

A Machine Learning Based Approach to Predict *Ostreopsis cf. ovata* Bloom Events from Meteo-Marine Forecasts

Ennio Ottaviani^a, Stefania Magri^{bc*}, Patrizia De Gaetano^c, Cecilia Cuneo^c, Mariachiara Chiantore^d, Valentina Asnaghi^d, Andrea Pedroncini^e

^aOnAIR s.r.l., Via Carlo Barabino, 16129 Genova, Italy

^bDICCA - Civil, Chemical and Environmental Engineering Dept. – Genova University, via Opera Pia 15 - 16145 Genova, Italy

^cARPAL (Agenzia Regionale per la Protezione dell'Ambiente Ligure), Via Bombrini 8, 16149 Genova, Italy

^dDISTAV Department of Earth, Environment and Life Sciences, Viale Benedetto XV - 16132 Genova, Italy

^eDHI Italia – Via Bombrini 11, 16149 Genova, Italy

stefania.magri@arpal.liguria.it

Ostreopsis cf. ovata, a benthic toxic marine dinoflagellate, has been recorded along Italian coasts since the '90, but large bloom events have been reported only in recent years. In 2005, a monitoring programme started along the Ligurian coast (North-western Mediterranean), where time series of cell abundances have been collected for several sites, together with a range of related environmental variables. Data of cell abundances in 15 sites, together with environmental data provided by meteo-marine forecasting models used by the Regional Environmental Agency (ARPAL), have been used to implement a predictive modelling tool, able to forecast *Ostreopsis* cells concentration threshold exceedance as a function of meteo-marine forecasts. Starting from the experience of the predictive model implemented in 2015, the Quantile Regression Forest (QRF) has been applied: the model has been trained on past data (from 2015 until 2017) and tested with data taken during the two last available years (2018 and 2019). The use of this extremely adaptable regression model to evaluate threshold exceedance has shown a good capacity to predict overcoming events at a given spatial location. This tool can help the Regional Agency in the decision making process, providing an alert when/where a given alarm threshold is exceeded in order to trigger the emergency procedures. This is a first step in defining a predictive sampling strategy able to better capture bloom events.

Keywords: classification, regression, machine learning, environmental data, *Ostreopsis ovata*

1. Introduction

Ostreopsis species are largely distributed from tropical to temperate marine areas worldwide and since the early part of the twenty-first century, *Ostreopsis* blooms regularly occur in the Mediterranean Sea during summer-autumn (Vila et al., 2001; Turki, 2005; Aligizaki and Nikolaidis, 2006; Mangialajo et al., 2008; Totti et al., 2010; Illoul et al., 2012; Ismael and Halim, 2012; Pfannkuchen et al., 2012). These phenomena seem to have recently increased, and the relationship with eutrophication and climate change has been hypothesized (Hallegraeff, 2008; Wells et al., 2020). Economic and social interests are affected by blooms, as they are responsible for respiratory and skin problems on humans (Tichadou et al., 2010; Del Favero et al., 2012) and benthic marine organisms (Pagliara and Caroppo, 2012; Gorbi et al., 2013; Carella et al., 2015).

Ostreopsis species have been, in the last decades, implicated in outbreaks causing health issues in countries along the northern Mediterranean coasts, particularly Spain, Italy and Greece. In 2007, the Italian Ministry of Health promoted the introduction of management actions related to the risk associated to *O. cf. ovata* blooms within the bathing water regulation, integrating surveys of *Ostreopsis* presence, and defining a three-phase monitoring plan: a routine phase, an alert phase and an emergency phase, identified by thresholds of cells concentrations in the water of, respectively, 10,000 cells/l (Alert) and 30,000 cells/l (Warning), followed by adequate communication to the public (Funari, 2014). In this context, applying cross-industry learning from the high hazard process sector, it seems advisable to perform a preventive environmental risk assessment, implementing suitable prediction criteria and performance indicators, addressing not only sea use and

