

Application of self-organizing maps to observed daily precipitation over the tropical and southern Pacific Ocean

M. Pike, B.R. Lintner, and M.J. Niznik

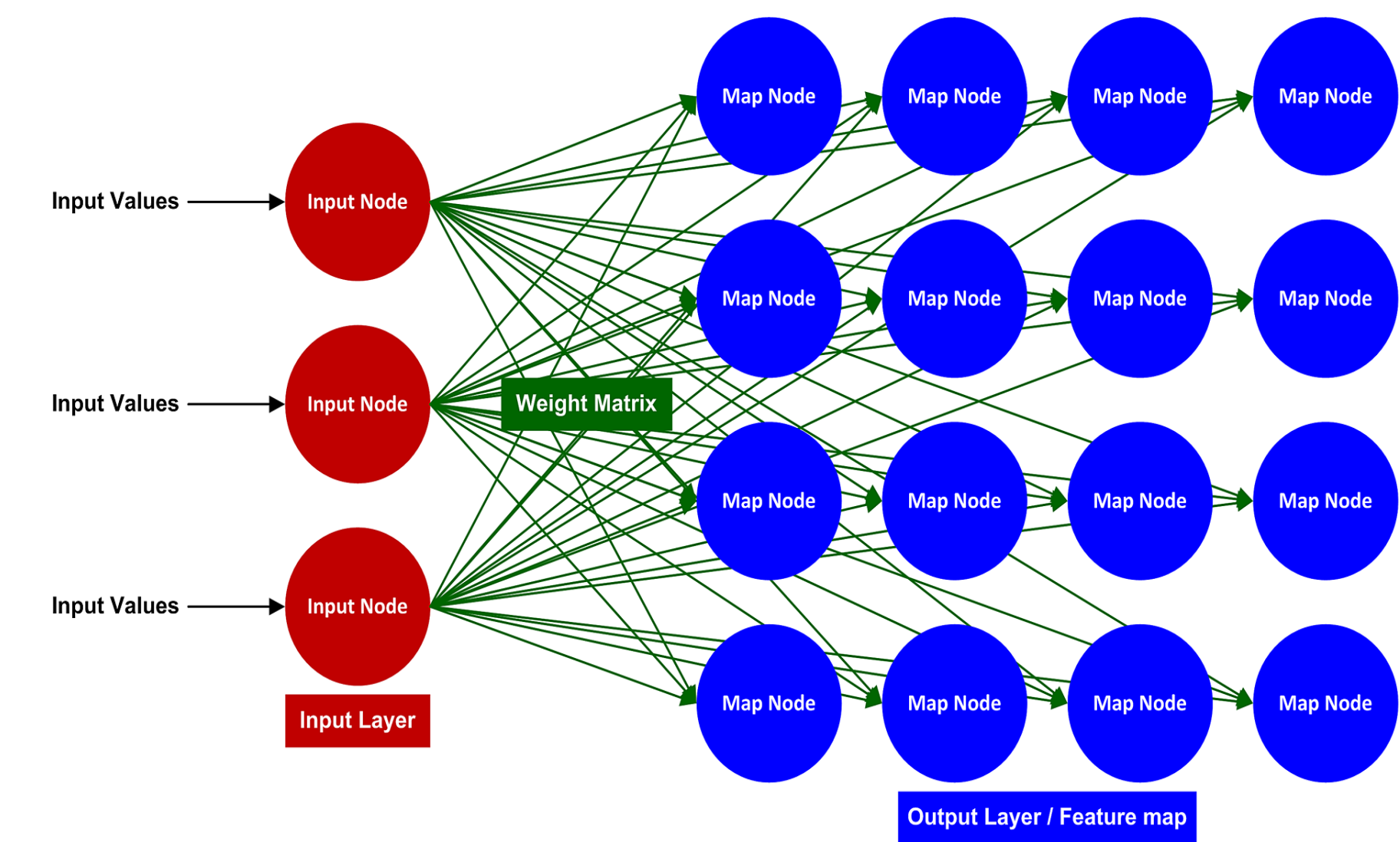
Department of Environmental Science, Rutgers, The State University of New Jersey

Abstract

Self-organizing maps (SOMs) comprise a class of artificial neural networks that aim to organize complex input data through computation of a set of M x N representative maps. Here we use an SOM routine to isolate the spatial patterns inherent in daily austral summer (December-January-February or DJF) rainfall over the tropical and southern Pacific Ocean basins from Tropical Rainfall Measuring Mission (TRMM) satellite observations. Applying a 2x2 SOM to all available DJFs from TRMM yields two maps that may be regarded as Intertropical Convergence Zone (ITCZ)-dominant, in which precipitation is more intense over the ITCZ region compared to the South Pacific Convergence Zone (SPCZ) region, while the remaining maps are SPCZ-dominant. The latter reflect a spatial translation of the SPCZ consistent with the previously described impact of the El Niño/Southern Oscillation (ENSO) or analogous low-frequency modes of variability on the SPCZ. Because of the pronounced impact of ENSO, we further consider SOMs computed separately for each of the El Niño and La Niña phases. This analysis suggests that while the overall position of the SPCZ is sensitive to the phase of ENSO, within each phase, similar high-frequency changes in SPCZ slope occur. Thus, while the mean position of the SPCZ may be dominantly controlled by ENSO phase, the distinct orientations within the same ENSO phase point to additional controls on SPCZ slope.

