χ^2 Analysis for OMC3

Nicholas Chapman

15 December 2010

1. The Data

The data were taken on two observing runs: February 2007 and September 2009. There are four overlapping fields to cover the filament OMC3. The number of individual fits images for each region and each year is shown in Table 1.

1.1. sharpinteg

For `sharpinteg` I used the following flags on all files:
`sharpinteg filename -r rgm.dat -f 1 -w -sil -em`
For file #36023 I also added the `-nc` flag to get it to process without checking the chop throw. For file #48359 I added the `-m 200` flag to mask out some correlated noise I observed in the Q and U maps.

1.2. sharpcombine

For `sharpcombine` I used the following flags:

```
sharp_combine point.txt combine.fits -hwp 91 -l 101 101 -sm 2 -ma 5 -ps 9.5 -pm
12 -bg 10 0 -ip -0.0034 0.00017 0.0036 0.0
```

2. χ^2 Analysis

The general idea of χ^2 analysis is to see if our data make sense. To do so, we divide up the data into bins, and see how well the results in each bin agree with the global average. In a perfect world, the reduced χ^2 would be right around 1. In reality, we always have something larger than 1. There are two possibilities for this. One, we have underestimated our errors, so we fix the χ^2 by inflating our error bars by the square root of the reduced χ^2. This is legitimate if the errors are statistical in nature. The other possibility is that there is some systematic error in our results. This is bad, because then we shouldn't be publishing until we understand it. If we plot the average χ^2 versus the average bin width, then we would expect χ^2 to be roughly constant for statistical errors, but to increase for long time scale systematic errors.

To look for this effect, I divided the data five ways. First, I binned all the 2007 and all the 2009 data separately to look for variations from year to year. Secondly, I made bins 6-8 FITS images wide. These data are called 'bin width=7' in the plots. Third, I made bins 4-5 FITS images wide called 'bin width=4'. Fourth and Fifth I attempted to make bins with approximately equal errors rather than equal number of files. 'bin width=σ' uses all the data and to make bins with roughly equal errors. The number of files per bin can vary from 4 to 16. 'bin width=$\sigma 2$' also has equal error, but I don't use all the data. Rather, I picked dither sets with roughly errors. This method is just a subset of the data used for 'bin width=4'.

The number of bins with each method is shown in Table 2. The data in each bin were then run through `sharp_combine` as described above to produce final maps.

The χ^2 analysis literally computes reduced χ^2 for each pixel and comparing to the value from the combined average, e.g. for Stokes Q:

$$\chi_r^2 = \frac{1}{N-1} \sum_{i=1}^{N} \frac{Q_i - Q_{avg}}{\sigma_{Q_i}^2} \tag{1}$$

Q_{avg} is the value found from combining all the data. To make a fair comparison between the different bin sizes, I chose the pixels 40-60 in both x and y, giving me 400 total pixels used in computing statistics. These pixels are good no matter the bin size, therefore we always have the same number of pixels.

Figures 1-4 show the χ^2 maps of Q and U for both the variable σ and constant σ. The box in each map shows the pixels we used in computing statistics. The contours are for $\chi^2 = 1, 3, 5, 7$,etc. Figure 5 shows the χ^2 trend with average number of files per bin. Note that 'bin=year' is also used as the rightmost datapoints for constant σ. This is because the errors across years are similar.

3.   Conclusions

Figure 5 doesn't show much of a trend. It looks fairly flat regardless of the binning method used. Remember, if there were systematic errors, then we would expect to see a positive slope. The point for OMC3_2 with average number of 4 (the leftmost point) is significantly higher than other points. I think this is because the χ^2 computed for Stokes U is much larger than for Stokes Q. This can be readily seen in Figure 4 under the bin=$\sigma 2$ column. However, I believe we can still treat our errors as statistical. Even if we believe this data point, it would suggest a negative trend, not a positive one. Furthermore, data for the other regions are approximately flat, even though these data were often taken at the same time as OMC3_2. Lastly, this might just simply be a case of poor data. When we use all the data (bin width=4), the data point for OMC3_2 is still high, but has dropped. *My final conclusion is that we can treat the errors as statistical, not systematic and thus inflate our errorbars accordingly.*

4. Final Catalog

4.1. Inflating the Errors

The most reliable estimate of χ^2 comes from constant σ bin width=σ. This uses all the data, but still uses four bins to provide better statistics. Now that I have established that our errors are statistical, I went back and recomputed the average χ^2 from the entire map. These values are: $\chi^2 = 3.59, 2.83, 3.33, 2.78$ for OMC3_1 through OMC3_4, respectively. Because we have so few bins, it seems best to apply a single correction for the entire map rather than on a pixel-by-pixel basis.

