NATIONAL TECHNICAL UNIVERSITY OF THENS

SCHOOL OF CIVIL ENGINEERING

DEPARTMENT OF WATER RESOURCES AND ENVIRONMENTAL ENGINEERING

"Stochastic investigation of short-term predictability of basic renewable energy resources with application on the non-connected island of Astypalea"

Diploma Thesis

Elli Klousakou

Supervisor: Koutsoyiannis Demetris, Professor, NTUA

Athens, 2020

«Στοχαστική διερεύνηση της βραχυπρόθεσμης προγνωσιμότητας βασικών πηγών ανανεώσιμης ενέργειας με εφαρμογή στο μη συνδεδεμένο νησί της Αστυπάλαιας»

Διπλωματική Εργασία

Έλλη Κλουσάκου

Επιβλέπων: Δημήτρης Κουτσογιάννης, Καθηγητής ΕΜΠ

Αθήνα, 2020

For my nephew

Ευχαριστίες

Η παρούσα Διπλωματική Εργασία σχεδιάστηκε, εξελίχθηκε και ολοκληρώθηκε κατά το Ακαδημαϊκό Έτος 2019-2020 σε συνεργασία με τον Τομέα Υδατικών Πόρων και Περιβάλλοντος της σχολής Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου.

Στο σημείο αυτό θα ήθελα να απευθύνω ένα μεγάλο ευχαριστώ από καρδιάς στον επιβλέποντα καθηγητή κο Δημήτρη Κουτσογιάννη που μου εμπιστεύτηκε το θέμα και ήταν πολύτιμος αρωγός καθ' όλη τη διάρκεια της εκπόνησής του, καθώς και για τις ευκαιρίες που μου έδωσε να εκπροσωπήσω το εργαστήριο και την ερευνητική ομάδα «ΙΤΙΑ» στο συνέδριο «European Geosciences Union" της Βιέννης κατά τα έτη 2018-2019. Ευχαριστώ ειλικρινά και θερμά τον μεταδιδάκτορα Παναγιώτη Δημητριάδη για την αποτελεσματική, συνεχή και αμέριστη βοήθειά του κατά την υλοποίηση του εν λόγω Διπλωματικού πονήματος. Θα ήθελα επίσης να ευχαριστήσω όλο το team του Τομέα Υδατικών Πόρων και Περιβάλλοντος της σχολής Πολιτικών Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου που όλο αυτό το διάστημα μου εξασφάλισε πρακτικά, ψυχολογικά και ηθικά άριστες συνθήκες για τη διεκπεραίωση της εργασίας μου.

Ένα μεγάλο ευχαριστώ οφείλω και τις πολύ καλές μου φίλες Κωνσταντίνα Μ., Εύα Β., Χαρά Β. και Κωνσταντίνα Α. για την υποστήριξή τους καθ' όλη τη διάρκεια εκπόνησης της εργασίας καθώς επίσης και για τι υπέροχες στιγμές που μοιραστήκαμε παρέα όλα αυτά τα χρόνια που είμαστε μαζί στη σχολή.

Τίποτε από όλα αυτά δεν θα είχε συμβεί χωρίς τους γονείς Κατερίνα και Δημήτρη που τους ευχαριστώ θερμά για όλη τη στήριξη και την υπομονή που έδειξαν, και κυρίως τις αδερφές μου Λίνα και Ιωάννα αλλά και το φίλο μου Νίκο που δεν σταμάτησαν λεπτό να πιστεύουν στους στόχους μου και να με ενθαρρύνουν συνεχώς να προχωρήσω.

Acknowledgments

The current diploma thesis was structured, developed and completed during the academic year 2019-2020 in cooperation with the Department of Water Resources Development and Environmental Engineering of the School of Civil Engineering of the National Technical University of Athens.

At this point I would like to sincerely thank professor Demetris Koutsoyiannis who entrusted me with this topic and for his precious contribution throughout the process. I am also very thankful to him for giving me the chance to participate twice in the" European Geosciences Union" representing the research team of " ITIA". I sincerely and wholeheartedly thank the post-doctor Panayiotis Dimitriadis for his effective, continuous and full support in the implementation of this Diplomatic Thesis. I would also like to thank the entire team of the Department of Water Resources and Environmental Engineering of School of Civil Engineering of the National Technical University of Athens for providing me with practical, psychological and ethical excellence throughout my work.

I am also very thankful for my very good friends Constantina M., Eva B., Hara B. and Constantina A. for their support throughout the preparation of this work as well as for the wonderful moments we shared together during our academic years.

None of this would have happened if it weren't for my parents, Katerina and Dimitris, whom I also thank for all the support and patience they showed during this process, and especially my sisters Lina and Ioanna and my friend Nikos who never stopped believing in my goals and they constantly encouraged me to keep going.

Εκτεταμένη Περίληψη (Extended abstract in Greek)

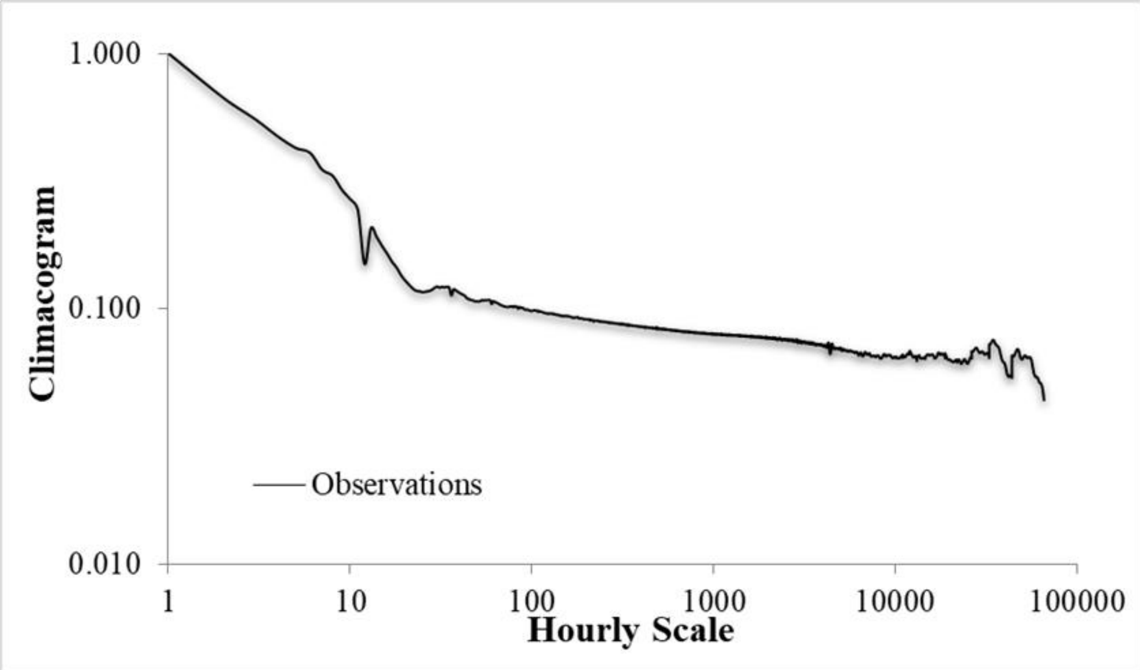
Πρόσφατες έρευνες αναφέρουν πως η ζήτηση της ενέργειας παγκοσμίως θα αυξηθεί αισθητά μέσα στις επόμενες δεκαετίες. Η αύξηση του πληθυσμού σε συνδυασμό με τη συνεχή ανάπτυξη πολλών χωρών καθώς και την ασταθή κοινωνικοπολιτική κατάσταση πολλών κρατών (κυρίως σε πιο υποανάπτυκτες χώρες) προμηνύουν τη διαρκή αύξηση των ενεργειακών αναγκών που έχουν ως αποτέλεσμα την υπερεκμετάλλευση των ορυκτών πόρων. Συνεπώς, εναλλακτικές μορφές ενέργειας αρχίζουν να υποκαθιστούν σταδιακά τους μέχρι τώρα διαδεδομένους πόρους και τρόπους παραγωγής ενέργειας.

Οι ανανεώσιμες πηγές ενέργειας (ΑΠΕ) χαρακτηρίζονται από μία πληθώρα πλεονεκτημάτων που τις καθιστούν αρκετά ελκυστικές πλέον. Ειδικά η ηλιακή και η αιολική ενέργεια φαίνεται πως κερδίζουν χρόνο με το χρόνο όλο και περισσότερο έδαφος στην παραγωγή της ενέργειας αφού δεν εκπέμπουν ρύπους στην ατμόσφαιρα, πηγάζουν από ανεξάντλητες πηγές και λόγω της ραγδαίας διάδοσής τους μειώνεται διαρκώς το κόστος παραγωγής τους. Εφόσον λοιπόν παρατηρείται αυτή η στροφή προς πιο οικολογικές μορφές παραγωγής ενέργειας, η διαχείριση την ενεργειακών συστημάτων είναι πλέον επιτακτική ανάγκη, συμπεριλαμβανομένων και όλων των αντίστοιχων μονάδων. Μέσω της βελτίωσης του σχεδιασμού αλλά και της λειτουργίας τον μονάδων ενέργειας θα είναι εφικτό να υπάρχει συνεχής, αξιόπιστη και οικονομικά πιο βέλτιστη ενέργεια.

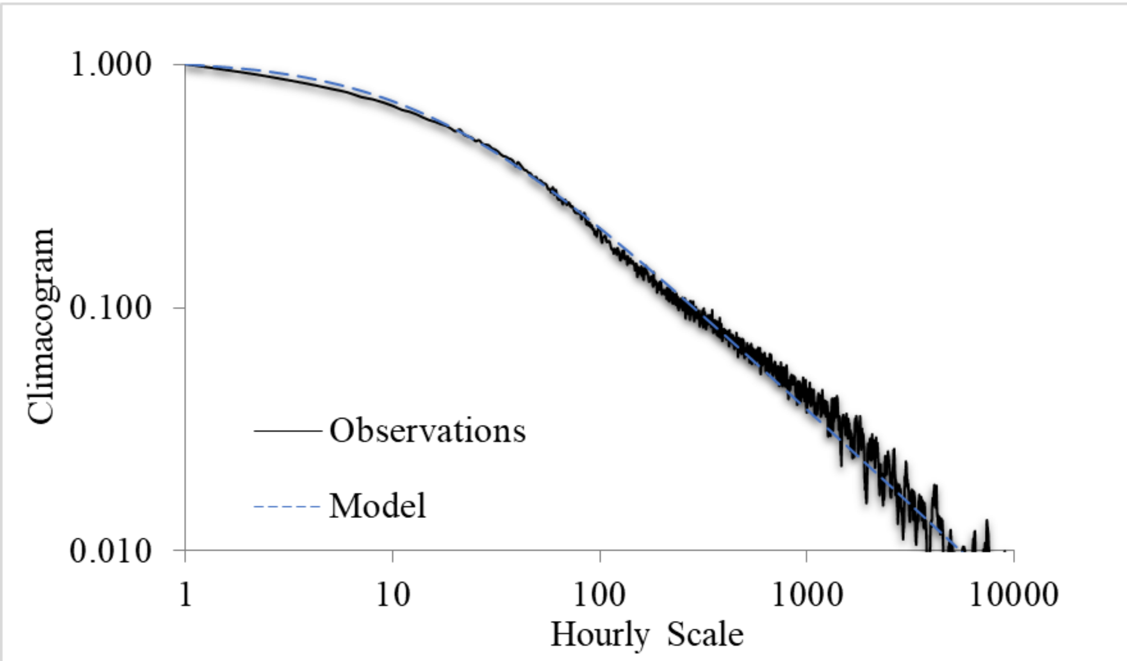
Παρά την αναγκαιότητα βελτιστοποίησης της διαδικασίας παραγωγής, τα ενεργειακά συστήματα μπορούν να αποβούν αρκετά περίπλοκα αφού αποτελούν πολυπαραμετρικά προβλήματα. Οι εναλλακτικές μορφές ενέργειας, συμπεριλαμβανομένου της αιολικής και της ηλιακής που θεωρούνται από τις πιο αξιοποιήσιμες, πηγάζουν από φυσικές υδρομετεωρολογικές και γεωφυσικές διεργασίες που είναι αρκετά απρόβλεπτες και εμφανίζουν μη ομαλή μεταβλητότητα. Συγκεκριμένα, όλες οι φυσικές διεργασίες που εμπλέκονται στην παραγωγή ενέργειας αναμένονται να παρουσιάσουν το φαινόμενο της μακροπρόθεσμης εμμονής (HK behavior). Αυτό σημαίνει πως η αβεβαιότητα εμφανίζεται στη φυσική διεργασία τόσο μακροπρόθεσμα όσο και σε μικρότερες χρονικές κλίμακες.

Αυτή η εγγενής αβεβαιότητα των στοιχείων της φύσης μεταφέρεται και στα αντίστοιχα ενεργειακά συστήματα γεγονός που οδηγεί στο συμπέρασμα πως επιβαρύνεται ακόμα περισσότερο η παραγωγή ενέργειας μέσω ανανεώσιμων πηγών. Μία αρκετά ισχυρή μέθοδος για την ποσοτικοποίηση της αβεβαιότητας στις διάφορες γεωφυσικές διεργασίες είναι η εφαρμογή του κλιμακογράμματος (η συνδιασπορά συναρτήσει της συναθροισμένης κλίμακας σε λογαριθμισμένους άξονες). Συγκριτικά με άλλες μεθόδους που μπορούν να διαχειριστούν την αβεβαιότητα (όπως το φάσμα ισχύος και η αυτοσυνδιασπορά), το κλιμακόγραμμα είναι αρκετά απλό στη εφαρμογή του και εμφανίζει ο μικρότερο στατιστικό σφάλμα. Πρόκειται για ένα αξιόπιστο στοχαστικό εργαλείο μέσω του οποίου εκτιμάται η παράμετρος Hurst που ουσιαστικά υπολογίζει το βαθμό της μη προβλεψιμότητας σε διάφορες διεργασίες- φυσικές και μη.

Μετά την διερεύνηση της συμπεριφοράς διάφορων φυσικών διεργασιών από τις οποίες μπορεί να παραχθεί ανανεώσιμη ενέργεια τα αποτελέσματα έδειξαν πως στις περισσότερες περιπτώσεις ο δείκτης Hurst H ήταν μεγαλύτερος του 0.5 αποδεικνύοντας την ύπαρξη της συμπεριφοράς της μακροπρόθεσμης εμμονής. Αυτό δείχνει πως δεν ακολουθούν ούτε ανέλιξη λευκού θορύβου ούτε μαρκοβιανή συμπεριφορά.



Σχήμα 1. *Κλιμακόγραμμα ηλιακής ακτινοβολίας.*



Σχήμα 2. *Κλιμακόγραμμα ταχύτητας ανέμου.*

Παρότι στην κοινή γνώμη επικρατεί η άποψη πως η αβεβαιότητα και ο ντετερμινισμός στη φύση δεν μπορούν να συνυπάρξουν κάτι τέτοιο τελικά δεν ισχύει και αποδεικνύεται πλέον και από πρόσφατες έρευνες. Στην πραγματικότητα όσο πιο ντετερμινιστική είναι μία διεργασία τόσο πιο διευρυμένο είναι και το παράθυρο προβλεψιμότητας. Το παράθυρο προβλεψιμότητας είναι ένα χρονικό διάστημα κατά το οποίο οι προγνώσεις που μπορούν να παραχθούν είναι μεγάλης ακρίβειας βάσει όμως ενός καθορισμένου σφάλματος.

Στην παρούσα έρευνα, για τον υπολογισμό του χρονικού περιθωρίου προγνωσιμότητας εφαρμόστηκε ένας στοχαστικός-αναλογικός αλγόριθμος και ένα μοντέλο «αναφοράς» (Benchmark model B2) σε διεργασίες ταχύτητας ανέμου και ηλιακής ακτινοβολίας, που βρέθηκαν από παγκόσμιες βάσεις δεδομένων, μιας και η αιολική και η ηλιακή ενέργεια αποτελούν τις πλέον αξιοποιήσιμες μορφές ανανεώσιμων ενεργειών. Για την εκτίμηση του παραθύρου προγνωσιμότητας χρησιμοποιήθηκε ο δείκτης απόδοσης Nash and Sutcliff *F*. Όσο ο δείκτης *F* λαμβάνει θετικές τιμές, οι προβλέψεις που υπολογίζονται από τους αλγορίθμους πρόγνωσης θεωρούνται μεγάλης αξιοπιστίας. Τη χρονική στιγμή που η πρόγνωση θα ισούται με τη μέση τιμή της χρονοσειράς, ο δείκτης *F* μηδενίζεται, οριοθετώντας έτσι το παράθυρο προβλεψιμότητας μιας διεργασίας˙ όταν $F<0$ τότε το μοντέλο είναι άχρηστο.

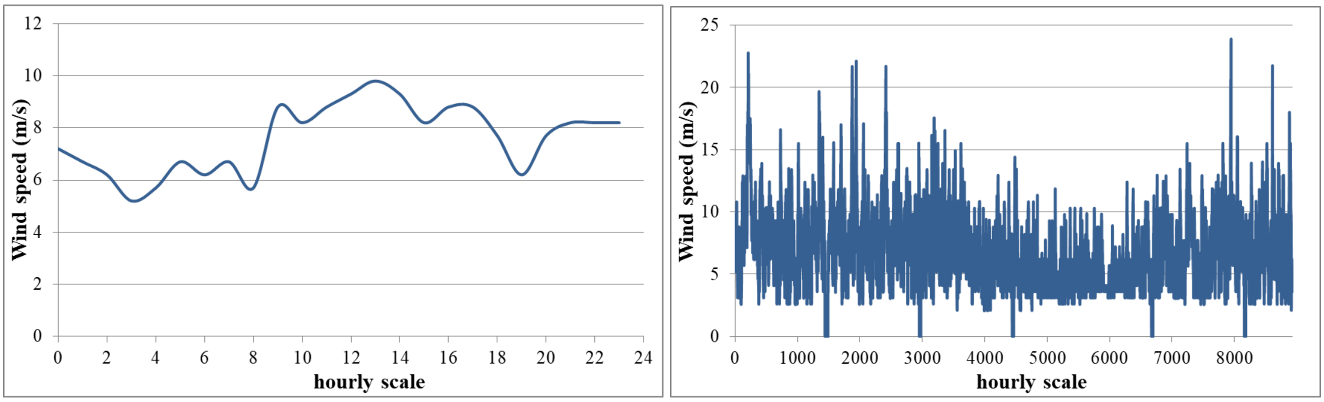
Το μοντέλο αναφοράς B2 που εφαρμόστηκε θεωρεί πως η πρόγνωση μίας τιμής ισούται με την προηγούμενη τιμή της ανεξάρτητα από τον χρονικό ορίζοντα που εξετάζεται. Ουσιαστικά το μοντέλο αυτό εφαρμόστηκε ώστε να ορίσει ένα ελάχιστο χρονικό περιθώριο κατά το οποίο οι προβλέψεις θα χαρακτηρίζονται αξιόπιστες και χρησιμοποιήθηκε ως σημείο αναφοράς για σύγκριση με τα αποτελέσματα του αναλογικού αλγορίθμου.

Το αναλογικό μοντέλο που χρησιμοποιήθηκε στη συνέχεια, είναι ένα αρκετά απλό μοντέλο αφού δεν στηρίζεται σε καμία μαθηματική σχέση. Η παραδοχή που ακολουθεί αυτό το μοντέλο είναι πως εξετάζοντας το παρελθόν μίας χρονοσειράς μπορούν να εντοπιστούν κάποια επαναλαμβανόμενα μοτίβα μέσα από τα οποία μπορεί να εκτιμηθεί η πρόγνωσης της εκάστοτε τιμής της διεργασίας. Για την εφαρμογή του αναλογικού μοντέλου προσδιορίστηκαν 3 παράμετροι:

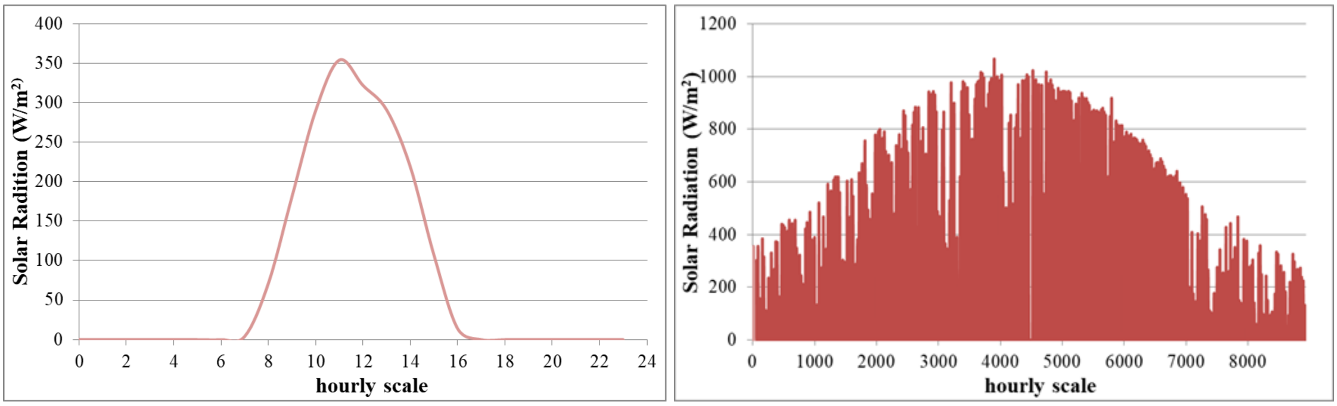
- το ποσοστό *p*= είναι το ποσοστό της χρονοσειράς στο οποίο θα γίνει η επαλήθευση της πρόβλεψης
- η παράμετρος h= το πλήθος των γειτόνων στους οποίους διερευνάται τιμή παρόμοια με αυτή που θέλουμε να προβλεφθεί
- το σφάλμα *g*= το σφάλμα που εκτιμάται για τον προσδιορισμό της πρόβλεψης.

Η εκτίμηση του σφάλματος έγινε μέσω στοχικής συνάρτησης ενός συντηρητικού-πεσιμιστικού σεναρίου και ενός πιο ευέλικτου σεναρίου μέσου σφάλματος για τη διερεύνηση του παραθύρου προβλεψιμότητας.

Κατά τη διερεύνηση της βραχυπρόθεσμης προγνωσιμότητας στις χρονοσειρές της ταχύτητας του ανέμου και της ηλιακής ακτινοβολίας λήφθηκε υπόψιν η διπλή κυκλοστασιμότητα, ημερίσια και εποχική, που εμφανίζουν οι διεργασίες εξαιτίας της διπλή περιοδικότητας της Γης. Γι αυτό το λόγω αποφασίστηκε να γίνει κανονικοποίηση των χρονοσειρών για να περιοριστεί η επίδραση της διπλή κυκλοστασιμότητας. Αρχικά, υπολογίστηκε η μέση τιμή και η τυπική απόκλιση της κάθε χρονοσειράς βάσει της ώρας και του μήνα και μετά οι εξεταζόμενες τιμές διαιρέθηκαν με την τυπική απόκλιση αφού αφαιρέθηκε η αντίστοιχη μέση τιμή.

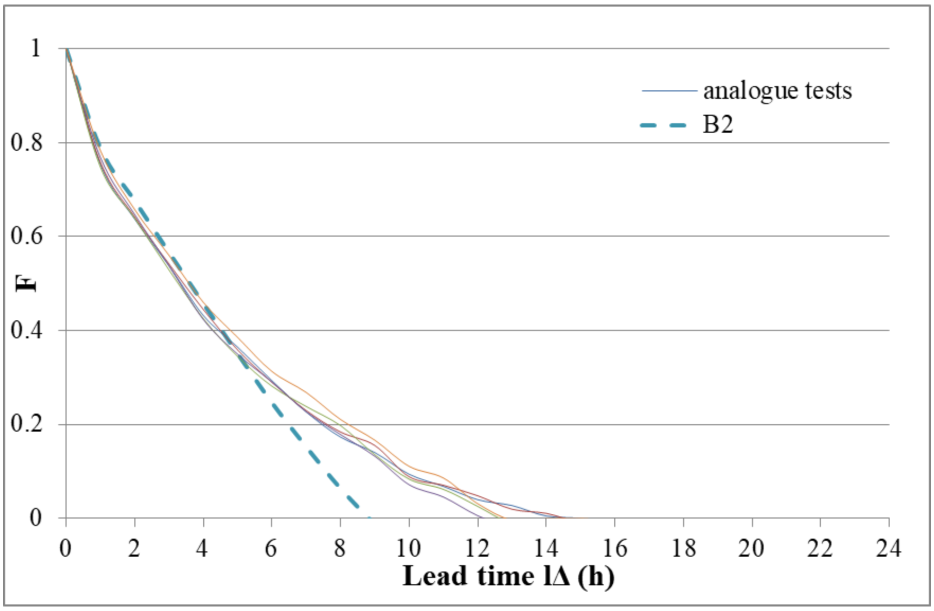


Σχήμα 3. *Απεικόνιση της ημερήσιας (αριστερά) και της εποχικής (δεξιά) περιοδικότητας του ανέμου.*

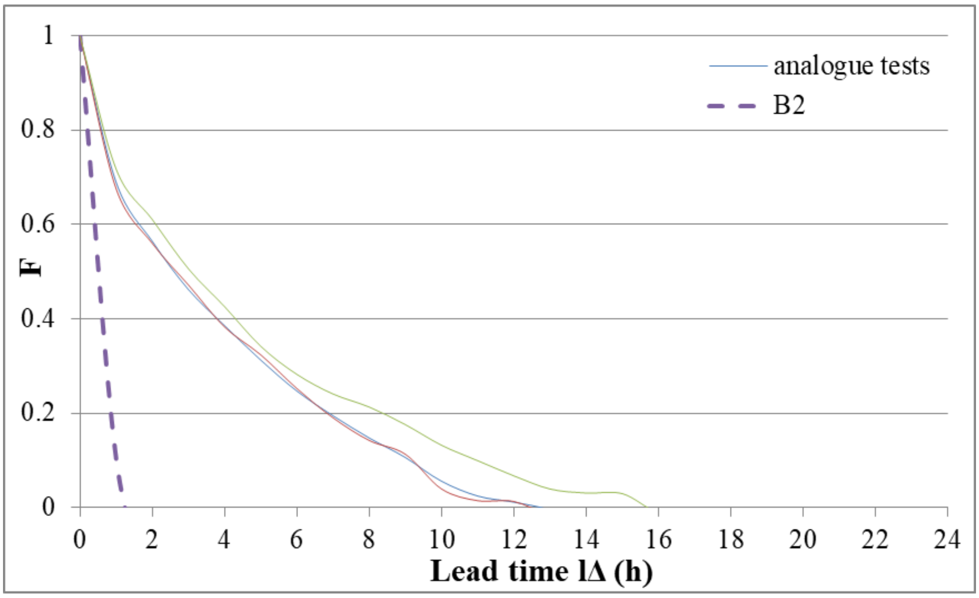


Σχήμα 4. *Απεικόνιση της ημερήσιας (αριστερά) και της εποχικής (δεξιά) περιοδικότητας της ηλιακής ακτινοβολίας.*

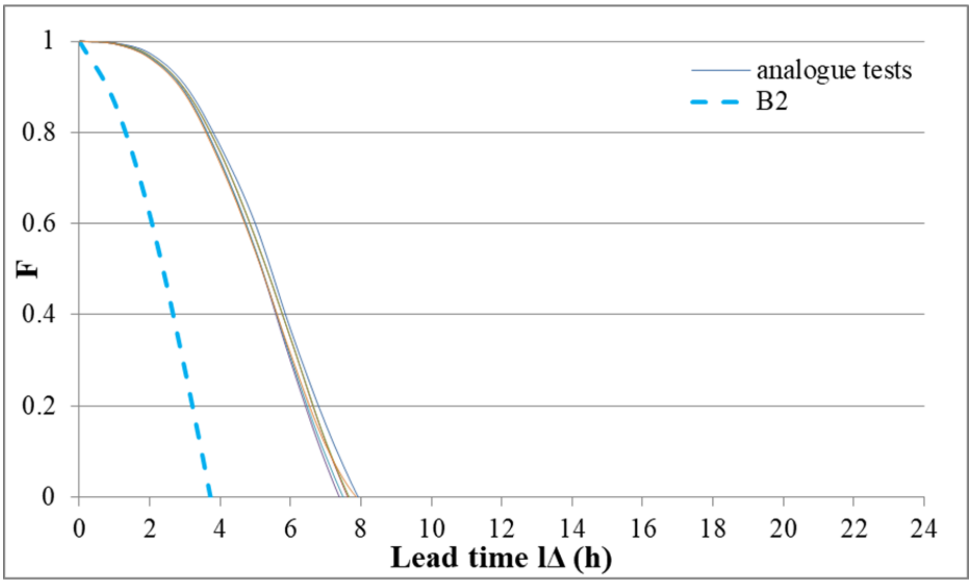
Η ανάλυση ευαισθησίας που πραγματοποιήθηκε για τη σύγκριση του αναλογικού αλγορίθμου και του B2 εφαρμόστηκε και στις αρχικές χρονοσειρές και στις κανονικοποιημένες προκειμένου να διερευνηθεί εάν η επίδραση της κυκλοστασιμότητας είναι αισθητή και στην προγνωσιμότητα. Τελικά το σενάριο μέσου σφάλματος σε συνδυασμό με την κανονικοποιημένη χρονοσειρά έδωσαν ευρύτερο παράθυρο προγνωσιμότητας 15h για τον άνεμο και 13h για την ηλιακή ακτινοβολία.



