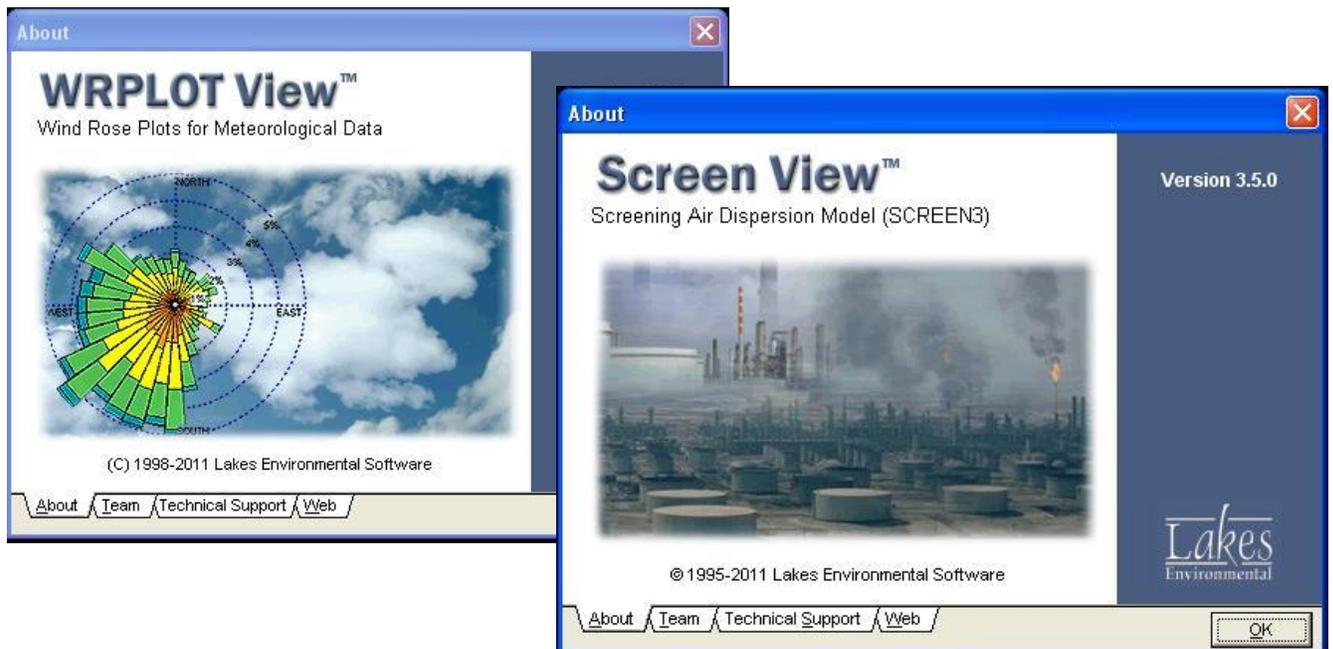


MANUAL PARA EL MANEJO DE SOFTWARE ESPECIALIZADO



**LENGUAJE R APLICADO AL
ANÁLISIS DE DATOS DE
CALIDAD DEL AIRE**

MIPA Sergio Ramos Herrera
MISA Elizabeth Magaña Villegas
MIA Jesús Manuel Carrera Velueta



INDICE

1 Software WRPLOT View Freeware 7.0.0	3
2 Software Screen View	13
3. R y el paquete OPENAIR.....	19



1 Software WRPLOT View Freeware 7.0.0

WRPlot es un software gratuito para la construcción de Rosas de Viento para datos meteorológicos. Proporciona una vista de gráficas de rosas de viento, análisis de frecuencias, así como gráficas para diversos formatos de datos meteorológicos.

Descarga e Instalación

Pasos para la instalación.

1. Descargar el archivo de instalación de la dirección: <http://www.weblakes.com/products/wrplot/>
2. Llenar el formato de registro para obtener el código de activación (Figura 1).
3. Instalar el WRPLOT View en tu computadora (Figura 2).
4. Una vez obtenido el código de activación, iniciar el WRPlot View (Figura 3).
5. Ingresar código de activación.
6. A partir de este momento se puede empezar a usar el WRPlot View (Figura 4).

The screenshot shows a web browser window displaying the registration page for WRPLOT View. The page has a navigation menu with options like 'Company', 'Software', 'Services', 'Training', 'Support', 'Download', and 'Buy'. The main content area is titled 'WRPLOT View™ - Freeware' and includes a 'Registration and Activation' section with a registration form. The form fields are filled with the following information:

First Name:	JManuel
Last Name:	Carrera
E-mail:	jmarrera@gmail.com
Confirm E-mail:	jmarrera@gmail.com
Company:	Universidad Juárez Autónoma de Tabasco
Address:	Km 0.5 Carretera Villahermosa-Cárdenas entronque
City:	Villahermosa
State/Province:	Tabasco / Centro
Country:	México
Postal Code:	86039

Additional elements on the page include a search bar, social media icons, and a sidebar with 'SOFTWARE' and 'FREEWARE' categories. The right sidebar contains information about the product, including a version 7.0.0 release notes link and resource links for modeling data and tips.

Figura 1. Formulario de registro para la obtención del código de activación.

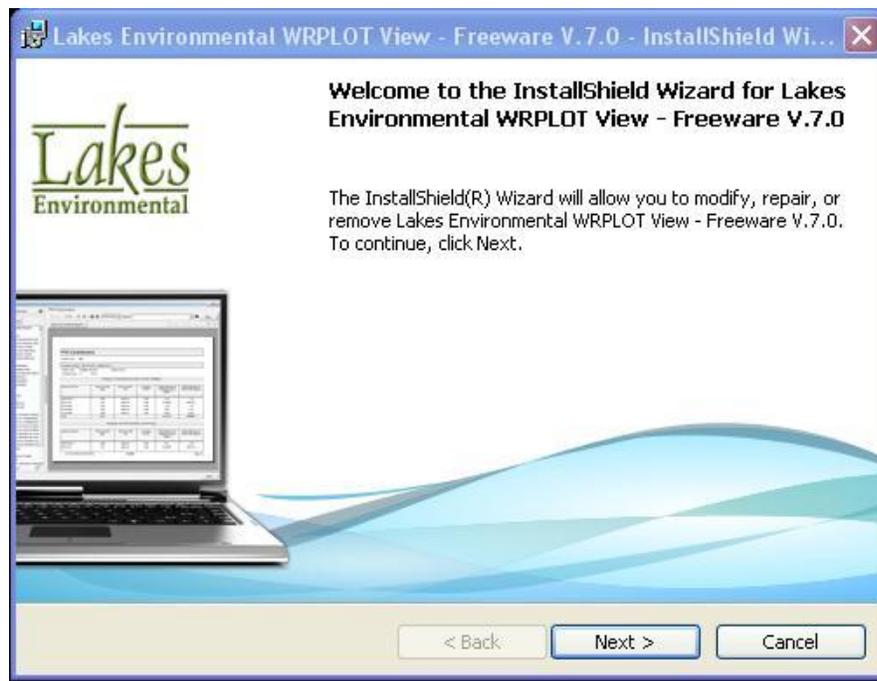
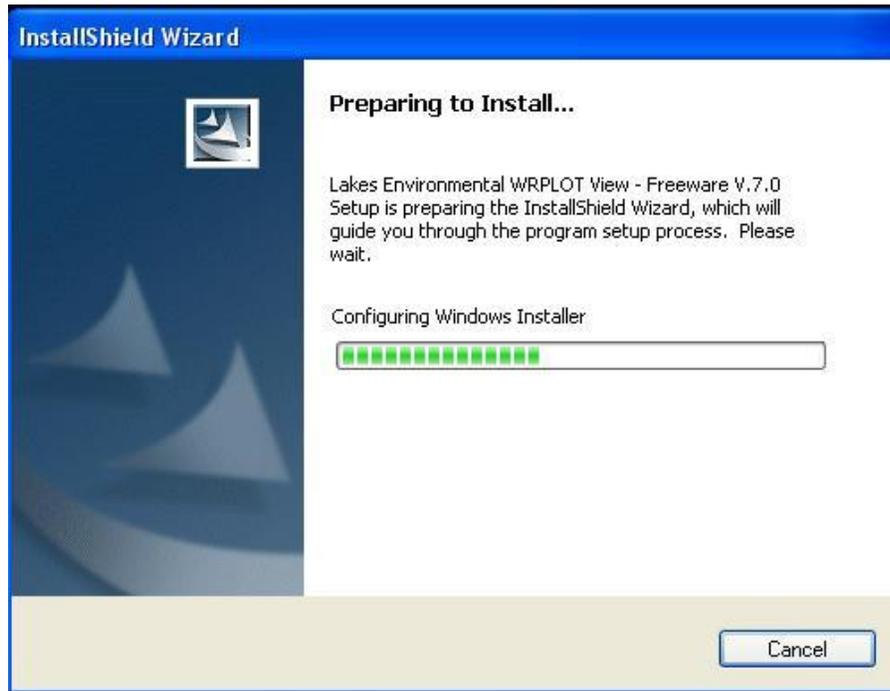


Figura 2. Algunas ventanas que se observarán durante la instalación del WRPlot.

Pulsar el botón de comando “Next >”, después de algunos instantes pulsar el botón “Finalizar”. El WRPlot está instalado.

