1 Motivación y antecedentes del programa.

Para hacer una buena predicción, es necesario conocer la variabilidad de la señal a predecir y los factores que le influyen. Además, la variabilidad del viento (y de la generación eólica en general) se puede entender como el error que se comete cuando se aplica el modelo de persistencia. La persistencia es el modelo más sencillo de predicción que existe y consiste en considerar que el viento no cambia a lo largo del tiempo.

La metodología que propongo utilizar con este programa es ir experimentando a través del análisis interactivo de series de datos. De esta forma se desvía el énfasis del formalismo matemático (definiciones de transformadas, modelos estocásticos, etc.) hacia el comportamiento del sistema (visualización de gráficas, caracterización cualitativa del comportamiento del viento, análisis de los factores físicos que influyen, etc.).

De hecho, he intentado que este programa compagine el análisis en el dominio de la frecuencia (un sistema estocástico normal se encuentra definido por su densidad espectral de varianza y su representación canónica es su transformada de Fourier) y del tiempo (que nos proporciona una visión más cercana y que es imprescindible para el control de la máquina y del sistema eléctrico, generar predicciones, etc.).

Mi intención es centrarme en la visualización de gráficas interactivas, donde los alumnos pueden introducir datos de las estaciones meteorológicas y luego el programa se encarga de generar las gráficas.

También se puede ver la comparación de las variaciones de dos series de datos meteorológicos en el dominio de la frecuencia (coherencia de las dos señales).

1.1 *Caracterización de la variabilidad en el dominio frecuencial.*

En un proceso estocástico, el valor observado en un instante depende del observado en instantes cercanos. Por ejemplo, el viento que hará dentro de una hora depende del que hace ahora. Las series temporales se pueden utilizar modelos autoregresivos integrales de media móvil (ARIMA) u otros análisis temporales cuando se está interesado en la evolución futura de la variable a partir de los valores pasados.

No obstante, existen situaciones en las que un análisis frecuencial es ventajoso respecto al análisis clásico de series en el dominio del tiempo. Por ejemplo, permite extraer mayor información sobre el comportamiento cíclico de las señales. Además está relacionado con las excepcionales propiedades teóricas de los procesos estocásticos normales. Por ejemplo, cada punto del espectro (o línea espectral) de un proceso normal es independiente de los demás. Es decir, las propiedades de una señal aleatoria normal se caracterizan únicamente por la varianza que presenta en cada frecuencia. Dicha caracterización se denomina representación canónica de la señal, pues es la forma más compacta de representar sus características. En otras palabras, cada línea del espectro es una variable normal independiente estadísticamente del resto. Por tanto, una señal aleatoria normal se puede obtener aplicando la transformada inversa de Fourier a un espectro aleatorio con la varianza correspondiente a la señal a esa frecuencia.

Dado que la evolución del tiempo viene condicionado por la suma de muchos factores, algunas variables meteorológicas pueden considerarse que siguen un comportamiento normal aproximadamente. Por ejemplo, la variabilidad del viento suele caracterizarse a través de la densidad espectral de su varianza ("Power Spectrum" en

nomenclatura inglesa, aunque yo me resisto a hacer una traducción más literal para equívocos con la potencia eléctrica). Los espectros del viento de Kaimal, de Van der Hoven, de Karman, de Davenport, del Consorcio Noruego del Petróleo, etc. son densidades espectrales de la varianza del viento caracterizada a partir del viento medido por esos autores. La densidad espectral de la varianza es el módulo medio de la transformada de Fourier al cuadrado.

Una ventaja del análisis frecuencial es que permite estudiar más fácilmente el efecto de considerar variables meteorológicas en emplazamientos cercanos. La varianza de la suma de potencias de varios aerogeneradores o del viento equivalente de una región se puede estimar a partir de la coherencia espacial del viento. Por tanto, la varianza espectral de esas señales corresponde a la varianza del viento, amortiguada a altas frecuencias para tener en cuenta el efecto del promediado espacial correspondiente a la cancelación parcial entre turbinas de fluctuaciones rápidas.

Este comportamiento simplifica el estudio de la variación de la generación eólica en función del tamaño de la región considerado. Es decir, el viento equivalente de una región se puede considerar una versión filtrada del viento medido en un punto. Aunque no se puede conocer el valor concreto del viento equivalente de una región en un instante a partir de las medidas en una sola torre meteorológica, si se puede estimar su variabilidad a partir de una sola torre, de las dimensiones de la región y de la coherencia del viento en la región. Es más, si solo se dispone de una estación meteorológica de medida, la estimación de máxima probabilidad del viento equivalente regional es el viento medido en la torre, suavizado a través de un filtro estimado a partir de las dimensiones y de la coherencia.

Cuando consideramos la suma de potencias de varios aerogeneradores o el viento equivalente de una región, es necesario aplicar un filtro debido al promediado espacial del viento que se puede estimar en función de las dimensiones del área estudiada.

1.2 *Caracterización de la variabilidad en el dominio temporal.*

Es muy conveniente caracterizar la evolución atmosférica a lo largo el tiempo porque el control de la energía eólica y del sistema eléctrico está íntimamente relacionado con la predicción meteorológica. Además, las condiciones climáticas varían a lo largo del tiempo y las variables atmosféricas sólo pueden considerarse procesos estacionarios normales mientras las características ambientales no varíen significativamente.

