The Typhoon Tracks Analysis using Tri-plots and Markov Chain

Tseng, John Chien-Han
Central Weather Bureau, Taipei, Taiwan

Pao, Hsing-Kuo
Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology, Taipei, Taiwan

Faloutsos, Christos
School of Computer Science, Carnegie Mellon University, PA, USA

Jun/28/2010
El Niño, Normal(neutral), La Niña years

**ENSO:** El Niño southern oscillation

**Anti ENSO:** La Niña

different anomaly of sea surface temperature area →
different convective area →
different generative regions of typhoon →
different typhoon tracks
The purpose…

To classify the typhoon trajectories in a period of time belonged to El Niño or La Niña events.
Related work

• (Lee et al., 2007): use MDL (minimum description length) to decide the ensemble (average) path of typhoon tracks.

• (Camargo et al., 2007a, b): use trajectories clustering based on Gaffney and Smyth (1999) with mixtures regression model and arrange 7 clusters of typhoon tracks.

• (Risi, 2004): use Markov chain to predict the typhoon paths.

• (Vlachos and Kollios, 2002): combine longest common subsequence to discover similar multidimensional trajectories

**Common point:** every typhoon trajectory is regarded as one sequence.

We think we probably could check typhoons by collecting a period of time of typhoon cases to be one sequence.
Outline

• Motivation
• Methodology and data
  - tri-plots and fractal dimension
  - Markov chain
  - the data
• Performance of Tri-plots
• Performance of Markov chain
• Conclusions and discussions
These two diagrams have same fractal dimension.

\begin{align*}
\text{in } S_1, \ N &= 8, \ \varepsilon = \frac{1}{3} \\
\text{in } S_2, \ N &= 8^2, \ \varepsilon = \left(\frac{1}{3}\right)^2
\end{align*}

It implies

\[ d = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} = \frac{\log(8^n)}{\log(3^n)} = \frac{n \log 8}{n \log 3} = \frac{\log 8}{\log 3} \]
Tri-plots: cross-plot, self-plot

Assuming there are two datasets $A$ and $B$, and the cross-plot function is defined as:

$$Cross_{A,B}(r) = \log(\sum_i C_{A,i} \cdot C_{B,i})$$

where $C_{A,i}$ ($C_{B,i}$) is the number of points from set $A$ ($B$) in the $i$-th cell, and $r$ is the distance of the pairs of points. Hence, the cross-plot function is proportional to the count of $A$-$B$ pairs within distance $r$, and the cross-plot is the figure of the cross-plot function versus $\log(r)$.

Also, the self-plot function is defined as

$$Self_{A}(r) = \log(\sum_i \frac{C_{A,i} \cdot (C_{A,i} - 1)}{2})$$
The typical tri-plots result

When we compare two datasets $A$ and $B$, ...

Adopt from Traina et al. (2001)
Markov chain and dissimilarity measures (1)

• Given a trajectory \( s = (x_1, \ldots, x_t, \ldots, x_T) \) and Euclidean pace is \( \lambda = \| x_{t+1} - x_t \| \)

• We define

\[
\Delta \lambda_t = \lambda_{t+1} - \lambda_t,
\]

\[
\Delta \theta_t = \theta_{t+1} - \theta_t
\]

where \( \theta_t \) is an angle between vector \( x_{t+1} - x_t \) and x-axis.

• According Markov chain, we assume

\[
P(\lambda_{t+1} | \lambda_t) \sim N(\lambda_t, \sigma^2_\lambda) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \exp\left( -\frac{(\lambda_{t+1} - \lambda_t)^2}{2\sigma^2_\lambda} \right),
\]

\[
P(\theta_{t+1} | \theta_t) \sim N(\theta_t, \sigma^2_\theta) = \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\left( -\frac{(\theta_{t+1} - \theta_t)^2}{2\sigma^2_\theta} \right)
\]
Markov chain and dissimilarity measures (2)

Given a model $m$, the log-likelihood $\ell(s; m)$ of a trajectory can be

$$
\ell(s; m) = \log L(s; m) = \log(P(x_1) \prod_{t=1} P(x_{t+1} | x_t))
$$

$$
= \log P(x_1) + \sum_{t=1} \log(p(x_{t+1} | x_t))
$$

We define the code length of the trajectory $s$ as the negative logarithm likelihood as

$$
c(s|m) = -\ell(s; m) = -\log L(s; m)
$$

The dissimilarity between two trajectories $s_1$ and $s_2$ is

$$
d(s_1, s_2) = \frac{1}{2} \left( \frac{c(s_1|m_2)}{c(s_2|m_2)} + \frac{c(s_2|m_1)}{c(s_1|m_1)} \right)
$$

After calculating all trajectories pairs, we define a dissimilarity matrix between the pace and the angle as

$$
D_{\text{target}} = \alpha D_{\Delta \lambda} + (1 - \alpha) D_{\Delta \theta}
$$
The Data