emergency planning, but also structural interventions for the long-term prevention (Fabiano et al., 2017). Several studies have highlighted the influence of environmental factors on bloom dynamics. Temperature has been identified as an important trigger affecting its distribution, though its role is not the same in all coastal areas around the world (Accoroni, 2016). Salinity turned out to be positively correlated with *O. cf. ovata* abundance (Carnicer, 2015), but the relationships between algal blooms and salinity are more complicated and other factors, such as nutrient levels, which are typically associated to low salinity waters, have to be considered. Studies have provided increasing evidence of a link between the nutrient enrichment of coastal waters (anthropogenic eutrophication) and harmful algal events (Glibert et al., 2010). However, there is very limited information on the relationships between nutrient concentrations and the occurrence of *Ostreopsis* blooms. Several studies considered hydrodynamic condition as a main factor affecting *Ostreopsis* bloom trends, highlighting that higher abundances are observed in sheltered sites compared to exposed ones (Barone, 2007; Totti et al., 2010). However, effects of hydrodynamics and turbulence may lead to opposite effects in each bloom phase. Despite the number of studies on *Ostreopsis* biology, ecology and toxin production, several aspects about the environmental concerns associated with this genus remain still unclear. Indeed, the complexity of conditions leading to blooms of this dinoflagellate is dependent on the specific area, thus site specific characteristics must be taken into account when implementing a predicting model.

As far as the North-Western Mediterranean is concerned, a preliminary descriptive model of the benthic *O. cf. ovata* bloom events was produced for the Ligurian Sea (Asnaghi et al., 2012), highlighting a relevant role of seawater temperature and hydrodynamics in driving the bloom. The percentage of variation explained by a relatively small set of variables was remarkably large (around 80%), highlighting the likelihood that a set of key seawater and meteorological variables could be used as good predictors in forecasting potential toxic events. These findings paved the way to building reliable predictive models based on a small set of variables (mostly meteorological) and a first predictive model of *O. cf. ovata* concentrations in seawater was built in by Asnaghi et al. (2017) using the Quantile Regression Forests (QRF).

Herein, starting from above experience, a machine learning based approach is proposed to predict the concentration of *O. cf. ovata* in seawater in 15 sites along the Ligurian coast, and the related threshold exceedance probability, using seawater and meteorological variables, derived from numerical forecasting models. The methodology provides a tool for local environmental and public health agencies, as well as a support for the agency in charge of monitoring in planning sampling activities to better capture bloom events.

2. Material and Methods

2.1 Theoretical bases of the model

A predictive model of *O. cf. ovata* concentrations in seawater was built in previous work (Asnaghi et al., 2017) using the Quantile Regression Forests (QRF), an ensemble machine learning method (Meinshausen, 2006), as the best option after the comparison of several modeling techniques (Ottaviani et al., 2015).

QRF is an inference scheme based on the Random Forests (RF) general method (Breiman, 2001). The basic RF algorithm grows an ensemble of trees in which, for each tree, only a sub-sample (“in-bag”) of the data is used for the training, and the remaining part (“out-of-bag”) is used for testing. In addition, only a random subset of predictor variables is considered to split point selection at each tree node. RF are useful to model complex input-output relations without assuming any information about its functional form and about the probability distribution of the variables. The size of the random subset is the single tuning parameter of the algorithm. Instead of a single output, the QRF provides a whole distribution of predicted values for each combination of input features, allowing the selection of the best quantile for prediction. The key difference between QRF and RF is that for each node in each tree, RF keeps only the mean of the observations that fall into this node and neglects all other information, while QRF keeps the value of all observations in this node, not just their mean, and assesses the conditional distribution based on this information.

2.2 Study area

Since 2005, regional monitoring surveys of *Ostreopsis cf. ovata* cell concentrations in seawater have been carried out by ARPAL (Regional Environmental Agency) from June to September at 14 sites along the Ligurian coast (Figure 1). Sampling sites are selected in sheltered areas, where the coast morphology and environmental characteristics are favorable for *Ostreopsis* proliferation: shallow water, poor water exchange (i.e., near coast protection structures); rocky-pebbly bottom; presence of macroalgae. Sampling is carried out according to the three-phase monitoring plan defined by regulation: twice a month for the routine phase, and repeated in case of exceeding the alert threshold. Data collected from 2015 to 2019 on the 14 sites monitored by ARPAL, together with data collected by Genoa University at one monitoring station along Genoa coast, were used.

