Hierarchical structure of correlations in a set of stock prices

Rosario N. Mantegna

Observatory of Complex Systems

Palermo University

In collaboration with:

Giovanni Bonanno

Fabrizio Lillo

Observatory of Complex Systems
http://lagash.dft.unipa.it

1. Hierarchical Structure of Correlations

- The correlation structure in a portfolio of stocks is pretty complex.

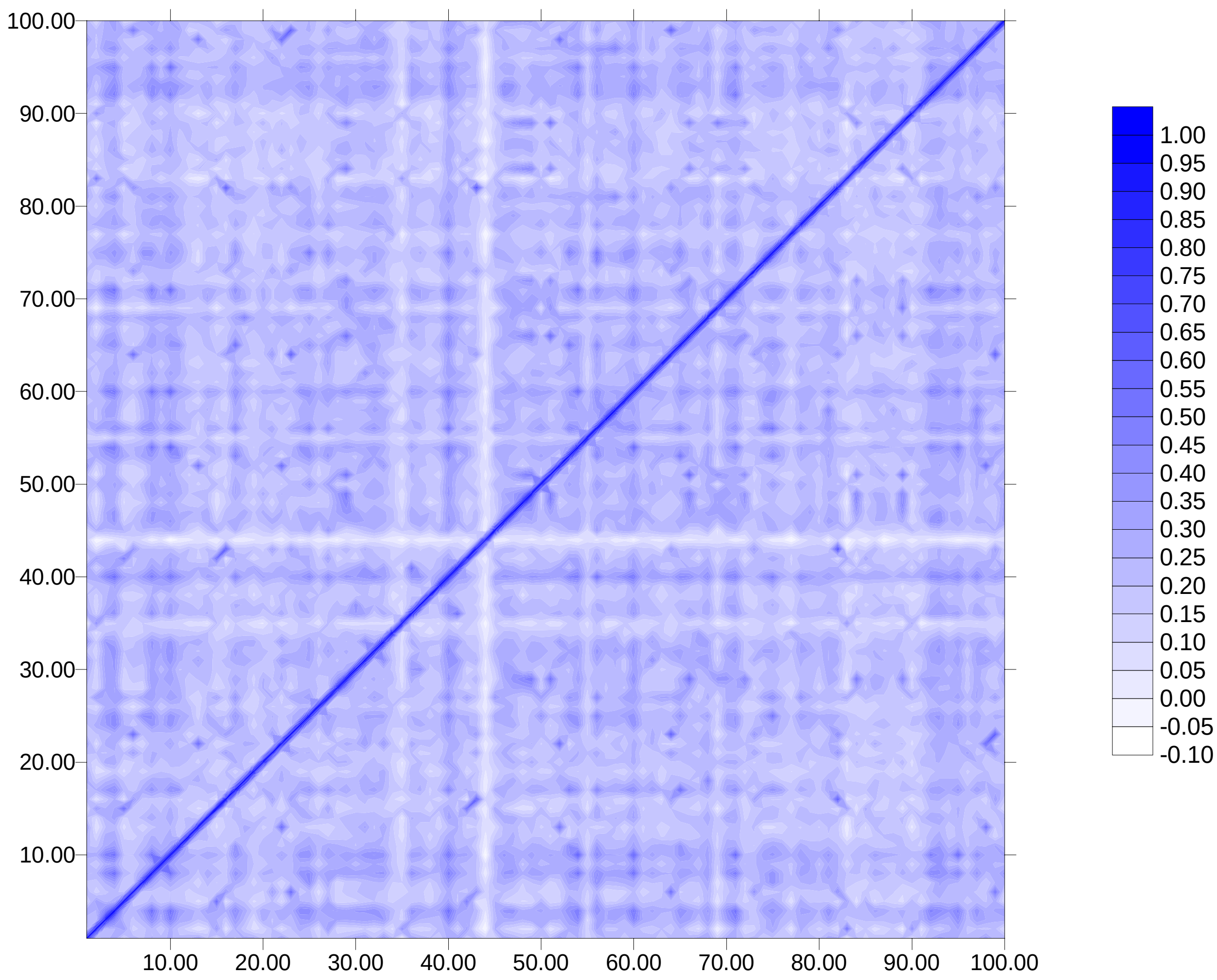
- One way to quantify correlation is done by considering the correlation coefficient

$$\rho_{ij} = \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{\langle S_i^2 - \langle S_i \rangle^2 \rangle \langle S_j^2 - \langle S_j \rangle^2 \rangle}}$$

$$S_i \equiv \ln Y_i(t) - \ln Y_i(t-1)$$

2. Hierarchical Structure of Correlations

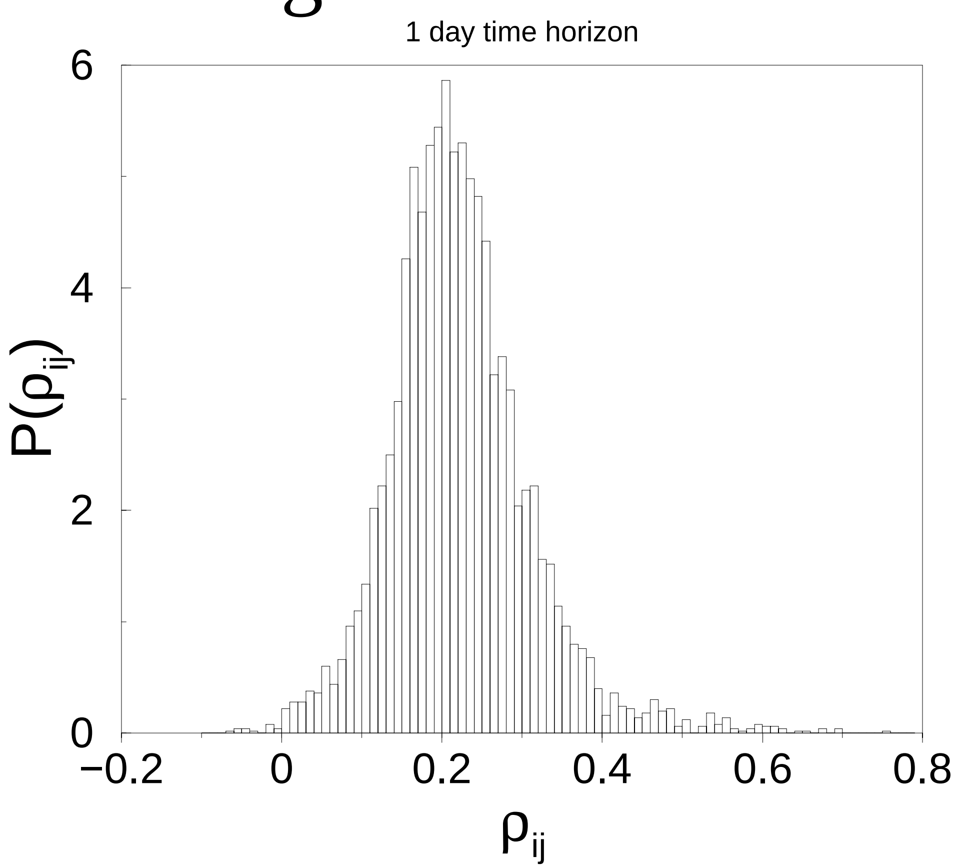
- Let us consider the set of 100 stocks used to compute the S&P 100 index



3. Hierarchical Structure of Correlations

- The matrix ρ is symmetrical and contains information about
- **n(n-1)/2** correlation coefficients.

