

Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Υπολογιστική Εργασία Χρονοσειρές 2018

Διαμαντή Μαρία 8133
mfdiamanti@ece.auth.gr

Ντζιώνη Δήμητρα 8209
dntzioni@ece.auth.gr

Θεσσαλονίκη, Φεβρουάριος 2018

Περιεχόμενα

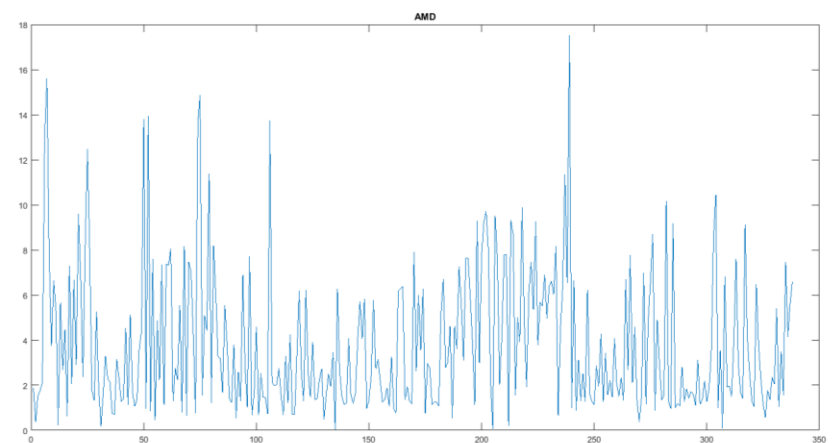
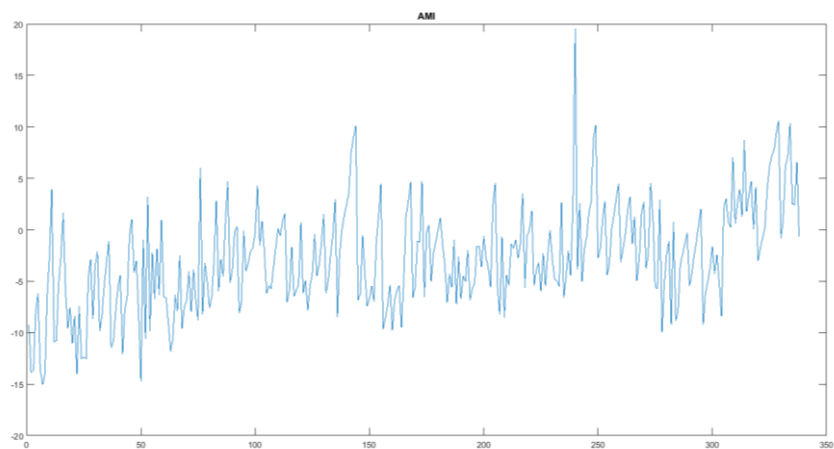
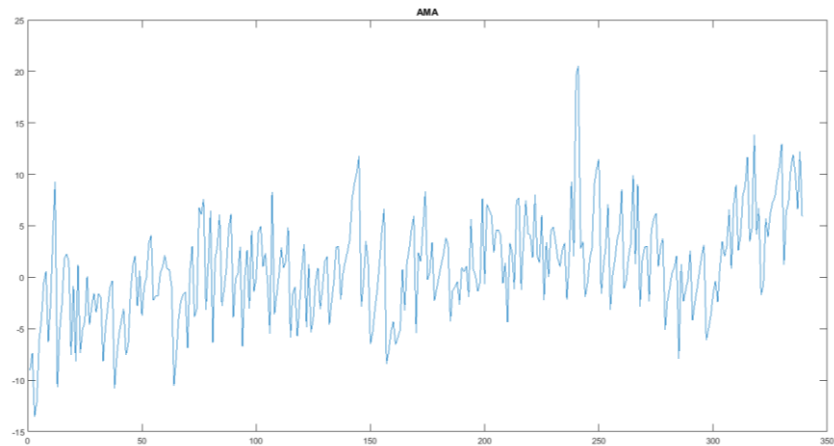
1	OKN Χρονοσειρά AMI dat29v1	3
1.1	Ερώτημα Πρώτο	3
1.2	Ερώτημα Δεύτερο	5
1.2.1	Απαλοιφή τάσης	5
1.2.2	Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.....	6
1.2.3	Διερεύνηση κατάλληλου γραμμικού μοντέλου	6
1.2.4	Πρόβλεψη ενός και δύο βημάτων μπροστά	7
2	OKN Χρονοσειρά AMI dat29v2	8
2.1	Ερώτημα Πρώτο	8
2.2	Ερώτημα Δεύτερο	10
2.2.1	Απαλοιφή τάσης	10
2.2.2	Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.....	11
2.2.3	Διερεύνηση κατάλληλου γραμμικού μοντέλου	11
2.2.4	Πρόβλεψη ενός και δύο βημάτων μπροστά	12
3	Σύγκριση Χρονοσειρών AMI των dat29v1 και dat29v2.....	13
4	OKN Χρονοσειρά TMI dat29v1	14
4.1	Ερώτημα Δεύτερο	14
4.1.1	Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.....	14
4.1.2	Διερεύνηση κατάλληλου γραμμικού μοντέλου	15
4.1.3	Πρόβλεψη ενός και δύο βημάτων μπροστά	16
5	OKN Χρονοσειρά TMI dat29v2	17
5.1	Ερώτημα Δεύτερο	17
5.1.1	Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.....	17
5.1.2	Διερεύνηση κατάλληλου γραμμικού μοντέλου	18
5.1.3	Πρόβλεψη ενός και δύο βημάτων μπροστά	18
6	Σύγκριση Χρονοσειρών TMI των dat29v1 και dat29v2	19
7	Μη γραμμική ανάλυση dat29v1.....	20
7.1	Διαγράμματα διασποράς στις δύο και στις τρεις διαστάσεις.....	20
7.2	Κατάλληλες παράμετροι ανακατασκευής χώρου κατάστασης με μέθοδο υστερήσεων	21
7.3	Υπολογισμός διάστασης συσχέτισης	23
7.4	Διερεύνηση κατάλληλου τοπικού μοντέλου	23

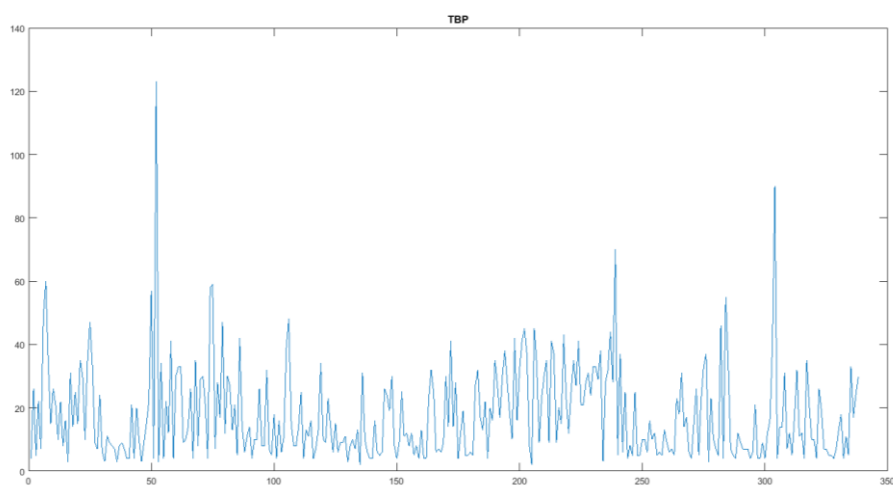
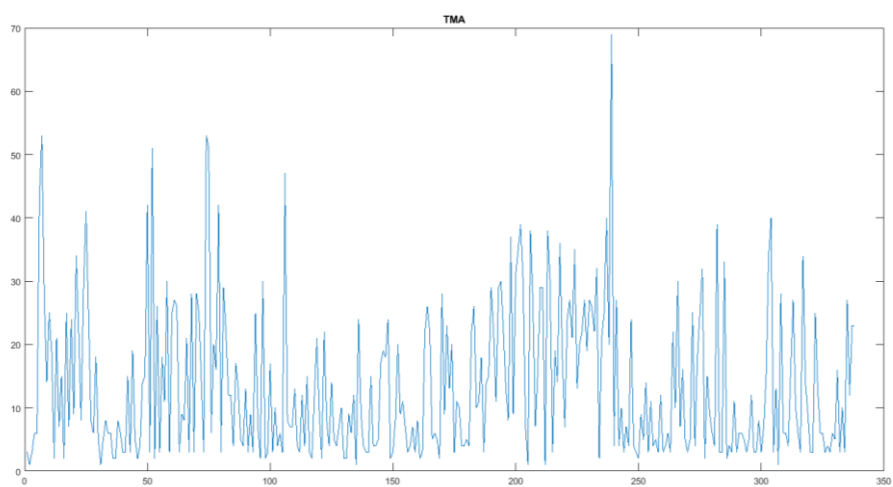
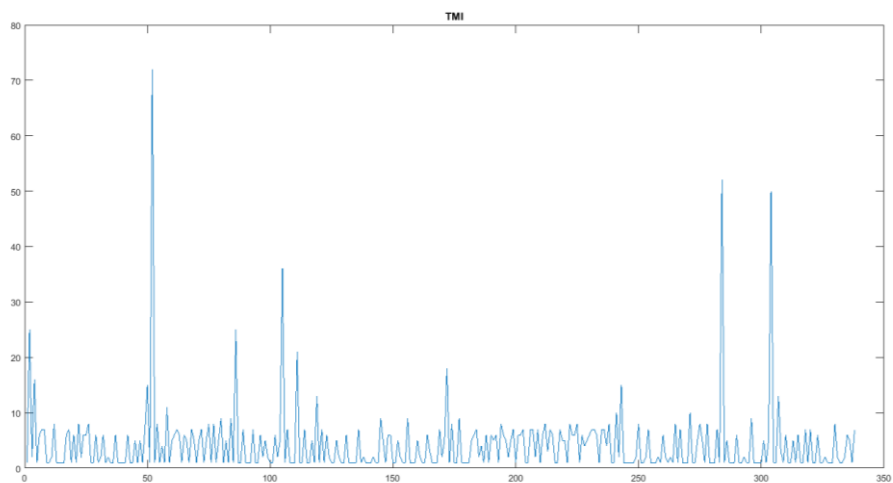
7.5	Διερεύνηση καταλληλότερων παραμέτρων ανακατασκευής χώρου κατάστασης.....	24
7.5.1	Υπολογισμός διάστασης εμβύθισης για υστέρηση $\tau = 1$	24
7.5.2	Υπολογισμός διάστασης συσχέτισης για $\tau = 1$	25
7.5.3	Διερεύνηση κατάλληλου τοπικού μοντέλου για $\tau = 1$	25
8	Μη γραμμική ανάλυση dat29v2.....	26
8.1	Κατάλληλες παράμετροι ανακατασκευής χώρου κατάστασης με μέθοδο υστερήσεων	27
8.2	Υπολογισμός διάστασης συσχέτισης	29
8.3	Διερεύνηση κατάλληλου τοπικού μοντέλου	30
8.4	Διερεύνηση καταλληλότερων παραμέτρων ανακατασκευής χώρου κατάστασης.....	31
8.4.1	Υπολογισμός διάστασης εμβύθισης για υστέρηση $\tau = 1$	31
8.4.2	Υπολογισμός διάστασης συσχέτισης για $\tau = 1$	31
8.4.3	Διερεύνηση κατάλληλου τοπικού μοντέλου για $\tau = 1$	32
9	Σύγκριση μη γραμμικής ανάλυσης OKN χρονοσειρών dat29v1 και dat29v2	33
10	Συμπεράσματα γραμμικής και μη γραμμικής ανάλυσης.....	33

1 ΟΚΝ Χρονοσειρά AMI dat29v1

1.1 Ερώτημα Πρώτο

Ζητούμενο του ερωτήματος αυτού είναι ο υπολογισμός των τοπικών ακροτάτων της χρονοσειράς dat29v1 μέσω της δοθείσας συνάρτησης `extremes.m`. Έπειτα, από τα ακρότατα αυτά εξάχθηκαν όλες οι απαιτούμενες χρονοσειρές (AMA, AMI, AMD, TMI, TMA, TBP), τα γραφήματα των οποίων παρουσιάζονται παρακάτω. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `grammiki_analysis - AMI.m`.





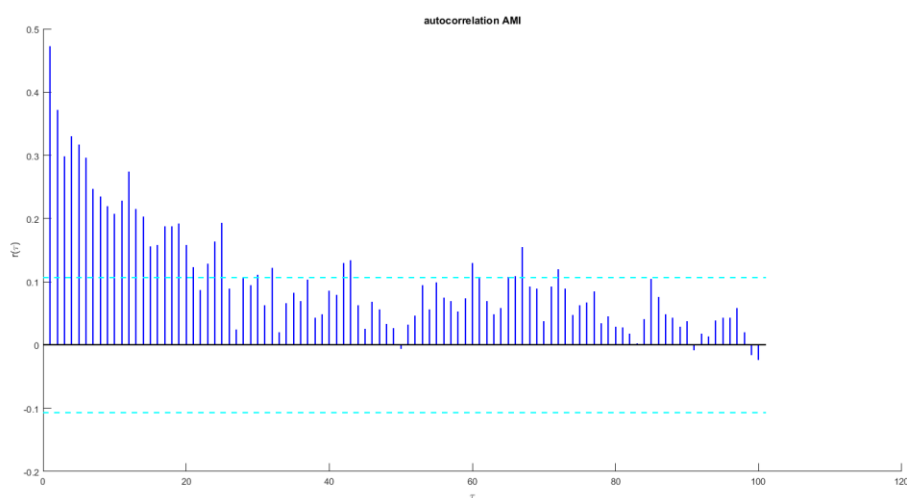
Η χρονοσειρά που επιλέγεται από τις παραπάνω για γραμμική ανάλυση είναι η **AMI**.

1.2 Ερώτημα Δεύτερο

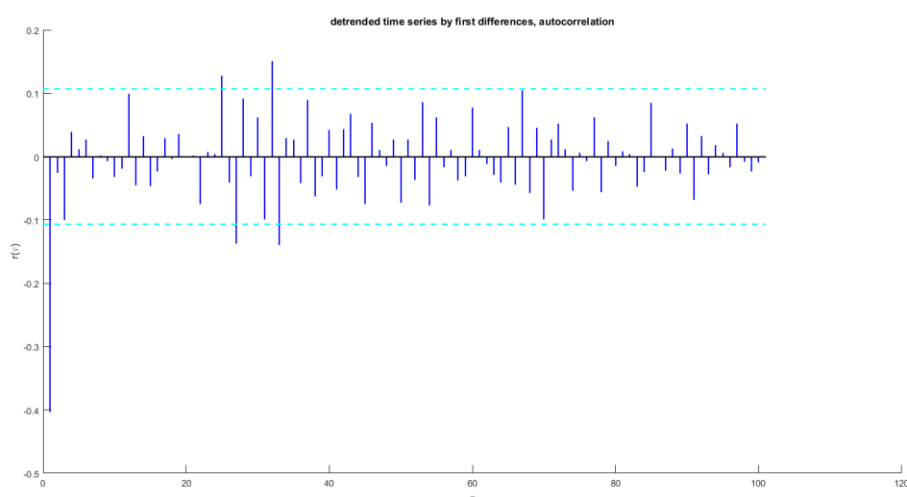
Στο ερώτημα αυτό πραγματοποιούμε γραμμική ανάλυση για τη χρονοσειρά AMI των δεδομένων dat29v1. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `grammiki_analysis - AMI.m`.

1.2.1 Απαλοιφή τάσης

Από το γράφημα της AMI, το οποίο παρατέθηκε στο προηγούμενο ερώτημα, παρατηρούμε ότι η χρονοσειρά αυτή παρουσιάζει τάση. Το γεγονός αυτό μπορεί να γίνει εμφανές και από το ακόλουθο γράφημα της συνάρτησης αυτοσυσχέτισης, η οποία βλέπουμε ότι μειώνεται πολύ αργά. Η αυτοσυσχέτιση υπολογίζεται με τη βοήθεια του δοσμένου κώδικα `autocorrelation.m`.



Για το λόγο αυτό, προχωράμε σε πρώτες διαφορές και λαμβάνουμε το ακόλουθο γράφημα αυτοσυσχέτισης.

