Calculation of the 68th percentile of the Absolute Pointing Error

Craig Stephenson

18 March 2014

Contents

1 Introduction

The Absolute Pointing Error (APE) plays a key role in the specification of the pointing requirements for the spacecraft. In the subsequent characterization of the operational performance, the 68th percentile of this quantity (the APE_{68})—upon which the requirements are placed—is estimated by means of calibration observations. For the Herschel mission, a two-stage process was used. First, a set of observations (comprising measurements of the mispointing about two perpendicular axes) was used to estimate the mean and the standard deviation of the underlying population distribution of offsets for each axis.[1] Then, the root-sum-square of these two standard deviations was computed. The resulting quantity, referred to in [13] as the $\mathrm{APE}^{\dagger}$, can thus be considered as providing an approximation to the estimated APE_{68} for the case where the mean offsets are neglected.

It is shown in this note how the approximation inherent in using the $\mathrm{APE}^{\dagger}$ may be avoided and how the estimated quantities (population means and standard deviations) may be used to compute an improved estimate of the APE_{68}, both with and without the mean offsets included. The results given in [13] are reproduced using the new method, thus allowing the effect of both the approximation and of excluding the mean offsets to be assessed.

Although the new method effectively eliminates the errors that were introduced through using the $\mathrm{APE}^{\dagger}$ as a proxy, the resulting values for the APE_{68} remain estimates, based as they are upon four estimated parameters. Therefore, to put the former errors into context, confidence intervals (and standard errors) for the estimated parameters are derived and these, in turn, are used to investigate the uncertainty in the resulting estimate of the APE_{68} for one long period of the mission (for which there were almost 200 calibration measurements).

In the past, results have also been presented for many much smaller data sets, corresponging to relatively short ranges of operational days (ODs). It is shown that much of the variation that has been observed between samples, both in the $\mathrm{APE}^{\dagger}$ and in the mean offsets, is consistent with the uncertainties that one would expect from using small samples. By means of Welch's t-test, it has been found that there are indeed some cases were the mean offsets change significantly between samples. However, as has been noted elsewhere [e.g. 13], such changes are likely to result from the highly attitude-dependent

[1] The population distributions are assumed to be normal.

nature of the star tracker bias errors. They need not reflect a concomitant change in the relative alignment between the instrument and the star tracker. To examine whether there are significant changes in the APE† (or the APE_{68}) between data sets, requires testing the data for homoscedasticity. One such test, Bartlett's test, has been used on two particular data sets to show that they are very likely heteroscedastic (indicating that the APE† had changed) or simply that the population distributions are non-normal (a very real possibility). With the operational mission now over, the APE_{68} is fixed and it could be argued that there is little value in pursuing these matters further. However, with work on-going aimed at a reduction of the ground-based Absolute Measurement Error (AME), it may be worthwhile to return to some of the issues touched on in this document, and perform a thorough set of hypothesis tests, in connection with the computation (estimation) of the AME_{68}.

According to the system requirements specification, the "pointing error specifications... are specified at a temporal probability level of 68%, which implies that [the] error will be less than the requirement for 68% of the time" [6, ch. 4, p. 16]. It is easy to see from Section 3, that the 68th percentiles which have been estimated to date (both in this document and elsewhere) are not in line with this definition. Firstly, the method computes the 68th percentile from a theoretical, infinite population of observations rather than from the actual set of observations. Secondly, the 68th percentile is computed with respect to the number of observations rather than with respect to the time. And finally, the method does not include the effect of spacecraft jitter (which strictly speaking contributes to the APE).

To remedy these problems an alternative method of computing the APE_{68}, based on the reconstructed attitude, is suggested in Section 5. However, it might still be argued that whilst the APE_{68} computed using this alternative method provides (or should provide) a more realistic representation of the accuracy which was actually achieved during the Herschel mission and whilst it may be in closer conformance with the requirements specification, it nevertheless does not provide a very good indicator of any improvements made to the pointing system. This is due to the fact that for much of the mission the APE was highly attitude-dependent, so that the value of the APE_{68} achieved was highly sensitive to the particular sequence of scientific observations chosen and to the time spent in each observation. For characterizing the operational performance of the pointing system the use of the APE_{68} is perhaps not altogether appropriate and a more detailed description

might be preferable.

2 Definition of APE_{68}

The Absolute Pointing Error (APE) is defined as the instantaneous separation between the actual and the commanded payload attitudes [see 5, Section 1_3.1.1]. For the Herschel Observatory, the APE (for each instrument) was decomposed into two components: the angular separation between the actual and the commanded boresight directions and the angular separation about the (actual) boresight direction. Each of these error indices may be modelled as a random process and confidence levels may be used both to impose requirements and to characterize the actual pointing performance.

As discussed in [5, Section 1_4], there are two distinct ways of attaching a meaning to the term 'confidence level'. The statistical distribution of the error index may refer either to: (i) the variation of the error index over an imaginary ensemble of spacecraft; or (ii) its variation over time.[2] Confidence levels may be applied to either (or both) of these distributions. In the System Requirements Specification, a temporal definition was employed and limits were given that each error index was required to not exceed (at 100% probability) for at least 68% of the time.[3] Since we only have access to calibration measurements from a single member of the ensemble of possible Herschel space missions (the actual Herschel mission) it makes sense to also use this temporal definition when characterizing the actual APE. We can nevertheless interpret the values we obtain as relating to the distribution over the ensemble (at any given time) by simply making the assumption of ergodicity. However, this does not affect the calculation, merely the interpretation of the results.

Furthermore, since all the calibration measurements relate to the direction of the boresight of the PACS instrument, it is only possible to derive confidence levels for a single APE index, namely the angular separation between the actual and the commanded PACS boresight direction. We will

[2] This statement is at least true when the error index is a continuous parameter process and the parameter used is the time. However, in our estimation of the APE_{68}, the error index is a discrete parameter process (random sequence) and the parameter used to index the family of random variables is the observation number.

