



Improving the performance of neural network based on group behavior with data (GMDH) using Harmonic Search Optimization (HSA) algorithm for simulating monthly river flow

Maryam Akbarichegeni¹ , Mohammad Ali Izadbakhsh^{2✉} , Saeid Shabanlou³ ,
Fariborz Yosefvand⁴ , Behrouz Yaghoubi⁵ , and Ahmad Rajabi⁶

1. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: ma.akbari817@gmail.com
2. Corresponding author, Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: izadbakhsh.mohammad.ali@gmail.com
3. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: saeid.shabanlou@gmail.com
4. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: fariborzosefvand@gmail.com
5. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: byaghubi@gmail.com
6. Department of Water Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran. E-mail: ahmad.rajabi1974@gmail.com

Article Info

ABSTRACT

Article type:
Research Article

Article history:
Received 05 October 2024
Received in revised form 10 January 2025
Accepted 01 February 2025
Available online 25 March 2025

Keywords:
GMDH,
neural network,
harmony search optimization
algorithm,
monthly river flow,
machine learning.

Objective: The aim of this study is to create a neural network model based on collective data encounter (GMDH) and improve it using Harmony Search Optimization (HSA) algorithm for simulating monthly river flow (HSA-GMDH).

Method: For this purpose, the rainfall and flow data of Gamasiab river of Kermanshah were used during a 20-year period (1370-1390). To develop the GMDH model, 80% of the data were used to train the model and other data were used to evaluate it. Also, the best input variables to the model were determined by trial and error method, based on this method, 3 datasets were formed (D1, D2, D3), then the performance of GMDH and HSA-GMDH models was evaluated with each of these datasets.

Results: The HSA-GMDH(D1) model performs better in the training and testing phase than the GMDH(D1) model. The HSA-GMDH(D2) model also performs better than the GMDH(d2) model. The HSA-GMDH(D3) model also performs better than the GMDH(D3) model. It is better to use an optimization algorithm such as HSA instead of the trial and error method to simulate the monthly flow of the river using the GMDH model.

Conclusions: Based on the results of this study, the HSA-GMDH model performs much better than the GMDH model, so it can be used as a powerful tool for simulating monthly river flow.

Cite this article: Akbarichegeni, M., Izadbakhsh, M. A., Shabanlou, S., Yosefvand, F., Yaghoubi, B., & Rajabi, E. (2025). Improving the performance of neural network based on group behavior with data (GMDH) using Harmonic Search Optimization (HSA) algorithm for simulating monthly river flow. *Advanced Technologies in Water Efficiency*, 5 (1), 64-77. <https://doi.org/10.22126/atwe.2025.11456.1146>



© The Author(s)

<https://doi.org/10.22126/atwe.2025.11456.1146>

Publisher: Razi University.

Introduction

One of the most important issues in the water resources management process is river flow modeling, because based on it, decisions can be made about important issues such as floods and droughts. One of the newest and best methods for modeling a variety of parameters in water resource management such as river flow are machine learning models. The aim of this study is to create a neural network model based on collective data encounter (GMDH) and improve it using Harmony Search Optimization (HSA) algorithm for simulating monthly river flow (HSA-GMDH).

Method

Area of study

The Gamasyab watershed is located in the west of the country and in the northern and central regions of the Zagros mountain range in the geographical range of 13 degrees 1 minutes to 13 degrees 3 minutes east longitude and 99 degrees 13 minutes to 91 degrees 13 minutes north latitude with an area of one million seventy three thousand hectares (Bazarafshan et al. 2013). This basin is limited from the north to the basin of Sirvan, Sefirod and Qara Chai Rivers, from the west to a part of the Karkheh watershed, and from the south to a part of the Dez river basin and a part of the Karkheh watershed. It is also limited to the Dez river basin from the evil side. The minimum height of this basin is 5539 meters above sea level and the maximum is 9699 meters. The rivers of the basin include the main tributary named Gamasyab, which originates from the northern and eastern parts of the basin through many sub-tributaries and joins the other tributaries of the Karkhe Dam basin at the end of the Kangavar Plain (Deh Kohnh et al., 2018).

GMDH neural network

GMDH is a method first developed by Ivakhenko as a multivariate analysis method for identifying and modeling complex systems. This model can be used without initial information to simulate complex systems. In fact, GMDH is a method to identify nonlinear systems with many input variables, which can be developed as a multilayer neural network. The parameters of this network are trained based on the least squares estimation approach.

Harmony Search Optimization Algorithm (HSA)

HSA is a kind of meta-innovative algorithm that was introduced in 2007 by Mahdavi et al., inspired by the behavior of musicians when playing a piece of music. Harmony means harmony, sound, and harmonious combination of sounds that are heard simultaneously and at the same time. Musicians use different musical instruments to produce a new music, and the beauty of music is the harmony between the notes of different instruments (Yang, 2009). In making music, the goal is to find the best harmony and produce the most beautiful music possible. Musicians try to choose better notes every performance so that the music evolves and becomes more beautiful every time (Mahdavi et al., 2007). Musicians remember the pieces played to compare the new piece with the previous pieces each time. From the simulation of the process that a musician goes through to harmonize a piece of music, an algorithm was extracted which we know as harmony search. This algorithm is inspired by music to reach the best answer. In the harmony algorithm, each data is a note and each answer vector is a harmony. The maximum number of layers and the maximum number of neurons in each layer of GMDH model are determined by trial and error method. Also, different parameters of HSA algorithm were determined by trial and error method. For this purpose, the rainfall and flow data of Gamasiab river of Kermanshah were used during a 20-year period (1370-1390). To develop the GMDH model, 80% of the data were used to train the model and

other data were used to evaluate it. Also, the best input variables to the model were determined by trial and error method, based on this method, 3 datasets were formed (D1, D2, D3), then the performance of GMDH and HSA-GMDH models was evaluated with each of these datasets.

Results

After developing the HSA-GMDH model, its results were compared with the GMDH model whose parameters were determined using the trial and error method. Based on the results, the HSA-GMDH(D1) model performs better in the training and testing phase than the GMDH(D1) model, so that the HSA-GMDH(D1) model has a lower RMSE value than the GMDH(D1) model, as well as the NSE and R^2 values. The HSA-GMDH(D2) model also performs better than the GMDH(D2) model and has a lower RMSE, NSE and R^2 value than the GMDH(D2) model. Like the developed models of D1 and D2 data, the HSA-GMDH(D3) model also performs better than the GMDH(D3) model, so that its RMSE value is lower than the GMDH(D3) model, and the NSE value and Its R^2 is higher than GMDH(D3) model. Therefore, based on the results of this study, it is better to use an optimization algorithm such as HSA instead of the trial and error method to simulate the monthly flow of the river using the GMDH model.

Conclusions

Finally, the performance of GMDH and HSA-GMDH models were compared. Based on the results of this study, the HSA-GMDH model performs much better than the GMDH model, so it can be used as a powerful tool for simulating monthly river flow.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data Availability Statement.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

Not applicable.

Conflict of Interest

The authors declare no conflict of interest.



بهینه سازی جریان ماهانه رودخانه گاماسیاب کرمانشاه با استفاده از شبکه عصبی مبتنی بر رفتار گروهی با داده ها (GMDH) بهبود یافته با الگوریتم بهینه سازی جست و جوی هارمونی (HSA)

مریم اکبری چگنی^۱، محمدعلی ایزدبخش^۲، سعید شعبانلو^۳، فریبرز یوسفوند^۴، بهروز یعقوبی^۵، و احمد رجبی^۶

۱. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: ma.akbari817@gmail.com

۲. نویسنده مسئول، گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: izadbakhsh.mohammad.ali@gmail.com

۳. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: saeid.shabanlou@gmail.com

۴. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: fariborzosefvand@gmail.com

۵. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: byaghubi@gmail.com

۶. گروه مهندسی آب، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران. رایانامه: ahmad.rajabi1974@gmail.com

چکیده

اطلاعات مقاله

هدف: این پژوهش باهدف توسعه و بهبود مدل شبکه عصبی مبتنی بر برخورد جمعی با داده ها (GMDH) برای شبیه سازی جریان ماهانه رودخانه انجام شد.

نوع مقاله: مقاله پژوهشی

روش پژوهش: برای این منظور، الگوریتم بهینه سازی جستجوی هارمونی (HSA) به منظور ارتقای عملکرد مدل GMDH استفاده شد و مدل ترکیبی HSA-GMDH ارائه گردید. پارامترهایی نظیر حداکثر تعداد لایه ها و حداکثر تعداد نوروها در هر لایه از طریق روش آزمون و خطا تعیین شدند. همچنین، پارامترهای مختلف الگوریتم HSA نیز با رویکرد آزمون و خطا تنظیم شدند. در این مطالعه، داده های بارش و جریان رودخانه گاماسیاب در استان کرمانشاه طی یک بازه زمانی ۲۰ ساله (۱۳۷۰-۱۳۹۰) به کار گرفته شد. برای توسعه مدل GMDH، از ۸۰ درصد داده ها جهت آموزش مدل و از ۲۰ درصد باقی مانده برای ارزیابی آن استفاده گردید. علاوه بر این، بهترین متغیرهای ورودی به مدل با روش آزمون و خطا شناسایی شدند و براین اساس، سه مجموعه داده (D1، D2 و D3) تشکیل شد. سپس عملکرد مدل های GMDH و HSA-GMDH با استفاده از هر یک از این مجموعه داده ها مورد ارزیابی قرار گرفت و در نهایت، مقایسه ای میان عملکرد این دو مدل صورت پذیرفت.