4.2. Getting the vectors

There are four partially overlapping regions. I extracted vectors from each map separately. Because the maps oversample $4\times$, we could only extract vectors for every 4th pixel. However, the question still arises, where to start counting? I see four options:

1. Start at the bottom left corner

2. Start at the first row with a non-NaN pixel, and the first column from the left.

3. Start at the first column with a non-NaN pixle, and the first row from the bottom.

4. Start at a pixel to ensure that the flux peak/object of interest/center of the map can have a vector

Because the maps have a large border of NaN pixels, option 1 doesn't seem like the best choice. Furthermore, because OMC3 is a filament, and we are concerned about the magnetic field along the full filament, then option 4 does not seem like a good choice either. I arbitrarily chose option 2, though option 3 should produce very similar maps. Because the first non-NaN pixel is different in every map, the offsets needed for `polsharp5` are different in every map. For OMC3a and OMC3b, the offsets are `offset=1,1` (pixel 42,30 and 46,30). For OMC3c, it is `offset=1,0` (pixel 50,29), and for OMC3d it is `offset=3,0` (pixel 56,29).

I extracted all vectors $\geq 2.5\sigma$ after applying the χ^2 correction mentioned above and using these offsets. At the same time I applied a correction for polarization efficiency. This amounts to dividing p and σp by 0.93 (increasing both by about 1.08). Thus, correcting for polarization efficiency does not impact S/N.

4.3. Mosaicing the vectors

To mosaic the vectors from each subregion into a single catalog, I position matched with a distance criterion of $9\rlap{.}''45$. This meant that any vector within $9\rlap{.}''45$ of another vector was averaged

together to produce a single output vector. I did the averaging in Stokes Q and U space, then converted back to polarization. I chose $9\rlap{.}''45$ because I found $9\rlap{.}''5$ led to some combining of vectors within a single region. I think the nominal resolution of 4 pixels actually works out to about $9\rlap{.}''48$. I had 87 total vectors, and 81 remained after averaging.

I also made a combined mosaic image of Stokes I. Many Bothans died to bring you this map. Most importantly, the world coordinates are not the same as in the orion1 map I got from Di Li and published in Li et al. (2007). I think there is a shift of roughly $8''$ in RA. My map has higher RA than those in Di Li's map. This is important because you will undoubtedly want to display the vectors on the much nicer looking ORI1 map from SharcII. The coordinates in my data file are self-consistent with the map I produced, but will not overlay properly with Di's map. To fix this problem, I did a quick bye-eye calibration. I compared the peak positions of ORI1_13 and ORI1_6 in both maps and computed the average change in coordinates needed for Di's map. This was $+7\rlap{.}''836$ in RA and $-0\rlap{.}''288$ in DEC. I then manually changed the CRVAL1 and CRVAL2 in the FITS header. See Figure 6 for a nice map of the field.

4.4. Included Files

The following files are included in the tarball:

- combined_stokesI.fits - This the combined intensity image for all four regions

- ori1_cal_fixed.fits - This is Di Li's Orion1 map with coordinates updated to agree with combined_stokesI.fits

- pol_3sigma.dat - All polarization vectors $\geq 3\sigma$

- pol_2.5sigma.dat - Polarization vectors with $2.5 \leq \sigma < 3$

Remember, you will need to rotate the vectors $90°$ to show the direction of the B field.

