

利用马尔可夫链理论建立波高相关性模式

董东璟¹ 吴立中¹ 高家俊^{1, 2} 庄士贤² 李汴军³

(1.水利及海岸工程研究所, 2.近海水文中心, 成功大学, 台南 台湾 中国; 3.台湾气象局海象中心, 台北 台湾 中国)

摘 要: 以台湾海峡之国光钻油平台与新竹海气象浮标的有效波高为分析对象, 通过移转机率矩阵建立马尔可夫链波高相关性模式。研究结果显示, 利用马尔可夫链相关性模式推估之波高的不确定程度低于传统回归相关分析所得, 而波高之估算误差受当时海况影响, 在恶劣的海况下, 本文模式之推估误差低于 20%。

关键词: 马尔可夫链理论; 波高相关性; 不确定分析

中图分类号: P714.1; P731.33; P751 **文献标识码:** A **文章编号:** 1001-6392(2002)03-0001-10

引 言

中国台湾四周环海, 与海洋有关之活动极为频繁, 诸如近岸港湾工程、船只航行、近远洋捕鱼, 岸边养殖、海底资源之钻勘、休闲游憩……等, 上述多数的海域活动需仰赖长期且稳定精确之海气象资料以提供必要的信息, 如海岸工程的规划需根据长期波浪观测资料, 规划可作业天数, 决定设计波浪; 港口营运则必须利用长期波浪资料评估每年船只不能进港时间, 不能作业天数以及疏散条件; 又如近年来海域游憩活动开始兴盛, 游艇的拥有者日益增多, 近海的海域信息为确保这些活动重要的因素之一。因此实时且精准的海气象水文观测资料, 已成为各类海域活动以及海岸工程之规划不可或缺之重要讯息。

海气象观测为一长期性之工作, 因海上气候变化万千, 使其困难度远高于陆上作业, 通常天候愈恶劣, 所测得资料就愈珍贵, 难度相对地就愈高。台湾成功大学近海水文中心 (Coastal Ocean Monitoring Center, COMC) 以从事台湾近海之海气象水文观测为主要任务, 近年来已逐渐于台湾沿海建立了以资料浮标 (Data Buoy)、观测桩 (pile station)、潮位站 (tidal station)、气象站等系统为主的环岛海气象观测网 (Kao 等, 1999)^[5], 同时也发展出资料品管方法 (董东璟, 1997)^[1] 来确保所测得之海气象数据质量。

现场海气象观测资料虽然常用来作为数值模型及物理模型的率定或验证对象, 然而现实中因受经济、环境、政治等因素影响, 并不易在海域上广泛地布置测点, 常见为了某一工程规划的需要而有短期测站的设置, 真正长期作业化 (operational) 的海气象观测站并不多。但是波浪统计分析上, 愈长资料的统计结果有愈小的不确定性, 代表性愈高, 本文基于此, 希望发展一个相关性分析方法, 试图藉由邻近的长期测站, 将短期观测数据繁衍延长, 以扩充海气象数据库。

统计学上的马尔可夫链理论 (Markov Chain Theory) 具有表现两相邻状态间演变之特性, 一般常被应用于时间序列分析 (Dynkin, 1965)^[3], 本文引用马尔可夫链概念, 将其应用于从事空间上相邻两测站之相关性分析。由于有效波高 (significant wave height) 为描述海况 (sea state) 最直接的要害之一, 因此为本文之分析对象。本研究希望藉由概率的信息来表达相邻两个测站之间波高状态的关系, 并将对本文所建立之相关性模式从事不确定性分析, 同时将分析结果与传统分析方法比较以评判本文所提方法之效能为何。

本文试图以统计理论从事波高相关性分析, 基本上欲从事分析的两个测站必须属于同一个统计上的均匀区 (homogeneous area), 波候特征类似的测站始具有统计分析的意义。基于上述对台湾沿海海域之波候描述, 本文选用位于台湾海峡内之国光平台与新竹浮标两个测站之波高数据为分析对象。国光平台为一个钻油平台, 离岸约 30 km, 接近台湾海峡中段处, 其水深约 60 m, 因其特殊位置, 使其所测得资料对深海气象以及波浪特性之研究有其重要性。新竹浮标则设置于新竹近海, 离岸约 2 km, 与国光平台相距约 25 km, 测站水深约 23 m。上述两测站的位置如图 1 所示。由于国光平台即将拆除, 深海的海气象资料即将中断, 因此若能利用新竹浮标建立两站相关性, 期能在国光平台拆除后, 仍能由统计方法繁衍国光平台资料, 或由新竹浮标测站推算国光平台处的海象信息。

