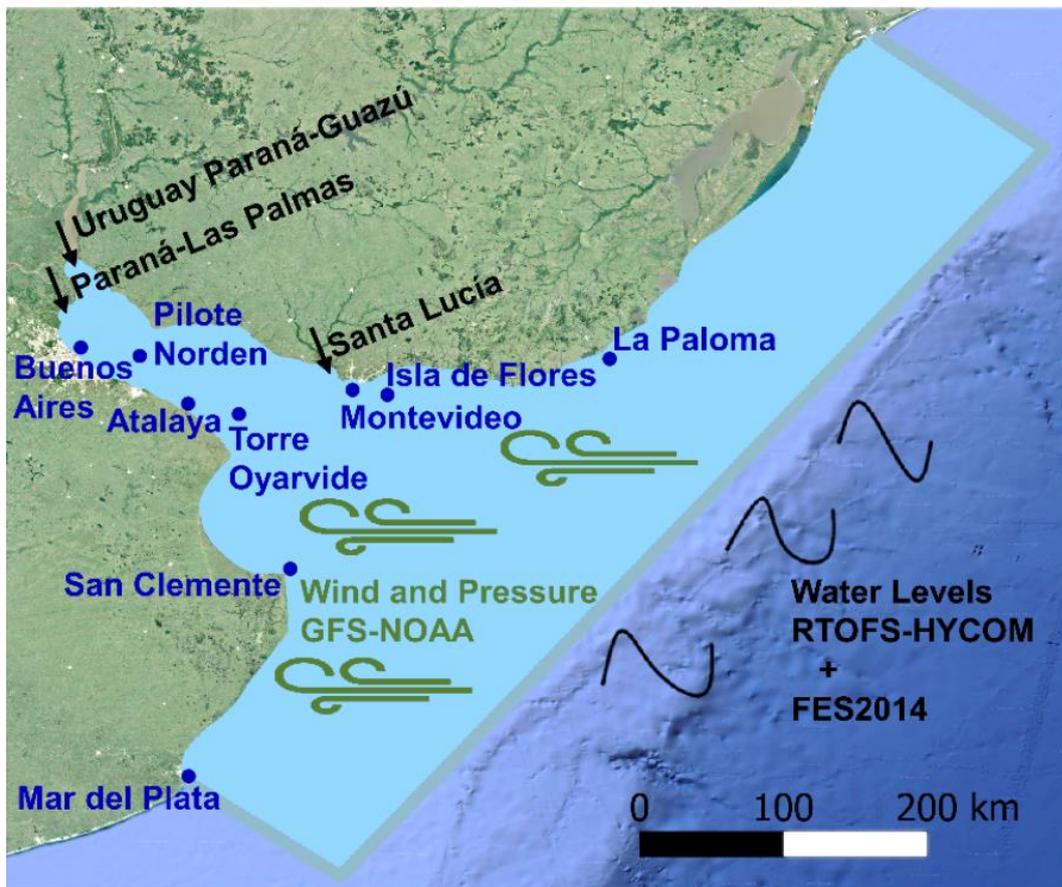


SISTEMA DE PRONÓSTICO DE LOS NIVELES DE MAREA EN TIEMPO REAL EN EL RÍO DE LA PLATA



TESIS DE MAESTRÍA EN MECÁNICA DE LOS FLUIDOS APLICADA
FEBRERO 2021

MAESTRANDO

ING. MARÍA BALLESTEROS ALMONACID

DIRECTORES

DRA. ING. MÓNICA FOSSATI

DR. ING. PABLO EZZATTI

Agradecimientos

Quiero agradecer a mi directora académica y tutora de esta tesis Mónica Fossati por incluirme en este proyecto. A Pablo Ezzatti por su rol mediador de co-tutor y por su punto de vista alternativo. Por sus propuestas durante las discusiones que llevaron a tomar las primeras decisiones agradezco a los miembros del IMFIA Pablo Santoro, Ismael Piedra-Cueva y Michelle Jackson. A los colaboradores del Instituto de Computación: Rodrigo Bayá, Renzo Marini y Danilo Espino. Agradezco también al equipo de soporte informático del IMFIA por tantas reparaciones y consultas resueltas. Gracias a mis profesores y compañeros de clase de Montevideo y Buenos Aires y en especial a los gurises con los que he ido compartiendo el día a día en la 236.

A mis padres por bancarme, a mi esposo por todo y a mis bebés por los *breaks* de bailes y risas.

Gracias a CSIC grupos y a la Fundación Julio Ricaldoni por sus fondos para este proyecto. Muy agradecida también a la posibilidad de contar con sala de lactancia en mi lugar de trabajo, para todos los que trabajan en avances de este tipo que benefician a toda la sociedad, muchos ánimos y a por más logros.

La investigación que da origen a los resultados presentados en el presente manuscrito recibió fondos de la Agencia Nacional de Investigación e Innovación bajo el código POS_NAC_2018_1_152257.

Índice

Resumen.....	9
I. Introducción	11
II. Estado del arte	15
II.1. Sistemas de pronóstico oceanográficos operacionales y sus componentes.....	15
II.1.1. Sistemas de observación de los océanos	16
II.1.1.1. Características generales	16
II.1.1.2. Satélites y oceanografía operacional	17
II.1.1.3. Sistemas de observación oceánica <i>In-Situ</i>	20
II.1.2. Modelos numéricos de la dinámica del océano	29
II.1.2.1. Ecuaciones gobernantes	29
II.1.2.2. Resolución numérica.....	31
II.1.2.3. Ecuaciones en 2 dimensiones. Ecuaciones promediadas en vertical.....	34
II.1.2.4. Modelos tridimensionales. La coordenada vertical	35
II.1.2.5. Modelos acoplados océano-atmósfera.....	37
II.1.3. Asimilación de datos.....	37
II.1.3.1. Introducción	37
II.1.3.2. Teorema de Bayes.....	40
II.1.3.3. Modelo state-space.....	41
II.1.3.4. Interpolación óptima (Optimal Interpolation)	42
II.1.3.5. Filtro de Kalman (Kalman Filter).....	43
II.1.3.6. Los métodos variacionales 3D-Var y 4D-Var	44
II.1.3.7. Filtro de Kalman Extendido (Extended Kalman Filter)	47
II.1.3.8. Filtro de Kalman por Ensamblés (Ensemble Kalman Filter)	48
II.1.3.9. Métodos no Gaussianos. Los Filtros de Partículas	50
II.1.3.10. La estimación de las matrices de covarianza de errores Q y R	51
II.1.4. Evaluación del pronóstico oceanográfico operacional.....	53
II.2. Pronósticos oceánicos globales	56
II.2.1. Descripción de los pronósticos disponibles.....	56
II.2.2. Detalles del modelo de pronóstico HYCOM	59
II.2.3. Detalles del modelo de pronóstico COPERNICUS.....	61
II.3. Sistemas de pronósticos costeros.....	62
II.3.1. Sistema Satbałtyk para el Mar Báltico.....	63
II.3.2. Estuario del Girona	66
II.3.3. Zona estuarina SSVBES	67

II.3.4. Río Tajo. Creación de un Módulo de Asimilación Secuencial para el MOHID mediante Filtro de Kalman Extendido	69
II.3.5. Mar Tirreno.....	70
II.3.6. Asimilación mediante 4D-LETKF para el modelo SMARA de la costa argentina	72
II.3.7. Modelo CROCO adaptado a simulaciones 2D en hindcast y pronóstico para el Río de la Plata	75
III. Generación del modelo de pronóstico.....	77
III.1. Características del área de estudio	77
III.2. Diseño del sistema de pronóstico.....	78
III.2.1. Configuración del modelo numérico	78
III.2.2. Condiciones de borde.....	79
III.3. Generación de la base de datos.....	81
III.3.1. Condiciones de borde oceánicas	82
III.3.2. Forzantes atmosféricos	84
III.3.3. Registros de los mareógrafos	84
III.4. Definición de la estructura del pronóstico	86
III.5. Implementación del sistema de pronóstico	87
III.6. Resultados del Sistema de pronóstico	90
IV. Evaluación	93
IV.1. Metodología	93
IV.2. Resultados	94
IV.2.1. Datos faltantes	94
IV.2.2. Comparación frente a datos de mareógrafos	96
IV.3. Análisis	116
V. Necesidades del sistema y recomendaciones.....	123
V.1. Necesidades del sistema.....	123
V.2. Introducción de la asimilación de datos en el Sistema de Pronóstico	124
VI. Conclusiones y líneas de trabajo futuro.....	131
VI.1. Conclusiones.....	131
VI.2. Publicaciones relativas a la tesis.....	132
VI.3. Futuras líneas de investigación.....	133
VII. Bibliografía	135
VIII. Anexo 1	141
IX. Anexo 2.....	149
X. Anexo 3.....	153
XI. Anexo 4.....	175

*A Martha Rodríguez Payret,
y quedé sola en nuestros almuerzos tempraneros.*

Resumen

El Río de la Plata y su Frente Marítimo es un cuerpo de agua complejo con características estuariales. Desde la mecánica de fluidos, configura un sistema dinámico y en cortos períodos de tiempo se producen variaciones muy significativas de las variables met-oceánicas, donde la marea meteorológica tiene gran incidencia.

La gran importancia del Río de la Plata para el Uruguay, y la región, potencia la necesidad de contar con una herramienta operacional (en tiempo real) que pronostique las mareas. Colaborando directamente en el abordaje de problemas asociados, por ejemplo, a la actividad portuaria, entre los cuales se destaca ayuda a la navegación. El sistema debe permitir apoyar la toma de decisiones en torno a diversas problemáticas: gestión y control del tráfico marítimo; eventuales derrames producidos por embarcaciones o vertidos de sustancias contaminantes; búsqueda y rescate de personas y/o materiales en accidentes marítimos; obras de ingeniería marítima en el dominio. Este tipo de herramientas se basan en modelos numéricos de flujo que permiten simular un número importante de procesos físicos. Dichas simulaciones involucran un conjunto de relaciones matemáticas y/o ecuaciones diferenciales con parámetros que deben ser calibrados en cada aplicación particular. El conocimiento del área de estudio alcanzado durante las investigaciones llevadas a cabo por el grupo de trabajo en los últimos años es la base que permite avanzar en una herramienta de pronóstico para el área.

En esta tesis se realiza, en primer lugar, una revisión profunda de las componentes de los sistemas de pronósticos operacionales (sistemas de observación, modelos numéricos, técnicas de asimilación de datos, etc.). También, se describen los modelos globales con pronósticos para el océano mundial y esfuerzos similares (pronósticos costeros y regionales). Por último, se resume el desarrollo de un sistema de base para la simulación operacional en el Río de la Plata y su Frente Marítimo a través de una primera herramienta a modo de prototipo para el pronóstico de los niveles. Se utiliza el modelo numérico TELEMAC-2D calibrado y validado en estudios anteriores e incorporando como condiciones del borde oceánico los pronósticos del modelo global *Real Time Ocean Forecast System*, RTOFS HYCOM (NOAA) y como forzantes atmosféricos los pronósticos del *Global Forecast System*, GFS (NOAA). Las salidas del modelo se comparan frente a los datos de mareógrafos disponibles en tiempo real y se difunden a través de la web <https://www.fing.edu.uy/imfia/pronostico-marea/> implementada para la publicación del sistema de pronóstico. Se presenta la evaluación para los pronósticos generados durante 2018 y 2019. Entre las posibles mejoras al prototipo generado se delinea la incorporación de la asimilación de datos medidos de mareógrafos a través de un esquema de filtro de Kalman por ensambles 4D-LETKF.

I. Introducción

El Río de la Plata y su Frente Marítimo es un cuerpo de agua complejo con características estuariales y marítimas localizado entre Argentina y Uruguay (Figura 1). En su orilla se encuentran las capitales de ambos países, así como los principales puertos, varias de las industrias más importantes y diversos centros turísticos.

Desde el punto de vista de la mecánica de fluidos, configura un sistema dinámico en el cual, en cortos períodos de tiempo, se producen modificaciones muy significativas de las variables meteorológicas, donde la marea meteorológica tiene gran incidencia en la dinámica de flujo (Santoro et al, 2013).

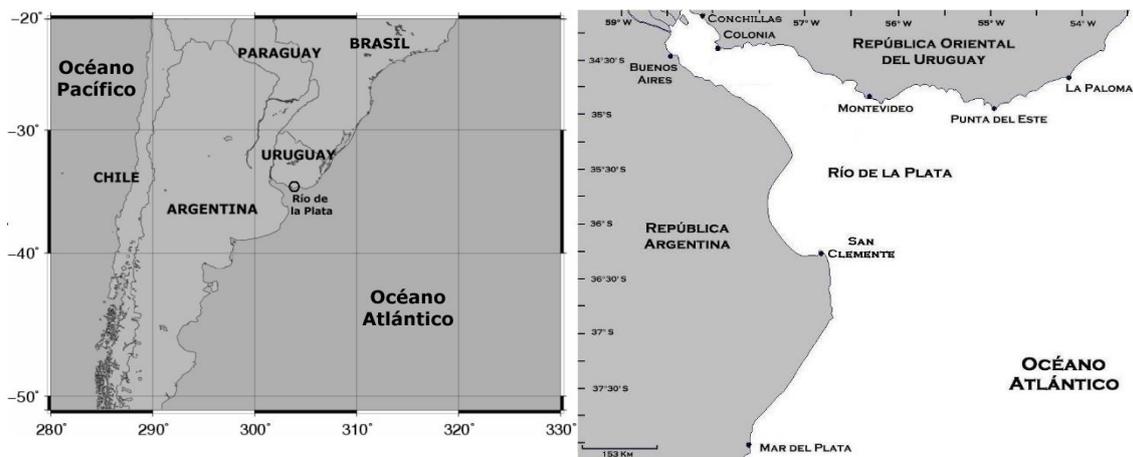


Figura 1 - Ubicación general del Río de la Plata. Extraído de Santoro, 2011.

Desde hace varios años en el Instituto de Mecánica de los Fluidos e Ingeniería Ambiental se trabaja en la simulación numérica de la hidrodinámica del Río de la Plata y su Frente Marítimo (Santoro et al., 2011). Gran parte de la experiencia acumulada se basa en la realización de simulaciones retrospectivas, las cuales han permitido estudiar la hidrodinámica del río (Santoro et al., 2013), la dinámica de frente salino y sus sedimentos (Fossati et al., 2014), y también evaluar exitosamente la respuesta hidrodinámica del sistema a distintas obras de ingeniería tales como dragados (Santoro et al., 2019), descargas de emisarios (Fossati et al., 2013), infraestructura portuaria, etc. Sin embargo, un sistema de pronóstico de las condiciones hidrodinámicas en tiempo real no existe todavía para la región.

Los sistemas de pronóstico de las condiciones hidrodinámicas se vinculan a problemas asociados a la actividad portuaria, entre los cuales se destaca la ayuda a la navegación. Los ejemplos de oceanografía operacional (en tiempo real) a nivel internacional incluyen modelos globales que proporcionan las condiciones del océano a nivel de gran escala y aplicaciones regionales o locales para áreas costeras. Estos sistemas sirven como apoyo en toma de decisiones en torno a diversas problemáticas en la medida en que ayudan a prever las corrientes y el movimiento de las masas de agua y así identificar y limitar las zonas de interés/riesgo. Algunos ejemplos son: gestión y control del tráfico marítimo; eventuales derrames producidos por embarcaciones o vertidos de sustancias contaminantes; búsqueda y rescate de personas y/o materiales en accidentes marítimos; obras de ingeniería marítima.

La información disponible en tiempo real en el Río de la Plata y su Frente Marítimo incluye la altura de marea medida en tres ubicaciones y difundida por la Administración Nacional de

Puertos (ANP), Uruguay y la información del Servicio de Hidrografía Naval (SHN), Argentina, que da un pronóstico de las ondas de tormenta con baja resolución para todo el Río de la Plata y publica niveles en hasta nueve estaciones. Como complemento, hay datos meteorológicos medidos y difundidos en tiempo real, y pronósticos diarios meteorológicos marinos. Si bien esta información es necesaria no es suficiente para realizar pronósticos en toda el área pues es extremadamente parcial. Así, los modelos numéricos aparecen como una herramienta adecuada para buscar soluciones con alto nivel de precisión en el espacio y tiempo.

El grupo de trabajo donde se desarrolla esta tesis tiene un carácter multidisciplinario e integra investigadores del Instituto de Mecánica de los Fluidos e Ingeniería Ambiental (IMFIA) y del Instituto de Computación, INCO. Durante la última década se han desarrollado estudios mediante la aplicación de diversos modelos numéricos que han permitido mejorar la comprensión de la dinámica de flujo en el Río de la Plata y su Frente Marítimo. Se incluyen el modelo en elementos finitos RMA integrado en vertical (Piedra-Cueva and Fossati, 2007; Fossati and Piedra-Cueva, 2008; Ezzatti et al., 2011), modelo hidrodinámico MOHID (Fossati and Piedra-Cueva, 2013) y sistema de modelos corrientes-oleaje-sedimentos TELEMAC (Santoro et al., 2017).

Como objetivo general de esta tesis se tiene la generación de un prototipo para un sistema de pronóstico operacional (en tiempo real) de niveles en el Río de la Plata. Este sistema toma como punto de partida el conocimiento del área de estudio de investigaciones anteriores y, para su definición, se deberán tomar decisiones en cuanto a su implementación: modelo numérico a utilizar, condiciones iniciales y de contorno, forzantes atmosféricos, etc.

Una propiedad de los modelos globales es servir de condiciones de contorno para aplicaciones regionales o costeras de forma que los pronósticos del modelo global se introducen en el borde del modelo local de mayor resolución. De este modo se aprovecha la capacidad de pronóstico del modelo global para una zona en la que se hace necesario un incremento de resolución. En relación a las condiciones iniciales para cada ciclo de pronóstico, la asimilación de datos busca mejorar la calidad de los pronósticos mediante la incorporación de mediciones durante la construcción de las condiciones iniciales de los modelos numéricos.

Los objetivos específicos de este trabajo de tesis son:

- 1) Revisar el estado del arte profundizando en las características y componentes de los sistemas de pronóstico oceánicos. Se debe incluir una introducción a los esquemas de asimilación de datos y también herramientas operacionales que publiquen en tiempo casi real las condiciones del océano global a nivel mundial y para zonas regionales o costeras. Esta exploración permitirá tanto establecer los requerimientos de los sistemas de pronóstico como la aplicabilidad de otros casos al sistema del Río de la Plata.
- 2) Diseño y desarrollo de un prototipo de herramienta de pronóstico de los niveles en Río de la Plata y su Frente Marítimo determinando condiciones de borde e iniciales apropiadas. Evaluación de la calidad del prototipo generado.
- 3) Delinear posibles mejoras al sistema de pronóstico y futuras líneas de investigación.

El presente informe se divide en seis apartados, comenzando por este apartado de introducción. En el apartado II se desarrolla el estado del arte que contiene una descripción de los sistemas de pronóstico operacionales y sus componentes, atendiendo al marco teórico y dedicando mayor detalle a una introducción a la asimilación de datos. A continuación se presentan los

modelos globales incluyendo una reseña de varios antecedentes internacionales de sistemas de pronósticos costeros. En el apartado III se describe la generación del modelo de pronóstico para los niveles de marea en el Río de la Plata y su Frente Marítimo. A continuación se presenta la evaluación del prototipo de pronóstico generado en el apartado IV. En el apartado V se enmarca la situación actual del prototipo generado y se realizan propuestas para futuras líneas de investigación. Finalmente se presentan conclusiones en el apartado VI.

II. Estado del arte

En este apartado se presentan, en primer lugar, los sistemas de pronóstico oceanográficos operacionales, describiendo cada uno de sus componentes y tomando como base la publicación de Schiller and Brassington (2011), *Operational Oceanography in the 21st Century*. A continuación, se introducen los modelos globales que estudian las condiciones del océano a nivel planetario. Se incluyen también en este apartado de estado del arte, estudios similares relativos a sistemas costeros de pronóstico operacionales.

II.1. Sistemas de pronóstico oceanográficos operacionales y sus componentes

La oceanografía operacional se refiere a la previsión de las condiciones del estado de los océanos a nivel de gran escala. El proyecto GODAE, Experimento Mundial de Asimilación de Datos Oceánicos, se creó en 1997 con el objetivo de coordinar e integrar la investigación asociada a sistemas globales y regionales de predicción y análisis de los océanos (<https://www.godae-oceanview.org/about/>). Los objetivos específicos de GODAE son: a) aplicar el estado del arte de modelos oceánicos y métodos de asimilación para producir pronósticos del océano abierto a corto plazo, condiciones de borde para extender la previsibilidad de subsistemas costeros y regionales y condiciones iniciales para los modelos de predicción climática; b) proporcionar análisis mundiales de la dinámica oceánica para mejorar su comprensión y mejorar la previsibilidad de su variabilidad, constituyendo la base para mejorar el diseño y la efectividad de un sistema de observación global de los océanos. En la publicación surgida del proyecto GODAE (Schiller and Brassington, 2011), se recoge el avance logrado en la materia hasta el momento, así como los objetivos a alcanzar en la próxima década. La más reciente publicación del GODAE (Chassignet et al., 2018) actualiza el avance alcanzado en la materia.

Los sistemas de pronóstico oceanográficos operacionales desarrollados en la actualidad siguen una estructura cíclica que implica: el manejo de los datos observacionales recientes, la realización de una fusión modelo-datos (asimilación de datos) consistente en la incorporación de mediciones durante la construcción de las condiciones iniciales; la realización de un modelo de pronóstico para generar los productos que incluyen la estimación del estado del océano, la realización de un diagnóstico del mismo y la estimación del error.

Este procedimiento se repite en un determinado cronograma regular o para ciertos eventos particulares. Todos estos elementos hacen que el sistema sea complejo, en el cual cada componente tiene aspectos específicos de diseño que impactan en el desempeño, el volumen de recursos necesarios y por ende en el costo total del sistema completo. En particular, la componente de modelación numérica en un sistema de pronóstico requiere un gran esfuerzo en sí mismo para proveer resultados de calidad.

La precisión de las simulaciones mediante modelos numéricos depende de la disponibilidad e idoneidad (precisión, resolución y duración) de las observaciones y de los modelos hidrológicos, oceánicos y meteorológicos que se utilizan para iniciar, forzar y realizar los cálculos. La calidad del pronóstico oceánico dependerá principalmente de la capacidad del modelo numérico oceánico para representar fielmente la física y la dinámica del océano. La modelación numérica

es una de las etapas en la cual se requiere inversión sostenida para infraestructura y organización, como acceso a plataformas de hardware que permitan alcanzar tiempos de cómputo razonables, mantenimiento del software e intercambio de datos.

La oceanografía operacional es la actividad rutinaria de medidas de los mares, océanos y atmósfera, sistemática y a largo plazo y su rápida interpretación y diseminación. El pronóstico oceánico se basa en la recolección en tiempo real de observaciones y su procesamiento mediante la rápida transmisión de los datos observados hacia los centros de procesamiento de datos, los cuales llevan a cabo el control de calidad y proveen los datos a los centros de pronóstico. Allí, computadoras de gran capacidad, utilizando modelos numéricos, asimilan las observaciones en los modelos para mejorar y generar condiciones iniciales para el pronóstico de los océanos; con pronósticos típicamente de 10 días en adelante (Chassignet et al., 2018). Como parte del control de calidad se establecen niveles de servicio hacia los usuarios y se vigila su cumplimiento. Con el propósito de facilitar el acceso a sus productos, los centros de pronóstico han desarrollado y mejorado las tecnologías de información y técnicas de visualización. Actualmente, los centros de pronóstico del océano cubren ambientes marinos desde escala global a costera incluyendo propiedades físicas y biogeoquímicas. El análisis y pronóstico de ecosistemas son áreas activas de investigación.

Los pronósticos de los centros de pronóstico son utilizados para generar productos de valor añadido como pueden ser rutas de pesca optimizadas, alertas de inundaciones costeras, predicción de floraciones de algas nocivas, etc. Debido a que las condiciones del océano están continuamente cambiando, el pronóstico final se debe distribuir de forma rápida, siendo la libre disponibilidad de los productos una de las insistencias del GODAE.

En las siguientes secciones se da un lineamiento sobre las distintas componentes de un sistema de pronóstico operacional: observaciones, modelos numéricos, asimilación de datos, evaluación y difusión de productos.

II.1.1. Sistemas de observación de los océanos

II.1.1.1. Características generales

Los sistemas de observación oceánica consisten en técnicas in situ y remotas por satélite para detectar, rastrear y predecir los cambios en los procesos físicos, químicos, geológicos y biológicos.

En el contexto de los sistemas de oceanografía operacional las redes de observación pueden dar servicio a dos tipos de uso (Canas, 2009):

- Validación de los resultados de modelos numéricos, dando la oportunidad de ensayar, validar y refinar los modelos.
- Establecimiento de las condiciones iniciales y una corrección regular de las variables de estado del modelo a través de la asimilación de datos.

Se requiere de un control de calidad efectivo que incluya un conjunto preestablecido y estandarizado de procedimientos, que determine de manera clara la validez de los datos previo a su utilización tanto para validación como para asimilación. La efectividad del control de calidad

de datos oceánicos es un problema complejo. Las observaciones son imperfectas y propensas a incluir errores.

En la actualidad, hay pocos estándares acordados para el control de calidad de datos en tiempo real y muy pocos casos donde los procedimientos y resultados de los centros oceanográficos han sido comparados (Schiller and Brassington, 2011). Con la innovación en instrumentos de medida deben desarrollarse nuevos modelos de error de observación para el control automatizado de la calidad de los nuevos tipos de datos, y los modelos de error existentes deben actualizarse para detectar y corregir nuevos modos de fallo del instrumento.

Normalmente, cuando se desarrolla un sistema de pronóstico, la disponibilidad de datos es limitada. Pero, en ocasiones, estos proyectos pueden llevar a justificar la incorporación de nuevos puntos de medida o la necesidad de instrumentos que permitan medir variables que no estaban disponibles. Dado el caso, se deberá estudiar el diseño del sistema de observación. A la hora de diseñar un nuevo sistema de observación lo primero que se debe saber es qué es lo que el sistema de observación pretende supervisar. Es poco probable que un sistema de observación optimizado para monitorear un aspecto específico de la circulación oceánica sea óptimo para monitorear todos los demás aspectos de la circulación. A su vez, no se debe perder de vista, que la cantidad y tipos de instrumentos que pueden considerarse durante el diseño de un nuevo sistema estarán determinados también por el presupuesto económico que se disponga. Una vez establecido el sistema de observación, es común que un mismo sistema sea utilizado por múltiples usuarios explotando diferentes métodos.

La densidad de observaciones requeridas (y su frecuencia) para monitorear un proceso dado, está dictada en gran medida por las escalas de longitud a las que deja de haber correlación de los campos que deben observarse (escalas de des-correlación de tiempo para la frecuencia). Esta propiedad determina cuán distantes pueden ser las observaciones sin que se pierdan las características importantes.

Otra característica relevante en el diseño de un sistema de observación es la co-variabilidad del océano: puede haber ubicaciones o cantidades que sean particularmente indicativas de todo el sistema que se va a observar. Esto ocurre, por ejemplo, en el estudio de los eventos de El Niño con el Índice de Oscilación Sur (SOI, por sus siglas en inglés). Este índice se calcula a partir de las variaciones en la diferencia de presión del aire entre Tahití y Darwin. Los períodos de SOI negativo sostenido generalmente corresponden a eventos de El Niño (Schiller and Brassington, 2011).

II.1.1.2. Satélites y oceanografía operacional

La oceanografía operacional está muy ligada al desarrollo de la oceanografía satelital. La capacidad de observar el océano global en tiempo casi real con alta resolución en el espacio y en el tiempo es, de hecho, un prerrequisito para el desarrollo de la oceanografía operacional global y sus aplicaciones. Existen dos tipos principales de técnicas satelitales para observar el océano. Las técnicas pasivas, que miden la radiación natural emitida por el mar o por la radiación solar reflejada. Las técnicas activas o de radar, que envían una señal y miden la señal recibida después de su reflexión en la superficie del mar. En ambos casos se debe tener en cuenta el paso por la atmósfera para aislar la señal de la superficie del mar (Schiller and Brassington, 2011).

Previo a la amplia disponibilidad de los satélites en los años 80, los oceanógrafos no tenían muchos datos. La aparición de la altimetría satelital (Figura 2) permitió el desarrollo de la asimilación de datos oceánicos y la oceanografía operacional global. Pero las observaciones del altímetro satelital se deben complementar con observaciones *In-Situ*, tanto para la calibración de las transformaciones de las medidas realizadas por el satélite cuando no se observan directamente variables geofísicas, como para monitorear la estabilidad de las misiones de larga duración.

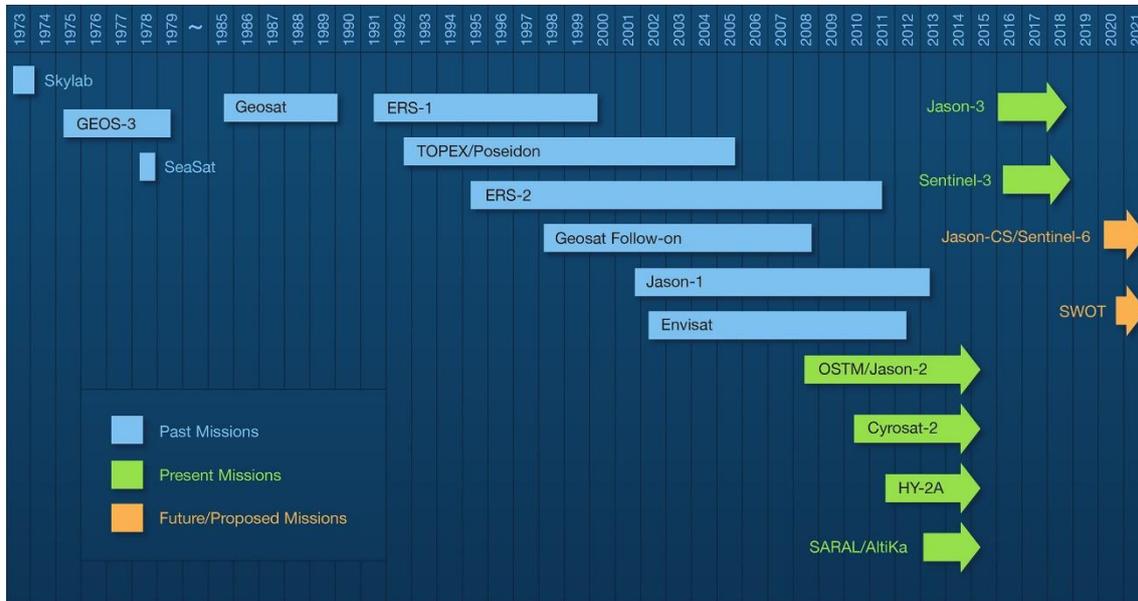


Figura 2 – Misiones de altímetros satelitales. Extraído de https://podaac.jpl.nasa.gov/Altimetric_Data_Information/

Los satélites proporcionan un servicio a largo plazo, continuo, global, de alta resolución en espacio y tiempo para los parámetros clave del océano: nivel del mar y circulación del océano, temperatura de la superficie del mar, color del océano, hielo marino, olas y vientos; y en los últimos tiempos también permiten conocer la salinidad de la superficie del mar (Schiller and Brassington, 2011).

En relación a la temperatura, el Grupo para la Temperatura de la Superficie del Mar de Alta Resolución (GHRST) (Donlon et al., 2009) se inició en 2001 y ha supervisado la introducción de productos SST (temperatura de la superficie del mar) estandarizados y verificados para múltiples usos, incluidos el análisis y la predicción de los océanos. El proyecto patrocinado por GODAE ha resultado en una red coordinada de centros que difunden datos de SST en tiempo real en un formato común a estándares acordados de una amplia gama de instrumentos de microondas e infrarrojos en órbitas polares y satélites geoestacionarios (Chassignet et al., 2018). Actualmente se dispone de datos en tiempo casi real de dos radiómetros, los datos de AVHRR (*Advanced Very High-Resolution Radiometer*) que son proporcionados por el NOAA y los datos del AMSR2 (*Advanced Microwave Scanning Radiometer*) a través de la web de la NASA.

En relación a la salinidad, los radiómetros basados en satélites de microondas de banda L proporcionan, por primera vez, mediciones de la salinidad de la superficie con un orden de precisión de 0.5 psu para promedios de 10 días en 50 km cuadrados y se espera que tengan un impacto positivo en los sistemas de pronóstico del océano (Le Traon et al., 2015).

La altimetría satelital condiciona la calidad de las predicciones oceánicas. Las medidas de esta plataforma se realizan a nivel global de forma que no dirigen su interés a ubicaciones concretas

sino que se efectúan siguiendo el recorrido del satélite y cubriendo todas las latitudes. El centro CMEMS proporciona en tiempo casi real los datos de las transectas medidas por los altímetros satelitales con un procesamiento básico (GLOBAL OCEAN ALONG-TRACK L3 SEA SURFACE HEIGHTS NRT) a través de su web <https://marine.copernicus.eu/services-portfolio/access-to-products/>. En la Tabla 1 se resumen las misiones llevadas a cabo por altímetros satelitales.

Altímetro	Fase	Fechas	Ciclos	Pases	Registros
GEOSAT	A	31 Mar 1985 - 30 Sep 1986	001 - 025	61152	104395824
	B	08 Nov 1986 - 30 Dec 1989	026 - 093		
	D	31 Mar 1985 - 30 Sep 1986	001 - 025		
ERS-1	A	01 Aug 1991 - 14 Dec 1991	001 - 046	47890	83644555
	B	14 Dec 1991 - 25 Mar 1992	047 - 081		
	C	14 Apr 1992 - 20 Dec 1993	083 - 101		
	D	24 Dec 1993 - 10 Apr 1994	103 - 138		
	E	10 Apr 1994 - 28 Sep 1994	139 - 140		
	F	28 Sep 1994 - 21 Mar 1995	141 - 143		
	G	24 Mar 1995 - 02 Jun 1996	144 - 156		
TOPEX	A	25 Sep 1992 - 11 Aug 2002	001 - 364	111960	264208509
	B	20 Sep 2002 - 08 Oct 2005	369 - 481		
	N	11 Aug 2002 - 20 Sep 2002	365 - 368		
POSEIDON	A	01 Oct 1992 - 12 Jul 2002	001 - 361	7472	15718917
ERS-2	A	29 Apr 1995 - 04 Jul 2011	000 - 169	148130	168866998
GFO-1	A	07 Jan 2000 - 17 Sep 2008	037 - 223	82623	150249181
JASON-1	A	15 Jan 2002 - 26 Jan 2009	001 - 260	102175	279080066
	B	10 Feb 2009 - 03 Mar 2012	262 - 374		
	C	07 May 2012 - 21 Jun 2013	382 - 425		
ENVISAT1	B	14 May 2002 - 22 Oct 2010	006 - 094	99633	259990495
	C	26 Oct 2010 - 08 Apr 2012	095 - 113		
JASON-2	A	04 Jul 2008 - 02 Oct 2016	000 - 303	98578	315107265
	B	13 Oct 2016 - 17 May 2017	305 - 327		
	C	11 Jul 2017 - 18 Jul 2018	332 - 355		
	D	25 Jul 2018 - 01 Oct 2019	356 - 383		
CRYOSAT2	A	14 Jul 2010 - 03 Sep 2020	004 - 135	105585	273751564
SARAL	A	14 Mar 2013 - 04 Jul 2016	001 - 035	78016	218850984
	B	04 Jul 2016 - 03 Sep 2020	036 - 079		
JASON-3	A	12 Feb 2016 - 03 Sep 2020	000 - 168	41801	138565610
S3A	A	01 Mar 2016 - 03 Sep 2020	001 - 062	46898	95804926
Total				1031913	2368234894

Tabla 1 - Base de datos de los altímetros satelitales. Extraído de <http://rads.tudelft.nl/rads/status.shtml> con fecha 03 de septiembre de 2020.

Como ya se mencionó, la calidad de las predicciones oceánicas de mesoescala en modelos de análisis y pronóstico está fuertemente condicionada a la disponibilidad de datos de altímetro de múltiples satélites (Chassignet et al., 2018). Es comúnmente aceptado que un análisis y

pronóstico en tiempo real requiere de al menos tres o preferiblemente cuatro altímetros orbitando simultáneamente en rutas apropiadamente seleccionadas (Le Traon, 2013).

Como referencia para las demás misiones y particularmente para el monitoreo climático, se precisa de un sistema de altímetros como las series Jason, que sea preciso y a largo plazo (Chassignet et al., 2018).

De las misiones previstas para la próxima década se espera una satisfactoria constelación de altímetros de esencial contribución a la oceanografía operacional que acompañe la creciente resolución de los modelos (Chassignet et al., 2018). Se destaca para el año 2022 el lanzamiento de la misión del altímetro de banda ancha SWOT (*Surface Water and Ocean Topography*) que será capaz de cubrir los océanos de la Tierra, grandes lagos, ríos y embalses al menos dos veces cada 21 días (<https://swot.jpl.nasa.gov/>).

II.1.1.3. Sistemas de observación oceánica *In-Situ*

Los sistemas de observación in situ comprenden tanto sistemas eulerianos (basados en ubicaciones fijas) como lagrangianos (cuya ubicación varía con el tiempo). A continuación se hace una breve introducción a las distintas plataformas de observación de variables físicas del océano: mareógrafos, buques de observación voluntaria, red XBT, buques de investigación, boyas a la deriva, boyas fijas (*moored bouys*), flotadores Argo, planeadores submarinos (*Gliders*) y radares de alta frecuencia. La libre disponibilidad de datos es una de las insistencias del GODAE y existen varias webs desde las que se puede tener acceso a gran cantidad de datos de medidas *in-situ*. De esta manera, entre los productos que pone a disposición el centro CMEMS a través de su web, se publican datos de observaciones *in-situ* en tiempo casi real y a nivel global de varias fuentes que incluyen: amarras, flotadores ARGO, XBTs, CTD. Aun así, durante esta investigación, se pudo comprobar que todavía hay observaciones que no han alcanzado este nivel de accesibilidad.

Mareógrafos

Los mareógrafos son instrumentos que miden de manera periódica los cambios del nivel del mar teniendo como nivel de referencia a un nivel en tierra. La Figura 3 ilustra su sistema de funcionamiento a partir de los registros de un flotador. Los mareógrafos constituyen las mediciones de tecnología más sencilla y a largo plazo teniendo como una de sus principales ventajas su fácil mantenimiento. A diferencia de otros sistemas de observación, los mareógrafos miden únicamente un parámetro físico (nivel) y en general se suelen ubicar en zonas costeras de fácil acceso donde se concentra la mayoría de estos aparatos.

De forma de coordinar las mediciones de mareógrafos y proporcionar una visión global de las medidas se creó el sistema GLOSS (*Global Sea Level Observing System*, <https://www.gloss-sealevel.org/>) que maneja una red de más de 290 estaciones mareográficas con cobertura mundial incluyendo datos provenientes de la contribución de más de 90 países. En la Figura 4 se puede ver en color verde las estaciones con medidas disponibles en tiempo real para octubre de 2020.

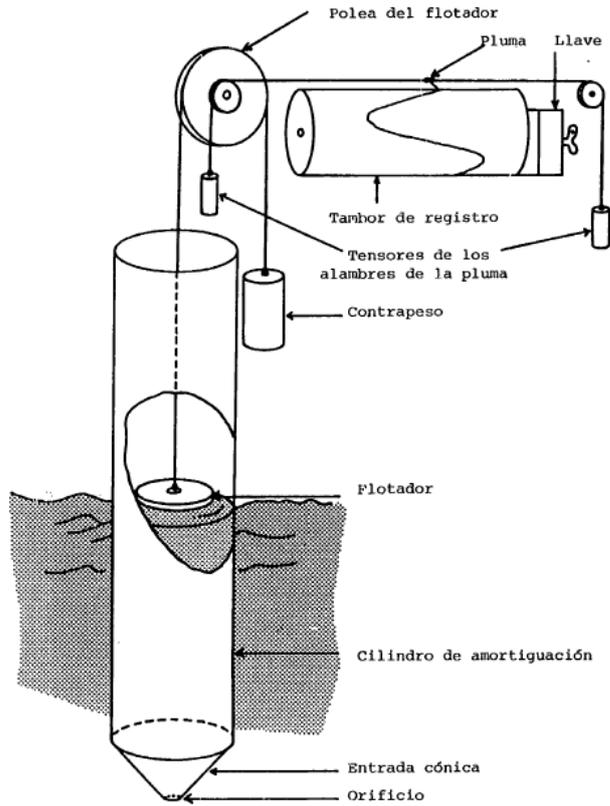


Figura 3 – Esquema de funcionamiento de un mareógrafo a flotador. Extraído de UNESCO (1985).

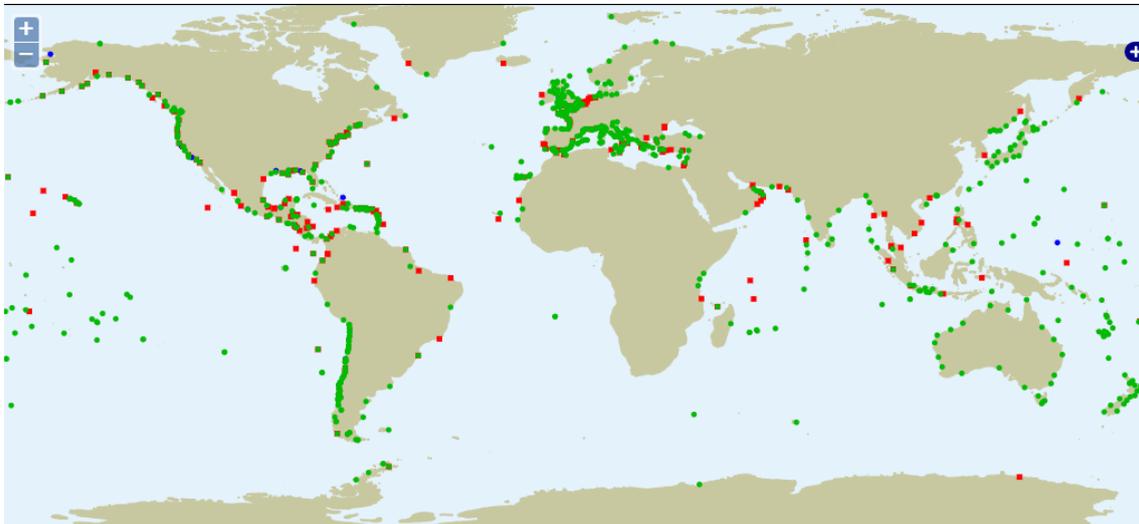


Figura 4 – Mareógrafos con medidas disponibles en la red GLOSS. Extraído de <http://www.ioc-sealevelmonitoring.org/map.php> en octubre 2020.

Buques de observación voluntaria

Las observaciones provenientes de buques de observación voluntaria (VOS por sus siglas en inglés) son el aporte que se obtiene de los diferentes buques durante su navegación debida a

distintas actividades. En el Esquema VOS los buques se categorizan según su función (buque seleccionado, complementario o auxiliar) y esta categorización determina el instrumental de medida para el registro de distintas variables meteorológicas con el que se equipará al buque. En general los reportes se realizan cada 3 ó 6 horas y como variables medidas se tienen el viento en superficie, temperatura del aire, humedad, temperatura de la superficie del mar (SST), presión atmosférica al nivel del mar, etc. También se reportan variables estimadas de manera visual entre las que se encuentran la nubosidad, el oleaje o la visibilidad.

Las medidas obtenidas por este sistema dan una amplia cobertura de parámetros marinos en superficie con alta resolución a lo largo de las rutas que repetidamente realizan los barcos pero, a diferencia de otros sistemas en los que el propósito principal es obtener medidas oceánicas, los buques de observación voluntaria no siempre obtienen los datos donde son requeridos. Otras desventajas que se atribuye a este sistema son que no se generan medidas en profundidad y que el buque no se detiene. En la Figura 5 se presenta el mapa de las observaciones de esta plataforma para octubre de 2020, se distingue entre los buques que utilizan TAC (códigos alfanuméricos tradicionales) o TDC (códigos basados en tablas), ya que esta categorización es relevante a la hora de procesar los datos.

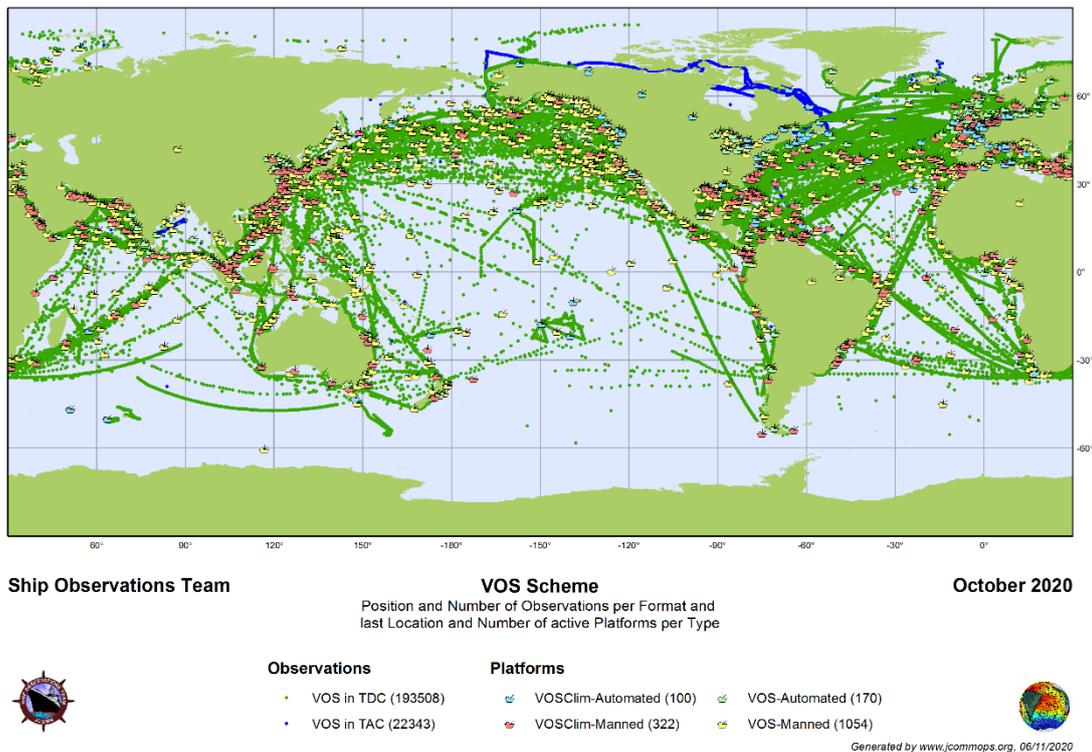
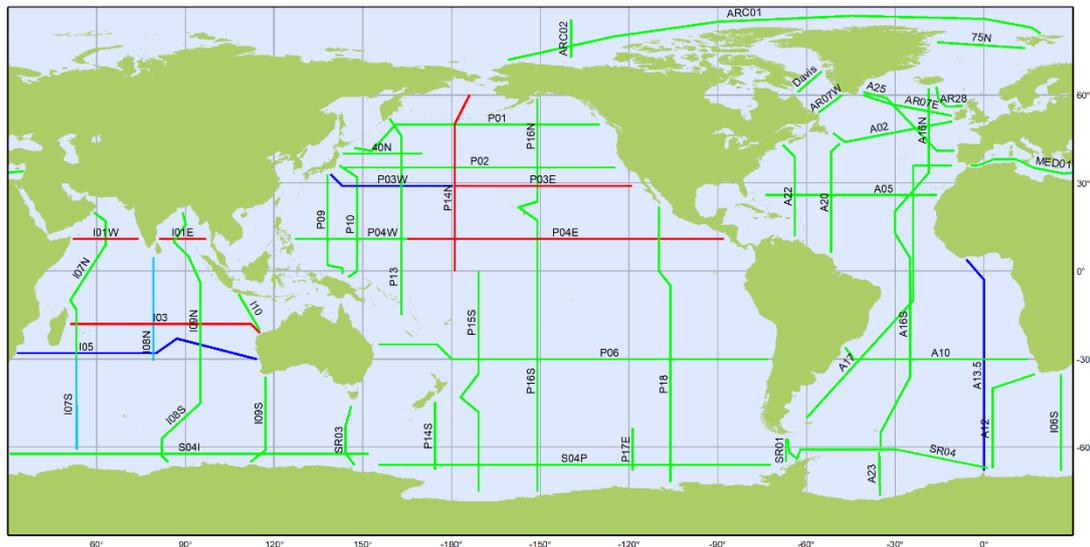


Figura 5 - Mapa de las observaciones provenientes de buques de observación voluntaria en octubre de 2020. Extraído de <http://www.ocean-ops.org/sot/>.

Red XBT

La red XBT tiene como objetivo medir la temperatura en el océano a lo largo de transectas fijas en todas las cuencas oceánicas. Estas transectas aparecen en la Figura 6 y han sido establecidas para atender los requerimientos de las comunidades internacionales científicas y operacionales

El programa de hidrografía repetida GO-SHIP (*Global Ocean Ship-based Hydrographic Investigations Program*, <http://www.go-ship.org/>) reúne los intereses de científicos y usuarios para desarrollar una red global de observación del clima oceánico y a través de acuerdos internacionales se monitorean de manera organizada las transectas de referencia que aparecen en la Figura 7 prestando atención a oceanografía física, ciclo del carbono, biogeoquímica marina y ecosistemas.



GO-SHIP Status of 2012-2023 Survey (55 Core Lines) December 2019

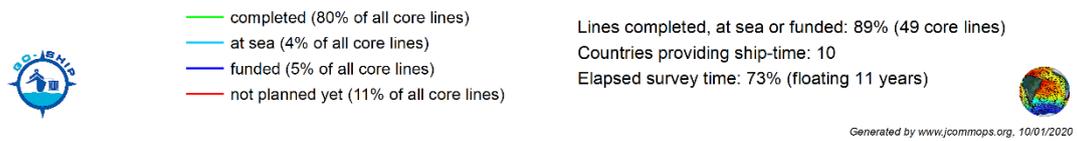


Figura 7 – Programa de hidrografía de a bordo repetida GO-SHIP. Estado de las mediciones en las transectas propuestas para la década 2012-2023. Extraído de https://www.go-ship.org/RefSecs/goship_ref_secs.html

Boyas a la deriva

Las boyas a la deriva son instrumentos que tras ser lanzados al mar son arrastrados por las corrientes de superficie y registran medidas durante su recorrido con una vida útil de unos 18 meses. Las boyas son rastreadas de manera satelital lo que permite determinar corrientes en superficie (Figura 8, panel izquierdo). Además, de manera general, se utilizan para medir la temperatura de la superficie del mar (SST) y la presión del aire a nivel de la superficie del mar y pueden contar con medidores de salinidad o viento. Las ventajas de este método son su cobertura global incluyendo zonas remotas, muestreo rápido y de bajo costo con tecnología robusta. Pero debido a su sistema de adquisición se obtienen únicamente variables de superficie y hay zonas que quedan sin muestrear. Un problema asociado a este método y que en general afecta a los instrumentos que permanecen largas temporadas en el océano es el *biofouling* o incrustación por colonización de organismos acuáticos. La independencia de estas medidas *In-situ* permite utilizarlas para validar y corregir medidas realizadas por satélites (por ejemplo, medidas de temperatura de la superficie del mar). En la Figura 8 (panel derecho) se presenta el mapa de las boyas a la deriva en operación en octubre de 2020.

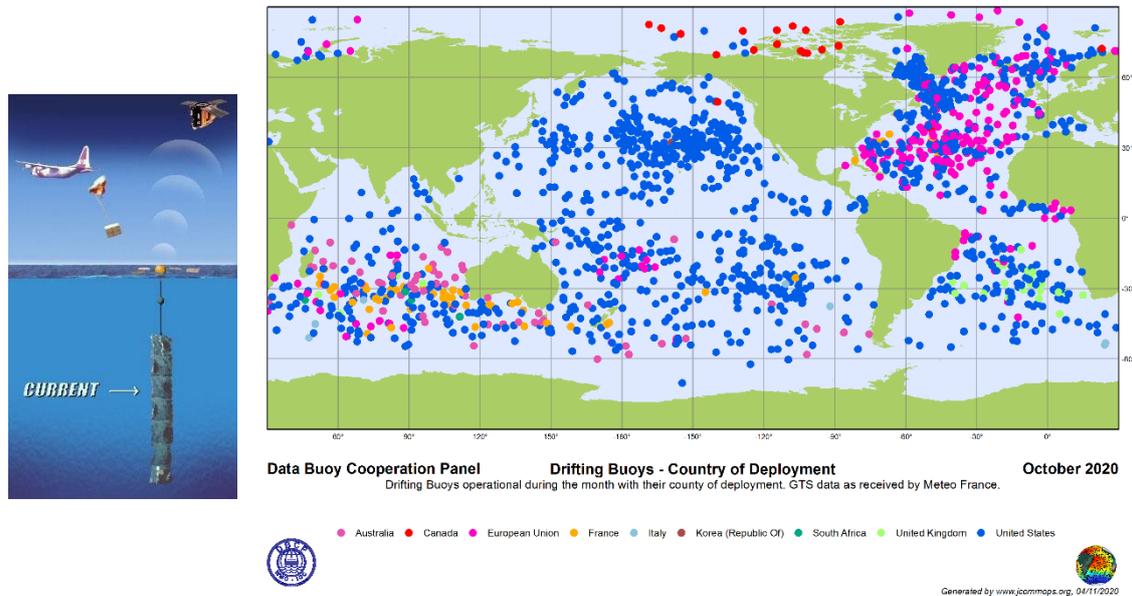


Figura 8 – Ilustración para representar el funcionamiento de las boyas a la deriva (panel izquierdo). Extraído de <https://cutterhealy.wordpress.com/>. Mapa de las boyas a la deriva en operación en octubre de 2020 (panel derecho). Extraído de <http://www.ocean-ops.org/dbcp/>.

Boyas fijas (moored bouys)

Las boyas fijas (Figura 9, panel izquierdo), a diferencia de las boyas a la deriva, son plataformas grandes y costosas que se encuentran ancladas en una posición fija desde la que transmiten los datos observados y donde generalmente permanecen varios años. Las variables que son capaces de medir incluyen parámetros en superficie como son el viento, temperatura de la superficie del mar y del aire, salinidad, así como la temperatura subsuperficial hasta unos 500 m de profundidad. Sus largas campañas de medida permiten disponer de series temporales que identifican cambios en el océano desde minutos a años pero también en la interacción que se da en la interfaz océano-atmósfera. Entre los problemas que afectan a este tipo de plataformas se incluyen la ocurrencia de vandalismo intencional o el daño accidental, y, en relación a los datos, que por su naturaleza euleriana, para poder discernir entre efectos locales y advectivos se requiere disponer de un conjunto de datos espaciales complementarios. En la Figura 9 (panel derecho) se presenta el mapa de las boyas y otras plataformas fijas en operación en octubre de 2020.

Flotadores Argo

Los instrumentos que se utilizan en el programa internacional Argo (<https://argo.ucsd.edu/>) siguen el ciclo representado en la Figura 10. Tienen una primera etapa de inmersión hasta los 1,000 m y a continuación van a la deriva durante unos 9 días. Seguidamente bajan hasta los 2,000 m o más y vuelven a la superficie donde transmiten la salinidad y temperatura registradas durante su recorrido, así como la ruta seguida. El ciclo se completa en unos 10 días. La mayor ventaja de esta plataforma es el monitoreo de zonas profundas del océano incluyendo zonas

remotas a un relativo bajo costo. En la Figura 11 se presenta el mapa de flotadores Argo en operación en octubre de 2020.

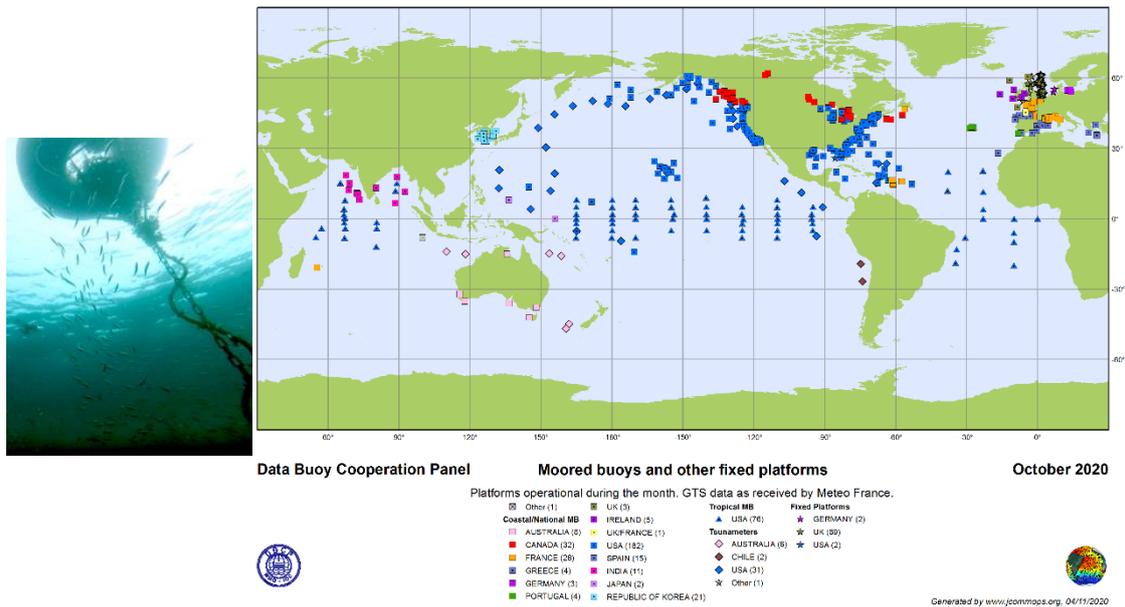


Figura 9 – Fotografía de una boya fija (panel izquierdo). Extraído de <https://www.mbari.org/mooring-maintenance-from-divers-perspective/>. Mapa de las boyas y otras plataformas fijas en operación en octubre de 2020 (panel derecho). Extraído de <http://www.ocean-ops.org/dbcp/>.

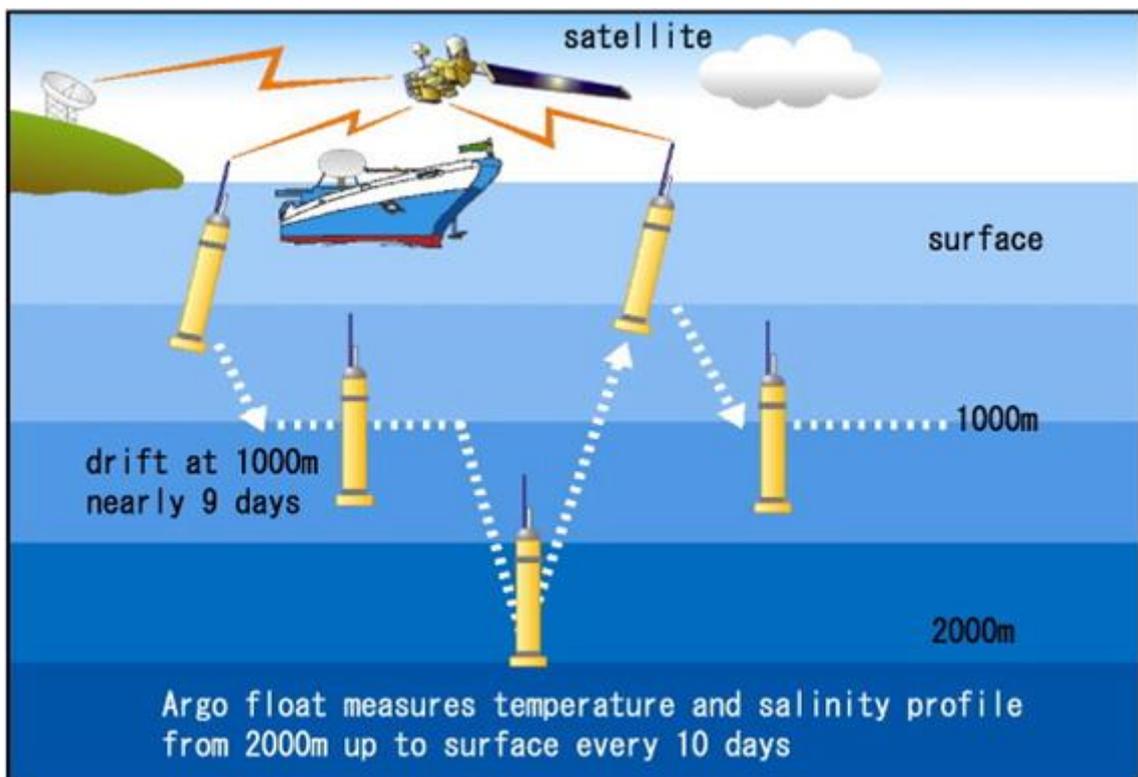


Figura 10 – Ilustración del ciclo realizado por los flotadores Argo. Extraído de https://www.jamstec.go.jp/e/about/press_release/20160129/

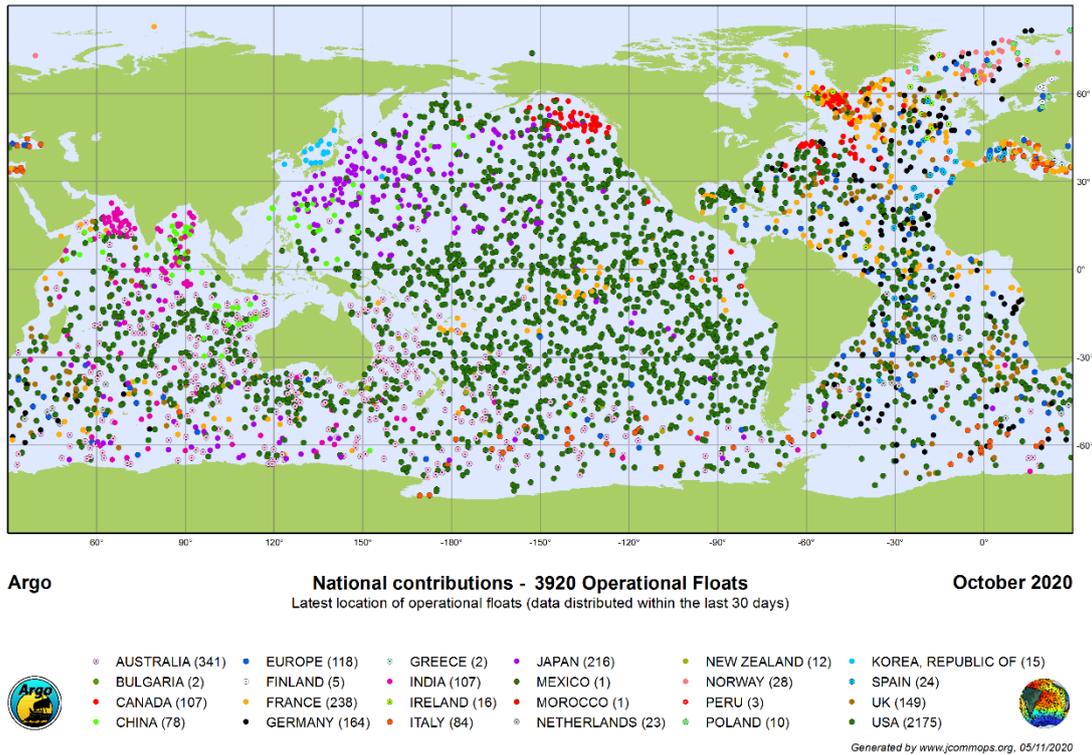


Figura 11 - Mapa de los flotadores Argo en operación en octubre de 2020. Extraído de <http://www.oceanops.org/argo/>.

Planeadores submarinos (*Gliders*)

Los planeadores submarinos (*Gliders*, Figura 12, <https://www.oceangliders.org/>) son vehículos submarinos autónomos de propulsión a muy bajo costo que se equipan con una variedad de sensores para medir variables del océano durante la ruta que han sido programados a seguir. A su salida a superficie envían los datos registrados y su recorrido. El sistema utilizado para desplazarse consiste esencialmente en cambiar su flotabilidad para que sus alas actúen convirtiendo movimiento vertical en horizontal, llegando a profundidades de unos 1,000 m. Las variables que registren dependerán de los sensores con que sean equipados, generalmente sensor CTD y también variables biogeoquímicas como el oxígeno disuelto. En la Figura 13 se presenta el mapa de planeadores submarinos en activo en octubre de 2020.

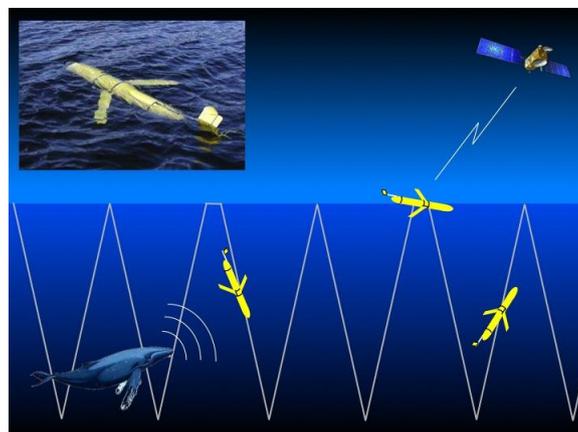


Figura 12 – Ilustración para esquematizar el funcionamiento de los planeadores submarinos (*Gliders*). Extraído de <http://dcs.whoi.edu/>

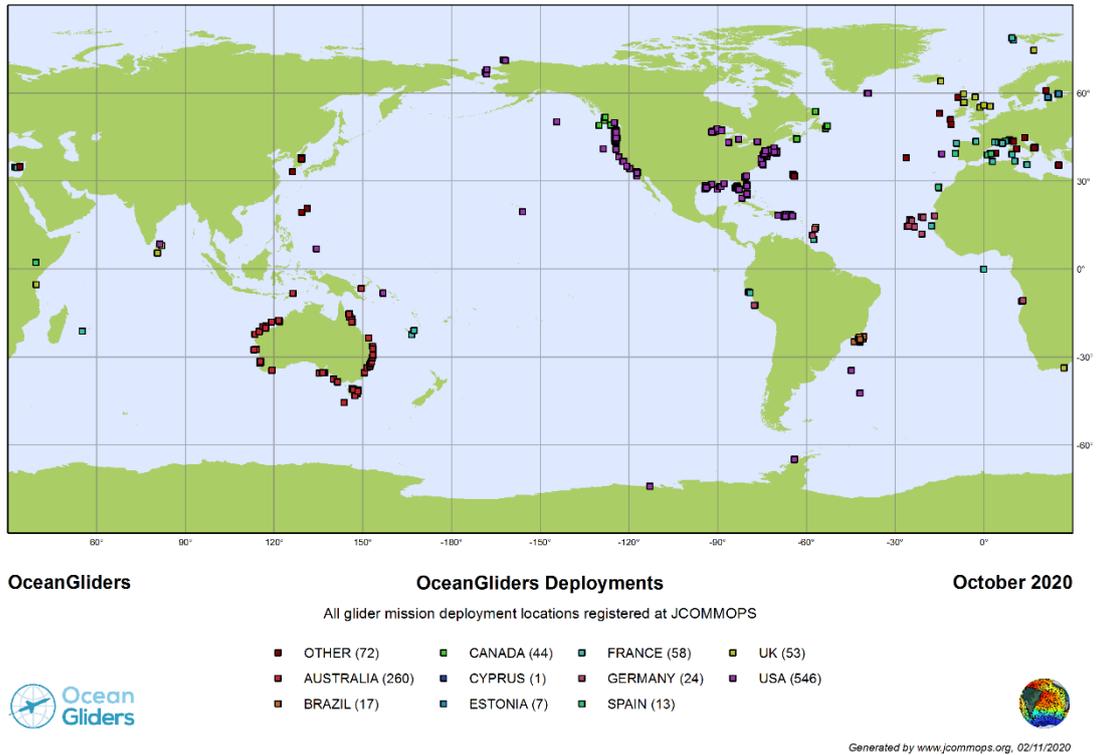


Figura 13 - Mapa de los planeadores submarinos (Gliders) en operación en octubre de 2020. Extraído de <https://www.ocean-ops.org/board?t=oceangliders>.

Radar HF

Los radares de alta frecuencia se utilizan para medir el oleaje y los campos de corrientes superficiales. Su funcionamiento (Figura 14) está basado en el efecto Doppler y son capaces de monitorear de manera continua áreas costeras con una cobertura de hasta 300 km de radio. Como desventaja aparece su elevado costo. En la Figura 15 se pueden ver las ubicaciones de los radares de la red global *The Global HF Radar Network* creada para aunar esfuerzos hacia un sistema global de medición de corrientes superficiales.

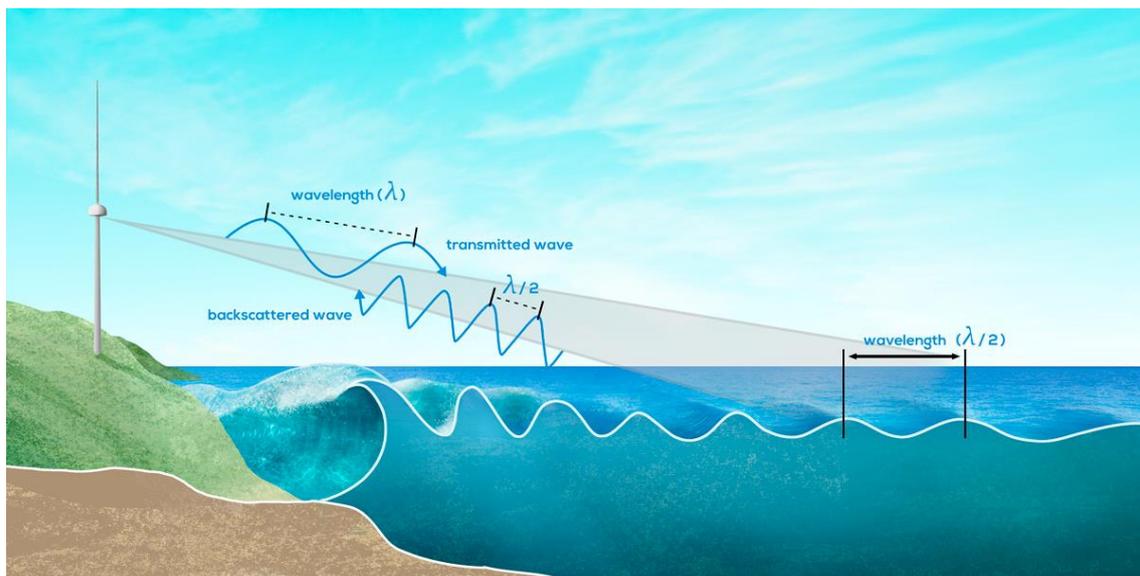


Figura 14 – Ilustración del funcionamiento del radar de alta frecuencia (HF). Extraído de <https://marinecue.it/5-tecnologie-utilizzate-per-studiare-gli-oceani/8503/>.

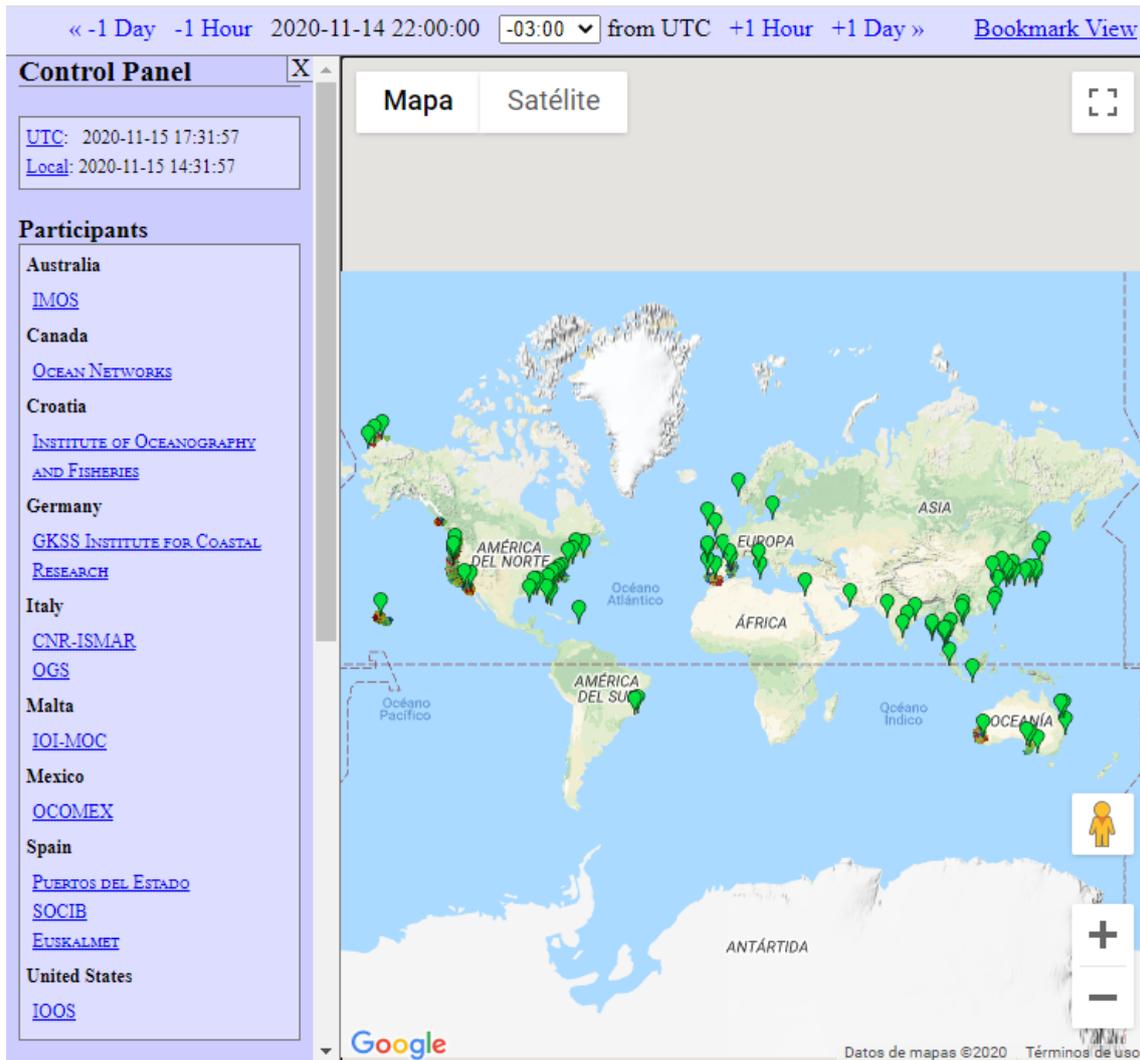


Figura 15 – Instituciones participantes y ubicación de los radares de HF de la red global The Global HF Radar Network en noviembre de 2020. Extraído de <http://global-hfradar.org>.

II.1.2. Modelos numéricos de la dinámica del océano

II.1.2.1. Ecuaciones gobernantes

A pesar de la creciente riqueza de observaciones satelitales e *in situ*, la cobertura de observaciones es todavía escasa para una parte amplia del océano. Los modelos numéricos son capaces de predecir las condiciones del océano a través de la resolución numérica de las ecuaciones diferenciales que gobiernan el movimiento de los fluidos. La simulación requiere tanto de la habilidad de representar de manera precisa los fenómenos que son resueltos como de parametrizar las escalas no resueltas, por ejemplo problemas de consideración de cierre de la turbulencia.

Generalmente, los modelos oceánicos de escala global y de cuenca utilizan las ecuaciones primitivas que relacionan las variables de velocidad, temperatura y salinidad y su evolución en

el espacio y el tiempo. Estas ecuaciones primitivas consisten en tres conjuntos de ecuaciones de balance:

- La ecuación de continuidad. Representa la conservación de la masa. En el campo de la oceanografía se considera que el fluido es incompresible (Winckler, 2018) y por lo tanto la conservación de la masa coincide con la conservación de volumen. La expresión de la ecuación de continuidad para fluido incompresible es:

$$\nabla \cdot \vec{u} = 0 \quad (1)$$

donde \vec{u} es el vector velocidad.

- La ecuación de conservación de momento. Para fluido incompresible (ecuaciones de Navier Stokes):

$$\frac{\partial \vec{u}}{\partial t} = -\nabla \cdot (\vec{u}\vec{u}) + \vec{g} + \nu \nabla^2 \vec{u} - \frac{1}{\rho} \nabla p - 2\vec{\Omega} \times \vec{u} + \vec{S} \quad (2)$$

donde \vec{u} es el vector velocidad, \vec{g} es la gravedad, ν es el coeficiente de difusión de momento, ρ es la densidad, p es la presión, $\vec{\Omega}$ es la velocidad angular de la tierra y \vec{S} las restantes fuerzas externas (fricción lateral y de fondo, tensión del viento, gradientes de presión atmosférica...).

- La ecuación de estado relaciona la densidad con la temperatura y la salinidad $\rho = \rho(T, S)$. El transporte de trazadores se resuelve a través de las ecuaciones de conservación para cada trazador (advección-difusión), así, la temperatura del sistema se obtiene de un balance de fuentes y sumideros de calor. Una ecuación similar se aplica a la salinidad en el océano. La ecuación de transporte de temperatura es:

$$\frac{\partial T}{\partial t} = -\nabla \cdot (T\vec{u}) + \nabla \cdot (D \cdot \nabla T) + Q \quad (3)$$

donde T es la temperatura, \vec{u} es el vector velocidad, D es el coeficiente de difusión de temperatura y Q representa las fuentes/sumideros.

Atendiendo a los fenómenos oceánicos que se quieran estudiar existen tres aproximaciones principales para representar las ecuaciones de Navier-Stokes de una manera más simplificada:

- Promediado de las ecuaciones: RANS (*Reynolds-Averaged-Navier-Stokes*). Se calcula el flujo medio y las fluctuaciones turbulentas se consideran menos importantes. Para ello cada cantidad física se representa en promedio en el tiempo de manera que cada variable es una suma de componente media y componente fluctuante. El promediado transforma los términos no lineales en términos lineales pero aparece un nuevo término que requiere el uso de un modelo de turbulencia (Popescu, 2014). Estas ecuaciones permiten calcular los valores medios pero no dan información sobre las actuaciones turbulentas, a excepción de aquellas cuya escala es bastante mayor al tamaño de la discretización espacial (Winckler, 2018). Las formulaciones más extendidas en su uso para inferir el efecto de las escalas de turbulencia no resueltas son: modelos de cero ecuación, modelos de una y dos ecuaciones y modelos del tensor de Reynolds, siendo uno de los más populares el modelo de dos ecuaciones $k - \varepsilon$ (Schäfer, 2006).

- Técnicas LES (*Large Eddy Simulation*, Figura 16). Se modela la turbulencia ignorando las escalas más pequeñas del flujo (utilizando un filtro pasa-bajo). Tiene un mayor costo computacional que RANS pero tienen la ventaja de resolver las escalas de turbulencia mayores. En la actualidad se reservan a investigación siendo el modelo de Smagorinsky uno de los más utilizados (Winckler, 2018; Schäfer, 2006).

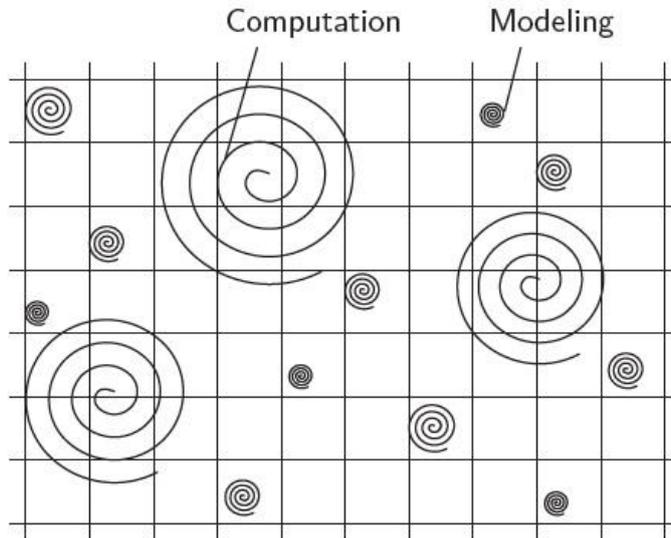


Figura 16 – Tratamiento de las estructuras turbulentas de escalas grandes y pequeñas con técnicas LES. Extraído de Schäfer (2006).

- Aproximaciones de las ecuaciones promediadas. La simplificación de las ecuaciones de Navier-Stokes se realiza a través de un promediado en una dirección, por ejemplo, promediado en vertical (Popescu, 2014).

II.1.2.2. Resolución numérica

Para resolver numéricamente las ecuaciones diferenciales que gobiernan la mecánica de fluidos, éstas se deben convertir en un sistema algebraico a través de numerosas aproximaciones. Una de las decisiones a adoptar durante la construcción del modelo numérico es la definición de la técnica de discretización: diferencias finitas, volúmenes finitos o elementos finitos. A continuación, se introducen las características de las técnicas de discretización utilizadas por los modelos numéricos del océano.

- Diferencias finitas:

El dominio continuo se discretiza a través de un mallado (generalmente uniforme) de manera que las ecuaciones se resuelven en su forma diferencial a través de un sistema algebraico para un número finito de nodos en los que se determinan los valores de las variables del problema (Popescu, 2014). Para cada nodo, las derivadas se aproximan a través de una combinación algebraica que relaciona los valores de la variable en nodos del entorno basándose en la serie de Taylor. Los métodos explícitos permiten estimar el estado del sistema en un momento posterior al actual, siendo este último conocido. Los métodos implícitos y semi-implícitos, en contraste, consideran tanto en estado actual como el

posterior, que es desconocido (Winckler, 2018). Los valores fuera de los nodos pueden ser obtenidos por interpolación. En la Figura 17, panel superior, se muestra el ejemplo de la representación de la función g . La función g varía en el espacio x a través del tiempo t . Los valores calculados (puntos negros) corresponden a los nodos de discretización del tiempo y el espacio (Popescu, 2014).

- Volúmenes finitos:

Este método fue inicialmente desarrollado para resolver problemas relacionados con cuerpos de agua y posteriormente se extendió a otras disciplinas (Popescu, 2014). Tiene como ventaja el poder utilizar geometrías arbitrarias o una discretización no estructurada del dominio. Las características de esta técnica de discretización son: el dominio se discretiza en elementos discretos llamados volúmenes de control; la ecuación diferencial a resolver se integra en una ecuación de equilibrio para cada volumen de control; las integrales obtenidas se aproximan mediante integración numérica; los valores de la función desconocida y sus derivadas se aproximan utilizando los valores en los nodos del volumen de control; la solución del problema se obtiene ensamblando las ecuaciones discretas obtenidas sobre un volumen de control, en un sistema algebraico de ecuaciones. Una propiedad importante de la técnica de volúmenes finitos es que el flujo numérico se conserva de una celda a otra (muy útil para la dinámica de fluidos). En la Figura 17, panel central, se muestra el ejemplo de la discretización para un dominio en 3 dimensiones centrado en la celda (Popescu, 2014).

- Elementos finitos:

El dominio se divide en subdominios que se denominan elementos finitos y que generalmente no están estructurados; en 2D, suelen ser triángulos o cuadriláteros, mientras que en 3D se utilizan con mayor frecuencia tetraedros o hexaedros. En la Figura 17, panel inferior, se presenta un ejemplo de subdivisión en 2D. La característica distintiva de esta técnica es que las ecuaciones se multiplican por una función de ponderación local del elemento antes de integrarse en todo el dominio (Ferziger and Peric, 2002). La función de ponderación suele tener la misma forma para todos los elementos, siendo, en los métodos más simples, una función de forma lineal dentro de cada elemento a partir de sus valores en las esquinas de los elementos. La integración a todo el dominio se realiza a través de matrices de rigidez y vectores de carga locales que se ensamblan para el sistema global del dominio. Para la solución del sistema se deben incorporar también las condiciones de contorno.

La aproximación de las ecuaciones diferenciales, de un sistema continuo espacio-tiempo por un dominio discreto, genera error de discretización o truncamiento, incluyéndose como fuentes de errores los que se listan a continuación: la selección de las coordenadas horizontales (esféricas u ortogonales), el mallado vertical y horizontal (el espaciamiento y su regularidad o no), la discretización del tiempo, la representación de las capas límite de la superficie y del fondo, la representación de la topografía del fondo, la ecuación de estado, la ecuación de transporte de trazador y momento, los modelos para representar los procesos de escalas no resueltas (de submalla), la modelación de la viscosidad y de la difusividad. Las fórmulas utilizadas para parametrizar los fenómenos no resueltos reflejan el conocimiento actual de esos fenómenos, como, por ejemplo los parámetros de viscosidad turbulenta y difusividad turbulenta en modelos del movimiento de gran escala para representar la interacción con el movimiento de mesoescala (Kowalik and Murty, 1993). Se hace notar que, debido al hecho de que la dimensión horizontal es normalmente mucho mayor a la dimensión

vertical, se tiende a considerar valores distintos para los parámetros en dirección horizontal y vertical.

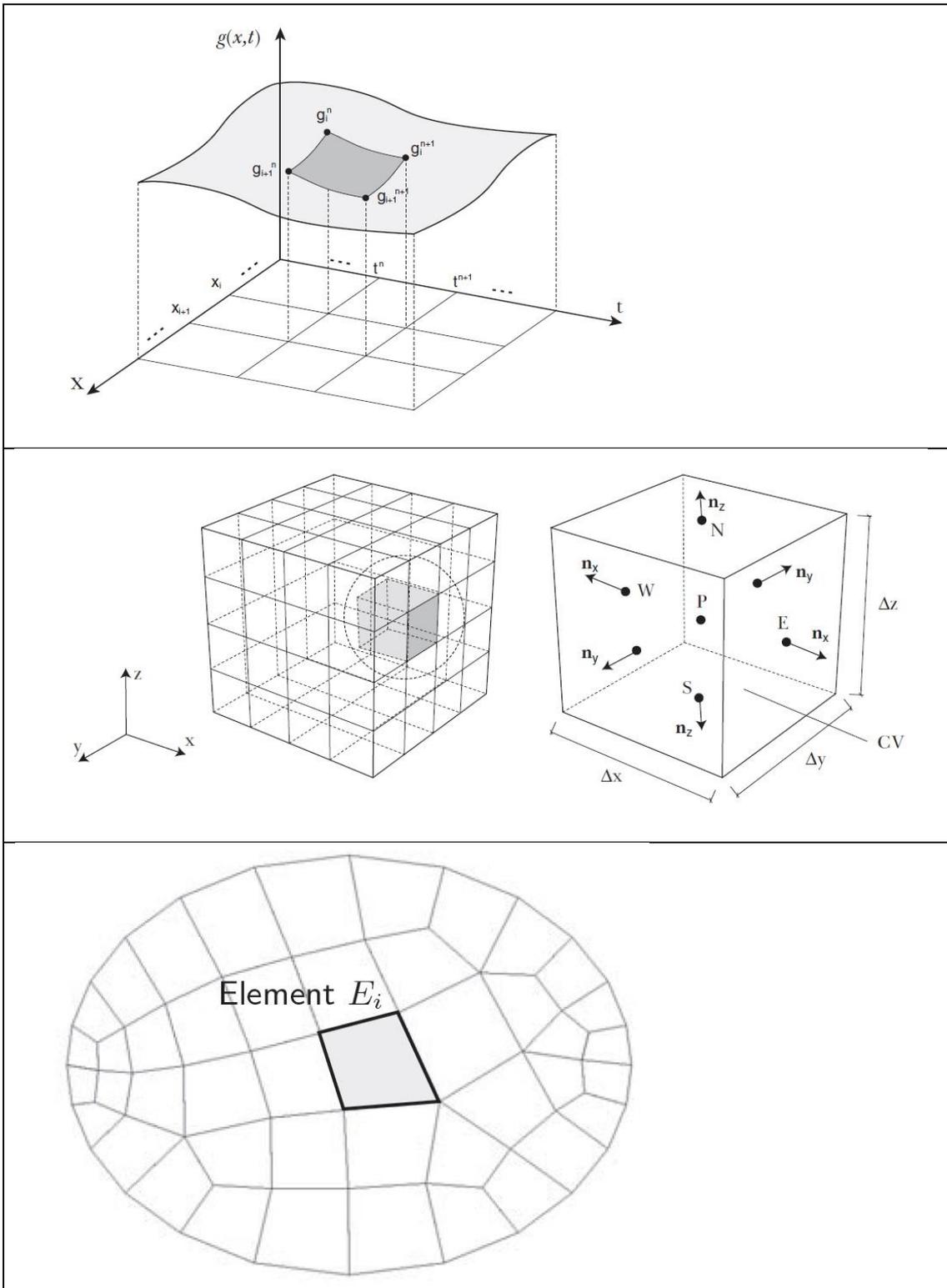


Figura 17 – Técnicas de discretización. Diferencias finitas (panel superior). Volúmenes finitos (panel central). Elementos finitos (panel inferior). Las 2 primeras extraídas de Popescu (2014) y la tercera extraída de Schäfer (2006).

Los primeros modelos operacionales de predicción a corto plazo del estado del océano (años 90 y principios de los años 2000) tenían una resolución horizontal de entre 1° y $1/4^\circ$ en el océano global y mayor en aplicaciones regionales. Este rango de resolución horizontal no es suficiente

para resolver la variabilidad de mesoescala, que tiene una escala de longitud de entre 10 km y 400 km y una escala de tiempo de entre semanas y días (Schiller and Brassington, 2011; Kantha and Clayson, 2000). Los sistemas de pronóstico del océano de gran escala de la actualidad centran su atención en resolver la turbulencia de mesoescala (*mesoscale eddies*), con una resolución de $1/10^{\circ}$ o mayor para el océano global y hasta 80 niveles o capas verticales y añadiendo pronósticos que atiendan la variabilidad biofísica. Se están desarrollando modelos globales de mayor resolución con los que se prevé llegar a una malla horizontal de $1/25^{\circ}$ en los próximos años (Chassignet et al., 2018; Schiller and Brassington, 2011).

Los requerimientos de los centros de predicción operacionales a la hora de producir pronósticos del océano global incluyen modelos numéricos del océano, información de pronósticos de forzantes atmosféricos en superficie, la batimetría oceánica, y el estado del océano reciente obtenido de las observaciones e introducido en el modelo a través de la asimilación de datos (Schiller and Brassington, 2011). Los requerimientos computacionales y operacionales de las simulaciones oceánicas para una apropiada resolución espacial y temporal son elevados requiriendo el uso de técnicas de computación de alto desempeño (HPC, por sus siglas en inglés) para proveer los pronósticos y servicios en tiempos razonables (Chassignet et al., 2018).

