LST-Binned Analysis of the EDGES Low-1 data.

Nivedita Mahesh, Steven Murray
Arizona State University

Update: 07/22/2020

1 Introduction

This report summarizes the efforts carried out in analysing the low-1 data of the EDGES system. The data was analysed in varying LST -averaged bins. This effort is the prelude to the work to be done for the dual band analysis.

Data sets: 1.) Data released with the Nature paper, 2.) Low-1 data from 2017-258 to 2018-095:.

Foreground models used:

- Linlog

$$T(\nu) = (\frac{\nu}{\nu_o})^{-\beta}(a_0 + a_2(log(\frac{\nu}{\nu_o}))^2 + a_3(log(\frac{\nu}{\nu_o}))^3 + a_4(log(\frac{\nu}{\nu_o}))^4..) \quad (1)$$

- Linpoly

$$T(\nu) = (\frac{\nu}{\nu_o})^{-\beta}(p_0 + p_1(\frac{\nu}{\nu_o}) + p_2(\frac{\nu}{\nu_o})^2 + p_3 log(\frac{\nu}{\nu_o})^3..) \quad (2)$$

- Physical

$$T(\nu) = (\frac{\nu}{\nu_o})^{-2.5}(a_0 + a_1 log(\frac{\nu}{\nu_o}) + a_2(log(\frac{\nu}{\nu_o}))^2) + e_1(\frac{\nu}{\nu_o})^{-4.5} + e_2(\frac{\nu}{\nu_o})^{-2} \quad (3)$$

2 Paper Data Analysis

The data that was released with the nature paper was binned in frequency and averaged over LST from 6hr to 18 hr. The spectra of this data is shown in figure 10.

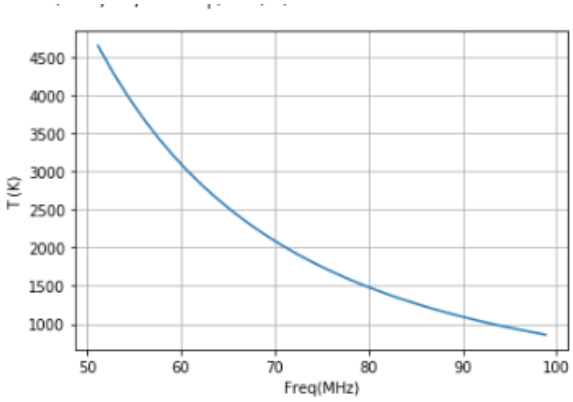

Figure 1: The sky temperature vs frequency for the low band data that was averaged from 6hr to 18 hr LST.

The data is restricted between the frequency range of 51.1 to 98.9 MHz in 125 bins.

To this data I fit different foreground models + the 21cm absorption model[bowman2018]. The estimation was done using polychord with nlives =1024 for each of the runs. The estimates of the parameters from each model is shown below.

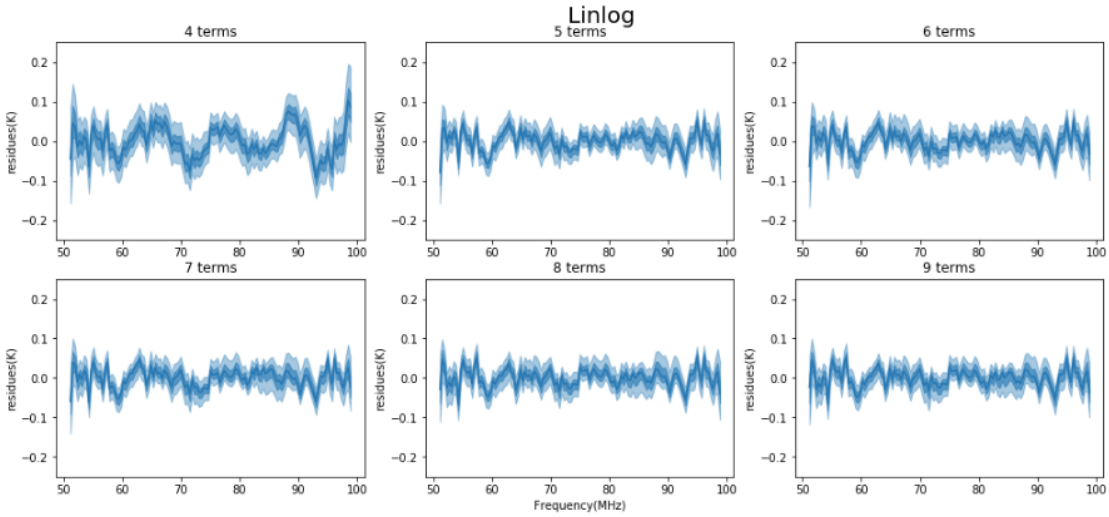

Figure 2: Residues vs frequency after fitting a **linlog model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of linlog terms used.

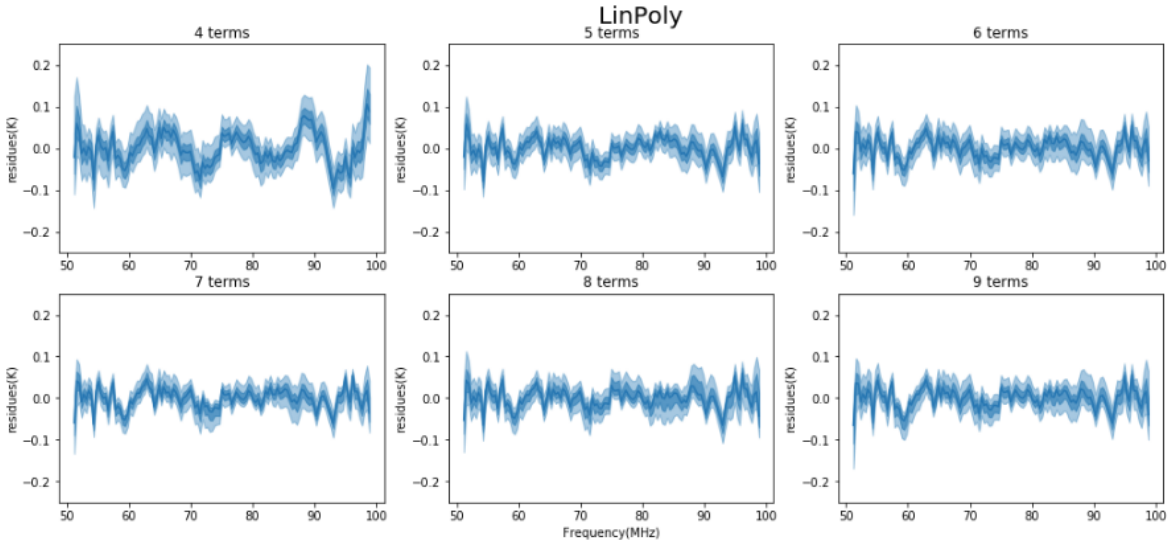

Figure 3: Residues vs frequency after fitting a **linpoly model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of linpoly terms used.

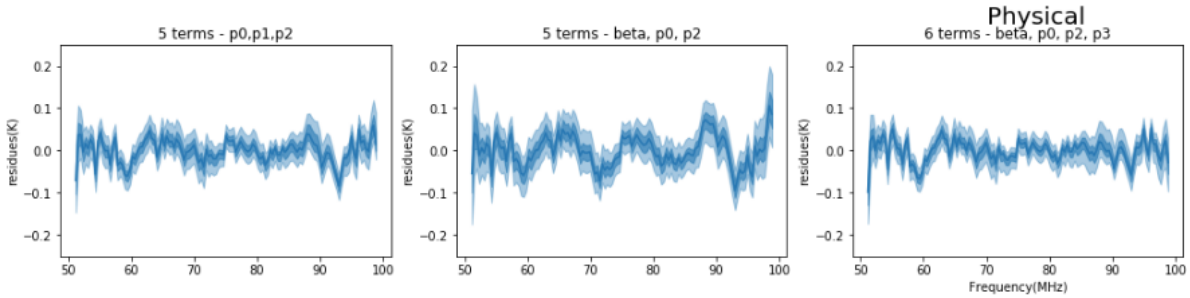

Figure 4: Residues vs frequency after fitting a **physical model** and a 21cm model to the data in figure 10. The left most plot corresponds to the equation in the paper. The middle plot fits for beta and removes the first order log term. The right most plot adds third order log term.

2.1 Residues:

The public release of the data was fit for 4 absorption parameters and varying foreground models and number of terms. The residues to the different fits are shown below:

The residues from Linlog and linpoly are seen to be similar for similar number of terms. On increasing the no.of terms more than 5 doesnt remove any additional features from the residues. These residues look comparable to the ones show in Bowman2018. The residues to the 5 term physical model (from the paper) is also similar to the other models. The 5 term physical model where p0 was replaced with beta did not result in similar residues. A third order log term had to be introduced to result in lowest residues.

There are some sharp features seen in residues(60 MHz, 93MHz) to >5 term linlog and linpoly. These could correspond to RFI. Another run of the parameter estimates will be done by flagging few of these channels.

2.2 Bayesian Evidence comparison

The Bayesian evidence of all the above runs is calculated and compared in figure 5.

