Real-time Modelling of Tsunami Data

Applied Physics Laboratory
Department of Statistics
University of Washington
Seattle, Washington, USA

http://faculty.washington.edu/dbp

collaborative effort with NOAA Center for Tsunami Research

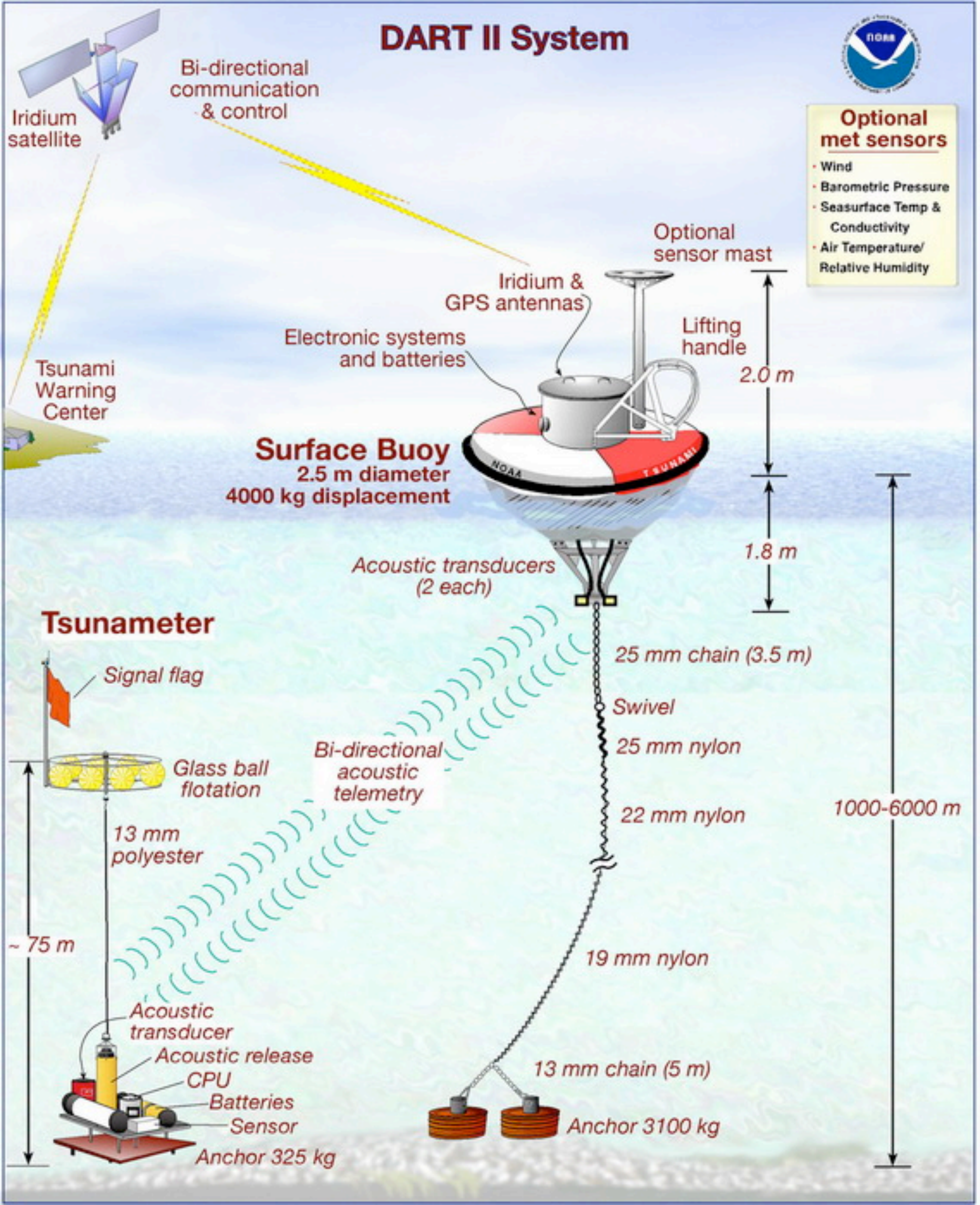
Background - I

- even before disasterous Sumatra tsunami in December 2004, destructive potential of earthquake-generated tsunamis was well-known
- due to rate at which a tsunami advances across the ocean, possible to lessen its effect on some coastal communities through advance warnings
- in USA, warning centers in Alaska and Hawaii are responsible for issuing timely bulletins about impending tsunamis
- seismometers give first indication of a potential tsunami event
  - starting time of earthquake event
  - location of epicenter
  - initial estimate of magnitude

Background - II

- seismic information alone is not enough to accurately forecast potential impact of tsnumai
- led to development of Deep-ocean Assessment and Reporting of Tsunamis (DART®) buoys

RONALD H BROWN
TSUNAMI


DART II System
Iridium satellite
Bi-directional communication & control
Optional met sensors
Wind
Barometric Pressure
Seasurface Temp & Conductivity
Air Temperature/ Relative Humidity
Optional sensor mast
Iridium & GPS antennas
Lifting handle
Electronic systems and batteries
2.0 m
Tsunami Warning Center
Surface Buoy
2.5 m diameter
4000 kg displacement
NOAA
TSUNAMI
1.8 m
Acoustic transducers (2 each)
Tsunameter
25 mm chain (3.5 m)
Signal flag
Swivel
25 mm nylon
Bi-directional acoustic telemetry
Glass ball flotation
22 mm nylon
1000-6000 m
13 mm polyester
~ 75 m
19 mm nylon
Acoustic transducer
Acoustic release
CPU
Batteries
Sensor
Anchor 325 kg
13 mm chain (5 m)
Anchor 3100 kg

Background - III

- prior to December 2004, there were six DART® buoys deployed in Pacific
- since then, many more have been added
- mindful of Percival's First Law

‘Real-time computer demonstrations are doomed to fail!’

