Swinar

ISSN: 2744-9343

A transparent ocean: why open data?

Data tendencies and FAIR Data

Nursery, CarlosArango

Winner of the CEMarin 2020 photography contest and cover of this edition of Sonar



#### Contributors to this edition

Andrés Fernando Osorio, PhD,

Executive Director CEMarin, Universidad Nacional de Colombia

Juan David Osorio, PhD,

CEMarin researcher, Universidad Nacional de Colombia

Adriana Marcela Ruiz Pineda,

CEMarin researcher, Universidad de Antioquia

Paula Cristina Sierra, PhD,

Invemar

Diana Ruiz Pino, PhD,

Sorbonne University

Dairo Escobar,

SiB Colombia Coordinator

Ana María Aguirre Cañas,

Director Colombia Connect

Natalia Ossa-Hernández,

CEMarin early stage researcher

Myriam Elizabeth Vargas,

CEMarin early stage researcher, Universidad de Los Andes & Justus Liebig University

Johann K. Delgado,

CEMarin alumni, Cornell Ornithology Lab Fellow

#### **Editorial staff CEMarin**

**Louise Lowe** 

Content Creator

**Ballantyne Puin** 

Communications Professional

Emmanuela Daza

Scientific Manager

Irene Lema

**Communications Director** 

Diseño Gráfico

Jorge Mario Ordóñez Duque

**Administrative Staff CEMarin** 

Marcela Sánchez - Administrative Director

Aída Acevedo - Administrative Coordinator

#### s#nar

Sonar is a publication of the Corporation Center of Excellence in Marine Sciences - CEMarin.

©2021. All Rights Reserved. Material in this issue may be photocopied by individual scientists for research or classroom use. Permission is also granted to use short quotes, figures, and tables for publication in scientific books and journals. For permission for any other uses, contact info@cemarin.org

Sonar is published bi-annualy by CEMarin, Carrera 21 No. 35-53, Bogotá, Colombia.

ISSN: 2744-9343















#### CONTENT

- 1 Editorial
- 2 Opinion
- 5 Data Tendencies: Challenges and Opportunities
- 8 Open and FAIR Data
- 11 COCAS: A strategy to face Climate Change and the need for data in Global South Oceans
- 13 Marine biodiversity and Open Data: backing the ocean we want
- 16 ColombiaConnect
- 17 Research Report I
- 18 Research Report II
- 20 Alumnus Report
- 22 CEMarin Activity Report

### EDITORIAL Andrés Fernando Osorio

## A transparent ocean- why open data?

his year marks the beginning of the Decade of Ocean Sciences for Sustainable Development (2021-2030). For those of us who live and work around the oceans, this represents multiple challenges and opportunities. This global initiative highlights the importance of Open Data. One of the objectives of the Decade is a transparent ocean with open access to data, information and technologies, and one of the ten challenges identified is the delivery of data, knowledge and technology to all. This is especially important in terms of articulating the dialogues between science, policy and society through access to data, information and communication.

To advance towards the Decade's mission of a transparent ocean and transformative ocean science solutions for sustainable development, we chose Open Data as the central theme of this issue. It occupies a central role in the most important problems facing marine ecosystems today. This subject is also key to achieving our own mission at CEMarin, and reflects the fact that "infrastructure, technology, models and data" is a cross-cutting sub-theme in our Strategic Research Plan.

Our researchers and member universities are increasingly using Big Data to complement field studies, test alternative scenarios, increase the statistical power of experiments, and infer past changes and predict future scenarios. To do this, in addition to the data that we can collectively generate, we often rely on data derived from public databases. It is both a challenge and an opportunity for marine science to transform data into knowledge and to connect, engage and empower all stakeholders, from the local to the global level, to address these issues.

However, the panorama at the national level in Colombia and at the international level, shows serious difficulties in an efficient and effective joint use of data. The Colombian government is committed to the idea of open data and it is true that there are a large number of public entities, organizations and systems involved in the collection and management of the wide-ranging data relevant for the marine sciences. Additionally, there are also many other actors involved in these processes, including NGOs, academia, and private corporations. As we explore in this issue, successfully articulating these diverse actors, integrating these systems, and ensuring open access remains a major challenge. Clearly, meaningful joint use of data at local, national, regional and global levels is also critical. From academia, a change of approach is also required regarding data management, its dissemination and analysis.

Open data is thus a key tool in this critical moment for marine science, and society as a whole, to better use, coordinate and integrate knowledge, models and data into decision-making processes, and in helping us understand the best management practices and possible future scenarios for our oceans. Understanding new global trends and concepts such as FAIR data, which should be systematically implemented across the world, is essential. Proper handling of data will undoubtedly be one of the great challenges of this decade and that is why we wanted to open this conversation. In this issue, we invite you to discover different perspectives on the subject of data in Colombia and the region.

<sup>1</sup>Professor Universidad Nacional de Colombia, Executive Director CEMarin

٦

#### **OPINION**

CEMarin researchers
Juan David Osorio y Adriana Ruiz

## Open data for a transparent ocean



Providence island post Hurricane Iota. Image: Irene Lema

transparent ocean with open access data is one of the objectives of the Decade of Ocean Sciences. Achieving this at the national and international levels is crucial, as the use of big data facilitates decision-making by allowing modeling, and consequently planning and prediction based on reliable information. What's more, it opens the door to integrating the variety of existing data sources, transforming data into knowledge, and connecting, engaging and empowering citizens, governments and industries.

In the case of Colombia, one of the challenges to be solved is precisely how to handle raw and processed data in a way that allows access to a lot of information that exists in isolation. Regarding the management of marine-coastal information in the country, the main public entities that collect, process and publish data in their information systems (IS) are the Institute of Marine and Coastal Research (Invemar), the Colombian Ocean Commission (CCO), the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM), the General Maritime Directorate (Dimar) and the National Authority for Aquaculture and Fisheries (AUNAP).

The IS that stores all environmental information is the Colombian Environmental Information System (SIAC) managed by IDEAM, which includes the Marine Environmental Information System (SIAM). The latter is in charge of collecting and structuring information associated with marine-coastal topics. SIAM is part of the Ocean Biogeographic Information System (OBIS) and the Global Facilitating Mechanism for Biodiversity Information (GBIF), two of the most successful systems in the world in open data. Additionally, the CCO, Dimar and AUNAP manage their own IS.

Despite the existence of these platforms, there remain difficulties in articulating them, limiting the interoperability and integration of the available information, as well as its social appropriation. An evaluation led by the National Planning Department (DNP) concluded that the IS have not had a major effect on decision-making processes and have not been used as inputs for the formulation and implementation of management, planning and ordering tools<sup>1</sup>.