- A statistical description of this information may be obtained by considering



4. Hierarchical Structure of Correlations

- The **correlation matrix** contains a large amount of economic information.

- **A key problem is:** How to extract this information?

- We propose to use a form of **cluster analysis** to discover the underlying **hierarchical structure**.

5. Hierarchical Structure of Correlations

The cluster analysis is performed in two steps:

1) By defining a **metric distance** starting from the **correlation coefficient**;

2) By extracting the **subdominant ultrametric** of the considered **metric distance**.

6. Hierarchical Structure of Correlations

A **metric distance** d_{ij} verifies the axioms

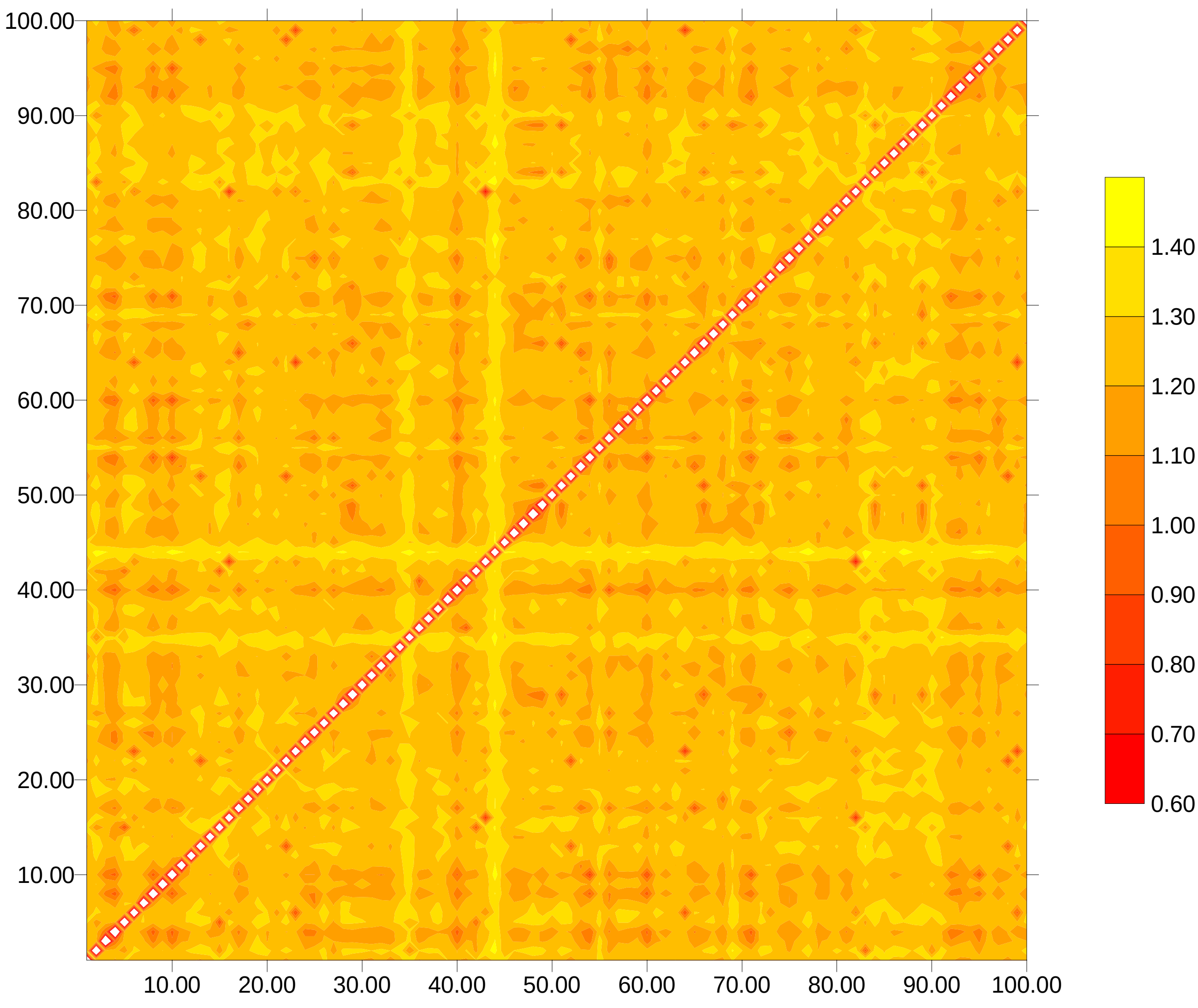
$$(i) \quad d_{ij} = 0 \iff i = j$$

$$(ii) \quad d_{ij} = d_{ji}$$

$$(iii) \quad d_{ij} \leq d_{ik} + d_{kj}$$

7. Hierarchical Structure of Correlations

The function $d_{ij} = \sqrt{2\left(1-\rho_{ij}\right)}$ verifies the axioms of a **metric distance.**



8. Hierarchical Structure of Correlations

An **ultrametric distance** verifies the axioms

$d^*(i,j)=0 \Leftrightarrow i \equiv j$

$d^*(i,j)=d^*(j,i)$

$d^*(i,j) \leq Max\{d^*(i,k),d^*(k,j)\}$.

To perform a cluster analysis we extract the **subdominant ultrametric distance matrix $d^<$.**

An example:

Starting from the **distance**

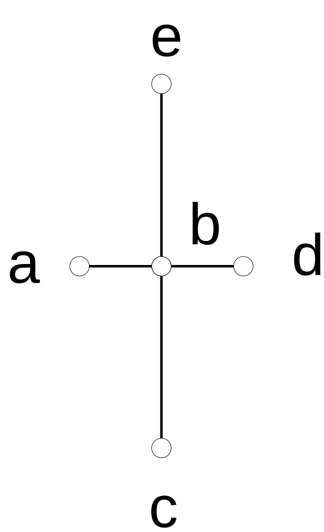
	a	b	c	d	e
a	0	0,85	0,94	0,89	0,96
b		0	0,88	0,78	0,93
c			0	0,95	1,01
d				0	0,97
e					0

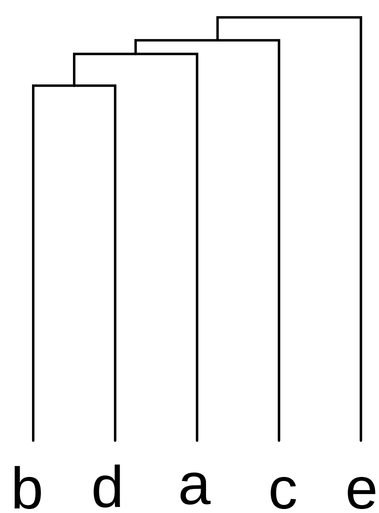
9. Hierarchical Structure of Correlations

