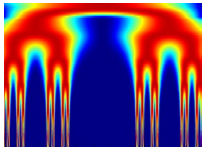
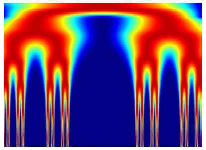


Multiscale Recurrence Analysis of Long-Term Nonlinear and Nonstationary Time Series

Yun Chen
Hui Yang, Ph.D.

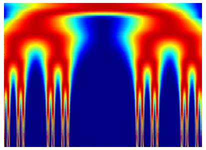
Industrial & Management Systems Engineering
University of South Florida

Relevant Publications



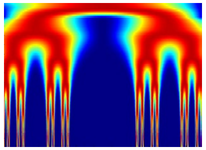
- **H. Yang***, “Multiscale Recurrence Quantification Analysis of Spatial Vectorcardiogram (VCG) Signals,” ***IEEE Transactions on Biomedical Engineering***, Vol. 58, No. 2, p339-347, 2011, DOI: 10.1109/TBME.2010.2063704
- Y. Chen† and **H. Yang***, "Multiscale recurrence analysis of long-term nonlinear and nonstationary time series," ***Chaos, Solitons and Fractals***, Vol. 45, No. 7, p978-987, 2012, DOI: 10.1016/j.chaos.2012.03.013
- **H. Yang***, S. T. S. Bukkapatnam, and L. G. Barajas, “Local recurrence model for performance prediction and prognostics in nonlinear and nonstationary systems,” ***Pattern Recognition***, Vol. 44, No. 8, p1834-1840, 2011, DOI: 10.1016/j.patcog.2011.01.010
- Y. Chen†, and **H. Yang***, “Wavelet packet analysis of disease-altered recurrence dynamics in the long-term spatiotemporal vectorcardiogram (VCG) signals”, *Proceedings of 2013 IEEE Engineering in Medicine and Biology Society Conference (EMBC)*, p. 2595-2598, July 3-7, 2013, Osaka, Japan. DOI: 10.1109/EMBC.2013.6610071

Outline



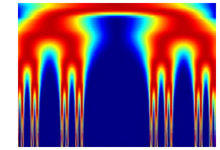
- ❑ Problem Statement
- ❑ Research Background
  - ➢ Wavelet transform
  - ➢ Recurrence analysis
- ❑ Research Methodology
  - ➢ Multiscale recurrence analysis
  - ➢ Feature selection
  - ➢ Randomized classification experiments
- ❑ Experimental Designs & Results
- ❑ Conclusions

Problem Statement

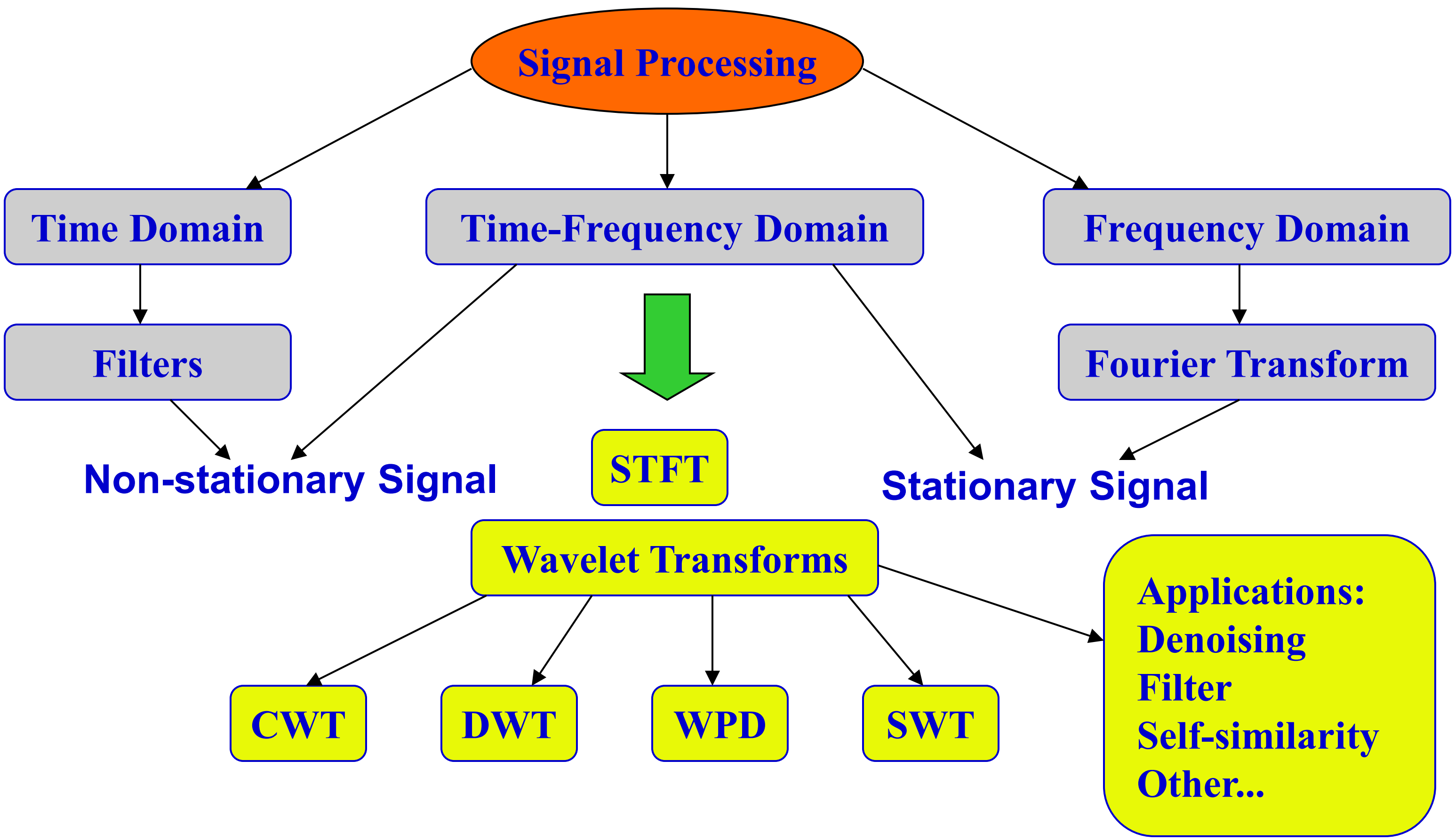


- **Challenges:**
  - **Data complexity: nonlinearity** and **nonstationarity**
  - **Enormous data torrents**
- **State of the Art:**
  - Conventional frequency-domain and linear system approaches
  - Nonlinear stochastic dynamics under highly nonstationary conditions
  - Nonlinear dynamic methods – **computational expensive**
- **Goals:**
  - Address the issues of nonlinearity, nonstationarity and large datasets
  - Extended and integrated into other nonlinear dynamic approaches
  - Disease-altered nonlinear dynamics

