Rapid assessment of populations trends of invasive species: Singular Spectrum Analysis (SSA)

Elias D. DANA

Abstract

Singular Spectrum Analysis (SSA) is a powerful analytical approach for biodiversity management. Its main advantages are due to its intuitive processing and visualization, since mathematical workflow is conceptually similar to the widely accepted Principal Components Analysis. Detailed analyses of population trends with mathematical tools are often difficult to achieve for managers by a number of reasons (large numbers or areas monitored, large number of species, insufficient statistics skills, strong knowledge level in demographic analyses, etc.). SSA has been used since the 1970's in signal processing to clarify signal vs. noisy information, but it has also been used in climate change analysis and other developmental areas. Besides, SSA is a rapid-learning method for technicians and managers with medium level of mathematical knowledge. Free software in Unix environment is available. Unfortunately, no free and friendly software is available for Windows SO. Although R package may offer solutions for really advanced users, it does not fit real work situations for managers of biological invasions. Caterpillar (Gistat Group, Ltd) is by now, the best option found by the author in terms of price, facility for results interpretation and time consumed in learning. The main disadvantage is the poor content of tutorial files.

Keywords

Trend analysis, alien species, exotic species, population analysis

Introduction

Population estimates of invasive species are often difficult to obtain due to particular conditions in many field and management situations (staff availability and costs, biological or site accessibility difficulties, etc.). Wildlife agencies, non-governmental organizations or research groups, usually face three problems, 1) exceeding the number of localities and ecological complexity, 2) repetition of census or population estimates to obtain temporal variations, 3) lack of specific and rapid analytical tools that allow obtaining results in an acceptable balance between time invested and the quality of the results. In this paper a different procedure based on solid mathematical principle is presented and the main advantages for implementation in management practices are provided.

What is Singular Spectrum Analysis?

Singular Spectrum Analysis (SSA) is a vectors-based method for analysis of time series of data. Types of data that can be subjected to this method are varied: repeated census data or population abundance estimates, fishing yields, fluctuations in biocoenoses indexes (such as Shannon's index, species richness), climate parameters and observational data aroused from volunteers' nets of observation in permanent stations, visitants' perception of management in natural areas, etc. Variation of any type of ecological or social indicator that can be represented along a time-axis (years, months, days) may be good candidates to be analyzed by SSA. As in any time-series analysis, the only pre-requisites to get a robust result are: 1) the series of data must be complete, i.e., the matrix that will be subjected to analyses must not have missed observations, 2) numerical data must have been taken in a similar way.

Point 2 is always controversial in any scientific analysis and one of the most critical aspects when multiyear researches are conducted. Consequently, interpretation of management results and their robustness are therefore often limited by the facts that, a) resource availability and management priorities frequently fluctuate in the medium term; b) resource

DOI: http://dx.doi.org/10.3391/mbi.2010.1.1.05

availability at a given time may determine the monitoring method and field sampling aspects.

Additionally, another frequent problem managers and technicians may have to deal with is knowledge acquisition and implementation. This happens in general, due to: 1) the need for multi-scale approaches, including monitoring of habitats, species and social uses of the territories involved, which increase difficulties to concentrate efforts and time to analyze specific management actions in a monographic approach, 2) the width of areas managed or 3) environmental urgencies that must be rapidly attended to.

As a consequence, absent data or methodological variations (for instance, effort invested in number of field data sampled) for data sampling can be common in time series analysis. Consequently, classic methods based on descriptive statistics are generally not suitable for interpretation of management results. Classic time series analyses usually require ordinary assumptions in data distribution that are hardly found in practice. However, conceptual understanding of mathematical subjacent principles and purchase of useful software may be difficult for many institutions or managers. SSA however, allows statistical analysis without excessive assumptions on data behaviour, analytical steps and interpretation are simpler when conducted by the appropriate software. SSA is an improved technique, compared to the classical approaches based on Fourier series transformations.

When variations in effort sampling are estimated for a given time series data by using simple data such as number of days invested, and optimally, the dependence relationship between the field data and effort is known, results obtained will be more robust. However, this may not be compulsory in all situations. For instance, partial field sample data may be complemented by estimates from other sources (such as fishermen or hunters' observations) to build the initial matrix that the analyst will work with.

The main use of SSA in the field of management includes situations in which time series variation is mainly influenced by seasonal change and/or by the management imposed to the habitat, species, etc. Chaotic non-deterministic changes can also be filtered by using this method. In all these cases, interpretation from a simple graphical drawing is very difficult. Increments and decrements within the series appear, and may be due to periodical (i.e., due to phenological reasons), catastrophic, and to variations in management conducted through months or years.