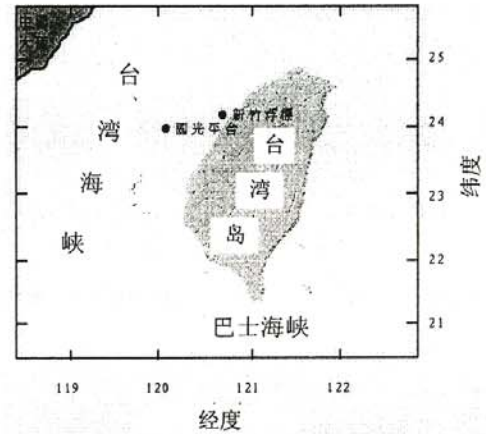


图 1 国光平台与新竹浮标测站位置图

Fig.1 Station site of Kuokwang oil platform and Hsinchu data buoy

1 台湾沿海波浪特性及本文所选用之测站

台湾近海海域的波候特征主要受季节与地理位置的影响, 在台湾东岸面临太平洋, 属开放性海域, 如图 2, 以波龄 (wave age, C_p/U_{10}) 作为判断波浪成熟度的参数 (Goda, 1985)^[4],

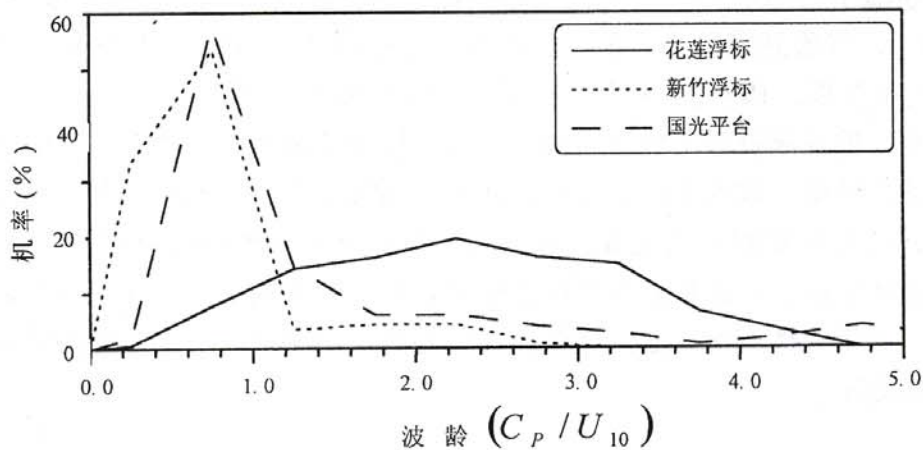


图 2 台湾东部花莲浮标测站与西部新竹浮标 (近海)、国光平台 (深海) 测站之波龄机率分布

Fig.2 Probability distribution of wave age at Hualien buoy, Hsinchu buoy (offshore) and Kuokwang platform (bathy)

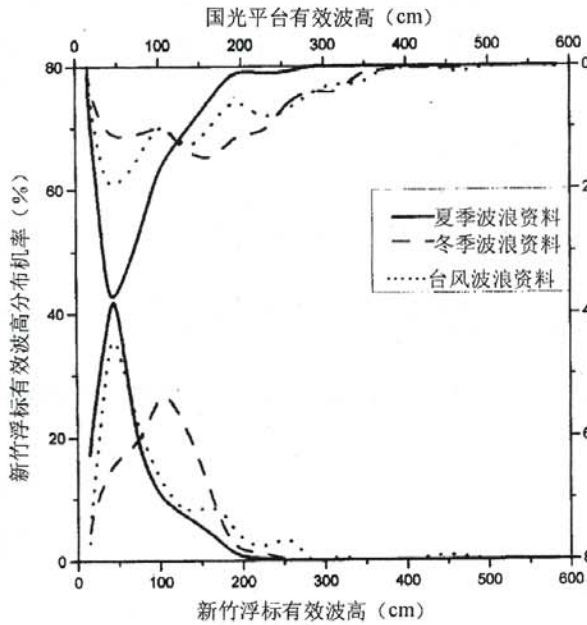


图 3 国光平台与新竹浮标各类型波高资料分布
 Fig.3 Distribution of wave height data at Kuokwang oil platform and Hsinchu data buoy