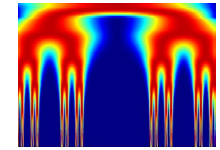
Time-Frequency Representation

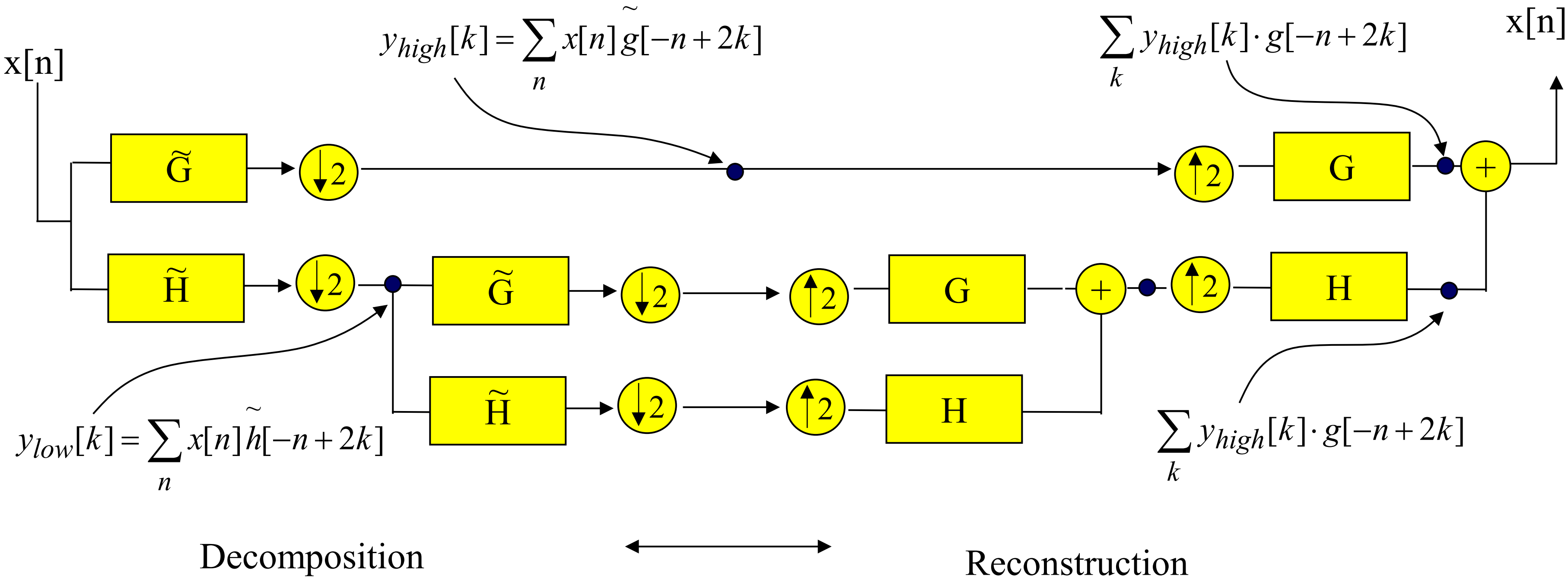


- Time info difficult to interpret in frequency domain
- Frequency info difficult to interpret in time domain
- Perfect time info in time domain, perfect freq. info in freq. domain



