Pointing of Optical Telescopes

S. R. Kulkarni

February 18, 2021–February 28, 2021

Executive Summary. I present an investigation of the pointing performance of the 48-inch Oschin telescope. I fit a classical 7-parameter model ("TPOINT" of P. Wallace) to the observed offsets. On applying the model the scatter is reduced to radius less than 30 arc-seconds (99%). The TPOINT model parameters show that the largest terms contributing to the error budget are due to collimation and flexure. Beyond that there are unexplained offsets arising in the far East (HA$>4\,$hr) and far North ($\delta > 60^\circ$).

1 Motivation

Accurate pointing increases the efficiency of a telescope. Astronomers will not be perusing finding charts when using telescope which points to arcsecond accuracy – whether they are doing imaging or spectroscopy. They can start long exposures without having a dull feeling in their stomach.

I had heard of Patrick Wallace and his pioneering work on accurate pointing of telescopes. I recently had the time to read the literature. I learnt that already in the seventies he had achieved precision pointing of the AAT and the UK 1.2-m Schmidt telescope at the level of only a few arc-seconds (rms, in each axis). This made me to pose the following question: "What is the limiting accuracy[1] of the three main telescopes at the Palomar Observatory?" I do recognize that the Palomar telescopes are of earlier vintage (P200: first light, 1949; P48: first light, 1951; P60: first light, 1972). Nonetheless, the question I pose is interesting: determine the ultimate accuracy of a telescope using the best algorithm(s).[2] The method should account for temperature variations, large slews and hysteresis (if present).

[1] By pointing I mean absolute pointing, sometimes also called as blind pointing.

[2] Physical model followed by empirical or Machine Learning, for instance.

I perused through the literature about pointing telescopes and collected articles of pedagogical value.[3] For data I decided to use the pointing data of the equatorial 48-inch Oschin telescope (P48) which houses a 16-CCD imager (see below). In order to reach its prime goal of covering the night sky every two nights, P48 observes typically 500 fields per night. Thus there is plenty of juicy pointing data. However, the complication is that the CCD camera is mounted on a hexapod which is dynamically adjusted in all three axes: tip, tilt and piston (focus). The usual pointing formulations do not include such dynamic terms.

For my first analysis, I elected to stick with a basic pointing model which accounts for (1) non-alignment of the axis of the fork of the telescope and the Earth's rotation, (2) deviation from orthogonality between the axes of the declination and hour angle motors, (3) lack of parellism between the optical and mechanical axes of the telescope (collimation) and (4) a simple model for the flexure of the telescope. This model is described in a classic paper by Wallace & Tritton (1979) which, in fact, was applied to the UK 48-inch Schmidt telescope ("U48") – a later version of P48. That simple model showed that the U48 could reliably point to better than 10 arc-seconds (in radius).

2 The Oschin 48-inch telescope

The Zwicky Transient Facility (ZTF) is based on two telescope: the Palomar 60-inch telescope (P60) carrying the SEDM) and P48. Weather permitting, P48 relentlessly images the sky every night. A "world-coordinate system" (WCS) is applied to the images. The two primary angles for P48, a telescope with an equatorial mount, are HA (h) and declination (DEC, δ). The difference between the requested and observed positions is recorded and reported as $\Delta\alpha^* \equiv \Delta\alpha\cos(\delta)$ (hereafter, `raoff`) and $\Delta\delta$ (`decoff`). While $\Delta\alpha = \Delta h$ it is important to note that astronomers use $\Delta\alpha*$ – the certain mistakes of tyros and occasional mistakes of experts (sigh).

3 First Look

Data. Richard Walters set up a simple API service[4] which yields offsets for a given night. Each field observed during a given night results in a data record: `time, ra, dec, ha, raoff, decoff`. The collection of records for a given night constitute our basic and essen-

[3] `https://sites.astro.caltech.edu/~srk/TP/TP.html` Papers quoted in this report can be found at this URL.

[4] `http://skyvision.caltech.edu:88/get_pointing_data?obsdate=YYYY-MM-DD` where YYYY, MM and DD have to be set to specific values, e.g. 2021-02-01

