A Hybrid Statistical Downscaling Approach Based on Nonstationary Time Series Decomposition

Xintong Li, Xiaodong Zhang
School of Environmental Science and Engineering
Shandong University
Outline

• The stationary hypothesis in SD
• Nonstationary Time Series Decomposition
• A Hybrid SD model: BEMD-rf
• The model performance
• Conclusion
Fine-scale

Large-scale

past

Statistical relationship

now

future scenarios

The underlying stationary hypothesis in SD

(Dixon, Lanzante et al. 2016)

Will this relationship be truly stationary in the future?
The climate is changing

A rising temperature

Severe drought and precipitation

• A changing climate make the hypothesis less reliable
• It is a big challenge to make SD method adapt to future non-stationary climate conditions
Nonstationary Time Series Decomposition

• Non-stationarity is an inbuilt trait embedded in climate system at different spatial and temporal scales

• A non-stationary time series decomposition was considered from the point of time-frequency analysis

  ➢ to separate interactions of different climate cycles
  
  ➢ capture stationary weather patterns
## A comparation among main time-frequency transforming

<table>
<thead>
<tr>
<th>Time-frequency Transforming</th>
<th>Fourier</th>
<th>Wavelets</th>
<th>Hilbert-Huang</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basis</td>
<td>a priori</td>
<td>a priori</td>
<td>adaptive</td>
</tr>
<tr>
<td>Frequency</td>
<td>convolution: global, uncertainty</td>
<td>convolution: regional, uncertainty</td>
<td>differentiation: local, certainty</td>
</tr>
<tr>
<td>Presentation</td>
<td>energy-frequency</td>
<td>energy-time-frequency</td>
<td>energy-time-frequency</td>
</tr>
<tr>
<td>Nonlinear Non-stationary</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Feature Extraction</td>
<td>no</td>
<td>discrete: no, continuous: yes</td>
<td>yes</td>
</tr>
<tr>
<td>Theoretical Base</td>
<td>theory complete</td>
<td>theory complete</td>
<td>empirical</td>
</tr>
</tbody>
</table>

* Frequency domain decomposition
Empirical Mode Decomposition (EMD, Huang et al., 1998)
a part of Hilbert-Huang transform, realizing a disintegration
\[ Y = IMF_1 + \cdots + IMF_n + \text{residual} \]

IMF, intrinsic mode functions
- IMFs can be seen as stationary components of original sequence in different time scales
- A residual component is a trend term

To overcome the shortcoming of EMD, losing synchronous time scales for multivariable decomposition, bivariate EMD (BEMD, Flandrin et al., 2004) was adopted.
A Hybrid SD model: BEMD-rf

Step 1: Data Collecting
- Ground max/min temperature, precipitation
- GCM max/min temperature, precipitation

Step 2: Data Partitioning
- Training datasets: 1950-1995
- Validating datasets: 1996-2005

Step 3: Nonstationary time series decomposition
- Bivariate Ensemble Empirical Mode Decomposition

Step 4: Select the optimal decomposition
- Ground IMFs
- GCM IMFs

Step 5: Training IMFs at multi-time scales
- Random Forest

Step 6: Nonstationary time series synthesis
- IMFs synthesis

Step 7: Model evaluation
- Downscaled max/min temperature, precipitation

Intrinsic Mode Functions (IMFs)
The model performance

- The decomposition of monthly *max temperature*

```
IMF1
IMF2
IMF3
IMF4
IMF5
IMF6
IMF7
IMF8
IMF9
Residual
```

Observation CanESM2

“Mode mixing” - loosing physical meaning
• Results of Significance Test

To ensure the physical meaning of decomposition

Good IMFs $\rightarrow$ with high information and less noises

Most components passed the test
- Downscaling performance

**Davis**

- The CanESM2 overestimates both the TMAX and TMIN
- The trends of downscaled TMAX and TMIN are consistent with those of observation
- A poor performance for precipitation, weak ability on depicting extreme values
- The proposed model has good generalization ability on topography
• Stationary analysis

- IMF with high variance contribution rate are stationary
- For other IMFs, they may be quasi-stationary
• Physical meaning of periods

