On the precision of sedimentation balance measurements

C. Kranenburg

Report no. 1 - 92

Hydromechanics Section
Department of Civil Engineering
Delft University of Technology
Delft, The Netherlands
1992
## CONTENTS

Abstract

1. Introduction .......................... 1
2. Precision of the distribution function $\tilde{f}$  
   2.1. Standard deviation .......................... 2
   2.2. Locally linear least-squares fit of $\hat{\gamma}$ 2
   2.3. Covariance of $\tilde{\gamma}$ and $\tilde{g}$ 3
   2.4. Number of data values and resolution 4
3. Simulation .......................... 6
4. Conclusions .......................... 8
Acknowledgements ......................... 9
References ................................ 9
Notation ................................ 9
Figures
Abstract

The settling velocities of fine sediments can be determined using a sedimentation balance essentially consisting of a settling tube at the bottom of which the weight of the sediment deposited is measured as a function of time. The distribution of settling velocities then is calculated from an equation, which is known as Oden’s formula. In this note the influence of random errors in the measured weight on the precision of the calculated distribution of settling velocities is analyzed. It is shown that the number of digitally sampled data values needed to calculate the distribution function increases with the ratio of time to sampling time interval, and that a compromise has be sought between precision and resolution.
1. **Introduction**

The distribution of settling velocities in a sample of particulate material can be determined using a sedimentation balance. This instrument essentially consists of a settling tube at the bottom of which the deposited sediment is weighed as a function of time. Several methods to introduce the sample into the settling tube exist. The method employed for samples of fine sediments (mud) is the following: after the sample has been added to the fluid (usually water) in the settling tube, it is mixed thoroughly over the total height of the fluid column. When the mixing is completed, the sediment particles start to settle and the time-dependent weight of the deposited sediment is measured (e.g., Kuijper et al., 1992).

The cumulative distribution function of the settling times follows from the measured weight. Let $\gamma(t)$ be the measured weight at time $t$ nondimensionalized by the measured weight for $t \to \infty$. At an arbitrary time $t$ all material having a settling time less than $t$ has settled. The derivative $d\gamma(t)/dt$ represents the particles with settling times greater than $t$. Of these particles a fractional amount $td\gamma(t)/dt$ has settled after the start at $t = 0$ of the measurement. At time $t$ the fractional amount $f(t)$ of deposited material with settling times less than $t$ therefore is

$$f(t) = \gamma(t) - t \frac{d\gamma(t)}{dt} \quad (1.1)$$

which expression is known as Oden's equation. The function $f(t)$ is the cumulative distribution of the settling times.

In practical applications $\gamma$ and $d\gamma/dt$ have to be estimated from measured data. Such estimates are denoted by an overbar so that

$$\bar{f}(t) = \bar{\gamma}(t) - t \frac{d\bar{\gamma}(t)}{dt} \quad (1.2)$$

---

1 The settling time is defined here as the height of the fluid column divided by the settling velocity. Settling times $t$ and settling velocities $w$ are related by $wt = h$, where $h$ is the height of the fluid column.
Some aspects of the accuracy of the method outlined are analysed in this note.

2. Precision of the distribution function $\bar{f}$

2.1 Standard deviation

Inaccuracies in the function $\bar{f}$ calculated using Eq. 1.2 are caused by systematic and random errors in the measured weight. Since systematic errors, such as those caused by temperature variations or evaporation, depend on the specific properties of the instrument used, only random errors in digitally sampled values of the weight are considered herein. The measured weight including a random error is denoted by $\gamma$. Possible random errors in the measured times $t$ are assumed to be negligible.

Denoting the derivative $d\gamma/dt$ by $\bar{g}$, the standard deviation of the measured function $\bar{f}$ as given by Eq. 1.2 then is given by

$$s^2_f = \left[ \frac{\partial f}{\partial \gamma} \right] s^2_\gamma + \left[ \frac{\partial f}{\partial \bar{g}} \right]^2 s^2_\bar{g} + 2 \frac{\partial f}{\partial \gamma} \frac{\partial f}{\partial \bar{g}} C_{\gamma \bar{g}}$$

(2.1)

where $s_z$ is the standard deviation of a quantity $z$, and $C_{\gamma \bar{g}}$ is the covariance of $\gamma$ and $\bar{g}$. The fractional weight $\gamma$ is defined as the ratio of weight to final weight. The latter weight can be determined with sufficient precision through averaging so that the error in $\gamma$ at time $t$ approximately equals the error in the weight at that time divided by the final weight. The random process is assumed to be stationary, which implies that the standard deviation $s_\gamma$ of the measured fractional weights $\gamma$ is independent of time.

