



Research paper

Numerical correlation of speleothem stable isotope records using a genetic algorithm



J. Pawlak*, H. Hercman

Institute of Geological Sciences, Warsaw Research Centre, Polish Academy of Sciences, Twarda 51/55, 00-818 Warsaw, Poland

ARTICLE INFO

Article history:

Received 21 April 2014

Received in revised form

20 December 2015

Accepted 22 December 2015

Available online 8 January 2016

Keywords:

Genetic algorithms

Correlation

Stable isotopes

Speleothems

ABSTRACT

The correlation of geological records is commonly based on subjective and arbitrary interpretation by a researcher. Geological data are usually collected as a series on a depth scale, and then chronologies of the profiles are constructed. The correlation of geological records based on matching quality indicators requires an efficient method of searching for the best position. *GenCorr* is a software based on a genetic algorithm, for correlation of records with or without accompanying chronological control. Where their age information is available, points on both records may be moved in the range limited by the age uncertainties. *GenCorr* has been developed by the authors in Borland RAD Studio 2007 environment. The authors tested the impact of matched record length, noise level, sampling resolution, or additional linear components and hiatuses on the correlation results using artificial test data. Finally, the genetic algorithm was applied to stable isotope records obtained from two stalagmites collected from the Demianova Cave System (Low Tatras, Slovakia). The results show that the genetic algorithm correlation method is similar to that used in pattern recognition by the human eye without prior knowledge of the ages. The genetic algorithm is able to correlate in multidimensional space and is free from human biases, but we accept that there are also occasions where additional expert opinion may be also important. *GenCorr* can be run on every PC with MS Windows operating system, the package was not tested in Mac environment. *GenCorr* is available as a freeware from the authors. For downloading, please contact: dzeq@twarda.pan.pl.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The basic problem in the comparison of geological data is determination of the age relationship between profiles. In some cases, numerical methods can be applied for age estimation of a few layers in the compared profiles. However, in many cases, where ages exceed the range of a dating method, or where the chronology is compromised by open system behaviour and/or post-depositional contamination by daughter or parent isotopes, age-depth modelling is not possible. In these cases, the correlation is the only possibility for age estimation. Several numerical method approaches for correlation problem have been published. The dynamic programming method has been used to estimate transfer functions of probability density functions for speleothems dated by radiocarbon and U-series methods (Goslar et al., 2000). Lisiecki and Lisiecki (2002), when correlating marine palaeoclimatic records,

used a dynamic programming technique to produce records on a common depth scale. Haam and Huybers (2010) used the Monte Carlo method to search for the match with the best covariance between a $\delta^{18}\text{O}$ record from Dongge Cave speleothem and the Intcal98 $\Delta^{14}\text{C}$ records. Based on Monte Carlo methods, Fohlmeister (2012) proposed an algorithm that uses ages from more than one speleothem from the same cave system to establish a single more accurate age-depth model. Spötl et al. (2006) did the same using non-parametric smoothing splines after identifying common nodes in two records. Mudelsee et al. (2012) used the Iscam software (Fohlmeister, 2012) and Stal-Age (Scholz and Hoffmann, 2011) to create the paleoenvironmental record taking into account the impact of age uncertainties. However, in many cases, the correlation of geological records is still a subjective and arbitrary interpretation by a scientist. In contrast, dendrochronology and varve chronology use correlation based on objective, numerical indicators. There is a strictly defined factor for matching quality control (Baillie, 1995). Dendrochronology is based on correlating the sequences of tree rings from tree trunks belonging to the same

* Corresponding author.

E-mail address: dzeq@twarda.pan.pl (J. Pawlak).

species. This method is accurate because a living tree produces one tree ring every year and hiatuses are rare. Therefore, the tree ring sequence is equivalent to the time scale. Typical geological records contain hiatuses. Sedimentation rates depend on many environmental factors, such as local relief, precipitation, temperature, and vegetation activity. For these reasons, sedimentation rates usually differ for the layers deposited in different areas during the same time interval, and the depth scale is not uniform over time.

In this study, the authors propose the genetic algorithm as a tool for correlating the records. Genetic algorithms are a general method that may be used for solving optimization problems, as well as a very efficient searching technique (Holland, 1975; Goldberg, 1991; Back, 1996). Genetic algorithms have been used previously for solving different problems in geological sciences. For example, Gleadow and Brown (1999) applied the genetic algorithm to solve the problem of reconstruction of the cooling history of rocks using FT methods. Another application of the genetic algorithm into the geology was the correlation of fossils records (Zhang and Plotnick, 2006). In this paper, the authors present a genetic algorithm correlation method and test it using simulated and observed climate data sets. For the latter, the authors use stable isotope data ($\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values) from cave speleothems, which are widely used for high-resolution palaeoclimatic reconstructions (Baker et al., 1993; Frisian et al., 2003; Roberts et al., 1998; Hu et al., 2008; Yuan et al., 2004; Wang et al., 2001, 2005). Additionally, speleothems can be reliably dated by U-series methods (Richards & Dorale, 2003; Hercman and Goslar, 2002) and these can be used for verification of the correlation results obtained by genetic algorithms. The approach presented in this paper could be used to correlate any records of geological data based on sediment, mineral, or ice accumulation.