Table 1. Number of FITS Images per Region

Region	Feb. 2007	Sept. 2009	Total
OMC3_1	21	8	29
OMC3_2	8	16	24
OMC3_3	8	12	20
OMC3_4	12	8	20

Table 2.   Number of Bins for each Binning Method per Region

		Bin Width			
Region	Year	7	4	σ	$\sigma 2$
OMC3_1	2	4	7	4	4
OMC3_2	2	3	6	4	4
OMC3_3	2	3	5	4	4
OMC3_4	2	3	5	4	4

Table 3.   Average χ^2 in each region.

		Bin Width			
Region	Year	7	4	σ	$\sigma 2$
OMC3_1	3.44	4.64	3.71	4.62	3.85
OMC3_2	5.62	3.51	6.07	3.74	7.45
OMC3_3	2.87	2.66	3.16	3.64	3.71
OMC3_4	2.78	2.78	2.85	3.06	2.78
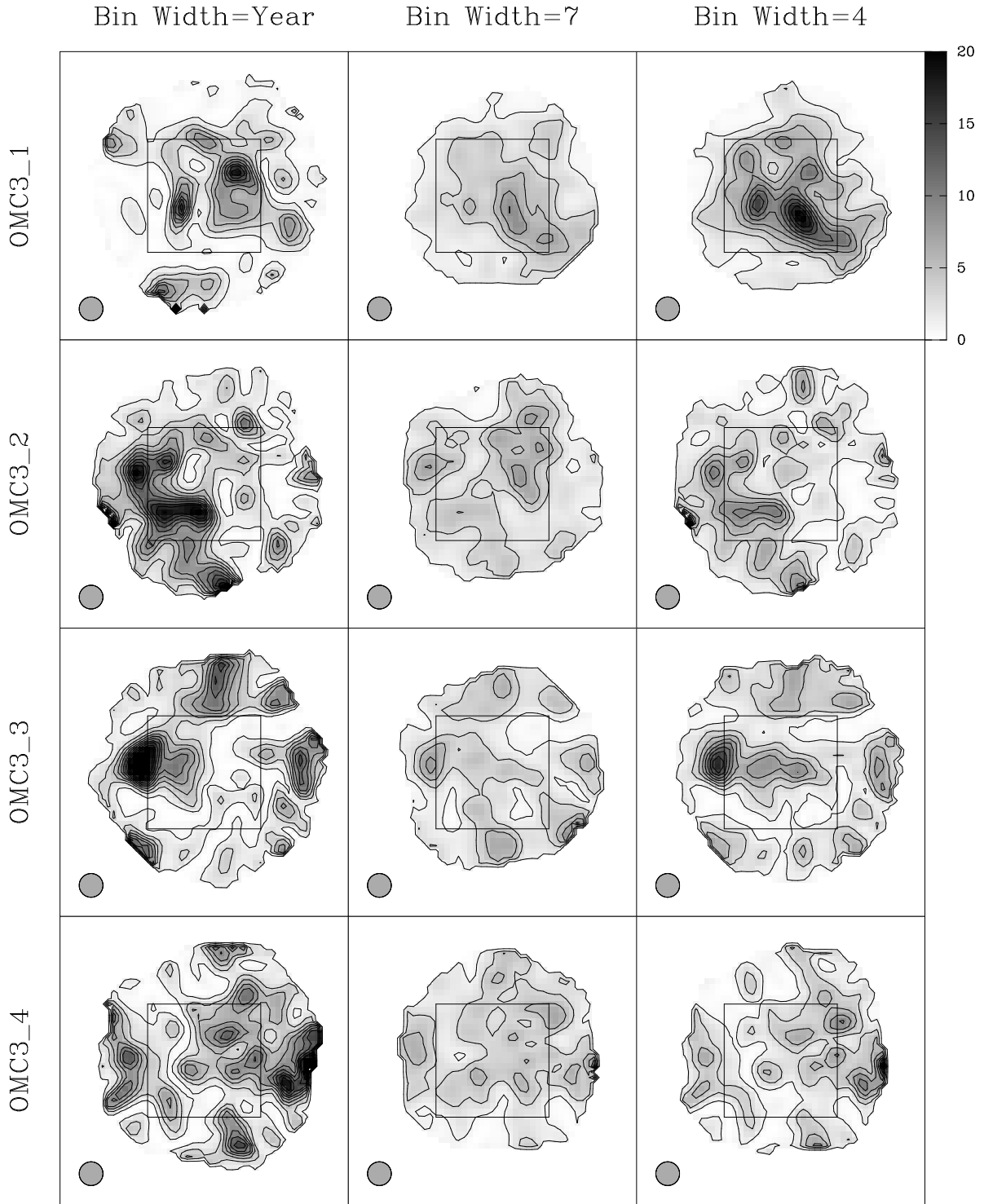
Fig. 1.