A note on spectra …

https://selipot.github.io

selipot.github.io
Google Google Scholar The New York Times NHC Active Tr...cal Cyclones CloudDrift Twitter Latest Posts – Shane Elipot ECMWF | Tropical cyclones Connected Pa

On the inappropriate use of variance-preserving spectra

BY SHANE ELIPOT

JANUARY 20, 2020

2 COMMENTS

I have been surprised by the continuing use of *variance-preserving spectra* in the oceanographic literature as a tool to diagnose periodicities in time series (this does not seem to happen in the purely climate literature … we, oceanographers, are often a bit behind). Since I don't think I know much, I often seek out the help of a couple of statistician friends that deal with time series (like all the time). I have casually asked them: "Why do you think people use variance-preserving spectra?". Their answer was revealing: they did not even know what a variance-preserving spectrum was! Once I explained to them (see below), they were even more confused. They also wondered what was wrong with us (oceanographers). On a more serious note, what concerns me is that we collectively spend millions of tax-payer dollars to obtain amazing time series of oceanographic variables (such as oceanic volume transport), and sometimes analyze them with the wrong tools, unfortunately. We need to collectively do better.

So, I have decided to write a little note to demonstrate why variance-preserving spectra should not be used to make conclusive statements about periodicity in a time series, and to suggest alternative methods. I have recently become a research assistant professor, and thus I am trying to get a bit more in the teaching side of things (I do not teach at my

Photo Credit: Molly Baringer

ROSENSTIEL SCHOOL
COOPERATIVE INSTITUTE for
MARINE & ATMOSPHERIC STUDIES

Deriving uncertainties for the Global Drifter Program hourly product

Case studies

Shane Elipot, July 16, 2024, OceanUQ Summer School

Goal is to go over 3 case studies

of uncertainty derivations and utilizations

- Elipot et al. (2016), *A global surface drifter data set at hourly resolution,* JGR: Oceans, with Rick Lumpkin, Renellys C. Perez, Jonathan M. Lilly, Jeffrey J. Early, and Adam Sykulski doi: 10.1002/2016JC011716

  -> **Examples 1 & 2**

- Elipot et al. (2022), *A dataset of hourly sea surface temperature from drifting buoys,* Scientific Data, with Adam Sykulski, Rick Lumpkin, Luca Centurioni, and Mayra Pazos doi: 10.1038/s41597-022-01670-2