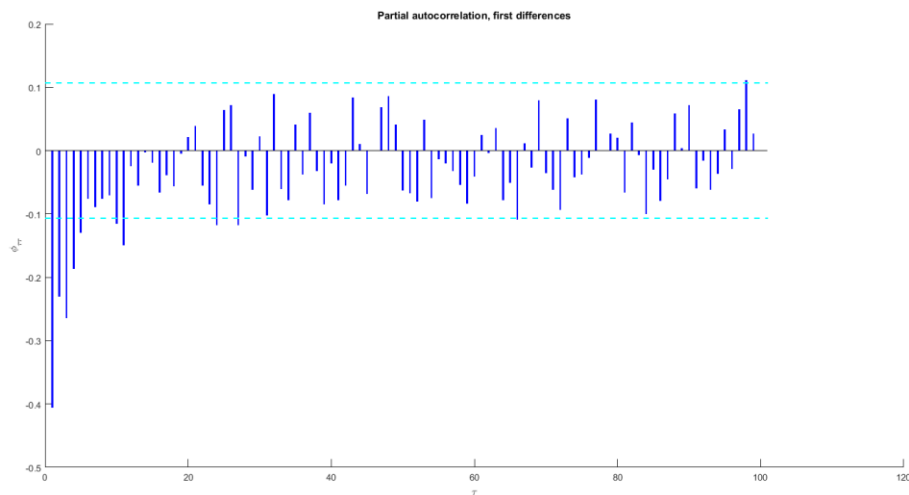


Όπως είναι εμφανές, η τάση αυτή έχει εξαλειφθεί και μπορούμε πλέον να συνεχίσουμε κανονικά την υπόλοιπη γραμμική μας ανάλυση, επάνω στη χρονοσειρά των πρώτων διαφορών.

1.2.2 Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

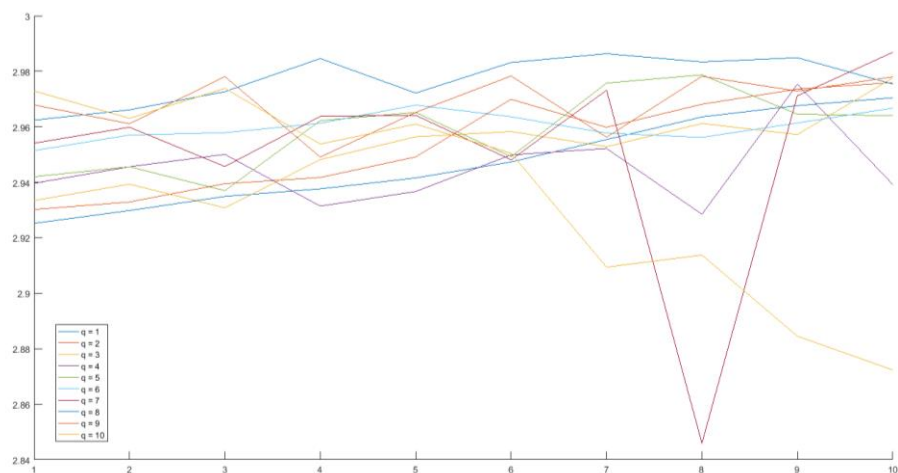
Το γράφημα της αυτοσυσχέτισης είναι αυτό που παρουσιάστηκε προηγουμένως.

Υπολογίζουμε, λοιπόν, και τη μερική αυτοσυσχέτιση της χρονοσειράς χωρίς τάση, η οποία φαίνεται στο ακόλουθο γράφημα. Για το σκοπό αυτό έγινε χρήση της συνάρτησης `parautocor.m`.



1.2.3 Διερεύνηση κατάλληλου γραμμικού μοντέλου

Σκοπός του ερωτήματος αυτού είναι να διερευνήσουμε και εν τέλει να προσαρμόσουμε ένα γραμμικό μοντέλο, $ARMA(p,q)$, το οποίο ταιριάζει καλύτερα στη χρονοσειρά μας. Προκειμένου να βρεθεί ένα ικανοποιητικό μοντέλο, χρησιμοποιούμε την μετρική AIC και δημιουργούμε ένα γράφημα αυτής που περιέχει όλους τους συνδυασμούς των τάξεων p και q , με μέγιστες τάξεις $p=q=10$ που απεικονίζεται παρακάτω. Στο γράφημα αυτό, για κάθε μια τιμή της τάξης q υπολογίζεται η μετρική AIC για όλες τις τάξεις p . Από αυτούς τους 100 συνδυασμούς τάξεων που προκύπτουν επιλέγουμε εκείνον το συνδυασμό p και q με το ελάχιστο AIC. Αναφέρεται πως έγινε χρήση του δοσμένου αρχείου `fitARMA.m`.

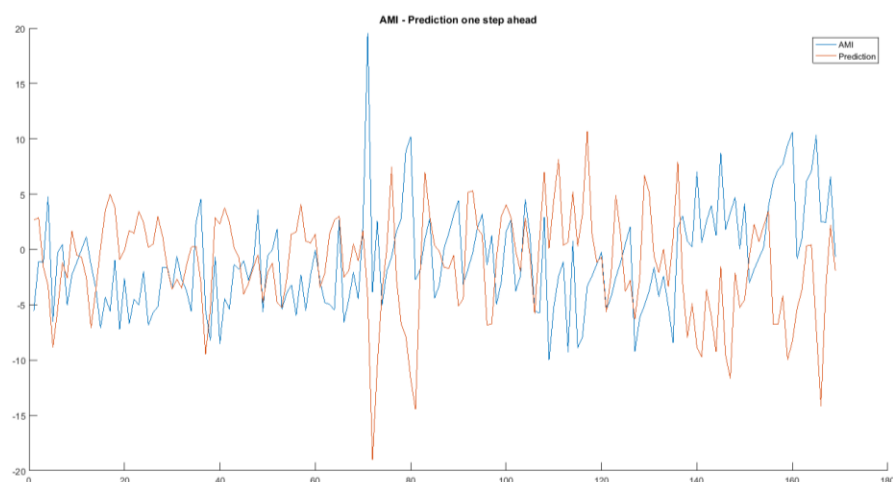


Μέσω κώδικα υπολογίζουμε το βέλτιστο συνδυασμό και καταλήγουμε ότι το μοντέλο που ταιριάζει καλύτερα στη χρονοσειρά μας είναι το ARMA(8,7), πράγμα που φαίνεται και από το γράφημα.

1.2.4 Πρόβλεψη ενός και δύο βημάτων μπροστά

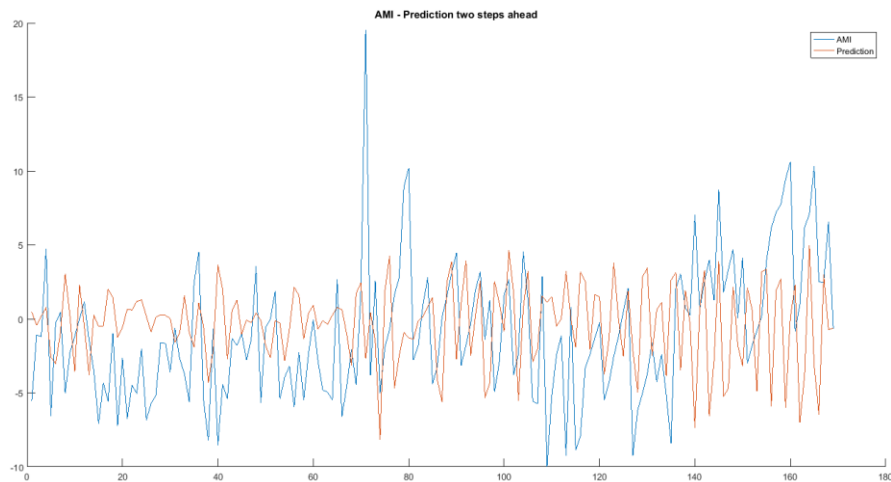
Στο ερώτημα αυτό της εργασίας κάνουμε πρόβλεψη ενός και δύο βημάτων μπροστά επάνω στη χρονοσειρά χωρίς τάση. Επισημαίνεται ότι χρησιμοποιήθηκε το δεύτερο μισό της χρονοσειράς ως test set στο αρχείο predictARMAnormse.m. Έπειτα, στόχος μας είναι να ανάγουμε τις προβλέψεις αυτές στην αρχική μας χρονοσειρά με τάση, δηλαδή την AMI. Κάνουμε, λοιπόν, την αντίστροφη διαδικασία από αυτήν των πρώτων διαφορών και παίρνουμε τις τελικές επιθυμητές προβλέψεις πάνω στη χρονοσειρά AMI. Στη συνέχεια, ακολουθούν γραφήματα των προβλέψεων για ένα και για δύο βήματα μπροστά συγκριτικά με τη χρονοσειρά AMI. Επίσης, σημειώνονται και τα αντίστοιχα NRMSEs, με βάση τη συνάρτηση normse.m.

Πρόβλεψη ενός βήματος μπροστά



Παρατηρούμε από το γράφημα ότι η πρόβλεψη δεν είναι καλή, πράγμα που επιβεβαιώνεται και από το NRMSE: $\text{normseAMI1} = 1.5974$. Όπως γνωρίζουμε, τιμές NRMSE γύρω από το 1 δίνουν προβλέψεις χειρότερες και από το εάν προβλέπαμε με τη μέση τιμή της χρονοσειράς, ενώ τιμές NRMSE κοντά στο μηδέν δηλώνουν ότι η πρόβλεψη είναι ικανοποιητική. Γενικότερα, όσο μικρότερη είναι η τιμή του NRMSE τόσο το καλύτερο.

Πρόβλεψη δύο βημάτων μπροστά

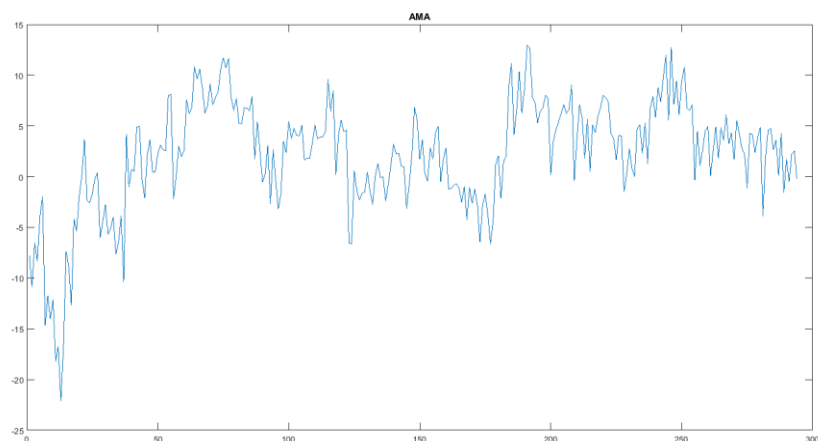


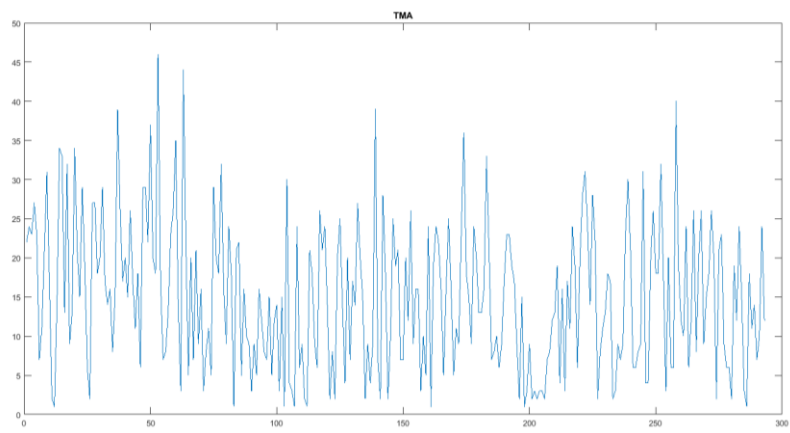
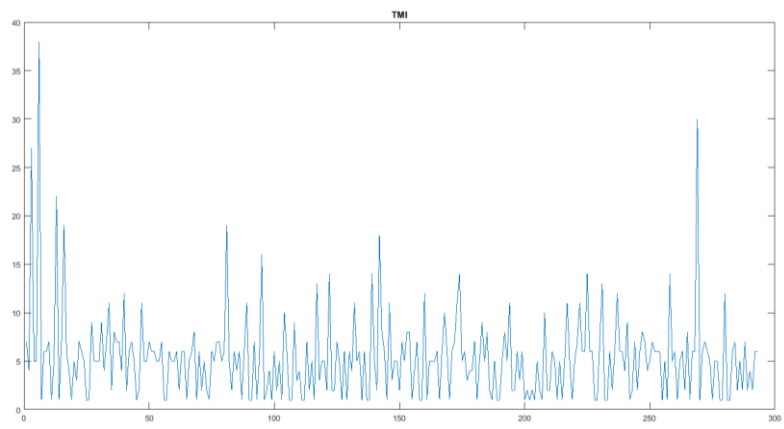
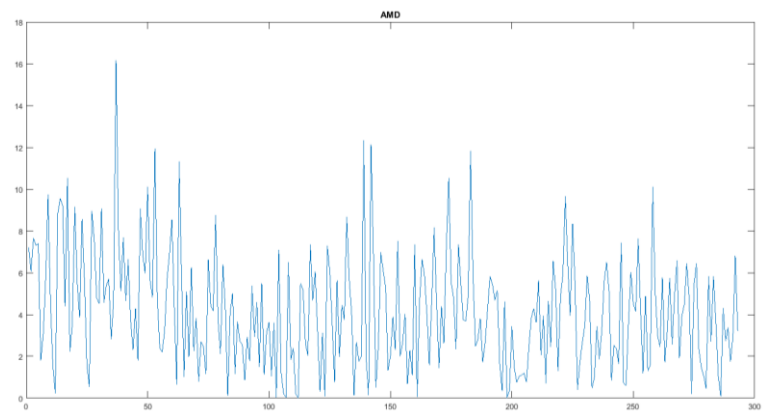
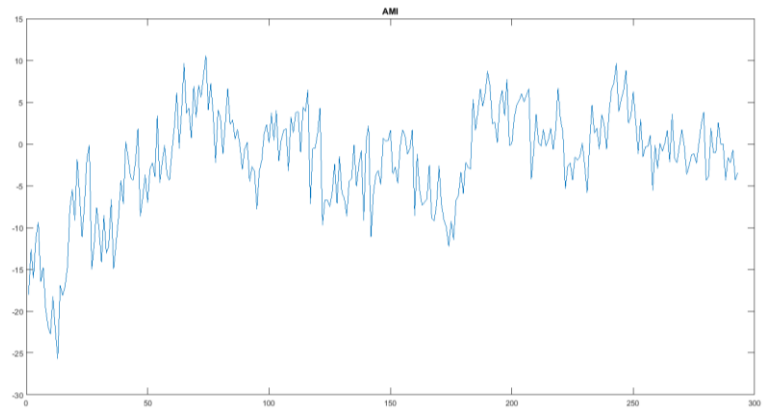
Και για τα δύο βήματα μπροστά οι προβλέψεις δεν είναι καλές. Παρομοίως το NRMSE είναι κοντά στη μονάδα και συγκεκριμένα ισούται με $\text{nrmseAMI2} = 1.2182$.

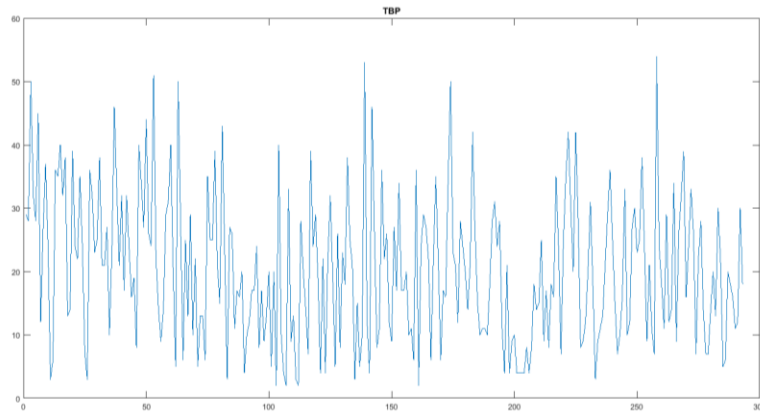
2 ΟΚΝ Χρονοσειρά AMI dat29v2

2.1 Ερώτημα Πρώτο

Ζητούμενο του ερωτήματος αυτού είναι ο υπολογισμός των τοπικών ακροτάτων της χρονοσειράς dat29v2 μέσω της δοθείσας συνάρτησης `extremes.m`. Έπειτα, από τα ακρότατα αυτά εξάχθηκαν όλες οι απαιτούμενες χρονοσειρές (AMA, AMI, AMD, TMI, TMA, TBP), τα γραφήματα των οποίων παρουσιάζονται παρακάτω. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `grammiki_analysis - AMI.m`.