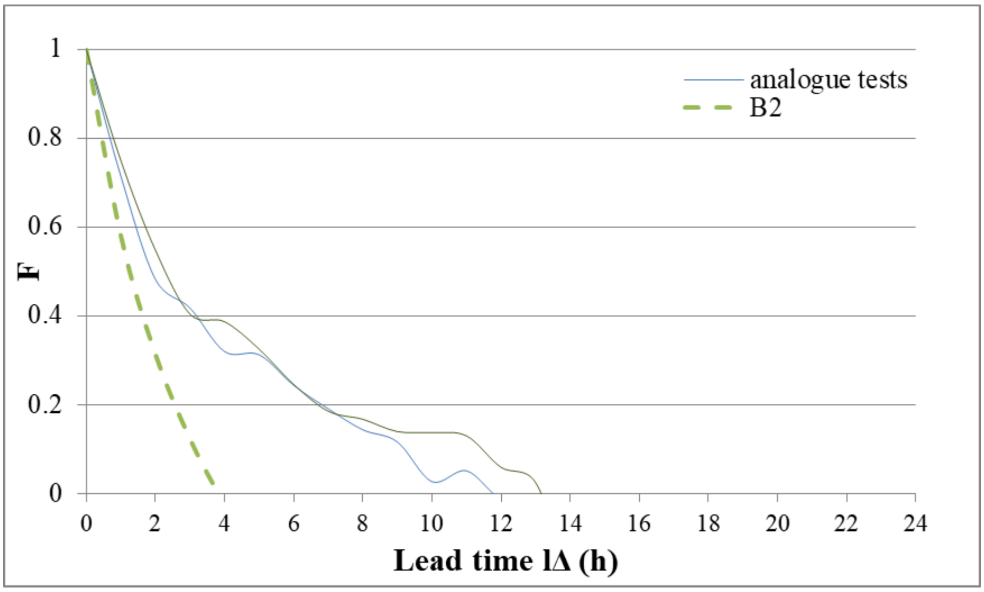
Σχήμα 5. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και B2 ($g_{average}$=0.5) χρονοσειράς ανέμου.*



Σχήμα 5. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και B2 ($g_{average}$=0.5) κανονικοποιημένης χρονοσειράς ανέμου.*



Σχήμα 5. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και Β2 ($g_{average}$=0.5) χρονοσειράς ηλιακής ακτινοβολίας.*



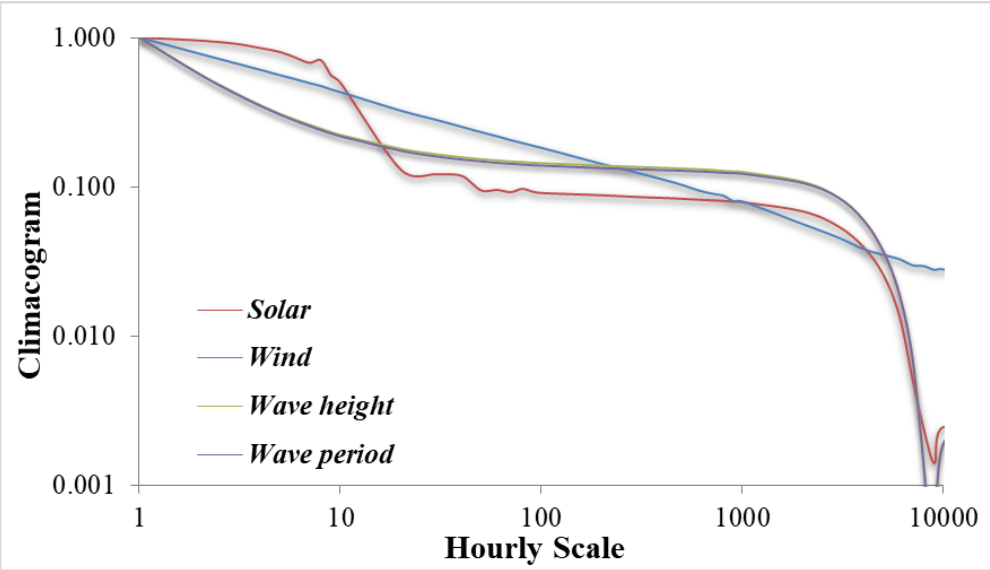
Σχήμα 6. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και Β2 ($g_{average}$=0.5) κανονικοποιημένης χρονοσειράς ηλιακής ακτινοβολίας.*

Πολλά νησιά στον κόσμο δεν είναι συνδεδεμένα με το κεντρικό σύστημα ενέργειας της ηπειρωτικής χώρας και εξαρτώνται από εγκαταστάσεις που τροφοδοτούνται από ορυκτούς πόρους γεγονός που μπορεί να επιφέρει αισθητή οικονομική επιβάρυνση αλλά να είναι επιβλαβές και για το περιβάλλον. Η συνεχής ανάπτυξη της εφαρμογής των ανανεώσιμων ενεργειών, δημιούργησε την ιδέα διαμόρφωσης ενός ενεργειακού συστήματος στην Αστυπάλαια (υδροηλεκτρική, αιολική, ηλιακή, ωκεάνια ενέργεια) που αποτελείται από ανανεώσιμες πηγές ενέργειας. Από τη στιγμή που δεν υπήρχαν σταθμοί μέτρησης στο νησί, δημιουργήθηκαν συνθετικές χρονοσειρές μέσω στοχαστικής μελέτης.

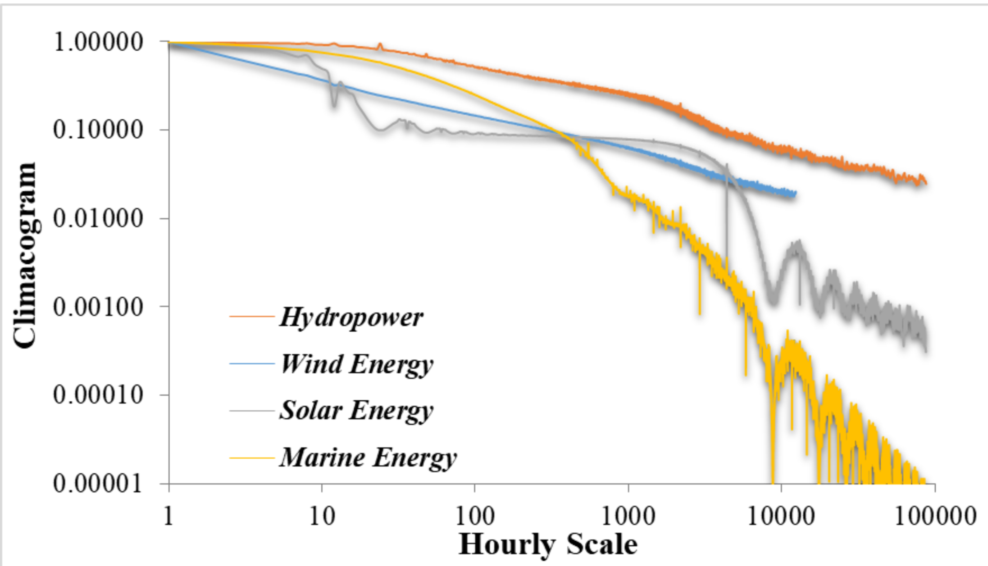
Αφού εξετάστηκαν διάφορες φυσικές διεργασίες μέσω του κλιμακογράμματος διαπιστώθηκε για ακόμα μία φορά πως η εγγενής αβεβαιότητα των φυσικών διεργασιών είναι εμφανής και μάλιστα διαφέρει από διεργασία σε διεργασία (διαφορετική κλίση του κλιμακογράμματος σε μεγαλύτερες κλίμακες).

Το αντίστοιχο κλιμακόγραμμα των παραγόμενων ανανεώσιμων ενεργειών έδειξε πως η «έμφυτη» αβεβαιότητα των διεργασιών μεταφέρεται και στο αντίστοιχο ενεργειακό σύστημα (ηλιακή , αιολική, υδροηλεκτρική και ωκεάνια ενέργεια). Αυτή η αβεβαιότητα μπορεί να είναι υπαίτια για πιθανές τεχνικές βλάβες των μονάδων παραγωγής αλλά να δυσχεραίνει τη λειτουργία και τη διαχείριση του συστήματος. Επιπλέον, θα πρέπει να ληφθεί υπόψη ότι καμία από τις προαναφερθείσες ανανεώσιμες ενέργειες δεν έχει πάντα επαρκές και συνεχόμενο ενεργειακό αποθεματικό. Για το λόγο αυτό σε προηγούμενη έρευνα διερευνήθηκε η εγκατάσταση και η λειτουργία ενός εφεδρικού συστήματος ενέργειας τροφοδοτούμενο από βιομάζα και γεωθερμία.

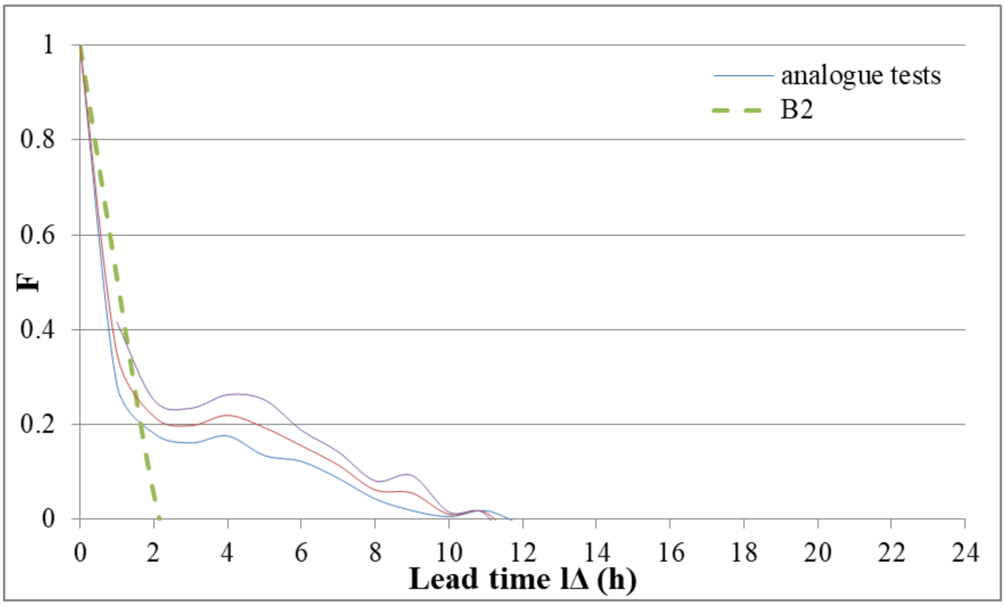
Η Αστυπάλαια σαν νησί έχει πληθώρα φυσικών διεργασιών που θα μπορούσαν να αξιοποιηθούν για την παραγωγή εναλλακτικών μορφών ενέργειας Η βελτιστοποίηση του ενεργειακού συστήματος της Αστυπάλαιας καθίσταται πλέον αναγκαία για την ομαλή λειτουργία και το σωστό προγραμματισμό των μονάδων παραγωγής ενέργειας. Γι αυτό το λόγω διερευνήθηκε η βραχυπρόθεσμη προγνωσιμότητα σε διεργασίες ανέμου και ηλιακής ακτινοβολίας στην Αστυπάλαια ώστε να προσδιοριστεί το χρονικό περιθώριο προβλεψιμότητας αυτών των χρονοσειρών όταν επρόκειτο να αξιοποιηθούν για την παραγωγή ενέργειας.



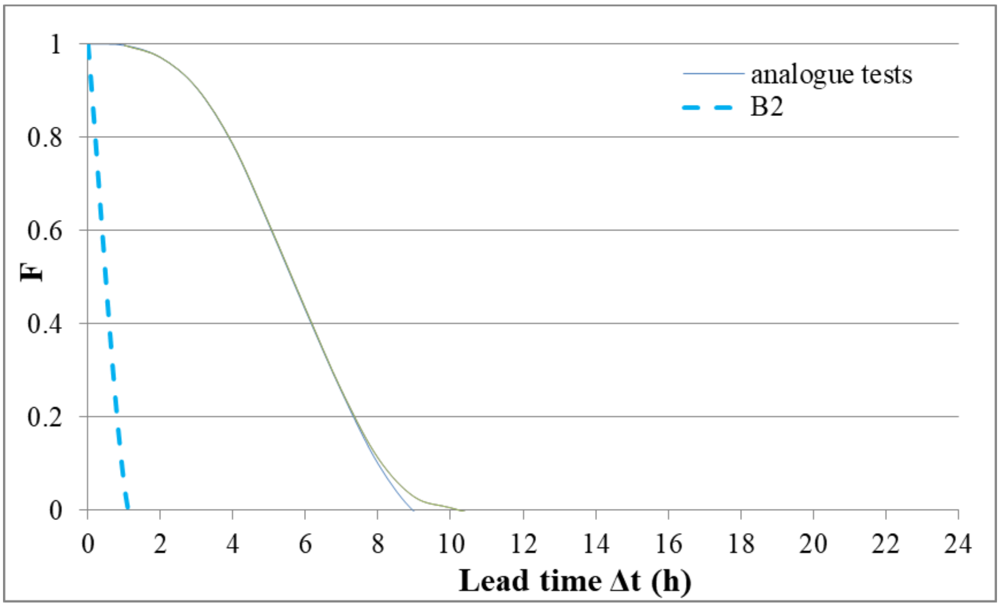
Σχήμα 7. *Κλιμακόγραμμα διάφορων φυσικών υδρομετεωρολογικών διεργασιών που εξετάζονται στο νησί της Αστυπάλαιας*



Σχήμα 8. *Κλιμακόγραμμα διάφορων ανανεώσιμών ενεργειών που εξετάζονται στο νησί της Αστυπάλαιας*



Σχήμα 9. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και Β2 ($g_{average}$=0.5) χρονοσειράς ανμεου στην Αστυπάλαια.*



Σχήμα 10. *Ανάλυση ευαισθησίας αναλογικού μοντέλου και Β2 ($g_{average}$=0.5) χρονοσειράς ηλιακής ακτινοβολίας στην Αστυπάλαια.*

Abstract

Energy demand worldwide is predicted to increase in the coming decades due to various economic and social reasons. Consequently, since fossil fuels deposits decrease dramatically due to overexploitation, renewable energy resources initiate to gradually substitute them. However, natural processes, from which renewable energy derives from, are characterized by an inherent degree of uncertainty that is introduced to the energy system. A recent study suggests that unpredictability and determinism may coexist in nature, defeating the prevailing theory that randomness is a component to predictability. The current thesis investigates the predictability time-window of wind speed and solar radiation processes by applying sensitivity analysis that compares a stochastic algorithm to a naïve one. A pilot application of a hybrid renewable energy system in the Greek non-connected island of Astypalea is also discussed, for which the significance of uncertainty is examined by exploring the uncertainty of the input of natural processes and by defining the predictability window in specific processes, as well.

Contents

1. Introduction

Recent studies indicate that energy demand worldwide is predicted to increase in the coming decades. Global population is constantly proliferating; countries that keep developing their economy multiply energy needs; the political scene is unstable in many countries, especially in the underdeveloped ones (https://www.bbc.com/bitesize/guides/zpmmmp3/revision/1). Consequently, since fossil fuels deposits decrease dramatically due to overexploitation, renewable energy resources initiate to gradually substitute them, providing a more feasible alternative. The continuous advances in renewable energy technology along with the gradual reduction of installation costs, pave the way towards a wider adaptation of renewable energy (Figure 1.1).

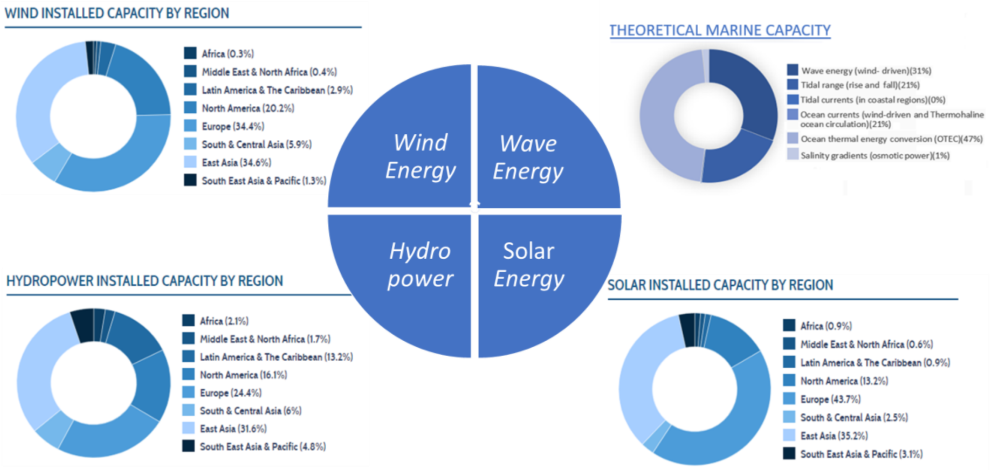


Fig.1.1: *Distribution of renewable energy resources installed capacity worldwide.*
[Source: https://www.worldenergy.org/]

According to the Directive 2009/28 / EC of the European Parliament (Article 2), wind, solar, geothermal and marine energy, hydroelectric power, biomass, gases released from landfills, wastewater treatment gases are considered as energy from renewable non-fossil sources. Renewable energy resources can be beneficial for both humanity and the environment. The continuing energy supply combined with fewer maintenance requirements (compared to fossil fuels) and the diminution of greenhouse emissions can urge people to adopt a "more green" approach to improve the energy balance (https://news.energysage.com/advantages-and-disadvantages-of-renewable-energy/).

The application of this approach has begun lately to be implemented more and more often through the shift to wind and solar energy. Both of them are considered as attractive renewable energy options. Wind is an element that can create renewable energy through wind turbines. The motion of a wind turbine's blades turns wind into kinetic energy which is then converted to electricity through special magnets installed inside the turbine. Solar energy is produced using the energy that solar radiation generates. Solar irradiance is absorbed by cells inside the solar panel. The captivated irradiance turns to DC (direct current) electrical power and then transforms to AC (alternative current) electrical power in the interior of the photovoltaic system which is the form of electricity that can feed an electrical system.

Both Wind and Solar energy are sustainable, abundant and free of pollutant atmospheric emissions compared to hydrocarbons and that is why they are the fastest-growing energy sources worldwide and are subsequently becoming increasingly popular. Hence, energy systems management - the operation and organization of all units involved in both energy production and consumption - is not only crucial but essential too. A long-term energy strategy should be applied to achieve optimization in energy systems operation and energy systems in general in order to have continuous and reliable power.

However, management of renewable energy resources systems can be quite complex. Aeolic and Solar energy derive from phenomena related to atmospheric processes whose behaviour is uncertain. Provided the aforementioned, unpredictability is inherent to renewable energy systems (Klousakou et al., 2018) and that makes energy management even harder. Researches have proven that the degree of uncertainty on the long-term can be identified among different natural processes and quantified in natural systems (Klousakou et al., 2018).

A robust yet innovative stochastic tool, the climacogram (a log-log plot of variance of the averaged process versus averaging time scale), can be applied in order to quantify the degree of unpredictability in a natural process through the Hurst parameter coefficient (H=0~1) (Koutsoyiannis, 2017). Additionally, the widely accepted opinion that unpredictability is a component of predictability in nature is invalid and misleading; determinism and unpredictability actually coexist. However, the greater the uncertainty -and thus the unpredictability- the shorter the horizon in which a process can be accurately predicted. If a process follows a more deterministic behaviour, the predictability window widens (Dimitriadis et al., 2016). The robust characterization of the predictability of the natural processes related to renewable energy production, or equivalently, of the related uncertainty, is a key first step for the design and management of energy systems, and it may also inform more efficient and reliable investments in the energy sector (while as a consequence firmer energy could be produced).

Prediction of processes is feasible despite the context of the process or situation examined. Hence, focusing on renewable energy, prediction of wind speed and solar radiation can be improved by using simple yet accurate models that rely on forecasting methods despite the variability contained in the processes, by minimizing the statistical error of the prediction to a degree. Forecasting is necessary for wind farms and solar arrays planning in order to optimize power generation. This research examines the performance of the analogue model, a simple yet reliable forecasting model according to recent studies. Since Solar and Wind energy represent the largest share in renewable energy systems, this research focuses on the examination of the behaviour of the corresponding natural processes from which energy derives-solar radiation and wind speed timeseries.

Due to the reasons mentioned above, humanity has been integrating alternative energy resources in energy production to cover daily needs. This means that energy production depends on natural elements that derive from either atmospheric or geophysical phenomena which successively leads to the conclusion that electricity production inherits a degree of uncertainty.

Forecasting is a vital solution that helps to cope with the problem of unpredictability in power generation. By calculating a sufficient amount of time in which predictions can be accurate, energy systems optimization will be more feasible. Operators could have the ability to adapt to extreme weather conditions and even identify when to commit or decommit power generators. In cases where energy production is impossible due to special circumstances reserves exploitation could be optimized to facilitate the process. That could also decrease the expanses in power generation creating a more profitable business.

This diploma thesis focuses on the improvement of management of wind and solar energy by investigating an optimum time window in which wind speed and solar irradiance forecasting respectively can be fairly accurate. Despite the fact that those natural processes present significant variability, it is attempted to apply a version of the analogue-ensemble model and create new timeseries with the estimated predicted data.

2. Identification of uncertainty on natural and energy demand processes

2.1 Methodology: Uncertainty quantification based on the second-order dependence structure

A very common characteristic that has been identified among geophysical and atmospheric processes is the Hurst-Kolmogorov behaviour; the more complex the processes, the greater the variability.

The British Hydrologist Edwin Hurst (1880-1978) was the first one to investigate the phenomenon of long-term persistence after studying the Nile River water fluctuation in order to determine the optimum size for a dam to be constructed. Afterwards, the Russian Soviet mathematician Andrey Kolmogorov (1903-1987) expanded the initial research focusing on the water turbidity.

A recent study has compared the results of roulette with the results of the investigation of climatological features in order to identify the Hurst Kolmogorov behaviour. Although the roulette can be described as deterministic, its behaviour was very unpredictable short-term after taking into consideration the long-term bias. However, its uncertainty faded on the long-term as the horizon widened, after it was described as a white noise process.

On the other hand different results were retrieved in the case natural processes. Assuming that the climate can be simulated and explained through chaotic equations, the study showed that climatological phenomena can be unpredictable in all timescales. To sum up, the phenomenon when a process reveals long term persistence is called "the Hurst-Kolmogorov behaviour" (Koutsoyiannis, 2010).

It is noted that a Hurst-Kolmogorov (HK) behaviour (indicating long-term persistence across scales) is expected in all geophysical and hydrometeorological processes (Koutsoyiannis, 2017), and interestingly, it has been shown that processes with completely different deterministic nature (such as hydrometeorological ones or experimental turbulence (Chardavellas et al., 2018; and references therein) exhibit certain stochastic similarities (i.e., long-term persistence), revealing a common stochastic nature.

A robust index that depicts the degree of uncertainty presented in a process long-term is the "Hurst" parameter (Dimitriadis and Koutsoyiannis, 2015). It is applied to define the long-term memory of the examined timeseries. Not only can it quantify the degree of unpredictability inherited in a process, but it simultaneously incorporates the phenomenon of the long-term persistence. The Hurst coefficient is defined as the measure that examines the long-term memory of the respective timeseries. It also presents the rate of change of the timeseries autocorrelation as the lag between pairs of values increases. However, although it would be expected that if the lag widens, the autocorrelation in the process of interest deteriorates, the reverse effect occurs; as the lag increases, the autocorrelation is still above zero. This effect witnesses that nature has inherit persistence which explains the preservation of the correlation to higher values.

For persistent processes the Hurst exponent appears to be H>0.5. Specifically:

- H= 0.5→ There is no correlation between the value on time t and the value on time $t+1$. That means that is shows a totally random change (i.e., white noise)
- H= (0.5, 1] →Persistence: Clustering effects

There are various methods to estimate the Hurst parameter, based on the climacogram, the power spectrum, the autocovariance, etc. (Dimitriadis and Koutsoyiannis, 2015); the climacogram is defined as the log-log plot of variance of the averaged process versus averaging time scale, whereas the power spectrum of a random process describes the distribution of its variance over the frequency.

Scientific studies (Dimitriadis, 2017) have shown that the climacogram is a more valid method (Dimitriadis and Koutsoyiannis, 2015), among the three, for estimating the Hurst parameter and so, it is very useful in detecting long term change or in general the dependence across scale. As far as the power spectrum or autocovariance is concerned, their estimation from data may distort the true behavior of the processes and thus, may lead to false interpretation. Furthermore, the climacogram has the smallest estimation error compared to the other two methods (with the power spectrum presenting the largest one) and a simple expression for statistical bias. The autocovariance and power spectrum are also prone to discretization errors as their values can never be equal with the true value in continuous time. Finally, the power spectrum has a complicated definition (based on the Fourier transformation of the autocovariance), which involves computational cost. Therefore, the climacogram can calculate the Hurst parameter more reliably (Figure 2.1).

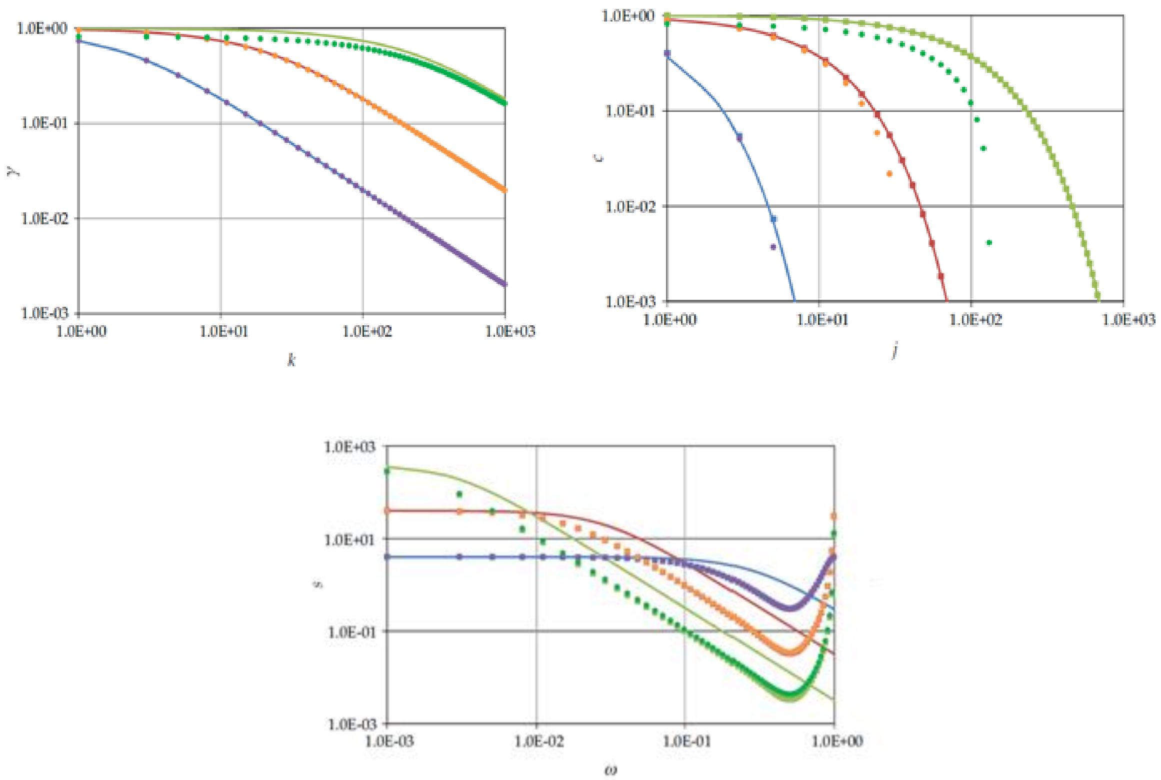


Fig.2.1: *True values in continuous and discrete time and expected values of the climacogram (top left), autocovariance (top right) and power spectrum (bottom).*
[Source: Dimitriadis and Koutsoyiannis, 2015]

A flexible climacogram-based stochastic model, GHK model, was applied in this investigation in each sample in order to estimate the degree of the induced uncertainty of the parent processes. This is a Hurst-Kolmogorov evolution that maximizes entropy production on large time scales but minimizes entropy for small ones. It was selected because it combines markovian behaviour that appears on shorter scales with the long term persistence:

$$\gamma(k) = ((1 + 1/q)/(1 + k/q))^{(2-2H)} \qquad (\textit{\textbf{Eq.2.1}})$$

where γ is the standardized climacogram, k the time scale (here represented as an integral multiple of the finest available time scale), H is the Hurst parameter, and q is a scale-parameter. Due to the effect of the statistical bias, for a robust estimation of the Hurst parameter, we require either a single long-length timeseries or multiple shorter timeseries in case where the former is limited.

2.2 Uncertainty in processes related to renewable energy resources (solar, wind, marine energy and hydropower)

Solar radiation data were accumulated from the "National Renewable Energy Laboratory" (NREL)-"National Solar Radiation Database (NSRDB)". This database includes over 1020 stations that are spread all over the United States of America. However, only a small percentage of them have actual measurements of solar irradiance hourly data. Specifically, 40 stations provide solar radiation hourly data that are measured in the surface of the Earth. The length of the timeseries is 15 years-approximately from the 1st of January 1991 to the 31st of December 2005. It should be noted that all stations have a modeled hourly solar radiation measurements that were generated from the processing of satellite images, as well as a measured hourly external radiation (which is proved to follow a deterministic behavior) (Koudouris, 2017). Since this study focuses on Solar Energy, only actual measurements are taking into consideration to conduct the current research, while the extraterrestrial data is ignored. In terms of the solar energy, renewable energy can be produced through solar radiation and solar panels. After collecting the related data from multiple stations located in the United States of America the climacogram was applied to investigate the behaviour of the process long-term.

It is noted that the climacogram of solar radiation timeseries depicts double periodicity caused by the diurnal and seasonal cycles of the sun (which will be extensively explained in another chapter later), which can be dealt with double cyclo-stationary models (Koudouris, 2017). For this process the Hurst parameter was estimated 0.81 (Figure 2.3).

As far as wind speed measurements are concerned, the corresponding timeseries were collected from a vast database consisting 15000 meteorological stations from across the globe (figure 2.2). Approximately half of them are still operating until this very day. Due to the extreme amount of data that were very difficult to manage, data used in a previous research were also used in the current study. The more adequate data were retrieved from the meteorological station installed in Massachusetts Institute of Technology (M.I.T.). The timeseries of wind speed measurements includes hourly data from approximately 72 years, where only 0.8% of the data were zeros (483.716 hours counting from the year 1943 until 2015. Another basic advantage that led to the choice of this particular timeseries is that only a subtle percentage of data were not documented and subsequently considering that the timeseries had documented data almost throughout its whole length the results could be more realistic.

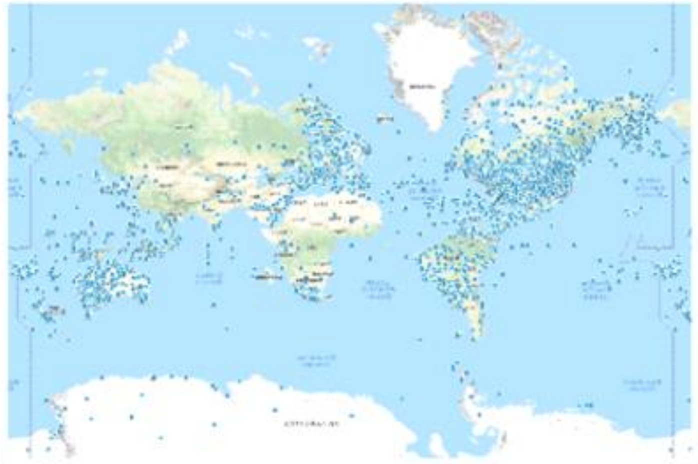


Fig.2.2: *Depiction of the exact location of meteorological stations used for wind speed data. [Source: Gkolemis, 2019]*

Additionally, for an illustration of the uncertainty in the wind speed process (related to the wind energy), the climacogram was estimated for the hourly wind station of MIT (Gkolemis, 2018), with a Hurst parameter equal to 0.61 (Figure 2.4). The climacogram was also applied to precipitation (related to hydropower; the oldest form of renewable energy) (Chalakatevaki, 2018), and the Hurst parameter was estimated 0.54 (Figure 2.5).

