THE OFFICIAL MAGAZINE OF THE OCEANOGRAPHY SOCIETY

Oceanography

CITATION

Otero, P., M. Ruiz-Villarreal, G. González-Nuevo, and J.M. Cabanas. 2014. Operational oceanography, end users, and social network sites: An exploratory analysis. *Oceanography* 27(3):138–147, http://dx.doi.org/10.5670/oceanog.2014.63.

DOI

http://dx.doi.org/10.5670/oceanog.2014.63

Operational Oceanography, End Users, and Social Network Sites

An Exploratory Analysis

BY PABLO OTERO,
MANUEL RUIZ-VILLARREAL,
GONZALO GONZÁLEZ-NUEVO,
AND JOSE MANUEL CABANAS

ABSTRACT. Many oceanographic products are currently being disseminated in a systematic and routine manner to end users. In recent years, data producers have gained insight into the specific requirements of the scientific community. However, there is still a lack of perception of the interests of the broader and non-expert public. This study analyzes the interests and needs of potential end users of operational oceanography by mining Web search engine and social media data. Results show an increasing number of people searching for operational oceanography-related products, with seasonality in these searches depending on the kind of variable. Information on currents is searched more during winter, waves during spring, and tides and temperature during summer. Moreover, the ranking of specific interests of the general public differs from the requirements of the fisheries and applied environmental scientists reported by the International Council for the Exploration of the Sea Working Group on Operational Oceanographic Products for Fisheries and Environment. The general public is more interested in temperature, wave conditions, and sea ice, whereas the highest priority of a group of scientists was temperature, currents, and salinity. An understanding of the terminology used by non-expert clients and their priorities will help institutions involved in curating and disseminating oceanographic data sets to better design their Web portals and applications.

INTRODUCTION

Over the past decade, the scientific community has put great effort into providing end users with reliable oceanographic data and products in a systematic and routine manner. This activity, known as operational oceanography, is mainly supported by measurements from research ships and from various instruments such as buoys, platforms, gliders, and autonomous underwater vehicles (AUVs), as well as satellite and model-derived data. A measure of success is the number of end users of these data and products, which is highly dependent on their rapid interpretation and dissemination. From a technical point of view, the current state of the art is the result of a combination of the fast growth of computing power used by numerical forecasting models in an increasingly demanding way, and of the development and expansion of the networks that allow transmission of increasing amounts of data at faster speeds from instruments and platforms to data assimilation centers, onward to intermediary value-adding organizations, and, finally, to end users. From an organizational point of view, the driving force for the development of the operational oceanography capability has been the ability to integrate data from dispersed coastal observatories into larger international frameworks, including the Global Earth Observation System of Systems (GEOSS), through its oceanographic component Global Ocean Observing System (GOOS) and Copernicus (http://copernicus.eu), previously known as GMES, Global Monitoring for Environment and Security. See Ruhl et al. (2011) for an extensive list of national and international projects and programs with topics that relate to ocean observatory development.

As a result of all these developments, anyone with an Internet connection has access to many physical oceanographic variables, such as sea level, temperature, salinity, and currents. Advances in biogeochemical modeling and the incorporation of specific sensors into observatories have led to increasing availability of chemical and biological variables. Consequently, many oceanographic products are accessible, including ecological quantities, coastal flood warnings, maps of pollutant dispersion, warnings of harmful algal blooms (HABs), and surface drift predictions.

In order to balance end-user demand and producer effort, the environmental products released should be based on dialogues between data producers and users (Polfeldt, 2006). With this aim, Berx et al. (2011) investigated the demand for operational oceanography products coming from the fisheries and environmental scientific community, one of the perceived user groups for such products. By using the Working Group on Operational Oceanographic products for Fisheries and Environments (WGOOFE) of the International Council for the Exploration of the Sea (ICES) as a platform, they issued a questionnaire to this user community. As a result, the top five data products on the scientists' wish list were temperature, salinity, currents, primary productivity, and algal blooms, preferably including hindcast (historic) data, aggregated in monthly or annual time scales, and with spatial resolution of 10 km and greater. The authors concluded that the real-time and high-resolution data that providers had been delivering do not satisfy scientific needs. Therefore, the perception of data providers should be realigned. But, can this conclusion be extrapolated to other groups of end users? Can a wider and non-expert public take advantage of these operational products and, if so, which products do the general public want?

To gain insight into end-user needs, the Iberian Margin Ocean Observatory (http://www.marnaraia.org), a coastal modeling and observational infrastructure along the Atlantic margin of the Euro-region Galicia (Spain)-Northern Portugal (see a description in Otero et al., 2011), organized a series of meetings with different potential users of its products: the shellfish and aquaculture sector (Vigo, Spain, June 2012), final producers of Web and mobile applications interested in meteorological and oceanographic data (Santiago de Compostela, Spain, October 2012), and the tourism industry as well as groups related to nautical activities (Vigo, Spain, June 2013). From this initiative, data producers confirmed that requirements differed slightly among the various user groups. The shellfish and aquaculture sectors were interested mainly in daily temperature, salinity, and HAB warnings, with a temporal horizon of about one week. While the state-of-the-art product encompasses temperature and salinity horizons typically from three to five days, there are currently no reliable HAB predictions. Sailors require hourly wind and sea surface current vector maps with high spatial resolution. Some surfers prefer the product to be a simple summary of wave and wind conditions so that they can quickly check them from their mobile devices, with detailed information available upon request. As a final example of the diversity of needs, scuba divers are interested in near-real-time turbidity and bottom current conditions. It seems that our existing data systems do not always meet the specific needs of these sectors. Unfortunately, this beneficial dialogue among producers and users was only established with a small number of individuals, preventing any kind of quantification and extrapolation of the conclusions.