Eq. 2.1 shows that if $s^2_f$ is to be less than some preselected value, the standard deviation $s_\bar{g}$ has to be $O(t^1)$ for large times (the covariance is shown in Section 2.3 to vanish). The calculation of the derivative $\bar{g}$ from measured data therefore requires special attention.

2.2 Locally linear least-squares fit of $\bar{g}$

One way to obtain precise estimates of $g$ is to locally approximate the function $\gamma(t)$
by a linear relation, thereby determining the two constants in such a relation by means of a least-squares fit.

Suppose that in a certain time interval the weight is sampled digitally according to

\[ \hat{y}_{i,j} = \tilde{y}(t_i + j\Delta t), \quad i = 1, 2, \ldots \quad j = \pm 1, \pm 2, \ldots \pm n \tag{2.2} \]

where \( t_i \) is the settling time at which \( \tilde{y} = f_i \) is to be calculated, \( n \) is an integer, and \( \Delta t \) a sampling time interval that does not depend on \( j \) (but may be made dependent on \( i \)).

A locally linear approximation \( \tilde{y} \) can be written as

\[ \tilde{y}(t) = \tilde{y}_i + \overline{g}_i (t - t_i), \quad |t - t_i| \leq n\Delta t \tag{2.3} \]

The sum \( Q \) of squared deviations is given by

\[ Q = \sum_{j=-n}^{n} (\hat{y}_{i,j} - \tilde{y}_i - \overline{g}_i j \Delta t)^2 \tag{2.4} \]

The least-squares fit is obtained by setting \( \partial Q / \partial \tilde{y}_i = 0 \) and \( \partial Q / \partial \overline{g}_i = 0 \), which gives\(^2\) (e.g., Bendat and Piersol, 1971)

\[ \tilde{y}_i = \frac{1}{2n+1} \sum_{j=-n}^{n} \hat{y}_{i,j} \quad (\text{mean value}) \tag{2.5} \]

\[ \overline{g}_i = \frac{3}{n(n+1)(2n+1)} \sum_{j=-n}^{n} j \hat{y}_{i,j} \]

These equations provide only estimates of \( \tilde{y}_i \) and \( \overline{g}_i \). Before calculating standard deviations

\(^2\) Note that \( \sum_{j=-n}^{n} j^2 = \frac{1}{3} n(n+1)(2n+1) \).
of these quantities, their covariance is considered.

2.3. Covariance of \( \tilde{\gamma} \) and \( \tilde{g} \)

Denoting a random error in a measured weight \( \gamma_{ij} \) by \( x_{ij} \) so that \( \tilde{\gamma}_{ij} = \gamma_{ij} + x_{ij} \), Eqs. 2.5 show that the covariance \( C_{\tilde{\gamma}\tilde{g}} \) is given by

\[
C_{\tilde{\gamma}\tilde{g}} = \frac{3}{n(n+1)(2n+1)^2 \Delta t} \left( \sum_{j=-n}^{n} x_{ij} \sum_{j=-n}^{n} jx_{ij} \right)
\]

(2.6)

where the pointed brackets denote ensemble averaging. Assuming that the errors \( x_{ij} \) are uncorrelated, this average becomes

\[
\left( \sum_{j=-n}^{n} x_{ij} \sum_{j=-n}^{n} jx_{ij} \right) = \sum_{j=-n}^{n} j\langle x_{ij}^2 \rangle
\]

(2.7)

For a stationary random process \( \langle x_{ij}^2 \rangle \) is independent of \( i + j \), and as a consequence the covariance vanishes (because \( \Sigma j \) vanishes). If the process is instationary, the covariance as given by Eq. 2.6 will be \( O(1/n) \) because

\[
\sum_{j=-n}^{n} j\langle x_{ij}^2 \rangle < n^2\langle x^2 \rangle_{\text{max}}
\]

(2.8)

Consequently, the covariance will then be small if \( n \gg 1 \).