Por tanto, la representación frecuencial de estos procesos es una aproximación a la realidad que puede utilizarse para obtener las características del proceso y, posteriormente, obtener una representación temporal que pueda ser utilizada finalmente para la predicción y el control de la energía eólica a partir de los valores medidos en el pasado.

La variabilidad en el dominio del tiempo se puede abordar con la teoría clásica de series temporales y modelos autoregresivos integrales de media móvil (ARIMA). No obstante, estos métodos están íntimamente ligados a las características espectrales de la señal. Como el programa *Leerficheros* no está centrado en la predicción sino en el análisis y caracterización, el ajuste de modelos ARIMA no está implementado actualmente.

Desde el punto de vista del control de redes eléctricas es muy importante conocer la probabilidad de cambios abruptos en la generación eólica (especialmente

para la programación de la reserva rodante y la seguridad de funcionamiento del sistema eléctrico).

Por una parte, la probabilidad de estos eventos extremos (aumentos/disminuciones y picos/valles abruptos en señales atmosféricas) puede deducirse del contenido frecuencia de la señal si el comportamiento es normal o se puede aproximar por un proceso normal transformado.

Por otra parte, las variaciones en series temporales son detectadas en el programa *Leerficheros* utilizando Wavelets de tipo Morlet. Utilizando esta familia y su complementaria con un dominio bastante corto se pueden detectar ráfagas de viento tipo rampa y pico a distintas escalas temporales y presentarlas en dos escalogramas. Para acelerar el cálculo de la transformada wavelet continua se ha utilizado un filtrado recursivo de la señal mediante medias móviles iterativas y la diferenciación entre puntos no contiguos.

La generación eólica puede desviarse del comportamiento normal en situaciones excepcionales: condiciones atmosféricas extremas, perturbaciones en la red, etc. Como herramientas mixtas en el tiempo-frecuencia, *Leerficheros* muestra el espectrograma (transformada corta de Fourier) de las variables atmosféricas. El espectrograma es la representación de la evolución del espectro de la misma forma que las barras de un ecualizador de sonido (que representa el contenido frecuencial de la música) varían a lo largo de los pasajes de una obra musical. El espectrograma puede servir para comprobar cuánto se aproxima la señal real a un proceso normal y puede servir también para identificar aquellos periodos durante los cuales la señal presenta variaciones atípicas que son características de condiciones meteorológicas extremas o comportamientos singulares del sistema.

Cuando el comportamiento de la señal no es normal, se puede utilizar la teoría general de las ecuaciones diferenciales estocásticas para modelizar el comportamiento aleatorio del sistema. Dado que la turbulencia muestra un cierto comportamiento exponencial, se puede modelizar el sistema con el método aproximado de las cadenas de Markov. Este método es análogo a la aproximación de una ecuación en derivadas parciales con una secuencia de diferencias finitas. Las variables que afectan a la evolución atmosférica se discretizan en un número finito de estados. Posteriormente se estima las probabilidades de transición a partir de la frecuencia observada, obteniéndose una matriz de transición de estado para el caso de una sola variable de estado y un tensor para el caso de dos variables o más. Una vez estimada las probabilidades de transición del sistema discretizado, se puede utilizar la interpolación para aproximar mejor la dinámica del sistema continuo.

El programa *Leerficheros* discretiza el viento (u otras variables del fichero de entrada) y calcula las matrices de transición entre los distintos niveles. En esta aproximación tan simple se considera que la evolución del sistema sólo depende del último valor medido, mientras que en la realidad habría que considerar otras variables de estado adicionales ya que la evolución atmosférica depende de muchos factores. No obstante, la aproximación con una cadena de Markov simple es muy ilustrativa e introduce al usuario a modelos de mayor orden. Además, la densidad probabilidad de transición para el caso simple se puede representar gráficamente y analizar de forma intuitiva sin necesidad de conocer la teoría markoviana[1].

[1] Los modelos de Markov permiten realizar un control óptimo de sistemas estocásticos utilizando procesos de decisión de Markov para obtener la programación dinámica óptima.

2 El Programa.

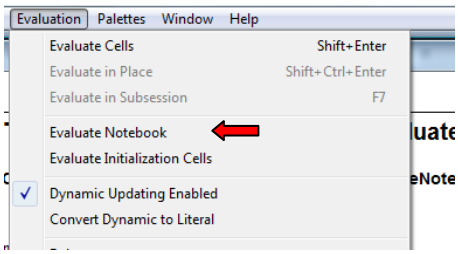
2.1 *Comienzo.*

El programa *Leerficheros* carga los datos de de los ficheros y compara las variaciones de dos series de datos. Por conveniencia, se ha implementado como un cuaderno en el programa de cálculo Mathematica™. De esta forma, se simplifica la realización de algunos cálculos y la interacción con el usuario.

Para poder utilizar el programa interactivo es necesario tener instalado en el ordenador la versión de pago del programa Mathematica™.

2.2 *Apertura de los ficheros de cálculos.*

Una vez ejecutado *Leerficheros* en versión de pago del programa Mathematica™, seleccionamos en el menú Evaluation, Evaluate Notebook.



A partir de este momento el programa pedirá los dos archivos a comprar y una serie de parámetros:

1.- *Lines in the header*: Hay que indicar el número de líneas antes de la línea del primer valor de la serie de datos.
2.- *Number of records*: Línea a la que empieza la lista de datos.
3.- *Start Date*: Hay que indicar la fecha de referencia. Siempre el valor 1/1/10.
4.- *Name of the Site*: Nombre del emplazamiento.
5.- *Unis*: Es para introducir las unidades de cada columna de datos. Se puede dejar así y luego modificarlas más adelante o modificarlo en este momento.
6.- *Number decimal separator*: Indicar el separador de decimales en los valores introducidos.
7.- *Columnas loaded from file*: Número de columnas que se desean cargar del fichero. Todas por defecto.

