $n$  بعد خواهد بود:  $W_1, W_2, W_3, \dots, W_n$ .



شکل ۳- ساختار یک SOM دوبعدی ۵\*۵ (کالته ۲۰۰۸).

الگوریتم یادگیری SOM بدین صورت است که قبل از آموزش، هر نورون باید وزندهی اولیه شود. معمولاً مقادیر تصادفی کوچکی به این وزن ها تخصیص داده می شود. وزن های اولیه در SOM معمولاً بین ۰ و ۱ مقداردهی می شوند ( $0 < W < 1$ ). سپس برداری از داده های آموزشی به تصادف انتخاب و به شبکه داده می شود. در قدم بعدی هر نورون بررسی می شود تا نورونی که دارای مشابه ترین وزن ها به بردار ورودی است پیدا شود. نورون برنده معمولاً به عنوان بهترین واحد انطباق یا  $BMU^1$  شناخته می شود. یک راه برای تعیین BMU، جستجوی همه نورون ها و محاسبه فاصله اقلیدسی بین بردار وزن های هر نورون و بردار ورودی

<sup>1</sup> Best matching unit

داده‌های ماهیانه مشاهدات سطح آب زیرزمینی در ۴۹ پیزومتر طی یک بازه ۱۳ ساله آبی (۱۳۸۱-۱۳۹۴)، جهت خوشه‌بندی آماده گردید و به صورت یک ماتریس  $۴۹ \times ۱۵۶$  مرتب شد. در این ماتریس هر پیزومتر معرف یک نمونه و هر نمونه دارای ۱۵۶ ویژگی است. ویژگی‌ها معرف مشاهدات سطح آب زیرزمینی مربوط به سال‌های آبی و ماه‌های سال هستند. پس از آماده‌سازی داده‌ها و نرمال کردن آن‌ها از طریق تابع `Som_normalize`، نوبت به انتخاب ابعاد شبکه و آموزش نقشه می‌رسد. هر نورون در شبکه می‌تواند حداقل یک نمونه را به خود اختصاص دهد و معرف یک خوشه باشد. بنابراین با توجه به تعداد پیزومترها می‌توان ابعاد شبکه را  $۷ \times ۷$  انتخاب کرد اما از آنجایی که بعضی نورون‌ها هرگز نورون برنده نیستند و نمی‌توانند معرف خوشه باشند، بنابراین بهتر است تعداد نورون‌های شبکه بیشتر از تعداد نمونه‌ها در نظر گرفته شود. ابعاد شبکه به دلخواه  $۸ \times ۷$  انتخاب شد. جهت مقداردهی و آموزش یک SOM، الگوریتم‌های مختلفی در جعبه ابزار SOM وجود دارد که برخی از آن‌ها صرفاً برای مقداردهی به شبکه، برخی برای آموزش و برخی دیگر برای مقداردهی و آموزش بکار گرفته می‌شوند. در این مطالعه از الگوریتم (`SOM_MAKE`) که می‌تواند برای دو هدف مقداردهی و آموزش نقشه بکار گرفته شود استفاده شد. شکل ۴ تعداد پیزومترهایی که هر نورون به خود اختصاص داده است را نشان می‌دهد. این تصویر نمایانگر این است که پیزومترهایی که الگوی نوسان مشابهی دارند توسط نورون یکسان گروه‌بندی شده‌اند و نورون‌هایی که هیچ پیزومتری به آنها اختصاص پیدا نکرده است با عدد صفر کدگذاری شده‌اند.

یادگیری باید در طول زمان کاهش یابد بلکه اثر یادگیری نیز باید متناسب با فاصله یک نورون از BMU باشد. در واقع در لبه‌های بیرونی همسایگی BMU، مقدار یادگیری بسیار ناچیز است. برای دستیابی به این هدف، رابطه اصلاح وزن‌ها طبق رابطه زیر باید کمی تغییر یابد:

$$W(t+1) = W(t) + \theta(t)L(t)(X(t) - W(t)) \quad [۱۰]$$

$\theta$  نمایانگر مقدار تاثیر فاصله یک نورون از BMU

روی یادگیری آن است که در طول زمان کاهش می‌یابد و با رابطه ۱۱ بیان می‌شود که در آن  $dist$ ، فاصله نورون از BMU و  $\sigma$  عرض تابع همسایگی است:

$$\theta(t) = e^{\left(\frac{-dist^2}{2\sigma^2(t)}\right)} \quad t = 1, 2, 3, \dots \quad [۱۱]$$

