

1 Introduction

The results here are some things which are alluded to in the main manuscript, but which would be far too much detail for there. All analyses were run with COSAR https://github.com/markmuetz/cosar_analysis, and omnium <https://github.com/markmuetz/omnium>. All figures should be compared against Fig. S1.

List of Figures

S1	Production run: Favour lower troposphere on, seed 391137	2
S2	DJF	3
S3	MAM	4
S4	JJA	5
S5	SON	6
S6	Favour lower troposphere off	8
S7	Seed 106586	10
S8	Seed 12042	11
S9	Seed 707637	12
S10	Seed 725154	13
S11	Low CAPE	16
S12	High CAPE	17
S13	Low shear	18
S14	High shear	19
S15	Low favour factor	20
S16	High favour factor	21
S17	K-means scores	22

List of Tables

1	RWPs, showing the match up between production and the different seeds	9
2	RWPs, showing the match up between production and the various sensitivity tests	14

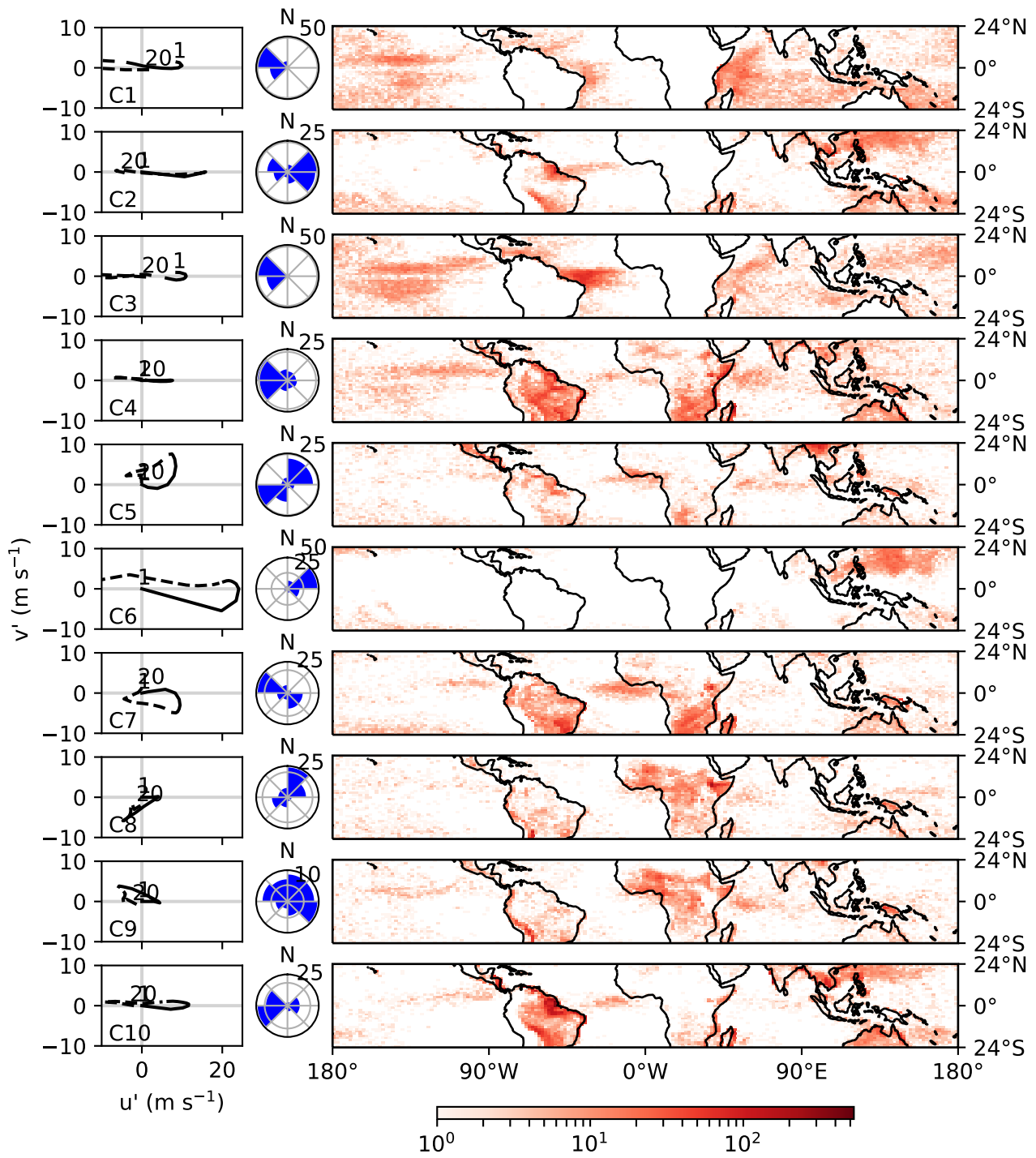


Figure S1. Production run: Favour lower troposphere on, seed 391137

2 Seasonal geographical locations

These are figures showing the geographical heatmaps for each of the seasons DJF, MAM, JJA, SON.

