Computers & Geosciences 64 (2014) 136–151

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Comparing large number of metaheuristics for artificial neural networks training to predict water temperature in a natural river

Adam P. Piotrowski [a,*], Marzena Osuch [a], Maciej J. Napiorkowski [b], Pawel M. Rowinski [a], Jaroslaw J. Napiorkowski [a]

[a] Institute of Geophysics, Polish Academy of Sciences, Ks. Janusza 64, 01-452 Warsaw, Poland
[b] Environmental Engineering Faculty, Warsaw University of Technology, Poland

ARTICLE INFO

Article history:
Received 26 September 2013
Received in revised form
6 December 2013
Accepted 20 December 2013
Available online 4 January 2014

Keywords:
Metaheuristics
Streamwater temperature forecasting
Climate change
Artificial neural networks
Differential evolution
Particle swarm optimization

ABSTRACT

Nature-inspired metaheuristics found various applications in different fields of science, including the problem of artificial neural networks (ANN) training. However, very versatile opinions regarding the performance of metaheuristics applied to ANN training may be found in the literature.

Both nature-inspired metaheuristics and ANNs are widely applied to various geophysical and environmental problems. Among them the water temperature forecasting in a natural river, especially in colder climate zones where the seasonality plays important role, is of great importance, as water temperature has strong impact on aquatic life and chemistry. As the impact of possible future climate change on water temperature is not trivial, models are needed to allow projection of streamwater temperature based on simple hydro-meteorological variables.

In this paper the detailed comparison of the performance of nature-inspired optimization methods and Levenberg–Marquardt (LM) algorithm in ANNs training is performed, based on the case study of water temperature forecasting in a natural stream, namely Biala Tarnowska river in southern Poland. Over 50 variants of 22 various metaheuristics, including a large number of Differential Evolution, as well as some Particle Swarm Optimization, Evolution Strategies, multialgorithms and Direct Search methods are compared with LM algorithm on ANN training for the described case study. The impact of population size and some control parameters of particular metaheuristics on the ANN training performance are verified. It is found that despite widely claimed large improvement in nature-inspired methods during last years, the vast majority of them are still outperformed by LM algorithm on the selected problem. The only methods that, based on this case study, seem competitive to LM algorithm in terms of the final performance (but not speed) are Differential Evolution algorithms that benefit from the concept of Global and Local neighborhood-based mutation operators. The streamwater forecasting performance of the neural networks is adequate, the major prediction errors are related to the river freezing and melting processes that occur during winter in the mountainous catchment under study.

1. Introduction

Nature-inspired computing methods are plentiful today (De Castro, 2007). Some of them are developed to simulate natural phenomena (e.g. Artificial Life, see Farmer and Belin, 1992), the others serve as regression or classification models (e.g. Artificial neural networks (ANN) (Haykin, 1999)) or found wide applicability as optimization tools (e.g. Evolutionary Algorithms). Different types of nature-inspired methods may be applied together, for example Evolutionary Algorithms are sometimes used to find ANN architecture or to calibrate ANN parameters. Both ideas have been discussed in numerous studies over the last two decades (Whitley et al., 1990; Yao, 1993, 1999; Branke, 1995; Yao and Liu, 1997; Cantu-Paz and Kamath, 2005; Ding et al., 2013) and the second one is the main topic of the present paper. For the discussion on application of nature-inspired methods to create optimal ANN architecture the interested reader is referred to the above papers and the studies by Islam et al. (2009) or Hunter et al. (2012).

Despite the popularity of the nature-inspired computing, the ANNs are usually trained by means of the classical gradient-based algorithms. The nature-inspired optimization methods are relatively rarely applied for such task – what was noted by Branke in 1995 and is still valid today. The reason is that the gradient-based and Newton′s methods, or the approaches that merge the advantage of both like Levenberg–Marquardt algorithm (LM) (Hagan and Menhaj, 1994), are quick, easy to implement and frequently lead to

* Corresponding author. Tel.: +48 22 6915 858; fax: +48 22 6915 915.
E-mail address: adampp@igf.edu.pl (A.P. Piotrowski).

0098-3004/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cageo.2013.12.013

good enough solutions. However, it is commonly claimed that other optimization methods may be needed to train ANNs as: (1) the gradient or Newton's-based methods are not global search approaches and are prone to stick in a local optimum, (2) in some cases the objective function used to train ANNs may be non-differentiable. Hence, the temptation to train nature-based ANN models by means of nature-based optimizers is easily understandable. Moreover, the problem of ANN training is even sometimes suggested as a good test to assess the performance of different metaheuristics (He et al., 2009).

Among a large number of various ANN applications, the geophysical and environmental ones, like evaluation of the ozone profile in the atmosphere (Jaroslawski, 2013), air quality modeling (Hoi et al., 2013), groundwater level forecasting (Shiri et al., 2013), modeling pollutant transport in rivers (Piotrowski et al., 2007), suspended sediment estimation (Kisi and Shiri, 2012) or precipitation estimation (Hsu et al., 1997) are frequently of great importance for practitioners. This study aims at the application of both ANNs and metaheuristics to forecasting, or projection, of water temperature in a natural river. The methodology is applied to the case study of the Biala Tarnowska river located in the southern Poland, in the Polish part of the Carpathian Mountains. The selected river is a right-bank tributary of the Dunajec River. It is a typical mountainous river with slopes up to 8.6‰ in the upper part and about 0.9‰ in the lower part. The area of the catchment above Koszyce Wielkie gauging station is 956.9 km^2. Biala Tarnowska river is characterized by large fluctuations in water levels (up to 8 m in the lower reaches), what increases the difficulty of water temperature prediction.