- The Japan Meteorological Agency (JMA) typhoon tracks data from 1950 to 2009. The time resolution is about 3~6 hrs.
- The National Center for Environment Prediction (NCEP) reanalysis II data from 1980 to 2009.
- The El Niño, La Niña, or neutral events based on Niño 3.4 index are published by NOAA. (http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml)
Framework; flow chart

- Typhoon trajectories
- NCEP high level wind
- Topo effect

- Tri-plots Feature selection

- Self-plots Distribution cluster
- Cross-plots Distances

- ISOMAP

- Markov chain Pace vector, angle

- Disimilarity measure

- ISOMAP

- SSVM

- ISOMAP

- SSVM
ISOMAP (Isometric Feature Mapping)

This method was proposed by Tenebaum et al. (2000, Science)
SVM (Support Vector Machines), SSVM (Smooth Support Vector Machines)
### El Niño, La Niña, Neutral events

<table>
<thead>
<tr>
<th>ENSO</th>
<th>La Niña</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002/Apr~2003/Feb</td>
<td></td>
<td>1996/Apr~1997/Apr</td>
</tr>
<tr>
<td>2004/Jul~2005/Feb</td>
<td></td>
<td>2001/Mar~2002/Apr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2003/Apr~2004/May</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2005/Mar~2006/Jul</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2008/Jun~2009/Jul</td>
</tr>
</tbody>
</table>
Outline

• Motivation
• Methodology and data
  - tri-plots and fractal dimension
  - Markov chain
  - the data
• Performance of Tri-plots
• Performance of Markov chain
• Conclusions and discussions
Feature selection

First, we use four events to select the features:

- **ENSO**: 1982.5~1983.7
  1986.8~1988.2 (long period)
- **La Niña**: 1998.7~2000.7 (long period)
  2007.8~2008.6
1986/88 and 1998/00 Typhoon track (long period)
1982/83 and 2007/08
Typhoon track
(short period)
### Table 3.1: The different features compositions in the tri-plots experiments

<table>
<thead>
<tr>
<th>Feature Composition</th>
<th>((\lambda, \phi))</th>
<th>((\lambda, \phi, p))</th>
<th>((\lambda, \phi, u, v))</th>
<th>((\lambda, \phi, u, v, p))</th>
<th>((\Delta\lambda, \Delta\phi))</th>
</tr>
</thead>
<tbody>
<tr>
<td>The slope of self-plot</td>
<td>(1.735, 1.749)</td>
<td>(1.772, 1.877)</td>
<td>(3.118, 2.942)</td>
<td>(3.309, 2.887)</td>
<td>(1.645, 1.696)</td>
</tr>
<tr>
<td></td>
<td>(1.749, 1.777)</td>
<td>(1.772, 2.127)</td>
<td>(3.118, 3.108)</td>
<td>(3.309, 3.040)</td>
<td>(1.645, 1.696)</td>
</tr>
<tr>
<td></td>
<td>(1.772, 1.777)</td>
<td>(2.214, 2.127)</td>
<td>(3.275, 3.108)</td>
<td>(3.439, 3.040)</td>
<td>(1.676, 1.696)</td>
</tr>
<tr>
<td></td>
<td>(1.772, 1.735)</td>
<td>(2.214, 1.877)</td>
<td>(3.275, 2.942)</td>
<td>(3.439, 2.887)</td>
<td>(1.676, 1.696)</td>
</tr>
<tr>
<td>The slope of cross-plot</td>
<td>1.900</td>
<td>3.273</td>
<td>3.628</td>
<td>3.597</td>
<td>1.851</td>
</tr>
<tr>
<td></td>
<td>2.023</td>
<td>2.820</td>
<td>3.798</td>
<td>4.534</td>
<td>1.661</td>
</tr>
<tr>
<td></td>
<td>1.982</td>
<td>3.051</td>
<td>3.794</td>
<td>4.679</td>
<td>1.402</td>
</tr>
<tr>
<td></td>
<td>1.946</td>
<td>2.787</td>
<td>3.765</td>
<td>4.460</td>
<td>1.322</td>
</tr>
<tr>
<td><strong>Note</strong></td>
<td>Fig. 3.1</td>
<td>Fig. 3.2</td>
<td>Fig. 3.3</td>
<td>Fig. 3.4</td>
<td>not shown</td>
</tr>
</tbody>
</table>

