Journal of Marine Systems
February 2014, Volume 130, Pages 1–3
http://dx.doi.org/10.1016/j.jmarsys.2013.11.010

Time series analysis of marine data: A key knowledge at the crossroads of marine sciences

I. Puillat[a, *], M. Prevosto[a], H. Mercier[b], S. Thomas[c]

[a] Ifremer Centre de Brest, REM/RDT, CS 10070, 29280 Plouzané, France
[b] CNRS, UMR6523 LPO, Ifremer Centre de Brest, CS 10070, 29280 Plouzané, France
[c] IUEM-European Institute for Marine Studies, Rue Dumont Durville, 29280 Plouzané, France

*: Corresponding author : Ingrid Puillat, Tel.: + 33 2 29 00 85 09. ; email address : ingrid.puillat@ifremer.fr

Keywords: Time series ; Marine data ; Signal processing ; Statistics

Introduction

In our time the idea of a major environmental degradation emerged, at both local and global scales, in the face of the recurrent human pollution. Consequently for the sake of a sustainable humanity development, of the ethic and of the ecology, the protection and the monitoring of our environment have become a major stake. Many scientific and technical tools contributed to improve the environmental knowledge, and helped in remote and in situ observations, and forecast modelling. In situ environmental sensor systems have been designed to be increasingly sustainable even in a hostile environment such as the deep ocean. In a similar way, the infrastructures hosting those sensors are now thought and built to be permanent. In marine sciences these considerations gave birth to the “Observatory” concept: a long-term infrastructure dedicated to both bottom and water column in situ observations. In open-ocean those observatories are managed in the infrastructure European projects EMSO and FIXO3 (www.esonet-emso.org, http://www.fixo3.eu/). This community gathers about 55 partners from about 15 countries in Europe and is working in close association with other international observatories communities: Neptune Canada; OOI in USA, DONET in Japan, etc. All scientists of these communities are getting a similar product: longer and longer time series data. They all agree on the need to acquire deep sea time series data as reviewed by (Ruhl et al., 2011): “…*such observatories will contribute to answering major ocean science questions including*: *How can monitoring of factors such as seismic activity*, *pore fluid chemistry and pressure*, *and gas hydrate stability improve seismic*, *slope failure*, *and tsunami warning*? *What aspects of physical oceanography*, *biogeochemical cycling*, *and ecosystems will be most sensitive to climatic and anthropogenic change*? *What are natural versus anthropogenic changes*? *Most fundamentally*, *how are marine processes that occur* *at differing scales related?*". Similarly in coastal oceanography, time series are also acquired and the involved scientists are got together in the JERICO (www.jerico-fp7.eu) consortium to harmonize the coastal infrastructures from the sensors to the data diffusion. From the coast to the open-ocean all are facing a common challenge: the analysis of this increasing data flow. In order to facilitate the data processing from the archiving to the distribution, Neptune Canada developed a great data management system which offers at the end of the network a data portal and tool: Oceans 2.0. Nevertheless the challenge to analyse this increasing data flow in an optimal way is not yet tackled.

Consequently the idea raised to help the community from coastal to open sea areas, at international level, with the organization of a conference involving the here-above cited scientific communities (not only) in order to share the experiences in time series analysis, to outline the gaps and needs for the future. This international event, titled "Time series analysis in marine sciences and applications for industry" (Logonna-Daoulas, Bretagne, 17-21 September 2012) included a 2-day training session followed by a 3-day conference. It gathered more than 100 attendees from 54 institutions amongst 24 countries with purpose i) to integrate the scientific community and research activities at the crossroads of marine sciences (physical oceanography, marine chemistry, marine biology, ecology, geology and ocean engineering); ii) to share the rapidly developing knowledge; iii) to enhance cross discipline interactions and collaborations. This conference produces several outcomes amongst which this special issue dedicated to time series analysis in marine sciences and gathering 11 papers. Considering that time-series analysis is the future for marine science to understand ocean processes and their dynamics in most of the marine research fields the conference was organized according to marine research fields. Meanwhile, as the initial objective was to exchange knowledge on time series analysis methods, this special edition is organized according to two topical analysis methods. The first section presents articles focused on the study of the variability and information contained in the observed signal (1), and the second section introduces articles dealing with the study of extreme events (2).

1. The study of the variability contained in the observed signal

In sciences of the Earth and especially in marine sciences, physical, biological and biogeochemical parameters show huge variability on a wide range of scales, from seconds to thousand years. The time series that are recorded in such systems are often highly nonlinear, non-stationary, multiscale, and noisy. Such complex series need adequate and specific methods for their analysis. A brief review of methods and their applications published in this special issue is performed hereafter.

