CITIZEN SCIENCE AND MACHINE LEARNING TO FORECAST *OSTREOPSIS* cf. *OVATA* BLOOMS IN THE ITALIAN SOUTHERN ADRIATIC SEA

Maddalena de Virgilio, Patrizia Borrello, Emanuela Spada, Pasquale Cataldo, Salvatore Cifarelli, Giuseppe Garofoli, Grazia Lamberti, Bernard Degryse, Ennio Ottaviani

Abstract: The toxic benthic dinoflagellate *Ostreopsis* cf. *ovat*a causes harmful algal blooms in Apulia region of Southern Italy.
In this study the volunteers of a citizens' observatory engaged with public research centres to apply a machine learning approach and develop a predictive modelling tool able to forecast *O. ovata* blooms.
We applied the Quantile Regression Forest to draw up two models named Model4Cities and Citizens'Model.
Model4Cities was trained with data of cell abundance detected by the Regional Agency of Environmental Protection of Apulia from 2010 to 2022 in the cities of Bisceglie, Molfetta, Giovinazzo and Bari where the microalgae proliferates at high rates.
Citizen'sModel was trained with data of cell abundance detected by citizens at two sites within the coastline of Molfetta from 2016 to 2022.
Both models show a good capacity to forecast *O. ovata* concentrations as function of meteorological open data with 81 % and 89 % of prediction accuracy.

Keywords: *Ostreopsis ovata*, forecast, citizen science, machine learning, Apulia Region

Maddalena de Virgilio, CNR-IBBR, Institute of Bioscience and Bioresources, Italy, maddalena.devirgilio@ibbr.cnr.it, 0000-0001-5777-3651
Patrizia Borrello, ISPRA - Italian Institute for Environmental Protection and Research, Italy, patrizia.borrello@isprambiente.it, 0009-0004-6217-5786
Emanuela Spada, ISPRA - Italian Institute for Environmental Protection and Research, Italy, emanuela.spada@isprambiente.it, 0009-0002-1494-5171
Pasquale Cataldo, CNR-IBBR, Institute of Bioscience and Bioresources, Italy, pasquale.cataldo@ibbr.cnr.it
Salvatore Cifarelli, Osservatorio del Mare a Molfetta, Italy, cifarellisalvatore07@gmail.com
Giuseppe Garofoli, Osservatorio del Mare a Molfetta, Italy, mfilati@tiscalinet.it
Grazia Lamberti, Osservatorio del Mare a Molfetta, Italy, grazia.lamberti@hotmail.it
Bernard Degryse, Osservatorio del Mare a Molfetta, Italy, bdegryse@yahoo.com
Ennio Ottaviani, University of Genoa, Italy, ennio.ottaviani@edu.unige.it, 0000-0002-2538-5985

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Maddalena de Virgilio, Patrizia Borrello, Emanuela Spada, Pasquale Cataldo, Salvatore Cifarelli, Giuseppe Garofoli, Grazia Lamberti, Bernard Degryse, Ennio Ottaviani, *Citizen science and machine learning forecasting* Ostreopsis cf ovata *blooms*, pp. 881-891, DOI: 10.36253/979-12-215-0556-6.76

Introduction

Harmful Algal Blooms (HABs) are detrimental to humans, marine organisms and the environment [1].

Ostreopsis cf. *ovata* Fukoyo, 1981 belonging to the genus Ostreopsis Schmidt (Schimdit 1902) is one of the most common microalgae that produces palytoxin-like toxins (ova-toxin) [2, 3].

Ova-toxin represents a threat for human health via entry into the food chain, inhalation or direct contact [4,5,6].

Since the 1990s, *O. ovata*-generated HABs have been reported in the Mediterranean Sea particularly in Spain, France and Italy, sometime affecting more than 200 individuals [7,8,9,10].

O. ovata has been detected all over the Italian coast [3, 4, 11, 12].

Since summer 2007, in Apulia region of Southern Italy, the Regional Agency of Environmental Protection (ARPA-Puglia) performs a regular monitoring of the microalgae from June to September, included in the national surveillance monitoring (L 979/82, DM 30/3/2010, DM 19/4/2018). The aim of this monitoring is to provide alarms through public information in order to prevent the intoxication of beach users (DM 19/4/2018).

