

河口海岸模型验证方法及其性能比较研究*

李谊纯, 林振良, 李庆

(北部湾大学建筑工程学院, 钦州市港口航道与近海工程重点实验室, 广西钦州 535011)

摘要:模型验证是河口海岸模型建立和应用的关键步骤,传统的方法包括主观评价、可视化技术、误差估计及统计检验等量化方法,但各种验证方法的性能优劣尚无明确的结论。为探讨各种方法的性能,本文对常用的几种量化方法的性能进行比较,并提出了一种基于统计检验的定量的模型验证方法。研究表明:即使在模型数据的误差明显时,*Skill*、相关系数和决定系数也可能保持较大的值,因此使用这些量化指标进行模型验证需谨慎。相对于绝对误差、均方根误差、*Skill*等常用量化指标,统计检验法既能有效地量化模型模拟的误差,亦能有效减小模拟结果中的异常值对整体模拟效果的影响。由于误差分布的原因,统计检验方法不适用于往复流的流向验证。

关键词:河口海岸模型 验证 置信区间 统计模型 假定检验

中图分类号:P751,P76 文献标识码:A 文章编号:1005-9164(2019)06-0634-07

0 引言

模型验证是确定河口海岸模型的模拟结果能否有效地模拟真实物理过程并进一步应用的最重要步骤。可视化方法一般是指将潮位、潮流、盐度、含沙量等模拟数据和实测数据针对某一参量(如时间)绘制于同一图中,进而比较二者的符合程度,判断模拟结果是否达到要求。此方法已广泛应用于海洋模拟^[1-3]。然而仅通过图形,实测数据和模拟数据之间并没有明确的定量化比较,在某种程度上这种方法也是基于验证者的主观判断。由于人们通常把每个模拟数据点的偏差看作是该点与实测数据曲线的最近距离^[4],所以一个看似令人满意的模拟结果很可能包含相当大的误差。一些量化指标如均方根误差、平均

绝对误差等,计算了每一对模拟数据与实测数据之间的偏差。这些方法直接给出了模拟的效果,但在使用这些方法时也存在一些不足。相关系数也常用于河口海岸模型的验证,它给出了模拟数据和实测数据之间的相关性,但不能量化误差的大小。目前,相关系数多与其他量化指标共同使用。*Skill*是一个常用的模型验证指标,一般与相关系数或均方根误差同时使用,是目前应用最为广泛的验证方法^[2-3]。然而在实际应用中,*Skill*值的变幅相当大,其值也不能直观表示误差的大小。此外,*Skill*并无明确的表征模拟结果可接受的阈值。统计检验在河口海岸模型验证中的应用不多。在河口海岸模型验证中,若将模拟数据和实测数据分别作为横纵轴画于一图,如果二者之间

* 钦州市科学研究与技术开发计划项目(20189922)资助。

【作者简介】

李谊纯(1977—),男,博士,工程师,主要从事河口海岸学研究,E-mail:yichlee@163.com。

【引用本文】

DOI:10.13656/j.cnki.gxkx.20191231.009

李谊纯,林振良,李庆.河口海岸模型验证方法及其性能比较研究[J].广西科学,2019,26(6):634-640.

LI Y C, LIN Z L, LI Q. Comparative study on the performance of estuarine and coastal model verification methods [J]. Guangxi Sciences, 2019, 26 (6): 634-640.

没有偏差,那么数据点将位于直线 $y=x$ 上^[5]。由于模拟数据不可能完全与观测数据一致,所以需要对这些数据点进行回归分析以衡量模拟的效果。 R 检验、 t 检验及 F 检验对于一元线性回归的显著性检验具有相同的效果,可用来检验拟回归模型的线性显著性^[6]。本文首先对目前常用的模型验证指标的性能及敏感性进行比较。在此基础上,对统计检验方法进行了归纳和扩展,给出了基于统计检验的模型验证方法。结果显示该方法对量化模型数据的误差具有很好的适应性和准确性。