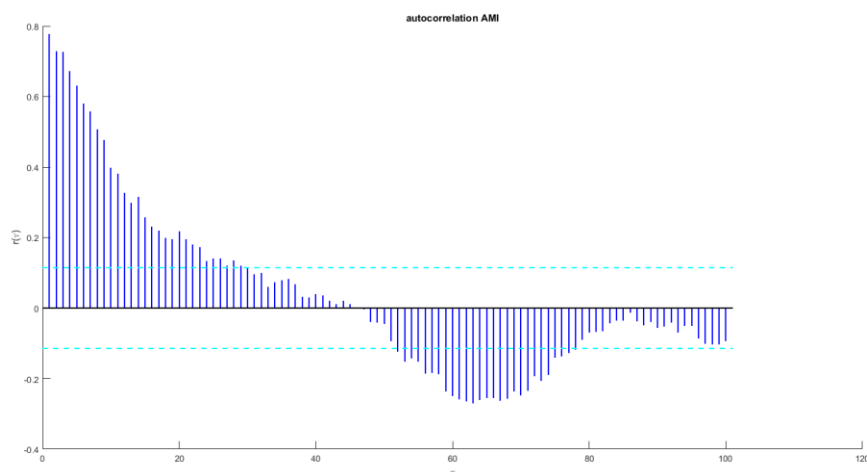
Η χρονοσειρά που επιλέγεται από τις παραπάνω για γραμμική ανάλυση είναι η **AMI**.

2.2 Ερώτημα Δεύτερο

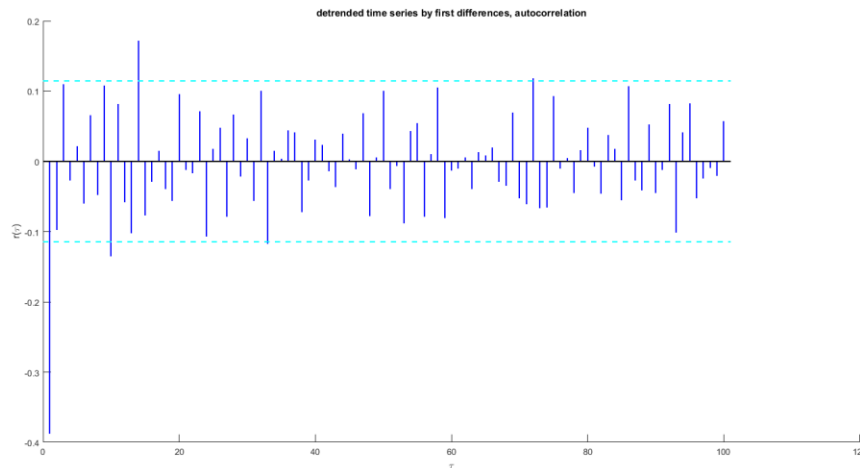
Στο ερώτημα αυτό πραγματοποιούμε γραμμική ανάλυση για τη χρονοσειρά AMI των δεδομένων dat29v2. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `grammiki_analysis – AMI.m`.

2.2.1 Απαλοιφή τάσης

Από το γράφημα της AMI, το οποίο παρατέθηκε στο προηγούμενο ερώτημα, παρατηρούμε ότι η χρονοσειρά αυτή παρουσιάζει τάση. Το γεγονός αυτό μπορεί να γίνει εμφανές και από το ακόλουθο γράφημα της συνάρτησης αυτοσυσχέτισης, η οποία βλέπουμε ότι μειώνεται πολύ αργά. Η αυτοσυσχέτιση υπολογίζεται με τη βοήθεια του δοσμένου κώδικα `autocorrelation.m`.



Για το λόγο αυτό, προχωράμε σε πρώτες διαφορές και λαμβάνουμε το ακόλουθο γράφημα αυτοσυσχέτισης.

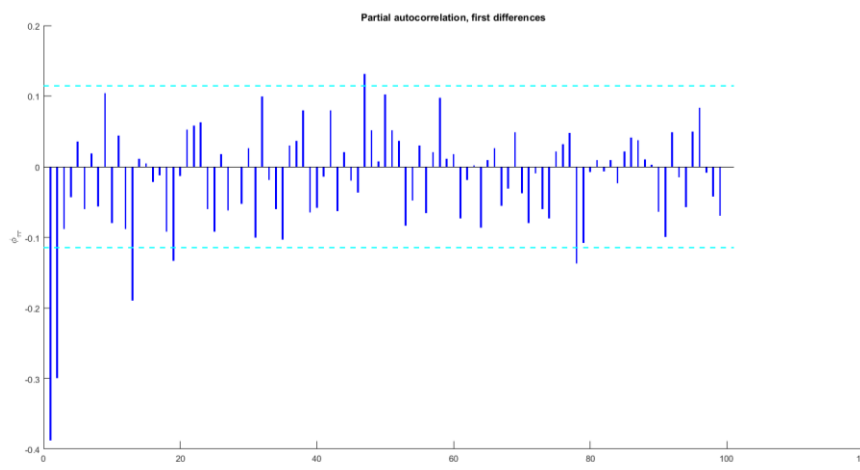


Όπως είναι εμφανές, η τάση αυτή έχει εξαλειφθεί και μπορούμε πλέον να συνεχίσουμε κανονικά την υπόλοιπη γραμμική μας ανάλυση, επάνω στη χρονοσειρά των πρώτων διαφορών.

2.2.2 Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

Το γράφημα της αυτοσυσχέτισης είναι αυτό που παρουσιάστηκε προηγουμένως.

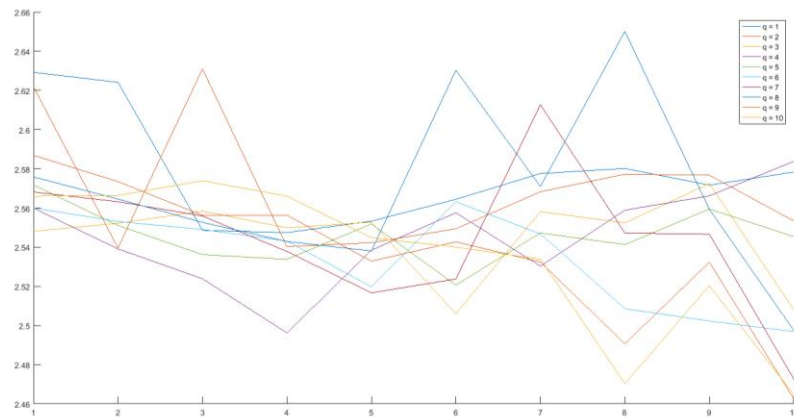
Υπολογίζουμε, λοιπόν, και τη μερική αυτοσυσχέτιση της χρονοσειράς χωρίς τάση, η οποία φαίνεται στο ακόλουθο γράφημα. Για το σκοπό αυτό έγινε χρήση της συνάρτησης `parautocor.m`.



2.2.3 Διερεύνηση κατάλληλου γραμμικού μοντέλου

Σκοπός του ερωτήματος αυτού είναι να διερευνήσουμε και εν τέλει να προσαρμόσουμε ένα γραμμικό μοντέλο, $ARMA(p,q)$, το οποίο ταιριάζει καλύτερα στη χρονοσειρά μας. Προκειμένου να βρεθεί ένα ικανοποιητικό μοντέλο, χρησιμοποιούμε την μετρική AIC και δημιουργούμε ένα γράφημα αυτής που περιέχει όλους τους συνδυασμούς των τάξεων p και q , με μέγιστες τάξεις $p=q=10$ που

απεικονίζεται παρακάτω. Στο γράφημα αυτό, για κάθε μια τιμή της τάξης q υπολογίζεται η μετρική AIC για όλες τις τάξεις p . Από αυτούς τους 100 συνδυασμούς τάξεων που προκύπτουν επιλέγουμε εκείνον το συνδυασμό p και q με το ελάχιστο AIC. Αναφέρεται πως έγινε χρήση του δοσμένου αρχείου `fitARMA.m`.

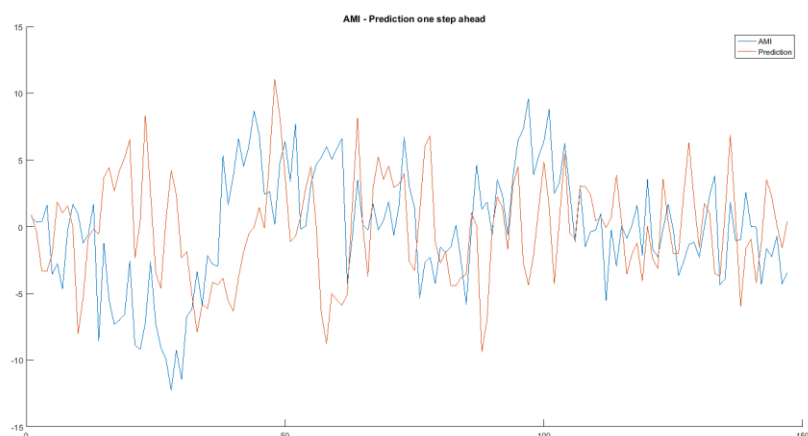


Μέσω κώδικα υπολογίζουμε το βέλτιστο συνδυασμό και καταλήγουμε ότι το μοντέλο που ταιριάζει καλύτερα στη χρονοσειρά μας είναι το $ARMA(10,9)$, πράγμα που φαίνεται και από το γράφημα.

2.2.4 Πρόβλεψη ενός και δύο βημάτων μπροστά

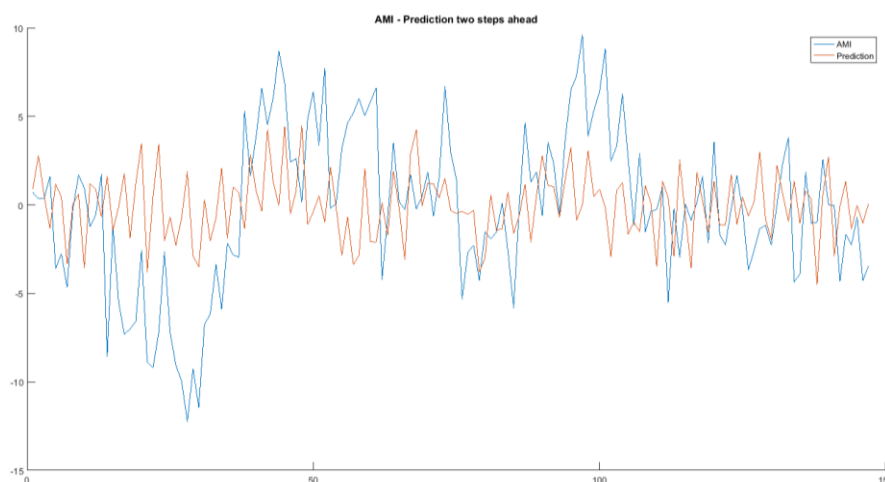
Στο ερώτημα αυτό της εργασίας κάνουμε πρόβλεψη ενός και δύο βημάτων μπροστά επάνω στη χρονοσειρά χωρίς τάση. Επισημαίνεται ότι χρησιμοποιήθηκε το δεύτερο μισό της χρονοσειράς ως `test set` στο αρχείο `predictARMArmse.m`. Έπειτα, στόχος μας είναι να ανάγουμε τις προβλέψεις αυτές στην αρχική μας χρονοσειρά με τάση, δηλαδή την AMI. Κάνουμε, λοιπόν, την αντίστροφη διαδικασία από αυτήν των πρώτων διαφορών και παίρνουμε τις τελικές επιθυμητές προβλέψεις πάνω στη χρονοσειρά AMI. Στη συνέχεια, ακολουθούν γραφήματα των προβλέψεων για ένα και για δύο βήματα μπροστά συγκριτικά με τη χρονοσειρά AMI. Επίσης, σημειώνονται και τα αντίστοιχα NRMSEs, με βάση τη συνάρτηση `rmse.m`.

Πρόβλεψη ενός βήματος μπροστά



Παρατηρούμε από το γράφημα ότι η πρόβλεψη δεν είναι καλή, πράγμα που επιβεβαιώνεται και από το NRMSE: $\text{nrmseAMI1} = 1.3473$.

Πρόβλεψη δύο βημάτων μπροστά



Και για τα δύο βήματα μπροστά οι προβλέψεις δεν είναι καλές. Παρομοίως το NRMSE είναι κοντά στη μονάδα και συγκεκριμένα ισούται με $\text{nrmseAMI2} = 1.0352$.

3 Σύγκριση Χρονοσειρών AMI των dat29v1 και dat29v2

Από όλη την παραπάνω ανάλυση καταλήγουμε στο συμπέρασμα ότι το γραμμικό μοντέλο που προσαρμόσαμε και στις δύο χρονοσειρές AMI των δεδομένων dat29v1 και dat29v2 δεν είναι αρκετό ώστε να γίνουν ορθές προβλέψεις. Σίγουρα θα πρέπει να γίνει έλεγχος και για άλλα μοντέλα, και πιθανόν μη γραμμικά και για τις δύο αυτές χρονοσειρές. Η χρονοσειρά AMI των δεδομένων dat29v2 φαίνεται να παρουσιάζει ελαφρώς πιο ισχυρές συσχετίσεις από αυτή των δεδομένων dat29v1 . Αυτό μπορεί να φανεί από το διάγραμμα της αυτοσυσχέτισής της με τάση, και επομένως και από το διάγραμμα της αυτοσυσχέτισής της χωρίς τάση, εφόσον οι τιμές της αυτοσυσχέτισης για διάφορες υστερήσεις είναι πιο κοντά στο 1 ή στο -1 (ισχυρή θετική ή αρνητική συσχέτιση) συγκριτικά με αυτή των δεδομένων dat29v1 .

Όσον αφορά την προσαρμογή των χρονοσειρών σε κάποιο γραμμικό μοντέλο, η χρονοσειρά AMI των δεδομένων dat29v2 έχει μικρότερη τιμή AIC για τις επιλεχθείσες τάξεις p και q του μοντέλου ARMA συγκριτικά με τη χρονοσειρά AMI των δεδομένων dat29v1 . Αυτό γίνεται εύκολα εμφανές από τα γραφήματα των μετρικών AIC για κάθε μια από τις χρονοσειρές. Η μικρότερη τιμή AIC υποδηλώνει καλύτερη προσαρμογή.

Τέλος, όσον αφορά την προβλεψιμότητα, για μια ακόμη φορά η AMI των δεδομένων dat29v2 παρουσιάζει ελάχιστα καλύτερη συμπεριφορά από αυτήν των δεδομένων dat29v1 . Το γεγονός αυτό μπορούμε να το συμπεράνουμε από τις μικρότερες τιμές που παίρνει το κριτήριο NRMSE, και οι οποίες αναφέρθηκαν προηγουμένως.

4 OKN Χρονοσειρά TMI dat29v1

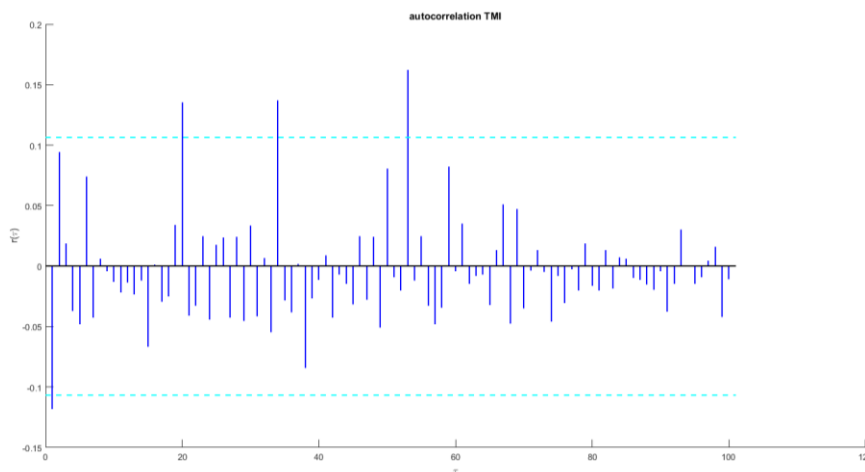
Για τη διεξαγωγή του ερωτήματος 4 της εργασίας επιλέχθηκε να γίνει γραμμική ανάλυση των χρονοσειρών **TMI** των δεδομένων dat29v1 και dat29v2, έτσι ώστε να συγκριθούν όλα τα παραπάνω συμπεράσματα για τις δύο AMI χρονοσειρές με τις δύο TMI χρονοσειρές. Ακολουθεί η γραμμική ανάλυση, η οποία βρίσκεται στο αρχείο grammiki_analysi – TMI.m.