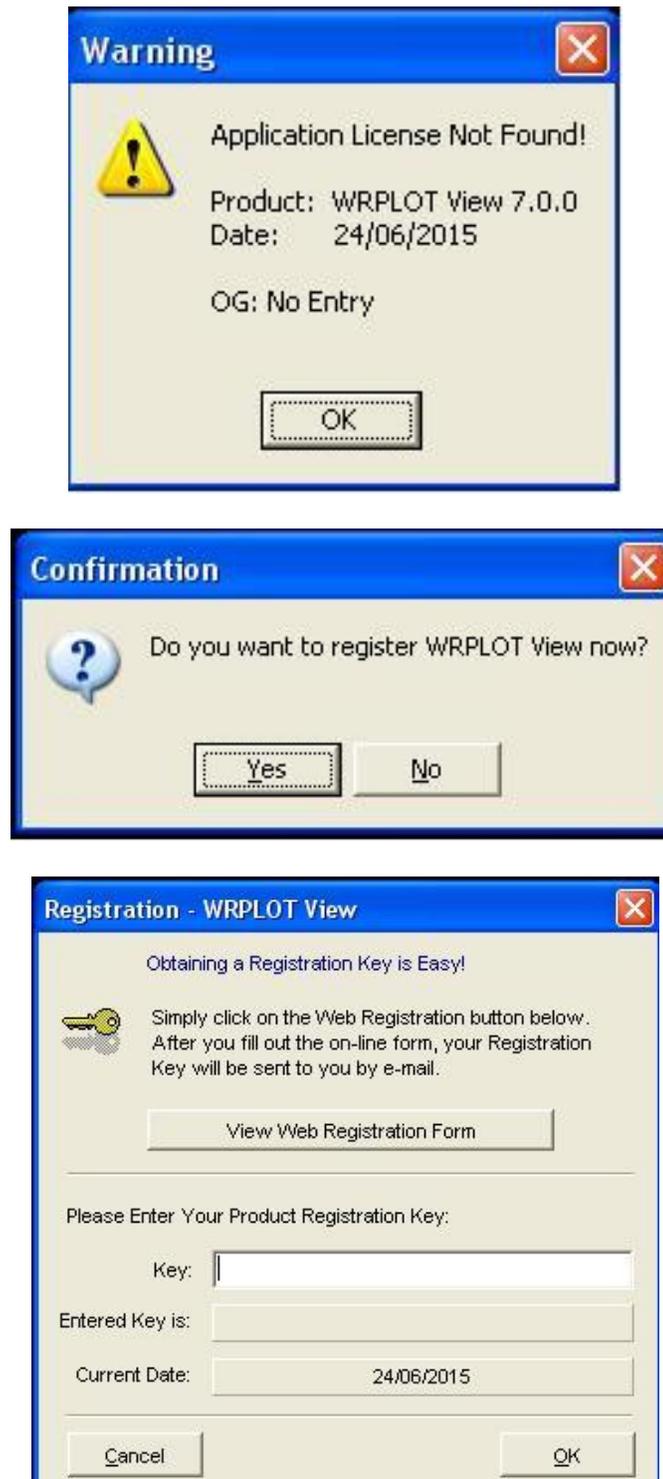


Figura 3. Ventanas que se mostrarán al iniciar WRPlot View. En la última ventana se ingresa el código de activación.

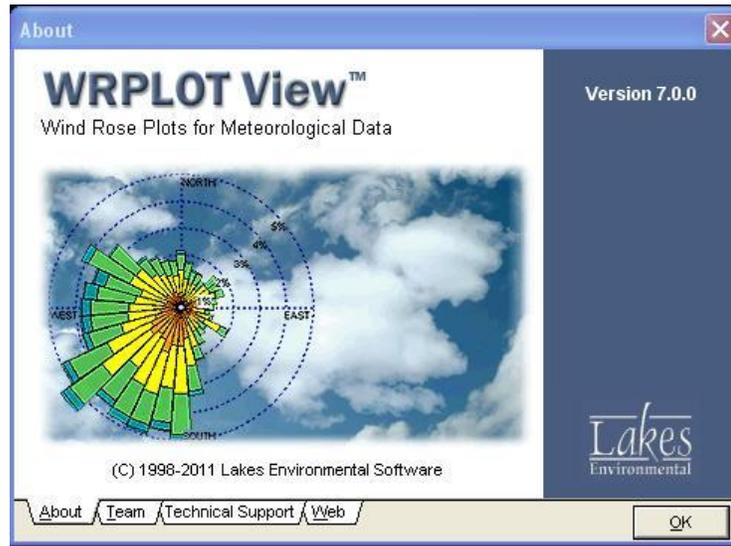


Figura 4. Pantalla de Inicio, una vez ingresado el código de activación.

Preparación de la Base de Datos.

La base de datos debe estar integrada por 7 columnas de datos: id estación, año, mes, día, hora, dirección del viento, velocidad del viento.

LAKES FORMAT							
* -----							
* Rows with * at the beginning of the line are recognized							
* as comments and are ignored by WRPLOT View							
*							
* STNID (5)							
* YEAR (4)							
* MONTH (2)							
* DAY (2)							
* HOUR (2) 0-23							
* WDIR (3) deg							
* WSPD (3) knots							
* -----							
14826	1988	1	1	0	260	17	
14826	1988	1	1	1	260	13	
14826	1988	1	1	2	260	14	
14826	1988	1	1	3	250	16	
14826	1988	1	1	4	250	15	
14826	1988	1	1	5	240	15	
14826	1988	1	1	6	240	14	

Una Rosa de Vientos representa la frecuencia de ocurrencia de los vientos en cada una de las direcciones de vientos especificadas y para cada clase de velocidad de viento para una localidad dada y durante un período de tiempo. La Figura 5 muestra la ventana principal del software.

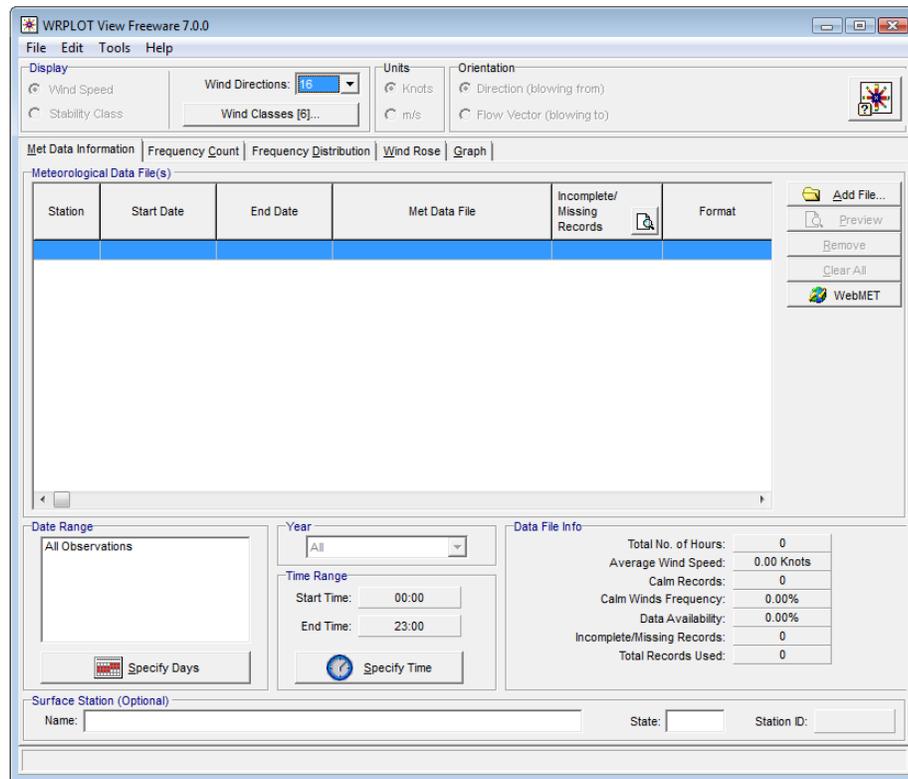


Figura 5. Ventana principal de WRPLOT View.

La Tabla 1 contiene un listado de las estaciones meteorológicas automáticas en Tabasco, cuya información esta accesible desde el portal del Servicio Meteorológico Nacional ().

Tabla 1. Estaciones meteorológicas automáticas en Tabasco

Estación	Longitud	Latitud	asnm(m)	Municipio
Paraíso	93° 09' 20"	18° 25' 23"	4	Paraíso
Sanchez Magallanes	93° 51' 52"	18° 17' 44"	5	Cárdenas
Huimanguillo	93° 23' 47"	17° 51' 04"	20	Huimanguillo
Villahermosa	92° 56' 00"	17° 59' 00"	6.5	Centro
Jonuta	92° 08' 16"	18° 05' 27"	9	Jonuta
Pantanos de Centla	92° 38' 47"	18° 24' 22"	2	Centla
Trez Brazos	92° 36' 26"	18° 23' 05"	11	
Dos Patrias	92° 47' 58"	17° 36' 42"	25	
San Pedro	91° 08' 59"	17° 46' 18"	50	
C.Usumacinta	91° 13' 44"	17° 17' 23"	123	

La Tabla 2 muestra un ejemplo típico de la estructura o campos de la base de datos de la estación meteorológica de Paraíso en formato Excel, descargada desde la página web del servicio meteorológico nacional (<http://smn.cna.gob.mx>). Para obtener el archivo siga el siguiente procedimiento: *observando el tiempo* → *estaciones meteorológicas automáticas* → en el mapa seleccionar Tabasco → *periodo: 60 min* → *aceptar* → seleccione la estación e importe datos.



Tabla 2. Datos de paraíso correspondientes al mes de septiembre de 2012.

dd/mm/aaaa hh:mm	DVS (°)	DVr (°)	VVs (m/s)	VVr (m/s)	T (°C)	HR (%)	Pb (mb)	Prec (mm)	RS (W/m ²)
01/09/2012 00:00	316	59	0.52	11.7	28.8	68	1011.8	0	5.7
01/09/2012 01:00	47	57	2.5	7.9	28.1	72	1011.7	0	0
01/09/2012 02:00	61	91	2.65	10	27.9	74	1011.8	0	0
01/09/2012 03:00	79	84	4.12	13.4	27.3	75	1012.8	0	0
01/09/2012 04:00	99	79	3.15	13.8	27.2	77	1013.9	0	0
01/09/2012 05:00	145	197	3.67	12.7	27.1	78	1014.7	0	0

Para procesar esta base de datos en WRPLOT es necesario agregarle cuatro campos o columnas: año, mes, día y hora. Esto se logra en EXCEL aplicando las funciones que llevan esos mismos nombres a la columna que lleva la fecha de la medición.

Para importar una base de datos en Excel desde el programa WRPlot, vaya al barra de menus y seleccione el menú *Herramientas (Tools)* → *Importar desde Excel (Import from Excel)*, se desplegará la ventana de la Figura 6. Para cargar la base de datos presione en la carpeta amarilla y seleccione el archivo. Luego en el área de datos de campo indique en que columnas se encuentran el año, mes, día y hora, velocidad y dirección del viento. También complete la parte de información de la estación. Luego guarde el archivo en la dirección que desee o simplemente presione *importar (import)* para que el archivo se salve con extensión *.sam.