By considering the shortest distances

b	d	0.78
a	b	0.85
b	c	0.88
a	d	0.89
b	e	0.93
a	c	0.94
c	d	0.95
a	e	0.96
d	e	0.97
c	e	1.01

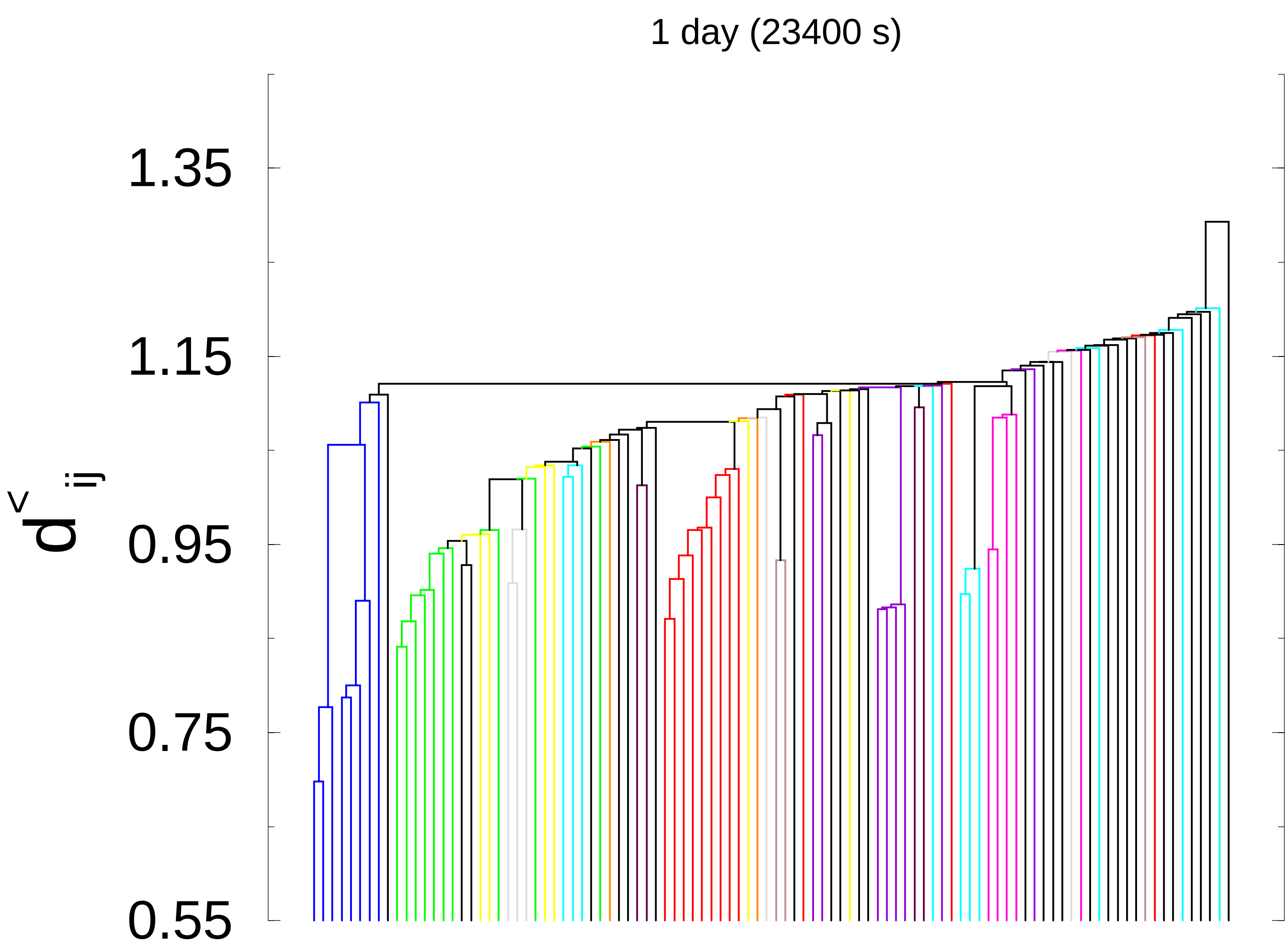
It is possible to obtain the **Minimum Spanning Tree** and an associated **Hierarchical Tree**





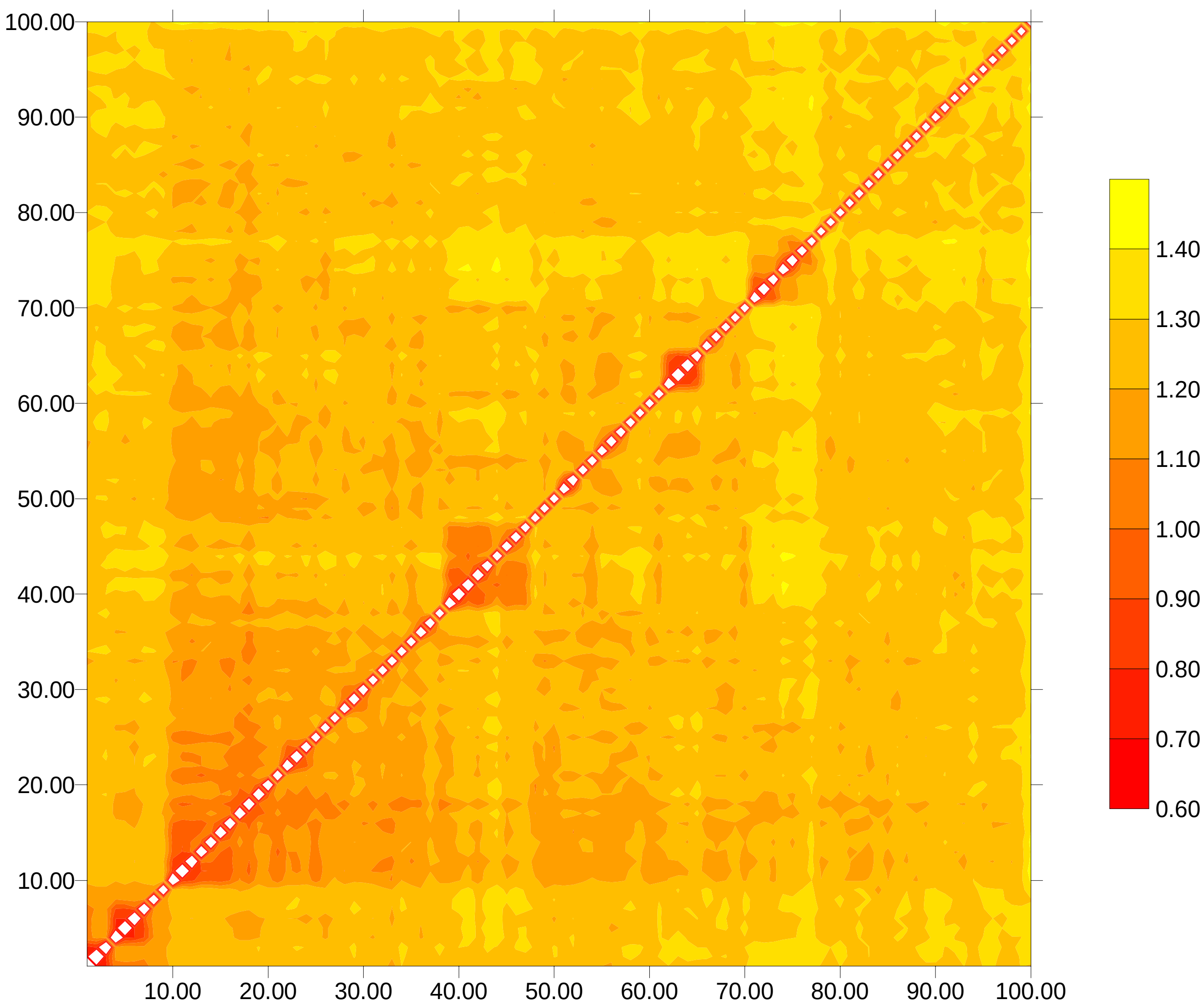
10. Hierarchical Structure of Correlations