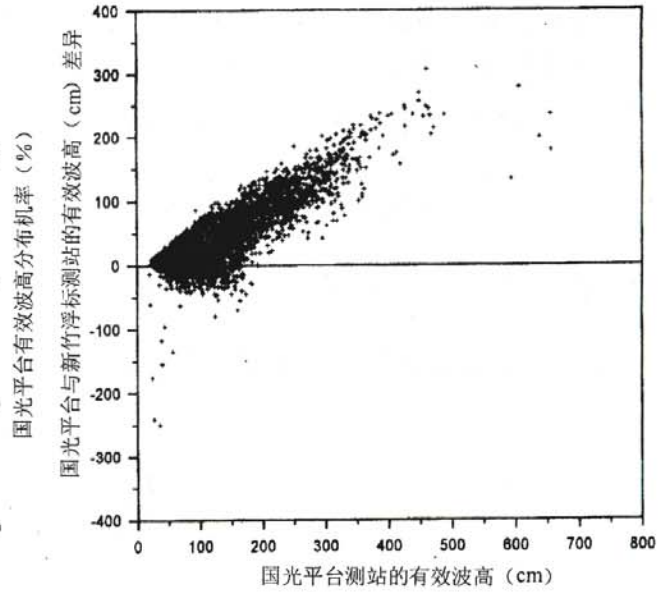


图 4 国光平台与新竹浮标测站的波高差异随波高之变化
 Fig.4 Change of difference of significant wave height with wave height at Kuokwang oil platform and Hsinchu data buoy

波龄式中 C_p 和 U_{10} 分别表示波速和风速。分析结果可看出台湾东西岸海域有截然不同之波浪成分，位于台湾东岸的花莲测站波浪波龄大部分大于 1.4，属于以涌浪成分为主的波浪。而西岸的台湾海峡为一风域受限的海域，即使在冬季寒潮期间，波浪亦难以无限制地成长，不论近岸（新竹浮标站）或深海（国光平台站）处，均量测到以风浪成分为主的波浪。

本文选用 1997 年 5 月至 1998 年 12 月期间，国光平台与新竹浮标同时都有波高观测纪录的资料为分析对象，而本文以下所指的波高均为有效波高。依据台湾气候特性，将分析资料分类为夏季、冬季与台风季节三种型态，其有效样本分别为 1 662、1 006 和 343 笔。图 3 绘制了三类资料的有效波高分布图，结果发现，国光平台夏季有效波高分布在 50~100 cm 之间的机率最大，冬季则是以 100~250 cm 出现的机率较高，新竹浮标也有相似之现象，因该测站位于近海，故仅是分布的量值小于国光平台所测得，由该图已大略能看出国光平台与新竹浮标波高资料可能有某关系存在。而由图 3 还观察到两测站波高间隐约相差某一特定值，这一猜测亦从图 4 得到了验证。图 4 说明了随着有效波高的增大，国光平台与新竹浮标间波高差异会随之增加，且只分布在某一特定波高差值范围内，换句话说，当波高愈大时，国光平台波高大于新竹浮标波高的程度愈高。

2 以马尔可夫链理论建立波高相关性模式

2.1 马尔可夫链理论

马尔可夫链 (Markov Chain) 为一序列的机率过程，考虑了两物理量在时间尺度上的机率成分。机率是用来衡量某一现象可能发生的程度，对于不能确定结果的决策抉择问题，机率通常是决定最佳决策的主要依据，譬如目前台湾的天气预报以降雨机率来预报隔日下雨的

可能程度,大部分的降雨径流预报模式亦都从事不确定分析,计算预报值的变异数,告知预报结果的变化范围。机率推算的方法很多,其中马尔可夫链为一个具有描述前后时刻不同状态特性的机率过程,其具有下列性质:

a) 每一次试验结果为状态 s_1, s_2, \dots, s_n 中之某一个。

b) 若某次试验之结果为一状态 s_i , 若 s_i 与另一试验(状态 s_j)之间的关系以机率表示为机率 p_{ij} , 则机率 p_{ij} 只与 s_i 以及 s_j 有关, 且 p_{ij} 称为由状态 s_i 转变至状态 s_j 之移转机率(transition probability), $P = [p_{ij}]$ 称为此马尔可夫链的移转机率矩阵(transition matrix), 定义如下:

$$P_{ij} = \frac{f_{i,j}}{\sum_{j=1}^m f_{i,j}} \quad i=1, 2 \dots n, \quad j=1, 2 \dots m \quad (1)$$

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \dots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix} \quad (2)$$

式中 $f_{i,j}$ 为在状态 i 的情况下转变为状态 j 之发生频率, 董东璟(1997)^[1]曾利用马尔可夫链理论来建立海气象资料品管系统中时间连续性的检定规范。本研究则计划将时间尺度的马尔可夫链引用至空间尺度上, 探讨同一时间空间上两测站波高的马尔可夫链关系。