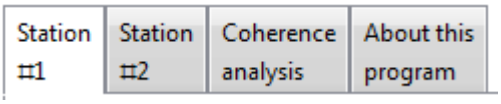
Al terminar con el primer fichero, comienza con el segundo igual y después muestra las cabeceras de los archivos. Si se ha utilizado el programa vientos, no hace falta modificar nada y tan solo hay que dar a todo ok. Las unidades se pueden cambiar luego.

2.3 *Solapas del programa*

El programa se ha organizado por solapas para poder acceder a la información rápidamente. El funcionamiento de las solapas es análogo al comportamiento de las solapas de otros programas como navegadores, editores de textos, etc.

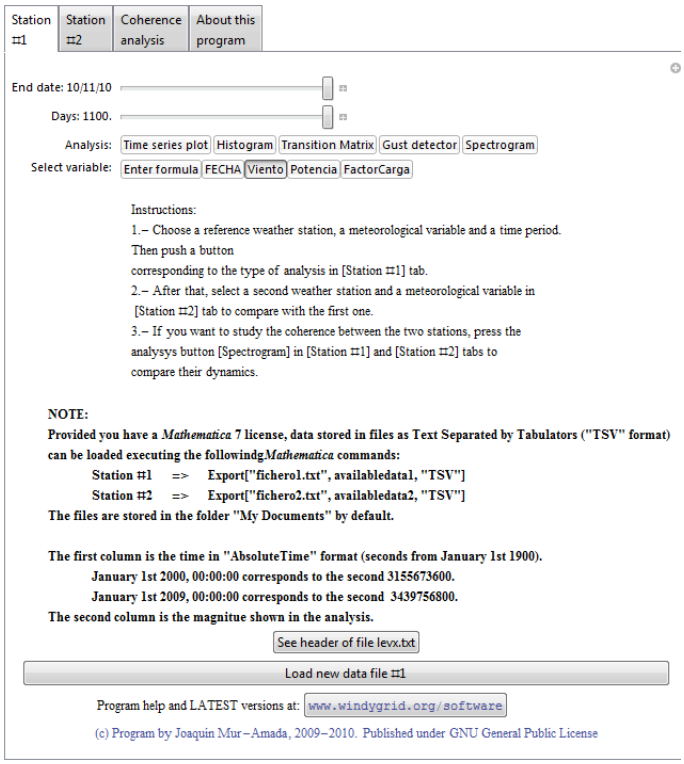
Cada solapa muestra una pantalla con información relacionada estrechamente. Las solapas actualmente disponibles son:

El Programa.



2.3.1 Station #1.

En esta solapa se puede visualizar un análisis de los datos facilitados por las series de datos del primer archivo cargado. Con el control "Period center" se selecciona la fecha del centro del periodo que se va a estudiar. Si el control deslizante está a la izquierda, se utilizarán los primeros datos disponibles y si está a la derecha (opción por defecto), se utilizarán los últimos datos disponibles.



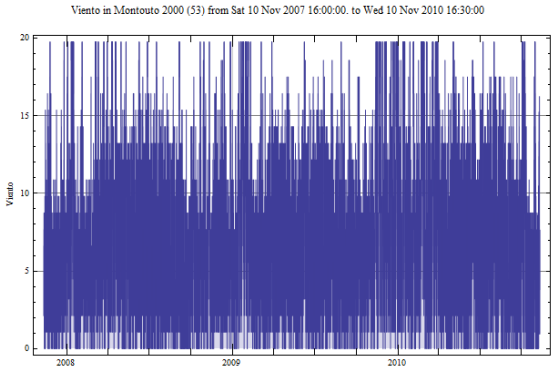
Con el control "Days" se selecciona la duración del periodo bajo estudio. Los periodos muy largos requieren más tiempo de análisis.

La variable a estudiar se selecciona en la lista "*Select a variable*" y dependen del número de columnas del archivo de entrada, que corresponden a diferentes variables.

Los tipos de análisis son:

2.3.1.1 Evolución de la variable en el tiempo (Time series plot).

El primer paso en un estudio de datos meteorológicos es visualizarlos para detectar posibles anomalías en la serie. Esto lo podemos realizar pulsando el botón "*Time series plot*".

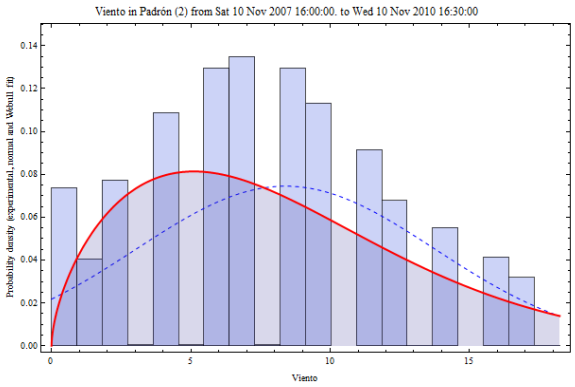


Por defecto, se muestran todos los datos disponibles. Las gráficas de más de una semana resultan poco informativas debido a las variaciones intradía.

