**Documento explicativo para el procesado de los Langleys. Métodos Brewer / Dobson.**

Se explica en este documento cual es el proceso y cuales las funciones involucradas en el análisis de los Langleys plots para el Brewer. Las funciones involucradas en el proceso se hallan todas bajo el directorio /matlab/Langley/Langley\_dev.

Se trabaja con dos metodologías diferentes, correspondientes a la usada en la red Dobson y a la que se puede considerar clásica:

* Classic Algorithm :

Siendo , donde (βi Rayleigh coefficient, m3 ozone airmass y m2 Rayleigh airmass). Si asumimos Ozono constante, entonces podremos hacer ,  es decir, la ETC será el corte con el eje de ordenadas para m3=0.

* Dobson Algorithm: En este caso se aplica la estrategia ***1/m3*** usada por los Dobsons, resultando

Lo que se obtiene ahora, a través de la pendiente ETCc, es la corrección a la ETC **estimada** ETCo. Aunque ambas metodologías serían esencialmente equivalentes (*Comparison of Plotting Methods for Solar Radiometer Calibration, S. M. ADLER-GOLDEN and J. R. SLUSSER: “the choice of the traditional Langley plot versus the alternative method is more a matter of taste than accuracy”*), las ventajas del método Dobson frente al método Brewer serían:

1. Las observaciones están mejor distribuidas frente a 1/m3 que frente a m3.
2. El error cometido al determinar la pendiente es inferior al error cometido al determinar el corte en m3=0 (aunque esto dependerá de la *m3* máxima usada).

Puesto que se ha respetado igual estructura para las matrices de datos independientemente de que datos sean los usados, podremos trabajar bien con los ***sumarios*** bien con las ***medidas individuales*** (5 por cada sumario). Además, podremos trabajar o bien con las medidas de un instrumento particular o bien con las medidas sincronizadas entre diferentes instrumentos. El tiempo de sincronización para los sumarios es Tsync = 5 minutos, mientras que en el caso de las medidas individuales se usa un tiempo de sincronización Tsync = 0.5 minutos en el caso de trabajar con medidas simultáneas. Notar que este parámetro es definido *hardcode* en todos los scripts, es decir, en caso de querer modificarlo, habría que hacerlo directamente en el código.

Tabla : Sumario de variables devueltas por la función read\_bdata.m

|  |  |
| --- | --- |
| **RAW DATA (as read from B files)** | |
| **ozone.raw0** | Salidas raw, tomadas directamente del fichero B (med. individuales) -> B config (1). Ratios & individual counts |
| **ozone.dsum** | Salidas raw, tomadas directamente del fichero B (sumarios) -> B config (1). Ratios |
| **RECALCULATED DATA (using defined configuration constants)** | |
| **ozone.raw** | Salidas recalculadas (med. individuales). Raw pmt counts & counts / second recalc (1 & 2 configs). No MS9`s (no ratios) |
| **ozone.ozone\_ds** | Ind. measurements, recalculated (1 & 2 configs). Ratios & MS9’s |

Tenemos entonces las siguientes cuatro posibilidades: sumarios, sumarios simultáneos (entre distintos instrumentos), medidas individuales y finalmente medidas individuales simultáneas (entre distintos instrumentos). Al igual que ocurría con los dos métodos posibles de análisis de los langley plots, también en este caso se encuentran resultados prácticamente idénticos para todos las posibilidades, de modo que debería prevalecer un criterio de economía, quedándonos con el caso más sencillo (medidas individuales).

