HYDROLOGY PROJECT

Technical Assistance

World Bank & Government of The Netherlands funded

Training module # SWDP - 30

How to validate rating curve

New Delhi, November 1999

CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas,
New Delhi – 11 00 16 India
Tel: 68 61 681 / 84 Fax: (+ 91 11) 68 61 685
E-Mail: dhvdelft@del2.vsnl.net.in

DHV Consultants BV & DELFT HYDRAULICS

with
HALCROW, TAHAL, CES, ORG & JPS

Table of contents

1. Module context

While designing a training course, the relationship between this module and the others, would be maintained by keeping them close together in the syllabus and place them in a logical sequence. The actual selection of the topics and the depth of training would, of course, depend on the training needs of the participants, i.e. their knowledge level and skills performance upon the start of the course.

2. Module profile

Title	:	How to validate rating curve
Target group	:	Assistant Hydrologists, Hydrologists, Data Processing Centre Managers
Duration	:	One session of 60 minutes
Objectives	:	After the training the participants will be able to: • Validate the rating curve
Key concepts	:	• Graphical validation • Statistical validation
Training methods	:	Lecture, software
Training tools required	:	Board, OHS, Computer
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	***General*** Overhead: Text: Rating curve validation (1) Overhead: Text: Rating curve validation (2)	5 min	 OHS 1 OHS 2
2	***Graphical validation tests*** Overhead: Figure 2.1: Rating curve fitting Overhead: Table 2.1: Results of curve fitting Overhead: Figure 2.2: Results of validation (1) Overhead: Figure 2.3: Results of validation (2) Overhead: Figure 2.4: Period-flow deviation scatterdiagram Overhead: Figure 2.5: Stage-flow deviation diagram Overhead: Figure 2.6: Cumulative deviation plot	20 min	 OHS 3 OHS 4 OHS 5 OHS 6 OHS 7 OHS 8 OHS 9
3	***Numerical validation tests*** Overhead: Text:Use of Student's 't' test to check gaugings Overhead: Table 2.2 Example of 't' test Overhead: Text: Absence from bias test (signs) Overhead: Text: Absence from bias test (values) Overhead: Text: Goodness of fit test	20 min	 OHS 10 OHS 11 OHS 12 OHS 13 OHS 14
4	***Exercise*** Validate the gaugings of August 1997 for stations Rakshewa and Khamgaon. Existing rating curve is based on data 1/6-3/8/97		

4. Overhead/flipchart master

5. Handout

Add copy of Main text in chapter 8, for all participants.

6. Additional handout

These handouts are distributed during delivery and contain test questions, answers to questions, special worksheets, optional information, and other matters you would not like to be seen in the regular handouts.

It is a good practice to pre-punch these additional handouts, so the participants can easily insert them in the main handout folder.

7. Main text

Contents

How to validate rating curve

1. General

- **Validation of a rating curve is required both after the relationship has first been fitted and subsequently when new gaugings have been carried out, to assess whether these indicate a change in rating. Validation is also used to assess the reliability of historical ratings.**

  Current meter gauging is carried out with variable frequency depending on previous experience of the stability of the control and of the rating curve. As a minimum it is recommended that six gaugings per year are carried out even with a station with a stable section and previously gauged over the full range of level. At unstable sections many more gaugings are required. The deviation of such check gaugings from the previously established relationship is computed and any bias assessed to determine whether they belong to the same population as the previous stage-discharge relationship.

- **Graphical and numerical tests are designed to show whether gaugings fit the current relationship equally and without bias over the full range of flow and over the full time period to which it has been applied**. If they do not, then a new rating should be developed as described in Module 29, but taking into account the deficiencies noted in validation.

- **Validation will be carried out at Divisional offices or at the State Data Processing Centre.**

2. Graphical validation tests

2.1 General

Graphical tests are often the most effective method of validation. These include the following:

- Stage/discharge plot with the new gaugings
- Period/flow deviation scattergram
- Stage/flow deviation scattergram
- Cumulative deviation plot of gaugings.
- Stage/discharge plots with gaugings distinguished by season

Judgements based on graphical displays are often indicative rather than prescriptive - a judgement on the part of the data processor is still required.