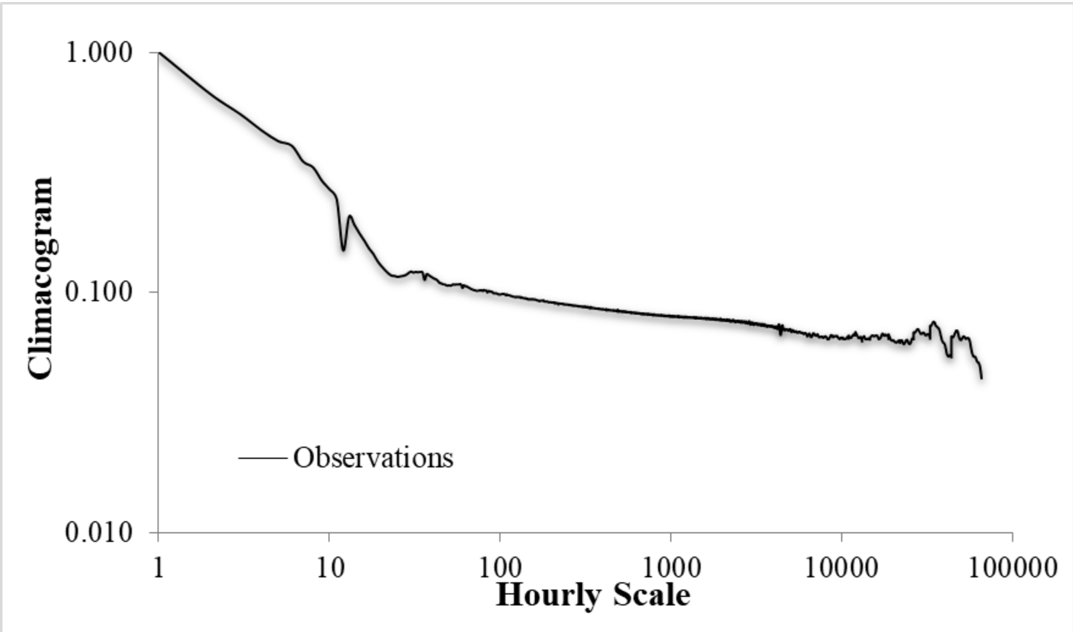


Fig.2.3: *Averaged climacogram of solar radiation from 40 stations.*

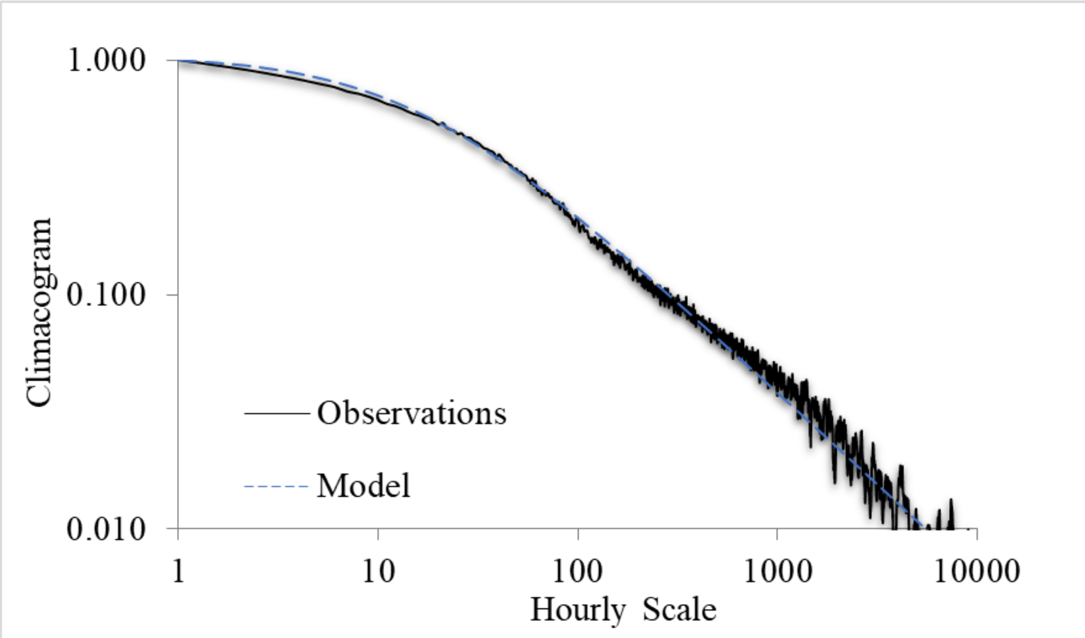


Fig.2.4: *Wind speed climacogram.*

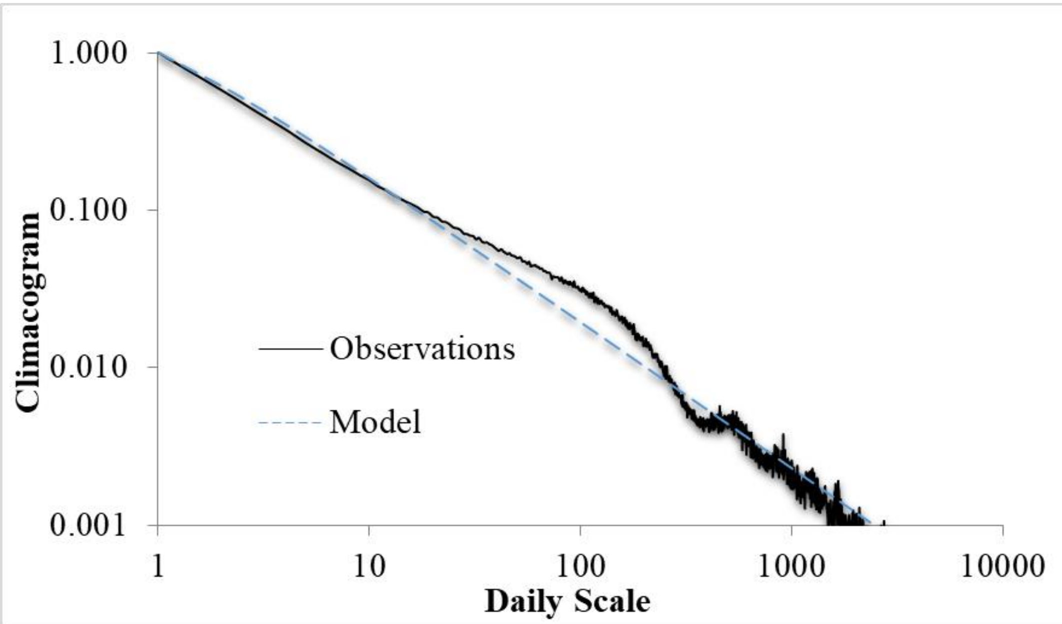


Fig.2.5: *Precipitation climacogram.*

The exploitation of geothermal energy began in the early 20th century. Up to now, its use has been extended because of the increase of the energy demand. In order to investigate the uncertainty for this process, temperature data were collected from a 25 borehole in Idaho, U.S.A. (https://openei.org/doe-opendata/). It is noted that it was extremely difficult to retrieve borehole temperature data since there are very few databases that offer them. After a long research, two different measurements (one in 17-01-2012 and the other 4 on 21-01-2012) provided spatial-series in order to define the Hurst parameter related to geothermal power. In both measurements (Figure 7-Figure8) the Hurst parameter was estimated 0.75 (Figure 2.6-2.7).

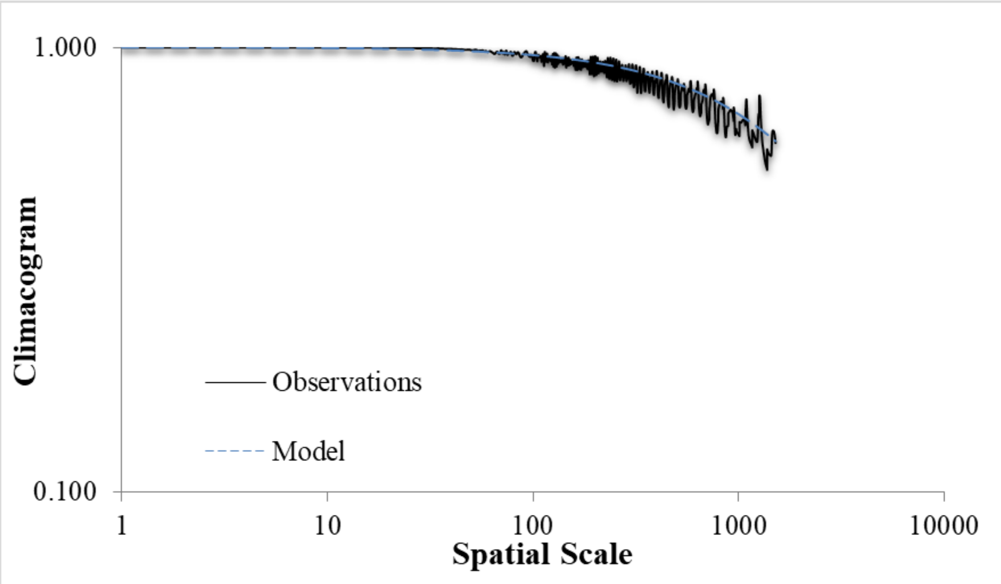


***Fig.2.6:** Climacogram for temperature borehole data measured in 17-01-2012.*

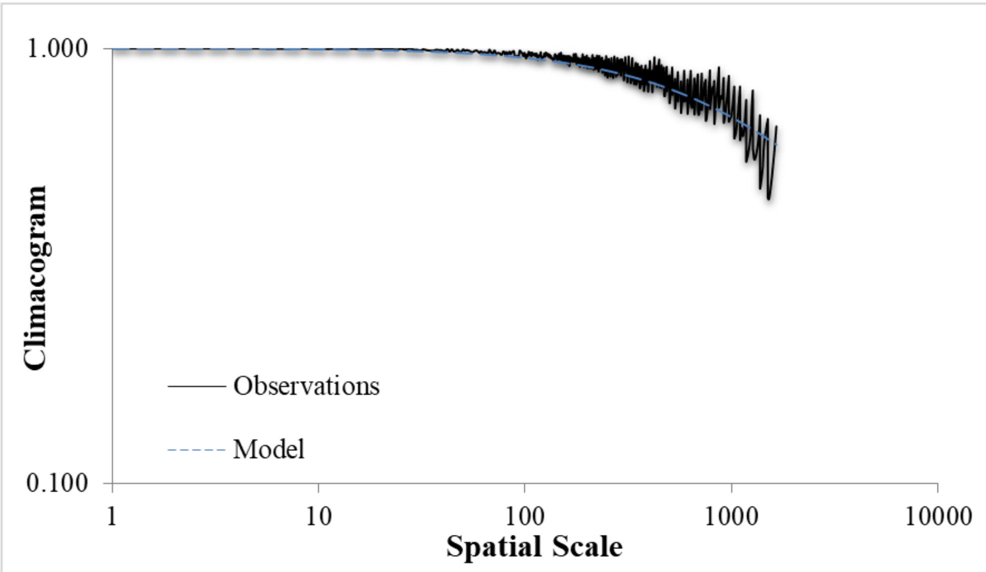


***Fig.2.7:** Climacogram for temperature borehole data measured in 21-01-2012.*

An alternative but not extensively applied form of renewable energy is the ocean (or else marine) energy. Although marine energy is a very promising energy source, it has limited application; this also justifies the limited research made so far on the uncertainty of the related processes. Nevertheless, the marine energy is expected to have a significant contribution in the near future, with a theoretical potential energy production estimated at 7500 EJ/year (Edenhofer et al., 2012). Marine Energy components can be classified into two categories (Edenhofer et al., 2012). The first category includes tides, waves and currents which are generated by gravitational forces.

NOAA (National Oceanic and Atmospheric Administration- www.noaa.gov) is an online database that documents the behaviour of o oceanic and atmospheric processes. Not only does it offer open data for natural processes but it also produces hydrometeorological forecasting covering a wide area with its measurements. Timeseries presenting data for currents and tidal behaviour were acquired thanks to the NOAA database.

As far as tides are concerned, 22 stations across the state of Florida in the USA (https://tidesandcurrents.noaa.gov/products.html) provided an adequate number of timeseries to define the Hurst parameter, which varied from 0.50 to 0.87 among the different stations and was 0.53 in average (Figure 2.8). Data related to currents, were gathered once again from the state of Florida, and the Hurst parameter was found 0.80 (Figure 2.9).

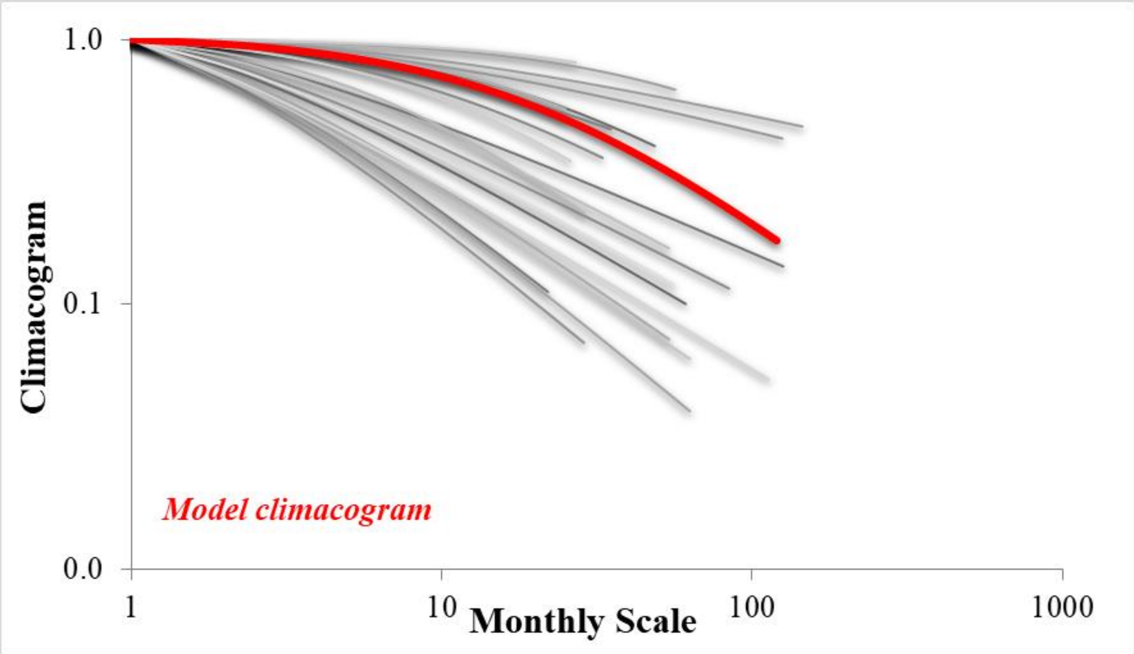


Fig.2.8: *Climacogram for tides.*

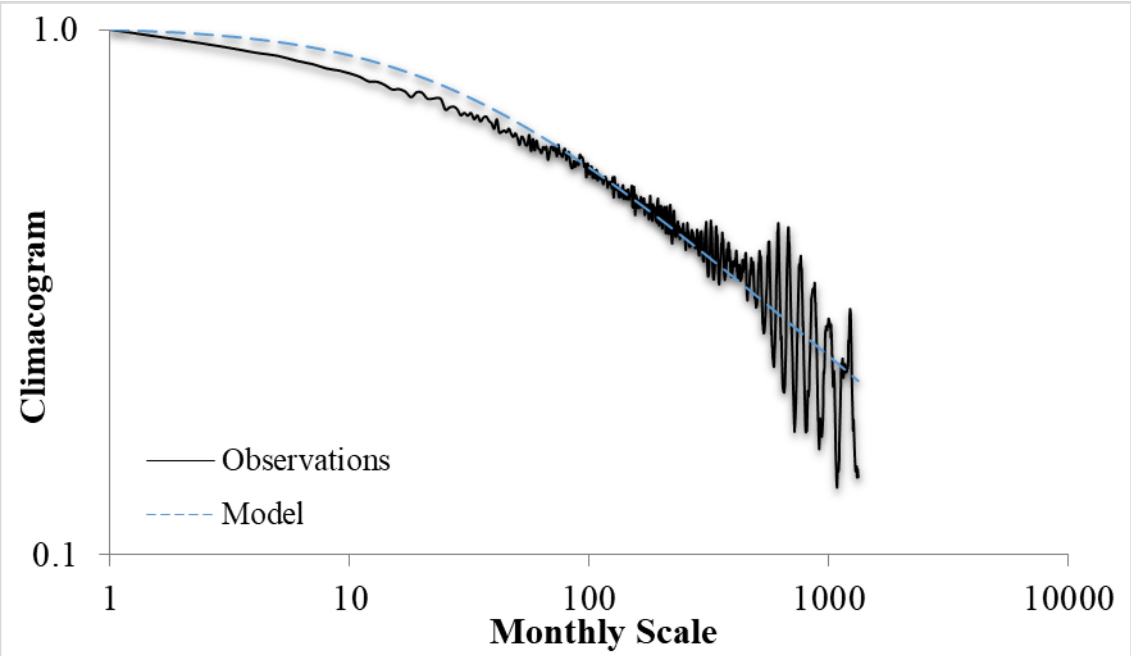


Fig.2.9: *Climacogram for currents.*

The PacIOOS (Pacific Islands Ocean Observing System) Observatory database covers an large area in the Pacific Ocean that includes the State of Hawaiʻi (Kauaʻi, Oʻahu, Maui County, Hawaiʻi Island); the territories of Guam, the Commonwealth of the Northern Mariana Islands , and American Samoa; the Freely Associated States of the Federated States of Micronesia, the Republic of the Marshall Islands, and the Republic of Palau; and the Minor Outlying Islands of Howland, Baker, Johnston, Jarvis, Kingman, Palmyra, Midway, and Wake. It is a vast database that can provide adequate timeseries of data related to the ocean such as waves or currents. It was proven to be of significant importance since all the timeseries retrieves from PacIOOs were almost full with only a subtle percentage of data missing through the length of the timeseries, thus it also documentσ measurements till this current day.

In order to investigate the stochastic behaviour of waves, four different variables were examined from a station in Hilo, Hawaii (data provided by PacIOOS (www.pacioos.org), which is a part of the U.S. Integrated Ocean Observing System (IOOS®), funded in part by National Oceanic and Atmospheric Administration (NOAA) Award #NA16NOS0120024.): height, direction, mean period and peak period. The Hurst parameter was calculated 0.75 for height (Figure 2.10), 0.87 for direction (Figure 2.11), 0.84 for peak period (Figure 2.12) and 0.66 for mean period (Figure 2.13).

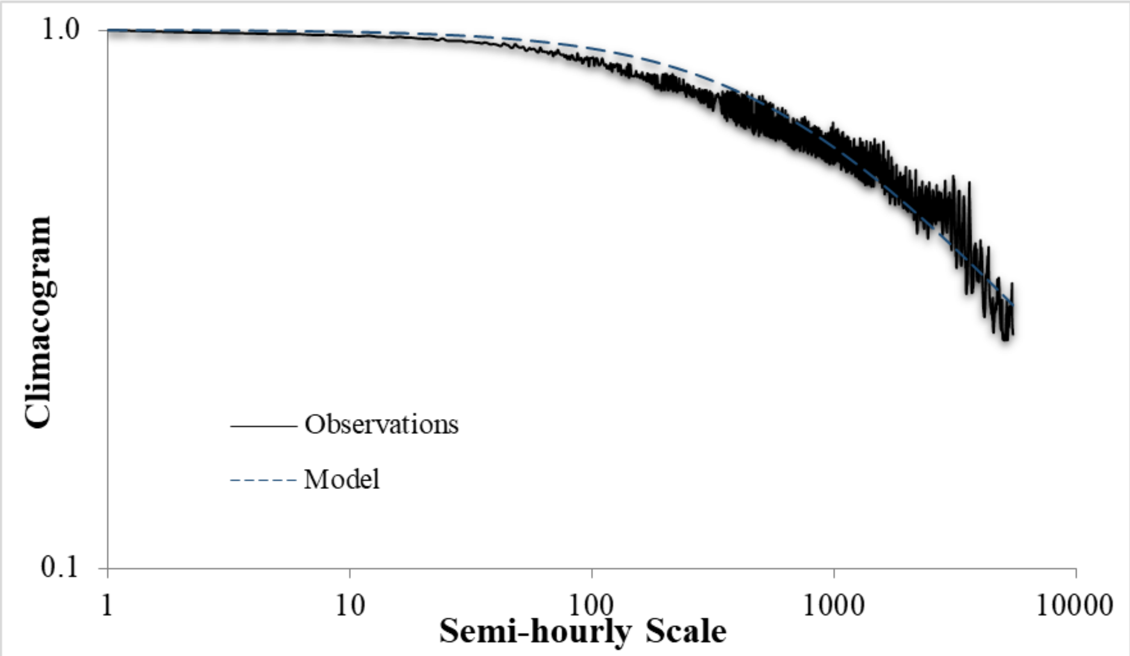


Fig.2.10: *Climacogram for wave height.*

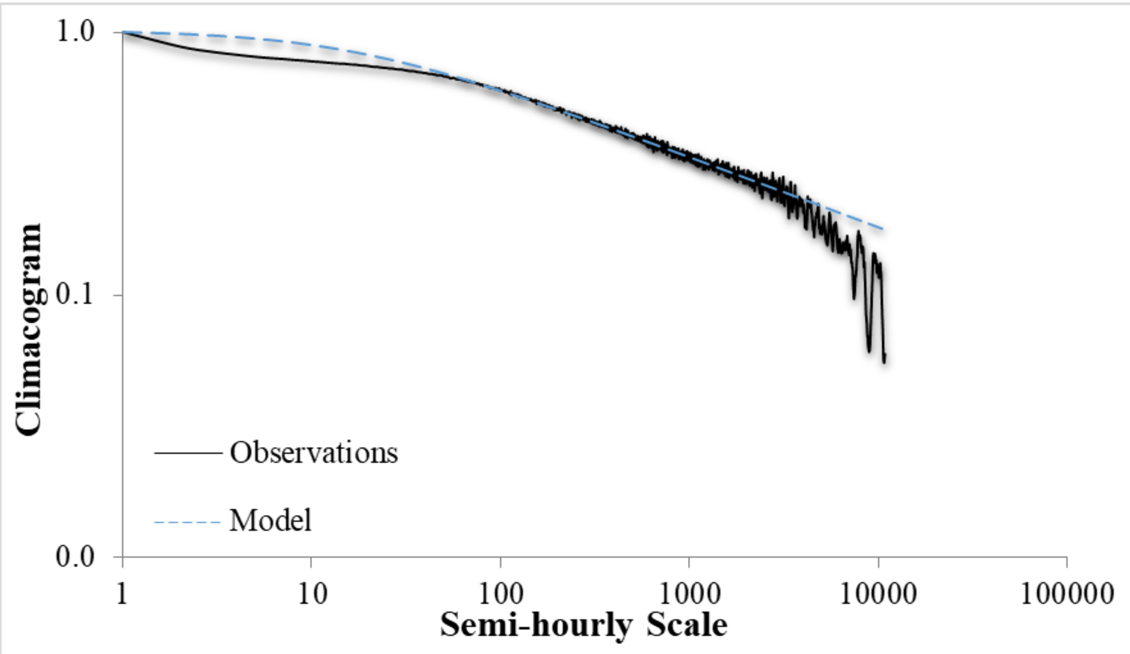


Fig.2.11: *Climacogram for wave direction.*

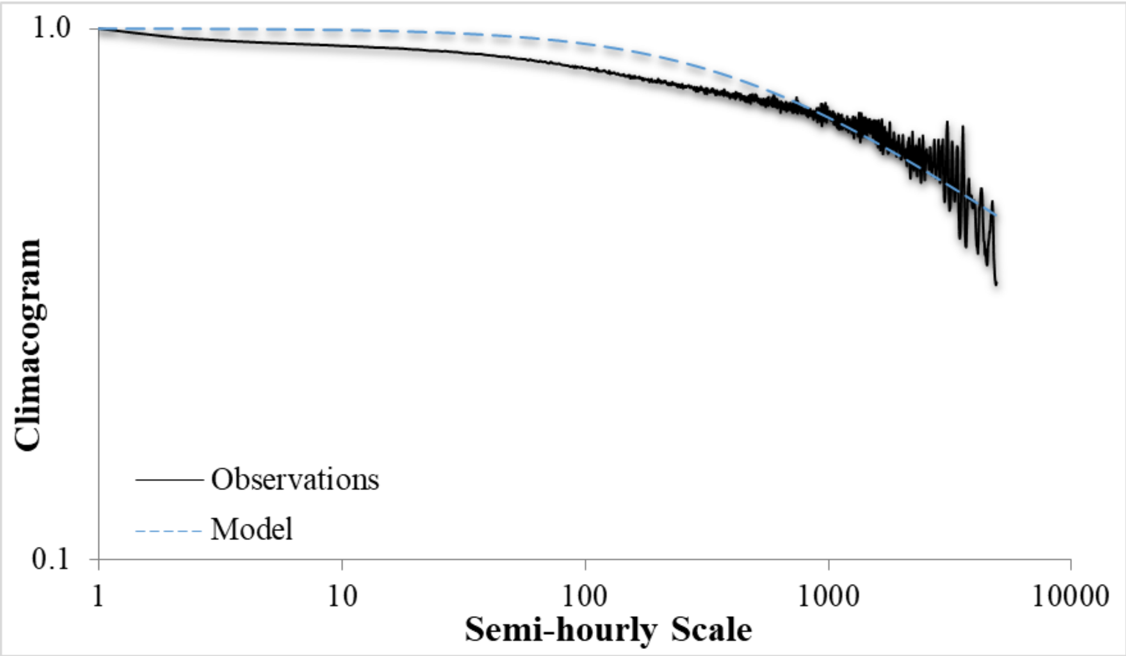


Fig.2.12: *Climacogram for waves' peak period.*

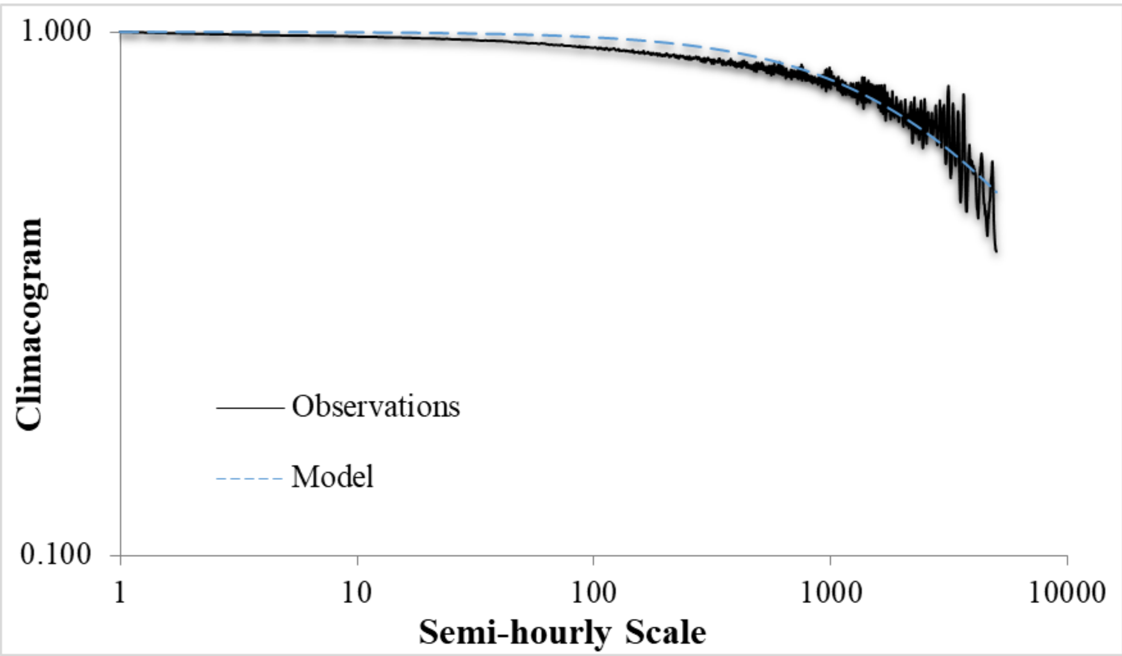


Fig.2.13: *Climacogram for waves' mean period.*

The second category of marine energy is expressed through sources related to chemical processes. Specifically, energy can be produced from temperature differences either from warm surface water (22-27 °C) or very cold water (4-7 °C) at a depth of approximately 1km; this form of energy is called "Ocean Thermal Energy" (Etemadi et al., 2012). It refers to new energy sources which can be characterized as innovative. People are not really aware of energy production through salinity and temperature differences since it is still in the experimental state. Finding data related to the respective natural processes proved to be quite inconvenient since there were not enough available databases. Thankfully two different stations located in Florida provided all the necessary data (SSW28", DSW40"; https://catalog.data.gov/dataset/temperature-salinity-and-waterchemistry-data-from-the-comprehensive-environmental-monitoring-p06c6a). Respectively, in both stations the Hurst parameter was estimated close to 0.55 (Figure 2.14-2.15).

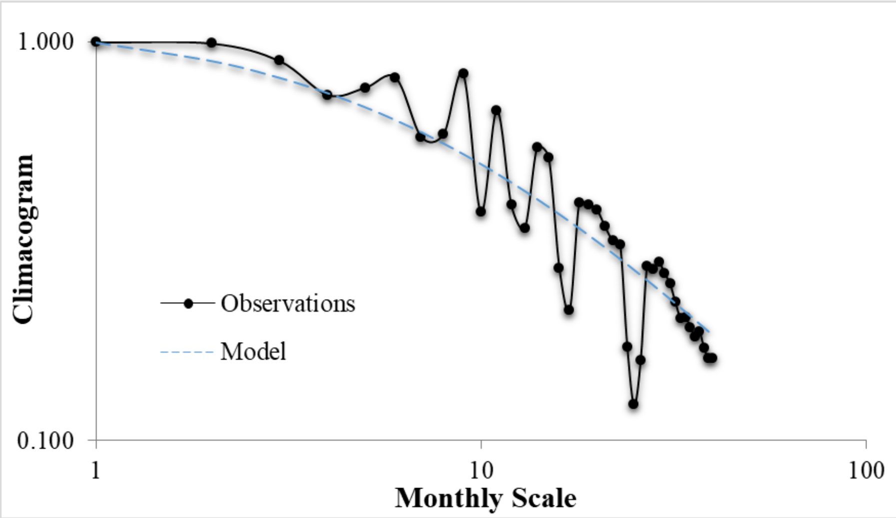


Fig.2.14: *Climacogram for water temperature in Florida (SSW28").*

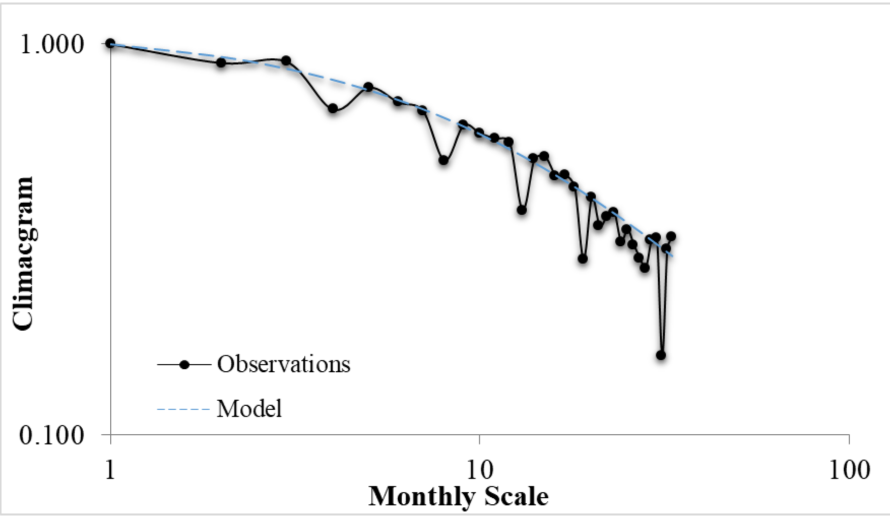


Fig.2.15: *Climacogram for water temperature in Florida (DSW40").*

Moreover, salinity differences can also generate renewable energy. "Salinity Gradient Energy" (S.G.E) (Schaetzle et al., 2015), often known as the blue energy, relies on the energy that is released when two solutions with different salinities come in contact and mix. The same database used for temperature timeseries was also used in the case of salinity. In both climacograms (each one referring to the respective stations like the temperature data in Figures 2.14-2.15) the Hurst parameter was found again 0.55 (Figures 2.16-2.17).