1 常用的模型验证量化指标

在河口海岸模型验证中,最常用的量化指标包括均方根误差(RMSE)、平均误差(MAE)平均相对误差(MA%E)等^[4]。Willmott^[7]给出的 *Skill* 是目前河口海岸模型中应用最为广泛的量化指标。

$$RMSE = \sqrt{\sum \frac{(x_{oi} - x_{mi})^2}{n}}, \quad (1)$$

$$MAE = \frac{\sum |x_{oi} - x_{mi}|}{n}, \quad (2)$$

$$MA\%E = \frac{\sum |x_{oi} - x_{mi}| / x_{oi}}{n}, \quad (3)$$

$$Skill = 1 - \frac{\sum_{i=0}^n |x_{oi} - x_{mi}|^2}{\sum_{i=0}^n (|x_{oi} - \bar{x}_o| + |x_{mi} - \bar{x}_o|)^2}, \quad (4)$$

式(1)–(4)中,下标 o, m 分别代表实测和模拟, i 为数据点号,上划线代表时间平均。对于潮流流速的验证, Schnab 等^[8] 和 Beletsky 等^[9] 给出了量化参数 F_n :

$$\|\vec{u}_{oi}, \vec{u}_{mi}\|_2 = \left(\frac{1}{n} \sum |\vec{u}_{oi} - \vec{u}_{mi}|^2\right)^{\frac{1}{2}}, \quad (5)$$

$$F_n = \frac{\|\vec{u}_o, \vec{u}_m\|_2}{\|\vec{u}_o, 0\|_2}, \quad (6)$$

其中式(5)为傅立叶测度(Fourier Norm)。Anderson 等^[10] 给出了一系列的流速验证的参数指标,式(7)–(10),上划线代表时间平均。这些指标分别对应了不同的流速统计量。 ϵ_1 量化了潮流的总流的模拟值与实测值之间的偏差, ϵ_2 和 ϵ_3 分别量化潮流流速中周期性变化部分及余流部分的模拟误差, ϵ_4 为实测潮流和模拟潮流的动能比。

$$\epsilon_1 = \frac{\sum [(u_o - u_m)^2 + (v_o - v_m)^2]}{\sum [(u_o)^2 + (v_o)^2]}, \quad (7)$$

$$\epsilon_2 = \frac{\sum [(u_o - \bar{u}_o - u_m + \bar{u}_m)^2 + (v_o - \bar{v}_o - v_m + \bar{v}_m)^2]}{\sum [(u_o - \bar{u}_o)^2 + (v_o - \bar{v}_o)^2]}, \quad (8)$$

$$\epsilon_3 = \frac{[(\bar{u}_o - \bar{u}_m)^2 + (\bar{v}_o - \bar{v}_m)^2]}{[(\bar{u}_o)^2 + (\bar{v}_o)^2]}, \quad (9)$$

$$\epsilon_4 = \frac{\sqrt{\sum [(u_o)^2 + (v_o)^2]}}{\sqrt{\sum [(u_m)^2 + (v_m)^2]}}. \quad (10)$$

2 统计检验

将模拟数据和实测数据分别置于 x 轴和 y 轴,则可以通过线性回归得到直线:

$$\hat{y}_i = a + bx_i (i = 1, 2, 3, \dots, n), \quad (11)$$

其中 a 和 b 为线性回归参数。如果模型数据和观测数据完全符合,则 $a=0, b=1$ 。回归方程的有效性可以通过假定检验(R 检验、 t 检验或 F 检验,式(12)–(15))来评价。

R 检验:

$$R = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}, \quad (12)$$

其中 $S_{yy} = \sum (y_i - \bar{y})^2, S_{xx} = \sum (x_i - \bar{x})^2, S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$ 。对于置信水平 α 及样本容量 n ,如果 $|R| \geq R_{1-\frac{\alpha}{2}}(n-2)$,则线性关系是显著的。对于 $b=1$ 或 $a=0$ 的显著性,可通过 t 检验或 F 检验($F = t^2$)完成^[11]:

对于 $b, H_0: b=1, H_1: b \neq 1$:

$$t = (b - \beta) \frac{\sqrt{(n-2)S_{xx}}}{\sqrt{S_E}} \sim t(n-2), \quad (13)$$

对于 $a, H_0: a=0, H_1: a \neq 0$:

$$t = a \frac{\sqrt{(n-2)}}{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)S_E}} \sim t(n-2), \quad (14)$$