2.2 Stage/discharge plot with new gaugings

The simplest means of validating the rating curve with respect to subsequent gaugings is to plot the existing rating curve with the new check gaugings. This is shown in the example for Station Pargaon. A rating curve is established for the period 30/6 – 3/8/97, see Figure 2.1. It shows a proper fit of the data to the existing rating curve, of which the numerical results are shown in Table 2.1. New data are available for the period 4-23/8/97. The new data with the existing rating curve are shown in Figure 2.2. From this plot it is observed that the new gaugings do not match with the existing curve. In Figure 2.3 the new gaugings are shown with the rating curve and its the 95% confidence limits (derived as t-times the standard error S_e). From this plot it can be judged if most check gaugings lie inside the confidence limits and thus whether they can be judged acceptable with respect to deviation. It is expected that 19 out of 20 observations will lie inside the limits if the standard error is considered at 5% significance level. However, except insofar as one can see whether all the new points lie

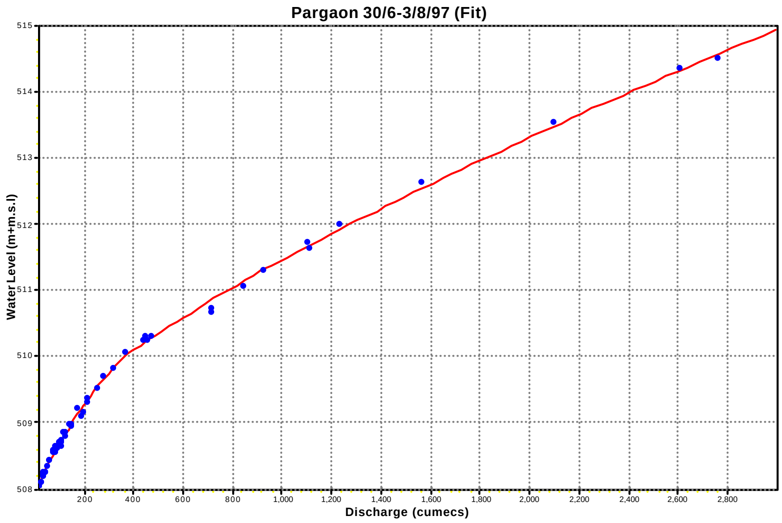

Figure 2.1 Rating curve for station PARGAON established based on data for the period 30/6-3/8/97

Table 2.1 Results of rating curve fitting

Analysis of stage-discharge data
Station name : PARGAON
Data from 1997 6 30 to 1997 8 2
Single channel

Given boundaries for computation of rating curve(s)

interval	lower bound	upper bound	nr. Of data
1	508.000	510.200	49
2	509.900	515.000	12

Power type of equation q=c*(h+a)**b is used

Boundaries / coefficients

lower bound	upper bound	a	b	c
508.00	509.98	-507.730	1.489	.1043E+03
509.98	515.00	-508.490	1.475	.1936E+03

Number	W level M	Q meas M3/S	Q comp M3/S	DIFf M3/S	Rel.dIFf 0/0	Semr 0/0
54	508.050	22.830	19.116	3.714	19.43	4.44
56	511.240	820.010	861.213	-41.203	-4.78	2.58
57	510.740	711.230	640.582	70.648	11.03	2.96
62	512.620	1566.740	1568.961	-2.221	-.14	3.12
63	514.510	2757.370	2735.225	22.145	.81	4.60
64	514.360	2609.830	2635.279	-25.448	-.97	4.49
65	513.530	2098.120	2104.625	-6.505	-.31	3.84
66	512.010	1235.470	1239.485	-4.015	-.32	2.71
67	511.730	1103.470	1096.850	6.620	.60	2.59

Overall standard error = 6.061

Statistics per interval

Interval	Lower bound	Upper bound	Nr.of data	Standard error
1	508.000	509.981	48	6.55
2	509.981	515.000	12	4.01