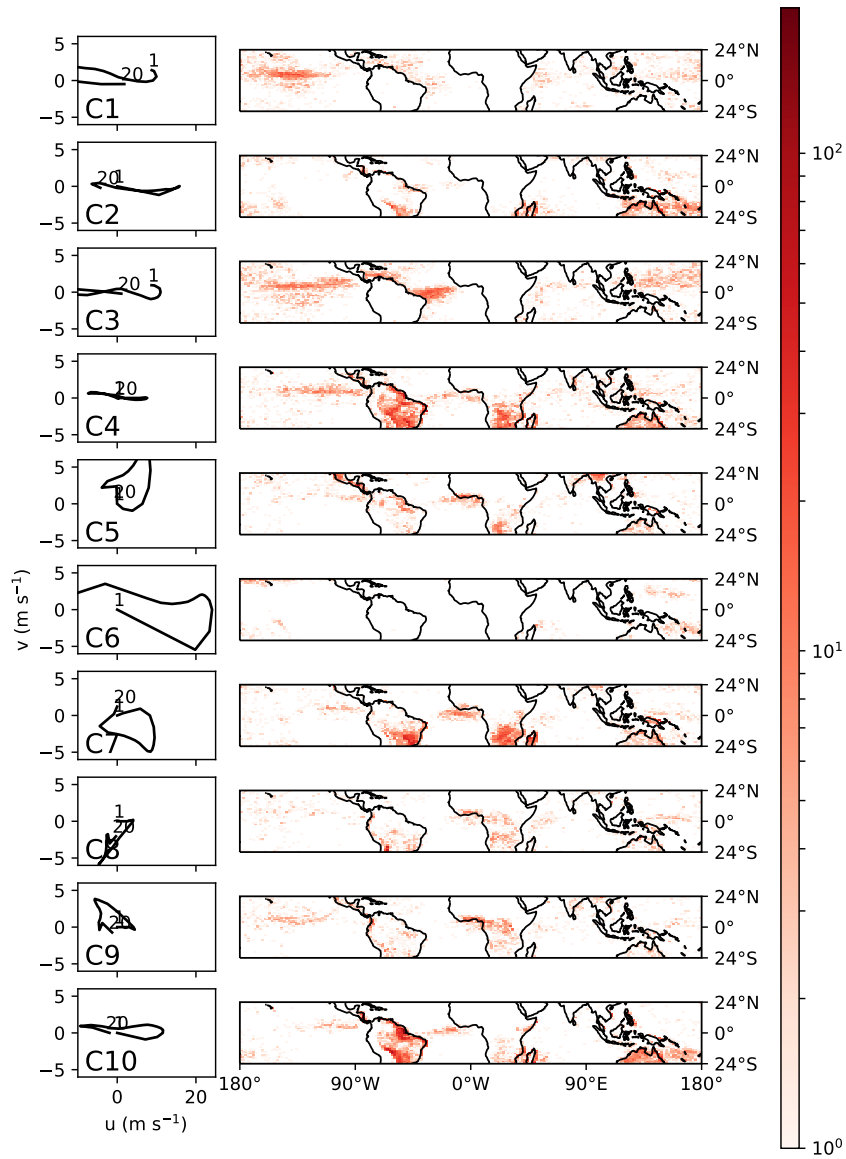


Figure S2. DJF

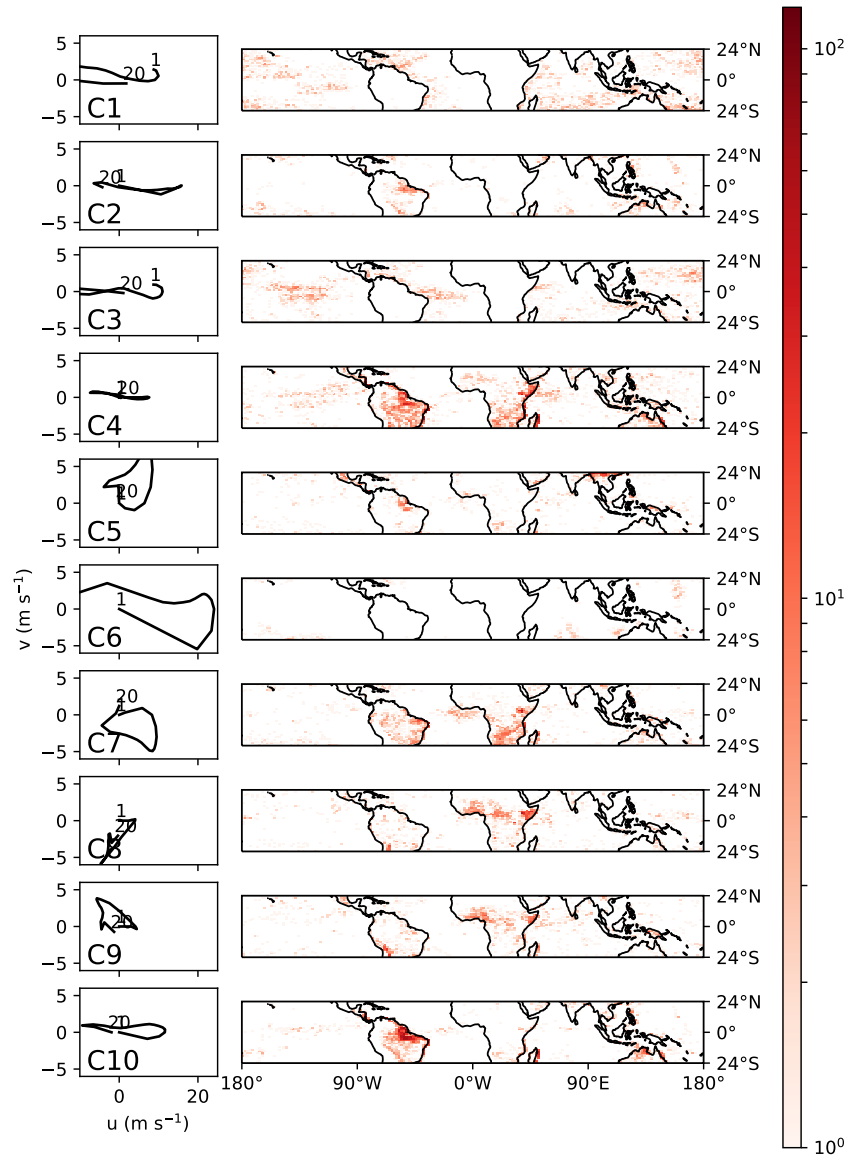


Figure S3. MAM

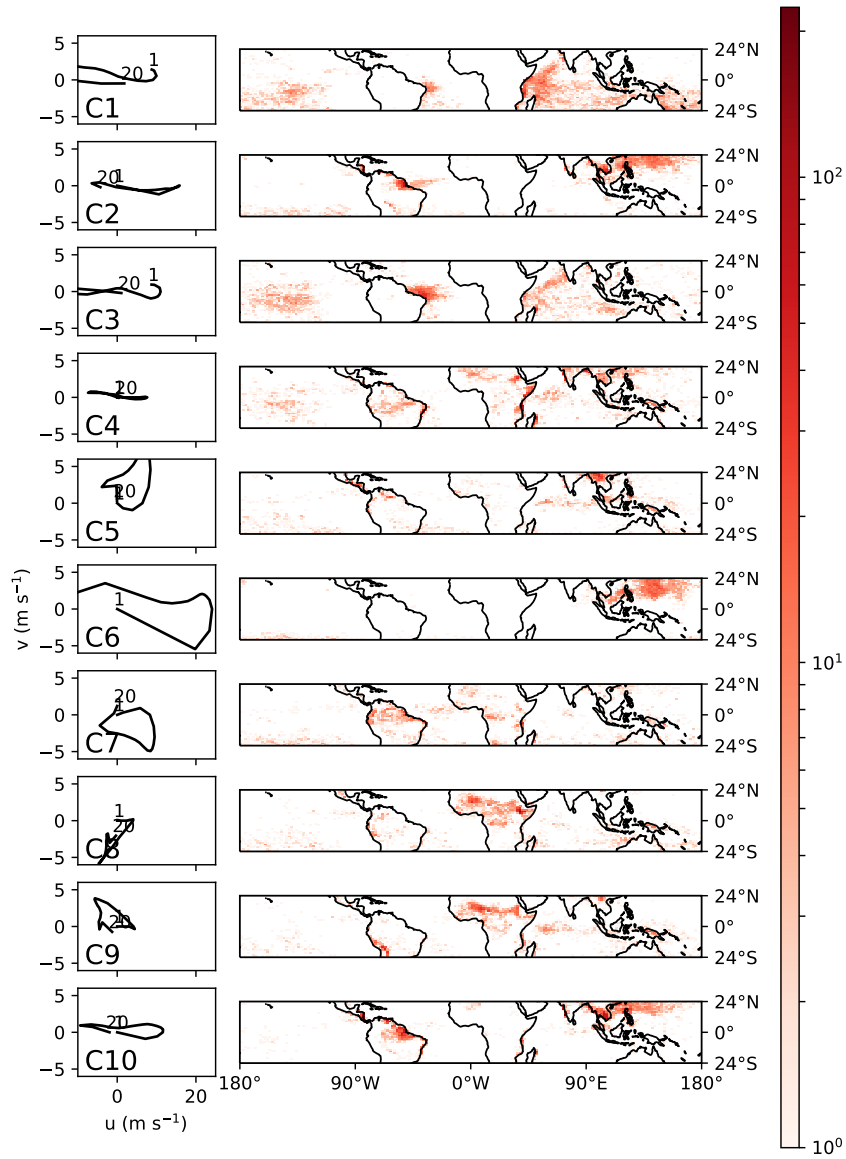


Figure S4. JJA

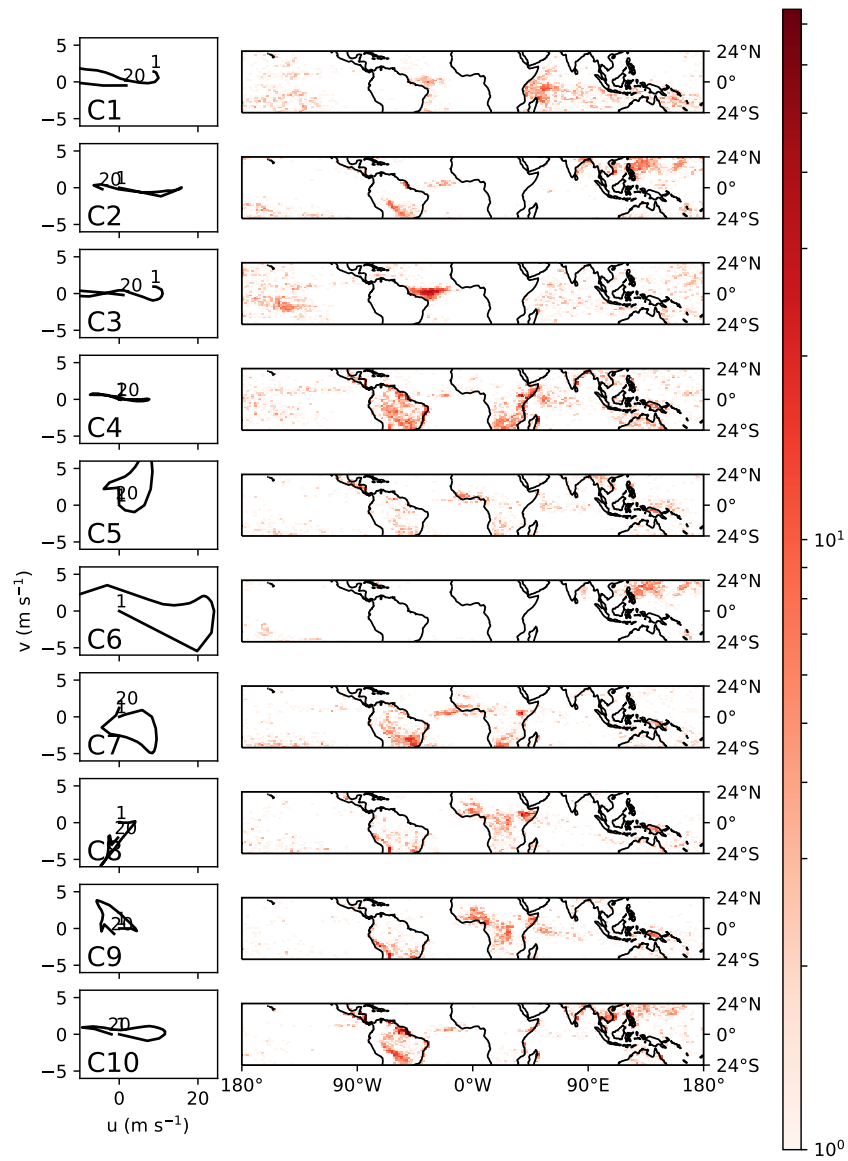


Figure S5. SON

3 Sensitivity to running with favour lower troposphere on/off

This section shows some justification for running with favour lower troposphere on. The original reason for running with this setting was to try to minimize the effect of the higher-latitude upper level winds in the clustering procedure. However, I wondered if this was still necessary after restricting my definition of the tropics to 23N - 23S (from 30N - 30S). I believe that it is after running this sensitivity test.

The means by which the lower troposphere is favoured is described in detail in the main document; briefly, it is achieved by weighting the contribution from above 500 hPa less than the contribution from below. The justification for doing this is that the lower-level winds are more important for the organization of convection.