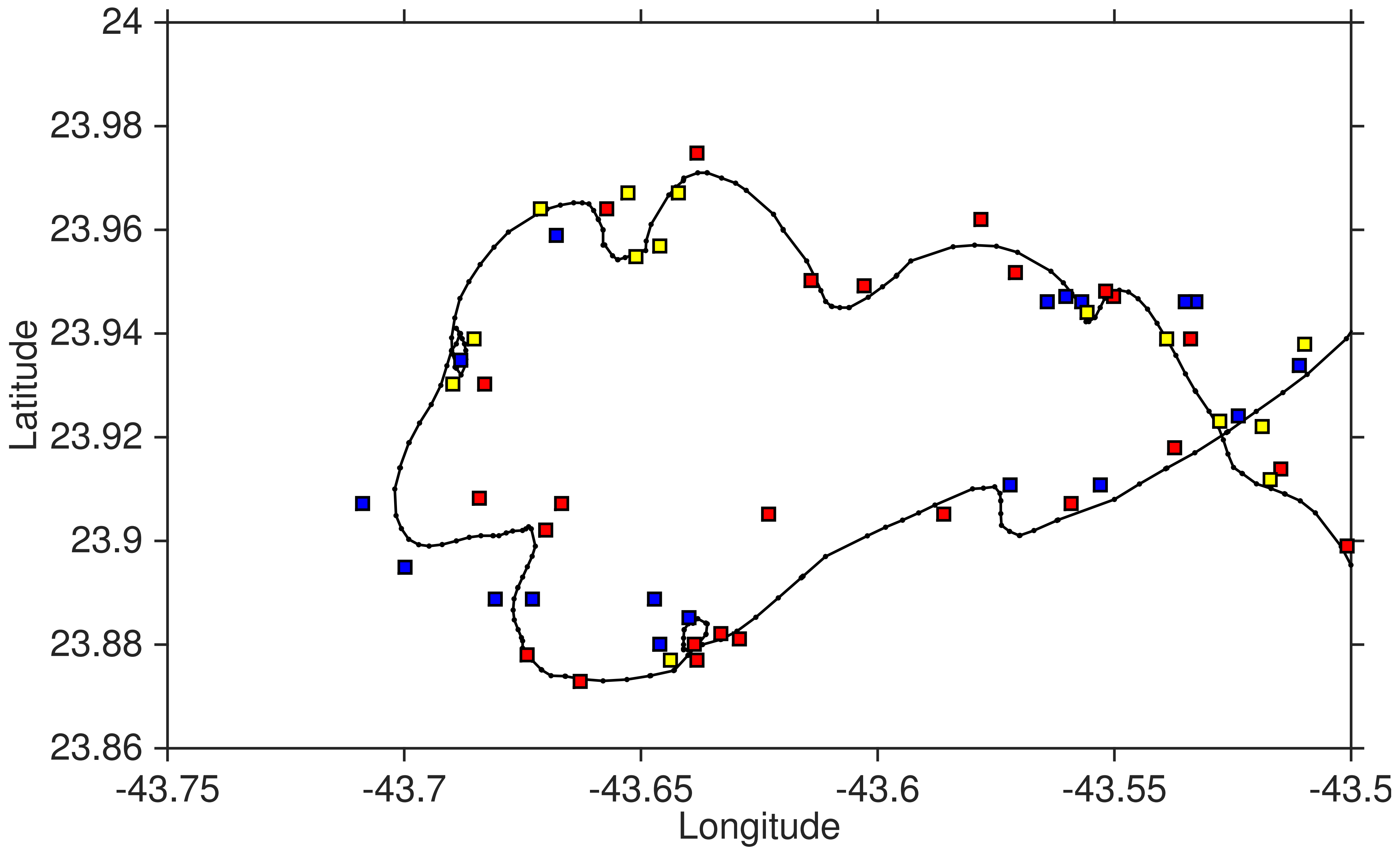
  -> **Example 3**

Example 1:
Characterizing errors and using uncertainty estimates

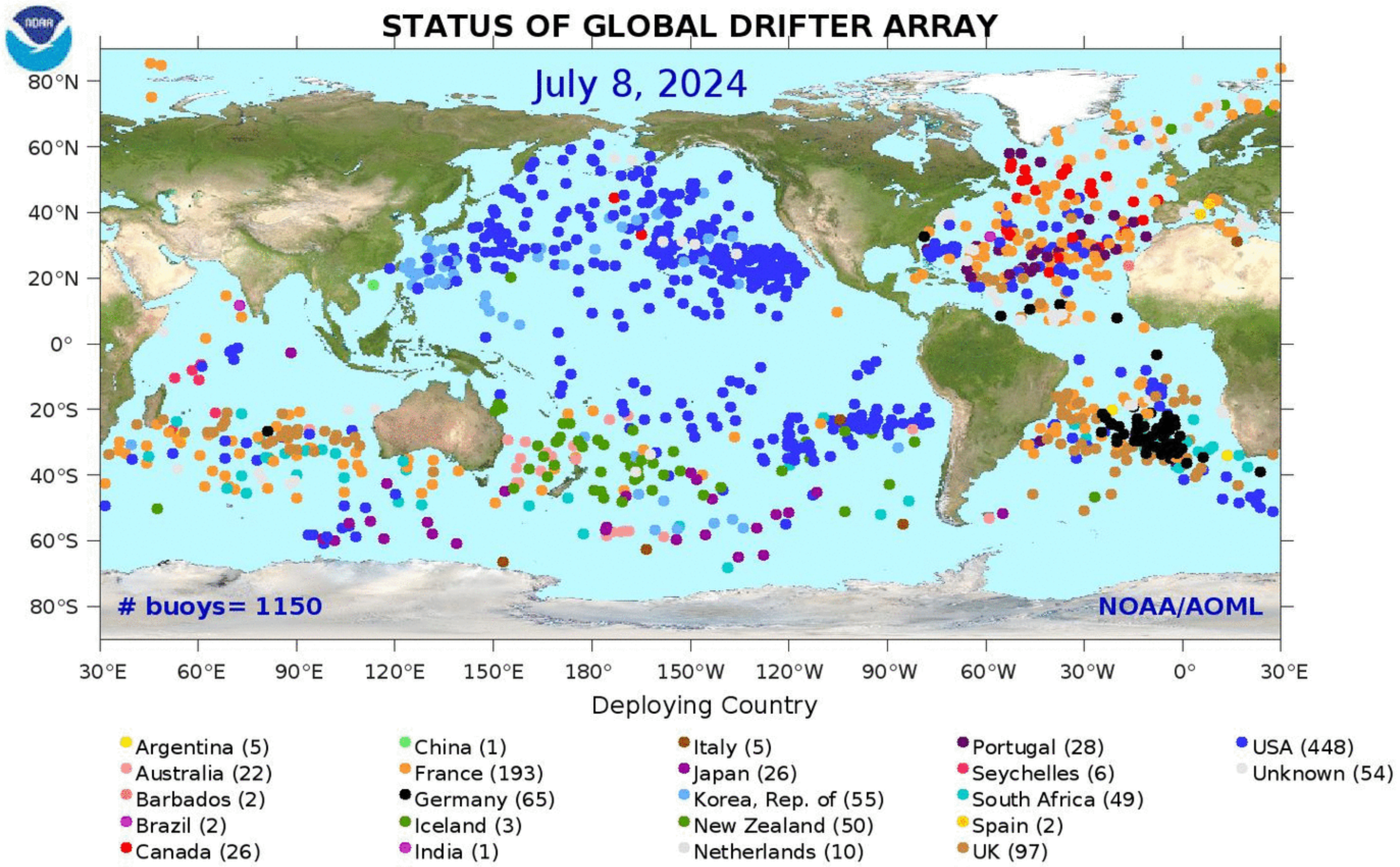
A global surface drifter data set at hourly resolution

Elipot et al. 2016

- <u>Goal</u>: produce time series of drifter positions and velocities estimates, with uncertainty estimates, at regular, hourly intervals.
- <u>Obstacle</u>: the observations are uneven in time and **uncertain**.

