

海洋工程中极值水位估计的一种条件分布联合概率方法*

方 国 洪

(中国科学院海洋研究所)

王 骥 贾绍德 焦宗凤 叶 安 乐

(国家海洋局海洋科技情报研究所)

(青 岛 海 洋 大 学)

一、前 言

在海洋工程建设(如港口建设、沿岸工程和海上平台建筑等)中,必须对正常条件(即作业条件)和极端条件(即恶劣环境条件)下的海洋环境参数作出估计。对于海洋水位而言,如《港口工程规范》(以下简称《规范》)中的设计水位和乘潮水位等即属于前者,校核水位即属于后者。正常条件下的水位参数比较容易确定,因为这种条件下的观测数据比较容易获得,有比较丰富的已知资料。极值水位是指若干年内才有可能出现的高水位或低水位,例如在《规范》中要求50年一遇的水位作为校核水位。所以如依靠直接观测,必须具备至少几十年的观测资料,这一要求只有在有长期验潮站的港口才能达到。如资料系列不够长,则必须采用一定的方法对它们的数值作出估计。

在有相当长期观测资料的情况下,常常采用极值分布的方法,即每年取一个最高(或最低)水位,然后拟合某种理论或经验的分布曲线,从而得出不同重现期的极高(或极低)水位。这种方法所得结果比较可靠(但也不是没有问题,详见第五节),但对资料要求高,特别在新建港口或海上工程,常常不能采用。

在有数年观测资料的情况下,可以采用同步差比法(见《规范》)。当拟建港和主港潮汐性质和风暴潮特性相近,特别是两地距离较近时,这种方法常常能给出良好的结果,不失为一种有效的方法,但其应用受到一定的条件限制。

在只有短期资料的情况下,《规范》要求首先计算出设计水位,然后在设计水位上加减一个常数推出校核水位。这个常数是依据附近长期验潮站资料得出的,因而也只有在附近有长期验潮站的条件下使用。

近年来,国外提出了一种“联合概率法”(Pugh and Vassie, 1978, 1980)。这种方法把水位中的潮汐水位(这里指的是可用调和方法预报的潮位,其中也包括周期性的气象潮)和余水位(即实测水位减去预报水位)分开,分别求其分布,然后再用联合概率法回来求合成水位的分布,从而得出极值水位高度。这种方法的优点是能够充分利用水位观

* 中国科学院海洋研究所调查研究报告第2146号。
收稿日期: 1991年11月25日。

测资料,因而可以在观测时间较短的情况下得出比传统方法更稳定的结果。在 Pugh 和 Vassie 所提出的方法中,潮汐和余水位被看作是互相独立的。这一假定在深水海区较接近实际,但在某些浅海区则与实际有较大出入。例如 Pugh 和 Vassie 发现,对位于泰晤士河的绍森德港,由联合概率法按上述假定得出的百年一遇的高水位比由传统极值分布方法得出的数值大约高 60cm; 而传统方法的结果与 49 年期间观测到的最高水位较接近。Pugh 和 Vassie 在后来的文章中对绍森德港的问题作了专门处理,得出较合理的结果,但基本上是采用经验的方法,在一般实际中使用是困难的。

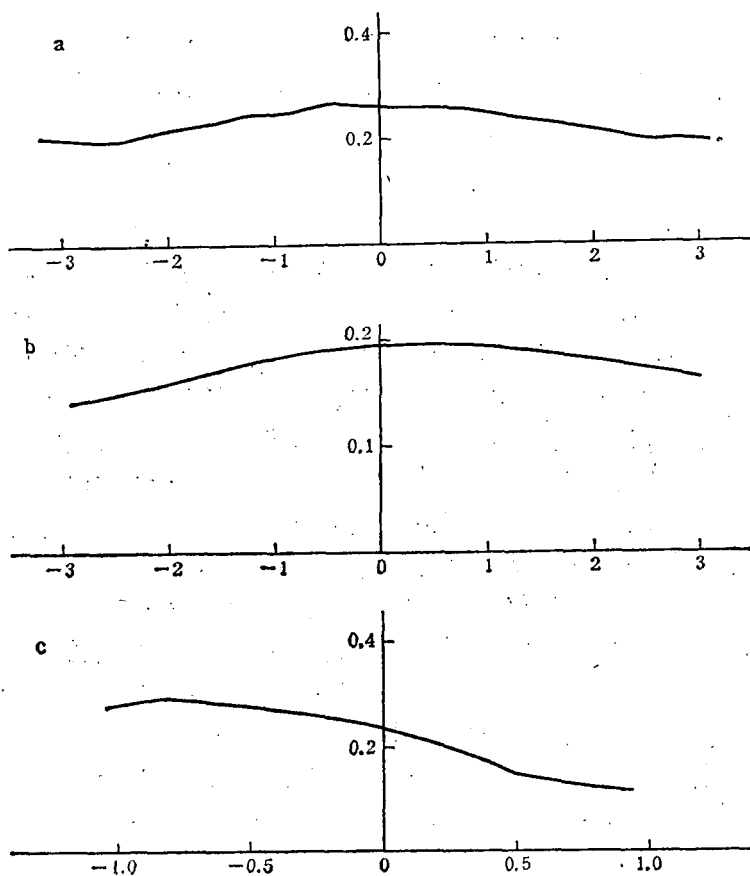


图 1 英国绍森德(a)、我国厦门(b)和秦皇岛(c)的余水位标准差(纵坐标,单位: m)和潮汐水位(横坐标,单位: m)的关系

正如 Pugh 和 Vassie 指出的,潮汐和风暴潮的非线性相互作用是造成绍森德港计算水位偏高的原因。实际分析表明,在潮汐高水位和低水位条件下余水位的标准差明显小于中水位条件下的余水位标准差(图 1a)。类似的现象在我国的厦门亦表现得相当典型(图 1b)。

潮汐和风暴潮的相互作用曾经受到许多海洋工作者的注意。我国具有宽广的大陆架,潮汐和风暴潮都十分显著,它们之间的相互作用也就十分明显。例如,丁文兰和刘风树(1986)对 6208 号台风潮与天文潮相互作用进行数值模拟,发现在渤海考虑风暴潮与天

文潮相互作用后所得增减水与实测余水位基本一致, 而不考虑这种相互作用的单纯风暴潮与实测余水位差别较大。在其他的一些研究中表明, 在相同风场情况下, 不同天文潮高度所对应的风增水不同, 高潮位情况下的风增水小于低潮位下的风增水。这些都说明潮汐水位与余水位不是相互独立的。潮汐和风暴潮的相互作用是引起不同潮汐水位条件下余水位分布差异的一个重要原因, 但是并不是唯一因素。实际上, 虽然风暴潮的发生有其随机性, 但是其统计特征是有一定规律的。例如我国冬季多寒潮, 且风暴频繁, 故余水位标准差在冬季一般比夏季大。这种现象在北方港口特别明显。如果潮汐本身也具有明显季节性变化的话, 则潮汐与余水位就表现出一定的统计上的相关关系。在我国北方潮汐即具有显著年变化, 它主要是由年周期分潮 S_2 引起的。这个分潮使得冬季水位变低, 夏季变高。北方冬天潮汐水位偏低和余水位标准差偏大都是起源于频繁的寒潮。一个因素同时引起两种效应, 这两种效应当然不会是互相独立的。图 1c 是秦皇岛余水位标准差与潮汐水位之间的关系。从图中可以明显看出低潮汐水位对应着大的余水位标准差。

图 1 给出的是一些相当典型的例子, 目的是说明在有些港口和海区确实不能采用余水位和潮汐相互独立的假定。但对相当多的港口和海区, 这种假定还是可以接受的。

为了考虑潮汐和余水位的这种相关性, 我们曾将联合概率法予以推广, 提出了条件分布联合概率法(方国洪、王骥, 1987)。这种方法不论潮汐和余水位是否相互独立均可应用, 具有更强的适应性。本文根据该文的基本想法, 选择了我国沿海资料较长, 并具有较好代表性的 10 个长期验潮站共 286 年的潮汐逐时记录, 进行了分析研究, 对该文提出的方法作了修改和补充。

二、条件分布联合概率方法

1. 潮汐和余水位联合概率分布与总水位分布之间的关系

我们按一般习惯把水位分成三部分。首先是平均水位, 它包含长期变化趋势。第二部分为可由潮汐预报得出的高度, 也就是前面所称的潮汐水位, 它主要是由天体引潮力引起, 有时也被简单地叫做天文潮, 但实际上它也包含由气象条件的周期性变化所引起的水位周期性变化部分。第三部分是总的水位扣除前两部分后的余差, 也就是前面所称的余水位。这个余差除了包含由天气过程等所引起的海面高度非周期性变化外, 还包含潮汐和非潮汐因素的耦合效应, 甚至还可能包含在实际预报中未能考虑到的潮汐因素, 如一部分小的浅水分潮等。在联合概率法里要尽可能把余水位中的潮汐部分消除掉, 这部分传统上也称为非潮汐水位。

严格地讲, 在研究多年一遇水位时, 平均水位的长期变化不应当被简单地忽略, 当重现期较长时尤其如此。人们推测, 人类活动的增加加强了温室效应, 因而气温将升高。这又将导致海水温度的升高和极地冰盖的融化, 从而将使海水水位升高。对未来几十年至一百年海水水位升高的速度尚无一致的看法。如果按最严重的估计, 水位升高的数值是不应当忽视的。本文将不涉及这个问题, 仅仅研究相对于平均水位的各种水位高度。在工程设计中, 如果工程结构使用年限长, 则应当考虑平均水位的长期变化。由于气候和海

洋环境的演变,潮汐水位和余水位的统计特征亦会有长期变化,但是人们对它的了解比对未来海平面变化的了解更少。不过似乎可以认为,这种变化的重要性要比海平面变化的重要性来得小而可不必考虑。

对于充分长的一段时间区间内,假定我们无倾向性地在任意一个时刻观测水位高度,则观测到的潮汐水位和余水位均是随机的。于是可以把它们都看成是随机变量,并分别用 X 和 Y 记之,而总的水位(按前面约定,已减去平均水位)用 Z 记之,且有

$$Z = X + Y \quad (2.1)$$

如二元随机变量 (X, Y) 的联合概率密度函数为 $f(x, y)$ 则 Z 的分布函数为

$$\begin{aligned} F(z) &= \iint_{x+y \leq z} f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{z-x} f(x, y) dy \right] dx \end{aligned} \quad (2.2)$$

在实际估计联合概率密度函数时,一个比较方便的办法是通过条件概率密度函数 $f(y|x)$ 得到。此时

$$f(x, y) = f(y|x)f'(x) \quad (2.3)$$

其中

$$f'(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (2.4)$$

为 X 的边际密度函数。在实际工作中 $f'(x)$ 很容易得出,如像 Pugh 和 Vassie(1978)所作的那样,可对 19 年逐时潮高推算值进行统计得出。

如果 X 和 Y 互相独立,则 $f(y|x)$ 与 x 无关,且有

$$f(y|x) = f''(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad (2.5)$$

其中 $f''(y)$ 为 Y 的边际密度函数。在这种情况下,式(2.2)简化为

$$F(z) = \int_{-\infty}^z \left[\int_{-\infty}^{\infty} f''(t-x)f'(x) dx \right] dt \quad (2.6)$$

该式是 Pugh 和 Vassie (1978, 1980) 计算的出发点。如上一节所述,我们将采用更一般的关系式(2.2)。

2. 数值计算方法

在实际工作中,常常不能得出各种分布函数或密度函数的解析形式,因而必须采用离散化了的数值形式表示。即使我们能够用某种解析形式表示这些函数,通常也不能得出式(2.2)中 $F(z)$ 的显式解析式。因而在实际工作中总是需要利用数值的方法。