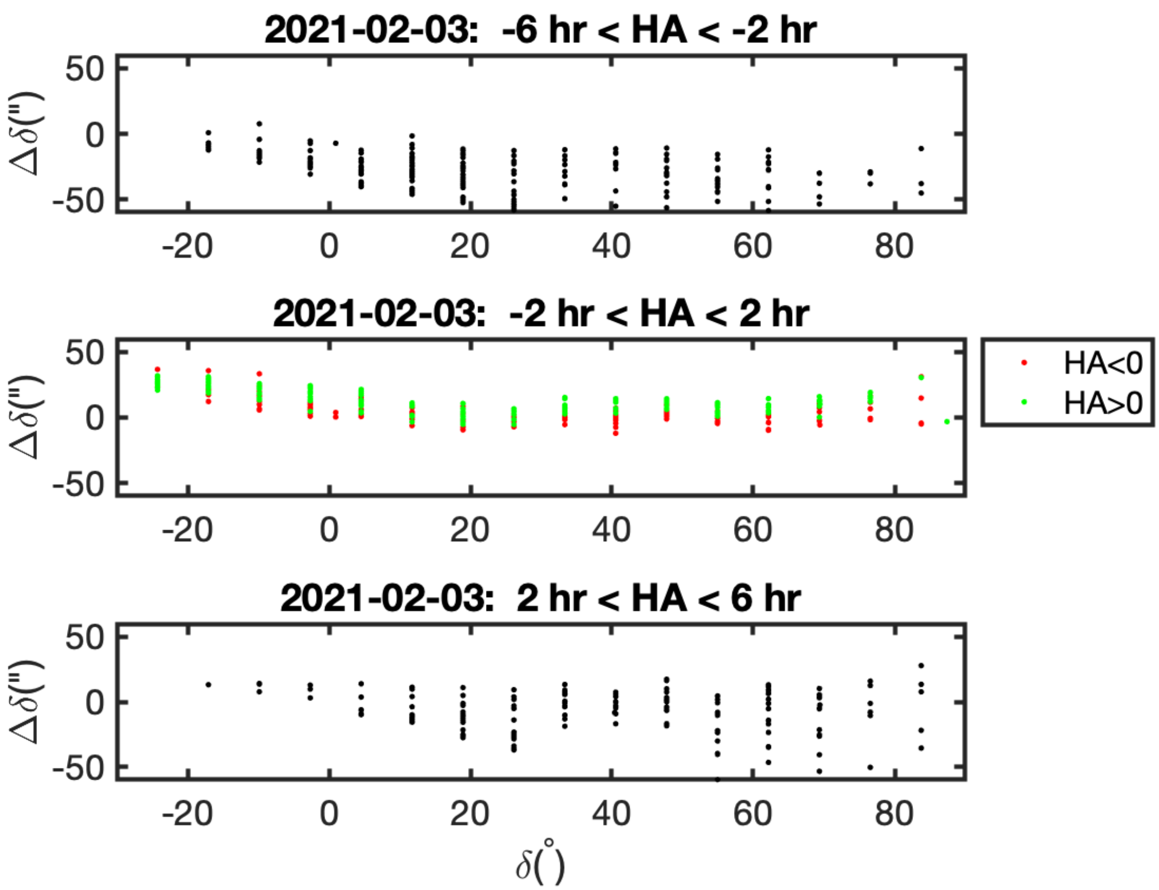

Figure 1: Run of $\Delta\delta$ with respect to δ for bands of HA.

tial data. Arguable refined models would benefit from having pressure and temperature (top of dome, bottom of dome, ambient) data.

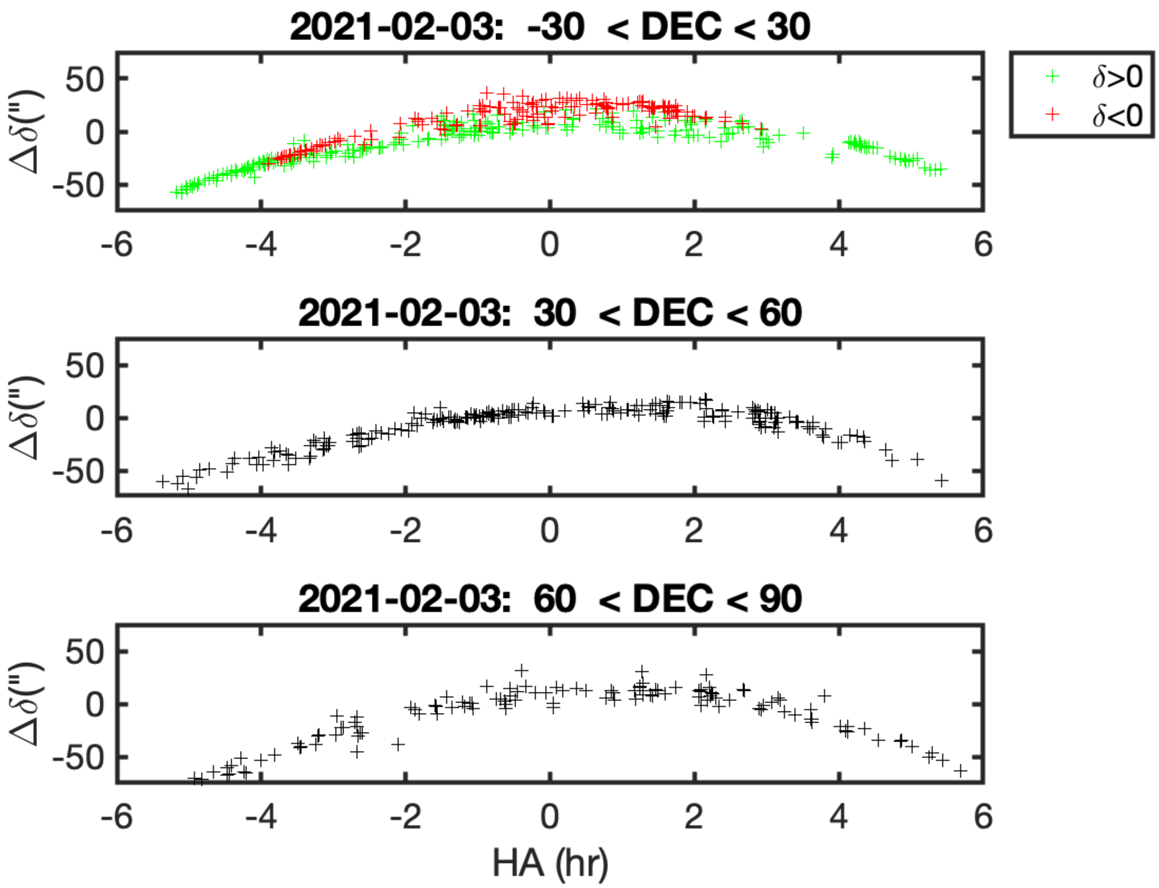

Figure 2: Run of $\Delta\delta$ with respect to HA for bands of DEC.

For our first look, I display the measured offsets as a function of hour angle h and δ – or a total of four plots (Figures 1- 4). In order to interpret the data more sensibly we will review the data in the framework of a physical model which is described in §A. A neophyte is well advised to review the Appendix before proceeding further.

From Equation 8 we see that $\Delta\delta$ does not depend on the declination. In contrast to this expectation, as can be seen from Figure 1 we see that $\Delta\delta$ is about 20 arc-seconds at low declinations and easing to zero at the equator. Next, let us review the run of $\Delta\delta$ versus HA Figure 4. According to Equation 8 we that there are two terms: $\mathrm{ME} + \mathrm{FO}\cos(h)$ and $\mathrm{MA}\sin(h)$. In the top panel ($-30^\circ < \delta < 30^\circ$) we see a clear segregation between $\delta > 0$ and $\delta < 0$. From Figure 1 we already deduced this dependence on δ. The overall shape