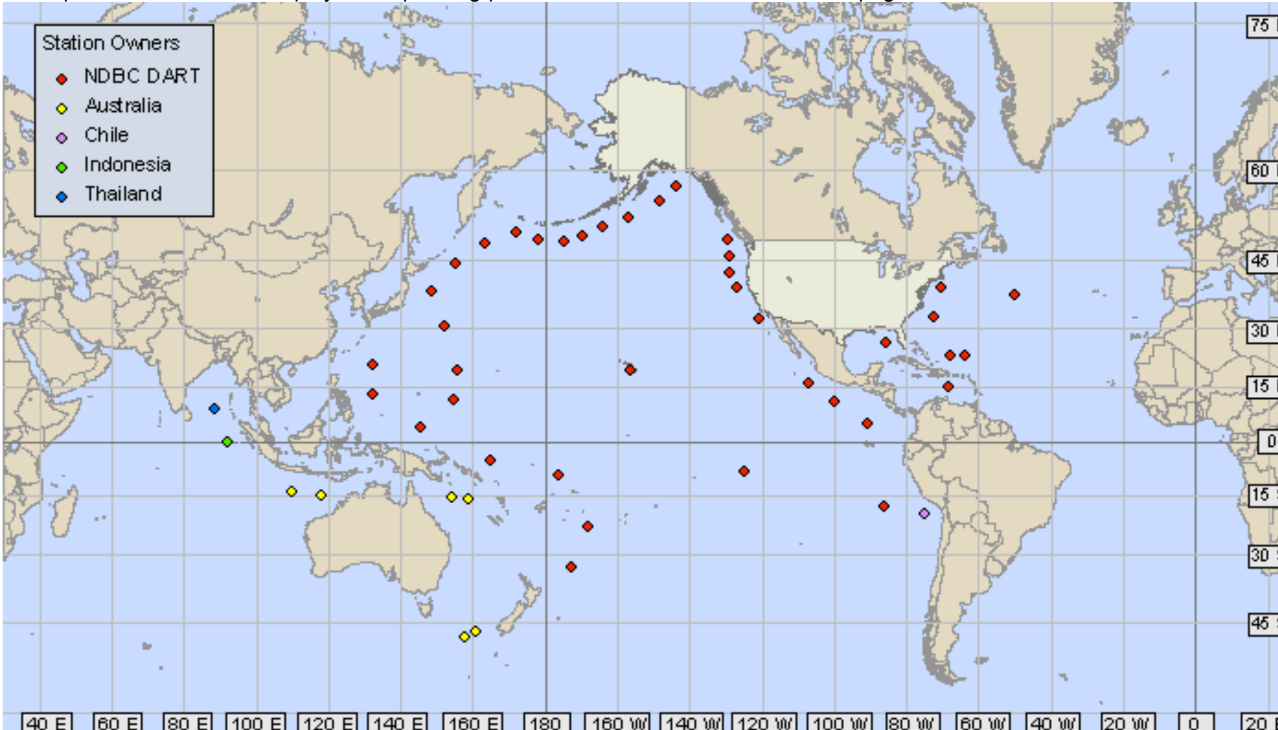
let's look at network of buoys and recent data by going to

`http://www.ndbc.noaa.gov/dart.shtml`

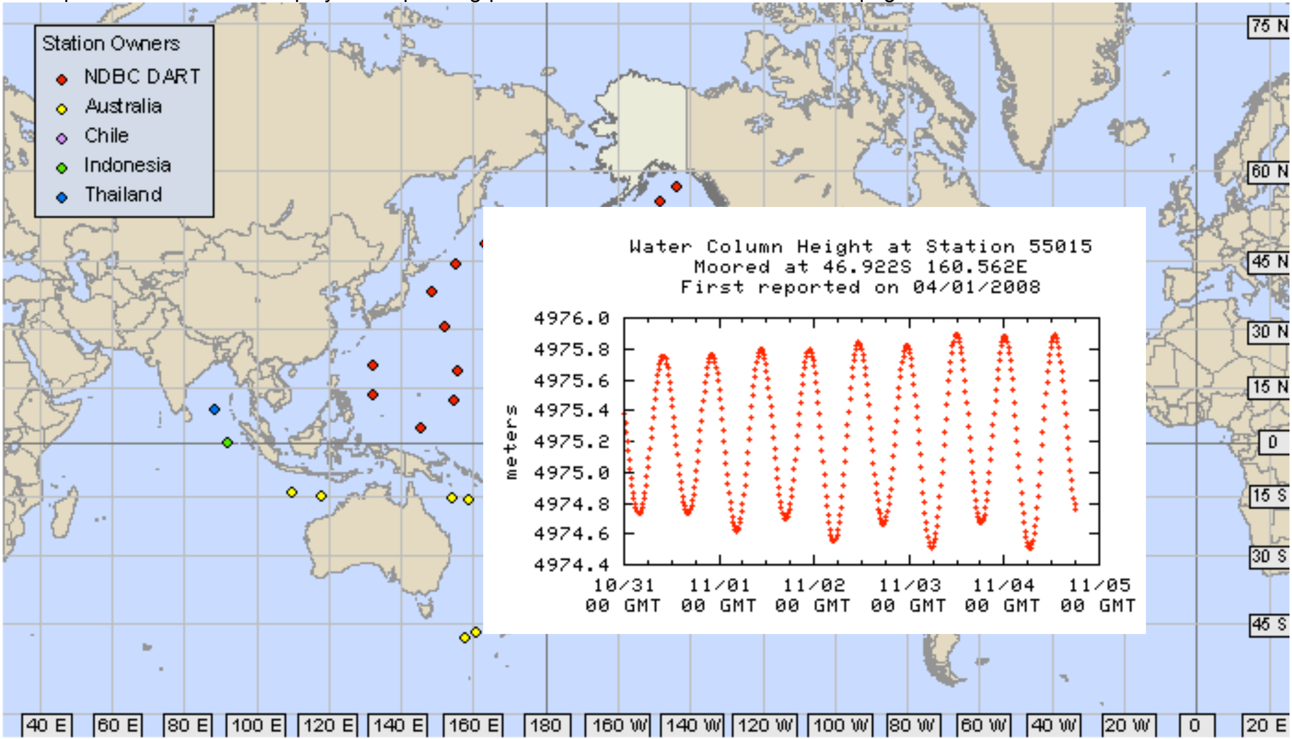
and movie of Nov. 2006 Kuril Island event downloadable from

`http://nctr.pmel.noaa.gov/kuril20061115.html`

Place pointer on station to display corresponding plot or click on station to view station page.



