**Informe Técnico Nº 5**

**- Julio 2022 -**

**ACTUALIZACIÓN DE LA BASE DE DATOS Y LOS PROGRAMAS UTILIZADOS EN EL PRONÓSTICO DIARIO DE HIELO MARINO**

**Lic. Mariana Scilingo**

**Introducción**

En la Antártida se encuentran trece bases administradas por Argentina, las cuales son importantes para llevar a cabo actividades relacionadas con la investigación científica. En el planeamiento de las campañas antárticas, imprescindibles para garantizar la permanencia y el abastecimiento de las bases, es muy importante conocer las condiciones del hielo marino (HM) con anticipación. Es por esta razón que el Servicio de Hidrografía Naval (SHN) lleva adelante un pronóstico que provee información de la concentración de HM tanto diaria como mensual, que se publica en la página del SHN a principio de cada mes (http://www.hidro.gov.ar/smara/SB/sb.asp).

El desarrollo del pronóstico de HM se basa en un modelo estadístico que aplica el análisis de Componentes Principales a los datos satelitales disponibles de concentración de HM. La base de datos satelitales diaria se actualiza anualmente y la mensual cuando termina cada mes. Cuenta hasta el momento con una extensión de más de 40 años. La permanente ampliación de esta base permite mejorar los resultados del pronóstico debido a su carácter climatológico. Además se requiere el desarrollo de programas computacionales con el fin de generar un pronóstico de HM que luego pueda ser visualizado en campos diarios y mensuales. Continuamente se trabaja en la actualización y el reajuste de los programas para disminuir el error en las predicciones. Complementariamente se trabaja en modernizar la visualización de la información.

En este trabajo se mostrarán las principales mejoras implementadas en el pronóstico diario y en su base de datos. También se mostrarán gráficos de los campos diarios obtenidos.