A first article is dedicated to the characterization of ambient acoustic noise versus anthropogenic sounds and their effect on marine ecosystems such as marine mammals. The reader will find an example of sound signal acquired during 42 months by 4 hydro-acoustic stations (van der Schaar et al., 2013, this issue).

The classical Fourier spectral analysis is not forgotten, inspired by turbulence, which in this issue is applied to characterise geo-hazards events related to sea bottom seismology. Indeed, Monna et al (2013, this issue) present time series analysis of seismic, gravimetric and magnetic measurements acquired on long-term observatories deployed in the Mediterranean Sea.

There are also other famous classical methods which are used mostly in biological oceanography, with statistical tests, principal component analysis (PCA), factor analysis (see e.g. (Gotelli and Ellison, 2004); both methods relying on an assumption of linearity of the system. In this issue, an illustration of statistical tests applied to temperature series covering the 1948-2010 period at Station M in the Norwegian Sea is presented Lorentzen (2013, this issue). Factor analyses, including PCA, are well known in ecology to study communities' pattern variability and are traditionally devoted to describe spatial characteristics. The temporal dynamic has been introduced in PCA leading to the Empirical Orthogonal Function (EOF) method in the middle of the XXth century (Lorenz, 1956) . The reader will find an application of EOF decomposition in Dragon et al. (2013, this issue) who examined the effects of an extreme calving event on regional sea ice distribution. They studied the high-frequency variability of sea ice concentrations for 1992-2011 years with passive microwave satellite data.

More recently new methodologies that allow the modelling of directional processes such as the distance-based Moran's Eigenvector Maps were developed (dbeMEM, (Dray et al., 2006; Legendre and Legendre, 2012). DbeMEM method is illustrated in the following pages to reveal the temporal structure of the benthic community observed in the field of view of a camera located in the Neptune Canada network in the Barkley Canyon (Matabos et al., 2013, this issue). Indeed, by combining high frequency ecology data from the camera with high frequency environment parameters such as temperature, salinity, fluorescence, turbidity and currents Matabos et al. related species and community patterns to environment parameters at different time scales using periodograms, wavelets and multivariate methods such as canonical redundancy analysis and dbeMEM. Jointly, Barkley canyon images are analysed by Doya et al. (2013, this issue) to study the diel rhythms in sablefish behavior and in other species. In Doya et al., the study of time series of current data help establishing the internal tidal regimes and were related to Chi-square periodograms of ecological data extracted from images. This article puts forward a promising perspective for the chronobiology.

We can cite the Dynamic Factor Analysis (DFA), a dimension reduction technique designed for time series, and traditionally applied in econometrics and psychology. It is now applied in fisheries studies (Zuur et al., 2003). In this issue, this method is applied to relate fishing pressure on commercial catch rates of clams with environmental variables such as upwelling index, NAO indices, SST, local wind, river discharges, off Portugal coast (Baptista and Leitão, 2013, this issue). The influence of local climatic features off Portuguese coasts on clams' catches is emphasized.

We mention finally the Empirical Mode Decomposition (EMD) associated to Hilbert Spectral Analysis (HSA), introduced by Norden Huang in 1998 (Huang et al., 1998) and developed later in Wimereux to characterize intermittency (Huang et al., 2008; Huang et al., 2011). Such approach is particularly adapted to nonlinear and non-stationary time series, and is rather robust even when there is a deterministic energy injection. It has already been applied to several marine time series (Schmitt et al., 2009) of which temperature and dissolved oxygen time series acquired on MAREL Carnot station, located in the coastal waters off Boulogne sur Mer (France), in the Eastern English Channel (Huang and Schmitt, 2013, this issue). EMD is applied to consider the correlation between two non-stationary time series, displaying the correlation at different scales and different locations using mode correlation analysis

2. The study of extreme events

For more than 30 years, engineers in charge of the design of industrial structures at sea, ships and protection from flooding, have been interested in both long and short term accurate characterisation of sea states and in the probabilities of the very rare “meteocean” events that the structures or coastal defences could encounter during their service life and that could induce their mechanical ruin or overtopping (Cunnane, 1973; Muir and El-Shaarawi, 1986).

