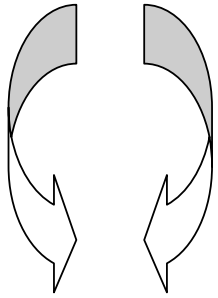
FRACTIONAL VECTOR AUTOREGRESSIVE MODELS FOR ENVIRONMENTAL DATA

Mariagrazia Granturco – Eugenia Nissi
Dipartimento di Metodi Quantitativi e Teoria Economica
Università degli Studi "G. D'Annunzio"
Viale Pindaro, 42 –65127 Pescara
m.granturco@dmqte.unich.it - nissi@dmqte.unich.it

TIES –2002 Annual Conference – Genova 18-22 giugno 2002

AIMS OF THE PAPER



Time series of pollutants usually show long memory dependence

In this setting, the aims of the paper is to analyze times series of the concentration of carbon monoxide in the Bergamo District by means of multivariate ARFIMA Models

Structure of the talk:

- *ARFIMA Models: methodological issues*
- *VAR Models: methodological issues*
- *Proposed methodology for multiple time series:*
  - *Estimation of the common long memory parameter;*
  - *Estimation of VAR model on the residual short memory component*
- *Case of study*

DEFINITIONS OF LONG MEMORY TIME PROCESSES

- **THE CORRELATION DECAY TO ZERO WITH AN HYPERBOLIC RATE**

$$\lim_{n\to\infty} \rho_{(k)} = c \cdot k^{-\alpha} \Rightarrow \sum_{k=-\infty}^{+\infty} |\rho(k)| = \infty$$

con: $\alpha \in (0\ ,\ 1)$

$c > 0$

- **THE SPECTRAL DENSITY FUNCTION IS NOT FINITE AT ZERO FREQUENCY**

$$\lim_{\lambda\to 0} f(\lambda) = \lim_{\lambda\to 0}\left[\frac{\sigma^2}{2\pi}\sum_{k=-\infty}^{+\infty}\rho(k)\cdot e^{-i\lambda k}\right] = c\,\lambda^{-\alpha}$$

TEMPORAL PROCESSES WITH LON MEMORY

STATIONARY INCREMENTS OF SELF SIMILARITY PROCESS

$$X(t) = Y(t)-Y(t-1): \; Y(t) = c^{-H}Y(tc)$$

IF:

- $H \in (½,1)$: **INVERTIBLE AND STATIONARY PROCESS WITH LONG MEMORY**
- **X(t)~N(μ,Σ): FRACTIONAL GAUSSIAN NOISE WITH SELF SIMILARITY PARAMETER *H***

ARFIMA(0,d,0) PROCESS

$$X(t) = (1-B)^{-d}\,\varepsilon(t)$$

IF:

- $d \in (0,½)$: **INVERTIBLE AND STATIONARY PROCESS WITH LONG MEMORY**
- **X(t)~N(μ,Σ):** ARFIMA PROCESS WITH FRACTIONAL DIFFERENCE OF ORDER ***d***

ASIMPTOTICALLY EQUIVALENT PROCESSES:

$$H = d + ½$$

⇩

Generalization : ARFIMA PROCESSES (p,d,q)

$$\phi(B)\,(1-B)^{d}\,X(t) = \theta(B)\varepsilon(t)$$

TEST FOR LONG MEMORY

$H_0 = x_1, x_2, \ldots, x_1 = I(0)$

$H_1 = x_1, x_2, \ldots, x_1 = I(1)$

➢ RESCALED RANGE TEST

$$R/S = \frac{\max_{1\le k\le n}\left[\sum_{j=1}^{k}(x_j - \bar{x}_n)\right] - \min_{1\le k\le n}\left[\sum_{j=1}^{k}(x_j - \bar{x}_n)\right]}{\text{1)}\ \sqrt{\frac{1}{n}\sum_{j=1}^{t}(x_j - \bar{x}_n)^2}} \quad \rightarrow^{d} = V(t)$$

CLASSIC R/S

$$\text{2)}\ \sqrt{S_x^2 + 2\sum_{j=1}^{q} w_j(q)\,\hat{\gamma}_j}$$

MODIFIED R/S

➢ KPSS TEST

$$\eta = \frac{1}{n^2\hat{\sigma}^2_n(q)} \sum_{j=1}^{n} \begin{cases} \text{1)}\ (x_t - \bar{x})^2 & \rightarrow^{d} \int_0^1 V_2(r)^2 dr \\ \text{2)}\ (x_t - \alpha - \beta t)^2 & \rightarrow^{d} \int_0^1 V_1(r)^2 dr \end{cases}$$

➢ LOBATO-ROBINSON TEST

$$LB = \sqrt{m}\,\frac{\sum_{j=1}^{m} v_j{}^{k} I(\lambda_j)}{\sum_{j=1}^{m} I(\lambda_j)} \sim N(0,1)$$

$$v_j = \ln(j) - \frac{1}{m}\sum_{j=1}^{m}\ln(i)$$

➢ RESCALED VARIANCE TEST

$$V/S = \frac{VAR(S_1, \ldots, S_n)}{n^2\hat{\sigma}^2_n(q)} \rightarrow^{d} V(t)$$

ESTIMATION OF THE LONG MEMORY PARAMETER H

ESTIMATION FROM CORRELOGRAM:

$$\lim_{k\to\infty} \rho(k) = c \cdot k^{1-2H}$$

ESTIMATION FROM SAMPLE VARIANCE:

$$\lim_{n\to\infty} \left[VAR(\overline{X})\right] = c \cdot n^{2H-2}$$

ESTIMATION FROM R/S REGRESSION:

$$\lim_{n\to\infty} \left[Log(R/S)\right] = A + H \cdot Log(n)$$

ESTIMATION FROM PERIODOGRAM:

$$\lim_{\lambda\to 0} f(\lambda) = c \cdot \lambda^{1-2H}$$

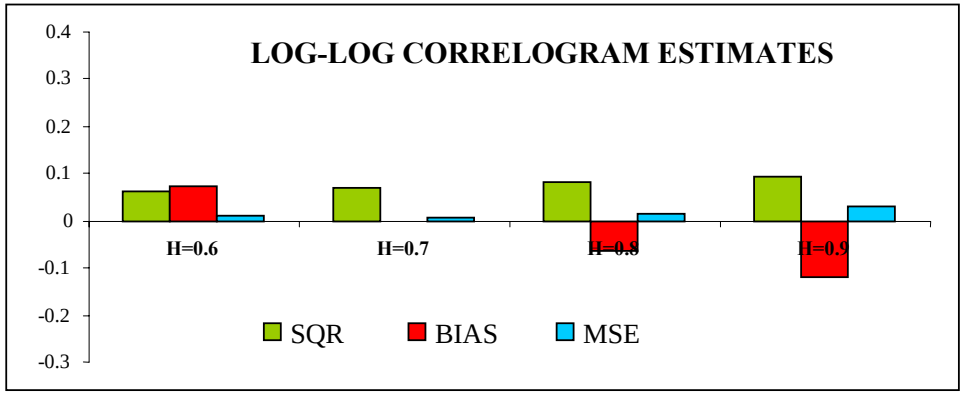
GPH ESTIMATION:

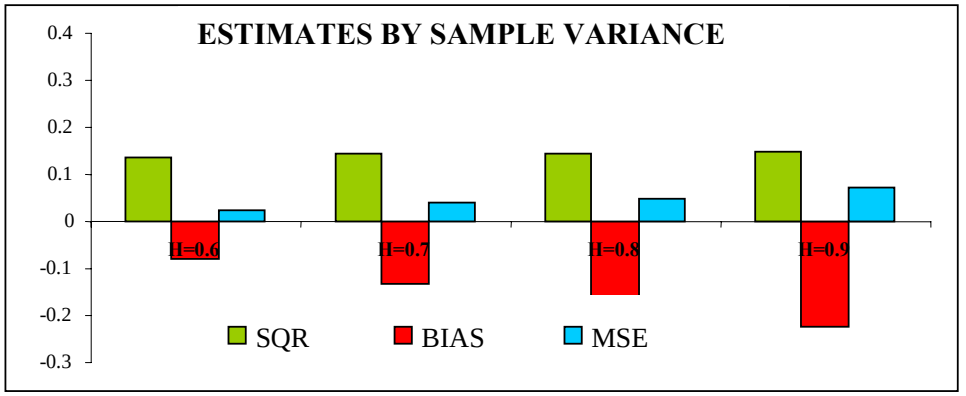
$$Log[I(\lambda)] = Log\left\{\frac{\sigma^2 f_a(0)}{2\pi}\right\} - d\ Log\left\{4\sin^2\frac{\lambda}{2}\right\} + \frac{\log I(\lambda)}{\log f_X(\lambda)}$$

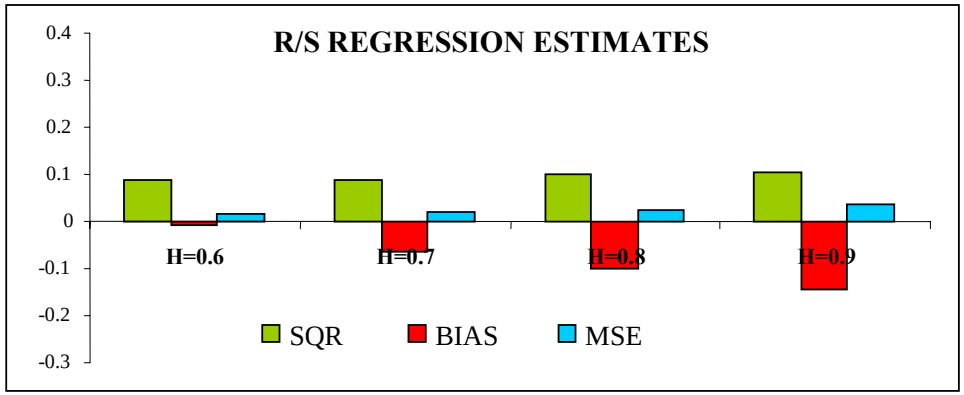
APPROXIMATE MLE (WHITTLE):

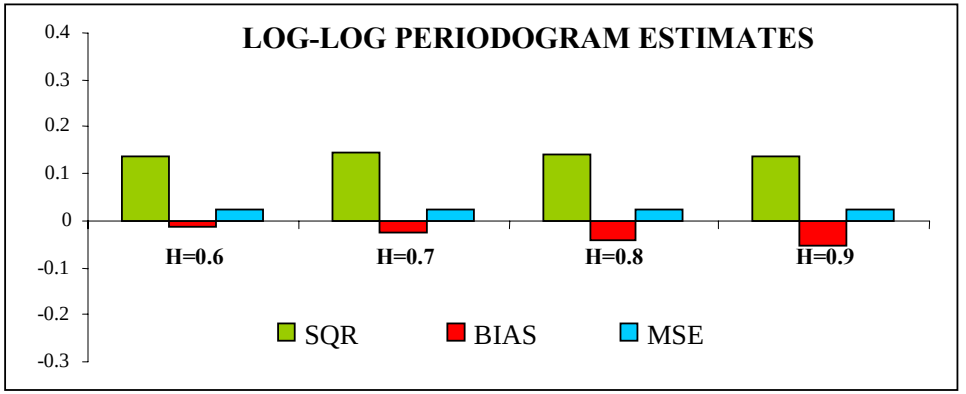
$$L_W^* = n^* \log(2\pi) - n^* \log\left[\frac{1}{n^*}\sum_{j=1}^{n^*}\frac{I^*(\lambda_j)}{f(\lambda,\boldsymbol{\kappa}^*)}\right] - \sum_{j=1}^{n^*}\log f(\lambda,\boldsymbol{\kappa}^*) - n^*$$

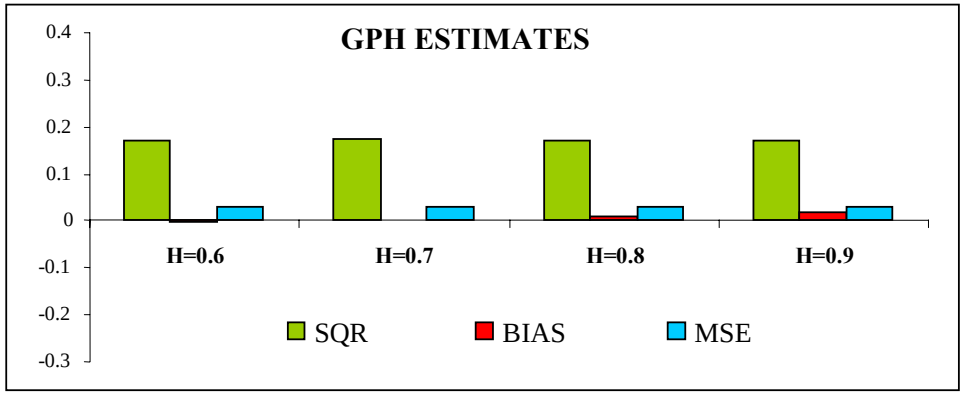
SIMULATION RESULTS

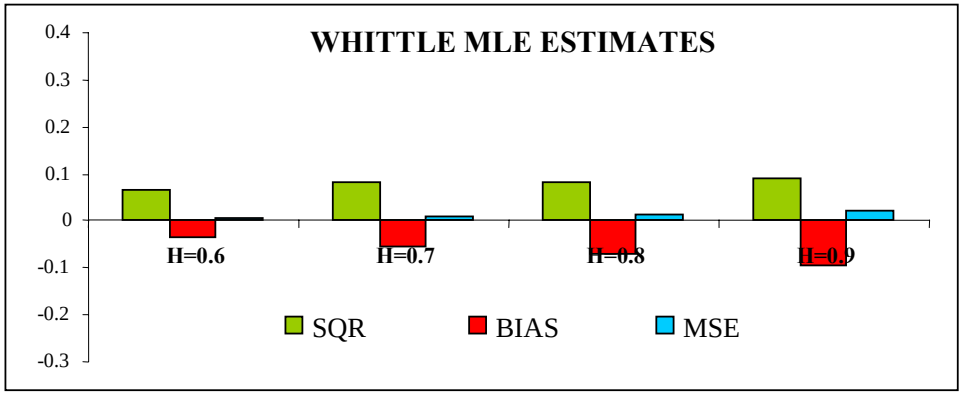












VAR MODEL OF ORDER p

$$y(t) = v + A(1)\, y(t-1) + \ldots + A(p)\, y(t-p) + \varepsilon(t) \qquad t = \pm 1, \pm 2 \ldots$$

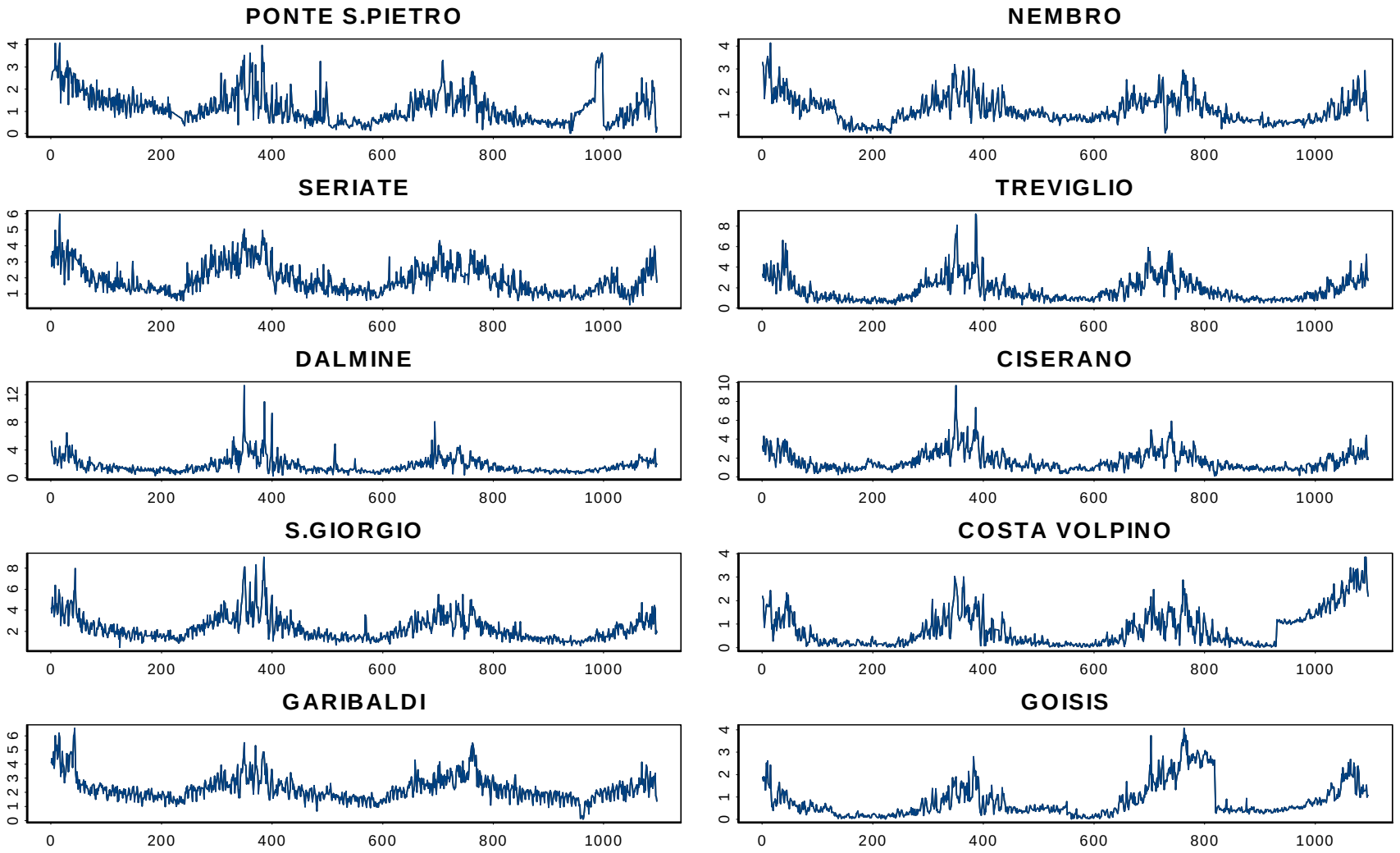
$y(t)$: ($K \times 1$) multivariate random vector

v: ($K \times 1$) vector of intercept terms

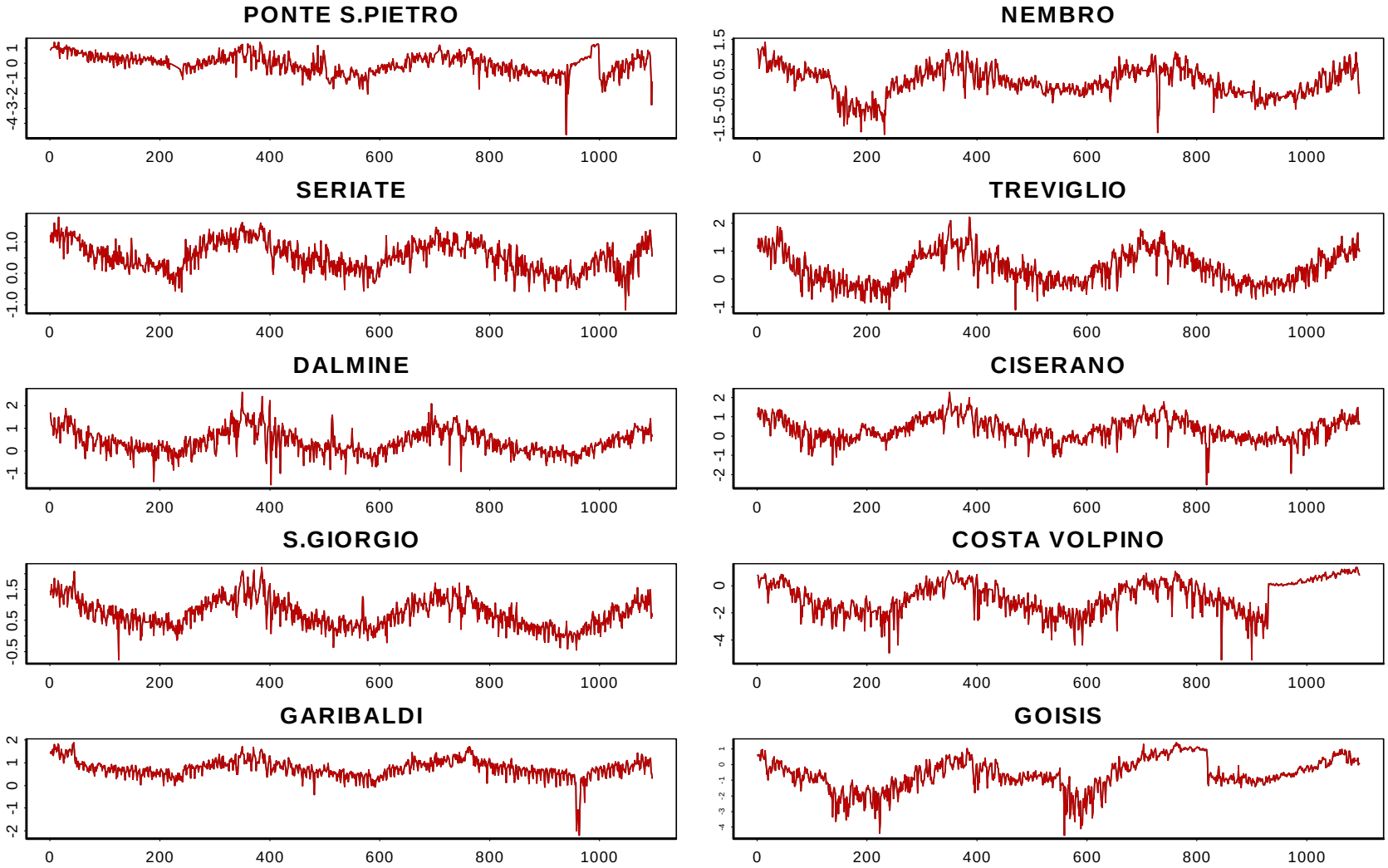
$A(1),\ldots A(p)$: ($K \times K$) matrices of coefficients

$\varepsilon(t)$: K-dimensional white noise process with non singular covariance matrix

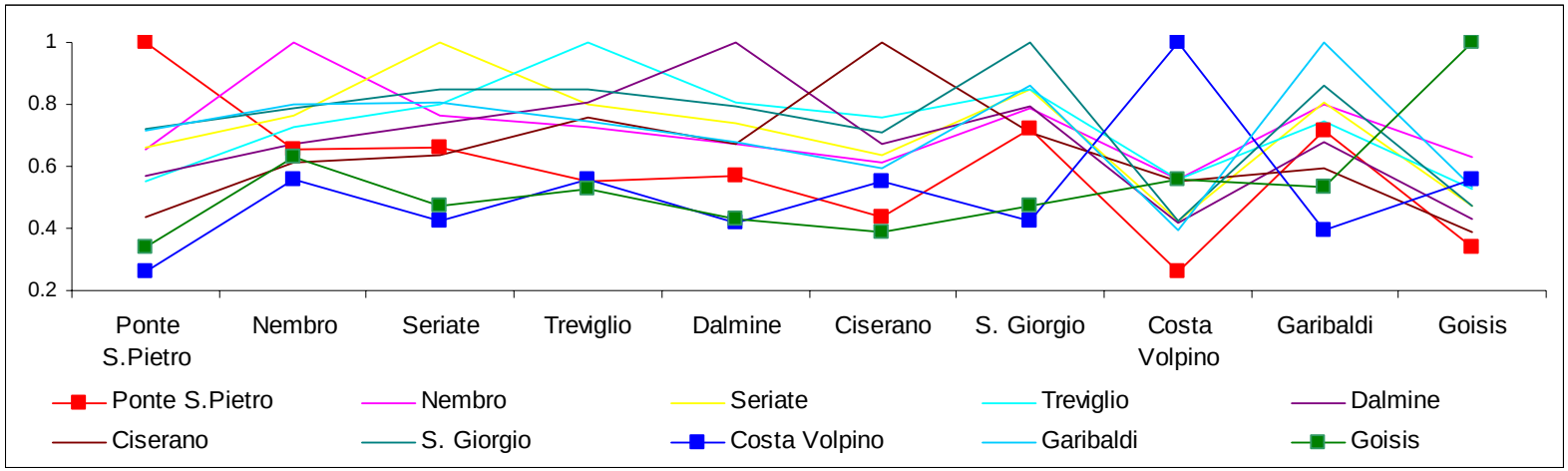
CONCENTRATION OF CARBON MONOXIDE (CO) IN BERGAMO DISTRICT



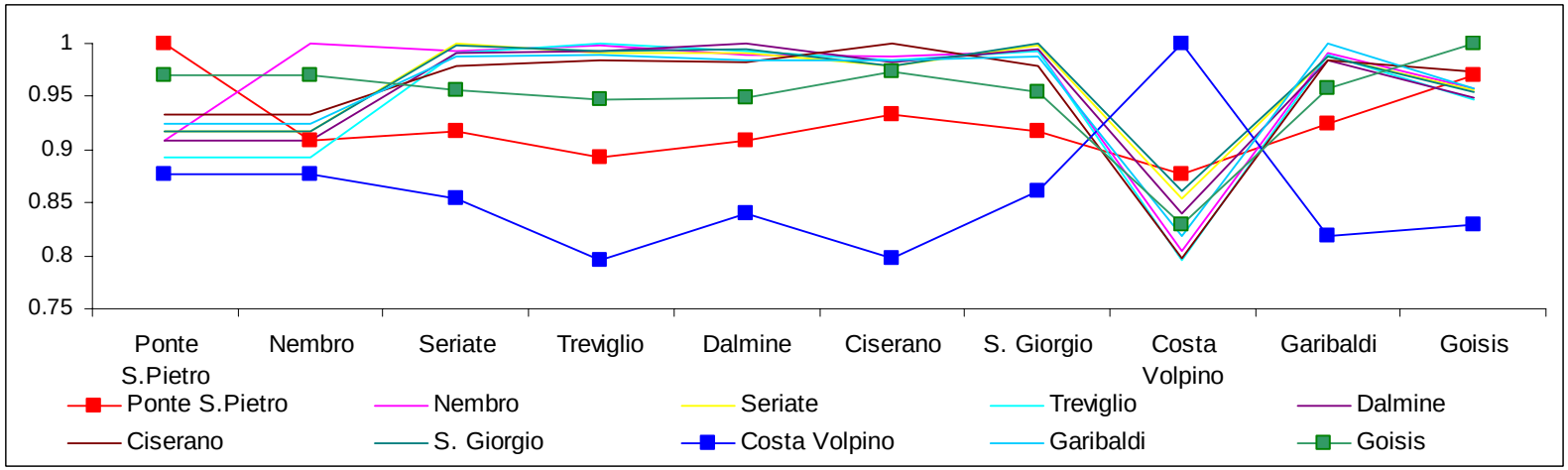
LOGARITM OF DATA



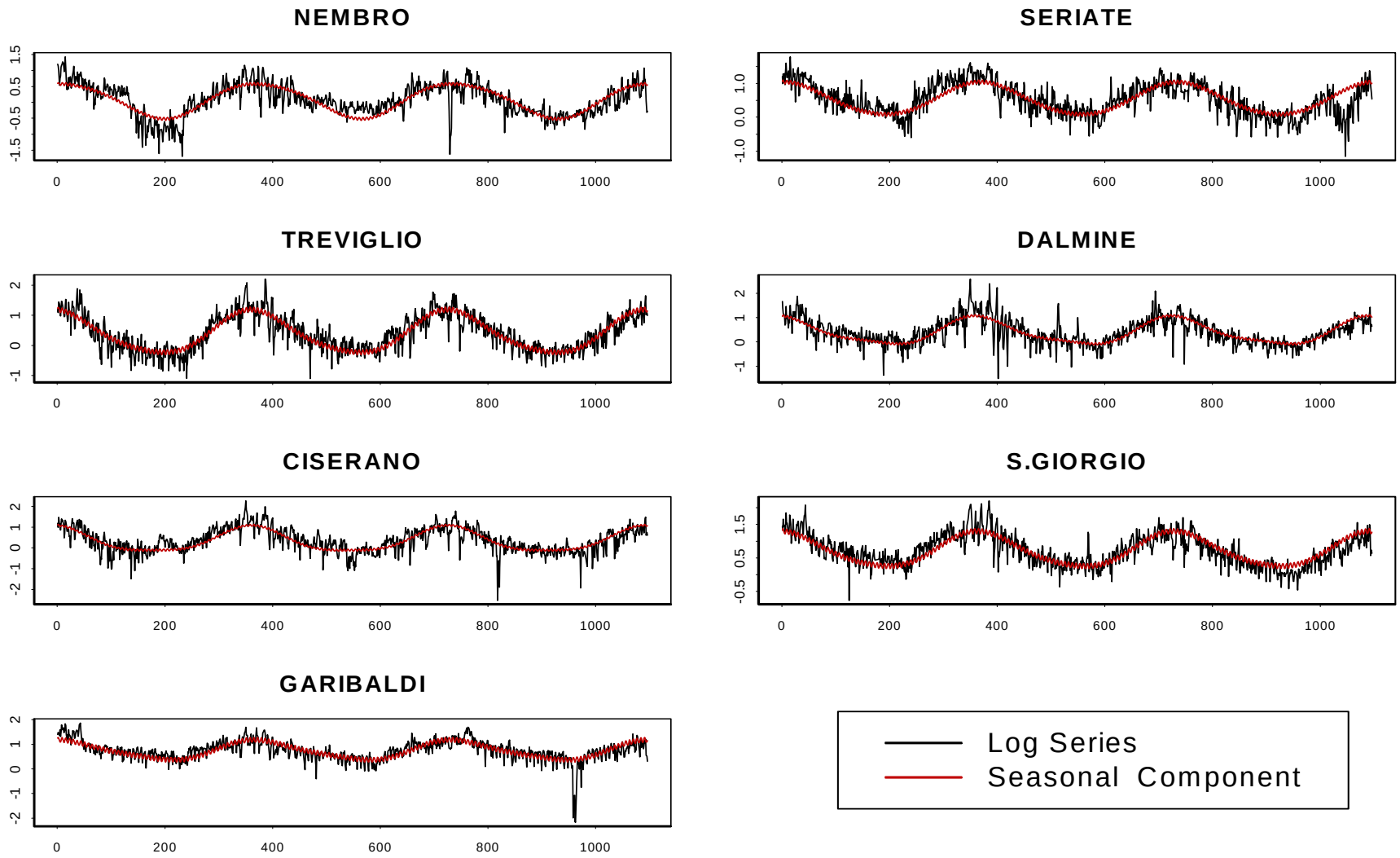
RAW TIME SERIES CORRELATIONS



PERIODOGRAM CORRELATIONS

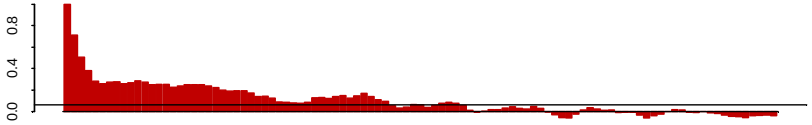


SEASONAL COMPONENT OF LOG DATA

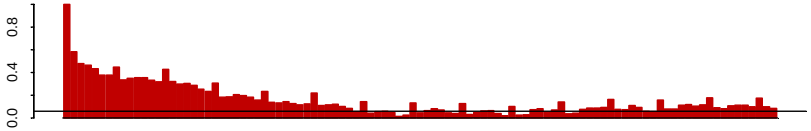


ACF OF DESEASONALISED LOG-SERIES

NEMBRO



SERIATE

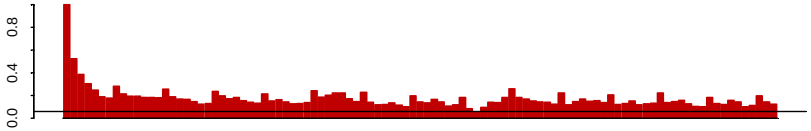


TREVIGLIO

DALMINE

CISERANO

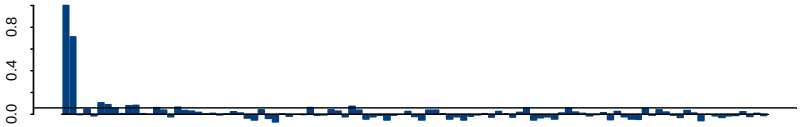
S.GIORGIO



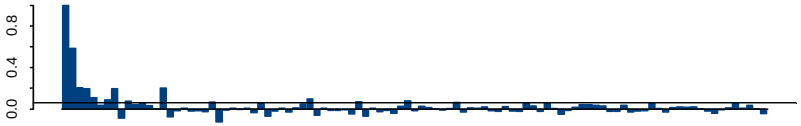
GARIBALDI

PACF OF DESEASONALISED LOG-SERIES

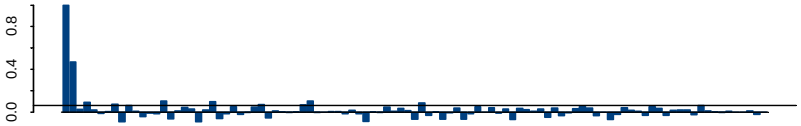
NEMBRO



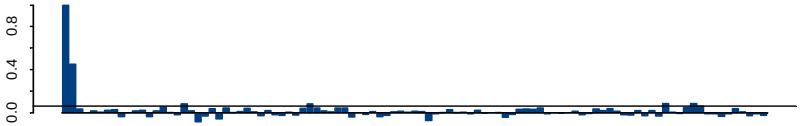
SERIATE



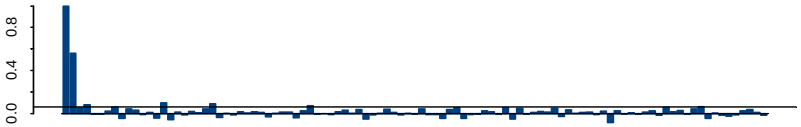
TREVIGLIO



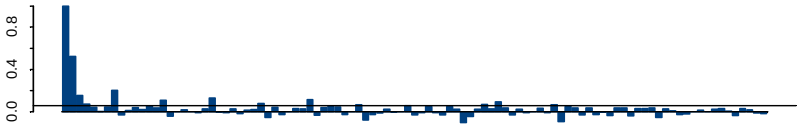
DALMINE



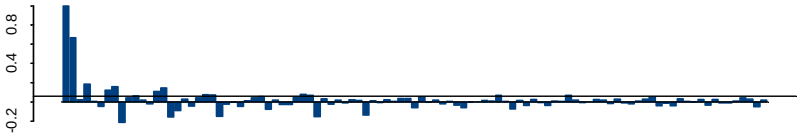
CISERANO



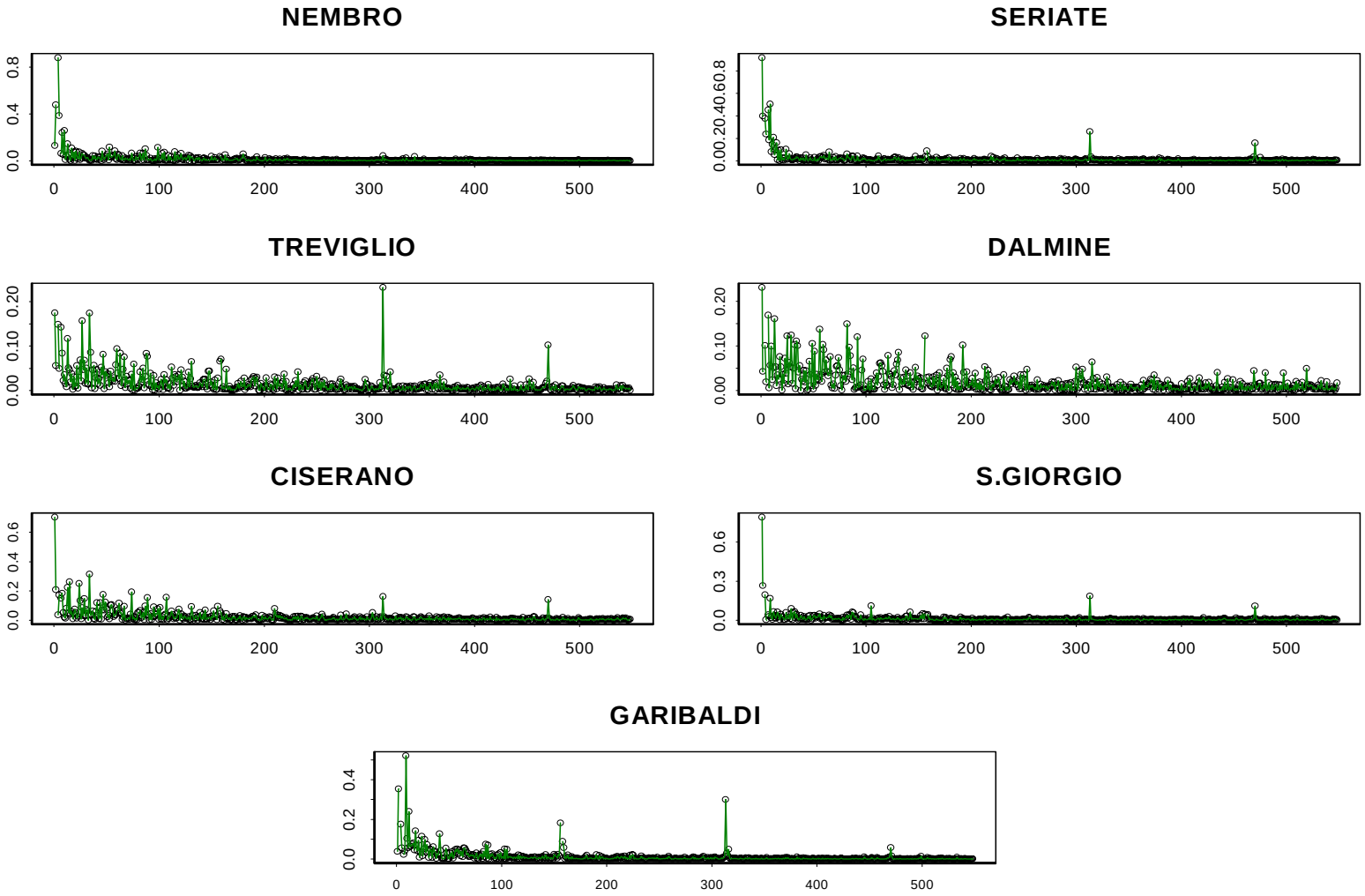
S.GIORGIO



GARIBALDI



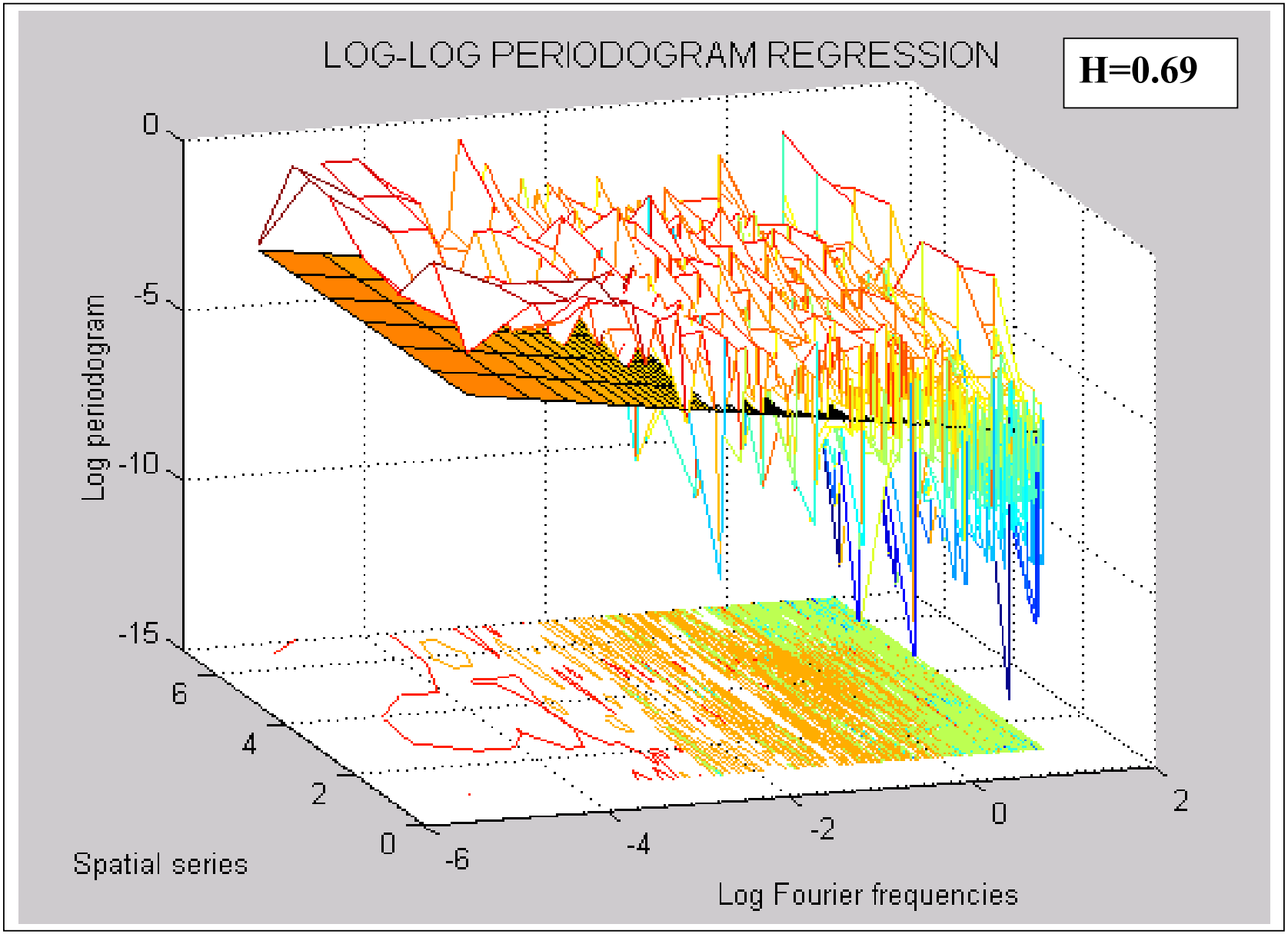
PERIODOGRAM OF DESEASONALISED LOG-SERIES

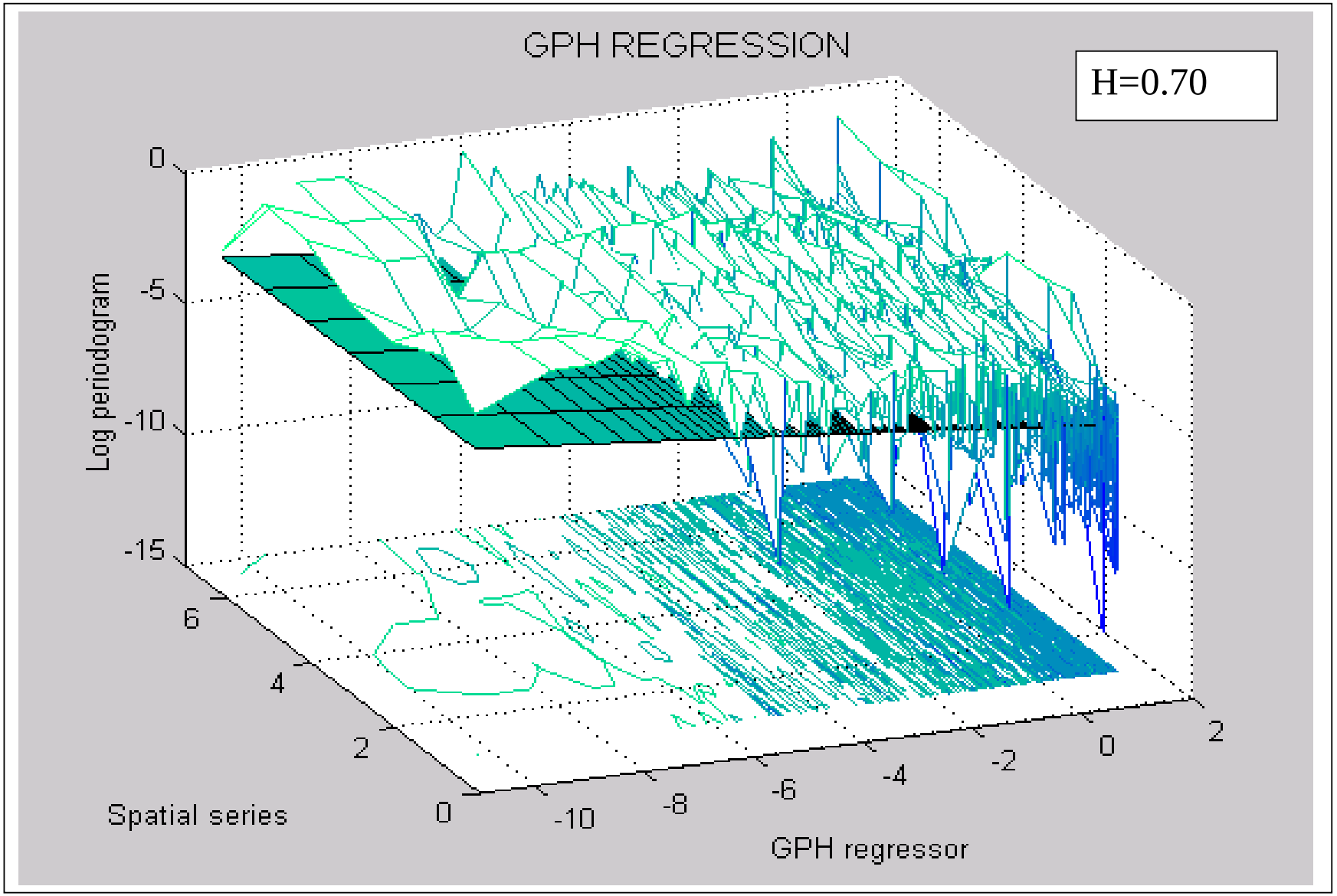


TEST FOR LONG MEMORY

1 Serie	(1)	0	4	8	10	12	26	50
V/S	(2)	1.106	0.276	0.198	0.177	0.161	0.104	0.074
R/S	(3)	127.527	63.69	53.947	51.053	48.684	39.098	32.907
KPSS_trend	(4)	-0.011	0.041	0.027	0.024	0.022	0.013	0.009
KPSS_mu	(5)	-0.02	0.071	0.047	0.042	0.037	0.023	0.016
	(6)	10	15	20	25	30	35	40
Lob-Rob	(7)	-0.222	-1.01	-1.845	-2.668	-3.469	-4.441	-5.326
2 serie		0	4	8	10	12	26	50
V/S		3.288	0.85	0.573	0.5	0.446	0.27	0.19
R/S		204.354	103.934	85.321	79.721	75.3	58.553	49.098
KPSS_trend		-0.003	0.039	0.025	0.022	0.02	0.012	0.01
KPSS_mu		-0.045	0.605	0.371	0.317	0.278	0.159	0.109
		10	15	20	25	30	35	40
Lob-Rob		-0.714	-1.427	-2.459	-3.496	-4.566	-5.603	-6.677
3 Serie		0	4	8	10	12	26	50
V/S		0.713	0.208	0.167	0.158	0.152	0.134	0.115
R/S		102.046	55.088	49.425	48.089	47.131	44.248	41.013
KPSS_trend		-0.028	0.028	0.021	0.02	0.019	0.016	0.013
KPSS_mu		-0.029	0.029	0.021	0.02	0.019	0.016	0.014
		10	15	20	25	30	35	40
Lob-Rob		-0.265	-0.466	-0.91	-1.168	-1.146	-1.011	-1.431
4 Serie		0	4	8	10	12	26	50
V/S		0.671	0.199	0.166	0.159	0.154	0.138	0.124
R/S		107.914	58.786	53.719	52.558	51.76	48.973	46.42
KPSS_trend		0	0.003	0.003	0.003	0.002	0.002	0.003
KPSS_mu		-0.023	0.2	0.158	0.149	0.144	0.125	0.11
		10	15	20	25	30	35	40
Lob-Rob		-0.278	-0.416	-0.714	-0.81	-0.94	-0.84	-1.296
5 Serie		0	4	8	10	12	26	50
V/S		1.731	0.472	0.364	0.339	0.321	0.26	0.214
R/S		158.637	82.86	72.766	70.228	68.298	61.512	55.776
KPSS_trend		-0.005	0.038	0.028	0.026	0.024	0.021	0.018
KPSS_mu		-0.032	0.252	0.182	0.167	0.156	0.122	0.098
		10	15	20	25	30	35	40
Lob-Rob		-0.826	-1.067	-1.669	-1.971	-2.482	-2.729	-3.423
6 Serie		0	4	8	10	12	26	50
V/S		3.695	1.009	0.748	0.676	0.62	0.418	0.281
R/S		217.189	113.496	97.718	92.859	88.998	73.013	59.941
KPSS_trend		0	0.007	0.005	0.005	0.005	0.004	0.004
KPSS_mu		-0.041	0.56	0.382	0.337	0.305	0.194	0.126
		10	15	20	25	30	35	40
Lob-Rob		-1.685	-2.744	-3.747	-4.792	-5.669	-6.434	-7.482
7 Serie		0	4	8	10	12	26	50
V/S		0.613	0.156	0.11	0.099	0.091	0.066	0.057
R/S		122.553	61.791	51.987	49.3	47.306	40.252	37.225
KPSS_trend		-0.002	0.016	0.01	0.009	0.008	0.006	0.006
KPSS_mu		-0.021	0.143	0.093	0.082	0.074	0.051	0.043
		10	15	20	25	30	35	40
Lob-Rob		0.114	-0.014	-0.271	-0.679	-1.041	-1.602	-2.251

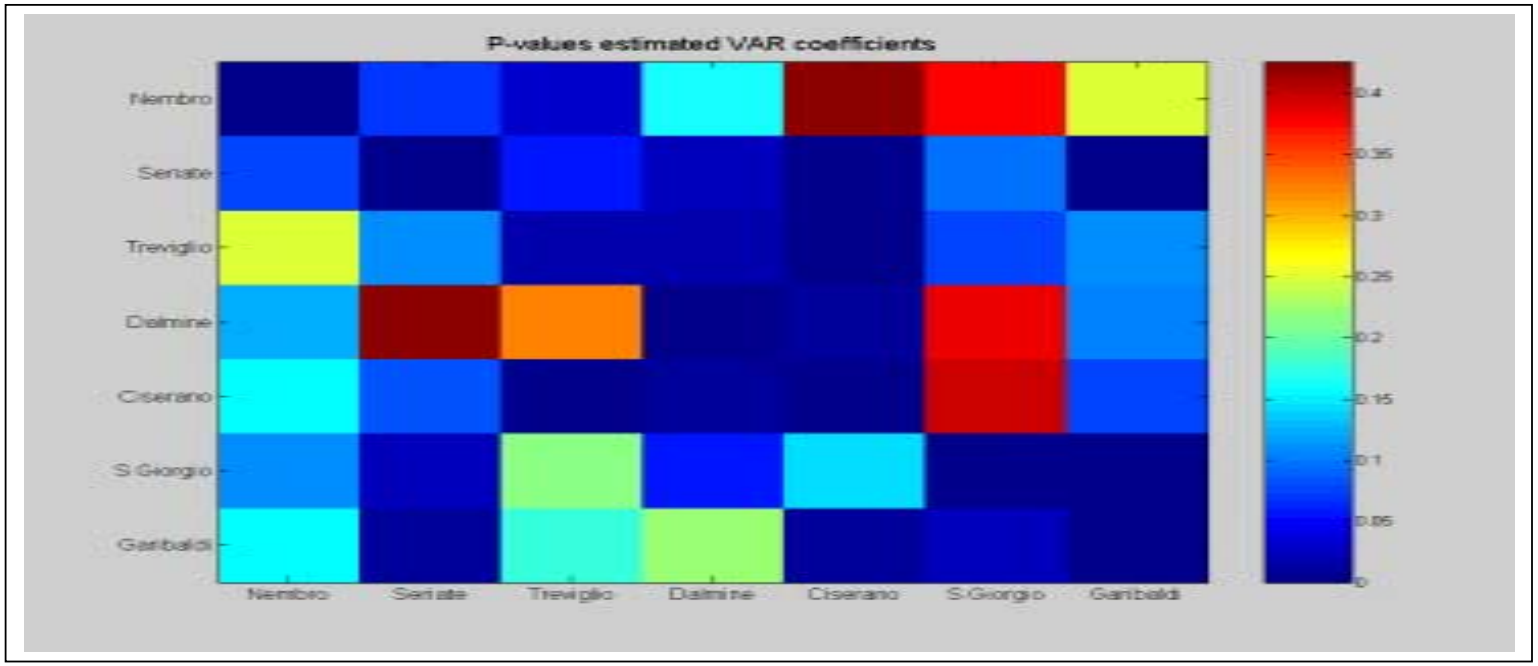
(1): Truncation lag of correlogram
(2): Critical value V/S test= 0.1869
(3): Critical value R/S test= 1.747
(4): Critical value KPSS1 test= 0.146
(5): Critical value KPSS2 test= 0.146
(6): Critical value LR test=1.96
(7): Truncation frequency of periodogram


LOG-LOG PERIODOGRAM REGRESSION
H=0.69
Log periodogram
0
-5
-10
-15
Spatial series
6
4
2
0
Log Fourier frequencies
-6
-4
-2
0
2

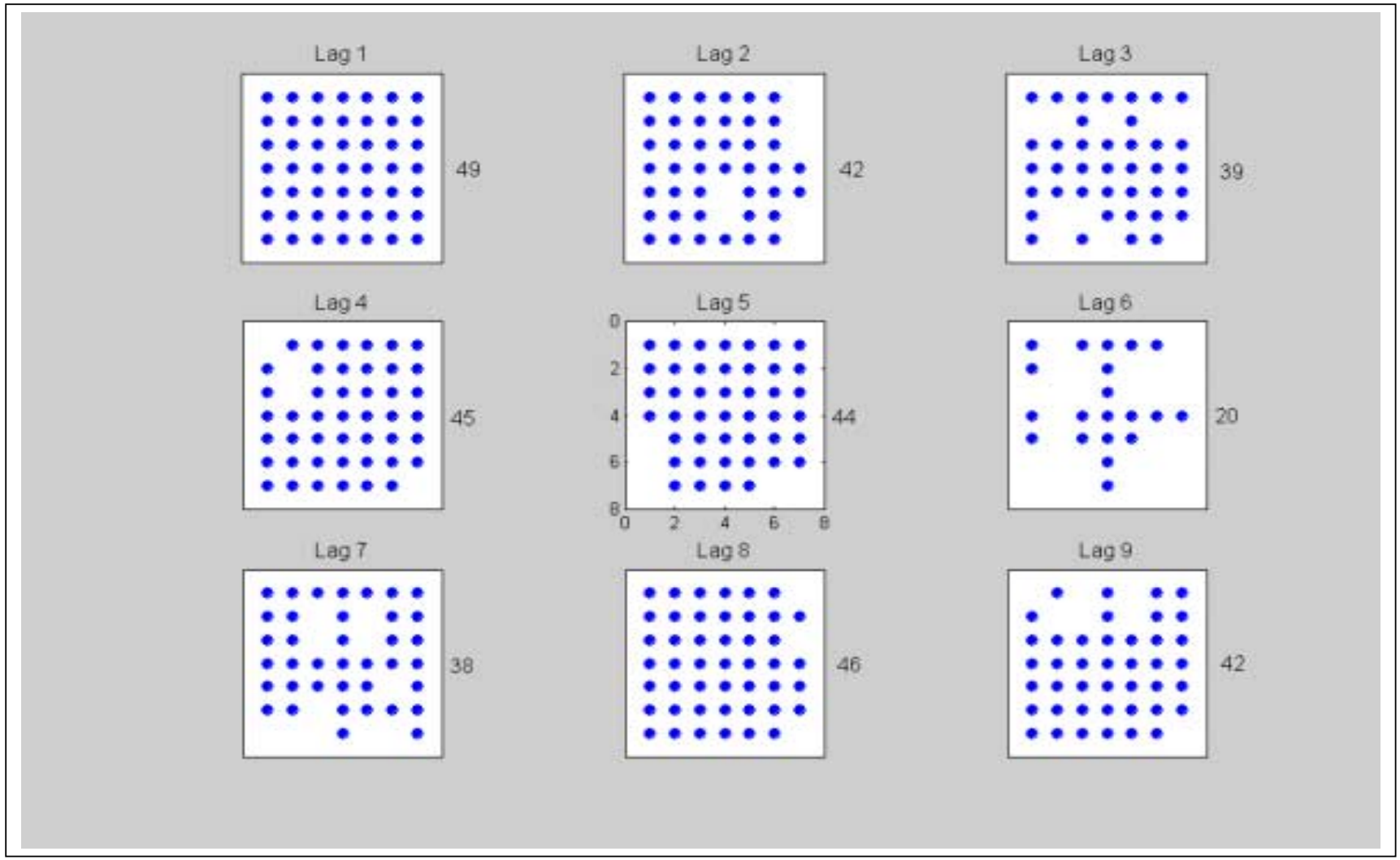

GPH REGRESSION
H=0.70
Log periodogram
0
-5
-10
-15
Spatial series
6
4
2
0
GPH regressor
-10
-8
-6
-4
-2
0
2

VAR(1) MODEL ON THE RESIDUAL SHORT MEMORY COMPONENT

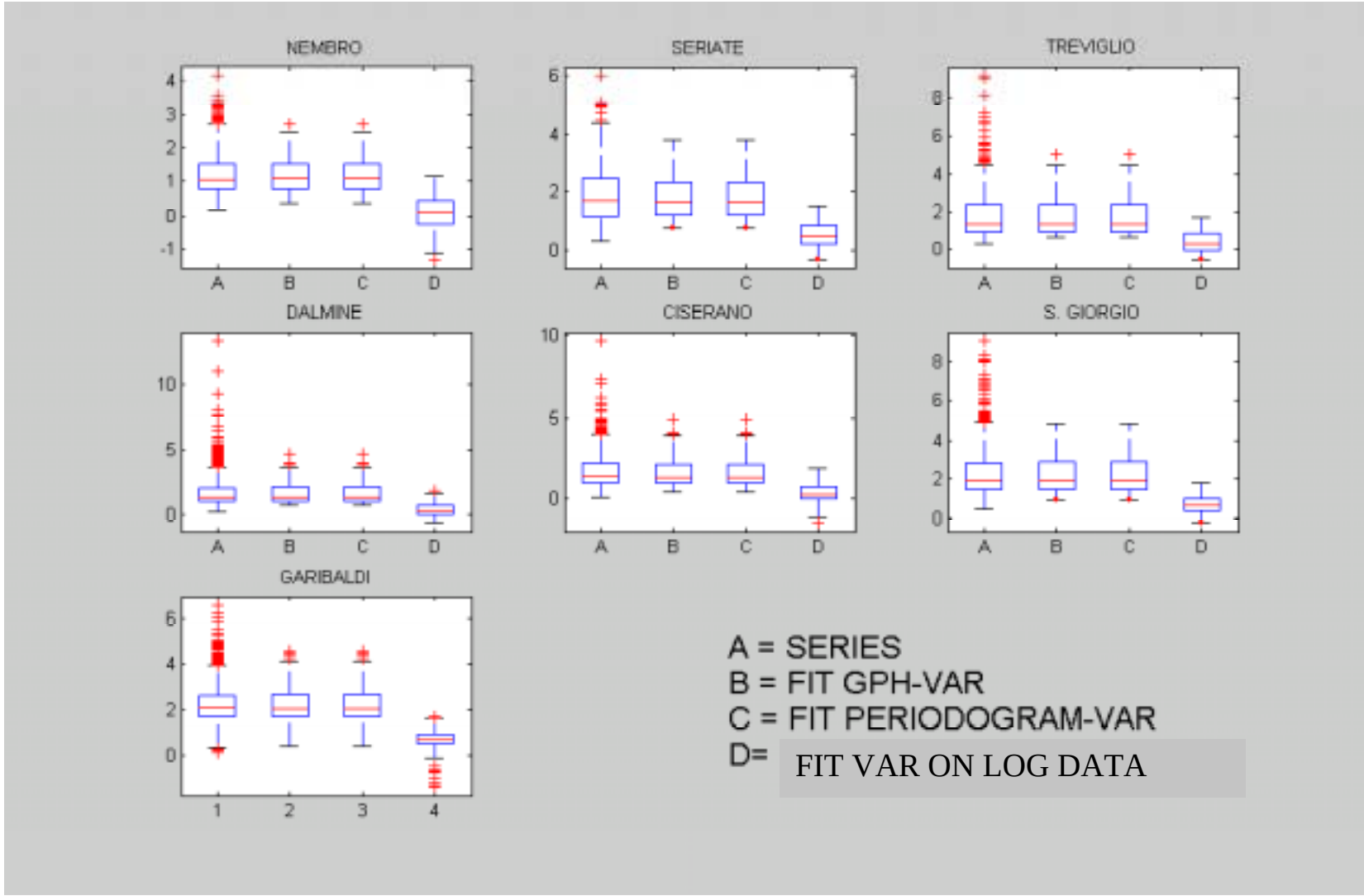
	Nembro Lag(-1)	Seriate Lag(-1)	Treviglio Lag(-1)	Dalmine Lag(-1)	Ciserano Lag(-1)	S. Giorgio Lag(-1)	Garibaldi Lag(-1)
Nembro	0.4720	-0.0465	0.0680	0.0223	0.0055	-0.0137	0.0244
Seriate	0.0509	0.1647	-0.0672	0.0514	0.0891	-0.0653	0.1266
Treviglio	-0.0261	-0.0474	0.0981	0.0606	0.1153	0.0753	0.0542
Dalmine	0.0523	-0.0087	0.0243	0.1758	0.0952	0.0185	0.0659
Ciserano	-0.0454	-0.0633	-0.1408	0.0730	0.3983	0.0167	0.0742
S. Giorgio	0.0399	-0.0652	-0.0302	0.0384	0.0326	0.1420	0.1331
Garibaldi	0.0302	-0.0699	-0.0334	0.0168	0.0631	-0.0835	0.4887



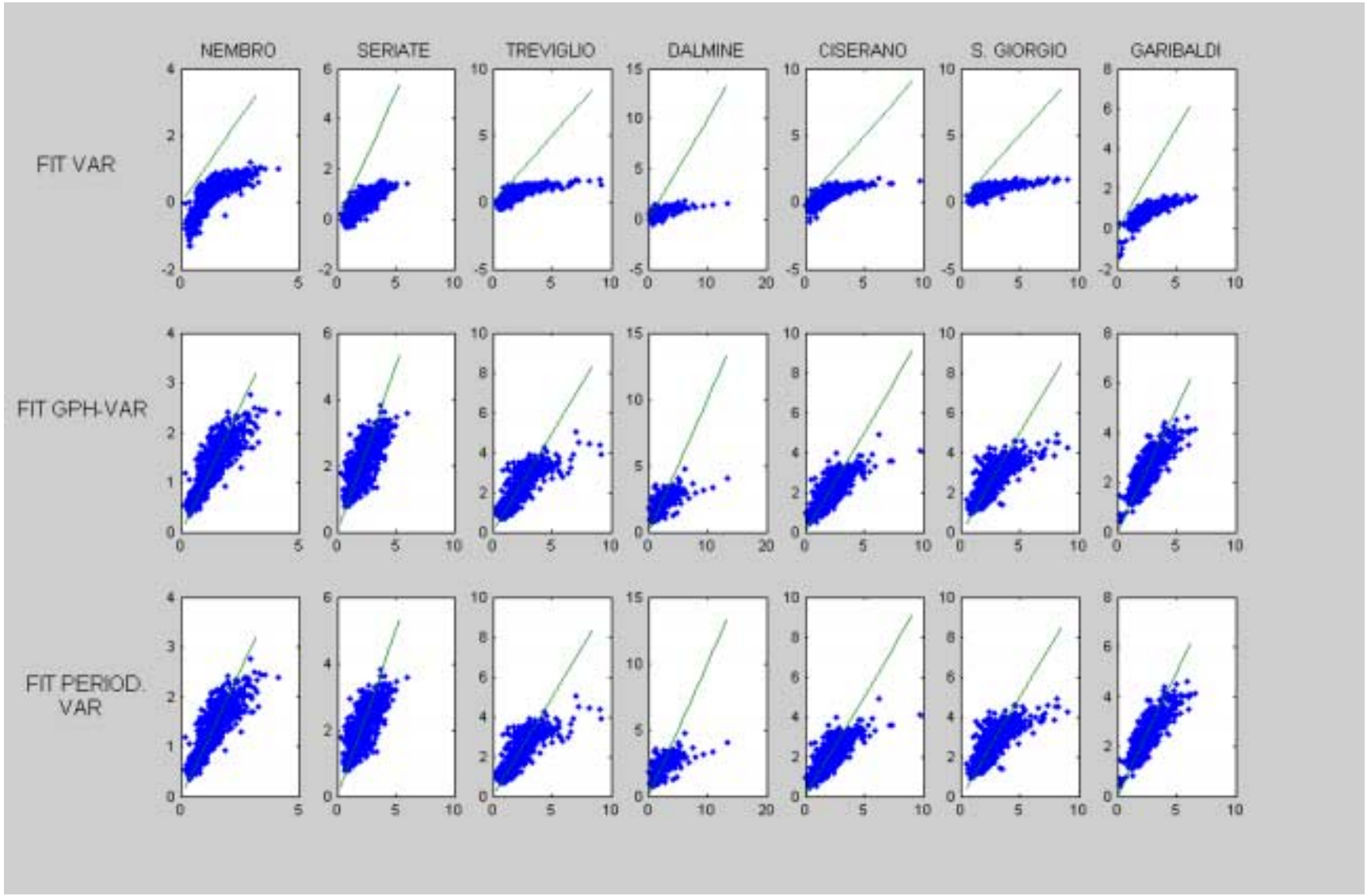
NON SIGNIFICANT ACF COEFFICIENTS FOR RESIDUALS FROM VAR(1) MODEL



GOODNESS OF FIT FOR ESTIMATED MODELS



SCATTER PLOT FOR ESTIMATED MODELS



FORECASTING FROM ESTIMATED MODELS