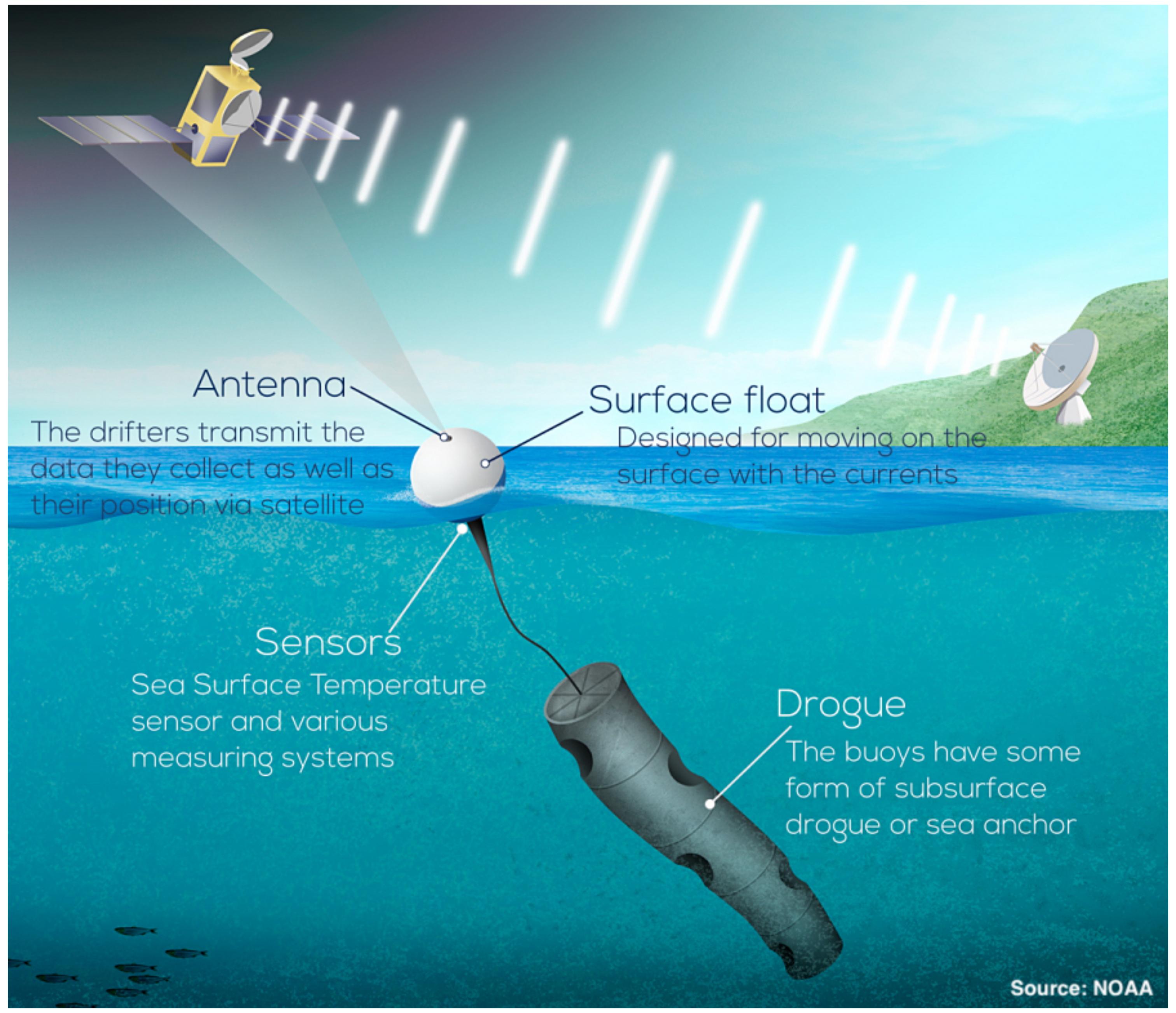


The Global Drifter Program (GDP)



