
Στοχαστικά Σήματα και Συστήματα

Θέμα μαθήματος: Railway-data-50

Όνομα : Χρήστος

Επώνυμο : Σταμόπουλος

ΑΜ : 1057168

Έτος : Δ'

Διδάσκων Καθηγητής: Σ.Δ. Φασόης

26.6.2020

Περιεχόμενα

Τμήμα Α – Προκαταρκτικά	3
Α1. Προκαταρκτικά.	3
Α2. Κανονικοποίηση.	4
Α3. Προκαταρκτική ανάλυση στο πεδίο του χρόνου.....	5
1. Εξέταση Κανονικότητας Σήματος.....	5
2. Εξέταση Αυτοσυμμεταβλητότητας.	6
Τμήμα Β: Μη παραμετρική ανάλυση στο πεδίο χρόνου.	8
Β1: Μη παραμετρική ανάλυση στο πεδίο του χρόνου.	8
Β2: Μη παραμετρική ανάλυση στο πεδίο συχνοτήτων.	10
1. Διακριτός Μετασχηματισμός Fourier (Discrete Fourier Transform, DFT) ...	10
2. Περιοδόγραμμα.	12
3. Εκτίμηση φάσματος μέσω της μεθόδου Welch	13
4. Σύγκριση των μεθόδων Welch – Blackman-Tukey - Περιοδογράμματος ως προς την εκτίμηση του φάσματος.	17
Β3: Έλεγχος στασιμότητας.	18
Τμήμα Γ – Παραμετρική ανάλυση ARMA.	23
Γ1: Προκαταρκτικά.....	23
Γ2: Εκτίμηση μοντέλων ARMA.	23
Γ3: Έλεγχος εγκυρότητας του επιλεγέντος μοντέλου ARMA.	27
1. Τμήμα ελέγχου εγκυρότητας.....	27
2. Τμήμα εκτίμησης.....	29
Γ4: Ανάλυση του επιλεγέντος μοντέλου ARMA.	33
Γ5: Πρόβλεψη βάσει του επιλεγέντο μοντέλου ARMA.	39

Τμήμα Α – Προκαταρκτικά.

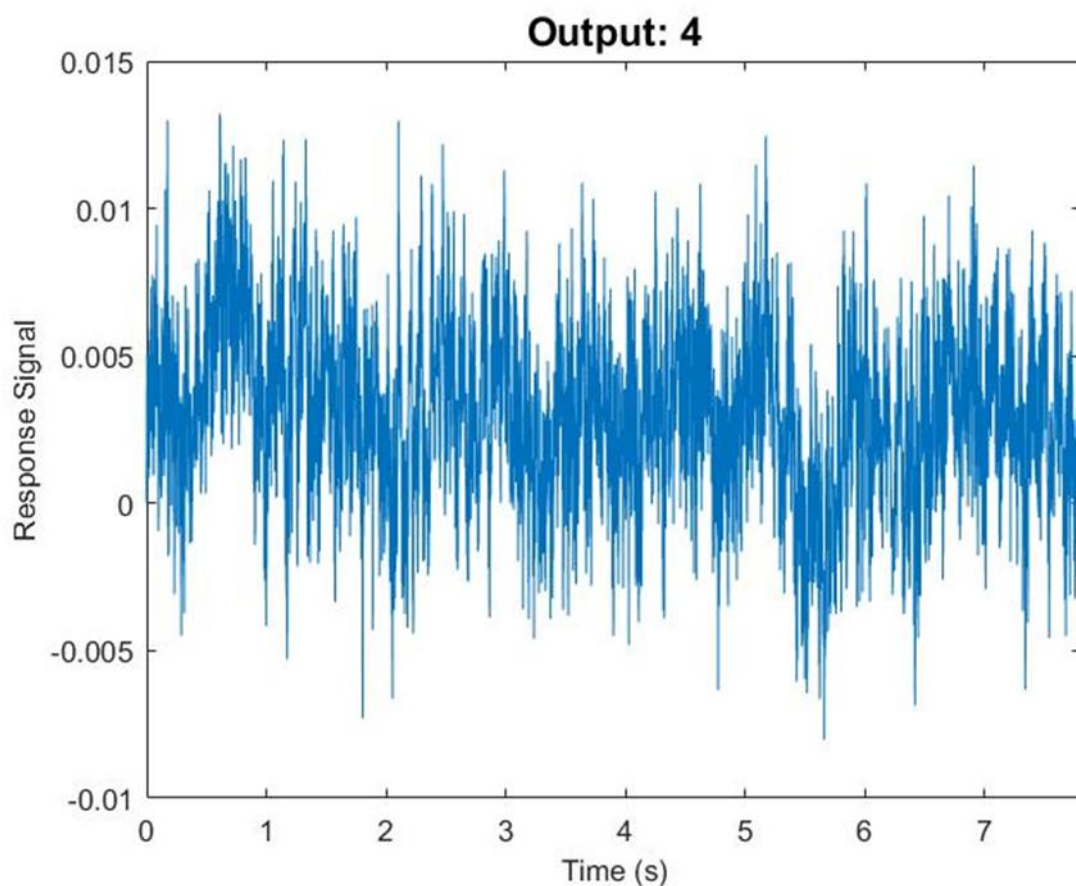
A1. Προκαταρκτικά.

Το σήμα που επιλέχθηκε για την ανάλυση είναι το **‘Railway-data-50’**. Έτσι έχουμε:

Μήκος Σήματος N : 4000 samples/τιμές.

Περίοδος Δειγματοληψίας Ts : $T_s = 1/f_s = 1/512 = 0.00195$ sec.

Παρακάτω παρουσιάζεται το σήμα που μετρήθηκε από τον 4^ο αισθητήρα, ο οποίος λαμβάνει τιμές σχετικά με την κατακόρυφη επιτάχυνση του car body.



Εικόνα 1: Το σήμα Railway-data-50 συναρτήσει του χρόνου.

Στην παραπάνω εικόνα απεικονίζεται το σήμα συναρτήσει του χρόνου.

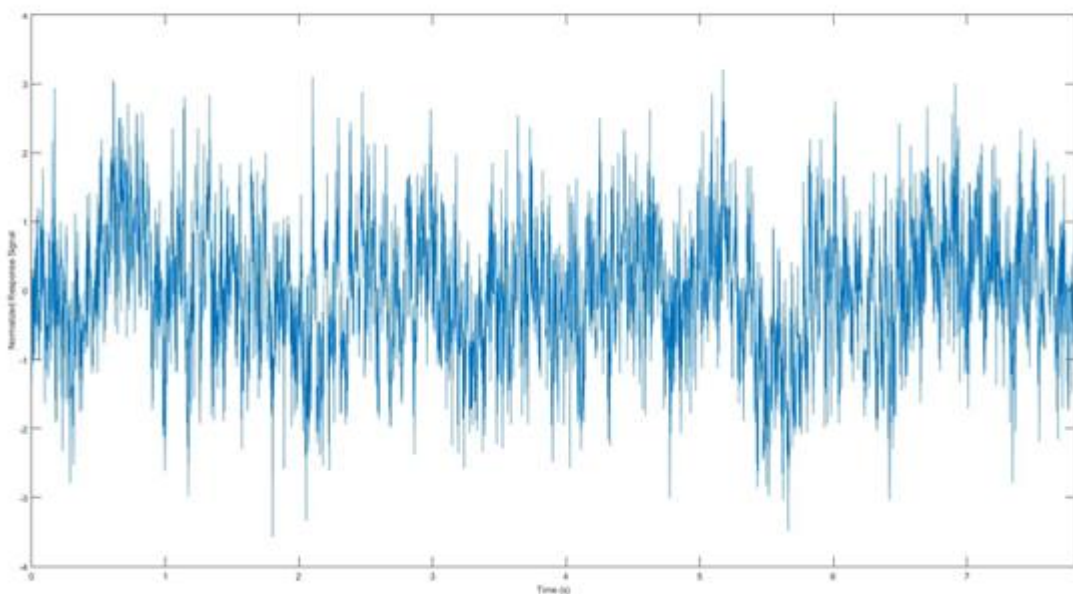
- **Παρατηρήσεις:**

Σύμφωνα με το διάγραμμα, εύκολα παρατηρείται πως η μέση τιμή του σήματος παραμένει σταθερή και δεν μεταβάλλεται με τον χρόνο. Επομένως, λόγω αυτής της σταθερότητας της μέσης τιμής διαπιστώνουμε πως το σήμα είναι στάσιμο.

A2. Κανονικοποίηση.

Σύμφωνα με την προηγούμενη αιτιολόγηση, το σήμα είναι στάσιμο και οι τιμές στις οποίες διακυμαίνεται είναι αρκετά μικρότερες της μονάδας. Με αυτόν τον τρόπο, απαραίτητη θα ήταν η εφαρμογή της κανονικοποίησης του σήματος για να είναι πιο ευκρινές. Έτσι, με την βοήθεια του προγράμματος Matlab, κεντράρουμε το σήμα – με την εντολή `detrend` – και το κανονικοποιούμε. Παρακάτω παρατίθεται το αποτέλεσμα της διαδικασίας.

Output: 4



Εικόνα 2: Το κεντραρισμένο και κανονικοποιημένο σήμα συναρτήσει του χρόνου.

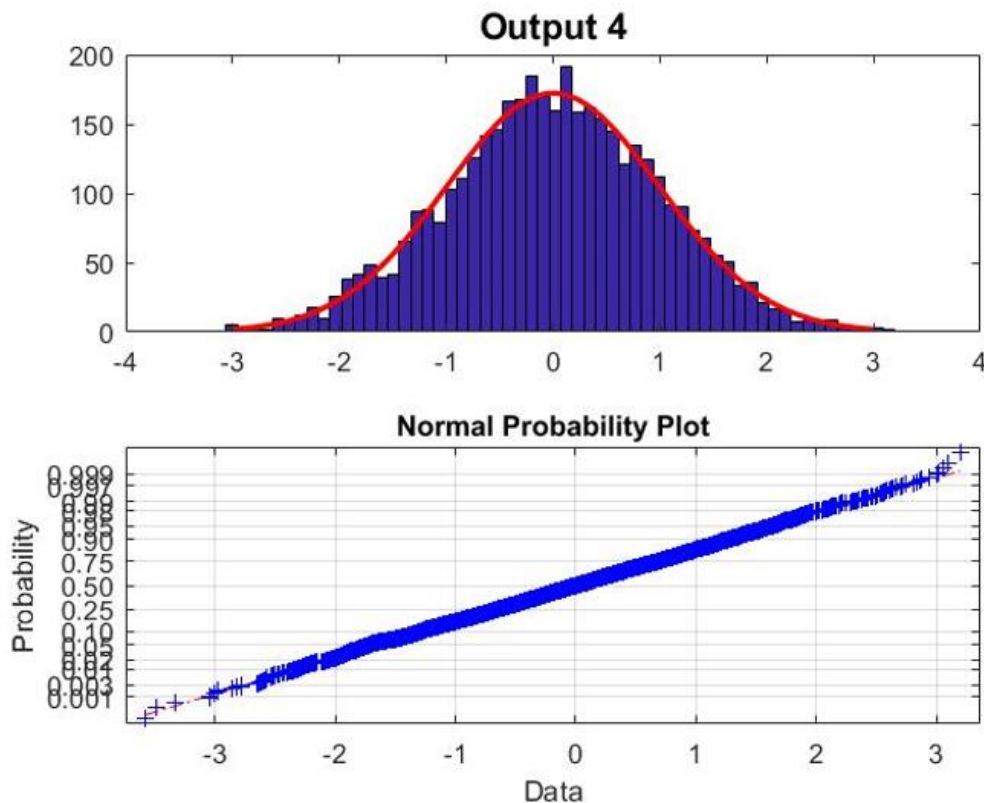
- **Παρατηρήσεις:**

Λόγω της κακής ανάλυσης του Print Screen στην παραπάνω εικόνα, βλέπουμε πως η κανονικοποίηση του σήματος φτάνει σε τιμές κοντά στο 3.5. Τέλος, εύκολα φαίνεται πως, με την εφαρμογή της κανονικοποίησης, το σήμα 'διαβάζεται' πιο εύκολα.

A3. Προκαταρκτική ανάλυση στο πεδίο του χρόνου

1. Εξέταση Κανονικότητας Σήματος.

Η εξέταση της κανονικότητας του σήματος, γίνεται μέσω της συνάρτησης πυκνότητας πιθανότητας του ιστογράμματος (εμπειρική συνάρτηση πυκνότητας πιθανότητας) και normal probability plot.



Εικόνα 3: Συνάρτηση Πυκνότητας Πιθανότητας και το Normal Probability Plot.

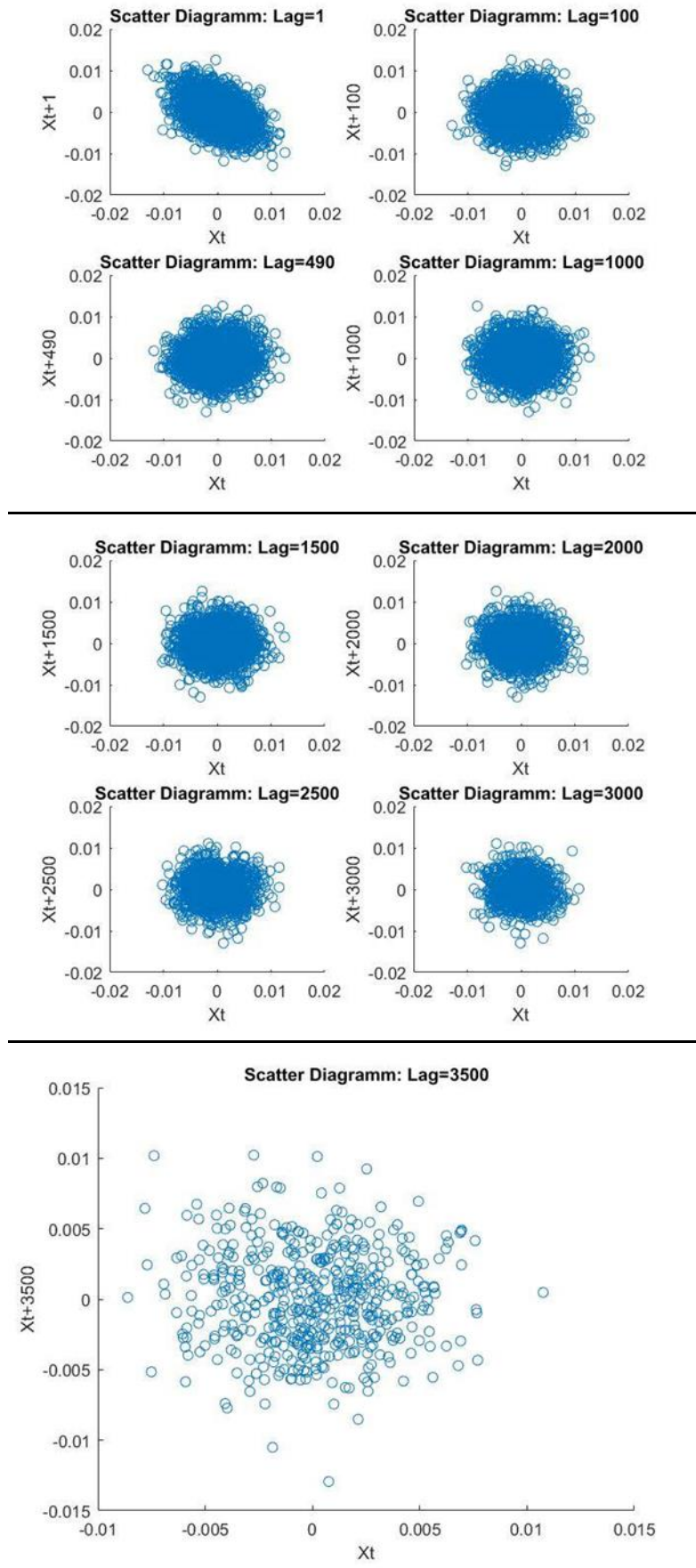
○ Παρατηρήσεις:

1. Αρχικά, όσον αφορά το 1^ο διάγραμμα από την εικόνα 3, παρατηρούνται σημεία όπου οι τιμές του σήματος βρίσκονται πάνω ή και κάτω από την κόκκινη γραμμή, η οποία υποδηλώνει την κανονική κατανομή. Αυτό σημαίνει ότι, το σήμα μας ακολουθεί την γκαουσιανή θεωρώντας τις αποκλίσεις από την κανονικότητα ως αποδεκτές.
2. Σχετικά με το normal probability plot, παρατηρείται πως στις άκρες οι τιμές του σήματος αποκλίνουν από την κόκκινη γραμμή. Με αυτόν τον τρόπο, υπολογίζεται η λοξότητα από τον τύπο : $\mu = E\{X_{t3}\} = 1.16 \cdot 10^{-7}$

2. Εξέταση Αυτοσυμμεταβλητότητας.