如所周知,潮汐水位变化有一定的范围,即存在一个最小值 $x_{\min}$ 和最大值 $x_{\max}$, 这些值一般不宜只由观测期间的最高和最低天文潮位确定,而应当至少推算 19 年潮汐值,取其最小值和最大值,或分别略小和略大于这些值。风暴潮大小常常被看作是不受限制的,例如在极值分布中就可以把拟合曲线(或直线)无限制地延长出去。但是实际上风暴潮不可能是无限的。当然,对于风暴潮我们也不应当只把实际观测到的最大增减水取作其变化范围,而应当取得大一些。在本工作中我们取 $\pm 10\text{m}$, 这显然是够大的范围。范围取得太大只会增加一些计算量而不会带来计算误差。我们记风暴潮的取值范围为 $(y_{\min},$

$y_{\max}$)。在数值计算中需将取值范围划分为若干个区间, 设潮汐分为 J 个等宽区间, 区间的宽度为 $\Delta x = (x_{\max} - x_{\min})/J$, 各个区间的中心位置为 $x_1, x_2, \dots, x_J$, 它们分别代表了区间 $(x_1 - \Delta x/2, x_1 + \Delta x/2), (x_2 - \Delta x/2, x_2 + \Delta x/2), \dots, (x_J - \Delta x/2, x_J + \Delta x/2)$ 。同样, 如余水位取值范围分为 K 个区间, 则区间宽度 $\Delta y = (y_{\max} - y_{\min})/K$, 各区间的中心位置为 $y_1, y_2, \dots, y_K$, 它们分别代表了区间 $(y_1 - \Delta y/2, y_1 + \Delta y/2), (y_2 - \Delta y/2, y_2 + \Delta y/2), \dots, (y_K - \Delta y/2, y_K + \Delta y/2)$ 。于是, 在 (x, y) 平面上, 由 $x = x_{\min}, x = x_{\max}, y = y_{\min}$ 和 $y = y_{\max}$ 四条直线所围的区域便被划分为 $J \times K$ 个方块。

如联合概率密度函数为 $f(x, y)$, 则出现在第 (j, k) 个方块的概率为

$$p_{j,k} = \int_{x_j - \Delta x/2}^{x_j + \Delta x/2} \int_{y_k - \Delta y/2}^{y_k + \Delta y/2} f(x, y) dy dx \quad (2.7)$$

在 $x = x_j$ (代表 x 落在区间 $x_j \pm \Delta x/2$ 内) 条件下, y 值出现在区间 $y_k \pm \Delta y/2$ 内的概率 $p_{k|i}$ 由下式确定:

$$p_{k|i} = \int_{y_k - \Delta y/2}^{y_k + \Delta y/2} f(y | x_j) dy \quad (2.8)$$

与前面边际分布相对应, 有

$$p'_j = \sum_{k=1}^K p_{j,k} \text{ 和 } p''_k = \sum_{j=1}^J p_{j,k} \quad (2.9)$$

前者表示 x 出现在区间 $x_j \pm \Delta x/2$ 内的频率 (不管 y 为何值); 后者类似。与式 (2.3) 对应的离散化关系式为

$$p_{i,k} = p_{k|i} p'_i \quad (2.10)$$

在应用式 (2.2) 时, 最好取 $\Delta x = \Delta y$ 。此时总水位 Z 的取值范围分为 I 个区间, $I = J + K - 1$, 区间宽度 $\Delta z = \Delta x = \Delta y$ 。第 1 个区间中心点的数值为 $z_1 = x_1 + y_1$, 总水位出现在 $z_1 \pm \Delta z/2$ 范围内的概率近似为 $p_{1,10}$ 。第 i 个区间中心点的数值为 $z_i = x_1 + y_1 + (i-1)\Delta z$, 总水位出现在 $z_i \pm \Delta z/2$ 范围内的概率为

$$q_i = \sum_{j+k=i+1} p_{j,k} = \sum_{j=1}^i p_{j,i-j+1} \quad (2.11)$$

而 Z 取值大于 $z_L + \Delta z/2$ 的概率 F_L 为

$$F_L = P(Z > z_L + \Delta z/2) = \sum_{i=L+1}^I q_i \quad (2.12)$$

如果这个值很小, 则总水位 $z_L + \Delta z/2$ 所对应的重现期为

$$T_L = 1/F_L \quad (2.13)$$

为了说明以上各量的意义, 我们给出一个例子。设现有 200 个水位观测值, 它们减去预报潮汐水位后可得 200 个余水位值, 每个余水位都对应着一个潮汐水位。如果取潮汐和余水位的间隔都是 0.1m, 则可得出现在各个方块内的观测个数如表 1(a) 所示, 表中数字除以总记录个数即得出现于各方块内的频率, 即前面的 $p_{j,k}$, 如表 1(b) 所示。

如果想求出总水位 $Z = X + Y$ 出现频率的分布, 则应沿着从左下角到右上角的斜线将 $p_{j,k}$ 相加。例如欲知总水位等于 -0.1 的出现频率, 则因为当 $x = 0.1, y = -0.2$ 或

表 1 潮汐和余水位联合分布例子

(a) 出现次数							
k y_k	1 -0.2	2 -0.1	3 0.0	4 0.1	5 0.2	合计	
j x_j							
1 -0.2	0	5	8	3	0	16	
2 -0.1	0	2	10	15	4	31	
3 0.0	2	20	30	10	3	65	
4 0.1	3	8	20	11	4	46	
5 0.2	3	2	12	6	2	25	
6 0.3	2	5	3	7	0	17	

(b) 出现频率 $p_{i,k}$							
1 -0.2	0.000	0.025	0.040	0.015	0.000	0.080	
2 -0.1	0.000	0.010	0.050	0.075	0.020	0.155	
3 0.0	0.010	0.100	0.150	0.050	0.015	0.325	
4 0.1	0.015	0.040	0.100	0.055	0.020	0.230	
5 0.2	0.015	0.010	0.060	0.030	0.010	0.125	
6 0.3	0.010	0.025	0.015	0.035	0.000	0.085	

(c) 总水位的出现频率 q_i										
i z_i	1 -0.4	2 -0.3	3 -0.2	4 -0.1	5 0.0	6 0.1	7 0.2	8 0.3	9 0.4	10 0.5
q_i	0.000	0.025	0.060	0.180	0.280	0.190	0.155	0.065	0.045	0.000

(d) 总水位的超值分布 F_i										
i $z_i + \Delta z/2$	1 -0.35	2 -0.25	3 -0.15	4 -0.05	5 0.05	6 0.15	7 0.25	8 0.35	9 0.45	10 0.55
F_i	1.000	0.975	0.915	0.735	0.455	0.265	0.110	0.045	0.000	0.000

$x = 0.0, y = -0.1$ 或 $x = -0.1, y = 0.0$ 或 $x = -0.2, y = 0.1$ 都能得出 $z = -0.1$, 故出现频率应当是 $0.015 + 0.100 + 0.050 + 0.015 = 0.180$ 。超值分布 F_i 很容易由各间隔的出现频率得出, 它们的结果见表 1(c) 和 (d)。

下面考察一下假定余水位与潮汐无关时的计算与结果。此时 Y 取 $-0.2, -0.1, \dots, 0.2$ 的频率被认为与 X 无关, 其值应当分别为 $0.050, 0.210, 0.415, 0.260, 0.0657$ 。同时表 1(b) 应换作表 2(a), 而总水位的出现频率和超值分布则如表 2(b) 和 2(c) 所示。

我们可以看到, 在表 1 中当 X 取两端的值 -0.2 和 0.3 时, 余水位的分布比较集中于中等数值, 因而总水位 Z 出现两端数值的频率亦低 [表 1(c) 中 $z_i = -0.4$ 和 0.5 时 $q_i = 0$]。 X 和 Y 相互独立这一假定的后果是把两端和中等潮汐情况下出现的余水位分布看成一样, 于是总水位 Z 在两端出现的频率增高了 [表 2(b) 中 $z_i = -0.4$ 和 0.5 时 q_i 分别

表 2 在潮汐和余水位假定相互独立时的分布

(a) 出现频率 $p_{j,k}$						
j	k y_k	1 -0.2	2 -0.1	3 0.0	4 0.1	5 0.2
	x_j					
1	-0.2	0.004	0.017	0.033	0.021	0.005
2	-0.1	0.008	0.033	0.064	0.040	0.010
3	0.0	0.016	0.068	0.135	0.085	0.021
4	0.1	0.012	0.048	0.096	0.060	0.015
5	0.2	0.006	0.026	0.052	0.032	0.008
6	0.3	0.004	0.018	0.035	0.022	0.006

(b) 总水位的出现频率 q_i										
i z_i	1 -0.4	2 -0.3	3 -0.2	4 -0.1	5 0.0	6 0.1	7 0.2	8 0.3	9 0.4	10 0.5
q_i	0.004	0.025	0.082	0.165	0.234	0.221	0.151	0.082	0.030	0.006

(c) 总水位的超值分布 F_i										
i $z_i + \Delta z/2$	1 -0.35	2 -0.25	3 -0.15	4 -0.05	5 0.05	6 0.15	7 0.25	8 0.35	9 0.45	10 0.55
F_i	0.996	0.971	0.889	0.724	0.490	0.269	0.118	0.036	0.006	0.000

为 0.004 和 0.006]。

从上述例子可以看到,如果认为潮汐和余水位相互独立,则各个余水位值对总水位的分布均产生影响,特别大的增减水,不论它们出现在潮汐高水位或低水位,均对总水位的极值分布产生影响,而年极值统计中只利用了少数资料,这表明,联合概率法能够比极值法更充分地利用水位观测资料。如果潮汐和余水位确实是相互独立的话,联合概率法的结果应具更好的代表性。但是如果余水位的分布与潮汐水位有关,则表 2 这个计算过程不会取得良好效果。

表 1 计算过程能够考虑到不同潮汐水位条件下余水位的不同分布,这是其优点。但是应当注意到,在这个计算过程中,中、低潮汐水位情况下出现的大增水对总水位的高水位极值分布不起作用;而中、高潮汐水位情况下出现的大减水对总水位的低水位极值分布也不起作用。也就是说,在极值分布计算中大部分大的增减水信息并未利用起来,这是这种计算过程的缺点。为了克服这个缺点,本文并不完全采用这种方法,列出表 1 计算过程只是为了说明基本原理。在介绍实际计算过程之前,我们先研究一下我国沿海水位分布的特点,特别是在不同潮汐水位情况下余水位分布的异同。

三、中国沿海余水位的统计分布特征

由于余水位的资料数量很大,这里选择了资料序列较长,具有代表性的 10 个长期验潮站。各站的站名、位置及观测资料年限见图 2。这 10 个站资料总长度为 286 年。葫芦岛和秦皇岛代表渤海情况,秦皇岛具有混合全日潮性质;青岛代表黄海;吴淞代表东海河口站;绿华山代表东海;厦门代表台湾海峡;汕头代表南海受陆架影响较大的沿岸站;黄埔代表南海河口站;榆林代表南海面对开阔海域的沿岸站;涠洲代表北部湾,这里全日潮占主要地位。