Place pointer on station to display corresponding plot or click on station to view station page.



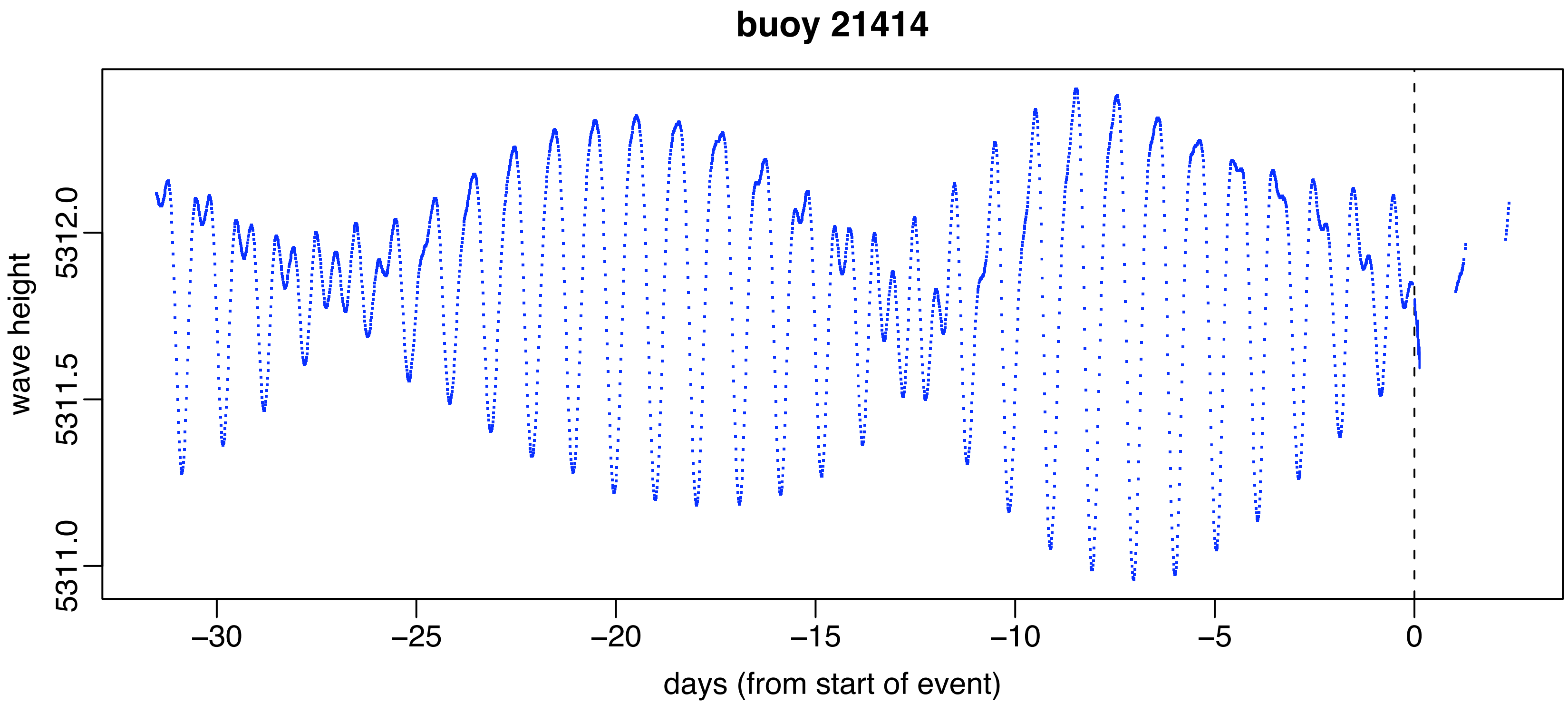
Format of Data

- buoy records one pressure measurement every 15 sec (represents an average over 15 sec)
- 15-min stream: when nothing is going on, buoy transmits one 15-sec measurement every 15 min
- 15-sec stream: when triggered by a seismic event, transmits all recorded measurements
- 1-min stream: during most of tsunami event, transmits averages of four consecutive 15-sec measurements

Short-term Inundation Forecast for Tsunamis (SIFT)

- SIFT is a computer application that uses DART® data and precomputed geophysically-based predictions to assess magnitude of tsunami while tsunami event is evolving
- SIFT currently used at warning centers in Alaska and Hawaii, but is still under development at NOAA Center for Tsunami Research
- statistical issues involved in processing data within SIFT
  - signal extraction: need to remove tidal component before precomputed predictions can be used
  - assessment of uncertainty: need error bars on estimates of magnitude of tsunami
  - variable selection: help operators select predictors

Buoy 21414 Data for Nov. 2006 Kuril Islands Event



Approaches for Signal Extraction (Detiding)

- three approaches, each with certain strengths and weaknesses
  1. empirical orthogonal functions (Tolkova, 2009)
  2. localized polynomial fits
  3. two-stage approach, the second of which involves Kalman filter/smoother
- will concentrate on third approach

Model for Data with Tides

- assume measured data at time t can be expressed as

$$x(t) = \mu(t) + \sum_{l=1}^{L} c_l(t) + \epsilon(t)$$

  where

  - $\mu(t) + \sum_l c_l(t)$ represents tides;
  - $\mu(t)$ is 'unpredictable' tidal component (has slowly varying mean level);
  - $c_l(t) = C_l \cos(2\pi f_l t + \phi_l)$ is lth part of 'predictable' tidal component (sinusoid with amplitude A_l, frequency f_l and phase ϕ_l); and
  - $\epsilon(t)$ is detided data

Model for 15-Min Stream Prior to Tsunami: I

- let $t = 0$ denote starting time (in days) of tsunami events
- can write $C_l \cos(2\pi f_l t + \phi_l) = A_l \cos(2\pi f_l t) + B_l \sin(2\pi f_l t)$, so assume data $x(t)$ at times $t < 0$ can be expressed as

$$x(t) = \mu + \sum_{l=1}^{L} A_l \cos(2\pi f_l t) + B_l \sin(2\pi f_l t) + \epsilon_t$$

  where

  - μ is an unknown overall mean level;
  - f_l is a known tidal frequency;
  - A_l and B_l are unknown amplitudes; and
  - ϵ_t is a residual term