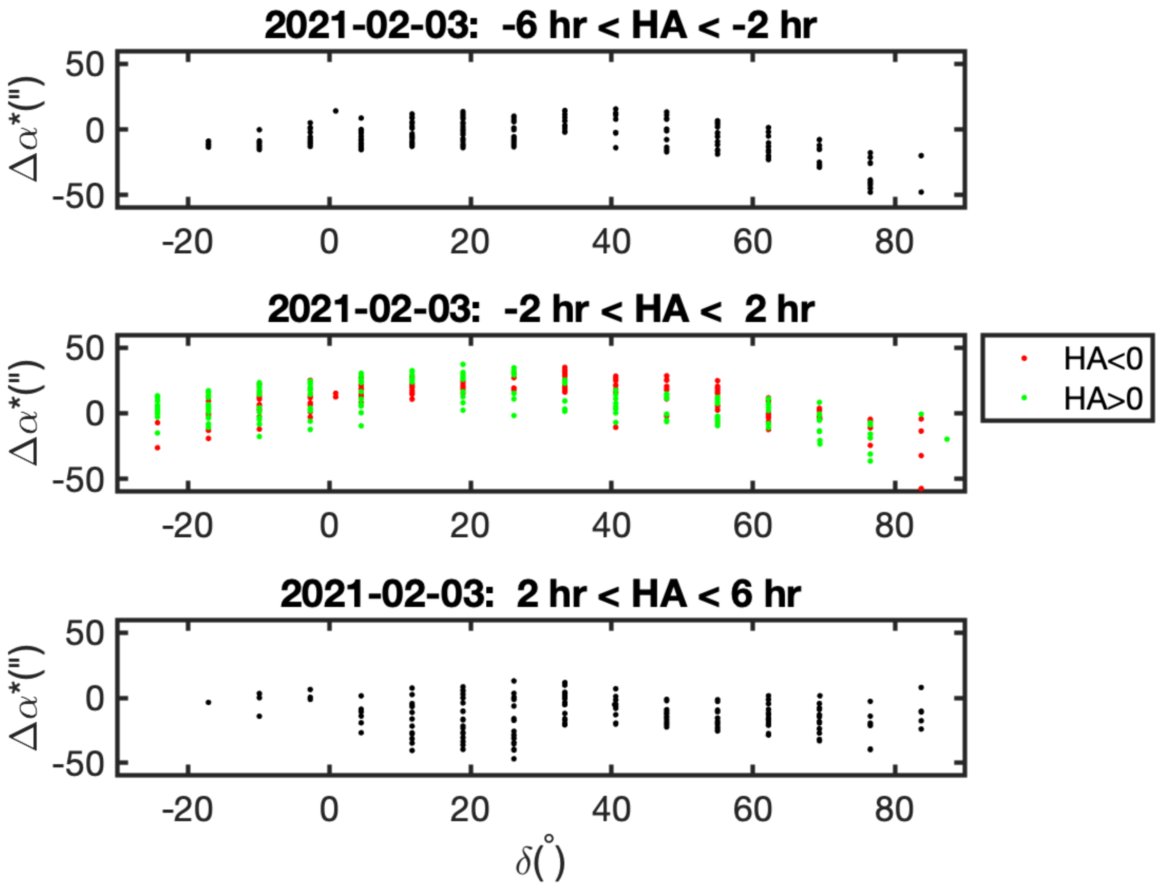

Figure 3: Run of $\Delta\alpha*$ with respect to δ for bands of HA.

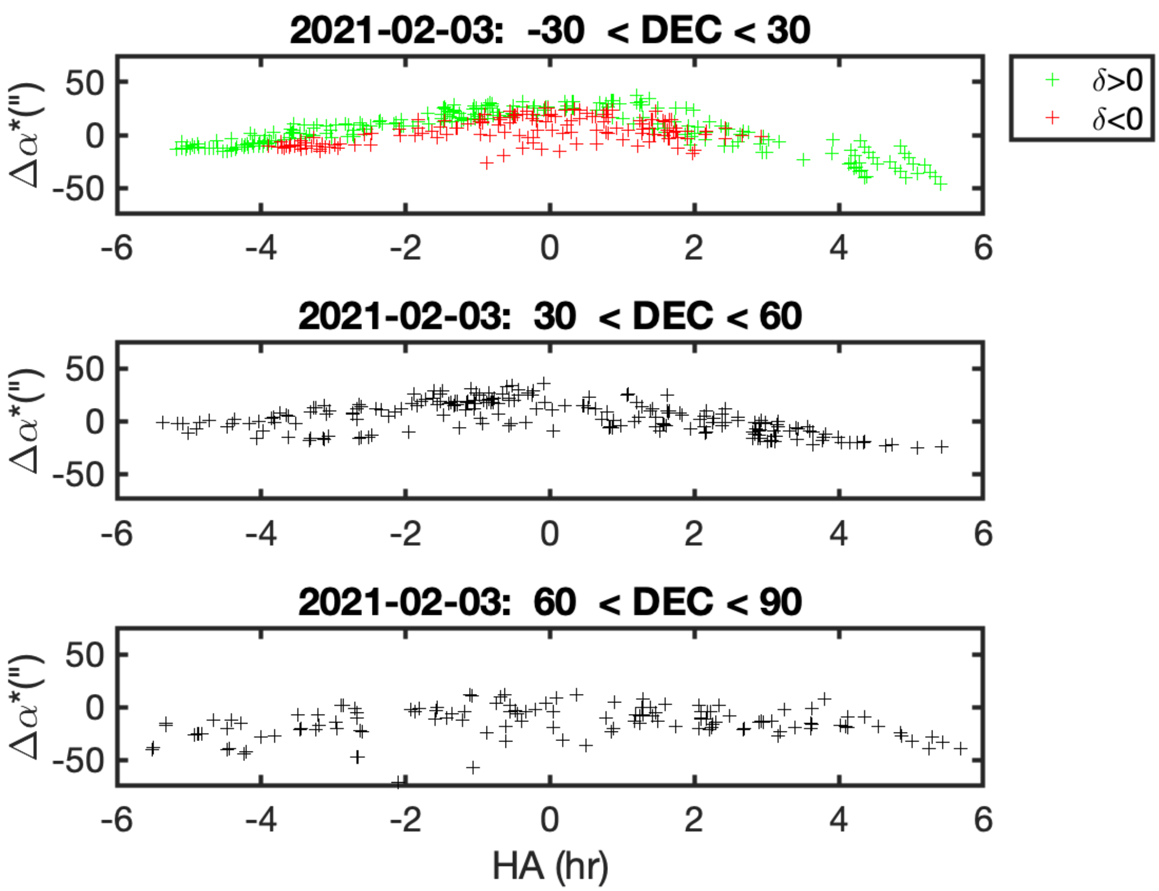

Figure 4: Run of $\Delta\alpha*$ with respect to HA for bands of DEC.

indicates a cosine term. So the MA component must be smaller than ME+FO. Finally, there is mysterious hump at large positive HA and large deviations when the telescope points to high declinations.