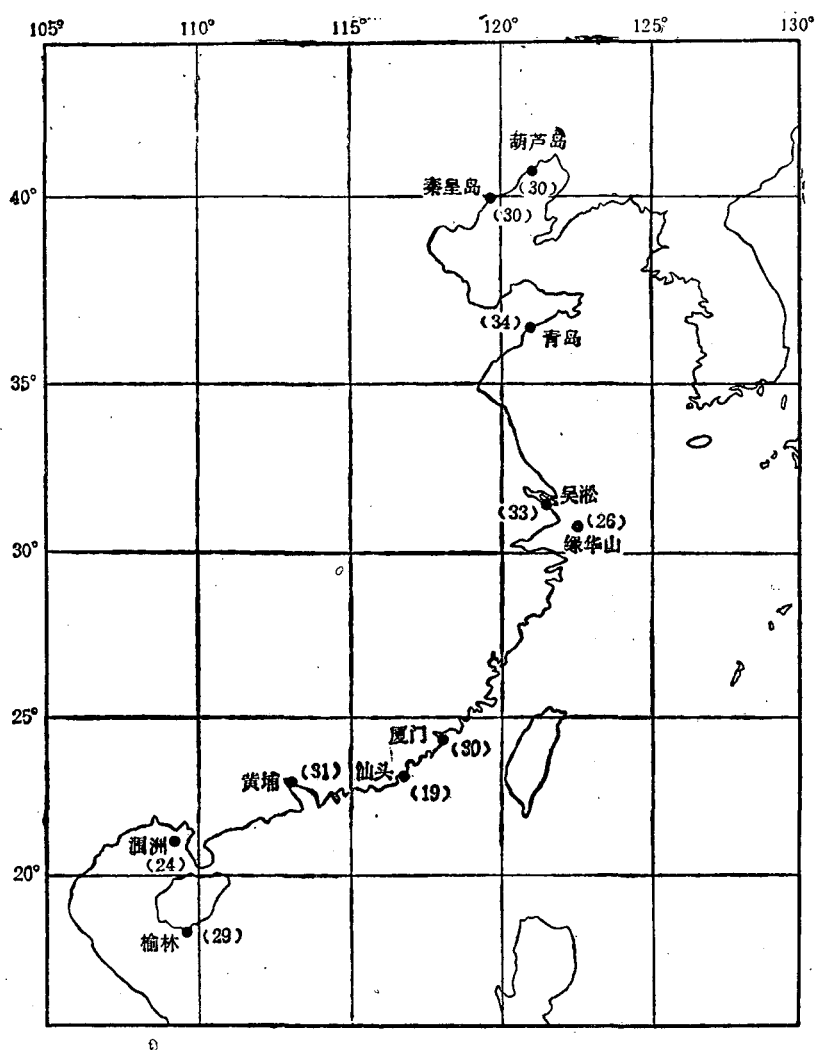


图 2 本文重点研究的验潮站站位
图中数字为资料长度(年)

1. 潮汐和余水位的分离

对于潮汐,我们采用调和分潮表示式:

$$x(t) = \sum f_i H_i \cos(\omega_i t + V_i + u_i - g_i) \quad (3.1)$$

其中 t 为时间; i 代表分潮; H 和 g 为调和常数——振幅和迟角; ω 为分潮角速度; f 为交点因子; V 为分潮初相角; u 为相角的交点订正。

首先对各年资料进行调和分潮分析,求出潮汐调和常数。采用的方法根据王骥和方国洪

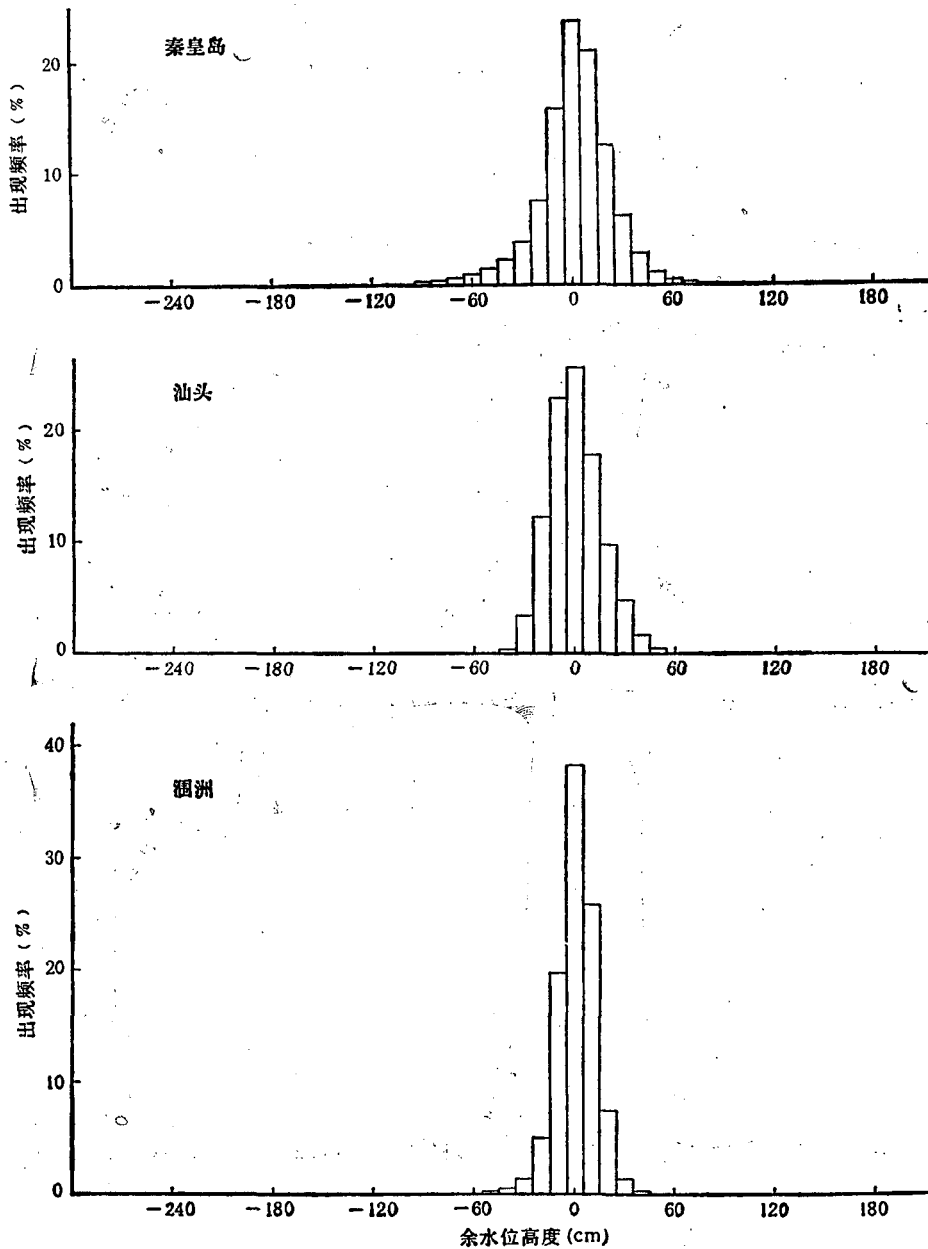


图3 秦皇岛、汕头和涠洲不同余水位高度的出现频率

(1981)。分潮个数为 111 个。然后用求得的各年潮汐调和常数分别后报当年潮位。将实测水位减去潮汐后报值得出余水位值。这样, 便把实测水位序列分解为一个潮汐序列和一个对应的余水位序列。

我们之所以用每年分析所得潮汐常数后报当年潮汐, 而不用多年平均潮汐常数后报, 是因为实际上各年分析所得潮汐常数有一定变化, 特别是润洲的变化较大, 这是由于摩擦非线性产生的, 这里不拟详细讨论。用当年分析结果后报有利于更好消除潮汐部分。

2. 余水位统计分布的一般特征

我们先不分潮汐水位高低, 把各站所有余水位放在一起统计。

将余水位按 0.1m 间隔统计出出现在各间隔内的频率 (即上节中 p_k''), 画出频率分布直方图。作为例子, 图 3 示出秦皇岛、汕头和润洲三站。

另外, 我们按下列公式计算了余水位分布的各阶统计数字特征:

平均值

$$m = \frac{1}{N} \sum_{n=1}^N y_n \quad (3.2)$$

标准差

$$\sigma = \mu_2^{1/2} \quad (3.3)$$

偏度

$$s = \mu_3 / \sigma^3 \quad (3.4)$$

峰度

$$k = \mu_4 / \sigma^4 \quad (3.5)$$

表 3 余水位分布的统计数字特征

站名	σ	s	k
葫芦岛	26.75	-1.00	6.38
秦皇岛	22.94	-1.09	7.24
青 岛	16.65	-0.48	5.83
吴 淞	16.90	0.55	6.16
绿华山	13.26	0.14	4.52
厦 门	18.33	0.54	4.35
汕 头	16.33	0.73	5.83
黄 埔	17.35	0.37	5.82
榆 林	11.42	0.48	5.77
润 洲	11.93	-0.58	6.85

其中 μ_r 为第 r 阶中心矩:

$$\mu_r = \frac{1}{N} \sum_{n=1}^N (y_n - m)^r \quad (3.6)$$

上面 y_n 为第 n 个余水位观测值, N 为观测值个数。

要注意的是, 在通常的数理统计中峰度被取作 $k = \mu_4 / \sigma^4 = 3$ 。这是由于正态分布的 μ_4 / σ^4 等于 3, 被取作标准。我们这里的峰度值没有减去 3。因而在这个定义下, 正态分布的 $k = 3$ 。

葫芦岛等 10 个验潮站余水位的统计数字特征列于表 3。由该表可以看到下列规律：

(1) 总的来看,余水位的标准差在我国沿海有自北向南减小的趋势,而河口站的数值要比处于开阔海域的验潮站大。造成这种特点的原因可能有两点:一是我国南方海区较为接近深海区而北方海区则处于陆架海的内部,同时在同样风力作用下浅海区易于造成较大增减水;二是北方受寒潮影响较大,南方受台风影响较大,前者出现频率远高于后者。

(2) 葫芦岛、秦皇岛和青岛等北方港口偏度为负,南方港口则多为正偏。这主要是由于前者受寒潮影响,常出现大的减水,特别这三个港位于陆地南侧,北风所引起的减水特别显著。南方主要受台风影响,出现大的增水的机会比出现大的减水的机会多。涠洲的情况与葫芦岛等地有点类似,寒潮在这里主要是东北风,台风在到达北部湾之前也在北部湾先引起北风,这都能导致北部湾北部的减水。