[3] For example, for pointed observations, it was required that for 68% of the time the angles between and around the boresight directions should remain below $3.7''$ and $3.0'$ respectively [6, ch. 4, p. 17].

treat this (strictly speaking) realization of a random process as a random variable, which we shall denote by ΔR, and write its (cumulative) distribution function over the period of interest as $F_{\Delta\mathrm{R}}(\Delta R)$. The 68th percentile of our APE index, APE_{68}, is then (by definition) $\mathrm{APE}_{68} \equiv F^{-1}_{\Delta\mathrm{R}}(0.68)$.[4]

3 Estimation of APE_{68} and APE'_{68}

3.1 Estimation of pointing offset distributions

The raw data used in the calculations described below were obtained during various calibration campaigns performed throughout the mission. For some of the assessments described in [13], such as the comparison of the pointing performance between mission periods,[5] there were a large number, n, of measurements available; for other assessments the number of available measurements was much smaller (in the worst-case, $n = 17$). The data consist of measurements of the mispointing of the PACS boresight; each measurement consisting of a pair, $(\Delta Y_i, \Delta Z_i)$, of angles giving the offsets in the directions of the spacecraft (ACA-frame) y- and z-axes.[6] Since each calibration observation contributes a single pair of offsets to the APE computation and these offsets result from averaging many individual measurements made within the observation, it follows that the computation (estimation) of the APE_{68} does not include the contribution from short-term variations of the spacecraft attitude, such as the high-frequency jitter. Furthermore, it is clear that it is impossible to use these calibration data to calculate a 68th percentile that conforms with the definition given in the requirements specification. That is, since the measurements are indexed with respect to the (calibration) observation number, and not the time, the probability level that is calculated

[4] In much of the post-launch Herschel documentation, the confidence level APE_{68} is referred to simply as the APE.

[5] For the purpose of assessing the pointing performance, the mission has been divided into five distinct periods [see 13]. These periods, numbered 1–5, correspond to the intervals between the following events: Launch (OD 1), reduction of STR baseplate temperature (OD 320); uplink of '1D-correction' (OD 762); uplink of '2D-correction' (OD 866); uplink of 'full correction' (OD 1011); and depletion of liquid helium (OD 1446). (A further measure was taken during period 5 to improve the pointing performance: in OD 1032 the 'tracking' flag was disabled for 73 stars in the on-board catalogue.)

[6] For the discussion presented in this note, the sign convention used for these offsets is irrelevant. Clearly, this decomposition of the mispointing into the two angles, ΔY_i and ΔZ_i, relies on the assumption that the pointing offset is small.

cannot be temporal.

For each data set $\{(\Delta Y_i, \Delta Z_i), i = 1, \ldots, n\}$, it was assumed that the measurements come from distributions that are normal, i.e. $\Delta\mathrm{Y} \sim N(\mu_y, \sigma_y^2)$ and $\Delta\mathrm{Z} \sim N(\mu_z, \sigma_z^2)$ and unbiased estimates of μ_y, μ_z, σ_y^2 and σ_z^2 were calculated according to:

$$m_y = \frac{1}{n}\sum_{i=1}^{n} \Delta Y_i, \quad s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\Delta Y_i - m_y)^2,$$
$$m_z = \frac{1}{n}\sum_{i=1}^{n} \Delta Z_i, \quad s_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\Delta Z_i - m_z)^2.$$

The resulting values of m_y, m_z, s_y and s_z are given in [13] where they are referred to as $\langle\Delta Y\rangle$, $\langle\Delta Z\rangle$, σ_Y and σ_Z.[7]

3.2 Calculation of APE_{68}

Having characterized the distributions ΔY and ΔZ through the estimation of μ_y, μ_z, σ_y^2 and σ_z^2, it remains to use these quantities to obtain an estimate of APE_{68}.[8]

The angles ΔY_i and ΔZ_i are assumed to be small,[9] so that

$$\Delta R_i = \sqrt{(\Delta Y_i)^2 + (\Delta Z_i)^2}.$$

It follows that the distribution function of ΔR is given by

$$F_{\Delta\mathrm{R}}(\Delta R) = \iint\limits_{D_{\Delta R}} p_{\Delta\mathrm{Y},\Delta\mathrm{Z}}(\Delta Y, \Delta Z)\, d(\Delta Y)\, d(\Delta Z), \tag{1}$$

where $p_{\Delta\mathrm{Y},\Delta\mathrm{Z}}$ is the joint (probability) density function of ΔY and ΔZ and $D_{\Delta R} = \{(\Delta Y, \Delta Z) : (\Delta Y)^2 + (\Delta Z)^2 \leq (\Delta R)^2\}$. If we also assume that the random variables ΔY and ΔZ are independent, so that

$$p_{\Delta\mathrm{Y},\Delta\mathrm{Z}}(\Delta Y, \Delta Z) = p_{\Delta\mathrm{Y}}(\Delta Y)\, p_{\Delta\mathrm{Z}}(\Delta Z) = \frac{1}{2\pi\sigma_y\sigma_z} e^{-\frac{1}{2}\left(\frac{(\Delta Y-\mu_y)^2}{\sigma_y^2} + \frac{(\Delta Z-\mu_z)^2}{\sigma_z^2}\right)}, \tag{2}$$

[7] Note that s_y and s_z are biased estimates of σ_y and σ_z. That is, $\mathrm{E}(\mathrm{s}_y) \neq \sigma_y$ and $\mathrm{E}(\mathrm{s}_z) \neq \sigma_z$, where s_y and s_z denote the random variables corresponding to s_y and s_z.

[8] Although the estimates of μ_y, μ_z, σ_y^2 and σ_z^2 are all unbiased, it does not follow that the resulting estimate of APE_{68} will be unbiased.

[9] See footnote 6.

then, upon setting $y = \Delta Y$ and $z = \Delta Z$, (1) becomes

$$\begin{aligned} F_{\Delta \mathrm{R}}(\Delta R) &= \frac{1}{2\pi\sigma_y\sigma_z} \iint\limits_{D_{\Delta R}} e^{-\frac{1}{2}\left(\frac{(y-\mu_y)^2}{\sigma_y^2}+\frac{(z-\mu_z)^2}{\sigma_z^2}\right)}\, dy\, dz \\ &= \frac{1}{2\pi\sigma_y\sigma_z} \int\limits_{y=-\Delta R}^{\Delta R} dy\, e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \int\limits_{z=-\sqrt{(\Delta R)^2-y^2}}^{\sqrt{(\Delta R)^2-y^2}} dz\, e^{-\frac{(z-\mu_z)^2}{2\sigma_z^2}} \\ &= \frac{1}{2\sigma_y\sqrt{2\pi}} \int\limits_{y=-\Delta R}^{\Delta R} dy\, e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}} \left[\mathrm{erf}\left(\frac{\sqrt{(\Delta R)^2-y^2}-\mu_z}{\sigma_z\sqrt{2}}\right)\right. \\ &\qquad \left. - \mathrm{erf}\left(\frac{-\sqrt{(\Delta R)^2-y^2}-\mu_z}{\sigma_z\sqrt{2}}\right)\right], \end{aligned} \tag{3}$$

where erf is the error function. The function $F_{\Delta \mathrm{R}}$ is monotonically strictly increasing and continuous, so it is a straightforward matter to use an iterative method to find $\mathrm{APE}_{68} \equiv F_{\Delta \mathrm{R}}^{-1}(0.68)$ numerically. (In practice we use our estimates m_y, m_z, s_y^2 and s_z^2 to find an estimate for APE_{68}, i.e. $\widehat{\mathrm{APE}}_{68}$.)