其中 $S_E = \sum (y_i - \hat{y}_i)^2$ 。上述检验仅对于 a 或 b 分别进行, Dent 和 Blackie 定义了一个统计量 F_D 用以同时对 $b=1$ 和 $a=0$ 进行检验^[5]:

$$F_D = \frac{(n-2) \left[n(a)^2 + 2n\bar{x}_m(a)(b-1) + \sum x_{mi}^2(b-1)^2 \right]}{2n(S)^2}, \quad (15)$$

其中 \bar{x}_m 为模型数据的平均值, S 为标准方差。需要

指出的是,该检验结果仅意味着线性关系是显著的,但不能确定此线性回归方程是一个合适的回归模型^[11]。

这些统计检验并未给出误差的直接量化,而模型模拟的效果需要对模型数据的绝对误差或相对误差进行量化。如《海岸与河口潮流泥沙模拟技术规范》^[12]中规定:高低潮位的误差小于 0.1 m、涨落潮的平均流速误差小于 10%。为了将统计检验与模型数据的误差联系起来,可采用一个新的验证流程,其基本思想是:对于式(11),若强制 $a=0$ 进行回归,则回归直线的斜率代表了平均相对误差;若强制 $b=0$ 进行回归,则回归直线的截距代表了平均绝对误差。

3 数值试验

3.1 数据构建

为研究前述验证指标的性能,构建了两组时间序列过程。首先将潮位过程表示如下:

$$\zeta_o(i) = A \cos\left(\frac{\pi}{12.5}i\right), i = 1, 2, 3, \dots, 25, \quad (16)$$

$$\zeta_m(i) = \zeta_o(i) (1 + f \cdot \text{Rand}(i)), i = 1, 2, 3, \dots, 25, \quad (17)$$

其中: $\zeta_o(i)$ 代表实测数据,振幅(A)为 2 m,频率($\omega = \pi/12.5$ h), $\zeta_m(i)$ 为模型数据。 $\text{Rand}(i)$ 为 $-1.0 \sim 1.0$ 的随机数序列。系数 f 用以控制误差的大小,变化范围为 $0.02 \sim 1.00$ 。时间序列的长度是 25 h。图 1 给出了两个基于式(16)–(17)的时间序列过程。对于式(7)–(11)涉及的流速过程,将式(16)中的振幅 A 分别取 1 m/s 和 0.2 m/s 对应于潮流椭圆的长轴和短轴流速振幅。此外,为研究模型验证指标对模型数据中异常值的适应性,构建数据序列:

$$\zeta_m(i) = \zeta_o(i) + \delta(j)A, i, j = 1, 2, 3, \dots, 25, \quad (18)$$

其中:当 $j=i$ 时, $\delta(j)=1$;当 $j \neq i$ 时, $\delta(j)=0$ 。

3.2 敏感性分析

基于 3.1 中数据序列,图 2 给出了潮位和潮流的相关系数(CC)、Skill 和均方根误差对于系数 f 的变化。在图 2a 中,均方根误差随系数 f 线性增加,这跟此数值试验的假设是一致的。相关系数呈现单调递减的趋势。当系数 f 为 0 时相关系数为 1;系数 f 为 1 时,相关系数为 0.83。随着系数 f 的增加,相关系数的变幅越大。Skill 的变化趋势与相关系数非常类似,当系数 f 从 0 逐渐增加到 1 时, Skill 从 1 减小至

0.89。系数 $f=1$ 代表了平均相对误差约为 50%,最大相对误差约为 100%。所以可以看出,不论是相关系数还是 Skill,其对误差的变化并不敏感。在模拟