### تعیین پیزومترهای نماینده

هر خوشه دارای مرکزی است که این مرکز به عنوان نماینده خوشه مطرح می‌شود. نماینده خوشه نمادی از اعضای خوشه است و در جایی از خوشه قرار می‌گیرد که فاصله اقلیدسی بین اعضای خوشه تا این نماینده کمترین مقدار ممکن است. هر چه این فاصله کمتر باشد، در نتیجه اعضای خوشه بیشترین شباهت<sup>۱</sup> را به نماینده دارند و انتخاب مرکز خوشه به عنوان نماینده خوشه به درستی انجام گرفته است. چون مرکز خوشه‌ها از قبل مشخص نیست، در هر مرحله یک پیزومتر به عنوان مرکز خوشه در نظر گرفته می‌شود. فاصله دیگر پیزومترها از این پیزومتر محاسبه می‌شود و در نهایت عددی که برای هر پیزومتر محاسبه می‌شود بیانگر مقدار فاصله ای است که این پیزومتر تا مرکز خوشه دارد. کمترین مقدار فاصله نشان‌دهنده بیشترین شباهت اعضای خوشه و در نتیجه انتخاب صحیح نماینده است.

### نتایج و بحث

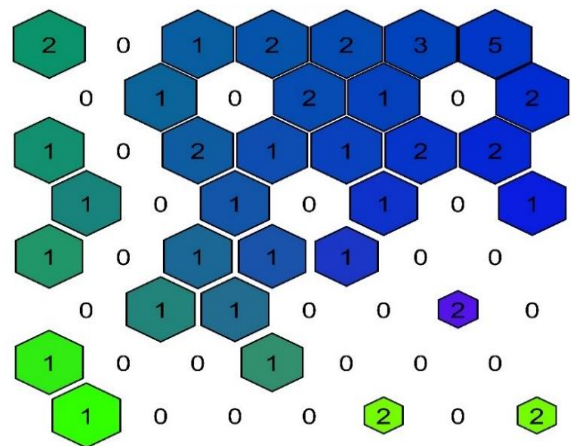
#### آماده‌سازی و خوشه‌بندی پیزومترها

<sup>2</sup> Neighbor weight distances

<sup>1</sup> Similarity

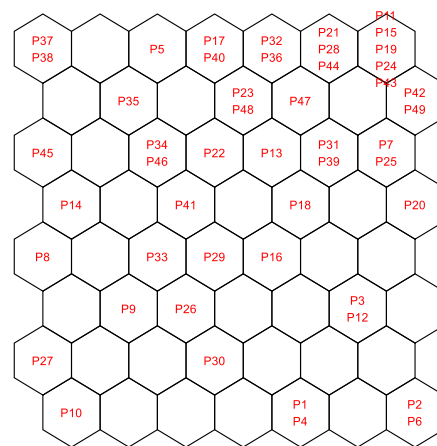
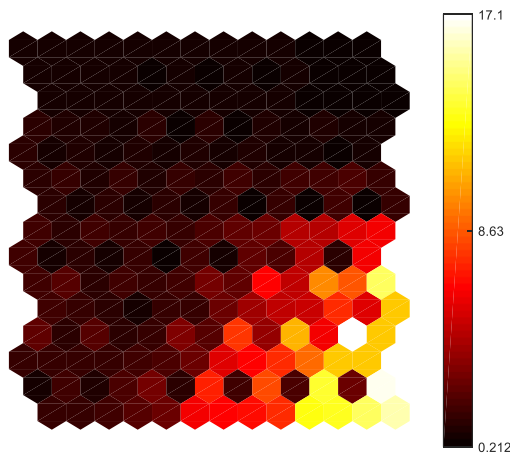


نورون‌های برنده هستند که یک یا چند پیزومتر به آن‌ها اختصاص پیدا کرده است. از آنجایی که تعداد این BMU ها بیانگر تعداد خوشه‌هایی است که در گام اول شناسایی شده است (هان و همکاران ۲۰۱۶) اما تعداد بهینه خوشه‌ها باید لزوماً تعیین شود. به وسیله شکل سمت چپ که فواصل بین نورون‌های همسایه حاصل شده به وسیله لایه خروجی SOM دوبعدی را نشان می‌دهد می‌توان به ساختار خوشه‌ای نقشه پی برد. مقادیر قرار گرفته در بالای نوار رنگی نشان دهنده فاصله زیاد بین نورون‌های همسایه است و بنابراین مرز خوشه‌ها را نشان می‌دهد. با توجه به این شکل می‌توان متوجه شد که تعداد بهینه خوشه‌ها می‌تواند ۴ باشد.



شکل ۴-تعداد پیزومترهای اختصاص داده شده به هر نورون.

شکل ۵، فواصل وزنی همسایه<sup>۱</sup> در نقشه خروجی و همچنین تعداد BMU ها را نشان می‌دهد. BMU ها همان



شکل ۵- (سمت راست) تعداد BMU ها و (سمت چپ) فواصل وزنی همسایه را نشان می‌دهد.