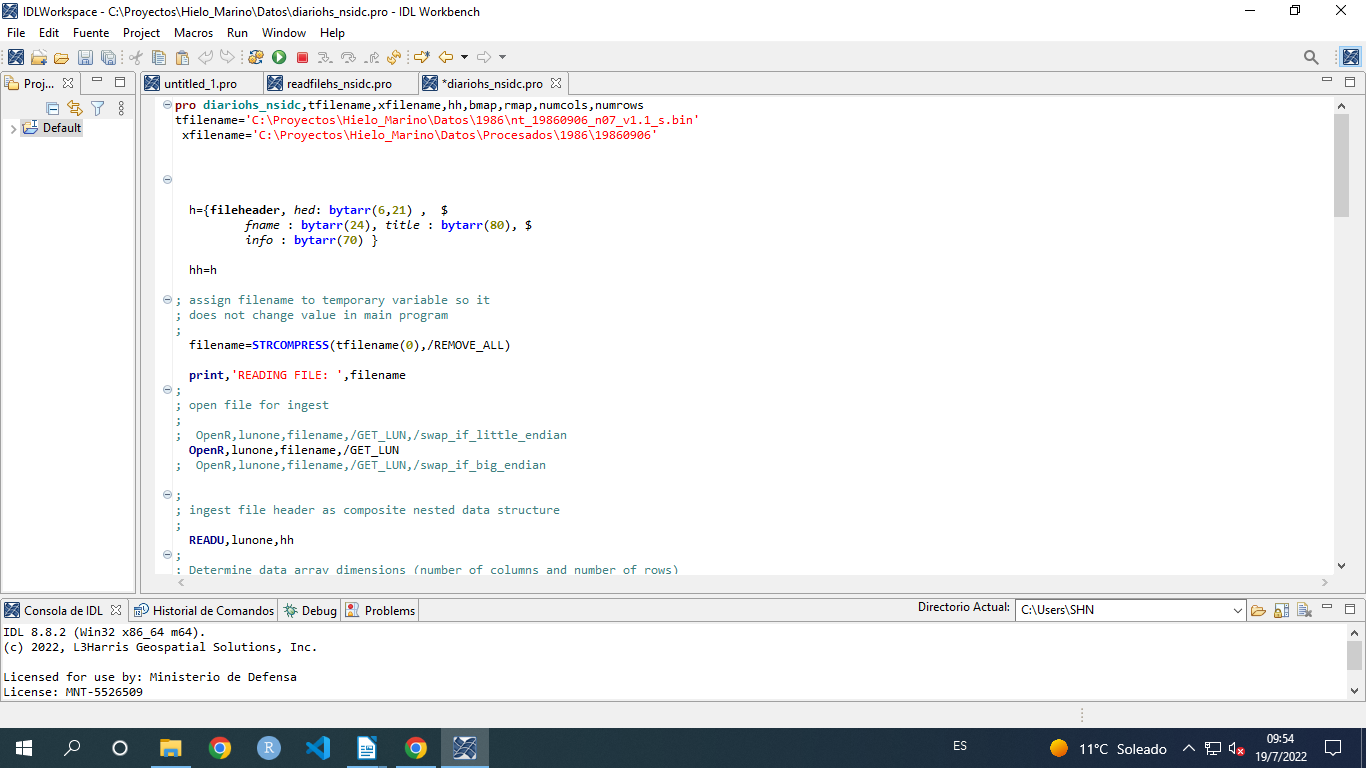
**Actualización de la base de datos**

Se utiliza el set de datos *Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data* (Cavalieri,1996), que provee datos de concentración de HM antárticos (<https://nsidc.org/data/NSIDC-0051/versions/1>). Es generado con el *Team Algorithm* de la NASA a partir de información satelital y distribuido por el Centro de Hielo y Nieve de los Estados Unidos (NSIDC, por sus siglas en inglés).

Los datos se proporcionan en la proyección estereográfica polar con una resolución espacial de 25 x 25 km. Comienzan en octubre de 1978 y se extienden hasta diciembre de 2021. Cada año NSIDC agrega datos del año anterior, luego de que estos hayan pasado un control de calidad. La actualización permite sumar nuevos datos diarios y mensuales a la base de datos que alimenta el modelo y el pronóstico diario.

Para poder leer lo datos es necesario procesarlos con programas en un software de IDL (Interactive Data Language), que tiene un lenguaje de programación que permite descomprimir y decodificar la información. La licencia de IDL fue adquirida gracias a un subsidio otorgado al proyecto PIDDEF que tiene como objetivo desarrollar el pronóstico de HM basándose en redes neuronales. Actualmente se está trabajando en mejorar la performance de los programas en IDL para poder procesar los datos diarios en menor tiempo. En la Figura 1 se muestra la interfaz de IDL con parte del código que se utiliza para procesar los datos diarios. Primero se abren los archivos binarios, se interpreta el encabezado y luego se identifican todos los valores que no sean de concentración de hielo válidos (tierra, costa, datos faltantes, puntos sobre el polo). El archivo de salida es utilizado por un programa que genera una base de datos en R con los valores diarios de HM de todos los años para cada punto de la Antártida.

Recientemente se terminaron de procesar los años 2018, 2019, 2020 y 2021 y ya fueron incorporados en las corridas mensuales y diarias. También se sumaron datos faltantes de HM de los años anteriores, que no habían sido procesados. Actualmente se cuentan con 14 GB de datos diarios procesados.