2.3.1.2 Histograma (Histogram).

Para estudiar la variación estática del viento o el potencial eólico de un emplazamiento, se utiliza la frecuencia de ocurrencia del viento. En bastantes emplazamientos, la distribución estadística medida se puede aproximar a una distribución Weibull. El histograma también se puede realizar a otras variables meteorológicas, aunque en esos casos, la distribución normal puede ajustarse mejor a las observaciones.

En este caso se nos mostrará un histograma con la frecuencia de observación de la variable seleccionada, comparado con la distribución estadística normal y Weibull ajustada a partir de la media y varianza de los datos. Gráficamente están representadas por una línea discontinua para la distribución normal y continua para la de Weibull.



2.3.1.3 Matriz de transición (Trasition Matrix).

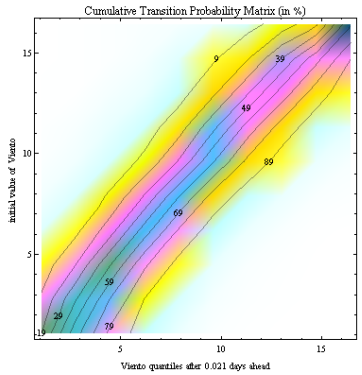
La matriz de transición representa la probabilidad de tener un viento después de un intervalo de tiempo, considerando únicamente la información del viento medido.

La escala vertical es el último valor medido y la escala horizontal valor que se podría obtener después de un intervalo de tiempo. Es decir, cada punto en la figura corresponde a un cambio desde un valor actual a otro valor posible un tiempo después

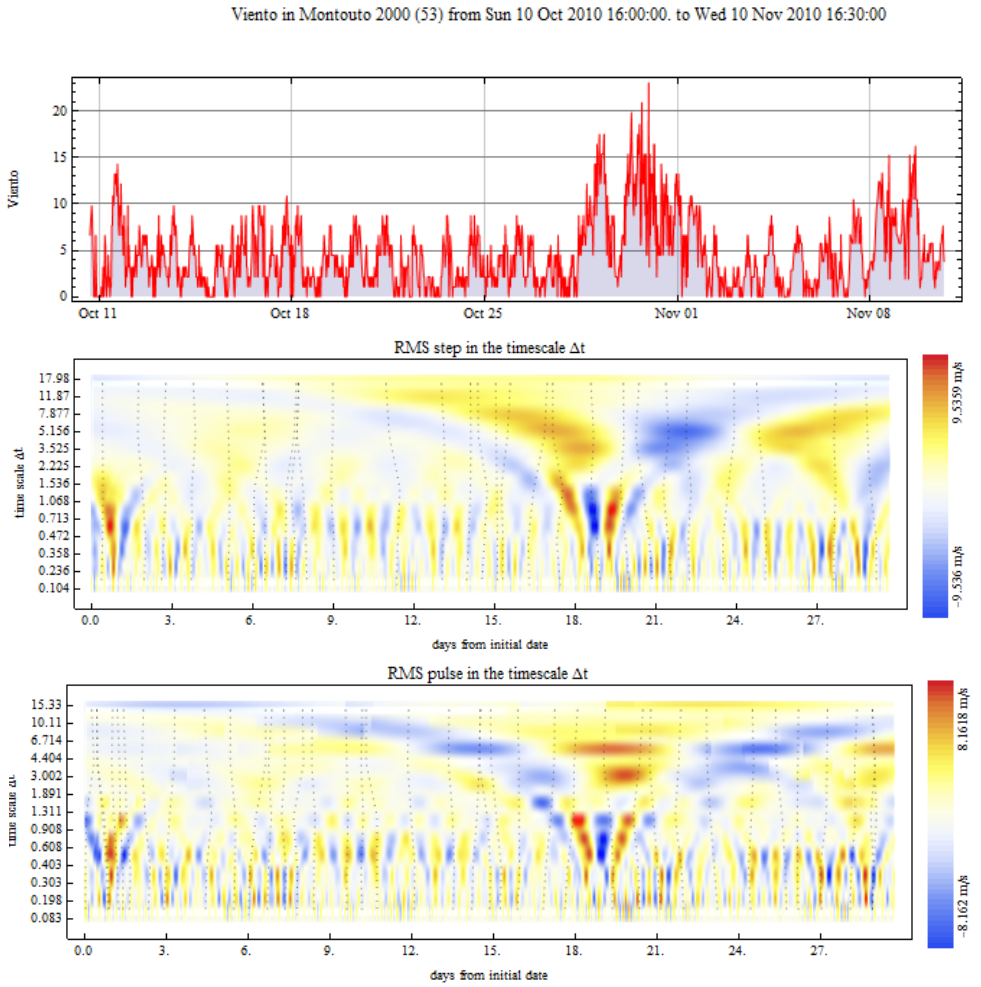
Por conveniencia, la densidad de probabilidad de cada cambio del valor se ha sombreado con colores. Por tanto, el sombreado es más oscuro en los cambios que se ha producido un mayor número de veces.

Sobreimpuesto al sombreado se representa la matriz de probabilidad acumulada mediante líneas que unen los cuantiles de probabilidad. Por ejemplo, entre las líneas 9%

y 89% tenemos el 80% de las transiciones del valor observadas. La distancia entre las líneas cuantiles en un corte horizontal correspondiente al valor actual, nos da la información del rango de valores que esperamos después de un intervalo de tiempo.