https://www.aoml.noaa.gov/phod/gdp/index.php

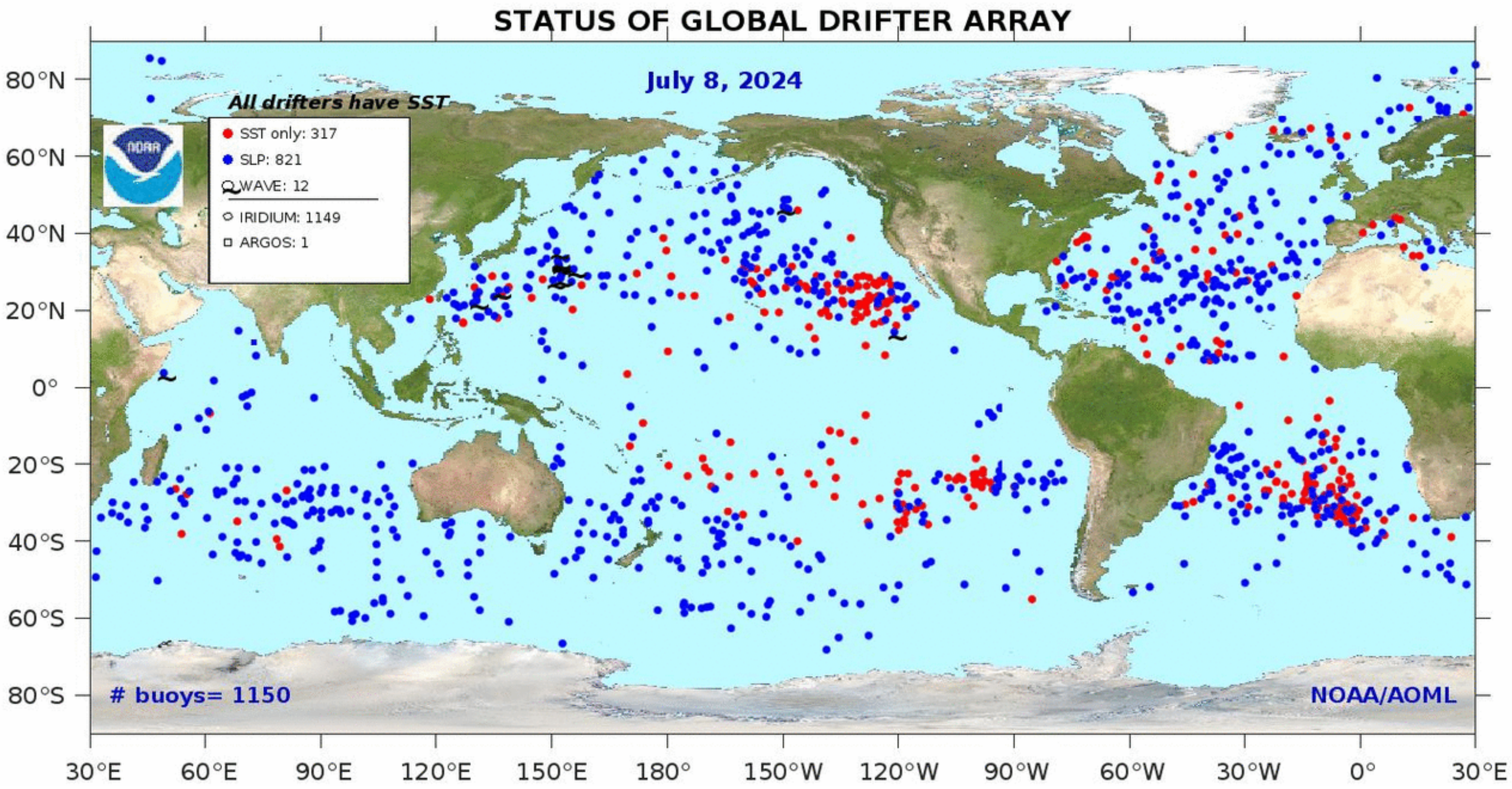
Tracking system: Argos or GPS



https://www.aoml.noaa.gov/phod/gdp/index.php

- Funded by the U.S. National Oceanic and Atmospheric Administration (NOAA), the GDP maintains, with international partners, a global array of satellite-tracked drifters to meet the need for an accurate and globally dense set (5°x5°) of in situ observations of **near-surface currents, sea surface temperature (SST), and Sea Level Pressure.**

The Global Drifter Program (GDP)