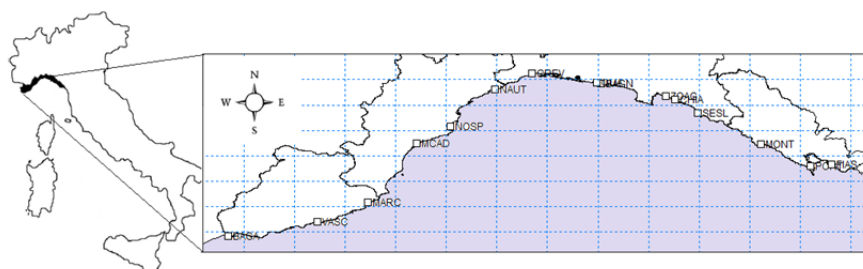


Figure 1. Distribution of the 15 sampling sites along the Ligurian coast line.

2.3 Environmental input data

On the basis of previous studies from the authors (Asnaghi et al., 2012, Asnaghi et al., 2017) and literature (Accoroni and Totti, 2016), for the present study, the environmental variables considered for algae concentration prediction were: seawater surface temperature, salinity, current speed and direction, air temperature, wind speed and direction, atmospheric pressure. All variables were derived from models operated by ARPAL and successfully validated (Vairo et al., 2017a), with the aim of setting up an operative predictive model. Meteorological data were obtained from the Limited Area Atmospheric Model MOLOCH, while seawater variables were obtained from the hydrodynamic model of the Ligurian Sea, developed with the three-dimensional finite volume MIKE 3 Flow Model FM (DHI, 2019). The hydrodynamic model is based on a flexible mesh approach, with the finest horizontal resolution of 50 m along the coast of Metropolitan city of Genova (**Errore. L'origine riferimento non è stata trovata.**) and approximately 500m in the remaining western and eastern Ligurian coast, and a hybrid vertical discretization system (σ and z) to account for the stratification effects. For environmental risk management, especially related to coastal phenomena, availability of high resolution models can highly influence results reliability both in the atmospheric environmental compartment (Vairo et al., 2014) and in sensitive sea environments (Vairo et al. 2017b, Magri et al. 2019).

2.4 Construction of the predictive model

A regression model was built using 10 features as predictors: 1-station ID, 2-day of the year, 3-sea surface temperature, 4-air temperature, 5-salinity, 6-atmospheric pressure, 7-East-West component of the wind speed, 8-North-South component of the wind speed, 9-East-West component of the surface sea current, 10-North-South component of the surface sea current.

All features were coded as continuous variables, apart from station ID that is a categorical variable with 15 unordered possible values. The response variable was base ten logarithms of the concentration of *Ostreopsis cf. ovata* in seawater. The explicit use of the station ID as a predictor is related with possible difference in behavior among the stations, due to their physical or geographic characteristics. In line of principle it should be possible to train a single model without this information or 15 different models for the 15 sites, but these options have different drawbacks. The first ignores the relevant differences among sites, surely affecting the blooms, the second uses too little data for training each model properly. The choice of a mixed model, using the station ID as a predictor, includes geographical information and allows information sharing among sites.

Data collected from 2015 to 2019 were used. Each tree was grown using a bootstrapped sample containing about 60% of all the data, with 5 features tried at each split. A total of 504 entries were present in the whole dataset, with 52 candidate bloom events collected during summer with the following distribution: 266 entries (and 32 bloom) in 2017 and before; 119 entries (and 11 bloom events) in 2018; 119 entries (and 9 bloom events) in 2019.

