Accepted Manuscript in Estuarine Coastal and Shelf Science 249, 5 February 2021, 107112

Gutiérrez-Estrada, J.C., Pulido-Calvo, I., Peregrín, A., García-Gálvez, A., Báez, J.C., Bellido, J.J., Souviron-Priego, L., Sánchez-Laulhé, J.M., López, J.A. 2021. Integrating local environmental data and information from non-driven citizen science to estimate jellyfish abundance in Costa del Sol (southern Spain). Estuarine, Coastal and Shelf Science 249, 107112

Integrating local environmental data and information from non-driven citizen science to estimate jellyfish abundance in Costa del Sol (southern Spain)

[1]Gutiérrez-Estrada, J.C., [1]Pulido-Calvo, I., [2]Peregrín, A., [1]García-Gálvez, A., [3,4]Báez, J.C., [5,6]Bellido, J.J., [5,6]Souviron-Priego, L., [7]Sánchez-Laulhé, J.M., [5]López, J.A.

[1]*Dpto. Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Campus de El Carmen, Universidad de Huelva, 21007 Huelva, Spain.*

[2]*Centro de Estudios Avanzados en Física, Matemáticas y Computación, Universidad de Huelva, 21007 Huelva, Spain.*

[3]*Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Málaga, Puerto Pesquero s/n, 29640 Fuengirola, Málaga, Spain.*

[4]*Investigador asociado de la Facultad de Ciencias de la Salud, Universidad Autónoma de Chile, Chile.*

[5]*Aula del Mar de Málaga, Calle Pacífico 80, E-29004, Málaga, Spain.*

[6]*Departamento de Biología Animal, Facultad de Ciencias, Universidad de Málaga. E-29071. Málaga, Spain.*

[7]*Agencia Estatal de Meteorología, Centro Meteorológico de Málaga, Spain.*

Abstract

Tourism, fishing and aquaculture are key economic sectors of Costa del Sol (southern Iberian Peninsula). The management of these activities is sometimes disturbed by the onshore arrival and stranding of jellyfish swarms. In the absence data on the occurrence of these organisms, it may be interesting to explore data from non-driven systems, such as social networks. The present study show how data in text format from a mobile app called Infomedusa can be processed and used to model the relationship between estimated abundance of jellyfish on the beaches and local environmental conditions. The data retrieved from this app using artificial intelligence procedures (transition network or TN algorithm), were used as input for GAM models to estimate the abundance of jellyfish based on wind speed and direction. The analysis of data provided by Infomedusa indicated that only 30.39% of messages provided by the users had information about absence/presence of jellyfishes in the beaches. On the other hand, the TN processing capacity showed an accuracy level to discriminate messages with information on absence/presence of jellyfish slightly higher than 80%. GAM models considering the wind direction and speed of previous day explained between 37% and 77% of the variance of jellyfish abundance estimate from Infomedusa data. In conclusion, this approach may contribute to the development of a system for predicting the onshore arrival of jellyfish in the Costa del Sol.

Keywords: Gelatinous organisms, Transport, Alborán Sea, Artificial intelligence, Transition networks

Introduction

The monitoring of jellyfish populations is extremely important, given that their presence in specific coastal areas at certain times of the year may interact negatively with artisanal fisheries (Graham et al., 2003; Knowler, 2005; Nastav et al., 2013) and aquaculture farms (Halsband et al., 2018). Direct impacts on social and economic sectors associated with tourism have also been acknowledged (Ghermandi et al., 2015; Nunes et al., 2015; Enriquez and Bujosa-Bestard, 2020), as in the case of Costa del Sol.

With its 185 km of coast, the economy of this coastal area in the south of the Iberian Peninsula is mainly based on sun and sand tourism and, to a lesser extent on artisanal fisheries and aquaculture. In the case of tourism, data from recent years show that the mean number of tourists increased by 17% from 2015, reaching a total of 13 million visitors in 2019, most of whom were sun and sand tourists, with an estimated economic impact of €14,442 million. The convergence of factors such as the apparent increase in jellyfish abundance and certain environmental conditions can lead to jellyfish strandings on the beaches, which may be visually unpleasant and alarming for bathers and even harmful to human health (Ghermandi et al., 2015).

On the other hand, while fishing contributes significantly less to the economy of Costa del Sol (€22.5 million in 2019) (Peláez, 2020), an increase of jellyfish populations may worsen the economy impact on this sector, which is already very sensitive to changes in the environment, regulations and laws. Further, aquaculture is an emerging sector in Costa del Sol, with net revenues of more than €10 million (Junta de Andalucía, 2012), that might suffer large losses due to jellyfish blooms.

Proper planning of the activities associated with these economic sectors is critical for their regular operation. For this reason, to enable adequate planning and implementation of appropriate measures, more data is needed to characterise jellyfish population dynamics and to assess their relationship with variations of environmental factors. Taking this into account, the gathering and analysis of information from non-official sources may be valuable.

Since the first voluntary collaborations of citizens in programmes for recording and processing scientific data, there have been multiple examples showing how this approach can greatly contribute to advance the knowledge of environmental sciences in general, with public involvement in projects related to climate change, invasive species, conservation biology, ecological restoration and monitoring of water quality, and the improvement of management practices for ecosystems disturbed by human activities in particular (Silvertown, 2009). Specifically, networks of committed citizens are useful to record the presence of species around the world in an easy and economical way (Conrad and Hilchey, 2011; Palmer et al., 2017; Snyder et al., 2019; Bellido et al., 2020).

The systems in which the use of data provided by volunteers has been notably successful are those in which the gathering of data necessary to understand their functioning is not feasible from technical and/or economic points of view by standard procedures (Hilchey, 2011; Hidalgo-Ruz and Thiel, 2013). Programmes for the monitoring of changes in populations of gelatinous organisms (like jellyfish) in coastal waters are a clear example of this (UNEP, 1983; Boero et al., 2008; Baumann and Schernwski, 2012; Nordstrom et al., 2019).

A common feature of most of these systems is that the recording and acquisition of data is vertical in nature, that is, the programmes are normally designed as part of an experimental analysis, by technicians and scientists who are specialists in the field of interest. For example, Pires et al. (2018) examined the hypothesis that the wind pattern is the key factor that regulates the transport of *Velella velella* in the coastal waters of Portugal. For this purpose, they analysed data on occurrence provided by observers through the GelAvista citizen science programme (http://gelavista.ipma.pt/), developed and designed by the Portuguese Institute for Sea and Atmosphere (IPMA). Similarly, Komninos (2019) developed a program called Swymn seeking to monitor jellyfish populations around Greek coasts, in which the tasks and contributions were clearly defined, and communicated to final users through the app designed for the purpose.

An alternative to vertical procedures for obtaining information relies on the metanalysis of the information uploaded by users in a non-driven manner to various different mobile apps and social media sites such as Facebook, Google+ and Twitter, among others (Cho and Park, 2012; Jovanovic and Vukelic, 2015; Langeneck et al., 2017; Shiffman, 2018; Sammons, 2019; Bargnesi et al., 2020). The information collected through these sources has, after processing, been successfully used in some species conservation programmes such as the ECOCEAN (Davies et al., 2012) and Angel Shark (Meyers et al., 2017) projects.

The aim of this study was to test the possibility of using data on the presence of jellyfish along Costa del Sol, obtained from a mobile app developed horizontally and not serving as a scientific monitoring programme, called Infomedusa (https://play.google.com/store/apps/details?id=es.infomedusa&hl=es; https://infomedusa.es/). With this objective, we first developed an algorithm to filter and classify the information provided by users. Subsequently, a database of the processed data was integrated with environmental parameters to generate models that would characterize the factors favouring the onshore arrival of jellyfish.

Material and methods

Study area

Costa del Sol extends approximately 185 km along the coast of the provinces of Málaga and Cádiz (Southern Spain), running from La Línea de la Concepción (Cádiz) (5º20'W 36º10'N) on the West, to Nerja (Málaga) (3º52'W 36º45'N) on the East. It has traditionally been divided into two zones, "Western" and "Eastern" Costa del Sol, with the city of Málaga, the capital of both zones, at the centre (Figure 1).

The study area is located in the north-west of the Alborán Sea, which comprise the westernmost part of the Mediterranean Sea. From a hydrological point of view, Alborán Sea is characterised by the presence of two superficial anticyclonic gyres at its eastern and western halves. This is result of the incoming of the Atlantic water jet through the Strait of Gibraltar (Sánchez-Garrido et al., 2013; Oguz et al., 2014). On the other hand, the winds have a clear seasonal regimen. In winter-spring, the winds have a northwestly component. In contrast, south-east component winds are predominant in summer time (Mercado et al., 2012).

In the present study, we have analysed a total of 149 beaches, all in the province of Málaga, between Ancha de Casares beach (5º13'W 36º22'N) on the West and Las Alberquillas beach (3º48'W 36º44'N) on the East. Along this stretch of coast, there are 14 localities with a variable number of beaches in their municipality (Table 1).

Infomedusa App

Infomedusa is a citizen science mobile app for iOS and Android operating systems, developed by the organisation Aula del Mar de Málaga (https://www.auladelmar.info/). Conceived in 2013 as an information system for bathers, it has been downloaded more than 100,000 times since its development. Seeking to create a non-scientific network of local observers along the coast, the app was updated in 2015, including a public chat in which any user is able to collaborate by reporting (text and photographs) from the beach which they are visiting (Figure 2). In this way, registered users are able to access the comments from other users who frequent the same beaches, and obtain real time information on the presence or absence of jellyfish or other similar organisms on the shore, although the app does not have predictive power. These characteristics (doesn't designed with scientific purposes and aimed at a large sector of the population) are typical of a horizontal and non-driven app.