The national panorama is even more complex due to the existence of NGOs, universities and corporations, among other organizations, which also collect and process data. Their results are generally published in scientific articles or other platforms, while access to raw data is usually restricted. Furthermore, an interdisciplinary field like marine sciences, the involvement of a variety of organizations and platforms that address the socioeconomic aspect is also required.

To understand how these problems affect the day-to-day life of the marine sciences in Colombia, we interviewed two CEMarin researchers, Adriana Marcela Ruíz Pineda and Juan David Osorio, experts in food security and sovereignty and hydrodynamics and coastal protection, respectively. We asked them to tell us about their experiences in obtaining and handling data for their investigations.

- CEMarin: What data do you usually need for your research and how has your experience of acquiring it been?
- Adriana Marcela Ruíz Pineda: The focus of my research is food security and sovereignty, and the nutritional and health status of the populations that inhabit coastal areas. The data that I use usually comes from macro research funded by the national government or local and departmental authorities, especially when biochemical data analysis is required, as this is expensive and complex because it is about evaluating humans. Data referring to socioeconomic and demographic indicators is obtained from official bodies, for example, the National Administrative Department of Statistics (DANE), the DNP and the Ministries. In general, it is easy to obtain the information that these organizations publish on their websites, but accessing other information that they do not make publicly available is a complex and fruitless process. Access to research databases funded by national, departmental or local governments is relatively possible through a formal application process to the institution that coordinated the research, however this is a long process, sometimes only the internal procedures of the institution can last months.
- **Juan David Osorio:** My research focuses on the study of coral reefs as natural coastal protection structures. The data that I use are time series of hydrodynamic, meteorological and topo-bathym-

etry variables typical of the area. I get the data from secondary information or direct measurements in the field. Acquiring information has always been difficult when we depend on state institutions or private companies. Even in the case of public institutions, some charge for the information and others make it available but with problems and gaps in spatial and/or temporal resolution. Additionally, in some cases it is necessary to establish collaboration agreements in order to access it.

- **CEMarin:** If you obtain raw data in the field, laboratory or otherwise, what do you do with this information and do you consider it important that this data is made public?
- AMRP: Much of my research is derived from secondary data, so what I have to share are the results. These are shared through scientific publications and in other media such as the scientific dissemination magazines of the University of Antioquia and booklets aimed at the communities involved in the projects. However, I believe that the data should be public, especially if it belongs to research financed with public money. The entities that coordinate these investigations or databases should establish the procedures, requirements and schedules to access their information, and report it publicly.
- JDO: So far I use the raw data for my research and I give it to whoever requests it once the publications are ready. We are currently working on several articles where the data is released precisely for public use. The intention is that the data is made available for further investigation based on said information, and not necessarily in my area of work.
- CEMarin: Finally, do you have any suggestions on how Colombian government entities could articulate, among themselves and with academia, to guarantee access to knowledge and open and quality data?
- AMRP: The Ministry of Science and Technology could coordinate the information produced by government entities, centralizing and publishing generalities about their information, grouping it by knowledge area and adding links that lead to the repositories where it can be downloaded. The Ministry could also coordinate research information from universities and other actors, not necessarily managing databases but managing, perhaps



Image: CEMarin archive

on a web page or in an app, general research data, so that researchers can consult, based on their field of work, what investigations have been carried out in this regard and thus be able to contact those who have already advanced on that path.

We must improve the dissemination of results beyond scientific publications that are slow and use scientific language that ordinary people do not understand. We must establish mechanisms that allow us to get closer to people and to show them the results of our analyses with the information they provide us. These new ways of communicating must have academic value, we must value newsletters, pamphlets or a series of radio programs that explain our results in the same way as a scientific article in a Q1 journal.

- **JDO:** Once the commitments of the projects and the confidentiality clauses have been met, I believe that the data should remain public after completing the research, especially if they are projects financed with public resources. There is currently no policy or format for storing this information, which would be incredibly useful for future research. Generally, the data is stored without a clear protocol that indicates key aspects such as dates, the units of the variables, or the collection method. This makes it difficult for third parties to analyze the data or assess its

quality. The articulation of data by academia is a difficult but not impossible task that requires the integration of various disciplines and the knowledge associated with the protocols for the management of databases and information. This is a matter of having resources - time and money - for something that generally no one does at the end of a project: organize and put the information in a repository, following defined benchmark protocols. I think it is a task that begins in the working group of each researcher and expands to other universities or institutes that may follow their example.

— Although the major challenges shown in this panorama are clear, the government's Bioceanic Conpes

proposes specific actions that consolidate the marine-coastal information of Colombia. The Administrative Department of the Presidency of the Republic will strengthen the National Oceanic and Coastal Information System (SINOC) of the CCO; Dimar will create and lead the maritime, river and coastal spatial data infrastructure; the DNP will include in its Terridata platform, in coordination with the Colombian Navy and the CCO, standardized and comparable coastal-marine indicators to be viewed at the municipal, departmental and regional levels; and the DANE will update and strengthen the estimates of the added value of marine economic activities.

At CEMarin, we know that making progress on this issue requires complex inter-institutional management and the design of multidisciplinary solutions. Data collection must be coordinated and standardized, consolidating the country's marine-coastal information, and guaranteeing free online access to data, direct from the ocean to the cloud. To achieve the transparent ocean promoted by the Decade of Ocean Sciences, it is hugely important that open data and free information become a bridge between academic knowledge, civil society and informed decision-making.

s#nar

<sup>1</sup> https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/3990.pdf (p. 24)



#### Paula Cristina Sierra Correa, PhD

paula.sierra@invemar.org.co

Head of Research and Information Invemar

# Data Tendencies: Challenges and Opportunities

nvemar is the Institute for Marine and Coastal Research in Colombia, linked to the Ministry of Environment and Sustainable Development, and we are responsible for the Marine Environmental Information System (SIAM) for Colombia. The SIAM is part of an ecosystem of subsystems that make up and contribute information to the Colombian Environmental Information System (SIAC), and includes many components and modules focused on specific topics. All these systems are available to the public and can be accessed through our website (https://siam.invermar.org.co).

We at Invemar administer the SIAM but this does not mean that we handle or produce all the data, rather that we build alliances with many institutes and entities that want to deposit their data on our servers, we offer them custody of their data while they maintain the right to define its accessibility. Additionally, we lead technical inputs for the development of public policies and tools on marine-coastal issues, making the data available to decision-makers.