The proposed methodology for the monitoring divides the dataset in a training part (the past years) and a testing part (the current year) and should be repeated yearly in order to maintain its predictive capability. However, bloom events in last years are quite rare and so to get a more reliable assessment of performance we used both 2018 and 2019 as a cumulated test case (20 blooms). However, the use of RF allows also an assessment of model performance even without this data splitting, thanks to the OOB mechanism (each tree of the forest is tested only on data outside its training set). This kind of analysis gives good hints about model tuning and feature relative importance. The QRF model is trained in regression mode, aiming to minimize the Mean Square Error (MSE) between real values and predictions, but it has been tested in classification mode, using 10,000 cells/l as a threshold. This was done in order to give maximum flexibility in changing the threshold without affecting the training, so the same model could work with another threshold as well. Moreover, prediction of bloom events contains not only a single value for concentration, but also a confidence interval of the estimate, and this is another useful feature of the regression model. The confidence interval

allows the evaluation of the probability of exceeding the threshold, enforcing the confidence of the prediction itself. However, we are not directly interested in exploiting the concentration values for this kind of study, therefore the main output is only classification, even if model capability is broader.

3. Results and discussion

3.1 True and false positive rate

In order to better describe the characteristics of the model, the probability at which an *in situ* concentration above 10,000 cells/l was correctly predicted ('True bloom' condition, that is the True Positive Rate, TPR) was computed for each quantile of the predictions in the OOB mode. Similarly, the probability of erroneously predicting bloom events when measured concentrations were below the specified threshold of 10,000 cells/l ('False bloom' condition, that is the False Positive Rate, FPR) was computed. Results were then plotted as probability versus quantile (Figure 2). The curves indicate that there is a range of quantiles in which TPR is significantly larger than FPR, and this suggests an 85-quantile as a good value for the model. For the present study, it was chosen to accurately predict the highest possible number of bloom events over a season, and, concurrently, adjusting the accepted probability of generating false alarms. However, other good tradeoffs are possible, depending on the real costs of false positive and false negative predictions.

3.2 Feature relative importance

Using the whole dataset we computed also the relative importance of each feature. For each feature, we permuted the values of this feature across each observation in the whole dataset and measured how worse the MSE becomes after the permutation. Results are shown in Figure 2.

The plot shows a clear strong importance of the site (n.1) and the day of the year (n.2), that is seasonality, and this is largely expected. Apart from these, sea surface temperature (n.3) continues to be the main driver for blooms. Other variables exerted a less clear impact on model prediction, and this may be dependent on the specific measure adopted for quantifying importance

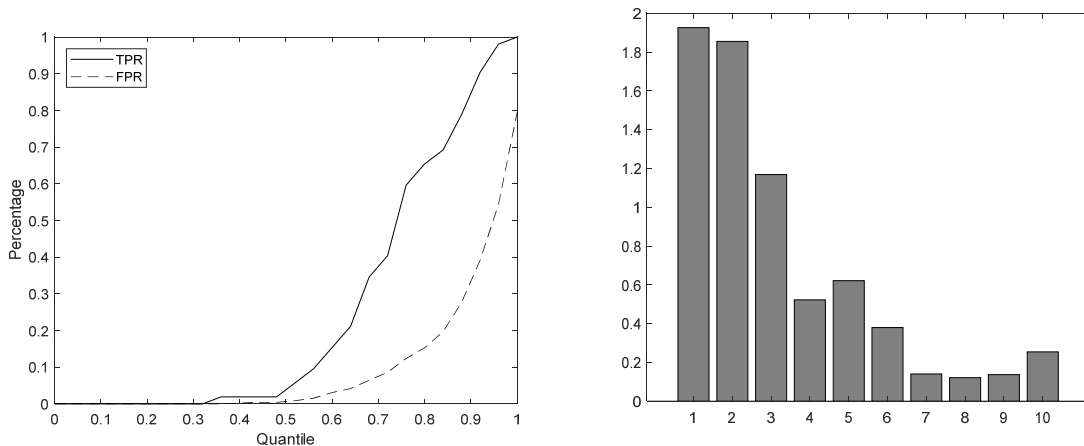


Figure 2. TPR and FPR curves vs quantile (left figure) and Features importance plot (right hand figure).

