Anonymous Referee #1
Received and published: 8 June 2018

The paper is a very useful document describing and characterizing the most important calibration sites for photometers, Izana and Mauna Loa. It is really well written, simple and schematic, and provides all the necessary information for scientists involved in photometer science. Therefore I consider it suitable for the publication.

Below few minor revisions to be done:
1. page 4, lines 17-19: it is stated that "Possible instrument fluctuations due to shipping are controlled by using always a couple of masters that travel together and rigorous comparison of master instruments at the inter-calibration sites". This is a good solution for the stability of the master instruments. However the other equipments shipped with a round trip for calibrating against the master can suffer of the same problem, and come back operative with a calibration no more perfect as the one done in the calibration site. Has this point never been checked or studied ?
For masters (which are well controlled) it is very rare that they change calibration during transport. For field instruments we carefully check the data on site after deployment in the field. Any anomalous behavior like AOD dependence on airmass, negative AOD, anomalous fluctuations in the triplets, etc. would be indicative of some problem during transport or installation. A set of flags is operational in AERONET to detect all kind of (known) instrument malfunctions from the data.

2. page 5 lines 12-15, add that the CAELIS software will be better described in section 4.1
Added.

3. Page 8 line 5: in the text it is stated that the error bar is the year-to-year variability,but in the caption of Figure 3 it is stated that the monthly mean is. Please clarify.
In this figure we show the number of suitable Langley calibrations in a certain month over a multiannual period. For a given month, e.g. January, each year has a different number of Langley days. Therefore we provide the average and the standard deviation that arise from the year to year variability. A clarification has been added in the caption.

4. section 4.2, Figure 4 : It would be interesting to highlight if the yearly variability has a sort of seasonal dependence. In this case it could depend on internal temperature, not well kept constant, or in the parametrization assumed in the Lambert-Beer Law. Which correction ar assumed for gases absorption? it would be interesting they are described briefly in the text.
It's not very likely that residual temperature dependence could result in the slight seasonal dependence of figure 4, because the internal temperature of the PFR is continuously monitored and the optical elements (photodiode and interference filters) are temperature controlled to better than 0.1 K. The expected small seasonal changes in atmospheric gases are not likely to bias the Langley plots either, because Langleys are not affected by the amount of the absorber but would be by a systematic diurnal cycle.
No gases absorption correction is assumed in the Langley analysis to obtain the extraterrestrial signals (only total optical depth is derived). For the aerosol optical depth climatology in section 3 (figs. 1 and 2) we used the standard corrections of GAW-PFR and AERONET networks (see references in the manuscript). AERONET basically uses climatology tables for ozone and NO2, whereas ozone from OMI is used in GAW-PFR. This information has been added to the text.

5. Caption Figure 1: fulfilled
Corrected.

Anonymous Referee #2

Received and published: 12 June 2018

Review of Toledano et al., "Assessment of Sun photometer Langley calibration at the high-elevation sites Mauna Loa and Izana"

General comments
In general, analyzing long and high quality time series of different instruments operated independently in two different networks is valuable work for the atmospheric physics community and appropriate for publication in ACP.
The manuscript describes the environmental variables at Mauna Loa and Izana, compiles climatologies of aerosol optical depth and finally aims at an analysis of the Langley calibration uncertainty for these two sites.
So the manuscript covers a wide scope, however, the scientific impact is weakened by a lack of a rigorous, in-depth analysis. In particular, the statistical (uncertainty) analysis includes several issues. A general indicator for this weakness is that uncertainty and accuracy are often used synonymously, systematic and statistical errors are not treated separately. Without any additional and more detailed discussion, accuracy should be replaced by uncertainty throughout the manuscript.

Specific comments
Section 2.1
P3, L4. "…because the AOD is very low and stable". In fact, I believe, the AOD variability is the actual criterion rather than just low AOD (although typically, both are correlated). This slight misconception appears again later in the manuscript.
Yes, the misconception has been removed. The very low AOD makes it possible that it is very stable in absolute sense, but the crucial point for a Langley plot is the stability. We state this at the very beginning of the introduction (P1, L16).

P4, L3. Just as a question, I wonder why the data in Table 1 are not displayed as e.g. a bar chart? This would probably even save space and convey the information much easier. Then again, I would argue that cloudy periods are mutually exclusive from AOD measurement periods, so cloud statistics do not add any information for the conclusions here, if later, statistics on Langley days are shown anyways. One interesting insight from cloud information could be the probability of suitable Langley conditions in cloud free conditions.
The low cloudiness is an important characteristic of the observatories for solar radiation observations. Previously to the Langley day analysis, we try to show the main features of the station locations for them to be so unique, including this cloud statistics and the aerosol optical depth climatology. The current table contains 4 pieces of information per station, a bar chart would probably be too busy.
The probability of suitable Langley conditions in cloud free conditions, as derived from the database, is about 83% at Mauna Loa and 79% at Izaña. This is seasonal dependent for Izaña, where the probability is about 60% in July and August due to Saharan dust.

Section 2.2
P5, L10. Table 2. Again, just a suggestion, but I believe that visual timelines of the instrument deployment (e.g. in the style of a Gantt chart) would be a lot more efficient than just printing numbers in Table 2. Note that "table" should be capitalized when followed by a number.
"Table" has been capitalized as commented. A figure would be more visual for the timelines but the exact information about number of deployment days would be lost. Moreover there is some visualization in Figure 4, so we would prefer to leave it as is now.

Section 3
P5, L27pp. If a quantitative comparison of time series is the goal here then the same period should be compared (rather than 1994-2016 versus 2000-2014), otherwise the discussion about other causes for differences is problematic. Regarding the cloud screening methods, as far I am aware of, at least part of the cloud screen for the PFR is based on Smirnov's method. What systematic difference can be expected from the differences of the methods, i.e. is Wehrli's method more stringent and therefore filters more data points (possibly with a bias of higher AOD)?
We try to show that long-term AOD climatologies as derived by GAW-PFR and AERONET are equivalent, despite the different periods and cloud-screening. Both networks accumulate enough observations and come to nearly identical climatology, that's the message we intend, rather than a detailed instrumental / quantitative comparison. Such approach has been accomplished by other publications, some of which are cited for reference.
As explained in Wehrli (2008) one step of the three-step method in GAW-PFR is a temporal filter inspired in Smirnov method, although the different sampling method and thresholds make the results to be probably different. To what extent the cloud-screening methods differ would need a dedicated analysis that is out of the scope of this manuscript. Some comparison and scoring can be found in Kazadzis et al. (2018a).
The reviewer is right that a quantitative comparison in the sense of the WMO threshold of 0.005 + 0.010/airmass does not apply for a climatology (no airmass can be easily attributed to a monthly mean). Therefore we have removed the sentence.

P6, L21. In fact, it looks more like a bimodal distribution, rather than log-normal.
The AOD histograms for Mauna Los and Izaña are closer to log-normal than normal distribution, as explained in the paper by O'Neill. It's true that a bi-modal distribution would also produce the geometric and arithmetic means to be separated. To avoid that possible confusion, the sentence has been changed to:
The geometric mean is often more suitable for AOD statistics, because the lognormal probability distribution is a better reference than a normal distribution for most aerosol types (O'Neill et al., 2000).

P7, L3. Why is the AOD in Fig. 2 not shown for 500 nm, as in Fig. 1? Also, what is the reason of using a log scale here and not in Fig. 1? In addition, I would like to suggest that for Fig. 2 histograms would be better suited to reveal the distributions.
Fig. 2 is based on AERONET data and it is shown to highlight in terms of daily means (each point is a day) the occurrence of enhanced background at both locations in spring and episodic dust at Izaña. The plot is nearly the same with 500nm but doesn't need to be directly comparable with Fig. 1. Wavelength channels are highly correlated anyway. Log-scale in Fig. 2 is used to facilitate the visualization of the different daily mean values.
We have changed the plot to 500nm for the sake of consistency. Histograms do not provide much new information, so they have not been included. They are shown here anyway, and could be included in the supplement if necessary.
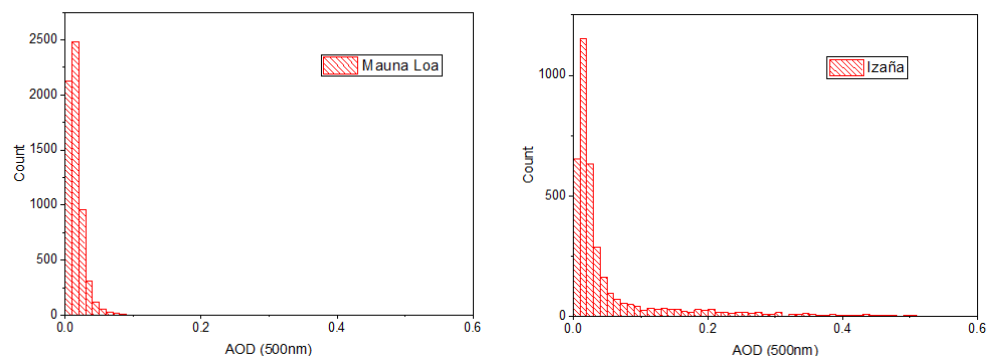
Section 4.1

P7, L28. Surely the criteria also affect the number of suitable Langley plots and hence are relevant in the "climatological sense".

Yes, it's true that they affect but we indicate that it's not a critical (significant) change. The impact is not large because Langley days in most cases are selected no matter what (reasonable) criteria are used. But there are always some cases in the borderline that can be in or out depending on the threshold. Obviously different thresholds would yield to different number of Langley plots. This sentence has been added for clarity: "Other thresholds were tested and revealed no significant changes in the analysis."

Section 4.2

This complete section should be improved by reducing confusing and irrelevant sentences and sharpening the statistical argumentation.

P8, L17. Of course there is no physical measurement without uncertainty. Is the "noise caused by changes in atmospheric transmittance having a hyperbolic (…) dependence" mainly due to residual AOD variations, which affect the slope and/or y-intersect of the Langley plots?

Yes, the main reason for changes in transmittance is the AOD variations. A change of 0.005 in AOD (at any wavelength) is clearly possible and that's the reason to use the high-altitude stations: the less aerosol, the lower the absolute magnitude of variation (for a total aerosol optical depth of 0.01, even 50% relative change would only produce a variability of 0.005). Pressure, water vapor, NO2 or ozone variations can also contribute, but the change needed in these components for a significant modification of the extraterrestrial constant (ETC) would be too large except for pressure in the UV channels (see table below). For instance, a 0.5% modification in ETC (at 500nm) would need a change of 35hPa during the morning or afternoon, 150DU of ozone or 1DU of NO2, all of these unrealistic change rates. Similarly, only a change >20mm during the morning or afternoon in the water vapor column would significantly affect 1020 and 1640nm channels. However a change of 5hPa in pressure during the morning or afternoon would be noticeable for ETC in 340 and 380nm.

	Needed change to produce V0 change of 0.5%					
Channel	Pressure	Ozone	Water vapor	NO2	CO2	CH4
(nm)	(hPa)	(DU)	(cm)	(DU)	(ppm)	(ppb)
340	7	89		0.5		
380	11	1235		0.3		
440	21	1556		0.4		
500	35	158		1.1		
675	120	122		24.9		
870	334					
1020	635		2.1			
1640	4387		3.8		486	1011

Moreover, Marenco (2007) demonstrated that with Langley plots alone it is impossible to identify atmospheric variations having a diurnal periodicity with extreme at noon. This has been added to the text. Some information about the presence of systematic errors has been added to section 4.3 (see corresponding answer below).

P8, L30. This a confusing paragraph. The sentence "Should the instrument degradation…" can be safely omitted. What is the significance of the sentence "…instrumental issues can be discarded…"? In fact, the linear trend is small (but detectable) and has been correctly taken into account.
We have deleted the sentence as suggested.
About the "instrumental issues can be discarded": we wanted to emphasize that the uncertainty in extraterrestrial signals is mainly a result of atmospheric changes, not due to instrumental issues. However, as the reviewer indicates, this possibility has been already explained and the instrumental drift correctly taken into account, therefore we have removed the sentence too.

P9, L3, Fig. 5. For the y-label, change "density" to "N" and also a heading would help like in Fig. 2, indicating site and instrument. Also, why has the analysis been done for Mauna Loa for 14 years and 3 instruments, while for Izana, only for 4 years and 1 instrument?
The changes to the plots have been done.
This kind of analysis is facilitated by long deployment periods, because we try to evaluate drift in instrument extraterrestrial signals. The three long deployments of GAW-PFR data in Mauna Loa (see Fig. 4) are optimal. At Izaña there were multiple changes in GAW-PFR instruments, so we chose the longest deployment available, in this case AERONET #244.