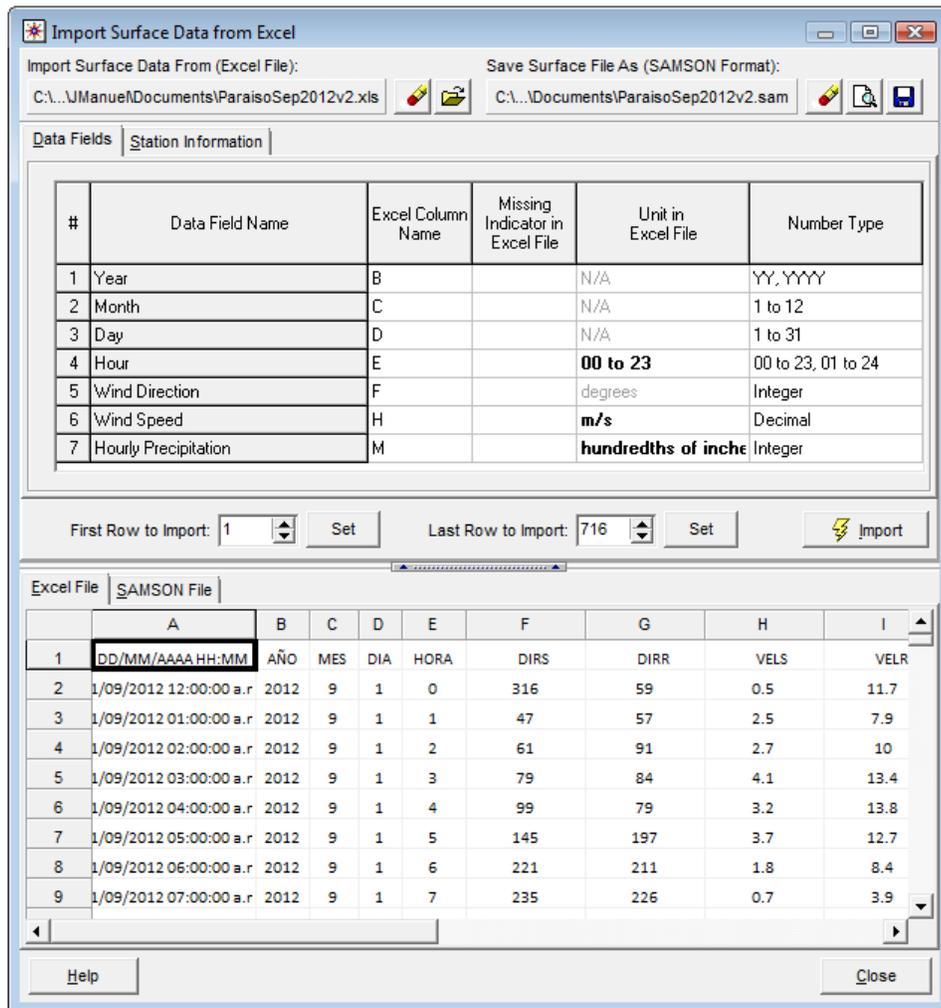


Figura 6. Ventana para importar datos desde Excel.

Para obtener la rosa de vientos cierre y regrese a la ventana principal y cargue el archivo presionando *Add File* y buscando el archivo con extensión *.sam (Figura 7).

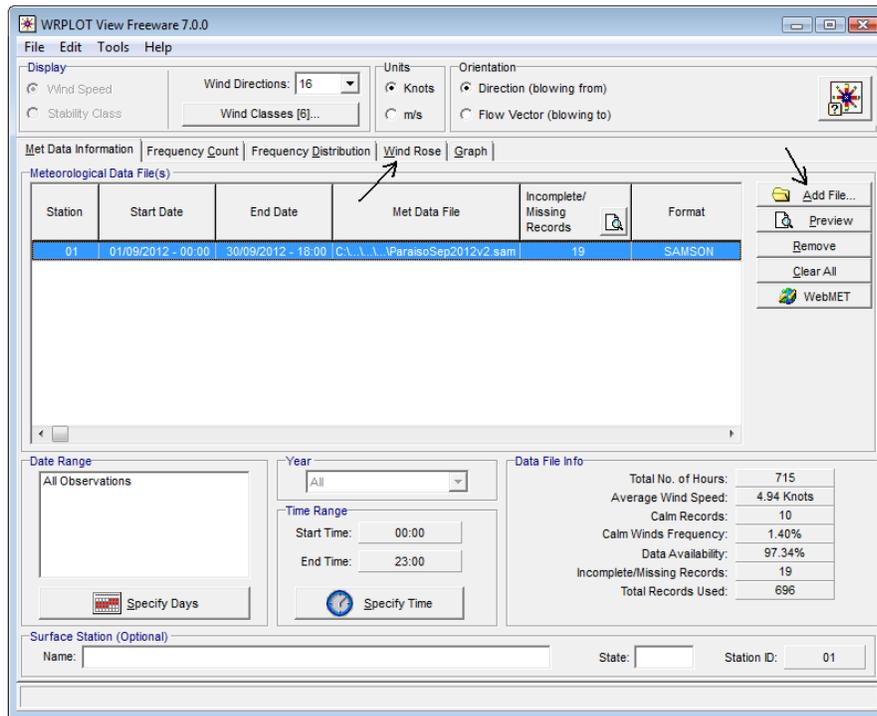


Figura 3. Para cargar internamente el archivo

Puede visualizar la rosa de vientos seleccionando la pestaña *Wind Rose* (Figura 7). Esta se mostrará en lo que se conoce el área de entrada (Figura 8).

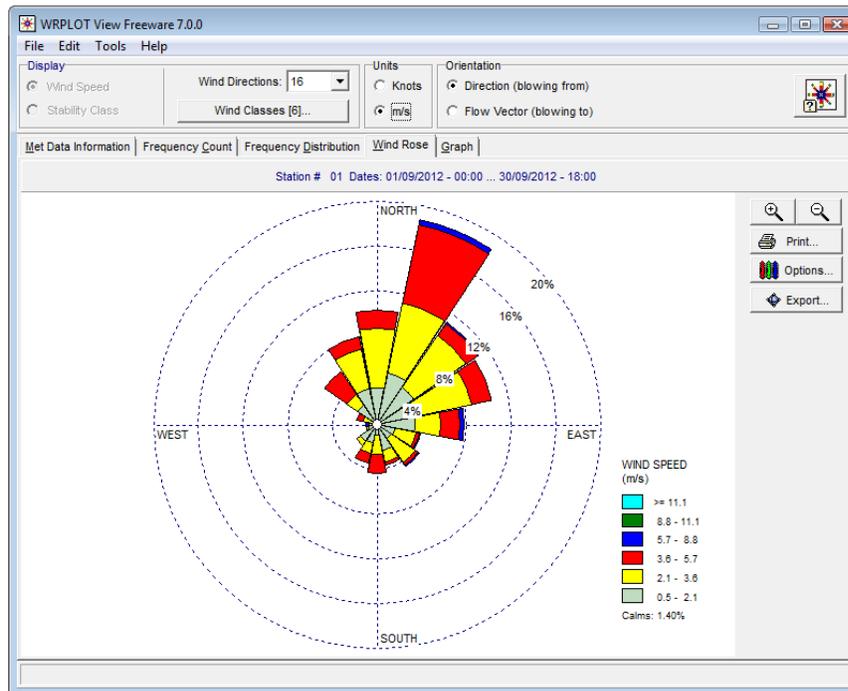


Figura 8. Rosa de vientos de paraíso en Septiembre de 2012

A continuación se muestran los resultados con una base de datos más extensa. La información corresponde a la estación meteorológica de paraíso del 5 de mayo de 2011 al 30 de septiembre de 2012. La base de datos *Paraiso2011_2012.xls* está disponible en la página web: <http://sites.google.com/site/dacbiolsergioramos>.

Después de exportar la base de datos y definir todos los parámetros correspondientes explicados antes, la rosa de vientos para todo el periodo considerado se presenta en la Figura 9. Se extrae de esta misma que el viento dominante fue del norte noreste (NNE) y que mayormente estos vientos estuvieron en el rango de 0.5 m/s a 5.7 m/s. Información cuantitativa se puede visualizar en la pestaña *Frequency Distribution* de la ventana principal del software.

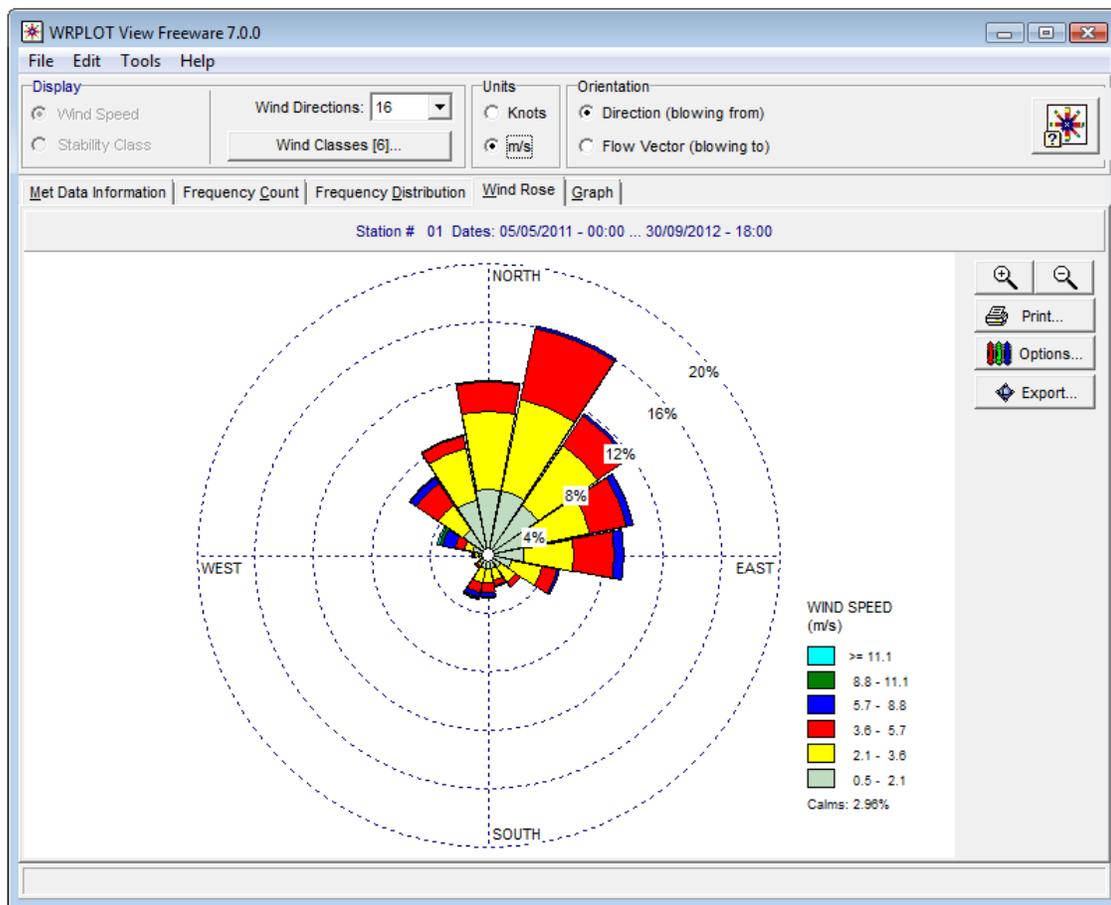


Figura 9. Rosa de vientos de paraíso 2011-2012

La Figura 10 muestra la frecuencia por clase de vientos o por rango de los vientos que soplaron en la estación de Paraíso en el periodo considerado. Esta se obtuvo con la pestaña *Graph*. Se puede concluir que los vientos más frecuentes estuvieron en el rango de 2.1 m/s a 3.6 m/s.