— χ^2 map of Stokes Q for variable σ. A $10''$ beam is shown in the bottom left. The box denotes the region we used for computing statistics. Contours are $\chi^2 = 1, 3, 5,$ etc.
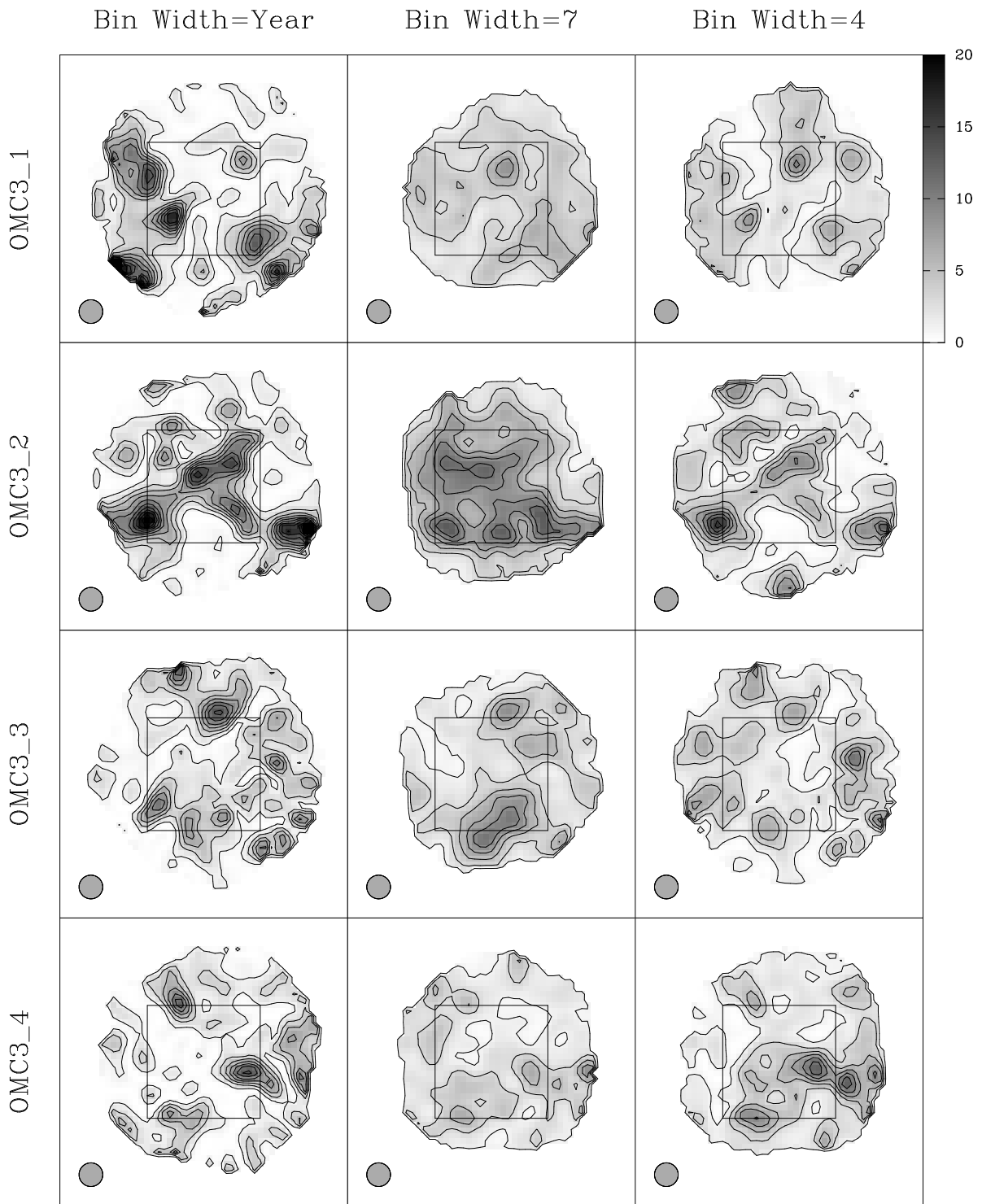
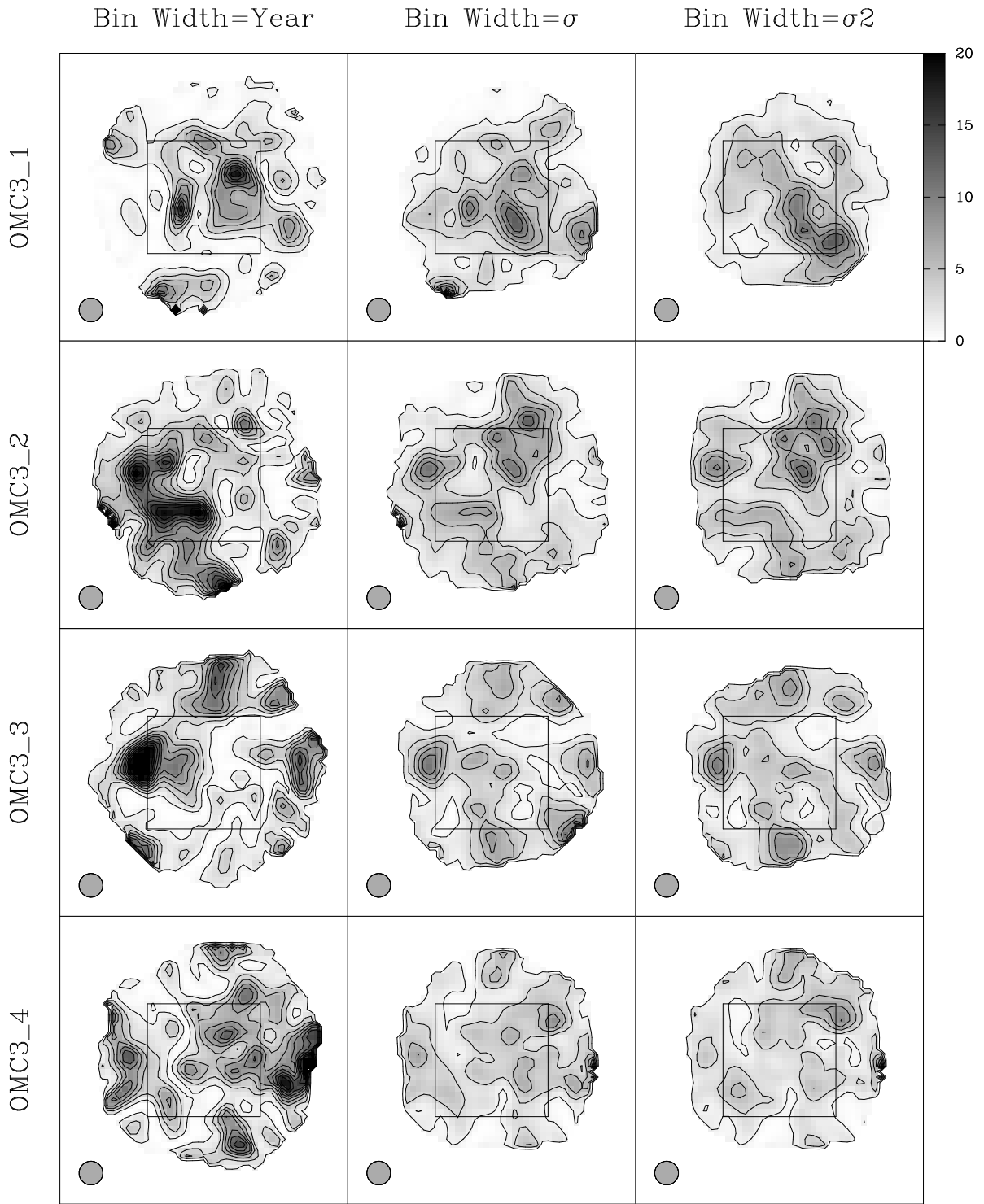Fig. 2.— χ^2 map of Stokes U for variable σ. Contours, etc are the same as in Figure 1.