Since Infomedusa was updated in 2015, reports have been generated based on the chats entered by users. In 2018 (between 25 January and 3 October), a total of 40,276 comments were left. Of these, we selected and validated 9,963 (24.7%).

Environmental data

The dominant species in mass strandings of jellyfish along the coast of Málaga during the study period was the mauve stinger (*Pelagia noctiluca*) (Bellido et al., 2020), and hence, we focused the analysis on this species. Further, the use of terms such as swarms or mass strandings of jellyfish in this study refers to high concentrations of *Pelagia noctiluca* near the shore, regardless of their origin, aggregation, population growth or change in distribution (Hamner and Dawson, 2009).

Two of the factors that strongly influence the presence of jellyfish in coastal waters are wind speed and its direction (Goy et al., 1988; Pourjomeh et al., 2017; MacAli et al., 2018). For this study, we have used hourly data on wind speed and direction from four weather stations of the Spanish Meteorological Agency (AEMET, http://www.aemet.es) in the study period (Figure 1).

Algorithm for data retrieval from the Infomedusa app

A rapid and efficient procedure or algorithm for extracting the information entered by users is key to enabling predictive models to be added to the app. In this study, data were processed through a transition network (TN) specifically designed for the domain of interest (Figure 3).

TNs are artificial intelligence algorithms that have been used in different fields of knowledge in an effective manner (Woods, 1973; Anderson, 1982; De Carolis et al., 1996; Gangopadhyay, 2001; Sidhom and Hassoun, 2002; Stehno and Retti, 2003; Gutiérrez-Estrada et al., 2005). They are syntax analysis procedures, in which labels or terms may be associated by means connections that define the transitions and relationships between labels, used to extract information from texts generated in natural language (Thorne et al., 1968; Woods, 1970). The labels define different components of the structure of a subject-verb relationship with specific functions, for example, an ARTICLE (‘a’, ‘an’, ‘the”) or a VERB (‘see’, ‘have’, …).

The TN for processing the information retrieved from Infomedusa only handles information in Spanish, this being the language used in 99.5% of the conversations recorded in the app. It has eight nodes or labels identified as ‘Name’, ‘Adverb of quantity’, ‘Phrase’, ‘Adverb of negation’, ‘Main verb’, ‘Verb *haber* [Spanish verb used for compound tenses]’, ‘Auxiliary verb’ and ‘Reflexive’ linked through 20 subnetworks. Each label is associated with a list or glossary composed of equivalent terms. Overall, the glossary contains 557 word roots related to the domain of interest, of which 322 correspond to the 'Adverb of quantity' label, which is divided into five categories: 'None' (24); 'Few/little' (62); 'Some' (47); 'Many/much' (55); and 'Very many/very much' (134). This allows the TN to differentiate between different quantities of jellyfish on the beach, and in this way, obtain a relative abundance value.

To validate true/false presence/absence, two researchers examined independently each message one by one. For each message, each researcher assigned a value=0 if the message does not contained information about jellyfish and value=1 if the message had information about jellyfish. In this last case, the researchers also evaluated the 'Adverb of quantity' reported by the user and assigned a relative abundance in function of the term used (None=0, Few/little=1, Some=2, Many/much=3, Very many/very much=4). Later, the messages were processed with the TN to extract the information about presence/absence and relative abundance. Then, the TN accuracy was assessed using the Rand index (Rand, 1971) as follows:

$$Precision_{Presence} = P_p = \frac{True\ presence}{True\ presence + False\ presence} \quad (1)$$

$$Precision_{Absence} = P_a = \frac{True\ absence}{True\ absence + False\ absence} \quad (2)$$

Modelling techniques and data processing

To explore the relationship between the qualitative abundance of jellyfish along the coast and wind conditions, we used generalised additive models (GAMs) (Hastie and Tibshirani, 1990). These models are an extension of a generalised linear model, differing from this type of model in two key aspects: (i) the distribution of the dependent or response variable may (explicitly) be non-normal and non-continuous; and

(ii) the values of the independent variables are estimated based on a linear combination of predictor variables connected to the dependent variable by a link function. In this way, in GAMs each independent variable is associated with a non-parametric general function, rather than individual parameters as in a multiple linear regression. This makes it possible to establish nonlinear relationships between independent and dependent variables (Schimek, 2000).

Particularly, in the present work we used a normal distribution function for the dependent variable and identity link function. Four degrees of freedom for the cubic spline smoother covariates set were applied to each continuous predictor variable. All GAM models were calibrated using Statistica 7.0 (StatSoft, Inc.).

Beaches were grouped in five sets according to the geographical location and the volume of data available such that all the groups were balanced (Table 1). Before calibrating the models, the time series grouped by beaches were smoothed by a 7-day bandpass filter. Freón et al. (2003) and Gutiérrez-Estrada et al. (2009) recommend this smoothing process to remove high-frequency noise due to gaps in the time series.

Results

Efficiency of the TN

Out of the 9963 comments, the TN extracted 3028 that contained information concerning the presence/absence and/or relative abundance of jellyfish. This represents a relatively small percentage (30.39%) of the information provided by users (Table 1). The highest rate of inclusion of this type of information was observed among users of the beaches in Fuengirola (38.69%), while only 24.51% of the comments from users visiting Vélez-Málaga beaches provided useful information.

The TN had an accuracy of 80.1%. Among the 19.9% of occasions on which the TN provided an incorrect answer, 12.7% were false positives and 87.3% false negatives. Hence, the TN had an accuracy of 81.1% (P_p=0.811) for positive cases (ones that report the jellyfish presence) and 80.0% (P_a=0.800) for negative cases (ones that report the jellyfish absence).

The sources of errors, listed in order of importance were: (i) user spelling errors; ii) user syntactic errors; iii) word roots not included in the database; iv) contingent comments associated with questions posed by users or referring to the past or the future; (v) syntactic constructions not included in the TN; and (vi) comments made in other languages. Taking into account only errors due to use of syntactic constructions not included in the TN, the accuracy increases to 89.2%. That is, 9.1% of the errors were attributable to syntactic subnetworks beyond the scope of the TN.

Estimation of the relative presence of jellyfish on the coast. GAMs

A total of 707 valid comments were available for the beaches of the locality of Málaga between 4 April and 24 July 2018 (Figure 4a). For this set of data, a GAM including wind speed and direction on the previous day yielded statistically acceptable results with a total deviance of 14.66 and explained variance of slightly over 77.5%.

The analysis of the partial residuals indicates that jellyfish strandings on the beaches of the locality of Málaga tend to occur when there are southern winds (180º) at a speed above 14 km/h. As the wind shifts to the east, the presence of jellyfish decreases (Figures 4b and 4c). The same effect is also observed on Group 2 beaches, for which there is a similar amount of information to that obtained for Málaga (Figure 5a). For this set of beaches, the model explains more than 75% of the variance taking into account wind direction and speed on the previous day (Figures 5b and 5c).

In the case of the group 3 beaches (Figures 6a), with a similar level of information to that available for the aforementioned groups, wind direction and speed have a significantly less marked effect than in these other groups. In this case, the model was only able to explain 37% of the variance. The onshore abundance of jellyfish is highest when the wind blows perpendicular to the coast (170º) and decreases drastically when it veers to the East (Figures 6b, 6c).

The situation is markedly different for beaches in groups 4 and 5. In the case of group 4 (locality of Marbella), with a total 554 valid comments (Figure 7a), a model that, as in the aforementioned cases, considers wind direction and speed during the previous day, indicates that jellyfish strandings on the beaches of Marbella are more likely with easterly winds at a speed of 10 to 14 km/h (Figures 7b,7c). This model explains more than 41% of the variance. In the case of the group 5 beaches (localities of Casares, Manilva and Estepona; N=242 records; Figure 8a), the model indicates that wind direction and speed had greater weight than in the case of group 4 beaches (explaining 70% of the variance). The winds associated with jellyfish reaching the coast are mainly from the northeastern quadrant, specifically between 100º and 140º, and light with speeds of less than 5.5 km/h (Figures 8b, 8c).

Figure 9 shows the time series of normalised relative abundance of jellyfish on the beaches from groups 3, 4 and 5. Although the data series are not particularly large, we can observe an alternation in cycles between the beaches located towards the west and those towards the East.

Discussion

The main objectives of the present study were to retrieve valid data on jellyfish strandings on the beaches along the coast of the province of Málaga using a non-driven app as well as to assess whether this information can be used as a source of data to identify local factors that may drive these gelatinous organisms towards the coast.

Information and data extraction methods

Our results indicate that the information provided by people who use the app may be very useful for the characterisation and estimation of jellyfish arrival, although this information is strongly determined by the climatic conditions associated with the holiday periods, as well as the data extraction method used.

The automatization of the process is a key factor if the goal is that these types of app have a forecast capacity. Accordingly the design of a TN has shown to have statistically acceptable sensitivity and enabled us to efficiently extract information on the presence and relative abundance of jellyfish on the beaches. This is in line with was has been reported has been in other specific fields regarding the use of this type of approach. Gutiérrez-Estrada et al. (2005) used a TN to extract information from disease diagnosis reports written in natural language by fish farm workers. Based on this information, these authors managed to develop an expert system to estimate the presence of diseases in different fish culture tanks more accurately than that achieved by fish farm technicians themselves.