3.3 Model validation

Model validation was performed by generating predictions of concentrations of *Ostreopsis cf. ovata* in the water column using meteorological and hydrodynamic data generated from June 2018 to September 2019 by the MIKE 3 and MOLOCH models (238 samples) but with a QRF model trained with data up to September 2017 (266 samples). These values were then compared to actual measurements of concentration of *O. cf. ovata* and a confusion matrix of bloom detection has been computed.

Table 1: Confusion matrix of predictions.

	Predicted	
Observed	Bloom	No bloom
Bloom	True bloom alert (13)	Undetected bloom (7)
No bloom	False bloom alert (56)	No bloom (162)

This kind of test represents a possible operating strategy of the model and shows the capability to deal with data collected in different years without degradation. With this data, in fact, TPR is 65% and FPR is 25%, not too different from TPR/FPR estimated with the whole dataset in OOB mode. The results of this test suggest that further analyses are needed and the present methodology should be useful to guide the sampling strategy. Yet, both estimates are not indicative of the detection performance of the full *Ostreopsis* monitoring chain.

4. Conclusions

Ostreopsis cf ovata blooms regularly occur in the Ligurian Sea during summer since 2005. Due to the associated risk for health problems on humans, a monitoring plan, carried on by ARPAL for the Ligurian Region, was introduced by the Italian Ministry of Health as part of the risk management actions. Survey campaigns follow a three-phase monitoring plan: a routine phase, that considers, from June to September, at least two sampling events per month, an alert phase and an emergency phase, identified by thresholds of cells concentrations in the water of, respectively, 10,000 cells/l (Alert) and 30,000 cells/l (Emergency), which require more intense sampling activities, followed by an adequate communication to the public. Monitoring campaigns are time intensive, and are carried out at a frequency that may not be appropriate to detect the variability of *Ostreopsis* dynamics. The routine phase samples are not always able to catch local blooms, while extra sampling during alert and emergency phases should be frequent enough to be able to identify the exact time period when conditions are actually associated with health risk for the population. However, *Ostreopsis* concentration alone is not sufficient to identify the duration of the health emergency, as environmental conditions, such as wave, wind, rain and temperature, can highly influence bloom effects over a small scale (in space and time). The use of a predictive model to support sampling strategy in both routine phase, and in the alert/emergency phase, could lead to a reduction of costs and to an optimization of the sampling strategy to provide a higher quality service of early warning. The proposed predictive model, being based on data generated by meteorological and hydrodynamic forecasting models, can be used in an operation way, providing important information of predicted *Ostreopsis* concentration as well as probability of threshold exceedance, thus supporting ARPAL in optimizing monitoring campaign designs, and possibly limit actual sampling to cases of predicted bloom events. Additionally, this predictive tool can help the Regional Agency in the decision-making process, providing an alert when/where a given alarm threshold is exceeded in order to trigger the emergency procedures. In this vision, the use of the Quantile Regression Forests (QRF) allows mimicking a more or less risk inclined approach for coastal managers.