The cause of *O. ovata* blooms are not completely understood and the relationship with eutrophication and climate change has been hypothesised [13,14].

Pollutants may also serve as a "Go" signal favouring the growth of one planktonic species over the others [15, 16, 17].

Noteworthy, stretches of the coast impacted by intense anthropic exploitation are in general characterized by high proliferation rate of the benthic microalgae. As a consequence, human standings at the urban beaches are more exposed to the risk of intoxication [18].

However, as shown by separate studies, meteo-marine parameters play a pivotal role in permitting *Ostreopsis* proliferation.

A preliminary descriptive model of *O. ovata* blooms events in the Ligurian sea highlights a relevant role of seawater temperature and hydrodynamics [19].

In a machine learning based approach, data of *O. ovata* concentration and data provided by weather forecasting model operated by ARPA-Liguria were used to develop a predictive tool to predict HABs in this region of Nothern-Eastern Italy [20].

More recently, the statistical correlation between weather data freely available in open access databases (iLMeteo, seatmperature.info) and *O. ovata* concentrations detected in Molfetta allowed the identification of dew point as a new parameter that can be used to forecast blooms of the microalgae [21].

In this study the approaches citizen science and machine learning were combined to predict *O. ovata* abundances.

To this purpose the ensamble machine learning method Quantile Regression Forest (QRF) was applied to built two distinct models named Model4Cities and Citizens'Model [22].

Both models were trained to predict *O. ovata* cell densities in seawater as function of three meteorological parameters such as air temperature, seawater temperature and dewpoint deposited in iLMeteo and seatemperature.info open access archives.

Model4Cities was trained with data of *O. ovata* concentrations disclosed on the institutional website of ARPA-Puglia from summer 2010 to summer 2022 in 4 cities of Apulia where the microalgae proliferates at high rates (i.e. Bisce glie, Molfetta, Giovinazzo and Bari).

Citizens'Model was trained with data of *O. ovata* concentrations detected by citizens from summer 2016 to summer 2022 at two public beaches in Molfetta (i.e. Prima Cala and Gavetone).

Both models show a good capacity to forecast *O. ovata* concentrations and the related threshold exceedance probability as function of meteorological and marine open data with 81 % and 89 % prediction accuracy respectively, and are the base of developing an early-warning system for bloom events.