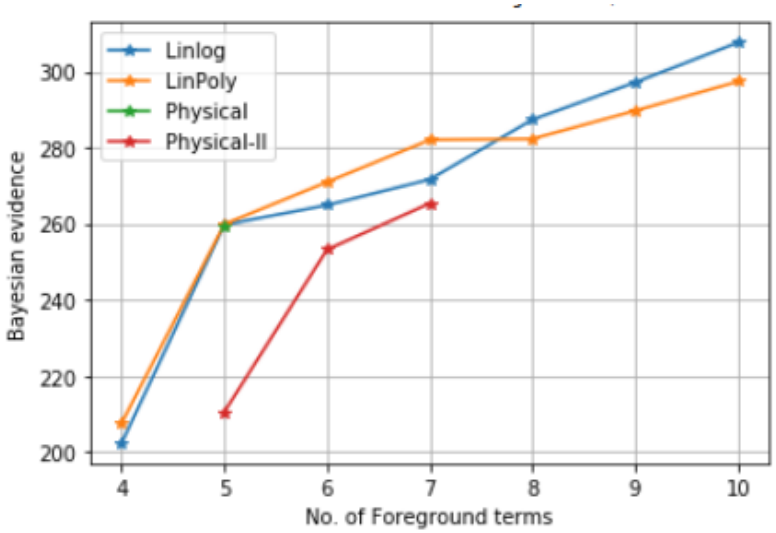

Figure 5: . Bayesian evidence vs no.of foreground terms for the 3 different models used on the data.

The Linpoly is seen to have the best evidence for 7 foreground terms. The Linlog doesn't indicate asymptotic behaviour for increasing foreground terms. The Physical model from the paper is seen to have similar evidence for 5 terms as the linlog and linpoly form. The modified Phyical model is seen to give increasing evidence for increasing number of terms.

2.3 Parameter estimation:

Foreground Model	a21 [mK]	ν_o [MHz]	τ	w [MHz]	β
Linlog5	580	78.2	6.2	18.8	-2.568
Linlog6	650	78.1	5.5	18.7	-2.568
Linlog7	590	78.1	6	18.8	-2.568
Linlog8	550	78.2	6.1	19.0	- 2.569
Linlog9	500	78.2	6.5	18.8	-2.568
LinPoly5	920	77.9	3.8	19.2	-2.65
LinPoly6	680	78.1	5.4	18.9	-4.7
LinPoly7	700	78.2	5	19.0	-2.95
LinPoly8	950	78.2	4	19.2	-2.95
LinPoly9	700	78.3	5	19.2	-3.48
Physical5	1800	79.7	1.9	22.7	
Physical5 - extend p0 prior	670	78.6	5	19.2	
Physical5 - beta,p0,p2	1750	79.5	2.0	22.7	-2.57
Physical6 - beta,p0,p2,p3	520	78.4	7.0	18.7	-2.58

3 Raw data averaged between 6 and 18hr GHA.

The raw low-1 data between days 2017-258 and 2018-095 was averaged over GHA of 6hr to 18 hr but kept in its raw resolution. The spectra of this data is over-plotted with the data in figure 10 and shown below.

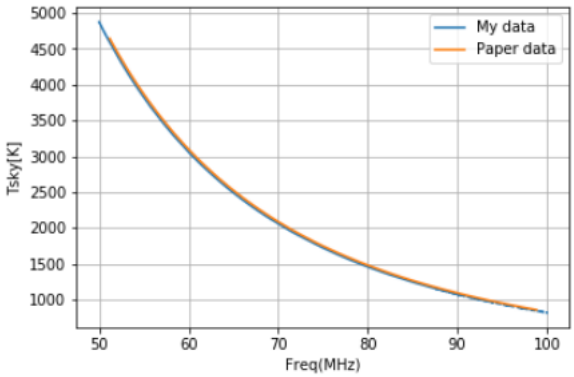

Figure 6: The sky temperature vs frequency for the low band data that was averaged from 6hr to 18 hr GHA.

The data covers the complete frequency range of 50 to 100 MHz in 8193 bins.

To this data I fit different foreground models + the 21cm absorption model[bowman2018]. The estimation was done using polychord with nlives =1024 for each of the runs. The estimates of the parameters from each model is shown below.

3.1 Residues:

The data was fit for 4 absorption parameters and varying foreground models and number of terms. The residues to the different fits are shown below:

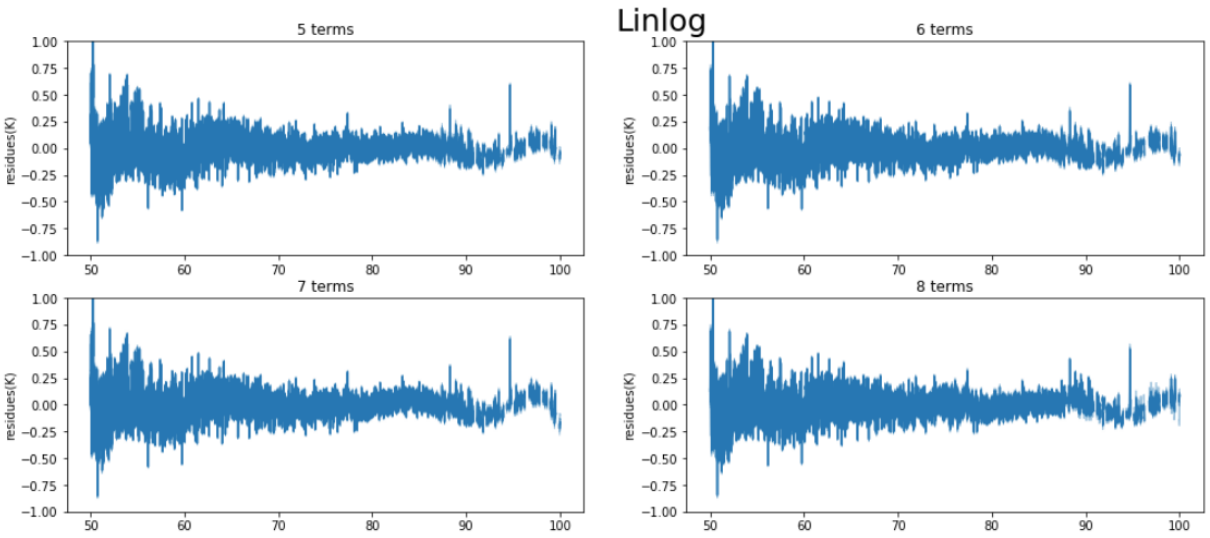

Figure 7: Residues vs frequency after fitting a **Linlog model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of Linlog terms used.

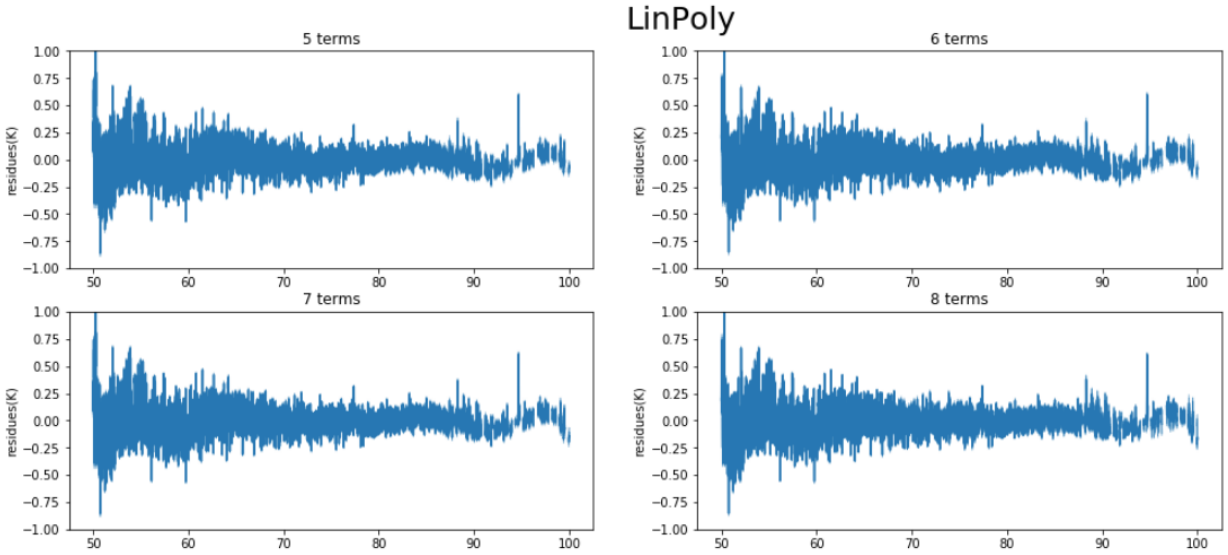

Figure 8: Residues vs frequency after fitting a **linpoly model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of linpoly terms used.

The residues from Linlog and linpoly are seen to be similar for similar number of terms. On increasing the no.of terms more than 5 doesnt remove any additional features from the residues.

RFI is seen all the residue plots for frequencies >92MHz.

3.2 Bayesian Evidence comparison

The Bayesian evidence of all the above runs is calculated and compared in figure 13.