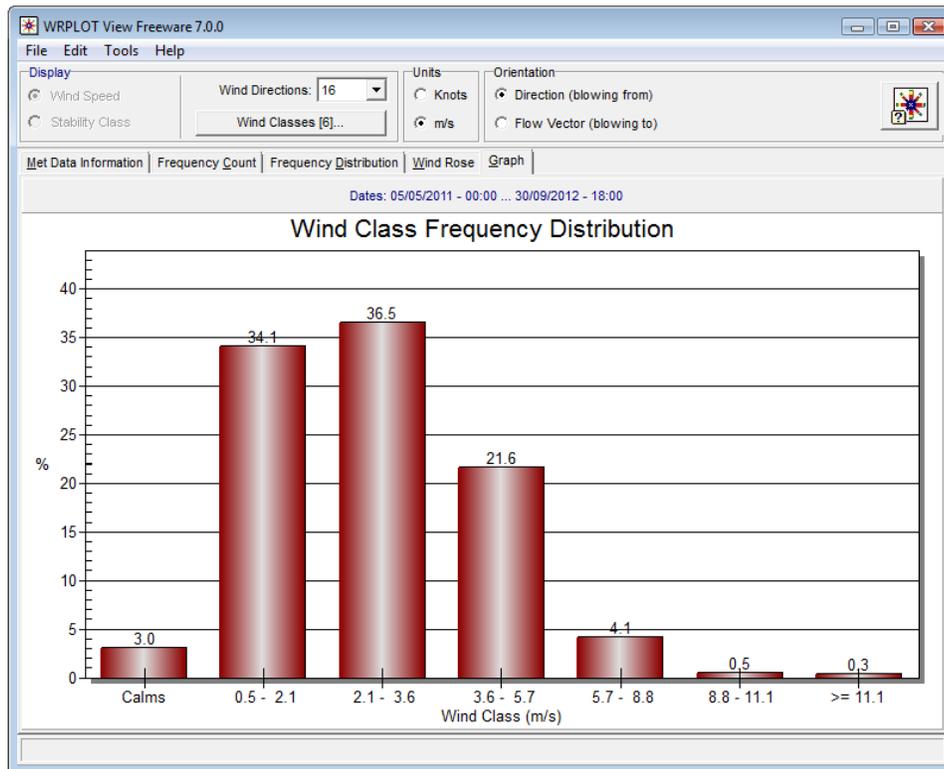


Figura 10. Frecuencia de los vientos para el periodo considerado

La Figura 11 muestra cómo se puede obtener la rosa de vientos anual, mensual, para ciertos días específicos o para ciertas horas durante el día. Primero debe seleccionar la pestaña *Met Data Information*, luego seleccionar el año en cuestión en el menú de cascada *year*. Se puede especificar un intervalo de tiempo durante el día con la opción *Specify Time* o seleccionar días específicos del año con la opción *Specify Days*.

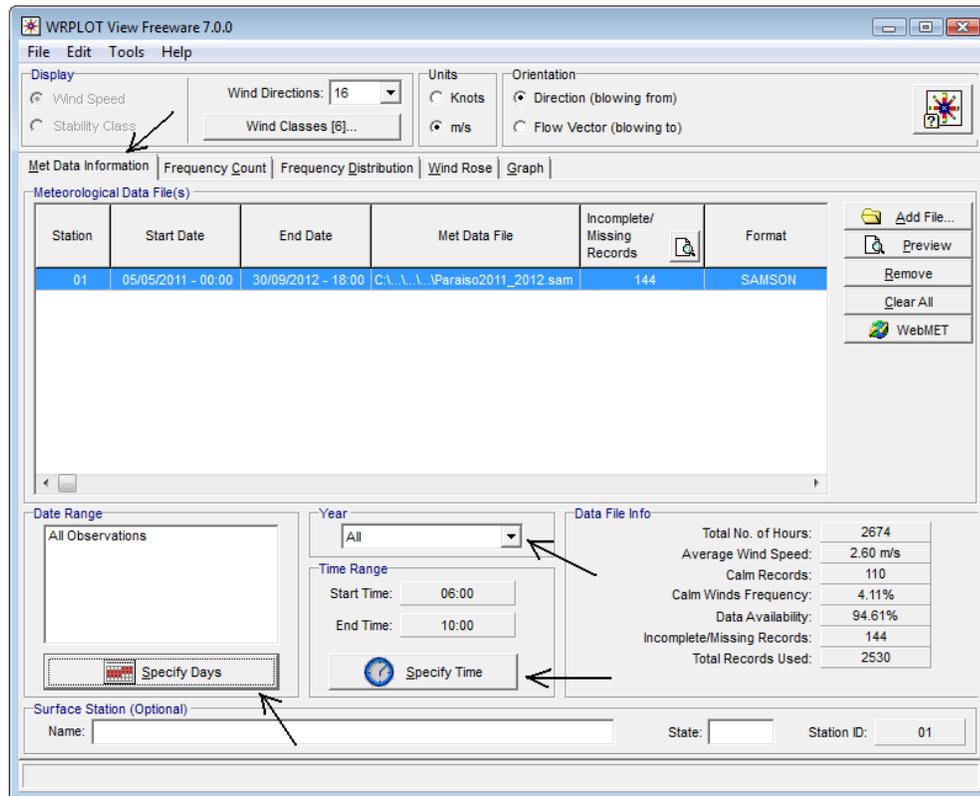


Figura 11. Opciones para seleccionar fechas específicas y horas al día para la rosa de vientos.

2 Software Screen View

Screen View es un modelo de sondeo amigable de la dispersión de contaminantes en el aire a partir de diversas fuentes. Este software está basado en el código abierto del modelo SCREEN₃, desarrollado por la Agencia de Protección Ambiental de los Estados Unidos (EPA, por sus siglas en inglés). La Figura 2.1 presenta la ventana de acceso al software desarrollado por la empresa weblake. A continuación se muestra con un ejemplo sencillo el uso de este software.

El modelo SCREEN₃ puede ser usado para estimar los peores casos de concentraciones a nivel del suelo de una sola fuente contaminante, así como las concentraciones en zonas de cavidad y las concentraciones debido a la inversión de ruptura y la fumigación de la costa.

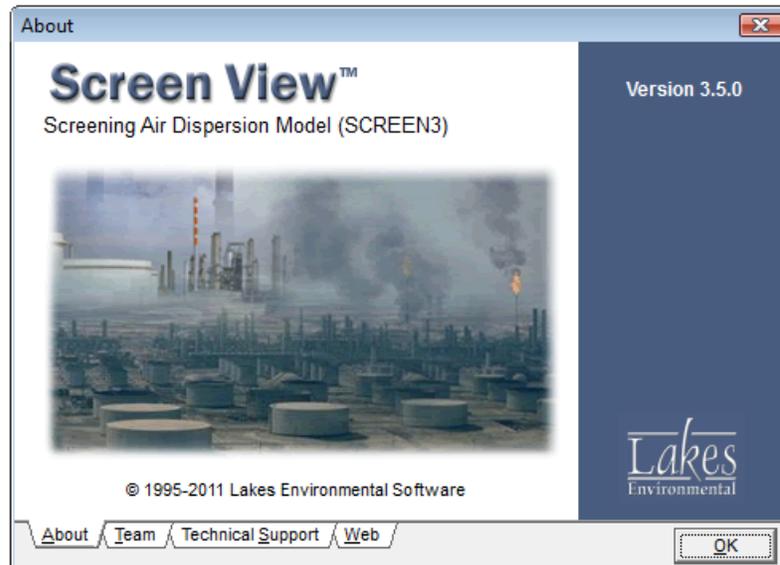


Figura 2.1 Ventana de presentación del software

Ejemplo 2.1. Un complejo termoeléctrica ubicado cerca de la costa cuenta con 6 unidades generadoras de vapor que consumen anualmente 3.7 Mm^3 de combustóleo. En su perímetro y hasta una distancia de 60 km se localizan varias zonas urbanas. En las primeras horas del día (de 8 de la mañana a las 6 de la mañana) el viento dominante es de tierra a mar y tiene una velocidad media de 4 m/s; y entre la 1 y las 6 de la tarde el viento dominante va de mar a tierra y su velocidad promedio es de 5 m/s. La emisión de SO_2 y de partículas preocupan más debido a su efecto en la salud humana. Calcule la concentración anual de SO_2 en la zona aledaña al complejo. Este complejo cuenta con tres chimeneas, cada una de 120 metros de altura y 5.5 metros de diámetro interior. La velocidad de salida de los gases de combustión oscila entre 22 y 23 m/s, con una temperatura de salida entre 425 y 428 K. La siguiente Tabla muestra las tasa de emisión en cada chimenea.

Unidad	Emisiones (ton/año)		
	SO_2	NO_x	PM_{10}
1	89,288	7,426	5,795
2	91,031	7,570	5,908
3	76,248	6,342	4,948

Solución con Screen View. A continuación se muestra el procedimiento para estimar las concentraciones debido a la chimenea número 1.



Se crea un nuevo proyecto en la pestaña New. Este se guarda con un nombre y la extensión *.scr por default en la carpeta Screen View (Figura 2.2). Inmediatamente se despliega la ventana de datos (Figura 2.3) dónde el usuario puede capturar la información pertinente.

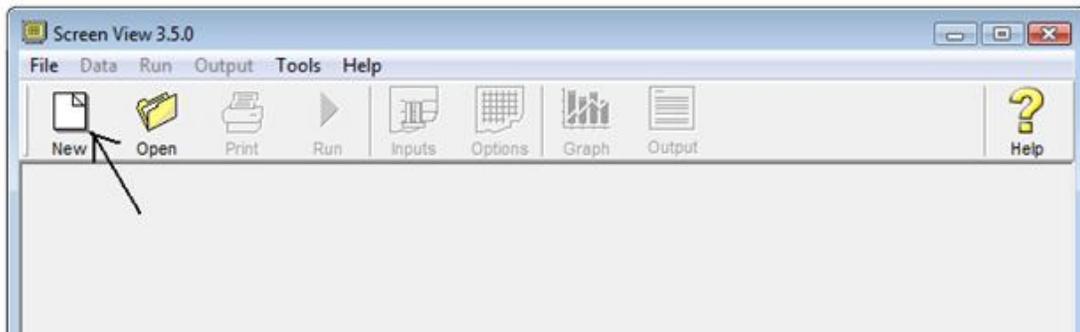


Figura 2.2

Las chimeneas se modelan como fuentes puntuales, que es la opción por default señalada. En esta ventana puede introducir la información básica de la fuente. En este caso la que corresponde a la unidad 1 y al SO₂. Se seleccionan coeficientes de dispersión para zona urbana y receptores a ras de suelo (Figura 2.3).