2.4 Number of data values and resolution

The standard deviations of \( \tilde{\gamma}_i \) and \( \tilde{g}_i \) can be calculated as indicated by Eq. 2.1. Assuming again that the random errors are uncorrelated, one obtains
\begin{align*}
\sigma_i^2 &= \frac{1}{2n+1} \sigma_y^2 \\
\sigma_{\delta i}^2 &= \frac{3}{n(n+1)(2n+1)\Delta t^2} \sigma_y^2 \\
(2.9) \\

\text{Substituting these relationships in Eq. 2.1 and using the result of Section 2.3 give}
\begin{align*}
\sigma_j^2 &= \frac{1}{2n+1} \left[ 1 + \frac{3}{n(n+1)} \left( \frac{t}{\Delta t} \right)^2 \right] \sigma_y^2 \\
(2.10) \\

\text{Assuming a gaussian distribution of random errors in the calculated function } \bar{f} \text{, error bars can be constructed once } \sigma_j \text{ is known.}

\text{Defining a relative precision of the function } \bar{f} \text{ by}
\begin{align*}
\sigma_j &= ks_y \\
(2.11) \\

\text{where } k \text{ is a coefficient to be chosen, it is found that the first term in brackets in Eq. 2.10 is negligible, if}
\begin{align*}
\left( k \frac{t}{\Delta t} \right)^{4/3} &\gg 1 \\
(2.12) \\

\text{This condition will be satisfied in most cases (e.g., } k \geq 1 \text{ and } t/\Delta t \gg 1). \text{ Assuming furthermore that } n \text{ is large so that } n(n+1) = (2n+1)^2/4, \text{ Eqs. 2.10 and 2.11 give as an estimate of the number of digitally sampled data values}
\begin{align*}
2n+1 &\approx \left( \frac{12}{k^2} \right)^{1/3} \left( \frac{t}{\Delta t} \right)^{2/3} \\
(2.13)
\end{align*}
\end{align*}
This expression shows that, in order to obtain a certain precision (represented by the coefficient \( k \)), the number of data values has to increase with \( t/\Delta t \).

The time interval, \( \tau \), over which the averaging process takes place is given by \( \tau = 2n\Delta t \), or

\[
\frac{\tau}{t} = \left( \frac{12}{k^2} \right)^{\frac{1}{3}} \left( \frac{\Delta t}{t} \right)^{\frac{1}{3}} \quad (n \gg 1 \text{ and } \tau/t < \frac{1}{2})
\]  

(2.14)

The filtering time \( \tau \) increases with \( \Delta t/t \). It also increases as \( k \) decreases so that the function \( \tilde{f} \) is increasingly smoothed for increasing precision. Consequently, a practical compromise has to be sought between precision and resolution.

Some values of \( 2n + 1 \) and \( \tau/t \) are listed in Table 2.1.

<table>
<thead>
<tr>
<th>( \frac{t}{\Delta t} )</th>
<th>( 2n + 1 )</th>
<th>( \tau/t )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( k = 1 )</td>
<td>( k = 5 )</td>
</tr>
<tr>
<td>( 10^2 )</td>
<td>49</td>
<td>17</td>
</tr>
<tr>
<td>( 10^3 )</td>
<td>229</td>
<td>79</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>1063</td>
<td>363</td>
</tr>
<tr>
<td>( 10^5 )</td>
<td>4933</td>
<td>1687</td>
</tr>
</tbody>
</table>

Table 2.1. Values of \( 2n + 1 \) and \( \tau/t \).

3. Simulation

The approach discussed was checked by simulating the measuring procedure numerically. In particular the influence of the linear fit to the data is difficult to assess otherwise. The simulation was made by prescribing an analytical expression for the weight function \( \gamma(t) \), to which (small) random numbers were added that represented the measuring errors. These random numbers were taken from a uniform distribution with zero mean. The weight \( \gamma(t) \) was given by
\[
\gamma(t) = \frac{t/t_0}{[1 + (t/t_0)\beta]}^{-1}
\] (3.1)

in which \(t_0\) is a time constant, and \(\beta\) a coefficient representing the width of the distribution of settling times (the width decreases as \(\beta\) increases). Eq. 3.1 implies realistic behaviour for \(t \to 0\) and \(t \to \infty\). The distribution function \(f\) given by Eqs. 1.1 and 3.1 is

\[
f(t) = [\gamma(t)]^{\beta + 1}
\] (3.2)

Approximate values \(f_i\) of the distribution function were calculated using Eqs. 1.2, 2.5 and 2.13 for preselected values of the coefficient \(\beta\), time \(t_0\), relative precision \(k\) and the sampling time interval \(\Delta t\). This interval was prescribed according to an expression of the form

\[
\Delta t_i = \Delta t_1 \left(\frac{t_i}{t_1}\right)^\alpha 
\] (3.3)

where \(\alpha\) is a coefficient \((0 \leq \alpha \leq 1)\). The case where \(\alpha = 0\) represents a constant sampling time interval, which requires a large number of data values for large \(t\) (see Table 2.1). The sampling time intervals are proportional to the sampling times \(t_i\) for \(\alpha = 1\). The number of data values required and the parameter \(\tau/t\) then do not depend on \(t_i\).

