

5. Tratamiento de la base de datos

El control de calidad de los datos de la estación de la ETSI ha formado parte de los objetivos de varios de los proyectos desarrollados en la Escuela Superior de Ingenieros, supervisados por el Grupo de Termodinámica. Estos proyectos han centrado su análisis en períodos anteriores al que en este caso se quiere estudiar, pero cabe esperar que aquellos criterios aplicados anteriormente que obtuvieron resultados satisfactorios sean útiles en esta ocasión. Por ese motivo y analizando toda la información de la que disponemos, en este proyecto principalmente seguiremos las directrices indicadas por la “Baseline Surface Radiation Network” [1], una organización con gran prestigio internacional y una larga trayectoria en la toma de medidas de radiación y su control de calidad. Además, completaremos el proceso añadiendo otros criterios de evaluación que se han ido adoptando a lo largo de estos años como consecuencia de la experiencia en el control de calidad de datos de nuestra estación en particular.

5.1 Descripción de los filtros

Cada medida de irradiancia tomada será validada según unos criterios, que llamaremos filtros. No se intentará corregir los valores de la base de datos, sino descartar aquellos que no sean fiables para nuestro estudio. No todos los filtros que aplicaremos serán del todo excluyentes, algunos únicamente nos proporcionarán información sobre la calidad de los datos. A continuación explicaremos en qué consisten los filtros que serán aplicados a la base de datos:

1. Filas completas:

El sistema de adquisición de datos por diversos motivos puede dejar de registrar valores durante algún tiempo y reiniciarse automáticamente en cualquier instante del día. Al reiniciarse pone a cero el valor de la radiación acumulada y continúa funcionando normalmente. En estos casos la irradiación al final del día será menor que la real porque le faltará un tramo. Además, puede dejar en el instante en que se reinició, una fila de datos sin completar. Esto supone un problema a la hora de manejar la matriz con cualquier programa porque puede que éste sólo lea los valores hasta ese instante o incluso no poder leerla considerándolo como un error en la aplicación. En la siguiente figura vemos un ejemplo de este fallo detectado en uno de los archivos:

Archivo	Edición	Formato	Ver	Ayuda							
06:49:27	55	139	171	305	88	189	495	1087	06:21		
06:49:33	55	140	172	306	88	189	495	1090	06:21		
06:49:38	55	140	172	307	88	189	495	1093	06:21		
06:49:44	56	140	172	307	88	190	496	1095	06:21		
06:49:49	56	140	172	308	88	190	496	1098	06:21		
06:49:54	56	141	173	309	88	191	496	1101	06:22		
06:50:00	56	141	173	310	88	191	496	1103	06:22		
06:50:05	56	141	173	311	88	192	497	1106	06:22		
06:50:11	56	142	173	312	88	192	497	1109	06:22		
06:50:16	56	142	174	313	88	193	497	1112	06:22		
06:50:22	55	142	174	314	88	193					
06:59:57	63	0	213	0	89	0	547	0	06:32		
07:00:06	63	1	213	2	89	1	547	6	06:32		
07:00:12	64	1	213	3	89	1	548	9	06:32		
07:00:17	64	1	214	4	89	2	548	11	06:32		
07:00:22	64	2	214	6	89	2	548	15	06:32		
07:00:28	64	2	214	7	89	3	548	18	06:32		

Figura 5.1 Archivo de radiación registrado el día 218 del 2001.

Todos los archivos que presenten este fallo en alguna de sus filas serán revisados visualmente uno a uno para localizar el instante o instantes en los que se producen.

Una vez detectado este tipo de fallo en todos los archivos, hemos tomado la decisión de eliminar aquellas filas que se encuentren incompletas para poder continuar utilizando la información que contienen, ya que el descartar una medida instantánea en un archivo que recoge medidas cada 5 segundos durante todo el día no resulta significativo y facilita el manejo de los datos mediante programas informáticos.

2. Días completos

El sistema de adquisición de datos está programado para que empiece a tomar medidas cierto tiempo antes de la salida del sol y que continúe otro tiempo después de la puesta del mismo. Esta tolerancia que se le introduce al principio y fin de la toma de datos está justificada por dos motivos. El primero es que podemos registrar medidas de irradiancia difusa y global antes del orto y después del ocaso, debido a fenómenos de refracción que se producen en el horizonte. El segundo es que el reloj interno del sistema de control puede desajustarse y calcular mal los instantes límites. Por estos motivos, es interesante clasificar los días en completos e incompletos, sobretodo si los valores de radiación que nos interesan son los diarios.

En la siguiente figura vemos un archivo que presenta este tipo de error:

Time	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Time
10:51:56	136	0	683	0	757	0	722	0	10:19	10:19
10:52:04	136	1	684	7	759	8	721	7	10:19	10:19
10:52:09	136	2	683	10	758	11	709	11	10:19	10:19
10:51:38	149	2	695	10	772	11	710	11	10:19	10:19
10:51:43	150	3	695	14	773	15	710	14	10:19	10:19
10:51:48	151	4	697	17	773	19	711	18	10:19	10:19
10:51:53	150	4	698	21	774	23	711	21	10:19	10:19
10:51:58	150	5	698	23	775	26	712	24	10:19	10:19
10:52:03	152	6	698	27	775	30	711	28	10:19	10:19
10:52:08	156	6	697	30	775	34	711	31	10:19	10:19
10:52:13	153	7	698	34	775	38	712	35	10:19	10:19
10:52:18	152	8	700	37	776	42	713	38	10:20	10:20
10:52:23	153	9	701	41	777	45	714	42	10:20	10:20
10:52:27	152	9	702	44	779	49	715	46	10:20	10:20
10:52:32	150	10	703	48	779	53	715	49	10:20	10:20
10:52:37	150	11	703	52	780	57	715	53	10:20	10:20

Figura 5.2 Archivo de radiación registrado el día 73 del 2000.