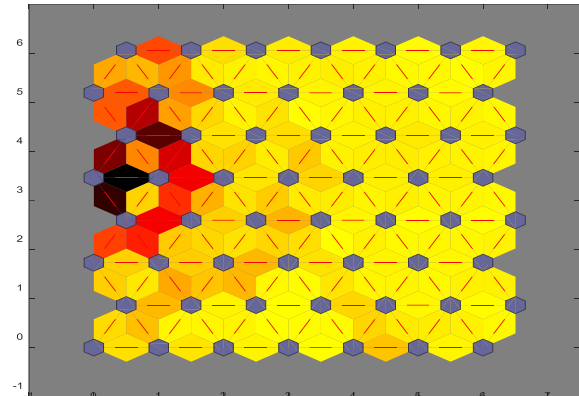
داده شده به آن‌ها کمتر است، در نتیجه پیزومترهای این نورون‌ها در گروه‌های مختلفی قرار داده می‌شوند. برعکس هر چه رنگ‌ها روشن‌تر باشند، فاصله بین نورون‌ها کمتر و شباهت بین الگوی نوسانات پیزومترهای اختصاص داده شده به آن‌ها بیشتر است. بنابراین پیزومترهای این نورون‌ها در گروه‌های یکسان

شکل ۶ نیز نمایش دیگری از فواصل وزنی همسایه است که توسط SOM دوبعدی به دست آمده است. هگزگونال‌های آبی نورون‌های تشکیل دهنده ابعاد شبکه و رنگ‌ها نشانگر فواصل بین این نورون‌ها هستند. هر چه رنگ‌ها تیره‌تر باشند، فاصله بین نورون‌ها زیاده‌تر و شباهت بین الگوی نوسانات پیزومترهای اختصاص

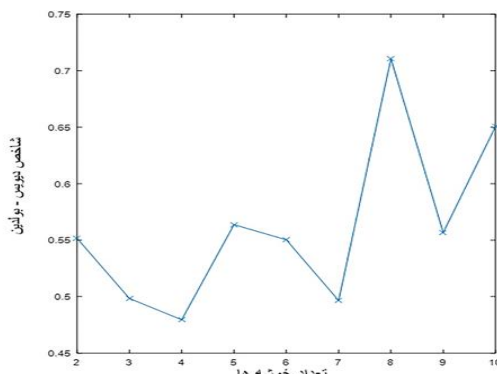
### تعیین تعداد خوشه بهینه

الگوریتم K-means با مقادیر K، از ۲ تا ۱۰ بر روی مجموعه داده‌ها اجرا شد. بقیه مقادیر الگوریتم همان مقادیر پیش‌فرض است و تمام ویژگی‌های نمونه‌های مجموعه داده در خوشه‌بندی به عنوان ورودی انتخاب شدند. در شکل ۷ سمت چپ، محور افقی تعداد خوشه‌ها یا به عبارتی مقدار K و محور عمودی، مقدار شاخص دیویس-بولدین را نشان می‌دهد. هر چه مقدار این معیار کوچک‌تر باشد، بیانگر مجموعه خوشه‌های مترادفی است که به خوبی از هم جدا شده‌اند (قهرمان ۲۰۱۴). مقادیر شاخص دیویس-بولدین برای تعداد خوشه‌های مختلف را نیز می‌توان در جدول ۱ مشاهده کرد. با توجه به شکل ۷ و جدول ۱ می‌توان نتیجه گرفت که تعداد ۴ خوشه بهینه برای این دسته از پیزومترها مناسب است. برای رسیدن به شبیه‌سازی‌های قوی، ۲۰۰ اجرا توسط این الگوریتم خوشه‌بندی صورت گرفت و بر این اساس، فراوانی‌ها برای ۱ تا ۱۰ خوشه مورد انتظار محاسبه شد (در نظر گرفتن ۲۰۰ اجرا به این دلیل می‌باشد که از این عدد به بعد در فراوانی هر خوشه تغییری حاصل نمی‌شود). با توجه به بیشترین فراوانی مرتبط به هر خوشه می‌توان تعداد خوشه بهینه را مشخص کرد و همان طور که در نقشه فواصل وزنی همسایه نیز نمایان بود، تعداد ۴ خوشه بهینه تعیین شد (شکل ۸).

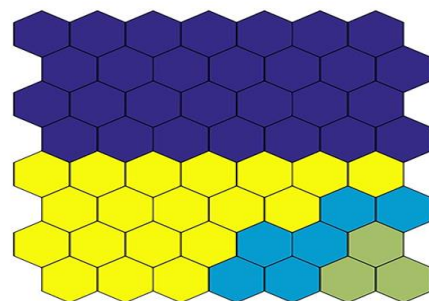
قرار داده می‌شوند (نورانی و همکاران ۲۰۱۵). طبق این شکل، هگزاگونال‌های تیره لایه خروجی را تقریباً به ۴ قسمت تقسیم می‌کنند.