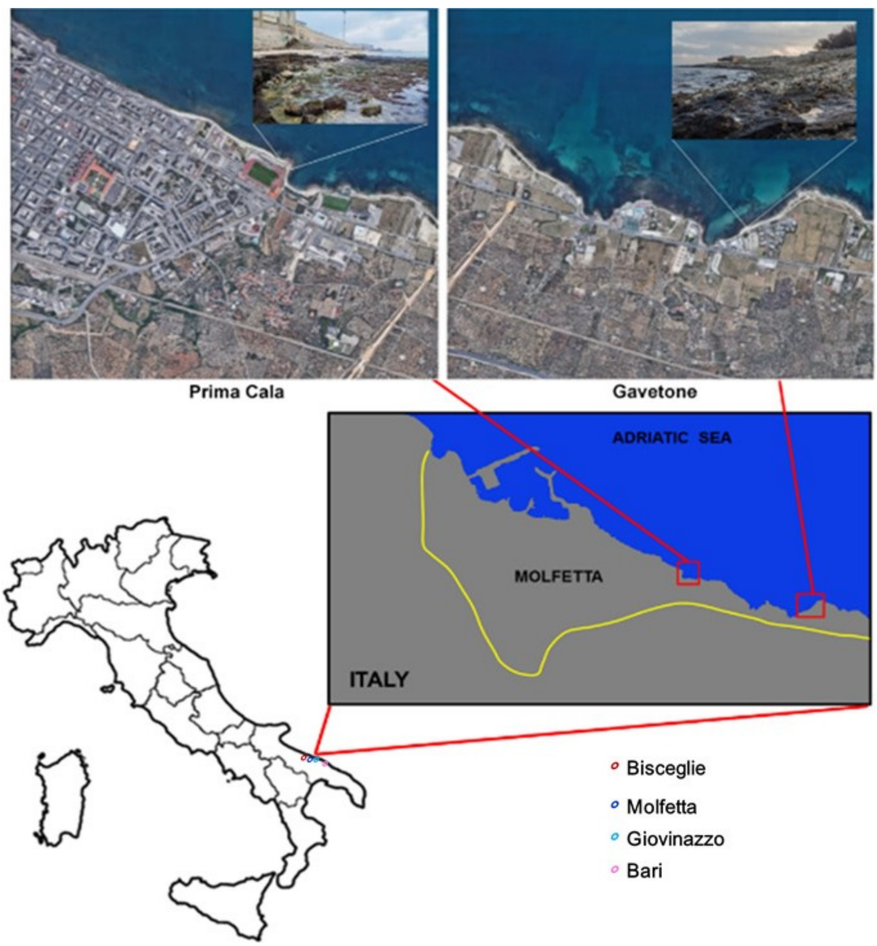

Figure 1 – Distribution of the 4 sampling sites along the coastline of Apulia region: Bisceglie, Molfetta, Giovinazzo, Bari, Molfetta-Prima Cala, Molfetta-Gavetone.

Materials and Methods

Study site

Since summer 2007, ARPA-Puglia carries out a survey along the coast of Apulia region from June to September, sampling seawater every two weeks.

In this study we have analysed the data of *O. ovata* abundance detected by ARPA-Puglia in the cities of Bisceglie (41°14'26,016" N;16°31' 35,4" E), Molfetta (41°12'0,576" N; 16°36'59,076"E), Giovinazzo (41°10'25,32"N; 16°42'47,16"E) and Bari (41°6'29,16N; 16°56'15"E) from summer 2010 to summer 2022 (Figure 1). These 4 cities are located along a coastal stretch permanently subjected to the impact of large urban centres including the metropolitan city of Bari which itself accounts for 316 015 inhabitants and 2692 inhabitants km^{-2}.

In Molfetta, in a previous study, two sampling stations where chosen, namely Prima Cala (41◦11'5 N;16◦36'54"E) and Gavetone (41◦11'37N;16◦38'11"E) that are two public beaches crowded in summertime and separated from each other by a distance of 5 km. Prima Cala is the sampling station closer to the urban centre than Gavetone [21]. Seawater sampling in Molfetta was performed by citizens on a weekly basis.

Citizen science observatory/community-based environmental monitoring

Osservatorio del Mare a Molfetta (OMM, observatory of the Sea in Molfetta) is a community of volunteers promoting through the participative methodology of citizen science: i) the protection of coastal and marine ecosystem at urban level; iì) the public's rights to access to environmental information in order to empower citizens to make key decisions relating to environmental issues. In addition, OMM provides incentives to scientists to engage with citizens to establish the absence or presence of a cause-and-effect association between marine biodiversity loss and habitual pollution/consumption practices [21,23].

In this study, using their own devices, trained OMM-volunteers collected values of meteorological parameters and *O. ovata* concentrations from open access databases (www.arpapuglia.it; www.ilmeteo.it; www.seawatertemperature.info).

Meteorological parameters and statistical analysis

We have compared and statistically correlated the concentration of *O. ovata* bottom sampled by ARPA-Puglia from summer 2010 to summer 2022 with the mean of 7 meteorological parameters measurements during a period of 15 days in the past respecting to the date of seawater sampling at the cities of Bisceglie, Molfetta, Giovinazzo and Bari. The statistical correlation between the same weather parameters and the values of *O. ovata* density previously detected by citizens in two sampling station in Molfetta, i.e. Prima Cala and Gavetone, from July 2016 to December 2020 is described in de Virgilio et al. 2021[21]. The weather parameters that have been used herein were: average air temperature, seawater temperature, relative humidity (RH, proportion of humidity present in the air expressed in percent), air pressure (also known as atmospheric pressure or barometric pressure is the force exerted by the weight of the air per unit of area of the surface of Earth), dew point (at constant air pressure, it is the temperature at which air become saturated with water vapour, i.e. RH = 100%), maximal wind (or wind flow speed is the maximal speed of the wind), average wind (the mean of speed of the wind) and sea water temperature.

Statistical analysis was performed with the GraphPad Prism software using an ANOVA model followed by Tukey's post-test. Values $p< 0.05$ were considered significant, and $p< 0.01$ highly significant. Correlation between two variables was calculated using the Spearman's rank correlation coefficient.

Construction of the predictive model

In this study we applied the Quantile Regression Forest (QRF), an ensemble machine learning method, to build two predictive models of *O. ovata* concentration in sea water, using MATLAB as programming framework [22].