2.2 马尔可夫链移转机率矩阵

移转机率矩阵藉由机率表示两波浪测站波高之间的关系。表1是所计算得之国光平台与新竹浮标冬季波高资料的移转机率, 由表1中各值可知新竹浮标发生某波高时, 在国光平台有可能发生各种波高之机率大小。从表1中对应状态之发生机率值, 可看出较高的机率值出现在表中左下三角型区域, 其所代表的意义是通常国光平台处测得的波高大于新竹浮标测站所测得的机率大, 此结果与当地之海域特性是一致的。另外, 每个状态的最大可能发生机率大部分均出现在移转机率表中邻近对角线处, 说明了这两个测站波高性质的同构型, 也增显了建立这两个测站间波高间相关性模式的可行性, 本文亦分析了夏季与台风类型的波浪资料, 也同样获得了类似的结果。

2.3 马尔可夫链波高相关性模式

前文说明了利用移转机率矩阵来建立两个测站间的波高机率关系, 一般而言, 利用上述移转机率矩阵表即可推知未测地点的波高状态, 然而为了应用上的方便, 本文引用期望值的观念, 导出马尔可夫链之波高相关性模式, 如下式:

$$\hat{H}_K = \sum_{j=1}^n \left[H_K^{j,L} + \frac{H_A - H_A^L}{H_A^u - H_A^L} \times (H_B^{j,U} - H_B^{j,L}) \right] \times P_j \quad (3)$$

式中: H_A 为已知的新竹浮标测站波高; H_A^L 为已知新竹浮标测站波高所属状态的波高

表 1 两站冬季波高资料之马尔可夫链转移机率矩阵

Tab.1 Markov's transition matrix of winter wave height data at Kuokwang oil platform and Hsinchu data buoy

		新竹浮标波高状态 (cm)										
		状态 1	状态 2	状态 3	状态 4	状态 5	状态 6	状态 7	状态 8	状态 9	状态 10	
		0~50	51~75	76~100	101~125	126~150	151~200	201~250	251~300	301~400	>400	
国 光 平 台 波 高 状 态 cm	状态 1	0~50	89.6	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	状态 2	51~75	9.2	24.6	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	状态 3	76~100	0.7	23.0	6.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	状态 4	101~125	0.4	24.1	16.1	4.1	0.0	0.0	0.0	0.0	0.0	0.0
	状态 5	126~150	0.1	13.4	23.0	21.6	6.4	0.9	0.0	0.0	0.0	0.0
	状态 6	151~200	0.0	12.8	35.4	41.9	28.8	7.9	0.0	0.0	0.0	0.0
	状态 7	201~250	0.0	1.1	17.4	24.3	41.7	22.8	0.0	0.0	0.0	0.0
	状态 8	251~300	0.0	0.0	0.6	7.2	12.8	32.5	15.4	0.0	0.0	0.0
	状态 9	301~400	0.0	0.0	0.0	0.9	10.3	34.2	30.8	100.0	0.0	0.0
	状态 10	>400	0.0	0.0	0.0	0.0	0.0	1.8	53.8	0.0	0.0	0.0

区间下限, H_A^U 则为该区间之上限; 而 $H_B^{j,U}$ 、 $H_B^{j,L}$ 则分别为欲推求之国光平台测站各状态 j ($j=1, 2 \dots n$) 的波高区间上下限; P_j 为各状态之机率值, \hat{H}_B 为推估得之国光平台测站波高。

3 马尔可夫链波高相关性模式之误差分析

为了解马尔可夫链波高相关性模式估算邻近测站波高之误差程度, 将本文模式的推估结果与实测值比较分析, 图 5 为冬季波高资料的分析结果, 图中推估值与理论值之相关系数为 0.74, 无因次的均方根误差 (root mean square error, RMSE) [定义如式 (4)] 为 28%, 同样的资料利用回归相关分析则均方根误差为 35%。无因次的均方根误差的定义如式 (4), 其中 H_{Kuo} 为国光平台实测波高, \hat{H}_{Kuo} 为藉由马尔可夫链模式推估得之国光平台波高。

$$RMSE = \sqrt{\frac{1}{n-1} \sum \left(\frac{H_{Kuo} - \hat{H}_{Kuo}}{H_{Kuo}} \right)^2} \quad (4)$$