2. Methods

Correlation of geological records without a time scale requires the possibility of taking different depositional rates into account. Fig. 1 presents the basic idea of geological record correlation using simulated proxy climate records. Fig. 1A.1 presents a hypothetical record of climatic change (K) in a time scale (T) that is recorded in two profiles (Fig. 1A.2), both of which exhibit different growth rates in a depth scale (D). Finding the proper position between the records from these two profiles is impossible if the procedure of correlation is based only on the movement of one record relative to the other (Fig. 1A.3). Only the independent movement of every point allows us to find the correct position (Fig. 1A.4). Such a procedure requires an efficient method, such as genetic algorithms, to explore the huge set of possible positions.

2.1. Genetic algorithm – basic concept

A genetic algorithm is a simulator of biological evolution (Holland, 1975) and allows the use of evolutionary rules to solve complex optimisation problems. Possible solutions are coded as a population of individuals, which are represented as sequences of bits. One bit in the individual sequence is defined as a gene. The individuals in the population are different; some individuals are closer to the optimal solution than others. The population is processed by the genetic algorithm in two stages: selection and replication (Fig. 1B). The selection of individuals is based on their heuristic function values, the shape of which depends on the problem to be solved. Individuals with better heuristic function values have a higher chance of survival than the other individuals. There are a few possible methods of selection for use in a genetic algorithm (Back, 1996). The most popular and effective are the tournament selection and the roulette selection (Miller and

Goldberg, 1995). Simply, the tournament selection procedure works in two stages: (1) two or more individuals are selected from the processed population randomly, (2) an individual with the best heuristic function value is chosen from the two individuals selected in stage 1. In the roulette method, individuals are placed on a roulette wheel. The surface, which is occupied by one individual, depends on the heuristic function value for this individual. Individuals with better values occupy more space and have a greater chance of being chosen for replication. The roulette method is more random than the tournament method. The tournament method is considered to be the most effective method of selection. The selected individuals are then replicated. Replication involves the creation of a new population based on the selected individuals (Holland, 1975; Goldberg, 1991; Back, 1996). Replication is realised by crossing and mutation (Fig. 1B). During crossing, new individuals are created from genes taken from the parent individuals. There are several crossing methods. The most popular is the uniform crossing method. In this method, every gene from a new individual has a chance to take its value from either one parent or the other (Brad and David, 1995). Crossing only mixes genes, but does not change their values. For altered gene values, we use the mutation methods. Mutation adds noise to one randomly selected gene. This noise is described by a normal distribution with 0 as a mean value. After several processing repetitions of the population cycle, individuals with better heuristic function values are created. After many repetitions, the chance of creating an individual that represents the optimal solution increases. Theoretically, the genetic algorithm can work to infinity. Practically, stopping the genetic algorithm at the optimal moment is necessary. There are two basic ways to stop a run of a genetic algorithm: after an assigned number of cycles or when the evolutionary progress is smaller than a minimum accepted value.

2.2. Procedure of correlation by genetic algorithms in GenCorr software

GenCorr software has been developed in the CodeGear™ RAD Studio 2007 Version 11.0.2902.10471 environment. It can be run on PC with any Windows operating system. GenCorr requires records data files in text format. The files should contain four columns separated by tabs. The first column is reserved for age, depth, or distance from the base data, the second one for paleoenvironmental data, the last two columns are reserved for the data error (1 sigma). It is assumed that the first record is the base record (with a known age–depth relationship) and the second is a matched record (with an unknown position). The base record should represent a longer time interval than the matched record. Our procedure sets a matched record into the time scale of the base record using the genetic algorithm. The genetic algorithm works by processing a population of individuals. In this correlation procedure, an individual is defined as one possible position of a matched record relative to the base record age scale. One gene describes the position of a point on the matched record in the age scale of the base record. The genetic algorithm creates an initial population at the beginning of the process. A large diversity of individuals at the beginning stage is important to the proper functioning of the genetic algorithm. To generate as much diversity as possible, two methods are used for the creation of the initial population. The first method is based on drawing a sequence of genes with the length of the matched record. The values of the drawn genes are between $-0.2*L$ and $1.2*L$, where L is a length of time from the base record. The drawn genes are sorted according to the superposition rule (i.e., points located deeper in the profile are older). Finally, an individual that describes the position of the matched record in the time scale of base record is created. This method usually creates

Download English Version:

<https://daneshyari.com/en/article/4724857>

Download Persian Version:

<https://daneshyari.com/article/4724857>

[Daneshyari.com](https://daneshyari.com)