شکل ۶- فواصل وزنی همسایه.

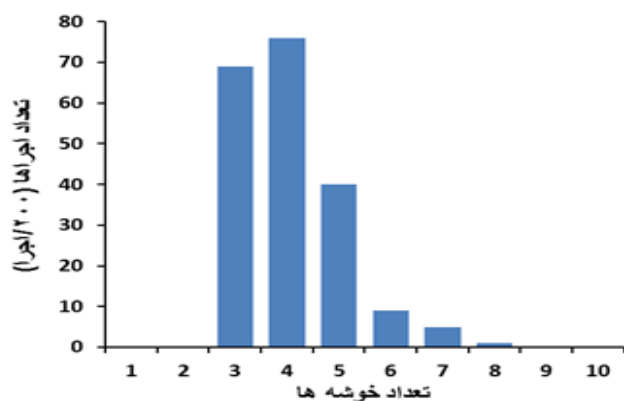


شکل ۷- (سمت راست) تقسیم شدن اندازه لایه خروجی به ۴ ناحیه که هر ناحیه نشان‌دهنده یک خوشه است. (سمت چپ) شاخص دیویس-بولدین و تعداد بهینه خوشه‌ها.



جدول ۱- مقادیر شاخص دیویس-بولدین برای تعداد خوشه‌های مختلف.

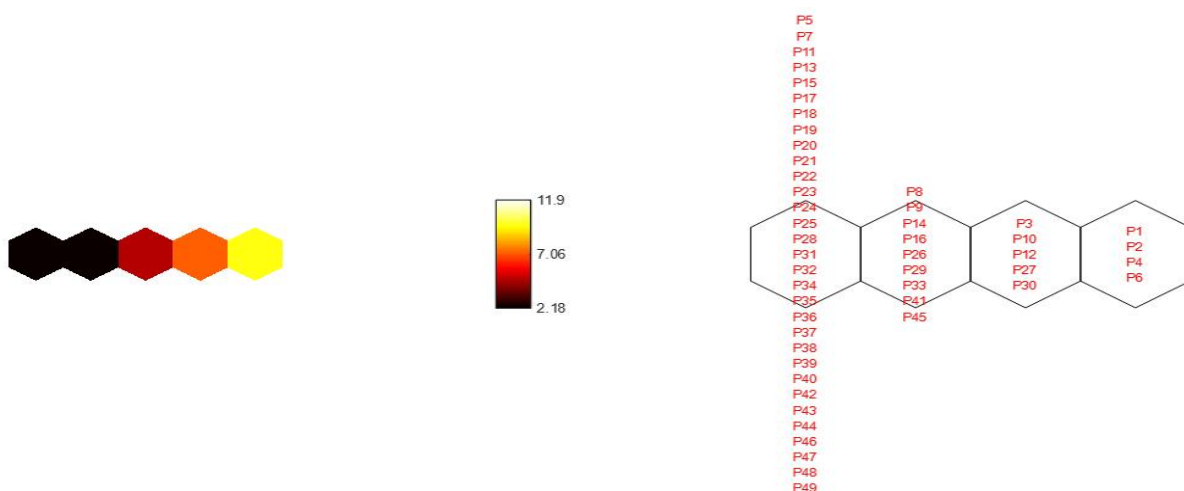
۱۰	۹	۸	۷	۶	۵	۴	۳	۲	تعداد خوشه
۰/۶۵۰۶	۰/۵۵۶۷	۰/۷۱۰۶	۰/۴۹۶۶	۰/۵۵۰۳	۰/۵۶۳۷	۰/۴۷۹۶	۰/۴۹۸۴	۰/۵۵۱۶	شاخص دیویس-بولدین



شکل ۸- نمودار تعداد بهینه خوشه‌ها.

شده است. با توجه به این شکل، از آنجایی که همه نوروں‌های لایه خروجی SOM برنده هستند و هیچ نوروں بازنده‌ای وجود ندارد می‌توان نتیجه گرفت که تعداد بهینه خوشه به درستی تعیین گردیده است. پس از تعیین صحیح تعداد بهینه خوشه‌ها و مشخص شدن پیژومترهای اختصاص داده شده به هر خوشه (جدول ۲)، نقشه موقعیت قرارگیری این ۴ خوشه بهینه تعیین شده از پیژومترهای انتخابی در منطقه مورد مطالعه تهیه شد (شکل ۱۰).