As for illustration of our words, the reader can refer Olagnon et al., (2013, this issue), who publish a study on swell offshore West Africa, mainly from Namibia to Nigeria. This location is submitted to complex wave conditions with the presence of several swells and wind sea almost at the same time, each of them depicting different frequencies and directions. This is leading to an increase number of parameters necessary to track for a suitable modelling. In order to face this complexity Olagnon et al. defined and assess parametric spectral models for individual sea state swell spectra and then for events (remote storms) based on long term statistics of observed swell climate. These works can apply for instance to fatigue assessment of an offshore structure, or to the coastal erosion of a shoreline but its application could be more difficult with a ship that would take different routes. Indeed along its routes a vessel would meet several sea conditions. To answer this question, Podgorski and Rychlick (2013, this issue) present a model that helps to estimate the reliability of a ship according to the variability of encountered sea states as the expected vessels damages can be assessed according to the significant wave height, ship speed and heading angle.

Similarly in hydrology flood defence, energy production, assessment of low flow, etc., ocean engineering is concerned with extreme values study. In the 70’s and 80’s, the time series were often short and sparse and so did not permit complex and accurate analysis of the extremes. Along the last century, extremes have been studied first through annual maximum samples and empirical distribution, fitted by Gumbel, Frechet or Weibull probability laws, then through peak-over-threshold (POT) analyses, making better use of available observational material provided sampling conditions are respected to insure a correct statistical inference, using Generalized

Pareto distributions. Formal *relationships between the two ways have been established.* In this issue an extremes model is applied to joint estimation of extreme storm peak significant wave height and peak period in the Northern Sea (Ewans and Jonathan, 2013, this issue). More and more, thanks to the increasing quality of the numerical models (in meteorology, hydrodynamics and hydrology), long time series of several decades are available and cover all the oceans with grids constantly refined. So, robust methods are necessary for the automatic process of the huge amount of data (see for example (Caires and Sterl, 2005)). These methods are constantly improved, taking into account seasonality, directional analysis, covariates, regionalization as described by (Anderson et al., 2001) and in (Haigh et al., 2012) for sea level and in the very complete review by (Jonathan and Ewans, 2013).

Conclusions

This special edition is gathering 11 presentations given during an international conference, over the 30 ones presented. Of course the here published articles cannot exhaustively show either applications of methods dedicated to study the variability contained in the observed signal or applications of methods dedicated to extreme events. Nevertheless we can add for instance the use of wavelets analysis applied to sediment transport investigation in estuarine processes (Jones and Chang, 2012). In physical oceanography the reader can refer to the analysis of the Meridional Overturning Circulation (MOC) variability from a 6-year high frequency time series of water transport (Frajka-Williams et al., 2013; Rayner et al., 2011) and to studies of the California Current in the climate systems after a 50-year time series (Cornuelle et al., 2012). Presentations allowed to be published and proceedings of the conference are available at: http://www.europolemer.eu/Europole-Mer-From-2007-to-2012/Scientific-Meetings/Gordon-like-conferences/Time-series-analysis.

The attractiveness of such a conference is strongly associated with its topic: time series analysis and to the fact that scientists are facing an intensifying flow of time series data (e.g.: 1D data from fixed point, 2 D data from repeat sections and modeling, satellites and video images). Indeed, data acquisition in environmental sciences strongly increased during the XXth century in parallel to technological progresses and societal needs to monitor the environment. A great step ahead occurred during the 90s when computers started to become common equipment in research infrastructures and at home, facilitating the data archiving and processing. For instance the French Coriolis data centre recorded 12 408 observation locations (defined by latitude, longitude, date & time and water depths) in 1970, 43 041 observation locations in 1980, then showing a multiplication factor of 4.5 between 1980 and 1990 and of 10 between 90s and 2000. The number of data points exponentially increased after 2000 due to real time data transmission and the development of operational oceanography that helped including regional data in addition to the global ones usually acquired. Extrapolating from this tendency, referring to the experience with Neptune Canada observatory (60 terabytes, i.e. 10^{12}, per year, https://marinemetadata.org/news/neptunenature) and to the US OOI observatory forecasting 3 petabytes, i.e. 10^{15}, per year, and taking into account the present and future technological progress in terms of sensor capabilities and data transmission bandwidth, our generation of scientists will be soon flooded by a huge in situ data flow, like in many scientific fields. This is already the case if one considers numerical models, satellites and remote data flows which for instance are generating a data flux of 100 terabyte per year to the CERSAT (Laboratory of Oceanography From Space in IFREMER (French Research Institute for Exploitation of the Sea). Moreover, the new generation of satellites such as Sentinel -1 can provide 1 petabyte of data per year, per satellite. Given the increasing size of existing environmental databases, a question is rising: will the data utilisation by scientists be optimum?