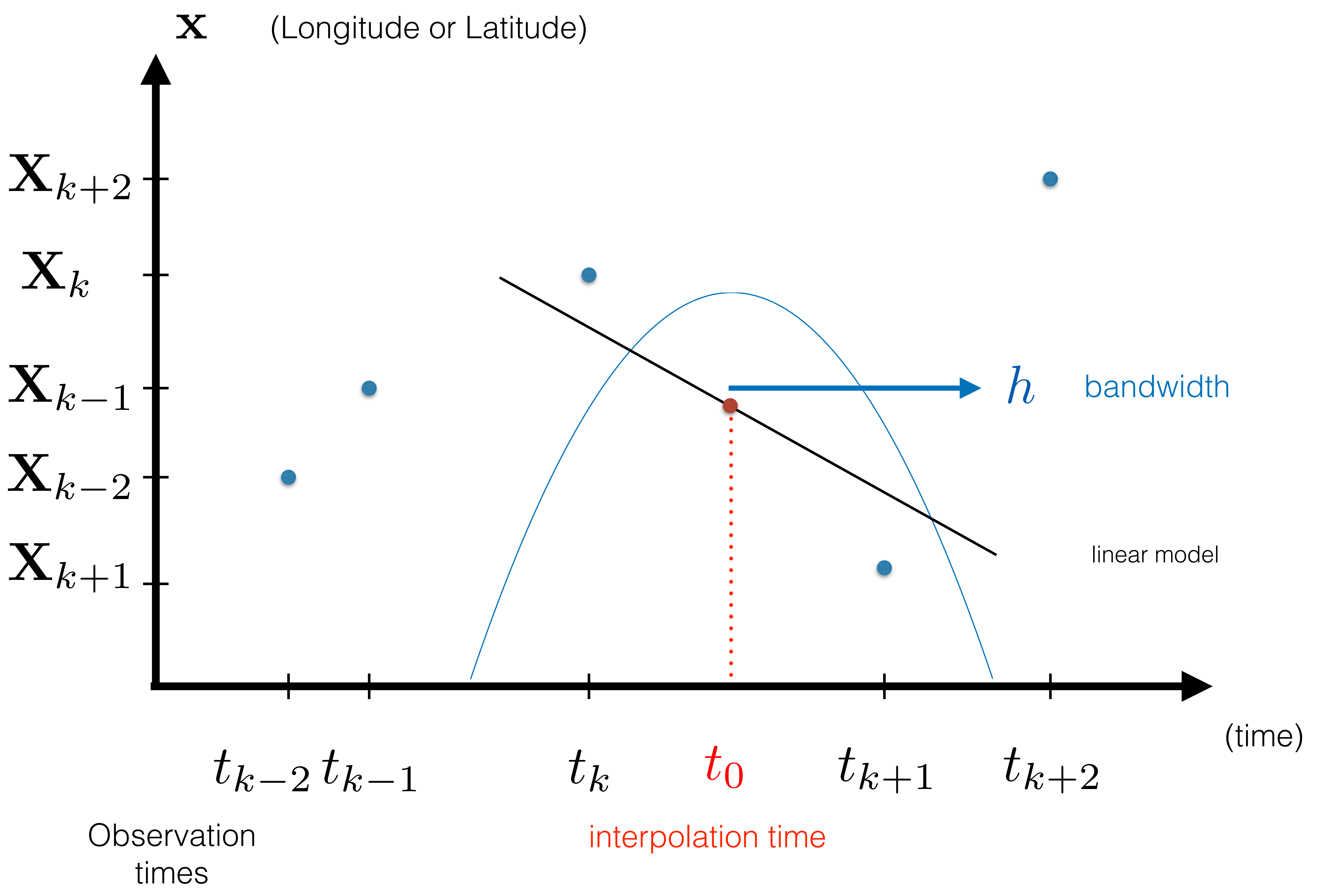


https://www.aoml.noaa.gov/phod/gdp/index.php

- The GDP data are transmitted to **operational users** via the **WMO Global Telecommunication System (GTS)** and later distributed to **research users** in the form of **quality-controlled (QC), delayed mode products**. The release of QC data products is led by the Data Assembly Center (DAC) of the GDP at the NOAA Atlantic Oceanographic and Meteorological Laboratory (AOML), Miami FL.

What were we trying to do?

Fit a linear model within a sliding window to obtain **estimates** of longitude or latitude at hourly intervals



What were we trying to do?

Mathematical models for longitude and latitude:

$$\begin{aligned}\phi(t;\boldsymbol{\beta}^{\phi}) &= \beta_0^{\phi} + \beta_1^{\phi}(t - t_0)\\ \theta(t;\boldsymbol{\beta}^{\theta}) &= \beta_0^{\theta} + \beta_1^{\theta}(t - t_0)\end{aligned}$$

We seek a set of parameters $\boldsymbol{\beta} = \left[\boldsymbol{\beta}^{\phi}, \boldsymbol{\beta}^{\theta}\right] = [\beta_0^{\phi}, \beta_1^{\phi}, \beta_0^{\theta}, \beta_1^{\theta}]$

that maximizes a weighted probability of the observed data, or *weighted likelihood:*