تاریخ دریافت: ۱۴۰۳/۰۷/۱۴

تاریخ بازنگری: ۱۴۰۳/۱۰/۲۱

تاریخ پذیرش: ۱۴۰۳/۱۱/۱۳

تاریخ انتشار: ۱۴۰۴/۰۱/۰۵

کلیدواژه ها:

شبکه عصبی GMDH،
الگوریتم بهینه سازی جستجوی
هارمونی،
جریان ماهانه رودخانه،
یادگیری ماشین.

یافته ها: بر اساس نتایج حاصل مدل GMDH با دیتاست D1 بهترین عملکرد را در مرحله آموزش و تست دارد و با دیتاست D2 بدترین عملکرد را در مرحله آموزش و تست دارد همچنین این مدل با دیتاست D3 نسبت به مدل D1 عملکرد ضعیف تری دارد؛ اما نسبت به دیتاست D2 عملکرد بهتری دارد. بر اساس نتایج حاصل هر سه مدل GMDH توسعه یافته با HSA نسبت به مدل های GMDH توسعه یافته با روش سعی و خطا کارایی بهتری دارند. بر اساس نتایج حاصل مدل HSA-GMDH(D1) نسبت به مدل GMDH(D1) عملکرد بهتری در مرحله آموزش و تست دارد. مدل HSA-GMDH(D2) نیز نسبت به مدل GMDH(D2) عملکرد بهتری دارد. مدل HSA-GMDH(D3) نیز نسبت به مدل GMDH(D3) عملکرد بهتری دارد.

نتیجه گیری: نتایج این پژوهش نشان داد که مدل HSA-GMDH عملکرد بهتری نسبت به مدل GMDH دارد؛ بنابراین، مدل پیشنهادی HSA-GMDH به عنوان یک ابزار قدرتمند می تواند در شبیه سازی جریان ماهانه رودخانه مورد استفاده قرار گیرد.

استناد: اکبری چگنی، مریم؛ ایزدبخش، محمدعلی؛ شعبانلو، سعید؛ یوسفوند، فریبرز؛ یعقوبی، بهروز؛ و رجبی، احمد. (۱۴۰۴). بهینه سازی جریان ماهانه رودخانه گاماسیاب کرمانشاه با استفاده از شبکه عصبی مبتنی بر رفتار گروهی با داده ها (GMDH) بهبود یافته با الگوریتم بهینه سازی جست و جوی هارمونی (HSA). فناوری های پیشرفته در بهره وری آب، ۵ (۱)، ۶۴-۷۷.

<https://doi.org/10.22126/ATWE.2025.11456.1146>

شر: دانشگاه رازی.



© نویسندگان

مقدمه

یکی از چالش‌های اصلی در مدیریت منابع آب، محدودیت‌های فنی و مالی در نصب تجهیزات و ایستگاه‌های هیدرومتری در تمام بخش‌های رودخانه‌ها است. این تجهیزات که برای اندازه‌گیری دقیق پارامترهایی همچون دبی، سرعت جریان و سایر مشخصه‌های هیدرولوژیکی به کار می‌روند، نیازمند سرمایه‌گذاری قابل توجه، نگهداری مستمر و دسترسی به مناطق صعب‌العبور هستند. علاوه بر این، در برخی مناطق به دلیل شرایط جغرافیایی دشوار، دسترسی به این نقاط و نصب تجهیزات هیدرومتری تقریباً غیرممکن می‌شود. این محدودیت‌ها موجب شده است که تخمین دقیق دبی در کل حوزه آبریز به یک مسئله اساسی تبدیل شود. تخمین دقیق دبی حوضه آبریز نقش مهمی در برنامه‌ریزی و مدیریت منابع آب ایفا می‌کند. این اطلاعات برای پیش‌بینی سیلاب‌ها، برنامه‌ریزی آبیاری کشاورزی، تخصیص منابع آب برای مصارف شرب و صنعتی و همچنین مدیریت خشک‌سالی ضروری هستند. در نبود داده‌های دقیق و گسترده از سراسر رودخانه، تخمین‌های مبتنی بر مدل‌سازی می‌توانند جایگزین ارزشمندی برای جمع‌آوری داده‌های مستقیم و گسترده از ایستگاه‌های هیدرومتری باشند. با این حال، مدل‌های سنتی اغلب نمی‌توانند پیچیدگی‌ها و تغییرات غیرخطی سیستم‌های هیدرولوژیکی را به‌طور دقیق شبیه‌سازی کنند. در سال‌های اخیر، با پیشرفت فناوری و افزایش دسترسی به داده‌های بزرگ (Big Data)، استفاده از روش‌های مبتنی بر هوش مصنوعی برای مدل‌سازی جریان رودخانه و تخمین دبی حوضه آبریز به‌طور چشمگیری افزایش یافته است. این روش‌ها با بهره‌گیری از یادگیری ماشین، قادر به تجزیه و تحلیل حجم زیادی از داده‌ها و استخراج الگوهای پیچیده و غیرخطی هستند که در روش‌های سنتی کمتر قابل دستیابی بودند. مدل‌هایی مانند شبکه‌های عصبی مصنوعی (ANN)، سیستم‌های استنتاج فازی و الگوریتم‌های یادگیری عمیق توانسته‌اند نتایج بسیار دقیق‌تری در پیش‌بینی جریان رودخانه ارائه دهند (شیانگ و همکاران^۱، ۲۰۲۱). شبکه‌های عصبی مصنوعی یکی از انواع روش‌های هوش مصنوعی هستند که می‌توان از آن‌ها برای اهداف مختلفی از جمله مدل‌سازی پارامترهای مختلف یا، پیش‌بینی مقدار یک متغیر در آینده، خوشه‌بندی متغیرهای مختلف و اهداف فراوان دیگری استفاده کرد (مزرعه و همکاران^۲، ۲۰۲۳).

ادیات موضوع و پیشینه پژوهش

یکی از مهم‌ترین مسائل در مدیریت منابع آب تجزیه و تحلیل جریان رودخانه است (گویامل و لی^۳، ۲۰۲۰)، برای این کار می‌توان از مدل‌های فیزیکی و مدل‌های مبتنی بر داده استفاده کرد، برای توسعه مدل‌های فیزیکی علاوه بر هزینه زیاد، به انواع زیاد تجهیزات و همچنین کاربران خبره نیاز است (مزرعه و همکاران^۴، ۲۰۲۴). به همین دلیل در دهه‌های اخیر بیش‌تر مطالعات در جهت بهبود روش‌های مبتنی بر داده متمرکز شده‌اند. در رابطه با پژوهش حاضر، مطالعاتی در داخل و خارج از کشور صورت گرفته است که در ادامه تعدادی از این مطالعات ارائه شده است.

۱. مطالعات داخلی

سبزواری و همکاران (۱۴۰۲) کارایی مدل ANN بهبودیافته با الگوریتم کرم شبتاب را برای پیش‌بینی مقدار جامدات محلول در آب (TDS) را بری کردند. آن‌ها برای توسعه این مدل از متغیرهای دبی (Q)، سدیم (NA)، منیزیم (Mg)، کلسیم (Ca)، سولفات (SO₄)، کلرید (Cl)، بی‌کربنات (HO₃)، هدایت الکتریکی (EC) و جامدات محلول رودخانه در بازه زمانی قبل ((TDS)_(t-1)) استفاده کردند. بر اساس نتایج حاصل از این مطالعه زمانی که وزن‌های مدل ANN با الگوریتم کرم شبتاب تعیین شوند نسبت به زمانی که وزن‌های این مدل به صورت تصادفی تعیین شوند عملکرد بهتری دارد؛ بنابراین برای مدل‌سازی TDS با استفاده از مدل ANN بهتر است از یک الگوریتم فرا ابتکاری مانند الگوریتم کرم شبتاب استفاده شود.

اسماعیلی و همکاران (۱۴۰۲) برای اولین بار بارش بلندمدت در شهر بابلسر را مدل سازی کردند. آن ها برای این کار از مدل برنامه ریزی بیان ژن (GEP) و روش تبدیل موجک (WT) برای توسعه یک مدل هیبریدی جدید (WT-GEP) استفاده کردند. سپس نتایج حاصل را با نتایج حاصل از مدل GEP مقایسه کردند. بر اساس نتایج حاصل از این مطالعه مدل WT-GEP نسبت به مدل GEP عملکرد بهتری دارد.

