

**Inversion Methods for Determining  
Tsunami Source Amplitudes  
from DART Buoy Data**

Don Percival

Applied Physics Laboratory  
University of Washington, Seattle

NOAA-sponsored collaborative effort

overheads for talk available at

<http://faculty.washington.edu/dbp/talks.html>

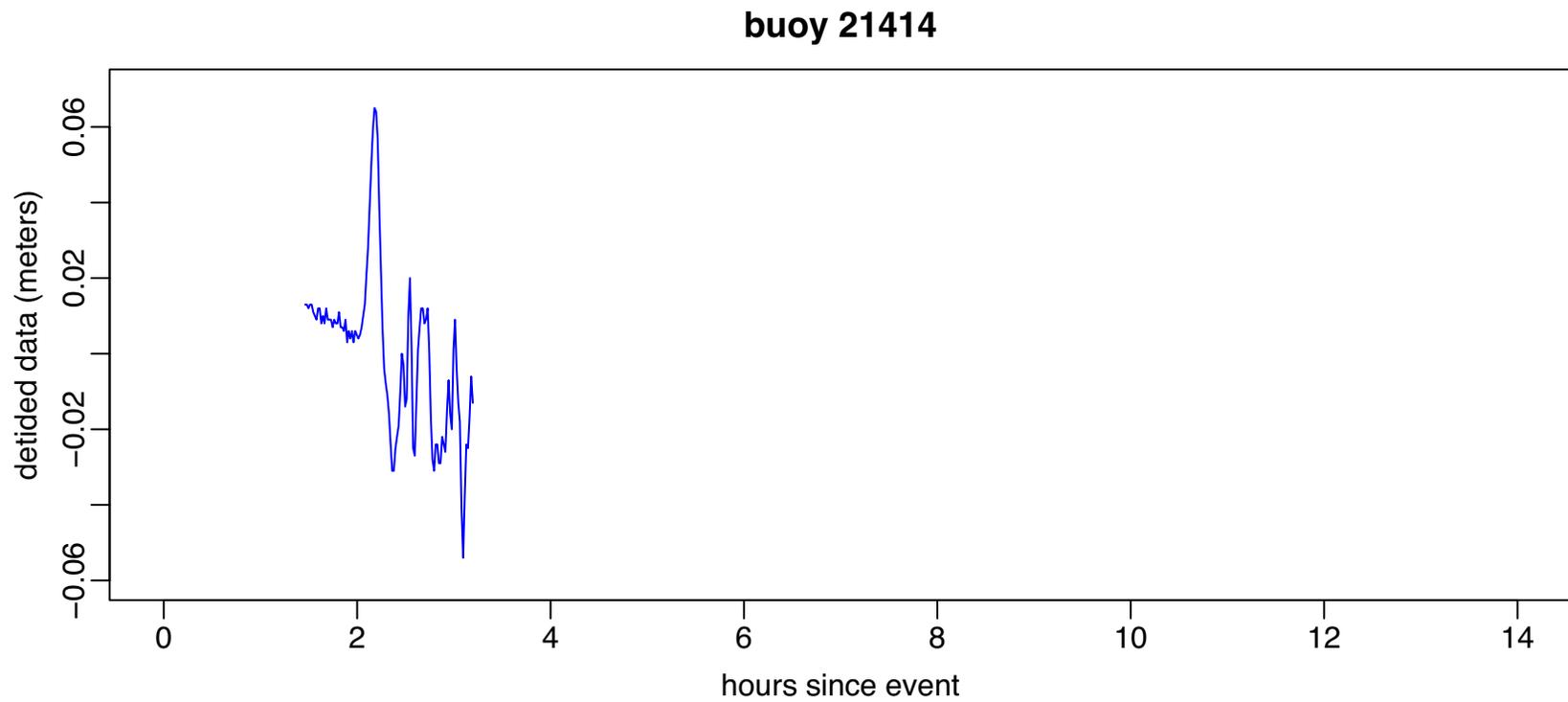
## Overview: I

- scientific problem: given data from DART buoys and models for unit magnitude earthquakes from various tsunami source locations, determine actual magnitudes (slips) and location(s) of actual earthquake
- will describe elements of basic inversion algorithm
- start with detided DART buoy data, noting need for detrending
- look at model for single buoy, noting need for interpolation
- introduce least squares criterion by looking at estimation of slip for single source model based upon data from one buoy
- consider assessing statistical variability in estimated slip
- look at effect of using varying amounts of buoy data
- considering adding data from a second DART buoy

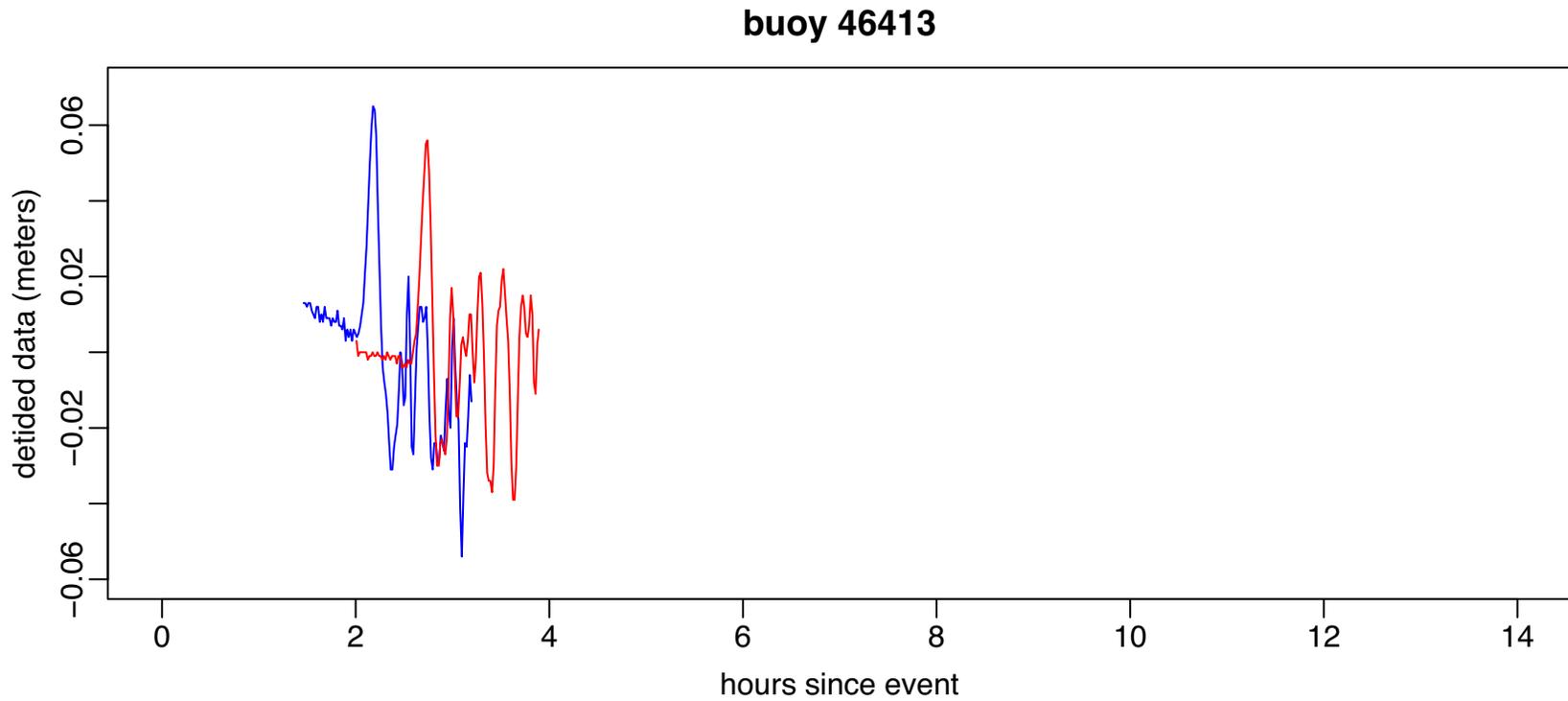
## Overview: II

- look at using more than one source
- discussion of various ‘bells and whistles’ currently implemented
  - imposing constraints on slips
  - allowing shifting and stretching of source models
- discussion of work in progress
  - use of statistical tests to select sources
- demo of R implementation of algorithm (if time permits)
- will motivate basic ideas behind algorithm using the Kuril Island event of Nov 2006

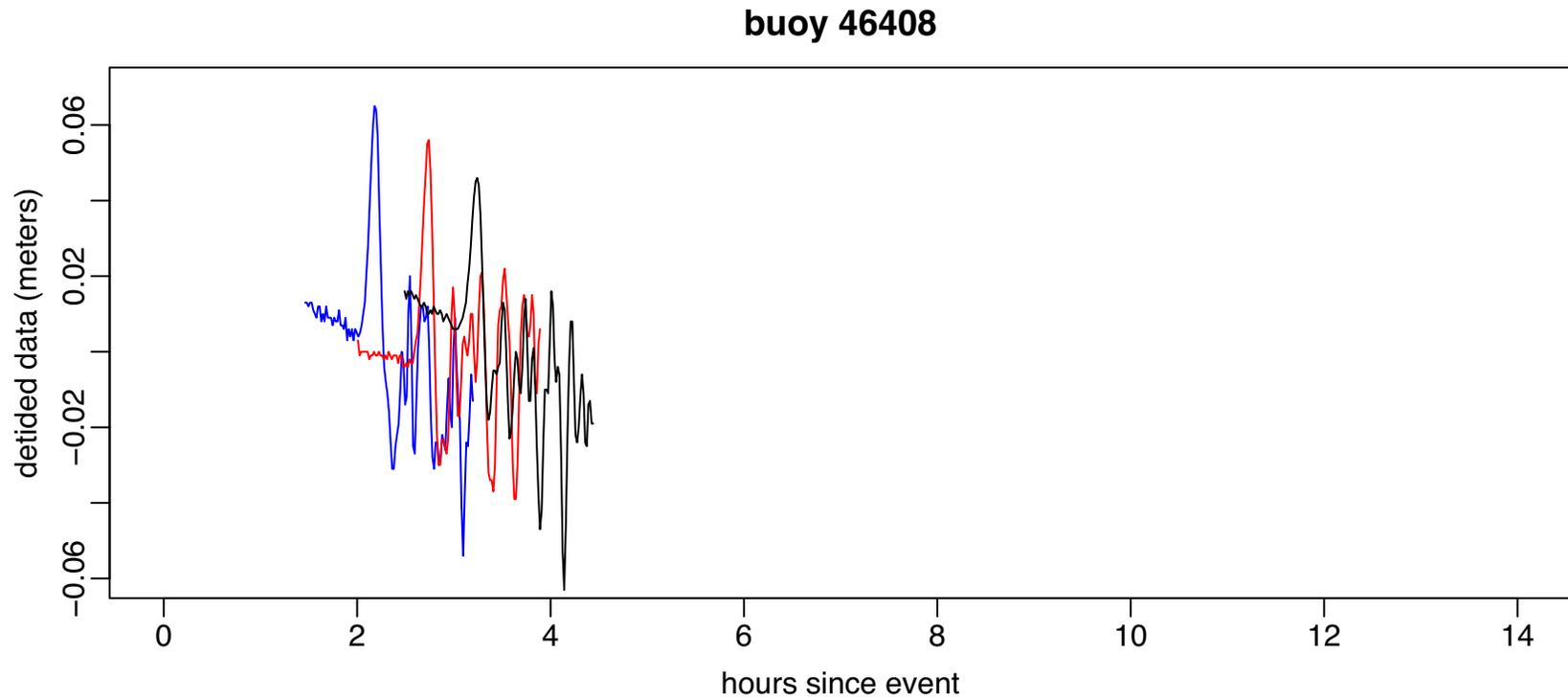
# Detided DART Buoy Data for Kuril Island Event



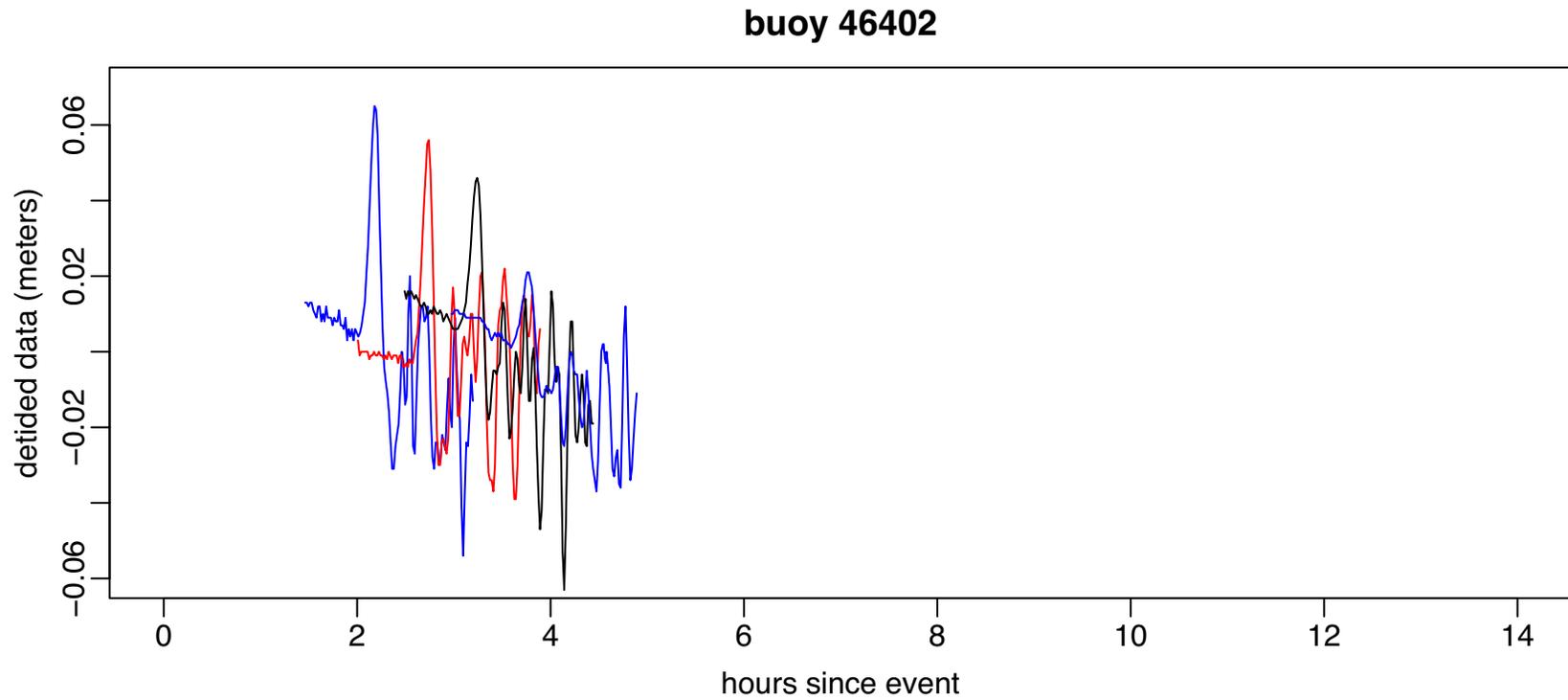
# Detided DART Buoy Data for Kuril Island Event



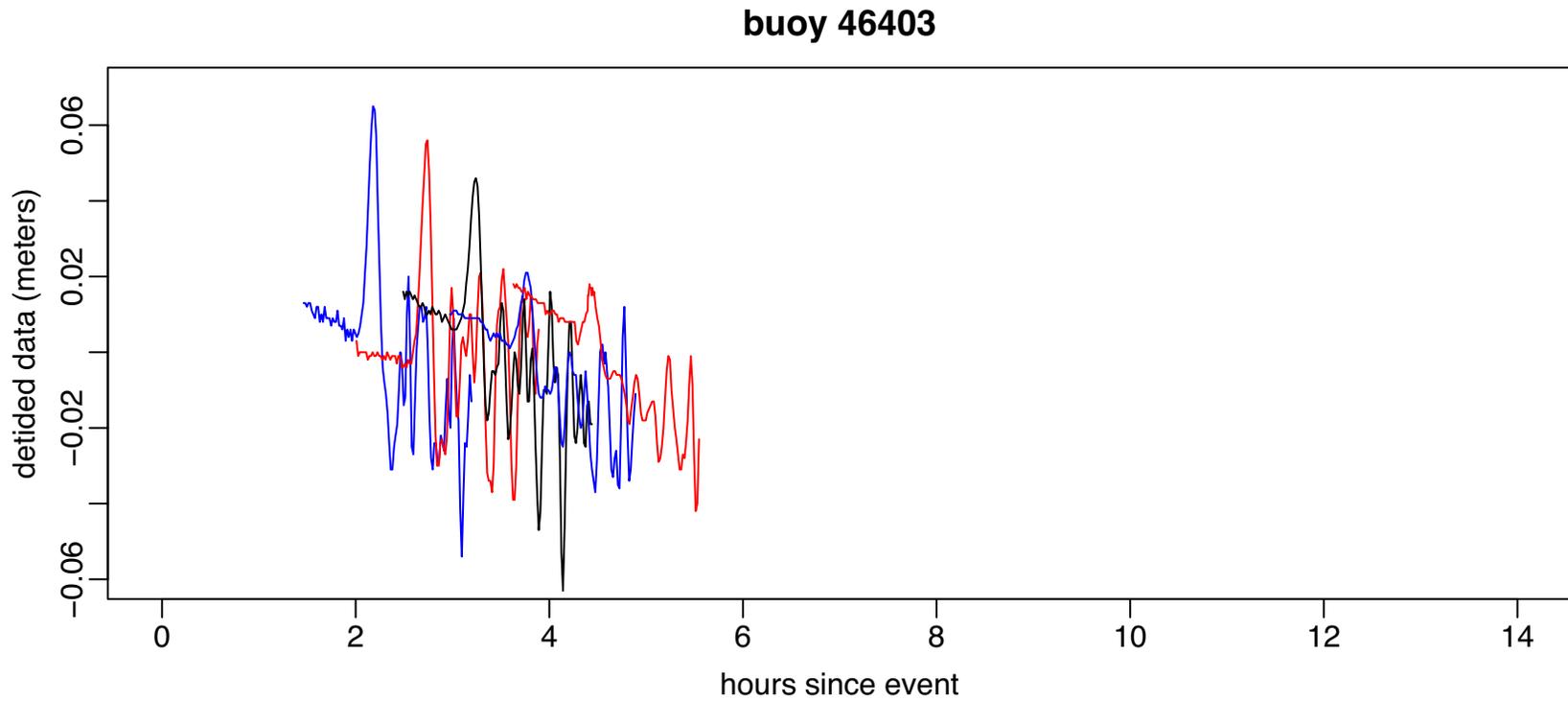
# Detided DART Buoy Data for Kuril Island Event



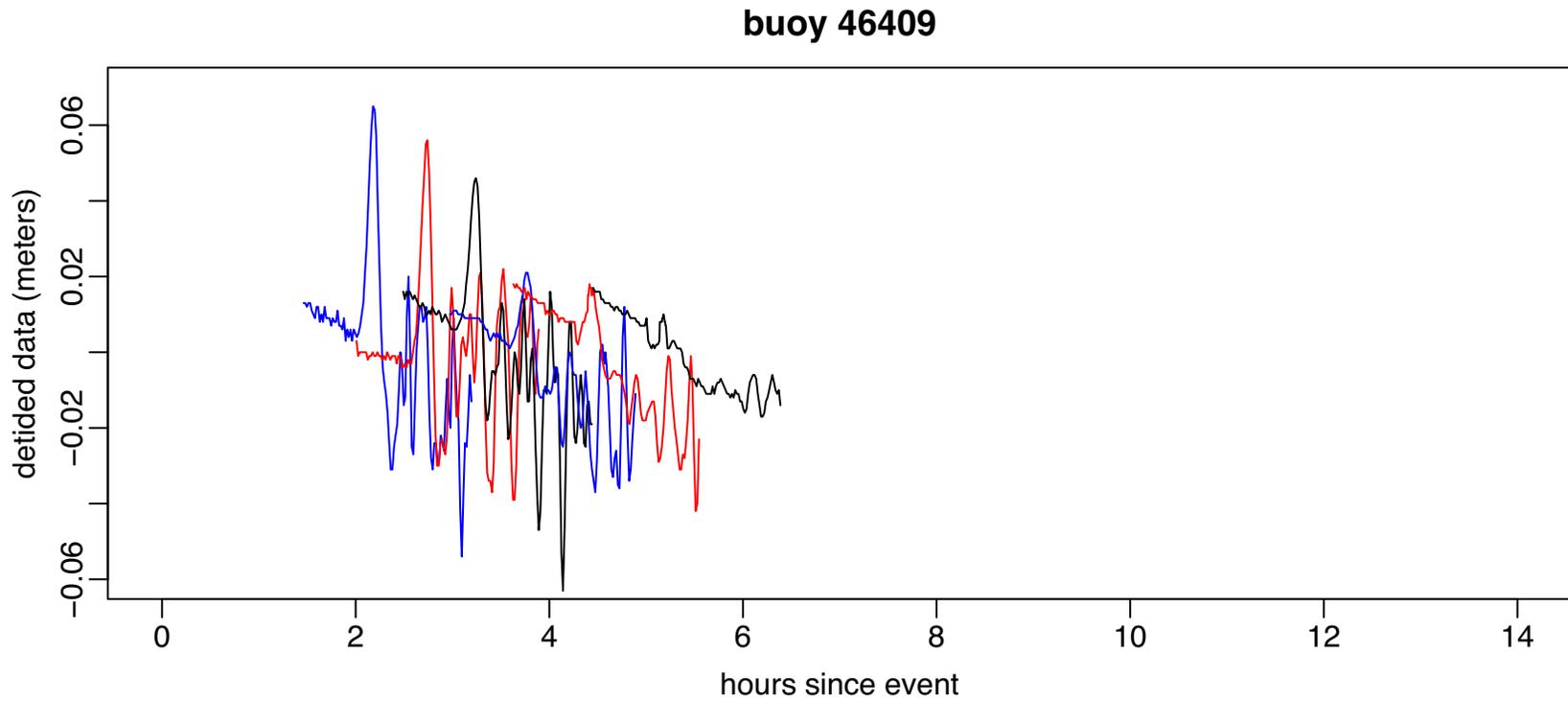
# Detided DART Buoy Data for Kuril Island Event



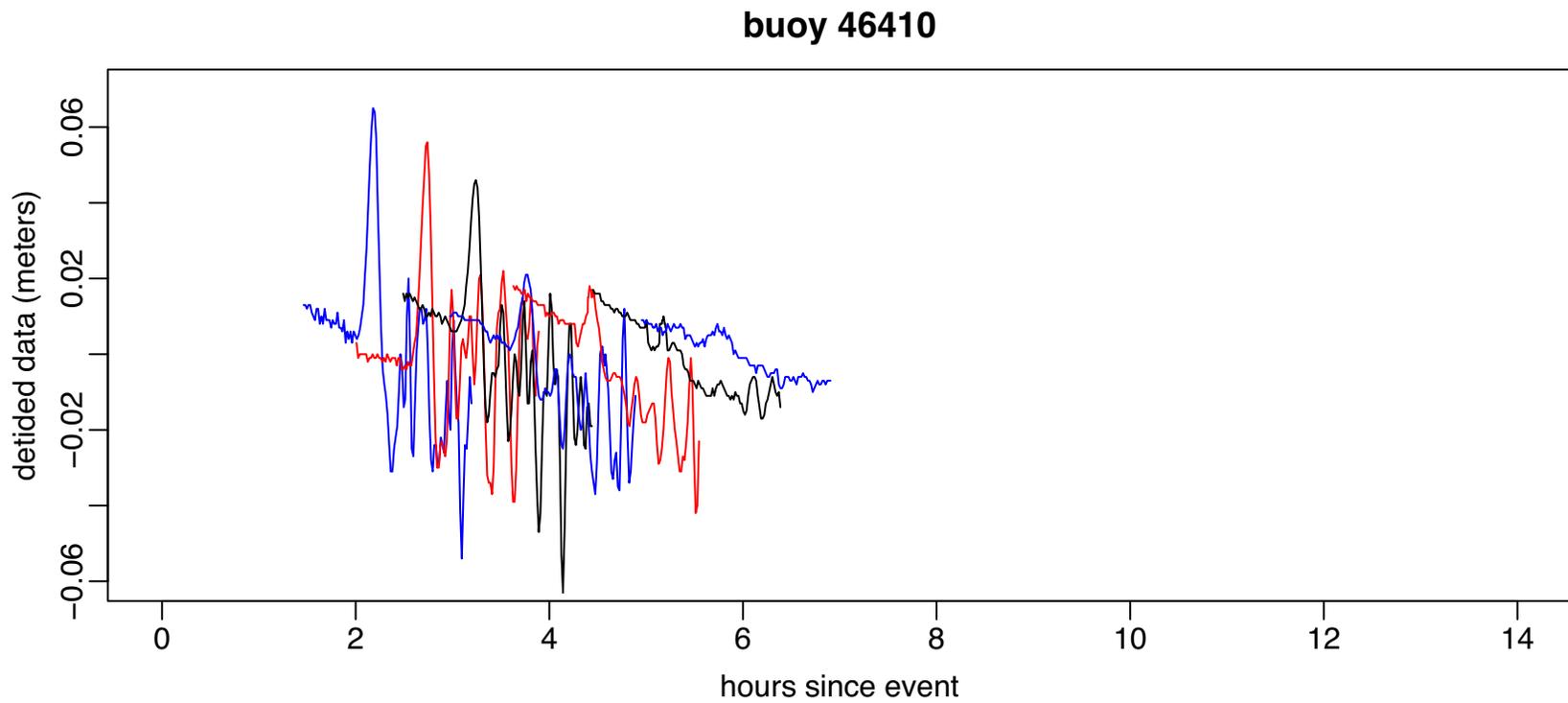
# Detided DART Buoy Data for Kuril Island Event



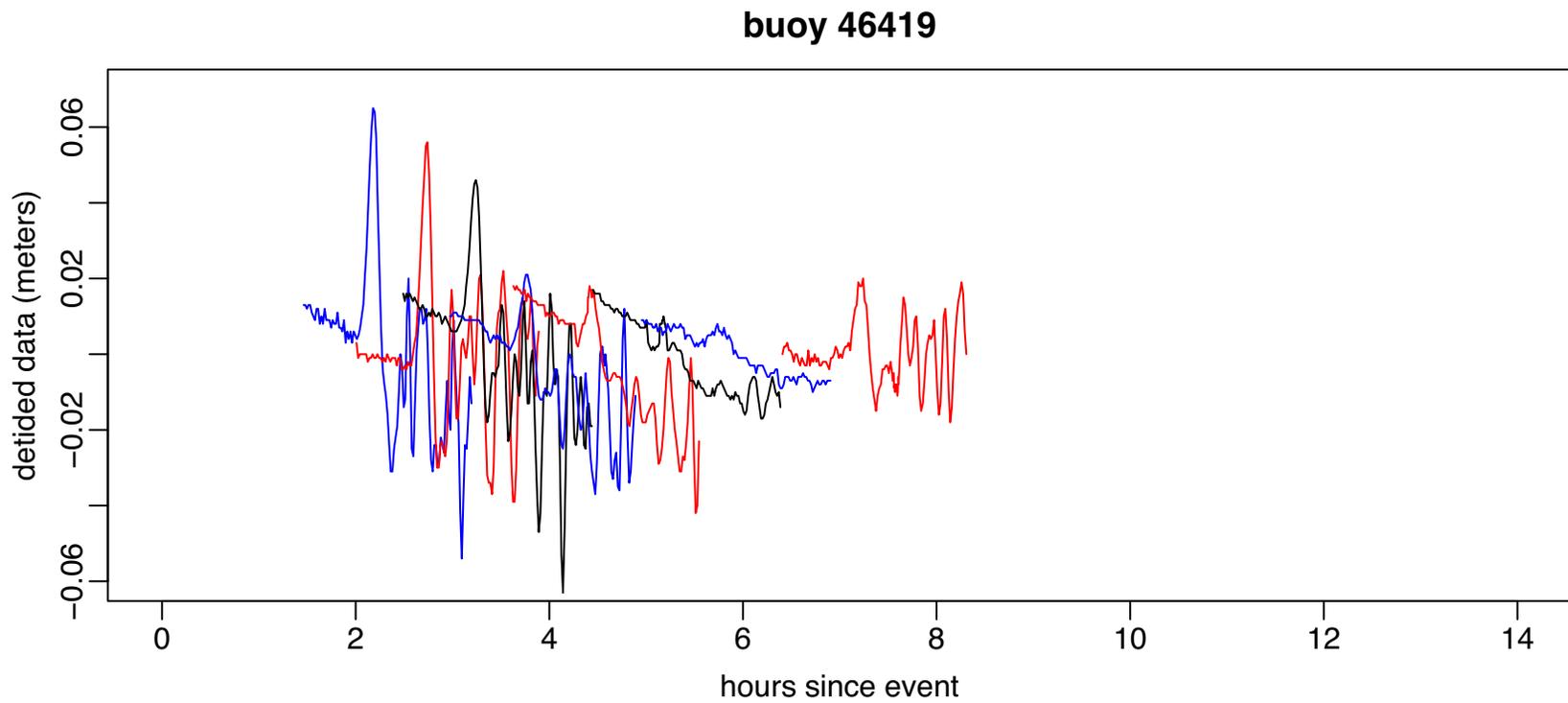
# Detided DART Buoy Data for Kuril Island Event



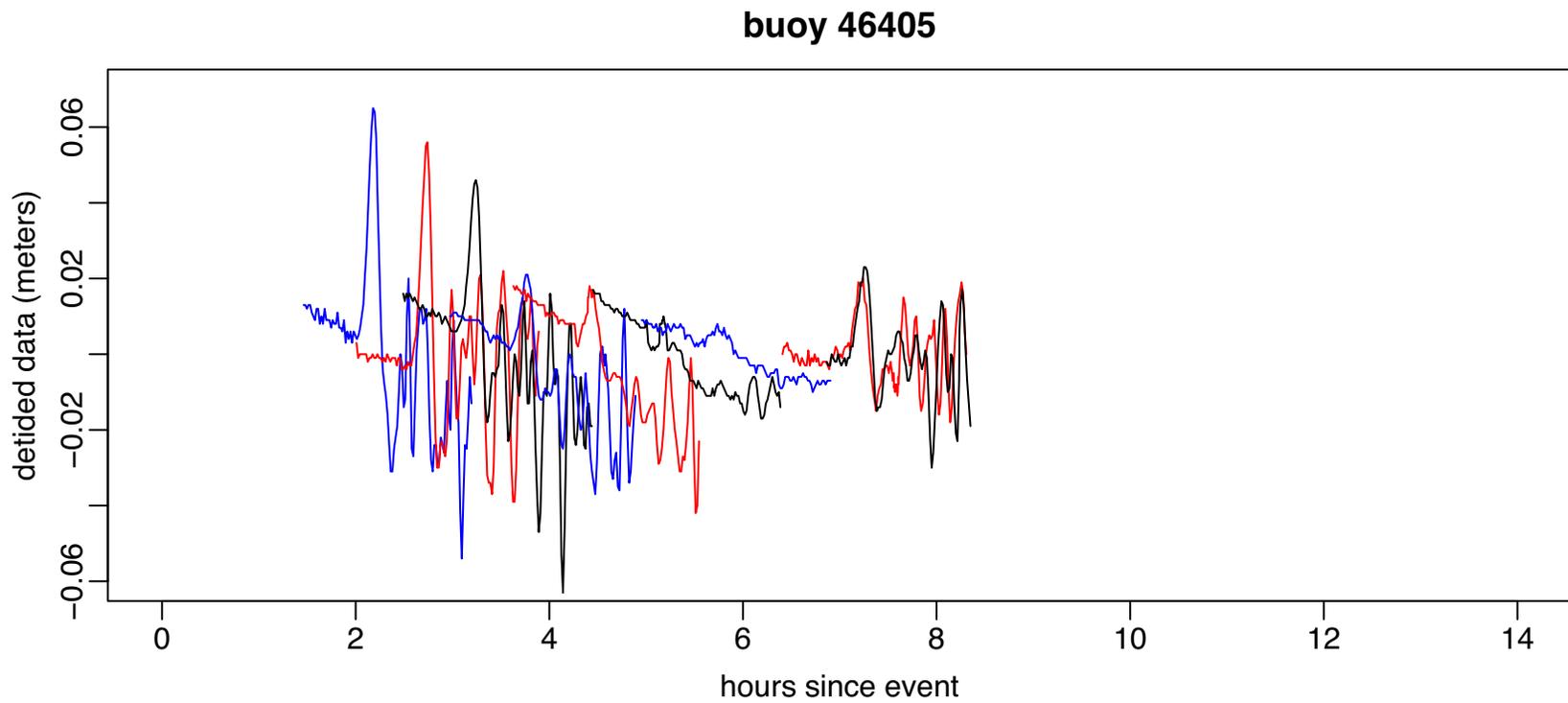
# Detided DART Buoy Data for Kuril Island Event



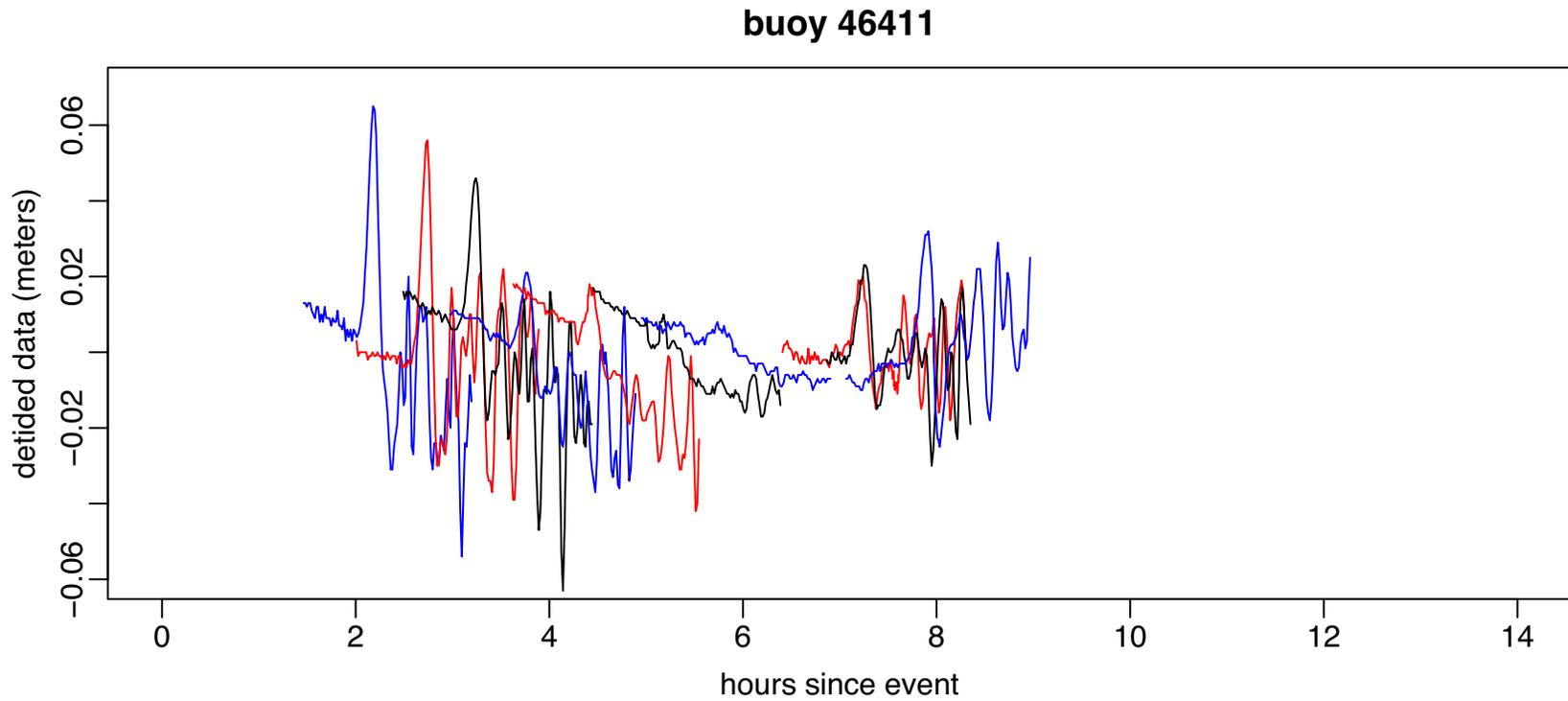
# Detided DART Buoy Data for Kuril Island Event



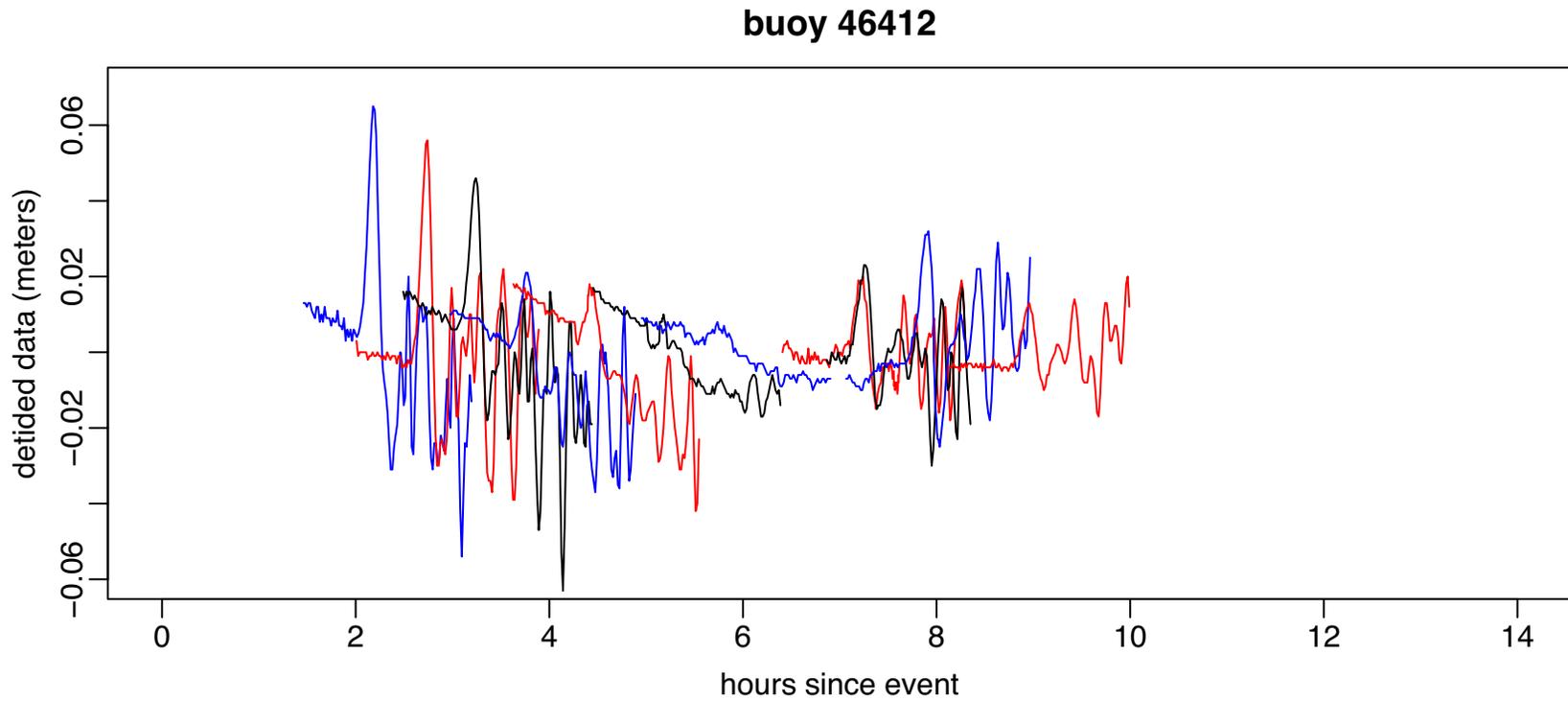
# Detided DART Buoy Data for Kuril Island Event



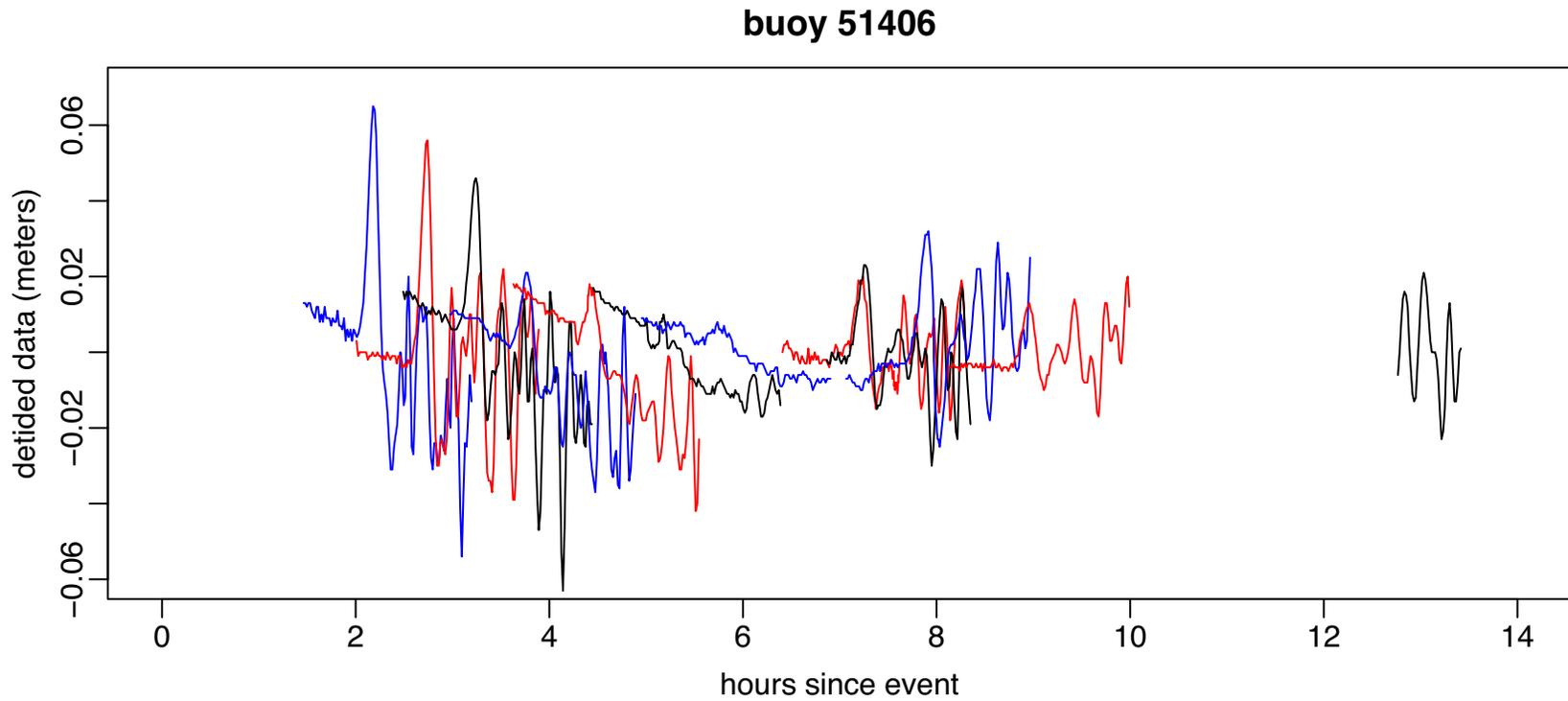
# Detided DART Buoy Data for Kuril Island Event



# Detided DART Buoy Data for Kuril Island Event



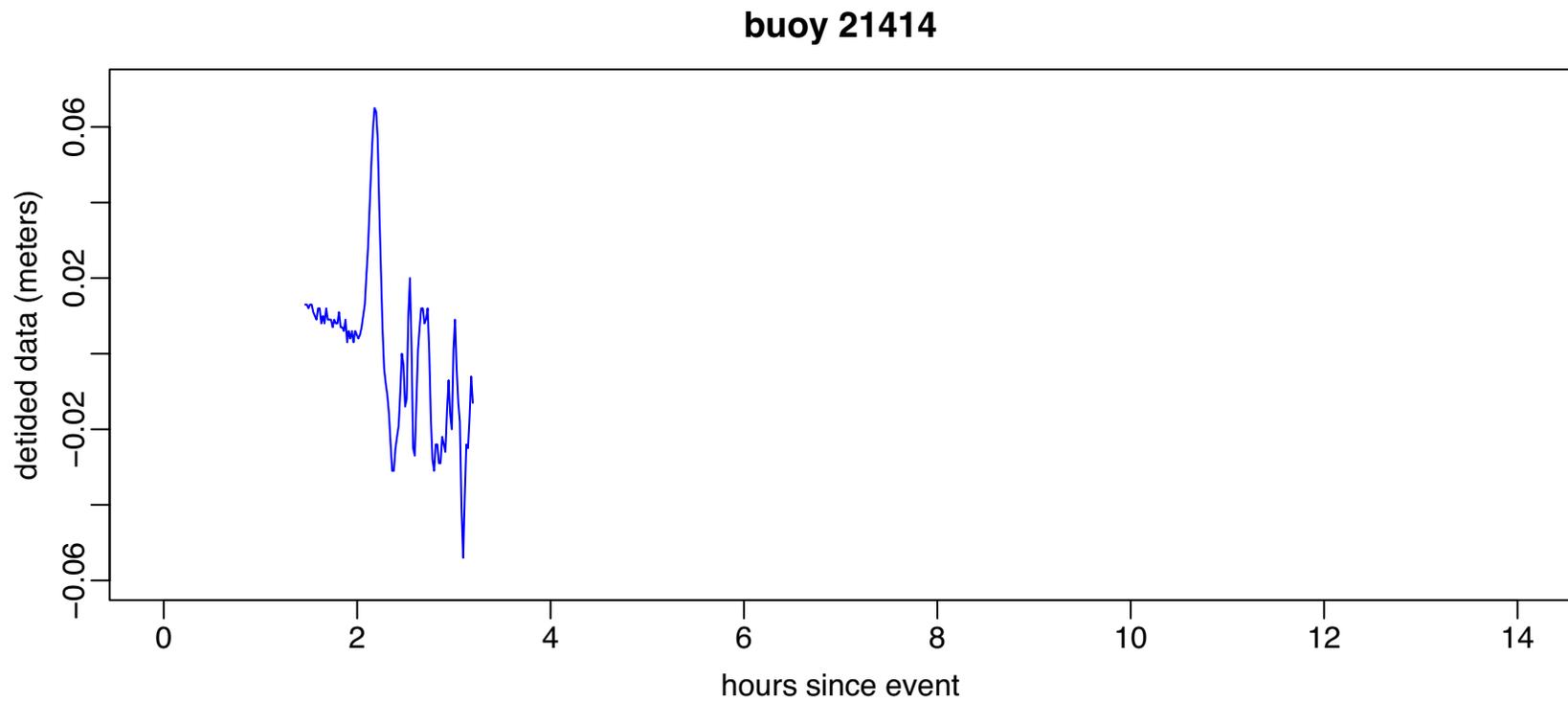
# Detided DART Buoy Data for Kuril Island Event



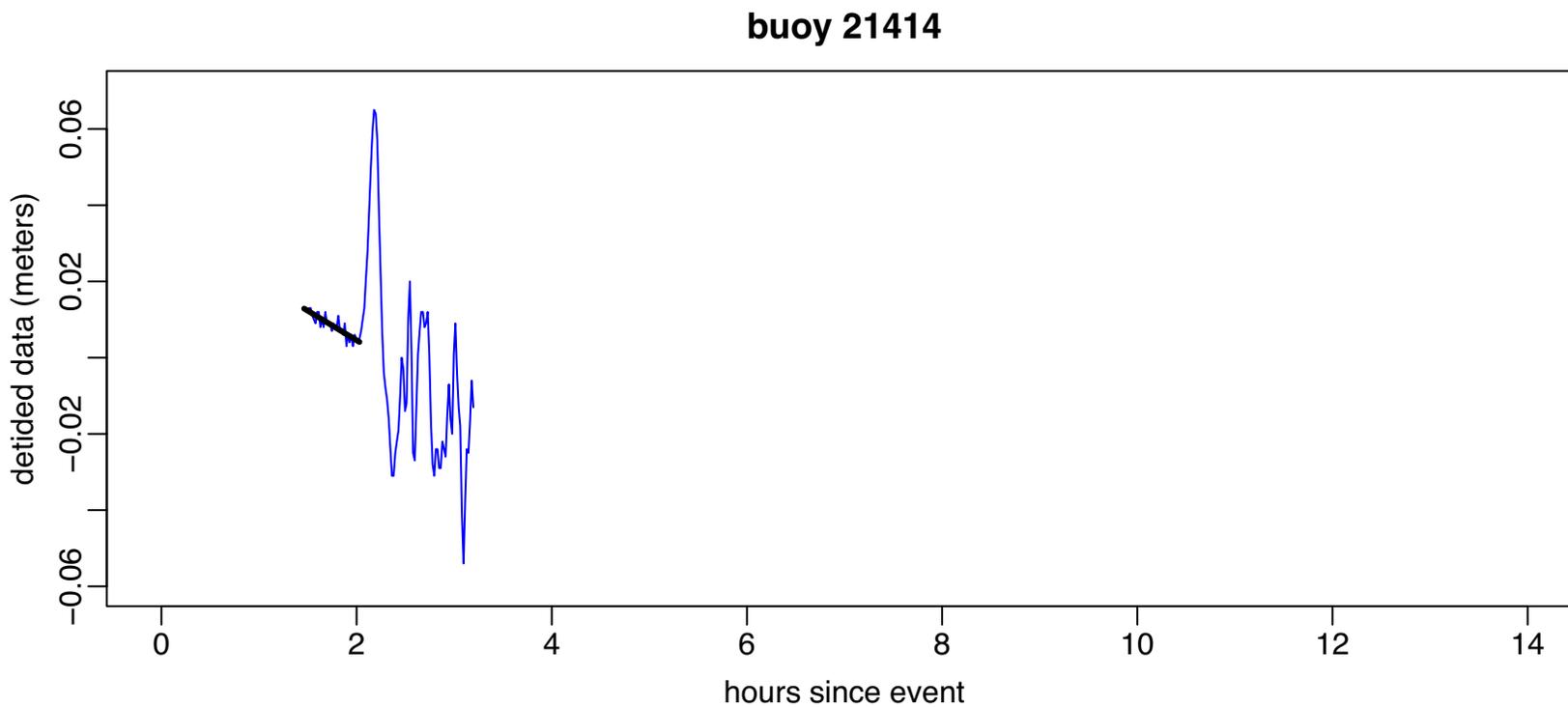
# Detrending

- inversion procedure assumes data have been successfully detided, which is not the case here
- data subjected to simple detrending procedure
  - identify region before start of first wave
  - fit line to this data using least squares procedure
  - extend fitted line through all the data
  - subtract extended line from data, yielding detrended data
- detrended data used as input to inversion procedure

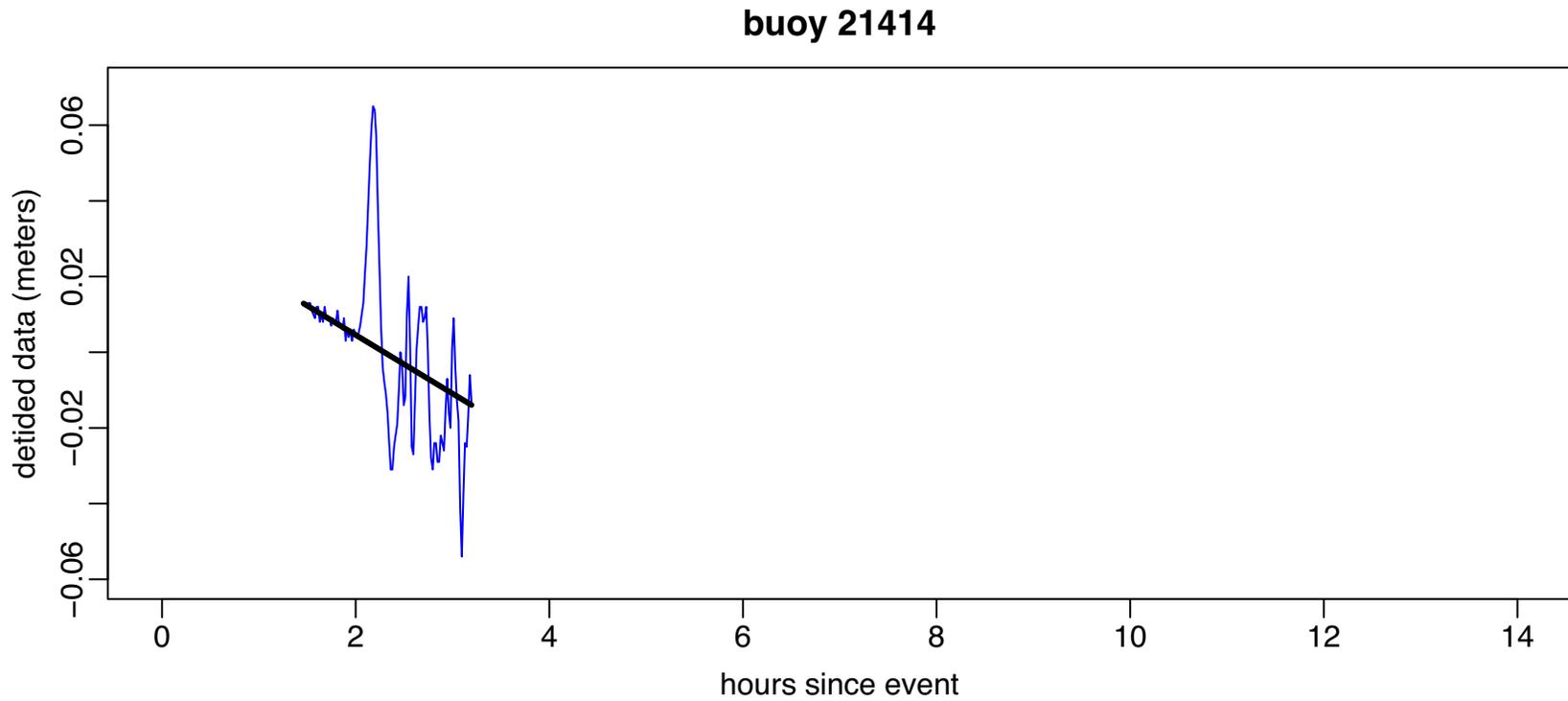
# Detided DART Buoy Data for Kuril Island Event



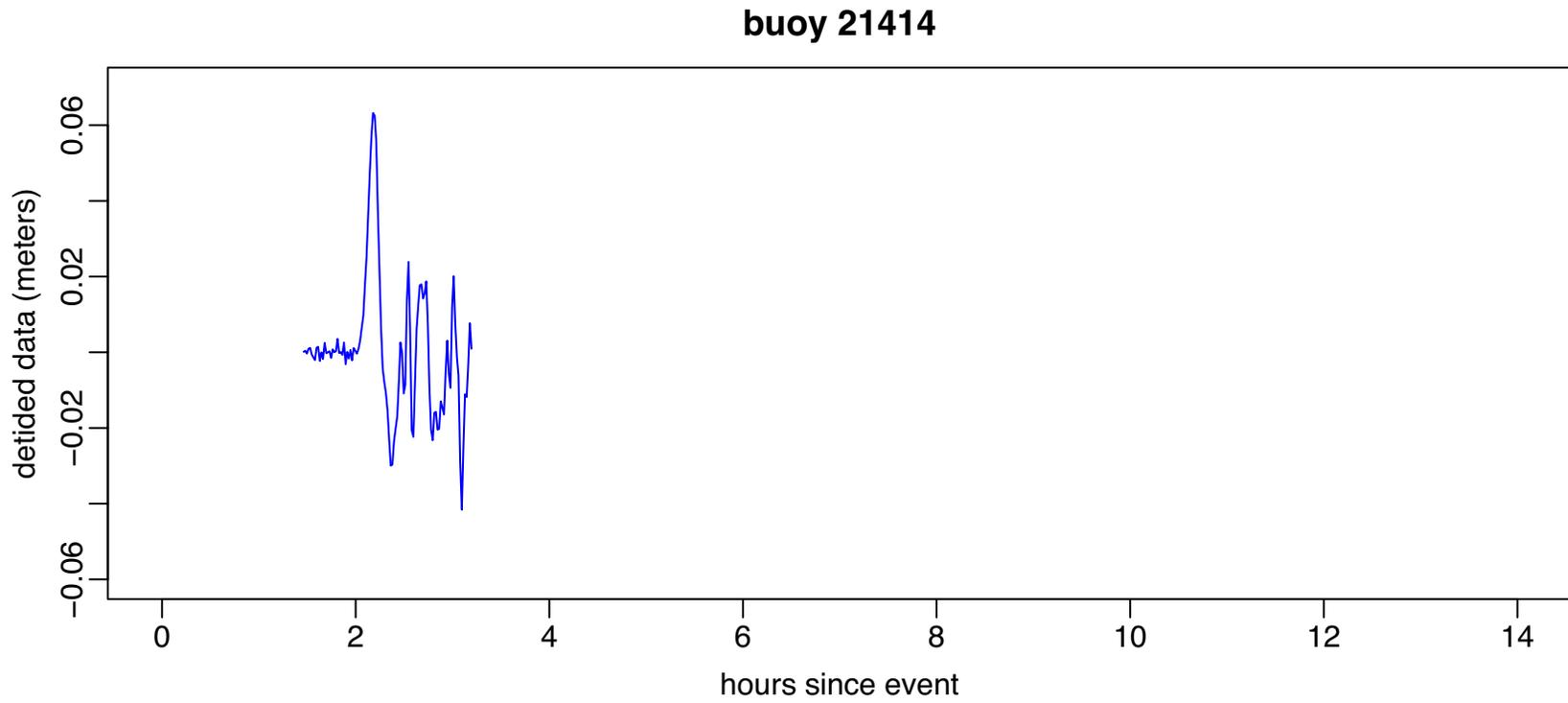
# Line Fitted to Bouy Data Before First Wave



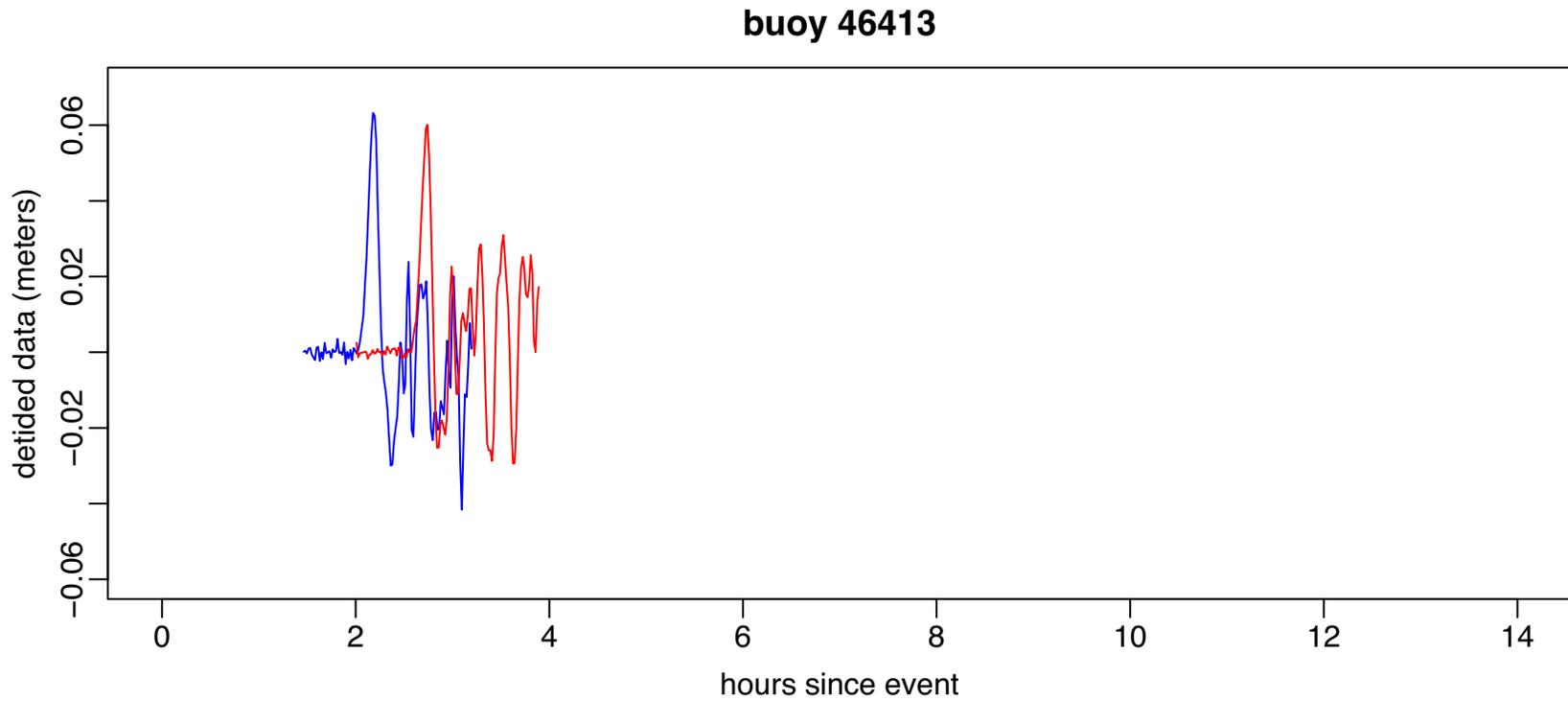
# Extending Fitted Line Through All Data



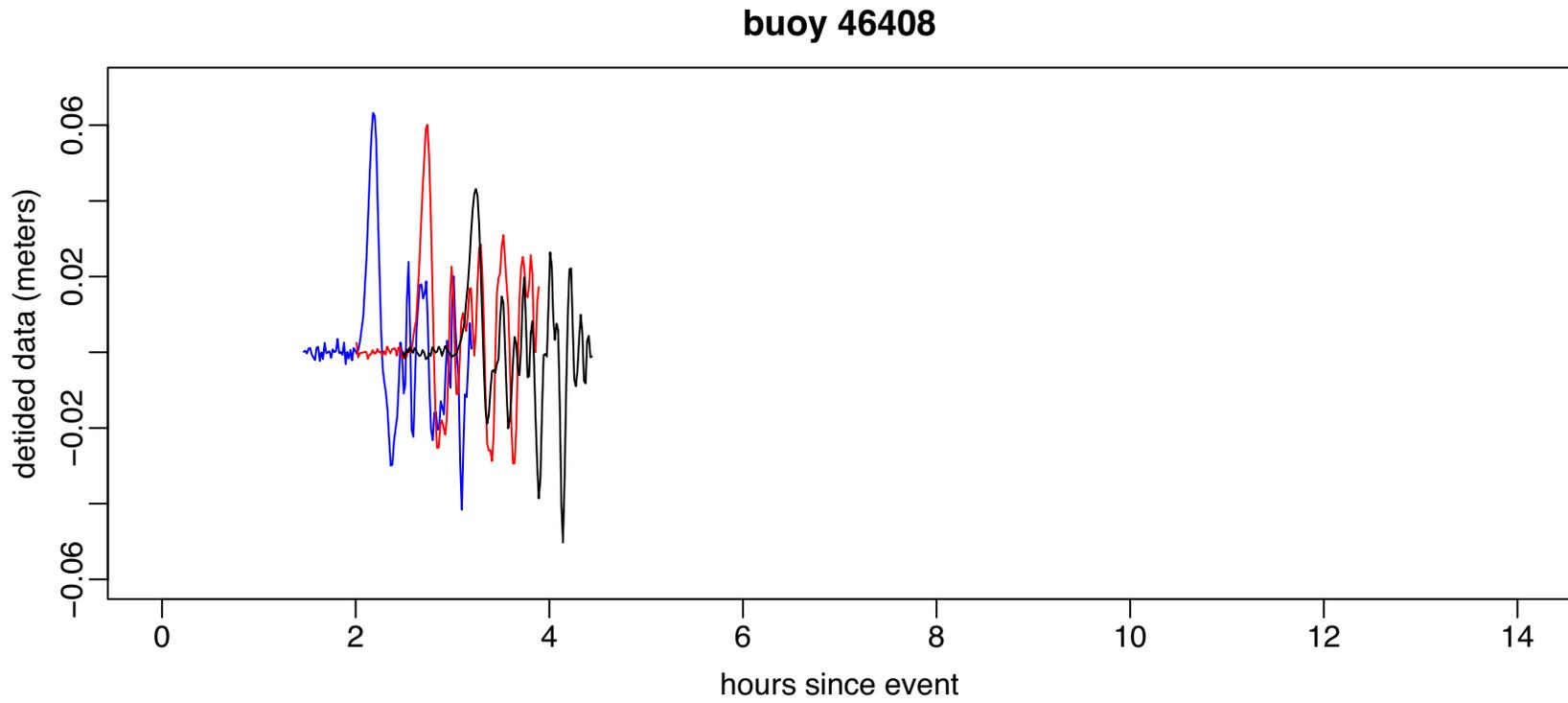
# Detrended DART Buoy Data for Kuril Island Event



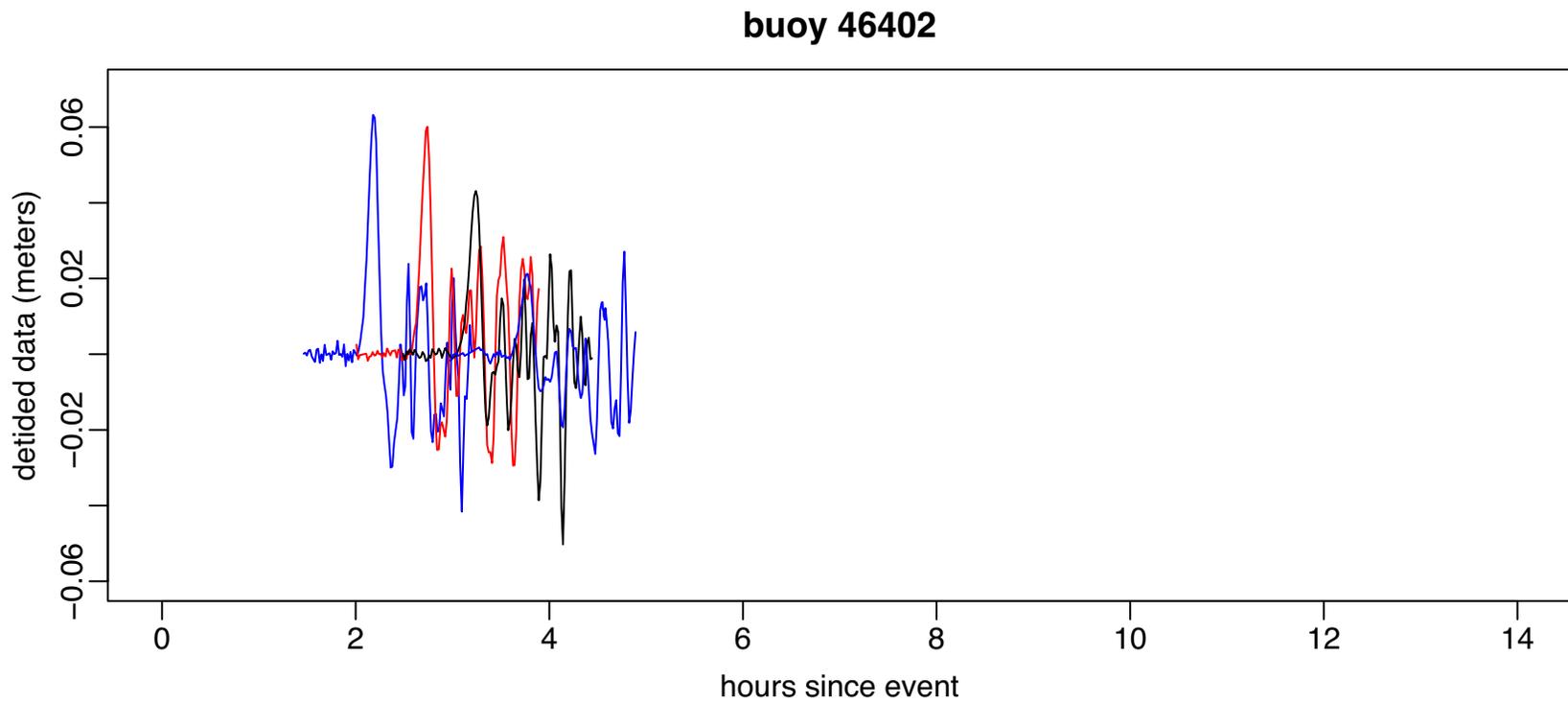
# Detrended DART Buoy Data for Kuril Island Event



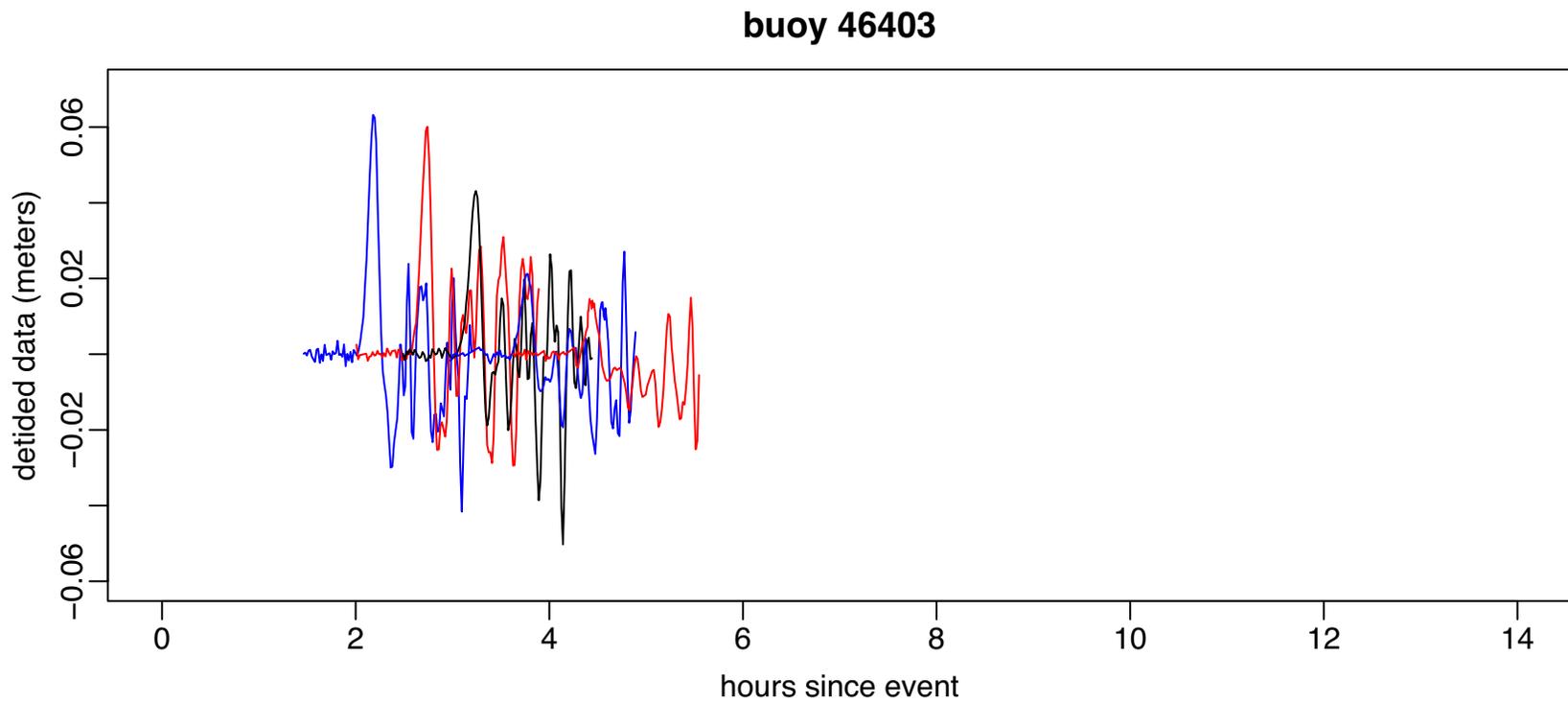
# Detrended DART Buoy Data for Kuril Island Event