另外, 由图 5 也看出了当波高较小时 (约 1.5 m 以下), 马尔可夫链模式推估结果有高估的情形, 因此为了完整地讨论三种类型波高资料的误差程度与分布, 以及探讨推估误差与波高的关系, 本文定义无因次误差 (Non-Dimensional Error, NDE) 来讨论, 如式 (5):

$$NDE = \frac{H_{Kuo} - \hat{H}_{Kuo}}{H_{Kuo}} \times 100\% \quad (5)$$

式中各符号意义与式 (4) 相同。NDE 值若为正表示马尔可夫链模式有低估的情形, NDE 若为负则表示马尔可夫链相关性模式的推估值高估。

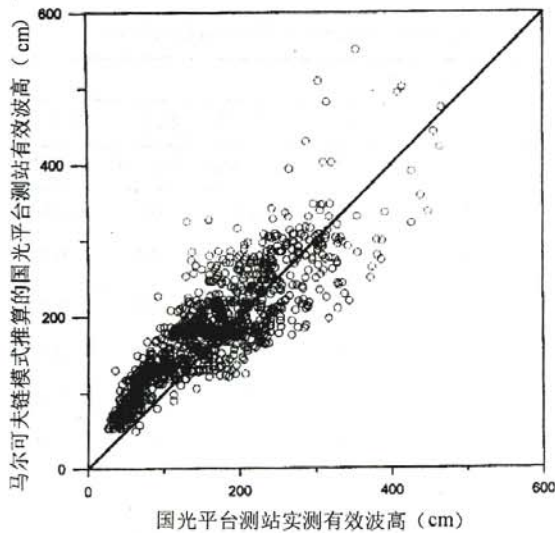


图5 国光平台实测波高与马尔可夫链模式推估值之比较 (冬季资料)

Fig.5 Comparison of observed wave height with Markov Chain wave height (winter wave data)

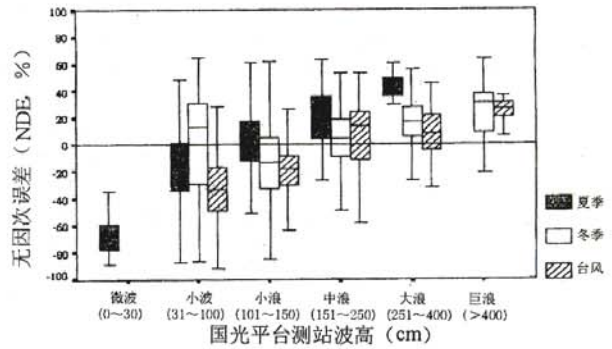


图6 各季节资料以马尔可夫链模式推估国光平台波高误差分布须盒图

Fig.6 Box plot of error distribution of wave height at Kuokwang oil platform estimated with Markov Chain using the data for each season

将夏季、冬季与台风三种类型波高资料的推估误差分布以图6之须盒图 (Box Plot) 表示, 盒子中的横直线代表误差分布之均值 (mean), 盒形上下边界代表误差分布之75%与25%统计值, 而其上下须表示出误差中之最大值与最小值。由图6不但可定性地了解误差的分布, 也可读出误差的量值。分析结果显示不论何种类型的波高资料, 马尔可夫链模式推估国光平台波高均有随着波高愈大, 推估误差愈小之趋势, 而在小波高时有高估之情形, 大波高时则有低估之现象。若分别对各种不同季节资料之推估结果计算其绝对误差平均值, 结果显示冬季资料之平均误差15%为最小, 夏季平均误差36%为最大。由于冬季与台风期间之大波高为工程设计之重要参考, 本文计算有效波高1.0 m以上之恶劣海况资料, 即图6中的“小浪”、“中浪”、“大浪”与“巨浪”等组别, 其平均误差分别为21% (夏季资料)、16% (冬季资料) 和18% (台风资料), 此误差程度虽尚可接受, 但仍须后续研究继续改进减低其推估误差。

4 马尔可夫链相关模式与回归相关分析的不确定性程度比较

4.1 马尔可夫链模式之不确定分析

因自然界现象具有随机性, 真正的结果往往无法精准地预测, 通常即使知道某一随机变量值亦无法准确地预测另一变量值。利用一数学模式来推测或预估某一随机变量值, 所估算得者仅为在该应变条件下可能之平均结果, 该推估值应有一变动范围, 并非没有误差, 亦即有其推估之不确定性 (uncertainty), 造成此不确定性的原因很多, 如样本数目不足、量测误差、模式考虑不周延等^[6]。不确定性程度的量化通常以估算值的标准偏差 (standard deviation) 来表示, 标准偏差能表达出估算值之分布是否集中于某一范围或是资料点呈现出分散的情况, 本文藉此来表示波高推估之可信任程度。