Figure S6 below shows why favouring the lower troposphere is done. In Fig. S1, which is run with the production settings, the RWPs are well defined and are not dominated by higher-level winds. The opposite is true in Fig. S6. In particular, in C9 and to a lesser extent C5 there is a strong signal from the high-level winds. Thus it makes sense to carry on using the favour lower troposphere setting. Also, there is not a one-to-one correspondence between these RWPs and those with the production settings.

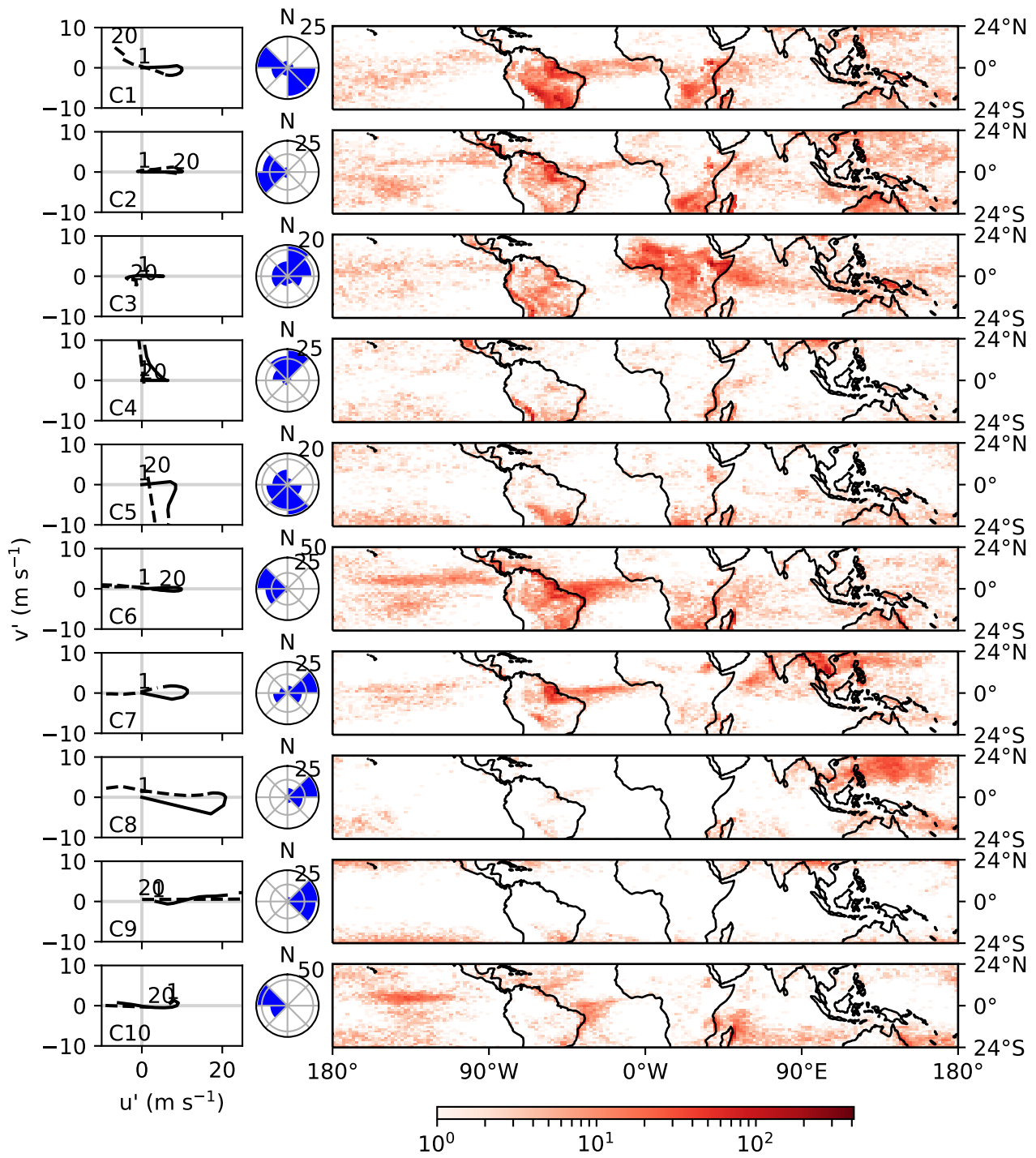


Figure S6. Favour lower troposphere off

4 Random seeds

Using five different random seeds leads to the same set of RWPs, with only slight differences. One-to-one correspondence in all cases.

391137 (prod)	106586	12042	707637	725154
C1	8	7	7	7
C2	4	4	3	1
C3	9	3	2	8
C4	3	6	4	9
C5	2	10	6	5
C6	7	8	10	4
C7	10	1	5	2
C8	6	2	9	6
C9	1	9	1	10
C10	5	5	8	3

