

波高趨勢預測之時間數列模型研究

林演斌¹ 馬瀾嘉²

¹ 國立成功大學近海水文中心計畫經理

² 國立成功大學統計學系副教授

摘要

海洋各種參數之變化趨勢與預測資訊，是海洋管理之重要參考依據。本研究蒐集蘇澳浮標 2000~2011 年波高平均值、最大值資料，引用統計學之時間數列方法，試著找出隱藏在資料中之趨勢資訊，並建立統計的預測模型。本研究以 ADF 檢定分析年波高數列之平穩性及其趨勢，並進行 ARIMA(2,1,1)、ARIMA(2,1,2)、ARIMA(0,1,1)、ARMA(2,2)等預測模型之建模與測試，其中 ARIMA(2,1,2)、ARMA(2,2)係以經自然對數轉換後之數列建模。根據實證趨勢分析結果顯示：蘇澳浮標 2000~2011 年之年波高平均值不具備時間趨勢及最大值數列有時間趨勢，年波高平均值及最大值數列皆為非平穩型數列。依據建模與測試結果得知：ARIMA(2,1,1)、ARIMA(2,1,2)模型分別適合未轉換及自然對數轉換後之年波高平均值，以 ARIMA(2,1,1)較佳。ARIMA(0,1,1)、ARMA(2,2)模型分別適合未轉換及自然對數轉換後之年波高最大值，以 ARMA(2,2)較佳。另外，不同的採樣方式適合不同之預測模型。與 Agrawal and Deo (2002)使用 ARMA(1,1)模型比較，本研究之數列均不適合 ARMA(1,1)模型。以自然對數轉換之波高資料而言，與 Stefanakos (1999)提議之 ARMA(2,2)模型比較，本研究亦發現自然對數轉換之年波高最大值適合該模型。

關鍵詞：資料浮標、示性波高、時間數列分析與預測

A study on time series analysis and forecasting of wave heights

Yen-Pin Lin* Mi-Chia Ma

* Manager, Coastal Ocean Monitoring Center, National Cheng Kung University

ABSTRACT

Trends and forecasts of a variety of parameters in the ocean are very important to marine management. This study gathers data of wave heights of Suao buoy from 2000 to 2011. This study converts data into information for marine management by utilizing the methodology of time series analysis and forecasting of statistics. The ADF test is utilized to analyze stationary and trends. ARIMA(2,1,1), ARIMA(2,1,2) and ARIMA(0,1,1), ARMA(2,2) models are established and tested. Series of wave heights are transformed by Napierian logarithm before fitting ARIMA(2,1,2) and ARMA(2,2) models. Based on results of the ADF test, it shows that: There is no time trend for annual averages but there exist time trend for maxima of significant wave heights. Series of annual averages and maxima are nonstationary. Based on analytical results of forecasting models, it shows that: ARIMA(2,1,1) and ARIMA(2,1,2) model are suitable for untransformed and transformed series of annual averages. ARIMA(2,1,1) is better than ARIMA(2,1,2). ARIMA(0,1,1) and ARMA(2,2) model are suitable for untransformed and transformed series of annual maxima. ARMA(2,2) is better than ARIMA(0,1,1). In addition, series sampled by different ways is fitted by different model. Sampled wave heights data do not fit ARMA(1,1) well which was utilized by Agrawal and Deo (2002).

Transformed annual maximum wave heights fit ARMA(2,2) well which was used by Stefanakos (1999).

Keywords: Data buoy; Significant wave height; Time series analysis and forecasting

一、前言

台灣海岸地區人口密集，隨著經濟發展，促使海岸地區利用日益增加，例如電廠之設置、濱海工業區之開發等，對於海岸地區之利用不勝枚舉。海岸地區之變化受到自然現象與人為因素之影響，其中波浪為影響海岸變化之重要因子，其趨勢預測可直接及間接地作為海岸管理參考。

海洋領域專家發展之波浪模式(model)，已能預測未來 72 小時之波浪變化。而波浪為時間數列，是一種隨機過程(random process)。統計學家發展之時間數列分析與預測方法亦已被普遍應用於管理、經濟、社會科學與工程等領域，因此本研究將之應用於分析以往波浪變化趨勢，及建立預測模型，檢驗預測結果之準確度。