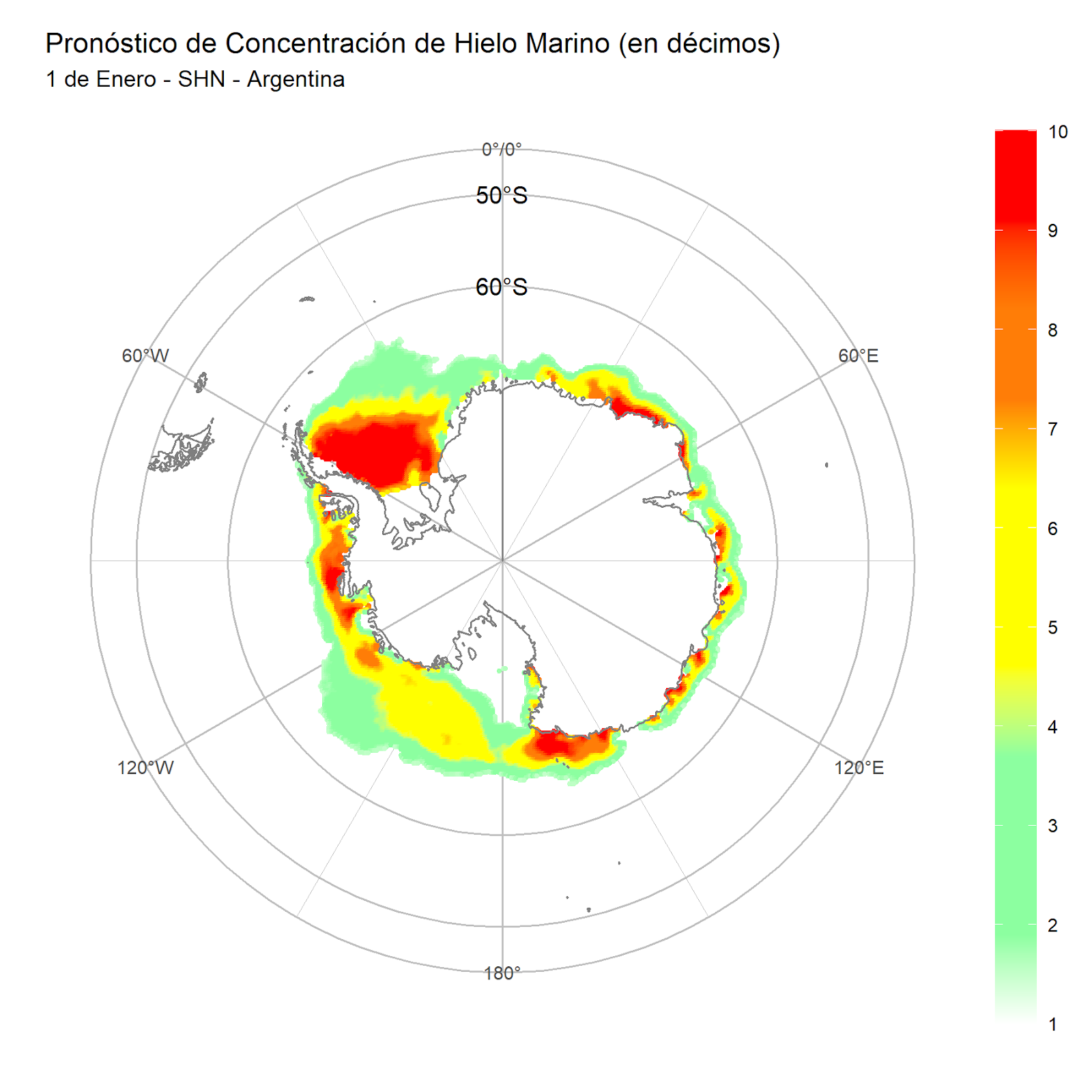


**Figura 1** - Interfaz de IDL que permite editar el código utilizado para procesar los datos satelitales de concentración de HM.