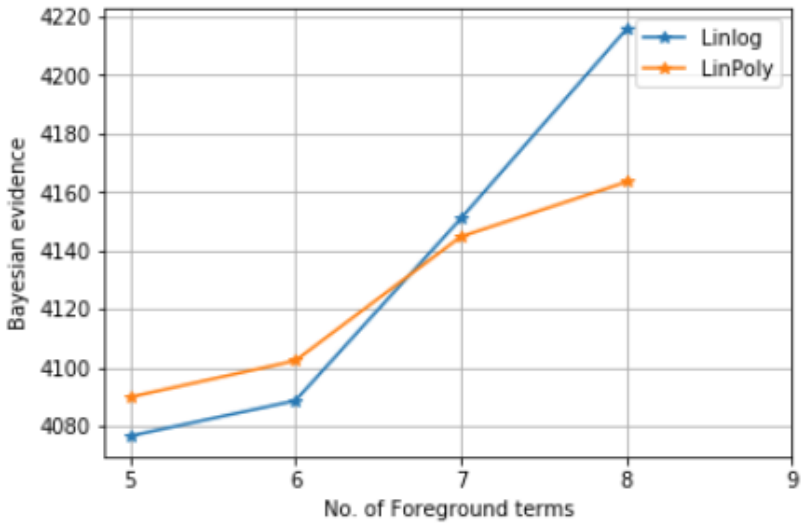

Figure 9: . Bayesian evidence vs no.of foreground terms for the 2 different models used on the data.

3.3 Parameter estimation:

Foreground Model	**a21 [mK]**	ν_o **[MHz]**	τ	**w [MHz]**	β
Linlog5	400	78.4	13	19.1	-2.580
Linlog6	450	78.3	12	19.0	-2.580
Linlog7	2200	78.4	2.3	19.3	-2.575
Linlog8	472	72.3	1.12	8.95	-2.58
LinPoly5	470	78.2	10	18.8	-3.25
LinPoly6	610	78.2	6.2	19.0	-2.72
LinPoly7	950	78.5	4.2	19.3	-3.15
LinPoly8	1500	78.8	3.0	19.6	-4.35

4 Raw data averaged between 12 and 17hr GHA.

The raw low-1 data between days 2017-258 and 2018-095 was averaged over 5 hours of GHA from 12 to 17hr (this is much less data compared to 12hr in the previous section) but kept in its raw frequency resolution.

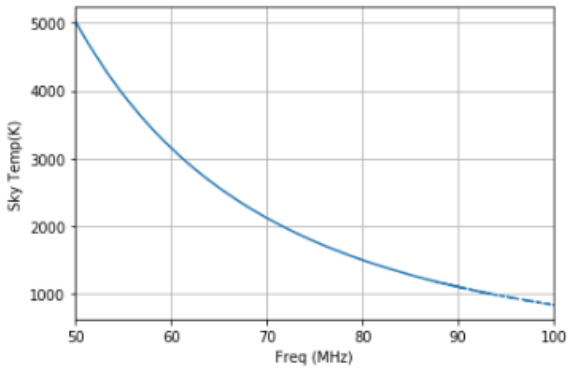

Figure 10: The sky temperature vs frequency for the low band data that was averaged from 12hr to 17 hr GHA.

The data covers the complete frequency range of 50 to 100 MHz in 8193 bins.

To this data I fit different foreground models + the 21cm absorption model[bowman2018]. The estimation was done using polychord with nlives =1024 for each of the runs. The estimates of the parameters from each model is shown below.

4.1 Residues:

The data was fit for 4 absorption parameters and varying foreground models and number of terms. The residues to the different fits are shown below:

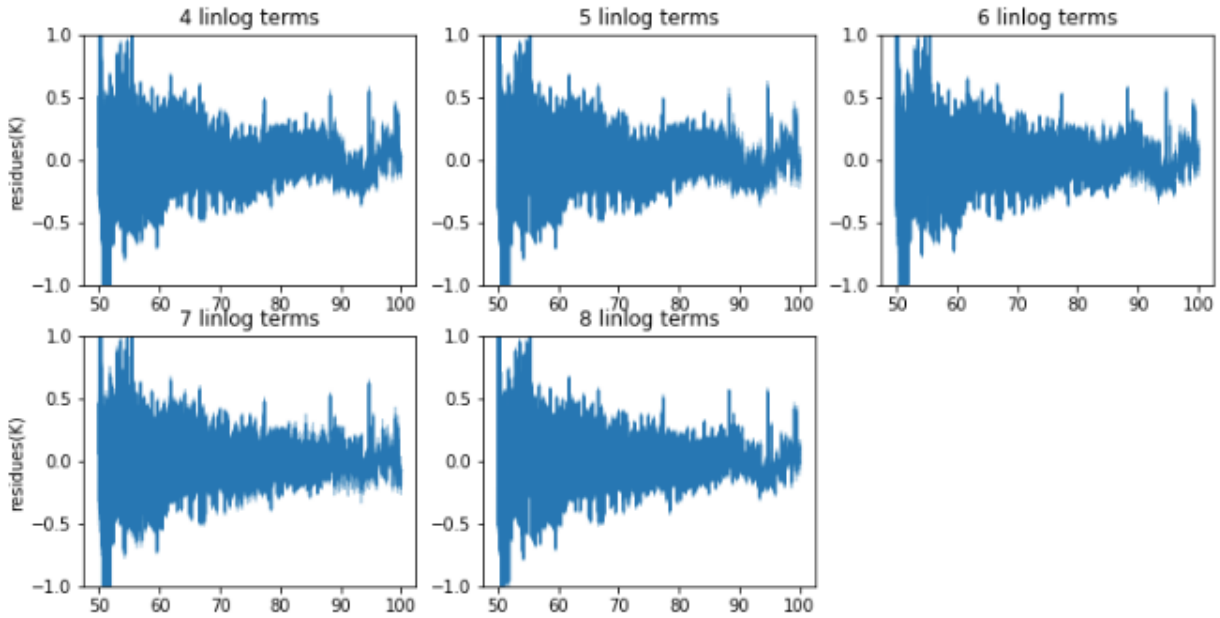

Figure 11: Residues vs frequency after fitting a **Linlog model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of Linlog terms used.

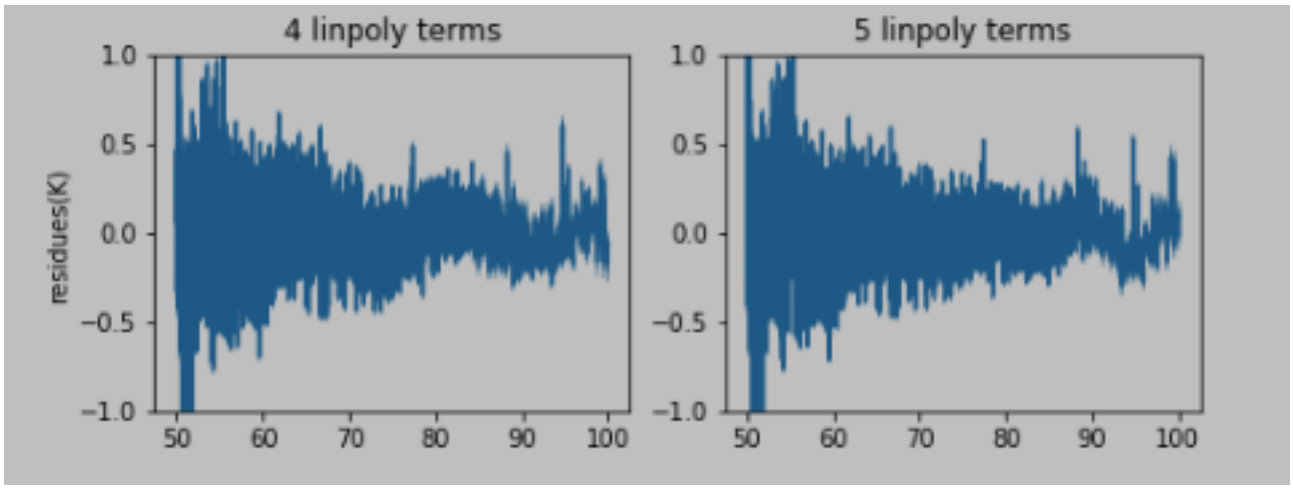

Figure 12: Residues vs frequency after fitting a **Linpoly model** and a 21cm model to the data in figure 10. Each subplot corresponds to different number of Linpoly terms used.

The residues to the 4 term and 5 term model fits looks similar for Linlog. The residues to the 7 term fit is seen to have the least structure.

RFI is seen all the residue plots for frequencies >92MHz.

4.2 Bayesian Evidence comparison

The Bayesian evidence of all the above runs is calculated and compared in figure 13.