Los sistemas regionales o costeros que no resuelven el estado del océano global precisan, además, el pronóstico de las condiciones del borde oceánico. Estas condiciones de contorno pueden ser proveídas por modelos globales, de modo que el modelo regional o costero sea un subescalado de mayor resolución del modelo global (Schiller and Brassington, 2011). En las regiones costeras con influencia de agua dulce se deberá tener en cuenta también la entrada de los ríos y, en algunas costas, la descarga de agua subterránea directamente al océano costero o el aporte de agua dulce de numerosos arroyos y ríos pequeños puede ser importante, por lo que también deberán estar incluidos en la modelación (Schiller and Brassington, 2011).

II.1.2.3. Ecuaciones en 2 dimensiones. Ecuaciones promediadas en vertical

Existen dos tipos de modelos bidimensionales (2D). Los modelos bidimensionales en horizontal (2DH) son los que se utilizan para simular mayormente ríos anchos y someros, estuarios, la costa, lagos, fiordos, etc. Permiten conocer las propiedades en planta del flujo dado que reproducen los fenómenos en el plano horizontal (x, y) . Los modelos bidimensionales en vertical (2DV) se utilizan en situaciones en las que únicamente se quiere conocer la distribución transversal del flujo y basta con resolver el flujo en el plano vertical (x, z) ó (y, z) . A continuación se introducen únicamente los modelos 2DH.

Los modelos 2DH se basan en el supuesto de que el flujo en el plano horizontal predomina por sobre la dimensión vertical. La batimetría se introduce para todo el espacio en planta y se considera que no hay flujo a través del fondo ni de la superficie libre. Para cada paso de tiempo, se calcula la profundidad total en la sección y las velocidades promediadas en la columna vertical de agua. Para su resolución se utilizan ecuaciones del movimiento promediadas verticalmente de manera que se obtienen velocidades horizontales promediadas en vertical. La integración en vertical de las ecuaciones de momento asumiendo presión hidrostática (la aceleración vertical

es despreciable en comparación a la aceleración de la gravedad), que las velocidades horizontales son constantes con la profundidad y condiciones barotrópicas (despreciando la estratificación en la densidad) resultan en las ecuaciones de onda larga, conocidas también como las ecuaciones de aguas someras (NSWE, por sus siglas en inglés) (Winckler, 2018; Kowalik and Murty, 1993).

Con un sistema de referencia cartesiano (x, y) en el plano horizontal, las ecuaciones de continuidad y momento promediadas en vertical se expresan como:

$$\frac{\partial h}{\partial t} + \frac{\partial(hU)}{\partial x} + \frac{\partial(hV)}{\partial y} = S_h \quad (4)$$

$$\frac{\partial U}{\partial t} = -\frac{\partial(UU)}{\partial x} - \frac{\partial(UV)}{\partial y} - \frac{g\rho}{\rho_0} \frac{\partial\eta}{\partial x} + \frac{1}{h} \frac{\partial}{\partial x} \left(hv \frac{\partial U}{\partial x} \right) + \frac{1}{h} \frac{\partial}{\partial y} \left(hv \frac{\partial U}{\partial y} \right) + fV + S_x \quad (5)$$

$$\frac{\partial V}{\partial t} = -\frac{\partial(UV)}{\partial x} - \frac{\partial(VV)}{\partial y} - \frac{g\rho}{\rho_0} \frac{\partial\eta}{\partial y} + \frac{1}{h} \frac{\partial}{\partial x} \left(hv \frac{\partial V}{\partial x} \right) + \frac{1}{h} \frac{\partial}{\partial y} \left(hv \frac{\partial V}{\partial y} \right) - fU + S_y \quad (6)$$

donde el eje x representa la dirección zonal (positiva hacia el este), el eje y representa la dirección meridional (positiva hacia el oeste), h es la profundidad total del agua, U la componente zonal de la velocidad promediada en vertical, V la componente meridional de la velocidad promediada en vertical, g la aceleración de la gravedad, ρ es la densidad y ρ_0 es una densidad de referencia, η es la elevación de la superficie libre, ν el coeficiente de difusión de momento, f es el parámetro de Coriolis ($f = 2\Omega \sin(\varphi)$, con Ω el módulo de la velocidad angular de la tierra y φ la latitud), S_h son las posibles fuentes (y sumideros) de masa en la ecuación de continuidad, S_x y S_y son la componente zonal y meridional del resto de las fuerzas externas (fricción lateral y de fondo, tensión del viento, gradientes de presión atmosférica, etc.).

La simplificación del problema desde tres a dos dimensiones reduce el costo computacional en general al 1% del esfuerzo de cálculo que necesita el modelado 3D (Winckler, 2018). Como contraparte, se pierde información sobre la estructura vertical del flujo. Estas ecuaciones se pueden utilizar para el modelado de una serie de ondas largas, es decir, mareas, tsunamis, mareas meteorológicas (*storm surge*), corrientes oceánicas (en la medida que no exista estratificación vertical), y ondas a escala oceánicas, como las de Kelvin y Rossby.

II.1.2.4. Modelos tridimensionales. La coordenada vertical

Los modelos 3D, o tridimensionales, permiten determinar las tres componentes espaciales de la velocidad en base a las ecuaciones 3D de Navier-Stokes. Estos modelos permiten calcular la velocidad y la presión en tres dimensiones. Entre los modelos 3D destacan las ecuaciones RANS (*Reynolds Averaged Navier-Stokes equations*) y las técnicas LES (*Large Eddy Simulation*) y DNS (*Direct Numerical Simulation*).

Los modelos que utilizan RANS y LES ya fueron introducidos como una simplificación a las ecuaciones de Navier-Stokes en II.1.2. La técnica de resolver las ecuaciones de Navier-Stokes con escalas del orden de los procesos viscosos (comparables a la escala de Kolmogorov) se denomina *Direct Numerical Simulation* (DNS) y es aplicable, con los recursos computacionales actuales, a dominios pequeños, del orden de decímetros, ya que no hacen simplificaciones en

relación a la turbulencia si no que la resuelven de forma exacta (Winckler, 2018; Ferziger and Peric, 2002). En la mecánica de fluidos ambiental, el tamaño de las escalas espaciales hace impracticable el uso de DNS.

En la representación de un flujo tridimensional, la selección de coordenadas es arbitraria y depende de la geometría del problema en particular. Usualmente se utilizan coordenadas cartesianas con ejes ortogonales entre sí que se cortan en un punto origen, es decir, en el punto de referencia del sistema. Otras alternativas (coordenadas esféricas, cilíndricas o curvilíneas) pueden ser adecuadas para geometrías específicas (Winckler, 2018). Una característica de los fluidos en rotación y estratificados, como es el caso del océano, es que lo lateral domina frente a lo vertical. Por tanto es común la orientación de los ejes horizontales de manera ortogonal a la dirección vertical local que viene determinada por la gravedad. La elección del tipo de coordenada vertical determina uno de los aspectos más importantes en el diseño de un modelo oceánico. Los problemas prácticos de representación y parametrización están a menudo directamente relacionados con la elección de coordenadas verticales (Schiller and Brassington, 2011). Actualmente, existen tres coordenadas verticales principales en uso, ninguna de las cuales proporciona una utilidad universal lo que ha llevado a enfoques híbridos. La Figura 18 muestra los tres regímenes habituales del océano atendiendo a la zona en la que cada coordenada vertical maximiza su utilidad y a continuación se da una breve descripción de cada una de estas zonas.

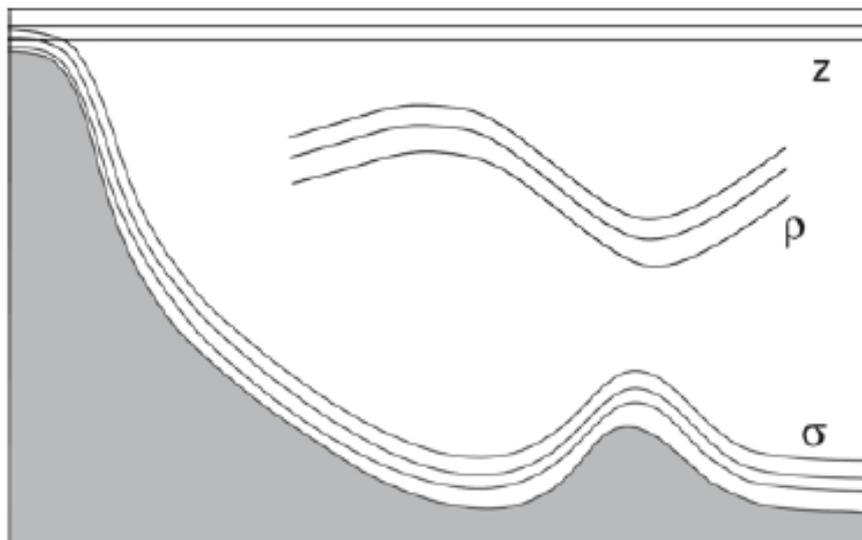


Figura 18 -Esquema de una cuenca oceánica. Se ilustran los tres regímenes del océano en relación con las consideraciones de una coordenada vertical apropiada. Extraído de Schiller and Brassington (2011).

En primer lugar está la capa superficial mezclada. Esta región es generalmente turbulenta y dominada por transferencias de momento, calor, agua dulce y trazadores. Normalmente está muy bien mezclado en la vertical a través de procesos convectivos/turbulentos tridimensionales. Son procesos no hidrostáticos, por lo que requieren de una resolución horizontal y vertical muy alta para ser representados explícitamente, es decir, una cuadrícula vertical a horizontal cerca de la unidad que se consigue utilizando la coordenada z . Los procesos de transporte de trazadores en el interior del océano predominantemente ocurren a lo largo de direcciones de densidad constante, por lo que la utilización de una coordenada vertical isopícnica (dependiente de la presión) ρ que coincida con estas capas de densidad constante resulta la más adecuada en esta zona. El tercer régimen incluye las regiones

dominadas por las corrientes impulsadas por densidad (*overflow*) y capas turbulentas de fondo (aguas profundas del océano mundial). Las coordenadas σ de seguimiento del terreno son particularmente adecuadas para modelar flujos a lo largo de la plataforma continental.

II.1.2.5. Modelos acoplados océano-atmósfera

Los intercambios de calor, agua dulce y momento entre los océanos y la atmósfera tienen un rol decisivo en el sistema climático global. Como parte del cambio climático antropogénico, que implica el calentamiento global y un ciclo hidrológico más severo, se esperan cambios en los flujos de calor y agua dulce en la interfaz océano-atmósfera. El conocimiento actual de muchos aspectos de la interacción océano-atmósfera es todavía básico pese a la mejora que supuso la llegada de la observación satelital que registra variables relacionadas con el flujo en la interfaz océano-atmósfera como son la temperatura de la superficie del mar y la velocidad del viento (Schiller and Brassington, 2011).

La forma más simple de modelación del océano incluye los forzantes atmosféricos en la superficie pero no permite la retroalimentación hacia la atmósfera. Esto puede tener un impacto limitado en un pronóstico de 15 días, pero el acoplamiento a un modelo atmosférico es esencial en el pronóstico desde escalas temporales estacionales a interanuales de eventos como El Niño (Schiller and Brassington, 2011).

En un modelo dinámico acoplado océano-atmósfera la interacción entre los dos sistemas se modela en lugar de prescribirse. El acoplamiento requiere una continua (o intermitente) interacción de doble sentido entre el océano y la atmósfera. Esta interacción se da a través de los flujos de calor, agua dulce y momento con el fin de que, a través de una simulación precisa de las interacciones aire-mar y aire-tierra, se pueda extender la habilidad de los modelos atmosféricos para generar pronósticos a mayor plazo del que se consigue actualmente sin los océanos (Kantha and Clayson, 2000). Trabajar con modelos acoplados permite, también, generar pronósticos oceánicos a más largo plazo.

El oleaje representa la interfaz entre la atmósfera y el océano y algunos sistemas de pronóstico ya incluyen el oleaje costero como parte de sus servicios. Se incluye la introducción del efecto de la rugosidad que proporciona el oleaje en los modelos atmosféricos así como la utilización de modelos acoplados atmósfera-océano para la predicción del clima y los huracanes. Así, los modelos de huracán ya incluyen un modelo meteorológico completo, un modelo oceánico completo y un modelo de oleaje completo (Schiller and Brassington, 2011). En los estudios de pronóstico del clima terrestre de décadas, o de más largo alcance temporal, se requiere, que además del océano y la atmósfera, los modelos sean capaces de generar simulaciones globales acopladas en las que se incluya el hielo marino (Kantha and Clayson, 2000).

II.1.3. Asimilación de datos

II.1.3.1. Introducción

La modelación oceánica tradicional consiste en resolver las ecuaciones del modelo de la manera más precisa posible y, después, comparar los resultados con los datos de observaciones.

Siguiendo este procedimiento, se han alcanzado niveles cualitativos satisfactorios. Sin embargo, debido al hecho de que generalmente se comprueba que los errores de los modelos exceden a los errores de los datos observados, resulta insatisfactorio disponer de una solución del modelo que no esté influenciada por los datos observados (Bennett, 1992). La asimilación de datos es el proceso de cálculo combinando información tanto de las observaciones como de la dinámica del modelo (Schiller and Brassington, 2011). En la Figura 19 se muestra una ilustración que representa los pilares de la oceanografía operacional con asimilación de datos.

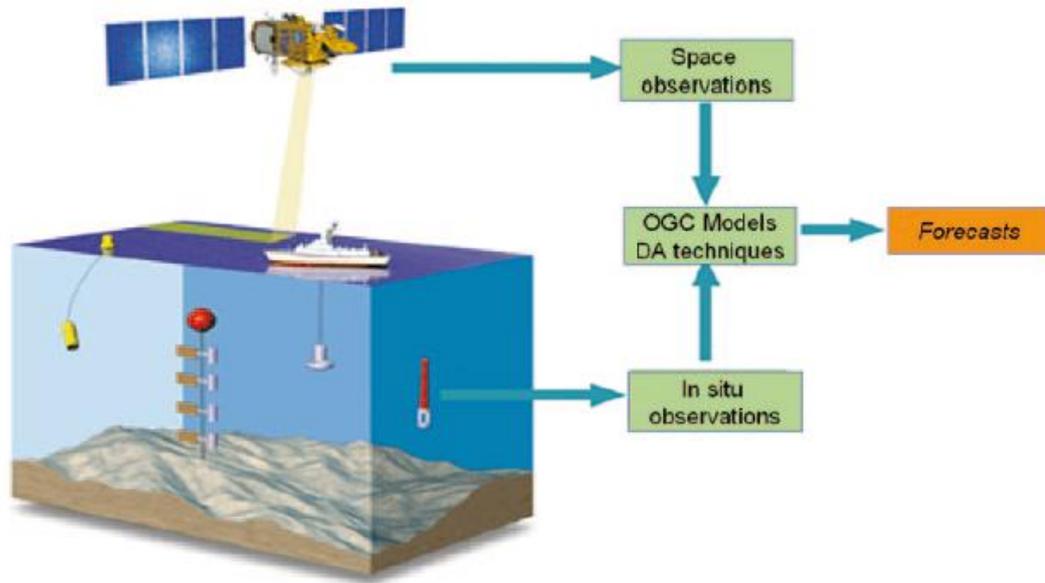


Figura 19 – Los tres pilares de la oceanografía operacional en tiempo cuasi-real: las observaciones satelitales, las medidas in-situ y los modelos de circulación oceánica combinados a través de los sistemas de asimilación de datos. Extraído de Schiller and Brassington (2011).

Para hacer uso de la información incorporada en observaciones y modelos, es necesario comprender las características de esta información. En particular, se debe reconocer que tanto las observaciones como los modelos tienen errores y estos deben tenerse en cuenta (Lahoz et al., 2010). Las observaciones tienen errores instrumentales: aleatorios (que se reducen al promediar) o sistemáticos (también conocidos como sesgo), y errores de representatividad (discrepancia entre las escalas detectadas por las observaciones y por el modelo). El uso de los modelos también introduce errores. Estos errores son inherentes a la construcción de los modelos. Los modelos pueden estar incompletos debido a la falta de comprensión o a la omisión de procesos para hacer que el problema sea manejable.

Otra característica clave de las observaciones (o mediciones) es que son discretas en el espacio y el tiempo, y por tanto la información proporcionada por las observaciones tiene “huecos”. Conceptualmente, interesa extender la información suministrada por las observaciones y completar esos vacíos (Lahoz et al., 2010), lo cual se realiza considerando un modelo como los mencionados anteriormente.

El progreso en la asimilación de datos oceánicos ha sido posible en gran medida gracias a los avances en computación de los últimos 30 años. Sin embargo la teoría y las técnicas de asimilación de datos tienen una larga historia con raíces matemáticas en el campo de la

probabilidad y la estimación, la teoría inversa y el cálculo clásico de variaciones. Las raíces operacionales de la asimilación de datos están estrechamente vinculadas a la comunidad de predicción meteorológica, que durante mucho tiempo ha abordado el problema de cómo suavizar e interpolar mediciones dispersas para optimizar las predicciones meteorológicas posteriores (Schiller and Brassington, 2011). En la Figura 20 se ofrece un ejemplo de cómo la asimilación de datos añade valor a las observaciones y al modelo; se denomina análisis a la estimación del estado tras la asimilación.

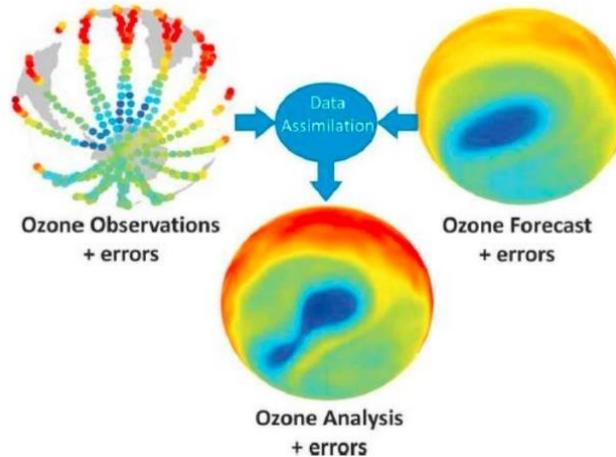


Figura 20 – Representación simplificada de cómo funciona la asimilación de datos y añade valor a la información de las observaciones y del modelo. Extraído de Lahoz et al. (2010).

La asimilación de datos en oceanografía se ha desarrollado en varias direcciones, cada una con diferente objetivo o énfasis. Los temas u objetivos principales pueden resumirse brevemente como sigue (Schiller and Brassington, 2011):

- Interpolación, extrapolación y suavizado. El propósito es la estimación del estado del océano a partir de datos escasos o irregularmente distribuidos. Con el objetivo de utilizar toda la información de medidas disponible aparecen las técnicas de interpolación cuando los datos son dispersos, las técnicas de extrapolación cuando hay escasez de datos y las técnicas de suavizado cuando hay gran densidad de datos e incluso datos redundantes. Estas técnicas son utilizadas tanto para la asimilación en modelos de pronóstico como para modelos de reanálisis (estudios retrospectivos que incluyen la asimilación de los datos disponibles, Lahoz et al., 2010). Esta visión de la asimilación de datos es la base para la mayoría del trabajo de asimilación de datos oceánicos.
- Calibración de parámetros. Se utilizan técnicas de asimilación de datos para la calibración de parámetros durante la generación de un modelo. El propósito es desarrollar el modelo más preciso mediante el ajuste de los parámetros inciertos o desconocidos (como por ejemplo parámetros escalares implicados en los submodelos de turbulencia) de forma que las predicciones sean lo más congruente posible con los datos disponibles. La dificultad de estos procedimientos radica en que generalmente incluyen la resolución de problemas de inversión fuertemente no lineales.
- Test de hipótesis. Este grupo incluye el empleo de técnicas de asimilación para testear el modelo dinámico y el sistema de observación de un sistema de predicción oceánica y también para el diseño de un nuevo sistema de observación.

Otra forma de categorizar las técnicas de asimilación de datos (Figura 21) es atendiendo al momento en el que se adquieren las observaciones con respecto al análisis (Carrassi et al., 2018):

predicción, filtrado y suavizado. La técnica de predicción tiene en cuenta las observaciones del pasado para corregir las condiciones iniciales de un modelo de pronóstico; en la técnica de filtrado se añade la observación del estado actual y de manera secuencial se produce la corrección de las condiciones iniciales y el pronóstico entre los instantes de observación; y las técnicas de suavizado tienen información de observaciones pasadas y futuras al momento del análisis por lo que suele identificarse con las condiciones de los reanálisis.

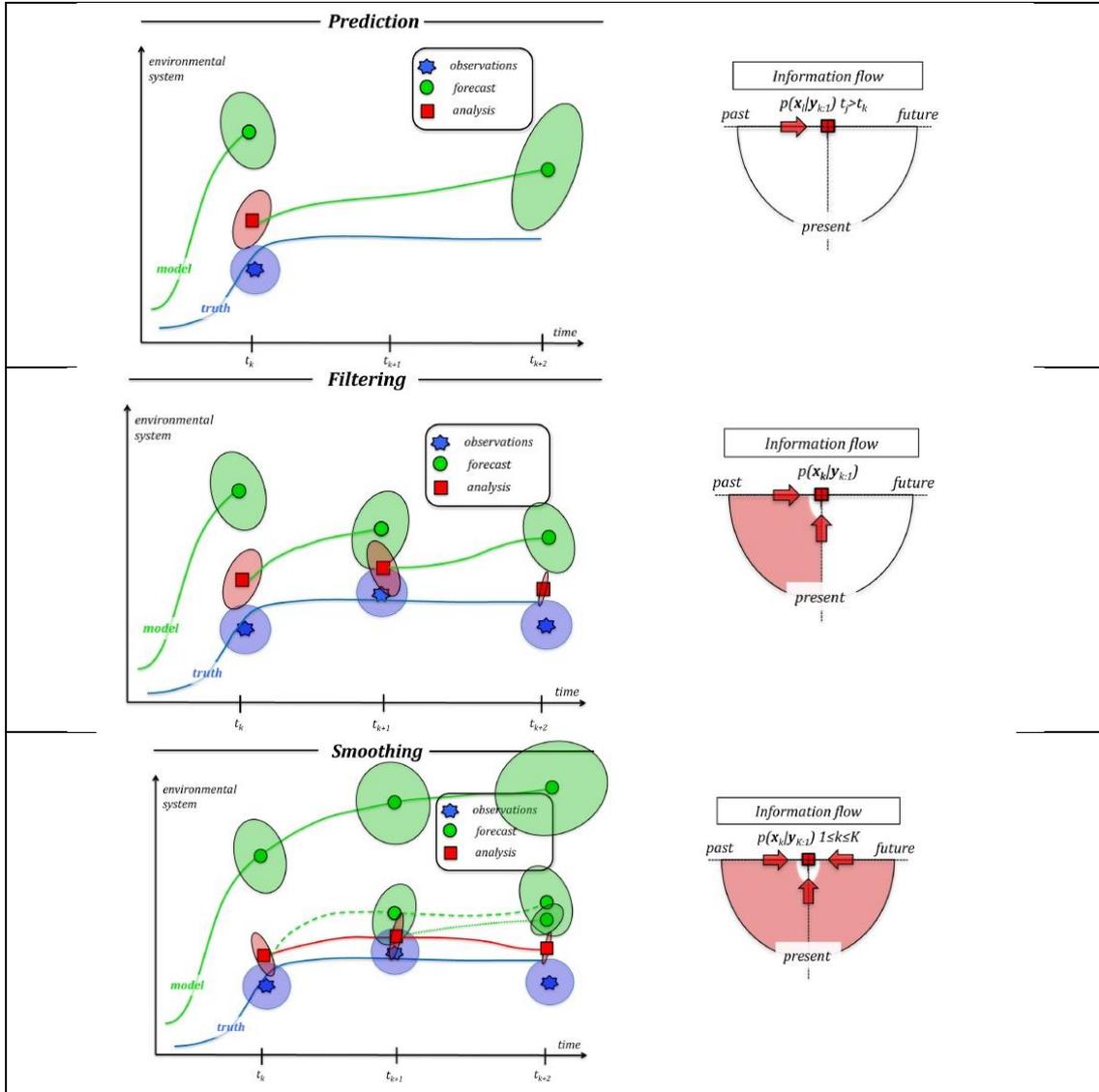


Figura 21 – Técnicas de asimilación de datos: predicción (prediction), filtrado (filtering) y suavizado (smoothing). La evolución de la señal verdadera (desconocida) se representa por una línea azul, los elipsoides representan las funciones de probabilidad de observación (azul), pronóstico (verde) y análisis (rojo), y la estrella, círculo y cuadrado respectivamente podrían corresponder a la media de esta función de probabilidad, aunque su elección dependerá del modelo de asimilación elegido. Extraído de Carrasi et al. (2018).

II.1.3.2. Teorema de Bayes

El Teorema de Bayes es uno de los conceptos básicos a partir del cual se pueden derivar las técnicas más complejas de asimilación de datos. Expresa de forma concisa la base de la asimilación de datos de manera que la información se expresa en términos de probabilidad (Carrasi et al., 2018). La Ecuación (7) presenta la fórmula matemática.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \quad (7)$$

Se asume:

x : variables de estado (en general, variables de un modelo).

y : variables observadas.

$p(x)$: función de distribución (pdf, del inglés *probability density function*) de las variables de estado (pdf *prior o background*). Concentra toda la información de las variables de estado antes de la asimilación de las observaciones.

$p(y)$: función de distribución (pdf) marginal de las observaciones, es decir, la distribución de y cualquiera que sea el valor del estado. Esta distribución es independiente de x , se trata como una constante de normalización.

$p(y|x)$: probabilidad condicional o verosimilitud de las observaciones, dadas las variables de estado (*likelihood*). Es decir, lo que se hubiese observado si se conociera el estado real.

$p(x|y)$: pdf de las variables de estado dadas las observaciones (o pdf posterior).

En el Teorema de Bayes se combina información del modelo y las medidas a partir de estimadores de los errores en ambos: condiciones iniciales, de borde, inhomogeneidades del modelo y los errores de las medidas. Esto se suele simplificar considerando errores Gaussianos (Schiller and Brassington, 2011).

II.1.3.3. Modelo state-space

El estado del sistema x (variables aleatorias) en el tiempo k , se asume que evoluciona desde el estado en un tiempo anterior $k - 1$ (se asumen tiempos discretos) a través de un modelo dinámico que representa las ecuaciones diferenciales de evolución del sistema \mathcal{M} (Carrassi et al., 2018). Este modelo será el que proporcione la distribución del *background*:

$$x_k = \mathcal{M}(x_{k-1}) + \eta_k \quad (8)$$

El estado del sistema es ordenado en un vector $x \in \mathbb{R}^N$, la dimensión de N viene dada por la elección de la cantidad de puntos que discreticen el espacio multiplicada por la cantidad de variables que resuelva el modelo (Schiller and Brassington, 2011). Así, para una malla de V puntos en los que se tengan en cuenta 4 variables (por ejemplo salinidad, velocidad este y norte y altura de la superficie libre), el tamaño de los vectores será de $x_k \in \mathbb{R}^N$, con $N = 4 \cdot V$

Se asume que el error del modelo, η_k , se distribuye como una normal de media 0 y matriz de covarianza Q_k (aunque en un principio y de manera más general puede estar incluido en el modelo y no como error aditivo).

Para las variables aleatorias observadas en el tiempo k se tiene la ecuación estocástica de la observación (Carrassi et al., 2018):

$$\mathbf{y}_k = \mathcal{H}(\mathbf{x}_k) + \mathbf{v}_k \quad (9)$$

El operador \mathcal{H} es el modelo que mapea, o transforma, las variables de estado a variables observadas. Por ejemplo, cuando se mide radiancia y se interpola a los valores de temperatura de una grilla. Las dimensiones del vector \mathbf{y} corresponden al número de observaciones, M , en el tiempo k . Se asume que el error observacional \mathbf{v}_k se distribuye como una normal de media 0 y matriz de covarianza \mathbf{R}_k .

II.1.3.4. Interpolación óptima (Optimal Interpolation)

La estrategia de interpolación óptima considera errores Gaussianos, por lo que las variables de estado pasan a estar representadas por su media y su matriz de covarianza. Se asume que se conoce una distribución *a priori* para las variables del estado \mathbf{x} , al que se llama $\mathbf{x}^b \in \mathbb{R}^N$, estado del *background*, con matriz de covarianza de errores asociada $\mathbf{B} \in \mathbb{R}^{N \times N}$. Las observaciones para ese instante están representadas por el vector $\mathbf{y} \in \mathbb{R}^M$ y la matriz de covarianza de los errores de las observaciones $\mathbf{R}_k \in \mathbb{R}^{M \times M}$. La ponderación del *background* y las observaciones viene representada en la Ecuación (10).

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{W}(\mathbf{y} - \mathcal{H}(\mathbf{x}^b)) \quad (10)$$

\mathbf{x}^a : estado analizado.

\mathbf{x}^b : estado del *background*.

\mathbf{y} : observaciones.

El primer pronóstico del estado o *background*, puede venir dado por un modelo de pronóstico desde un estado pasado. Para poder compararlo con las medidas, el pronóstico del modelo se interpola a la ubicación de las observaciones y, si las variables medidas son diferentes, se convierte de las variables del modelo a las variables observadas (como las radiaciones satelitales o las reflectividades del radar). Por lo tanto, la primera aproximación a las observaciones es $\mathcal{H}(\mathbf{x}^b)$, donde \mathcal{H} es el operador de observación que realiza la interpolación y transformación necesarias de las variables del modelo al espacio de observación (Kalnay et al., 2003).

La diferencia entre las observaciones y la primera aproximación del modelo $\mathbf{y} - \mathcal{H}(\mathbf{x}^b)$ se denota "incrementos de observación" o "innovaciones". El análisis \mathbf{x}^a se obtiene al agregar las innovaciones al pronóstico del modelo (primera aproximación) con pesos \mathbf{W} que se determinan con base en las covarianzas de error estadístico estimadas del pronóstico y las observaciones (Kalnay et al., 2003).

Considerando la combinación que obtenga la menor covarianza de error para el análisis (matriz \mathbf{P}^a), se obtiene la matriz de pesos \mathbf{W} que determina el análisis en la interpolación óptima, Ecuación (11). En la fórmula, el operador \mathcal{H} puede ser no lineal, pero \mathbf{H} es el tangente lineal de \mathcal{H} .

$$\mathbf{W} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}^{-1})^{-1} \quad (11)$$

La matriz de pesos W pasa a denominarse matriz de ganancia de Kalman: K y la matriz de covarianza de error del análisis que se obtiene es:

$$P^a = (I - KH)B \quad (12)$$

Las aplicaciones del algoritmo de interpolación óptima exigen de la especificación de las matrices BH^T y HBH^T . Esto limita los usos de este método a operadores de observación H muy simples correspondientes a interpolaciones sencillas. Así también, la matriz de covarianza de los errores del *background* B para este método es una simplificación constante que no se actualizará en el siguiente paso de tiempo y la ganancia K se modifica aún más para tener una estructura de bloque mediante el uso de innovaciones solo en pequeñas regiones alrededor de los puntos de la malla (Lahoz et al., 2010).

II.1.3.5. Filtro de Kalman (Kalman Filter)

El problema de filtrado (Figura 21, panel central) es el más común en aplicaciones geofísicas, y se caracteriza por un procesamiento secuencial, en el que las observaciones se utilizan a medida que están disponibles (Carrassi et al., 2018).

El prototipo de análisis secuencial, *Kalman Filter* (Figura 22), es el más extendido en su uso. En general, las observaciones se agrupan y asimilan para corregir las condiciones iniciales, que proporciona la simulación anterior, previo a cada corrida del modelo. Esto hace que se genere un estado del océano que no es continuo y no es consistente con la dinámica del océano o las condiciones de borde en los tiempos de análisis (Schiller and Brassington, 2011).

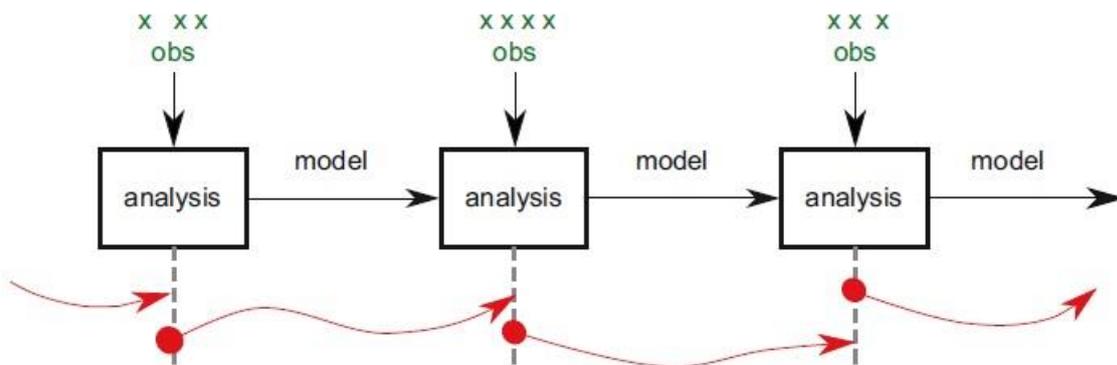


Figura 22 - Filtro de Kalman, análisis secuencial. Extraído de Schiller and Brassington (2011).

En *Kalman Filter* se tienen las ecuaciones (13) a (17) para la obtención de la media y covarianza del error del estado pronosticado y del estado analizado. Cada asimilación de datos se realiza en dos pasos: pronóstico y análisis (Carrassi et al., 2018).

El primer paso, de pronóstico, aplica el modelo al estado de análisis del tiempo anterior x_{k-1}^a , con lo que se obtiene el estado pronosticado por el modelo, el *background*, x^b (se adopta en la nomenclatura x^f , para indicar que se trata de un pronóstico o *forecast*). Aplicando el modelo a la matriz de covarianza de los errores del estado de análisis del tiempo anterior, P_{k-1}^a , se obtiene la matriz de covarianza para los errores del *background* o B (pasa a llamarse matriz de covarianza de los errores del estado pronosticado P^f). El paso de análisis actualiza la matriz de

ganancia de Kalman y obtiene el estado de análisis y la matriz de covarianza de errores del estado de análisis.

El *Kalman Filter* asume la simplificación de modelo lineal \mathcal{M} , (se adopta la nomenclatura \mathbf{M} para indicar que se puede escribir en forma de matriz $\mathbf{M} \in \mathbb{R}^{N \times N}$ y multiplicará al vector de estado y a la matriz \mathbf{P}_{k-1}^a).

Paso de pronóstico:

$$\mathbf{x}_k^f = \mathbf{M}_{k:k-1} \mathbf{x}_{k-1}^a \quad (13)$$

$$\mathbf{P}_k^f = \mathbf{M}_{k:k-1} \mathbf{P}_{k-1}^a \mathbf{M}_{k:k-1}^T + \mathbf{Q}_k \quad (14)$$

Paso de análisis:

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (15)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f) \quad (16)$$

$$\mathbf{P}_k^a = (\mathbf{I}_k - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^f \quad (17)$$

Tomando como entradas las matrices \mathbf{Q}_k , \mathbf{R}_k , \mathbf{H}_k y $\mathbf{M}_{k:k-1}$ para $k \geq 1$ y la condición inicial para la media y la covarianza de error : $\mathbf{x}_0^a = \mathbf{x}_0$, $\mathbf{P}_0^a = \mathbf{P}_0$, el *Kalman Filter* es capaz de obtener la media y la matriz de covarianza para todo tiempo $k \geq 1$ (Carrassi et al., 2018).

Los métodos de asimilación de datos secuenciales como las ecuaciones del *Kalman Filter* son generalmente inviables para los sistemas de predicción oceánica, particularmente cuando la covarianza de análisis $\mathbf{P}^a \in \mathbb{R}^{N \times N}$ debe construirse o evolucionarse explícitamente. Afortunadamente, se han desarrollado muchos enfoques para manejar los operadores del álgebra lineal asociados o hacer aproximaciones subóptimas a las ecuaciones de filtro completas. La aplicación a modelos \mathcal{M} no lineales y operadores \mathcal{H} también no lineales son un tema clave para los que se desarrollan extensiones al *Kalman Filter* (Schiller and Brassington, 2011).

II.1.3.6. Los métodos variacionales 3D-Var y 4D-Var

Los métodos variacionales parten del teorema de Bayes y definen una función de costo $\mathcal{J}(\mathbf{x})$ cuya minimización global corresponde con la máxima verosimilitud del análisis.

Las resoluciones de los sistemas de ecuaciones resultantes en los métodos variacionales no son directas, si no que se realizan a partir de incrementos iterativos. Ver Carrassi et al. (2018) o Kalnay et al. (2003) para más detalle. De este modo, los métodos 3D-Var llegan a la misma formulación que se obtiene para *Optimal Interpolation*, y, con una formulación alternativa para

la matriz de ganancia de Kalman $K = [B^{-1} + H^T R^{-1} H]^{-1} H^T R^{-1}$, se obtiene el análisis de 3D-Var como se expresa en las ecuaciones (18) y (19).

$$x^a = x^b + \delta x \quad (18)$$

$$[B^{-1} + H^T R^{-1} H] \delta x = H^T R^{-1} (y - \mathcal{H}(x^b)) \quad (19)$$

La diferencia entre *Optimal Interpolation* y el enfoque 3D-Var está en el método de solución: en *Optimal Interpolation*, los pesos W se obtienen para cada punto de malla o volumen de malla, utilizando simplificaciones adecuadas. En 3D-Var, se realiza la minimización de $J(x)$ directamente, lo que permite una flexibilidad adicional y un uso global simultáneo de los datos (Kalnay et al., 2003). Notar que en 3D-Var, se asume que las observaciones, análisis y *background* son válidos en el mismo tiempo.

El método 4D-Var, ver Figura 23, añade el tiempo; en particular, para la asimilación se cuenta con observaciones en una ventana de tiempo que será la ventana de asimilación. Aparece el modelo \mathcal{M} para hacer evolucionar el estado hacia tiempos posteriores y el modelo adjunto, representado como M^T , que permite la integración del modelo hacia atrás en el tiempo. Para un modelo lineal M , el adjunto coincide con la matriz traspuesta M^T (Schiller and Brassington, 2011).

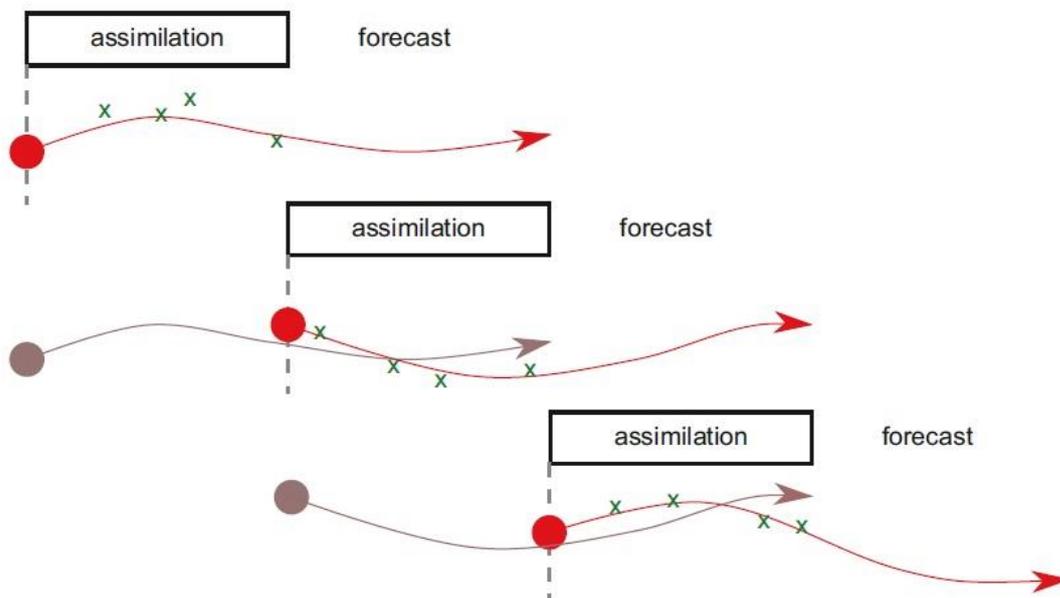


Figura 23 – 4D-Var. Extraído de Schiller and Brassington (2011).

En *Strong-Constraint* 4D-Var (Figura 24, panel central) se asume que el modelo es perfecto, (no aparece Q). La variable de control es el estado en el tiempo inicial. Las correcciones a partir de las observaciones se computan para el instante correspondiente a cada observación, pero se propagan hacia atrás para corregir las condiciones iniciales a través del modelo adjunto. El modelo pronostica el estado asimilado para toda la ventana a partir de las condiciones iniciales corregidas.

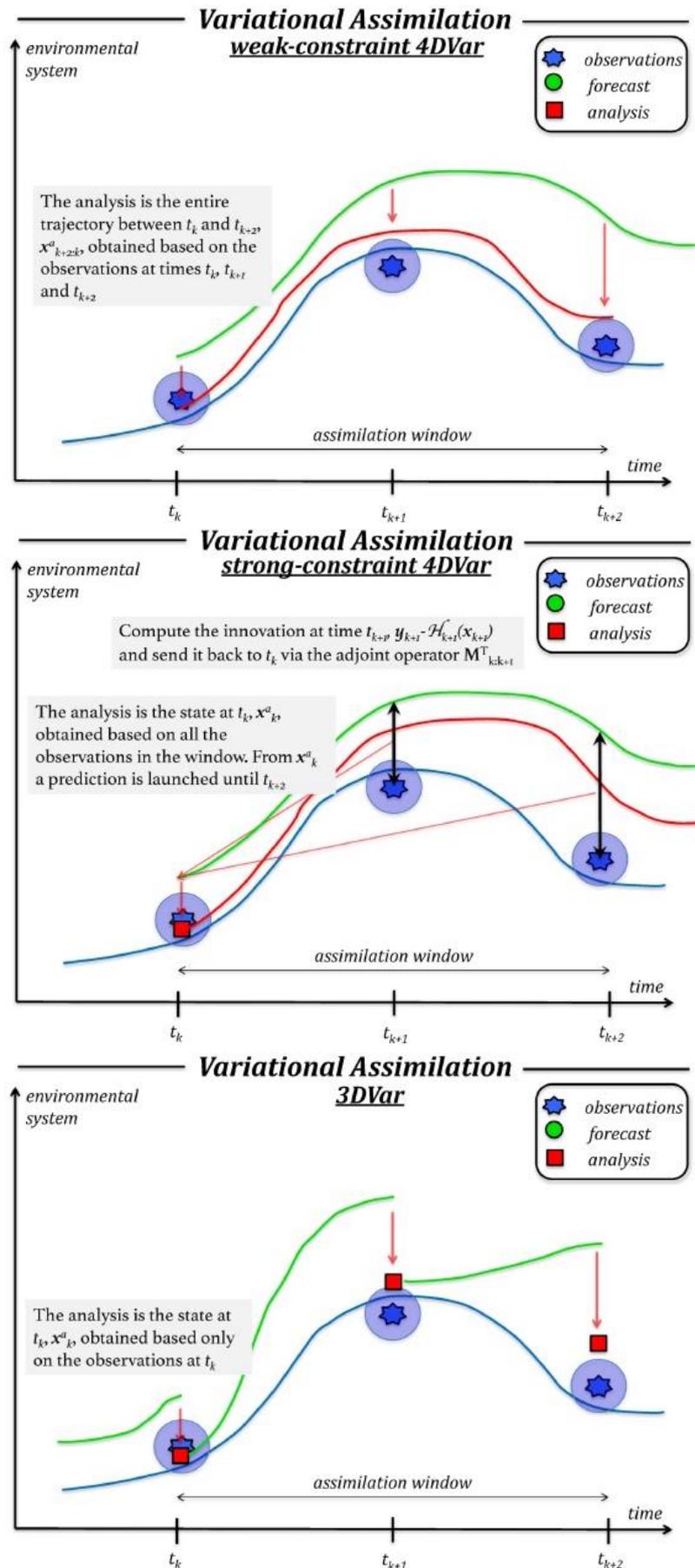


Figura 24 – Métodos variacionales. Weak-Constraint 4D-Var (panel superior), Strong-Constraint 4D-Var (panel central), 3D-Var (panel inferior). Extraído de Carrassi et al. (2018).

En *Weak-Constraint* 4D-Var (Figura 24, panel superior), se permite que el modelo no sea perfecto, la variable de control para la minimización es la trayectoria completa de x^a para todos los tiempos incluidos en la ventana de asimilación. La estimación en cada tiempo estará influenciada por las observaciones en la ventana de asimilación, tanto observaciones anteriores al tiempo como posteriores (Carrassi et al., 2018).

Un problema de los métodos variacionales es la hipótesis de covarianza del *background* B estacionaria, ya que no hay una ley para su evolución en el tiempo. Estos métodos tampoco computan explícitamente la incerteza del análisis representada por P^a para los otros métodos (Carrassi et al., 2018).

II.1.3.7. Filtro de Kalman Extendido (Extended Kalman Filter)

Una aproximación que hizo posible la implementación del *Kalman Filter* para su aplicación a las geociencias es el Filtro de Kalman Extendido (EKF por sus siglas en inglés). El EKF representa una expansión de primer orden del Filtro de Kalman y extiende su uso (de ahí la nomenclatura) para permitir operadores de las observaciones \mathcal{H} no lineales y también que el modelo que propaga el estado \mathcal{M} sea no lineal (Carrassi et al., 2018).

Al igual que el Filtro de Kalman, el Filtro de Kalman Extendido es secuencial. Si bien para el pronóstico del vector de estado se utiliza el modelo completo no lineal \mathcal{M} , la actualización de la matriz de covarianza de los errores del pronóstico utiliza el modelo tangente lineal M . La linearización se hace alrededor de la solución no lineal y por lo tanto será dependiente del estado.

En EKF se sigue asumiendo que todos los errores tienen distribución Gaussiana, pero al permitir la acción de una dinámica no lineal, estos no permanecerán Gaussianos y su exactitud acompañará el grado de no linealidad del modelo (Carrassi et al., 2018).

Los pasos secuenciales de pronóstico y análisis quedan de la siguiente manera (Kalnay et al., 2003):

Paso de pronóstico:

$$x_k^f = \mathcal{M}_{k:k-1}(x_{k-1}^a) \quad (20)$$

$$P_k^f = M_{k:k-1}P_{k-1}^aM_{k:k-1}^T + Q_k \quad (21)$$

Paso de análisis:

$$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1} \quad (22)$$

$$x_k^a = x_k^f + K_k \left(y_k - \mathcal{H}_k(x_k^f) \right) \quad (23)$$

$$P_k^a = (I_k - K_k H_k) P_k^f \quad (24)$$

A diferencia de los métodos variacionales, en el filtro de Kalman la matriz de covarianza de los errores del *background*, \mathbf{P}_k^f , sí que es evolucionada por la dinámica del modelo. Pero los cálculos se deben especificar para matrices de tamaño $N \times N$ para cada paso de tiempo. La dificultad de esto último, para problemas de grandes dimensiones, llevó a la aparición de Filtros de Kalman de rango reducido (*reduced-rank Kalman Filter*) para los que se realiza alguna transformación que reduzca el tamaño de las matrices $N \times N$ a tamaño $P \times P$ con $P < N$ (Kalnay et al., 2003).

II.1.3.8. Filtro de Kalman por Ensamblés (Ensemble Kalman Filter)

El principio en el que se basan los métodos por ensambles (EnKFs) es utilizar una muestra de P realizaciones del pronóstico $\mathbf{x}_{k,p}^f$ (p representa a un miembro del ensamble y va desde 1 hasta P) para estimar la covarianza de los errores del *background*, \mathbf{P}_k^f , directamente como la matriz de covarianza de los vectores del ensamble. La matriz de ganancia de Kalman, \mathbf{K}_k , puede ser estimada a partir del mismo ensamble, lo que permite calcular un ensamble de análisis, $\mathbf{x}_{k,p}^a$. A partir del ensamble de análisis, se puede estimar la covarianza del análisis, \mathbf{P}_k^a , y el proceso continúa (Schiller and Brassington, 2011).

En el Filtro de Kalman por Ensamblés se utilizan técnicas de estimación de rango reducido para aproximar los filtros clásicos y hacer la implementación factible en tiempo real (Lahoz et al., 2010). En este enfoque, un ensamble de ciclos de asimilación de datos P se lleva a cabo simultáneamente. Todos los ciclos asimilan las mismas observaciones reales, pero a fin de mantenerlas de manera totalmente independiente, se agregan diferentes conjuntos de perturbaciones aleatorias, $\mathbf{v}_{k,p}$, a las observaciones asimiladas en cada miembro del ensamble de asimilaciones de datos, $\mathbf{y}_{k,p} = \mathbf{y}_k + \mathbf{v}_{k,p}$ (Kalnay et al., 2003).

Con estos métodos, un ensamble que consiste en un pequeño número de vectores de análisis, $\mathbf{x}_{k,p}^a$, (P es mucho menor que el tamaño del estado N) se propaga simultáneamente por el modelo no lineal, \mathcal{M} , de un tiempo de observación al siguiente para proporcionar un ensamble de estados del *background*, $\mathbf{x}_{k,p}^f$. En este caso la forma de incluir un modelo imperfecto, no es mediante la matriz \mathbf{Q} en el cálculo de \mathbf{P}_k^f , sino mediante la adición a cada miembro del ensamble de una realización del error aleatorio $\boldsymbol{\eta}_{k,p}$. El ensamble del *background* se actualiza con las observaciones para dar un nuevo conjunto de vectores de análisis. El estado de análisis "óptimo" y su matriz de covarianza de error se determinan usando un filtro similar a los filtros clásicos (Lahoz et al., 2010).

Una ventaja de estos métodos es que no utilizan el tangente lineal del modelo ni del operador de las observaciones. La precisión de los estados estimados depende, sin embargo, de la propagación del ensamble, que debe ser suficiente para capturar el verdadero comportamiento del sistema (Lahoz et al., 2010).

Al igual que en la actualización de las matrices de covarianza de los errores del *background* y del análisis, en la actualización de la matriz de ganancia de Kalman, \mathbf{K}_k , también se evita la necesidad de matrices $N \times N$ como son el tangente lineal y el adjunto del operador de las observaciones aproximando $\mathbf{P}_k^f \mathbf{H}_k^T$ y $\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T$ por cálculos con los miembros del ensamble.

Los pasos secuenciales de pronóstico y análisis quedan de la siguiente manera, donde la barra superior representa el promedio del ensamble (Kalnay et al., 2003):

Paso de pronóstico:

$$\mathbf{x}_{k,p}^f = \mathcal{M}_{k:k-1}(\mathbf{x}_{k-1,p}^a) + \boldsymbol{\eta}_{k-1,p} \quad (25)$$

$$\mathbf{P}_k^f \approx \frac{\mathbf{1}}{P-1} \sum_{p=1}^P (\mathbf{x}_{k,p}^f - \overline{\mathbf{x}_k^f}) (\mathbf{x}_{k,p}^f - \overline{\mathbf{x}_k^f})^T \quad (26)$$

Paso de análisis:

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (27)$$

$$\mathbf{P}_k^f \mathbf{H}_k^T \approx \frac{\mathbf{1}}{P-1} \sum_{p=1}^P (\mathbf{x}_{k,p}^f - \overline{\mathbf{x}_k^f}) (\mathcal{H}_k(\mathbf{x}_{k,p}^f) - \overline{\mathcal{H}_k(\mathbf{x}_{k,p}^f)})^T \quad (28)$$

$$\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T \approx \frac{\mathbf{1}}{P-1} \sum_{p=1}^P (\mathcal{H}_k(\mathbf{x}_{k,p}^f) - \overline{\mathcal{H}_k(\mathbf{x}_{k,p}^f)}) (\mathcal{H}_k(\mathbf{x}_{k,p}^f) - \overline{\mathcal{H}_k(\mathbf{x}_{k,p}^f)})^T \quad (29)$$

$$\mathbf{x}_{k,p}^a = \mathbf{x}_{k,p}^f + \mathbf{K}_k (\mathbf{y}_{k,p} - \mathcal{H}_k(\mathbf{x}_{k,p}^f)) \quad (30)$$

$$\mathbf{P}_k^a \approx \frac{\mathbf{1}}{P-1} \sum_{p=1}^P (\mathbf{x}_{k,p}^a - \overline{\mathbf{x}_k^a}) (\mathbf{x}_{k,p}^a - \overline{\mathbf{x}_k^a})^T \quad (31)$$

El atractivo de este enfoque es que, en principio, puede aplicarse directamente a modelos lineales o no lineales. Además, incluso si los estadísticos del ruido del sistema no son Gaussianos, el campo de análisis satisface aproximadamente un criterio de varianza mínima, dentro de los límites de precisión de las estadísticas de muestra.

En la práctica surgen dos dificultades. Primero, que pueden ser necesarios una gran cantidad de miembros del ensamble para estimar con precisión elementos fuera de la diagonal de la matriz de covarianza de errores del *background* o pronóstico cuando la correlación entre las variables es pequeña, ya que en esos casos la correlación pasa a estar dominada por los errores de muestreo. Por esta razón, la covarianza de la muestra debe estar localizada para reducir correlaciones distantes. La operación de localización aumenta el rango efectivo de la covarianza. Consiste en elegir para cada punto de la grilla qué observaciones serán utilizadas (Lahoz et al., 2010), pero debe hacerse con una cuidadosa consideración de las correlaciones dinámicas que se quieren preservar (Schiller and Brassington, 2011). La otra dificultad principal es que los miembros del ensamble del *background* o pronóstico son no independientes después de que el Filtro Kalman se haya estado ejecutando. Esto puede contribuir a una pérdida de varianza y

bloqueo del filtro. Para evitarlo se han desarrollado varias estrategias para la inflación de covarianza y la reinicialización del filtro (Schiller and Brassington, 2011).

II.1.3.9. Métodos no Gaussianos. Los Filtros de Partículas

Los métodos variacionales y los Filtros de Kalman tienen dificultades para manejar problemas no lineales (Van Leeuwen et al., 2019). Se trata de soluciones basadas en aproximaciones Gaussianas que asumen que la función de distribución de los errores está completamente descrita por la media y la matriz de covarianza. Las operaciones involucradas en estas aproximaciones son computacionalmente exigentes debido a la necesidad de inversiones de matrices de covarianza en el paso de análisis, esto contrasta con la simplicidad del Teorema de Bayes que sirvió de base para su desarrollo (Carrassi et al., 2018).

Los filtros de partículas (PF, del inglés *Particle Filter*) son un tipo de métodos no Gaussianos que aproximan la función de distribución (pdf) por una distribución de una muestra (Van Leeuwen et al., 2019). Así, para una muestra de P partículas x_p (p representa a una partícula y va desde 1 hasta P) con peso para cada partícula w_p , la función de distribución de x queda representada como:

$$p(x) \approx \begin{cases} w_p & \text{si } x = x_p \\ 0 & \text{si } x \neq x_p \forall p \end{cases} \quad (32)$$

Un enfoque directo para el problema de la asimilación de datos sería usar métodos de Monte Carlo y extraer P muestras de la ecuación de Bayes con la expectativa de que en el límite asintótico, es decir, $P \rightarrow \infty$, se estimaría adecuadamente la densidad condicional $p(x|y)$. En un contexto secuencial, este enfoque se denomina filtro de partículas (PF) o Monte Carlo secuencial (Carrassi et al., 2018).

En el filtro de partículas secuencial estándar se tiene como punto de partida de una muestra de P partículas, inicialmente con pesos iguales que suman 1, $x_{p,k-1}$ que representan la función de distribución (pdf) de las variables de estado (pdf *prior* o *background*). Por conveniencia a este tiempo se le llama $k - 1$. Estas partículas se hacen evolucionar a través del modelo no lineal imperfecto (para cada partícula se añadirá una realización del error aleatorio) hasta el tiempo de las observaciones, tiempo k . La asimilación de datos se da al ponderar estas partículas a través de la verosimilitud o *likelihood*, $p(y|x)$, y se incluye una normalización de los pesos por su sumatoria para que mantengan el tamaño original (Van Leeuwen et al., 2019).

$$x_{p,k} = \mathcal{M}(x_{p,k-1}) + \eta_{p,k} \quad (33)$$

$$w_{p,k} = \frac{p(y_k|x_{p,k})}{\sum_{j=1}^P p(y_k|x_{j,k})} \quad (34)$$

$$p(x_k|y_k) \approx \begin{cases} w_{p,k} & \text{si } x = x_p \\ 0 & \text{si } x \neq x_p \forall p \end{cases} \quad (35)$$

Este esquema se representa en la Figura 25. La variable modelo corre a lo largo del eje vertical, el peso de cada partícula corresponde al tamaño de los puntos en este eje. El eje horizontal denota tiempo, con observaciones en un intervalo de tiempo de 10 unidades de tiempo. En este esquema las partículas de mayor verosimilitud tienen pesos grandes que aumentan a cada paso de tiempo dejando al resto de partículas con pesos insignificantes y haciendo que el filtro degenera en una sola partícula en pocos pasos de tiempo.

Para evitar mantener miembros con pesos insignificantes y otros con pesos enormes se introducen las técnicas de muestreo por importancia y remuestreo “resampling”. El método SIR aparece representado en la Figura 26 y mantiene partículas de igual peso, replicando las que obtuvieron mayor verosimilitud, pero perturbando a cada réplica con una realización distinta del error del modelo.

Estos filtros requieren de un enorme número de partículas para su convergencia en sistemas geofísicos, pero los filtros de partículas desarrollados en la actualidad han hecho un gran progreso a partir del uso de ideas del transporte (óptimo), el uso de localización y estrategias inteligentes de muestreo adaptativo. Además, se han desarrollado híbridos potentes entre filtros de partículas y filtros de Kalman por ensambles y métodos variacionales. De todo esto se puede encontrar una revisión actualizada en Van Leeuwen et al. (2019).

II.1.3.10. La estimación de las matrices de covarianza de errores Q y R

La confianza que se tenga en el *background* o *prior* proporcionado por el modelo y en las observaciones es lo que determina si el análisis será más parecido al estado proporcionado por el modelo o a la interpolación de las observaciones. En la mayoría de las técnicas de asimilación de datos actuales que consideran errores aditivos como ruidos Gaussianos, estas confianzas vienen representadas por la matriz de covarianza de los errores del *background*, B ó P^f , y por la matriz de covarianza de los errores de las observaciones R . Además, son estas matrices las que determinan la confianza que se tiene en el análisis, representada por la matriz de covarianza de los errores del análisis, P^a .

La matriz de covarianza de los errores del *background* puede ser una matriz constante, B , como en los métodos variacionales y en interpolación óptima, que deberá ser correctamente estimada, o puede ser evolucionada en el tiempo por el modelo como en los Filtros de Kalman, P^f . En este caso, para su determinación en cada paso de tiempo se hace evolucionar la matriz P^a a través del modelo y se añade la matriz de covarianza de los errores del modelo, Q . El éxito de la asimilación de datos depende de lo buena que sea la estimación que se haga de Q y R .

En Tandeo et al. (2018), se ejemplifica de manera ilustrativa para un problema autorregresivo de una variable que una buena estimación de Q nunca compensa una mala estimación de R , y viceversa; y la influencia que tiene la calidad de las estimaciones de Q y R en la matriz P^a y en el estado analizado. En la Figura 27 se ve el caso para cuando la matriz Q estimada es la verdadera pero la estimación de R es demasiado optimista para las observaciones, siendo un 10% de la verdadera R . Los puntos negros representan las observaciones y la línea roja el estado real, mientras que la línea negra representa el estado analizado. La confianza en el análisis se

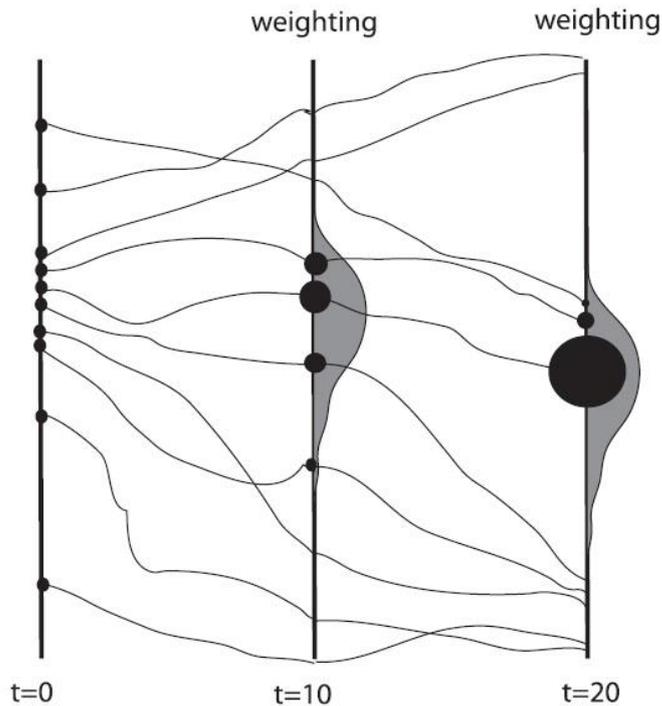


Figura 25 – Esquema filtro de partículas estándar. Extraído de Van Leeuwen (2009).

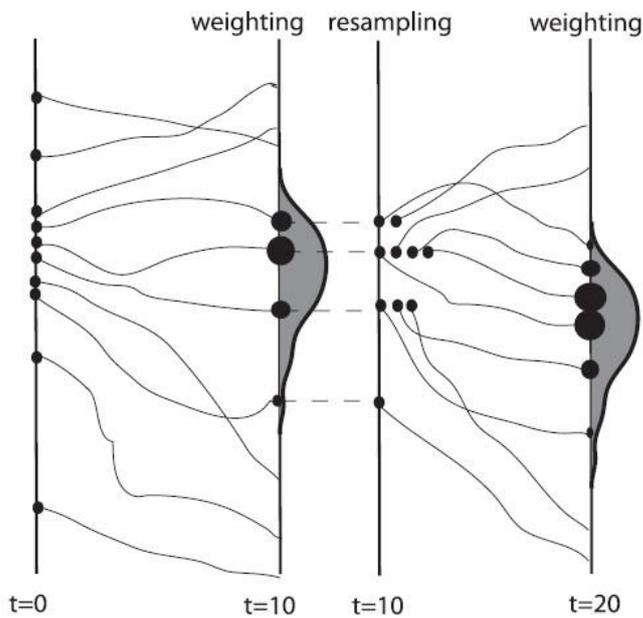


Figura 26 - Esquema del filtro de partículas con remuestreo, también llamado SIR (Sequential Importance Resampling). Todas las partículas tienen el mismo peso en el tiempo cero. En el momento 10, las partículas se ponderan de acuerdo con la verosimilitud (likelihood) y se vuelven a muestrear para obtener un conjunto de igual peso. Extraído de Van Leeuwen (2009).

representa mediante la franja gris. Se puede ver que la infraestimación de R genera que el análisis se acerque demasiado a las observaciones y que se tenga excesiva confianza en el resultado obtenido, franja gris muy estrecha. En este caso, se obtiene que la cobertura proporcionada por el análisis es de un 50% (siendo el óptimo, para cuando ambas matrices se estimaron correctamente, de un 95%) y un error RMSE de 0.85 (el óptimo es 0.71).

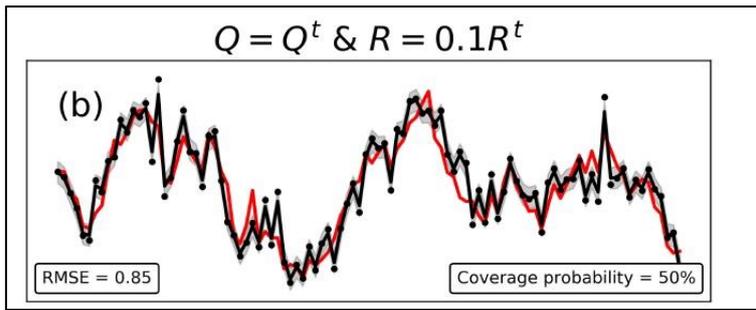


Figura 27 – Asimilación de datos para modelo autorregresivo. Estimación de Q y R . Extraído de Tandeo et al. (2018).

Para su aplicación a la asimilación de datos se han desarrollado metodologías robustas que estiman de manera simultánea las matrices Q y R . En la mayoría de ellas se hace uso de las innovaciones que son las diferencias entre las observaciones y los estados pronosticados por el *background*. En Tandeo et al. (2018) se realiza una revisión de estos métodos para las técnicas de asimilación por ensambles.

Las metodologías desarrolladas se pueden dividir en dos grupos atendiendo al criterio estadístico del que se sirven: los métodos basados en los momentos, en los que se asume que los momentos muestrales (media y covarianza) son iguales a los teóricos y los métodos basados en el *likelihood* o verosimilitud de las observaciones.

Teniendo en cuenta su definición, la matriz Q cuantifica las deficiencias del modelo para representar la física subyacente y la matriz R incluye los errores instrumentales y errores de representatividad. Se asume que no existe correlación entre los errores del modelo y los errores de las observaciones y además, se asume que tanto el modelo como las observaciones son insesgados (*unbiased*), algo que parece poco razonable para las aplicaciones a situaciones reales de asimilación de datos (Tandeo et al., 2018).

En los Filtros de Kalman por Ensamblados la matriz P^f se determina a través de los miembros del ensamble. De manera de añadir incerteza a su determinación a partir de un número finito de muestras se utiliza la inflación multiplicativa. Se trata de equiparar la estructura de Q a la de P^f , multiplicando esta última por un factor mayor que uno durante su determinación. De esta manera se incluye el error del modelo a través de una inflación multiplicativa en lugar de aditiva, que sería sumando Q . El problema de utilizar únicamente la inflación multiplicativa es que las correcciones sólo se pueden hacer en el subespacio definido por los miembros del ensamble y de manera óptima se añade la inflación aditiva para incorporar nuevas direcciones para corregir el modelo.

En los Filtros de Kalman por Ensamblados se utiliza la inflación multiplicativa para compensar los errores de muestreo y la inflación aditiva para los errores del modelo.

II.1.4. Evaluación del pronóstico oceanográfico operacional

La evaluación de los resultados de los sistemas de predicción oceánica consiste en conocer su capacidad para representar el mundo real. Esta evaluación es importante para informar a los usuarios cuan confiable son los productos del pronóstico y a su vez permite identificar áreas a

mejorar en los sistemas. La forma común de evaluar los resultados de los sistemas de pronóstico es a través de su comparación frente a observaciones y, por lo tanto, se deberá evaluar previamente la calidad de las observaciones y tener en cuenta que los modelos no son capaces de resolver todos los procesos incluidos en las observaciones. Como se dijo anteriormente, los resultados de un sistema de pronóstico se utilizan para generar subproductos de valor añadido (rutas de pesca, alertas de inundaciones, etc.), así que es importante que la evaluación incluya la amplia variedad de escalas de tiempo y de espacio que pueden llegar a utilizar las aplicaciones derivadas del sistema.

La calidad de los resultados de los sistemas oceanográficos operacionales está condicionada por: la calidad del modelo numérico (incluyendo resolución horizontal y vertical, parametrizaciones, etc.), la calidad de los campos de forzantes en superficie (o en el caso de un modelo acoplado, la calidad del modelo atmosférico), el esquema de asimilación de datos utilizado (tipo y número de observaciones que se utilizan en la asimilación, el esquema de asimilación en sí mismo, el control de calidad de las observaciones) y, en el caso de modelos regionales, las condiciones de contorno oceánicas cuyo impacto en los resultados puede ser muy significativo.

Algunos estadísticos que se utilizan para la comparación de los resultados de un sistema en una ubicación respecto a las observaciones disponibles en esa ubicación (adaptado de Hess et al., 2003) son:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Observado_i - Pronóstico_i)^2} \quad (36)$$

$$BIAS = \frac{1}{N} \sum_{i=1}^N (Observado_i - Pronóstico_i) \quad (37)$$

$$(38)$$

$$\begin{aligned} \text{Correlación Pearson} &= \frac{Cov(Pronóstico, Observado)}{s_x s_y} = \\ &= \frac{\frac{1}{N-1} \sum_{i=1}^N [(Pronóstico_i - \overline{Pronóstico}) \cdot (Observado_i - \overline{Observado})]}{\sqrt{\left[\frac{1}{N-1} \sum_{i=1}^N (Pronóstico_i - \overline{Pronóstico})^2 \right]} \cdot \sqrt{\left[\frac{1}{N-1} \sum_{i=1}^N (Observado_i - \overline{Observado})^2 \right]}} \end{aligned}$$

La diferencia de media, o sesgo (**BIAS**), se utiliza para evaluar la habilidad del modelo para representar la media observada. El error cuadrático medio (**RMSE**) se utiliza para evaluar la precisión. La correlación para evaluar la capacidad del modelo para reproducir los patrones de las observaciones.

La determinación de la habilidad (*skill*) de un modelo de pronóstico dependerá de la aplicación a la que dé servicio, un ejemplo de este tipo de estadísticos es el cálculo de la precisión del pronóstico frente a otro pronóstico de referencia. Con formulación modificada del error cuadrático medio se presenta el índice de acuerdo **Willmott skill**. El mayor valor del índice de acuerdo es **Willmott skill = 1**, y significa un acuerdo perfecto entre modelo y

observación, mientras que el menor valor indica un completo desacuerdo **Willmott skill** = 0. La utilización de **RMSE** es limitada ya que no se espera que los errores sigan una distribución normal (Hess et al., 2003), y por lo tanto como complemento se calculan la frecuencia central **CF(ζ)**, **POF(ζ)** frecuencia positiva de *outlier* y **NOF(ζ)** frecuencia negativa de *outlier*. La frecuencia central indica con qué frecuencia el error ($Pronóstico_i - Observado_i$) está dentro de los límites aceptables $\pm\zeta$.

$$\begin{aligned} \text{Willmott skill} &= & (39) \\ &= 1 - \frac{\frac{1}{N} \sum_{i=1}^N (Pronóstico_i - Observado_i)^2}{\frac{1}{N} \sum_{i=1}^N (|Pronóstico_i - \overline{Observado}| + |Observado_i - \overline{Observado}|)^2} \end{aligned}$$

CF(ζ) Frecuencia Central. Porcentaje de los errores que se encuentran entre los límites $\pm\zeta$

POF(ζ) Frecuencia positiva de *outlier*. Porcentaje de los errores que son mayores que ζ

NOF(ζ) Frecuencia negativa de *outlier*. Porcentaje de los errores que son menores que $-\zeta$

MDPO(ζ) Máxima duración de *outlier* positivo. Un evento de *outlier* positivo consiste en 2 o más ocurrencias consecutivas de error mayor que ζ . El estadístico **MDPO** es la longitud de tiempo del evento más largo

MDNO(ζ) Máxima duración de *outlier* negativo. Un evento de *outlier* negativo consiste en 2 o más ocurrencias consecutivas de error mayor que ζ . El estadístico **MDPO** es la longitud de tiempo del evento más largo

Cuando se van a comparar varios pronósticos, una buena forma de poder visualizar varios aspectos de su comportamiento es a través del diagrama de Taylor, que grafica un punto por cada pronóstico en función de tres estadísticos: el coeficiente de correlación de Pearson, el error cuadrático medio (**RMSE**) y la desviación estándar.

Una de las técnicas de validación más aplicada a los resultados de niveles de la superficie del mar es la técnica de análisis de armónicos. Se realiza el análisis armónico de los resultados de nivel de agua en una ubicación y se compara con el análisis armónico de las observaciones disponibles en esa ubicación. Este análisis también se utiliza para comparar la componente de señal de la marea astronómica. Una vez identificada la señal de marea astronómica se pueden realizar comparaciones en las que previamente se elimine esta componente.

Para la validación de modelos utilizando mapas de datos con variabilidad espacial y temporal se utiliza el análisis de componentes principales (EOF, por sus siglas en inglés). Con esta técnica se identifican modos de variabilidad espacial, su evolución temporal y la importancia relativa de cada patrón. Los modos producidos por este análisis son ortogonales en el espacio y el tiempo y, además, el primer maximiza la varianza explicada sobre el conjunto de datos total. Estos modos no están vinculados físicamente y se derivan de análisis estadísticos. Los resultados del análisis EOF se pueden usar para comparar la variabilidad entre los campos del modelo y las observaciones y también para la reconstrucción de los campos del modelo o de las mediciones con solo algunos de los principales modos de variabilidad.

En sistemas que incluyen asimilación de datos, una verificación que permite comprobar que la asimilación está funcionando como es esperable consiste en comprobar que el análisis obtenido

tras la asimilación es más cercano a las observaciones. Así, siguiendo la nomenclatura del apartado anterior, se espera que la diferencia $\mathbf{y} - \mathcal{H}(\mathbf{x}^a)$ sea menor que $\mathbf{y} - \mathcal{H}(\mathbf{x}^b)$. También se puede realizar la evaluación mediante estadísticos que comparan con las observaciones, pero, se debe tener la precaución de que las observaciones que se estén asimilando no podrán ser utilizadas para este fin (algunos datos de observaciones que no están disponibles en tiempo real pueden ser utilizados en evaluaciones a posteriori).

Durante el desarrollo de un nuevo sistema de pronóstico oceánico es común que se realicen comparaciones con los sistemas existentes de manera de validar las mejoras del nuevo pronóstico. Esto se hace con simulaciones de períodos de varios años en el pasado. Cuando ya se ha pasado la etapa de validación puede pasar a ser operacional y entonces se realizan evaluaciones sistemáticas de verificación y también evaluaciones específicas para aplicaciones concretas que, junto con la validación, permiten a los usuarios conocer la precisión esperable del sistema.

Para que los usuarios sean capaces de comprender los resultados de los sistemas es esencial realizar un post-proceso de las salidas numéricas que sintetice la información. Esto puede incluir promedios en el espacio y en el tiempo, interpolación, o la generación de los subproductos de valor añadido relevantes para un uso concreto. Los datos post-procesados suelen ser evaluados directamente y es ahí donde se puede identificar el impacto que ha tenido el post-procesamiento.

II.2. Pronósticos oceánicos globales

II.2.1. Descripción de los pronósticos disponibles

Los modelos globales resuelven la hidrodinámica del océano global proporcionando pronósticos para los próximos días. A continuación se hace una breve reseña de los sistemas globales que operan en el marco del GODAE a partir de los reportes anuales del año 2019 (<https://www.godae-oceanview.org/documents/q/category/govst/system-reports/national-systems-reports-2019/>). En los apartados II.2.2 y II.2.3 se introducen los modelos que han sido considerados para esta investigación, Modelo HYCOM y Modelo COPERNICUS, por ser los que proporcionan pronósticos para la zona de estudio que podrían utilizarse como condiciones de contorno en un subescalado regional para el Río de la Plata. Estos modelos tienen una resolución horizontal suficiente para resolver de la turbulencia de mesoescala (*mesoscale eddies*) y dan pronósticos en la frecuencia de 1 dato por hora para los próximos días para la variable de altura de la superficie del mar.

La Agencia Meteorológica Japonesa (JMA, por sus siglas en inglés) utiliza el modelo de circulación general del océano MRI.COM para, cada 5 días, producir pronósticos estacionales (7 meses) para el océano global en su sistema MOVE/MRI.COM-G2 y pronósticos regionales de manera diaria en MOVE/MRI.COM-WNP y MOVE/MRI.COM-Seto. Los pronósticos regionales alcanzan los 30 días de rango para WNP (*Western North Pacific*) y los 10 días para Seto (*Seto Inland Sea*). La asimilación de datos incluye datos de temperatura y salinidad de sistemas de observación *In-Situ*: VOS XBTs, Argo y datos de amarres tropicales; datos de los altímetros satelitales disponibles y además, los sistemas regionales asimilan SST (temperatura de la

superficie del mar) de varias fuentes: satelital, barcos y boyas. En WNP también se asimila concentración del hielo marino. El modelo es tridimensional con coordenada vertical híbrida (z-sigma) y malla horizontal de 1° de longitud y 0.5° de latitud y resuelve las variables hidrodinámicas temperatura, salinidad y velocidad. El sistema de asimilación empleado se llama 3DVAR-IAU. La agencia también cuenta con varios productos de reanálisis y para los pronósticos estacionales se encuentra en desarrollo un sistema de asimilación acopada océano-atmósfera. Se puede encontrar más información y la forma de acceso a los productos en <https://ds.data.jma.go.jp/gmd/goos/data/rtrdb/jma-pro.html>.

El Centro Europeo de Previsiones Meteorológicas a Medio Plazo (ECMWF, por sus siglas en inglés) cuenta con varios sistemas acoplados para los que se realizan simulaciones del océano. El producto ORAS5 (*Ocean ReAnalysis System 5*) forma parte del sistema operacional OCEAN5, que es un sistema acoplado océano-hielo_marino; ORAS5 se utiliza como condición inicial para los demás sistemas de pronóstico de mediano y largo plazo. HRES, ENS, Extended-ENS y S5 son sistemas de pronóstico acoplado atmósfera-océano-hielo_marino-oleaje-tierra. Estos sistemas utilizan el modelo de circulación general del océano NEMO V3.4 (*Nucleus for European Modelling of the Ocean*) de resolución horizontal 1/4° y 75 capas verticales para producir pronósticos de alcance desde 10 días hasta 7 meses con una resolución para la atmósfera desde unos 9 km hasta aproximadamente 36 km; los resultados son determinísticos para HRES y un ensamble de 51 miembros para el resto. Los pronósticos de los modelos oceánicos tienen un uso exclusivamente interno y no se distribuyen. Se puede encontrar mayor información sobre los productos acoplados en <https://www.ecmwf.int/en/forecasts/documentation-and-support>.

El servicio meteorológico nacional del Reino Unido, Met Office, opera el modelo FOAM (*Forecasting Ocean Assimilation Model*) para producir pronósticos para el océano global con 1/4° de resolución horizontal y 1/12° para las aplicaciones regionales del Norte Atlántico, el Océano Índico y el Mediterráneo. El sistema cuenta con 75 capas verticales para el dominio global y 50 capas en los regionales y se utiliza el modelo de circulación general del océano NEMO V3.6. Este modelo se usa como parte de un sistema océano-atmósfera (GLO-CPL) y también es utilizado por la Marina Real del Reino Unido para varias aplicaciones como modelos acoplados físico-biogeoquímicos. Utilizando los datos observados disponibles de SST del grupo GHRSSST (www.ghrsst.org) se genera OSTIA (*Operational Sea Surface Temperature and Ice Analysis*), un producto de mapas diarios de temperatura de la superficie del mar sin zonas con datos faltantes para el océano global. OSTIA cuenta con una resolución horizontal de 0.05° y para su generación se utilizan los datos disponibles *in-situ* y satelitales de radiómetros infrarrojos y microondas. También de manera diaria se genera GMPE (siglas de GHRSSST *Multi-Product Ensemble*) para analizar e inter-comparar los distintos análisis y productos observacionales de SST que son generados por varios grupos a partir de los datos proporcionados por GHRSSST; se publican estimadores de la incerteza de las distintas fuentes para una malla global de resolución horizontal 0.25°. Los resultados de FOAM pueden ser solicitados para investigación pero no están abiertos al público. Los productos GLO-CPL, OSTIA y GMPE están disponibles para los usuarios del Servicio de la Unión Europea de Monitoreo del Ambiente Marino, COPERNICUS (CMEMS, por sus siglas en inglés) en <http://marine.copernicus.eu>.

El Centro Nacional de Pronósticos del Ambiente Marino (NMEFC) de China genera pronósticos diarios de la hidrodinámica del océano global para los próximos 7 días con resolución horizontal de 1/12° utilizando el modelo NEMO V3.6 y para aplicaciones regionales se utiliza el modelo ROMS (*Regional Ocean Modeling System*) y se aumenta la resolución. Las variables tridimensionales horarias que es capaz de resolver el sistema incluyen campos de temperatura

y salinidad del océano, altura de la superficie del mar, velocidad, densidad del hielo marino. Los dominios regionales que se estudian son el mar de Bohai, los mares de China, el noroeste del Pacífico y el norte de la India. Los resultados se manejan únicamente de manera interna. En <http://www.nmefc.cn/nr/cont.aspx?itemid=2&id=7149> se puede encontrar una descripción del sistema y en <http://www.nmefc.cn/hailiu/quanqiu.aspx> se encuentra un visualizador web simplificado de los pronósticos globales.

En la República de Corea del Sur, la agencia KHOA (*Korea Hydrographic and Oceanographic Agency*) y el Instituto KIOST (*Korea Institute of Ocean Science and Technology*) operan sistemas de pronóstico del océano. Así, KHOA, de manera diaria, opera y evalúa frente a datos medidos el Sistema Coreano de Observación y Pronóstico del Océano (KOOFS por sus siglas en inglés) para dominios regionales, sub-regionales, áreas costeras y puertos. El instituto KIOST con frecuencia semanal genera predicciones para el noroeste del océano pacífico.

La Red brasileña de Modelación y Observación Oceanográfica (REMO) tomando como base el modelo HYCOM (*HYbrid Coordinate Ocean Model*, Estados Unidos) genera de manera diaria pronósticos regionales con resolución horizontal desde los $1/4^\circ$ hasta $1/24^\circ$ y alcance de hasta 6 días. Sólo se distribuyen los productos de resolución horizontal más baja y, además, se pueden visualizar figuras de los pronósticos de los campos de SST, SSH, corrientes en superficie y secciones verticales de temperatura y velocidades en la web del Centro Hidrográfico de la Marina Brasileña, CHM, <https://www.marinha.mil.br/chm/dados-do-smm-modelagem-numerica-tela-de-chamada/modelagem-numerica>.

El Consorcio para la Estimación de la Circulación y el Clima del Océano (ECCO, por sus siglas en inglés, <http://ecco.jpl.nasa.gov>) produce reconstrucciones del océano en tiempo pasado similares a reanálisis pero con la importante diferencia de que son dinámica y cinemáticamente consistentes por lo que permiten cálculos de balance de manera precisa. Con la última actualización, ECCOv4r4, se tienen simulaciones desde 1992 hasta 2017 de la hidrodinámica del océano cuasi-global (temperatura, salinidad, velocidades) con la finalidad de comprender los cambios en el océano que incluyen aumento del nivel del mar, pérdida de hielo marino, eventos del Niño, y los cambios en el ciclo del carbono y el agua.

El centro CMCC italiano (<https://www.cmcc.it/data-services-and-products/forecasting-systems>) a través del sistema GOFS16 (*Global Ocean Forecast System*) produce a diario desde julio de 2017 pronósticos para los próximos 7 días de las variables de la hidrodinámica del océano a nivel global con una resolución horizontal de $1/16^\circ$ y 98 capas verticales. Para ello se utiliza el modelo de circulación general del océano NEMO V3.6 acoplado con el modelo de hielo marino LIM2. En la web <http://gofs.cmcc.it/> se publican mapas de los resultados en 2D y la Unidad de Diseminación de CMCC dispone los datos bajo demanda. Con una resolución horizontal de $1/24^\circ$ y 141 niveles verticales se generan pronósticos para el Mar Mediterráneo y parte del Atlántico para los próximos 10 días.

El centro francés Mercator Ocean involucra a los principales centros europeos de oceanografía operacional y gestiona el Servicio de la Unión Europea de Monitoreo del Ambiente Marino, COPERNICUS, CMEMS. En <https://marine.copernicus.eu/services-portfolio/access-to-products/> se encuentran los 174 productos disponibles para el océano. Se incluyen modelos globales y regionales, observaciones *In-Situ* pasadas y también procesadas en tiempo casi-real, productos de reanálisis, análisis y pronóstico (actualizaciones diarias y semanales, con alcance hasta los 10 días), en 2 dimensiones y en 3 dimensiones, obteniéndose variables físicas de la hidrodinámica del océano (temperatura, salinidad, altura de la superficie del mar, velocidad, concentración y espesor del hielo marino, etc.) así como variables biogeoquímicas (clorofila a, fitoplancton, oxígeno disuelto, nitrato, fosfato, PH, etc.), viento en superficie y oleaje. Se publican 4 productos

de pronóstico del océano global: PHY_001_024, WAV_001_027, PHY_CPL_001_015 y BIO_001_028. El primero, PHY_001_024, es el modelo operacional de pronóstico de las condiciones de la hidrodinámica del océano que será introducido en la sección II.2.3, de manera diaria se pronostican 10 días para una malla de resolución horizontal $1/12^\circ$ y 50 niveles verticales; el segundo, WAV_001_027, proporciona pronósticos de oleaje a diario para los próximos 5 días; el tercero, PHY_CPL_001_015 es el sistema océano-atmósfera que se nombra en la reseña referente al servicio meteorológico de Reino Unido; el cuarto, BIO_001_028, es un sistema de pronóstico para los próximos 10 días de las variables biogeoquímicas del océano con actualización semanal y una resolución horizontal de $1/4^\circ$ y 50 capas verticales. Recientemente se ha puesto a disposición un muy completo visualizador web del catálogo de productos publicados en CMEMS <https://cmems.lobelia.earth/>.

Los Centros Nacionales de Predicción Ambiental (NCEP) de los Estados Unidos generan pronósticos a diario para los próximos 8 días través del sistema RTOFS (*Real Time Ocean Forecast System*), este sistema será descrito en la sección II.2.2 y tiene una malla de resolución horizontal $1/12^\circ$ y 41 capas verticales. También de manera operacional se generan dos sistemas de pronóstico para huracanes acoplados a modelos regionales del océano: el sistema HWRF (*Hurricane Weather Research and Forecast*) y el sistema HMON (*Hurricanes in a Multiscale Ocean-coupled Nonhydrostatic*).

Otros sistemas de pronóstico regionales incluidos en GODAE son el sistema TOPAZ del Centro de Pronóstico Marino del Ártico (MFC por sus siglas en inglés) que genera pronósticos para el Océano Atlántico Norte y el Ártico; el sistema CONCEPTS de Canadá para el Norte Atlántico y el Ártico, que llega hasta resolución $1/36^\circ$ en algunas regiones y cuenta con un pronóstico global de $1/4^\circ$ de resolución; y el sistema INDOFOS del Centro de Servicios de Información Oceánica de la India (INCOIS) para el Océano Índico.

II.2.2. Detalles del modelo de pronóstico HYCOM

El modelo Global RTOFS se basa en una configuración global de HYCOM (*HYbrid Coordinate Ocean Model*, <https://www.hycom.org/>) desarrollado por el NOAA (Administración Nacional Oceánica y Atmosférica de EEUU), el Servicio Meteorológico Nacional de EEUU en asociación con la Marina de EEUU como parte del GODAE. Es un sistema de pronóstico en tiempo real de circulación general oceánica que proporciona datos actualizados a diario del estado del océano (2 días de análisis) y pronóstico para los próximos 8 días.

Se trata de un modelo 3D con una resolución horizontal de $1/12^\circ$ ($\sim 0.0833^\circ$); esta resolución horizontal permite resolver la turbulencia de mesoescala. En vertical tiene coordenada híbrida con 41 capas de modo de utilizar cada tipo de coordenada dónde sus propiedades son más ventajosas. Así, se utiliza coordenada z para las zonas costeras, que incluyen profundidades desde los 5 m, pasando por coordenada sigma para profundidades medias y llegando a una coordenada isopícnica (dependiente de la presión) para el océano abierto. La batimetría que utiliza es una mejora de los datos de NRL DBDB2 de la armada de los EEUU que incorpora varios paquetes de datos batimétricos de alta resolución.

El sistema de pronóstico es forzado con los campos de predicción meteorológica global del NCEP. Incluye la asimilación de perfiles de salinidad y temperatura *in situ* procedentes de una variedad de fuentes, así como la asimilación de datos de detección remota de temperatura, altura de la superficie del océano y concentración de hielo marino (<http://polar.ncep.noaa.gov/global/about/>).

Los campos 3D se publican con resolución temporal de 1 dato cada 24 horas para los próximos 8 días para las variables de salinidad, temperatura, velocidad este y norte. Los resultados de las predicciones publicados con resolución temporal de 1 dato por hora son únicamente 2D y para los próximos 3 días e incluyen, a nivel global, variables de salinidad, temperatura, velocidad este y norte y altura de la superficie del océano sin incorporar la marea astronómica. Estos datos son publicados el día de la simulación y se encuentran disponibles en la web hasta el final del día siguiente. El modelo cuenta con un visualizador web para los pronósticos de los campos globales en <https://polar.ncep.noaa.gov/global/nc/>. En la Figura 28 se muestra un ejemplo de nivel proporcionado por el visualizador web y en la Tabla 2 se resumen las características de este modelo global.

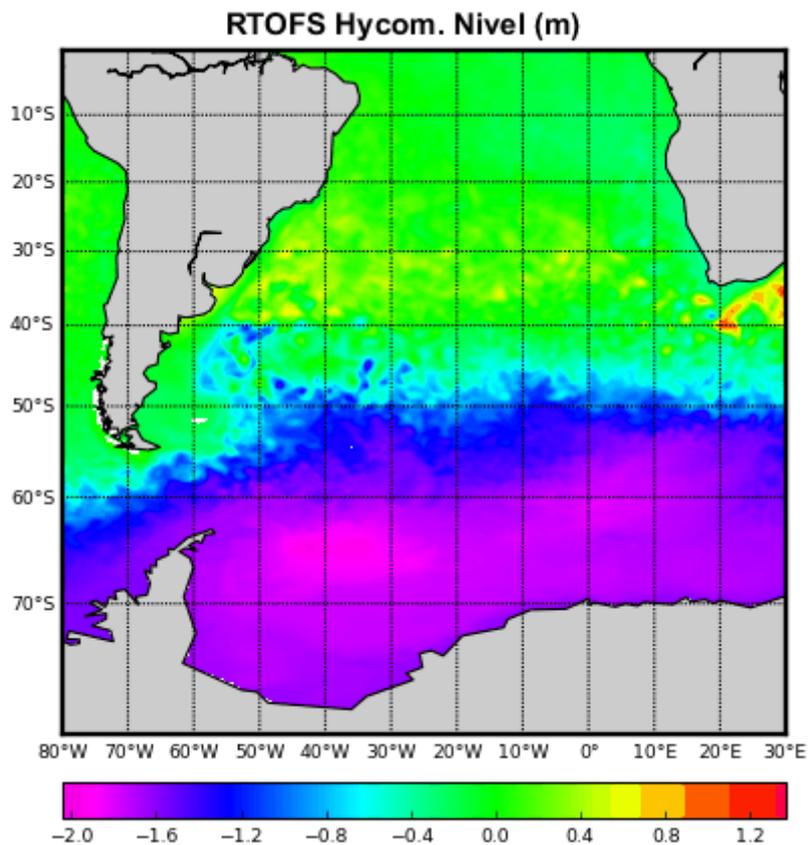


Figura 28 - RTOFS HYCOM. Visualizador web para el nivel (m). Extraído de <https://polar.ncep.noaa.gov/global/nc/>

Modelo Global RTOFS HYCOM	
Desarrollador	Administración Nacional Oceánica y Atmosférica de EEUU (NOAA)
Referencia	http://polar.ncep.noaa.gov/global/about/
Pronóstico con frecuencia 1dato/hora	3 días de pronóstico GMT0 2D, resolución horizontal 1/12º(~0.0833º)
Variables	Salinidad, temperatura, velocidad este y norte y altura de la superficie del océano (sin incorporar marea astronómica)
Forzante atmosférico	Campos de predicción meteorológica global del NCEP
Batimetría	Mejora de los datos de NRL DBDB2

Asimilación	Perfiles de salinidad y temperatura in situ (variedad de fuentes) Datos de detección remota de temperatura, altura de la superficie del océano y concentración de hielo marino
--------------------	---

Tabla 2 – Características principales del modelo de pronóstico global RTOFS HYCOM

II.2.3. Detalles del modelo de pronóstico COPERNICUS

El modelo COPERNICUS, *Global Sea Physical Analysis and Forecasting*, GLOBAL_ANALYSIS_FORECAST_PHY_001_024, desarrollado por el Servicio de Monitoreo del Medio Ambiente Marino de la Comisión Europea. Es el sistema global de análisis y pronóstico de océanos operacional de Mercator que proporciona predicción horaria para los próximos 10 días actualizados a diario. También proporciona resultados de análisis para los últimos 14 días (*hindcast*) con actualización semanal y disponibilidad desde el año 2006.