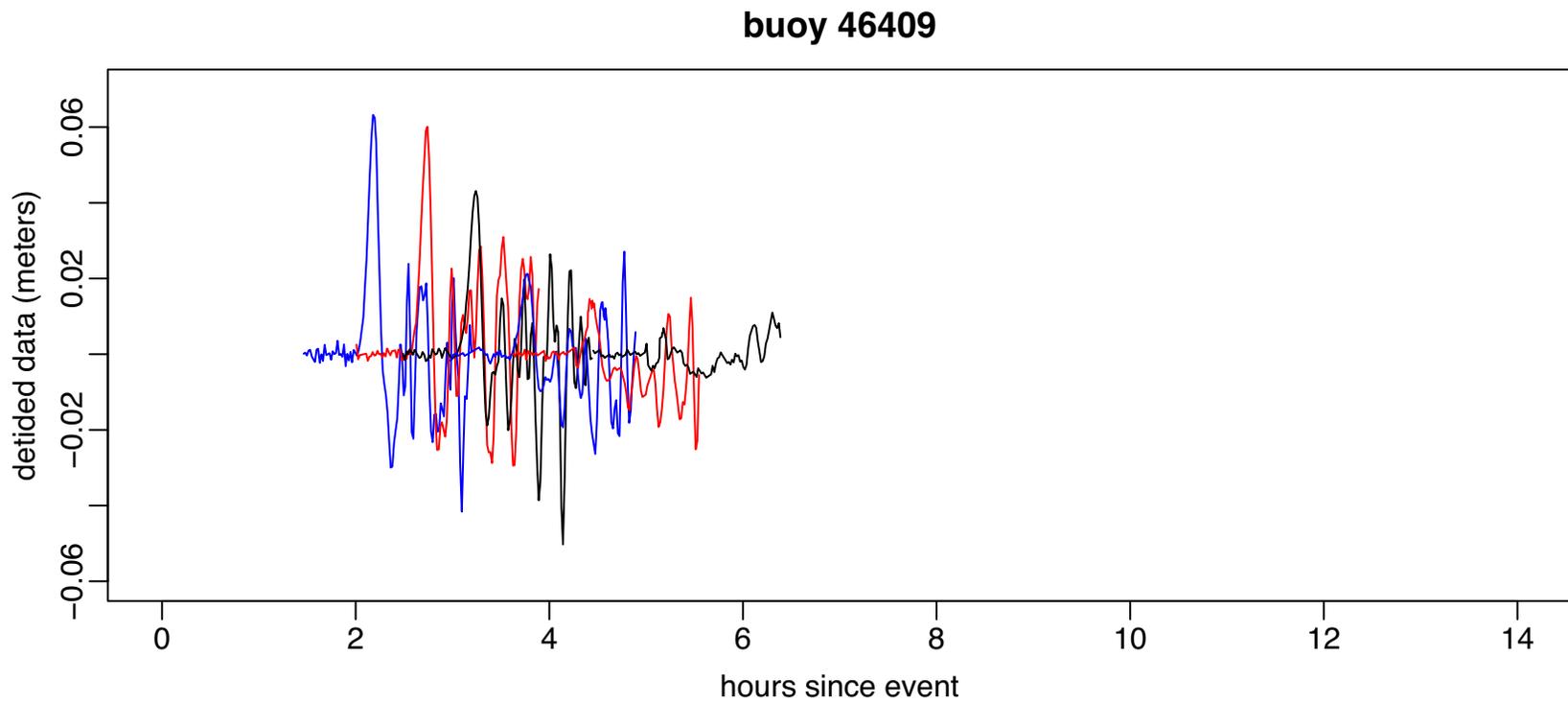
# Detrended DART Buoy Data for Kuril Island Event



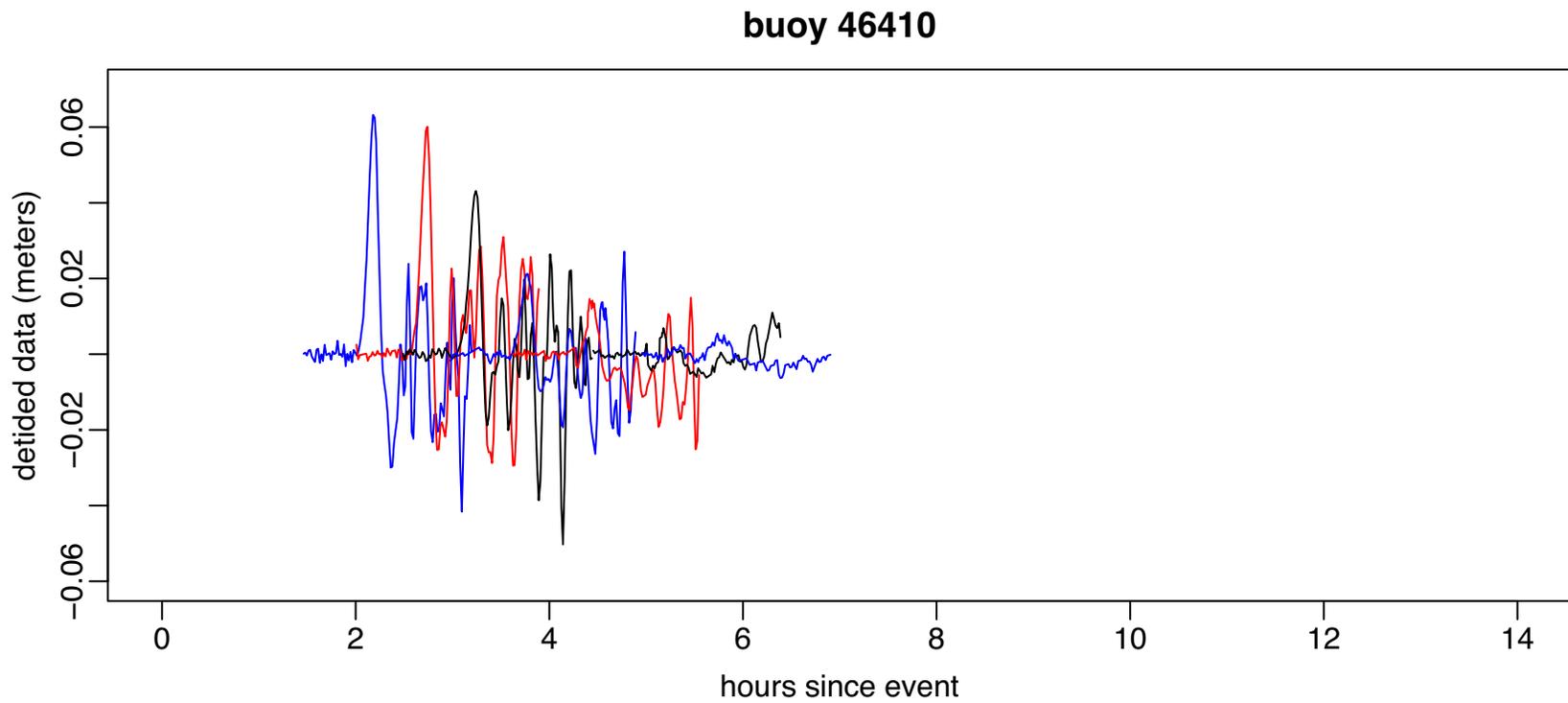
# Detrended DART Buoy Data for Kuril Island Event



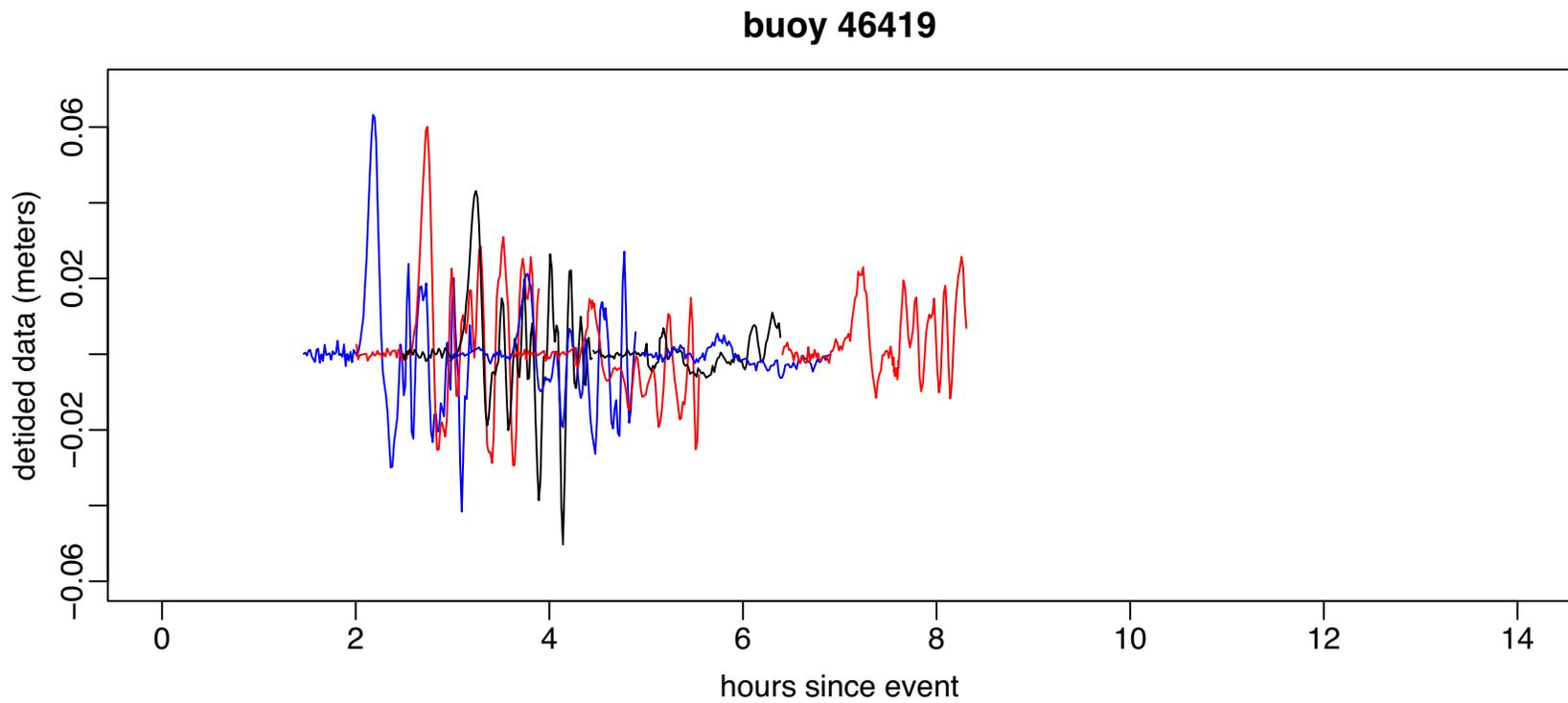
# Detrended DART Buoy Data for Kuril Island Event



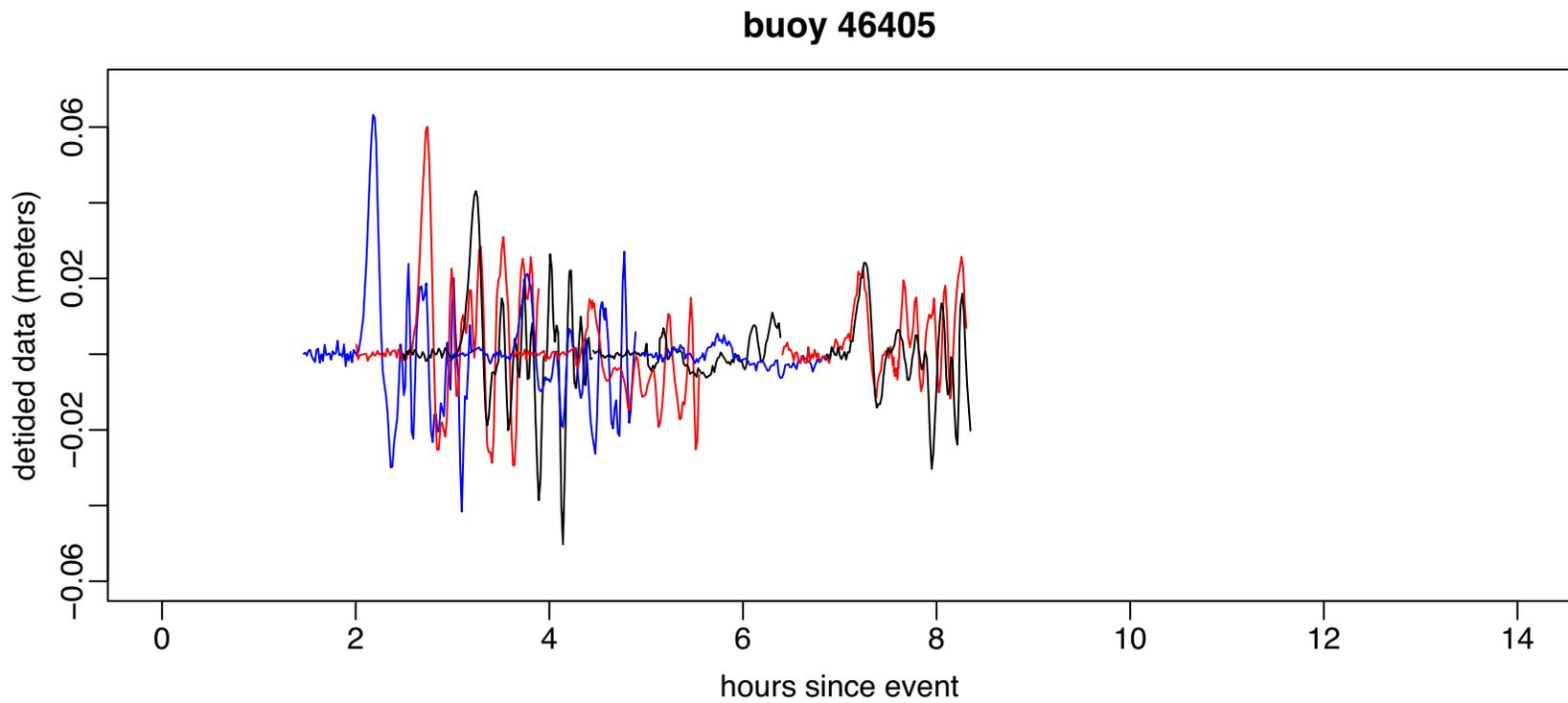
# Detrended DART Buoy Data for Kuril Island Event



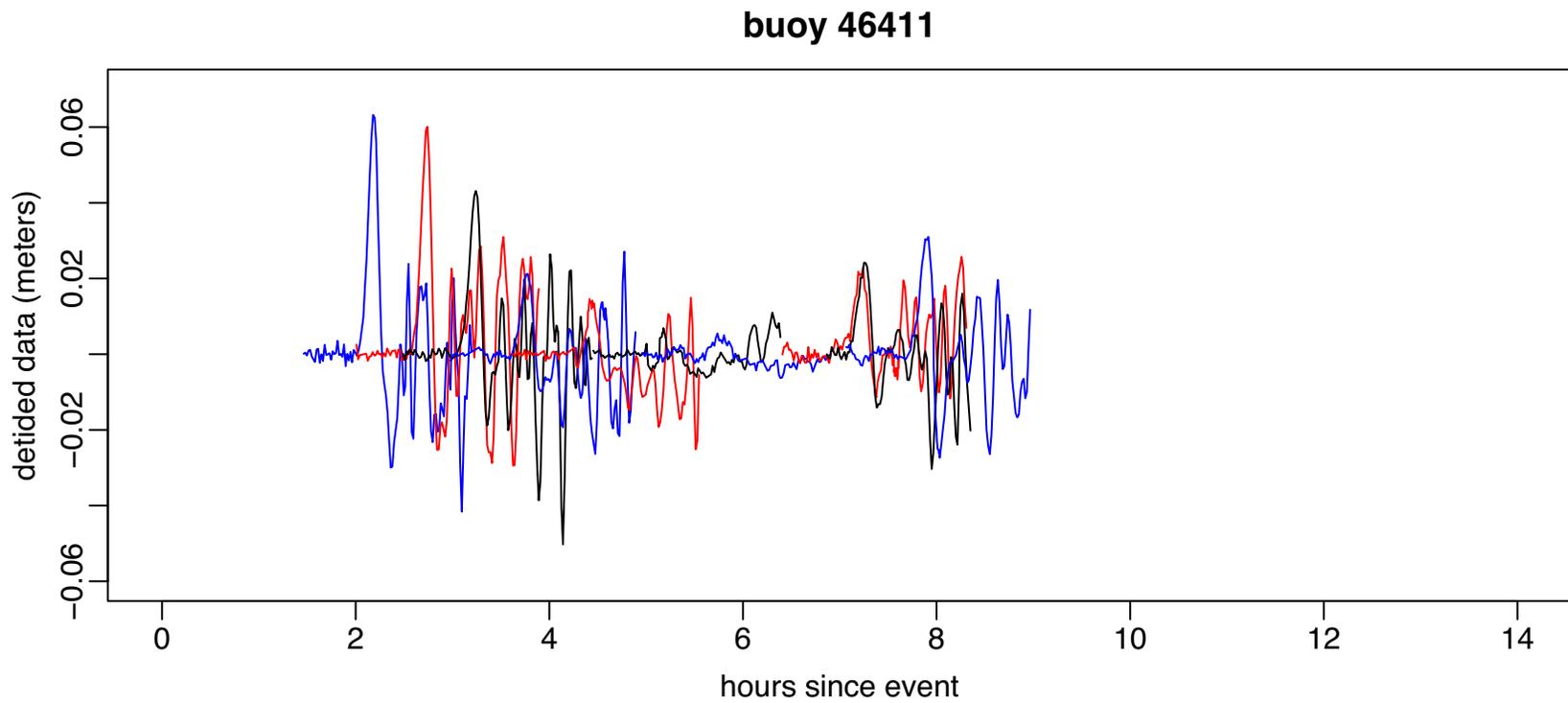
# Detrended DART Buoy Data for Kuril Island Event



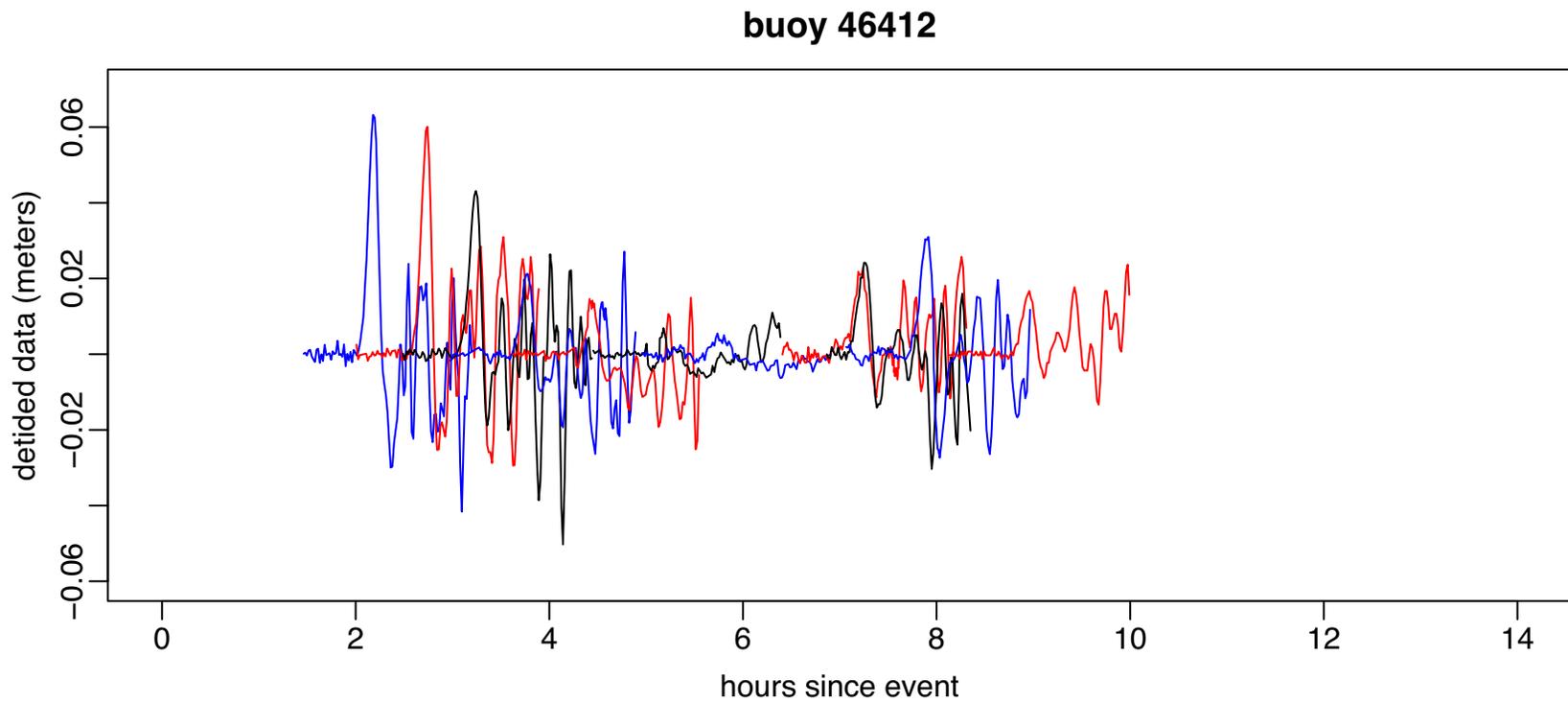
# Detrended DART Buoy Data for Kuril Island Event



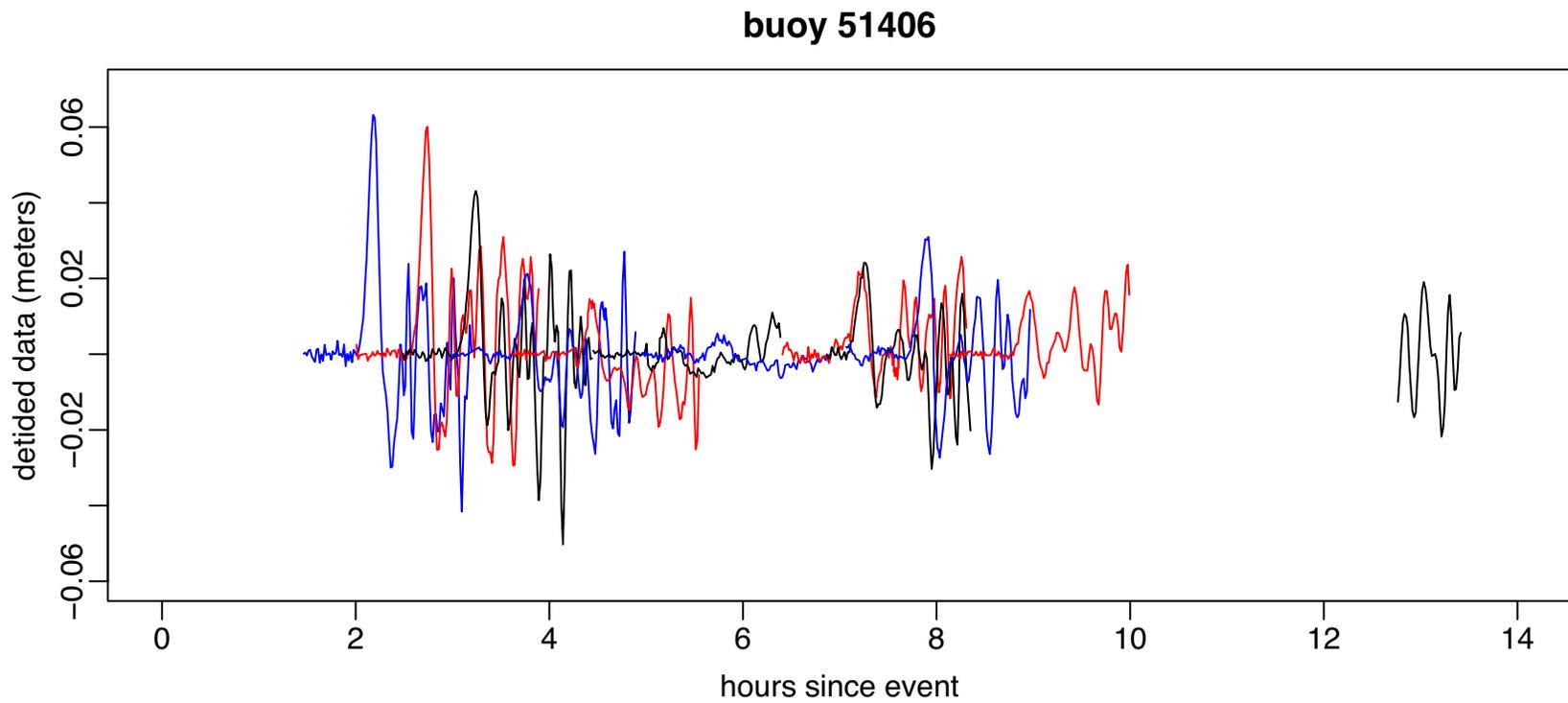
# Detrended DART Buoy Data for Kuril Island Event



# Detrended DART Buoy Data for Kuril Island Event



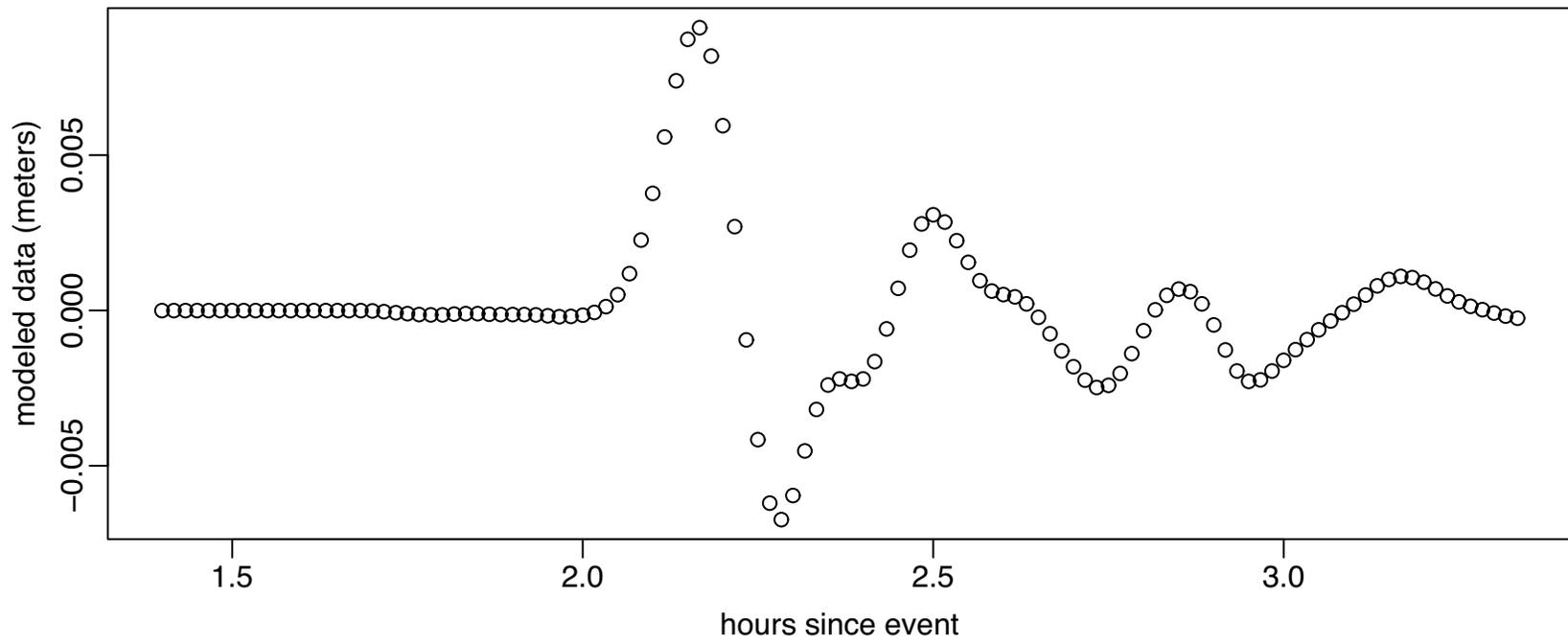
# Detrended DART Buoy Data for Kuril Island Event



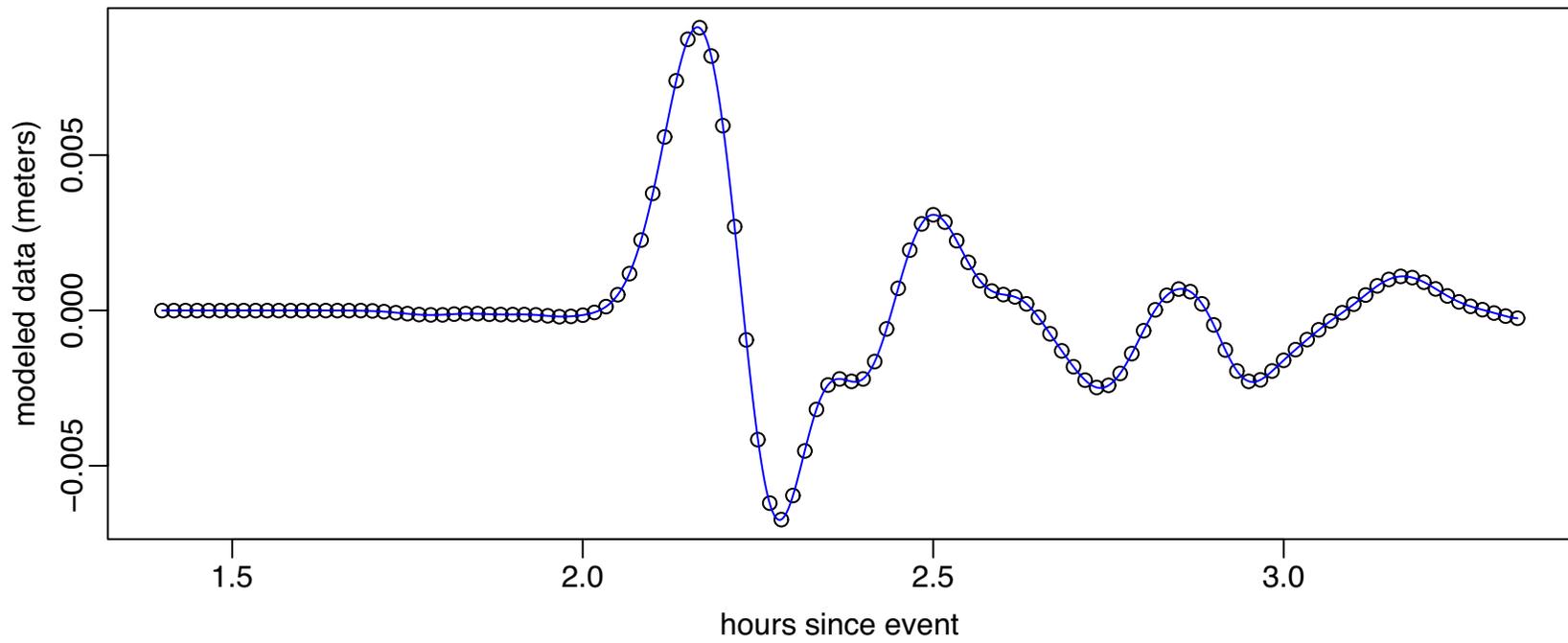
## Models for DART Buoys

- consider source locations for Kuril Island event
- for a given source location (e.g., a12), can generate a model of what we would expect to see at each buoy if the earthquake came from just that source
- each model is generated over a grid of discrete times, which might or might not correspond to the times at which DART buoy data are collected
- use cubic spline to interpolate model, so can regard model  $g(t)$  and its first derivative  $g'(t)$  as being defined for all times  $t$
- as an example, consider model from a12 source for buoy 21414

# a12 Source Model for Buoy 21414

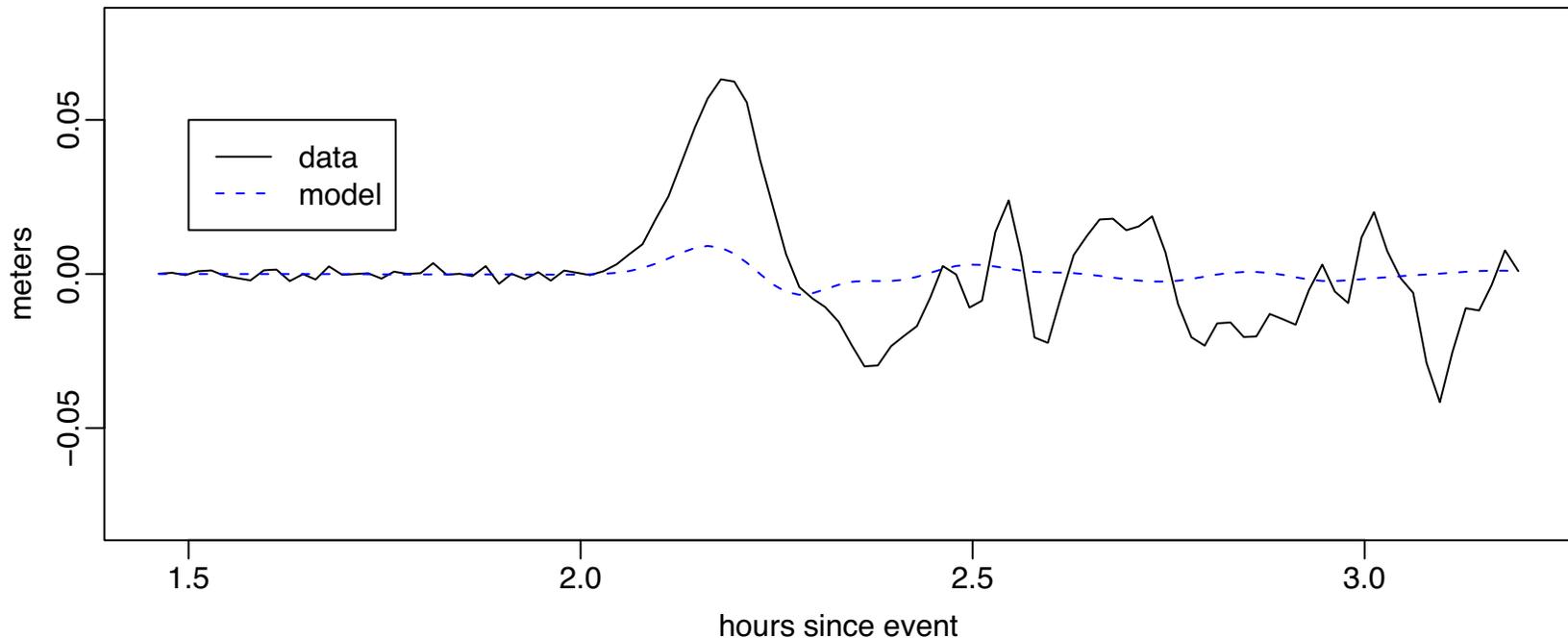


# a12 Source Model for Buoy 21414



## Fitting Models to DART Buoy Data: I

- model from a12 source for buoy 21414 generated under assumption of a unit magnitude for the earthquake



- poor match, so multiple model  $g(t)$  by  $A$  to get a better fit, where  $A$  is interpreted as earthquake magnitude (the 'slip')

## Fitting Models to DART Buoy Data: II

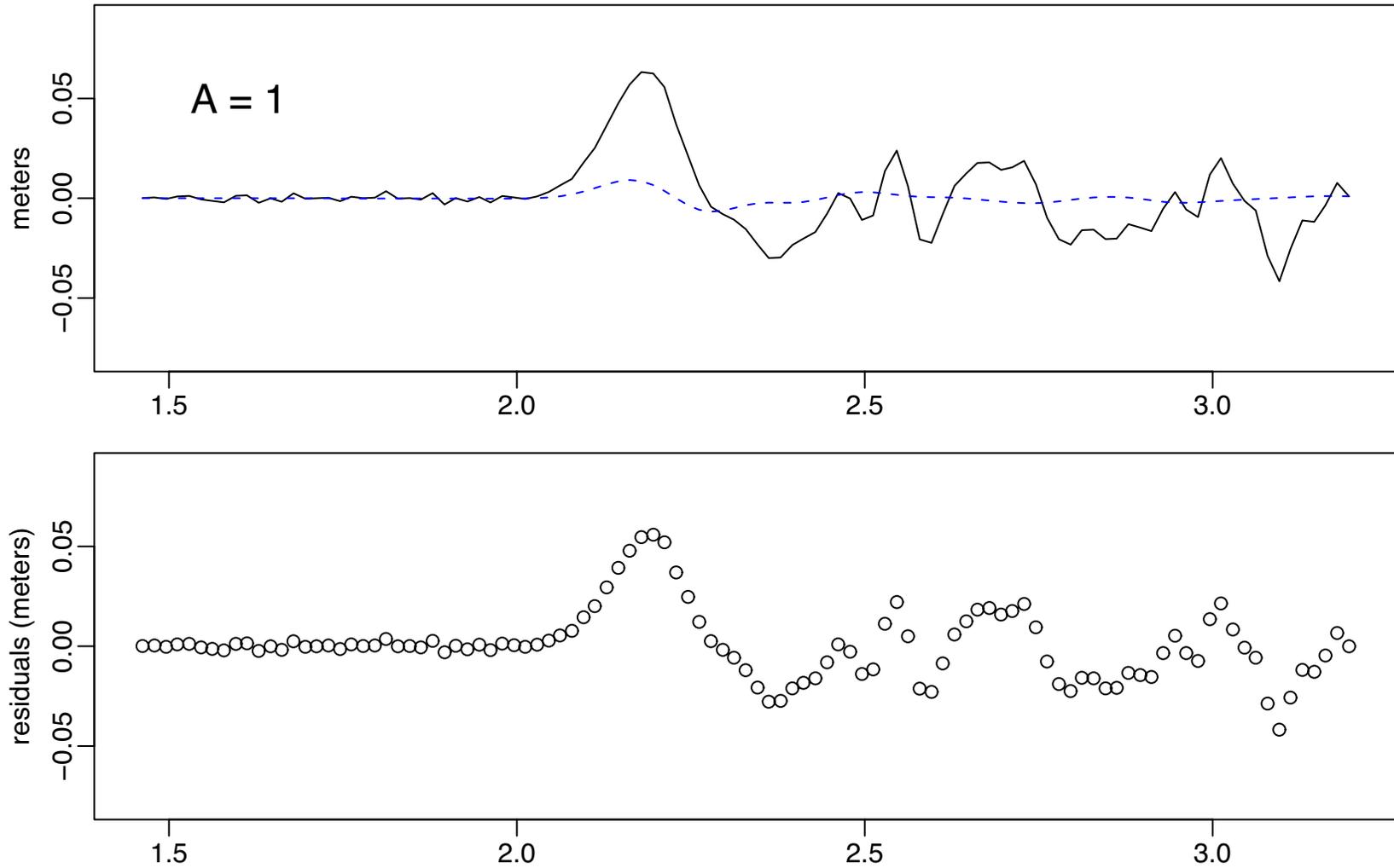
- let  $x_t$  represent the DART buoy data at time  $t$
- we entertain the model

$$x_t = A \cdot g(t) + e_t,$$

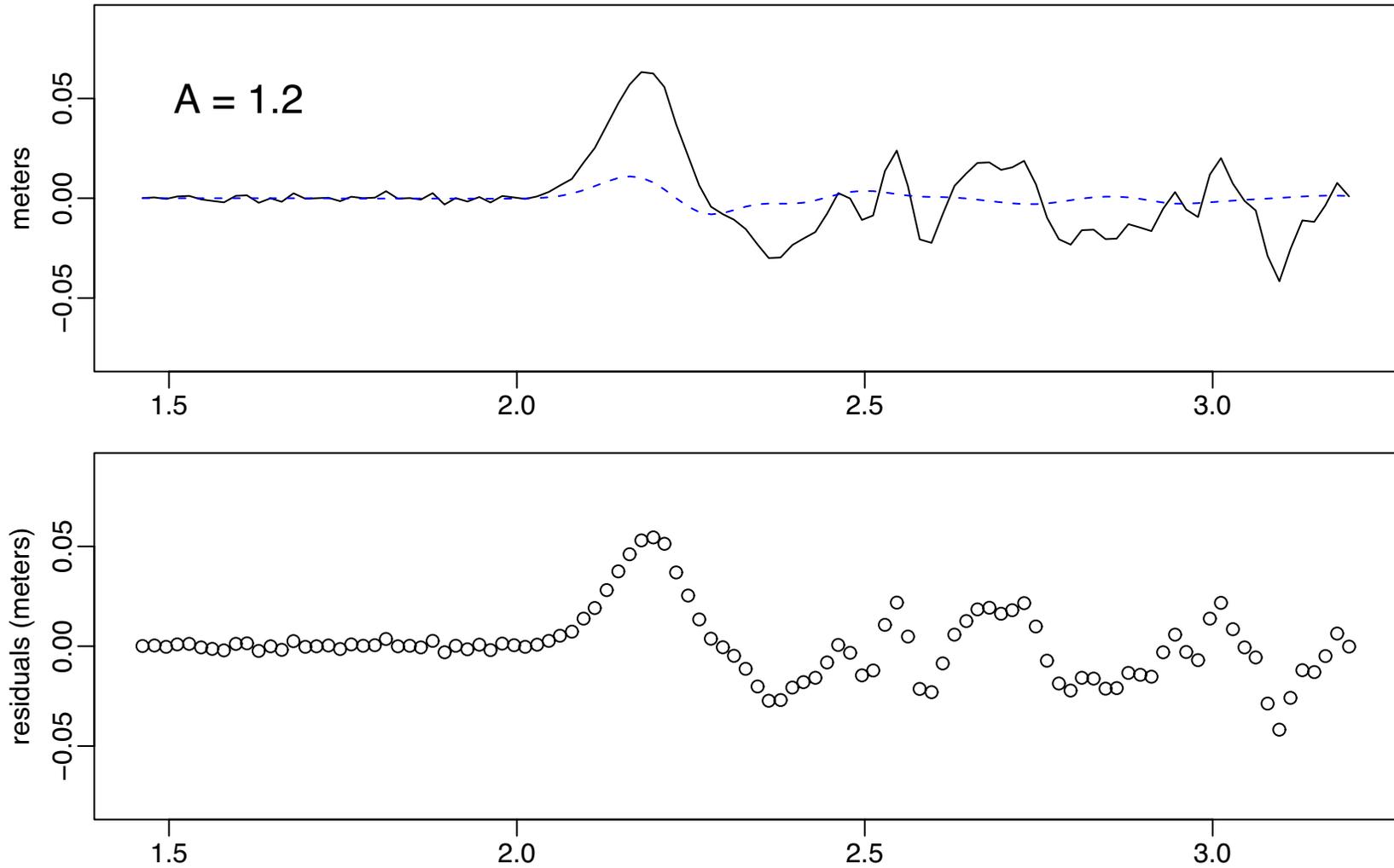
where  $e_t$  is a residual term (mismatch between data and model)

- if we let  $A$  range over a grid of values, then we can compute corresponding residuals  $e_t = x_t - A \cdot g(t)$  for any given  $A$
- as an example, let us compute residuals from fit of a12 source to buoy 21414 data for  $A = 1.0, 1.2, 1.4, \dots, 10.0$

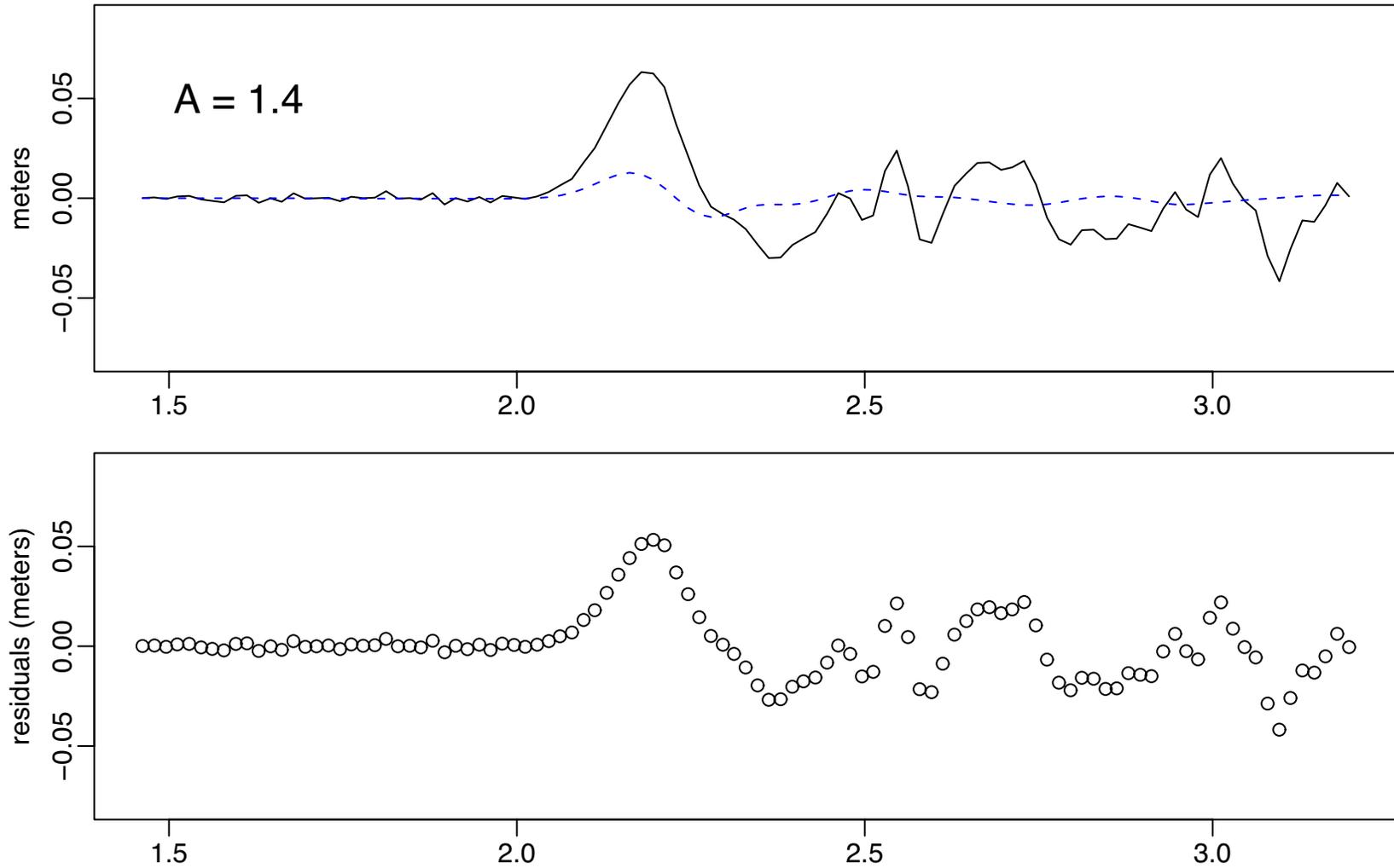
# Matching a12 Model to 21414 Data by Varying Slip



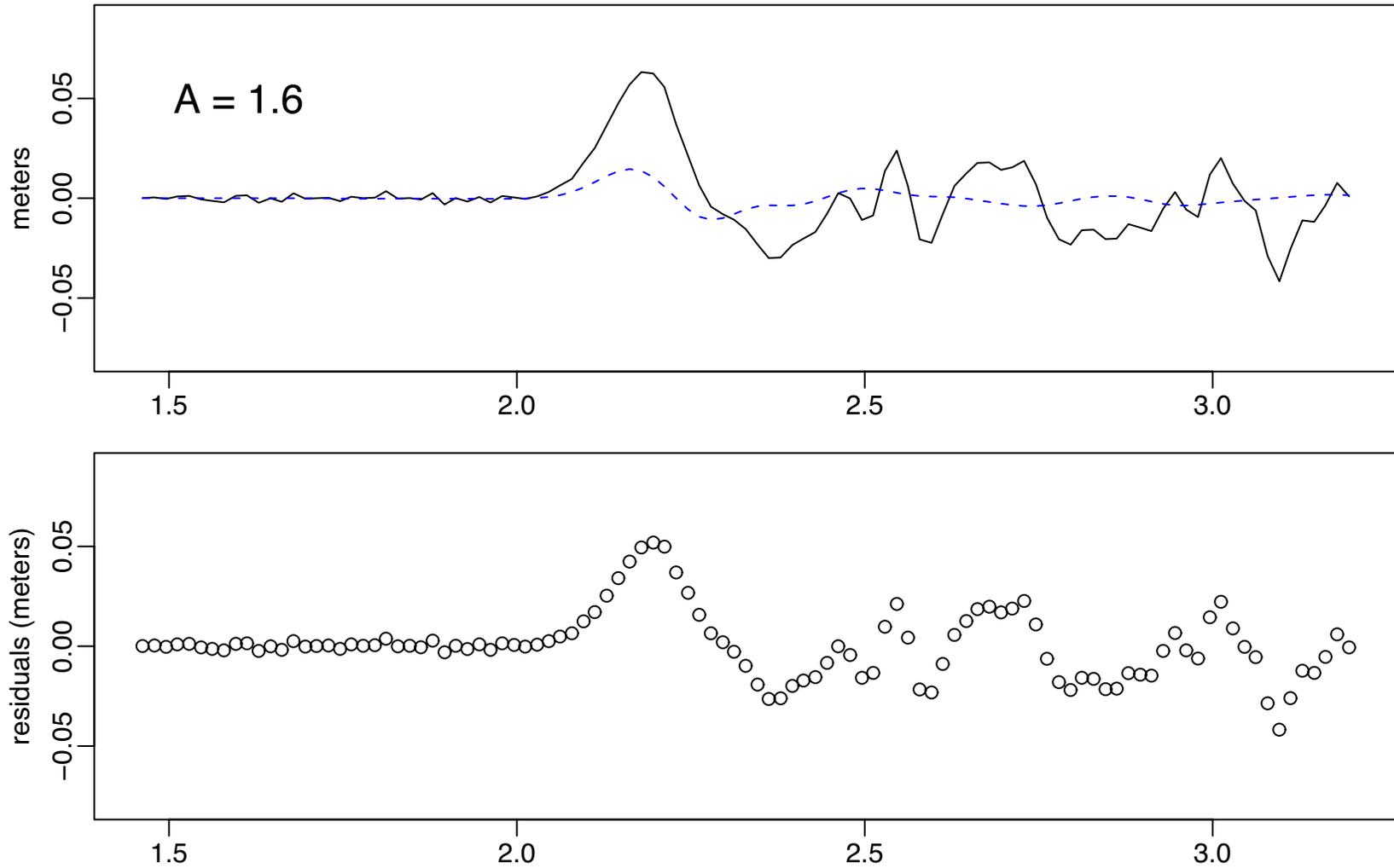
# Matching a12 Model to 21414 Data by Varying Slip



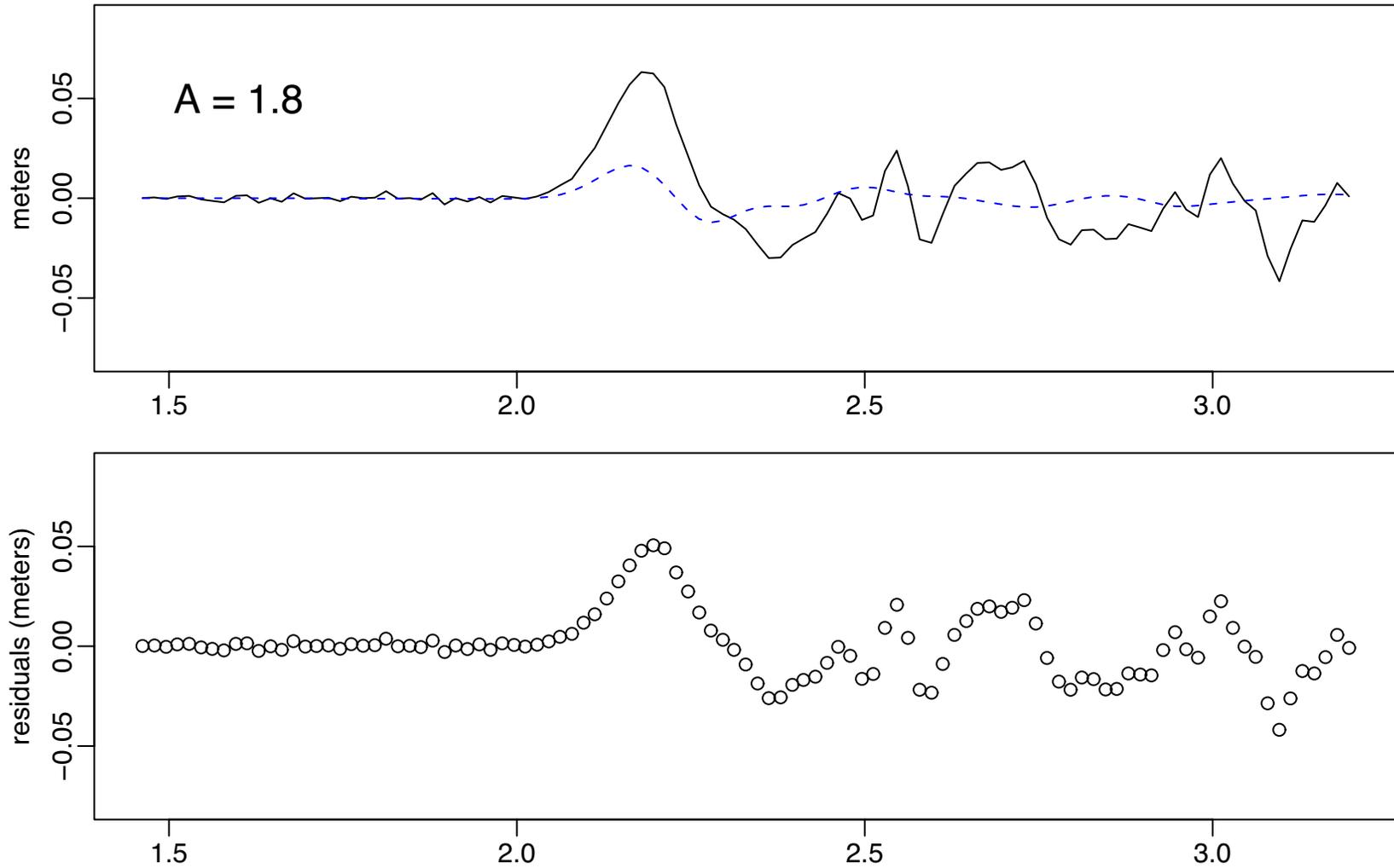
# Matching a12 Model to 21414 Data by Varying Slip



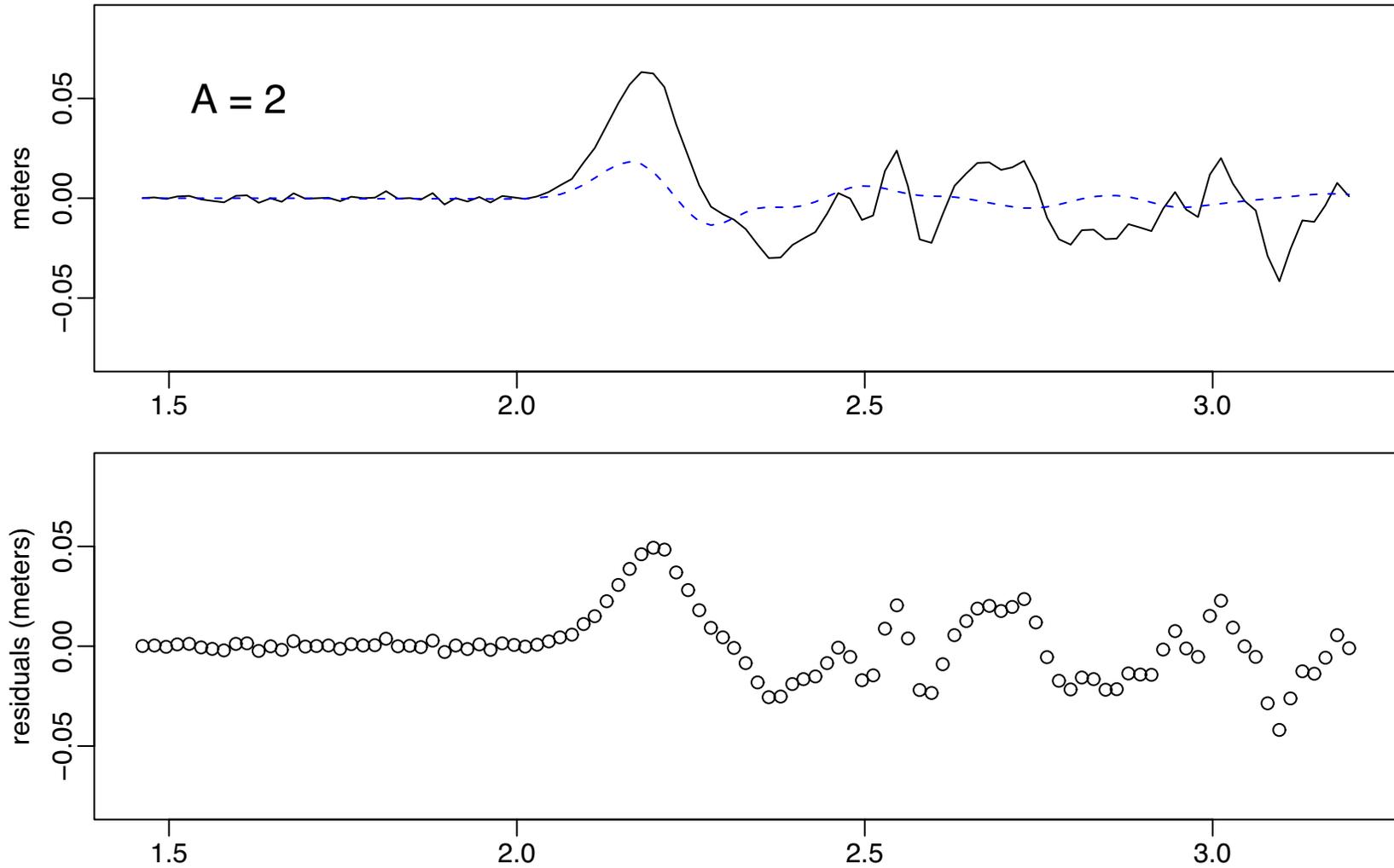
# Matching a12 Model to 21414 Data by Varying Slip



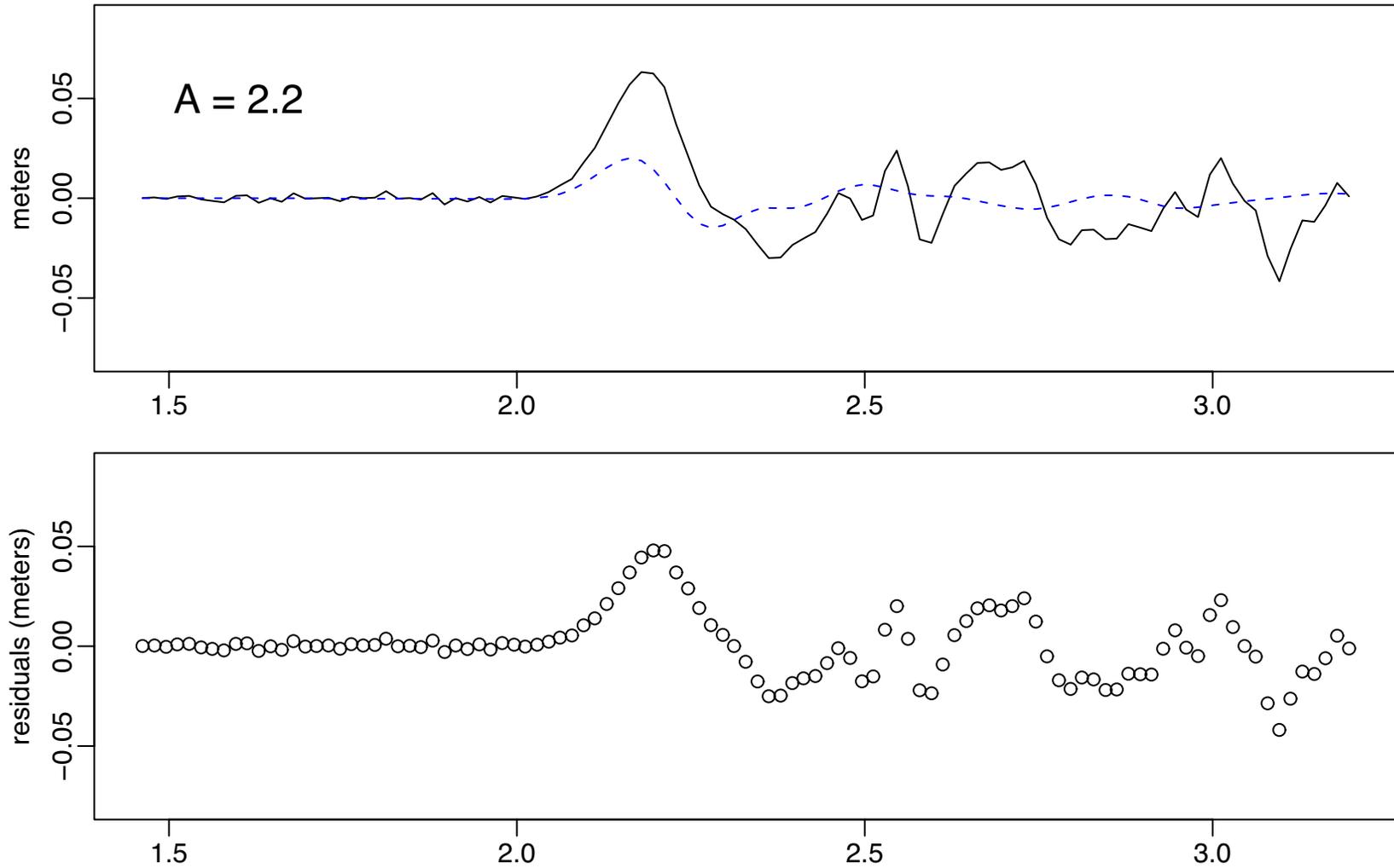
# Matching a12 Model to 21414 Data by Varying Slip



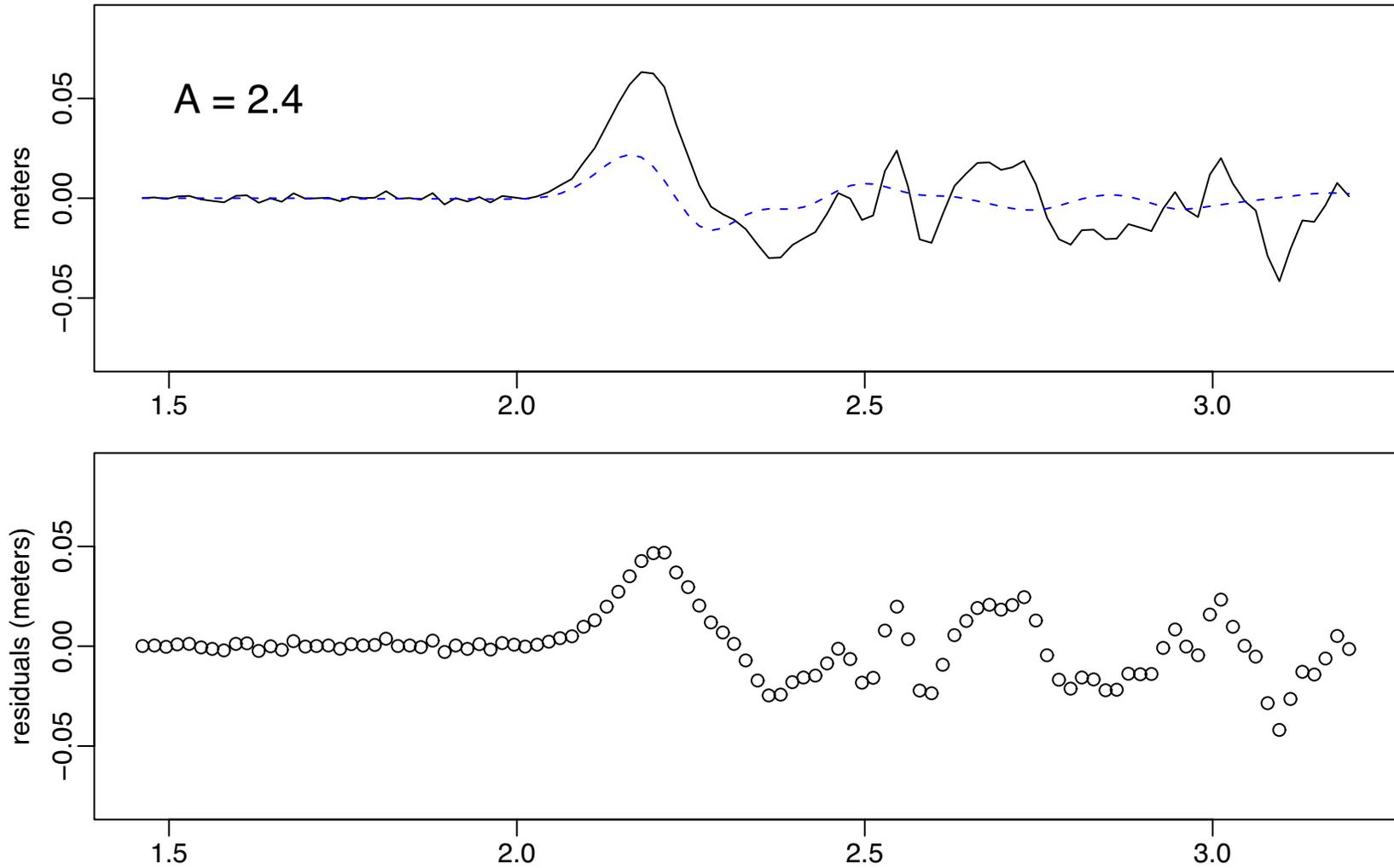
# Matching a12 Model to 21414 Data by Varying Slip



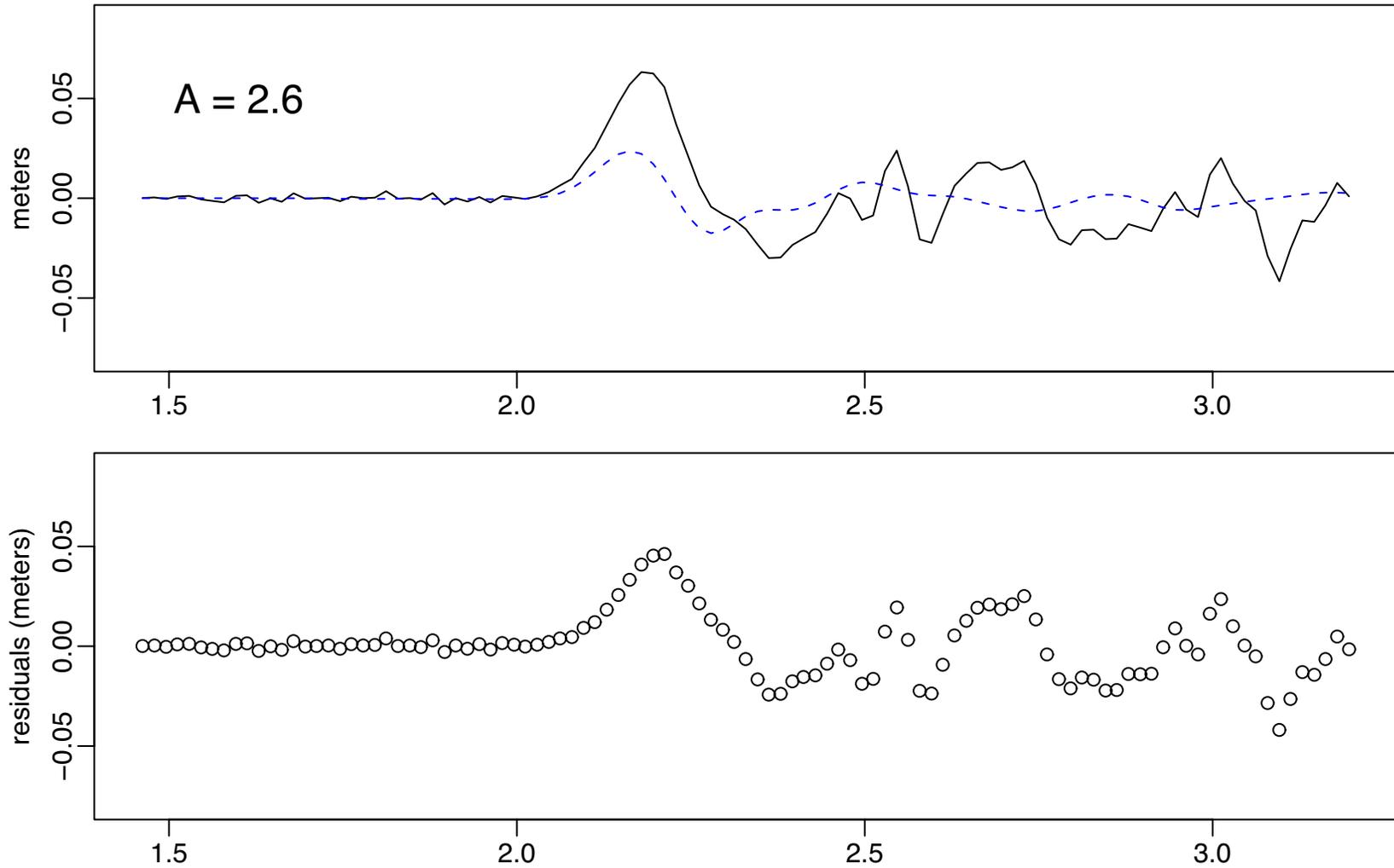
# Matching a12 Model to 21414 Data by Varying Slip



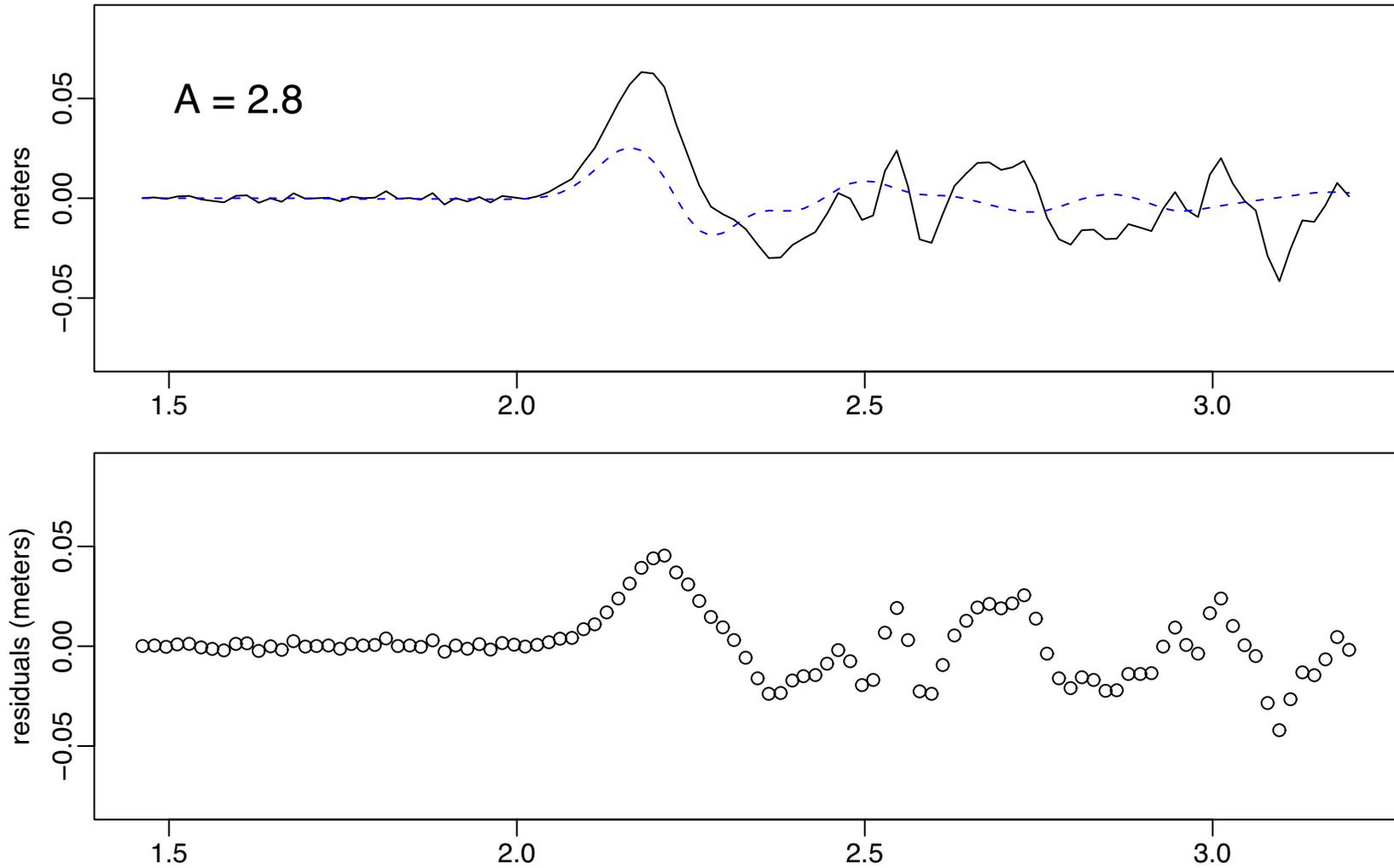
# Matching a12 Model to 21414 Data by Varying Slip