3. Saltos.

Como ya se ha indicado en el punto anterior, en determinadas ocasiones el sistema de adquisición de datos se reinicia y comienza de nuevo a funcionar correctamente un tiempo después. El tiempo que tarda en volver al correcto funcionamiento introduce un error a la hora de calcular valores de irradiación, cuya importancia depende de la magnitud temporal del salto y del intervalo de integración de los valores para el cálculo de la irradiación. Es decir, no influirá igual un salto de 50 minutos en el cálculo de la irradiación diaria que en la horaria. Tomaremos en un principio la decisión de descartar aquellos días que presenten saltos superiores a 30 minutos para el cálculo de valores diarios y superiores a 10 minutos para valores horarios. En la siguiente figura podemos ver el registro de datos de un día en el que se produce este tipo de fallo:

Time	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Time
14:09:52	339	6020	398	11845	395	11428	92	13257	13:52	13:52
14:09:57	335	6022	388	11847	384	11430	81	13258	13:52	13:52
14:10:03	332	6023	385	11849	379	11432	79	13258	13:52	13:52
14:10:08	331	6025	385	11851	378	11434	81	13258	13:52	13:52
14:10:13	332	6027	390	11853	382	11436	88	13259	13:52	13:52
14:10:19	333	6029	393	11855	385	11438	92	13259	13:52	13:52
14:10:24	334	6031	397	11858	389	11441	96	13260	13:52	13:52
14:10:30	334	6032	398	11860	390	11443	97	13260	13:52	13:52
14:10:35	334	6034	397	11862	389	11445	92	13261	13:52	13:52
14:10:41	331	6036	392	11864	385	11447	87	13261	13:52	13:52
15:04:16	247	6036	530	11864	587	11447	525	13261	13:53	13:53
15:04:23	239	6038	496	11868	550	11452	483	13266	14:46	14:46
15:04:29	236	6039	486	11871	535	11455	479	13268	14:46	14:46
15:04:34	236	6040	503	11873	545	11457	524	13271	14:46	14:46
15:04:39	241	6042	559	11877	592	11461	610	13274	14:46	14:46
15:04:45	248	6043	599	11880	634	11464	655	13278	14:47	14:47
15:04:50	253	6045	609	11883	652	11468	660	13282	14:47	14:47
15:04:56	254	6046	614	11886	660	11471	670	13285	14:47	14:47
15:05:01	257	6047	626	11890	671	11475	689	13288	14:47	14:47
15:05:06	259	6049	635	11893	681	11479	693	13293	14:47	14:47
15:05:12	259	6050	621	11896	672	11482	668	13296	14:47	14:47

Figura 5.3 Ejemplo de salto >30minutos en el día 255 del 2002

4. Físicamente posibles

Este procedimiento intenta detectar grandes errores en los valores producidos durante la medición. Es el primer filtro que se debe aplicar para controlar las variables de radiación. Los valores de irradiancia, que estén dentro de los intervalos definidos en la tabla 5.1, son considerados “físicamente posibles”. Este procedimiento es tan básico y fundamental que puede ser aplicado a cada valor de irradiancia independientemente de la magnitud del resto de las variables radiométricas e incluso meteorológicas.

Límite inferior	Irradiancia	Límite superior
0	I_{g0}	I_{CS}
0	I_{d0}	I_0+10
0	I_D	I_{CS}

Tabla 5.1 Tabla de intervalos “físicamente posibles”.

En la tabla superior, I_{CS} es la constante solar definida en la expresión [3.22] e I_0 es la irradiancia solar antes de entrar en contacto con la atmósfera o también llamada irradiancia extraterrestre.

En este filtro existen dos tipos de límites, constantes y variables. Para la global horizontal y para la directa normal los límites considerados son constantes y representan el máximo y el mínimo valor que son concebibles en las actuales condiciones climáticas de la tierra. Para la difusa horizontal, el límite superior es la radiación extraterrestre sobre superficie horizontal, esta magnitud se calcula usando la expresión [3.24] donde vemos que su valor depende del factor de corrección de la distancia Tierra-Sol [3.2 y 3.3] variable según el día del año y de la altura solar [3.21] que es función del instante temporal en el que nos encontremos y de la latitud del lugar.

5. Extremadamente raros

En esta prueba los límites del intervalo son más restrictivos que los del filtro anterior. Los valores de irradiancia que traspasen estos límites pueden ocurrir realmente bajo cortos períodos de tiempo o bajo situaciones extremadamente raras. Los límites de los intervalos para la prueba “extremadamente raros” se muestran a continuación:

Irradiancia	Límite Superior
I_{g0}	I_0 si $\theta_z < 80^\circ$ $I_0 + 0.56 (\theta_z - 93.9)^2$ si $\theta_z \geq 80^\circ$
I_{d0}	700 W/m^2
I_D	$I_{CS} E_0 0.9^m$

Tabla 5.2 Tabla de intervalos “extremadamente raros”.