从图 3 亦可以看出,秦皇岛的余水位频率分布图显得比较矮宽,而涠洲则较高窄。同时秦皇岛和涠洲大的减水出现频率比大的增水要高,而汕头则相反。

(3) 峰度的数值似乎有在渤海、黄海和南海较高,东海较低的趋势。但是无论在哪个海区,峰度均明显高于正态分布的峰度。

余水位的分布与正态分布有明显差异,这一点应特别引起注意,尤其在研究极值分布时。

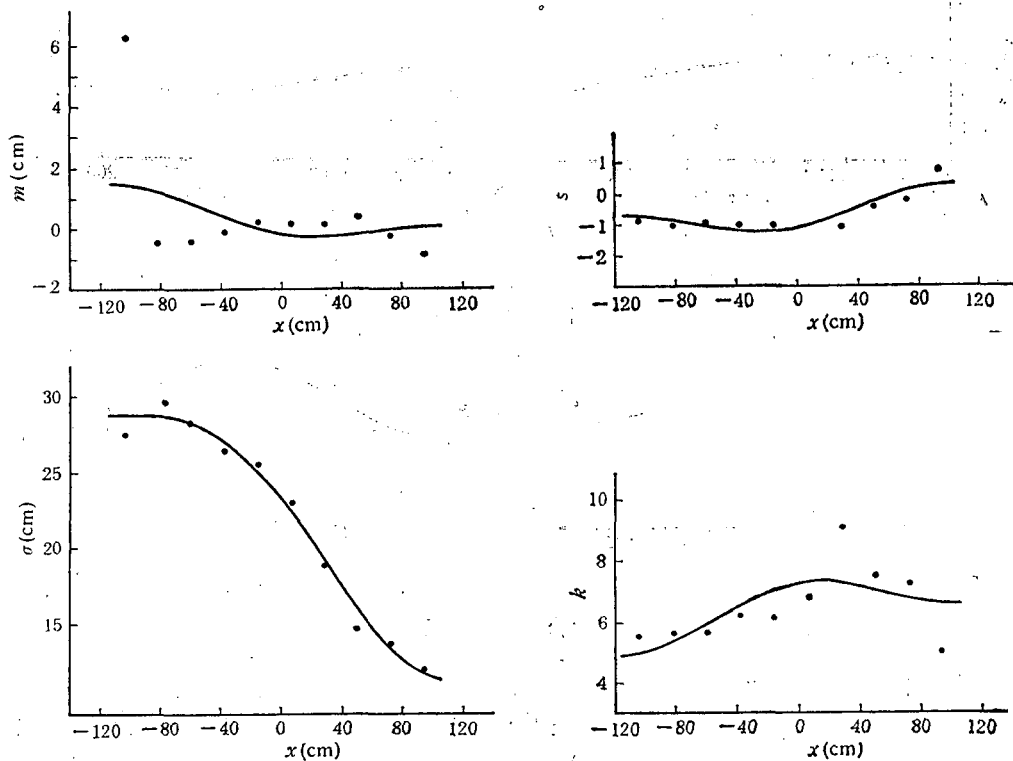


图 4a 秦皇岛余水位分布数字特征与潮汐的关系

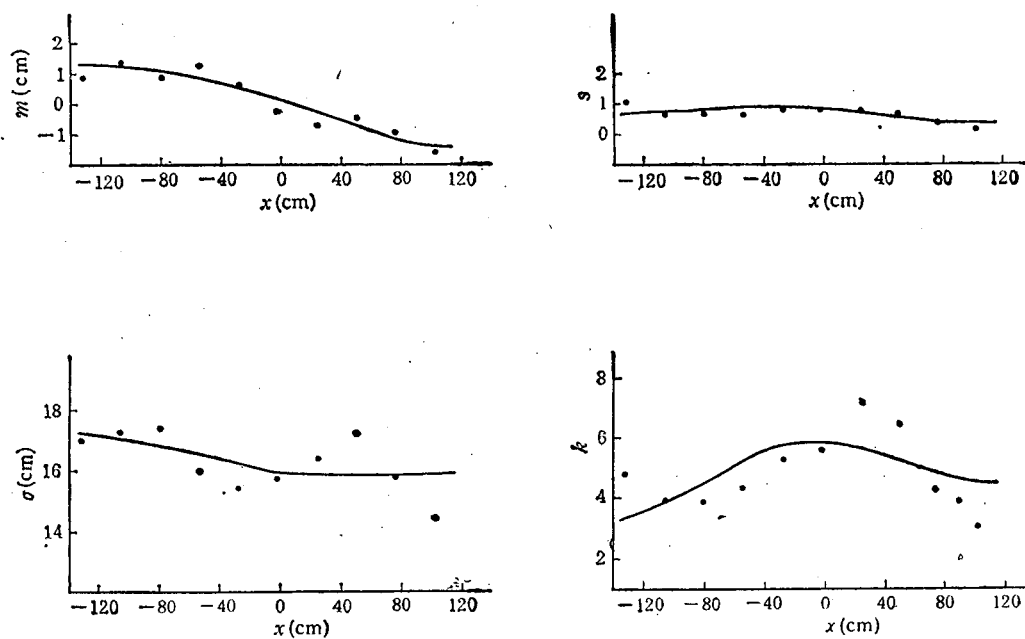


图 4b 汕头余水位分布数字特征与潮汐的关系

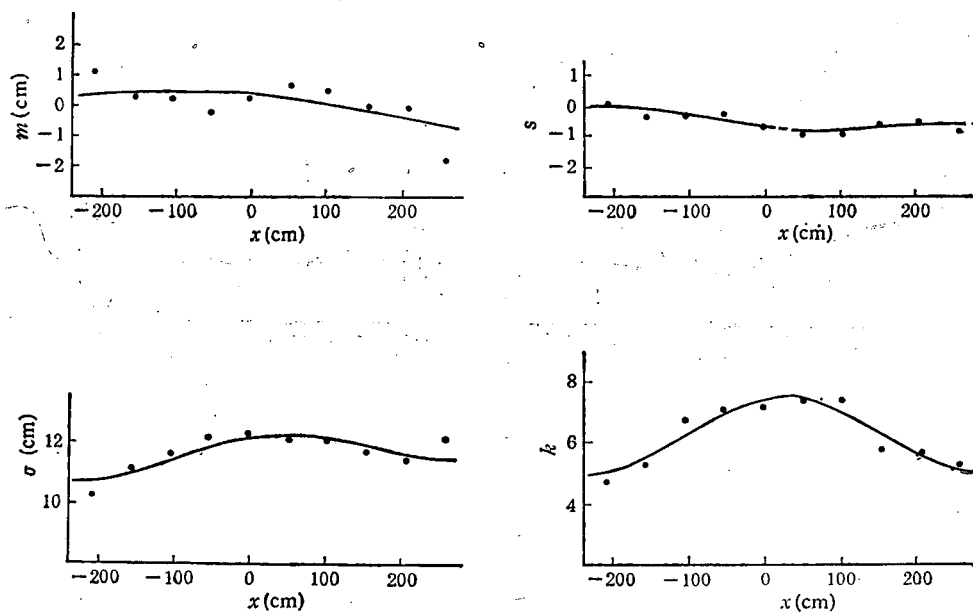


图 4c 涇洲余水位分布数字特征与潮汐的关系

3. 余水位统计分布与潮汐的关系

为了考虑余水位统计分布与潮汐的关系, 我们按各站出现的最低和最高潮汐水位的数值而把潮汐水位分成 10 个区间, 即取 $J = 10$ (J 的定义见上节)。余水位的序列按照每个值所对应的潮汐值也被分成 10 个序列, 对这 10 个序列分别进行统计可得不同潮汐

表 4 不同潮汐条件下余水位分布的数字特征

j	x (cm)	N	m (cm)	σ (cm)	s	k
葫芦岛						
1	-155	4 261	-1.35	29.65	-0.74	4.08
2	-114	29 227	-1.47	29.30	-0.73	4.53
3	-73	49 879	-0.71	25.93	-0.95	6.32
4	-32	41 001	-0.61	26.40	-1.04	6.83
5	10	39 990	-0.31	27.89	-1.05	6.69
6	51	36 744	0.03	26.41	-1.00	6.80
7	92	28 931	0.32	26.88	-1.08	6.55
8	133	19 137	0.95	25.74	-1.14	6.90
9	174	8 191	2.50	20.28	-1.20	9.95
10	215	1 228	3.80	14.36	0.07	3.62
全部		258 589	-0.26	26.75	-1.00	6.38
秦皇岛						
1	-104	386	6.25	27.52	-0.88	5.51
2	-82	5 296	-0.39	29.63	-1.08	5.62
3	-60	16 534	-0.45	28.30	-0.97	5.60
4	-38	34 705	-0.13	26.43	-1.04	6.23
5	-16	55 295	0.20	25.57	-1.08	6.08
6	6	62 097	0.13	22.96	-1.09	6.81
7	28	47 544	0.12	18.89	-1.09	9.03
8	50	25 581	0.37	14.76	-0.42	7.51
9	72	8 858	-0.26	13.56	-0.21	7.25
10	94	1 632	-0.88	11.96	0.78	5.02
全部		257 928	0.07	22.94	-1.09	7.24
青 岛						
1	-240	1 563	3.85	16.96	-0.26	4.43
2	-190	12 286	-0.77	16.69	-0.46	5.28
3	-140	31 913	-0.90	16.20	-0.39	5.43
4	-90	43 774	-0.70	16.26	-0.41	5.74
5	-40	44 099	-0.11	16.66	-0.46	5.82
6	10	42 289	0.58	17.19	-0.48	5.73
7	60	47 782	0.68	17.57	-0.58	5.77
8	110	46 449	0.28	16.70	-0.59	6.08
9	160	24 172	-0.15	15.19	-0.41	6.55
10	210	3 604	0.45	13.41	-0.13	6.68
全部		297 931	0.00	16.65	-0.48	5.83

续表 4

j	x (cm)	N	m (cm)	σ (cm)	s	k
吴 淞						
1	-173	3 140	13.51	16.90	0.20	3.68
2	-128	28 623	-0.64	16.37	0.29	5.01
3	-83	55 540	-1.89	17.01	0.43	5.89
4	-38	50 557	0.80	17.33	0.70	6.77
5	8	43 749	1.23	17.10	0.65	6.94
6	53	41 493	0.19	16.61	0.56	6.18
7	98	35 059	-0.69	16.29	0.58	6.04
8	143	21 675	-0.26	16.37	0.56	5.47
9	188	8 393	0.91	16.43	0.67	5.91
10	233	1 028	2.17	17.15	1.78	14.06
全部		289 257	0.00	16.90	0.55	6.16
绿华山						
1	-221	1 908	0.77	15.32	0.47	3.95
2	-172	12 298	-0.68	14.58	0.31	4.11
3	-123	26 542	-1.44	14.19	0.31	4.54
4	-74	35 960	-0.99	13.66	0.29	4.47
5	-25	34 478	-0.22	13.15	0.18	4.10
6	25	34 503	0.25	12.79	0.13	4.21
7	74	35 465	0.70	12.62	0.03	4.58
8	123	26 221	1.54	12.35	0.04	4.98
9	172	13 244	2.44	12.70	-0.28	6.70
10	221	1 952	3.64	12.98	-0.16	6.57
全部		222 571	0.11	13.26	0.14	4.52
厦 门						
1	-292	3 634	2.53	14.17	0.70	4.21
2	-226	19 176	1.60	15.37	0.83	4.82
3	-160	39 026	1.85	16.99	0.98	5.47
4	-94	41 772	0.31	18.31	0.87	4.66
5	-28	33 318	-1.10	19.20	0.74	4.32
6	38	30 914	-1.26	19.85	0.53	3.99
7	104	35 645	0.15	19.56	0.28	3.82
8	170	36 983	-0.09	18.39	0.14	3.76
9	236	19 862	-1.91	17.43	0.11	3.86
10	302	2 589	-3.79	16.53	0.10	3.40
全部		262 919	0.01	18.33	0.54	4.35

续表 4

j	x (cm)	N	m (cm)	σ (cm)	s	k
汕 头						
1	-132	489	0.94	17.01	1.08	4.87
2	-106	3 845	1.40	17.28	0.67	3.94
3	-80	10 886	0.96	17.42	0.65	3.89
4	-54	16 160	1.30	16.01	0.67	4.39
5	-28	21 513	0.69	15.44	0.83	5.28
6	-2	33 331	-0.21	15.75	0.77	5.61
7	24	45 469	-0.65	16.42	0.78	7.16
8	50	25 918	-0.48	17.21	0.69	6.39
9	76	6 880	-0.97	15.81	0.41	4.32
10	102	639	-1.56	14.44	0.17	3.09
全部		165 130	-0.03	16.33	0.73	5.83
黄 埔						
1	-139	2 366	4.95	18.34	0.27	3.46
2	-106	21 826	-1.86	19.63	0.51	4.68
3	-73	33 919	-1.08	18.25	0.33	4.90
4	-40	42 684	0.24	17.53	0.55	8.50
5	-7	47 384	0.84	16.77	0.37	6.06
6	27	49 298	0.85	16.57	0.33	4.98
7	60	38 538	0.41	16.49	0.28	5.09
8	93	23 495	-0.92	16.77	0.20	7.08
9	126	9 823	-1.82	17.68	0.50	4.34
10	159	1 402	-4.22	19.30	0.49	3.58
全部		270 735	-0.01	17.35	0.37	5.82
榆 林						
1	-104	280	-0.81	11.20	-0.84	5.61
2	-81	4 573	0.49	10.08	0.01	4.84
3	-58	19 003	-0.69	10.96	0.21	5.46
4	-35	43 127	-0.16	11.46	0.46	5.97
5	-12	66 365	-0.15	11.36	0.38	5.61
6	12	52 421	0.13	11.70	0.59	5.70
7	35	37 002	0.42	11.42	0.44	5.53
8	58	22 948	0.40	11.15	0.57	5.98
9	81	7 145	0.13	12.08	1.08	7.31
10	104	934	0.50	12.56	1.07	5.36
全部		253 798	0.02	11.42	0.48	5.77