本文利用马尔可夫链模式由新竹浮标波高推估之国光平台波高与实测国光平台波高取离差, 再做加权平均, 将所得波高差值分散性的量为实测波高与推估波高间之变异数 (variance

)。将变异数开根号即可得标准差如式 (6) 所示：

$$S_{Kuo} = \sqrt{\sum_{j=1}^{10} \left\{ H_{Kuo}^{jL} + \frac{H_{Hsi} - H_{Hsi}^L}{H_{Hsi}^u - H_{Hsi}^L} \times (H_{Kuo}^u - H_{Kuo}^{jL}) \right\} - \hat{H}_{Kuo}}^2 \times P_j \quad (6)$$

式中各符号代表意义如同 2.3 节所述。

4.2 回归相关分析及其不确定性分析

4.2.1 回归相关分析

传统上常用回归分析来探讨两物理量之关系，本文以此做为马尔可夫链波高相关性模式之比较对象。由于波浪传递所需的时间，理论上，两测站间的波高相关性应有一时间延迟 (time lag)，因此在对两站波高资料做回归分析之前，本文利用交叉相关函数 (cross correlation function) 来推求稽延时间时，两地波高有最高之相关程度，交叉相关系数定义如式 (7)、式 (8)：

$$r_{xy} = \frac{C_{xy}(k \times \Delta t)}{\sigma_x \sigma_y} \quad (7)$$

$$C_{xy}(k \times \Delta t) = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y}), \quad k = 0, 1, 2, \dots \quad (8)$$

上式中 r_{xy} 为两测站间的交叉相关函数， C_{xy} 为两站波高共变异数 (covariance)， n 为样本数； k 为时间稽延项数，实际的时间稽延为 $k \times \Delta t$ 。

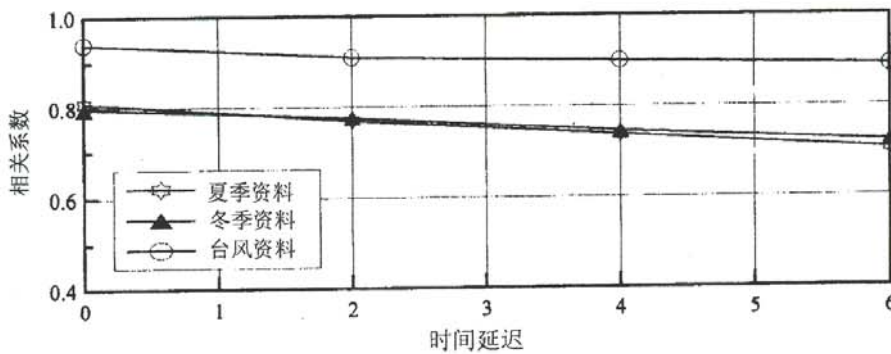


图 7 国光平台与新竹浮标之不同稽延时间的波高交叉相关系数

Fig.7 Cross correlation coefficient of wave height with different time lag at Kuokwang oil platform and Hsinchu data buoy

图 7 结果发现各季节波高资料均以零小时之稽延时间为最佳，且相关系数大于 0.75，表示了在同一时刻之国光平台与新竹浮标波高有最佳相关性。

本文以线性 (linear)、对数 (log)、指数 (exponential) 及幂次 (power) 等 4 种常用之回归模式来探讨国光平台与新竹浮标波高资

表 2 各季节波高资料回归分析成果表

Tab.2 Regressive results of wave height data

	夏季波高资料	冬季波高资料	台风波高资料
线性	0.80 / 33 %	0.87 / 32 %	0.94 / 31 %
对数	0.79 / 41 %	0.83 / 42 %	0.89 / 38 %
指数	0.79 / 43 %	0.86 / 37 %	0.85 / 45 %
幂次	0.83 / 34 %	0.91 / 30 %	0.93 / 27 %