۲. مطالعات خارجی

کامبالیمات و دکا^۱ (۲۰۲۱) از ترکیب مدل SVM و تبدیل موجک (WT) برای پیش بینی دبی ماهانه رودخانه استفاده کردند. برای این کار از داده های جریان ماهانه رودخانه استفاده کردند. در این مطالعه از مقادیر مشاهداتی دبی جریان از یک تا سه ماه قبل برای پیش بینی دبی ماه جاری به عنوان ورودی مدل ها استفاده شد. نتایج این مطالعه نشان داد که در هر دو ایستگاه مدل هیبرید WT-SVM عملکرد بهتری نسبت به مدل SVM دارد.

رضایی و وادیاتی^۲ (۲۰۲۰) عملکرد مدل های SVM و شبکه عصبی مصنوعی (ANN) و سیستم استنتاج فازی عصبی - تطبیقی (ANFIS) برای پیش بینی دبی روزانه رودخانه در من کوهستانی نیمه خشک شمال غرب چین ارزیابی کردند. برای این مطالعه از داده های سال ۲۰۰۱ تا ۲۰۰۳ و ۲۰۰۹ تا ۲۰۱۱ استفاده کردند. بر اساس نتایج حاصل از این مطالعه به خوبی می توان از این مدل ها برای مدل سازی جریان رودخانه استفاده کرد، اما بررسی دقیق تر نتایج مدل ها نشان می دهد که مدل SVM نسبت به سایر مدل ها عملکرد بهتری دارد.

چن و همکاران^۳ (۲۰۲۰) میانگین دبی جریان روزانه رودخانه های بهشت آباد (دائمی) و جونگان (غیردائمی) در شمال غرب ایران با چهار مدل شامل الگوریتم برنامه سازی بیان ژن (GEP)، WT-ANN، ANN و WTt-GEP (WGEP) را برای سال های ۱۹۹۹-۲۰۰۸ مدل سازی کردند. ورودی های مدل ها شامل دبی رودخانه تأخیر داده شده از یک تا سه روز قبل و برای بارش از یک تا سه روز قبل بودند. بر اساس نتایج حاصل از این مطالعه مدل GEP نسبت به سایر مدل ها عملکرد بهتری دارد. همچنین از میان الگوهای ورودی، الگوهای مبتنی بر دبی جریان و ترکیب دبی جریان و بارش با تفاوت ناچیز به یکدیگر، بهترین الگوها به دست.

صمدیانی فر و همکاران^۴ (۲۰۱۹) برای بهبود دقت مدل سازی جریان روزانه رودخانه با استفاده از مدل SVM، برای تعیین پارامترهای این مدل از الگوریتم بهینه سازی الگوریتم مگس میوه (FOA) تعیین کردند. سپس نتایج حاصل از مدل FOA-SVM را با نتایج مدل درخت تصمیم (M5) مقایسه کردند. آن ها برای ارزیابی نتایج از معیارهای MAE، R، معیار اطلاع بیزی-شوارتز (BIC) استفاده کردند. بر اساس نتایج حاصل از این مطالعه مدل FOA-SVM نسبت به مدل M5 عملکرد بهتری دارد.

از دیگر تحقیقات در زمینه موضوع این تحقیق می توان به تحقیقات عزیز و همکاران (۲۰۲۳)؛ امیری و همکاران (۲۰۲۱)؛ شعبانلو (۲۰۱۸)؛ اسماعیلی و همکاران (۲۰۲۱)؛ فلاحی و همکاران (۲۰۲۳)؛ پناهی و همکاران (۲۰۲۲)؛ جلیلیان و همکاران (۲۰۲۲)؛ مرادی و همکاران (۲۰۲۳) و امان جلیلی و همکاران (۲۰۲۳) اشاره کرد.

روش پژوهش

۱. منطقه مورد مطالعه

حوضه آبخیز گاماسیاب در غرب کشور و در مناطق شمالی و مرکزی رشته کوه زاگرس در محدوده جغرافیایی ۱۳ درجه ۱ دقیقه تا ۱۳ درجه ۳ دقیقه طول شرقی و ۹۹ درجه ۱۳ دقیقه تا ۹۱ درجه ۱۳ دقیقه عرض شمالی با مساحت یک میلیون

1. Kambalimath & Deka

2. Rezaei & Vadiati

3. Chen et al

4. Samadianfard et al

هفتادوسه هزار هکتار قرار دارد (بذرافشان و همکاران، ۱۳۹۳). این حوضه از شمال به حوضه رودخانه‌های سیروان، سفیدرود و قره چای، از مغرب به قسمتی از حوزه آبخیز کرخه و از جنوب به قسمتی از حوضه رودخانه دز و قسمتی از حوزه آبخیز کرخه محدود می‌شود. همچنین از سمت شرق به حوضه رودخانه دز محدود می‌گردد. حداقل ارتفاع این حوضه ۵۵۳۹ متر از سطح دریا و حداکثر آن ۹۶۹۹ متر است. رودخانه‌های حوضه شامل سرشاخه اصلی به نام گاماسیاب بوده که از بخش شمالی و شرقی حوضه به وسیله سرشاخه‌های فرعی زیادی سرچشمه گرفته و در انتهای دشت کنگاور به دیگر سرشاخه‌های حوضه سد کرخه می‌پیوندند (ده کهنه و همکاران، ۱۳۹۸).

۲. شبکه عصبی GMDH^۱

GMDH روشی است اولین بار که توسط ایواخنکو^۲ به عنوان یک روش تجزیه و تحلیل چندمتغیره برای شناسایی و مدل‌سازی سیستم‌های پیچیده توسعه داده شد. می‌توان از این مدل بدون داشتن اطلاعات اولیه برای شبیه‌سازی سیستم‌های پیچیده استفاده کرد (فارلو^۳، ۲۰۲۰). درواقع GMDH روشی برای شناسایی سیستم‌های غیرخطی با متغیرهای ورودی زیاد است که می‌توان آن را به صورت یک شبکه عصبی چندلایه توسعه داد. آموزش پارامترهای این شبکه بر اساس رویکرد تخمین حداقل مربعات انجام می‌شود. این مدل به طور گسترده‌ای در زمینه‌های مختلف تحقیقاتی و کاربردی برای برخورد با عدم قطعیت‌ها، سیستم‌های خطی و غیرخطی توسط محققان مختلفی مورد استفاده قرار گرفته است. ارتباط بین متغیرهای ورودی - خروجی در این مدل را می‌توان با سری ولترا که مشابه با چندجمله‌ای گسسته شده کولموگروف - گابور است نشان داد که در رابطه (۱) ارائه شده است.

$$f(X) = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ijk} x_i x_j + \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n a_{ijk} x_i x_j x_k + \dots \quad (1)$$

در این رابطه $X(x_1, x_2, x_3, \dots, x_m)$ متغیر ورودی و $W(w_1, w_2, w_3, \dots, w_m)$ بردار وزن‌ها است، m نیز تعداد متغیرها ورودی به شبکه عصبی GMDH است. تابع انتقال یا محرک این نرون‌ها می‌تواند به صورت یک چندجمله‌ای خطی یا غیرخطی به صورت زیر بیان شود که در رابطه (۲) نشان داده شده است.

$$y = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_2^2 + w_5 x_1 x_2 \quad (2)$$

در الگوریتم GMDH برای تولید نرون‌های لایه اول، تمام متغیرهای ورودی باهم ترکیب شده و تمام ترکیبات دوتایی متغیرهای ورودی در نظر گرفته می‌شوند. سپس بخشی از متغیرها به عنوان ورودی به لایه بعد منتقل می‌شوند. از روش‌های مختلفی مانند روش تجزیه مقادیر منفرد (SVD)^۴ می‌توان برای محاسبه ضرایب این رابطه استفاده کرد (فارلو، ۲۰۲۰) اما در این مطالعه از الگوریتم بهینه‌سازی جستجوی هارمونی (HSA)^۵ برای تعیین این ضرایب استفاده شد.

۳. الگوریتم بهینه‌سازی جستجوی هارمونی (HSA)^۵

HSA نوعی الگوریتم فرا ابتکاری است که در سال ۲۰۰۷ توسط مهدوی و همکاران^۶ با الهام از رفتار نوازندگان هنگام نواختن یک قطعه موسیقی معرفی شد. هارمونی به معنای هم آهنگی، هم صدایی و ترکیب خوش آهنگ صداهایی که به صورت هم‌زمان و در یک لحظه شنیده می‌شوند، است. نوازندگان برای تولید یک موسیقی جدید از آلات موسیقی مختلفی استفاده می‌کنند و زیبایی موسیقی به هماهنگی بین نت‌های سازهای مختلف است (یانگ^۷، ۲۰۰۹). در ساخت یک موسیقی، هدف یافتن بهترین هارمونی و تولید زیباترین موسیقی ممکن است. نوازندگان در هر بار اجرا سعی می‌کنند نت‌های بهتری انتخاب کنند تا در هر بار اجرا، موسیقی تکامل بیشتری یافته و زیباتر گردد (مهدوی و همکاران، ۲۰۰۷). نوازندگان قطعه‌های نواخته شده را به خاطر می‌سپارند تا در هر بار قطعه جدید را با قطعات قبلی مقایسه کنند. از شبیه‌سازی فرایندی که یک نوازنده برای هارمونیک کردن یک قطعه

1. Group method of data handling
2. Ivakhnenko
3. Farlow
4. Singular value decomposition
5. Harmony Search Algorithm
6. Mahdavi et al
7. Yang

موسیقی طی می کند، الگوریتمی استخراج شده که آن را بانام جستجوی هارمونی می شناسیم. این الگوریتم برای رسیدن به بهترین جواب، از موسیقی الهام گرفته شده است (جیم^۱، ۲۰۰۹). در الگوریتم هارمونی هر یک از داده ها را یک نت و هر بردار جواب یک هارمونی است. مراحل توسعه الگوریتم جستجوی هارمونی را می توان به پنج گام تقسیم کرد (مهدوی و همکاران، ۲۰۰۷).