The present study attempts to analyze the interests and needs of the potential end users of operational oceanography from a broader viewpoint, by mining social media with Google Trends (http://www.google.com/trends), Facebook's Advertiser Tools (https://www.facebook.com/advertising), and the Twitter Streaming application programming interfaces (APIs; https://dev.twitter.com/docs/streaming-apis). In this way, we can learn about what people search for on the Web, how they do it, what their interests and hobbies are, and what they are talking about. These results will be compared to the ranking of operational data products required by environmental scientists to provide a different and helpful point of view to data providers.

EXPLORING THE WEB

Search engines are the main gateways for accessing information from the Internet. Among the various search engines, Google is the most popular, with a market share during the first half of 2013 of 70.98% in desktop platforms and 90.97% in mobile devices (data from Market Share Statistics for Internet Technologies, http://www.netmarketshare.com).

Google Trends is a public Web facility of Google Inc. that provides a real-time daily and weekly time series index of the volume of queries that users enter into Google Search in a given geographic area and in a particular time period (the data go back to January 1, 2004). The query index is the total query volume for the search term in question divided by the total number of queries in that region during the time period being examined. The maximum query share in the time period specified is normalized to 100. The system also eliminates repeated queries from a single user over a short time period so that the level of interest is not artificially impacted. It is important to note that Google Trends data are computed using only a small sample of all information stored in their servers; hence, the results for the same query may vary a little from day to day.

In recent years, several studies have proven the usefulness of Google Trends to many subjects, including surveillance, prevention, and control of emerging diseases (Brownstein et al., 2009; Ginsberg et al., 2009; Hulth et al., 2009); research related to financial market fluctuations (Preis et al., 2010); interest in developing countries (Preis et al., 2012); unemployment forecasting (Askitas and Zimmermann, 2009; Baker and Fradkin, 2011); and prediction of consumer behavior (Goel et al., 2010). In this section, we apply this emerging tool to investigate the Web search share of terms related to operational oceanography. Understanding the how users query Google will allow producers of operational products to better meet the needs of their target public.

Google Trends designates a certain threshold of traffic for search terms so that those with low volumes will not appear. This happens, for example, for those searches simultaneously including the terms "operational" and "oceanography." It is expected that these general query terms are restricted to expert or academic spheres and, hence, their volume is too small to meet the threshold. That the threshold is not met does not imply a lack of the general public's interest in near-real-time and forecast conditions in the coastal sea and, consequently, in operational oceanographic products.

Figure 1 compares the share of queries that combine the following terms: "ocean," "sea," and "marine" with "weather" or "forecast." The figure also shows the query share combining the Spanish terms "mar" (sea) and "tiempo" (weather), more searched than other combinations using the Spanish terms "océano" (ocean), "predicción," and "previsión" (forecast), something that happened independent of the diacritical marks. We have excluded searches that reference the weather of large cities (> 25,000 inhabitants) whose names include "mar," such as Mar del Plata (Argentina), Viña del Mar (Chile), and Premiá del Mar and Roquetas del Mar (Spain).

In general terms, the interest increases or remains over time, with the exception of those searches combining "marine" and "weather." However, a decreasing share trend does not necessarily mean fewer queries because the number of Internet users has grown dramatically in recent years.

Worldwide queries have a seasonal trend, with primary peaks in July/August, secondary peaks in December/January, and minima in February/March. The existence of two seasonal peaks could be explained by the fact that the Northern Hemisphere search volume is greater

Pablo Otero *(otero_pablo@hotmail.com) is Research Scientist and* **Manuel Ruiz-Villarreal** *is Senior Scientist, both at the Instituto Español de Oceanografía, Centro Oceanográfico de A Coruña, A Coruña, Spain.* **Gonzalo González-Nuevo** *is Research Scientist and* **Jose Manuel Cabanas** *is Senior Scientist, both at the Instituto Español de Oceanografía, Centro Oceanográfico de Vigo, Vigo, Spain.*

than that of the Southern Hemisphere during their respective summers. In general terms, "weather" is used more often in searches than "forecast." Moreover, the term "marine" is used more than "sea" and "ocean." However, the way in which people search for this kind of information strongly differs among countries and in time. For example, "marine weather" and "marine forecast" are the main ways of searching for this kind of information in New Zealand, and this occurs mainly during the austral summer. In fact, New Zealand has the highest query share of these terms and, therefore, we assume that interest in this kind of information is relatively high in New Zealand compared to other countries. In contrast, Internet users in the UK (in this particular case, secondary seasonal peaks are also observed during winter) and in Ireland more often make queries using the term "sea," whereas users in Australia prefer the term "ocean." In the United States, currently "ocean weather" is the preferred option, followed by "marine forecast." However, prior to 2007, the main term US users employed for their search was "marine weather," revealing that preferences in terminology also change in time. The Spanish terms have been included to show the rising trend in recent years, especially in Chile and Argentina from 2011 on, although Spain is still the country with the highest relative interest among the Spanish-speaking community.