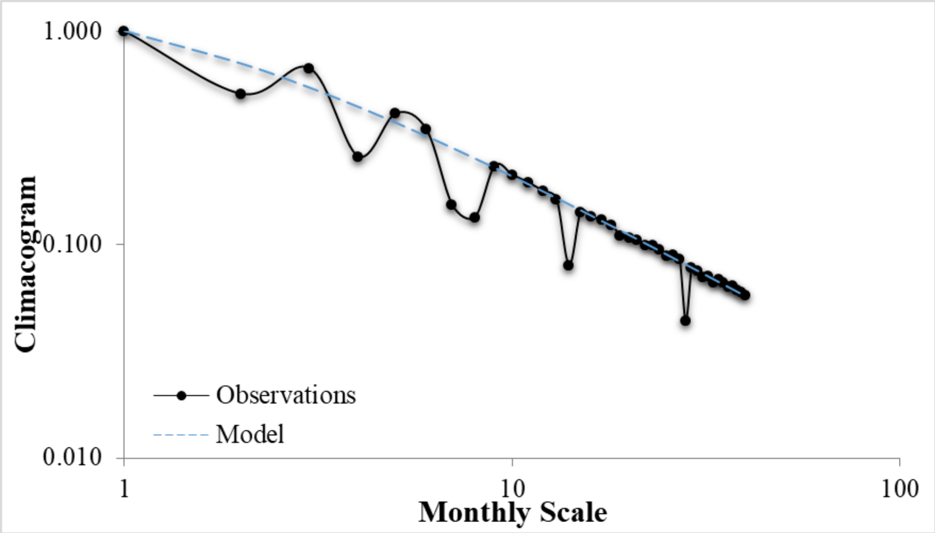


Fig.2.16: *Climacogram for salinities in Florida (SSW28").*

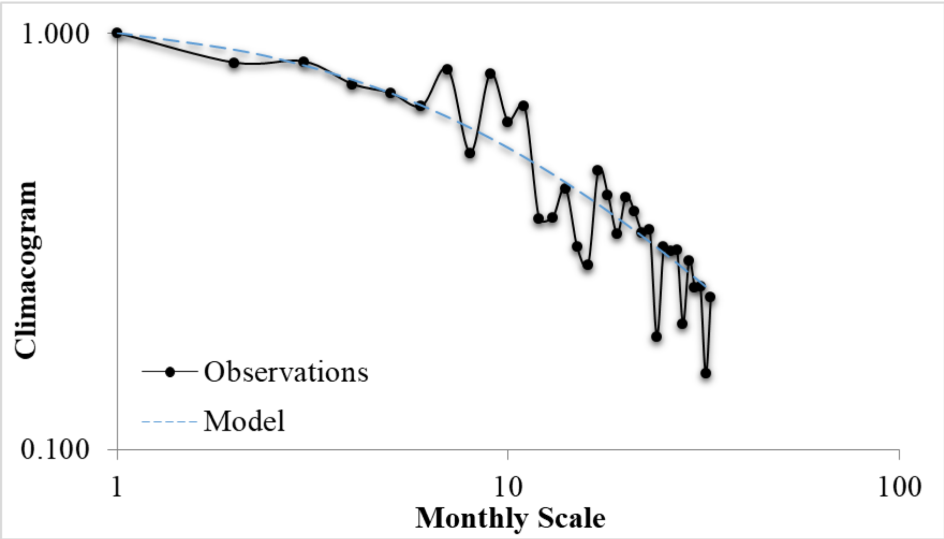


Fig.2.17: *Climacogram for water temperature in Florida (DSW40").*

3.1 Typical forecasting algorithms applied on hydrometeorological processes

3.1 Necessity of forecasting

After producing multiple climacograms that refer to natural processes from which renewable energy resources can be produced, the values of the Hurst exponent that were calculated to quantify unpredictability in those processes pointed that nature can be quite unpredictable.

However, although uncertainty rules most natural processes, natural systems may instil both randomness and determinism in their processes. A dominant understanding that unpredictability is a component of predictability, suggests that there is a “virus of randomness” infecting only specific phenomena. Recent studies indicate that this hypothesis is inaccurate and misleading (Dimitriadis et al., 2016); in fact, determinism and unpredictability can actually coexist, although the latter has a wider time-window of predictability than the former, in which predictability dominates unpredictability.

Since solar and wind energy seem to be very appealing worldwide due to their advantages, this chapter focuses in the investigation of solar irradiance and wind speed behaviour in order to define their predictability time widow.

Energy production may reserve risks that could have both practical and financial consequences to the system:

- Energy production means can be very expensive during their installation. However, the cost may increase due to constant preservation processes in order to preserve the lifespan of the operator.
- An unexpected failure on the operators could postpone energy production if there is no adequate back-up plan.
- Reservoir of energy surplus installation is also a considerable financial factor that can affect the operation of an energy system if it is not exploited wisely and when truly necessary. Storage of energy deposits should be explored through holistic studies and not just financial optimization considering cost as the only parameter of management.
- Natural Processes from which renewable energy resources derive can be quite advantageous due to their ability for infinite power supply. Nevertheless, their inherent uncertainty could render them disadvantageous.
- Unreliable forecasting cannot promise the safe operation of the system since it is based on erroneous data.

Short-term prediction of natural processes could be proved essential in energy production and it may orient to a more sufficient operation of energy systems. Forecasting was being applied hundreds of years ago, even when there were no respective technological means to investigate it properly. Reaching today it is already applied in multiple sectors. Many functions have become quite complex because of their multi-varied dependence. Predictions can be an indirect mean to reduce the uncertainty of a function's nature. That means that more emphasis should be given in prediction generation due to its crucial necessity.

3.2 A brief history of forecasting

The human kind had conceived the idea of forecasting centuries ago willing to fulfil its need for preparation, programming and precaution. In order to explain the change of seasons and various weather phenomena ancient civilizations starting observing the changes in the behaviour of natural elements and processes such as the movement of the sun, precipitation or motion of clouds. More specifically, humans relied on visual proof and weather wisdom that they tried to record. In 650 BC, Babylonians pursued to produce short-term weather forecasting based on the behaviour of clouds or through astrology. Later in 300 BC, Chinese people were able to create a calendar depicting a year that included 24 festivals. The grouping occurred according to weather conditions (https://earthobservatory.nasa.gov/).

Ancient Greeks were also pioneers in weather forecasting. Although ancients Greeks used mythology and the action of gods or divine creatures to explain changes in the weather, the Greek philosopher Aristotle (384-322 BC) developed theories about the explanation of natural phenomena such as winds, thunders or lightning, hurricanes etc. In fact, he compiled his theories in a four-volume dissertation named "*Meteorologica (Μετεωρολογικά)* ". Although some of his sayings in the book were inaccurate, it was centuries later that they were rejected by modern ones. The successor to Aristotle in the "Peripatic School" (a school of philosophy in Athens, Greece) was Theophrastus, who also documented his investigation in weather forecasting in a book with the name "*Weather Signs (Τα Μερομήνια)".*

Optical observations and speculations were inadequate for forecasting. A huge obstacle on weather predictions generation was the lack of suitable equipment and measurement tools. During the Renaissance, the need for scientific training and development of technological means was obvious. By the end of the 17^{th} century, the contribution of new appealing theories developed on forecasting and the invention of instruments capable of measuring natural processes characteristics enhanced the improvement of predictions. In fact, they started documenting more data thanks to the new technological means and they were able to create databases, maps or even wind patterns.

The importance of immediate evolution of forecasting appears studying the contemporary history. The English Rodger FitzRoy (1805-1865) was another pioneer who lived during the 19^{th} century. He actually introduced the word "*forecast"* to the scientific community. At that time, wild storms near British coasts costed the lifes of many men. After a deadly shipwreck, FitzRoy was chosen to manage the issue of warnings during the storms. Thus, he started gathering data from multiple locations near the Coasts of Britain through a telegraph system

that quickly expanded. Consequently, with the accumulation of data, he started producing weather forecasting to warn ships on time (https://earthobservatory.nasa.gov/).

Forecasting is a need that exists from ancient years and is essential for multiple sectors in human activities. The radical technological development nowadays has improved forecasting methods and means helping producing forecasting. Thanks to computational systems and by automation it is more feasible to manage situations or prevent future difficulties.

3.3 Forecasting algorithms

Nowadays forecasting can be described more accurately as a necessity. It is widely applied to manage different operations and situations (from weather prediction to economic issues). Therefore, many methods with various properties have been developed. According to studies the choice of the most suitable technique depends mostly on the experience of the researcher, the availability of historical data, the nature of the predicted elements, the availability of time to fulfil the forecast and the cost of the forecasting method (Chambers et al., 1971).

As far as natural processes are concerned, meteorologists apply the next methods to make weather forecasting: the Climatology Method, the Analogue Method, the Persistence and Trends method and Numerical Weather Prediction Method (NWP) (ww2010@atmos.uiuc.edu). Climatologists create predictions via supercomputers focusing on numerical or statistical procedures (Alonzo, 2017).

In wind speed and wind power generation forecasting there are four prevailing techniques that are applied to make predictions according to the literature and they derive from the generalized methods above; Persistence Method, Statistical Method, Physical Method and Hybrid Methods (Soman et al., 2010) (Figure 3.1):

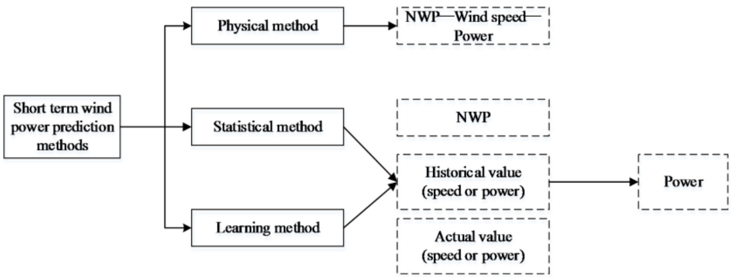


Fig.3.1: *Short-term wind power prediction methods.*
[Source: Xu et al., 2015]

Persistence and Trend method: The persistence method is considered one of the easiest methods to create a forecast and it relies on the past. It is based on the principle that a process will be the same in the near future (the process at time Δt will be the same as time $t+\Delta t$), providing that the environmental conditions remain unchanged. For example, if wind speed is the process examined, tomorrow's wind speed will be the same as today's data (Soman et al., 2010). It is a very reliable method for short term forecasting. Persistence and Trend Method is usually used to benchmark against different prediction methods in order to improve the forecasting method results.

Physical Method: This method is applied to produce short to medium weather forecast (including wind speed forecast). The physical approach uses parameterization of the lower atmosphere based on the information given from weather service nearby the area of interest (Soman et al., 2010). Then the prediction of wind speed is produced through Numerical Weather Prediction (NWP) models where it should emphasize that they are not capable of providing accurate wind speed forecast when the topography of the area investigated is complex (Cassola et al., 2012). When it comes to wind power generation, wind prediction is taken one step further by fitting wind data to the correct hub high of wind turbines and then applying the power curve of turbines or the logarithmic power law to convert wind process to Aeolic power (Foley et al., 2010 and references therein).

Statistical Methods: When it is required to create prediction for short horizons, statistical method can be proven reliable. However, as the time window of forecasting widens, the prediction error also increases impeding the results given by these methods (Chang 2014). Statistical methods process many data and ignore the meteorological conditions when it is destined for energy demand forecasting. A wide variety of statistical techniques such as autoregressive (AR), moving average (MA), autoregressive moving average model, (ARMA) and auto regressive integrated moving average model (ARIMA) are applied to create a statistical relationship between natural processes predicted values and historical data (Foley et al., 2010 and references therein). The more subtle the difference the smaller the statistical error; the forecasting becomes more accurate (Soman et al., 2010).

Hybrid Methods: All the above methods can be mixed to generate natural processes forecasting despite their differences. For example a common application of a hybrid method is the combination of physical and statistical methods. Short-term and medium-term models can form another hybrid model (Soman et al., 2010). The hybrid approach makes wind speed forecasting methods quite versatile.

The choice of the forecasting methodology, when it comes to solar irradiance prediction, depends on the horizon that the researcher wants to foresee (Figure 3.2). The accuracy of the forecasting method relies on the predicted horizon. Those methods are classified into two categories according to studies (Voyant et al. 2016):

The ***first category*** refers to methods that are more valid for short-term predictions for approximately up to 6h. Statistical methods and extrapolation that process Satellite Image techniques and Ground based image techniques can derive predicted data (Voyant et al. 2016). A way to predict solar radiation through those methods is to calculate the accumulated light that penetrates a cloud after it is reflected on a satellite chosen for the research or by measuring the shadow of clouds through spatial resolutions (Arya et al., 2015). Numerical

Weather Prediction models can achieve the generation of solar radiation forecasting usually using real-time measurements.

The ***second category*** includes integrated Numeric Weather prediction models that can achieve fairly accurate prediction with a wider time horizon up to 2d and sometimes even longer than that. When applying Numerical Weather Prediction it is very often that post-processing will be needed and satellite data may also be involved.

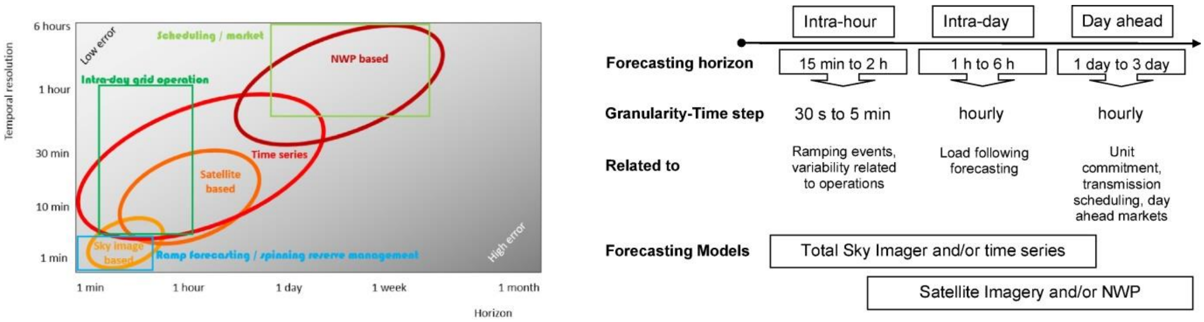


Fig.3.2: (a) Forecasting error versus forecasting models (left) and (b) Relation between forecasting horizons, forecasting models and the related activities (right).
[Source: Voyant et al., 2016]

So far, various types of models have been applied weighting the prediction according to the models' accuracy on a variety of levels during a research. Many factors can affect the selection of the method used such as the relevance and availability of historical data, the degree of accuracy desirable, the time period the researcher aims to foresee, the parameters included (Jackson et al., 2004) and the cost/value of the forecast (Chambers et al., 1971). It is essential that the behavior of the process or any effect of bias should be investigated in order to minimize the estimating error and increase the credibility of the model. The majority of forecast models can be categorized into three dominating types: stochastic prediction models, analogue models and machine learning models:

Stochastic (Chaotic) Forecast Models: A stochastic process is defined as a large cluster of random variables (chaotic behaviour), usually following a non-linear pattern, manifesting dynamic evolution or describing periodic phenomena (non-stationary processes). Consequently stochastic forecast models can efficiently represent a phenomenon that includes randomness due to its data variability. Stochastic forecast models are extensively used in many sectors such as optimization of energy systems, investigation of hydrometeorological phenomena even in biology and medical studies.

Deterministic models can be characterized as the opposites of stochastic modelling. They receive data as inputs and reproduce the same results no matter how many times the model is applied. A strict restriction on deterministic modeling is that factors with uncertainty are considered external parameters.

On the other hand, the principle followed by stochastic models is that they recognize the variability of certain factors and incorporate it to the results. Thus, they examine different scenarios by combining multiple factors and parameters. Every time the parameters change, the output is modified too offering changeable results. Natural processes appear to have a significant degree of variability which implies uncertainty. Stochastic modelling can be an essential solution to create simulations of natural systems. However, it is important that prior to the construction of the stochastic model it is necessary to study the marginal and dependence properties of the process of interest (Koutsoyiannis and Langousis, 2011).

Although natural processes are mostly defined by unpredictability, the paradox is that they exhibit some autocorrelation on short timescales and subsequently they could be also characterized as "chaotic deterministic models" (Breuer, 2011), implying that chaos theory can provide an alternative to either stochastic or deterministic models.

Machine Learning (Artificial Neural Networks): As technology constantly evolves, artificial intelligence is becoming more and more prevalent in many scientific sectors, including forecasting. Through artificial intelligence, algorithms can improve their performance and minimize their statistical error to a degree by evaluating the situation they examine (Makridakis et al., 2018). Machine learning models rely the way they operate on artificial intelligence.

Machine learning (ML), a mix of statistical models and mathematical optimization, is a new method that computational systems exploit to estimate a prediction (Ahmed et al., 2010). An algorithm tries to trace the relationship and the correlation the variable examined has with the data inserted as input. Machine learning is directly correlated to statistics, but there is an important difference. Results from statistical methods are based on the examination of a specific sample depending on the nature of the research while ML models generalize predictive patterns. Many researchers have combined statistical and machine learning methods, creating a new field in science called "Statistic Learning" (James et al., 2013).

By inserting the input of interest, the machine learning models are able to predict a future situation or even make a decision. In fact, some machine learning techniques, such as artificial neural network, are inspired by the human brain's neurological networks (Solomatine, 2010). ML algorithms, combining new software with the human intelligence and specialization, have multiple sectors of application, every sector of the economy such as workforce planning, real-estate market, stock/shares, but can be easily used in natural systems (Sharma et al., 2011, Tomasino et al., 2009,), biophysical parameters, medical investigations or even human functions and human behaviour examination as well. A very important feature of machine learning forecasting models is that they are based on past states.

This means that the performance of a ML model improves if there are adequate historical data and especially a database updated with recent measurements.

ML can be divided into two main categories: supervised and unsupervised learning. The former category uses a database that includes both the input and the desired outcome in order to create the prediction algorithm; when the output is limited to a certain group of values the machine learning is called classification model, but when the output of the prediction is included in a range of values the algorithm is called regression algorithm. The latter subgroup (unsupervised learning) performs its tasks only by inserting the input solely without the output to the model (Brownlee, 2016).

Analogue Models: The Analogue Model, or Analogue-based Ensemble Model, does not use any mathematical expressions between the variables. On the contraire, it focuses exclusively on repeating patterns and patterns' changes that appear in the past. It is important to note that the Analogue Model is capable to produce mostly short term forecast.

A recent study suggests that the analogue model relies on the creation of a Probability Distribution Function (PDF). After calculating the PDF of a process from the historical data the forecast is then generated. In order to verify the credibility and accuracy of the forecast, current predicted data are compared to the historical values of the process examined (Written by Laura Clemente-Harding as a part of the Warner Internship for Scientific Enrichment (WISE) Teaching Material Development Version 1.0)

Another approach of the Analogue model indicates that the prediction can be produced by examining historical data of a process (Dimitriadis et al., 2016). By investigating past states in a process, called neighbours, the analogue model can estimate a prediction in the lead time of interest. Consequently, in order to examine timeseries and time projection that is based on the past, adequate historical data are of significant importance (Fraedrich et al., 2003). Analogue models may often require post-processing in order to minimize the estimating error and enhance both the reliability and the statistical consistency (L.D. Monache and S. Alessandrini, 2015). Analogue models can be efficiently applied to natural processes or any other kind of timeseries that has documented past data.

Hybrid models: Many cases may allow or may even impose that the optimum mean for prediction is a model that focuses on more than one forecasting method. For example, a previous study has successfully presented the amalgamation of an analogue model with a machine learning model called NLAPA (non-linear analogue predictor analysis). This hybrid method maintains the characteristics of both sub-methods, proving that it can outperform by producing satisfying outputs in hydrometeorology and climate downscaling (Cannon, 2007). Neural networks (machine learning) are not only functional with the analogue model. A stochastic approach of the machine learning algorithms, which is another hybrid method, can also guarantee a good performance in specific cases, including biology and medicine researches, compared to other models (Tian et al., 2003).

The choice of the model depends on multiple parameters according to the case in question. Each forecasting model can outperformed compared to the others. Yet, there are significant disadvantages that can define the reliability of the model and the simplicity that each model can stand.

To begin with, the analogue model is very easy to construct and that casts it quite popular (Koutsoyiannis et al., 2010). Still, its performance deteriorates as processes are examined in the long term; the shorter the horizon, the better the prediction (Fraedrich et al., 2003). Moreover, in order to optimize the performance of the analogue model, calibration is needed to increase statistical consistency and reliability and to eliminate the error (L.D. Monache and S. Alessandrini, 2015) and also even if the analogue-ensemble based algorithms produces valid predictions, consistency is not granted in deterministic, natural processes (Koutsoyiannis et al., 2010).

A very important step in forecasting is to pinpoint the degree to which a prediction algorithm is efficient and whether it is feasible to ameliorate their performance. Baseline and benchmark models are applied as comparison indexes to compare if the behaviour of the model improves or deteriorates.

Baseline models: The forecasting models also include baseline models. Baseline models are very basic algorithms that can predict data of interest. They are created in order to compare their results with the output of another more demanding forecasting model.

Baseline testing is often applied first in order to set "a point of departure" that can be exceeded in order to optimize the performance of the model compared to it. It compares present performance of application with its own previous performance intending to clarify whether changes in modelling affected the accuracy of predictions. It is used as an index to help the researcher proceed in any modifications needed to improve the output of the model.

It is often used in machine learning. Machine learning can regularly present different scenarios-output when modifying the parameters. Consequently, a new baseline model is created to figure the quality of the performance.

Benchmark models: Benchmark modelling is another procedure to evaluate the capabilities of a forecasting model. For instance, it can be the point of comparison with a forecasting model; or it can also perform as a comparison indicator with market or industry data.

Benchmark models should run at the beginning of the procedure of forecasting to define a threshold for comparison. They essentially set the point which the forecasting model must overcome to be more reliable.

After comparing forecasting models performance, some interesting conclusions came up. As far as machine learning/neural networks models are concerned, they have good performance in the near future, just like the analogue model. However, when estimating predictions in the long-term, its behaviour deteriorates by either overestimating or underestimating the predicted value (Solomatine, 2003). This occurs because all statistical methods are based on the principle that processes progress following a certain repeating pattern especially on shorter horizons (Chambers et al., 1971). Consequently, lack of recent data in the processes in question may distort the prediction in both machine learning and analogue models. Additionally, both analogue-ensemble based models and NLAPA models cannot predict more values than the length of the examined timeseries and consequently post-processing step is needed (Cannon, 2007).

The comparison between the hybrid model NLAPA and analogue models indicates that it shows the ability of matching the performance of analogue models that rely on linear dimension reduction or predictor scaling algorithms when the relationship between the data used for the forecasting model and the predicted data is linear, and outreaching the performance of the aforementioned models when there is a nonlinear relationship. The outcome of experiments performed on real-world datasets proved that NLAPA (the hybrid model) can outperform in complicated synoptic- to local-scale relationships, like precipitation, while preserving inter-variable relationships and spatial relationships between sites (Cannon, 2007).

Finally, a recent study suggests that since nature is dominated by randomness, a stochastic approach can be an optimal solution for forecasting natural processes behavior due to the theory of maximum entropy (Koutsoyiannis et al., 2010). Additionally, after comparing stochastic forecast models to other types of forecasting models it was concluded that not only can they outperform and predict accurate results of future values in short univariate timeseries like machine learning models (Papacharalampous, 2019), but also in hydrosystems predicted values of stochastic models can be more valid than the results from analogue models (Koutsoyiannis, 2011).

4. Sensitivity analysis of forecasting algorithms

4.1 Sensitivity analysis of wind speed and solar radiation processes for calculation of predictability time-window

The short-term investigation was focused on wind speed (in which H=0.61 and q=15) and solar radiation processes (where H=0.81 and q=2) since Wind and Solar energy are the most popular renewable energy resources worldwide. Two models were applied for the calculation of the short-term predictability time horizon. The first one was a benchmark model while the second was a more sophisticated stochastic. More information about the models will be elaborated later on this chapter.

The data related to wind processes were once again from the MIT database since it provided full and wide-lengthen timeseries. When examining solar irradiance data it appeared that processing data from 40 stations would be time consuming and complex. Across the 40 stations, station No35 in Eugene Airport, also known as Mahlon Sweet Fields, located in Lane County, Oregon U.S.A was chosen to apply tests related to the investigation. After examining the data from all stations, data from Eugene Airport database were the most accurate since it contained the most recorded data with the least gaps. Thus, it was efficient to produce more efficient results since a reliable database was exploited.

Both wind and solar processes may reflect the phenomenon of double periodicity which is attributed to the orbitals that Earth performs, creating the changing from day to night and the sequence of seasons within a year (Figures 4.1, 4.2, 4.3, 4.4). Therefore, wind speed and solar radiation are prone to double cyclo-stationarity, both diurnal and seasonal (Koudouris, 2017; Gkolemis, 2019). More analytically, when the statistical parameters (mean and standard deviation) of a process maintain unchanged over time, the process is characterized as stationary (Koutsoyiannis, 2013). Researches have already examined the hourly and monthly statistic parameters of solar irradiance and wind processes, proving that it can affect the behavior of the respective natural processes through time but it can also be harnessed by applying certain models and simulating them through specific distributions.

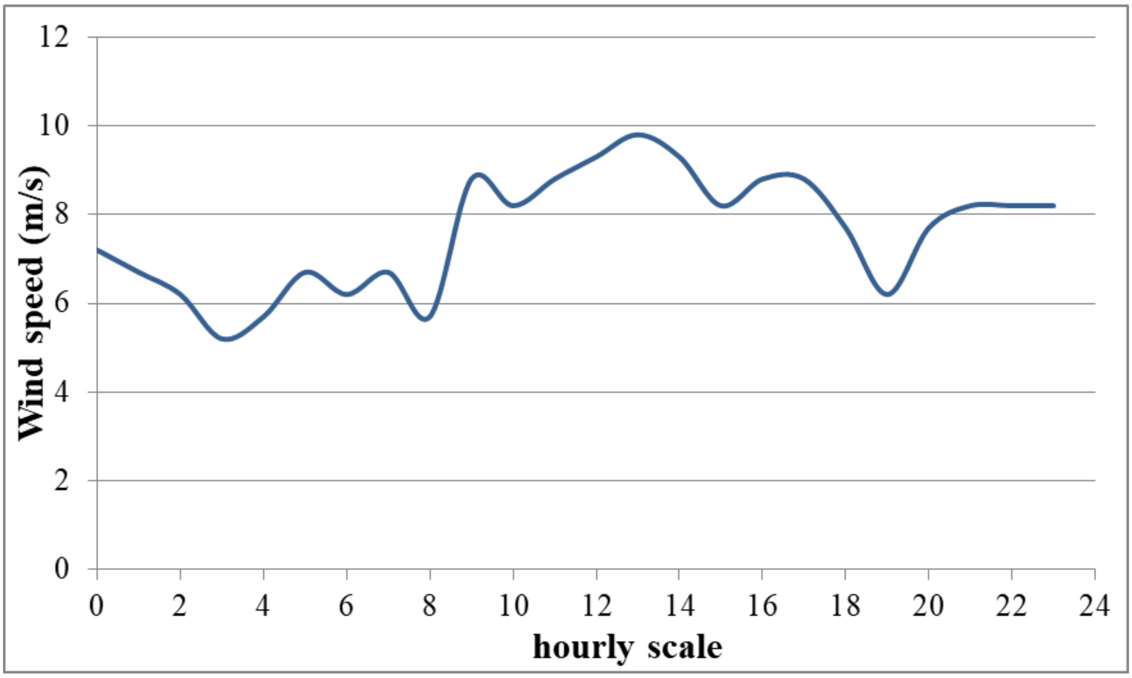


Fig.4.1: *Illustration of diurnal periodicity effect on wind speed process.*

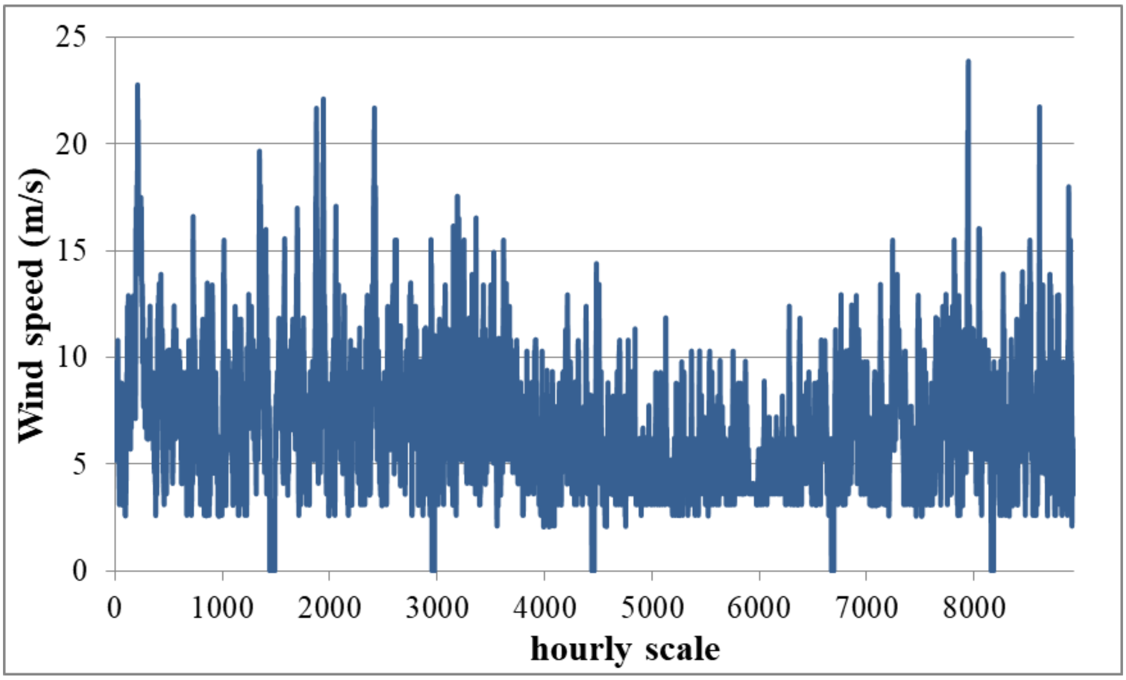


Fig.4.2: *Illustration of seasonal periodicity effect on wind speed process.*

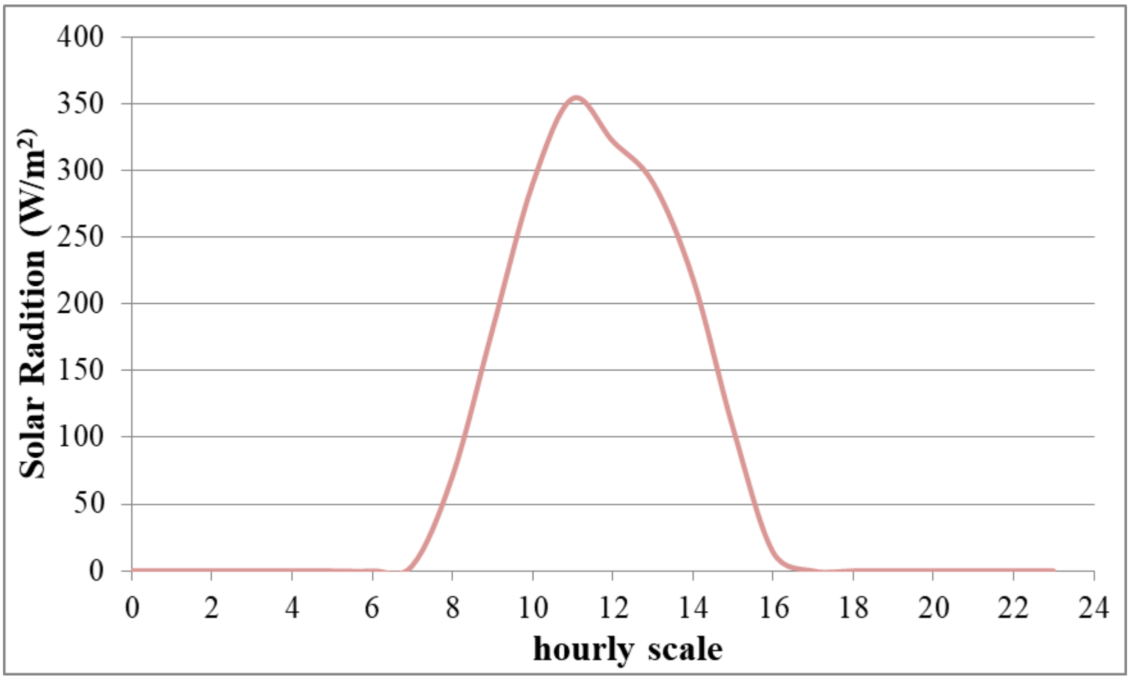


Fig.4.3: *Illustration of diurnal periodicity effect on solar irradiance process.*

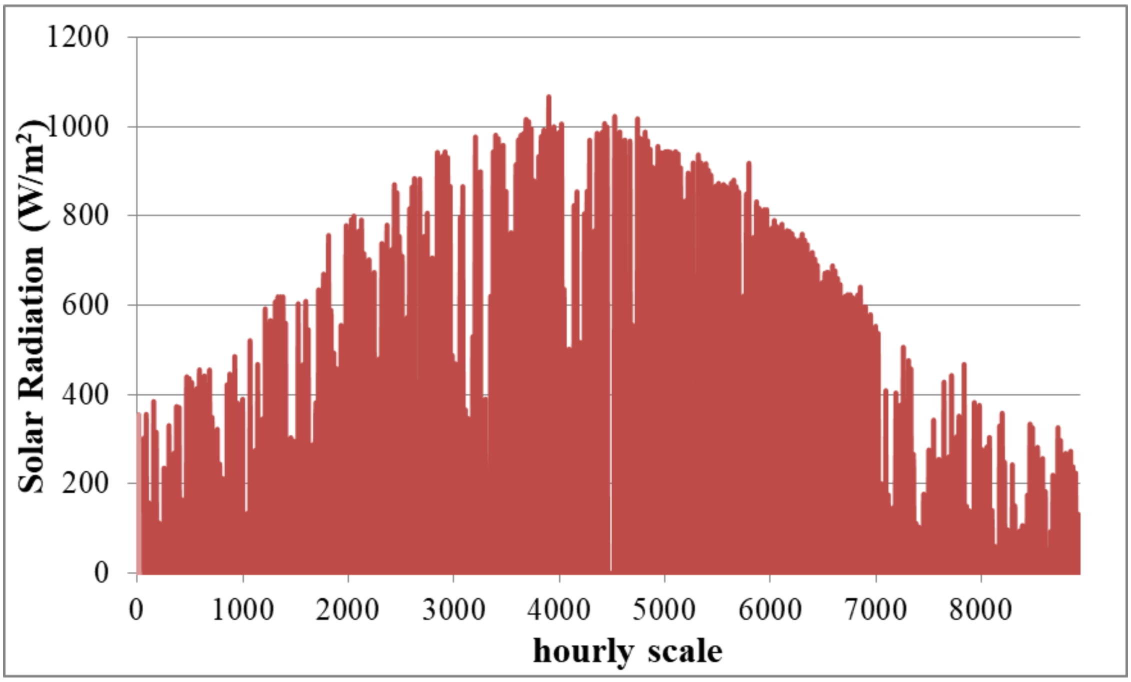


Fig.4.4: *Illustration of diurnal periodicity effect on solar irradiance process.*

In order to examine whether double cyclo-stationarity would sway the performance of the forecasting methods, a simple method was applied to restrict this effect; After calculating the statistical parameters (μ=mean: central tendency and σ=the standard deviation: the square root of variance) of each of the two timeseries of interest on the hourly and monthly scale (since the effect refers to diurnal and seasonal periodicity), the simplified expression below was implemented in each value; the mean was subtracted at each value examined and the difference was divided by the standard deviation:

$$\acute{s} = \frac{s_d - \mu_{i_j}}{\sigma_{i_j}} \qquad (\textbf{\textit{Eq.4.1}})$$

where $\acute{s}$ = value of interest (in this case either wind speed or solar radiation), μ=mean and σ=standard deviation. The *i, j* indexes refer to the respective hour and month (i= hour, varying from 1 to 24 and j= month, varying from 1 to 12).