Se trata de un modelo 3D con una resolución horizontal de $1/12^\circ$ ($\sim 0.0833^\circ$) y 50 niveles en vertical con resolución decreciente de forma que en los primeros 100 m cuenta con 22 capas. La batimetría utilizada en el sistema es una combinación interpolada de las bases de datos ETOPO1 y GEBCO8. El sistema es forzado por los campos atmosféricos del Sistema de Pronóstico Integrado ECMWF (Centro Europeo de Previsiones Meteorológicas a Medio Plazo). Los datos asimilados incluyen datos del altímetro satelital, perfiles de temperatura y salinidad *In-Situ*, así como la concentración de hielo marino satelital https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_PHY_001_024.

El sistema mantiene públicos datos de análisis desde 2006 y una predicción de 10 días. Para el acceso a los datos se requiere la identificación del usuario y contraseña y para la creación del usuario se solicitan datos básicos. Los campos en 3D se publican con resolución temporal de 1 dato cada 24 horas para los próximos 2 días para las variables de salinidad, temperatura, velocidad este y norte. Los resultados de las predicciones publicados con resolución temporal de 1 dato por hora se proporcionan para los próximos 10 días pero únicamente 2D e incluyen, a nivel global, variables de temperatura, velocidad este y norte y altura de la superficie del océano con respecto al nivel de reposo (sin tener en cuenta la marea astronómica). Se hace notar que el modelo COPERNICUS publica sus datos horarios a la hora GMT0 +30 min. En el visualizador web del catálogo de productos publicados en CMEMS se pueden encontrar, para este producto, campos de las variables en frecuencias horarias, diarias y medias horarias y mensuales <https://cmems.lobelia.earth/>. En la Figura 29 se muestra un ejemplo del visualizador para el nivel en el dominio global y en la Tabla 3 se resumen las características del modelo.

Modelo COPERNICUS GLOBAL_ANALYSIS_FORECAST_PHY_001_024	
Desarrollador	Servicio de Monitoreo del Medio Ambiente Marino de la Comisión Europea
Referencia	http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_PHY_001_024

Pronóstico con frecuencia 1dato/hora	10 días de pronóstico GMT0+30minutos 2D, resolución horizontal 1/12º(~0.0833º)
Variables	Temperatura, velocidad este y norte y altura de la superficie del océano (sin incorporar marea astronómica)
Forzante atmosférico	Campos atmosféricos del Sistema de Pronóstico Integrado ECMWF
Batimetría	Combinación interpolada de las bases de datos ETOPO1 y GEBCO8
Asimilación	Datos del altímetro satelital Perfiles de temperatura y salinidad in situ Concentración de hielo marino satelital

Tabla 3 – Características del modelo de pronóstico global COPERNICUS.

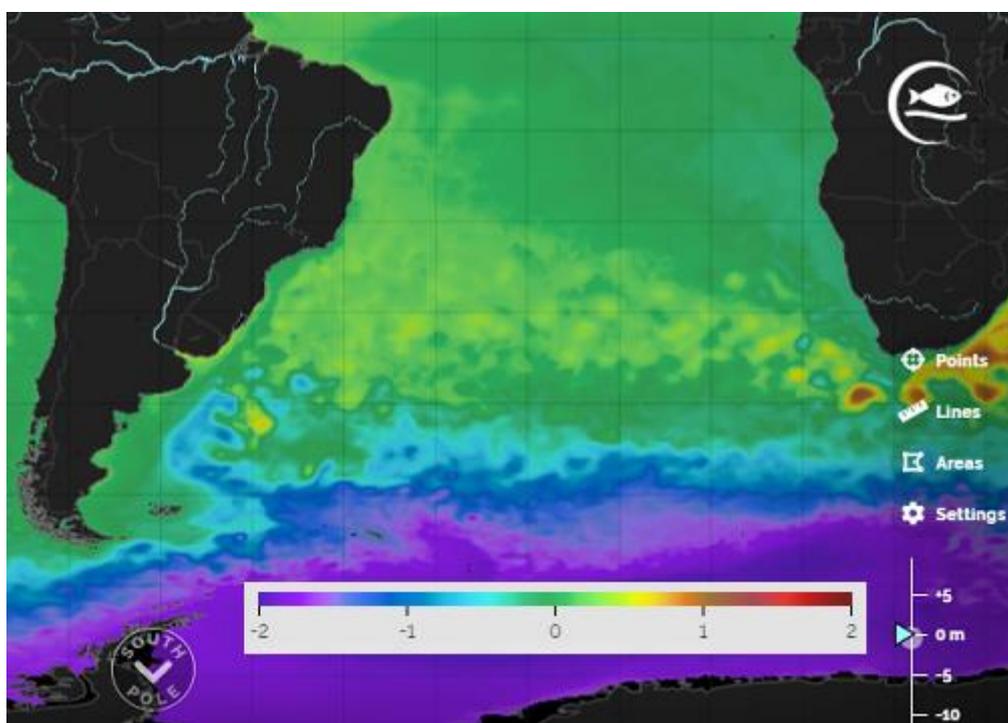


Figura 29 – COPERNICUS PHY 001 024. Visualizador web para el nivel (m). Extraído de <https://marine.copernicus.eu/>

II.3. Sistemas de pronósticos costeros

En este apartado se presentan varios antecedentes relevantes vinculados a pronósticos hidrodinámicos costeros, separados en pronósticos operacionales y no operacionales, los cuales se incluyen por tener aspectos relevantes para este trabajo. En II.3.1 se presenta el sistema Satbaŭtyk para el Mar Báltico que realiza pronósticos en 3D y se encuentra operativo con una página web que se actualiza a diario. Los pronósticos del estuario del río Gironda (II.3.2), Francia, utilizan el modelo TELEMAC-2D y sirven como dato de entrada para un sistema de alerta frente a inundaciones. En II.3.3 se hace una reseña del sistema SOFS que pronostica las condiciones 3D en la zona estuarina SSVBES, Brasil, con condiciones de contorno provenientes de un modelo

global. En II.3.4 se indica cómo se desarrolló un módulo de asimilación de datos para el modelo MOHID y su aplicación a los pronósticos del Río Tajo en Portugal. El apartado II.3.5 introduce los resultados de la experiencia en el Golfo de Nápoles en el Mar Tirreno, Italia, en su asimilación de datos satelitales y de radar de alta frecuencia. En los siguientes apartados se presentan dos estudios que incluyen a la zona del Río de la Plata. En II.3.6 se describe un esquema de asimilación de datos sencillo de implementar y computacionalmente eficiente, que fue incluido en el sistema de pronóstico de mareas que utiliza el Servicio de Hidrografía Naval Argentino para su costa este. Finalmente, en II.3.7, se muestra los primeros pasos hacia el desarrollo de un modelo de pronóstico/hindcast pre-operacional para el Río de la Plata de la reciente publicación del centro CIMA de la Universidad de Buenos Aires.

II.3.1. Sistema Satbałtyk para el Mar Báltico

El sistema para el Mar Báltico (Figura 30), Satbałtyk System, utiliza el modelo PM3D (*Parallel Model 3D*) en modo operacional publicando campos y series temporales de predicciones de la hidrodinámica en su página web: satbałtyk.iopan.gda.pl/

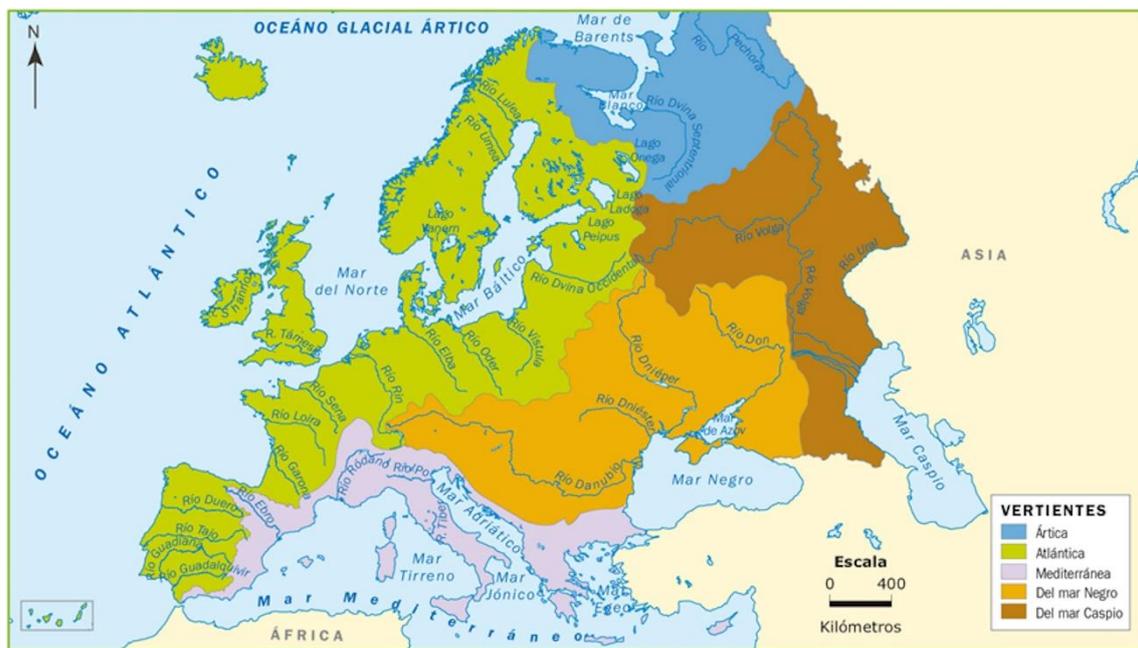


Figura 30 - Hidrografía de Europa. Localización del Mar Báltico.

El modelo (denominado eco-hidrodinámico) fue desarrollado en el Instituto de Oceanografía de la Universidad de Gdańsk en Polonia (IOUG, por sus siglas en inglés). Además de los campos pronosticados a 48 horas de nivel del mar, corrientes en superficie, temperatura y salinidad del agua, el modelo pronostica campos de los siguientes nutrientes: nitratos, amonio, fosfatos, silicatos, nitrógeno total y fósforo, oxígeno disuelto y biomasa de fitoplancton.

El mallado horizontal es cartesiano mientras que el mallado vertical utiliza coordenada sigma con 18 capas. El modelo permite definir sub-áreas con diferente densidad de malla y se define dentro del área modelada del Mar Báltico de resolución 1 NM (Milla Náutica, aproximadamente 1.8 km) de resolución un subdominio con resolución 0.5 NM (Figura 31). Se incluye la entrada de agua dulce de 150 ríos a través de una climatología mensual calculada con los datos desde

1990 hasta 2000. Como forzantes atmosféricos se tienen los campos de pronósticos atmosféricos del modelo UM (*Unified Model*) del centro ICM (Centro interdisciplinario de los modelos matemáticos y computacionales de la Universidad de Varsovia, Polonia) con una resolución de 4 km de la velocidad del viento, presión atmosférica, temperatura del aire y presión de vapor. Se tiene también como entrada la energía solar a partir de los diagnósticos del modelo SolRad. El modelo asimila datos satelitales de temperatura de la superficie del mar, SST, de los radiómetros AVHRR y MODIS que previamente son procesados por el propio sistema (Kowalewski and Kowalewska-Kalkowska, 2017). El proceso de asimilación es similar a una interpolación, dado que la cobertura de los datos incluye todo el dominio, excluyendo los puntos nublados y excesivamente cálidos y corrigiendo no sólo el valor de la capa superior, si no también capas inferiores a través de una función lineal. Más detalles sobre el método de asimilación se pueden encontrar en Kowalewska-Kalkowska and Kowalewski (2019).

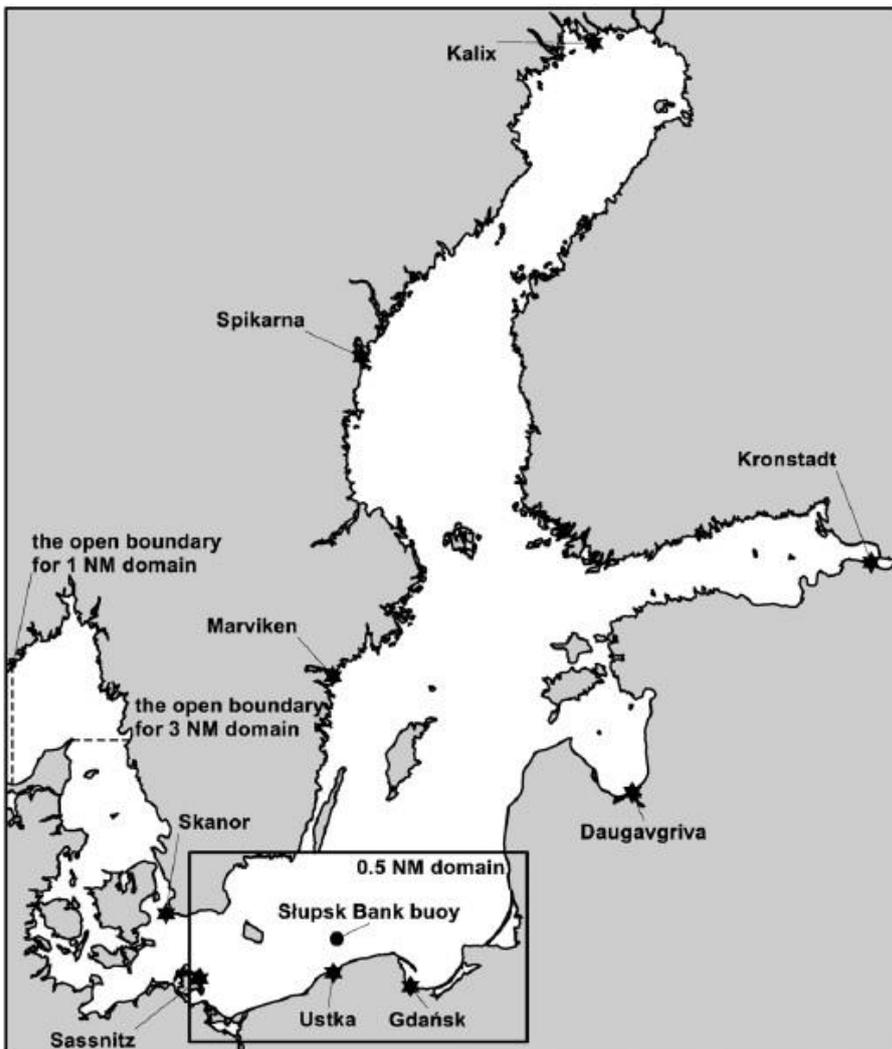


Figura 31 – Dominio del modelo PM3D para el Mar Báltico. Extraído de Kowalewski and Kowalewska-Kalkowska (2017).

El sistema publica los pronósticos generados en su página web; se incluyen mapas con resolución temporal de 6 horas para las próximas 48 horas de nivel, temperatura del agua (a distintas profundidades: 0, 3, 5, 10, 20 y 30 m), corrientes en superficie y salinidad en superficie. En la Figura 32 se muestra un ejemplo del campo de corrientes en superficie pronosticado y en la Figura 33 un ejemplo del campo de salinidad en superficie.

La web publica resultados de otros modelos que incluyen variables meteorológicas y también oceánicas: temperatura de la superficie del mar (observaciones satelitales completadas por los resultados del modelo hidrodinámico), nubosidad, flujo de radiación solar en la superficie del mar, concentración de la clorofila en superficie (observada por el satélite y completada con los resultados del modelo ecohidrodinámico, EcoSat), etc. Además, permite visualizar gráficas con los datos observados disponibles, así, por ejemplo para la fecha de septiembre de 2020 se incluyen tres flotadores Argo que miden temperatura del agua, salinidad, profundidad, presión del agua y oxígeno (para todas las variables se publica 1 medida cada 2 días en hora UTC), cuatro estaciones meteorológicas y una estación UV.

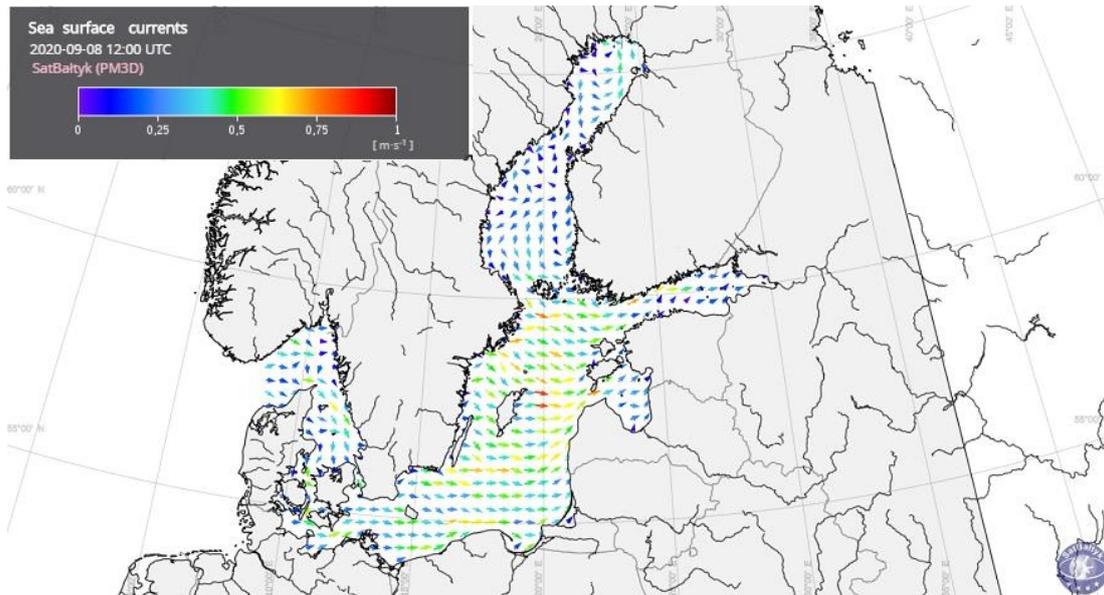


Figura 32 – Pronóstico de corrientes en superficie. Modelo PM3D para el Mar Báltico. Extraído de satbaltyk.iopan.gda.pl/

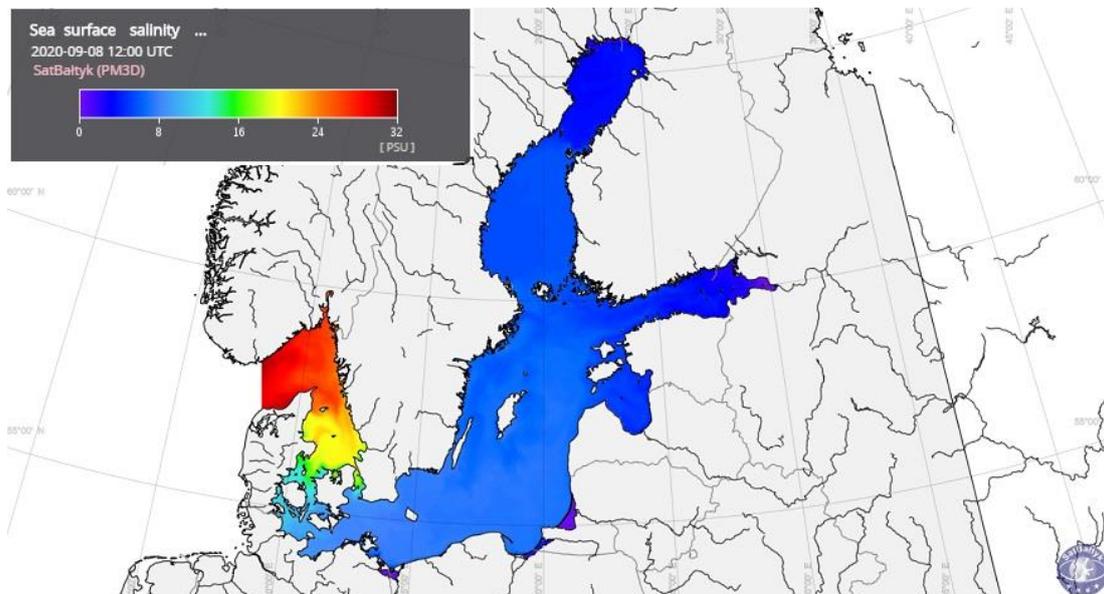


Figura 33 - Pronóstico de salinidad en superficie. Modelo PM3D para el Mar Báltico. Extraído de satbaltyk.iopan.gda.pl/

En Kowalewski and Kowalewska-Kalkowska (2017) se analiza la sensibilidad a la resolución de malla del modelo y también el tiempo de cálculo al introducir la computación en paralelo. Se

comparan los resultados del modelo frente a los registros disponibles en 9 mareógrafos, para ello, se incluyen en la evaluación, el coeficiente de correlación y el error cuadrático medio. También se dan resultados de frecuencia central (CF), frecuencia positiva de outlier (POF) y frecuencia negativa de outlier (NOF). El CF ilustra cómo el nivel de agua simulado difiere de las lecturas en no más de ± 0.15 m; el POF dice con qué frecuencia es más alto el simulado que el nivel observado en más de 0.3 m; y el NOF con qué frecuencia es más bajo el simulado que el nivel medido en más de 0.3 m.

II.3.2. Estuario del Gironda

El Servicio Meteorológico Nacional de Francia, Météo-France, utiliza el modelo TELEMAC-2D para aguas someras con elementos finitos para, en forma operacional, obtener los pronósticos de las condiciones hidrológicas del estuario Gironda en Francia (Figura 34, panel izquierdo). Este sistema envía sus resultados al Servicio de información sobre riesgo de inundaciones de los principales ríos de Francia para que, en caso necesario, puedan emitir alerta de inundaciones en el estuario en su página web (Figura 35): <https://www.vigicrues.gouv.fr/niv2-bassin.php?CdEntVigiCru=23>

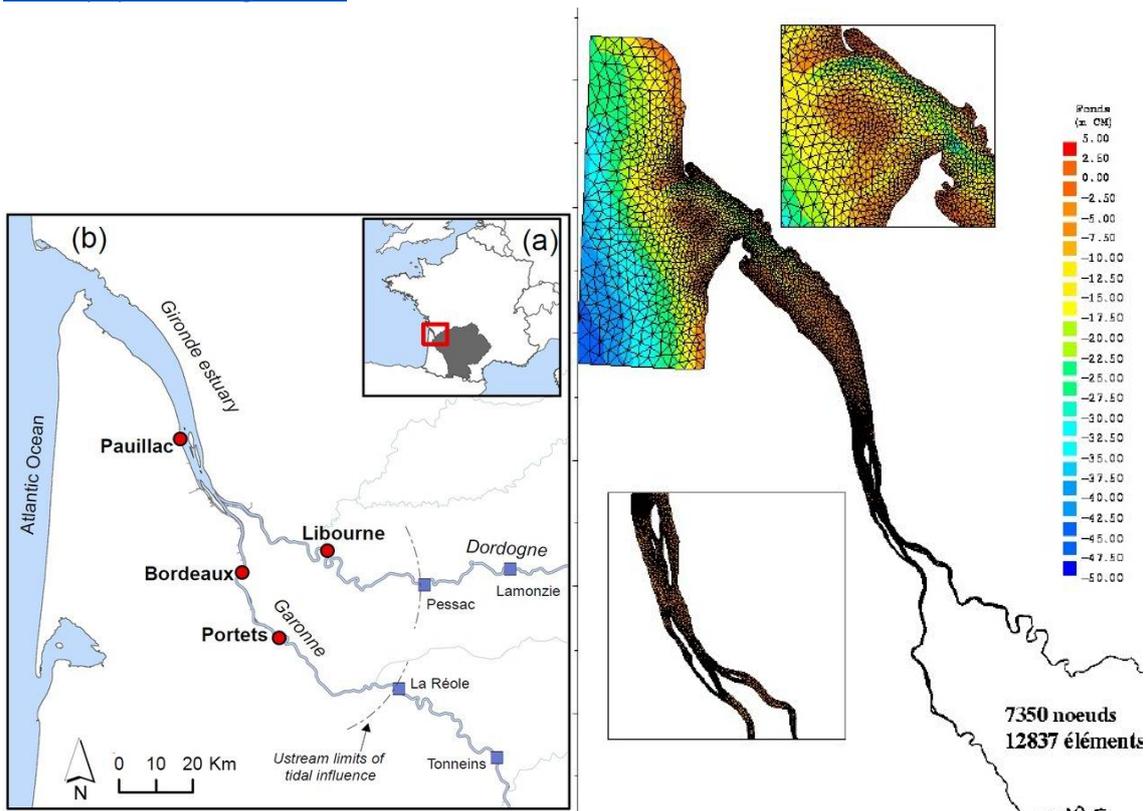


Figura 34 - Panel izquierdo: estuario del Gironda en Francia. En (a) se muestra la cuenca de sus principales tributarios, los ríos Garonne y Dordogne. Extraído de Jalón-Rojas et al. (2015). Panel derecho: malla del sistema de pronóstico del estuario del río Gironda. Extraído de Paradis et al. (2009).

El sistema pronostica a diario las condiciones hidrodinámicas de niveles y componentes de la velocidad promediada en vertical para las próximas 48 horas. Se utiliza un mallado triangular no estructurado que tiene elementos desde 5 km hasta 50 m sobre los ríos (Figura 34, panel derecho). El sistema recibe como condiciones de contorno oceánicas los pronósticos del modelo operacional Météo-France de niveles y corrientes. Como forzantes se tienen pronósticos de

campos de viento y presión atmosférica y para la entrada de agua dulce de los principales tributarios se calculan los aportes para las primeras horas y se consideran constantes para todo el pronóstico.

Los resultados no están abiertos al público. Se comparten de manera interna con los supervisores del pronóstico que son los encargados de emitir las alertas de inundaciones. Como parte del post-proceso se comparan de manera automática los pronósticos del día anterior con las observaciones de nivel disponibles, lo que permite controlar el comportamiento del modelo y poder estudiar los eventos que no están siendo bien representados.

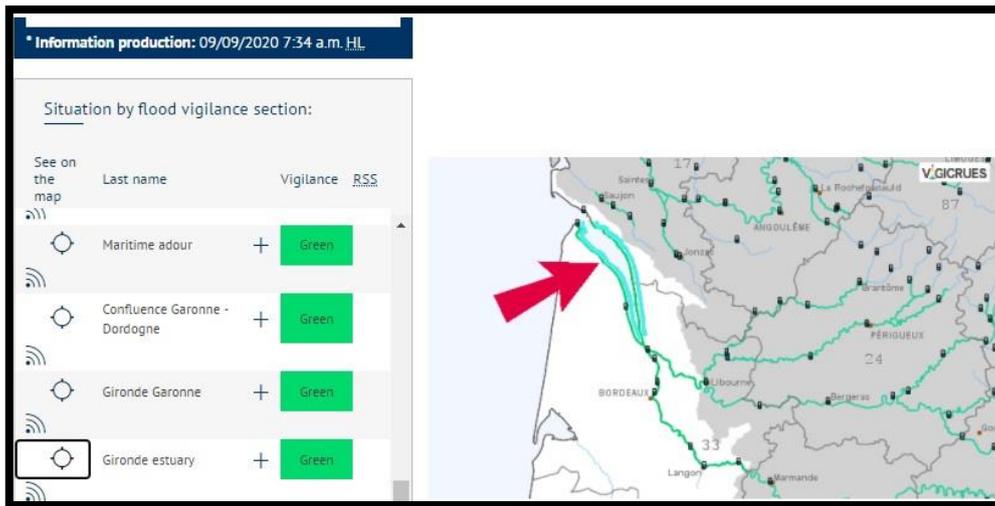


Figura 35 – Pronóstico sin alerta de inundación para el estuario del río Gironda, Francia. Extraído de <https://www.vigicru.es/niv2-bassin.php?CdEntVigiCru=23>

II.3.3. Zona estuarina SSVBES

El sistema SOFS (*Santos Operational Forecasting System*) genera el pronóstico diario de niveles y corrientes en 3D para los próximos 3 días para la zona estuarina SSVBES (*Santos-Sao Vicente-Bertioga Estuarine System*), Figura 36. El Laboratorio de Hidrodinámica Costera del Instituto de Oceanografía, Universidad de Sao Paulo, desarrolló el sistema SOFS a través de un sistema de modelos encajados de forma que la malla de mayor resolución SSVBES recibe las condiciones de borde de la malla de resolución más grosera SBB (*South Brazil Bight*), para ello se utiliza el modelo numérico POM-rain, <http://www.ccpo.odu.edu/POMWEB/userinfo.htm>.

En la Figura 38 se presentan los dominios encajados (Costa et al., 2020) utilizados en el sistema. Las condiciones de contorno del modelo SBB provienen del modelo global COPERNICUS presentado en II.2.3, *Global Sea Physical Analysis and Forecasting*, el modelo global proporciona las condiciones de contorno de nivel y corrientes al modelo SBB, la temperatura y salinidad provienen de una climatología estacional. El modelo SSVBES recibe la descarga de agua dulce como caudales constantes en 30 puntos (entre 0.5 m³/s y 30 m³/s), la temperatura, salinidad y los niveles provienen de SBB más la componente astronómica que se calcula para que su propagación genere la marea estudiada en un mareógrafo del estuario (estación de Praticagem). Como forzantes atmosféricos se introducen a ambos modelos encajados los pronósticos del

viento a 10 m y la presión en superficie del modelo WRF de CPTec (Centro de Previsión del Tiempo y Estudios Climáticos, Instituto Nacional de Investigación Espacial de Brasil).

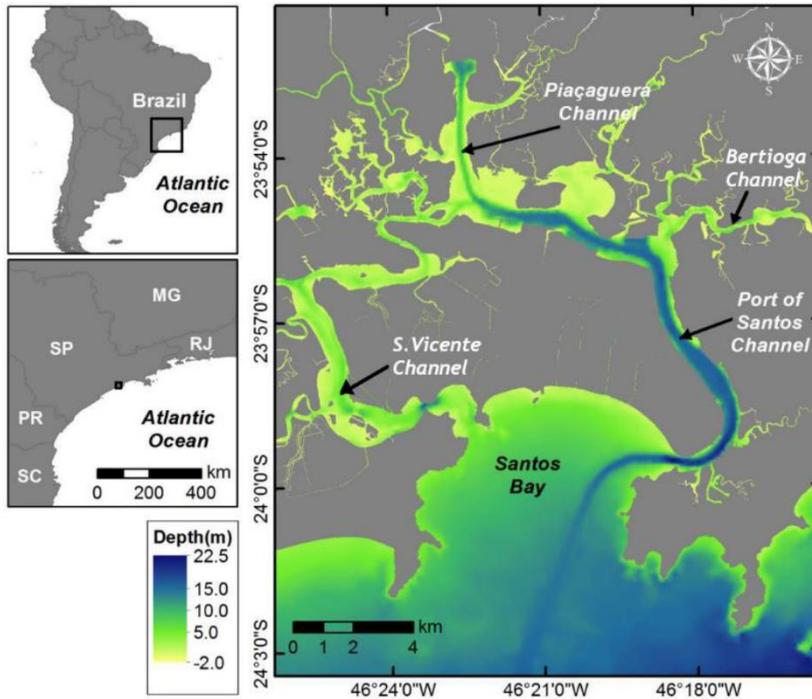


Figura 36 - Zona estuarina SSVBES (Santos-Sao Vicente-Bertioga Estuarine System). Extraído de Mendes et al. (2019).

Los pronósticos tienen una resolución de 1 dato por hora y de forma diaria se publica en la web <http://preamar.io.usp.br/en> una simplificación de corrientes en superficie y una gráfica con la evolución prevista para el nivel (Figura 37).

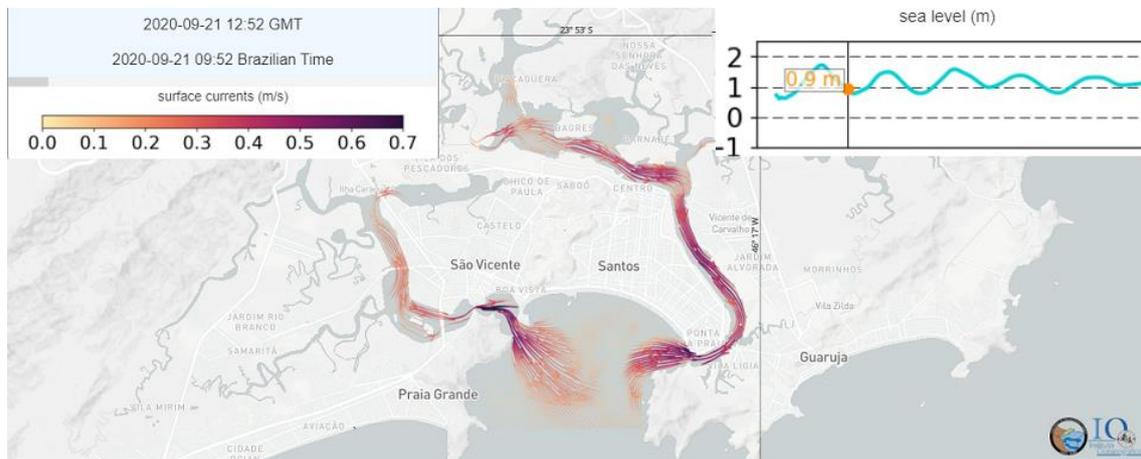


Figura 37 – Ejemplo de pronósticos de corrientes en superficie y evolución del nivel publicados en la web <http://preamar.io.usp.br/en>

Para el sistema de modelos encajados se utiliza un mallado horizontal ortogonal curvilíneo irregular con distancias desde 35 km en la zona oceánica de SBB a 200 m en la zona costera. Para SSVBES se tienen celdas con distancias desde más de 400 m hasta los 26.7 m para las zonas de los canales. La coordenada vertical es del tipo sigma con 21 niveles equiespaciados en SBB y 11 niveles en SSVBES.

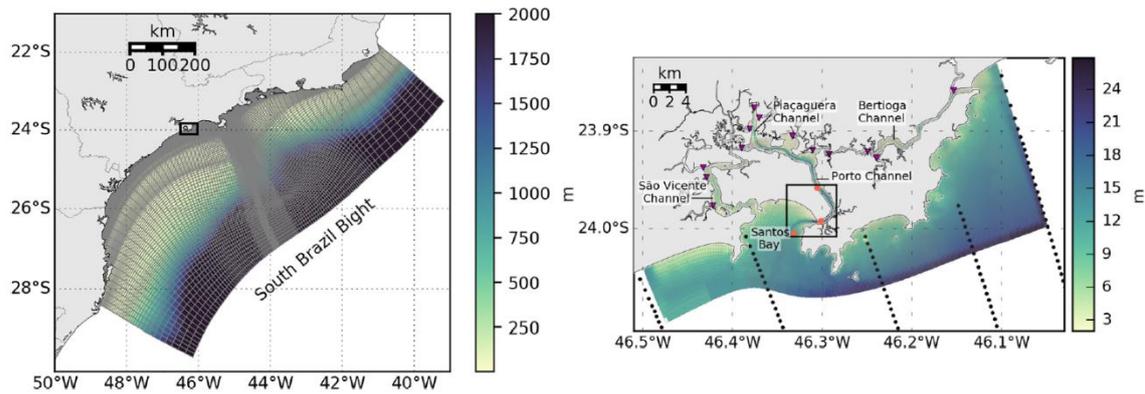


Figura 38 – Sistema SOFS de modelos encajados. Malla gruesa del SBB (izquierda) y malla fina del SSVBES (derecha). Extraído de Costa et al. (2020).

El Sistema de pronóstico fue validado mediante comparación con datos durante 7 eventos de nivel anormalmente alto producto de la combinación de marea astronómica y meteorológica. Se utiliza el estadístico *skill Willmott* definido en II.1.4 y el RMSE comparando niveles y corrientes modelados y medidos para evaluar el sistema. De forma de evaluar la bondad de los pronósticos de los forzantes atmosféricos se hacen simulaciones que comparan los resultados obtenidos con forzantes atmosféricos de reanálisis. El sistema muestra su potencial para pronosticar mareas meteorológicas y servir de apoyo a los servicios de emergencia y ayudar a la navegación en el Puerto de Santos. En un futuro esperan mejorar la capacidad de predecir los campos de temperatura y salinidad.

II.3.4. Río Tajo. Creación de un Módulo de Asimilación Secuencial para el MOHID mediante Filtro de Kalman Extendido

Para su aplicación al modelo de pronóstico del estuario del Río Tajo en Lisboa, Portugal, se desarrolla un módulo de asimilación de datos para el modelo MOHID en 2D como parte de la tesis de doctorado de Ângela Pereira de Matos Canas (Canas, 2009).

El módulo de asimilación secuencial creado en este trabajo es capaz de procesar el filtro SFEK en aplicaciones barotrópicas y baroclínicas de MOHID y el filtro SEEK en aplicaciones barotrópicas de MOHID de acuerdo con el algoritmo de filtro SEEK estándar de Pham et al. (1998) y el algoritmo de filtro SFEK descrito en Brasseur et al. (1999). El método de asimilación de datos elegido es el filtro SFEK (*Singular Fixed Extended Kalman filter*), la versión fija del filtro de Kalman Extendido Evolutivo Singular (*Singular Evolutive Extended Kalman*, SEEK) (Pham et al., 1998; Brasseur et al., 1999). Se trata de un esquema subóptimo del filtro de Kalman que considera una dimensión reducida de la matriz de covarianza de error de la solución modelo, informada por las Funciones Ortogonales Empíricas dominantes (*Empirical Orthogonal Functions*, EOF) de la covarianza de error de estimación del modelo.

El modelo para pronosticar la hidrodinámica en el estuario Río Tajo consiste en dos dominios encajados en 2D barotrópicos (Figura 39): un dominio más grande que contiene la costa portuguesa y gallega que proporciona las condiciones de borde de marea para un dominio más pequeño que contiene el estuario del Tajo y el área costera adyacente. El modelo es forzado con

marea, campos de viento espacialmente variables y flujos promedio para los ríos Tajo, Sorraia y Trancão.

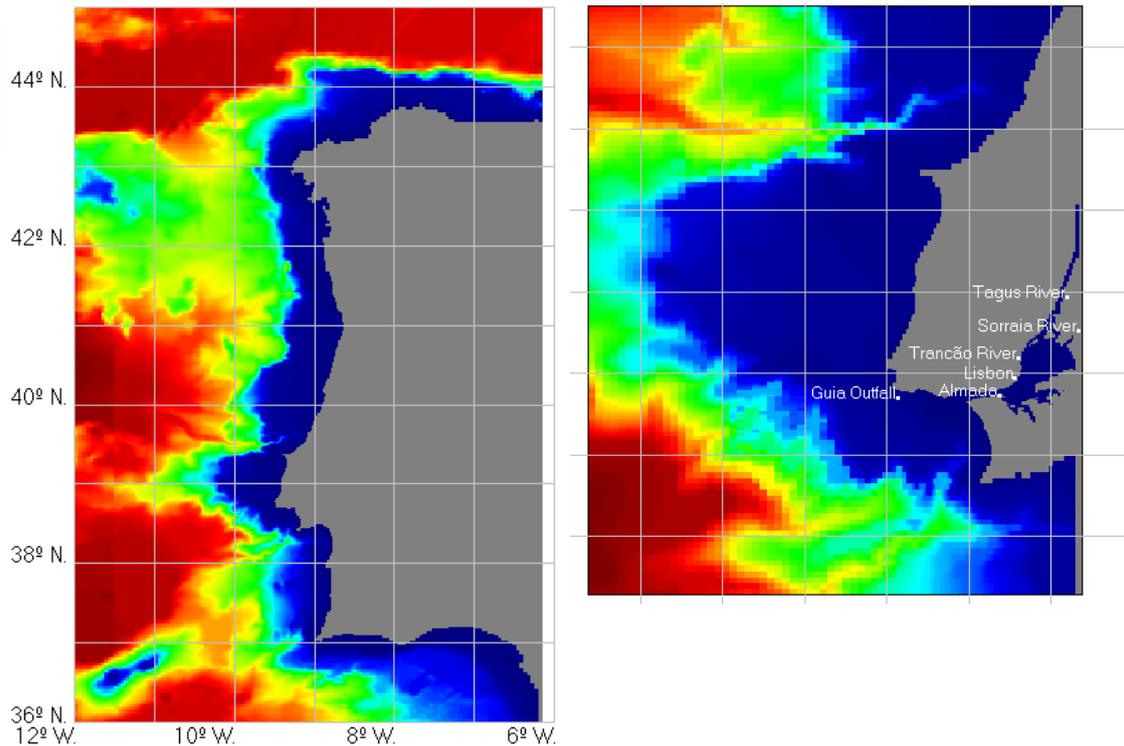


Figura 39 – Sistema de modelos encajados del MOHID 2D del Estuario del Río Tajo. Extraído de Canas (2009).

De forma de validar el módulo implementado se realiza un conjunto de ensayos de asimilación a partir de la definición de experimentos gemelos. Como solución verdadera del estado se tiene la mejor configuración del MOHID 2D a la que se llega en dicha tesis de doctorado, Modelo Verdadero. El Modelo Erróneo, al que se aplicará la asimilación de datos, es el modelo verdadero perturbado y se obtiene por la perturbación del nivel medio del mar impuesto en el límite del dominio más grande con un ruido blanco de desviación estándar de 0.1 m, produciendo un error de modelo no sesgado. Las mediciones se crean a partir de los resultados del nivel de agua de la simulación del Modelo Verdadero para ubicaciones de mareógrafos agregando un error aleatorio normalmente distribuido de media cero y matriz de covarianza de errores diagonal que se considera perfectamente estimada durante la aplicación del filtro.

Los resultados de los ensayos de asimilación realizados son que durante los primeros 15 días de simulación se obtienen mejoras en los valores del nivel del agua tanto en las ubicaciones de medición como en las que no son de medición. Por el contrario, para los campos de velocidad éstos generalmente se degradan por la asimilación en todos los ensayos, posiblemente debido a la débil representatividad de la base EOF y a la acumulación de errores en el dominio del modelo durante el período de simulación.

II.3.5. Mar Tirreno

En Iermano et al. (2016) se estudia la mejora al introducir la asimilación de datos a través de un esquema 4D-Var al modelo en 3D de circulación oceánica ROMS (*Regional Ocean Modeling System*) en la costa del Golfo de Nápoles y la zona adyacente del mar Tirreno (Figura 40). Las

variables que resuelve el modelo incluyen campos tridimensionales de velocidad, temperatura y salinidad. Los datos a asimilar son datos satelitales de temperatura de la superficie del mar (SST) y altura de la superficie del mar (SSH), y datos de campos de velocidad en superficie provenientes de un sistema de radares de alta frecuencia (HF). El esquema de asimilación utilizado permite identificar las variables o subespacios en el modelo que son influenciadas por la asimilación de los datos disponibles. Para ello se utiliza como indicador la influencia de la asimilación de datos en la simulación del transporte litoral a través de una sección.

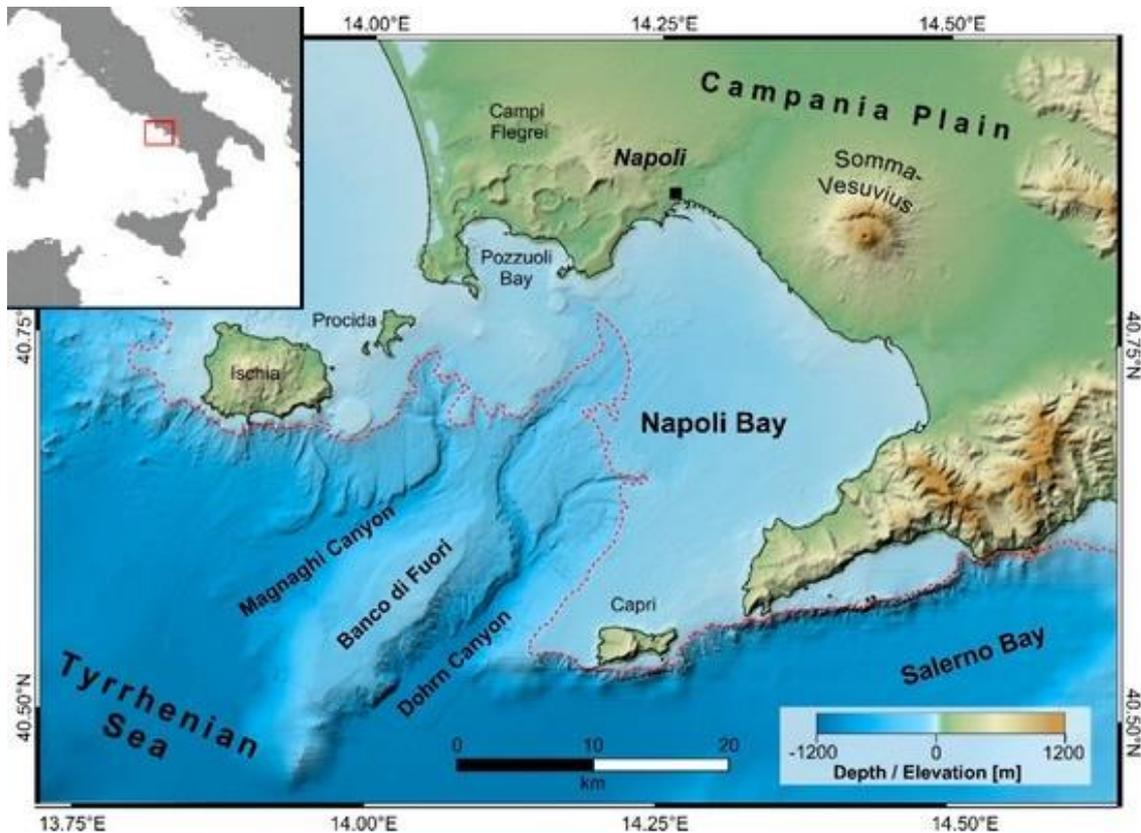


Figura 40 – Ubicación del Golfo de Nápoles en el mar Tirreno. Extraído de Aiello et al. (2020).

El modelo utiliza una malla de 3 km de resolución horizontal y 30 niveles verticales de capas sigma. El esquema de asimilación utilizado incluye en el vector de estado no sólo a las variables resultado del modelo, sino también a las condiciones de contorno y a los forzantes, por lo que éstos también podrán sufrir una modificación en sus valores en el estado analizado que se obtiene tras la asimilación de datos. Los forzantes atmosféricos de entrada al modelo son derivados de los datos de reanálisis Era Interim del centro europeo ECMWF. Como condiciones de contorno de entrada al modelo se tienen los datos de Mercator Océan NEMO (modelo antecesor al COPERNICUS descrito en II.2.3, Lellouche et al., 2018). La evaluación de la incorporación de la asimilación de datos se hace mediante dos experimentos idénticos, con y sin asimilación de datos, para el período octubre-diciembre de 2010.

La asimilación mediante *Strong-Constraint* 4D-Var se hace utilizando valores constantes y no correlacionados para los errores de las observaciones lo que genera una matriz de covarianza de errores, \mathbf{R} , diagonal. Seleccionan valores de desviación estándar de 0.02 m para los niveles, 0.4 °C para la temperatura y 0.07 m/s para las velocidades. La matriz de covarianza de los errores del *background*, \mathbf{B} , se basa en la varianza del modelo durante de las corridas del modelo sin asimilación entre 2007 y 2010. A modo de climatología se utiliza una \mathbf{B} diferente para cada mes

del año; de manera similar, con los valores de 2007 a 2010, se construyeron las matrices de covarianza de errores para los forzantes atmosféricos y las condiciones de borde oceánicas. En la construcción de las matrices se aplica el mecanismo de localización de forma de no tener en cuenta correlaciones entre ubicaciones distantes.

La validación del modelo con asimilación de datos se realiza de manera cualitativa mediante la comparación de los campos de salinidad en superficie (con y sin asimilación de datos) frente a los registros satelitales de clorofila-a (se utilizan estas observaciones por no ser datos asimilados). La comparación del transporte litoral a través de una sección en el modelo con y sin asimilación de datos permite discernir cuándo se están corrigiendo las condiciones iniciales proporcionadas por el *background*, o cuándo la asimilación de datos corrige los forzantes o las condiciones de contorno. En general, son las condiciones iniciales las que mayor diferencia proporcionan. En relación a cuáles son las observaciones asimiladas que generan mayores variaciones se observa que esto es variable en cada ciclo de asimilación pero obtuvieron que en promedio las observaciones del radar y de SSH tienen una relevancia similar para la asimilación.

En el estudio también se incluye la comparación de resultados con datos de salinidad y temperatura a varias profundidades obtenidos por muestreadores CTD. Estas comparaciones permiten conocer cómo la asimilación de datos en superficie se transfiere en profundidad. Para este caso de estudio los resultados no son buenos, y se indica que para que la asimilación de datos sea exitosa se requiere de mayor atención en cuanto a cómo se definieron las matrices de covarianza que afectan a los resultados al descender en la columna de agua.

En el artículo de Ieramo et al. (2016) se pone de manifiesto que la influencia de un modelo de mayor escala que proporciona condiciones de contorno a uno de menor escala se pierde al alejarse de los bordes, y que, por lo tanto, aunque el modelo de mayor escala incluya la asimilación de datos, la introducción de la asimilación de datos en el modelo regional (de menor escala) implica una mejora.

II.3.6. Asimilación mediante 4D-LETKF para el modelo SMARA de la costa argentina

En Etala et al. (2015) se aplica el Filtro Transformado de Kalman por Ensamblados (ETKF por sus siglas en inglés, Ensemble Transform Kalman Filter) presentado en Hunt et al. (2007) en su versión local y 4D (permitiendo asimilar las observaciones asíncronas a los tiempos de asimilación) para mejorar los niveles pronosticados por el modelo SMARA. El modelo SMARA (Servicio de Hidrografía Naval, SNH, de Argentina) es un modelo hidrodinámico regional 2D promediado en vertical de resolución $1/3^\circ$ que representa para la costa argentina y el estuario del Río de la Plata un número limitado de componentes de marea astronómica y la marea meteorológica. Es un modelo que no está funcionando de forma operacional.

El filtro ETKF es un filtro de Kalman por Ensamblados determinístico. Se calcula la media del ensamble analizado y los miembros del ensamble se actualizan por fórmulas separadas a partir de la matriz de covarianza de los errores del análisis que se obtiene previamente. La ventaja de este filtro es la facilidad de aplicación cuando hay pocos datos observados, ya que su formulación permite el cálculo por separado para cada uno de los puntos de la malla en las que se divide el dominio. Esto proporciona una gran facilidad para la aplicación de la localización,

eligiendo qué observaciones son utilizadas en cada celda. La variante 4D permite asimilar las observaciones que se realizan a tiempos diferentes a los tiempos de asimilación.

La formulación del ETKF se basa en que la matriz de perturbaciones del ensamble es capaz de transformar elementos desde el subespacio definido por el ensamble hacia el espacio del estado. Se considera conocida la matriz de covarianza de los errores de las observaciones, \mathbf{R} , y el filtro se centra en la obtención de la matriz de covarianza de los errores del *background*, \mathbf{P}^f , para el caso de modelo perfecto ($\mathbf{Q} = \mathbf{0}$).

La no consideración de errores del modelo lleva a ignorar las observaciones. El filtro ETKF incluye en su implementación la inflación multiplicativa durante la obtención de la matriz de covarianza de los errores del análisis para tener en cuenta que el modelo no es perfecto y también se nombra la posibilidad de inflación aditiva que es lo que se implementó en Etala et al. (2015).

La necesidad de tener en cuenta el efecto localizado de las observaciones se justifica diciendo que en una región local suficientemente pequeña, el sistema se comporta como un sistema inestable de reducida dimensión gobernado por la dinámica de las regiones cercanas. Además, si no se implementara la localización, el limitado tamaño de la muestra proporcionada por el ensamble produciría correlaciones espurias entre localizaciones lejanas en la matriz de covarianza del *background*. La aplicación de la localización se puede hacer de manera sencilla eligiendo qué observaciones son tenidas en cuenta para cada punto de la malla o, como se hace en Etala et al. (2015), a través de un decaimiento con la distancia del punto de la malla a las observaciones.

La formulación obtenida en ETKF reduce a $\mathbf{P}\mathbf{x}\mathbf{P}$ el rango de las matrices $\mathbf{N}\mathbf{x}\mathbf{N}$ a invertir durante el filtro de Kalman. A continuación, se presenta la forma de obtener la matriz de covarianza del *background* que realiza este método.

La matriz de perturbaciones $\mathbf{X}_k^f \in \mathbb{R}^{N \times P}$ contiene, para el tiempo k , las diferencias de cada miembro con la media del ensamble. Sus columnas, desde $p = 1$ hasta P , son la diferencia del miembro p con la media del ensamble, $(x_{k,p}^f - \overline{x_k^f})$. Con $P < N$, los miembros del ensamble definen el subespacio, \mathcal{S} , de las perturbaciones.

La matriz de covarianza de los errores del *background* en el subespacio \mathcal{S} es proporcional a la identidad $\mathbf{P}\mathbf{x}\mathbf{P}$, $\tilde{\mathbf{P}}_k^f \approx \frac{1}{P-1}\mathbf{I}$, obteniéndose la matriz de covarianza de los errores del *background* en el espacio del estado, \mathbf{E} , a través de la transformación por \mathbf{X}_k^f . Es decir, $\mathbf{P}_k^f \approx \mathbf{X}_k^f \tilde{\mathbf{P}}_k^f (\mathbf{X}_k^f)^T$.

De este modo, en la formulación de ETKF, se utiliza \mathbf{X}_k^f para transformar elementos del subespacio \mathcal{S} al espacio \mathbf{E} . De manera análoga a lo presentado para \mathbf{P}_k^f , se obtienen fórmulas para construir en el subespacio \mathcal{S} la matriz de ganancia de Kalman, \mathbf{K}_k , y la matriz de covarianza de los errores del análisis, \mathbf{P}_k^a .

Este método es capaz de tener en cuenta observaciones asíncronas a los tiempos de análisis. Aplicando el operador de las observaciones a cada miembro del ensamble se obtienen los vectores $\mathbf{y}_{k,p}^f$. La matriz de $\mathbf{Y}_k^f \in \mathbb{R}^{M \times P}$, contiene en sus columnas las diferencias $(y_{k,p}^f - \overline{y_k^f})$. La forma de tener en cuenta observaciones asíncronas en la versión 4D-LETKF es a través de un operador de las observaciones diferente para cada instante de observación. Así, para cada instante de observación, se generan los vectores $\mathbf{y}_{k,p}^f$ de ese instante de observación.

Concatenando los vectores obtenidos para cada instante se obtienen los vectores $y_{k,p}^f$ globales y la matriz de Y_k^f . Del mismo modo se forma la matriz de covarianza de los errores de las observaciones, R_k , que será diagonal por bloques al considerar las observaciones en distintos tiempos no correlacionadas. Se apunta la importancia de tener en cuenta las correlaciones temporales del modelo para no limitar la influencia de observaciones asíncronas con la localización. En Etala et al. (2015), para las observaciones asíncronas se aplica también un factor de decaimiento exponencial con el tiempo.

El prototipo presentado en Etala et al. (2015) muestra el potencial de la asimilación de los datos disponibles de mareógrafos y del altímetro satelital (en la Figura 41 se muestra el dominio de cálculo con las observaciones disponibles). Mediante un ensamble de 20 miembros ($P = 20$) se realiza la asimilación secuencial en ciclos de 6 horas según el filtro 4D-LETKF.

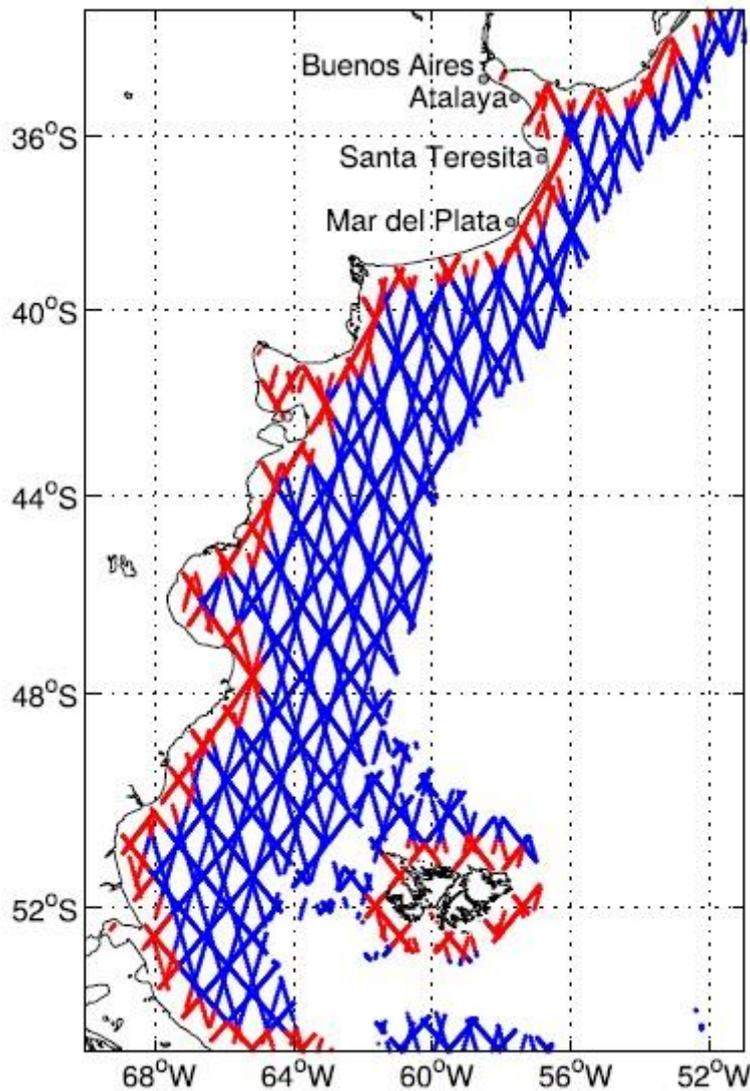


Figura 41 – Ubicación del punto de control (Buenos Aires) y de los mareógrafos y trazas de los altímetros satelitales a asimilar. Extraído de Etala et al. (2015).

La asimilación se realiza sin tener en cuenta la marea astronómica. Las observaciones de mareógrafos asimiladas son residuos horarios en Mar del Plata, Santa Teresita y Atalaya. Las últimas dos también se utilizan en validación, así como Buenos Aires. Los datos de altímetro satelital son SSH disponibles para la zona en septiembre de 2011.

Los experimentos realizados pueden incluir los datos de mareógrafos únicamente, o los datos de mareógrafos y altímetros satelitales y pueden ser con participación de inflación aditiva o no. También se hace variar la escala de localización y el error observacional (R). Las comparaciones se hacen durante eventos de intensidad moderada a fuerte que tuvieron lugar en la plataforma continental y en el estuario en septiembre de 2011.

El artículo de Hunt et al. (2007) tiene como objetivo principal presentar un método de asimilación que sea fácil de implementar y computacionalmente eficiente para grandes sistemas espacio-temporales caóticos, es decir, la atmósfera y el océano. Este algoritmo permite la utilización de modelo no lineal y de operador de las observaciones no lineal sin requerir del cálculo de sus tangentes lineales ni de sus adjuntos. Es particularmente eficiente cuando el número de miembros del ensamble es pequeño comparado con la dimensión del estado y la cantidad de observaciones.

En la aplicación para el modelo SMARA (Etala et al., 2015) la asimilación de datos se observa como más adecuada para escenarios de vientos del sur de gran *fetch*, siendo necesarios datos locales para los escenarios de vientos locales del sudoeste. Al tener en cuenta los datos de los altímetros satelitales se corrige la influencia exagerada de los mareógrafos aislados y se proporciona información en áreas que carecen de observaciones costeras.

Como conclusión se sugiere un impacto positivo a resaltar en la utilización de la asimilación de los datos disponibles para los pronósticos a corto plazo en la zona media y externa del Río de la Plata.

II.3.7. Modelo CROCO adaptado a simulaciones 2D en hindcast y pronóstico para el Río de la Plata

En Dinápoli et al. (2020) se modifica el código del modelo CROCO 3D (Coastal and Regional Ocean COmmunity model <http://www.croco-ocean.org>) hacia una versión 2D barotrópica. Mediante un sistema de modelos encajados, Figura 42, se realizan simulaciones retrospectivas para los niveles y las corrientes promediadas en vertical para el Río de la Plata. Este modelo no está actualmente operativo.

El modelo A, Figura 42, tiene una malla regular de unos 12 km de distancia entre puntos y es el que proporciona las condiciones de contorno al modelo B, de mayor resolución (aproximadamente 4 km). Las condiciones de contorno oceánicas del modelo A son la marea astronómica de 8 componentes principales diurnas y semidiurnas (armónicos) obtenidas del atlas de marea astronómica TPXO9. Como forzantes se tienen la presión en superficie, la velocidad del viento a 10 m de altura y la densidad del aire en superficie obtenidas de los reanálisis de ERA5 (*1-hourly COPERNICUS Climate Change Service; C3S*) del centro ECMWF (*European Centre for Medium-Range Weather Forecast*); en el modelo B la velocidad del viento es corregida a través de una función empírica calibrada con las observaciones de viento para permitir una mejor representación de los vientos regionales. El aporte de agua dulce son las observaciones diarias de los caudales de los ríos Paraná y Uruguay.

El modelo se evalúa mediante la comparación estadística de los niveles simulados frente a las medidas disponibles en 8 mareógrafos en situaciones normales principalmente en el período desde diciembre de 2002 hasta marzo de 2003. La evaluación de las velocidades de corriente

simuladas se hace frente a los datos de 11 ubicaciones de medidores de corriente provenientes de campañas de medida de varios meses de duración en los años 1996, 1998-1999, 2002-2003 y 2003-2004.

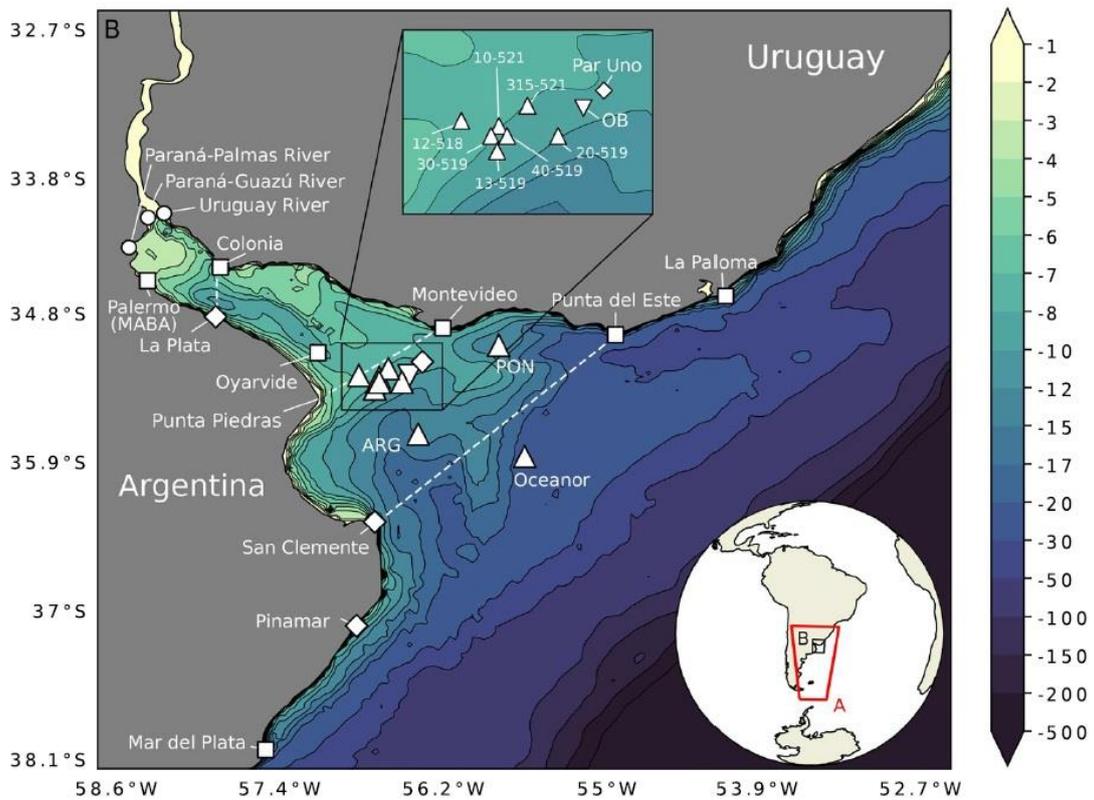


Figura 42 – Modelo A y modelo B para el sistema CRACO aplicado al Río de la Plata. Se representa la batimetría y se muestran las ubicaciones de medidas. Extraído de Dinápoli et al. (2020).

Se realiza la evaluación del modelo para situaciones de marea meteorológica extrema (positiva y negativa) mediante la comparación de la marea residual (diferencia entre el nivel del mar y la marea astronómica pronosticada por los armónicos) simulada frente a la marea residual medida en la estación de Palermo (Figura 42). El artículo muestra la evaluación en 10 situaciones más extremas de marea meteorológica, tanto positiva como negativa, de los últimos 20 años.

La aplicación del modelo en modo pronóstico se genera para estas situaciones de marea extrema y para ello se modifica la fuente de la que se extraen los forzantes meteorológicos, pasando a ser los pronósticos del HRES (6-hourly Atmospheric Model High Resolution 10 day) del centro ECMWF. Para estas aplicaciones se realiza un calentamiento previo del modelo utilizando los forzantes de reanálisis.

Como conclusiones se tiene que los primeros pasos presentados hacia el desarrollo de un modelo de pronóstico/hindcast pre-operacional son respaldados por la capacidad de los resultados obtenidos para simular incluso eventos extremos. Para las simulaciones en modo pronóstico se destaca la relevancia que toman las incertezas de los pronósticos de forzantes meteorológicos hacia los resultados obtenidos.

III. Generación del modelo de pronóstico

III.1. Características del área de estudio

El Río de la Plata recibe el aporte de la segunda cuenca más grande de Sudamérica (y la quinta en el mundo), con un área de 3,170,000 km². El aporte de agua dulce se realiza en más del 97% a través de los Ríos Paraná y Uruguay, siendo la descarga media de aproximadamente 22,000 m³/s. El conocimiento de la zona de estudio a través de proyectos anteriores permite describir las principales características de este cuerpo de agua y delinear los requerimientos que tendrá un modelo de pronóstico.

Los principales forzantes que influyen sobre la circulación del agua en el Río de la Plata son la descarga fluvial de los tributarios, la onda de marea oceánica (astronómica y meteorológica) y los vientos que soplan sobre la superficie del agua, pero las variaciones en los parámetros físico-químicos (particularmente la salinidad) afectan también la circulación al modificar la densidad del agua (Fossati et al., 2014). El efecto de los forzantes sobre el agua está condicionado, a su vez, por la configuración de la línea de costa y la batimetría de fondo. El aumento, primero paulatino y luego abrupto, en la sección transversal del río tiene gran influencia en sus características hidrodinámicas (Fossati et al., 2014).

Un sistema de pronóstico necesitará, por tanto, que el modelo numérico incluya secciones correspondientes a la entrada de agua dulce de los principales tributarios; pronósticos para las condiciones de borde no sólo de la componente de la marea meteorológica sino también para la marea astronómica; pronósticos para los forzantes meteorológicos que incluyan viento y es recomendable que, además de resolver las variables físicas de la hidrodinámica (nivel y velocidades), se incluya como trazador la salinidad. Asimismo, tanto para la comparación en tiempo cuasi-real como para una posible inclusión de la asimilación de datos en el sistema se debe disponer de datos medidos en tiempo real en el dominio.

La estrategia seguida en trabajos anteriores para tener en cuenta la componente de marea meteorológica en las condiciones de borde de un modelo regional ha consistido en un sistema de modelos encajados (Figura 43). De este modo, se parte de un primer nivel a escala regional, en dos dimensiones, que abarca gran parte del Océano Atlántico Sur y un segundo nivel a escala local del Río de la Plata, de mayor resolución y que puede ser en tres dimensiones para algunas aplicaciones. Al primer dominio se lo ha denominado AStide y al segundo dominio se le llama RPtide. El modelo AStide tiene como condiciones de borde el atlas de marea astronómica y vientos y presiones de reanálisis en la superficie libre por lo que es capaz de proporcionar las condiciones de contorno al modelo RPtide correspondientes a la marea total (astronómica más meteorológica). Para el dominio AStide se utiliza el modelo numérico MOHID mientras que para el dominio RPtide se han utilizado, según la aplicación, el modelo numérico MOHID (Alonso et al., 2017) o el sistema de modelos corrientes-oleaje-sedimentos TELEMAC (Santoro et al., 2017).

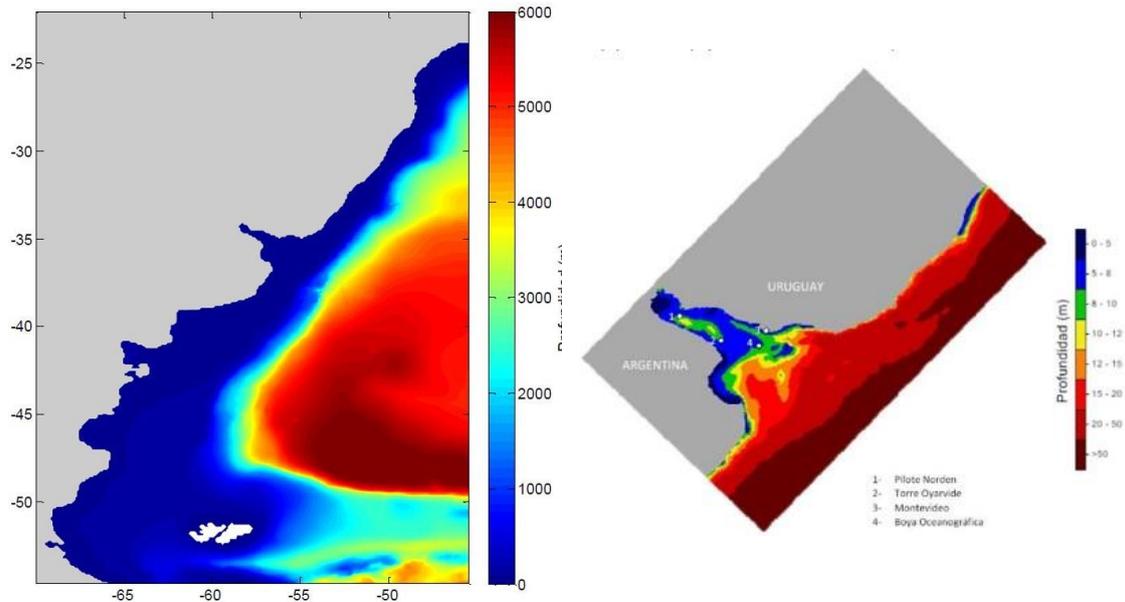


Figura 43 – Sistema de modelos encajados definidos para los estudios en el Río de la Plata; dominio del ASTide (panel izquierdo) y dominio del RPtide (panel derecho).

III.2. Diseño del sistema de pronóstico

III.2.1. Configuración del modelo numérico

Para el prototipo de pronóstico de niveles en el área de estudio desarrollado en este trabajo se manejaron como alternativas las implementaciones ya calibradas y validadas de estudios anteriores. Se selecciona como herramienta numérica el modelo TELEMAC-MASCARET (Hervouet, 2007) con la configuración definida para el *hindcast* de niveles desarrollado recientemente (Jackson et al., 2018). Se utiliza una malla bidimensional en elementos finitos (Figura 44) no-estructurada de resolución variable desde 3 km hasta 100 m en la zona interior. A través del sistema de ecuaciones de aguas someras presentado en II.1.2.3 y utilizando el modelo de turbulencia de Smagorinsky para las escalas no resueltas se obtienen la variación de la superficie libre, la velocidad promedio en vertical y se incluye como trazador activo la salinidad (según la ecuación de estado (40)). Si bien nuestro grupo de trabajo utiliza diversos modelos numéricos, se decidió emplear el modelo TELEMAC por tratarse de herramientas de código abierto, eficiente computacionalmente, en continuo desarrollo por grupos de reconocido prestigio internacional, y con una comunidad de usuarios extendida y con fluida comunicación (www.opentelemac.org).

$$\rho = \rho_0 + 0.749979S \quad (40)$$

La malla es identificada como mesh_21. La batimetría del área de interés fue generada a partir de datos batimétricos digitalizados extraídos del GEBCO (*General Bathymetric Chart of the Ocean*) para la zona del dominio del frente marítimo, y para la zona del Río de la Plata se utilizaron una carta batimétrica del proyecto FREPLATA (FREPLATA, 2004) y una serie de datos de batimetría provistos por el Servicio de Hidrografía Naval de la Argentina (SHN). Como plano

de referencia se toma el “Cero Wharton” situado a 0.91 m por debajo del nivel medio de las aguas del Puerto de Montevideo. Se definen en la malla tres secciones que representan los aportes fluviales más relevantes: la primera corresponde al río Santa Lucía; la segunda al ingreso de los ríos Uruguay y Paraná-Guazú, y la tercera al ingreso del caudal del río Paraná Las Palmas. Se realiza una división del caudal del río Paraná de forma que el 73 % del caudal del río Paraná corresponde al río Paraná Guazú y el 27 % al río Paraná Las Palmas.

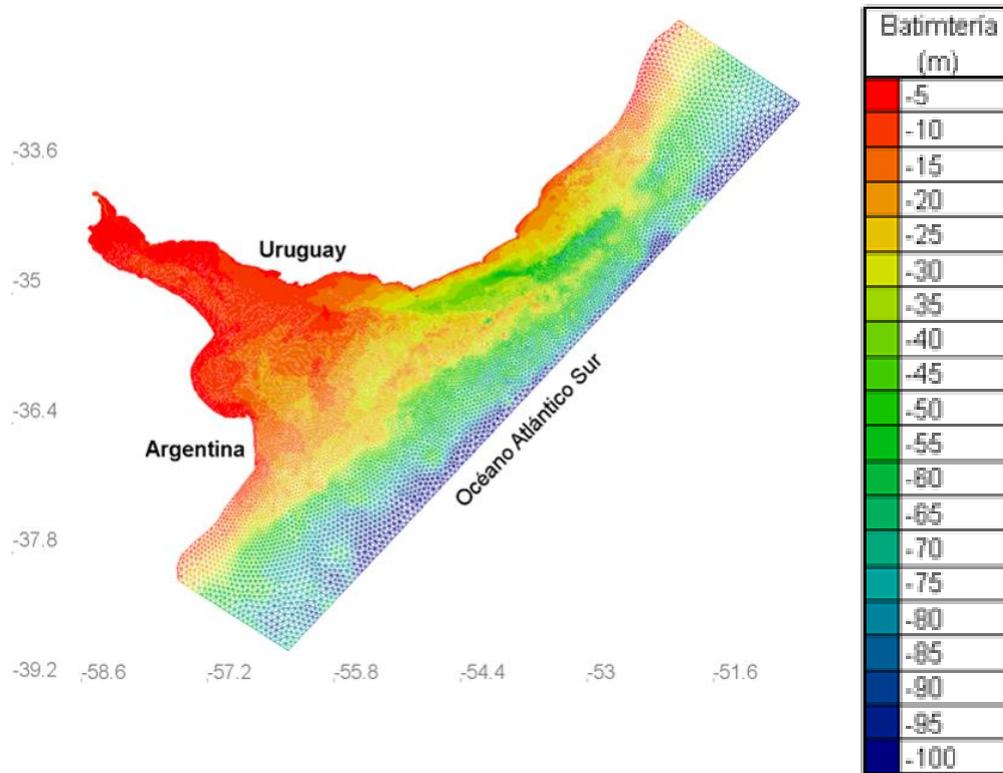


Figura 44 – Malla de cálculo del modelo TELEMAC utilizada en el sistema de pronóstico.

Para la calibración del modelo se utilizaron datos de niveles medidos en tres estaciones mareográficas, La Paloma, Montevideo y Colonia (Jakcson et al., 2018). Para la calibración se realizó la zonificación del Río de la Plata en tres zonas (Interior, Intermedia y Exterior) y se seleccionaron como parámetros de calibración el coeficiente de rugosidad de Manning variable según cada zona y un coeficiente asociado al arrastre del viento. Como resultado de la calibración se obtuvo el valor de Manning igual a 0.015 para las zonas Intermedia y Exterior, un número de Manning de 0.007 para la zona Interior y el coeficiente de arrastre C_D variable con la intensidad del viento según la formulación de Flather (Flather, 1976).

III.2.2. Condiciones de borde

Para implementar la configuración descrita anteriormente del modelo TELEMAC-2D en modalidad pronóstico se requiere contar con pronósticos de las condiciones de contorno en la frontera oceánica, de los forzantes atmosféricos y de los caudales fluviales.

Para la frontera oceánica es posible utilizar resultados del modelo global HYCOM como el modelo COPERNICUS. Ambos modelos globales fueron descritos en la sección II.1.2. Estos modelos globales no incorporan la marea astronómica. Por tanto, es necesario añadir la marea astronómica proveniente de un atlas de marea astronómica. Como alternativas se plantean el

atlas TPXO (http://volkov.oce.orst.edu/tides/tpxo8_atlas.html) y el atlas FES2014 (<https://www.aviso.altimetry.fr/en/data/products/auxiliary-products/global-tide-fes/description-fes2014.html>). Estos atlas se basan en la asimilación de grandes series de registros de altímetros satelitales (y en algunas ubicaciones también datos de mareógrafos) para la generación de las componentes de la marea astronómica.

El atlas TPXO es utilizado por la marina de los Estados Unidos de manera superpuesta a la SSH (altura de la superficie del mar) pronosticada por el RTOFS (HYCOM) para obtener la elevación total del nivel del mar. Los modelos TPXO son proporcionados en formato binario y en su web se ofrecen paquetes para su lectura en Fortran o Matlab. Se incluyen amplitudes de elevaciones de la superficie del mar y corrientes para ocho constituyentes armónicos primarios (M2, S2, N2, K2, K1, O1, P1, Q1), dos de períodos largos (Mf, Mm) y tres no lineales (M4, MS4, MN4) (además de 2N2 y S1 sólo para la reciente actualización TPXO9).

El atlas FES2014 (Carrere et al., 2015) es la última actualización del modelo oceánico global de cálculo de marea astronómica elaborado por LEGOS, Noveltis, CNES, CLS, Francia. A través de su web se puede solicitar acceso a los archivos para usuarios registrados y se obtiene además un programa en lenguaje C para obtener las elevaciones y corrientes para 34 constituyentes de marea (2N2, EPS2, J1, K1, K2, L2, La2, M2, M3, M4, M6, M8, Mf, MKS2, Mm, MN4, MS4, MSf, MSqm, Mtm, Mu2, N2, N4, Nu2, O1, P1, Q1, R2, S1, S2, S4, Sa, Ssa, T2).

Se exploró la disponibilidad en tiempo real de las fuentes de pronóstico de viento y presión atmosférica. El modelo global COPERNICUS utiliza como forzantes los campos atmosféricos del Sistema de Pronóstico Integrado ECMWF que no tienen libre disponibilidad (el centro ECMWF ofrece su venta). Los campos de forzantes meteorológicos globales del NCEP, NOAA, que utiliza el HYCOM, sí se encuentran con libre disponibilidad, siendo su producto de pronóstico el GFS (*Global Forecast System*) (<http://nomads.ncep.noaa.gov/>).

En el diseño del sistema de pronóstico del Río de la Plata se establecieron dos alternativas de modelos globales (HYCOM y COPERNICUS) para el pronóstico de las condiciones de borde a las que se debe añadir la marea astronómica de un atlas (TPXO o FES2014) y como forzante se utilizan los campos GFS del NOAA.

Para determinar qué tan bien representan los modelos globales la marea en la zona exterior del Río de la Plata, se realizaron diversas comparaciones de los niveles de dichos modelos con datos disponibles. Por un lado se realizó la comparación de los niveles obtenidos con el modelo regional disponible en el IMFIA (ASTide) con el modelo global COPERNICUS con la componente astronómica del modelo FES (Ballesteros et al., 2018). Esta comparación se realizó para los datos del año 2015 en los 353 puntos del borde oceánico del dominio RPtide calculando para las variables de nivel y componentes de la velocidad la desviación estándar de la serie en cada punto, la correlación Pearson y realizando un chequeo mediante visualización. Por otra parte los resultados de ambos modelos también se compararon con los registros de los mareógrafos ubicados en Mar del Plata y La Paloma, realizando nuevamente un chequeo visual mediante *scatterplots*, comparación de la desviación estándar, la correlación Pearson y el CRMSE. Los resultados obtenidos se presentan en el Anexo 1. Principalmente se observa que el modelo global tiene un comportamiento similar al modelo ASTide y por lo tanto es posible la incorporación de sus pronósticos como condiciones de borde en el dominio RPtide. En relación a las velocidades, el comportamiento del modelo global no se acerca del mismo modo que con los niveles al comportamiento del modelo ASTide y en este caso no se cuenta con datos de observaciones que permitan definir qué modelo destaca. Para el prototipo de pronóstico

desarrollado en esta tesis se incorporan al borde del dominio únicamente pronósticos de niveles del modelo global a los que se determinó agregar 0.91 m, además de la marea astronómica proveniente del atlas FES 2014, para llevar los niveles al cero Wharton, cero de referencia del prototipo al cual está referida la batimetría del modelo.

Por último, se exploró la posibilidad de utilizar registros de altímetros satelitales para evaluar los modelos globales. Se generaron comparaciones visuales y *scatter* para los datos del año 2015 en el dominio del Río de la Plata frente a los resultados del MOHID_RPtide y el COPERNICUS. Para ello a los resultados del modelo MOHID_RPtide se les restó la componente astronómica (Pawlowicz et al., 2002). Estas primeras exploraciones se ofrecen en el Anexo 2. El grado de avance en el estudio de estos datos sirve para conocer que cubren toda el área del Río de la Plata pero con una frecuencia temporal muy baja con un promedio de menos de un pase diario. La naturaleza de los datos descargados para esta primera exploración se asemeja a los datos de los modelos globales en el aspecto de que no incorporan marea astronómica, por lo que en caso de querer ser utilizados para la evaluación del modelo se deberá aportar esta componente a partir de un atlas. Para poder llevar a cabo la comparación de los resultados del prototipo de pronóstico frente a los datos de altímetro satelital se requiere de un mayor estudio de los datos que incluyan un nivel de referencia para cada punto del dominio frente al cero Wharton que en la exploración llevada a cabo desaparece al descartar la marea astronómica. Los datos del altímetro suponen medidas en ubicaciones en toda el área donde no se dispone de mareógrafos pero por su dispersión temporal por el momento no se incluirán en la evaluación del prototipo.

La exploración de la disponibilidad de campos de pronósticos (condiciones de borde oceánicas y forzantes atmosféricos) permite conocer que estos datos no se mantienen accesibles y que, en general, los pronósticos son sustituidos de forma diaria con nuevos datos. Es aquí cuando surge la necesidad de generar una base de datos de pronósticos. Desde finales de 2017 se están recopilando diariamente datos de pronóstico de los niveles y corrientes de los modelos globales, así como pronósticos de los forzantes atmosféricos (vientos y presiones en superficie) y datos medidos por los mareógrafos en el dominio. Esta base de datos permite disponer de los datos a los que se hubiese tenido acceso en tiempo real y de este modo poder evaluar un tiempo largo de simulaciones de pronóstico para cualquier modificación del modelo en las mismas condiciones en las que se hubiese hecho en tiempo real.

III.3. Generación de la base de datos

En este apartado se describen los detalles de los datos de pronósticos disponibles en tiempo real, tanto de modelos globales como de forzantes atmosféricos, y también de los mareógrafos que publican las observaciones en tiempo casi real. Se atiende a características técnicas como son su resolución horizontal pero también a características operacionales como son la hora a la que cada día se publican los datos. Se detalla el procedimiento para su incorporación a la base de datos.

En general, un requerimiento importante en la oceanografía operacional es la capacidad de cómputo. El objetivo de este proyecto es generar un sistema de pronóstico para una región de considerables dimensiones si se compara con los emprendimientos costeros existentes, siendo la cantidad de nodos para este primer prototipo en versión 2D para la mesh_21 de 37,534 puntos. Las primeras pruebas de este proyecto se desarrollaron en una computadora de

escritorio pero los requerimientos hicieron necesaria una migración para la que se tuvo en cuenta, además de la velocidad de los cálculos pensando en el desarrollo futuro, una gran disponibilidad de almacenamiento, la continuidad en la conexión a la red eléctrica y a Internet y la no interferencia con otros usuarios que suele generarse en las plataformas de uso compartido y que impedirían que el pronóstico se generase a la misma hora todos los días. Estas consideraciones llevaron a no utilizar el Centro Nacional de Computación (ClusterUY, www.cluster.uy) y, desde mediados de 2019 el desarrollo del sistema se ejecuta en un servidor que ha sido denominado OBRAS09 y tiene como características: Servidor Hp proliant DL 380 gen9. Procesador: dos Xeon E5-2680 v3 de 12 cores cada uno. Velocidad: 2.5Ghz. Memoria RAM: 128GB. Sistema operativo Ubuntu 18.04.4. Discos: 7 discos de 5TB en raid 5, total 28TB (aquí se almacena el pronóstico entre otras cosas) y 24 discos de 4TB en raid 10, total 81TB.

III.3.1. Condiciones de borde oceánicas

Para los modelos globales que están siendo considerados en este trabajo, HYCOM y COPERNICUS, los pronósticos con resolución temporal horaria disponibles corresponden a datos en dos dimensiones. Se estableció incluir en la base de datos únicamente los datos que tienen esta resolución temporal.

El modelo HYCOM publica sus pronósticos globales en formato NetCDF o también se puede utilizar una interfaz OpenDAP para el acceso a datos de una región (esta opción no permite descargar los datos con frecuencia temporal horaria, siendo la mejor resolución temporal disponible con la interfaz 1 dato cada 3 horas). Para este trabajo se descargan por tanto los archivos NetCDF. Independientemente del modo en que se descarguen los datos es muy recomendable suscribirse a la lista de notificaciones y avisos de problemas de los servicios del centro NCEP (<https://www.lstsrvc.ncep.noaa.gov/mailman/listinfo/ncep.list.nomads-ftpprd>) donde, entre otros, se dan avisos de los datos publicados por la agencia NOAA, en la que se encuentra el modelo del pronóstico RTOFS (HYCOM). Debido a que los datos no se distribuyen bajo suscripción, el estar incluido en esta lista permite recibir notificaciones de retraso en la diseminación de productos, mantenimiento, actualizaciones y lanzamiento de nuevos sistemas, seguimiento de la resolución de problemas, cambios en la ruta de acceso a los productos, etc. Se pueden explorar los datos disponibles para los distintos productos operacionales distribuidos por NOAA en la web de NOMADS (*NOAA Operational Model Archive and Distribution System*) <https://nomads.ncep.noaa.gov/>.

La web desde la que se descargan los datos del RTOFS (*Real Time Ocean Forecast System*) (HYCOM) es <https://nomads.ncep.noaa.gov/pub/data/nccf/com/rtofs/prod/>. De manera diaria se publican los pronósticos (*forecast*) y la estimación del estado actual (*nowcast*) y estos archivos se mantienen publicados únicamente hasta el final del día siguiente. Como parte de la base de datos de pronósticos se descargan diariamente los 244 archivos que contienen variables con frecuencia horaria, *rtofs_glo_2ds_TXXX_1hrly_CCCC.nc* (T=n para *nowcast*, f para *forecast*, XXX=horas, desde 000 hasta 048 para *nowcast*, desde 000 hasta 072 para *forecast*, CCCC=diag para variables de diagnóstico, prog para variables de pronóstico). Se aclara que las variables que se van a utilizar están en los archivos de *forecast* y que los archivos de *nowcast* corresponden a las 48 horas anteriores a la fecha de publicación, coincidiendo los archivos n048 con los f000. Esta descarga supone un tamaño de 107 GB y debe ser ejecutada cuando los archivos se encuentran disponibles, que suele ser cerca de las 10:15 AM GMT0. Los datos que se quieren conservar para

este proyecto son las condiciones de borde de la malla mesh_21, pero para su utilidad en posibles aplicaciones del grupo se decidió incluir a todo el Atlántico Sur. Así, el procesamiento que se hace a los archivos descargados es la extracción de variables, recortado y unificación en un solo archivo. Las variables que se extraen tienen una resolución horizontal de 1/12° y son las siguientes: elevación de la superficie del mar, salinidad en superficie, temperatura en superficie y componentes de la velocidad en superficie y promediadas en vertical.

El proceso de descarga y recorte sucede a diario en la computadora OBRAS09 como una tarea programada que inicia a las 11:00 GMT0 y para la que se emplea un *script* de Python que ubica los archivos descargados en una carpeta con la fecha del día y que luego llama a un *script* de Matlab que realiza la apertura de los archivos para extraer las variables, el recorte de las variables al dominio entre las longitudes [290, 315] y las latitudes [-22, -55] y la unificación del producto en un archivo .mat con la fecha del día (FECHAfn.mat, por ejemplo: 20191119fn.mat), con un tamaño de unos 350 MB. Finalmente los 244 archivos iniciales son borrados. Este proceso se completa en un tiempo variable de aproximadamente 2 horas.