Figure 2 compares the share of various queries with some classical physical oceanographic parameters: temperature, waves, tides, ice, currents, and salinity, listed here in descending order of popularity. All plotted values are normalized by the maximum of searches during the period, which corresponds with queries about waves in January 2005. The numbers of queries about temperature, tides, currents, and waves show strong seasonality. Whereas searches about temperature and tides peak during boreal summer, particularly during July, searches about currents are at their lowest levels during these months. The minimum query share concerning temperature and tides occurs in December. Information about waves is more searched during April and May, and registers a minimum in the volume of queries in December. Salinity is of minimum relative interest among the compared variables, with minimum queries during July, and no clear seasonal pattern of peaks. News about melting of the Arctic sea ice reported at the end of summer 2012, when the summer sea ice extent had reached an all-time minimum, significantly increased Google queries about sea ice (e.g., see the National Snow and Ice Data Center's press release at https://nsidc.org/news/press/20121002_MinimumPR.html). The query share about sea level (not shown in the figure) was, on average, double that of temperature. Interest in sea level is not seasonal, and peaks may be mainly related to news about global warming and climate change. Other terms evaluated, for example, zooplankton, algal bloom, chlorophyll, primary productivity, sea fronts, and turbidity, had fewer total searches compared to the variables discussed previously. Note that Figure 2 does not include searches in languages other than English. However, searches using Spanish or Portuguese are increasing in accordance with the trend pointed out for Figure 1.

The Figure 3 map identifies the main cities around the world whose

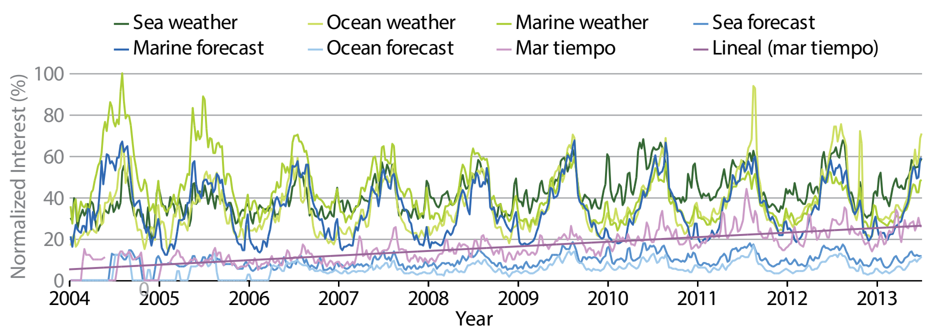

Figure 1. Comparison of the share of different Google queries from January 1, 2004 to July 1, 2013. Numbers represent search interest relative to the highest point on the chart. This doesn't convey absolute search volume. Blue colors include the term "forecast" and green colors the term "weather." The linear trend (r^2 = 0.63) in the case of the Spanish terms is also shown (purple line).

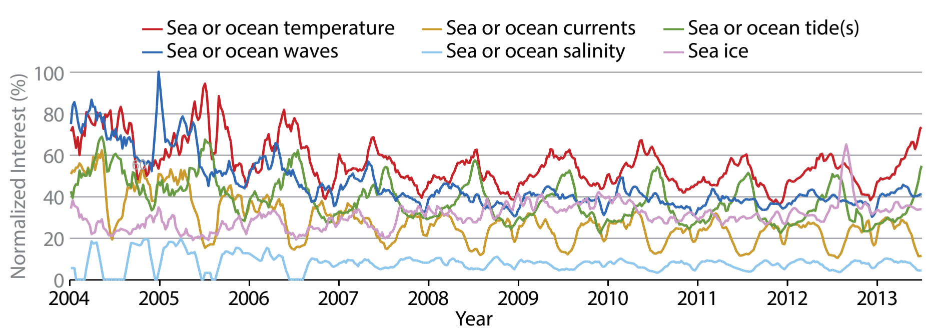

Figure 2. Comparison of the share of different Google queries concerning some physical oceanographic variables done from January 1, 2004 to July 1, 2013.

populations conduct a larger volume of searches on the ocean-related topics discussed above with respect to the total volume of searches in those cities. Again, it is important to note that Google does not provide the absolute number of searches; thus, the map only shows those cities with a large relative interest in these topics. It indicates query hotspots, including those conducted in Spanish, English, and Portuguese, all in the top-six listing of languages by number of native speakers (Ethnologue, 2013); we excluded Mandarin, Hindi, and Arabic from our study. Queries in English are the most popular, followed by searches in Spanish, which are, on average, 3.5 times greater than those in Portuguese.