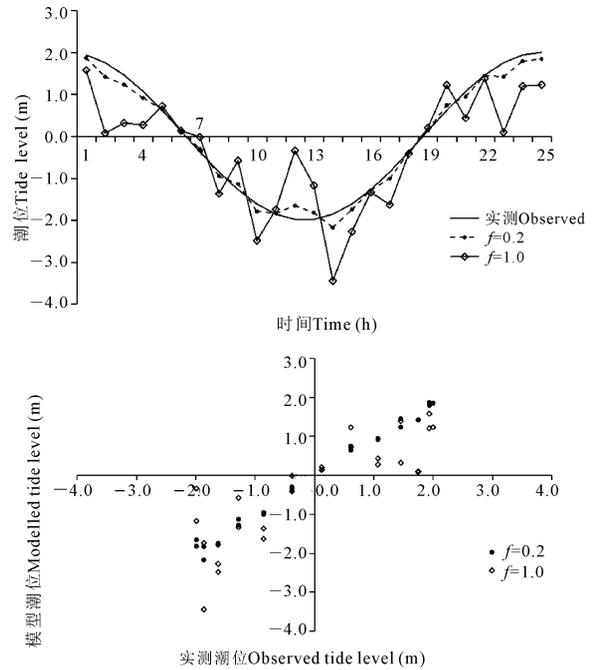


图 1 潮位实测数据和模拟数据示意图

Fig. 1 Schematic diagram of observed and simulated data for tide

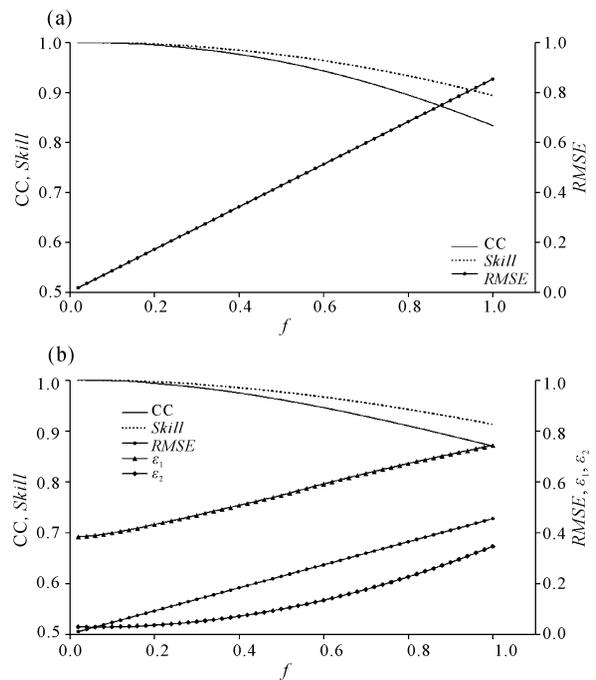


图 2 潮位(a)、流速(b)验证相关系数(CC)、Skill、均方根误差及 ϵ_1 、 ϵ_2 随 f 的变化

Fig. 2 Variations of tide level (a), velocity (b) verification correlation coefficient (CC), Skill, RMSE, ϵ_1 and ϵ_2 against factor f

数据的误差相当大时, 相关系数或 *Skill* 值仍可能出现较大的值。图 2b 给出了基于前节所述方法构造的潮流流速数据计算的式(7)–(9)的值。可以看出, 相关系数、*Skill* 及均方根误差呈现与图 2b 类似的变化趋势, 相关系数与 *Skill* 分别由 1 减小至 0.87 和 0.91。 ϵ_1 的增加约为 0.38 至 0.74, ϵ_2 增加约 0.03 至约 0.35。事实上, 对于不同的模型试验, 由于其背景流速值(式(7)–(9)中的分母)变化很大, ϵ_1 和 ϵ_2 在不同模型之间可能差别很大, $\epsilon_1 - \epsilon_4$ 在不同模型间不具有可比性。

综合而言, *Skill* 对偏差没有显著的敏感性。即使是系数 f 达到 1.0, 它的范围仍为 1.0~0.9, 相关系数也保持在 1.0 至 0.8 的区间。其他一些模拟显示的 *Skill* 和相关系数值可能小于上面的数值实验。如果随机数序列是均匀分布的, 则当 $f=1$ 时, 平均误差约为 50%。这说明虽然模拟数据和实测数据之间的偏差比较大, 但是 *Skill* 与相关系数仍可能为很大的值。虽然不能据此断定小的 *Skill* 值一定意味着模拟不理想, 但是在较大的误差情况下可能会有一个明显大的 *Skill*, 这可能会导致对模拟结果好坏的误判。在这个数值实验中, 相关系数显示了实测数据和模拟数据序列之间的相关性, 但是它不能量化误差, 不能单独用于河口海岸模型验证。根据以上分析, *Skill* 和相关系数可用于潮位, 盐度, 泥沙等标量验证, 也可用于对比潮流的流速与流向, 但使用 *Skill* 和相关系数的验证方法是否是一个合适的方法是需谨慎的。