P9, L14. The concept of "adding statistical uncertainty" is statistically confusing and the representation in Fig. 6 is suboptimal in many ways.
First, to avoid this confusion, I believe, simply the absolute uncertainty should be considered and plotted here. Also, how does the uncertainty of a one day Langley plot (as shown in Fig. 5) increase to "1% in total". Please clarify.
This was a mistaken approach, thanks for pointing it out. We assumed the linear trend to have the 0.3-0.5% uncertainty as given by the standard deviation in Fig 5. But the linear trend is almost the "truth" as the error gets divided by Square root N (and for this long term analysis, N is large). We have reformulated the analysis completely.
Fig. 5 actually provides the uncertainty of a single Langley plot calibration. We have now used a 2-sigma criterion (k=2) to provide the uncertainty at the 95% confidence level, i.e. 0.007 or 0.7% for Mauna Loa and 0.9% at Izaña. This is the Type A uncertainty for a single Langley plot. As we combine an increasing number of Langley plots, the standard deviation of the mean gets reduced as in Fig 6, which now shows uncertainty as a function of the number of Langley plots that are averaged, starting with 1 and up to 20. This has been plotted in log-log scale, with the theoretical line of slope= -0.5 plotted as reference. The data fit to slope of -0.40, not far for the expected. We have also included more data in the region between 1 and 10, as indicated.

Second, the statistical uncertainty is generally expected to decrease with square root N, the number measurements, in this case number of days. So the data would ideally be plotted in log-log scale to be able to compare it to a linear slope of -0.5. A deviation from that slope indicates additional error sources (short term drift of the instrument or changes in the signal). Third, the region between 1 and 10 days seems important, so more data points would be beneficial.
(See also answer above). As mentioned, the plot in log-log scale has a slope of -0.40, so it's not far from the expected value. The existence of additional error sources cannot be discarded, as

well as possible correlations. The long-term stability of the instruments is demonstrated but short-term drifts are possible due to small obstructions, residual temperature changes, etc.

P9, L24. Please explain why suitable days get reduced.
We meant that the number of suitable pm langleys is only 134 days per year at Mauna Loa.

P9,L28. Please clarify the "strong requirement" and include the variability of the AOD, rather than just low AOD.
The AOD requirement is AOD(500nm)<0.025. We agree that it is necessary to specify that the threshold in AOD is intended to reduce the possibility of AOD variability, rather than AOD itself. New sentence is: "The strong requirement of AOD(500nm) < 0.025 is needed to prevent AOD variability and achieve the low uncertainty required by AERONET and GAW-PFR."

P9, L35. Again, it is the lower variability of AOD and the wavelength dependence is caused by the Angstrom exponent >0.
The sentence has been changed to: "This wavelength dependence in uncertainty occurs due to lower AOD variability at the longer wavelengths."

Section 4.3
For a "deep assessment" a lot more factors should at least be mentioned. E.g. gas absorption of ozone at 500nm, how is the ozone considered, climatological values?
Or, e.g. what is the effect of different definitions of air mass? As mentioned, it becomes important for large air masses.
We have included a new paragraph in this section, providing the main factors that affect Langley plots, with a set of references that extensively describe the effect of finite bandwidth, contribution of diffuse light, airmass and systematic semidiurnal variations of aerosol. The quantification made e.g. by Reagan (1986) indicates that these errors should not exceed 0.1% with the specified instruments (field of view 1.2deg., GPS time, etc.) and Langley conditions at the high altitude stations.
About other components, there is literature about the semidiurnal pressure variation ("atmospheric tide", Le Blancq, 2011, about 1-2 hPa amplitude). The current processing accounts for pressure changes and they would mean too little change in Rayleigh optical depth anyway. The diurnal changes in water vapor (in the order of few mm amplitude), would not significantly affect the aerosol channels. Diurnal cycles in other components such as NO2, CO2 or ozone, linked to incoming radiation, vegetation activity, etc. are also of small amplitude and therefore are not expected to produce significant bias in Langley plot calibrations.

P10, L11, Fig. 7. 401 nm is not relevant here, so it should be omitted for clarity of the figure.
We have removed it for clarity, although we lose information on the changes in the blue channel. Most of the discussion is focused on the 500nm channel, but both PFR and Cimel have UV channels.

Considering the standard deviation for MLO in Fig. 5 of 0.3% it is not surprising that variations at the 0.4 % level are not significant and that there are "no correlations". Plotting error bars or bands for the Cimel AOD in Fig. 7, may visually reduce the expectation to detect correlations. Also, why not use a 2-day moving average? Maybe the dip around the 31.10. would actually be significant in the 500 nm Cimel calibrations.
The error bars have been added to the Cimel data at 500nm. The moving average does not improve much because the variations in the Cimel calibrations around 5.10 and 18.11 are not at all shown by Virgo data, even if the dip around 31.10 is a bit more clear (only for 500nm, not for 870nm).

P10, L17. "…averaging over several weeks". From Fig. 6, it looks like averaging more than 10 days does not significantly reduce the uncertainty.

The sentence was not very clear. We meant that, in order to accumulate 10-15 good Langleys, it's possible that 2-3 weeks of measurements are needed. The sentence has been rewritten.

P10, L18. Could the authors please explain the physical reason why turbulence at 12 km altitude and variations in the refractive index should have an effect on the AOD? Surely this would affect the imaging of stars, but does the blurring effect cause direct solar radiation to be scattered outside the FOV of the sun photometers?

The aerosol optical depth depends almost linearly with the scattering coefficient, which in turn depends on the refractive index and size distribution. For fine particles, the influence of the refractive index on the AOD can be large (see for instance the simulations in Wang & Rood, JGR2008).

P10, L27. Would a Brewer really be better suited for this study? Is the sensitivity of the instrument an issue here, or the stability?

Both sensitivity and stability are pertinent, as well as sampling speed and frequency, and other instrumental factors, corrections, etc. For the sake of clarity, we have removed the sentence about the Brewer instrument.

Section 5
P11, L23. At least in the conclusion, the statistical uncertainty should be clearly specified as 1-sigma standard deviation for a one day Langley plot. From Fig. 5, this was estimated to 0.3% (Mauna Loa) and 0.5% (Izana) . "…a single Langley plot will be typically within 1% of the mean". What exactly does the 1% signify? 95% confidence interval?

Following the changes in the statistical analysis, we have used now the 95% confidence interval, therefore 2-sigma. This has been clarified in the text: "Applying a 2-sigma criterion, the typical calibration uncertainty for a single Langley plot is ~0.7-0.9% (at the 95% confidence level). The necessary averaging of at least 10 Langley-derived extraterrestrial signals reduces the uncertainty to 0.25% at Mauna Loa and 0.4% at Izaña."

Why should the averaging be replaced by the temporal linear fit? Fitting a straight line would be a generalized method, including averaging as a special case with a line with slope zero.

The reference instruments are drifting by 0.1% per year. This is negligible for 2-3 week periods that are needed to accumulate 10-15 Langleys. Therefore averaging is a reasonable approach. If the deployment is very long (as in Figure 5), the linear fit is more adequate.

P11, L32. The discussion about the subtropical jet was not really conclusive.

The sentence has been softened to: "Furthermore, more investigations are needed to explore whether the subtropical jet above Izaña is a possible explanation for the increase in the Langley plot residuals in this station during the spring months."

Technical corrections
There are different rules about capitalization, but I think in the context of e.g. "direct sun measurements", "sun photometry" etc., the common practice is to not capitalize "sun".
P1,L11. "…this uncertainty being smaller…"
P3, L15. "…it is the reference observatory…"
P4, L18. "…direct sun signal…"
P6, L2. "Smirnov"
P8, L9. "a certain", or even better "a slight"?
P11, L14. "…we find a climatological average…"
P12, L4. "signal losses"

OK, thanks.

Assessment of Sun photometer Langley calibration at the high-elevation sites Mauna Loa and Izaña

Carlos Toledano[1], Ramiro González[1], David Fuertes[1,2], Emilio Cuevas[3], Thomas F. Eck[4,5], Stelios Kazadzis[6], Natalia Kouremeti[6], Julian Gröbner[6], Philippe Goloub[7], Luc Blarel[7], Roberto Román[1], África Barreto[8,3,1], Alberto Berjón[1], Brent N. Holben[4], and Victoria E. Cachorro[1]

[1]Group of Atmospheric Optics, University of Valladolid (GOA-UVa), Spain
[2]GRASP-SAS, Lille, France
[3]Izaña Atmospheric Research Center, Meteorological State Agency of Spain (AEMET), Tenerife, Spain
[4]NASA Goddard Space Flight Center, Greenbelt, MD, USA
[5]Universities Space Research Association, Columbia, MD, USA
[6]Physikalisch-Meteorologisches Observatorium Davos, World Radiation Center – PMOD/WRC, Davos, Switzerland
[7]Laboratory of Atmospheric Optics, University of Lille, Villeneuve d'Ascq, France
[8]Cimel Electronique, Paris, France

Correspondence to: Carlos Toledano (toledano@goa.uva.es)

Abstract. The aim of this paper is to analyze the suitability of the high-mountain stations Mauna Loa and Izaña for Langley plot calibration of Sun photometers. Thus the aerosol optical depth (AOD) characteristics and seasonality, as well as the cloudiness, have been investigated in order to provide a robust estimation of the calibration ~~accuracy~~uncertainty, as well as the number of days that are suitable for Langley calibrations. The data used for the investigations belong to AERONET and GAW-PFR networks, which maintain reference Sun photometers at these stations with long measurement records: 22 years at Mauna Loa and 15 years at Izaña. In terms of clear sky and stable aerosol conditions, Mauna Loa (3397m a.s.l.) exhibits on average of 377 Langleys (243 morning and 134 afternoon) per year suitable for Langley plot calibration, whereas Izaña (2373m a.s.l.) shows 343 Langleys (187 morning and 155 afternoon) per year. The background AOD(500nm) values, on days that are favorable for Langley calibrations, are in the range 0.01-0.02 throughout the year, with well defined seasonality that exhibits a spring maximum at both stations plus a slight summer increase at Izaña. The statistical analysis of the long-term determination of extraterrestrial signals yields to a calibration uncertainty of ~~~0.2~~0.25-0.5%, ~~being this uncertainty~~ this uncertainty being smaller in the visible and near infrared and larger in the ultraviolet wavelengths. This is due to atmospheric variability ~~that~~ produced by changes in several factors, mainly the aerosol optical depth. The uncertainty cannot be reduced based only on quality criteria of individual Langely plots and averaging over several days is shown to reduce the uncertainty to the needed levels for reference sun photometers.

1 Introduction

The Langley plot method (Shaw, 1983) is widely used for absolute calibration of Sun photometers. The main requirement for the method to be successful is the atmospheric transmittance stability during the period in which direct Sun observations

are acquired at varying solar elevations. Apart from the original (classic) approach, several variations have been developed (e.g. Herman et al., 1981; Forgan, 1994; Campanelli et al., 2004). These are mostly intended to reduce the uncertainty and calibration error in case of changes in the atmospheric transmittance during the observation period.

Sun photometer networks like the AErosol RObotic NETwork (AERONET, Holben et al., 1998), the Global Atmospheric Watch – Precision Filter Radiometer (GAW-PFR, Wehrli, 2005), Skyradiometer Network (SKYNET, Nakajima et al., 1996), use the Langley plot method to calibrate the direct Sun channels, i.e. obtain extraterrestrial signals (V_0), with the aim of calculating aerosol optical depth (AOD). Although some networks (e.g. SKYNET) perform Langleys 'on site' (Campanelli et al., 2007), networks like AERONET and GAW only use high altitude stations to provide accurate absolute calibration with the Langley plot method in the so-called master instruments. The calibration is later transferred to field instruments by comparison in a calibration platform.

The AERONET network currently has 3 calibration centers: Goddard Space Flight Center (GSFC, in Greenbelt, Maryland), Laboratory of Atmospheric Optics (LOA, in Lille/Carpentras, France) and Group of Atmospheric Optics (GOA, in Valladolid, Spain). The GSFC master instruments are calibrated at the Mauna Loa Observatory, in Hawaii. The LOA and GOA masters are calibrated at Izaña Observatory. The GAW-PFR network is managed by the Physikalisch Meteorologisches Observatorium Davos, World Radiation Center (PMOD/WRC) at Davos (Switzerland). It uses a triad of reference (PFR) instruments at Davos which are considered by the World Meteorological Organization (WMO-GAW) as the reference instrument triad for AOD measurements. It also operates permanent reference instruments at Izaña and at Mauna Loa, that return periodically (every six months) to PMOD/WRC and are compared with the reference triad (Kazadzis et al., 2018b).