EXPLORING SPECIFIC INTERESTS OF END USERS

In the previous section, we showed the search terms that people use, although this information does not necessarily reflect their real interests or hobbies. Being able to gather this information would help to more effectively target the public and increase the use of operational oceanography information. A large fraction of the public uses Facebook, a social network created in 2004 that had grown to more than 1,100 million users as of September 2013, exceeding Google's number of site visits. Four billion pieces of content are shared each day on Facebook, and it is now integrated into over 10 million websites and applications (Facebook Inc., 2012). Facebook's astounding ubiquity has generated a rising number of social research efforts to assess its impact on social life (Wilson et al., 2012).

Facebook Advertiser (https://www.facebook.com/advertising) is a powerful tool usually used to obtain quick upper bounds on market sizes. Segmentation among precise interests is possible thanks to the terms that users employ when sharing content through this social network. As explained by the Facebook team, interests are determined by what people are connected to on Facebook, such as pages and apps. Facebook provides the terminology used to make reference to a certain interest and directly translates it into several languages.

Hence, the comparison among countries is straightforward. In this study, this tool is used to mine the number of Facebook users who demonstrated interest in specific variables concerning the sea or the ocean. Table 1 shows the data extracted for the whole Facebook network, restricted to a reduced list of countries that were selected for different reasons. The United States, Canada, UK, Ireland, New Zealand, and Australia are included for their high query shares about operational oceanography-related terms in the Google search engine, Spain for being the main representative of the rising interest among the Spanish-speaking countries (see Exploring the Web section), and, finally, France for its long tradition in oceanography-related studies.

Because Facebook infers interest from the shared content in the social net, these absolute values are not fully representative of the users' concerns. However, the potential of these data is that we can sort specific interests in order of relevance. Considering the whole community, more than 3.5% of users demonstrated an interest in the sea and more than 1.2% in the ocean, although there were differences among countries. Among the list of oceans, the Pacific Ocean holds the greatest interest among the global Facebook community. If we remove New Zealand and Australia from the analysis, the Atlantic Ocean is of greatest interest to all the remaining countries. Table 1 also shows the interest about specific zones of the coastal ocean, from the generic term "coast" to the less popular and more specific "intertidal zone." It is worth mentioning the high interest in all zones among the Australian Facebook users compared with the others nationalities. Concerning specific oceanographic variables, about 0.5% of Facebook users

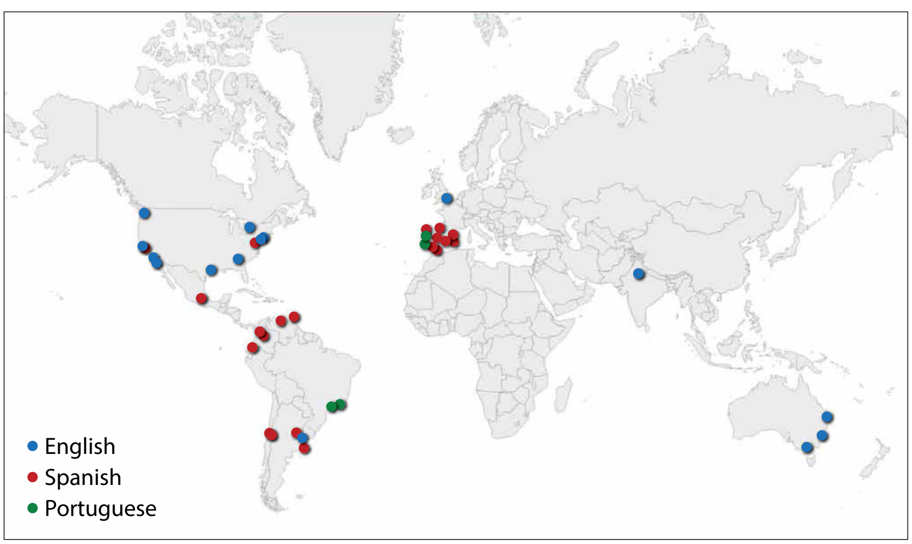

Figure 3. Cities with a larger Google query share concerning physical oceanographic variables (updated September 2013), such as temperature, salinity, currents, waves, and sea level. Colors indicate the language used during the search: English (blue), Spanish (red), and Portuguese (green).

have demonstrated some interest in tides, a value that rises to ~ 2% in the case of US users. The next variable in this list of interests is "sea level," followed by "sea currents," "sea surface temperature," "salinity," "algal blooms," and "red tides" (the last term is most used by the general public in Spain). Interest in the terms "eddies" and "oceanic fronts" is too low to include in the table.

TWITTER ANALYSIS

Twitter Inc. provides an online social networking and microblogging service that allows users to send and read comments (tweets) about any topic within a 140-character limit. Launched in July 2006, Twitter rapidly gained worldwide popularity. In March 2013, more than 200 million active users created over 400 million tweets each day (reported from the Twitter blog at https://blog.twitter.com/2013/celebrating-twitter7). Twitter users can read the latest news, exchange ideas, and share interests. Moreover, the simplicity of the system allows posting of comments by robot-like devices, opening the possibility of sending automated messages in near-real time that contain data collected by, for example, a coastal buoy. Thus, we anticipate that operational oceanography will benefit from using this social network.