4 Programs

I have written a bash script for downloading the data and Matlab programs for display and analysis. Below ">>" stands for Matlab prompt and "$" for Unix prompt.

```
1. $ getp48p 2021-02-04 %generates file P2021_02_04.csv
>> INFILE='P2021_02_04.csv';   %nightly file
>> P48p_readin   ... creates structure P (housing time,ra,dec,ha,raoff,decoff)

2. Visualize the measured offsets as a function of HA and dec
>> display__measured_offsets

3. Deduce the seven TPOINT coefficients from structure P
>> tpoint  ... the 7 parameters are stored in structure MODEL
               ... also written out as P2101_02_04.model

4. Plot a summary of residuals after applying MODEL
>> plot_summary_postfit_residuals

5. As with measured offsets visualize residuals as a function of HA and dec
>>display_residual_offsets
```

5 TPOINT fit for one night

The 7-parameters, all in arc-seconds, are as follows (for one night)

```
ME: 10.2
MA: -1.0
NP:  6.9
CH: 101.6
FO: 73.1
ID: -13.5
IH:  41.9
```

It appears that the optical collimation error (CH) and flexure dominate the error generation.In Figure 5 I display the "before" and "after" view of the offsets. The telescope

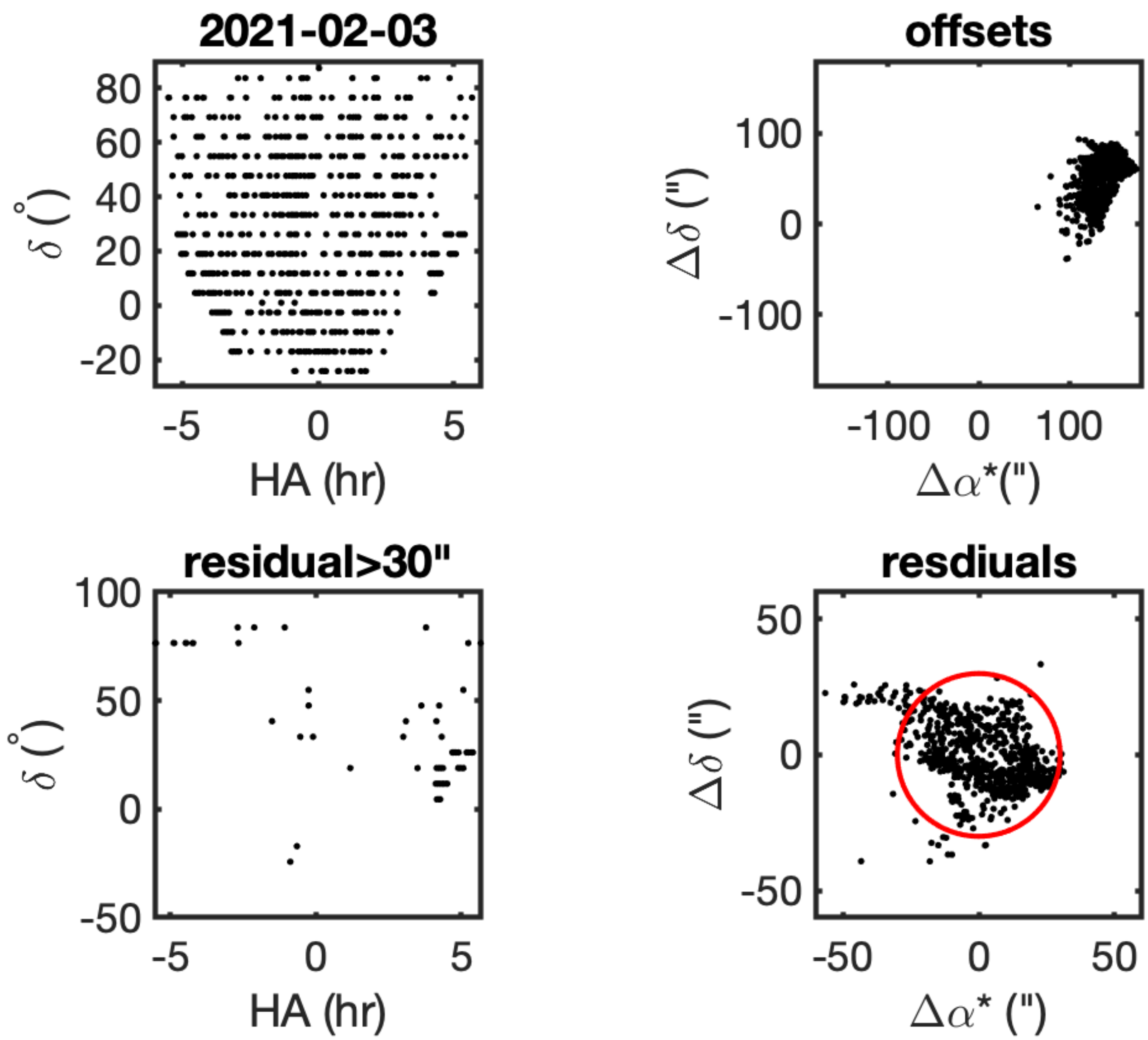

Figure 5: Reading clockwise, from top left. [1] Each dot is an observation in the HA-δ plane. [2] The measured offsets; here $\Delta\alpha* = \Delta\alpha\cos(\delta)$. The plane is α-δ plane. [3] The residuals following fitting the offsets to a TPOINT model. The radius of the red circle is 30 arc-seconds. [4] The observations which generate post-residual terms outside the red circle are plotted on HA-δ plane. Notice the excess at large HA and in the North.

generates larger offsets in the far East (large HA) far North.