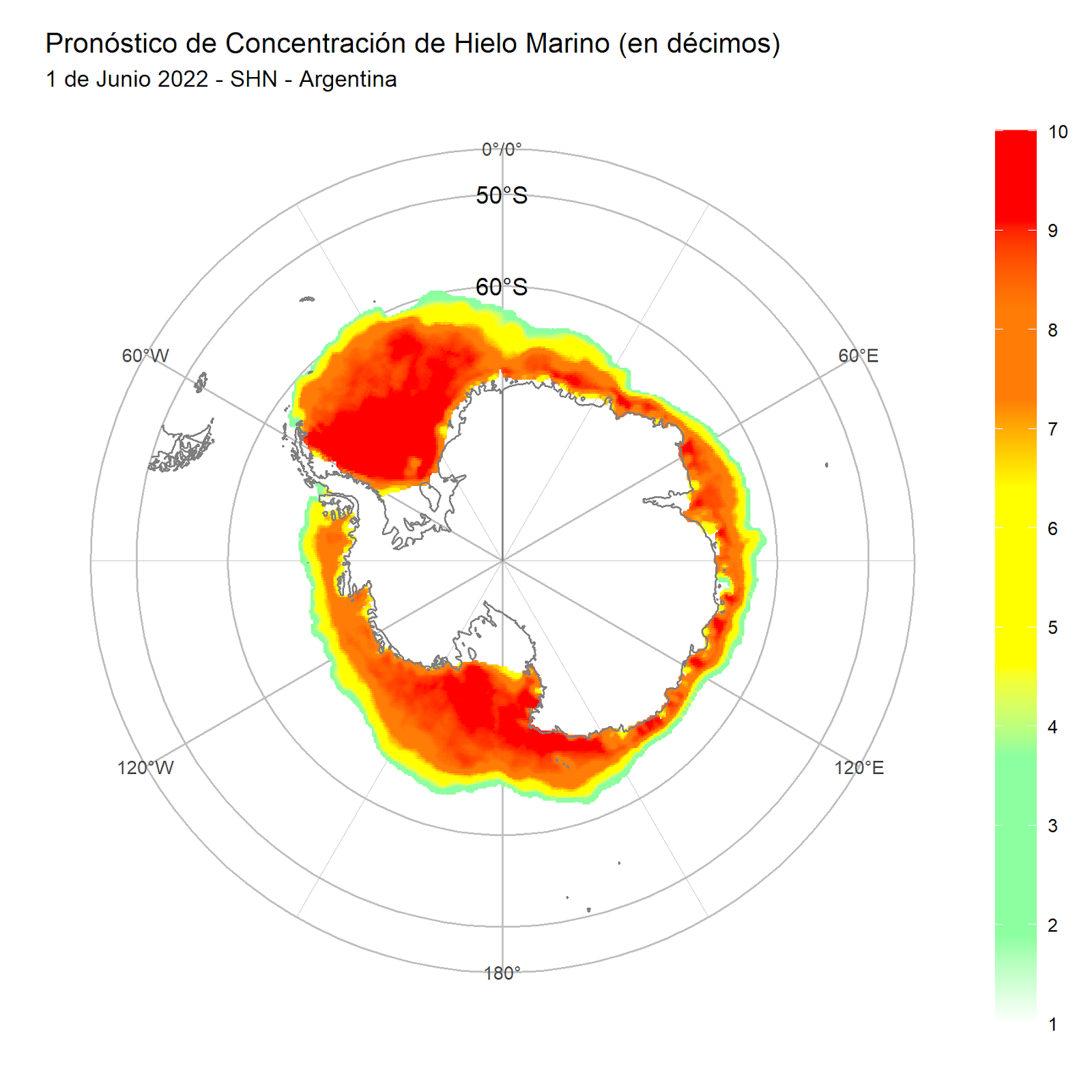
**Pronóstico climatológico de HM**

El modelo de pronóstico climático de concentración de HM y de variables meteorológicas se basa en técnicas de Machine Learning: análisis de componentes principales con aprendizaje supervisado y propagación hacia atrás para una continua actualización del sistema de pronóstico.

El modelo calcula campos patrones de concentración de HM antártico a partir de la base de datos procesados utilizando el análisis por componentes principales. Luego genera una predicción de los patrones más probables que podrían ocurrir en los próximos meses. El pronóstico diario compone campos de concentración de HM con campos reales de la base de datos para meses y años específicos asociados a los patrones más probables. El pronóstico mensual compone campos de concentración de HM, anomalías de HM y también de variables meteorológicas (temperatura, precipitación y presión) a partir de los campos reales asociados a los patrones más probables. La relación entre los campos de HM y las variables meteorológicas ha sido probada en Barreira (2012). El pronóstico se corre mensualmente y general campos mensuales a tres meses en el caso del pronóstico mensual y campos diarios a un mes en el caso del pronóstico diario que son graficados y posteriormente analizados.