For our portfolio the **hierarchical tree** is



11. Hierarchical Structure of Correlations

We can now re-analyze the **distance matrix** by using the sequence order sorted out by the cluster analysis



The interpretation of the distance matrix is now more direct

12. Hierarchical Structure of Correlations

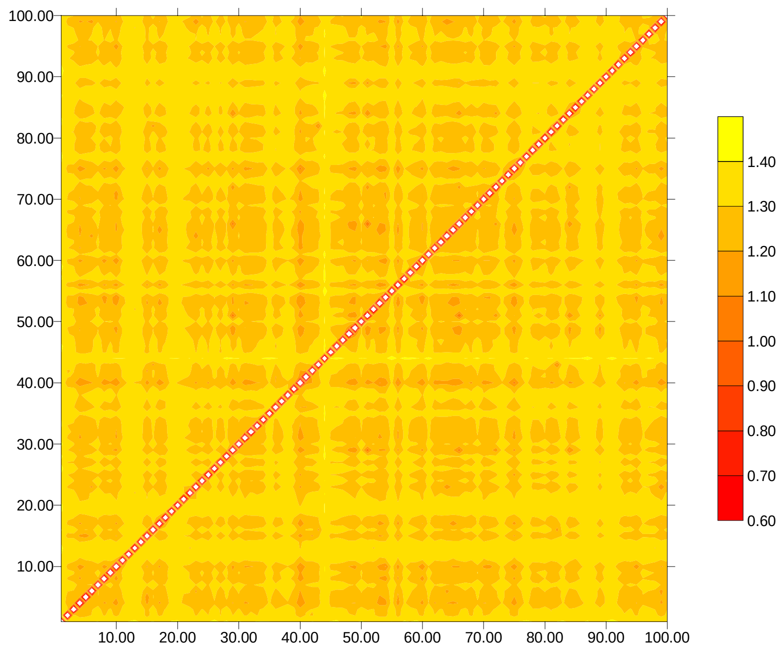
What do we learn by performing this approach?

Cluster analysis is a useful technique able to detect relevant **economic information**

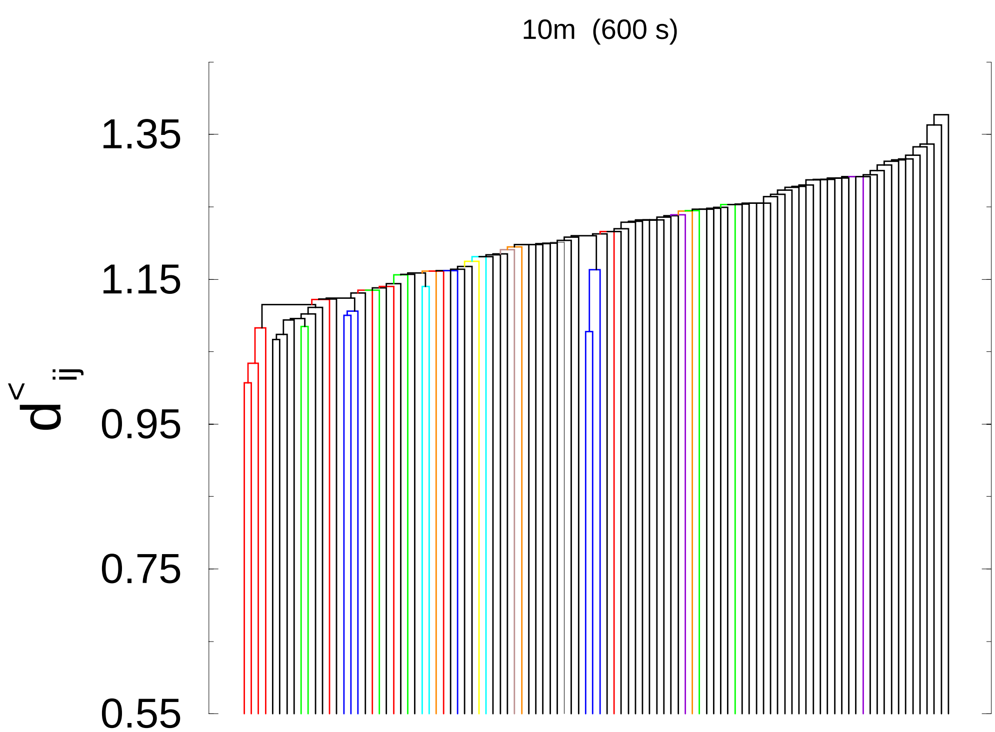
Next question:
Is this **information** the same at all **time horizons**?