Table 1. RWPs, showing the match up between production and the different seeds

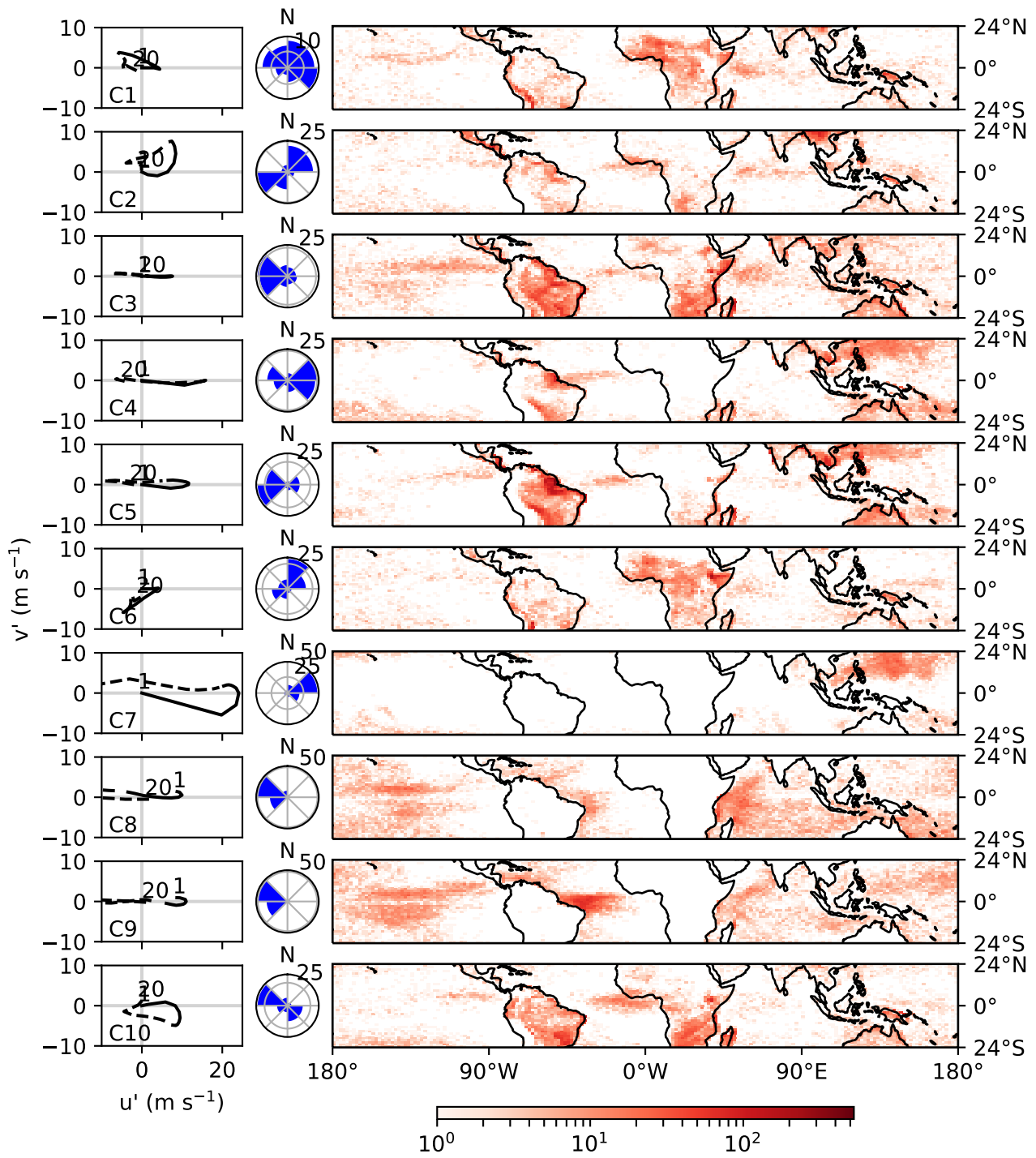


Figure S7. Seed 106586

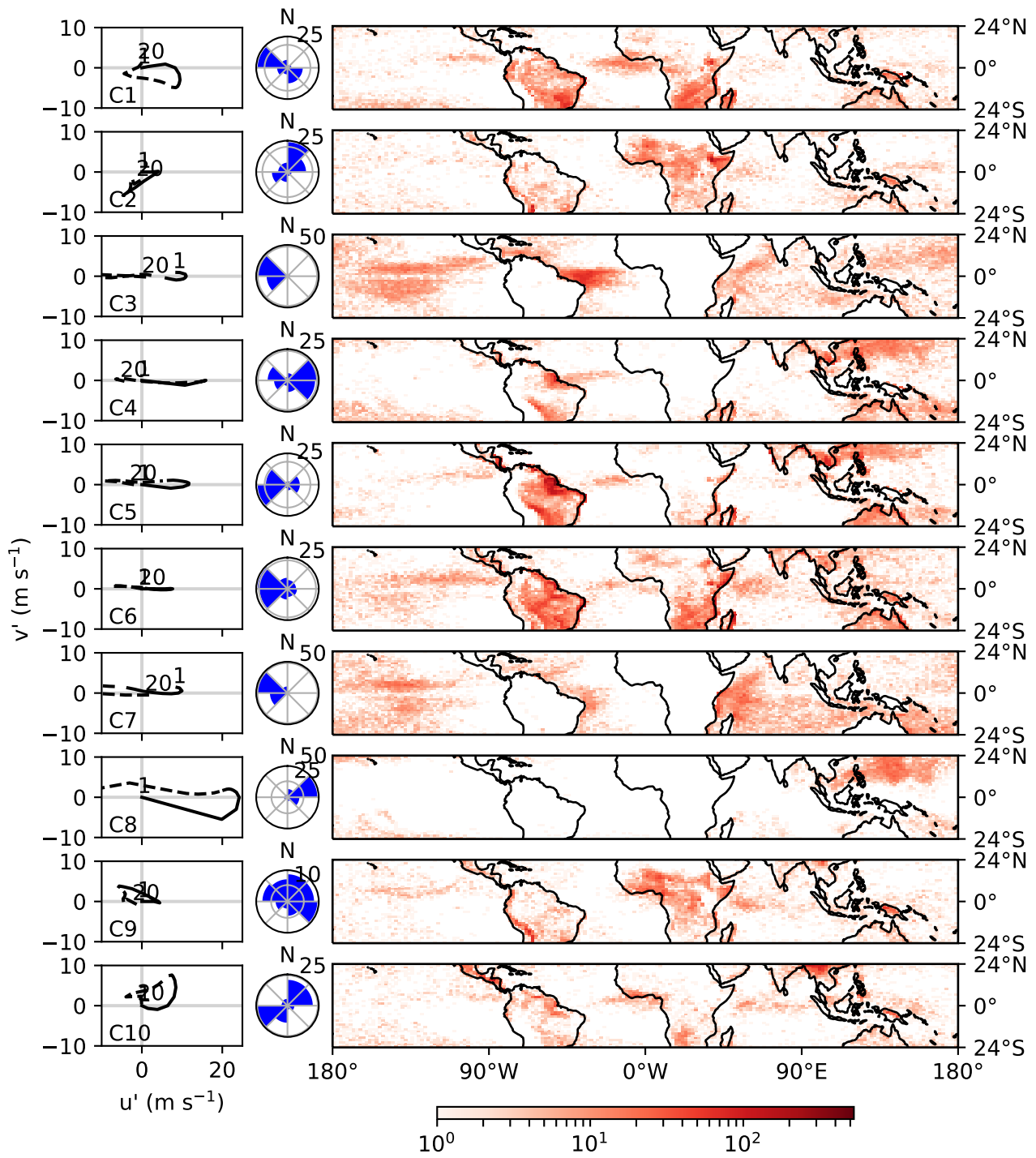


Figure S8. Seed 12042

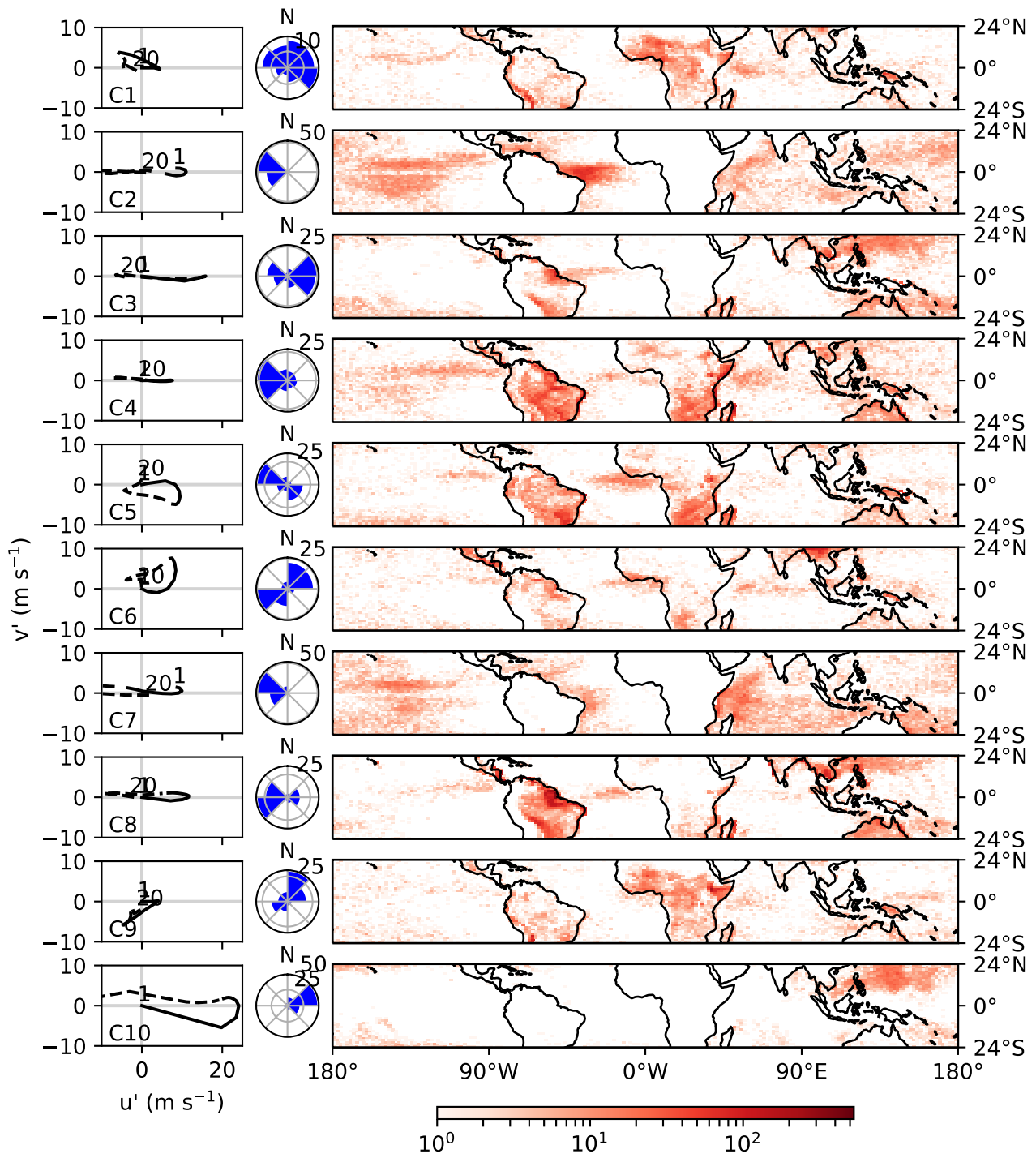


Figure S9. Seed 707637

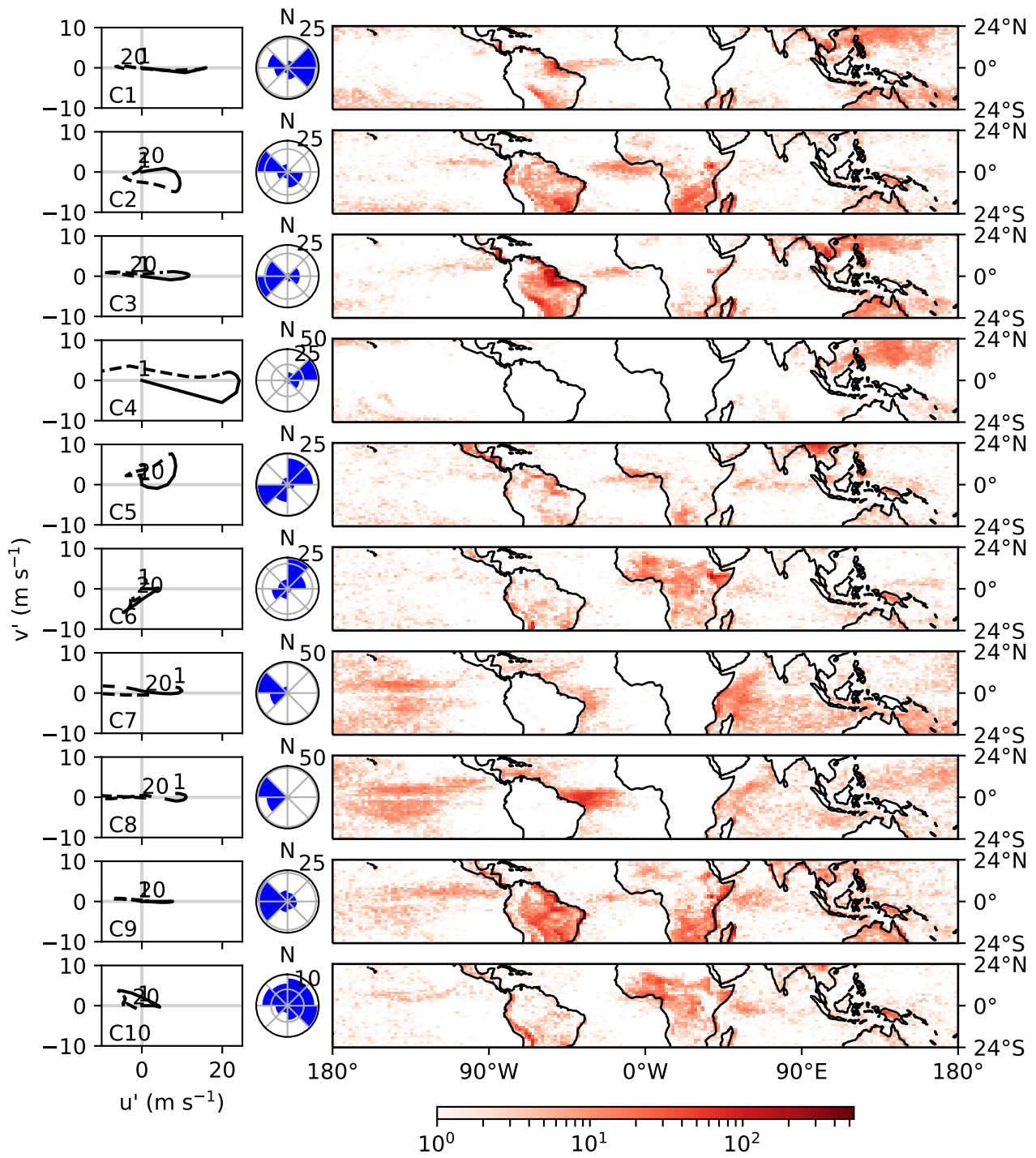


Figure S10. Seed 725154

5 Sensitivity to parameters

prod	low CAPE	high CAPE	low shear	high shear	low fav.	high fav.
C1	6	6	10	5	5	10
C2	1	5	2	6	4	9
C3	3	9	3	4	6	5
C4	10	8	6	7	7	4
C5	9	10	5	8	2	6
C6	5	7	8	10	10	7
C7	7	3	4	9	3	2
C8	2	2	1	3	1	1
C9	8	4	7	1	8	8
C10	4	1	9	2	9	3

Table 2. RWPs, showing the match up between production and the various sensitivity tests

I ran six sensitivity tests:

1. low CAPE: CAPE less than 75 J kg^{-1}
2. high CAPE: CAPE less than 125 J kg^{-1}
3. low max. shear: max. shear less than 65th percentile
4. high max. shear: max. shear less than 85th percentile
5. low favour factor: favour factor = 3
6. high favour factor: favour factor = 5

The match up between production and each of the sensitivity tests was good. In all cases, there was a one-to-one correspondence between the 10 RWP in production and the sensitivity tests (see Table 2). For “high CAPE”, the form of C6 was slightly different. “high CAPE” and “high shear” runs had noticeably fewer profiles included. “low favour factor” and “high favour factor” had respectively 8 and 6 PCs.