En las Figuras 2 y 3 se muestran campos de concentración de HM, que fueron generados con el pronóstico diario en diciembre de 2021 y en mayo de 2022 para el mes siguiente.   
 

**Figura 2** - Campo de concentración de HM pronosticado para el 1 de enero de 2022.



**Figura 3** - Campo de concentración de HM pronosticado para el 1 de junio de 2022

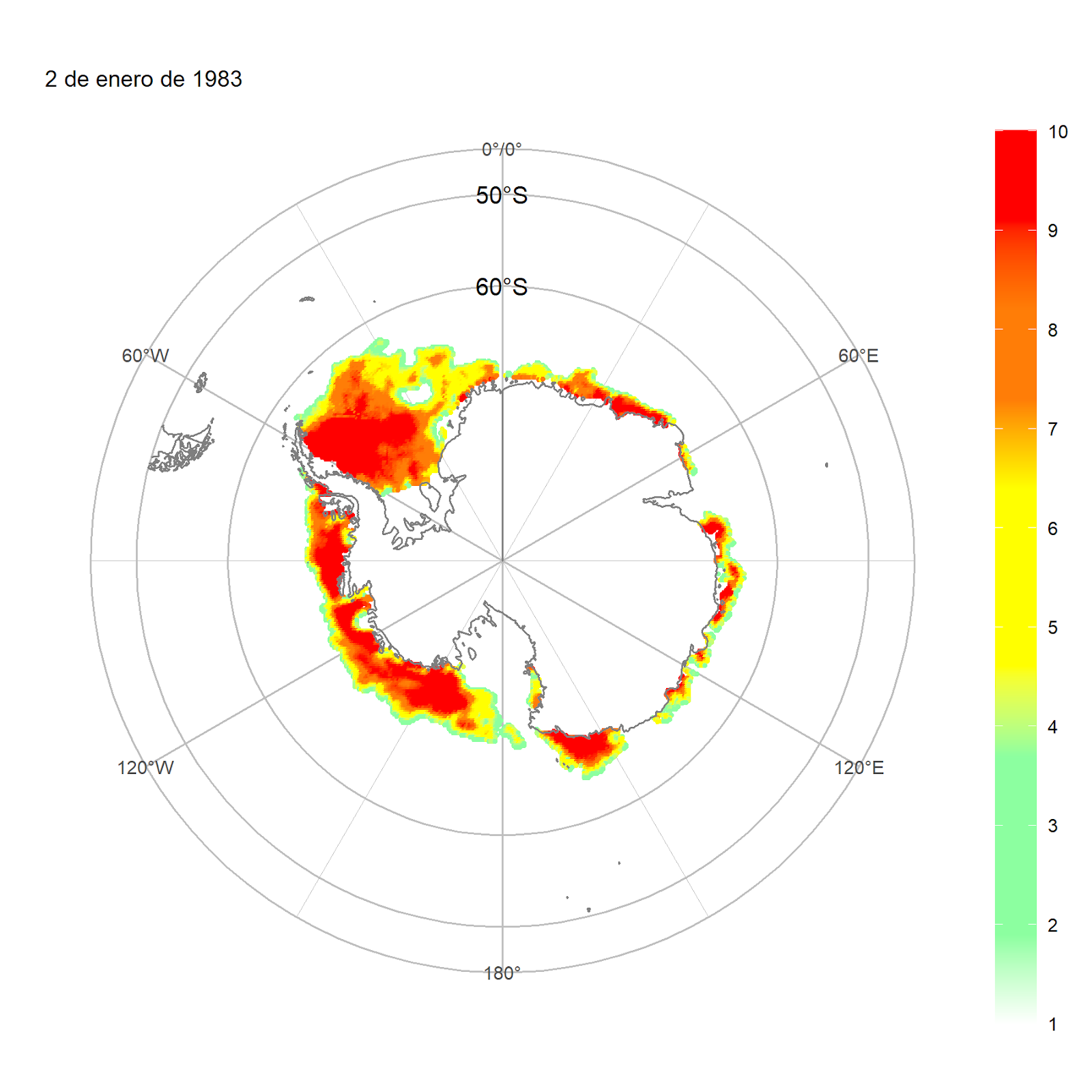
**Actualización del pronóstico diario**

Los datos de concentración de HM provienen de distintas misiones satelitales que fueron mejorando la resolución temporal y espacial. Se cuenta con datos cada dos días entre octubre de 1978 y agosto de 1987, y con datos diarios desde agosto de 1987 hasta la actualidad. Esta diferencia en la cobertura temporal no permite que el pronóstico pueda realizar predicciones usando los años previos al 88.