13. Hierarchical Structure of Correlations

A ten minutes time horizon

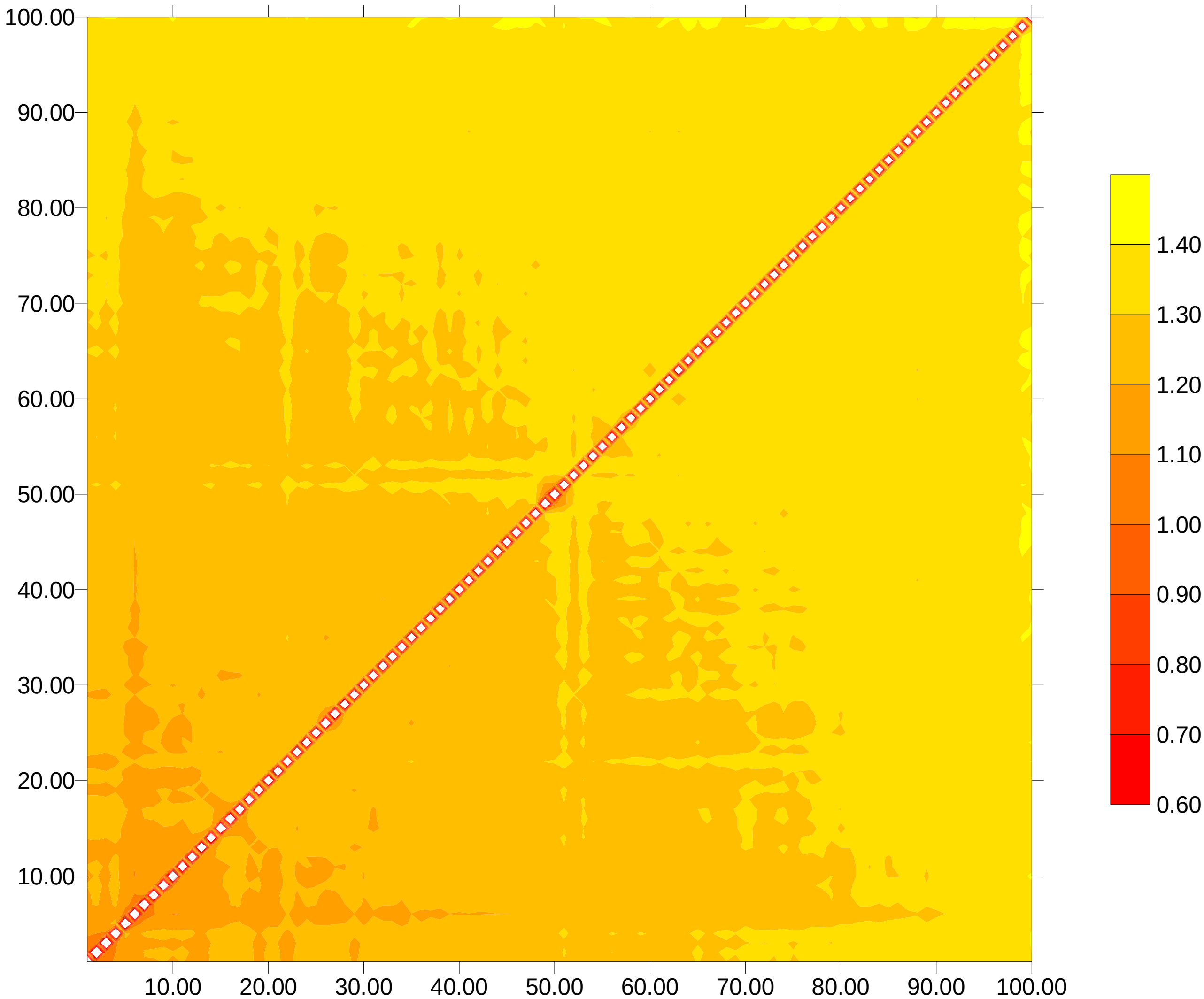


Hierarchical tree



14. Hierarchical Structure of Correlations

600 s d_{ij} ordered by the MST



Note that:

1) **Correlations are weaker**;

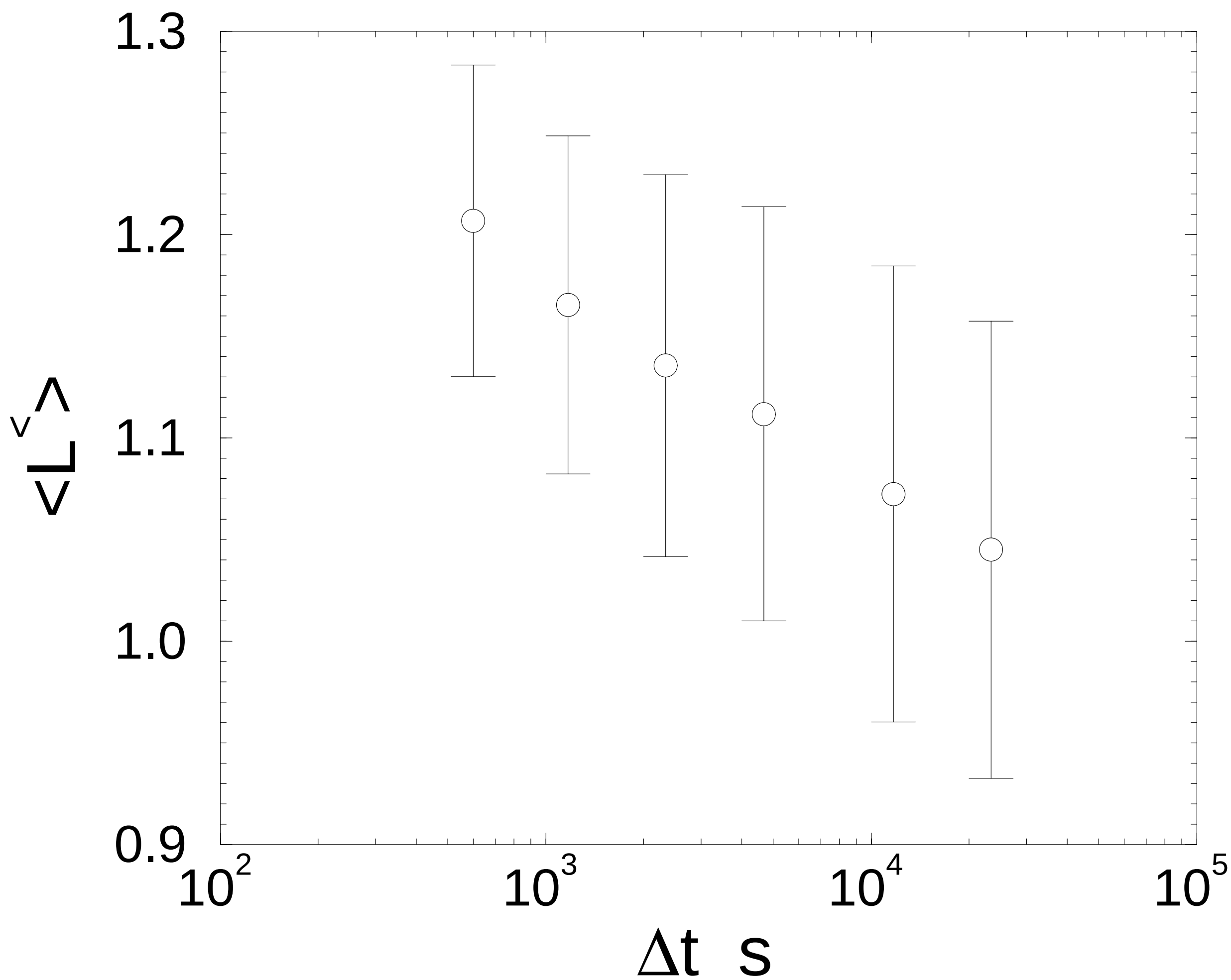
2) **The structure is less complex**

15. Hierarchical Structure of Correlations

Can we use the hierarchical modeling to interpret and quantify the structuring of **economic information** in a financial market?