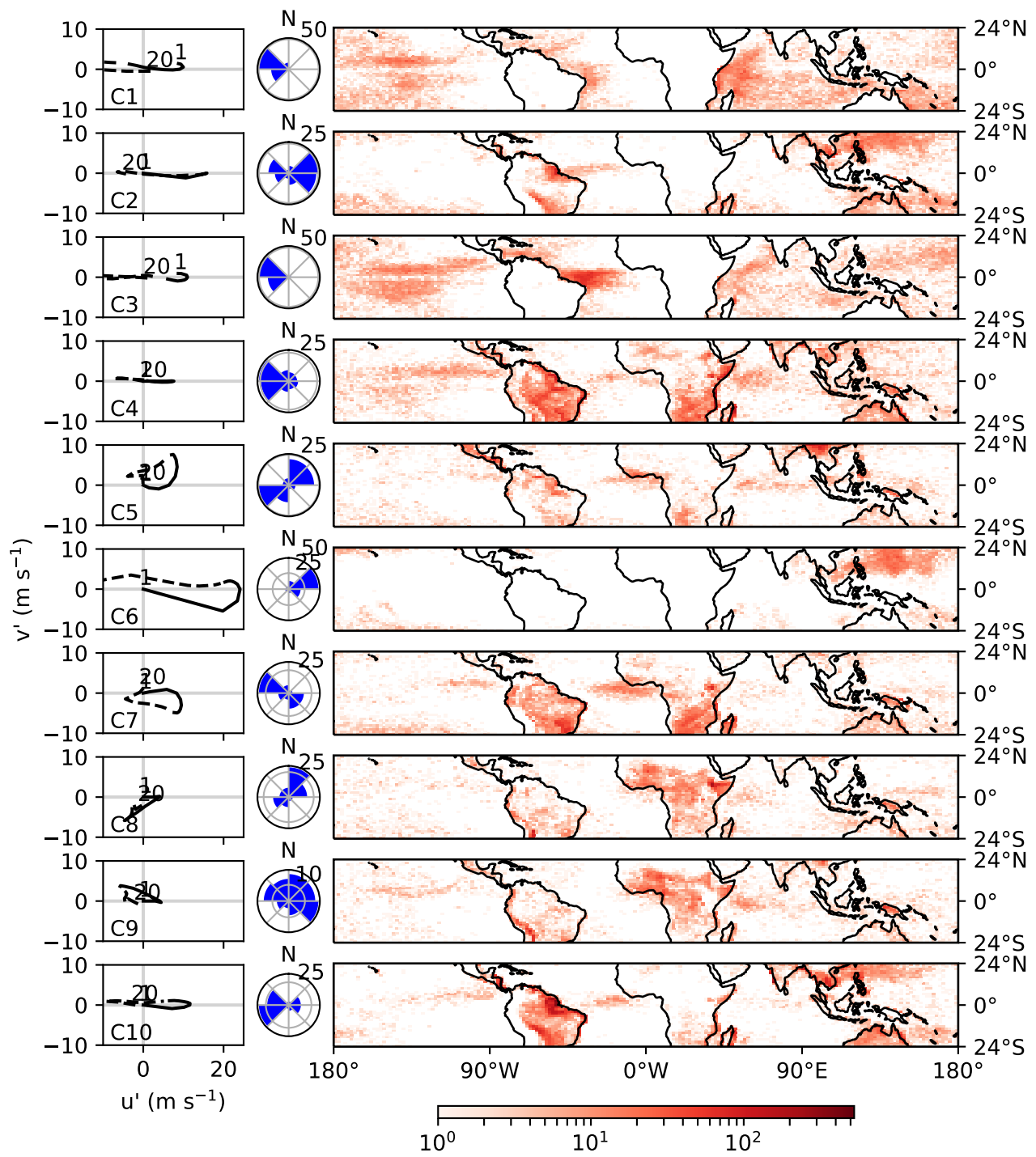


Figure S11. Low CAPE

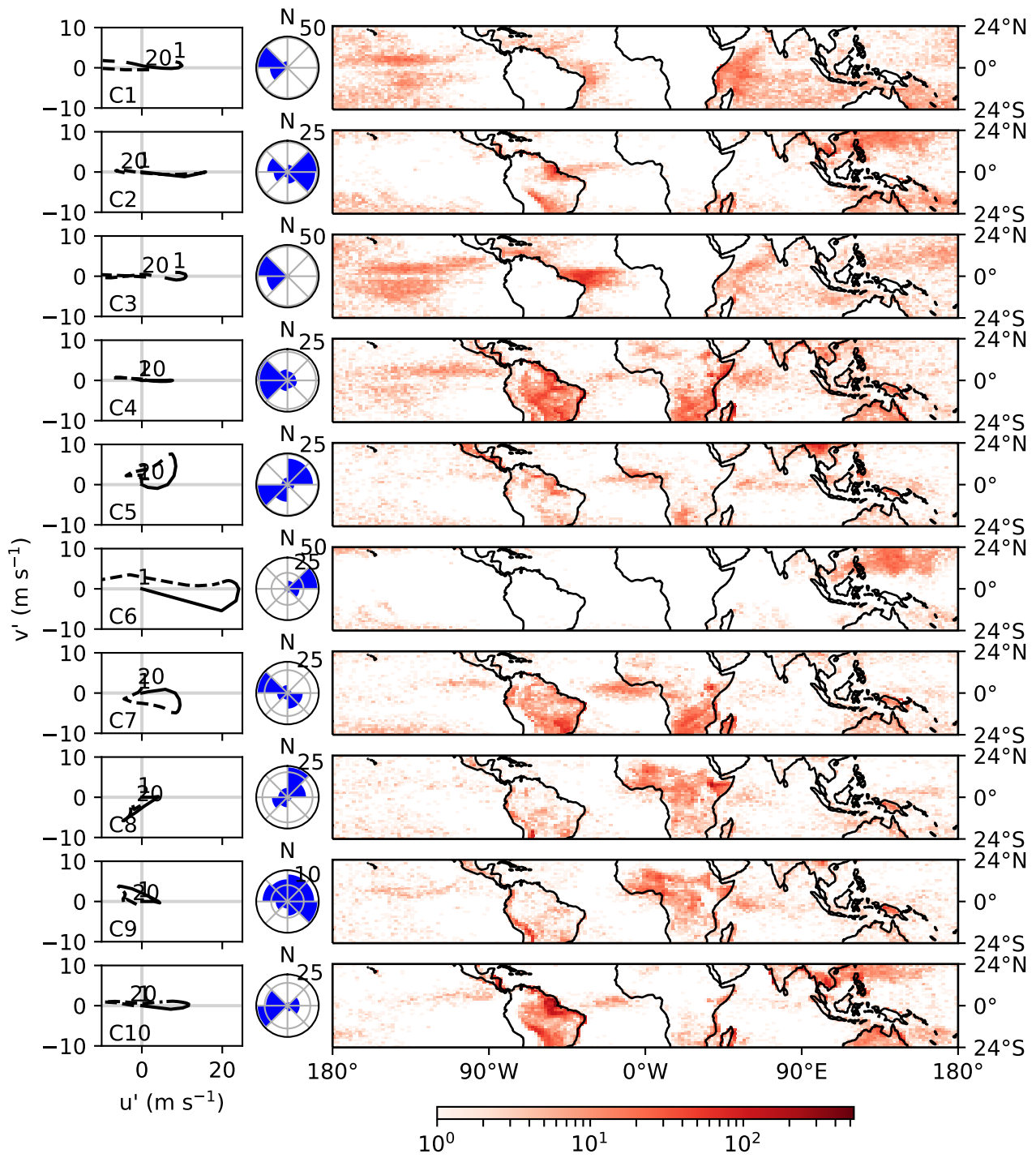


Figure S12. High CAPE

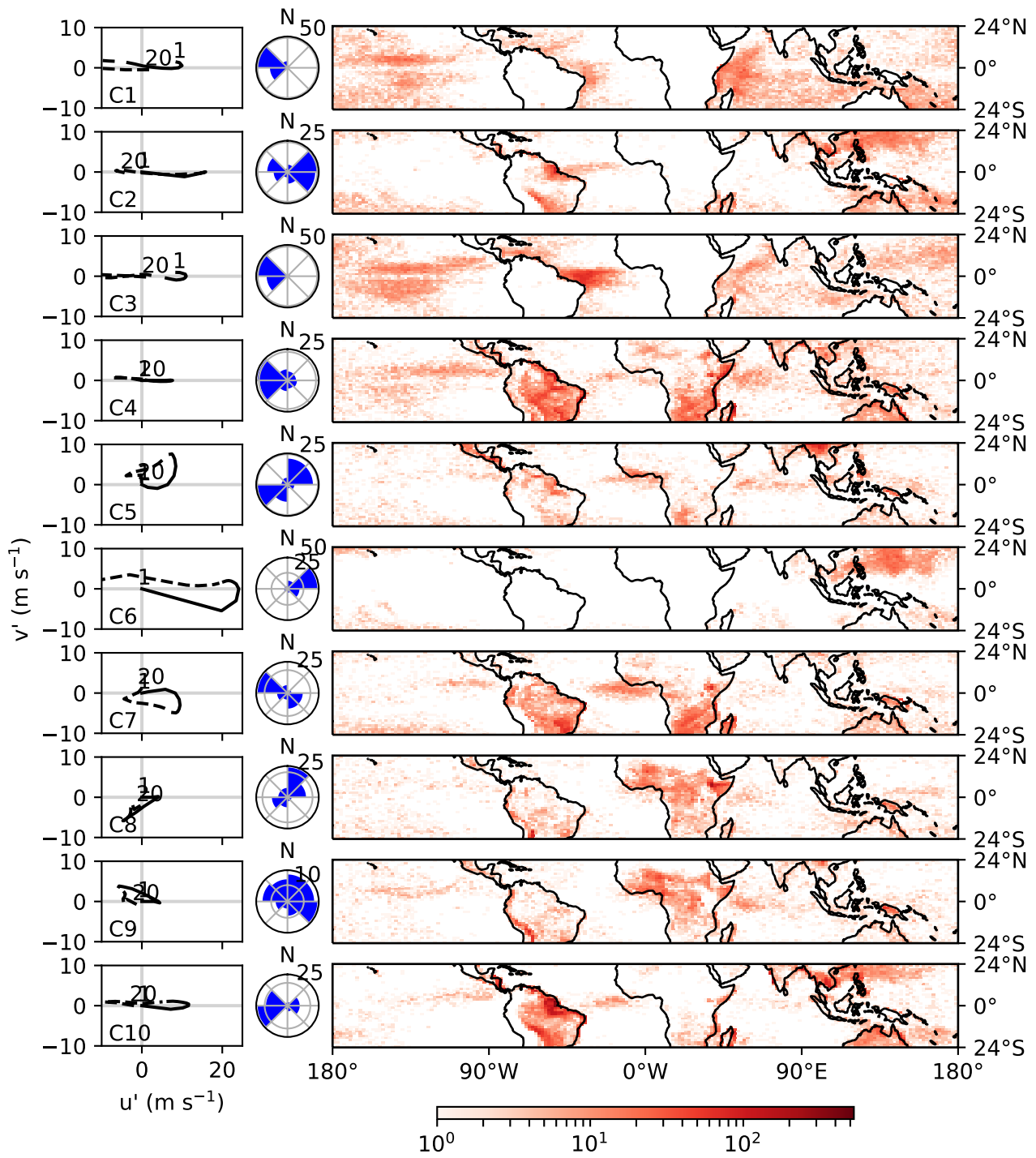


Figure S13. Low shear