گام اول: مقداردهی اولیه مسئله بهینه سازی و پارامترهای اولیه است. به این معنی که ابتدا مسئله به صورت زیر تعریف می شود:

$$\text{Minf}(x). x_i \text{ is in } X_i, i=1,2,3...N \quad (3)$$

در این مرحله مقدار پارامترها، اندازه حافظه هارمونی HMS^۲، سرعت بررسی حافظه هارمونی HMCR^۳، سرعت تنظیم قطعات (PAR^۴)، حداکثر تعداد تکرارها (NI^۵) تعیین می شود.

گام دوم: مقدار HMS با مقادیر اولیه ای که به طور تصادفی تولید می شوند، پر می گردد.

گام سوم: ایجاد یک هارمونی جدید بهبودیافته است، به این صورت که در این گام یک بردار هارمونی جدید بر اساس روش های در نظر گرفتن حافظه، انتخاب تصادفی و تنظیم زیرومی تولید می شود.

$$x'_i = \begin{cases} x_i \in \{x_i^1, x_i^2 \dots x_i^{HMS}\} & \text{HMCR} \\ x_i \in X_i & (1-HMCR) \end{cases} \quad (4)$$

حال مقداری که در مرحله قبل از حافظه انتخاب شده است با احتمال PAR به مرحله تنظیم زیرومی می رود.

$$x'_i = \begin{cases} \text{yes with probability PAR} \\ \text{No with probability PAR} \end{cases} \quad (5)$$

اگر درایه مورد نظر شانس تغییر داشته باشد با فرمول زیر تغییر می کند:

$$x'_i \leftarrow x'_i \pm \text{rand}() * bw \quad (6)$$

به این معنی که مقدار درایه با یک مقدار رندم بین -۱ و ۱ جمع می شود و سپس ضرب در مقدار پهنای باند می شود. پهنای باند همان مکانیسم تمرکز و تنوع است. به عبارتی هدف این است که در ابتدای الگوریتم تنوع زیاد باشد و هرچه به پایان الگوریتم نزدیک می شویم تنوع کاهش و تمرکز افزایش یابد. پس در هر تکرار پهنای باند را با یک ضریب کاهنده کاهش می دهیم. برای این کار تغییر در حافظه هارمونی با دو روش زیر انجام شود:

۱- در نظر گرفتن نت های موجود در حافظه یعنی تعدادی از درایه ها را بر اساس نت های موجود در حافظه تغییر می دهد.

۲- تنظیم زیرومی نت ها: یعنی یک تعداد از داده ها را انتخاب می کند و مقداری به آن ها اضافه می کند

گام چهارم: به روز کردن حافظه هارمونی است به این صورت که اگر بردار هارمونی جدید از بدترین هارمونی موجود در حافظه بهتر باشد، جایگزین آن می شود.

گام پنجم: تکرار شدن گام های ۵ و ۸ می باشد تا زمانی که شرط پایانی ارضا شود یا تکرارها تمام شود.

۴. معیارهای ارزیابی مدل

بعد از توسعه مدل ها از شاخص های جذر میانگین مربعات خطا (RMSE^۶)، R ضریب تعیین^۷ (R²)، ضریب NSE^۸ برای ارزیابی کارایی مدل های مختلف استفاده شد که در روابط (۹) الی (۱۱) ارائه شده است (چای و^۹، ۲۰۱۴؛ مک کوئن و همکاران^{۱۰}، ۲۰۰۶؛ و اوزر^{۱۱}، ۱۹۸۵).

1. Geem
2. Harmony memory size
3. Harmony memory considering rate
4. Pitch adjusting rate
5. Number of improvisations
6. Root Mean Squared Error
7. Coefficient of determination

$$RMSE = \left[N^{-1} \sum_{i=1}^N (p_i - o_i)^2 \right]^{\frac{1}{2}} \quad (7)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (o_i - p_i)^2}{\sum_{i=1}^N (o_i - \bar{o})^2} \quad (8)$$

$$R^2 = \left[\sum_{i=1}^N (p_i - \bar{p})(o_i - \bar{o}) \right]^2 \left[\sum_{i=1}^N (p_i - \bar{p})^2 \sum_{i=1}^N (o_i - \bar{o})^2 \right]^{-1} \quad (9)$$

در این رابطه N تعداد مشاهده‌ها است، p_i مقدار پیش‌بینی شده است، o_i مقدار مشاهده‌ای است و $\bar{p}$ و $\bar{o}$ به ترتیب متوسط مقادیر p_i و o_i هستند

یافته‌های پژوهش

۱. انتخاب متغیرهای ورودی به مدل

یکی از مهم‌ترین بخش‌های هر مدل یادگیری ماشین انتخاب بهترین متغیرهای مستقل برای مدل‌سازی متغیر وابسته است. زیرا هر چه تعداد متغیرهای مستقل بیشتر باشد مدل پیچیده‌تر و همگرایی آن دشوارتر می‌شود، از طرفی اگر تعداد متغیرهای مستقل کم باشد نمی‌توان متغیر وابسته را با دقت خوبی مدل‌سازی کرد. به این ترتیب در این مطالعه برای مدل تشخیص بهترین متغیرهای ورودی از روش NCA^۸ استفاده شد. برای توسعه این روش از exact, non and average fit method و Fletcher–Goldfarb–Shanno (BFGS) solver^۹ استفاده شد، همچنین مقدار پارامتر لامبدا با استفاده از روش سعی و خطا تعیین شد (یانگ و همکاران^{۲، ۱۲}). با استفاده از این روش سه نوع دیتاست برای تشکیل مدل GMDH تعیین شدند که در جدول (۱) نشان داده شده است. بر اساس نتایج حاصل از این روش دبی رودخانه و بارش در گام $t-1$ بر اساس همه روش‌ها پارامترهای مهمی برای شبیه‌سازی جریان رودخانه در گام زمانی t تعیین شده‌اند.

جدول ۱. ترکیب متغیرهای مختلف شناسایی شده با استفاده از روش NCA برای توسعه مدل GMDH

متغیرهای مستقل	روش حل	متغیر وابسته	نوع NCA	دیتاست
$Q_{t-1}, Q_{t-2}, Q_{t-4}, R_t, R_{t-1}, R_{t-3}, T_t, T_{t-3}, T_{t-5}$	BFGS	Q_t	Exact	D1
$Q_{t-1}, Q_{t-4}, Q_{t-5}, R_t, R_{t-1}, R_{t-2}, T_t, T_{t-1}, T_{t-2}$	BFGS	Q_t	Non	D2
$Q_{t-1}, Q_{t-2}, Q_{t-3}, R_t, R_{t-1}, R_{t-4}, T_{t-1}, T_{t-2}, T_{t-3}$	BFGS	Q_t	average	D3

۲. نتایج مدل GMDH

بعد از مشخص شدن بهترین متغیرها برای تشکیل ماتریس ورودی مدل GMDH توسعه داده شد، این مدل ابتدا با استفاده از داده‌های آموزش توسعه داده شد سپس با داده‌های تست ارزیابی شد. برای توسعه این مدل حداکثر تعداد لایه‌ها، حداکثر تعداد نورون‌های هر لایه و مقدار پارامتر SP^3 که بر اساس آن تعیین می‌شود چند درصد از نورون‌های هر لایه به لایه بعد منتقل شوند با روش سعی و خطا تعیین شدند. همچنین مقدار اولیه پارامترهای اندازه جمعیت (PS^4) و حداکثر تعداد تکرارها (MNI^5) با روش سعی و خطا تعیین شدند. مقدار حد پایین و بالا متغیرها HSA نیز به ترتیب ۱- و ۱ در نظر گرفته شد، این مراحل برای هر یک از مجموع داده‌های D1، D2 و D3 انجام شد و شاخص‌های خطا برای هر یک از مدل‌ها محاسبه شدند. بر اساس نتایج حاصل