Mauna Loa is a reference site for radiometric observations and calibrations. It was very early considered as an ideal place for calibration of Sun photometers using the Langley technique (Shaw, 1979), hence it hosts reference instruments of the main radiometric networks. Many studies have already reported the atmospheric aerosol characteristics at Mauna Loa (Bodhaine et al., 1981, 1992; Dutton et al., 1994; Andrews et al., 2011; Hyslop et al., 2013), to cite some. Numerous studies about aerosol characteristics at Izaña have also been conducted (e.g. Prospero et al., 1995; Rodríguez et al., 2011; García et al., 2016). Izaña is also commonly used for accurate Langley plot calibrations (even in Moon photometry, Barreto et al. (2013, 2016)), although the site performance has not yet been quantitatively evaluated in this sense.

After years of continuous Sun photometer observations at the Mauna Loa and Izaña observatories, long and high quality measurement records are available, and the quantification of the calibration performance can be accomplished with the support of robust datasets. Therefore, the aim of this paper is to analyze the capability of the two high-mountain stations Mauna Loa and Izaña for Langley plot calibration, in terms of aerosol characteristics, seasonality and cloudiness; and provide statistically robust ~~figures for calibration accuracy~~quantification of calibration uncertainty. The data used for the investigations belong to AERONET and GAW-PFR networks, both having reference instruments at these stations with long measurement records. Several factors and physical processes affecting the performance of the Langley plots are analyzed.

2 Sites and instrumentation

2.1 The Mauna Loa and Izaña observatories

The atmospheric stability required for the Langley plot method is more easily achieved in remote, high-elevation locations, especially because the AOD is very low and stable. Several characteristics make Izaña and Mauna Loa Observatories to be unique for this purpose.

The Izaña Observatory (Tenerife, Spain, 28°N, 16°W) is located at the top of a mountain plateau, 2373 m above sea level, about 15 km away from the Teide peak. It is run by the Meteorological State Agency of Spain (AEMET, see http://izana.aemet.es). Izaña is normally above a strong temperature inversion layer and therefore free of local anthropogenic influence. It is a World Meteorological Organization (WMO) Global Atmospheric Watch (GAW) program station as well as WMO-CIMO Testbed for Aerosols and Water Vapour Remote Sensing Instruments (http://testbed.aemet.es). It hosts reference instruments of several radiometric networks (e.g. Regional Brewer Calibration Centre, GAW-PFR, AERONET, PANDORA, etc.). Details of the Izaña facilities and activities are described in Cuevas et al. (2017b).

The Mauna Loa Observatory (Big Island, Hawaii, 19°N, 155°W) is located on the slope of Mauna Loa volcano, 3397 m above sea level. It was created in 1956 and run by the National Oceanic and Atmospheric Administration (NOAA, see https://www.esrl.noaa.gov/gmd/obop/mlo). It is the reference observatory for a wide set of atmospheric composition research programs (greenhouse gases, carbon cycle, aerosols, water vapor, ozone, trace gases, etc.) and has been continuously monitoring and collecting data related to the atmospheric change.

Both observatories are located in the free troposphere. The aerosol content above is very low (see section 3), as well as the water vapor column (PWV, precipitable water vapor) and the molecular (Rayleigh) optical depth, making it easier to ensure stable conditions during a Langley plot calibration. For instance the water vapor column at Izaña ranges from 0.2cm in winter to 0.7cm in summer (monthly averages, AERONET-derived, see table Table S1) whereas in the nearby site 'Santa_Cruz_Tenerife' located at sea level, the PWV ranges from 1.5cm to 2.5cm. The atmosphere is therefore very stable, especially in the mornings. In the afternoon, local convection can rise the boundary layer up to the Observatory level, especially at Mauna Loa. The strong inversion associated to the Trade Wind at Izaña very often prevents from boundary layer to reach the observatory (Carrillo et al., 2015).

Another important feature to assure the success of the Langley calibration, is to reduce as much as possible the time needed to acquire Sun observations at a wide optical air mass range, in order to avoid possible atmospheric changes. The latitude of Mauna Loa and Izaña, close to the tropics, make the air mass to change rapidly from 7 to 2, i.e. solar elevations from 8° to 30° approximately, lasting about 1:35h to 2:15h depending on the season (the duration is few minutes shorter for Mauna Loa). Just for comparison, at 37° latitude, the time in winter to change from air mass 7 to 2 is more than 3h. At higher latitudes, air mass 2 is not reached in winter.

The cloudiness is another main aspect in performing Langleys. Even thin high clouds perturb the Langley calibration dramatically. To evaluate the sky conditions with the same methodology at both locations, a cloud satellite product has been used. In particular, the cloud products (GDP-4.8 version) of the algorithms OCRA and ROCCIN (Loyola R. et al., 2010) from

GOME-2 onboard MetOp-A have been used to evaluate cloud fraction and cloud top height respectively. The cloud top height is a crucial parameter due to the high elevation of the observatories. The monthly mean cloud fraction and number of clear sky days, defined as cloud fraction < 0.1, have been evaluated over the period 2007-2014. If the cloud top height was lower than the site elevation, the cloud fraction was considered 0. The results are shown in Table 1. On average, Mauna Loa exhibits 24

5 clear sky days per month, whereas Izaña has 20. There is some seasonal variability, being the period between May and August the most sunny at both locations. However it is possible that very thin cirrus (optical depth < 0.1) are not detectable in these satellite products. This will be taken into account in the analysis of the Langley regressions (section 4).

Besides the necessary atmospheric conditions, the facility itself including permanent and trained staff, convenient access and easy logistics are also an important point to consider. Actually the capacity of the measurement platforms themselves is a

10 limitation given that many radiometric networks have reference instruments in these two observatories. This limitation together with the relatively expensive shipping to such remote locations, is the main reason for AERONET (and many other networks) to calibrate master instruments with the Langley method at Izaña and Mauna Loa, and then transfer the absolute calibration to field instruments in calibration platforms located in much more accessible facilities at GSFC, Carpentras, Davos, Valladolid, etc. As example, 15 to 20 calibrations of AERONET master instruments are accomplished every year at Izaña. Of course the

15 calibration ~~accuracy~~ uncertainty of the field instruments is therefore less than that of masters, but logistically it is not reasonable to ship several hundred instruments every year to Mauna Loa or Izaña. The AOD calibration ~~accuracy~~ uncertainty needed for field instruments (0.01 to 0.02 absolute error as recommended by Kazadzis (2016)) can be achieved by means of side-by-side inter-calibration (Holben et al., 1998; Eck et al., 1999). Possible instrument fluctuations due to shipping are controlled by using always a couple of masters that travel together and rigorous comparison of master instruments at the inter-calibration sites.

20 Ratio of ~~Sun direct~~ direct sun signals between the two masters must keep below 1% variability.

2.2 Instrumentation and datasets

The AERONET standard instrument is the Cimel-318, that has been extensively described (e.g. Holben et al., 1998). It is an automatic radiometer equipped with a 2-axis robot, that collects both direct Sun and sky radiance observations at selected wavelengths in the range 340 to 1640nm. Three generations of Cimels have been used in AERONET: the first (starting the

25 early 1990's) were analog instruments. After 2002 the digital version (Cimel 318N) came into play, and after 2013 the so-called Triple instruments (Cimel 318T, after Sun-Sky-Moon measurement capability) started to operate. All three types of instruments can still be found nowadays in AERONET.

The Precision Filter Radiometer of the GAW-PFR network is described in detail in Wehrli (2005). It uses four AOD channels at 368, 412, 500 and 862 nm and needs a separate solar tracker. It is designed for long-term stability, therefore the detectors are

30 behind a shutter except for the brief sampling periods and the instrument is stabilized in temperature and hermetically sealed, having internal atmosphere of pressurized dry nitrogen.

Both instruments use interference filters to select the wavelengths, with full width at half maximum of about 2-10nm (filters are narrower in the ultraviolet wavelengths). The PFR uses one optical path and detector per channel, allowing simultaneous (and continuous) observation in the 4 channels. Conversely, the Cimel has a single detector (or 2 in the case of instruments

with 1640nm channel) and the filters are mounted in a rotating filter wheel. The Cimel configuration allows more wavelength channels (up to 10) but they can only be measured sequentially. In automatic operation, the Cimel takes a triplet measurement (3 separate measurements in a 1-minute interval) every 15 minutes (or 3 minutes, in the high frequency sampling mode), although during the 'Langley sequence' –am or pm for air masses larger than 2– the Cimel measures at fixed solar elevations, with higher frequency.

The AERONET observations at Mauna Loa started in 1994. The observation period used in this study spans 20 years (1997-2016). Within this period, 210 deployments of 22 different master photometers were done. This gives an idea of the frequent swap of Cimel instruments, once per month on average. The AERONET measurements at Izaña started in 2003 and had 37 deployments (71 days on average, 16 different instruments) until January 2011, when instrument #244 was set as permanent reference. The GAW-PFR measurements started in 2000 and 2001 at Mauna Loa and Izaña respectively. The list of PFR radiometers deployed at each location is given in ~~table~~ Table 2. The high long-term stability of these radiometers will be shown in section 4.

The database tool 'CÆLIS' (Fuertes et al., 2017, www.caelis.uva.es), developed at the Group of Atmospheric Optics, University of Valladolid (GOA-UVa) since 2008, has been used to facilitate the organization and extraction of data. It consists of a relational database, a web interface and a real-time data processing module. The specific software used in this work will be better described in section 4.1. The 'demonstrat' software tool (Holben et al., 1998) was used to browse the AERONET data and construct the AERONET data sets at the two stations, given the frequent swap out of master instruments (every 3-4 months). Conversely the GAW-PFR data sets are composed by few instruments deployed for very long periods.

The two approaches have been therefore different, being AERONET priority to frequently recalibrate and maintain the master instruments, shipping them to the inter-calibration platforms, whereas GAW-PFR has prioritized the stability in the long-term observations, in order to facilitate the assessment of trends in the aerosol content, well in line with the GAW aims. However in the last years (since 2011) AERONET has a permanent instrument at Izaña, not involved in the rotation of masters between this site and the inter-calibration platforms.

3 Aerosol Climatology

The aerosol characteristics at Mauna Loa and Izaña observatories can be well established thanks to the long records of the AERONET and GAW-PFR networks. The very low aerosol optical depth is a general feature at Mauna Loa throughout the year. At Izaña, very clean days alternate with some desert dust intrusions, especially in spring and summer (Cuevas et al., 2017a). The overall statistics for aerosol optical depth at 500nm wavelength is provided in Figure 1 and Table S1. These are computed by averaging all available daily mean values in the investigated period within a certain month of the year. As indicated above, 20 years of continuous AERONET data are used for Mauna Loa and 13 years for Izaña. Version 2 AERONET data have been used in this analysis. A detailed description of the direct sun algorithm, including gaseous absorption spectral corrections, is provided in the AERONET website (see https://aeronet.gsfc.nasa.gov/new_web/publications.html). As for GAW-PFR data, 15 years are available at Mauna Loa and 14 years at Izaña. A detailed description of the AOD derivation is given by Kazadzis et al. (2018b). Both

are depicted in Figure 1. Although the measurement periods are different, the long-term averages of AERONET and GAW-PFR differ less than 0.01 for all months, with mean absolute difference of 0.0035 for the monthly means. ~~This difference also fulfills the WMO criterion for intercomparison (WMO, 2005), which is set to $0.005 + 0.010/airmass$ (Kazadzis et al., 2018a).~~

The cloud screening methodologies of AERONET and GAW differ, thus contributing to differences in monthly means. AERONET uses the algorithm by Smirnov et al. (2000), based on temporal variance as utilized by AERONET. GAW data are cloud screened following the methodology by Wehrli (2008). Other authors have accomplished extensive comparison of Cimel and PFR observations (Kim et al., 2008; Kazadzis et al., 2014, 2016, 2018a) with excellent results.

Regarding Mauna Loa (Fig. 1), the AOD (500nm) has a mean value of 0.016 (geometric mean 0.013), peaks in March with 0.028 and is minimum in August-September, with 0.011. The AOD (500nm) daily mean only exceeded 0.05 in 0.6% of the days. The monthly standard deviations indicate that the variability within each month is very low too. The largest variability is found from March through May, with monthly standard deviations about 0.015. The Ångström exponent AE(440-870nm), also given in Table S1, shows a mean value of 1.25, that is indicative of dominance by fine mode particles. The AE is slightly lower in May (1.02), indicating somewhat greater proportion of coarse mode particles. The spring peak in aerosol concentration at Mauna Loa is a well documented phenomenon and it is attributed to the advection of Asian dust (e.g. Bodhaine et al., 1981; Perry et al., 1999). The uncertainty in AE is very high at MLO since the uncertainty in AOD (about $0.002 - 0.003$) is quite large in relation to the ~ 0.01 measured AOD. Thus the AE values at MLO should be in general taken with caution.