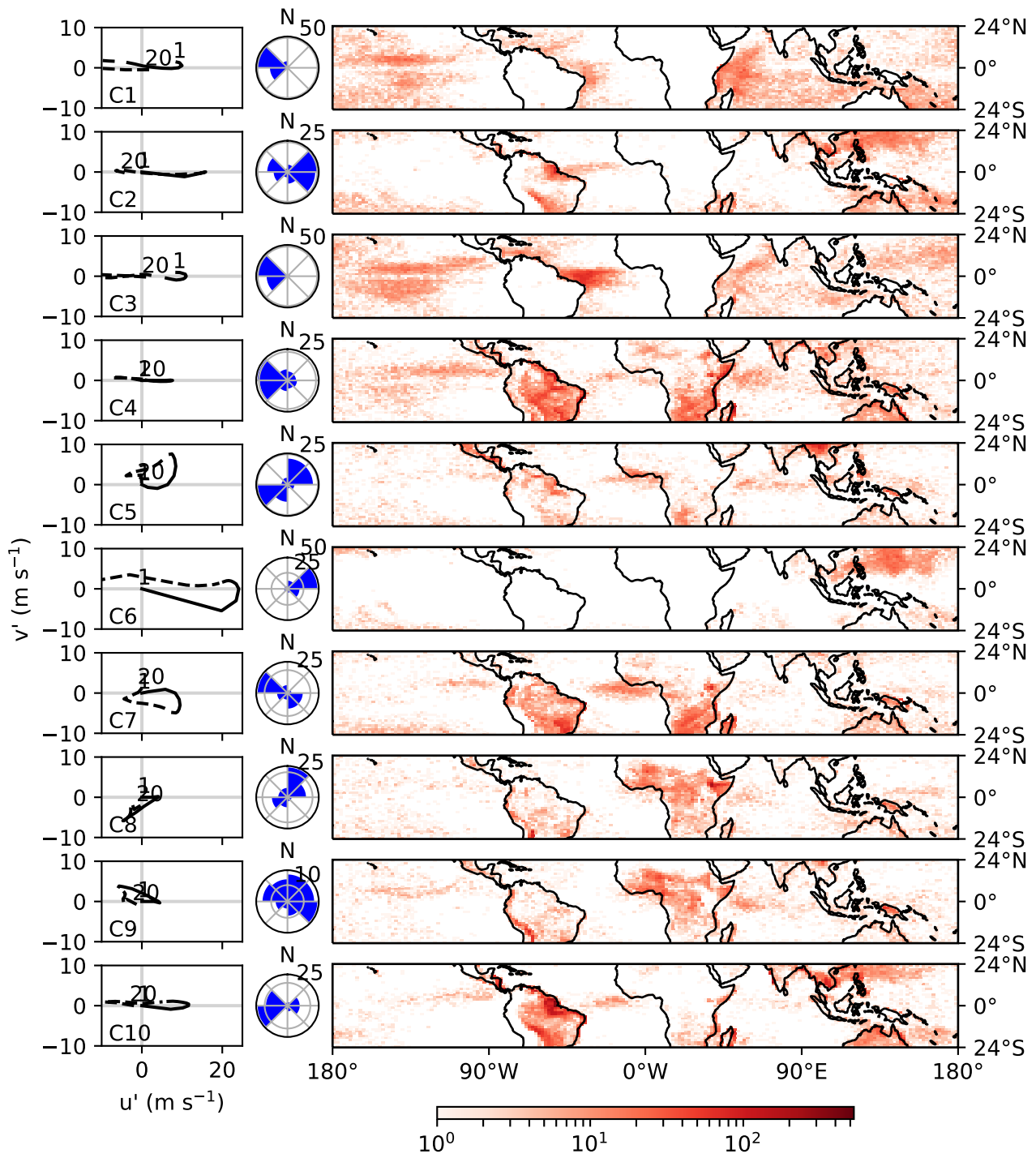


Figure S14. High shear

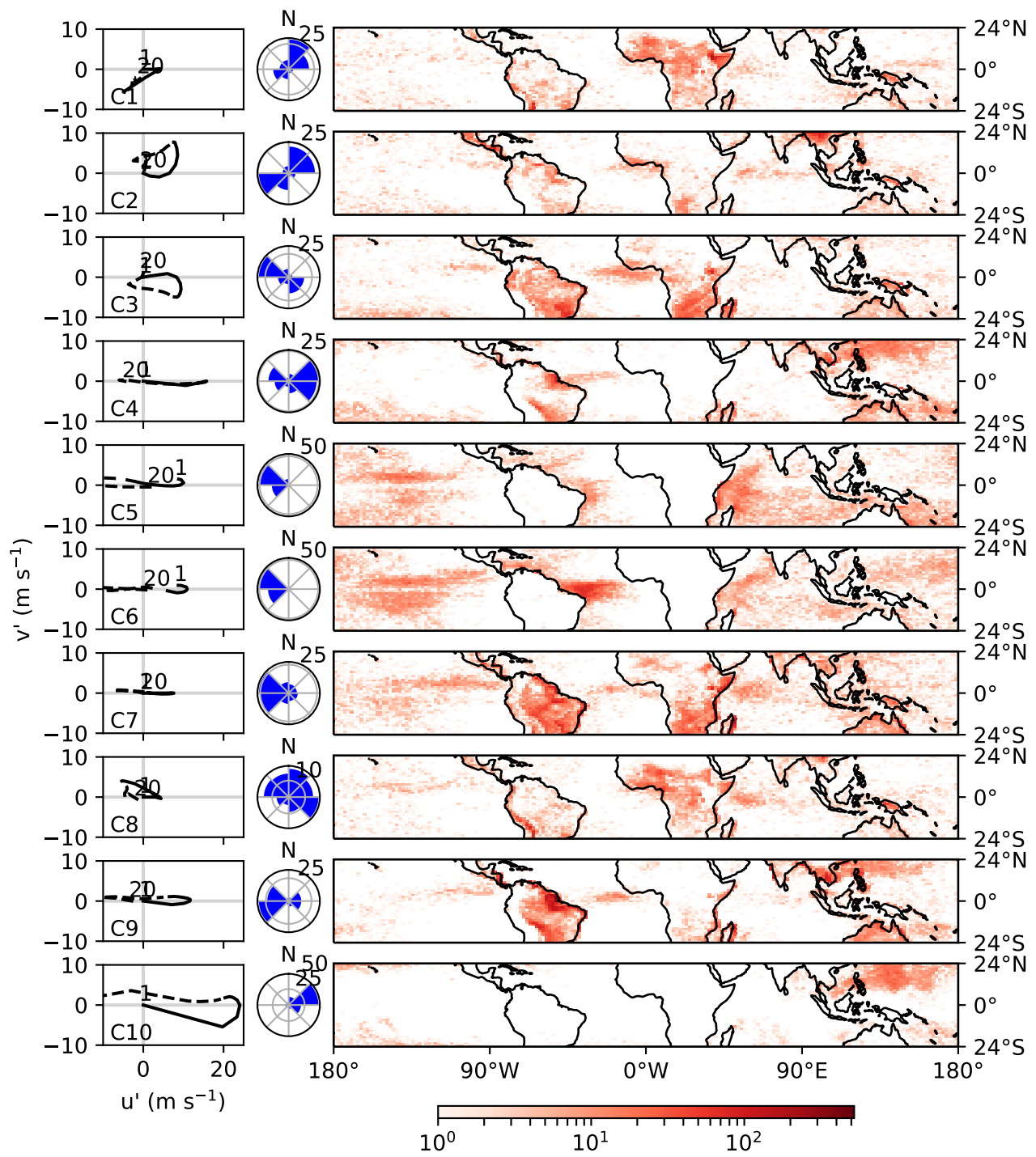


Figure S15. Low favour factor

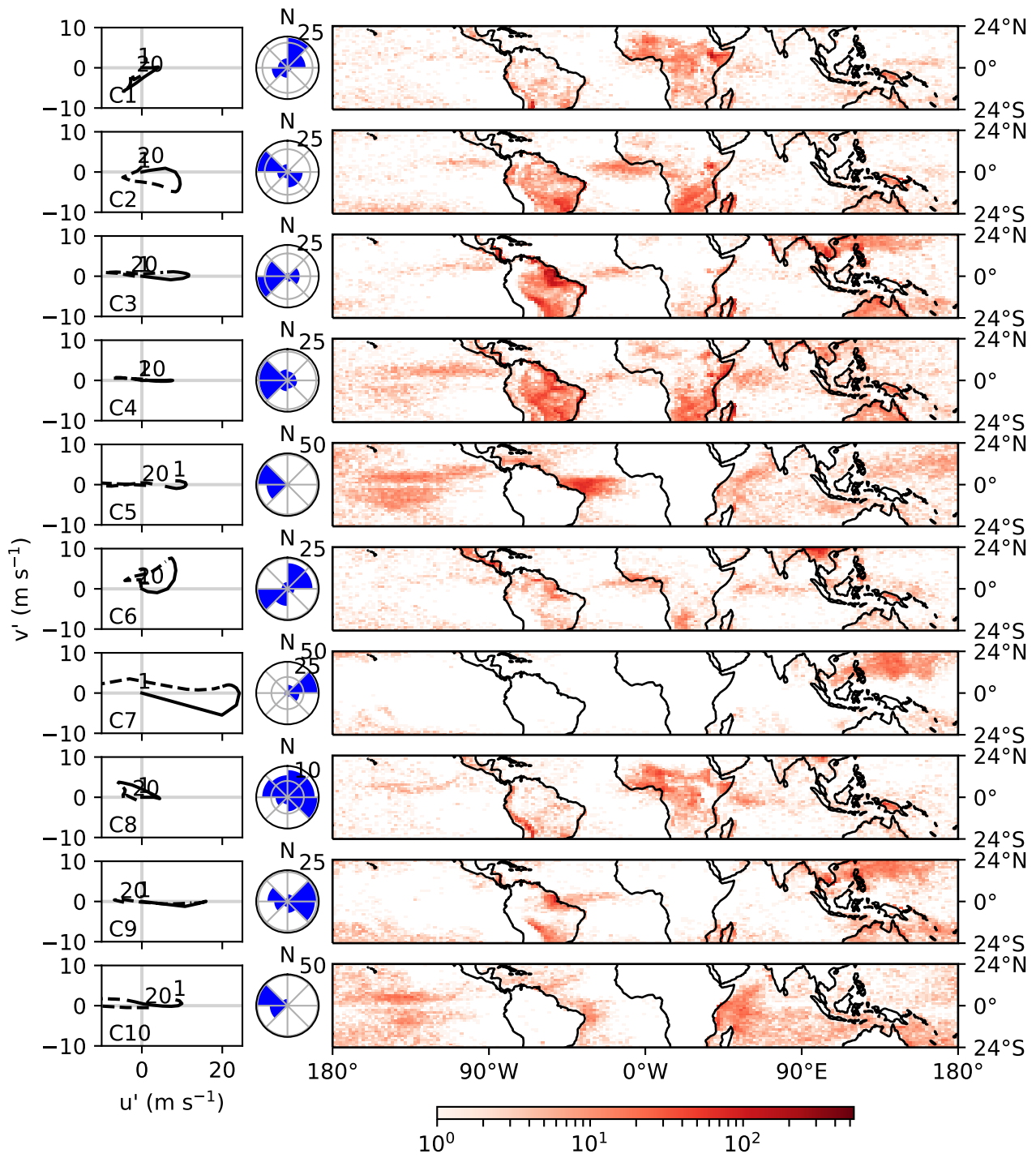