Bin Width=Year   Bin Width=σ Bin Width=$\sigma2$

Fig. 3.— χ^2 map of Stokes Q for constant σ. Contours, etc are the same as in Figure 1.

Bin Width=Year   Bin Width=σ Bin Width=$\sigma2$

OMC3_1

OMC3_2

OMC3_3

OMC3_4
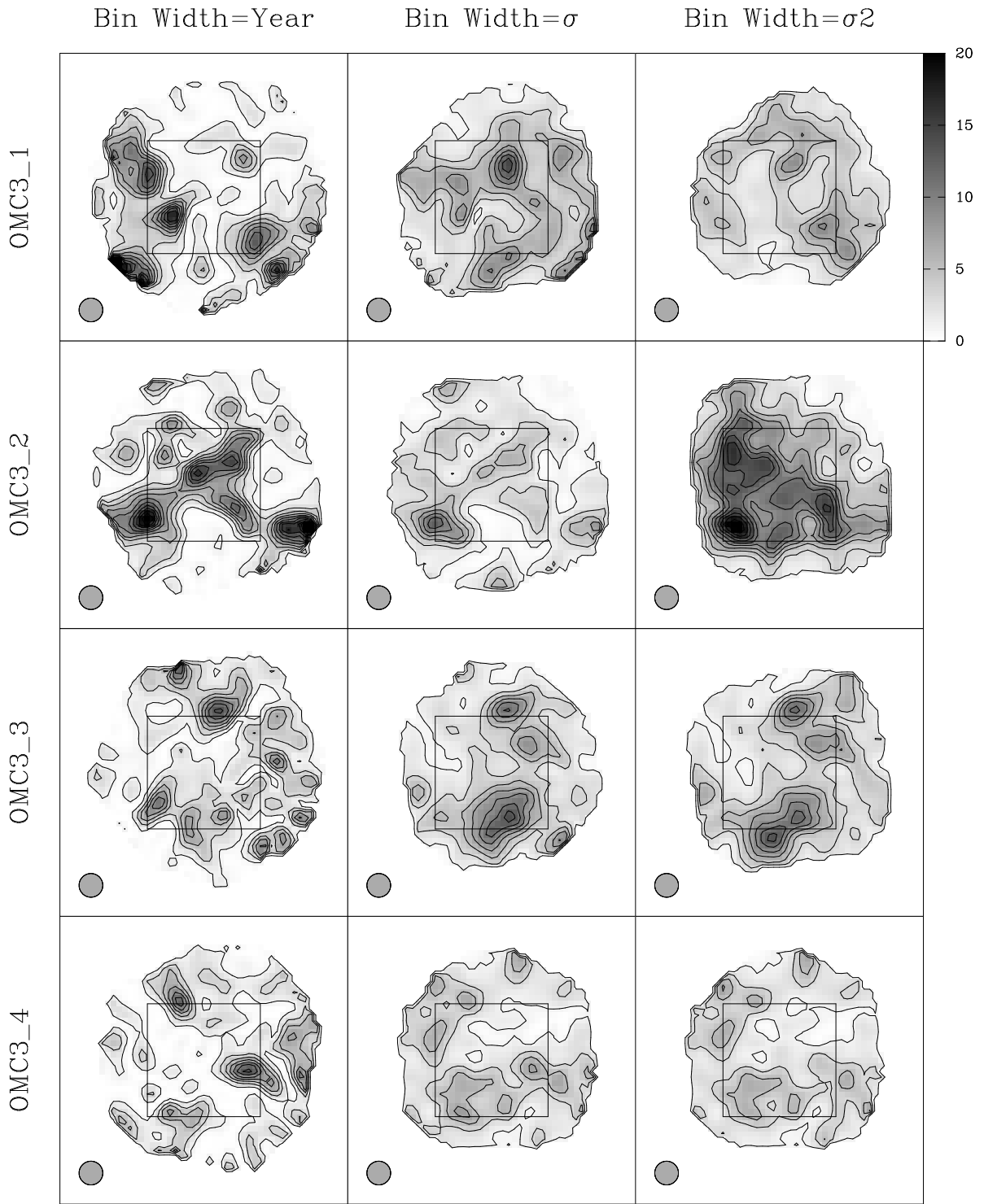
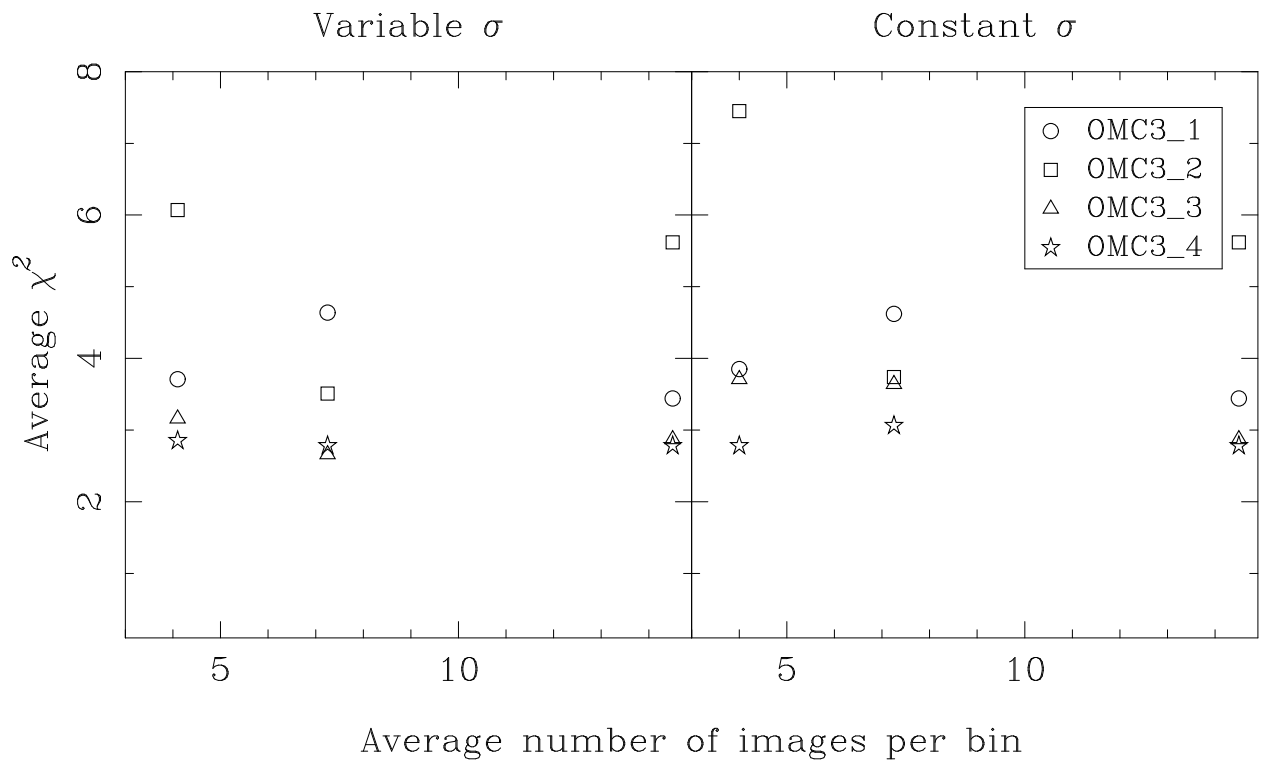Fig. 4.— χ^2 map of Stokes U for constant σ. Contours, etc are the same as in Figure 1.