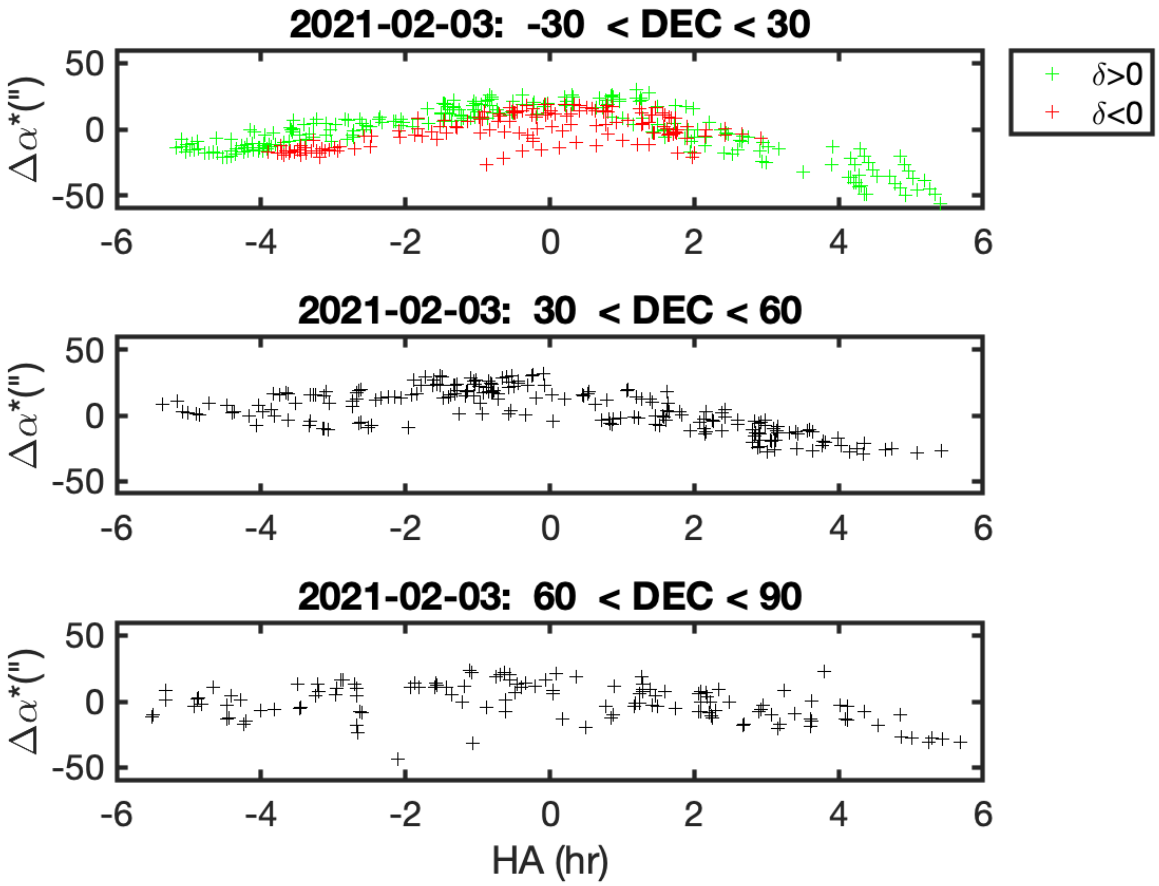

Figure 6:

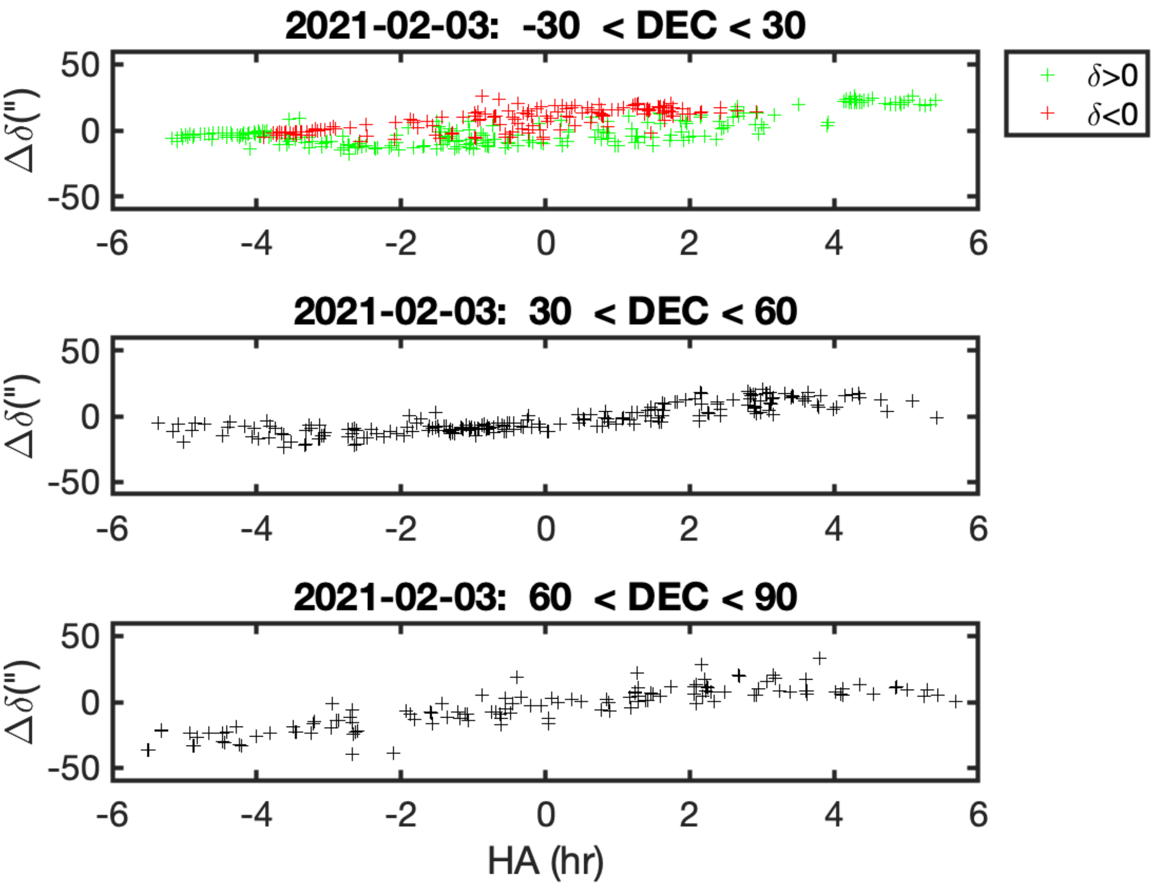
2021-02-03: -30 < DEC < 30
δ>0
δ<0
Δδ(")
2021-02-03: 30 < DEC < 60
Δδ(")
2021-02-03: 60 < DEC < 90
Δδ(")
HA (hr)