En la tabla superior I_{CS} es la constante solar [3.22], I_0 es la irradiancia solar extraterrestre [3.24], θ_z es el ángulo cenital [3.21], E_0 es la corrección de la distancia Tierra-Sol [3.2 y 3.3] y m es la masa de aire relativa de Kasten (1989):

$$m = \frac{1}{\cos \theta_z} \quad [5.1]$$

Los límites establecidos en la tabla 5.2 están basados en las siguientes asunciones:

Es posible registrar valores de irradiancia global superiores a la extraterrestre en los trópicos cuando nubes aisladas pasan cerca del sol. Pero este fenómeno no suele ocurrir durante más de tres o cuatro minutos. De modo que si la irradiancia global horizontal es mayor que el límite superior durante más tiempo, los valores tendrían que ser revisados. La irradiancia difusa puede alcanzar valores altos cuando el cielo está cubierto con nubes de alta transmitancia y el albedo del suelo es alto. Aunque nunca se ha observado un valor de 700 W/m^2 hasta la fecha. Si un valor excede este límite puede ser que un nuevo record haya sido registrado o que la cúpula de cristal del piranómetro esté cubierta por una pequeña capa de hielo.

La irradiancia directa normal rara vez registra un coeficiente de transmisión mayor de 0.9, ya que sólo la disminución de radiación solar debida a dispersiones de Rayleigh es del 9% al nivel del mar. En altitudes grandes y con una atmósfera seca, es posible observar un coeficiente de transmisión mayor que el valor propuesto. Cuando el valor exceda el límite superior, es recomendable examinar la altitud y el vapor de agua precipitable, asegurándose de que la estación se encuentre en una altitud superior a los 3000 m y que sobre ella hay menos de 1 cm equivalente de agua precipitable.

6. Variables cruzadas

Este filtro intenta recoger errores mucho menores, que pueden escapar de los filtros anteriores. El principio de esta prueba está basado en una relación matemática entre los valores de tres de las variables de radiación medidas.

A continuación, en la tabla 5.3 presentamos las restricciones que aplicaremos en este caso:

Límite inferior	Irradiancia	Límite Superior
$(I_{g0}-I_{d0})-50 \text{ Wm}^{-2}$	$I_D \cos \theta_z$	$(I_{g0}-I_{d0})+50 \text{ Wm}^{-2}$
$I_D \cos \theta_z - 50 \text{ Wm}^{-2}$	$I_{g0} - I_{d0}$	$I_D \cos \theta_z + 50 \text{ Wm}^{-2}$

Tabla 5.3 Tabla de intervalos “variables cruzadas”.

Con este filtro se trata de comprobar la afinidad entre la misma variable medida de dos formas diferentes. Como ya hemos visto, la irradiancia global sobre superficie horizontal es la suma de la difusa horizontal y la directa horizontal:

$$I_{g0} = I_{d0} + I_{D0} = I_D \cos \theta_z \quad [5.2]$$

O lo que es lo mismo:

$$I_D \cos \theta_z = I_{g0} - I_{d0} \quad [5.3]$$

La irradiancia global y difusa son medidas por piranómetros, mientras que la irradiancia directa normal es medida por el pirheliómetro con seguimiento en dos ejes. Aún medidas con distintos aparatos deberían cumplir la igualdad. Aunque 50 W/m^2 parece

ser un error improbable para la medida de irradiancia directa, los puntos exceden con mucha frecuencia este rango. Este error, bastante habitual, es causado principalmente por un mal funcionamiento del seguidor del pirheliómetro o un mal ajuste de la banda de sombra del piranómetro que mide la irradiancia difusa.

Con los filtros descritos en este apartado podríamos hacer dos grandes grupos, uno en el que incluiríamos aquellos que detectan errores que se producen como consecuencia de un mal funcionamiento del sistema de adquisición de datos y otro que podrían aplicarse para el control de calidad de cualquier base de datos de radiación. Dentro del primer grupo incluiríamos los filtros 1,2 y 3, que corrigen aquellos errores repetitivos que se han detectado durante el desarrollo de este y otros proyectos. En el segundo, los tres restantes son aquellos que utiliza la BSRN en el control de calidad de su base de datos. En todo momento, la aplicación de los filtros se complementará con la inspección visual de la representación gráfica de las variables en la medida de lo posible para corroborar los resultados y despejar dudas que puedan surgir.

5.2 Implantación de los filtros y análisis de resultados

Antes de aplicar los filtros, debe hacerse un análisis de los días que a simple vista pueden descartarse por encontrarse vacíos o generados de forma errónea, registrando en todo momento valores nulos, ya que estos podrían dificultar el proceso de filtrado al ejecutar los programas. Por este motivo los siguientes días no serán válidos:

Año 2000: (4 días)

Días: 76, 94, 166, 284.

Año 2001: (1 día)

Días: 347

Año 2002: (1 día)

Días: 216.

Año 2003: (4 días)

Días: 139, 216, 235, 310.

Año 2004: (4 días)

Días: 43, 98, 235, 326.

Año 2005: (2 días)

Días: 209, 219.