Invemar was born in the 1960s and since then it has collected an immense amount of information. Before we became part of the Ministry of the Environment the data focused mainly on biodiversity, but in the last 25 years, physical, chemical and socioeconomic data have also been incorporated. We have the Museum of Marine Natural History for Colombia, MAKURIWA, supported by the Information System on Marine Biodiversity (SIBM), which is part of the SIAM. Researchers can access the many unique organisms watched over by the museum, and which, in the end, represent a very specific kind of open data. This data is shared globally with the Oceanic Biodiversity Information System (OBIS) and Global Biodiversity Information Facility (GBIF), and at the national level with the Colombian Biodiversity Information System (SIB Colombia). We work for the collection, analysis, custody and making available of the data of the 50% of the country that is coastal/marine.

Conducting science in the oceans involves very specific conditions. In the case of data collection, the costs of research are much higher and highly-specialized personnel are required to obtain quality information. We have different areas within Invemar. and personally, I coordinate the area of information management and research with the aim that all the information generated in Colombia by the marine sciences is made visible and used. In addition to biodiversity, we have different areas within Invemar, including geosciences and marine climate, valuation, use and bioprospecting of resources, and environmental quality. Personally, I coordinate the area of information management and research with the aim that the information generated in Colombia for marine sciences is made visible and used.



Global Data: Image by Pete Linforth from Pixabay.

Clearly, there are many challenges to achieving this, and the Decade of Ocean Sciences for Sustainable Development, and its goal of a transparent ocean. One is to stay at the forefront of the field with respect to both data collection technologies, information systems (IS) and complex multiscale analyses of research elements. In this sense, it is relevant to mention that it is common to find reluctance on

the part of researchers to "let go of" their data, and make them available in the cloud. This is highly understandable, given the arduous processes they undertake to produce it. However, it is worth clarifying that, from the point of view of those who administer the IS, we are very aware that there may be special conditions. For example, while some data can be 100% open, there are cases in which we apply a permission for use, where those who generated the data authorize requests for its use. This is the data that is not published openly on our platforms. For this reason, requests to access data usually require supporting signatures from the entity in which the researcher works and declarations about the purposes of the data use, its confidentiality, and the citation of the source of the data if it is used in publications. On these issues of information transparency and its use, I believe that the country has made considerable progress and we continue to learn.

direct data, as well as to visualize and present the data in the appropriate way for different audiences, especially to attract the attention of other researchers, decision-makers and potential donors. We need data to be simple enough for those who require simple information, but have the scientific rigor demanded by researchers.



Data network with nodes: Image by Gerd Altmman from Pixabay.

One of the most important challenges facing the country regarding data management is the need to make researchers understand how important it is to standardize the collection of their data, and a strengthening in data analysis. Historically, we have been a country of "data collectors", but we are not so good at documenting this data and its metadata. This is something that can be taught and learned. It is clear that sometimes researchers do not know how to deliver their data and/or do not know how to request it or access it in a semi-automatic way, and this could be strengthened through academic training. We need that, from university, students know what they can do with their information in relation to ICTs. Academia must show students the range of possibilities of open data, free information and IS.

Another challenge is to train researchers in the use of new technologies to take advantage of the benefits they offer for the collection of direct and in-

So what can we do to face these challenges? First, we need to abide by existing standards for certain types of information, for example those related to remote sensing, navigability and access to ports, biodiversity and marine environmental quality, among many others. These standards facilitate the availability of data and the undertaking of complex comparative analyses from the local to the international, for example, in climate change analysis with a local focus, but global application. In Colombia, from the National Administrative Department of Statistics (DANE), the national government has promoted a statistics system that establishes standards for how data should be collected. In the case of the multiple IS, a great effort has been made in the country to improve their interoperability through their standardization. The National Technical Committee for the Coordination of Oceanic Data and Information (CTN Diocean) coordinated by Dimar in terms of data, can play a key role in

this process. Invemar contributes to this Committee as an Associated Unit of Data and Information of the Intergovernmental Oceanographic Commission (IOC) - UNESCO, as an information node (processed data).

Invemar is the Regional Training Center for Latin America on ocean-related issues, and among the

> many courses we offer is one on the publication and management of marine biodiversity data (OBIS), with the participation of experts of the IOC, promoting the knowledge and use of established international standards, such as Darwin Core and Ecological Metadata Language. The course has a focus on applying these standards in practice, we ask students to bring their own data in order to put it into the forms required by these standards, and to share them and analyze them with international data. Through education, we hope that the requests we receive from researchers will become increasingly detailed and specific, following the requirements of the international standards, just as we expect these to be applied to the data we receive and manage.

> Data science, and particularly big data, offers us many opportunities for the analysis of multiple data across different spatial and temporal scales, and by connecting

the different IS available a more complete analysis is carried out to be able to extract the information key to explaining a phenomenon. Taking advantage of ICT, at Invemar we offer an advisory service for researchers on the collection and use of data, and specifically on mechanisms of data extraction, for example, web scraping. Web scraping is an automated search for existing information on the internet and specific IS. Invemar has used it through the Clearing House Mechanism LAC, a platform that provides users with quick, direct access to relevant sources of information. This mechanism can replace the process of going to a library, searching for books, and reading those books to locate the desired information -or an equivalent process using digital sources- all through an algorithm. Regarding these data extraction mechanisms, Invemar is part of the Ocean InfoHub project for Latin America and the Caribbean, with the objective of contributing to the achievement of the transparent ocean established by the Decade of Ocean Sciences. These artificial intelligence systems can do many of the tasks that humans currently do, thus allowing researchers, with the appropriate training, to concentrate most of their efforts on analysis, the most important objective of our times in this field.

Given that ICTs are also expensive and complex, another challenge is securing an adequate, and greater investment by the state for the oceans, through the Ministry of Science, Technology and Innovation, to be able to maintain the IS and facilitate their interoperability, update the data, and make it available to researchers and the general public by employing the best data extraction mechanisms offered by ICT.

All of this goes hand in hand with academia and other data-generating actors, who must incorporate in their research projects the process of including metadata and documenting it well, instead of leaving these important actions as an afterthought upon completion of the investigations. This should be just as important as delivering the results to the financial sources of the projects and will ensure that the data can be correctly distributed and used in future research.