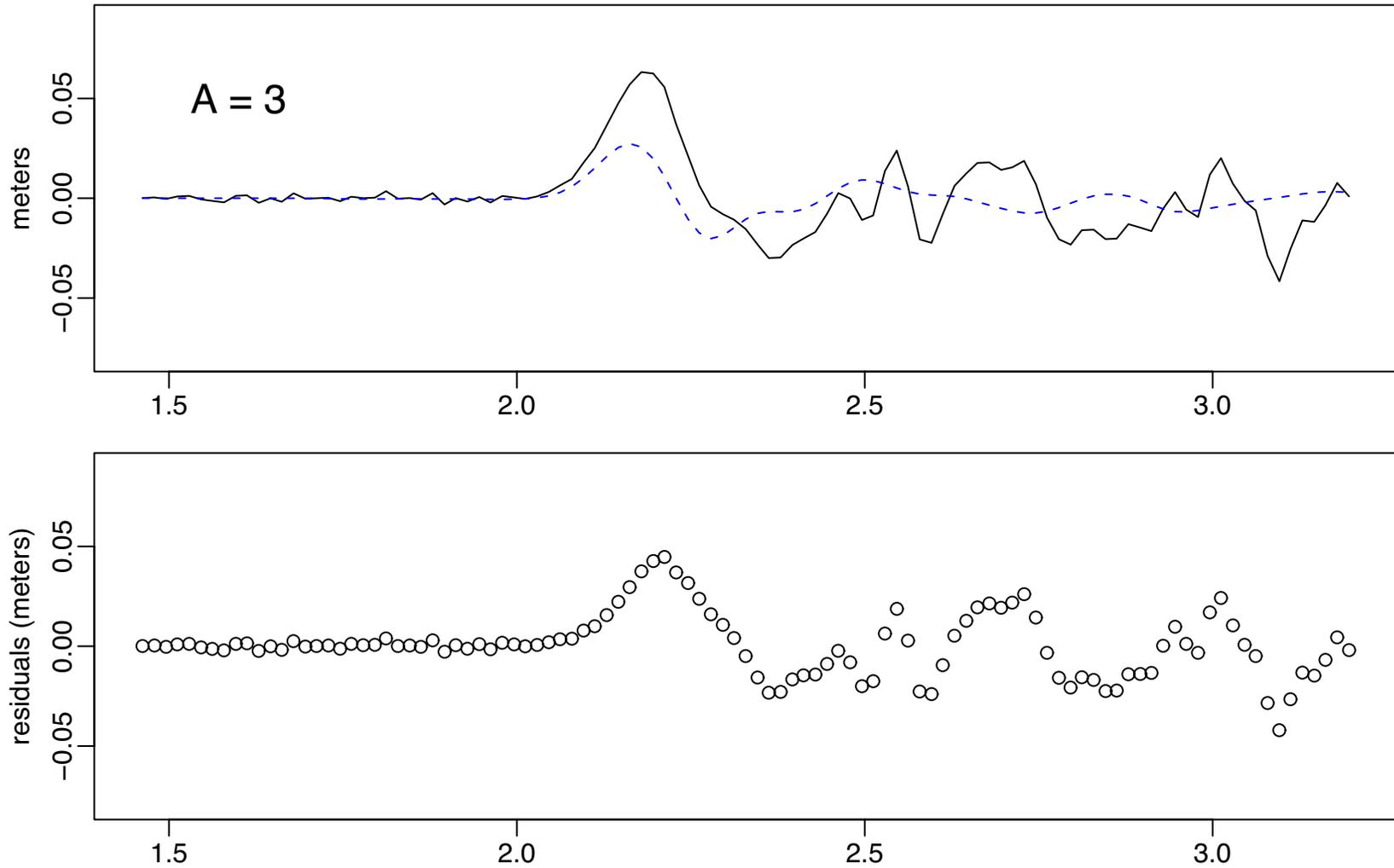
# Matching a12 Model to 21414 Data by Varying Slip



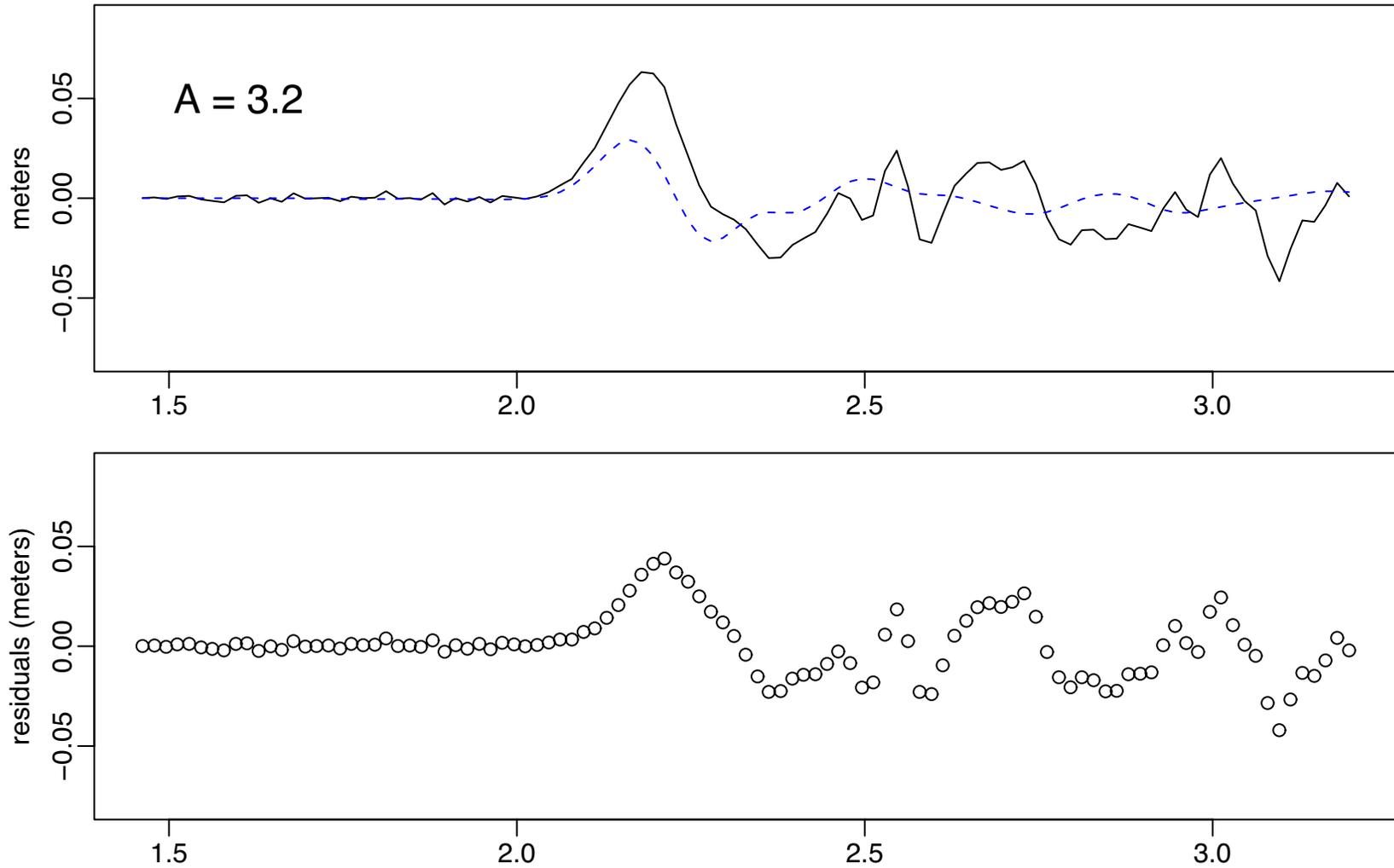
# Matching a12 Model to 21414 Data by Varying Slip



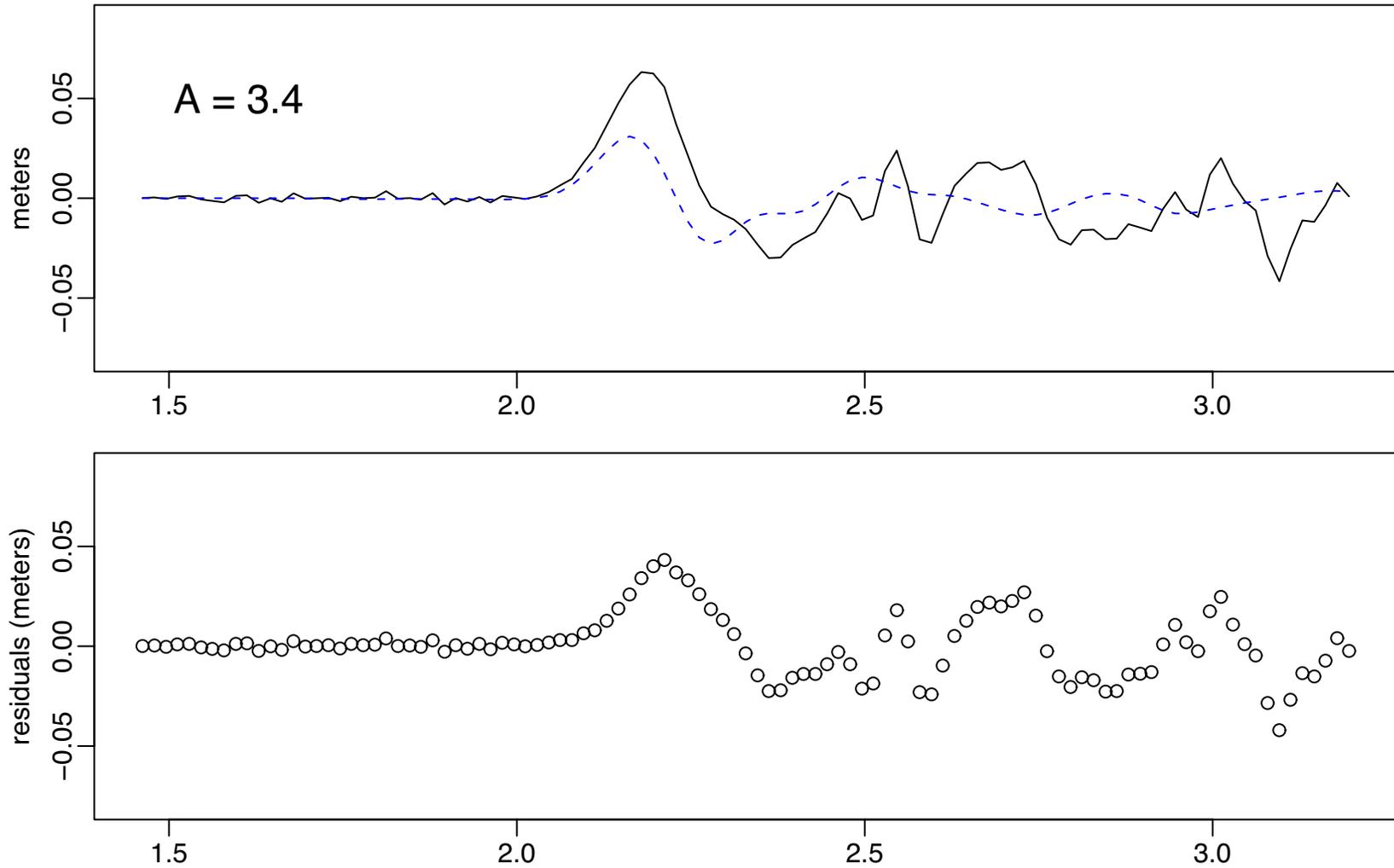
# Matching a12 Model to 21414 Data by Varying Slip



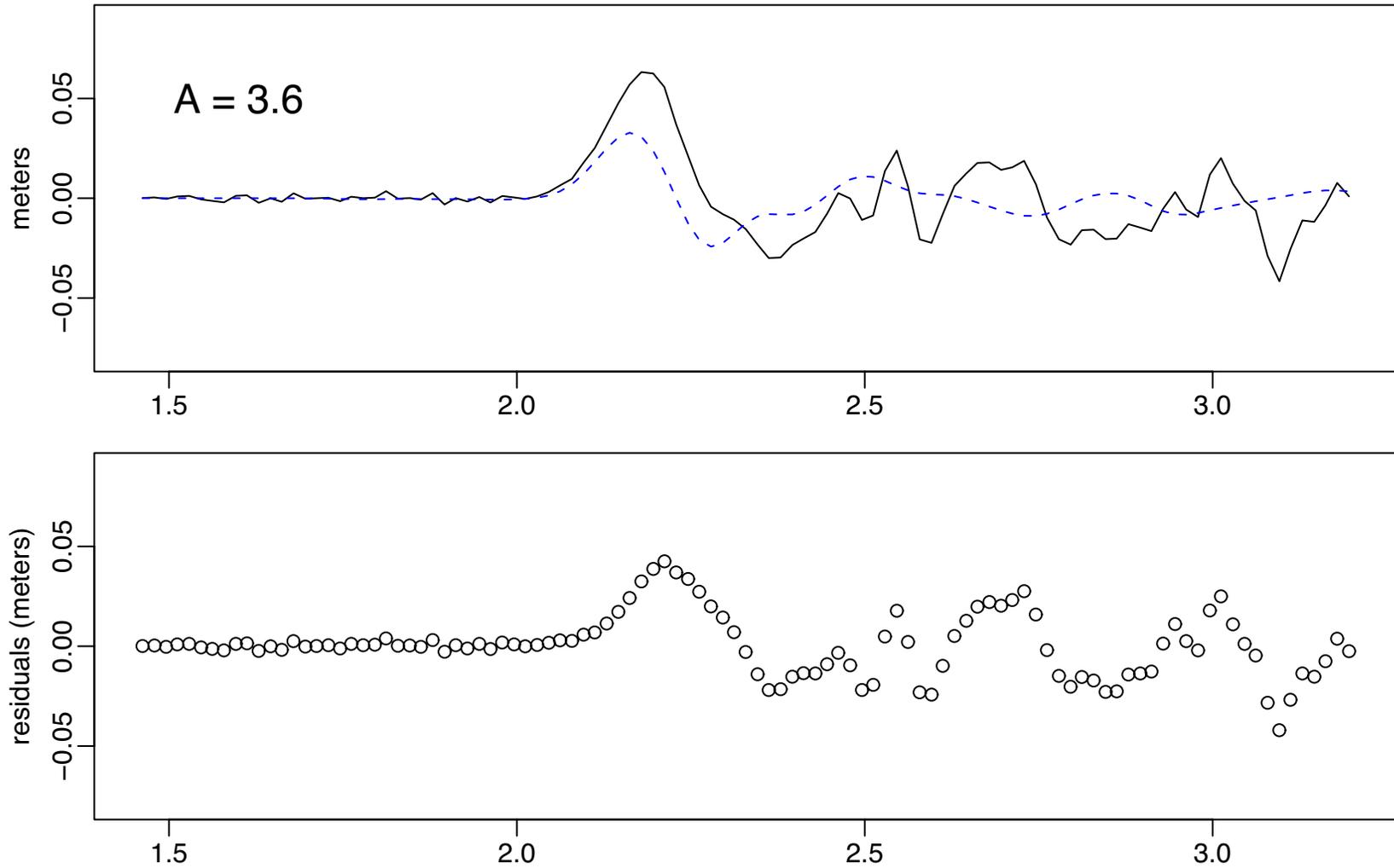
# Matching a12 Model to 21414 Data by Varying Slip



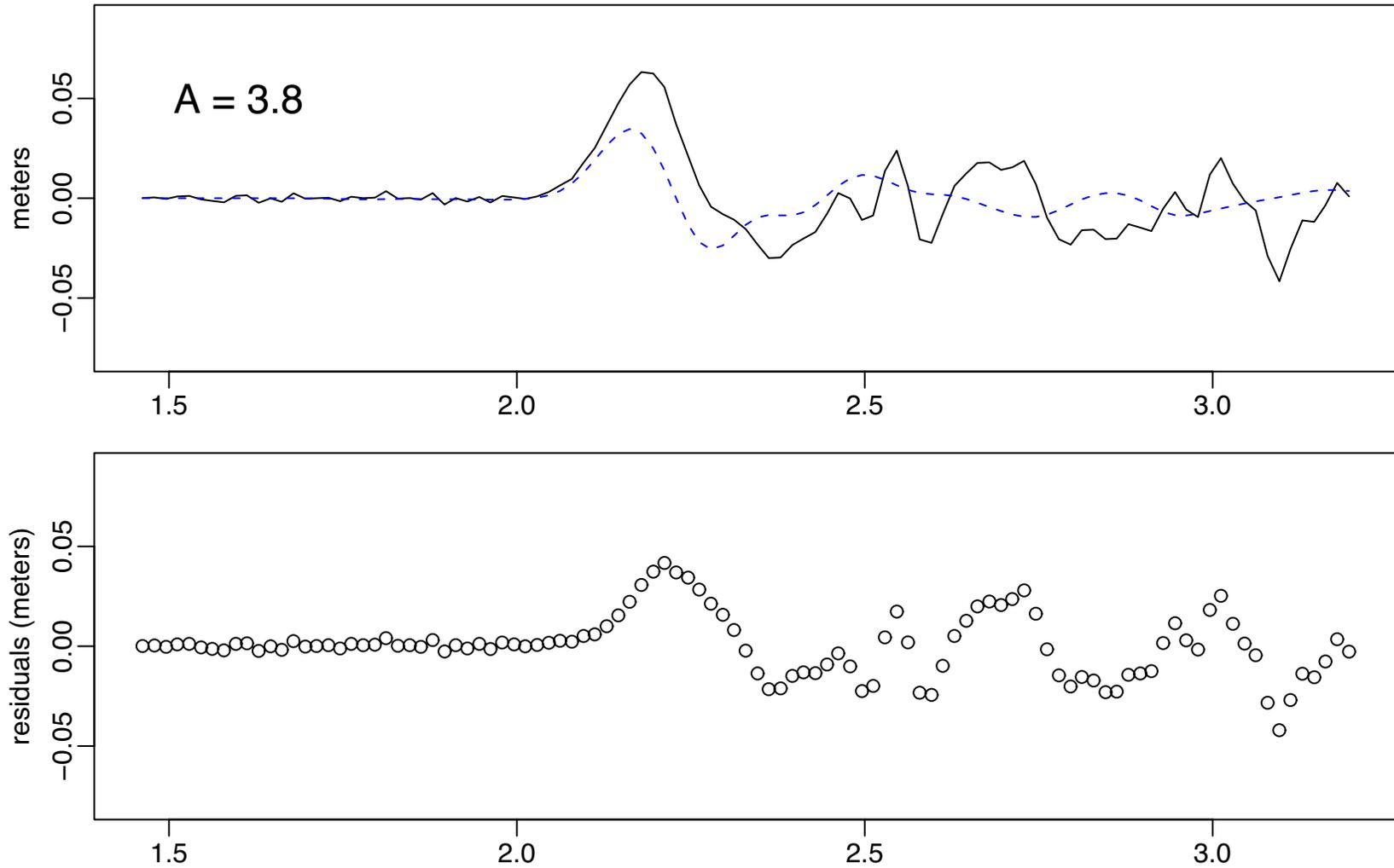
# Matching a12 Model to 21414 Data by Varying Slip



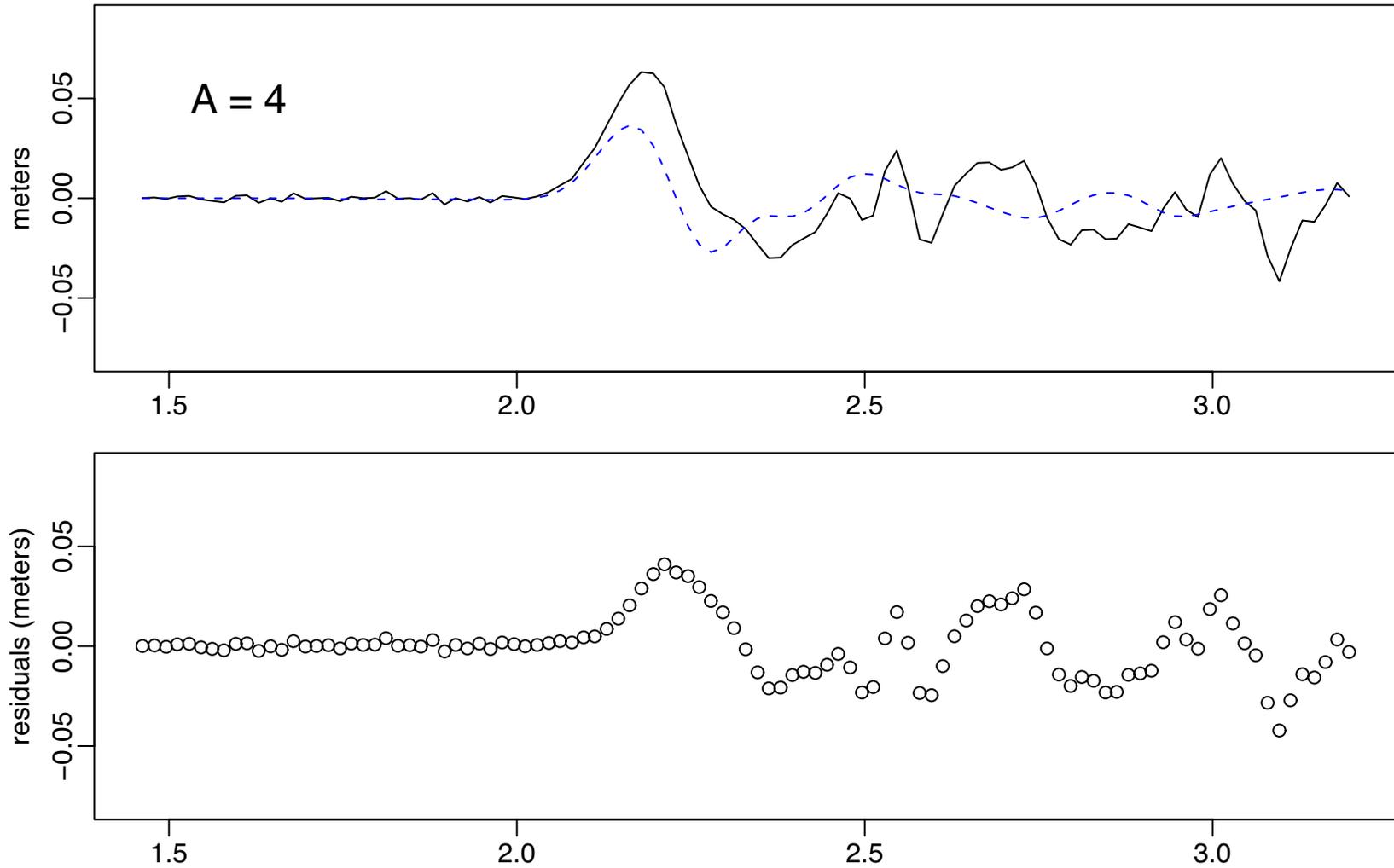
# Matching a12 Model to 21414 Data by Varying Slip



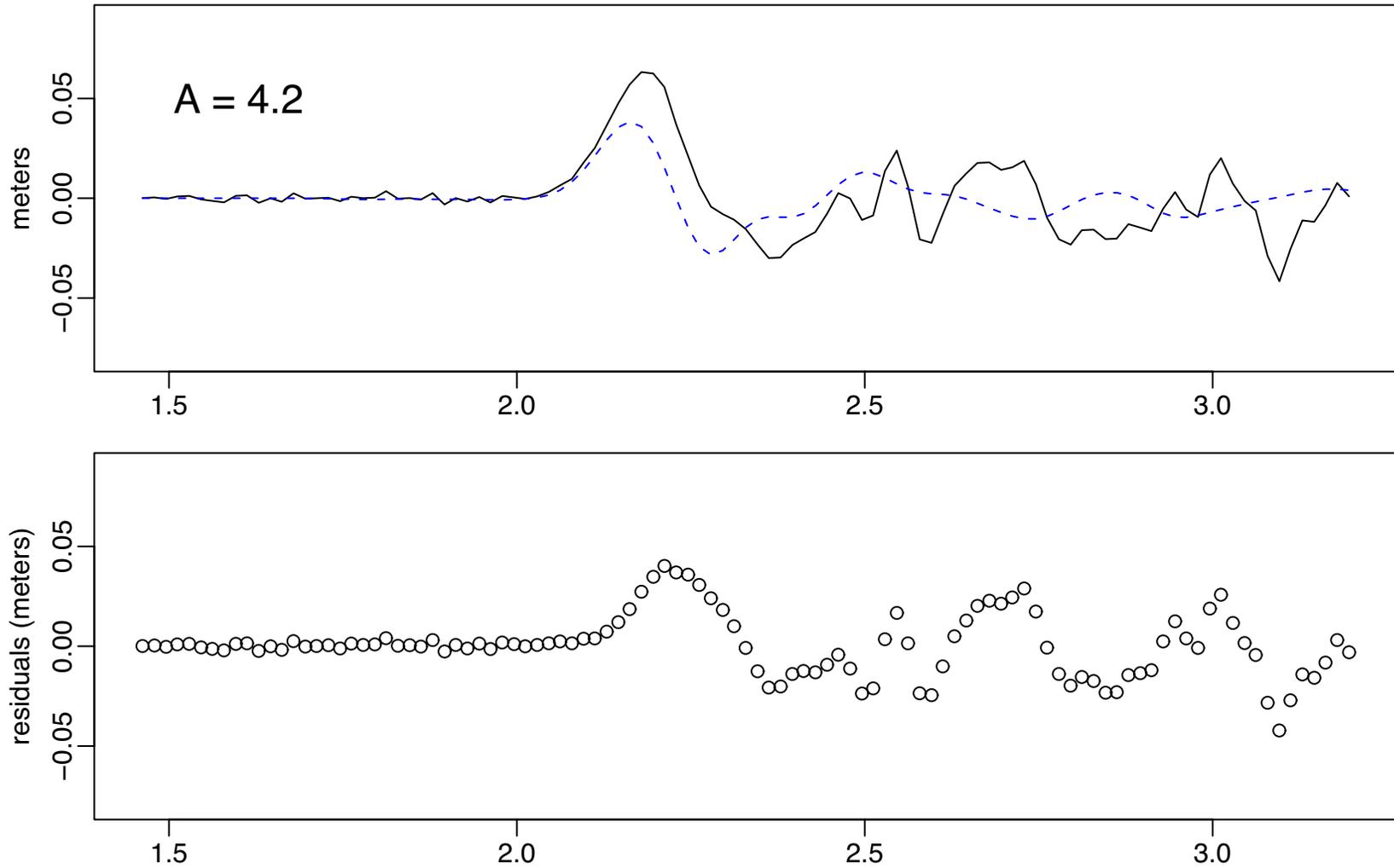
# Matching a12 Model to 21414 Data by Varying Slip



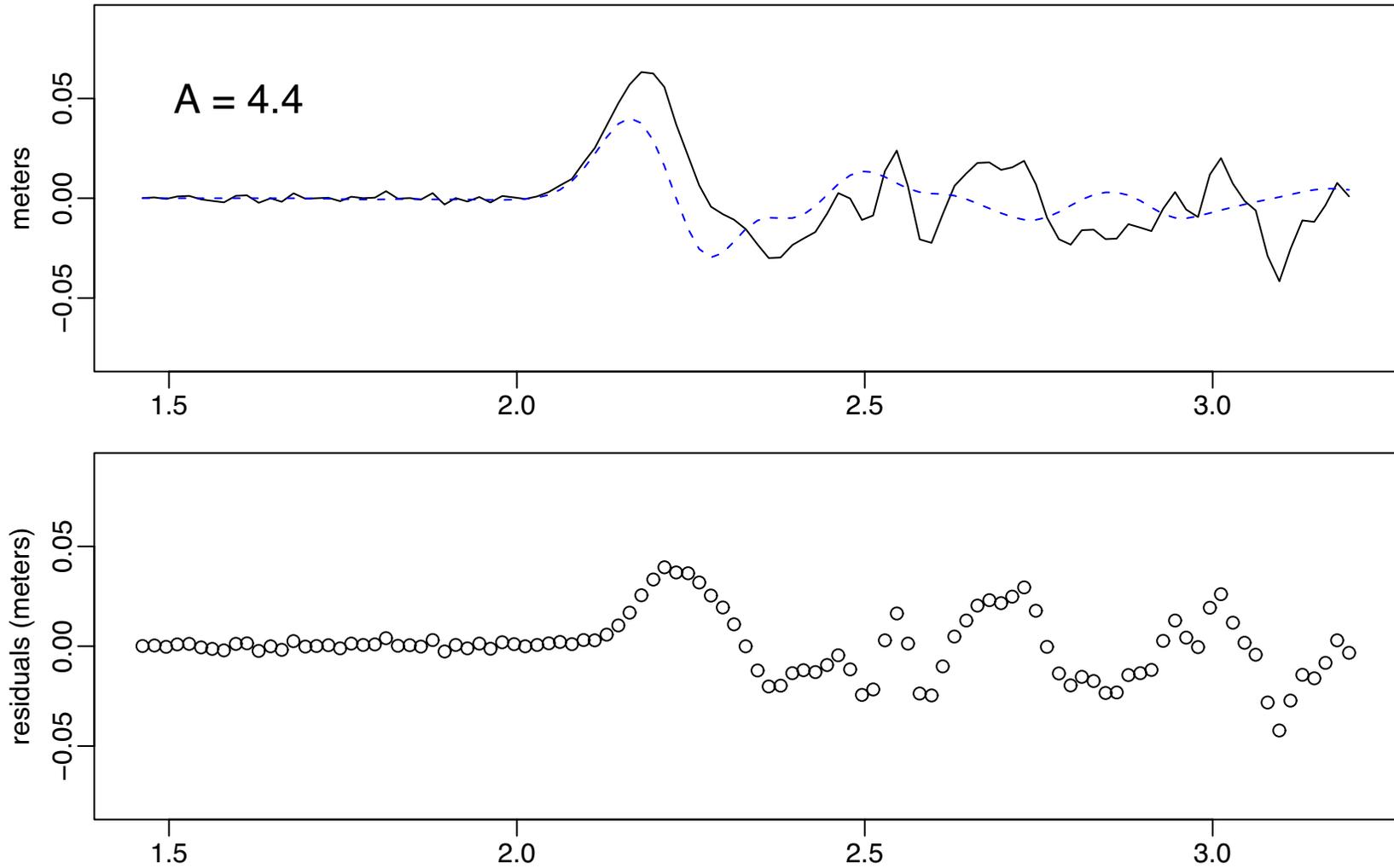
# Matching a12 Model to 21414 Data by Varying Slip



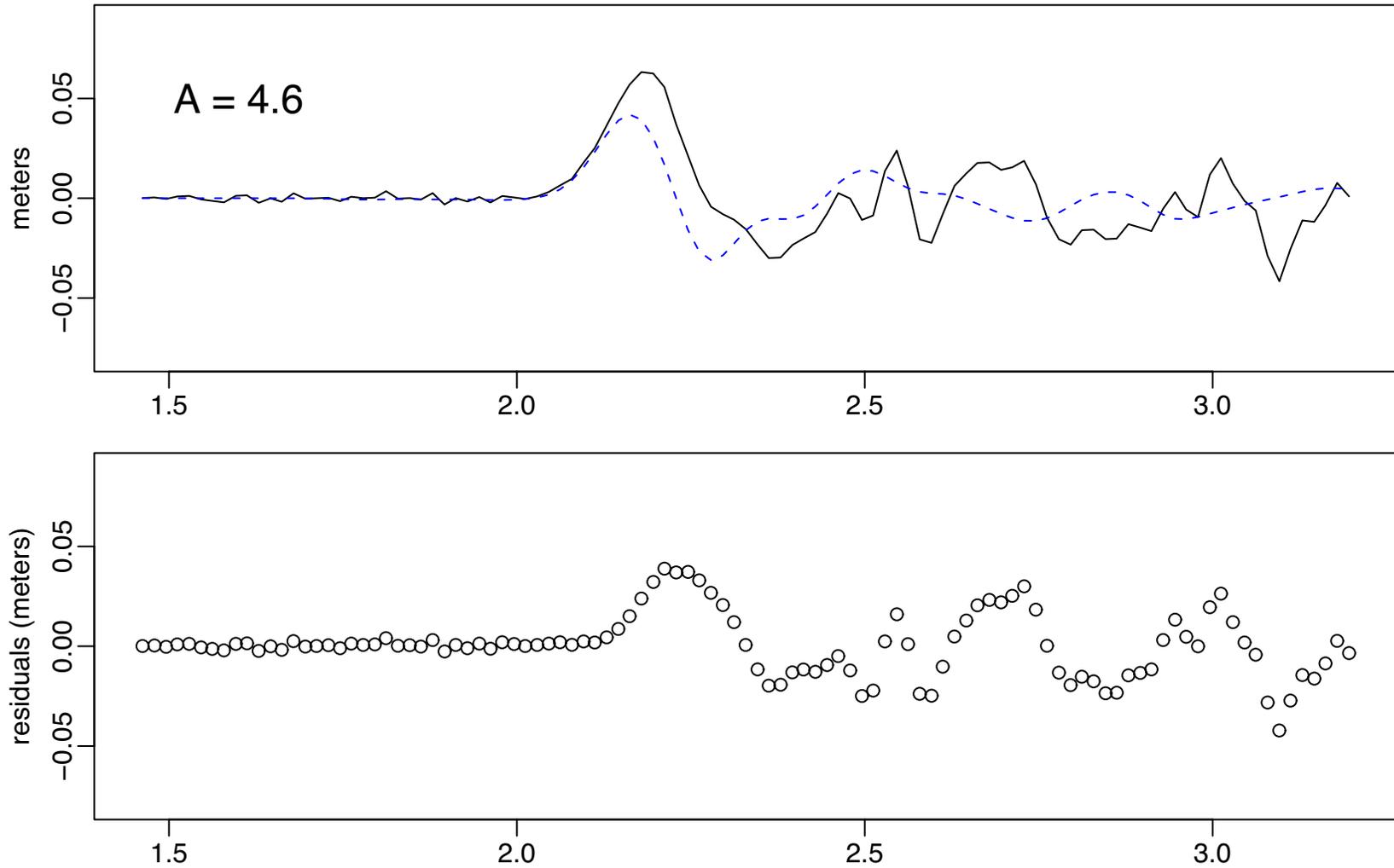
# Matching a12 Model to 21414 Data by Varying Slip



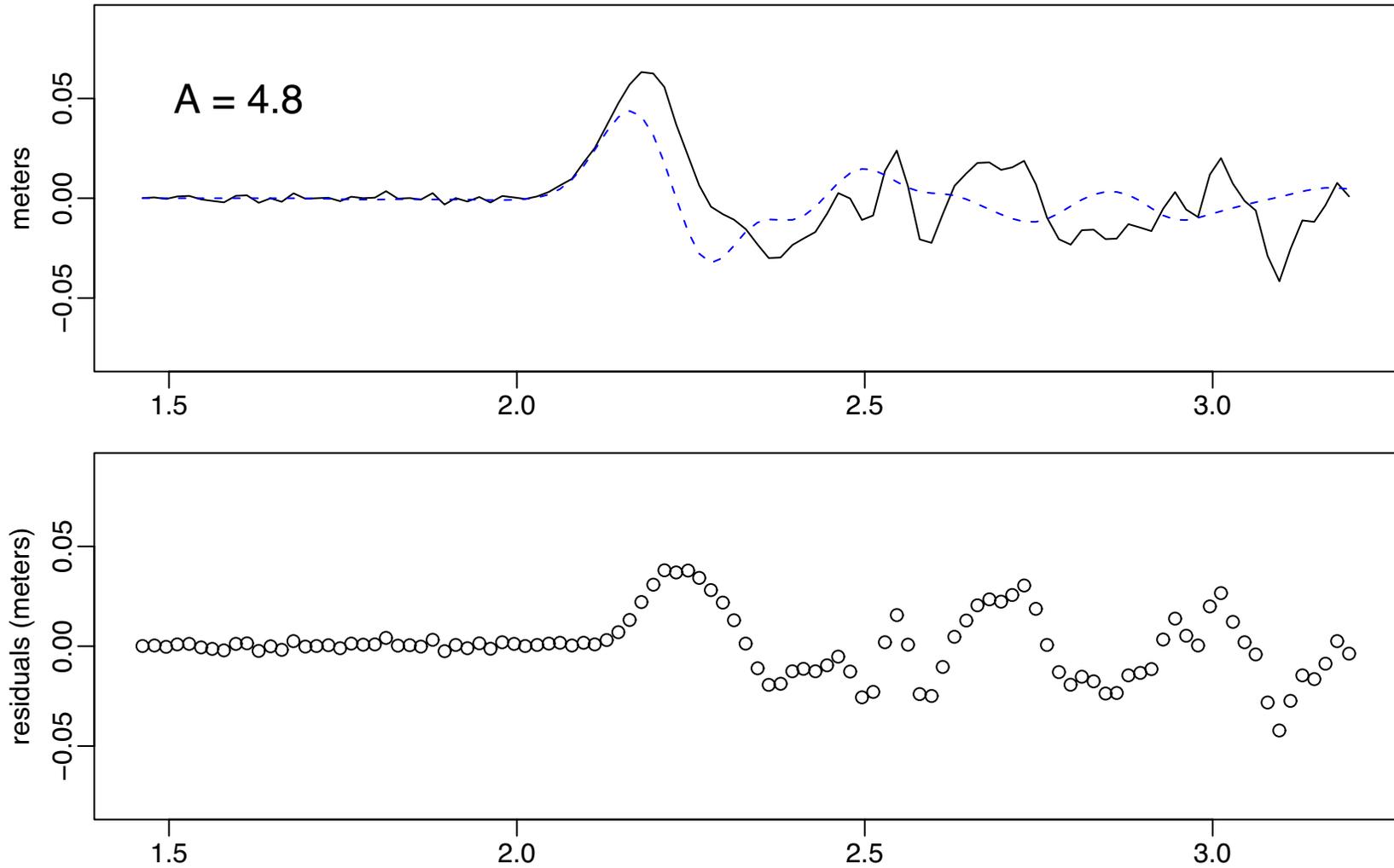
# Matching a12 Model to 21414 Data by Varying Slip



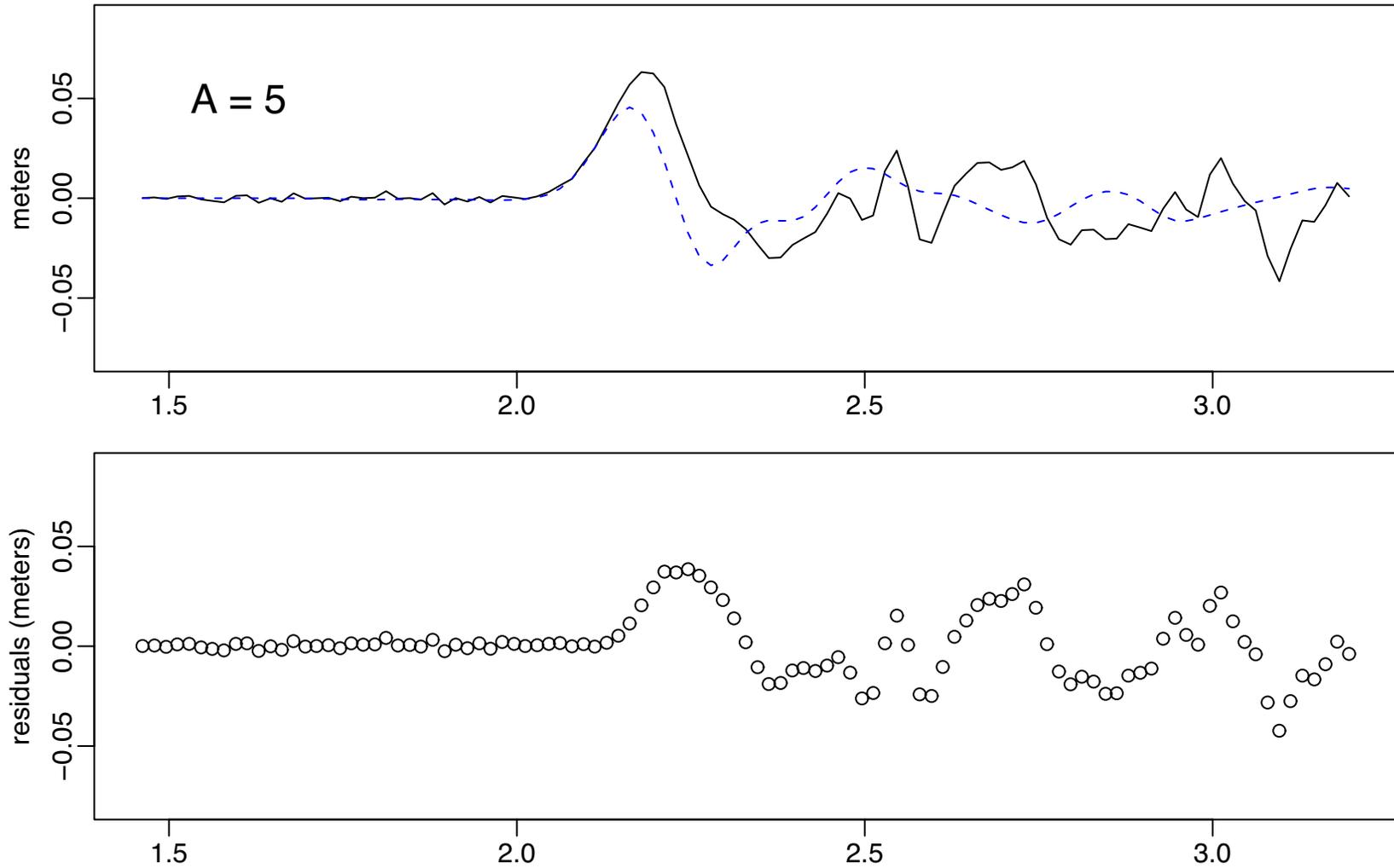
# Matching a12 Model to 21414 Data by Varying Slip



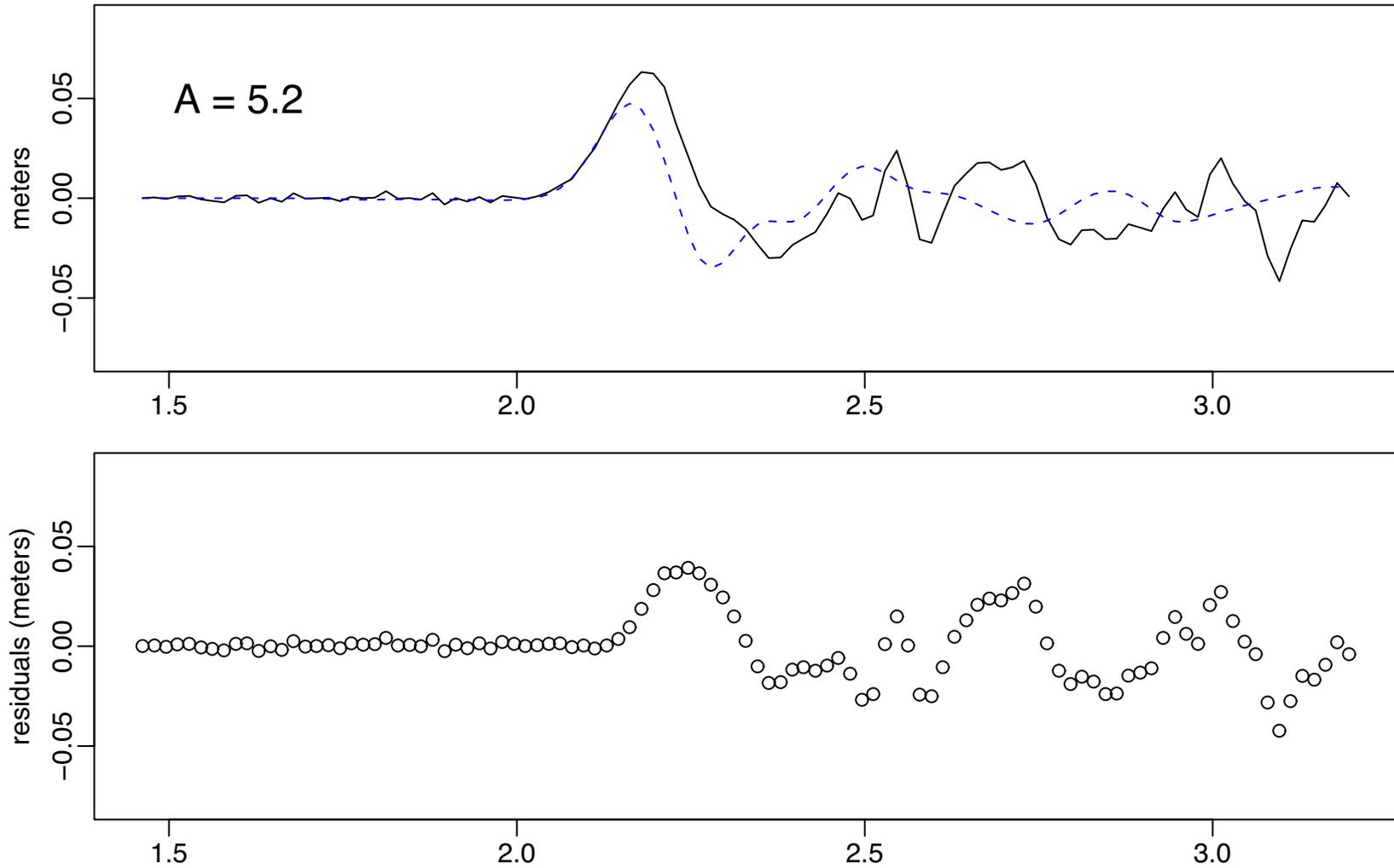
# Matching a12 Model to 21414 Data by Varying Slip



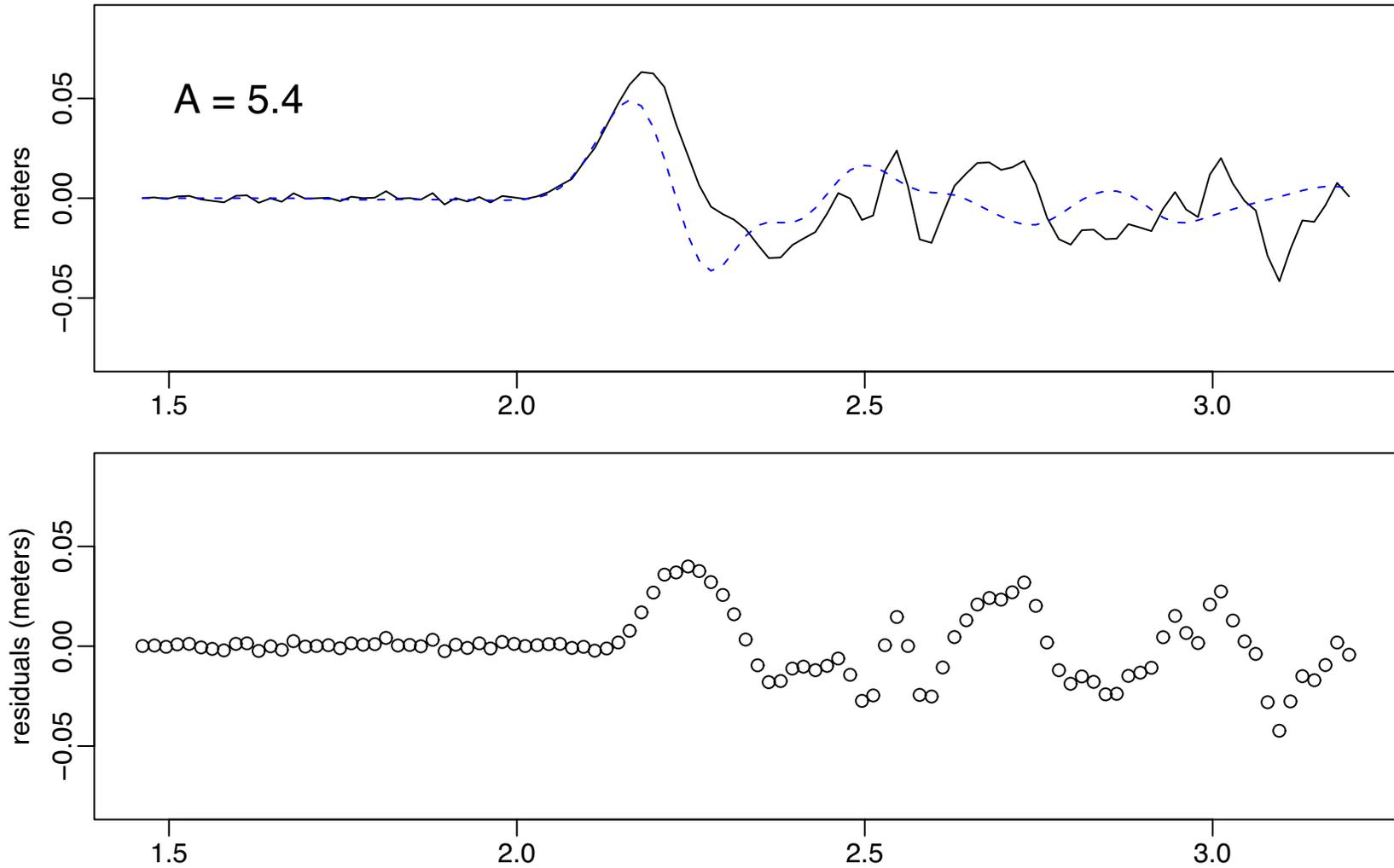
# Matching a12 Model to 21414 Data by Varying Slip



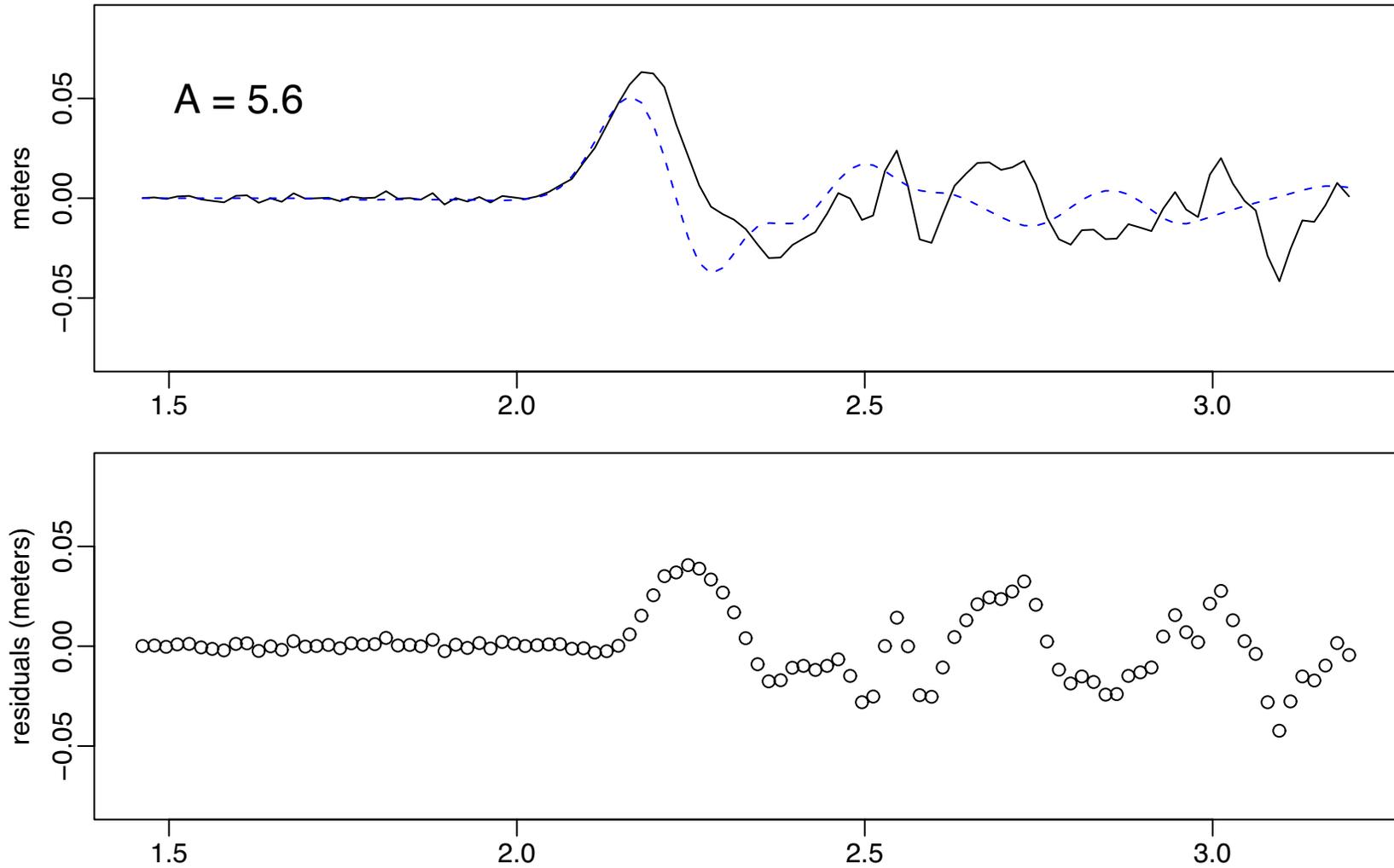
# Matching a12 Model to 21414 Data by Varying Slip



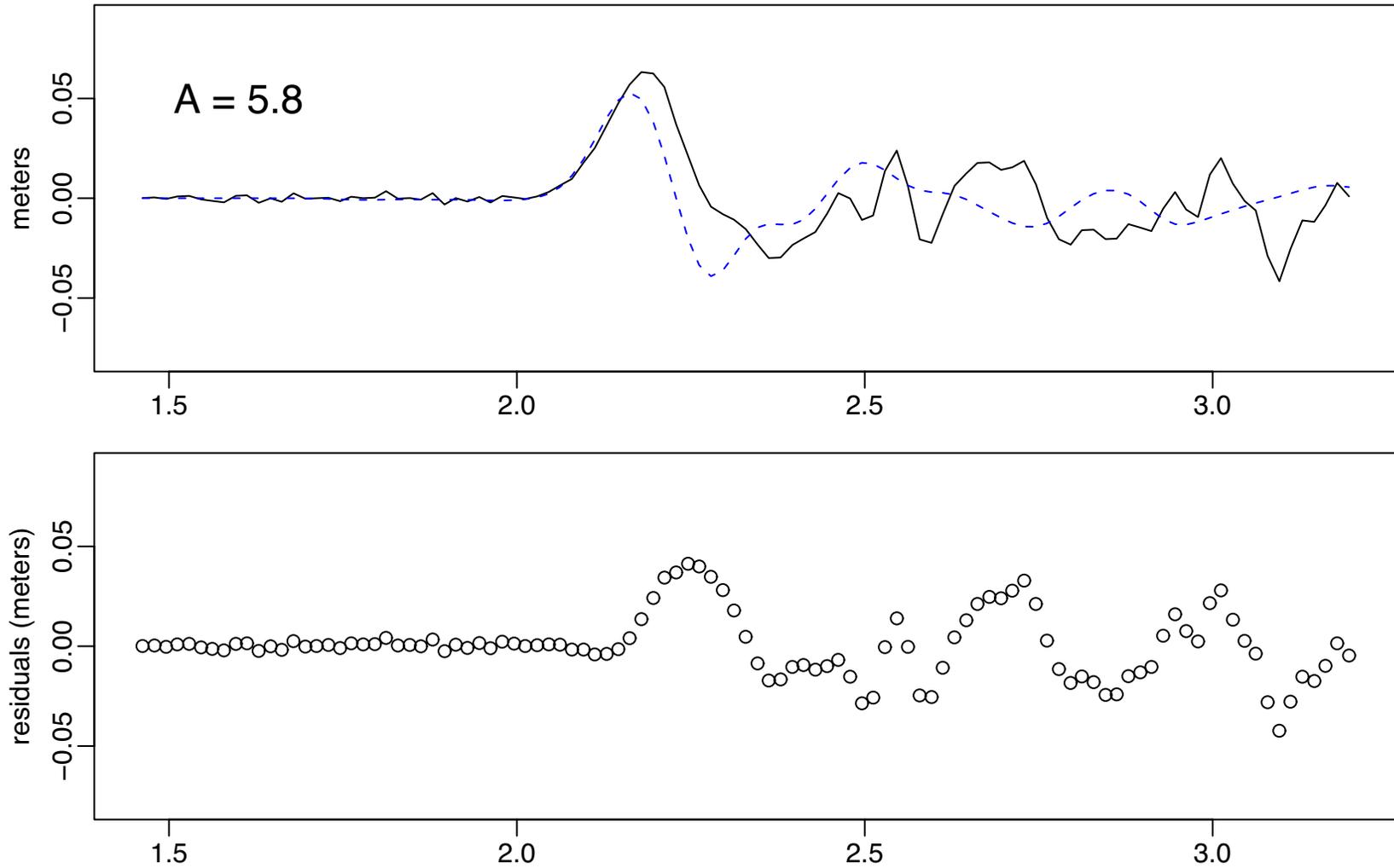
# Matching a12 Model to 21414 Data by Varying Slip



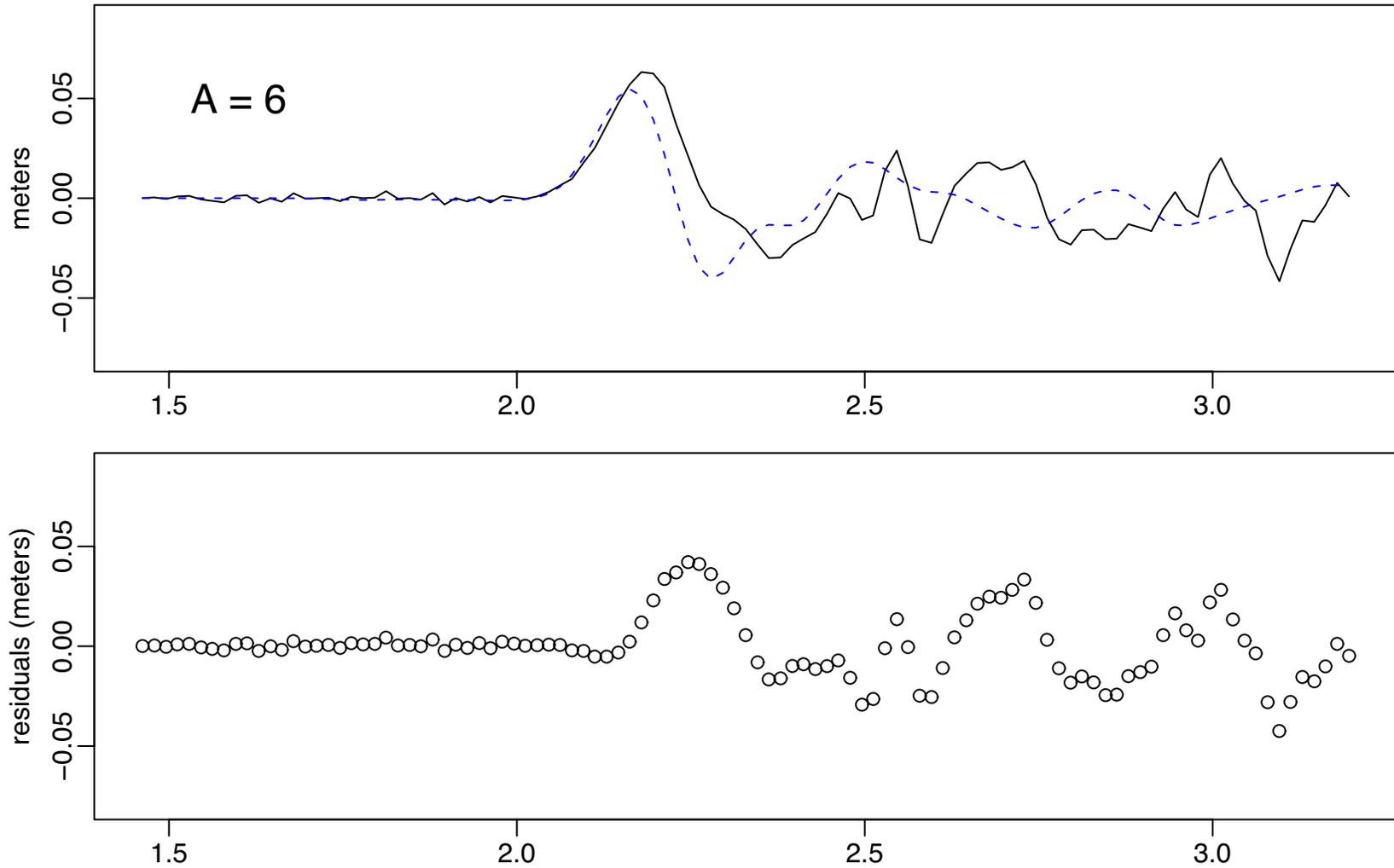
# Matching a12 Model to 21414 Data by Varying Slip



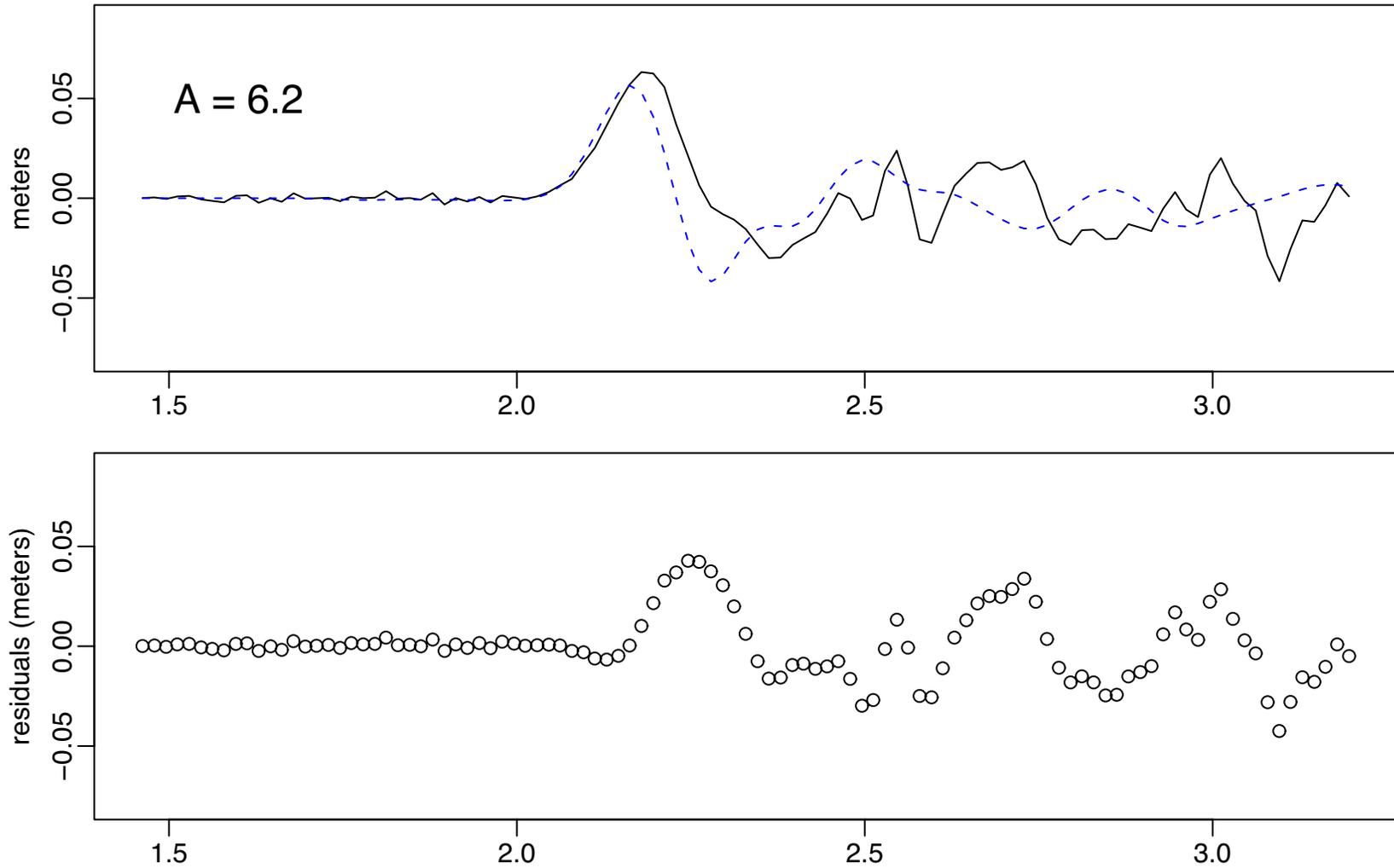
# Matching a12 Model to 21414 Data by Varying Slip



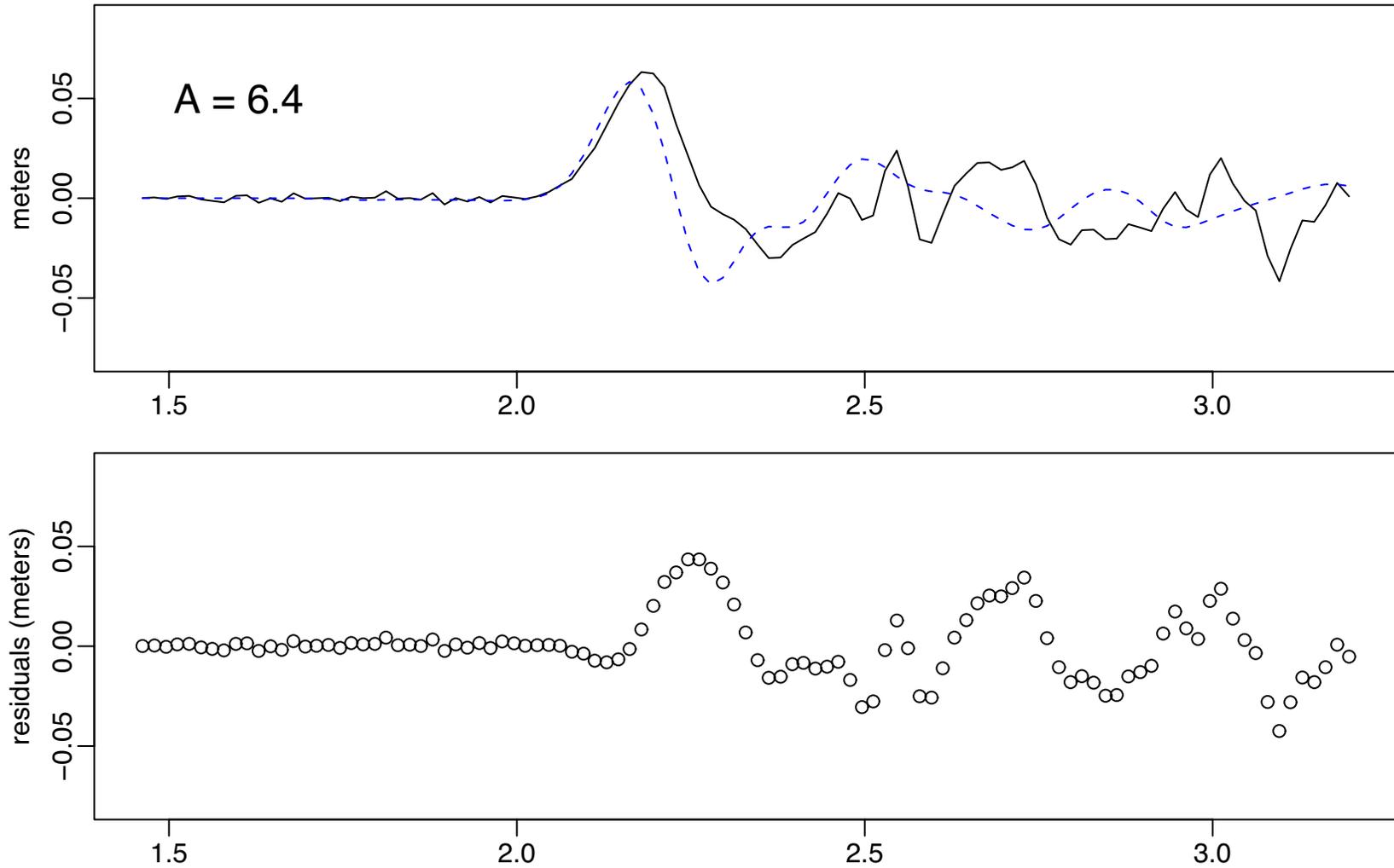
# Matching a12 Model to 21414 Data by Varying Slip



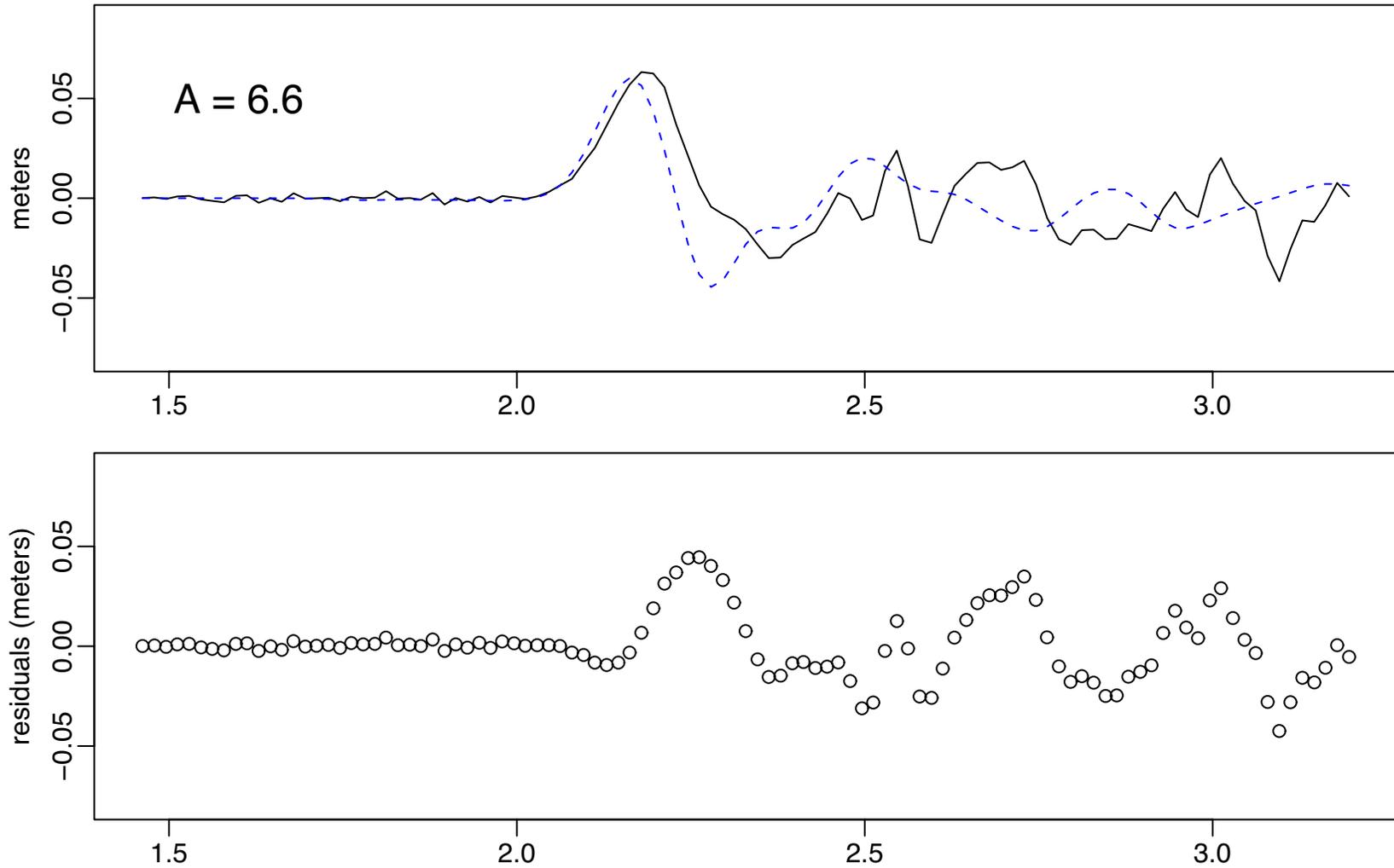
# Matching a12 Model to 21414 Data by Varying Slip