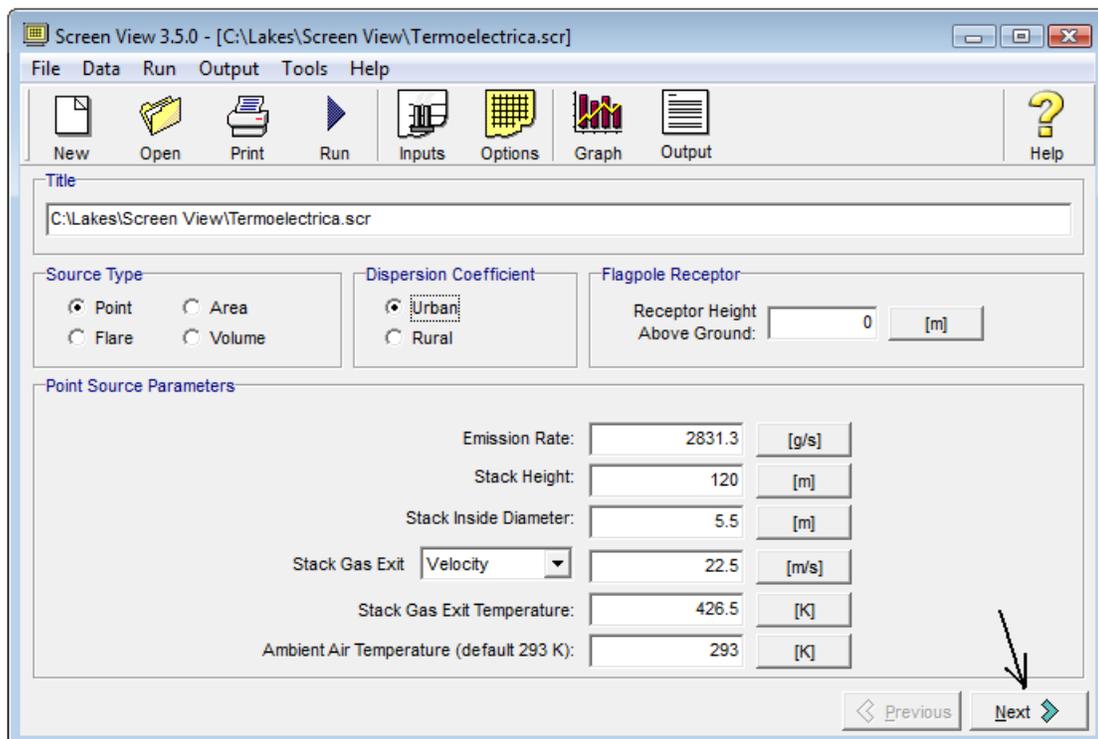


Figura 2.3 Ventana para introducir la información básica de la fuente

Al dar next la ventana cambia para permitir ingresar datos correspondientes al tipo de terreno y condición meteorológica (Figura 2.4). Si selecciona todas las condiciones meteorológicas, internamente se harán los cálculos que den la concentración del peor escenario; esta es la opción que se selecciono en este caso. Para introducir las distancias a las que desee calcular las concentraciones seleccione la pestaña *Automated Distances* (Figura 2.5).

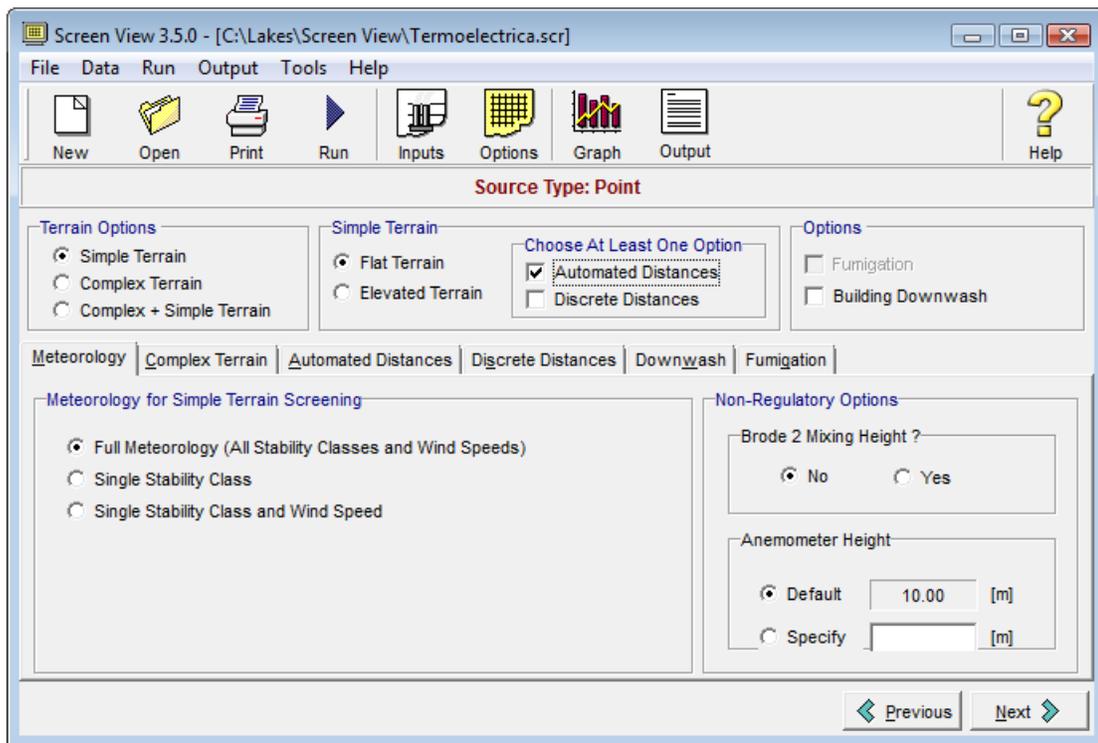


Figura 2.4 Datos correspondientes al tipo de terreno y condición meteorológica

Al dar next la ventana cambia a la de la Figura 2.6 dónde puede consultar detalles de su proyecto o correrlo con Run. Después de la corrida se le pregunta al usuario que tipo de salida desea. Hay dos opciones: gráfico o archivo de salida con los resultados de las concentraciones. La Figura 2.7 muestra el gráfico de concentraciones. Estas corresponden a concentraciones horarias y en el gráfico es posible distinguir la concentración máxima posible.

Si desea saber las condiciones meteorológicas que se asocian con esta concentración máxima o tener acceso a los niveles de concentración con la distancia entonces seleccione la opción archivo de salida. Este muestra un resumen de las condiciones bajo las cuales se hizo la corrida. Además este archivo puede guardarse y abrirse desde Excel. Esta última opción le permitirá modelar el comportamiento conjunto de las tres chimeneas, ya que desafortunadamente, el software no permite la entrada de múltiples fuentes puntuales. Para este caso se recomienda usar el principio de superposición. Es decir, calcule las concentraciones para las otras dos fuentes de manera separada y luego sume (con la ayuda de Excel) las concentraciones y obtenga un solo perfil para representar el efecto sobre los receptores viento abajo debido a las tres fuentes de dispersión.

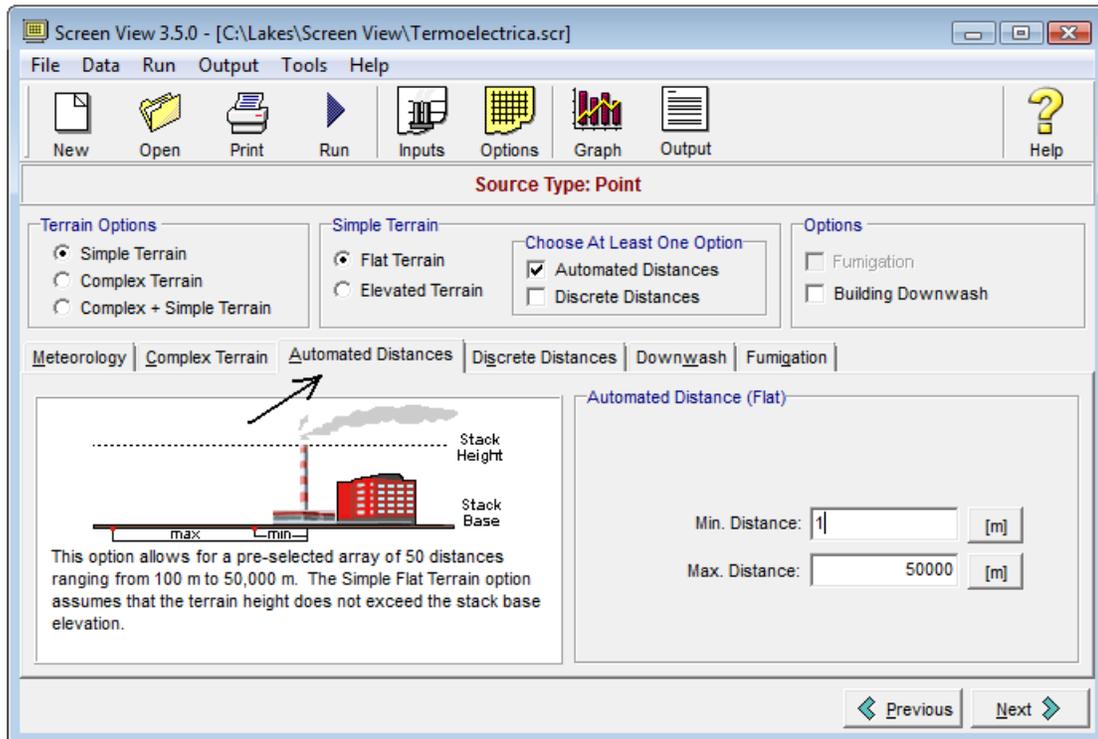


Figura 2.5 Ventana relativa a las distancias a las que se desea calcular las concentraciones

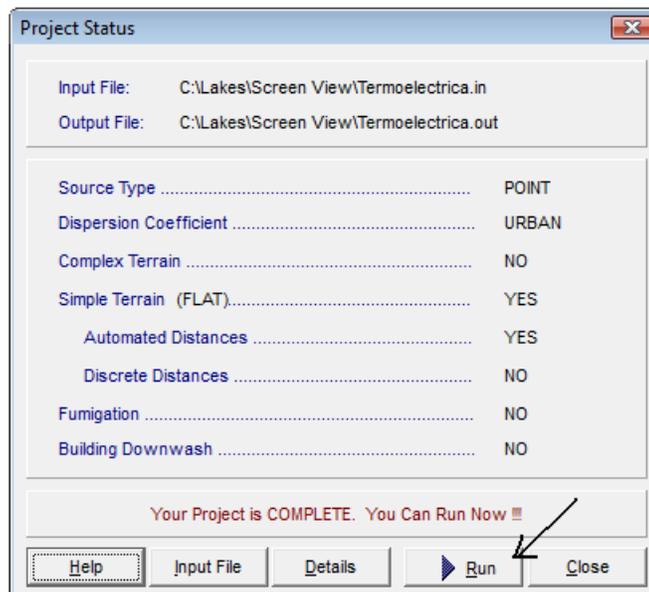


Figura 2.6. ventana para la corrida del proyecto de modelación.