8. Nash-Sutcliffe efficiency

9. Chai & Draxler

10. McCuen et al

11. Ozer

1. Neighborhood component analysis

2. Yang et al

3. Selection Pressure

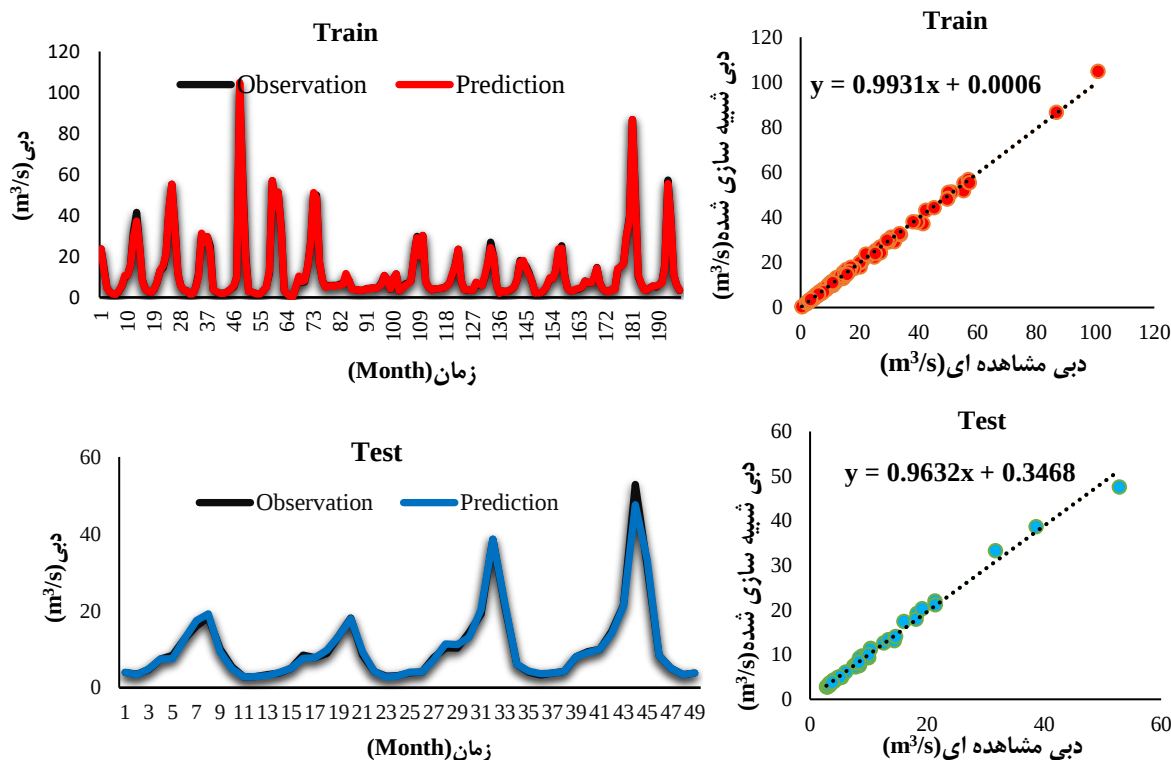
4. Population Size

5. Maximum Number Of Iterations

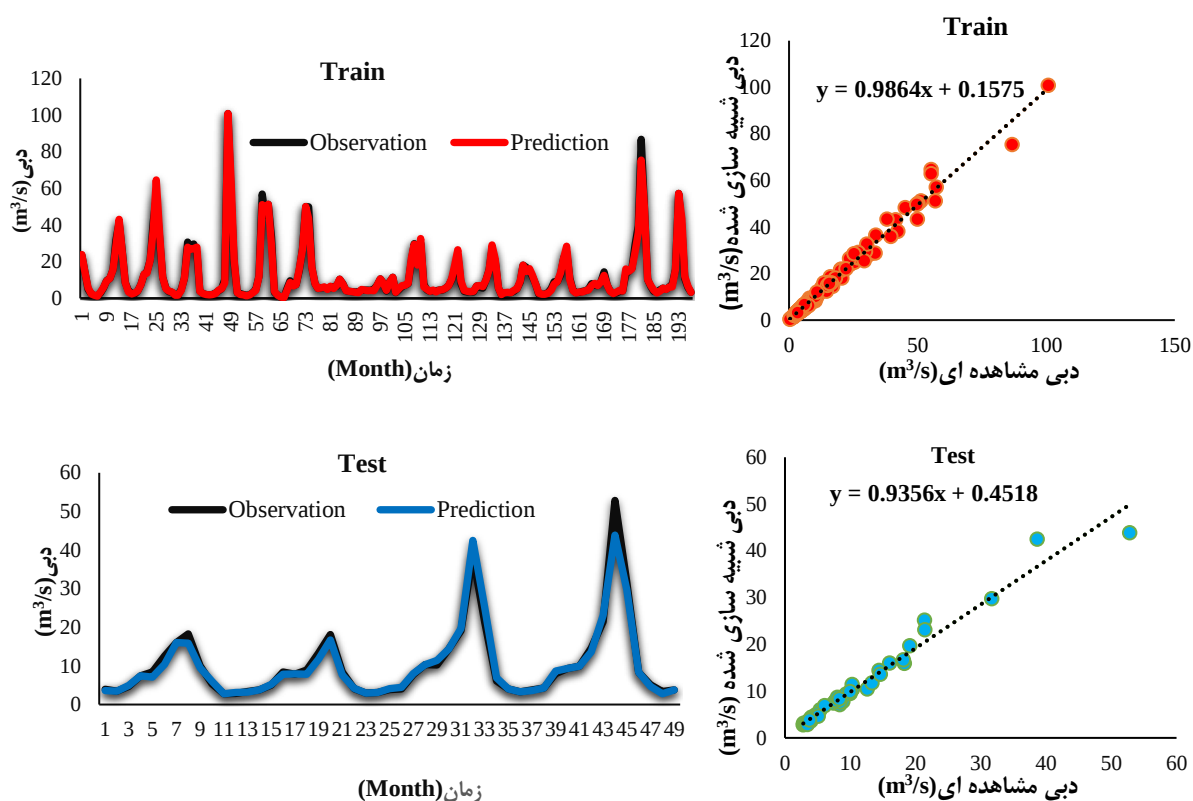
مدل GMDH با دیتاست D1 بهترین عملکرد را در مرحله آموزش و تست دارد (Train: $RMSE: 0.0.2632, NSE: 0.9921, R^2: 0.9984$ Test: $RMSE: 0.2010, NSE: 0.9926, R^2: 0.9987$ و با دیتاست D2 بدترین عملکرد را در مرحله آموزش و تست دارد (Train: $RMSE: 0.2070, NSE: 0.9237, R^2: 0.9850$, Test: $RMSE: 0.2604, NSE: 0.9259, R^2: 0.9857$). همچنین این مدل با دیتاست D3 نسبت به مدل D1 عملکرد ضعیفتری دارد (Train: $RMSE: 0.2485, NSE: 0.9506, R^2: 0.9885$, Test: $RMSE: 0.2430, NSE: 0.95126, R^2: 0.9887$) اما نسبت به دیتاست D2 عملکرد بهتری دارد. در جدول (۲) مقدار شاخص های خطا مدل های GMDH(D1), GMDH(D2), HSA-GMDH(D3) در مرحله آموزش و ارزیابی نشان داده شده اند. همچنین در شکل های (۱) تا (۳) مقدار جریان رودخانه واقعی و مقدار مدل سازی شده به وسیله مدل های مختلف نشان داده شده است.

جدول ۲. نتایج مدل ماهانه شبکه عصبی GMDH برای هر یک از دیتاست های D3، D2، D1 در مرحله آموزش و تست