References

- Accoroni S., Totti C., 2016, The toxic benthic dinoflagellates of the genus *Ostreopsis* in temperate areas: a review, *Advances in Oceanography and Limnology*, 7(1), 1-15.
- Aligizaki K., Nikolaidis G., 2006, The presence of the potentially toxic genera *Ostreopsis* and *Coolia* (Dinophyceae) in the North Aegean Sea, Greece., *Harmful Algae* 5, 717–730.
- Asnagli V., Pecorino D., Ottaviani E., Pedroncini A., Bertolotto R.M., Chiantore M., 2017, A novel application of an adaptable modeling approach to the management of toxic microalgal bloom events in coastal areas, *Harmful Algae*, 63, 184-192.
- Asnagli V., Bertolotto R., Giussani V., Mangialajo, L., Hewitt, J., Thrush, S., Moretto, P., Castellano, M., Rossi, A., Povero, P., Cattaneo-Vietti, R., 2012, Interannual variability in *Ostreopsis ovata* bloom dynamic along Genoa coast (North- Western Mediterranean): a preliminary modeling approach. *Cryptogamie Algologie*, 33 (2), 181-189.
- Barone R., 2007, Behavioural trait of *Ostreopsis ovata* (Dinophyceae) in Mediterranean rock pools: the spider's strategy, *Harmful Algae News*, 33, 1-3.
- Breiman L., 2001, Random forests, *Mach. Learn.*, 45, 5–32.
- Carella F., Sardo A., Mangoni O., Di Cioccio D., Urciuolo G., De Vico G., Zingone A., 2015, Quantitative histopathology of the Mediterranean mussel (*Mytilus galloprovincialis* L.) exposed to the harmful dinoflagellate *Ostreopsis cf. ovata*, *J. Invertebr. Pathol.*, 127, 130-140.
- Carnicer O., Guallar C., Andree KB., Diogène J., Fernández-Tejedor M., 2015, *Ostreopsis cf. ovata* dynamics in the NW Mediterranean Sea in relation to biotic and abiotic factors, *Environmental Research*, 143, 89-99.
- Del Favero G., Sosa S., Pelin M., D'Orlando E., Florio C., Lorenzon P., Poli M., Tubaro A., 2012., Sanitary problems related to the presence of *Ostreopsis* spp. in the Mediterranean Sea: a multidisciplinary scientific approach, *Ann. Ist. Super. Sanità*, 48, 407-414.
- Fabiano B., Vianello C., Reverberi A.P., Lunghi E., Maschio G., 2018, A perspective on Seveso accident based on cause-consequences analysis by three different methods, *Journal of Loss Prevention in the Process Industries*, 49, 18-35.