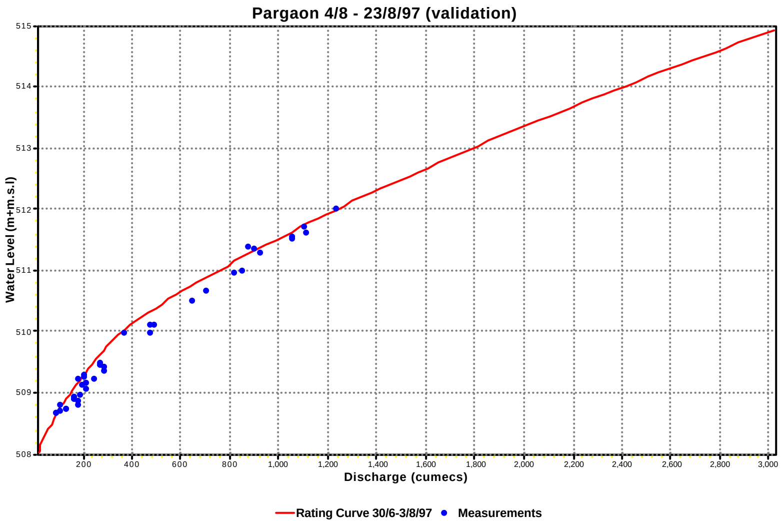

Figure 2.2 New gaugings at Pargaon station plotted against the existing rating curve

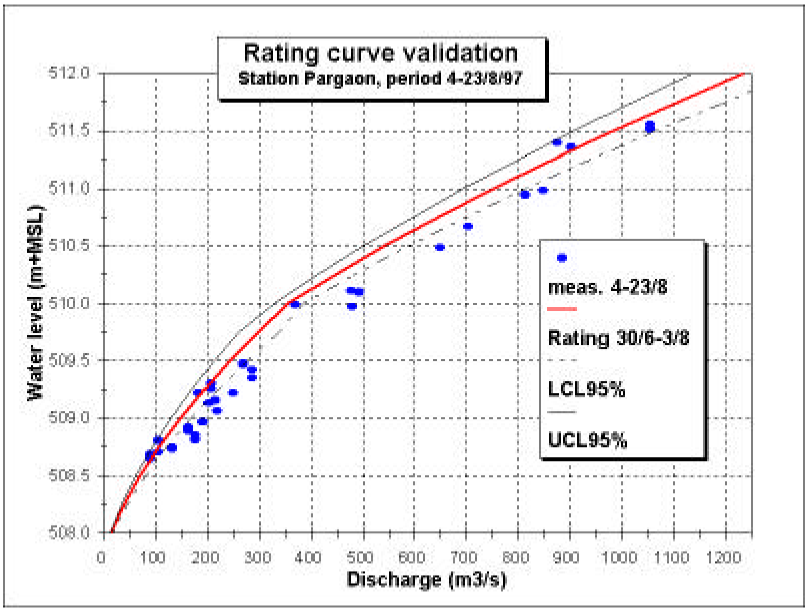

Figure 2.3 New gaugings at Pargaon station plotted against existing rating curve with 95% confidence limits

above or below the previous regression line, the graph does not specifically address the problem of bias. For example, if some 25 new gaugings may all lie scattered within 95% confidence limits, it does not show any significant change in behaviour. However, if these points are plotted and sequence of each observation is also considered and if upon that a certain pattern of deviation (with respect to time) is perceivable and significant then such situation may warrant new ratings for different periods of distinct behaviour. For the Pargaon-case the plot confirms earlier observations that the new gaugings significantly differ from the existing rating.

2.2 Period/flow deviation scattergram

A period/flow deviation scattergram (Figure 2.4) is a means of illustrating the negative and positive deviation of each current meter gauging from the present rating curve and whether there has been a gradual or sudden shift in the direction of deviations within the period to which the rating has been applied. It also shows whether recent additional gaugings show deviation from previous experience. In the example shown in Fig. 2.4, percentage deviations are very high; there are far more gaugings with positive than with negative deviations. The rating is therefore biased and a revision of the rating is strongly recommended.