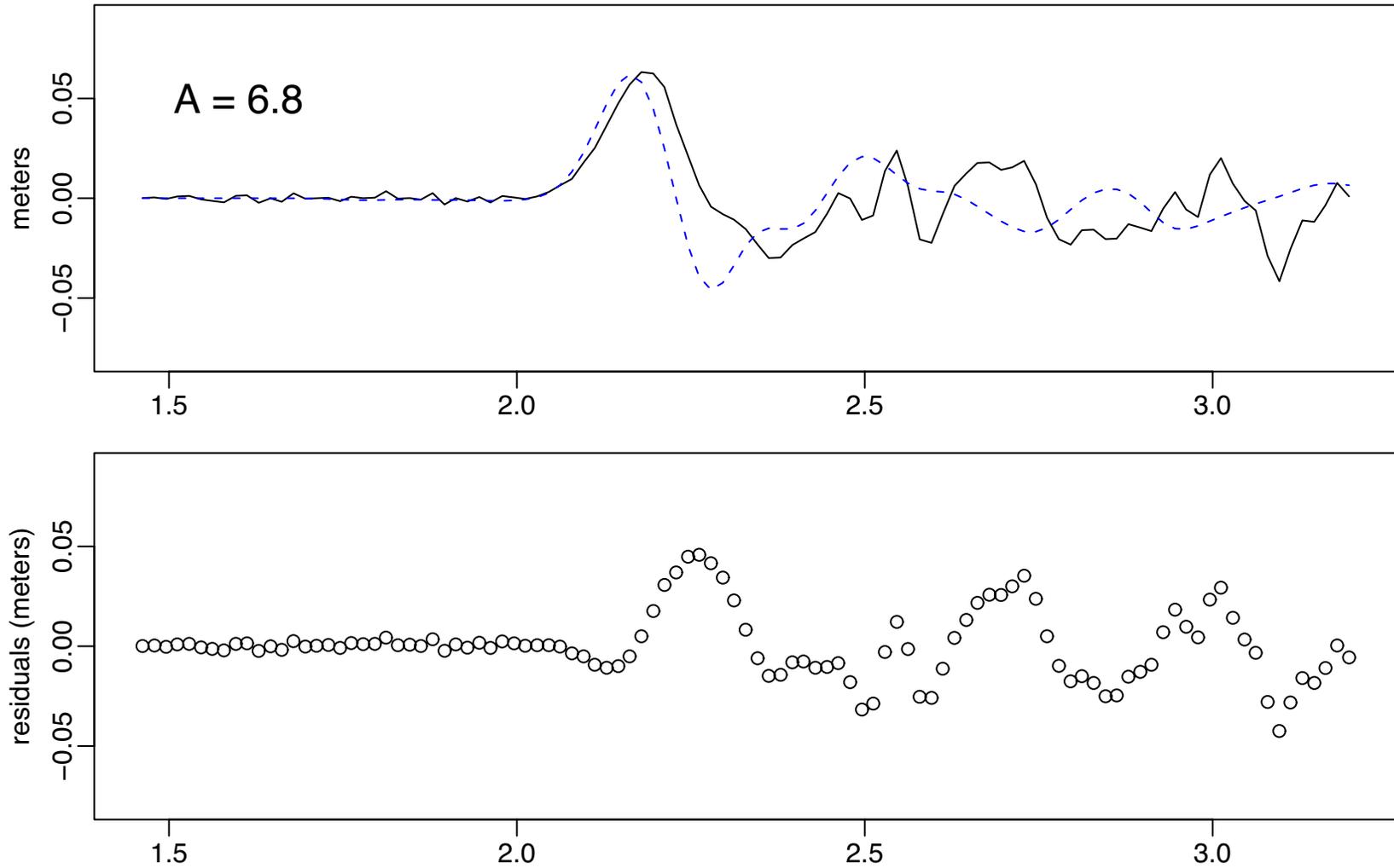
# Matching a12 Model to 21414 Data by Varying Slip



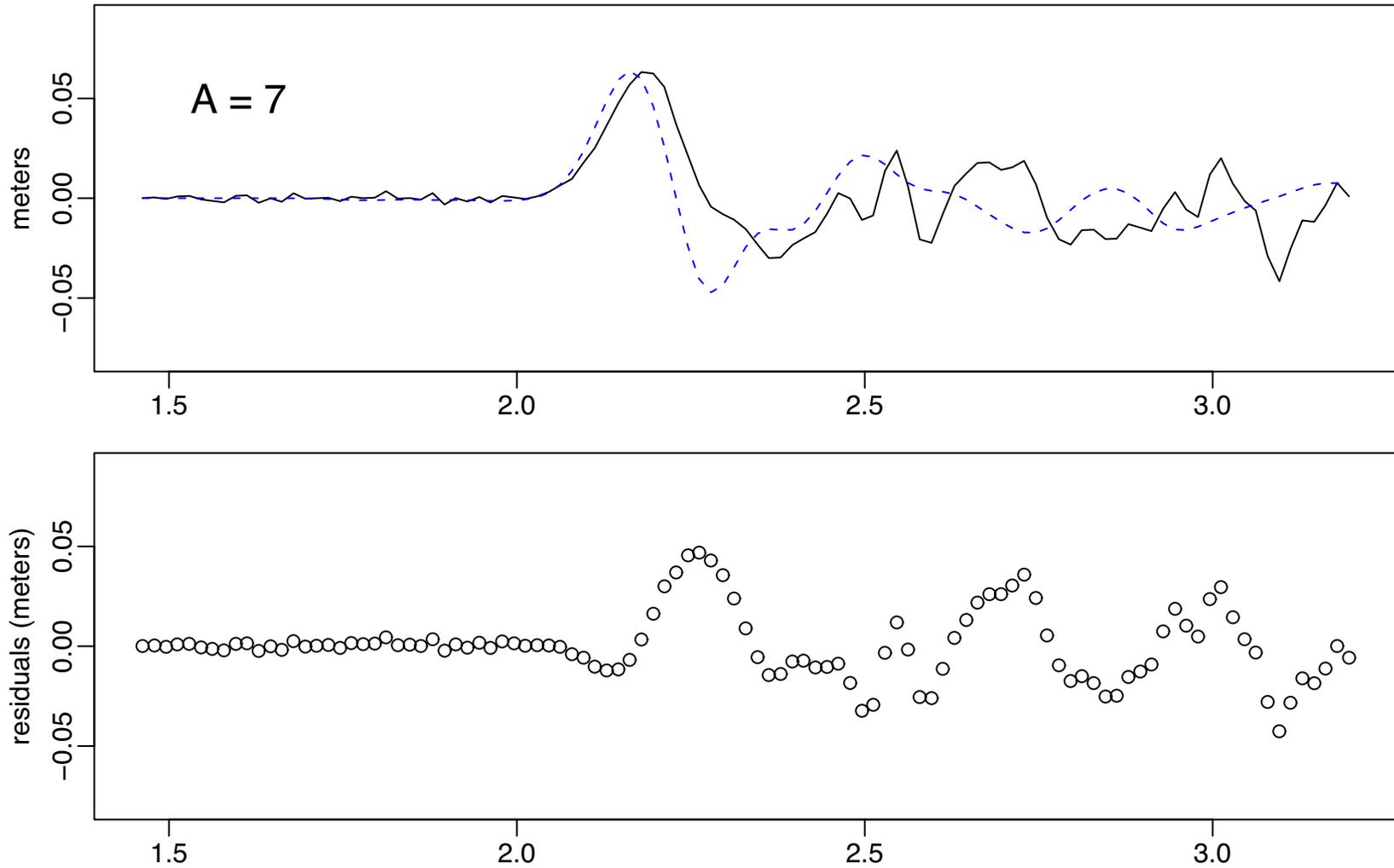
# Matching a12 Model to 21414 Data by Varying Slip



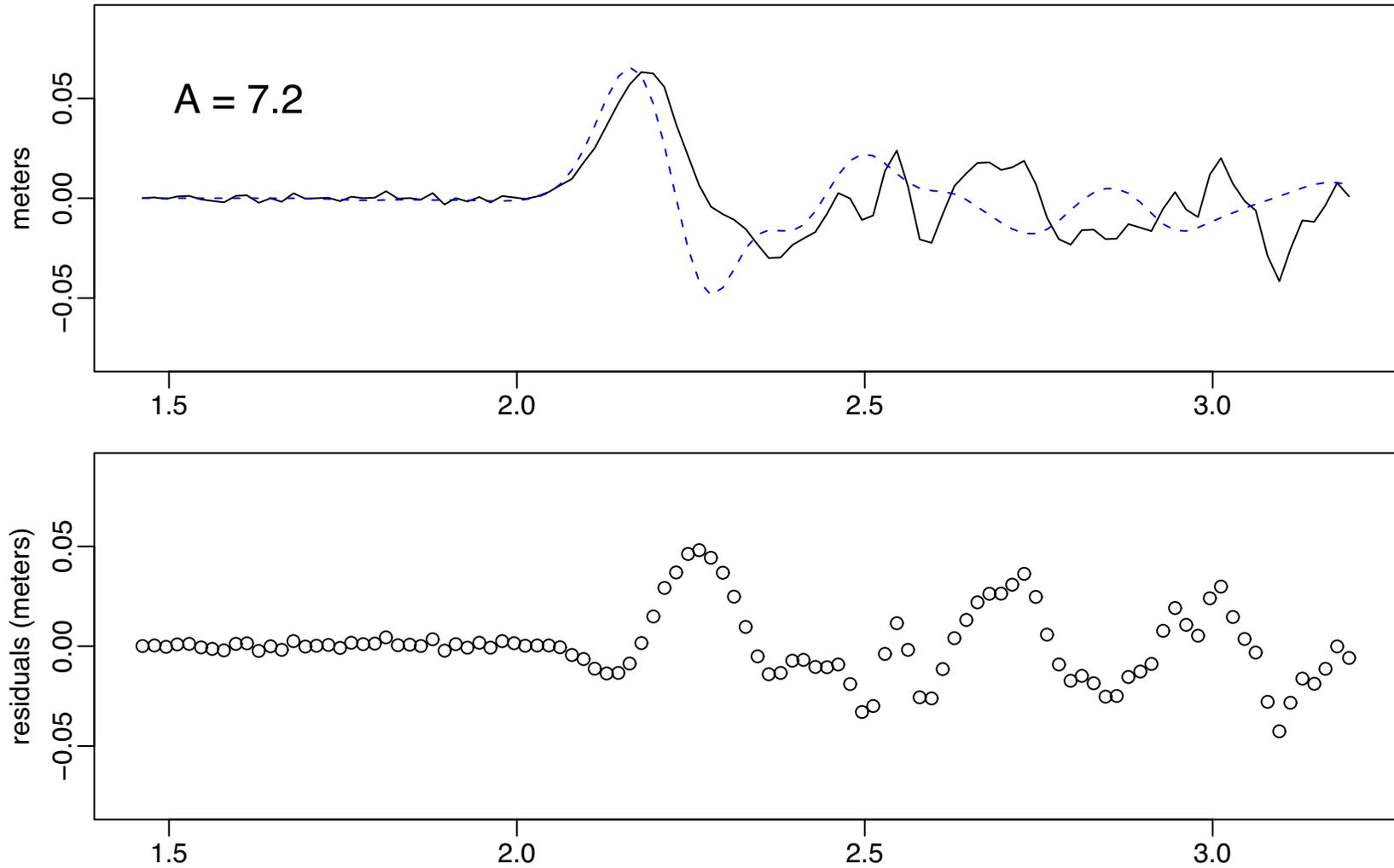
# Matching a12 Model to 21414 Data by Varying Slip



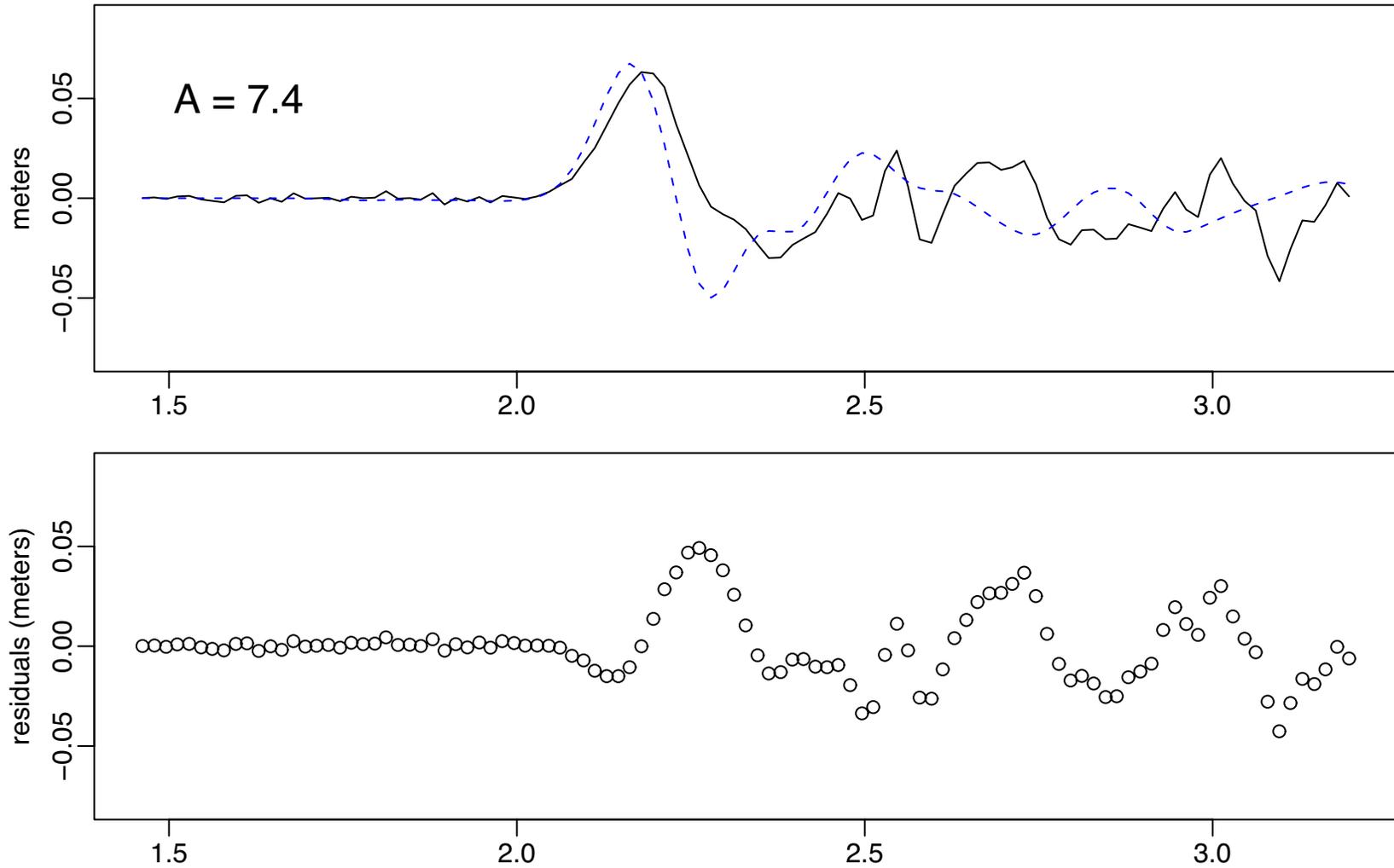
# Matching a12 Model to 21414 Data by Varying Slip



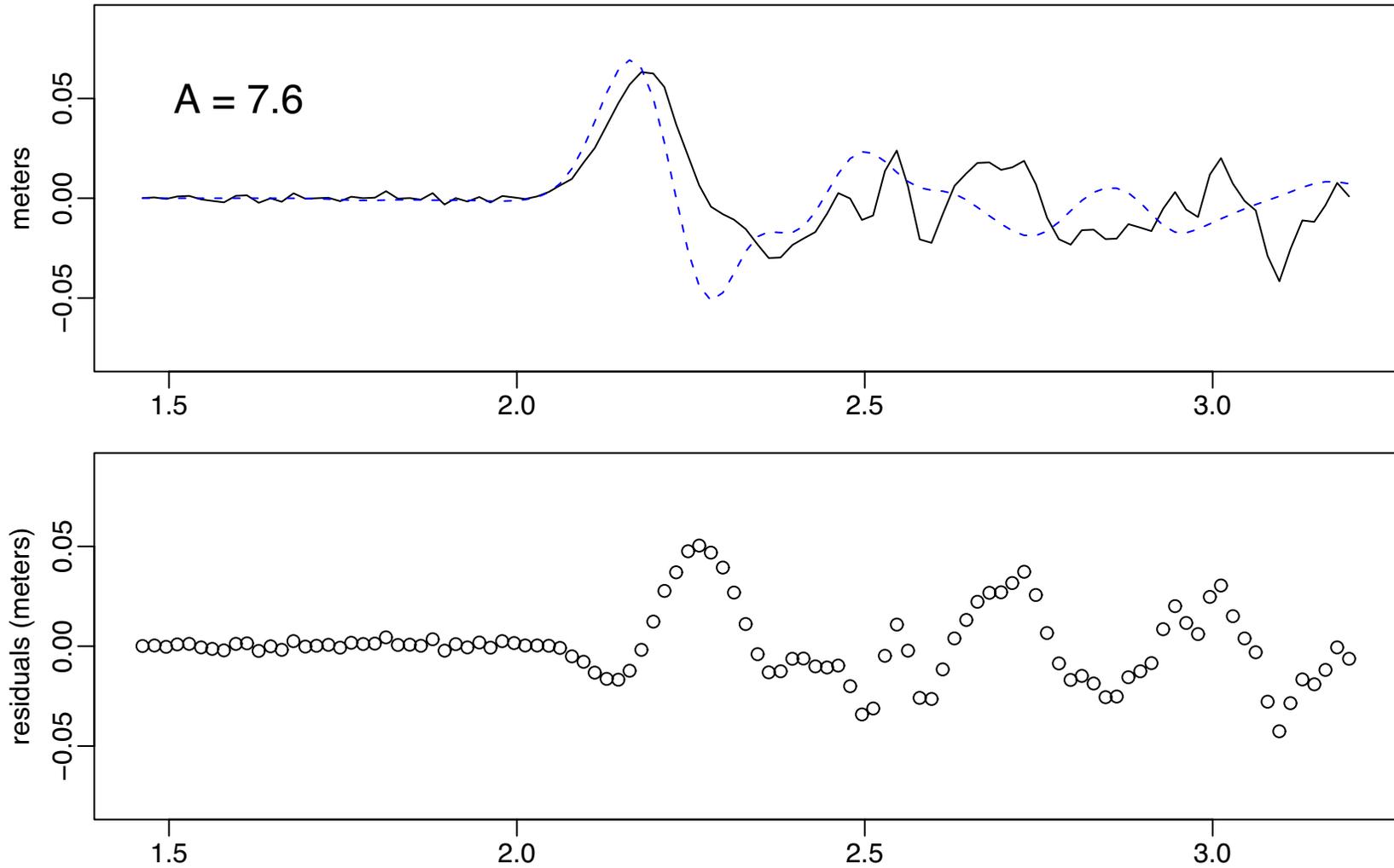
# Matching a12 Model to 21414 Data by Varying Slip



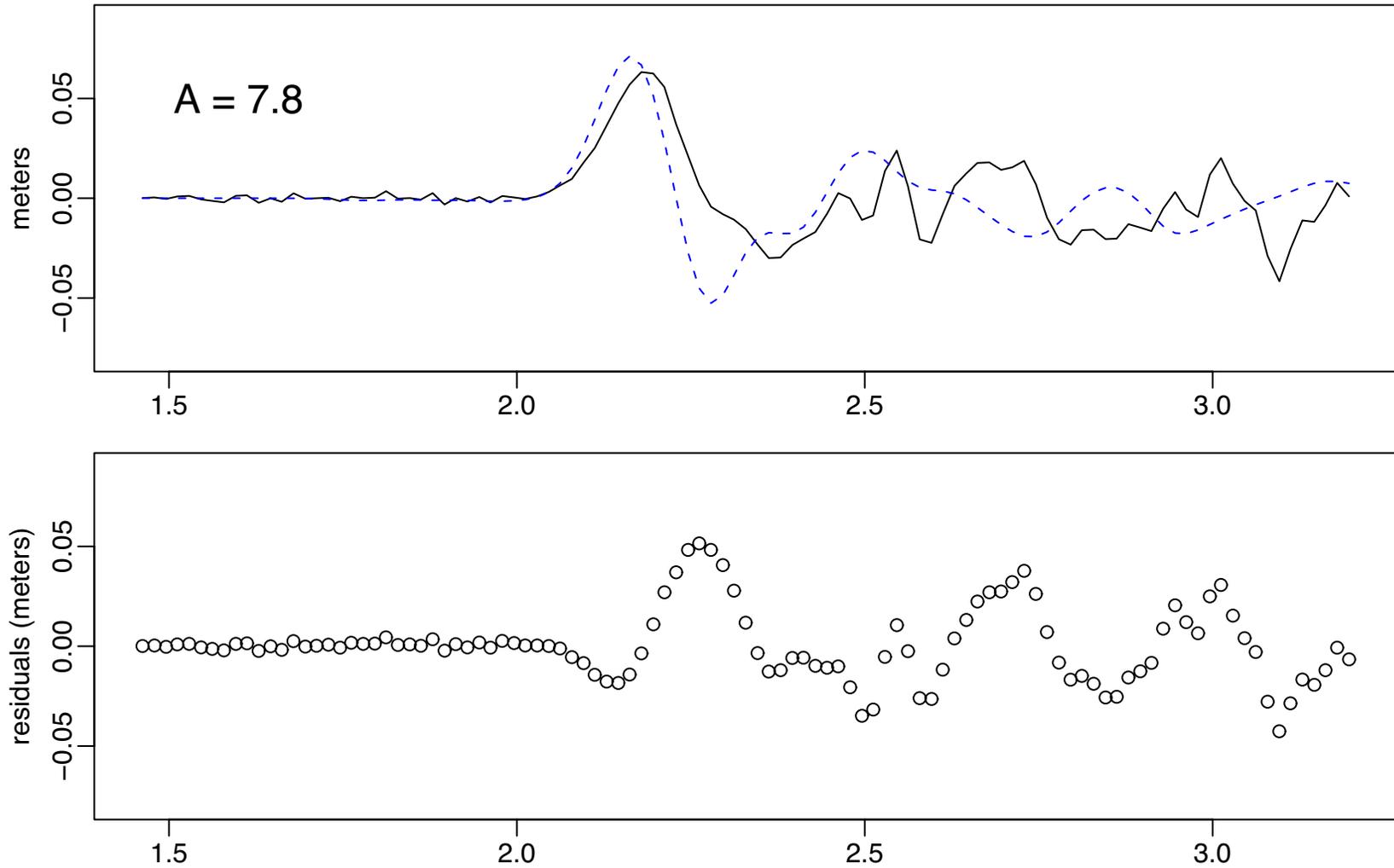
# Matching a12 Model to 21414 Data by Varying Slip



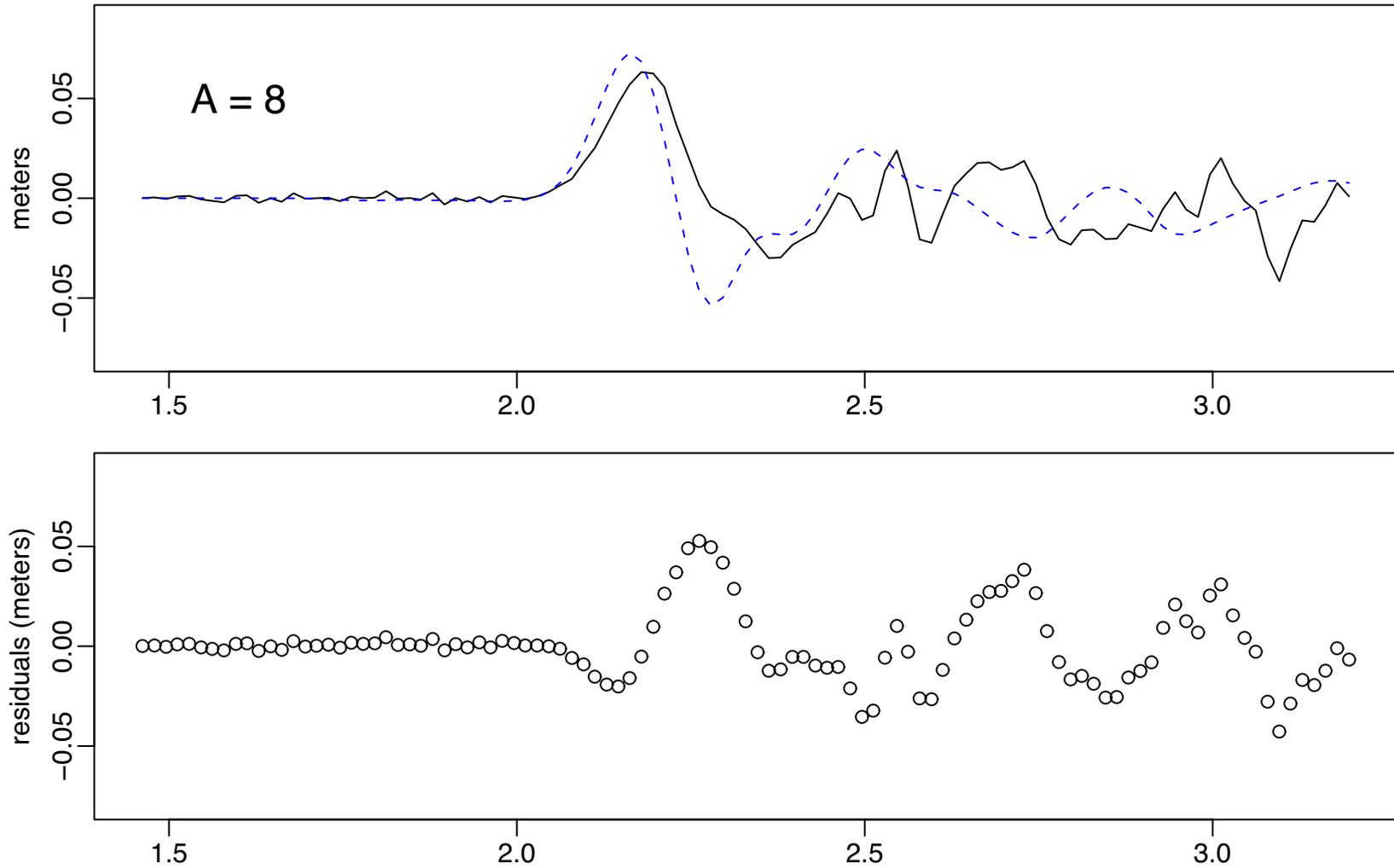
# Matching a12 Model to 21414 Data by Varying Slip



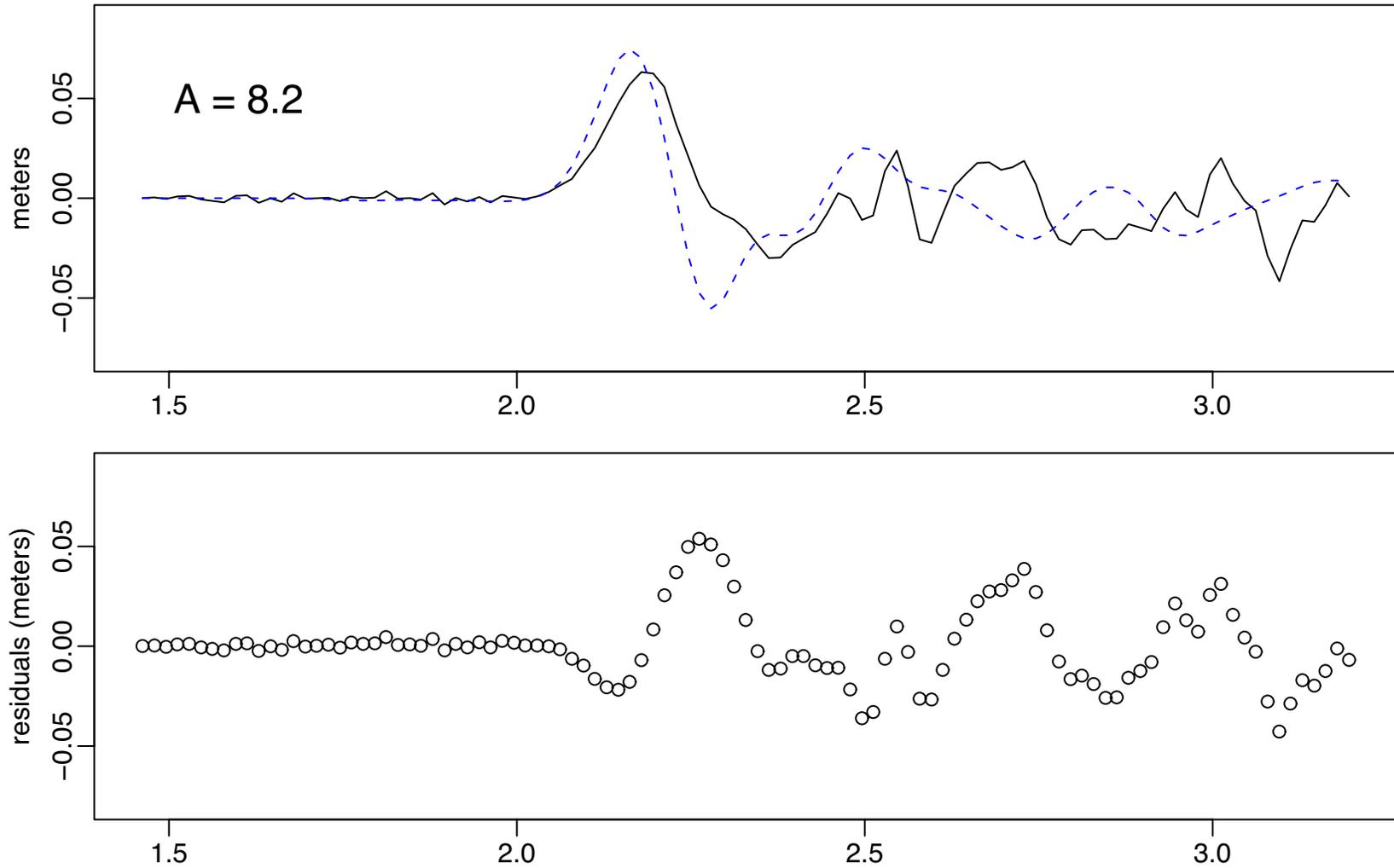
# Matching a12 Model to 21414 Data by Varying Slip



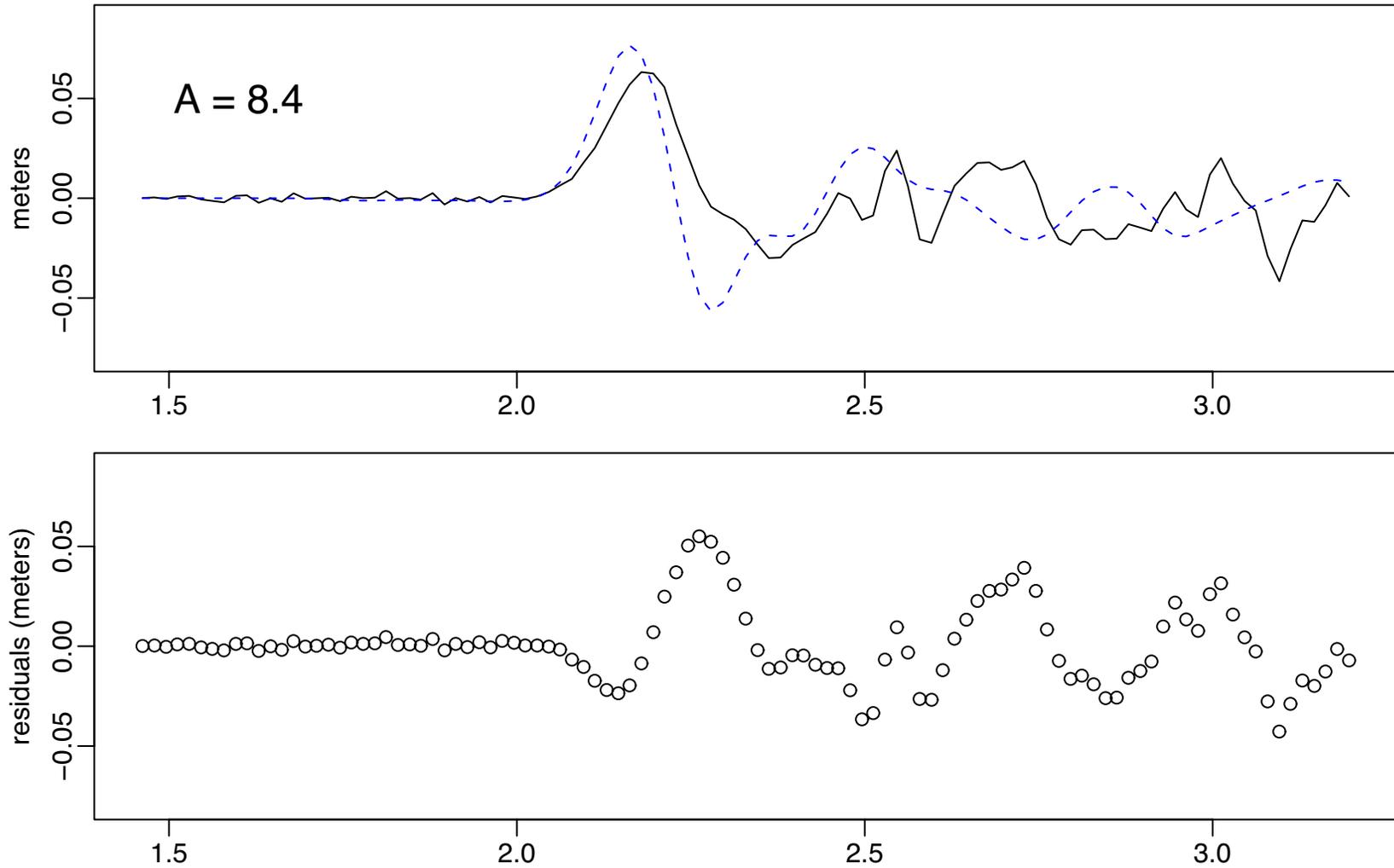
# Matching a12 Model to 21414 Data by Varying Slip



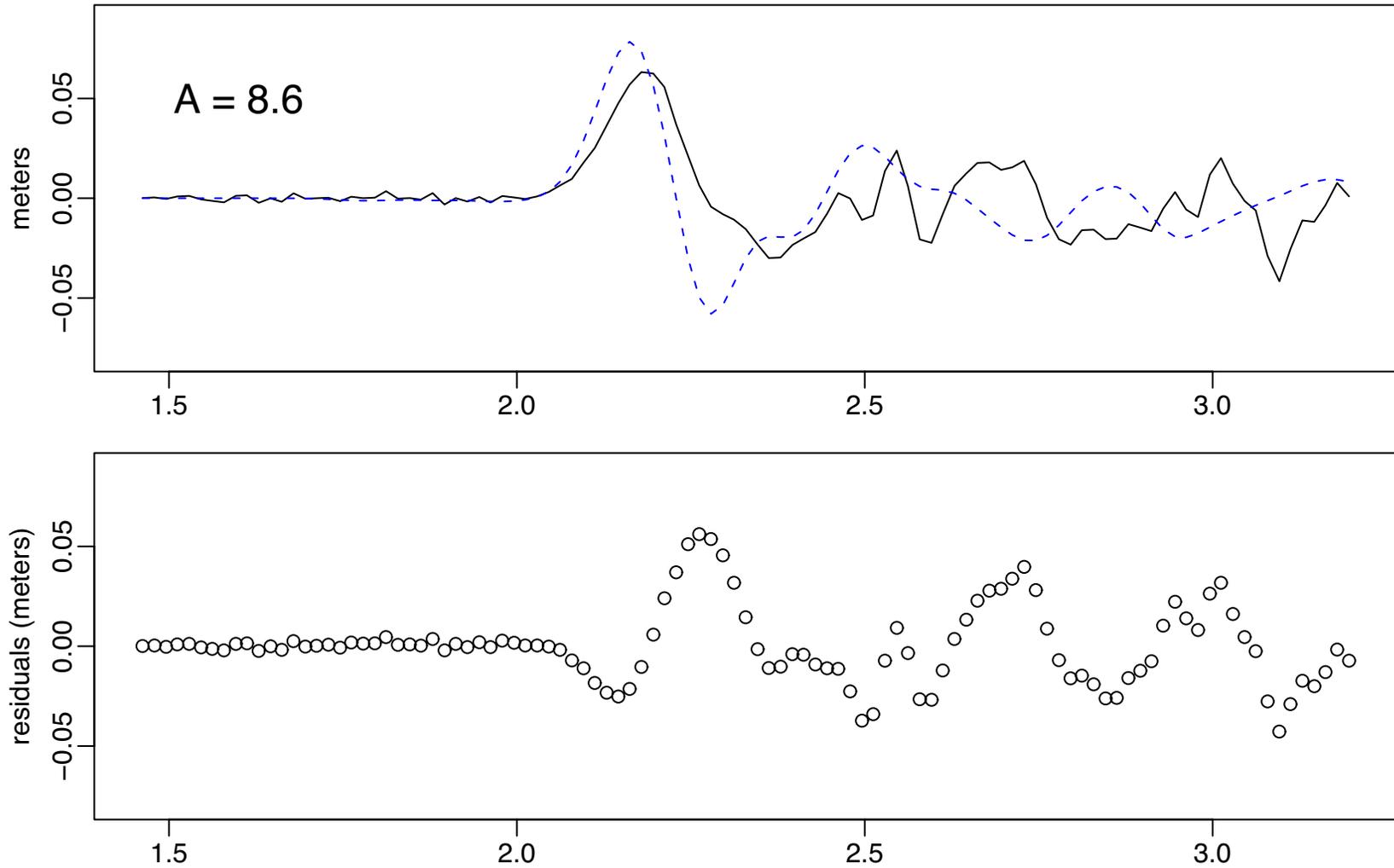
# Matching a12 Model to 21414 Data by Varying Slip



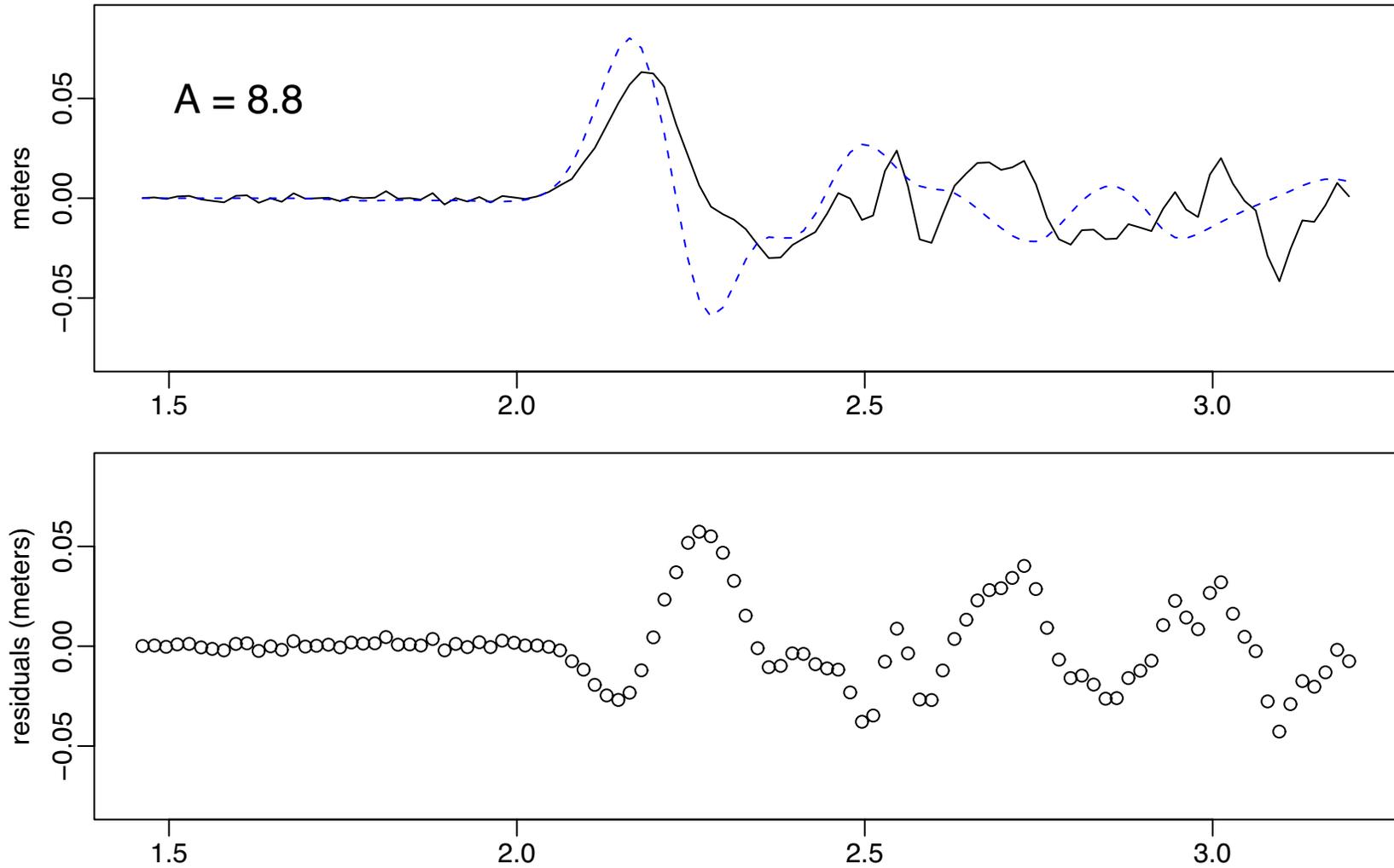
# Matching a12 Model to 21414 Data by Varying Slip



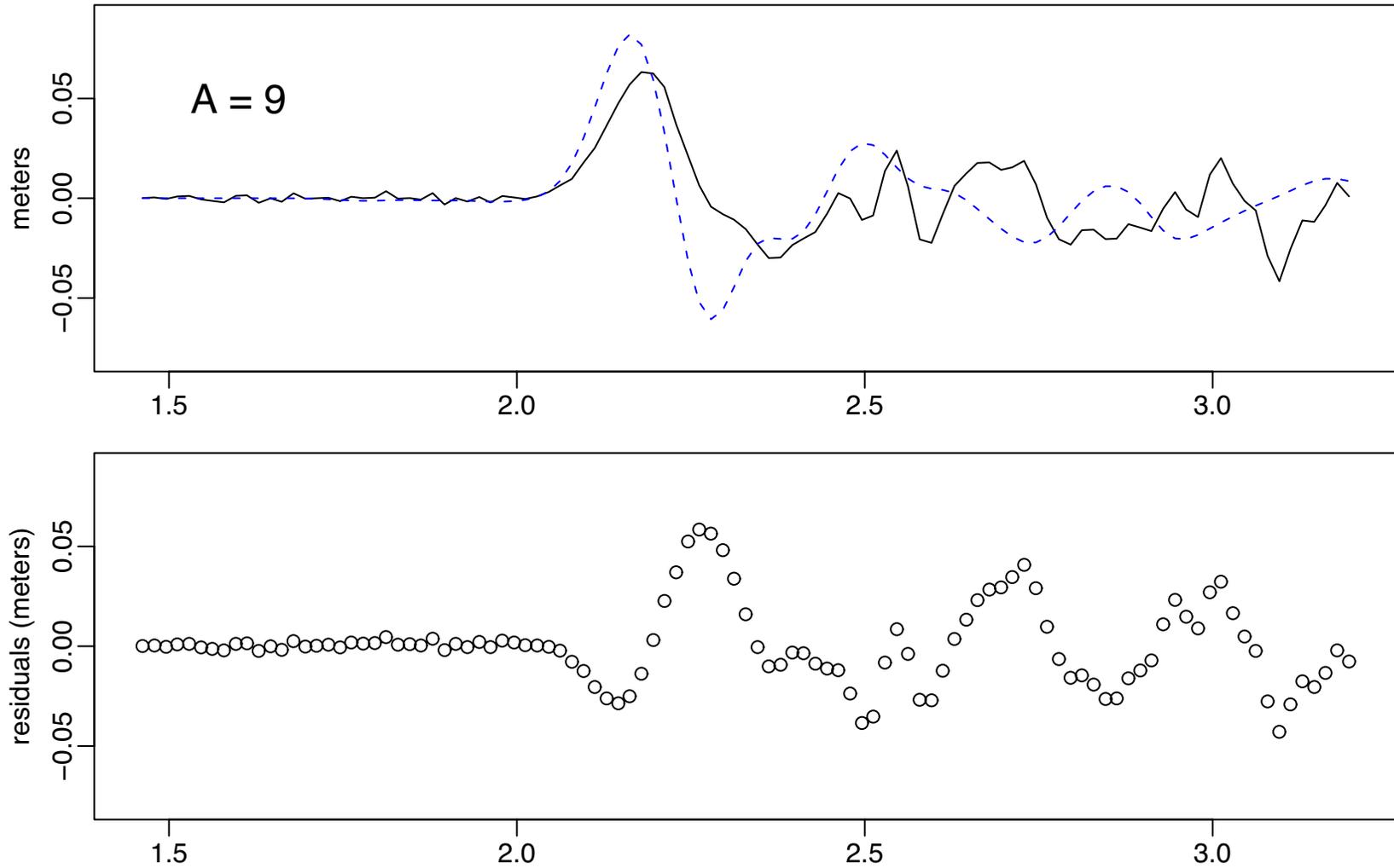
# Matching a12 Model to 21414 Data by Varying Slip



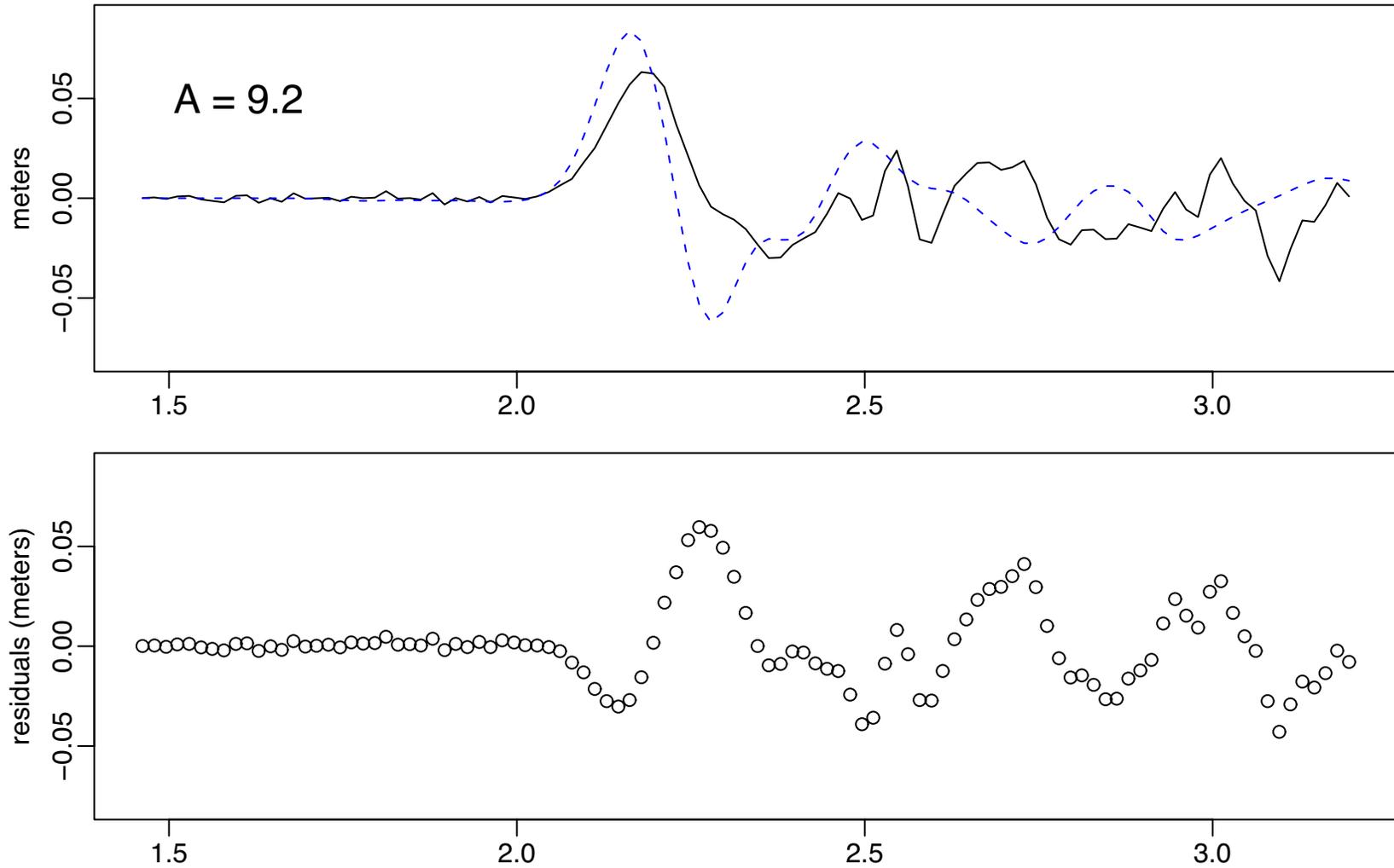
# Matching a12 Model to 21414 Data by Varying Slip



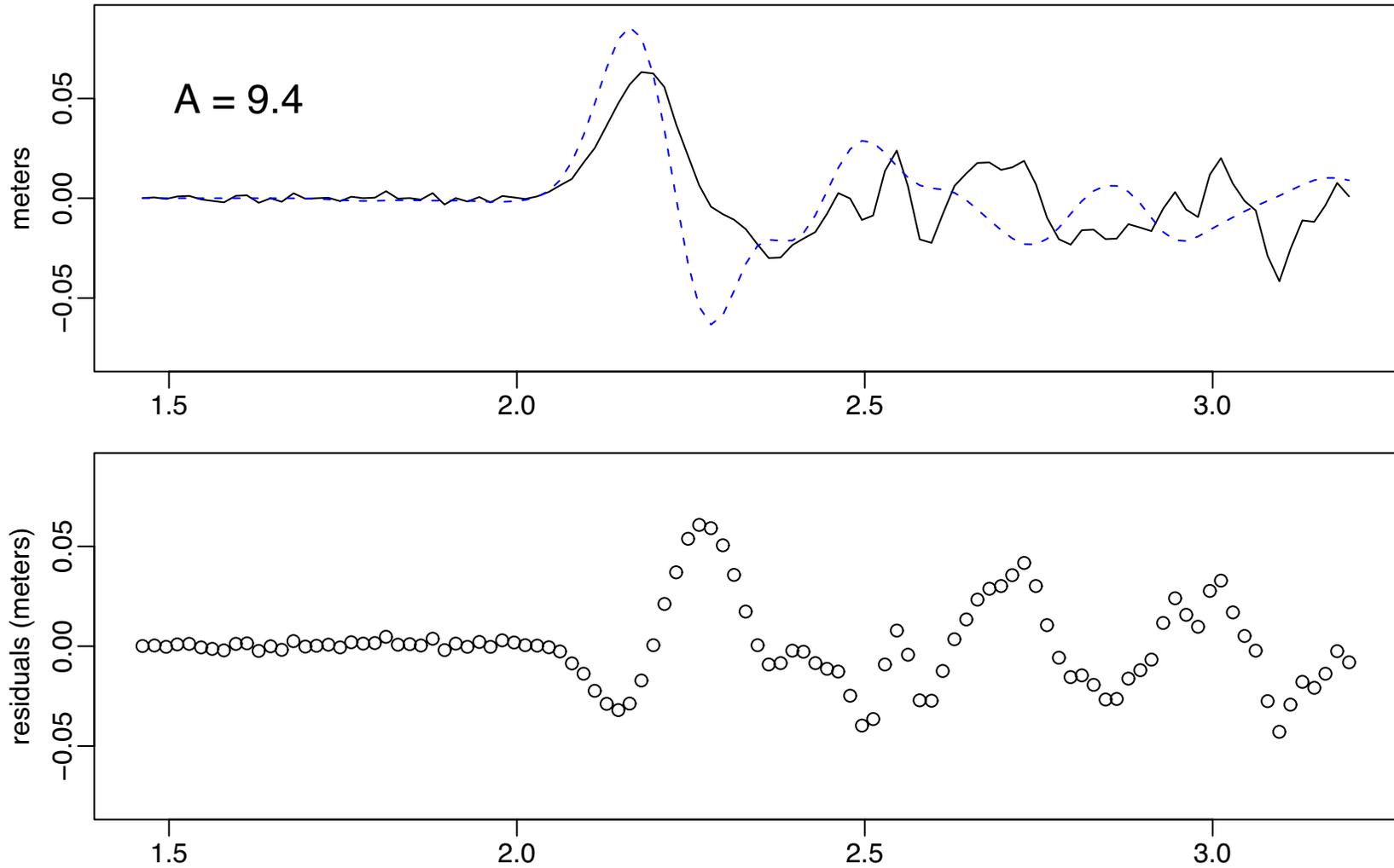
# Matching a12 Model to 21414 Data by Varying Slip



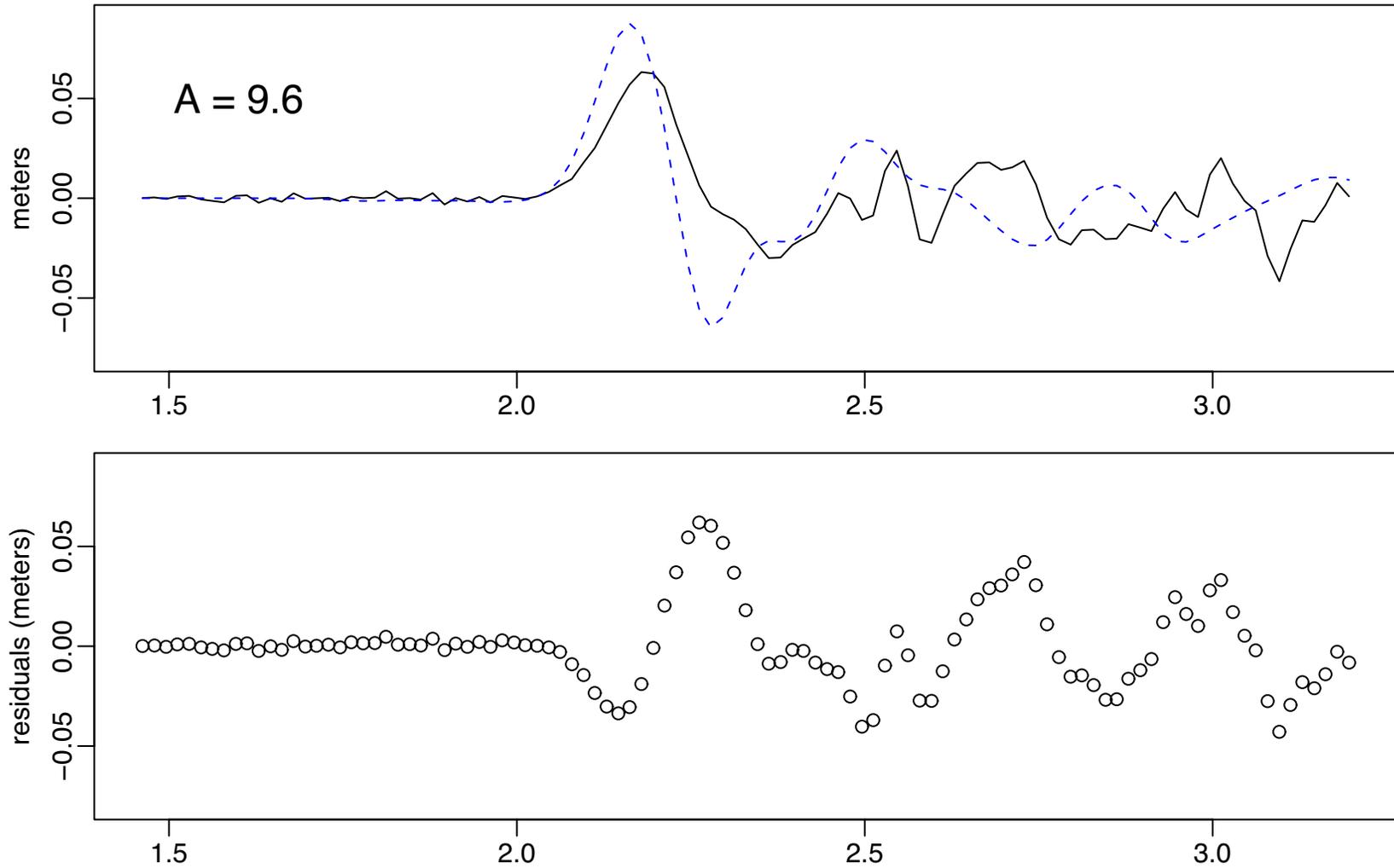
# Matching a12 Model to 21414 Data by Varying Slip



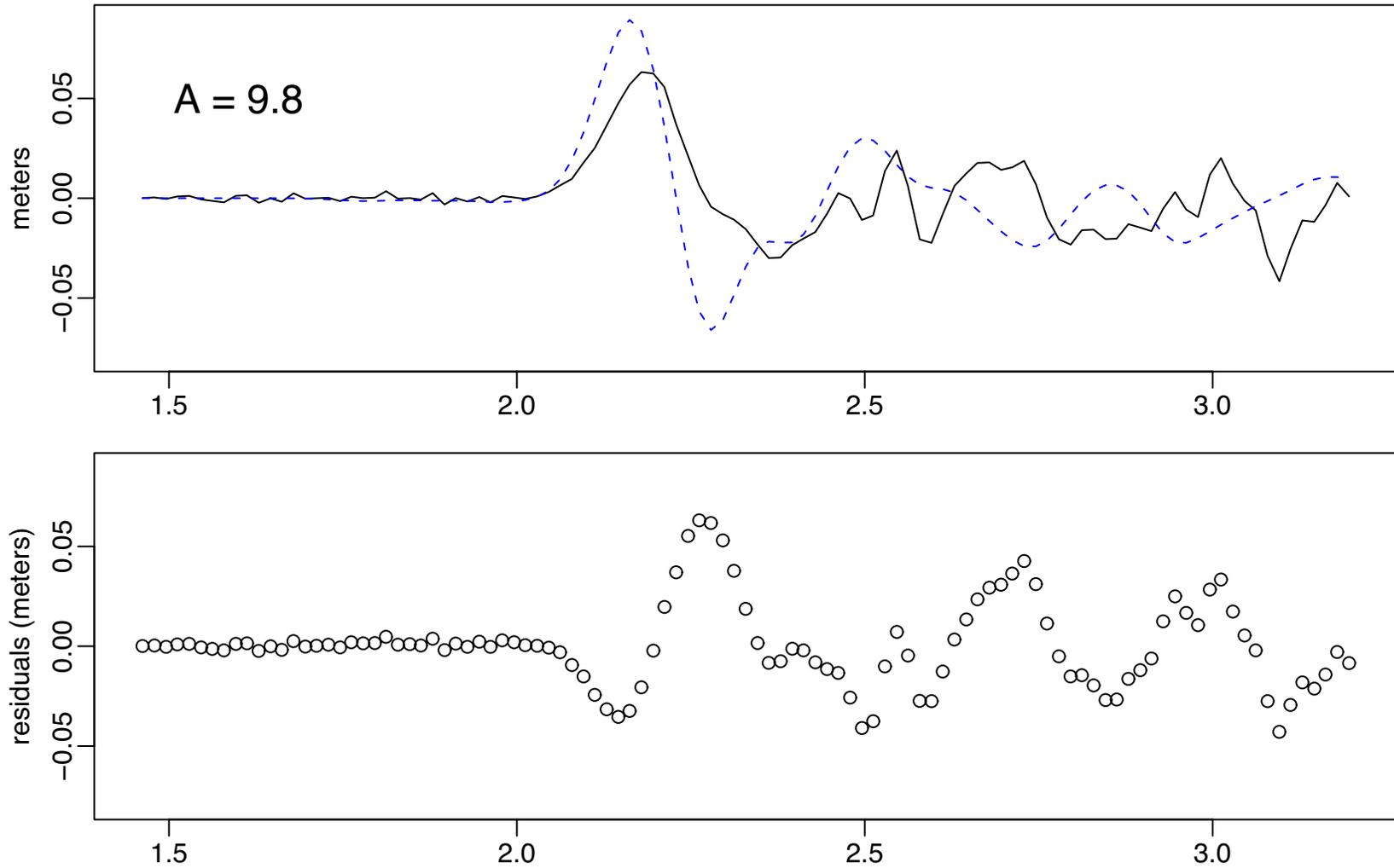
# Matching a12 Model to 21414 Data by Varying Slip



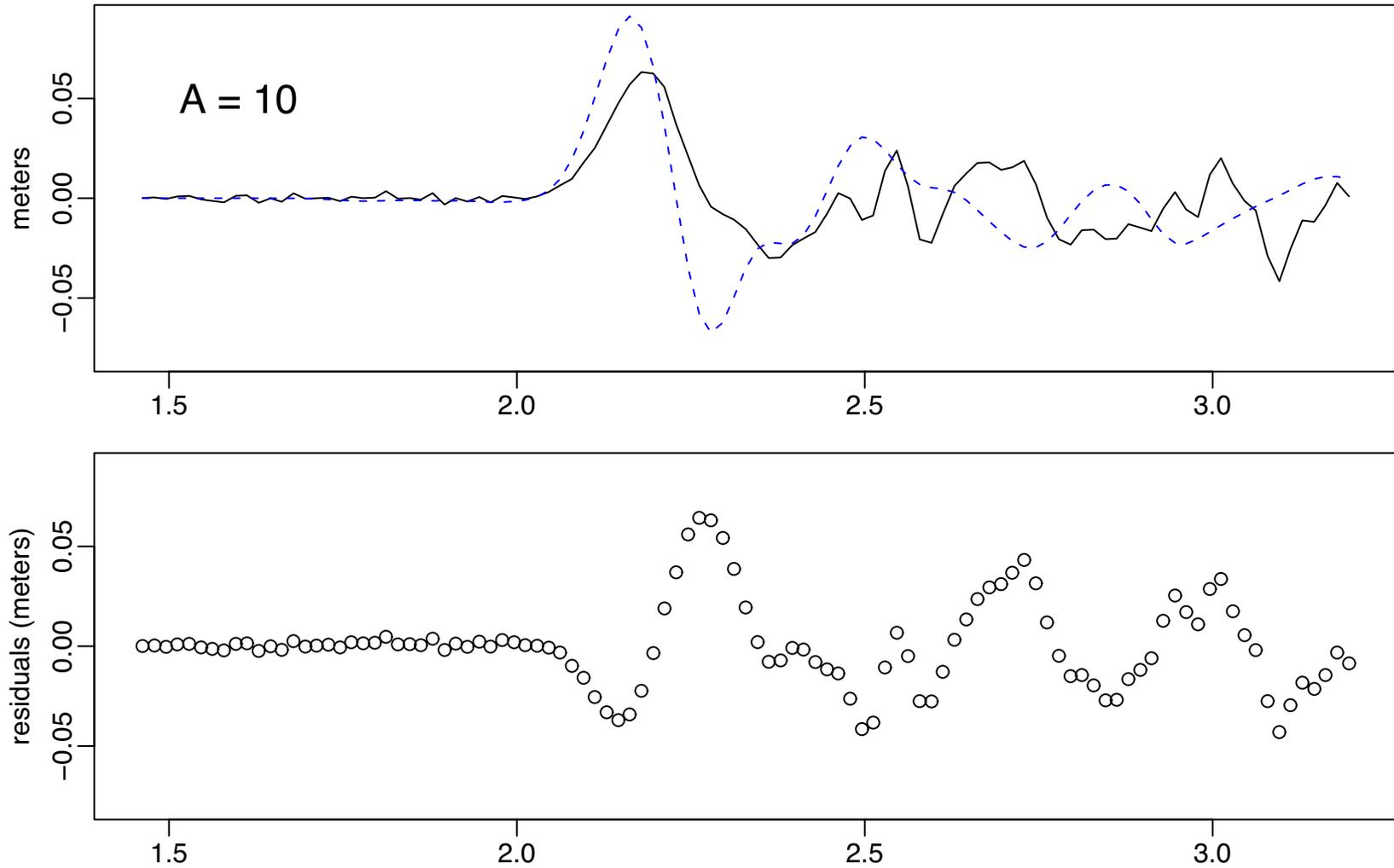
# Matching a12 Model to 21414 Data by Varying Slip



# Matching a12 Model to 21414 Data by Varying Slip



# Matching a12 Model to 21414 Data by Varying Slip



## Fitting Models to DART Buoy Data: III

- Q: what is the ‘best’ choice for  $A$ ?
- set  $A$  such that residuals  $e_t$  are ‘small’ by some measure
- many measures are possible – here are three common ones:

– make sum of squared residuals as small as possible:

$$\sum_t e_t^2 = \sum_t [x_t - A \cdot g(t)]^2 \equiv f_2(A)$$

– make sum of magnitudes of residuals as small as possible:

$$\sum_t |e_t| = \sum_t |x_t - A \cdot g(t)| \equiv f_1(A)$$

– make largest magnitude of residuals as small as possible:

$$\max_t |e_t| = \max_t |x_t - A \cdot g(t)| \equiv f_\infty(A)$$

## Fitting Models to DART Buoy Data: IV

- here is a specialized one:
  - make sum of squared residuals at peak and trough as small as possible:

$$e_{t_0}^2 + e_{t_1}^2 = [x_{t_0} - A \cdot g(t_0)]^2 + [x_{t_1} - A \cdot g(t_1)]^2 \equiv f_{pt}(A),$$

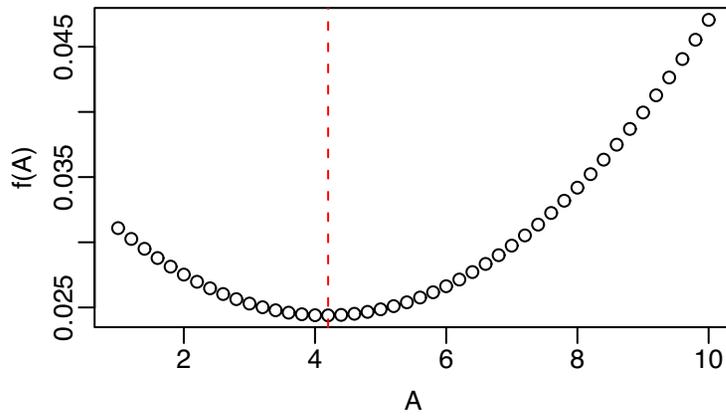
where  $t_0$  and  $t_1$  are such that

$$g(t_0) = \max_t \{g(t)\} \quad \text{and} \quad g(t_1) = \min_t \{g(t)\}$$

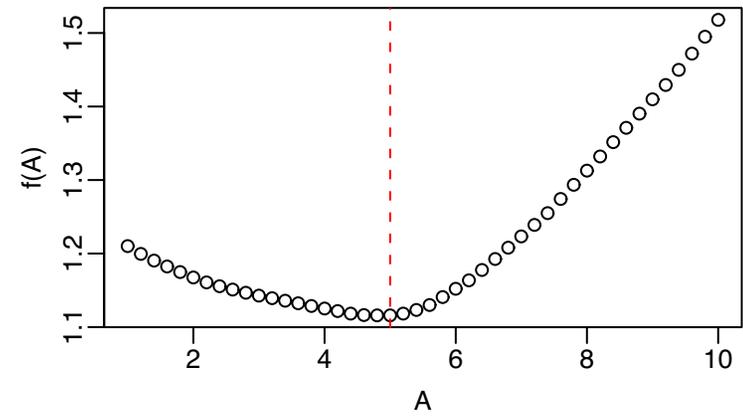
- let's look at plots of  $f(A)$  versus  $A$  for the four measures, where, as before,  $A = 1.0, 1.2, 1.4, \dots, 10.0$

# Four Residual Measures $f(A)$ versus Slip $A$

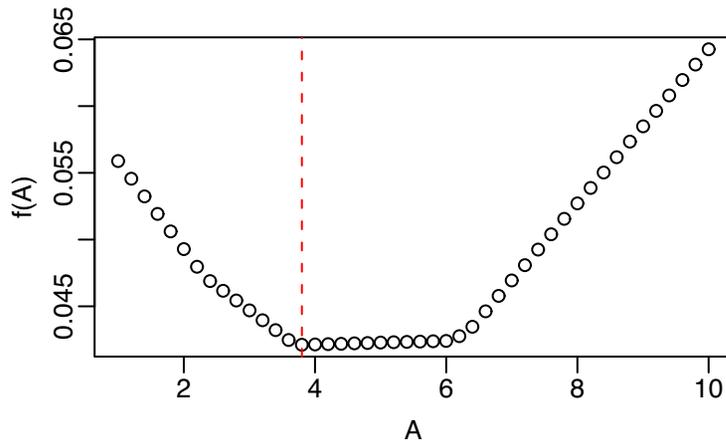
**f\_2**



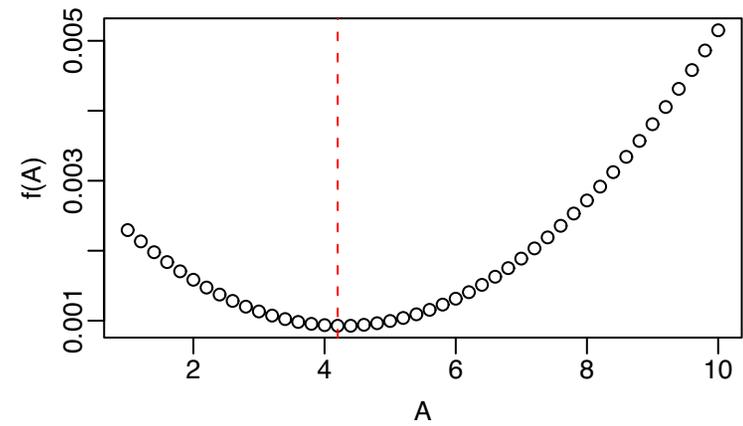
**f\_1**



**f\_inf**



**f\_pt**



## Fitting Models to DART Buoy Data: V

- estimated slips are  $\hat{A}_2 = 4.2$ ,  $\hat{A}_1 = 5$ ,  $\hat{A}_\infty = 3.8$  and  $\hat{A}_{pt} = 4.2$
- two measures based on least squares, i.e.,  $f_2(A)$  and  $f_{pt}(A)$ , have certain advantages, including:

- no need to do grid search because location of minimum of

$$f(A) \equiv \sum_t [x_t - A \cdot g(t)]^2$$

is given by a simple formula:

$$\hat{A}_{ls} = \frac{\sum_t x_t g(t)}{\sum_t [g(t)]^2}, \text{ here yielding } \hat{A}_2 \doteq 4.17 \text{ and } \hat{A}_{pt} \doteq 4.26$$