روش دیگری که می‌توان جهت اطمینان از دستیابی به تعداد خوشه‌های بهینه بکار برد، استفاده از یک SOM یک بعدی است. در این مرحله (که تحت عنوان مرحله دوم نامگذاری می‌شود)، یک SOM یک بعدی جهت طبقه‌بندی پیژومترها به تعداد خوشه‌های بهینه‌ای که در گام اول تعیین شد به کار برده می‌شود (نقشه‌ای به ابعاد ۴×۱ که ۴ نشانگر خوشه و ۱ نشانگر یک بعدی بودن نقشه است). فواصل وزنی همسایه و پیژومترهای اختصاص داده شده به هر خوشه در شکل ۹ نشان داده

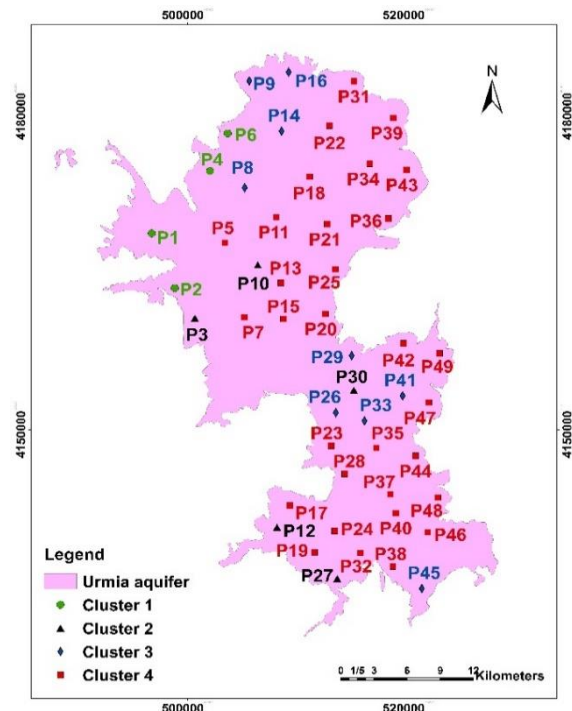


شکل ۹- (سمت راست) پیژومترهای اختصاص داده شده به هر خوشه. (سمت چپ) فواصل وزنی همسایه.

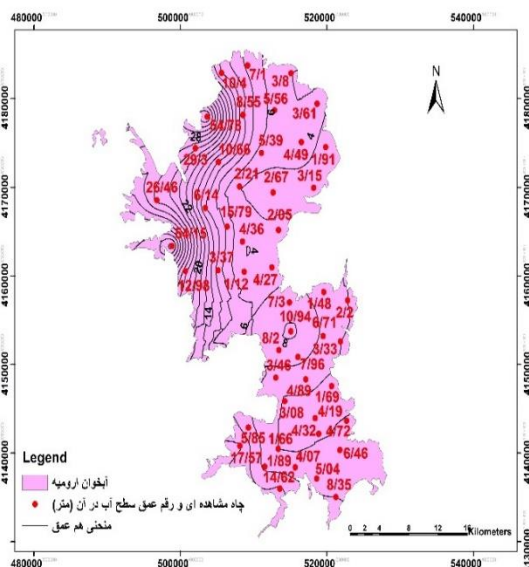
جدول ۲- اعضای هر خوشه.

اعضای هر خوشه	تعداد خوشه
P1, P2, P4, P6	1
P3, P10, P12, P27, P30	2
P8, P9, P14, P16, P26, P29, P33, P41, P45	3
P5, P7, P11, P13, P15, P17, P18, P19, P20, P21, P22, P23, P24, P25, P28, P31, P32, P34, P35, P36, P37, P38, P39, P40, P42, P43, P44, P46, P47, P48, P49	4

هم عمق می‌توان دلیل قرارگیری پیژومترها در هر خوشه را از نظر عمق برخورد به سطح آب زیرزمینی (عمق برخورد به سطح آب نشان‌دهنده عمق سطح ایستابی از سطح زمین است)، زمان نفوذ آب به آبخوان و میزان تبخیر و تعرق مورد بررسی قرار داد. بدین صورت که هر چه عمق برخورد به سطح آب زیرزمینی بیشتر باشد، تبخیر و تعرق کمتر و زمان نفوذ آب به داخل آبخوان بیشتر است. جدول ۳ عمق برخورد به سطح آب زیرزمینی را در پیژومترهای هر خوشه نشان می‌دهد. با توجه به این جدول، اعضاء خوشه یک در عمق برخورد به سطح آب بین ۲۶/۴۶ تا ۵۴/۷۸ متر، خوشه دو در عمق بین ۱۰/۹۴ تا ۱۷/۵۷ متر، خوشه سه در عمق ۶/۷۱ تا ۱۰/۶۶ متر و خوشه چهار در عمق ۱/۱۲ تا ۶/۴۶ متری واقع شده‌اند. در جدول ۴ نیز خوشه‌ها از نظر عمق برخورد به سطح آب زیرزمینی، زمان نفوذ آب به آبخوان و میزان تبخیر و تعرق با یکدیگر مقایسه شده‌اند.