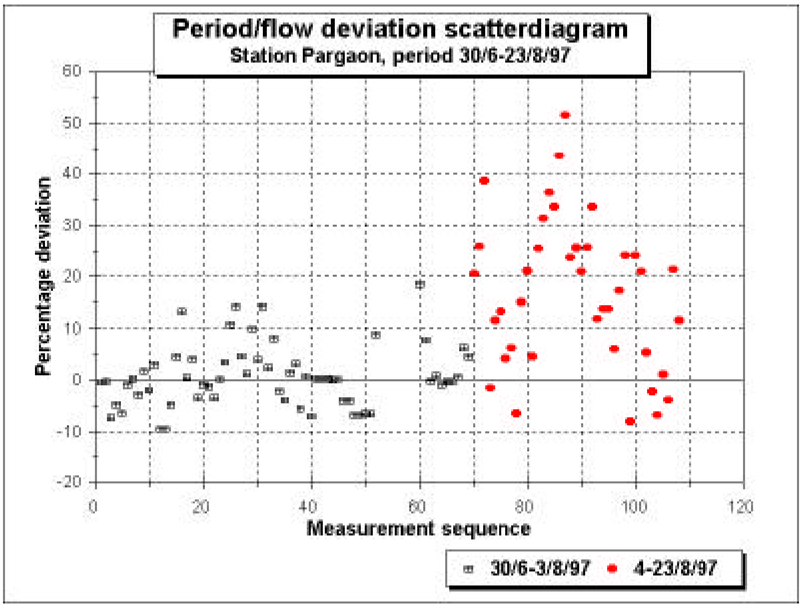

Figure 2.4 Period-flow deviation scatterdiagram for Pargaon rating curve data and new gaugings

2.3 Stage/flow deviation diagram

A similar scattergram plot shows the percentage deviation with stage (Figure 2.5) and is a means of illustrating whether over certain ranges of stage the relationship is biased. Most recent gaugings can also be placed within this context.

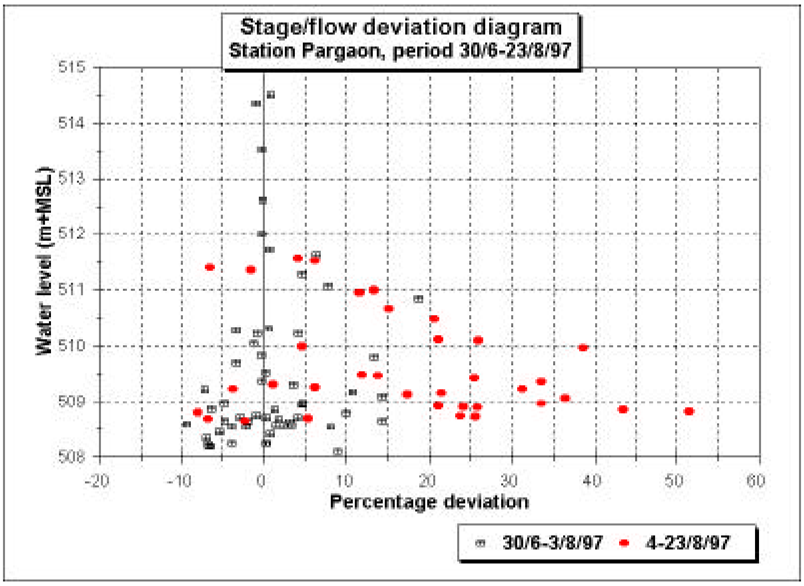

Figure 2.5 Stage-flow deviation scatterdiagram for Pargaon rating curve data and new gaugings

In the example shown in Fig. 2.5, there is some difference in deviation at different stages; particularly at the lower stages the differences are substantial. This plot also confirms the necessity for revision of the rating curve.

2.4 Cumulative deviation plot of gaugings

A plot of the cumulative deviation of gaugings from the rating curve give another indication of bias and whether that bias changes with time. Fig. 2.6 shows such a plot for the example of station Pargaon. From the upward trend of the line for the new gaugings it is concluded that the new gaugings produce consistently higher flow values for the same stages than before.

2.5 Stage discharge plots with gaugings distinguished by season.

It is sometimes helpful to separate gaugings between seasons to demonstrate the effect of varying weed growth or other seasonal factors on the stage discharge relationship. The effects of weed growth may be expected to be at a maximum in low flows before the onset of the monsoon; monsoon high flows wash out the weed which increases progressively from the end of the rains. The discharge for given level may thus differ from one month to another. This shows up more clearly in rivers where winter low flows are little affected by weed growth than summer low flows and thus show much smaller spread. Where an auxiliary gauge is available, a backwater rating curve (normal fall method) may be used. Otherwise a simple rating curve may be used in weed absent periods and Stout's shift method during periods of variability.