The low AOD makes it difficult to investigate any other aerosol optical and microphysical properties, in particular those derived from the inversion of sky radiances for the AERONET instruments using the Dubovik inversion code (Dubovik and King, 2000; Dubovik et al., 2006). Such properties, like single scattering albedo or complex refractive index, are generally not quality assured if AOD(440nm) is less than 0.4 (Holben et al., 2006). Given that the AOD stability is the main requirement for Langley calibrations, in-depth investigation of the aerosol properties is not in the scope of this work and will not be considered here.

The mean AOD (500nm) at Izaña Observatory is 0.054 (geometric mean 0.029), with important seasonal variability. The ~~difference between arithmetic and~~ geometric mean is ~~a good indicator of~~ often more suitable for AOD statistics, because the log-normal ~~distribution of the AOD data (O'Neill et al., 2000).~~ probability distribution is a better reference than a normal distribution for most aerosol types (O'Neill et al., 2000). Monthly means range from 0.02 –November through February– up to 0.14 in July and August (geometric means 0.07 on both months, see Fig. 1). The transport of Saharan dust over Izaña in Summer enhances the aerosol content and the variability, as indicated by the large monthly standard deviations up to 0.15 in July. The Ångström exponent, that has a mean value of 0.99, exhibits a clear decrease in the summer months down to 0.54 in August, confirming the predominance of coarse dust particles. Despite this variability, 25th percentile of AOD is < 0.03 in July and August, indicating a relevant portion of pristine days during the summer months.

In order to assess the dust event frequency over Izaña, the presence of dust has been investigated within the 13-year AERONET database. Following similar methodology that proposed by Toledano et al. (2007), dust events were identified by daily mean $AOD(870nm) > 0.05$ and $AE < 0.6$, which approximately correspond to the 75th and 25th percentiles of these magnitudes in the Izaña dataset. This simple approach results in the identification of 58 dust event days per year on

average. The seasonal distribution is not even. On the contrary, dust events are very rare from October to February (1-2 days per month), while July and August, on average, exhibit 16 and 17 dust event days respectively, which cause the higher AOD values observed in these months (Fig. 1). Similar results, even with slightly different methodology, were achieved by ~~Guirado-Fuentes et al. (2017)~~Guirado-Fuentes (2015).

5   The dust occurrence over Izaña in summer may yield to the incorrect conclusion that, during several months each year, the Langley calibrations are not possible in this station. But as it was previously indicated, dust events alternate with very clean (background) conditions. To demonstrate this important feature, all daily means of AOD (440nm) over 2004-2014 have been plotted as a function of the day of the year (Figure 2b). For comparison, Figure 2a displays the same plot for Mauna Loa. As can be seen, most of the daily observations at Izaña (about 75%) correspond to background values. Higher daily means,

10   corresponding to dust events, are evident from June to September. Dust events are less frequent and with lower AOD outside those months. Note that dust transport in winter occurs at much lower altitude than in summer, therefore the aerosol column above the observatory is minor in winter as compared to summer dust events, in which dust can reach 5 km height (Ansmann et al., 2011; Guirado-Fuentes, 2015; Cuevas et al., 2015). Izaña is therefore a privileged location for studying Saharan dust within the Saharan Air Layer.

15   Another feature of the AOD seasonal cycle is the increase of the background AOD (lowest values) from March to May, with maximum background of about day of the year equal to 120, i.e. beginning of May. This is not exactly in phase with the spring AOD peak at Mauna Loa (in April). The background AOD is in May about 0.016 (440nm), whereas the rest of the year it is as low as AOD=0.005. Interestingly, this enhanced background occurs both at Mauna Loa and Izaña (Figure 2), although it is unclear whether these two seasonal maxima have the same origin.

20 ## 4   Assessment of calibration capability

4.1   Langely plot analysis

In order to investigate the station capability for Langley calibration, a software tool has been developed and integrated in CÆLIS~~(Fuertes et al., 2017)~~. It performs two Langley plots for each available day (morning and afternoon, i.e. 'am' and 'pm') and stores the resulting extraterrestrial signal together with a set of regression statistics: correlation coefficient, standard devi-

25   ation of the fit (σ), number of valid points, air mass range, fitting error for slope and intercept, etc. The routine performs the linear fit from airmass 7 to 2[1], and analyzes the standard deviation of the fit. If the residual for some point is larger than 2σ, the point is eliminated and a new iteration starts until all points are within 2σ or the number of remaining points is less than 10. If $\sigma > 0.2$ or there is not enough number of points, the process stops.

    This type of automatic and iterative analysis, allows identifying whether a certain day is suitable for Langley plot calibration

30   according to pre-established quality thresholds. In our study, we have considered that for a certain period (morning or afternoon) within a particular day, the Langley calibration is possible if $\sigma < 0.006$, the number of valid points is $> 33\%$ of the initial

---

[1]This differs from the airmass range used in AERONET for Langley calibrations, i.e. 5 to 2, and 4 to 2 for the two UV channels (380 and 340 nm), thereby avoiding errors in optical airmass determination that increase significantly at larger airmass (Russell et al., 1993).

number of observations (Harrison and Michalsky, 1994) and AOD(500nm)< 0.025. These criteria can be chosen based on experience (Kiedron and Michalsky, 2016), but they are not critical in this study because we do not intend to perform the calibration of any particular instrument. For instance, for calibration purposes a higher threshold in σ should be used for the UV wavelengths. However our purpose here is to analyze the number of suitable Langley plots in a climatological sense. Other thresholds were tested and revealed no qualitative changes in the analysis. Other statistical indicators of the linear regression quality, such as the correlation coefficient, do not have enough sensitivity to be used for this purpose.

It is then straightforward to search the database for Langley periods fulfilling the indicated criteria. The results are given in Figure 3, in which the average number of Langley plots for each month is indicated, as well as the standard deviation resulting from the year-to-year variability. Morning and afternoon Langleys are given separately. It is common practice to use only mornings for Langley calibration, but in principle both periods are possible and therefore will be both considered in our study. Overall, Mauna Loa meets the selected criteria in 377 Langleys per year (243 'am' calibrations and 134 'pm'). This means, on average, about 20 morning Langleys and 11 afternoon Langleys per month. Izaña meets the criteria in 343 Langleys per year (187 'am' calibrations and 155 'pm'), which means 15 morning and 13 afternoon Langleys per month. There is ~~certain~~ a slight seasonality, with less suitable days in spring and fall at Mauna Loa and better conditions from May through September and December-January. At Izaña the dust events reduce the number of suitable days in July-August, and the best time of the year is May-June.

The AOD (500nm) for the selected 'Langley' days, is given in Fig. 1b, in which monthly averages are calculated for comparison with the overall climatology (Fig. 1a). This plot provides the seasonality of the background AOD values, that exhibits a spring maximum at both stations plus a slight summer increase at Izaña.

4.2   Calibration and statistical uncertainty

A major issue pointed out by many authors is that, despite the available Langley plots can be screened with very strict criteria, a certain variability, i.e. uncertainty in the extraterrestrial signals, remains (Kazadzis, 2016). The noise is caused by small changes in atmospheric transmission having a hyperbolic (solar air mass) dependence, thus they do not affect the linearity of the Langley plot but may change the result (Shaw, 1976; Cachorro et al., 2004). ~~That is also the reason not to use the correlation coefficient to discriminate Langley plots . This noise~~ These changes in atmospheric transmission are mainly due to AOD variations, which affect the slope and/or y-intersect of the Langley plots and cannot be identified with Langley plots alone (Marenco, 2007). For instance, a small variation of 0.005 in AOD could produce ~0.5% departure in the extraterrestrial signal. Other components like ozone, NO_2 or water vapor, would need unrealistically large changes during the Langley period to produce significant changes in the plot y-intersect. Only changes in pressure $> 5hPa$ during the Langley period would produce a significant change in the shorter wavelengths ($< 400nm$).

This uncertainty is well illustrated in Figure 4, in which the GAW-PFR data at Mauna Loa have been selected. ~~They~~ These data are very appropriate for this analysis due to the long deployment periods. We can see the daily extraterrestrial signals (500nm) obtained with the Langley plot method, after screening with the above mentioned criteria (section 4.1). Making the criteria even stricter reduces of course the number of available points, but does not reduce the variability much farther. That is

the reason why many authors propose (and it is common practice) averaging a sufficient number of Langley plots to be able to achieve a satisfactory calibration (Slusser et al., 2000; Kazadzis, 2016).

For long deployments, such as the PFR's in Figure 4, the temporal fit to the extraterrestrial signals V_0 resulting from the Langley plots is better than just averaging, because it will take into account slow degradation of the optical elements (filters, detectors), which is quite clear, although small, in the plot. For instance, PFR#27 degraded by 0.4% in 5.6 years ($-0.07\%year^{-1}$). This is a successful example in long-term instrumental stability. ~~Should the instrument degradation be faster, the statistical treatment would need to be adjusted accordingly. This can be produced by changes in filter transmission, etc. However we must highlight that the instruments used for our analysis exhibited minimum degradation, thus instrumental issues can be discarded to distort the statistics presented for the stations.~~

Once the slow temporal trend is taken into account, we can try to quantify the residuals in V_0 determination, as a quantification of the ~~accuracy~~ uncertainty of the Langley calibration at the site. The histogram of the V_0 values from the PFR (500nm wavelength), normalized to the long-term temporal trend, is provided in Figure 5a (morning Langleys only). The average of the V_0 distribution is 1.0 and the standard deviation is $\sigma = 0.0033$. In the plot we have superimposed a Gaussian distribution with the same mean and standard deviation (0.3%). The V_0 distribution does not pass a normality test mainly because the distribution has strong kurtosis (leptokurtic shape), with up to 81% of the data contained in $\pm1\sigma$, indicating that most of the values are very close to the average. ~~The standard deviation (0.3%) is therefore a reasonable (even conservative) estimation of the calibration uncertainty~~ Using expanded uncertainty (k=2), which for a normal distribution corresponds to a coverage interval of approximately 95%, the Type A[2] uncertainty of a single Langley plot at Mauna Loa ~~, that agrees with the uncertainty reported by Holben et al. (1998) for AERONET.~~ is therefore ~0.7%. The same analysis for Izaña was carried out with the data of Cimel #244, that is operated continuously since November 2011. The histogram of the residuals of the linear fit of V_0 is depicted in Figure 5b, with a relative standard deviation of 0.0046 (or 0.5%). The distribution of the residuals at Izana follows a Gaussian distribution (at 95% confidence level). Using the expanded uncertainty (k=2, 95% confidence level), the uncertainty of a single Langley plot at Izaña is therefore ~0.9%. This particular instrument (in the 500nm channel) degraded by 0.35% in 5 years (~~$-0.07\%year^{-1}$~~ $-0.07\%year^{-1}$), thus showing also high stability.

Furthermore, we can evaluate the statistical uncertainty of the V_0 determination as a function of the number of averaged Langley plots, with respect to the linear interpolation described before. For this purpose, we have computed moving averages between ~~5 and 30~~ 2 and 20 days (number of Langley V_0's), and compared them with the reference value obtained from the linear interpolation. The 15-day moving average is also plotted in Figure 4. We basically calculate the residuals between the moving averages and the linear temporal trend, and plot them as a function of the number of Langley plots that are averaged. The result can be interpreted as the ~~additional uncertainty that is added to the calibration when we average a limited~~ reduction in uncertainty as we average an increasing number of Langley-retrieved V_0's ~~, as compared to the temporal linear fit over a long period (>1 year).~~. Figure 6 shows the decrease in ~~this additional uncertainty as~~ the expanded uncertainty as a function of the

---

[2] The Type-A evaluation of standard uncertainty may be based on any valid statistical method for treating data (https://physics.nist.gov/cuu/Uncertainty/index.html). Standard uncertainty refers to the uncertainty for a coverage interval of 68% (±1 standard deviation, k=1), whereas expanded uncertainty, k=2, for a normal distribution corresponds to a coverage interval of approximately 95%.

number of averaged ~~V_0's increases. Note that using only one Langley plot will typically increase the calibration uncertainty by 0.5% (1% in total)even though the linear regression fulfills strict quality criteria~~Langley plots. The starting point is the uncertainty of one single Langley plot as described above. Note that the statistical uncertainty is generally expected to decrease with square root of the number measurements, in this case number of Langley plots. This is indicated in the plot as the red dashed line. In our case, the uncertainty reduction is close to this theoretical value (slope of -0.4 in log–log scale). If we average more than ~~20~~ 10 Langley plots, then ~~we reduce this additional uncertainty to~~ the uncertainty is <~~0.1%~~0.25%, in agreement with the calibration uncertainty at Mauna Loa reported for AERONET (Holben et al., 1998).