شکل ۱۰- نقشه موقعیت ۴ خوشه بهینه تعیین شده از پیژومترهای انتخابی در محدوده آبخوان ارومیه.



### بررسی وضعیت کلی هر خوشه

از ۴ خوشه بهینه تعیین شده می‌توان نتیجه گرفت که الگوی متفاوت نوسانات تراز آب زیرزمینی در دشت ارومیه وجود دارد. تهیه نقشه هم عمق سطح آب زیرزمینی می‌تواند عامل مهمی در تعیین ارتباط پیژومترها با یکدیگر و قرارگیری آن‌ها در خوشه‌های مختلف باشد (شکل ۱۱). طبق نقشه هم عمق سطح آب زیرزمینی، در دشت ارومیه عمق برخورد به سطح آب زیرزمینی از غرب به شرق کاهش می‌یابد و قسمت عمده دشت تحت پوشش منحنی ۴ متر می‌باشد. از طریق نقشه

شکل ۱۱- نقشه هم عمق سطح آب زیرزمینی: تراز رسم شده مربوط به فصل تر سال آبی ۹۴-۹۳ می‌باشد (به عنوان مثال).

جدول ۳- عمق برخورد به سطح آب زیرزمینی در پیزومترهای هر خوشه.

خوشه	اعضای هر خوشه	عمق برخورد به سطح آب زیرزمینی
1	P1, P2, P4, P6	26.46, 54.15, 29.3, 54.78
2	P3, P10, P12, P27, P30	12.98, 15.79, 17.57, 14.62, 10.94
3	P8, P9, P14, P16, P26, P29, P33, P41, P45	10.66, 10.4, 8.55, 7.1, 8.2, 7.3, 7.96, 6.71, 8.35
4	P5, P7, P11, P13, P15, P17, P18, P19, P20, P21, P22, P23, P24, P25, P28, P31, P32, P34, P35, P36, P37, P38, P39, P40, P42, P43, P44, P46, P47, P48, P49	6.14, 3.37, 2.21, 4.36, 1.12, 5.85, 5.39, 1.89, 4.27, 2.67, 5.56, 3.46, 1.66, 2.05, 3.08, 3.8, 4.07, 4.49, 4.89, 3.15, 4.19, 5.04, 3.61, 4.32, 1.48, 1.91, 1.69, 6.46, 3.33, 4.72, 2.2

جدول ۴- مقایسه عمق برخورد، زمان نفوذ آب و میزان تبخیر و تعرق در خوشه‌ها.

خوشه	عمق برخورد به سطح آب زیرزمینی	زمان نفوذ آب به آبخوان	میزان تبخیر و تعرق
یک	خیلی زیاد	خیلی زیاد	خیلی کم
دو	زیاد	زیاد	متوسط
سه	متوسط	متوسط	زیاد
چهار	کم	کم	خیلی زیاد

## تعیین پیزومترهای نماینده

پس از طبقه‌بندی پیزومترها نوبت به شناسایی پیزومترهای نماینده در هر خوشه می‌رسد. با توجه به اعضای هر خوشه و نحوه تعیین پیزومترهای نماینده که پیش‌تر ذکر گردید، فاصله هر عضو تا مرکز خوشه و تعیین پیزومتر نماینده در جدول ۵ مشخص گردیده

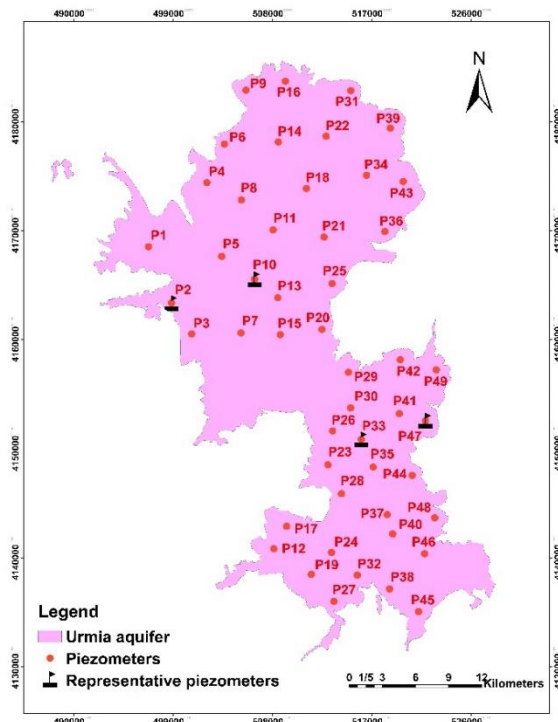
است. همچنین موقعیت قرارگیری پیزومترهای نماینده در روی نقشه نشان داده شده است (شکل ۱۲). پس از مشخص شدن پیزومترهای نماینده، آن‌ها را می‌توان جهت کاربردهای مختلف به خصوص مدل‌سازی سطح آب زیرزمینی به کار برد.