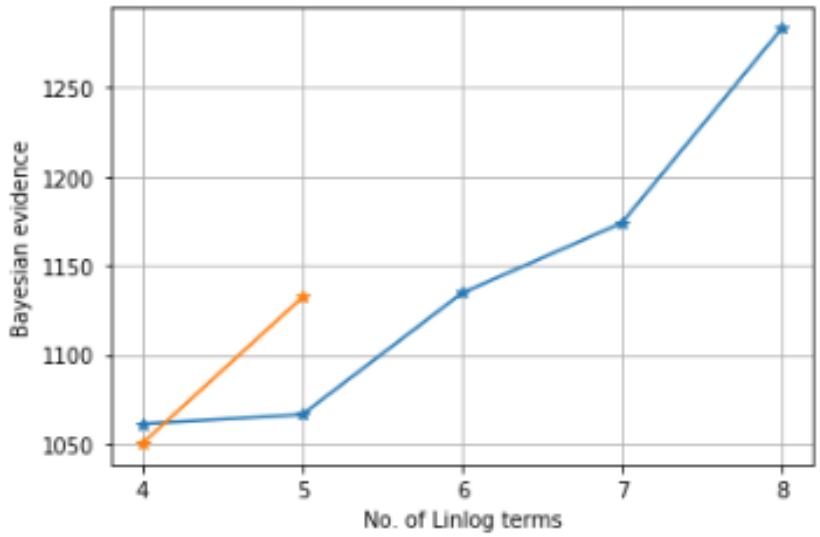

Figure 13: . Bayesian evidence vs no.of foreground terms of the linlog and linpoly foreground models used on the data.

4.3 Parameter estimation:

Foreground Model	**a21 [mK]**	ν_o **[MHz]**	τ	**w [MHz]**	β
Linlog4	820	78.6	5.8	20.8	-2.594
Linlog5	720	78.25	7	20.3	-2.595
Linlog6	1600	81.0	3.2	24.5	-2.592
Linlog7**	100	71.4	2.3	8	-2.6
Linlog8	650	72.4	1.12	10	-2.59
LinPoly4	637	77.8	7.3	19.57	-2.4
LinPoly5	1800	81.0	3.25	24.9	-3.1
LinPoly6					
LinPoly7					

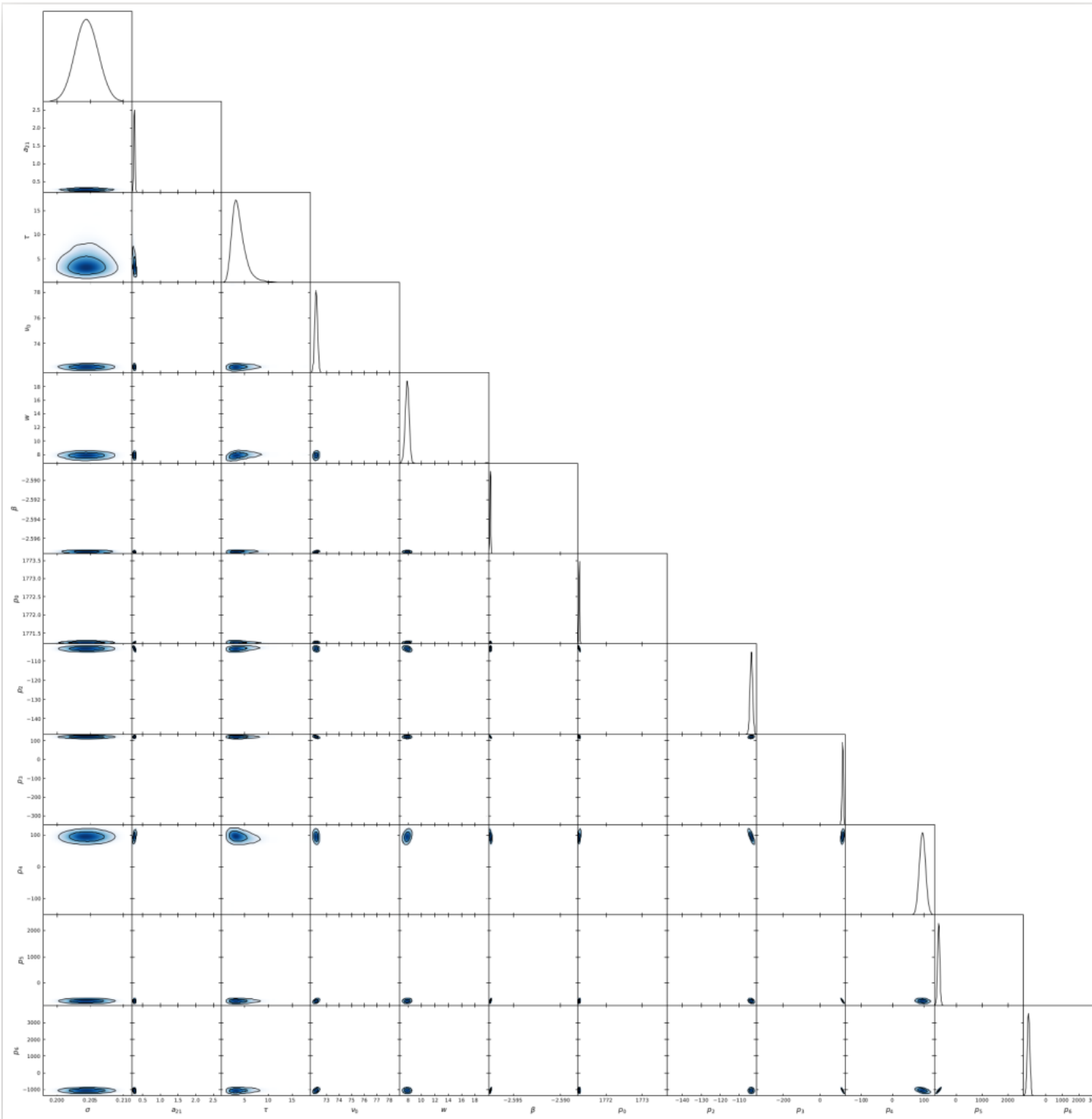

Figure 14: **Raw data- 12 to 17hr: 7 - Linlog.** Parameter estimation for 7 linlog terms and the four 21 cm model terms.

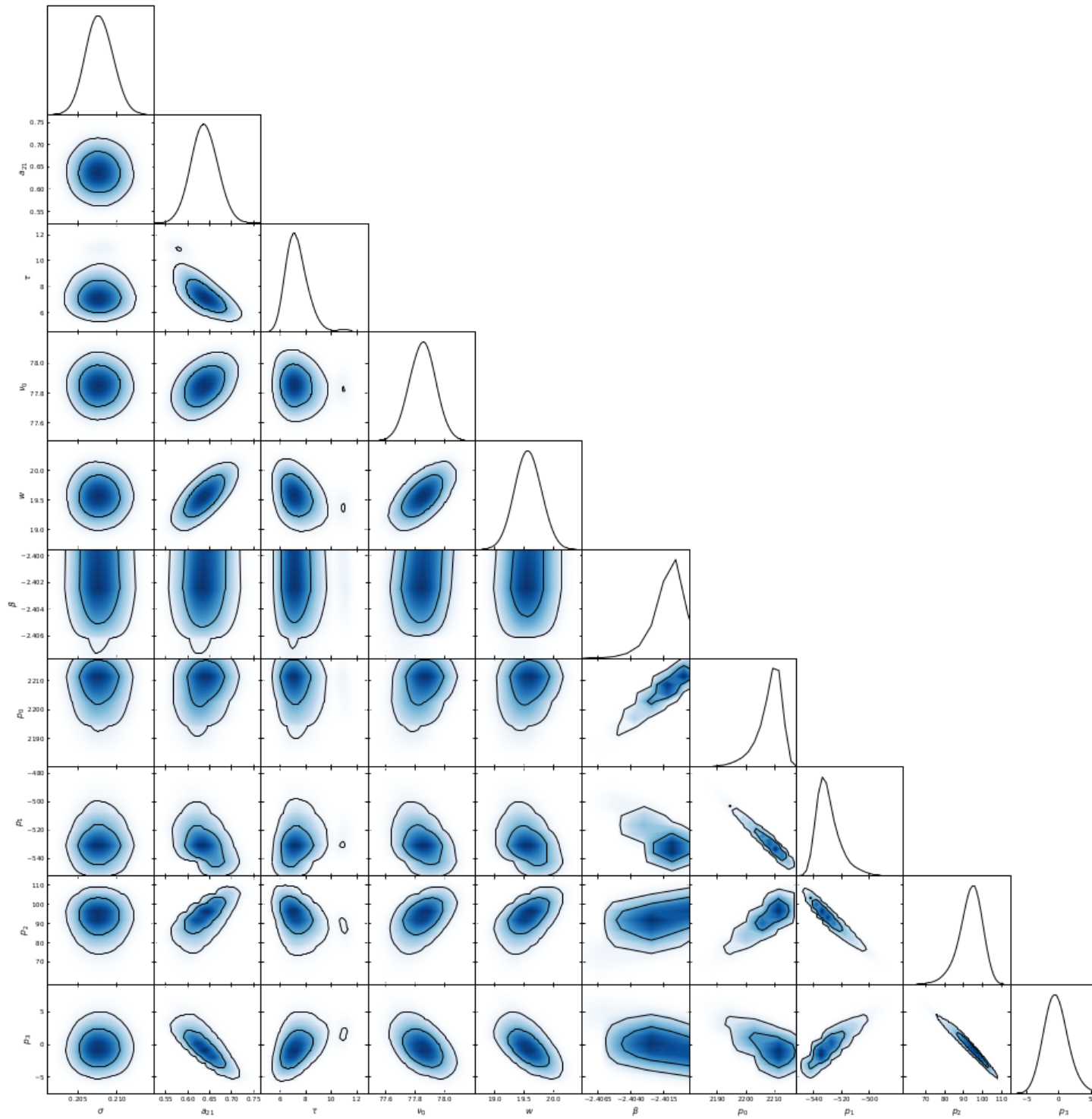

Figure 15: **Raw data- 12 to 17hr: 4 - Linpoly.** Parameter estimation for 4 linpoly terms and the four 21 cm model terms.

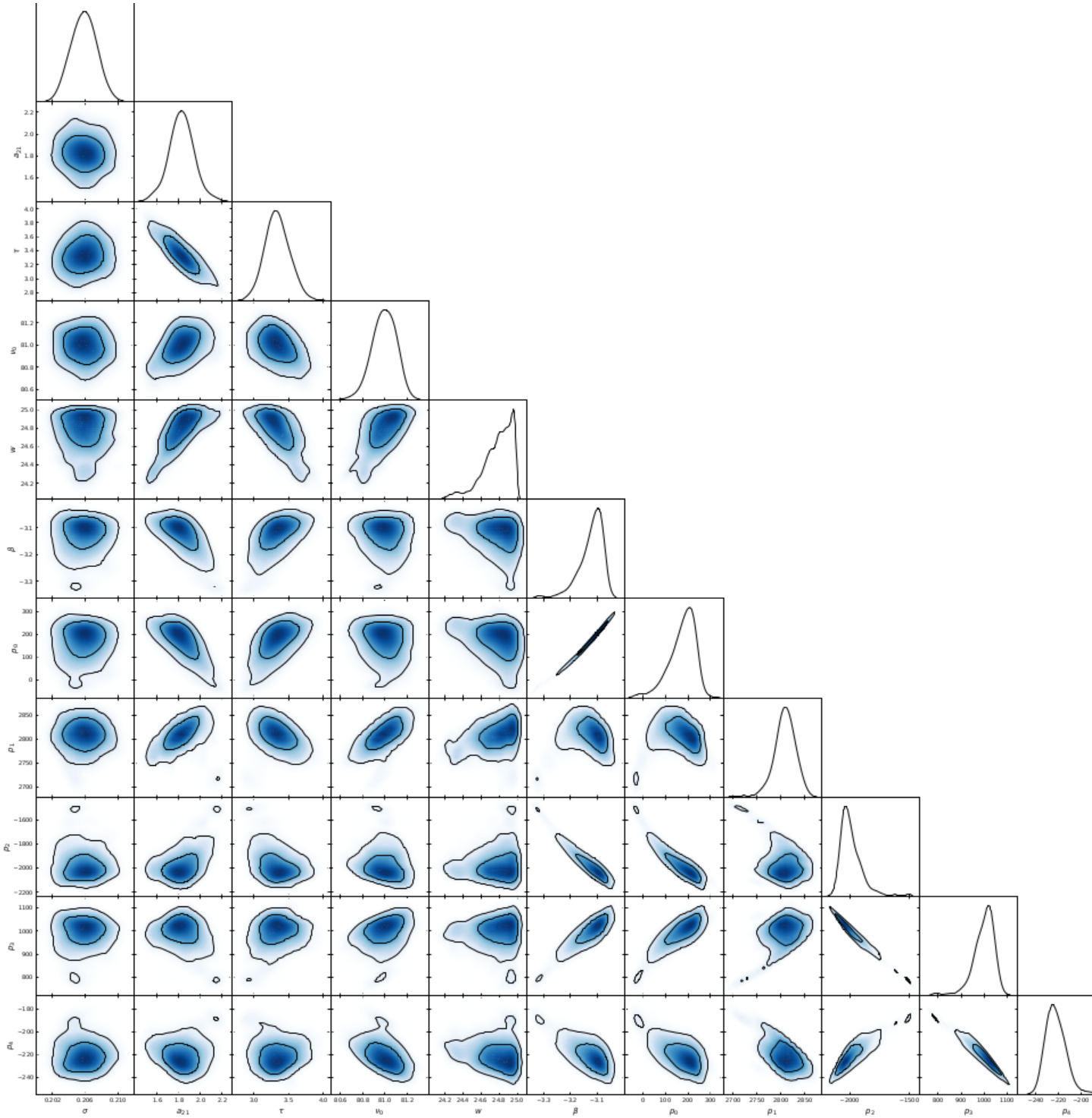

Figure 16: **Raw data- 12 to 17hr: 5 - Linpoly.** Parameter estimation for 5 linpoly terms and the four 21 cm model terms.

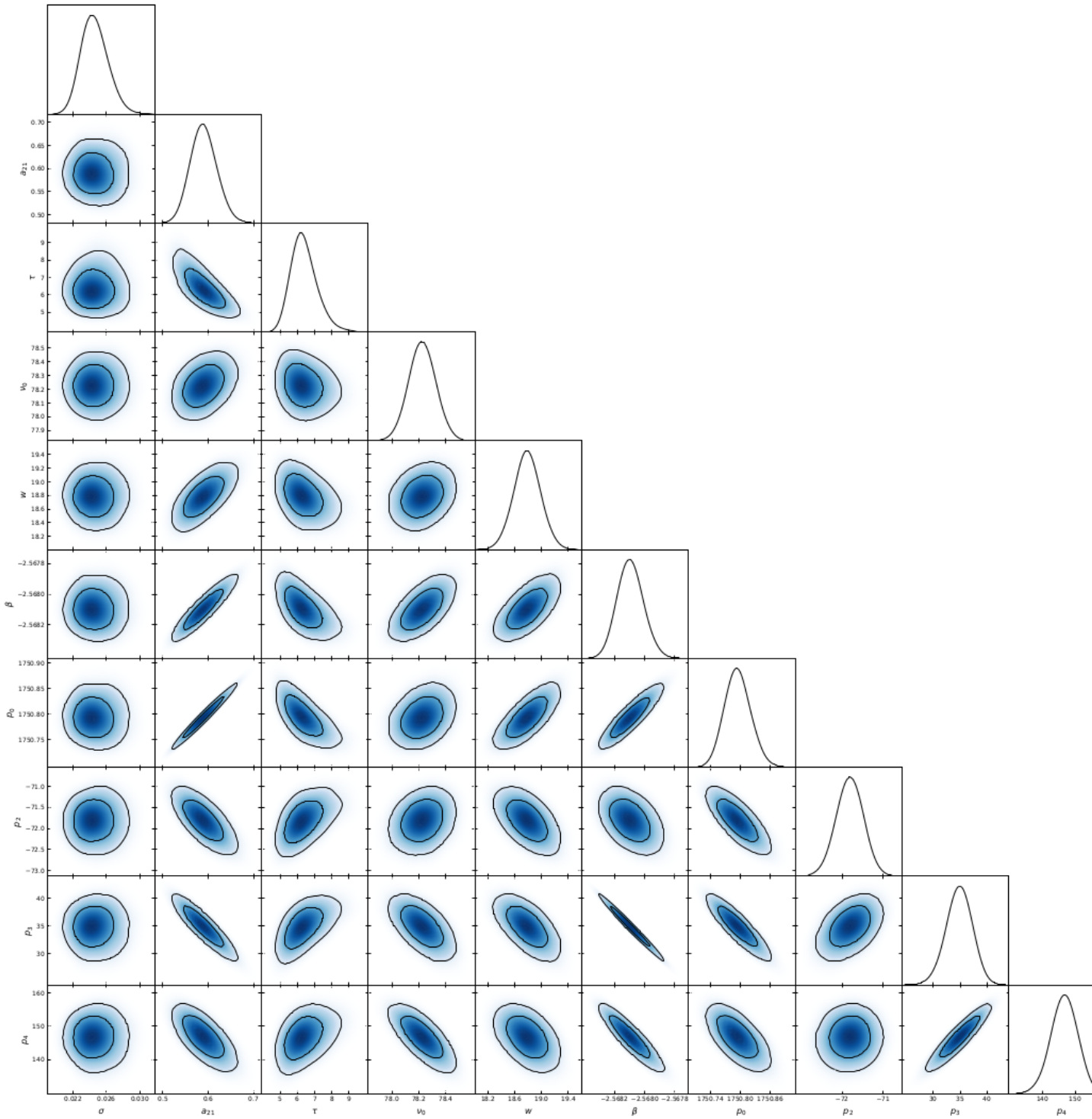

Figure 17: **Paper data: 5- Linlog.** Parameter estimation for 5 linlog terms and the four 21 cm model terms. Reasonable estimates of the foreground terms and the 21cm model parameters are in accordance with the paper.

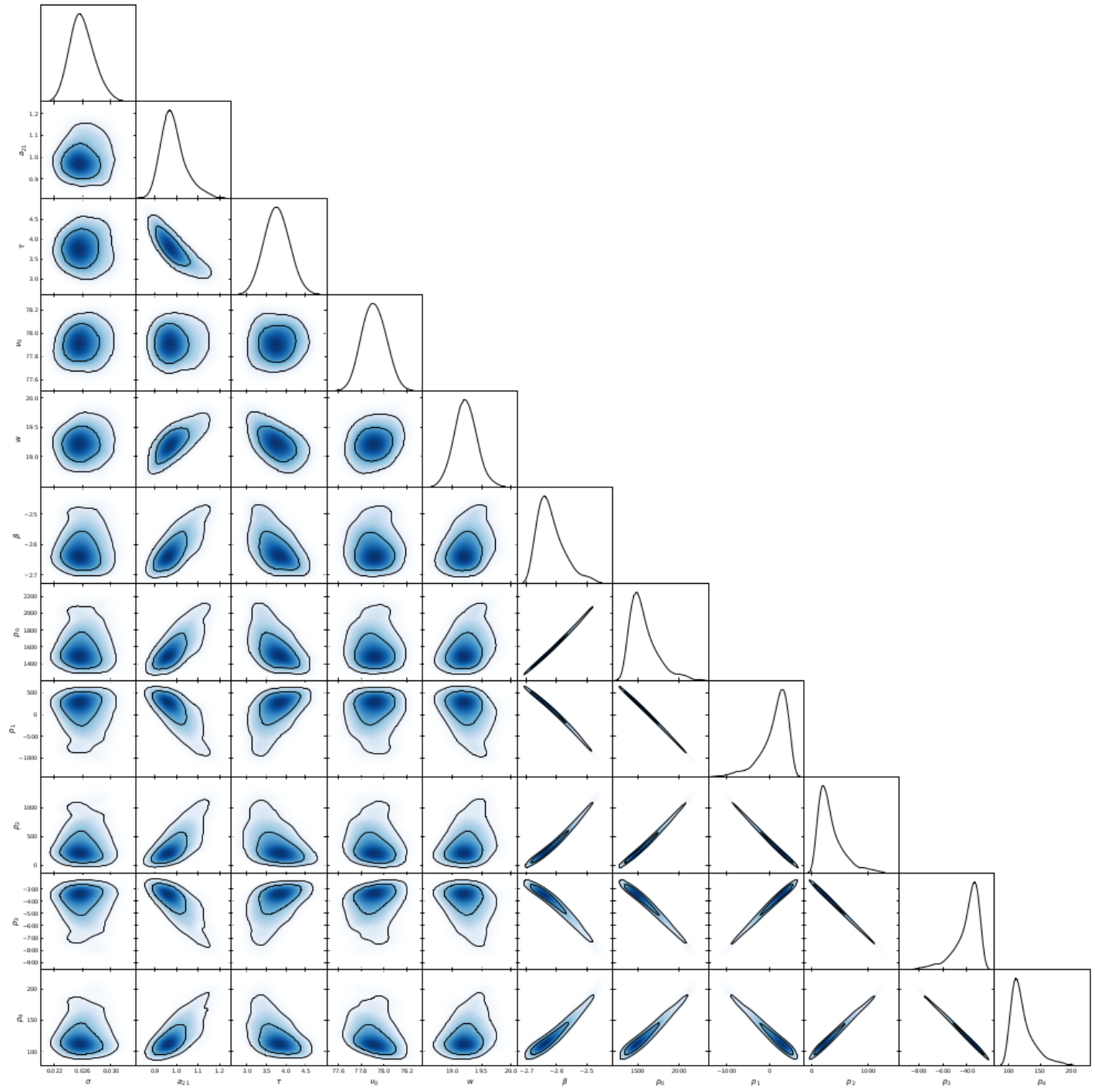

Figure 18: **Paper data: 5- LinPoly.** Parameter estimation for 5 poly terms and the four 21 cm model terms. The foreground polynomial coefficients are highly correlated with each other, while the 21 cm model parameters are largely uncorrelated except for the profile amplitude (A) and flattening (τ). This trend was seen in the nature paper figure 10.

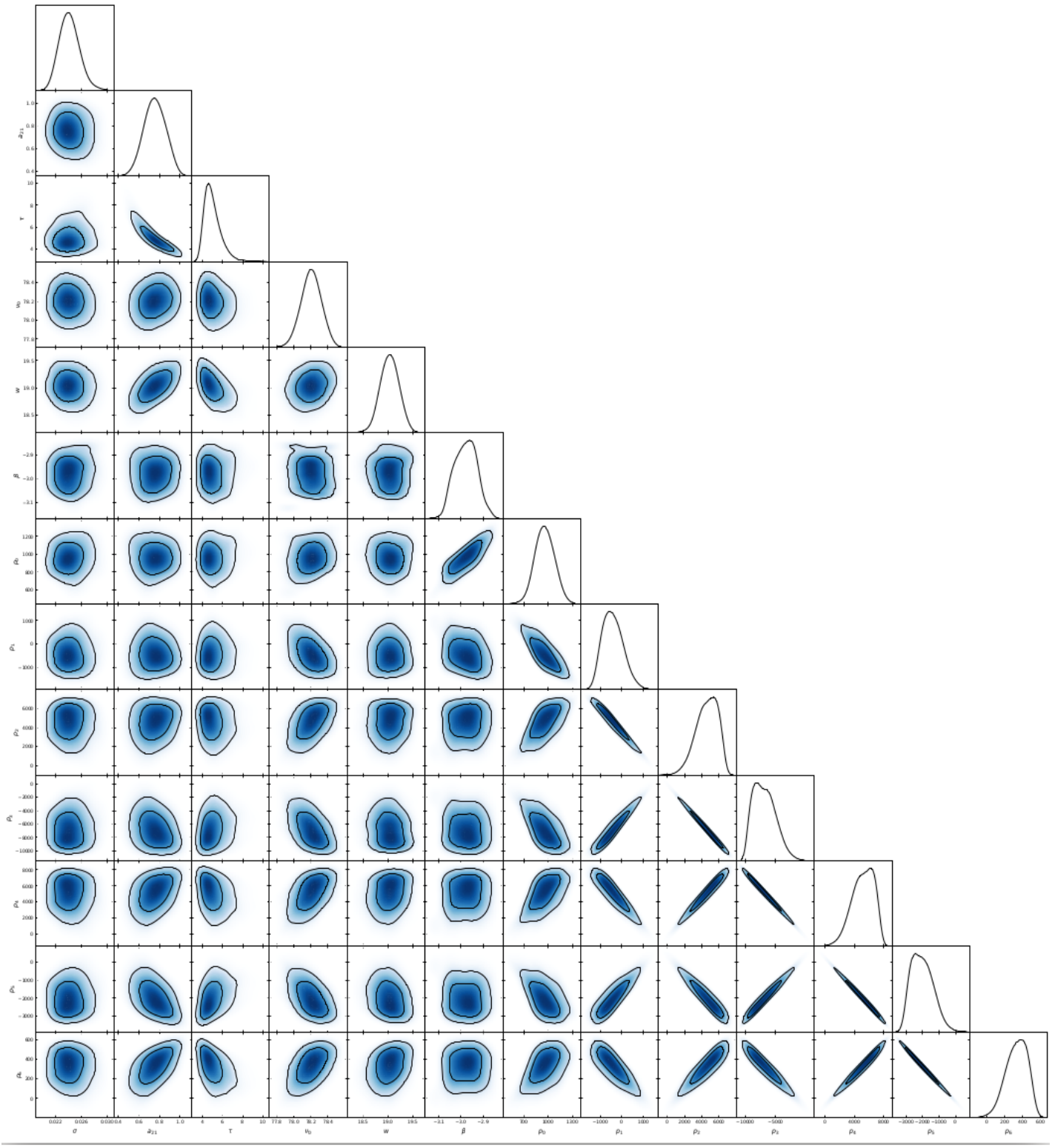

Figure 19: **Paper data: 7- LinPoly.** Parameter estimation for 7 poly terms and the four 21 cm model terms .Same trend as in figure 18. But the Amplitude of the feature agrees better with was reported in the paper. But beta is higher in value.

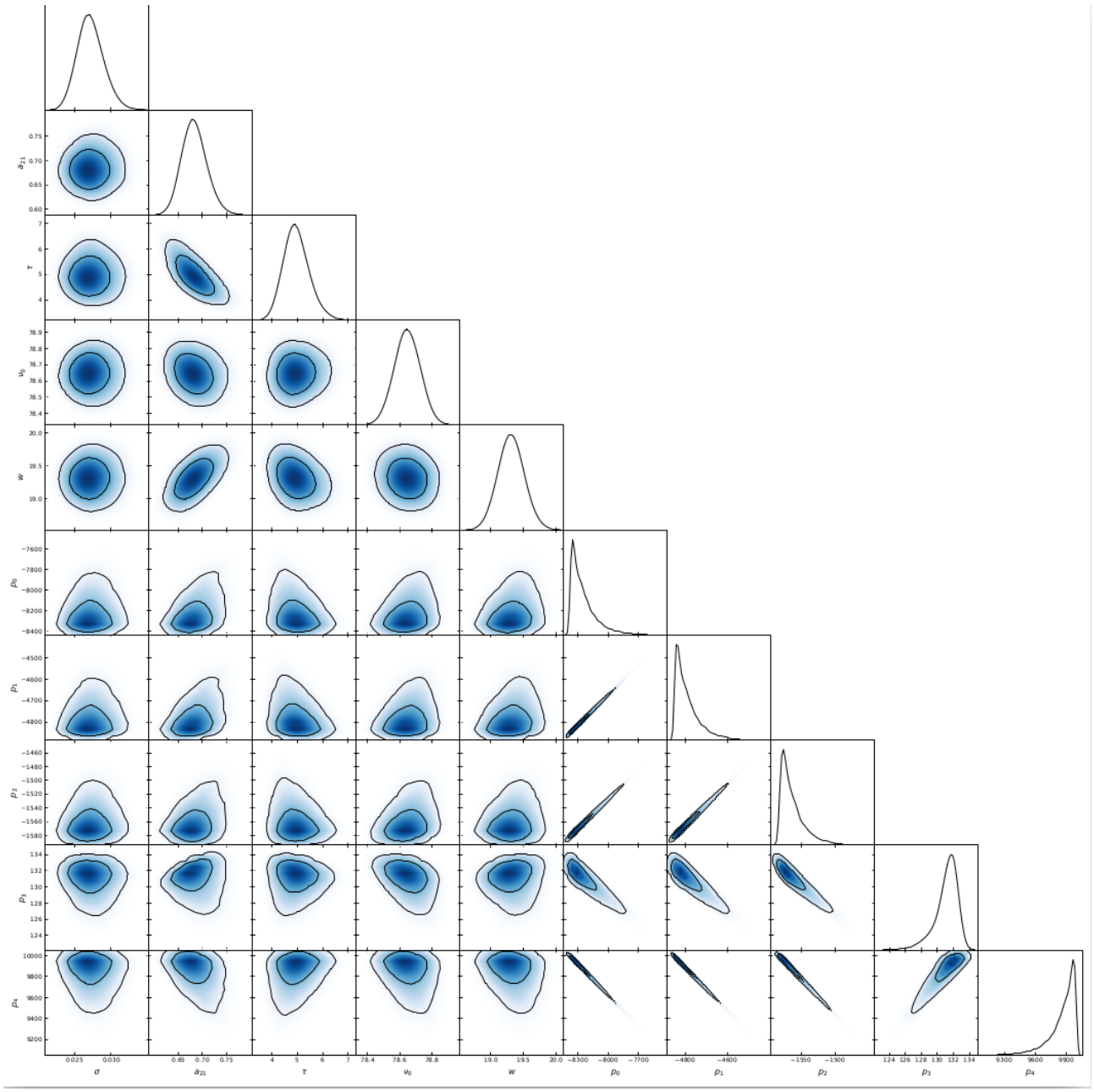

Figure 20: **Paper data: 5- physical.** Parameter estimation for 5 physical terms and the four 21 cm model terms. This is similar to **??** except that the prior on p0 was changed from 100 to 10,000 to +/-10,000.

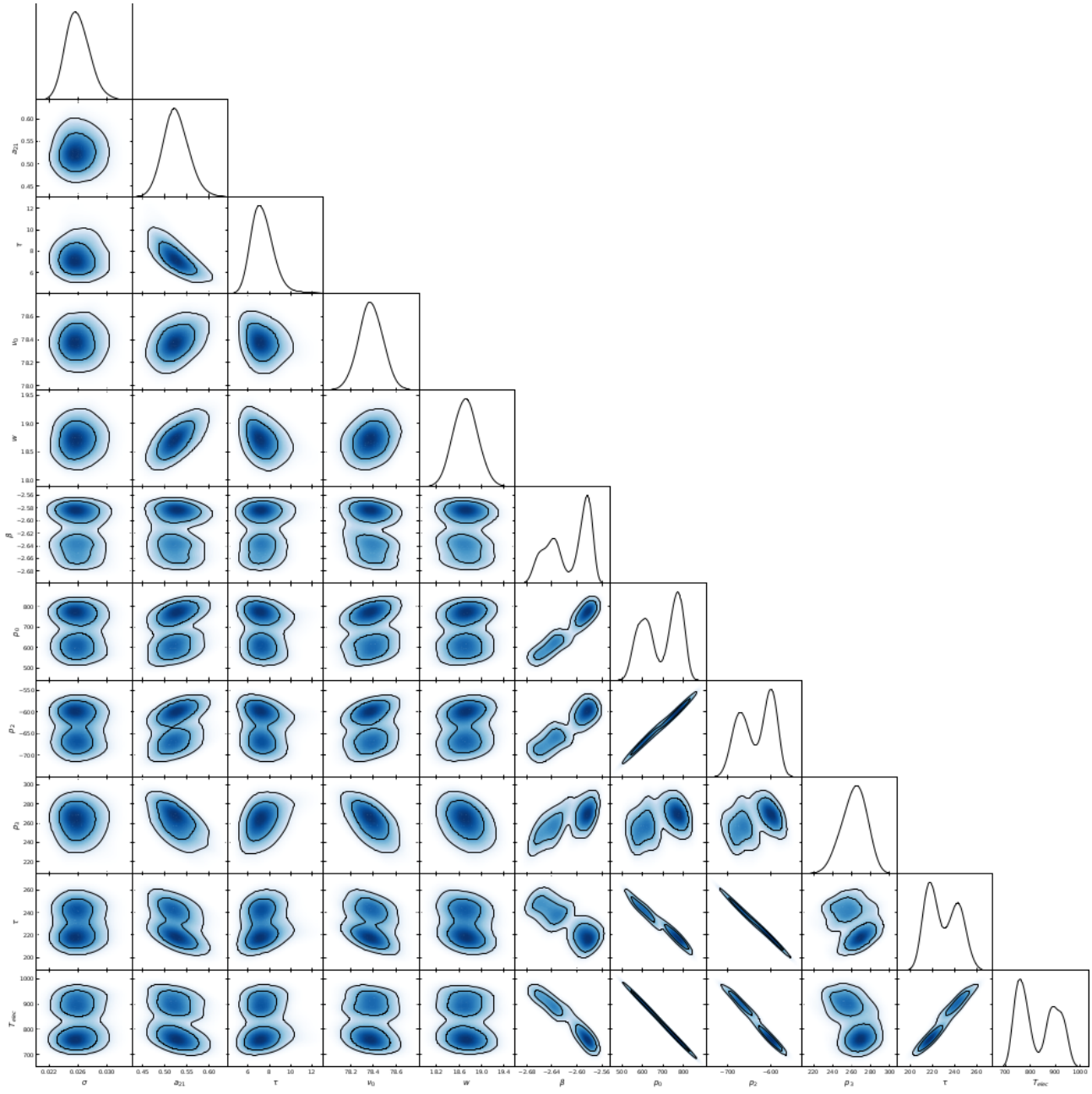

Figure 21: **Paper data: 6- Physical (modified).** Parameter estimation for 5 physical terms and the four 21 cm model terms. This is similar to 20 except that the term $p_0 * log(\frac{\nu}{\nu_o})$ was removed and $p_3 * (log(\frac{\nu}{\nu_o}))^3$ was added.

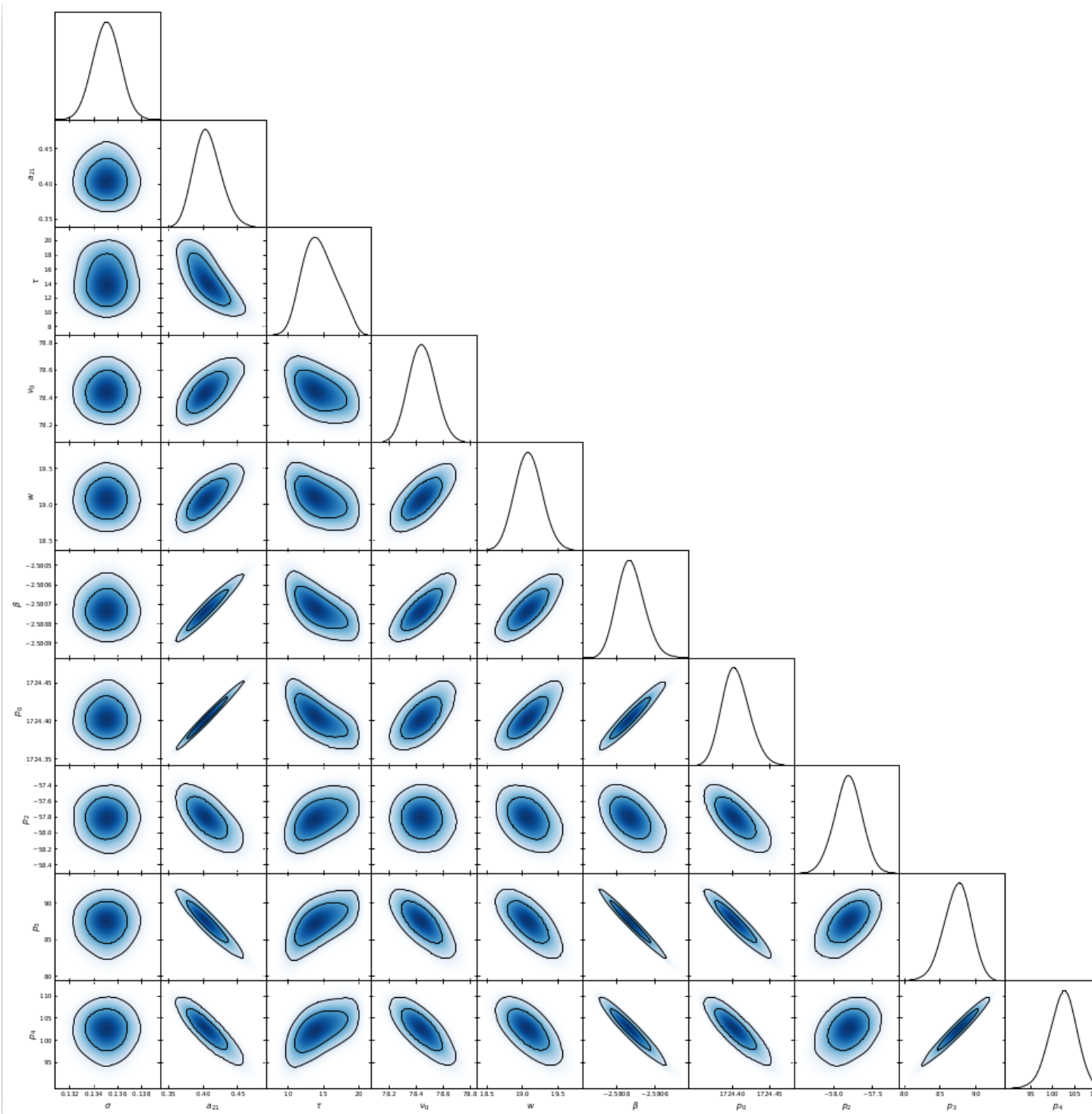

Figure 22: **Independent data: 5- Linlog.** Parameter estimation for 5 linlog terms and the four 21 cm model terms. Reasonable estimates of the foreground terms and the 21cm model parameters are in accordance with the paper. Except that the estimated τ value is high.

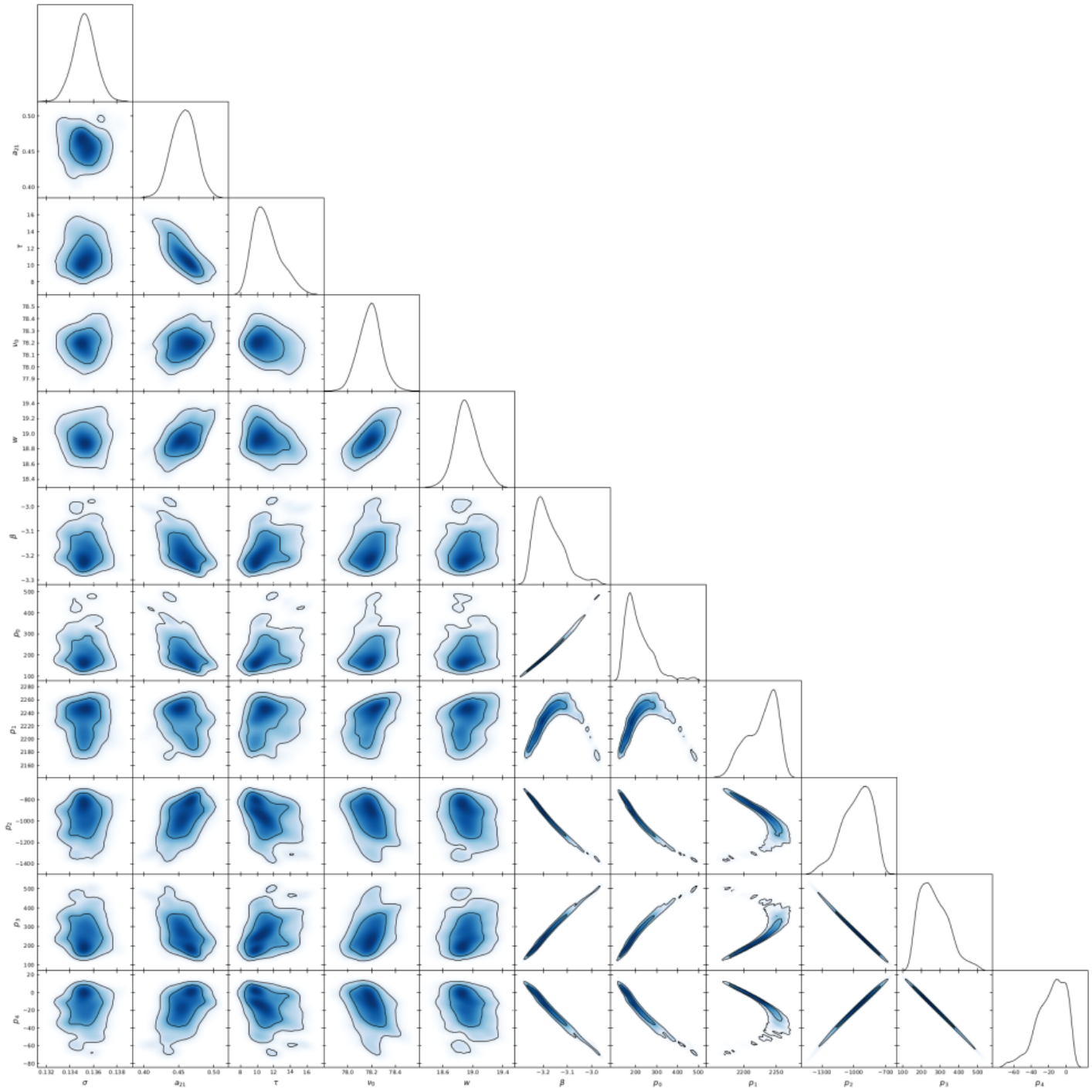

Figure 23: **Independent data: 5- Linpoly.** Parameter estimation for 5 poly terms and the four 21 cm model terms. Though the mean of the estimates seem reasonable the distribution indicates that the run did not go so well.

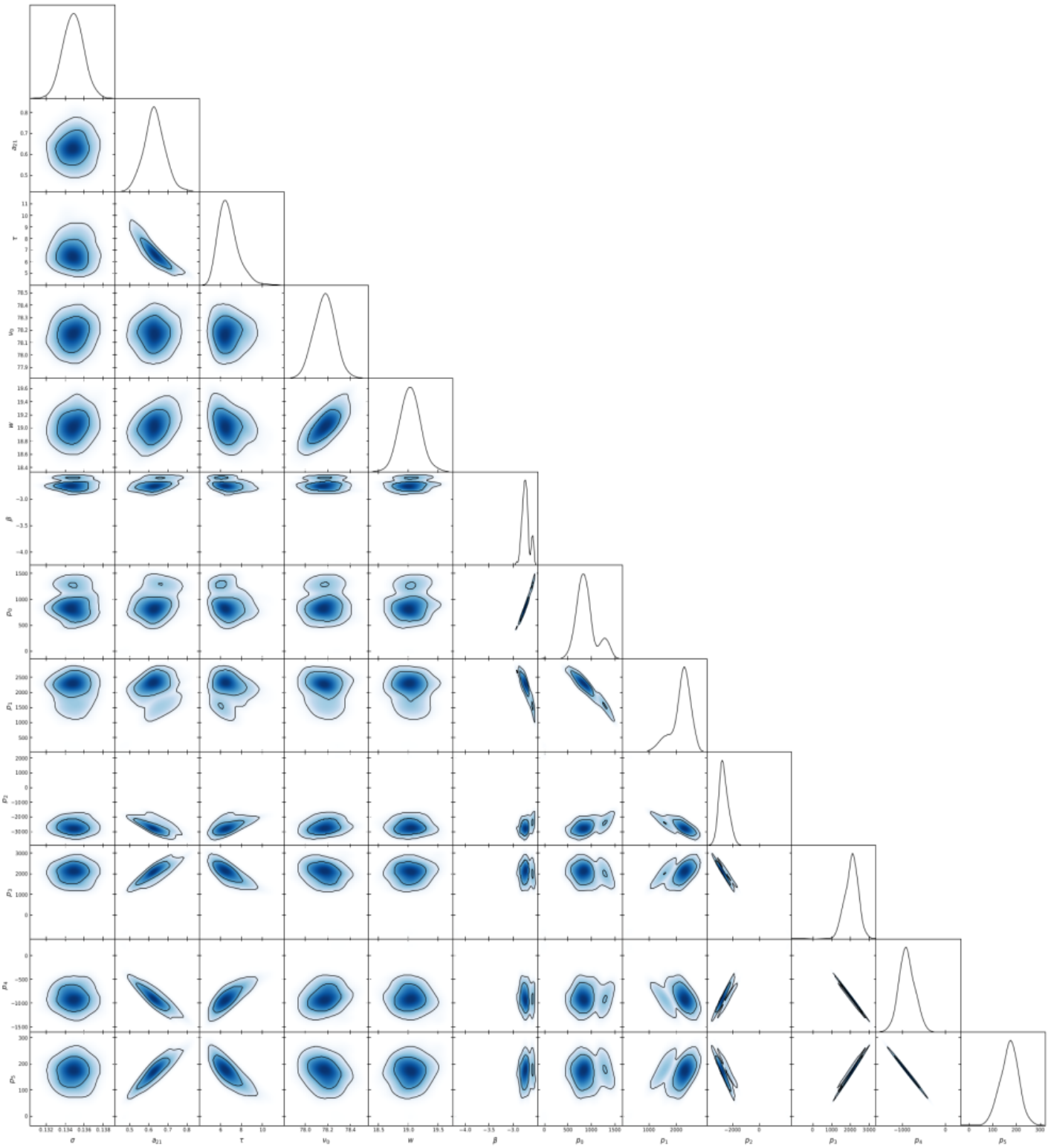

Figure 24: **Independent data: 6- Linpoly.** Parameter estimation for 6 poly terms and the four 21 cm model terms. The estimates are more reasonable that the previous run. But the foreground parameters are seen to have a bi-modal distribution.