El modelo COPERNICUS (mediante acceso de usuario registrado) permite la descarga de las variables deseadas en el dominio requerido y para el rango de fechas que se defina directamente mediante un *script* de Python que genera la solicitud al servidor. El producto es proporcionado en un archivo en formato NetCDF. Los datos también pueden ser solicitados de manera equivalente a través de la interfaz de su página web <https://marine.copernicus.eu/services-portfolio/access-to-products/> siguiendo los pasos de selección del producto (GLOBAL_ANALYSIS_FORECAST_PHY_001_024), añadir al carrito, selección del conjunto de datos (GLOBAL-ANALYSIS-FORECAST-PHY-001-024-HOURLY-T-U-V-SSH), selección del área geográfica delimitando el rango de longitud y latitud (la zona del Atlántico Sur, LON [-70, -45], LAT [-22, -55]), el rango temporal (se descarga un mes en total, 22 días de *nowcast* y los 10 días disponibles de pronóstico, *forecast*), el rango de la profundidad (para este producto está disponible únicamente en 2D, en superficie), las variables (se seleccionan todas las variables disponibles, temperatura en superficie, componentes de la velocidad y altura de la superficie del mar). Tras la solicitud, el archivo con los datos puede demorar varios minutos en generarse y comienza su descarga. Siguiendo estas pautas los archivos que se reciben tienen un tamaño de unos 775 MB e incluyen a las variables mencionadas con una resolución horizontal de 1/12° y con la particularidad de que su frecuencia horaria corresponde a la hora GMT0+30 minutos, es decir, 00:30, 01:30, 02:30... que es el formato seguido por este sistema. Con un *script* de Matlab se ordenan de manera similar al modelo HYCOM en un archivo con la fecha del día (FECHAfnaCOP.mat, por ejemplo 20191119fnaCOP.mat) de igual tamaño al original que es borrado. Este proceso se lanza desde OBRAS09 a las 18:00 GMT0 y se completa en un tiempo variable desde 3 minutos hasta más de 1 hora siendo el tiempo de espera variable según lo que demore el servidor (de COPERNICUS) en generar el archivo solicitado.

A diferencia del modelo HYCOM, en el modelo COPERNICUS con la llegada del nuevo pronóstico los datos del pronóstico del día anterior son sobrescritos y no se mantienen en la web. Esto hace que sólo puedan ser descargados en el día en el que son generados. El hecho de que los productos puedan ser recibidos únicamente por usuarios registrados permite distribuir avisos a los usuarios que suelen incluir la indisponibilidad, mantenimiento, retraso y llegada de nuevos productos. Este medio también es utilizado anualmente para contactar a los usuarios y solicitar que completen una encuesta de satisfacción que incluye preguntas sobre la utilización de los productos.

Para el primer prototipo del pronóstico se utiliza la marea astronómica pronosticada por el Atlas FES2014 (última actualización de este sistema). Debido a que los pronósticos de marea son determinísticos y que únicamente dependen de la ubicación, se generó en única vez el archivo que contiene la evolución de la marea astronómica para los puntos en el borde de la malla mesh_21 hasta el año 2028, MA_FES2014_borde_MESH21v4_2014_2028.mat.

III.3.2. Forzantes atmosféricos

Los únicos pronósticos de forzantes atmosféricos a los que se ha podido obtener libre acceso son los campos de predicción meteorológica global del NCEP (NOAA) que utiliza el HYCOM. Cada 6 horas (cuatro veces al día) el sistema actualiza los pronósticos con resolución espacial de 0.25° y para la resolución temporal de 1 dato por hora se pronostican 120 horas. Los archivos tienen formato GRIB, se publican en el servidor del GFS-NOAA http://nomads.ncep.noaa.gov/cgi-bin/filter_gfs_0p25.pl y se mantienen durante 10 días. Se pueden descargar de 3 formas: los archivos GRIB completos, a través de un filtro web especificando en la ruta los parámetros de dominio y variables solicitadas o a través de la interfaz web especificando manualmente las variables que se solicitan. Se utiliza la segunda alternativa para descargar las cuatro actualizaciones diarias de los 121 archivos (de hora de pronóstico 000 a 120) que contienen las variables de presión atmosférica en superficie (lev_surface var_PRES) y componentes de la velocidad del viento a 10 m de altura (lev_10_m_above_ground var_UGRD var_VGRD). Las variables seleccionadas son las que se consideran relevantes para la hidrodinámica de la zona (Santoro et al., 2011). La región para la que se descargan los pronósticos es la correspondiente al Atlántico Sur (LON [290, 315], LAT [-22, -55]).

Los archivos que contienen los forzantes del primer pronóstico diario (hora 00:00) se encuentran disponibles pasadas las 5:00 GMT0. Para su descarga en OBRAS09 se emplea una tarea programada que inicia a las 9:00 GMT0 y para la que se emplea un *script* de Python que descarga los 121 archivos en una carpeta con la fecha del día y luego llama a un *script* de Matlab que lee, extrae variables y unifica estos archivos en un .mat con la fecha y hora (FECHAnoaa00.mat, por ejemplo 20191119noaa00.mat) que tiene un tamaño de unos 15MB y borra los archivos originales que ocupaban unos 7.5MB. El proceso de descarga y unión se completa generalmente en menos de 5 minutos. Los pronósticos correspondientes a las 06:00, 12:00, 18:00 podrían utilizarse en sistemas que actualicen los forzantes varias veces al día. Actualmente se están descargando para la base de datos pero como por el momento no son utilizados de forma operacional se descargan todos juntos cuando está disponible el de las 18:00. Se generan para cada día con el nombre FECHAnoaa06.mat, FECHAnoaa12.mat, FECHAnoaa18.mat (por ejemplo 20191119noaa06.mat, 20191119noaa12.mat, 20191119noaa18.mat).

III.3.3. Registros de los mareógrafos

En relación a las observaciones a incluir en la base de datos se debe prestar atención a los mareógrafos de la zona que van a utilizarse para comparar el modelo de pronóstico actual pero también a las fuentes de datos que no mantienen disponibles los datos con el paso del tiempo y que, por lo tanto, no podrán ser obtenidas en caso de requerirse en otro momento para futuros

modelos o comparaciones. La exploración de los datos de altímetro satelital confirmó que es posible acceder en un futuro a los datos por lo que éstos no se incluyen por el momento en la base de datos. Los mareógrafos de Uruguay y Argentina que se incluyen en la base de datos son los que emiten mediciones en tiempo real, publicados por la Administración Nacional de Puertos (ANP) en Uruguay y al Servicio de Hidrografía Naval (SHN) en Argentina.

Los mareógrafos de Uruguay de los que se dispone de datos en tiempo real son Isla de Flores, Muelle Fluvial (Puerto de Montevideo) y La Paloma (desde mitad de 2018). A través de la web de ANP <http://mareografo.anp.com.uy/index.php> se puede acceder a los datos registrados únicamente mediante usuario autorizado (en el grupo de trabajo se cuenta con usuario autorizado); sin usuario se pueden visualizar las series medias únicamente. Los datos proporcionados por la ANP se encuentran referidos al cero Wharton. La web del Servicio de Hidrografía Naval (SHN) de Argentina publica el nivel de marea astronómica frente a nivel medido junto a los valores registrados en las últimas horas para las estaciones de Atalaya, Buenos Aires, La Plata, Mar del Plata, Torre Oyarvide, Pilote Norden, San Clemente, San Fernando y Santa Teresita. Estos datos son sobrescritos en cada actualización y no se mantienen disponibles. Cada estación tiene un plano de referencia local llamado plano de reducción.

Los registros de los mareógrafos publicados por ANP y SHN presentan, ocasionalmente, diversos errores que deben considerarse a la hora de utilizar esas mediciones. Por tanto, se definieron e implementaron una serie de controles de calidad a los mismos. Los errores típicos que se pueden encontrar en las series temporales de registros de mareógrafos (IOC/UNESCO, 2020) incluyen errores aleatorios como los errores de tipeo cuando los registros se transcriben de manera manual (columna equivocada, intercambio de cifras consecutivas, fecha incorrecta), errores electrónicos por ruido en las medidas (picos), problemas durante la transmisión del dato (huecos o datos faltantes), datos fuera del rango de medida (*outliers*), calibración del sensor (modificación del cero de referencia), etc. La depuración de nivel inicial que se lleva a cabo a las observaciones registradas por mareógrafos constituye un primer acercamiento al control de calidad de datos tomando como base los registros del 2018 y que periódicamente se deberá reevaluar para su actualización. En futuras versiones del sistema de pronóstico se mejorará el control de calidad de datos implementando su actualización automática. Como referencias bibliográficas se incluyen IOC/UNESCO (2020) y UNESCO (1993). En el Anexo 3 se describe el proceso de filtrado para cada fuente basado en las medidas de 2018. Para las medidas de la ANP se produce la descarga diaria y se aplica un filtrado para cada estación como control de calidad de datos. Las observaciones provenientes del SHN deben ser descargadas varias veces al día y el control de calidad de datos actual es únicamente eliminar los *outliers*.

Las observaciones publicadas en la web de ANP se descargan y procesan en OBRAS09 a través de tarea programada a las 8:00 GMT0. Utilizando un *script* de Python se descarga un archivo .csv para cada mareógrafo conteniendo los datos registrados en los últimos 4 días (1 dato cada 5 minutos) y se llama a un *script* de Matlab para que realice el filtrado y guarde el resultado en un .mat para cada estación (denominados FECHAIdF.mat, FECHAMVD.mat, FECHALPa.mat, por ejemplo 20191119IdF.mat, 20191119MVD.mat, 20191119LPa.mat). Los datos de la web del SHN se descargan cada hora a través de tarea programada en OBRAS09 mediante un *script* de Python. La frecuencia de las medidas disponibles es de un dato por hora con la particularidad de que su frecuencia horaria corresponde a la hora +45 minutos (por ejemplo 00:45, 01:45, 02:45, etc.), el *script* genera un archivo diario que es editado durante el día y contiene la acumulación de los datos publicados ese día para todos los mareógrafos argentinos en un .csv. El filtrado y unificación a través de un *script* de Matlab ocurre una vez al día a través de la tarea programada

en OBRAS09 a las 03:00GMT0 y se genera un archivo que contiene los datos medidos en los últimos 4 días en un .mat (FECHAArgentina.mat, por ejemplo 20191119Argentina.mat).

La obtención de los archivos de niveles depurados demora menos de un minuto para ambas fuentes y los archivos generados tienen un tamaño de menos de 30KB en total.

III.4. Definición de la estructura del pronóstico

La estructura que va a seguir el sistema está condicionada por la disponibilidad de los datos de entrada de fuentes externas ya que los datos que vayan a ser utilizados deberán haber sido incorporados a la base de datos previamente al inicio del pronóstico. Los datos de entrada son los forzantes atmosféricos, condiciones de borde oceánicas, caudales de agua dulce y, dado que se quiere incluir como salida del sistema una comparación frente a las observaciones recientes, los datos de los mareógrafos que se vayan a comparar también deben estar en la base de datos antes de iniciar el pronóstico. En la Figura 45 se presenta un esquema de los componentes del sistema de pronóstico generado, los cuales se describen en esta sección.

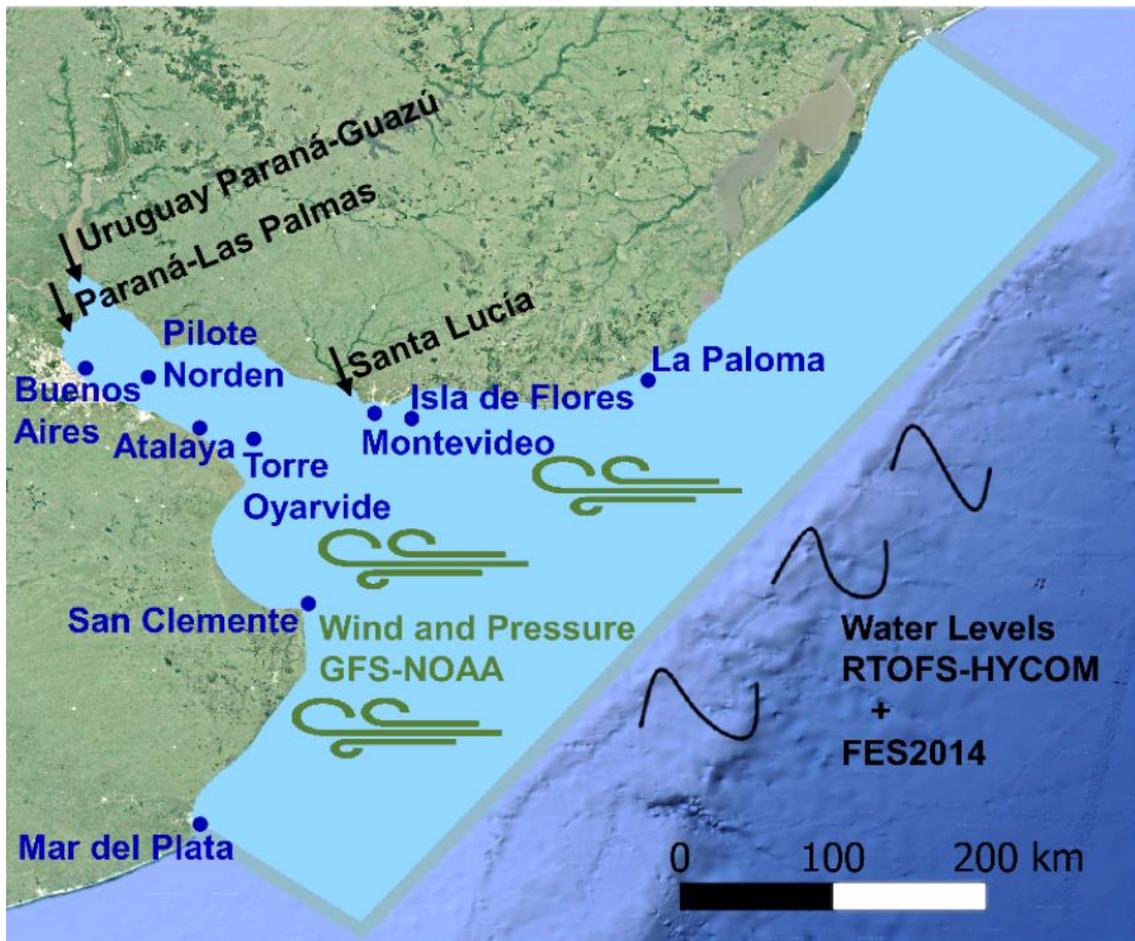


Figura 45 – Esquema de los componentes del sistema de pronóstico del Río de la Plata: condiciones de borde oceanográficas, forzantes atmosféricos e ingresos de agua dulce. Ubicación de los mareógrafos con datos disponibles en tiempo real.

Los campos de predicción meteorológica global del NCEP (NOAA), que utiliza el HYCOM, se encuentran con libre disponibilidad y su adecuada resolución y experiencia en la utilización de

los reanálisis NCEP en el grupo de trabajo hizo que fueran seleccionados para este primer prototipo. A partir de esta definición se determinó utilizar para las condiciones de borde oceanográficas el modelo HYCOM de manera que coincida el forzante atmosférico del modelo global con el del pronóstico del Río de la Plata. El pronóstico suministrado por HYCOM es de 72 horas mientras que la disponibilidad de forzantes atmosféricos llega hasta las 120 horas. Por tanto el pronóstico del Río de la Plata está condicionado por el modelo de marea global teniendo un horizonte de pronóstico de 72 horas.

Se imponen los niveles pronosticados con frecuencia horaria por HYCOM en los 190 puntos del borde de la malla mesh_21 (presentada en la Figura 44). Debido a que los pronósticos de HYCOM se actualizan una vez por día se decide que el sistema tenga también una frecuencia diaria y los pronósticos de forzantes atmosféricos que se utilizarán son los de la hora 00. Para la marea astronómica se agrega en los puntos del borde la marea pronosticada por el Atlas FES2014. La Figura 45 presenta esta selección de manera esquemática.

En relación a los aportes fluviales, no se cuenta actualmente con pronósticos por tanto se define utilizar caudales constantes correspondientes a la media histórica de los últimos 10 años para las 3 secciones de ingreso de la malla seleccionada, mesh_21. Estos valores resultan en 20,547 m³/s para el ingreso de los ríos Uruguay y Paraná-Guazú y 5,825 m³/s correspondiente al río Paraná Las Palmas, lo que totaliza un caudal de 26,372 m³/s afluente al Río de la Plata. Se incluye también el Río Santa Lucía con 180 m³/s.

Como parte del sistema de pronóstico se establece la evaluación del mismo considerando los datos disponibles en tiempo real de niveles medidos en mareógrafos. Así, se incluyen los 3 mareógrafos de la ANP, Uruguay, y 6 de los 9 mareógrafos del SHN, Argentina. No se incluyen los mareógrafos de San Fernando y La Plata por ubicarse en un afluente del Río de la Plata ni se incluye el mareógrafo de Santa Teresita ya que actualmente se encuentra fuera de servicio. Las comparaciones se realizarán, por tanto, utilizando las observaciones en Isla de Flores, Montevideo, La Paloma, Atalaya, Buenos Aires, Mar del Plata, Pilote Norden, San Clemente y Torre Oyarvide (ver en la Figura 45 la ubicación).

Como se mencionó anteriormente, para poder iniciar los cálculos con TELEMAC-2D se debe disponer de los datos de pronósticos y observaciones que van a ser utilizados previamente preprocesados y en la base de datos. Atendiendo a hora a la que se encuentran disponibles, la principal limitante es el acceso a los datos del sistema HYCOM. Basándose en los tiempos de descarga de HYCOM se establece que el sistema comience a ejecutarse a través de tarea programada en OBRAS09 a las 14:00GMT0. Para la implementación del sistema se deben generar los *scripts* que hagan la conversión desde los archivos de la base de datos hacia archivos que pueda leer el sistema TELEMAC-2D. Una vez ejecutado, debe existir un postproceso que permita la fácil lectura de resultados y una salida de comparación con las observaciones de las últimas horas. El sistema genera un pronóstico a 72 horas y se establece que cada pronóstico diario tome como condiciones iniciales el pronóstico a 24 horas del día anterior. Así que estos datos deberán extraerse como un input para el día siguiente.

III.5. Implementación del sistema de pronóstico

En este apartado se describen los pasos seguidos para la implementación del sistema de pronóstico para el Río de la Plata y su Frene Marítimo de forma que genere como resultados los

niveles para las próximas 72 horas y su comparación frente a los niveles registrados en las horas anteriores. Se parte de los archivos disponibles tras la descarga diaria e incorporación a la base de datos. Se realiza un preproceso para su lectura por parte de TELEMAC-2D y se ejecuta el programa que tiene en cuenta otros datos de entrada que no precisan de una actualización diaria, como son los caudales de agua dulce y los pronósticos de marea astronómica en el borde oceánico. Para la salida final se realiza un postproceso que permite la comparación frente a los datos de nivel medidos. A continuación se describen los pasos seguidos por el sistema de forma diaria a través de tarea programada en OBRAS09 a las 14:00GMT0 cuyo esquema se presenta en la Figura 46.

Preproceso Borde Oceánico: partiendo del archivo generado en la descarga diaria: FECHAfn.mat (por ejemplo 20191119fn.mat), a través de un *script* de Matlab se interpola a los puntos del borde de la malla de cálculo la variable de nivel, se añade la componente astronómica a partir del atlas FES2014, y se genera el archivo de condiciones de borde que leerá TELEMAC-2D: BC_SSE_FES2014_FECHA_mesh_21.txt (por ejemplo BC_SSE_FES2014_20191119_mesh_21.txt). En caso de que los datos de HYCOM no pudieran ser obtenidos, está implementado que el sistema de pronóstico del Río de la Plata trabaje únicamente con datos de marea astronómica como condición de borde.

Preproceso forzantes atmosféricos: partiendo del archivo generado en la descarga diaria: FECHAnoaa00.mat (por ejemplo 20191119noaa00.mat), a través de un *script* de Matlab se interpola a la malla de cálculo y se genera el archivo de forzantes que leerá TELEMAC-2D: FECHAnoaa_mesh_21.slf (por ejemplo 20191119noaa_mesh_21.slf). En caso de que no esté disponible el dato de forzante atmosférico el pronóstico no se realizará. Sin embargo, es de destacar que es poco frecuente que esto suceda. Para poder reanudar los pronósticos se deberá actuar y se podrán emplear los datos con frecuencia horaria en el momento en que estuvieran disponibles o los datos de una fuente externa como es <https://rda.ucar.edu/datasets/ds084.1/> (NCEP et al., 2015), que tiene el respaldo de los pronósticos del GFS pero con frecuencia de un dato cada 3 horas y la misma resolución espacial de 1/4°.

Ejecución TELEMAC: el modelo TELEMAC-2D se corre tomando como input el archivo de forzantes y el de condiciones de borde, así como las condiciones iniciales correspondientes al pronóstico a 24 horas generado en la corrida del día anterior. Se utiliza la opción de paralelización con 8 cores de forma de reducir los tiempos de cómputo. Se obtiene el archivo de resultados que incluye el pronóstico a 72 horas para todo el dominio de niveles y componentes Este y Norte de la velocidad integrada en vertical: HYCOM_SFECHA.slf (por ejemplo HYCOM_S20191119.slf).

Postproceso TELEMAC: Se extrae el pronóstico a 24 horas que servirá de condiciones iniciales para la próxima corrida: HYCOM_SFECHA_rst.slf (por ejemplo HYCOM_S20191119_rst.slf). Además, a través de un *script* de Matlab se extraen las series horarias de nivel en 22 puntos para las próximas 72 horas del pronóstico: HYCOM_SFECHA_series.mat (por ejemplo HYCOM_S20191119_series.mat). Este archivo servirá como dato de entrada para la comparación de los resultados del modelo de pronóstico con los datos medidos en los mareógrafos.

Resultados Estadísticos: Para la simulación de hace 72 horas (HYCOM_SFECHA-3_series.mat, por ejemplo HYCOM_S20191116_series.mat) se calculan los estadísticos del comportamiento del pronóstico durante las primeras 24 horas (1 día), de la hora 24 a la hora 48 (2 días) y de la hora 48 a la hora 72 (3 días). Estas comparaciones estadísticas se realizan a través de un *script* de

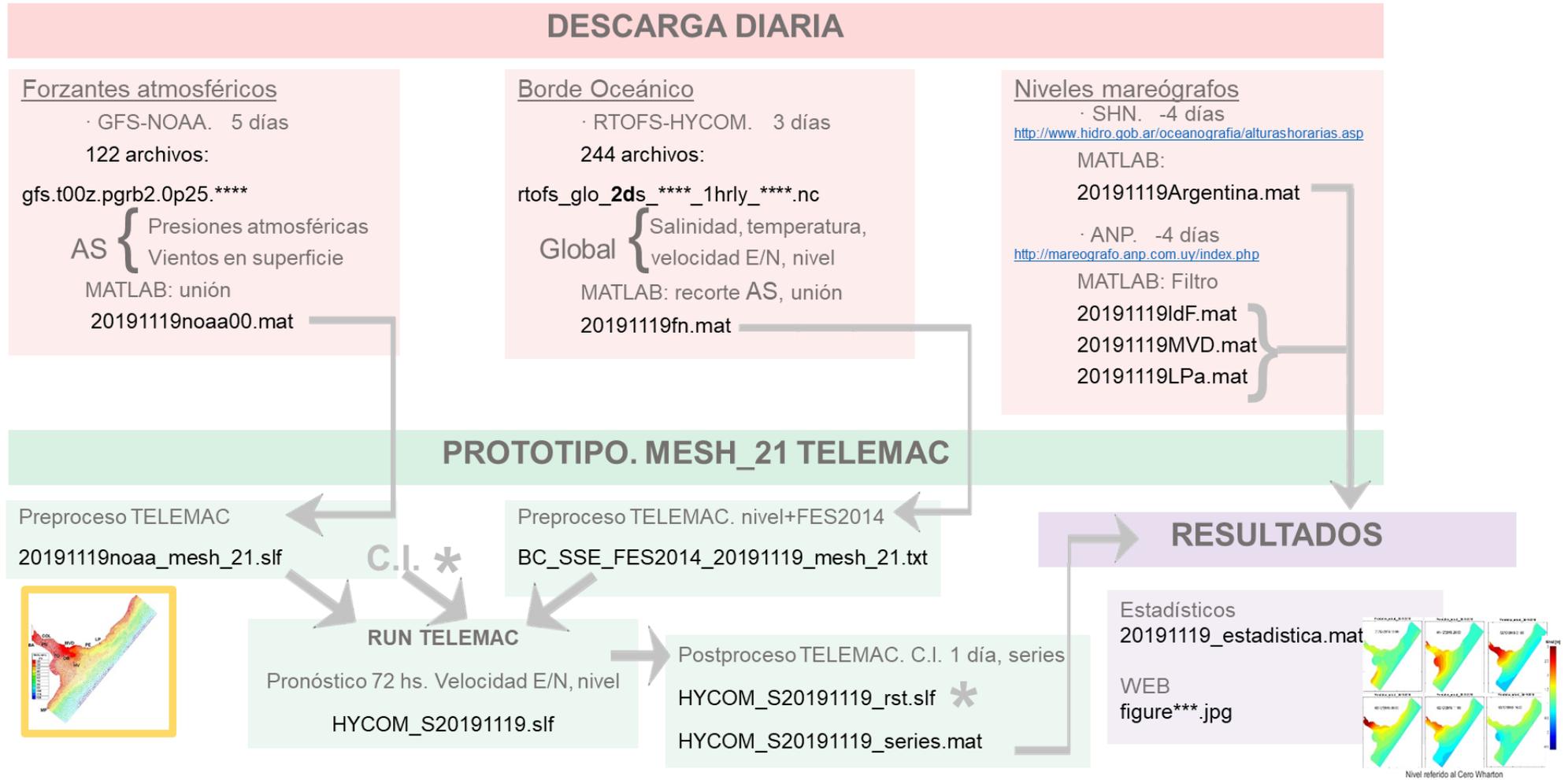


Figura 46 – Esquema del sistema de pronóstico generado.

Matlab y se obtiene el archivo FECHA-3_estadistica.mat (por ejemplo 20191116_estadistica.mat, que se obtuvo el día 20191119). Los estadísticos utilizados son el error cuadrático medio (RMSE), la correlación de Pearson y el sesgo (BIAS) y fueron descritos en el apartado II.1.4.

Resultados Difusión WEB: Se implementó la web <https://www.fing.edu.uy/imfia/pronostico-marea/> para la publicación del sistema de pronóstico. En su interfaz se muestra la salida gráfica para cada mareógrafo, y la evolución de los niveles horarios del pronóstico actual para el dominio de cálculo. A través de un *script* de Matlab, tomando como input el archivo generado por TELEMAT-2D (HYCOM_SFECHEA.slf, por ejemplo HYCOM_S20191119.slf) se generan figuras con los mapas que muestran cómo será la evolución de los niveles para las próximas 72 horas. Partiendo de las series de pronóstico de las últimas corridas (HYCOM_SFECHEA-3_series.mat, HYCOM_SFECHEA-2_series.mat, HYCOM_SFECHEA-1_series.mat, HYCOM_SFECHEA_series.mat, por ejemplo HYCOM_S20191116_series.mat, HYCOM_S20191117_series.mat, HYCOM_S20191118_series.mat, HYCOM_S20191119_series.mat) y de los datos de niveles medidos descargados, se generan figuras para 9 mareógrafos que muestran la evolución de los pronósticos en comparación con los niveles medidos. De esta manera, se presenta una comparación visual con los datos medidos del pronóstico del día (correspondiente a FECHA), del pronóstico generado hace 24 horas (correspondiente a FECHA-1), del pronóstico generado hace 48 horas (correspondiente a FECHA-2) y el generado hace 72 horas (correspondiente a FECHA-3). Además, estas figuras vienen acompañadas de un resumen de estadísticos para la corrida de hace 3 días (correspondiente a FECHA-3).

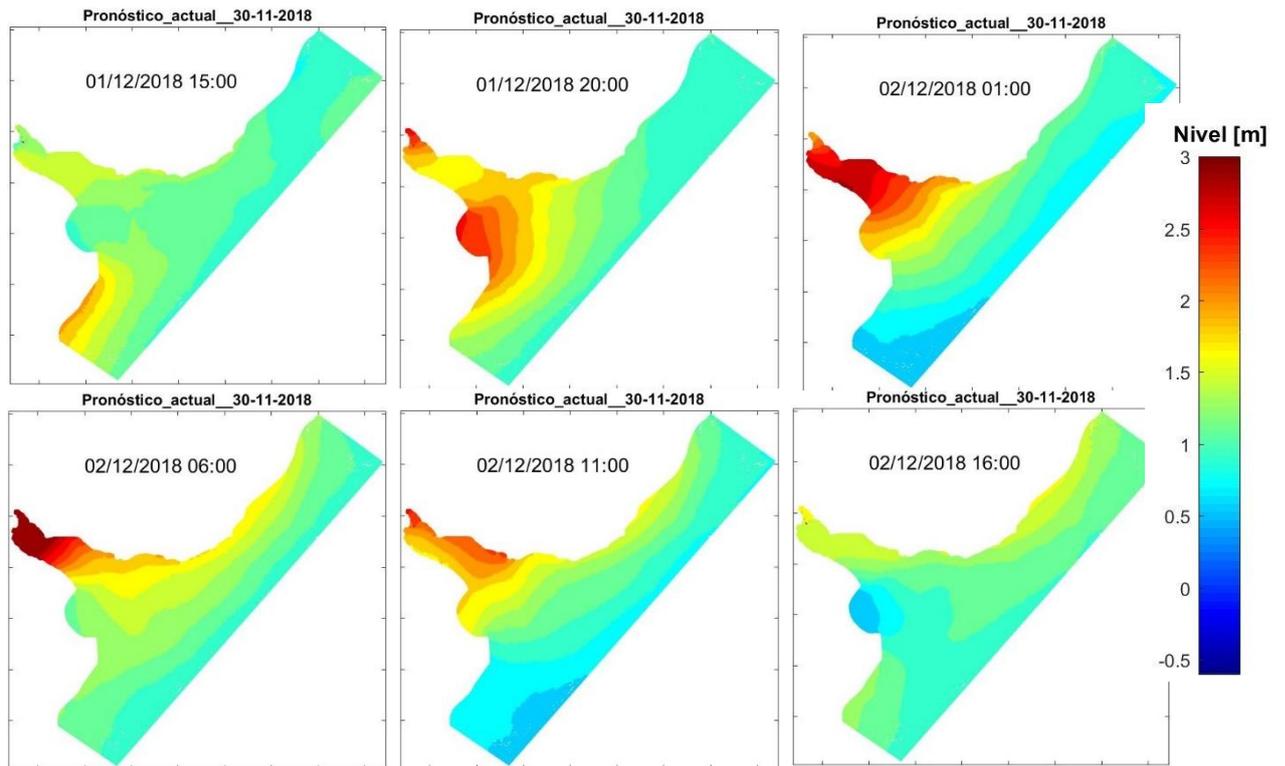
El sistema de pronóstico implementado inicia el 1° de enero de 2017 partiendo de una condición inicial de reposo, nivel inicial uniforme igual a 0.91 m y un valor de salinidad constante igual a 35 ppm. El primer año de pronóstico (correspondiente al 2017) se descarta para la evaluación de resultados teniendo en cuenta que la condición inicial de salinidad requiere que se simulen varios meses para independizarse de la misma.

III.6. Resultados del Sistema de pronóstico

El prototipo operacional obtiene el pronóstico de niveles y corrientes 2D para toda la malla de cálculo para los próximos 3 días. En la Figura 47 se muestra un ejemplo de los campos de niveles obtenidos para el pronóstico del 30-11-2018.

Como parte de los resultados y para su comparación con los mareógrafos disponibles se extraen series en 22 puntos que corresponden a: Atalaya, Buenos Aires, La Plata, Mar del Plata, Torre Oyarvide, Pilote Norden, San Clemente, San Fernando, Santa Teresita, Montevideo, Isla de Flores, La Paloma, Punta del Este, Colonia, Conchillas, Carmelo, Punta Brava, Punta Yeguas, Boya Oceanográfica, Hidrovía, Boya GS, Juan Lacaze. En la Figura 48 y Figura 49 se puede ver el ejemplo de salida gráfica para la comparación frente a los datos medidos por los 9 mareógrafos durante el pronóstico del 18-11-2019.

El sistema obtiene los resultados numéricos en OBRAS09 en aproximadamente 5 minutos. Las salidas gráficas llegan a demorar más de una hora siendo la limitante la generación de los 72 mapas con los campos para las próximas horas. La actualización de la web se puede encontrar disponible a las 16:00GMT0.



Nivel referido al Cero Wharton

Figura 47 – Ejemplo de la evolución temporal de los niveles para el pronóstico del 30-11-2018 en todo el dominio de cálculo.

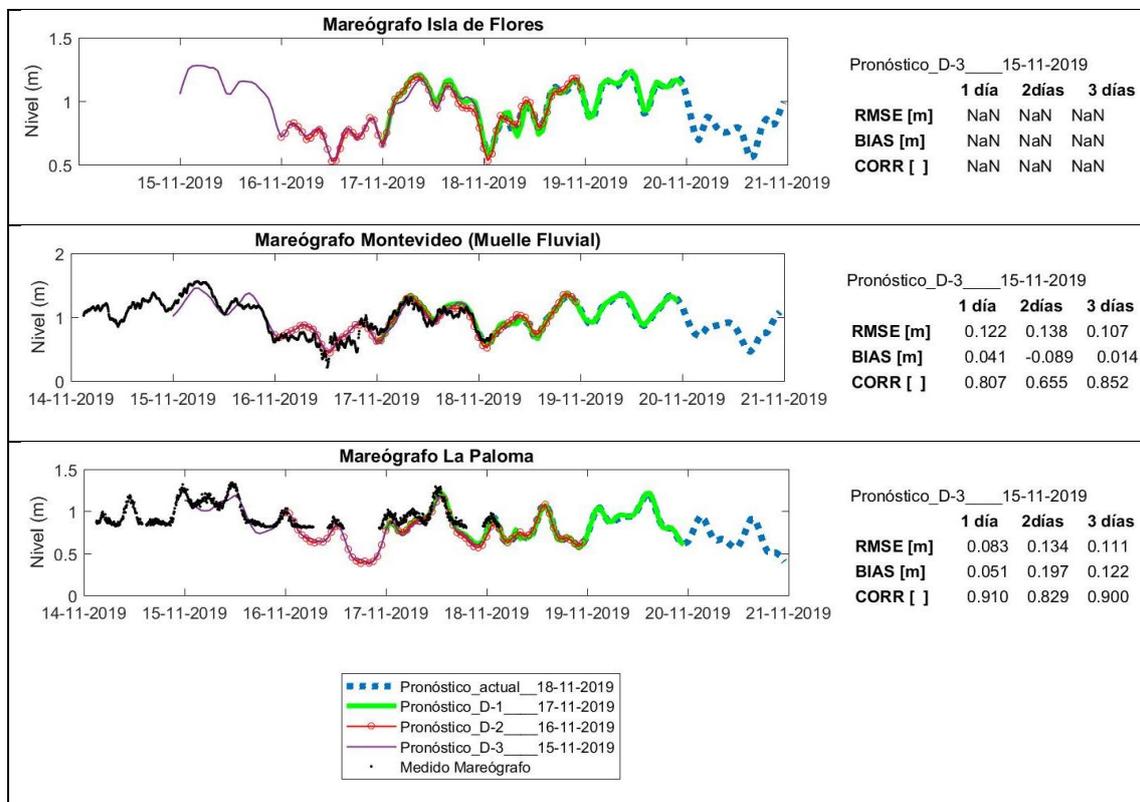


Figura 48 - Salida gráfica del pronóstico. Ejemplo de la comparación entre los resultados del modelo y los mareógrafos de Isla de Flores, Montevideo y La Paloma para el pronóstico del 18-11-2019.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

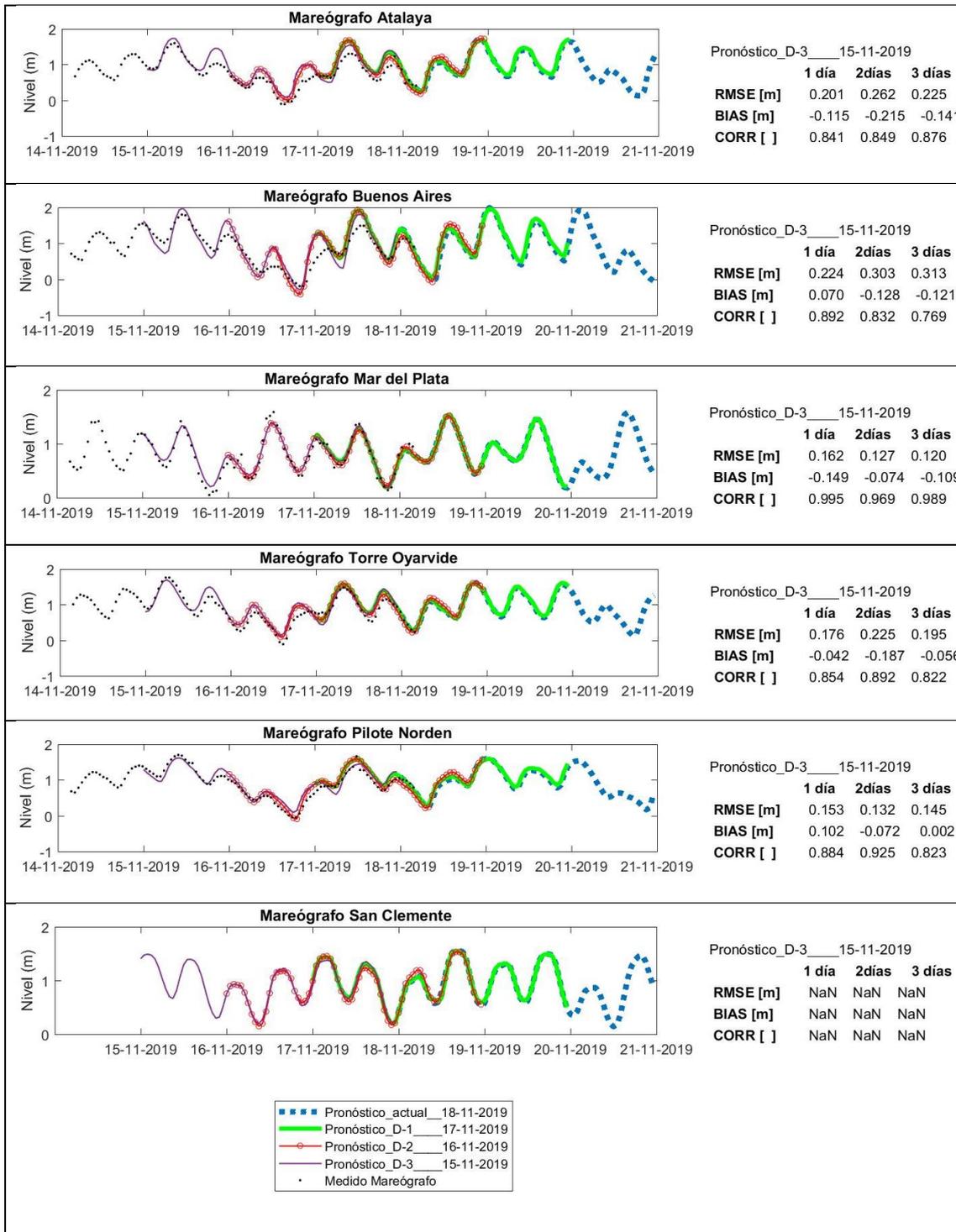


Figura 49 - Salida gráfica del pronóstico. Ejemplo de la comparación entre los resultados del modelo y los mareógrafos de Atalaya, Buenos Aires, Mar del Plata, Torre Oyarvide, Pilote Norden y San Clemente para el pronóstico del 18-11-2019.

IV. Evaluación

En este capítulo se describe la metodología para la evaluación del sistema de pronóstico generado, se presentan los resultados y se realiza un análisis de los mismos.

IV.1. Metodología

Para la evaluación del sistema se utilizan los pronósticos generados durante los años 2018 y 2019. Como se mencionó en el capítulo anterior, el sistema de pronóstico es forzado con niveles globales del modelo HYCOM y vientos horarios del GFS NOAA, cuando ambos están disponibles. En caso de que no estén disponibles los datos del modelo global HYCOM el sistema se ejecuta únicamente con marea astronómica procedente del atlas FES2014 y en caso de que no se cuente con el forzante horario del NOAA en tiempo real el sistema se ejecuta utilizando pronósticos con una menor resolución temporal (1 dato cada 3 horas). Para realizar la evaluación del pronóstico generado para los años 2018 y 2019 se utilizan todos los valores pronosticados pero se determina en primer lugar en qué momentos hubo faltantes de los forzantes globales para tener en cuenta en el análisis de calidad.

Luego, los resultados del modelo de pronóstico se evalúan respecto a los registros de nivel de los mareógrafos disponibles. Los mareógrafos de Uruguay, cuyos registros se obtienen de la ANP, tienen como plano de referencia el cero Wharton que corresponde con el cero del modelo de pronóstico generado en esta tesis. Para los datos proporcionados por el SHN, Argentina, en la Tabla 4 se presenta la conversión a cero Wharton (Santoro, 2011; <http://www.hidro.gov.ar/oceanografia/alturashorarias.asp>). Para el mareógrafo de San Clemente no se realiza ninguna conversión pues no se pudo acceder a la información de la estación.

Mareógrafo	Conversión a cero Wharton
Atalaya	SHN + 0.1367 m
Buenos Aires	SHN + 0.1367 m
Mar del Plata	SHN - 0.2175 m
Torre Oyarvide	SHN + 0.1367 m
Pilote Norden	SHN + 0.45 m
San Clemente	No se pudo determinar

Tabla 4 – Conversión a cero Wharton para los registros proporcionados por el Servicio de Hidrografía Naval (SHN), Argentina.

Se analizan los resultados de 2018 para las primeras 24 horas pronosticadas cada día (primer día de pronóstico) utilizando los valores que el modelo da como resultado para cada ubicación de mareógrafo y únicamente habiendo corregido el dato de mareógrafo para su conversión a cero Wharton. Se indican las diferencias encontradas en los resultados del pronóstico para el segundo y tercer día y también con lo obtenido para el 2019. Por otra parte, se realizan comparaciones restando la media anual que permiten ver el comportamiento insesgado del sistema y el efecto que tendría una mejor estimación de la media en los distintos puntos.

Las comparaciones que se realizan frente a datos de mareógrafos consisten en el graficado del diagrama de dispersión (*scatter plot*) y el cálculo de estadísticos cuya formulación se presentó

en el apartado II.1.4. Los estadísticos que se han calculado son: error cuadrático medio **RMSE**, diferencia de media, o sesgo **BIAS**, correlación **Correlación Pearson**, frecuencia central **CF(ζ)** con un umbral de $\zeta = 0.15m$, frecuencia positiva de *outlier* **POF(ζ)** con un umbral de $\zeta = 0.3m$ y frecuencia negativa de *outlier* **NOF(ζ)** con un umbral de $\zeta = 0.3m$. Estas comparaciones se realizan directamente frente a los valores que el modelo da como resultado para cada ubicación de mareógrafo. Se repiten las comparaciones de los estadísticos **CF(ζ)**, **POF(ζ)** y **NOF(ζ)** para los valores sin tener en cuenta la media anual (para ello se resta su media de ese año al valor del modelo y a los registros del mareógrafo también se resta su media anual).

Para cada comparación se separa el comportamiento del pronóstico durante las primeras 24 horas (1 día), de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Así, para cada estadístico se tienen 3 series (pronósticos a 1 día, 2 días y 3 días) para cada uno de los dos años evaluados (2018 y 2019). Cada una de estas series contiene los 365 valores de ese estadístico durante ese año y junto con esa serie se presenta su histograma y su media anual.

Cada una de las dos fuentes de datos de registros de mareógrafos utilizadas tiene una frecuencia horaria diferente que es también diferente a la frecuencia de los niveles que arroja el modelo por lo que se deben llevar a la misma frecuencia. Así, para la comparación frente a los registros de los mareógrafos de la ANP (Uruguay) dado que se tiene 1 dato cada 5 minutos, se procedió del siguiente modo. A partir de los datos conservados tras la depuración descrita en el Anexo 3, se obtiene la media de los valores registrados desde media hora antes hasta media hora después de la hora a comparar (por ejemplo, para la comparación de las 04:00 se hará la media de los datos conservados desde las 03:30 hasta la 04:30). Para la comparación frente a los registros obtenidos de la web del SHN, dado que la frecuencia de las medidas disponibles es de un dato por hora con la particularidad de que su frecuencia horaria corresponde a la hora +45 minutos (por ejemplo 00:45, 01:45, 02:45, etc.), se procedió del siguiente modo. A partir de los datos conservados tras la depuración descrita en el Anexo 3, se obtiene la media ponderada del registro anterior y posterior a la hora a comparar (por ejemplo, para la comparación de las 04:00 se hará la media de 3 veces el registro de las 03:45 y 1 vez el registro de las 04:45). Para ambas fuentes de datos la media se realiza siempre que haya al menos 1 dato disponible y en caso contrario se considera dato faltante. En el apartado de datos faltantes también se resumen los datos faltantes de los mareógrafos obtenidos de este modo.

IV.2. Resultados

IV.2.1. Datos faltantes

En las tablas Tabla 5 a Tabla 8 se señalan las fechas para las cuales no se contó con información de los forzantes globales. En relación a las condiciones de contorno oceánicas de HYCOM se dan un total de 53 datos faltantes para el año 2018 (Tabla 5) correspondientes a 11 intervalos y de 17 datos faltantes para el año 2019 (Tabla 6) correspondientes a 7 intervalos, lo que es equivalente a un 14.5% y a un 4.7% respectivamente. En estas ocasiones el sistema se ejecutó únicamente con pronóstico de marea astronómica para las condiciones de contorno oceánicas.

En relación a los forzantes atmosféricos, se dan un total de 75 datos faltantes para el año 2018 (Tabla 7) correspondientes a 4 intervalos y de 12 datos faltantes para el año 2019 (Tabla 8) correspondientes a 2 intervalos, lo que es equivalente a un 20.5% y a un 3.3% respectivamente. Para la ejecución del modelo se pudieron obtener para todos los casos de datos faltantes los pronósticos de forzantes atmosféricos con una resolución horaria menor (1 dato cada 3 horas) y la misma resolución espacial de 1/4º.

En la Tabla 9 se hace un resumen de los porcentajes de datos faltantes para los años 2018 y 2019 en el que se incluyen los datos de condiciones de contorno de HYCOM, los datos de forzantes de NOAA y, además, los registros de los 9 mareógrafos frente a los que se compara el modelo. Como se dijo anteriormente, en el caso de mareógrafos, los datos faltantes incluyen tanto datos a los que no se pudo acceder como datos eliminados por la depuración descrita en el Anexo 3 y el procedimiento que lleva los registros a una frecuencia horaria que coincide con la del modelo. Cuando hay datos faltantes en los mareógrafos esos instantes no se pueden incluir en el cálculo de estadísticos.

Intervalos con datos faltantes HYCOM. Año 2018		
Inicio	Fin	Duración
20180101	20180107	7 días
20180505	20180505	1 día
20180522	20180523	2 días
20180527	20180530	4 días
20180603	20180603	1 día
20180616	20180616	1 día
20180701	20180701	1 día
20180706	20180706	1 día
20180712	20180712	1 día
20180727	20180827	32 días
20181005	20181006	2 días
	Total	53 días

Tabla 5 – Intervalos con datos faltantes HYCOM. Año 2018

Intervalos con datos faltantes HYCOM. Año 2019		
Inicio	Fin	Duración
20190108	20190113	6 días
20190315	20190315	1 día
20190323	20190323	1 día
20190401	20190402	2 días
20190412	20190416	5 días
20190616	20190616	1 día
20190715	20190715	1 día
	Total	17 días

Tabla 6 – Intervalos con datos faltantes HYCOM. Año 2019

Intervalos con datos faltantes NOAA. Año 2018		
Inicio	Fin	Duración
20180101	20180212	43 días
20180515	20180517	3 días
20180724	20180820	28 días
20180822	20180822	1 día
	Total	75 días

Tabla 7 – Intervalos con datos faltantes NOAA. Año 2018

Intervalos con datos faltantes NOAA. Año 2019		
Inicio	Fin	Duración
20190101	20190110	10 días
20190415	20190416	2 días
	Total	12 días

Tabla 8 – Intervalos con datos faltantes NOAA. Año 2019

Resumen de % de datos faltantes		
	Año 2018	Año 2019
Condiciones de borde oceánicas		
HYCOM	14.5%	4.7%
Forzantes atmosféricos		
NOAA	20.5%	3.3%
Mareógrafos		
Isla de Flores	0.6%	10.5%
Montevideo	0.01%	0.06%
La Paloma	36.0%	22.5%
Atalaya	35.6%	62.0%
Buenos Aires	14.1%	2.0%
Mar del Plata	16.1%	4.3%
Torre Oyarvide	18.9%	6.8%
Pilote Norden	30.6%	11.0%
San Clemente	27.1%	68.4%

Tabla 9 - Resumen de datos faltantes 2018 y 2019. Condiciones de borde HYCOM, forzantes NOAA, mareógrafos.

IV.2.2. Comparación frente a datos de mareógrafos

La variación temporal para el año 2018 de los estadísticos se presenta a continuación en las figuras de la Figura 50 a la Figura 68. Se incluyen: diagrama de dispersión (*scatter plot*) (Figura 50), error cuadrático medio **RMSE** (Figura 51 a Figura 53), diferencia de media, o sesgo **BIAS** (Figura 54 a Figura 56), correlación **Correlación Pearson** (Figura 57 a Figura 59), frecuencia central **CF(0.15m)** (Figura 60 a Figura 62), frecuencia positiva de *outlier* **POF(0.3m)** (Figura 63 a Figura 65) y frecuencia negativa de *outlier* **NOF(0.3m)** (Figura 66 a Figura 68).

En el Anexo 4 se presentan los resultados obtenidos para el año 2019, así como los resultados tras restar la media anual de **CF(0.15m)**, **POF(0.3m)** y **NOF(0.3m)** para 2018 y 2019. El análisis de resultados aparece en el siguiente apartado IV.3.

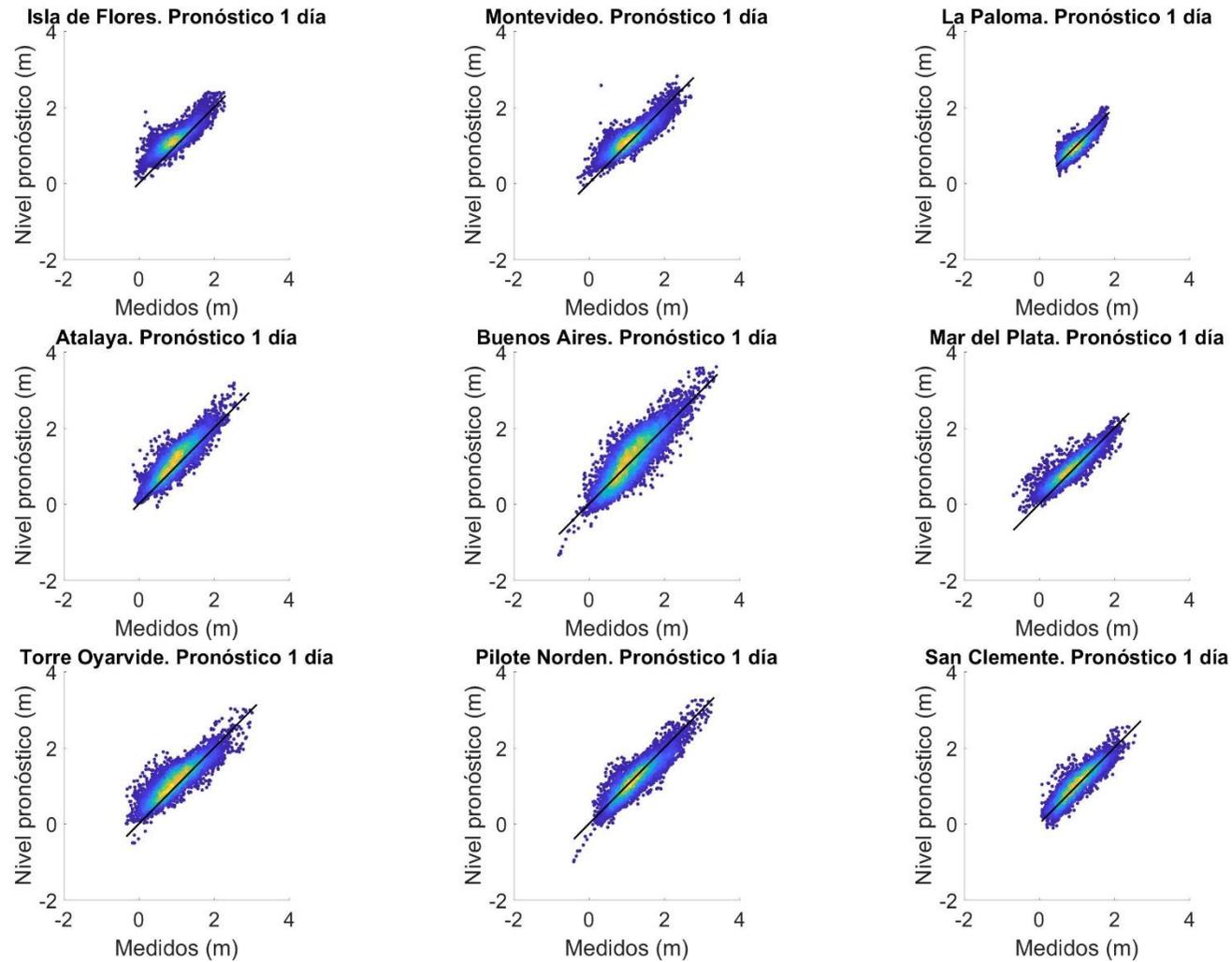


Figura 50 - Diagramas de dispersión para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día)

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

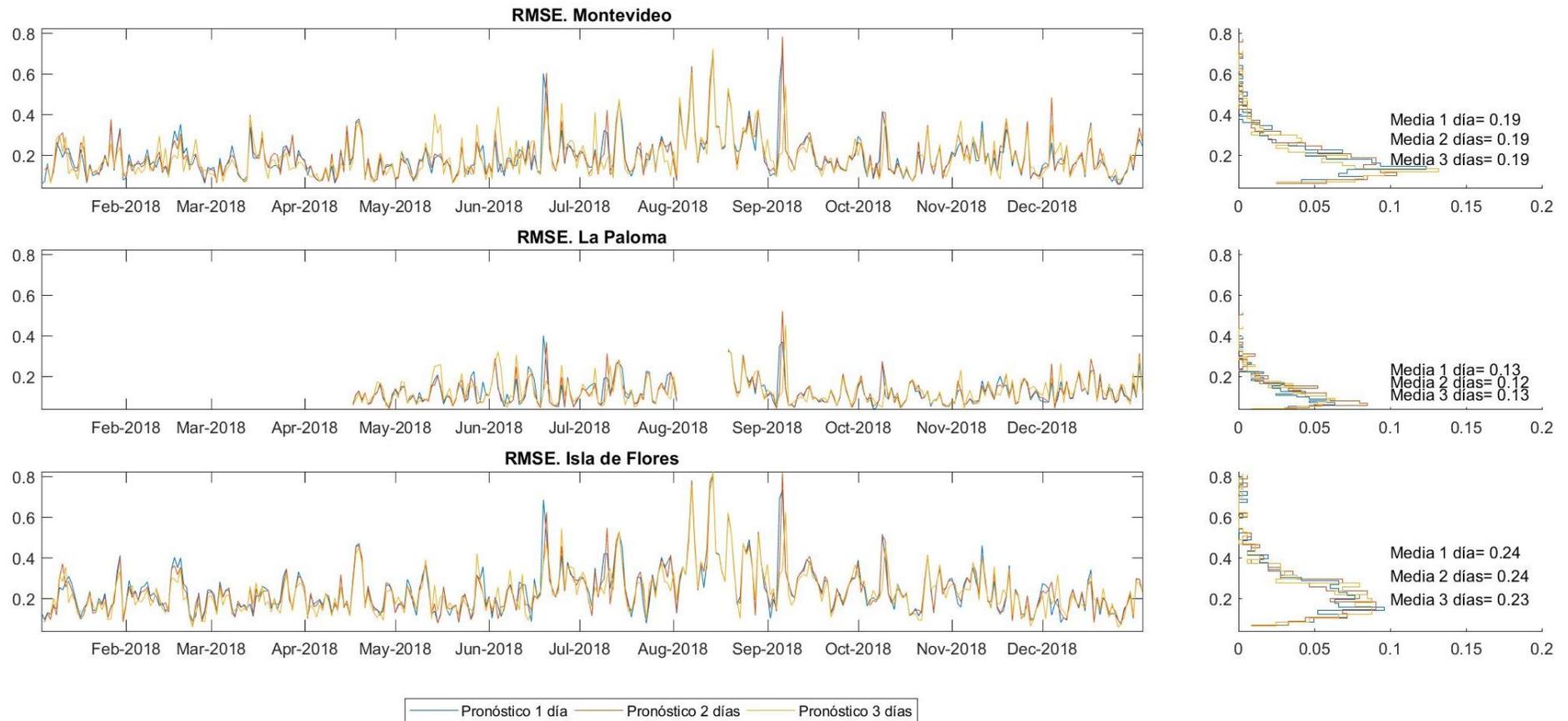


Figura 51 - Error cuadrático medio RMSE [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

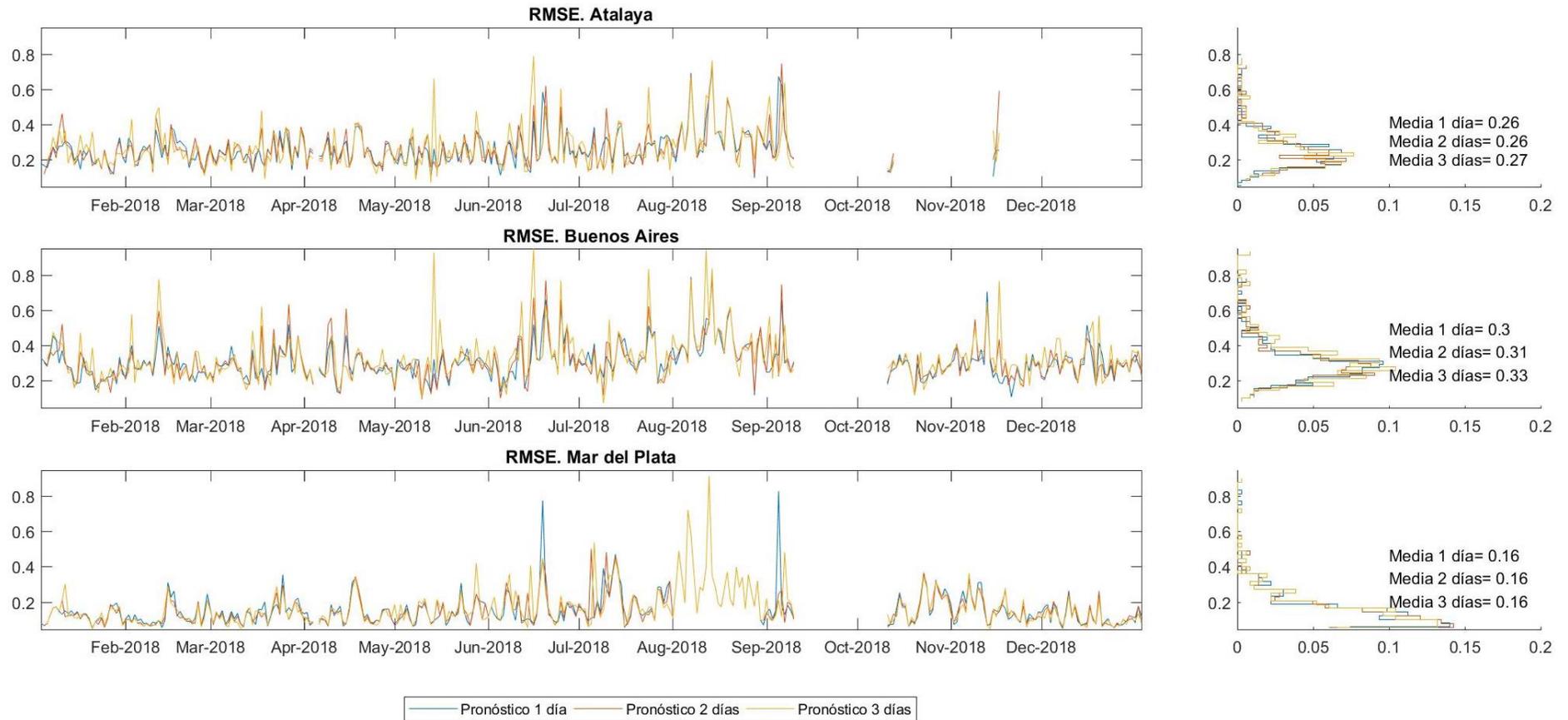


Figura 52 - Error cuadrático medio RMSE [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

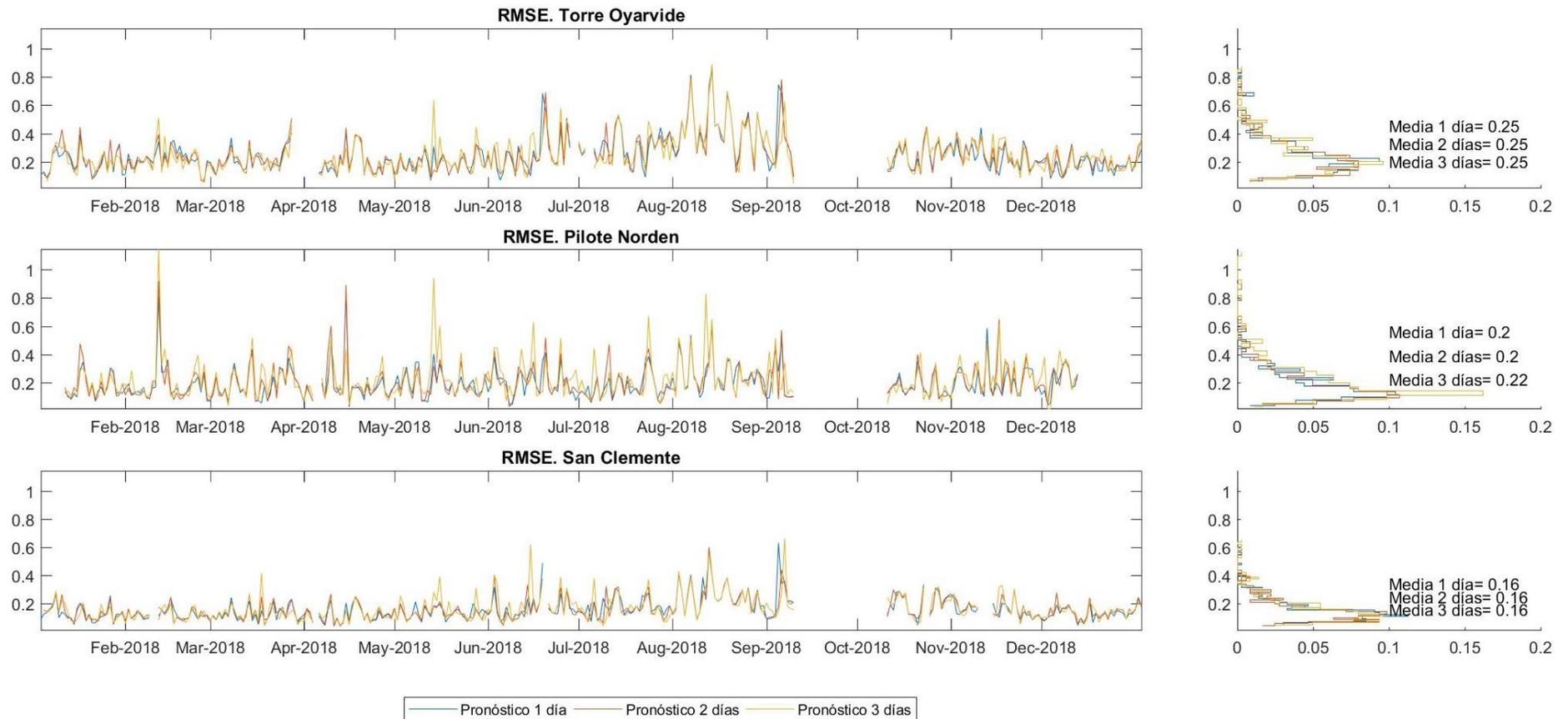


Figura 53 - Error cuadrático medio RMSE [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

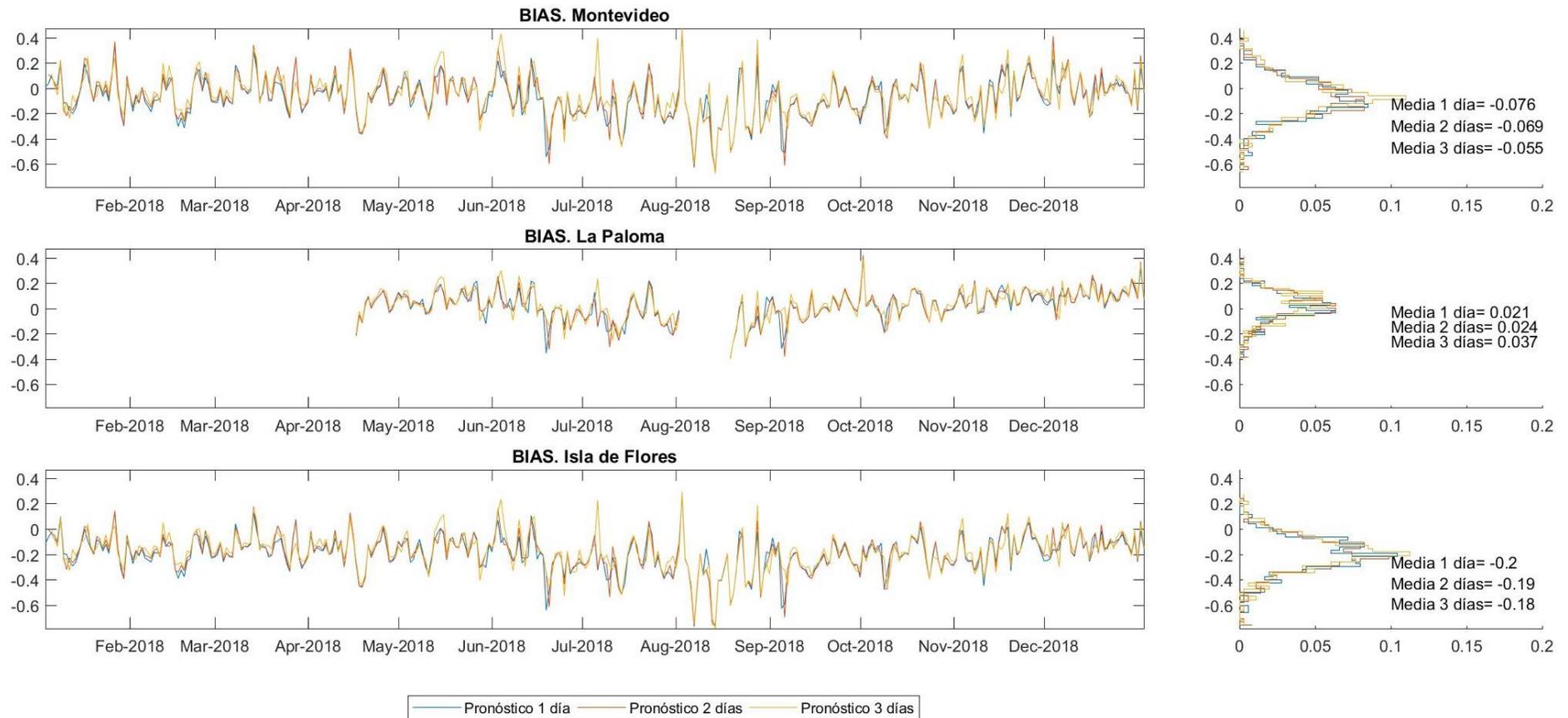


Figura 54 - Diferencia de media, o sesgo BIAS [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

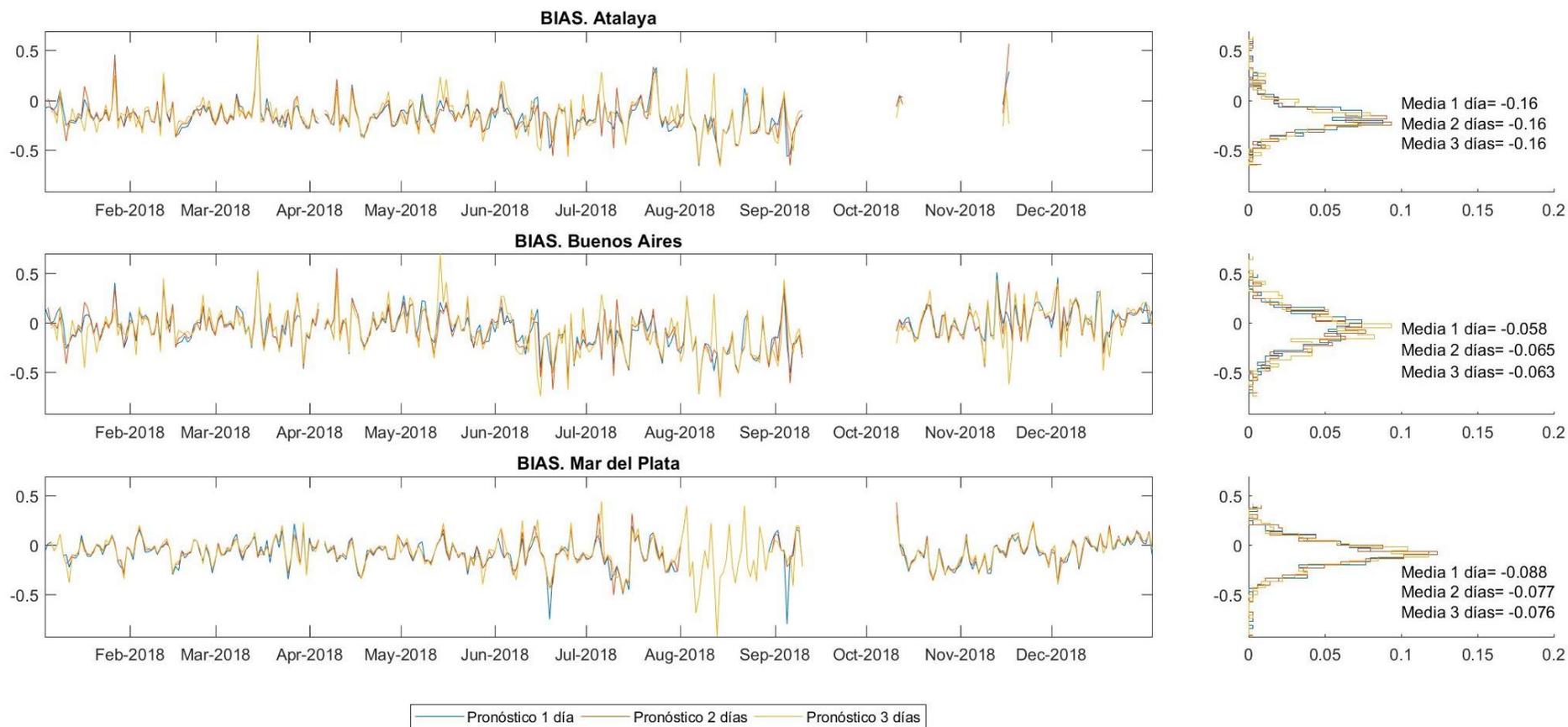


Figura 55 - Diferencia de media, o sesgo BIAS [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

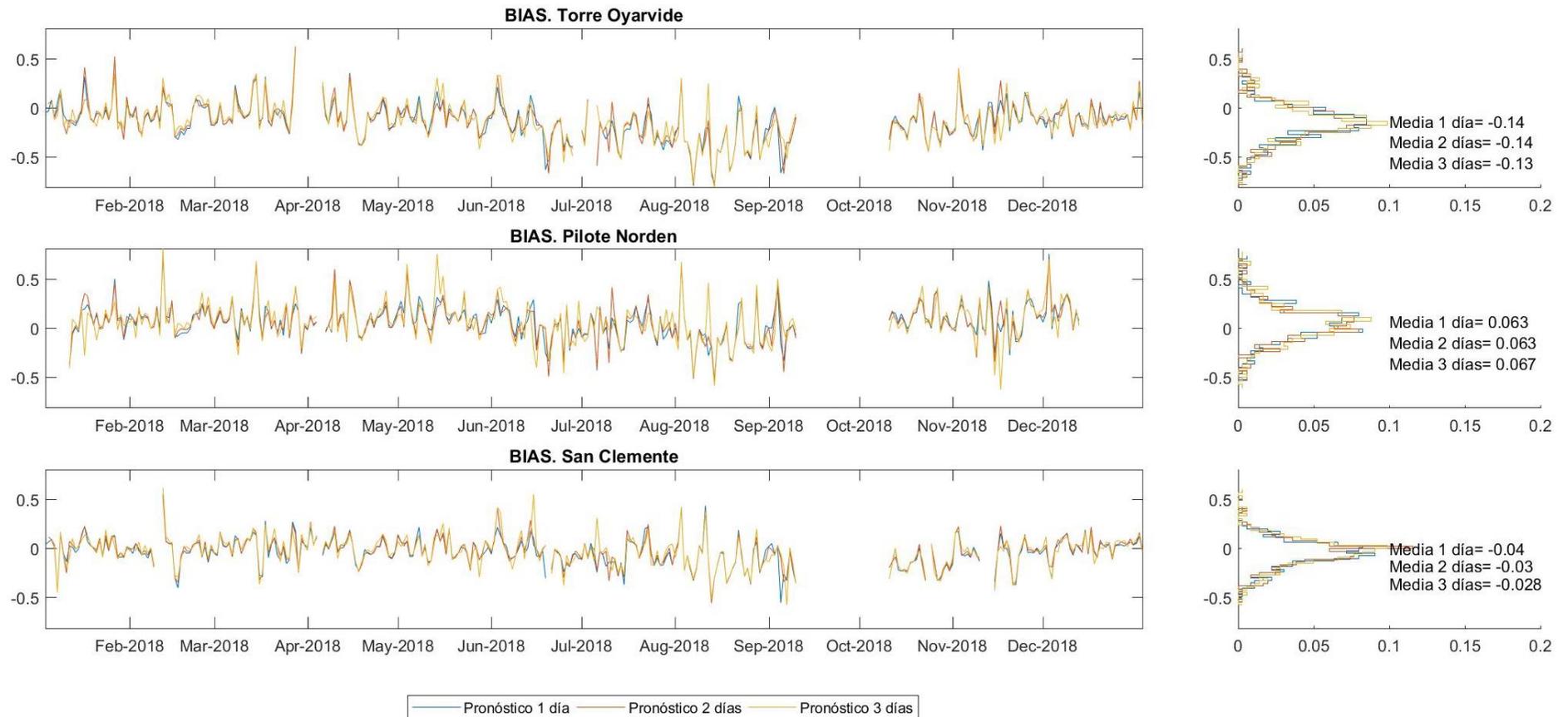


Figura 56 - Diferencia de media, o sesgo BIAS [m] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

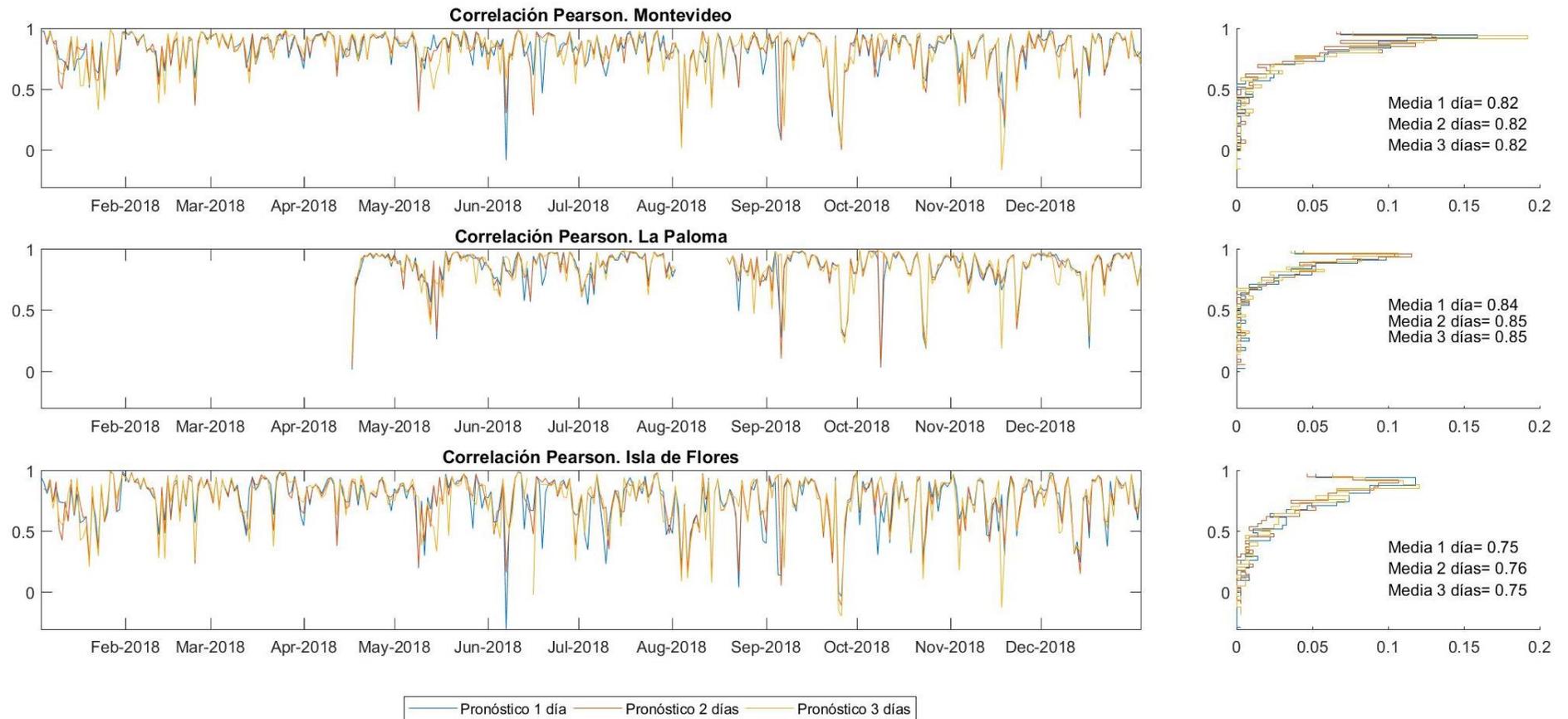


Figura 57 – Correlación Pearson [-] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

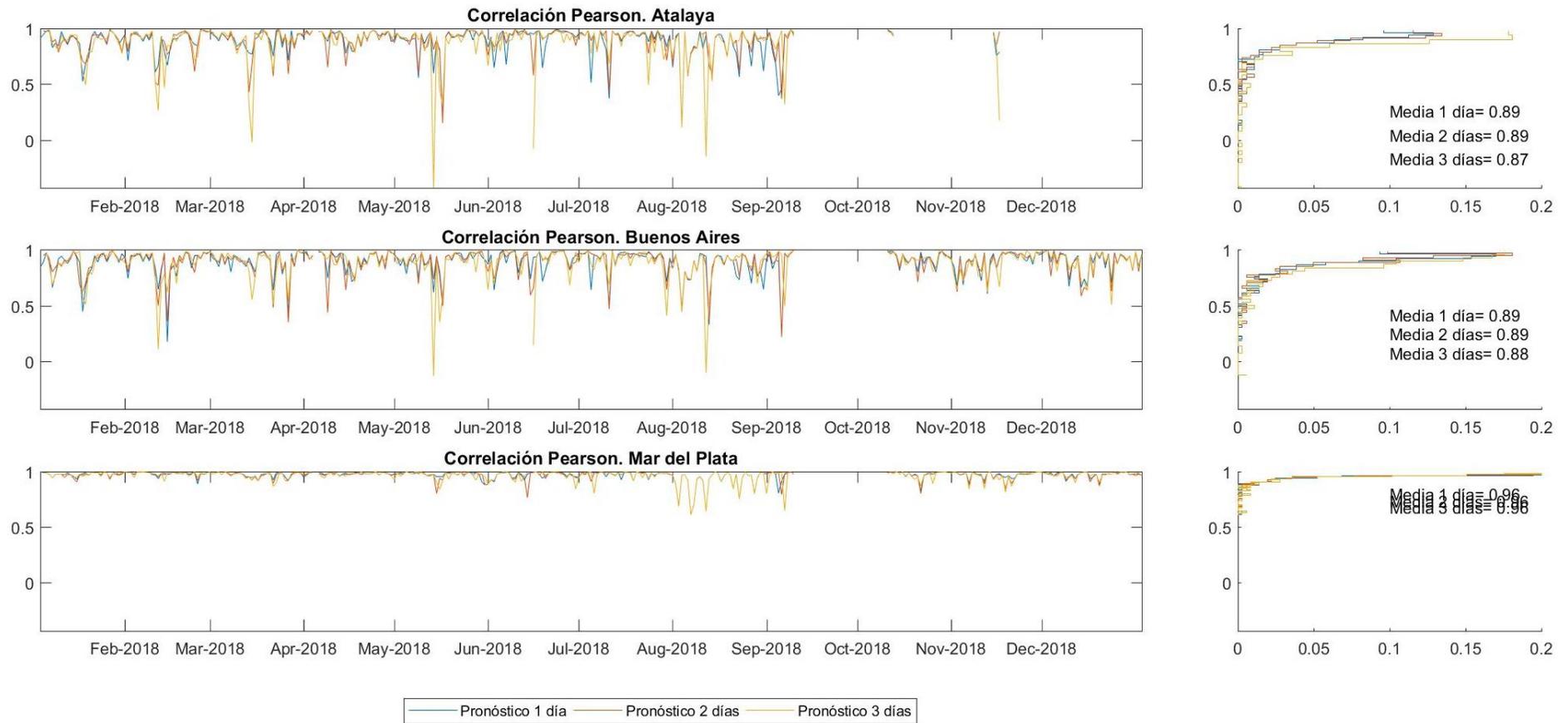


Figura 58 - Correlación Pearson [-] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

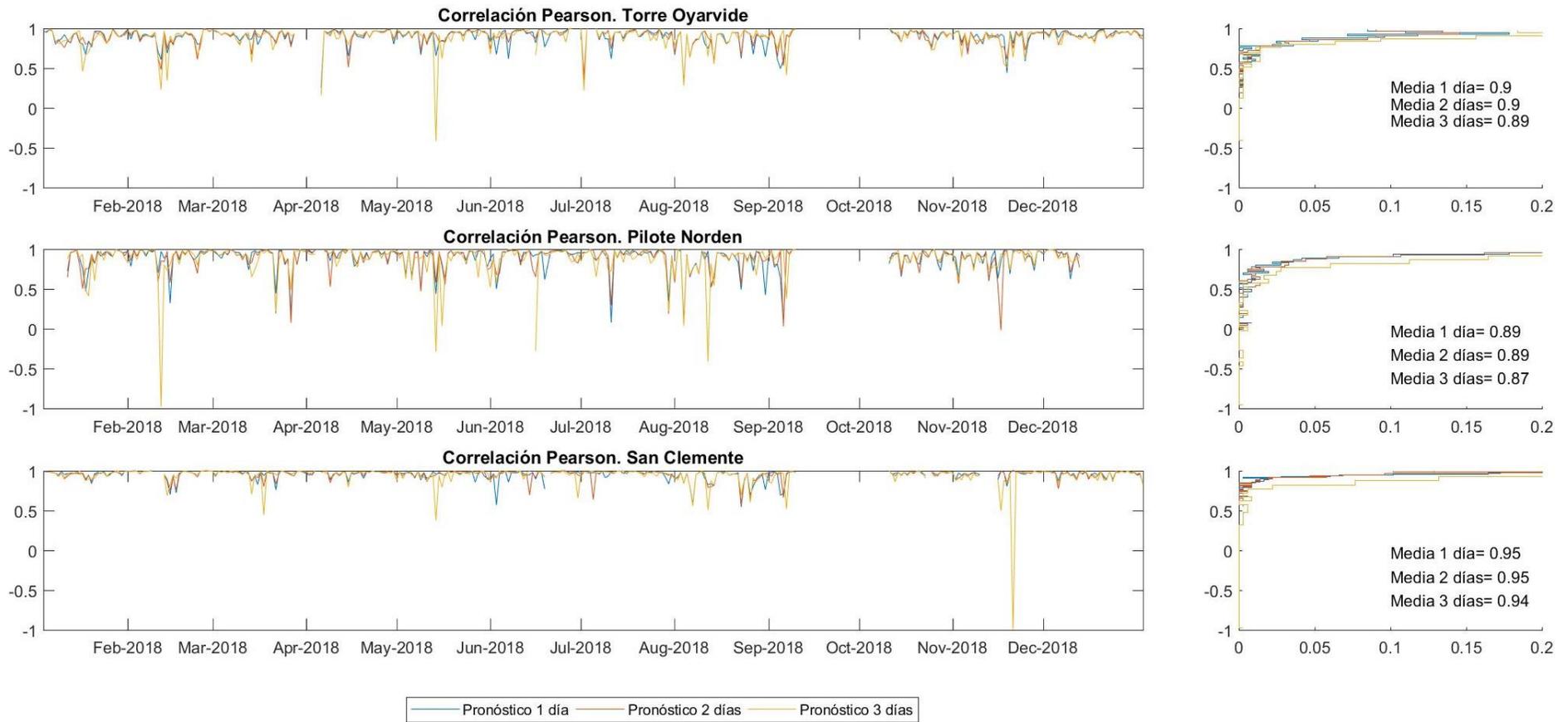


Figura 59 - Correlación Pearson [-] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