Discrete Wavelet Transform





G **Half band high pass filter**

H **Half band low pass filter**

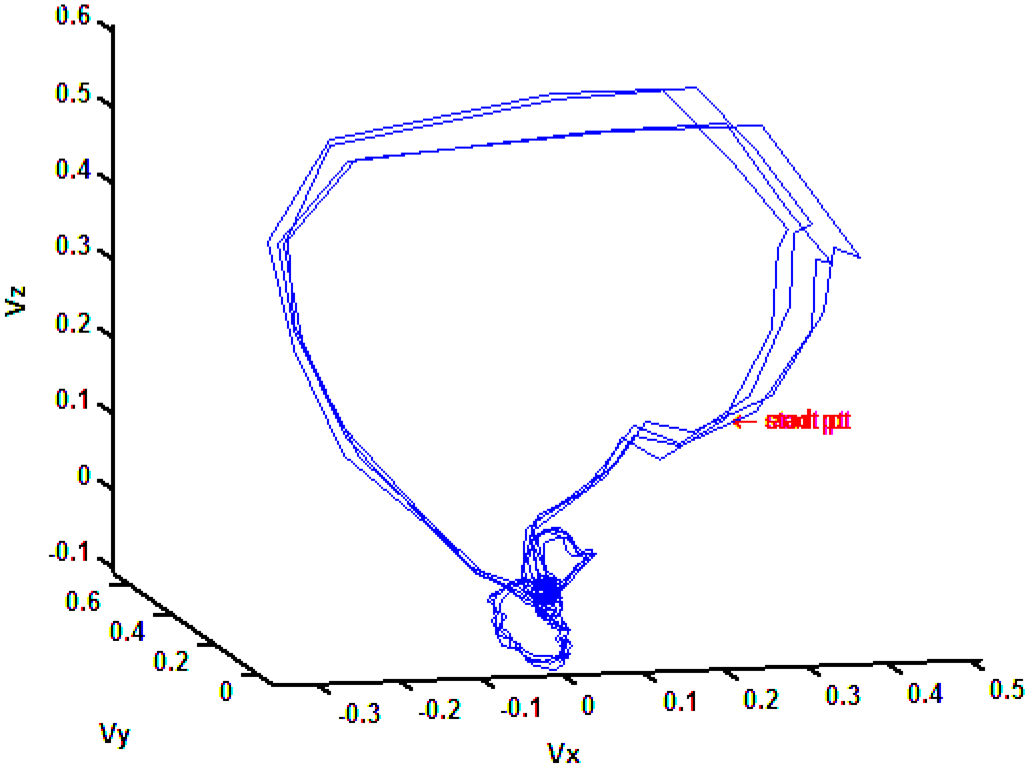
↓2 **Down-sampling**

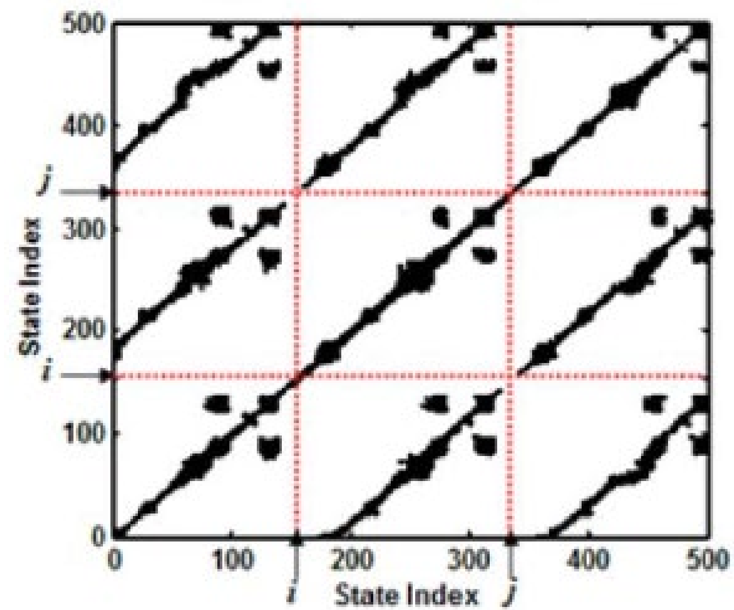
↑2 **Up-sampling**

2-level DWT decomposition & wavelet filter bank

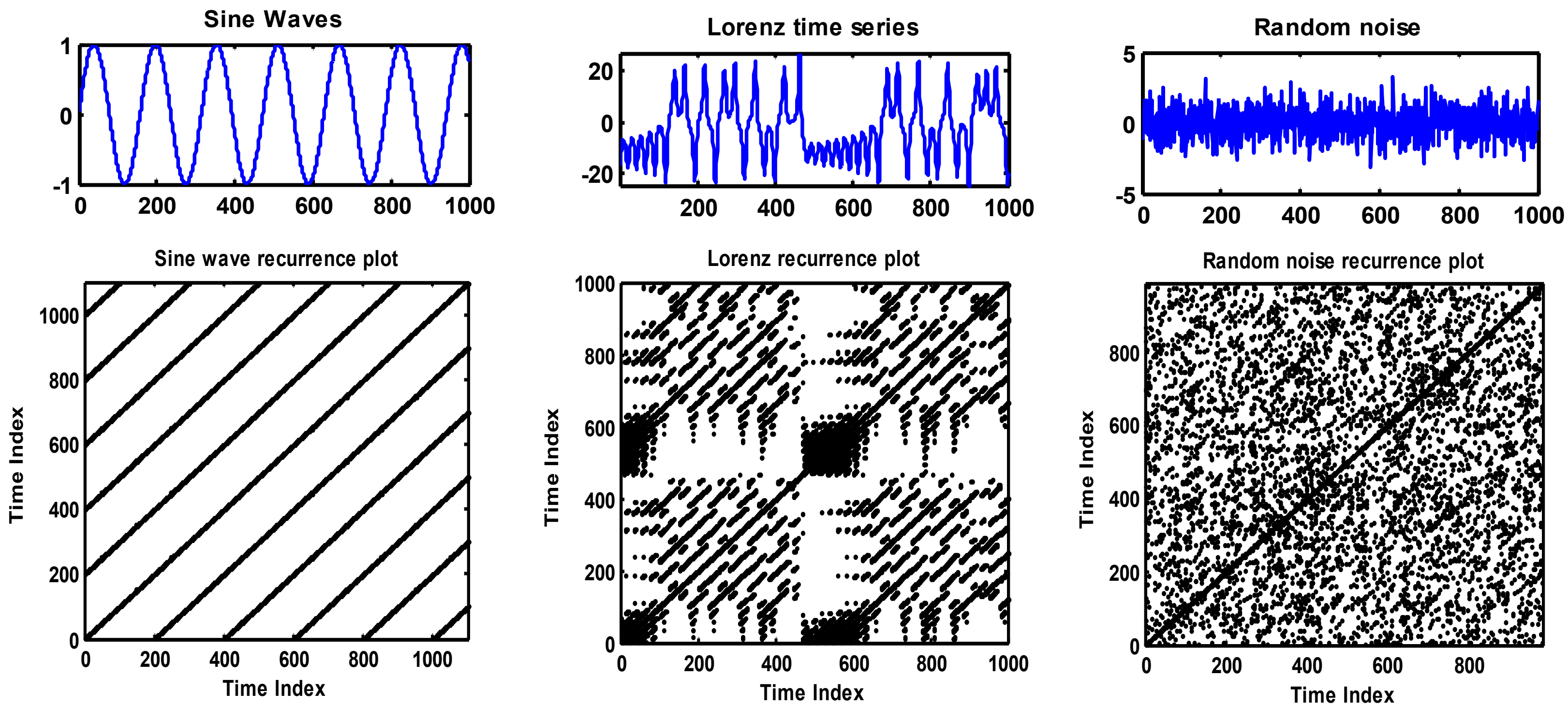
Recurrence Analysis

- Recurrence patterns of the dynamical systems
  - Recurrence plot: $R(i,j) = \Theta(\varepsilon - \|\boldsymbol{x}(i) - \boldsymbol{x}(j)\|)$





Structures in Recurrence Plots



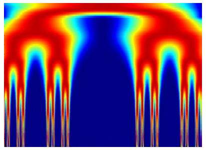
Linear System **Nonlinear System** **Randomized System**

- **Small-scale** structures
  - single dots, diagonal and vertical lines
- **Large-scale** structures
  - homogenous, periodic and disrupted visualization

Recurrence Features

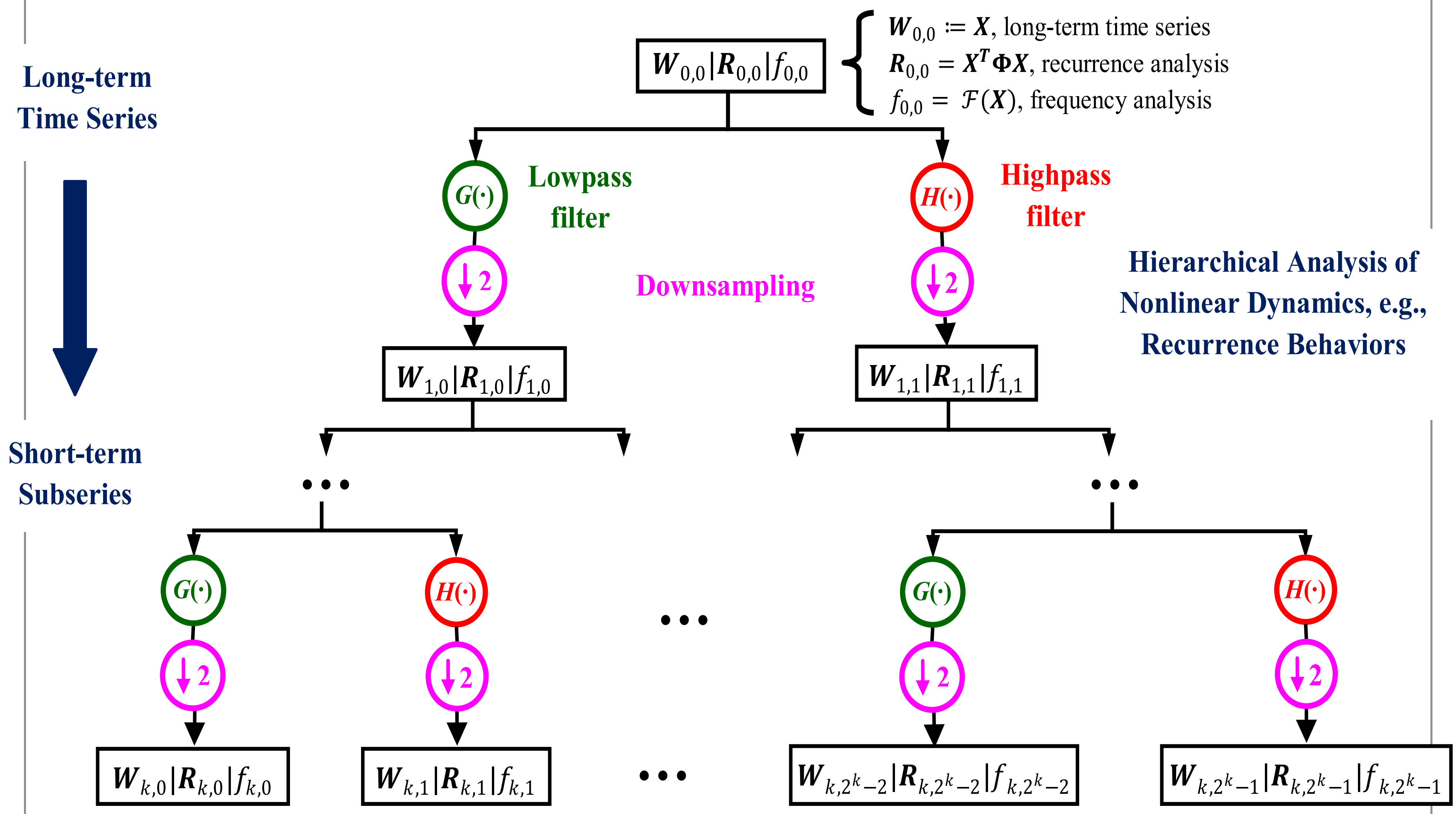
- Quantifying the *topological features of Recurrence Plots*
- **Statistical features** to quantify certain recurrence patterns from Threshold Recurrence Plot (Kantz, Marwan, and Kurths et al.):
  - Recurrence rate (%REC)
  - Determinism (%DET)
  - Linemax (LMAX)
  - Entropy (ENT)
  - Laminarity (%LAM)
  - Trapping time (TT)
- **Diagonal structures** (first four) and **vertical structures** (last two) in the threshold recurrence plot
- **Computational complexity:** square increase