Παρακάτω, για την εύρεση της δυναμικότητας, παρατίθεται η εικόνα 4, η οποία απεικονίζει τα διαγράμματα Scatter για $k = \text{lag} = 1, 100, 490, 1000, 1500, 2000, 2500, 3000, 3500$. Όταν το διάγραμμα έχει ελλειψοειδής μορφή και είναι πυκνό, σημαίνει πως για το συγκεκριμένο k υπάρχει μεγάλη εξάρτηση της τιμής X_{t+k} από την τιμή X_t . Πιο συγκεκριμένα, εάν αυτή η ελλειψοειδής μορφή έχει μια κλίση προς τα αριστερά τότε υπάρχει μια αρνητική συσχέτιση των X_{t+k} και X_t , ενώ αν συγκλίνει προς τα δεξιά υπάρχει μια θετική συσχέτιση των τιμών αυτών.

Απεναντίας, όσο πιο αραιό και πιο κυκλικό είναι το διάγραμμα, τόσο μικρότερη εξάρτηση έχουν οι τιμές X_{t+k} και X_t , μεταξύ τους.



Εικόνα 4: Διαγράμμα Scatter για διάφορες τιμές του κ .

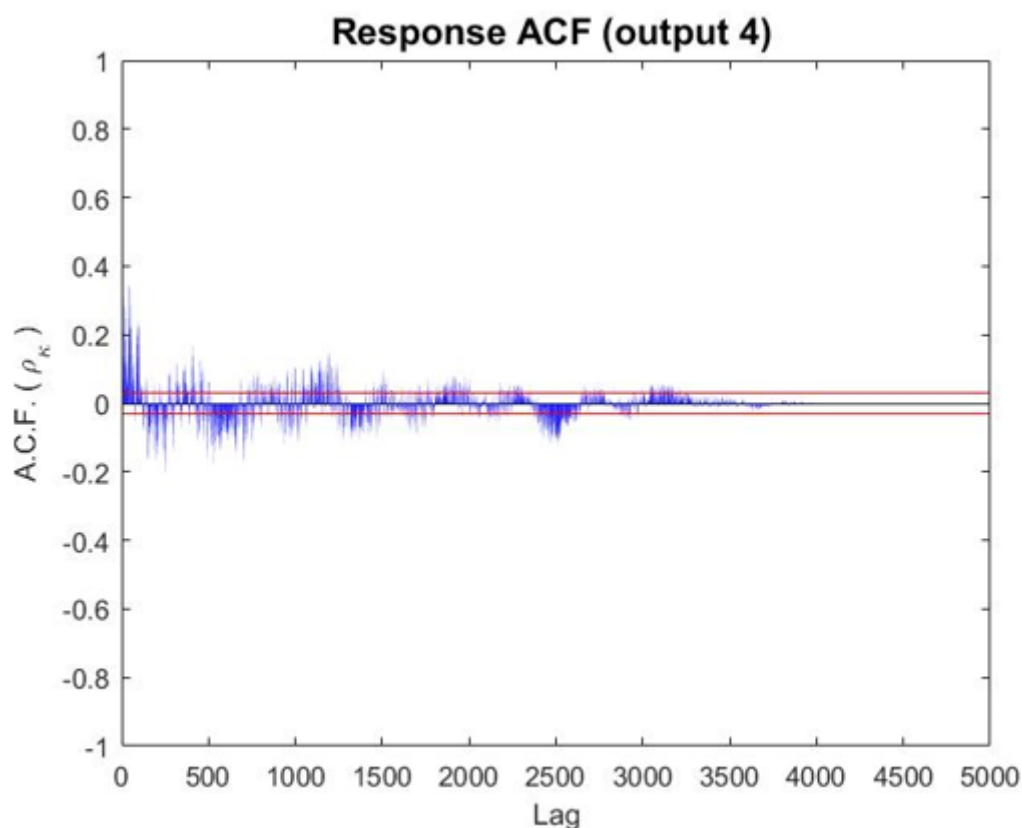
○ **Παρατηρήσεις:**

- ✓ **K=1:** Παρατηρείται ελλειψοειδής μορφή της εικόνας 4, μεγάλη πύκνωση αλλά και αρνητική κλίση. Άρα, έχουμε μεγάλη εξάρτηση και μια αρνητική συσχέτιση της τιμής X_{t+1} από την τιμή της X_t , το οποίο θεωρείται απολύτως λογικό.
- ✓ **K=1000-2500:** Στην συνέχεια όσο αυξάνεται το κ παρατηρείται και κυκλική μορφή του διαγράμματος Scatter για κάθε κ που εξετάζεται. Πιο συγκεκριμένα, για $\kappa = 1000 - 2500$, παρατηρείται σχεδόν κυκλική μορφή των διαγραμμάτων με συνέπεια να υπάρχει μικρή εξάρτηση της τιμής $X_{t+\kappa}$ από την τιμή της X_t . Αυτό μπορεί να φανεί και από το διάγραμμα της ανηγμένης αυτοσυμμεταβλητότητας ρ_κ που παρατίθεται παρακάτω στο ερώτημα B1. Επειδή οι τιμές της ανηγμένης αυτοσυμμεταβλητότητας για $\kappa = 1000 - 2500$ είναι κοντά στο όριο του λευκού θορύβου, αλλά όχι μέσα σε αυτό, τότε λογικό είναι να υπάρχει μικρή συσχέτιση μεταξύ των τιμών $X_{t+\kappa}$ και X_t .
- ✓ **K=3500:** Όσον αναφορά για $\kappa=3500$, δεν παρατηρείται εξάρτηση της τιμής X_{t+3500} από την τιμή της X_t διότι το διάγραμμα Scatter είναι πάρα πολύ αραιό, με συνέπεια να έχει χαθεί η αυτοσυμμεταβλητότητα.

Τμήμα Β: Μη παραμετρική ανάλυση στο πεδίο χρόνου.

B1: Μη παραμετρική ανάλυση στο πεδίο του χρόνου.

Σύμφωνα με το πρόγραμμα "Matlab" και την εντολή "corrcoef", δημιουργείται το διάγραμμα της ανηγμένης αυτοσυμμεταβλητότητας ρ_κ , το οποίο φαίνεται στην εικόνα 5.



Εικόνα 5: Διάγραμμα Ανηγμένης Αυτοσυμμεταβλητότητας.

ο **Παρατηρήσεις:**

Όπως φαίνεται από την εικόνα 5, η ανηγμένη αυτοσυμμεταβλητότητα ρ_k παίρνει τιμές μεταξύ των ορίων του λευκού θορύβου ($\pm 1.96/\sqrt{N}$), μετά από περίπου 3200 χρονικές στιγμές σε σχέση με την παρούσα χρονική στιγμή (lag = 3200).

Επίσης, όσον αφορά το επίπεδο σφάλματος 5%, παρατηρείται πως μετά από 3200 χρονικές στιγμές, τουλάχιστον το 95% των τιμών της ρ_k βρίσκονται μέσα στα όρια του λευκού θορύβου (κόκκινες οριζόντιες γραμμές). Με άλλα λόγια, οι τιμές του σήματος μέχρι την χρονική στιγμή $t+3200$ θα έχουν συσχέτιση με την τιμή του σήματος την χρονική στιγμή t .

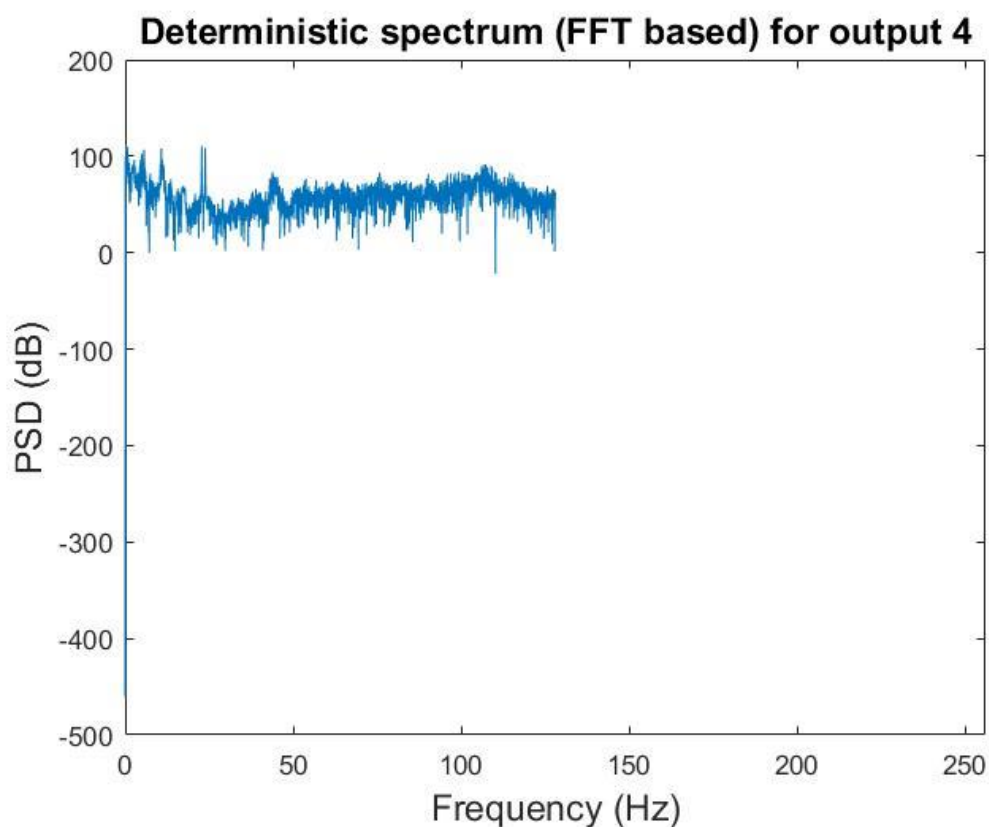
Επιπλέον, από την χρονική στιγμή $t+3200$ και μετά, πάλι δεν δεχόμαστε την λευκότητα. Αυτό συμβαίνει διότι, ακόμα και που όλα τα ρ_k βρίσκονται μέσα στην ζώνη του λευκού θορύβου παρατηρούμε πως, λόγω της

ημιτονοειδής μορφής τους, υπάρχει συσχέτιση γιατί θα' πρεπε τα ρ_k να είναι μεταξύ τους στατιστικά ασυσχέτιστα.

B2: Μη παραμετρική ανάλυση στο πεδίο συχνοτήτων.

1. Διακριτός Μετασχηματισμός Fourier (Discrete Fourier Transform, DFT)

Σύμφωνα με το πρόγραμμα "Matlab" και την εντολή "fft", υπολογίζεται ο διακριτός μετασχηματισμός Fourier του σήματος. Στην εικόνα 6 απεικονίζεται ο διακριτός μετασχηματισμός Fourier συναρτήσει της συχνότητας σε Hz (από 0 έως $f_s/2$).



Εικόνα 6: Διακριτός Μετασχηματισμός Fourier συναρτήσει Συχνότητας.

○ **Παρατηρήσεις:**

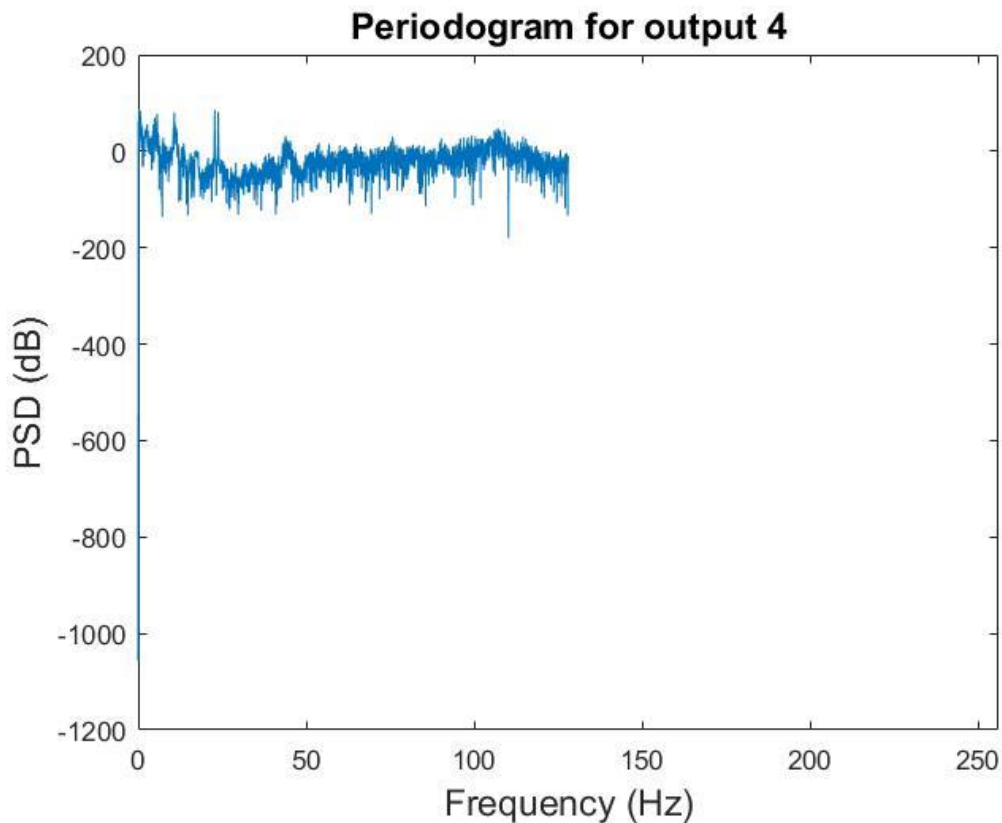
Από το παραπάνω διάγραμμα, παρατηρείται πως υπάρχουν τουλάχιστον 5 κορυφές του σήματος, δηλαδή τουλάχιστον 5 συχνότητες συντονισμού. Αυτό σημαίνει ότι υπάρχουν 5 ζευγάρια μιγαδικών πόλων, επομένως στο παραμετρικό μοντέλο θα πρέπει να είναι τουλάχιστον δέκατης τάξης. Με τον όρο τουλάχιστον εννοούμε πως, μπορεί να υπάρχουν κι άλλα πράγματα στο σήμα μας τα οποία δεν φαίνονται, όπως για παράδειγμα πραγματικοί πόλοι.