<table>
<thead>
<tr>
<th></th>
<th>IMF1</th>
<th>IMF2</th>
<th>IMF3</th>
<th>IMF4</th>
<th>IMF5</th>
<th>IMF6</th>
<th>IMF7</th>
<th>IMF8</th>
<th>IMF9</th>
<th>IMF10</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAVIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRCP</td>
<td>STA_IMF</td>
<td>2.9</td>
<td>4.4</td>
<td>7.4</td>
<td><strong>12.4</strong></td>
<td>23.2</td>
<td>33.6</td>
<td>74.7</td>
<td>134.4</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.9</td>
<td>4.4</td>
<td>7.2</td>
<td><strong>12.4</strong></td>
<td>21</td>
<td>39.5</td>
<td>56</td>
<td>134.4</td>
<td>224</td>
</tr>
<tr>
<td>TMAX</td>
<td>STA_IMF</td>
<td>2.7</td>
<td>4.9</td>
<td><strong>12</strong></td>
<td>18.16</td>
<td>29.2</td>
<td>56</td>
<td>112</td>
<td>224</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.7</td>
<td>4.9</td>
<td><strong>12</strong></td>
<td>18.16</td>
<td>29.2</td>
<td>56</td>
<td>112</td>
<td>224</td>
<td>336</td>
</tr>
<tr>
<td>TMIN</td>
<td>STA_IMF</td>
<td>2.6</td>
<td>4</td>
<td>6.9</td>
<td><strong>12</strong></td>
<td>19.2</td>
<td>29</td>
<td>56</td>
<td>134.4</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.6</td>
<td>4</td>
<td>6.7</td>
<td><strong>12</strong></td>
<td>20.4</td>
<td>32</td>
<td>74.7</td>
<td>134.4</td>
<td>224</td>
</tr>
<tr>
<td>TAHOE</td>
<td>STA_IMF</td>
<td>2.7</td>
<td>4.2</td>
<td>6.2</td>
<td><strong>12.2</strong></td>
<td>21</td>
<td>33.6</td>
<td>61.1</td>
<td>112</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.6</td>
<td>4</td>
<td>5.9</td>
<td><strong>12.2</strong></td>
<td>21</td>
<td>33.6</td>
<td>56</td>
<td>112</td>
<td>168</td>
</tr>
<tr>
<td>TMAX</td>
<td>STA_IMF</td>
<td>2.6</td>
<td>5.3</td>
<td><strong>12</strong></td>
<td>20.4</td>
<td>33.6</td>
<td>61.1</td>
<td>134.4</td>
<td>168</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.7</td>
<td>5.5</td>
<td><strong>12</strong></td>
<td>20.4</td>
<td>35.4</td>
<td>61.1</td>
<td>134.4</td>
<td>224</td>
<td>224</td>
</tr>
<tr>
<td>TMIN</td>
<td>STA_IMF</td>
<td>2.6</td>
<td>4.4</td>
<td><strong>12</strong></td>
<td>16</td>
<td>26.9</td>
<td>51.7</td>
<td>74.7</td>
<td>168</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>GCM_IMF</td>
<td>2.6</td>
<td>4.9</td>
<td><strong>12</strong></td>
<td>16.4</td>
<td>30.5</td>
<td>42</td>
<td>74.7</td>
<td>134.4</td>
<td>224</td>
</tr>
</tbody>
</table>

*The periods were averaged periods and not precise.

• A strong annual cycle control the local weather mode.
• For the decadal cycles, they may influenced by the sea surface temperatures of tropical and mid-latitude regions.

(Joseph et al. 2000)
Conclusions

a) The interactions of different climatic cycles were separated. BEMD can decompose observations and GCM simulations into synchronized IMFs which are correspond to physical processes of climate at different time scales.

b) The decomposed stationary components can capture most variations of temperature and precipitation. The proposed BEMD-RF model outperforms the downscaling model without BEMD and can regain the observation sequences.

c) Because precipitation is more sensitive to sub-diurnal and synoptic modes than temperature and the coarse time resolution of monthly data used in this study which cannot capture a cycle with fine time scale, the downscale results of temperature are better than those of precipitation.

d) The inherent weakness including mode mixing and problem of boundary condition of interpolating envelop of BEMD can affect the decomposition results which inevitable brings errors in the proposed downscaling method.
School of Environmental Science and Engineering
Shandong University

Xintong Li, PhD student
contact: lxt_sdu@outlook.com