4.1 Ερώτημα Δεύτερο

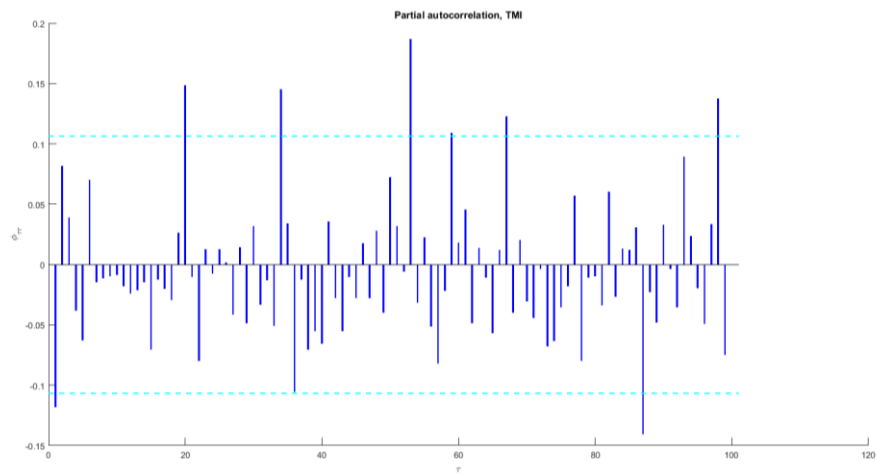
Στο ερώτημα αυτό πραγματοποιούμε γραμμική ανάλυση για τη χρονοσειρά TMI των δεδομένων dat29v1. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο grammiki_analysi – TMI.m.

4.1.1 Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

Από το γράφημα αυτοσυσχέτισης της TMI που ακολουθεί παρατηρούμε πως δεν υπάρχουν αργές μεταβολές της αυτοσυσχέτισης. Έτσι, συμπεραίνουμε ότι η χρονοσειρά αυτή δεν παρουσιάζει τάση και προχωράμε κανονικά στη γραμμική ανάλυση αυτής δίχως να πάρουμε τις πρώτες διαφορές. Η αυτοσυσχέτιση υπολογίζεται με τη βοήθεια του δοσμένου κώδικα autocorrelation.m.

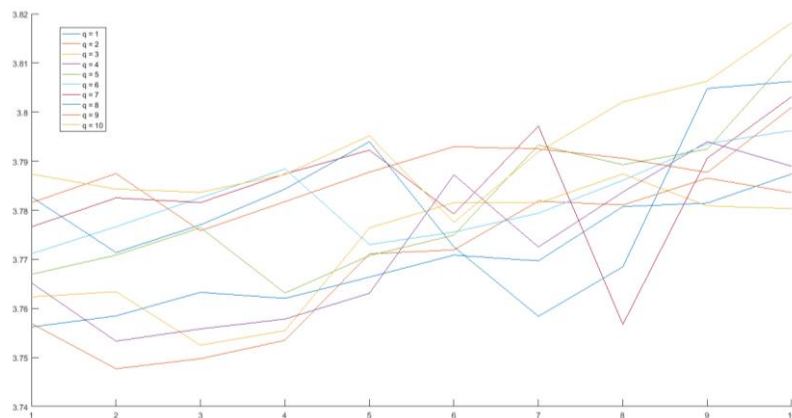


Υπολογίζουμε, λοιπόν, και τη μερική αυτοσυσχέτιση της χρονοσειράς, η οποία φαίνεται στο ακόλουθο γράφημα. Για το σκοπό αυτό έγινε χρήση της συνάρτησης parautocor.m.



4.1.2 Διερεύνηση κατάλληλου γραμμικού μοντέλου

Σκοπός του ερωτήματος αυτού είναι να διερευνήσουμε και εν τέλει να προσαρμόσουμε ένα γραμμικό μοντέλο, $ARMA(p,q)$, το οποίο ταιριάζει καλύτερα στη χρονοσειρά μας. Προκειμένου να βρεθεί ένα ικανοποιητικό μοντέλο, χρησιμοποιούμε την μετρική AIC και δημιουργούμε ένα γράφημα αυτής που περιέχει όλους τους συνδυασμούς των τάξεων p και q , με μέγιστες τάξεις $p=q=10$ που απεικονίζεται παρακάτω. Στο γράφημα αυτό, για κάθε μια τιμή της τάξης q υπολογίζεται η μετρική AIC για όλες τις τάξεις p . Από αυτούς τους 100 συνδυασμούς τάξεων που προκύπτουν επιλέγουμε εκείνον το συνδυασμό p και q με το ελάχιστο AIC. Αναφέρεται πως έγινε χρήση του δοσμένου αρχείου `fitARMA.m`.

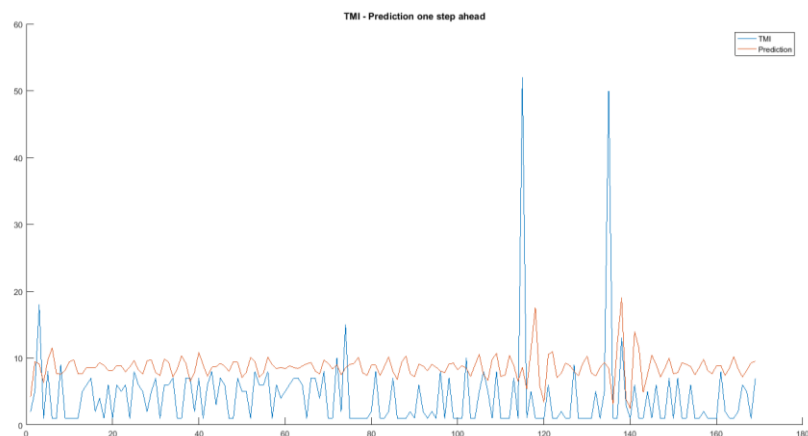


Μέσω κώδικα υπολογίζουμε το βέλτιστο συνδυασμό και καταλήγουμε ότι το μοντέλο που ταιριάζει καλύτερα στη χρονοσειρά μας είναι το $ARMA(2,2)$, πράγμα που φαίνεται και από το γράφημα.

4.1.3 Πρόβλεψη ενός και δύο βημάτων μπροστά

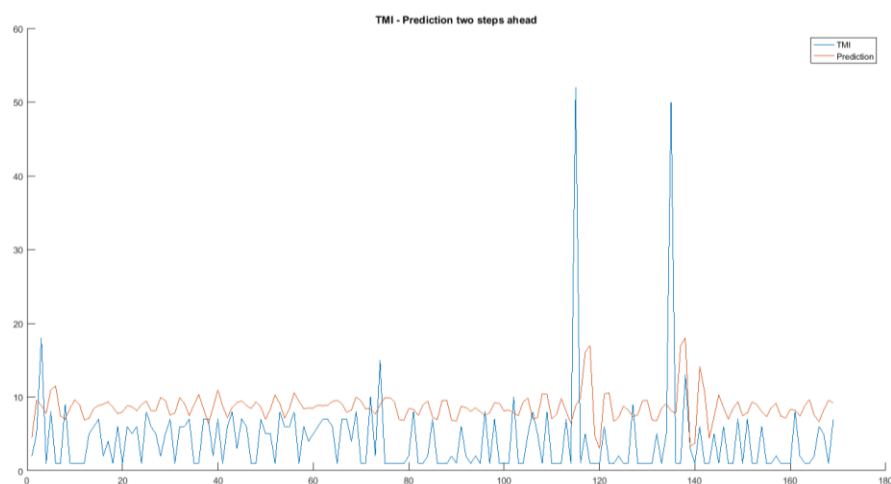
Στο ερώτημα αυτό της εργασίας κάνουμε πρόβλεψη ενός και δύο βημάτων μπροστά επάνω στη χρονοσειρά μας. Επισημαίνεται ότι χρησιμοποιήθηκε το δεύτερο μισό της χρονοσειράς ως test set στο αρχείο predictARMArmse.m. Στη συνέχεια, ακολουθούν γραφήματα των προβλέψεων για ένα και για δύο βήματα μπροστά συγκριτικά με τη χρονοσειρά TMI. Επίσης, σημειώνονται και τα αντίστοιχα NRMSEs, με βάση τη συνάρτηση nrmse.m.

Πρόβλεψη ενός βήματος μπροστά



Παρατηρούμε από το γράφημα ότι η πρόβλεψη για μια ακόμη φορά δεν είναι καλή, πράγμα που επιβεβαιώνεται και από το NRMSE: $\text{nrmseTMI1} = 1.2512$. Όπως γνωρίζουμε, τιμές NRMSE γύρω από το 1 δίνουν προβλέψεις χειρότερες και από το εάν προβλέπαμε με τη μέση τιμή της χρονοσειράς, ενώ τιμές NRMSE κοντά στο μηδέν δηλώνουν ότι η πρόβλεψη είναι ικανοποιητική.

Πρόβλεψη δύο βημάτων μπροστά



Και για τα δύο βήματα μπροστά οι προβλέψεις δεν είναι καλές. Παρομοίως το NRMSE είναι κοντά στη μονάδα και συγκεκριμένα ισούται με $\text{nrmseTMI2} = 1.2613$.

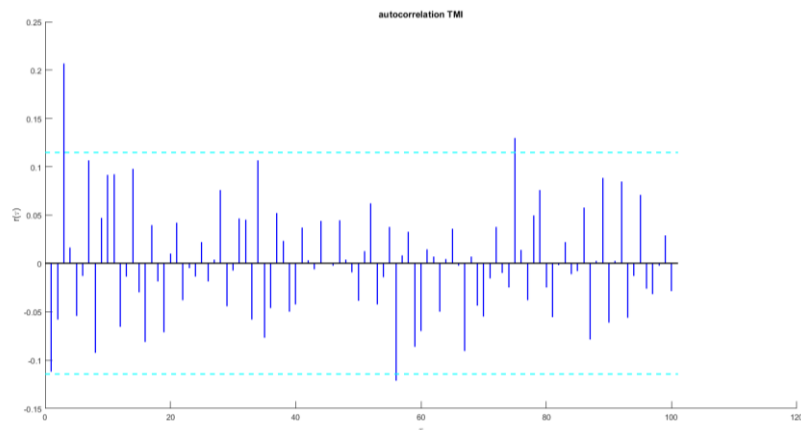
5 OKN Χρονοσειρά TMI dat29v2

5.1 Ερώτημα Δεύτερο

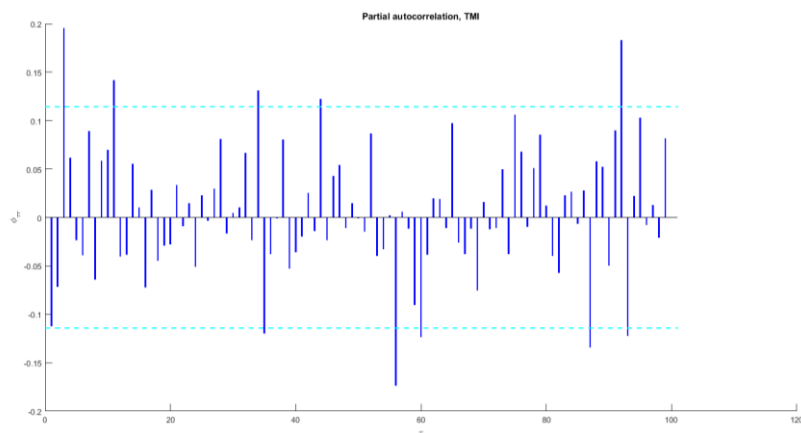
Στο ερώτημα αυτό πραγματοποιούμε γραμμική ανάλυση για τη χρονοσειρά TMI των δεδομένων dat29v2. Ο αντίστοιχος κώδικας βρίσκεται στο αρχείο `grammiki_analysis – TMI.m`.

5.1.1 Γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης

Από το γράφημα αυτοσυσχέτισης της TMI που ακολουθεί παρατηρούμε πως δεν υπάρχουν αργές μεταβολές της αυτοσυσχέτισης. Έτσι, συμπεραίνουμε ότι η χρονοσειρά αυτή δεν παρουσιάζει τάση και προχωράμε κανονικά στη γραμμική ανάλυση αυτής δίχως να πάρουμε τις πρώτες διαφορές. Η αυτοσυσχέτιση υπολογίζεται με τη βοήθεια του δοσμένου κώδικα `autocorrelation.m`.

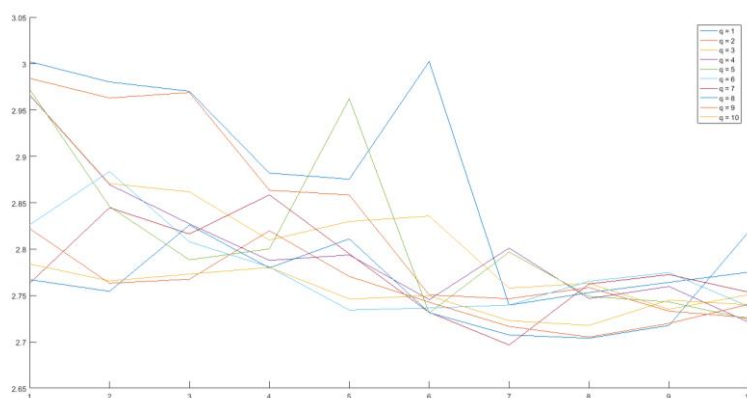


Υπολογίζουμε, λοιπόν, και τη μερική αυτοσυσχέτιση της χρονοσειράς, η οποία φαίνεται στο ακόλουθο γράφημα. Για το σκοπό αυτό έγινε χρήση της συνάρτησης `parautocor.m`.



5.1.2 Διερεύνηση κατάλληλου γραμμικού μοντέλου

Σκοπός του ερωτήματος αυτού είναι να διερευνήσουμε και εν τέλει να προσαρμόσουμε ένα γραμμικό μοντέλο, πιθανόν $ARMA(p,q)$, το οποίο ταιριάζει καλύτερα στη χρονοσειρά μας. Προκειμένου να βρεθεί ένα ικανοποιητικό μοντέλο, χρησιμοποιούμε την μετρική AIC και δημιουργούμε ένα γράφημα αυτής που περιέχει όλους τους συνδυασμούς των τάξεων p και q , με μέγιστες τάξεις $p=q=10$ που απεικονίζεται παρακάτω. Στο γράφημα αυτό, για κάθε μια τιμή της τάξης q υπολογίζεται η μετρική AIC για όλες τις τάξεις p . Από αυτούς τους 100 συνδυασμούς τάξεων που προκύπτουν επιλέγουμε εκείνον το συνδυασμό p και q με το ελάχιστο AIC. Αναφέρεται πως έγινε χρήση του δοσμένου αρχείου `fitARMA.m`.

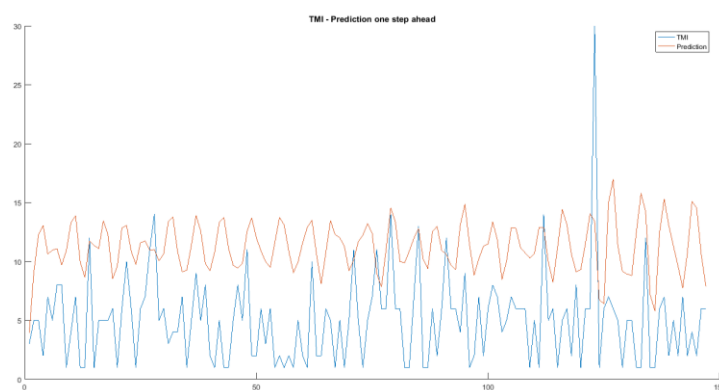


Μέσω κώδικα υπολογίζουμε το βέλτιστο συνδυασμό και καταλήγουμε ότι το μοντέλο που ταιριάζει καλύτερα στη χρονοσειρά μας είναι το $ARMA(7,7)$, πράγμα που φαίνεται και από το γράφημα.