Describo ahora brevemente como trabajar en cada uno de los casos. Para todos ellos, obtener finalmente valores para la constante extraterrestre implica 3 pasos:

1. Transformar las lecturas de los ficheros B en un formato común de trabajo: funciones ***langley\_data\_cell.m*** / ***langley\_data\_cell\_summ.m***
2. Aplicación o no de determinadas condiciones de filtrado de datos raw, función ***langley\_filter\_lvl1.m*** -> paso opcional, pero con mucha precaución ¡! (ver filtros)
3. Análisis de los langley plots usando los métodos clásico y alternativo. Función ***langley\_analisys.m***

Adicionalmente, en el caso de querer trabajar con medidas simultáneas entre varios equipos habrá que hacer uso, a continuación del paso 1), de las funciones ***langley\_indv\_sync.m*** / ***langley\_summ\_sync.m***

* Sumarios

Para trabajar de esta manera se asume que contamos con los sumarios ya generados, es decir, habrá que seguir el procedimiento ordinario ***read\_bdata -> sl\_report\_jday -> read\_cal\_config\_new -> test\_recalculation -> filter\_corr***

No voy a entrar en detalles acerca de cómo obtener las variables summary y summary\_old, pero si conviene tener en cuenta lo siguiente:

* **Aplicamos la función filter\_corr.m a los sumarios**. Entonces tendremos MS9 con y sin corrección por filtros, campos 8 y 9 de summary, respectivamente.
* **No aplicamos la función filter\_corr.m a los sumarios**. Entonces tendremos MS9 sin corrección por filtros (campos 8). El campo 9 es ahora la desviación estándar

Tabla 2: Salida de la función langley\_summ\_sync.m

|  |
| --- |
| ***time, lat, long, sza, m2, m3, flag,  NaN, tst, filt, temp*** ... % 1-11  ***NaN, NaN, NaN, NaN, NaN, NaN, NaN***  ... %  ***O3 old, NaN, NaN, NaN, NaN, NaN, MS9 old***  … % 25 ratios recalculated 1 (Rayleigh corr.!!)  ***NaN, NaN, NaN, NaN, NaN, NaN, NaN*** ... %  ***O3 old, NaN, NaN, NaN, NaN, NaN, MS9 old***  ... % 39 ratios recalculated 2 (Rayleigh corr.!!) |

Por defecto se trabaja con el campo 9, de modo que será importante aplicar siempre la función filter\_corr.m. Es decir, por defecto tendremos inicialmente las ratios MS9 no corregidas, y será en el paso 2 del esquema general anterior cuando apliquemos las correcciones por filtros.

Las funciones con las que trabajaremos en el caso de trabajar con los sumarios serían ***langley\_data\_cell\_summ.m*** para convertir los sumarios al formato común de los Langleys (variables-inputs: summary, summary\_old, config; variables-outputs: ozone\_lgl\_sum, cfg, lgl\_leg) y, en el caso de querer trabajar con medidas simultáneas entre instrumentos, ***langley\_summ\_sync.m*** (variables-inputs: ozone\_lgl\_sum y variable de definiciones Cal).

La salida de ambas funciones será un cell-array para cada instrumento. Cada una de estas celdas contiene a su vez una celda para cada día conteniendo una matriz de datos de 39 columnas con la estructura mostrada en la Tabla 2.

* Medidas Individuales

Trabajaremos en este caso con las ratios MS9 individuales, es decir, con las 5 componentes de cada sumario. No es directo obtener una variable con los sumarios que contenga estas ratios y además las cuentas recalculadas (o brutas) para cada canal, y es por esto por lo que combinamos en una las dos variables ozone\_raw (cuentas) y ozone\_ds (ratios). Notar que únicamente de este modo seremos capaces de calcular Langleys para cada canal (por ejemplo para AOD). A partir de las medidas individuales recalculamos el ozono, que se puede usar con dos objetivos: 1) depurar medidas de ozono, descartando todos los grupos de 5 medidas que no verifiquen el criterio desviación estándar < umbral (umbral sería un argumento opcional, que por defecto toma el valor 2.5 DU) y 2) obtener el ozono, que se podría usar en un siguiente paso para descartar condiciones de ozono diario muy variable.