二、文獻探討

Agrawal and Deo (2002)採用時間數列分析方法中的 ARMA(1,1)及 ARIMA(1,1,1)兩種模型預測波高變化，分別以 3 小時到 24 小時的時距進行預測，經檢驗預測資料與實測資料之相關係數(coefficient of correlation)，ARMA 介於 0.698~0.808，ARIMA 介於 0.727~0.824。以該研究的樣本而言，ARMA(1,1)及 ARIMA(1,1,1)於不同時距預測結果之差異不大。然而，在其他顯著非平穩型時間數列之樣本，實際情形可能不是如此。

波高為海況資料的一種，Monbet et al. (2007)說明海況並非為常態分布。若將其轉換為常態，則可用現存技術模擬常態的隨機過程，例如：ARMA 模型等。Monbet et al. (2007)並說明，由於波高極為接近對數常態分布，一般而言以自然對數來轉換波高，自然對數為一種轉換函數。一旦確定轉換函數，再以轉換後的時間數列套用 ARMA 等模型建模，曾被提議的模型包含 AR(1)、ARMA(2,2)、AR(20)、ARMA(4,4)等，其中 ARMA(2,2)為 Stefanakos (1999)所提議，係以自然對數為轉換函數。

三、研究方法

3.1 研究流程與假設

1. 預測模型研究流程

本研究以實測波高資料，引用時間數列方法分析 2000~2011 年波高之單變數時間數列模型預測結果。研究流程如圖 1 所示：

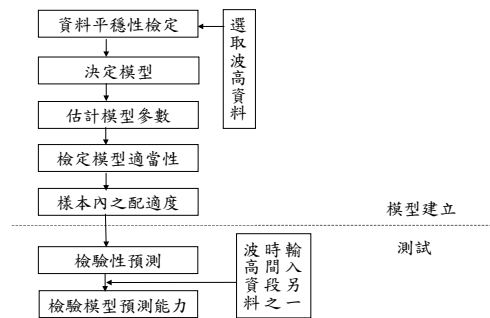


圖 1 單變數時間數列預測模型研究流程圖

2. 研究假設

本文欲驗證之假設(對立假設)為：

- H1: 假設年波高平均值數列時間趨勢項斜率不為 0。
- H2: 假設年波高最大值數列時間趨勢項斜率不為 0。
- H3: 假設擬合模型的殘差存在自我相關。
- H4: 假設擬合模型的殘差不是常態分布。

擬合模型的假設條件是否成立，模型殘差是否存在自我相關、是否為常態分布都將被檢驗。針對自然對數轉換後之資料，以上假設也都會再次驗證。

3.2 研究變數操作性定義與衡量

本研究旨在分析過去十二年波高之變化趨勢，並建立波高之單變數時間數列模型，作為未來實務應用參考。本研究探討之變數為波高，係以海上之設備，經由觀測作業之取樣、處理、分析等過程，再經過逐年分析平均值、最大值後而成為本研究之輸入資料。本節簡介波高觀測設備及資料之定義與

衡量。

1. 波高觀測設備

觀測設備為資料浮標(data buoy)，是一座海上自動海氣象環境監測站，為成功大學近海水文中心參考美國國家海洋大氣總署之觀測方法，依據台灣海域環境特性所改良設計而成，其觀測項目包含波高、週期、波向、風速、風向、陣風、氣溫、氣壓、水溫、海流等，並具擴充能力。本研究蒐集的資料是由資料浮標即時地傳輸至成功大學近海水文中心，通過資料品管程序後分析而得。資料傳輸流程如圖 2。

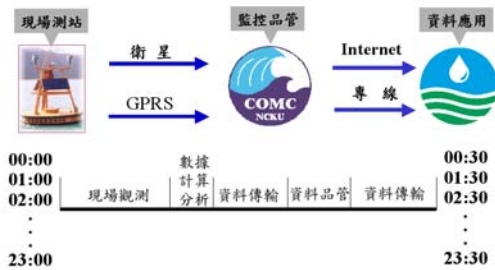


圖 2 資料傳輸流程

2. 資料之定義與衡量

實際海面波浪為不規則波，海洋領域專家為描述所觀測的波浪，定義示性波高(significant wave height)，本研究稱之波高為示性波高，係由一維波譜計算而得。