续表 4

j	x (cm)	N	m (cm)	σ (cm)	s	k
涠洲						
1	-209	1 186	1.12	10.30	0.00	4.72
2	-157	14 560	0.33	11.12	-0.38	5.30
3	-105	30 521	0.21	11.69	-0.29	6.76
4	-53	43 388	-0.24	12.08	-0.33	7.03
5	-1	40 275	0.24	12.28	-0.71	7.14
6	51	27 000	0.65	12.07	-0.92	7.34
7	103	22 405	0.46	12.06	-0.92	7.36
8	155	18 613	-0.15	11.70	-0.61	5.79
9	207	6 889	-0.13	11.48	-0.52	5.68
10	259	559	-1.83	12.06	-0.82	5.28
全部		205 396	0.17	11.93	-0.58	6.85

条件下的余水位的分布,同时可得出不同潮汐条件下余水位的统计特征量。这些特征量的计算公式同式(3.2)~(3.6),不过这时它们与潮汐有关,我们将用一个下标 $j = 1, 2, \dots, J$ 来区别它们。对葫芦岛等 10 个站,不同潮汐条件下特征量的计算结果见表 4。作为例子,我们取秦皇岛、汕头和涠洲三站作图,如图 4 所示,图中圆点为实测资料统计结果,曲线为拟合结果,拟合方法见下节。

由表 4 可知,作为余水位大小的最主要特征量——标准差在不同潮汐条件下可能有所不同。特别对北方沿岸,这一差异十分显著。其中最突出的是秦皇岛,低潮位情况下的余水位标准差比高潮位情况下大一倍以上。余水位分布的偏度也可能随着潮位的不同而不同,但一般来说,对同一个站基本情况差别尚不太大。例如北方港口一般都是负偏,南方港口(涠洲除外)一般都是正偏。当然也有个别例外,这种情况多出现在参加统计的数据个数极少(即 $j = 1$ 或 10 附近)的时候。峰度的变化性更大一些,但对同一个站,其值亦较接近。

四、极值分布的计算方法

在第二节中曾指出,直接利用不同潮汐条件下统计所得的余水位分布来计算极值分布会丢失许多有用的信息。但是如果不区分潮汐条件则有可能得出不符合实际的结果。为了解决这一矛盾,必须允许余水位的分布随潮汐而变化,但应当根据一些基本观测事实对其变化性予以一定限制。

一种理想的方式是设法给出余水位分布的一定模型。最好是能够用函数的形式表示,然后调整其参数使得它符合某一地点在某一潮汐条件下的余水位分布。方国洪和王骥(1987)提出用 Edgeworth 级数表示余水位分布就是基于这一想法。Edgeworth 级

数系以正态分布为基础,加上由 Hermite 多项式与偏度及峰度的乘积表示的订正项而构成。后来经过一系列试验未能取得理想的效果。其主要原因是实际的余水位分布与正态分布相差太大,因而订正项不能迅速收敛。后来我们又提出一种所谓复合 Pearson-III 型函数来拟合余水位分布。同时还试验了一种更加依赖经验统计的方法,称之为相似法。这里对这两种方法都加以介绍,尽管本文根据目前试验结果建议采用相似法。

在这两种方法中都用到余水位条件分布的数字特征,故先把数字特征的处理过程作一介绍。

1. 余水位条件分布数字特征的函数拟合

由实际观测资料统计得出的不同潮汐条件下余水位分布的数字特征既呈现一定的规律性,同时也包含一定的偶然性。特别当潮汐水位 X 在极端值附近时, $f'(x)$ 很小,可以用来统计余水位 Y 的条件数字特征的观测次数也很少,所得的估计量也常常不够稳定(参看表 4 和图 4)。为了消除这些偶然性,可以采用某种函数形式来拟合实测统计值。采用函数形式的另一个好处是,可以很容易求出在任意潮汐条件下余水位分布的数字特征值,它可以起到内插和外推的作用。

我们用 $q^{(p)}$, $p = 1, 2, 3, 4, \dots$ 来代表 $m, \sigma, s, k, \dots$ 。如前面已指出,潮汐水位只出现在有限区间 $(x_{\min}, x_{\max})$ 。通过前节叙述的统计过程,我们已经得到对应于 $x = x_1, x_2, \dots, x_J$ 时,由实测统计得出的 $q^{(p)} = q_1^{(p)}, q_2^{(p)}, \dots, q_J^{(p)}$ 。

对于任意的函数 $q(x)$ 可以将其表示成一组正交函数的和,例如可以把它表示为 Hermite 多项式之和。在目前情况下采用 Hermite 多项式有其特殊优越之处,因为该多项式是以误差函数为权函数而正交的,而潮汐的分布与正态分布有一定的类似性,即中等数值多而两端数值少。但是这种函数可以允许在区间的两端具有较大的变化率,故实测统计结果在两端的较大的偶然因素会保留下来,而且通过外推(因 x_1 和 $x_{\min}$, x_J 和 $x_{\max}$ 还具有 $\Delta x/2$ 的距离)还会将这些不稳定性进一步扩大。为了避免这一缺点,我们宁愿采用一种较简单的余弦级数。

为此我们令

$$x' = \frac{\pi(x - x_{\min})}{x_{\max} - x_{\min}} \quad (4.1)$$

则 $q^{(p)}$ 定义在区间 $(0, \pi)$ 内。今用函数

$$\hat{q}^{(p)}(x') = a_0^{(p)} + \sum_{m=1}^M a_m^{(p)} \cos mx' \quad (4.2)$$

来拟合 $q_j^{(p)}$, 这里要求 $M+1 \leq J$ 。上式我们只引入余弦函数而不含正弦函数,其目的是使得在 $x' = 0, \pi$ 时有 $d\hat{q}(x')/dx' = 0$, 即要求在 x 分布区域的两端数字特征变化很小。由实测值得出矛盾方程组:

$$a_0^{(p)} + \sum_{m=1}^M a_m^{(p)} \cos mx'_j = q_j^{(p)} \quad (j = 1, 2, \dots, J) \quad (4.3)$$

由于在 X 分布区域两端实测余水位数据个数明显偏少,在确定系数 a_m 时我们对方程 (4.3) 加权再求最小二乘解,即要求

$$\sum_{j=1}^J \left\{ W_j \left[a_0^{(p)} + \sum_{m=1}^M a_m^{(p)} \cos mx_j' - q_j^{(p)} \right] \right\}^2 = \min \quad (4.4)$$

其中权函数 W_j 应当与参与统计的观测个数 N_j 有关, N_j 越小, 权系数 W_j 亦应越小。但如果直接取 W_j 比例于 N_j , 则 $j=1$ 和 J 附近的实测数字特征所起的作用就太小。为此我们取

$$W_j = \arctg(N_j \times 0.005) \quad (4.5)$$

为直观起见, 我们把这个函数关系列出如下:

$N = 0$	10	100	1000	10000	∞
$W = 0$	0.05	0.46	1.37	1.55	1.57

由(4.4)式得出的法方程如下:

$$\begin{bmatrix} C_{00} & C_{01} & \cdots & C_{0M} \\ C_{10} & C_{11} & \cdots & C_{1M} \\ C_{20} & C_{21} & \cdots & C_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ C_{M0} & C_{M1} & \cdots & C_{MM} \end{bmatrix} \begin{bmatrix} a_0^{(p)} \\ a_1^{(p)} \\ a_2^{(p)} \\ \cdots \\ a_M^{(p)} \end{bmatrix} = \begin{bmatrix} F_0 \\ F_1 \\ F_2 \\ \cdots \\ F_M \end{bmatrix} \quad (4.6)$$

其中

$$C_{mn} = \sum W_j^2 \cos mx_j' \cos nx_j'$$

$$F_m = \sum W_j^2 q_j^{(p)} \cos mx_j'$$

式中 $\sum$ 代表 j 从 1 到 J 求和。由法方程(4.6)可以求得系数 $a_0^{(p)}, a_1^{(p)}, \dots, a_M^{(p)}$ 。再代入(4.2)式和(4.1)式可求出任意 x 值对应的 $q^{(p)}$ 值。由于高阶项的可靠性实际上是很差的, 我们只取 $M=2$ 。图 4 中的曲线即是拟合的结果。

2. 复合 Pearson-III 型分布

表 4 表明, 余水位的分布与正态分布有明显差异。首先它可以出现正偏和负偏, 当它是负偏时, 出现大的负值的可能性比出现大的正值可能性大; 反之亦然。其次它的峰度比正态的峰度明显地大, 亦即在相同标准差的情况下, 余水位出现大绝对值的可能性比正态分布期望的可能性大。这表明, 像正态分布那样, 其密度函数随着 y 的增大以 $\exp(-y^2/2)$ 的速度减小是太快了, 不适合余水位分布。能否找到一种函数, 它既可以是偏态的, 同时其随 y 而减小的速度又明显低于 $\exp(-y^2/2)$? Pearson-III 型(以后简称 P-III)函数能满足这种条件, 它是偏态的, 且差不多随 y 作 $\exp(-y)$ 的速度减小。但是 P-III 函数的定义域一端为有限, 这又不符合余水位的分布特征。为了克服这一点我们取正偏和负偏的 P-III 函数进行叠加, 称为复合 P-III 函数。

P-III 函数可写成(金光炎, 1964)

$$f(y) = \begin{cases} A(y-a)^{\alpha-1} \exp[-\beta(y-a)], & \text{对 } y > a \\ 0, & \text{对 } y \leq a \end{cases} \quad (4.7)$$

作为分布密度函数, 必须保证

$$\int_{-\infty}^{\infty} f(y) dy = 1$$

为此, A 值取

$$A = \beta^\alpha / \Gamma(\alpha) \quad (4.8)$$

其中 Γ 代表 Γ 函数。

由式(4.7)知服从 P -III 型分布随机变量的均值为

$$\bar{y} = \int_a^\infty y f(y) dy = a + \alpha/\beta \quad (4.9)$$

故中心化后, 随机变量 $Y' = Y - \bar{y}$ 的密度函数为

$$f_1(y') = \begin{cases} A(y' + \alpha/\beta)^{\alpha-1} \exp[-\beta(y' + \alpha/\beta)], & \text{对 } y' > -\alpha/\beta \\ 0, & \text{对 } y' \leq -\alpha/\beta \end{cases} \quad (4.10)$$

由上式得

$$\frac{df_1(y')}{dy'} = \frac{(y' + d)f_1(y')}{b'_0 + b'_1 y'} \quad (4.11)$$

其中

$$\begin{cases} b'_0 = -\alpha/\beta^2 \\ b'_1 = -1/\beta \\ d = 1/\beta \end{cases} \quad (4.12)$$

根据金光炎[1964, 式(6-4)], 各阶中心矩有下列递推关系:

$$\begin{aligned} \mu_{n+1} &= -nb'_0\mu_{n-1} - [(n+1)b'_1 + d]\mu_n \\ &= (n\alpha/\beta^2)\mu_{n-1} + (n/\beta)\mu_n \end{aligned}$$