Las funciones con las que trabajaremos serán ***langley\_data\_cell.m*** (variables-i*nputs:* ozone\_raw, ozone\_ds y config; variables-outputs: ozone\_lgl, cfg, lgl\_leg) para obtener variables con el formato común de los Langleys y además, en el caso de querer trabajar con medidas simultáneas entre instrumentos, ***langley\_indv\_sync.m*** (variables-inputs: ozone\_lgl y variable de definiciones Cal).Las salidas de la función ***langley\_data\_cell.m*** serán:

Tabla 3: Salida de la función langley\_data\_cell.m: variables ozone\_lgl y ozone\_lgl\_sum

|  |
| --- |
| ***'date' 'hg\_id' 'nds' 'sza' 'm2' 'm3' 'sza' 'saz' 'tst' 'filt' 'temp'*** *...% 1-11*  ***'f0' 'f1' 'f2' 'f3' 'f4' 'f5' 'f6' ...*** *% 12-18 count-rates recalculated 1 (Rayleigh uncorrected !!)*  ***'o3\_1' 'r1' 'r2' 'r3' 'r4' 'r5' 'r6' ...*** *% 19-25 ratios recalculated 1 (Rayleigh corrected !!)*  ***'F0' 'F1' 'F2' 'F3' 'F4' 'F5' 'F6' ...*** *% 26-32 count-rates recalculated 2 (Rayleigh uncorrected !!)*  ***'O3\_2' 'R1' 'R2' 'R3' 'R4' 'R5' 'R6' ...*** *% 33-39 ratios recalculated 2 (Rayleigh corrected !!)* |

Tabla : Salida de la función langley\_data\_cell.m: variable cfg

|  |
| --- |
| ***'Usage date', 'o3 Temp coef 1', 'o3 Temp coef 2', 'o3 Temp coef 3', 'o3 Temp coef 4', 'o3 Temp coef 5', 'O3 on O3 Ratio', 'ETC on O3 Ratio', 'Dead time (sec)', 'ND filter 0', 'ND filter 1', 'ND filter 2', 'ND filter 3', 'ND filter 4', 'ND filter 5'*** |

* ***ozone\_lgl***: serían las medidas individuales (5 por cada sumario), depuradas según el procedimiento estándar usado con los sumarios (normal O3\_std < 2.5, 100<O3<600, hg flag =1, n = 5).
* ***cfg***: Constantes de calibración en juego. Estructura con campos old y new (ver Tabla 4).
* ***lgl\_leg***: legenda para cada una de las variables de salida
* ***ozone\_lgl\_sum***: sumarios recalculados a partir de las medidas individuales anteriores. Los campos disponibles son los detallados en la Tabla 3. Obtendremos esta variable sólo en caso de pedirla.

General_.tif

Figure 1: Resultado del análisis de los Langley Plots para un instrumento particular y para 6 conjuntos de diferentes obtenidos a partir de datos simultáneos / no simultáneos con / sin aplicación de filtros de depuración. Métodos Brewer (arriba) y Dobson (abajo).

Como se ve, independientemente de si trabajamos bien con sumarios bien con medidas individuales, la estructura de los datos de análisis es idéntica. Esto nos permitirá desarrollar las funciones de análisis de forma que sean comunes para cualquier tipo de dato de entrada.

En la Figure 1 se muestran como ejemplo los resultados de los Langleys obtenidos de los sumarios calculados a partir de la función ***test\_recalculation.m*** y los obtenidos a partir de las medidas individuales, función ***langley\_data\_cell.m***. Los resultados en ambos casos son idénticos.

Paso a comentar a continuación primero la depuración de los datos y luego el análisis de los Langleys.

1. **Depuración de los datos, nivel 1 (langley\_filter\_lvl1.m)**

La función ***langley\_filter\_lvl1.m*** recibe como input los datos con el formato adecuado para el trabajo con los langley, en cualquiera de las formas comentadas anteriormente. Se trata de una única función muy flexible a la que podremos pasar una variedad de inputs opcionales. Los input opcionales se refieren a la depuración que queramos llevar a cabo, que por otra parte se aplica a “medios-días”, AM-PM, definidos a partir del tiempo solar verdadero (calculado a partir de la función ***sun\_pos.m***). Son los siguientes:

* ***airmass:*** rango de masas ópticas de trabajo. Puede ser un solo valor ó una pareja de valores. No filtro ([]) por defecto.
* ***N\_filt:*** número mínimo de medidas por filtro. 5 por defecto (NOT IMPLEMENTED).
* ***N\_hday:*** número mínimo de medidas para cada “medio-día”. 20 sumarios → 20 x 5 medidas individuales, por defecto.
* ***O3\_hday:*** máxima desviación estándar para el ozono promedio de cada “medio-día”. Se pretende con este filtro descartar los “medios-días” en los que el ozono no permanece aproximadamente constante. Por defecto no se usa este filtro.
* ***Date\_range:*** Rango de fechas a analizar. No filtro ([]) por defecto.
* ***F\_corr:*** Corrección a las ratios MS9 debidos a efectos de No-Neutralidad de los filtros de atenuación (ND). Pueden ser de dos formas: variable F\_corr generada por la función ***read\_cal\_config\_new.m***, o bien un vector 6-D (e.g. [0 0 0 11 NaN 0]). Es importante hacer notar lo siguiente: aún en el caso de que queramos trabajar con datos raw, es decir, sin aplicar ningún filtro de los comentados en esta sección, en el caso de que los filtros de atenuación del instrumento en cuestión no sean neutrales será necesario hacer pasar los datos de análisis por esta función, pero sólo con el argumento opciona ‘F\_corr’. La razón está en que se trabaja por defecto con el campo 9 de los sumarios, que aún después de haber aplicado la función ***filter\_corr.m***, se corresponde con las MS9’s no corregidas por filtros. Esto es así también para las medidas individuales.
* ***AOD:*** AOD lvl1.5 obtenido de la página web de AERONET (canal 340 nm). Una vez hemos obtenido el fichero con los datos de AOD, la lectura / procesado de los mismos se realiza internamente por medio de la función ***read\_aeronet\_ampm.m***. No filtro por defecto.
* ***Cloud:*** Fichero de texto con la información relativa a la existencia o no de nubes para todo el periodo analizado y para medios-días. Para obtener este fichero se utiliza la función ***cloud\_screening.m***, basada en el método de Long & Ackerman adaptado a Izaña (*García, R.D., O.E, García, E. Cuevas, V.E. Cachorro, P.M. Romero-Campos, R. Ramos and A.M. de Frutos, Solar radiation measurements compared to simulations at the BSRN Izaña station. Mineral dust radiative forcing and efficiency study, JGR-Atmospheres, Vol 119, 1-16, DOI: 10.1002/2013JD020301, 2014*). Es preciso disponer de los ficheros de la bsrn, en formato izo\_bsrn\_corr\_dddyy.txt, además del fichero con datos de AOD obtenido de AERONET (en este caso se trabaja con el canal 500 nm). No filtro por defecto.
* ***lgl\_days:*** Declarando este argumento a 1 obtendremos una tabla-resumen con las condiciones halladas para todos los días considerados (ver Tabla 5 como ejemplo). Por defecto no se genera dicha tabla (lgl\_days=0).
* ***Plots:*** En caso de definirlo a 1, nos dará un ploteo para cada medio-día mostrando la relación MS9-airmass junto con el ozono, y los residuos del ajuste, junto con algunos parámetros de diagnóstico. Se muestra un ejemplo en la Figure 2.
* ***Summ:*** Argumento opcional necesario para poder trabajar con sumarios. Los valores posibles son 0 (no sumarios, por defecto) ó 1 (sumarios). Este es un input adicional que, aunque opcional, es importante definir explícitamente, o al menos tener presente su significado. Como hemos dicho, la función ***langley\_filter\_lvl1.m*** admite como entrada bien los datos de langley obtenidos a partir de los sumarios bien los datos de langley obtenidos a partir de las medidas individuales. En este último caso existe un filtro adicional (digamos level 0) que no es más que el fitro rutinario aplicado a los sumarios, es decir O3\_std < 2.5, 100<O3<600, hg flag =1, n = 5. En el caso de que estemos trabajando con la variable ***langsumm\_sync\_data*** entonces este filtro inicial no es necesario, ya ha sido implementado, pero en el caso de que trabajemos bien con la variable ***ozone\_lgl*** bien con ***ozone\_lgl\_sum***, entonces debemos aplicarlos. Si se trata de la primera (***ozone\_lgl***), asignamos a ‘summ’ el valor 0, y lo que obtenemos son las medidas individuales pero se han descartado aquellas cuyo sumario asociado no cumpla con las condiciones anteriores. Si se trata de la segunda (***ozone\_lgl\_sum***) entonces asignamos a ‘summ’ el valor 1, y se aplicará el filtro de los sumarios. En el caso de la variable ***langsumm\_sync\_data*** también debemos hacer ‘summ ‘ = 1, sólo que ahora no pasará nada. En definitiva, siempre que trabajamos con sumarios debemos iniciar el valor del argumento opcional ‘summ’ a 1.