The set of streaming APIs offered by Twitter gives developers low latency access to Twitter's global stream of tweet data, which includes the tweet text along with metadata, such as time, geographical coordinates associated with the tweet (if GPS is enabled), and user information from the user profile (e.g., the user's real name). This information has been used by geoscience researchers to detect earthquakes (Sakaki et al., 2010) and by public health officials for surveillance of influenza outbreaks (e.g., Signorini et al., 2011) and drug (Paul and Dredze, 2013) and tobacco use (Myslín et al., 2013).

Table 1. Precise interests (‰) of Facebook users (updated on September 30, 2013) in different countries. The total population that is at least 13 years old (the minimum age to create an account under current Facebook rules) was obtained from the Health Nutrition and Population Statistics of the World Bank (http://databank.worldbank.org) for 2012. No data are shown when the number of people with a specific interest is below 1,000. Note that Facebook's users are shown in percent (%), whereas precise interests are per thousand (‰).

		Facebook Community	USA	Canada	UK	Ireland	New Zealand	Australia	France	Spain
POPULATION (*thousands)		1,110,000*	256,588*	29,953*	53,562*	3,179*	3,049*	18,964*	55,254*	40,024*
FACEBOOK USERS (%)			70	63	67	75	79	67	51	47
TERM	Sea	35.60	12.22	12.63	12.22	13.33	12.50	15.63	32.86	52.63
	Ocean	12.10	17.78	12.63	11.67	13.33	17.50	25.00	18.57	6.21
OCEAN	Pacific	7.50	7.11	5.58	1.83	3.75	39.17	12.50	2.29	2.63
	Atlantic	5.60	8.67	9.68	3.94	13.33	4.33	3.44	7.86	5.89
	Indian	1.10	0.22	0.17	0.38	2.17	0.56	1.41	1.00	0.37
	Arctic	0.51	0.11	0.37	0.03	–	–	–	0.04	0.93
	Antarctic	0.14	0.06	–	0.05	–	0.55	1.72	–	–
ZONE	Coast	6.00	4.89	5.89	3.44	6.67	9.17	11.09	2.93	6.95
	Continental shelf	0.31	0.21	0.10	0.19	–	–	0.55	–	0.67
	Estuary	0.12	0.11	0.27	0.19	1.08	0.70	0.70	–	–
	Intertidal zone	0.04	0.01	–	0.12	–	–	1.14	0.13	–
VARIABLES AND FEATURES	Tide	4.90	17.78	4.95	3.67	2.92	0.55	3.28	1.93	4.00
	Sea level	0.47	0.17	0.19	0.21	1.25	–	0.10	0.14	0.08
	Sea ice	0.37	0.04	0.03	0.04	–	–	–	0.66	0.07
	Ocean current	0.21	0.11	0.06	0.07	–	–	–	0.12	–
	Sea surface temperature	0.17	0.11	–	–	–	–	–	–	–
	Salinity	0.12	0.08	–	0.03	–	–	0.08	0.04	0.05
	Algal bloom / Red tide	0.09 / 0.08	0.29 / 0.12	– / –	– / –	– /–	0.81 / –	0.17 / –	– / –	– / 0.17

The API currently limits access to public tweets to a random 1% sample and to those posted during the previous week. For this study's three-month period (September 26 to December 26, 2013), we downloaded tweets that referenced some of the oceanographic variables analyzed in previous sections (terms and hashtags occurring in the messages, in English or Spanish). Search terms were refined by inspection of the downloaded tweets to ensure that they satisfied our constraints. For example, in case of the variable "temperature," tweets should also include in their body texts the term "sea," "ocean," or "marine." For tweets written in Spanish, those with references to cities with the word "mar" (sea) in their names were excluded, as explained earlier. When searching tweets with reference to sea state and waves, we excluded those with the term "tsunami" as well as the ones automatically posted by ocean buoys, such as those from the Caribbean Coastal Ocean Observing System (http://www.caricoos.org) or others posted by users via automated services.

We downloaded and analyzed a total of 22,439 tweets. Sea ice was the most cited topic of interest (35.3%), followed by waves (17.7%), temperature (15%), currents (14%), tides (13.7%), salinity (2.4%), and algal blooms (1.9%). The presence of geo-tagged tweets (when the user connects the GPS in the mobile device) would offer valuable information that could be compared to that shown in Figure 3. However, geo-tagged tweets represented only 1.1% of our entire sample, severely limiting the potential for large-scale analyses. The low number of geo-tagged tweets in our study is even lower than values of 2% of total tweets and 3% of Twitter users previously reported by Burton et al. (2012).

DISCUSSION

Operational oceanography has developed rapidly in the last decade, and an impressive number of real-time observations and amount of forecast data are now available through many Web services. Reaching the current state required many technological advances and strong computational efforts, but there is still a gap for users between their needs and their ability to obtain and manipulate the information (Berx et al., 2011). This situation exists despite the use of geospatial standards and the implementation of Web servers specifically designed to provide and serve scientific data sets (e.g., THREDDS data server; http://www.unidata.ucar.edu/software/thredds/current/tds). At this stage and with the aim of advertising the potential of the available data sets, the design of specific data products oriented to end users' needs could be highly valuable. This design should be properly targeted, on the one hand focusing on the scientific community, and on the other hand focusing on the general public with specific interests in the coastal sea—something that varies among countries and even among seasons (see Figures 1 and 2). Initial market research and proper design would increase the visibility of a website or a Web page in a search engine as a first stage toward attracting users. It is in this context that we see the crucial importance of the terminology discussed in the earlier section, Exploring the Web.