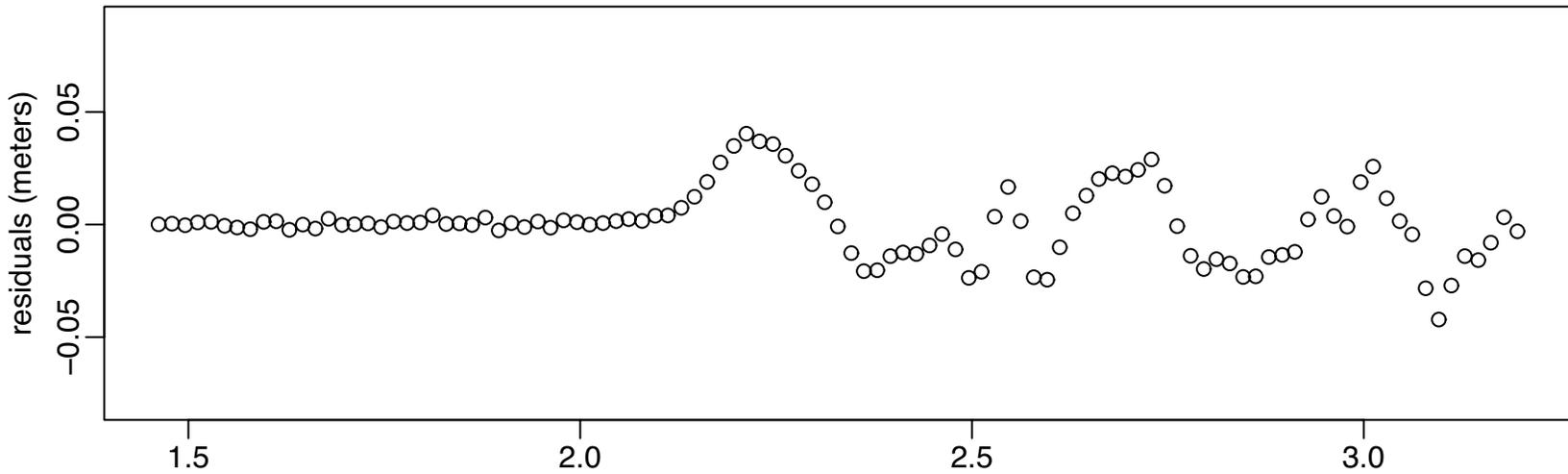
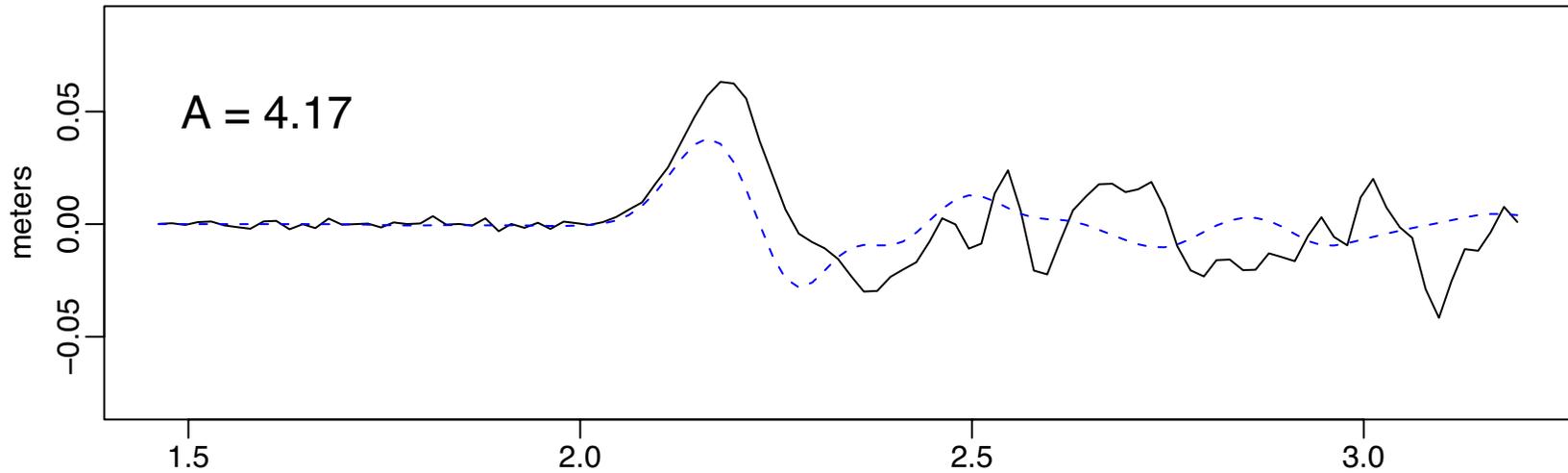
(follows from taking equation  $f'(A) = 0$  and solving for  $A$ )

- statistical variation in  $\hat{A}_{ls}$  easy to quantify

## Assessing Variability in $\hat{A}_{ls}$ : I

- reformulate model  $x_t = A \cdot g(t) + e_t$  in vector notation as  $\mathbf{x} = A\mathbf{g} + \mathbf{e}$ , where  $\mathbf{x}$  is column vector containing the  $x_t$ 's etc.
- least squares estimate of  $A$  is  $\hat{A}_{ls} = \mathbf{g}^T \mathbf{x} / \mathbf{g}^T \mathbf{g}$
- need to consider statistical properties of residuals  $e_t$
- if residuals were Gaussian (normally) distributed and uncorrelated with a common variance  $\sigma_e^2$ , then  $\hat{A}_{ls}$  is Gaussian distributed with mean  $A$  and variance  $\sigma_e^2 / \mathbf{g}^T \mathbf{g}$
- allows us to compute standard deviations (SDs) and to write  $\hat{A}_2 = 4.17 \pm 0.59$  and  $\hat{A}_{pt} = 4.26 \pm 2.69$  (note size of SDs)
- assumptions of uncorrelatedness and common variance are dicey, as can be seen from plot of residuals associated with  $\hat{A}_2$

# Least Squares Estimate $\hat{A}_2$ of Slip for a12 & 21414



## Assessing Variability in $\hat{A}_{ls}$ : II

- assumption of common variance of  $e_t$ 's not viable for data before first wave, but, because  $g(t) = 0$  there, these data have no effect on estimate  $\hat{A}_{ls} = \sum_t x_t g(t) / \sum_t [g(t)]^2$
- while assumption of common variance reasonable for data beginning at first wave, assumption of uncorrelatedness is not
- can model correlation using a first order autoregressive process:

$$e_t = \phi e_{t-1} + w_t,$$

where  $w_t$  is Gaussian white noise

- implies that correlation between  $e_t$  and  $e_{t+\tau}$  given by  $\phi^{|\tau|}$
- estimate of  $\phi$  via correlation between  $e_t$  &  $e_{t+1}$  yields  $\hat{\phi} \doteq 0.86$

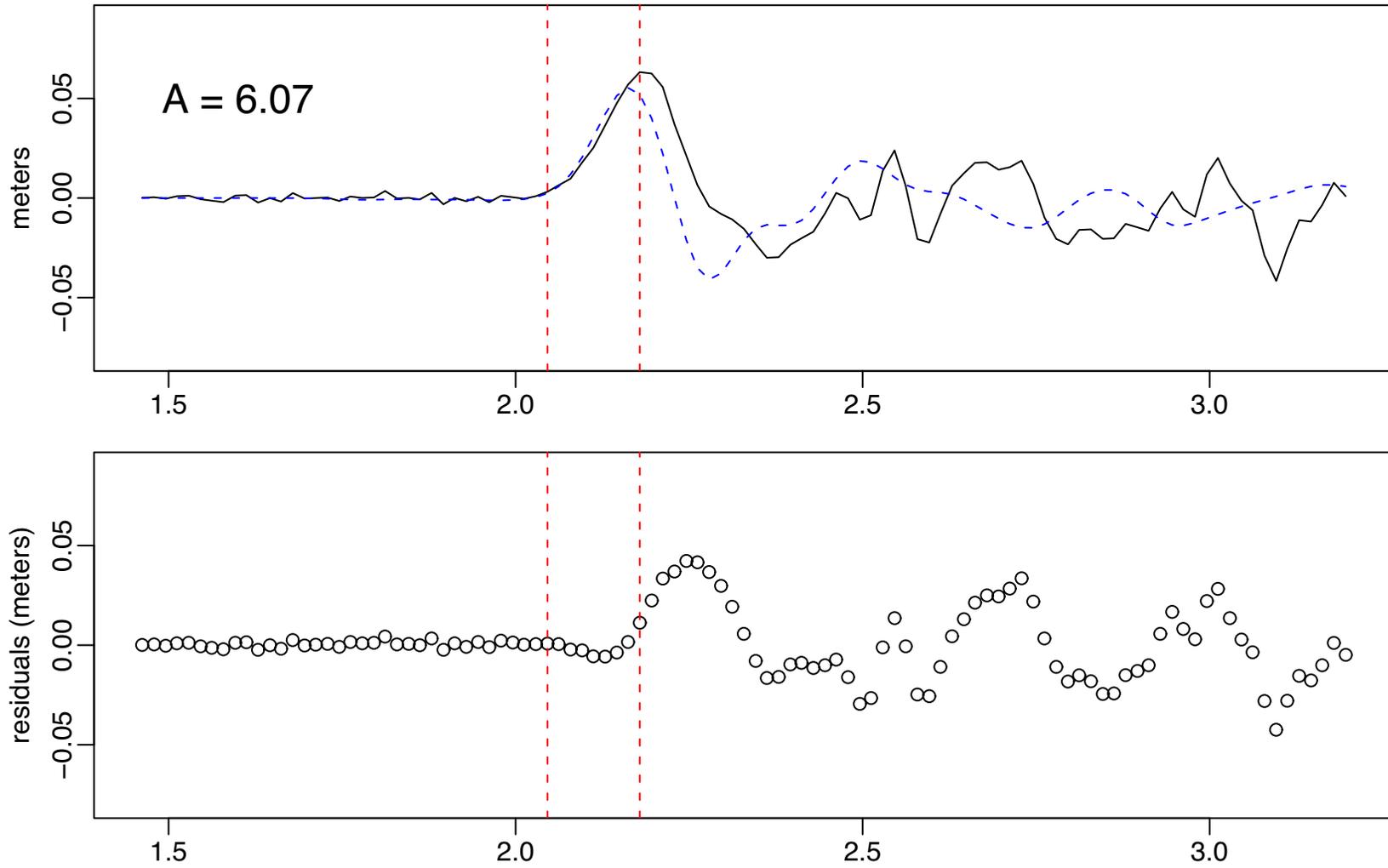
## Assessing Variability in $\hat{A}_{ls}$ : III

- theory says  $\hat{A}_{ls}$  is Gaussian distributed with mean  $A$  and variance  $\sigma_e^2 \cdot \mathbf{g}^T V \mathbf{g} / (\mathbf{g}^T \mathbf{g})^2$ , where  $V$  is matrix whose  $(j, k)$ th element is  $\phi^{|j-k|}$
- yields  $\hat{A}_2 = 4.17 \pm 1.33$ , which has larger SD than what was obtained under questionable assumptions ( $\hat{A}_2 = 4.17 \pm 0.59$ )

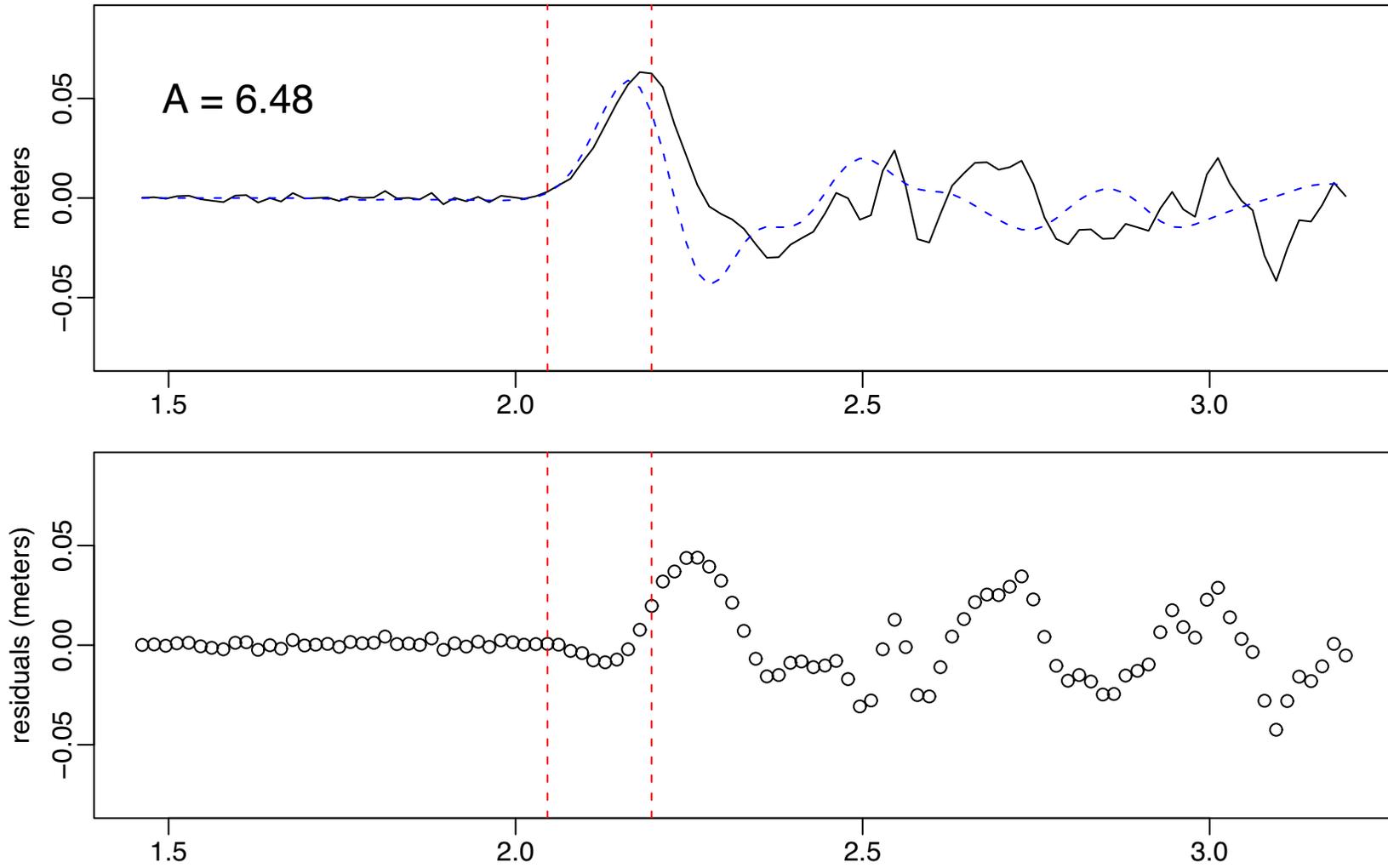
## Least Squares (LS) as Criterion for Estimating Slips

- inversion algorithm uses LS as criterion for estimating slip  $A$
- user must decide amount of DART data to use
- two extremes: all available data or just two data points
- use of more data should yield estimator  $\hat{A}$  with smaller variance *if* model is valid over entire range of data
- if model decreases in validity as time increases, should limit data to, say, first quarter wave or first full wave
- real-time constraints also dictate interest in use of limited amount of data
- starting with a quarter wave of data, let's look at LS fits involving varying amounts of data

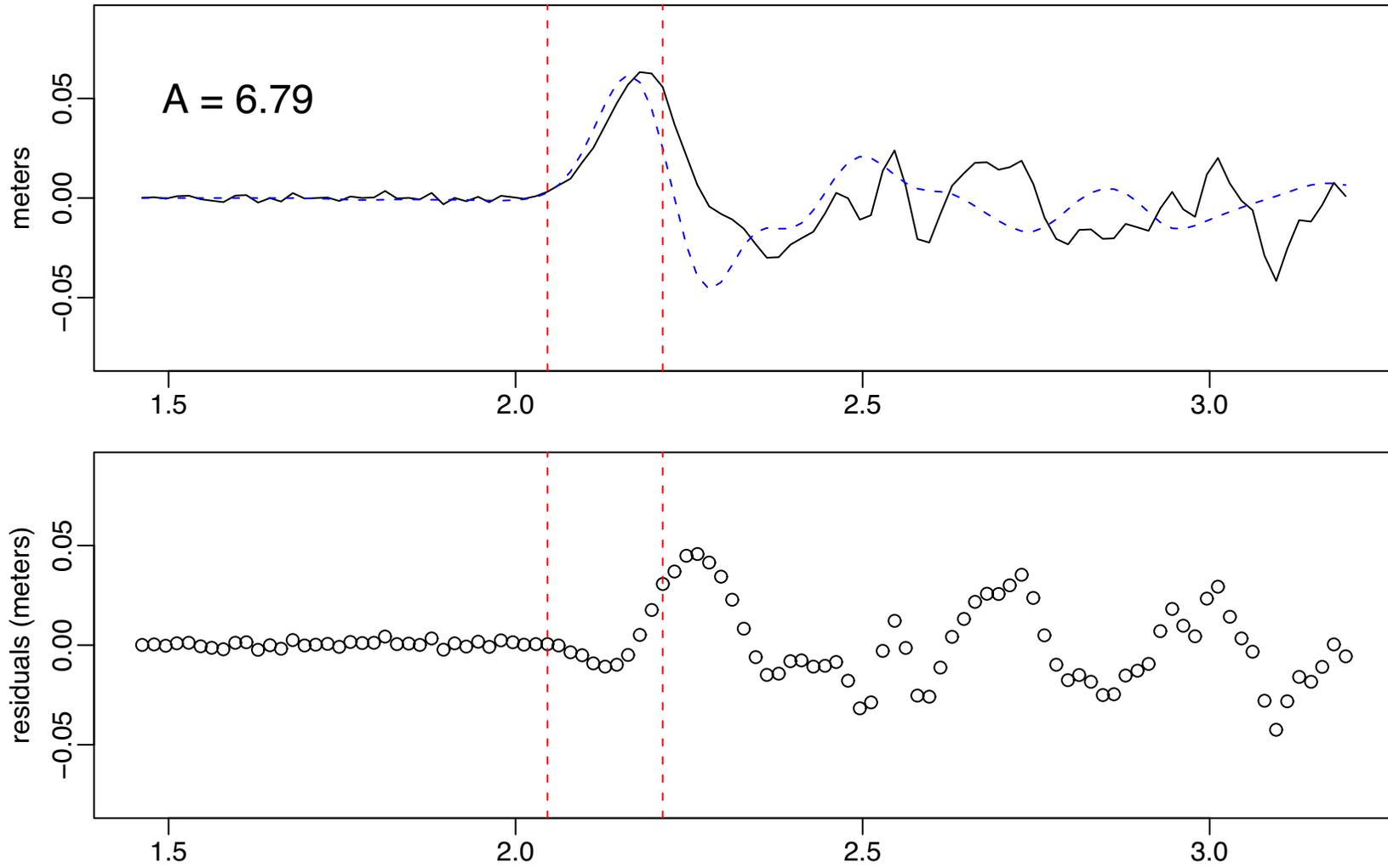
# LS Fit of a12 Model to Selected 21414 Data



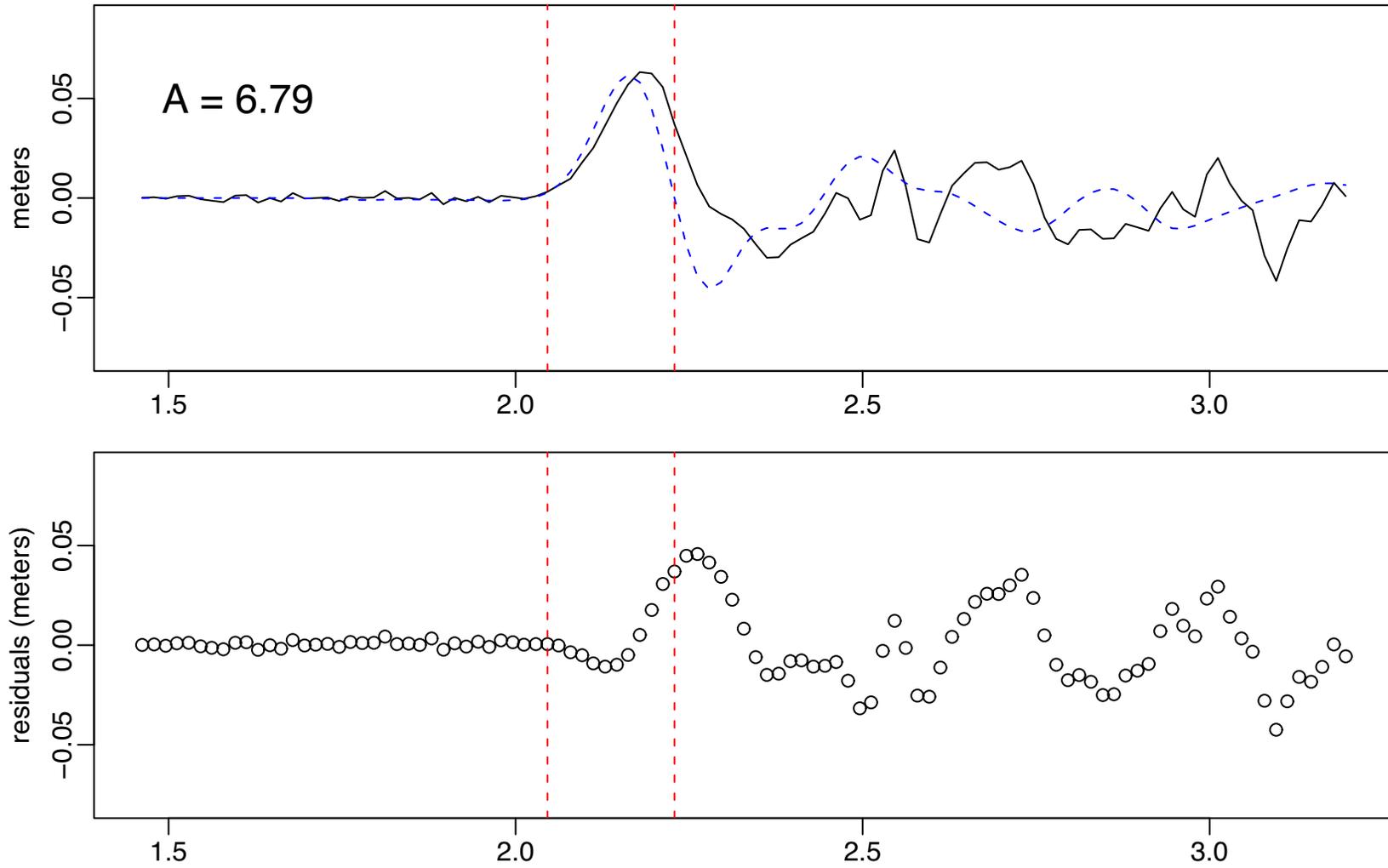
# LS Fit of a12 Model to Selected 21414 Data



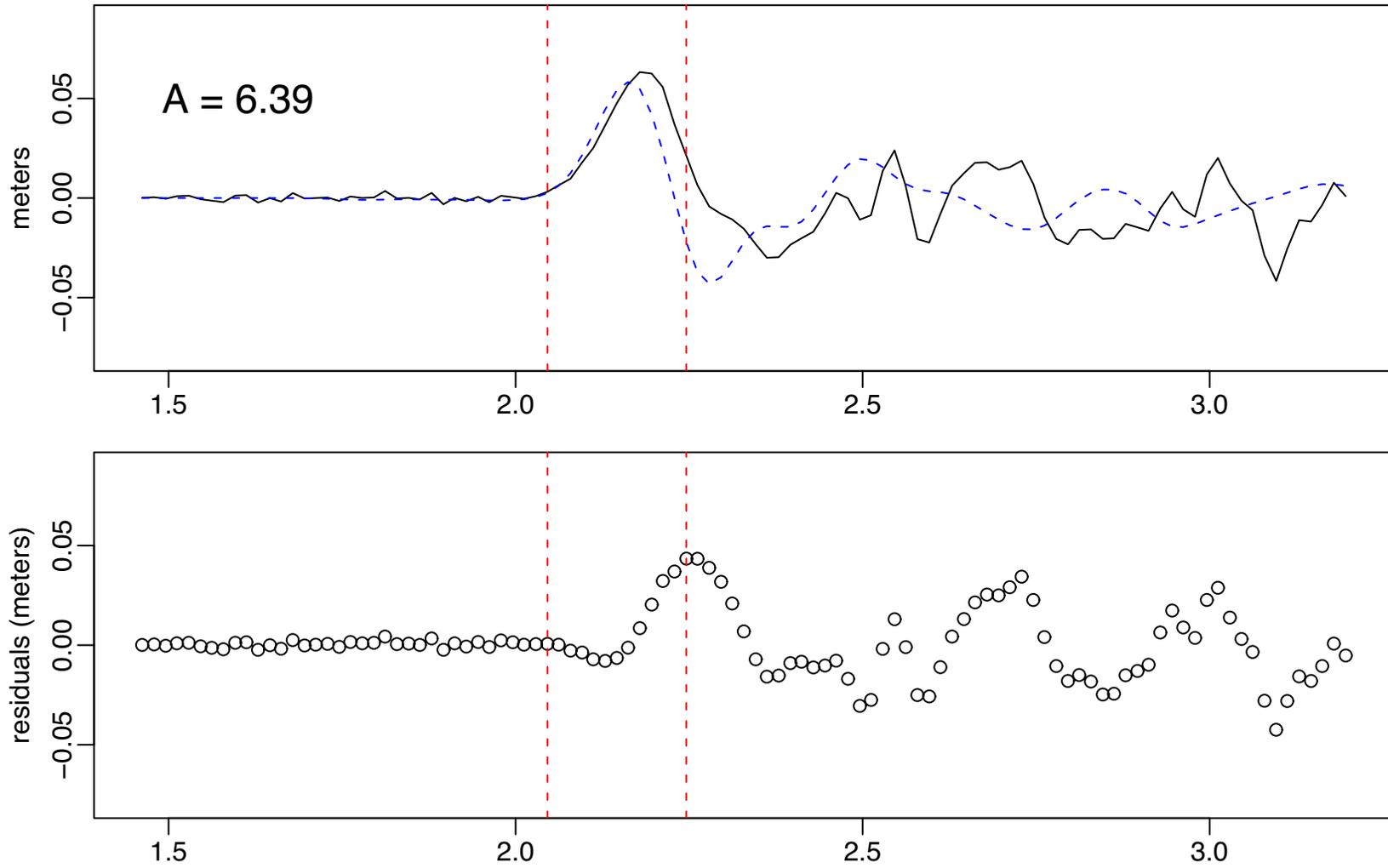
# LS Fit of a12 Model to Selected 21414 Data



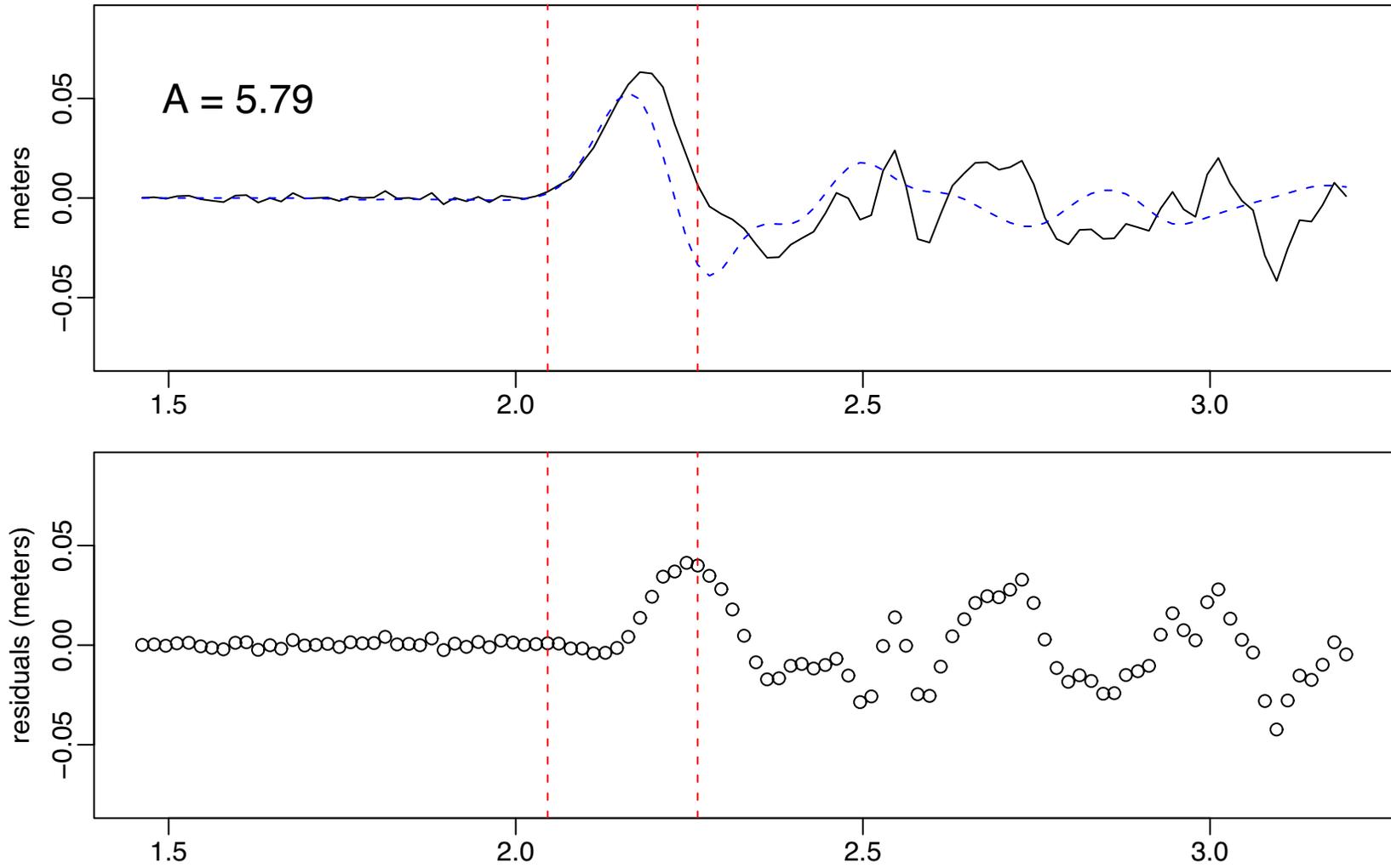
# LS Fit of a12 Model to Selected 21414 Data



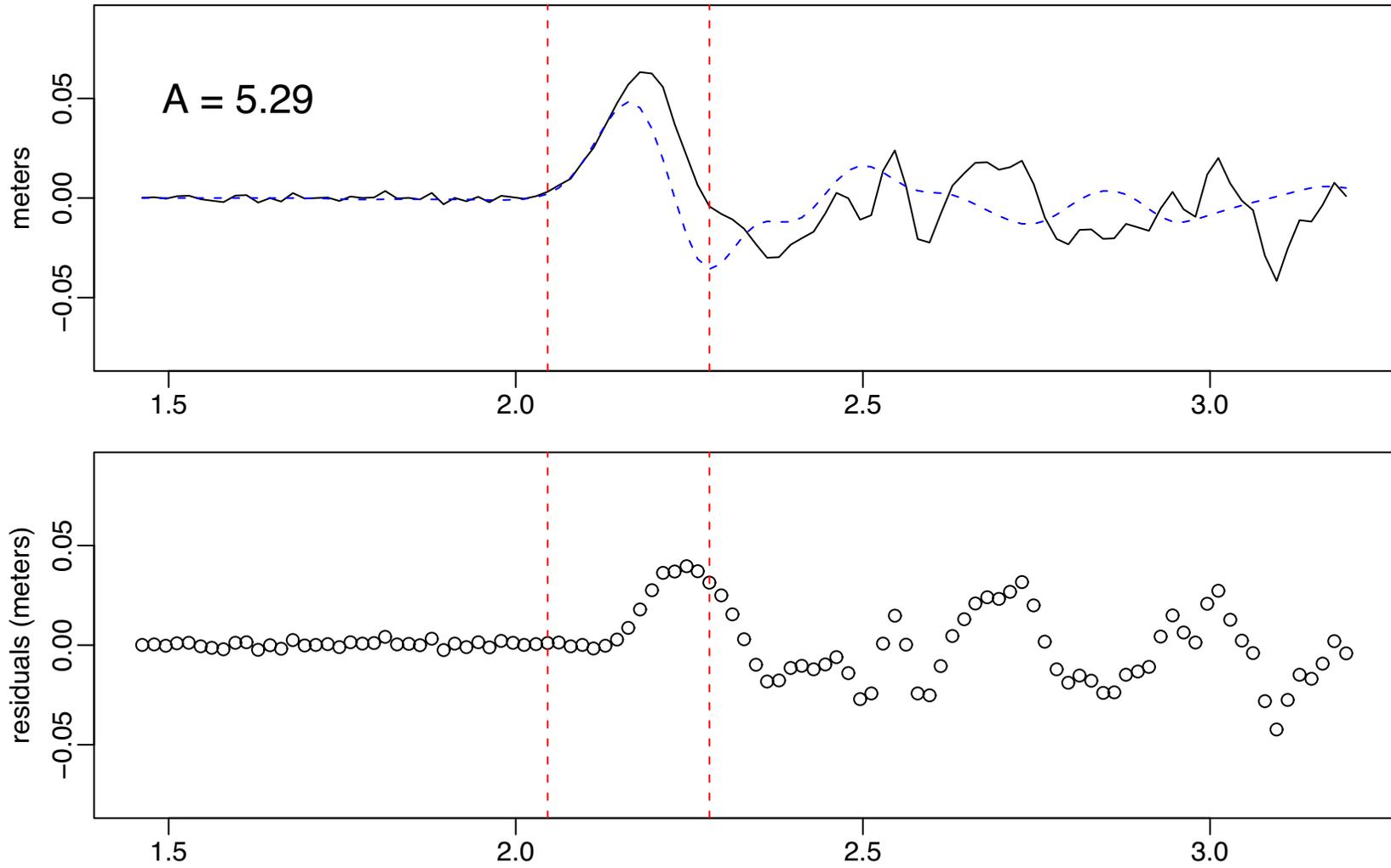
# LS Fit of a12 Model to Selected 21414 Data



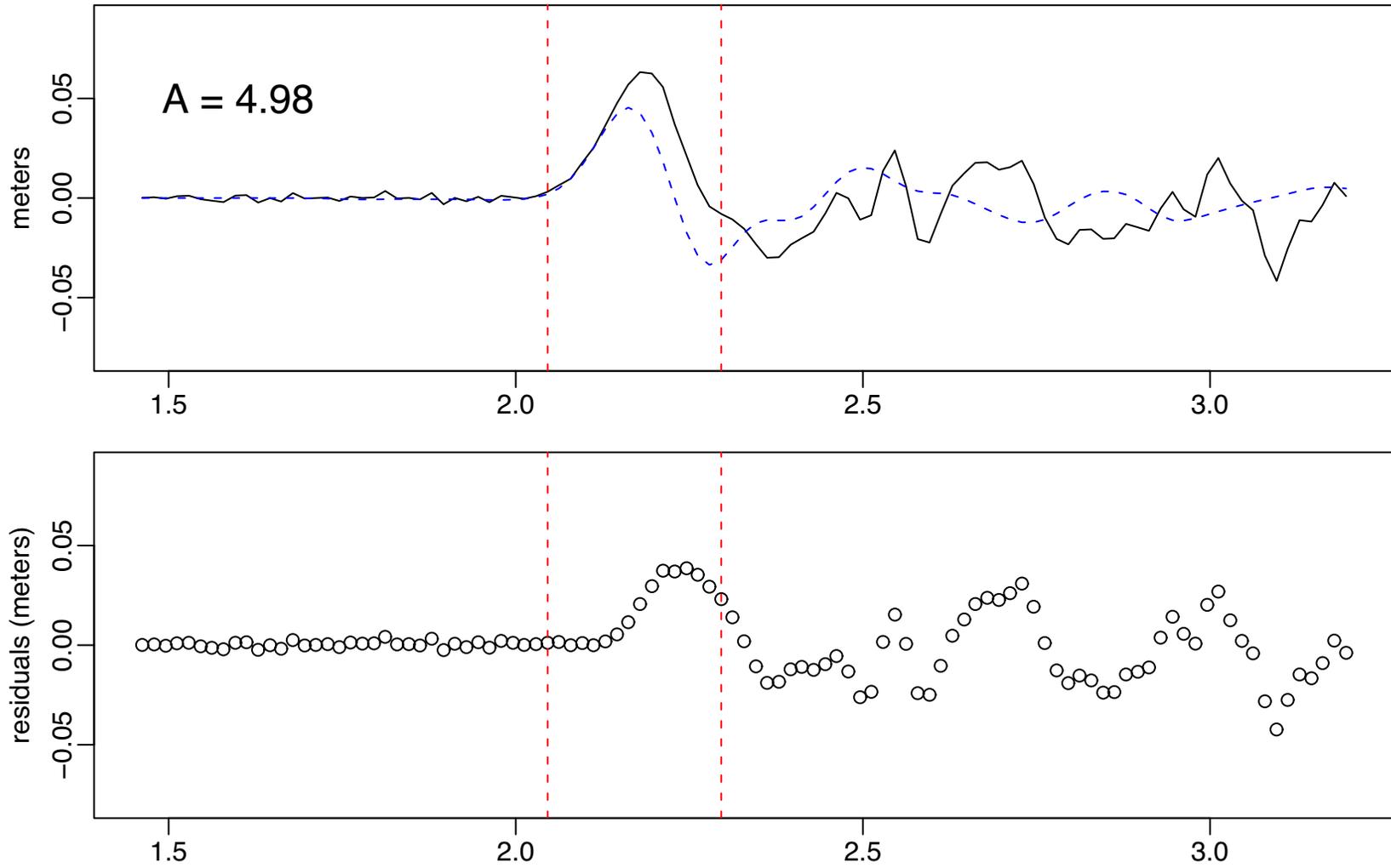
# LS Fit of a12 Model to Selected 21414 Data



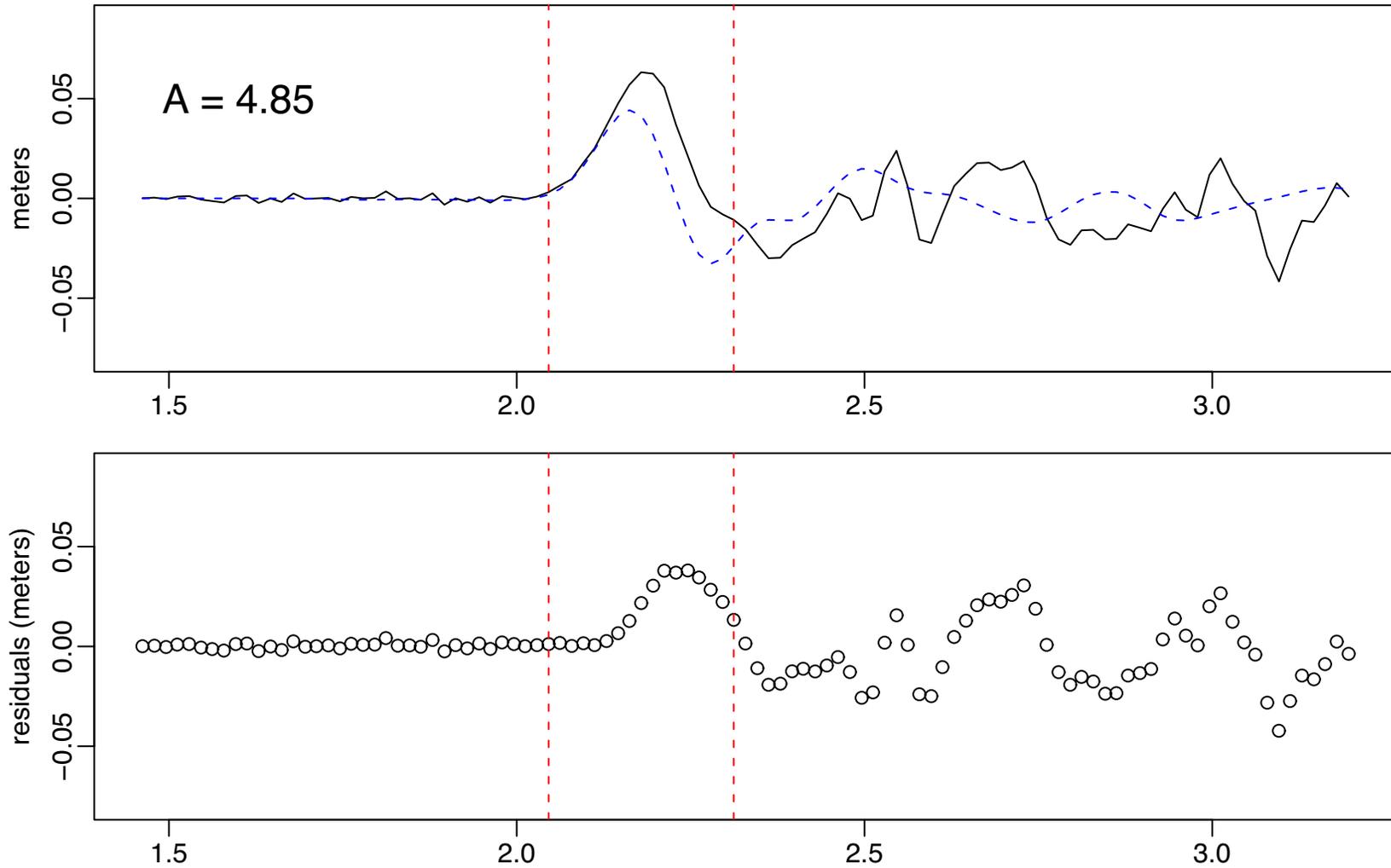
# LS Fit of a12 Model to Selected 21414 Data



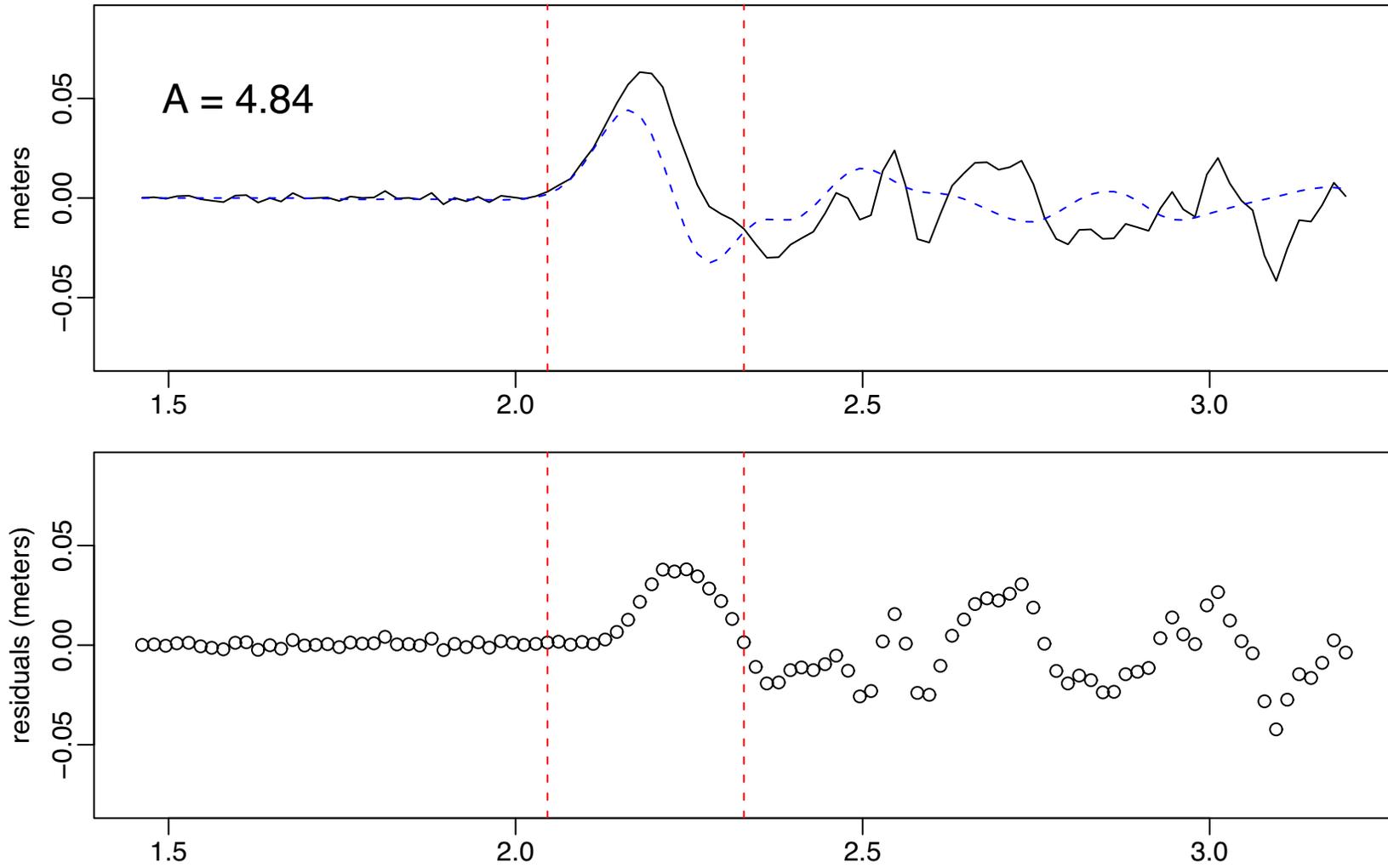
# LS Fit of a12 Model to Selected 21414 Data



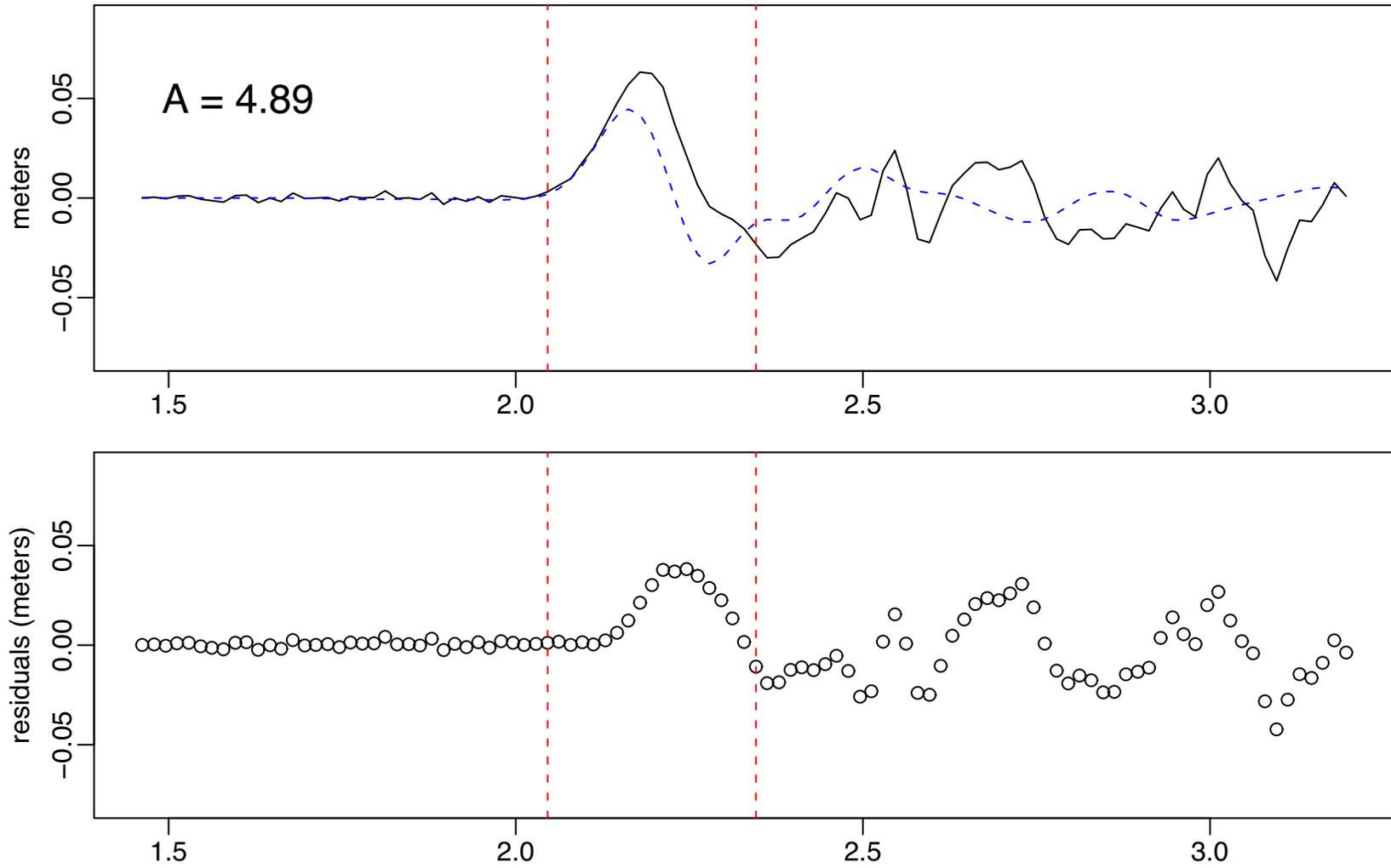
# LS Fit of a12 Model to Selected 21414 Data



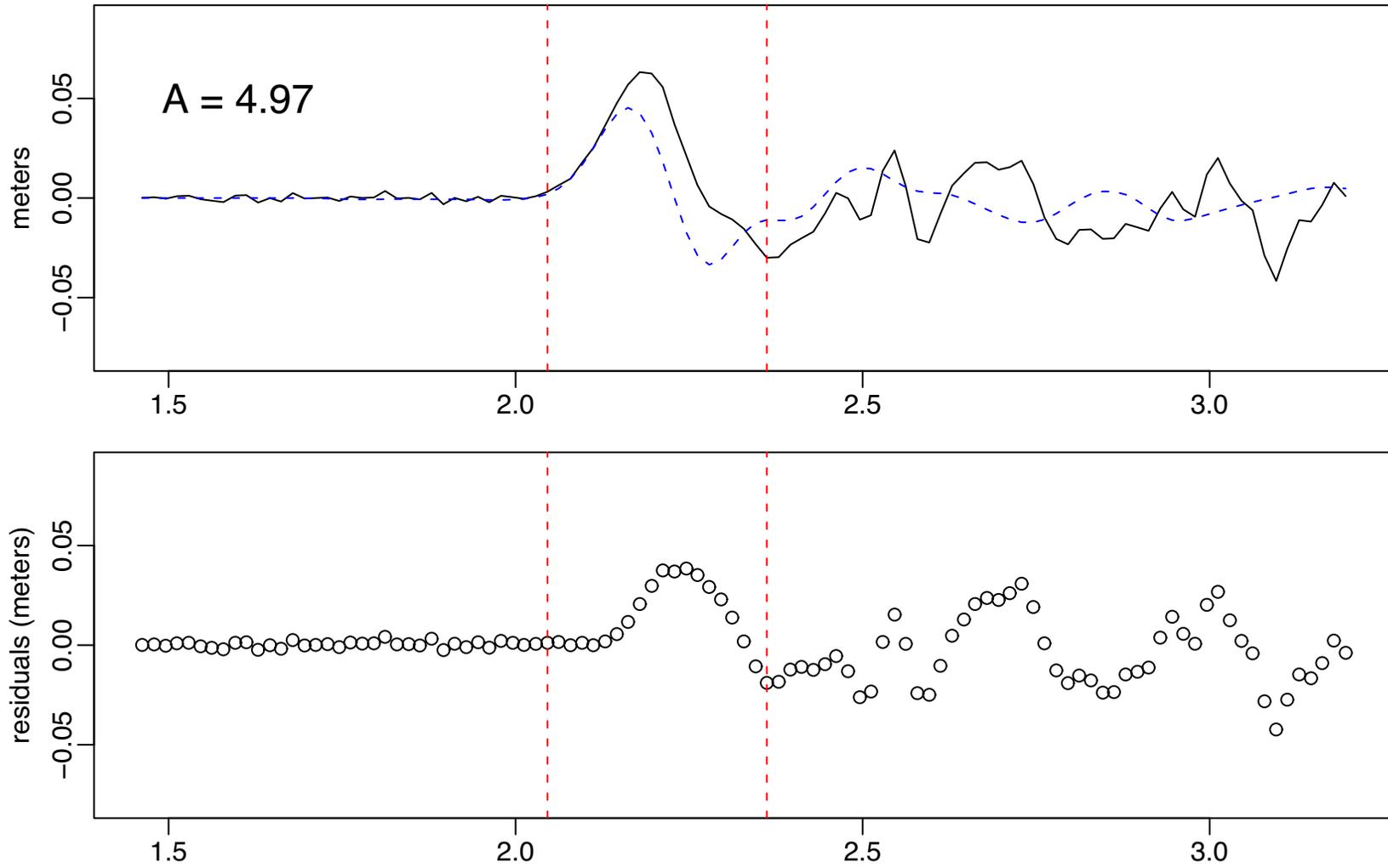
# LS Fit of a12 Model to Selected 21414 Data



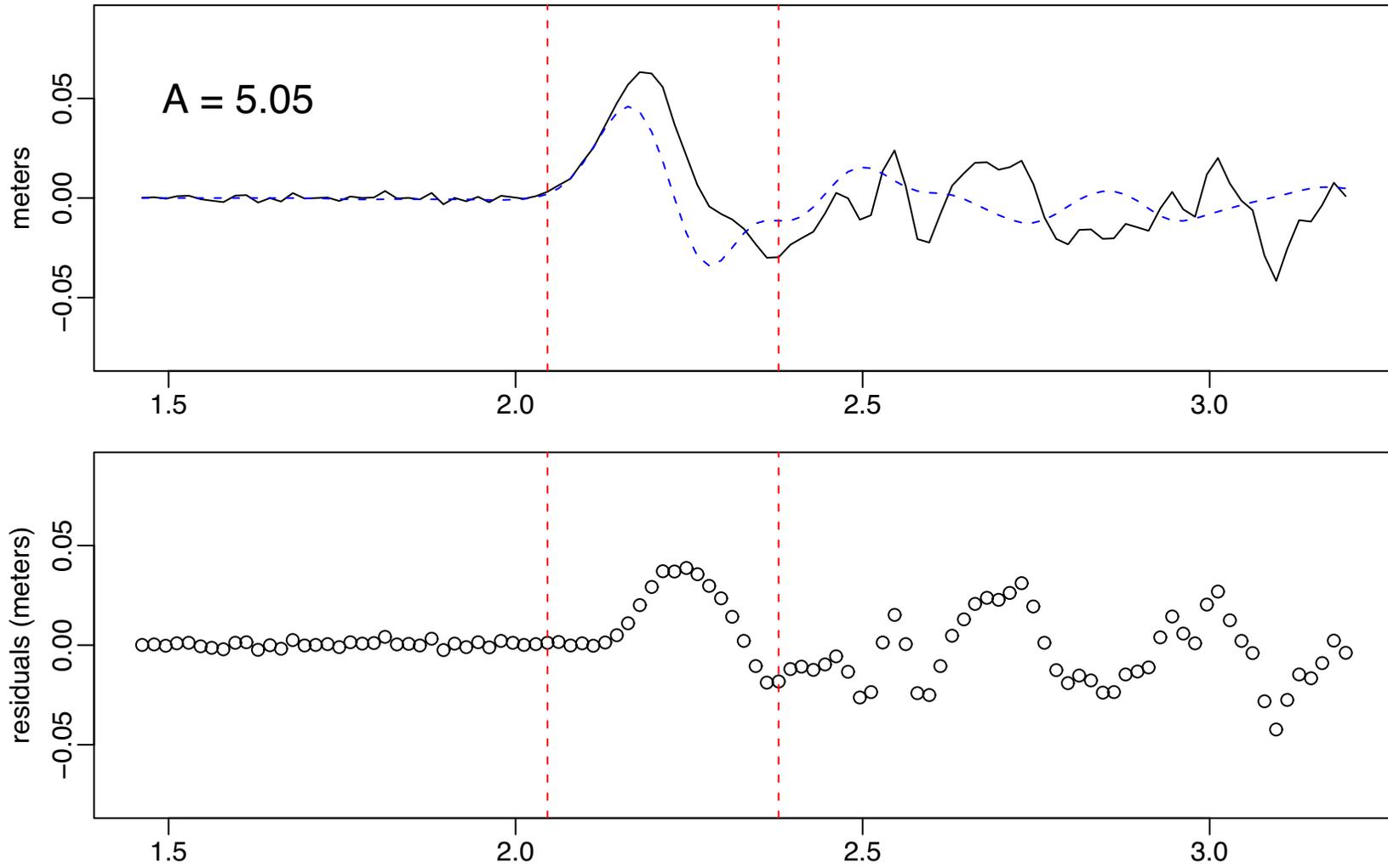
# LS Fit of a12 Model to Selected 21414 Data



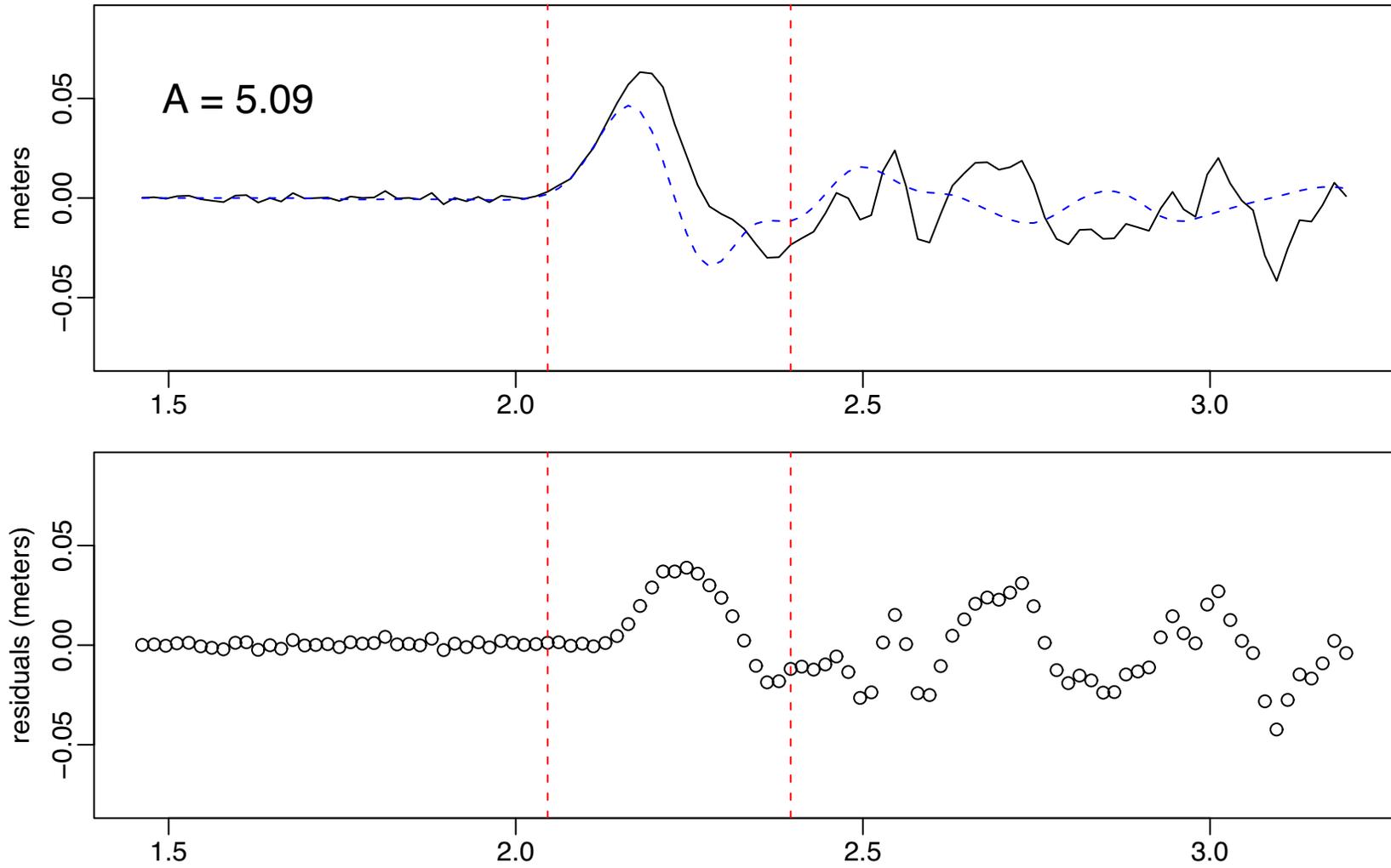
# LS Fit of a12 Model to Selected 21414 Data



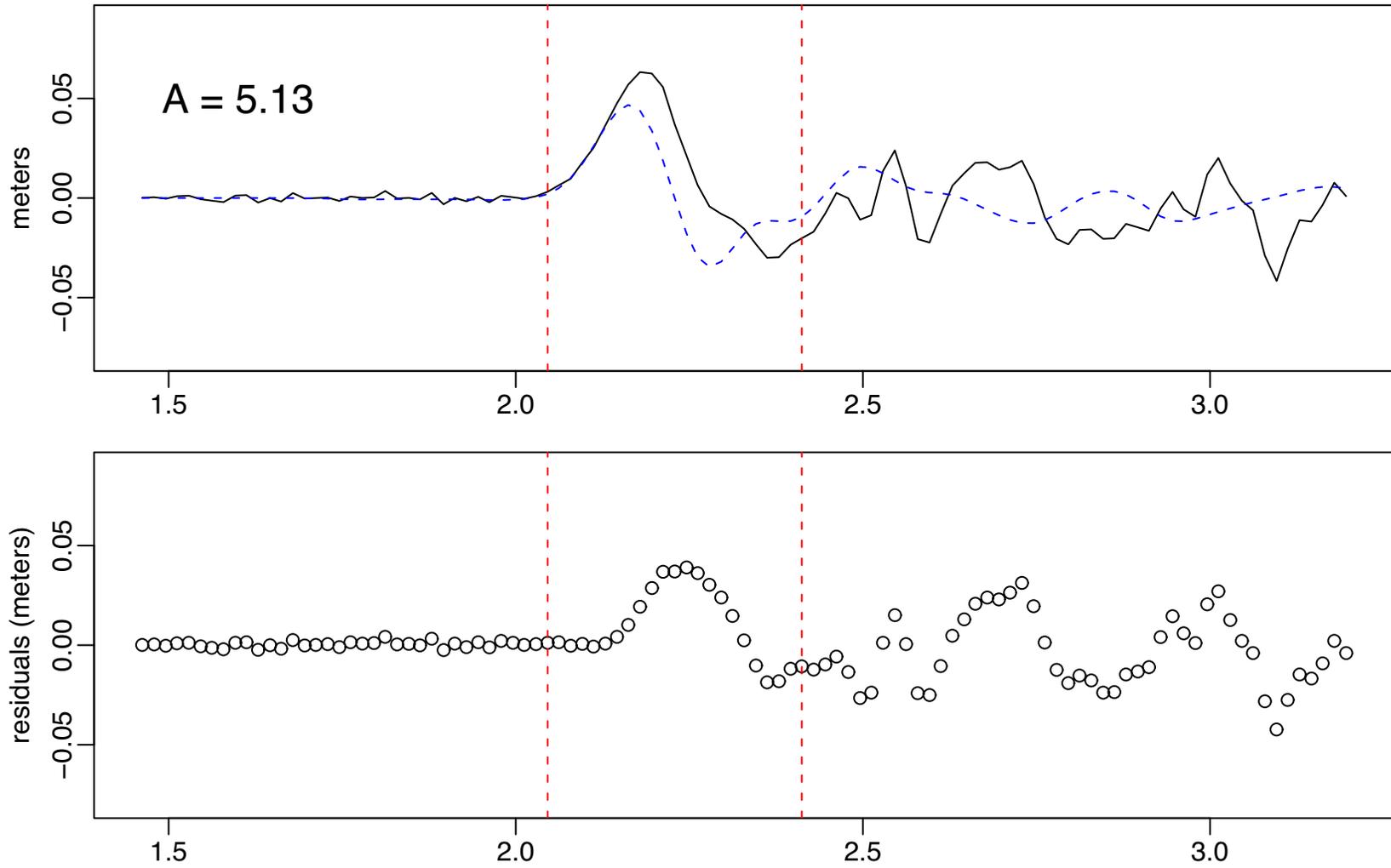
# LS Fit of a12 Model to Selected 21414 Data



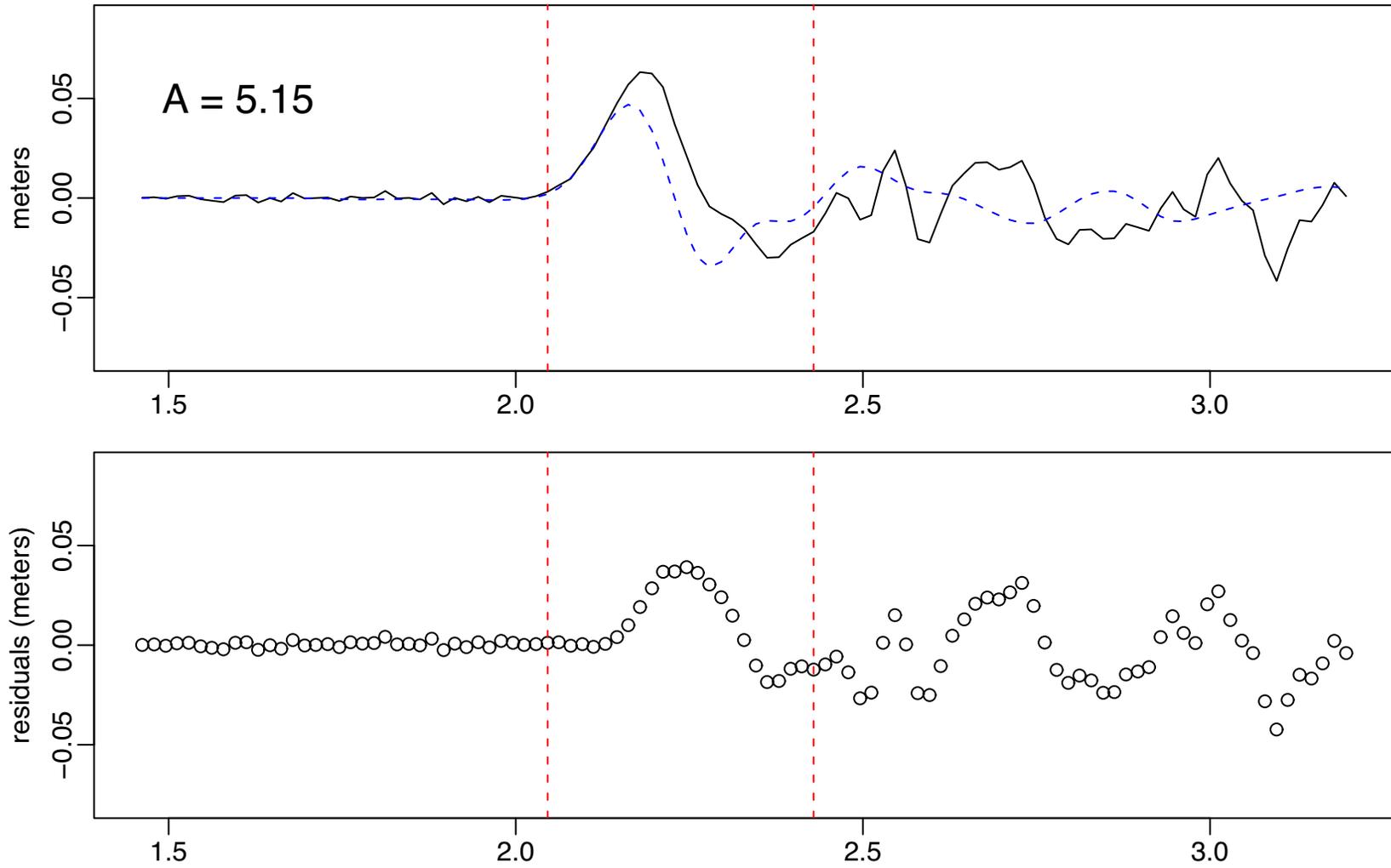
# LS Fit of a12 Model to Selected 21414 Data



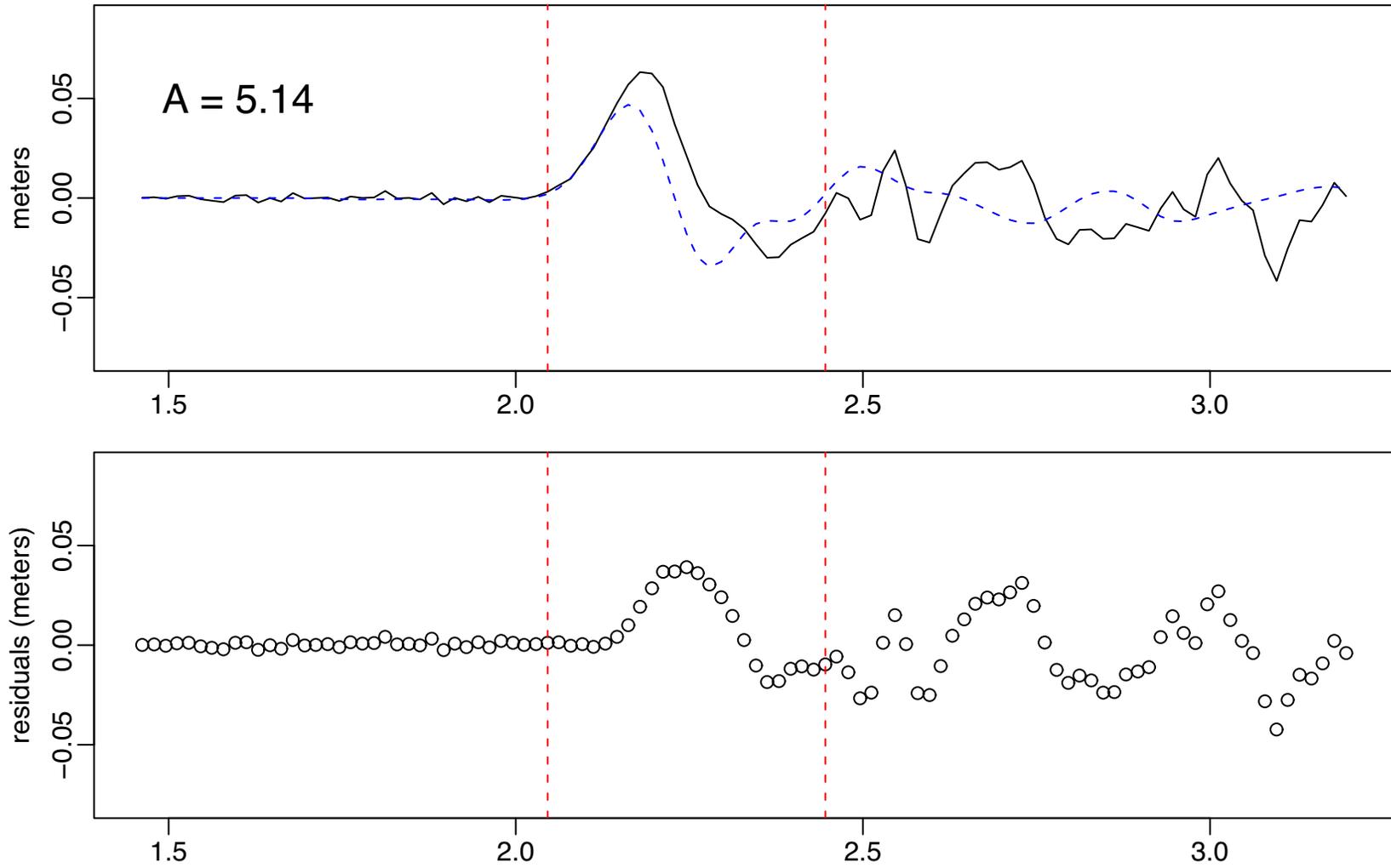
# LS Fit of a12 Model to Selected 21414 Data