To continue advancing the necessary protocols, and to stay at the forefront of marine and data sciences, it is highly important that all these processes are developed jointly between those responsible for the IS and the researchers who share and use the data. It is not simple or quick to apply international standards and it is true that in Colombia we continue, to some extent, to lag behind in this matter. Without doubt, changing people's way of thinking is not easy and it does not happen overnight, but I believe that we are experiencing a process of change, in which the value of sharing data is recognized. We no longer accept the old assumption of "who has the information has the power", from now on it is those who know how to take advantage of ICT, extract data and analyze the information who will have the power to make our oceans transparent and accessible.



# Open and FAIR Data



Ocean of data. Image: Jorge Mario Ordóñez Duque

ata is incredibly important for scientific research and society in general, but it is also clear that there are significant gaps regarding data collection and access, both in Colombia and globally. While it sounds paradoxical, the saying that we know more about the surface of the moon than we do about the ocean is totally true. We invited Captain Juan Camilo Forero, Executive Secretary of the Colombian Ocean Commission (CCO) and Dr Alejandro Orfila, researcher at the Spanish Higher Council for Scientific Research, to discuss the important theme of oceanographic data in Colombia, and global trends in this area, with our Executive Director, Dr Andrés Osorio.

Beginning with the question of why, in general terms, data is so important, according to Dr Orfila, we are totally unaware of what happens inside the ocean. It is essential to be able to measure its data and variables, as well as those of other marinecoastal ecosystems, in order to identify and explain the phenomena that we see today, for example, rising sea levels and increasing temperatures. Moreover, this is key to solving current and future problems: data is the basis of everything. In

the specific case of Colombia, Captain Forero commented that there is a huge amount of existing work generating data, from a range of actors and institutions. But he stressed that "the Decade also imposes a challenge on us: this transparency that the ocean must have regarding the availability of information, availability of data, which will help us to know, in an integrated way, a little more about this wonderful environment which, as the doctor said a moment ago, we know very little about." The transparent ocean objective of the Decade of Ocean Sciences goes hand in hand with the reality of the day-to-day

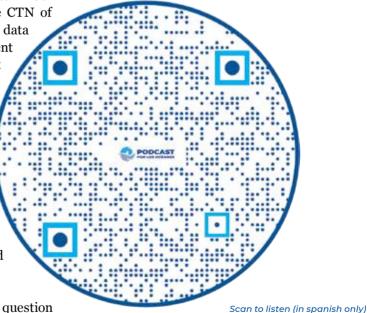
needs of researching and decision-making related to our oceans.

Both guests agree on the need for interinstitutional articulation and the interoperability of platforms in order to consolidate existing information and fill in the gaps, and above all to make the data available to everyone, and to researchers and decisionmakers in particular. Captain Forero told us that, in the Colombian context, this process is led by the National Technical Committee for the Coordination of Ocean Data and Information of the CCO (CTN). Dr. Orfila described the Spanish context: "There is no doubt that the policy from the State has always been that the data is public; the data that is financed with taxes, with public money, has to be public." He added that, it is recognized today that even a selfish researcher understands that sharing their data brings them more benefits than saving it privately, in terms of knowledge and the taking up of their projects by other actors.

This reflects the broader context of the European Union, where there is a 2018 mandate from the European Council regarding FAIR data, according to which data generated in EU member states with public resources -both raw and processed- must be public and follow quality controls and standardization according to international standards. The acronym FAIR refers to findability - that data can be found through mechanisms such as a doi, accessibility - that any actor, from the university or largest corporation to a citizen who pays taxes, can access the data, interoperability and reusability - that data is published according to international norms for standardization.

Regarding the challenges in open access to data, the guests told us about the short-term actions that can help address these. Captain Forero highlighted the standardization and integration of the large amount of information that already exists in Colombia, for example, the creation by the CTN of reference manuals on best practices for data management to incentivize the fulfilment of the principles embedded in the concept of FAIR data, even if different entities follow their own protocols. A fundamental objective of this initiative is to avoid the duplication of work by different actors and thus also reduce research costs. In the European case, Dr Orfila categorized this process as, to a certain extent, a natural evolution with the rise of information and communication technologies and an emphasis on transparency in all areas of public life and in all public entities.

Finally, the participants discussed the question of how, from academia and entities like CEMarin, where a key part of our mission is to promote collaboration and networking in marine sciences, a new culture and paradigm can be reached in which any data collected becomes, almost automatically, published data to achieve a real system of FAIR data in Colombia. Captain Forero suggested that universities and research centers need to demonstrate that the value of data goes far beyond a single project, and that data stored on a hard drive loses much of its value if no one can access and use it, as long as this goes hand in hand with a proper and responsible use of the data generated by others. In this last aspect, academia and public entities must work together; they have a crucial role in the care of the scientific heritage produced by the nation for the world, and in promoting the culture of open data. Dr Orfila added that academia and research are about learning from the knowledge produced by others, and also from their mistakes. Therefore, making data available to others contributes to this ongoing work, to a shared pact between all researchers and academics, and adds value to the projects of those who generate and share their data.



can to listen (in spanish only)





## COCAS: A strategy to face Climate Change and the need for data in Global South Oceans

iana Ruiz Pino is a researcher and specialist in the biogeochemical carbon cycle and oxygen in the ocean, a pioneer in the evolution of CO<sub>2</sub> and oxygen minimum zones (OMZ) in the ocean. She is author of more than 60 publications related to the role of the ocean in atmospheric CO<sub>2</sub> capture and its impacts on acidification and deoxygenation. She has contributed to the development of the first biogeochemical models that link the carbon cycle and marine ecosystems, used today for weather forecasting. Diana is currently responsible for the first platform of autonomous coastal buovs in southern latitudes: COCAS, the Coastal Ocean Observatory for Climate, CO2 and Acidification in the Global South. This project aims to carry out in real time, and for the first time at high frequency, the long-term monitoring of the impacts of climate change in the Exclusive Economic Zones of the Southern Hemisphere (GSEEZ). The GSEEZ are the most populated, inhabited and tourist marine regions. Much of the world's population depends on these zones' resources and wealth but they are also the regions that could suffer most from the impacts of climate change.

In the absence of robust data and time series that allow informed decisionmaking, COCAS focuses on the tropical and subtropical coastal areas of the ocean that have not been sufficiently sampled, mainly measuring physical, biogeochemical and biological marine parameters. Through a surveillance system consisting of fixed buoys fitted with sensors, placed in the air and in the sea, data of eight variables for the marine surface-atmosphere is recorded at different depths: temperature, humidity, pressure, rain, wind speed and direction, long and shortwave radiation. Data is also recorded for eleven variables for the water column: temper-

ature, salinity, currents, light intensity, turbidity, nitrates, oxygen, pH, dissolved carbon dioxide, estimates of the total biomass of phytoplankton and main functional groups of phytoplankton. The sensors are autonomous and the data is recorded automatically, either in each sensor or in a central brain, and transmitted in real time to the laboratories or users. More than twenty COCAS buoys are already deployed, while others are in development. On average, COCAS scientists spend more than four million Euros per year to maintain and use their coastal moorings.