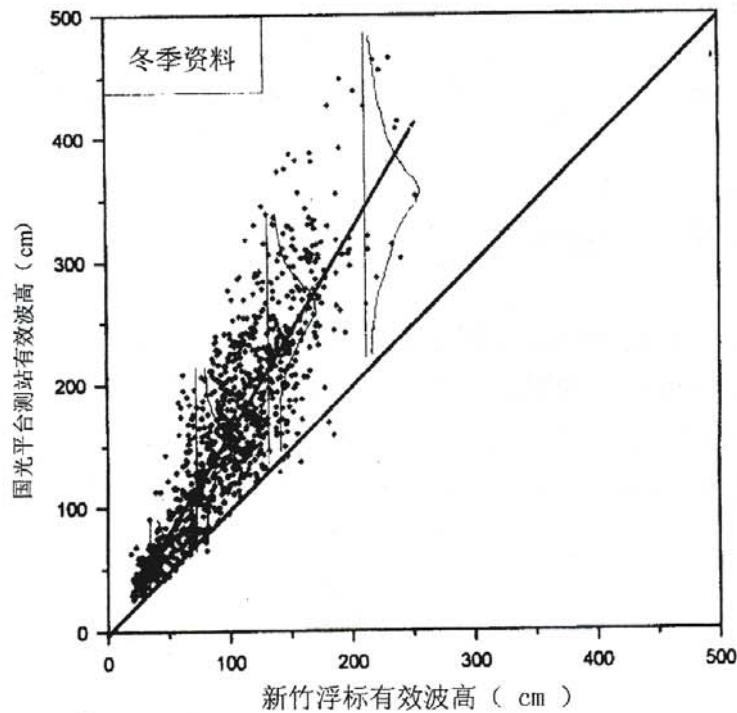


图 8 国光平台与新竹浮标夏季波高幂次回归关系

Fig. 8 Power regressive relation of summer wave height at Kuokwang oil platform and Hsinchu data buoy

料的回归相关模式，以相关系数和前文所定义之无因次均方根误差作为模式选择的评判指针。表 2 显示两站间之三种数据类型的波高资料均以幂次型态回归模式，优于其它回归模式，其中以夏季波高间之相关系数最低，冬季与台风期间资料相关系数均在 0.9 以上。图 8 为新竹浮标与国光平台冬季波高之回归结果。由图 8 中可看出推估值的变异程度随波高而变大，因此接续的下节将从事回归模式的不确定分析。

4.2.2 回归模式之不确定性分析

回归模式的不确定性分析亦是计算其变异数作为评判不确定程度之依据，通常回归模式之推估变异数估算方法有两种，由回归方程所求得之估算量散布范围均匀时，称此回归模式具有固定之变异数；相反的，若回归方程估算值具有明显之分散状况时，则称此回归模式的变异数不为常数。如图 8 所示，本文回归模式所估算的波高呈一放射状散开，说明该回归模式所得之估算值的变异数随自变量而增大，其标准差可表示成自变量（新竹浮标波高， H_{Hsi} ）的函数，如式（9）：

$$S = \sigma \cdot g(H_{Hsi}) \quad (9)$$

式中 σ 为常数； n 为资料笔数。参考 Ang (1984)^[5] 的做法，本文假设变异数较小时应占有较大重要性，应加权处理，令加权系数如式（10）：

$$w_i = \frac{1}{g(H_{Hsi})^2} \quad (10)$$

根据最小平方法令平方差总合 [式（11）] 最小时求得 σ 参数，如式（12）：

$$\Delta^2 = \sum w_i [\ln(H_{Kuo})_i - \ln(\alpha) - \beta \ln(H_{Hsi})_i]^2 \quad (11)$$

$$\sigma = \sqrt{\frac{\sum w_i [\ln(H_{Kuo})_i - \ln(\alpha) - \beta \ln(H_{Hsi})_i]^2}{n-2}} \quad (12)$$

式中 α 与 β 为本文最佳回归模式 $\hat{H}_{Kuo} = \alpha H_{Hsi}^\beta$ 中的参数。

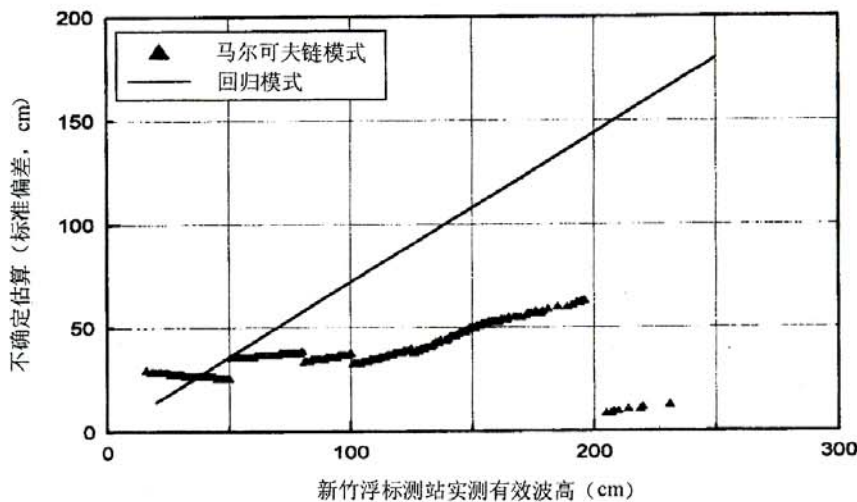


图9 马尔可夫链与回归分析冬季波高推估结果之不确定程度比较

Fig.9 Comparison of uncertainty between winter wave heights estimated with Markov Chain and regressive model