Understanding how non-expert users search the Web is a first step, but data producers should also be aware of the specific requirements of their target public. Figure 4 ranks some of the oceanographic variables required by fisheries and applied environmental scientists, according to an ICES survey (Berx et al., 2011). It is worth noting that this survey focused on a specific community of research scientists (those interested in the ecosystem approach to management of the marine environment), and therefore should not be considered a complete review of the user community of operational oceanographic products. Data have been normalized by

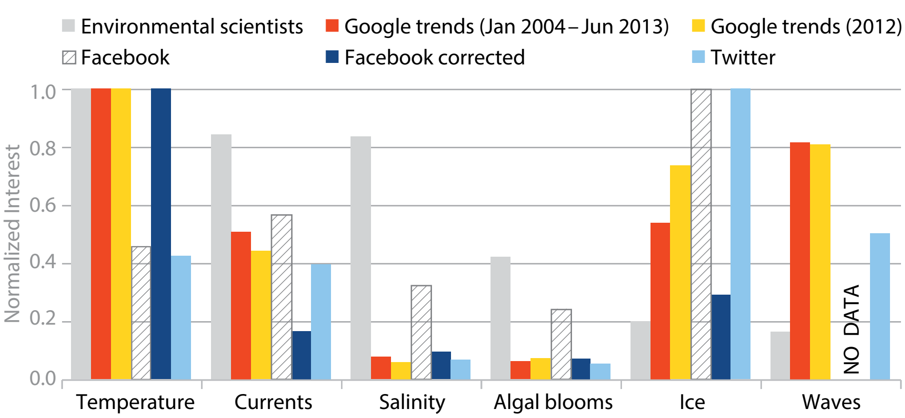

Figure 4. Rankings showing the variables required by fisheries and applied environmental scientists (data obtained from Berx et al., 2011; gray bars) compared to Google search queries from January 2004 to June 2013 (red), Google queries during 2012 (yellow), the specific interests of Facebook users (hatched), those interests corrected to account for the specific classification used by Facebook as explained in the text (dark blue), and related comments in Twitter (light blue). The classification of interests done by the Facebook team was too general in the case of waves, preventing inclusion of data concerning this variable. Data have been normalized by their corresponding maximum categories.

temperature, the variable most needed by scientists. The figure also shows the normalized query share for similar variables in the Google search engine during two periods: (1) from January 2004 to June 2013, and (2) the last complete year before this study was carried out (2012). In both cases, the query share has also been scaled by temperature. The specific interests obtained from the biographies of the Facebook users, scaled here by the proportion of users with specific interest in "sea ice," are also shown (hatched bars). This is the same variable used to scale relative importance from the whole of downloaded tweets. To facilitate the comparison, only those variables from the ICES survey that were unequivocally identified in the social networks have been included. For example, "primary productivity" occupies fourth place in the list of data products required by scientists, but the terminology that references this variable is too general because it includes terrestrial systems and also many biological aspects.

As expected, the sorted list of variables required by environmental scientists differs from the ranking of variables that attract the interest of the general public searching the Web. Scientists are interested in acquiring knowledge and using it to provide ecosystem advice and assessment, whereas those in the general public mostly gather information about sea conditions for their leisure activities, follow news about the sea or ocean, or comment and share their interests and concerns by way of social networks. In spite of the differences, "temperature" is a common starting point in this ranking among all kinds of users, with the exception of those using Facebook. Facebook classifies this specific interest as "sea surface temperature," whereas Google queries and tweets use the term "temperature" with no reference to depth. In an attempt to correct this mismatch, the queries in Google that used the term "temperature" have been checked. Queries that do not specify water depth are 7.5 times larger than those that make reference to the surface. Taking this ratio into account, we can correct the specific interest of the Facebook users with respect to this variable and generate a new classification (dark blue bars). We realize that this proportion may be not representative, but we are more interested in a ranked list than in the relative weight of these variables.

Information about currents is tied for second place with salinity in the list of scientists' requirements. In contrast, our analysis indicates that the general public is more interested in "sea ice" or "waves." The public's interest in sea ice must be understood in a climate change context. For example, the number of tweets referring to this variable was extremely high after the Intergovernmental Panel on Climate Change (IPCC) press release on September 27, 2013, that summarized the IPCC Working Group I Fifth Assessment Report (IPCC, 2013). Increasing interest in sea ice is also observed in the number of Google queries during 2012 in comparison with the whole available period. Thus, the popularity of specific oceanographic variables in queries and social networks reflects real-time concerns of the general public, as measured here in the case of sea ice. On the other hand, sea state conditions are very popular both in queries to Google and in Twitter, where many tweets make reference to the height or period of sea waves. Knowledge of sea state conditions is required for many nautical sports or activities, and some of these users have automated the posting of some ocean buoy records through this social network. Here, we have tried to minimize the number of these automated posts in our analysis; otherwise, this variable would be the most popular in Twitter, surpassing temperature. Sea and ocean currents occupy the fourth place among the interests of the general public, followed by salinity and algal blooms.