Como salida, la función langley\_filter\_lvl1.m devuelve una matriz de datos con igual estructura que la de entrada, pero depurada según los criterios anteriores. Además, nos devuelve en pantalla una tabla con los diferentes parámetros asociados a cada “medios-días”, como se muestra en las páginas siguientes.

General_DayLangley_122.tif

Figure :

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Diaj | AM | O3\_std | N | AOD | AOD\_std | Cld | PM | O3\_std | N | AOD | AOD\_std | Cld |
| 03-Apr-2014 | 93 | NaN | NaN | NaN | 0.0119 | 0.00236 | 0 | 1 | 1.45 | 134 | 0.0185 | 0.00579 | 1 |
| 04-Apr-2014 | 94 | NaN | NaN | NaN | 0.0371 | 0.01408 | 0 | 1 | 1.64 | 133 | 0.0255 | 0.00242 | 1 |
| 05-Apr-2014 | 95 | 1 | 1.26 | 61 | 0.0210 | 0.00188 | 1 | 1 | 1.41 | 131 | 0.0270 | 0.00535 | 1 |
| 06-Apr-2014 | 96 | 1 | 1.40 | 64 | 0.0233 | 0.00210 | 1 | 1 | 1.49 | 130 | 0.0371 | 0.00582 | 1 |
| 23-Apr-2014 | 113 | 1 | 1.30 | 80 | 0.0187 | 0.00118 | 1 | 1 | 1.63 | 102 | 0.0214 | 0.00358 | 1 |
| 28-Apr-2014 | 118 | 1 | 2.29 | 125 | 0.0175 | 0.00466 | 1 | NaN | NaN | NaN | 0.0255 | 0.00696 | 1 |
| 30-Apr-2014 | 120 | 1 | 1.89 | 130 | 0.0307 | 0.00428 | 1 | NaN | NaN | NaN | 0.0492 | 0.01031 | 0 |
| 01-May-2014 | 121 | 1 | 1.38 | 130 | 0.0441 | 0.00666 | 1 | NaN | NaN | NaN | 0.0501 | 0.00883 | 1 |
| 02-May-2014 | 122 | 1 | 1.45 | 135 | 0.0262 | 0.00296 | 1 | 1 | 1.42 | 126 | 0.0286 | 0.00562 | 1 |
| 03-May-2014 | 123 | 1 | 1.47 | 125 | 0.0262 | 0.00288 | 1 | 1 | 1.63 | 132 | 0.0296 | 0.00248 | 1 |
| 05-May-2014 | 125 | NaN | NaN | NaN | 0.0329 | 0.00550 | 1 | 1 | 1.68 | 123 | 0.0344 | 0.00257 | 1 |
| 07-May-2014 | 127 | 1 | 1.14 | 90 | 0.0700 | 0.00999 | 1 | 1 | 1.38 | 140 | 0.0933 | 0.00897 | 1 |

Tabla :