- Funari E., Manganelli M., Testai E., 2014, *Ostreopsis cf. ovata*: linee guida per la gestione delle fioriture negli ambienti marino costieri in relazione a balneazione e altre attività ricreative. Roma: Istituto Superiore di Sanità, Rapporti ISTISAN 14/19.
- Glibert P.M., Allen J.I., Bouwman A.F., Brown C.W., Flynn K.J., Lewitus A.J., Madden C.J., 2010, Modeling of HABs and eutrophication Status, advances, challenges. *J. Mar. Syst.*, 83, 262-275.
- Gorbi S., Avio G.C., Benedetti M., Totti C., Accoroni S., Pichierri S., Bacchiocchi S., Orletti R., Graziosi T., Regoli F., 2013, Effects of harmful dinoflagellate *Ostreopsis cf. ovata* exposure on immunological, histological and oxidative responses of mussels *Mytilus galloprovincialis*. *Fish Shellfish Immunol.*, 35, 941-950.
- Hallegraef G.M., 2008, Harmful algal blooms, coastal eutrophication and climate change, *Biol. Mar. Medit.*, 15(1), 6-15.
- Illoul H., Hernandez F.R., Vila M., Adjas N., Younes A.A., Bournissa M., Koroghli A., Marouf N., Rabia S., Ameer F.L.K., 2012, The genus *Ostreopsis* along the Algerian coastal waters (SW Mediterranean Sea) associated with a human respiratory intoxication episode, *Cryptogam. Algol.*, 33, 209-216
- Ismael A., Halim Y., 2012, Potentially harmful *Ostreopsis* spp. In the coastal waters of Alexandria-Egypt, *Mediterr. Mar. Sci.*, 13, 208-212.
- Magri S., Quagliati M., Di Gaetano P., Vairo T., Fabiano B., 2019, Fuel Spill After Ships Collision: Accident Scenario Modelling for Emergency Response, *Chemical Engineering Transactions*, 74, 1363-1368.
- Mangialajo L., Bertolotto R., Cattaneo-Vietti R., Chiantore M., Grillo C., Lemée R., Melchiorre N., Moretto P., Povero P., Ruggieri N., 2008, The toxic benthic dinoflagellate *Ostreopsis ovata*: Quantification of proliferation along the coastline of Genoa, Italy, *Mar. Pollut. Bull.*, 56, 1209-1214.
- Meinshausen N., 2006, Quantile regression forests, *J. Mach. Learn. Res.* 7, 983-999.
- Ottaviani E., Asnaghi V., Chiantore M., Pedroncini A., Bertolotto, R., 2015, Statistical methods for the analysis of *Ostreopsis ovata* bloom events from meteo-marine data, *Proceedings of CLADAG 2015 10th Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society, Cagliari (Italy)*, 8-10 Oct 2015, Volume: Book of Abstracts (ISBN: 978-88-8467-749-9).
- Pagliara P., Caroppo C., 2012, Toxicity assessment of *Amphidinium carterae*, *Coolia cfr. monotis* and *Ostreopsis cfr. ovata* (Dinophyta) isolated from the northern Ionian Sea (Mediterranean Sea), *Toxicon* 60, 1203-1214.
- Pfannkuchen M., Godrijan J., Maric Pfannkuchen D., Ivesa L., Kruzic P., Ciminiello P., Dell'Aversano C., Dello Iacovo E., Fattorusso E., Forino M., Tartaglione L., Godrijan M., 2012, Toxinproducing *Ostreopsis cf. ovata* are likely to bloom undetected along coastal areas, *Environ. Sci. Technol.*, 46, 5574-5582.
- Tichadou L., Glaizal M., Armengaud A., Grosseil H., Lemée R., Kantin R., Lasalle J.L., Drouet G., Rambaud L., Malfait P., de Haro L., 2010, Health impact of unicellular algae of the *Ostreopsis* genus blooms in the Mediterranean Sea: experience of the French Mediterranean coast surveillance network from 2006 to 2009, *Clinical Toxicology*, 48, 839-844.
- Totti C., Accoroni S., Cerino F., Cucchiari E., Romagnoli T., 2010, *Ostreopsis ovata* bloom along the Conero Riviera (northern Adriatic Sea): Relationships with environmental conditions and substrata, *Harmful Algae*, 9, 233-239.
- Turki S., Harzallah A., Sammari C., 2006, Occurrence of harmful dinoflagellates in two different Tunisian ecosystems: the lake of Bizerte and the Gulf of Gabes, *Cah. Biol. Mar.*, 47, 253-259.
- Ungaro N., Assennato G., Blonda M., Cudillo B., Petruzzelli M., Mariani M., Pastorelli A. M., Aliquò M. R., D'angela A., Aiello C., Ranier S., Occurrence of the potentially toxic dinoflagellate *ostreopsis ovata* along the apulian coastal areas (southern Italy) and relationship with anthropogenic pollution, *Fresenius Environmental Bulletin*, 19(9), 1813-1821.
- Vairo T., Curro' F., Scarselli S., Fabiano B., 2014, Atmospheric emissions from a fossil fuel power station: dispersion modelling and experimental comparison, *Chemical Engineering Transactions*, 36, 295-300.
- Vairo T., Magri S., Qualgliati M., Reverberi A., Fabiano B., 2017a, An Oil Pipeline Catastrophic Failure: Accident Scenario Modelling and Emergency Response Development, *Chemical Engineering Transactions*, 57, 373-378.
- Vairo T., Del Giudice T., Quagliati M., Barbucci A., Fabiano B., 2017b, From land- to water-use-planning: A consequence-based case-study related to cruise ship risk, *Safety Science*, 97, 120-133.
- Vila M., Garcés E., Masó M., 2001, Potentially toxic epiphytic dinoflagellate assemblages on macroalgae in the NW Mediterranean, *Aquat. Microb. Ecol.*, 26(1), 51-60.
- Wells M.L., Karlson B., Wulff A., Kudela R., Asnaghi V., Berdalet E., Cochlan W., Davidson K., De Rijcke M., Dutkiewicz S., Hallegraef G., Flynn K.J., Legrand C., Paer H., Silke J., Suikkanen S., Thompson P., Trainer V., Trick C., 2020, Future HAB Science: Directions and Challenges in a Changing Climate, *Harmful Algae*, 91.