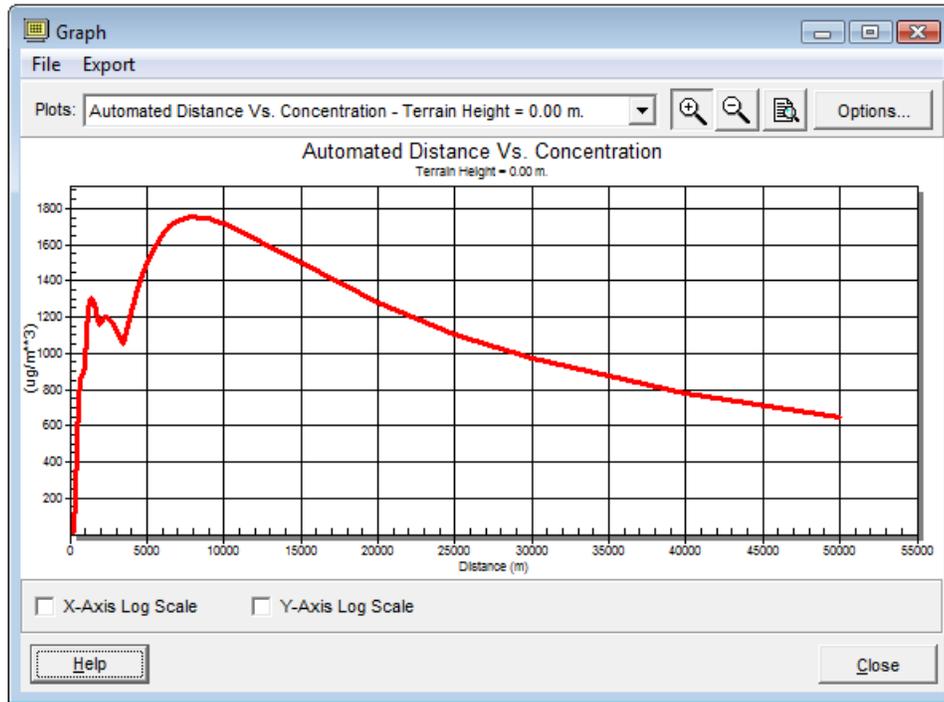


Figura 2.7 Salida gráfica del proyecto



3. R y el paquete OPENAIR

El paquete *openair* permite, entre otras cosas, calcular rosas de viento a partir de una base de datos meteorológicos horarios. Si ya está instalada en R, debe cargarse a la sesión con el comando *library* del siguiente modo:

Esquema básico de una instrucción en R:

instrucción (**objeto/conjunto de datos**, **comando**="variable")

3.1 Instrucciones previas al trabajo con archivos y datos:

require: se usa en cualquier sesión de trabajo en R para cargar la librería con la que se desea trabajar, por ej: `require(openair)`. También puede usarse `library(openair)`.

getwd o setwd: `getwd()` se usa para presentar en pantalla el directorio de trabajo de R. `setwd("C:/documentosR")`, establece un nuevo directorio de trabajo para guardar archivos en R.

list.files: presenta un listado de los archivos que se encuentren en el directorio de trabajo. Se puede especificar una lista en particular. Ej. `list.files(pattern=".txt")`.

Instrucciones básicas para el manejo de los conjuntos de datos

names: devuelve el encabezado de todas las variables de que dispone el archivo. Ej. `names(datos)`.

head: devuelve el encabezado del archivo con las primeras seis filas de datos de que dispone el archivo de datos. Ej. `head(datos)`.

tail: devuelve el encabezado del archivo con las últimas seis filas de datos de que dispone el archivo de datos. Ej. `tail(datos)`.

summary: proporciona el resumen de datos para el parámetro PM10 del archivo de datos datos, sin mostrar los datos de las demás variables. Ej. `summary(datos$PM10)`.

str: proporciona la estructura de los datos del archivo datos incluyendo la tipología de cada uno de los campos. Ej. `str(datos)`.

save, load, write: permite salvar los datos cargados previamente, volverlos a cargar después de guardados y generar un nuevo archivo de texto con el conjunto de datos cargados anteriormente, respectivamente.

Ej. `save(datos, file="datos.dat")`,
`load(datos.dat)`
`write("datosNuevos.txt", now.names=TRUE)`

rm, remove: elimina el objeto datos de nuestro proyecto. Ej. `rm(datos)` o `remove(datos)`.



3.2 Tipos de Variables:

Variables para comandos lógicos: **TRUE** o **FALSE** se usan para activar o desactivar comandos en R y Openair.

Variables numéricas simples: Son variables que reciben un valor numérico simple. Para valores con parte decimal se usa el punto, ya que la coma esta reservada para separar comandos. Ej. **angulo = 45**.

Variables de texto o variables string: Este tipo de variables como entradas caracteres o conjuntos de caracteres. En este caso el texto se escribe siempre entre comillas, “ ”. Ej. **sexo = “Femenino”**.

Variables vectoriales o multivariantes: Hay comandos que requieren como variables que guarden un vector o un conjunto de dos o más variables, numéricas o de texto. Para ello se usa la instrucción **c(i, ii, iii, ...)** siendo **c** el complemento a la instrucción. Ej. **valores = c(3 4 5 6)**.

Variables dentro de una secuencia determinada: **seq(x,y,z)**, sirve para determinar secuencias de datos donde x representa el mínimo de la secuencia, y el máximo y z el intervalo. Permite establecer escalas o rangos de datos personalizados.

Instrucciones básicas para el manejo de los datos

Funciones matemáticas de R

- **sum(datos\$r, na.rm = TRUE)**, devuelve la suma de los datos contenidos en la columna r de la matriz datos.
- **sqrt(variable/numero)**, devuelve la raíz cuadrada del dato entre parentesis.
- **log(variable)** o **log10(variable)**, devuelve el logaritmo de variable.
- **exp(variable)**, devuelve la función exponencial de variable.

Funciones estadísticas básicas de R

- **mean(datos\$pm10, na.rm=TRUE)**, devuelve la media total de los datos de particulas PM10.
- **max(datos\$pm10, na.rm=TRUE)**, devuelve el valor máximo del conjunto de datos. Se puede emplear también la función min para encontrar el valor mínimo.
- **quantile(datos\$pm10, na.rm=TRUE)**, devuelve los cuartiles propios del conjunto de datos 0%(mínimo), 25%, 50% (mediana), 75%, 100% (máximo)
- **quantile(datos\$pm10, c(.30, .45, .65, .95), na.rm=TRUE)**, devuelve cuartiles personalizados, elegidos por el usuario para el 30%, 45%, 65% y el 95%.
- **median(datos\$pm10, na.rm=TRUE)**, devuelve la mediana de las particulas pm10.
- **sd(datos\$pm10, na.rm=TRUE)**, devuelve la desviación estandar.
- **var(datos\$pm10, na.rm=TRUE)**, devuelve la varianza de los datos de PM10.



Comandos básicos para Gráficas en R

El comando **PLOT** genera una gráfica a partir de los datos indicados en la función. Algunos parámetros comunes a ésta y otras funciones gráficas son:

- **main**, permite incluir un título en la gráfica.
- **sub**, incluye un título en la parte inferior de la gráfica
- **xlab**, **ylob**, incluye títulos en los ejes.
- **xlim**, **ylim**, establece los límites en los ejes.
- **col**, permite cambiar los colores de la gráfica

Ej. `plot(datoscont$no2, xlab="horas", ylab="concentración", ylim=c(0,400), main="Evolución del NO2", col="blue", col.main="green", col.lab="red")`.

Rosas de Viento

A continuación se explicará la forma de obtener la rosa de vientos y gráficos de las series de tiempo de las variables a partir del archivo de datos meteorológicos horarios *ParaisoNov2015.dat* del 2 de enero al 22 de mayo de 2015. Este archivo contiene la información meteorológica de la estación de Paraíso, Tabasco en el mes de noviembre de 2014. La base de datos contiene varios campos, entre los que se destacan, la velocidad del viento (VV) en Km/h que se ubica en la cuarta columna y su dirección (wd) en grados desde el norte.

date	wd	DIRR	ws	VELR	TEMP	HR	PB	PREC	RAD
02/01/201518:00	124	115	17.77	34.9	28.8	65	1013.2	0	763.8
02/01/201519:00	127	125	16.03	33.8	29.4	64	1011.9	0	640.8
02/01/201520:00	129	131	13.09	23.4	29.9	62	1011.1	0	509.3
02/01/201521:00	160	153	7.58	16.2	30.3	62	1011.0	412	
02/01/201522:00	115	101	8.5	21.6	29.6	62	1010.9	0	115.2
02/01/201523:00	107	105	8.54	20.9	28.1	68	1011.1	0	22.7
03/01/201500:00	97	104	7.25	13.7	26.9	73	1011.3	0	-1
03/01/201501:00	99	102	7.56	13.7	26.3	76	1011.9	0	-1
03/01/201502:00	96	92	7.52	18.7	25.7	79	1012.0	-1	
03/01/201503:00	97	87	4.77	12.6	24.9	82	1012.2	0	-1
03/01/201504:00	97	105	7.43	17.6	24.6	83	1012.2	0	-1
03/01/201505:00	97	122	6.55	13.3	24	86	1012.2	0	-1
03/01/201506:00	95	101	8.87	17.3	23.4	91	1011.8	0	-1
03/01/201507:00	124	121	13.67	24.8	23.8	90	1011.4	0	-1
03/01/201508:00	123	132	14.93	26.3	23.6	91	1011.3	0	-1
03/01/201509:00	122	128	12.91	24.1	23.2	92	1011.1	0	-1
03/01/201510:00	155	162	10.36	19.4	22.7	94	1011.6	0	-1
03/01/201511:00	159	147	8.26	16.6	22	98	1012.0	0	
03/01/201512:00	163	182	7.22	13.3	22.2	99	1012.6	0	0.7
03/01/201513:00	153	159	8.46	15.1	22.3	98	1013.4	0	44.7
03/01/201514:00	180	203	7.57	12.6	23	96	1014.2	0	195.7
03/01/201515:00	161	167	7.22	13.7	24.5	88	1014.6	0	453.8
03/01/201516:00	149	129	5.34	12.2	26.7	79	1014.5	0	674.2
03/01/201517:00	259	306	3.85	16.2	28.2	70	1013.7	0	755.2



Se carga la base de datos en R.

```
> datos=read.table("Paraiso2015.dat", header=TRUE)
> head(Datos)
```

El campo de vientos se pasa de km/h a m/s

```
> datos$ws=datos$ws*(10/36)
> head(datos)
```

Si el campo de velocidades tiene un identificador diferente a ws, por ejemplo VV, se puede renombrar como ws con la función *names* cómo se indica en las líneas de abajo. Datos[4] señala la columna donde se ubica el campo VV, por ejemplo.

```
> names(datos)[4]="ws"
```