3.3 Calculation of APE'_{68}

The mean offsets, μ_y and μ_z, were initially interpreted as indicating the misalignment of the PACS instrument [13, p. 8]. Since a known misalignment could easily be corrected by updating the appropriate Spacecraft Instrument Alignment Matrix (SIAM), it was also of interest to find the reduced value of APE_{68} which would result following such an update. We will denote this new value of APE_{68} by APE'_{68} and define it to be 68th percentile of the random variable

$$\Delta \mathrm{R}' \equiv \sqrt{(\Delta \mathrm{Y} - \mu_y)^2 + (\Delta \mathrm{Z} - \mu_z)^2},$$

with distribution function

$$F_{\Delta \mathrm{R}'}(\Delta R') = \iint\limits_{D_{\Delta R'}} p_{\Delta \mathrm{Y},\Delta \mathrm{Z}}(\Delta Y + \mu_y, \Delta Z + \mu_z)\, d(\Delta Y)\, d(\Delta Z). \tag{4}$$

Since $F_{\Delta \mathrm{R}'}(\Delta R') = F_{\Delta \mathrm{R}}(\Delta R)|_{\mu_y=\mu_z=0}$, we may use (3) to compute $F_{\Delta \mathrm{R}'}(\Delta R')$. However, a simpler expression for $F_{\Delta \mathrm{R}'}(\Delta R')$ is obtained upon substituting

(2) into (4):

$$
\begin{aligned}
F_{\Delta \mathrm{R}'}(\Delta R') &= \frac{1}{2\pi\sigma_y\sigma_z} \iint\limits_{D_{\Delta R'}} e^{-\frac{1}{2}\left(\frac{y^2}{\sigma_y^2}+\frac{z^2}{\sigma_z^2}\right)}\, dy\, dz \\
&= \frac{2}{\pi} \int\limits_{\theta=0}^{\pi/2} d\theta \int\limits_{r=0}^{g(\Delta R',\theta)} dr\, r e^{-\frac{r^2}{2}}, \qquad (y = \sigma_y\, r\, \cos\theta,\ z = \sigma_z\, r\, \sin\theta) \\
&\qquad\qquad \text{where} \quad g(\Delta R', \theta) = \Delta R'(\sigma_y^2 \cos^2\theta + \sigma_z^2 \sin^2\theta)^{-\frac{1}{2}}, \\
&= \frac{2}{\pi} \int\limits_{\theta=0}^{\pi/2} \left[1 - e^{-\frac{g(\Delta R',\theta)^2}{2}}\right] d\theta \\
&= 1 - \frac{2}{\pi} \int\limits_{\theta=0}^{\pi/2} e^{-\frac{(\Delta R')^2}{2(\sigma_y^2\cos^2\theta + \sigma_z^2 \sin^2\theta)}}\, d\theta.
\end{aligned}
\tag{5}
$$

As a check on this result, we note that if $\sigma_y = \sigma_z = \sigma$ say, the right-hand side of (5) may be integrated a second time, giving—as would be expected—the distribution function, $F_{\Delta \mathrm{R}'}(\Delta R') = 1 - e^{-\frac{(\Delta R')^2}{2\sigma^2}}$, for a Rayleigh distribution with scale parameter σ. (In this particular case, $\mathrm{APE}'_{68} = F^{-1}_{\Delta \mathrm{R}'}(0.68) = \sigma\sqrt{-2\ln 0.32} \approx 1.51\sigma$.)

In the general case, where $\sigma_y \neq \sigma_z$, we again note that $F_{\Delta \mathrm{R}'}$ is monotonically strictly increasing and continuous, and so use an iterative method in conjunction with (5) to solve $F_{\Delta \mathrm{R}'}(\mathrm{APE}'_{68}) = 0.68$ numerically. (In practice we use our estimates s_y^2 and s_z^2 to find an estimate for APE'_{68}, i.e. $\widehat{\mathrm{APE}}'_{68}$.)

3.4 Approximation of $\widehat{\mathrm{APE}}'_{68}$ by $\mathrm{APE}^{\dagger}$

In the various documents where the pointing performance has been reported [e.g. 12, 13] equations (3) and (5) have not been used. Instead the $\widehat{\mathrm{APE}}'_{68}$ has been approximated by the root-sum-square of the two estimated standard deviations, a quantity referred to in [13, p. 12] as the $\mathrm{APE}^{\dagger}$, i.e.

$$
\mathrm{APE}^{\dagger} \equiv \sqrt{s_y^2 + s_z^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (\Delta Y_i - m_y)^2 + (\Delta Z_i - m_z)^2}.
$$

From Section 3.3, it is seen immediately that when $s_y^2 = s_z^2$, this approximation introduces an additional error of $(\sqrt{-2\ln 0.32} - \sqrt{2})s_y \approx 0.10\, s_y$. The accuracy of this approximation when $s_y \neq s_z$ is investigated in Section 4.2.

4 Accuracy and interpretation of the results

4.1 Accuracy of APE_{68} and APE'_{68} estimates

The iterative method of computing the APE_{68} and APE'_{68}, by means of (3) and (5) respectively, can in principle be performed to whatever accuracy is desired and so this calculation does not introduce any errors, beyond those made in assuming that the measurement errors come from independent normal distributions, i.e. $\Delta\mathrm{Y} \sim N(\mu_y, \sigma_y^2)$ and $\Delta\mathrm{Z} \sim N(\mu_z, \sigma_z^2)$. To test this assumption properly we would have to return to the original measurement data and to apply a test for normality, such as the Anderson–Darling test, Kuiper's test or the Shapiro–Wilk test [2, 10, 14]. In the absence of such testing, it is only possible to get an indication of the normality of these distributions from an examination of the histograms in [12, 13]. It is further known that roll errors about the ACA x-axis will lead to some correlation between the ΔY and ΔZ distributions.