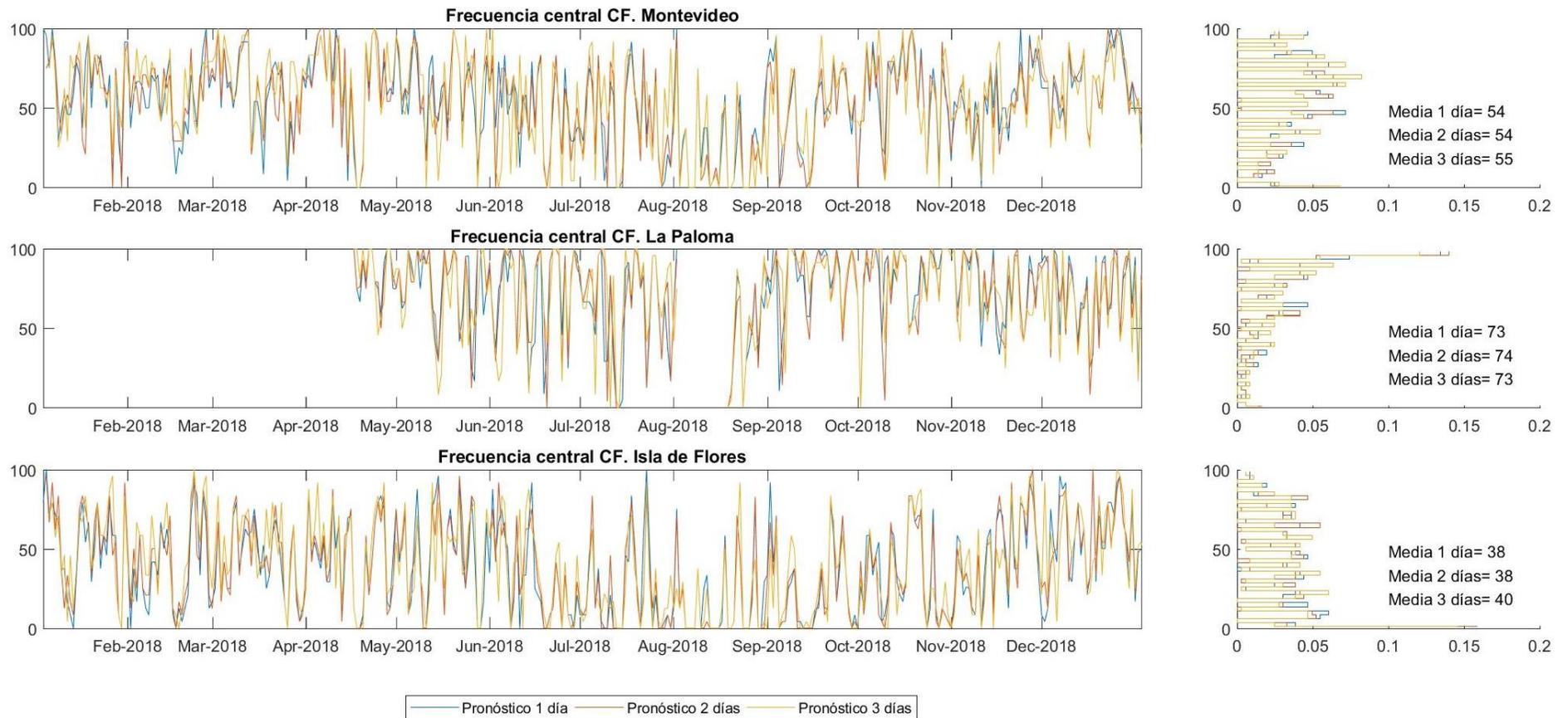


Figura 60 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

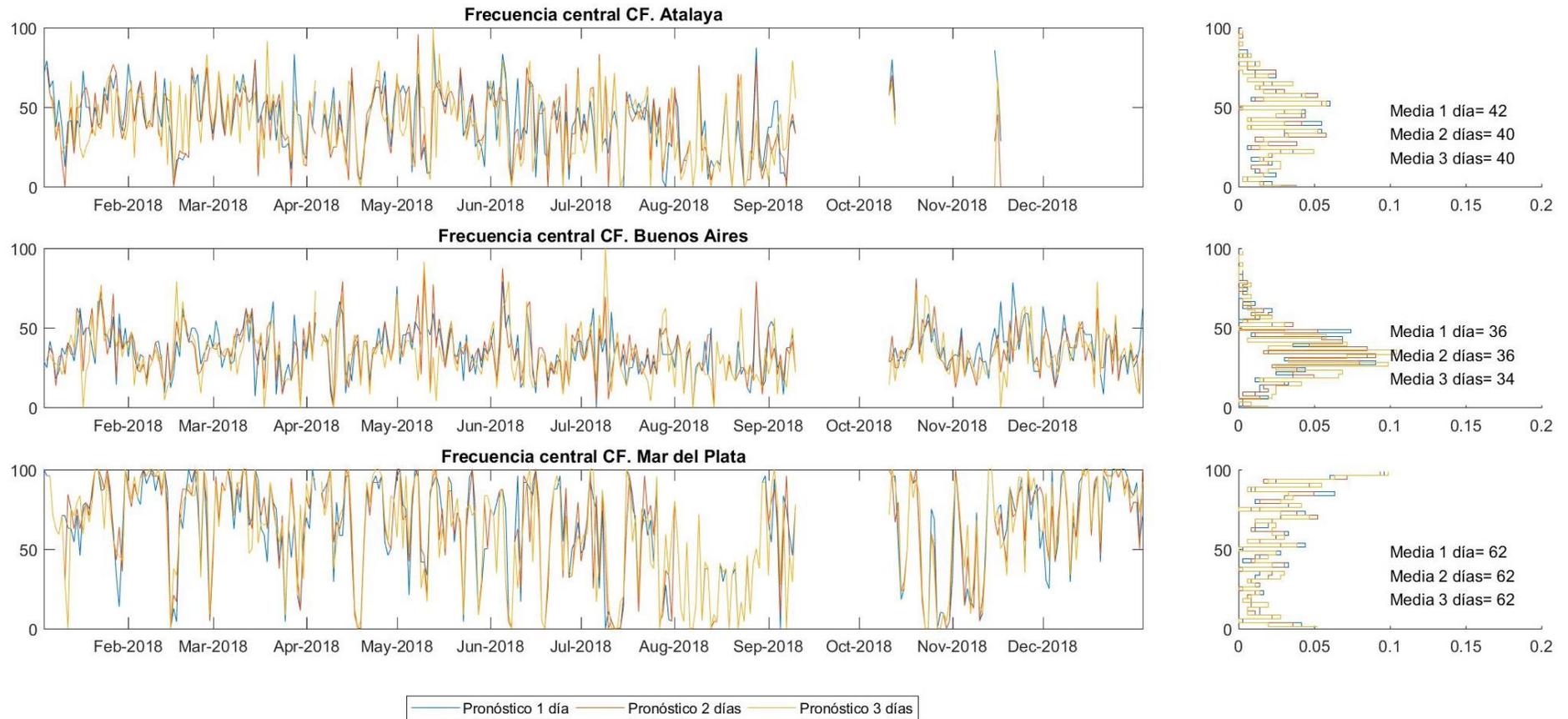


Figura 61 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

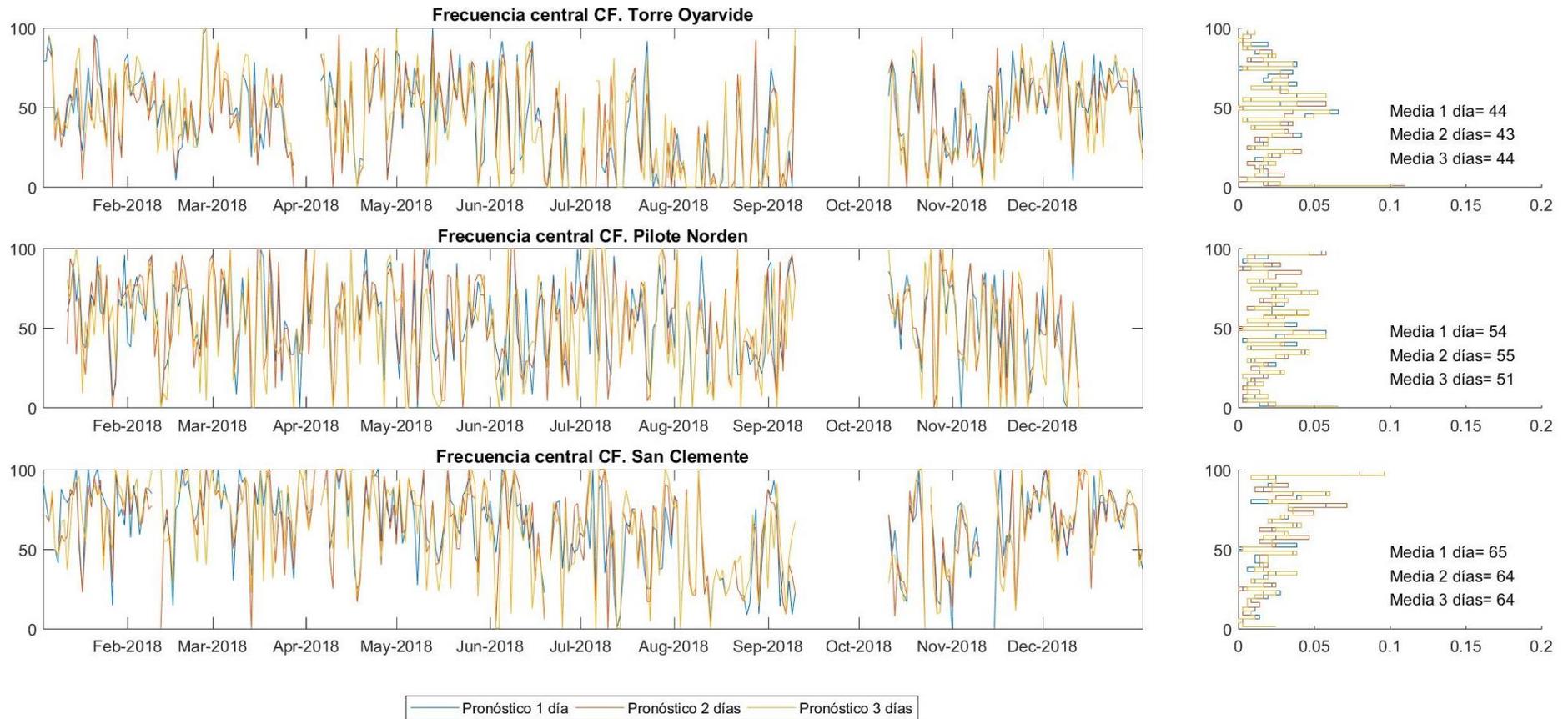


Figura 62 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

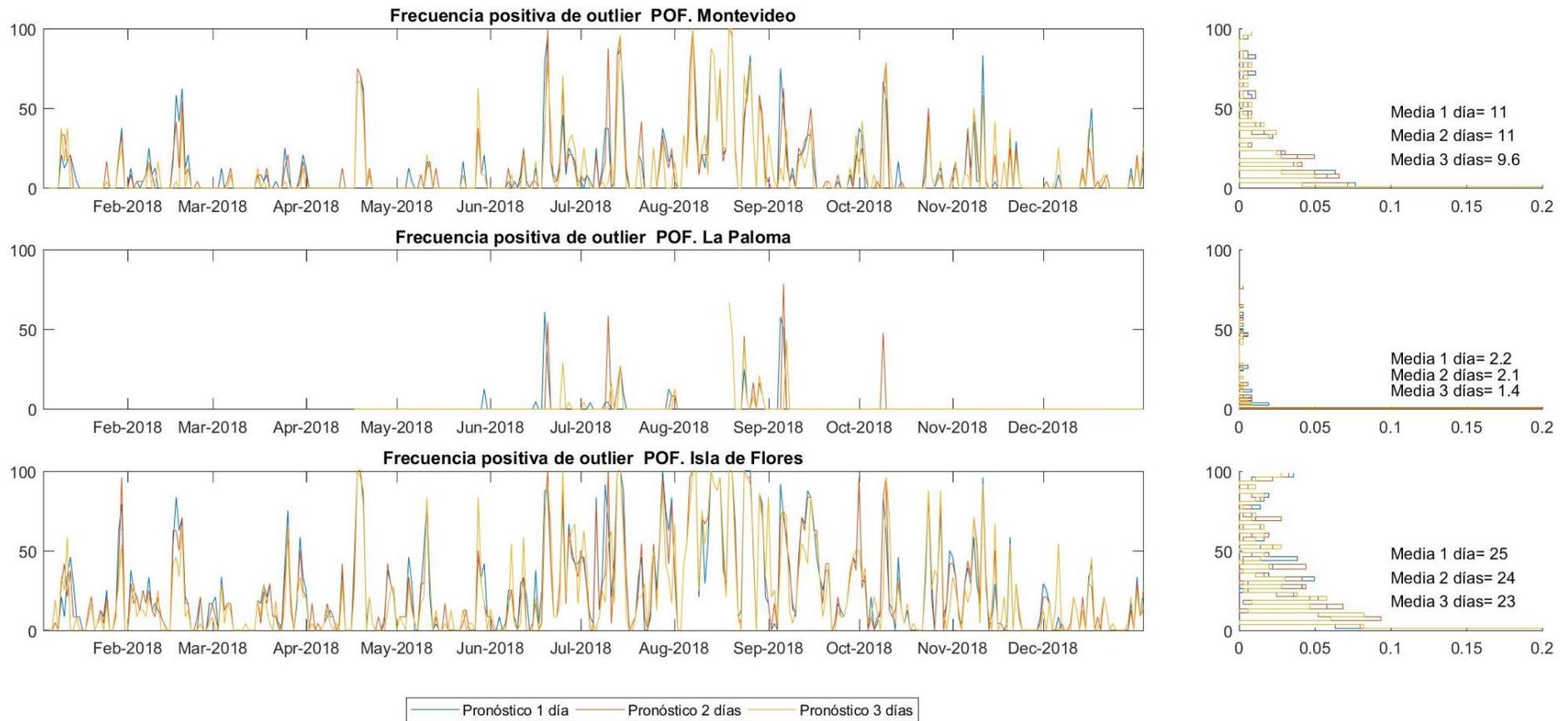


Figura 63 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

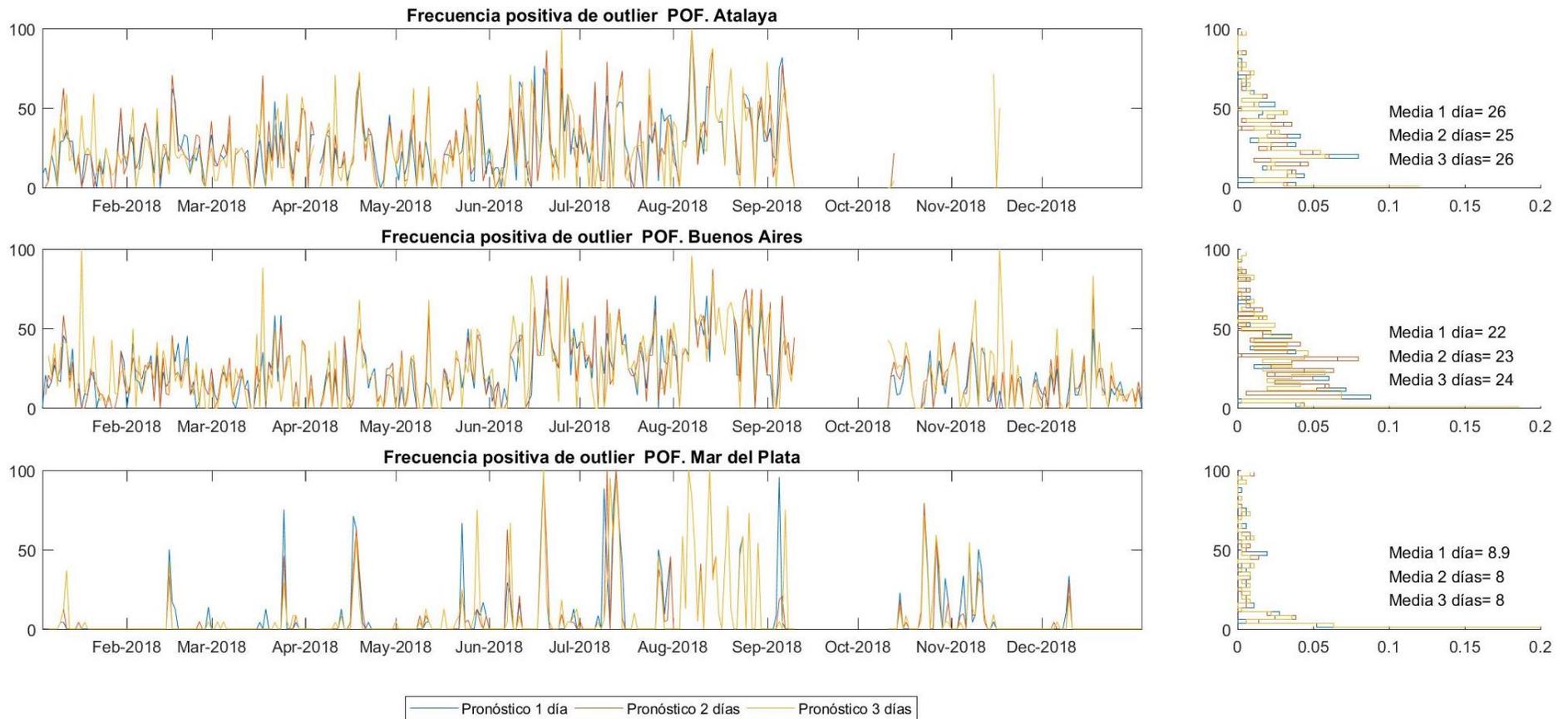


Figura 64 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

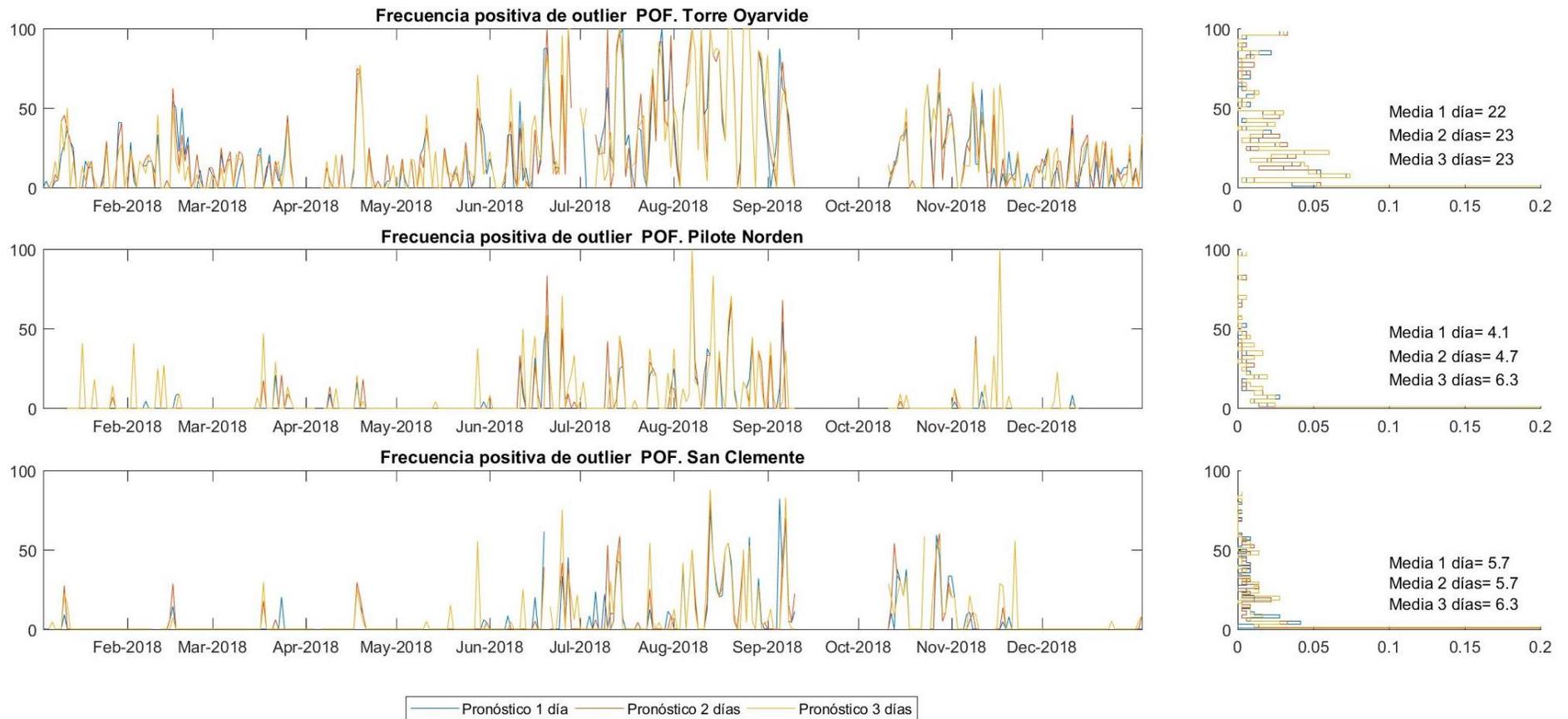


Figura 65 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

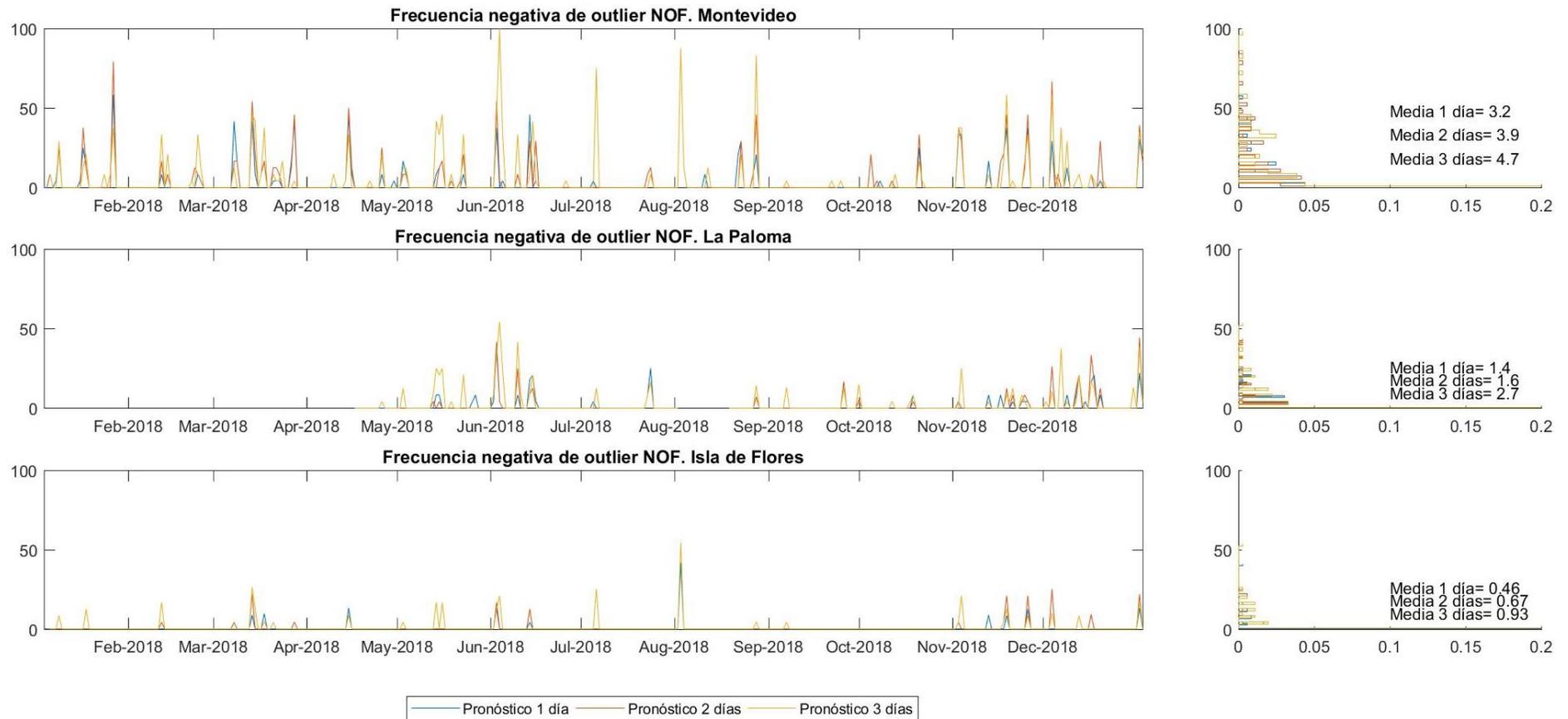


Figura 66 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

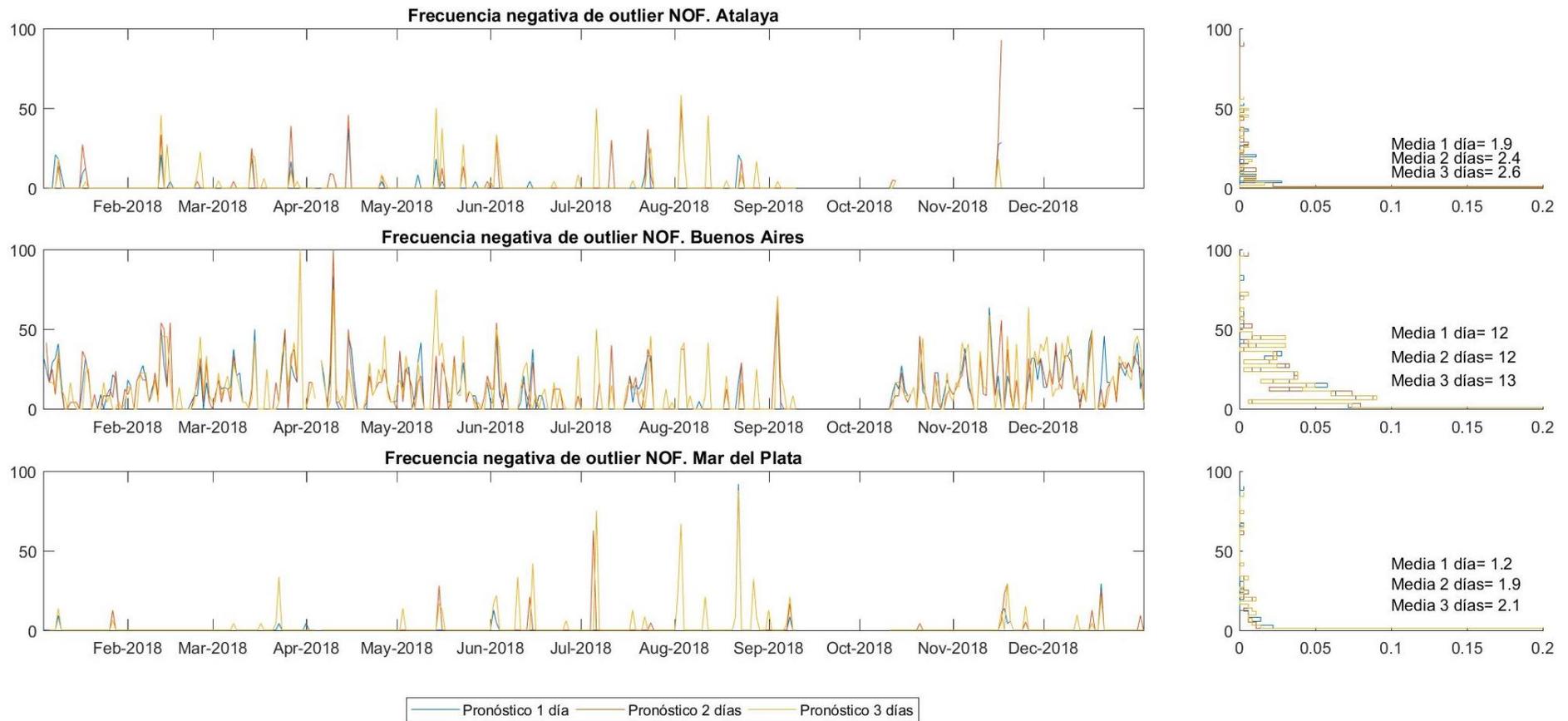


Figura 67 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

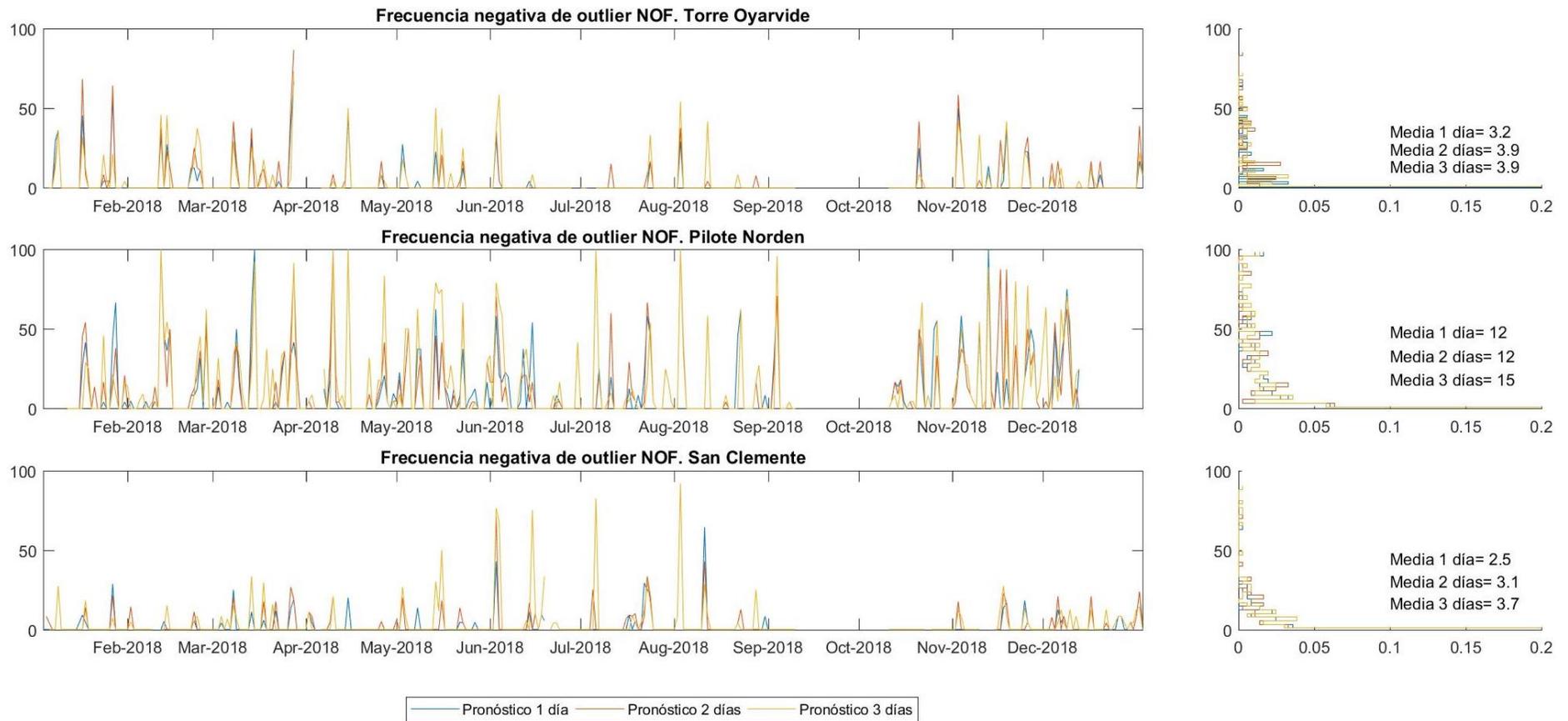


Figura 68 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

IV.3. Análisis

Aquí se analizan los resultados de la comparación de los niveles que arroja el prototipo de pronóstico generado frente a los registros de mareógrafos.

Los resultados que se obtendrían si el modelo fuera capaz de pronosticar de un modo perfecto serían $RMSE = 0$, $BIAS = 0$, $Correlación\ Pearson = 0$, $CF(\zeta) = 1$, $POF = 0$ y $NOF = 0$. Además, en la representación de los diagramas de dispersión todos los puntos se encontrarían alineados a 45° (línea negra representada en los diagramas). En la Tabla 10 y en la Tabla 11, se presenta a modo de resumen la media anual (para el año 2018) de los estadísticos calculados para cada estación considerando las primeras 24 horas de pronóstico (1 día).

Año 2018. Pronóstico 1 día			
	RMSE [m]	BIAS [m]	CORR [-]
Mareógrafos			
Isla de Flores	0.24	-0.20	0.75
Montevideo	0.19	-0.08	0.82
La Paloma	0.13	0.02	0.84
Atalaya	0.26	-0.16	0.89
Buenos Aires	0.30	-0.06	0.89
Mar del Plata	0.16	-0.09	0.96
Torre Oyarvide	0.25	-0.14	0.90
Pilote Norden	0.20	0.06	0.89
San Clemente	0.16	-0.04	0.95

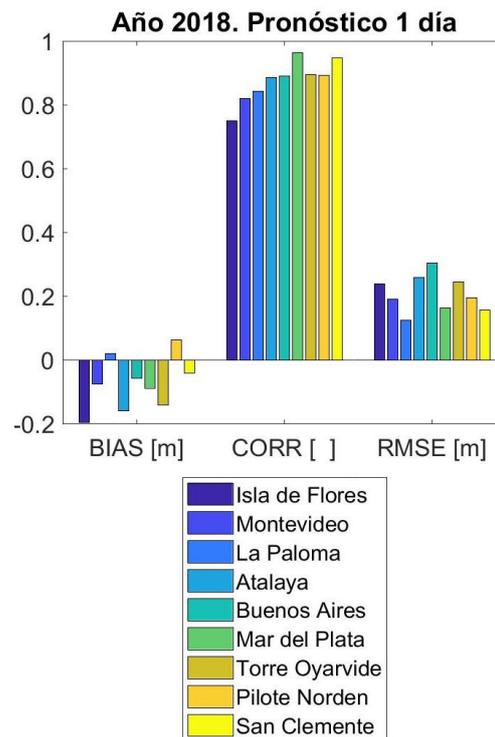


Tabla 10 - BIAS, RMSE, Correlación Pearson para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día). Tabla de valores y gráfico de barras.

Año 2018. Pronóstico 1 día			
	CF	POF	NOF
Mareógrafos			
Isla de Flores	37.8%	25.4%	0.5%
Montevideo	54.1%	11.5%	3.2%
La Paloma	73.3%	2.2%	1.4%
Atalaya	41.9%	25.6%	1.9%
Buenos Aires	36.2%	22.1%	12.5%
Mar del Plata	61.5%	8.9%	1.2%
Torre Oyarvide	43.7%	22.3%	3.2%
Pilote Norden	54.3%	4.1%	11.9%
San Clemente	64.5%	5.7%	2.5%

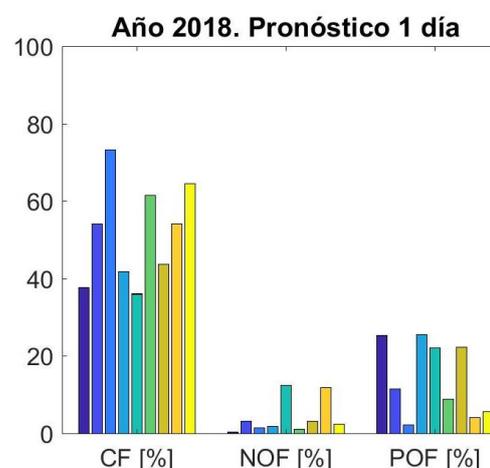


Tabla 11 - Frecuencia Central (CF), Frecuencia Positiva de Outlier (POF), Frecuencia Negativa de Outlier (NOF) para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día). Tabla de valores y gráfico de barras.

Para las 9 estaciones mareográficas se obtienen valores menores a 0.30 m para la media anual del error cuadrático medio **RMSE**, pero, el análisis de la serie anual de 2018 de este estadístico revela una gran variabilidad con jornadas en las que se superan los 0.80 m (por ejemplo, Mar del Plata a mediados de agosto de 2018). Durante la primera mitad del año, este estadístico se mantuvo en un rango de valores similar para todos los meses en las distintas estaciones y no se produjeron picos destacables con la excepción de la estación de Pilote Norden que a mediados de febrero y de nuevo a mediados de abril presenta un pico cercano a los 0.80 m. En el mes de junio se presenta un pico de error en todas las estaciones y durante el mes de agosto y principios de septiembre también se dan varios picos simultáneamente para todas las estaciones. A continuación se tiene un hueco de 1 mes de duración para los datos de las estaciones del SHN. Durante ese mes, desde mediados de septiembre hasta mediados de octubre, el comportamiento de los mareógrafos de la ANP presenta valores similares a la primera mitad del año. En los últimos meses del año se destaca un único pico de error para la estación de Buenos Aires que coincide con un aumento en el error de Pilote Norden.

El aumento en los errores que se da para todas las estaciones durante agosto podría deberse a la falta de datos de HYCOM como condición de contorno (del 27 de julio al 28 de agosto). En la mayor parte de este intervalo tampoco se dispone de datos de forzantes NOAA (se utilizaron datos con menor resolución horaria), pero, este primer análisis señala mayor dependencia a los datos faltantes de HYCOM ya que no se destacan grandes errores desde enero hasta mitad de febrero cuando no se disponía de NOAA (se utiliza una menor resolución horaria) y sí de HYCOM.

En la Figura 69 (panel izquierdo) se presentan sobre el mapa de la zona de estudio los valores de del error cuadrático medio **RMSE**, que se vieron en la Tabla 11. La variabilidad espacial del error cuadrático medio **RMSE** presenta valores más bajos en promedio en la comparación frente a los mareógrafos más oceánicos, La Paloma, Mar del Plata y San Clemente. Los errores más altos se dan en las estaciones de Buenos Aires y Atalaya. La comparación de este estadístico entre estaciones puede no ser una buena representación del comportamiento espacial del sistema de pronóstico, pues durante su cálculo se tiene en cuenta el rango que abarcan los datos en cada estación. Así, en La Paloma, que se obtiene un **RMSE** bajo, el rango en el que se mueven los datos también es bajo y lo contrario ocurre en Buenos Aires.

En relación a la **Correlación Pearson** calculada con los valores de 2018 se obtienen valores promedio por encima de 0.89 para todas las estaciones de SHN. El menor valor, 0.75, se da para la estación de Isla de Flores. Su dispersión espacial se puede ver en la Figura 69 (panel derecho) en la que se destaca una correlación muy cercana a la unidad para las estaciones más oceánicas de Argentina: Mar del Plata y San Clemente. Cuando se analiza la serie anual se puede ver que para estas estaciones (Mar del Plata y San Clemente), se dan valores apartados de la unidad para el período del 27 de julio al 28 de agosto (detectado con ausencia de condiciones de contorno provenientes del modelo global HYCOM). Los otros cuatro mareógrafos de SHN (Atalaya, Buenos Aires, Torre Oyarvide y Pilote Norden) presentan series similares para el año 2018 en relación a los valores diarios de correlación. Así, los histogramas de este estadístico son muy parecidos y suelen coincidir las jornadas con baja correlación para estos cuatro mareógrafos. Las estaciones de Montevideo e Isla de Flores se encuentran ubicadas muy próximas en el espacio, esto se refleja en una serie de correlación similar pero, se aprecia que los valores de Isla de Flores son generalmente más bajos que en Montevideo. La estación de La Paloma no tiene una correlación tan alta como las otras estaciones oceánicas (San Clemente y Mar del Plata) y aparecen algunas jornadas con baja correlación que no son detectadas en otras estaciones.

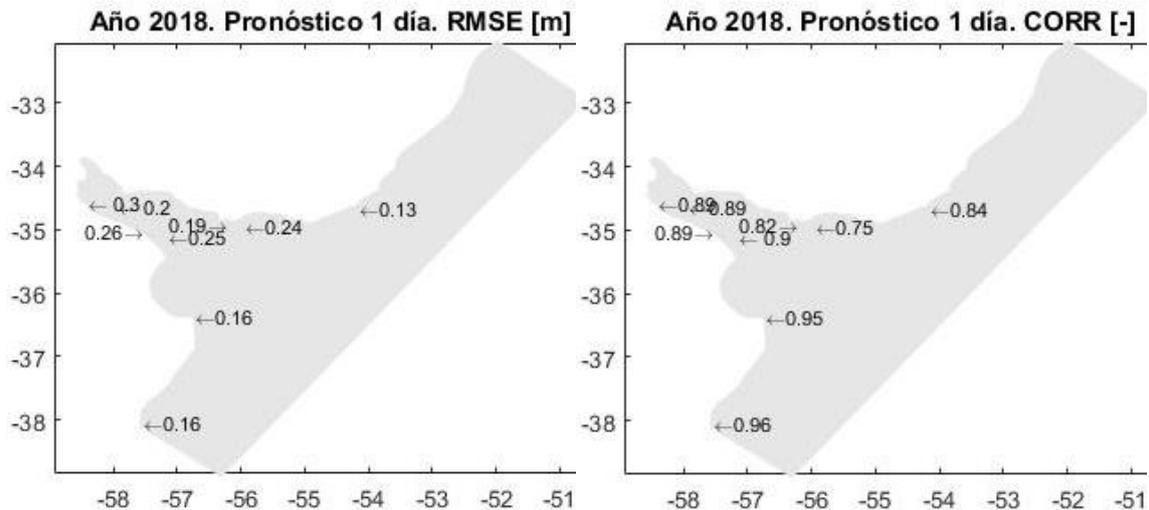


Figura 69 – Variación espacial del RMSE promedio anual y de la Correlación Pearson promedio anual para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día).

Como se mencionó en el apartado II.1.4, no se espera que los errores en las series de niveles sigan una distribución normal, y es por esto que el error cuadrático medio **RMSE**, tiene un uso limitado. Los estadísticos presentados en la Tabla 11, frecuencia central **CF(ζ)**, frecuencia positiva de *outlier* **POF(ζ)** y frecuencia negativa de *outlier* **NOF(ζ)** complementan la información de los estadísticos anteriores.

La diferencia de media, o sesgo **BIAS**, fue calculado de forma diaria del mismo modo que los otros estadísticos y se observa que efectivamente la distribución que sigue su histograma no es normal. Recordando el modo en el que la diferencia de media se calculó, valor observado menos valor modelado, los valores positivos en este estadístico representan una subestimación del modelo y los valores negativos una sobreestimación del modelo. Así, por ejemplo, al analizar el histograma del sesgo en 2018 para la estación de Montevideo se observa que su media es negativa (el modelo sobreestima), y que además, los valores más alejados de la media se encuentran hacia el lado negativo, lo que indica que los errores más grandes son por sobreestimación del nivel.

La media anual del sesgo **BIAS** es una representación de un error sistemático en el nivel de referencia de cada estación. La discusión de las posibles fuentes de origen de una mala estimación en el nivel medio se hará en el siguiente apartado V. Este error no influye en los estadísticos desarrollados en los párrafos anteriores (**RMSE** y **Correlación Pearson**) pero cuando se evalúan **CF(ζ)**, **POF(ζ)** y **NOF(ζ)** cobra gran importancia. En la Figura 70 se ilustra la repercusión de una mala estimación en el nivel medio en la estación de Isla de Flores que presenta un sesgo medio anual para 2018 de -0.20 m. En el panel izquierdo se muestra el diagrama de dispersión de los niveles pronosticados frente a los niveles medidos y en el diagrama de dispersión del panel derecho a los datos se les resta su media anual. Se puede ver que al restar la media anual, panel derecho, ambas series pasan a tener media cero. Se incluyen en ambos diagramas de dispersión bandas verdes que limitan los valores contabilizados para la frecuencia central **CF(0.15m)**: cuando no se resta la media anual (panel izquierdo) se obtiene el valor presentado en la Tabla 11, 37.8%, cuando se resta la media anual (panel derecho) el valor pasa a ser 64.5%. El corrimiento de la media afecta del mismo modo a la frecuencia positiva de *outlier* (en los diagramas se representa con banda roja superior el límite que deja por encima a los valores contabilizados para **POF(0.3m)**) y a la frecuencia negativa de *outlier* (en los diagramas se representa con banda roja inferior el límite que deja por debajo a los valores

contabilizados para **NOF(0.3m)**). El corrimiento de media reduce el valor de la frecuencia positiva de *outlier* desde 25.4% hasta 4.4% y aumenta el valor de frecuencia negativa de *outlier* que tenía un valor muy bajo de 0.5% a 3.9%. Se calcularon los valores de estos estadísticos respecto a la media anual (Anexo 4) y sus valores medios se presentan en la Tabla 12.

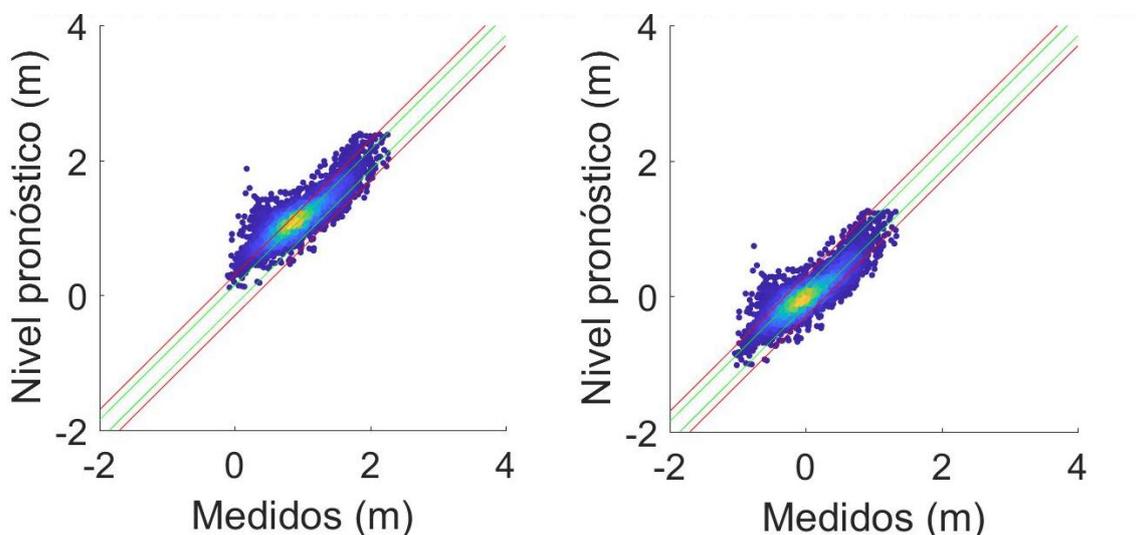


Figura 70 – Diagrama de dispersión para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) para el mareógrafo de Isla de Flores (panel izquierdo). Diagrama de dispersión con los valores respecto a la media anual (panel derecho). Se presentan en verde las bandas que determinan CF(0.15m) y en rojo la banda superior para POF(0.3m) y la banda inferior para NOF(0.3m).

Año 2018. Pronóstico 1 día			
	CF	POF	NOF
Mareógrafos			
Isla de Flores	64.5%	4.4%	3.9%
Montevideo	58.1%	5.9%	6.1%
La Paloma	74.4%	3.3%	0.7%
Atalaya	53.3%	9.9%	5.9%
Buenos Aires	36.9%	17.1%	16.1%
Mar del Plata	69.1%	4.2%	2.7%
Torre Oyarvide	53.2%	8.8%	8.4%
Pilote Norden	58.6%	7.0%	6.7%
San Clemente	65.0%	4.3%	3.3%

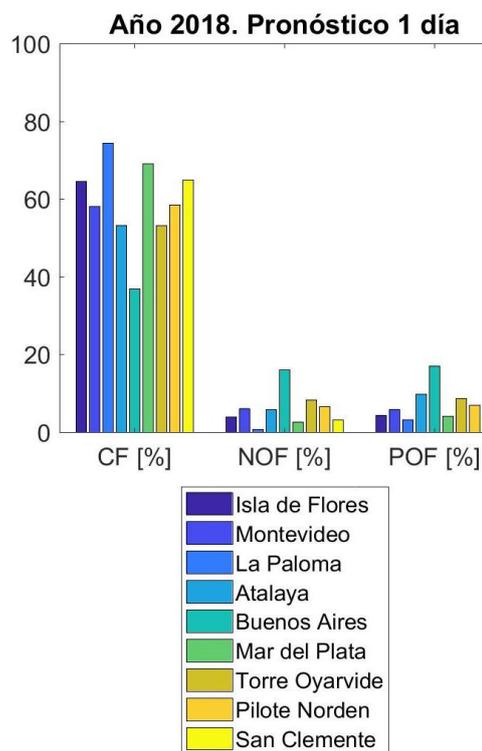


Tabla 12 - Frecuencia Central (CF), Frecuencia Positiva de Outlier (POF), Frecuencia Negativa de Outlier (NOF) para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día). Valores obtenidos respecto a la media anual. Tabla de valores y gráfico de barras.

Las estaciones que tenían un menor sesgo anual medio (en valor absoluto), La Paloma y San Clemente, son las que menos modifican sus valores de **CF(ζ)**, **POF(ζ)** y **NOF(ζ)** al restar la media anual. Los valores altos de frecuencia central coinciden con valores bajos de frecuencia

positiva de *outlier* y frecuencia negativa de *outlier*. Al analizar las series anuales de estos estadísticos se puede ver que las jornadas con valores cercanos al 100% de frecuencia central ocurren con frecuencia para La Paloma, Mar del Plata y menos en Pilote Norden. Las estaciones en las que generalmente se obtienen valores más bajos son Buenos Aires y Atalaya. Otra observación es que la variabilidad temporal de estos estadísticos es muy grande presentándose frecuentemente para todas las estaciones jornadas con muy buenos resultados seguidas de jornadas con muy malos resultados.

Cuando se comparan los datos del modelo para el segundo y tercer día de pronóstico se obtienen valores similares al primer día para todas las estaciones y para todos los estadísticos presentados en su valor medio anual. Sin embargo, al analizar la serie de resultados diarios se aprecian eventos para los que se obtienen peores resultados con una anticipación de 3 días. Esto se puede ver a mediados de mayo (11-05-2018) para Buenos Aires, Atalaya, Torre Oyarvide, Pilote Norden y en menor medida Montevideo, La Paloma e Isla de Flores. Se observa que para este evento se obtiene para el pronóstico a 3 días, en comparación con los correlativos pronósticos a 2 días y a 1 día, un mayor **RMSE**, una menor **Correlación Pearson** y un mayor sesgo **BIAS**. Que, por la forma en la que está calculado este último estadístico, indica que con una anticipación de 3 días el modelo subestima en mayor medida el nivel pronosticado para este evento, en concordancia con una mayor frecuencia negativa de *outlier* **NOF(0.3m)**, y que en los pronósticos generados en los siguientes días esta subestimación se redujo. En la Figura 71 se muestran los niveles pronosticados y medidos para este evento en Buenos Aires y Torre Oyarvide.

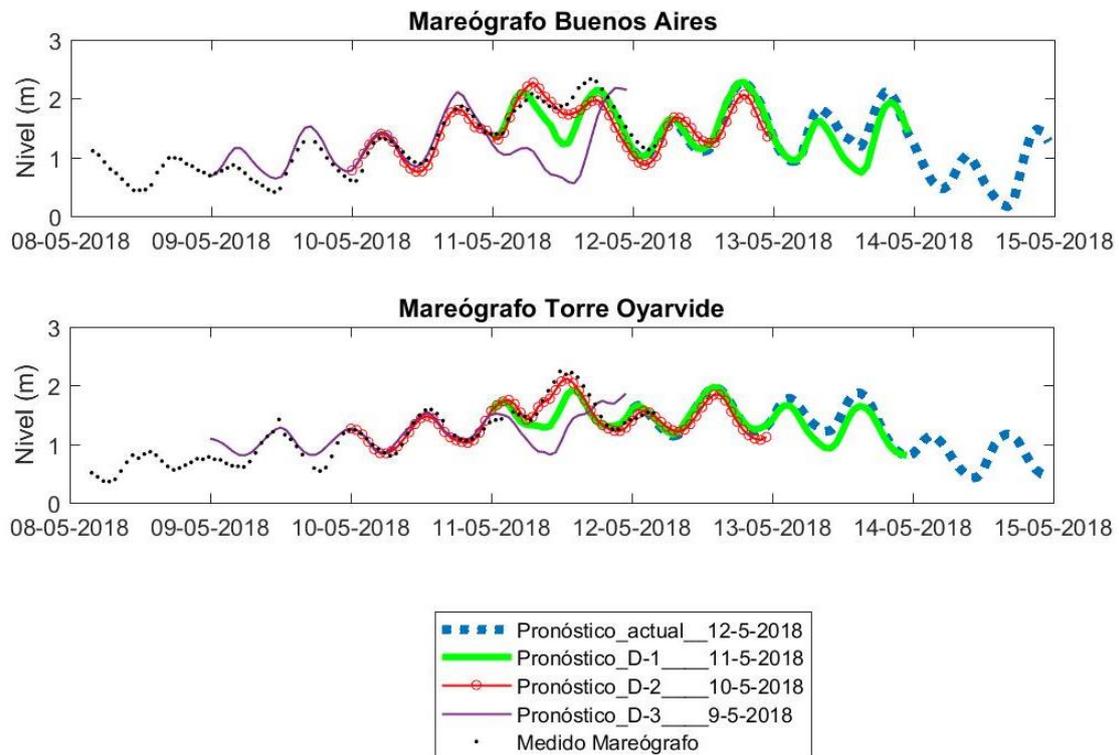


Figura 71 - Salida gráfica del pronóstico. Comparación entre los resultados del modelo y los datos del mareógrafo para Buenos Aires y Torre Oyarvide para el pronóstico del 12-05-2018.

Otro evento en el que se mejora el pronóstico desde los 3 días de anticipación se da a mediados de noviembre (14-11-2018) para las estaciones de Buenos Aires, Pilote Norden y en menor medida Torre Oyarvide. En este caso, con una anticipación de 3 días, el modelo está

sobreestimando: sesgo negativo y mayor frecuencia positiva de *outlier*. En la Figura 72 se muestran los niveles pronosticados y medidos para este evento en Buenos Aires y Torre Oyarvide.

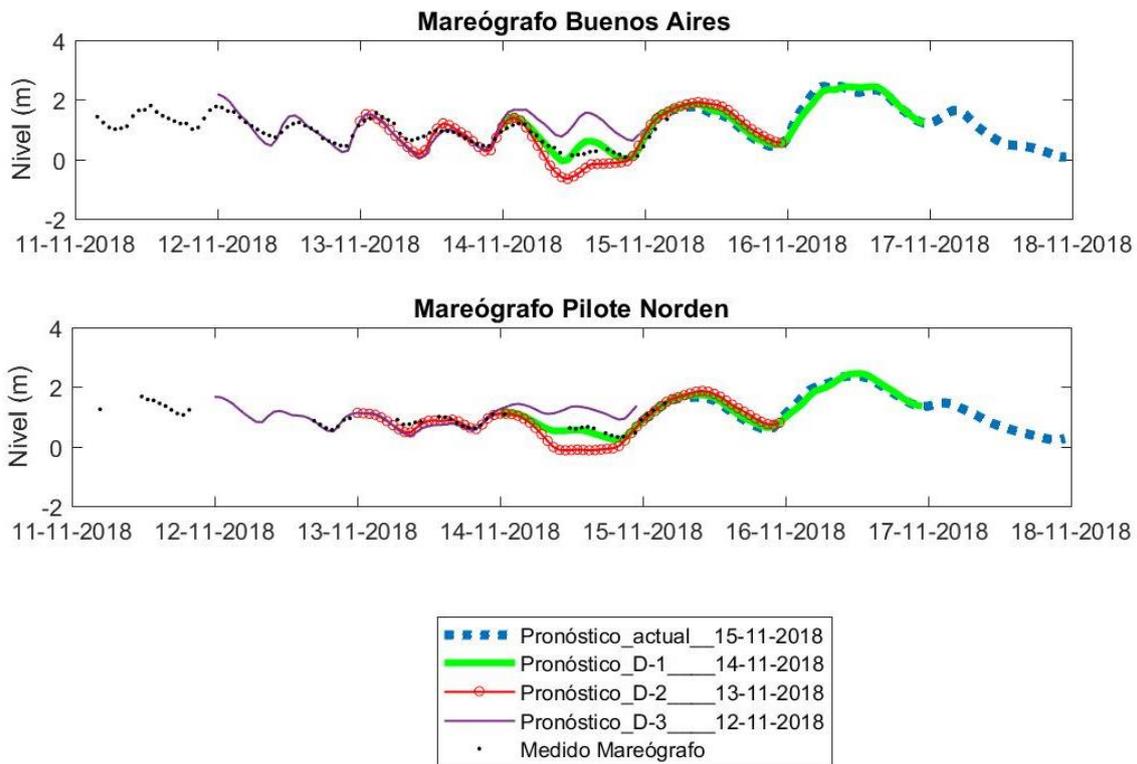


Figura 72 - Salida gráfica del pronóstico. Comparación entre los resultados del modelo y los datos del mareógrafo de Buenos Aires y Pilote Norden para el pronóstico del 15-11-2018.

Los resultados obtenidos durante el segundo año de comparación (año 2019) siguen las mismas tendencias que en el primer año. En primer lugar, esto significa que el filtro de mareógrafos (Anexo 3) que fue generado con los datos del año 2018, sirve para filtrar los datos del 2019. En la Tabla 13 y Tabla 14 se resumen las medias diarias para los estadísticos calculados para 2019. Atendiendo a la frecuencia central, **CF(0.15m)**, se obtiene una mejora con respecto al año anterior para todas las estaciones a excepción de Mar del Plata, para la que los valores son similares.

Año 2019. Pronóstico 1 día			
	RMSE [m]	BIAS [m]	CORR [-]
Mareógrafos			
Isla de Flores	0.21	-0.17	0.76
Montevideo	0.16	-0.04	0.84
La Paloma	0.12	0.09	0.83
Atalaya	0.23	-0.16	0.92
Buenos Aires	0.28	-0.05	0.91
Mar del Plata	0.16	-0.11	0.97
Torre Oyarvide	0.22	-0.13	0.92
Pilote Norden	0.18	0.06	0.92
San Clemente	0.15	-0.02	0.93

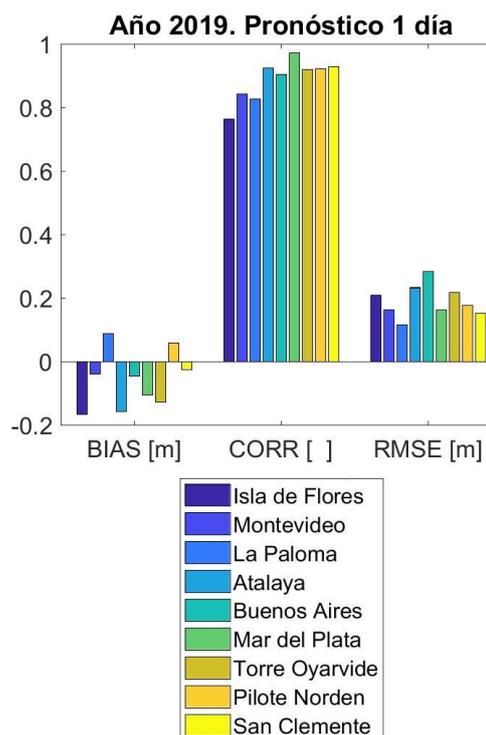


Tabla 13 - BIAS, RMSE, Correlación Pearson para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día). Tabla de valores y gráfico de barras.

Año 2019. Pronóstico 1 día			
	CF	POF	NOF
Mareógrafos			
Isla de Flores	68.3%	1.9%	3.1%
Montevideo	63.5%	3.5%	4.6%
La Paloma	74.7%	1.4%	0.2%
Atalaya	53.9%	3.9%	10.0%
Buenos Aires	37.9%	15.5%	15.7%
Mar del Plata	69.9%	1.8%	2.0%
Torre Oyarvide	57.8%	5.4%	6.3%
Pilote Norden	60.6%	4.1%	5.4%
San Clemente	67.4%	1.9%	3.9%

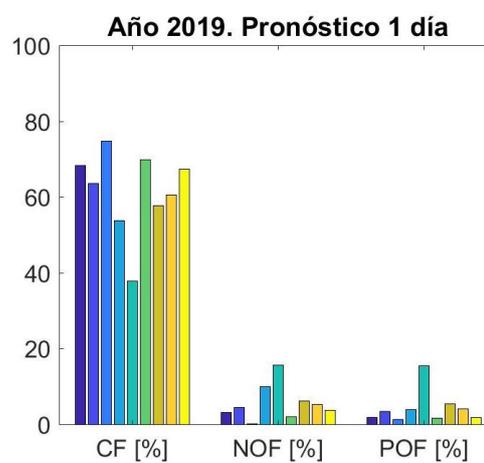


Tabla 14 - Frecuencia Central (CF), Frecuencia Positiva de Outlier (POF), Frecuencia Negativa de Outlier (NOF) para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día). Valores obtenidos respecto a la media anual. Tabla de valores y gráfico de barras.

V. Necesidades del sistema y recomendaciones

En este apartado, en primer lugar, se discute brevemente las necesidades detectadas durante la evaluación del prototipo desarrollado y, seguidamente se incluye una propuesta para la asimilación de datos.

V.1. Necesidades del sistema

La evaluación del sistema pone de manifiesto la importancia de disponer en tiempo real de los datos que se utilizan para constituir las condiciones de borde. En caso de no disponer de los datos de NOAA, el sistema queda desactualizado hasta que no se descarguen los datos de forzantes atmosféricos. En el caso de que los datos faltantes sean los de HYCOM el sistema se ejecuta únicamente con marea astronómica como condición de borde y los pronósticos obtenidos parecen ser de menor calidad. Con la migración del sistema al servidor OBRAS09 (finales de 2019) se pudo ver que las ocasiones de datos faltantes se redujeron drásticamente, aun así, atendiendo a los resultados de la evaluación conviene estudiar alternativas a que el sistema se ejecute únicamente con marea astronómica. Una posibilidad es la utilización de los pronósticos de condiciones de contorno del día anterior (HYCOM+FES2014) y, para poder alcanzar el mismo horizonte temporal de 72 horas, se deberá disponer de los pronósticos del día anterior hasta las 96 horas. Esta posibilidad es útil cuando se tiene únicamente un día de datos faltantes. Si se quiere poder implementar se deberán descargar los pronósticos a partir de la hora 72 que tienen una menor resolución horaria (un dato cada 3 horas) y que por el momento no se han llegado a incluir en la base de datos.

La comparación frente a datos de mareógrafos permite conocer que en algunas estaciones se dan valores elevados de sesgo. Estos errores no podrían ser corregidos a través de una aplicación de asimilación de datos al sistema, ya que, como se dijo anteriormente, durante la asimilación se asume que tanto el modelo como las observaciones son insesgados (*unbiased*). Es por esto que una necesidad del sistema es establecer el origen de estos sesgos para intentar corregirlos. La media anual del sesgo **BIAS** es una representación de un error sistemático en el nivel de referencia de cada estación. Su origen puede ser una mala información sobre la conversión al cero del modelo (cero Wharton en nuestro caso), pues los mareógrafos disponen de una regla local y mediante procedimientos topográficos se genera la conversión entre ceros distantes. La metodología que permite la conversión de los ceros locales a otros ceros de referencia precisa de un buen modelo del geoide en la zona de estudio que con los avances tecnológicos ha generado ajustes en estas conversiones. Así, se puede observar un gran interés en relacionar el cero de Riachuelo (cero del mareógrafo de Buenos Aires) con el cero Wharton, que ha llevado a obtener hasta la cifra de la décima de milímetro, actualizando el valor presentado por Santoro (2011) de 0.1364 m a los 0.1367 m actuales. Sin embargo, la conversión del cero de San Clemente no se puede obtener, al igual que ocurrió en Santoro (2011) y la conversión para Pilote Norden es la misma que se pudo obtener por Santoro (2011) que se recomienda utilizar como una aproximación, por no haberse realizado teniendo en cuenta el geoide si no únicamente nivelación satelital.

Otra fuente que puede dar origen a un sesgo sistemático es que el propio modelo no esté calculando en cero Wharton para todos los puntos. El sistema Telemac-2D (Hervouet, 2007)

permite, entre otras opciones, introducir las coordenadas en sistema Mercator (x,y) o en coordenadas esféricas de latitud y longitud. Cuando se utilizan coordenadas esféricas se asume como origen de la coordenada vertical el elipsoide WGS84, al que se referencia la batimetría. El programa devolverá los datos en este sistema de coordenadas pero para los cálculos internos en TELEMAC todo es convertido a un sistema de coordenadas Mercator. La forma en la que se incorporan los datos de HYCOM como condiciones de contorno es interpolando los valores del pronóstico global a las ubicaciones de los puntos del borde de la malla de cálculo (mesh_21), añadiendo la marea astronómica calculada por FES2014 en estos puntos y sumando 0.91 m. Este valor constante de 0.91 m se generó durante las calibraciones del modelo AStide-RPtide y su incorporación a este modelo de pronóstico se decidió durante las exploraciones descritas en el Anexo 1.

V.2. Introducción de la asimilación de datos en el Sistema de Pronóstico

En este capítulo se esquematizan los pasos a seguir para la introducción de la asimilación de datos en la generación de las condiciones iniciales del Sistema de Pronóstico de Niveles para el Río de la Plata. Se presenta la incorporación mediante 4D-LETKF para asimilar niveles de mareógrafos y, en su caso, del altímetro satelital, a través de este filtro de Kalman por Ensamblados. En el apartado V.2.2 se especifican los pasos para la asimilación de datos de dos mareógrafos.

V.2.1. Asimilación mediante 4D-LETKF

El Filtro Transformado de Kalman por Ensamblados (ETKF por sus siglas en inglés, Ensemble Transform Kalman Filter) que se utilizó para la mejora del modelo SMARA de la costa argentina podría ser utilizado para obtener mejoras en el actual sistema operacional de pronóstico del Río de la Plata y su frente marítimo.

Para aplicar este filtro se deberá calibrar el número P de miembros del ensamble, haciendo pruebas con distintos valores y buscando la convergencia de la solución. Para el tiempo k se tienen los P miembros del ensamble del *background*, $x_{k,p}^f$, con $p = 1$ hasta P . Cada uno de los vectores $x_{k,p}^f$ es un vector columna que contiene concatenados los valores que toman cada una de las variables del estado para ese miembro del ensamble. Así, para una malla de V puntos en los que se tengan en cuenta 4 variables (salinidad, velocidad este y norte y altura de la superficie), el tamaño de los vectores será de $x_{k,p}^f \in \mathbb{R}^N$, con $N = 4 \cdot V$.

Las observaciones disponibles (mareógrafos y altímetro satelital) corresponden a la variable de altura de la superficie. Los mareógrafos miden la altura total en referencia al cero del mareógrafo, mientras que la altura ssh (*sea surface height*) no tiene en cuenta la marea astronómica. A la hora de construir el operador de las observaciones H_k , se debe tener la precaución de que su aplicación a un vector del estado proporcione los valores de las variables observadas. La matriz R_k , será la matriz de covarianza de error de las observaciones, que deberá ser calibrada. Para las medidas de mareógrafos se puede considerar que no hay correlación

entre distintas ubicaciones, lo que resultará en una \mathbf{R}_k diagonal, pero en el caso de las medidas obtenidas por un mismo pase del satélite es recomendable considerar correlación en sus errores.

La sencilla forma mediante concatenación propuesta por Hunt et al. (2007) para la inclusión de observaciones asíncronas hace que, para tener en cuenta observaciones que ocurren en un tiempo que se encuentra entre los instantes de asimilación $k - 1$ y k (llamemos a este tiempo, tiempo j) el operador de las observaciones sea el correspondiente a ese tiempo \mathbf{H}_j y se aplique a los valores que los miembros del ensamble toman en ese instante $\mathbf{x}_{j,p}^f$. Por lo que, el algoritmo generado para la aplicación de este filtro deberá no solo ser capaz de extraer los valores del estado para los tiempos k , sino también para tiempos intermedios j cuando aparezcan observaciones asíncronas a asimilar. La matriz \mathbf{R}_j , es la matriz de covarianza de error de las observaciones del tiempo j .

Aplicando el operador de las observaciones a cada miembro del ensamble se obtienen los vectores $\mathbf{y}_{k,p}^f \in \mathbb{R}^M$ (con M el número de observaciones de ese instante). La matriz $\mathbf{Y}_k^f \in \mathbb{R}^{M \times P}$ contiene en sus columnas las diferencias $(\mathbf{y}_{k,p}^f - \overline{\mathbf{y}}_k^f)$. Cuando se disponga de observaciones asíncronas, $\mathbf{y}_{k,p}^f$ se formará concatenando $\mathbf{H}_k \cdot \mathbf{x}_{k,p}^f$ y $\mathbf{H}_j \cdot \mathbf{x}_{j,p}^f$ (se concatenarán así para todos los j disponibles entre $k - 1$ y k), y de igual modo su media $\overline{\mathbf{y}}_k^f$ y la matriz \mathbf{Y}_k^f . También se concatenarán las matrices \mathbf{R}_k y \mathbf{R}_j lo que resultará en una matriz de covarianza de errores de las observaciones diagonal por bloques, ya que se asumen que los errores de observaciones entre distintos tiempos asíncronos no están relacionados.

El vector de observaciones \mathbf{y}_k incluirá del mismo modo la concatenación de todas las observaciones.

Cuando se aplica localización el esquema de cálculo se repite para cada uno de los puntos de la malla. La localización puede implementarse tanto para observaciones lejanas en el espacio como para observaciones lejanas en el tiempo. Una forma de tener en cuenta la localización es mediante localización suavizada de forma que un factor que decae desde 1 hasta 0 al aumentar la distancia entre la observación y el punto de la malla multiplique a los valores de la matriz \mathbf{R}_k^{-1} .

De manera de no infraestimar la incerteza del modelo en la obtención del *background*, en Hunt et al. (2007) se recomienda la inclusión de la inflación multiplicativa durante la obtención de la matriz de covarianza de los errores del análisis. Aunque la posibilidad de inflación aditiva también se nombra en este artículo y es lo que se implementó en Etala et al. (2015). Se pueden estudiar ambas posibilidades. La forma de incluir inflación multiplicativa es mediante un factor $\rho > 1$ durante el cálculo.

El resumen de implementación del 4D-LETKF es el siguiente. Al aplicar localización los pasos del 3 al 8 se realizan para cada punto de malla.

- 1) Aplicación del operador de las observaciones a cada uno de los miembros del ensamble. Se obtienen los vectores $\mathbf{y}_{k,p}^f$ y la matriz de \mathbf{Y}_k^f .
- 2) Obtener la media del ensamble $\overline{\mathbf{x}}_k^f$, y la matriz de perturbaciones \mathbf{X}_k^f que contiene las diferencias de cada miembro con la media del ensamble $(\mathbf{x}_{k,p}^f - \overline{\mathbf{x}}_k^f)$.

- 3) Selección de las filas de \overline{x}_k^f , y de X_k^f correspondientes al punto de malla dado.
- 4) Calcular $C_k = (Y_k^f)^T \cdot R_k^{-1}$. Aquí es donde se aplicará la localización.
- 5) Calcular la matriz $\tilde{P}_k^a \in \mathbb{R}^{P \times P}$, $\tilde{P}_k^a = [(P-1)I/\rho + C_k \cdot Y_k^f]^{-1}$. Aquí, $\rho > 1$ es el factor de inflación multiplicativa.
- 6) Calcular la matriz $W_k^a \in \mathbb{R}^{P \times P}$, $W_k^a = [(P-1)\tilde{P}_k^a]^{1/2}$
- 7) Calcular el vector $\bar{w}_k^a = \tilde{P}_k^a \cdot C \cdot (y_k - \overline{y}_k^f)$. Añadir \bar{w}_k^a a cada columna de W_k^a , las columnas de la matriz resultante son los vectores de análisis $w_{k,p}^a$
- 8) Multiplicar las filas seleccionadas de X_k^f por cada $w_{k,p}^a$ y añadir las filas seleccionadas de \overline{x}_k^f , para obtener las filas correspondientes al punto de malla dado de los miembros del ensamble analizado, $x_{k,p}^a$
- 9) Tras realizar los pasos del 3 al 8 para cada punto de malla se tendrán todas las filas de cada miembro $x_{k,p}^a$

La inicialización del ensamble puede hacerse de manera arbitraria con un ensamble de estados físicamente plausibles y se espera que el filtro sea capaz de sincronizarse con las observaciones tras varios ciclos de asimilación.

El ensamble analizado son las condiciones iniciales para que el modelo (TELEMAC-2D) evolucione cada miembro hacia el tiempo de asimilación siguiente, y en forma de pronóstico hacia tiempos posteriores. Si se quiere presentar un único pronóstico será por tanto la media del ensamble para cada tiempo.

V.2.2. Prototipo 1. Pasos para la asimilación de 2 mareógrafos

En este apartado se especifica la implementación del prototipo 1 que incluye técnicas de asimilación de datos. Específicamente, medidas de los mareógrafos de Montevideo y Buenos Aires, iniciando en el año 2019, y trabajo mediante un esquema 4D-LETKF en un ciclo de 6 horas (horas 00:00, 06:00, 12:00, 18:00), el modelo numérico de base es el TELEMAC-2D con la malla que utiliza actualmente el sistema de pronóstico (malla mesh_21). Se hace notar que en la redacción de este apartado, para los valores que deben ser calibrados, se especifica un valor inicial con criterios básicos para esta implementación primaria y exploratoria.

Los datos facilitados por la Administración Nacional de Puertos (ANP), Uruguay para el mareógrafo de Montevideo tienen una frecuencia de un dato cada 5 minutos. Se propone asimilar los datos medidos cada 1 hora (horas 00:00, 01:00, 02:00, 03:00, 04:00, 05:00, 06:00, 07:00, 08:00, 09:00, etc.). En un ciclo de asimilación cada 6 horas se tendrán 6 medidas a asimilar, 5 de ellas serán medidas asíncronas y una de ellas estará sincronizada con el tiempo de asimilación. Así, para la asimilación de las 00:00, se asimilarán como medidas asíncronas los registros de las 19:00, 20:00, 21:00, 22:00, 23:00 del día anterior, y como medida síncrona la registrada a las 00:00.

Los datos obtenidos del Servicio de Hidrografía Naval (SHN), Argentina para el mareógrafo de Buenos Aires tienen una frecuencia de un dato cada 1 hora publicados a la hh:45. Se propone

asimilar todos los datos disponibles que serán siempre asíncronos al tiempo de asimilación. Así, para la asimilación de las 00:00, se asimilarán como medidas asíncronas los registros de las 18:45, 19:45, 20:45, 21:45, 22:45, 23:45 del día anterior.

Por lo tanto, en cada ciclo de asimilación se podrán asimilar hasta 12 medidas dependiendo de la disponibilidad de los datos. Se recomienda previo a la asimilación utilizar el mismo filtro para eliminar datos erróneos que utiliza el sistema de pronóstico actual (descrito en el Anexo 3) de manera de mitigar el efecto de datos erróneos.

La matriz de covarianza de los errores de las observaciones $\mathbf{R}_k \in \mathbb{R}^{M \times M}$ con $M = 12$ cuando no haya datos medidos faltantes. Una aproximación usual es que se trate de una matriz diagonal. Para este caso se podría tener un valor diferente para los elementos correspondientes a mareógrafos distintos, pero como primera propuesta se considera el mismo valor, de modo que $\mathbf{R}_k = r \cdot \mathbf{I}$, con $r = 0.004m^2$. Esta estrategia representa una desviación estándar de 2cm.

La cantidad de miembros del ensamble debe ser calibrada. Las primeras pruebas pueden empezar con un ensamble de $P = 10$ miembros. La inicialización del ensamble se hace con estados físicamente plausibles que corresponden a estados cualesquiera de los disponibles del sistema de pronóstico actual, lo que llevará a que la matriz de covarianza de los errores del *background* sea inicialmente muy grande y tras varios ciclos de asimilación se espera que el filtro sea capaz de sincronizarse con las observaciones. A la hora de calibrar el número de miembros del ensamble P esto debe ser tenido en cuenta, y buscar la convergencia al aumentar P tras varios ciclos de asimilación. Una posibilidad para la inicialización de los miembros $\mathbf{x}_{k=0,p=1:10}^f$ es escoger los estados correspondientes a las condiciones iniciales (hora 00:00) de los pronósticos de los días 01/02/2018, 01/03/2018, 01/04/2018, 01/05/2018, 01/06/2018, 01/07/2018, 01/08/2018, 01/09/2018, 01/10/2018, 01/11/2018. Estas condiciones iniciales coinciden con el pronóstico a 24 horas del día anterior.

Debido a que se va a realizar una asimilación en 4D, para la inicialización del ensamble se deberán tener también los valores del día anterior que los estados presentaban en los puntos de la malla correspondientes a los mareógrafos de Montevideo y de Buenos Aires. Estos valores se extraen de la evolución del pronóstico a 24 horas del día anterior y corresponderán a las horas 19:00, 20:00, 21:00, 22:00, 23:00 para Montevideo y 18:45, 19:45, 20:45, 21:45, 22:45, 23:45 para Buenos Aires. Como ejemplo, para el primer miembro que se extrajo de las condiciones iniciales del pronóstico del 01/02/2018, se deberán tener los valores en los mareógrafos del pronóstico del 31/01/2018 para las horas que se han detallado.

Este ensamble del tiempo inicial $\mathbf{x}_{k=0,p=1:10}^f$, se toma como el ensamble del *background* para el 01/01/2019 a las 00:00 que es cuando se ha determinado que se inicia el ciclo de asimilación.

Con el ensamble inicial se puede entrar en el ciclo de asimilación:

- 1) Aplicación del operador de las observaciones a cada uno de los miembros del ensamble. Se obtienen los vectores $\mathbf{y}_{k,p}^f$ y la matriz de \mathbf{Y}_k^f cuyas columnas son las diferencias de cada vector con la media, es decir $(\mathbf{y}_{k,p}^f - \overline{\mathbf{y}_k^f})$.

En el caso de que no haya datos medidos faltantes, el operador de las observaciones aplicado a cada uno de los miembros del ensamble producirá vectores $\mathbf{y}_{k,p=1:10}^f \in \mathbb{R}^M$, con $M = 12$ si no hay datos faltantes, la cantidad de observaciones (6 correspondientes al mareógrafo de Montevideo y 6 correspondientes al mareógrafo de Buenos Aires).

Los elementos de los \mathbf{P} vectores serán:

$$y_{k,p=1:10}^f = \begin{pmatrix} \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{5horas} \\ \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{4horas} \\ \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{3horas} \\ \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{2horas} \\ \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{1hora} \\ \text{nivel en MVD para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{5horas} - \mathbf{15min} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{4horas} - \mathbf{15min} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{3horas} - \mathbf{15min} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{2horas} - \mathbf{15min} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{1hora} - \mathbf{15min} \\ \text{nivel en BAs para el miembro } \mathbf{p} \text{ en el tiempo } \mathbf{k} - \mathbf{0horas} - \mathbf{15min} \end{pmatrix}$$

- 2) Obtener la media del ensamble $\overline{x_k^f}$, y la matriz de perturbaciones X_k^f que contiene las diferencias de cada miembro con la media del ensamble $(x_{k,p}^f - \overline{x_k^f})$.

La cantidad de puntos de la malla mesh_21 es de $V = 37,534$ y se tienen en cuenta 4 variables (salinidad, velocidad este y norte y altura de la superficie), el tamaño de los vectores será de $x_{k,p=1:10}^f \in \mathbb{R}^N$, con $N = 4 \cdot V = 150,136$.

Los siguientes pasos, del 3 al 8, se deben hacer para cada punto de la malla.

- 3) Selección de las filas de $\overline{x_k^f}$, y de X_k^f correspondientes al punto de malla dado.

Este paso consistirá en quedarse con los 4 elementos que corresponden a las variables del punto de malla de $\overline{x_k^f}$, y de las 4 filas de la matriz X_k^f .

- 4) Calcular $C_k = (Y_k^f)^T \cdot R_k^{-1}$. Aquí es donde se aplicará la localización.

Si existieran datos faltantes se deberá actualizar el tamaño de la matriz R_k . Cuando no haya datos faltantes $R_k \in \mathbb{R}^{M \times M}$ con $M = 12$ y en el caso de aplicar localización los elementos de su matriz inversa irán multiplicados por el factor de peso de localización correspondiente al punto de malla.

Este peso puede calcularse como se hace en Etala et al. (2015):

$$\omega(dist) = e^{-dist^2/2L^2},$$

donde $dist$ es la distancia del punto de malla a Montevideo para la mitad superior de R_k^{-1} y la distancia del punto de malla a Buenos Aires para la mitad inferior.

Además, para tener en cuenta la distancia temporal, se puede aplicar también el peso en función de esta distancia multiplicando a los elementos de R_k^{-1} ,

$$\omega(t) = e^{-t^2/2T^2},$$

La primera fila de \mathbf{R}_k^{-1} irá multiplicada por el valor correspondiente a $t = 5horas$, etc.

Los valores de las constantes L y T deberán ser calibrados. Para una primera implementación podría no tenerse en cuenta la localización.

- 5) Calcular la matriz $\tilde{\mathbf{P}}_k^a \in \mathbb{R}^{P \times P}$, $\tilde{\mathbf{P}}_k^a = \left[(P-1)\mathbf{I}/\rho + \mathbf{C}_k \cdot \mathbf{Y}_k^f \right]^{-1}$. Aquí, $\rho > 1$ es el factor de inflación multiplicativa.

Como factor de inflación multiplicativa inicial puede tomarse $\rho = 1.15$. Este valor se calibrará con el criterio de que los miembros del ensamble del *background* mantengan cierta variabilidad correspondiente al error del modelo.

No se presentan especificaciones para la implementación de los siguientes pasos por tratarse de operaciones con las matrices y vectores ya generados en los apartados anteriores.

- 6) Calcular la matriz $\mathbf{W}_k^a \in \mathbb{R}^{P \times P}$, $\mathbf{W}_k^a = \left[(P-1)\tilde{\mathbf{P}}_k^a \right]^{1/2}$
- 7) Calcular el vector $\bar{\mathbf{w}}_k^a = \tilde{\mathbf{P}}_k^a \cdot \mathbf{C} \cdot (\mathbf{y}_k - \bar{\mathbf{y}}_k^f)$. Añadir $\bar{\mathbf{w}}_k^a$ a cada columna de \mathbf{W}_k^a , las columnas de la matriz resultante son los vectores de análisis $\mathbf{w}_{k,p}^a$
- 8) Multiplicar las filas seleccionadas de \mathbf{X}_k^f por cada $\mathbf{w}_{k,p}^a$ y añadir las filas seleccionadas de $\bar{\mathbf{x}}_k^f$, para obtener las filas correspondientes al punto de malla dado de los miembros del ensamble analizado, $\mathbf{x}_{k,p}^a$
- 9) Tras realizar los pasos del 3 al 8 para cada punto de malla se tendrán todas las filas de cada miembro $\mathbf{x}_{k,p}^a$

Una vez obtenidos los miembros del ensamble analizado $\mathbf{x}_{k,p}^a$ se tienen las condiciones iniciales para generar el pronóstico de cada uno de los miembros. A través del modelo TELEMAT-2D con las condiciones iniciales que proporciona el ensamble analizado se obtiene la evolución de los miembros para las próximas 6 horas. Como forzantes se tienen los pronósticos del GFS-NOAA de las 00, 06, 12 y 18 horas pero las condiciones de contorno para el borde oceánico que se obtienen de los pronósticos del RTOFS-HYCOM son actualizadas únicamente a las 00 horas, por lo que serán esos pronósticos los que se utilicen para 4 ciclos de asimilación.

VI. Conclusiones y líneas de trabajo futuro

A continuación se resumen las principales conclusiones arribadas durante el desarrollo de la tesis, se listan las publicaciones asociadas al trabajo y se identifican las líneas de trabajo futuro que permiten extender y consolidar el presente esfuerzo.

VI.1. Conclusiones

Durante el desarrollo de esta tesis se ha profundizado en el estudio detallado de las componentes necesarias para la generación de un sistema de pronóstico operacional: sistemas de observación, modelos numéricos, asimilación de datos, evaluación y difusión del pronóstico. La revisión de la bibliografía y la exploración de las distintas webs en las que se presentan los productos de observaciones y de pronósticos de las variables oceánicas han servido de base para la toma de decisiones con respecto a las componentes apropiadas para la generación del sistema de pronóstico de niveles para el Río de la Plata y su Frente Marítimo.

El prototipo generado de sistema operacional logra cumplir con el objetivo general que se había marcado para este trabajo y permite disponer de una primera herramienta para el dominio. El prototipo junto con la revisión del estado actual de los sistemas de pronóstico operacionales presentada en este documento constituyen las bases para hacer frente a nuevos trabajos en la línea de la oceanografía operacional cumpliendo, por tanto, con los objetivos específicos que motivaron este proyecto y abriendo nuevos caminos de desafíos para el grupo de trabajo.

La herramienta desarrollada tiene las siguientes características: se utilizó el modelo numérico en elementos finitos TELEMAC-2D implementado y calibrado para la zona con malla no-estructurada de resolución variable desde 3 km hasta 100 m en la zona interior. Las condiciones de borde oceánicas son los pronósticos del modelo global RTOFS-HYCOM (NOAA/NCEP de 1/12° de resolución) combinados con el atlas de marea astronómica FES2014. Como forzantes se imponen vientos y presión atmosférica del pronóstico GFS de la NOAA de 1/4° de resolución. El aporte de caudal fluvial de este primer prototipo se considera constante e igual a la media anual distribuida en 3 secciones. Para la evaluación de la calidad del modelo se utilizan datos de niveles disponibles en tiempo real. Con el prototipo desarrollado se puede saber cómo serán los niveles en nuestras aguas para un horizonte temporal de 3 días, siendo actualizadas las condiciones cada 24 horas.

La generación de este sistema de pronóstico que toma como condiciones de contorno oceánicas los pronósticos de un modelo global supone no sólo el subescalado que aumenta la resolución en la zona regional costera, sino también la inclusión en el dominio de zonas que el modelo global no consideraba cubiertas por agua. Al utilizar el modelo global HYCOM, este aumento de dominio llega a suponer prácticamente toda la zona interior del Río de la Plata. Además, la adición de la marea astronómica a la condición de contorno oceánica permite que los resultados del modelo sean niveles de marea total y, por tanto, directamente comparables frente a la marea registrada por los mareógrafos.

Se destaca además que el prototipo de sistema de pronóstico generado constituye en la actualidad el primer pronóstico de niveles que funciona de manera operacional para toda el área del Río de la Plata. De los antecedentes que contienen al Río de la Plata en su dominio, el reciente

estudio presentado por Dinápoli et al. (2020) incluye los lineamientos para una ejecución en tiempo real que sería capaz de proporcionar pronósticos de forma operacional pero que por el momento no se está realizando. En la propuesta de Etala et al. (2015), se ensayó el efecto que tendría la asimilación de datos de nivel (mareógrafos y altímetro satelital) para toda la costa argentina y se incluye el Río de la Plata. La incorporación de los datos medidos a través de la asimilación de datos es una de las bases de la oceanografía operacional y en Etala et al. (2015) el sistema sobre el que se ejecutaron las pruebas de asimilación de datos tiene en cuenta la marea total, pero se trata de un modelo cuya resolución horizontal es menor a la de los sistemas de pronóstico globales de la actualidad ($1/3^{\circ}$ frente a $1/12^{\circ}$), que ya incluyen la asimilación de los datos disponibles. Si se compara el prototipo de sistema de pronóstico generado en esta tesis con los pronósticos de sistemas operacionales de otras áreas costeras se observa gran variedad en cuanto a las decisiones que se han tomado para la generación de los distintos sistemas: resolución horizontal, resolución horaria, variables pronosticadas, difusión de los resultados pronosticados, inclusión de la asimilación de datos en tiempo real, y dado el caso, a través de qué técnica y qué variables son las que se asimilan, etc. Así, por ejemplo, en Kowalewska-Kalkowska and Kowalewski (2019) se da la asimilación de datos únicamente para la SST satelital y se opta por un modelo en tres dimensiones con resolución horaria de 1 dato cada 6 horas para las próximas 48 horas con difusión en tiempo real para todas las variables. Sin embargo los pronósticos del sistema de Jalón-Rojas et al. (2015) son en dos dimensiones y no están abiertos al público. El otro sistema que se ha introducido que se encuentra operacional es Costa et al. (2020) que tiene difusión de sus resultados en forma simplificada y es en tres dimensiones sin asimilación de datos.

Por último, los resultados de la evaluación experimental realizada para validar el prototipo indican la importancia de disponer en tiempo real de los datos que se utilizan para constituir las condiciones de borde oceánicas, aluden a una pérdida de capacidad de pronóstico hacia el interior del Río de la Plata y ponen de manifiesto algunas necesidades del sistema como son la investigación de los posibles orígenes de sesgos sistemáticos. Buscando mejorar la calidad de los resultados, en el trabajo se delinea la incorporación de técnicas de asimilación de datos, específicamente medidas de mareógrafos, a través de un esquema de filtro de Kalman por ensambles 4D-LETKF.

VI.2. Publicaciones relativas a la tesis

En este apartado se listan las actividades de difusión científica realizadas en el marco de la tesis.

- Ballesteros, M., Fossati, M., Santoro, P., Ezzatti, P., 2018. Sistema operacional de pronóstico de corrientes y mareas en el Río de la Plata. Condiciones de borde. XXVIII Congreso Latinoamericano de Hidráulica, Buenos Aires, Argentina.
- Ballesteros, M., Fossati, M., Santoro, P., Ezzatti, P., 2019. Prototipo de un Sistema Operacional de Pronóstico de Corrientes y Mareas para el Río de la Plata. II Congreso Agua Ambiente y Energía, Montevideo, Uruguay.
- Ballesteros, M., Santoro, P., Fossati, M., Ezzatti, P., 2019. Development of a water level forecasting system in the Río de la Plata estuary. LAPECO: 1st Latin America Physics of Estuaries and Coastal Oceans conference, Florianópolis, Brasil.

VI.3. Futuras líneas de investigación

Este trabajo se desarrolló con la filosofía de disponer, en base al conocimiento del área, de un primer sistema de pronóstico sobre el cual construir diferentes mejoras y/o extensiones. Luego de realizar el pronóstico surgen varios caminos identificados, en este apartado se identifican los principales.

Una línea de trabajo es la asimilación de datos, que podría tomar la forma descrita en el apartado anterior en una versión inicial y cuyos resultados serán punto de partida para la mejora de este campo. El estudio de las experiencias de sistemas similares ayuda a determinar posibles mejoras cuya exploración resulta de interés.

Las posibilidades de mejora en relación a las condiciones de borde oceánicas que toma este primer prototipo pueden ser incluir, además de los niveles, las corrientes pronosticadas por el modelo global (se deberá estudiar su superposición con las corrientes provenientes del atlas de marea astronómica), así como explorar la utilización de los pronósticos de COPERNICUS en sustitución de los de HYCOM.

Para los campos de forzantes atmosféricos se dan alternativas para su adecuación al área del dominio que incluyen su ajuste teniendo en cuenta observaciones locales de forma similar a como se hizo en Dinápoli et al. (2020), pero también la posibilidad de un subescalado de estos campos atmosféricos que aumenten la resolución de la malla del modelo global.

En relación a los datos de entrada de caudales de agua dulce, por el momento se utilizan caudales constantes. Una posible línea de investigación para el futuro es la consideración de caudales variables a lo largo del año que pueden provenir de una climatología o del pronóstico en base a los caudales medidos en tiempo real (inicialmente podría ser la permanencia del caudal medido durante el tiempo de pronóstico).

Los requerimientos del sistema actual no tienen en cuenta las necesidades de los usuarios reales. El contacto con posibles usuarios determinará áreas a mejorar que pueden incluir desde el requerimiento de una hora del día a la que el pronóstico deba haber sido distribuido (lo que supondrá prestar atención a los tiempos de cómputo y descarga de datos), pasando por resoluciones de detalle en subáreas de interés, nuevas ubicaciones de extracción de series de variables o incluso nuevas variables (por ejemplo corrientes) y llegando a solicitar subproductos como son recomendaciones para navegantes, avisos a puertos, etc. Además, el contacto con posibles usuarios es también una posible fuente de datos medidos frente a los que comparar el sistema.

En relación a las observaciones frente a las que comparar los resultados del modelo, o como posibles datos a incorporar mediante asimilación de datos, en este momento existen pocas medidas que se encuentren disponibles en tiempo real. La posibilidad que ofrecerá el futuro altímetro satelital de banda ancha (<https://swot.jpl.nasa.gov/>) deberá ser explorada. Aun así, si se quiere implementar un sistema en tres dimensiones y también si se quieren incluir nuevas variables como pueden ser el pronóstico de las corrientes o de la salinidad, la necesidad de nuevas medidas que permitan la evaluación de la calidad de los resultados en tiempo real aumenta.

VII. Bibliografía

Aiello, G., Iorio, M., Molisso, F., Sacchi, M., 2020. Integrated Morpho-Bathymetric, Seismic-Stratigraphic, and Sedimentological Data on the Dohrn Canyon (Naples Bay, Southern Tyrrhenian Sea): Relationships with Volcanism and Tectonics. *Geosciences* 2020, 10, 319.

Alonso, R.; Jackson, M.; Santoro, P.; Fossati, M.; Solari, S.; Teixeira, L., 2017. Wave and tidal energy resource assessment in Uruguayan shelf seas. *Renewable Energy, Elsevier*. 114, 18–31.

Ballesteros, M., Fossati, M., Santoro, P., Ezzatti, P., 2018. Sistema operacional de pronóstico de corrientes y mareas en el Río de la Plata. Condiciones de borde. XXVIII Congreso Latinoamericano de Hidráulica.

Bennett, A. F., 1992. *Inverse methods in physical oceanography*, Cambridge University Press, Cambridge.

Brasseur, P., Ballabrera-Poy, J., Verron, J., 1999. Assimilation of altimetric data in the mid-latitude oceans using the Singular Evolutive Extended Kalman filter with an eddy-resolving, primitive equation model, *Journal of Marine Systems*, 22, pp. 269-294.

Canas, A., 2009. Modelling and data assimilation techniques for operational hydrodynamic forecast in Tagus Estuary. PhD thesis. Technical University of Lisbon, Portugal.

Carrasi, A., Bocquet, M., Bertino, L., Evensen, G., 2018. Data Assimilation in the Geosciences: An overview on methods, issues and perspectives. *WIREs Clim Change*, 9(5):e535.