### Table 3.1 (continued)

<table>
<thead>
<tr>
<th>Feature Composition</th>
<th>((\Delta\lambda, \Delta\phi, \Delta\varphi))</th>
<th>((\lambda, \phi, \Delta\lambda, \Delta\phi))</th>
<th>((\lambda(t - 1), \phi(t - 1), \lambda(t), \phi(t), \lambda(t + 1), \phi(t + 1)))</th>
</tr>
</thead>
<tbody>
<tr>
<td>The slope of self-plot</td>
<td>(2.110, 2.166)</td>
<td>(2.551, 2.804)</td>
<td>(1.673, 2.054)</td>
</tr>
<tr>
<td></td>
<td>(2.110, 2.019)</td>
<td>(2.551, 2.666)</td>
<td>(1.673, 2.324)</td>
</tr>
<tr>
<td></td>
<td>(2.159, 2.019)</td>
<td>(2.769, 2.666)</td>
<td>(1.901, 2.409)</td>
</tr>
<tr>
<td></td>
<td>(2.159, 2.166)</td>
<td>(2.769, 2.804)</td>
<td>(1.901, 1.954)</td>
</tr>
<tr>
<td>The slope of cross-plot</td>
<td>2.224</td>
<td>3.763</td>
<td>1.947</td>
</tr>
<tr>
<td></td>
<td>2.094</td>
<td>3.785</td>
<td>1.747</td>
</tr>
<tr>
<td></td>
<td>2.092</td>
<td>3.871</td>
<td>2.023</td>
</tr>
<tr>
<td></td>
<td>2.148</td>
<td>3.665</td>
<td>1.781</td>
</tr>
<tr>
<td><strong>Note</strong></td>
<td>not shown</td>
<td>not shown</td>
<td>not shown</td>
</tr>
</tbody>
</table>

*Note: use this idea in Markov chains*
Experiment 1: just use lon. and lat. $(\lambda, \phi)$
Experiment 2: \((\lambda, \phi, p)\)
Experiment 3: use lon., lat., u, v $(\lambda, \phi, u, v)$ 4-D

PCP self- and cross- plots

```
"tytrackhlwenso8283.ninp-tytrackhlwanina0708.ninp.pcp-ll" + 11.18 + x (3.628)
"tytrackhlwenso8283.ninp-tytrackhlwanina8283.ninp.pcp-ll" + 10.80 + x (3.118)
"tytrackhlwanina0708.ninp-tytrackhlwanina0708.ninp.pcp-ll" + 10.48 + x (2.942)
"tytrackhlwanina9800.ninp-tytrackhlwanina9800.ninp.pcp-ll" + 11.40 + x (3.108)
```

PCP self- and cross- plots

```
"tytrackhlwenso8688.ninp-tytrackhlwanina9800.ninp.pcp-ll" + 12.40 + x (3.794)
"tytrackhlwanino8688.ninp-tytrackhlwenso8688.ninp.pcp-ll" + 12.08 + x (3.275)
"tytrackhlwanina9800.ninp-tytrackhlwanina9800.ninp.pcp-ll" + 11.40 + x (3.108)
"tytrackhlwanina0708.ninp-tytrackhlwanina0708.ninp.pcp-ll" + 10.48 + x (2.942)
```

---

---

---

---
Experiment 4: use lon., lat., u, v, minp \((\lambda, \phi, u, v, p)\) 5-D

PCP self- and cross- plots

PCP self- and cross- plots

PCP self- and cross- plots

PCP self- and cross- plots

ack3dhlwens0823.ninp-tytrack3dhlwlanina0708.ninp.pcp-ll''
9.90 + x^*(3.597)

rack3dlwens0823.ninp-tytrack3dhlwlanina9800.ninp.pcp-ll''
10.75 + x^*(4.534)

ack3dhlwens0868.ninp-tytrack3dhlwlanina9800.ninp.pcp-ll''
11.37 + x^*(4.679)

rack3dlwens0868.ninp-tytrack3dhlwlanina9800.ninp.pcp-ll''
11.06 + x^*(3.439)

ckdhlwlanina9800.ninp-tytrack3dhlwlanina0708.ninp.pcp-ll''
10.19 + x^*(3.040)

ckdhlwlanina0708.ninp-tytrack3dhlwlanina9800.ninp.pcp-ll''
9.45 + x^*(2.887)
Topography effect, the geodistance calculation

- Use land, ocean (1, 0) grid (lon., lat.) data to be one mask function as follows:

\[
M(x = (\text{lon}, \text{lat})) = \begin{cases} 
1, & \text{if } x \text{ at land} \\
0, & \text{if } x \text{ at ocean}
\end{cases}
\]

- Define Gaussian function

\[
G(x) = e^{-\frac{(x-x_{tc})^2}{2\sigma^2}}
\]

where \(x=(\text{lon.}, \text{lat.})\), means one specific grid point;
and \(x_{tc}=(\text{lon.}, \text{lat.})\), means the typhoon center.