Figure S16. High favour factor

6 K-means elbow plot

There is no obvious 'kink' in the scores for different number of clusters.

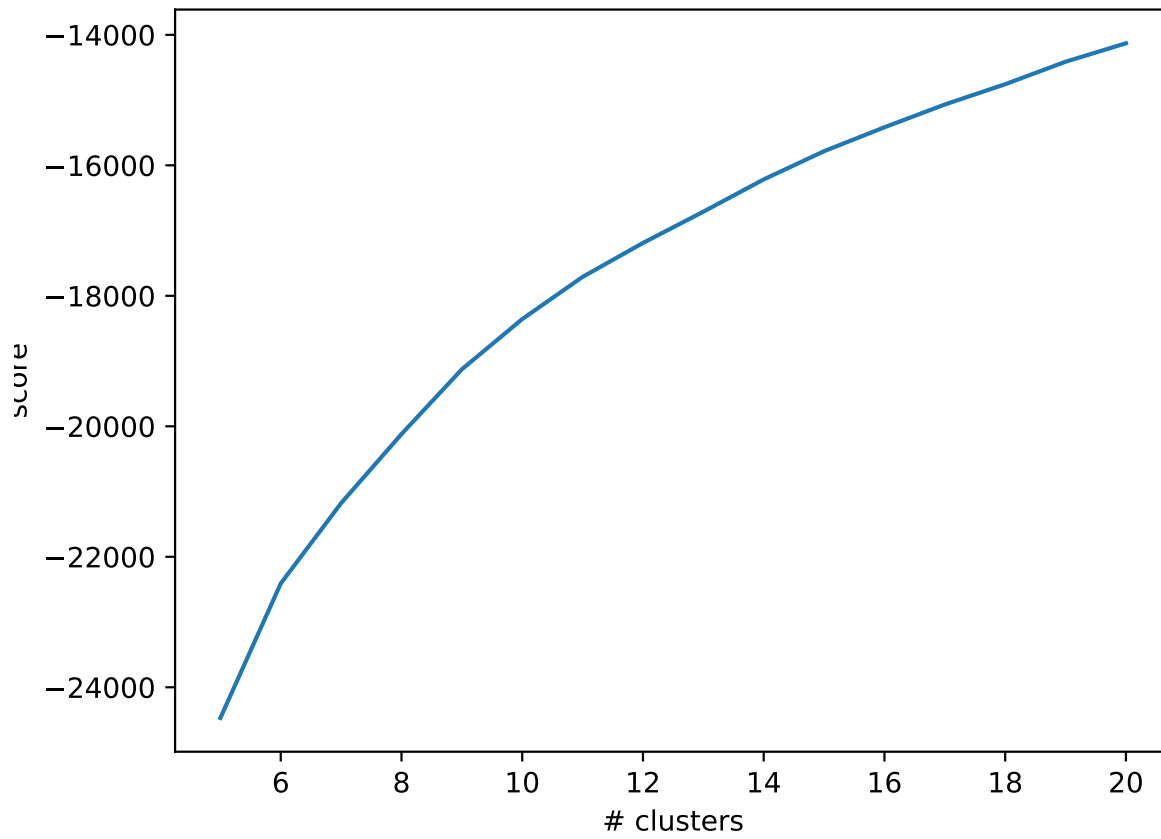


Figure S17. K-means scores