We selected the most important environmental predictors deduced from the statistical correlation (i.e. sea water temperature, average air temperature and dew point) and included into the model the values collected up to 15-days before the date of sampling. In this way, we aimed to capture the temporal evolution of the environmental variables with expected influence on concentration dynamics.

After a rank correlation analysis, we selected three time lags (2, 5 and 9 days) in order to cover the observed time span in days with a small number of values. Correlation between concentration and environmental variables peaked at 9 days in the past.

Doing so, the final vector used to define the regression model included 11 predictors:

- station ID (categorical)
- day of the year (numeric with range 1-365),
- 3 values for sea water temperatures corresponding to the 2nd, 5th and 9th day before test date;
- 3 values for average air temperatures corresponding to the 2nd, 5th and 9th day before test date;
- 3 values for dew point temperatures corresponding to the 2nd, 5th and 9th day before test date.

All features were encoded as continuous variables, except for station ID that was treated as a categorical variable with a number of unordered possible values corresponding to the different sites used in the training. We considered the station ID as a predictor since physical or geographical characteristics of the site may cause differences in the behaviour of *O. ovata* detected in bottom sampled sea water. The response variable was processed by the base-10 logarithm in order to deal with the large variation in concentration dynamics.

A total of 468 and 216 entries were present for the whole dataset of Model4Cities and Citizens' Model, respectively. The candidate blooms were defined with a threshold of 100 000 cell/L on *O. ovata* concentration in bottom sampled sea water.

Results were cross-validated using the out-of-bag (OOB) approach, as a common practice with tree-based models. In this way, a global assessment of the predictive capability of the model was generated. In a second step we also checked the temporal effectiveness of the method by splitting the dataset by year and using data of 2021 and 2022 in Model 4cities and data of 2020 and 2022 in Citizens' Model as an independent test data set.

Accuracy has been evaluated using MAE error metrics for the log-concentration.

Results

Correlations with meteorological parameters

We have previously demonstrated that at two distinct site along the coast of Molfetta, namely Prima Cala and Gavetone, the values of three specific weather parameters such as seawater temperature, dew point and, at a latter extent, air temperature exhibit a strong positive correlation with *O.* cf. *ovata* concentrations detected by volunteers [21].

In a like manner, as shown in table 1, calculation of Spearman's coefficient showed that the proliferation of the microalgae in the cities of Bisceglie, Molfetta, Giovinazzo and Bari detected by ARPA-Puglia from summer 2010 to summer 2022 positively correlates better with seawater temperature (ρ=0,52 at Bisceglie, ρ=0,55 at Molfetta, ρ=0,48 at Giovinazzo, ρ=0,59 at Bari) and dew point (ρ= 0,47 at Bisceglie, ρ= 0,53 at Molfetta, ρ= 0,48 at Giovinazzo, ρ= 0,52 at Bari) than air temperature (ρ= 0,24 at Bisceglie, ρ=0,35 at Molfetta, ρ= 0,4 at Giovinazzo, ρ= 0,37 at Bari).

Correlation coefficient very close to zero (and not significant) demonstrate that there is no relationship between high concentrations of *O. ovata* and the other weather parameters such as relative humidity, air pressure, average and maximal wind speed at the 4 sites considered in this study (Table 1).

Table 1 – Correlation between concentrations of *O. ovata* and weather parameters. Values of Spearman's correlation coefficient (ρ) are shown. Statistical significance is also indicated. p<0,05 is considered significant and p<0,001 highly significant.

Weather parameters	Bisceglie	Molfetta	Giovinazzo	Bari
Air temperature °C	0,24 p=0,0162	0,35 p=0,0007	0,4 p<0,0001	0,37 p<0,0001
Seawater temperature °C	0,52 p<0,0001	0, 55 p<0,0001	0,49; p<0,0001	0,59; p<0,0001
Dew point °C	0,47; p<0,0001	0,53 p<0,0001	0,48 p<0,0001	0,52 p<0,0001
Relative humidity (%)	0,1 p>0,05	0,06 p>0,05	-0,04 p>0,05	0 p>0,05
Air pressure (mb)	-0,03 p>0,05	-0,07 p>0,05	-0,02 p>0,05	-0,08 p>0,05
Wind average (km/h)	0,01 p>0,05	0,11 p>0,05	0,13 p>0,05	0,11 p>0,05
Wind maximal (km/h)	0,07 p>0,05	0,04 p>0,05	0,09 p>0,05	0,06 p>0,05

True and false positive rate

Data of *O. ovata* concentration collected from 2010 to 2022 by Arpa-Puglia were used to train Model4Cities where a total of 468 entries were present in the whole dataset including 144 candidate blooms throughout Bisceglie, Molfetta, Giovinazzo and Bari.