The aforementioned authors reported the relatively high efficiency of TNs, mainly limited by the specific sphere of discourse of the problems addressed. On the other hand, they also identified problems inherent to these systems, which in our case led to 20% of the information not being properly retrieved. For example, there is the clear problem of the expressiveness of the notation used, this means that a significant source of errors was found to lie in the absence of TN routes for the non-fixed expressions of users. This effect is observed mainly in records that first describe weather conditions on the beach, and after using the conjunction "and", go on to describe the presence or absence of jellyfish. This problem is also found in excessively long entries and those that have syntactic sense within a context of a conversation that runs across various different records. De Carolis et al. (1996) identified this problem as a consequence of the gap between the strategic and tactical generation of text, which is ultimately determined by the non-driven nature of the app.

Estimation of the relative abundance of jellyfish along the coast

Several factors hinder the estimation of the relative abundance of jellyfish that arrive to the coasts: (i) absence of targeted scientific data (Tomlinson et al., 2018); (ii) the lack of knowledge of the biological cycles of the species of jellyfish of interest; and as a consequence, (iii) the relationship of the transport mechanism of jellyfish with climatic indices and other environmental factors, such as sea temperature, salinity, currents and wind, in relevant biogeographic and macroecological regions (Bellido et al., 2020).

The macroecological processes of Costa del Sol are part of the oceanographic dynamics of the Alborán Sea. In terms of oceanic circulation, the Alborán Sea is characterised by two anticyclonic gyres created by the flow of surface water from the Atlantic Ocean through the Strait of Gibraltar (Sánchez-Garrido et al., 2013; Oguz et al., 2014). The island of Alborán, to the north of Cape Three Forks, separates the gyre zones. This configuration generates a cyclonic cell that affects the coast of the province of Málaga, promoting Ekman transport and the upwelling of cold and nutrient-rich waters close to the coast (Tintore et al., 1988; Fiala et al., 1994). As a consequence, there is a higher concentration of plankton that tends to attract plankton-eating animals such as jellyfish (Bellido et al., 2020).

On the other hand, the regional circulation patterns in the Alborán Sea alone would not explain the mass strandings of jellyfish on the beaches of Costa del Sol. Bellido et al. (2020) indicate that certain climatic conditions at the regional level strongly influence local surface temperature and salinity, causing the mixed layer depth to vary, reducing Ekman transport and nutrient concentration. These authors associate growth of the population of *Pelagia noctiluca* in the Alborán Sea with the behaviour of two climate indices, namely, the North Atlantic Oscillation (NAO) and Artic Oscillation (AO). Nonetheless, though the use of these indices may facilitate a holistic approach to the abundance of these organisms in the Alborán Sea, it seems evident that local conditions play a key role in the increase of mass strandings of jellyfish on beaches along the coast of Málaga.

The results of our study clearly support this latter hypothesis. As reported by other authors (Bieri, 1977; Graham et al., 2001; Marshall and Burchardt, 2005), under certain regional determinants, local factors such as wind speed and direction tend to favour the arrival of jellyfish to the beaches. In our case, models which are relatively simple in terms of the number of variables entered (wind speed and direction on the previous day) locally explained between 37% and 77% of the variance observed. These results are in line with those obtained by other authors (Bingham and Albertson, 1974). Pires et al. (2018) analysed the abundance and transport mechanisms of *Velella velella* near the coasts of Portugal and found that disruption of upwelling caused by local winds were decisive in the transport of large numbers of individuals towards the coast.

The results of the present study suggest that the presence of swarms of *Pelagia noctiluca* along the eastern coast of the province of Málaga, from Nerja to the city of Málaga, may be explained by south-easterly winds. When these winds dominate and there is an abundance of jellyfish in the Central and Eastern Alborán Sea, jellyfish may reach the coast, often distributed from the east to the west.

With regards to the western coastal zone of the province of Málaga, the weakening of the regional conditions controlled by the intensity of the western anticyclonic gyre tends to favour disruption of the upwelling process, as well as the formation of eddy currents which may transport jellyfish that had previously accumulated in the gyre towards the coast (Sánchez-Garrido et al., 2013; Oguz et al., 2014; Bellido et al., 2020). Under these regional conditions, wind speed and direction can be the key factors that determine the timing and intensity of jellyfish arrival to the shores. This would explain why all the models indicate that the arrival of these organisms is likely when the wind blows perpendicular to the coastline at the local level. On the other hand, under these conditions, our results (Figure 9) support the hypothesis of Bellido et al. (2020) concerning jellyfish movement dynamics that suggests that jellyfish are first transported to the westernmost beaches, close to the Strait of Gibraltar (Group 5), and from there, progressively dragged by currents and pushed by winds towards the eastern beaches (groups 3 and 4).

Conclusions

The results of our study show how data from a non-driven system can be processed and used to obtain statistically acceptable relationships between relative abundance patterns of jellyfish along the beaches of the province of Málaga and local environmental factors. That is, information from citizen science projects may not only add to sets of data recorded under the supervision of scientific and technical programmes, but also provide robust results in the absence of programmes specifically designed for data acquisition.

In the future, the availability of larger volumes of data collected through Infomedusa may help to clarify relative abundance patterns of jellyfish along Málaga's beaches and identify the physical mechanisms involved in the transport of these organisms towards the coast. This is a first step towards the development and deployment of an early warning system that would allow proper planning and management of tourism and fishing along the coast of Málaga.

Acknowledgements

The authors wish to express their gratitude to the Department of Economy and Knowledge of the regional Government of Andalusia (Spain) and to the European Social Fund (ESF) for providing financial support through the 2014-2020 Operational

Program of Youth Employment to Ana García-Gálvez (SNGJ-JPI-051) for a research placement in the Department of Agroforestry Sciences at the University of Huelva (Spain). Part of this work was supported by grant from the Spanish Ministry of Science under project TIN2017-89517-P.

Author statement

Gutiérrez-Estrada J.C.: Conceptualization, Methodology, Software, Validation, Formal Analysis; Data Curation, Writing-Original Draft; **Pulido-Calvo, I.:** Conceptualization, Methodology, Formal Analysis, Writing-Review and Editing; **Peregrín, A.:** Methodology, Software; **García-Gálvez, A.:** Validation, Data Curation; **Báez, J.C.:** Investigation; Writing-Review and Editing; **Bellido, J.J.:** Software, Investigation, Resources; **Souviron-Priego, L.:** Software, Investigation, Resources; **Sánchez-Laulhé, J.M.:** Resources; Data Curation; **López, J.A:** Software, Investigation, Resources

References

Anderson, K.R. 1982. Syntactic analysis of seismic waveforms using augmented transition network grammars. Geoexploration 20(1-2): 161-182.

Bargnesi, F., Lucrezi, S., Ferretti, F. 2020. Opportunities from citizen science for shark conservation, with a focus on the Mediterranean Sea. The European Zoological Journal 87(1): 20-34.

Baumann, S., Schernewski, G. 2012. Occurrence and public perception of jellyfish along the German Baltic coastline. Journal of Coast Conservation 16: 555-566.

Bellido, J., Báez, J., Souviron-Priego, L., Ferri-Yáñez, Salas, C., López, J., Real, R. 2020. Atmospheric indices allow anticipating the incidence of jellyfish coastal swarms. Mediterranean Marine Science 21(2): 289-297.

Bieri, R. 1977. The ecological significance of seasonal occurrence and growth rate of *Velella* (Hydrozoa). Publications of the Seto Marine and Biology Laboratory 24(1/3): 63-76.

Bingham, F.O., Albertson, H.D. 1974. Observation on beach strandings of the Physalia (Portuguese man-of-war) community. Veliger 17: 220-224.

Boero, F., Féral, J.P., Azzurro, E., Cardin, V., Riedel, B., Despalatovic, M., Munda, I., Moschella, P., Zaouali, J., Fonda-Umani, S., Theocharis, A., Wiltshire, K., Briand, F. 2008. Climate warming and related changes in Mediterranean marine biota. Executive summary of CIESM workshop 35.

Cho, S.E., Park, H.W. 2012. Government organizations' innovative use of the Internet: The case of the Twitter activity of South Korea's Ministry for Food, Agriculture, Forestry and Fisheries. Scientometrics 90(1): 9-23.

De Carolis, B., Derosis, F., Grasso, F., Rossiello, A., Berry, D.C., Gillie, T., 1996. Generating recipient-centered explanations about drug prescription. Artificial Intelligence in Medicine 8: 123-145.

Fiala, M., Sournia, A., Claustre, H., Marty, J.C., Prieur, L., Vétion, G. 1994. Gradients of phytoplankton abundance, composition and photosynthetic pigments across the Almeria-Oran front (SW Mediterranean Sea). Journal of Marine Systems 5(3-5): 223-233.

Fréon, P., Mullon, C., Voisin, B., 2003. Investigating remote synchronous patterns in fisheries. Fisheries Oceanography 12 (4/5): 443-457.

Gangopadhyay, A. Conceptual modeling from natural language functional specifications. Artificial Intelligence in Engineering 15(2): 207-218.

Gatt, M.P., Deidun, A., Galea, A., Gauci, A. 2018. Is citizen science valid tools to monitor de occurrence of jellyfish? The spot the jellyfish case study from Maltese Islands. Journal of Coastal Research 85: 316-320.