In the following section of this chapter, several prediction models were tested in order to identify the degree of predictability short-term. The tool that was used to determine the time-window of predictability throughout the two processes was the Nash and Sutcliffe coefficient.

The Nash–Sutcliffe efficiency coefficient is an index that determines the efficiency of model outputs. Its contribution helped examine whether the predictions produced from the Benchmark models and the analogue model were accurate and to what extent. The Nash and Sutcliffe coefficient (abbreviated as *F*), expressed from the mathematical equation below, varies from $-\infty$ to 1.

$$F = 1 - \frac{\sum_{d=1}^{n} \sum_{i=0}^{bd} \left(\tilde{s}(i) - s_d(i)\right)^2}{\sum_{d=1}^{n} \sum_{i=0}^{bd} \left(\bar{s}(i) - s_d(i)\right)^2} \qquad (\textbf{\textit{Eq.4.2}})$$

where $\tilde{s}$= the prediction on lead time , s_d = data of interest , where d is an index for the sequential number of the experiments, $\bar{s} = timeseries\ mean$, the double sum is in case where n timeseries similar to the original one are applied, i denotes time, bd is the total number of recorded frames in the d^{th} experiment.

Before applying the analogue model which is the main forecasting model that it is examined in the current thesis, two naive benchmark models where applied in order to compare their results with the performance of the ensemble model.

The first Benchmark model, abbreviated as B1, follows the hypothesis that the predicted state $\tilde{s}$ of the examined value of interest is equal to the mean state of the timeseries $\bar{s}$:

$$\tilde{s}\big((t+l)\big)\Delta = \bar{s} \qquad (\textbf{\textit{Eq.4.3}})$$

where tΔ is the present time in s (t →dimensionless time), lΔ the lead time of prediction in s (l > 0) and Δ the sampling frequency (in this case Δ=1).

Since the prediction constantly coincides with the mean of the timeseries applying the B1, the F coefficient will always get zero values as an output. Consequently the B1 model was applied to illustrate that whenever the Nash and Sutcliffe coefficient produced from each forecasting model is equal to zero, F=0, the model predictions are as accurate as the mean. It should be noted that by the time F≤0, the model accuracy aggravates significantly, indicating that the model is useless from that point. That means that a model's accuracy can be defined while the F coefficient varies from 0 to 1, 0≤F≤1. For that reason, it was considered redundant to depict B1 on diagrams since the horizontal axis indicated the limit of efficiency. Therefore no negative values will be depicted while illustrating the Nash and Sutcliffe coefficient.

Another naïve benchmark model was tested for natural processes forecasting and for short it will be abbreviated B2. This benchmark model could calculate aν indicative, minimum horizon of predictability. Specifically, its main use was to basically compare it with the stochastic model applied later, to ascertain if there is any room for improvement on the predictions. This model follows the principle that the prediction $\tilde{s}$ of a past state is the current value s, ignoring the how long the timeseries can be:

$$\tilde{s}\big((t+l)\big)\Delta = s \qquad (\textbf{\textit{Eq.4.4}})$$

where again tΔ is the present time in s (t →dimensionless time), lΔ the lead time of prediction in s (l > 0) and Δ the sampling frequency (Δ=1) (just in the B1).

B2 was tested for both wind speed process (Figure 4.5) and solar radiation process (Figure 4.6), including both the initial (as retrieved from the database, with real time data) and the standardized timeseries (after reducing the effect of double cyclo-stationarity). The results can be illustrated on the next figures:

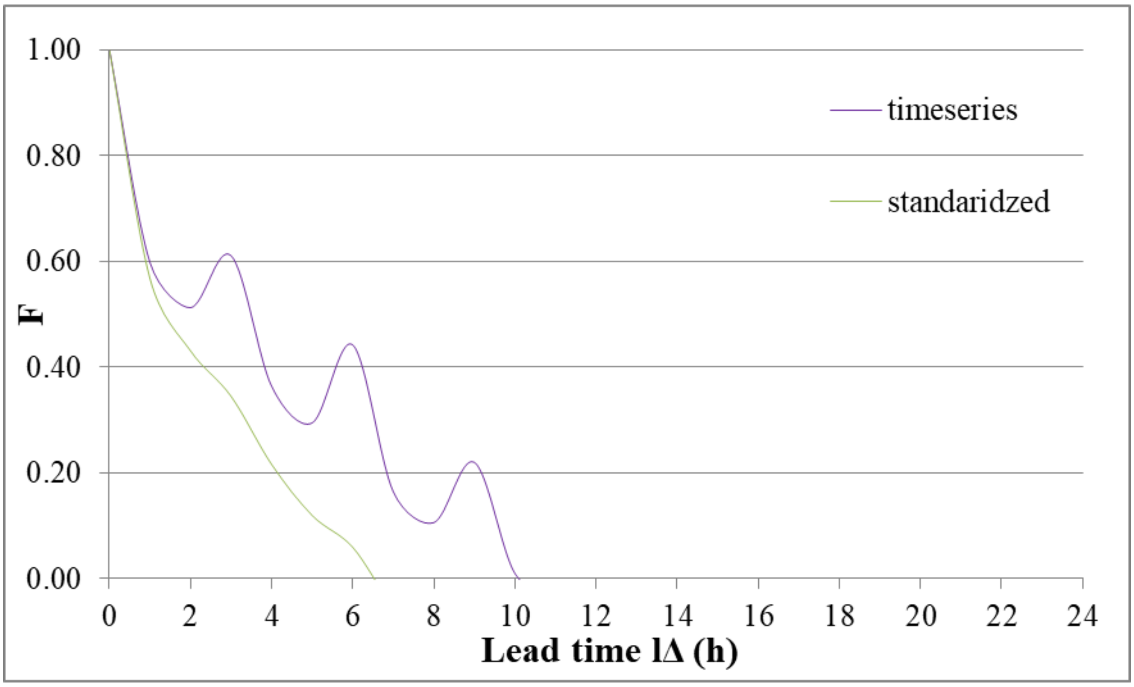


Fig.4.5: *Application of B2 benchmark model on wind speed timeseries both the initial and the standardized one.*

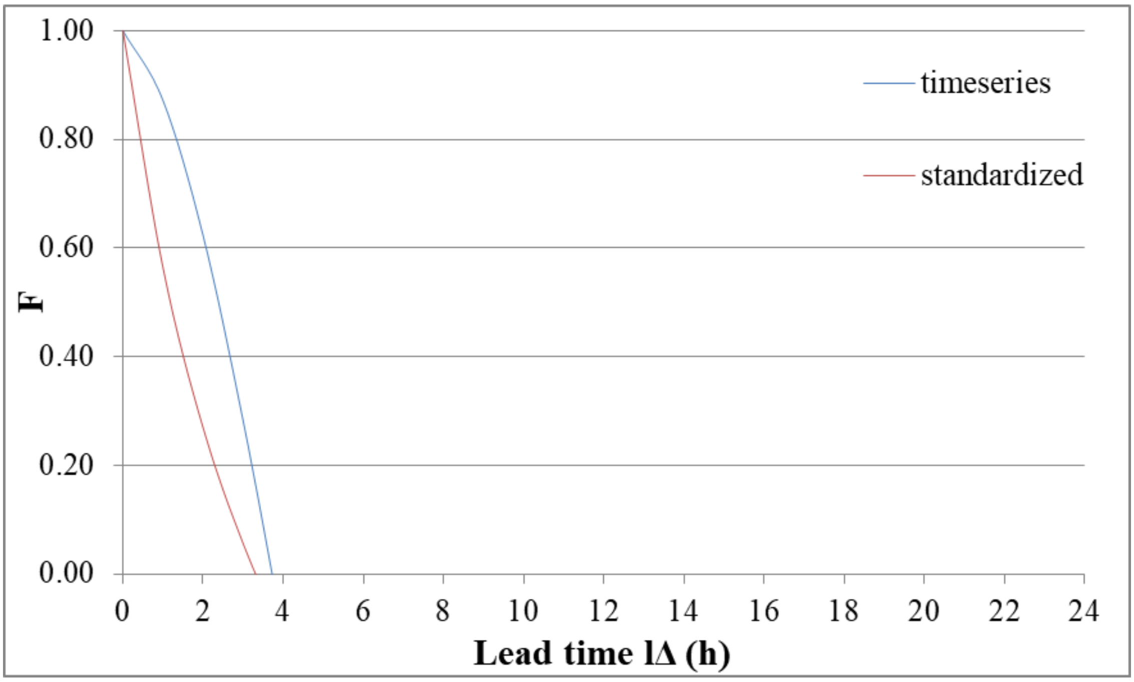


Fig.4.6: *Application of B2 benchmark model on solar radiation timeseries both the initial and the standardized one.*

As far as the wind speed process is concerned, the B2 model seems to outperform on the initial timeseries calculating a time-window of 11h, while the standardized one reduces the horizon of predictability to approximately 7.5h. On the other hand, although solar irradiance is supposed to be affected more conspicuously by double cyclo-stationarity, the application of the B2 depicted that before after the standardization on the original timeseries the predictability time-window was approximately the same, approaching 4 hours.

The B2 was actually tested in order to compare it with the stochastic analogue-ensemble forecasting model which is a very popular and quite simple model applied for forecasting. It has already been applied in multiple processes and chaotic systems, including natural processes and other kinds of experiments (it has also been applied on a recent study where the predictability window was aimed to be calculated after throwing a dice) (Dimitriadis et al., 2016). Therefore, it was chosen to create short-term predictions of natural processes when it comes to power generation. The application of the model on natural processes has proven so far that it corresponds efficiently.

The analogue model applied (also known as deterministic data-driven model) (Dimitriadis et al., 2016) is based on the hypothesis that natural processes include repeating patterns in the past, following an almost deterministic behavior. Thus, in order to produce the predicted values backdated, natural processes data is examined.

First of all, the analogue model uses a percentage, *p*, of the timeseries in order to make validation of the predicted states. The remaining percentage of the timeseries is the "past" which is investigated in order to pinpoint the repeating past states. The prediction of future states $s((t+l)\Delta)$, depends on *h* past states $s((t-r+1)$, $r \in [1,h]$. Each past state is a group of *k* similar states of the timeseries which are also called neighbors or analogues of each value examined $s_j\left((t_j-r+1)\Delta\right)$. Afterwards, all the neighbors are compared to the value of interest with the condition that each neighbor ($j \rightarrow \in [1,k]$) for all *r* conforms to this mathematical expression :

$$\left\|s_j\left((t_j-r+1)\Delta\right)-s\left((t-r+1)\Delta\right)\right\| \leq g \qquad \textbf{(\textit{Eq.4.5})}$$

where *g* is the error defined to ensure the degree of reliability of the forecasting.

Once all the neighbors are examined with the criteria that they do not differ longer than the predefined threshold, the predicted state $\tilde{s}$ at lead time *lΔ* can be generated as:

$$\tilde{s}\big((t+l)\Delta\big) = \frac{1}{k} * \sum_{j=1}^{k} s_j\left(\left(t_j + l\right)\Delta\right). \qquad \textbf{(\textit{Eq.4.6})}$$

For the efficiency of the analogue model when applied on the relative natural processes, the percentage p was estimated at 30% for the part of the validation of the prediction.

The lead time in which the time window of predictability would be detected included 24 hours. The selection of the time range in which the time window would be detected was based on the hypothesis that short –term forecasting could optimize energy systems operation by introducing a day-to-day operational planning since natural processes from which renewable energy resources derive from. Additionally, after an extend research on the literature it was concluded that a 24-hour ahead forecasting has been a subject of interest on the scientific community before aiming the optimization of energy systems' performance(Cococcioni et al.,2011, Adeshina et al., 2017). Multiple and long preliminary tests evaluating the estimating error were performed. For the wind speed timeseries investigation the estimating error was calculated g=0.5, relying on the accuracy of the corresponding measuring instruments of wind. The same tests were followed on the definition of the threshold of solar irradiance where it was calculated g=0.5 too.

Since the forecasting model would be applied on the standardized timeseries too, the same procedure was repeated to calculate the estimating error $\acute{g}$ followed by subtracting mean from each error and eventually their difference was divided by the standard deviation. Additionally, the error was also properly rounded to match the new range of data on the standardized processes. Consequently, the results were: $\acute{g}_w$=0.15 for the wind timeseries and $\acute{g}_s = 0.1$ for solar process.

An additional attempt to optimize properly the computational error of the analogue model led to the application of an objective function. By applying pessimistic and average threshold scenario on the timeseries the results were modified accordingly, following the behavior of each process explored.

A sensitivity analysis on Matlab was applied to examine the reliability of the forecast over a number of multiple neighbors h of the past states of the value tested. After initializing the application of the stochastic model on the long timeseries, the computational burden was vast for Matlab to process since the duration of each test would last over month, delaying the investigation dramatically. Consequently, a pilot application of the deterministic-driven model was performed on a year of each corresponding timeseries.

The criteria for selecting a reliable year which would correspond sufficiently was to present similar statistical parameters with the original long timeseries and also include as many documented data as possible, avoiding to generate inaccurate forecasting due to lack of measurements. Since the statistical parameters, mean and standard deviation, of the original wind speed timeseries were calculated μ= 5.45 and σ=2.48 respectively, a year with similar features (μ=5.77 and σ=2.78) was finally chosen, including a subtle percentage of gaps appearing measurements. Additionally, the statistical parameters of the solar radiation process was estimated μ=159.42 and σ=250.54 while in the year that will eventually be explored the statistical parameters are μ=159.09 and σ=248.53 also presenting an insignificant lack of data that could possibly distort the results. When evaluating this decision over the examination of the standardized timeseries, the results were quite satisfying; on the standardized wind speed process, the mean was calculated μ=0 and the standard deviation σ=0.98 while in the single-year standardized timeseries μ=0.91 and σ=1.06. Furthermore, solar radiation standardized timeseries statistical parameters were μ=0 and σ=0.91 while in the selected year presented μ=0.05 and σ=0.94.

Since the timeseries of interest were significantly modified, the benchmark model B2 (Figures 4.7-4.8) was tested once more on the single-year timeseries, presenting new results on the figures below.

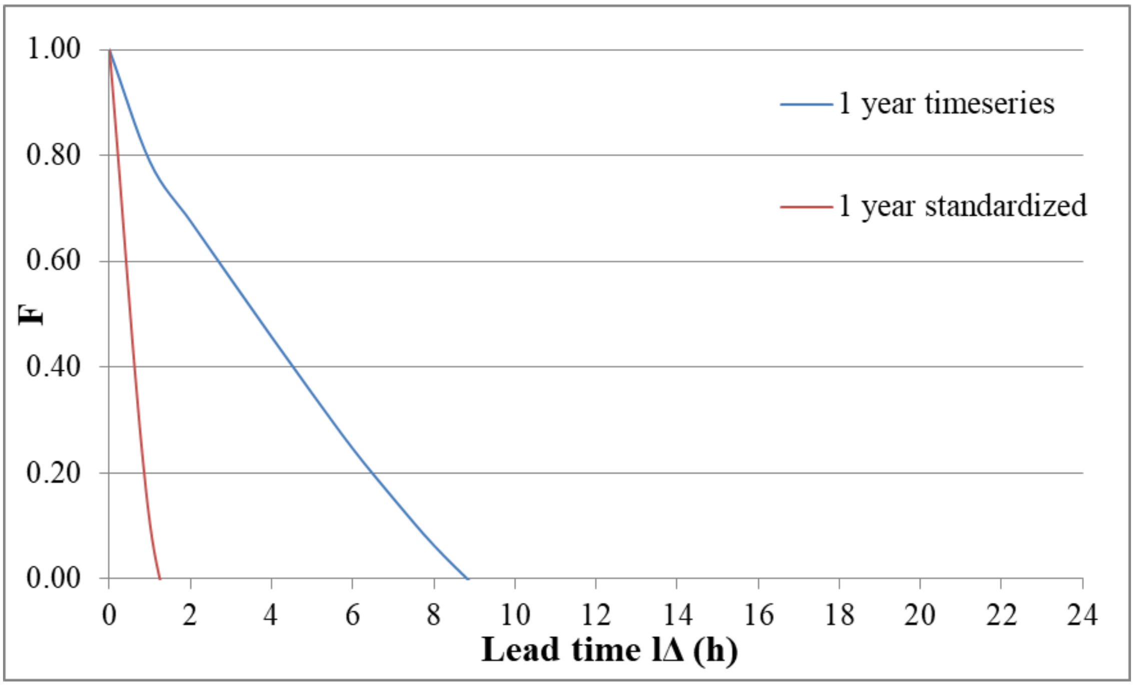


Fig.4.7: *Application of B2 benchmark model on wind speed single-year timeseries both the initial and the standardized one.*

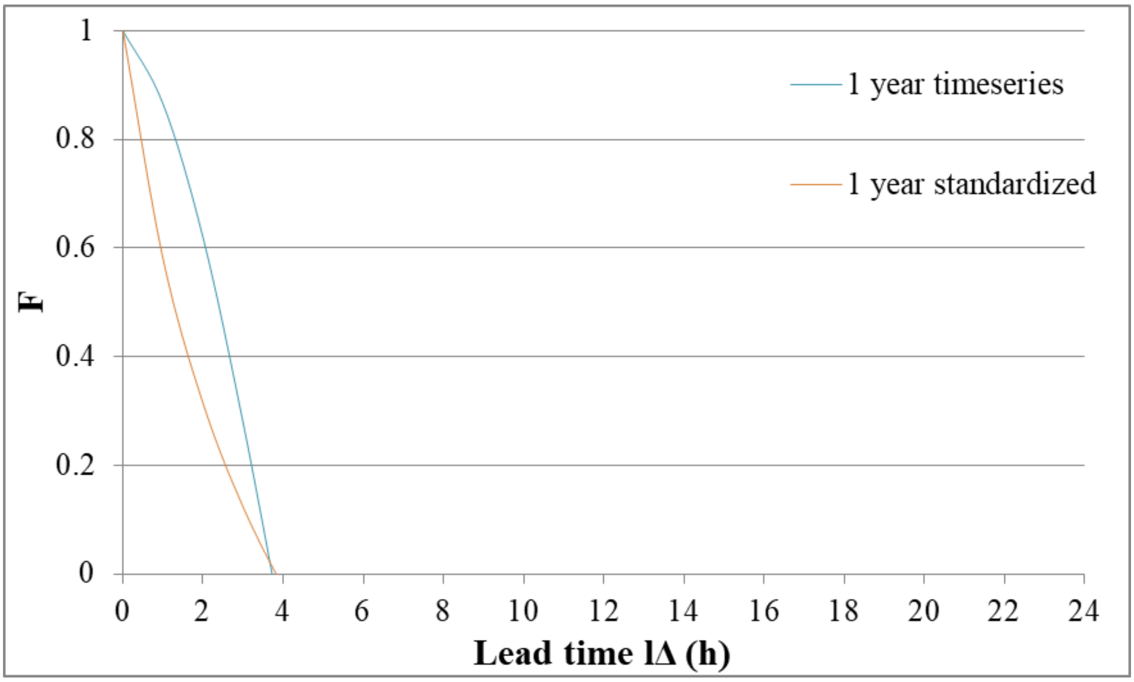


Fig.4.8: *Application of B2 benchmark model on solar radiation single-year timeseries both the initial and the standardized one.*

While the application of B2 on the full-length timeseries of wind speed calculated the time window about 11 hours, when reducing the data to one year the predictability window is remotely limited to 9 hours. However the results from the single-year timeseries deteriorates dramatically to 1.5 h compared to the standardized timeseries with total data. Furthermore, the performance of the B2 model when tested on single-year solar irradiance timeseries did not change significantly, calculating the time window to approximately 4 hours while in this case the standardized single-year timeseries subtly outperformed over both the initial total timeseries and the single-year timeseries before the standardizations.

Furthermore, the next step to the identification of the time horizon where determinism conquers unpredictability in the natural processes of interest is the presentation of the sensitivity analysis performed on the analogue model after testing it for various analogues (neighbors) varying from 3 to 20 (Figures 4.9~4.16).

Once more when the application of the analogue model started, the computational burden was significant as well leading to tests that lasted very long (Table 4.1). Taking into consideration that the model was applied for four timeseries (wind speed process, wind speed standardized process, solar radiation process, solar radiation standardized process) each one tested for both scenarios of the error, the time needed to complete the tests was multiplied by eight.

Table 4.1: Indicative duration of the application of the analogue model when tested on the single-year timeseries

Neighbors (Analogues)	**Duration**	
h	***hours***	***minutes***
3	2	25
4	2	30
5	2	35
6	3	0
7	3	10
8	3	30
9	3	40
10	4	0
11	4	10
12	4	25
13	4	50
14	4	0
15	5	10
16	5	30
17	6	0
18	6	15
19	6	0
20	6	30

Total time ~ 78 h --> 3.25d

When applying the sensitivity analysis by testing the analogue model and the second benchmark model, the results generated where representative only for a specific number of neighbours. While increasing the number of analogues h, the forecasting would produce more accurate predictions. However, the limited length of the timeseries (since only one year was tested eventually) started to affect the sensitivity analysis after a given point. By significantly increasing the amount of the analogues, which would be explored to define the predicted state, the length of the timeseries would be further limited, generating less and worse predictions. Consequently the *F* coefficient would approach the value of 1, implying that the accuracy of the prediction was improved which is misleading since the results where the abrupt improvement of the F exponent were fictional. The following figures show the results of the sensitivity analysis that gave realistic results, omitting the amount of neighbours that generated fictional predictions

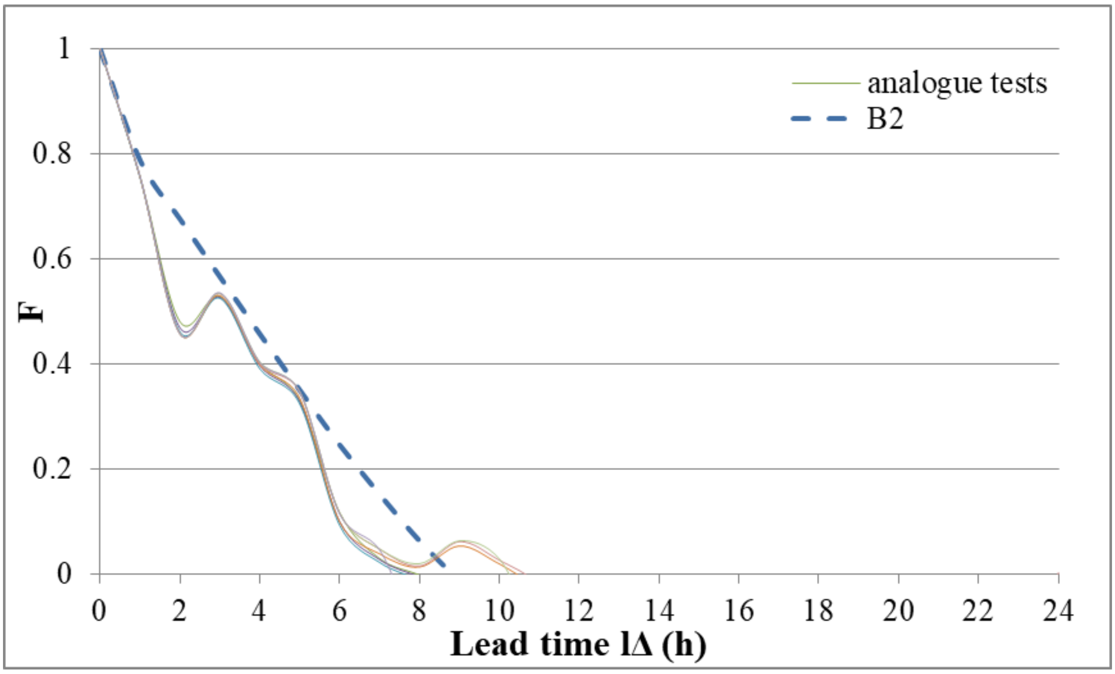


Fig.4.9: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on wind speed from the station in the U.S.A.*

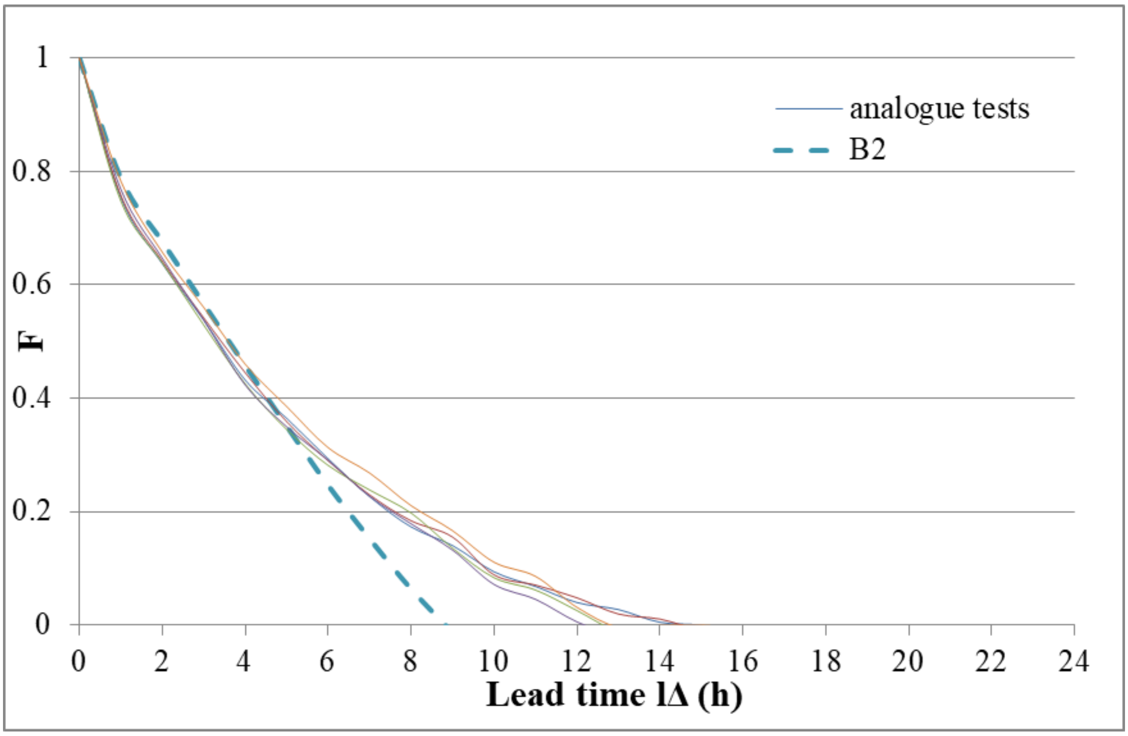


Fig.4.10: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on wind speed from the station in the U.S.A.*