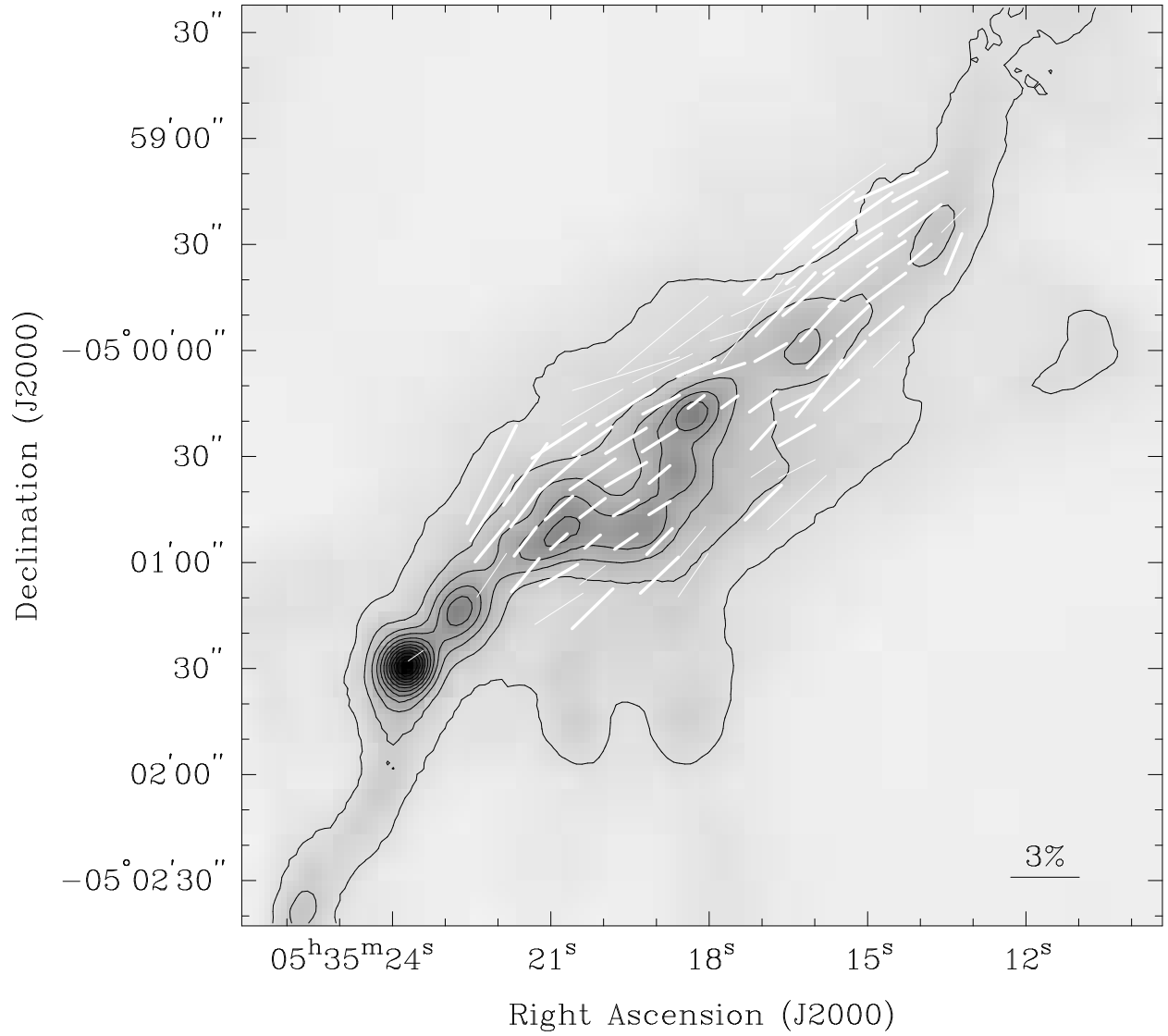Fig. 5.— Trend of average χ^2 versus average number of images per bin.

Fig. 6.— Map of the OMC3 filament from Li et al. (2007) with the E-field vectors overlaid. Thick lines are $\geq 3\sigma$ while thin lines are those with $2.5 \leq \sigma < 3$. Contours start at 5 Jy/beam in steps of 5.