波高資料分析流程如下：

- (1) 擷取波浪儀器訊號：於每小時以 2Hz 頻率擷取 1024 個訊號。
- (2) 計算一維波譜。
- (3) 分析不規則波代表值：波高= $4.004\sqrt{m_0}$ ， m_0 為一維波譜曲線與頻率軸包圍之面積，代表波浪能量。

3.3 時間數列分析方法簡介

1. 平穩性與趨勢分析

判斷時間數列是否平穩有許多方法，若僅以肉眼檢視圖形趨勢，無法將結果量化。因此本研究依林(2006)說明時間數列分析的單位根檢定(unit root test)，以檢定數列是否平穩與具有時間趨勢。該法

最早由 Dickey and Fuller (1979)所提出，其目的為檢定 AR(1)模型 $y_t = \phi y_{t-1} + \varepsilon_t$ ， $t=1, \dots, T$ 之數列是否平穩， y 為待分析之變數， ϕ 為參數， ε 為白噪音，下標代表時刻，其虛無假設為 $\phi=1$ (數列含單位根)，對立假設為 $\phi < 1$ (數列為平穩)，表示如下：

$$H_0: \phi = 1 \quad \text{vs} \quad H_a: \phi < 1 \quad (1)$$

令 $\Delta y_t = y_t - y_{t-1}$ ，則 $\Delta y_t = (\phi - 1)y_{t-1} + \varepsilon_t$ 。若令 $\gamma = \phi - 1$ ，則虛無假設： $\gamma = 0$ ，對立假設： $\gamma < 0$ 。之後 Dickey and Fuller 將模型增加時間趨勢及自我迴歸項如下：

$$\Delta y_t = a_0 + a_1 t + \gamma y_{t-1} + \sum_{i=1}^p \beta_i \Delta y_{t-i} + \varepsilon_t \quad (2)$$

其中 $a_0 + a_1 t$ 為典型的時間趨勢模型， a_0 為截距項， $a_1 t$ 為時間趨勢項， $\sum_{i=1}^p \beta_i \Delta y_{t-i}$ 為自我迴歸項，此項之目的在確保 ε_t 不具自我相關性，其中 p 為增加 $\beta_i \Delta y_{t-i}$ 的項目數， β_i 為係數。Dickey and Fuller 並提出多種 ADF 檢定(Augmented Dickey-Fuller test)統計量，利用 ADF 檢定統計量檢查估計模型參數是否顯著，並刪除不顯著者。

由於 ADF 檢定假設與 DF 檢定相同，在此將 DF 檢定依不同模型、樣本數 T 為 100 之檢定假設與臨界值(critical value)，列舉如下：

表 1 DF 模型之檢定假設與臨界值($T=100$)

模型	檢定假設	$\alpha=0.05$		$\alpha=0.01$	
		臨界值	臨界值	臨界值	臨界值
$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$	$H_0: \gamma = 0$	-1.95	-2.60	-2.89	-3.51
$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$	$H_0: \gamma = 0$	-2.89	-3.51	4.71	6.70
	$H_0: a_0 = \gamma = 0$	4.71	6.70	-3.45	-4.04
$\Delta y_t = a_0 + a_1 t + \gamma y_{t-1} + \varepsilon_t$	$H_0: \gamma = 0$	-3.45	-4.04	4.88	6.50
	$H_0: a_0 = a_1 = \gamma = 0$	4.88	6.50	6.49	8.73
	$H_0: a_1 = \gamma = 0$	6.49	8.73		

2. 單變數平穩型時間數列預測方法

在以 ADF 檢定判定時間數列為平穩型後，始能套用平穩型時間數列預測方法。常見單變數平穩型時間數列模型有 p 階自我迴歸過程(autoregressive process) AR(p)、 q 階移動平均過程(moving-average process) MA(q)、(p, q)階混合自我迴歸與移動平均過程(mixed autoregressive-moving average process) ARMA(p, q)如下：

$$\text{AR}(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (3)$$

$$\text{MA}(q): y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (4)$$

$$\text{ARMA}(p,q): y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (5)$$

在本研究中， y_t 為 t 時刻之波高，或經自然對數轉換之波高， μ 為移動平均過程之平均數， $\phi_i, i=1, \dots, p$ 為 AR 參數， ε_t 為 t 時刻之白噪音， $\theta_i, i=1, \dots, q$ 為權重。