Una vez contabilizados los días de cada año con los que inicialmente podemos contar, seguiremos el orden que presentamos a continuación para aplicar los filtros:

1. Filas completas:

Es muy complejo saber en qué archivos aparecen huecos, ya que no parece viable revisarlos visualmente uno a uno, así que aprovecharemos que al ejecutar el programa desarrollado para el cálculo de los valores horarios (que explicaremos en el apartado 7.1) presenta en pantalla un mensaje de error al encontrar un hueco en estos archivos para averiguar en qué días aparecen. Posteriormente, analizaremos visualmente estos archivos y localizaremos en qué filas se encuentran para eliminarlas. Estos días no serán descartados ya que se ha corregido el fallo que introducían. A continuación, enumeramos los días en los que hemos reparado este tipo de fallo:

Año 2000: 120, 224, 325, 353, 327, 62, 90, 81 (8 días).

Año 2001: 113, 159, 176, 218, 228, 254, 26, 295, 296, 298, 302 (10 días).

Año 2002: 102, 134, 14, 156, 179, 185, 198, 217, 270, 91, 98 (11 días).

Año 2003: 111 (1 día).

Año 2004: 174 (1 día).

Año 2005: (0 días).

2. Días completos

Para saber los días que no cumplen este requisito hemos desarrollado un programa en Matlab que genera un archivo “.dat” para cada año que indica a qué hora se produce la salida y la puesta de sol y a qué hora se comienza y terminan a registrar datos para cada uno de los días con los que contamos. El programa compara estos valores, y si la salida y puesta de sol no se encuentran dentro del intervalo de valores registrados los considera incompletos. Si esto se cumple aparecerán como días completos. A continuación, mostramos un ejemplo de los archivos generados con este filtro:

Dias	hss	1ªhreg.	hps	últ.h.reg	
1 Año 2005					
2 Dias no válidos o incompletos:					
3	Dias	hss	1ªhreg.	hps	últ.h.reg
4	90	5.79	9.35	18.21	18.68
5	94	5.71	14.52	18.29	18.77
6	110	5.40	9.13	18.60	19.08
7	129	5.08	4.58	18.92	10.07
8	131	5.05	10.37	18.95	19.45
9	193	4.80	10.35	19.20	19.70
10	206	4.94	4.45	19.06	8.52
11	207	4.96	8.52	19.04	19.53
12	215	5.07	9.05	18.93	19.42
13	223	5.20	7.78	18.80	19.28
14	312	6.88	10.75	17.12	17.58
15	318	6.98	8.07	17.02	17.48
16 Dias válidos o completos:					
17	Dias	hss	1ªhreg.	hps	últ.h.reg
18	1	7.26	6.77	16.74	17.23
19	2	7.26	6.75	16.74	17.23
20	3	7.25	6.75	16.75	17.25

Figura 5.4 Imagen del archivo “completo2005.dat”.

En la figura, la primera columna corresponde a los días, la segunda (hss) es la hora de salida del sol, la tercera es la primera hora a la que comienza a tomar valores el sistema de adquisición de datos (1ªhreg.), la cuarta es la hora de puesta del sol (hps) y la quinta la hora a la que se registra el último valor (últ.h.reg.).

Como se puede observar en la figura 5.4 también sacamos un registro de las variables para los días completos, el motivo es que se puede registrar difusa y global antes de la salida y después de la puesta de sol. De modo que si el sistema de adquisición de datos comienza a tomar medidas, por ejemplo, justo a la salida del sol analicemos ese día con más detalle y decidamos si lo consideraremos completo o no.

Los archivos generados con el programa completo.m para el período de estudio se presentan en el anexo I.

3. Saltos

Para poder detectar los saltos en los archivos, desarrollamos un programa en Matlab que compara la hora de cada valor registrado con la del anterior y si se produce un salto mayor de un tiempo especificado lo almacena en un fichero “.dat”. Este programa lo ejecutaremos para todos los días de cada año generando de esta manera un archivo para cada año con los resultados obtenidos. En función del uso que le vayamos a dar a los datos la duración del salto nos afectará en mayor o menor medida, por eso este filtro lo ejecutaremos de forma diferente para el modelo diario que para el horario. Hemos decidido descartar aquellos días que presenten saltos mayores de 30 minutos para el desarrollo del modelo diario. Para la obtención del modelo horario, saltos de 29 minutos pueden introducir un error bastante elevado a la hora de calcular la radiación horaria, por ese motivo en este caso se ha decidido descartar todos aquellos días que presenten saltos mayores que 30 minutos y además, de los restantes, aquellos que presenten saltos entre 10 y 29 minutos más de 5 veces al día. Los días que, en este último caso, tengan saltos menos de 5 veces al día serán analizados con más detalle para ver si deben o no

ser descartados. Otro motivo por el cual hemos incorporado esta nueva condición es que el sistema de adquisición de datos almacena el vector de tiempo solar 5 segundos después que el resto de las variables, como podemos observar en la siguiente figura:

	0	10	20	30	40	50	60	70		
1347	09:34:35	87	2528	763	18067	522	12861	848	23458	09:12
1348	09:34:41	87	2529	764	18071	522	12864	849	23462	09:12
1349	09:34:46	87	2529	764	18075	523	12867	848	23467	09:12
1350	09:34:51	86	2529	765	18079	523	12869	849	23471	09:12
1351	09:34:56	86	2530	765	18083	523	12872	848	23476	09:12
1352	09:42:22	86	2540	783	18178	544	12938	855	23578	09:12
1353	09:51:27	87	2540	801	18178	569	12938	859	23579	09:19
1354	09:51:35	87	2541	801	18188	569	12945	859	23590	09:29
1355	09:51:40	87	2542	802	18192	569	12948	860	23594	09:29

Figura 5.5 Ejemplo del desfase que incorpora el sistema de adquisición de datos en el vector de tiempo solar.