Resumen

El Análisis de Espectro Singular (Singular Spectrum Analysis or SSA) es una potente aproximación analítica para el manejo de la biodiversidad. Sus principales ventajas se centran en que, al basarse en álgebra matricial, los pasos para realizar este tipo de análisis son conceptualmente semejantes a los bien conocidos Análisis de Componentes Principales. Los análisis detallados de tendencias poblacionales con métodos matemáticos son habitualmente difíciles de realizar por los gestores debido a numerosas razones como (amplio número de áreas en seguimiento, altos números de especies, necesidades de altos conocimientos en análisis demográficos, etc.). El Análisis de Espectro Singular se ha empleado desde los años 70 en el análisis de señales para aclarar el balance entre información de señal vs. ruido, estudios climáticos y otras áreas de desarrollo. Se trata de un método cuyo aprendizaje resulta comparativamente rápido para técnicos y gestores con nivel medio de conocimientos matemáticos. Existe disponible Software Gratuito (Freeware) de aplicación directa para Unix, pero no para entornos Windows. Los paquetes en R son útiles solo para usuarios con conocimientos avanzados en programación en R. *Caterpillar* (Gistat Group, Ltd) es, por ahora, la mejor opción encontrada por el autor por precio y facilidad de manejo, aunque el tutorial incluido es pobre e insuficiente.

Palabras clave

Análisis de tendencias, alóctonas, exóticas, análisis de poblaciones

SSA allows a manager: a) to decompose the effects caused by these three processes, b) estimate their relative percentage importance, c) detect and inspect the specific underlying general trend. It can decompose a short, noisy time series into a (variable) trend, periodic oscillations, other statistically significant components that are aperiodic, and noise.

For general users it seems to be easier to learn, since its flow work is similar to the classical Principal Components Analysis. SSA is a well known method in other technical fields, such as digital signal processing, climate change analyses and nonlinear dynamics during the 1970's to 1990's (see Ghil & Vautard 1991, Vautard et al. 1992 and references therein included). A complementary advantage is that forecasting is possible, although this must be used with caution. As with any predictive module, predicted results obtained from forecasting should be considered more as a general prediction than as an infallible prediction. Simple general information for readers not familiar with the methods can be obtained from internet search. More specific information can be freely accessed from Loeuille and Ghil (2004). Finally, the book by Golyandina et al. (2001) reflects a great effort made by authors to synthesize concepts and present the techniques in an easier manner than that usually found in mathematics books.

Numerical data obtained from managed invasive populations are not essentially different to other data types for which these methods have proved to be robust. However, the technique is still underexploited for population trend analysis or other biological fluctuations.

What about free software?

Despite the time since the launch of SSA, the only free software for SSA and other related analytical methods is Spektra, provided by SSA -MTM Group, University of California (and participation of researchers from other institutions) at the site http://www.atmos.ucla.edu/tcd/ssa/. The Toolkit freely offered has been ported to SUN, Linux, DEC, SGI, and to Mac OS X. This limits potential use for users of other Operative Systems. Unfortunately, although R-package may offer solutions for really advanced users, it does not fit real work situations for managers of biological invasions due to its complexity and is time consuming to learn programming commands. The other existing software tested by the author is 'Caterpillar', developed by Gistat Group Ltd at a reasonable price, especially if it is compared to the prices of common Statistical Packages. This program is the best option found by the author in terms of price, facility for result interpretation and time spent in learning. The main disadvantages are the poor content of tutorial files.

References

Ghil M, Vautard M (1991) Interdecadal oscillations and the warming trend in global temperature time series. Nature 350 (28 March): 324-327.

Golyandina N, Nekrutkin V, Zhiglhavsky A (2001) Analysis of time series structure: SSA and related techniques. Chapman & Hall/CRC, Boca Ratón.

Loeuille N and Ghil M (2004) Intrinsic and climatic factors in North-American animal population dynamics. BMC Ecology 4:6 (http://www.biomedcentral.com/1472-6785/4/6, accessed 23 December 2009)

Vautard R, Yiou P, Ghil M (1992) Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. Physica D 58 95-126.

Author

DANA, Elias D.

edana@ual.es

Grupo Investigación Transferencia I+D en el Área de Recursos Naturales. Universidad de Almería (Almeria, España)