# Introducción a la aplicación web Pampero



# Guzmán López\* Asistente en manejo de datos oceanográficos Proyecto FREPLATA URU/09/G31

Diciembre, 2013

#### Resumen

La aplicación web Pampero está diseñada para permitirle al usuario visualizar y analizar de manera interactiva datos de viento. Fue desarrollada en el marco del Proyecto FREPLATA URU/09/G31 dentro del "Programa de Monitoreo y Evaluación y Sistema de Información Integrado y establecido para la toma de decisiones y la Gestión del Río de la Plata y su Frente Marítimo". El objetivo era generar una herramienta que permitiese a los usuarios procesar datos de viento provenientes de la estación meteorológica instalada en la Estación de Monitoreo Ambiental del Río de la Plata Boya oceanográfica. Todo el software y librerías empleadas en esta aplicación son gratuitas, y la mayor parte del software empleado es libre. Para su desarrollo se utilizó fundamentalmente el software R (estadística computacional y gráficos) y la librería Shiny (marco de aplicaciones web para R).

\*contacto: <a href="mailto:guzilop@gmail.com">guzilop@gmail.com</a>

# Índice de contenido

1 - Introducción	3
1.1 - Justificación	3
2 - La aplicación web Pampero	3
2.1 - Descripción técnica	3
2.2 - Requerimientos e instalación	4
2.2.1 - Modo local	4
2.2.2 - Modo red	
2.3 - Utilizando Pampero	
2.3.1 - Descripción de la interfaz general de usuario (GUI)	6
2.3.2 - Cargar datos desde un archivo de texto	
2.3.3 - Cargar datos desde un archivo de Microsoft Excel	8
2.3.4 - Cargar datos desde Base de datos	g
2.3.5 - Tabla de datos	10
2.3.6 - Rosa de los vientos	11
2.3.7 - Serie de tiempo	12
2.3.8 - Resumen de los datos	14
2.3.9 - Acerca de esta aplicación web	14
3 - Código fuente	15
3.1 - script server.R	15
3.1 - script ui.R	15
1 - Peferencias	15

## 1 - Introducción

## 1.1 - Justificación

El proyecto FREPLATA URU/09/G31 a través de su Programa de Monitoreo y un Sistema Binacional Integrado de Información para apoyar la toma de decisiones, busca conformar y consolidar un Sistema de Información Ambiental Costero Marino (SIACM) a nivel nacional para el Río de la Plata y su Frente Marítimo (RPFM). Con esta intención, es que ha evaluado las necesidades de información e intereses por parte de las instituciones competentes en la gestión sobre el RPFM con miras de fomentar las capacidades institucionales y la articulación de estrategias. En tal sentido se contrató a tres asistentes en manejo de datos oceanográficos para que trabajen en el diagnóstico de la información existente y el diseño de bases de datos del Servicio de Oceanografía Hidrografía y Meteorología de la Armada (SOHMA) y de la Dirección Nacional de Recursos Acuáticos (DINARA). Para facilitar el acceso, visualización y análisis de los datos de viento provenientes de la estación meteorológica instalada en la estación de monitoreo ambiental Boya oceanográfica se desarrolló la aplicación web Pampero. Buscando que usuarios no expertos puedan acercarse a los datos de viento mediante una interfaz general amigable e interactiva.

# 2 - La aplicación web Pampero

## 2.1 - Descripción técnica

La aplicación web Pampero consta de un directorio con dos scripts: *server.R* y *ui.R*. En el script *server.R* se asignan las entradas para los cálculos a realizarse y se le asignan expresiones reactivas, y en el script *ui.R* se define la interfaz general de usuario de la aplicación. Además, dichos scripts también incluyen información adicional y accesos a

recursos que requiera la aplicación para su funcionamiento. Por más información, en la sección 4 – Código fuente de esta guía se pueden descargar ambos scripts.

## 2.2 - Requerimientos e instalación

La aplicación web Pampero puede ejecutarse localmente o a través de la red con conexión a Internet.

#### 2.2.1 – Modo local

Para correr la aplicación web Pampero localmente se requiere instalar la siguiente lista de software y librerías de R:

- R (versión > 2.15.2)
- librería "Shiny, Web Application Framework for R" (versión 0.8.0)
- librería "openair, Open-source tools for the analysis of air pollution data" (versión 0.9-0)
- librería "gnumeric, Read data from files readable by gnumeric" (versión 0.7-2)
- librería "DBI, R Database Interface" (versión 0.2-7)
- librería "RmySQL, R interface to the MySQL database" (versión 0.8-0)
- librería "googleVis, Interface between R and the Google Chart Tools" (versión 0.4.7)
- Administrador de Sistemas de Bases de datos Relacionales: MariaDB (5.5.34) o
   MySQL (versión 5.5.32)
- Explorador web (ej. Mozilla Firefox, Chromium, Google Chrome, etc).
- Adobe Flash Player plugin.