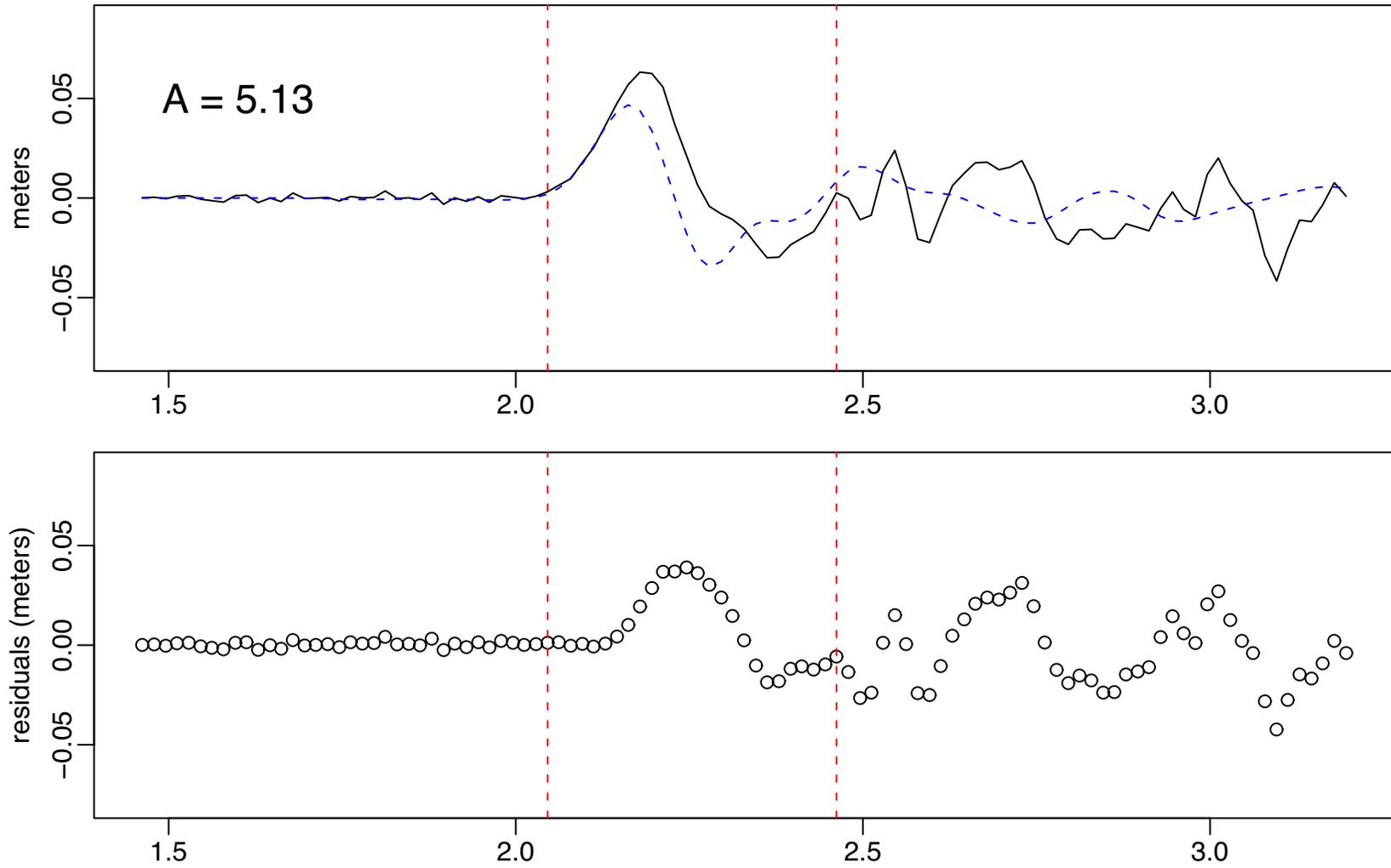
# LS Fit of a12 Model to Selected 21414 Data



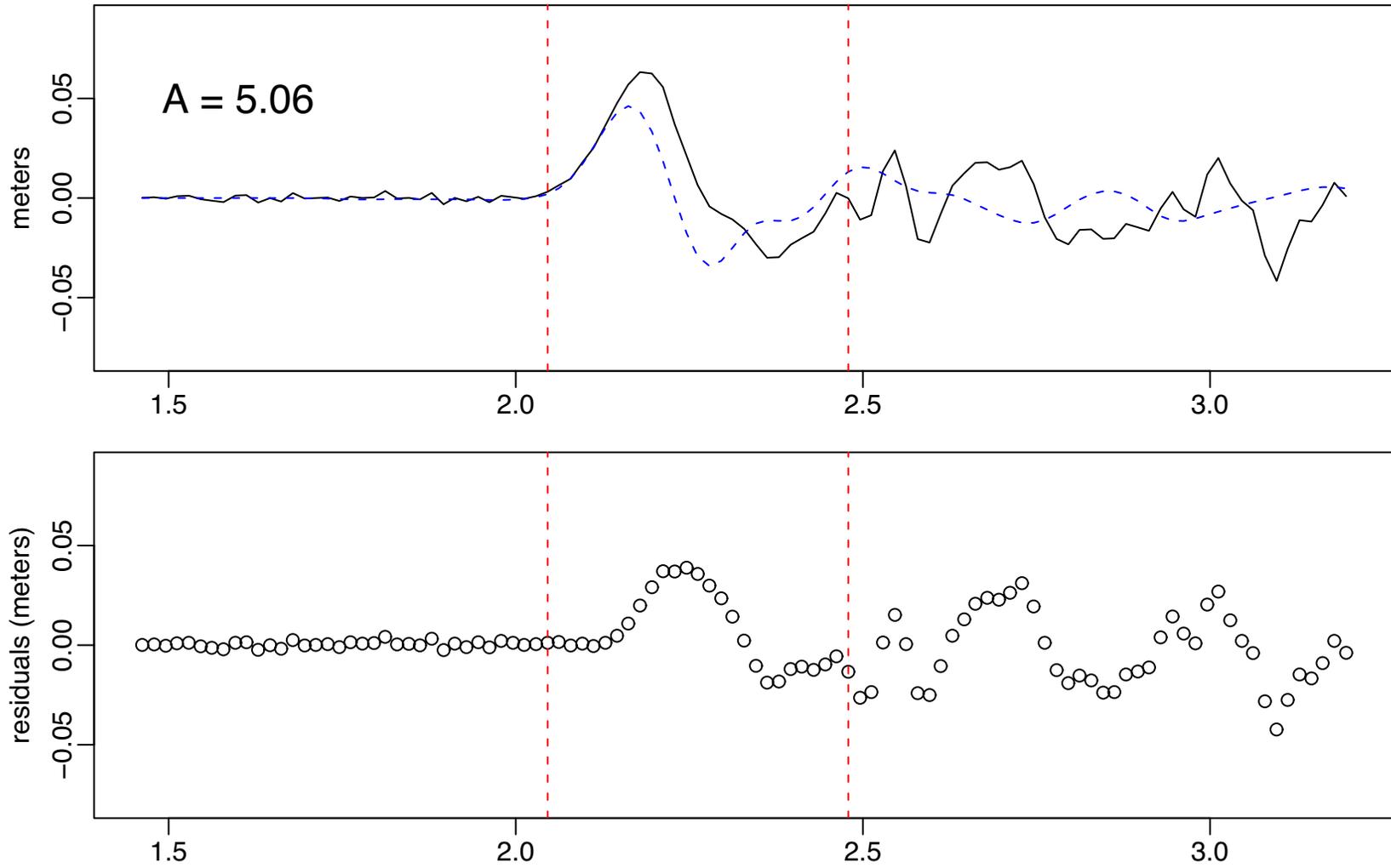
# LS Fit of a12 Model to Selected 21414 Data



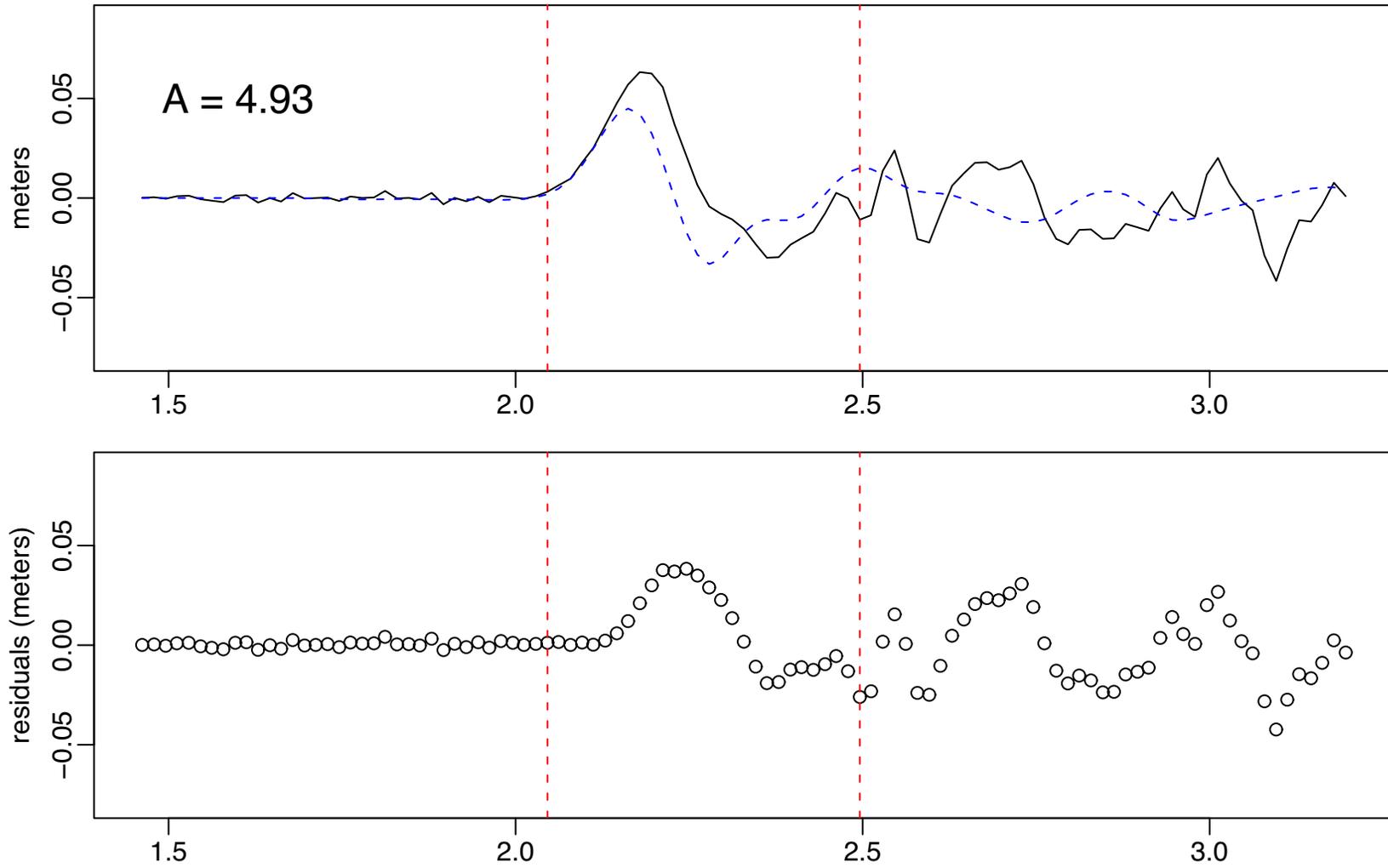
# LS Fit of a12 Model to Selected 21414 Data



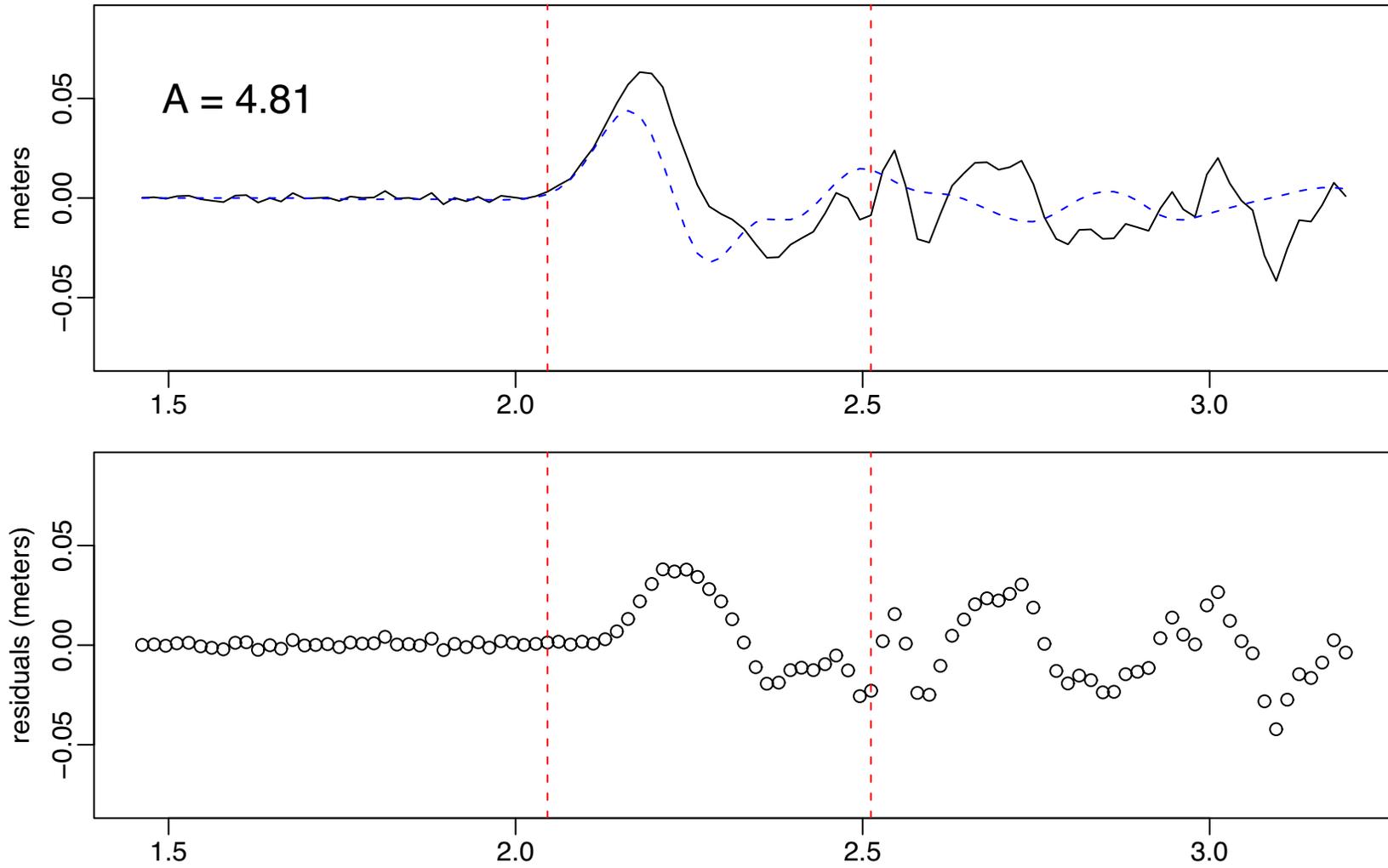
# LS Fit of a12 Model to Selected 21414 Data



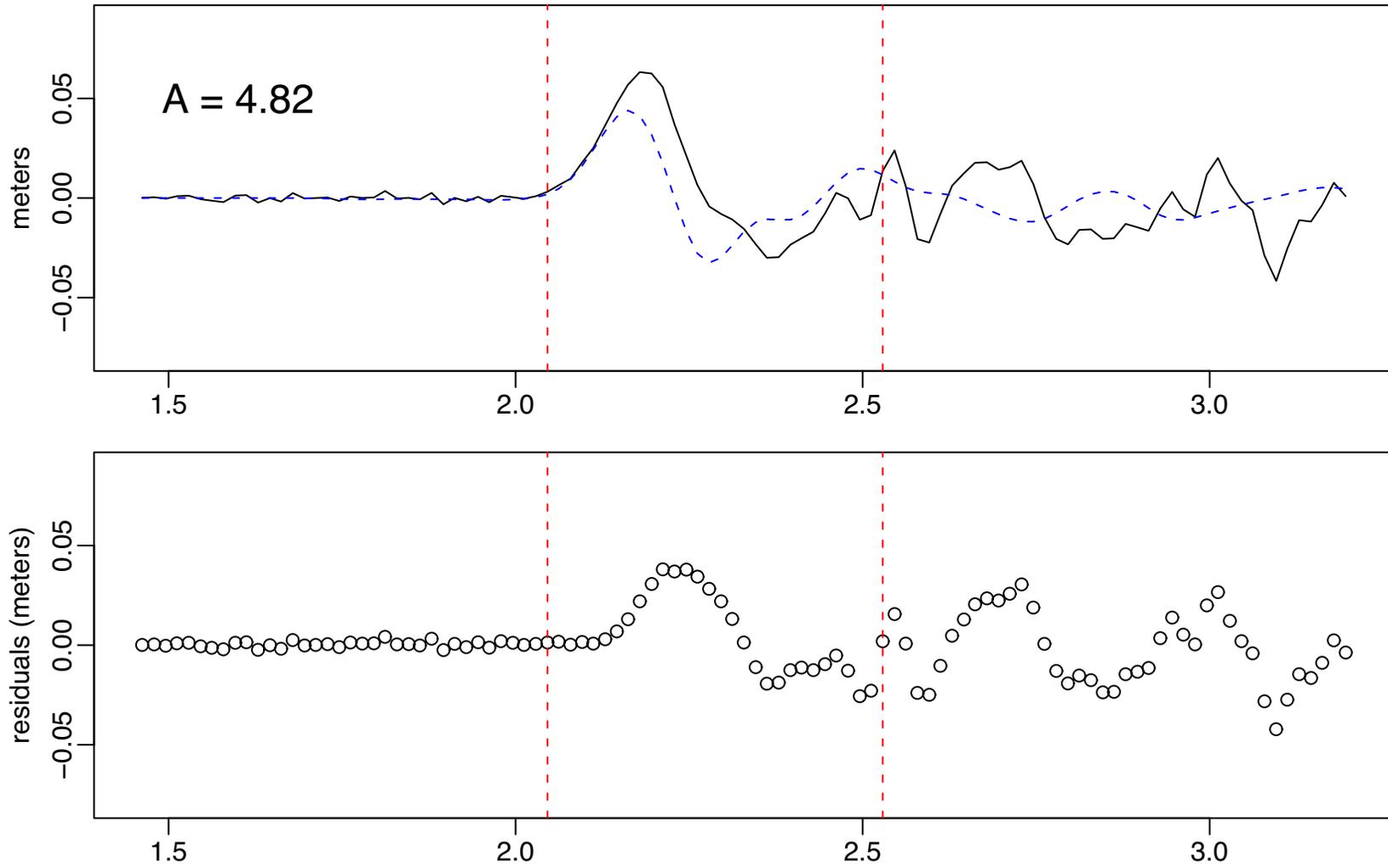
# LS Fit of a12 Model to Selected 21414 Data



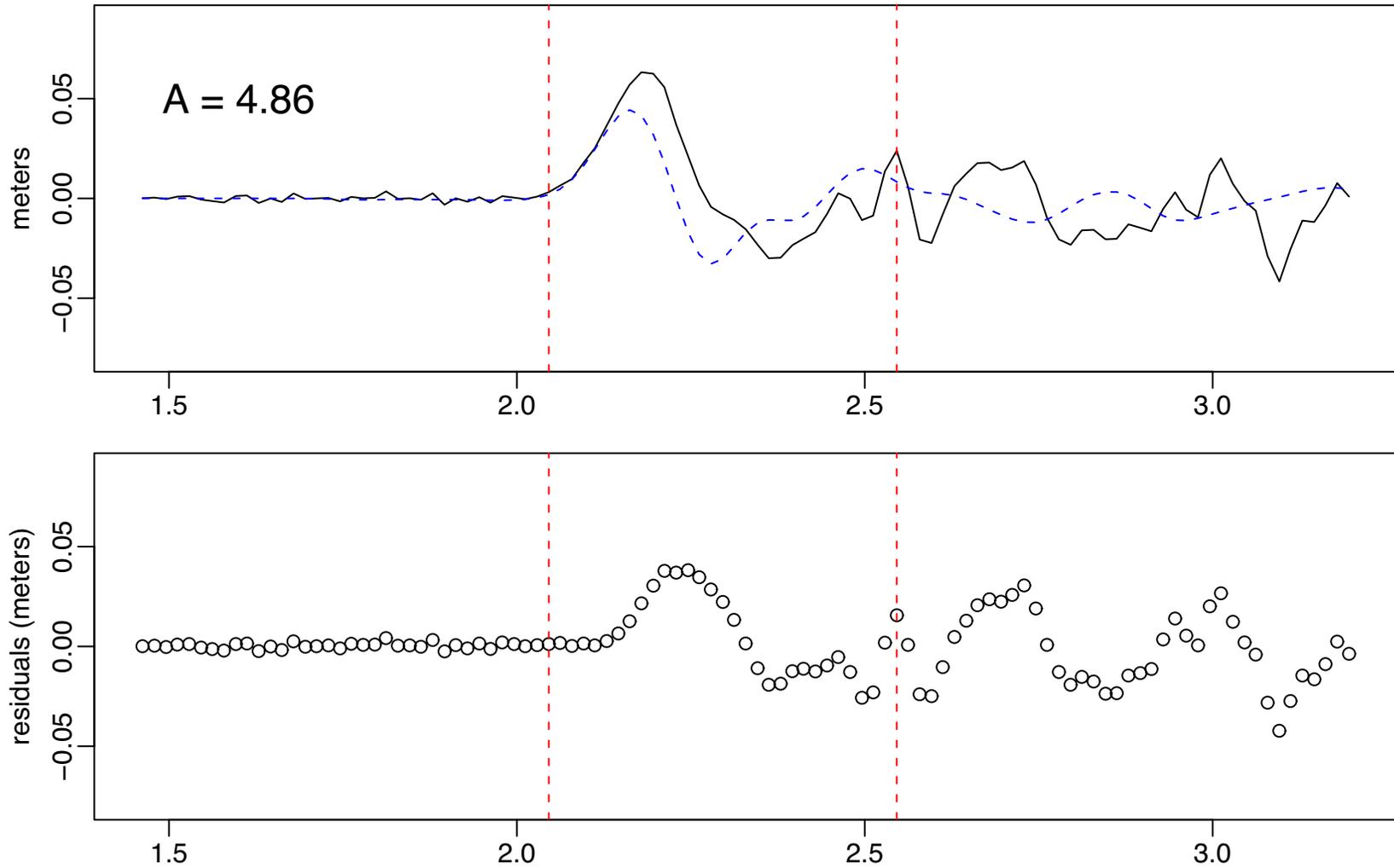
# LS Fit of a12 Model to Selected 21414 Data



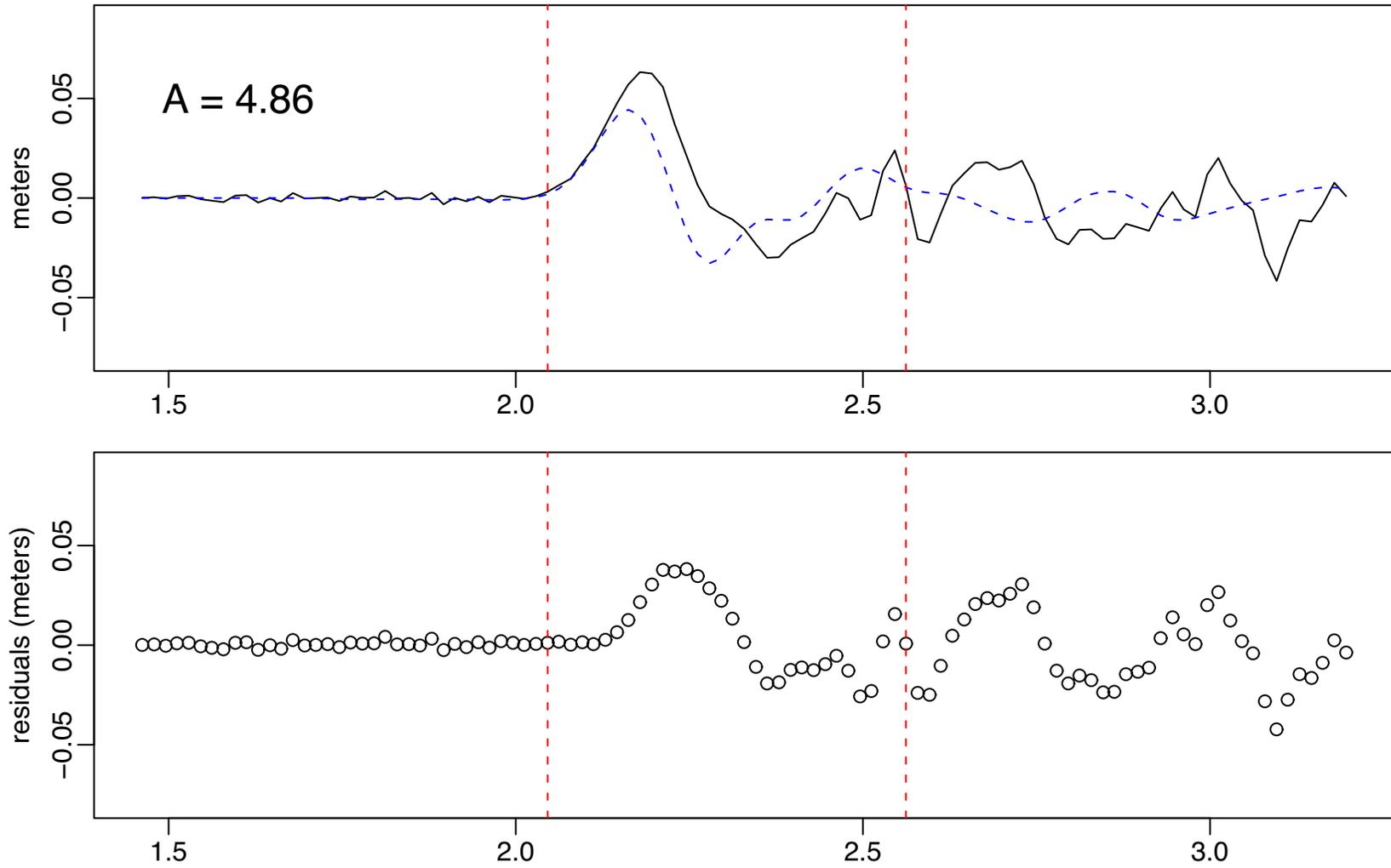
# LS Fit of a12 Model to Selected 21414 Data



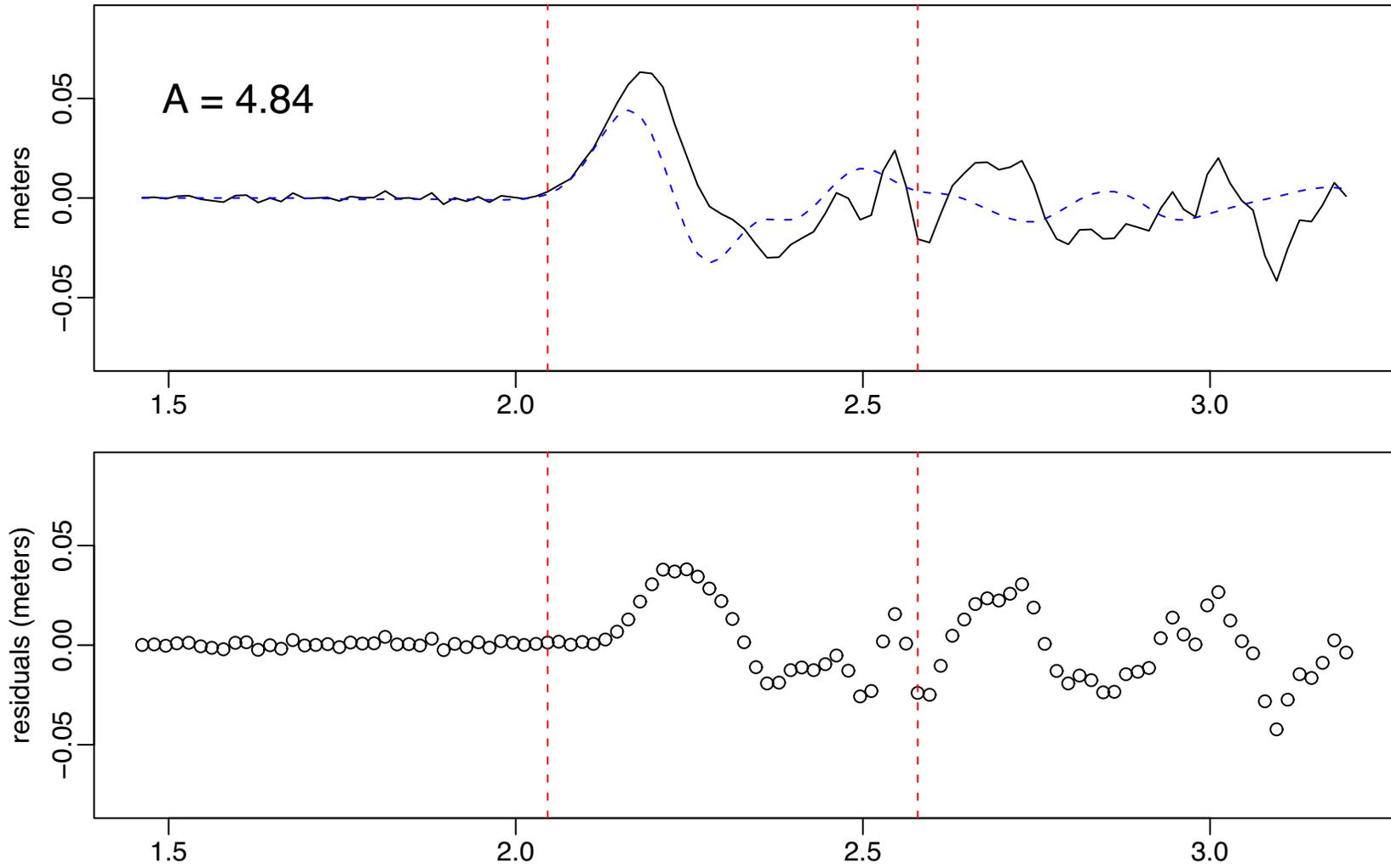
# LS Fit of a12 Model to Selected 21414 Data



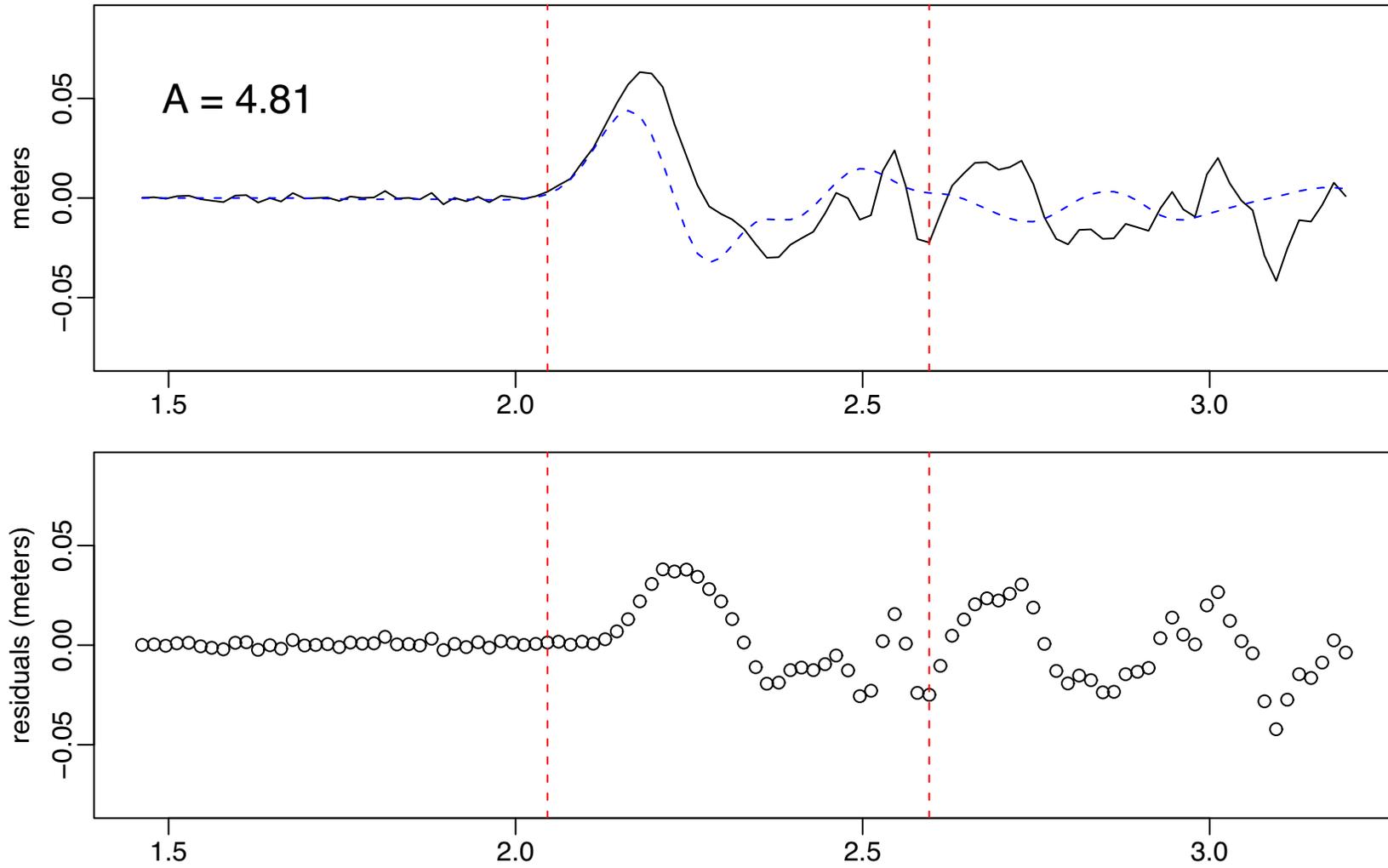
# LS Fit of a12 Model to Selected 21414 Data



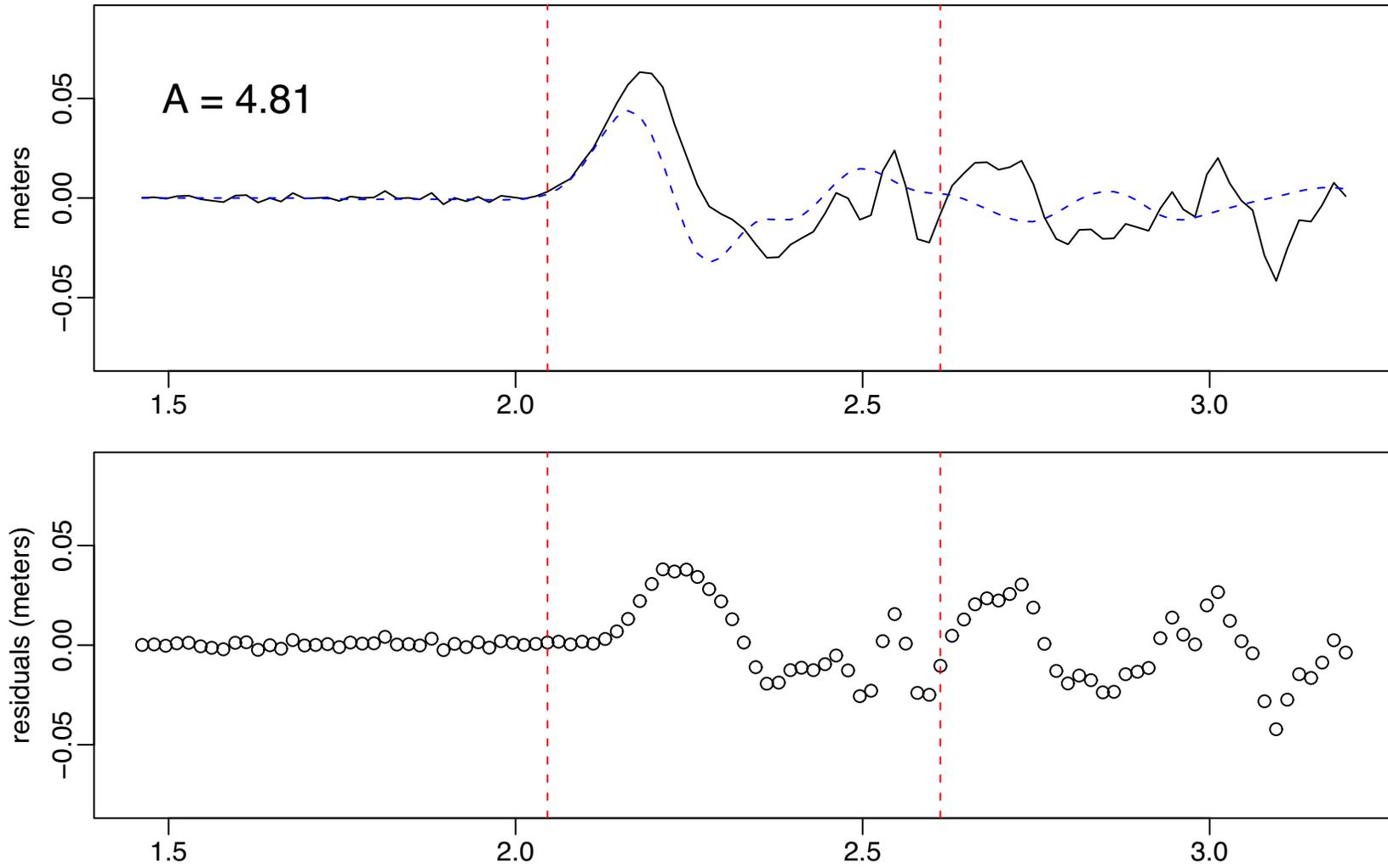
# LS Fit of a12 Model to Selected 21414 Data



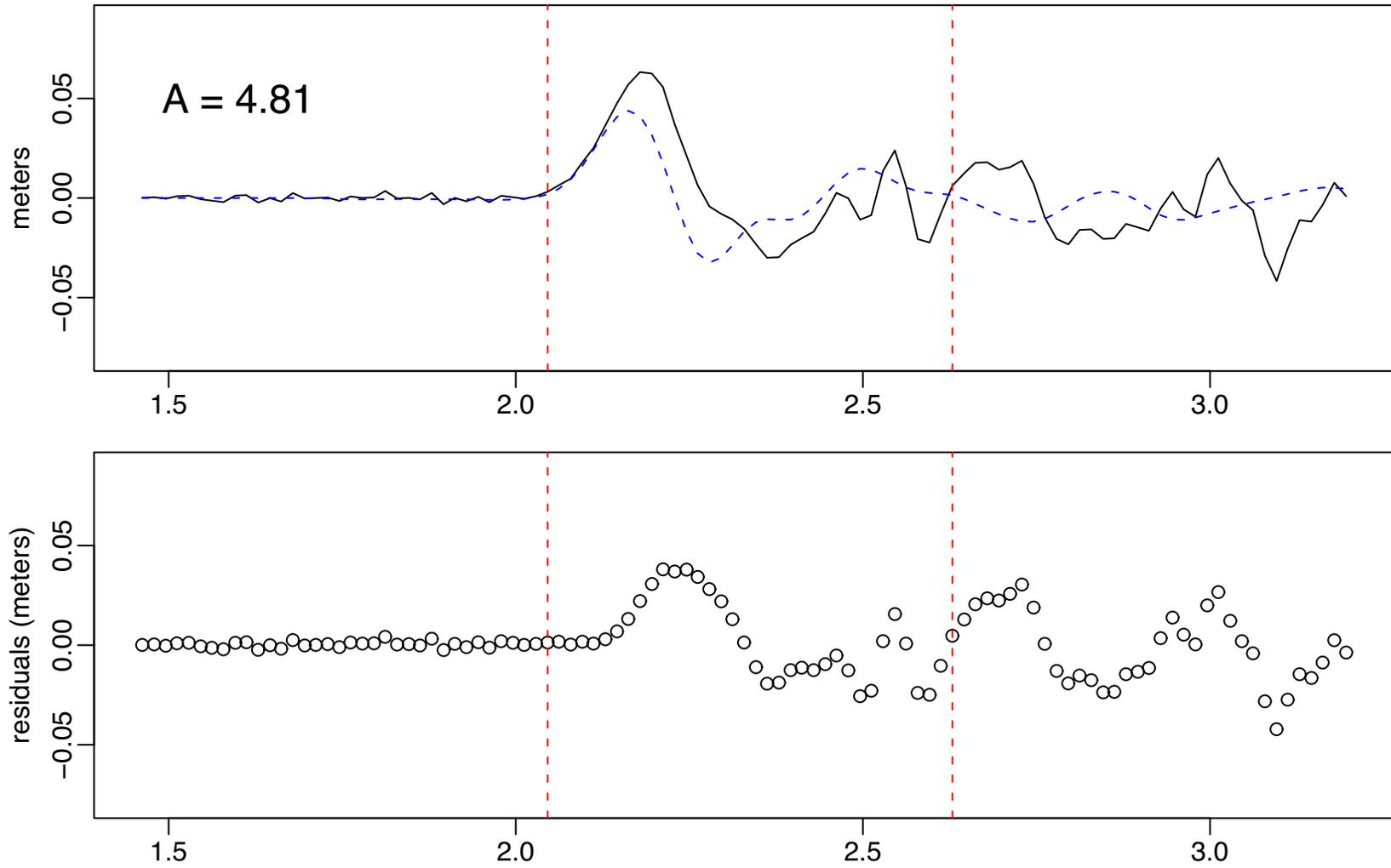
# LS Fit of a12 Model to Selected 21414 Data



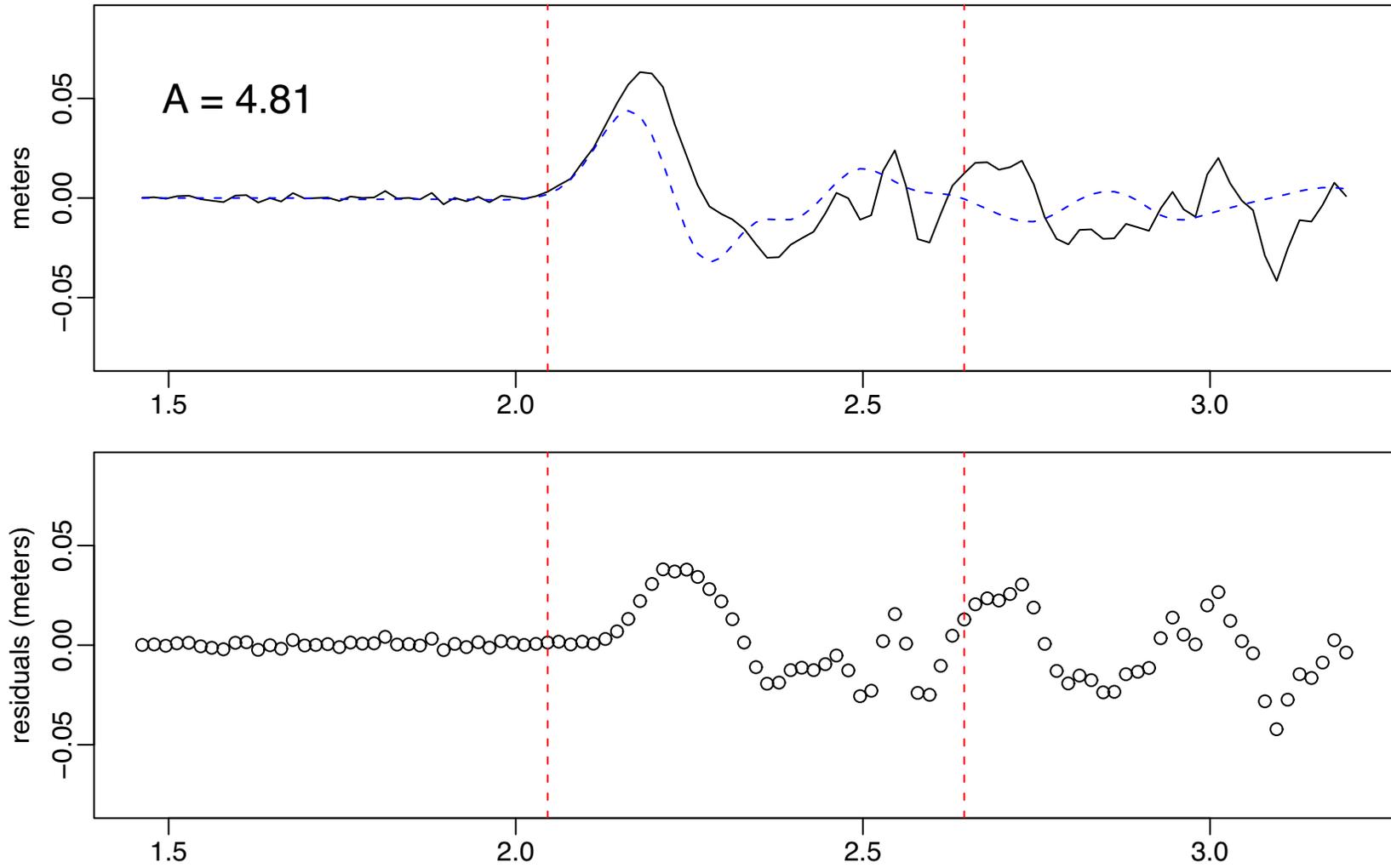
# LS Fit of a12 Model to Selected 21414 Data



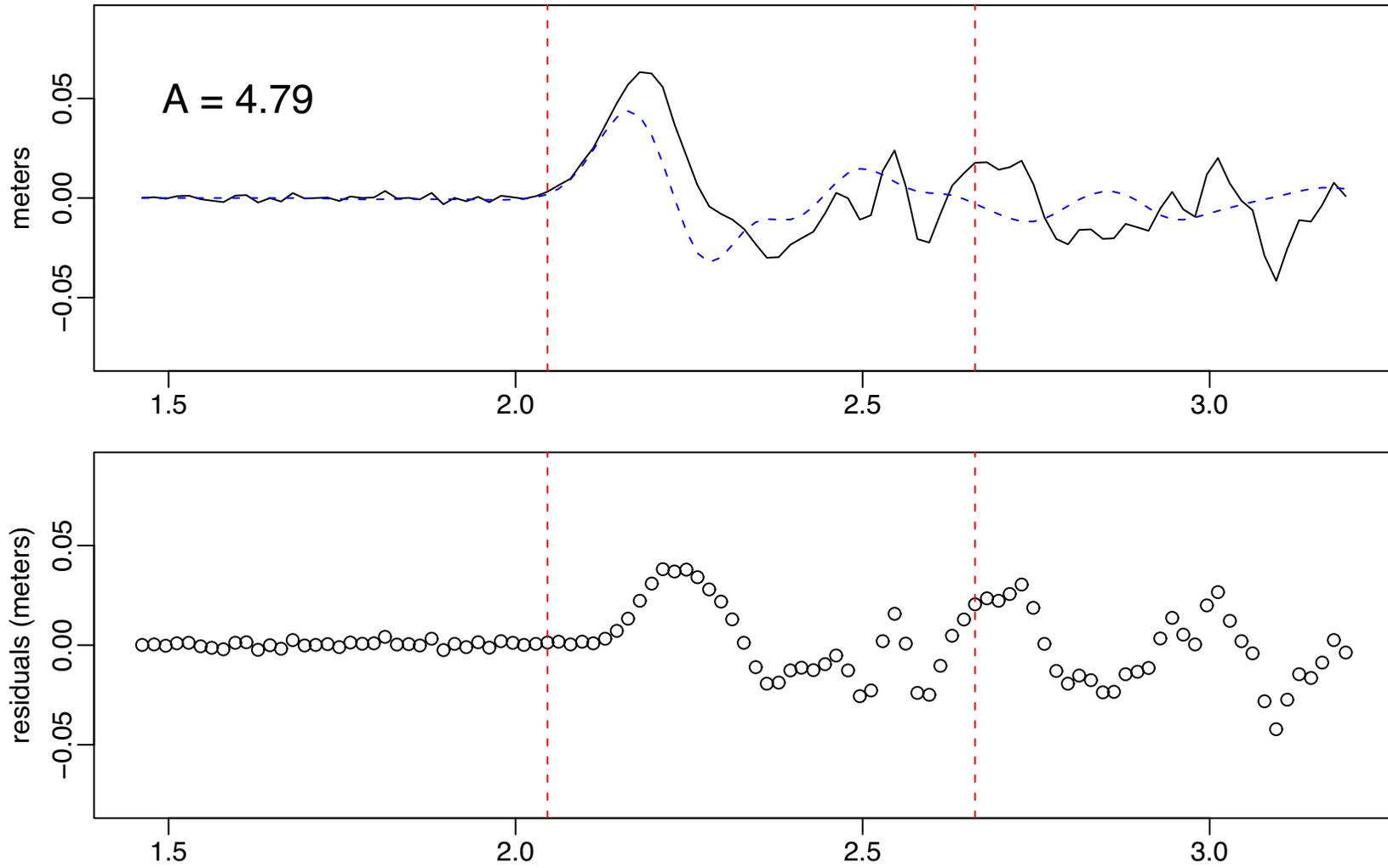
# LS Fit of a12 Model to Selected 21414 Data



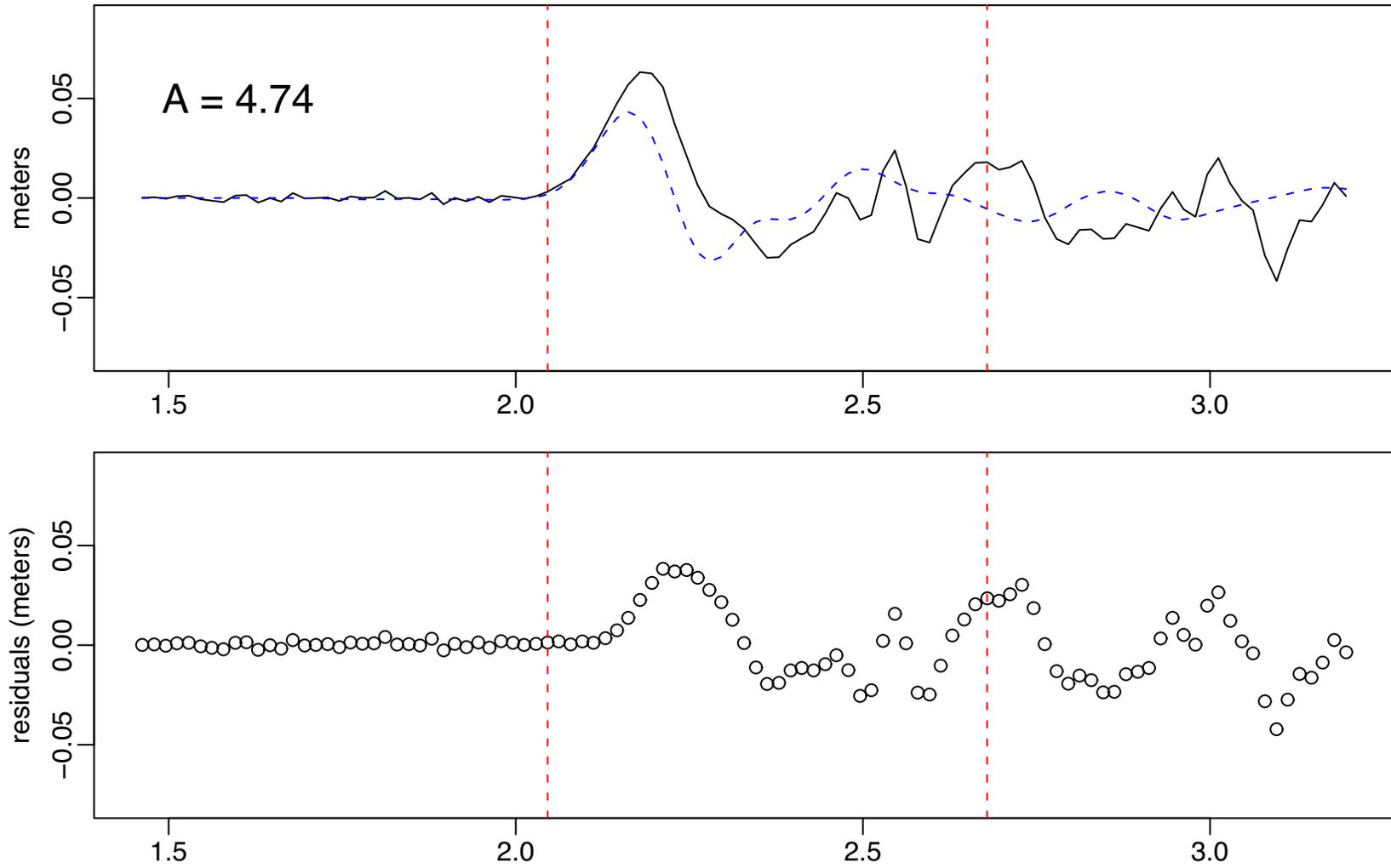
# LS Fit of a12 Model to Selected 21414 Data



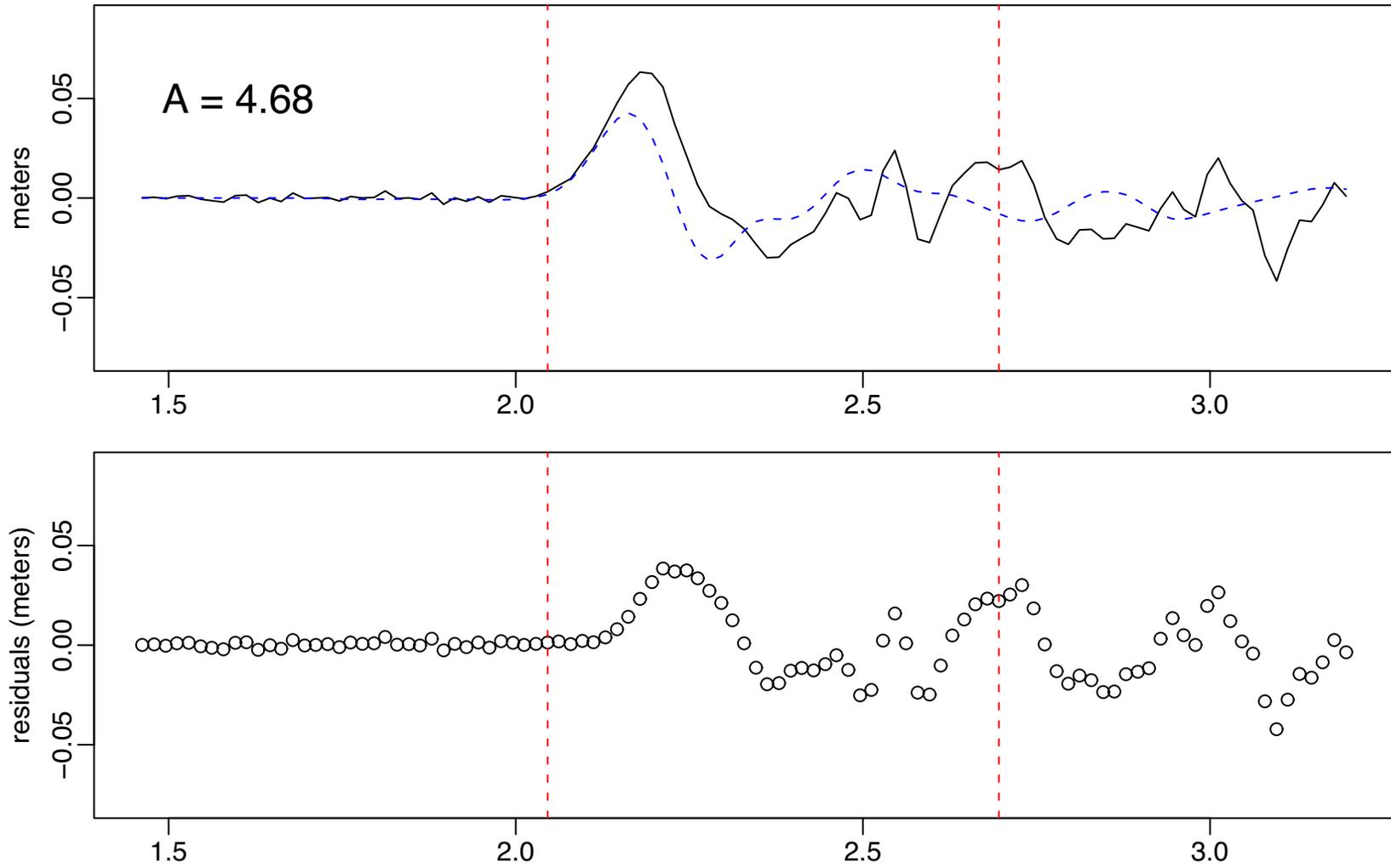
# LS Fit of a12 Model to Selected 21414 Data



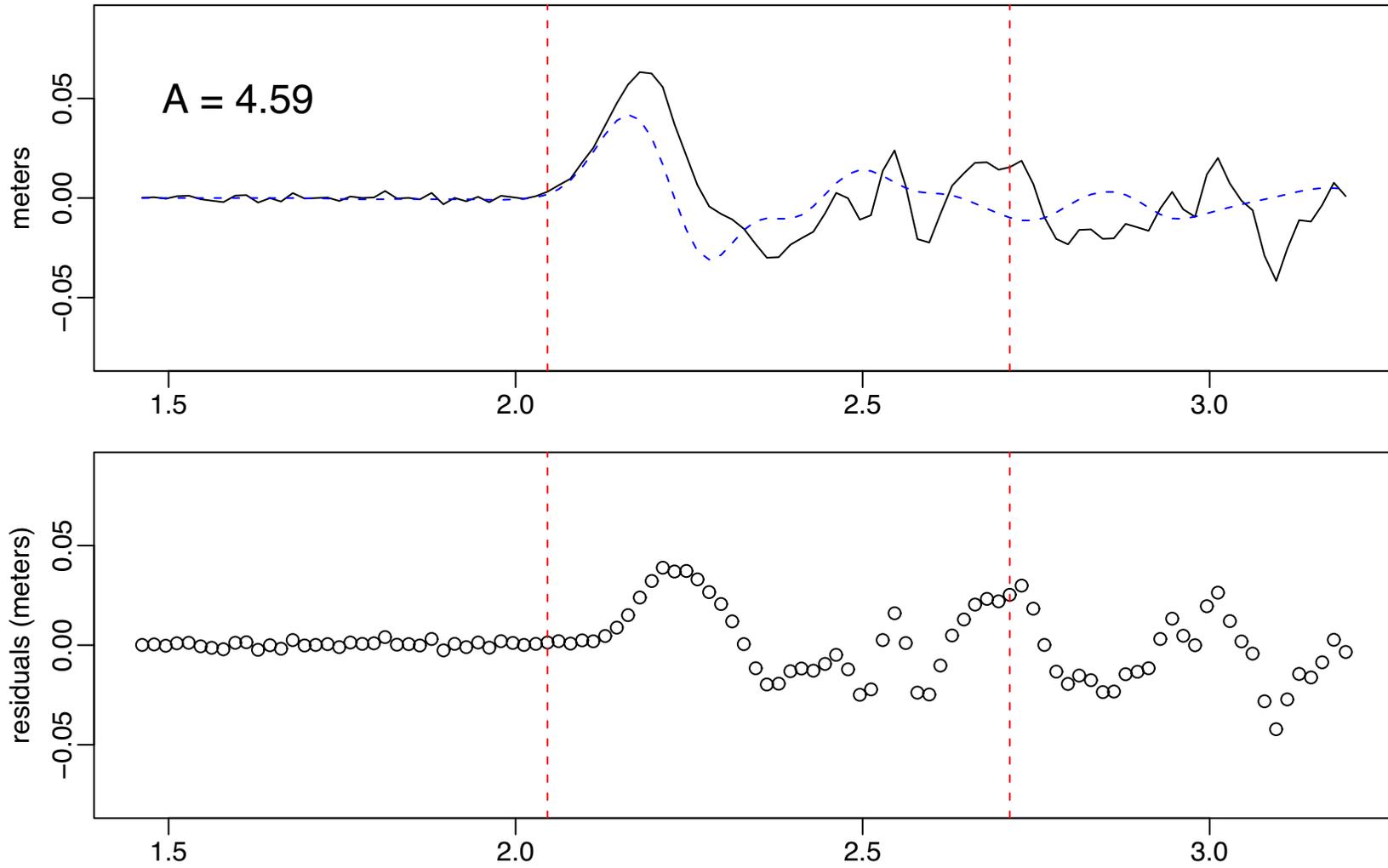
# LS Fit of a12 Model to Selected 21414 Data



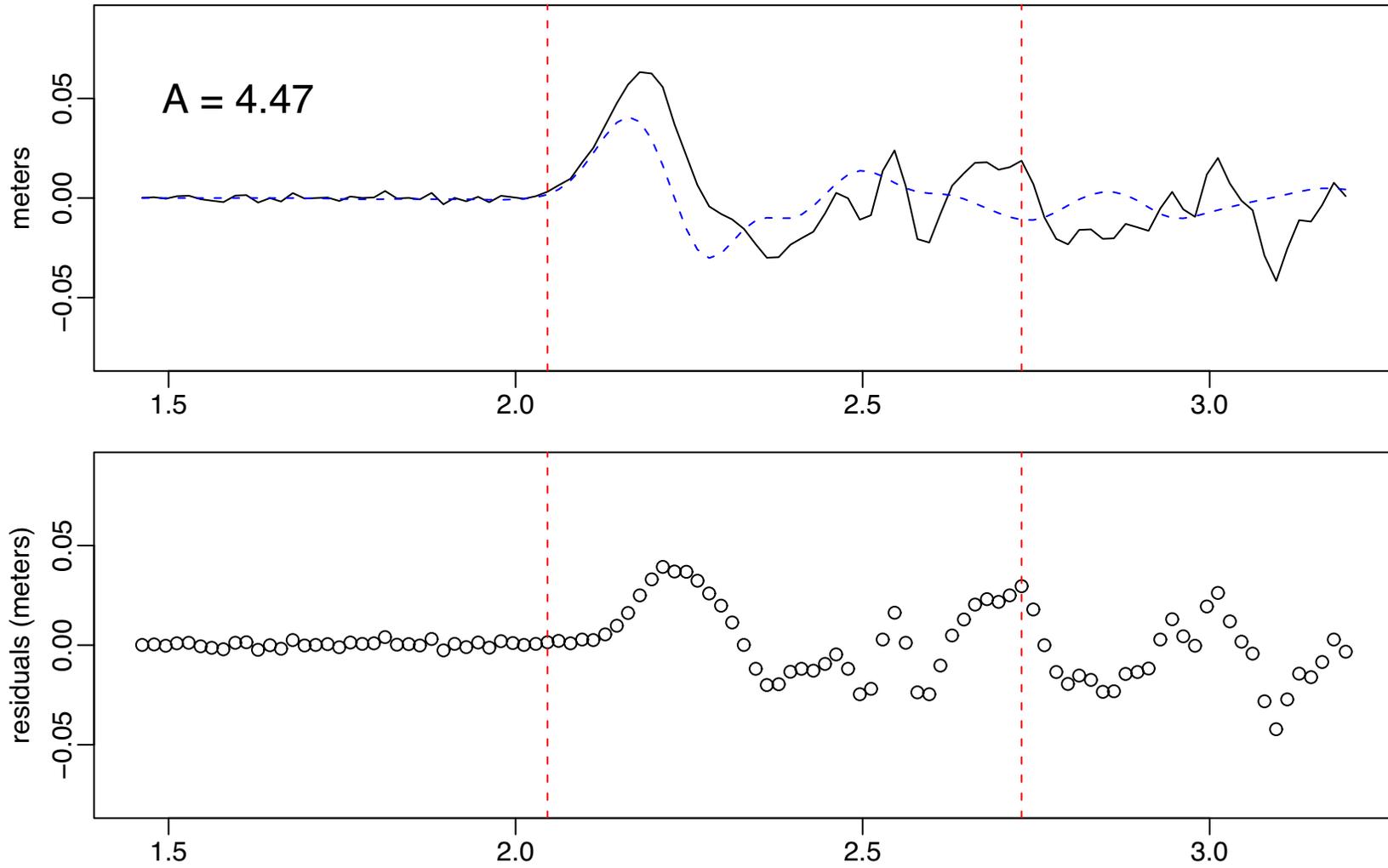
# LS Fit of a12 Model to Selected 21414 Data



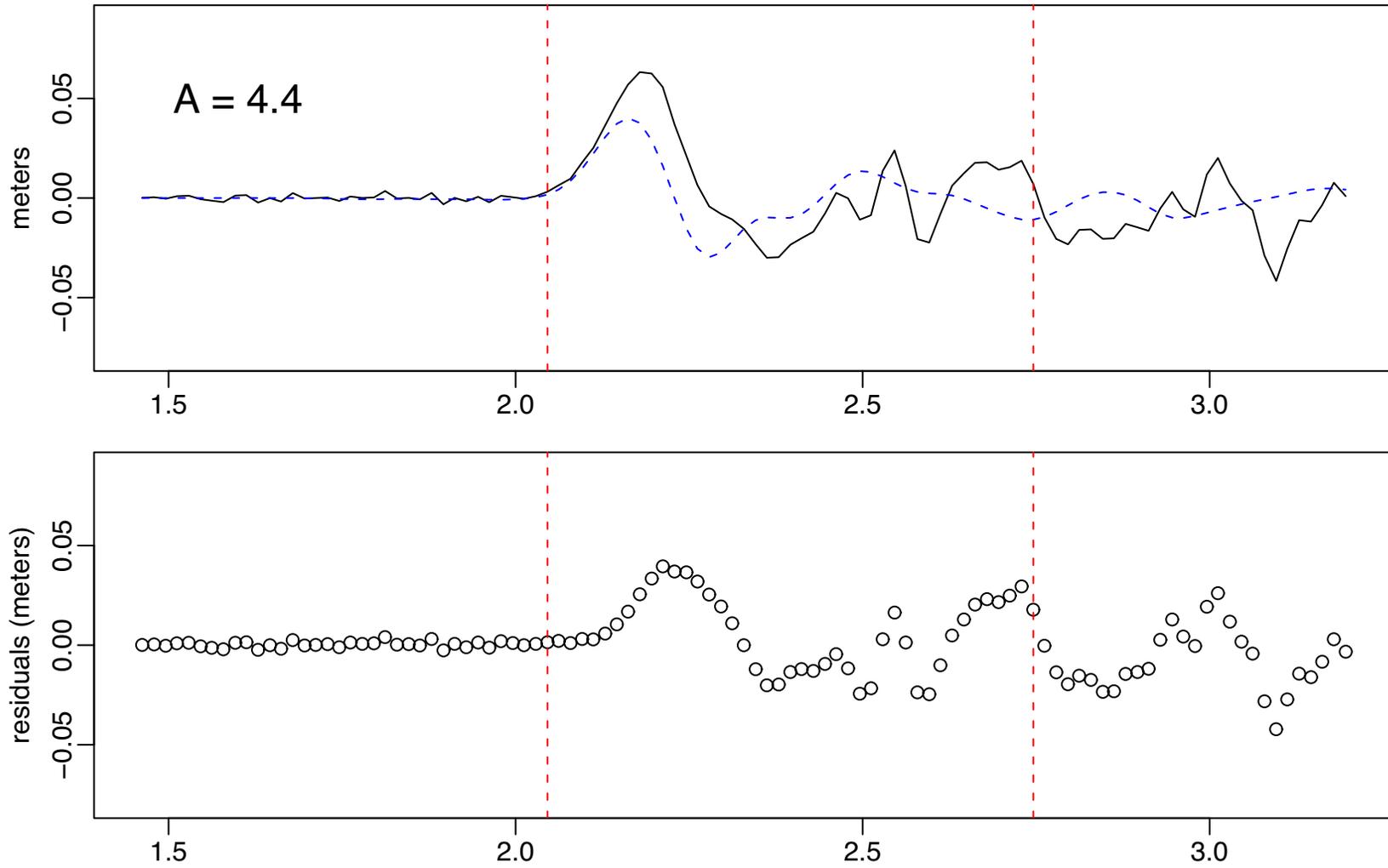
# LS Fit of a12 Model to Selected 21414 Data



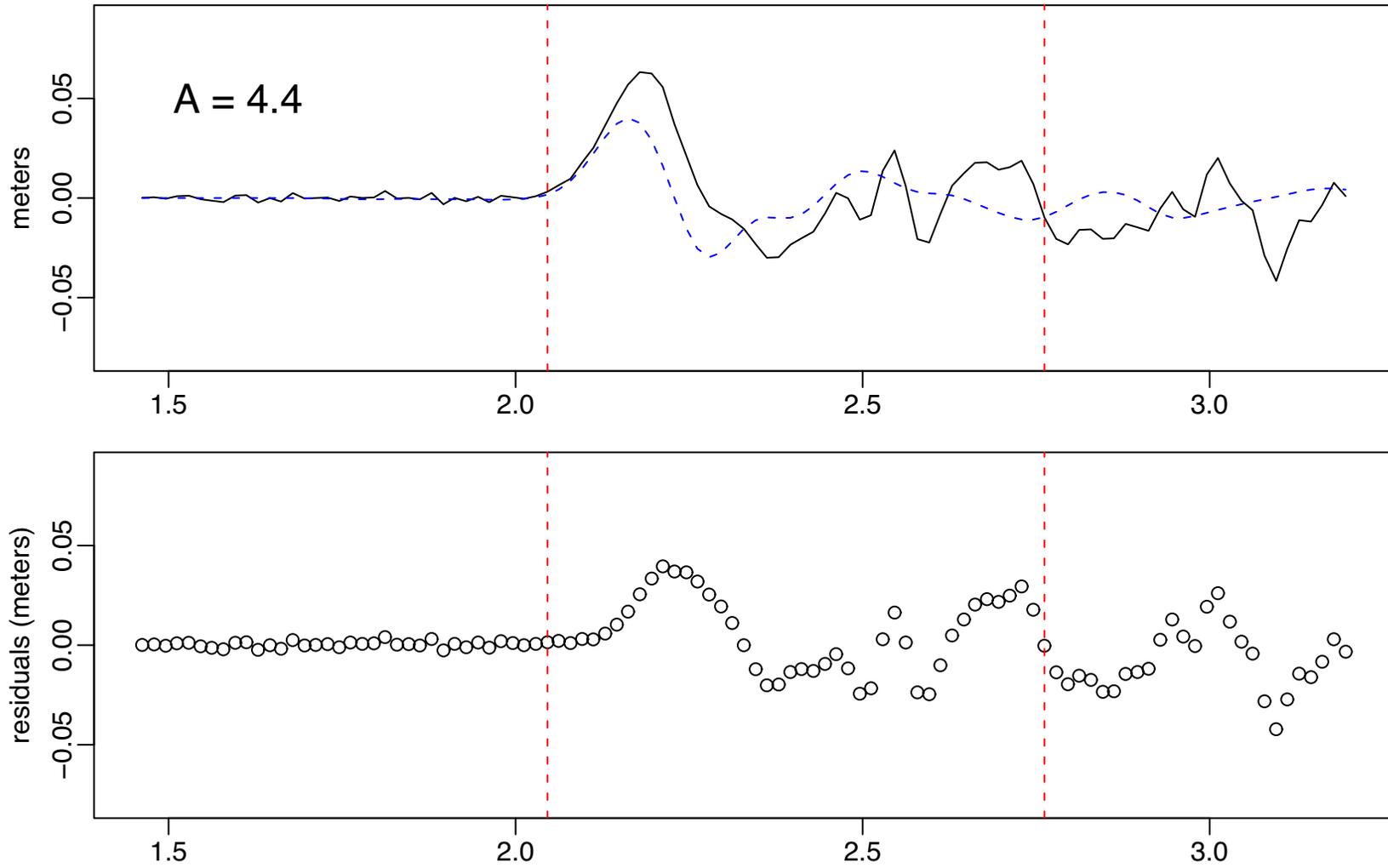
# LS Fit of a12 Model to Selected 21414 Data