These considerations are in agreement with the fourth paradigm definition in "The Fourth Paradigm: Data-Intensive Scientific Discovery" (Hey et al., 2009). According to their words the first three paradigms were experimental, theoretical and (more recently) computational science. This book of essays argues that a fourth paradigm of scientific discovery is at hand: the analysis of massive data sets. The basic idea is that our capacity for collecting scientific data has far outstripped our present capacity to analyse it, then our focus should be on developing technologies that will interpret this "Deluge of Data". In the face of this intensifying data flow, a suitable suite of scientific workflow products is needed to facilitate the archiving, assimilation, visualisation, modelling and interpretation of the data-contained information. A lot of energy has already been invested on the harmonization of data archiving. As a result a great leap forward has been made and data base systems offer more and more data to a wider community of users. Nevertheless, data visualisation, analysis and interpretation are still too specific to the concerned research fields and are still too often working on homemade recipes. In order to fill this gap, semi-automated analysis of the data systems could be integrated in the data base systems but first the expertise to interpret and valid the analysis requires a high level scientific knowledge.

Acknowledgements

Guest editors, organizing committee of the conference “Time series analysis in marine sciences and applications for industry”, are very grateful to the hereafter listed agencies for co-funding the conference and publication of this special edition: Europôle mer (FR), Labex Mer (FR), Ifremer –Edrome Institut Carnot (FR), ONRG (USA), CNRS-INSU (FR), Conseil général du Finistère (FR), Conseil régional de Bretagne (FR), Brest Métropole Océane (FR).

By publishing this special edition we also acknowledge the review of about 30 referees who kindly contributed in less than 12 months to this special edition.

Finally special thanks are devoted to Nadine Renier (IEUM, Brest) for her deep involvement in the conference organisation.

References:

Anderson, C.W., Carter, D.J.T., Cotton, D., 2001. Wave Climate Variability and Impact on Offshore Design Extremes, *Report for Shell International and the Organization of Oil & Gas Producers*, p. 90.

Baptista, V., Leitão, F., 2013, this issue. Commercial catch rates of the clam Spisula solida reflect local environmental coastal conditions. Journal of Marine Systems.

Caires, S., Sterl, A., 2005. 100-Year Return Value Estimates for Ocean Wind Speed and Significant Wave Height from the ERA-40 Data. Journal of Climate 18, 1032–1048.

Cornuelle, B., Kim, S.Y., Mazloff, M., Verdy, A., 2012. The california current in the climate system: analysis of observations, in: Puillat (Ed.), Time-series analysis in marine science and applications for industry, Logonna-Daoulas (17-21 sept. 2012, Brittany, France), p. 48.

Cunnane, C., 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. Journal of Hydrology 18, 257-271.

Doya, C., Aguzzi, J., Pardo, M., Matabos, M., Company, J.B., Costa, C., Mihaly, S., Canals, M., 2013, this issue. Diel behavioral rhythms in sablefish (Anoplopoma fimbria) and other benthic species, as recorded by the Deep-sea cabled observatories in Barkley canyon (NEPTUNE-Canada). Journal of Marine Systems.

Dragon, A.-C., Houssais, M.-N., Herbaut, C., Charrassin, J.-B., 2013, this issue. A note on the intraseasonal variability in an Antarctic polynia: Prior to and after the Mertz Glacier calving. Journal of Marine Systems.

Dray, S., Legendre, P., Peres-Neto, P., 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196, 483-493.

Ewans, K., Jonathan, P., 2013, this issue. Evaluating environmental joint extremes for the offshore industry using the conditional extremes model. Journal of Marine Systems.

Frajka-Williams, E., Johns, W., Meinen, C., Beal, L., Cunningham, S., 2013. Eddy impacts on the Florida Current. Geophysical Research Letters 40, 349-353.

Gotelli, N.J., Ellison, A.M., 2004. A primer of ecological statistics. Sinauer Associates Publishers, Sunderland, Mass.

Haigh, I.D., Nicholls, R., Wells, N., 2012. A comparison of the main methods for estimating probabilities of extreme still water levels. Coastal Engineering 57, 838-839.

Hey, T., Tansley, S., Tolle, K., 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences 454, 903.

Huang, Y., Schmitt, F.G., 2013, this issue. Time dependent intrinsic correlation analysis of temperature and dissolved oxygen time series using empirical mode decomposition. Journal of Marine Systems.