- use linear least squares to estimate μ, A_l and B_l – denote these estimates as $\hat{\mu}$, $\hat{A}_l$ and $\hat{B}_l$

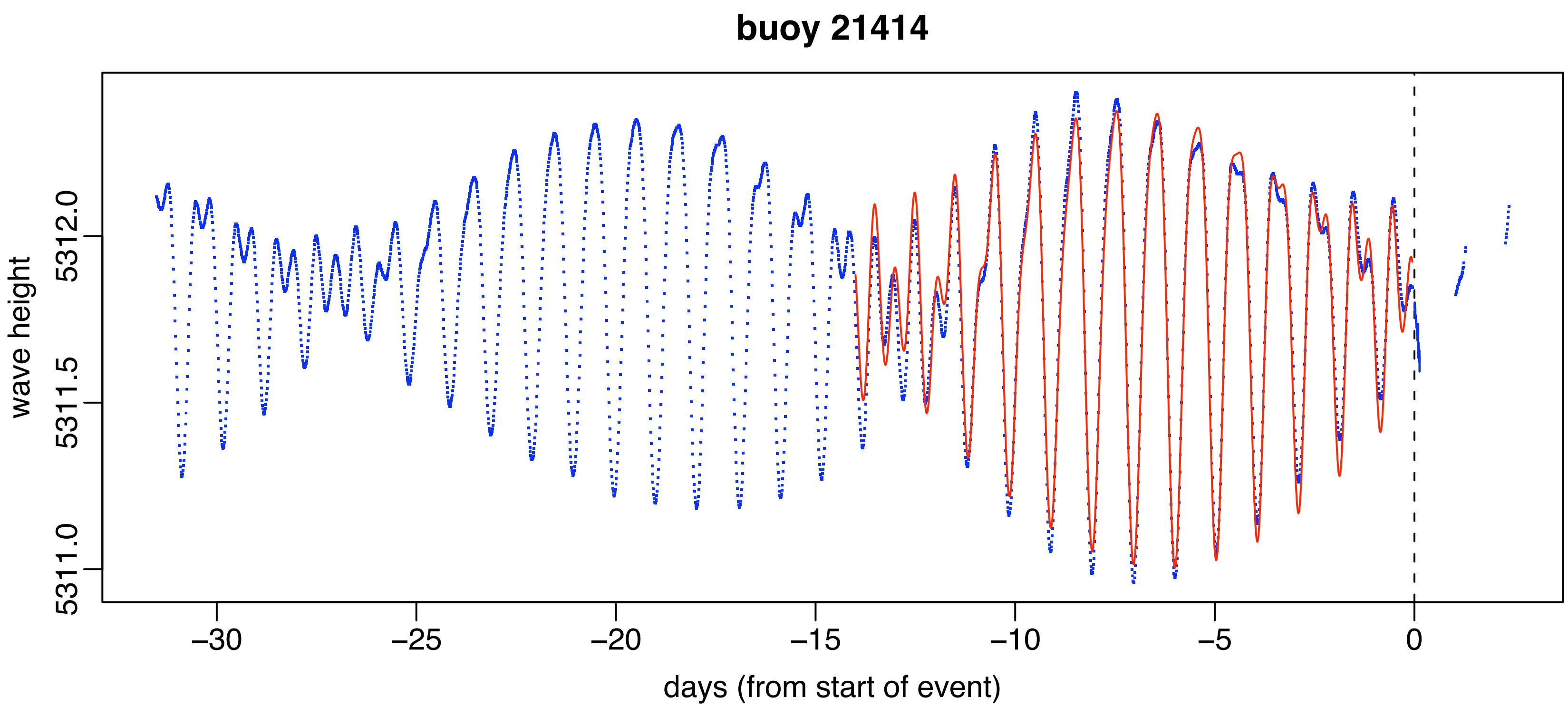
Model for 15-Min Stream Prior to Tsunami: II

- if using 14 days of prior data, let $L = 4$, and set f_1, f_2, f_3 and f_4 to so-called M2, S2, O1 and K1 frequencies
  (other rules apply if use, e.g., 1 day or 29 days of prior data)
- estimate tidal component at time t using

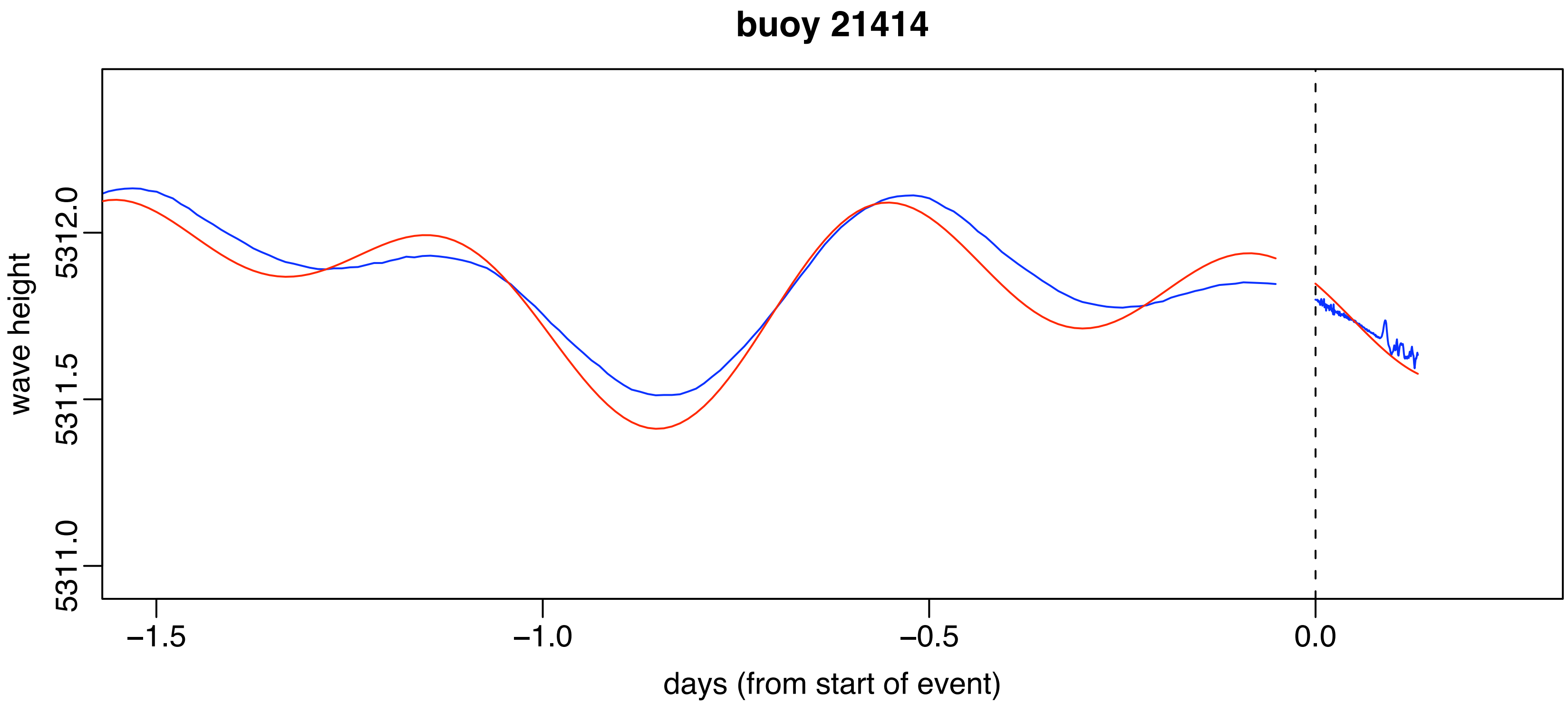
$$\hat{x}(t) = \hat{\mu} + \sum_{l=1}^{L} \hat{A}_l \cos(2\pi f_l t) + \hat{B}_l \sin(2\pi f_l t)$$