Outline



- ❑ Problem Statement
- ❑ Research Background
  - ➢ Wavelet transform
  - ➢ Recurrence analysis
- ❑ **Research Methodology**
  - ➢ Multiscale recurrence analysis
  - ➢ Feature selection
  - ➢ Randomized classification experiments
- ❑ Experimental Designs & Results
- ❑ Conclusions

Hierarchical Nonlinear Dynamics



Completeness of MRA

- Given a time series $\boldsymbol{X} = \{x_1, x_2, \cdots, x_N\}^T$
- Embedded state space $\boldsymbol{x}(i) = \left(x_i, x_{i+\tau}, \cdots, x_{i+\tau(M-1)}\right)$
- Recurrence distance matrix ⇦ time series:

$$UR_{0,0}^2(i,j) = \|\boldsymbol{x}(i) - \boldsymbol{x}(j)\|^2 = \sum_{m=0}^{M-1} \left|x_{i+m\tau} - x_{j+m\tau}\right|^2$$

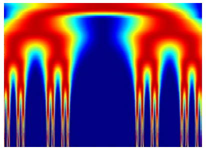
$$= \sum_{m=0}^{M-1} \left(\boldsymbol{X}^T \Phi_{i+m\tau, j+m\tau} \boldsymbol{X}\right)$$

$\Phi_{i,j}$ positive semidefinite, 1 in the ii^{th} and jj^{th} elements, -1 in the ij^{th} and ji^{th} elements and 0 otherwise.

- Recurrence distance matrix ⇦ wavelet subseries:

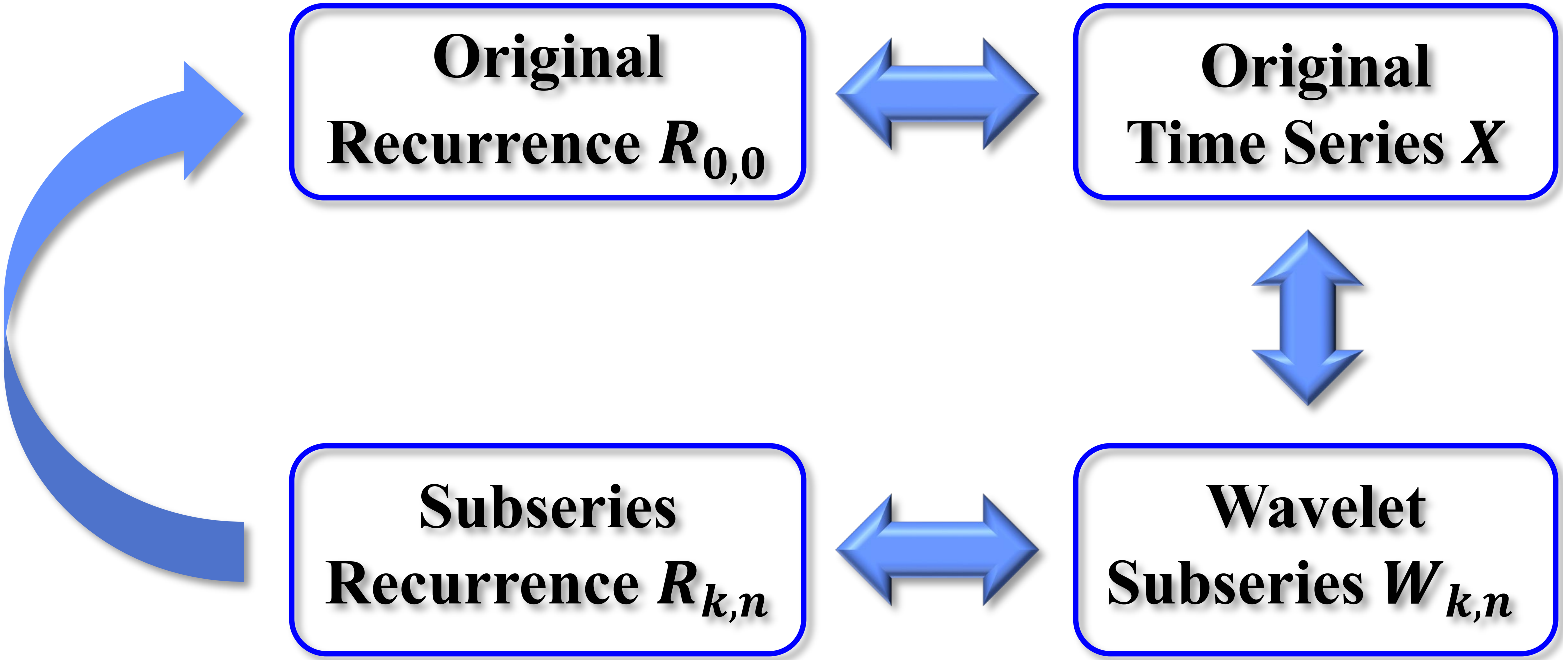
$$UR_{0,0}^2(i,j) = [\boldsymbol{W}_{k,2^k-1}^T \cdots \boldsymbol{W}_{k,0}^T] \mathcal{W}_k \left(\sum_{m=0}^{M-1} \Phi_{i+m\tau, j+m\tau}\right) \mathcal{W}_k^T \begin{bmatrix} \boldsymbol{W}_{k,2^k-1} \\ \vdots \\ \boldsymbol{W}_{k,0} \end{bmatrix}$$

Multiscale Recurrence Analysis

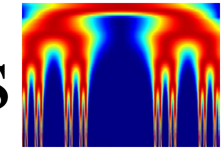


- Reconstruct time series from recurrence plot
  - **Y. Hirata et al. (2008); M. Thiel et al. (2004)**