3.3 统计检验方法

3.3.1 与误差直接计算方法的一致性

如前所述, 基于式(11)–(15)的方法不能与模拟数据的误差直接联系起来。若是预设式(11)中的截距(斜率), 则斜率(截距)在一定意义上则代表了平均相对(绝对)误差。对于这两种情况, 相应的斜率和截距可分别表示如下

$$b = \frac{\sum x_{oi} x_{mi}}{\sum x_{oi}}, \quad (19)$$

$$a = \frac{\sum y_i - \sum x_i}{n}. \quad (20)$$

利用式(13)、(14)对式(19)、(20)进行显著性检验, 检验斜率(b)或截距(a)是否显著不为 1 或 0。若在某个显著性水平 α 时检验通过, 则 b 的置信区间可以表示为

$$(b_{low}, b_{high}) = \left(b - t_{\frac{\alpha}{2}} \frac{\sqrt{S_E}}{\sqrt{(n-2)S_{xx}}}, b + t_{\frac{\alpha}{2}} \frac{\sqrt{S_E}}{\sqrt{(n-2)S_{xx}}} \right), \quad (21)$$

其中, b_{low} 和 b_{high} 分别为置信区间的上、下限。当截距预设 0 时, $(b-1)$ 则代表了平均相对误差。最大相对误差可定义为

$$E_b = \text{Max}(\text{Abs}(b_{low} - 1), \text{Abs}(b_{high} - 1)) \times 100\%, \quad (22)$$

其中 Max 和 Abs 分别为最大值和绝对值。类似的, 对于预设斜率 $b=1$ 的情况, 截距 a 的置信区间可表示为

$$(a_{low}, a_{high}) = \left(a - t_{\frac{\alpha}{2}} \frac{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)S_E}}{\sqrt{(n-2)}}, a + t_{\frac{\alpha}{2}} \frac{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)S_E}}{\sqrt{(n-2)}} \right), \quad (23)$$

类似式(22), 最大绝对误差可定义为

$$E_a = \text{Max}(\text{Abs}(a_{low}), \text{Abs}(a_{high})). \quad (24)$$

利用 3.1 中生成的数据对式(19)–(24)进行数值试验。计算结果如图 3 和表 1 所示。其中相对误差和绝对误差的阈值分别设为 10% 和 0.1 m。

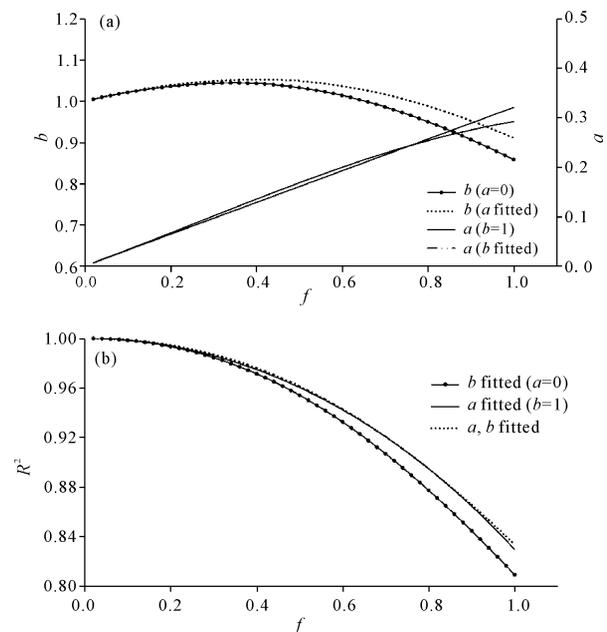


图 3 斜率 b 和截距 a (a), 决定系数 R^2 (b) 随 f 的变化
Fig. 3 Variations of slope b and intercept a and coefficient of determination R^2 against factor f