- for $t < 0$, can compare estimated tide to actual data
- for $t > 0$, can detide data from 1 min and/or 15 sec streams by subtracting off $\hat{x}(t)$

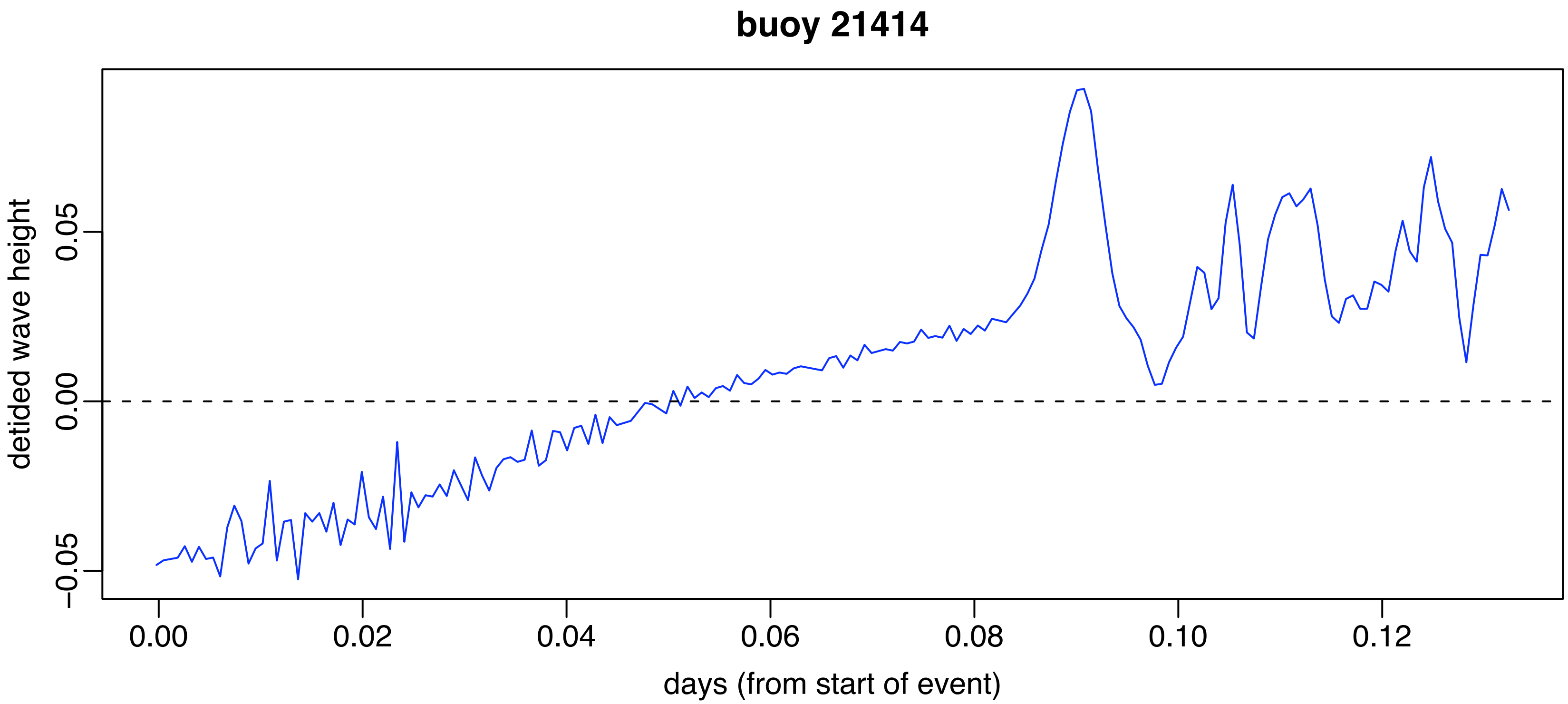
14 Day Model (M2, S2, O1 and K1)