جدول ۵- فاصله هر عضو تا مرکز خوشه و انتخاب پیزومتر نماینده.

خوشه	فاصله هر پیزومتر تا مرکز خوشه	پیزومتر نماینده
1	P1(982.30), P2(953.40), P4(977.64), P6(1091.80)	P2
2	P3(250.84), P10(220.59), P12(264.28), P27(241.29), P30(236.86)	p10
3	P8(227.47), P9(254.30), P14(178.63), P16(233.02), P26(212.93), P29(180.93), P33(172.79), P41(233.72), P45(265.11)	p33
4	P5(583.08), P7(571.06), P11(550.33), P13(579.76), P15(832.02), P17(585.17), P18(723.79), P19(585.08), P20(650.48), P21(500.80), P22(622.13), P23(514.62), P24(614.12), P25(597.03), P28(584.39), P31(461.15), P32(496.63), P34(759.36), P35(732.31), P36(534.29), P37(697.37), P38(651.25), P39(498.62), P40(512.95), P42(553.93), P43(582.71), P44(543.79), P46(785.74), P47(455.50), P48(497.76), P49(530.41)	P47

### نتیجه‌گیری کلی

تعداد و تراکم بسیار بالای چاه‌های بهره‌برداری در سطح آبخوان ارومیه در کنار پیزومترهای موجود در سطح این آبخوان تهیه اطلاعات کمی و کیفی فراوانی را فراهم می‌آورد. از طرفی دیگر، تهیه اطلاعات دوره‌ای از این موقعیت‌ها نیازمند صرف هزینه و زمان بسیاری است که به همین دلیل بهینه‌سازی شبکه پایش در این آبخوان را به امری ضروری تبدیل کرده است. در همین راستا، نتایج بدست آمده از بکارگیری الگوریتم خوشه‌بندی SOM برای بهینه‌سازی شبکه پایش کمی در آبخوان ارومیه نشان می‌دهد که در این روش، تطابق بالایی بین خوشه‌های شناسایی شده و تغییرات سطح آب زیرزمینی در محل هر پیزومتر وجود دارد. این نوع شبکه عصبی قادر به کشف الگوهای مناسبی از داده‌ها است که می‌تواند به شناسایی ویژگی‌های اعضای هر خوشه کمک کند. بنابراین، به نظر می‌رسد از الگوریتم خوشه‌بندی SOM می‌توان در تحلیل رفتار کمی آبخوان (تغییرات تراز آب زیرزمینی) و یافتن پیزومترهای نماینده معدود استفاده کرده و به این ترتیب امکان ارزیابی قابل قبول رفتار آبخوان در زمان کمی و با صرف کمترین هزینه فراهم آورد.



شکل ۱۲- نقشه موقعیت پیزومترهای نماینده.

### منابع مورد استفاده

Adamowski J and Chan HF, 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology* 407(1-4): 28-40.

Aghabeigi N, 2012. Providing a hybrid model for clustering on infertility data. M.Sc. Thesis of Management Information System, Department of Industrial Engineering, KN Toosi University of Technology.

Amiri V, 2016. Modeling and determination of saltwater interface in the aquifer of west side of the Urmia Lake using geophysical, hydrochemical and isotopic methods. Ph.D. Thesis of Hydrology, Faculty of Earth Sciences, Kharazmi University.

Amiri V, Nakhaei M and Lak R, 2017. Using radon-222 and radium-226 isotopes to deduce the functioning of a coastal aquifer adjacent to a hypersaline lake in NW Iran. *Journal of Asian Earth Sciences* 147: 128-147.

Amiri V, Nakhaei M, Lak R and Kholghi M, 2016a. Geophysical, isotopic, and hydrogeochemical tools to identify potential impacts on coastal groundwater resources from Urmia hypersaline Lake, NW Iran. *Environmental Science and Pollution Research* 23(16): 16738-16760.

Amiri V, Nakhaei M, Lak R and Kholghi M, 2016b. Investigating the salinization and freshening processes of coastal groundwater resources in Urmia aquifer, NW Iran. *Environmental Monitoring and Assessment* 188(4): 233.

