

Supplement of Hydrol. Earth Syst. Sci., 23, 669–690, 2019
<https://doi.org/10.5194/hess-23-669-2019-supplement>
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.



Supplement of

A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry

Tim Busker et al.

Correspondence to: Tim Busker (tbusker@live.nl) and Ad de Roo (ad.de-roo@ec.europa.eu)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

Variance of lake level and volume estimates

In this supplement, we derive the variance equations of lake level and volume estimators. Section S1 defines the level-area linear model. Section S2 reports the equations of least-squares parameter estimates and related properties. In Sect. S3, we derive the variance equations for the lake level estimate (used to compute prediction intervals around level estimates) and for its expected value (used to compute confidence intervals around the fitted area-level regression line). Finally, in Sect. S4 and S5, we derive the variance equations for the lake volume and volume difference estimates.

S1 Level-area linear model

We define the following stochastic linear model of lake water level:

$$h_i = \beta + \alpha A_i + \varepsilon_i \quad (\text{S1})$$

where h_i and A_i are observed lake water level (m) and area (km^2) at time step i ; α and β are the slope (m km^{-2}) and intercept (m) parameters; and the residuals $\{\varepsilon_i | i = 1, \dots, n\}$ (m) are uncorrelated and zero-mean normal distributed random variables with variance σ^2 :

$$\varepsilon_i \sim N(0, \sigma^2) \quad (\text{S2})$$

S2 Least squares parameter estimates

Least squares estimates of α and β are obtained by minimising the sum of squared residuals:

$$\hat{\alpha}, \hat{\beta} = \arg \min_{\alpha, \beta} \sum_{i=1}^n (h_i - \beta - \alpha A_i)^2 \quad (\text{S3})$$

The obtained estimates are

$$\hat{\alpha} = \frac{\text{cov}[h, A]}{\text{var}[A]} \quad (\text{S4})$$

and

$$\hat{\beta} = \bar{h} - \bar{A}\hat{\alpha} \quad (\text{S5})$$

where var and cov are the variance and covariance operators, and $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. It can be shown that the estimators $\hat{\alpha}$ and $\hat{\beta}$ are unbiased, i.e. $\text{E}[\hat{\alpha}] = \alpha$ and $\text{E}[\hat{\beta}] = \beta$, where E is the expected value operator. The residuals $\hat{\varepsilon}_i$ resulting from the above least squares parameter estimates have the following properties:

$$\sum_{i=1}^n \hat{\varepsilon}_i = 0 \quad (\text{S6})$$

$$\sum_{i=1}^n A_i \hat{\varepsilon}_i = 0 \quad (\text{S7})$$

The variances of the parameter estimates are

$$\text{var}[\hat{\alpha}] = \frac{\sigma^2}{n \text{var}[A]} \quad (\text{S8})$$

and

$$\text{var}[\hat{\beta}] = \bar{A}^2 \text{var}[\hat{\alpha}] \quad (\text{S9})$$

The covariance of $\hat{\alpha}$ and $\hat{\beta}$ is

$$\text{cov} [\hat{\alpha}, \hat{\beta}] = -\bar{A} \text{var} [\hat{\alpha}] \quad (\text{S10})$$

The proofs of Eq. (S4-S10) are not shown here for the sake of brevity.

S3 Variance of lake level estimates

The variance of the lake level estimate can be used to compute the uncertainty bounds for the lake level predictions obtained by applying the linear model (S1). These bounds are referred to as prediction intervals (PI) in the manuscript. The variance can be expressed as the variance of the linear combination of random variables $\hat{\alpha}$, $\hat{\beta}$ and ε_i (lake area A_i is treated as a deterministic variable):

$$\text{var} [\hat{h}_i] = \text{var} [\hat{\beta}] + A_i^2 \text{var} [\hat{\alpha}] + \text{var} [\hat{\varepsilon}_i] + 2A_i \text{cov} [\hat{\alpha}, \hat{\beta}] + 2 \text{cov} [\hat{\alpha}, \hat{\varepsilon}_i] + 2 \text{cov} [\hat{\beta}, \hat{\varepsilon}_i] \quad (\text{S11})$$

It can be proved that $\text{cov} [\hat{\alpha}, \hat{\varepsilon}_i] = \text{cov} [\hat{\beta}, \hat{\varepsilon}_i] = 0$. The proof, which follows from Eq. (S6, S7), is not shown. Substituting Eq. (S9, S10) into (S11), we obtain

$$\text{var} [\hat{h}_i] = \text{var} [\hat{\alpha}] \left(\bar{A}^2 + A_i^2 - 2A_i\bar{A} \right) + \sigma^2 \quad (\text{S12})$$

which can be rewritten by substituting Eq. (S8) and $\bar{A}^2 + A_i^2 - 2A_i\bar{A} = \text{var} [A] + (A_i - \bar{A})^2$:

$$\text{var} [\hat{h}_i] = \sigma^2 \left(1 + \frac{1}{n} + \frac{(A_i - \bar{A})^2}{n \text{var} [A]} \right) \quad (\text{S13})$$

which is a convex function of A_i minimised at $A_i = \bar{A}$.

The variance of the expected value of the lake level ($\text{E} [\hat{h}_i] = \hat{\beta} + \hat{\alpha}A_i$) are used to compute confidence intervals (CI) around the area-level regression line:

$$\text{var} \left[\text{E} [\hat{h}_i] \right] = \text{var} [\hat{\beta}] + A_i^2 \text{var} [\hat{\alpha}] + 2A_i \text{cov} [\hat{\alpha}, \hat{\beta}] = \sigma^2 \left(\frac{1}{n} + \frac{(A_i - \bar{A})^2}{n \text{var} [A]} \right) \quad (\text{S14})$$

S4 Variance of lake volume estimates

Given the lake volume estimate equation

$$\hat{V}_i = \frac{A_i}{2} (\hat{\alpha}A_i + \hat{\varepsilon}_i) \quad (\text{S15})$$

and the assumptions used in Sect. S3, the variance of the lake volume estimate is

$$\text{var} [\hat{V}_i] = \frac{A_i^2}{4} (A_i^2 \text{var} [\hat{\alpha}] + \sigma^2) \quad (\text{S16})$$

S5 Variance of lake volume difference estimates

Given the reference lake area A_0 , we compute the change with respect to the corresponding volume V_0 as

$$\Delta \hat{V}_i = \hat{V}_i - \hat{V}_0 = \frac{A_i^2 - A_0^2}{2} \hat{\alpha} + \frac{A_i}{2} \hat{\varepsilon}_i - \frac{A_0}{2} \hat{\varepsilon}_0 \quad (\text{S17})$$

from which we can derive the variance of the volume difference. Recalling that $\text{cov} [\hat{\varepsilon}_i, \hat{\varepsilon}_j | i \neq j] = 0$ and $\text{cov} [\hat{\varepsilon}_i, \hat{\alpha}] = 0$, we get the following:

$$\text{var} [\Delta \hat{V}_i] = \frac{(A_i^2 - A_0^2)^2}{4} \text{var} [\hat{\alpha}] + \frac{A_i^2 + A_0^2}{4} \sigma^2 \quad (\text{S18})$$