Water temperature in natural rivers may vary due to natural or semi-natural processes (Poole and Berman, 2001; Caissie, 2006) or due to human-made thermal pollutions (Vega et al., 1998; Kalinowska et al., 2012); the first issue is of interest in the present paper. Forecasting water temperatures of natural rivers without autoregressive inputs is of large importance for aquatic life and chemistry (St-Hilaire et al., 2012). Models are frequently needed to evaluate the temperature in the natural rivers mostly due to two reasons: (1) the water temperature is often not measured on site, (2) they are essential in assessment of the possible impact of climate change on the future biological and chemical processes in rivers (Jeong et al., 2012). Note that the relation between climate change and the change of water temperature in natural rivers is not trivial (van Vliet et al., 2011). Among deterministic (Caissie et al., 2007) or simple data-driven models (St-Hilaire et al., 2012; Grbic et al., 2013) ANNs are well established as a promising modeling tool for such task (Sahoo et al., 2006; Chenard and Caissie, 2008; Daigle et al., 2009) and are claimed to outperform other empirical model types (Sahoo et al., 2009).

Since the work by Holland (1975), from among nature-inspired metaheuristics Evolutionary Computation (EC), i.e. methods inspired by biological evolution, have become extremely popular. Plenty of optimization approaches pertain to the concept of EC today (see e.g. the studies of Onwobulu and Babu (2004) and Boussaid et al. (2013)). Apart from EC, there is also a fast growing community of optimization algorithms inspired by the behavior of the swarms of animals (including Particle Swarm Optimization (PSO), Eberhart and Kennedy, 1995, and Ant Colony Optimization (ACO), Dorigo et al., 1996), as well as by other biological (for example Biogeography-based Algorithm, Simon, 2008) or non-biological phenomena (e.g. Simulated Annealing, Kirkpatrick et al., 1983, Gravitation Search Algorithm, Rashedi et al., 2009, or Harmony Search, Geem et al., 2001). But one must acknowledge that not all novel metaheuristics turn out successful and useful for the practical problems (Crepinsek et al., 2012) and that the attention should be paid to possible pitfalls when proposing the novel methods or using the already known ones to solve novel tasks (Weise et al., 2012). It should also be noted that large number of modern metaheuristics in fact share similarities with the older, well-known Direct Search methods like Nelder–Mead simplex (NMA) (Nelder and Mead, 1965), the Rosenbrock algorithm (RA) (Rosenbrock, 1960) or the Controlled Random Search (Price, 1977). To summarize, nowadays one is equipped with a multitude of methods, but the choice of the optimization metaheuristics that allow for successful and efficient solution of a particular problem is usually not trivial. It is hoped that the present paper may help to find the best metaheuristics, or at least to filter out the ones that are not promising in ANN training for geophysical and environmental applications.

Since the difference between nature- and non-nature-inspired metaheuristics is rather vague, there is little theoretical or practical reason to restrict the attention to the nature-based methods only, even if their popularity is in the climax today. Although different metaheuristics have been applied to ANN training in the past, in large number of papers, including geophysical or environmental ones, only a single optimization method is used, without a broad comparison with other metaheuristics or methods that use the information about relevant derivatives. However, as the search for proper metaheuristics for ANN training started in the 1980s, a number of papers from different fields of science that perform a rigorous comparison must be mentioned, even if: (1) in most of them the number of compared optimization methods is too low to draw more general conclusions and (2) different studies frequently lead to contradictious findings. For example Schaffer (1994) and Alba and Chicano (2004) showed that the selected Genetic Algorithms (GA) (Holland, 1975) perform poorer than the back-propagation method, whereas research of Montana and Davis (1989), Sexton et al. (1998), Sexton and Gupta (2000) and Huang et al. (2009) led to the opposite conclusion. A discussion on the differences among GA's used in some of these papers, which could clarify the contradictory conclusions, may be found in Montagno et al. (2002), the broader debate on application of GA's to ANN training is given in classical studies by Whitley et al. (1990) and Yao (1999).

The application of the so-called Memetic Algorithms (MAs) (Krasnogor and Smith, 2005), i.e. methods that are composed of a global search metaheuristic and a local search method (that may, or may not use the information about respective derivatives) to ANN training seems less disputable. Alba and Chicano (2004) showed that the hybridization of GA with LM may lead to better results than using LM, back-propagation or GA separately when ANNs are trained for classification problems. Delgado et al. (2006) claimed the superiority of the newly proposed MA, composed of GA and a quasi-Newton local search operator, over a few gradient-based methods for recurrent ANNs. Socha and Blum (2007) showed that both ACO algorithm and GA perform poorer than LM or the simple back-propagation algorithm when applied to ANN training for classification problems, but the MA composed of ACO and LM may lead to some improvement over LM or back-propagation algorithms used alone. Also Zhang et al. (2007) and Han and Zhu (2013) claimed that the MA created by hybridizing PSO with the back-propagation outperforms its both internal algorithms used separately for ANN training. Although sometimes the ANNs are trained in two steps, first by global search heuristic and then the solution found is further refined by gradient-based method (e.g. Cannon, 2012), such approach is considered highly inefficient and cannot be considered a kind of MA paradigm. It must also be added, that when the local search method is gradient-based, the whole MA approach looses some of its practical importance for ANN training, as it cannot be used for non-differentiable functions.

ANN training by means of other metaheuristics leads to versatile conclusions, like in the case of GA's. Leerink et al. (1995) and Ismail