To this end we investigate the **subdominant ultrametric** structure $d^{<}_{ij}$ at different **time horizons** Δt

We quantify the compactness of the detected hierarchical structure by measuring the total length $L^{<}$ of the MST



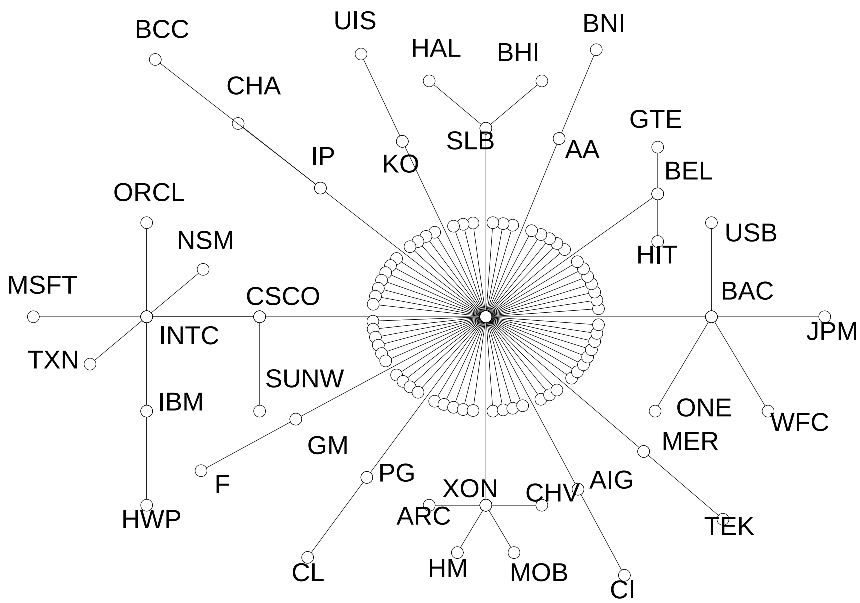
17. Hierarchical Structure of Correlations

Our empirical results show that the **minimum spanning tree** and the **hierarchical tree** become more structured as time horizon increases

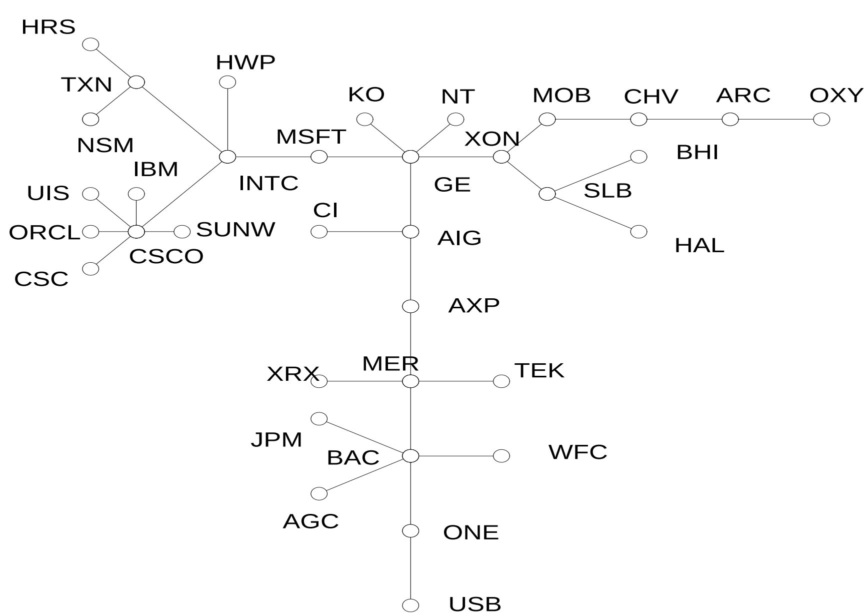
The analysis performed at different **time horizons** shows that the market is ``learning” which is the most appropriate degree of pair correlation.

<u>18. Hierarchical Structure of Correlations</u>

MST obtained for $\Delta t=600$ seconds

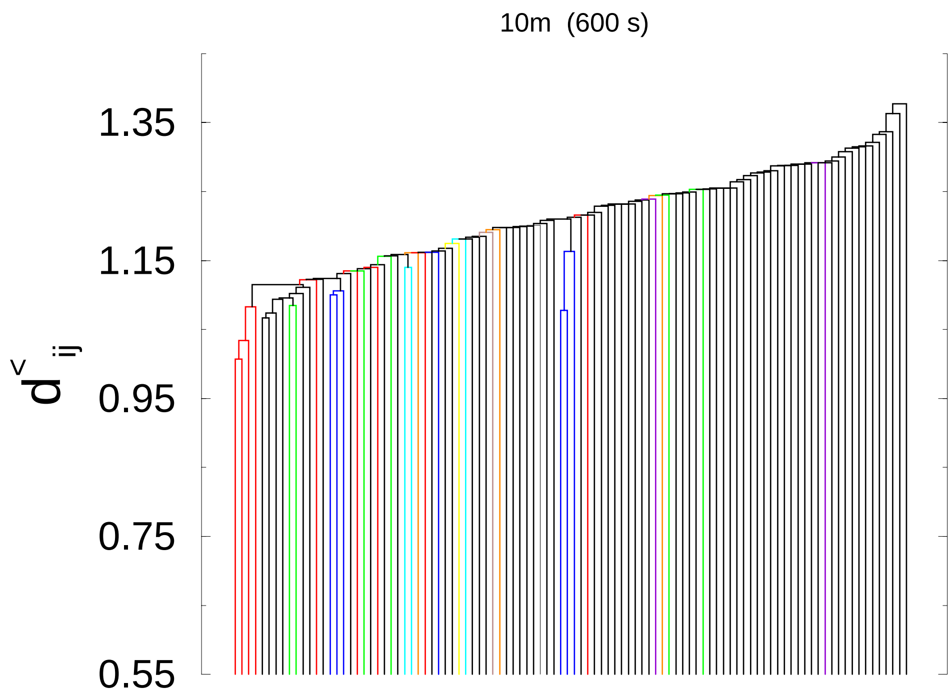


MST obtained for $\Delta t=1$ day

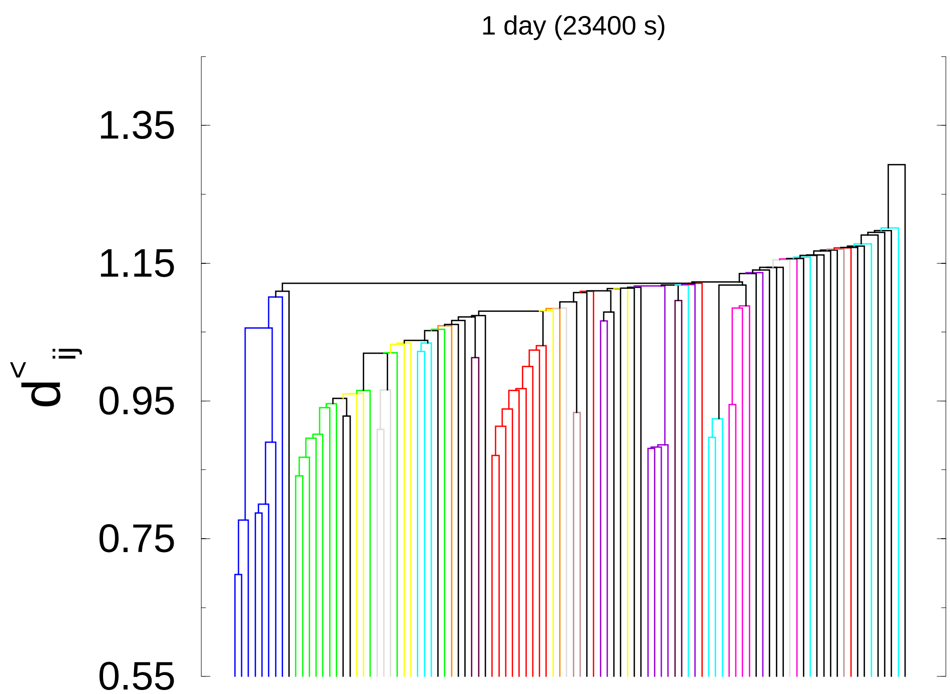


19. Hierarchical Structure of Correlations

HT obtained for Δt=600 seconds

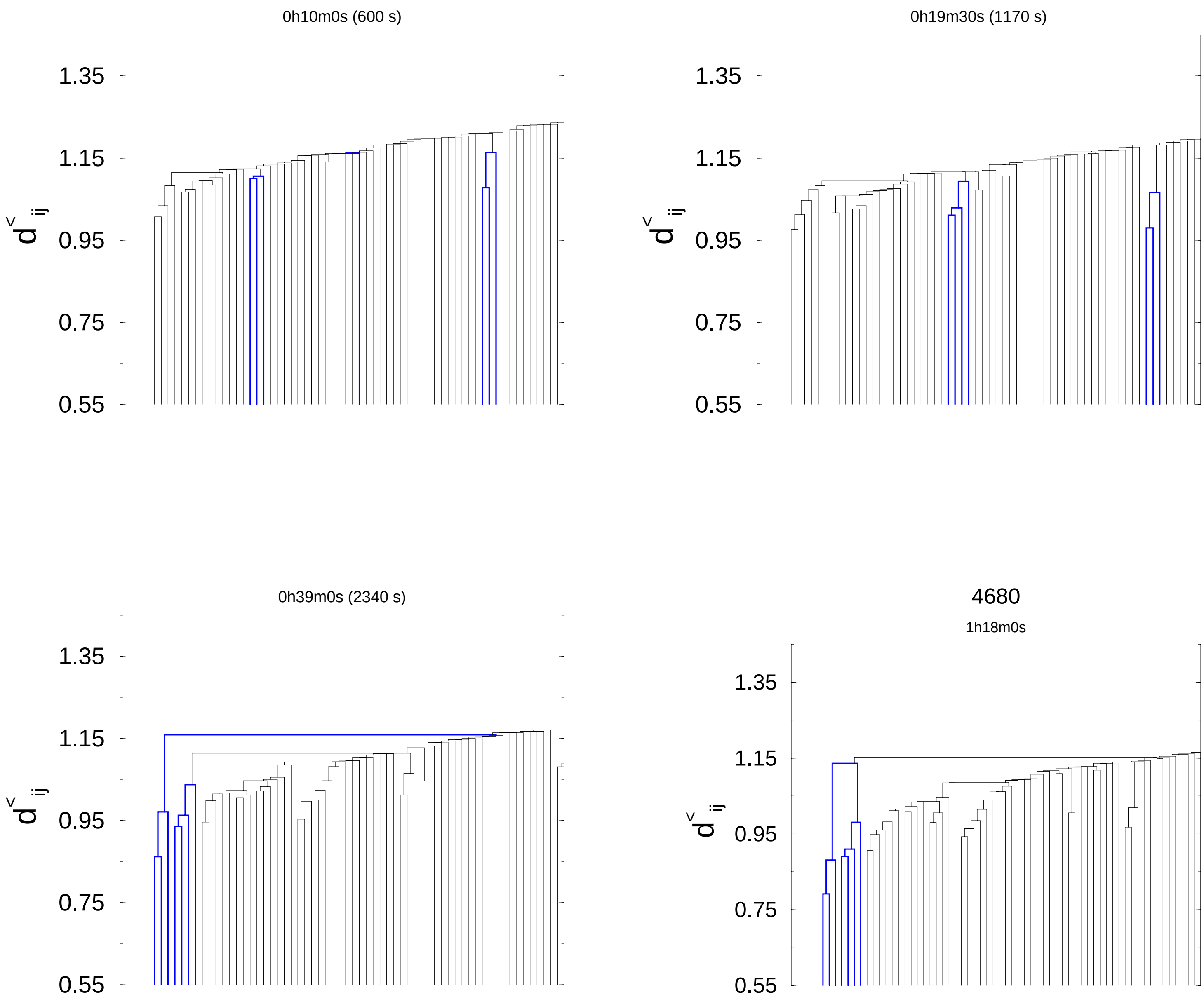


HT obtained for Δt=1 day

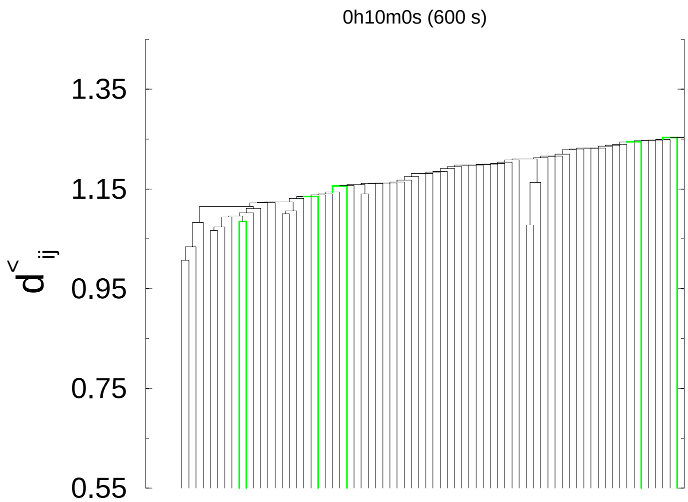


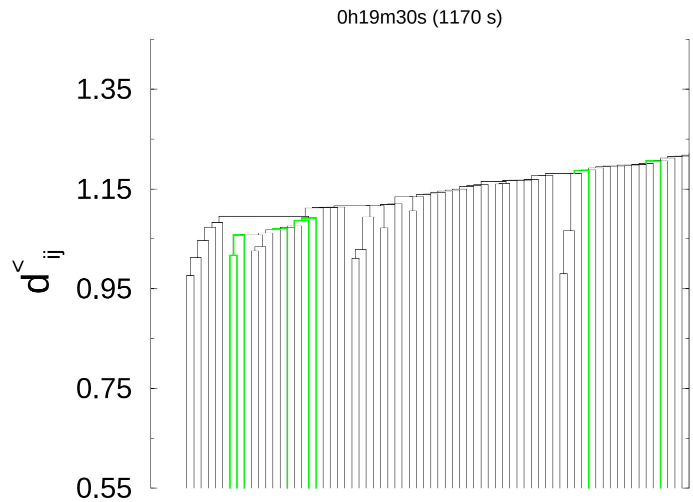
20. Hierarchical Structure of Correlations