14 Day Model and Extrapolation



Data Detided using 14 Day Model



Second-Stage Detiding via Kalman Filter/Smoother

- after first-stage detiding, can write

$$x(t) - \hat{x}(t) = \mu(t) + \epsilon(t),$$

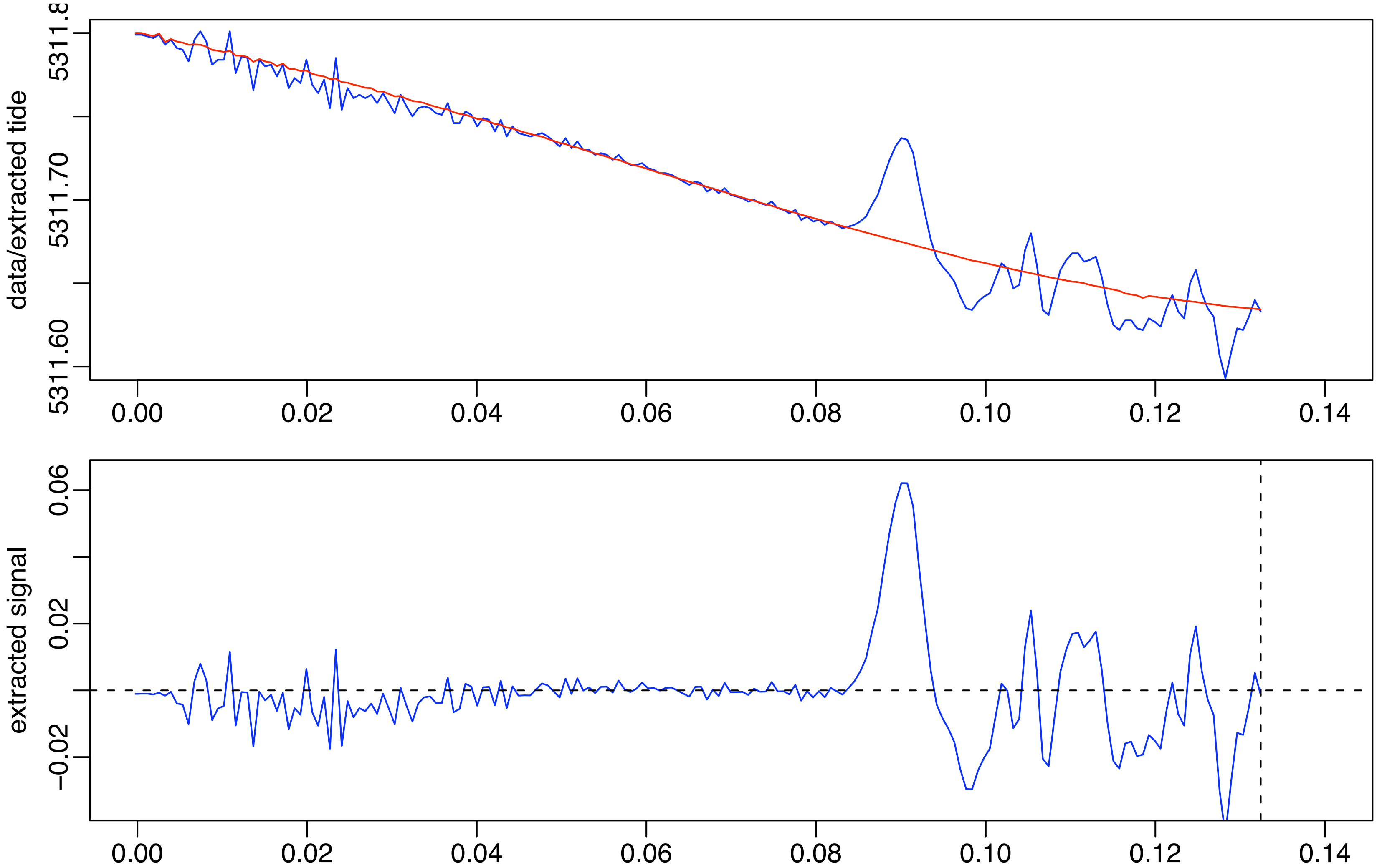
where $\mu(t)$ is 'unpredictable' tidal component

- extract $\mu(t)$ by formulating a state-space model, which allows use of Kalman filter/smoother:

$$\begin{aligned} \mu(t) &= \mu(t-1) + v(t) \\ v(t) &= v(t-1) + \zeta(t) \end{aligned}$$

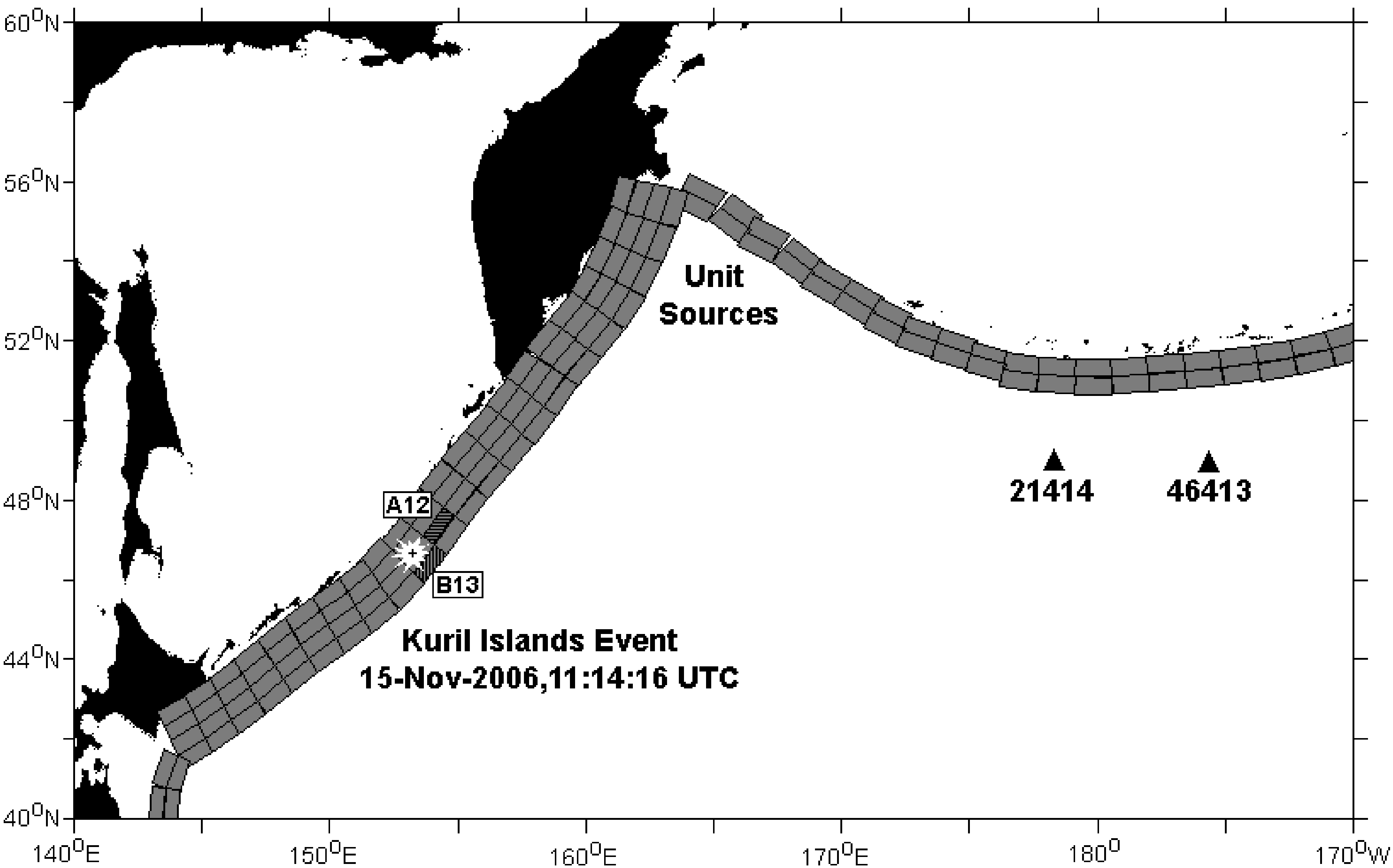
where $\zeta(t)$ is Gaussian white noise with mean zero and variance σ_ζ^2 (can be set using historical data)