3. 單變數非平穩型時間數列分析方法

非平穩型時間數列模型則透過取差分，將 y_t 轉為平穩型時間數列，再做 ARMA 模型分析。以 B 為後移運算子，一次差分可表示為 $\Delta y_t = (1-B)y_t = y_t - y_{t-1}$ ，二次差分可表示為 $\Delta^2 y_t = (1-B)^2 y_t = (1-2B+B^2)y_t = y_t - 2y_{t-1} + y_{t-2}$ ， d 次差分可表示為 $\Delta^d y_t = (1-B)^d y_t$ 。(p,d,q)階自我迴歸整合移動平均模型 (autoregressive integrated moving average model) ARIMA(p,d,q) 如下：

$$\text{ARIMA}(p,d,q): \phi_p(B)W_t = \theta_q(B)\varepsilon_t \quad (6)$$

其中 $\phi_p(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$ ， $W_t = \Delta^d (y_t - \mu)$ ， $\theta_q(B) = (1 - \theta_1 B - \dots - \theta_q B^q)$ 。

四、結果與討論

本研究以蘇澳浮標波高作為主要觀測變數，選取 2000~2011 年之年波高平均值、年波高最大值作為趨勢分析及建模研究用，分別取得 12 筆資料。假設檢定的顯著水準訂為 0.05。

4.1 平穩性與趨勢分析

1. 年波高平均值

首先，以式(7)之模型檢定殘差無自我相關，
$$\Delta y_t = a_0 + a_1 t + \gamma y_{t-1} + \beta_1 \Delta y_{t-1} + \varepsilon_t \quad (7)$$

檢定變數 y_{t-1} 的係數之 ADF 檢定值為 -2.46 (p 值 = 0.33)，大於 0.05，表示檢定結果為不顯著，無法拒絕 $\gamma = 0$ 之虛無假設，代表數列含有單位根，為非平穩型。

接著檢定 a_1 是否為 0，經檢定可知 a_1 之 p 值為 0.77，大於 0.05，表示檢定結果為不顯著，無法拒絕 $a_1 = 0$ 之虛無假設，代表年波高平均值數列不具有時間趨勢。然後去除式(7)之 $a_1 t$ ，成為式(8)：

$$\Delta y_t = a_0 + \gamma y_{t-1} + \beta_1 \Delta y_{t-1} + \varepsilon_t \quad (8)$$

再檢定變數 y_{t-1} 的係數之 ADF 檢定值為 -2.82 (p

值 = 0.09)，大於 0.05，表示檢定結果為不顯著，無法拒絕 $\gamma = 0$ 之虛無假設，代表數列含有單位根，為非平穩型。

進一步檢定 a_0 ，其 p 值為 0.03，小於 0.05，表示檢定結果為顯著，拒絕 $a_0 = 0$ 之虛無假設。

2. 年波高最大值

首先，以式(7)之模型檢定殘差無自我相關，檢定變數 y_{t-1} 的係數之 ADF 檢定值為 -3.18 (p 值 = 0.14)，p 值大於 0.05，表示檢定結果為不顯著，無法拒絕 $\gamma = 0$ 之虛無假設，代表數列含有單位根，為非平穩型。

接著檢定 a_1 是否為 0，經檢定可知 a_1 之 p 值為 0.01，小於 0.05，表示檢定結果為顯著，拒絕 $a_1 = 0$ 之虛無假設，年波高最大值數列具有時間趨勢。

4.2 預測模型分析

1. 平穩性檢定

年波高平均值為非平穩型，因此先將數列進行差分，並檢定為平穩後再建模。年波高最大值數列亦為非平穩型，亦進行差分再建模。另外，為建立 ARMA(2,2) 模型，先以自然對數轉換年波高平均值、最大值數列，再檢定平穩性，若為非平穩，再進行差分，直到檢定為平穩，檢定步驟同 4.1 節。結果年波高平均值數列經過一次差分後為平穩型、年波高平均值數列經自然對數轉換再一次差分後數列為平穩型、年波高最大值數列經過一次差分後為平穩型、年波高最大值數列經自然對數轉換後為平穩型。