However, the accuracy of our estimated quantities, $\widehat{\mathrm{APE}}_{68}$ and $\widehat{\mathrm{APE}}'_{68}$, will also depend on the accuracy of the estimates of μ_y, μ_z, σ_y^2 and σ_z^2, i.e. on the accuracy of m_y, m_z, s_y^2 and s_z^2.[10] Based on our assumption that the parent distributions are normal, we have:

$$\begin{aligned} \mathrm{m}_y &\sim N\left(\mu_y, \frac{\sigma_y^2}{n}\right), \quad \mathrm{s}_y^2 \sim \frac{\sigma_y^2}{n-1}\chi^2_{n-1}, \\ \mathrm{m}_z &\sim N\left(\mu_z, \frac{\sigma_z^2}{n}\right), \quad \mathrm{s}_z^2 \sim \frac{\sigma_z^2}{n-1}\chi^2_{n-1}, \end{aligned} \tag{6}$$

where m_y, m_z, s_y^2, s_z^2 are the random variables corresponding to m_y, m_z, s_y^2 and s_z^2. Noting that $\mathrm{var}\{\chi^2_{n-1}\} = 2(n-1)$, it follows that the estimated

[10] The accuracy of $\widehat{\mathrm{APE}}'_{68}$ depends only on the accuracy of s_y^2 and s_z^2.

standard errors in the means and the variances are:

$$\begin{aligned} \widehat{\sigma}_{\mathrm{m}_y} &= \frac{\widehat{\sigma}_y}{\sqrt{n}}, \quad \widehat{\sigma}_{\mathrm{s}_y^2} = \widehat{\sigma_y^2}\sqrt{\frac{2}{n-1}}, \\ \widehat{\sigma}_{\mathrm{m}_z} &= \frac{\widehat{\sigma}_z}{\sqrt{n}}, \quad \widehat{\sigma}_{\mathrm{s}_z^2} = \widehat{\sigma_z^2}\sqrt{\frac{2}{n-1}}, \end{aligned} \tag{7}$$

where $\widehat{\sigma}_y$, $\widehat{\sigma}_z$, $\widehat{\sigma_y^2}$ and $\widehat{\sigma_z^2}$ are our estimates of the population standard deviations and variances. (If required, values for the standard errors in the standard deviations can also be calculated; see Appendix A.) Since the estimated means, m_y, m_z, come from normal distributions, the relation between their estimated standard errors and the confidence intervals for μ_y and μ_z is well-known. For example, the 68% confidence interval for μ_y is

$$m_y - \widehat{\sigma}_{\mathrm{m}_y} < \mu_y < m_y + \widehat{\sigma}_{\mathrm{m}_y}. \tag{8}$$

The estimated variances, s_y^2 and s_z^2, come from scaled χ^2-distributions. Therefore, the calculation of confidence intervals for σ_y^2 and σ_z^2 is more involved. For example, a 68% confidence interval for σ_y^2 (allowing 34% probabilty on either side of the expected value) is:

$$\frac{\widehat{\sigma_y^2}}{n-1} F^{-1}_{\chi^2_{n-1}}(0.16) < \sigma_y^2 < \frac{\widehat{\sigma_y^2}}{n-1} F^{-1}_{\chi^2_{n-1}}(0.84), \tag{9}$$

where $F_{\chi^2_{n-1}}$ is the distribution function for a chi-squared distribution with $n-1$ degrees of freedom. These limits can be found iteratively, using

$$F_{\chi^2_{n-1}}(x) = P\left(\frac{n-1}{2}, \frac{x}{2}\right),$$

where P is the lower (regularized) incomplete gamma function.

Example 1 (period 2)

As an illustration we consider the estimation of the APE_{68} and APE'_{68} for period 2. Using the data in [13, Table 4, top line], we have $n = 196$ and:

$$\begin{aligned} m_y &= -0.40'', \quad s_y = 1.17'', \\ m_z &= -0.13'', \quad s_z = 2.05''. \end{aligned}$$

Ignoring the very slight bias introduced by using s_y and s_z as estimates of σ_y and σ_z, we may take:

$$\widehat{\sigma}_y = s_y = 1.17'', \quad \widehat{\sigma_y^2} = s_y^2 = 1.17^2 = 1.37''^2,$$
$$\widehat{\sigma}_z = s_z = 2.05'', \quad \widehat{\sigma_z^2} = s_z^2 = 2.05^2 = 4.20''^2.$$

From (7) it follows that:

$$\widehat{\sigma}_{\mathrm{m}_y} = 0.08'', \quad \widehat{\sigma}_{\mathrm{s}_y^2} = 0.14''^2,$$
$$\widehat{\sigma}_{\mathrm{m}_z} = 0.15'', \quad \widehat{\sigma}_{\mathrm{s}_z^2} = 0.43''^2,$$

and from (8) and (9), together with the equivalent inequalities for the z-axis, we obtain the 68% confidence intervals:

$$-0.48'' < \mu_y < -0.32'',$$
$$-0.28'' < \mu_z < 0.02'',$$

and

$$1.23''^2 < \sigma_y^2 < 1.51''^2,$$
$$3.78''^2 < \sigma_z^2 < 4.62''^2.$$

(Although the above confidence intervals for σ_y^2 and σ_z^2 were calculated by iteratively determining that $F^{-1}_{\chi^2_{195}}(0.16) = 175.4$ and $F^{-1}_{\chi^2_{195}}(0.84) = 214.6$, it is seen that in both cases, almost identical results would have been obtained if we had simply used the standard errors $\widehat{\sigma}_{\mathrm{s}_y^2}$ and $\widehat{\sigma}_{\mathrm{s}_z^2}$.)