Kalman Filtering/Smoothing – 14 Day Model



Modelling of Detided DART® Buoy Data: I

- geophysically-based propagation predictions are based on a simplification provided by so-called 'unit sources'
- unit sources are rectangles of area 100×50 km^2 covering portions of globe from which tsunami-generating earthquakes can occur

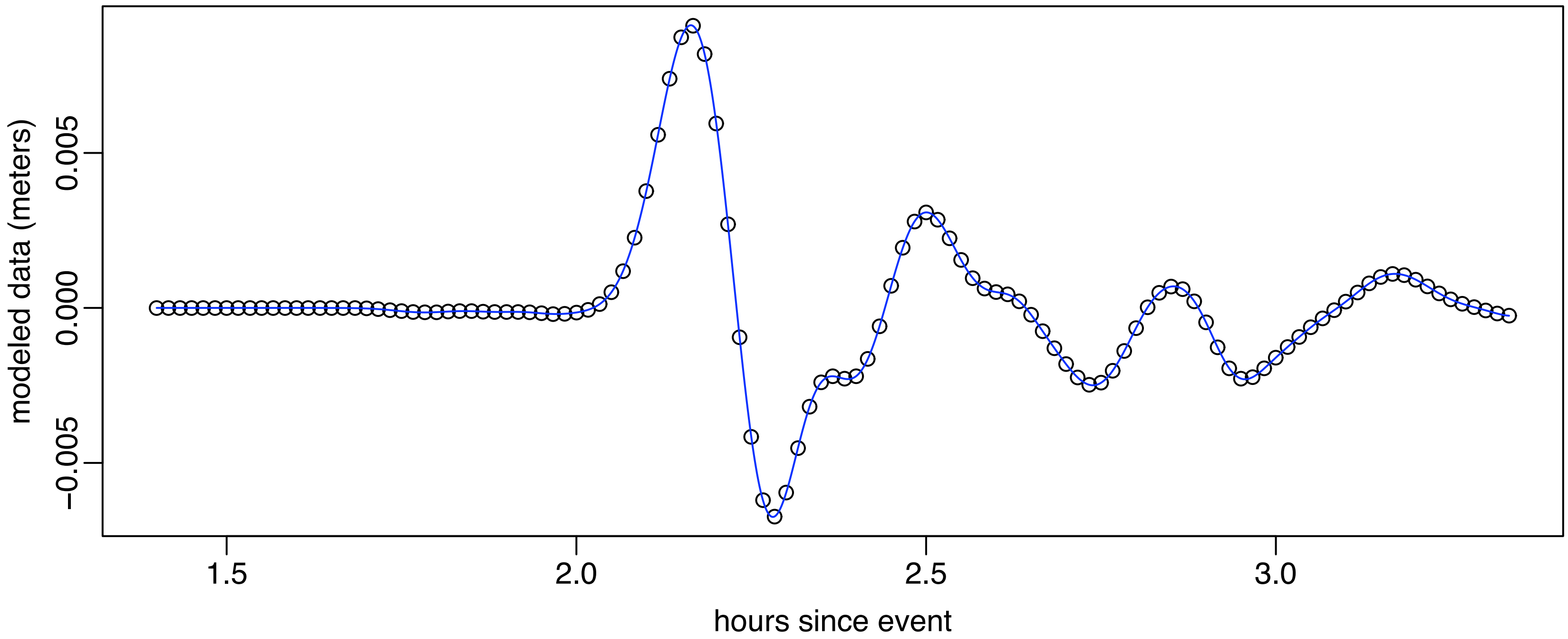


60°N
56°N
52°N
48°N
44°N
40°N
140°E
150°E
160°E
170°E
180°
170°W
Unit Sources
21414
46413
A12
B13
Kuril Islands Event
15-Nov-2006,11:14:16 UTC

Modelling of Detided DART® Buoy Data: II

- propagation predictions must be precomputed (not enough time to compute them once a tsunami is in progress)
- database has been established with precomputed predictions for each pairing of particular unit source and particular buoy
- predictions assume earthquake is located in center of unit source and is of magnitude 7.5
- entry in database predicts what will observed over time at a particular buoy given earthquake from a particular unit source