或

$$\mu_{n+1} = (n/\beta)[(\alpha/\beta)\mu_{n-1} + \mu_n] \quad (4.13)$$

因 $\mu_0 = 1$, $\mu_1 = 0$, 由上式可推得各阶中心矩如下:

$$\begin{cases} \mu_2 = \alpha/\beta^2 \\ \mu_3 = 2\alpha/\beta^3 \\ \mu_4 = 3\alpha(\alpha+2)/\beta^4 \\ \mu_5 = 4\alpha(5\alpha+6)/\beta^5 \\ \dots \end{cases} \quad (4.14)$$

如 $\alpha > 0$, $\beta > 0$, 则式(4.10)具有正偏态性质。

将式(4.10)中 y' 换成 $-y'$, 则可得纵坐标轴为对称轴的与 $f_1(y')$ 相对称的负偏 P -III 函数:

$$f_2(y') = \begin{cases} A(\alpha/\beta - y')^{\alpha-1} \exp(\beta y' - \alpha), & \text{对 } y' < \alpha/\beta \\ 0, & \text{对 } y' \geq \alpha/\beta \end{cases} \quad (4.15)$$

它的各阶中心矩的递推关系为

$$\mu_{n+1} = (n/\beta)[(\alpha/\beta)\mu_{n-1} - \mu_n] \quad (4.16)$$

以 $\mu_0 = 1$, $\mu_1 = 0$ 代入, 有

$$\begin{cases} \mu_2 = \alpha/\beta^2 \\ \mu_3 = -2\alpha/\beta^3 \\ \mu_4 = 3\alpha(\alpha+2)/\beta^4 \\ \mu_5 = -4\alpha(5\alpha+6)/\beta^5 \\ \dots\dots \end{cases} \quad (4.17)$$

根据上述,复合 P -III 函数可取

$$f(y) = \begin{cases} A_1(-y+a_1)^{\alpha_1-1}e^{\beta_1(y-a_1)}, & \text{对 } y < a_2 \\ A_1(-y+a_1)^{\alpha_1-1}e^{\beta_1(y-a_1)} \\ + A_2(y-a_2)^{\alpha_2-1}e^{-\beta_2(y-a_2)}, & \text{对 } a_2 \leq y < a_1 \\ A_2(y-a_2)^{\alpha_2-1}e^{-\beta_2(y-a_2)}, & \text{对 } y \geq a_1 \end{cases} \quad (4.18)$$

这里假定 $a_1 > a_2$. 为保证

$$\int_{-\infty}^{\infty} f(y) dy = 1$$

应有

$$A_1\Gamma(\alpha_1)/\beta_1^{\alpha_1} + A_2\Gamma(\alpha_2)/\beta_2^{\alpha_2} = 1 \quad (4.19)$$

式(4.18)共有 7 个可供选择的独立参量。在实际工作中可能会由于参数太多,稳定性较差。为了减少参量,我们首先给出一个限制,即要求正偏和负偏的 P -III 函数都是中心化了的。此时,

$$\begin{cases} a_1 = \alpha_1/\beta_1 \\ a_2 = -\alpha_2/\beta_2 \end{cases} \quad (4.20)$$

在满足(4.19)式和(4.20)式条件下,(4.18)式仍包含 5 个独立参量。

式(4.18)也可写作

$$f(y) = f_1(y) + f_2(y) \quad (4.21)$$

其中

$$f_1(y) = \begin{cases} A_1(\alpha_1/\beta_1 - y)^{\alpha_1-1}e^{-\alpha_1+\beta_1 y}, & \text{对 } y < \alpha_1/\beta_1 \\ 0, & \text{对 } y \geq \alpha_1/\beta_1 \end{cases} \quad (4.22)$$

$$f_2(y) = \begin{cases} 0, & \text{对 } y < -\alpha_2/\beta_2 \\ A_2(y + \alpha_2/\beta_2)^{\alpha_2-1}e^{-\alpha_2-\beta_2 y}, & \text{对 } y \geq -\alpha_2/\beta_2 \end{cases} \quad (4.23)$$

本文将采用矩方法确定复合 P -III 分布中的各参量。记

$$\mu_n^{(i)} = \int_{-\infty}^{\infty} y^n f_i(y) dy \quad (i=1,2) \quad (4.24)$$

则 $f(y)$ 的 n 阶中心矩 μ_n 为

$$\mu_n = \int_{-\infty}^{\infty} y^n f(y) dy = \mu_n^{(1)} + \mu_n^{(2)} \quad (4.25)$$

进一步对函数(4.18)作限制,取

$$\begin{cases} A_1 = \frac{1}{2} \beta_1^{\alpha_1} / \Gamma(\alpha_1) \\ A_2 = \frac{1}{2} \beta_2^{\alpha_2} / \Gamma(\alpha_2) \end{cases} \quad (4.26)$$

则有

$$\begin{cases} \mu_0 = 1 \\ \mu_1 = 0 \\ \mu_2 = \frac{1}{2} (\alpha_1/\beta_1^2 + \alpha_2/\beta_2^2) \\ \mu_3 = -\alpha_1/\beta_1^3 + \alpha_2/\beta_2^3 \\ \mu_4 = \frac{1}{2} [(3\alpha_1^2 + 6\alpha_1)/\beta_1^4 + (3\alpha_2^2 + 6\alpha_2)/\beta_2^4] \\ \mu_5 = 2[-\alpha_1(5\alpha_1 + 6)/\beta_1^5 + \alpha_2(5\alpha_2 + 6)/\beta_2^5] \end{cases} \quad (4.27)$$

假定随机变量已标准化,则

$$\alpha_1/\beta_1^2 + \alpha_2/\beta_2^2 = 2 \quad (4.28)$$

引入新参变量 λ ,

$$\lambda = (\alpha_2/\beta_2^2)/(\alpha_1/\beta_1^2) \quad (4.29)$$

它代表了 $f_1(y)$ 和 $f_2(y)$ 在构成方差中所占的相对比重,并由以后与实际分布的比较来确定。

由式(4.28)和(4.29)得

$$\begin{cases} \alpha_1 = 2\beta_1^2/(1+\lambda) \\ \alpha_2 = 2\lambda\beta_2^2/(1+\lambda) \end{cases} \quad (4.30)$$

代入式(4.27)后两式,得

$$\mu_3 = 2(-d_1 + \lambda d_2)/(1+\lambda) \quad (4.31)$$

$$\mu_4 = 6 \left[\frac{1+\lambda^2}{(1+\lambda)^2} + \frac{1}{1+\lambda} d_1^2 + \frac{\lambda}{1+\lambda} d_2^2 \right] \quad (4.32)$$

其中

$$d_i = 1/\beta_i, (i=1,2) \quad (4.33)$$

由式(4.31)和(4.32)解得

$$\begin{cases} d_1 = (\lambda c - \mu_3)/2 \\ d_2 = (c + \mu_3)/2 \end{cases} \quad (4.35)$$

其中

$$c = \left\{ \frac{4}{\lambda} \left[\frac{\mu_4}{6} - \frac{\mu_3^2}{4} - \frac{1+\lambda^2}{(1+\lambda)^2} \right] \right\}^{\frac{1}{2}} \quad (4.36)$$

注意这里 μ_3 和 μ_4 是三阶和四阶标准中心矩。

式(4.36)中 $(1+\lambda^2)/(1+\lambda)^2 \geq 1/2$ 。因此不等式

$$\frac{\mu_4}{3} \geq 1 + \frac{\mu_3^2}{2} \quad (4.37)$$

是存在实数 d 的必要条件。特别当 $\mu_3 = 0$ 时式(4.37)要求 $\mu_4 \geq 3$ 。又根据式(4.33)–(4.35),可知必须有 $\mu_4 > 3$ 。这表明,欲以上述复合 P-III 分布拟合余水位分布,则后者的峰度必须大于正态分布的峰度。这当然可以满足,我们之所以采用复合 P-III 分布,就是因为余水位分布的峰度显著大于正态分布。在 $\mu_3 \neq 0$ 时,峰度还要更大一些。我们考察表 4 中列出的实际余水位的数字特征,可以发现均满足式(4.37)。

在式(4.37)满足的条件下我们可令

$$\frac{\mu_4}{3} - \frac{\mu_3^2}{2} - 1 = \varepsilon \quad (4.38)$$

若 $\varepsilon \geq 1$, 则 λ 可取任意正值; 若 $1 > \varepsilon > 0$, 则 λ 取值范围为

$$\frac{(1 - \varepsilon^{\frac{1}{2}})^2}{1 - \varepsilon} < \lambda < \frac{(1 + \varepsilon^{\frac{1}{2}})^2}{1 - \varepsilon} \quad (4.39)$$

将满足式(4.39)的 λ 值(若 $\varepsilon > 1$, 则任意正值)代入式(4.36), (4.35), (4.33) 和 (4.30)可得 α_i, β_i 。但只有进一步满足

$$\alpha_i > 1, \beta_i > 0 \quad (i = 1, 2) \quad (4.40)$$

的那些解才有意义。对于所有这些解, 式(4.27)中前 5 个式子都自动满足。在所有的解中我们选择由式(4.27)第 6 式算得的 μ_i 值与实测值最接近的那一组。

上面叙述的是如何由三至五阶标准中心矩确定标准化复合 P-III 密度函数 $f(y)$ 。又由式(4.1)和(4.2)可得出潮汐 X 等于任一数值 x 时的余水位各阶矩, 因而可得此时标准化的余水位条件密度函数 $f(y|x)$ 。计算总水位的分布可以采用第二节叙述的过程。首先用潮汐调和常数推算 19 年逐时潮汐水位, 以每 10cm 为间隔统计各潮汐水位 x_i 的出现频率 p_{i0} 。 $x = x_i$ 时余水位平均值为 m_i , 标准差为 σ_i , 则余水位出现在 $y_k \pm \Delta y/2$ (实际计算中取 $\Delta y = 10\text{cm}$)范围内的频率为

$$\begin{aligned} p_{k|i} &= \int_{(y_k - \Delta y/2 - m_i)/\sigma_i}^{(y_k + \Delta y/2 - m_i)/\sigma_i} f(y|x_i) dy \\ &\approx f\left(\frac{y_k - m_i}{\sigma_i} \middle| x_i\right) \Delta y / \sigma_i \end{aligned} \quad (4.41)$$

最后由式(2.11)–(2.13)计算不同水位高度所对应的重现期。通过内插可以反过来得出不同重现期所对应的水位高度。由于我们前面用的是逐时观测值, 因而由式(2.13)得出的重现期的单位也是小时, 实际应用中则化作以年为单位。

3. 相似法

相似法假定对同一地点、不同潮汐条件下的余水位分布可以有差别, 但差别只表现在均值和标准差的不同, 其标准化的分布则相同。在这一假定下, 余水位分布的偏度和峰度以及更高阶标准中心矩都将不会随潮汐而改变。符合上述条件的分布我们称之为相似。这当然只是一种近似, 但是从表 4 可以看出, 实际余水位的条件数字特征大体上符合这个条件。此外, 我们以后在求极值分布时还要对标准差作某些调整, 以考虑偏度和峰度随潮汐的变化所带来的影响。

我们进一步假定, 对同一个站, 余水位所有条件分布均与边际分布相似。亦即,

$$f(y_i|x_i) = \frac{\sigma}{\sigma_i} f''(y'') \quad (4.42)$$

其中

$$y'' = \frac{\sigma}{\sigma_i} (y - m_i) + m \quad (4.43)$$

这样, 就可以用边际分布来求出条件分布。不过, 在实际工作中我们没有采用函数形式来

表示密度函数,而是直接采用数值形式的统计结果。具体作法如下:

(1) 不分潮汐情况,对余水位进行统计,求得经验分布。在实际统计中,我们采用水位间隔为 $\Delta y = 1\text{cm}$ 。对每个余水位值 y_k'' ,可得到对应的出现频率 p_k'' 。同时算出边际分布平均值 m 和标准差 σ 。

(2) 用潮汐调和常数推算 19 年逐时潮汐水位,确定其取值范围 $(x_{\min}, x_{\max})$,并以 $\Delta x = 10\text{cm}$ 为间隔统计各潮汐水位 x_j 的出现频率 p_{j0} 。

(3) 按本节第 1 小节的方法计算出 $a_0^{(p)}$, $a_1^{(p)}$ 和 $a_2^{(p)}$, $p = 1, 2, 3, 4$ 。计算对应各潮汐水位等于 x_j 时的余水位平均值和标准差。其中平均值按式(4.2)计算;标准差计算公式则作一些订正,以考虑偏度和峰度的影响。即

$$\begin{aligned} \sigma_j = a_0^{(2)} [1 - \alpha(a_2^{(3)} \cos x_j' + a_1^{(3)} \cos 2x_j') \\ + \beta(a_1^{(4)} \cos x_j' + a_2^{(4)} \cos 2x_j')] \\ + a_1^{(2)} \cos x_j' + a_2^{(2)} \cos 2x_j' \end{aligned} \quad (4.44)$$

式中方括号内的订正是基于如下考虑作出的: 正的 $a_2^{(3)}$ 表明低潮位和高潮位时余水位的偏度比平均值高,而正的偏度能引起极端低水位和极端高水位的增高,而这种效果与低潮位时余水位标准差减少、高潮位时余水位标准差增大的效果一样,故可引入与 $-a_2^{(3)} \cos x_j'$ 成比例的项对标准差进行校正,以考虑偏度变化带来的影响。其余各项的理由相同。式中校正系数 α 和 β 是经验性的。本工作中取 $\alpha = \beta = 0.05$ 。

(4) 对各个 x_j , 算出 m_j 和 σ_j 后,以 $\Delta y = 10\text{cm}$ 为间隔,计算对应于各 y_k 值的条件频率 $p_{k|j}$:

$$p_{k|j} = \sum_{k'} p_{k'}'' \quad (4.45)$$

式中右端系对余水位边际分布的出现频率求和, k' 满足

$$\frac{\sigma}{\sigma_j} (y_k - \Delta y/2 - m_j) + m \leq y_k'' < \frac{\sigma}{\sigma_j} (y_k + \Delta y/2 - m_j) + m \quad (4.46)$$

(5) 按式(2.10)~(2.13)计算水位的重现期。

五、结果讨论

表 5 列出了葫芦岛等 10 个验潮站采用 Gumbel 极值分布拟合法,独立联合概率法和本文两种方法所得 50 年一遇高、低水位的结果。按《规范》规定,50 年一遇的水位为校核水位。

《规范》还推荐了一种“同步差比法”。如果某拟建港有 5 年以上,但不到 20 年的实测水位,则可选定附近一个已知校核水位的港口作主港,假定拟建港和主港的校核水位之比(以平均海面为起算)等于两港某一期实测年最高(低)水位平均值之比,从而可推算出拟建港的校核水位。为了与这个方法也作一比较,我们把由该法计算所得的校核水位亦列出于表 5 中。用同步差比法推算表中各港的校核水位时,分别由地理位置最接近、具有 20 年以上验潮记录的港口作主港: 葫芦岛、秦皇岛、青岛、吴淞、绿华山、厦门、汕头、黄埔、榆林和涠洲各港校核水位分别由秦皇岛、塘沽、乳山口、金山嘴、长涂、东山、北津、东方

表 5 由长期资料用不同方法所得 50 年一遇水位值

港 口		方 法				
		Gumbel 法	独立联合 概率法	复合 P-III 法	相似法	同步差比法
高水位	葫芦岛	283	316*	299	293	342*
	秦皇岛	172	213*	185	186	159
	青 岛	297	288	291	283	289
	吴 淞	344	337	345	348	338
	绿华山	284	281	296	286	316*
	厦 门	385	391	368	377	391
	汕 头	296	311	196*	313	185*
	黄 埔	246	311*	250	308*	303*
	榆 林	170	168	182	174	159
低水位	涠 洲	303	309	299	305	331*
	葫芦岛	-303	-328*	-356*	-319	-367*
	秦皇岛	-286	-257*	-331*	-292	-240*
	青 岛	-319	-327	-323	-320	-398*
	吴 淞	-227	-245	-240	-236	-223
	绿华山	-280	-290	-289	-290	-282
	厦 门	-354	-370	-354	-358	-383*
	汕 头	-180	-178	-180	-177	-171
	黄 埔	-199	-261*	-215	-271*	-192
	榆 林	-143	-142	-152	-139	-139
	涠 洲	-262	-281	-266	-269	-241*
与 Gumbel 法比较的均方差**			26(18)	29(30)	23(9)	43(43)
与 Gumbel 法比较的绝均差**			19(14)	17(18)	13(8)	30(30)

注：水位值为相对于平均海面，单位：cm。

* 与 Gumbel 法结果相差 20cm 以上。

** 括号外数值为取所有站进行统计的结果，括号内为不计黄埔站的结果。

和北海推算。由于需要同期资料，年数比图 1 中少一些，但一般都在 15 年以上。

评论各种方法所得结果的合理性时，一般都希望有一个客观的标准。对于多年一遇水位，由于工程应用的习惯，作比较时常常把由极值分布所得结果作为标准。如果观测系列足够长，这样做是比较合适的。目前的情况是，我们所研究的极值水位对应的重现期是 50 年，而各站实际观测长度都是二三十年，由极值分布得出的极值水位值均是外推的结果，因此不能保证每个站的估计值都合理。但另一方面，本文研究了 10 个站，共有 286 年资料，从总体上讲，由极值分布所得的估计值应当说是客观合理的，可以作为比较的标准。

表 5 列出其它四种方法结果与 Gumbel 法结果比较的标准差和绝均差。每种有两个数据，一个是把所有 10 个站都统计在内，另一个是除黄埔外的 9 个站的统计值。关于黄埔港的数据将在稍后讨论。从所列数据可看出无论考虑或不考虑黄埔，均以相似法与 Gumbel 法最接近。因此我们认为，相似法所得结果优于独立联合概率法和复合 P-III 法，更明显优于同步差比法。

如前所述,独立概率法的优点是能够充分利用余水位的信息,缺点是忽视不同潮汐条件下的余水位分布的差异,在余水位分布与潮汐关系不明显的海区能够取得良好效果,但是在余水位分布与潮汐关系密切的海区则效果不好。复合 P-III 法能够充分考虑不同潮汐条件下余水位分布的差异,但对余水位信息未能充分利用。例如用该法计算时,极高水位差不多完全决定于潮汐高水位本身以及在潮汐高水位条件下余水位的分布,那些在中等或低潮水位下发生的风暴潮数据基本不影响高水位计算结果。相似法看来是一种适当的折中,因为该法中的基本分布是由全部余水位资料得出的,而不同潮汐条件下余水位分布的差异又能通过余水位标准差的变化而予以考虑。同步差比法效果最差,只有在拟建港和主港足够接近时才能采用。

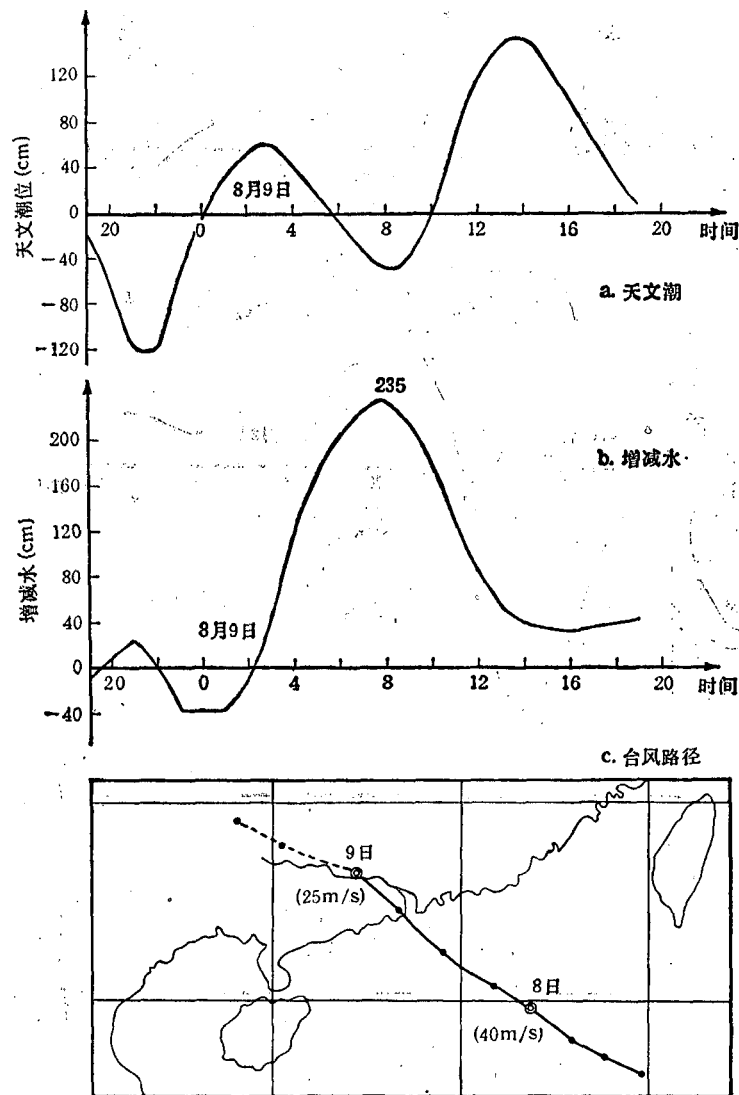


图 5 6411 号台风路径及其在黄埔港引起的风暴潮
 (c 图中日期表示该日 08 时台风中心位置,括号内数字为最大风速,图 6 同)

如前所述, Gumbel 法从总体上看应当是可信的。但并不是对每个站都能够提供良好的结果。黄埔港可以作为一个例证。按 Gumbel 法计算得出的 50 年一遇高水位为 246cm, 低水位为 -199cm。而用相似法得出的结果分别是 308cm 和 -271cm。我们检查了余水位的情况, 发现在黄埔风暴潮并不小。1952—1984 年的资料中, 最大增水为 235cm, 最大减水为 212cm。就是说用 Gumbel 法得出的 50 年一遇高水位仅比 31 年内观测到的最大增水高 11cm, 50 年一遇低水位值反而比 31 年内观测到的最大减水还高 13cm。而黄埔最高和最低天文潮位分别为 175cm 和 155cm。我们曾怀疑这些大的增减水数据是否合理, 特地对其过程进行了检查, 结果确认这些数据是可信的。