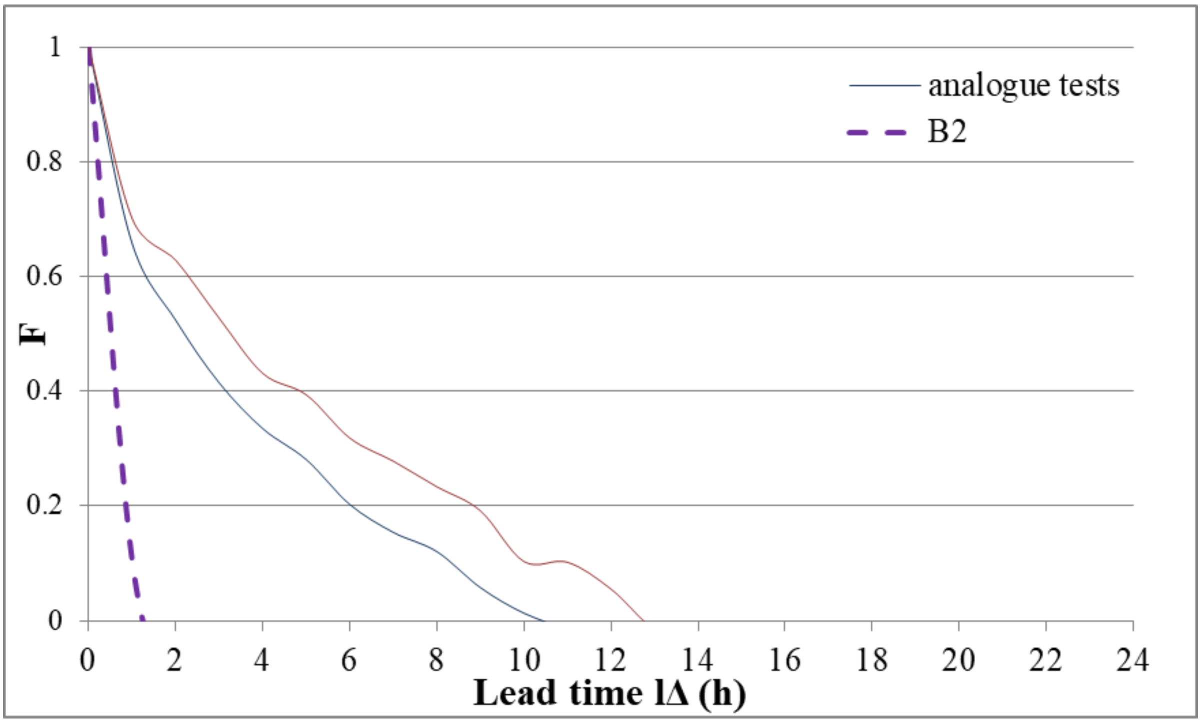


Fig.4.11: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on standardized wind speed from te station in the U.S.A.*

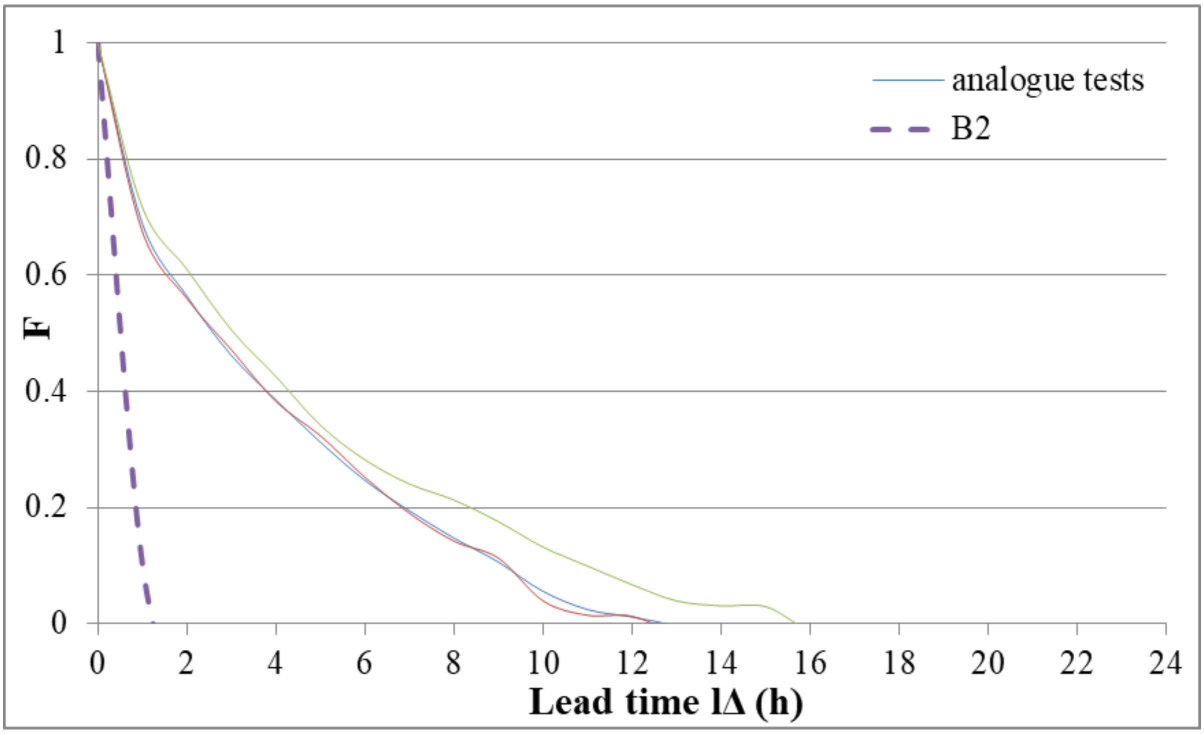


Fig.4.12: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on standardized wind speed from the station in the U.S.A.*

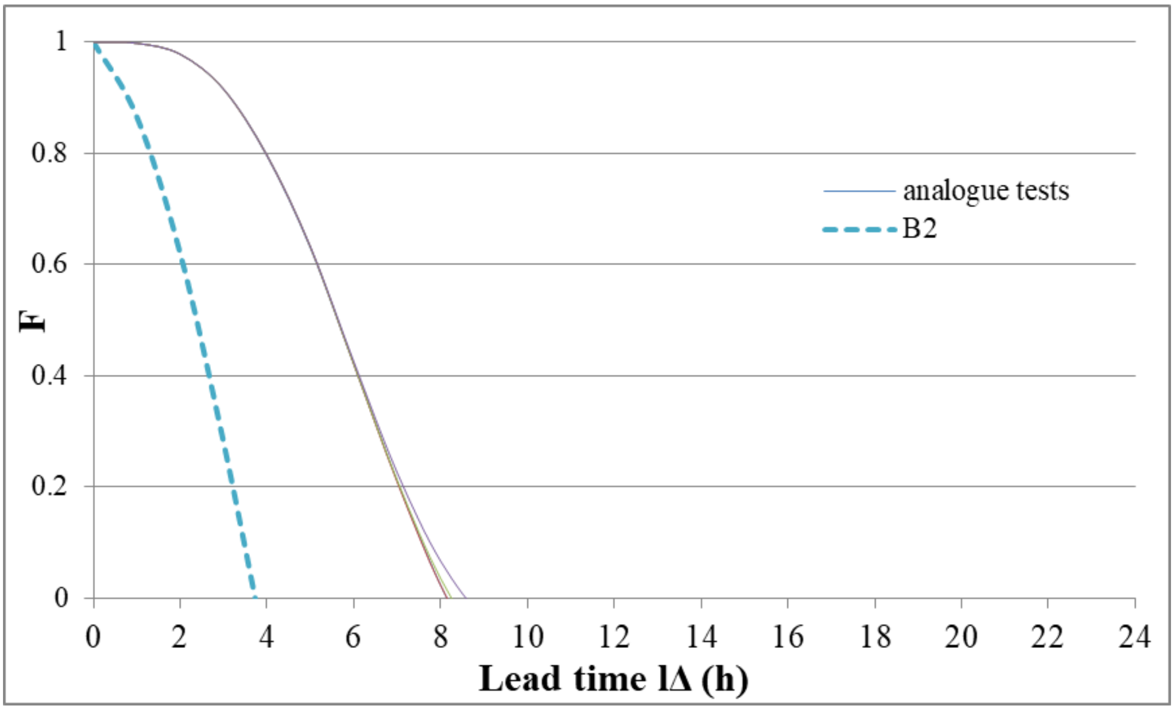


Fig.4.13: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on solar radiation from the station in the U.S.A.*

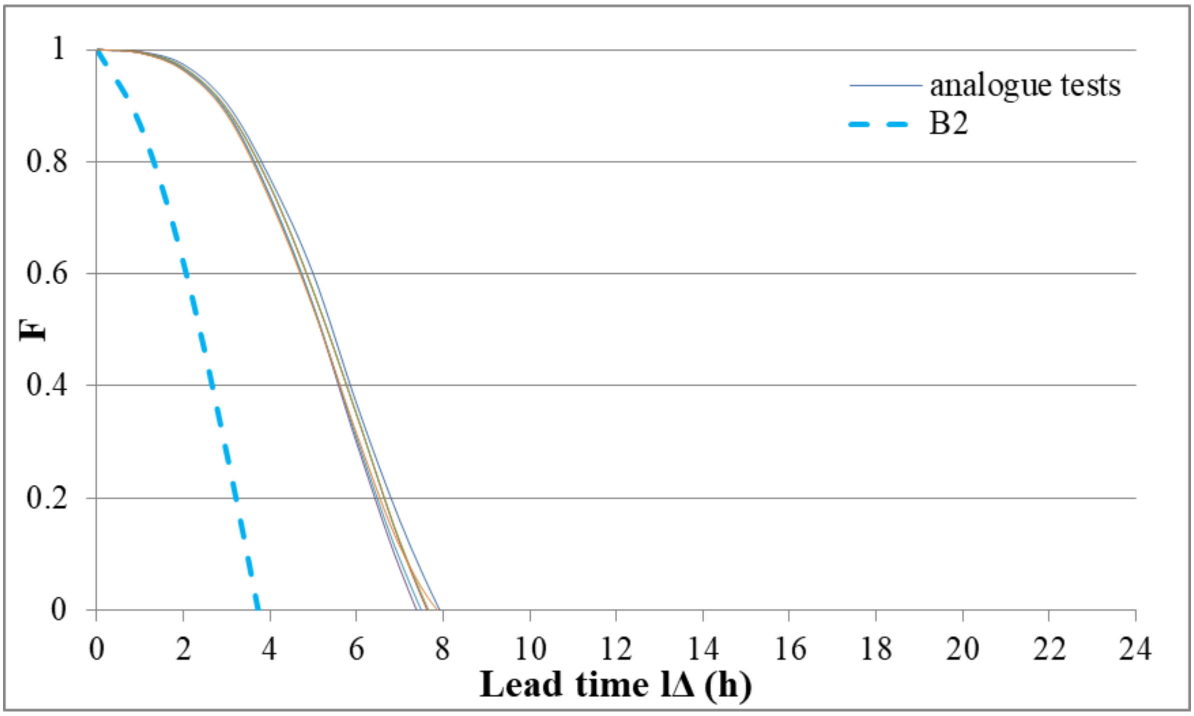


Fig.4.14: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on solar radiation from the station in the U.S.A.*

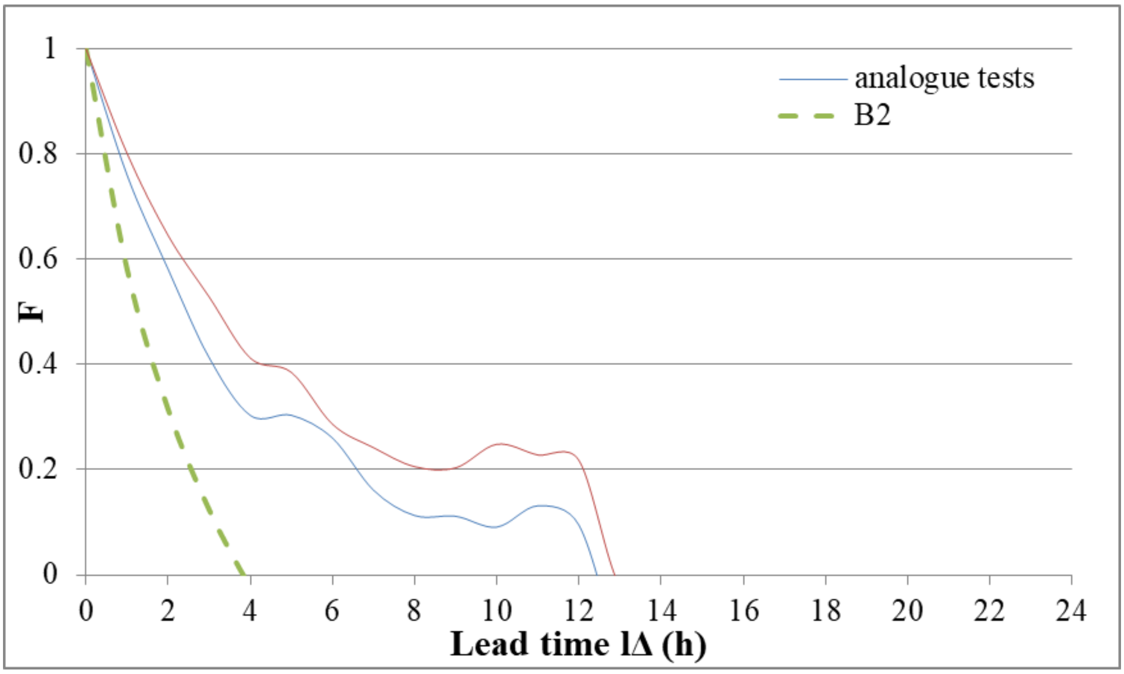


Fig.4.15: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on standardized solar radiation from the station in the U.S.A.*

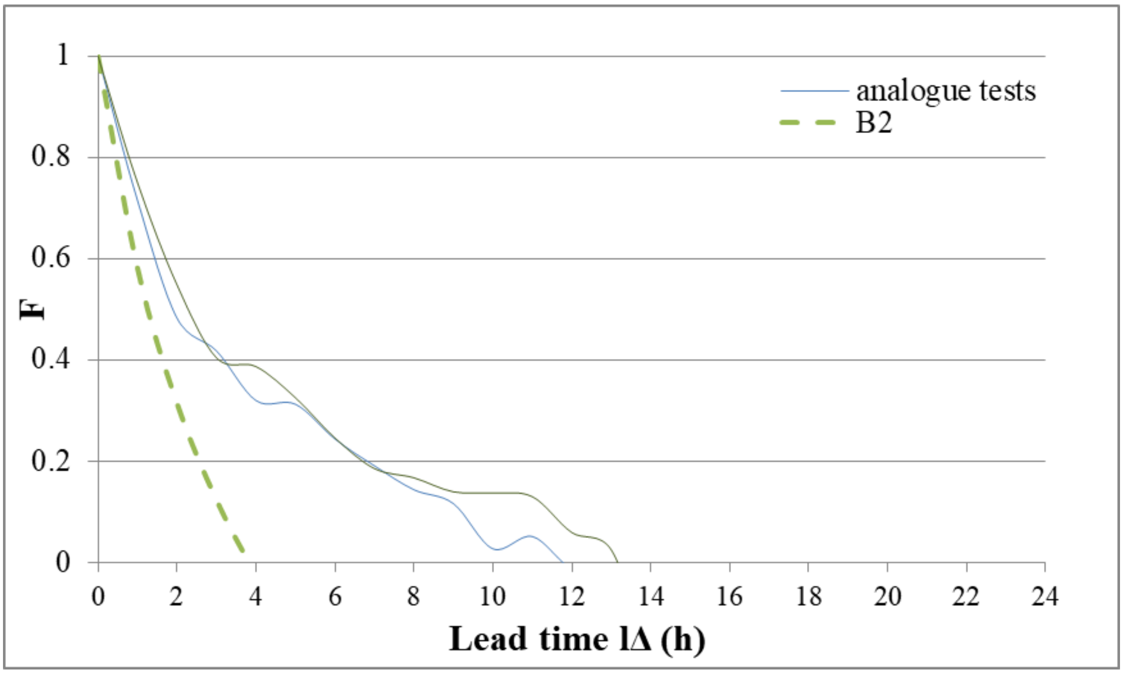


Fig.4.16: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on standardized solar radiation from the station in the U.S.A.*

As far as wind speed process is concerned, the analogue model outperformed over the B2 that proved it could foresee 9 hours, providing a wider time-window of predictability. In the case where the error was more pessimistic (Figure 4.9), the predictability window can reach up to 11 hours. However, the average error may perform a prediction of approximately 14 hours (4.10) which is a respectable improvement of the analogue performance compared to the other tests.

Moreover, the standardized wind process, where the double cyclo-stationarity effect is restricted, the average error on the sensitivity analysis maximized the lead time of prediction up to 16 hours (Figure 4.11), while when using pessimistic error would predict accurately 13 hours (Figure 4.12), taking into account that it is a more conservative condition. Once more the analogue model performed more efficiently against the benchmark model that had a limited time window of just 1 hour.

On the solar radiation sensitivity analysis, the alteration of the error did not seem to affect significantly the performance of the model considering that in both tests the prediction reached the 8 hours (Figure 4.14), although the more conservative error generated a slightly longer time horizon of predictability (Figure 4.13). Contrarily, the naïve benchmark model would foresee only 4 hours ahead.

Finally, the investigation of the predictability window on the standardized solar irradiance process gave interesting results. As presented in the sensitivity analysis of the non-standardized timeseries, the selection of the error scenario, did not alternate the duration of valid forecasting since they both calculated a time window lasting approximately 13 hours (Figures 4.15-4.16). This last test verified that in both processes, wind speed and solar radiation, the analogue-ensemble model seems to be more efficient compared to B2.

4.2 Exploration of predictability horizon through the climacogram

As it has already been noted in previous chapters, through the investigation of the climacogram it is feasible to determine the degree of inherent uncertainty on hydrometeorological and geophysical processes. Applying the sensitivity analysis on wind speed and solar radiation processes proved that determinism can actually prevail over unpredictability for a limited amount of time. Focusing on the exploration of the climacogram (defined as the log-log plot of variance of the averaged process versus averaging time scale) on shorter scales could provide a preliminary calculation of the predictability time-window

It is reminded that sensitivity analysis tests denoted that the analogue ensemble- model combined with the average threshold displayed more satisfactory results. Thus, only those are being evaluated and the conclusions are presented in the table below. When examining the climacogram in shorter scales, it is observed that when variance begins to take values lower than approximately 0.5, the correlation between the values (from either wind speed or solar radiation processes) deteriorates in terms of the climacogram, indicating that determinism starts to fade from that point on. Sensitivity analysis verifies that the time window can be detected within that period of time, where there is adequate correlation between the values, since the Nah and Sutcliffe efficiency coefficient takes values greater than 0 (zero).

Another aspect to study predictability is to observe that when the *F* coefficient aggravates taking negative values, so does the correlation between values of the examined natural processes. This hypothesis proved to be valid by investigating either wind speed process or solar radiation behavior. The results can be shown on the graphs and tables below where the climacogram is compared with the Nash and Sutcliff efficiency coefficient F.

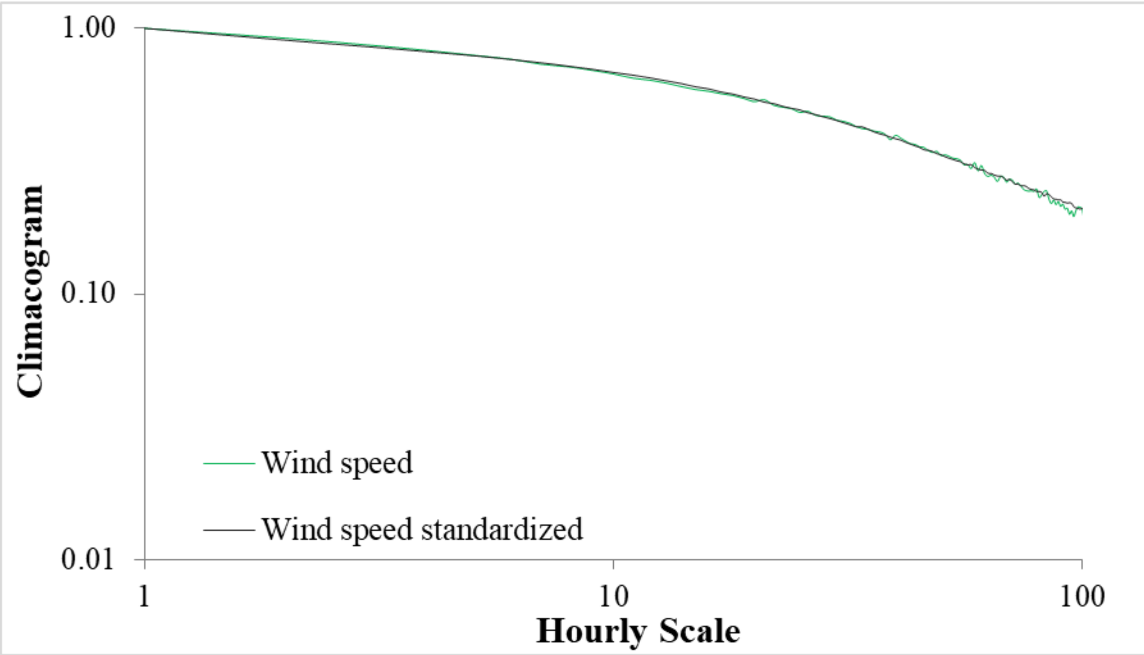


Fig.4.17: *Climacogram of wind speed focusing on short-term behavior.*

Table 4.2: *Climacogram and Nash and Sutcliffe efficiency coefficient correlation in wind speed process.*

	Wind speed data MIT		**Wind speed standarized data MIT**	
	climacogram (γ)	**Nash and Sutcliffe efficiency coefficcient (F)**	**climacogram (γ)**	**Nash and Sutcliffe efficiency coefficcient (F)**
3h	*0.67*	*0.56*	*0.85*	*0.51*
6h	*0.52*	*0.31*	*0.76*	*0.28*
9h	*0.45*	*0.17*	*0.70*	*0.18*
12h	*0.39*	*0.03*	*0.65*	*0.07*
15h	*0.36*	*<0*	*0.58*	*0.029*
18h	*0.34*	*<0*	*0.54*	*<0*
21h	*0.32*	*<0*	*0.51*	*<0*
24h	*0.30*	*<0*	*0.48*	*<0*

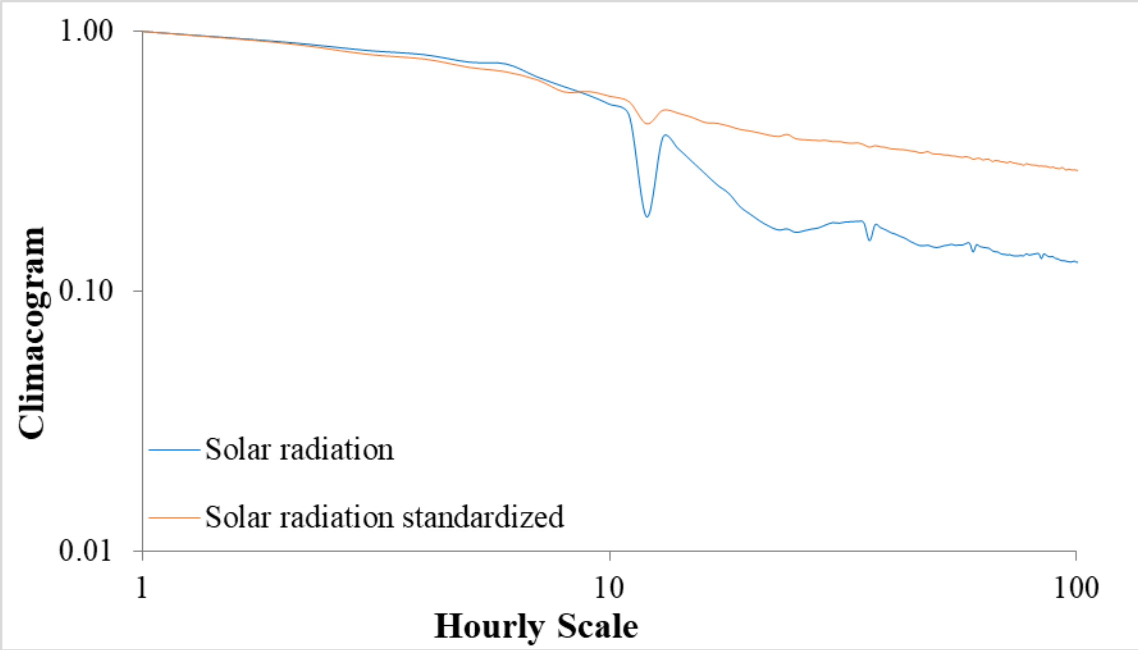


Fig.4.18: *Climacogram of solar radiation focusing on short-term behavior.*

Table 4.3: *Climacogram and Nash and Sutcliffe efficiency coefficient correlation in solar radiation process.*

	Solar radiation data from station No35		Solar radiation standarized data from station No35	
	climacogram (γ)	**Nash and Sutcliffe efficiency coefficcient (F)**	**climacogram (γ)**	**Nash and Sutcliffe efficiency coefficcient (F)**
3h	*0.85*	*0.9*	*0.82*	*0.4*
6h	*0.75*	*0.34*	*0.70*	*0.25*
9h	*0.57*	*<0*	*0.57*	*0.14*
12h	*0.19*	*<0*	*0.44*	*0.059*
15h	*0.32*	*<0*	*0.47*	*<0*
18h	*0.24*	*<0*	*0.43*	*<0*
21h	*0.19*	*<0*	*0.40*	*<0*
24h	*0.17*	*<0*	*0.40*	*<0*

5. Investigating the degree of predictability in the case study of Astypalea, Greece

5.1 The significance of uncertainty for the green energy autonomy in non-connected Greek islands

Most of the Greek islands, especially remote Aegean islands, are not connected to the electricity network of the mainland and depend only on oil-fuelled power plants, which have a high oil import cost and a high environmental impact. The continuous advances in renewable energy technology along with the gradual reduction of installation costs, pave the way towards a wider adaptation of renewable energy. Effective planning of a renewable energy resources system requires the investigation of renewable resources and energy demand. Such a case study was implemented in the remote island of Astypalea with the expectation to create a sufficient and autonomous energy system on a non-connected island.

The non-connected island of Astypalea is located in the Aegean Sea in Greece and it is the fourth largest island that belongs to the Dodecanese islands (Figure 5.1). It covers an area of 97 km². Its coastline extends to 110 km. It is located west of Nisyros and east of Anafi (Cyclades). The recorded distance of Astypalea from the island of Kos is 23 nm, from the island of Rhodes 96nm and from Piraeus 117 nm. The airport of Astypalea connects the island directly with Athens, Leros, Kos and Rhodes.

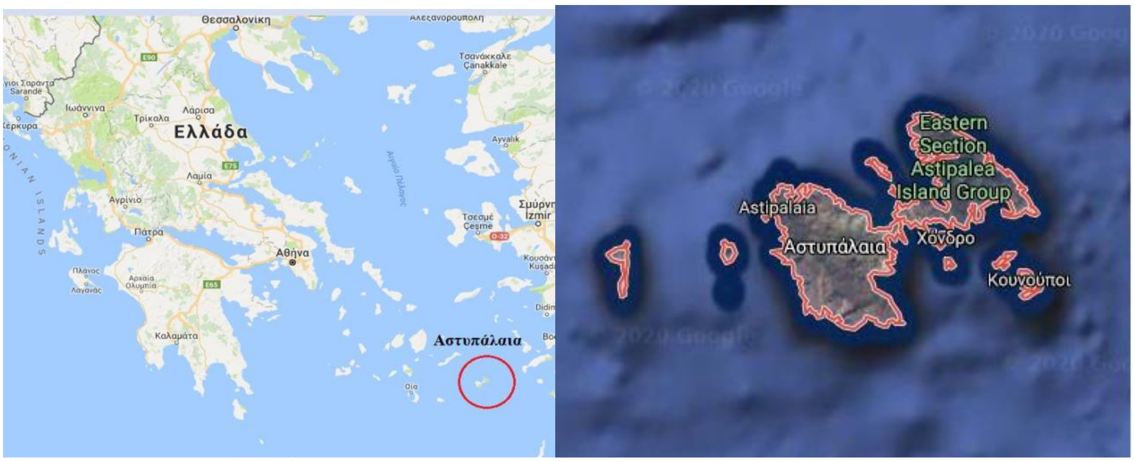


Fig.5.1: *Location of the island of interest Astypalea.*
[Source: earth.google.com- maps.gogle.com]

While the municipality of the island belongs to the patchwork of Dodecanese islands, Astypalea geographically is closer to the Cyclades. Thus, it operates as a geographical and cultural “bridge” that connects the two patchworks.

The island of Astypalea counts about 1.334 residents (according to demographical data). However, it can annually attract a considerable amount of 20.000 tourists from all over the world. As a result, the greater percentage of energy is consumed to cover residential needs as well as the needs of tourism enterprises. Consequently during the summer season energy demands increases dramatically. Electricity system of Astypalea that is not interconnected to the mainland, , use low-efficiency autonomous oil stations (diesel and fuel oil) installed since the 1960s and 1970s, with local production plants operating at maximum load to meet energy demands especially during the summer months due to sharp increase of the number of tourists.

An explicit factor of great significance is the high cost of energy on the non-connected islands compared to the expanses on the mainland. The average cost of electricity production for the autonomous power stations according to the official data of HEDNO in August 2017, stood at 336.96 € / MWh, which is about seven times the continental system limit value, which was around 50 € / MWh (Table 5.1).

Table 5.1: *Average power generation cost for non-connected islands.*
(Source: HEDNO, 2016)

Island	Average Production Cost (€/MWh)
Astypalea	286,06
Anafi	390,2
Antikythera	945,72
Ikaria	379,83
Symi	245,12
Rhodes	173,77
Milos	146,20
Paros	114,49

The non-connected island of Astypalea is a case of great interest that has triggered other researchers in the past as well. Since there are no installed stations on the island that measure meteorological data, the data used in this case study were obtained from a previous investigation that was presented in the EGU 2017, Vienna (Chalakatevaki et al, 2017) where an electricity mix in the island of Astypalea was created. Measurements from meteorological stations located in islands and regions nearby the island were exploited to produce synthetic timeseries which are applied in this study.

In the case study of Astypalea six renewable resources were examined (solar, wind, marine, hydropower, biomass and geothermal) in order to create a sufficient energy mix (Chalakatevaki, 2017). The hydrometeorological variables were treated as stochastic processes with deterministic components. In the absence of meteorological stations at the exact location, the diurnal and seasonal periodicity of temperature and dew point were estimated from nearby stations (Hadjimitsis et al., 2017), as well as the energy demand and the cross-correlations among all processes (Koskinas et al., 2017).

The possibility of hydropower production in the island was based on an existing water basin, and a simulation framework for 10 small-scale reservoir systems was modelled (Papoulakos et al., 2017), with respect to high solar (Koudouris, 2017), wind and wave energy (Moschos et al., 2017). The potential of deploying agricultural residues, as well as cultivating energy crops with low irrigation demands for biomass energy production, was investigated. Taking into account that Astypalea is located at the Volcanic Arc of southern Aegean Sea, the potential of implementing measurements in order to verify the possibility of geothermal energy was also considered (Chalakatevaki et al., 2017).

It is noted that when designing an energy system based only on weather-related renewable energy resources, i.e. solar, wind, marine and hydroelectric energy, the peak hourly demand is not satisfied due to energy imbalance and a great amount of energy surplus, uncontrollable and unsynchronized with the demand, is produced. Therefore, biomass and geothermal resources were added to the energy mix in order to cover the remaining deficit and a pumped-storage system was used to store electric energy surplus and satisfy the peak deficits (Chalakatevaki et al., 2017).

The implementation of the renewable energy mix resulted in a high installed capacity since the installed power of each renewable energy source does not always synchronize with the demand and a rather high installation cost. The examined solutions still had a high implementation and maintenance cost, and therefore, it was reasonable to consider using thermal stations.

In the near future, it is expected that the cost of renewable resources will be further reduced and the proposed solutions could be more attractive from a financial, societal as well as environmental point of view.