This process of researching and monitoring the variables in COCAS has been planned in such a way as to generate a link between science and the needs of society. In this way, COCAS manages to integrate marine sciences with end users, ranging from local communities to national and international public and private organizations. This joint work has enabled the identification of science-based solutions to environmental problems and ecosystem services, as well as fostering the use and development of low-cost adapted sensors and platforms.

Among the commitments of the COCAS group is to develop the long-term capacity of new generations of scientists and end users with a gender balance, based on transformative and hybrid science (that is, science that takes an active role in social changes and which fuses modern and traditional knowledge)<sup>1</sup>.

Obtaining marine data of international quality is not only one of COCAS' priorities, but also a necessity for studies on temporal variability and the impacts of climate change on the ocean. This quality of data implies both the use of sensors and buoys

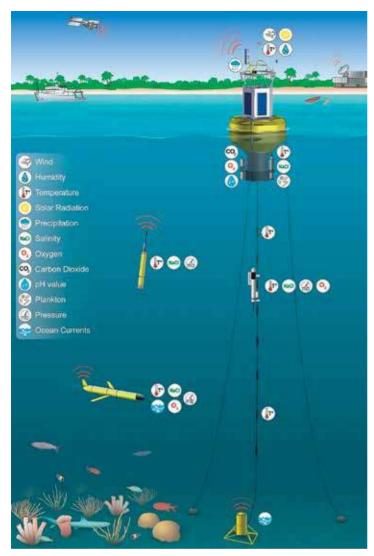


Image. Diagram of a COCAS buoy. Taken from: ECO UN Ocean Decade 2021

adapted to local conditions, and the implementation of calibration, validation and intercomparison protocols.

Ocean and atmosphere are two reservoirs "without borders" and where impacts of both natural and anthropic origin imply that the variability of the related regions is observed and taken into account. In the current context of the climate emergency, open data is a priority of all large-scale and new scientific programs. Education, training and knowledge transfer, not only between countries but also from academia to industry and society, mean that the search for solutions to current societal problems will benefit greatly from the sharing of data acquired by COCAS. Current databases used to validate climate

model forecasts are lacking in data for the southern oceans and GSEEZ. Reducing uncertainties and improving model forecasts requires not only the collection of quality data, but also its sharing and its openness to international use.

In this sense, COCAS promotes the creation and exploitation of a joint database of marine data obtained on the eastern and western edges of the Atlantic, Pacific and Indian oceans, and the Caribbean and Mediterranean seas. From a scientific point of view, with this strategy it will be possible to identify, understand and compare the origin and propagation of climatic and environmental signals and impacts that affect previously very distant populations and ecosystems. The comparison of the responses of resurgences or coral ecosystems, of fishervrich areas located to the east or west of the same ocean or in different oceans, or even the impact of the same signal, for example the Atlantic oscillation or the El Niño phenomenon, is an important element of COCAS' objectives. To achieve all these objectives, the operation of this project is based on the creation of a consortium of researchers, end users and stakeholders working in South-North and South-South synergy and benefiting each other nationally and internationally.

Today, more than 12 countries located in the southern hemisphere, 5 European Union member states, the United States, the United Kingdom and China have joined this project. Colombia, with coastal areas in the Caribbean and the Pacific, is now also part of COCAS.

In 2021, COCAS was honored as a project within the United Nations Ocean Decade program. The participation of COCAS in this program will help the long-term existence of the platform and to ensure the continuous obtaining of quality data over several decades, and thus ultimately allowing a clear vision of future evolutions.



# Marine biodiversity and open data: backing the ocean we want

he Decade of Ocean Sciences has knocked on our door in search of allies that allow us all, as a society, to build sustainability around this great ecosystem. With great satisfaction and hope, I find in this decade a commitment to information systems, open data, and the participatory processes of various actors, facilitating a better knowledge of the biodiversity of our oceans. In the current context of an unprecedented decline in biodiversity, data on the presence and abundance of species is an essential tool for planning, implementing and monitoring conservation and sustainable use strategies.

These types of commitments are not unknown in Colombia. As a country, we have been undertaking joint efforts for more than 20 years to build the best possible inventory of our biodiversity and make it available to anyone, regardless of their geographical location. Without a doubt, this is a task of immense complexity for a megadiverse country like ours.

So how have we been doing this? Let's start with two strategic information systems, the Information System on Biodiversity of Colombia - SiB Colombia¹ and the Information System on Marine Biodiversity - SiBM². These systems have the mission of facilitating the publication, access and use of data on biodiversity in our country, and in recent years they have been working together under the same data management model that brings users, publishers and producers of information closer together to support new research, education and decision-making processes related to the knowledge, conservation and sustainable use of biodiversity.



Image. National Biodiversity Information Systems.

¹https://sibcolombia.net/

<sup>&</sup>lt;sup>2</sup>https://siam.invemar.org.co/sibm

These systems, which operate as a large national network, are a reality thanks to the participation of hundreds of organizations and individuals who share data and information under the principles of free access, transparency, cooperation, recognition and shared responsibility. This is materialized in an open access policy called Create Share Transform, which works as a guide for those interested in understanding the dynamics of open access, including key concepts like open data, free licensing and participatory science around the data on biodiversity<sup>3</sup>. Additionally, the systems work in alliance with the World Biodiversity Information Infrastructure - GBIF<sup>4</sup> and the Ocean Biodiversity Information System<sup>5</sup>, global systems with country nodes.



Image. Diagram of articulation between Colombia's SiB and other national and international systems in which the data shared by publishing organizations is also visible.

A key question regards who publishes data through these systems? To date, SiB Colombia has 1,568 datasets published by 162 organizations. These publishers are mostly universities, government agencies, research centers and institutes, NGOs, the business sector, networks and initiatives of specialists and participatory science. Altogether, there are about 11.5 million individual pieces of data available for consultation and download. Of this, only 2% correspond to marine-coastal data, provided by 20 national organizations and with a prominent participation of the Institute for Marine and Coastal Research - Invemar, the main data publisher. If we continue to detail this data, we can quickly see that the five species with the most data are an alga and

four hard corals (Lobophora variegata, Orbicella faveolata, Agaricia agaricites, Orbicella annularis and Siderastrea siderea), among many other interesting metrics that we can dive into.