We have also tried to quantify the differences that can be found between morning ('am') and afternoon ('pm') Langley plots in terms of ~~accuracy~~uncertainty. The criteria applied to select afternoon Langley plots are exactly the same as above, but the number of suitable data ~~gets reduced to~~ is only 134 days per year at Mauna Loa (a factor 1.8 less). The standard deviation of the V_0's gets also higher for 'pm' Langleys ($\sigma = 0.0045$). At Izaña the decrease of 'pm' successful Langleys is not that large, with 155 days per year (a factor 1.2 less), and the standard deviation of the V_0's increases up to 0.006.

The strong requirement ~~in AOD~~of AOD(500nm)< 0.025 is needed to ~~achieve the high accuracy~~ prevent AOD variability and achieve the low uncertainty required by AERONET and GAW-PFR. A recent work by Barreto et al. (2014) included moderate, but stable throughout the day, AOD up to 0.3 in the Langley plot calibrations, that were used to recover a long-term aerosol optical depth data set at Izaña (spanning 1976-2012) from an astronomical spectrometer. The AOD uncertainty in that case gets increased but it is worth mentioning that, depending on the instrument or the intended application, the set of criteria (for instance in AOD) used to select Langley calibrations can be changed.

Finally it must be noted that the uncertainty estimations have been done for the 500nm wavelength. The standard deviation of the V_0's in a typical $\sim 20 - 30$ Langley series is larger in the UV, at $\sim 0.4 - 0.5\%$, and smaller in the NIR wavelengths (870, 1020, 1640 nm) at $\sim 0.1 - 0.2\%$. This wavelength dependence in uncertainty occurs due to lower AOD variability at the longer wavelengths. For the UV the higher variance might be also due to filter blocking issues and also possibly to temperature effects for AERONET Cimels that have not been accounted for in the UV wavelengths (in addition to higher AOD in the UV range) .

4.3 Additional uncertainty sources

~~In order to make a deep assessment of the calibration accuracyusing the Langley plot method, we have~~ There are several other factors that can be considered in the analysis of Langley plot uncertainty. A number of authors analyzed e.g. the effect of finite bandwidth of the sun photometer channels (Box, 1981; Thomason et al., 1982), structured vertical aerosol and uncertainty in airmass determination (Thomason et al., 1983; Forgan, 1988; Russell et al., 1993), diffuse light contributions to the radiometer signal (Reagan et al., 1986), and systematic diurnal variation of optical depth (Schmid and Wehrli, 1995; Marenco, 2007). Actually these factors produce systematic errors, although the current instrument performance in terms of field of view, signal stability and time accuracy, together with the limitation of airmass range and the very low aerosol content in the selected Langley conditions for Mauna Loa and Izaña, make these errors to be <0.1% (Reagan et al., 1986)[3]. Systematic semidiurnal

---

[3]These are Type B uncertainty estimates, not based on the statistical analysis of series of observations

cycles in other components like pressure (i.e. atmospheric tide, Le Blancq (2011)), ozone, water vapor or NO_2, have very small amplitude and do not affect the aerosol wavelength channels significantly.

We have investigated other possible sources contributing to the uncertainty of the Langley plot method. First, we have analyzed the variability of the solar extraterrestrial irradiance, which is assumed as constant in our previous analysis. The measurements of the space-based photometer run by PMOD/WRC as part of the VIRGO Experiment on the ESA/NASA SOHO Mission (Fröhlich et al., 1995) were used for this purpose. The VIRGO data series comprises more than 20 years of total and spectral (in three bands) solar irradiance. It clearly shows the 11-year cycle in solar irradiance, which is in the order of 0.1%. Given the frequency of recalibration (at least 2-3 times per year) of the GAW-PFR and AERONET reference instruments, this solar cycle should not be an issue for AOD calculations.

However short-term variations in spectral solar irradiance can be as large as 0.5% (at 402 nm) in few weeks during high solar activity, as it is the case of the episode in October-November 2003, depicted in Fig. 7 for the ~~three~~ two Sunphotometer wavelengths (~~402,~~ 500 and 862 nm). We analyzed the extraterrestrial signal provided by the PFR and the Cimel from the ground during this event, unsing the Langley plot method. The resulting (normalized) V_0's, also included in Fig. 7, are however rather noisy and do not correlate with the space-based data. Either the atmospheric variability or the instrument precission prevent the detection of this kind of abrupt changes in solar irradiance even from high altitude stations, at least with these particular instruments. Averaging ~~Langley calibration over several weeks~~ several Langley calibrations is shown again necessary to overcome this possible uncertainty.

Another source of uncertainty that has been analyzed is the presence of the subtropical jet above Izaña in spring, which introduces strong turbulence around 12 km height. This phenomenon is well known by the astronomers of the nearby Canary Astrophysics Institute, since it produces blurring and twinkling of stars due to turbulent mixing in the Earth's atmosphere, that causes variations of the refractive index. To investigate this, we have analyzed the V_0 repeatability as in Fig. 5 but making monthly statistics, in order to check for any seasonality in the quality of the calibrations. The result is shown in Fig. 8. The variability of the Langley plots, as evaluated from the standard deviation of the V_0's (500nm wavelength), is somewhat larger in March and October-November, as compared to the rest of the year. According to Fig. 3 in (Rodríguez-Franco and Cuevas, 2013), March-April are the months with stronger winds in the upper troposphere above Izaña station, but the V_0 variability is not conclusive to confirm or discard the hypothesis. This assessment will need further investigations, ~~for instance using other instruments with very high sensitivity like the Brewer spectrophotometer, which is routinely operated at Izaña and calibrated for AOD with the Langley plot method (Lopez-Solano, 2017). But at~~ but at least we can conclude that noisier Langley plots are to be expected at Izaña in March and fall. At Mauna Loa the standard deviation of the Langley plots is only higher in April, in coincidence with the higher mean AOD in this month.

11

5   Summary and conclusions

The main aerosol optical depth characteristics of the high elevation sites Mauna Loa and Izaña have been analyzed in detail, in order to quantify the characteristics of these locations for Langley plot calibration of Sun photometers. For this purpose, we used long-term records of AERONET and GAW-PFR reference Sun photometers.

The aerosol monthly ~~climatology~~ climatologies derived from both ~~networks agrees~~ types of network instruments agree within 0.0035 optical depth at 500 nm (at both sites), and shows very low aerosol concentrations. For background conditions used in Langley calibrations, AOD (500nm) ranges from 0.01 to 0.02 for both stations. The seasonality is characterized by a spring maximum at Mauna Loa and the occurrence of Saharan dust events in summer at Izaña. Despite the different network operation (frequent swap of AERONET masters, long deployments for GAW-PFR), and data processing schemes (including cloud-screening), they are both shown to be successful in accurate aerosol monitoring in such pristine locations.

The analysis of cloudiness was accomplished by means of the cloud products OCRA and ROCCIN from GOME-2. On average, Mauna Loa and Izaña exhibit 24 and 20 clear sky days per month, respectively (very thin cirrus clouds are not included in these statistics). Therefore the clouds are not an obstacle for acquiring sufficient number of Langley plots. If we look for days fulfilling also the requirement of high atmospheric stability needed for accurate Langley plot calibration, we ~~yield to~~ find a climatological average of 243 morning and 134 afternoon periods per year at Mauna Loa (about 20 morning and 11 afternoon Langleys per month). Izaña meets the criteria in 187 morning and 155 afternoon periods (about 15 morning and 13 afternoon Langleys per month on average). These conditions were established for Langley plots having standard deviation of the residuals $\sigma < 0.006$, number of valid points $> 33\%$ of the initial number of direct Sun observations and AOD(500nm)< 0.025. Dust events at Izaña (especially in summer), reduce the number of available calibration days during those months but do not prevent from having a sufficient number of clean days for Langley calibration (13 morning Langleys in August at the minimum).

Despite the strict criteria used to select individual Langley plots, a certain noise derived from small changes in Atmospheric transmission (mainly changes in AOD), results in the time series of extraterrestrial signals to have a certain variability. This dispersion has been used to statistically provide a quantification of the calibration ~~accuracy, conservatively estimated as uncertainty. The standard deviation of the Langley-derived extraterrestrial signals is~~ ~0.3% at Mauna Loa and ~0.5% at Izaña for 500 nm, regardless of the data set (GAW-PFR or AERONET). ~~The necessary averaging of Langley-derived extraterrestrial signals may be replaced by a temporal linear fit in case of long deployments. With these criteria,~~ Applying a expanded uncertainty (k=2) criterion, the typical calibration uncertainty for a single Langley plot ~~will be typically within 1% of the mean.~~ is 0.7-0.9% (at the 95% confidence level). The necessary averaging of at least 10 Langley-derived extraterrestrial signals reduces the statistical uncertainty to 0.25% at Mauna Loa and and 0.4% at Izaña.

Due to convective activity, morning Langley plots more often fulfill the prescribed stability conditions than afternoons. The probability to have changes in atmospheric transmission is larger in the afternoons and therefore the noise in extraterrestrial signal determination is also larger as compared to the mornings. This effect has been quantified in terms of reduction in the number of available accurate Langley plots: at Mauna Loa, a factor 1.8 less afternoon Langleys; and smaller reduction (a factor 1.2 less) is found at Izaña. It has also been found that fast variations in solar extraterrestrial irradiance, up to 0.5% in few

weeks, are not easily detectable from the ground with this kind of instruments. Furthermore, more investigations are needed to explore whether the subtropical jet above Izaña is ~~pointed out as~~ a possible explanation for the increase in the Langley plot residuals in this station during the spring months.

With this analysis we can conclude that the high-altitude stations Mauna Loa and Izaña meet the GAW-PFR and AERONET
5   network requirements in terms of ~~accuracy~~ uncertainty , i.e. 0.2-0.5% in calibration factors or 0.002-0.005 in AOD (for $airmass = 1$). The long-term operation and maintenance of reference instruments at these unique locations is shown to be key in accurate aerosol monitoring worldwide. The stability of the reference instruments has also been proved to be very high, with signal ~~loses~~ losses due to degradation of optical components below 0.1% per year over long periods.

References

Andrews, E., Ogren, J., Bonasoni, P., Marinoni, A., Cuevas, E., Rodríguez, S., Sun, J., Jaffe, D., Fischer, E., Baltensperger, U., Weingartner, E., Coen, M. C., Sharma, S., Macdonald, A., Leaitch, W., Lin, N.-H., Laj, P., Arsov, T., Kalapov, I., Jefferson, A., and Sheridan, P.: Climatology of aerosol radiative properties in the free troposphere, Atmospheric Research, 102, 365 – 393, doi:10.1016/j.atmosres.2011.08.017, 2011.

Ansmann, A., Petzold, A., Kandler, K., Tegen, I., Wendisch, M., Müller, D., Weinzierl, B., Müller, T., and Heintzenberg, J.: Saharan Mineral Dust Experiments SAMUM–1 and SAMUM–2: What have we learned?, Tellus, 63B, this issue, 2011.

Barreto, A., Cuevas, E., Damiri, B., Guirado, C., Berkoff, T., Berjón, A. J., Hernández, Y., Almansa, F., and Gil, M.: A new method for nocturnal aerosol measurements with a lunar photometer prototype, Atmospheric Measurement Techniques, 6, 585–598, doi:10.5194/amt-6-585-2013, 2013.

Barreto, A., Cuevas, E., Pallé, P., Romero, P. M., Guirado, C., Wehrli, C. J., and Almansa, F.: Recovering longterm aerosol optical depth series (1976–2012) from an astronomical potassium-based resonance scattering spectrometer, Atmos. Meas. Tech., 7, 4103–4116, 2014.

Barreto, A., Cuevas, E., Granados-Muñoz, M.-J., Alados-Arboledas, L., Romero, P. M., Gröbner, J., Kouremeti, N., Almansa, A. F., Stone, T., Toledano, C., Román, R., Sorokin, M., Holben, B., Canini, M., and Yela, M.: The new sun-sky-lunar Cimel CE318-T multiband photometer – a comprehensive performance evaluation, Atmospheric Measurement Techniques, 9, 631–654, doi:10.5194/amt-9-631-2016, 2016.