Data collected by citizens at Molfetta from 2016 to 2022 were used to train Citizen'sModel where a total of 216 entries were present in the whole data set with 81 candidate blooms throughout Prima Cala and Gavetone.

In both models, the probability at which an in situ *O. ovata* concentration above 100 000 cells/L in bottom sampled water was correctly predicted ("True bloom condition", i.e. True Positive Rate, TPR) was computed for the 50th quantile of the prediction in OOB mode. On the other way round, the probability of wrongly predicted bloom event when detected concentrations were below the specific threshold of 100 000 cells/L ("False bloom" conditions, i.e. False Positive Rate, FPR) was computed. Results were then plotted in Receiving Operating Characterisic (ROC) curves as TPR versus FPR for both models.

As shown in Figure 2A, the values of the Area Under Curve (AUC= 0,719; 0,933; 0,868; 0,902; 0,854 at Bisceglie, Molfetta, Giovinazzo, Bari and Mean, respectively) indicate a good capacity of Model4Cities to discriminate between positive and negative classes. In a similar trend, as shown in Figure 2B, the AUC values of Citizen'Model ROC curves (AUC=0,918; 0,957; 0,938 at Prima Cala, Gavetone and Mean, respectively) indicate that this model is endowed of an even better class separation capacity than Model4Cities.

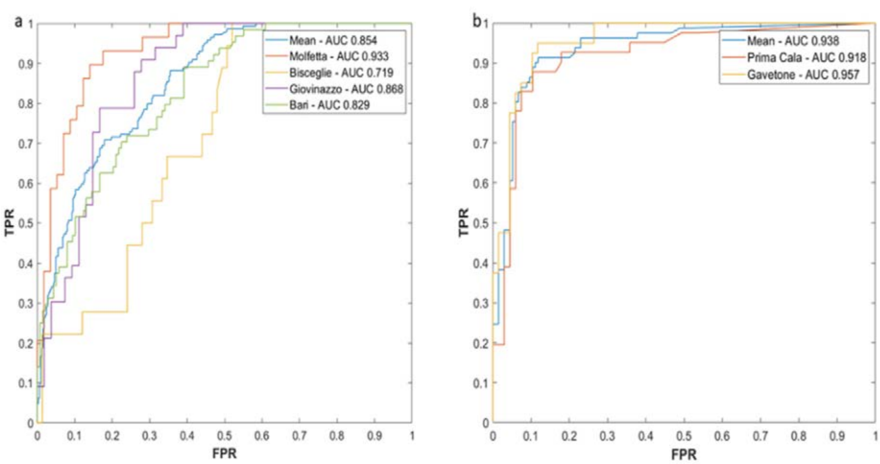

Figure 2 – ROC curves and AUC values for Model4Cities (a) and Citizen's Model (b). AUC > 0,5 indicates an acceptable class separation capacity, AUC=1 indicates optimal class separation capacity.

Confusion matrices were computed to provide a quantitative measurement of the accuracy and the errors made by the entire training datasets of both QRFs, comparing the predictions values with the actual measurements of *O. ovata* abundance detected by Arpa Puglia in Model4Cities or by citizens in Citizens'Model. These tests evidenced a good level of performance of both models with an accuracy of 81 % in Model4Cities (81 % for Bisceglie, 88 % for Molfetta, 80 % for Giovinazzo, 77 % for Bari) and an accuracy of 89 % in Citizens'Model (0,88 % at Prima Cala, 0,9 % at Gavetone).