图 3a 中, 对于 a 和 b 均拟合(Fitted)的情况, 在 f 为 0.4 左右时, 斜率 b 达到最大, 约为 1.05; 当系数 f 为 1 时, b 约为 0.91。截距 a 随着 f 的增加而

单调增加,从0增至约0.29。当 a 预设0时, b 的变化呈现与之前明显的差异:当 f 小于0.3时, b 几乎与 a 和 b 均拟合的情况相同;当 f 大于0.3时,随着 f 的增大,二者差异增大。当系数 f 为1时, b 约为0.81。当 b 的预设1时, a 的值与 a 和 b 均拟合

表1 统计检验试验结果

Table 1 Results of the statistic tests

	试验1 Test 1	试验2 Test 2	试验3 Test 3	试验4 Test 4		试验5 Test 5	
				$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.05$	$\alpha=0.01$
接受域 Acceptance region	全域 All region	全域 All region	全域 All region	$f \leq 0.22$	$f \leq 0.18$	$RMSE \leq 0.12 \text{ m}$	$RMSE \leq 0.10 \text{ m}$
拒绝域 Rejection region	—	—	—	$f > 0.22$	$f > 0.18$	$RMSE > 0.12 \text{ m}$	$RMSE > 0.10 \text{ m}$

表1给出了相关统计检验的结果。将实测数据和模拟数据画于同一图中,横轴和纵轴分别代表实测数据和模拟数据。线性拟合后,首先利用 R 检验, F 检验和 t 检验线性关系的显著性进行检验(试验1)。而后,利用式(13)和(14)检验 b 是否显不等于1(试验2)以及 a 是否显著不同于0(试验3)。结果表明,无论置信水平为0.05或0.01,即使系数 f 达到一个相当大的值,这些检验都能通过。在此基础上,按式(22)或式(24)计算统计意义上的最大的相对误差和绝对误差;进一步将区间限与判断模拟结果是否可接受的阈值进行比较并判定模拟结果是否可接受。在给定显著水平($\alpha=0.05, 0.01$)的情况下,试验4是针对阈值为10%的检验。在试验4中,当 $\alpha=0.05$ 时 f 的临界值为0.22,当 $\alpha=0.01$ 时 f 的临界值为0.18。 f 值平均约为0.20,这意味着20%的最大误差。如果误差是均匀分布的,那么也意味着10%的平均误差。试验5是针对绝对误差取阈值为0.1m的检验。在试验5中,对于 $\alpha=0.05$ 和 $\alpha=0.01$, $RMSE$ 的临界值分别为0.12m和0.10m,与假设的阈值近似一致。由于表1中的结果具有统计意义,计算得到的临界值并不等于严格等于预设的阈值。据此试验,可以认为 $\alpha=0.05$ 和 $\alpha=0.01$ 都是可接受的显著性水平。

3.3.2 对异常数据的包容性

为检验前述统计检验方法对异常数据的包容性,利用式(19)得到的数据对式(21)~(22)进行计算。在该数据序列中,只有某一数据点(j 点)增大一个振幅,其他值保持不变。这种情况下,如果采用最大相对误差或最大绝对误差都会得到模型结果不可接受的结论。即使是均方根误差,亦达到0.4m。对于置

信水平 $\alpha=0.05$ 和 $\alpha=0.01$,式(24)得到的最大统计误差约为0.25m和0.31m,远小于式(19)数据序列直接计算的最大误差。与均方根误差相比,亦明显较小。可以认为,此统计检验方法虽不能自动识别和剔除异常数据点,但对其导致的干扰具有明显的弱化效果。

3.3.3 在瓯江模型中的应用

图4为瓯江口北口七里港附近某测点2005年6月洪季的一次实测的流速和流向过程与物理模型试验结果的对比。在此,对流速和流向过程分别应用式

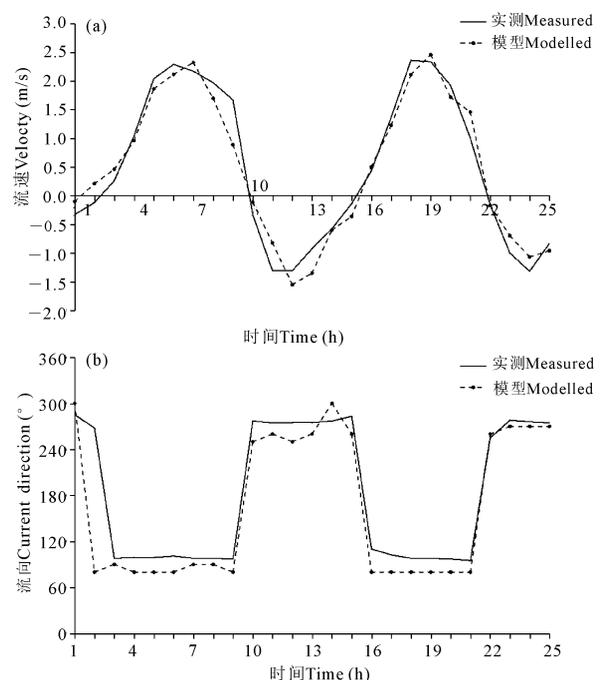


图4 流速(a)和流向(b)的实测与模型试验的数值对比
Fig. 4 Measured numerical comparison of flow velocity (a) and flow direction (b) with model test

(21)–(24)。表 2 给出了相关统计量的值。由图 4 可以看出, 流向过程模型值偏小, 流速过程在直观上符合较好。如前所述, 由于每个点的偏差常被直观看做是模型点与实测数据线的最近距离, 所以看似令人满意的模拟结果可能包含相当大的误差。表 2 列出了相关系数、Skill 等指标的计算值。对于流速, 相关系数和 Skill 分别为 0.975 和 0.987。一般而言, 这样的数值已可以表征很好的符合度。然而, 平均误差和均方根误差分别为 0.24 m/s 和 0.29 m/s, 最大误差甚至达到 0.78 m/s。可见, 虽然由相关系数和 Skill 判断, 模型结果是可以接受的, 但是平均误差约为 20%, 远大于《海岸与河口潮流泥沙模拟技术规范》^[12]所规定的阈值 10%。流向的结果与流速过程类似。结合 3.2 节的相关系数和 Skill 对误差的敏感性分析, 可以认为采用相关系数结合 Skill 量化模

表 2 统计检验试验结果

Table 2 Results of statistic tests

项目 Item	最大误差 Maximum error	MAE	RMS	CC	Skill	相对统计误差 (%) Relative statistical error (%)		绝对统计误差 (m/s) Absolute statistical error (m/s)	
						$\alpha=0.05$	$\alpha=0.01$	$\alpha=0.05$	$\alpha=0.01$
						流速 Velocity	0.78 m/s	0.24 m/s	0.29 m/s
流向 Direction	188°	23°	41°	0.92	0.949				

4 结论

本文基于数值试验对一系列河口海岸模型验证方法进行了总结和比较。之后, 对统计检验方法进行了分析和扩展。结果表明, 即使模型误差显著, Skill 和相关系数也可能得出较大的量值, 这将导致模拟数据和实测数据符合良好的错误结论, 最小二乘回归拟合的判定系数对误差的响应也不明显。统计检验是模型验证中的一种有效方法。针对斜率(b)和截距(a)的一般统计检验, 如 R 检验, t 检验和 F 检验, 只检验线性关系的显著性而不是量化误差。当预设截距为 0 或斜率为 1 时, 最小二乘拟合的系数可用于估计相对误差或绝对误差。统计误差于直接计算的误差具有一定的一致性, 但比直接计算的误差更稳定。由于往复流流向的误差不一定满足正态分布假设, 统计检验法对其不适用。

参考文献

[1] PURKIANI K, BECHERER J, FLÖSER G, et al. Numerical analysis of stratification and destratification

型的验证效果存在一定的不准确性。表 2 中同时列出了基于式 (21)–(24) 所得统计误差的计算结果。当 $\alpha=0.05$ 时绝对统计误差为 0.15 m/s, 相对统计误差为 12.9%。由于统计误差要求是在某个置信水平下误差的区间范围, 因此就数值上直观而言, 统计误差的值比直接计算误差似乎表明模拟结果具有更好的符合性。针对更多数据的检验尚待后续研究。此外, 值得注意的是, 前述统计误差的计算方法及其对于流速或潮位过程得出的结论对于流向过程不一定适用, 这是由于往复流的流向误差的分布与潮位或流速误差的分布不同, 转流时的误差(图 4 第 2 个数据点)可能导致误差的分布不满足正态分布假设。利用统计误差对旋转流流向进行验证尚待进一步的研究。

processes in a tidally energetic inlet with an ebb tidal delta [J]. Journal of Geophysical Research: Oceans, 2015, 120(1):225-243.

- [2] GRÄWE U, FLÖSER LÖSER G, GERKEMA T, et al. A numerical model for the entire Wadden Sea: Skill assessment and analysis of hydrodynamics [J]. Journal of Geophysical Research: Oceans, 2016, 121(7): 5231-5251. DOI:10.1002/2016JC011655.
- [3] LIN W B, WANG Y G, RUAN X H, et al. Modeling residual circulation and stratification in Oujiang river estuary [J]. China Ocean Engineering, 2012, 26(2): 351-362.
- [4] MAYER D G, BUTLER D G. Statistical validation [J]. Ecological Modelling, 1993, 68(1/2): 21-32.
- [5] DENT J B, BLACKIE M J, HARRISON S R. Systems simulation in agriculture [M]. US, Barking: Applied Science Publishers, 1979.
- [6] 周纪芾, 蒯诗松. 概率论与数理统计[M]. 第 3 版. 北京: 中国统计出版社, 2007.
- [7] WILLMOTT C J. On the validation of models [J]. Physical Geography, 1981, 2(2): 184-194.

- [8] SCHWAB D J, CLITES A H, MURTHY C R, et al. The effect of wind on transport and circulation in lake St. Clair [J]. *Journal of Geophysical Research Atmospheres*, 1989, 94(C4): 4947-4958.
- [9] BELETSKY D, SCHWAB D J. Modeling circulation and thermal structure in Lake Michigan: Annual cycle and interannual variability [J]. *Journal of Geophysical Research: Oceans*, 2001, 106(C9): 19745-19771.
- [10] ANDERSON E J, SCHWAB D J. Relationships between wind-driven and hydraulic flow in Lake St. Clair and the St. Clair River Delta [J]. *Journal of Great Lakes Research*, 2011, 37(1): 147-158.
- [11] 王惠文. 偏最小二乘回归方法及其应用 [M]. 北京: 国防工业出版社, 1999.
- [12] 中华人民共和国交通运输部天津水运工程科学研究所. 海岸与河口潮流泥沙模拟技术规程 [S]. 北京: 人民交通出版社, 2010.

Comparative Study on the Performance of Estuarine and Coastal Model Verification Methods

LI Yichun, LIN Zhenliang, LI Qing

(Qinzhou Key Laboratory of Waterway and Coastal Engineering, College of Civil Engineering and Architecture, Beibu Gulf University, Qinzhou, Guangxi, 535011, China)

Abstract: Model verification is a key step in the establishment and application of estuarine and coastal models. Traditional methods are quantitative methods including subjective evaluations, visualization technologies, error estimations and statistical tests. However, the performance of various verification methods has not been clearly defined. In order to explore the performance of various methods, this paper compares the performance of several commonly used quantitative methods and a quantitative model verification method based on statistical tests is proposed. It shows that even when the errors of model data are significant, the values of *Skill*, correlation coefficient and determination coefficient might maintain relatively large, so cautions should be exercised when using these quantitative indicators for model verification. Compared with commonly used quantitative indicators such as absolute error, root mean square error and *Skill*, the statistical test method can not only effectively quantify the error of model simulation, but also effectively reduce the influence of abnormal values in the simulation results on the overall simulation effect. Due to the distribution of error, the statistical test method might not be suitable for flow direction verification of alternating current.

Key words: estuary coastal model, verification, confidence interval, statistical model, hypothesis test

责任编辑: 符支宏



微信公众号投稿更便捷

联系电话: 0771-2503923

邮箱: gxkx@gxas.cn

投稿系统网址: <http://gxkx.ijournal.cn/gxkx/ch>