Published	Data
Instituto de Investigaciones Marinas y Costeras - Invemar	200.719
Universidad Nacional de Colombia	4.476
Universidad del Magdalena	3.948
Universidad Industrial de Santander	3.312
Moam Monitoreos Ambientales S.A.S	1.781
Universidad del Valle	1.106
Anadarko Colombia Company	893
Stratos Consultoría Geológica	849
Universidad de los Andes	666
Gobernación de San Andrés y Providencia	623
Fundación Malpelo y Otros Ecosistemas Marinos	566
Pontificia Universidad Javeriana	458
Fundación Universidad de Bogotá Jorge Tadeo Lozano	326
Fundación Colombia Azul	291
Universidad Pedagógica y Tecnológica de Colombia	271
Universidad de Antioquia	197
Asociación Calidris	172
Universidad del Sinú Cartagena	149
WWF Colombia	98
Coralina	65

Table: Top 20 national organizations in terms of publishing marine-coastal data.

	Published Data
Lobophora variegata	6.269
Orbicella faveolata	4.289
Agaricia agaricites	4.120
Orbicella annularis	4.055
Siderastrea siderea	3.970
Montastraea cavernosa	3.851
Orbicella franksi	3.671
Porites astreoides	3.622
Agaricia tenuifolia	3.569
Pocillopora damicornis	3.492

Table. Top 10 species in terms of data available from national organizations.

Another important question is what data can be published? Data on biodiversity comes from multiple sources: biological collections, inventories, monitoring, EIA, DNA barcodes, camera traps, remote sensors, participatory science, agrobiodiversity, graduate theses, among others. However, almost all of these can be structured for publication in three types of data sets, each with different degrees of complexity: Biological Records, Sampling

<sup>&</sup>lt;sup>3</sup>https://sibcolombia.net/acceso-abierto/

<sup>&</sup>lt;sup>4</sup>https://www.gbif.org/es/country/CO/summary

<sup>&</sup>lt;sup>5</sup>https://obis.org/node/d2f71b1b-9138-4aba-ad8f-8327ac3d04le

Events, and Species Lists. To be able to publish them, it is necessary to standardize these processes and seek the greatest completeness possible, so that the data can be used in different ways, and thus, contribute to answering different questions. Sharing as much data and information as possible through SiB Colombia is highly valuable, even data that may be considered incomplete contributes to knowledge about our biodiversity and serves to respond to important questions.

Next, we must ask what are the advantages of publishing data through these systems? An important part of the mission of these national systems is to promote a culture in which all people recognize the benefits of publishing open data about biodiversity, both for the publisher and for society in general.

- Making data visible and accessible through SiB Colombia, SiBM and other information infrastructures such as GBIF and OBIS, contributes to global knowledge about biodiversity and, therefore, to solutions that promote its conservation and sustainable use, as proposed by the Decade of the Oceans.
- It makes it easy to integrate data sets around the world and discover new opportunities for collaboration between data owners and other researchers.
- It enables individuals and organizations to gain appropriate visibility and recognition for their work as generators and curators of biodiversity data. Recognition will be higher if you are the author of a peer-reviewed data article, gaining traditional academic recognition for publishing biodiversity data sets. Additionally, these publications are recognized by the Colombian National System of Science, Technology and Innovation (SNCTI) as research products.
- It facilitates the tracking of the use and citations of data sets published and accessed through SiB Colombia and similar infrastructures (GBIF, OBIS).
- Some funding agencies now require that the researchers and organizations that receive their funds have a data management plan that includes its open publication at the end of a project, through platforms such as SiB Colombia.
- It contributes to the goals of the Post-2020 Global Framework for Biodiversity, especially those associated with invasive species, protected areas, threatened species and knowledge management.

- It contributes to the evaluation of SDG 14 (underwater life) and 15 (life in terrestrial ecosystems).
- It allows intergovernmental bodies such as GEO-BON and IPBES to build models and assessments of biodiversity respectively, at different scales.

#### What are the current challenges?

The concern to understand the factors that contribute to biodiversity loss has motivated the scientific community, civil society, policy makers and the business sector to focus efforts to consolidate an inventory of species that allows identifying patterns and changes over time and space. For more than 20 years, these information systems have been consolidated so that anyone, anywhere, has open access to data on biodiversity at the species level. We have built these systems to a point where they now have a national reputation as the "most comprehensive, well-known, and openly available access point to the country's species data." The interconnected challenges of biodiversity loss and climate change have increased the demand for, and urgency of, data for research, policy and decision-making from the local to the global scale.

These infrastructures must continue growing to include more and varied types of data and better computing services, in order to provide the biodiversity information required by research and national policies. This will make it easier for the community to continue to grow with the participation of new organizations, experts, and data users. The urgent need for solutions to multiple crises forces these information systems to build a truly national network capable of facilitating the publication of all potentially available biodiversity data through a robust infrastructure.

Finally, I would invite all generators and users of data on marine biodiversity to create a community around these information networks, where there is certainty that through open data we can facilitate the construction, use and democratization of knowledge, improving decision-making and people's quality of life. With the Decade of the Oceans, we as a country have a unique opportunity to achieve this, significantly increasing the mobilization of marine-coastal data as a catalyst for the generation of better knowledge.



#### COLOMBIACONNECT

# ColombiaCONNECT and CEMarin: united for a best practice project



he objective of ColombiaCONNECT is to promote the fair and sustainable use of biological resources in the country by building a network that links the areas of bioeconomy, biodiversity and studies on peace and armed conflict. This work began in 2020 when the German Federal Ministry of Higher Education and Research (BMBF, for its acronym in German) awarded 1 million euros to the consortium, which is composed of Colombian-German institutions. With these resources, the network has implemented four best practice projects, one of which is led by CEMarin, a founding member of ColombiaCONNECT.

The best practice project headed by CEMarin is called "Participatory aquaponics in Urabá fishing communities: a path to gender equality and environmental and socioeconomic sustainability" and is led by Dr. Jenny Leal Flórez, a CEMarin researcher. Dr. Leal and her team carry out, with the community of Turbo, Urabá, the assembly of two aquaponics systems that are operated by fisherwomen from the region using local fish species. Aquaponics is a closed production system that integrates the technique of aquaculture with hydroponics, that is, it is a combination of the production of fish and the production of vegetables without soil through a common means: water.

The selection of this community as the place to carry out this project responds to its specific characteristics. 63.5% of the population in Turbo has been a victim of the armed conflict, 15% live in a situation of extreme poverty, 39% do not have their basic needs satisfied and there is a depletion of fishing resources due to pollution and overexploitation. These circumstances are mixed with structural conditions that relegate women to supporting tasks that, in most cases, are not remunerated. For this reason, both CEMarin and ColombiaCONNECT considered that this was the appropriate place where the fields of interest that link the network converged.

This project is developed on three levels. First, in the social aspect, it includes a strong component of training with the fisherwomen in order to promote a knowledge exchange that strengthens the community in this area. Second, on a technical level, the community, ColombiaCONNECT and CEMarin have the objective of implementing aquaponic systems in a socially, environmentally and financially sustainable way. And finally, on an economic level, this project of excellence contains a market study and a business plan that will provide an economic assessment of ecosystem services.

In this way, ColombiaCONNECT and CEMarin are putting into practice the principles that were established as the network's guiding framework: keeping the interest of the community at the center of our work, using scientific knowledge to improve the quality of life in the places where we work, and promoting the fair and sustainable use of biological resources. We hope this is the first of many other best practice projects that we will undertake together.

#### **RESEARCH REPORT I**



CEMarin early stage researcher: Natalia Ossa-Hernández

(Doctoral student in Marine Sciences, Universidad del Valle)

Director & CEMarin researcher: José J. Tavera PH.D (Professor, Universidad del Valle)

# Species diversification in marine fishes in the Pacific and Caribbean basins

arine ecosystems are exposed to factors that alter their functions on different time scales. Anthropogenic factors caused by human intervention have shown important recent

impacts, and there are also natural ones of great magnitude that have occured over thousands or millions of years, such as the appearance of geographic barriers. The latter alters the structure and configuration of the ecosystems, for example, breaking the connection between two preexisting environments -the case of the Isthmus of Panama- isolating populations that experienced a continuous flow before the appearance of the barrier (Velasco 2018). The emergence of the Isthmus of Panama has been extensively studied, especially regarding how it altered the terrestrial and marine habitats of organisms and environmental conditions. However, somewhat less attention has been paid to the changes in the evolutionary history of the marine organisms on each side of the barrier after its emergence, and the differential effect it may have had on different taxa with different natural histories.







Photos: Natalia Ossa-Hernández

It has been shown that geographical barriers alter the dynamics of species diversification by isolating populations. Still, there are not many studies focused on patterns of diversification rates after the barrier appearance and if there is a predictable effect. There is a need to investigate whether patterns are more similar among closer and unrelated taxa in the same or different environments, or whether they depend on an interaction between the

> species trait and the environment (Day et al. 2008). In our project, we are collecting molecular and morphological data to explore the link between the emergence of the Isthmus of Panama and the diversification rates and modes of diversification in fishes of the families Carangidae, Centropomidae, Gerreidae, and Haemulidae. Localities like Gorgona Island. Guapi. Buenaventura and the Santa Marta region have been visited to collect specimens. For specimens with distributions outside Colombia, the sequences are taken from NCBI. Identifying the mode of species diversification is fundamental to understanding how biodiversity changes over evolutionary time, and how the direction and magnitude of morphological traits are related to diversification rates. In Colombia, we have the ideal scenario to conduct this research as we have access to both oceanic basins. Our research will provide basic knowledge essential for un-

derstanding how geography and ecology have interacted in the dynamics of species diversification and how they have been able to modulate rates of diversification and morphological evolution in the Pacific and Caribbean basins.

#### **RESEARCH REPORT II**

## CEMarin early stage researcher: Myriam Elizabeth Vargas Morales

(Doctoral student in Economy, Universidad de los Andes and Justus Liebig University, Giessen)

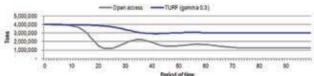
Director & CEMarin researcher: Jorge Higinio Maldonado Ph.D (Professor, Universidad de los Andes)

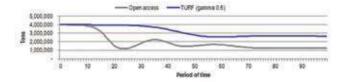
# Evaluating strategies for poverty reduction and the improvement of fishing performance in marine-coastal communities in developing countries

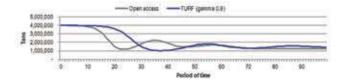
here are many factors that contribute to the sustained conditions of vulnerability and poverty in rural coastal households. These include high exposure to different shocks and economic and natural disasters; the communities' strong dependence on depleted or overexploited ecosystems and natural resources; limited access to alternative means of income generation; and the deficit of civil infrastructure, public services and access to markets. Designing appropriate policies to face these conditions requires understanding how fisheries management and poverty reduction mechanisms affect the communities' well-being and the fish stocks on which they depend.

One mechanism of fisheries management is Territorial Use Rights in Fisheries (TURF). These are management systems in which it is assumed that their implementation creates an incentive system among TURF users to exploit and manage resources in a sustainable way, and thus leading to the recovery of the resource as well as increased benefits for the agents who extract them. However, the successful consolidation of this result may be limited when rights are granted over the distribution area of a resource that cannot be retained. In this research, the effect of TURF is analyzed in terms of a free access regime, using a dynamic bioeconomic theoretical model and considering a) that the resource is mobile, and b) that the fleets that extract it can be economically and technologically differentiated. The results so far show that in a context in which TURFs fail to generate sufficient incentives for self-organization, TURFs tend to favor the sustainability of biomass over free access, although this by itself it does not rule out the danger of serious resource depletion.

In the context of poverty reduction, the literature usually highlights two mechanisms for rural households: the receipt of cash transfers and the decision to diversify their income. Although both have been extensively investigated in economic literature,







Illustrations: Biomass under free access and TURF with different spatial assignments of TURF (gamma) and fishing fleets differentiated by catchability coefficient and production costs.

their implications for rural coastal households remain unclear. Therefore, this research explores how monetary transfers affect the well-being of coastal households by analyzing a production model of fishing households where households are both producers and consumers, and therefore, their production decisions are governed by the rules of minimization of costs and their consumption decisions by the norms of maximization of utility. Theoretical results show that cash transfers have positive effects on the proportion of fish caught used by the household for self-consumption and the demand for other consumer goods, while there is no direct relationship between the transfers and the supply of fish or household labor.

Similarly, by determining the degree of income diversification for a rural coastal community in the Colombian Caribbean using Simpson's Diversification Index (SDI) and the instrumental variable technique, the effect between the SDI and per capita labor income of households is being established. Preliminary results show a positive effect of the SDI on per capita household income, with heterogeneous effects between non-fishing and fishing households, and for households located in the second and fourth income quartiles.

The evaluation of these three strategies offers elements of the socioeconomic dynamics of population groups that are highly dependent

on fishing resources. Thus, they can be incorporated into the design of policies aimed at the wellbeing of these communities, and therefore complementing the usual perspective that their well-being can only be achieved by improving the efficiency of their fishing efforts.

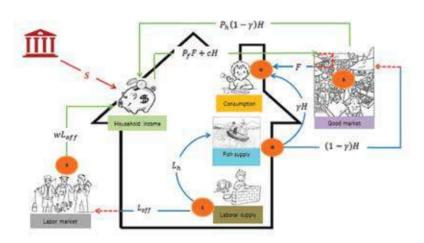


image: Fishing household production model with direct cash transfers





Photos A & B: Opening workshop to collect information used to identify the effect of income diversification on the per capita labor income of households



#### **CEMarin ALUMNUS REPORT**

#### Alumnus CEMarin: Johann K. Delgado

(Doctoral student at Cornell DeFrees Hydraulics Lab) jkd78@cornell.edu

Coastal Solutions Fellow at Cornell Lab of Ornithology

n 2015, the Paris Agreement (PA) was signed by 196 countries, opening the door to a new transition towards a greener and more climate-adapted world. This was a huge milestone for climate negotiations that started 25 years back. The PA highlighted the need to increase adaptation capacity in vulnerable developing countries where climate impacts would be most severe.

Colombia has already started moving towards a more resilient future. At a national and regional scale, the country is building climate adaptation plans that allow communities to develop better decision-making processes that account for climate hazards. However, at a local scale and especially for coastal areas, some constraints arise, reducing the effectiveness of climate adaptation plans. For instance, communities located along the Pacific coast are generally relatively small, remote, and socioeconomically vulnerable. Hence, adaptive capacities are restricted, leading communities to inaction

and their exposure to imminent consequences such as coastal floods and climate-driven displacements.

Broadly speaking, the success of coastal adaptation plans depends on coping with two major challenges. First, researchers need to develop local and accurate scientific information, with which policymakers and planners can structure projects to reduce risks, protect assets, and conserve coastal and marine ecosystems. Second, perhaps the uphill battle in this process is to achieve an efficient cross-sector collaboration among government, the private sector and communities, to advocate together for a real transformation.

On the Pacific coast of Colombia, a clear case combines all the aspects mentioned above: Punta Soldado Island, Buenaventura. There, its local Afrodescendant community has already suffered at least three forced relocations and experienced the dieback of tens of hectares of mangrove forest



Photo: Mangrove dieback on Punta Soldado Island, 2019



Photo: Infrastructure damage on the Punta Soldado Island, 2019

due to rapid beach erosion and anomalous coastal floods. In an effort to increase local knowledge, our research team is estimating changes in the magnitude and frequency of coastal floods, measuring rates of coastline retreat over the last three decades, and monitoring ecosystems such as mangroves, beaches, and mudflats. When combined, this information allows us to understand climate hazards and co-create better climate adaptation strategies for the island. Moreover, these insights become academic material for training the new generation of Punta Soldado's climate champions. However, improving collaboration and engagement with the local and regional governments is the remaining

challenge of this project. We hypothesize that due to the complexity of the social and economic problems that Buenaventura constantly faces, governments might deal with these issues with more priority than addressing climate change.

In conclusion, Punta Soldado Island is indirectly becoming a window to an imminent future in which a changing climate will force us to make decisions. The levels of understanding and organization that we can reach will define the impacts of such changes. Fulfilling international agreements will be only possible by achieving local and real transformations.

#### **CEMarin ACTIVITY REPORT**

n the months since our last edition, we at CE-Marin have been busy. Although we are all better-adapted to the challenges of working virtually, the first half of 2021 has brought further difficulties due to the complicated sociopolitical situation in Colombia. Nevertheless, we have been working hard on many projects and events aligned with our mission and Strategic Research Plan, and in collaboration with other leaders in the field of marine sciences at the national, regional and international levels.

Some of our current standout projects include:

- Participating in the HIDRIA-CEMarin Consortium, developing the consultation process for the development of the National Ocean Policy, Strategy and Action Plan of Panama. In this project we contribute our technical, scientific and academic experience in the field of marine sciences, and the set of knowledge, processes, abilities and skills required, through experts in the areas of the environment, marine biodiversity and coastal engineering.
- A detailed study to define risk for the Archipelago of San Andrés, Providencia and Santa Catalina regarding hurricanes in the most probable categories, a crucial project following severe destruction on the islands by Hurricane Iota in November 2020. We are working jointly with Coralina (the Corporation for Sustainable Development of the Archipelago) and the National University of Colombia. Field work was carried out in March and June 2021 and we continue modeling the data and developing a communications strategy to ensure we can make a robust contribution to future decision-making in this area.
- Contributing to the Climate Adaptation Plan on Punta Soldado Island in the municipality of Buenaventura, Colombia, in collaboration with the Coastal Solutions Program of Cornell University, USA, and the National University of Colombia. The objective of our contribution has been to obtain and transfer technical and

scientific knowledge to local communities and decision-makers in the Bay of Buenaventura regarding coastal vulnerability scenarios and nature-based adaptation strategies, allowing them to increase their adaptive capacity in the face of climate change scenarios that include sea level rise and extreme waves.

Additionally, at CEMarin we have actively organized and participated in several events, most recently in relation to the annual World Oceans Day on 8<sup>th</sup> June 2021, and in collaboration with diverse organizations such as the Colombian Ocean Commission (CCO), Agenda del Mar, the University of Antioquia, the CES University, and the Universidad del Norte, among others. On 3<sup>rd</sup> June 2021, the World Oceans Forum organized by the EPM Foundation and the CCO was held in Medellín and included the participation of CEMarin representatives and researchers in its panels of experts. If you missed the live transmission, you can find the recording online via the Facebook page of the EPM Foundation.

Our annual International Oceans Workshop, a set of academic activities and events focused on the collective construction of strategies to promote ocean conservation and socially sustainable and equitable environmental development in Colombia has been postponed until 3<sup>rd</sup> and 4<sup>th</sup> November 2021 due to the complicated current panorama in the country. However, we are enthusiastic about the robust agenda we have planned under the title "Voices and Leaderships: Community and Academia for the Oceans". You can find out more and register for the event at https://tallerinternacionaldelosoceanos.online/

Finally, we are excited to have continued developing our "Podcast for the Oceans", most recently with a special series of five episodes focused on the UN Decade of Ocean Science and which include the participation of a range of national and international experts. You can find our Podcast por los océanos (available in spanish only) on Spotify and Apple Podcasts, don't forget to subscribe for future updates.