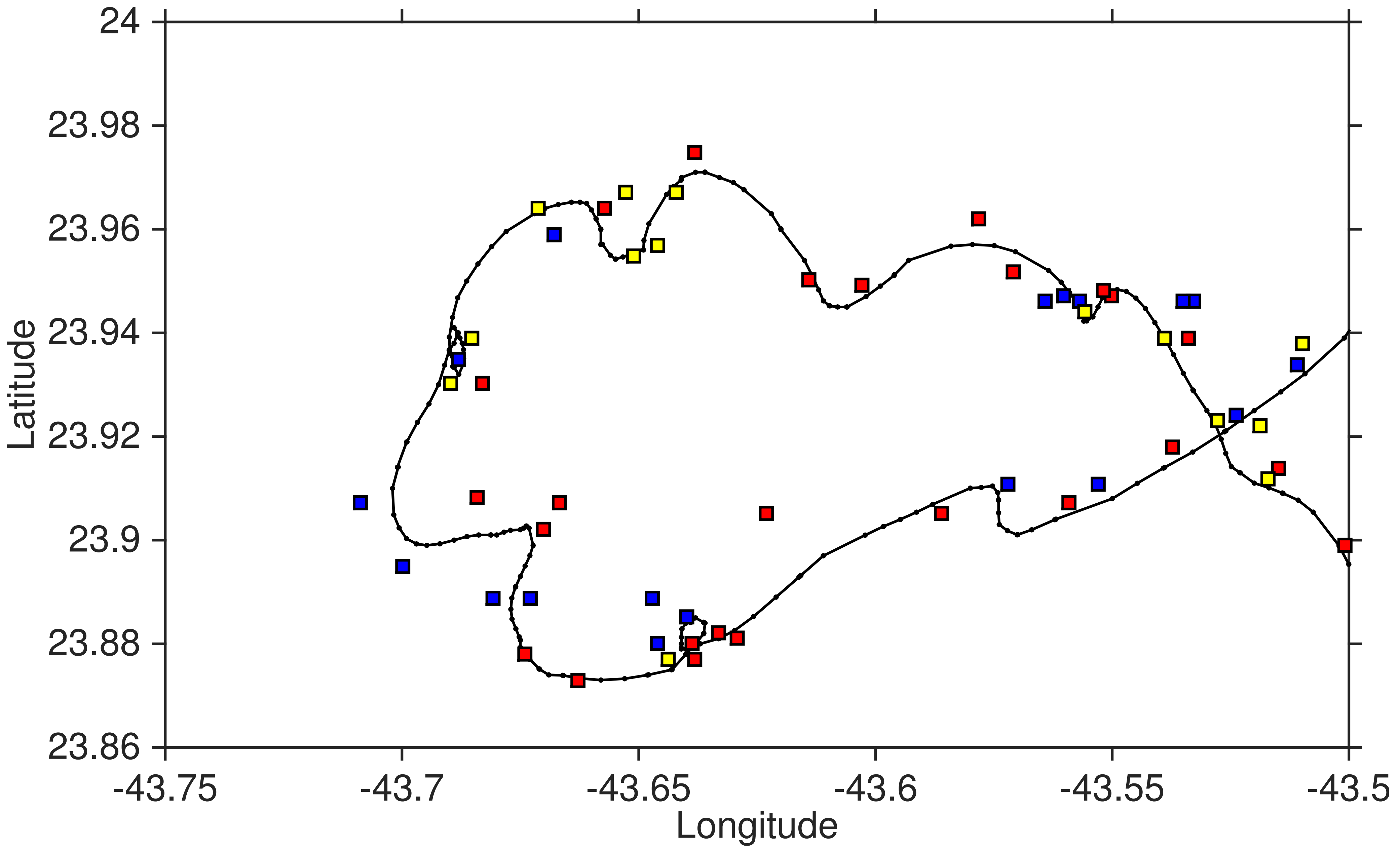
$$L(\boldsymbol{\beta}) = \prod_{k=1}^{N} \left\{p\left[\Phi_k, \Theta_k, \phi(t_k;\boldsymbol{\beta}^{\phi}), \theta(t_k;\boldsymbol{\beta}^{\theta})\right]\right\}^{w_k}$$

p is the PDF describing whether a location observation (Φ_k, Θ_k)

will yield the true location $(\phi(t_k), \theta(t_k))$ ➡ That's the pdf of the **errors**!

Characterizing drifter position errors

by fitting PDF models



$$\Delta\phi = \Phi - \phi$$

$$\Delta\theta = \Theta - \theta$$

error is observation minus truth

Black dots: *GPS* positions taken as "truth"

Colored squares: *Argos* **position estimates** with location **error** characterized by a class, associated with an equivalent radius error *r* assumed to be representative of **one standard deviation of the errors**:

Class 3: r < 250 m

Class 2: r = 250-500 m

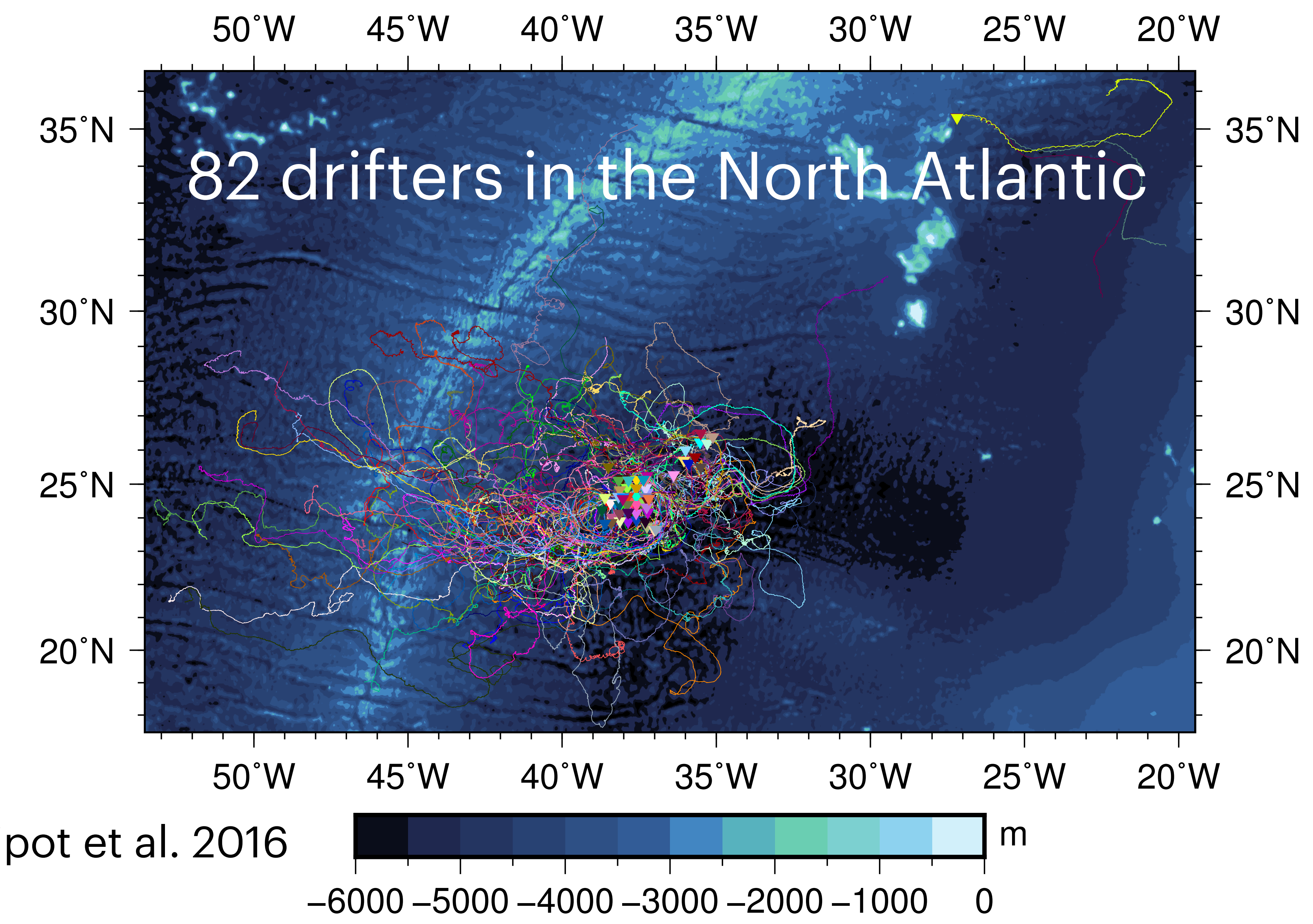
Class 1: r = 500-1500 m

Class 0: r > 1500 m

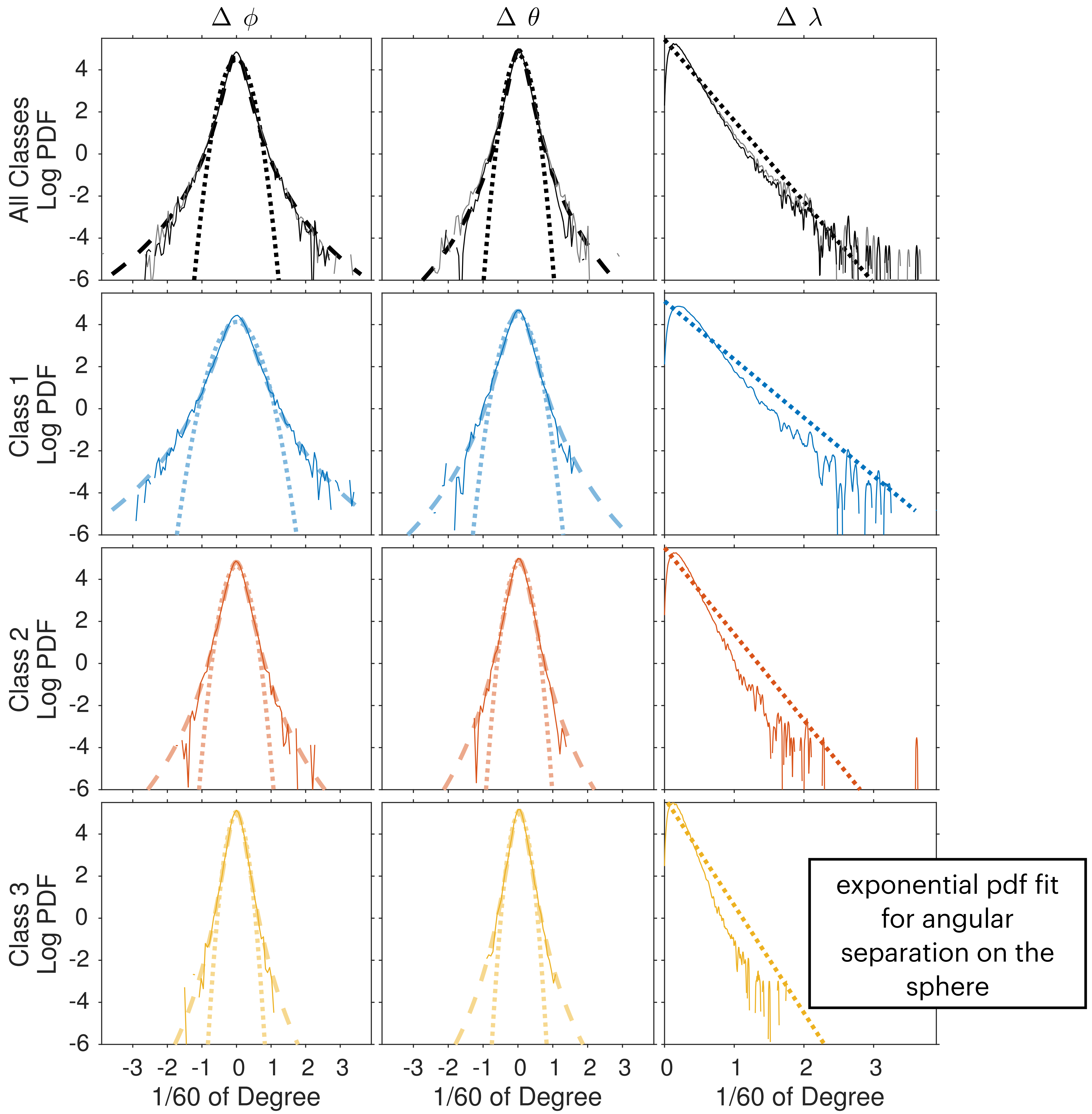
Characterizing drifter position errors

by fitting PDF models

Elipot et al. 2016



Elipot et al. 2016



empirical PDF of errors

Gaussian PDF fits

t location-scale PDF fits

Characterizing drifter position errors