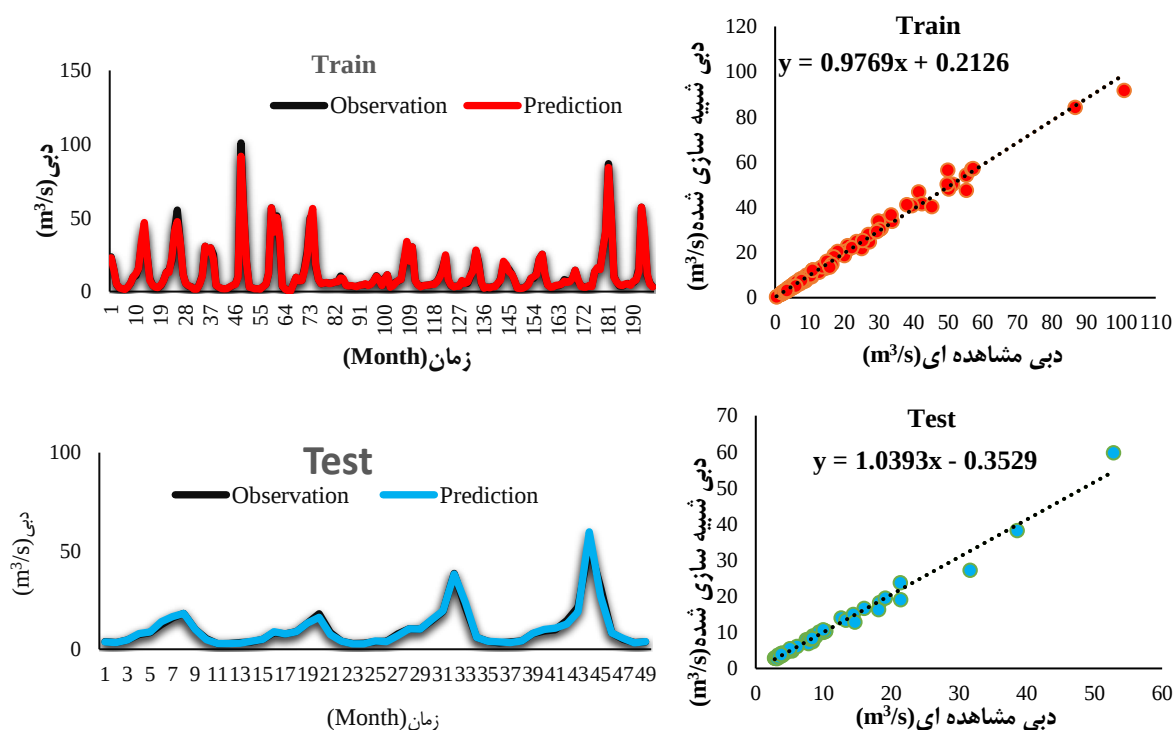
مدل	آموزش			تست		
	RMSE	NSE	R ²	RMSE	NSE	R ²
GMDH(D1)	0.1647	0.9444	0.9497	0.1575	0.9445	0.9498
GMDH(D2)	0.1951	0.9352	0.9471	0.1920	0.9371	0.9477
GMDH(D3)	0.1863	0.9425	0.9484	0.1809	0.9429	0.9487



شکل ۱. نتایج مدل GMDH(D1) برای مرحله آموزش و تست



شکل ۲. نتایج مدل GMDH(D2) برای مرحله آموزش و تست



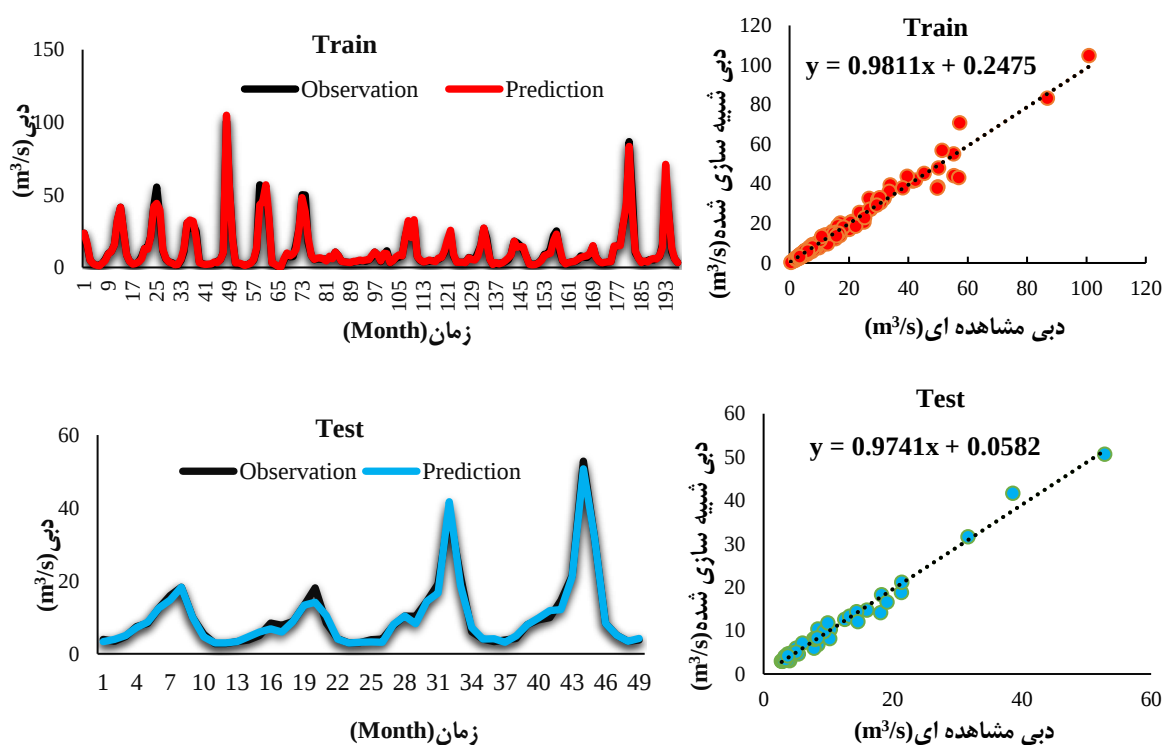
شکل ۳. نتایج مدل GMDH(D3) برای مرحله آموزش و تست

۳. نتایج مدل HSA-GMDH

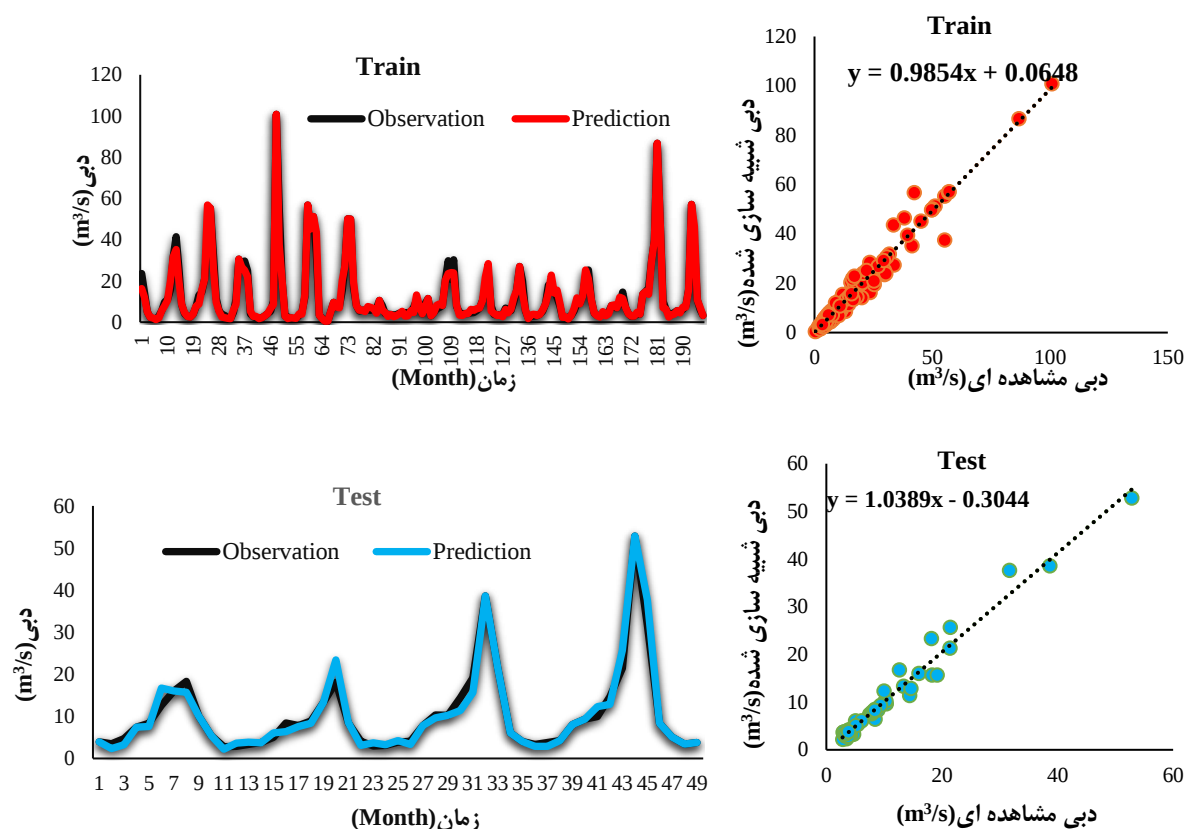
در این مرحله از الگوریتم HSA برای بهبود کارایی مدل‌ها GMDH(D1)، GMDH(D2)، GMDH(D3) استفاده شد. به این ترتیب سه مدل HSA-GMDH(D1)، HSA-GMDH(D2)، HSA-GMDH(D3) توسعه داده شدند. در جدول (۲) مقدار شاخص‌های خطای مدل‌های HSA-GMDH(D1)، HSA-GMDH(D2)، HSA-GMDH(D3) در مرحله آموزش و ارزیابی نشان داده شده‌اند همچنین در شکل‌های (۶) تا (۸) مقدار جریان رودخانه واقعی و مقدار مدل‌سازی شده به وسیله مدل‌های مختلف نشان داده شده است. بر اساس نتایج حاصل هر سه مدل GMDH توسعه یافته با HSA نسبت به مدل‌های GMDH توسعه یافته با روش سعی و خطا کارایی بهتری دارند.