2.3.1.4 Detector de rachas (Gust Detector).



Con el programa leerfichero, podemos verla evolución temporal de una variable junto con un detector de duración y amplitud de dicho cambio (RSM Step) y el pulso de dicho cambio (RMS Step). La duración de dicho cambio o pulso corresponde con el eje vertical mientras que la amplitud del cambio o pulso se muestra en la tonalidad.

La señal se filtra con un determinado tiempo de respuesta correspondiente a la escala temporal vertical y se obtiene la diferencia de la variable entre puntos separados

por el intervalo correspondiente a la escala temporal. De esta forma se obtiene el valor de la variación del viento en una escala temporal sin que le afecte las oscilaciones más rápidas y más lentas.

Según la leyenda podemos ver las rampas de bajad y de subida de las rachas, siendo las zonas más rojas las rampas de subida y las azules las rampas de bajada. Cuanto más intenso sea el color, mayor es el salto en valor absoluto.

2.3.1.5 Espectrograma (Spectrogram).

Las fluctuaciones de una magnitud aleatoria estacionaria (cuyas propiedades no varían en el tiempo) se puede realizar calculando la densidad espectral de varianza o PSD (del inglés "*Power Spectral Density*"). La densidad espectral de varianza indica en qué frecuencias se reparte la varianza de la señal. Esto permite cuantificar que porción de la varianza de la magnitud se debe a las oscilaciones diurnas, semidiurnas y de unos cuantos días.

En la práctica, las propiedades del viento y de otras variables meteorológicas cambian en función de la situación atmosférica. Por ello, es conveniente dividir series largas de tiempo en porciones de unos pocos días y observar la distribución de la varianza en la frecuencia. Esto permite constatar si se producen patrones diferentes a lo largo del tiempo.

El espectrograma consiste en una representación del contenido frecuencial de la señal en función del tiempo. Se suele estimar aplicando la transformada de Fourier a pequeñas porciones de la serie temporal. Esto permite estudiar señales en el dominio del tiempo y de la frecuencia conjuntamente. Se emplea habitualmente en audio y en tratamiento de señales de frecuencia variable. Su uso para la energía eólica es novedoso.

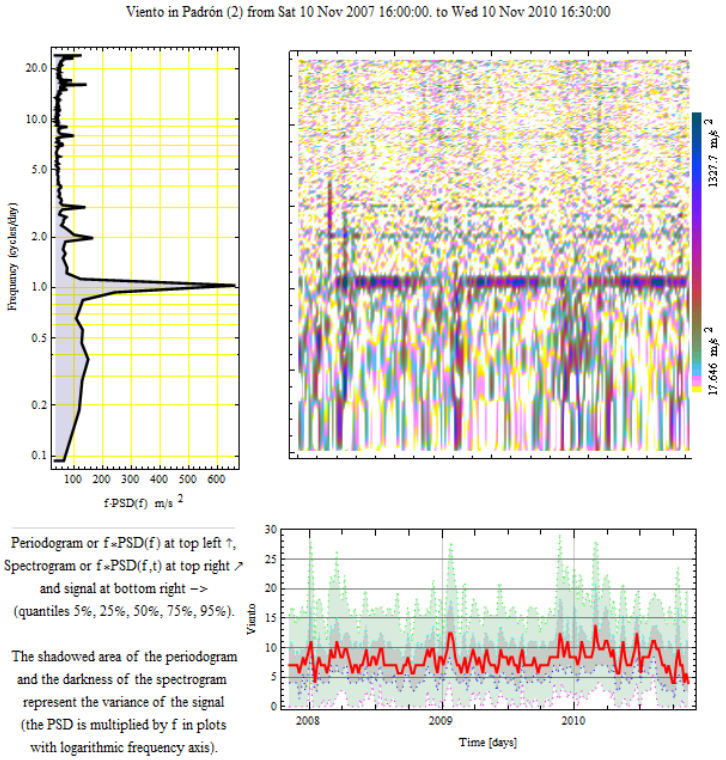
La disposición que se ha elegido para en esta aplicación consiste en:

- el periodograma en la esquina superior derecha;
- el recorrido intercuantil junto con la media y el cuantil 5% y 95% en la esquina inferior derecha;
- el espectrograma, propiamente dicho, en la esquina superior derecha.

El espectrograma muestra las oscilaciones que contiene una señal a lo largo del tiempo. El promedio del espectrograma, es decir, el promedio de las oscilaciones de la señal es el periodograma.

Se ha utilizado la escala logarítmica para la frecuencia porque se observan mejor las oscilaciones de baja frecuencia en las variables meteorológicas. Cuando se utiliza la escala logarítmica en la frecuencia, se suele representar la densidad espectral de la varianza multiplicada por la frecuencia, $f \cdot PSD$, ya que así el área sombreada del periodograma y la oscuridad en el espectrograma representan la varianza de la señal. Es decir, los puntos más oscuros del espectrograma indican que se ha producido una oscilación en el tiempo y frecuencia correspondiente. Las zonas blancas del espectrograma indican ausencias de la frecuencia correspondiente al eje vertical en el instante de tiempo indicado por el eje horizontal.

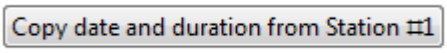
Por ejemplo, los emplazamientos cercanos al mar suelen mostrar oscilaciones semidiurnas (frecuencia $f = 2$ ciclos/día, con 2 picos de viento y 2 periodos de menor viento por día) mientras que en la zona del valle del Ebro las oscilaciones intradía más habituales son de frecuencia $f = 3$ ciclos/día (3 picos locales de viento y 3 periodos de menor viento por día).