Bodhaine, B. A., Mendonca, B. G., Harris, J. M., and Miller, J. M.: Seasonal variations in aerosols and atmospheric transmission at Mauna Loa Observatory, Journal of Geophysical Research: Oceans, 86, 7395–7398, doi:10.1029/JC086iC08p07395, 1981.

Bodhaine, B. A., Harris, J. M., Ogren, J. A., and Hofmann, D. J.: Aerosol optical properties at Mauna Loa Observatory: Long-range transport from Kuwait?, Geophysical Research Letters, 19, 581–584, doi:10.1029/92GL00524, 1992.

Box, M. A.: Finite bandwidth and scattered light effects on the radiometric determination of atmospheric turbidity and the solar constant, Appl. Opt., 20, 2215–2219, doi:10.1364/AO.20.002215, http://ao.osa.org/abstract.cfm?URI=ao-20-13-2215, 1981.

Cachorro, V. E., Romero, P. M., Toledano, C., Cuevas, E., and de Frutos, A. M.: The fictitious diurnal cycle of aerosol optical depth: A new approach for "in situ" calibration and correction of AOD data series, Geophysical Research Letters, 31, L12 106, doi:10.1029/2004GLO19651, 2004.

Campanelli, M., Nakajima, T., and Olivieri, B.: Determination of the solar calibration constant for a sun-sky radiometer: proposal of an in-situ procedure, Appl. Opt., 43, 651–659, doi:10.1364/AO.43.000651, 2004.

Campanelli, M., Estellés, V., Tomasi, C., Nakajima, T., Malvestuto, V., and Martinez-Lozano, J. A.: Application of the SKYRAD improved Langley plot method for the in situ calibration of CIMEL sun-sky photometers, Applied Optics, 46, 2688–2702, 2007.

Carrillo, J., Guerra, J. C., Cuevas, E., and Barrancos, J.: Characterization of the Marine Boundary Layer and the Trade-Wind Inversion over the Sub-tropical North Atlantic, Boundary-Layer Meteorology, 158, 311–330, doi:doi:10.?1007/?s10546-015-0081-1, 2015.

Cuevas, E., Camino, C., Benedetti, A., Basart, S., Terradellas, E., Baldasano, J. M., Morcrette, J. J., Marticorena, B., Goloub, P., Mortier, A., Berjón, A., Hernández, Y., Gil-Ojeda, M., and Schulz, M.: The MACC-II 2007–2008 reanalysis: atmospheric dust evaluation and characterization over northern Africa and the Middle East, Atmospheric Chemistry and Physics, 15, 3991–4024, doi:10.5194/acp-15-3991-2015, 2015.

Cuevas, E., Gómez-Peláez, A., Rodríguez, S., Terradellas, E., Basart, S., García, R., García, O., and Alonso-Pérez, S.: The pulsating nature of large-scale Saharan dust transport as a result of interplays between mid-latitude Rossby waves and the North African Dipole Intensity, Atmospheric Environment, 167, 586 – 602, doi:10.1016/j.atmosenv.2017.08.059, 2017a.

Cuevas, E., Milford, C., Bustos, J. J., del Campo-Hernández, R., García, O. E., García, R. D., Gómez-Peláez, A. J., Guirado-Fuentes, C.,
5   Marrero, C., Prats, N., Ramos, R., Redondas, A., Reyes, E., Rodríguez, S., Romero-Campos, P. M., Schneider, M., Belmonte, J., Yela, M., Almansa, F., Barreto, A., López-Solano, C., Basart, S., Terradellas, E., Afonso, S., Bayo, C., Berjón, A., Bethencourt, J., Carreño, V., Castro, N. J., Cruz, A. M., Damas, M., De Ory-Ajamil, F., García, M., Gómez-Trueba, V., González, Y., Hernández, C., Hernández, Y., Hernández-Cruz, B., Jover, M., León-Luís, S. F., López-Fernández, R., López-Solano, J., Rodríguez, E., Rodríguez-Franco, J. J., Rodríguez-Valido, M., Sálamo, C., Sanromá, E., Santana, D., Santo Tomás, F., Sepúlveda, E., Sierra, M., and Sosa, E.: Izaña Atmospheric
10   Research Center Activity Report 2015–2016., Tech. rep., State Meteorological Agency of Spain (AEMET), Madrid, Spain and World Meteorological Organization, Geneva, Switzerland. (Eds.) Cuevas, E., Milford, C. and Tarasova, O. NIPO: 014-17-012-9, WMO/GAW Report No. 236, 2017b.

Dubovik, O. and King, M.: A Flexible Inversion Algorithm for Retrieval of Aerosol Optical Properties from Sun and Sky Radiance Measurements., J. Geophys. Res., 105, 20,673–20,696, 2000.

15   Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W. J., Léon, J.-F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, Journal of Geophysical Research, 111, D11 208, doi:doi:10.1029/2005JD006619, 2006.

Dutton, E. G., Reddy, P., Ryan, S., and DeLuisi, J. J.: Features and effects of aerosol optical depth observed at Mauna Loa, Hawaii: 1982–1992, Journal of Geophysical Research: Atmospheres, 99, 8295–8306, doi:10.1029/93JD03520, 1994.

20   Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: The wavelength dependence of the optical depth of biomass burning, urban and desert dust aerosols., J. Geophys. Res., 104, 31 333–31 350, 1999.

Forgan, B. W.: Bias in a solar constant determination by the Langley method due to structured atmospheric aerosol: comment, Appl. Opt., 27, 2546–2548, doi:10.1364/AO.27.002546, http://ao.osa.org/abstract.cfm?URI=ao-27-12-2546, 1988.

Forgan, B. W.: General method for calibrating Sun photometers, Appl. Opt., 33, 4841–4850, doi:10.1364/AO.33.004841, 1994.

25   Fröhlich, C., Romero, J., Roth, H., Wehrli, C., Andersen, B. N., Appourchaux, T., Domingo, V., Telljohann, U., Berthomieu, B., Delache, P., Provost, J., Toutain, T., Crommelynck, D., Chevalier, A., Fichot, A., Däppen, W., Gough, D. O., Hoeksema, T., Jiménez-Gómez, M., Herreros, J., Roca-Cortés, T., Jones, A. R., Pap, J., and Willson, R. C.: VIRGO: Experiment for Helioseismology and Solar Irradiance Monitoring, Sol. Phys., 162, 101–128, 1995.

Fuertes, D., Toledano, C., Torres, B., R., G., Berjón, A., Cachorro, V., and de Frutos, A.: CAELIS: Software for assimilation, management
30   and processing data of an atmospheric measurement network, Geoscientific Instrumentation, Methods and Data Systems (in preparation), 2017.

García, R. D., García, O. E., Cuevas, E., Cachorro, V. E., Barreto, A., Guirado-Fuentes, C., Kouremeti, N., Bustos, J. J., Romero-Campos, P. M., and de Frutos, A. M.: Aerosol optical depth retrievals at the Izaña Atmospheric Observatory from 1941 to 2013 by using artificial neural networks, Atmospheric Measurement Techniques, 9, 53–62, doi:10.5194/amt-9-53-2016, 2016.

35   Guirado-Fuentes, C.: Caracterización de las propiedades de los aerosoles en columna en la región subtropical, Ph.D. thesis, Universidad de Valladolid, 2015.

Guirado-Fuentes, C., Cuevas, E., Cachorro, V. E., Toledano, C., Alonso-Pérez, S., A. Barreto and, Y. H., Yela, M., and de Frutos, A.: Aerosol characterization in the free troposphere and the marine boundary layer over the subtropical Eastern North Atlantic, (in preparation), 2017.

Harrison, L. and Michalsky, J.: Objective algorithms for the retrieval of optical depths from ground-based measurements, APPLIED OPTICS, 33, 5126–5132, 1994.

Herman, B. M., Box, M. A., Reagan, J. A., and Evans, C. M.: Alternate approach to the analysis of solar photometer data, APPLIED OPTICS, 20, 2925–2928, 1981.

5 Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – a federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Holben, B. N., Eck, T. F., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., and Dubovik, O.: Aeronet's Version 2.0 quality assurance criteria, Proc. SPIE, 6408, 64 080Q8 – 14, doi:10.1117/12.706524, 2006.

10 Hyslop, N., Trzepla, K., Wallis, C., Matzoll, A., and White, W.: Technical note: A 23-year record of twice-weekly aerosol composition measurements at Mauna Loa Observatory, Atmospheric Environment, 80, 259–263, 2013.

Kazadzis, S.: WMO/GAW Aerosol Measurement Procedures, Guidelines and Recommendations. 2nd Edition, Tech. rep., World Meteorological Organization, 2016.

Kazadzis, S., Veselovskii, I., Amiridis, V., Gröbner, J., Suvorina, A., Nyeki, S., Gerasopoulos, E., Kouremeti, N., Taylor, M., Tsekeri, A.,
15 and Wehrli, C.: Aerosol microphysical retrievals from precision filter radiometer direct solar radiation measurements and comparison with AERONET, Atmospheric Measurement Techniques, 7, 2013–2025, doi:10.5194/amt-7-2013-2014, 2014.

Kazadzis, S., Kouremeti, N., Gröbner, J., and Cuevas, E.: Comparisons of Aerosol Optical Depth measurements from different global Networks: the 4th filter radiometer comparison, in: CIMO TECO 2016. Instruments and Observing Methods Report No. 125., 2016.

Kazadzis, S., Kouremeti, N., Diémoz, H., Gröbner, J., Forgan, B. W., Campanelli, M., Estellés, V., Lantz, K., Michalsky, J., Carlund, T.,
20 Cuevas, E., Toledano, C., Becker, R., Nyeki, S., Kosmopoulos, P. G., Tatsiankou, V., Vuilleumier, L., Denn, F. M., Ohkawara, N., Ijima, O., Goloub, P., Raptis, P. I., Milner, M., Behrens, K., Barreto, A., Martucci, G., Hall, E., Wendell, J., Fabbri, B. E., and Wehrli, C.: Results from the Fourth WMO Filter Radiometer Comparison for aerosol optical depth measurements, Atmospheric Chemistry and Physics, 18, 3185–3201, doi:10.5194/acp-18-3185-2018, 2018a.

Kazadzis, S., Kouremeti, N., Nyeki, S., Gröbner, J., and Wehrli, C.: The World Optical Depth Research and Calibration Center (WORCC)
25 quality assurance and quality control of GAW-PFR AOD measurements, Geoscientific Instrumentation, Methods and Data Systems, 7, 39–53, doi:10.5194/gi-7-39-2018, 2018b.

Kiedron, P. W. and Michalsky, J. J.: Non-parametric and least squares Langley plot methods, Atmospheric Measurement Techniques, 9, 215–225, doi:10.5194/amt-9-215-2016, 2016.

Kim, S.-W., Yoon, S.-C., Dutton, E. G., Kim, J., Wehrli, C., and Holben, B. N.: Global Surface-Based Sun Photometer Network for Long-
30 Term Observations of Column Aerosol Optical Properties: Intercomparison of Aerosol Optical Depth, Aerosol Science and Technology, 42, 1–9, doi:10.1080/02786820701699743, 2008.

Le Blancq, F.: Diurnal pressure variation: the atmospheric tide, Weather, 66, 306–307, doi:10.1002/wea.857, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/wea.857, 2011.

Lopez-Solano, J. e. a.: Aerosol optical depth in the European Brewer Network, Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-1003,
35 2017.

Loyola R., D. G., Thomas, W., Spurr, R., and Mayer, B.: Global patterns in daytime cloud properties derived from GOME backscatter UV-VIS measurements, International Journal of Remote Sensing, 31, 4295–4318, 2010.

Marenco, F.: On Langley plots in the presence of a systematic diurnal aerosol cycle centered at noon: A comment on recently proposed methodologies, Journal of Geophysical Research: Atmospheres, 112, D6, doi:10.1029/2006JD007248, 2007.

Nakajima, T., Tonna, G., Rao, R., Boi, P., Kaufman, Y., and Holben, B.: Use of sky brightness measurements from ground for remote sensing of particulate polydispersions, Applied Optics, 35, 15, 2672–2686, 1996.

5 O'Neill, N. T., Ignatov, A., Holben, B. N., and Eck, T. F.: The lognormal distribution as a reference for reporting aerosol optical depth statistics; Empirical tests using multi-year, multi-site AERONET sunphotometer data, J. Geophys. Res., 27, 3333–3336, 2000.