Some results of the simulations made are shown in Figs. 1-7. The parameters that were varied are the precision \(k\) \((k = 5 \text{ and } k = 25)\) and those listed in Table 3.1. Parameters not varied were: \(t_0 = 200s\) and \(t_1 = 50s\). The same set of random numbers was used in all simulation runs.

The figures show the functions \(\gamma\) and \(f\), and the deviations \(\delta f = \bar{f} - f\) of the approximate distribution function from the exact function. The deviations \(\delta f\) were normalised by \(s_f = k s_\gamma\). Ideally, the quantity \(\delta f/s_f\) should be of order (plus or minus) one, and should not show a systematic dependence on the precision \(k\) or time.
Table 3.1. Values of \( \alpha \), \( \beta \) and \( \Delta t_1 \).

<table>
<thead>
<tr>
<th>Run no.</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \Delta t_1 ) (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{2} )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>( \frac{1}{2} )</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

As a whole, Figs. 1b-7b indeed seem to bear out the above arguments. The erratic nature of the deviations shown in these figures indicates that these deviations are caused mainly by the random errors in \( \gamma \), and that the influence of the linear approximation is minor. This conclusion becomes even more obvious when comparing Figs. 1b and 4b, 2b and 5b, and 3b and 6b. In each pair of figures the deviations for \( k = 25 \) are the same on the plotted scale, and those for \( k = 5 \) differ only little in most cases. Apparently the differences in the function \( \gamma(t) \) have little influence. Figs. 6b and 7b show some exceptions. The filtering times \( r \) for \( k = 5 \) (0.21t and 0.36t, respectively) are large in these cases so that, with respect to Fig. 3b, the influence of the linear approximation does become noticeable.

The small deviations in Fig. 7b for \( k = 25 \) are due to the rounding off upwards to one of the integer \( n \) given by Eq. 2.13.

4. **Conclusions**

The following conclusions can be drawn from this work:

- The locally linear least-squares fit is suited to accurately calculate the distribution of settling times from measured data.
- The required number of data values increases with precision and ratio of time to sampling time interval.
The resolution decreases with precision. As a consequence, a compromise between resolution and precision has to be sought.

Acknowledgements
This work was funded in part by the Commission of the European Communities, Directorate General for Science, Research and Development under MAST Contract no. 0035 (G6 Morphodynamics research programme). The writer is indebted to Mr. A.M. den Toom who carried out the numerical simulations.

References

Notation
\[ C_{yz} \] covariance of quantities \( y \) and \( z \)
\[ f \] cumulative distribution function of settling times
\[ g = \frac{dy}{dt} \] derivative of weight function
\[ h \] height of fluid column
\[ i,j \] integers
\[ k = \frac{s_j}{s_y} \] relative precision
\[ n \] integer
\[ Q \] sum of squared deviations
\[ s_z \] standard deviation of quantity \( z \)
\[ t \] settling time
\[ t_0 \] time constant
\[ w \] fall velocity
\[ x \] random error in measured weight
\[ \alpha, \beta \] coefficients
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>weight divided by final weight</td>
</tr>
<tr>
<td>$\delta f$</td>
<td>$\bar{f} - f$, error in function $f$</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>sampling time interval</td>
</tr>
<tr>
<td>$\tau$</td>
<td>$2n\Delta t$, filtering time interval</td>
</tr>
</tbody>
</table>

- approximate value
- measured value
- linear approximation
Fig. 1a. Weight $\gamma$ and distribution function $f$, run no. 1.

Fig. 1b. Errors in distribution function $f$, run no. 1.
Fig. 2a. Weight $\gamma$ and distribution function $f$, run no. 2.

Fig. 2b. Errors in distribution function $f$, run no. 2.
Fig. 3a. Weight $\gamma$ and distribution function $f$, run no. 3.

Fig. 3b. Errors in distribution function $f$, run no. 3.
Fig. 4a. Weight $\gamma$ and distribution function $f$, run no. 4.

Fig. 4b. Errors in distribution function $f$, run no. 4.
Fig. 5a. Weight $\gamma$ and distribution function $f$, run no. 5.

Fig. 5b. Errors in distribution function $f$, run no. 5.
Fig. 6a. Weight $\gamma$ and distribution function $f$, run no. 6.

Fig. 6b. Errors in distribution function $f$, run no. 6.
Fig. 7a. Weight $\gamma$ and distribution function $f$, run no. 7.

Fig. 7b. Errors in distribution function $f$, run no. 7.