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Diaj | O3\_std | N | AOD | AOD\_std | Cld | O3\_std | N | AOD | AOD\_std | Cld |
| 03-Apr-2014 | 93 | 1.26 | 15 | 0.0119 | 0.00236 | 0 | 1.45 | 134 | 0.0185 | 0.00579 | 1 |
| 04-Apr-2014 | 94 | 1.32 | 68 | 0.0371 | 0.01408 | 0 | 1.64 | 133 | 0.0255 | 0.00242 | 1 |
| 05-Apr-2014 | 95 | 1.26 | 61 | 0.0210 | 0.00188 | 1 | 1.41 | 131 | 0.0270 | 0.00535 | 1 |
| 06-Apr-2014 | 96 | 1.40 | 64 | 0.0233 | 0.00210 | 1 | 1.49 | 130 | 0.0371 | 0.00582 | 1 |
| 07-Apr-2014 | 97 | 1.07 | 38 | 0.0235 | 0.00186 | 0 | 1.39 | 135 | 0.0273 | 0.00220 | 0 |
| 08-Apr-2014 | 98 | 1.34 | 110 | 0.0251 | 0.00624 | 0 | 1.75 | 150 | 0.0352 | 0.01073 | 0 |
| 09-Apr-2014 | 99 | 1.40 | 125 | 0.0306 | 0.01141 | 0 | 1.71 | 130 | 0.0195 | 0.00000 | 0 |
| 11-Apr-2014 | 101 | 1.39 | 143 | 0.8015 | 0.12882 | 0 | NaN | NaN | NaN | NaN | 0 |
| 14-Apr-2014 | 104 | 1.65 | 19 | 0.0226 | 0.01309 | 0 | 1.36 | 93 | 0.0325 | 0.00389 | 0 |
| 15-Apr-2014 | 105 | 1.74 | 135 | 0.0305 | 0.00324 | 0 | 1.97 | 124 | NaN | NaN | 0 |
| 16-Apr-2014 | 106 | 1.34 | 125 | 0.0337 | 0.00523 | 0 | 1.95 | 133 | 0.0302 | 0.00561 | 0 |
| 17-Apr-2014 | 107 | 1.49 | 146 | 0.0284 | 0.00830 | 0 | 1.55 | 118 | 0.0274 | 0.00583 | 0 |
| 18-Apr-2014 | 108 | 1.43 | 120 | 0.0550 | 0.01360 | 0 | 1.48 | 115 | 0.0868 | 0.01402 | 0 |
| 19-Apr-2014 | 109 | NaN | NaN | 0.0735 | 0.04278 | 0 | 2.18 | 35 | NaN | NaN | 0 |
| 20-Apr-2014 | 110 | 1.65 | 81 | NaN | NaN | 0 | 2.02 | 80 | 0.0203 | 0.00568 | 0 |
| 21-Apr-2014 | 111 | 4.26 | 125 | 0.0249 | 0.01333 | 0 | 2.52 | 128 | 0.0150 | 0.00845 | 1 |
| 22-Apr-2014 | 112 | 1.83 | 115 | 0.0229 | 0.00329 | 0 | 2.46 | 127 | 0.0259 | 0.00309 | 0 |
| 23-Apr-2014 | 113 | 1.30 | 80 | 0.0187 | 0.00118 | 1 | 1.63 | 102 | 0.0214 | 0.00358 | 1 |
| 24-Apr-2014 | 114 | 1.37 | 67 | 0.0232 | 0.00170 | 0 | 2.20 | 120 | 0.0294 | 0.00381 | 0 |
| 28-Apr-2014 | 118 | 2.29 | 125 | 0.0175 | 0.00466 | 1 | 3.38 | 70 | 0.0255 | 0.00696 | 1 |
| 29-Apr-2014 | 119 | 1.39 | 125 | 0.0374 | 0.00558 | 0 | 2.23 | 124 | 0.0462 | 0.01539 | 0 |
| 30-Apr-2014 | 120 | 1.89 | 130 | 0.0307 | 0.00428 | 1 | 1.80 | 108 | 0.0492 | 0.01031 | 0 |
| 01-May-2014 | 121 | 1.38 | 130 | 0.0441 | 0.00666 | 1 | 3.33 | 130 | 0.0501 | 0.00883 | 1 |

1. **Obtención de la constante ETC (langley\_analys.m)**

La función ***langley\_analys.m*** realiza una regresión simple de MS9 frente a la masa óptica para obtener la constante ETC para cada *medio-día*. Los argumentos de entrada son, como necesarios, los siguientes:

1. **lgl\_data:** puede ser cualquiera de las variables discutidas hasta ahora (***langsumm\_sync\_data,*** ***ozone\_lgl***, ***ozone\_lgl\_sum, langindv\_sync\_data***

…)

1. **brw:** puesto que los datos de entrada van a ser cell-arrays, uno para cada instrumento, con este parámetro identificamos el instrumento que nos interesa estudiar.
2. **Cal:** esta es una variable-estructura normalmente resultado de la evaluación del fichero de definiciones (por ejemplo ***calizo2013\_setup.m***). En todo caso de aquí sólo se toman las configuraciones 1 y 2.

|  |
| --- |
| General_.tif |
| General_.tif |

Figure 3: Efecto del filtrado adicional ‘res\_filt’. Se eliminan los datos cuyo residuo sea mayor que 1.5 veces la desviación estándar del total de residuos. Datos brutos (no se ha aplicado ningún otro filtro). Método Brewer (arriba) y método Dobson (abajo).

Además de los argumentos obligatorios anteriores, existe uno opcional ***‘res\_filt’.*** Este parámetro controla si se realiza o no un filtro adicional de depuración sobre los datos de la regresión, descartando aquellos que verifican la condición |residuos| > 1.5 x σ(residuos). Es decir, se eliminan aquellos datos cuyo residuo sea mayor que 1.5 veces la desviación estándar del total ([Michalsky](http://www.esrl.noaa.gov/gmd/grad/neubrew/docs/publications/Harrison_ApplOpt1994.pdf)). Una vez hecho esto, se vuelve a calcular la regresión con los nuevos datos depurados.

Como posibles salidas, hasta ahora la función devuelve las variables ***resp*** y **stats** para cada uno de los métodos Brewer y Dobson (cuatro salidas en total)***.*** La información contenida en cada una de ellas sería:

* ***resp***: Matriz de datos 3-D. La dimensión 3 se corresponde con cada una de las configuraciones usadas (2). La dimensión 2 se corresponde con los campos Date, Langley AM y Langley PM. Por último, tendremos una fila (dimensión 1) para cada día analizado (con al menos AM o PM válidos)
* ***stats***: En este caso se trata de una estructura con los campos r, ci y rs.
  + ***stats.r***: Matriz cell-array (un elemento para cada día procesado). Para cada día tenemos una matriz de datos 3-D, similar a la variable ***resp:*** La dimensión 3 se corresponde con cada una de las configuraciones usadas (2). La dimensión 2 se corresponde con los campos Date, airmass, ND filter, residuos AM y residuos PM. Tendremos una fila (dimensión 1) para cada dato analizado.
  + ***stats.ci:*** Intervalo de confidencia al 95% para las constantes ETC determinadas. Se trata de una matriz de datos 3-D. La dimensión 3 se corresponde con cada una de las configuraciones usadas (2). La dimensión 2 se corresponde con los campos Date, lower CI para AM ETC, higher CI para AM ETC, lower CI para PM ETC y higher CI para PM ETC. Por último, tendremos una fila (dimensión 1) para cada día analizado (con al menos AM o PM válidos)
  + ***stats.rs***: R-square statistics. La estructura es como hasta ahora: matriz de datos 3-D. La dimensión 3 se corresponde con cada una de las configuraciones usadas (2). La dimensión 2 se corresponde con los campos Date, rs AM y rs PM. Tendremos una fila (dimensión 1) para cada día analizado (con al menos AM o PM válidos).