"DATA PRODUCERS INTERESTED IN TARGETING THE GENERAL PUBLIC SHOULD DESIGN SITES THAT USE TERMINOLOGY FAMILIAR TO THEIR AUDIENCES AND PRIORITIZE END-USER INTERESTS, AND THEY SHOULD FACILITATE INTERACTION WITH THE PUBLIC THROUGH SOCIAL NETWORK SITES."

The popularity of tweets reporting sea state conditions encourages us to include a ranking of interests of Facebook users in nautical activities or sports (Table 2). This list offers data producers an idea of the potential market

Table 2. Precise interests (‰) of Facebook users by country concerning sports and activities at sea (updated on September 30, 2013). No data are shown when the number of people with a specific interest is below 1,000.

	Facebook Community	USA	Canada	UK	Ireland	New Zealand	Australia	France	Spain
Surf	33.70	43.33	24.21	20.56	35.83	62.50	67.19	30.00	29.47
Fishing	29.30	67.78	49.47	34.44	37.50	63.33	67.19	18.57	11.58
Scuba diving	10.40	20.00	11.58	10.56	10.83	14.17	14.06	4.29	3.26
Sailing	8.00	15.56	12.63	13.89	15.00	17.50	21.88	6.36	2.53
Kitesurfing	3.70	4.00	2.74	4.00	5.08	4.75	5.31	7.00	5.26
Windsurfing	2.70	2.78	1.68	2.44	2.50	1.75	2.03	5.93	2.32
Stand-up paddle surfing	1.90	4.44	1.47	0.72	1.17	2.92	3.13	1.50	2.42
Submarine fishing	0.74	0.91	0.27	0.17	–	3.50	2.34	0.54	0.67
Sea kayaking	0.21	0.47	0.81	0.56	1.25	1.00	0.47	0.11	0.10

size of consumers of oceanographic products and can help them adapt their information for these sectors. As an example, surfing is the nautical sport with the highest popularity, particularly in Australia and New Zealand, where ~ 6–7% of Facebook users express interest in it. Hence, we would expect high interest in tides and wave conditions. As another example, French providers should be aware of the high popularity of wind- and kitesurfing in their countries compared to others.

Finally, data providers should be sensitive to the technology and kinds of devices used to access the information, which will completely determine user experience. Access from mobile browsers is rising. As an example, during summer 2013, a Web service developed by the Iberian Margin Ocean Observatory (RAIA), which provides information on sea surface temperature off the beaches of Northwest Iberia (http://playas.ieo.es), received more than 35% of its visits from mobile browsers (19.31% from Android, 14.66% from iOS Apple, and the rest from other mobile browsers). Thus, complex input forms should not be included in future designs for the mobile interface, although dedicated and more complex designs remain useful for desktop users and scientists.

SUMMARY AND CONCLUSIONS

The results of this study show an increasing number of people seeking nowcast and forecast marine products as revealed by the query share in the Google search engine, particularly among the Spanish-speaking community. This interest contrasts with the need for hindcast data among the community of fisheries and applied environmental scientists. Data providers should be aware of differences among users as they attempt to fully satisfy user demand. As a starting point, they need to understand the terminology used by the general public when searching for this operational oceanography information on the Web, and they should be aware that specific terms differ among countries. For example, people from New Zealand prefer the term "marine" as the first option in their searches, people from United States prefer "ocean," and people from the UK prefer "sea." In general, the term "weather" is preferred to "forecast." In addition, the terminology may vary over time for a specific country or region. For example, US queries before 2007 mainly combined the terms "marine" and "weather," while recently the terms "ocean" and "weather" are used more frequently. The number of searches also shows seasonality, with peaks in July and August and minima in February and March, although secondary peaks can be observed during the austral summer. Also, the list of oceanographic variables that is of greater interest among the general public differs from the list that is important to environmental scientists. Whereas the public is more interested in temperature, wave conditions, and sea ice, the scientists focus on temperature, currents, and salinity. Moreover, the interest of the general public in a specific variable is seasonal: information about currents is searched more during winter, about waves during spring, and about tides and temperature during summer. The specific interests expressed in the social networks are determined by societal concerns, news, and trends, as observed by the high number of tweets about sea ice after climate change related news was released. Data producers interested in targeting the general public should design sites that use terminology familiar to their audiences and prioritize end-user interests, and they should facilitate interaction with the public through social network sites. This final effort in the value chain of operational oceanography would benefit the individuals and institutions involved in creating and curating these scientific data sets.

ACKNOWLEDGEMENTS

We are grateful to Luz M. García and two anonymous reviewers for their helpful comments. Pablo Otero and the RAIA Coastal Observatory have been financed by the Programa Operativo de Cooperación Transfronteriza España-Portugal POCTEP 2007-2013 (0313-RAIA-1-E).

REFERENCES

Askitas, N., and K.F. Zimmermann. 2009. Google econometrics and unemployment forecasting. *Applied Economics Quarterly* 55:107–120, http://dx.doi.org/10.3790/aeq.55.2.107.