Para ejecutar la aplicación Pampero se debe ejecutar la siguiente sentencia en la línea de comandos de R: shiny::runApp(appDir='/directorio/Pampero/'). El directorio corresponde a la ubicación donde se encuentra la carpeta "Pampero" con los archivos server.R y ui.R.

Nota: el explorador web debe tener habilitado el Adobe Flash Player plugin y permitir JavaScript.

#### 2.2.2 - Modo red

Para correr la aplicación web Pampero a través de la red se requiere configurar y mantener un servidor Linux. Además se necesita instalar la siguiente lista de software y librerías de R en el servidor Shiny:

- R (versión > 2.15.2)
- librería "Shiny, Web Application Framework for R" (versión 0.8.0)
- Shiny Server software (http://www.rstudio.com/shiny/server/)
- librería "oce, Analysis of Oceanographic data" (versión 0.9-12)
- librería "gnumeric, Read data from files readable by gnumeric" (versión 0.7-2)
- librería "DBI, R Database Interface" (versión 0.2-7)
- librería "RmySQL, R interface to the MySQL database" (versión 0.8-0)
- librería "googleVis, Interface between R and the Google Chart Tools" (versión 0.4.7).
- Administrador de Sistemas de Bases de datos Relacionales: MariaDB (5.5.34) o
   MySQL (versión 5.5.32)

Utilizando un servidor Linux los usuarios acceden a la aplicación web Pampero a través

Introducción a la aplicación web Pampero

Página 5

de una página web.

Otra alternativa es acceder mediante el repositorio GitHub donde se almacena el código fuente ejecutando la siguiente sentencia en la línea de comandos de R: shiny::runGitHub(repo='Pampero', username='guzmanlopez').

**Nota:** como instalar Shiny Server (<a href="http://www.rstudio.com/shiny/server/install-opensource">http://www.rstudio.com/shiny/server/install-opensource</a>) y como configurar Shiny Server (<a href="http://rstudio.github.io/shiny-server/latest/">http://rstudio.github.io/shiny-server/latest/</a>).

## 2.3 - Utilizando Pampero

La aplicación web Pampero está diseñada para permitirle al usuario visualizar y analizar de manera interactiva datos de viento.

## 2.3.1 - Descripción de la interfaz general de usuario (GUI)

La aplicación web Pampero consta de un panel lateral ubicado a la izquierda de la pantalla donde se especifican los controles de entrada y de un panel principal ubicado a la derecha del panel lateral con una serie de pestañas [Tabla], [Rosa de los vientos], [Serie de tiempo], [Resumen] y [Acerca de esta APP] donde se muestran las salidas (Figura 1).

#### 2.3.2 - Cargar datos desde un archivo de texto

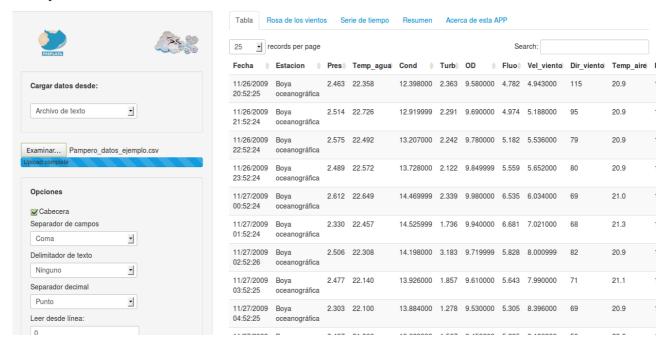
Para cargar datos de un archivo de texto, dirigirse a la pestaña [**Tabla**] del panel principal y en **Cargar datos desde:** seleccionar **Archivo de texto**. Hacer clic en el botón <**Examinar**> y seleccionar el archivo deseado (Figura 1).