جدول ۳. نتایج مدل ماهانه شبکه عصبی HSA-GMDH برای هر یک از دیتاست‌های D1، D2، D3 در مرحله آموزش و تست

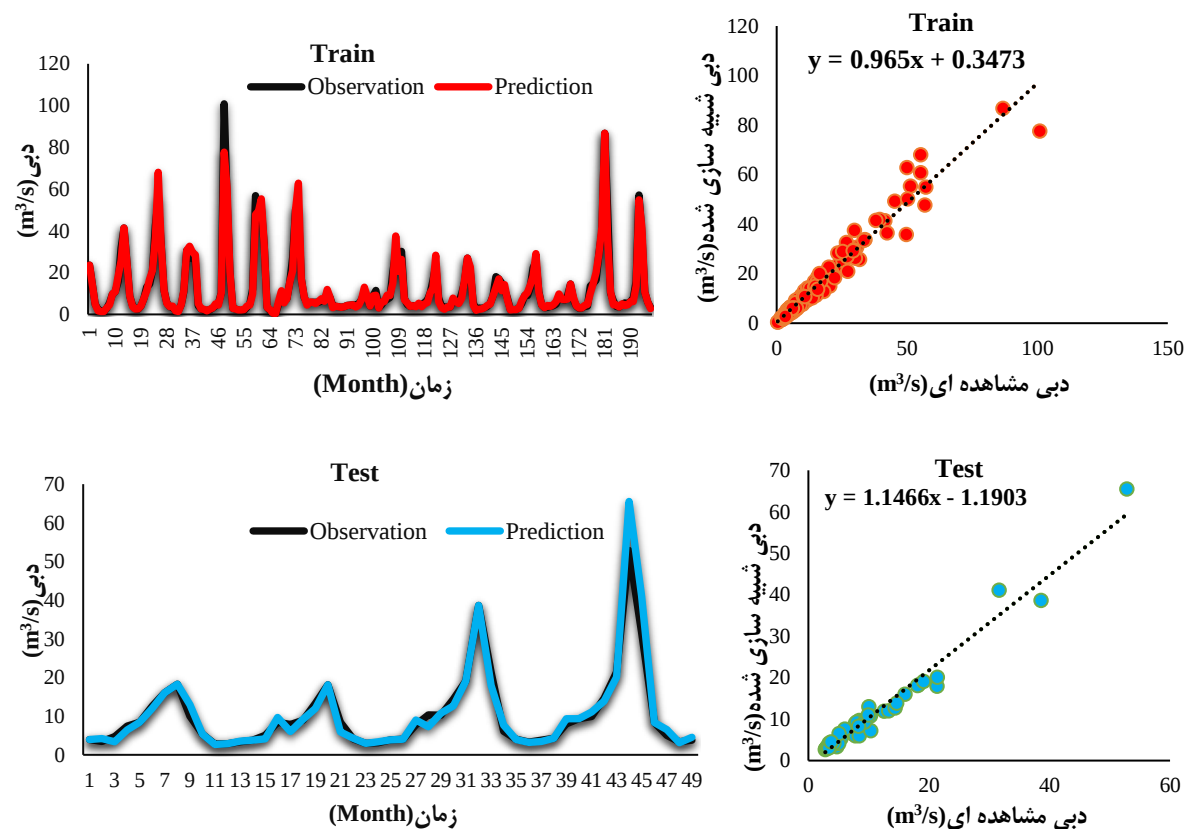
مدل	آموزش			تست		
	RMSE	NSE	R ²	RMSE	NSE	R ²
HSA-GMDH(D1)	0.2067	0.9247	0.9456	0.2028	0.9276	0.9464
HSA-GMDH(D2)	0.2236	0.9030	0.9390	0.2187	0.9036	0.9392
HSA-GMDH(D3)	0.2146	0.9121	0.9435	0.2118	0.9140	0.9445



شکل ۶. نتایج مدل HSA-GMDH(D1) برای مرحله آموزش و تست



شکل ۷. نتایج مدل HSA-GMDH(D2) برای مرحله آموزش و تست



شکل ۸. نتایج مدل HSA-GMDH(D3) برای مرحله آموزش و تست

۴. مقایسه نتایج مدل های GMDH و HSA-GMDH

بعد از توسعه مدل HSA-GMDH نتایج آن با مدل GMDH که پارامترهای آن با استفاده از روش سعی و خطا تعیین شدند مقایسه شد. بر اساس نتایج حاصل مدل HSA-GMDH(D1) نسبت به مدل GMDH(D1) عملکرد بهتری در مرحله آموزش و تست دارد به طوری که مدل HSA-GMDH(D1) نسبت به مدل GMDH(D1) مقدار RMSE کمتر و همچنین مقدار NSE و R^2 بیشتری دارد. مدل HSA-GMDH(D2) نیز نسبت به مدل GMDH(d2) عملکرد بهتری دارد و مقدار RMSE کمتر NSE و R^2 بیشتری از مدل GMDH(D2) دارد؛ مانند مدل های توسعه یافته بری مجموع داده های D1 و D2 مدل HSA-GMDH(D3) نیز نسبت به مدل GMDH(D3) عملکرد بهتری دارد، به طوری که مقدار RMSE آن نسبت به مدل GMDH(D3) کمتر و مقدار NSE و R^2 آن نسبت به مدل GMDH(D3) بیشتر است؛ بنابراین بر اساس نتایج حاصل از این مطالعه برای شبیه سازی جریان ماهانه رودخانه با استفاده از مدل GMDH بهتر است به جای روش سعی و خطا از یک الگوریتم بهینه سازی مانند HSA استفاده شود.

بحث

یکی از مهم ترین مسائل در مدیریت منابع آب تجزیه و تحلیل جریان رودخانه است. مدل سازی و پیش بینی دبی رودخانه از مهم ترین اهداف در فرایند مدیریت منابع آب است، به این دلیل همواره تلاش می شود از روش های دقیقی برای مدل سازی آن و پیش بینی آن استفاده کرد. برای توسعه مدل های فیزیکی علاوه بر هزینه زیاد، به انواع زیاد تجهیزات و همچنین کاربران خبره نیاز است؛ در نتیجه برای این کار می توان از مدل های فیزیکی و مدل های مبتنی بر داده استفاده کرد. هدف این مطالعه مدل سازی جریان ماهانه رودخانه گاماسیاب کرمانشاه برای یک دوره ۲۰ ساله است. برای این کار از شبکه عصبی GMDH استفاده شد.

نتیجه گیری

برای این کار از شبکه عصبی GMDH استفاده شد، در این مدل برای هر نورون یک چند جمله ای توسعه داده می شود، هر یک از این چند جمله ای ها وزن هایی دارند که باید با استفاده از یک روش دقیق تعیین شوند. از روش های مختلفی می توان برای این کار استفاده کرد اما در این مطالعه از الگوریتم HSA استفاده شد. برای توسعه این مدل از اطلاعات بارش و دبی در ماه های قبل استفاده شد. با استفاده از روش سعی و خطا سه مجموع داده تشکیل شدند که با استفاده از آن ها این مدل بهترین نتیجه را داشت، سپس برای هر مجموعه عملکرد هر مدل HSA-GMDH ارزیابی شد. برای ارزیابی کارایی مدل های مختلف از شاخص های RMSE، NSE و R^2 استفاده شد. بر اساس نتایج حاصل از توسعه مدل GMDH و HSA-GMDH برای مجموع داده های مختلف، مدل HSA-GMDH نسبت به مدل GMDH در مرحله آموزش و تست عملکرد بهتری دارد؛ بنابراین برای شبیه سازی جریان ماهانه رودخانه بهتر است به جای روش GMDH از روش HSA-GMDH استفاده شود.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

نویسندگان اصول اخلاقی را در انجام و انتشار این پژوهش علمی رعایت نموده اند و این موضوع مورد تأیید همه آنهاست.

تعارض منافع

بنا بر اظهار نویسندگان این مقاله تعارض منافع ندارد.