***APENDICE:***

%% ---- langley from summaries ----

xlim\_brw=[1500 1700]; xlim\_dbs=[-70 70];

airm\_rang=[1.35 4.30]; cfgs=2;

ozone\_lgl\_sum={}; cfg\_sum={};

for i=[Cal.n\_ref br]

[ozone\_lgl\_sum{i},cfg\_sum{i},lgl\_leg] = langley\_data\_cell\_summ(summary{i},summary\_old{i},config{i});

end

% ---- summaries ----

[summ\_brw\_raw{br} summ\_dbs\_raw{br}] = langley\_analys(ozone\_lgl\_sum,br,Cal);

% ---- summaries dep. ----

ozone\_lgl\_sum\_dep{br} = langley\_filter\_lvl1(ozone\_lgl\_sum{br},'airmass',airm\_rang,'summ',1,'O3\_hday',2.5);

[summ\_brw\_dep{br} summ\_dbs\_dep{br}] = langley\_analys(ozone\_lgl\_sum\_dep,br,Cal,'res\_filt',0);

% ---- summaries dep sync ----

langsumm\_sync\_data = langley\_summ\_sync(ozone\_lgl\_sum,Cal);

langsumm\_sync\_dep{br} = langley\_filter\_lvl1(langsumm\_sync\_data{br},'airmass',airm\_rang,'summ',1,'O3\_hday',2.5);

[summ\_brw\_dep\_sync{br} summ\_dbs\_dep\_sync{br}] = langley\_analys(langsumm\_sync\_dep,br,Cal);

%% ---- langley from Indiv. Measurements ----

for ii=[Cal.n\_ref br]

[ozone\_lgl{ii},cfg\_indv,leg,ozone\_lgl\_sum{ii}] = langley\_data\_cell(ozone\_raw{ii},ozone\_ds{ii},config{ii});

end

% ---- Indiv. Measurements ----

ozone\_lgl\_{br}=langley\_filter\_lvl1(ozone\_lgl{br},'F\_corr',F\_corr{br});

[indv\_brw\_raw{br} indv\_dbs\_raw{br}] = langley\_analys(ozone\_lgl\_,br,Cal,'res\_filt',1,'plot\_flag',0);

% ---- langley from Indiv. Measurements (depured) ----

ozone\_lgl\_dep{br}=langley\_filter\_lvl1(ozone\_lgl{br},'plots',0,...

'F\_corr',F\_corr{br},'airmass',airm\_rang,'O3\_hday',2.5);

[indv\_brw\_dep{br} indv\_dbs\_dep{br}] = langley\_analys(ozone\_lgl\_dep,br,Cal,'res\_filt',1,'plot\_flag',0);

% ---- langley from Indiv. Measurements sync. (depured) ----

langindv\_sync\_data = langley\_indv\_sync(ozone\_lgl,Cal);

ozone\_lgl\_dep{br}=langley\_filter\_lvl1(langindv\_sync\_data{br},'plots',0,'F\_corr',F\_corr{br},...

'airmass',airm\_rang,'O3\_hday',2.5);

[indv\_brw\_dep\_sync{br} indv\_dbs\_dep\_sync{br} stats\_brw{br}] = langley\_analys(ozone\_lgl\_dep,br,Cal,'res\_filt',1,'plot\_flag',0);

Optional

  a) Data selected michalsky

<http://www.esrl.noaa.gov/gmd/grad/neubrew/docs/publications/Harrison_ApplOpt1994.pdf>

    same output but adding  a flag

  b) AOD from aeronet from the same period

  c) Color index

  d) Cloud observations / Cloud camera

 2) Data selection 1

    “Variables”  for selection  level 1

           b) n air mass range  -> default 1.15-3.5

           c) n observations  ( equally distributed better) -> default 25

           d) number of measurements by filter                 -> default 5

           e) Filter correction                                              -> default from config

3)    Regression  (MS9-ETCR)/n  vs 1/n ->  Linregress -> Store all the outputs of linregress  !!. In addition:  airmass range (min, max) , n obs

 output 1;  daily (AM/PM) langleyplot  (1/nu)  am/pm of the langley colors by filter  by flag

                 R6 vs time

Summary output 1:

**Simultaneous ETC correction for all the brewer per day per config**

**ETC  correction instrument vs instrument**

**ETC  correction instrument vs AOD**

4) Recalculation

        ozone mean and std recalculated with ETC corrected

5)  Filter #2  for regression results

                  n obs

                  ozone mean and std during langley

                  R coeff corr

                  Error in slope