5.1.3 Πρόβλεψη ενός και δύο βημάτων μπροστά

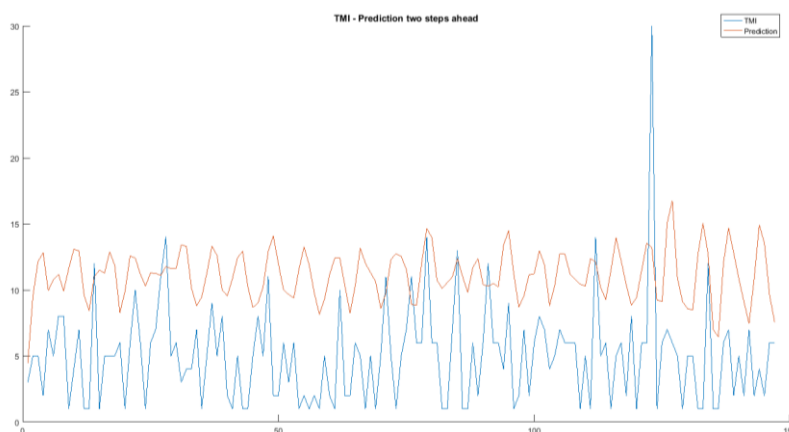
Στο ερώτημα αυτό της εργασίας κάνουμε πρόβλεψη ενός και δύο βημάτων μπροστά επάνω στη χρονοσειρά μας. Επισημαίνεται ότι χρησιμοποιήθηκε το δεύτερο μισό της χρονοσειράς ως test set στο αρχείο `predictARMArmse.m`. Στη συνέχεια, ακολουθούν γραφήματα των προβλέψεων για ένα και για δύο βήματα μπροστά συγκριτικά με τη χρονοσειρά TMI. Επίσης, σημειώνονται και τα αντίστοιχα NRMSEs, με βάση τη συνάρτηση `rmse.m`.

Πρόβλεψη ενός βήματος μπροστά



Παρατηρούμε από το γράφημα ότι η πρόβλεψη δεν είναι καθόλου καλή, πράγμα που επιβεβαιώνεται και από το NRMSE: $\text{nrmseTMI1} = 1.9383$, το οποίο ισούται με 2 ουσιαστικά.

Πρόβλεψη δύο βημάτων μπροστά



Και για τα δύο βήματα μπροστά οι προβλέψεις είναι λάθος. Παρομοίως το NRMSE είναι κοντά στην τιμή 2 και συγκεκριμένα ισούται με $\text{nrmseTMI2} = 1.9035$.

6 Σύγκριση Χρονοσειρών TMI των dat29v1 και dat29v2

Από όλη την παραπάνω ανάλυση καταλήγουμε στο συμπέρασμα ότι το γραμμικό μοντέλο που προσαρμόσαμε και στις δύο χρονοσειρές TMI των δεδομένων dat29v1 και dat29v2 δεν είναι αρκετό ώστε να γίνουν ορθές προβλέψεις. Σίγουρα θα πρέπει να γίνει έλεγχος και για άλλα μοντέλα. Η χρονοσειρά TMI των δεδομένων dat29v2 φαίνεται να παρουσιάζει ελαφρώς πιο ισχυρές συσχετίσεις από αυτή των δεδομένων dat29v1 . Αυτό μπορεί να φανεί από το διάγραμμα της αυτοσυσχέτισής της με τάση, και επομένως και από το διάγραμμα της αυτοσυσχέτισής της χωρίς τάση, εφόσον οι τιμές της αυτοσυσχέτισης για διάφορες υστερήσεις είναι πιο κοντά στο 1 ή στο -1 (ισχυρά θετική ή αρνητική συσχέτιση) συγκριτικά με αυτή των δεδομένων dat29v1 . Τα ίδια ακριβώς συμπεράσματα είχαμε βγάλει και τις εξαρτήσεις της χρονοσειράς AMI των ίδιων δεδομένων dat29v2 .

Όσον αφορά την προσαρμογή των χρονοσειρών σε κάποιο γραμμικό μοντέλο, η χρονοσειρά TMI των δεδομένων dat29v2 έχει μικρότερη τιμή AIC για τις επιλεχθείσες τάξεις p και q του μοντέλου ARMA συγκριτικά με τη χρονοσειρά TMI των δεδομένων dat29v1 , όπως ισχύει ακριβώς και με τις αντίστοιχες χρονοσειρές AMI. Αυτό γίνεται εύκολα εμφανές από τα γραφήματα των μετρικών AIC για κάθε μια από αυτές. Όσο μικρότερη η τιμή του AIC τόσο καλύτερη η προσαρμογή στο γραμμικό μοντέλο.

Τέλος, όσον αφορά την προβλεψιμότητα, αυτή τη φορά η χρονοσειρά TMI των δεδομένων dat29v1 παρουσιάζει καλύτερη συμπεριφορά από αυτήν των δεδομένων

dat29v2. Το γεγονός αυτό μπορούμε να το συμπεράνουμε από τις τιμές του κριτηρίου NRMSE, οι οποίες αναφέρθηκαν αναλυτικά για κάθε μια χρονοσειρά προηγουμένως και οι οποίες πλησιάζουν την τιμή 2 για την TMI των dat29v2. Η προβλεψιμότητα, λοιπόν, είναι το μοναδικό σημείο στο οποίο διαφέρουν τα συμπεράσματά μας όσον αφορά τις χρονοσειρές AMI με τις TMI. Αντίθετα, οι AMI και TMI φαίνεται να εμφανίζουν τις ίδιες ιδιότητες ως προς την ύπαρξη ισχυρών συσχετίσεων και ως προς την προσαρμοστικότητα σε κάποιο γραμμικό μοντέλο.

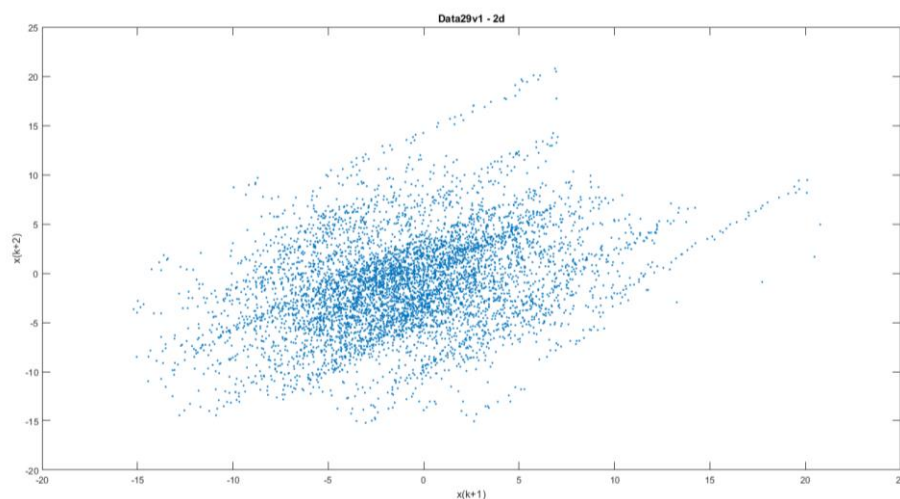
7 Μη γραμμική ανάλυση dat29v1

Στο ερώτημα αυτό της εργασίας διερευνούμε την πιθανή προσαρμογή της OKN χρονοσειράς των δεδομένων dat29v1 σε ένα μη γραμμικό δυναμικό σύστημα, το οποίο πιθανόν να περιέχει και θόρυβο. Η μη γραμμική μας ανάλυση, λοιπόν, αποτελείται από τα ακόλουθα βήματα. Ο κώδικας που αντιστοιχεί στη διαδικασία αυτή βρίσκεται στο αρχείο `mi_grammiki_analysi.m`.

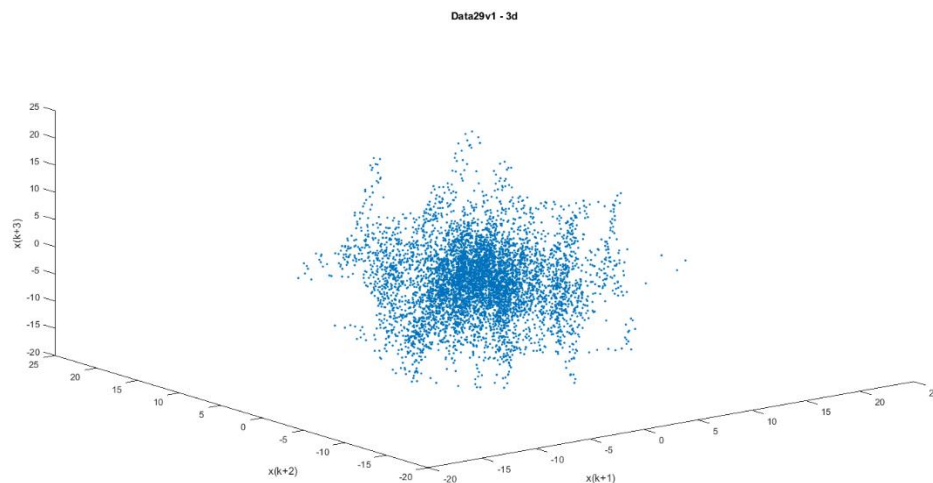
7.1 Διαγράμματα διασποράς στις δύο και στις τρεις διαστάσεις

Παρακάτω παρουσιάζονται τα γραφήματα διασποράς της χρονοσειράς σε δύο και τρεις διαστάσεις. Στο σημείο αυτό χρησιμοποιήσαμε τη δοσμένη συνάρτηση `embeddelays.m`. Από τα γραφήματα αυτά βγάζουμε το συμπέρασμα ότι τα δεδομένα μας δε βρίσκονται σε ένα επίπεδο.

Γράφημα διασποράς 2D

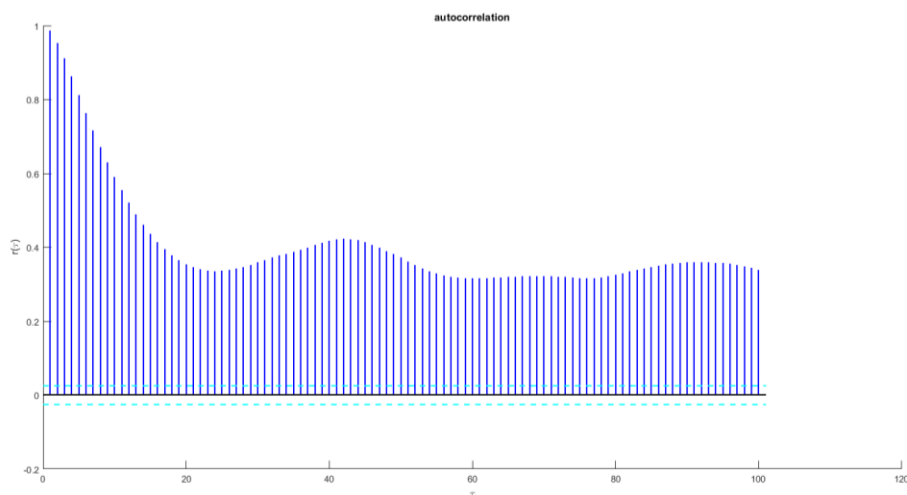


Γράφημα διασποράς 3D



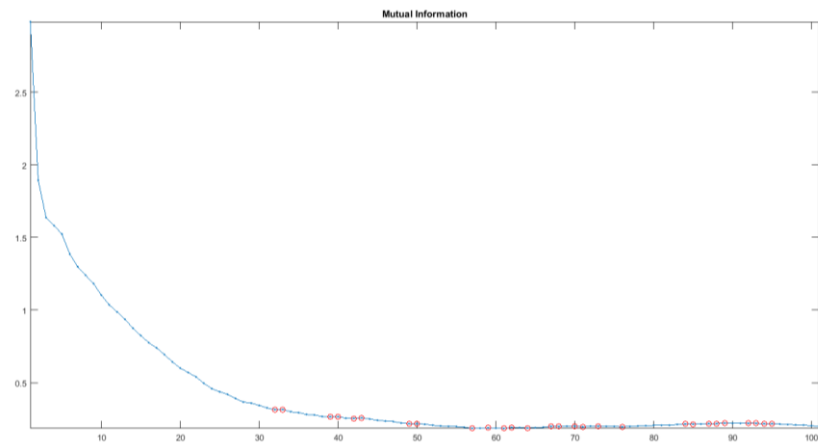
7.2 Κατάλληλες παράμετροι ανακατασκευής χώρου κατάστασης με μέθοδο υστερήσεων

Προκειμένου να κατασκευάσουμε ένα καινούργιο χώρο κατάστασης διάστασης m , πρέπει να προβούμε σε υπολογισμό της διάστασης εμβύθισης m και της υστέρησης τ . Όσον αφορά την υστέρηση τ , προσπαθήσαμε να υπολογίσουμε μια ορθή τιμή για αυτήν μέσω της συνάρτησης αυτοσυσχέτισης των δεδομένων, η οποία παρουσιάζεται παρακάτω. Για το σκοπό αυτό, έγινε χρήση της δοσμένης συνάρτησης `autocorrelation.m`.



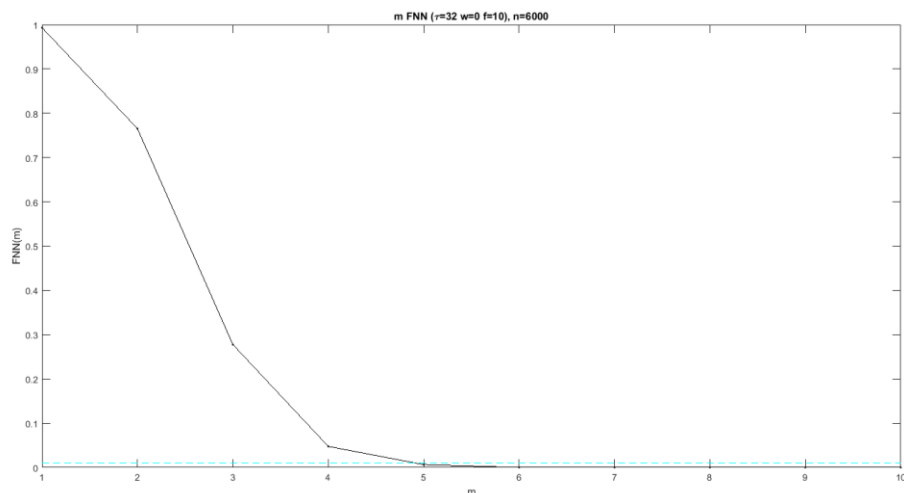
Ωστόσο, όπως είναι εμφανές από το γράφημα αυτό δε μπορούμε να βγάλουμε συμπέρασμα για την τιμή της υστέρησης τ , εφόσον δε συγκλίνει γρήγορα στο μηδέν. Επίσης, δεν υπάρχει τιμή υστέρησης που να αντιστοιχεί σε αυτοσυσχέτιση $1/e = 0,3679$. Για το λόγο αυτό, χρησιμοποιούμε τη συνάρτηση του mutual information και κάνοντας χρήση του αρχείου `mutualinformation.m` παίρνουμε το ακόλουθο γράφημα. Επιπλέον, χρησιμοποιήθηκε και η συνάρτηση `extremes.m` έτσι ώστε να

σημειώσουμε τα τοπικά ακρότατα, από τα οποία εμείς θέλουμε το πρώτο τοπικό ελάχιστο.



Από το παραπάνω διάγραμμα φαίνεται ότι η τιμή της υστέρησης είναι η $\tau = 32$. Παρόλα αυτά, επειδή το αποτέλεσμα είναι «περίεργο» θα κάνουμε μια διερεύνηση για το εάν αυτή η τιμή είναι όντως η κατάλληλη. Ένας ακόμη λόγος που μας ωθεί στη διερεύνηση αυτή είναι το γεγονός ότι γνωρίζουμε πως για χρονοσειρές από απεικονίσεις, δηλαδή από διακριτά συστήματα, ή από συνεχή συστήματα που μετρήθηκαν με μεγάλο τ , συνήθως θέτουμε $\tau=1$.