En la sección de **Opciones** se deben especificar las características del archivo para que los datos sean cargados correctamente en el panel principal. La opción **Cabecera** se

debe seleccionar si el archivo cargado posee columnas con títulos. La opción Separador de campos hace referencia a qué símbolo define las celdas, y estos pueden ser *Coma*, *Punto y coma*, o *Tabulador*. La opción Delimitador de texto hace referencia a cómo se delimitan los campos de texto y estos pueden ser *Ninguno*, *Comillas dobles*, o *Comillas simples*. Según la configuración regional del archivo seleccionado la opción Separador decimal puede ser *Punto* o *Coma*. La opción *Leer desde línea:* define desde que línea de texto se comienza a leer el archivo y es un valor numérico entero que debe definir el usuario. Por último se debe establecer el *Formato de Fecha y Hora*, cuyas opciones varían en función del orden de aparición de los campos (día, mes y año o mes, día y año), de los delimitadores de campos de la fecha ("/" o "-") y de si el año está por extenso (ej. "2009" o "09").

Las opciones establecidas por defecto en Pampero para abrir archivos de texto son: con cabecera, campos separados por comas, sin delimitador de campos de texto, con punto como separador decimal, leer desde la línea cero (no excluir ninguna línea de texto) y formato de fecha y hora como "mes/día/año en extenso y hora".

Los archivos cargados deben tener los datos de fecha y hora en la primer columna y los datos de la estación meteorológica en la segunda columna como se muestra en el ejemplo de la Figura 1.



**Figura 1 -** Captura de pantalla de como cargar datos desde un archivo de texto con Pampero. A la izquierda con fondo gris oscuro se ubica el panel lateral donde se configuran las entradas a la aplicación (sección Opciones). A la derecha, donde se visualizan los datos se ubica el panel principal donde se muestran la salidas de la aplicación, en este caso una tabla con los datos cargados.

## 2.3.3 - Cargar datos desde un archivo de Microsoft Excel

Para cargar datos de un archivo de Microsoft Excel, dirigirse a la pestaña [**Tabla**] del panel principal y en **Cargar datos desde:** seleccionar **Archivo de Microsoft Excel**. Hacer clic en el botón **Examinar** y seleccionar el archivo deseado (solo terminación ".xls").

En la sección de **Opciones** se deben especificar las características del archivo para que los datos sean cargados correctamente en el panel principal. La opción **Cabecera** se debe seleccionar si el archivo cargado posee columnas con títulos. En la opción **Nombre de hoja** el usuario debe escribir el nombre de la planilla en el campo de texto vacío. El nombre de la planilla es aquel que aparece en la región inferior cuando el

archivo es abierto con un programa que abre hojas de cálculo (ej. Microsoft Excel o LibreOffice Calc). En general, por defecto toma el nombre de "Sheet1" u "Hoja1". El nombre debe ser escrito sin comillas. Por último se debe establecer el *Formato de Fecha y Hora*, cuyas opciones varían en función del orden de aparición de los campos (día, mes y año o mes, día y año), de los delimitadores de campos de la fecha ("/" o "-") y de si el año está por extenso (ej. "2009" o "09").

Las opciones establecidas por defecto en Pampero para abrir archivos de Microsoft Excel son: con cabecera y formato de fecha y hora como "mes/día/año en extenso y hora".

Los archivos cargados deben tener los datos de fecha y hora en la primer columna como se muestra en el ejemplo de la Figura 1.

**Nota:** la opción cargar datos desde un archivo de Microsoft Excel es experimental y puede presentar fallas como no lograr cargar el archivo elegido. Es un complemento ya que siempre es posible convertir el archivo de Microsoft Excel a texto (formato no restrictivo) y ser cargado con la opción de cargar datos desde un archivo de texto.

#### 2.3.4 - Cargar datos desde Base de datos

Para cargar datos desde una Base de datos, dirigirse a la pestaña [**Tabla**] del panel principal y en **Cargar datos desde:** seleccionar **Base de datos** (Figura 2). Luego se debe seleccionar una **Estación meteorológica** que puede ser **Boya oceanográfica**, **Pilote Norden** o **Torre Oyarvide** y el lapso de tiempo para consultar en **Fechas**, **Hora inicial** y **Hora Final**. Apretando el botón **Enviar consulta**> obtenemos los datos elegidos. Los datos consultados pueden descargarse como una planilla de texto (.csv) apretando el botón **Descargar datos**>.

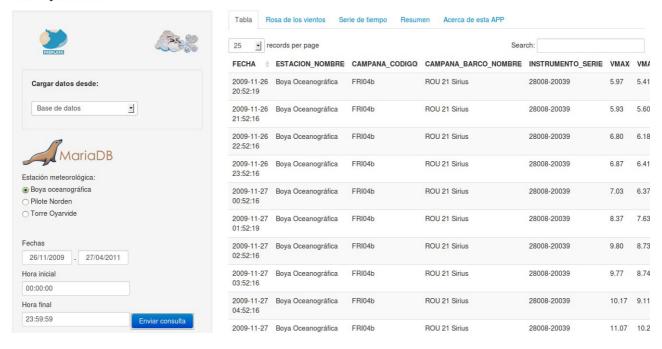


Figura 2 - Captura de pantalla de como cargar datos desde la Base de datos con Pampero.