Στην συνέχεια, υπολογίζεται το συχνοτικό βήμα – διακρισιμότητα ως εξής:

$$R_{\text{FFT}} = \frac{2\pi}{T,N} = \frac{2\pi*512}{4000} = 0,8 \text{ Hz}$$

2. Περιοδόγραμμα.

Σύμφωνα με το πρόγραμμα “Matlab”, υπολογίζεται το περιοδόγραμμα του σήματος συναρτήσει της συχνότητας (απο 0 έως $f_s/2$), το οποίο απεικονίζεται παρακάτω.



Εικόνα 7: Περιοδόγραμμα Σήματος συναρτήσει Συχνότητας.

ο Παρατηρήσεις:

Το Περιοδόγραμμα, που απεικονίζεται στο παραπάνω διάγραμμα, σε σχέση με το διάγραμμα Διακριτού Μετασχηματισμού Fourier φαίνεται μετατοπισμένο προς τα κάτω. Αυτό συμβαίνει διότι, το περιοδόγραμμα είναι στην ουσία το τετράγωνο της τιμής του μετασχηματισμού Fourier προς το μήκος των τιμών N .

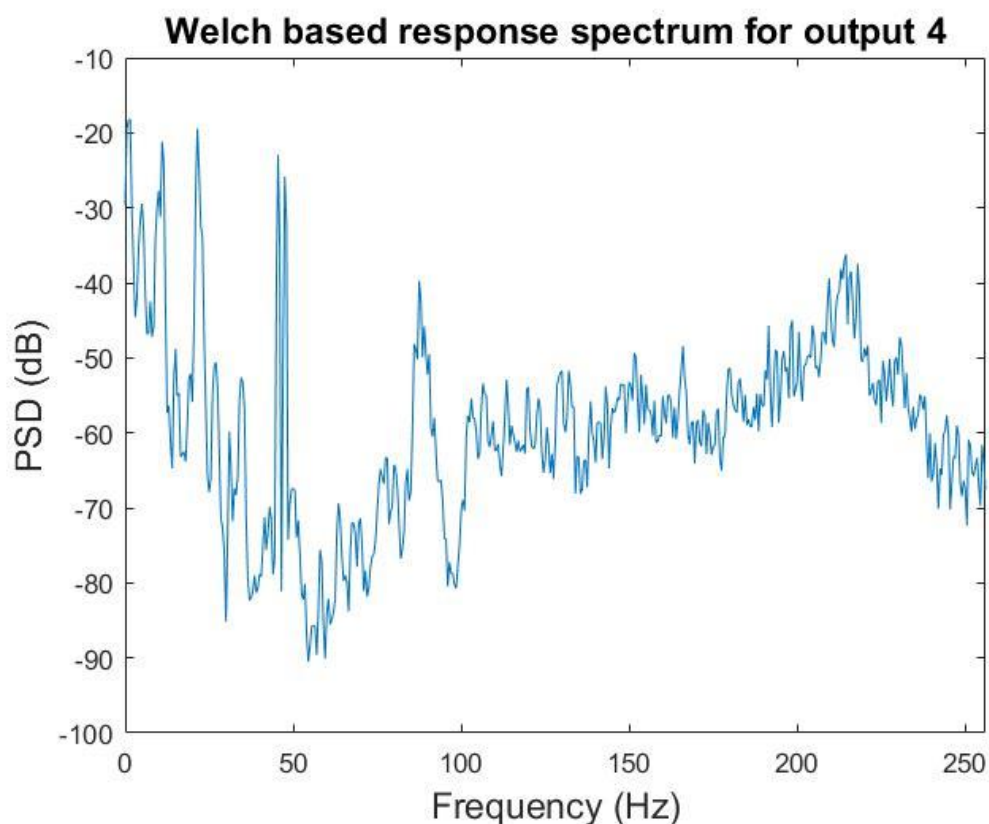
Έτσι, επειδή το Περιοδόγραμμα και το διάγραμμα του Διακριτού Μετασχηματισμού Fourier είναι λογαριθμικής κλίμακας και το εκθετικό έρχεται μπροστά στον λογάριθμο, με αυτόν τον τρόπο, στο γράφημα το μόνο που πρέπει να αλλάζει είναι η κλίμακα.

Τέλος, να προσθέσουμε πως, ο Μετασχηματισμός Fourier της απόκρισης είναι λογαριθμικής κλίμακας λόγω των decibel που είναι λογαριθμικής μορφής.

Σημείωση : Ο καμπύλη του μετασχηματισμού Fourier και του Περιοδογράμματος, δεν συνεχίζει μέχρι το τέλος του διαγράμματος ($f=f_s/2$) αλλά σταματά στην μέση ($f=f_s/4$). Κάποιο πρόβλημα θα υπάρχει με το πρόγραμμα «Matlab» το οποίο δεν μπορώ να εξηγήσω.

3. Εκτίμηση φάσματος μέσω της μεθόδου Welch

Έπειτα, παρουσιάζεται το διάγραμμα 8 στο οποίο απεικονίζεται η εκτίμηση του φάσματος σύμφωνα με την μέθοδο Welch.



Εικόνα 8: Εκτίμηση Φάσματος με μέθοδο Welch.

ο Παρατηρήσεις:

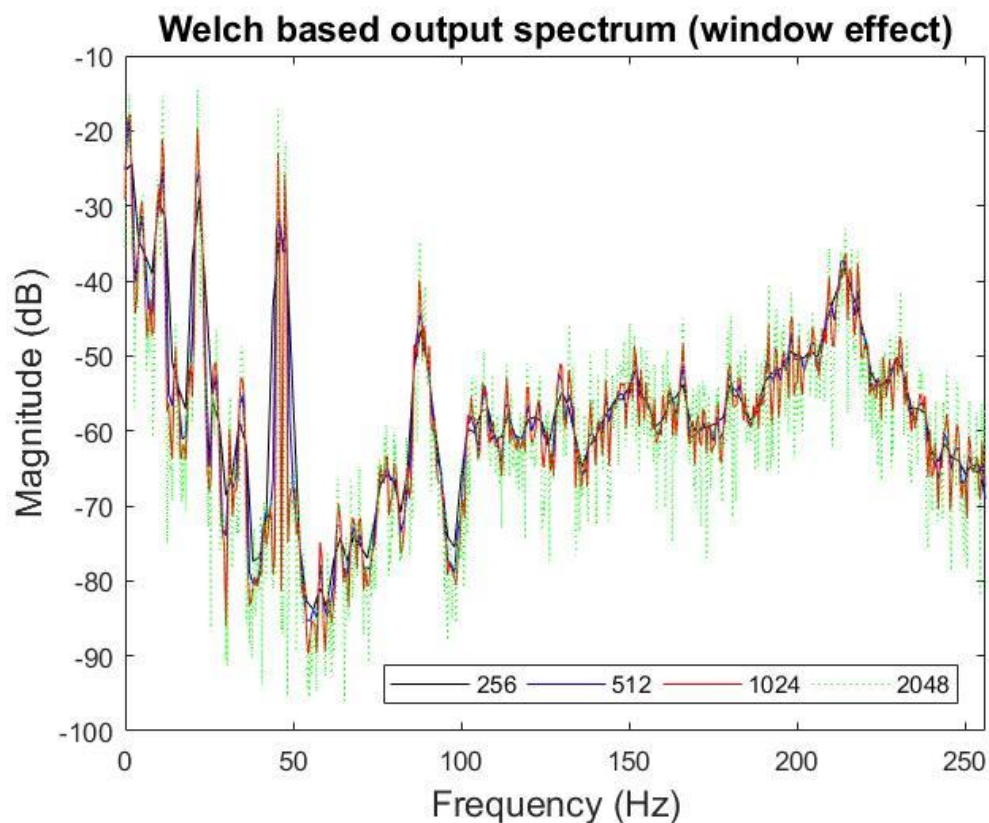
Για την κατασκευή του παραπάνω διαγράμματος, χρησιμοποιήθηκε ως μήκος παραθύρου το $L=1024$ τιμές δεδομένων. Επίσης, χρησιμοποιήθηκε overlap μεταξύ των παραθύρων της τάξεως του 95%.

Επειδή γνωρίζουμε πως, οι θεωρητικές μέγιστες τιμές του overlap κυμαίνονται γύρω στο 50% με 75%, στην προκειμένη επιλέχθηκε το 75% ως μέγιστη δυνατή λύση.

Στην συνέχεια, υπολογίζεται το συχνοτικό βήμα – διακρισιμότητα ως εξής:

$$R_{\text{Welch}} = \frac{2\pi}{T,N} = \frac{2\pi \cdot 512}{1024} = 3,14 \text{ Hz}$$

Παρακάτω, στο Διάγραμμα 9, απεικονίζεται η εκτίμηση του φάσματος μέσω της μεθόδου Welch, για διάφορα μήκη παραθύρου L. Το L παίρνει τις τιμές 256, 512, 1024, 2048.



Εικόνα 9: Εκτίμηση φάσματος μέσω μεθόδου Welch για μήκη παραθύρων 256,512,1024 και 2048.

○ Παρατηρήσεις:

Με βάση το παραπάνω διάγραμμα, φαίνεται πως για μικρά μήκη παραθύρου (L=256 ή L=512), δεν υπάρχει ικανοποιητική προσέγγιση του φάσματος αφού όπως παρατηρείται, συγκριτικά με τα μεγάλα μήκη παραθύρου, δεν απεικονίζονται όλα τα όρη και όλες οι κοιλότητες που απεικονίζονται όσο αυξάνουμε το L. Επίσης, η διακρισιμότητα για μικρό L αυξάνει, παρόλο που συνάρτηση γίνεται ομαλή, δεν το επιδιώκουμε για

την εκτίμηση του φάσματος. Έτσι, είναι λογικό να προβούμε σε μεγάλα μήκη παραθύρου.

Συγκρίνοντας τις καμπύλες για $L=1024$ και για $L=2048$ παρατηρείται έντονη στοχαστικότητα της εκτίμησης του φάσματος, όταν $L=2048$, λόγω της μεγάλης ταλαντωτικής συμπεριφοράς. Οπότε, παρότι για $L=2048$ έχουμε μικρότερη διακρισιμότητα από ότι για $L=1024$, δεν πρέπει να επιλέξουμε μήκος παραθύρου 2048 διότι υπάρχει μεγάλη στοχαστικότητα στην εκτίμηση του σφάλματος. Τελικά, για να έχουμε μικρή διακρισιμότητα που μας βοηθήσει στην εκτίμηση, αλλά και για να μην υπάρχει μεγάλη στοχαστικότητα, επιλέγεται ως κατάλληλο μήκος παραθύρου το $L=1024$.

Στην συνέχεια, υπολογίζεται το συχνοτικό βήμα – διακρισιμότητα για μήκη παραθύρου 256, 512, 1024 και 2048.

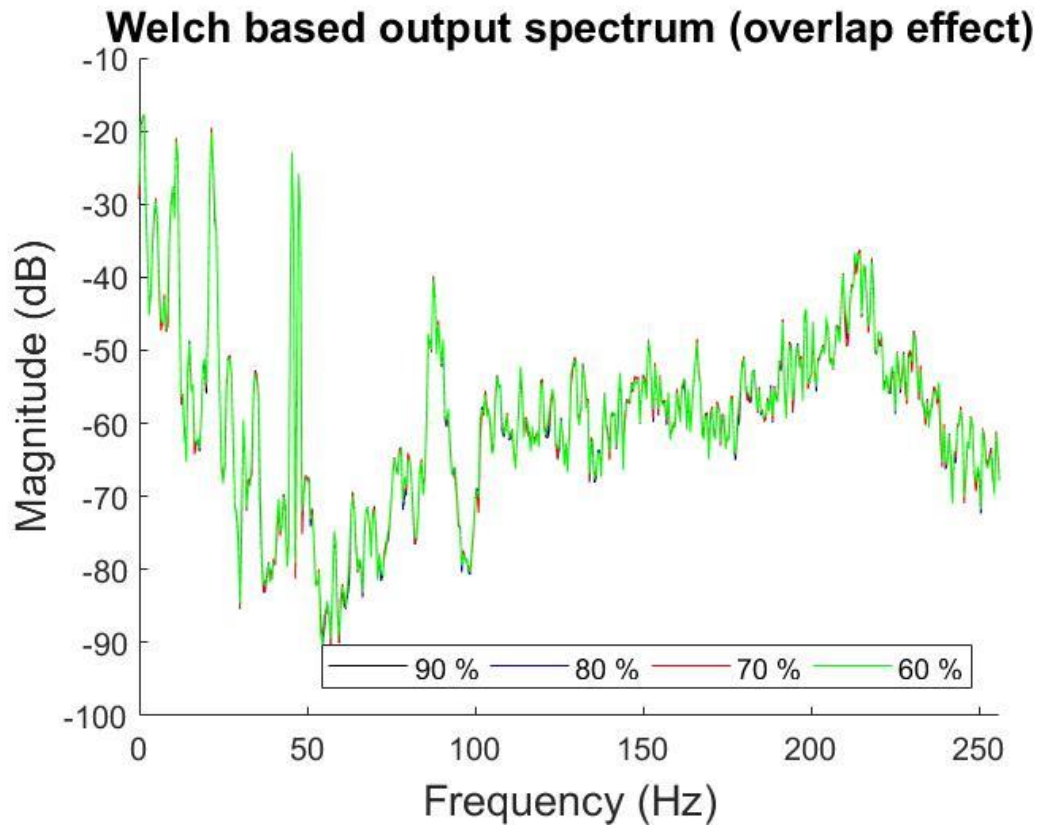
$$R_{\text{Welch}=256} = \frac{2\pi}{T,L} = \frac{2\pi \cdot 512}{256} = \mathbf{12.57 \text{ Hz}}$$

$$R_{\text{Welch}=512} = \frac{2\pi}{T,L} = \frac{2\pi \cdot 512}{512} = \mathbf{6.28 \text{ Hz}}$$

$$R_{\text{Welch}=1024} = \frac{2\pi}{T,L} = \frac{2\pi \cdot 512}{1024} = \mathbf{3.14 \text{ Hz}}$$

$$R_{\text{Welch}=2048} = \frac{2\pi}{T,L} = \frac{2\pi \cdot 512}{2048} = \mathbf{1.57 \text{ Hz}}$$

Παρακάτω παρατίθεται η εκτίμηση του φάσματος μέσω της μεθόδου Welch, για τις διάφορες τιμές του overlap. Το overlap είναι της τάξεως του 60%, 70%, 80% και 90%.