2. 決定模型

本研究先以一次差分後之年波高平均值數列之第 1~9 筆資料搭配 ARMA(1,1) 模型建模，但 ARMA(1,1) 模型殘差存在自我相關，因此增加自我迴歸項，改以 ARMA(2,1) 模型建模，成為 ARIMA(2,1,1) 模型。並以自然對數轉換再經一次差分後之年波高平均值數列之第 1~9 筆搭配 ARMA(2,2) 模型建模，為 ARIMA(2,1,2) 模型。

另外，先以一次差分後之年波高最大值數列搭配 ARMA(1,1) 模型建模，但自我迴歸參數不顯著，改以第 1~8 筆搭配 MA(1) 模型建模，為 ARIMA(0,1,1) 模型。並以自然對數轉換後年波高最大值數列之第 1~8 筆搭配 ARMA(2,2) 模型建模。估計模型參數如

表 2。

表 2 模型參數估計值

檢定係數	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (0,1,1)	ARMA (2,2)
ϕ_1	-0.253	-0.219	-	0.610
ϕ_2	-0.587	-0.102	-	0.385
θ_1	-0.940	-0.436	-0.889	-0.133
θ_2	-	-0.558	-	-0.862

3.檢定模型適當性

本節檢定模型殘差是否無自我相關且為常態分布，分別引用殘差 Q 檢定與 JB 檢定(Jarque-Bera test) 之方法，以年波高平均值套用 ARIMA(2,1,1)說明。

以 ARIMA(2,1,1)模型檢定時間間隔 1~5 時期之殘差是否無自我相關，依據殘差 Q 檢定結果，全部間隔時期之顯著性均大於 0.05 之顯著水準，代表殘差已無自我相關。

圖 3 為 ARIMA(2,1,1)模型殘差之 JB 檢定結果，可知 JB 統計量為 0.71，大於 0.05，結果不顯著，無法拒絕殘差為常態分布之虛無假設。綜合以上檢定結果，顯示 ARIMA(2,1,1)模型適合年波高平均值數列。

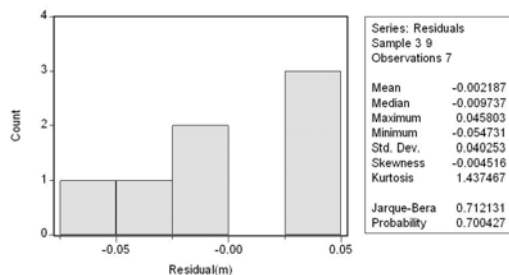


圖 3 ARIMA(2,1,1)模型 JB 檢定結果

依照以上檢定方法，亦可得 ARIMA(2,1,2)模型適合自然對數轉換後之年波高平均值數列。ARIMA(0,1,1)模型適合年波高最大值數列。另外，ARMA(2,2)模型適合自然對數轉換後之年波高最大值數列。

4.樣本內之配適度

為從適合的模型中選擇較適當之模型，本研究引用統計方法中之配適度與預測力作為衡量指標。

配適度指標為 AIC(Akaike Information Criterion)、SBC(Schwartz Bayesian information)，兩

者皆為誤差平方和(Error Sum of Squares, SSE)之函數，其值愈小代表配適度愈佳。從表 3 可知適合年波高平均值之 ARIMA(2,1,1)及 ARIMA(2,1,2)模型比較結果，ARIMA(2,1,1)模型之配適度較 ARIMA(2,1,2)模型佳。

另外，年波高最大值之 ARIMA(0,1,1)及 ARMA(2,2)模型比較結果，ARMA(2,2)模型之配適度較 ARIMA(0,1,1)模型佳。

表 3 模型配適度

配適度指標	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (0,1,1)	ARMA (2,2)
AIC	-2.881	-1.495	5.054	1.110
SBC	-2.904	-1.526	5.064	0.971