Aparentemente esto no supone un problema grave ya que el intervalo de tiempo que hay entre dos medidas consecutivas es de 5 segundos, pero puede llegar a convertirse en un error de mayor gravedad si se produce un salto en la toma de datos de gran magnitud. Por ejemplo, si queremos calcular los valores horarios integrando los instantáneos a partir de este vector de tiempos, y se produce un salto justo en el límite de la integración, puede adjudicarse a la hora anterior un valor que pertenece a un instante muy avanzado de la hora siguiente. Por eso, además del filtro para saltos mayores de 30 minutos es necesario incorporar para valores horarios un filtro de 10 minutos.

En la siguiente figura vemos un ejemplo de los archivos que se obtienen al aplicar el filtro para saltos mayores de 30 minutos:

Dias	tiempo(h)	salto(h)	N°rep	
4	125	20.01	0.51	2
5	144	18.53	0.53	7
6	145	19.35	0.57	10
7	146	20.21	0.61	8
8	147	20.26	0.64	9
9	148	9.86	0.67	5
10	156	8.93	0.58	1
11	205	4.90	4.40	1
12	239	18.99	0.51	3
13	240	19.18	0.53	6
14	241	19.22	0.53	7
15	242	18.85	0.54	8
16	243	18.56	0.56	8
17	244	17.45	0.56	5
18	245	17.81	0.56	6
19	246	19.35	0.58	9

Figura 5.6 Imagen del archivo "saltos2002.dat".

Como se ve en la figura 5.5, no sólo registramos los días que no pasan este filtro sino que en la segunda columna se indica el instante en el que se ejecuta el mayor salto registrado durante todo el día, en la tercera su valor y la cuarta nos informa del número de veces que se han encontrado saltos mayores que el límite impuesto.

Todos los archivos generados con el programa saltos.m se presentan en el anexo II.

4. Físicamente posible

Para ver qué días no pasan este filtro se ha desarrollado un programa en Matlab, que genera un archivo “.dat” para cada año, comprueba si todos los valores almacenados en los ficheros “arad” cumplen las condiciones impuestas en la tabla 5.1. El archivo generado nos informa de los días que no pasan el filtro, del número de instantes al día que toma valores fuera de los márgenes esperados y por último el máximo valor que se registra como erróneo a lo largo del día. A continuación vemos una imagen que muestra un ejemplo de fichero generado:

Año2000							
Dias que no pasan el filtro "fisicamente posible"							
	Dia	Ig	N°rep.	Idif	N°rep.	IrD	N°rep.
4	1	0	0	388	977	0	0
5	2	0	0	994	946	0	0
6	3	0	0	676	992	0	0
7	4	0	0	816	1127	0	0
8	5	0	0	1327	1096	0	0
9	6	0	0	1160	1057	0	0
10	7	0	0	301	1052	0	0
11	8	0	0	680	933	0	0
12	9	0	0	1032	882	0	0
13	10	0	0	0	0	0	0
14	12	0	0	1267	1509	0	0
15	19	0	0	336	982	0	0
16	20	0	0	258	1002	0	0
17	25	0	0	285	1553	0	0
18	26	0	0	697	1127	0	0
19	27	0	0	1009	1340	0	0
20	28	0	0	598	1107	0	0
21	29	0	0	730	1558	0	0
22	30	0	0	661	1556	0	0

Figura 5.7 Imagen del archivo “fisicamenteposibles2000.dat”.

Para evitar descartar días por fallos que van a ser subsanados más adelante, el filtro se aplicará para instantes en que la altura solar sea mayor que 10° . Ya que es posible registrar valores extraños cuando el sol está tan bajo por fenómenos de refracción.

Los días en los que aparece en los ficheros que la radiación global en algún instante no ha sido físicamente posible serán descartados. Los días en que la radiación difusa no es físicamente posible en algún instante, sólo serán descartados para los casos en que calculemos los K_d utilizando esta variable, al igual que haremos en el caso de la directa al calcular los K_{dD} .

Todos los archivos generados con el programa fisicamenteposible.m para el período de estudio se adjuntan en el anexo III.

5. Extremadamente raros

Par aplicar este filtro también desarrollaremos un programa en Matlab, similar al del caso anterior, generando archivos “.dat “ con la misma estructura pero imponiendo en este caso que las medidas registradas se encuentren dentro de los límites definidos en la

tabla 5.2. Los archivos generados con este programa serán como el que mostramos a continuación:

3	Day	Ig	N°rep.	Idif	N°rep.	IrD	N°rep.
4	57	1298	567	0	0	0	0
5	59	928	13	0	0	0	0
6	61	997	1728	0	0	0	0
7	62	1004	480	0	0	0	0
8	72	1072	12	0	0	0	0
9	82	1246	72	0	0	0	0
10	85	1197	8	0	0	0	0
11	88	1126	7	0	0	0	0
12	92	1122	68	0	0	0	0
13	93	1295	248	0	0	0	0
14	94	1296	230	0	0	0	0
15	96	1323	103	0	0	0	0
16	98	1250	66	0	0	0	0

Figura 5.8 Imagen del archivo "extremadamentraros2000.dat".