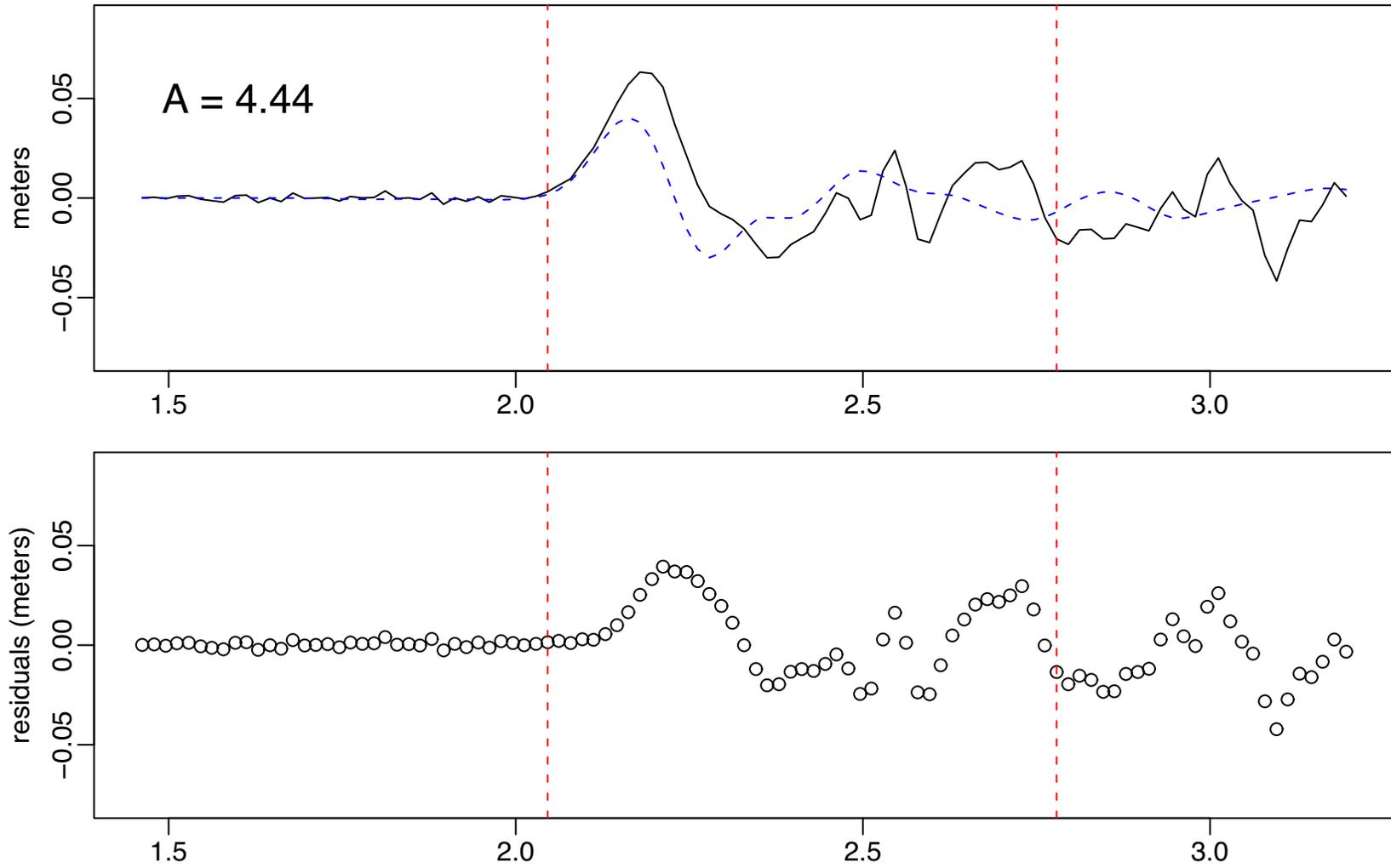
# LS Fit of a12 Model to Selected 21414 Data



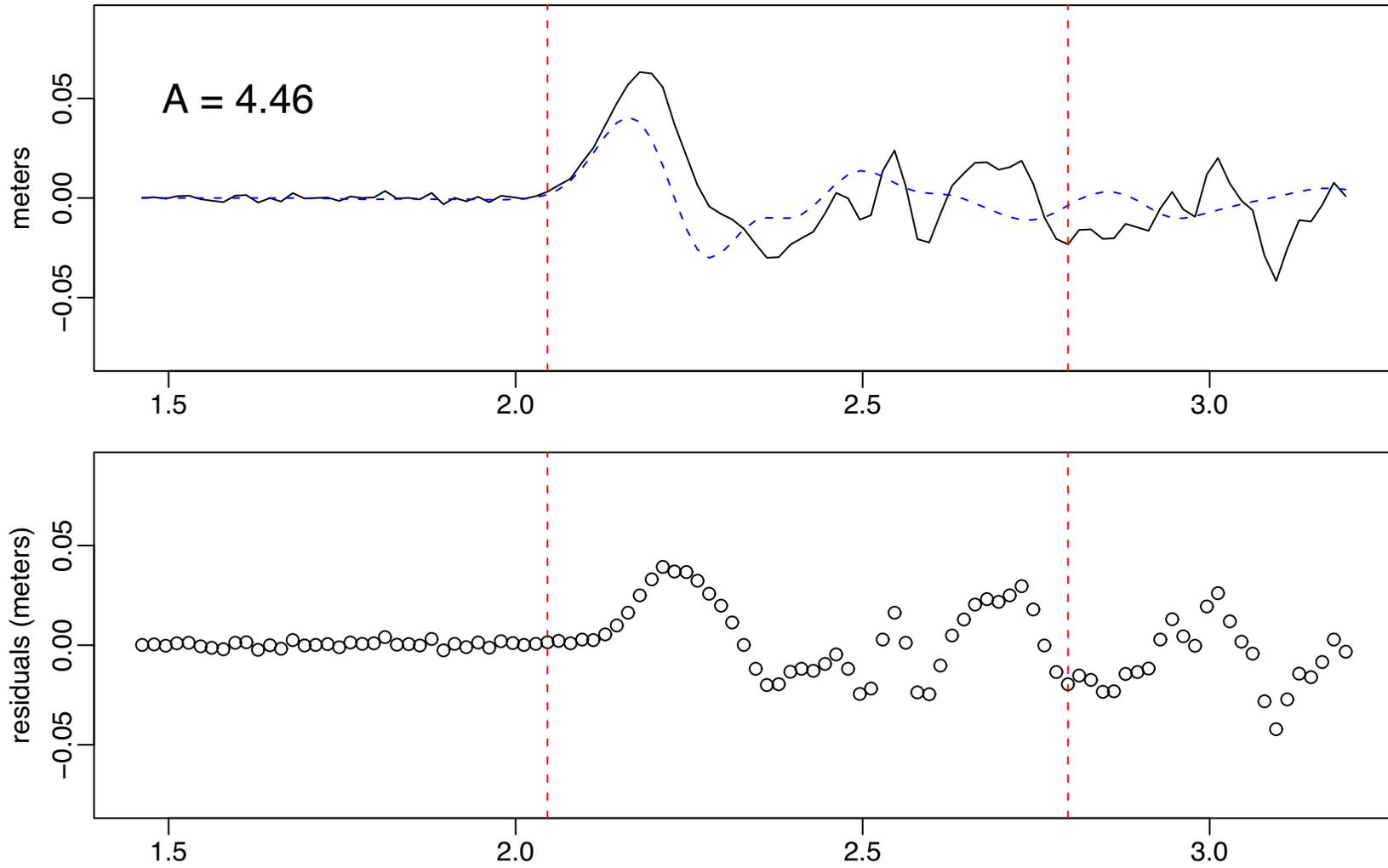
# LS Fit of a12 Model to Selected 21414 Data



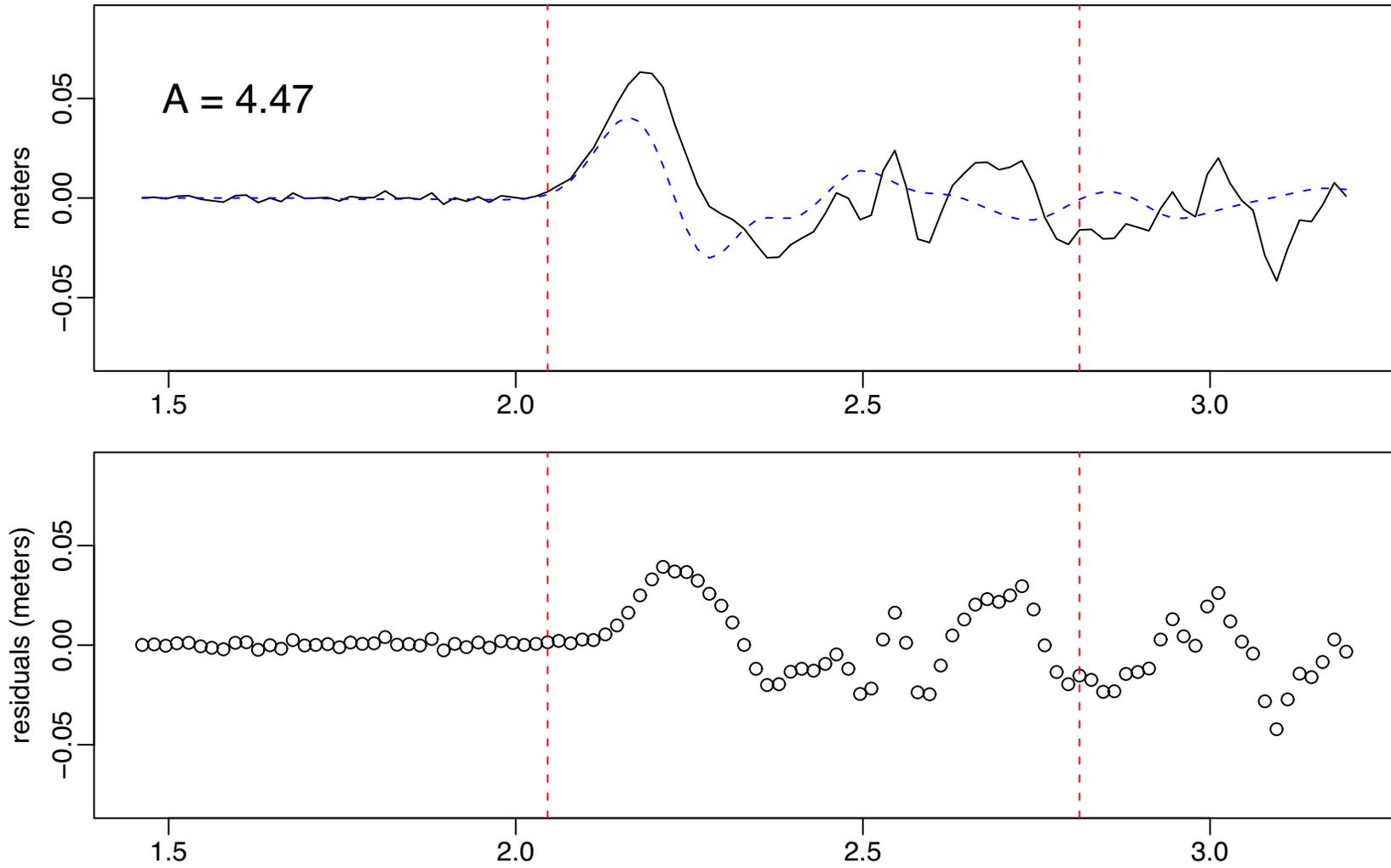
# LS Fit of a12 Model to Selected 21414 Data



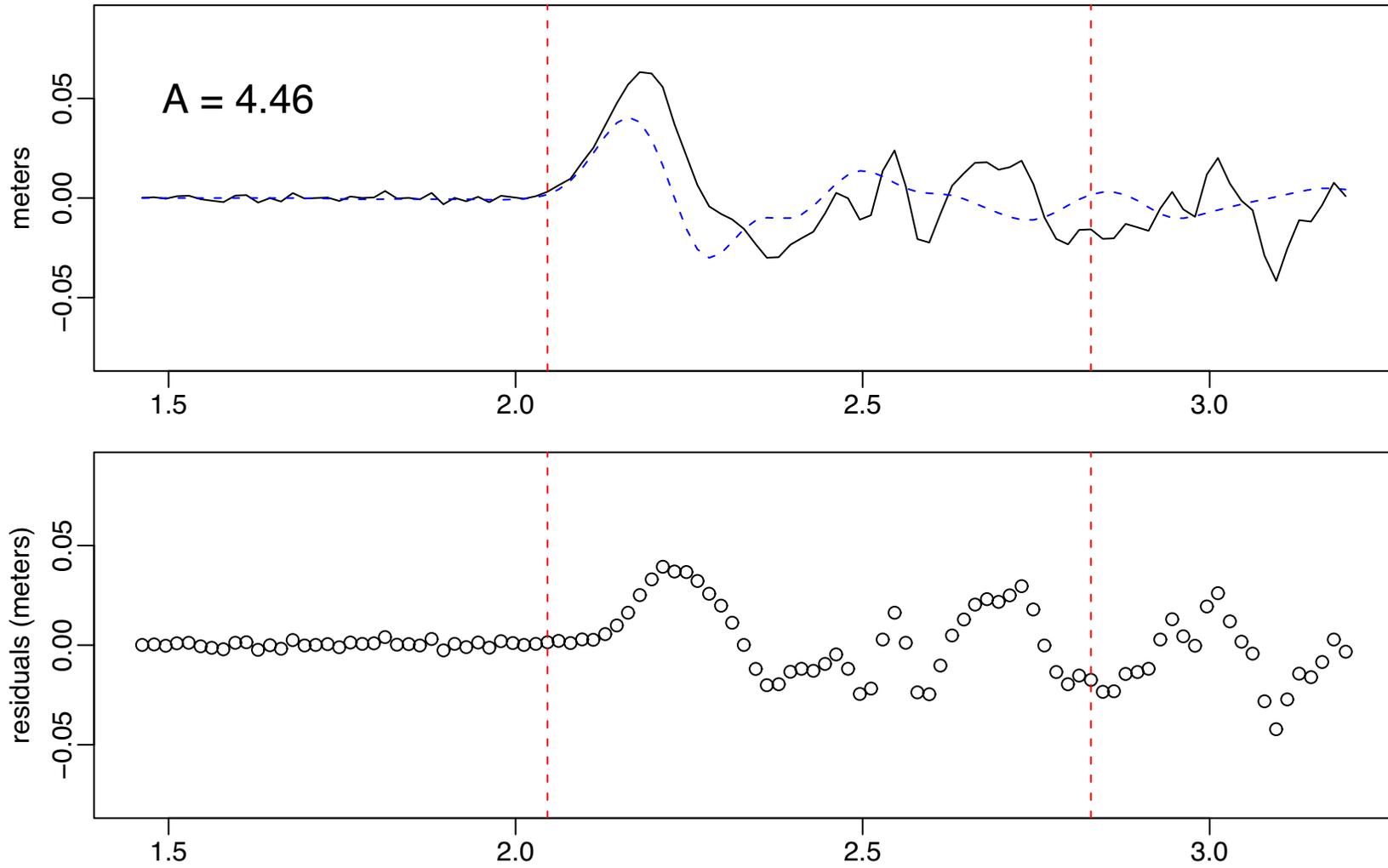
# LS Fit of a12 Model to Selected 21414 Data



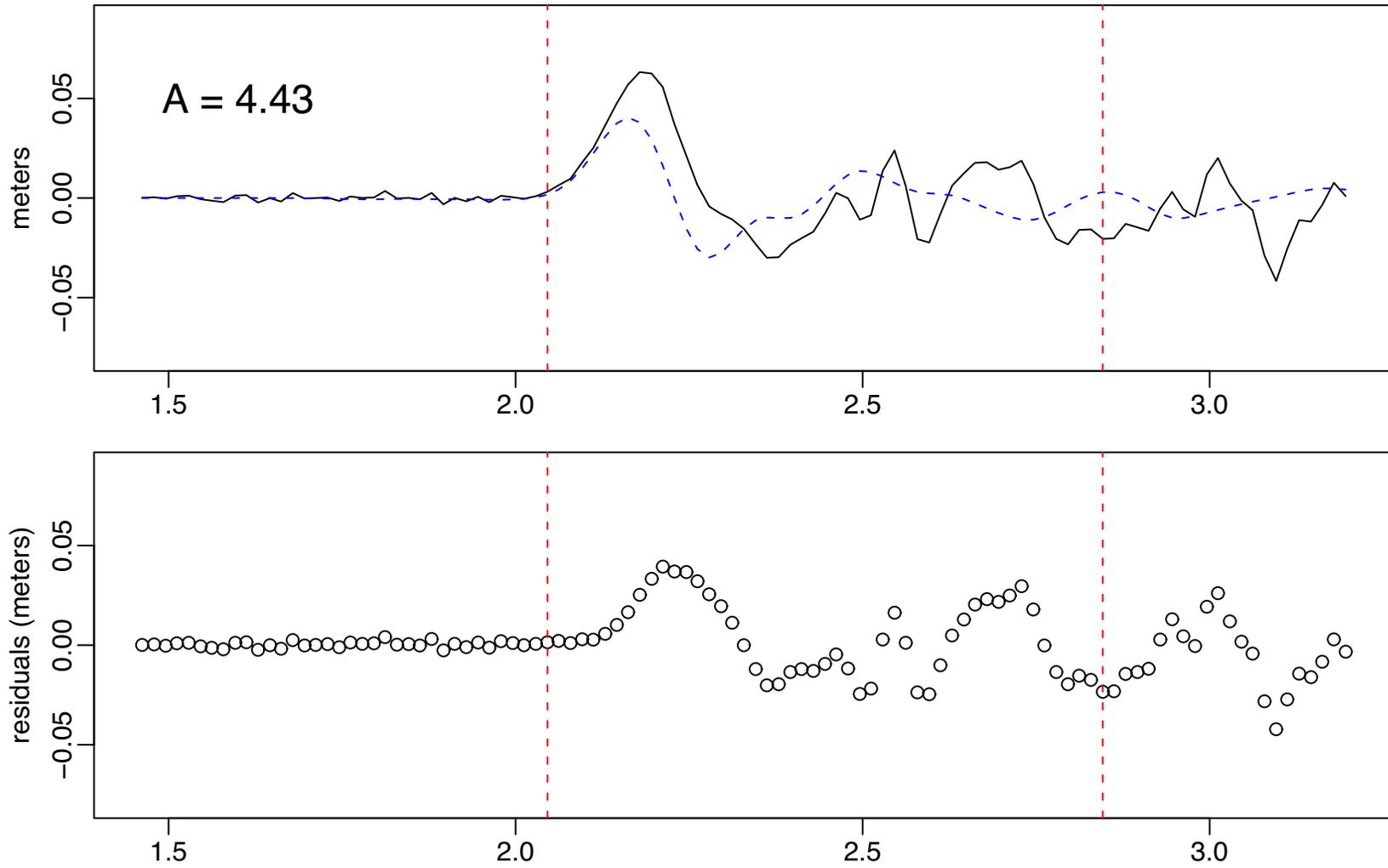
# LS Fit of a12 Model to Selected 21414 Data



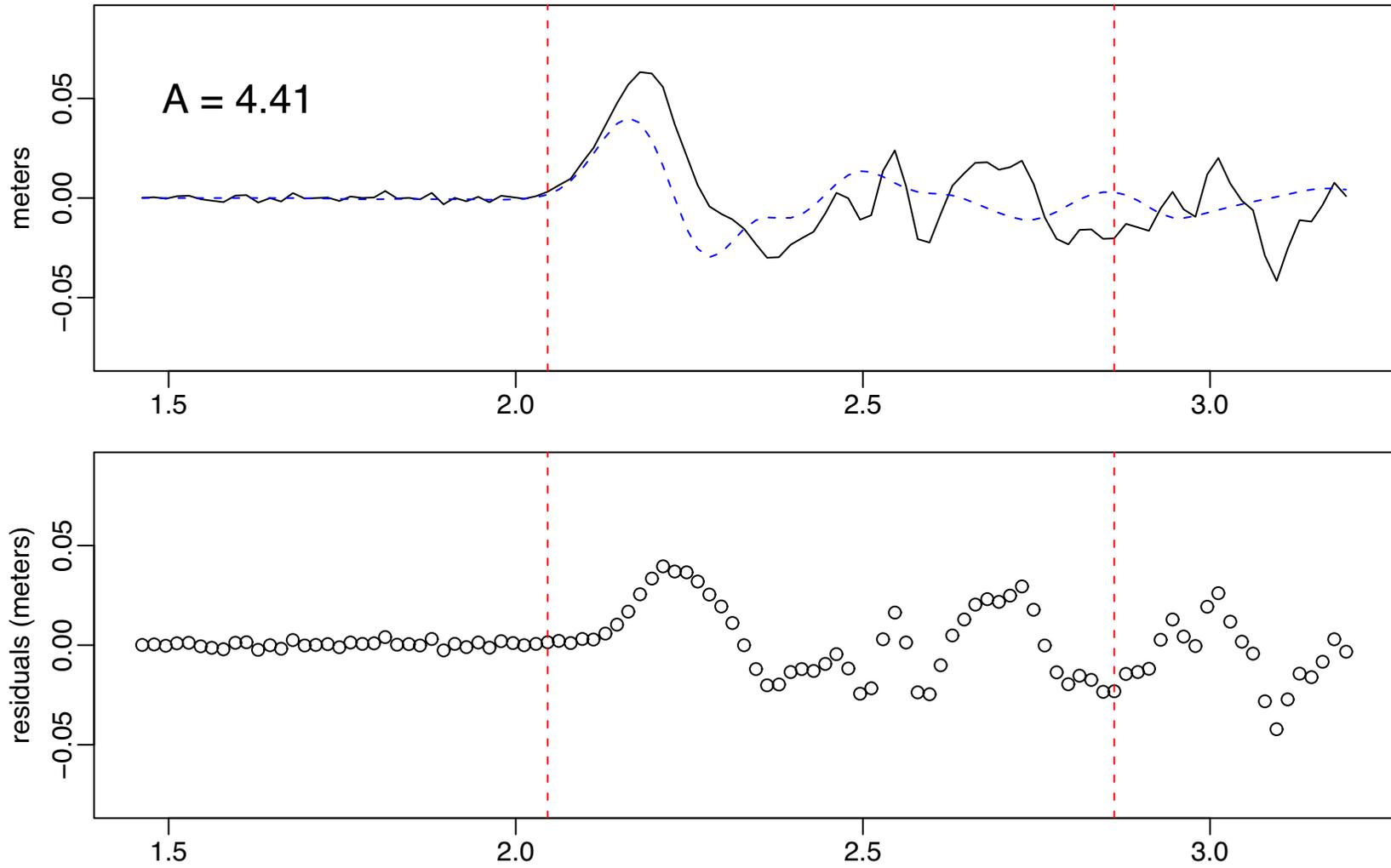
# LS Fit of a12 Model to Selected 21414 Data



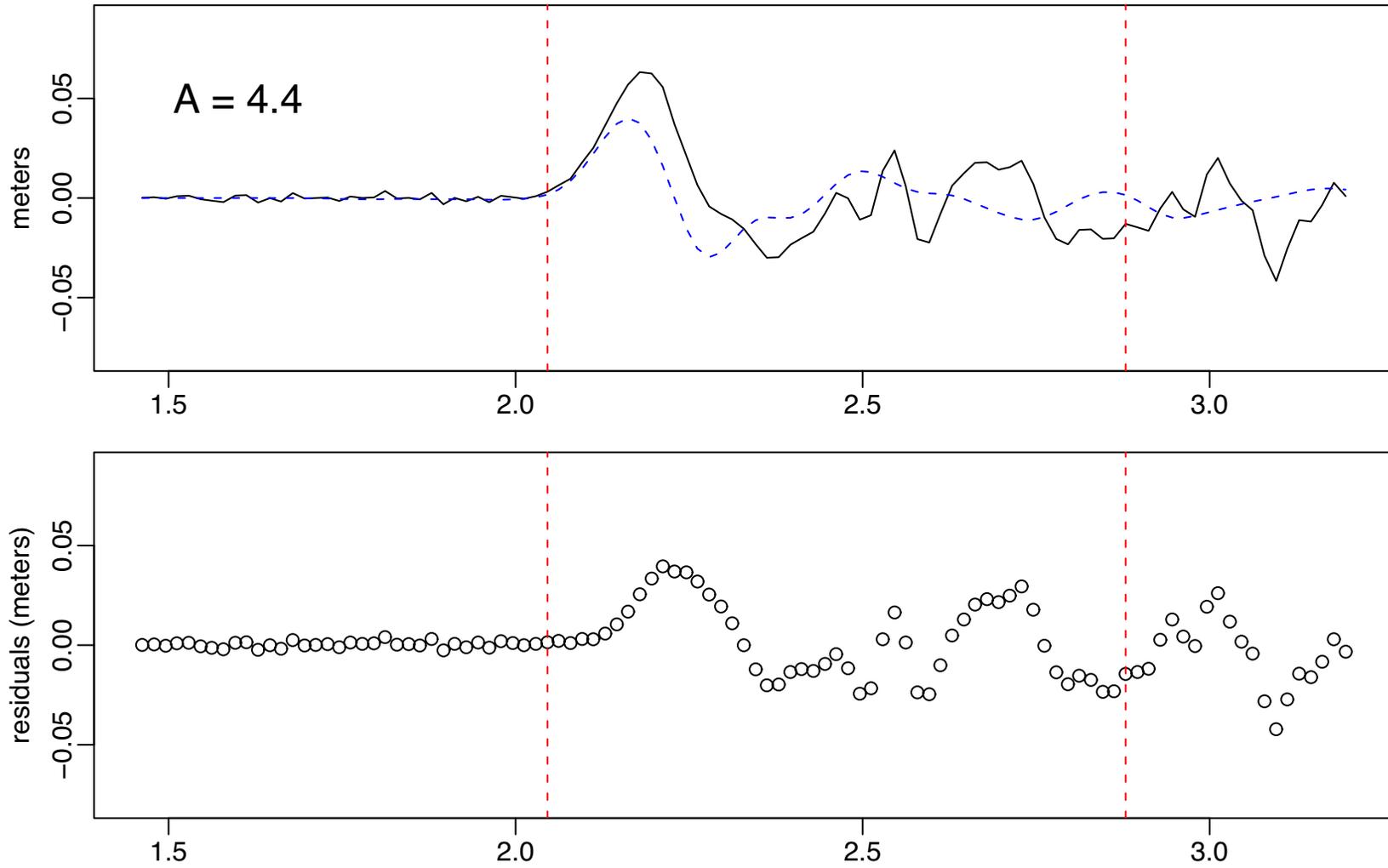
# LS Fit of a12 Model to Selected 21414 Data



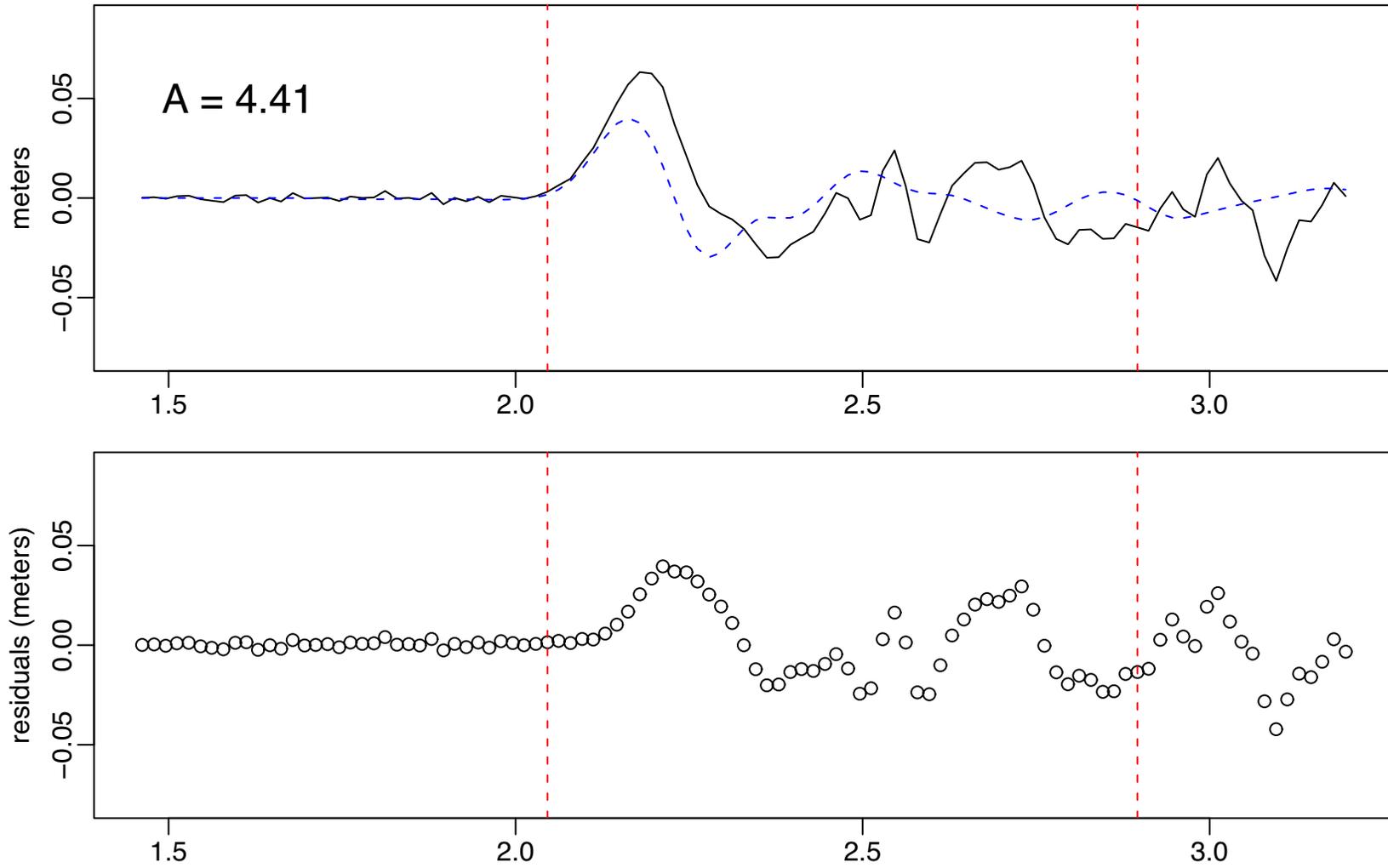
# LS Fit of a12 Model to Selected 21414 Data



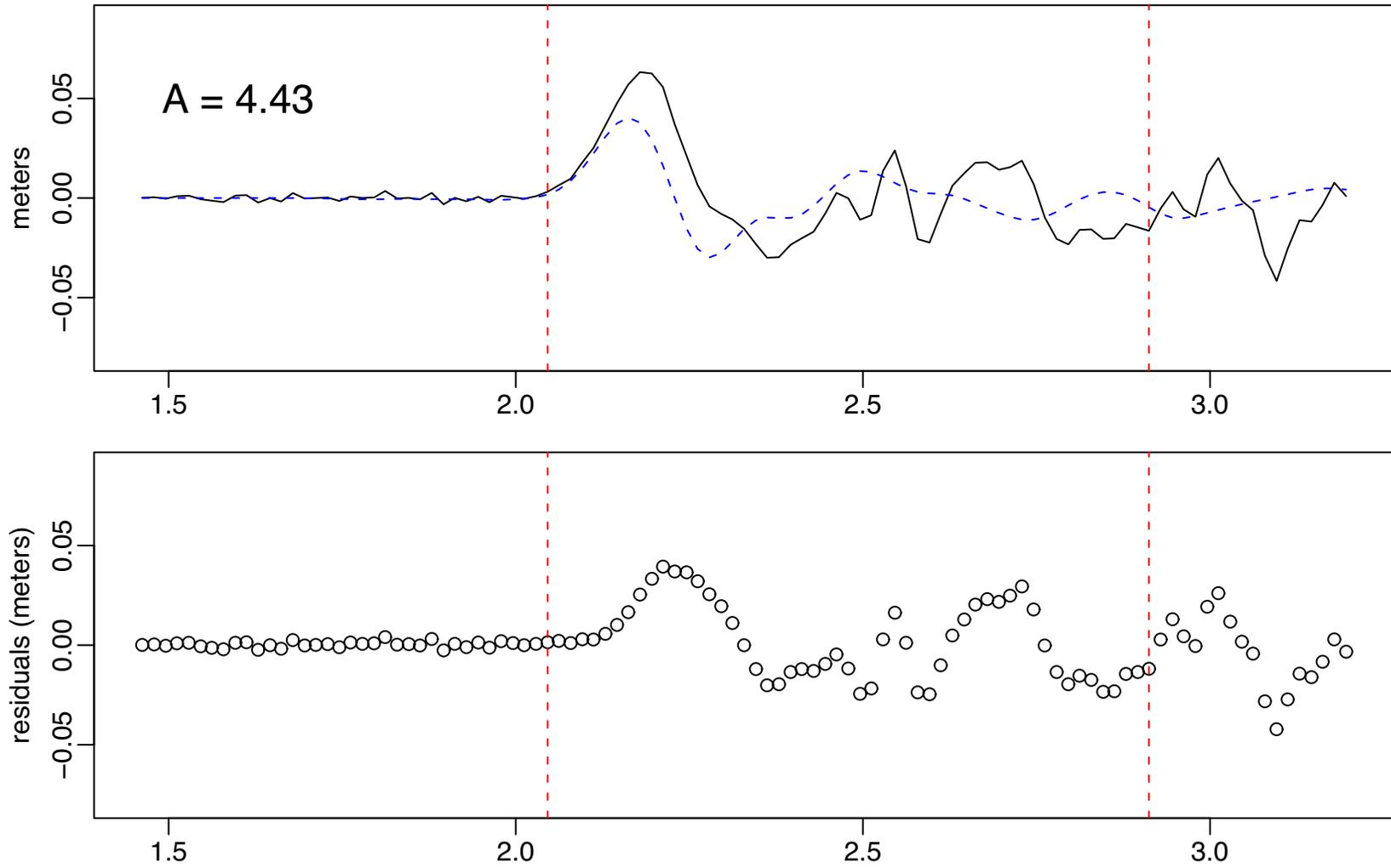
# LS Fit of a12 Model to Selected 21414 Data



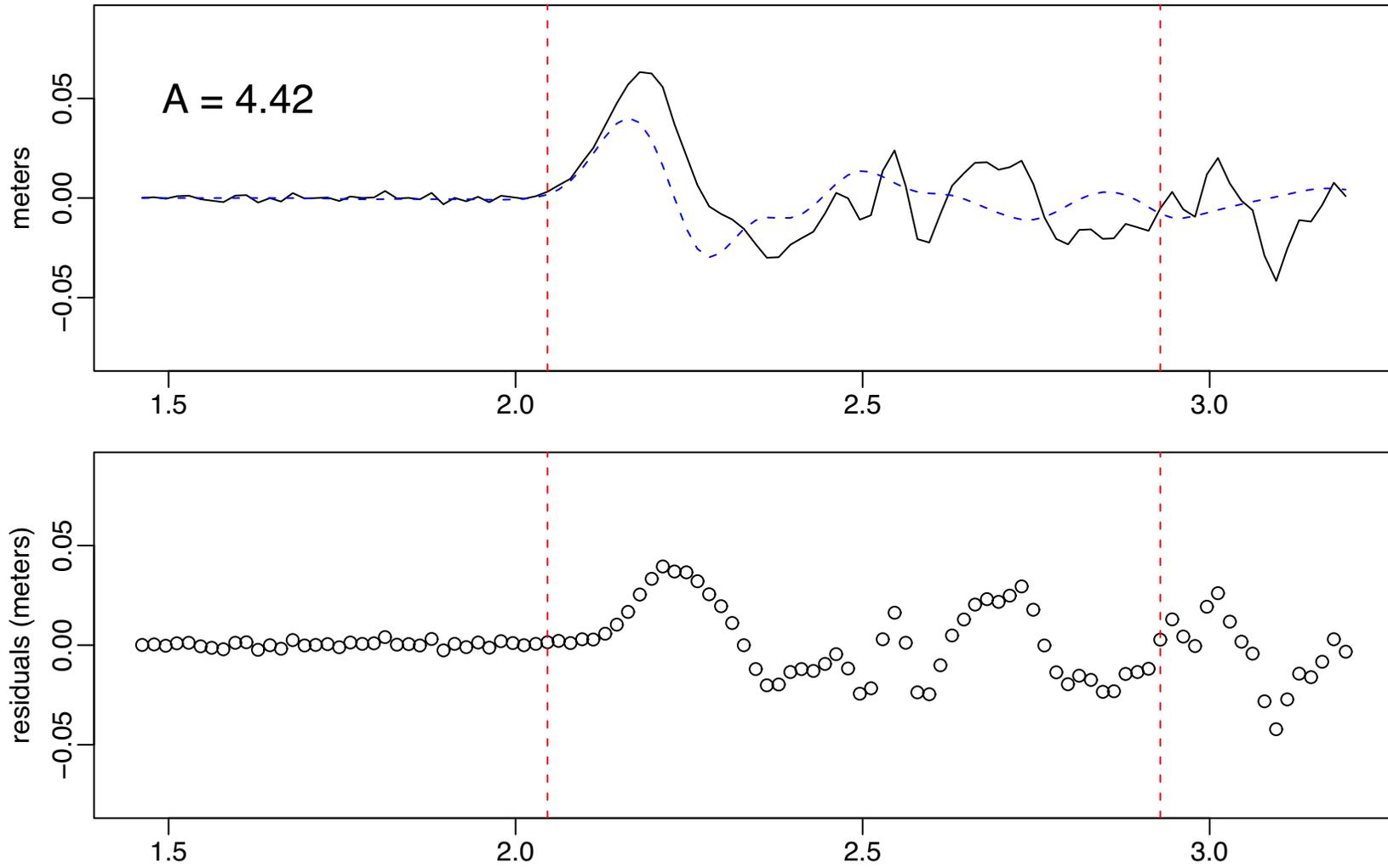
# LS Fit of a12 Model to Selected 21414 Data



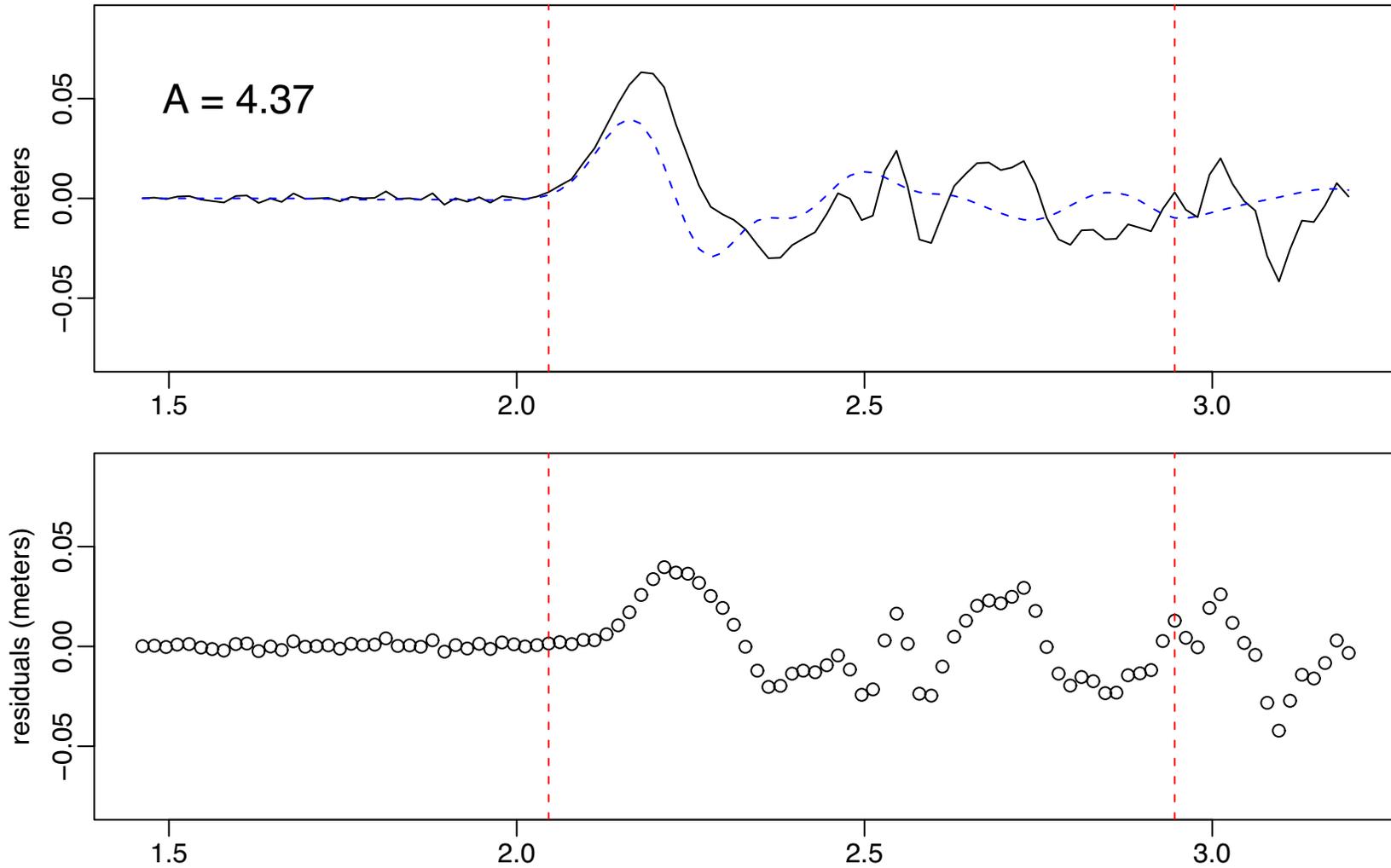
# LS Fit of a12 Model to Selected 21414 Data



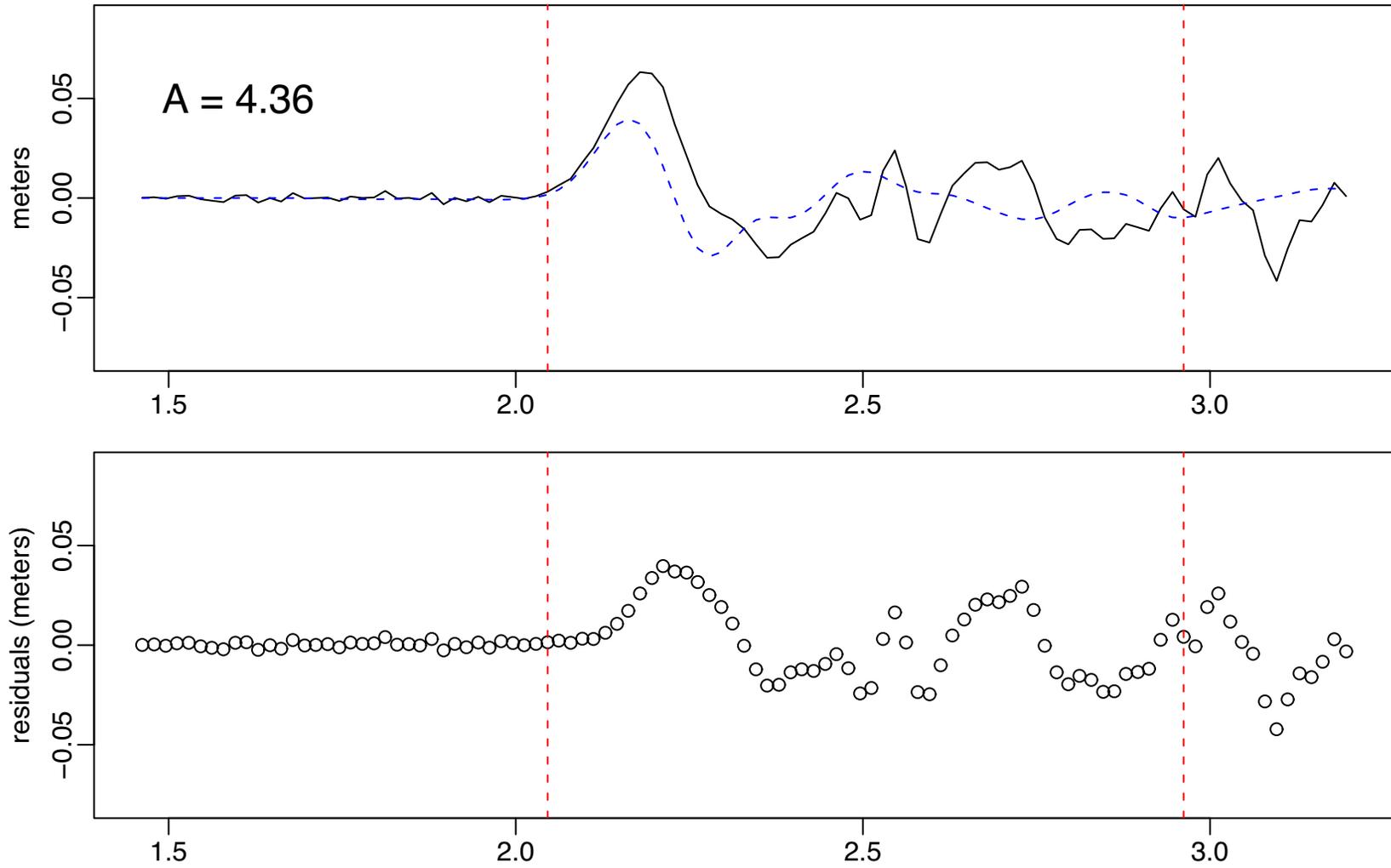
# LS Fit of a12 Model to Selected 21414 Data



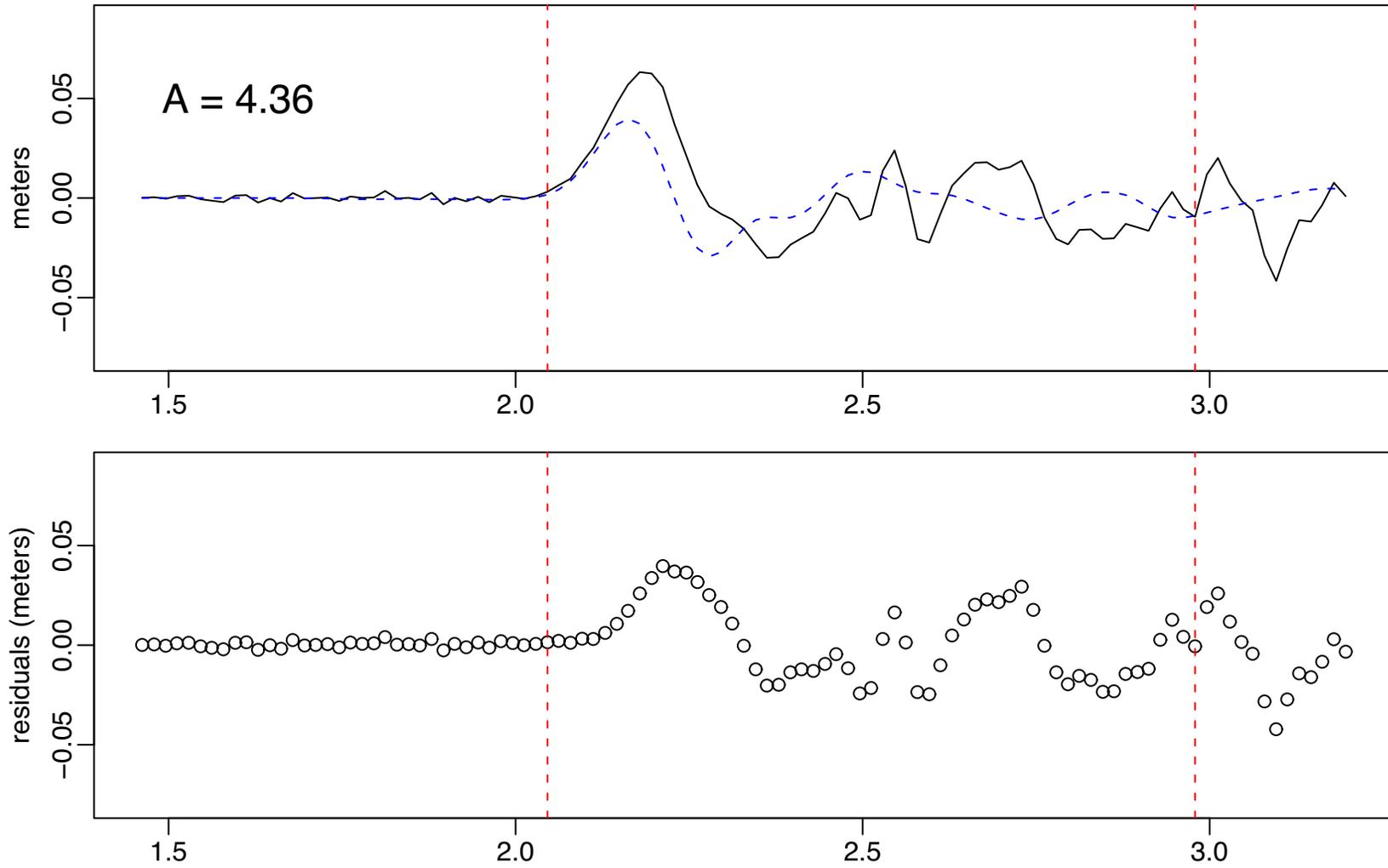
# LS Fit of a12 Model to Selected 21414 Data



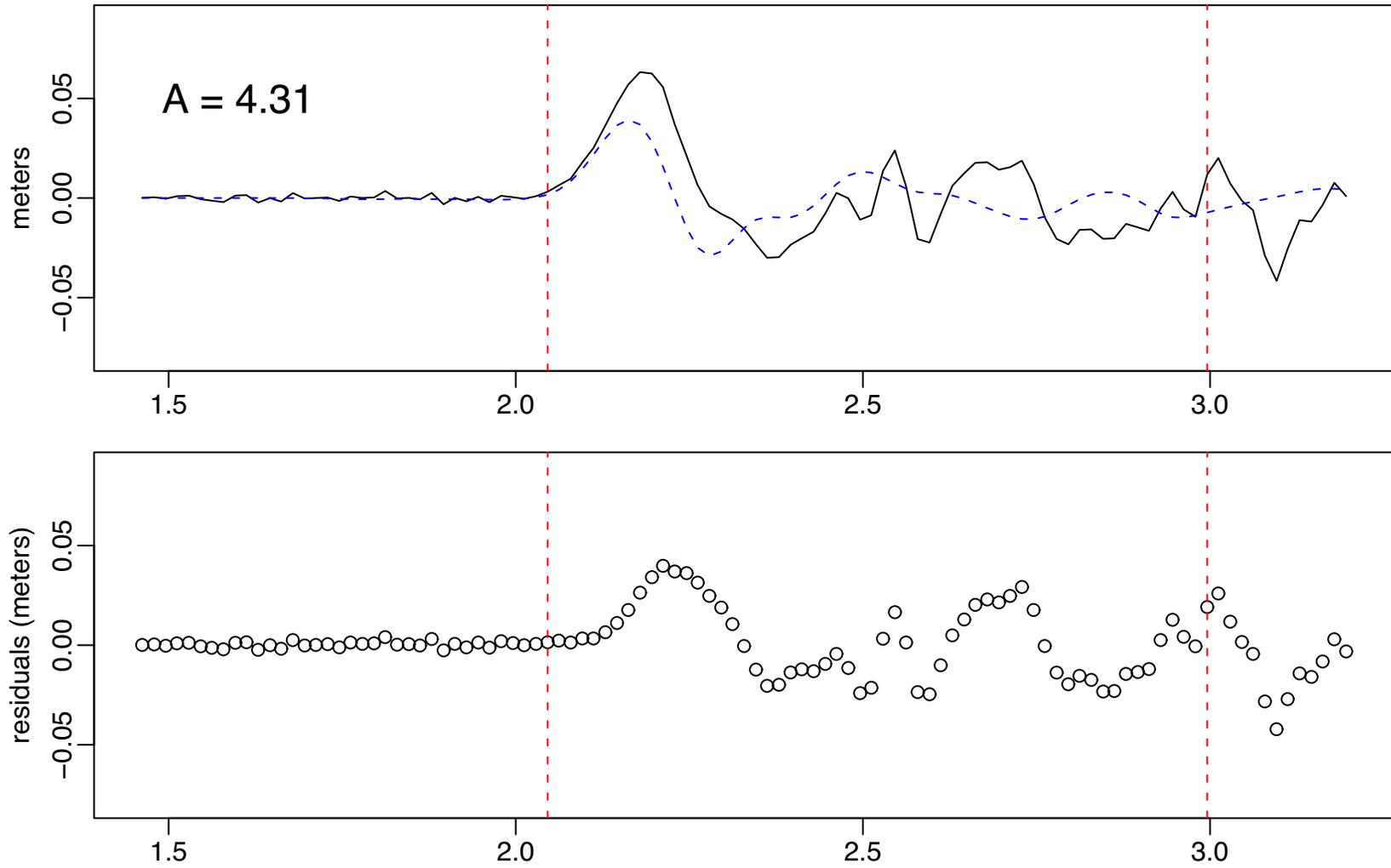
# LS Fit of a12 Model to Selected 21414 Data



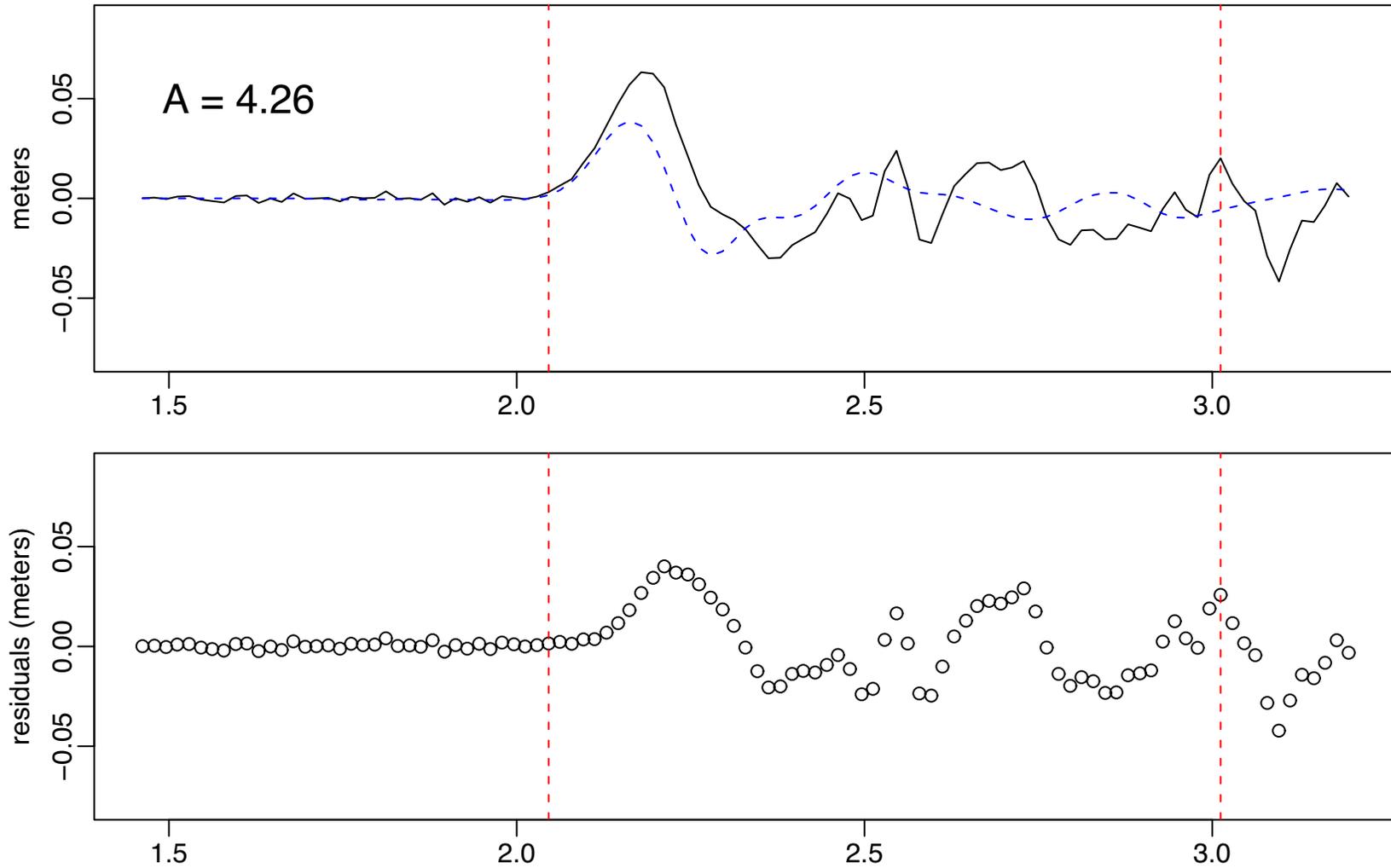
# LS Fit of a12 Model to Selected 21414 Data



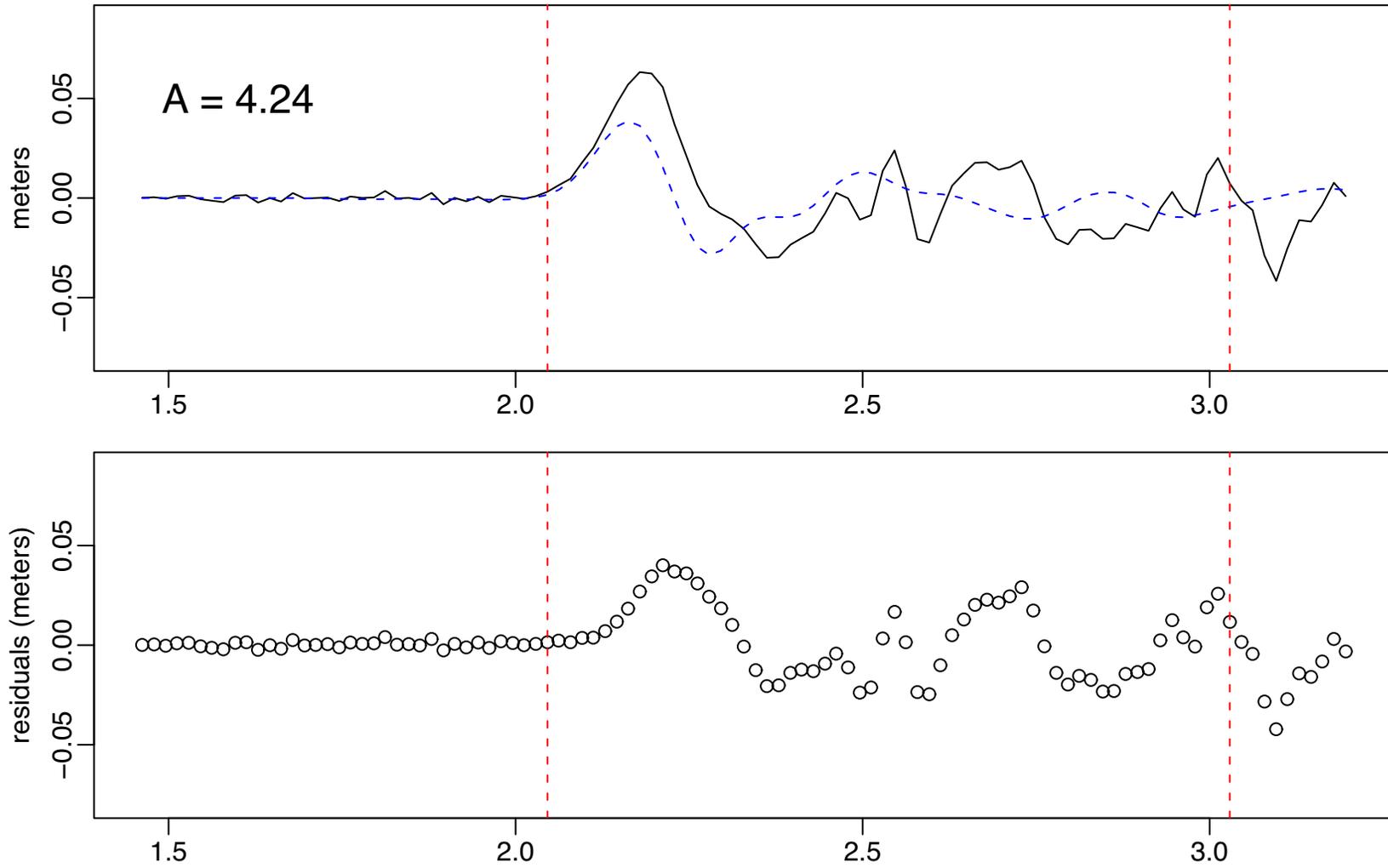
# LS Fit of a12 Model to Selected 21414 Data



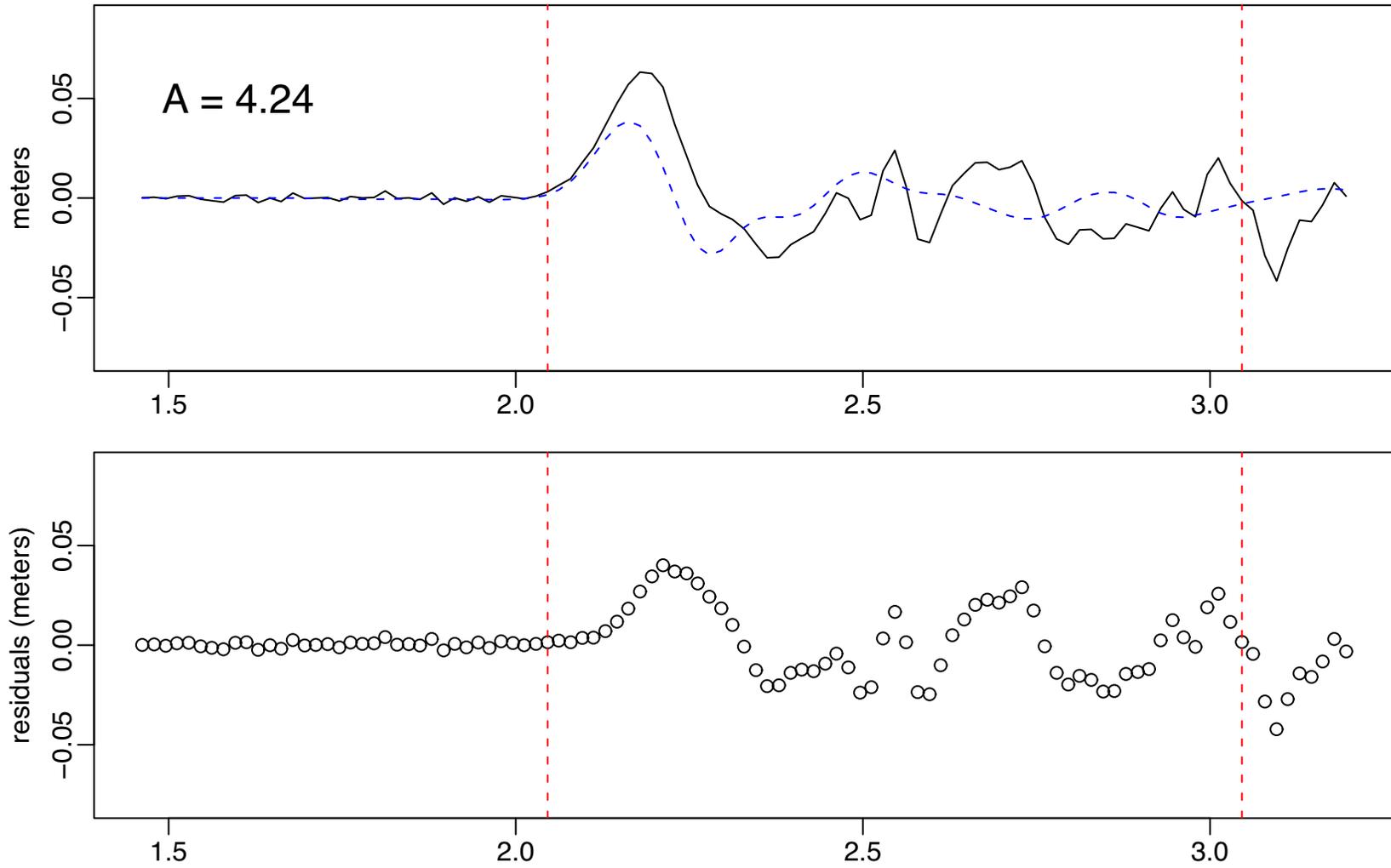
# LS Fit of a12 Model to Selected 21414 Data



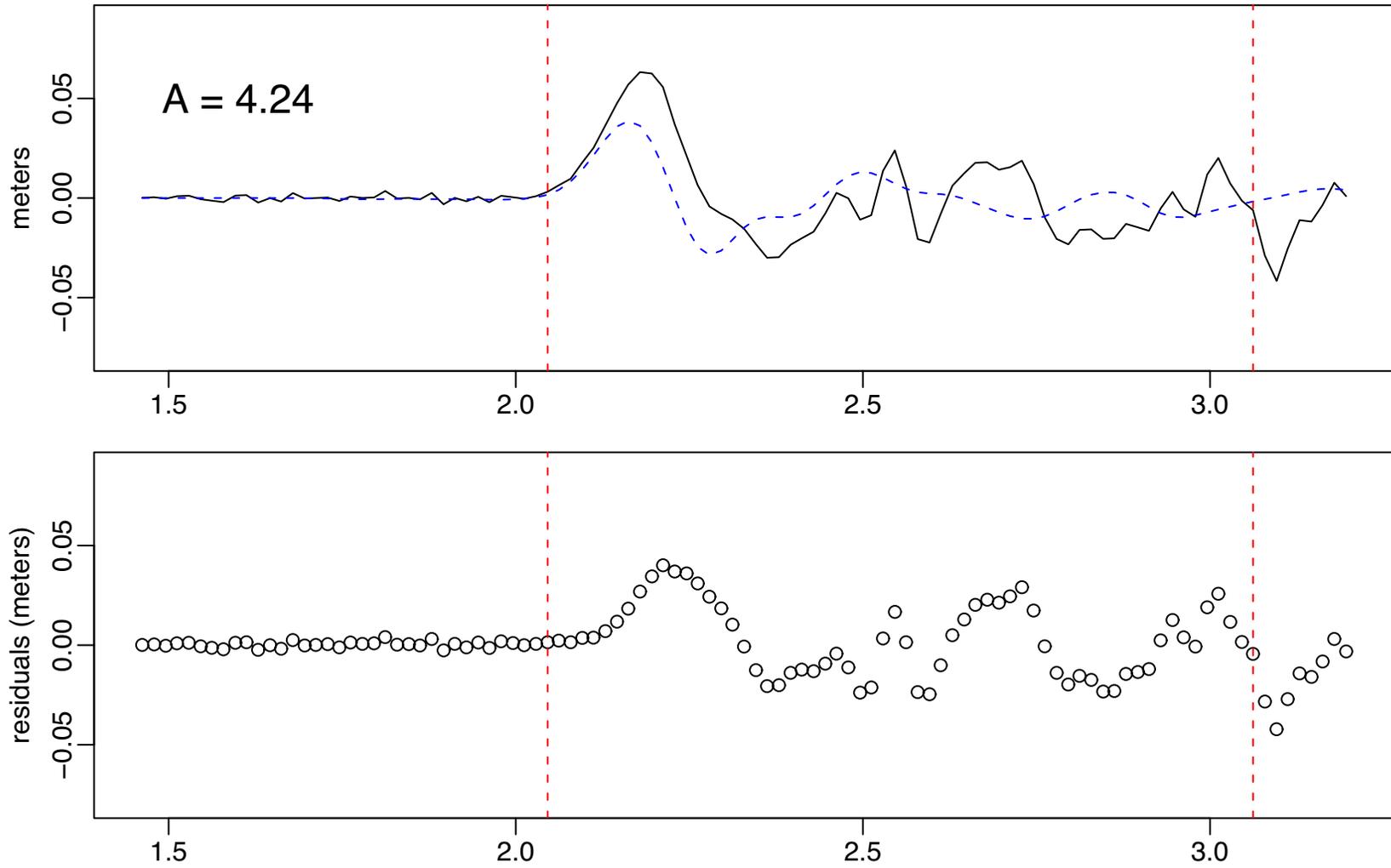
# LS Fit of a12 Model to Selected 21414 Data



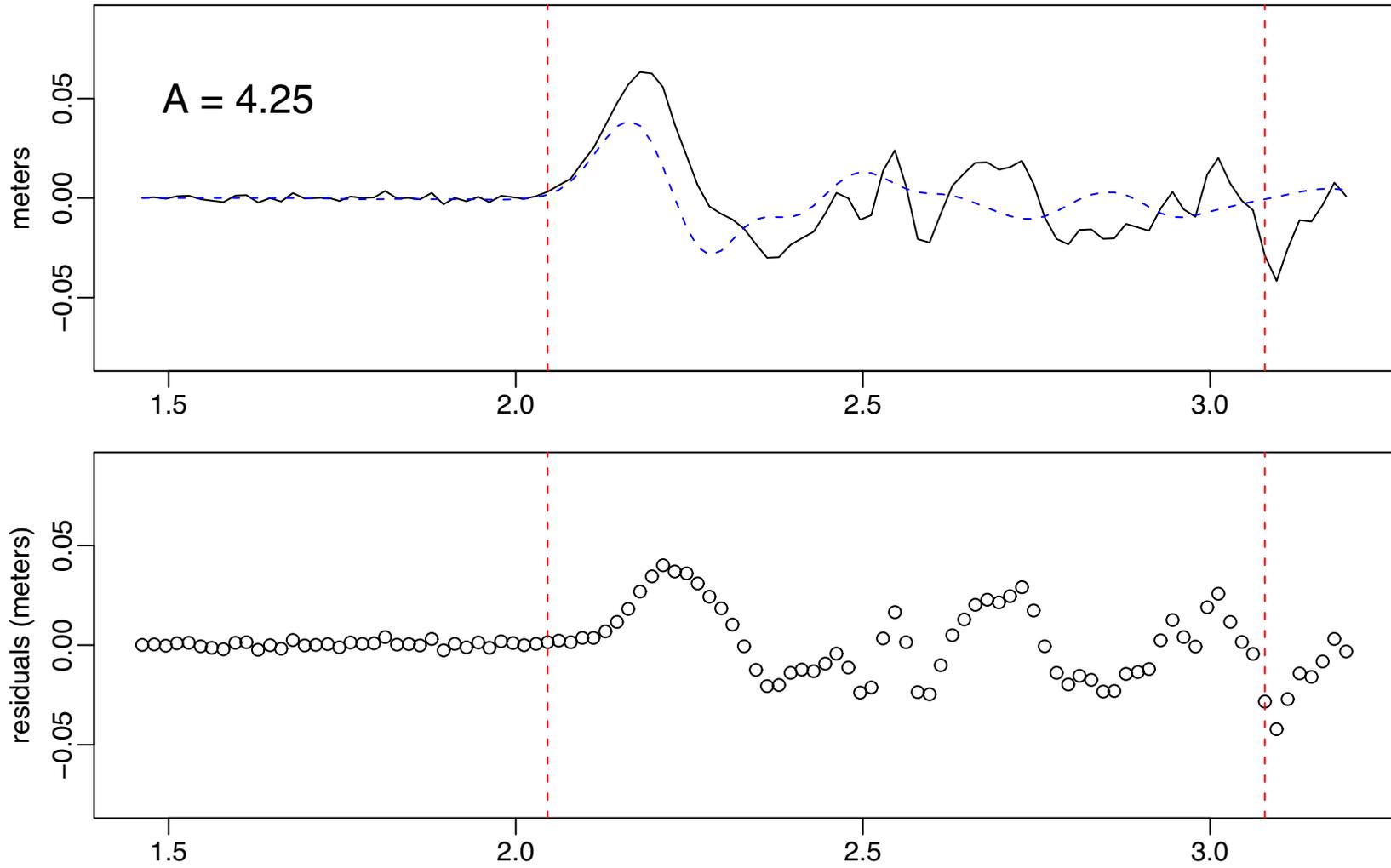
# LS Fit of a12 Model to Selected 21414 Data