2.3.2 Station #2

Uno de los objetivos del programa es comparar las oscilaciones entre dos series de datos. La serie de referencia (*Station #1*) se compara con una segunda serie (*Station #2*) en el dominio tiempo-frecuencia utilizando el ratio de los espectrogramas.

La solapa de la segunda serie es análoga a la primera, salvo que en la segunda estación (*Station #2*). A la hora de comparar, es necesario realizar dicha comparación en el mismo rango de fechas de tiempo. Para ajustar dicho rango basta con pulsar el botón:



Este botón aparecerá una vez seleccionado un rango de fechas y un tipo de estudio en ambos.

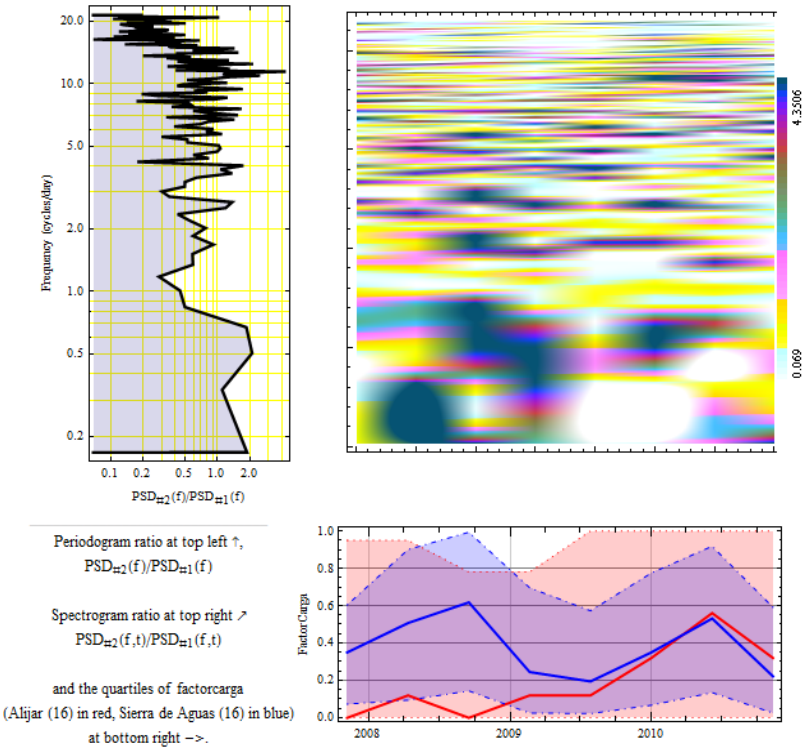
2.3.3 Coherence Analysis

2.3.3.1 Spectrogram Module Ratio

Para señales deterministas, se suele calcular la función de transferencia entre las dos magnitudes dividiendo el espectro de las señales. Utilizando la teoría clásica de sistemas lineales invariantes en el tiempo, las variables meteorológicas se podrían aproximar, a lo sumo, a un sistema multidimensional donde, en vez de utilizar una función de transferencia, habría que utilizar una matriz de funciones de transferencia y en donde las incertidumbres de modelado y medida se considerarían ruido.

Un primer paso sería calcular el ratio de los espectros de la señal de salida (magnitud seleccionada en la segunda estación) respecto a la señal de entrada (magnitud seleccionada en la estación principal). En la práctica, este ratio es variable tanto en módulo como en fase y se representa en una gráfica de intensidad análoga a la utilizada para el espectrograma.

El programa permite seleccionar variables distintas para cada estación meteorológica, e incluso se puede seleccionar la misma estación en la solapa *Station #1* y *Station #2* para comparar las oscilaciones de distintas variables meteorológicas en una misma estación



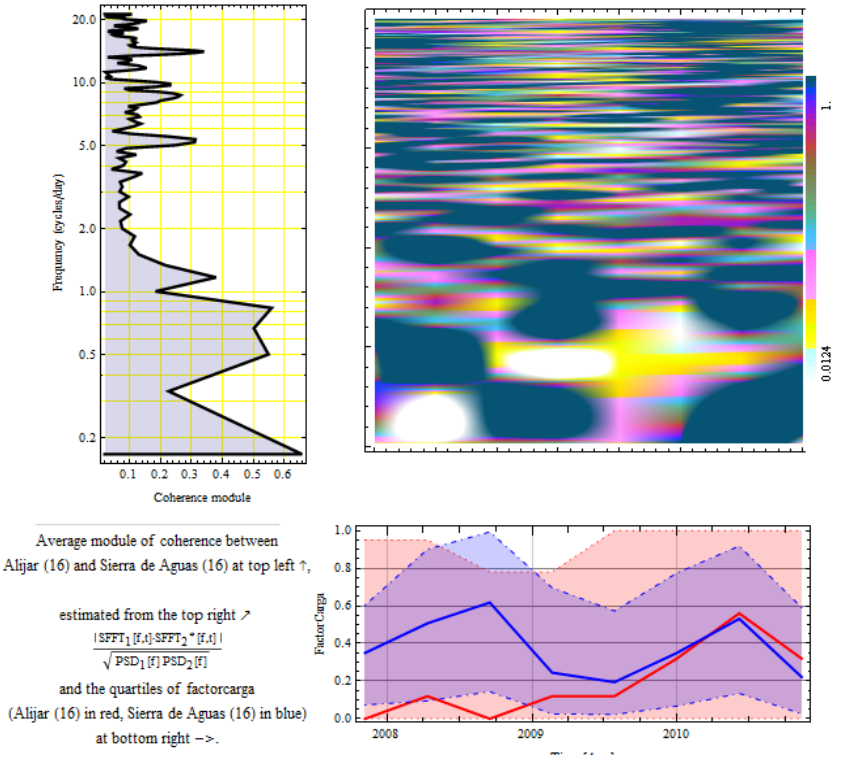
.