Model Validation

Model4Cities validation was performed by generating predictions of *O. ovata* concentrations in bottom sampled water as function of the value of weather parameters occurring in summers 2022 and 2023 (77 samples) with a QRF trained with data spanning from summer 2010 to summer 2021 (431 samples) in Bisceglie Molfetta, Giovinazzo and Bari (mean AUC value=0,72). The values of predicted *O. ovata* concentrations were then compared with the actual cell concentrations detected by ARPA-Puglia and a confusion matrix of *O. ovata* concentrations threshold exceedance prediction was computed. With this data set, the accuracy of prediction of Model4Cities for summer 2022 and 2023 was 0,79 % throughout Bisceglie, Molfetta, Giovinazzo and Bari.

In a similar manner, the accuracy of predictions produced by Citizen'sModel in Molfetta at Prima Cala and Gavetone was validated. In this case, data corresponding to summers 2020 and 2022 (72 samples) were tested with a QRF trained from summer 2016 to summer 2019 (144 samples, mean AUC value=0,87). The values of predicted *O. ovata* concentrations were then compared with the actual cell densities detected by citizens in Molfetta in summers 2020 and 2022. The confusion matrix computed on this dataset indicates an accuracy of predictions for summers 2020 and 2022 of 81 % throughout Prima Cala and Gavetone.

Validation results were quite similar to those obtained with OOB procedure; therefore we are confident about the good generalization ability of the two models.

Discussion

Herein, we report the results of a collaboration established between volunteers of a citizens' observatory and public research centres with the aim of developing a forecasting system of *O. ovata* blooms in 4 cities of Apulia region of Southern Italy where the toxic microalgae proliferates at high rates affecting citizens standing on the coast.

Although it is well known that *O. ovata* represents a certain harm to human health and marine fauna, the dynamics of its blooms are still poorly understood [24].

Noteworthy, Pfannkuchen and colleagues have demonstrated that *Ostreopsis* blooms might remain undetected within conventional monitoring programs with high potential to affect human health [25]. Therefore, the development of an efficient predictive model to forecast *O. ovata* HABs is needed to reduce their impact on citizens' health.

In Apulia, since summer 2007, ARPA-Puglia performs a regular monitoring of the proliferation of the dinoflagellate in bottom and column sampled seawater every two weeks from June to September [26], included in the national surveillance monitoring (L 979/82, DM 30/3/2010, DM 19/4/2018).

Concomitantly, trained volunteers of a citizens' observatory monitored the microalgae from summer 2016 to summer 2022 at two different sites along the coast of Molfetta sampling seawater on a weekly basis [21].

In this study, we applied a machine learning approach to draw up two predictive models of *O. ovata* cell densities.

Herein, we show that a small set of meteorological parameters, such as air temperature, seawater temperature and dew point are sufficient to predict *Ostreopsis* HABs in the coastline stretch of Apulia including the cities of Bisceglie Molfetta, Giovinazzo and Bari.

In fact, both QRF exhibited a good level of performance as shown by the values of accuracy of predictions which were 81% and 89% for Model4Cities and Citizens'Model, respectively.

In both QRFs the response variable was the ten logarithm of *Ostreopis* cells concentration in bottom sampled seawater.

In the monitoring plan of ARPA Puglia, 30,000 cells/L is the threshold of *O. ovat*a concentration in the water column to disclose to the public warning about the onset of risks to human health (http://www.arpa.puglia.it, ISTISAN 14/9). Yet, we chose to consider values of cell concentration detected in bottom sampled water since they are less influenced by hydrodynamism in respect to values of cell density occurring in column water. Therefore, the most appropriate response variable would be the value of cell concentration in bottom sampled seawater **[**19].

In fact, *O. ovata* is a benthic species and its occurrence in the water column is due to the detachment from the bottom substrate and consequent resuspension in the water column as a result of water turbulence and/or of exceeding cells density on the bottom substrate [4, 19].

In addition, there is a tight coupling between the magnitude of *O. ovata* density in the water column and bottom sampled seawater: the higher is the concentration on the substrate, the higher is the abundance in the water column.

Predicted concentrations higher than the threshold value of 100 000 cells/L were considered as bloom events. Indeed this threshold value, chosen on the base of personal experience and previous observation [26], provided a high level of replicability in the prediction of blooms and increased the statistical inference of the two models.

Conclusion

We have performed a study on *O. ovata* proliferation dynamics using 10-years environmental data deposited in open access archives.

O. ovata blooms may remain undetected by conventional monitoring programs.

A small set of weather parameters such as seawater temperature, air temperature and dew point of previous days are sufficient to predict *O. ovata* blooms with good accuracy.