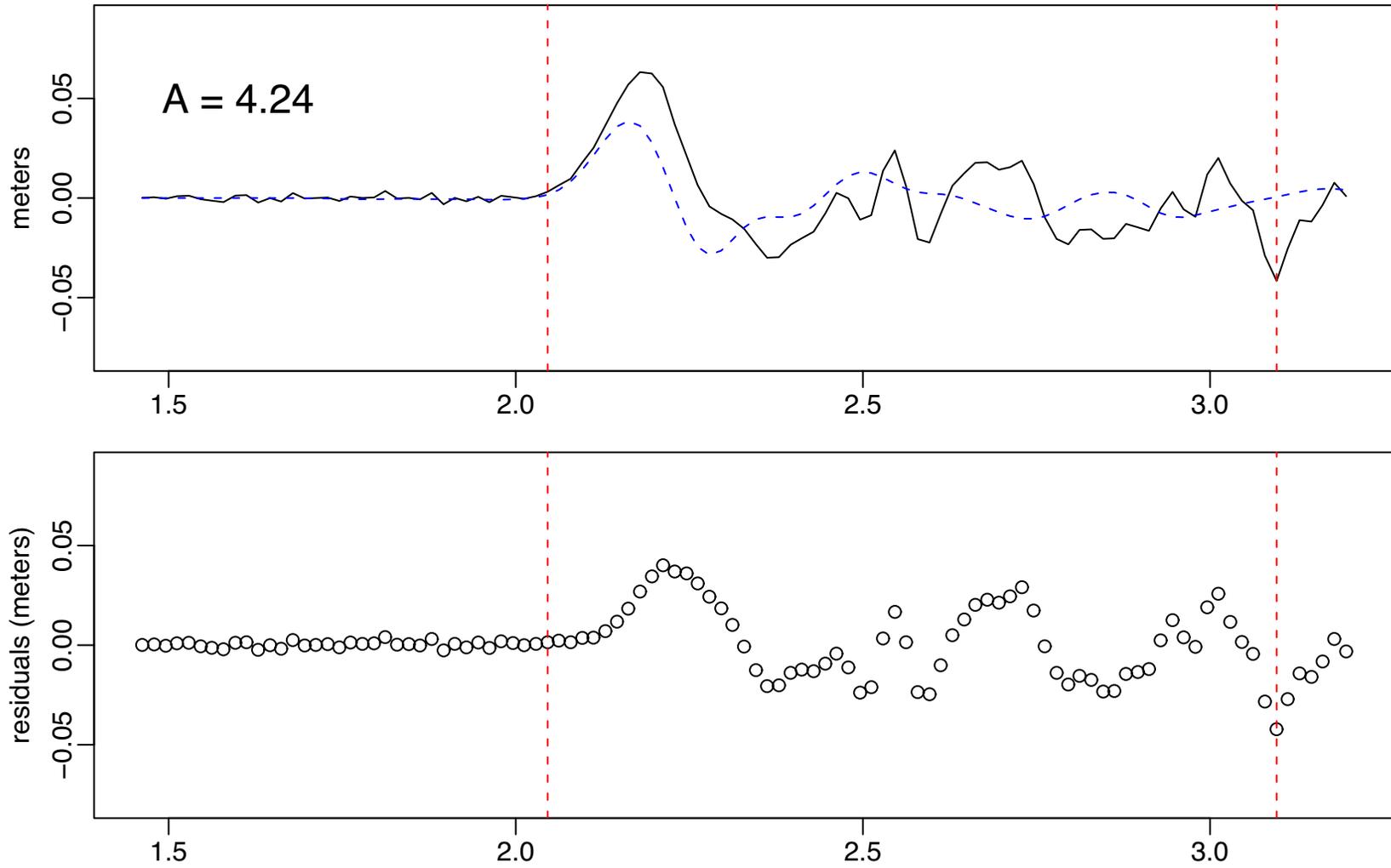
# LS Fit of a12 Model to Selected 21414 Data



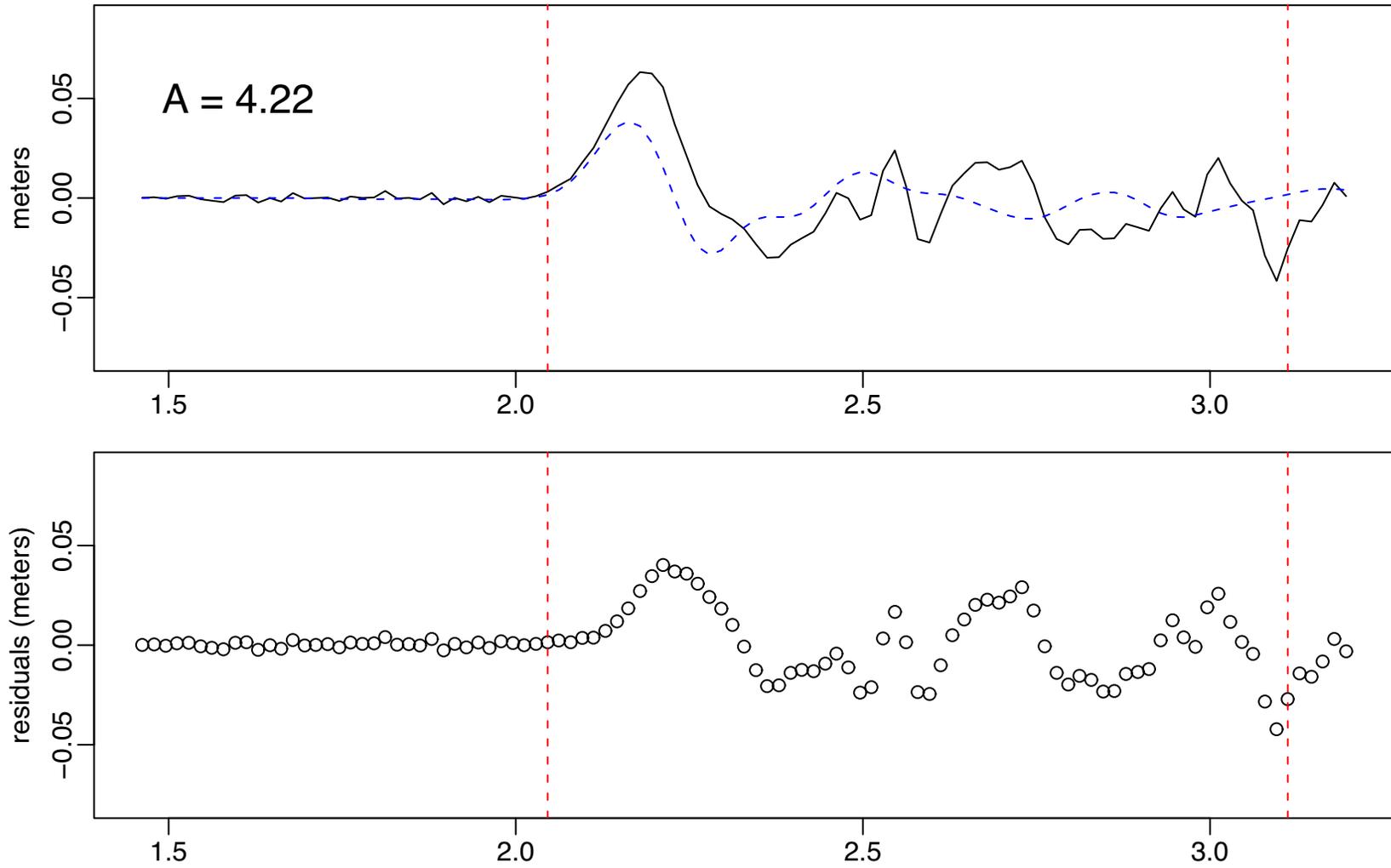
# LS Fit of a12 Model to Selected 21414 Data



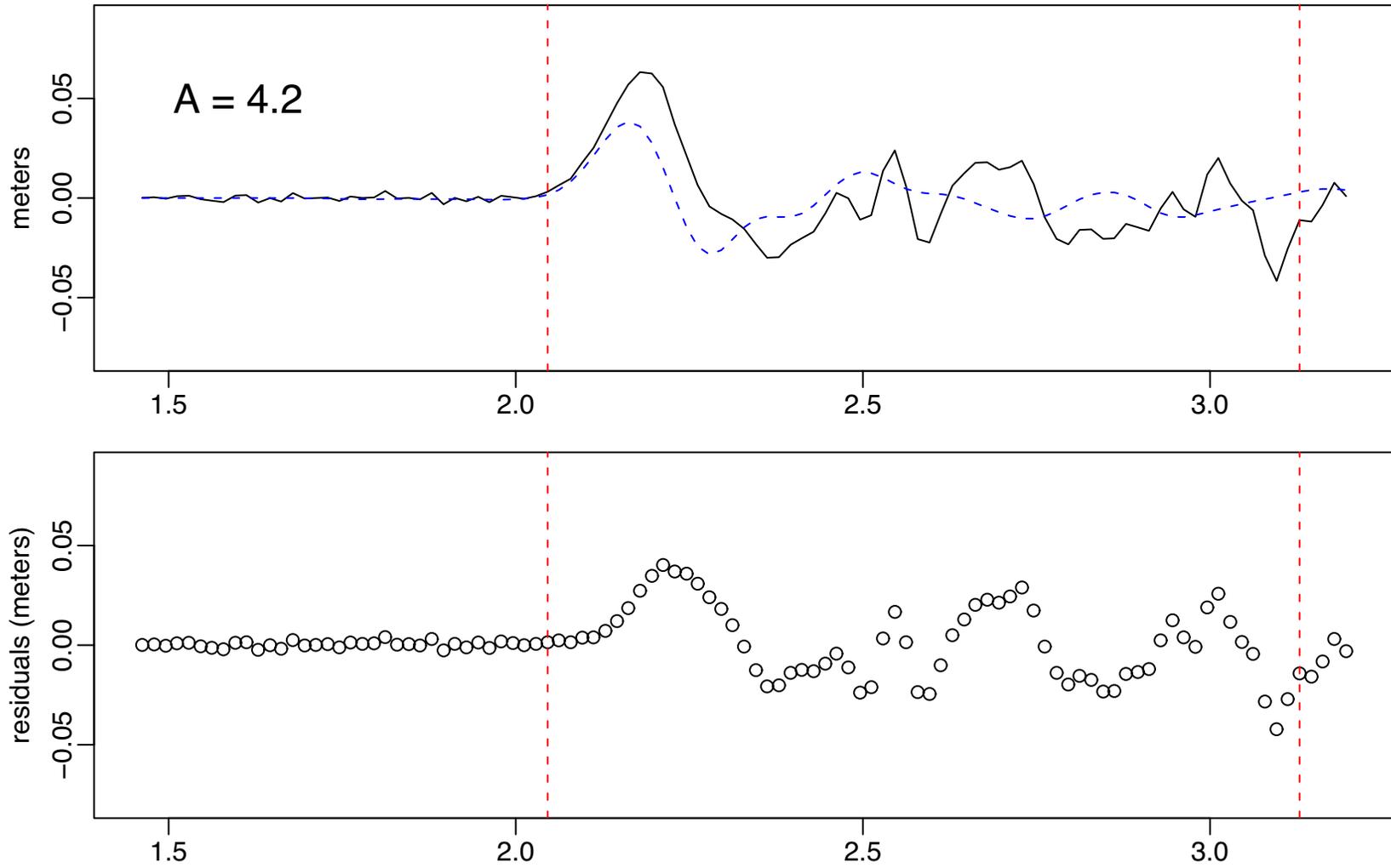
# LS Fit of a12 Model to Selected 21414 Data



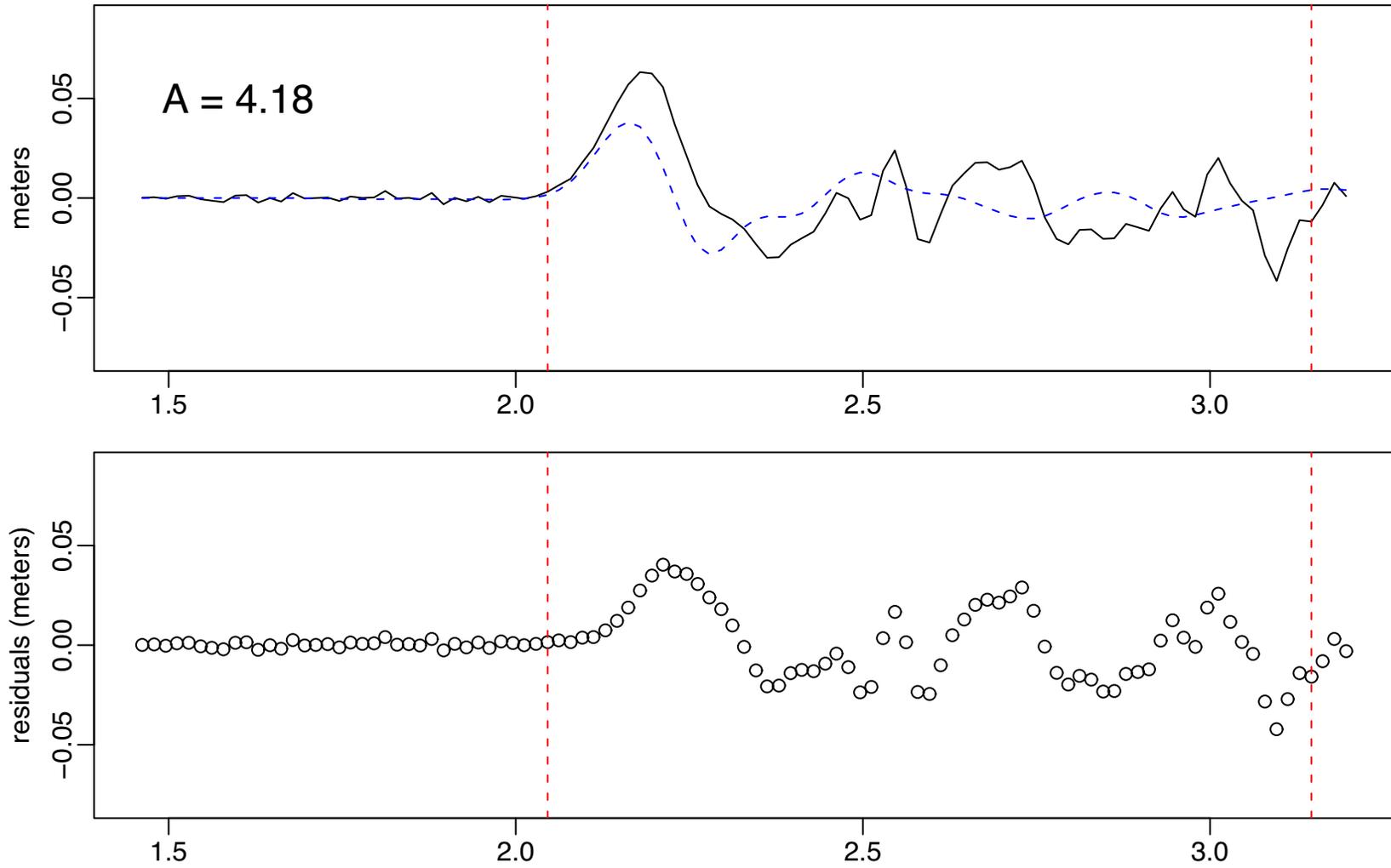
# LS Fit of a12 Model to Selected 21414 Data



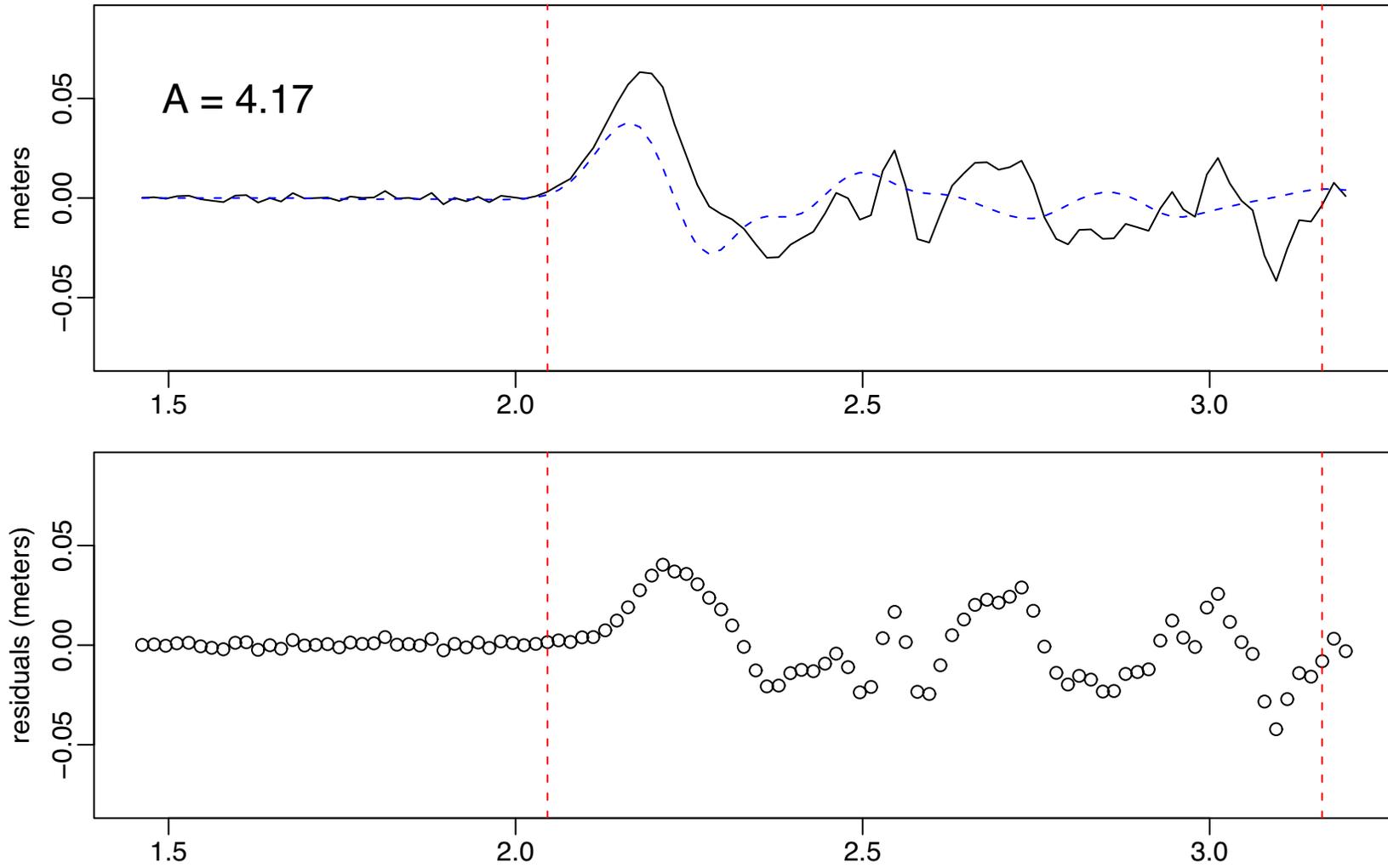
# LS Fit of a12 Model to Selected 21414 Data



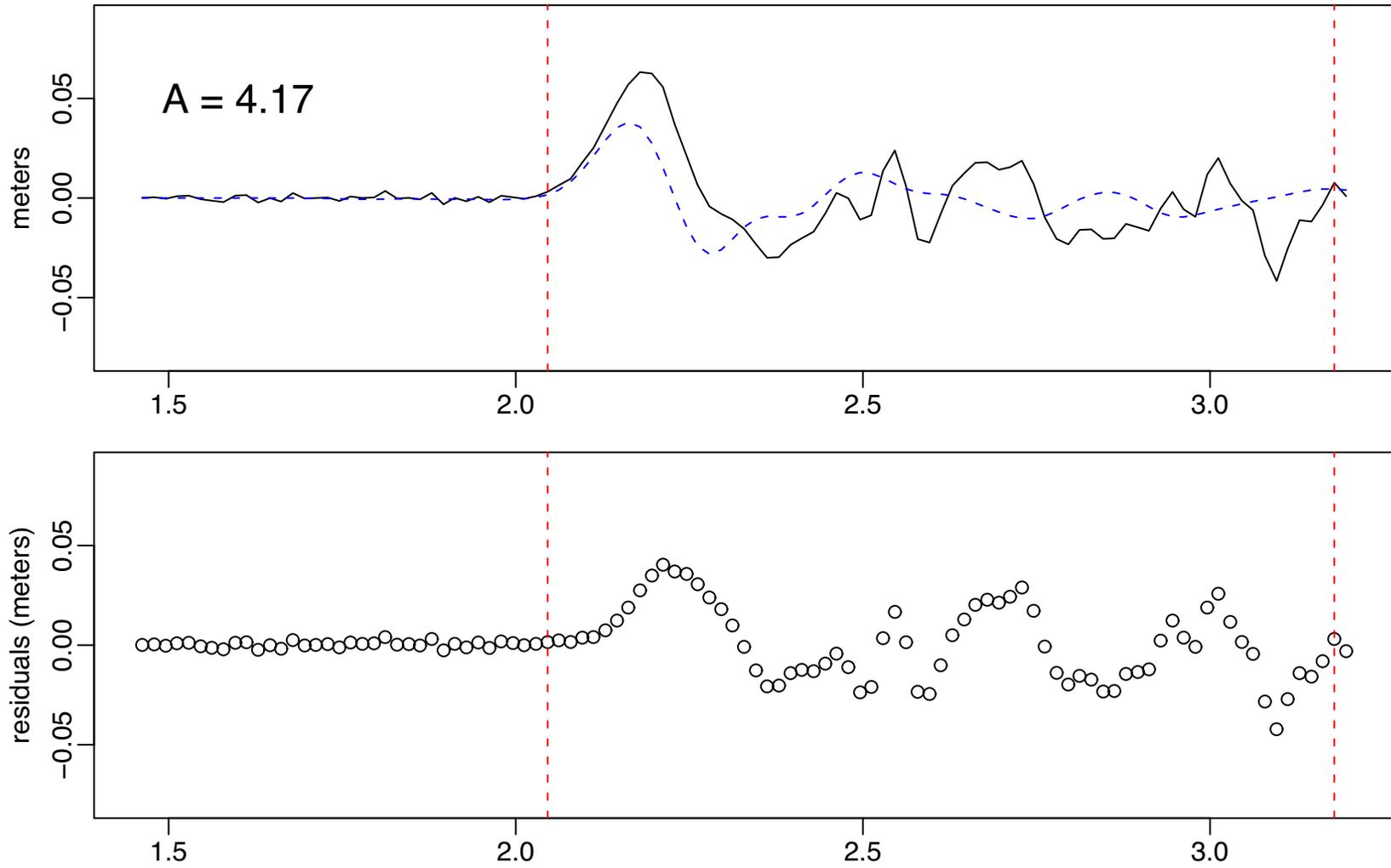
# LS Fit of a12 Model to Selected 21414 Data



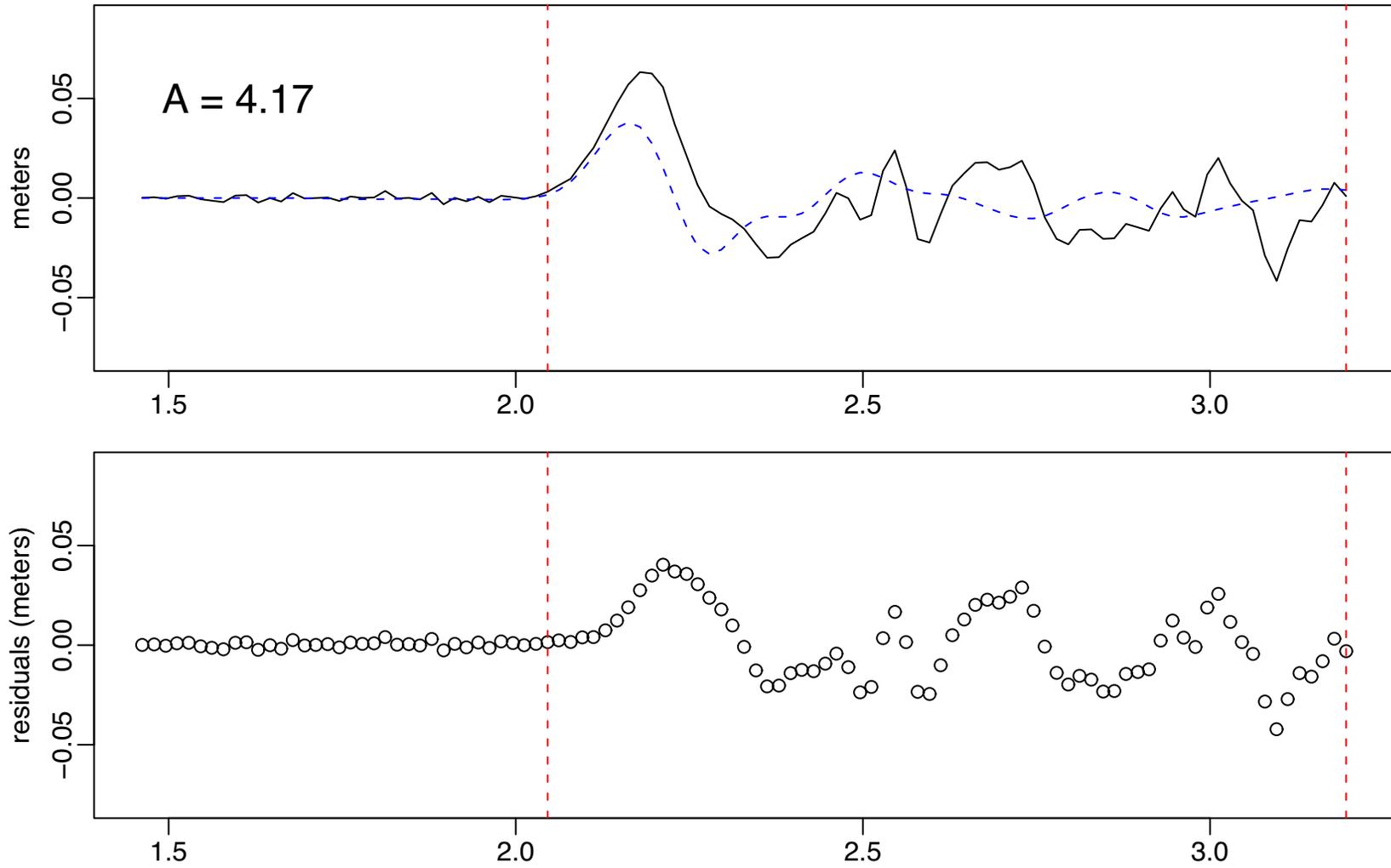
# LS Fit of a12 Model to Selected 21414 Data



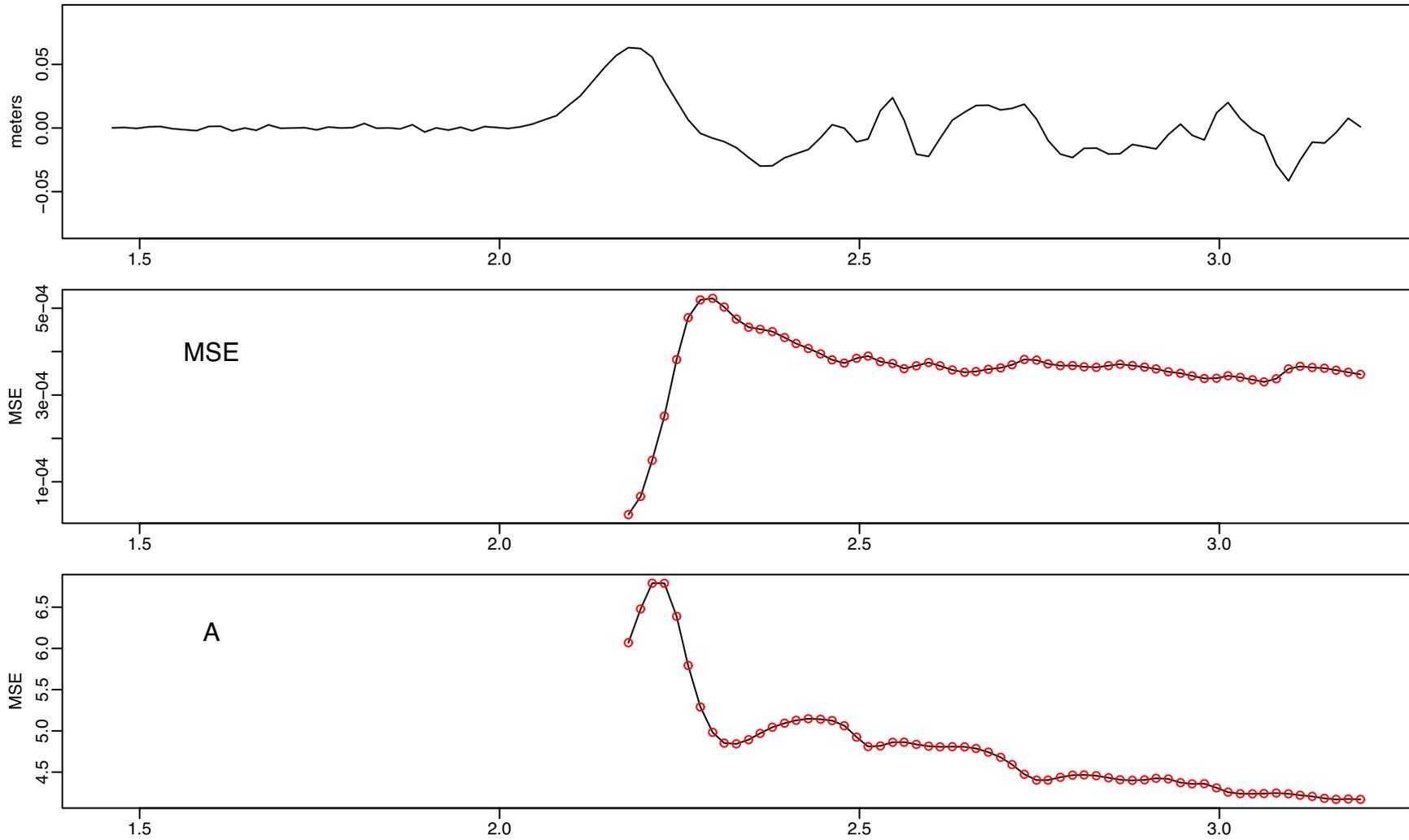
# LS Fit of a12 Model to Selected 21414 Data



# LS Fit of a12 Model to Selected 21414 Data



# Mean Squared Errors and $\hat{A}$ for Selected 21414 Data



## Incorporating Data from a Second Buoy: I

- so far, have modeled data from buoy 21414 in terms of an earthquake coming from source a12
- in vector notation, we have  $\mathbf{x} = A\mathbf{g} + \mathbf{e}$
- in preparation for looking at additional buoys and additional sources, let's rewrite model as  $\mathbf{x}_1 = A_{a12} \cdot \mathbf{g}_{1,a12} + \mathbf{e}_1$ , where
  - $\mathbf{x}_1$  is a vector containing data from first buoy (here 21414)
  - $A_{a12}$  is a scalar representing slip associated with source a12
  - $\mathbf{g}_{1,a12}$  is a vector containing unit slip model for what first buoy should see from earthquake originating at source a12
  - $\mathbf{e}_1$  is a vector of residuals (represents combination of measurements errors and model inaccuracies)

## Incorporating Data from a Second Buoy: II

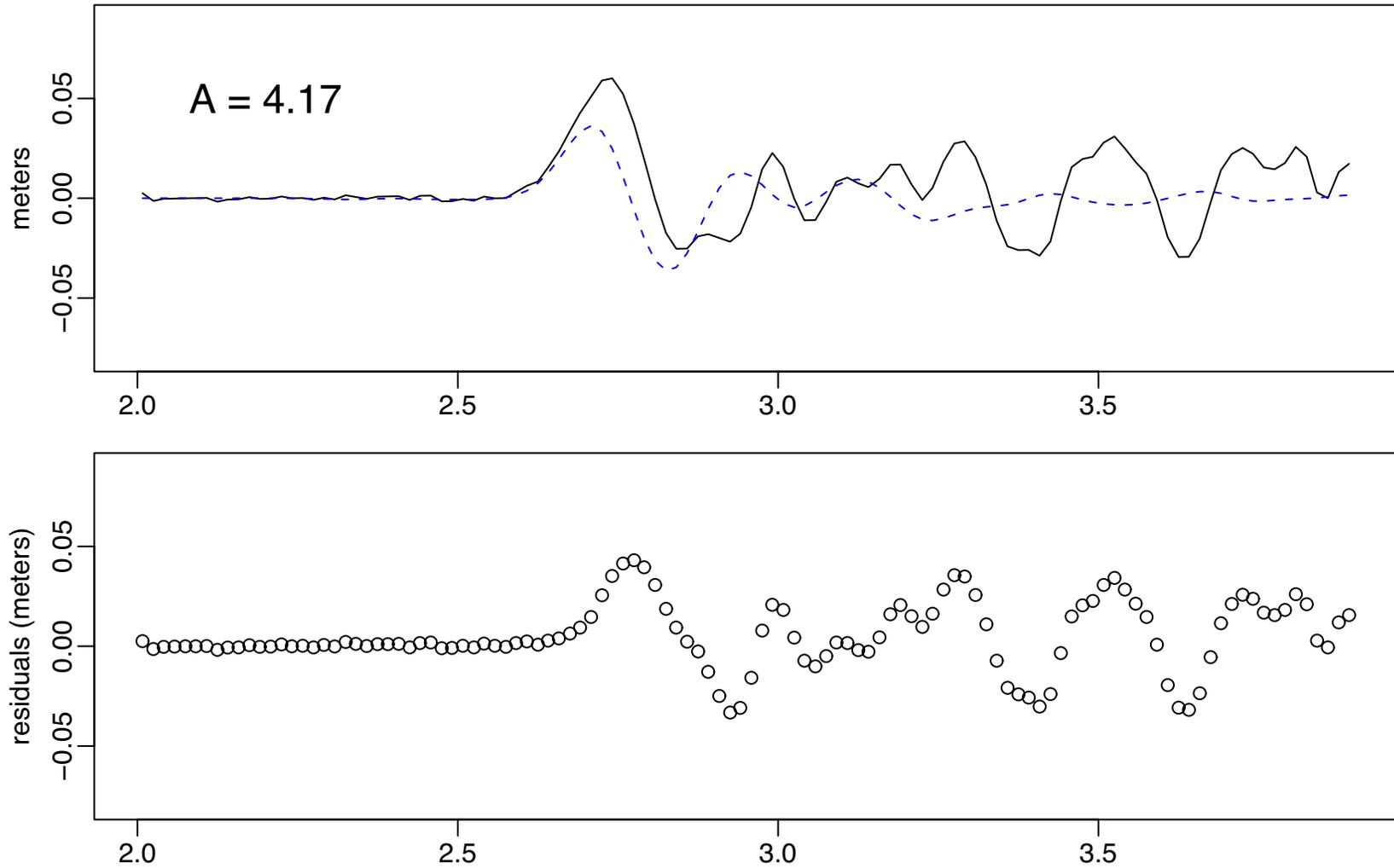
- in this notation, LS estimator of  $A_{a12}$  is given by

$$\hat{A}_{a12} = \mathbf{g}_{1,a12}^T \mathbf{x}_1 / \mathbf{g}_{1,a12}^T \mathbf{g}_{1,a12} = \left( \mathbf{g}_{1,a12}^T \mathbf{g}_{1,a12} \right)^{-1} \mathbf{g}_{1,a12}^T \mathbf{x}_1$$

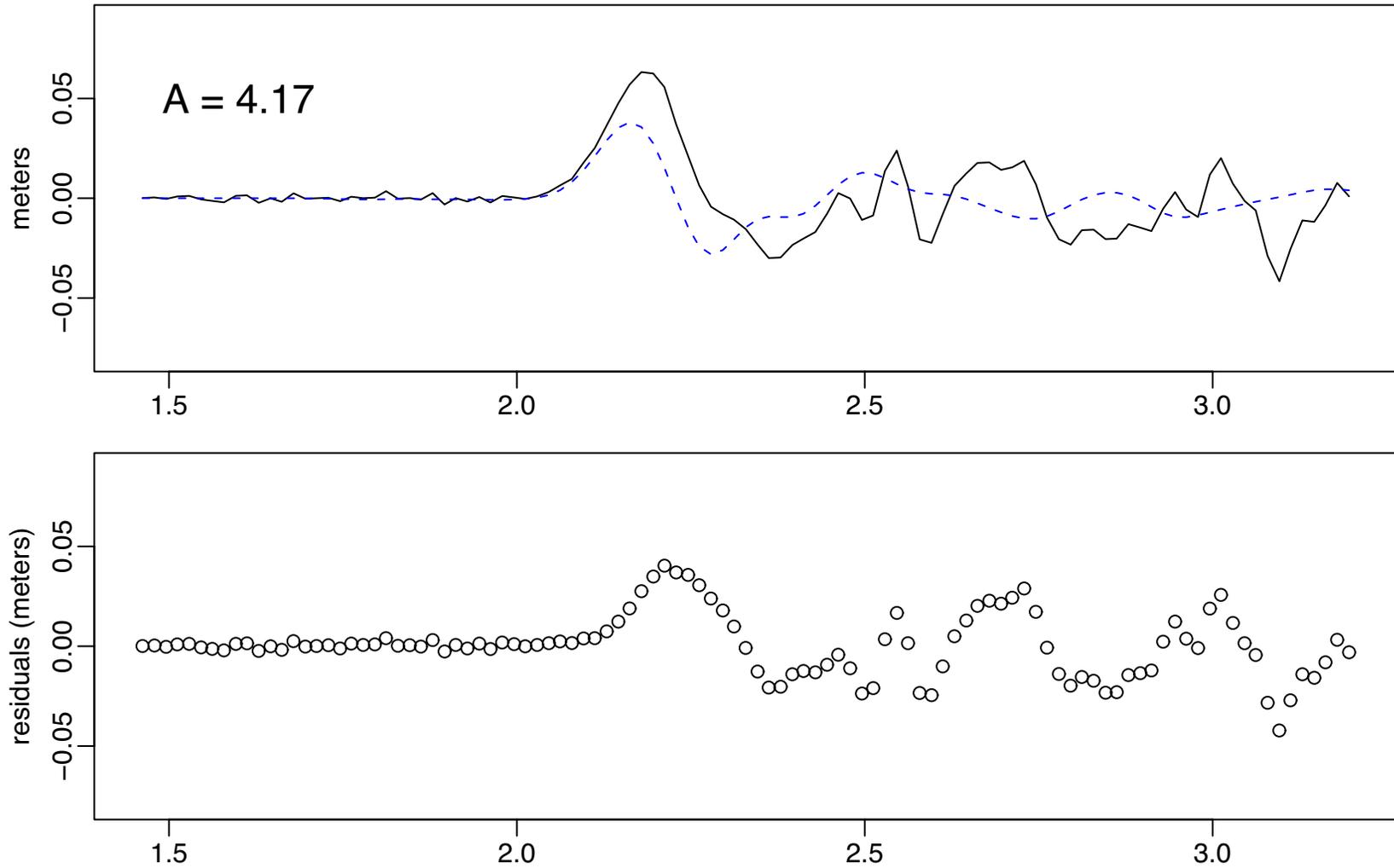
(last expression of interest for generalizations to come)

- now consider data  $\mathbf{x}_2$  from a second buoy (46413)
- model this data as  $\mathbf{x}_2 = A_{a12} \cdot \mathbf{g}_{2,a12} + \mathbf{e}_2$
- note that, while  $\mathbf{g}_{2,a12}$  for buoy 46413 is different from  $\mathbf{g}_{1,a12}$  for buoy 21414, both models have the same slip  $A_{a12}$
- given our estimate  $\hat{A}_{a12} \doteq 4.17$  based upon just  $\mathbf{x}_1$ , let's see how well  $\mathbf{x}_2$  and  $\hat{A}_{a12} \cdot \mathbf{g}_{2,a12}$  match up ('cross-validation')

# a12 Slip from 21414 Data Applied to 46413 Data



# a12 Slip from 21414 Data Applied to 21414 Data



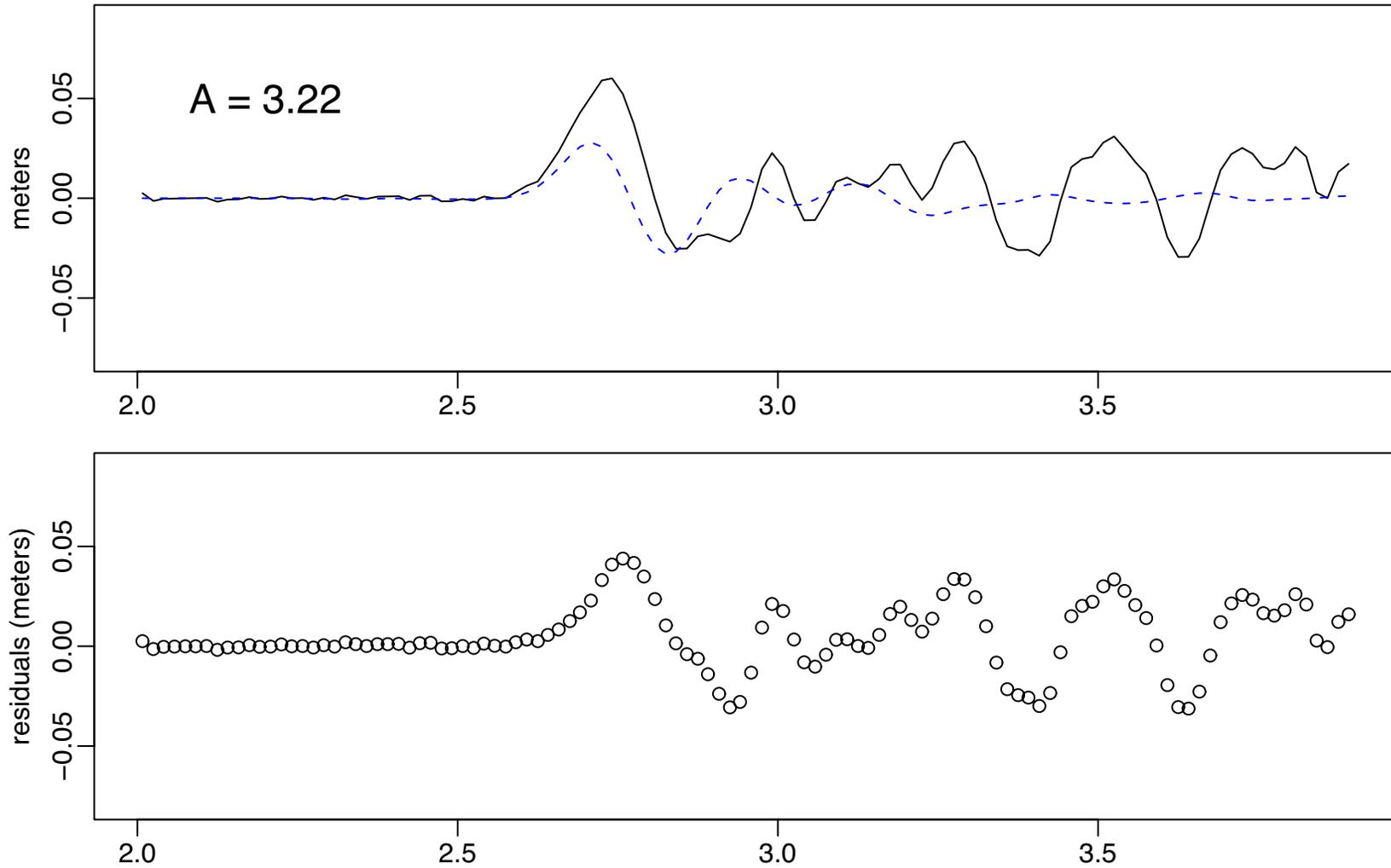
## Incorporating Data from a Second Buoy: III

- can also estimate  $A_{a12}$  using just data from 46413:

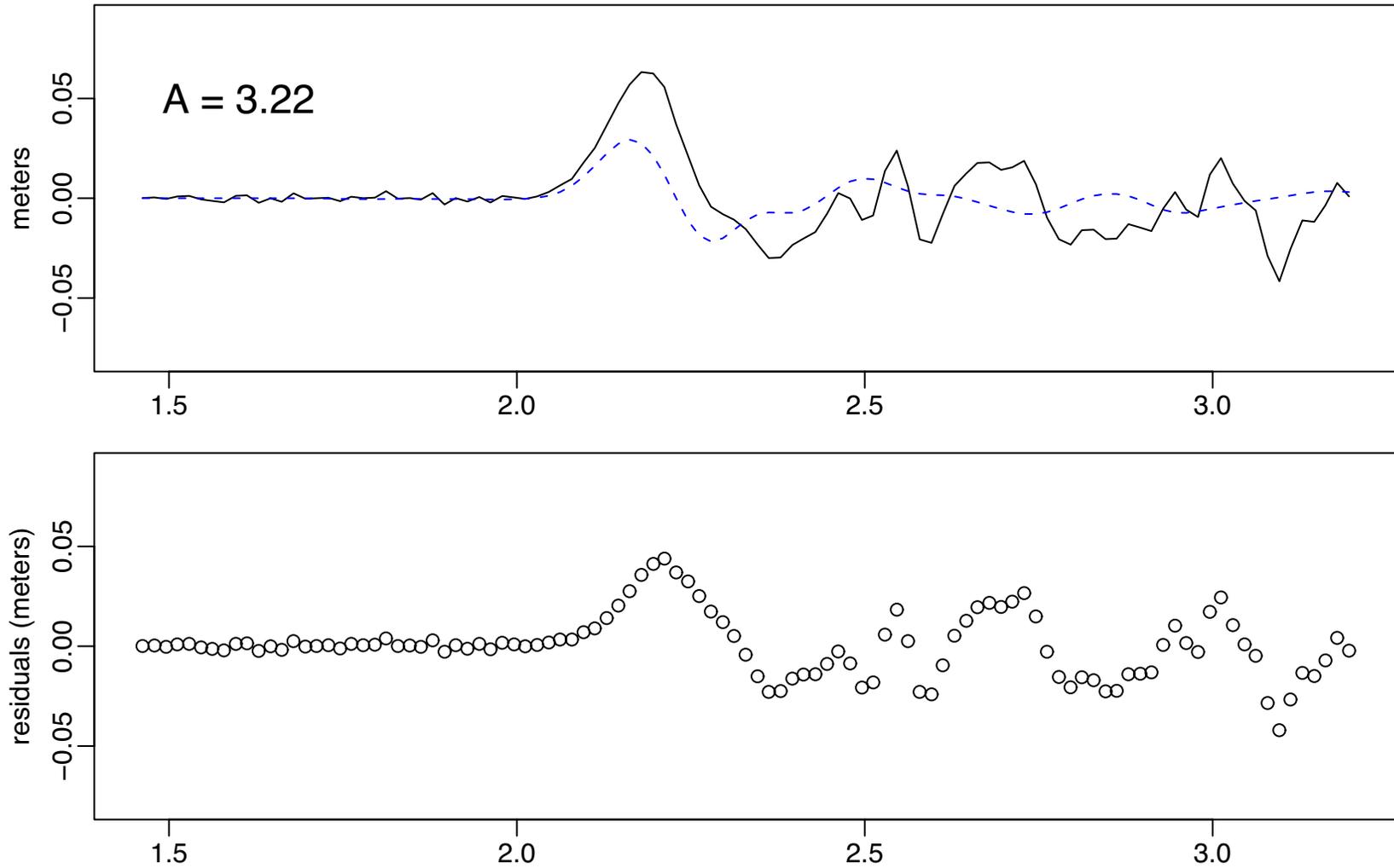
$$\hat{A}_{a12} = \left( \mathbf{g}_{2,a12}^T \mathbf{g}_{2,a12} \right)^{-1} \mathbf{g}_{2,a12}^T \mathbf{x}_2$$

- yields  $\hat{A}_{a12} \doteq 3.22$ , whereas we had  $\hat{A}_{a12} \doteq 4.17$  from 21414
- can look at plots corresponding to the ones we had before

# a12 Slip from 46413 Data Applied to 46413 Data



# a12 Slip from 46413 Data Applied to 21414 Data



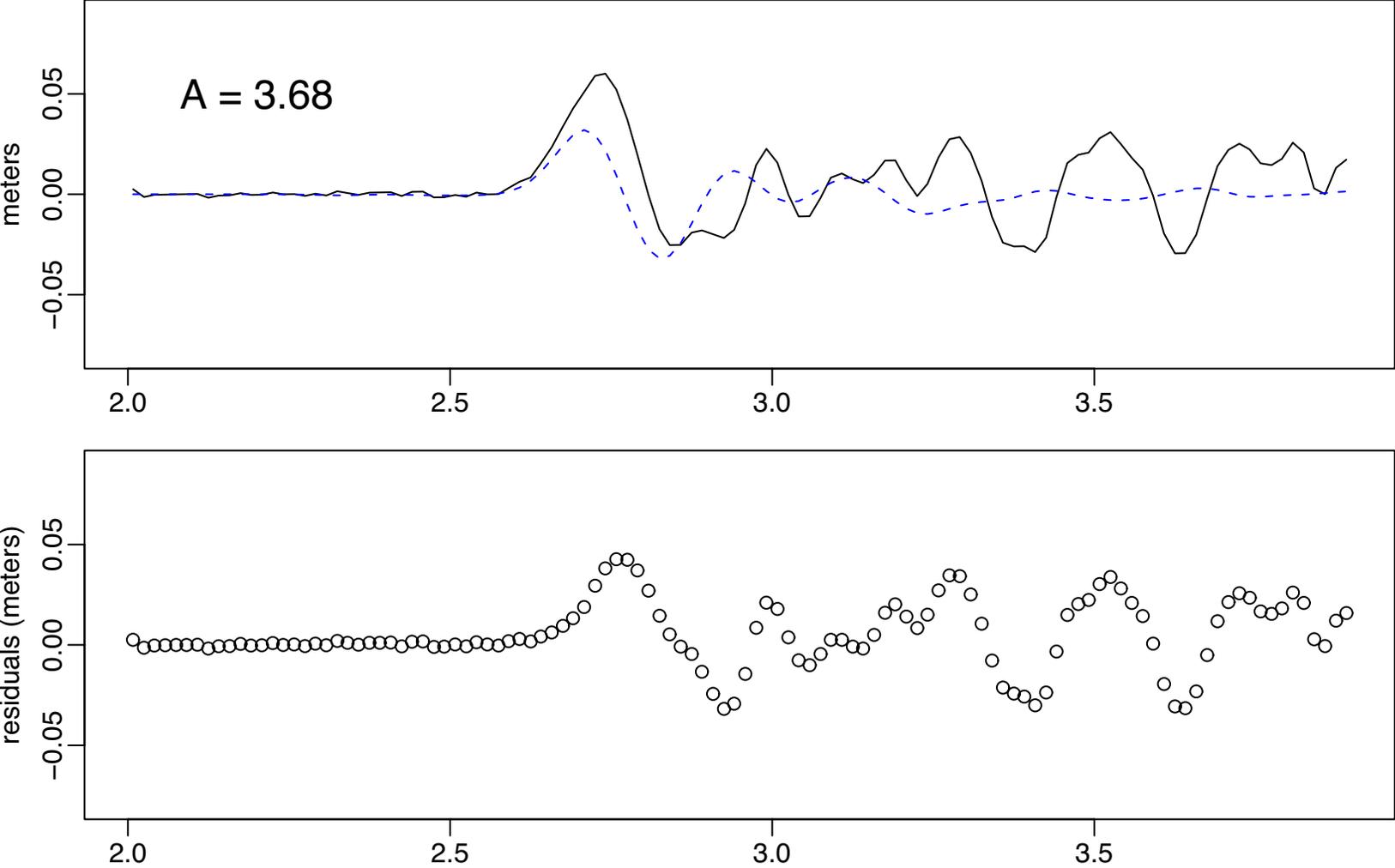
## Incorporating Data from a Second Buoy: IV

- another approach is to use data from both buoys to get a joint estimate for  $A_{a12}$
- joint model is  $\mathbf{x}_{1:2} = A_{a12} \cdot \mathbf{g}_{1:2,a12} + \mathbf{e}_{1:2}$ , where
  - $\mathbf{x}_{1:2}$  is a vector formed by stacking  $\mathbf{x}_1$  on top of  $\mathbf{x}_2$
  - $A_{a12}$  is a scalar representing slip associated with source a12
  - $\mathbf{g}_{1:2,a12}$  is a vector formed by stacking  $\mathbf{g}_{1,a12}$  on top of  $\mathbf{g}_{2,a12}$
  - $\mathbf{e}_{1:2}$  is a vector of residuals
- LS estimator of  $A_{a12}$  now takes the form

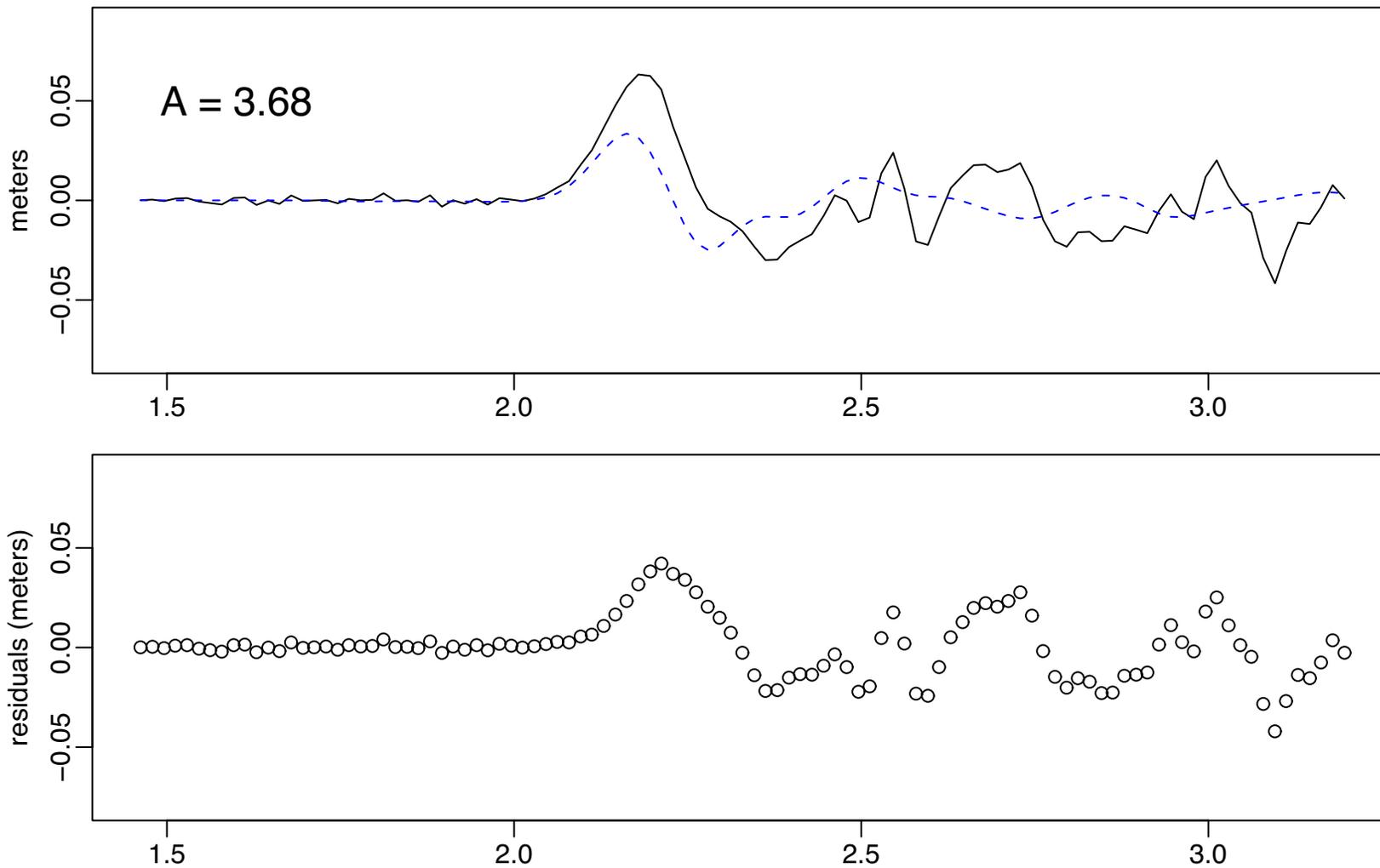
$$\hat{A}_{a12} = \left( \mathbf{g}_{1:2,a12}^T \mathbf{g}_{1:2,a12} \right)^{-1} \mathbf{g}_{1:2,a12}^T \mathbf{x}_{1:2}$$

- yields  $\hat{A}_{a12} \doteq 3.68$  (cf. 4.17 from 21414 and 3.22 from 46413)
- can look at plots corresponding to the ones we had before

# a12 Slip from Both Buoys Applied to 46413 Data



# a12 Slip from Both Buoys Applied to 21414 Data



## Using Linear Combinations of Sources: I

- so far, have modeled data in terms of a single source (a12)
- in vector notation, our model is  $\mathbf{x}_{1:2} = A_{a12} \cdot \mathbf{g}_{1:2,a12} + \mathbf{e}_{1:2}$
- suppose earthquake is actually a linear combination of two sources, namely, a12 and b13
- our model is now  $\mathbf{x}_{1:2} = A_{a12} \cdot \mathbf{g}_{1:2,a12} + A_{b13} \cdot \mathbf{g}_{1:2,b13} + \mathbf{e}_{1:2}$
- can reexpress this model as  $\mathbf{x}_{1:2} = G\mathbf{A} + \mathbf{e}_{1:2}$ , where
  - $G$  is a matrix with two columns, namely,  $\mathbf{g}_{1:2,a12}$  and  $\mathbf{g}_{1:2,b13}$
  - $\mathbf{A}$  is a vector with two elements, namely,  $A_{a12}$  and  $A_{b13}$
- LS estimator of  $\mathbf{A}$  is given by  $\hat{\mathbf{A}} = (G^T G)^{-1} G^T \mathbf{x}_{1:2}$ , which is similar in form to

$$\hat{A}_{a12} = \left( \mathbf{g}_{1:2,a12}^T \mathbf{g}_{1:2,a12} \right)^{-1} \mathbf{g}_{1:2,a12}^T \mathbf{x}_{1:2}$$

## Using Linear Combinations of Sources: II

- complication: models from two sources can be very similar!
- in worse case scenario, have  $\mathbf{g}_{1:2,b13} = \alpha \mathbf{g}_{1:2,a12} \equiv \alpha \mathbf{g}$
- in this case,  $G = [\mathbf{g}, \alpha \mathbf{g}]$  and

$$G^T G = \begin{bmatrix} \mathbf{g}^T \mathbf{g} & \alpha \mathbf{g}^T \mathbf{g} \\ \alpha \mathbf{g}^T \mathbf{g} & \alpha^2 \mathbf{g}^T \mathbf{g} \end{bmatrix},$$

which has a determinant of zero, so  $(G^T G)^{-1}$  does not exist

- instead of using  $\hat{\mathbf{A}} = (G^T G)^{-1} G^T \mathbf{x}_{1:2}$ , can handle this case by solving equation  $G^T G \hat{\mathbf{A}} = G^T \mathbf{x}_{1:2}$  with help of a singular value decomposition (SVD) of the matrix  $G^T G$
- in general, use of SVD yields protection against problems of numerical stability in computing  $\hat{\mathbf{A}}$

## Using Linear Combinations of Sources: III

- using sources a12 and b13 to model data from buoys 21414 and 46413, LS estimates of slips are

$$\hat{\mathbf{A}} \equiv [\hat{A}_{a12}, \hat{A}_{b13}]^T \doteq [2.61, 3.81]^T$$

- fitted model and residuals for 21414 are given by

$$\mathbf{f}_1 \equiv \hat{A}_{a12} \cdot \mathbf{g}_{1,a12} + \hat{A}_{b13} \cdot \mathbf{g}_{1,b13} \quad \text{and} \quad \mathbf{e}_1 = \mathbf{x}_1 - \mathbf{f}_1$$

likewise, fitted model and residuals for 46413 are given by

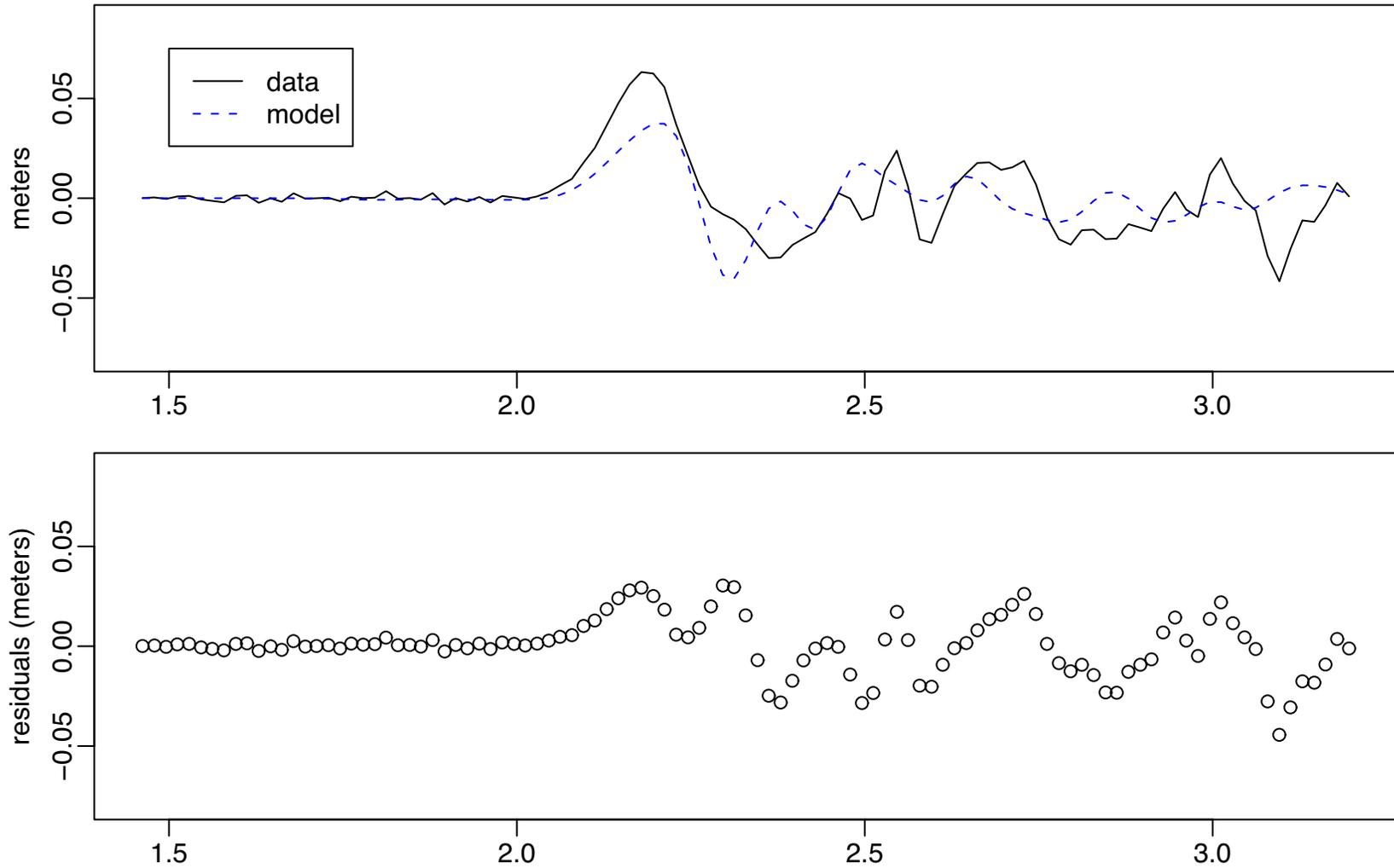
$$\mathbf{f}_2 \equiv \hat{A}_{a12} \cdot \mathbf{g}_{2,a12} + \hat{A}_{b13} \cdot \mathbf{g}_{2,b13} \quad \text{and} \quad \mathbf{e}_2 = \mathbf{x}_2 - \mathbf{f}_2$$

- can use this model to predict what a third buoy should see:

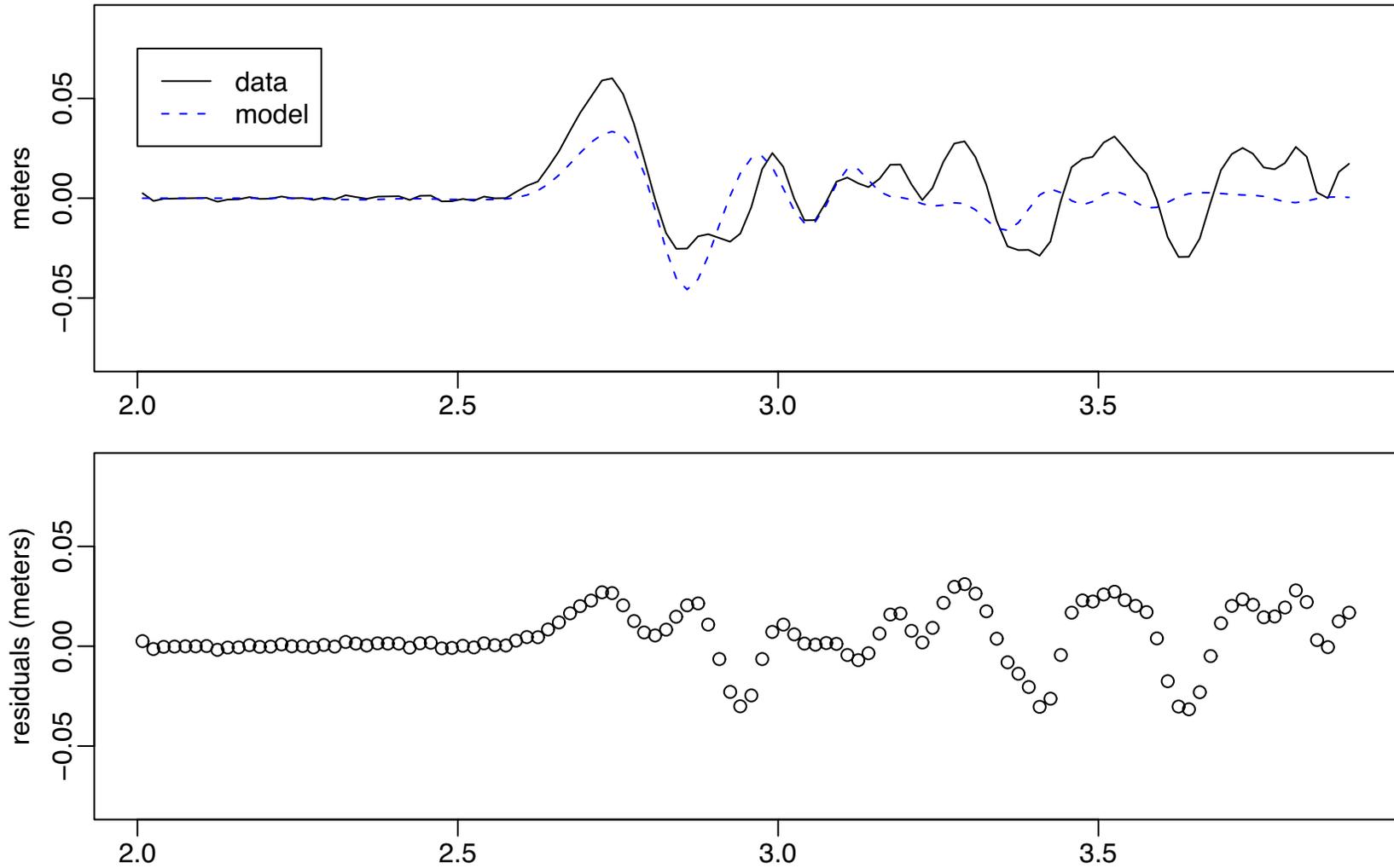
$$\mathbf{f}_3 \equiv \hat{A}_{a12} \cdot \mathbf{g}_{3,a12} + \hat{A}_{b13} \cdot \mathbf{g}_{3,b13}$$

(‘cross-validation’)

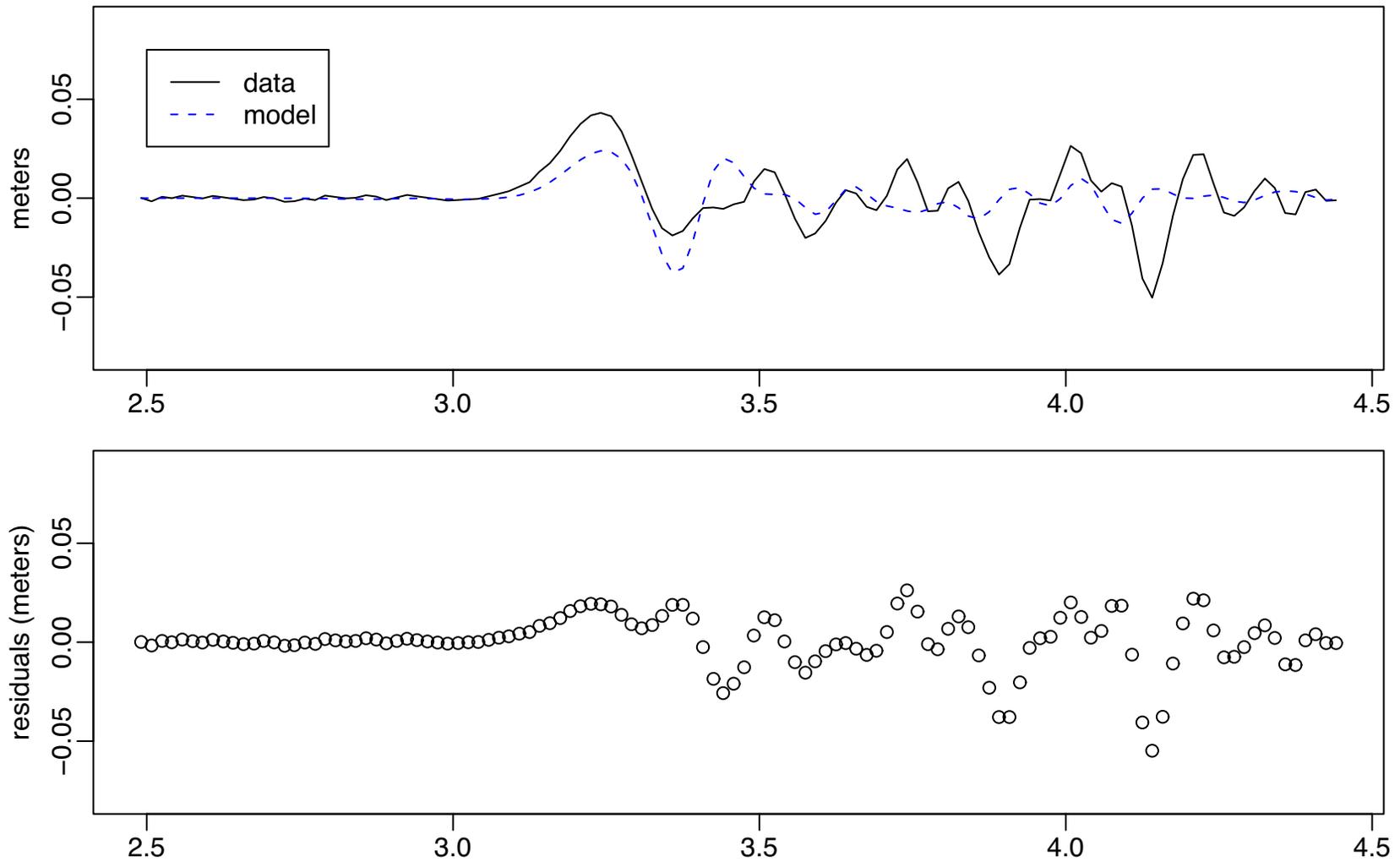
# Data, Fitted Model and Residuals for 21414



# Data, Fitted Model and Residuals for 46413



# Data, Fitted Model and Residuals for 46408



## Bells & Whistles

- current implementation of inversion algorithm allows for
  - constraints on slips (either  $A \geq 0$  or  $A \leq 0$ )
  - shifting of source models, i.e., use of

$$\tilde{g}(t) = g(t - a),$$

where  $a$  is a shift that can be constrained to interval  $[a_l, a_u]$

- stretching/shrinking of source models, i.e., use of

$$\tilde{g}(t) = g(t/b),$$

where  $b$  is a stretch/shrink factor that can be constrained to interval  $[b_l, b_u]$

- shifting and stretching/shrinking together, i.e., use of

$$\tilde{g}(t) = g([t - a]/b)$$

with constraints on both  $a$  and  $b$

## Future Directions

- inversion algorithm requires choice of sources as part of input
- seismic information might suggest, say, eight sources
- currently user can do a joint fit and then manually select sources
- want to investigate use of statistical tests to select sources
- two approaches: step-up and step-down
- step-up approach starts with one source and uses statistical tests to add other sources one at a time
- step-down approach starts with, say, eight sources and uses statistical tests to remove sources one at a time
- idea is that these approaches might provide guidance on source selection for users

## Demo of R Implementation

- R is an interpretive statistical language freely available from

`http://www.r-project.org/`

under the General Public License (GPL)

- R is popular in the statistical community for
  - testing out new ideas in statistics,
  - performing statistical analysis and
  - creating graphics
- inversion algorithm has been bread-boarded in R