Para que el programa pueda usar la base completa, se realizó una interpolación lineal entre los datos para obtener concentraciones diarias. Cada archivo procesado tiene 104912 puntos entre 39,4ºS y 89,8ºS con porcentaje de concentraciones de HM. La interpolación de las concentraciones de cada punto entre dos días seguidos disponibles entre 1978 y 1987 permitió obtener una base de datos extendida que cubre todo el período satelital. Desde mayo de 2022 las corridas del pronóstico diario han implementado la base generada por interpolación.

El programa de pronóstico diario tiene que ser capaz de trabajar con una gran cantidad de datos. La base disponible se va ampliando a medida que se incorporan nuevos años procesados. Actualmente se está trabajando en modificar el procedimiento que sigue el programa en asimilar la base de datos debido a limitaciones en la memoria. De esta manera se optimizará la asimilación de bases cada vez más grandes. La reciente base de datos ampliada debido a la interpolación y debido a la incorporación de nuevos años procesados pudo ser asimilada en el pronóstico diario gracias a algunas mejoras en el código que la levanta, evitando que el tamaño de la misma produzca problemas en las corridas.

En la Figura 4 se muestra un ejemplo de un campo de HM interpolado a partir de dos campos reales consecutivos que ya es utilizado en las nuevas corridas. Corresponde al 2 de enero de 1983.



**Figura 4** - Campo de concentración de HM del 2 de enero de 1983 obtenido por interpolación de los campos reales de HM de los días 1 y 3 de enero del mismo año.

**Apéndice**

A continuación se muestra una parte del código que crea una base de datos en R a partir de datos satelitales para el año 1978. Los datos reales son cada dos días y se crea una nueva base con datos diarios interpolados.

lat <- read.table(file="pss25lats.txt",col.names = "Lat")

lon <- read.table(file="pss25lons.txt",col.names = "Long")

area <- read.table(file="pss25area.txt",col.names = "Area")

# 1978

a1978 <- list.files(path = "Procesados/1978/", full.names = TRUE, recursive = TRUE)

datos\_1978 <- read.table(a1978[1])

for (i in 2:length(a1978)){

datos\_1978[,i] <- read.table(file=a1978[i])

}

datos\_1978 <- cbind(lat,lon,area,datos\_1978)

datos\_1978 <- datos\_1978[- which(datos\_1978[,5]==-7777.0) , ]

datos\_1978 <- datos\_1978[- which(datos\_1978[,5]==-8888.0) , ]

date <- seq(as.Date("1978-10-26"), as.Date("1978-12-31"), by="days")

d <- dim(datos\_1978)

datos\_1978\_i <- datos\_1978[4]

for (i in 5:d[2]) {

datos\_1978\_i <- cbind(datos\_1978\_i,rowMeans(cbind(datos\_1978[,i-1], datos\_1978[,i])),datos\_1978[i])

}

colnames(datos\_1978\_i) <- date

datos\_1978 <- cbind(datos\_1978[,1:3],datos\_1978\_i)

##### Lista ####

datos\_lista <-list()

datos\_lista[[1]] <- datos\_1978

**Referencias**

Barreira, Sandra. 2012. Variabilidad espacial y temporal del hielo marino antártico. Relación con la circulación atmosférica. Editorial Académica Española. ISBN: 978-3-659-02325-5

Cavalieri, D. J., C. L. Parkinson, P. Gloersen, and H. J. Zwally. 1996, updated yearly. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. Doi: https://doi.org/10.5067/8GQ8LZQVL0VL.