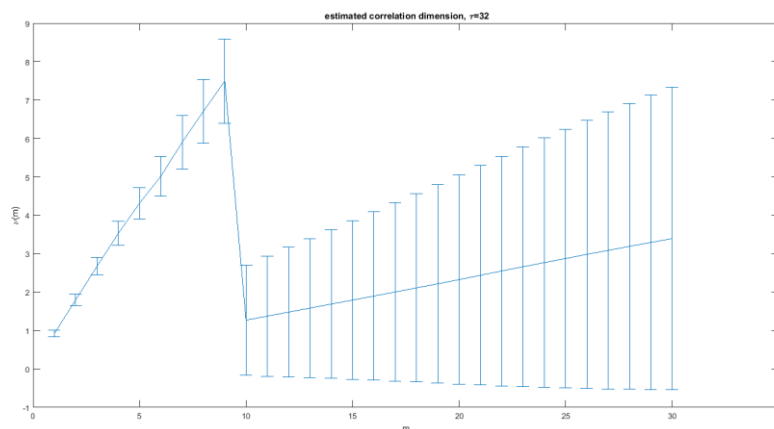
Όσον αφορά τη διάσταση εμβύθισης m για $\tau = 32$, επιλέξαμε να ακολουθήσουμε την πιο συχνά χρησιμοποιούμενη μέθοδο των ψευδών κοντινότερων γειτόνων, για την οποία έγινε χρήση της συνάρτησης `falsenearest.m`.



Από το παραπάνω γράφημα καταλήγουμε ότι η διάσταση εμβύθισης ισούται με $m = 5$, γιατί μετά από την τιμή αυτή το ποσοστό των ψευδών γειτόνων γίνεται επαρκώς μικρό.

7.3 Υπολογισμός διάστασης συσχέτισης

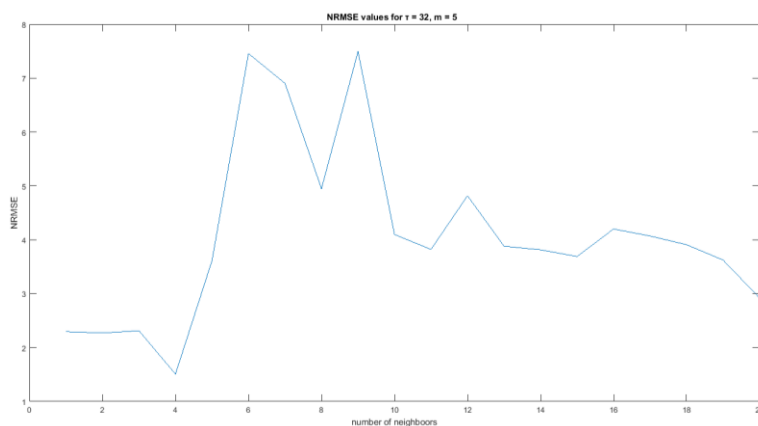
Παρακάτω παρουσιάζεται το γράφημα της διάστασης συσχέτισης για διαφορετικές διαστάσεις εμβύθισης και κατάλληλη παράμετρο υστέρησης $\tau = 32$. Προκειμένου να εξάγουμε το διάγραμμα αυτό κάνουμε χρήση της συνάρτησης `correlationdimension.m`.



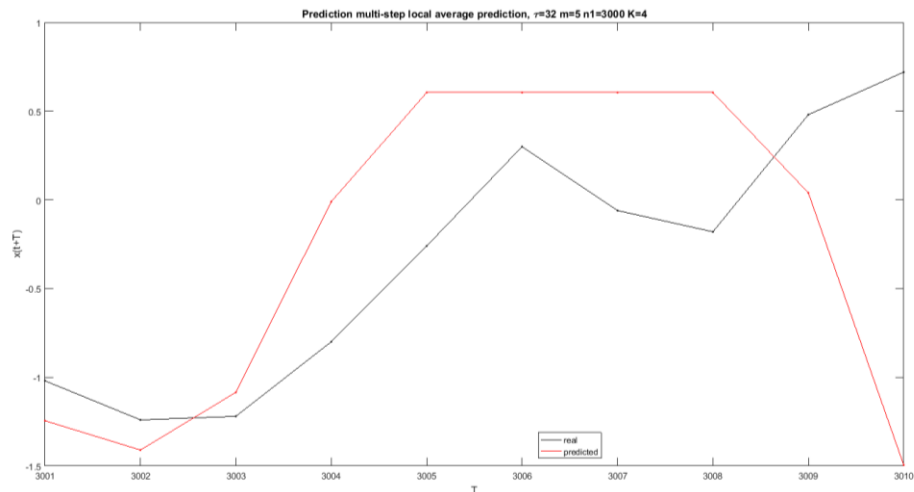
Από το γράφημα αυτό επιβεβαιώνεται η παρατήρηση από τις σημειώσεις του μαθήματος, σύμφωνα με την οποία η διάσταση συσχέτισης επηρεάζεται άμεσα από την επιλογή της υστέρησης τ και της διάστασης εμβύθισης m . Μάλιστα για μεγάλες τιμές τ , όπως στη δική μας περίπτωση όπου $\tau = 32$, βλέπουμε ότι το γράφημα δε συγκλίνει. Οπότε δε μπορούμε να βγάλουμε κάποιο συμπέρασμα για τη διάσταση συσχέτισης ν .

7.4 Διερεύνηση κατάλληλου τοπικού μοντέλου

Στο ερώτημα αυτό, για δεδομένες τιμές της διάστασης εμβύθισης m και της υστέρησης τ , συγκεκριμένα για $\tau = 32$ και $m = 5$ που υπολογίστηκαν νωρίτερα, υπολογίζουμε το πλήθος των κοντινότερων γειτόνων, για τους οποίους το NRMSE είναι ελάχιστο. Εκτελούμε, λοιπόν, μια επανάληψη για πλήθη γειτόνων από 1 μέχρι 20 και κάνουμε χρήση της συνάρτησης `localpredictmultistep.m`. Από τη διαδικασία αυτή εξάγουμε το ακόλουθο διάγραμμα και βρίσκουμε ότι το ελάχιστο NRMSE ισούται με 1,5073, και το οποίο αντιστοιχεί σε 4 κοντινότερους γείτονες.



Επιπλέον, μέσω της συνάρτησης αυτής έχουμε και το ακόλουθο διάγραμμα, το οποίο μας απεικονίζει την πραγματική μας χρονοσειρά αλλά και αυτήν που προβλέφθηκε για 10 βήματα μπροστά. Είναι φανερό πως τα αποτελέσματα της πρόβλεψης δεν είναι καλά, πράγμα που καταδεικνύεται και από την προαναφερθείσα τιμή του NRMSE για 4 κοντινότερους γείτονες, η οποία είναι 1,5073. Τιμές NRMSE γύρω από τη μονάδα ξέρουμε ότι υποδηλώνουν πως η πρόβλεψή μας παρομοιάζεται με πρόβλεψη με βάση τη μέση τιμή της χρονοσειράς.



Να επισημάνουμε στο σημείο αυτό, ότι έγινε επιλογή του μοντέλου τοπικού μέσου όρου (local average model) στη συνάρτηση `localpredictmultistep.m` δηλώνοντας την αντίστοιχη παράμετρο της συνάρτησης με μηδέν. Δοκιμάστηκε, επίσης, και η μέθοδος των προβολών PCR, η οποία έδωσε χειρότερα αποτελέσματα από αυτή του τοπικού μέσου όρου. Για το λόγο αυτό, τα αποτελέσματα αυτά δε συμπεριλήφθηκαν στην αναφορά μας. Τέλος, να αναφέρουμε και το γεγονός ότι η μέθοδος των Ordinary Least Squares δεν έτρεχε στο μηχάνημά μας.

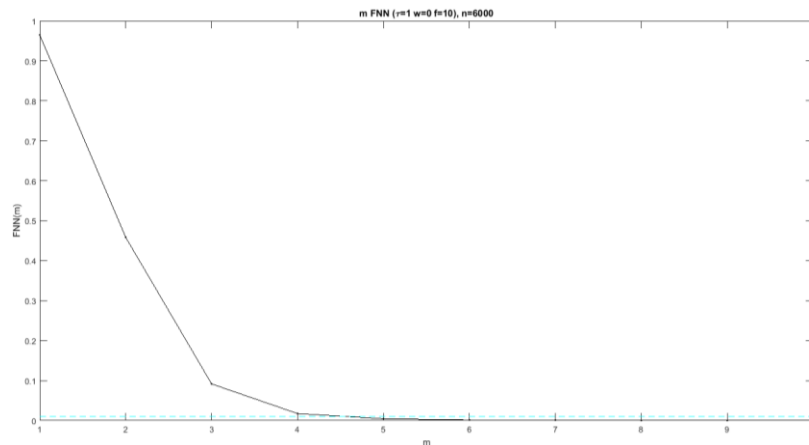
7.5 Διερεύνηση καταλληλότερων παραμέτρων ανακατασκευής χώρου κατάστασης

Για το λόγο που αναφέρθηκε και νωρίτερα, το γεγονός δηλαδή ότι η υστέρηση $\tau = 32$ είναι πολύ μεγάλη και λόγω του ότι τα δεδομένα μας είναι συνεχή, προχωράμε σε επιλογή υστέρησης $\tau = 1$ και εκτελούμε όλα τα προηγούμενα βήματα μη γραμμικής ανάλυσης, προκειμένου να ελέγξουμε εάν τα αποτελέσματα με αυτήν την επιλογή θα είναι καλύτερα ή χειρότερα.

7.5.1 Υπολογισμός διάστασης εμβύθισης για υστέρηση $\tau = 1$

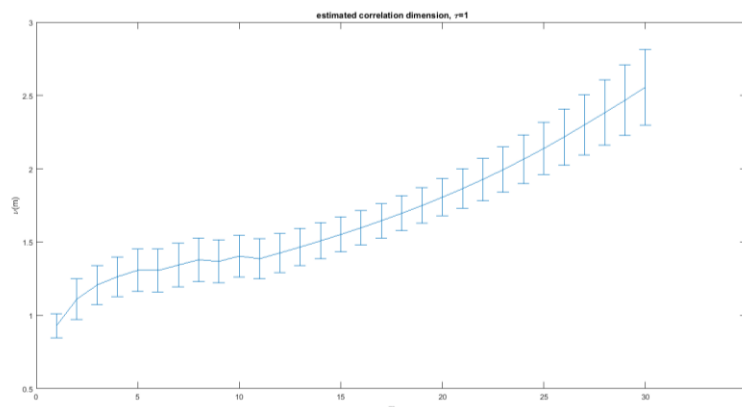
Κατά τα γνωστά, για τον υπολογισμό της διάστασης εμβύθισης ακολουθούμε τη μέθοδο των ψευδών κοντινότερων γειτόνων και χρησιμοποιούμε τη συνάρτηση `falsenearest.m`. Το γράφημα που λαμβάνουμε είναι το ακόλουθο, το οποίο μας

υποδεικνύει ότι η διάσταση εμβύθισης ισούται με $m = 5$, γιατί μετά από την τιμή αυτή το ποσοστό των ψευδών γειτόνων γίνεται επαρκώς μικρό.



7.5.2 Υπολογισμός διάστασης συσχέτισης για $\tau = 1$

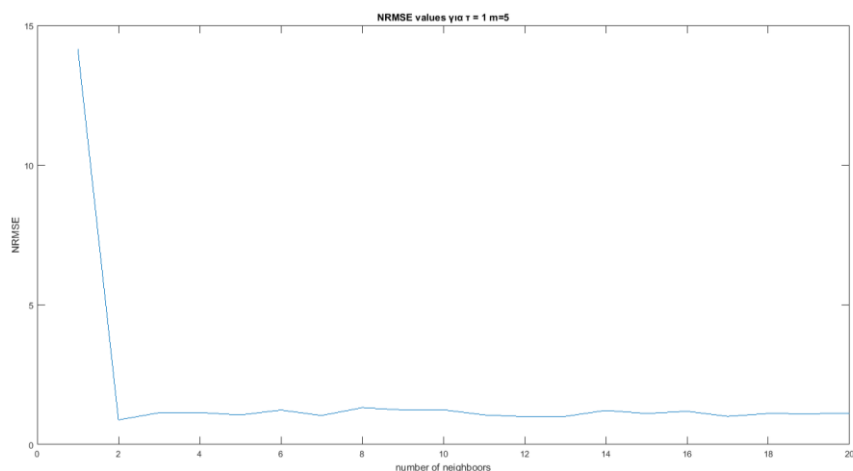
Παρακάτω παρουσιάζεται το γράφημα της διάστασης συσχέτισης για διαφορετικές διαστάσεις εμβύθισης και κατάλληλη παράμετρο υστέρησης $\tau = 1$. Προκειμένου να εξάγουμε το διάγραμμα αυτό κάνουμε χρήση της συνάρτησης `correlationdimension.m`.



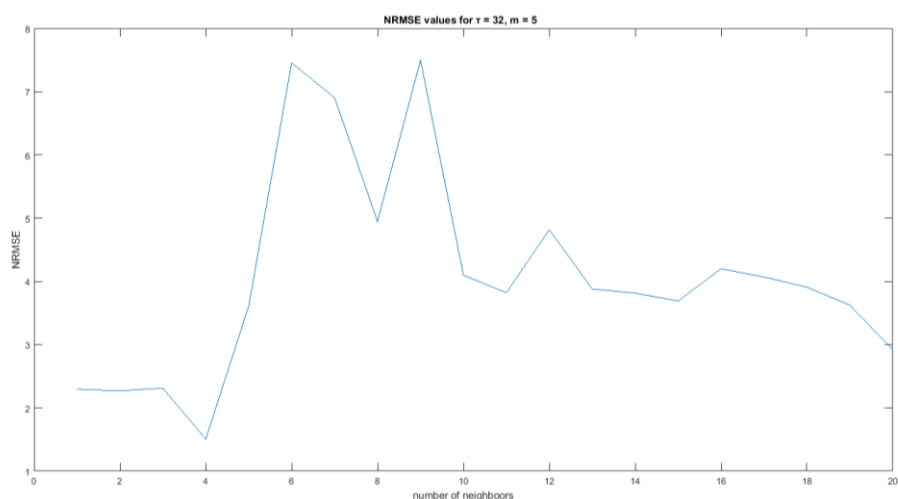
Και πάλι, δεδομένης της νέας μας επιλογής $\tau = 1$, το γράφημα της διάστασης συσχέτισης δε συγκλίνει σε κάποια τιμή.

7.5.3 Διερεύνηση κατάλληλου τοπικού μοντέλου για $\tau = 1$

Στο ερώτημα αυτό, για δεδομένες τιμές της διάστασης εμβύθισης m και της υστέρησης τ , συγκεκριμένα για $\tau = 1$ και $m = 5$ που υπολογίστηκαν νωρίτερα, υπολογίζουμε το πλήθος των κοντινότερων γειτόνων, για τους οποίους το NRMSE είναι ελάχιστο. Εκτελούμε, λοιπόν, μια επανάληψη για πλήθη γειτόνων από 1 μέχρι 20 και κάνουμε χρήση της συνάρτησης `localpredictmultistep.m`. Από τη διαδικασία αυτή εξάγουμε το ακόλουθο διάγραμμα και βρίσκουμε ότι το ελάχιστο NRMSE ισούται με 0,8808, και το οποίο αντιστοιχεί σε 2 κοντινότερους γείτονες.



Επιπλέον, μέσω της συνάρτησης αυτής έχουμε και το ακόλουθο διάγραμμα, το οποίο μας απεικονίζει την πραγματική μας χρονοσειρά αλλά και αυτήν που προβλέφθηκε για 10 βήματα μπροστά. Είναι φανερό πως τα αποτελέσματα της πρόβλεψης σαφώς καλύτερα από το προηγούμενα με υστέρηση $\tau = 32$, πράγμα που καταδεικνύεται και από την προαναφερθείσα τιμή του NRMSE για 2 κοντινότερους γείτονες, η οποία είναι 0,8808. Συνεπώς, η επιλογή μας να δοκιμάσουμε τιμή υστέρησης $\tau = 1$ ήταν σωστή, δίνοντάς μας καλύτερα αποτελέσματα.