4.3 马尔可夫链与回归分析之不确定性比较

本文提出之马尔可夫链模式与回归相关分析的波高估算不确定程度比较如图9，由于回归模式之变异数并非定值，造成其不确定程度随波高而增大，如图中斜直线，而马尔可夫链模式之不确定则与各状态之样本数目有关，如图中新竹浮标波高在200~250 cm状态之标准偏差反而异常小之原因系因为该组之样本数不足所致，更丰富的历史资料将可提供更可靠的结果。图9的分析结果显示回归分析的不确定程度随波高增加而急剧增大，而马尔可夫链模式推估值之标准偏差则在50 cm上下，不确定程度低于回归模式，事实上此即为建立马尔可夫链模式时考虑各波高状态发生机率所获得之效应。

5 结论

本文引用已有的马尔可夫链观念，加入相应的分析与公式推导后，建立了一个新的波高相关性模式，经过与传统回归相关分析的比较，获得以下结论：

a) 马尔可夫链模式推估误差与海况有高度相关，推估误差随波高增大而减小，而在小波高之平静海况下本文模式略为高估，在波涛汹涌之大波高时，则略有低估之情形，大波高期间之波高平均推估误差低于20%。

b) 马尔可夫链波高相关性模式的推估误差量值虽与回归分析误差相近，无大幅改进，但以马尔可夫链波高相关性模式推估波高之不确定程度则远低于传统回归分析所得，显示了

马尔可夫链模式确实为进行相关性分析可行的方法之一。

本文成文期间承蒙台湾气象局提供相关资料, 特此致谢。大连理工大学俞聿修教授对于本文的建议, 令本文更加充分, 在此一并致上最高谢意。

参考文献

- [1] 董东璟, 庄士贤, 高家俊. 海气象观测资料品管系统之建立 [C]. 台湾第十九届海洋工程研讨会论文集, 1997, 477~484
- [2] Alfredo H-S Ang and Wilson H Tang. Probability Concepts: in Engineering Planning and Design [R]. John Wiley & Sons Inc., 1986
- [3] Dynkin E B. Markov Processes Volume I [M]. Academic Press, 1965
- [4] Goda Y. Random seas and design of maritime structure [M]. University of Tokyo Press, 1985
- [5] Kao C C, Laurence Z H, Chuang Y P Lin *et al.* An Introduction to the Operational Data Buoy System in Taiwan [C]. in: Wind & Wave Climate. Proceeding of MEDCOAST Conference, Antalya, Turkey, 1999, 33~39
- [6] Charles T H. Statistical Methods in Hydrology [M]. Iowa State University Press, 1982

作者简介: 董东璟, 男 (1973—), 硕士, 台湾成功大学水利及海洋工程研究所博士班研究生。专业及研究方向为波浪观测实务、风浪理论与波浪统计分析; 博士研究则主攻波浪观测中的波浪遥感分析。发表论文 4 篇, 会议论文 13 篇。

Development of Wave Height Correlation Model by Using Markov Chain Theory

DONG Dongjing¹, WU Lizhong¹, GAO Jiajun^{1, 2}, ZHUAN Shixian² and LI Bianjun³

(1. Department of Hydraulics and Ocean Engineering, 2. Coastal Ocean Monitoring Center, Cheng-Kung University, Tainan, Taiwan, China; 3. Marine Meteorology Center in Taiwan Weather Bureau, Taibei, Taiwan, China)

Abstract In order to extend the short-term wave records, the statistical correlation model of wave heights between two close wave stations is investigated and established by using the Markov Chain theory. The field wave data from the obsolescent Kuokwang oil platform, which locates at the middle of the Taiwan Straits and the operational Hsinchu data buoy station, which locates at a distance of 25 km eastern from the platform are analysed in this study. As a comparison result, the Markov Chain wave height correlation model has lower uncertainty results than the traditional regression model. The estimated errors are highly dependent on the wave height and wave category. For the interested high wave height, the error is less than 20 %.

Keywords Markov Chain theory; wave height correlation; uncertainty analysis