#### 2.3.5 - Tabla de datos

Una vez que se hayan cargado los datos y se visualizan los campos correctamente en el panel principal (Figura 1 y Figura 2), Pampero permite generar algunos filtros básicos de los mismos. Se puede elegir visualizar entre 10, 25, 50 o 100 registros de datos por página con al opción **records per page**, permite buscar un valor determinado en el campo **Search**, y se pueden ordenar la tabla de datos por columna en orden ascendente o descendente haciendo clic sobre el título de la columna. La búsqueda también se puede hacer por columna introduciendo un valor en los campos vacíos que están al final de cada una. Con los botones **Previous**> y **Next**> se pasa a la hoja previa o a la siguiente respectivamente. Con los botones numerados entre **Previous**> y **Next**> se accede directamente a la hoja elegida.

Nota: los filtros establecidos en esta sección son meramente informativos y no son aplicados en

procesos posteriores ni sobre los datos originales. Pampero no modifica ni escribe sobre el archivo cargado.

#### 2.3.6 - Rosa de los vientos

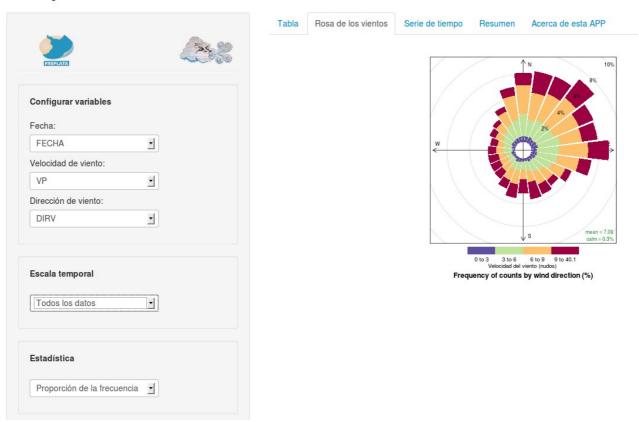
En la pestaña [Rosa de los vientos] es posible visualizar gráficos de rosas de los vientos para los datos cargados (Figura 3). La rosa de los vientos es un diagrama que combina la frecuencia en que el viento sopla de una determinada dirección junto con su intensidad (velocidad de viento) para un lapso de tiempo determinado. Además reporta el promedio de la velocidad de viento y el porcentaje del tiempo en que se observó un estado de calma según la escala Beaufort (velocidad de viento menor a 1 nudo).

Para visualizar la rosa de los vientos se debe seleccionar las entradas desde los datos cargados en la sección Configurar variables (Fecha:, Velocidad de viento: y Dirección de viento:). Luego en la sección Escala temporal podemos elegir visualizar una rosa de los vientos a partir de *Todos los datos* cargados o los datos cargados discriminados por año (*Anual*), estación (*Estacional*), mes (*Mensual*) u hora (*Horaria*). En la sección Estadística podemos elegir el cálculo aplicado a cada clase establecida. Si elegimos *Frecuencia* visualizamos el porcentaje en que el viento sopló de esa dirección, si elegimos *Contribución al promedio* visualizamos cuánto contribuye cada clase al promedio general de la velocidad de viento para el período y si elegimos *Conteo absoluto* visualizamos la cantidad total de datos para cada clase.

En la sección **Configuración de plot** podemos cambiar el intervalo (**Intervalo:**, entre 1 y 10) de cada clase de velocidad del viento, la cantidad de divisiones que queremos visualizar para del rango de la velocidad de viento (**Quiebres:**, entre 1 y 10) y las clases de dirección de viento (**Ángulo:**, entre 1 grado y 45 grados).

Los valores por defecto de la rosa de los vientos son todos los datos en la escala temporal, proporción de la frecuencia en estadística, intervalo de 3 unidades, 4 quiebres y un angulo de 15 grados para definir las clases de direcciones de viento.

# Pampero APP



**Figura 3 -** Captura de pantalla de Rosa de los vientos para datos cargados desde la Base de datos y con las opciones establecidas por defecto.

## 2.3.7 - Serie de tiempo

En la pestaña [**Serie de tiempo**] podemos visualizar la serie temporal de una o más variables de los datos cargados (Figura 4). Para ello se debe seleccionar las entradas desde los datos cargados en la sección **Configurar variables** (**Fecha:** y **Variables**).