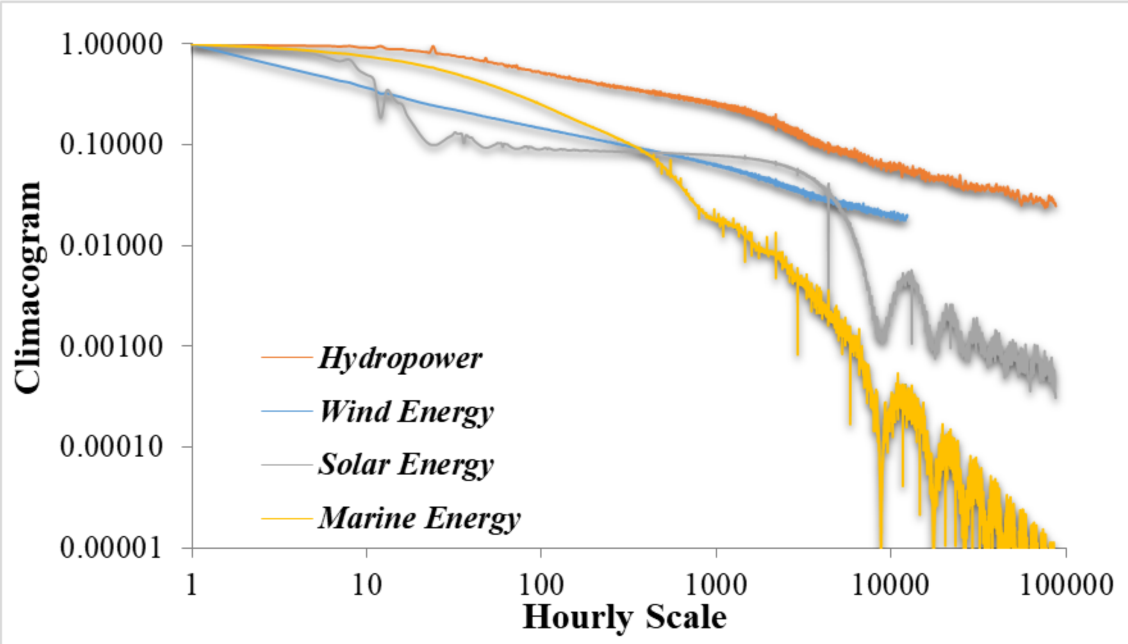


Fig.5.2: *Climacogram demonstrating the uncertainty among the various renewable energy technologies examined in the case study of Astypalea Island.*

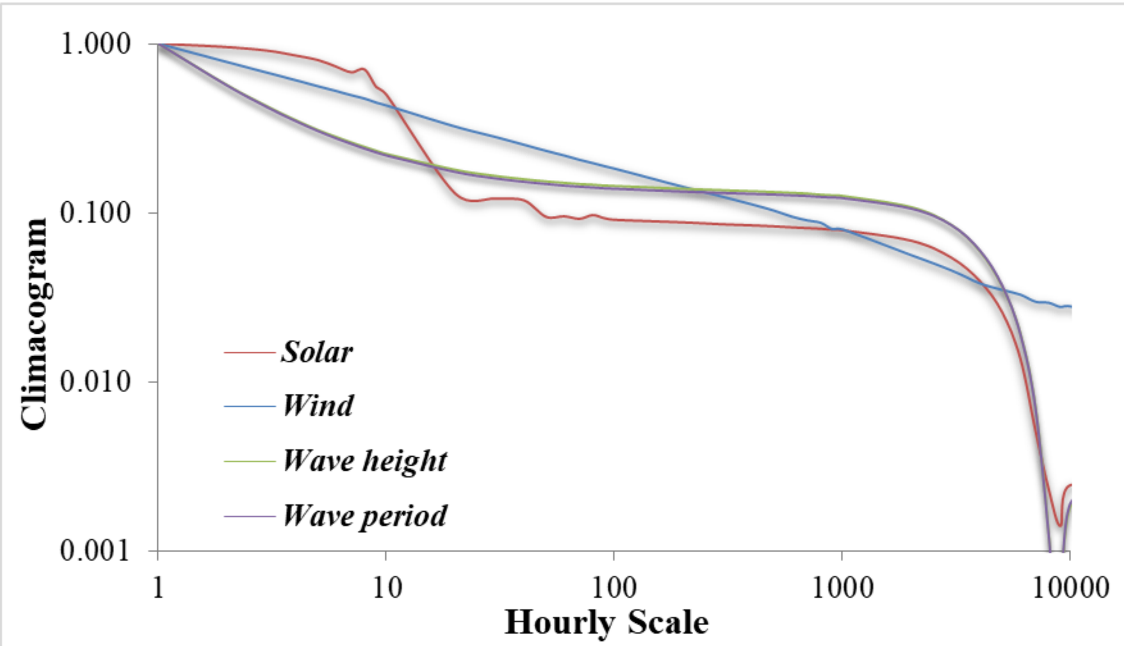


Fig.5.2: *Climacogram demonstrating the uncertainty among hydrometeorological processes investigated in the case study of Astypalea Island.*

5.2 Sensitivity analysis of wind speed and solar radiation processes for calculation of predictability time-window in the island of Astypalea

The remote island of Astypalea faces many challenges since it is not connected to the main energy distribution system of the mainland. The concept of creating a microgrid (Lasseter and Paigi, 2004, Lee et al., 2016) (a micro-scale, autonomous energy system that is often suggested for remote islands and non-connected areas aiming to support the increase in energy demand, reduce the effect possible problems such as a blackout or a mechanical failure, improve energy distribution and even limit the operational cost of the energy system) on the island by redesigning its energy system with renewable energy resources triggered the idea of investigating the predictability window of wind speed and solar irradiance processes and compare them to the results from processes from global databases that were evaluated in the previous chapter. Double cyclo-stationarity was taken into consideration for the processes in Astypalea as well and thus a standardized process is also created to reduce the effect following the procedure already described on chapter 4.1.1.; after calculating the statistical parameters of each timeseries, the mean was subtracted at each value examined and the difference was divided by the standard deviation (Equation 4.1).

Since the climacogram can provide a preliminary calculation of the time period where predictability conquers uncertainty when explored on shorter time scales, the first step in the case study of Astypalea was to make an early estimation of the predictability time window that will be later verified after applying the sensitivity analysis.

The wind speed climacogram (Fig.5.3) displays higher correlation for approximately 10 hours according to the climacogram, indicating that predictability on that period of time is sufficient till that point compared to the next hours. Thus it is expected that the predictability time window will last about 10 hours for both the original and the standardized timeseries (Table 5.2).

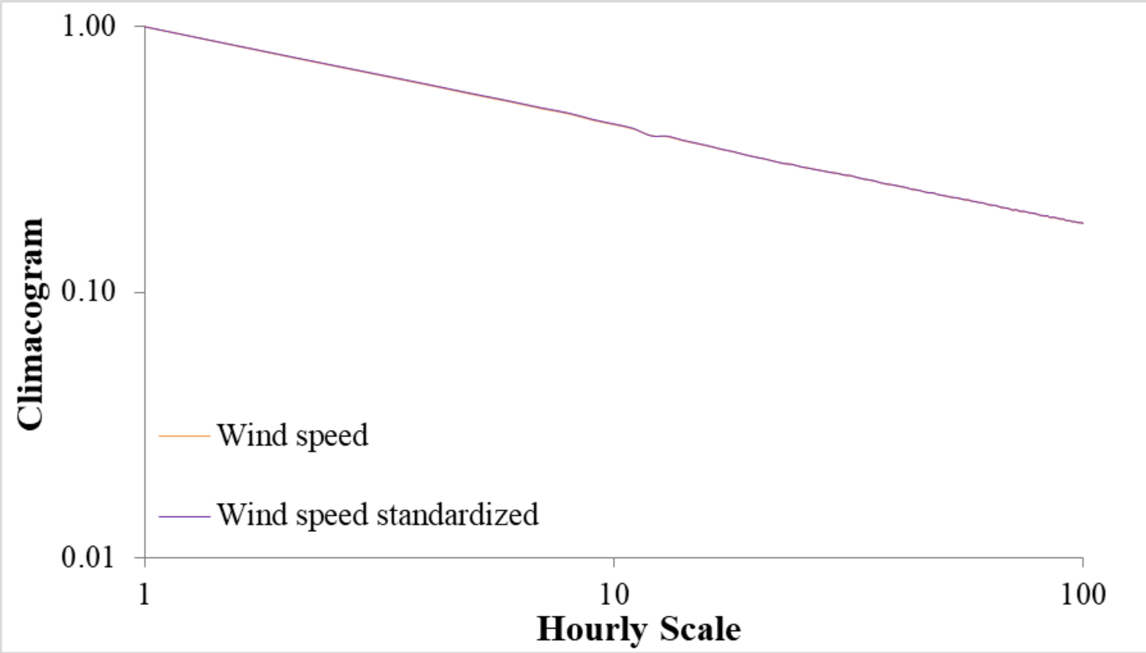


Fig.5.3: *Climacogram of wind speed in Astypalea focusing on short-term behavior.*

Table 5.2: *Astypalea's wind speed climacogram values calculated for 24-hours.*

	Wind data from Astypalaia	**Wind standarized data from Astypalaia**
3h	*0.67*	*0.67*
6h	*0.52*	*0.52*
9h	*0.45*	*0.45*
12h	*0.39*	*0.39*
15h	*0.36*	*0.36*
18h	*0.34*	*0.34*
21h	*0.32*	*0.32*
24h	*0.30*	*0.30*

A far as the investigation of Astypalea's solar radiation process is concerned, the short-term display of the climacogram (Figure 5.4) presents sufficient behaviour with high correlation for approximately 10 hours, which deteriorates from that point on. Consequently solar radiation process can be predictable for at least 10 hours. Additionally the climacogram of the standardized timeseries presumes that the predictability time-window could reach up to 10 hours implying that predictability within that period of time will be accurate (Table 5.3).

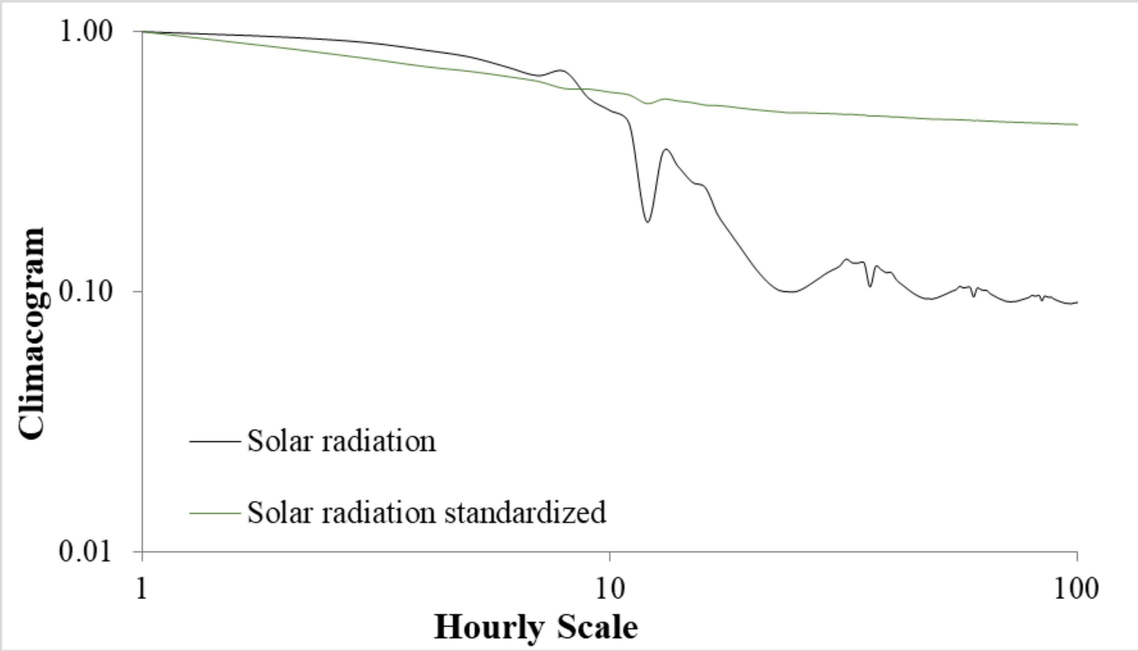


Fig.5.4: *Climacogram of solar radiation in Astypalea focusing on short-term behavior.*

Table 5.3: *Astypalea's solar radiation climacogram values calculated for 24-hours.*

	Solar data fromAstypalaia	Standarized solar data from Astypalaia
3h	*0.91*	*0.91*
6h	*0.73*	*0.73*
9h	*0.55*	*0.59*
12h	*0.19*	*0.24*
15h	*0.26*	*0.53*
18h	*0.17*	*0.51*
21h	*0.12*	*0.50*
24h	*0.10*	*0.49*

Since the computational burden would delay the process again, it was decided that trial testing on a single year of the timeseries would be performed in order to obtain representative results through the sensitivity analysis. The criteria were the same followed in the previous chapter, where a year with similar statistical parameters (mean and standard deviation) to the original synthetic timeseries would be selected for the investigation. The initial wind speed timeseries had $\mu = 5.58$ and $\sigma=3.94$ and the standardized had $\mu=-0.1$ and $\sigma=0.92$. Thus, the most representative year would display $\mu = 5.85$ and $\sigma=4.03$ and after the standardization the statistical parameters would be $\mu=-0.08$ and $\sigma=0.94$. In solar radiation process of Astypalea, before the standardization the statistical parameters were calculated $\mu=203.48$ and $\sigma=288.09$ while after the standardization were altered accordingly; $\mu=0$ and $\sigma=1$. The year with the statistical parameters closer to the original one's displayed $\mu=204.60$ and $\sigma=290,91$ $\mu=0.02$ and $\sigma=1.05$ after the standardization.

The first stage for calculating the predictability time window for processes in Astypalea was to apply the second Benchmak model, B2, in order to set a threshold on the prediction horizon. B1 was also taken into consideration in order to present which values of the F efficiency coefficient were valid for this experiment. It is reminded that B2 assumes that the predicted state $\tilde{s}$ is equal to the current state s (Equation 4.4).

The time horizon of predictability was calculated approximately 1h for the non-standardized wind speed timeseries while after reducing cyclo-stationarity the predictability window slightly expanded to approximately 1.8 hours. The application of B2 on solar radiation presented a time window of predictability reaching 1 hour for both the original and the standardized single year timeseries.

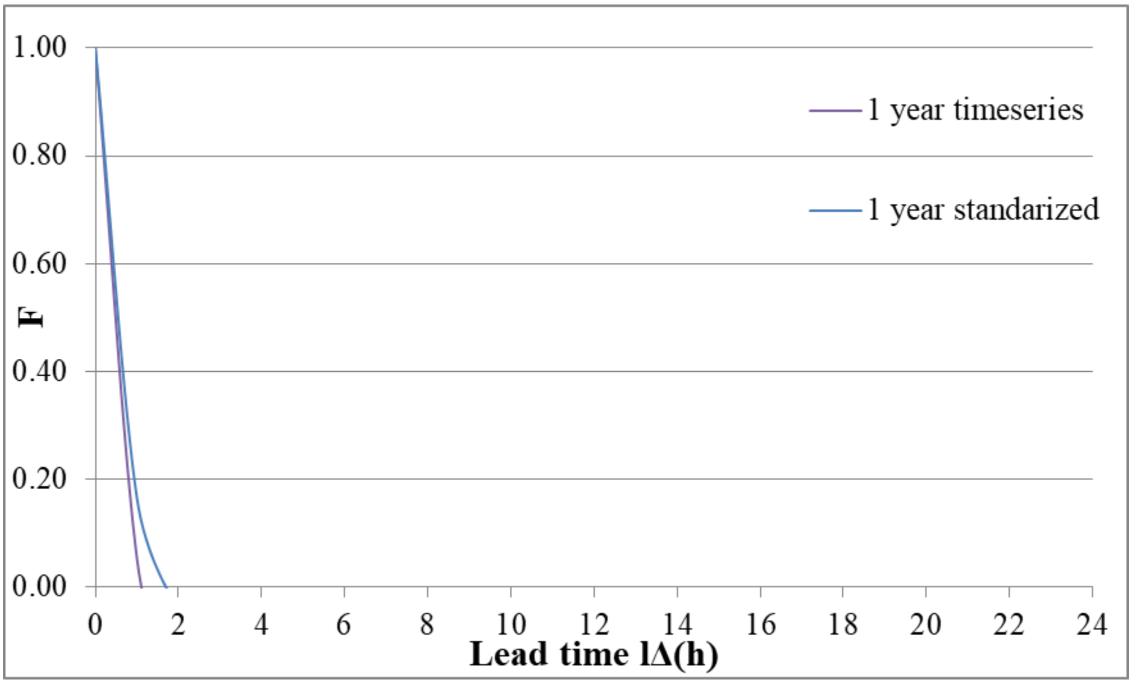


Fig.5.5: *Application of B2 benchmark model on wind speed single-year timeseries in Astypalea, both the initial and the standardized one.*

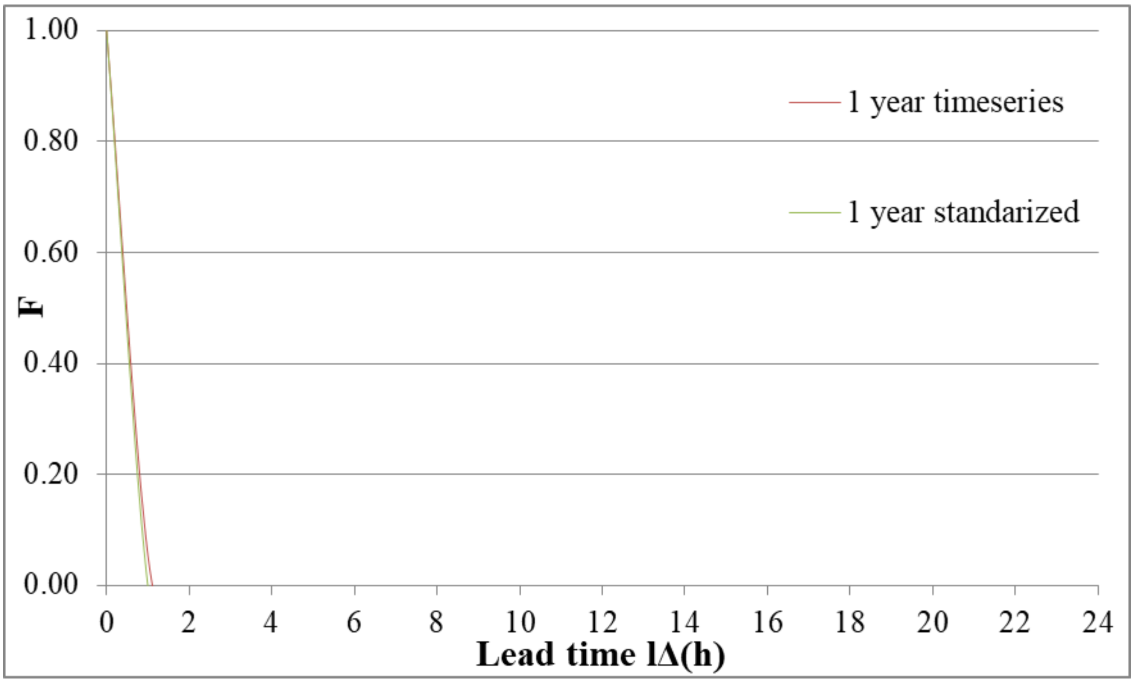


Fig.5.6: *Application of B2 benchmark model on solar radiation single-year timeseries in Astypalea, both the initial and the standardized one.*

After performing the test with the B2 forecasting model, the analogue ensemble model is applied. The parameters (p, h) used in the case of Astypalea are the same applied in Chapter 4.1.1 on processes retrieved from U.S.A. databases following the same principles. The estimating error was calculated g=0.5 for both solar radiation and wind speed process while after the standardization it was modified $\acute{g}_s = 0.1$ and $\acute{g}_w = 0.15$ respectively. A pessimistic and an average error will be tested in both processes to examine the outcome. The results presented in the figures below depict the time window calculated from several neighbors despite the fact that predictions were investigated for analogues 3~20. Just like in the case of wind speed timeseries from the MIT station and the solar radiation from Eugene Airport, while increasing the amount of neighbors on the forecasting model, the limited timeseries started affecting the results of the predictions and thus tests generating realistic results will be displayed.

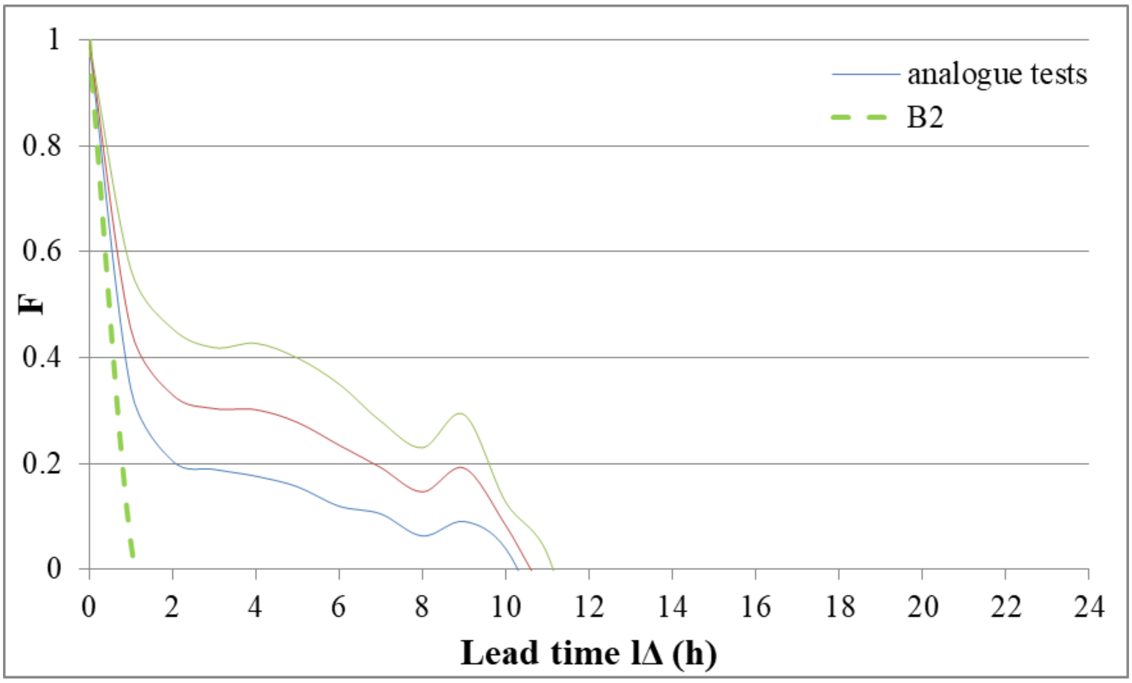


Fig.5.7: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on wind speed in Astypalea.*

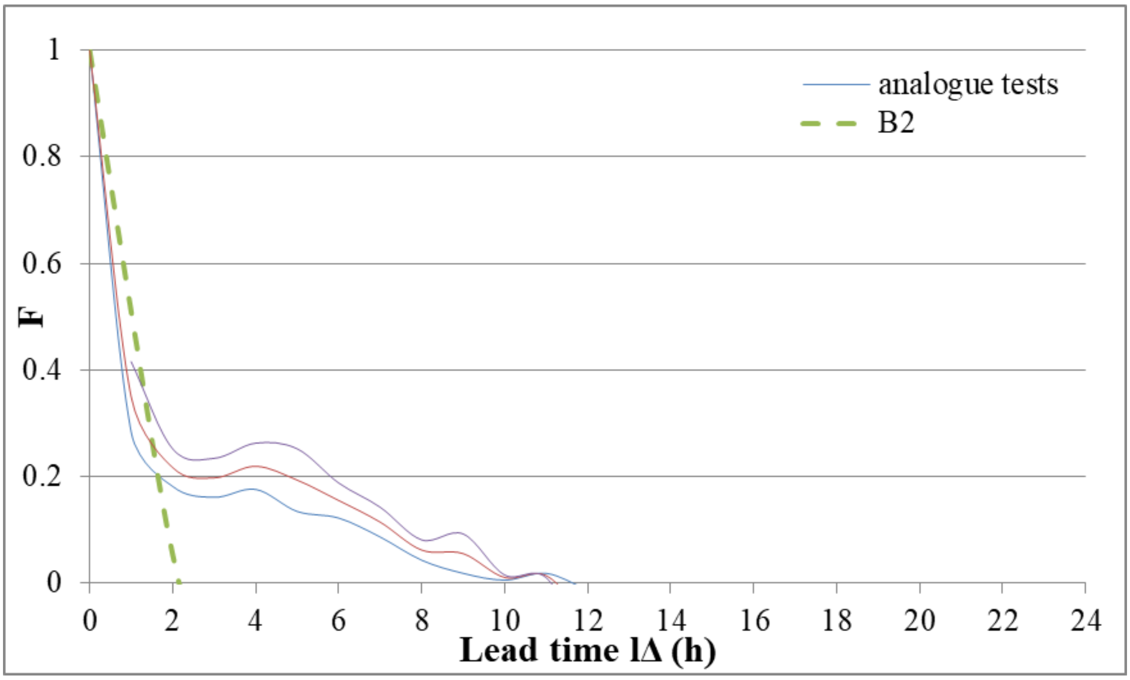


Fig.5.8: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on wind speed in Astypalea.*

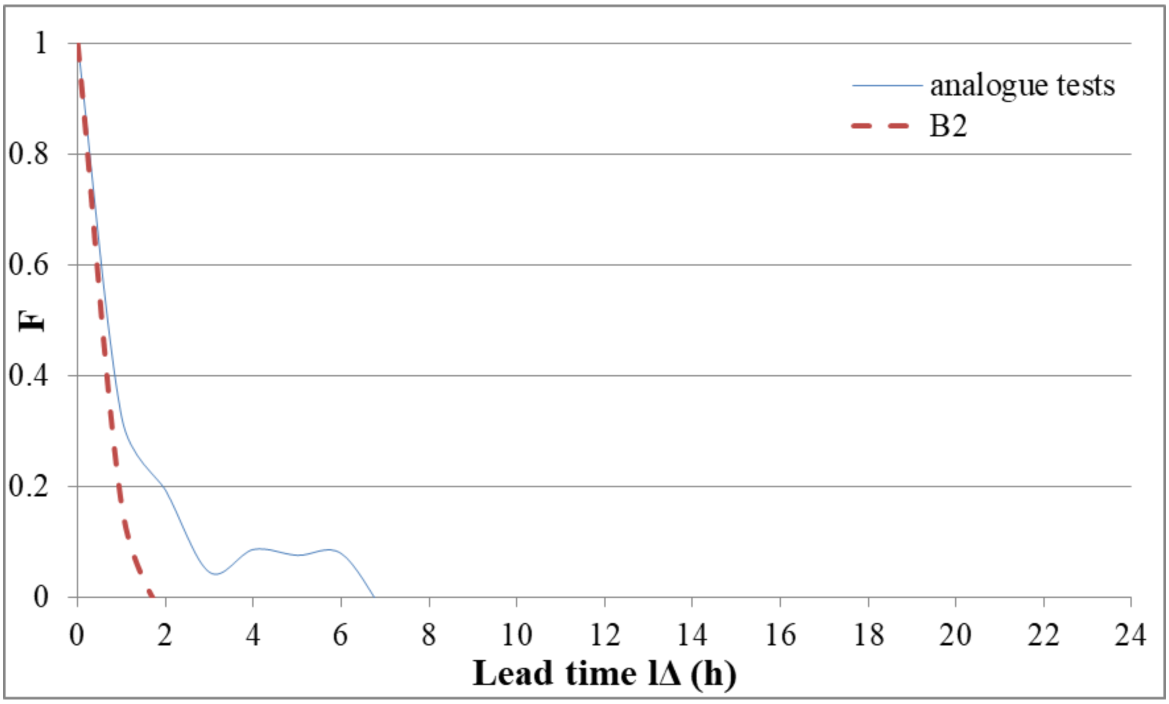


Fig.5.9: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on standardized wind speed in Astypalea.*

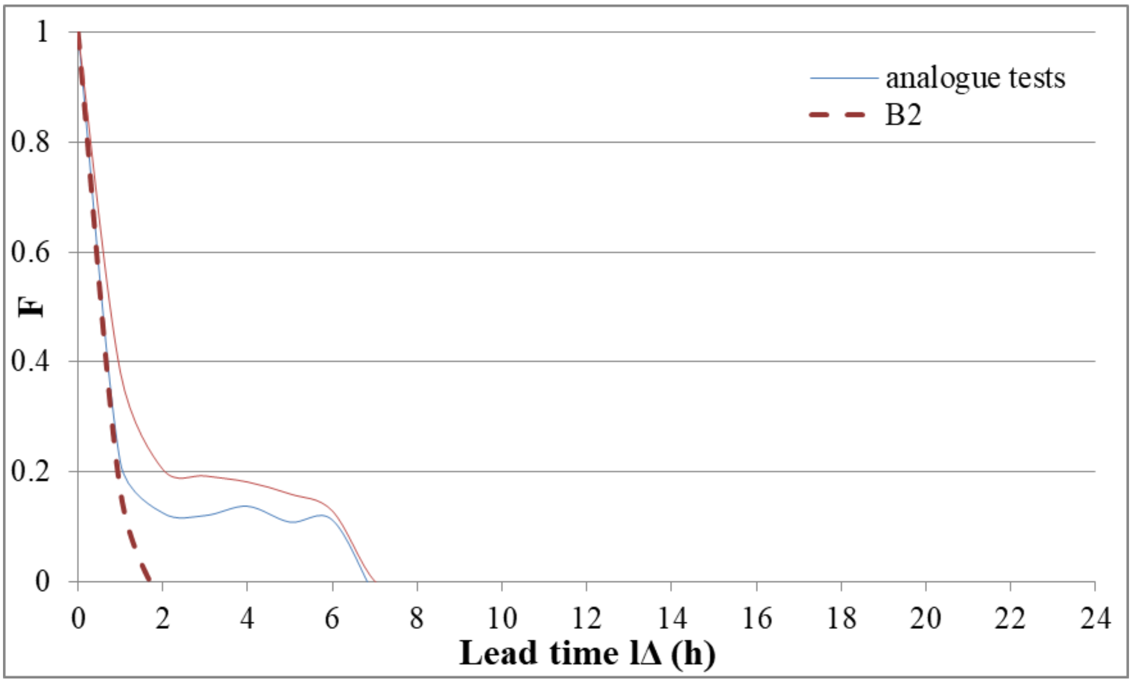


Fig.5.10: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on standardized wind speed.*

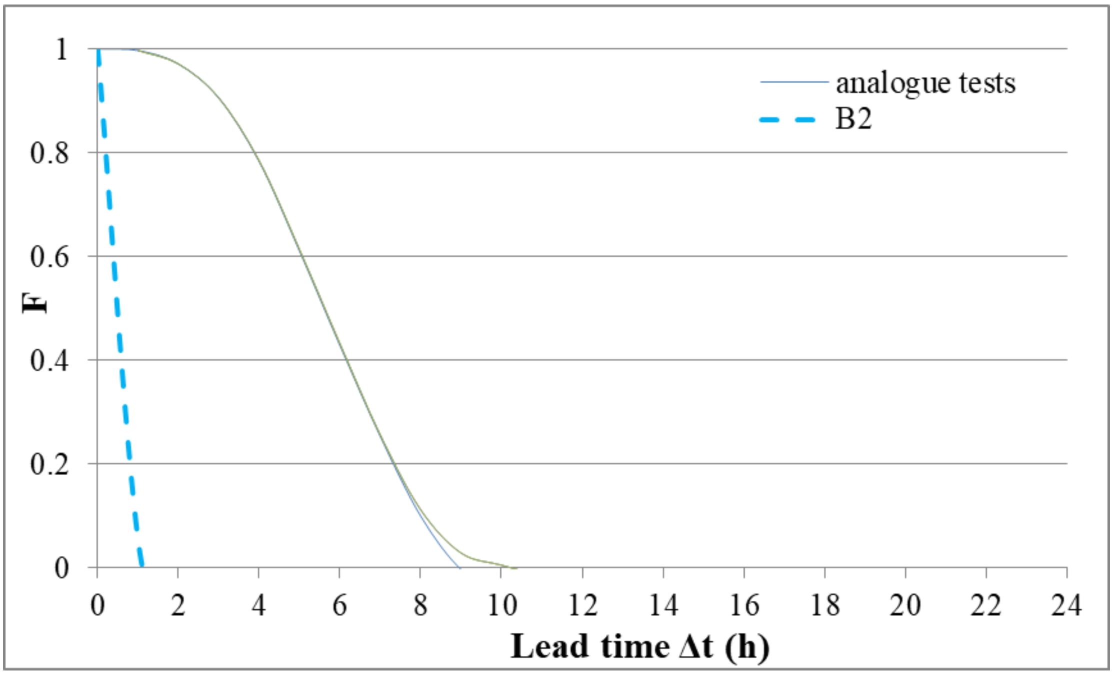


Fig.5.11: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on solar radiation in Astypalea.*

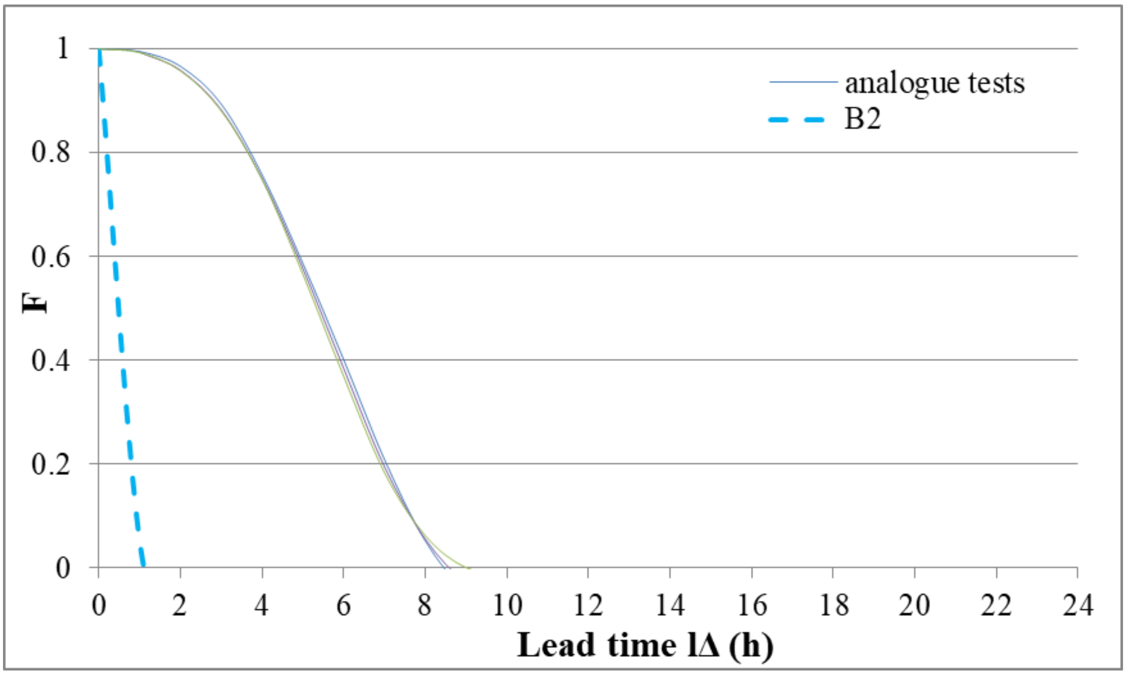


Fig.5.12: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on solar radiation in Astypalea.*

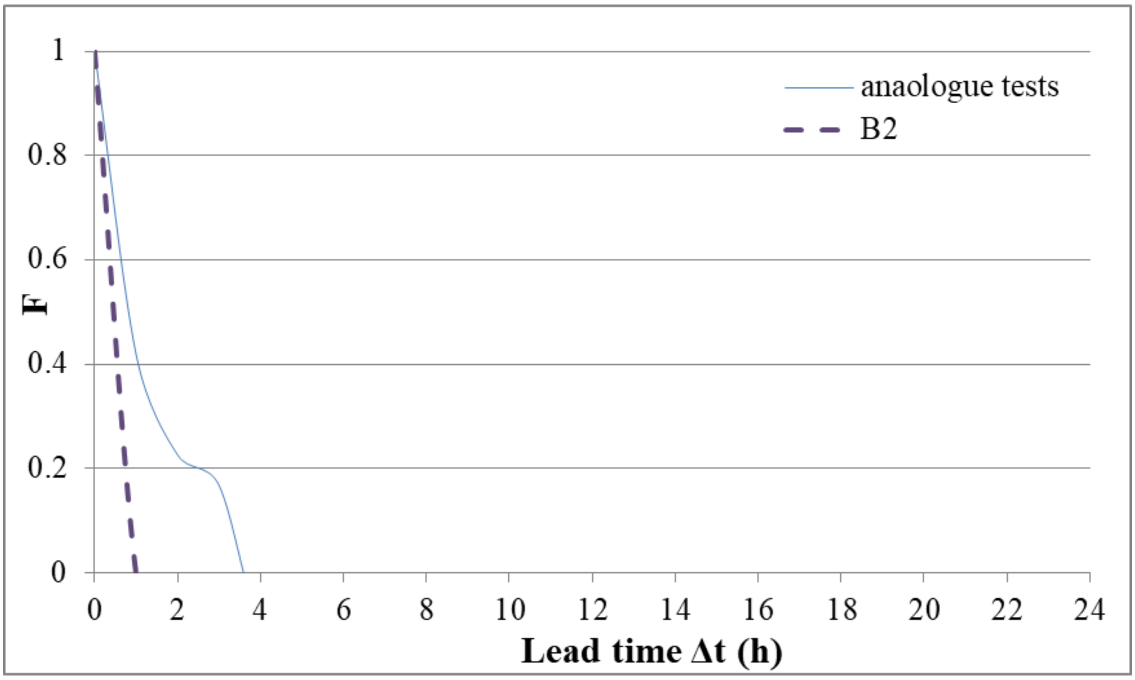


Fig.5.13: *Sensitivity analysis of analogue model (g_{max}=0.5), compared to B2 on standardized solar radiation.*

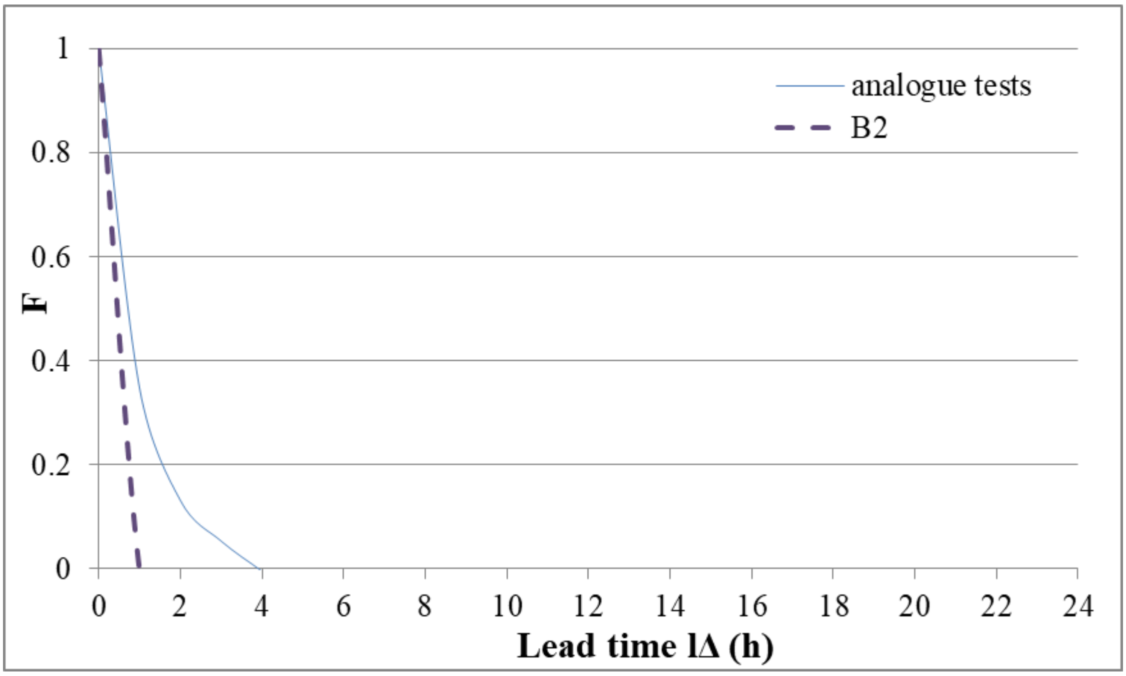


Fig.5.14: *Sensitivity analysis of analogue model ($g_{average}$=0.5), compared to B2 on standardized solar radiation.*

After the sensitivity analysis was finished, the results could be evaluated. To begin with, in wind speed process timeseries, the predictability time-window calculated through the analogue model was wider than the one calculated from B2 which could efficiently predict approximately 1 hour. However although the pessimistic error is considered more conservative, the estimated predictability window was 10 hours (Figure 5.7). However, when testing the average error, the time-horizon where predictions can be considered accurate reached 9 hours (Figure 5.8), which means that it is reduced by 1 hour from the test on the same model using the conservative error.

Continuing the evaluation on the wind speed timeseries, after its standardization, the analogue model outperformed over the B2 model. The analogue model generated accurate predictions for approximately 3.8 hours using pessimistic error scenario (Figure 5.9). About the same results were obtained from the application of the analogue model while tested with an average error slightly improved by 0.2 hours showing that the time window could reach up to 4 hours (Figure 5.10). Similar to the non-standardized wind timeseries, the B2 model accuracy reached 1 hour.

Furthermore, the B2 model once again did not seem to improve prediction even after the standardization of solar radiation process. Both the initial and the standardized timeseries achieved a prediction that could not exceed 1 hour. Additionally, the alteration of the estimating error during the application of the analogue-ensemble algorithm did not seem to significantly affect the length of the predictability window. In the non-standardized timeseries the predictability window was calculated 10 hours using the more conservative error (Figure 5.11), while during the application of the average error the horizon was reduced by 1 hour, reaching 9 hours (Figure 5.12). The modified timeseries where cyclo-stationarity was detained, predictability window was notably worse compared to the one estimated in the initial timeseries, foreseeing only 4 hours with sufficient accuracy (Figures 5.13-5.14).

In retrospect, the initial hypothesis expressed in the beginning of this section - that the predictability window for both wind speed and solar radiation process would extend for approximately 10 hours - should be reminded. The calculation of the predictability time-window through the sensitivity analysis proved that in solar radiation determinism may prevail for up to 10 hours while in wind speed the predictability horizon is estimated to 12 hours. In conclusion, it is once again verified that the exploration of the climacogram on shorter scales can provide an early estimation of the time horizon within which the predictions generated through certain methods can be valid.

Natural processes may “carry” an inherent degree of uncertainty. The identification of this degree is of great significance since hydrometeorological and geophysical processes may be exploited for the generation of renewable energy, given the fact that those natural processes are infinite sources.

However, the uncertainty may be inherited to the energy system causing operational malfunctions and problems on both the organization and planning. Consequently by "harnessing" the problem of unpredictability throughout multiple processes, energy management could be optimized. By defining the predictability window of each process, planning to a more regular base could be applied for the organization and the improvement of natural systems. This could contribute to reduce the operational cost significantly and start exploiting reserves more optimal.

The illustration of the climacogram of natural processes followed by the respective energy they produce, focusing again on wind speed and solar radiation process (Figures 5.15-5.16) (since they are the most appealing worldwide) can prove that they have a very similar behavior. Even long-term, the correlation between the process and the corresponding energy remains high.

Consequently, taking into account that each natural process is highly correlated to the respective produced energy, the predictability time-window calculated for the process should be very close to the predictability window of energy. Since the time window in the processes so far does not exceed 20 hours, a day-to-day operational planning should be designed. Thus renewable energy systems could start replacing fossil fuels turning to an alternative and more "green" approach on power generation that could offer constant power supply that derives from inexhaustible sources.

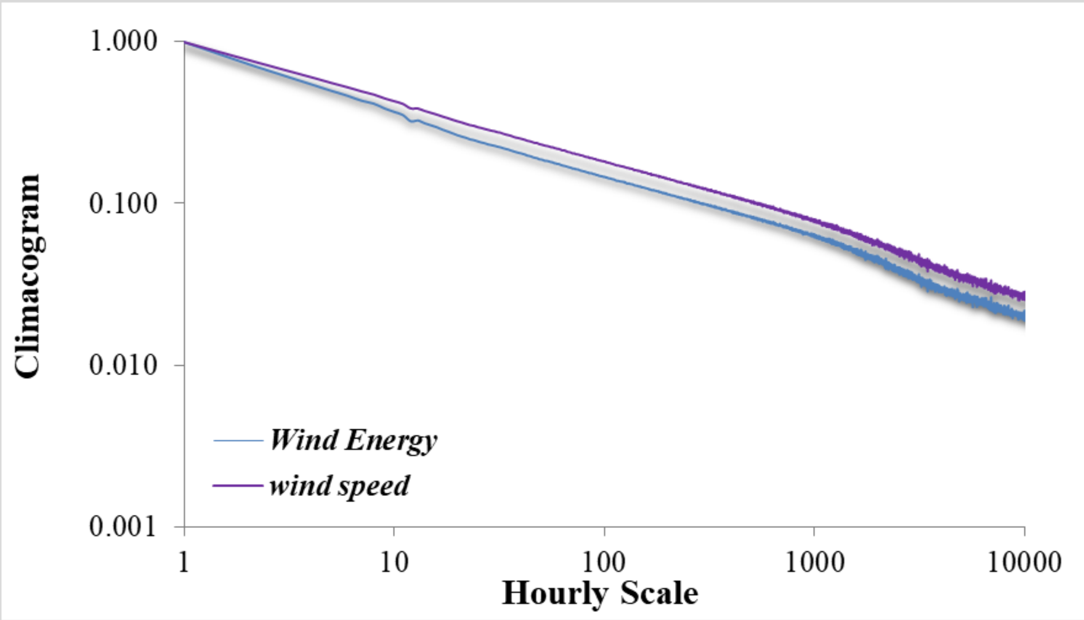


Fig.5.15: *Illustration of the correlation of wind speed process to Wind energy in Astypalea through the climacogram.*

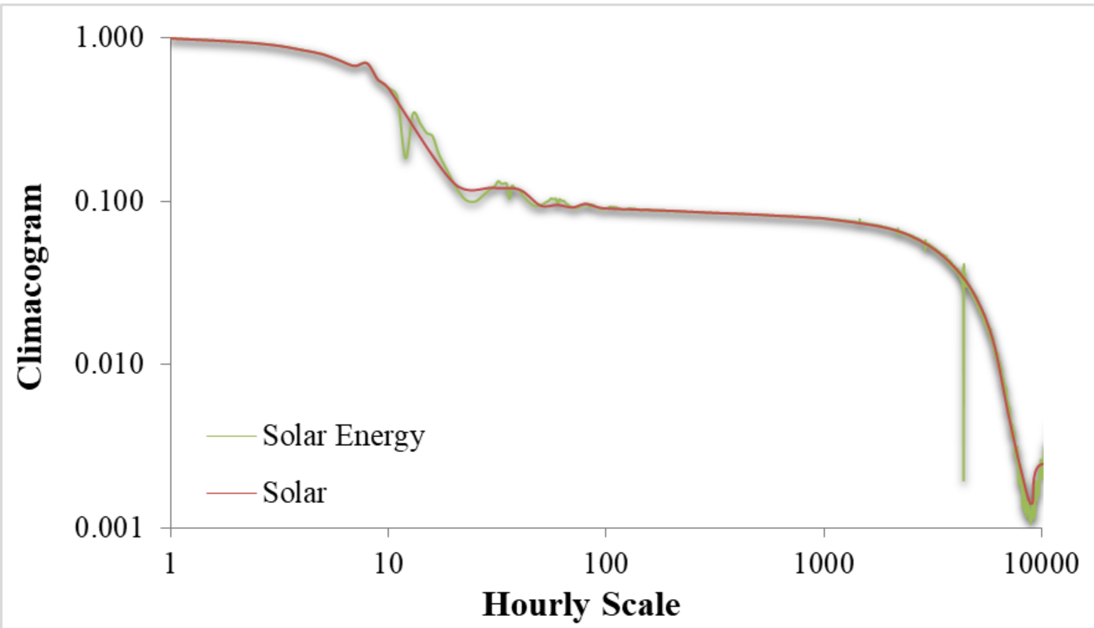


Fig.5.16: *Illustration of the correlation of solar radiation process to Solar energy in Astypalea through the climacogram.*

6. Conclusions

- Since the Hurst parameter is greater than 0.5, the 10 examined processes associated with renewable energy resources systems are governed by a great degree of uncertainty. Yet the degree of uncertainty significantly varies across the examined processes. Specifically, we find solar radiation and wave height to exhibit large Hurst parameters, while others such as precipitation exhibit relatively lower ones.

- The Hurst parameter varies across different time scales, which is also supported by recent research results and it is estimated greater than 0.5 ($H > 0.5$, even using a downward biased estimator through the climacogram) for all of the examined processes. Consequently, several natural processes (marine, precipitation, sun and wind) exhibit the Hurst-Kolmogorov behaviour and not a Markovian or a white noise one.

- After applying the sensitivity analysis among wind speed and solar radiation processes, the results displayed that the analogue algorithm outperformed compared to the B2 model expanding the unpredictability time window up to approximately 10 hours.

- The performance of the analogue model was tested for a pessimistic error scenario (conservative) and an average one, more flexible estimating error. The application of the average threshold could foresee validly more hours compared to the more conservative error, calculating a wider predictability window for processes retrieved from the area of the U.S.A.

- The standardized timeseries, where double cyclo-tationarity was aimed to be reduced, estimated a longer predictability window compared to the non-standardized processes implying that the limitation of the effect of cyclo-stationarity may generate more accurate forecasting.

- Through the climacogram-based analysis it is revealed that in long time scales, the larger the Hurst parameter the shorter the predictability window while additionally the greater the q parameter the wider the predictability time horizon; solar radiation process is characterized by a larger degree of uncertainty and through the sensitivity analysis the predictability window would last 8 hours while in wind speed process the Hurst parameter was lower than the solar radiation but its predictability window was wider reaching up to 15 hours.

- The degree of variability of the different processes at the time-scales of interest are crucial for the design and operation of a renewable energy system, where the operation rules of each energy source should be specified in a way that ensures reliability and optimal performance of the whole system. For example, a general operational rule of such systems is that energy from the least reliable sources or with the smallest predictability time window, e.g. solar energy, is consumed in priority.

- Recent analyses concerning the renewable energy design and management for the non-connected island of Astypalea have illustrated how the uncertainty of several renewable energy sources can be efficiently managed through stochastic analysis. Therefore, stochastic analysis is essential in the renewable energy management both for the analysis of predictability of the related natural processes and for the analysis of the systems dynamics and optimal design and operation under increased uncertainty.

The stochastic aspect of renewable energy systems still remains a relatively underexplored field. Further research will focus on the exploration of the uncertainty of the related processes based on more extended datasets as well as on the quantification of the propagation of uncertainty from the natural processes to the final energy production.

7. References

Alani Adeshina, Osunmakinde Isaac, "Short-Term Multiple Forecasting of Electric Energy Loads for Sustainable Demand Planning in Smart Grids for Smart Homes", MDPI Journal, DOI: 10.3390/su9111972, Sustainability 2017, 9, 197 2, Published October 2017.

Alessandrini S., Delle Monache L., Sperati S., Nissen J.N., "A novel application of an analog ensemble for short-term wind power forecasting", Elsevier, Renewable Energy 76, 768-781, December 2014.

Alonzo Ben. "Types of Meteorology", sciencing.com, meteorology-8031, Published March 2020.

Breuer L., "Introduction to Stochastic Processes", Data Science Central, Retrieved November 2011.

Brownlee Jason, "Supervised and Unsupervised Machine Learning Algorithms", Machine Learning Algorithms, March 2016.

Chambers C. John, Mullick K. Satinder, Smith D. Donald, "How to Choose the Right Forecasting Technique", Harvard Business, Issue, Published July 1971

Chang, W. (2014) A Literature Review of Wind Forecasting Methods. Journal of Power and Energy Engineering, **2**, 161-168. doi: 10.4236/jpee.2014.24023, April 2014.

Wen-Yeau Chang, "A Literature review of Wind Forecasting Methods, Journal of Power and Energy Engineering, 2, 161-168, published April 2014

Cannon Alex, "Nonlinear analog predictor analysis: A coupled neural network/ analog model for climate downscaling", Science Direct, Elsevier, DOI: 10.1016/j.neunet.2007.04.002, 2007.

Chalakatevaki M., Stamou, P., Karali, S., Daniil, V., Dimitriadis, P. Tzouka, K., Iliopoulou, T., Koutsoyiannis, D., Papanicolaou P., and Mamassis, N.: Creating the electric energy mix in a non-connected island, Energy Procedia, 125, 425–434, doi:10.1016/j.egypro.2017.08.089, 2017.

Chalakatevaki, M., Klousakou, E., Dimitriadis, P., Iliopoulou, T., and Koutsoyiannis, D.: Stochastic investigation of the uncertainty of hydrometeorological processes by means of the climacogram, EGU General Assembly, Vienna, Austria, 8-13 April 2018, Geophysical Research Abstracts, EGU2018-17714-1, 2018.

Chardavellas, E., Dimitriadis, P., Papakonstantis, I., Koutsoyiannis, D., and Papanicolaou, P.: Stochastic similarities between the microscale of vertical thermal jet and macroscale hydrometeorological processes, EGU General Assembly, Vienna, Austria, 8-13 April 2018, Geophysical Research Abstracts, EGU2018-17803, 2018.

Cococcioni Marco, D'Andrea Eleonora, Lazzerini Beatrice, "24-Hour-Ahead forecasting of energy production in solar PV systems",Conference Paper, Research Gate, DOI: 10.1109/ISDA.2011.6121835, Published November 2011.

Dimitriadis Panayiotis, Koutsoyiannis Demetris, "Climacogram versus autovariance and power spectrum in stochastic modeling for Markovian and Hurst-Kolmogorov processes", Stoch Environ Res Risk Assess, DOI 10.1007/s004 77-015-1023-7, Springer-Verlag Berlin Heidelberg 2015.

Dimitriadis, P., Koutsoyiannis, D., and Tzouka, K.,:Predictability in dice motion: how does it differ from hydrometeorological processes?, Hydrological Sciences Journal, 61 (9), 1611–1622, doi:10.1080/02626667.2015.1034128,2016.

Dimitriadis Panayiotis, "Hurst-Kolmogorov dynamics in hydroclimatic processes and in the microscale of turbulence", Thesis for the degree of the Doctor of Philosophy, National Technical University of Athens, published online December 2017.

Etemadi, A., Emdadi, A., AsefAfshar, O. and Emami, Y.: Electricity generation by the ocean thermal Energy, Energy Procedia, 12, 936-943, , 2011.

Edenhofer, O., R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner, et al. (Eds.): Renewable energy sources and climate change mitigation, Summary for policy makers and technical summary, Special report of the intergovernmental panel on climate change, Edition 2011, IPCC, USA, Cambridge, UK/New York: Cambridge University, 2012.

A.M. Foley, Member IEEE, P.G. Leahy, Member IEEE and E.J. McKeogh, Member IEEE, "Wind Power Forecasting & Prediction Methods", IEEE, 9th International Conference on Environment and Electrical Engineering, Prague, Czech Republic, May 2010.

Fraedrich Klaus, Raible Christoph, Sielmann Frank, "Analog Ensemble Forecasts of Tropical Cyclone Tracks in the Australian Region", Weather and Forecasting, Volume 18, American Meteorogical Society, 2003.

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, "An Introduction to Statistical Learning", Springer. p. vii., 2013.

Gkolemis, A., Dimitriadis, P., Karakatsanis, G., Iliopoulou, T., and Koutsoyiannis, D.: A stochastic investigation of the intermittent behaviour of wind; application to renewable energy resources management, EGU General Assembly, Vienna, Austria, 8-13 April 2018, Geophysical Research Abstracts, EGU2018-15979-3, 2018.

Gkolemis A., Stochastic investigation and simulation of wind speed - Application in management of wind works, Diploma thesis, Department of Water Resources and Environmental Engineering – National Technical University of Athens, Athens, February 2019.

Hadjimitsis, E., Demetriou, E., Sakellari, K., Tyralis, H., Dimitriadis, P., Iliopoulou, T. and Koutsoyiannis, D.: Investigation of the stochastic nature of temperature and humidity for energy management, Geophysical Research Abstracts, EGU General Assembly, Vienna, Austria, 23-28 April 2017, EGU2017-10164-5, 2017.

Iliopoulou, T., Papalexiou, S. M., Markonis, Y., and Koutsoyiannis, D.: Revisiting long-range dependence in annual precipitation. Journal of Hydrology, 556, 891-900, 2016
Jackson Charles, Sen Mrinal, Stoffa Paul, "An Efficient Stochastic Bayesian Approach to Oprimal Parameter and Uncertainty Estimation for Climate Model Predictions", Journal of Climate, Volume 17, American Meteorological Society, Published July 2004.

Karakatsanis, G., Roussis, D., Moustakis, Y., Gournari, P., Parara, I., Dimitriadis, P. and Koutsoyiannis, D.: Energy, variability and weather finance engineering, Energy Procedia, 125, 389-397, 2017.

Klousakou, E., Chalakatevaki, M., Tomani, R., Dimitriadis, P., Efstratiadis, A., Iliopoulou, T., Ioannidis, R., Mamassis, N., and Koutsoyiannis, D.: Stochastic investigation of the uncertainty of atmospheric processes related to renewable energy resources, EGU General Assembly, Vienna, Austria, 8-13 April 2018, Geophysical Research Abstracts, EGU2018-16982-2, 2018.

Klousakou E., Chalakatevaki M., Dimitriadis P., Iliopoulou T., Ioannidis I., Karakatsanis G., Efstratiadis A., Mamasis N., Tomani R., Chardavellas E. and Koutsoyiannis D.: A preliminary stochastic analysis of the uncertainty of natural processes related to renewable energy resources, Published by Copernicus Publications on behalf of the European Geosciences Union, Advances in Geosciences, 45, 193-199, August 2018.

Koskinas, A., Zacharopoulou, E., Pouliasis, G., Engonopoulos, I., Mavroyeoryos, K., Deligiannis, E., Karakatsanis, G., Dimitriadis, P., Iliopoulou, T., Koutsoyiannis, D. and Tyralis, H.: Simulation of electricity demand in a remote island for optimal planning of a hybrid renewable energy system, EGU General Assembly, Vienna, Austria, 23-28 April 2017, Geophysical Research Abstracts, EGU2017-10495-4, 2017.

Koudouris, G., Dimitriadis, P., Iliopoulou, T., Mamassis, N. and Koutsoyiannis, D.: Investigation on the stochastic nature of the solar radiation process, Energy Procedia, 125, 398–404, 2017.

Koudouris Giannis, Dimitriadis Panagiotis, Iliopoulou Theano, Mamassis Nikos, Koutsoyiannis Demetris, "A stochastic model for the hourly solar radiation process for application in renewable resources management'', Published by Copernicus Publications on behalf of the European Geosciences Union, Advances in Geosciences, 45, 139-145, August 2018.

Koutsoyiannis D., Langounis A., "Precipation", Elsevier B.V., 2011
Koutsoyiannis, D.: Generic and parsimonious stochastic modelling for hydrology and beyond, Hydrological Sciences Journal, 61 (2), 225–244, doi:10.1080/02626667.2015.1016950, 2016.

Koutsoyiannis, D.: Entropy production in stochastics, Entropy, 19 (11), 581, doi:10.3390/e19110581, 2017.

Makridakis Spyros, Spiliotis Evangelos, Assimakopoulos Vassilios, "Statistical and Machine Learning forecasting methods: Concerns and ways forward", Plus One, Published March 2018.

Manache Delle Luca, Alessandrini Stefano, "Analog-Based Ensemble Model Output Statistics", Junk Et Al, Monthly Weather Review, ForWind-Center for Wind Energy Research, DOI: 10.1175/MWR-15-0095.1, American Meteorogical Society, Published July 2015.

Moschos, E., Manou, G., Dimitriadis, P., Afendoulis, V., Koutsoyiannis, D., and Tsoukala, V.: Harnessing wind and wave resources for a Hybrid Renewable Energy System in remote islands: a combined stochastic and deterministic approach, Energy Procedia, 125, 415–424, doi: 10.1016/j.egypro.2017.08.084, 2017.

Ahmed Nesreen, Atiya Amir, Gayar Neamat, El-Shishiny Hisham, "An Empirical Comparison of Machine Learning Models for Time Series Forecasting", Econometric Reviews, 29:56, 594-621, DOI: 10.1080/07474938.2010.481556, Published September 2010.
Papacharalampous Georgia, Tyralis Hristos, Koutsoyiannis Demetris, "Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes'', Stochastic Environmental Research and Risk Assessment, DOI: 10.1007/s00477-018-1638-6, Springer-Verlag GmbH Germany, part of Springer Nature 2019.

Papoulakos, K., Pollakis, G., Moustakis, Y., Markopoulos, A., Iliopoulou, T., Dimitriadis, P., Koutsoyiannis, D., and Efstratiadis, A.: Simulation of water-energy fluxes through smallscale reservoir systems under limited data availability, Energy Procedia, 125, 405–414, 2017.

Saurabh S. Soman, Hamidreza Zareipour, Senior Member, IEEE, Om Malik, Life Fellow, IEEE, and Paras Mandal, Member, IEEE, "A Review of Wind Power and Wind Speed Forecasting Methods With Different Time Horizons", 2010

Schaetzle, O., and Cees, J. N. Buisman: Salinity gradients energy: current state and new trends, Engineering 2015, 1(2), 164-166, doi 10.15302/j-eng-2015046, 2015.

Sharma Navin, Sharma Pranshu, Irwin David, Shenoy Prashant, "Predicting Solar Generation from Weather Forecasts Using Machine Learning", Virtual Power Plants, Distributed Generation, Microgrids, Renewables and Storage (IEEE SmartGridComm), Department of Computer Science, University of Massachusetts Amherst, 978-1-4577-1702-4/11, IEEE, 2011.

Solomatine Dimitri, Dulal Khada, "Model trees as an alternative to neural networks in rainfall-runoff modelling", Hydrological Sciences Journal, 48:3, 399-411, DOI: 10.1623/hysj.48.3.399.45291, Published January 2010.

Tomasino Mario, Zanchettin Davidde, Travesco Pietro, "Long-range forecasts of River Po discharges based on predictable solar activity and a fuzzy neural network model.'', Hydrological Sciences Journal, 49:4, -684, DOI: 10.1623/hysj.49.4.673.54431, published 15 December 2009.

Tyralis, H., Dimitriadis, P., Koutsoyiannis, D., O'Connell, P., Tzouka, K. and Iliopoulou, T.: On the long-range dependence properties of annual precipitation using a global network of instrumental measurements, Advances in Water Resources, 111, 301-318, 2017a.

Tyralis, H., Karakatsanis, G., Tzouka, K., and Mamassis, N.: Exploratory data analysis of the electrical energy demand in the time domain in Greece, Energy Procedia, 134, 902-918, 2017b.

Voyant, Cyril & Notton, G. & Kalogirou, Soteris & Nivet, Marie-Laure & Paoli, Christophe & Motte, Fabrice & Fouilloy, Alexis. (2017). Machine Learning methods for solar radiation forecasting: a review. Renewable Energy. 105. 10.1016/j.renene.2016.12.095.