3. Numerical validation tests

3.1 Use of Student's 't' test to check gaugings

A test such as Student's "t" test may be used to decide whether check gaugings can be accepted as part of the homogeneous sample of observations making up the existing stage-discharge curve. Such a test will indicate whether or not the stage-discharge relation requires re-calculation or the section requires recalibration.

In this test, the 't' statistic is calculated as the ratio of the mean deviation and the standard error of the difference of the means as:

$$t = \bar{d}_1 / S \qquad (1)$$

where $\bar{d}_1$ is the mean deviation of the new gaugings (in percent) from the existing curve

and S is the standard error of the difference in the means expressed as:

$$S = a\sqrt{(N + N_1)/NN_1} \qquad (2)$$

where N is the number of gaugings used to derive the existing rating

and N_1 is the number of new gaugings

a is given by the following expression:

$$a = \sqrt{\frac{\sum(d^2) + \sum(d_1 - \bar{d}_1)^2}{N + N_1 - 2}} \qquad (3)$$

where $S(d)^2$ is the sum of the squares of the percent differences for the old gaugings from the existing rating.

If this computed value of *'t'* = $\bar{d}$ / *S* is greater than the critical value of *t'* for ($N + N_1 - 2$) degrees of freedom at 95% probability level then further action must be considered. Either the development of a new rating or a request to field staff for additional check gaugings. The critical values of Students *'t'* statistic at the 95% confidence level can be obtained from the standard tables available for the Student's 't' distribution. It should be noted that rating changes are more frequent and more noticeable in the low flow range. Review and validation is therefore done with respect to each range and, unless there is evidence to the contrary, unaffected ranges should retain the old rating but with the range limits adjusted for the new intersection.

As an example, the validation of the new gaugings at station Pargaon is shown in Table 3.1. The results of the 't'-test is seen to support earlier observation of significant deviation.

Table 3.1 Results of validation

```
Validation stage-discharge data
 Station name :  PARGAON
 Data from 1997  8  4 to 1997  8 22

 Procedure    : Standard
 Equation type: Power
 Interval  Boundaries           Parameters:
     1       508.000   509.981  -507.730     1.489   104.324
     2       509.981   515.000  -508.490     1.475   193.614

 Data used to estimate parameters:
 Interval  St. error of est.  Number of data
    1          6.546            48
    2          4.013            12

 Number   W level    Q meas    Q comp      DIFf  Rel.dIFf      Semr
                M      M3/S      M3/S      M3/S       0/0       0/0

     81   509.990   368.450   352.261    16.189      4.60     15.48
     82   509.420   285.790   227.843    57.947     25.43     15.51
     83   509.220   247.940   188.880    59.060     31.27     11.81
     84   509.060   217.610   159.488    58.122     36.44     10.23
     85   508.960   189.660   141.964    47.696     33.60     10.57
     86   508.850   177.270   123.482    53.788     43.56     12.38
     87   508.810   177.270   116.972    60.298     51.55     13.39
     88   508.740   130.970   105.863    25.107     23.72     15.53
     89   508.730   130.970   104.309    26.661     25.56     15.88
     90   508.920   163.570   135.149    28.421     21.03     11.05
     91   508.890   163.570   130.107    33.463     25.72     11.55
     92   509.350   285.790   213.934    71.856     33.59     14.11
     93   509.480   268.610   239.991    28.619     11.93     16.73
     94   509.460   268.610   235.915    32.695     13.86     16.32


 Overall standard error =    14.729

 Statistics per interval
 Interval Lower bound Upper bound Nr.of data Standard error
        1     508.000     509.981        27           25.78
        2     509.981     515.000        15           13.18

 Results of student T-test on absence of bias
 Interval Degrees of freedom 95% T-value Actual T-value Result
        1                 73       1.993          7.673 Reject
        2                 25       2.060          2.920 Reject
```

3.2 Test for absence from bias in signs

A well-balanced rating curve must ensure that the number of positive and negative deviations of the observed values from the rating curve is evenly distributed. That is, the difference in number between the two should not be more than can be explained by chance fluctuations. The test is employed to see if the curve has been established in a balanced manner so that the two sets of discharge values, observed and estimated (from the curve), may be reasonably supposed to represent the same population.