Recurrence Plot ⟺ Time Series



Wavelet Preservation of Dynamics



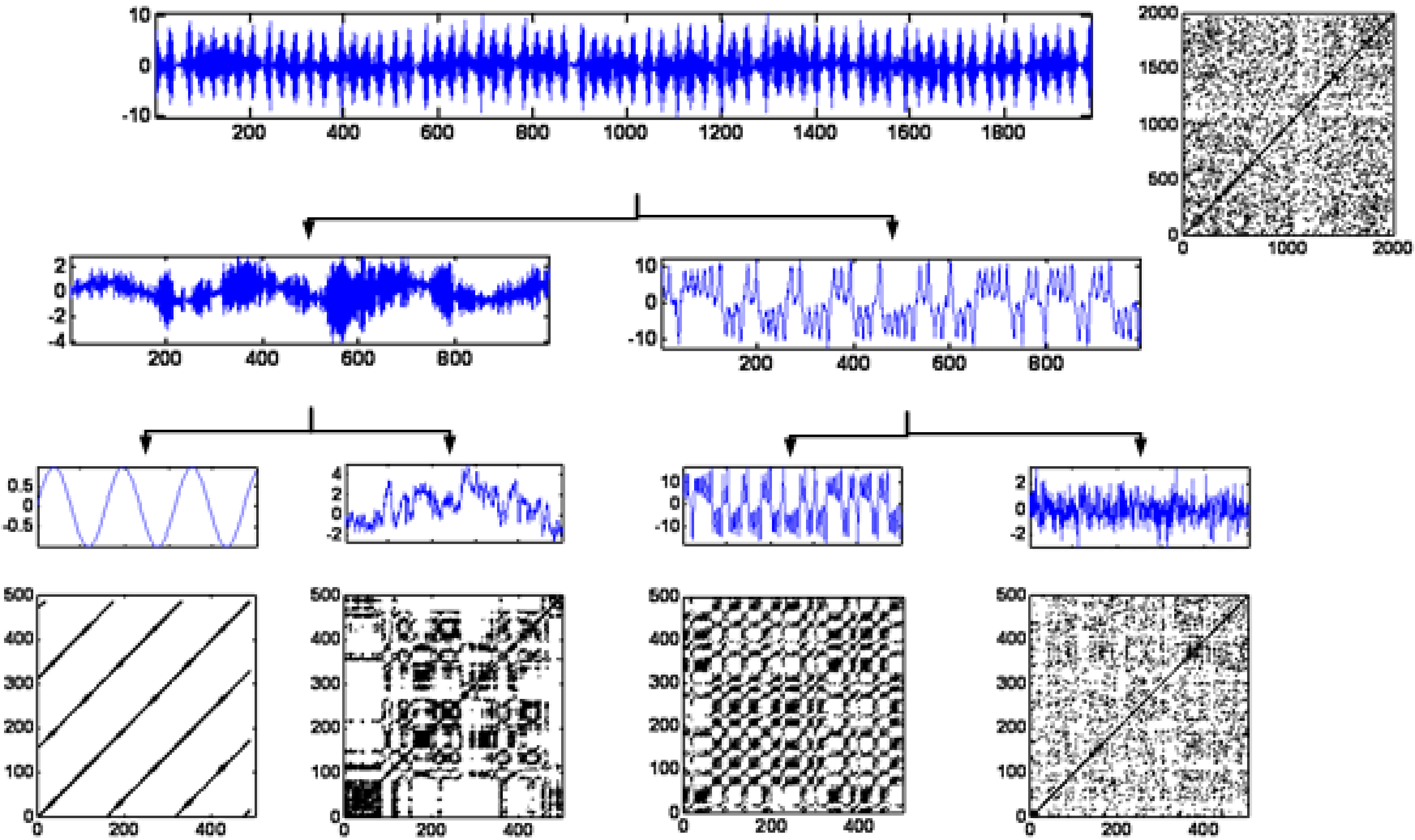
- Time-delay state space: $\boldsymbol{w}(i) = \left(w_i, w_{i+\tau}, \cdots, w_{i+\tau(M-1)}\right)$
- **Gram matrix:** $G(i,j) \equiv \boldsymbol{w}(i) \cdot \boldsymbol{w}(j)$
- Multidimensional scaling

$$UR^2(i,j) = [\boldsymbol{w}(i) - \boldsymbol{w}(j)] \cdot [\boldsymbol{w}(i) - \boldsymbol{w}(j)]$$

$$G(i,j) = -\frac{1}{2}\left[UR^2(i,j) - \frac{1}{N}\sum_{k=1}^{N} UR^2(i,k) - \frac{1}{N}\sum_{k=1}^{N} UR^2(k,j) + \frac{1}{N^2}\sum_{g=1}^{N}\sum_{h=1}^{N} UR^2(g,h)\right]$$

- Gram matrix is a square matrix: $G = U\Lambda U^T$
  - Λ is a diagonal matrix formed from the eigenvalues of G
  - U is a matrix of the corresponding eigenvectors of G
- Gram matrix is positive semidefinite
  - $G = U\sqrt{\Lambda}\sqrt{\Lambda}U^T = U\sqrt{\Lambda}\left(U\sqrt{\Lambda}^T\right)^T = U\sqrt{\Lambda}(U\sqrt{\Lambda})^T$
- **ISOMETRY: $\boldsymbol{U\sqrt{\Lambda}}$ and $\boldsymbol{w_i}$**

Multiscale Recurrence Analysis

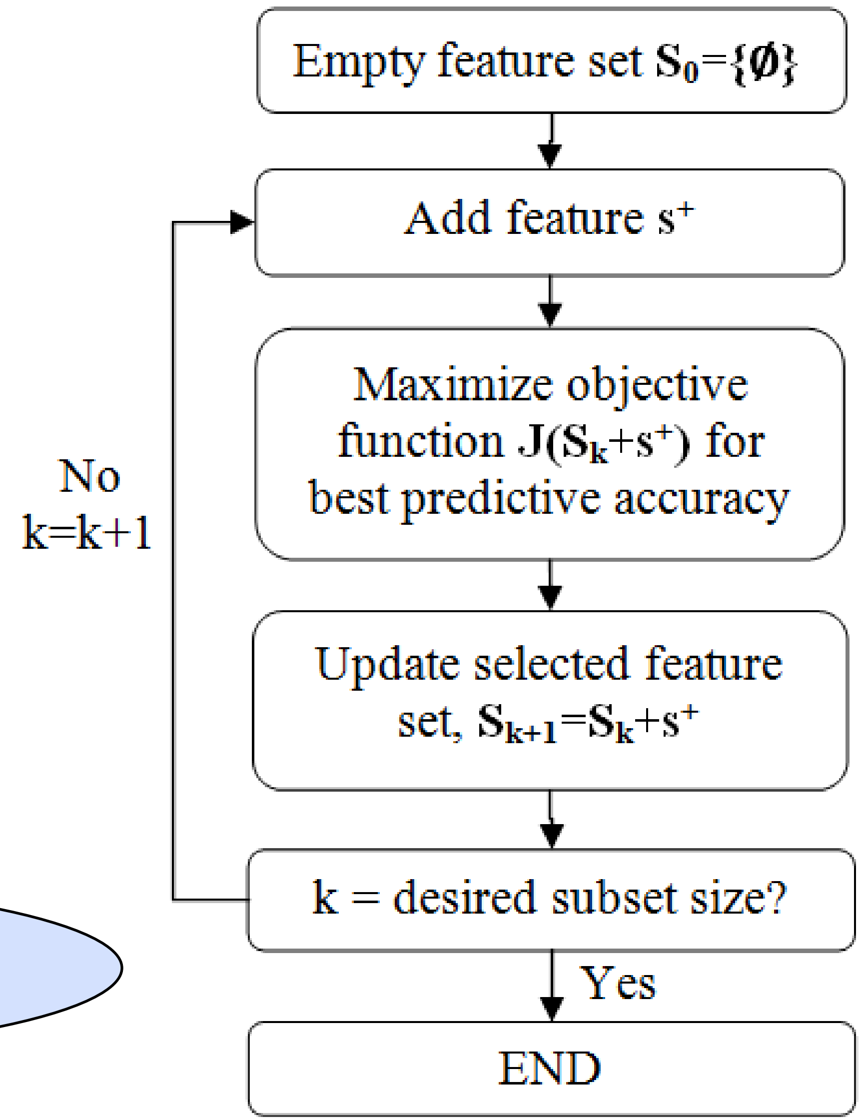


Feature Selection

- ❑ Recurrence features
  - ➢ RR, DET, LMAX, ENT, LAM, TT are extracted for each of the wavelet subseries
  - ➢ k^{th} level: 2^k number of wavelet subseries
  - ➢ Selected level: *m* to *n*
- ❑ Total feature size:

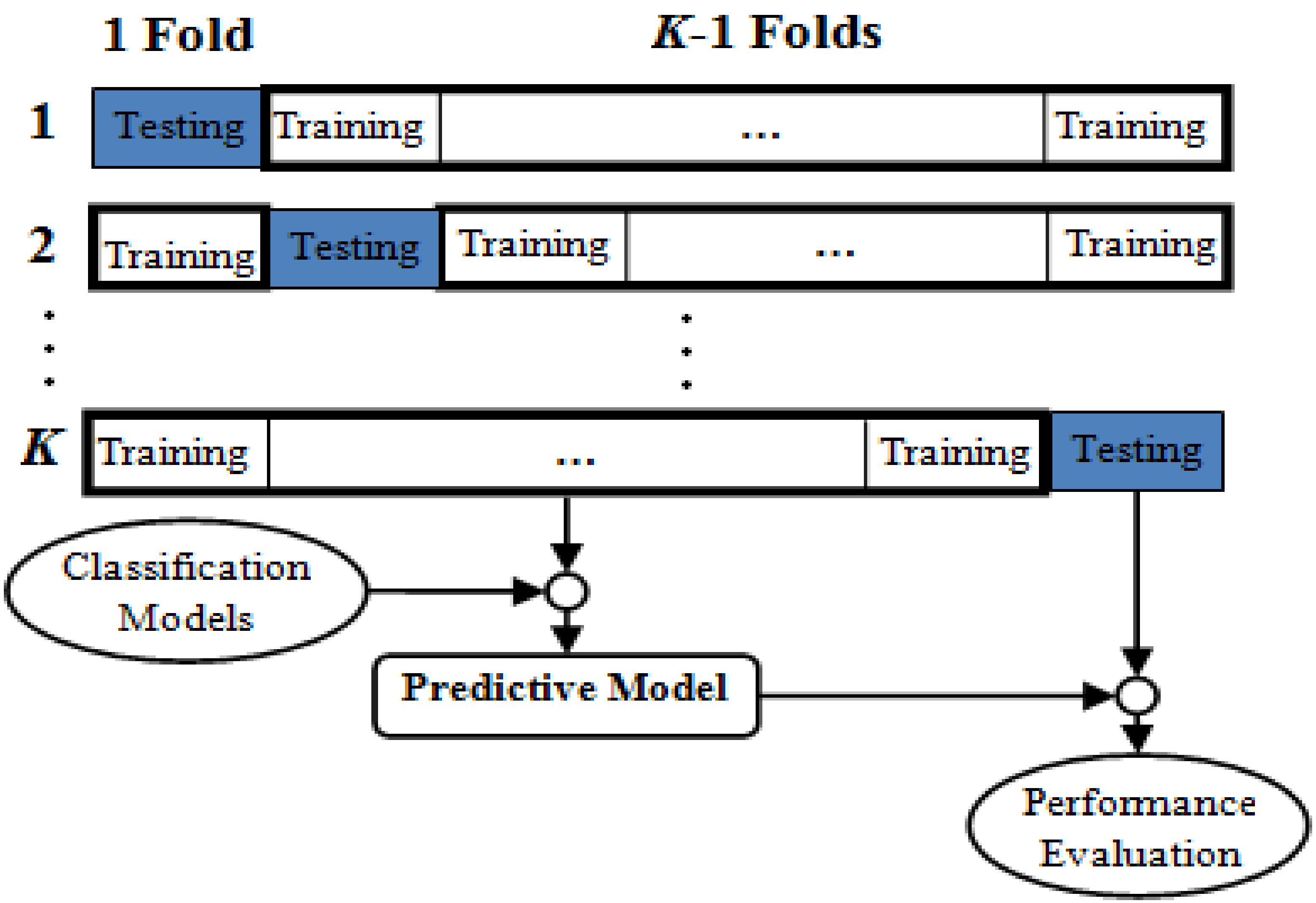
$$\sum_{k=m}^{n} 6 \times 2^k$$

Curse of dimensionality



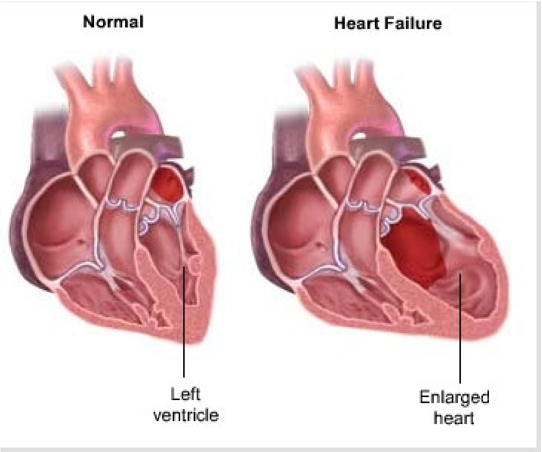
Cross-validation

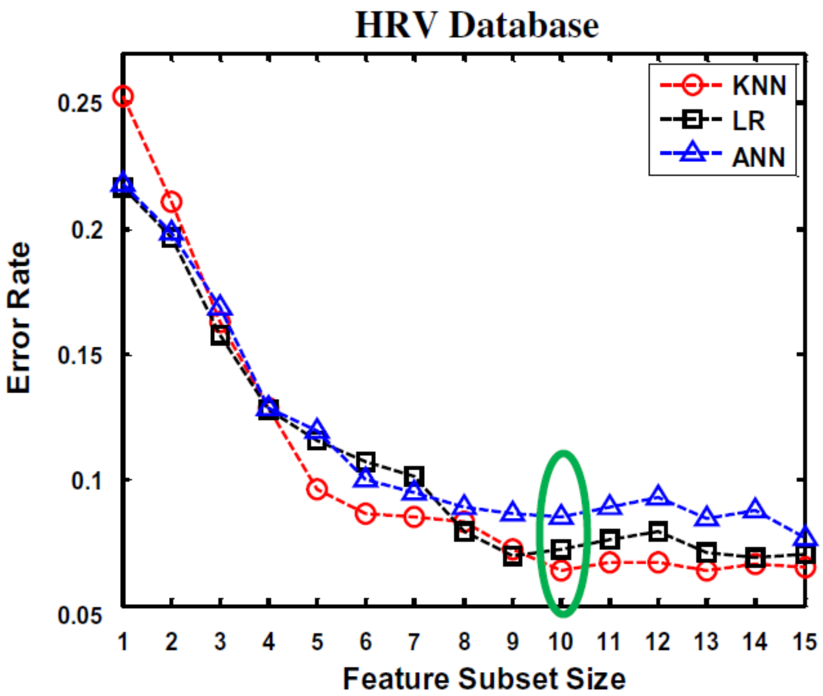
❑ K-fold cross-validation & Random sub-sampling



Dataset – Heart Rate Variability

- Dataset – 24-hour heart rate variability (**HRV**)
  - 54 Health control (**HC**)
  - 29 congestive heart failure (**CHF**)
- Classification models;
  - K-nearest neighbor (**KNN**)
  - Logistic regression (**LR**)
  - Artificial neural network (**ANN**)
- Feature selection
  - Selected level: **6** to **9**
  - Total: $\sum_{k=6}^{9} 6 \times 2^k = 5760$
- Select **10** features in order to prevent overfitting.