Una vez importados los datos de contaminación atmosférica que queremos, es preciso indicar a R y esencialmente a Openair, cual es el campo que contiene los datos de fecha y qué formato tiene, pues será este campo el que utilice como index para la posterior representación gráfica y estudio estadístico de los datos en la mayoría de las funciones a ejecutar. Esto se lleva a cabo mediante la función *as.POSIXct*, que convierte la fecha al formato adecuado para su manejo en R, y la función *strptime*, que le dice a R en qué formato está la fecha que se va a utilizar. Un ejemplo de su uso sería el siguiente: ☐

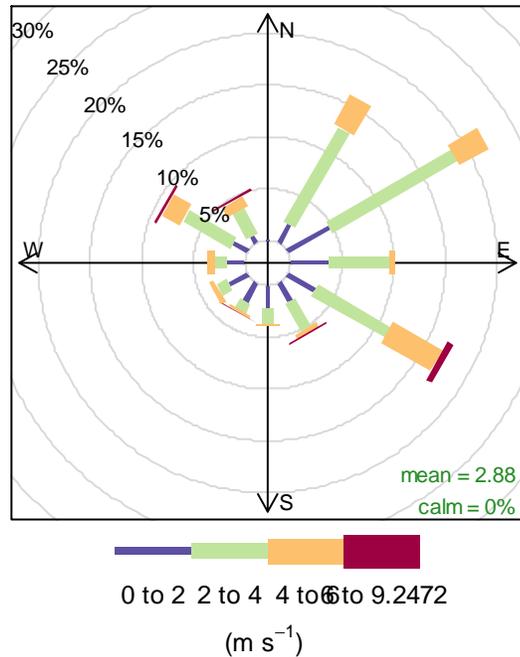
```
Ej. datos$date = as.POSIXct(strptime(datoscont$date, format = "%d/%m/%Y %H:%M", "UTC")),
```

donde le indicamos a Openair que la fecha de los datos *datos\$date*, se corresponderán con la posición fecha, y con el formato día/mes/año y hora:minuto "%d/%m/%Y %H:%M", y se encuentran en GMT, aunque también le podemos decir que están en UTC, si así fuese.

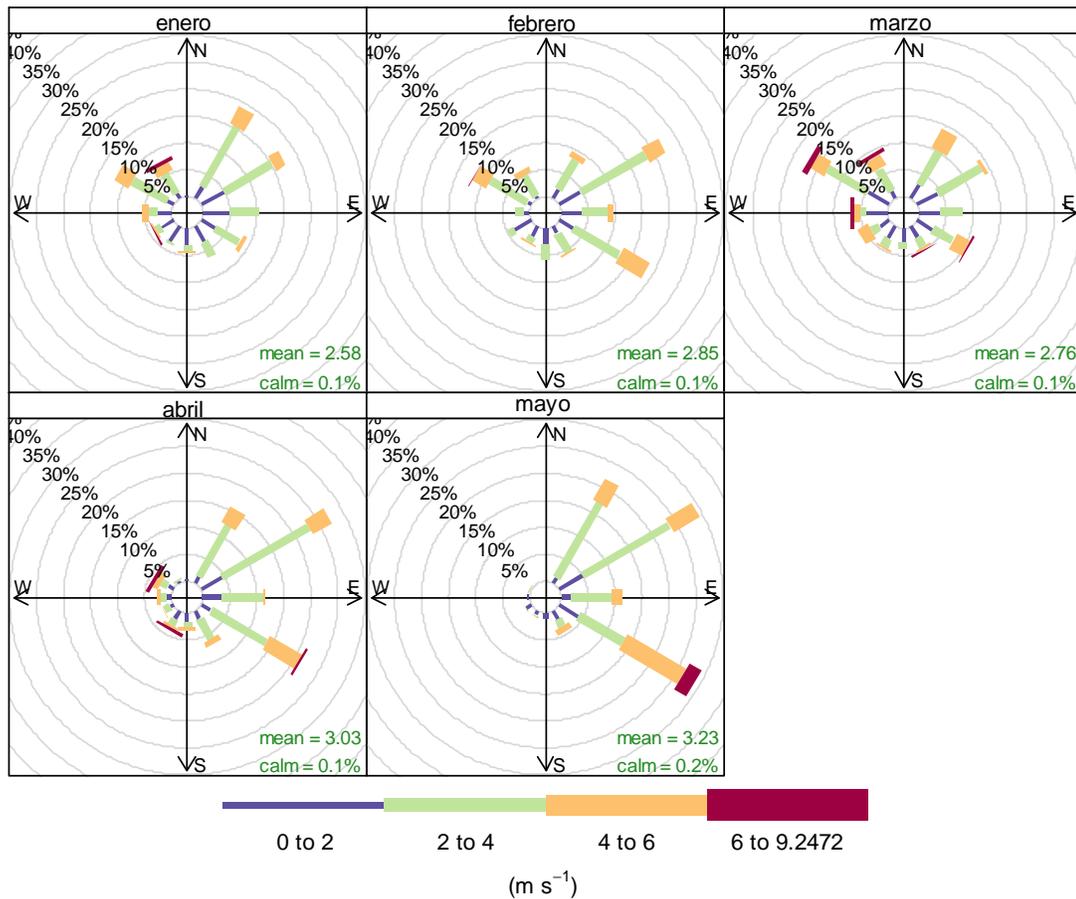
Concepto	Código	Tipo	Formato	Rango
DIAS				
... del mes	%d	número	nn	01 – 31
... de la semana	%a	texto	Abreviado	Sab – Vie
	%A	texto	Completo	Sábado – Viernes
	%w	número	nn	0 – 6 (0 = domingo)
... del año	%j	número	nnn	001 -366
SEMANAS				
	%U	número	nn	0 – 53 (usa domingo como primer día)
	%W	número	nn	0 – 53 (usa lunes como primer día)
MESES				
	%b	texto	Abreviado	Ene – dic
	%B	texto	Completo	Enero – Diciembre
	%m	número	nn	01 – 12
AÑOS				
	%y	número	nn	00 – 99
	%Y	número	nnnn	1900 – 2999
HORAS				
	%l	número	nn	01 – 12
	%H	número	nn	00 – 23
	%p	texto	Completo	AM/PM (a usar sólo con %l)
MINUTOS				
	%M	número	nn	00 – 61
SEGUNDOS				
	%S	número	nn	00 – 61

Por lo tanto, se le da formato a la fecha para pedir los gráficos de series de tiempo
`> datos$date=as.POSIXct(strptime(datos$date, format="%d/%m/%y%H:%M", "UTC"))`
 Se construye la Rosa de Vientos con el comando `windRose`. También se muestra la rosa de vientos para los diferentes días de la semana.

`> windRose(datos)`



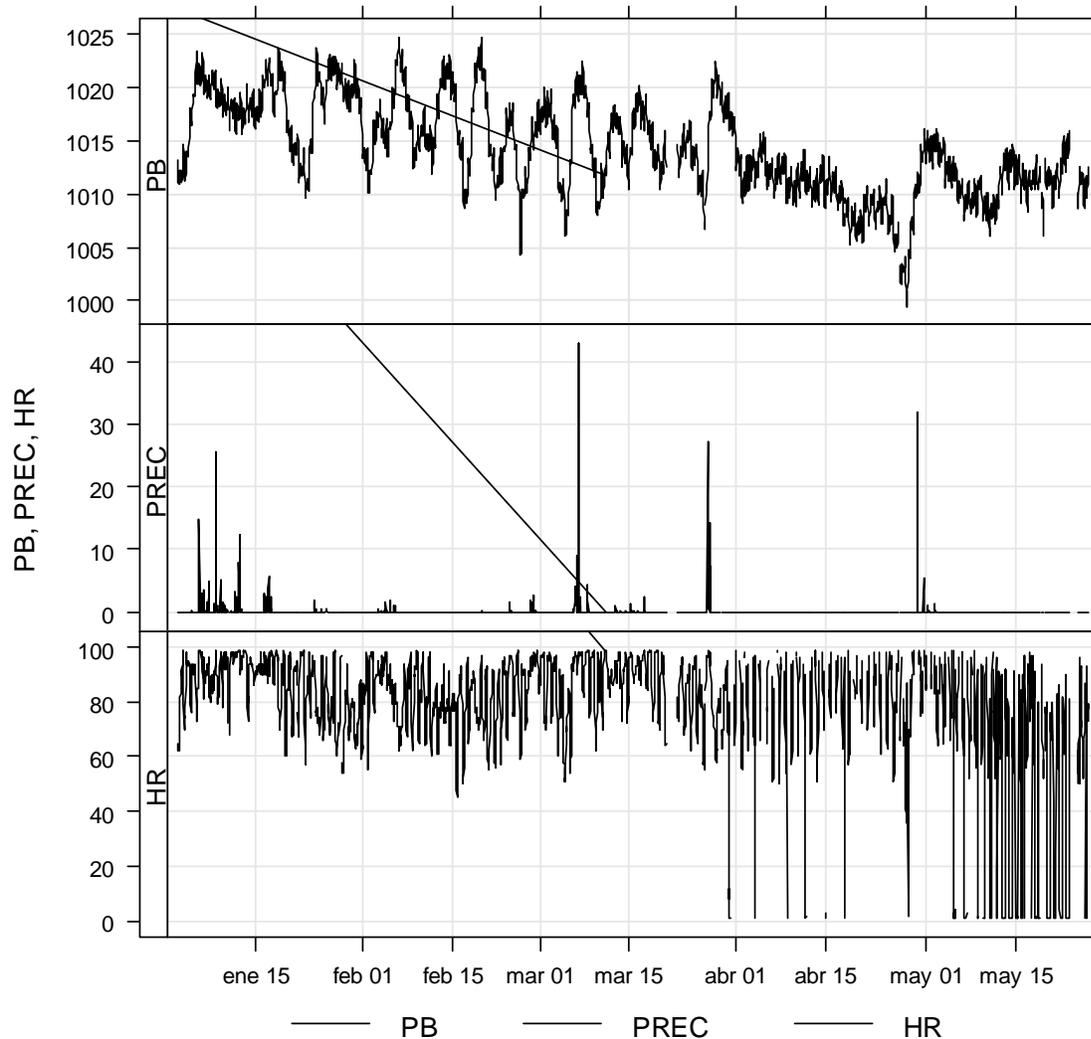
`> windRose(datos, type="weekday")`



Frequency of counts by wind direction (%)

Las siguientes instrucciones obtienen las graficas temporales la velocidad del viento, la temperatura y la humedad relativa de los valores horarios originales, de los promedios diarios, del máximo diario y del mínimo diario.

```
> timePlot(datos, pollutant=c("PB", "PREC", "HR"))
```



También se pueden obtener las gráficas por promedio diario, o más específicamente por máximos diarios, mínimos diarios, etcétera.

```
> timePlot(datos, pollutant=c("ws", "TEMP", "HR"), avg.time = "day")
> timePlot(datos, pollutant=c("ws", "TEMP", "HR"), avg.time = "day", statistic="max")
> timePlot(datos, pollutant=c("ws", "TEMP", "HR"), avg.time = "day", statistic="min")
```

Las opciones de avg.time son: sec, min, hour, day, week, quarter y year.