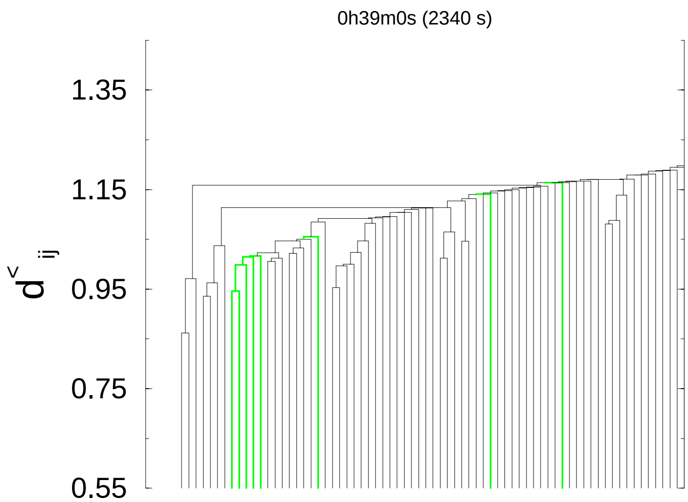
Energy sector

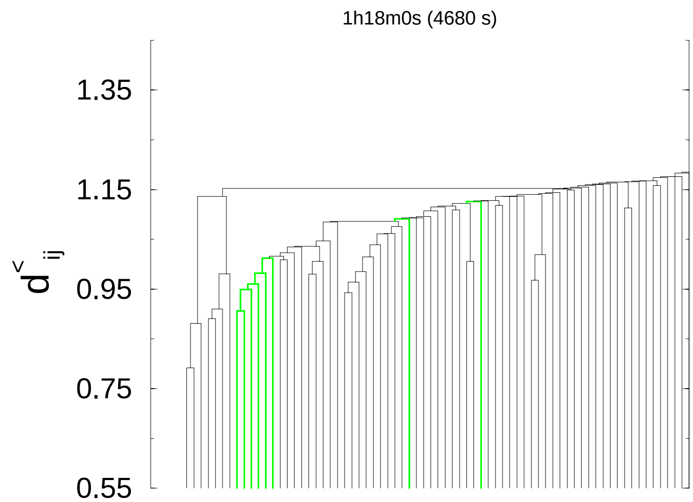


Financial sector

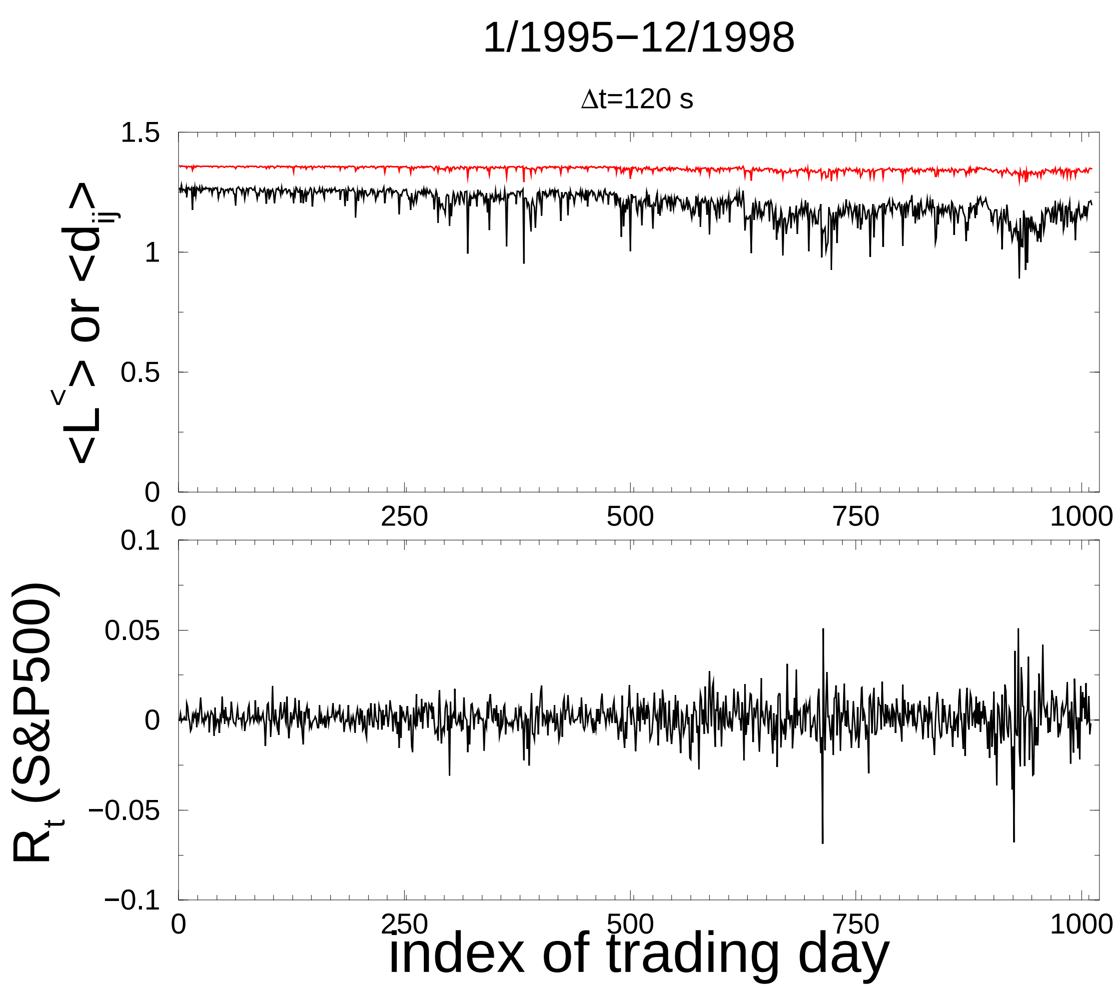








Is the hierarchical modeling more sensitive than customary methods in detecting the correlations of a stock portfolio?



Conclusions

Cluster analysis based on the **subdominant ultrametric** provides a direct way to detect **economic information** present into stock price time series without any external assumption.

The structuring of **economic information** occurs at different **time horizons** and increases by when the **time horizon** increases.