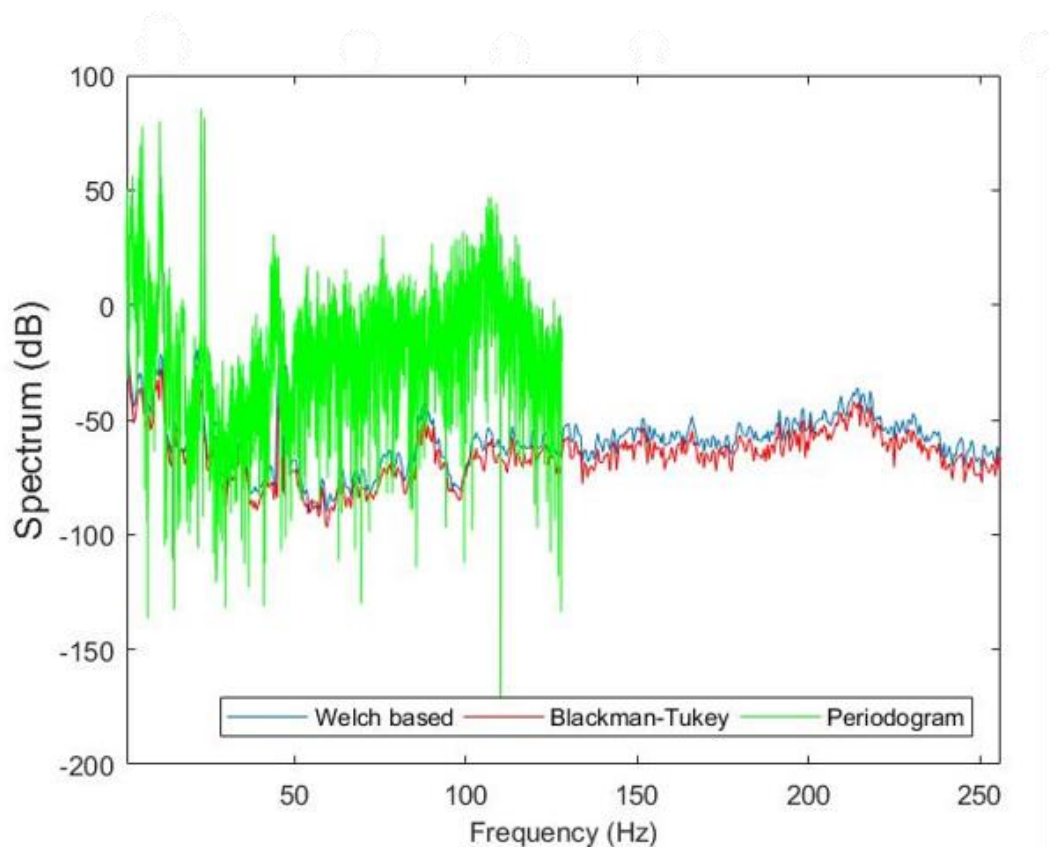
Εικόνα 10: Εκτίμηση φάσματος μέσω Welch για overlap 60%, 70%, 80%, 90%.

○ **Παρατηρήσεις:**

Όπως φαίνεται παραπάνω, η εκτίμηση του φάσματος επηρεάζει ελάχιστα έως και καθόλου το overlap. Αυτό συμβαίνει διότι, όταν τα segments γίνονται "πάρα πολύ" overlap συμπίπτει πολύ το ένα στο άλλο. Ουσιαστικά, είναι σαν να επαναλαμβάνει τον εαυτό του και δεν έχει καμία καινούργια πληροφορία. Επομένως, η αύξηση του overlap συνεισφέρει σε πολύ μικρό βαθμό, στην καλύτερη εκτίμηση του φάσματος.

4. Σύγκριση των μεθόδων Welch – Blackman-Tukey – Περιοδογράμματος ως προς την εκτίμηση του φάσματος.

Στο Διάγραμμα 11, απεικονίζεται η εκτίμηση του φάσματος σύμφωνα με την μέθοδο Welch, Blackman-Tukey, και του Περιοδογράμματος.



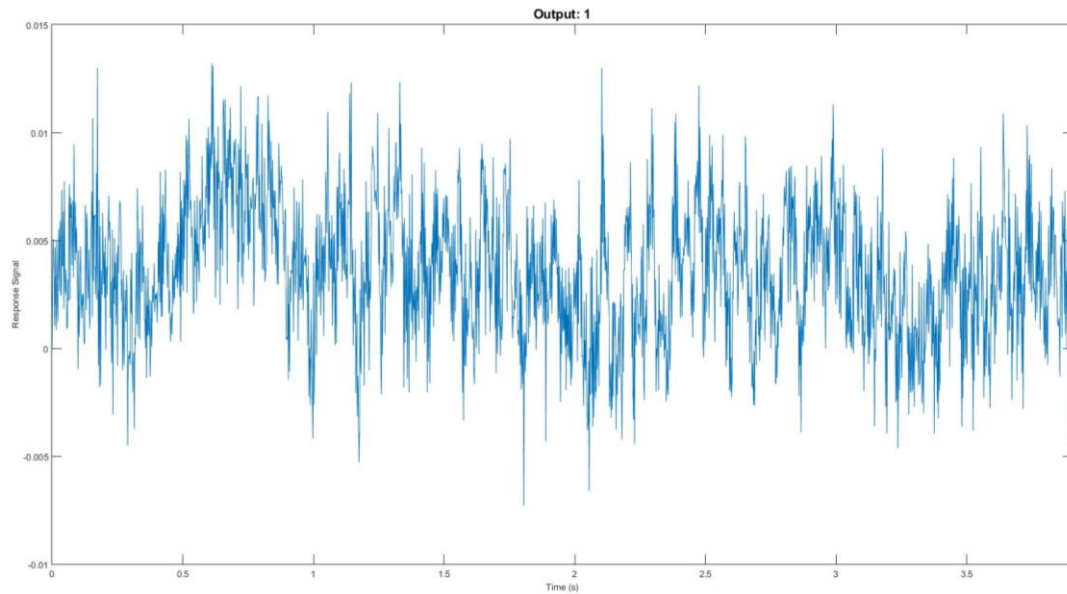
Εικόνα 11: Εκτίμηση φάσματος με μέθοδο Welch, Blackman-Tukey και Περιοδογράμματος.

ο Παρατηρήσεις:

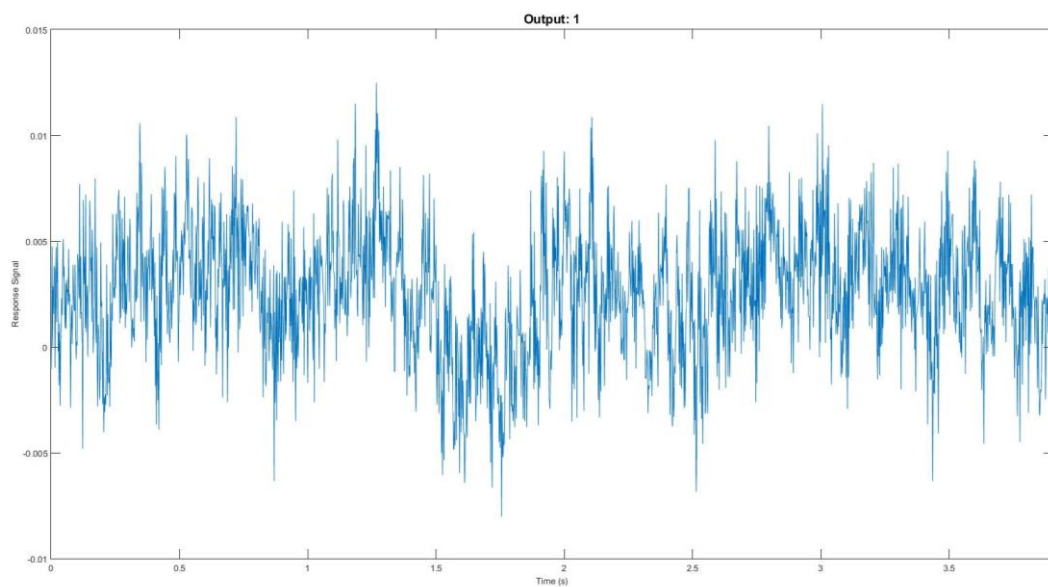
Συγκρίνοντας τις μεθόδους Welch και Blackman-Tukey, παρατηρείται μεγάλη ομοιότητα μεταξύ των καμπυλών. Συγκρίνοντας την εκτίμηση του Περιοδογράμματος με τις άλλες δύο εκτιμήσεις, φαίνεται πως υπάρχει μεγάλη στοχαστικότητα συγκριτικά με τις εκτιμήσεις κατά Welch και Blackman-Tukey αντίστοιχα. Βέβαια, και οι τρεις εκτιμήσεις, μας δίνουν μια πρώτη εντύπωση για την ελάχιστη τάξη του μοντέλου ARMA(n,n). Με λίγα λόγια, και οι τρεις μέθοδοι εκτιμούν το φάσμα να έχει την ίδια μορφή.

B3: Έλεγχος στασιμότητας.

Τα διαγράμματα που ακολουθούν απεικονίζουν το σήμα μας σε δύο ίσα μέρη. Το πρώτο μέρος έχει 1 έως 2000 samples και το δεύτερο έχει 2001 έως 4000 samples.



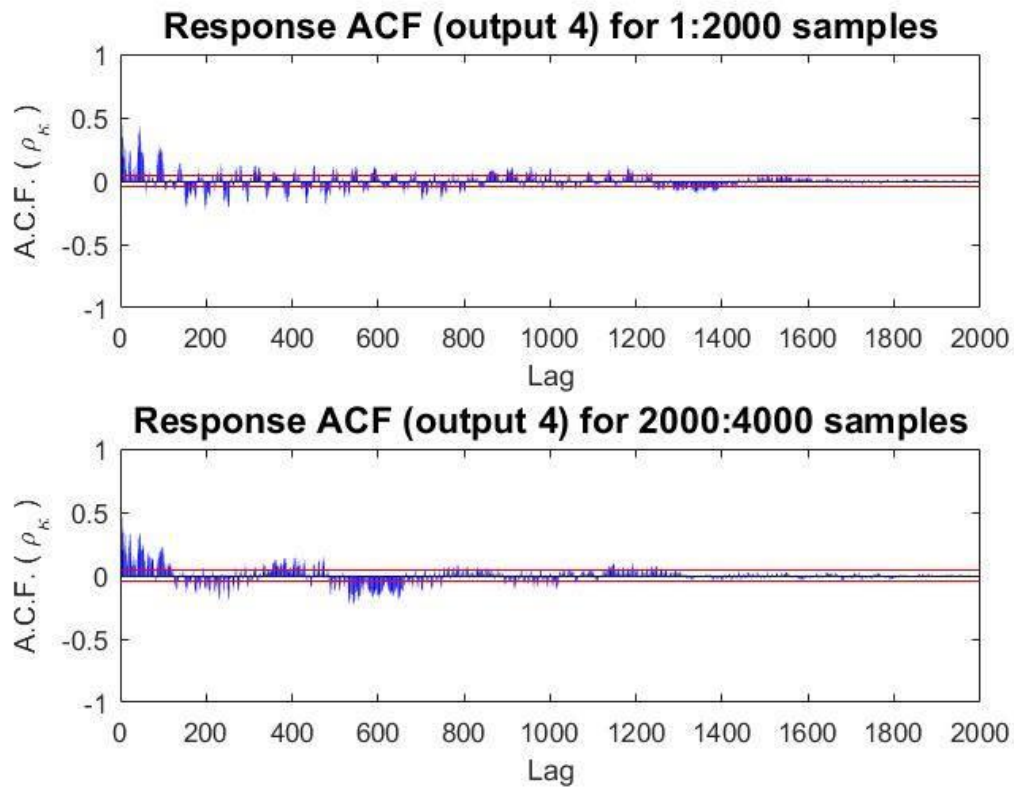
Εικόνα 12: 1-2000 samples.



Εικόνα 13: 2001-4000 samples.

Στην συνέχεια, θα υπάρξει εκτίμηση της ανοιγμένης αυτοσυμμεταβλητότητας για κάθε τμήμα του σήματος και θα

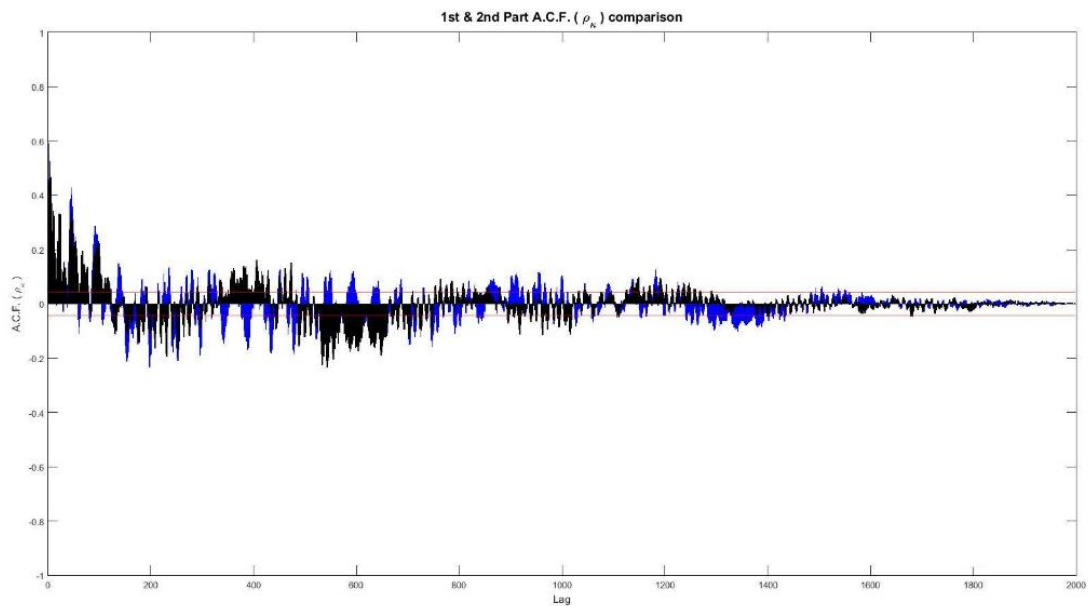
παρουσιαστεί ένα γράφημα στο οποίο υπολογίζεται η συνάρτηση της ανοιγμένη αυτοσυμμεταβλητότητα συναρτήσει του lag για καθένα από τα δύο μέρη του σήματος.



Εικόνα 14: Συνάρτηση Ανοιγμένης Αυτοσυμμεταβλητότητας συναρτήσει του lag.

Στην συνέχεια, θα υπάρξει σύγκριση των τιμών του κάθε τμήματος της ανοιγμένης αυτοσυμμεταβλητότητας σε ένα διάγραμμα, όπως φαίνεται παρακάτω.

○ Παρατηρήσεις:

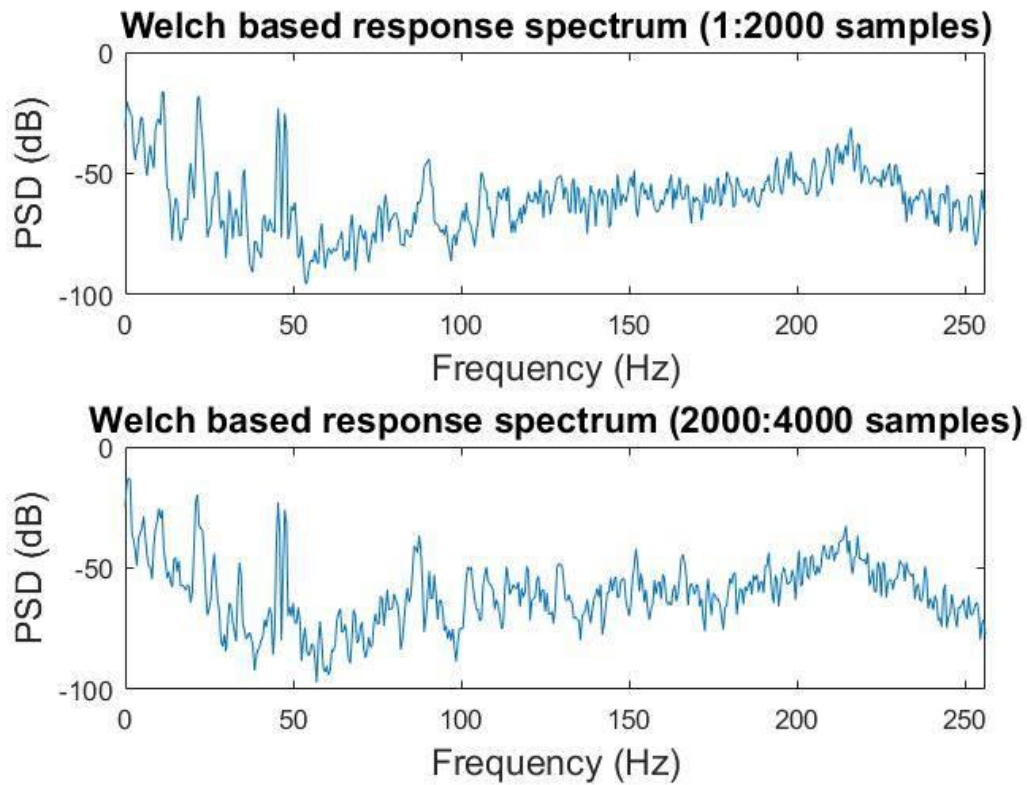


Εικόνα 15: Σύγκριση ACF των δύο τμημάτων.

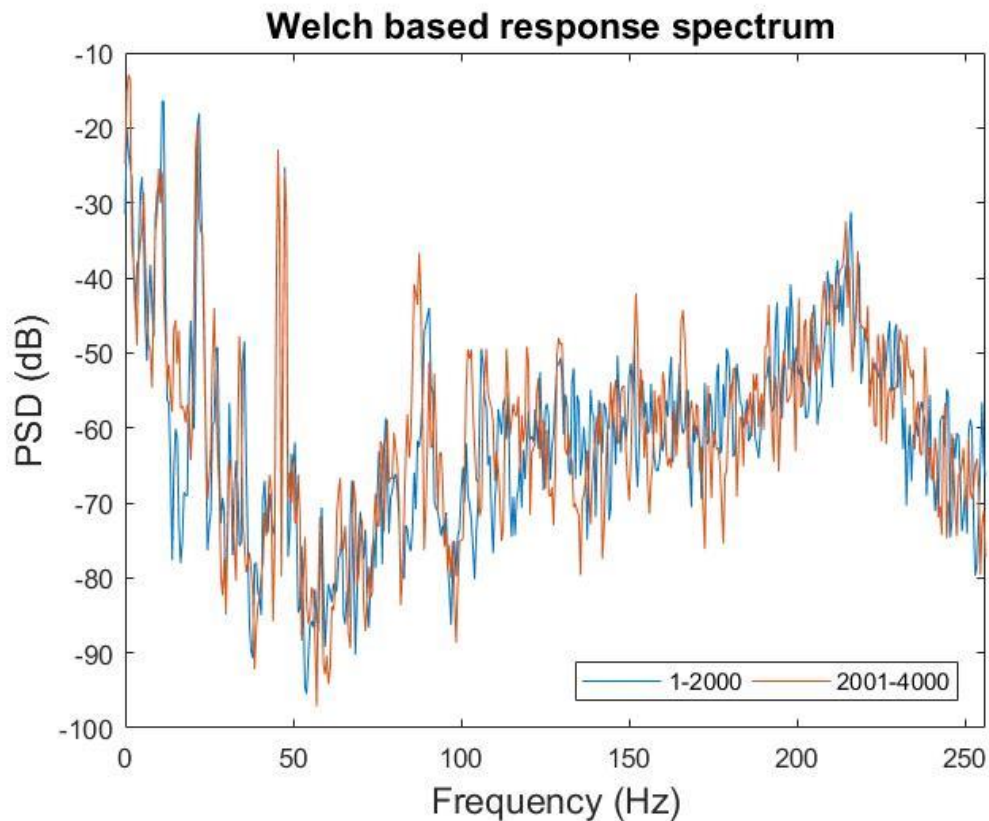
Σύμφωνα με την εικόνα 15, στο μαύρο χρώμα είναι οι τιμές των ACF του 2^{ου} μέρους του σήματος, ενώ στο μπλέ χρώμα φαίνονται οι τιμές ACF του 1^{ου} μέρους του σήματος για τις οποίες δεν συμπίπτουν με το 2^ο μέρος. Έτσι, αν και υπάρχει μια ημιτονειδής μορφή των ACF κάθε τμήματος, δείχνει να έχουν διαφορά φάσης. Αυτό σημαίνει πως, το σήμα μας φαίνεται να έχει μια μορφή μη στασιμότητας. Όμως, ακόμα και έτσι, δεν δεχόμαστε την μη στασιμότητα του σήματος και θεωρούμε πως το σήμα μας είναι στάσιμο με χαρακτηριστικά μη στασιμότητας.

Σημειώνεται όμως πως στο Διάγραμμα 5, η συσχέτιση των τιμών σταματά για lag = 3200 περίπου. Εάν στο Διάγραμμα 15 το lag κυμαινόταν από 0-3500, θα έπρεπε να έχουμε αδυναμία της συσχέτισης των τιμών για lag > 3200. Οπότε η συσχέτιση κανονικά δεν χάνεται για lag=1400 παρότι οι τιμές βρίσκονται κάτω από τα όρια της λευκότητας, αλλά θα χάνονται για lag 3200 περίπου.

Στην συνέχεια, υπολογίζεται η εκτίμηση του φάσματος με την μέθοδο Welch, και για τα δύο μέρη που χωρίστηκε το σήμα. Στο Διάγραμμα 16 φαίνεται η εκτίμηση του φάσματος για το κάθε ένα μέρος, ενώ στο Διάγραμμα 17, γίνεται η συγκριση των δύο αυτών εκτιμήσεων



Εικόνα 16: Εκτίμηση Φάσματος Welch για τα δύο μέρη.



Εικόνα 17: Σύγκριση των δύο εκτιμήσεων του φάσματος για τα δύο μέρη.

○ **Παρατηρήσεις:**

Παρατηρείται πως οι δύο εκτιμήσεις του φάσματος για τα δύο μέρη του σήματος είναι εξαιρετικά κοντά η μία με την άλλη. Όμως, η εκτίμηση του φάσματος στο 2^ο μέρος ακολουθεί μια πιο στοχαστική μορφή συγκριτικά με το 1^ο μέρος. Παρόλα αυτά, μπορεί να θεωρηθούν ίδιες. Έτσι, μπορούμε να υποθέσουμε πως υπάρχει στασιμότητα του συγκεκριμένου σήματος, ως προς το φάσμα.

Τμήμα Γ – Παραμετρική ανάλυση ARMA.

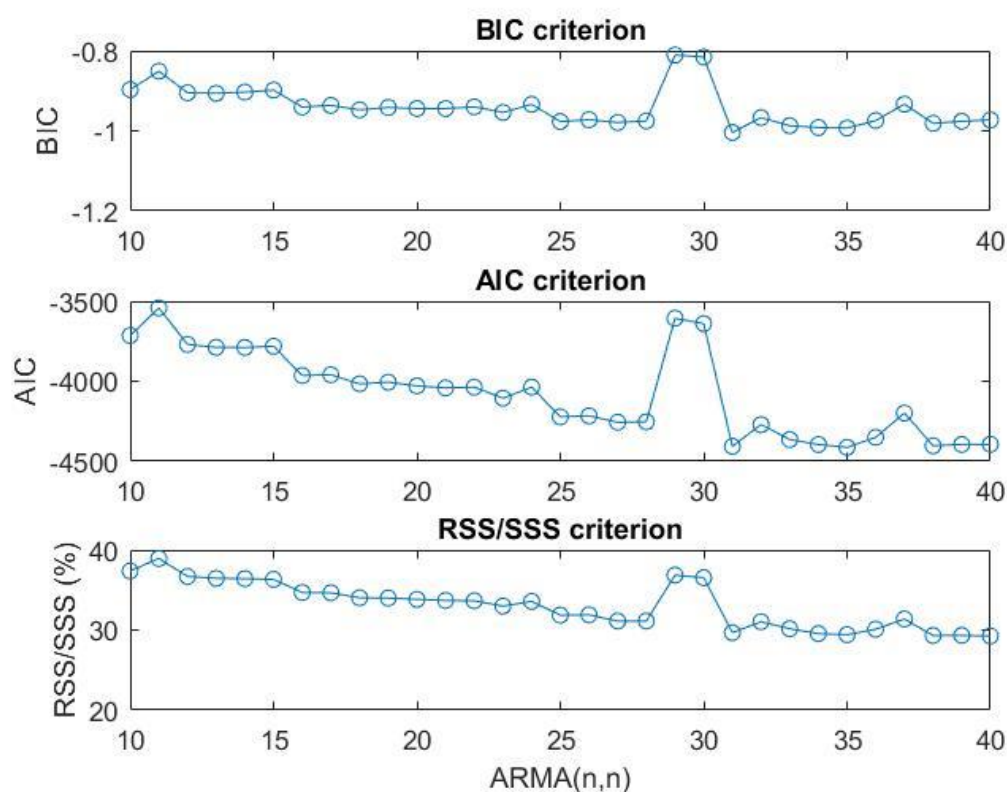
Γ1: Προκαταρκτικά.

Το σήμα χωρίζεται σε δύο τμήματα. Το πρώτο τμήμα ονομάζεται τμήμα εκτίμησης και έχει μήκος N_e τιμές, ενώ το δεύτερο τμήμα ονομάζεται τμήμα ελέγχου εγκυρότητας, με μήκος N_v τιμών. Δεδομένου ότι $N_e \gg N_v$, λαμβάνεται αυθαίρετα ότι $N_e = 3300$ δείγματα και $N_v = 700$ δείγματα.

Ταυτόχρονα, είναι λογικό να ισχύει ότι : $N_e + N_v = 3300 + 700 = 4000$ δείγματα ($=N$).

Γ2: Εκτίμηση μοντέλων ARMA.

Σύμφωνα με το πρόγραμμα "Matlab", υπολογίζονται τα κριτήρια BIC, AIC και RSS/SSS τα οποία απεικονίζονται στο Διάγραμμα.



Εικόνα 18: Κριτήρια BIC, AIC και RSS/SSS.

○ Παρατηρήσεις:

- ✓ BIC: Σύμφωνα με την καμπύλη του κριτηρίου BIC, παρατηρείται πως, υπάρχει ελάχιστο για $n=31$. Αυτό, μας οδηγεί στο συμπέρασμα πως, το μοντέλο που θα επιλέξουμε θα είναι τάξης ARMA(31,31).
- ✓ AIC: Με αντίστοιχο τρόπο, παρατηρείται πως η καμπύλη του κριτηρίου AIC έχει και αυτή με την σειρά της ελάχιστο στο $n=31$. Επομένως, το μοντέλο που επιλέγεται είναι το ARMA(31,31). Παρότι το κριτήριο AIC τείνει να υπερεκτιμήσει την τάξη του μοντέλου AR, στο συγκεκριμένο παράδειγμα παρατηρείται πως τα μοντέλα BIC και AIC προβαίνουν σε ίδιες εκτιμήσεις.
- ✓ RSS/SSS: Παράλληλα, από την καμπύλη RSS/SSS παρατηρούνται τοπικά μέγιστα για $n=11,14,30,36,40$. Αυτό προκύπτει λόγω ύπαρξης μηδενιστών έξω από τον μοναδιαίο κύκλο για κάθε ένα από τα μοντέλα ARMA(n,n), όπου το n λαμβάνει τις τιμές 11,14,30,36,40. Όμως αφού επιλέχθηκε το μοντέλο ARMA(31,31), δεν χρειάζεται με την μέθοδο αντιστρεψιμότητας να τοποθετήσουμε τους συγκεκριμένους μηδενιστές εντός του μοναδιαίου κύκλου, επειδή δεν θα επιλεγθεί κανένα από τα μοντέλα ARMA(n,n) με $n=11,14,30,36,40$.

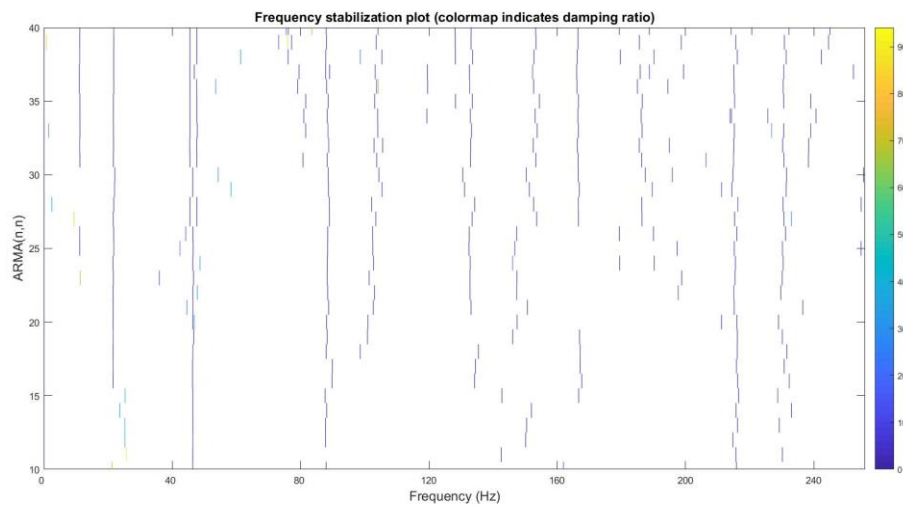
Τέλος, για το εκτιμηθέν μοντέλο παρατηρείται, πως στο διάγραμμα RSS/SSS, η στοχαστικότητά του είναι της τάξεως του 30%.

Επιπροσθέτως, όσον αφορά το SPP έχουμε :

$$SPP = \frac{Ne}{(n+n+1)} = \frac{3300}{31+31+1} =$$
$$= \mathbf{SPP=52.4}$$

Επομένως, αφού συνίσταται $SPP > 15$ τότε, η παραπάνω σχέση ισχύει.

Στην συνέχεια, παρατίθεται το διάγραμμα απεικόνισης σταθεροποίησης συχνοτήτων.

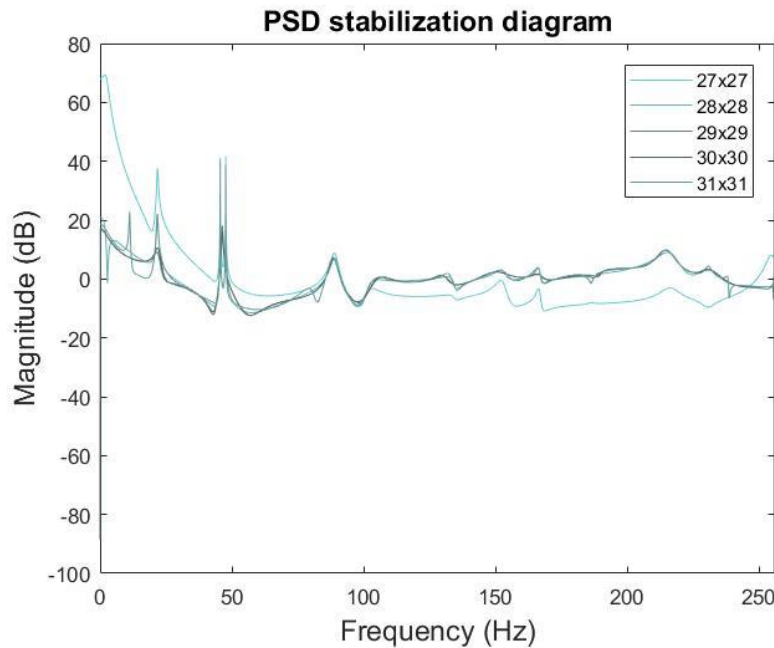


Εικόνα 19: Γράφημα Σταθεροποίησης Συχνοτήτων.

○ **Παρατηρήσεις:**

Όπως φαίνεται παραπάνω, για τιμές $n=31$ παρατηρείται πως υπάρχει σύγκλιση των πραγματικών συχνοτήτων. Εξαιτίας του σκούρου χρώματος σύμφωνα με την βαθμονόμηση, είναι κατανοητό πως έχουμε μεγάλες τιμές του λόγου απόσβεσης ζ , για τις συγκεκριμένες συχνότητες, και άρα δεν έχουμε ταλαντωτική συμπεριφορά του υπό ανάλυση, σήματος.

Έπειτα, απεικονίζεται η σταθεροποίηση του φάσματος.



Εικόνα 20: Γράφημα Σταθεροποίησης Φάσματος.

○ Παρατηρήσεις:

Όπως προκύπτει από το γράφημα της σταθεροποίησης του φάσματος, αυξάνοντας την τάξη του ARMA(n,n) μοντέλου, μέχρι το n να ισούται με 31, δημιουργούνται κορυφές στο εύρος συχνότητας 0-50 Hz, οι οποίες δεν δημιουργούνται για χαμηλή τάξη του μοντέλου. Αυτό υποδεικνύει πως τα μοντέλα ARMA(n,n), με $n < 31$, δεν είναι κατάλληλα να προσεγγίσουν το συγκεκριμένο σήμα, δηλαδή υπάρχει σύγκλιση.

Επίσης, αυξάνοντας την τάξη του μοντέλου, δημιουργούνται και επιπλέον κοιλάδες οι οποίες προέρχονται από την ύπαρξη ζεύγους μηδενιστών κοντά ή έξω από τον μοναδιαίο κύκλο. Εάν κάποιο ζεύγος μηδενιστών βρίσκεται έξω από τον κύκλο, τότε πρέπει να εισχωρήσει μέσα σε αυτόν σύμφωνα με την μέθοδο της αντιστρεψιμότητας.

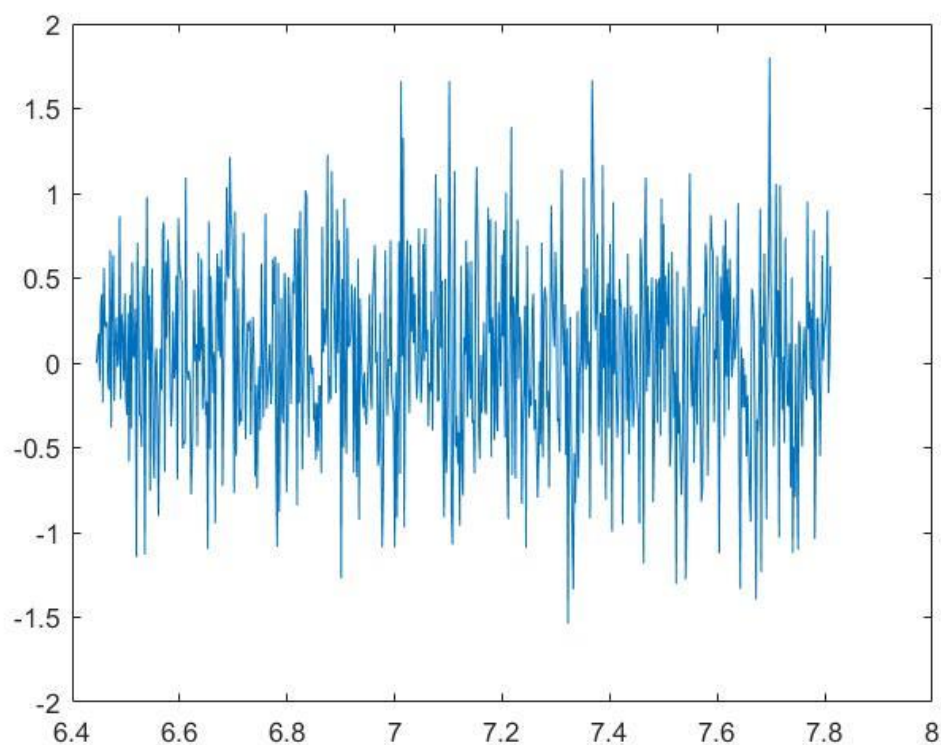
Τελικά, με βάση τις παραπάνω παρατηρήσεις που έγιναν, θεωρείται πως το μοντέλο ARMA(31,31) είναι το καταλληλότερο ώστε να προσεγγίσει όσο το δυνατόν αποτελεσματικότερα το σήμα μας.

Γ3: Έλεγχος εγκυρότητας του επιλεγέντος μοντέλου ARMA.

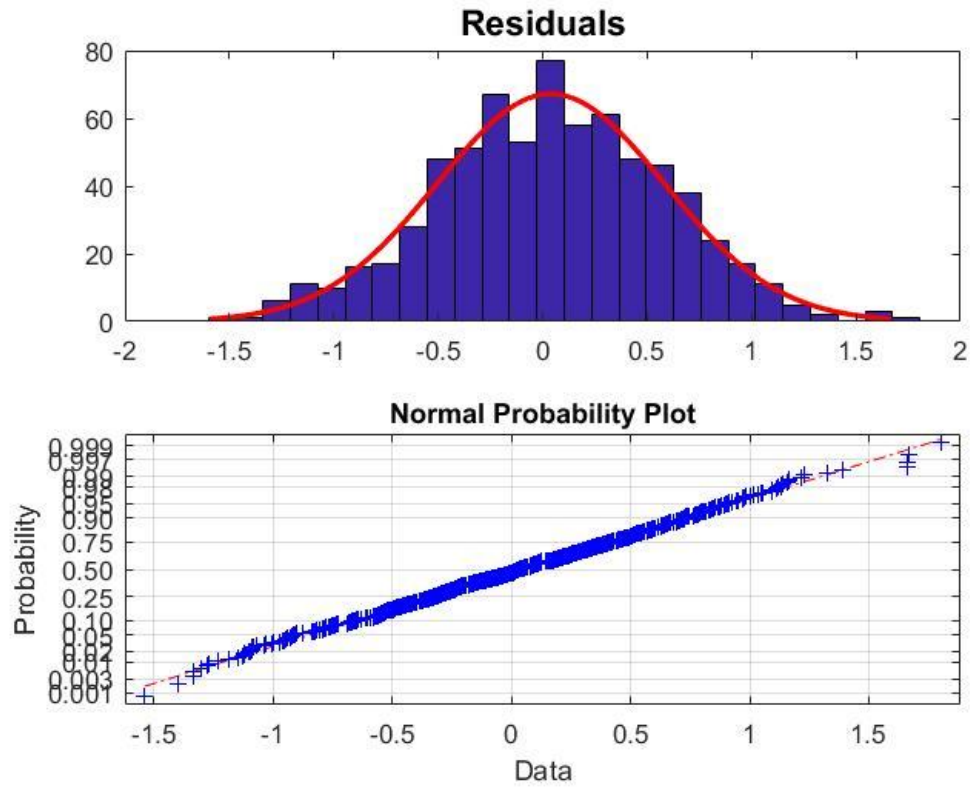
1. Τμήμα ελέγχου εγκυρότητας.

Αρχικά, για την παρουσίαση των διαγραμμάτων που ζητούνται στο συγκεκριμένο ερώτημα, απαραίτητη προϋπόθεση είναι ο υπολογισμός των υπολοίπων e_t που προκύπτουν από το τμήμα ελέγχου εγκυρότητας. Στην συνέχεια, για τον έλεγχο της εγκυρότητας του μοντέλου, ελέγχουμε την κανονικότητα και την λευκότητα των υπολοίπων e_t .

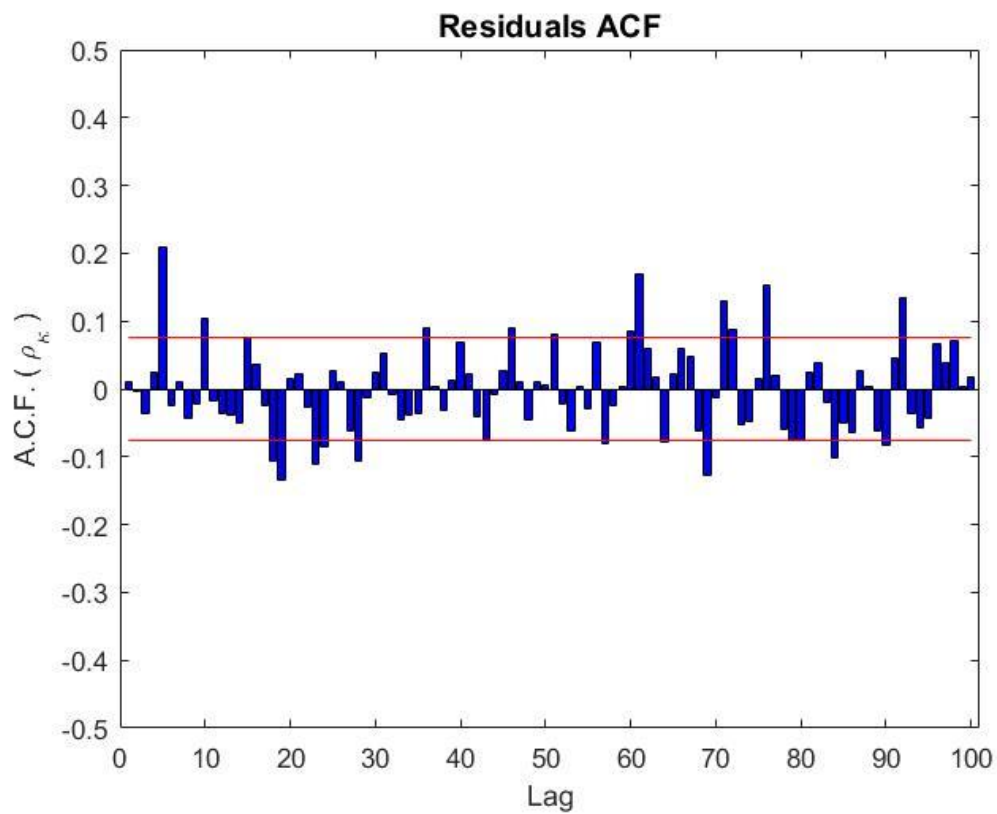
Παρακάτω, παρουσιάζονται στα διαγράμματα 21, 22, 23 τα υπόλοιπα συναρτήσει του χρόνου, η συνάρτηση πυκνότητας πιθανότητας των υπολοίπων και το αντίστοιχο Normal probability plot και η ανηγμένη αυτοσυμμεταβλητότητα ρ_k των υπολοίπων συναρτήσει του lag.



Εικόνα 21: Τα υπόλοιπα του παραμετρικού μοντέλου συναρτήσει του χρόνου.



Εικόνα 22: Συνάρτηση Πυκνότητας Πιθανότητας των υπολοίπων και το Normal Probability Plot.



Εικόνα 23: Η ανηγμένη αυτοσυμμεταβλητότητα ρ_k των υπολοίπων.

○ Παρατηρήσεις:

✓ Διάγραμμα υπολοίπων συναρτήσει του χρόνου.

Παρατηρείται πως τα υπόλοιπα έχουν σταθερή μέση τιμή ίση με 0. Έτσι, το διάγραμμα θεωρείται στάσιμο.

✓ Συνάρτηση πυκνότητας πιθανότητας υπολοίπων και Normal probability plot.

Βλέπουμε ότι, τα υπόλοιπα ακολουθούν με μεγάλη ακρίβεια την κανονική κατανομή. Αν και υπάρχουν αποκλίσεις στην αρχή και το τέλος, όσον αφορά το Normal probability plot θεωρούνται πολύ μικρές και αναμενόμενες με αποτέλεσμα να μην απορρίπτεται η κανονικότητα των υπολοίπων e_t .

✓ Ανηγμένη αυτοσυμμεταβλητότητα ρ_k των υπολοίπων.

Τέλος, παρατηρείται πως γύρω στις 12 τιμές βρίσκονται έξω από τα όρια της λευκότητας. Με αυτόν τον τρόπο, εύκολα συμπεραίνουμε πως, το μοντέλο μας δεν είναι το κατάλληλο για επίπεδο σφάλματος 5%.

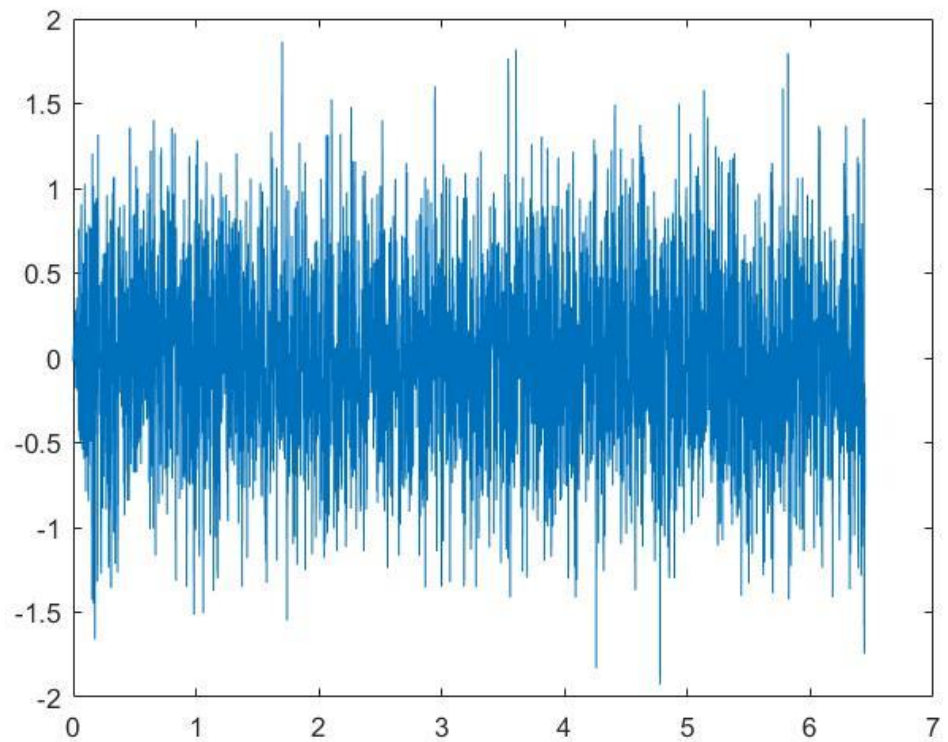
Παρόλα αυτά, λαμβάνοντας υπόψη την στοχαστικότητα η οποία είναι της τάξεως του 30%, θεωρείται λογική η ανακριβής εκτίμηση της ρ_k με οποιοδήποτε μοντέλο ARMA(n,n), λαμβάνοντας έτσι τιμές που απέχουν από την πραγματικότητα.

Έτσι, εμπιστευόμενοι κυρίως το κριτήριο BIC, μπορεί να θεωρηθεί πως, το μοντέλο ARMA(31,31) είναι το καταλληλότερο μοντέλο στο συγκεκριμένο πρόβλημα, παρότι της παραβίασης του φράγματος του 5% σφάλματος.

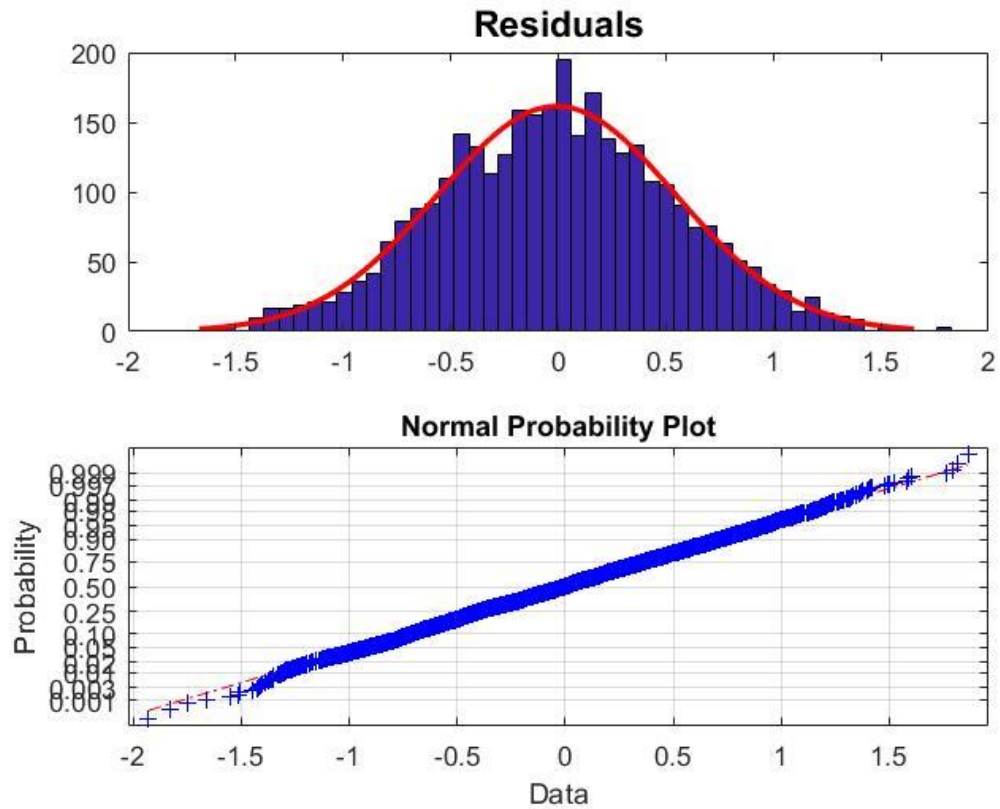
2. Τμήμα εκτίμησης.

Με τον ίδιο τρόπο με παραπάνω, υπολογίζοντας τα υπόλοιπα e_t που προκύπτουν από το τμήμα εκτίμησης και ελέγχοντας την κανονικότητα και την λευκότητα των υπολοίπων e_t για τον έλεγχο

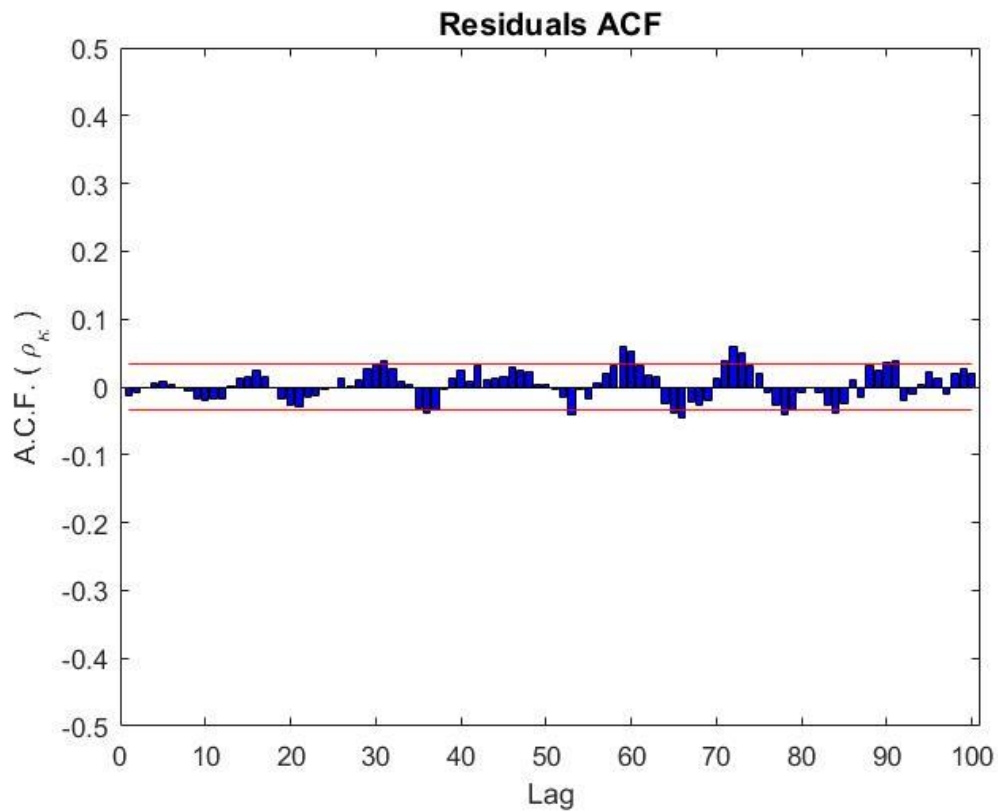
την εγκυρότητα του μοντέλου παρατίθενται τα αντίστοιχα διαγράμματα με παραπάνω.



Εικόνα 24: Υπόλοιπα παραμετρικού μοντέλου συναρτήσει του χρόνου.



Εικόνα 25: Συνάρτηση Πυκνότητας Πιθανότητας υπολοίπων και το Normal Probability Plot.



Εικόνα 26: Ανηγγμένη Αυτοσυμμεταβλητότητα ρ_k των υπολοίπων.

○ Παρατηρήσεις:

✓ Διάγραμμα υπολοίπων συναρτήσει του χρόνου.

Παρατηρείται πως τα υπόλοιπα έχουν σταθερή μέση τιμή ίση με 0. Έτσι, το διάγραμμα θεωρείται στάσιμο.

✓ Συνάρτηση πυκνότητας πιθανότητας υπολοίπων και Normal probability plot.

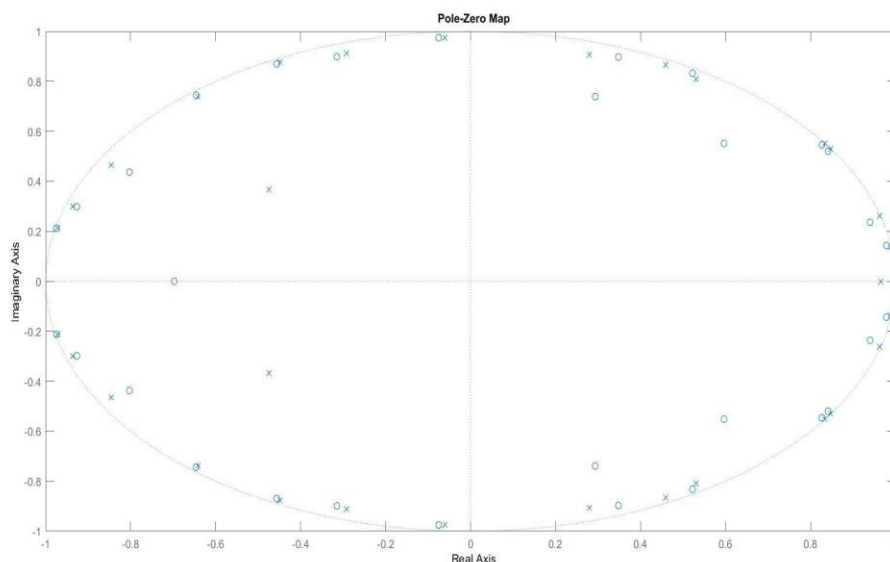
Βλέπουμε ότι, τα υπόλοιπα ακολουθούν με μεγάλη ακρίβεια την κανονική κατανομή. Αν και, υπάρχουν αποκλίσεις στην αρχή και το τέλος, όσον αφορά το Normal probability plot, θεωρούνται πολύ μικρές και αναμενόμενες με αποτέλεσμα να μην απορρίπτεται η κανονικότητα των υπολοίπων e_t .

✓ Ανηγμένη αυτοσυμμεταβλητότητα ρ_k των υπολοίπων.

Τέλος, παρατηρείται πως γύρω στις 5 τιμές βρίσκονται έξω από τα όρια της λευκότητας. Αυτό σημαίνει πως, το μοντέλο μας είναι κατάλληλο για επίπεδο σφάλματος 5%.

Γ4: Ανάλυση του επιλεγέντος μοντέλου ARMA.

Στα παρακάτω διάγραμμα παριστάνονται οι θέσεις των πόλων και των μηδενιστών σε σχέση με τον μοναδιαίο κύκλο.

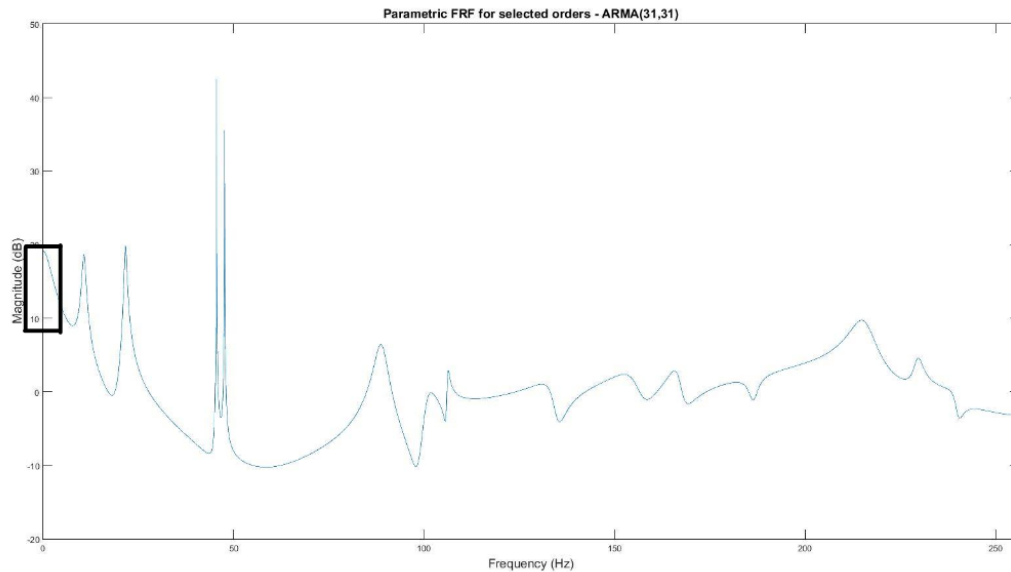


Εικόνα 27: Απεικόνιση Πόλων και Μηδενιστών στον μοναδιαίο κύκλο.

○ Παρατηρήσεις:

Διαπιστώνεται πως, αρκετοί πόλοι και μηδενιστές βρίσκονται πολύ κοντά στον μοναδιαίο κύκλο, επομένως υπάρχει ευστάθεια στο σύστημά μας. Επιπροσθέτως, οι απότομες κορυφές με τις απότομες κοιλάδες είναι πολύ κοντά η μία στην άλλη στο διάγραμμα εκτίμησης φάσματος. Αυτό σημαίνει ότι, ζευγάρια πόλων βρίσκονται πολύ κοντά σε ζευγάρια μηδενιστών. Το γεγονός αυτό επιβεβαιώνεται, στο παραπάνω διάγραμμα (εικόνα 27), αφού αρκετά είναι τα ζευγάρια πόλων και μηδενιστών που είναι αρκετά κοντά μεταξύ τους. Τέτοιου είδους εξέλιξη ήταν αναμενόμενη, καθώς σε τέτοια ταλαντωτικά συστήματα είναι συνηθισμένη αυτή η συμπεριφορά πόλων και μηδενιστών.

Επιπλέον, άξιο σχολιασμού είναι ότι, ένας πραγματικός πόλος βρίσκεται κοντά στο συν 1. Αυτό επιβεβαιώνεται στο διάγραμμα εκτίμησης φάσματος με μία κορυφή πολύ κοντά στο μηδέν, όπως φαίνεται παρακάτω.



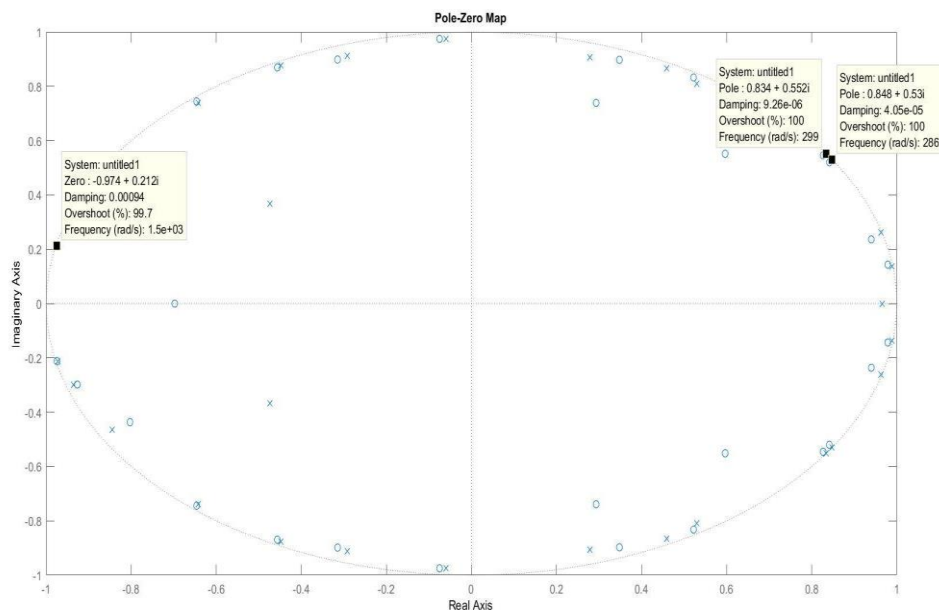
Εικόνα 28: Απεικόνιση πραγματικού πόλου στο διάγραμμα εκτίμησης φάσματος.

Οπότε, αν σκεφτούμε πως, το μοντέλο random walk έχει πόλο κοντά στο συν 1 και είναι μη στάσιμο, τότε θα τολμούσαμε να πούμε πως το σύστημά μας τείνει σε μία μορφή μη στασιμότητας. Σε περίπτωση μη στασιμότητας, αυτό, δεν θα ήταν λογικό και μπορεί να οφειλόταν σε προεξεργασία που έχει γίνει στο σήμα, στο μετρητικό κλπ.

Όμως, στο συγκεκριμένο πρόβλημα και σύμφωνα με τις παρατηρήσεις που έγιναν στο ερώτημα B3 για την εικόνα 13, ότι το σήμα μας είναι στάσιμο με χαρακτηριστικά μη στασιμότητας αυτό επιβεβαιώνεται στην εικόνα 26 και 27. Επομένως, κατά πάσα πιθανότητα ο πόλος στο συν 1 δεν οφείλεται στα σφάλματα που προαναφέρθηκαν παραπάνω στην παράγραφο, αλλά τείνει να είναι μέρος στην φυσική του προβλήματος.

Πάντως, σε καμία περίπτωση, το μοντέλο μας δεν θα το χαρακτηρίζαμε μοντέλο ARIMA, δηλαδή μοντέλα ARMA με χαρακτηριστικά random walk.

Στην συνέχεια, θα παρουσιαστούν ενδεικτικά κάποιους από τους μηδενιστές που δημιουργούν στην εκτίμηση τους φάσματος απότομες κοιλάδες αλλά και κάποιους από τους πόλους που δημιουργούν απότομες κορυφές.



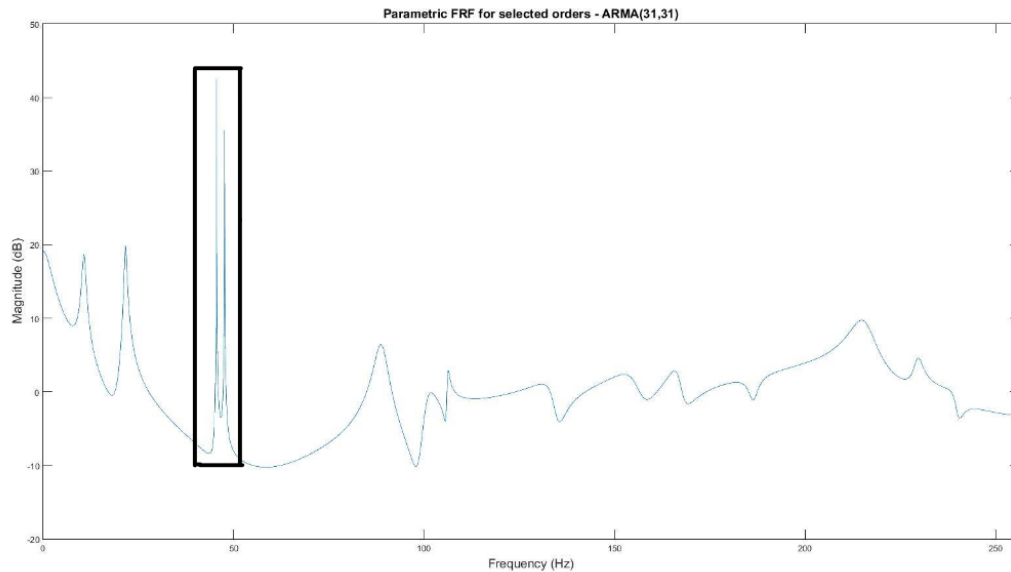
Εικόνα 29: Απεικόνιση μηδενιστών που δημιουργούν απότομες κοιλάδες και πόλων που δημιουργούν απότομες κορυφές στο διάγραμμα εκτίμησης του φάσματος του παραμετρικού μοντέλου.

Πιο συγκεκριμένα:

- Όσον αφορά, τους πόλους που απεικονίζονται με μαύρο χρώμα στο παραπάνω διάγραμμα, ισχύει ότι:
 - Για την απότομη κορυφή στην συχνότητα $286/2\pi = 46$ Hz του διαγράμματος της εκτίμησης του φάσματος του παραμετρικού μοντέλου, ευθύνεται ο πόλος που φαίνεται στο Διάγραμμα 25, με συχνότητα 286 rad/s.
 - Επιπλέον, για την απότομη κορυφή στην συχνότητα $299/2\pi = 48$ Hz του διαγράμματος της εκτίμησης του φάσματος του παραμετρικού μοντέλου, ευθύνεται ο πόλος που φαίνεται στο Διάγραμμα 25, με συχνότητα 299 rad/s.

Οι συγκεκριμένοι πόλοι βρίσκονται ελάχιστα έξω από τον μοναδιαίο κύκλο, όπου για την επίλυση αυτού του προβλήματος απαραίτητο θα ήταν η χρήση της μεθόδου αντιστρεψιμότητας έτσι ώστε να επιτευχθεί η επαναφορά των πόλων στον μοναδιαίο κύκλο. Ωστόσο οριακά δεν δεχόμαστε την αστάθεια.

Παρακάτω παρουσιάζονται οι απότομες κορυφές της εκτίμησης τους φάσματος στα 46 και 48 Hz.

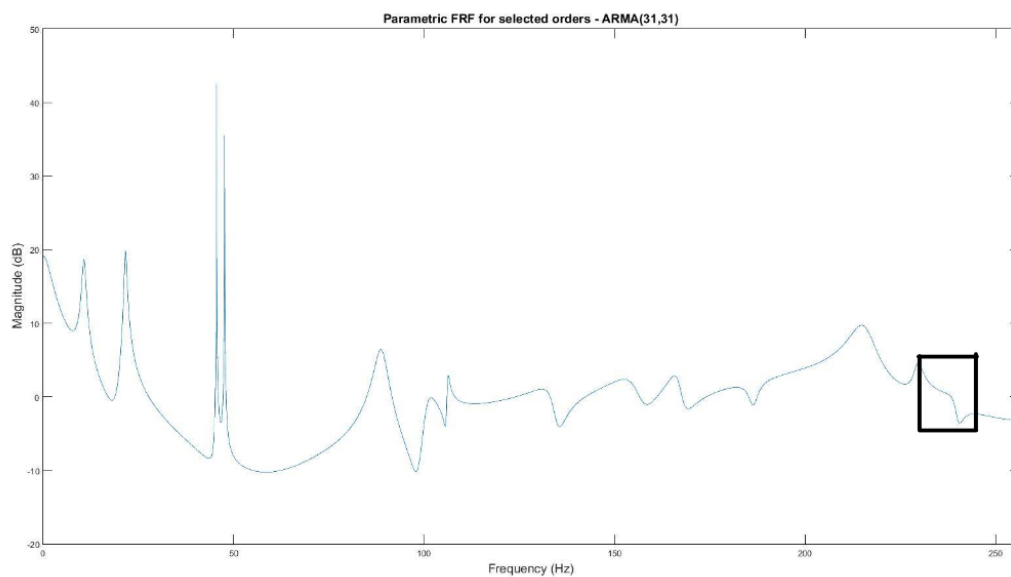


Εικόνα 30: Απεικόνιση πόλων που δημιουργούν απότομες κορυφές στα 46 και 48 Hz στο διάγραμμα εκτίμησης φάσματος.

➤ Όσον αφορά, τον μηδενιστή ισχύει ότι:

- Για την απότομη κοιλάδα στην συχνότητα $1500/2\pi = 239$ Hz του διαγράμματος της εκτίμησης του φάσματος του παραμετρικού μοντέλου, ευθύνεται ο μηδενιστής που φαίνεται στο Διάγραμμα 25, με συχνότητα 1500 rad/s.

Παρακάτω παρουσιάζεται η απότομη κοιλάδα της εκτίμησης τους φάσματος στα 239 Hz.

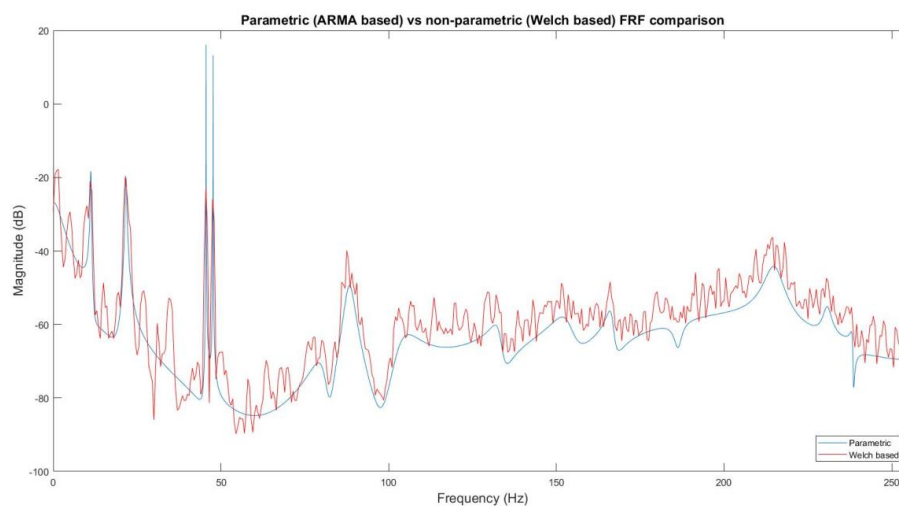


Εικόνα 31: Απεικόνιση μηδενιστή που δημιουργεί απότομη κοιλάδα στα 239 Hz στο διάγραμμα εκτίμησης φάσματος.

Παρακάτω, ακολουθεί ο Πίνακας 1, ο οποίος παρουσιάζει τις φυσικές συχνότητες, για τους αντίστοιχους λόγους απόσβεσης (ζ). Με σκούρα γράμματα παρατίθενται οι πραγματικές συχνότητες, ενώ με κανονικά οι ψευδοσυχνότητες.

Πίνακας 1	
Φυσικές Συχνότητες	Λόγοι Απόσβεσης
11.2109	1.2886
21.6666	1.0766
45.5464	0.0041
47.6357	0.0009
80.7713	3.3659
88.3152	1.7835
103.7729	4.1078
133.0024	1.4320
153.3486	2.3373
166.4935	0.7379
186.2991	0.9643
206.4213	20.1360
214.9819	1.3525
230.7333	0.6028
238.3439	0.1447

Στην συνέχεια, παρατίθεται το Διάγραμμα 32, το οποίο μας παρουσιάζει την σύγκριση του φάσματος, μεταξύ του παραμετρικού μοντέλου ARMA(31,31) και του μη-παραμετρικού μοντέλου Welch.



Εικόνα 32: Η σύγκριση του παραμετρικού μοντέλου ARMA(31,31) με το μη-παραμετρικό μοντέλο Welch.

- **Παρατηρήσεις:**

Παρατηρείται πως το παραμετρικό μας μοντέλο ακολουθεί με ικανοποιητική ακρίβεια, την εκτίμηση κατά Welch. Αυτό σημαίνει ότι το συγκεκριμένο μοντέλο είναι πολύ καλό για την παραμετρική μοντελοποίηση.

Διαφορές υπάρχουν βέβαια, και πιο συγκεκριμένα στις απότομες κορυφές που δημιουργούνται στα 46 και 48 Hz αντίστοιχα. Σε αυτές τις συχνότητες το μοντέλο μας έχει, όπως τονίστηκε και παραπάνω, δύο πόλους οι οποίοι βρίσκονται ελάχιστα έξω από τον μοναδιαίο κύκλο. Πάντως, αν εξαιρεθεί το γεγονός των πόλων που βρίσκονται έξω από τον μοναδιαίο κύκλο, το συγκεκριμένο μοντέλο μοντελοποιεί παραμετρικά αρκετά σωστά το πρόβλημά μας.

Διασπορά Λευκού Θορύβου.

$$\hat{\sigma}_W^2 = \frac{RSS}{N - \ell}$$

$$= 0.0039$$

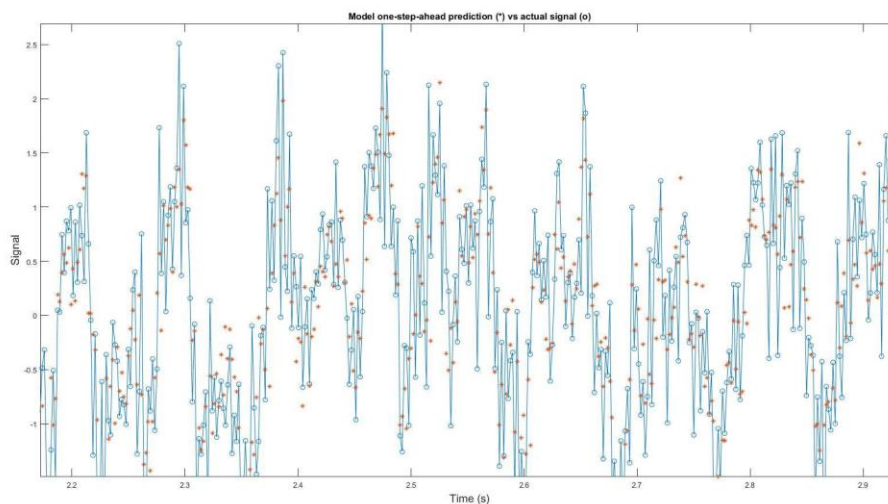
- **Παρατηρήσεις:**

Ικανοποιητικό αποτέλεσμα αφού επιτυγχάνεται η μικρή διασπορά του Λευκού Θορύβου.

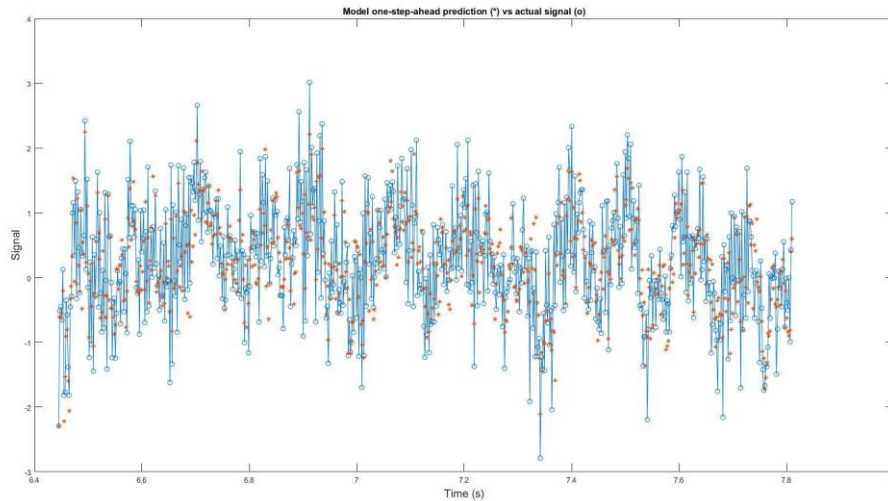
Γ5: Πρόβλεψη βάσει του επιλεγέντος μοντέλου ARMA.

Στο Διάγραμμα 33 φαίνονται οι προβλέψεις του ενός βήματος, μαζί με τις πραγματικές τιμές, για το τμήμα εκτίμησης. Έχει γίνει κατάλληλο ζουμ στο διάγραμμα, έτσι ώστε να φαίνονται καλύτερα οι προβλέψεις και οι πραγματικές τιμές.

Στο Διάγραμμα 34 φαίνονται οι προβλέψεις του ενός βήματος, μαζί με τις πραγματικές τιμές, για το τμήμα ελέγχου εγκυρότητας.



Εικόνα 33: Οι προβλέψεις ενός βήματος και οι πραγματικές τιμές, για το τμήμα εκτίμησης (σε μεγέθυνση).



Εικόνα 34: Οι προβλέψεις ενός βήματος και οι πραγματικές τιμές για το τμήμα ελέγχου εγκυρότητας.

○ Παρατηρήσεις:

Από τα δύο προηγούμενα διαγράμματα παρατηρείται ότι οι προβλεφθείσες τιμές έχουν μικρή, αλλά αισθητή απόκλιση από τις πραγματικές τιμές, είτε εξετάζουμε το τμήμα εκτίμησης είτε το τμήμα ελέγχου εγκυρότητας. Μεγαλύτερες αποκλίσεις παρουσιάζονται από όσο φαίνεται στο διάγραμμα 34 και συγκεκριμένα στο τμήμα ελέγχου εγκυρότητας. Αυτές οφείλονται κατά κανόνα στην θεώρηση ότι το σήμα μας είναι στάσιμο και ότι όλες οι μεταβλητές ακολουθούν κανονική κατανομή. Επίσης ακόμα και εδώ επιβεβαιώνονται τα χαρακτηριστικά μη στασιμότητας για το σήμα μας. Το συγκεκριμένο γεγονός δημιούργησε κάποιες διαφοροποιήσεις στο μοντέλο μας από το πραγματικό. Παρόλα αυτά όμως, θεωρούμε πως οι αποκλίσεις αυτές των δύο τμημάτων είναι αναμενόμενες, αφού δεν γίνεται να έχουμε ταύτιση της πραγματικής με την προβλεφθείσα τιμή πάντοτε. Έτσι λοιπόν, θεωρούμε πως το μοντέλο είναι αποδεκτό, και υπάρχει ικανοποιητική πρόβλεψη του ενός βήματος.