La siguiente instrucción presenta los datos suavizados de la temperatura.

```
> timePlot(datos, pollutant="HR", avg.time = "hour", smooth=TRUE)
```



Para obtener las tendencias de los valores horarios

```
> timeVariation(datos, pollutant="TEMP")
```

La instrucción *cutData* permite agregar a los datos un nuevo campo clasificador. Por ejemplo, el mes o el día de la semana. Esto permite crear gráficos de cajas por día de la semana por ejemplo como en la siguiente instrucción.

```
> datos = cutData(datos, type="weekday")
> head(datos)
> attach(datos)
> boxplot(TEMP~weekday)
```

El comando *selectRunning* crea un subconjunto de datos donde una de las variables se mueve dentro de cierto rango o está por arriba de valor límite (*threshold*)

```
> HRn = selectRunning(datos, pollutant="HR", threshold=80)
> head(HRn)
```

La Función de Representación Geográfica: GoogleMapsPlot

Utiliza los mapas de Google para realizar representaciones espaciales de la calidad del aire. La función *GoogleMapsPlot* se basa en mapas estáticos simples de Google (API de Google Maps). La función *GoogleMapsPlot* resulta muy útil para representar datos sobre distintos tipos de planos, permitiendo comprobar la distribución geográfica de las concentraciones de un determinado contaminante.

La función *GoogleMapsPlot* representa, como si fuese un gráfico X-Y habitual, las coordenadas de latitud y longitud geográfica que le demos en nuestro conjunto de datos,

Una vez representados los pares de coordenadas en la gráfica, lo que hace es añadir de fondo el mapa de Google Maps que se adapte a las coordenadas que tengamos en nuestros ejes X e Y. De esta forma, sólo con un conjunto de datos que disponga de la ubicación de las estaciones y los comandos básicos de la función, se representaría la ubicación de los puntos de control.

```
Ej. GoogleMapsPlot(estaciones, latitude="lat", longitude="long")
```

donde lo que hace la función es representar los distintos pares de coordenadas de las estaciones de que disponga el archivo de datos que hemos cargado con el nombre de "estaciones".

Si a esta configuración básica de representación le añadimos a cada punto de ubicación la posibilidad de que represente un parámetro, solicitando a la función que lo obtenga desde nuestro conjunto de datos, como por ejemplo el PM10, nos encontramos que para cada punto de la gráfica (par de datos de longitud y latitud), Openair modifica la representación del punto de



ubicación, variándolo de color y tamaño para representar la concentración dada del parámetro solicitado en el punto de coordenadas establecido, tal y como podemos ver en el siguiente ejemplo:

Ej. `GoogleMapsPlot(datoscontub, latitude="lat", longitude="long", pollutant="pm10")`

donde lo que hace la función es representar los distintos pares de coordenadas de las estaciones de de que disponga el archivo de datos que hemos cargado, para que disponga de las coordenadas de ubicación, modificando los puntos en tamaño y color para incluir la representación de la concentración de PM10 dada para cada punto.

Un aspecto importante a tener en cuenta, tal y como puede comprobarse en el ejemplo adjunto, es que GoogleMapsPlot no dispone actualmente de ninguna capacidad de cálculo estadístico o integración de los datos por lo que representa para cada punto todos los datos que encuentre del parámetro.

Esto supone que no basta con añadir las coordenadas geográficas a nuestro conjunto de datos sino que, tal y como podemos imaginar, debemos tratar previamente los mismos, antes de ejecutar la función de representación geográfica, de forma que la representación se haga para los datos pretendidos tanto desde el punto de vista de integración temporal como desde el punto de vista de su desagregación, si queremos aplicar el comando `type`.



Variaciones sobre la función:

Introducir las coordenadas geográficas a representar: latitude, longitude

La función requiere obligatoriamente que se le indiquen los campos de conjunto de datos que disponen de las coordenadas geográficas que posicionan cada uno de los puntos a representar para lo cual utiliza los comandos `latitude` y `longitude`, a los que habrá que indicarles los campos que disponen de los datos de latitud y longitud, respectivamente. La coordenada se puede proporcionar a Openair en formato numérico decimal, no siendo necesaria su transformación a grados, minutos y segundos.

Añadir un parámetro o contaminante al punto: pollutant

Es posible suministrar un campo adicional relativo a un parámetro o contaminante que se representará para cada punto de ubicación dispuesto. La representación de dicho parámetro se llevará a cabo modificando el tamaño y/o color del punto en cuestión, en función de los comandos que establezca el usuario para la función.

Modificar los márgenes de la gráfica: xlim, ylim

La función `GoogleMapsPlot` adopta automáticamente los valores de los Ejes X e Y que mejor se adapten a la representación de los puntos que se hayan suministrado, ajustando los límites superior e inferior de ambos ejes a un cuadrado que contenga todos los puntos con la máxima resolución posible del plano de fondo.

En cualquier caso, se podrá modificar dichos valores con los comandos `xlim` o `ylim`, a los que tendrá que dotarse con los valores mínimo y máximo que se desea para las coordenadas de longitud (para el Eje X) o de latitud (para el Eje Y). Se debe tener en cuenta que, en la actual versión de Openair, el área de representación de la gráfica es siempre cuadrada, por lo que el usuario sólo podrá modificar el Eje X o el Eje Y, adaptándose automáticamente el otro eje a los valores que correspondan.

Desagregar los datos para representar distintas gráficas: type

La función `GoogleMapsPlot` dispone del comando `type` que le permite desagregar los datos para representarlos en distintas gráficas. Si bien esta función conserva todas las características del comando `type`, vistas en gráficas anteriores, se debe recordar al usuario que el conjunto de datos a utilizar debe encontrarse previamente tratado, lo cual incluye también las posibles desagregaciones que se vayan a hacer con dicho comando. De lo contrario, la gráfica realizará para el mismo punto y la misma gráfica varias representaciones (tantas como valores adopte el parámetro para dicha ubicación), tal y como ocurría en la gráfica de la página anterior.

Es importante recordar siempre que la función `GoogleMapsPlot` basa su representación en los pares de coordenadas y no en el campo `date`, como otras gráficas. Se trata de representar coordenadas con la posibilidad de incluir una tercera variable, y no de representar sólo parámetros contaminantes.



Etiquetar los puntos geográficos representados: labels

Si el conjunto de datos, además de las coordenadas geográficas, aporta un campo de texto descriptivo de los puntos, es posible indicar a la gráfica, mediante el comando `labels`, el campo a utilizar para identificar y etiquetar cada uno de los puntos.

Personalizar los colores de la gráfica: cols

Al igual que en otras gráficas de Openair, el comando `cols` permite personalizar los colores de representación del parámetro o contaminante, ya sea utilizando cualquiera de los paquetes de colores predeterminados de Openair, o los colores que el usuario considere más adecuados.

En caso de utilizar el paquete predeterminado “greyscale”, no sólo se afectará a la escala de colores de representación del parámetro o contaminante, sino que se transformará toda la gráfica a escala de grises, con inclusión del plano de fondo de Google.

Tamaño con que se representa el punto en la gráfica: cex, cex.range

Si representamos en la gráfica tan sólo ubicaciones, el valor por defecto del tamaño del punto es `NULL` y se establece en función del rango de representación geográfica que se obtenga en la gráfica. En todo caso, se puede formar a los puntos a un tamaño distinto, utilizando para ello el comando `cex`, al que se le podrá asignar el tamaño que el usuario estime más adecuado, mediante un número entero.

Si el punto a representar en la gráfica busca a su vez representar el valor o concentración de un parámetro del conjunto de datos, la función `GoogleMapsPlot` varía por defecto su tamaño y color en función de la concentración o valor que adquiera el parámetro en cada caso. En este caso, el valor de `cex` es igual al valor de la concentración, lo cual puede afectar a la calidad de la representación, especialmente en aquellos casos en los que la ubicación de los puntos es muy próxima. En estos casos, el usuario puede optar por anular la representación del valor por tamaños, estableciendo un valor único con el comando `cex`, o escalar el valor de representación del punto mediante el comando `cex.range`, al que se le asignará el valor de escala que se quiera en función de las concentraciones a representar, como por ejemplo: `cex.range=c(1,3)`.

Eliminar la representación de los ejes en la gráfica: axes

Mediante el comando lógico `axes` es posible forzar a la gráfica a que represente los Ejes X e Y, opción por defecto, o hacer que los elimine de la gráfica, mediante el comando `axes=FALSE`.

Modificar el tipo de mapa que se muestra en la gráfica: matype

Por defecto, la función `GoogleMapsPlot` utilizará el plano del terreno de que dispone Google como mapa de base en función de las coordenadas que se disponga. Sin embargo, la opción de terreno “terrain” no es la única disponible, y el usuario podría precisar del uso de otros planos como: los mapas de satélite o fotografía aérea “satellite”, muy útiles cuando la representación se va a hacer para redes de calidad del aire de núcleos urbanos, los mapas de carreteras o vías de tráfico “roadmap” o “mobile”, o un plano híbrido entre el de terreno y el de carreteras “hybrid”.



Establecer el centro del plano: center

De forma automática, y en función de las coordenadas por puntos que le hayamos dado, la función `GoogleMapsPlot` establece el centro de la imagen del plano para ubicarla dentro del cuadrado que sirve para representar los puntos de coordenadas de la gráfica. Sin embargo, el usuario podría requerir de dicha imagen esté desplazada del centro del cuadrado, por lo que podrá personalizar las coordenadas que sirven para centrar la imagen con el comando `center`, al que se deberá asignar tanto la latitud como la longitud, en este orden, tal y como marca el siguiente ejemplo: `center=c(40.2,-4.5)`

Modificar el nivel de transparencia de la representación: plot.transparent

El comando `plot.transparent` nos permite regular el nivel de transparencia con que se representa el punto geográfico en la gráfica, pudiendo adoptar valores que vayan de `plot.transparent=0`, para hacer el punto invisible, a `plot.transparent=1`, para hacer la representación totalmente opaca.

Además de las opciones de configuración vistas hasta el momento, la función gráfica `GoogleMapsPlot` permite el uso de las opciones de configuración clásica de R para gráficas `plot`, tales como `xlab`, `ylab`, `main`, `sub`, etc. e incluso opciones de configuración vistas en otras gráficas como `strip` o `layout`, para influir sobre el modo de representación de la gráfica cuando se procede a su desagregación con `type`, tal y como podemos ver en el ejemplo final.

Referencias Bibliográficas

Folios-Pliego, F. (2012). Manual básico: Lenguaje R aplicado al Análisis de Datos de Calidad del Aire