年波高平均值經一次差分後數列和以 ARIMA(2,1,1)模型擬合之預測結果如圖 4 所示。

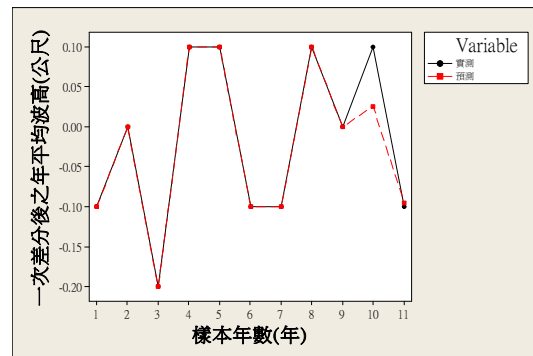


圖 4 ARIMA(2,1,1)模型預測結果

5.檢驗模型預測能力

從表 4 比較年波高平均值兩模型之平均絕對值誤差率(Mean Absolute Percentage Error, MAPE)，可知 ARIMA(2,1,1)模型之預測能力較 ARIMA(2,1,2)模型佳。

從表 4 比較年波高最大值兩模型之 MAPE，可知 ARMA(2,2)模型之預測能力較 ARIMA(0,1,1)模型佳。

表 4 模型預測能力比較表

預測能力指標	ARIMA (2,1,1)	ARIMA (2,1,2)	ARIMA (0,1,1)	ARMA (2,2)
樣本外樣本數	2	2	3	4
MAPE(%)	39.59	78.96	156.5	7.89

五、結論與限制

上述資料時間數列之趨勢與模型實證分析結果詳如表 5。由實證檢定結果可知，2000~2011 年蘇澳浮標之年波高平均值為非平穩型、不具有時間趨勢，但最大值數列為非平穩型、具有時間趨勢。

表 5 本研究假設實證結果之彙整

研究假設	實證結果
假設 H1：年波高平均值式(7) $a_1 \neq 0$ 。	不成立
假設 H2：年波高最大值式(7) $a_1 \neq 0$ 。	成立
假設 H3：擬合模型的殘差存在自我相關。	不成立
假設 H4：擬合模型的殘差不是常態分布。	不成立

ARIMA(2,1,1)模型適合年波高平均值數列，ARIMA(2,1,2)模型適合自然對數轉換之年波高平均值數列，不論是配適度或預測能力，皆以 ARIMA(2,1,1)較佳。

ARIMA(0,1,1)模型適合年波高最大值數列，ARMA(2,2)模型適合自然對數轉換之年波高最大值數列，不論是配適度或預測能力，皆以 ARMA(2,2)較佳。

與 Agrawal and Deo (2002)使用 ARMA(1,1)模型比較，本研究採樣的數列均不適合 ARMA(1,1)模型。以自然對數轉換之波高資料而言，Stefanakos (1999)提議使用 ARMA(2,2)模型，本研究發現自然對數轉換之年波高最大值適合該模型。

本研究之限制有：

- (1) 本研究僅蒐集 12 年實測資料，於分析時 ADF 檢定之臨界值是基於殘差呈常態的假設，若在小樣本殘差不呈常態，臨界值僅能供參考。
- (2) 在分析過程中發現部分檢定統計量接近 0.05 顯著水準，顯示檢定結論不十分明朗，有待未來觀測更長時期資料後再分析。
- (3) 本研究所得結論僅適用於研究區域。

謝誌

承蒙經濟部水利署長期投入經費建立與維持近海水文觀測並進行各年度之統計分析，使本研究得以從中蒐集資料並順利完成，謹致謝忱。

參考文獻

1. 林茂文(2006)*時間數列分析與預測：管理與財經之應用*，華泰文化事業股份有限公司，第三版。

2. Agrawal J.D., Deo M.C. (2002) "On-line wave prediction," *Marine Structures*, Vol. 15, Issue 1, pp. 57-74.
3. Dickey, D.A., Fuller, W.A. (1979) "Distribution of the estimators for autoregressive time series with a unit root," *Journal of American Statistical Association*, Vol. 74, No. 366, pp. 427-431.
4. Monbet V., Ailliota P., Prevostob M. (2007) "Survey of stochastic models for wind and sea state time series," *Probabilistic Engineering Mechanics*, Vol. 22, pp. 113-126 .
5. Stefanakos C.N. (1999) "Nonstationary stochastic modelling of time series with applications to environmental data," *Ph.D. thesis*, NTAU.