This test is performed by counting observed points falling on either side of the curve. If Q_i is the observed value and Q_c the estimated value, then the expression, Q_i - Q_c , should have an equal chance of being positive or negative. In other words, the probability of Q_i - Q_c being positive or negative is ½. Hence, assuming the successive signs to be independent of each other, the sequence of the differences may be considered as distributed according to the binomial law (p+q) N, where N is the number of observations, and p and q, are the probabilities of occurrence of positive and negative values are ½ each. The expected number of positive signs is positive signs is Np. Its standard deviation is √(Npq). **The "t" statistic is then found by dividing the difference between the actual number of positive signs N_1 and expected number of positive signs Np by its standard deviation√(Npq):**

$$t = \frac{|N_1 - Np| - 0.5}{\sqrt{Npq}}$$

The resulting value is compared with the critical value of "t" statistic for 5% significance level for the degrees of freedom equal to the total number of stage discharge data. If the value of the critical "t" statistic is more than that obtained for the observed data then it can be considered that the data does not show any bias with respect to sign of the deviations between observed and computed discharges.

3.3 Test for absence from bias in values

This test is designed to see if a particular stage discharge curve, on average, yields significant under estimates or over estimates as compared to the actual observations on which it is based. (Compare the graphical test using the period/flow deviation and stage /flow deviation scattergrams) The percentage differences are first worked out as:

$$P = 100\,(Q_i - Q_c)\,/\,Q_c \qquad (4)$$

If there are N observations and P_1, P_2, P_3, …, P_N are the percentage differences and P_{av} is the average of these differences, the standard error of P_{av} is given by:

$$Se = \sqrt{\frac{(\sum P_i - P_{av})^2)}{N(N-1)}} \qquad (5)$$

The average percent P_{av} is tested against its standard error to see if it is significantly different from zero. The "t" statistic for in this case is computed as:

$$t = (P_{av} - 0) / S_e \qquad (6)$$

If the critical value of "t" statistic for 5% significance level and N degrees of freedom is greater than the value computed above then it may be considered that there is no statistical bias in the observed values with respect to their magnitudes as compared with that obtained by the rating curve.

The percentage differences have been taken as they are comparatively independent of the discharge volume and are approximately normally distributed about zero mean value for an unbiased curve.

3.4 Goodness of fit test

Due to changes in the flow regime, it is possible that long runs of positive and/or negative deviations are obtained at various stages. This may also be due to inappropriate fitting of the rating curve. This test is carried out for long runs of positive and negative deviations of the observed values from the stage-discharge curve. The test is designed to ensure a balanced fit in reference to the deviations over different stages. (Compare the graphical tests using stage/flow deviation scattergram and cumulative deviation plot of gaugings)

The test is based on the number of changes of sign in the series of deviations (observed value minus expected or computed value). First of all, the signs of deviations, positive or negative, in discharge measurements in ascending order of stage are recorded. Then starting from the second sign of the series, “0” or “1” is placed under sign if the sign agrees or does not agree respectively with the sign immediately preceding it. For example,

```
+ - + + + + - - - - + + + + -
  1 1 0 0 0 1 0 0 0 1 0 0 0 1
```

If there are N numbers in the original series, then there will be (*N – 1*) numbers in the derived series 11000100010001

If the observed values are regarded as arising from random fluctuations about the values estimated from the curve, the probability of a change in sign could be taken to be ½. However, this assumes that the estimated value is the median rather than the mean. If N is fairly large, a practical criterion may be obtained by assuming successive signs to be independent (i.e. by assuming that they arise only from random fluctuations), so that the number of “1”s (or “0”s) in the derived sequence of (*N – 1*) members may be judged as a binomial variable with parameters (*N – 1*) and ½.

From the above derived series, the actual number of changes of sign is noted. The expected number of changes of sign is computed by multiplying total possible numbers (i.e. *N –1*) with the probability of change of sign (i.e. ½). The statistical significance of the departure of the actual number of change of signs from the expected number is known by finding the “t” statistic as follows:

$$t = \frac{\left| N' - (N-1)\,p \right| - 0.5}{\sqrt{(N-1)\,pq}} \qquad (7)$$

where *N’* denotes the actual number changes of sign.

If the critical value of “t” statistic, for (*N – 1*) degrees of freedom, is more than that computed above then it can be considered to be having adequate goodness of fit. Otherwise, the results will indicate that there is significant bias in the fitted curve with respect to long runs of positive or negative deviations.