Baker, S., and A. Fradkin. 2011. What drives job search? Evidence from Google search data. Technical report, Stanford University, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1811247.

Berx, B., M. Dickey-Collas, M.D. Skogen, Y-H. De Roeck, H. Klein, R. Barciela, R.M. Forster, E. Dombrowsky, M. Huret, M. Payne, and others. 2011. Does operational oceanography address the needs of fisheries and applied environmental scientists? *Oceanography* 24(1):166–171, http://dx.doi.org/10.5670/oceanog.2011.14.

Brownstein, J.S., C.C. Freifeld, and L.C. Madoff. 2009. Digital disease detection: Harnessing the Web for public health surveillance. *New England Journal of Medicine* 360:2,153–2,157, http://dx.doi.org/10.1056/NEJMp0900702.

Burton, S.H., K.W. Tanner, C.G. Giraud-Carrier, J.H. West, and M.D. Barnes. 2012. "Right time, right place" health communication on Twitter: Value and accuracy of location information. *Journal of Medical Internet Research* 14(6), e156, http://dx.doi.org/10.2196/jmir.2121.

Ethnologue. 2013. *Ethnologue: Languages of the World*, 17th ed. Published by SIL International, http://www.ethnologue.com.

Facebook, Inc. 2012. Form 10-K Annual Report 2012. Retrieved from Securities and Exchange Commission EDGAR (Electronic Data Gathering, Analysis, and Retrieval system) database website: http://www.sec.gov/edgar.shtml.

Ginsberg, J., M.H. Mohebbi, R.S. Patel, L. Brammer, M.S. Smolinski, and L. Brilliant. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457:1,012–1,014, http://dx.doi.org/10.1038/nature07634.

Goel, S., J.M. Hofman, S. Lahaie, D.M. Pennock, and D.J. Watts. 2010. Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences of the United States of America* 7:17,486–17,490, http://dx.doi.org/10.1073/pnas.1005962107.

Hulth, A., G. Rydevik, and A. Linde. 2009. Web queries as a source for syndromic surveillance. *PLoS ONE* 4(2):e4378, http://dx.doi.org/10.1371/journal.pone.0004378.

IPCC (Intergovernmental Panel on Climate Change). 2013. *Climate Change 2013: The Physical Science Basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, eds, Cambridge University Press, Cambridge, United Kingdom, and New York, NY, USA, 1,535 pp.

Otero, P., M. Ruiz-Villarreal, L. Garcia-Garcia, M. Marta-Almeida, M. Cobas, G. González-Nuevo, and J.M. Cabanas. 2011. Walking on the sea side: Modeling and observational efforts of the Iberian Margin Ocean Observatory (RAIA). OCEANS, 2011 IEEE – Spain, June 6–9, 2011, http://dx.doi.org/10.1109/Oceans-Spain.2011.6003564.

Paul, J.M., and M. Dredze. 2013. Drug extraction from the Web: Summarizing drug experiences with multi-dimensional topic models. Pp. 168–178 in *Proceedings of NAACL-HLT 2013*. Atlanta, Georgia, 9–14 June 2013, N13-1017, Association for Computational Linguistics, http://aclweb.org/anthology//N/N13.

Polfeldt, T. 2006. Making environment statistics useful: A Third World perspective. *Environmetrics* 17:219–226, http://dx.doi.org/10.1002/env.747.

Preis, T., D. Reith, and H.E. Stanley. 2010. Complex dynamics of our economic life on different scales: Insights from search engine query data. *Philosophical Transactions of the Royal Society A* 368:5,707–5,719, http://dx.doi.org/10.1098/rsta.2010.0284.

Preis, T., H.S. Moat, H.E. Stanley, and S.R. Bishop. 2012. Quantifying the advantage of looking forward. *Scientific Reports* 2:350, http://dx.doi.org/10.1038/srep00350.

Myslín, M., S.H. Zhu, W. Chapman, and M. Conway. 2013. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. *Journal of Medical Internet Research* 15(8), e174, http://dx.doi.org/10.2196/jmir.2534.

Ruhl, H.A., M. André, L. Beranzoli, M.N. Çağatay, A. Colaço, M. Cannat, J.J. Dañobeitia, P. Favali, L. Géli, M. Gillooly, and others. 2011. Societal need for improved understanding of climate change, anthropogenic impacts, and geo-hazard warning drive development of ocean observatories in European Seas. *Progress in Oceanography* 91:1–33, http://dx.doi.org/10.1016/j.pocean.2011.05.001.

Sakaki, T., M. Okazaki, and Y. Matsuo. 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors. Pp. 851–860 in *WWW'10: Proceedings of the 19th International Conference on World Wide Web*. ACM, New York, USA, http://dx.doi.org/10.1145/1772690.1772777.

Signorini, A., A.M. Segre, and P.M. Polgreen. 2011. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS ONE* 6(5):e19467, http://dx.doi.org/10.1371/journal.pone.0019467.

Wilson, R.E., S.D. Gosling, and L.T. Graham. 2012. A review of Facebook research in the social sciences. *Perspectives on Psychological Science* 7:203, http://dx.doi.org/10.1177/1745691612442904.