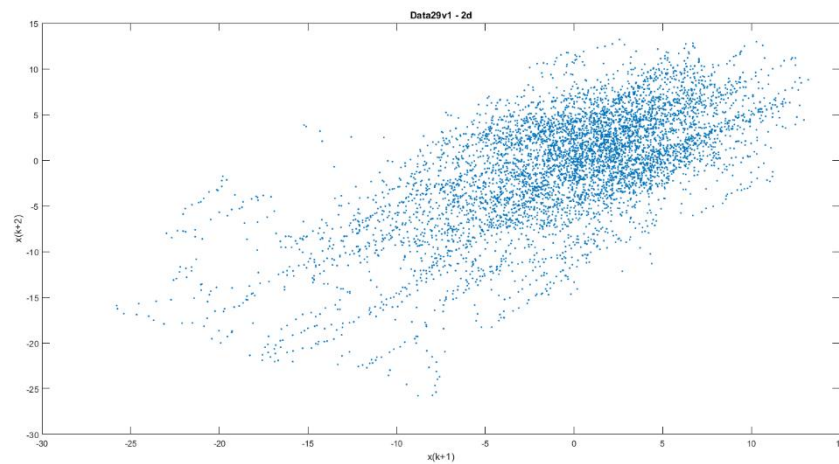


8 Μη γραμμική ανάλυση dat29v2

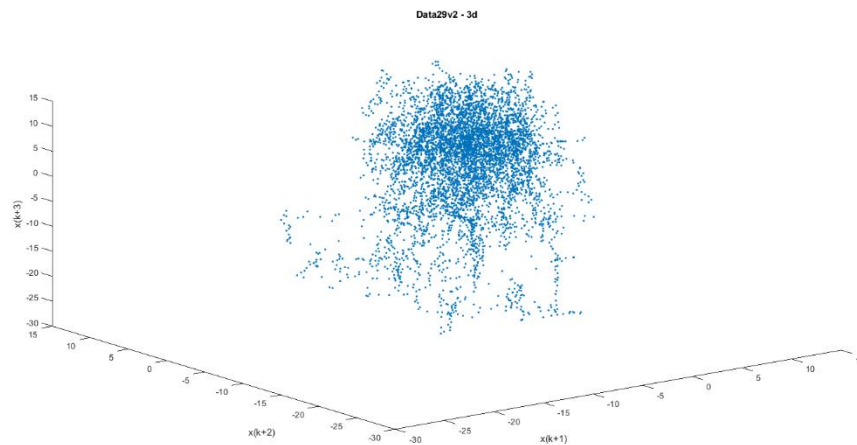
Στο σημείο αυτό, επαναλαμβάνουμε την προαναφερθείσα μη γραμμική ανάλυση για την OKN χρονοσειρά των δεδομένων dat29v2. Τα βήματα που ακολουθούμε είναι τα ακόλουθα. Ο κώδικας που αντιστοιχεί στη διαδικασία αυτή βρίσκεται στο αρχείο `mi_grammiki_analysi.m`.

Παρακάτω παρουσιάζονται τα γραφήματα διασποράς της χρονοσειράς σε δύο και τρεις διαστάσεις. Στο σημείο αυτό χρησιμοποιήσαμε τη δοσμένη συνάρτηση `embeddelays.m`. Από τα γραφήματα αυτά βγάζουμε το συμπέρασμα ότι τα δεδομένα μας ακόμη μια φορά δε βρίσκονται σε ένα επίπεδο.

Γράφημα διασποράς 2D

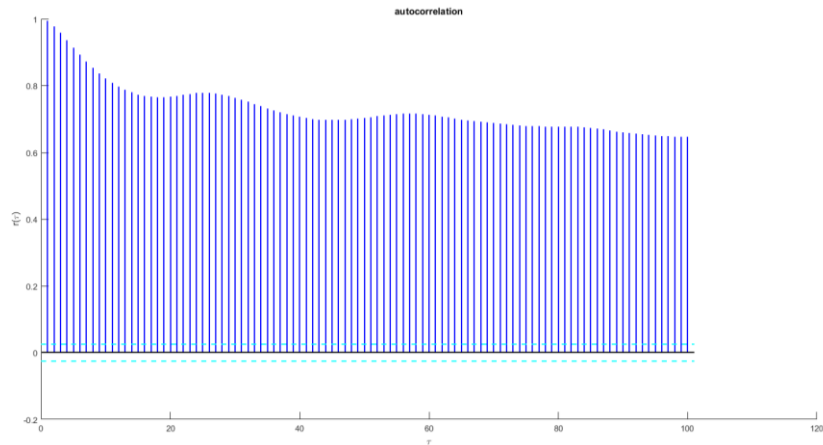


Γράφημα διασποράς 3D

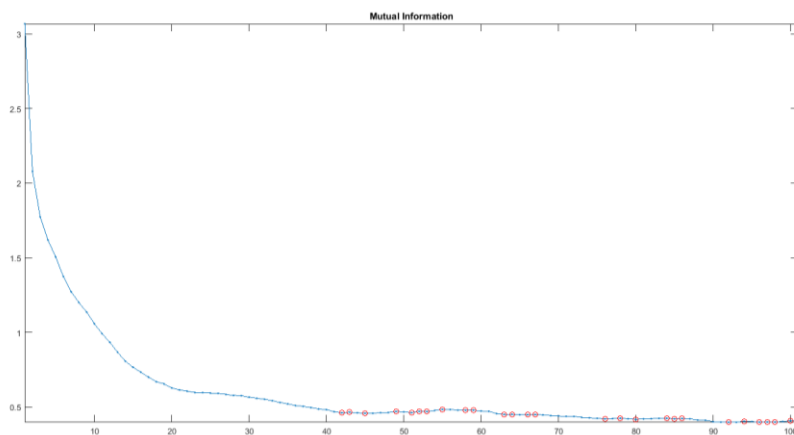


8.1 Κατάλληλες παράμετροι ανακατασκευής χώρου κατάστασης με μέθοδο υστερήσεων

Προκειμένου να κατασκευάσουμε ένα καινούργιο χώρο κατάστασης διάστασης m , πρέπει να προβούμε σε υπολογισμό της διάστασης εμβύθισης m και της υστέρησης τ . Όσον αφορά την υστέρηση τ , προσπαθήσαμε να υπολογίσουμε μια ορθή τιμή για αυτήν μέσω της συνάρτησης αυτοσυσχέτισης των δεδομένων, η οποία παρουσιάζεται παρακάτω. Για το σκοπό αυτό, έγινε χρήση της δοσμένης συνάρτησης `autocorrelation.m`.

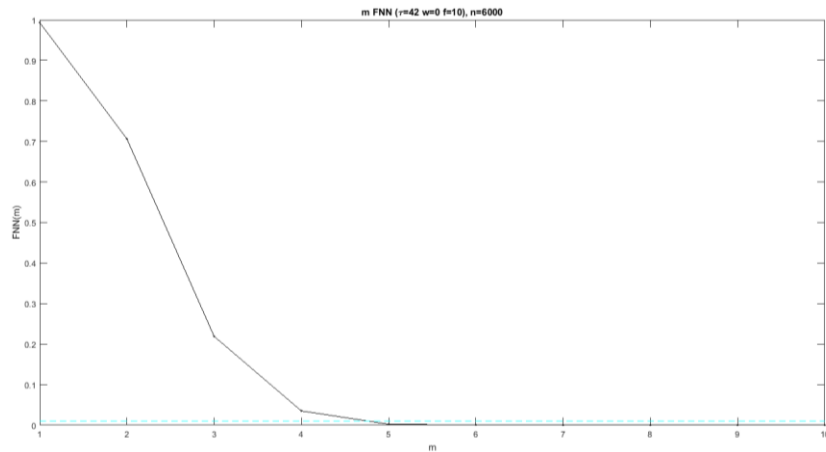


Ωστόσο, όπως είναι εμφανές από το γράφημα αυτό δε μπορούμε να βγάλουμε συμπέρασμα για την τιμή της υστέρησης τ , εφόσον δε συγκλίνει γρήγορα στο μηδέν. Επίσης, δεν υπάρχει τιμή υστέρησης που να αντιστοιχεί σε αυτοσυσχέτιση $1/e = 0,3679$. Για το λόγο αυτό, χρησιμοποιούμε τη συνάρτηση του mutual information και κάνοντας χρήση του αρχείου `mutualinformation.m` παίρνουμε το ακόλουθο γράφημα. Επιπλέον, χρησιμοποιήθηκε και η συνάρτηση `extremes.m` έτσι ώστε να σημειώσουμε τα τοπικά ακρότατα, από τα οποία εμείς θέλουμε το πρώτο τοπικό ελάχιστο.



Από το παραπάνω διάγραμμα φαίνεται ότι η τιμή της υστέρησης είναι η $\tau = 42$. Παρόλα αυτά, επειδή το αποτέλεσμα είναι «περίεργο» θα κάνουμε μια διερεύνηση για το εάν αυτή η τιμή είναι όντως η κατάλληλη. Ένας ακόμη λόγος που μας ωθεί στη διερεύνηση αυτή είναι το γεγονός ότι γνωρίζουμε πως για χρονοσειρές από απεικονίσεις, δηλαδή από διακριτά συστήματα, ή από συνεχή συστήματα που μετρήθηκαν με μεγάλο τ_s , συνήθως θέτουμε $\tau=1$.

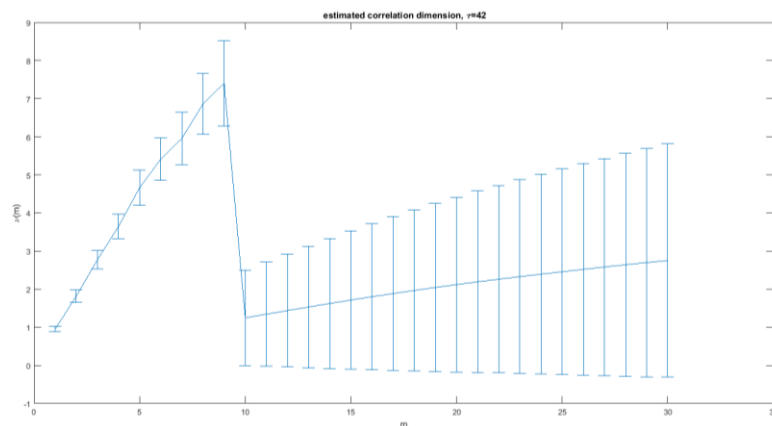
Όσον αφορά τη διάσταση εμβύθισης m για $\tau = 42$, επιλέξαμε να ακολουθήσουμε την πιο συχνά χρησιμοποιούμενη μέθοδο των ψευδών κοντινότερων γειτόνων, για την οποία έγινε χρήση της συνάρτησης `falsenearest.m`.



Από το παραπάνω γράφημα καταλήγουμε ότι η διάσταση εμβύθισης ισούται με $m = 5$, γιατί μετά από την τιμή αυτή το ποσοστό των ψευδών γειτόνων γίνεται επαρκώς μικρό.

8.2 Υπολογισμός διάστασης συσχέτισης

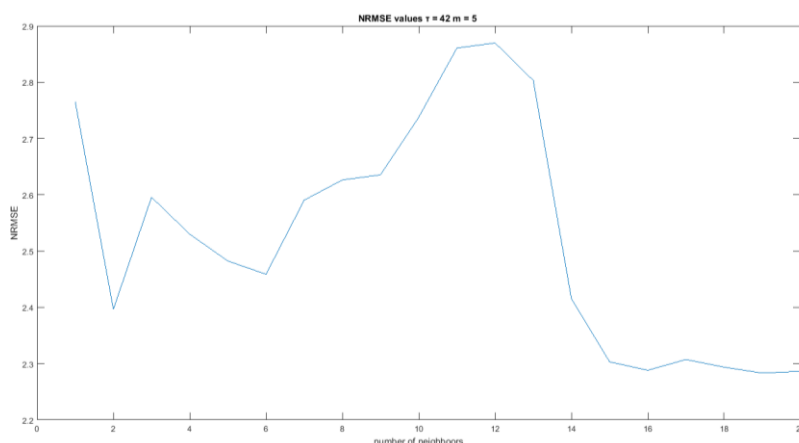
Παρακάτω παρουσιάζεται το γράφημα της διάστασης συσχέτισης για διαφορετικές διαστάσεις εμβύθισης και κατάλληλη παράμετρο υστέρησης $\tau = 42$. Προκειμένου να εξάγουμε το διάγραμμα αυτό κάνουμε χρήση της συνάρτησης `correlationdimension.m`.



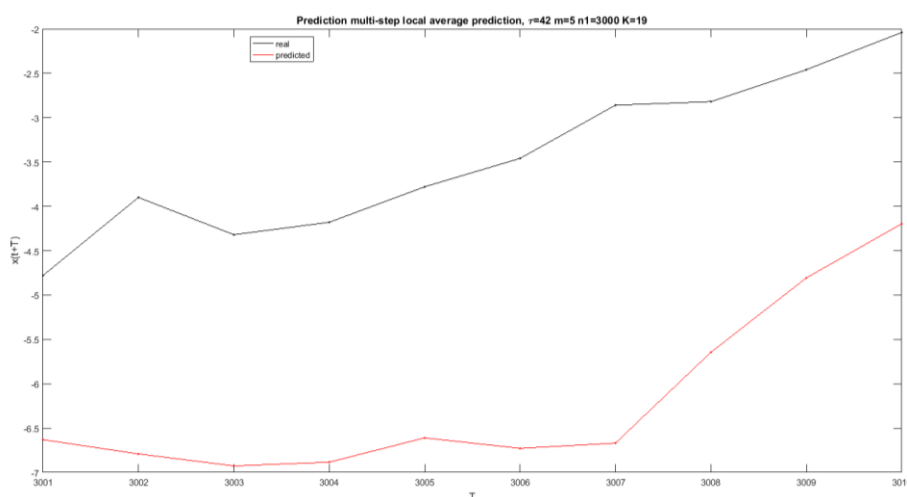
Από το γράφημα αυτό επιβεβαιώνεται η παρατήρηση από τις σημειώσεις του μαθήματος, σύμφωνα με την οποία η διάσταση συσχέτισης επηρεάζεται άμεσα από την επιλογή της υστέρησης τ και της διάστασης εμβύθισης m . Μάλιστα για μεγάλες τιμές τ , όπως στη δική μας περίπτωση όπου $\tau = 42$, βλέπουμε ότι το γράφημα δε συγκλίνει. Οπότε δε μπορούμε να βγάλουμε κάποιο συμπέρασμα για τη διάσταση συσχέτισης ν .

8.3 Διερεύνηση κατάλληλου τοπικού μοντέλου

Στο ερώτημα αυτό, για δεδομένες τιμές της διάστασης εμπύθισης m και της υστέρησης τ , συγκεκριμένα για $\tau = 42$ και $m = 5$ που υπολογίστηκαν νωρίτερα, υπολογίζουμε το πλήθος των κοντινότερων γειτόνων, για τους οποίους το NRMSE είναι ελάχιστο. Εκτελούμε, λοιπόν, μια επανάληψη για πλήθη γειτόνων από 1 μέχρι 20 και κάνουμε χρήση της συνάρτησης `localpredictmultistep.m`. Από τη διαδικασία αυτή εξάγουμε το ακόλουθο διάγραμμα και βρίσκουμε ότι το ελάχιστο NRMSE ισούται με 2,2832, και το οποίο αντιστοιχεί σε 19 κοντινότερους γείτονες.



Επιπλέον, λαμβάνουμε και το ακόλουθο διάγραμμα από το οποίο μπορούμε οπτικά να δούμε την πραγματική μας χρονοσειρά, αλλά και τη χρονοσειρά που προβλέψαμε, και να κρίνουμε κατά πόσο η πρόβλεψή μας ήταν καλή. Είναι εμφανές ότι η πρόβλεψη δεν είναι καθόλου καλή, πράγμα που το διαπιστώνει κανείς εύκολα και από την προαναφερθείσα τιμή του NRMSE η οποία ξεπερνά την τιμή 2.



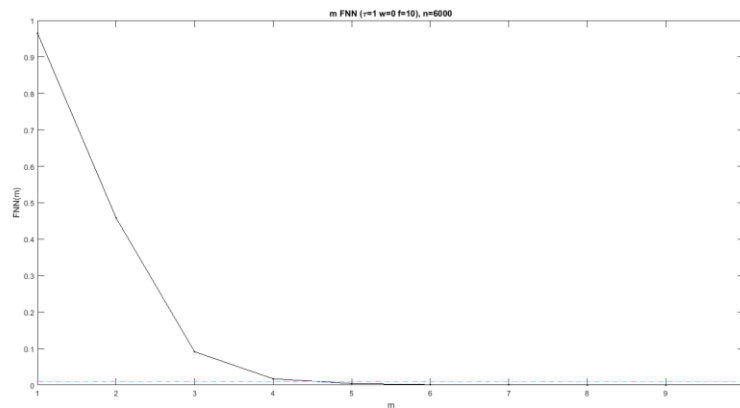
Εύλογα επομένως, προχωράμε προς τη διερεύνηση περισσότερο αντιπροσωπευτικών παραμέτρων ανακατασκευής του χώρου κατάστασης. Πιο συγκεκριμένα, διερευνούμε την τιμή υστέρησης $\tau = 1$.