Ghermandi, A., Galil, B., Gowdy, J., Nunes, P.A.L.D. 2015. Jellyfish outbreak impacts on recreation in the Mediterranean Sea: welfare estimates from a socioeconomic pilot survey in Israel. Ecosystem Services 11: 140-147.

Graham, W.M., Pagès, F., Hamner, W.M. 2001. A physical context for gelatinous zooplankton aggregations: a review. Hydrobiologia 451: 199-212.

Graham, W.M., Martin, D.L., Felder, D.L., Asper, V.L., Perry, H.M. 2003. Ecological and economic implications of a tropical jellyfish invader in the Gulf of Mexico. Biological Invasions 5: 53-69.

Gutiérrez-Estrada, J.C., De Pedro Sanz, E., López-Luque, R., Pulido-Calvo, I. 2005. SEDPA, an expert system for disease diagnosis in eel rearing systems. Aquacultural Engineering 33:110-125.

Gutiérrez-Estrada, J.C., Yáñez, E., Pulido-Calvo, I., Silva, C., Plaza, F., Bórquez, C. 2009. Pacific sardine (*Sardinops sagax*, Jenyns 1842) landings prediction. A neural networks ecosystemic approach. Fisheries Research 100: 116-125.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, Boca Raton.

Hilchey, C.C.C. 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. Environmental Monitoring and Assessment 176: 273-291.

Jovanovic, V., Vukelic, D. 2015. Using geosocial analysis for real-time monitoring the marine environments. Journal of Environmental Protection and Ecology 16(4): 1344-1352.

Knowler, D. 2005. Reassessing the costs of biological invasions: *Mnemipsis leidyi* in the Black sea. Ecological Economics 52(2): 187-199.

Komninos, A. 2019. Pro-social behavior in crowdsourcing systems: experiences from a field deployment for beach monitoring. International Journal of Human-Computer Studies 124: 93-115.

Langeneck, J., Marcelli, M., Bariche, M., Azzurro, E. 2017. Social networks allow early detection of non indigenous species: first record of the red drum *Sciaenops ocellatus* (Actinopterygii: Perciformes: Sciaenidae) in Italian waters. Acta Adriatica 58(2): 365-370.

Marshall, H.G. Burchardt, L. 2005. Neuston: its definition with a historical review regarding its concept and community structure. Archive of Hydrobiology 164: 429-448.

Mercado, J.M., Cortés, D., Ramírez, T., Gómez, F. 2012. Decadal weakening of the wind-induced upwelling reduces the impact of nutrient pollution in the Bay of Málaga (western Mediterranean Sea). Hydrobiologia 680: 91-107.

Meyers, E.K.M., Tuya, F., Barker, J., Jiménez-Alvarado, D., Castro-Hernández, J.J., Haroun, R., Rödder, D. 2017. Population structure, distribution and habitat use of the critically endangered Angelshark, *Squatina squatina*, in the Canary Islands. Aquatic Conservation Marine and Freshwater Ecosystems May: 1-12.

Nastav, B., Malej, M., Malej, Jr., Malej, A. 2013. Is it possible to determine the economic impact of jellyfish outbreaks on fisheries? A case study – Slovenia. Mediterranean Marine Science 14(1): 214-223.

Nordstrom, B., James, M.C., Martin, K., Worm, B. 2019. Tracking jellyfish and leatherback sea turtle seasonality through citizen science observers. Marine Ecology Progress Series 620: 15-32.

Nunes, P.A.L.D, Loureiro, M.L., Piñol, L., Sastre, S., Voltaire, L., Canepa, A. 2015. Analyzing beach recreationists' preferences for the reduction of jellyfish blooms: economic results from a stated-choice experiment in Catalonia, Spain. PLoS ONE 10(6): e0126681.

Oguz, T., Macias, D., Garcia-Lafuente, J., Pascual, A., Tintore, J., 2014. Fueling plankton production by a meandering frontal jet: a case study for the Alboran Sea (Western Mediterranean). PLoS ONE 9(11): e111482.

Pires, R.F.T., Cordeiro, N., Dubert, J., Marraccini, A., Relvas, P., dos Santos, A. 2018. Untangling Velella velella (Cnidaria: Anthoathecatae) transport: a citizen science and oceanographic approach. Marine Ecology Progress Series 529: 241-251.

Rand, W.M. 1971. Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association 66: 846–850.

Sánchez-Garrido, J.C., García-Lafuente, J., Álvarez-Fanjul, E., Sotillo, M.G., de los Santos, F.J., 2013. What does cause the collapse of the Western Alboran Gyre? Results of an operational ocean model. Progress in Oceanography 116: 142-153.

Sammons, S.M. 2019. Bridging the gap between scientists and anglers: the Black Bass conservation committee's social media outreach efforts. Fisheries 44(1): 37-41.

Schimek, M.G. 2000. GAM spline algorithms: a direct comparison. In: Bethlehem J.G., van der Heijden P.G.M. (eds) COMPSTAT. Physica, Heidelberg.

Shiffman, D.S. 2018. Social media for fisheries science and management professionals: how to use it and why you should. Fisheries 43(3): 123-129.

Sidhom, S., Hassoun, M. 2002. Morpho-syntactic parsing for a text mining environment: an NP recognition model for knowledge visualization and information retrieval. Knowledge Organization 29(3): 171-180.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology and Evolution, 24(9): 467-471.

Snyder, J.T., Whitney, M.M., Dam, H.G., Jacobs, M.W., Baumann, H. 2019. Citizen science observations reveal rapid, multi-decadal ecosystem changes in eastern Long Island Sound. Marine Environmental Research 146: 80-88.

Stehno, B., Retti, G. 2003. Modelling the logical structure of books and journals using augmented transition network grammars. Journal of Documentation 59(1): 69-83.

Thorne, J. Barley, P., Dewar, H. 1968. The syntactic analysis of English by machine. In: Machine Intelligence vol. 3, New York. American Elsevier Publishing Co., Inc.: 281-310.

Tintore, J., La Violette, P.E., Blade, I., Cruzado, A. 1988. A study of an intense density front in the Eastern Alboran Sea: the Almeria-Oran front. Journal of Physical Oceanography 18(10): 1384-1397.

Tomlinson, B., Maynou, T., Sabetés, A., Fuentes, V., Canepa, A., Sastre, S. 2018. Systems approach modeling of the interactive effects of fisheries, jellyfish and tourism in the Catalan coast. Estuarine, Coastal and Shelf Science 201: 198-207.

UNEP. 1983. Mediterranean action plan. Long terms programme for pollution monitoring and research in the Mediterranean (Med POL – PHASE II). Workshop on jellyfish blooms in the Mediterranean, Athens.

Woods, W.A. 1970. Transition networks grammars for natural language analysis. Communications of the Association for Computing Machinery 13(10): 591-610.

Woods, W.A. 1973. Progress in natural language understanding: an application to Lunar geology. Proceedings of the AFIPS Conference 42.

Tabla 1. Number of comments registered in Infomedusa APP by region. The orientation of the beaches, beach set and number of beaches for each set is indicated

City	Total comments	Valid comments	Beach orientation	Set	Number of beaches
Málaga	2822	707 (25.05%)	South and southeast	1 - blue	18
		Total=707			Total=18
Algarrobo	54	17 (31.48%)	South	2 - orange	2
Nerja	516	141 (27.32%)	South	2 - orange	16
Rincón de la Victoria	860	298 (34.65%)	South	2 - orange	4
Torrox	255	77 (30.19%)	South	2 - orange	10
Vélez-Málaga	1122	275 (24.51%)	South	2 - orange	7
		Total=808			Total=39
Benalmádena	635	234 (36.64%)	Southeast	3 - red	17
Fuengirola	396	153 (38.69%)	Southeast	3 - red	9
Mijas	294	102 (34.69)	Southeast	3 - red	14
Torremolinos	687	228 (33.19%)	Southeast	3 - red	4
		Total=717			Total=44
Marbella	1579	554 (35.09%)	South	4 – green	26
		Total=554			Total=26
Casares	58	20 (34.48%)	Southeast	5 - grey	2
Estepona	303	88 (29.04%)	Southeast	5 - grey	12
Manilva	382	134 (35.08%)	Southeast	5 - grey	16
		Total=242			Total=20

Figure captions

Figure 1. Study area. Locations of beach sets analysed belonging to the main Costa del Sol cities are shown (see table 1). Also, the locations of the Manilva, Marbella Cabopino, Fuengirola and Málaga meteorological stations are indicated.

Figure 2. Snapshots of the main and secondary windows of the Infomedusa app. An example of the type of messages left by the users is also shown.

Figure 3. Transition network (TN) of Infomedusa APP and schematic representation of the message: 'Right now on the arroyo de la miel beach I haven't seen a single jellyfish'.

Figure 4. GAM results for Málaga city beaches (set 1, blue). a) Message frequency distributions registered by Infomedusa from Wednesday April 4 to Tuesday July 24; b) Partial residual analysis of the GAM model considering wind direction and wind speed as model inputs; c) Location of the beaches considered in the analyses, and respective wind intensities and directions favouring (red arrows) or preventing (blue arrows) the onshore arrival of jellyfish. The wind rose corresponds to the mean daily wind direction and wind speed registered in the Málaga meteorological station from Wednesday April 4 to Tuesday July 24.

Figure 5. GAM results for Algarrobo, Nerja, Rincón de la Victoria, Torrox and Vélez-Málaga cities beaches (set 2, orange). a) Message frequency distributions registered by Infomedusa from Wednesday April 28 to Tuesday July 24; b) Partial residual analysis of the GAM model considering wind direction and wind speed as model inputs; c) Location of the beaches considered in the analyses, and respective wind intensities and directions favouring (red arrows) or preventing (blue arrows) the onshore arrival of jellyfish. The wind rose corresponds to the mean daily wind direction and wind speed registered in the Málaga meteorological station from Wednesday April 4 to Tuesday July 24.

Figure 6. GAM results for Benalmádena, Fuengirola, Mijas and Torremolinos cities beaches (set 3, red). a) Message frequency distributions registered by Infomedusa from Saturday May 12 to Wednesday July 25; b) Partial residual analysis of the GAM model considering wind direction and wind speed as model inputs; c) Location of the beaches considered in the analyses, and respective wind intensities and directions favouring (red arrows) or preventing (blue arrows) the onshore arrival of jellyfish. The wind rose corresponds to the mean daily wind direction and wind speed registered in the Fuengirola meteorological station from Saturday May 12 to Wednesday July 25.

Figure 7. GAM results for Marbella city beaches (set 4, green). a) Message frequency distributions registered by Infomedusa from Saturday June 6 to Tuesday July 24; b) Partial residual analysis of the GAM model considering wind direction and wind speed as model inputs; c) Location of the beaches considered in the analyses, and respective wind intensities and directions favouring (red arrows) or preventing (blue arrows) the onshore arrival of jellyfish. The wind rose corresponds to the mean daily wind direction and wind speed registered in the Marbella Cabopino meteorological station from Saturday June 6 to Tuesday July 24.

Figure 8. GAM results for Marbella city beaches (set 5, grey). a) Message frequency distributions registered by Infomedusa from Sunday May 20 to Tuesday July 24; b) Partial residual analysis of the GAM model considering wind direction and wind speed as model inputs; c) Location of the beaches considered in the analyses, and respective wind intensities and directions favouring (red arrows) or preventing (blue arrows) the onshore arrival of jellyfish. The wind rose corresponds to the mean daily wind direction and wind speed registered in the Manilva meteorological station from Sunday May 20 to Tuesday July 24.

Figure 9. Abundances (normalized values) though time at western (set 5) and eastern (sets 3 and 4) beaches. Red/blue arrows indicate high/low occurrence of jellyfish in eastern beaches, days later of being detected in the western beaches.

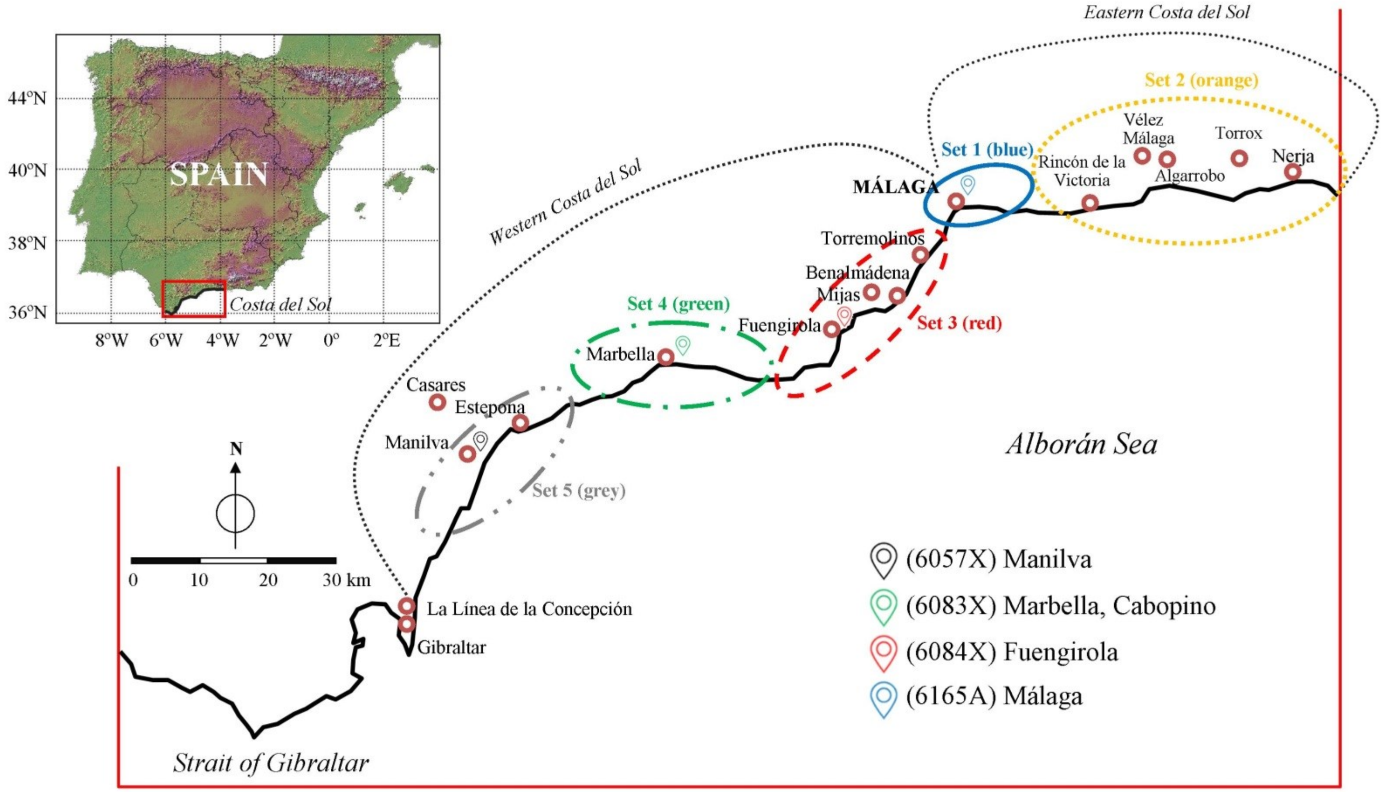

44ºN
40ºN
38ºN
36ºN
8ºW
6ºW
4ºW
2ºW
0º
2ºE
SPAIN
Costa del Sol
Eastern Costa del Sol
Set 2 (orange)
Vélez Málaga
Torrox
Set 1 (blue)
Rincón de la Victoria
Algarrobo
Nerja
MÁLAGA
Western Costa del Sol
Torremolinos
Benalmádena
Mijas
Set 3 (red)
Set 4 (green)
Fuengirola
Marbella
Casares
Estepona
Manilva
Set 5 (grey)
Alborán Sea
N
0
10
20
30 km
La Línea de la Concepción
Gibraltar
Strait of Gibraltar
(6057X) Manilva
(6083X) Marbella, Cabopino
(6084X) Fuengirola
(6165A) Málaga