En este caso el filtro tampoco se aplicará para instantes en que la altura solar sea mayor que 10° , por el mismo motivo que en el caso anterior.

Los resultados obtenidos al aplicar este filtro han de ser analizados con mayor detalle.

Si observamos los ficheros generados, veremos que son pocos los días que debido a la radiación difusa o a la directa se sobrepasan los límites establecidos. Por este motivo, y porque no existe ninguna explicación física que justifique que estas dos variables alcancen valores de esta magnitud, los días en que esto ocurra serán eliminados. Sin embargo cabe la posibilidad de que el límite de la radiación global se exceda en algún instante del día, como ya explicamos en el apartado anterior. Y además, observamos que para todos los años se excede durante la misma época y únicamente los días nublados, que es cuando puede aparecer este fenómeno. Al no encontrar nada que justifique los resultados obtenidos, no eliminaremos los días que sobrepasen el límite superior del filtro "extremadamente raros" para los valores de radiación global.

Todos los archivos generados con el programa extremadamentraros.m para el período de estudio se adjuntan en el anexo IV.

6. Variables cruzadas

Para aplicar este filtro también desarrollaremos un programa en Matlab que genera un archivo con los resultados obtenidos. En este caso el programa calculará para cada instante el valor de la irradiancia directa sobre superficie horizontal de dos formas diferentes como vimos en el apartado 5.1. Si la diferencia que existe entre el valor de dicha variable obtenida por los dos caminos es superior a 50 W/m^2 en valor absoluto, se considerará que para ese día no se ha pasado el filtro. En el archivo que generará el programa, también se indicará el número de veces que no se cumple esta condición y la mayor diferencia obtenida para cada caso. A continuación, mostramos un ejemplo de archivo generado a partir de este programa:

1	Año2001					
2	Dias que no pasan el filtro "Variables cruzadas"					
3	Dia	Ig0-Id0 (W/m2)	ID*cos(z) (W/m2)	Max.dif (W/m2)	N°exc	N°def
4	1	60	18	42	0	0
5	100	769	725	44	0	0
6	101	523	494	29	0	0
7	102	421	386	35	0	0
8	103	445	405	40	0	0
9	104	578	535	43	0	0
10	105	528	472	56	1	0
11	106	444	368	76	144	0
12	107	501	241	260	564	29
13	108	458	272	186	108	11
14	109	369	255	114	47	1
15	11	299	163	136	201	0

Figura 5.9 Imagen del archivo "variables cruzadas2000.dat".

Como podemos ver en la tercera fila, la segunda columna representa el valor de la irradiancia directa sobre superficie horizontal calculada a partir de la global y la difusa y la segunda a partir de la directa normal. En ambas columnas se almacena el valor de la irradiancia directa sobre superficie horizontal para el instante del día en que aparece la mayor diferencia, cuyo valor se encuentra en la cuarta columna. La quinta y la sexta nos informan del número de veces que las variables difieren entre ellas una cantidad mayor de 50 W/m^2 por defecto y cuantas veces por exceso:

$$\text{Por exceso: } (I_{g0} - I_{d0}) - I_{D} \cos \theta_z > 50 \text{ W/m}^2 \quad [5.4]$$

$$\text{Por defecto: } (I_{g0} - I_{d0}) - I_{D} \cos \theta_z < 50 \text{ W/m}^2 \quad [5.5]$$

Hemos hecho esta distinción, a la hora de contabilizar los errores para poder adjudicar cada uno al aparato de medidas responsable. De esta forma, si los errores se producen por exceso, esto implica que los valores de directa son menores de lo que deberían ser, y es muy probable que se haya desalineado el pirheliómetro. Si ocurre todo lo contrario, y el error es por defecto, que significa que la difusa ha registrado valores más elevados de lo normal, lo más probable es que no se haya corregido la posición de la banda de sombra y que las medidas de irradiancia difusa sean cercanas a la de irradiancia global. Una vez obtenido los resultados, nos damos cuenta que son muchos los días en que se sobrepasa el margen de error considerado. Así que decidimos no eliminarlos todos, sino únicamente aquellos que no cumplan la condición de forma extrema. Para ello, tendremos que hacer un análisis más detallado de los errores encontrados y ver si pueden motivar el rechazo del día.

Hay casos en los que claramente se reconoce el error a simple vista y otros que deben complementarse con las representaciones gráficas de las variables.

Entre los casos que pueden verse a simple vista encontramos aquellos donde el número de veces que se registran fallos es muy elevado y siempre con la misma tendencia (por defecto o por exceso) y aquellos que son menos elevados en número pero más en magnitud. En estos dos casos los días se han considerado no válidos. El resto de resultados sólo nos va a informar de la calidad de los datos y nos ayudará a tomar decisiones más adelante a la hora de considerar como válidos algunos valores K_d y K_t .

A continuación mostramos dos ejemplos de los casos que a simple vista nos informa de un mal funcionamiento de los aparatos de medida:

	0	1.0	2.0	3.0	4.0	5.0	6.0
53	151	405	343			62	66
54	152	883	832			51	6
55	153	860	31			829	9631
56	154	860	1			859	9624
57	155	858	0			858	9544
58	156	738	0			738	8858
59	157	725	5			720	5907
60	158	705	5			700	5245
61	159	330	2			328	711
62	16	268	168			100	998
63	160	401	2			399	2076
64	161	247	1			246	1187
65	162	733	5			728	8479
66	163	684	0			684	8649
67	164	732	3			729	9120
68	165	753	0			753	7467
69	166	798	0			798	9640
70	167	307	0			307	1020
71	168	721	701			20	0
72	169	792	773			19	0

Figura 5.10 Imagen del archivo “variables cruzadas2004.dat”, mal funcionamiento del pirheliómetro.