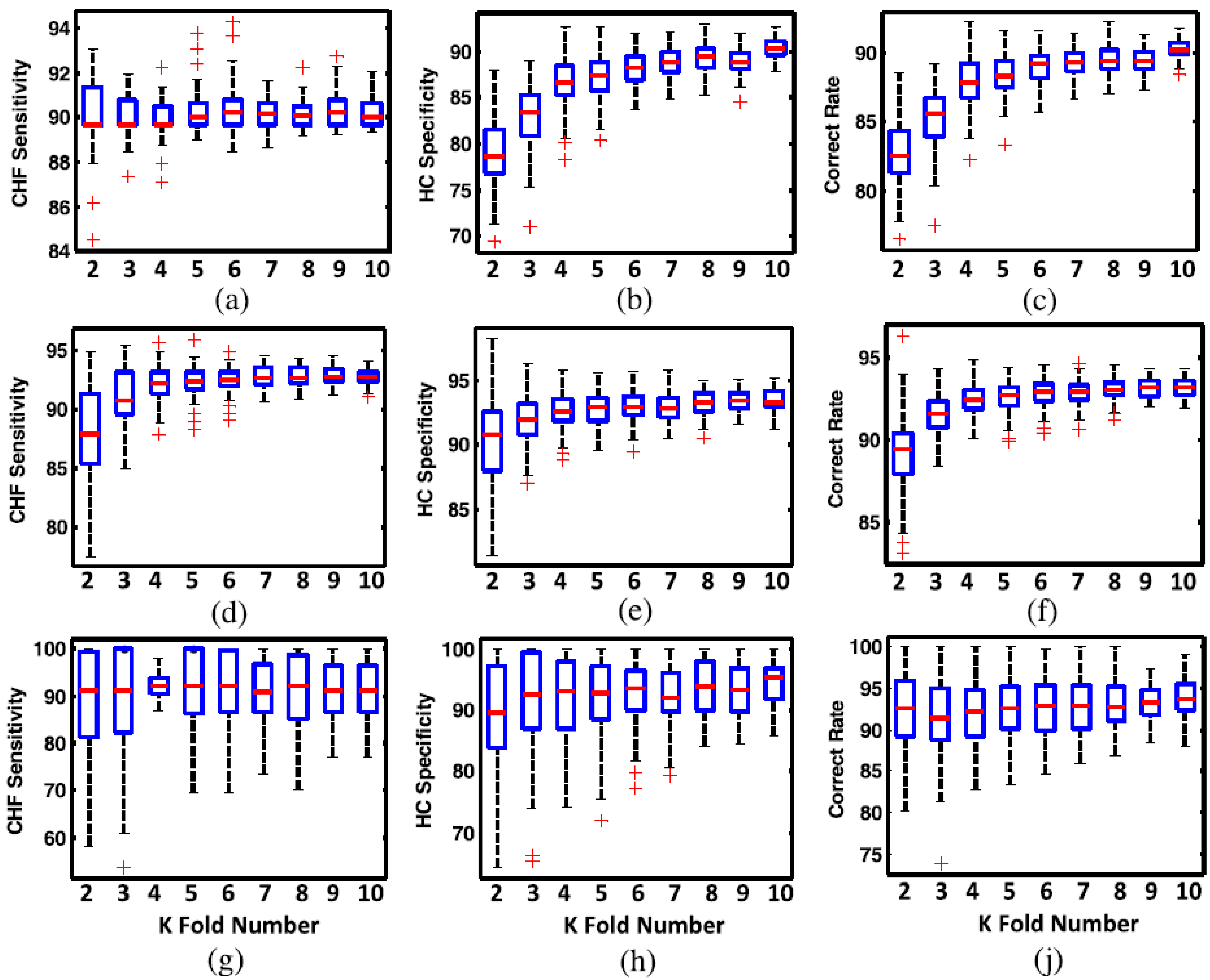


Performance Evaluation (1)