FIGURE 1

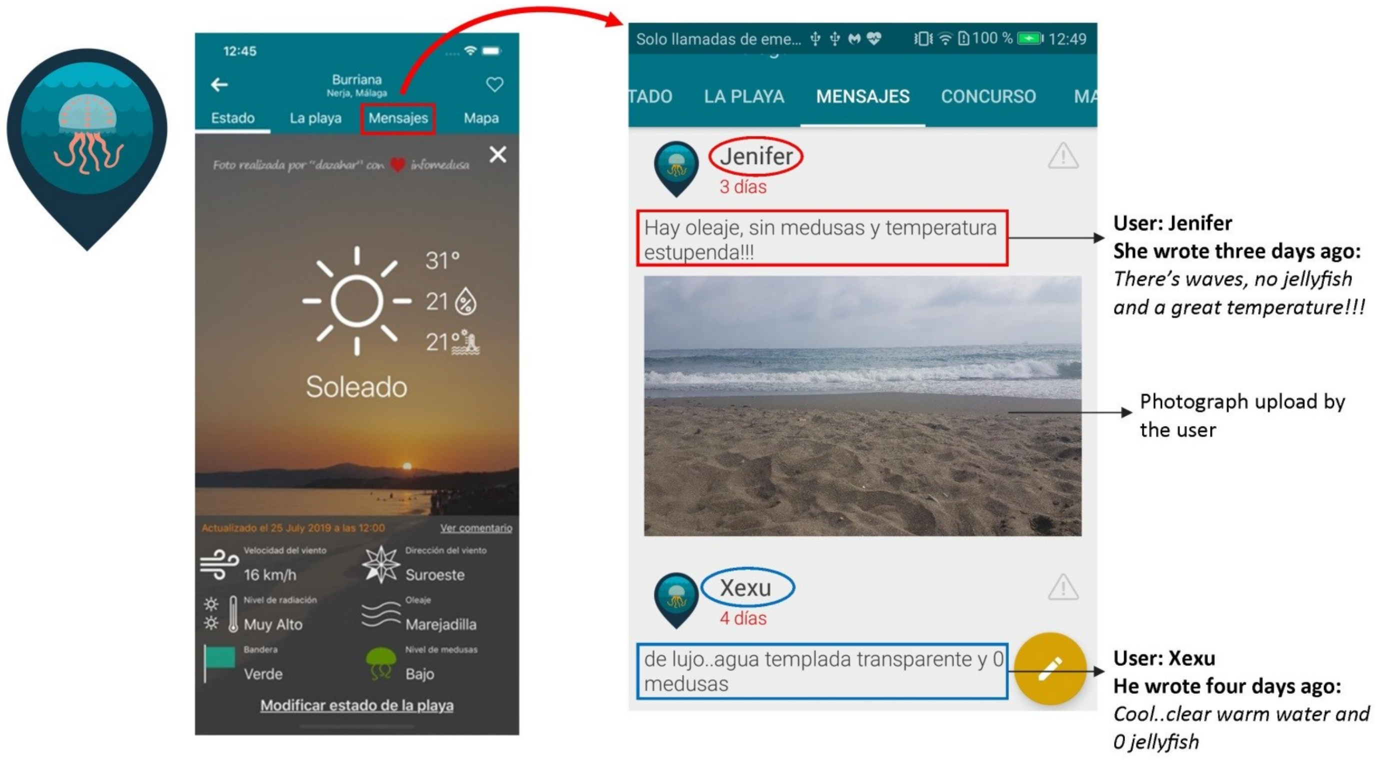
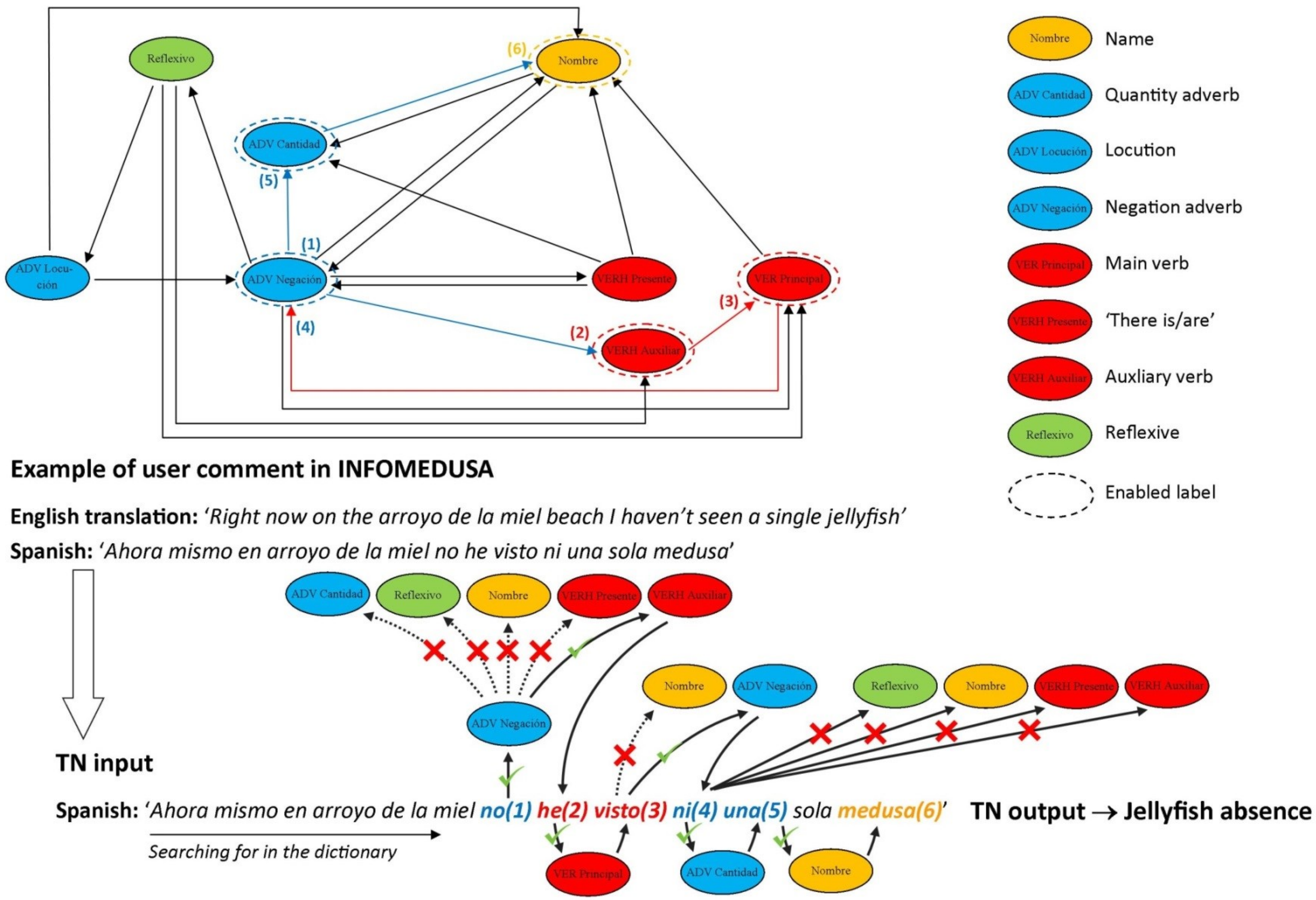
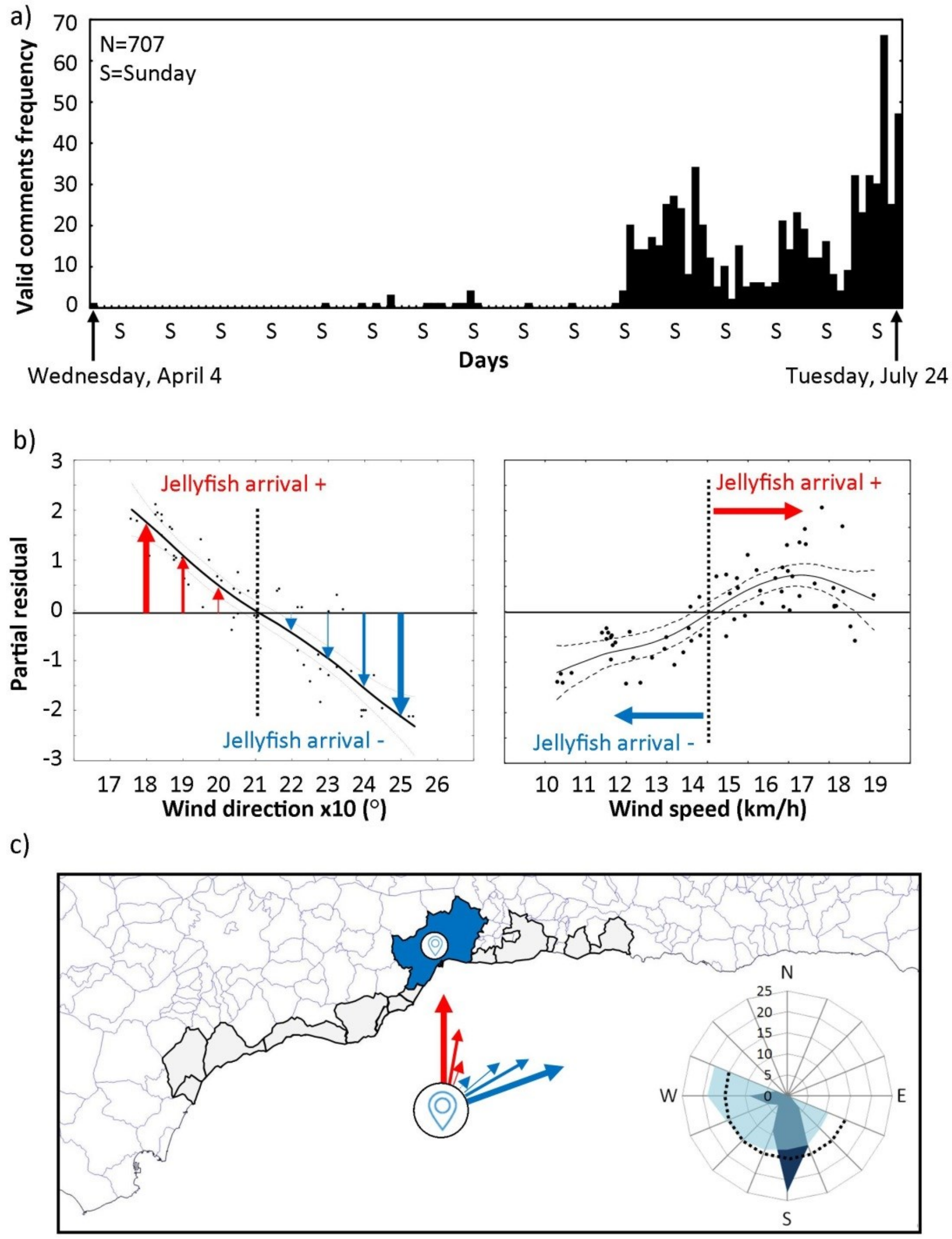

12:45
Burriana
Nerja, Málaga
Estado
La playa
Mensajes
Mapa
31°
21
21°
Soleado
16 km/h
Suroeste
Muy Alto
Marejadilla
Verde
Bajo
Modificar estado de la playa
TADO
LA PLAYA
MENSAJES
CONCURSO
Jenifer
3 días
Hay oleaje, sin medusas y temperatura estupenda!!!
Xexu
4 días
de lujo..agua templada transparente y 0 medusas
User: Jenifer
She wrote three days ago:
There's waves, no jellyfish and a great temperature!!!
Photograph upload by the user
User: Xexu
He wrote four days ago:
Cool..clear warm water and 0 jellyfish

FIGURE 2



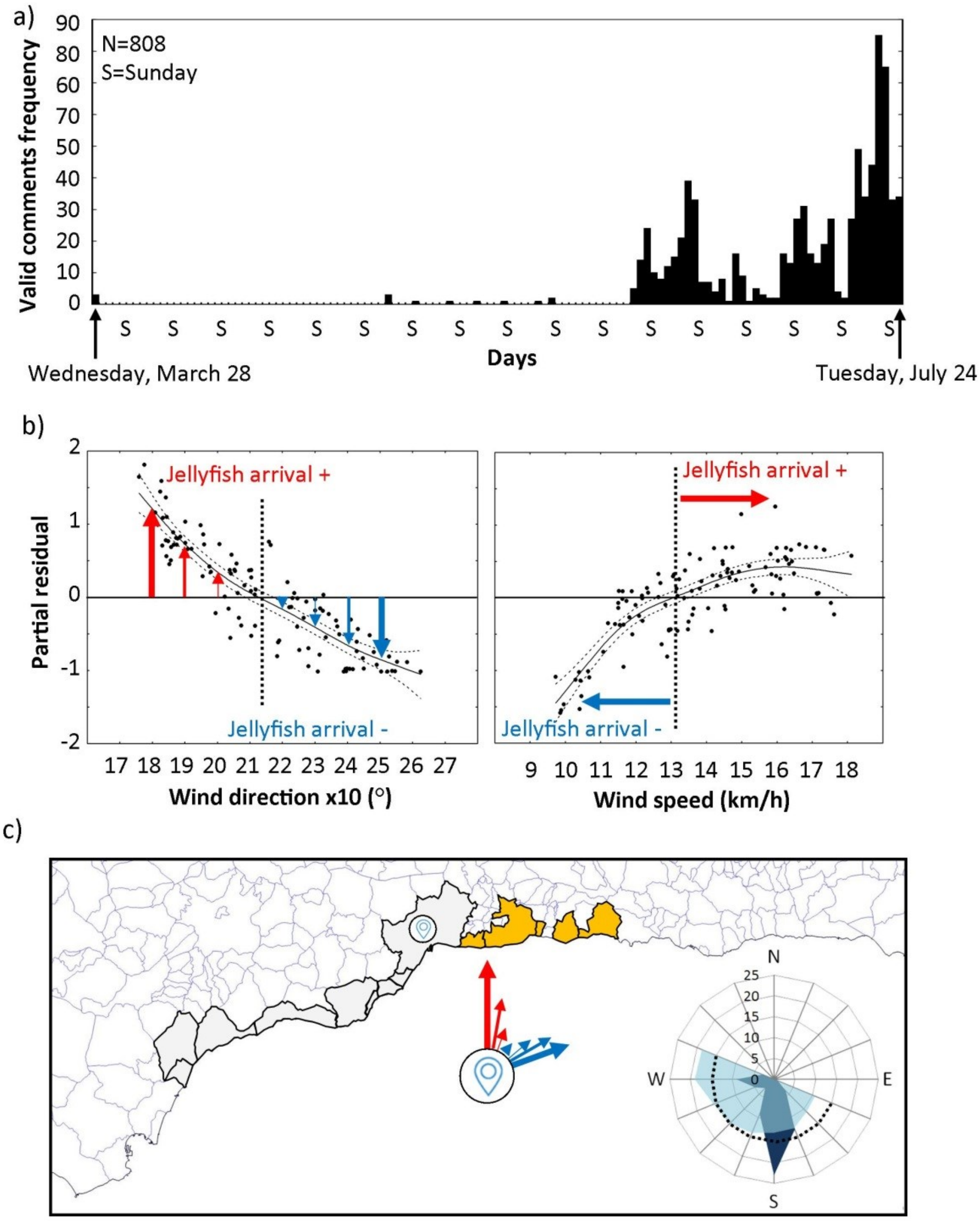
Reflexivo
Nombre
ADV Cantidad
ADV Locución
ADV Negación
VERH Presente
VER Principal
VERH Auxiliar
(1)
(2)
(3)
(4)
(5)
(6)
Name
Quantity adverb
Locution
Negation adverb
Main verb
'There is/are'
Auxliary verb
Reflexive
Enabled label
Example of user comment in INFOMEDUSA
English translation: 'Right now on the arroyo de la miel beach I haven't seen a single jellyfish'
Spanish: 'Ahora mismo en arroyo de la miel no he visto ni una sola medusa'
TN input
Spanish: 'Ahora mismo en arroyo de la miel no(1) he(2) visto(3) ni(4) una(5) sola medusa(6)'
Searching for in the dictionary
TN output → Jellyfish absence

FIGURE 3



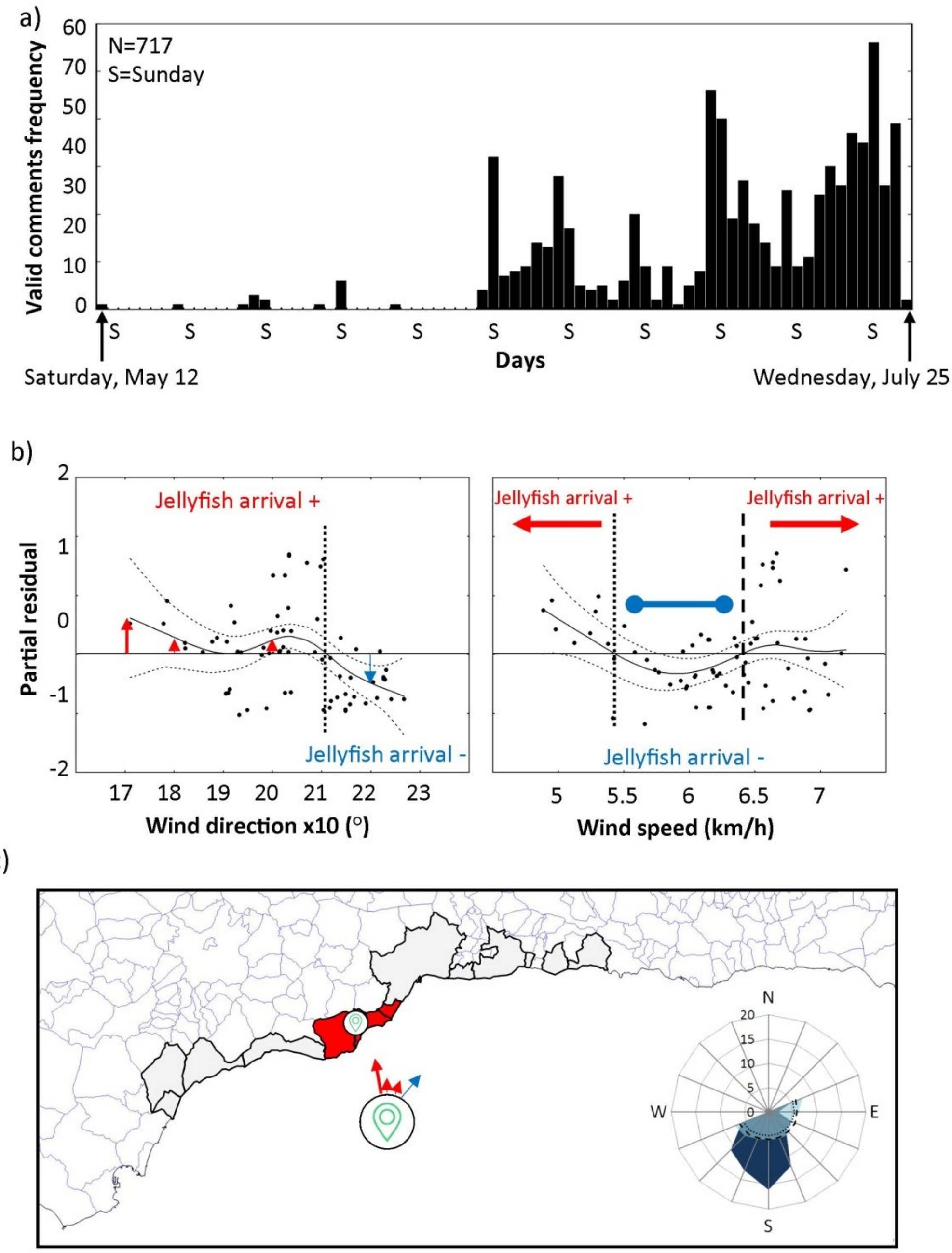
a)
N=707
S=Sunday
Valid comments frequency
70
60
50
40
30
20
10
0
S S S S S S S S S S S S S S S S
Days
Wednesday, April 4
Tuesday, July 24
b)
Jellyfish arrival +
Jellyfish arrival -
Partial residual
3
2
1
0
-1
-2
-3
17 18 19 20 21 22 23 24 25 26
Wind direction x10 (°)
Jellyfish arrival +
Jellyfish arrival -
10 11 12 13 14 15 16 17 18 19
Wind speed (km/h)
c)
N
W
E
S
25
20
15
10
5
0