Despite there being no simple relationship between these confidence intervals and those for the APE_{68} and APE'_{68}—the best that could be done would probably be to use Monte Carlo simulations—we can obtain some idea of the likely errors by calculating the APE_{68} and APE'_{68} for those values of μ_y, μ_z, σ_y^2 and σ_z^2 (within the 68% confidence limits) which produce the best (lowest) and worst (highest) values. The results are shown in Table 1. The nominal values of the APE_{68} and APE'_{68} (those calculated using the expected values of μ_y, μ_z, σ_y^2 and σ_z^2) are $2.50''$ and $2.45''$ respectively. It is seen that, for period 2, there is little difference between the APE_{68} and the APE'_{68} and with 196 observations the results are likely to be reasonably accurate.

	$\sigma_y^2 = 1.23$ $\sigma_z^2 = 3.78$	$\sigma_y^2 = 1.51$ $\sigma_z^2 = 4.62$
$\mu_y = 0$ $\mu_z = 0$	2.32	2.57
$\mu_y =$ -0.32 $\mu_z = 0$	2.35	2.60
$\mu_y =$ -0.48 $\mu_z =$ -0.28	2.41	2.65

Table 1: APE_{68} and APE'_{68} for period 2 (arcsec.)

Example 2 (cycles 23–31)

As a second illustration, we consider the estimation of the APE_{68} and APE'_{68} using the 51 measurements from cycles 23–31. According to [13, Table 3], we have:

$$m_y = -0.58'', \quad s_y = 1.33'',$$
$$m_z = -0.30'', \quad s_z = 2.26''.$$

Using the above values of s_y and s_z to derive our best estimates of σ_y, σ_z, σ_y^2 and σ_z^2 (the values of s_y and s_z for the entire period 2 are fairly similar), we obtain the standard errors:

$$\widehat{\sigma}_{\mathrm{m}_y} = 0.19'', \quad \widehat{\sigma}_{\mathrm{s}_y^2} = 0.35''^2,$$
$$\widehat{\sigma}_{\mathrm{m}_z} = 0.32'', \quad \widehat{\sigma}_{\mathrm{s}_z^2} = 1.02''^2.$$

From these we obtain the 68% confidence intervals:

$$-0.77'' < \mu_y < -0.39'',$$
$$-0.62'' < \mu_z < \quad 0.02'',$$

and

$$1.42''^2 < \sigma_y^2 < 2.12''^2,$$
$$4.08''^2 < \sigma_z^2 < 6.13''^2.$$

The estimated value for the APE'_{68} using the data from these cycles is $2.73''$ (see Table 2). However, it is clear by simply comparing the above confidence

intervals for σ_y^2 and σ_z^2 with those in example 1, that there is no reason to believe that the APE'_{68} was in fact any higher during cycles 23–31 than during the rest of period 2. Similarly, there is no reason to suspect that the mean offsets were any different.

If the same calculation is performed for cycles 15–22, a period for which a low value of $\mathrm{APE}^{\dagger}$ was recorded (i.e. $1.99''$), it is found that:

$$0.57''^2 < \sigma_y^2 < 0.87''^2,$$
$$2.55''^2 < \sigma_z^2 < 3.93''^2.$$

indicating that σ_y^2 may have been lower during these cycles. To investigate this further, we apply Bartlett's test to the samples from cycles 15–22 and cycles 23–31.[11] For the y-axis, we obtain a p-value of 0.003, indicating that there is very strong evidence that either σ_y^2 has changed between these two periods or that the population distributions are non-normal. (In comparison, the p-value for the z-axis is 0.12, so that there is little reason to believe that σ_z^2 has changed.) Other tests for homoscedasticity exist which do not rely on the population distributions being normal, but in general these require a knowledge of the individual measurements within each sample.

4.2 Accuracy of $\mathrm{APE}^{\dagger}$

In practice, the $\mathrm{APE}^{\dagger}$ (see Section 3.4) has been used as a proxy for the $\widehat{\mathrm{APE}}'_{68}$. To investigate the accuracy of this 'approximation', the values of these two quantities have been computed for each of the data sets (i.e. values of m_y, m_z, s_y and s_z) given in [13].[12] The results are given, along with the corresponding values of the $\widehat{\mathrm{APE}}_{68}$ in Table 2.[13] It can be seen that the error introduced through the use of the $\mathrm{APE}^{\dagger}$ is typically $0.1''$, which is less than the uncertainty in the $\widehat{\mathrm{APE}}'_{68}$, even for a case where there are many observations such as that considered in example 1 above.

[11] For the comparison of two samples, of size n_1 and n_2 and variance s_1^2 and s_2^2, Bartlett's test statistic is:

$$X^2 = \frac{(N-2)\ln(s_p^2) - (n_1-1)\ln s_1^2 - (n_2-1)\ln s_2^2}{1 + \left(\frac{1}{n_1-1} + \frac{1}{n_2-1} - \frac{1}{N-2}\right)/3},$$

where $s_p^2 = \frac{1}{N-2}[(n_1-1)s_1^2 + (n_2-1)s_2^2]$ and $N = n_1 + n_2$. The associated p-value is given

Table no. in [13].	$m_y(''')$	$m_z(''')$	$s_y(''')$	$s_z(''')$	$\mathrm{APE}^{\dagger}(''')$	$\widehat{\mathrm{APE}}'_{68}(''')$	$\widehat{\mathrm{APE}}_{68}(''')$
2	—	—	1.09	1.56	1.90	2.01	2.01
	-0.45	-1.71	1.23	1.51	1.95	2.07	2.84
	-0.35	-1.67	1.23	1.72	2.11	2.23	2.89
	-0.57	-1.75	0.93	1.22	1.53	1.63	2.61
3	-0.83	-0.17	0.85	1.80	1.99	2.04	2.27
	-0.58	-0.30	1.33	2.26	2.62	2.73	2.83
	-0.08	0.19	1.05	2.03	2.29	2.35	2.36
	-0.16	0.56	1.06	1.67	1.98	2.07	2.15
	-0.01	-0.58	1.10	1.19	1.62	1.73	1.84
	0.27	-0.57	1.12	2.05	2.34	2.42	2.50
	-0.62	0.55	1.28	2.65	2.94	3.01	3.15
4	-0.40	-0.13	1.17	2.05	2.36	2.45	2.50
	-0.26	0.90	0.92	1.12	1.45	1.54	1.84
	-0.27	0.54	0.70	0.65	0.96	1.02	1.22
	-0.13	0.20	0.79	0.80	1.12	1.20	1.23
5	0.38	0.77	1.24	2.01	2.36	2.47	2.62
	0.61	0.76	0.52	0.74	0.90	0.95	1.44
7	-0.34	-0.86	0.73	1.15	1.36	1.43	1.73
8	0.25	0.30	0.94	1.30	1.60	1.69	1.74

Table 2: Comparison of $\mathrm{APE}^{\dagger}$ with $\widehat{\mathrm{APE}}'_{68}$ (and $\widehat{\mathrm{APE}}_{68}$).

4.3 APE'_{68} or APE_{68}?

As noted in Section 3.3, it was initially believed that the measured mean offsets, m_y and m_z, would provide estimates of the instrument misalignments. It was therefore assumed that following each set of calibrations the SIAMs would be updated to reflect the newly-estimated alignments. This would then have had the effect of reducing the estimated value of the APE_{68} to $\widehat{\mathrm{APE}}'_{68}$ for the subsequent observations. (For observations prior to any such SIAM update, the value of the APE_{68} would of course remain unchanged, although

by $p = 1 - F_{\chi^2_1}(X^2) = 1 - P\left(\frac{1}{2}, \frac{X^2}{2}\right)$, where P is the lower incomplete gamma function.

[12]The data in the third row of Table 2 of [13, p. 12] are inconsistent: $\sqrt{1.23^2 + 1.72^2} \neq 2.24$. I have assumed that it is the root-sum-square value that is incorrect.

[13]To provide a check on the results, the computation of $\widehat{\mathrm{APE}}'_{68}$ was made both with (3), setting $\mu_y = \mu_z = 0$, and with (5).

the 68th percentile of the ground-based Absolute Measurement Error, the AME_{68}, would be reduced.)

However, at a certain point in the mission (during period 2) there was some doubt that the measured mean offsets between calibrations were actually reflecting real changes in the instrument alignments and a decision was taken to no longer update the SIAMs. Consequently, it would seem that the justification for using the estimate of the APE'_{68} (or its proxy the $\mathrm{APE}^{\dagger}$)— even as an indication of the Absolute Measurement Error—ends at this point in the mission and thereafter a more realistic measure is provided by the estimated APE_{68}.

The estimation of the APE_{68} is itself not without problems. As already described in [13, 15, 17], for much of the mission a large contribution to the APE came from bias errors introduced by the star tracker measurements. These bias errors are extremely attitude-dependent, resulting as they do from the particular locations of the guide stars within the field-of-view of the star tracker or from thermoelastic distortion, which it is conjectured may arise if reflected sunlight impinges on the star tracker mounting when observing at extreme values of the solar aspect angle. Since all the calibration observations associated with a given observation cycle (or OD range) were constrained to use a very similar attitude [13, pp. 10–11], it follows that the measured attitude errors for a short OD range are not very representative of those which were actually experienced during the scientific observations, when the spacecraft may have been at quite different attitudes. This may also provide an explanation for the occasionally large changes in the mean offsets that were measured for different observation cycles.

For example, consider the five sets of measurements shown in Table 3 (taken from [13, Table 3]). To investigate whether the population means, μ_y and μ_z, have changed between sets corresponding to adjacent ranges (of observation cycles) we apply Welch's t-test.[14] The null-hypothesis is that the population means are equal, so we use a two-tailed test. The resulting values of the test statistic, t_{ij}, the degrees-of-freedom, ν_{ij}, and the p-value,

[14] To allow for the fact that the population variances may be unequal, Welch's version of Student's t-test has been used. It would be interesting to test the data for homoscedasticity.

p_{ij}, for each case and for each axis are given in Table 4.[15] It is seen that

Set	Cycles	n	m_y ($''$)	m_z ($''$)	s_y^2 ($''^2$)	s_z^2 ($''^2$)
1	15–22	45	-0.83	-0.17	0.72	3.24
2	23–31	51	-0.58	-0.30	1.77	5.11
3a	32–36	40	-0.08	0.19	1.10	4.12
3b	33–36	34	-0.16	0.56	1.12	2.79
4	37–39	22	-0.01	-0.58	1.21	1.42

Table 3: Measured means and variances

Case		y-axis			z-axis		
i	j	t_{ij}	ν_{ij}	p_{ij}	t_{ij}	ν_{ij}	p_{ij}
1	2	-1.11	86	0.27	0.31	93	0.76
2	3a	-2.00	89	0.05	-1.09	87	0.28
2	3b	-1.61	80	0.11	-2.01	82	0.05
3a	4	-0.24	42	0.81	1.88	60	0.07
3b	4	-0.51	44	0.61	2.98	53	0.00(4)

Table 4: Welch's test: test statistic, degrees-of-freedom and p-value

there are two borderline cases where the null-hypothesis could be rejected at the 0.05 level and one case (the comparison of m_z from sets 3b and 4), where it would be rejected even at the 0.005 level. So only in these three cases (particularly the last) is there good evidence for concluding that there has been a real change in mean value of the offsets.[16] However, as noted above,

[15] Let m_k, s_k^2 and n_k be the mean, variance and size of sample (set) k, then

$$t_{ij} = \frac{m_i - m_j}{\sqrt{\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}}}, \quad \nu_{ij} \approx \frac{\left(\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}\right)^2}{\frac{s_i^4}{n_i^2\nu_i} + \frac{s_j^4}{n_j^2\nu_j}},$$

where $\nu_k = n_k - 1$. The p-values were obtained using an on-line calculator: `http://onlinestatbook.com/lms/calculators/t_dist.html`.

[16] An application of Welch's t-test to the data sets for ODs 731 and 733 [13, Table 3], produces p-values of 0.03 (y-axis) and 0.15 (z-axis). It appears that the explanation for this significant (at the 0.05 level) difference in the means for the y-axis offsets is that the second set of calibrations was performed with the redundant star tracker.

it is likely that much of the mean offsets is due to star tracker bias errors associated with the particular attitudes used for the calibration observations.

Work is currently underway to improve the accuracy of the ground-based pointing reconstruction. This accuracy is reflected in the AME_{68}, the estimation of which may be performed in an identical manner to that of the APE_{68}, with the exception that the ground-based reconstructed attitude replaces the target attitude, i.e. $(\Delta Y_i,\ \Delta Z_i)$ become the offsets between the ground-based estimate of the PACS boresight and its 'true' (PACS measured) direction. It is expected that, following the implementation of these improvements, the measured mean offsets will provide improved estimates of the instrument alignments.

5 Alternative method of computing APE_{68}

The method which has been described in Section 3 assumes that the measured pointing offsets, $\{\Delta Y_i : i = 1, \ldots, n\}$ and $\{\Delta Z_i : i = 1, \ldots, n\}$, are realizations of the (independent) random variables $\Delta\mathrm{Y} \sim N(\mu_y, \sigma_y^2)$ and $\Delta\mathrm{Z} \sim N(\mu_z, \sigma_z^2)$ and uses the two samples to estimate μ_y, μ_z, σ_y^2 and σ_z^2. The APE_{68} that is then computed is the 68th percentile of $\Delta R = \sqrt{(\Delta Y)^2 + (\Delta Z)^2}$.

However, it might be argued that the 68th percentile we are interested in is not that associated with some theoretical infinite populations, but rather that directly related to the actual (finite number of) observations corresponding to a given period of the mission (perhaps the entire mission) and of perhaps a certain type (pointed, scan, ...). Moreover, to conform with the requirements specification [6, ch. 4, p. 16], the APE should include the spacecraft jitter and the 68th percentile which we should be calculating is with respect to the time, not with respect to the number of observations. An alternative approach to the calculation, which remedies these three issues, is simply to use the ground-based reconstructed attitude to find how $\Delta R(t)$ varies with the time, t, over each of the observations within a given set, and from this to construct the distribution function for ΔR and hence compute the APE_{68}.

Let the period for which we wish to calculate the APE_{68} contain m observations (of the chosen type) corresponding to the intervals $[t_s^{(1)}, t_e^{(1)}]$, $[t_s^{(2)}, t_e^{(2)}]$, ..., $[t_s^{(m)}, t_e^{(m)}]$, and let the distribution function of $\Delta\mathrm{R}$ for the ith observation be $F_{\Delta\mathrm{R}}^{(i)}(\Delta R)$. Then the distribution function for the entire period is given

by:

$$F_{\Delta\mathrm{R}}(\Delta R) = \frac{1}{T}\sum_{i=1}^{m} F_{\Delta\mathrm{R}}^{(i)}(\Delta R)\,[t_e^{(i)} - t_s^{(i)}],$$

where $T = \sum_{i=1}^{m}[t_e^{(i)} - t_s^{(i)}]$.[17]

The distribution function $F_{\Delta\mathrm{R}}^{(i)}(\Delta R)$ for each individual observation could be constructed empirically by comparing the target instrument boresight direction (which might be approximated using the on-board filtered ACA-attitude and the SIAM applied during operations) with the true instrument boresight direction (as approximated using the ground-based reconstructed ACA-attitude and an improved estimate of the instrument alignment) over the interval $t_s^{(i)} \leq t \leq t_e^{(i)}$ and finding the fraction of the observation for which the error ΔR is less than certain user-selected values.[18]

Where the calibration measurements could still play a role is in obtaining the improved estimate of the instrument alignments. The mispointing offsets may be decomposed into three components: (i) the offsets due to the bias errors in the star tracker attitude measurements; (ii) the offsets due to the jitter in the ACA-frame attitude; and (iii) the offsets due to the misalignment between the ACA-frame and the instrument (PACS) reference frame. The new ground-based pointing reconstruction should be able to account for most of the bias in the star tracker attitude measurements (i.e. the part resulting from errors in the measured star positions) and the jitter [see 3, 4, 7, 8, 15, 16].[19] To obtain improved estimates of the instrument alignments we might therefore compare the reconstructed boresight direction with the boresight direction as measured by PACS. It is hoped that it may also be possible to model the effects of thermoelastic distortion by correlating this relative alignment with a temperature measured on the spacecraft somewhere in the viscinity of the star tracker supporting struts.

[17] To treat ΔR as a random variable, it is simply necessary to consider it to be a function defined on the probability space associated with choosing times randomly and uniformly from $\bigcup_{i=1}^{m}[t_s^{(i)}, t_e^{(i)}]$.

[18] Perhaps, as a first approximation—particularly for pointed observations early in the mission (when star tracker bias errors dominated)—a constant value for ΔR could be computed for each observation. In this case, $F_{\Delta\mathrm{R}}(\Delta R)$ simplifies to a weighted sum of Heaviside step functions.

[19] A small bias in the star tracker attitude measurements will also result from errors (e.g. due to proper motion) in the catalogue positions of the tracked stars.

A Standard errors in the estimated standard deviations

From (6), estimates of the standard errors in the estimated population standard deviations, s_y and s_z, are given by:

$$\begin{aligned} \widehat{\sigma}_{\mathrm{s}_y} &= \frac{s_y}{\sqrt{n-1}}\sqrt{\mathrm{var}\{\chi_{n-1}\}} \\ \widehat{\sigma}_{\mathrm{s}_z} &= \frac{s_z}{\sqrt{n-1}}\sqrt{\mathrm{var}\{\chi_{n-1}\}}, \end{aligned} \tag{10}$$

where[20]

$$\mathrm{var}\{\chi_{n-1}\} = n - 1 - 2\left[\frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}\right]^2. \tag{11}$$

To obtain an idea of how these standard errors vary when n is large, we derive an asymptotic series for $\sqrt{\mathrm{var}\{\chi_{n-1}\}}$.[21]

Let

$$f(m) = \frac{1}{m^{1/2}}\frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma\left(m\right)},$$

where Γ is the gamma function. It may be shown [9, p. 602],[22] that the function f has the asymptotic expansion

$$f(m) \sim 1 - \frac{1}{8m} + \frac{1}{128m^2} + \frac{5}{1024m^3} + \cdots, \quad \text{as } m \to \infty.$$

Since an asymptotic power series can be raised to an arbitrary power [11, p. 298], it follows that

$$[f(m)]^2 \sim 1 - \frac{1}{4m} + \frac{1}{32m^2} + \frac{1}{128m^3} + \cdots, \quad \text{as } m \to \infty.$$

Therefore, from (11), we obtain:

$$\begin{aligned} \mathrm{var}\{\chi_{n-1}\} &= (n-1)\left(1 - \left[f\left(\frac{n-1}{2}\right)\right]^2\right) \\ &\sim \frac{1}{2} - \frac{1}{8(n-1)} - \frac{1}{16(n-1)^2} + \cdots, \quad \text{as } n \to \infty \end{aligned}$$

[20] See `http://en.wikipedia.org/wiki/Chi_distribution` (accessed on 18 March 2014).

[21] In fact, it will be seen that the first few terms of this series can be used to provide a very good approximation for the (normalized) standard errors even when n is small.

[22] Graham et al. use the generalized factorial power $m^{\overline{1/2}}$ for $\frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m)}$ [see 9, p. 211].

and hence

$$\sqrt{\mathrm{var}\{\chi_{n-1}\}} \sim \left(\frac{1}{2} - \frac{1}{8(n-1)} - \frac{1}{16(n-1)^2} + \cdots\right)^{\frac{1}{2}}$$
$$\sim \frac{\sqrt{2}}{2}\left(1 - \frac{1}{8(n-1)} - \frac{9}{128(n-1)^2} + \cdots\right), \quad \text{as } n \to \infty.$$

So finally from (10) we have:

$$\widehat{\sigma}_{s_y} \sim \frac{s_y}{\sqrt{2(n-1)}}\left(1 - \frac{1}{8(n-1)} - \frac{9}{128(n-1)^2} + \cdots\right), \quad \text{as } n \to \infty, \quad (12)$$

with a similar expression for $\widehat{\sigma}_{s_z}$.[23] Figure 1 compares the values of the normalized standard error, $\widehat{\sigma}_{s_y}/s_y$, computed using the MATLAB gamma function, with the approximate values obtained from (12).

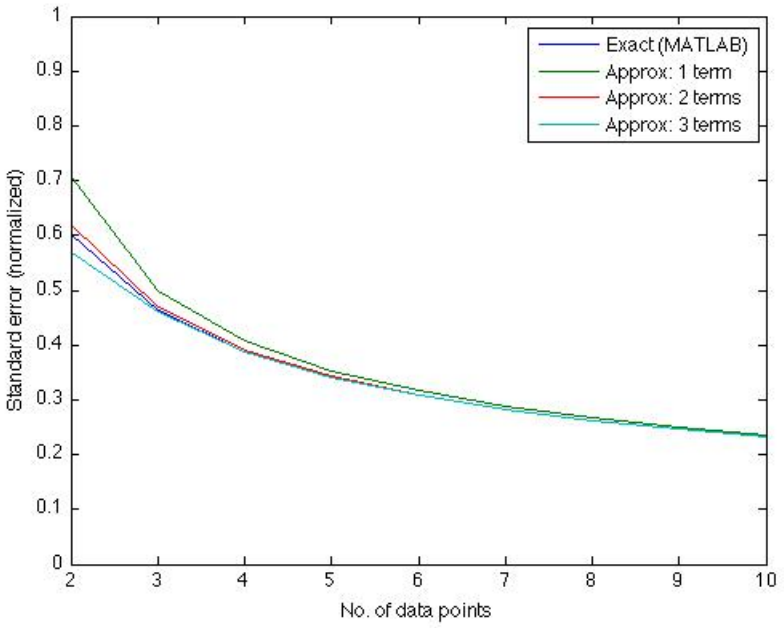

Figure 1: Approximation of $\widehat{\sigma}_{s_y}/s_y$ using (12)

[23] The first terms of these asymptotic expansions are unchanged by the use of an unbiased estimator [see 1].

References

[1] S. Ahn and J. A. Fessler. Standard Errors of Mean, Variance, and Standard Deviation Estimators. Jul 2003. URL `http://web.eecs.umich.edu/~fessler/papers/files/tr/stderr.pdf`.

[2] T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2):193–212, 1952.

[3] H. Aussel. Towards an accurate reconstruction of Herschel pointing. Technical Report SAp-PACS-HA-0729-11, Issue 1.0, CEA Saclay, Nov 2011.

[4] H. Aussel. PACS pointing jitter : the GYR view. Technical Report SAp-PACS-HA-0730-12, Issue 1.0, CEA Saclay, Aug 2012.

[5] D. G. Dungate. ESA Pointing Error Handbook. Technical Report ESA-NCR-502, volume 1, VEGA, Feb 1993.

[6] ESA Herschel/Planck Project Team. System Requirements Specification [SRS]. Technical Report SCI-PT-RS-05991, Issue 3/3, ESA, Jul 2004.

[7] H. Feuchtgruber. Herschel STR-A CCD Sub-Pixel Structure. Technical Report PICC-ME-TN-041, Issue 1.0, Herschel PACS, Mar 2012.

[8] H. Feuchtgruber. Reconstruction of the Herschel Pointing Jitter. Technical Report PICC-ME-TN-042, Draft, Revision 1.0, Herschel PACS, Sep 2012.

[9] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison–Wesley, second edition, 1994.

[10] N. H. Kuiper. Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series A*, 63:38–47, 1960.

[11] H. Poincaré. Sur les intégrales irrégulières des équations linéaires. *Acta Mathematica*, 8:295–344, 1886.

[12] M. Sánchez-Portal, B. Altieri, D. Lutz, A. Marston, U. Klaas, M. Nielbock, and the Herschel Pointing Working group. Herschel Pointing Calibration Report. Technical Report HERSCHEL-HSC-DOC-1515, Issue 1.0, Herschel Science Centre, Nov 2009.

[13] M. Sánchez-Portal, A. Marston, B. Altieri, H. Aussel, H. Feuchtgruber, U. Klaas, D. Lutz, B. Merín, M. Nielbock, M. Oort, G. Pilbratt, M. Schmidt, C. Stephenson, M. Tuttlebee, and The Herschel Pointing Working Group. The Pointing System of the *Herschel* Space Observatory: Description, Calibration, Performance and Improvements. *Experimental Astronomy*, 2014. Draft version submitted on 11 Nov 2013.

[14] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3 and 4):591–611, Dec 1965.

[15] C. Stephenson. Improvement of the Herschel star tracker attitude measurements. Sep 2013.

[16] C. Stephenson. Herschel: gyro-based attitude reconstruction. Oct 2013.

[17] M. Tuttlebee. HERSCHEL/PLANCK Star Tracker Performance Assessment and Calibration. Technical Report PT-CMOC-OPS-RP-6435-HSO-GF, Issue 1.0, ESOC Flight Dynamics, Aug 2013.