8.4 Διερεύνηση καταλληλότερων παραμέτρων ανακατασκευής χώρου κατάστασης

Για το λόγο που αναφέρθηκε και νωρίτερα, το γεγονός δηλαδή ότι η υστέρηση $\tau = 42$ είναι πολύ μεγάλη και λόγω του ότι τα δεδομένα μας είναι συνεχή, προχωράμε σε επιλογή υστέρησης $\tau = 1$ και εκτελούμε όλα τα προηγούμενα βήματα μη γραμμικής ανάλυσης, προκειμένου να ελέγξουμε εάν τα αποτελέσματα με αυτήν την επιλογή θα είναι καλύτερα ή χειρότερα.

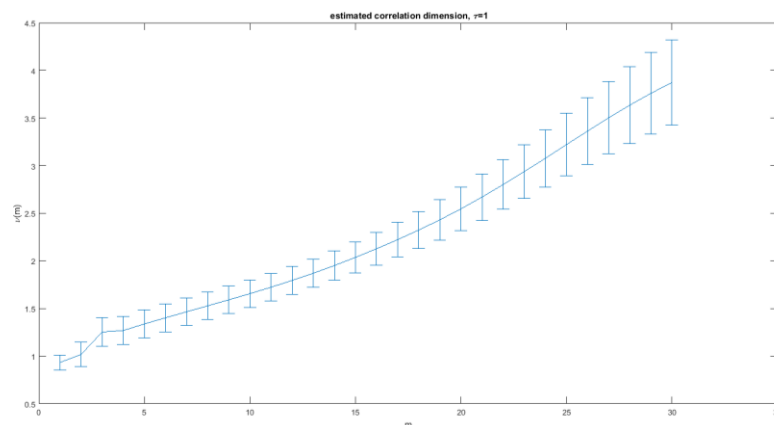
8.4.1 Υπολογισμός διάστασης εμβύθισης για υστέρηση $\tau = 1$

Κατά τα γνωστά, για τον υπολογισμό της διάστασης εμβύθισης ακολουθούμε τη μέθοδο των ψευδών κοντινότερων γειτόνων και χρησιμοποιούμε τη συνάρτηση `falsenearest.m`. Το γράφημα που λαμβάνουμε είναι το ακόλουθο, το οποίο μας υποδεικνύει ότι η διάσταση εμβύθισης ισούται με $m = 5$, γιατί μετά από την τιμή αυτή το ποσοστό των ψευδών γειτόνων γίνεται επαρκώς μικρό.



8.4.2 Υπολογισμός διάστασης συσχέτισης για $\tau = 1$

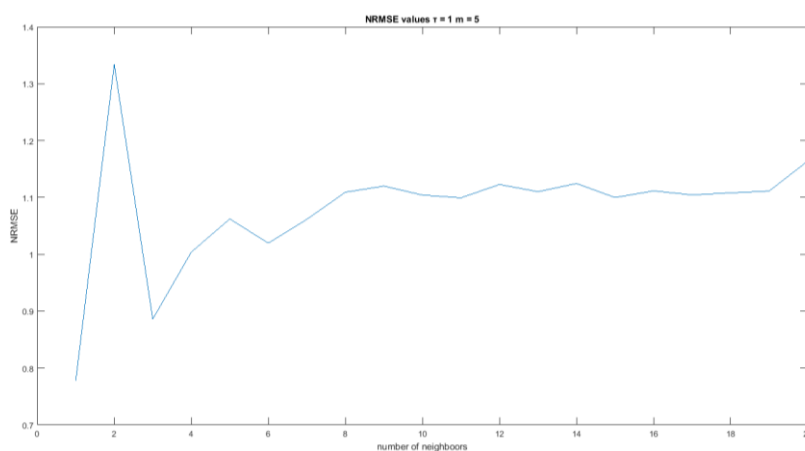
Παρακάτω παρουσιάζεται το γράφημα της διάστασης συσχέτισης για διαφορετικές διαστάσεις εμβύθισης και κατάλληλη παράμετρο υστέρησης $\tau = 1$. Προκειμένου να εξάγουμε το διάγραμμα αυτό κάνουμε χρήση της συνάρτησης `correlationdimension.m`.



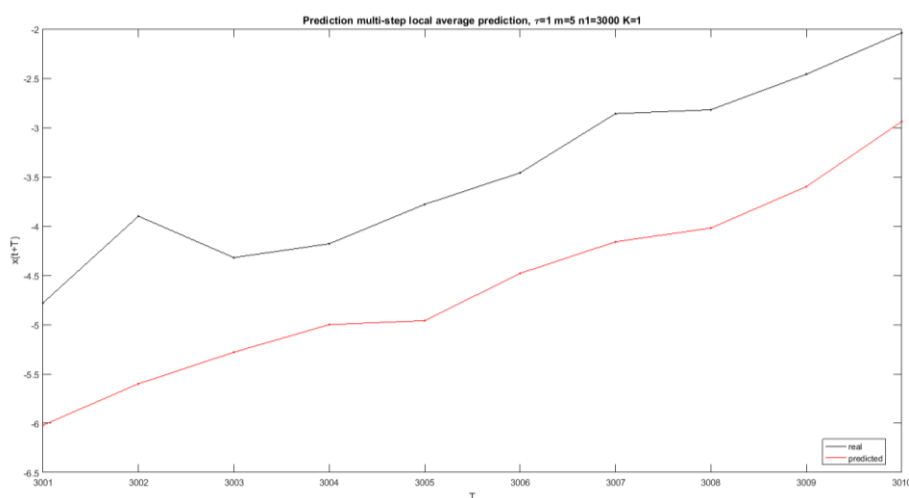
Παρόλο που κάναμε αλλαγή στην τιμή της υστέρησης $\tau = 1$, δε φαίνεται να υπάρχει κάποια βελτίωση ως προς τη σύγκλιση σε κάποια τιμή διάστασης συσχέτισης.

8.4.3 Διερεύνηση κατάλληλου τοπικού μοντέλου για $\tau = 1$

Στο ερώτημα αυτό, για δεδομένες τιμές της διάστασης εμβύθισης m και της υστέρησης τ , συγκεκριμένα για $\tau = 1$ και $m = 5$ που υπολογίστηκαν νωρίτερα, υπολογίζουμε το πλήθος των κοντινότερων γειτόνων, για τους οποίους το NRMSE είναι ελάχιστο. Εκτελούμε, λοιπόν, μια επανάληψη για πλήθη γειτόνων από 1 μέχρι 20 και κάνουμε χρήση της συνάρτησης `localpredictmultistep.m`. Από τη διαδικασία αυτή εξαγάγουμε το ακόλουθο διάγραμμα και βρίσκουμε ότι το ελάχιστο NRMSE ισούται με 0,7776, και το οποίο αντιστοιχεί σε 1 κοντινότερο γείτονα.



Όπως και προηγουμένως, παραθέτουμε το διάγραμμα όπου συγκρίνονται η πραγματική μας χρονοσειρά αλλά και αυτή την οποία προβλέψαμε. Συγκριτικά με το προηγούμενο διάγραμμα που αφορούσε τις τιμές $\tau = 42$ και $m = 5$, το παρόν είναι αρκετά καλύτερο, καθώς φαίνονται οι δύο καμπύλες να πλησιάζουν καλύτερα η μια την άλλη. Παρόλα αυτά, δε μπορούμε να πούμε κιόλας ότι η πρόβλεψη είναι καλή γενικά.



9 Σύγκριση μη γραμμικής ανάλυσης OKN χρονοσειρών dat29v1 και dat29v2

Από όλη την παραπάνω ανάλυση καταλήγουμε στο συμπέρασμα ότι το μη γραμμικό μοντέλο που προσαρμόστηκε αρχικά και στις δύο χρονοσειρές των δεδομένων dat29v1 και dat29v2 δεν είναι αρκετό ώστε να γίνουν ορθές προβλέψεις. Για το λόγο αυτό τροποποιήθηκε η τιμή της υστέρησης τ , ώστε να οδηγηθούμε σε καλύτερα αποτελέσματα. Η σύγκριση συνεπώς των δύο χρονοσειρών θα γίνει για τιμή υστέρησης $\tau = 1$ και στις δύο περιπτώσεις.

Όσον αφορά την προσαρμογή και την προβλεψιμότητα των χρονοσειρών σε κάποιο τοπικό μοντέλο, αυτή έγινε όπως αναφέρθηκε και νωρίτερα με το τοπικό μοντέλο μέσου όρου. Για τις επιλεχθείσες τιμές των τ , m ο βέλτιστος αριθμός των γειτόνων που προέκυψε ήταν 1 και για τις δύο χρονοσειρές. Με βάση, λοιπόν, τις τιμές του κριτηρίου NRMSE που προκύπτουν, η χρονοσειρά των δεδομένων dat29v1, παρουσιάζει ελαφρώς καλύτερη συμπεριφορά από αυτήν των δεδομένων dat29v2, καθώς έχει μικρότερη ελάχιστη τιμή NRMSE.

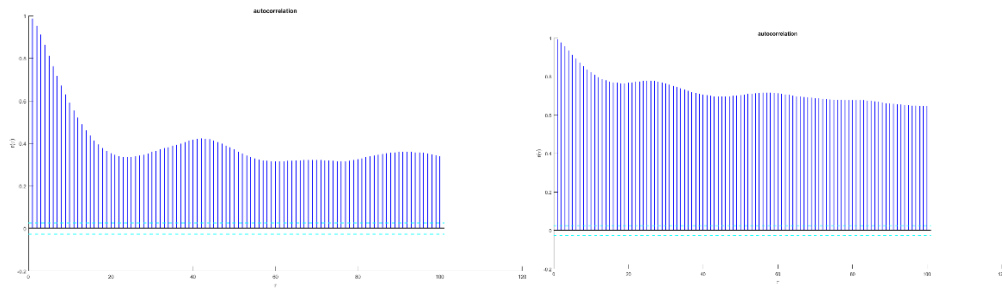
Στο σημείο αυτό, αξίζει να αναφέρουμε ακόμη πως καθώς αυξάνεται το πλήθος των γειτόνων, οι τιμές του κριτηρίου NRMSE για τη χρονοσειρά των δεδομένων dat29v1 συγκλίνουν στην τιμή 1. Σε αντίθεση, οι αντίστοιχες τιμές για τη χρονοσειρά των δεδομένων dat29v2 παρουσιάζουν μια συνεχή αύξηση, καθώς το πλήθος των γειτόνων μεγαλώνει. Αυτό μας οδηγεί στο συμπέρασμα πως η χρονοσειρά των δεδομένων dat29v1 θα μπορεί να προβλέπεται και για μεγαλύτερο πλήθος γειτόνων βάσει της μέσης τιμής, ενώ δεν ισχύει το ίδιο και για την άλλη χρονοσειρά.

10 Συμπεράσματα γραμμικής και μη γραμμικής ανάλυσης

Προηγουμένως συγκρίθηκαν δύο παραγόμενες χρονοσειρές από τις αρχικές που δόθηκαν, αφού πρώτα έγινε γραμμική ανάλυση. Όπως ήδη αναφέρθηκε, η χρονοσειρά TMI αποδίδει καλύτερα, καθώς οι τιμές του κριτηρίου NRMSE είναι καλύτερες από ότι η AMI. Ωστόσο, όπως είναι φανερό, στη μη γραμμική ανάλυση παρουσιάστηκαν και για τις δύο χρονοσειρές ακόμη χαμηλότερες τιμές για το κριτήριο του NRMSE από ότι στη γραμμική. Το γεγονός αυτό μας οδηγεί στο συμπέρασμα πως οι δοθείσες χρονοσειρές μπορούν να προσαρμοστούν και να προβλεφθούν καλύτερα όταν τις αναλύουμε με μη γραμμικό τρόπο.

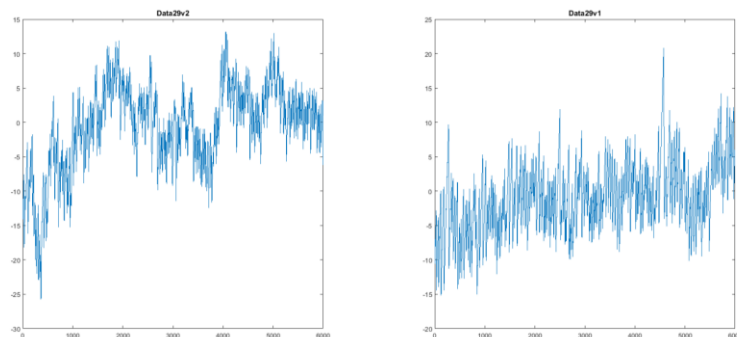
Παρατηρώντας τις δύο χρονοσειρές, data29v1 και data29v2 και δεδομένου ότι η μέθοδος που χρησιμοποιήθηκε κατά το πείραμα είναι αξιόπιστη ως προς το θόρυβο, θα μπορούσαμε να πούμε πως οι δύο χρονοσειρές δεν ανήκουν στην ίδια κατηγορία. Πιο αναλυτικά, από την απεικόνιση τους βλέπουμε πως η data29v2 παίρνει πιο γρήγορα ακραίες τιμές, ενώ σε αντίθεση η data29v1 φτάνει πιο ομαλά στις ακραίες τιμές της.

Όσον αφορά την αιτιοκρατικότητα ή τη στοχαστικότητα των χρονοσειρών dat29v1 και dat29v2, εξάγουμε το συμπέρασμα ότι πρόκειται για αιτιοκρατικές χρονοσειρές. Αυτό μπορούμε να το δούμε από τα διαγράμματα αυτοσυσχετίσεων για κάθε μία από αυτές. Το αριστερό διάγραμμα αναφέρεται στα δεδομένα dat29v1 και το δεξί στα dat29v2. Παρατηρούμε ότι οι συσχετίσεις ενός σημείου με το επόμενο και το προηγούμενό του είναι σημαντικές, επομένως δε φαίνεται να υπάρχει στοχαστικότητα.



Επίσης, μπορούμε να βασιστούμε ακόμη και στο ίδιο το paper του ακόλουθου συνδέσμου <https://www.sciencedirect.com/science/article/pii/S0010480997914415> στο οποίο αναφέρεται ότι οι χρονοσειρές αυτές φαίνεται να παράγονται από μια ντετερμινιστική, χαμηλής διάστασης και πιθανώς χαοτική γεννήτρια. Συνεπώς, και ως προς τη γραμμικότητα και μη γραμμικότητα μπορούμε να πούμε ότι οι χρονοσειρές μας προσδιορίζονται από μη γραμμικά δυναμικά συστήματα, εφόσον οι τιμές του κριτηρίου NRMSE είναι πολύ χαμηλότερες από τις τιμές των χρονοσειρών που εξετάστηκαν στο κομμάτι της γραμμικής ανάλυσης της εργασίας. Η διάσταση των χρονοσειρών είναι χαμηλή εφόσον το παρατηρούμενο μέγεθος είναι μόνο ένα, καθώς και η πολυπλοκότητά τους είναι χαμηλή, διότι τα μοντέλα που προσαρμόσαμε είχαν μικρή υστέρηση, διάσταση εμβύθισης και πλήθος γειτόνων.

Τέλος, τα δεδομένα dat29v1 και dat29v2 αποτελούν χρονοσειρές από την εξέταση δύο διαφορετικών ατόμων, από τα οποία το ένα είναι υγιές και το άλλο ασθενής. Το γεγονός αυτό μπορούμε να το αποφανθούμε εύκολα παρατηρώντας και συγκρίνοντας τη μορφή των δύο χρονοσειρών, η οποία φαίνεται να είναι αρκετά διαφορετική, όπως παρακάτω.



Διαφορετικά, θα μπορούσαμε να είχαμε χρησιμοποιήσει το κριτήριο που αναφέρεται στο paper, το οποίο επισημαίνει ότι η μέση τιμή της διάστασης συσχέτισης ενός ασθενούς είναι μικρότερη από αυτήν ενός υγιούς ανθρώπου.