- Amiri V, Nakhaei M, Lak R and Kholghi M, 2016c. Assessment of seasonal groundwater quality and potential saltwater intrusion: a study case in Urmia coastal aquifer (NW Iran) using the groundwater quality index (GQI) and hydrochemical facies evolution diagram (HFE-D). *Stochastic Environmental Research and Risk Assessment* 30(5): 1473-1484.
- Anonymous, 2018. Restoration and balancing plan of the groundwater resources, Operational plan for the restoration of groundwater resources in the Urmia study area. Iran Water Resources Management, Tehran.
- Chang FJ, Chang LC, Huang CW and Kao IF, 2016. Prediction of monthly regional groundwater levels through hybrid soft-computing techniques. *Journal of Hydrology* 541: 965-976.
- Chen IT, Chang LC and Chang FJ, 2018. Exploring the spatio-temporal interrelation between groundwater and surface water by using the self-organizing maps. *Journal of Hydrology* 556: 131-142.
- Dehghan N, Zare A, Peyvandi P and Hadizadeh M, 2012. K-mean clustering method for grouping the body form. Pp. 2-3. 8th National Conference on Textile Engineering. May, Yazd University, Iran. (In Persian with English abstract)
- Filzmoser P, Garrett RG and Reimann C, 2005. Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences* 31(5): 579-587.
- Ghahraman B and Kamran D, 2014. Adopting hierarchial cluster analysis to improve the performance of k-mean algorithm. *Water and Soil Science- University of Tabriz* 28(3): 471-480.
- Ghannadpur S, Hezarkhani A and Rudpeyma T, 2016. Estimation of anomalous values by using an appropriate combination between mahalanobis distances and three widely used data mining methods, case study: Parkam. *Journal of Aalytical and Numerical Methods in Mining Engineering* 7(13): 45-57. (In Persian with English abstract)
- Ghazanfari M, Alizadeh S and Tymurpoor B, 2008. *Data Mining and Knowledge Discovery*. Publish Center of Iran University of Science & Technology, Tehran.
- Han JC, Huang Y, Li Z, Zhao C, Cheng G and Huang P, 2016. Groundwater level prediction using a SOM-aided stepwise cluster inference model. *Journal of Environmental Management* 182: 308-321.
- Han J, Pei J and Kamber M, 2011. *Data Mining: Concepts and Techniques*. Elsevier, USA.
- Hosseini Z and Nakhaie M, 2015. Estimation of groundwater level using a hybrid genetic algorithm-neural network. *Pollution* 1(1): 9-21.
- Kalteh AM, Hjorth P and Berndtsson R, 2008. Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software* 23(7): 835-845.
- Kawamura A, Ishihara S, Amaguchi H and Takasaki T, 2016. Fluctuation patterns of groundwater levels in Tokyo caused by the Great East Japan Earthquake. Pp. EPSC2016-3932. EGU General Assembly Conference Abstracts. 17-22 April, Vienna, Austria.
- Kohonen T, 1997. *Self-Organizing Maps*. Springer-Verlag, Berlin, Heidelberg.
- Kohonen T, 1998. The Self-Organizing Map. *Neurocomputing* 21: 1-6.
- Moosavi V, Vafakhah M, Shirmohammadi B and Behnia N, 2013. A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods. *Water Resources Management* 27(5): 1301-1321.
- Moosavi Z, 2015. Modeling the turning away behavior of actual Mellat Bank customers using machine learning techniques. M.Sc. Thesis of Decision Science and Knowledge Engineering, Faculty of Management, Kharazmi University.
- Nohegar A, Kazemi M, Ahmadi J, Gholami H and Mahdavi R, 2018. Using univariate and multivariate methods to detect outliers in sediment fingerprinting method, case study: Tange Bostanak watershed. *Journal of Watershed Engineering and Management* 9(4): 398-412. (In Persian with English abstract)
- Nourani V, Alami MT and Vousoughi FD, 2015. Wavelet-entropy data pre-processing approach for ANN-based groundwater level modeling. *Journal of Hydrology* 524: 255-269.
- Nourani V, Alami MT and Vousoughi FD, 2016. Self-organizing map clustering technique for ANN-based spatiotemporal modeling of groundwater quality parameters. *Journal of Hydroinformatics* 18(2): 288-309.
- Nourani V, Mehrvand M and Baghanam AH, 2014. Implication of SOM-ANN based clustering for multistation rainfall-runoff modeling. *Journal of Urban and Environmental Engineering* 8(2): 198-210.

Tsai MJ, Abrahart RJ, Mount NJ and Chang FJ, 2014. Including spatial distribution in a data-driven rainfall-runoff model to improve reservoir inflow forecasting in Taiwan. *Hydrological Processes* 28(3): 1055-1070.

Vadiati M, Asghari-Moghaddam A, Nakhaei M, Adamowski J and Akbarzadeh AH, 2016. A fuzzy-logic based decision-making approach for identification of groundwater quality based on groundwater quality indices. *Journal of Environmental Management* 184: 255-270.