The standard deviation in the Gaussian function is \(\sigma = 2000km\)

- The distance is calculated by

\[
dist = \sum_{(80E,20S)}^{(120W,80N)} (x-x_{tc})M(x)G(x)
\]
The Distribution(1)

• The slope and intercept of self-plot or cross-plot can be regarded as the point in the space.
• If the datasets are quite different each other, then the distribution will be more dispersive or separative.
• We use the distribution of self-plot to decide what kind of feature we will use in classification. Then, the distances between different cross-plots points are the distances of ISOMAP. After ISOMAP, we use SSVM to do classification.
The Distribution

The dist. of (m,b) by \((\lambda, \phi)\)

The dist. of (m,b) by \((\lambda, \phi, p)\)

The dist. of (m,b) by \((\lambda, \phi, u, v, p)\)

The dist. of (m,b) by \((\lambda, \phi, u, v, p, \text{topo})\)
The neutral year events...
Why we do not use 500 hPa wind...

The dist. of \((m,b)\) by \((\lambda, \phi, u500, v500, p, \text{topo})\)
9596, 8586 events

Almost same slope and same intercept…
9798, 0203 events

Same slope but different intercept...
8889, 9800 events

Same slope but different intercept…
Figure 3.7: The 2-D structure from ISOMAP with $k=6$ based on cross-plots experiments.
Table 3.2: Error table of the SSVM based on cross-plots and ISOMAP+

<table>
<thead>
<tr>
<th>$k$</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSVM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training error</td>
<td>0.268</td>
<td>0.032</td>
<td>0.116</td>
<td>0.025</td>
<td>0.023</td>
<td>0.027</td>
</tr>
<tr>
<td>Testing error</td>
<td>0.334</td>
<td>0.271</td>
<td>0.358</td>
<td>0.289</td>
<td>0.320</td>
<td>0.327</td>
</tr>
</tbody>
</table>
Outline

• Motivation
• Methodology and data
  - tri-plots and fractal dimension
  - Markov chain
  - the data
• Performance of Tri-plots
• Performance of Markov chain
• Conclusions and discussions
The distribution of $\Delta \lambda$ and $\Delta \theta$
The 2-D structure from ISOMAP with $k=4$ based on Markov Chain experiments with $\alpha=0.6$ in dissimilarity matrix calculation.
Table 4.1: Error table of the SSVM based on Markov chains and ISOMAP\((k=4)\).

<table>
<thead>
<tr>
<th>α</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSVM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training error</td>
<td>0.246</td>
<td>0.184</td>
<td>0.243</td>
<td>0.272</td>
<td>0.122</td>
<td>0.195</td>
</tr>
<tr>
<td>Testing error</td>
<td>0.271</td>
<td>0.267</td>
<td>0.308</td>
<td>0.284</td>
<td>0.248</td>
<td>0.264</td>
</tr>
</tbody>
</table>
Figure 4.x: The 2-D structure from ISOMAP with $k=4$ based on Markov Chain experiments with $\alpha=0.8$ in dissimilarity matrix calculation.
Threshold($\Delta p>10$) + ISOMAP + SSVM (2)

Table 4.1: Error table of the SSVM based on Markov chain and ISOMAP($k=4$).

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSVM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training error</td>
<td>0.173</td>
<td>0.052</td>
<td>0.135</td>
<td>0.048</td>
<td>0.161</td>
<td>0.168</td>
<td>0.154</td>
<td>0.037</td>
<td>0.031</td>
<td>0.004</td>
<td>0.045</td>
</tr>
<tr>
<td>Testing error</td>
<td>0.277</td>
<td>0.253</td>
<td>0.248</td>
<td>0.223</td>
<td>0.287</td>
<td>0.287</td>
<td>0.280</td>
<td>0.196</td>
<td>0.181</td>
<td>0.197</td>
<td>0.248</td>
</tr>
</tbody>
</table>
Conclusions and discussions (1)

• We have about 70% accuracy to classify the typhoon tracks belonged to El Niño or La Niña events based on tri-plots and Markov chain.
• The features of classification in tri-plots are \((\lambda, \phi, u, v, p, \text{topo})\), which can be used or considered in typhoon databases or typhoon tracks predictions.
• Either tri-plots or Markov chain, according to the distributions, afford the objective view to cluster the typhoon tracks of the different annual year events.
• The different features; for example, how to describe the surrounding circulations of typhoon like invariant moments, and the new features should be considered and designed in the following experiments.
• The different probability relation in Markov chain like Chi-square distribution is worth of rechecking again, even currently we cannot get the better results. Also, we think Markov chain still have more potential ability for classification.
• How to implement these methods to other traditional Meteorological data.
Conclusions and discussions (2)

• The labeled events are too less (7 ENSO, 5 La Nina), even the records of individual event seems enough (>=1,000). At the same time, we do not want to slice one typhoon case into different trajectory sequence, so we just split the original 12 events into 60 events, and the number of the events is still too less.

Thanks!