In this study, advances in knowledge have been achieved to improve the forecasting of *O. ovata* blooms in 4 cities of Apulia region of Southern Italy.

Model4Cities and Citizens'Model represent a cost-effective and reliable tool to forecast *O. ovata* concentrations and pave the way to the development of an early-warning system of bloom events.

Since both models are based on weather forecasts stored in open access archives, they can be easily scaled up at national and international level in all the areas affected by *O. ovata* HABs.

References

[1] Faimali M., Giussani V., Piazza V., Garaventa F., Corrà C., Asnaghi V., Privitera D., Gallus L., Cattaneo-Vietti R., Mangialajo L., Chiantore M. (2012) - *Toxic effects of harmful benthic dinoflagellate Ostreopsis ovata on invertebrate and vertebrate marine organisms*. Mar. Env. Res. 76, 97-107.

[2] Fukuyo Y. (1981) - *Taxonomical study on benthic dinoflagellates collected in coral reefs.* Bull. Jpn. Soc. Sci. Fish. 47, 967-978.

[3] Penna A., Fraga S., Battocchi C., Casabianca S., Perini F., Samuela C., Casabianca A., Riobò P., Giacobbe M.G., Totti C., Accoroni S., Vila M., René A., Scardi M., Aligizaki K., Nguyen-Ngoc L., Vernesi C. (2012) - *Genetic diversity of the genus Ostreopsis Schmidt: phylogeographical considerations and molecular methodology applications for field detection in the Mediterranean Sea*. Cryptogam. Algol. 33: 153-163.

[4] Totti C., Accoroni S., Cerino F., Cucchiari E., Romagnoli T. (2010) - *Ostreopsis ovata bloom along the Cornero riviera (northern Adriatic Sea): relationships with environmental conditions and substrata*. Harmful Algae 9: 233-239.

[5] Vila M., Arin L., Battocchi C., Bravo I., Fraga S., Penna A., Reñé A., Riobó P., Rodriguez F., Sala M.M., Camp J., de Torres M., Franco J.M. (2012) - *Management of Ostreopsis blooms in recreational waters along the Catalan coast (nw MediterraneanSea): cooperation between a research project and a monitoring program.* Cryptogam: Algol. 33: 143-152.

[6] Pelin M., Forino M., Brovedani V., Tartaglione L., Dell'Aversano C., Pistocchi R., Poli M., Sosa S., Florio C., Ciminiello P., Tubaro A. (2016) - *Anatoxin-a, a palytoxin analogue isolated from Ostreopsis cf. ovata Fukuyo: cytotoxic activity and ELISA detection.* Environ. Sci. Technol. 50: 1544-1551.

[7] Gallitelli M., Ungaro N., Addante L.M., Procacci V., Silveri N.G., Sabba C. (2005) - *Respiratory illness as a reaction to tropical algal blooms occurring in a temperate climate.* JAMA. 293, 2599-2600.

[8] Kermarec F., Dor F., Armengaud A., Charlet F., Kantin R., Sauzade D., Giannetti S., de Haro L. (2008) - *Les risques sanitaires liés à la présence d'Ostreopsis ovata dans les eaux de baignade ou d'activités nautiques*. ERS. 7, 357-363.

[9] Blanfuné A., Boudouresque C.F., Grossel H., Thibaut T. (2015) - *Distribution and abundance of Ostreopsis spp. and associated species (Dinophyceae) in the northwestern Mediterranean: the region and the macroalgal substrate matter.* Environ. Sci. Poll. Res. 22, 12332-12346.

[10] Ninčević Gladan Ž., Arapov J., Casabianca S., Penna A., Honsell G., Brovedani V., Pelin M., Tartaglione L., Sosa S., Dell'Aversano C., Tubaro A., Žuljević A., Grbec B., Čavar M., Bužančić M., Bakrač A., Skejić S. (2019) - *Massive Occurrence of the Harmful Benthic Dinoflagellate Ostreopsis cf. ovata in the Eastern Adriatic Sea.* Toxins. 11, 300.

[11] Mangialajo L., Ganzin N., Accoroni S., Asnaghi V., Blanfuné A. (2011) - *Trends in Ostreopsis proliferation along the Northern Mediterranean coasts*. Toxicon. 57, 408-420.