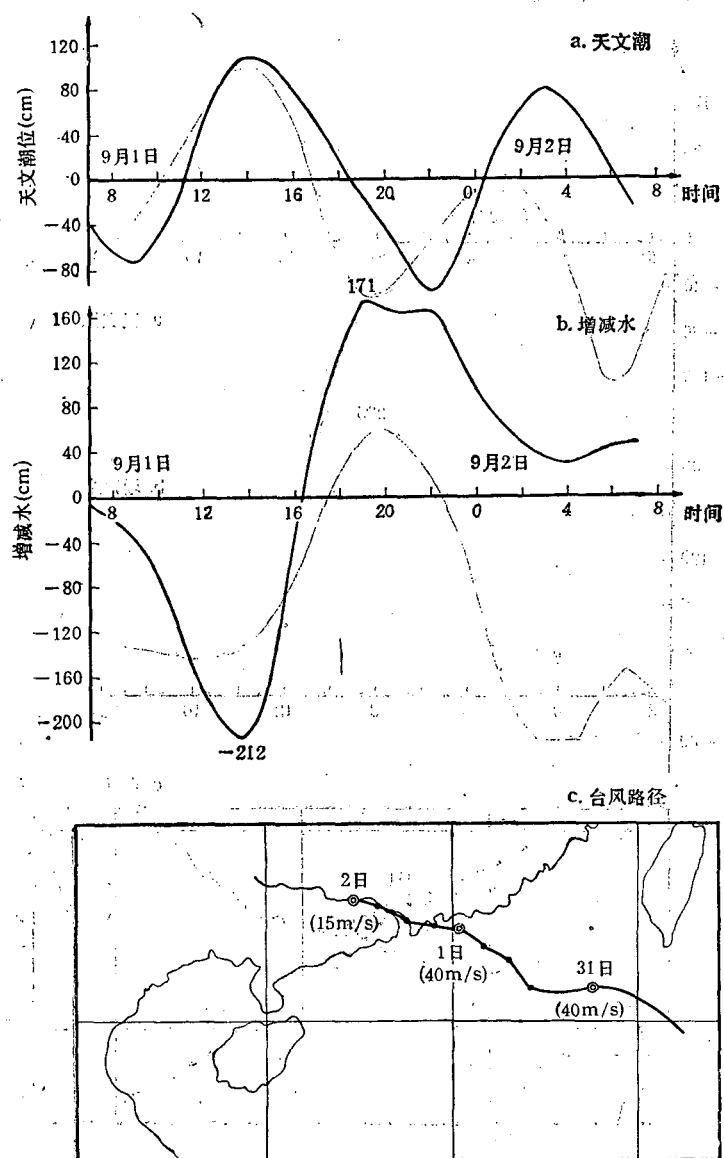


图 6 6213 号台风路径及其在黄埔港引起的风暴潮

图 5 是 6411 号台风引起的黄埔站增水过程, 图中 a 是预报潮汐曲线, b 是增减水曲线, c 示台风路径。该号台风 8 月 8 日靠近珠江口, 风力 12 级以上。约 9 日 0 时在珠江口偏西海岸登陆。从台风路径看, 登陆以前珠江口风向偏东, 无明显增减水, 登陆后转西南风, 故 8 月 9 日 8 时增水达最大, 与风场情况一致。恰巧这时正值天文潮为低潮, 总水位最高发生在 9 时, 其值为平均海面上 197cm, 没有引起灾害, 对极值分布的统计不产生影响。

黄埔最大的减水发生在 1962 年 9 月 1 日 14 时, 也是由台风引起, 潮汐增减水及台风过程见图 6。此次 6213 号台风风力在 12 级以上, 与前面的台风略有不同, 是在珠江口以东登陆, 故在登陆之前, 珠江口为偏北风, 完全可能引起如图 6b 所示的大减水。又是非常凑巧, 最大的减水正好发生在天文高潮, 故总水位并不很低。当台风中心越过珠江口后接着引起增水, 这个增水也不算小, 但又恰巧是天文潮较低的阶段, 亦未引起高水位。此外, 还有一次大增水 (231cm) 出现在 1966 年 7 月 14 日, 亦正值天文低潮。由此可见, 黄埔实测年最高水位不太高, 年最低水位不太低, 是有其偶然性的。这是海洋水位统计中采用极

表 6 由 5 年资料用不同方法所得 50 年一遇水位值及其平均值和标准差

港 口		年 份	Gumbel 法	相似法	年 份	同步差比法
吴	高 水 位	61—65	398	331	59—62	356
		66—70	311	306	66—70	338
		71—75	375	309	71—75	320
		76—80	322	307	76—81	342
		平均值	351	313	平均值	339
		标准差	36.1	10.3	标准差	12.8
淞	低 水 位	61—65	—227	—238	59—62	—228
		66—70	—270	—238	66—70	—230
		71—75	—222	—230	71—75	—217
		76—81	—212	—230	76—81	—218
		平均值	—233	—234	平均值	—223
		标准差	22.2	4.0	标准差	5.8
绿 华 山	高 水 位	58—62	307	292	60—64	338
		63—67	302	291	65—69	337
		68—72	250	278	70—74	316
		73—77	291	281	75—80	305
		78—82	257	275	80—83	284
		平均值	281	283	平均值	316
		标准差	23.5	6.9	标准差	20.3
	低 水 位	58—62	—284	—296	60—64	—281
		63—67	—285	—296	65—69	—287
		68—72	—284	—286	70—74	—284
		73—77	—295	—288	75—78	—277
		78—82	—269	—287	80—83	—281
		平均值	—283	—291	平均值	—282
		标准差	8.3	4.5	标准差	3.3

续表 6

港 口		年 份	Gumbel 法	相似法	年 份	同步差比法
厦 门	高 水 位	64—68	395	376	64—68	393
		69—73	414	379	69—73	387
		74—78	374	376	74—78	387
		79—83	400	376	79—83	398
		平均值	396	377	平均值	391
		标准差	14.4	1.3	标准差	4.6
	低 水 位	64—68	-365	-359	64—68	-368
		69—73	-351	-360	69—73	-376
		74—78	-349	-356	74—78	-384
		79—83	-381	-358	79—83	-383
		平均值	-362	-358	平均值	-378
		标准差	12.8	1.5	标准差	6.4
榆 林	高 水 位	54—58	213	188	56—60	156
		59—63	162	146	61—65	150
		64—68	134	147	66—70	145
		69—73	209	180	71—75	175
		74—78	144	149	76—80	170
		平均值	172	162	平均值	159
		标准差	32.8	18.2	标准差	11.5
	低 水 位	54—58	-195	-150	56—60	-147
		59—63	-152	-132	61—65	-151
		64—68	-137	-130	66—70	-145
		69—73	-132	-128	71—75	-129
		74—78	-119	-122	76—80	-122
		平均值	-147	-132	平均值	-139
		标准差	26.2	9.4	标准差	11.2
潮 州	高 水 位	70—74	305	300	70—74	309
		75—79	259	298	75—79	315
		79—83	295	299	80—83	311
		平均值	286	299	平均值	312
		标准法	19.8	0.8	标准差	2.5
	低 水 位	70—74	-262	-269	70—74	-241
		75—79	-227	-265	75—79	-243
		80—83	-263	-271	80—83	-236
		平均值	-251	-268	平均值	-240
		标准差	16.7	2.5	标准差	2.9

注：1. 表中年份如：61—65 为 1961—1965 年，余同。

2. 水位值为相对于平均海面，单位：cm。

值分布的一个缺点。从上面分析来看，应当说，对黄埔港，相似法（以及独立联合概率法）要比 Gumbel 法合理。

另外值得一提的是汕头的极高水位，相似法结果（313cm）比 Gumbel 法明显高。实际观测到的最高水位也恰巧为 313cm，发生在 1969 年 7 月 28 日 11 时。当时潮位为

31cm,并不算高;但风暴增水为 282cm。这样看来,相似法的结果应当是可信的。

所以,用长期观测资料得出的相似法结果是令人满意的。

对于较短期资料,方法的优劣主要决定于其结果的稳定性。如果用不同年份的较短期资料,所得结果差异较大,则这个方法就不能认为是可靠的。

在表 5 的 10 个验潮站中,我们选出吴淞、绿华山、厦门、榆林和温州 5 个站用于比较不同方法的稳定性。其它验潮站没有选用,是因为不同方法所得结果差别较大。特别是这些站用同步差比法所得结果与 Gumbel 法结果相差太大(两种方法校核高、低水位差合起来超过 60cm),即使稳定性很好亦无实际意义。

以大约 5 年为一组,采用 Gumbel 法、相似法和同步差比法进行计算,所得结果示于表 6。Gumbel 法和相似法各组的资料年份完全一样;同步差比法也尽可能一样,但受主港资料的限制,有少数组略有差别。

正如预料, Gumbel 法的稳定性最差,因而不能根据 5 年资料用该法推算极值水位。相对而言,同步差比法的稳定性要好得多。因此,如果拟建港与主港距离足够近,可以根据 5 年以上资料应用同步差比法计算极值水位。但是,同步差比法的稳定性还是显著低于相似法。

通过上面的比较和讨论,我们认为相似法的效果是良好的。如果观测资料时段较长(20 年以上),有时相似法的结果可能比 Gumbel 法好,因而可以作为参考之用。如观测资料时段较短(20 年以下),除非在足够近处能够找到已知极值水位的主港,一般宜以相似法代替同步差比法。

当然,相似法的计算量明显大于极值分布法和同步差比法。使用电子计算机,这个问题不难解决。

参 考 文 献

- 方国洪、王骥,1987,计算潮位频率分布的一个条件概率方法,港口工程技术规范学术论文集,全国水运工程标准技术委员会、交通部水运规划设计院编,115—118。
- 王骥、方国洪,1981,不完整逐时潮汐观测记录的分析,海洋学报,3(2):193—210。
- 丁文兰、刘风树,1986,6208号台风潮与天文潮相互作用的初步分析,第一届潮汐与海平面学术讨论会论文集,国家海洋局海洋科技情报研究所,56—65。
- 金光炎,1964,水文统计的原理与方法,水利电力出版社,79,85。
- Pugh, D. T. and J.M. Vassie, 1978, Extreme sea levels from tide and surge probability, Proceedings of the 16th Coastal Engineering Conference, 911—929.
- Pugh, D. T. and J. M. Vassie, 1980, Applications of the joint probability method for extreme sea level computations, Proceedings of the Institution of Civil Engineers, Part 2, 959—975.

A METHOD FOR ESTIMATING EXTREME SEA LEVELS BASED ON TIDE-SURGE JOINT PROBABILITY IN OCEAN ENGINEERING*

Fang Guohong

(Institute of Oceanology, Academia Sinica)

Wang Ji, Jia Shaode and Jiao Zongfeng

(Institute of Marine Scientific and Technological Information, SOA)

Ye Anle

(Qingdao Ocean University)

ABSTRACT

Extreme sea levels are traditionally estimated from extrapolation of observed annually maximum (or minimum) sea level heights, by fitting an extreme-value distribution. This approach needs long-term observation, which is often not available for some newly developing areas.

Pugh and Vassie (1978) proposed combining the probability distributions of tide and surge to obtain the probability distribution of total (tide plus surge) sea level. In that study, tide and surge were regarded as independent. Later, Pugh and Vassie (1980) studied the extreme sea level at Southend and concluded that the interaction between tide and surge was important for estuarine areas.

In the present study the dependence of the probability distribution of surge on tidal level is examined. It is shown that for some Chinese coastal areas the statistical characteristics of surge are strongly related to tidal level. For example, in the Bohai Sea the standard deviation of surge in the case of lower tidal level can be twice as large as that for higher tidal level. This is attributed to the meteorological effects. The strong northerly monsoon in winter can cause larger surges and lower the tidal level at the same time.

In order to take into account the dependence of the surge distribution on tidal level, the conditional distribution of surge versus tide is introduced. The marginal distribution of tide can be simply obtained through the statistics of the predicted tidal heights for a period of 19 years. The distributions of surge corresponding to different tidal levels are considered as "similar", that is, the normalized distributions of surge corresponding to different tidal levels are considered as identical to each other and are the same as the normalized marginal distribution of surge. Thus, the conditional distribution of surge can be calculated from the normalized marginal distribution and the standard deviation of surge corresponding to different tidal level. From the joint distribution of tide and surge the distribution of total sea level can be obtained.

The extreme sea levels obtained from the joint probability method are compared with those obtained from the traditional method. It is shown that general consistency is achieved if the observational sea level series are long, and the former results are much more stable than the latter if the observational sea level series are short.

* Contribution No. 2146 from the Institute of Oceanology, Academia Sinica.