Figure 7:

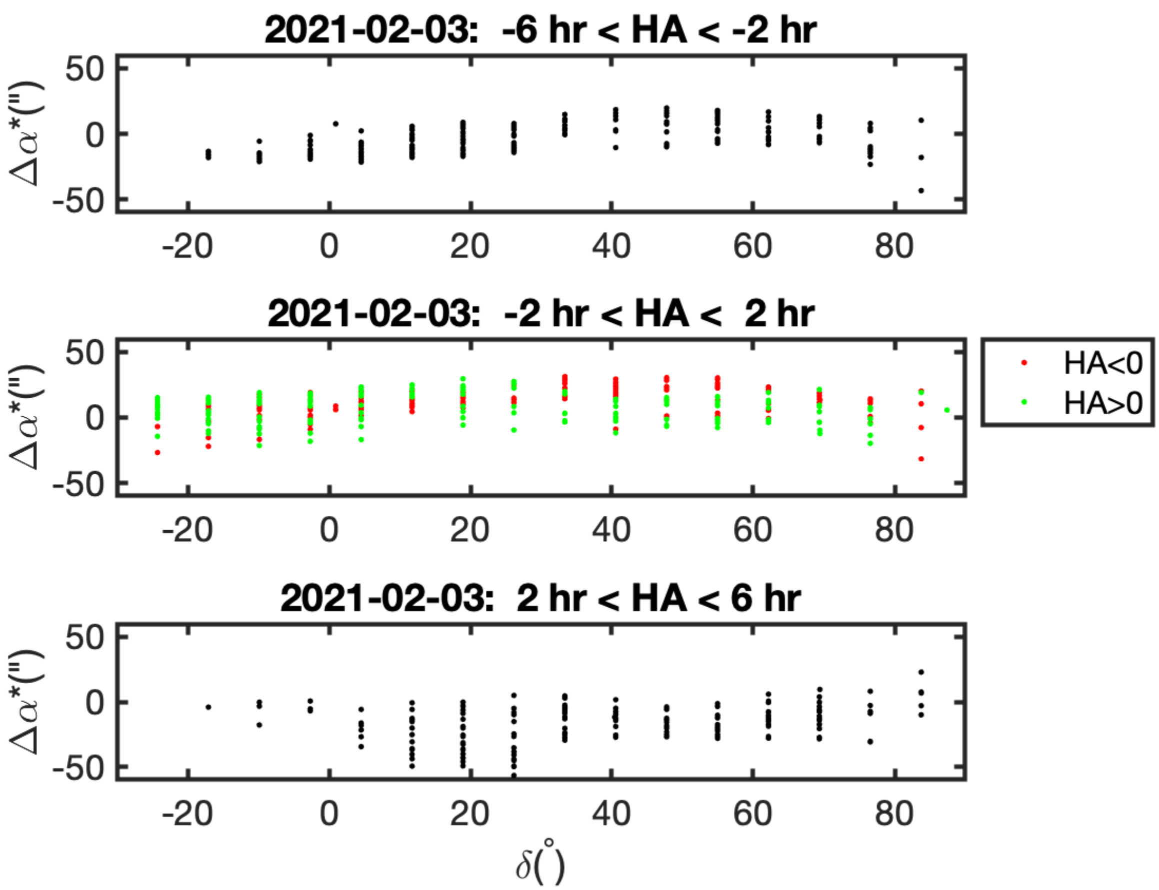
2021-02-03: -6 hr < HA < -2 hr
2021-02-03: -2 hr < HA < 2 hr
2021-02-03: 2 hr < HA < 6 hr
Δα*(")
δ(°)
HA<0
HA>0
50
0
-50
-20
0
20
40
60
80

Figure 8:

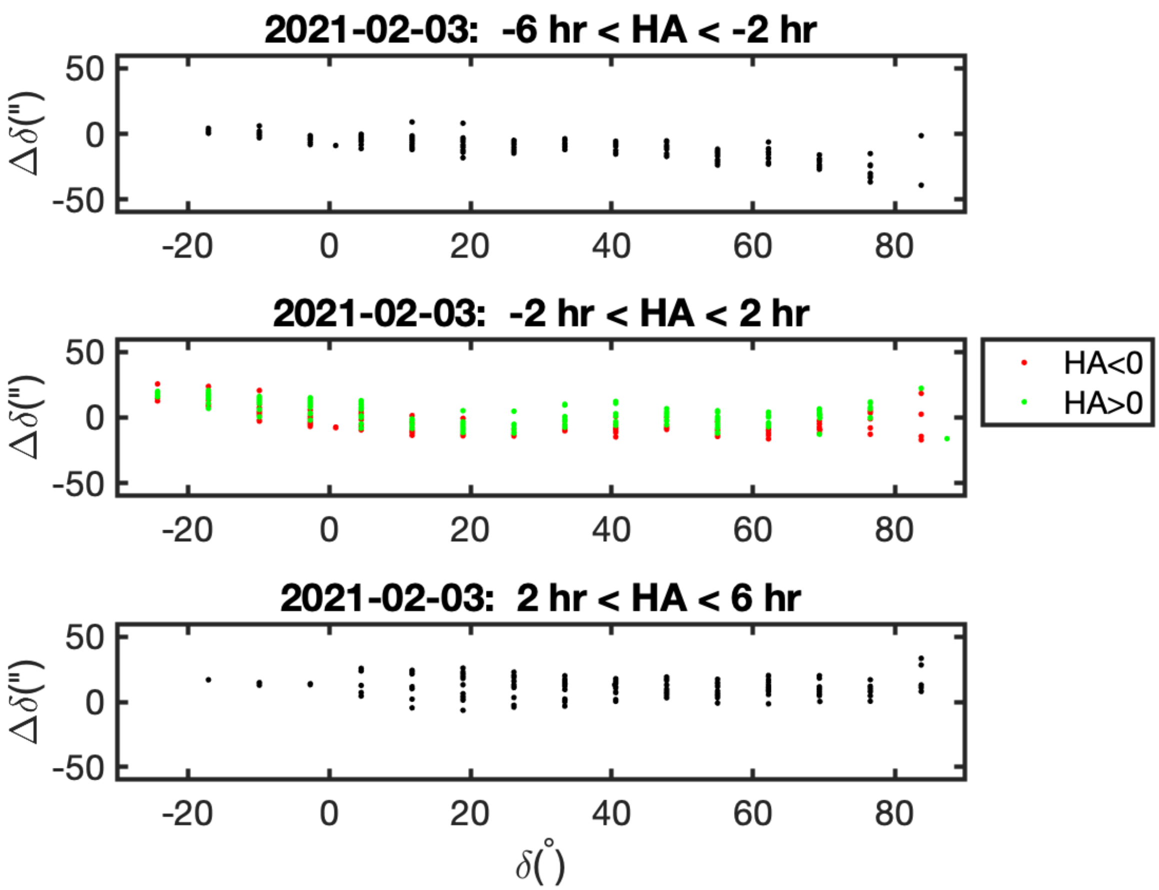
2021-02-03: -6 hr < HA < -2 hr
2021-02-03: -2 hr < HA < 2 hr
HA<0
HA>0
2021-02-03: 2 hr < HA < 6 hr
Δδ(")
δ(°)
50
0
-50
-20
0
20
40
60
80

Figure 9:

6 Run of TPOINT parameters

In the three figures below the run of the 7-parameters for the month of February can be found (Figures 10, 11 and 12). From these figures it appears that CH and FO dominate the error budget. It remains to be seen if the pole wander (ME,MA) and wander of NP is real or a consequence of CH and FO.

A possible next step would be to construct an "average" model by averaging, in some fashion, the seven parameters obtained nightly. One simple approach is to take the median of each parameter. This is the approach taken in `build_TP_model` (below). This model then can be applied to all the nights.

```
A.  Redo steps 1--4 for nights.
$ cat *.model > a; mv a all.model
>> build_TP_model  .. build a MODEL based on many nights of data

B. Review the post-fit residuals for each night after applying model
>> INFILE='P2021_02_04.csv'; p48p_readin; plot_summary_postfit_residuals
```

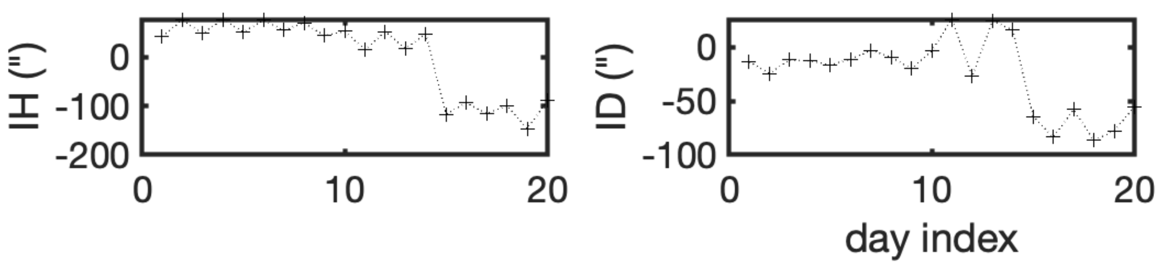

Figure 10: On my request Walters made a "bore-sight" fix on the evening of 19 February 2021 – which explains the jump on day 15.

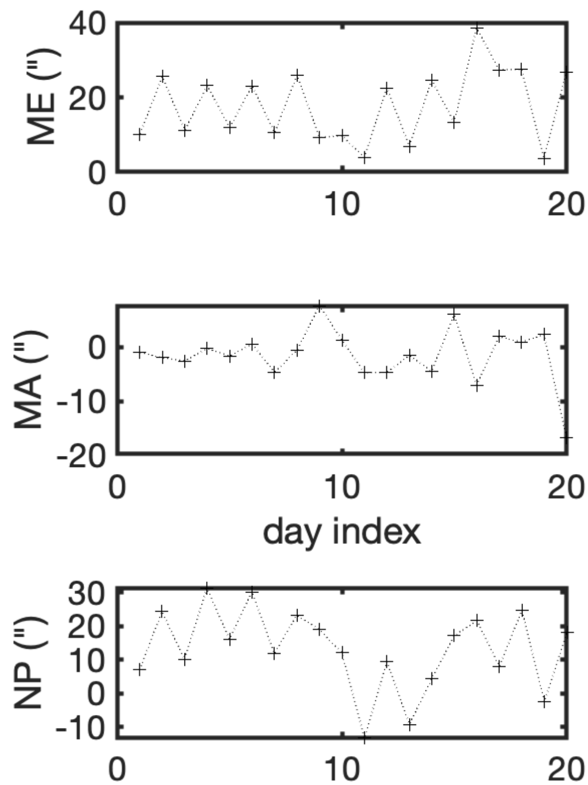

Figure 11:

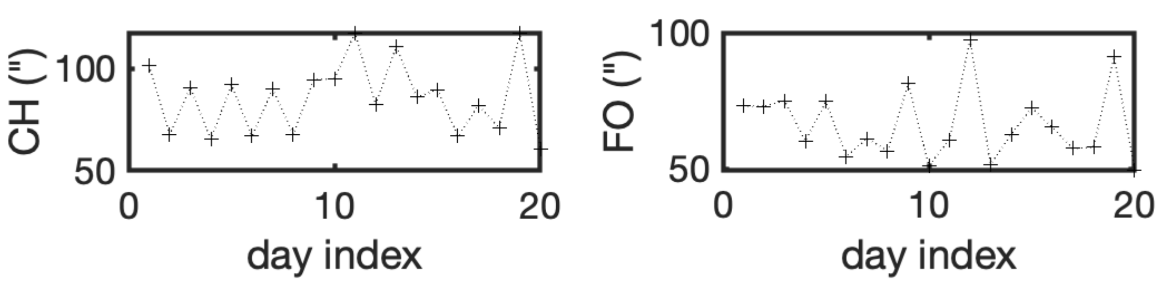

Figure 12: The values of collimation (CH) and overhang due to flexure (FO) are considerable.

A The basic 7-parameter pointing model

Wallace & Tritton (1976) present a basic geometrical model for an equatorial telescope (which, with rotation, can be adopted for an alt-az telescope also). The model with seven parameters admits of three principal errors and allows for a rudimentary model for flexure error.

Misalignment of the polar axis. The polar axis of the telescope is defined by the mount (the giant yoke in the case of P200). The offset of the polar axis with respect to the rotation axis of the Earth has two components: in elevation (ME) and in azimuth (MA). You can raise or lower the pole of the mount to change the elevation angle or rotate the yoke to address the error in azimuth. The misalignment leads to errors in both HA and δ.

$$\Delta h = \tan(\delta)(\mathrm{ME}\sin(h) - \mathrm{MA}\cos(h)) \tag{1}$$
$$\Delta\delta = \mathrm{ME}\cos(h) + \mathrm{MA}\sin(h) \tag{2}$$

Non-perpendicularity of the RA and DEC axes. Ideally these two axes are supposed to be perpendicular to each other. Let NP the angle by which the two axis deviate from being perpendicular.

$$\Delta h = \mathrm{NP}\tan\delta \tag{3}$$

Optical Collimation Error. So far we have two of errors associated with the mount of the telescope. However, the observations are made via telescope optics. As with polar-axis misalignment there are two components: CH and CE.

$$\Delta h = \mathrm{CH}\sec(\delta) \tag{4}$$
$$\Delta\delta = \mathrm{CE} \tag{5}$$

Flexure. This can be complicated term, requiring understanding of the mechanical structure of the telescope. Wallace advocates the following as a start.

$$\Delta\delta = \mathrm{FO}\cos(h) \tag{6}$$

with no error in h.

Zero point errors. In modern telescopes there are no clock errors so IH=0. However there could be zero point offsets in the encoders (both axies). Next, the declination error of optical collimation error can be absorbed into the zero point error for declination, ID. Finally, the “bore-sigh” errors are absorbed by IH and ID.

Note that these errors are in hour angle and declination. In particular, while Δh is the error the angular error in RA is $\Delta h \cos(\delta)$. Inversely, Δh and $\Delta\delta$ should be used in fitting the observed *angular offsets* to this model one should use Δh (and not $\Delta\alpha$) and $\Delta\delta$.

$$\Delta h = \tan(\delta)\big[\mathrm{ME}\sin(h) - \mathrm{MA}\cos(h) + \mathrm{NP}\big] + +\mathrm{CH}\sec(\delta) + \mathrm{IH} \quad (7)$$

$$\Delta\delta = \cos(h)\big[\mathrm{ME} + \mathrm{FO}\big] + \mathrm{MA}\sin(h) + \mathrm{ID} \quad (8)$$

A.1 Least squares formulation

Linear least squares is the quest for parameters that minimize a model which is linear in parameters (but not necessary in x_j). Outside astronomy (for example, machine learning), the popular term for linear least-squares is "linear regression". The basic equation of linear least squares is of the form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}_0 + \boldsymbol{\epsilon} \quad (9)$$

where y_i is the ith measurement and ϵ_i, the corresponding noise; here i ranges from 1 to n. The linear (unknown) parameters are given by the column vector $\boldsymbol{\beta}$ of size $(m, 1)$. Some jargon: $\mathbf{y}$ is called as the "response" vector, β_0 is the intercept, $\boldsymbol{\beta}$ is the slope and $\mathbf{X}$ is the "predictor" matrix. Usually the intercept is included in $\boldsymbol{\beta}$ (with X_{1i} set 1). When $m > 1$ this problem is called "multiple linear regression". Finally, the least squares is optimal only when $\boldsymbol{\epsilon}$ results from Gaussian processes. "Multi-variate (multiple)" linear regression is when $\mathbf{y}$ has more than one column. Note that the size of β is the same as that of $\mathbf{y}$.

The best-fit values for the parameters is given by the well-known formula

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

The post-fit value and the corresponding residuals are given by

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}, \qquad \boldsymbol{e} = \mathbf{y} - \hat{\mathbf{y}}.$$

Linear algebra assures you that the resulting sum of the squares of the residuals, $\boldsymbol{e}\boldsymbol{e}'$, is the smallest value possible.

At first blush, after reviewing Equations 7 and 8, it would appear that we should formulate a "multi-variate multiple" linear regression model. However, note that $\Delta\delta$ and Δh share some common linear parameters. It took me two days to realize that the simpler solution is to convert this problem into the standard multiple linear regression model. *One* approch is to represent the measurements as follows:

$$\mathbf{y} \equiv \begin{pmatrix} \Delta h_1 \\ \Delta \delta_1 \\ \Delta h_2 \\ \Delta \delta_2 \\ ... \end{pmatrix}$$

where the subscript is the index of the measurements, $i = 1, 2, ..n$. This approach is not unique. For instance, we could have listed all of Δh first and followed it by $\Delta \delta$. Consistent with this approach $\boldsymbol{\beta}$ is formulated to include both overlapping (MA, MH) and non-overlapping (IH, ID, FO, NP, CH) parameters. We (arbitrarily) construct this 7×1 vector as follows:

$$\boldsymbol{\beta} \equiv \begin{pmatrix} \text{ME} \\ \text{MA} \\ \text{NP} \\ \text{CH} \\ \text{FO} \\ \text{ID} \\ \text{IH} \end{pmatrix}.$$

Having specified $\mathbf{y}$ and $\boldsymbol{\beta}$ we now have fully specified the rules for constructing the response matrix, $\mathbf{X}$. For instance, consider a pair or rows of $\mathbf{X}$ with row index $K = 2i - 1$ and $K' = K + 1$. Equation 7 specifies the coefficients for row K while Equation 8 specifies that for row K':

$$\mathbf{X} = \begin{bmatrix} . & . & . & . & . & . & . \\ \sin(h_K)\tan(\delta_K) & -\cos(h_K)\tan(\delta_K) & \tan(\delta_K) & \sec(\delta_K) & 0 & 0 & 1 \\ \cos(h_{K'}) & \sin(h_{K'}) & 0 & 0 & \cos(h_{K'}) & 1 & 0 \\ . & . & . & . & . & . & . \end{bmatrix}.$$

The size of $\mathbf{X}$ matrix is $(2n, 7)$ whereas that of $\mathbf{y}$ is $(2n, 1)$.