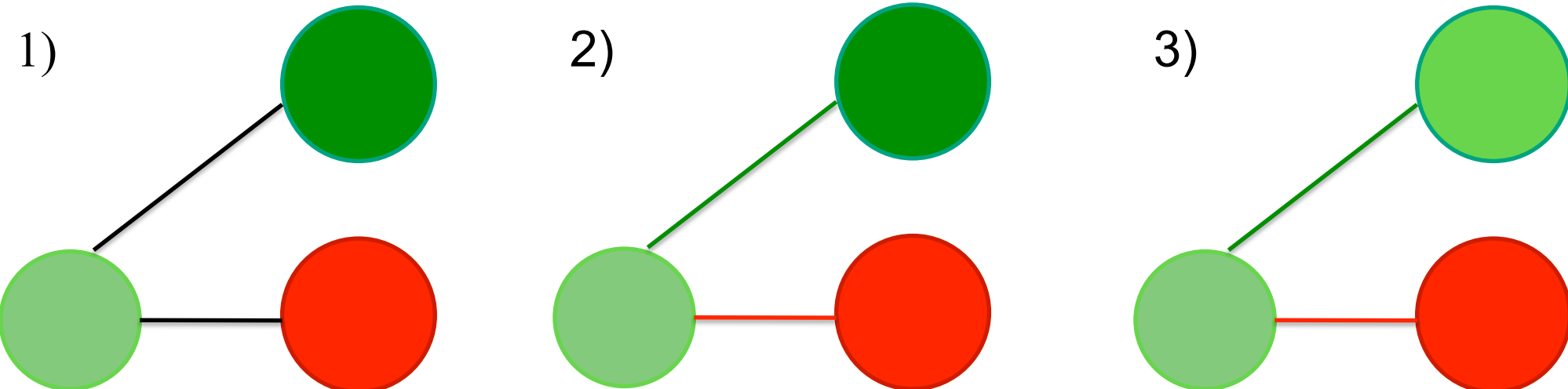
Methods

Schematic overview of SOMs



Initially, map nodes in the output layer are seeded with random weight values of the same dimensions as the data in the input layer. As the SOM iterates through the data along the time axis, the weights of each daily snapshot of precipitation is compared to the weights of each of the map nodes using the Euclidean Distance formula. The result for each of these comparisons using this formula is a single integer representing proximity in N-space; a smaller value indicating greater similarity between the weights, and thus greater similarity between the data at the two nodes. The map node returning the smallest value is known as the ‘winning node’, which then shifts its weights an arbitrary amount to more closely resemble the daily snapshot in the input layer. In addition to the weights of the ‘winning node’ adjusting, the weights of the winning node’s immediate neighbors may also be adjusted, but to a lesser degree. This process is called *cooperative learning*. The result is evident in the final output where the most similar nodes are clustered together, and the most dissimilar nodes are located furthest apart, typically at opposite corners of the output array. Below is a simple illustrative example.

SOM Example



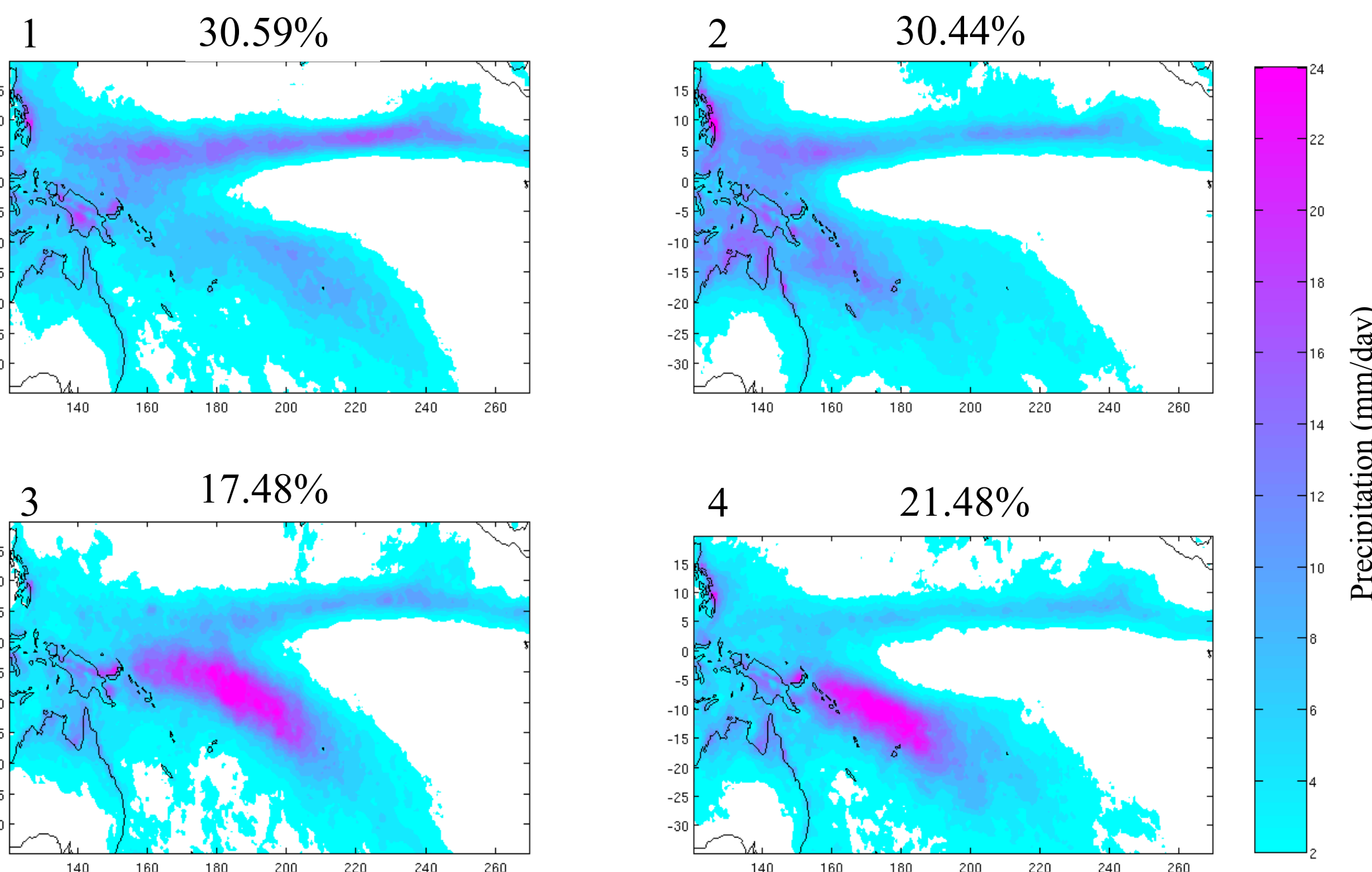
- Given an SOM of two map nodes, red and green, and a data point, light green, each containing three weighted values for amounts of red, blue and green (red, blue, green)
- Given:
 - red = (6,0,0)
 - green = (0,0,6)
 - light green = (3,6,3)
- (1) SOM computes Euclidean distance between light green and red and light green and green