Huang, Y., Schmitt, F.G., Lu, Z., Liu, Y., 2008. Scaling analysis of time series using empirical mode decomposition and Hilbert spectral analysis. TRAITEMENT DU SIGNAL 25, 481-491.

Huang, Y.X., Schmitt, F.G., Hermand, J.P., Gagne, Y., Lu, Z.M., Liu, Y.L., 2011. Arbitrary-order Hilbert spectral analysis for time series possessing scaling statistics: comparison study with detrended fluctuation analysis and wavelet leaders. Phys Rev E Stat Nonlin Soft Matter Phys 84, 016208.

Jonathan, P., Ewans, K., 2013. Statistical modelling of extreme ocean environments for marine design: A review. Ocean Engineering 62, 91-109.

Jones, C., Chang, G., 2012. A deeper look at estuarine processes - Wavelets and circulation., in: Puillat (Ed.), Time series analysis in marine science and applications for industry., Logonna-Daoulas (17-21 sept. 2012, Brittany, France), p. 48.

Legendre, P., Legendre, L., 2012. Numerical Ecology, 3rd Edition ed. Elsevier.

Lorentzen, T., 2013, this issue. Statistical analysis of temperature data sampled at station-m in the Norwegian sea. Journal of Marine Systems.
Lorenz, E., 1956. Empirical orthogonal functions and statistical weather prediction. Tech. Rep. 1, Statistical Forecasting Project. Massachusetts Institute of Technology, Department of Meteorology., Cambridge, MA, p. 49pp.
Matabos, M., Bui, A.O.V., Mihály, S., Aguzzi, J., Juniper, S.K., Ajayamohan, R.S., 2013, this issue. High-frequency study of epibenthic megafaunal community dynamics in Barkley Canyon: A multi-disciplinary approach using the NEPTUNE Canada network. Journal of Marine Systems.
Monna, S., Falcone, G., Beranzoli, L., Chierici, F., Cianchini, G., De Caro, M., De Santis, A., Embriaco, D., Frugoni, F., Marinaro, G., Montuori, C., Pignagnoli, L., Qamili, E., Sgroi, T., Favali, P., 2013, this issue. Underwater geophysical monitoring for European Multidisciplinary Seafloor and water column Observatories. Journal of Marine Systems.
Muir, L.R., El-Shaarawi, A.H., 1986. On the calculation of extreme wave heights: A review. Ocean Engineering 13, 93-118.
Olagnon, M., Kpogo-Nuwoklo, K.A., Guédé, Z., 2013, this issue. Statistical processing of West Africa wave directional spectra time-series into a climatology of swell events. Journal of Marine Systems.
Podgórski, K., Rychlik, I., 2013, this issue. A model of significant wave height for reliability assessment of a ship. Journal of Marine Systems.
Rayner, D., Hirschi, J., Kanzow, T., Johns, W., Wright, P., Frajka-Williams, E., Bryden, H., Meinen, C., Baringer, M., Marotzke, J., Beal, L., Cunningham, S., 2011. Monitoring the Atlantic meridional overturning circulation. Deep-Sea Research Part Ii-Topical Studies in Oceanography 58, 1744-1753.
Ruhl, H.A., Andre, M., Beranzoli, L., Cagatay, M.N., Colaco, A., Cannat, M., Danobeitia, J.J., Favali, P., Geli, L., Gillooly, M., Greinert, J., Hall, P.O.J., Huber, R., Karstensen, J., Lampitt, R.S., Larkin, K.E., Lykousis, V., Mienert, J., Miguel Miranda, J., Person, R., Priede, I.G., Puillat, I., Thomsen, L., Waldmann, C., 2011. Societal need for improved understanding of climate change, anthropogenic impacts, and geo-hazard warning drive development of ocean observatories in European Seas. Progress in Oceanography 91, 1-33.
Schmitt, F.G., Huang, Y., Lu, Z., Liu, Y., Fernandez, N., 2009. Analysis of velocity fluctuations and their intermittency properties in the surf zone using empirical mode decomposition. Journal of Marine Systems 77, 473-481.
van der Schaar, M., Ainslie, M.A., Robinson, S.P., Prior, M.K., André, M., 2013, this issue. Changes in 63 Hz third-octave band sound levels over 42 months recorded at four deep-ocean observatories. Journal of Marine Systems.
Zuur, A.F., Tuck, I.D., Bailey, N., 2003. Dynamic factor analysis to estimate common trends in fisheries time series. *Canadian Journal of Fisheries and Aquatic Sciences* 60, 542-552.