Las oscilaciones por debajo de la unidad o con colores claros indican que las oscilaciones tienen mayor magnitud en la estación #1 o de referencia , mientras que las oscilaciones por encima de 1 o colores oscuros indican que las oscilaciones han sido más acusadas en la estación #2 o secundaria.

Se han añadido la gráfica temporal del viento para ambas estaciones y los recorridos intercuantiles (cuantiles 25% y 75%).

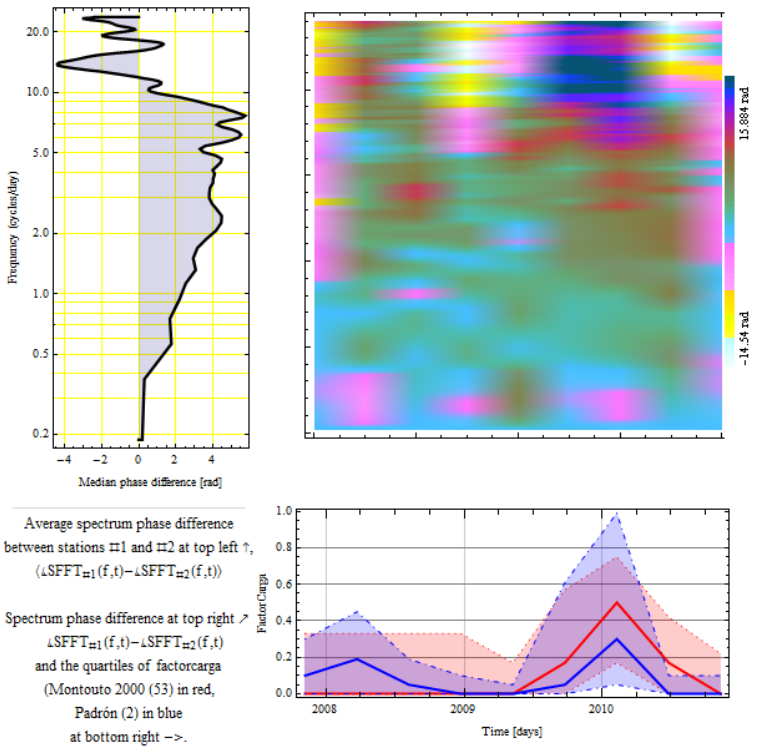
2.3.3.2 Coherence Module.

La coherencia a la frecuencia *f* de las señales no es constante en el tiempo. En el gráfico se puede ver zonas oscuras, que indican que se ha observado en ambas estaciones fluctuaciones significativas en el instante y frecuencia correspondiente. Las zonas blancas indican que no se han observado fluctuaciones significativas de las características correspondientes a las coordenadas en los ejes tiempofrecuencia en, al menos, una de las dos estaciones.



2.3.3.3 Coherence Phase.

La fase de la coherencia indica el retraso medio τ entre las oscilaciones de cada serie temporal de frecuencia *f*. La fase se suele plegar al rango [- π ,+ π] o [0,2 π] por conveniencia, pero el retraso o adelanto entre las oscilaciones pueden exceder un periodo de oscilación (τ > 1/*f*).