$$d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$

- (2) Distance to green data point is smaller; green is ‘winning node’
- (3) Green node’s weights change an arbitrary amount to better represent the data level

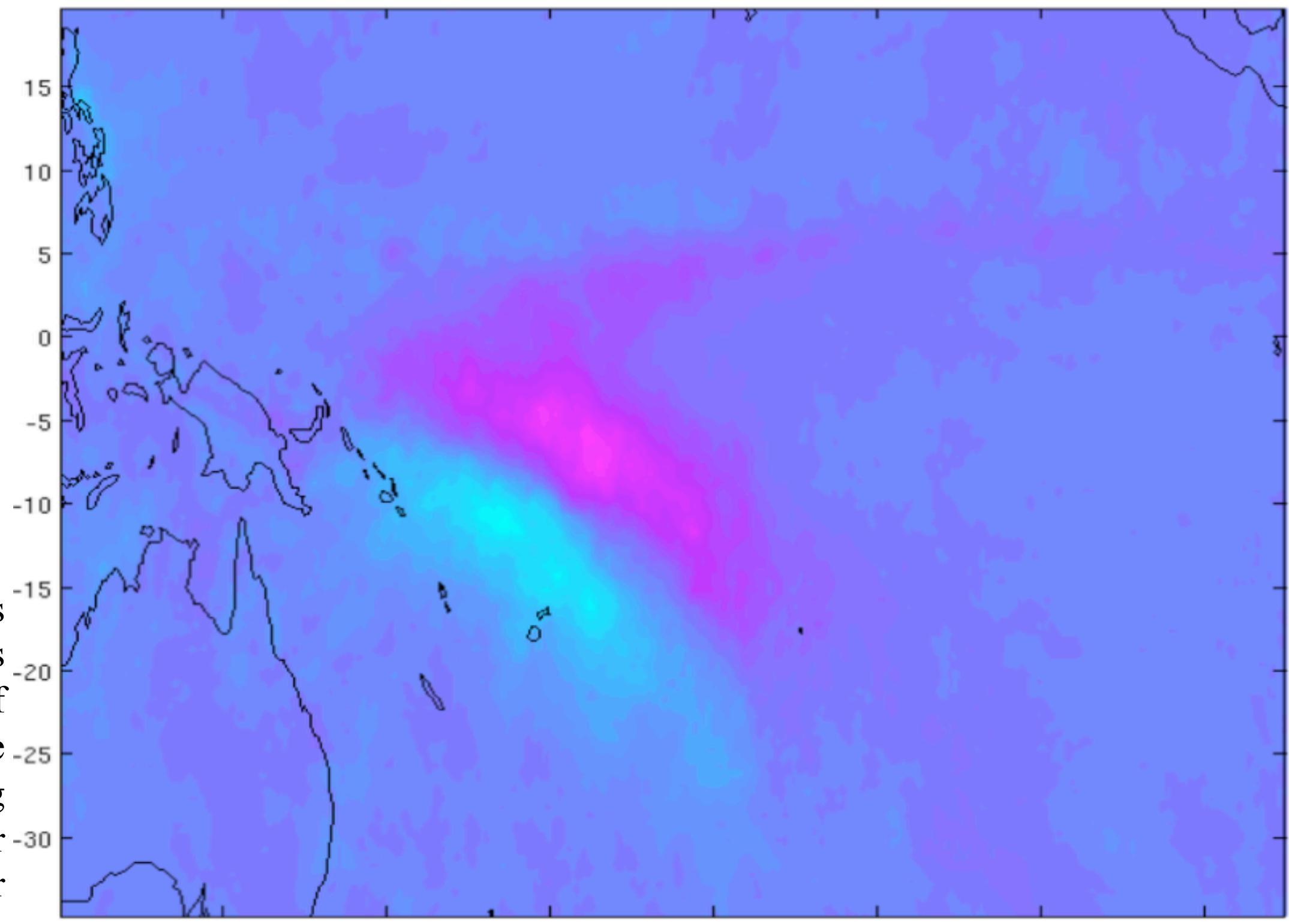
Results

DJF TRMM 1999-2013 SOM Analysis



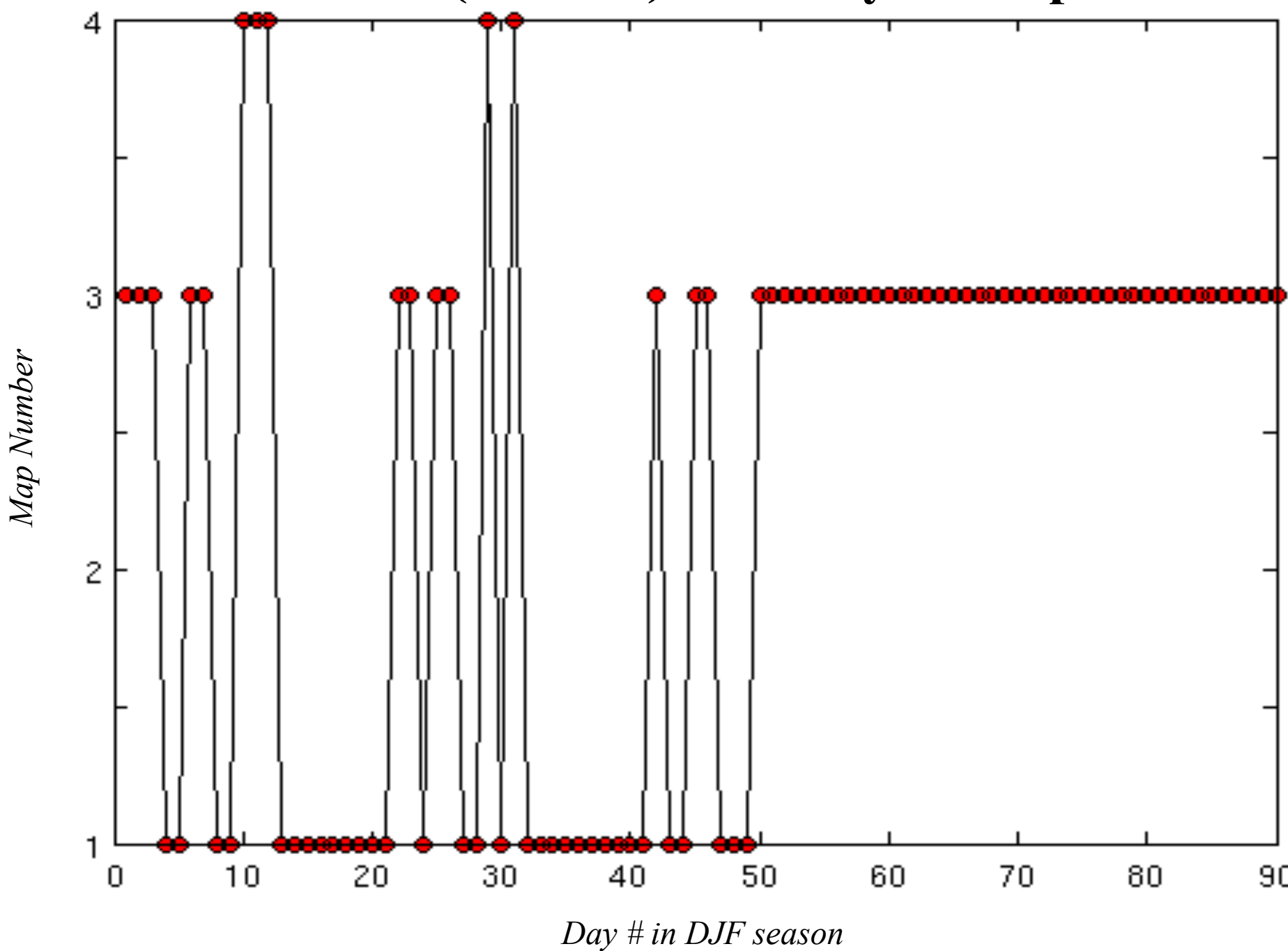
Maps 1 and 3 show an El Niño-like precipitation pattern based on the position and orientation of the SPCZ, while maps 2 and 4 indicate a more La Niña-like signature.

Difference of Maps 3 and 4



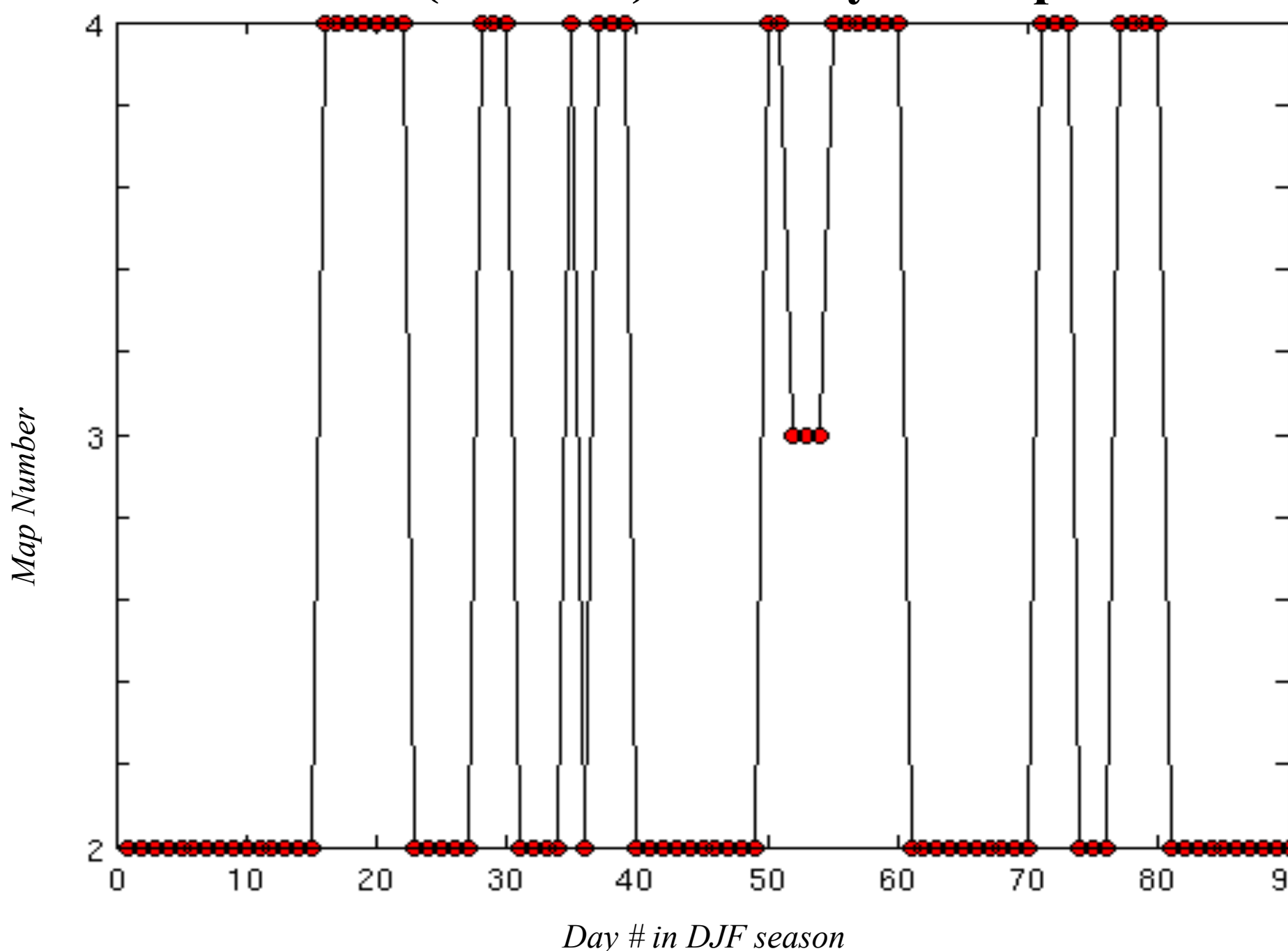
The difference of maps 3 and 4 more clearly illustrates the shift of the SPCZ between positive and negative ENSO phases: the light blue/cyan represents precipitation associated with La Niña conditions, while the pink is precipitation associated with El Niño conditions.

2010 (El Niño) DJF Day vs Map



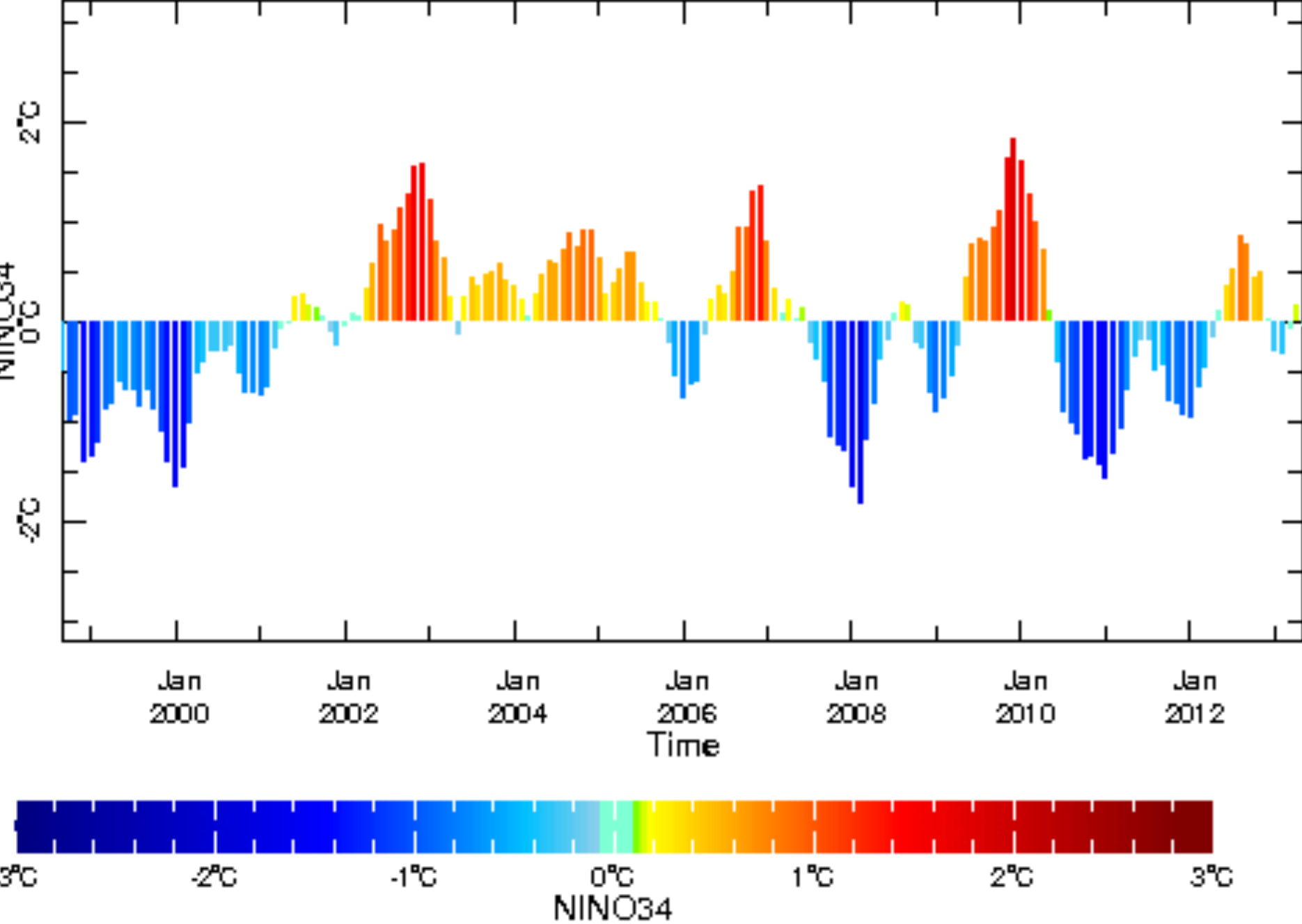
Days in Dec 2009 – Feb 2010 are primarily binned into maps 1 or 3, and when transitioning to a new map, usually go from 1 to 3, or 3 to 1

2011 (La Niña) DJF Day vs Map

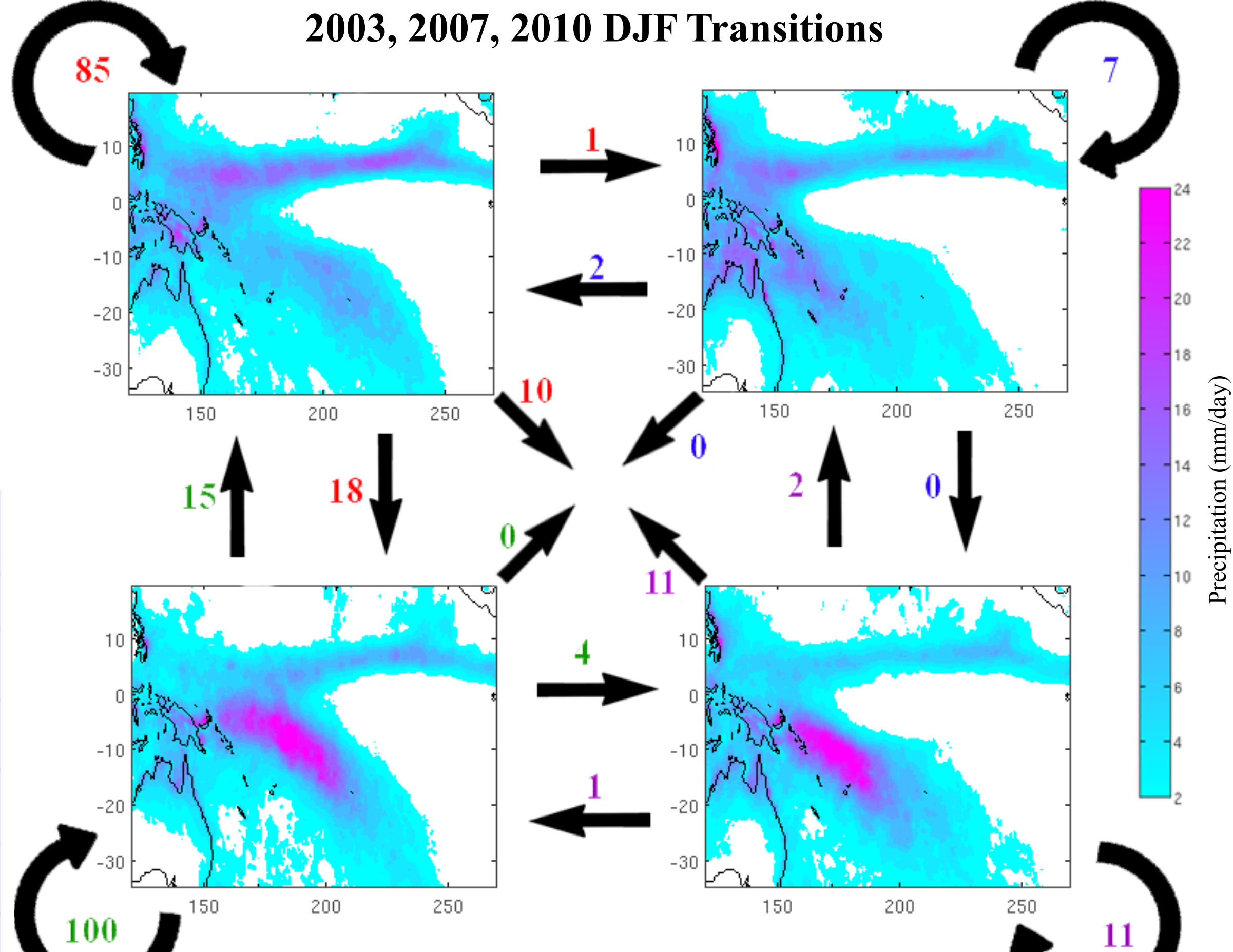


Days in Dec 2010– Feb 2011 are primarily associated with maps 2 and 4, with the dominant transitions from 2 to 4 or 4 to 2.

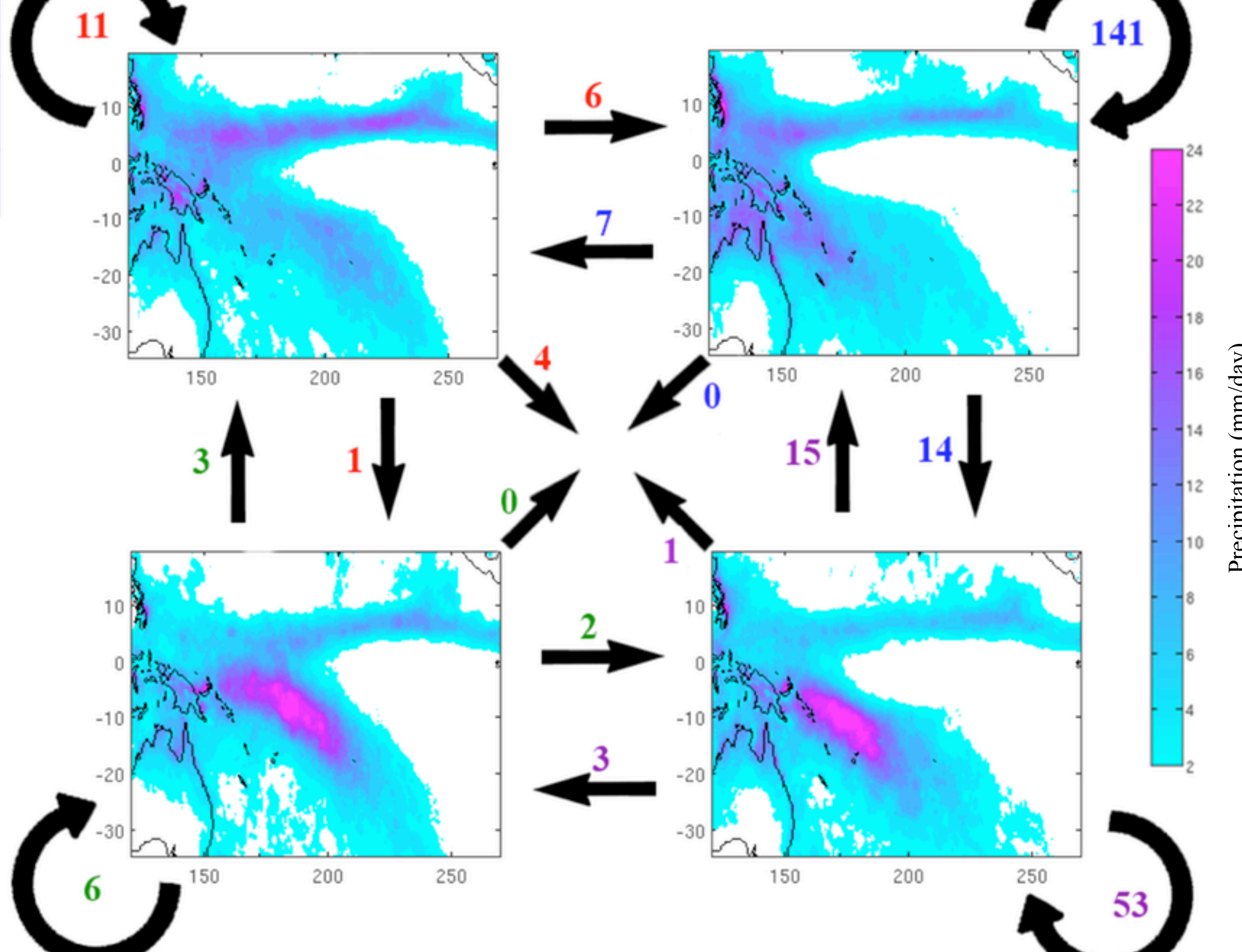
ENSO indices in Niño 3.4 region 1998-2013



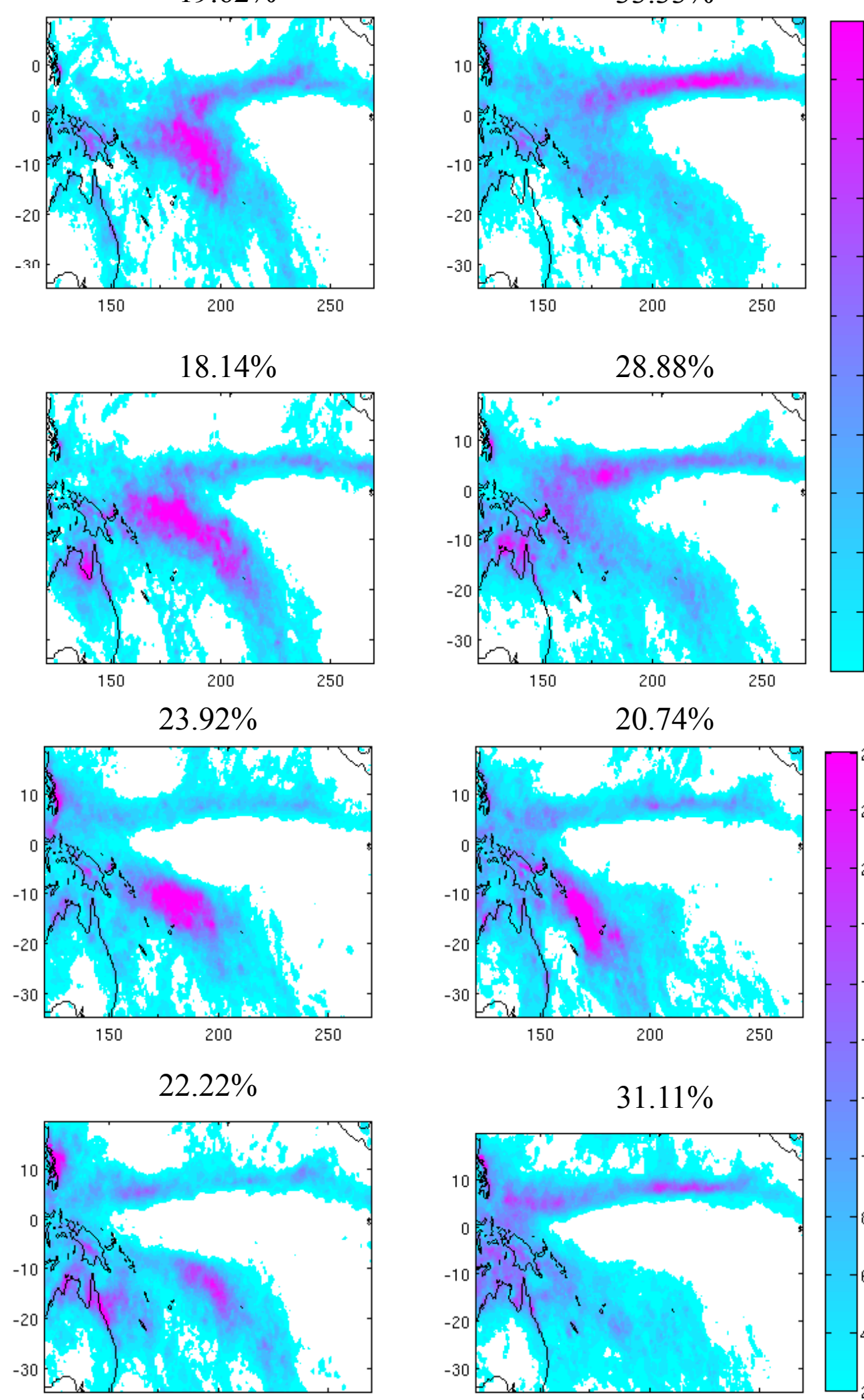
2003, 2007, 2010 DJF Transitions



2000, 2008, 2011 DJF Transitions



By considering the DJF data for the three strongest El Niño(-like) years, and three strongest La Niña(-like) years (regardless of whether minimum SST for an official ENSO event is met), and plotting the day to day transitions for the 270 days in each data set, it is evident that the SOM is correctly converting an otherwise noisy data set into four distinctive modes, with most of the transitions dominated by two output maps.



SOM Routine 2003, 2007, 2010 DJF

When the SOM is run on the data from the three most El Niño(-like) years, a precipitation pattern expected with a warm central Pacific develops, with an SPCZ shifted to the north and east. Strong SPCZ activity can be seen in maps 1 and 3, and strong ITCZ activity in maps 2 and 4.

SOM Routine 2000, 2008, 2011 DJF

When the SOM is run on the data from the three most La Niña(-like) years, a precipitation pattern expected with a cold central Pacific develops, with an SPCZ shifted to the west. Strong SPCZ activity can be seen in maps 1 and 2, and strong ITCZ activity in maps 3 and 4.