Perry, K. D., Cahill, T. A., Schnell, R. C., and Harris, J. M.: Long-range transport of anthropogenic aerosols to the National Oceanic and Atmospheric Administration baseline station at Mauna Loa Observatory, Hawaii, Journal of Geophysical Research: Atmospheres, 104, 18 521–18 533, doi:10.1029/1998JD100083, 1999.

10 Prospero, J. M., Schmitt, R., Cuevas, E., Savoie, D. L., Graustein, W. C., Turekian, K. K., Volz-Thomas, A., Díaz, A., Oltmans, S. J., and Levy, H.: Temporal variability of summer-time ozone and aerosols in the free troposphere over the eastern North Atlantic, Geophysical Research Letters, 22, 2925–2928, doi:10.1029/95GL02791, 1995.

Reagan, J. A., Thomason, L. W., Herman, B. M., and Palmer, J. M.: Assessment of Atmospheric Limitations on the Determination of the Solar Spectral Constant from Ground-Based Spectroradiometer Measurements, IEEE Transactions on Geoscience and Remote Sensing, 15 GE-24, 258–266, doi:10.1109/TGRS.1986.289645, 1986.

Rodríguez, S., Alastuey, A., Alonso-Pérez, S., Querol, X., Cuevas, E., Abreu-Afonso, J., Viana, M., Pérez, N., Pandolfi, M., and de la Rosa, J.: Transport of desert dust mixed with North African industrial pollutants in the subtropical Saharan Air Layer, Atmospheric Chemistry and Physics, 11, 6663–6685, doi:10.5194/acp-11-6663-2011, 2011.

Rodríguez-Franco, J. J. and Cuevas, E.: Characteristics of the subtropical tropopause region based on long-term highly resolved sonde records 20 over Tenerife, Journal of Geophysical Research, 118, 10 754–10 769, 2013.

Russell, P. B., Livingston, J. M., Pueschel, R. F., Reagan, J. A., Browell, E. V., Toon, G. C., Newman, P. A., Schoeberl, M. R., Lait, L. R., Pfister, L., Gao, Q., and Herman, B. M.: Post-Pinatubo optical depth spectra vs. latitude and vortex structure: Airborne tracking sunphotometer measurements in AASE II, Geophysical Research Letters, 20, 2571–2574, doi:10.1029/93GL03006, 1993.

Schmid, B. and Wehrli, C.: Comparison of sunphotometer calibration by Langley technique and standard lamp, Appl. Opt., 34, 4500–4512, 25 1995.

Shaw, G.: Error analysis of multi-wavelength sun photometry, Pure and Applied Geophysics, 114, 1–14, 1976.

Shaw, G. E.: Aerosols at Mauna Loa: Optical Properties, Journal of the Atmospheric Sciences, 36, 862–869, 1979.

Shaw, G. E.: Sun photometry, Bull. Am. Meteorol. Soc., 64, 4–10, 1983.

Slusser, J., Gibson, J., Bigelow, D., Kolinski, D., Disterhoft, P., Lantz, K., and Beaubien, A.: Langley method of calibrating UV filter 30 radiometers, Journal of Geophysical Research: Atmospheres, 105, 4841–4849, doi:10.1029/1999JD900451, 2000.

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-Screening and Quality Control Algorithms for the AERONET Database, Remote Sens. Environ., 73, 337–349, 2000.

Thomason, L. W., Herman, B. M., Schotland, R. M., and Reagan, J. A.: Extraterrestrial solar flux measurement limitations due to a Beer's law assumption and uncertainty in local time, Appl. Opt., 21, 1191–1195, doi:10.1364/AO.21.001191, 1982.

35 Thomason, L. W., Herman, B. M., and Reagan, J. A.: The Effect of Atmospheric Attenuators with Structured Vertical Distributions on Air Mass Determinations and Langley Plot Analyses, Journal of the Atmospheric Sciences, 40, 1851–1854, 1983.

Toledano, C., Cachorro, V. E., de Frutos, A. M., Sorribas, M., Prats, N., and de la Morena, B. A.: Inventory of African desert dust events over the southwestern Iberian Peninsula in 2000-2005 with an AERONET Cimel Sun photometer, Journal of Geophysical Research, 112, D21 201, doi:Doi:10.1029/2006JD008307, 2007.

Wehrli, C.: GAW–PFR: A network of Aerosol Optical Depth observations with Precision Filter Radiometers. In: WMO/GAW Experts workshop on a global surface based network for long term observations of column aerosol optical properties, Tech. rep., GAW Report No. 162, WMO TD No. 1287, http://library.wmo.int/pmb_ged/wmo-td_1287.pdf, 2005.

Wehrli, C. J.: Remote Sensing of Aerosol Optical Depth in a global surface network, Ph.D. thesis, ETH ZURICH, 2008.