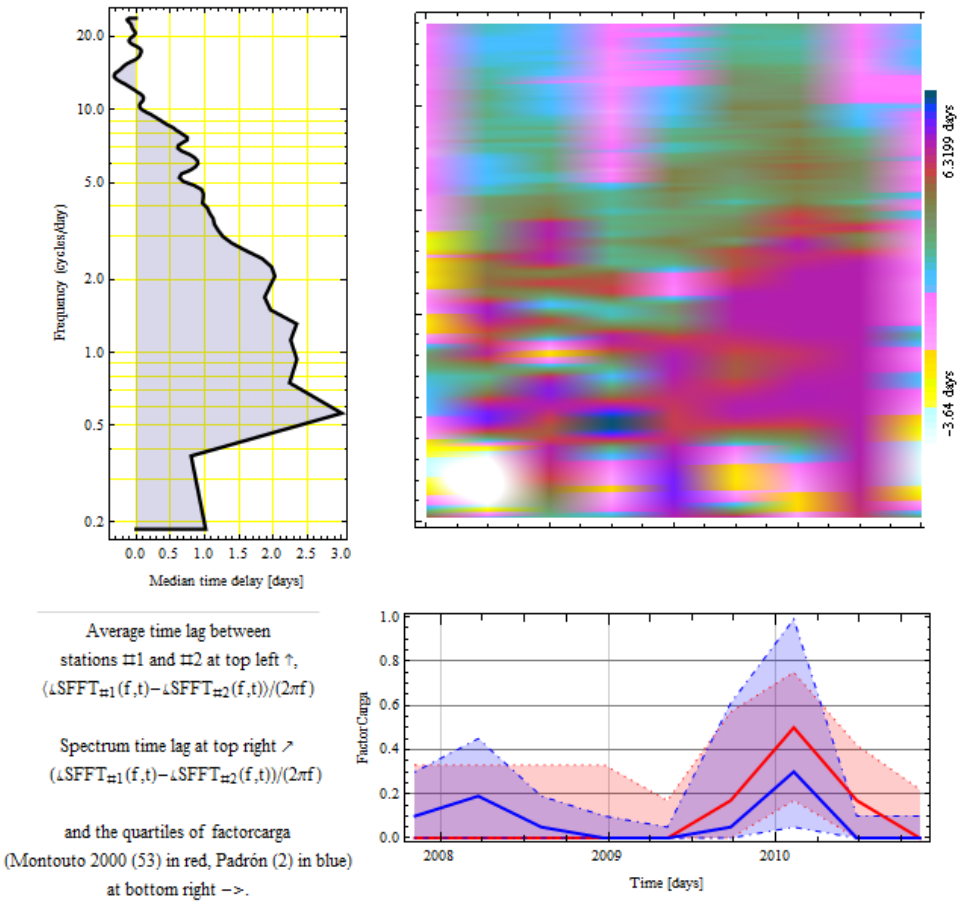


En tal caso, la fase puede exceder del rango [- π ,+ π] o [0,2 π] y es importante desplegar la frecuencia para obtener una estimación real del retraso entre oscilaciones utilizando la fórmula $\tau = \phi /(2 \pi f)$. Los algoritmos básicos detectan saltos entre dos líneas espectrales de más de π radianes, pero en presencia de ruido y fenómenos no lineales es conveniente utilizar métodos más robustos. Para estimar la fase real del espectrograma se pueden utilizar algoritmos 2-D como los que se utilizan en tratamiento de imágenes.

La gráfica siguiente muestra el desfase entre los espectros estimado a la derecha y la fase de la coherencia a la izquierda. Por conveniencia se ha añadido los cuantiles de las dos señales en la parte inferior.

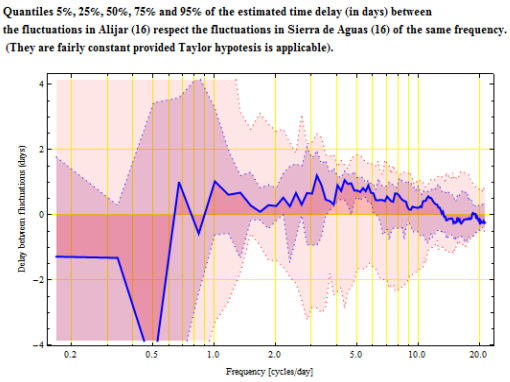
2.3.3.4 Fluctuation Time Delay

Es la igual que el punto anterior pero en vez de ver el desfase en radianes, vemos el desfase en días.



2.3.3.5 Time Delay Quantiles

Como su nombre indica, muestra los cuantiles del Time Delay.

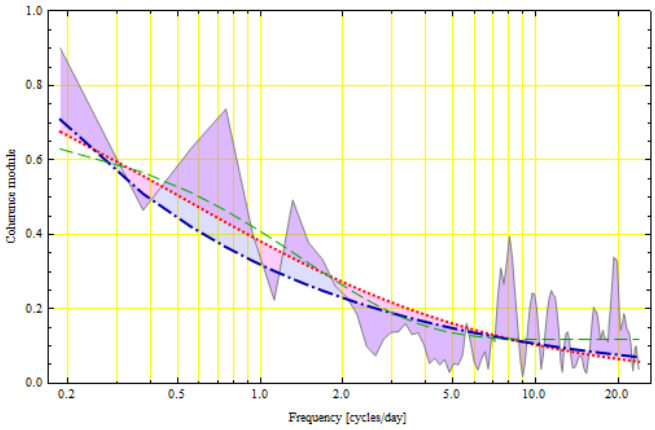
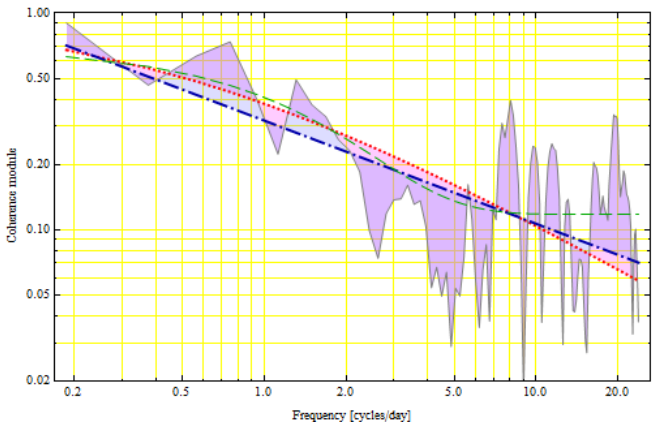


2.3.3.6 Coherence Module Model Fitting.

Se representa gráficamente, la el módulo coherencia entre las estaciones según diferentes modelos.

	Module of coherence[f]	Appearance
Measured coherence	$\frac{\lvert SFFT_1[f,t]\cdot SFFT_2^*[f,t]\rvert}{\sqrt{PSD_1[f]\,PSD_2[f]}}$	thin, grey solid line
Fractional model	$\frac{0.786208}{0.786208+1.27193\,f^{0.727006}}$	thick, dotted red line
Power law model	$\frac{0.319245}{f^{0.476406}}$	dot–dashed blue line
Exponential model	$0.117715+0.583323\,e^{-0.695145\,f}$	thin, dashed green line

Comparision of the models in semi–logarithmic plot:



Comparision of the models in double–logaritmic plot: