APPENDIX A

CHARACTERIZING THE ZONAL STRUCTURE WITH AN HARMONIC FIT

Having selected a subset of density measurements in some way, I fit a function to it to quantify the variation in density with longitude. A linear combination of harmonic functions is a natural choice as the fitted function. I fit:

$$\rho = a_0 + a_1 \cos(\lambda) + b_1 \sin(\lambda) + \cdots + a_k \cos(k\lambda) + b_k \sin(k\lambda) \tag{A.1}$$

Where ρ is density, λ is east longitude, and the a_i and b_i are model parameters. This is a wave-k model. There must be enough data points in the subset to constrain the fit, so there must be at least $2k+1$ data points. It is not possible to fit fewer data points with a wave-k model. If k is smaller than an optimal value, then the model does not capture all the information present in the data. If k is larger than an optimal value, then the model parameters are poorly constrained. I shall discuss the optimal value of k later in this section.

In general, when fitting a model to data, one should incorporate measurement uncertainties directly into the fitting procedure. However, as I now outline, that is not appropriate here. The uncertainties quoted for each density measurement in the PDS archive are not formal uncertainties in the strictest sense (Tolson et al., 2000; Keating et al., 2001a). As described in Tolson et al. (2000), these uncertainties include contributions from an empirical "quality indicator". However, one might still consider using them as if they were formal 1 σ uncertainties. There is a problem with this approach. The quoted uncertainties each refer to the uncertainty of an individual measurement. Should a later measurement at exactly the same longitude obtain the same density within the quoted uncertainty? Not necessarily. The conditions within the experiment, the martian upper atmosphere, are beyond our control. They vary, due to weather, between measurements. The quoted uncertainties are very relevant to each individual measurement. They are less relevant to the repeatability of each measurement at some later time. The repeatability of measurements on a timescale of one sol is addressed in Section 2.4. I find that sol-to-sol variability is five times greater than the few percent measurement uncertainties at 130 km altitude.

The measurement uncertainties are simply inappropriate for direct incorporation into the fitting procedure in the manner of page 162 of Bevington (1969). The appropriate uncertainties are those between repeat measurements taken over the duration of the experiment. These uncertainties cannot be calculated *a priori* using only knowledge of the measurement instrument and measurements from a single aerobraking pass. I must use the mean square error between the measurements and model predictions as a proxy for the uncertainty in the repeatability of each measurement following Equation 8.29 of Bevington (1969). The individual measurement uncertainties do not become comparable to the sol-to-sol variabilities until altitudes greater than 150 km. Since measurements at altitudes above 150 km are rarely used in this chapter, I neglect the measurement uncertainties throughout this chapter.

With the above assumption about dealing with uncertainties, m density measurements, and n model parameters, the basic model follows Equation 7.18 of Neter and Wasserman (1974):

$$\underline{Y} = \underline{\underline{X}}\,\underline{p} + \underline{\epsilon} \tag{A.2}$$

Where $\underline{Y}$ is an m-element vector of density observations, $\underline{\underline{X}}$ is an m-by-n

matrix, $\underline{p}$ is an n-element vector of model parameters, and $\underline{\epsilon}$ is an n-element vector of uncorrelated random variables of mean 0 and standard deviation σ.

Using the model outlined in Equation A.1:

$$\begin{aligned} X_{i0} &= 1 \\ X_{i1} &= \cos(\lambda_i) \\ X_{i2} &= \sin(\lambda_i) \\ &\vdots \\ X_{in-2} &= \cos\left(\frac{(n-1)\lambda_i}{2}\right) \\ X_{in-1} &= \sin\left(\frac{(n-1)\lambda_i}{2}\right) \end{aligned} \tag{A.3}$$

$$\begin{aligned} p_0 &= a_0 \\ p_1 &= a_1 \\ p_2 &= b_1 \\ &\vdots \\ p_{n-2} &= a_{\frac{n-1}{2}} \\ p_{n-1} &= b_{\frac{n-1}{2}} \end{aligned} \tag{A.4}$$

The index i labels an individual density measurement. The n variables $(1, \cos(\lambda), \sin(\lambda), \ldots)$ being fitted are linearly independent. The least squares solution for the model parameters, $\underline{p}$, follows Equation 7.21 of Neter and Wasserman (1974):

$$\underline{p} = \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \underline{\underline{X}}^T \underline{Y} \tag{A.5}$$

Where a superscript T indicates the transpose of a matrix and the superscript -1 indicates the inverse of a matrix. Note that the constant density term, a_0 or p_0, is not formally identical to the mean density. Hence I refer to it as the constant density term rather than as the zonal mean term. However, the two are usually similar.

The model predictions, $\underline{\hat{Y}}$, for atmospheric densities at the longitudes at which measurements have been made follow Equation 7.23 of Neter and Wasserman (1974):

$$\underline{\hat{Y}} = \underline{\underline{X}}\, \underline{p} \tag{A.6}$$

The covariance matrix for $\underline{p}$ follows Equation 7.39 of Neter and Wasserman (1974):

$$\underline{\underline{Cov}}\left(\underline{p}\right) = \frac{\left(\underline{\hat{Y}} - \underline{Y}\right)^T \left(\underline{\hat{Y}} - \underline{Y}\right)}{m - n} \left(\underline{\underline{X}}^T \underline{\underline{X}}\right)^{-1} \tag{A.7}$$

1 σ uncertainties in the model parameters are given by:

$$\sigma_{p_j} = Cov_{jj} \tag{A.8}$$

Using sine and cosine terms is useful for forming a linear model. However, it is more useful to interpret paired sine and cosine terms as being a single sinusoid with an amplitude and a phase. I define the phase of a given harmonic as the longitude of its first peak east of 0°. Thus the phase of the wave-q harmonic must lie between 0° and 360°E/q. q labels a certain harmonic.

$$a_q \,\cos\left(q\lambda\right) + b_q \,\sin\left(q\lambda\right) = R_q \,\cos\left(q\left[\lambda - \alpha_q\right]\right) \tag{A.9}$$

Where R_q is the amplitude of the qth harmonic and $q\alpha_q$ is the phase of the qth harmonic. Trigonometry yields the following solutions for R_q and α_q in terms of a_q and b_q.

$$R_q = \left(a_q^2 + b_q^2\right)^{1/2} \tag{A.10}$$

$$\alpha_q = \frac{1}{q}\tan^{-1}\left(b_q/a_q\right) \tag{A.11}$$

Using the usual formula, Equation 4.9 of Bevington (1969), for transforming uncertainties of uncorrelated variables:

$$\sigma_{R_q}^2 = \sigma_{a_q}^2\left(\frac{\partial R_q}{\partial a_q}\right)^2 + \sigma_{b_q}^2\left(\frac{\partial R_q}{\partial b_q}\right)^2 \tag{A.12}$$

$$\sigma_{\alpha_q}^2 = \sigma_{a_q}^2\left(\frac{\partial \alpha_q}{\partial a_q}\right)^2 + \sigma_{b_q}^2\left(\frac{\partial \alpha_q}{\partial b_q}\right)^2 \tag{A.13}$$

$$\sigma_{R_q}^2 = \sigma_{a_q}^2\left(\frac{a_q}{R_q}\right)^2 + \sigma_{b_q}^2\left(\frac{b_q}{R_q}\right)^2 \tag{A.14}$$

$$\sigma_{\alpha_q}^2 = \frac{b_q^2\sigma_{a_q}^2 + a_q^2\sigma_{b_q}^2}{q^2\left(a_q^2 + b_q^2\right)^2} \tag{A.15}$$

A phase of $3° \pm 15°$ is a perfectly reasonable and meaningful result. An amplitude of 3 kg km^{-3} $\pm$ 15 kg km^{-3} means that the harmonic is not present in any statistically significant sense and that its corresponding phase is meaningless.

I generally normalized the zonally-varying terms in each wavefit (R_q) by their constant density term (R_0). This facilitates a comparison of the strength of the zonal structure between different seasons or altitudes. I use the symbol R_q^N to label normalized harmonic amplitudes.

$$R_q^N = \frac{R_q}{R_0} \tag{A.16}$$

$$\sigma^2_{R_q^N} = \sigma^2_{R_q} \left(\frac{\partial R_q^N}{\partial R_q} \right)^2 + \sigma^2_{R_0} \left(\frac{\partial R_q^N}{\partial R_0} \right)^2 \tag{A.17}$$

$$\sigma^2_{R_q^N} = \frac{R_q^2 \, \sigma^2_{R_0}}{R_0^4} + \frac{\sigma^2_{R_q}}{R_0^2} \tag{A.18}$$

A 1 σ uncertainty about the fitted function can be calculated at any given longitude following Equation 7.54a of Neter and Wasserman (1974):

$$\sigma^2_{fit} = \left(1 + \underline{X}^T_{new} \left(\underline{\underline{X}}^T \underline{\underline{X}} \right)^{-1} \underline{X}_{new} \right) \frac{\left(\underline{\hat{Y}} - \underline{Y} \right)^T \left(\underline{\hat{Y}} - \underline{Y} \right)}{m - n} \tag{A.19}$$

where $\underline{X}_{new}$ is an n-element vector whose elements are calculated identically to those of a single column of $\underline{\underline{X}}$, as in Equation A.3. The longitude used to calculate $\underline{X}_{new}$ is that at which the 1 σ uncertainty is desired. Formally, this is the 1 σ uncertainty on what a new observation at that longitude might be. The predicted value of that observation is as in Equation A.6.

Finally, I must quantify the goodness (or otherwise) of the fit. Well-characterized model parameters are useless in any predictive sense if the fit is poor. Since I do not know the appropriate uncertainties I cannot use a χ^2 test. This problem is discussed on page 192 of Bevington (1969).

I used an F-test instead and followed Equation 7.30 in Neter and Wasserman (1974). If there were at least a minimum number of data points, defined later in this section, used in the fit *and* this F-test showed a probability of 90% or greater that not all model parameters beyond the constant density term should be zero, then I accepted the fit as good. If not, then I concluded that there was no justification for modelling the density with anything beyond a constant density term. Bad fits generally occurred in regions where there were significantly fewer data points than usual, which might be due to data dropouts or a high rate of periapsis precession through a given latitude range. Since so many of the bad fits were due to insufficient data rather than the zonal structure being well-characterized as merely a constant density term, bad fits are generally ignored in the Figures either by omission or by interpolating nearby good fits into that region to estimate what the fit would be like if there were more data. If fewer than the minimum number of data points were available, then I did not attempt a fit.

To perform an F-test, I calculated:

$$F = \frac{\left(\overline{Y}\underline{I} - \underline{Y}\right)^T \left(\overline{Y}\underline{I} - \underline{Y}\right) - \left(\hat{\underline{Y}} - \underline{Y}\right)^T \left(\hat{\underline{Y}} - \underline{Y}\right)}{\left(\hat{\underline{Y}} - \underline{Y}\right)^T \left(\hat{\underline{Y}} - \underline{Y}\right)} \frac{m-n}{n-1} \tag{A.20}$$

Where the scalar $\overline{Y}$ is the mean of $\underline{Y}$ and $\underline{I}$ is an n-element vector with all elements equal to 1. This result for F is then compared against the requisite statistical distribution, using IDL's F_PDF function, to find the probability that all model parameters beyond the constant density term are zero. If this probability exceeds 10%, then I declare that the fit is bad.

An F-test can, in theory, be used to determine whether to include any higher-order terms in the fitting function as discussed on page 200 of Bevington (1969). One might consider beginning with a constant density model, then adding progressively higher harmonics until the latest addition failed to provide significant improvement. However, if the zonal structure is dominated, say, by wave-3 and wave-4 terms and has only small contributions from wave-1 and wave-2 terms, then this approach would recommend truncating the fitting function before reaching the dominant harmonics.

For consistency, the same basic model has to be applied to all subsets of the data to which I attempt to fit a harmonic model. It is inconsistent to fit, say, a full wave-4 model to one portion of the data and, say, a constant density term plus wave-3 and wave-4 terms to another portion of the data. I chose to truncate my fits at wave-4 throughout Chapter 2. I chose to use a wave-4 model after extensive experimentation on subsets of data from the Daytime Precession and Polar Crossing parts of Phase 2. Models with fewer harmonics seemed to me to have significantly worse fits, on the whole. Models with more harmonics did not have significantly better fits, though the increase in the number of free parameters without a corresponding increase in the number of data points to control them led to increased uncertainties in the solution for the model parameters. Lacking a formal technique for arriving at this conclusion, I acknowledge that other choices are valid.

A wave-4 model has 9 free parameters and so cannot be fitted to less than 9 data points. If there are only slightly more than 9 data points in a data subset, then all model parameters are predicted with such large uncertainties as to be useless. I did not attempt any fits to data subsets with less than 15 data points. If the model is fit with only slightly more data points than free parameters, then the model parameters have large uncertainties regardless of whether the model is good or not. If more data points are used, then large uncertainties in model parameters mean that the data are more complicated than the simple model. Thus using many data points increases the meaningfulness of the uncertainties on the model parameters.