	0	1.0	2.0	3.0	4.0	5.0	6.0
111	230	112	229			117	0
112	231	109	336			227	46
113	232	110	438			328	0
114	233	87	386			299	0
115	234	-6	518			524	0
116	235	-21	623			644	0
117	236	-73	615			688	5
118	237	-107	669			776	2
119	238	-132	723			855	66
120	239	-151	727			878	1
121	240	-169	774			943	0
122	241	-147	727			874	0
123	242	-100	552			652	3
124	243	-153	659			812	0
125	244	-157	698			855	0
126	245	-128	550			678	0
127	246	-152	666			818	0
128	247	-188	759			947	0
129	248	-79	525			604	5
130	249	284	198			86	7

Figura 5.10 Imagen del archivo “variables cruzadas2000.dat”, mal funcionamiento de la banda de sombra.

Estos ficheros y el resto que completan el período de estudio podemos encontrarlos en el anexo V.

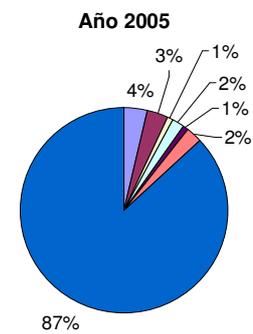
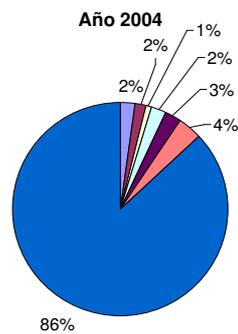
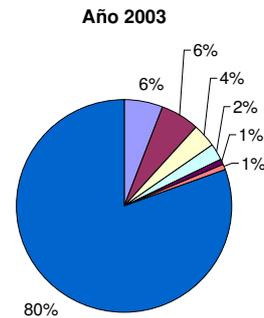
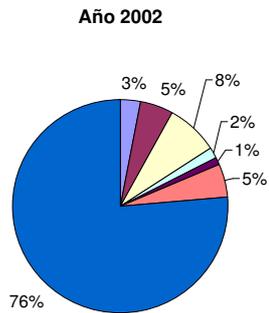
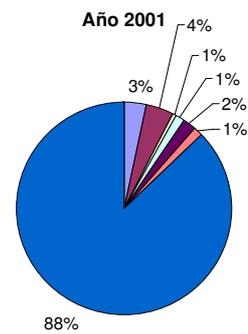
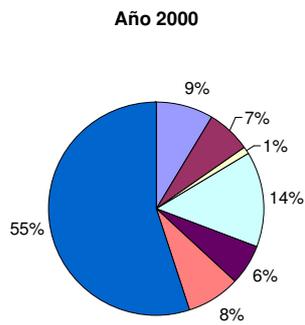
A continuación, presentamos una tabla que nos informa del número de días que han pasado cada uno de los filtros y el porcentaje de días que supone respecto a los que deberían haber sido almacenados en la base de datos:

- Para el uso de valores diarios:

➤ **Difusa:**

	Días no registrados	Incompletos	saltos 30 min.	físicamente imposibles	Extremad. raros	variables cruzadas	Días válidos
2000	32	24	4	53	22	30	201
2001	12	15	2	5	8	5	318
2002	11	18	29	6	4	18	279
2003	21	22	13	9	3	3	294
2004	8	6	3	8	10	13	318
2005	13	12	3	7	4	9	317
						Total	1727

Tabla 5.4 Número de días que no han pasado cada uno de los filtros.

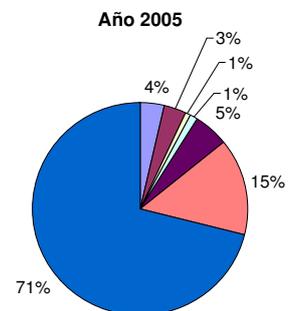
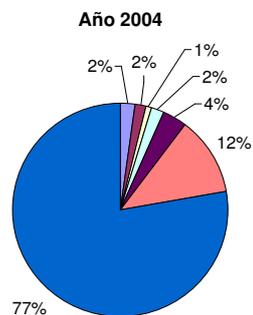
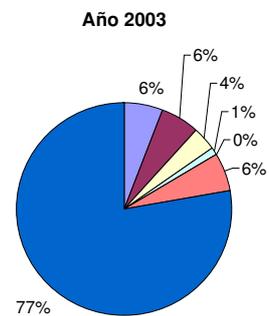
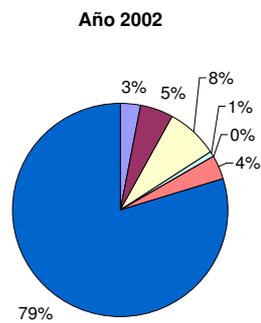
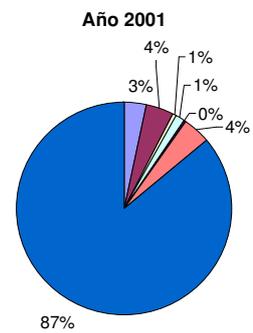
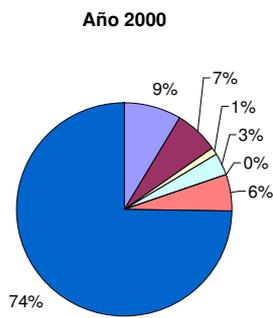