KNN

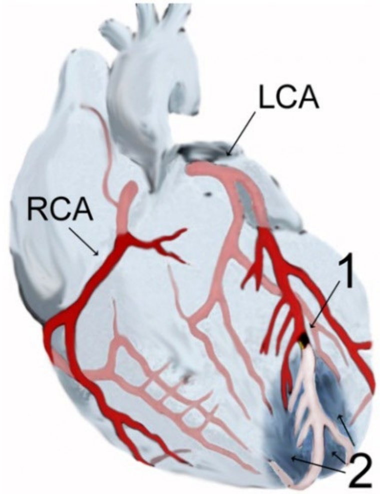
LR

ANN

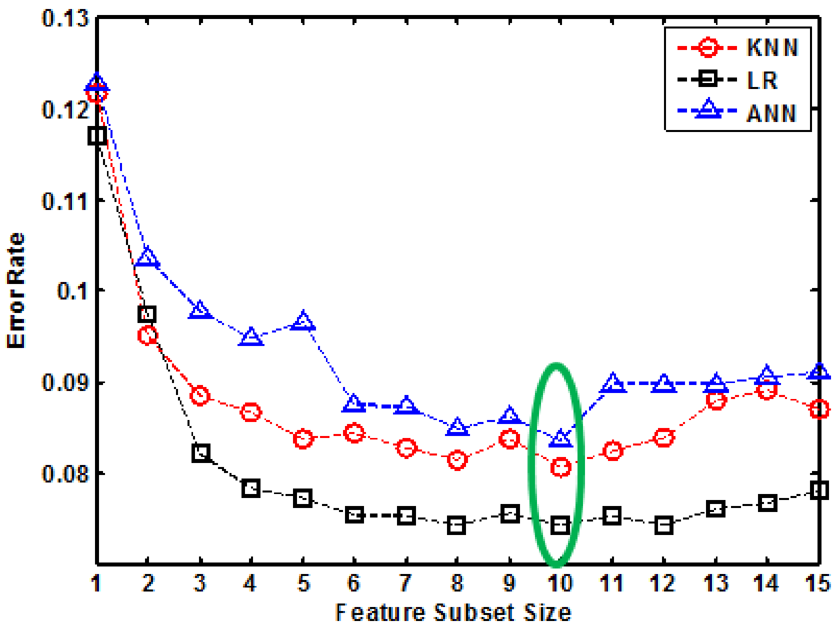


Dataset – Vectorcardiogram

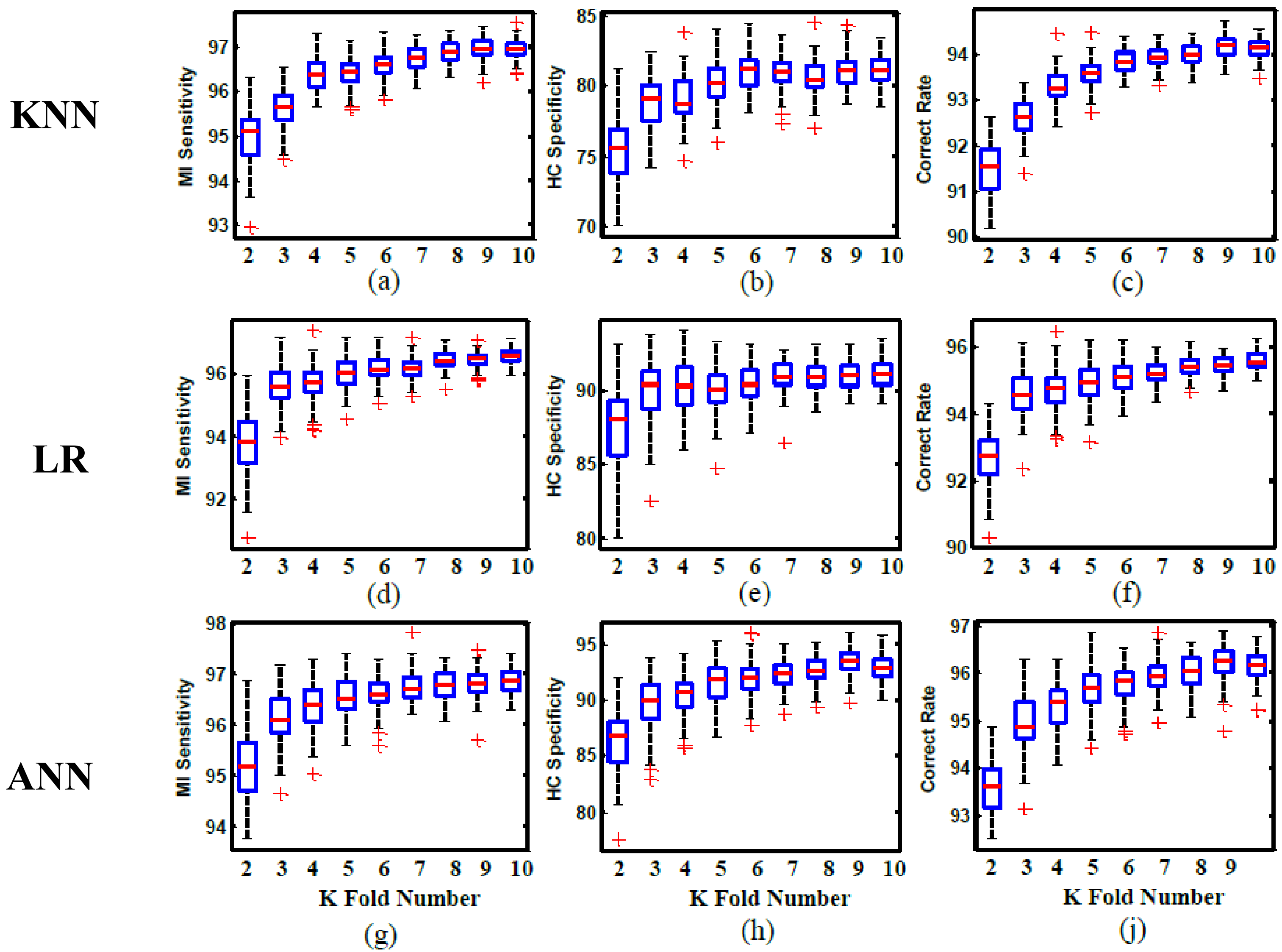
- Dataset – Vectorcardiogram (**VCG**)
  - 80 Health controls (**HC**)
  - 368 myocardial infarctions (**MI**)
- Classification models:
  - K-nearest neighbor (**KNN**)
  - Logistic regression (**LR**)
  - Artificial neural network (**ANN**)
- Feature selection
  - Selected level: **4** to **5**
  - Total: $\sum_{k=4}^{5} 6 \times 2^k = 288$
- Select **10** features in order to prevent overfitting.



MI (2), after occlusion (1) of a branch of LCA, RCA

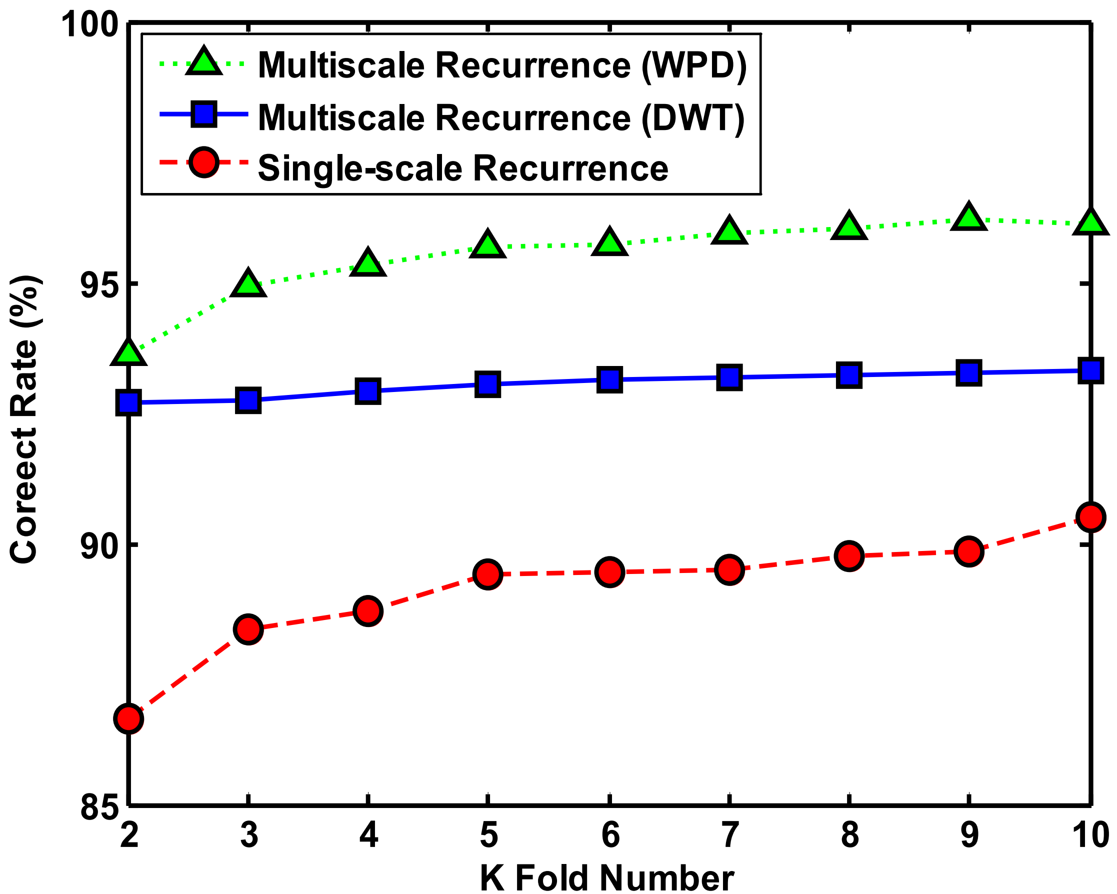


Performance Evaluation (2)

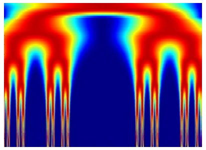


Single-scale VS. Multiscale

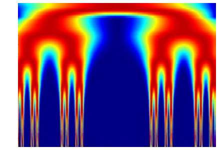
- Comparison of classification correct rates between single-scale and multi-scale (i.e., DWT and WPD) recurrence analysis



Conclusions



- **Challenges:**
  - **Data complexity: nonlinearity** and **nonstationarity**
  - **Enormous data torrents**
- Multiscale recurrence analysis
  - **Large size of dataset** – dyadic subsampling
  - **Nonstationarity** – wavelet decomposition
  - **Nonlinearity** – recurrence analysis
- Discriminant analysis
  - **HRV** database: **92.1%** (sensitivity) and **94.7%** (specificity)
  - **VCG** database: **96.8%** (sensitivity) and **92.8%** (specificity)
- Single-scale vs. multiscale recurrence analysis



END
Questions?