Carrere L, Lyard F, Cancet M, Guillot A (2015) FES 2014, a new tidal model on the global ocean with enhanced accuracy in shallow seas and in the Arctic region. *EGU Gen Assem Conf Abstr*

Chassignet, E. P., Pascual, A., Tintoré J., Verron, J., 2018. *New Frontiers in Operational Oceanography*. GODAE OceanView, 815pp, <https://doi.org/10.17125/gov2018>

Costa, C. G. R., Leite, J. R. B., Castro, B. M., Blumberg, A. F., Georgas, N., Dottori, M., Jordi, A., 2020. An operational forecasting system for physical processes in the Santos-Sao Vicente-Bertioga Estuarine System, Southeast Brazil. *Ocean Dynamics* 70:257–271. <https://doi.org/10.1007/s10236-019-01314-x>

Dinápoli, M.G., Simionato, C.G., Moreira, D., 2020. Development and validation of a storm surge forecasting/hindcasting modelling system for the extensive Río de la Plata Estuary and its adjacent Continental Shelf. *Natural Hazards*. Springer. <https://doi.org/10.1007/s11069-020-04079-5>.

Donlon, C.J., Casey, K.S., Robinson, I.S., Gentemann, C.L., Reynolds, R.W., Barton, I., Arino, O., Stark, J., Rayner, N., LeBorgne, P., Poulter, D., Vazquez-Cuervo, J., Armstrong, E., Beggs, H., Llewellyn-Jones, D., Minnett, P.J., Merchant, C.J., Evans, R., 2009. The GODAE High-Resolution Sea Surface Temperature Pilot Project. *Oceanography* 22(3):34–45, doi:10.5670/oceanog.2009.64.

Etala, P., Saraceno, M., Echevarria, P., 2015. An investigation of ensemble-based assimilation of satellite altimetry and tide gauge data in storm surge prediction, *Ocean Dyn*, vol. 65, no. 3, pp. 435-447.

Ezzatti, P., Fossati, M., Piedra-Cueva, I., 2011. An efficient version of the RMA-11 model. CLEI Electronic Journal. Vol. 14, Number 2, Paper 4.

Ferziger, J.H., Peric', M., 2002. Computational Methods for Fluid Dynamics, third ed. Springer, Berlin, Heidelberg, New York.

Fletcher R. A., 1976. Results from a storm surge prediction model of the North-West European continental shelf for April, November and December 1973. Institute of Oceanography, UK. Report nº24.

Fossati, M. and Piedra-Cueva, I., 2008. Numerical modelling of residual flow and salinity in the Río de la Plata. Applied Mathematical Modelling. Vol. 32, Issue 6, pp. 1066-1086.

Fossati, M. and Piedra-Cueva, I., 2013. A 3D hydrodynamic numerical model of the Río de la Plata and Montevideo's coastal zone, Applied Mathematical Modelling. Vol 37, pp. 1310–1332.

Fossati, M., Fernández, M., Piedra-Cueva, I., 2013. Determination of a submarine outfall discharge location based on a 3D hydrodynamic-lagrangian high resolution model implemented for the Río de la Plata in South America. Ocean modelling for coastal management - Case studies with MOHID.

Fossati, M., Santoro, P., Mosquera, R., Martínez, C., Ghiardo, F., Ezzatti, P., Pedocchi, F., Piedra-Cueva, I., 2014. Dinámica de flujo, del campo salino y de los sedimentos finos en el Río de la Plata. RIBAGUA - Revista Iberoamericana del Agua, Vol. 1, pp. 48-63.

FREPLATA, 2004. Análisis diagnóstico Transfronterizo del Río de la Plata y su Frente Marítimo. Documento Técnico. Proyecto Protección Ambiental del Río de la Plata y su Frente Marítimo. Proyecto PNUD/GEF/RLA/99/G31.

Fu, L-L., Chelton, D.B., Le Traon, P-Y., Morrow, R., 2010. Eddy dynamics from satellite altimetry. Oceanography 23, 14–25. (doi:10.5670/oceanog.2010.02).

Hervouet, J. M., 2007. Hydrodynamics of free surface flows: modelling with the finite element method, John Wiley & Sons Ltd, Chichester, UK.

Hess, K. W., Gross, T. F., Schmalz, R. A., Kelley, J. G. W., Aikman, F., Wei, E., Vincent, M. S., 2003. NOS standards for evaluating operational nowcast and forecast hydrodynamic model systems. NOAA Technical Report NOS CS 17, Silver Spring, MD.

Hunt, B. R., Kostelich E. J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. Physica D., 230, 112-126.

Iermano, I., Moore, A. M., Zambianchi, E., 2016. Impacts of a 4-dimensional variational data assimilation in a coastal ocean model of southern Tyrrhenian Sea. Journal of Marine Systems. vol. 154, pp. 157-171.

Intergovernmental Oceanographic Commission of UNESCO, 2020. Quality Control of in situ Sea Level Observations: A Review and Progress towards Automated Quality Control. Paris, France, Intergovernmental Oceanographic Commission of UNESCO, 71pp. (IOC Manuals and Guides, 83 Vol.1)

Jackson, M., Santoro, P., Fossati, M., 2018. Pronóstico retrospectivo (hindcast) de nivel de mar en la costa de Uruguay. XXVIII Congreso Latinoamericano de Hidráulica

Jalón-Rojas, I., Schmidt, S., Sottolichio, A., 2015. Turbidity in the fluvial Gironde Estuary (southwest France) based on 10-year continuous monitoring: Sensitivity to hydrological conditions. *Hydrol Earth Syst. Sci.*, 19, 2805–2819.

Kalnay, E., 2003. *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge.

Kantha, L. H. and Clayson, C. A., 2000. *Numerical Models of Oceans and Oceanic Processes*. In: *International Geophysics Series*, vol. 66. Academic Press, San Diego. 940 pp.

Kowalewski, M. and Kowalewska-Kalkowska, H., 2017. Sensitivity of the Baltic Sea level prediction to spatial model resolution, *Journal of Marine Systems*, 173, 101–113, doi:10.1016/j.jmarsys.2017.05.001

Kowalewska-Kalkowska, H. and Kowalewski, M., 2019. Combining Satellite Imagery and Numerical Modelling to Study the Occurrence of Warm Upwellings in the Southern Baltic Sea in Winter, *Remote Sensing*, 11, art. no. 2982, <https://doi.org/10.3390/rs11242982>

Kowalik, Z. and Murty, T. S., 1993. *Numerical Modeling of Ocean Dynamics*. Singapore: World Scientific.

Lahoz W. A., Khattatov B., Ménard R. (eds.), 2010. *Data Assimilation: making sense of observations*. Springer ISBN: 9783540747024.

Le Traon, P.-Y., 2013. From satellite altimetry to Argo and operational oceanography: three revolutions in oceanography. *Ocean Science*. 9, 901–915. <http://dx.doi.org/10.5194/os-9-901-2013>.

Le Traon, P.-Y., Antoine, D., Bentamy, A., Bonekamp, H., Breivik, L.A., Chapron, B., Corlett, G., Dibarboure, G., DiGiacomo, P., Donlon, C., Faugère, Y., Font, J., Girard-Arduin, F., Gohin, F., Johannessen, J.A., Kamachi, M., Lagerloef, G., Lambin, J., Larnicol, G., Le Borgne, P., Leuliette, E., Lindstrom, E., Martin, M.J., Maturi, E., Miller, L., Mingsen, L., Morrow, R., Reul, N., Rio, M.H., Roquet, H., Santoleri, R., Wilkin, J., 2015. Use of satellite observations for operational oceanography: recent achievements and future prospects. *Journal of Operational Oceanography*, 8:sup1, s12-s27, DOI: 10.1080/1755876X.2015.1022050

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., Le Traon, P.-Y., 2018. Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system. *Ocean Science*, 14, 1093–1126, <https://doi.org/10.5194/os-14-1093-2018>.

Martínez, C., Fossati, M., Silva, J. P., Dufrechou, E., Santoro, P., Ezzatti, P., Piedra-Cueva, I., 2015. Towards a 3D Hydrodynamic numerical modeling system for long term simulations of the Río de la Plata dynamic. In: *Proceedings of the 36th IAHR World Congress, Delft - The Hague*.

Mendes, J., Leitão, P., Chambel Leitão, J., Bartolomeu, S., Rodrigues, J., Dias, J. M., 2019. Improvement of an Operational Forecasting System for Extreme Tidal Events in Santos Estuary (Brazil). *Geosciences*, 9, 511.

NCEP, NWS, NOAA, and DOC: NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, <https://doi.org/10.5065/D65D8PWK>, 2015.

- Paradis, D., Ohl, P., Daniel, P., 2009. Operational storm surges forecasting in an estuary. *Marine Geodesy*; 32:166-177
- Pawlowicz, R., Beardsley, B., Lentz, S., 2002. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Computers Geosciences*. 28, 929–937.
- Pham, D., Verron, J., Roubaud, M., 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography, *Journal of Marine Systems*, 16, pp. 323-340.
- Piedra-Cueva, I. and Fossati, M., 2007. Residual currents and corridor of flow in the Rio de la Plata. *Applied Mathematical Modelling*. Vol. 31, pp. 564–577.
- Popescu, I., 2014. *Computational Hydraulics: Numerical methods and modelling*. IWA Publishing, London, UK. ISBN 9781780400457
- Santoro, P., 2011. Estudio de la marea meteorológica en el Río de la Plata. MSc thesis. Universidad de la República, Uruguay.
- Santoro, P., Fernández, M., Fossati, M., Cazes, G., Terra, R., Piedra-Cueva, I., 2011. Pre-operational forecasting of sea level height for the Río de la Plata. *Applied Mathematical Modelling*. Vol. 35, pp. 2462–2478.
- Santoro, P. E., Fossati, M., Piedra Cueva, I., 2013. Study of the meteorological tide in the Río de la Plata. *Continental Shelf Research*. Vol. 60, pp. 51-63.
- Santoro, P., Fossati, M., Tassi, P., Huybrechts, N., Pham Van Bang, D., Benoit, M., Piedra-Cueva, I., 2015. Hydrodynamic and fine sediment transport numerical modelling, application to the Río de la Plata and Montevideo Bay. XXII TELEMAC-MASCARET User Club, Daresbury Laboratory, UK, pp. 119-126.
- Santoro, P., Fossati, M., Tassi, P., Huybrechts, N., Pham Van Bang, D., Piedra-Cueva, I., 2017. A coupled wave-current-sediment transport model for an estuarine system: Application to the Río de la Plata and Montevideo Bay. *Applied Mathematical Modelling*. Vol. 52, pp. 107–130.
- Santoro, P., Fossati, M., Tassi, P., Huybrechts, N., Pham Van Bang, D., Piedra-Cueva, I., 2019. Effect of self-weight consolidation on a hydro-sedimentological model for the Río de la Plata estuary. *International Journal of Sediment Research* Vol. 34, pp. 444–454.
- Schäfer. M., 2006. *Computational engineering: introduction to numerical methods*. Springer, Berlin.
- Schiller, A. and Brassington, G. B., 2011. *Operational Oceanography in the 21st Century*. Springer ISBN: 9789400703315.
- Tandeo, P., Ailliot, P., Bocquet, M., Carrassi, A., Miyoshi, T., Pulido, M., Zhen, Y., 2018. Joint Estimation of Model and Observation Error Covariance Matrices in Data Assimilation: a Review. arXiv preprint arXiv:1807.11221., in review, 2020.
- UNESCO, 1993. *Manual of Quality Control Procedures for Validation of Oceanographic Data. Manual and Guides*, 26, 437 pp.
- UNESCO, 1985. *IOC Manual and Guides Nro. 14*, 83pp.
- Van Leeuwen, P. J., 2009. Particle filtering in geophysical systems. *Monthly Weather Review*, 137, pp.4089-4114.

Van Leeuwen P. J., Künsch H. R., Nerger L., Potthast R., Reich S., 2019. Particle filters for high-dimensional geoscience applications: A review. *Quarterly Journal of the Royal Meteorological Society*, 2019. ISSN 1477870X. doi: 10.1002/qj.3551.

Winckler, P., 2018. *Introducción al modelado de procesos costeros*. Primera Edición. Valparaíso, Chile: Universidad de Valparaíso.

VIII. Anexo 1

Se evalúan para el año 2015, los resultados del modelo numérico de marea en el Atlántico Sur ASTide (Martinez et al., 2015) en los puntos del borde oceánico del modelo local RPtide (Figura 73), con los valores de niveles del modelo global COPERNICUS+FES2014 (FES2014 incorpora la componente astronómica). Las comparaciones realizadas fueron: visualización de series, desviación estándar de la serie para cada punto y correlación de Pearson. Se comparan, además los registros de los mareógrafos de Mar del Plata y La Paloma con las simulaciones de COPERNICUS y ASTide mediante *scatter*, cálculo de desviación estándar, correlación de Pearson y CRMSE.

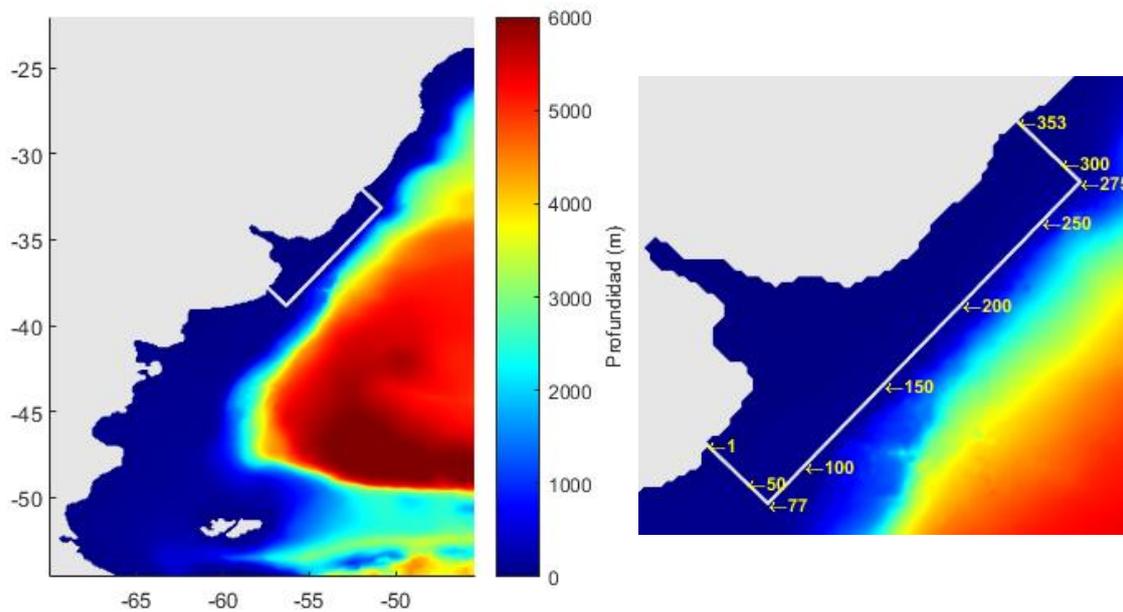


Figura 73 – Dominio del modelo ASTide (panel izquierdo). Zoom al dominio del modelo local RPtide con la numeración de los puntos del borde (panel derecho).

El modelo ASTide tiene como condiciones de borde el atlas de marea astronómica y vientos y presiones de reanálisis en la superficie libre por lo que es capaz de proporcionar las condiciones de contorno al modelo RPtide correspondientes a la marea total (astronómica más meteorológica). El cero de referencia del sistema es el cero Wharton, que pasa 0.91 m por encima del Cero Oficial que coincide con el nivel medio de las aguas del puerto de Montevideo. Durante la calibración del sistema ASTide-RPtide se determinó que al nivel proporcionado por el atlas de marea astronómica se sumara 0.91 m antes de introducirlo como condición de contorno al ASTide. En la Figura 74 se puede ver en color azul el nivel medio para el año 2015 para cada punto del borde del RPtide dónde recibe los datos generados por el ASTide.

La referencia del modelo global COPERNICUS (y también HYCOM) es el geoides. Una superficie de geopotencial constante con la que coincidiría el nivel medio del mar si el océano estuviera en reposo. El geoides es también la referencia del atlas de marea astronómica FES2014 que se suma a los modelos globales para obtener la marea total. En la Figura 74 se puede ver en color naranja el nivel medio para el año 2015 para cada punto del borde del RPtide de los niveles generados por la suma de COPERNICUS+FES2014. Se hace notar que los valores medios de esta suma coinciden con los valores medios de únicamente el modelo COPERNICUS y únicamente el modelo FES2014.

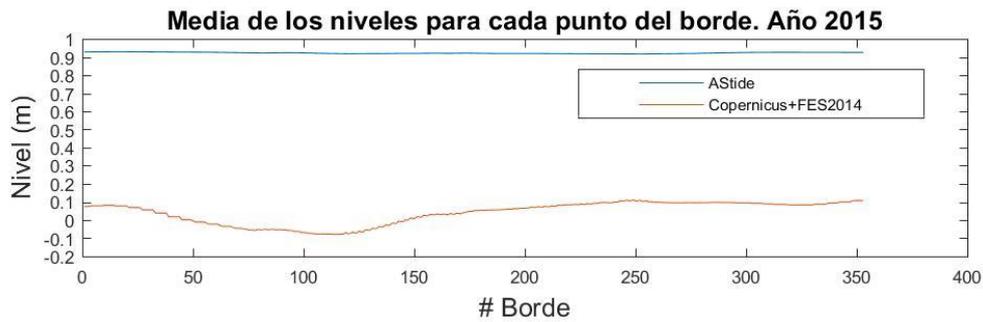


Figura 74 – Media de los niveles en los puntos del borde del RPtide. Modelos AStide y COPERNICUS+FES2014. Año 2015.

Tras la comparación de niveles medios se determina sumar 0.91 m a los niveles generados de COPERNICUS+FES2014. En la Figura 75 se presenta, a modo de ejemplo, la serie de niveles durante el mes de marzo de 2015 correspondiente al borde #100 (Figura 73). Se presentan comparaciones frente a los datos medidos en las estaciones de Mar del Plata y La Paloma, que aunque se encuentran dentro del dominio del RPtide son las estaciones más cercanas a su borde. En las figuras Figura 76 y Figura 77 se presenta la comparación de niveles para 15 días de octubre de 2015 para Mar del Plata y La Paloma, respectivamente. Los resultados de desviación estándar y correlación Pearson para todos los puntos del borde se presentan en la Figura 78 y en la Figura 79, y para las estaciones de Mar del Plata y La Paloma se pueden ver los resultados de correlación Pearson y CRMSE en los diagramas de dispersión de cada modelo (Figura 80 y Figura 81).

Atendiendo a estas comparaciones, no se destaca la representación que da un modelo frente al otro. Ambos modelos, Mohid AStide y COPERNICUS dan una representación similar de lo registrado por los mareógrafos.

Se generaron comparaciones de las componentes de la velocidad obtenidas por cada uno de los modelos para los puntos del borde y para las ubicaciones de los mareógrafos, no pudiendo contar con mediciones para su comparación. En las figuras desde Figura 82 hasta Figura 89 se presentan las comparaciones de serie, media, desviación estándar y correlación Pearson.

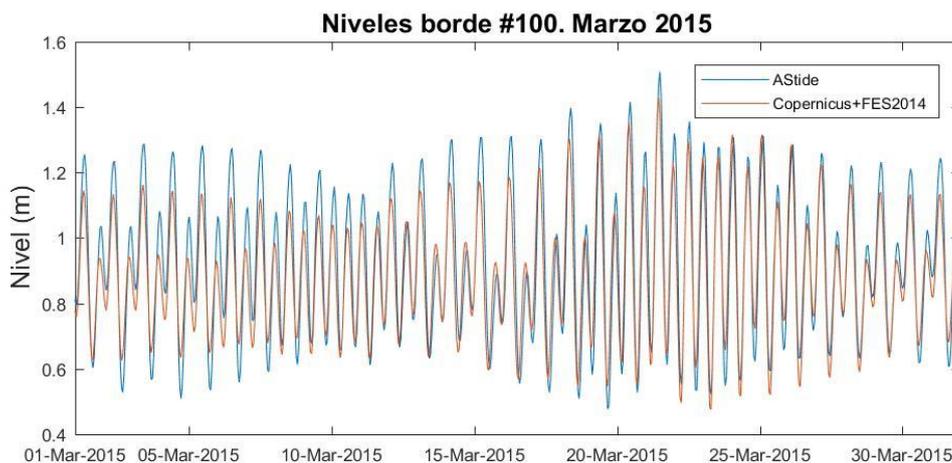


Figura 75 – Serie de los niveles obtenidos por los modelos AStide y COPERNICUS+FES2014 en el centésimo punto del borde del dominio RPtide. Marzo 2015.

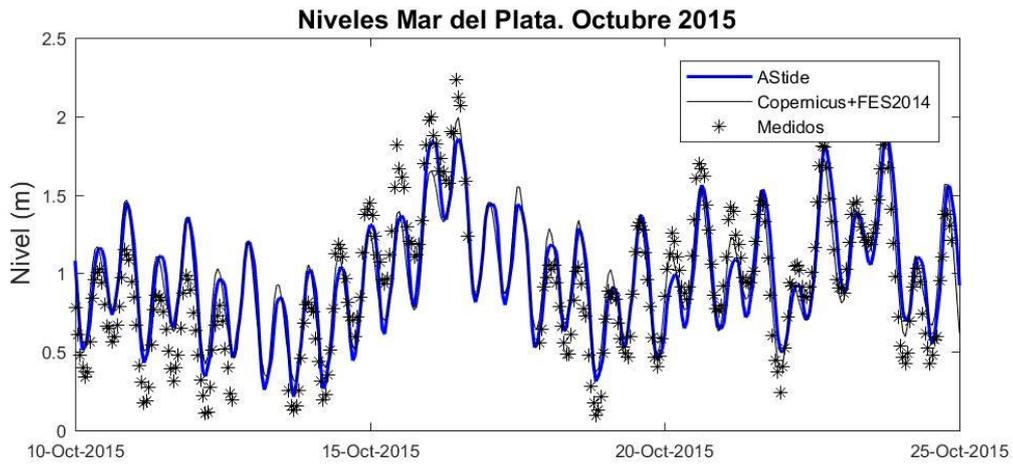


Figura 76 – Serie de los niveles obtenidos por los modelos AS tide y COPERNICUS+FES2014 y serie de niveles medidos en Mar del Plata. Octubre 2015.

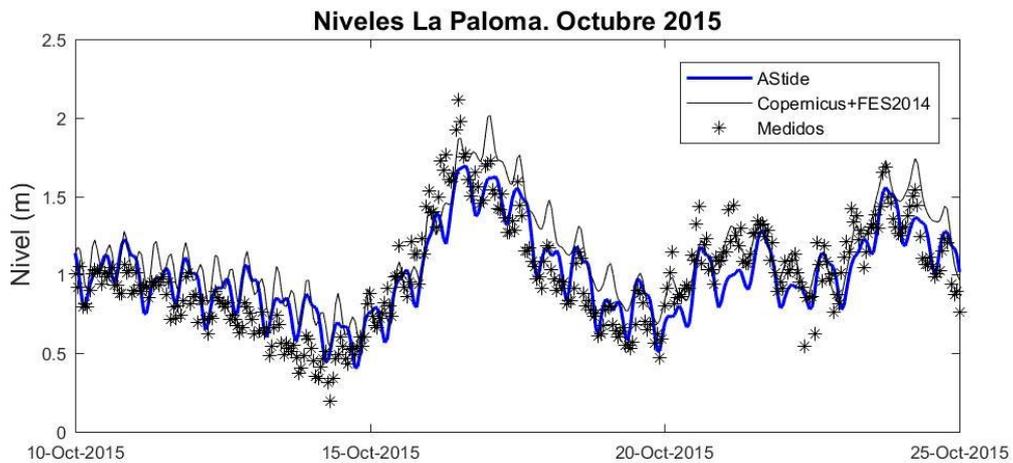


Figura 77 – Serie de los niveles obtenidos por los modelos AS tide y COPERNICUS+FES2014 y serie de niveles medidos en La Paloma. Octubre 2015.

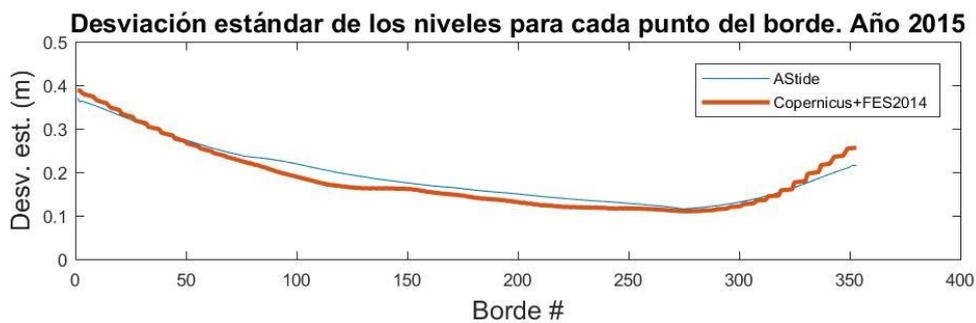


Figura 78 – Desviación estándar de los niveles en los puntos del borde del RPTide. Modelos AS tide y COPERNICUS+FES2014. Año 2015.

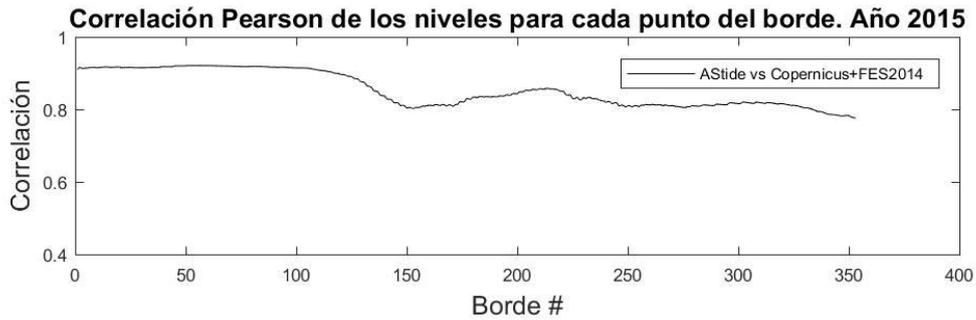


Figura 79 – Correlación Pearson de los niveles en los puntos del borde del RPTide. Modelos AS tide y COPERNICUS+FES2014. Año 2015

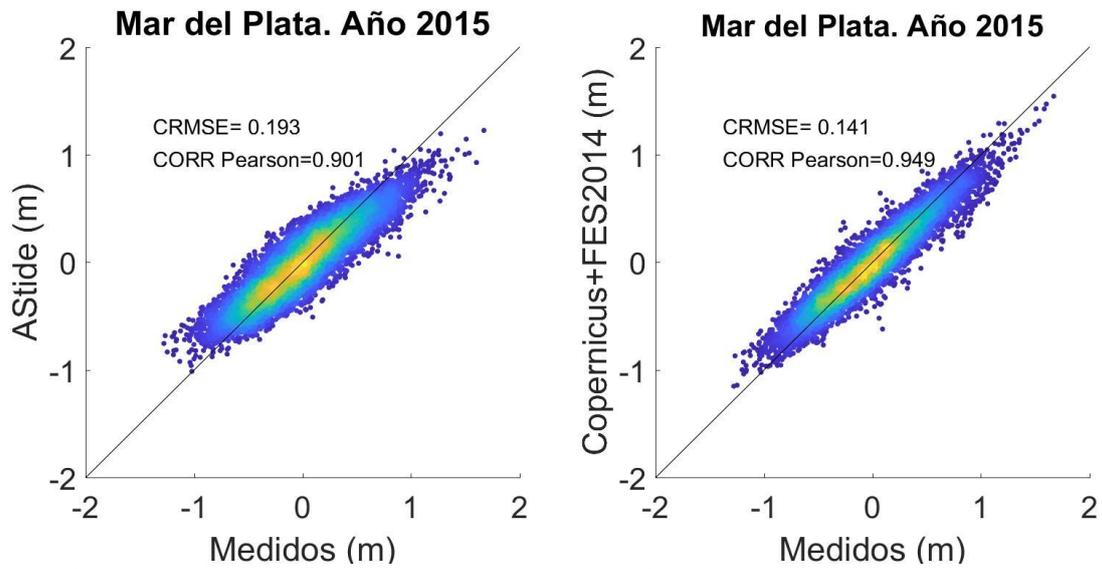


Figura 80 - Scatter, CRMSE y correlación Pearson. Mar del Plata – Año 2015.

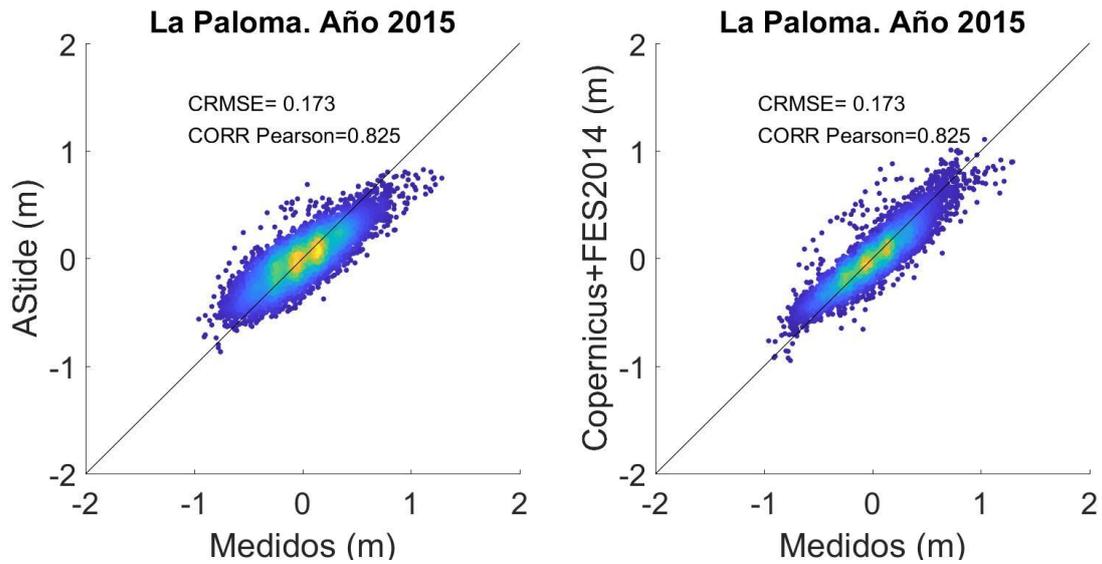


Figura 81 - Scatter, CRMSE y correlación Pearson. La Paloma – Año 2015.

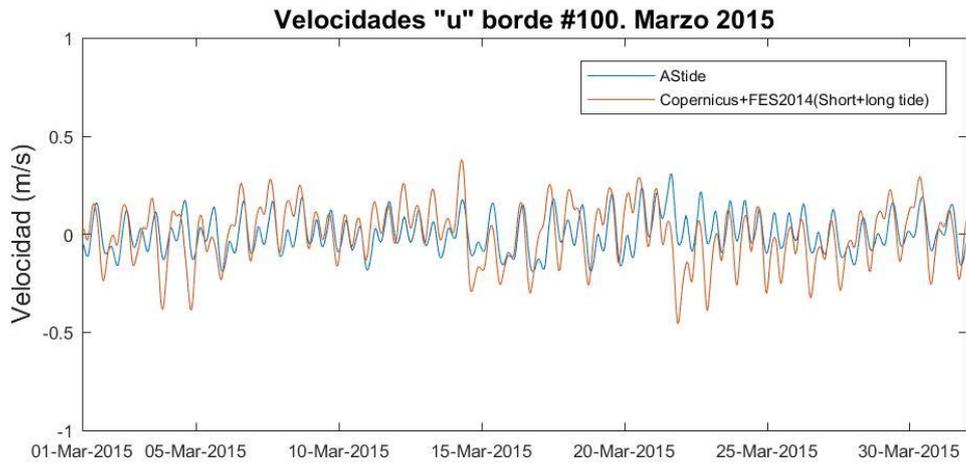


Figura 82 – Velocidades "u" borde 100 – Año 2015. Zoom de 15 días de duración

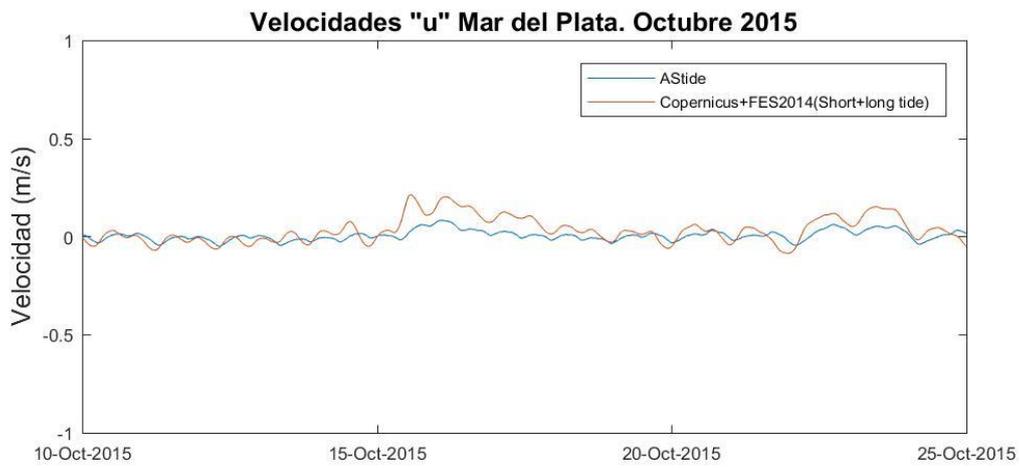


Figura 83 – Velocidades "u" Mar del Plata – Año 2015. Zoom de 15 días de duración

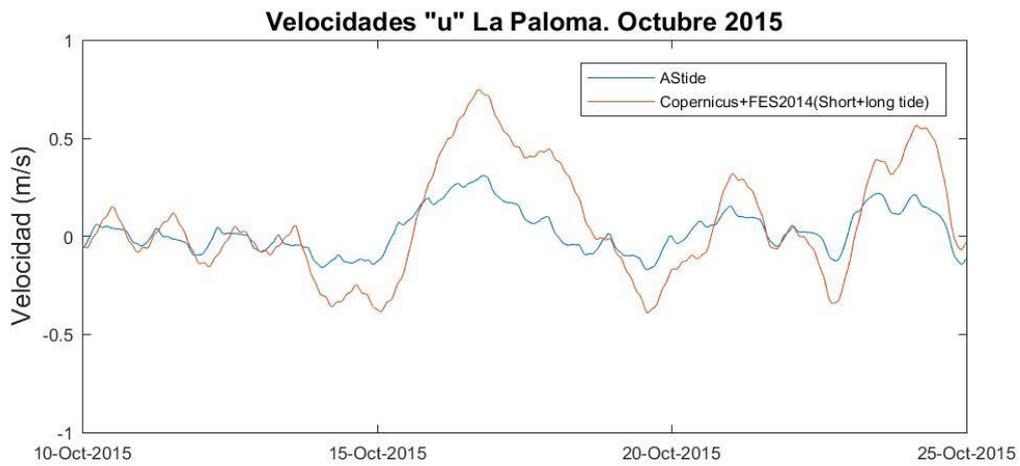


Figura 84 – Velocidades "u" La Paloma – Año 2015. Zoom de 15 días de duración

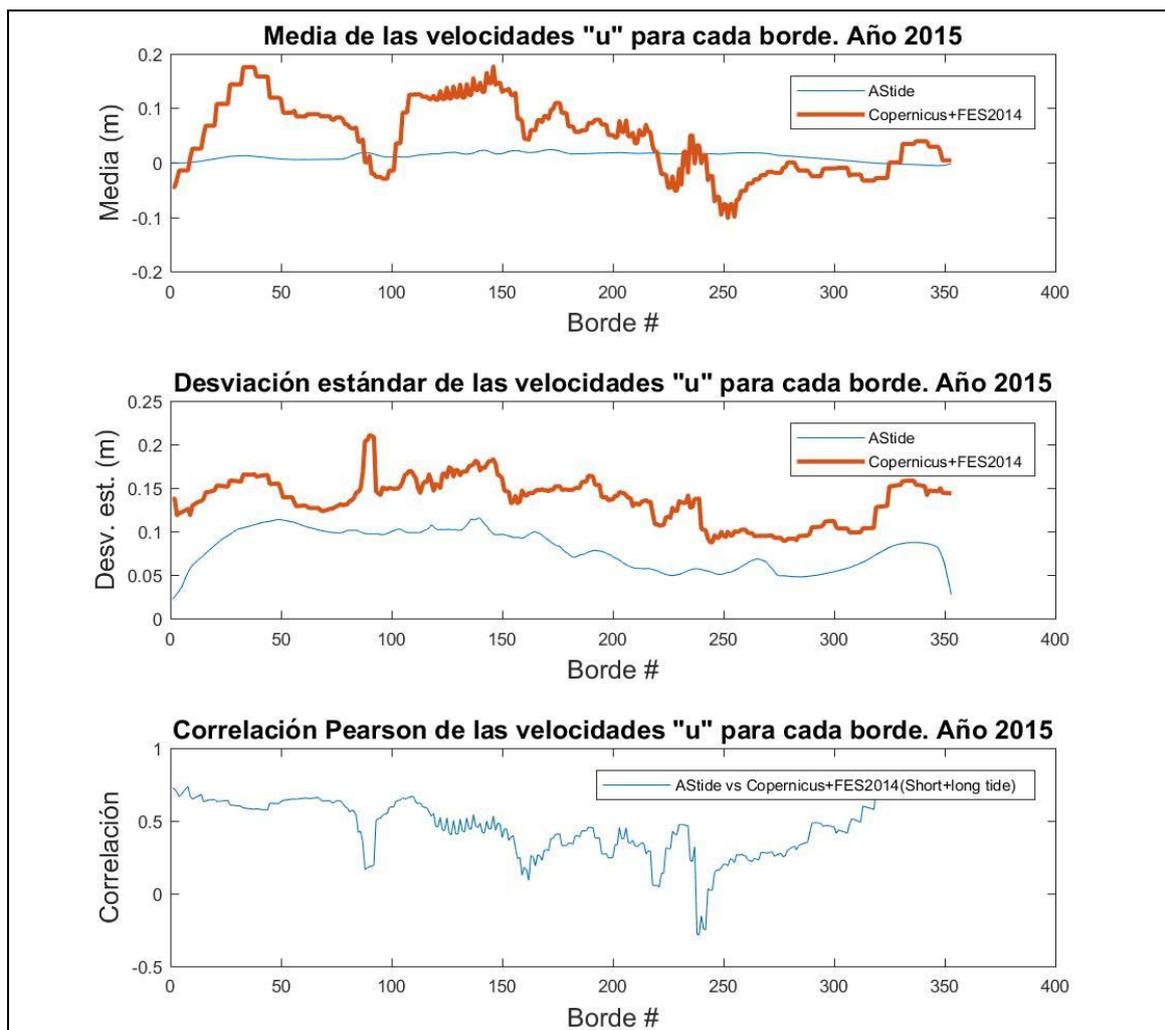


Figura 85 – Análisis de las velocidades “u”. Año 2015 – Media, desviación estándar y correlación de Pearson de la serie para cada punto del borde del RPtide

	Velocidades “u”	
	Mar del Plata. Año 2015.	La Paloma. Año 2015
Desviación estándar AStide	0.0245	0.0940
Desviación estándar COPERNICUS	0.0680	0.2659
Corr. Pearson AStide_COPERNICUS	0.5364	0.7915

Tabla 15 - Análisis de las velocidades “u”. Año 2015 – Desviación estándar y correlación de Pearson de la serie para las estaciones del Mar del Plata y La Paloma

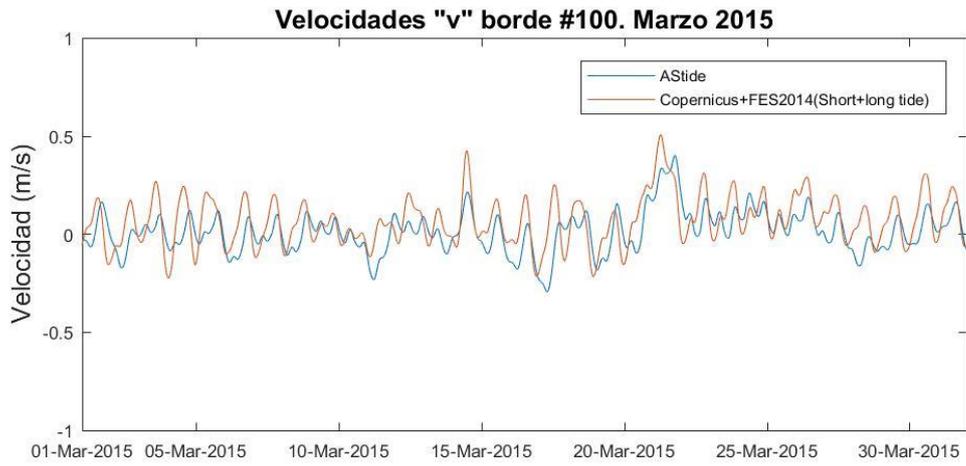


Figura 86 – Velocidades "v" borde 100 – Año 2015. Zoom de 15 días de duración

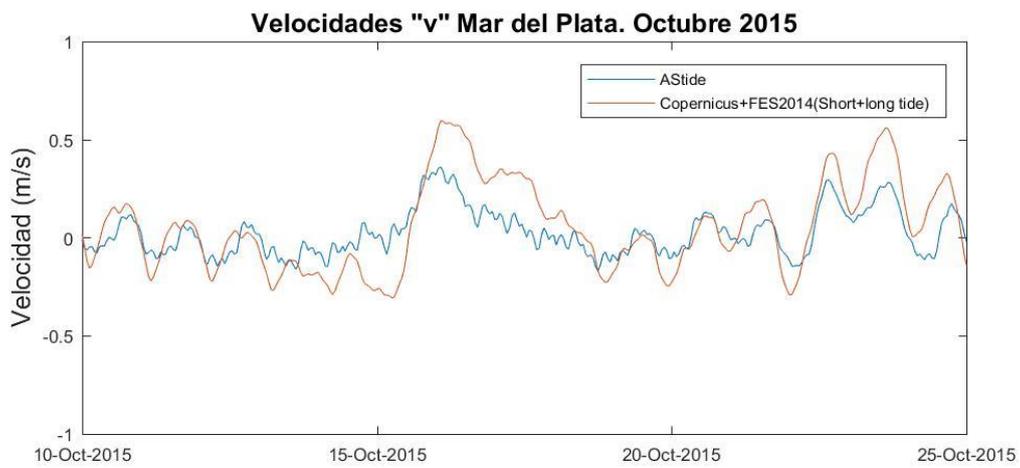


Figura 87 – Velocidades "v" Mar del Plata – Año 2015. Zoom de 15 días de duración

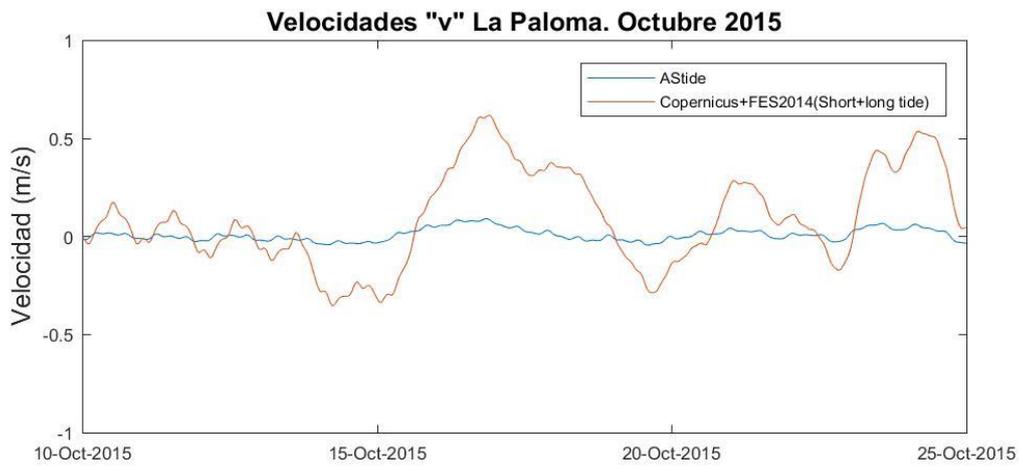


Figura 88 – Velocidades "v" La Paloma – Año 2015. Zoom de 15 días de duración

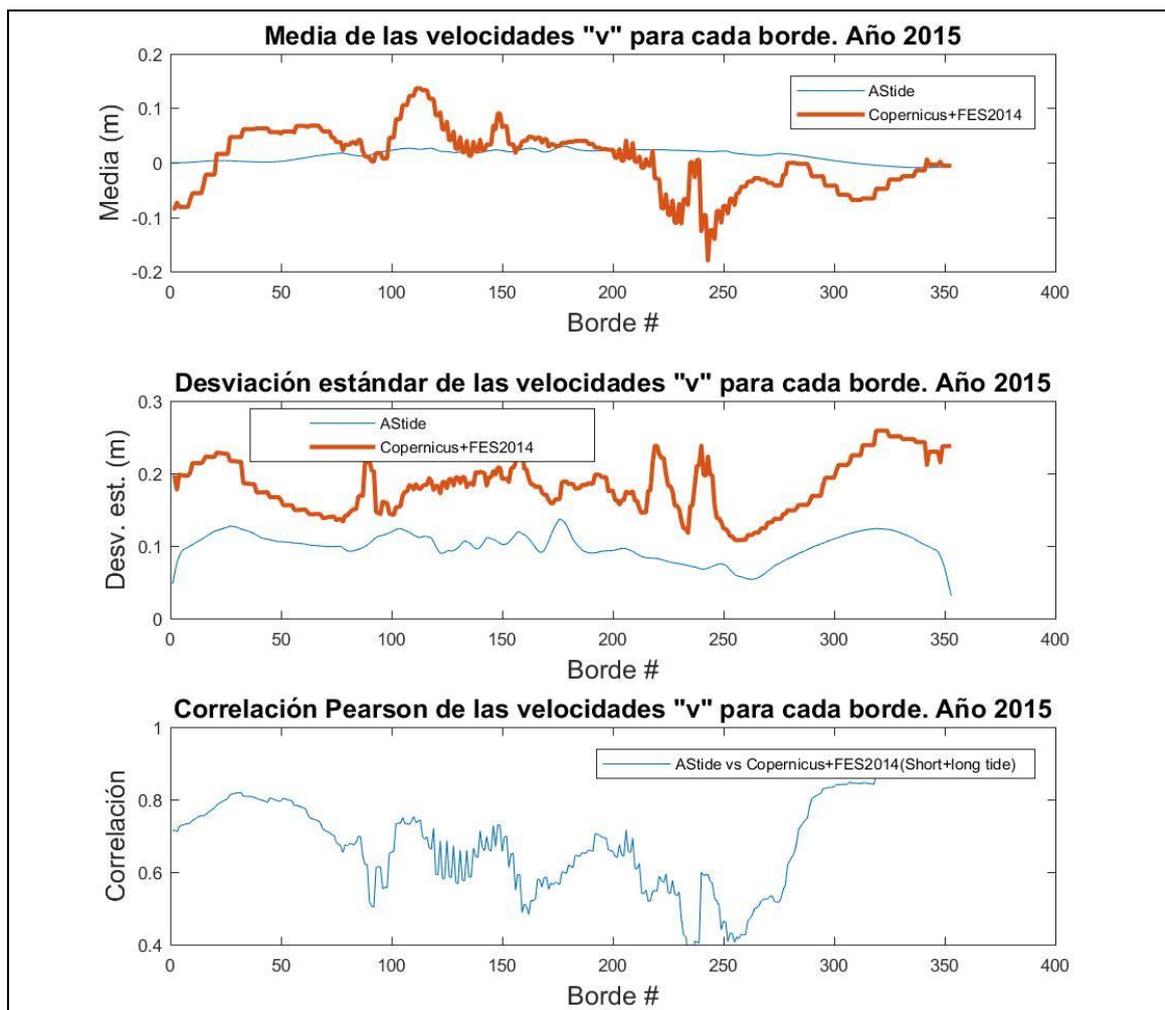


Figura 89 – Análisis de las velocidades “v”. Año 2015 – Media, desviación estándar y correlación de Pearson de la serie para cada punto del borde del RPtide

	Velocidades “v”	
	Mar del Plata. Año 2015.	La Paloma. Año 2015
Desviación estándar AStide	0.1081	0.0259
Desviación estándar COPERNICUS	0.2135	0.2385
Corr. Pearson AStide_COPERNICUS	0.6964	0.7834

Tabla 16 - Análisis de las velocidades “v”. Año 2015 – Desviación estándar y correlación de Pearson de la serie para las estaciones del Mar del Plata y La Paloma

IX. Anexo 2

Se exploran los datos procedentes del altímetro satelital SARAL sin procesar/sin validar que se podrían obtener en tiempo real desde la web <https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/ssha.html#c6652>. Se realizan comparaciones visuales para el año 2015 de los registros frente a los resultados de los modelos COPERNICUS y MOHID_RPtide. Para ello, al resultado del modelo MOHID_RPtide se le extrae la componente astronómica (Pawlowicz et al., 2002).

Para el año 2015 se dieron un total de 342 pases del altímetro en la zona del Río de la Plata (entre las coordenadas lat $[-40,-30]$, lon $[-60,-50]$). El recorrido de estos pases se puede ver en la Figura 90, panel izquierdo, se marca en color azul el dominio del RPtide. En la Figura 91 se presentan las alturas registradas por los altímetros frente a los niveles de los modelos COPERNICUS y RPtide, a modo de ejemplo se muestran las 5 trazas registradas entre el 13 y el 19 de diciembre (trazas 326 a 330, Figura 90, panel derecho).

Esta primera exploración sirve para conocer la distribución espacial y temporal de los datos del altímetro. Se observa una cobertura global para el área de estudio incluyendo la zona interior del Río de la Plata. La frecuencia temporal con la que el altímetro pasa por el dominio estudiado (entre las coordenadas lat $[-40,-30]$, lon $[-60,-50]$) es menor a una vez al día en promedio, pero puede llegar a 2 veces al día (por ejemplo el 13 de diciembre de 2015). Cuando atendemos únicamente al dominio del RPtide la frecuencia disminuye.

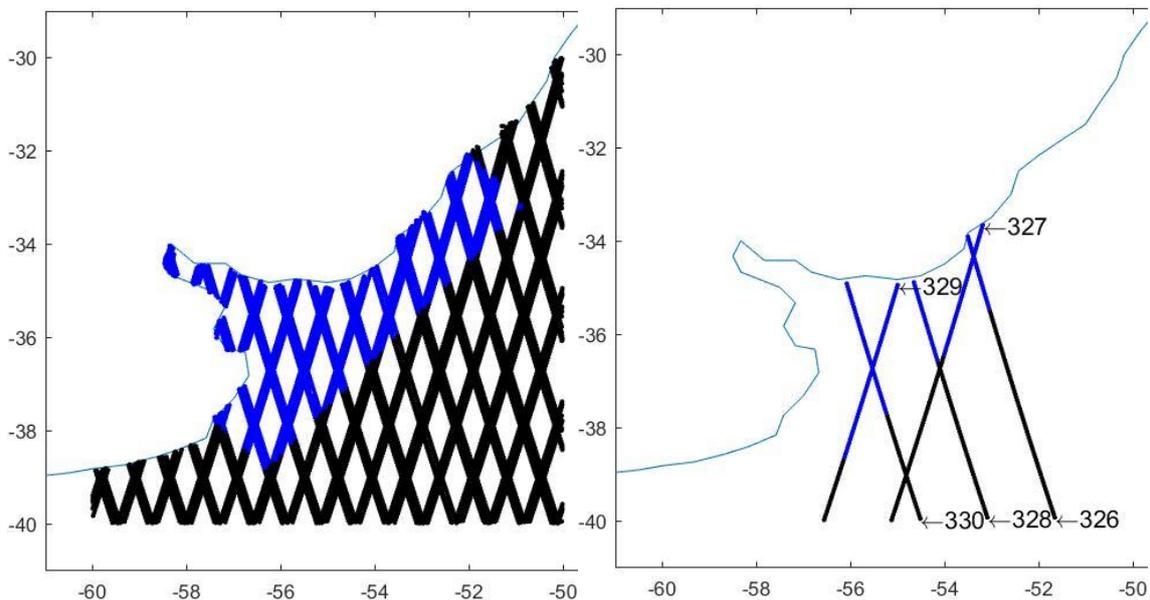


Figura 90 - Pases del altímetro SARAL entre las coordenadas lat $[-40,-30]$, lon $[-60,-50]$. En el panel izquierdo se pueden ver los 342 pases del año 2015 y en el panel derecho los 5 pases entre el 13 y el 19 de diciembre de 2015. Se marca en color azul el dominio del RPtide.

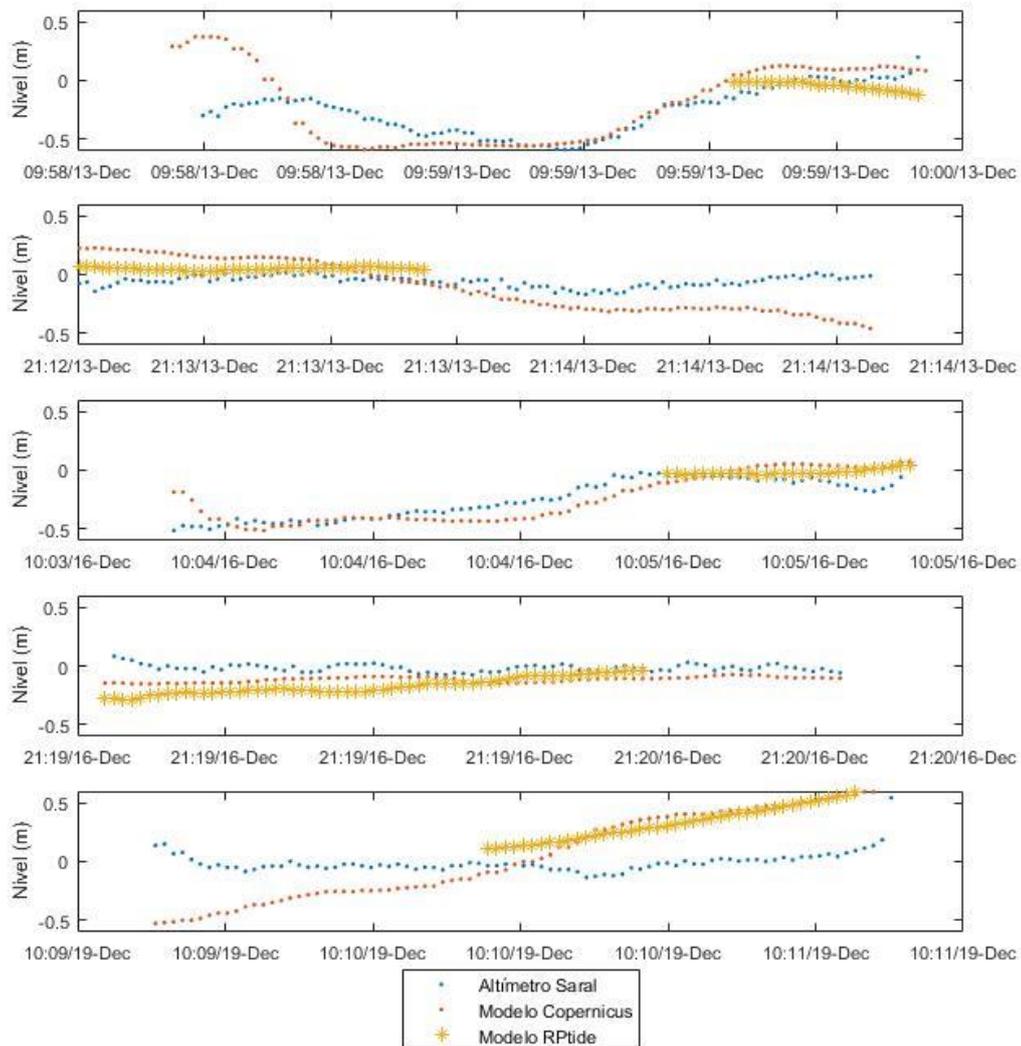


Figura 91 – Comparación registros Altímetro SARAL, modelo COPERNICUS y modelo MOHID_RPtide. Año 2015 del 13 al 19 de diciembre.

En la Figura 92 se pueden ver los valores que se obtienen en los puntos de la malla del modelo COPERNICUS cercanos a las ubicaciones de las trazas del Altímetro satelital SARAL. En la Figura 93 aparece el mismo registro para el altímetro pero para su comparación frente al MOHID_RPtide se visualiza únicamente el dominio de este último y al resultado del modelo se le extrae la marea astronómica. De modo de obtener una síntesis de lo visualizado por los mapas anteriores, en la Figura 94 (panel izquierdo) se muestra un *scatter* de los valores de la Figura 92 y en la Figura 94 (panel derecho) se superpone el resultado de ambos modelos únicamente para el dominio del RPtide.

Las comparaciones frente a los niveles de los modelos sirven para conocer que los datos del altímetro se asemejan a los del modelo COPERNICUS en que no incorporan marea astronómica, pero, aunque el modelo global incluye los datos del altímetro entre las fuentes de los datos que asimila, se encuentran diferencias en muchas de las trazas comparadas.

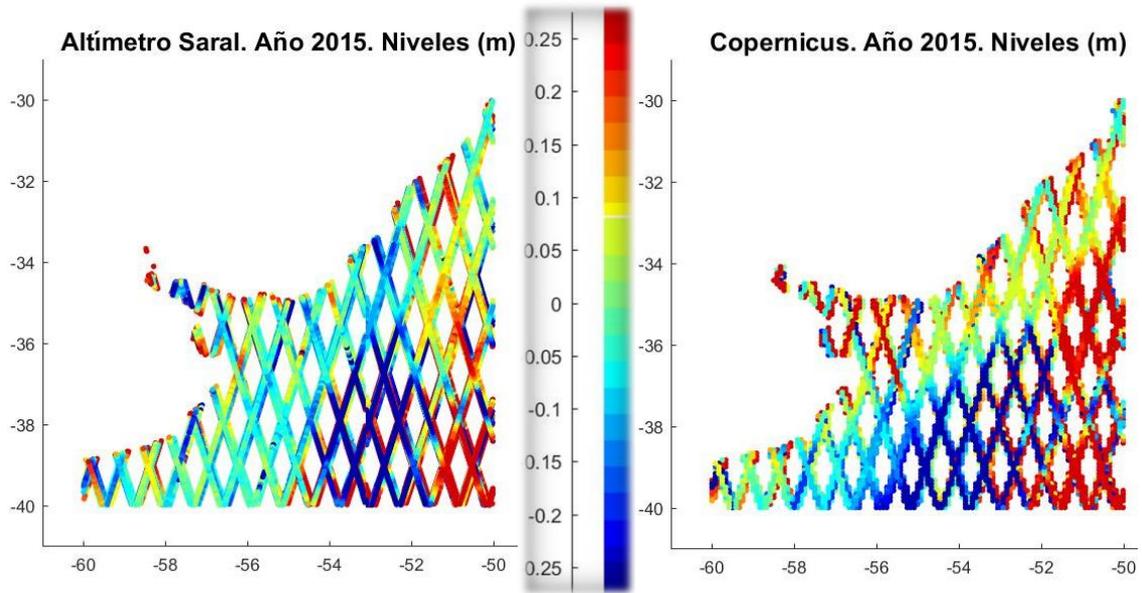


Figura 92 - Comparación de los niveles del Altímetro SARAL y el modelo COPERNICUS. Año 2015.

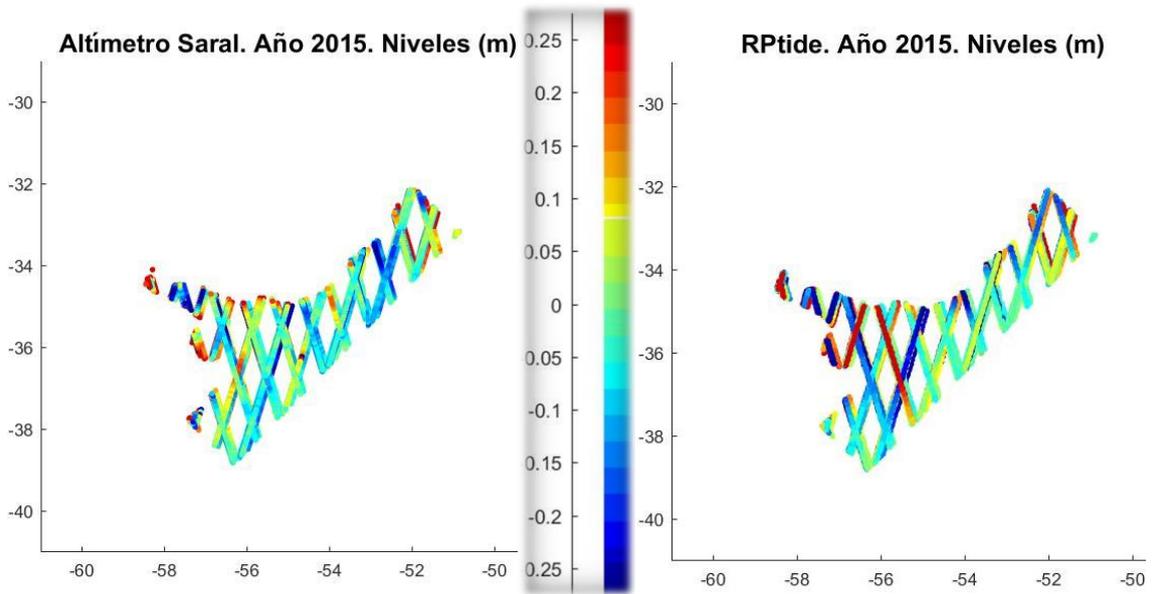


Figura 93 - Comparación de los niveles del Altímetro SARAL y el modelo MOHID_RPtide. Año 2015.

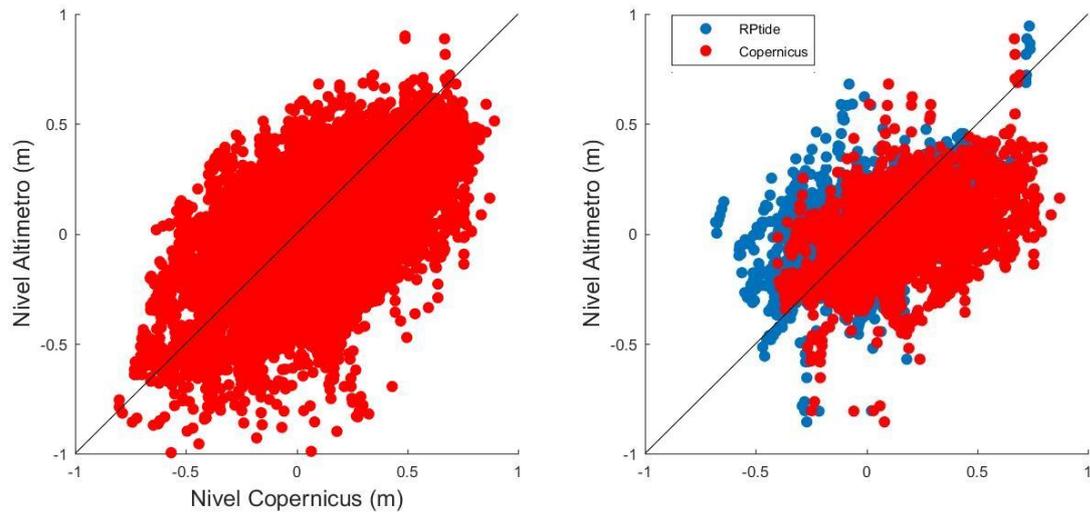


Figura 94 - Scatter del nivel registrado por el altímetro satelital SARAL en el año 2015 frente al modelo COPERNICUS (panel izquierdo), frente a los modelos COPERNICUS y MOHID_RPtide en el dominio del RPtide(panel derecho).

X. Anexo 3

Filtrado de los datos de niveles de mareógrafos

Las medidas de nivel disponibles en tiempo real corresponden a los mareógrafos de La Paloma, Montevideo (Muelle Fluvial) e Isla de Flores en Uruguay y de Atalaya, Buenos Aires, La Plata, Torre Oyarvide, Pilote Norden, San Clemente, San Fernando, Mar del Plata, y Santa Teresita en Argentina (ver ubicación en la Figura 95). En este apartado se describe el preproceso que se crea de manera automática a los datos descargados para su inclusión en la base de datos. El mismo fue desarrollado al generar la depuración y control de los datos disponibles desde el 1/1/2018 al 31/12/2018. La Tabla 17 presenta la cantidad de datos faltantes y la cantidad de datos disponibles luego del filtrado para cada estación y en los siguientes epígrafes se describe la metodología de depuración aplicada a los datos de cada estación.

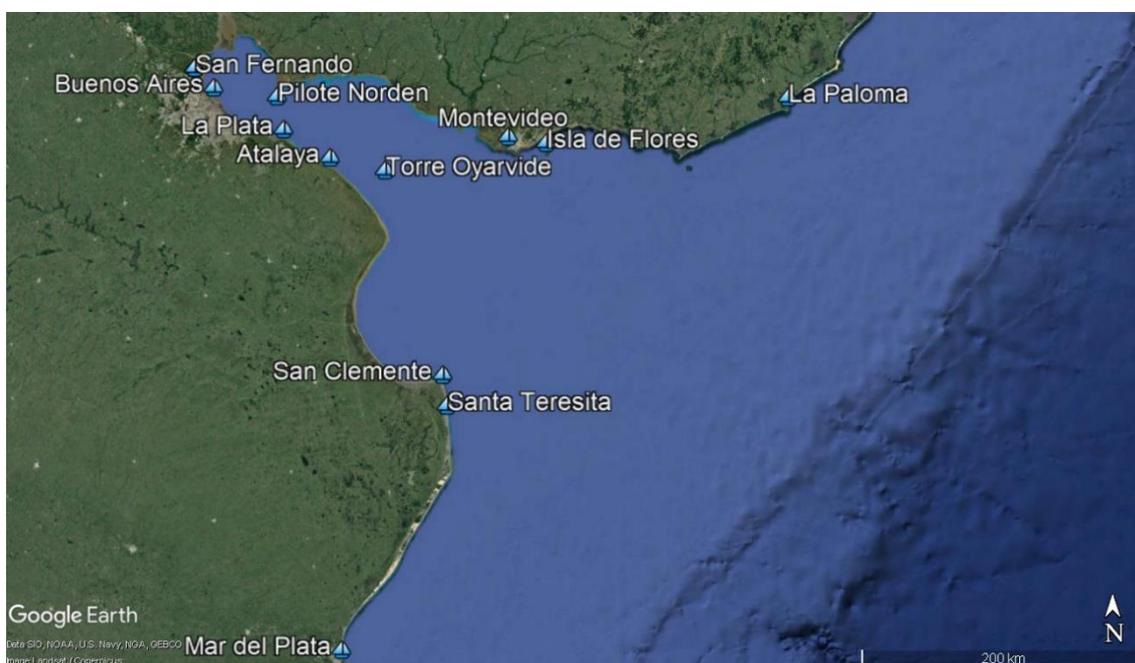


Figura 95 –Ubicación mareógrafos.

Estación	% de datos faltantes	% de datos tras el filtrado
Montevideo	0.05%	96.64%
Isla de Flores	0.02%	93.38%
La Paloma	32.73%	59.54%
Atalaya	34.26%	65.72%
Buenos Aires	12.66%	87.33%
La Plata	14.29%	85.71%
Mar del Plata	13.90%	86.08%
Torre Oyarvide	16.83%	83.16%
Pilote Norden	27.29%	72.67%
San Clemente	22.59%	77.41%
San Fernando	76.43%	23.57%
Santa Teresita	100%	0%

Tabla 17 – Porcentaje de datos faltantes para el año 2018 y porcentaje de datos obtenidos luego del proceso de depuración para el año 2018 para los mareógrafos de Uruguay y Argentina.

X.1. Montevideo (Muelle Fluvial)

Estos datos son facilitados por la Administración Nacional de Puertos (ANP), Uruguay, a través de su web: <http://mareografo.anp.com.uy/>. Con una frecuencia de un dato cada 5 minutos se puede acceder a las mediciones de Nivel de Agua Máximo, Nivel de Agua Medio y Nivel de Agua Mínimo, siendo los datos faltantes comunes a las 3 medidas. Para el período 1/1/2018 al 31/12/2018 se dan datos faltantes en 48 ocasiones, lo que es equivalente a un 0.05%. Estos datos corresponden a 4 intervalos que se presentan en la Tabla 18.

Intervalos con datos faltantes		
Inicio	Fin	Duración
20180325/02:00	20180325/02:55	60 min
20181008/08:45	20181008/09:40	60 min
20181016/08:55	20181016/09:50	60 min
20181022/08:40	20181022/09:35	60 min
	Total	4 horas

Tabla 18 – Intervalos con datos faltantes. Mareógrafo de Montevideo (Muelle Fluvial), 2018

Filtro 1: Filtrado por diferencia entre el registro máximo y mínimo

El objetivo de este filtro, Filtro 1, es eliminar los datos cuando se han presentado perturbaciones que no corresponden a la señal de marea. Se identifican estas perturbaciones con los casos en que la señal registrada máxima y la señal registrada mínima presentan diferencias importantes (la Figura 96 muestra un ejemplo el 8/11/2018 entre las 6:00 y las 12:00).

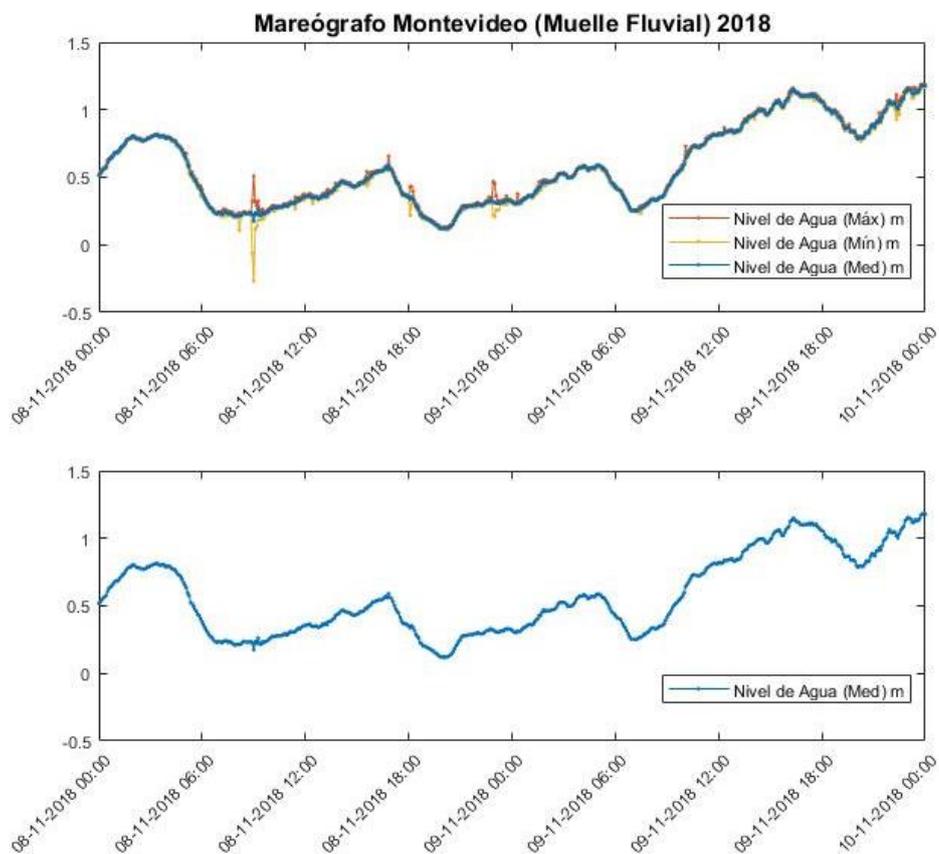


Figura 96 – Niveles mareógrafo Montevideo (Muelle Fluvial) 8-11-2018 a 10-11-2018. El 08-11-2018 entre las 6:00 y las 12:00 la señal registrada máxima y la señal registrada mínima presenta diferencias importantes.

Para el período 1/1/2018 al 31/12/2018 las diferencias entre el registro máximo y mínimo se muestran en la Figura 97 (panel izquierdo) y en la Figura 97 (panel derecho) se puede ver la distribución acumulada empírica de esta diferencia. Se eligió como límite una diferencia de 0.08 m a partir del cambio de tendencia en la distribución acumulada empírica. Así, los valores de Nivel de Agua Medio correspondientes con una diferencia registrada mayor serán depurados. Para el 2018 este filtro elimina 3,415 datos del mareógrafo de Montevideo lo que corresponde con un 3.25%.

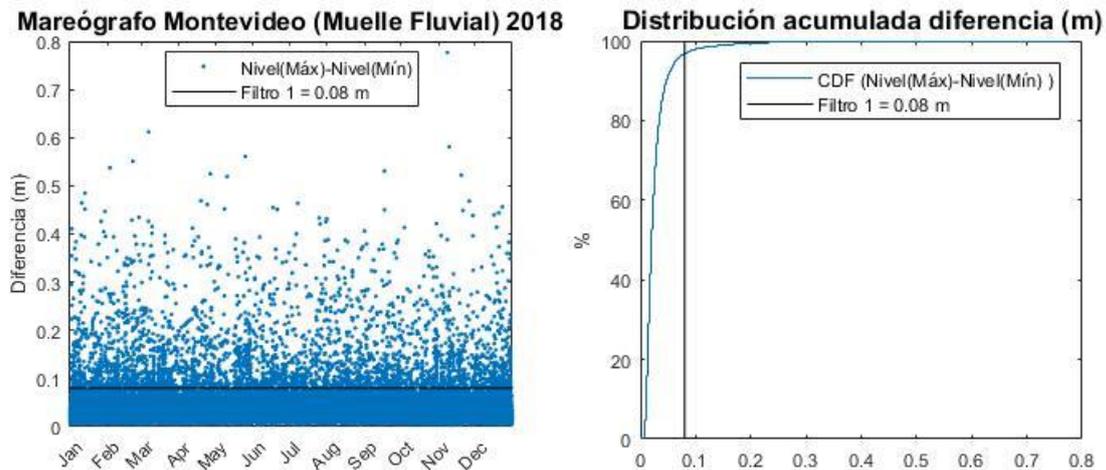


Figura 97 - Mareógrafo Montevideo (Muelle Fluvial) 2018. Diferencia entre el registro máximo y mínimo cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia entre el registro máximo y mínimo cada 5 minutos (panel derecho)

Filtro 2: Filtrado por apartamiento con el registro anterior

En la serie analizada se encontraron casos con un cambio muy brusco entre un valor y el siguiente. En la Figura 98 (panel izquierdo) se muestra la diferencia con el registro anterior, que para el período analizado alcanza los 0.34 m para el 9/10/2018 a las 9:15. En la Figura 99 se muestra este caso. Esto no hubiese sido depurado con el Filtro 1, ya que la diferencia entre el valor máximo y mínimo es pequeña. Se considera que este cambio brusco corresponde con una corrección de los datos hacia adelante. El objetivo de este filtro es eliminar los datos anteriores a un cambio brusco. Así se mantiene la corrección hacia adelante sin un salto irreal en los valores.

En la Figura 98 (panel derecho) se puede ver la distribución acumulada empírica de la diferencia con el valor anterior. Analizando la Figura 98 (panel izquierdo) se puede ver que para algunos eventos se dan varios casos seguidos de gran diferencia con el valor anterior y estos eventos no deben ser eliminados por el filtro. En la Figura 98 (panel derecho) se hace un zoom a la parte superior de la curva de forma de encontrar los eventos aislados que sí quieren ser eliminados. Se decide que el Filtro 2 eliminará los datos anteriores a un salto mayor a 0.15 m. Con este umbral, en el período analizado, se encuentran 14 casos. Se decide eliminar los datos correspondientes a la hora anterior al salto.

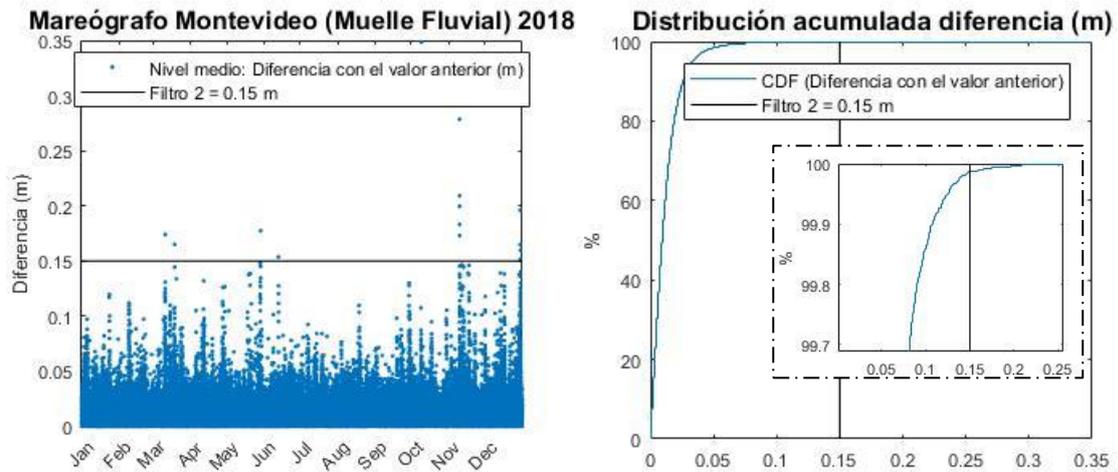


Figura 98 - Mareógrafo Montevideo (Muelle Fluvial) 2018. Diferencia con el valor anterior cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia con el valor anterior cada 5 minutos y zoom en la zona superior de la curva de distribución (panel derecho).

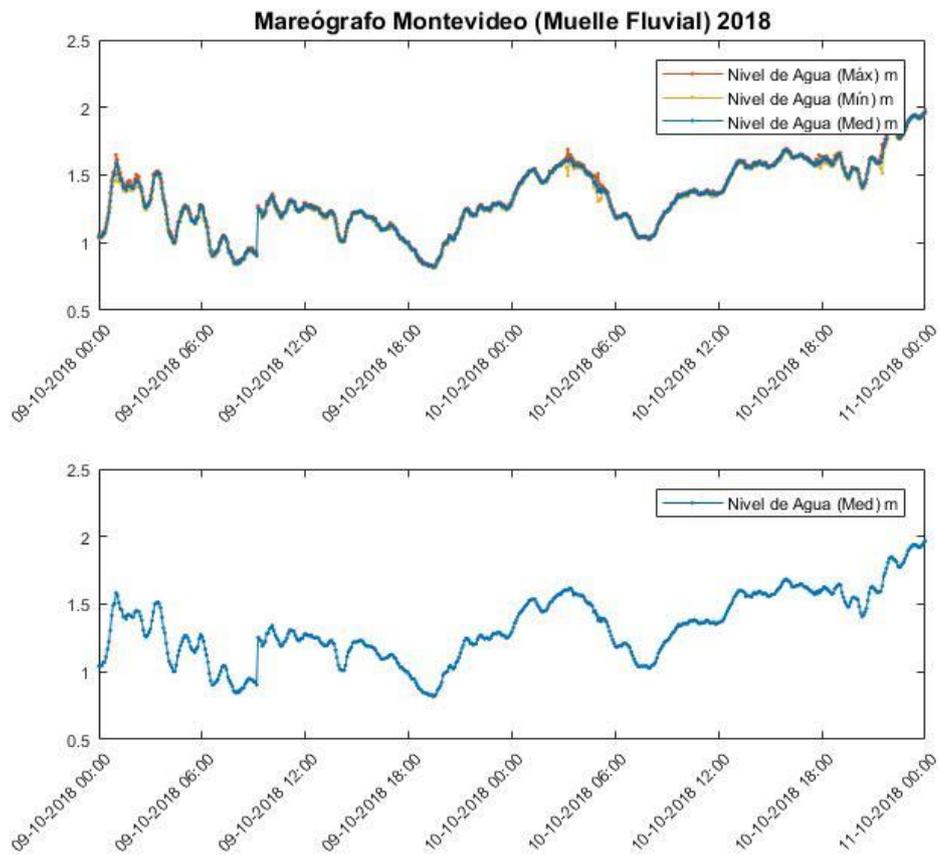


Figura 99 – Niveles mareógrafo Montevideo (Muelle Fluvial) 9-10-2018 a 11-10-2018. El 9-10-2018 a las entre las 6:00 y las 12:00 se produce un cambio muy brusco entre un valor y el siguiente (0.34 m).

Ejemplos de la aplicación de los filtros

En las siguientes figuras se muestran ejemplos de los datos que han sido depurados con los filtros para el mareógrafo de Montevideo. Así, el Filtro 1 se marca con una banda transparente negra y el Filtro 2 con una banda transparente roja.

Se observa que con la elección de estos dos filtros se conserva la mayoría de los datos. En el primer caso, el 9-10-2018 Figura 100, se eliminan los datos previos a un salto brusco en los registros. El caso de la Figura 101 muestra como el Filtro 2 se suma al Filtro 1 en un evento de gran variación en los datos.

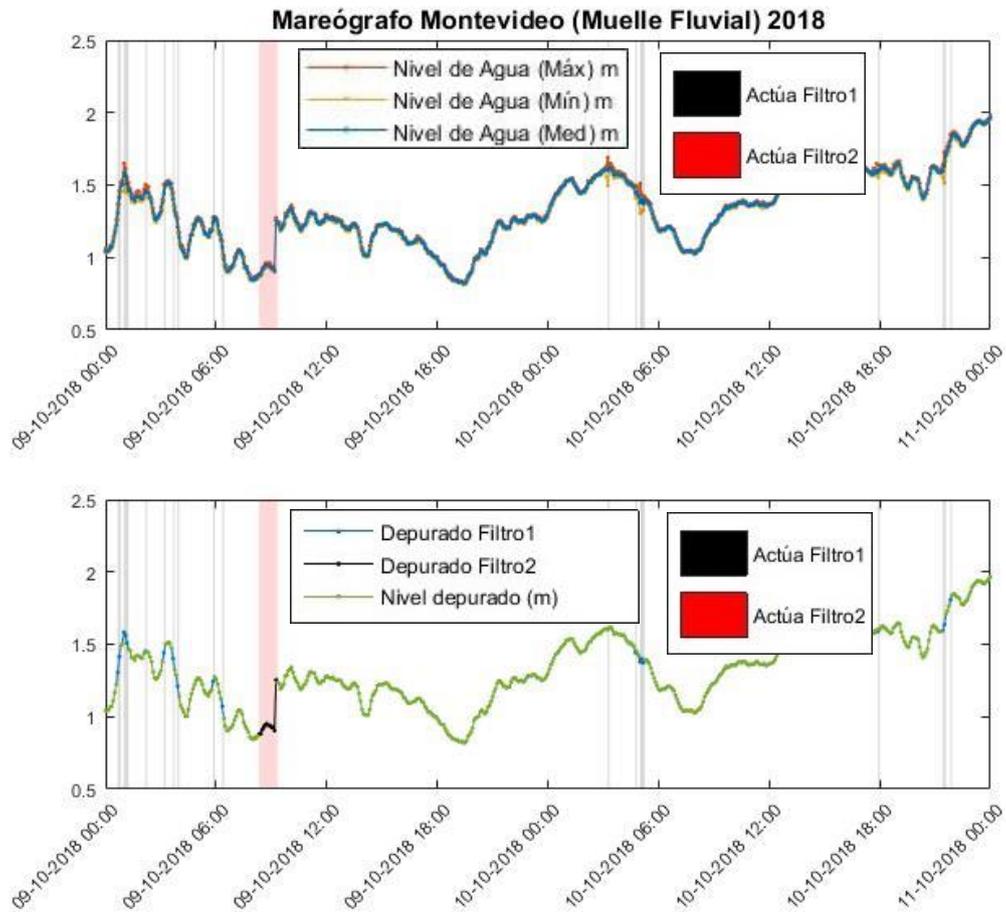


Figura 100 - Efecto filtrado de datos. Mareógrafo Montevideo (Muelle Fluvial) 9-10-2018 a 11-10-2018

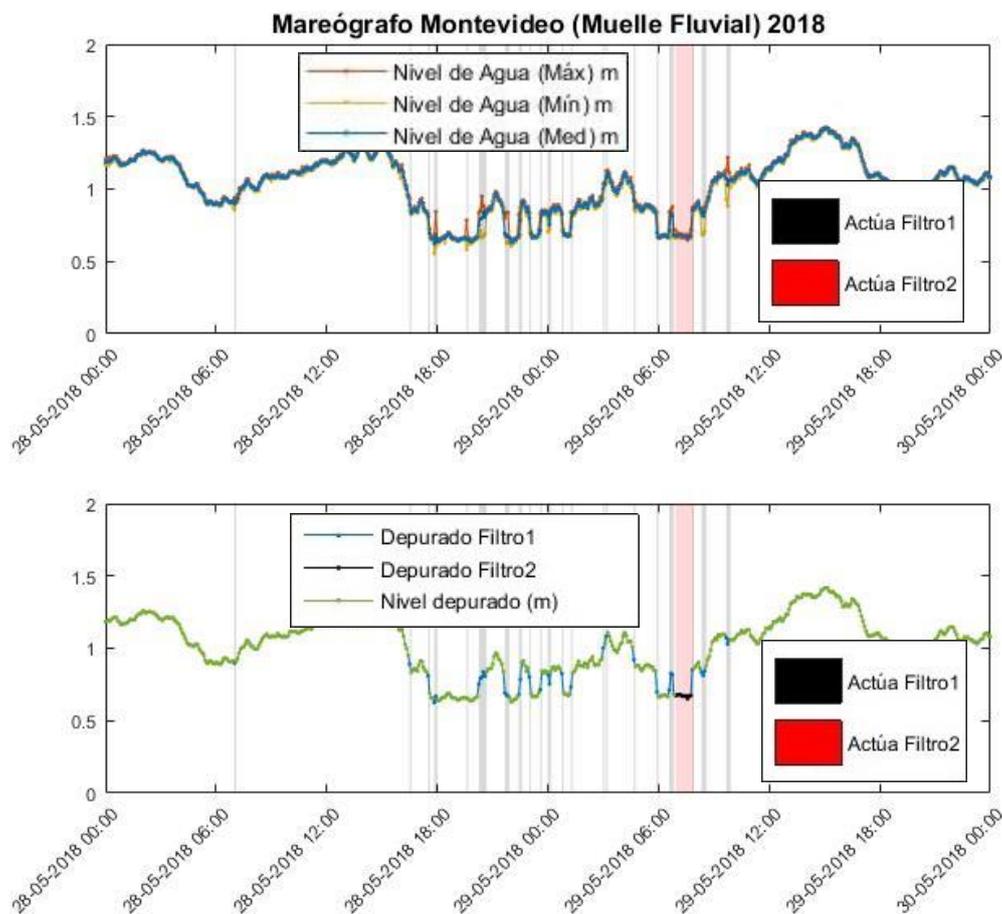


Figura 101 - Efecto filtrado de datos. Mareógrafo Montevideo (Muelle Fluvial) 28-05-2018 a 30-05-2018

X.2. Isla de Flores

Estos datos son facilitados por la Administración Nacional de Puertos (ANP), Uruguay, a través de su web: <http://mareografo.anp.com.uy/>. Con una frecuencia de un dato cada 5 minutos se puede acceder a las mediciones de Nivel de Agua Máximo, Nivel de Agua Medio y Nivel de Agua Mínimo, siendo los datos faltantes comunes a las 3 medidas. Para el período 1/1/2018 al 31/12/2018 se dan datos faltantes en 17 ocasiones, lo que es equivalente a un 0.02%. Estos datos corresponden a 5 intervalos que se presentan en la Tabla 19:

Intervalos con datos faltantes		
Inicio	Fin	Duración
20180306/17:00	20180306/17:00	5 min
20180325/02:00	20180325/02:55	60 min
20180430/09:05	20180430/09:10	10 min
20180430/13:45	20180430/13:45	5 min
20180906/12:00	20180906/12:00	5 min
	Total	1 hora 25 min

Tabla 19 – Intervalos con datos faltantes. Mareógrafo de Isla de Flores, 2018

A los datos del mareógrafo de Isla de Flores se aplican los mismos filtros que para el mareógrafo de Montevideo, descritos en el apartado anterior.

- Filtro 1: Filtrado por diferencia entre el registro máximo y mínimo. Toma como límite una diferencia de 0.08 m (Figura 102). Para el año 2018 se eliminan 6742 datos lo que corresponde con un 6.42%.
- Filtro 2: Filtrado por apartamiento con el registro anterior (Figura 103). Eliminará los datos de la hora anterior a un salto mayor a 0.15 m. Para el 2018 se dan 41 casos.

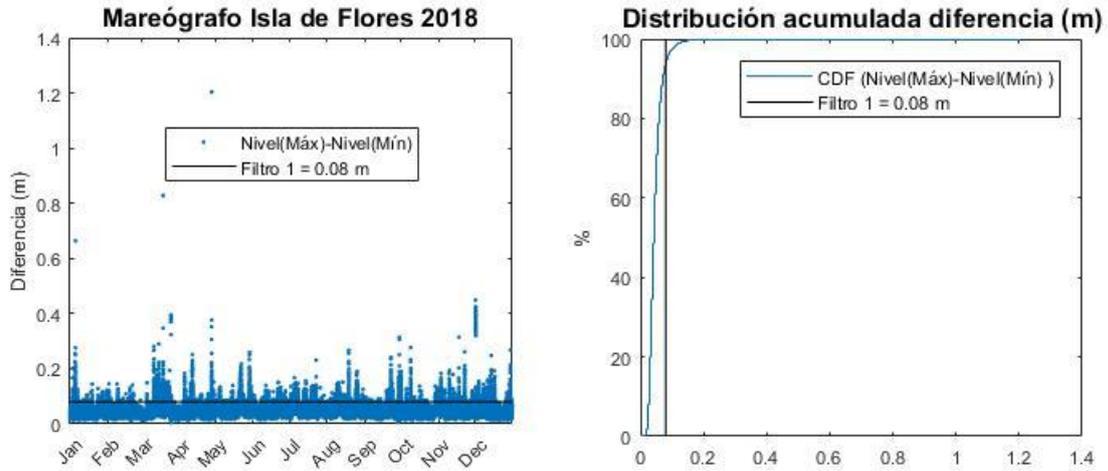


Figura 102 - Mareógrafo Isla de Flores 2018. Diferencia entre el registro máximo y mínimo cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia entre el registro máximo y mínimo cada 5 minutos (panel derecho)

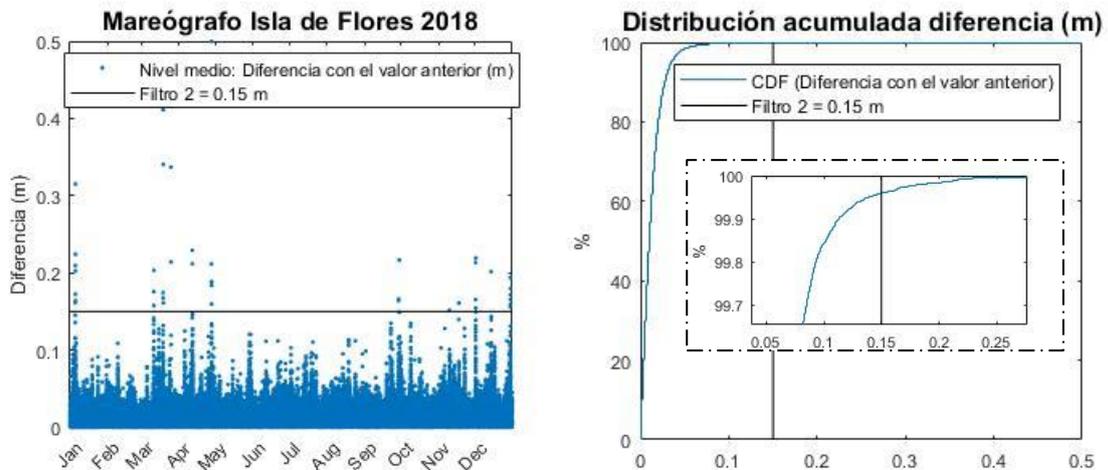


Figura 103 – Mareógrafo Isla de Flores 2018. Diferencia con el valor anterior cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia con el valor anterior cada 5 minutos y zoom en la zona superior de la curva de distribución (panel derecho).

Además, para los datos del mareógrafo de Isla de Flores se aprecia límites en el rango de medida. Como límite superior para el registro máximo se tomó 2.6 m y para el registro mínimo -0.2 m. Esto se ejemplifica respectivamente en las figuras Figura 104, 1/12/2018 de 16:00 a 21:30, y Figura 105, 29/9/2018 de 7:40 a 10:50. El Filtro 3 eliminará la hora central correspondiente a los datos fuera del rango de medida.

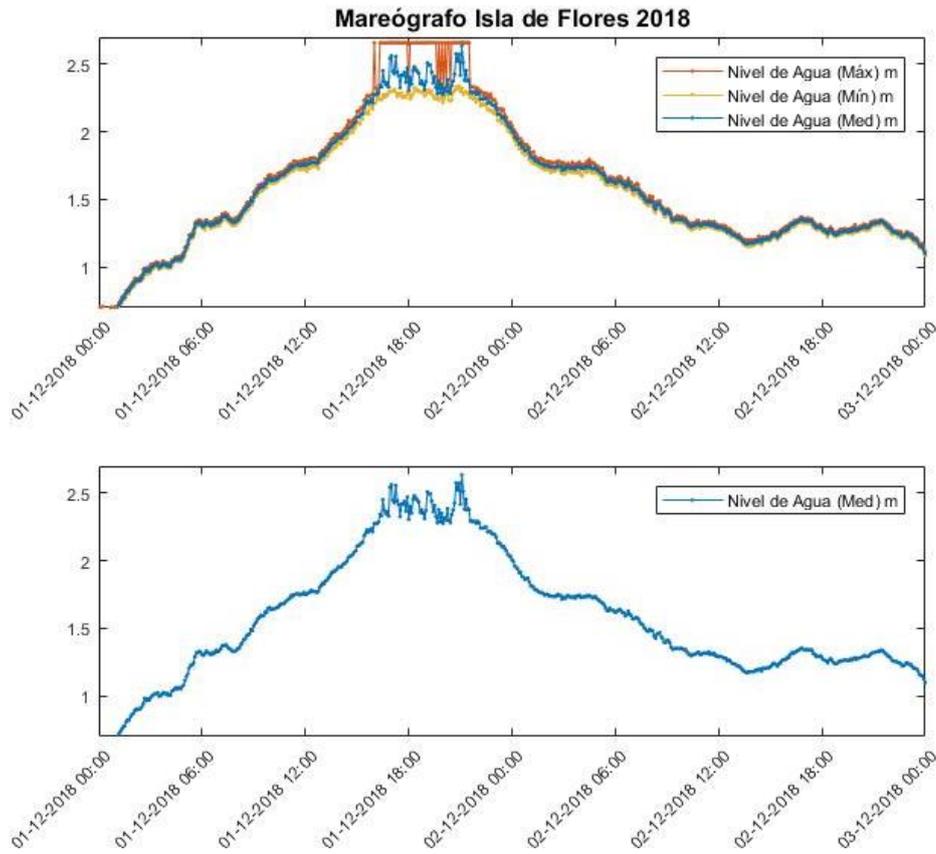


Figura 104 - Niveles mareógrafo Isla de Flores 1-12-2018 a 3-12-2018. El 01-12-2018 entre las 12:00 y las 24:00 la señal registrada máxima alcanza el límite superior del rango de medida del mareógrafo, 2.6 m.

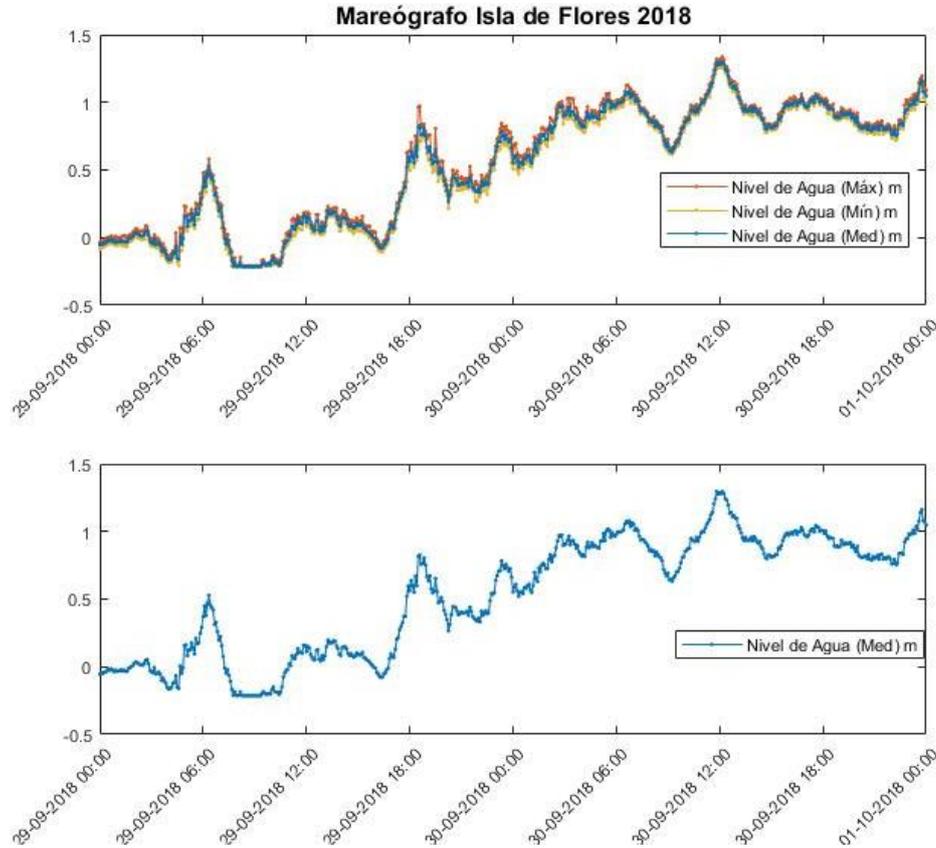


Figura 105 - Niveles mareógrafo Isla de Flores 29-9-2018 a 1-10-2018. El 29-09-2018 entre las 06:00 y las 12:00 la señal registrada mínima alcanza el límite inferior del rango de medida del mareógrafo, -0.2 m.

Ejemplos de la aplicación de los filtros

En las siguientes figuras se muestran ejemplos de los datos que han sido depurados con los filtros para los registros del mareógrafo de Isla de Flores. Así, el Filtro 1 se marca con una banda transparente negra, el Filtro 2 con una banda transparente roja y el Filtro 3 con una banda transparente amarilla.

Se puede ver que con la elección de estos tres filtros se mantiene casi toda la serie de datos. En el primer caso, el 27-04-2018 (Figura 106), se eliminan los datos previos a un salto brusco en los registros. El caso de la Figura 107 muestra como el Filtro 3 se hace necesario para eliminar los datos fuera del rango de medida del mareógrafo.

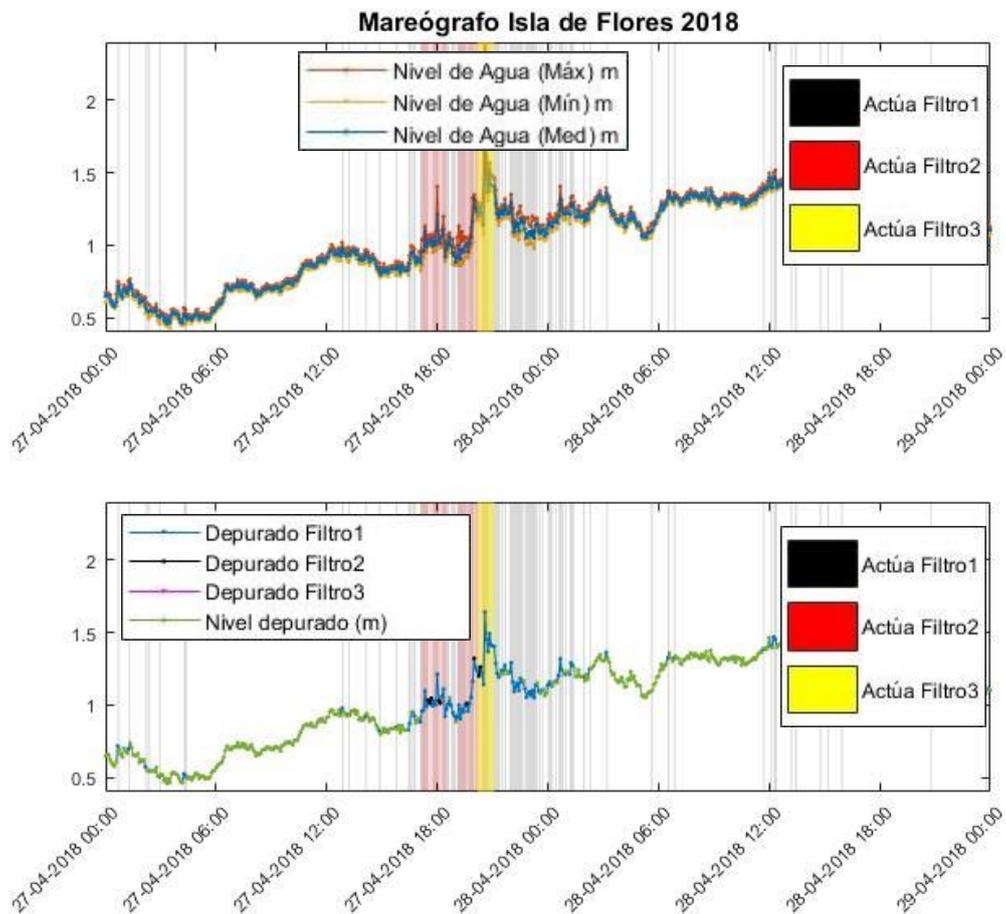


Figura 106 - Efecto filtrado de datos. Mareógrafo Isla de Flores 27-04-2018 a 29-04-2018

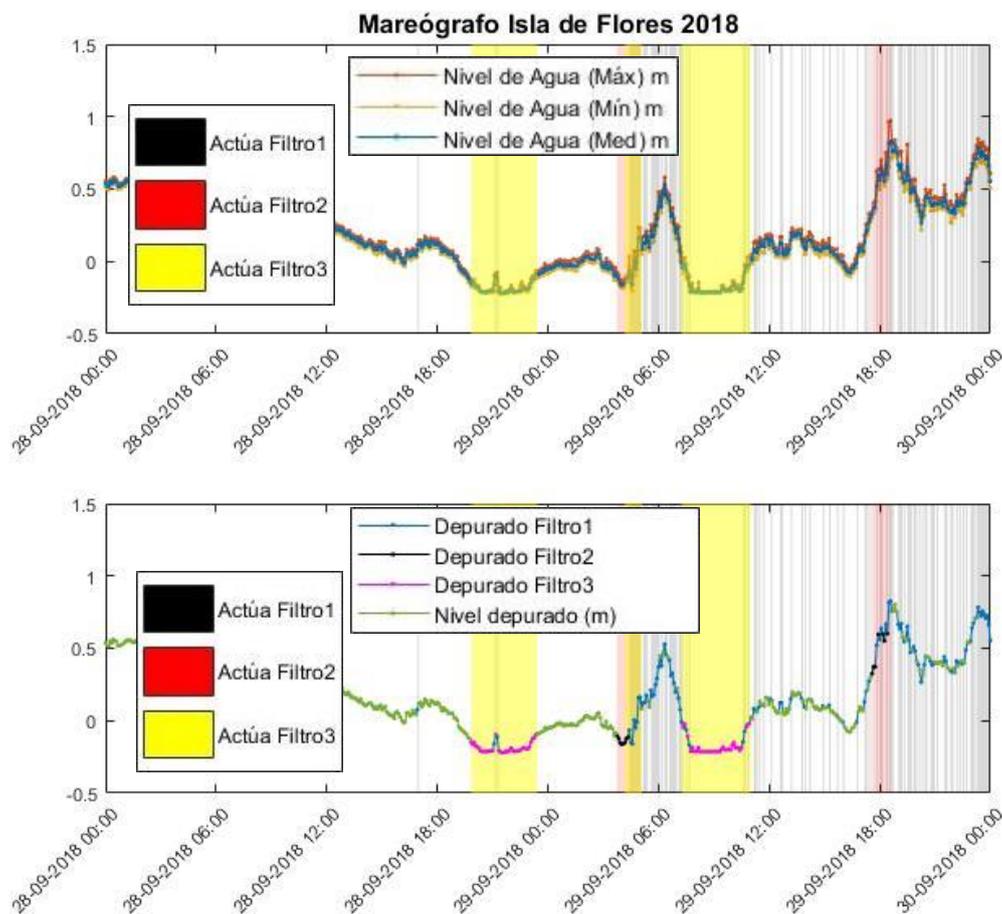


Figura 107 - Efecto filtrado de datos. Mareógrafo Isla de Flores 28-09-2018 a 30-09-2018

X.3. La Paloma

Estos datos son facilitados por la Administración Nacional de Puertos (ANP), Uruguay, a través de su web: <http://mareografo.anp.com.uy/>. Con una frecuencia de un dato cada 5 minutos se puede acceder a las mediciones de Nivel de Agua Máximo, Nivel de Agua Medio y Nivel de Agua Mínimo, siendo los datos faltantes comunes a las 3 medidas. Para el período 1/1/2018 al 31/12/2018 se dan datos faltantes en 34406 ocasiones, lo que es equivalente a un 32.73%. Estos datos corresponden a 6 intervalos que se presentan en la Tabla 20:

Intervalos con datos faltantes		
Inicio	Fin	Duración
20180101/00:00	20180414/12:20	103 días, 12 horas, 25 min
20180602/10:25	20180602/10:25	5 min
20180731/16:00	20180816/11:35	15 días, 19 horas, 40 min
20181008/08:45	20181008/09:40	60 min
20181016/08:55	20181016/09:50	60 min
20181022/08:40	20181022/09:35	60 min
	Total	119 días, 11 horas, 10 min

Tabla 20 – Intervalos con datos faltantes. Mareógrafo de La Paloma, 2018

A los datos del mareógrafo de La Paloma se aplican estos filtros que se desarrollan en los siguientes apartados.

- Filtro 1: Filtrado por diferencia entre el registro máximo y mínimo. Toma como límite una diferencia de 0.20 m. Para el año 2018 se eliminan 3794 datos lo que corresponde con un 5.39%.
- Filtro 2: Filtrado por apartamiento con el registro anterior. Eliminará los datos de la hora anterior a un salto mayor a 0.4 m. Para el 2018 se dan 2 casos.
- Filtro 3: Filtrado por datos fuera del rango de medida. Se eliminarán los datos de la hora central cuando el registro máximo sea mayor de 2 m y/o cuando el registro mínimo sea menor de 0 m.
- Filtro 4: Filtrado por cercanía entre el registro máximo y mínimo. Para este mareógrafo se vio la necesidad de añadir este filtro. Pues el Filtro 3 depura correctamente los datos altos fuera del rango de medida, pero se vio que no es así para los datos bajos fuera del rango de medida, debido a que el valor mínimo que es capaz de registrar el mareógrafo varía con el tiempo. El Filtro 4 sí consigue depurar este tipo de datos.

Además, en este período se encontraron valores extremos previamente a que el mareógrafo dejara de funcionar. Estos datos se eliminarían con el Filtro 3, pero para este análisis fueron eliminados de antemano. Se trata de valores por encima de 5 m o por debajo de -5 m y totalizan 317 datos.

Filtro 1: Filtrado por diferencia entre el registro máximo y mínimo

Para el período 1/1/2018 al 31/12/2018 la diferencia entre el registro máximo y mínimo se muestra en la Figura 108 (panel izquierdo) y la distribución acumulada de las diferencias entre el registro máximo y mínimo se muestran en la Figura 108 (panel derecho). Se eligió como límite una diferencia de 0.20 m. Así, los valores de Nivel de Agua Medio correspondientes con una diferencia registrada mayor serán depurados. Este filtro elimina 3794 datos lo que corresponde con un 5.39%.

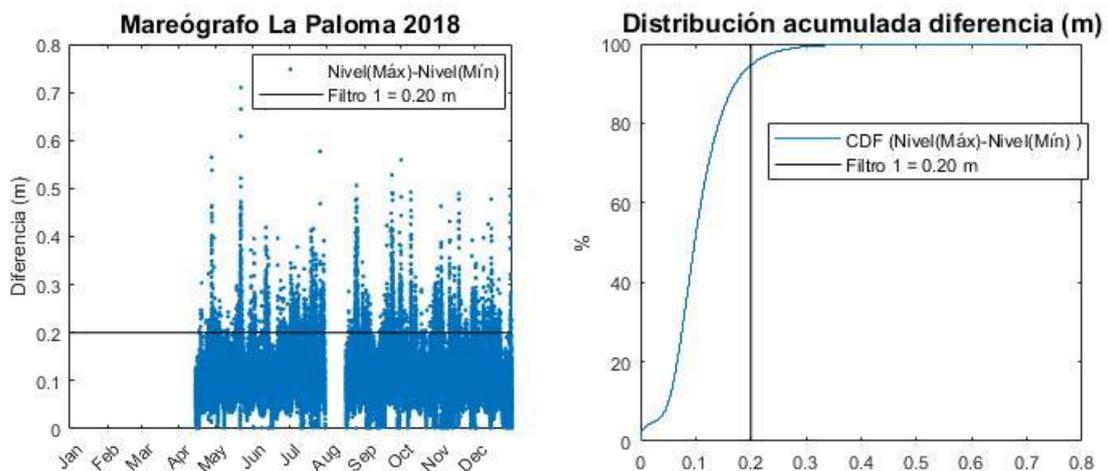


Figura 108 - Mareógrafo La Paloma 2018. Diferencia entre el registro máximo y mínimo cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia entre el registro máximo y mínimo cada 5 minutos (panel derecho)

Filtro 2: Filtrado por apartamiento con el registro anterior

En la Figura 109 (panel izquierdo) se muestra la diferencia con el registro anterior y en la Figura 109 (panel derecho) se puede ver la distribución acumulada. El Filtro 2 eliminará los datos de la hora anterior a un salto mayor a 0.4 m. Para el 2018 se dan 2 casos.

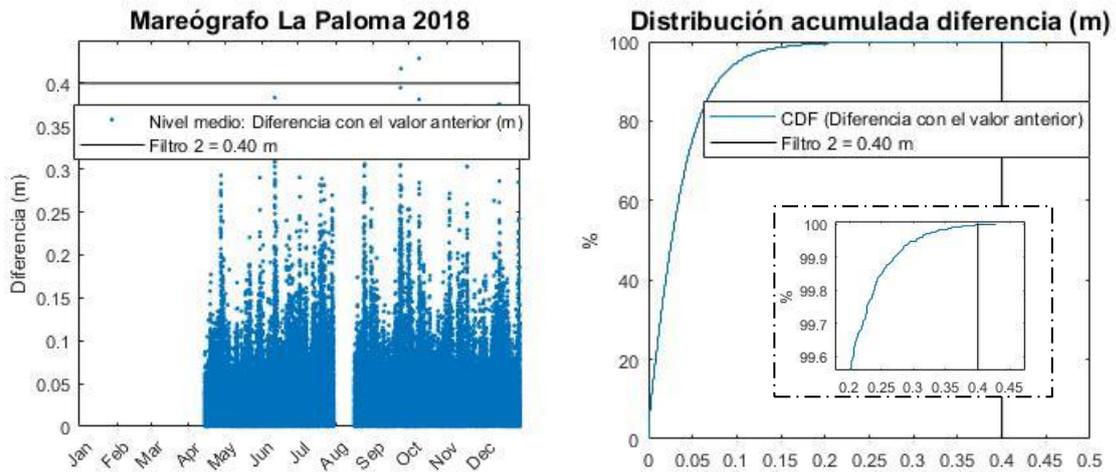


Figura 109 – Mareógrafo La Paloma 2018. Diferencia con el valor anterior cada 5 minutos (panel izquierdo). Distribución acumulada para la diferencia con el valor anterior cada 5 minutos y zoom en la zona superior de la curva de distribución (panel derecho).

Filtro 3: Filtrado por datos fuera del rango de medida

Cómo límite superior para el registro máximo se tomó 2.0 m, siendo la Figura 110, 2/12/2018 de 00:15 a 04:00, un ejemplo. Se vio que el rango de medida para el registro mínimo varía a lo largo del tiempo. Así, en la Figura 111, del 11/07/2018 a las 18:00 al 12/07/2018 a las 18:00 el dato mínimo registrado toma un valor de 0.44 m y en la Figura 112, el 27/12/2018 el dato mínimo registrado toma un valor de 0.70 m. Esto se depurará con el siguiente Filtro 4.

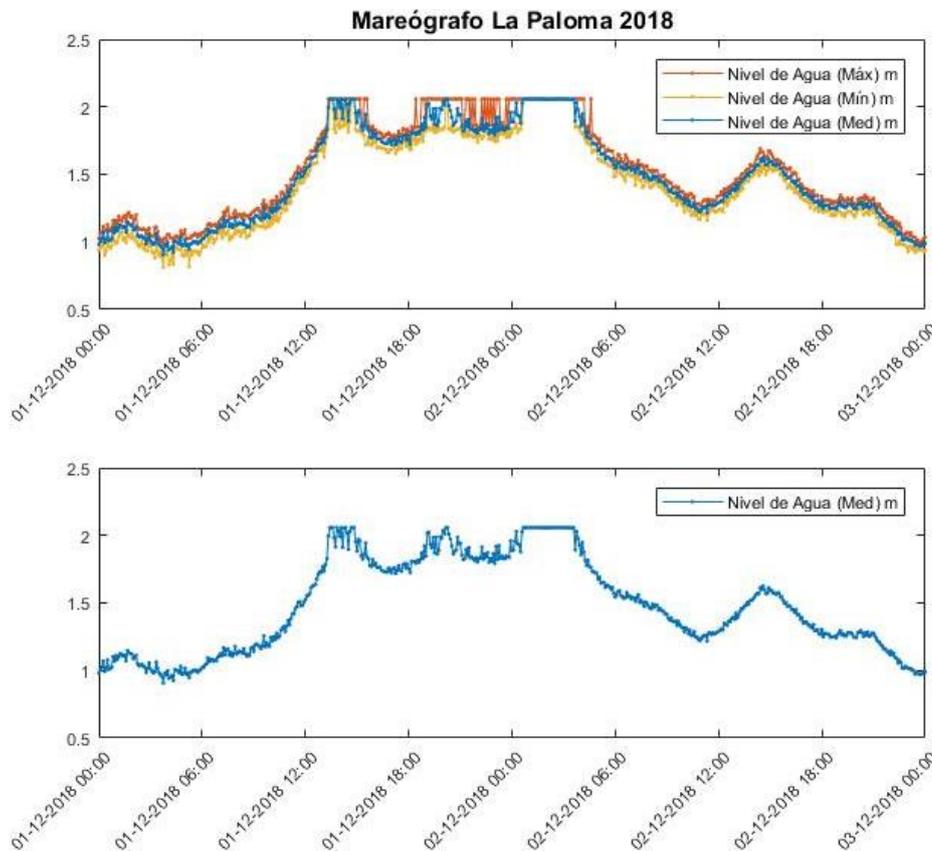


Figura 110 - Niveles mareógrafo La Paloma 01-12-2018 a 03-12-2018. A partir de las 12:00 del 01-12-2020 la señal registrada máxima alcanza el límite superior del rango de medida del mareógrafo, 2.0 m.

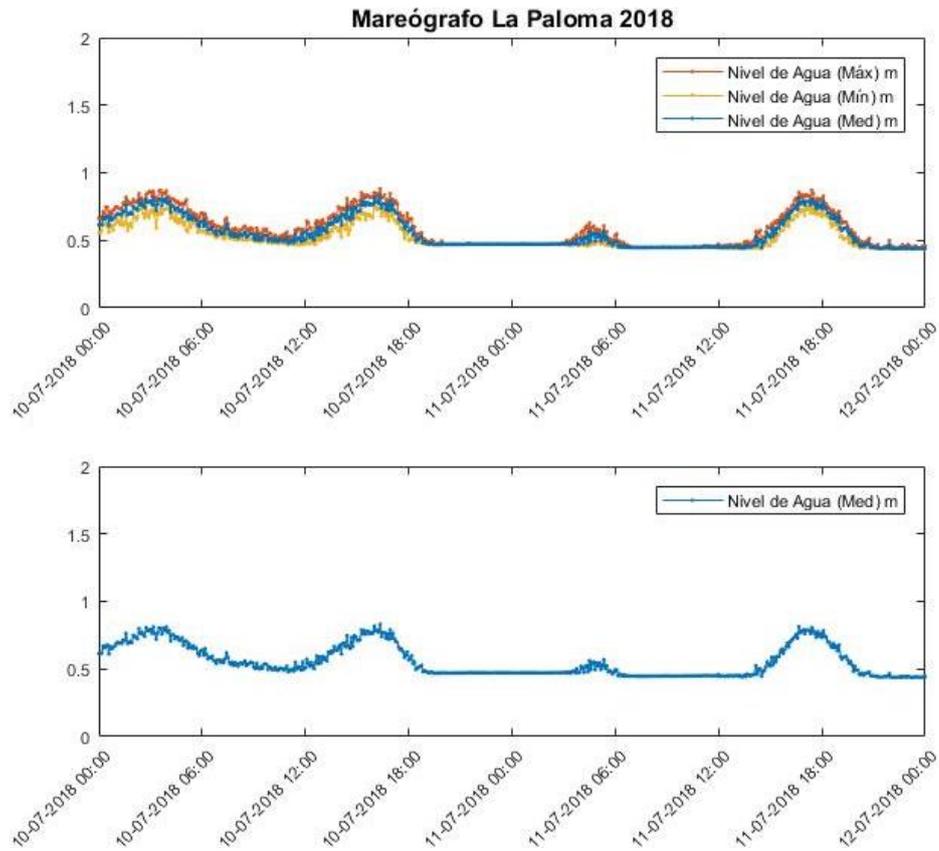


Figura 111 - Niveles mareógrafo La Paloma 10-07-2018 a 12-07-2018. A partir del 10-07-2018 a las 18:00 la señal registrada mínima alcanza un límite inferior de 0.44 m.

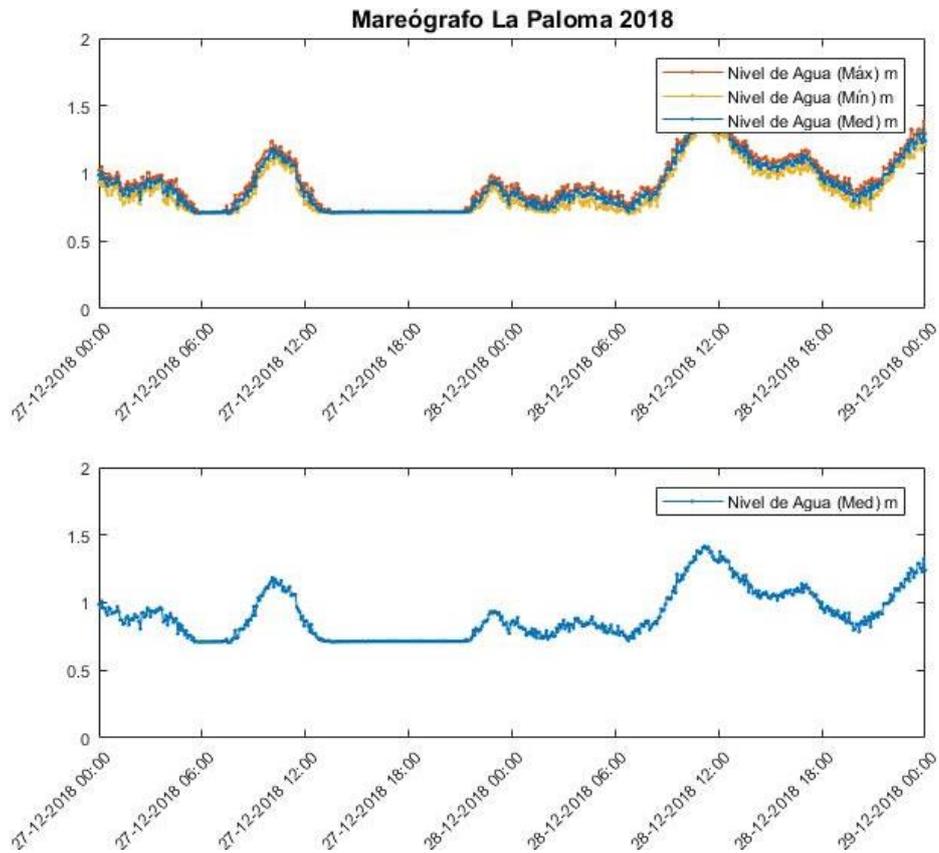


Figura 112 - Niveles mareógrafo La Paloma 27-12-2018 a 29-12-2018. A partir del 10-07-2018 a las 18:00 la señal registrada mínima alcanza un límite inferior de 0.70 m.

Para evitar valores extremos, que se vio que aparecen previo a que el mareógrafo deje de tomar registros, el Filtro 3 también eliminará los valores cuando el registro mínimo sea menor a 0 m.

Filtro 4: Filtrado por cercanía entre el registro máximo y mínimo

En la Figura 108, se vio que la distribución acumulada de la diferencia entre el registro máximo y mínimo tiene una zona inicial vertical: datos con baja diferencia entre el registro máximo y mínimo. Se comprobó que al depurar los datos que presentaban una baja diferencia entre el registro máximo y mínimo se depuraban los datos inferiores fuera del rango de medida ejemplificados en la Figura 111 y Figura 112. Se eligió depurar la hora central a los datos que presentaran una diferencia entre el registro máximo y el mínimo inferior a 0.008 m. En la Figura 113, se muestran los valores del registro mínimo que activan el Filtro 4, y en el siguiente apartado se dan algunos ejemplos.

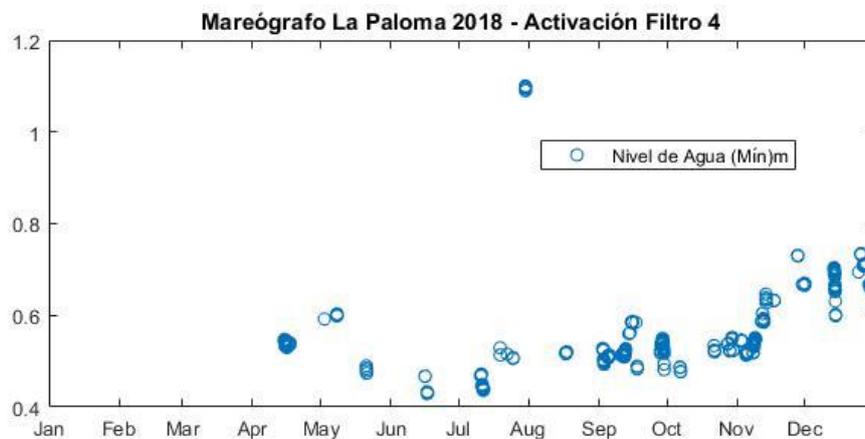


Figura 113 - Diferencia entre el registro máximo y el mínimo inferior a 0.008 m. Mareógrafo La Paloma 2018

Ejemplos de la aplicación de los filtros

En las siguientes figuras se muestran ejemplos de los datos que han sido depurados con los filtros para el mareógrafo de La Paloma. Así, el Filtro 1 se marca con una banda transparente negra, el Filtro 2 con una banda transparente roja, el Filtro 3 con una banda transparente amarilla y el Filtro 4 con una banda transparente rosa.

Se puede ver que con la elección de estos cuatro filtros se mantiene casi toda la serie de datos. En el primer caso, Figura 114, el 13/06/2018 el Filtro 3 se suma al Filtro 1 para depurar datos fuera de rango; el efecto de la actuación del Filtro 4 es nulo, sin embargo, la necesidad del Filtro 4 se ejemplifica en los casos de la Figura 115 del 16/06/2018 y del 17/06/2018. La Figura 116 muestra como el Filtro 4 logra depurar datos fuera del rango de medida del mareógrafo que fueron registrados como un período sin variación en el nivel.

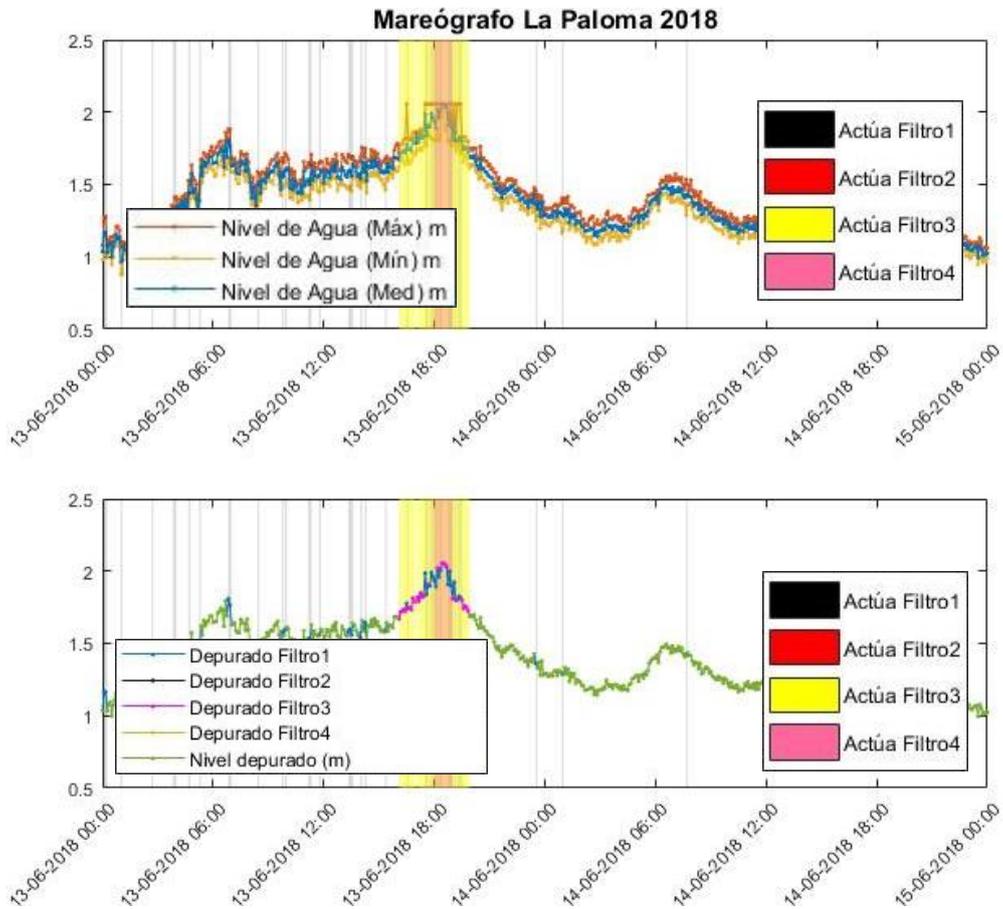


Figura 114 - Efecto filtrado de datos. Mareógrafo La Paloma 16-06-2018 a 15-06-2018

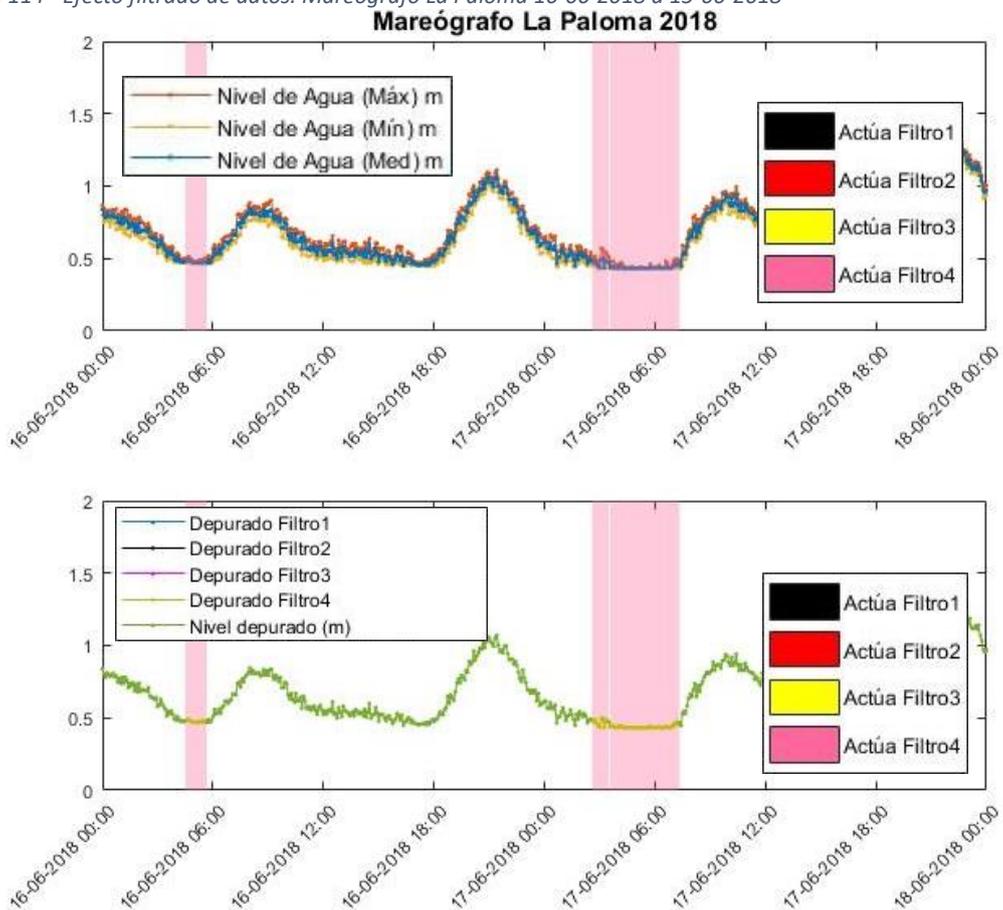


Figura 115 - Efecto filtrado de datos. Mareógrafo La Paloma 16-06-2018 a 18-06-2018

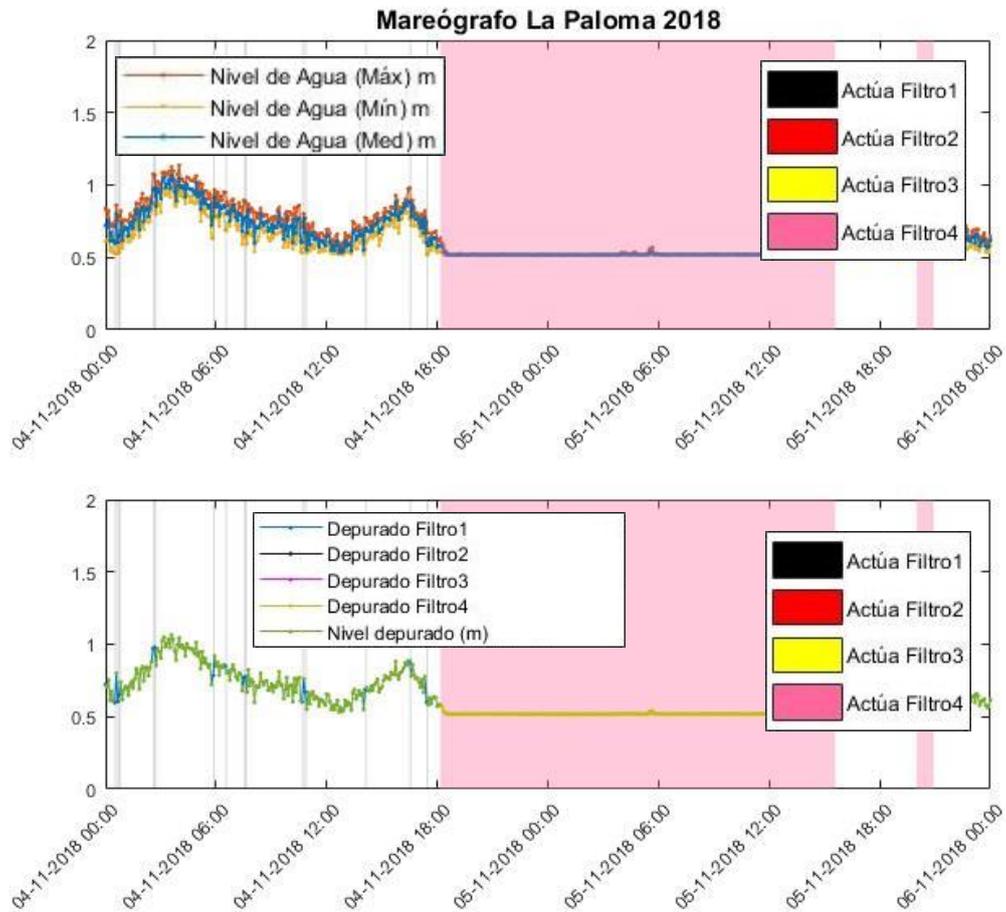


Figura 116 - Efecto filtrado de datos. Mareógrafo La Paloma 04-11-2018 a 06-11-2018

X.4. Mareógrafos Argentinos

Estos datos se obtienen de la web del Servicio de Hidrografía Naval (SHN), Argentina, <http://www.hidro.gob.ar/oceanografia/alturashorarias.asp>. Con una frecuencia de un dato cada 1 hora se publican los datos del nivel a la hh:45 (09:45 10:45 11:45...). Para el período 1/1/2018 al 31/12/2018 se presentan en las tablas siguientes los períodos con datos faltantes. La Tabla 21 es un resumen de los porcentajes por estación. En la Tabla 22 se da la distribución en duraciones de intervalos con datos faltantes para cada estación. En las tablas desde la Tabla 23 hasta la Tabla 31 se especifican los intervalos de duración mayor a 10 horas.

Estación	Cantidad de datos faltantes	% de datos faltantes
Atalaya	3001	34.26%
Buenos Aires	1109	12.66%
La Plata	1252	14.29%
Mar del Plata	1218	13.90%
Torre Oyarvide	1474	16.83%
Pilote Norden	2391	27.29%
San Clemente	1979	22.59%
San Fernando	6695	76.43%
Santa Teresita	8760	100%

Tabla 21 – Cantidad de datos faltantes. Mareógrafos Argentina, 2018

Duración	Cantidad			
	Atalaya	Buenos Aires	La Plata	Mar del Plata
1 hora	66	82	119	124
2 horas	11	6	29	20
3 horas	2	4	11	7
4 horas	7	6	7	6
5 horas	8	7	10	8
de 6 a 10 horas	7	11	14	12
de 10 horas a 1 día	5	4	5	6
de 1 a 14 días	5	2	1	1
mayor a 14 días	5	1	1	1
Total	125 días, 1 hora	46 días, 5 horas	52d., 4horas	52d., 18horas
	T. Oyarvide	Pilote Norden	S. Clemente	S. Fernando
1 hora	107	153	253	42
2 horas	18	31	58	5
3 horas	8	26	18	1
4 horas	10	8	23	0
5 horas	11	13	7	1
de 6 a 10 horas	15	37	18	6
de 10 horas a 1 día	7	8	15	1
de 1 a 14 días	2	5	5	1
mayor a 14 días	1	2	1	1
Total	61d., 10horas	99d., 6horas	82d., 11 h.	278d., 23 h.

Tabla 22 – Distribución de la duración de los intervalos con datos faltantes. Mareógrafos Argentina, 2018

Atalaya. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180327/06:45	20180328/04:45	23 horas
20180312/02:45	20180312/21:45	20 horas
20180401/12:45	20180403/04:45	1 día, 17 horas
20180701/12:45	20180702/03:45	16 horas
20180710/11:45	20180711/06:45	20 horas
20180712/17:45	20180713/04:45	12 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181010/20:45	20181016/22:45	6 días, 3 horas
20181017/00:45	20181103/15:45	17 días, 16 horas
20181103/17:45	20181109/19:45	6 días, 3 horas
20181109/21:45	20181110/23:45	1 día, 3 horas
20181111/01:45	20181112/12:45	1 día, 12 horas
20181114/13:45	20181202/02:45	17 días, 14 horas
20181202/08:45	20181217/00:45	14 días, 17 horas
20181217/02:45	20181231/23:45	14 días, 22 horas

Tabla 23 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Atalaya, 2018

Buenos Aires. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180312/11:45	20180312/21:45	11 horas
20180326/23:45	20180328/04:45	1 día, 6 horas
20180401/12:45	20180403/04:45	1 día, 17 horas
20180701/12:45	20180702/03:45	16 horas
20180712/17:45	20180713/04:45	12 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181130/09:45	20181130/19:45	11 horas

Tabla 24 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Buenos Aires, 2018

La Plata. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180312/10:45	20180313/00:45	15 horas
20180327/06:45	20180328/04:45	23 horas
20180401/12:45	20180403/04:45	1 día, 17 horas
20180701/12:45	20180702/03:45	16 horas
20180712/17:45	20180713/04:45	12 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181130/09:45	20181130/19:45	11 horas

Tabla 25 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo La Plata, 2018

Mar del Plata. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180127/13:45	20180128/11:45	23 horas
20180312/10:45	20180312/21:45	12 horas
20180327/06:45	20180328/05:45	1 día, 0 horas
20180401/12:45	20180403/04:45	1 día, 17 horas
20180701/12:45	20180702/03:45	16 horas
20180712/17:45	20180713/05:45	13 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181130/09:45	20181130/19:45	11 horas

Tabla 26 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Mar del Plata, 2018

Torre Oyarvide. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180312/11:45	20180312/21:45	11 horas
20180325/12:45	20180403/11:45	9 días, 0 horas
20180625/21:45	20180626/07:45	11 horas
20180627/05:45	20180628/00:45	20 horas
20180630/12:45	20180702/08:45	1 día, 21 horas
20180702/10:45	20180703/07:45	22 horas
20180712/17:45	20180713/04:45	12 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181030/23:45	20181031/09:45	11 horas
20181130/09:45	20181130/19:45	11 horas

Tabla 27 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Torre Oyarvide, 2018

Pilote Norden. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180101/00:45	20180109/10:45	8 días, 11 horas
20180208/21:45	20180209/07:45	11 horas
20180304/02:45	20180304/22:45	21 horas
20180305/06:45	20180306/07:45	1 día, 2 horas
20180312/10:45	20180312/21:45	12 horas
20180327/00:45	20180328/04:45	1 día, 5 horas
20180401/10:45	20180403/23:45	2 días, 14 horas
20180412/15:45	20180413/15:45	1 día, 1 hora
20180502/09:45	20180502/23:45	15 horas
20180701/12:45	20180702/03:45	16 horas
20180712/17:45	20180713/04:45	12 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181111/17:45	20181112/12:45	20 horas
20181130/09:45	20181130/19:45	11 horas
20181210/05:45	20181231/23:45	21 días, 19 horas

Tabla 28 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Pilote Norden, 2018

San Clemente. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180104/21:45	20180105/08:45	12 horas
20180206/10:45	20180209/18:45	3 días, 9 horas
20180213/19:45	20180214/09:45	15 horas
20180312/11:45	20180312/22:45	12 horas
20180327/06:45	20180328/04:45	23 horas
20180401/12:45	20180403/04:45	1 día, 17 horas
20180411/17:45	20180412/09:45	17 horas
20180514/16:45	20180515/09:45	18 horas
20180616/19:45	20180618/10:45	1 día, 16 horas
20180701/12:45	20180702/03:45	16 horas
20180710/18:45	20180711/07:45	14 horas
20180712/17:45	20180713/04:45	12 horas
20180814/17:45	20180815/09:45	17 horas
20180907/06:45	20181008/12:45	31 días, 7 horas
20181020/21:45	20181022/10:45	1 día, 14 horas
20181029/14:45	20181030/09:45	20 horas
20181107/19:45	20181112/18:45	5 días, 0 horas
20181113/21:45	20181114/08:45	12 horas
20181118/07:45	20181118/19:45	13 horas
20181130/09:45	20181130/19:45	11 horas
20181211/20:45	20181212/08:45	13 horas

Tabla 29 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo San Clemente, 2018

San Fernando. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180109/11:45	20181008/12:45	272 días, 2 horas
20181114/10:45	20181116/07:45	1 día, 22 horas
20181130/09:45	20181130/19:45	11 horas

Tabla 30 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo San Fernando, 2018

Santa Teresita. Intervalos con datos faltantes de duración mayor a 10 horas		
Inicio	Fin	Duración
20180101/00:45	20181231/23:45	365 días

Tabla 31 – Intervalos con datos faltantes de duración mayor a 10 horas. Mareógrafo Santa Teresita, 2018

Filtro 1: Filtrado de *outliers*

El único filtro que se realiza a los datos registrados por los mareógrafos argentinos consiste en eliminar los valores atípicos (*outliers*). Se eliminan los datos mayores a 6 m y los menores a -3 m. En la siguiente Figura 117 se muestran los niveles del año 2018 antes y después de aplicar el filtro.

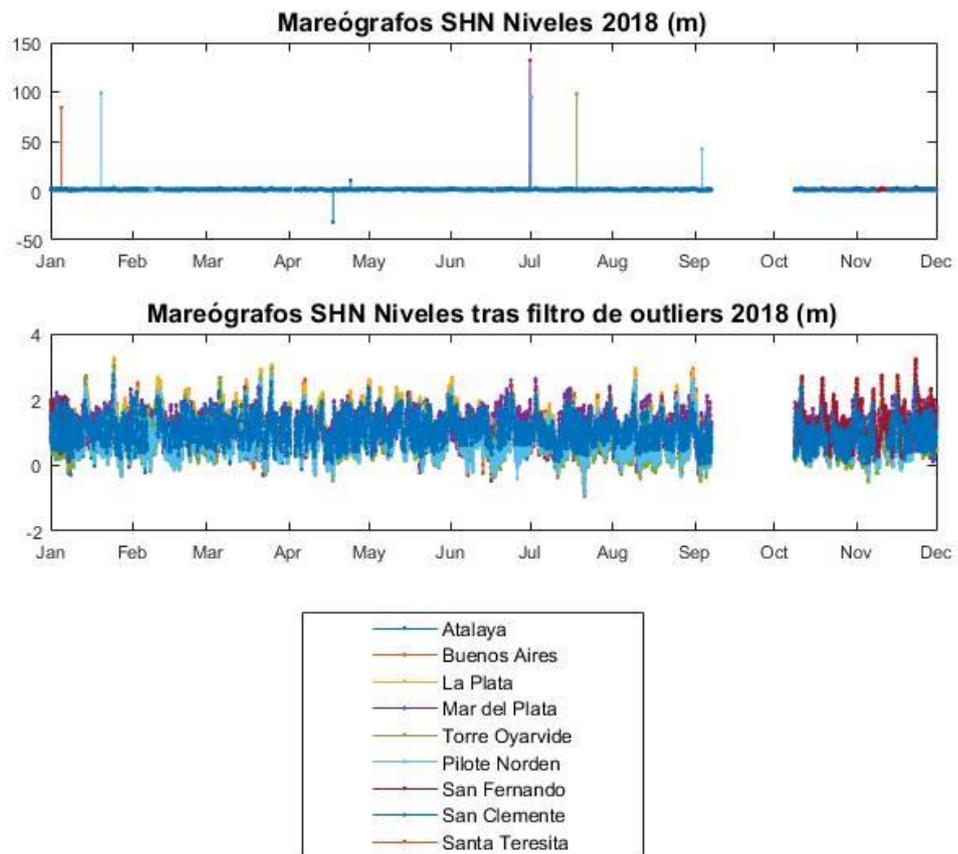


Figura 117 - Niveles mareógrafos argentinos 2018. Antes y después de la aplicación del filtro de outliers.

XI. Anexo 4

Estadísticos para el año 2019. Estadísticos sin media anual para el año 2018 y 2019.

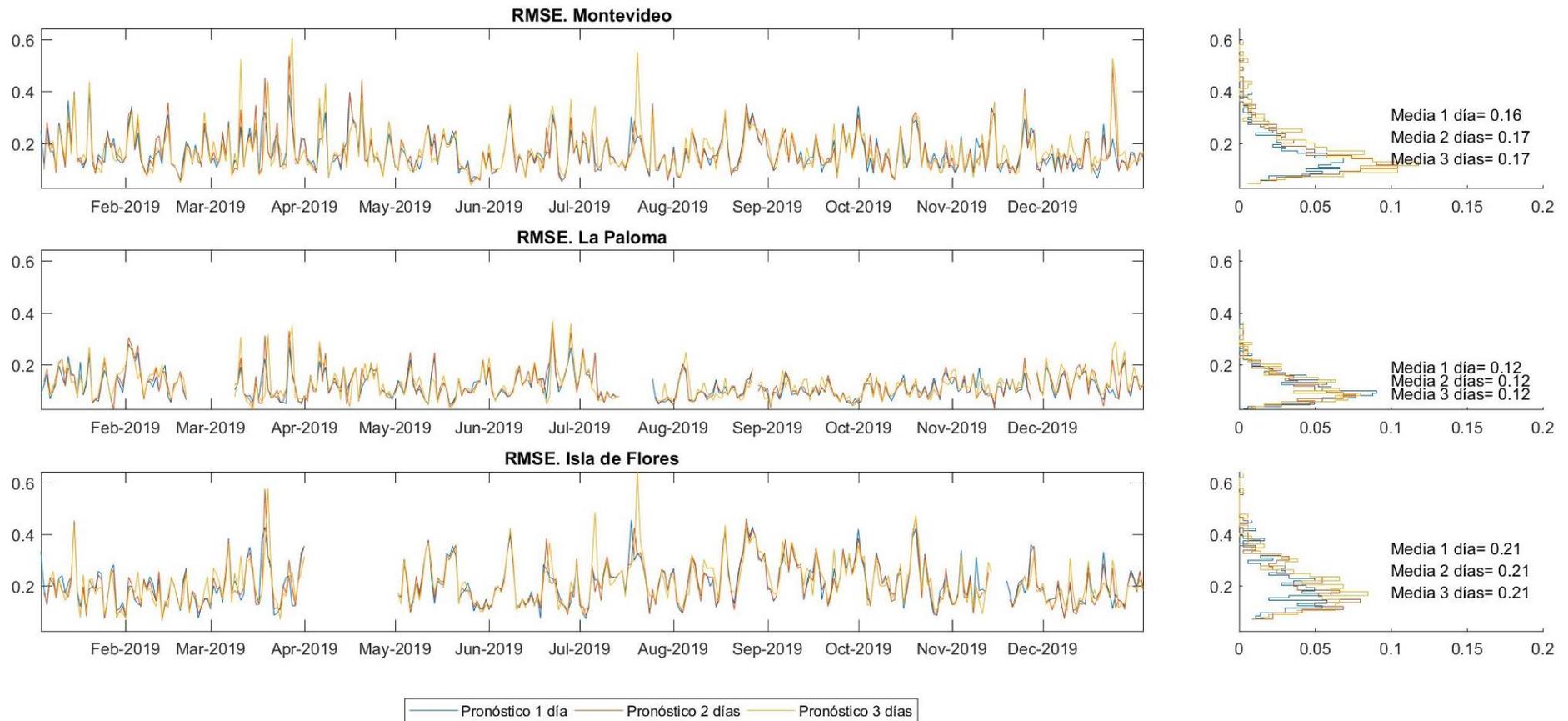


Figura 118 - Error cuadrático medio RMSE [m] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

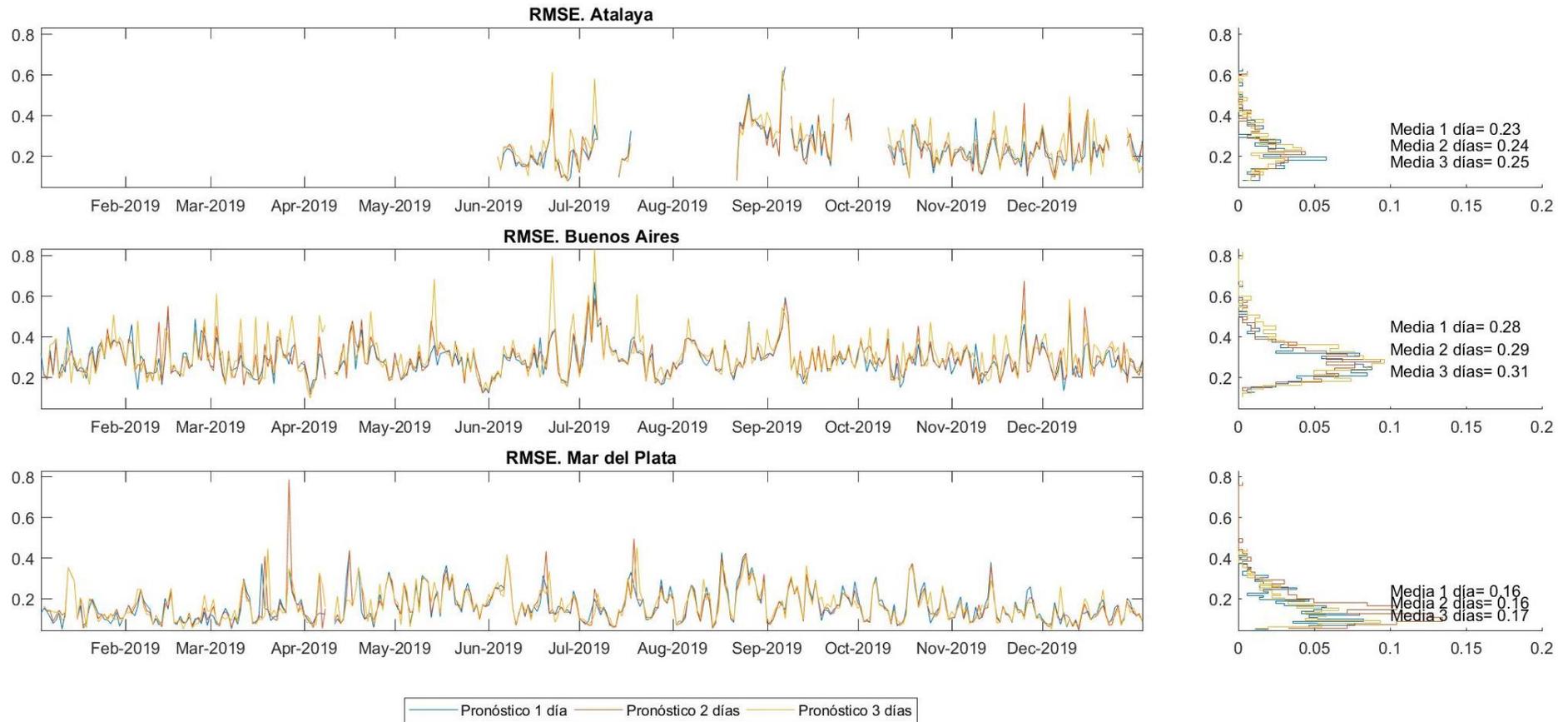


Figura 119 - Error cuadrático medio RMSE [m] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

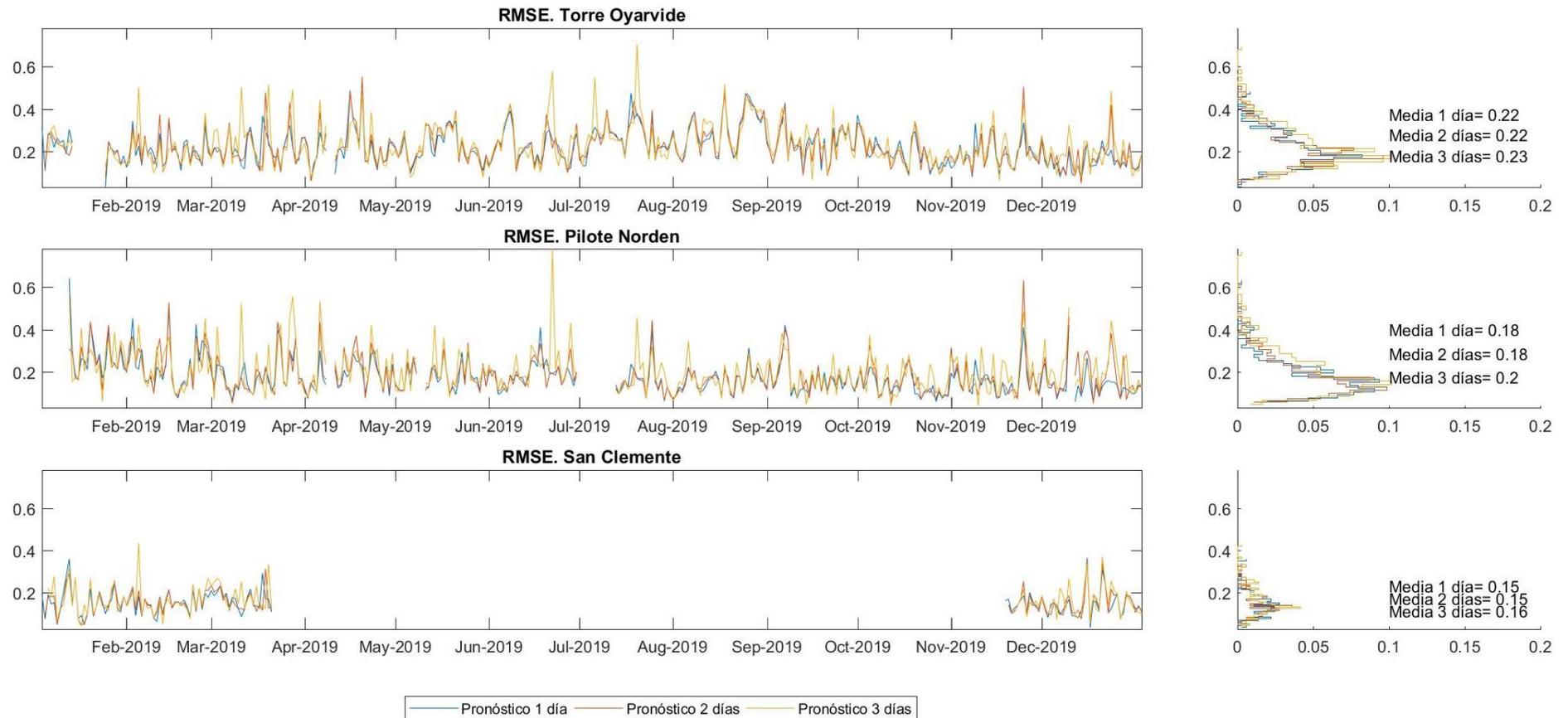


Figura 120 - Error cuadrático medio RMSE [m] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días) para el año 2018. Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

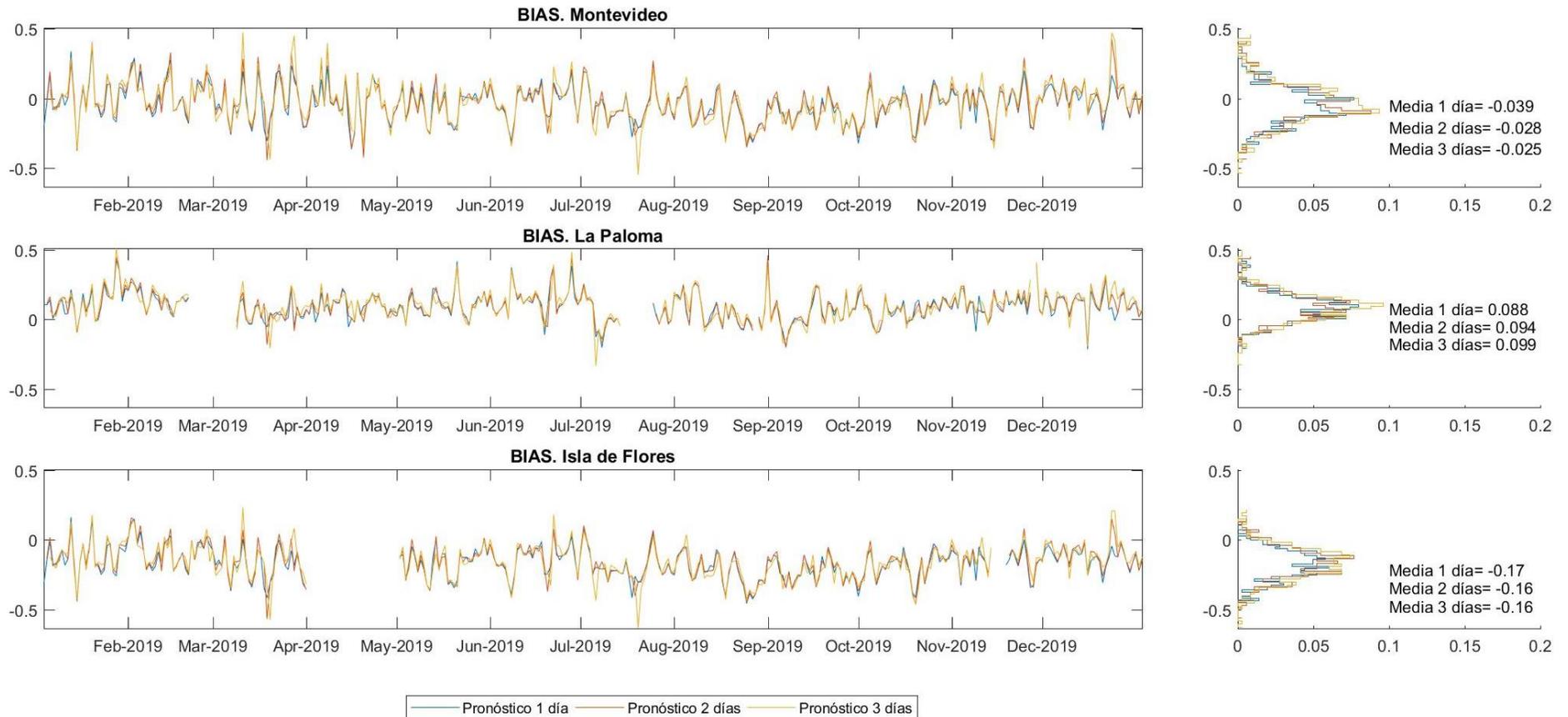


Figura 121 - Diferencia de media, o sesgo BIAS [m] para el año 2019 Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

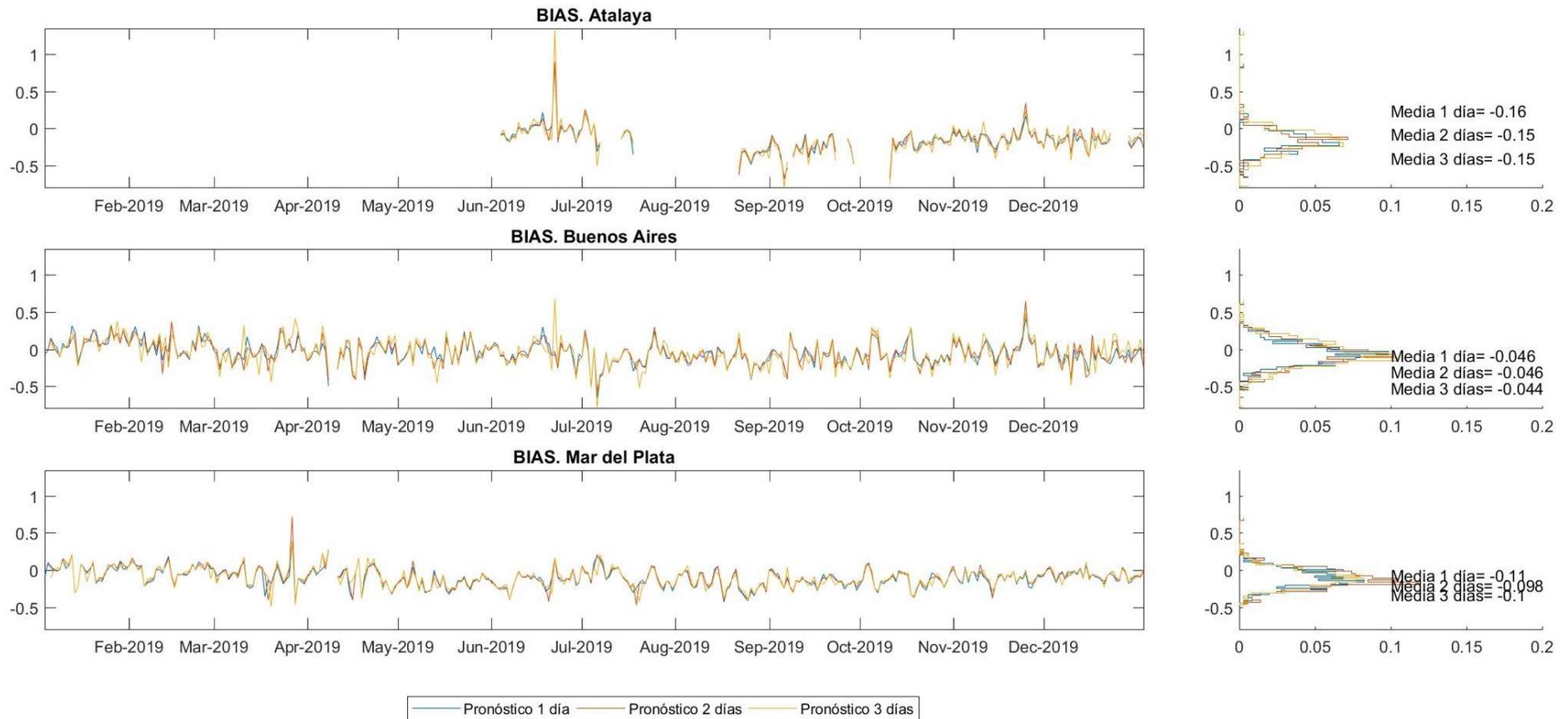


Figura 122 - Diferencia de media, o sesgo BIAS [m] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

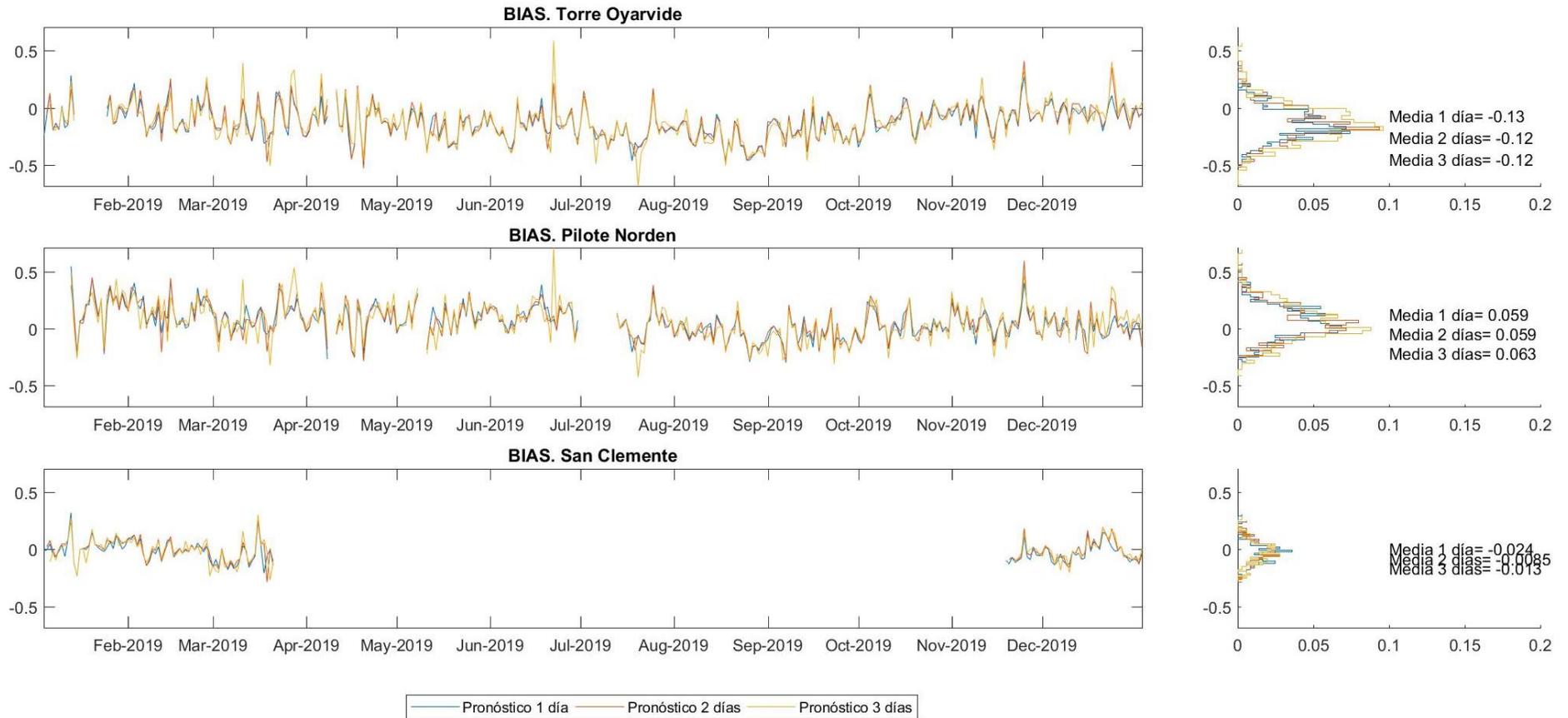


Figura 123 - Diferencia de media, o sesgo BIAS [m] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

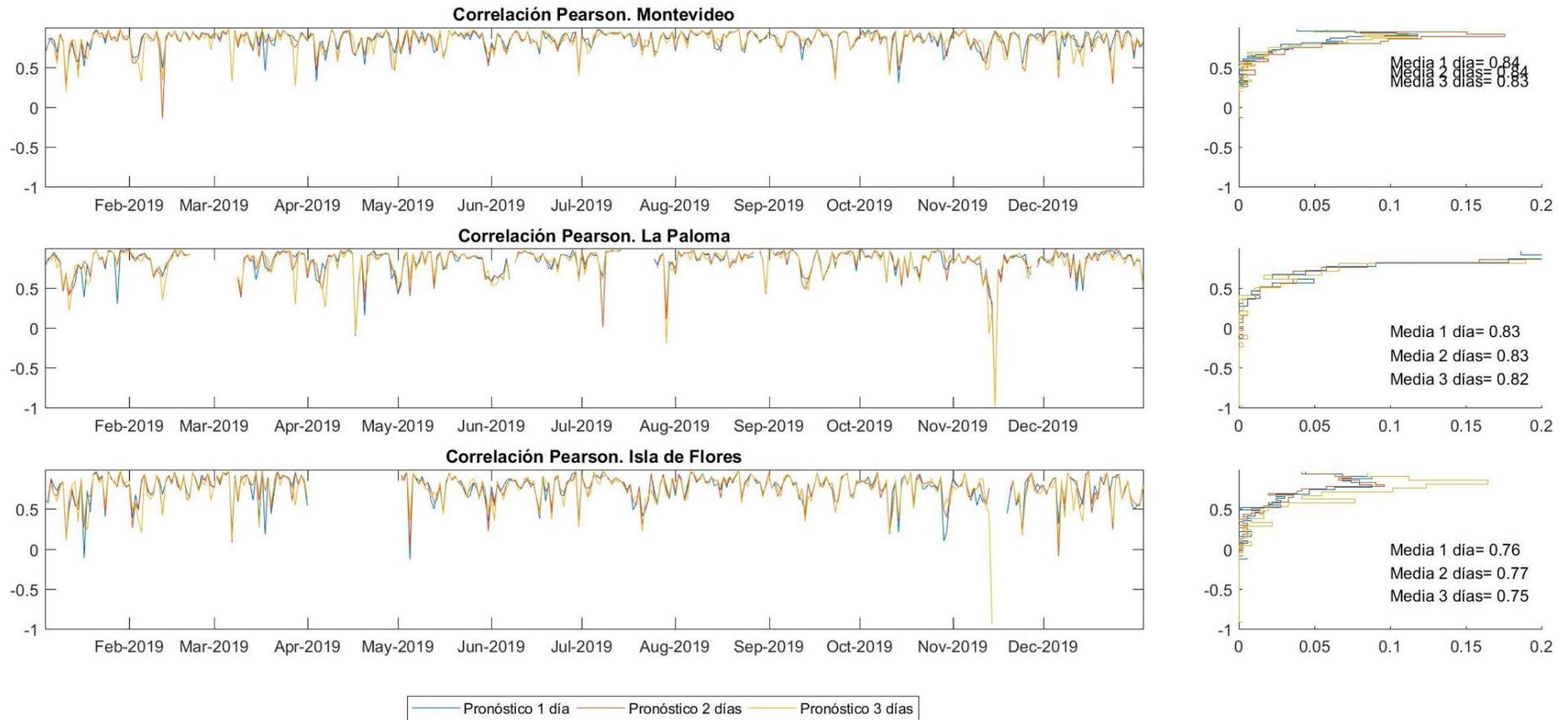


Figura 124– Correlación Pearson [-] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

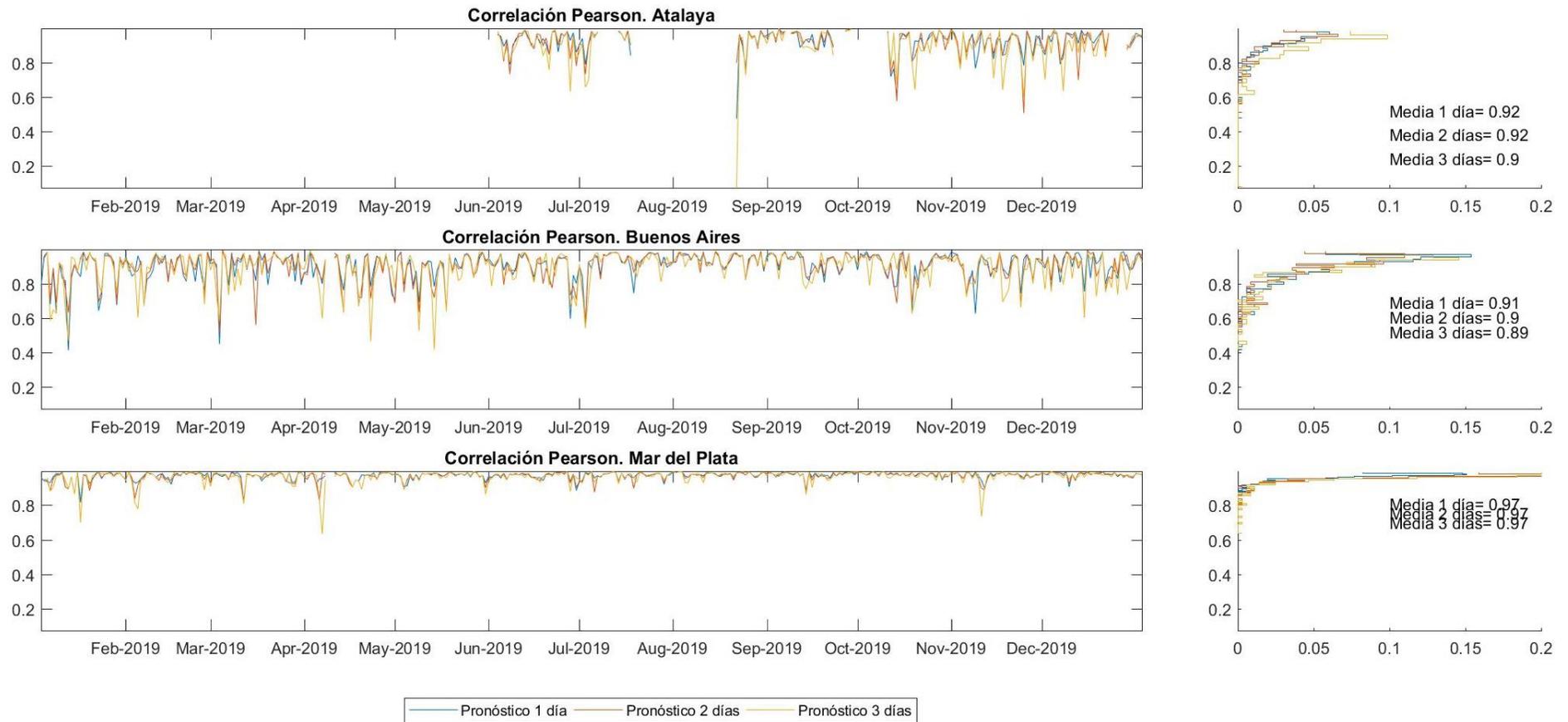


Figura 125 - Correlación Pearson [-] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

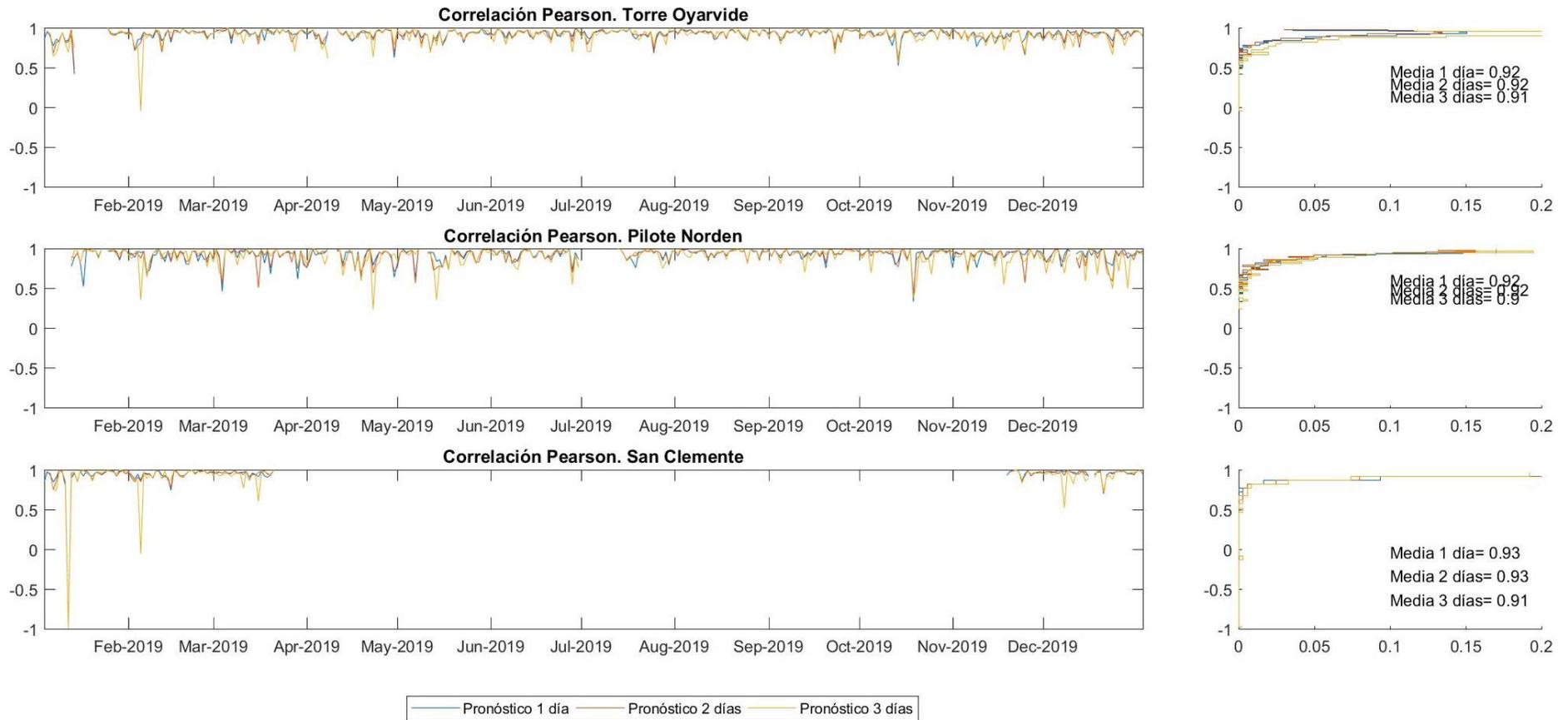


Figura 126 - Correlación Pearson [-] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

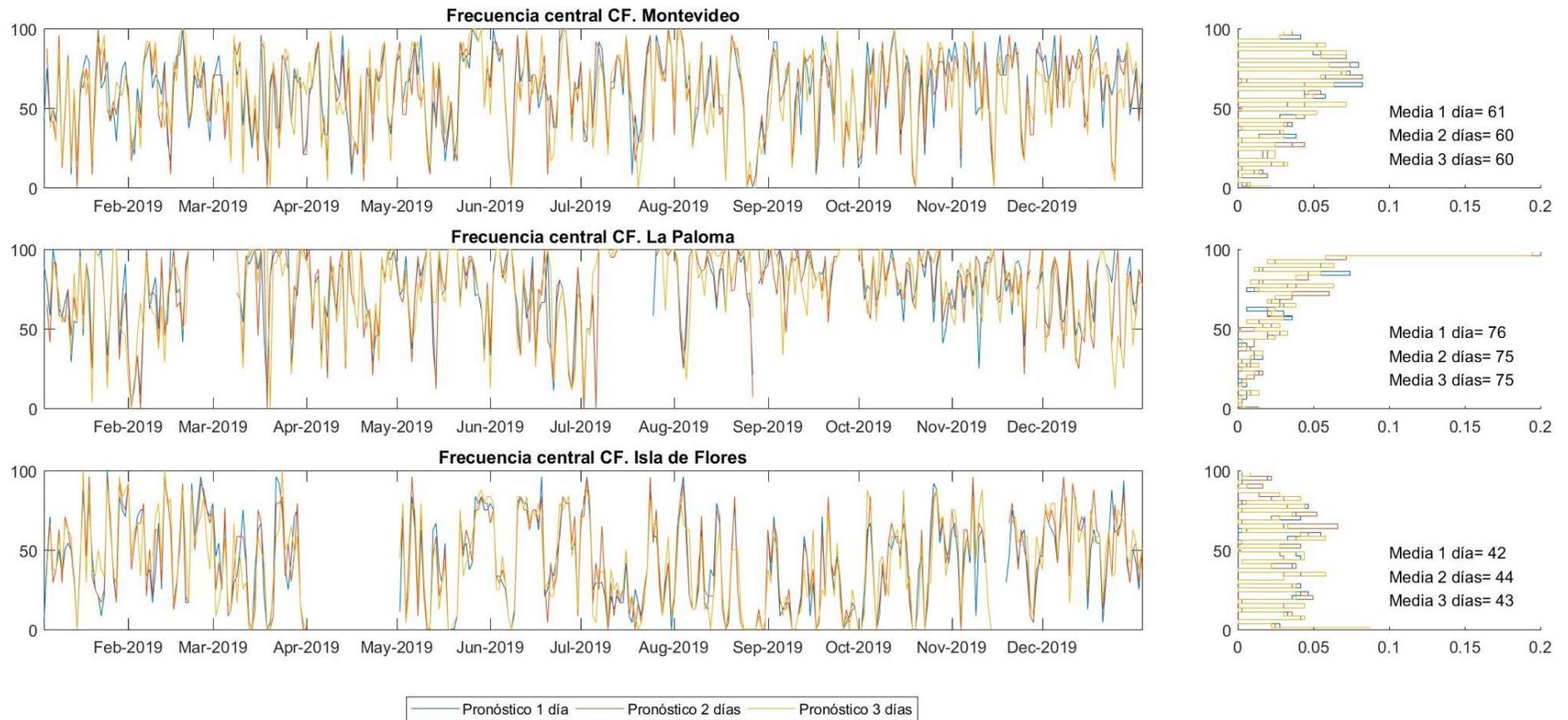


Figura 127 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

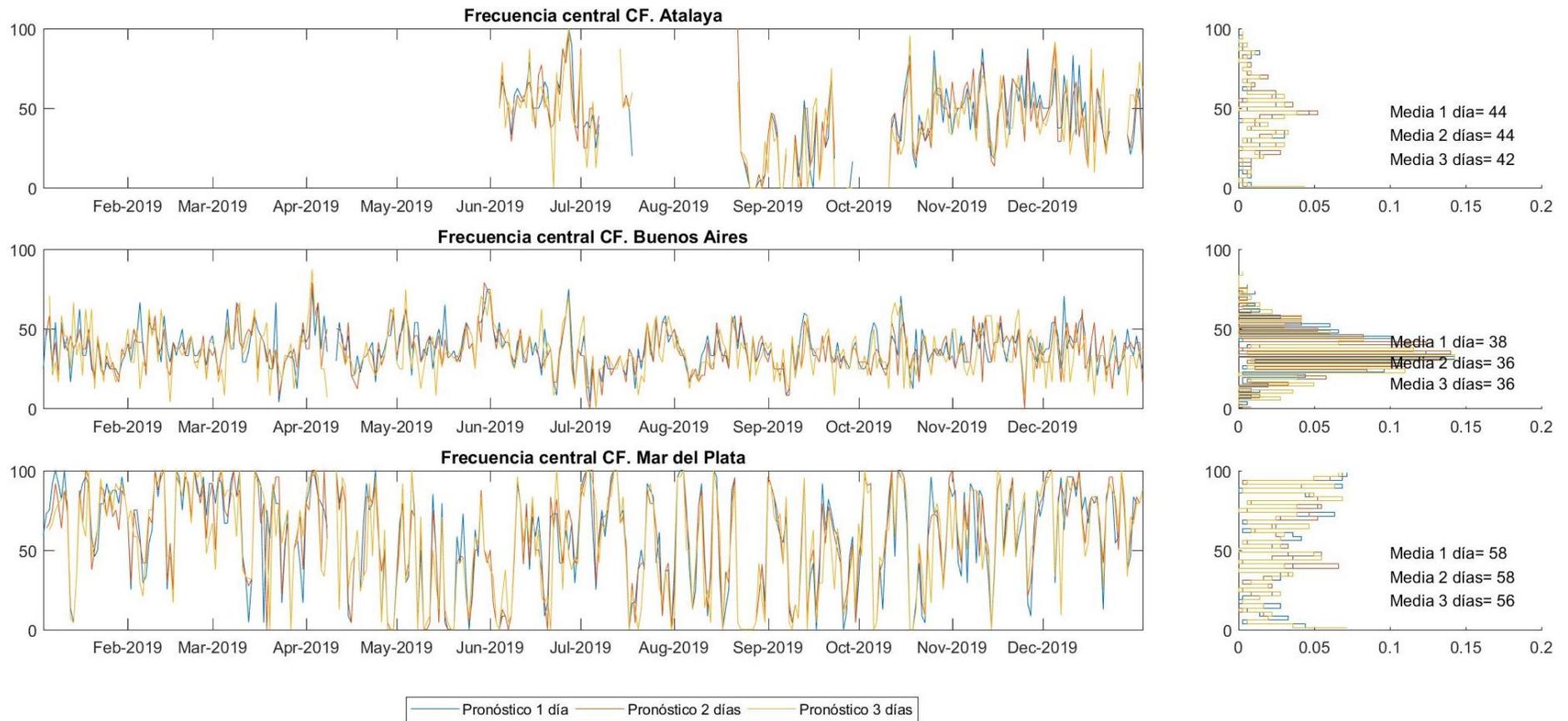


Figura 128 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

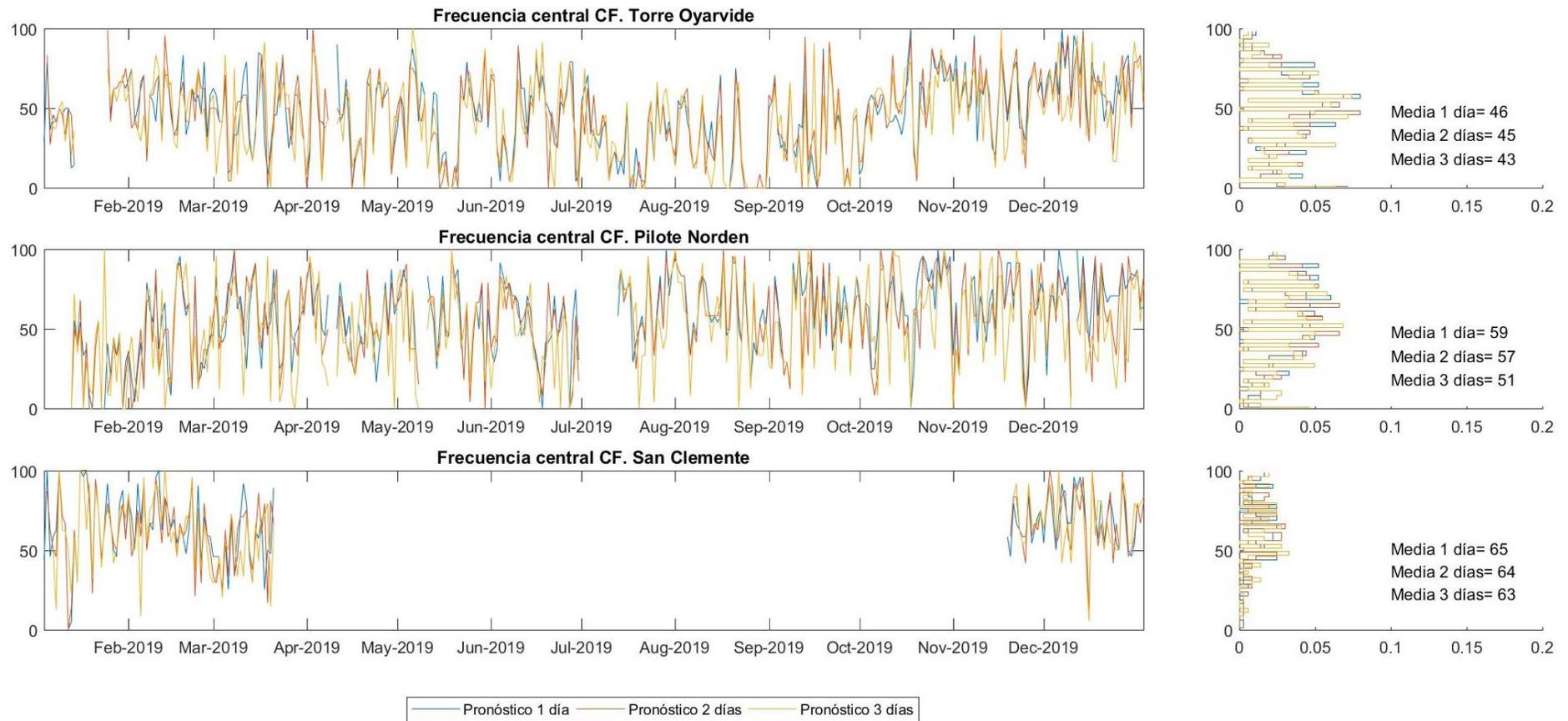


Figura 129 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

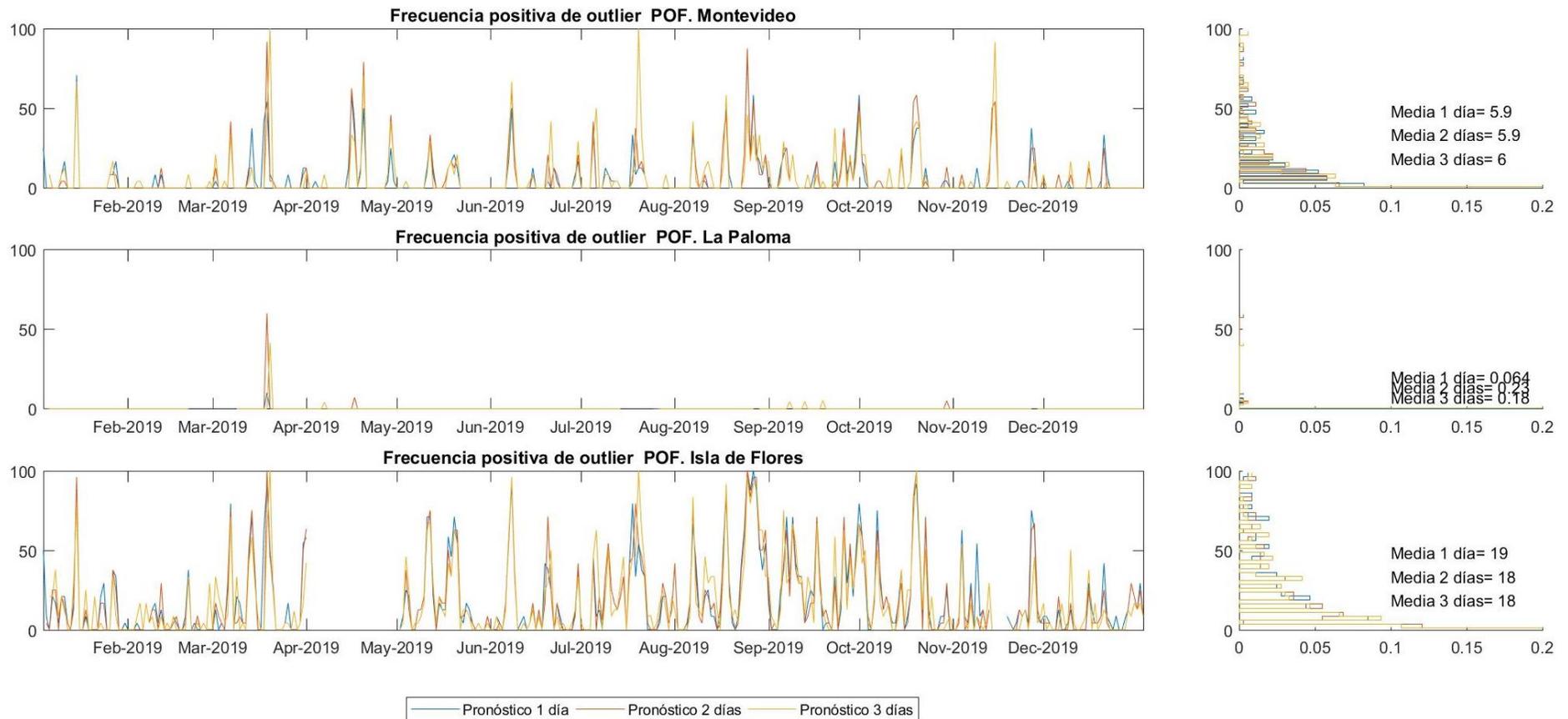


Figura 130 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

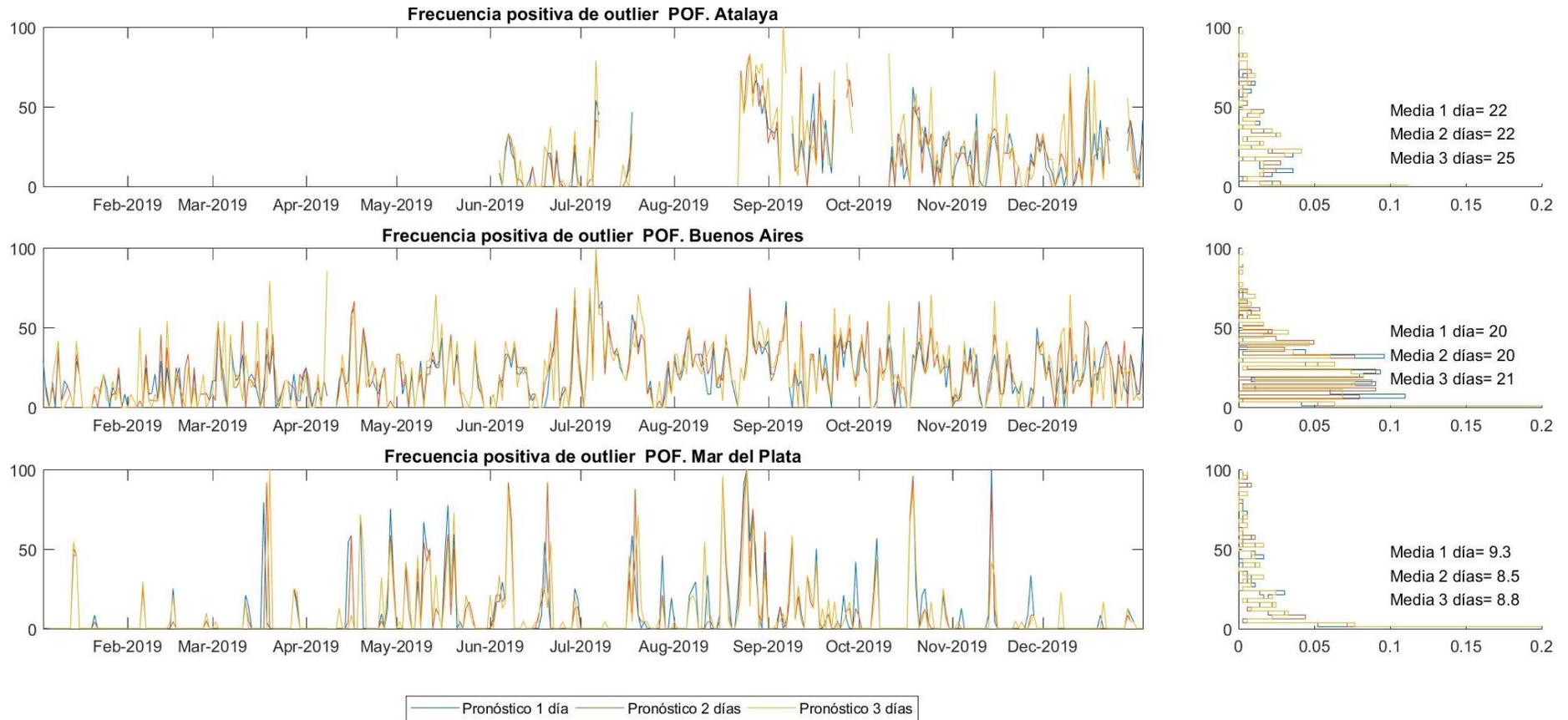


Figura 131 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

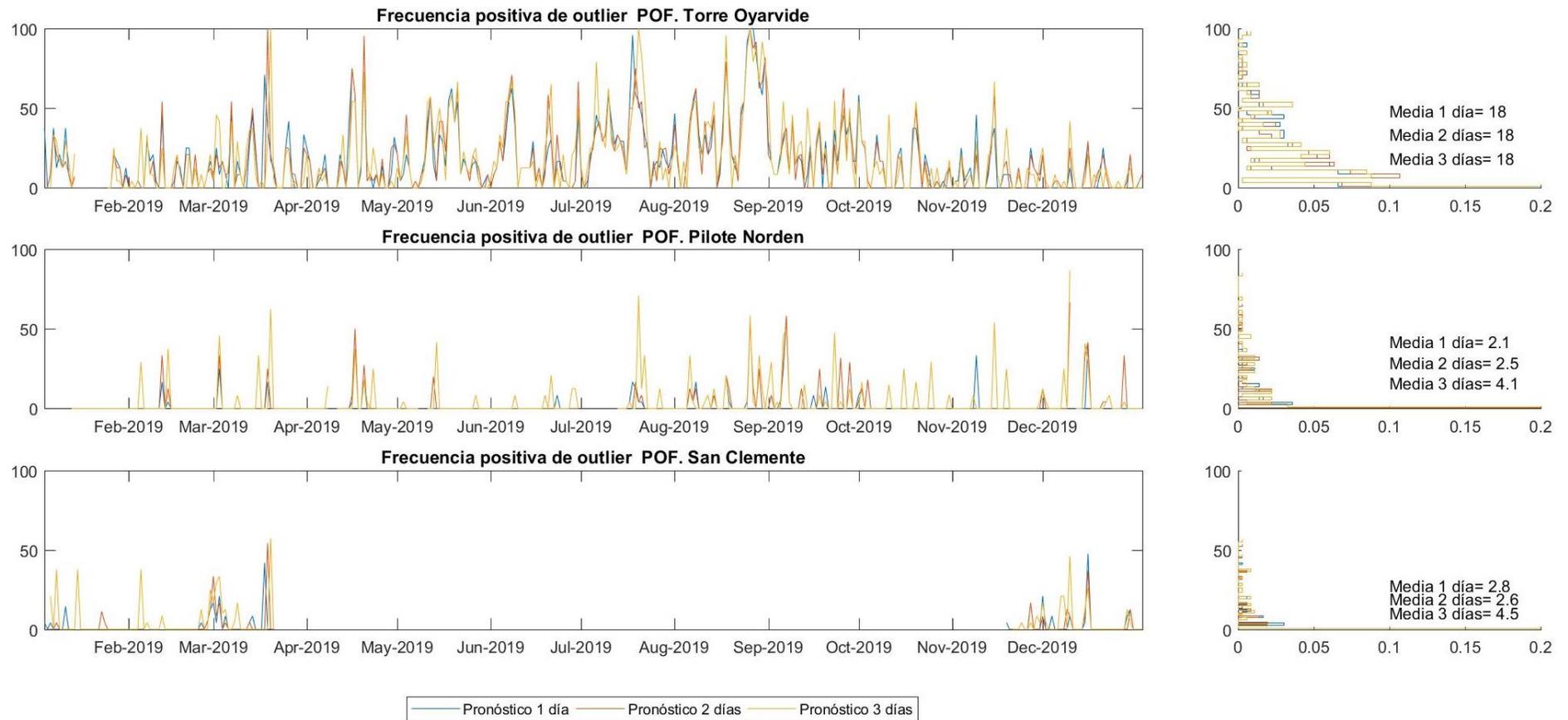


Figura 132 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

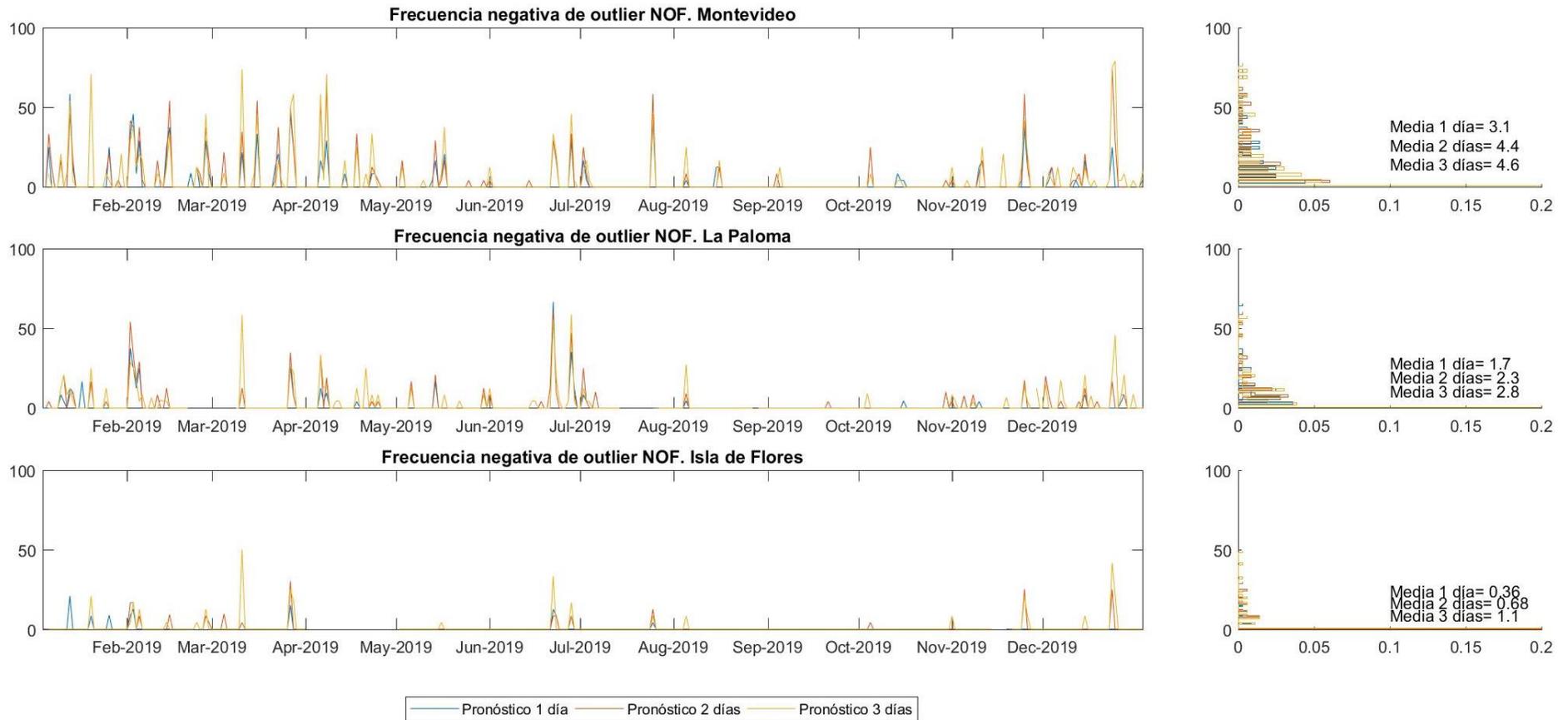


Figura 133 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

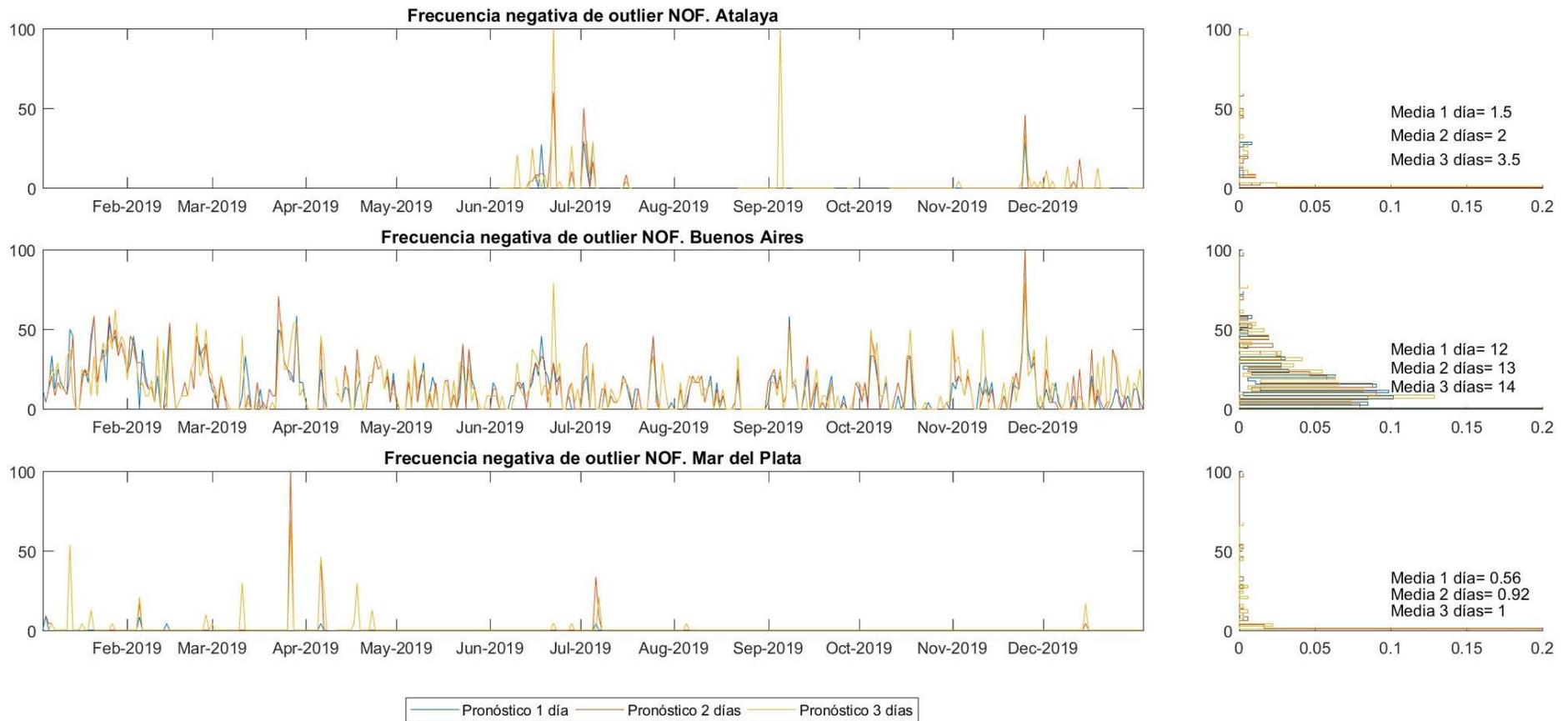


Figura 134 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

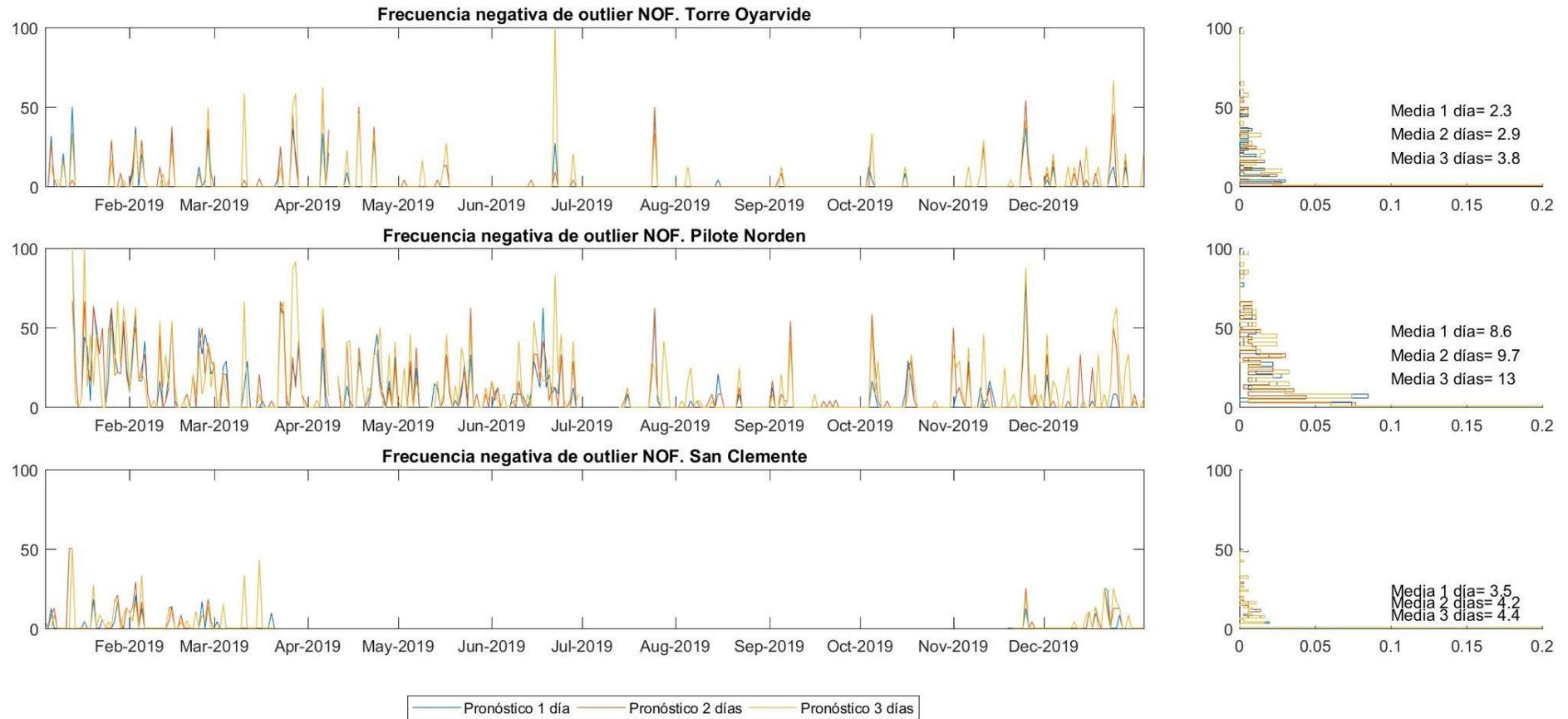


Figura 135 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

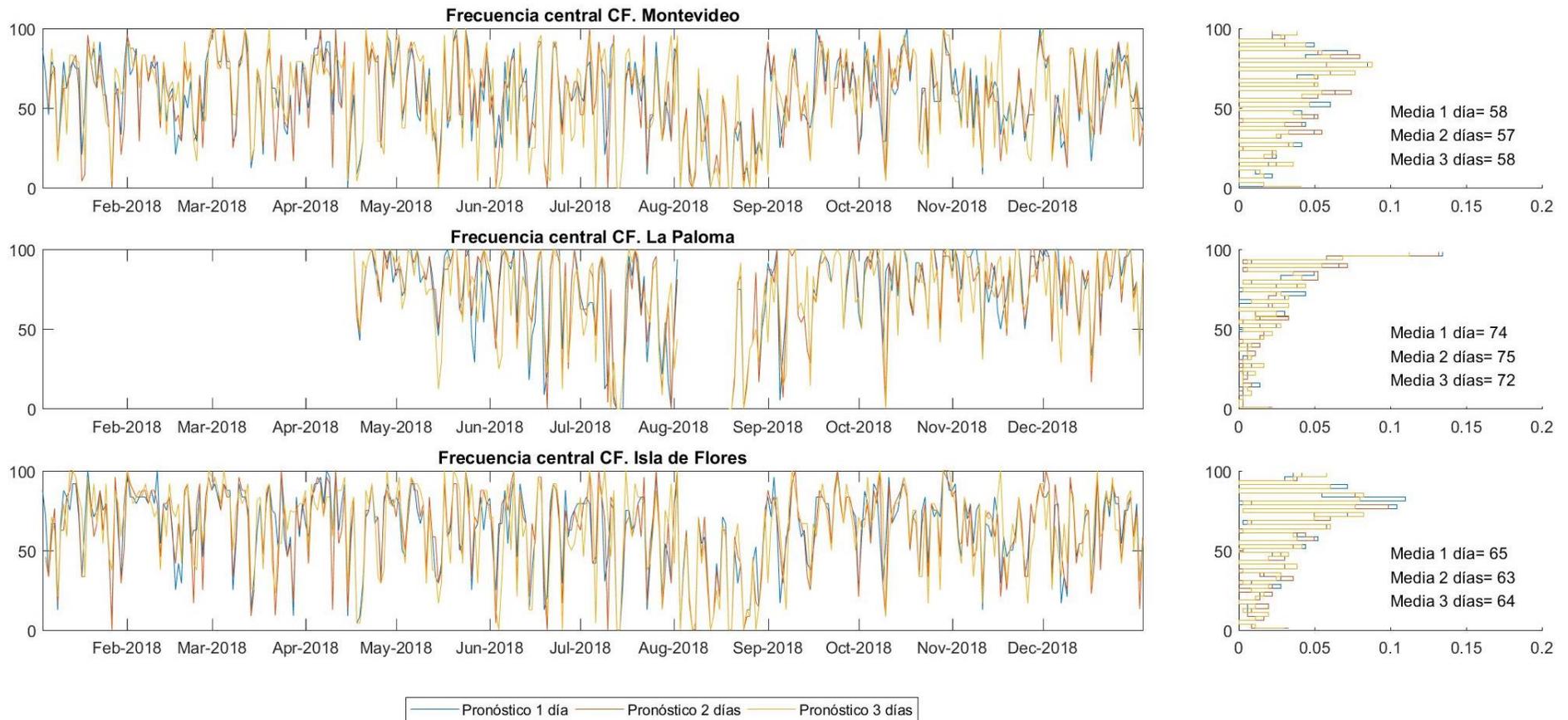


Figura 136 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

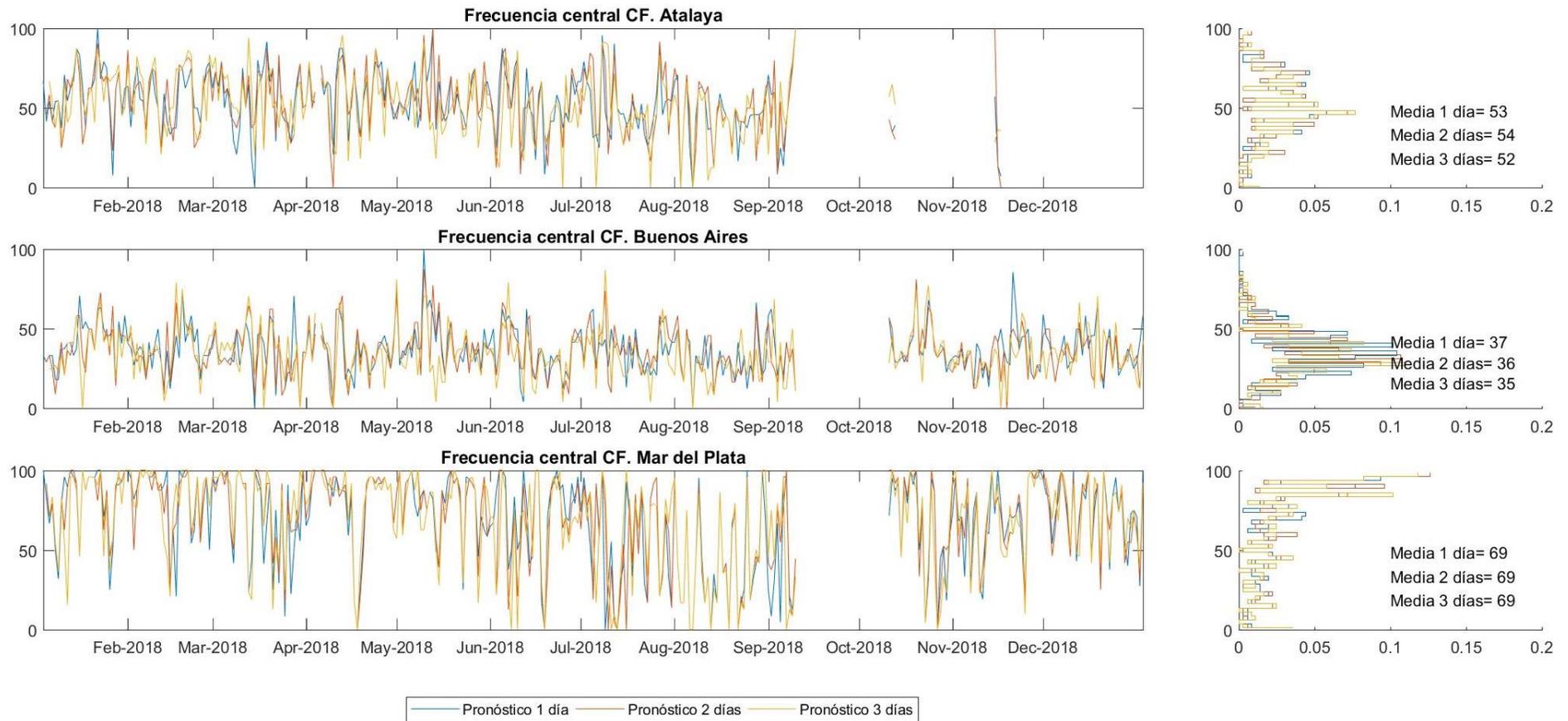


Figura 137 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

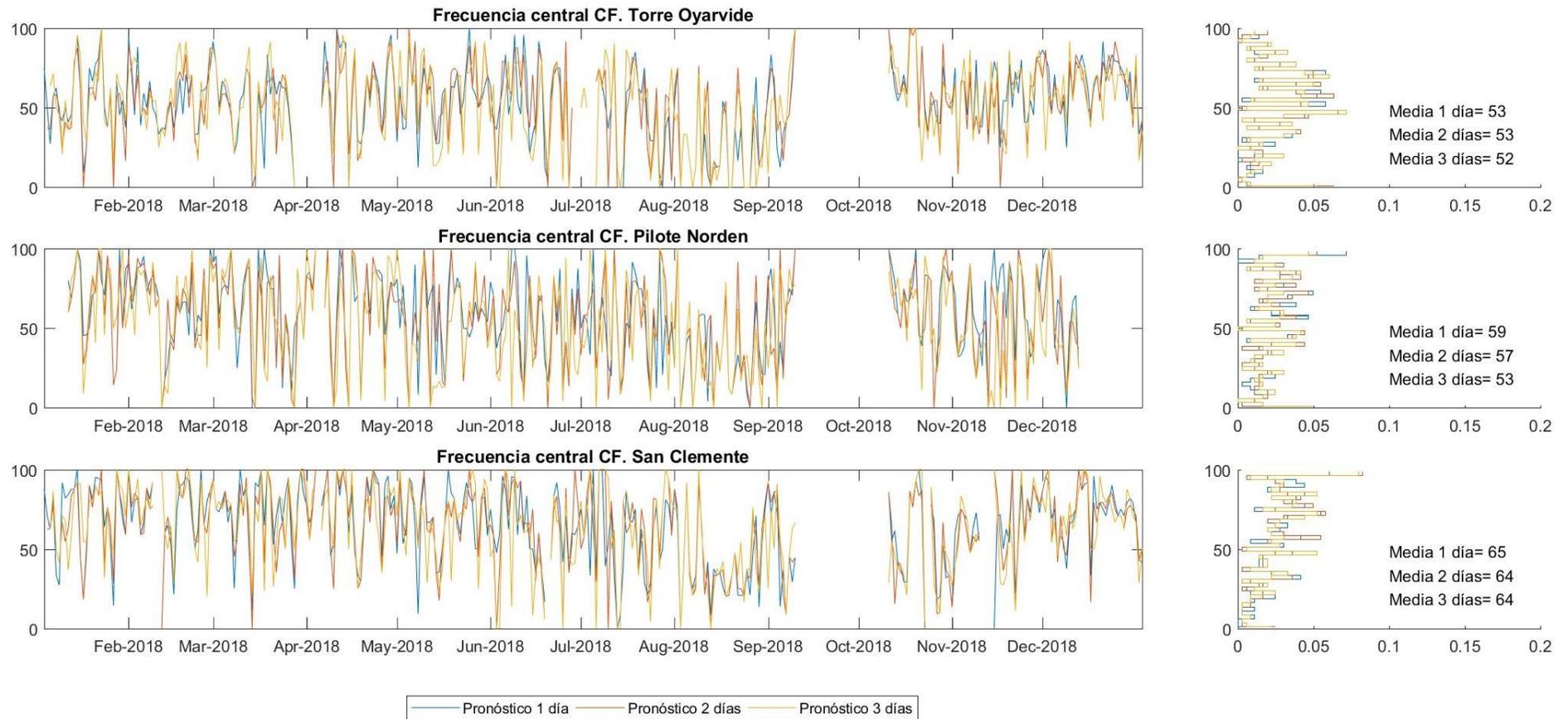


Figura 138 - Frecuencia central CF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

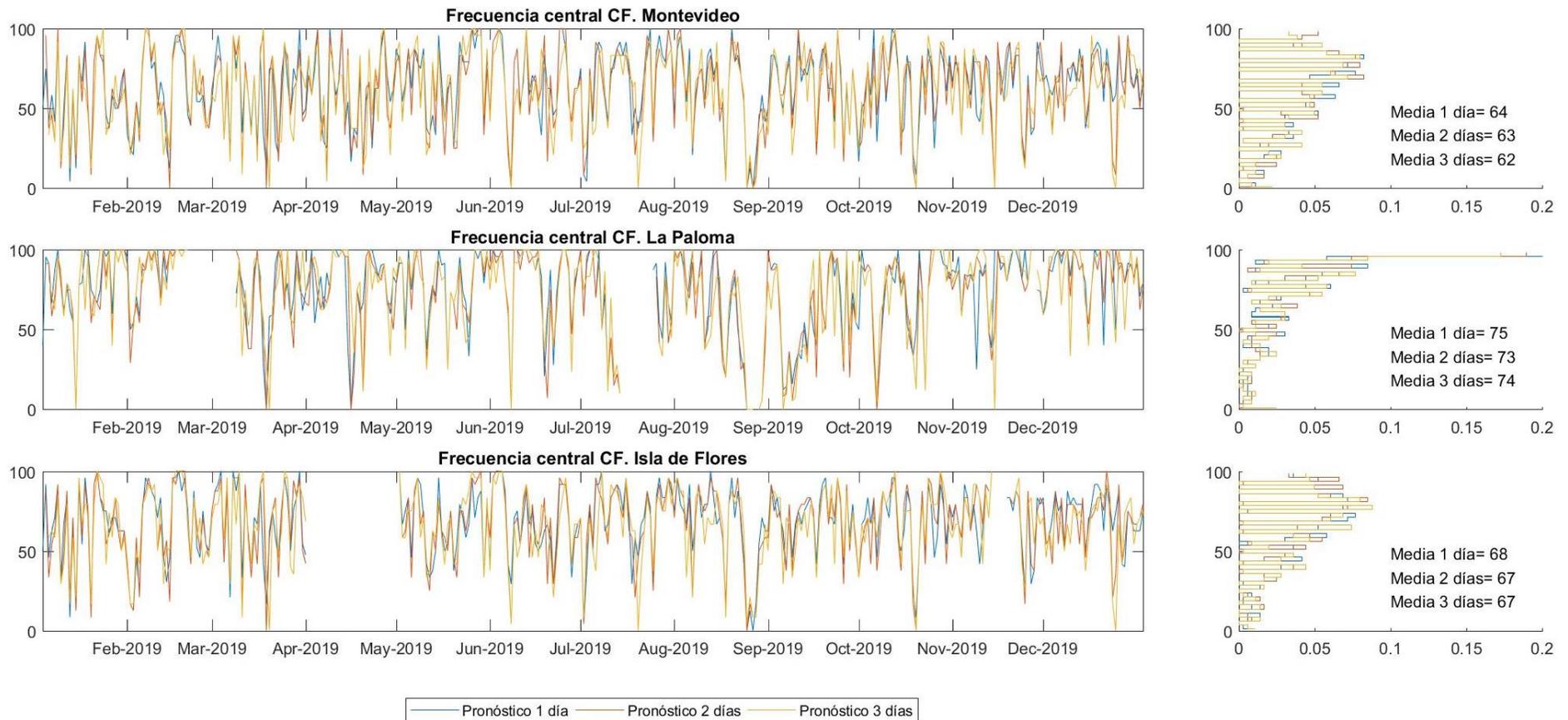


Figura 139 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

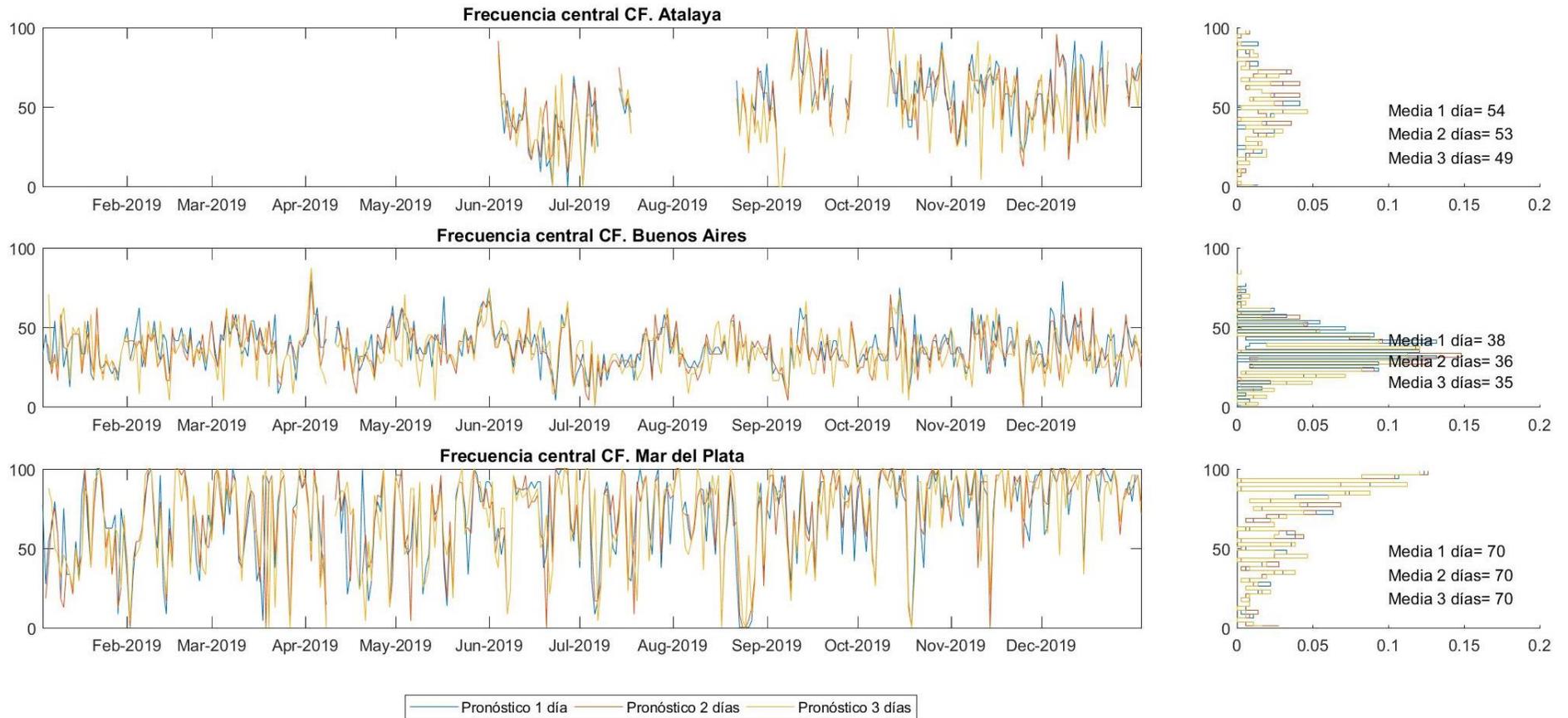


Figura 140 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

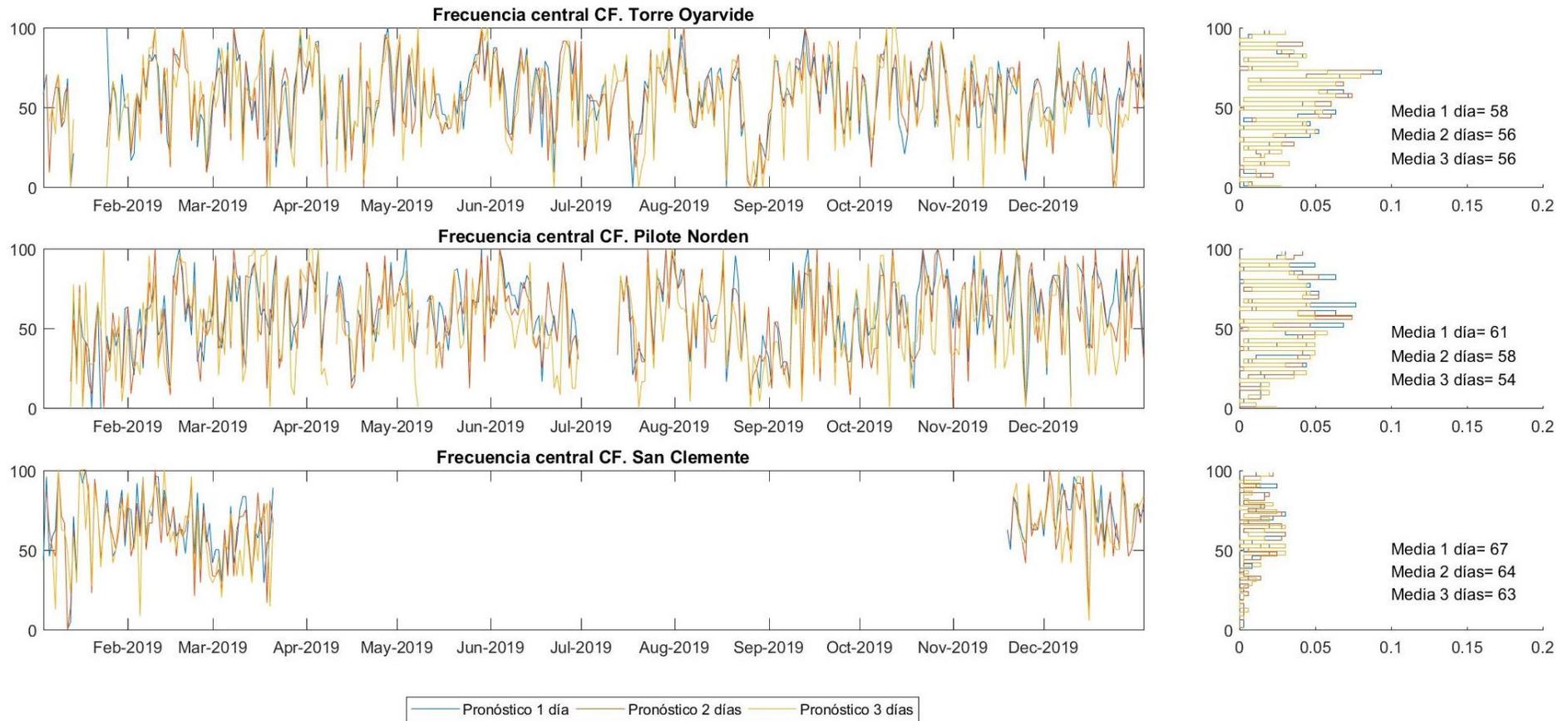


Figura 141 - Frecuencia central CF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

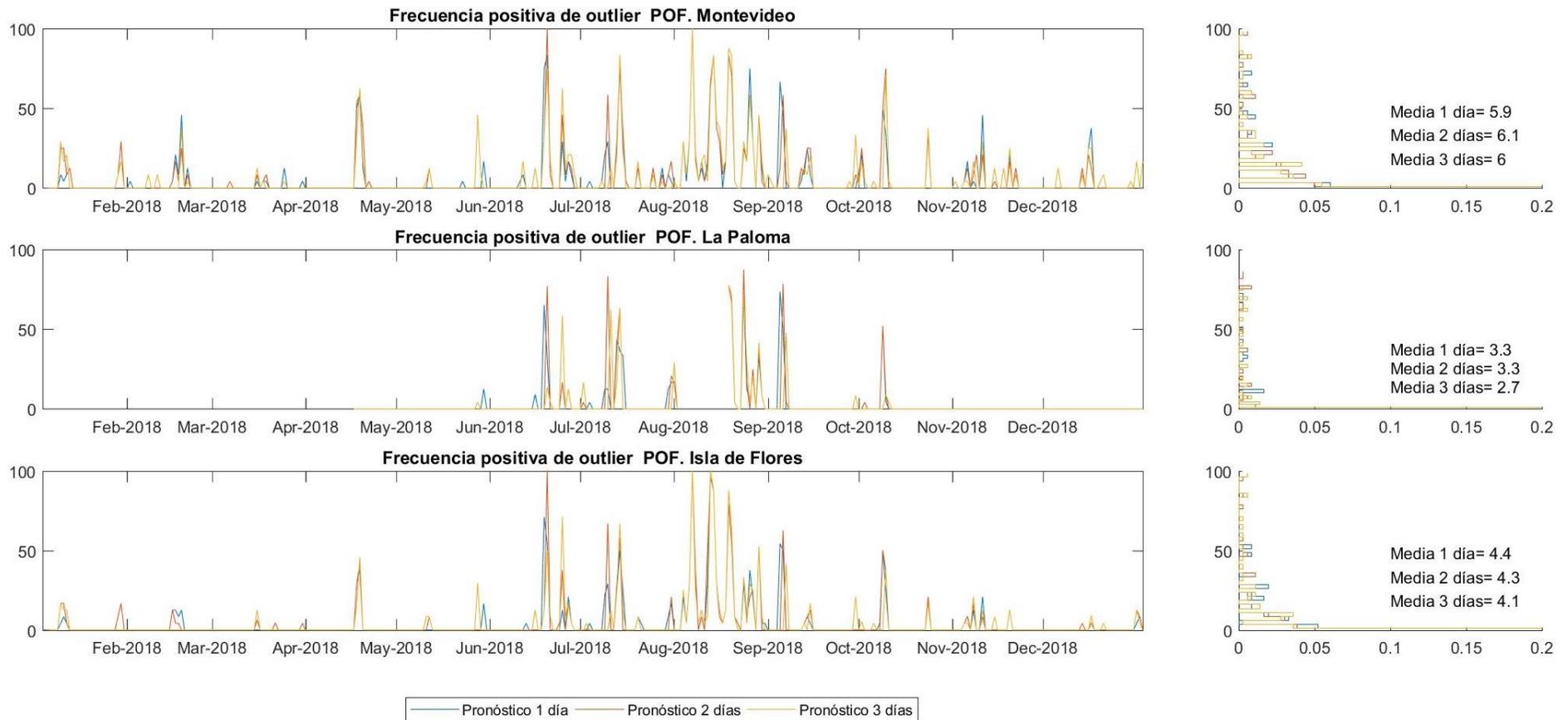


Figura 142 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

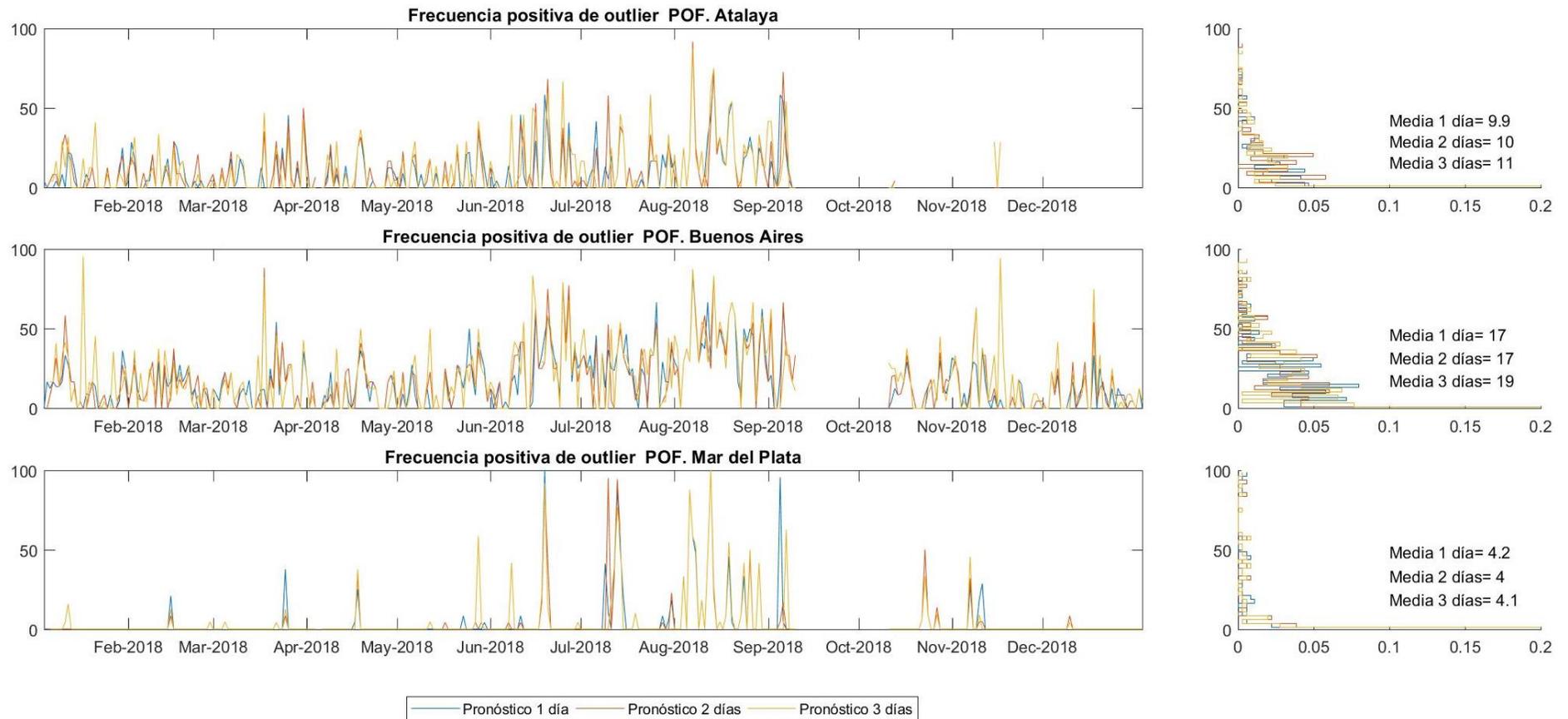


Figura 143 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

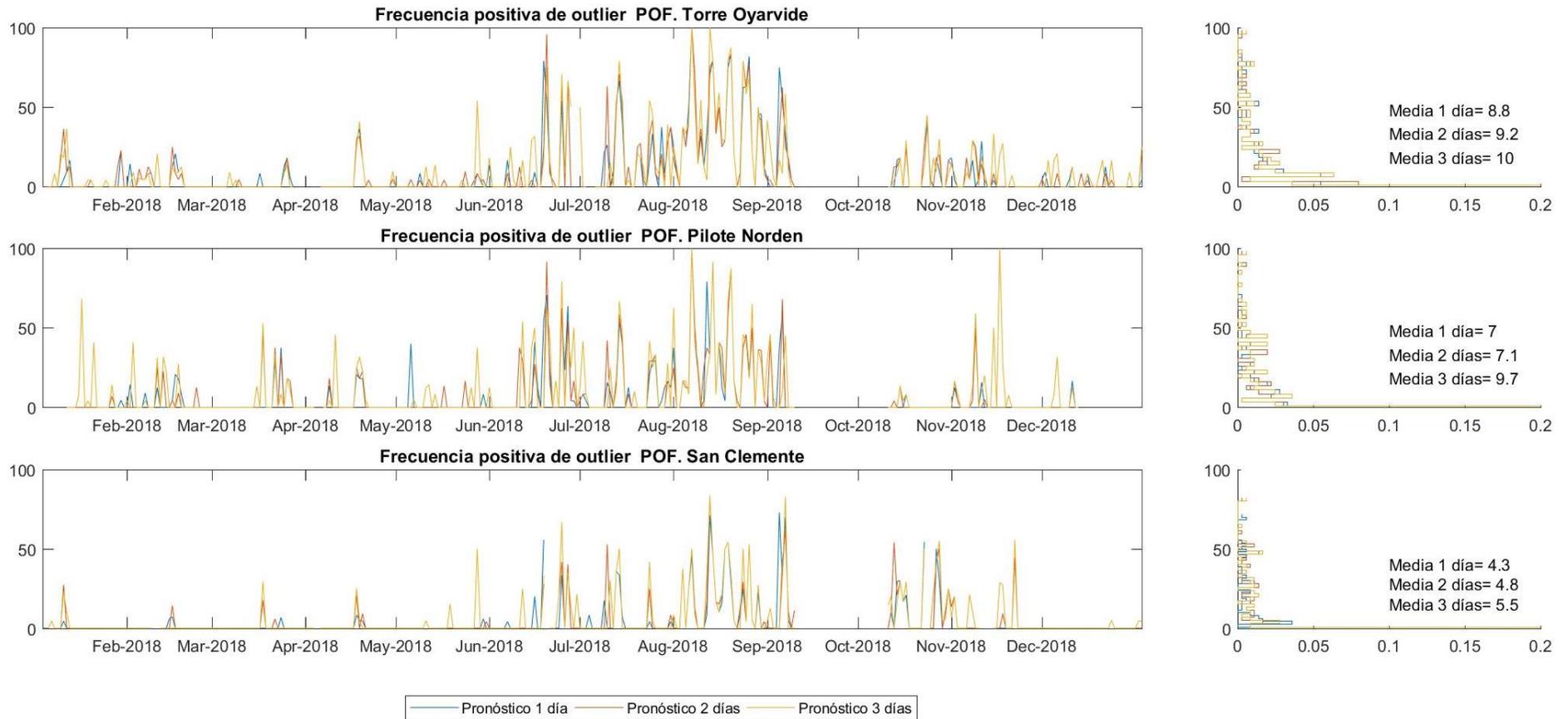


Figura 144 - Frecuencia positiva de outlier POF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

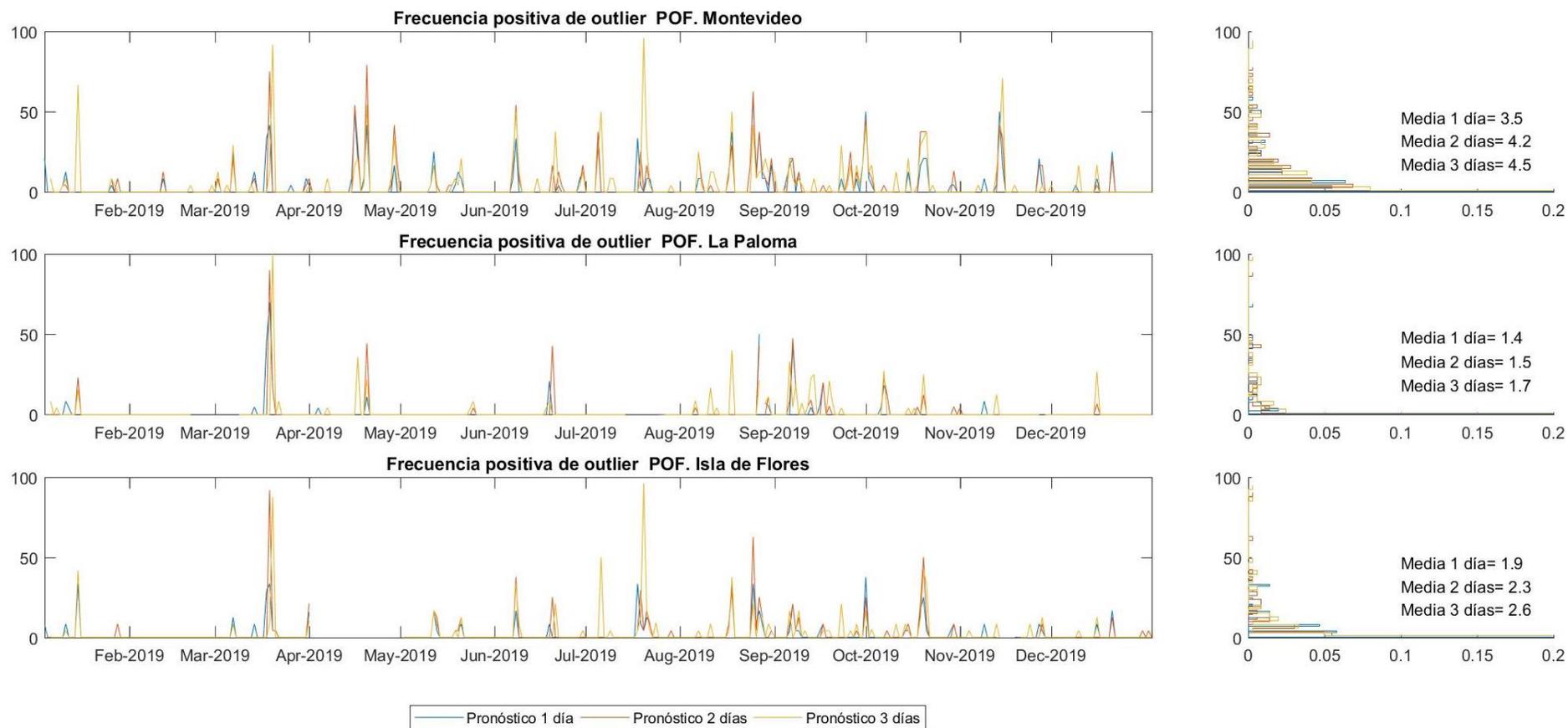


Figura 145 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

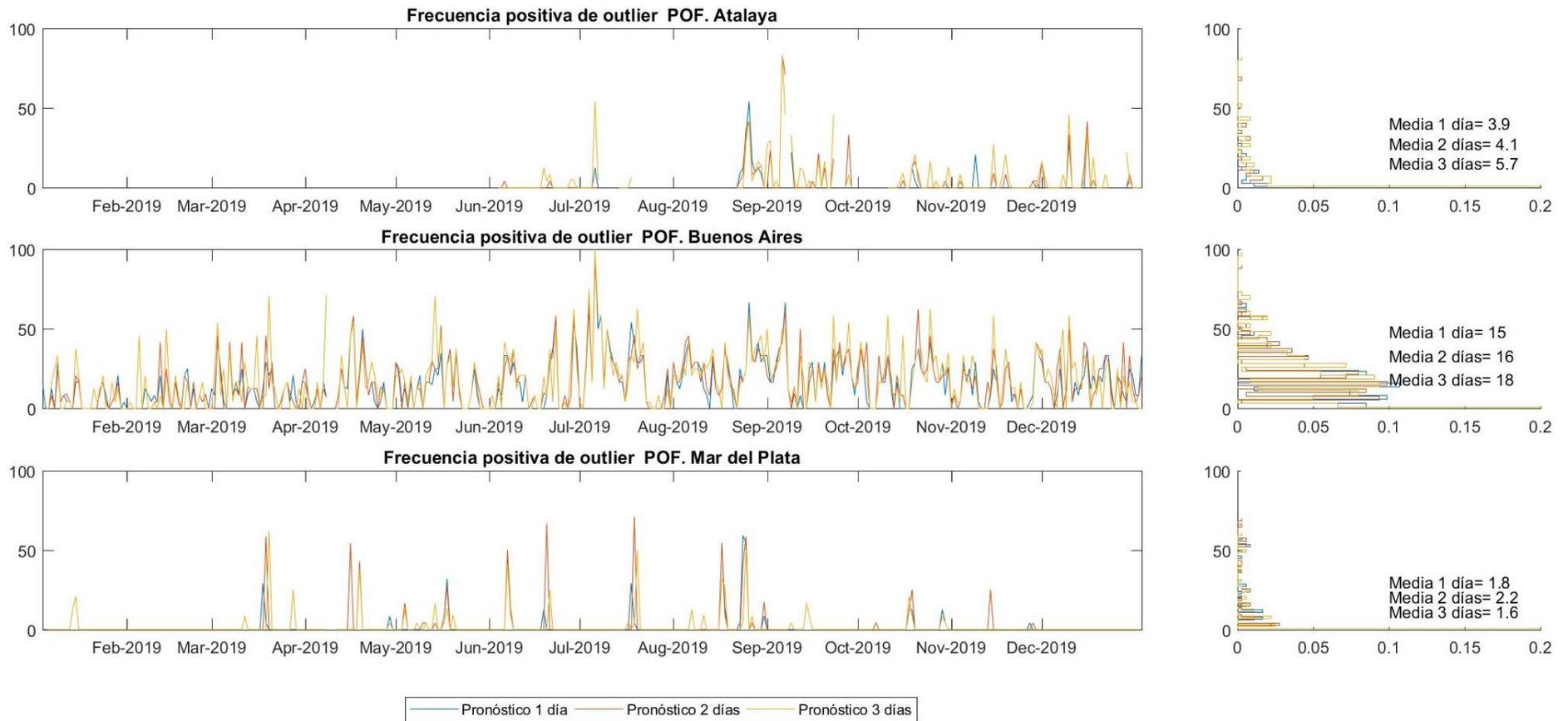


Figura 146 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

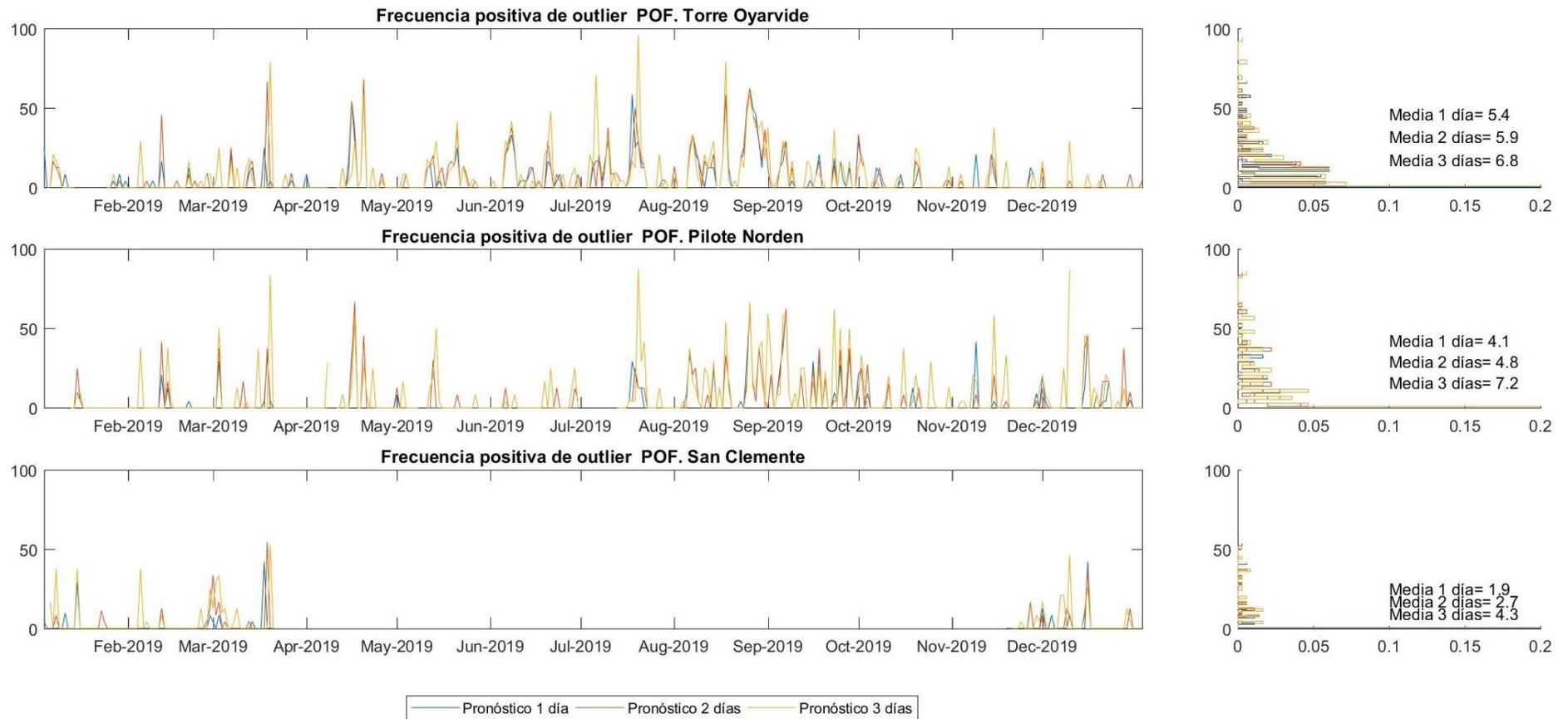


Figura 147 - Frecuencia positiva de outlier POF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

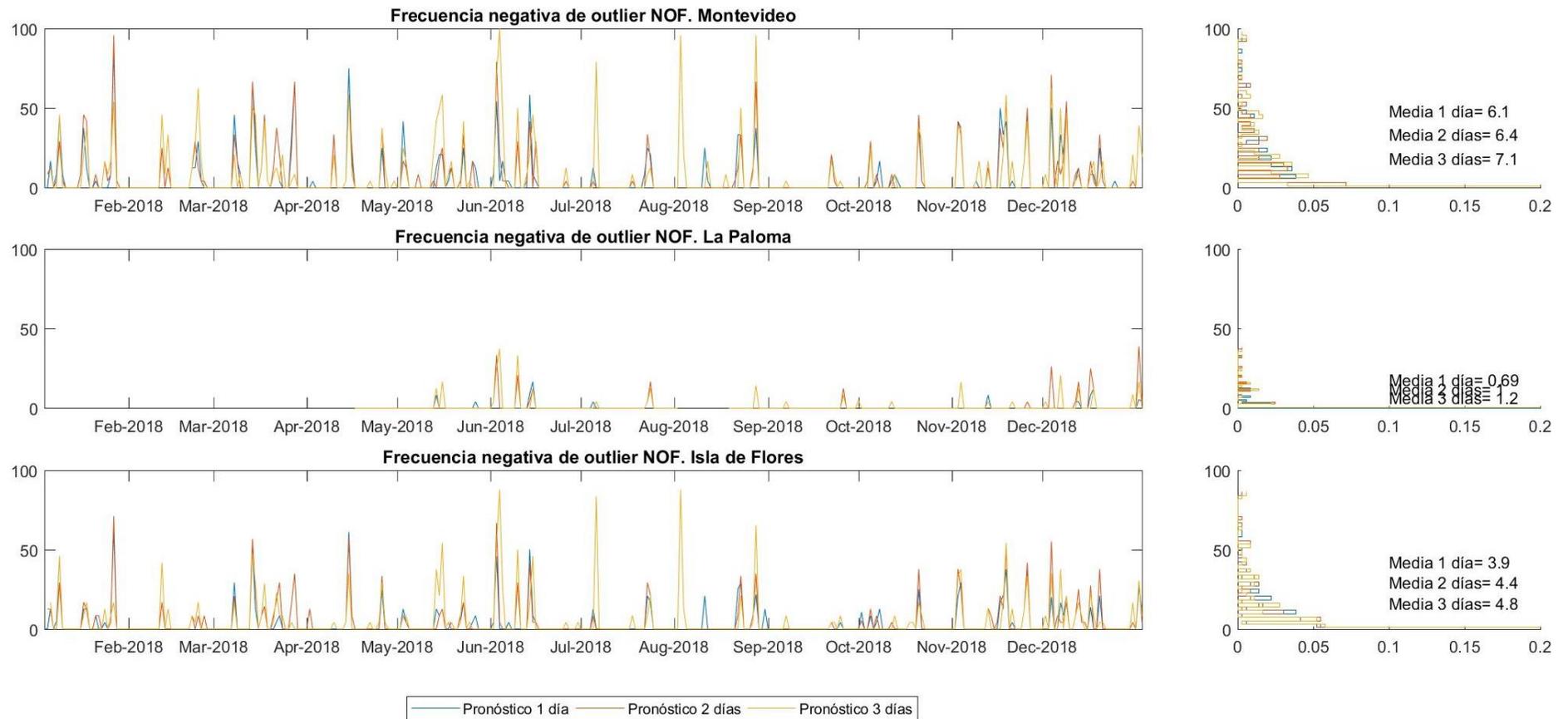


Figura 148 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

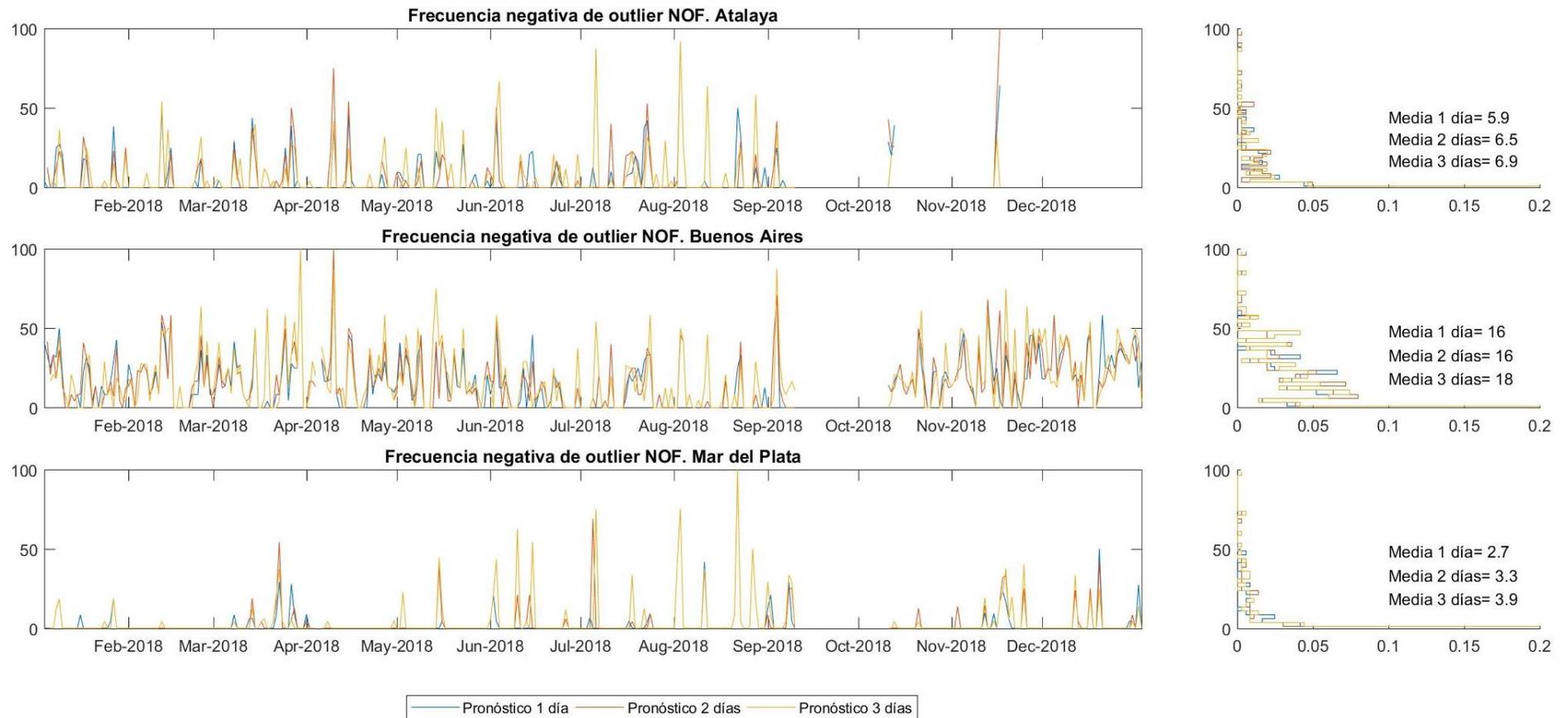


Figura 149 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

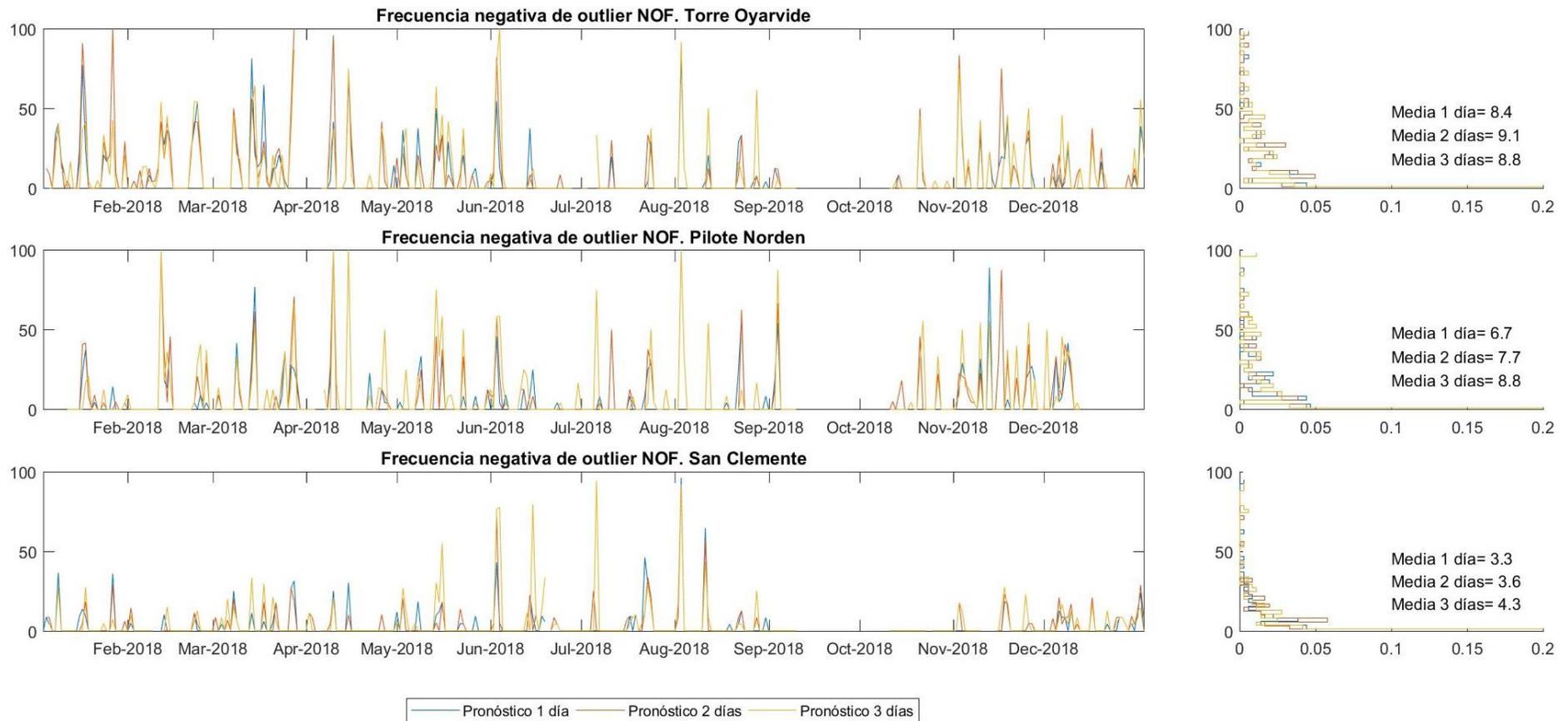


Figura 150 - Frecuencia negativa de outlier NOF [%] para el año 2018. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

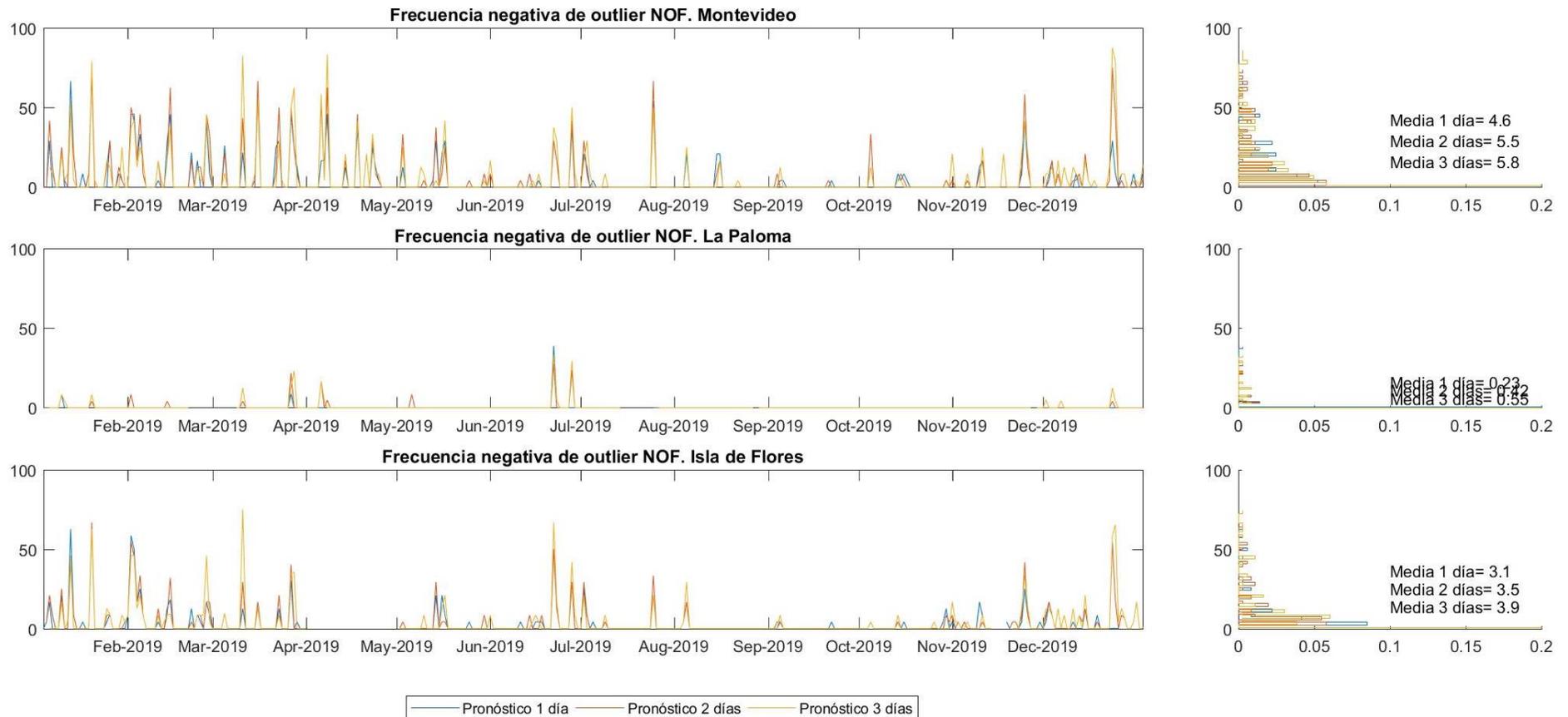


Figura 151 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Montevideo, La Paloma e Isla de Flores. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

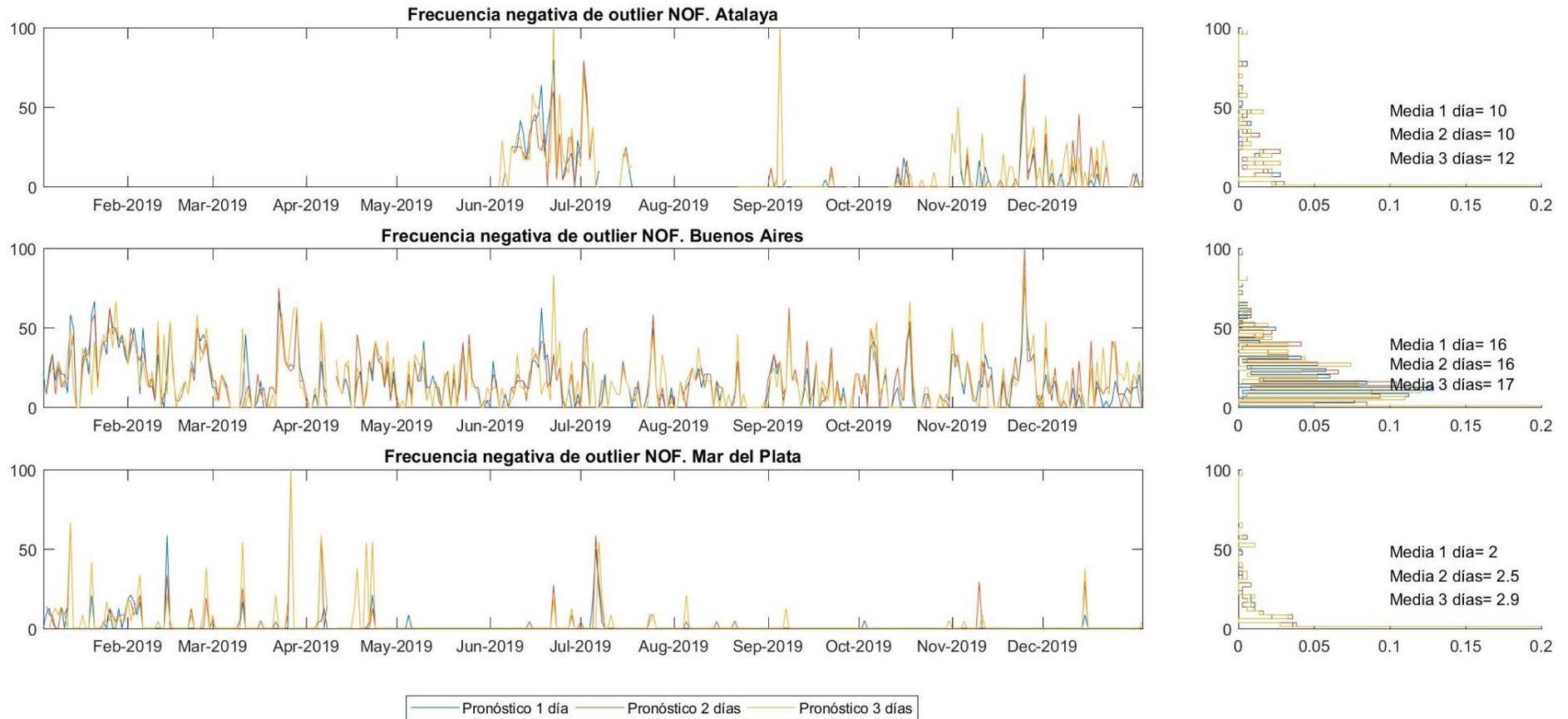


Figura 152 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Atalaya, Buenos Aires y Mar del Plata. Valores respecto a la media anual.

Sistema de pronóstico de los niveles de marea en tiempo real en el Río de la Plata

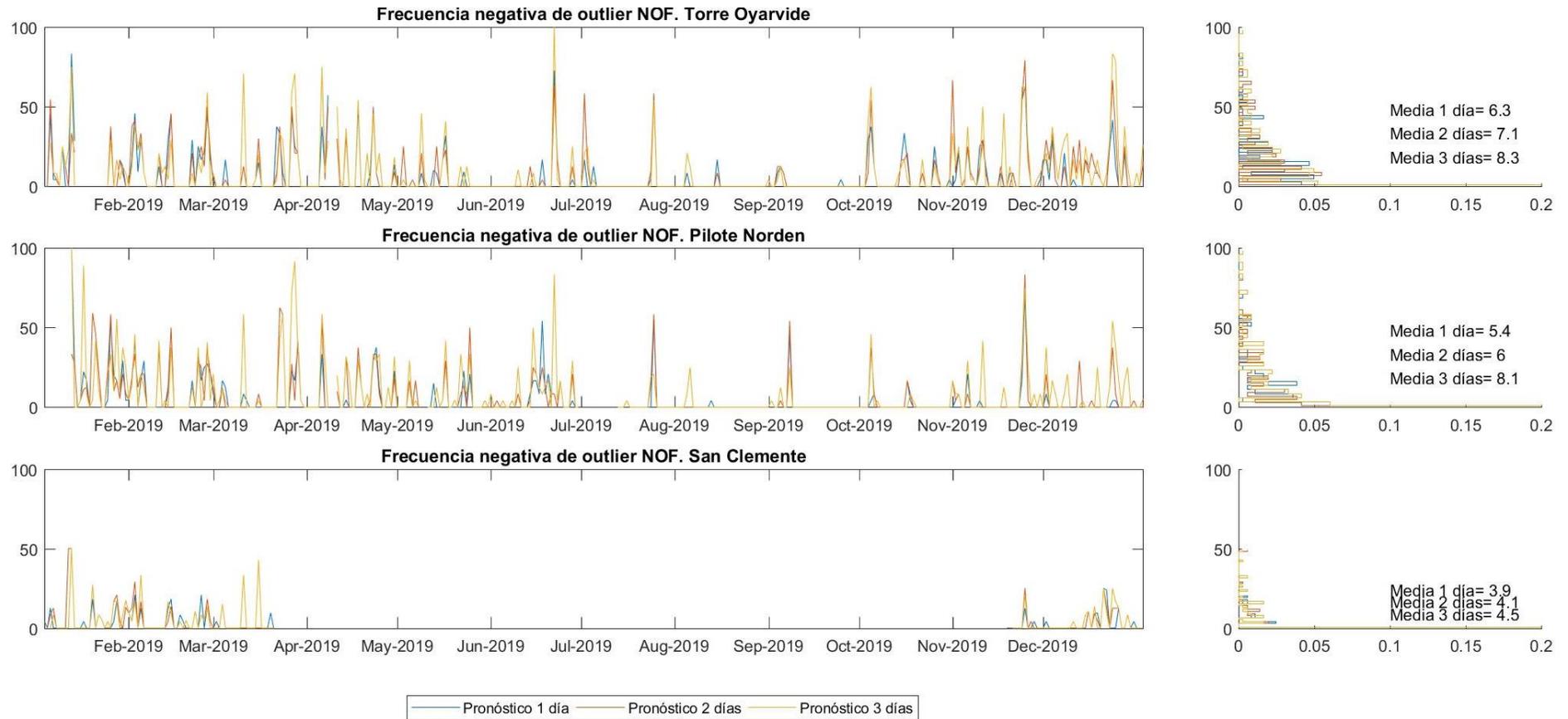


Figura 153 - Frecuencia negativa de outlier NOF [%] para el año 2019. Comportamiento del modelo durante las primeras 24 horas de pronóstico (1 día) de las 24 a las 48 horas (2 días) y de las 48 a las 72 horas (3 días). Mareógrafos de Torre Oyarvide, Pilote Norden y San Clemente. Valores respecto a la media anual.