Días no registrados	Incompletos	Saltos 30 min.	Físicamente imposibles
Extremadamente raros	Variables cruzadas	Días válidos	

➤ **Directa:**

	Días no registrados	Incompletos	Salto 30 min.	físicamente imposibles	Extremad. raros	variables cruzadas	Días válidos
2000	32	24	4	12	0	21	273
2001	12	15	2	5	1	16	314
2002	11	18	29	3	0	13	291
2003	21	22	13	4	0	21	284
2004	8	6	3	7	14	44	284
2005	13	12	3	4	20	53	260
						Total	1706

Tabla 5.5 Número de días que no han pasado cada uno de los filtros.



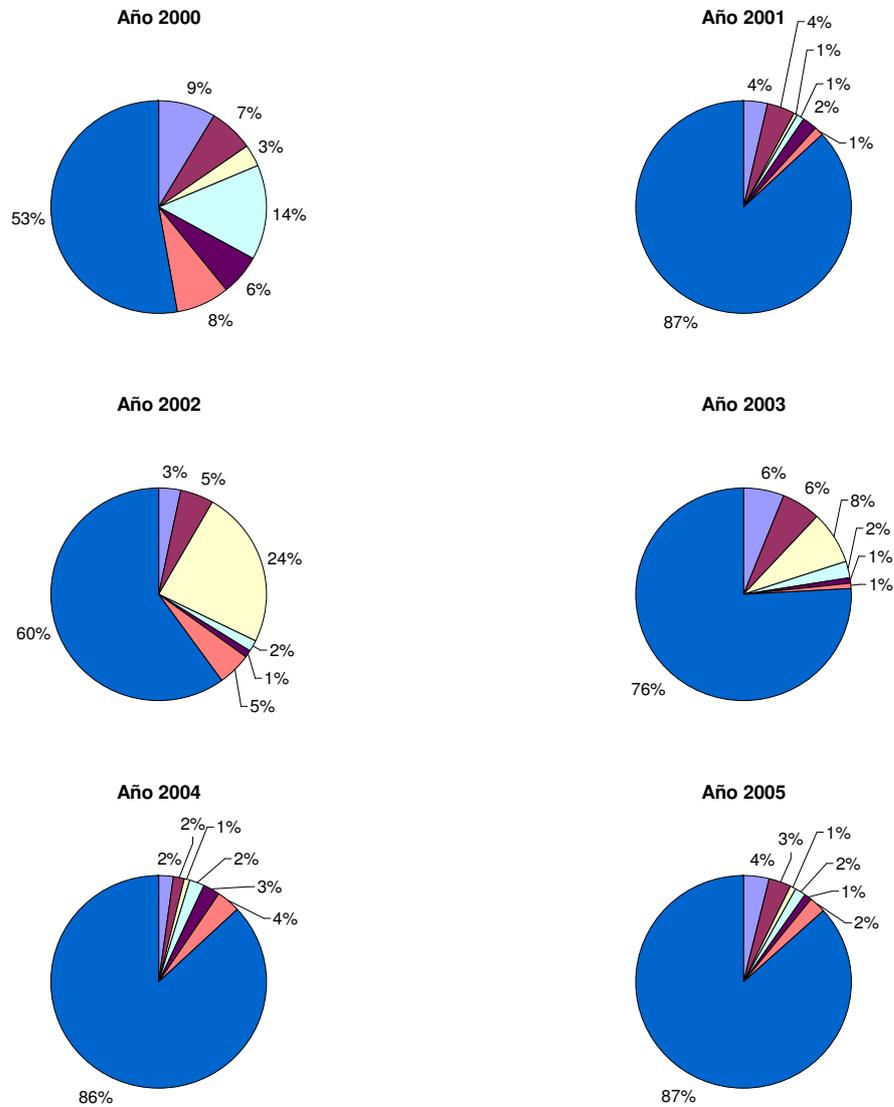
■ Días no registrados	■ Incompletos	■ Salto 30 min.	■ Físicamente imposibles
■ Extremadamente raros	■ Variables cruzadas	■ Días válidos	

• Para el uso de valores horarios:

➤ Difusa:

	Días no registrados	Incompletos	Salto 10 min.	físicamente imposibles	Extremad. raros	variables cruzadas	Días válidos
2000	32	24	12	53	22	30	193
2001	13	15	2	5	8	5	318
2002	12	18	88	6	4	18	220
2003	22	22	29	9	3	3	278
2004	8	6	3	8	10	13	318
2005	14	12	3	7	4	9	317
Total							1644

Tabla 5.6 Número de días que no han pasado cada uno de los filtros.



➤ **-Directa:**

	Días no registrados	Incompletos	Salto 10 min.	físicamente imposibles	Extremad. raros	variables cruzadas	Días válidos
2000	32	24	12	12	0	21	265
2001	12	15	2	5	1	16	314
2002	11	18	88	3	0	13	232
2003	21	22	29	4	0	21	268
2004	8	6	3	7	14	44	284
2005	13	12	3	4	20	53	260
						Total	1623

Tabla 5.7 Número de días que no han pasado cada uno de los filtros.