FIGURE 4



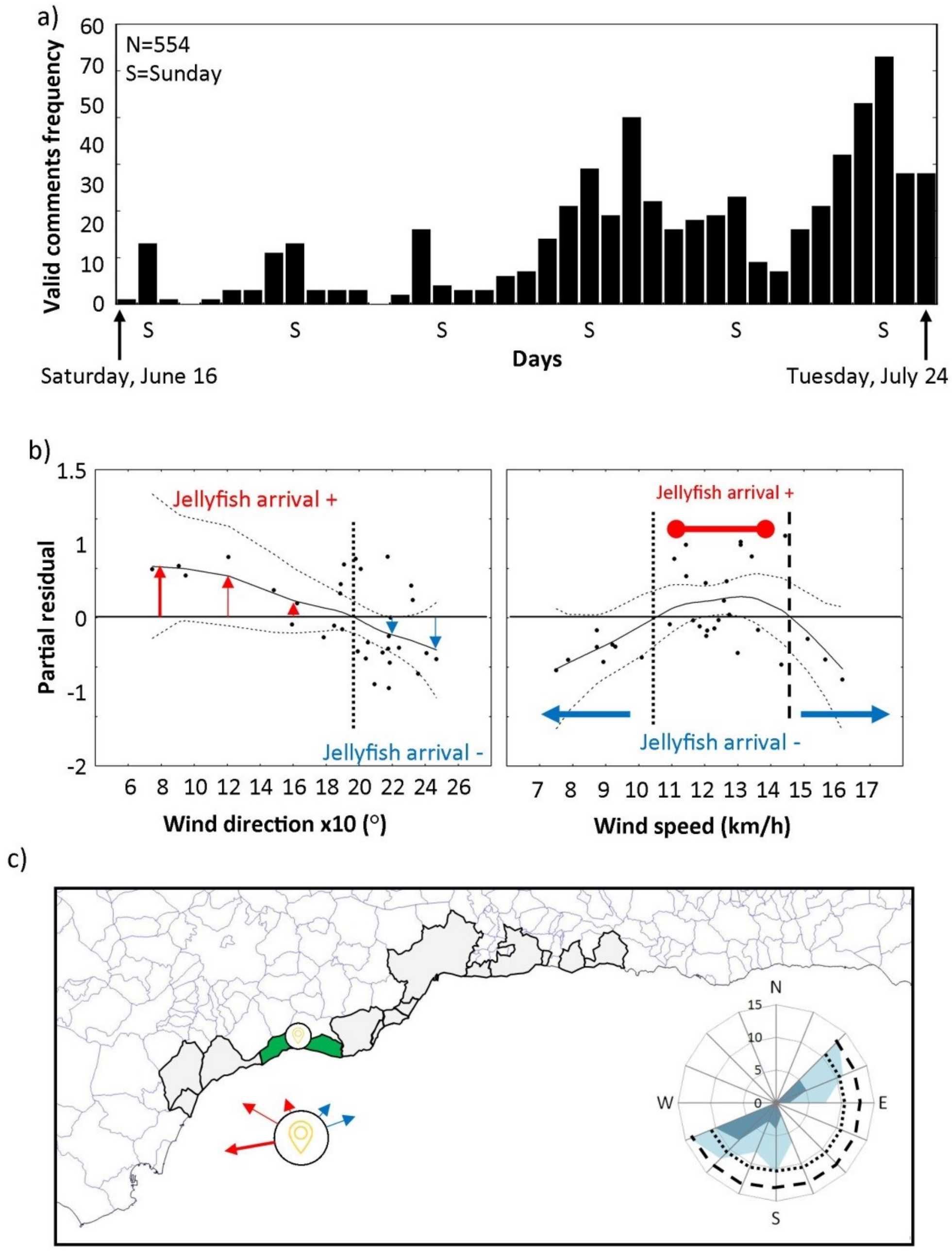
a)
N=808
S=Sunday
Valid comments frequency
Days
Wednesday, March 28
Tuesday, July 24
b)
Jellyfish arrival +
Jellyfish arrival -
Partial residual
Wind direction x10 (°)
Wind speed (km/h)
c)
N
W
E
S

FIGURE 5



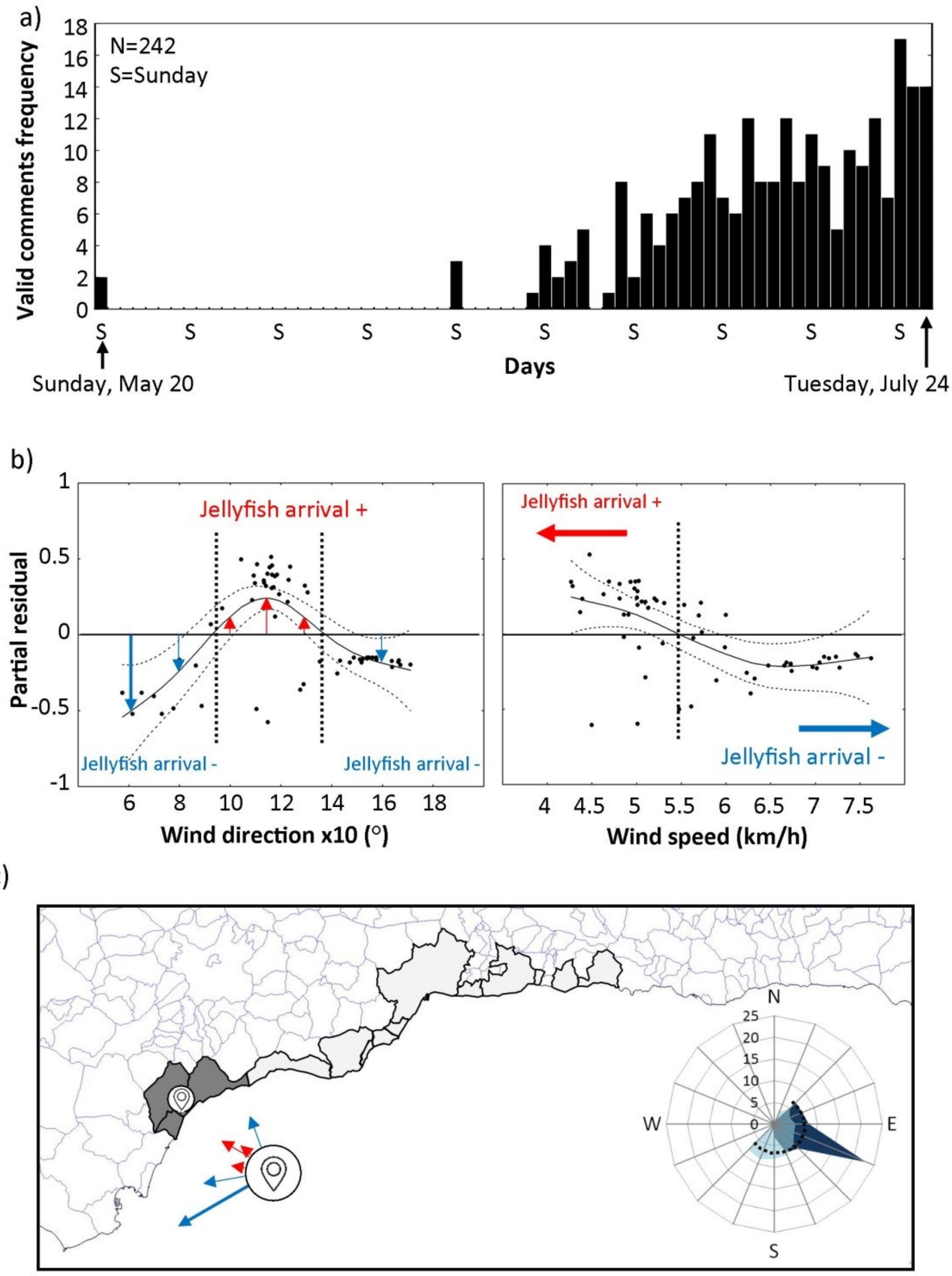
a)
N=717
S=Sunday
Valid comments frequency
Days
Saturday, May 12
Wednesday, July 25
b)
Partial residual
Jellyfish arrival +
Jellyfish arrival -
Wind direction x10 (°)
Wind speed (km/h)
c)
N
S
E
W

FIGURE 6



a)
Valid comments frequency
60
70
50
40
30
20
10
0
N=554
S=Sunday
S
S
S
S
S
S
Days
Saturday, June 16
Tuesday, July 24
b)
Partial residual
1.5
1
0
-1
-2
Jellyfish arrival +
Jellyfish arrival -
6 8 10 12 14 16 18 20 22 24 26
Wind direction x10 (°)
Jellyfish arrival +
Jellyfish arrival -
7 8 9 10 11 12 13 14 15 16 17
Wind speed (km/h)
c)
N
W
E
S
15
10
5
0

FIGURE 7



a)
N=242
S=Sunday
Valid comments frequency
Days
Sunday, May 20
Tuesday, July 24
b)
Jellyfish arrival +
Jellyfish arrival -
Jellyfish arrival -
Partial residual
Wind direction x10 (°)
Jellyfish arrival +
Jellyfish arrival -
Wind speed (km/h)
c)
N
E
S
W

FIGURE 8

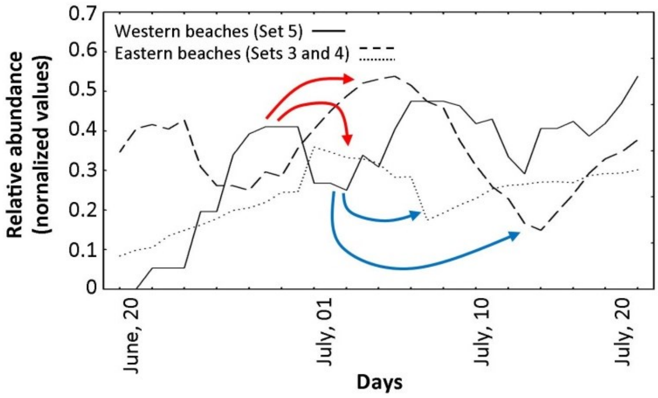

Western beaches (Set 5)
Eastern beaches (Sets 3 and 4)
Relative abundance (normalized values)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
June, 20
July, 01
July, 10
July, 20
Days

FIGURE 9