[12] ISPRA (2023) - *Monitoraggio della microalga potenzialmente tossica Ostreopsis cf. ovata lungo le coste italiane Anno 2022.* Rapporti 396/2023. https://www.isprambiente.gov.it/it/pubblicazioni/rapporti/

[13] Hallegraeff G.M. (2008) - *Harmful algal blooms, coastal eutrophication and climate change.* Biol. Mar. Medit, 15(1), 6-15.

[14] Wells M.L., Karlson B., Wulff A., Kudela R., Asnaghi V., Berdalet E., Cochlan W., Davidson K., De Rijcke M.,Dutkiewicz S., Hallegraeff G., Flynn K.J., Legrand C., Paper 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, 101632.

[15] Kim Y., Son J., Lee Y.-S., Wee J., Lee M., Cho K. (2020) - *Temperature-dependent competitive advantages of an allelopathic alga over non-allelopathic alga are altered by pollutants and initial algal abudance levels.* Sci. Rep. 10, 4419.

[16] Tester P.A., Litaker R.W., Berdalet E. (2020) - *Climate change and harmful benthic microalgae.* Harmful Algae. 91, 101655.

[17] Fabri-Ruiz S., Berdalet E., Ulses C., Somot S., Vila M., Lemée R., Irisson J.O. (2024) - *Harmful Ostreopsis cf. ovata Blooms Could Extend in Time Span with Climate Change in the Western Mediterranean Sea. In press* DOI: 10.2139/ssrn.4764250

[18] Meroni L., Chiantorea M., Petrillo M., Asnaghi V. (2018) - *Habitat effects on Ostreopsis cf. ovata bloom dynamics*. Harmful Algae. 80, 64-71.

[19] Asnaghi V., Bertolotto R., Giussani V., Mangialajo, L., Hewitt, J., Thrush, S., Moretto, P., Castellano, M., A., Povero, P., Cattaneo-Vietti, R. (2012) - *Interannual variability in Ostreopsis ovata bloom dynamic along Genoa coast (North- Western Mediterranean): a preliminary modelling approach.* Cryptogamie Algologique, 33 (2), 181-189.

[20] Ottaviani E., Magri S., De Gaetano P., Cuneo C. Chiantore M., Asnaghi V., Pedroncini A. (2020) - *A machine learning based approach to predict Ostreopsis cf. ovata bloom events from meteo-marine forecasts*. Chem. Eng. Trans. 82, 409-414.

[21] de Virgilio M., Cifarelli S., Garofoli G., Lamberti G., Massari V., Degryse B. (2021) - *Citizen science in the monitoring of Ostreopsis ovata blooms in southern Italy: A five-year study.* Mar Pol Bulletin. 173, 112981.

[22] Meinshausen N. (2006) - *Quantile regression forests*. J. Mach. Learn. Res. 7, 983–999.

[23] de Virgilio M., Cifarelli S., de Gennaro P., Garofoli G., Degryse B. (2020) - *A first attempt of citizen science in the genetic monitoring of a Posidonia oceanica meadowin the Italian Southern Adriatic Sea.*J. Nat. Conserv. 56, 125826.

[24] Tibiriçá C.E.J.A., Leite I.P., Batista T.V.V., Fernandes L.F., Chomérat N., Herve F., Hess P., Mafra L.L. (2019) - *Ostreopsis cf. ovata Bloom in Currais, Brazil: Phylogeny, Toxin Profile and Contamination of Mussels and Marine Plastic Litter.* Toxins. 11, 446.

[25] Pfannkuchen M., Godrijan J., Pfannkuchen D.M., Iveša L., Kružić P., Ciminiello P., Dell'Aversano C., Dello Iacovo E., Fattorusso E., Forino M., Tartaglione L., Godrijan M. (2012) - *Toxin-producing Ostreopsis cf. ovata are likely to bloom undetected along coastal area*. Environ. Sci. Technol. 46, 5574-5582.

[26] Ungaro N., Assennato G., Blonda M., Cudillo B., Petruzzelli M.R., Mariani M., Pastorelli A.M., Aliquò M.R., D'Angela A., Aiello C., Ranieri S. (2010) - *Occurrence of the potentially toxic dinoflagellate Ostreopsis ovata along the Apulian coastal areas (southern Italy) and relationship with anthropogenic pollution.* Fresenius Environ. Bull. 19, 1813-1821.