Model for Buoy 21414 and Unit Source a12



Modelling of Detided DART® Buoy Data: III

- will use buoy data to adjust predictions to handle earthquakes greater or less than 7.5 in magnitude
- consider case of 2 buoys (1, 2) and three unit sources (a, b, c)
- let $\mathbf{x}_1$ and $\mathbf{x}_2$ be relevant detided data from buoys 1 and 2
- let $\mathbf{g}_{1,a}$ etc. be prediction of what buoy 1 should see from earthquake at unit source a
- leads to following model for buoy data:

$$\mathbf{x} \equiv \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{1,a} & \mathbf{g}_{1,b} & \mathbf{g}_{1,c} \\ \mathbf{g}_{2,a} & \mathbf{g}_{2,b} & \mathbf{g}_{2,c} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \end{bmatrix} \equiv G\boldsymbol{\alpha} + \boldsymbol{\epsilon}$$

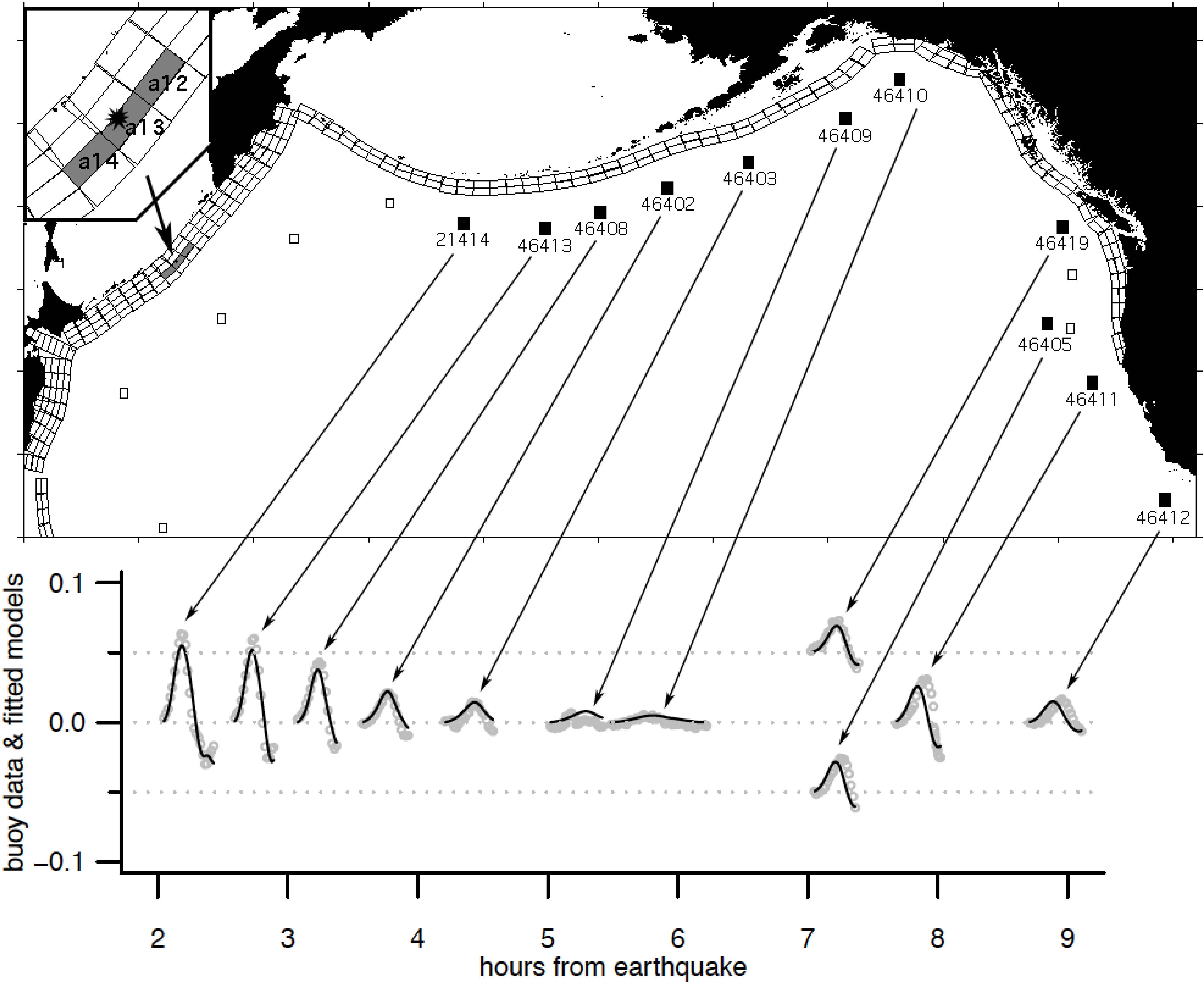
- error terms $\boldsymbol{\epsilon}_j$ assumed to obey first-order autoregressive model

Modelling of Detided DART® Buoy Data: IV

- solve for $\boldsymbol{\alpha}$ using constrained least squares:

$$\text{minimize } \|\boldsymbol{\epsilon}\|^2 = \|\mathbf{x} - G\boldsymbol{\alpha}\|^2 \text{ subject to } \boldsymbol{\alpha} \geq \mathbf{0}$$

- constraint needed to get physically reasonable solution $\hat{\boldsymbol{\alpha}}$
- note: if kth element of $\hat{\boldsymbol{\alpha}}$ is set to zero, kth unit source effectively eliminated from model
- sum of elements of $\hat{\boldsymbol{\alpha}}$ can be used to estimate tsunami magnitude T_M
- uncertainty in T_M assessable using covariance matrix for $\hat{\boldsymbol{\alpha}}$
- following example involves 11 buoys and 3 units sources



a12
a13
a14
21414
46413
46408
46402
46403
46409
46410
46419
46405
46411
46412
buoy data & fitted models
0.1
0.0
−0.1
2
3
4
5
6
7
8
9
hours from earthquake

Concluding Comments

- SIFT application currently relies on experienced operators to select appropriate unit sources
- desirable for future versions of SIFT to include statistically-oriented selection of unit sources (related to variable selection in linear regression)
- selection of data from buoy also currently done manually by operators – need for statistical guidance here also

Thanks to . . .

- Ross Darnell for invitation to speak
- colleagues at NOAA Center for Tsunami Research in Seattle
  - Don Denbo
  - Marie Eblé
  - Edison Gica
  - Hal Mofjeld
  - Mick Spillane
  - Elena Tolkova
  - Vasily Titov
- Crime Scene for 2nd place finish (got $13 in $2 sweep!!!)