منابع

- اسماعیلی، یوسف، رجبی، احمد، یوسفوند، فریبرز، و شعبانلو، سعید. (۱۴۰۲). تخمین بارش درازمدت شهر بابلسر با استفاده از برنامه‌ریزی بیان ژن بهینه‌یافته. نشریه علمی پژوهشی مهندسی آبیاری و آب ایران، ۱۳(۲)، ۱۹۷-۲۱۵. <https://doi.org/10.22125/iwe.2022.162632>
- بذرافشان، ام البنین، سلاجقه، علی، مهدوی، محمد، و فاتحی، احمد. (۱۳۹۳). بررسی کارایی مدل‌های هیبرید شبکه عصبی مصنوعی-استوکاستیک در پیش‌بینی خشکسالی‌های هیدرولوژیکی با استفاده از آماره کاپا (مطالعه موردی: حوزه آبخیز رودخانه گاماسیاب). مجله علوم و مهندسی آبخیزداری ایران، ۸(۲۷)، ۳۵-۴۸. <http://dorl.net/dor/20.1001.1.20089554.1393.8.27.5.8>
- سبزواری، فرحناز، یعقوبی، بهروز، و شعبانلو، سعید. (۱۴۰۲). کاربرد مدل هیبریدی شبکه‌های عصبی مصنوعی و الگوریتم کرم شب‌تاب برای پیش‌بینی مقدار جامدات محلول در آب رودخانه. مجله حفاظت منابع آب و خاک، ۱۳(۲)، ۱۳-۲۳. <https://doi.org/10.30495/wsrcj.2023.71849.11356>
- علیدادی ده کهنه، صابر، سلگی، اباذر، شهنی دارابی، مهنوش، و زارعی، حیدر. (۲۰۱۹). ارزیابی مدل‌های ژنتیکی جهت مدل‌سازی جریان رودخانه. نشریه علمی پژوهشی مهندسی آبیاری و آب ایران، ۹(۳)، ۱-۱۷. <https://doi.org/10.22125/iwe.2019.88628>

References

- AlidadiDehkohne, S., Solgi, A., Shehni darabi, M., & Zarei, H. (2019). Evaluation of genetic models for river flow modeling. *Irrigation and Water Engineering*, 9(3), 1-17. <https://www.doi.org/10.22125/iwe.2019.88628> (In Persian)
- Aman Jalili, A., Najarchi, M., Shabanlou, S., & Jafarinia, R. (2023). Multiobjective optimization of water resources in real time based on integration of NSGA-II and support vector machines. *Environ Sci Pollut Res*, 30, 16464–16475. <https://doi.org/10.1007/s11356-022-22723-4>
- Amiri, S., Rajabi, A., Shabanlou, S., Yosefvand, F., & Izadbakhsh, M.A. (2023). Prediction of groundwater level variations using deep learning methods and GMS numerical model. *Earth Sci Inform*, 16, 3227–3241. <https://doi.org/10.1007/s12145-023-01052-1>
- Azizi, E., Yosefvand, F., Yaghoubi, B., Izadbakhsh, M.A., & Shabanlou, S. (2023) Modelling and prediction of groundwater level using wavelet transform and machine learning methods: A case study for the Sahneh Plain, Iran. *Irrigation and Drainage*, 72(3), 747–762. <https://doi.org/10.1002/ird.2794>
- Bazrafshan, O., Salajegheh, A., Mahdavi, M., & Fatehi, A. (2015). A Study of Efficiency of the Hybrid model Artificial Neural Network Models - Stochastic in Hydrological Drought Forecasting Using kappa Statistics (Case Study: Gamasiab Watershed Basin). *Iranian Journal of Watershed Management Science and Engineering*, 8(27), 35-48. <http://dori.net/dor/20.1001.1.20089554.1393.8.27.5.8> (In Persian)
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific model development*, 7(3), 1247-1250. <http://dx.doi.org/10.5194/gmd-7-1247-2014>
- Chen, C., He, W., Zhou, H., Xue, Y., & Zhu, M. (2020). A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River Basin, northwestern China. *Scientific reports*, 10(1), 3904. <https://www.nature.com/articles/s41598-020-60698-9>
- Esmaili, F., Shabanlou, S., & Saadat, M. (2021). A wavelet-outlier robust extreme learning machine for rainfall forecasting in Ardabil City, Iran. *Earth Sci Inform*, 14, 2087–2100. <https://doi.org/10.1007/s12145-021-00681-8>
- Esmaili, Y., rajabi, A., yosefvand, F., & shabanlou, S. (2022). Estimation of Long-Term Rainfall in Babolsar City by Using the Optimized Gene Expression Programming. *Irrigation and Water Engineering*, 13(2), 197-215. <https://www.doi.org/10.22125/iwe.2022.162632> (In Persian)
- Fallahi, M.M., Shabanlou, S., Rajabi, A., Yosefvand, F., & IzadBakhsh, M.A. (2023). Effects of climate change on groundwater level variations affected by uncertainty (case study: Razan aquifer). *Appl Water Sci*, 13(143). <https://doi.org/10.1007/s13201-023-01949-8>
- Farlow, S. J. (2020). Self-organizing methods in modeling: GMDH type algorithms. CrC Press, Boca Raton, Florida. <https://doi.org/10.1201/9781003065227>
- Geem, Z. W. (2009). Music-inspired harmony search algorithm: theory and applications. Springer, Berlin/Heidelberg, Germany. <http://dx.doi.org/10.1007/978-3-642-00185-7>
- Guamel, I. A., & Lee, H. S. (2020). Watershed modelling of the Mindanao River Basin in the Philippines using the SWAT for water resource management. *Civil Engineering Journal*, 6(4), 626-648. <http://dx.doi.org/10.28991/cej-2020-03091496>
- Jalilian, A., Heydari, M., Azari, A., & Shabanlou, S. (2022) Extracting optimal rule curve of dam reservoir base on stochastic inflow. *Water Resour Manage*, 36, 1763–1782. <https://doi.org/10.1007/s11269-022-03087-3>

- Kambalimath S, S., & Deka, P. C. (2021). Performance enhancement of SVM model using discrete wavelet transform for daily streamflow forecasting. *Environmental Earth Sciences*, 80(3), 101. <https://link.springer.com/article/10.1007/s12665-021-09394-z>
- Mahdavi, M., Fesanghary, M., & Damangir, E. (2007). An improved harmony search algorithm for solving optimization problems. *Applied mathematics and computation*, 188(2), 1567-1579. <http://dx.doi.org/10.1016/j.amc.2006.11.033>
- Mazraeh, A., Bagherifar, M., Shabanlou, S., & Ekhlasmand, R. (2023). A Hybrid Machine Learning Model for Modeling Nitrate Concentration in Water Sources. *Water, Air, & Soil Pollution*, 234(11), 1-22. <http://dx.doi.org/10.1007/s11270-023-06745-3>
- Mazraeh, A., Bagherifar, M., Shabanlou, S., & Ekhlasmand, R. (2024). A novel committee-based framework for modeling groundwater level fluctuations: A combination of mathematical and machine learning models using the weighted multi-model ensemble mean algorithm. *Groundwater for Sustainable Development*, 24, 101062. <http://dx.doi.org/10.1016/j.gsd.2023.101062>
- McCuen, R. H., Knight, Z., & Cutter, A. G. (2006). Evaluation of the Nash–Sutcliffe efficiency index. *Journal of hydrologic engineering*, 11(6), 597-602. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:6\(597\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597))
- Moradi, E., Yaghoubi, B., & Shabanlou, S. (2023). A new technique for flood routing by nonlinear Muskingum model and artificial gorilla troops algorithm. *Appl Water Sci*, 13, 49. <https://doi.org/10.1007/s13201-022-01844-8>
- Ozer, D. J. (1985). Correlation and the coefficient of determination. *Psychological bulletin*, 97(2), 307. <https://psycnet.apa.org/doi/10.1037/0033-2909.97.2.307>
- Panahi, J., Mastouri, R., & Shabanlou, S. (2022). Insights into enhanced machine learning techniques for surface water quantity and quality prediction based on data pre-processing algorithms. *J Hydroinf*, 24(4), 875–897. <https://doi.org/10.2166/hydro.2022.022>
- Rezaei, K., & Vadiati, M. (2020). A comparative study of artificial intelligence models for predicting monthly river suspended sediment load. *Journal of Water and Land Development*. 45(IV-VI), 107-118. <http://dx.doi.org/10.24425/jwld.2020.133052>
- Sabzevari, F., Yaghoubi, B., & Shabanlou, S. (2023). Application of Hybrid Model of Artificial Neural Networks and Firefly Algorithm to Predict the Amount of TDS in River Water. *Journal of Water and Soil Resources Conservation*, 13(2), 13-23. <https://doi.org/10.30495/wsrcj.2023.71849.11356> (In Persian)
- Samadianfard, S., Jarhan, S., Salwana, E., Mosavi, A., Shamshirband, S., & Akib, S. (2019). Support vector regression integrated with fruit fly optimization algorithm for river flow forecasting in Lake Urmia Basin. *Water*, 11(9), 1934. <http://dx.doi.org/10.20944/preprints201905.0361.v1>
- Shabanlou, S. (2018). Improvement of extreme learning machine using self-adaptive evolutionary algorithm for estimating discharge capacity of sharp-crested weirs located on the end of circular channels. *Flow Measurement and Instrumentation*, 59, 63-71. <https://www.doi.org/10.1016/j.flowmeasinst.2017.11.003>
- Xiang, X., Li, Q., Khan, S., & Khalaf, O. I. (2021). Urban water resource management for sustainable environment planning using artificial intelligence techniques. *Environmental Impact Assessment Review*, 86, 106515. <http://dx.doi.org/10.1016/j.eiar.2020.106515>
- Yang, W., Wang, K., & Zuo, W. (2012). Fast neighborhood component analysis. *Neurocomputing*, 83, 31-37. <http://dx.doi.org/10.1016/j.neucom.2011.10.021>
- Yang, X.-S. (2009). Harmony search as a metaheuristic algorithm. Music-inspired harmony search algorithm: theory and applications, *Springer*, Berlin/Heidelberg, Germany. http://dx.doi.org/10.1007/978-3-642-00185-7_1