El plot **Series de tiempo** es un gráfico interactivo que cuenta con las siguientes opciones:

- Hacer zoom sobre el gráfico seleccionando una escala temporal prefijada por defecto en un minuto ("1"), cinco minutos ("5"), una hora ("1h"), un día ("1d"), cinco días ("5d"), un mes ("1m"), tres meses ("3m"), seis meses ("6m"), un año ("1y"), y la escala máxima que abarcan los datos cargados ("Max").
- Haciendo clic con el botón izquierdo del ratón sobre el panel principal del gráfico y moviendo el ratón se avanza o retrocede en el tiempo.
- Con la rueda del ratón aumentamos el zoom o lo disminuimos.
- Se puede seleccionar el rango de tiempo que se desee utilizando el panel inferior del gráfico, de manera de estirar o contraer el selector desde sus bordes o desplazándolo.
- Pasando el ratón sobre las gráficas nos arroja el valor que toman las variables.



**Figura 4 -** Captura de pantalla de Serie de tiempo para datos de velocidad promedio de viento (azul) y temperatura del aire (rojo).

#### 2.3.8 - Resumen de los datos

En la pestaña [**Resumen**] es posible obtener estadísticas descriptivas básicas ("Promedio", "Desvío estándar", "Mediana", "Valor máximo", "Valor mínimo", "Cantidad de datos") de cada una de las variables (Figura 5).

#### Pampero APP Rosa de los vientos Serie de tiempo Resumen Acerca de esta APP n VMAX 4.17 8.77 56.83 0 11865 9.13 VMAXE 7.71 41.25 11865 7.94 3.54 0 VMIN 6.35 6.49 3.09 38.72 0 11865 VP 40.1 7.04 3.17 6.9 0 11865 DIRV 149.42 359 11865 DIRVR 359 0 11865 147.3 98.58 119 TEMP 0 11865 18.22 4.74 18.6 28 PRES 1014.3 1040.9 0 11865 1014.2 27.16 HUM 77.7 95.5 0 11865 76.02 10.89

**Figura 5 -** Captura de pantalla de Resumen de los datos cargados a la aplicación donde se muestran estadísticas descriptivas básicas.

PRECA

#### 2.3.9 - Acerca de esta aplicación web

En la pestaña [Acerca de esta APP] es posible acceder a una descripción general de la aplicación (Figura 6). En la misma se especifica para qué sirve Pampero, en el marco del proyecto en que se estuvo desarrollando y el objetivo de su desarrollo. Incluye también las referencias bibliográficas a las distintas librerías de R utilizadas para correr la aplicación y un contacto con el autor del desarrollo de la aplicación para que si a un usuario le surgen dudas pueda consultarlas o pueda reportar "bugs".

0 11865

54.57



**Figura 6 -** Captura de pantalla de Acerca de esta APP. En esta parte se describe la aplicación web y se puede acceder mediante hipervínculos a esta guía de usuario, al código fuente y a las citas del software y librerías utilizadas.

# 3 - Código fuente

#### 3.1 - script server.R

https://github.com/guzmanlopez/Pampero/blob/master/server.R

#### 3.1 - script ui.R

https://github.com/guzmanlopez/Pampero/blob/master/ui.R

#### 4 - Referencias

**David A. James y Saikat DebRoy (2011).** RMySQL: R interface to the MySQL database. R package version 0.8-0. <a href="http://CRAN.R-project.org/package=RMySQL">http://CRAN.R-project.org/package=RMySQL</a>

David Carslaw y Karl Ropkins (2013). openair: Open-source tools for the analysis of

air pollution data. R package version 0.9-0. <a href="http://CRAN.R-project.org/package=openair">http://CRAN.R-project.org/package=openair</a>

**Karoly Antal. (2012).** gnumeric: Read data from files readable by gnumeric. R package version 0.7-2. <a href="http://CRAN.R-project.org/package=gnumeric">http://CRAN.R-project.org/package=gnumeric</a>

**Markus Gesmann & Diego de Castillo.** Using the Google Visualisation API with R. The R Journal, 3(2):40-44, December 2011.

**R Core Team (2013).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <a href="http://www.R-project.org/">http://www.R-project.org/</a>

**RStudio and Inc. (2013).** shiny: Web Application Framework for R. R package version 0.8.0. <a href="http://CRAN.R-project.org/package=shiny">http://CRAN.R-project.org/package=shiny</a>

R Special Interest Group on Databases (2013). DBI: R Database Interface. R package version 0.2-7. http://CRAN.R-project.org/package=DBI