WMO: WMO/GAW Experts Workshop on a Global Surface-based Network for Long Term Observations of Column Aerosol Optical Properties, GAW Report 162, WMO/TD-No 1287, Davos 2004, edited by: Baltensperger, U., Barries, L., and Wehrli, C., available at: http://library.wmo.int/pmb_ged/wmo-td_1287.pdf, 2005.
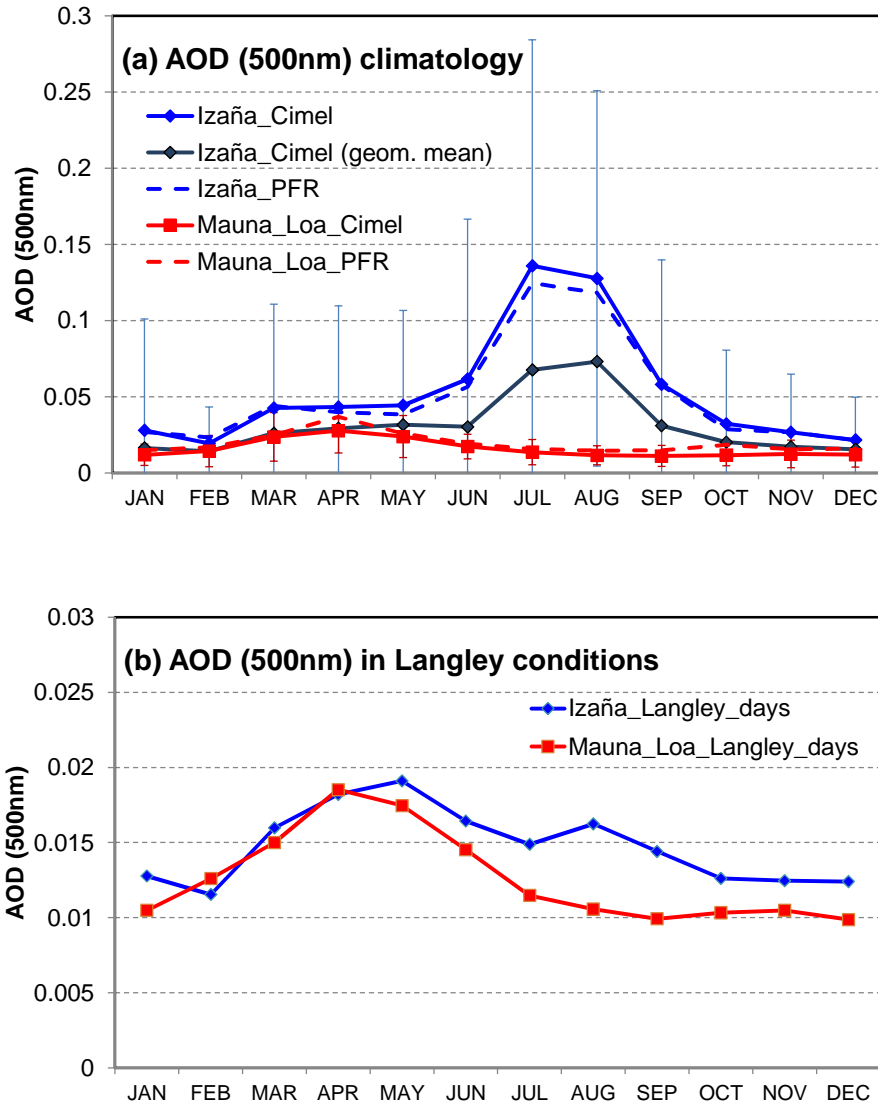
Figure 1. (a) Monthly mean aerosol optical depth (500nm) at Mauna Loa (1994-2016) and Izaña (2004-2016) for AERONET and GAW-PFR. Bars indicate ±1 monthly standard deviation. Black line indicates geometric mean values for AOD at Izaña (in contrast to the arithmetic mean for the other variables). (b) Monthly mean aerosol optical depth (500nm) for the days ~~fufilling~~ fulfilling the criteria for Langley calibration as given in section 4.1.
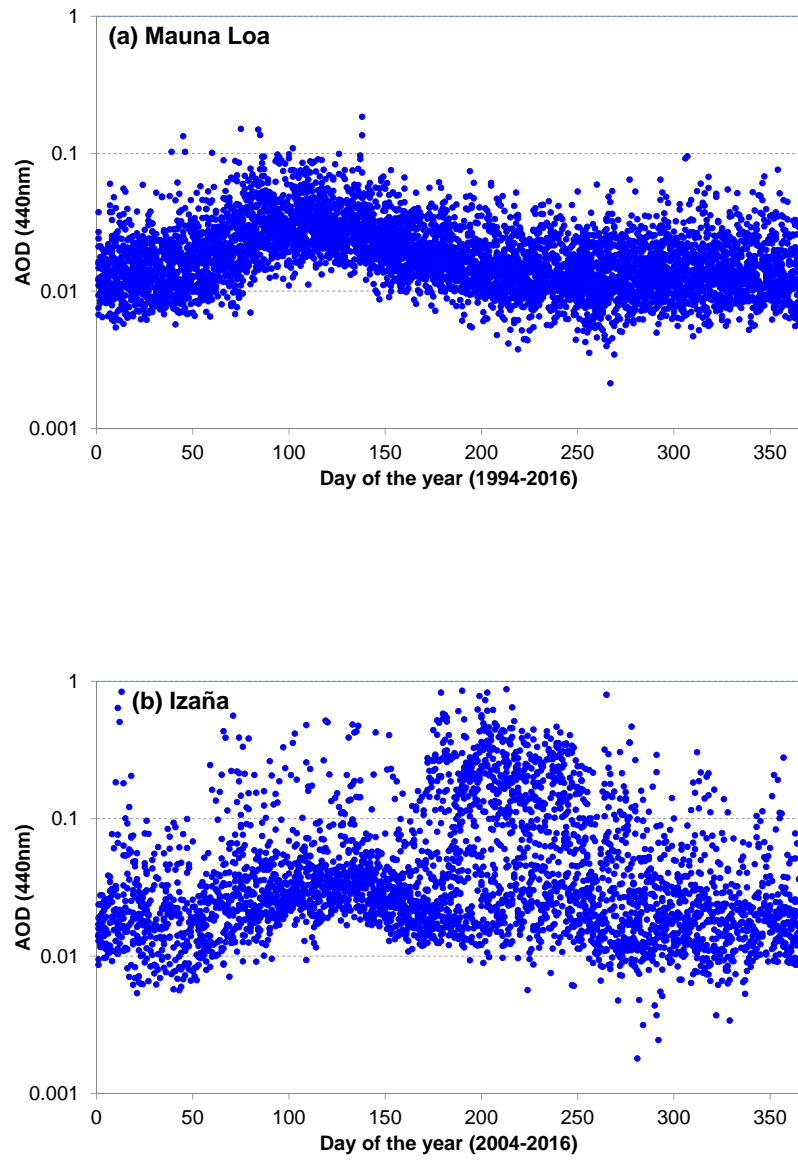
Figure 2. Daily means of aerosol optical depth (440nm) as a function of the day of the year at: (a) Mauna Loa (1994-2016) and (b) Izaña (2004-2016) using AERONET data.
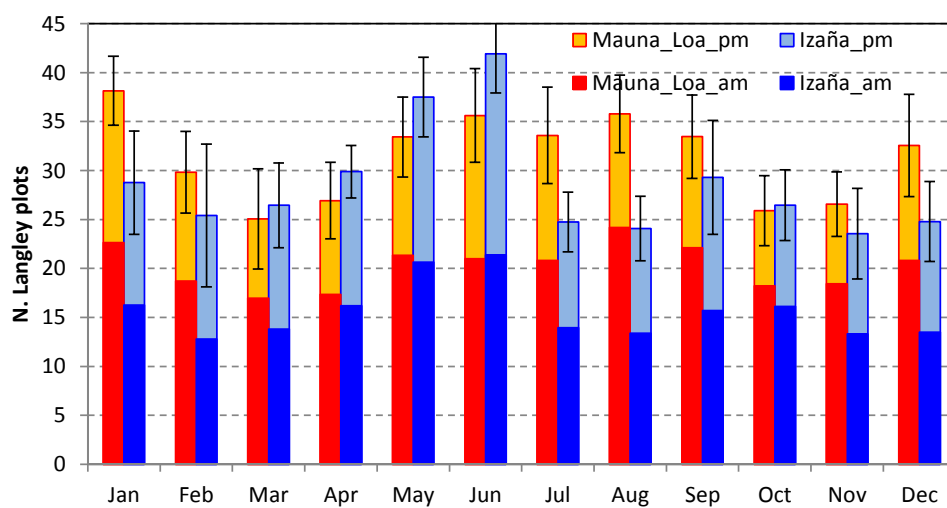
Figure 3. Mean number of suitable Langley calibrations per month at Mauna Loa and Izaña based on GAW-PFR and AERONET data (see text). Bars indicate ±1 standard deviation ~~within the~~ for each month due to year-to-year variability. Morning ('am') and afternoon ('pm') Langley plot calibrations are given separately.
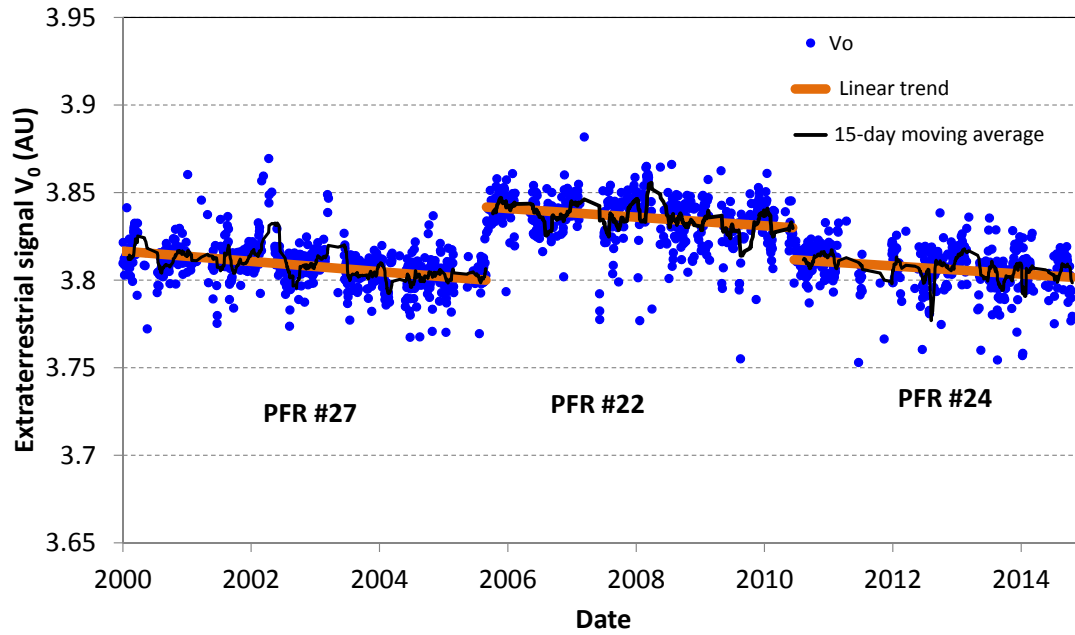
Figure 4. Daily extraterrestrial voltages (V_0) at 500nm wavelength obtained with the Langley plot method for the GAW-PFR at Mauna Loa (morning calibrations only). The temporal linear fit to the V_0's for each instrument deployment is superimposed, as well as the 15-day moving average. Note that these are instrument signals, i.e. depend on each particular instrument and are not directly comparable.
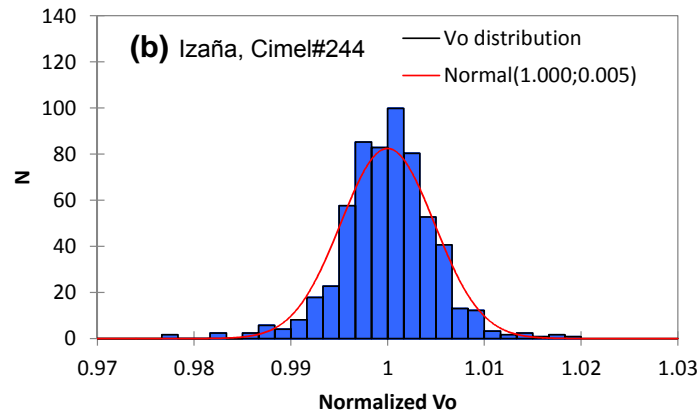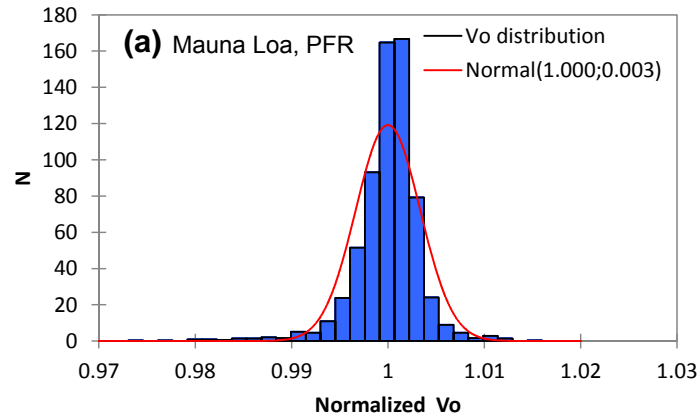
Figure 5. Histogram of daily extraterrestrial voltages (V_0) at 500nm wavelength normalized by the temporal trend: (a) At Mauna Loa using GAW-PFR data (2000-2014); (b) At Izaña using AERONET #244 (2012-2016). Red lines indicate a normal distribution (with the given parameters).
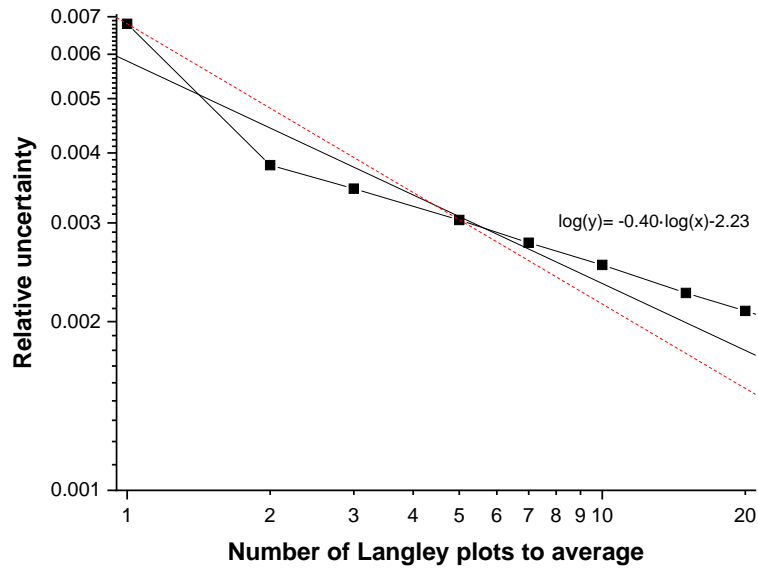
Figure 6. ~~Additional uncertainty added to~~ <u>Uncertainty of</u> the Langley plot calibration vs. number of Langley plot V_0's that are averaged, using GAW-PFR data (500nm) at Mauna Loa (2000-2014). <u>Black line indicates the fit to the experimental data in log–log scale (fit result provided). Red dashed line indicates the theoretical decrease in uncertainty with the square root of the number of measurements.</u>
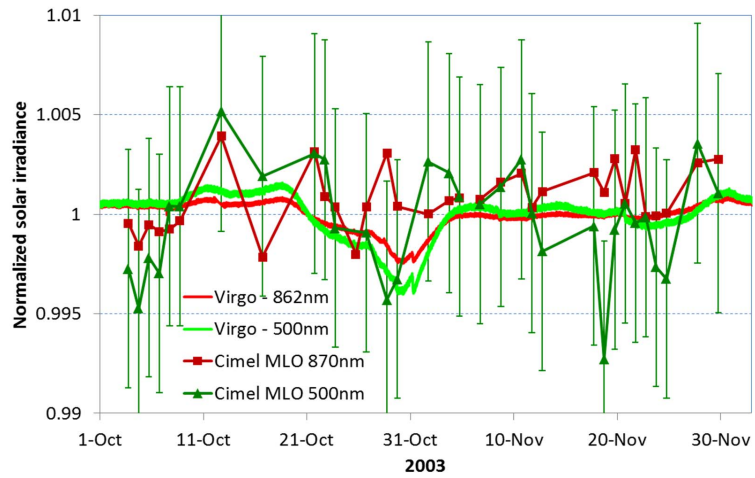
Figure 7. Solar extraterrestrial normalized irradiance as measured by the VIRGO space-based photometer during 2003-2004 at ~~three wavelengths: 402nm (blue),~~ 500nm (green) and 862nm (red) wavelengths. Cimel extraterrestrial signal derived from Langley plots at 500nm and 870nm in Mauna Loa for the same period are also depicted. Error bars for 500nm channel indicate the 0.7% typical uncertainty for this spectral channel.
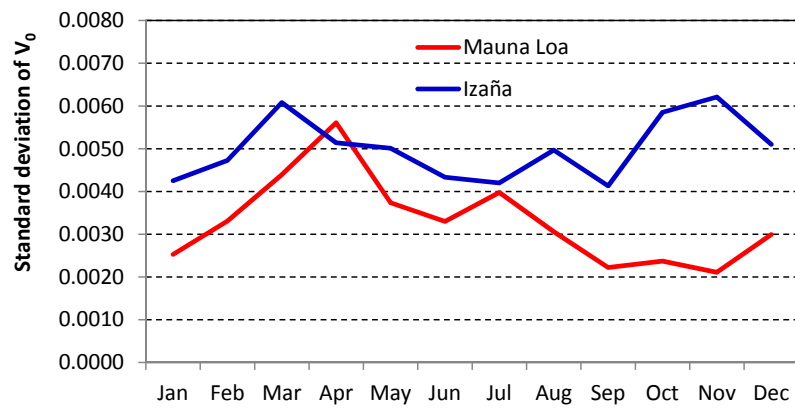
Figure 8. Standard deviation of V_0's (500nm wavelength) from Langley calibrations for every month at Izaña and Mauna Loa using AERONET data.

Table 1. Cloud fraction and number of clear sky days over Mauna Loa and Izaña observatories, derived from GOME-2 cloud products (Loyola R. et al., 2010) over 2007-2014 . Clear sky is defined as cloud fraction < 0.1. The number of investigated days within each month for the 8-year period is also provided.

	Mauna Loa				Izaña			
	Mean Cloud cover fraction	Frequency of cloud cover < 0.1 (%)	Mean N. days fraction< 0.1	N days	Mean Cloud cover fraction	Frequency of cloud cover < 0.1 (%)	Mean N. days fraction< 0.1	N days
Jan	0.06	88.9	28	162	0.13	60.67	19	178
Feb	0.11	75.0	21	164	0.12	67.96	19	181
Mar	0.14	70.8	22	171	0.14	61.22	19	196
Apr	0.11	76.1	23	155	0.11	58.48	18	171
May	0.06	81.9	25	171	0.08	68.85	21	183
Jun	0.05	85.6	26	160	0.05	80.56	24	180
Jul	0.03	86.2	27	159	0.06	76.24	24	181
Aug	0.02	91.2	28	159	0.08	66.47	21	173
Sep	0.07	79.2	24	149	0.14	53.29	16	167
Oct	0.09	76.9	24	156	0.16	58.48	18	171
Nov	0.12	72.9	22	155	0.15	57.74	17	168
Dec	0.19	68.2	21	157	0.16	60.34	19	174
YEAR	0.09	79.4	290	1918	0.11	64.34	235	2123

Table 2. Deployment periods of GAW-PFR instruments at Mauna Loa and Izaña.

(a) Mauna Loa			
Instrument	Start date	End date	N days
PFR #27	1-Jan-2000	1-Sep-2005	2070
PFR #22	2-Sep-2005	16-Jun-2010	1748
PFR #24	16-Jun-2010	31-Dec-2014	1659
(b) Izaña			
Instrument	Start date	End date	N days
PFR #25	9-Jun-2001	25-Feb-2009	2818
PFR #06	14-May-2009	1-Jan-2013	1328
PFR #21	2-Jan-2013	30-Apr-2014	483
PFR #06	1-May-2014	31-Dec-2014	244