by fitting PDF models

Normal PDF ⟶ $$p(z|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}$$

t location-scale PDF ⟶ $$p(z|\mu,\sigma,\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{1}{\nu}\left(\frac{z-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}}$$

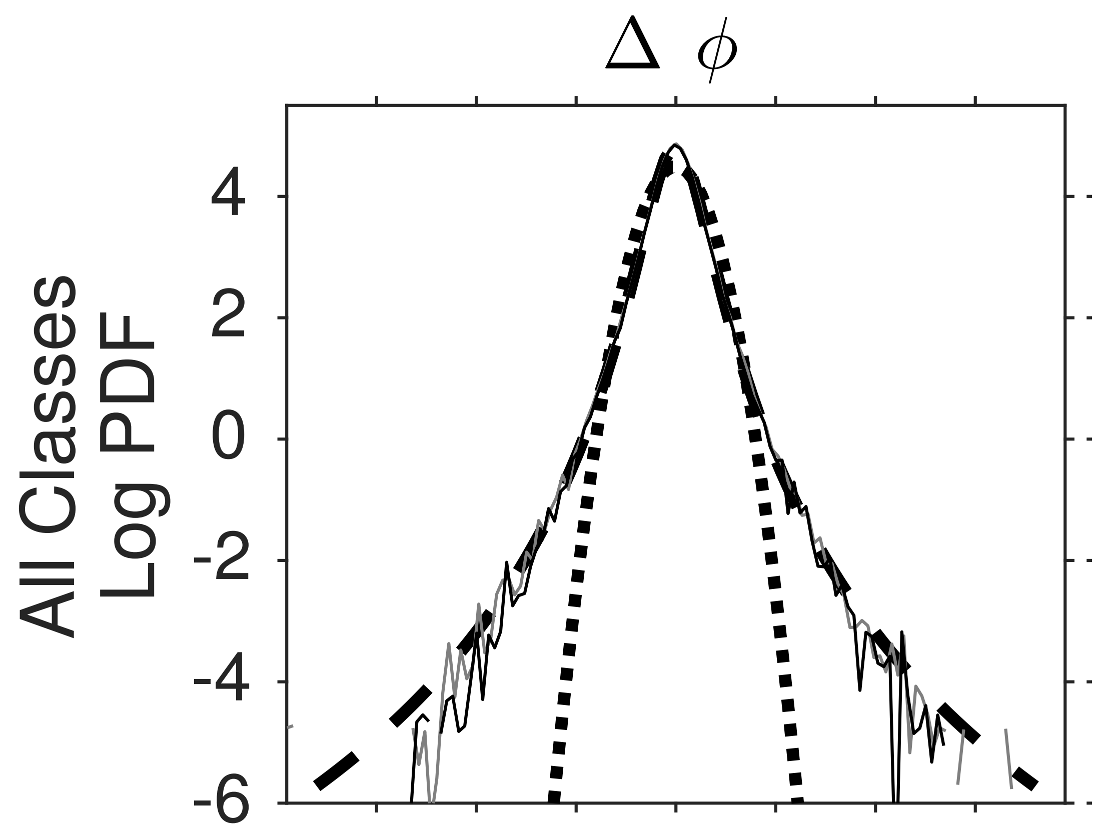
(nonstandardized Student's t PDF)

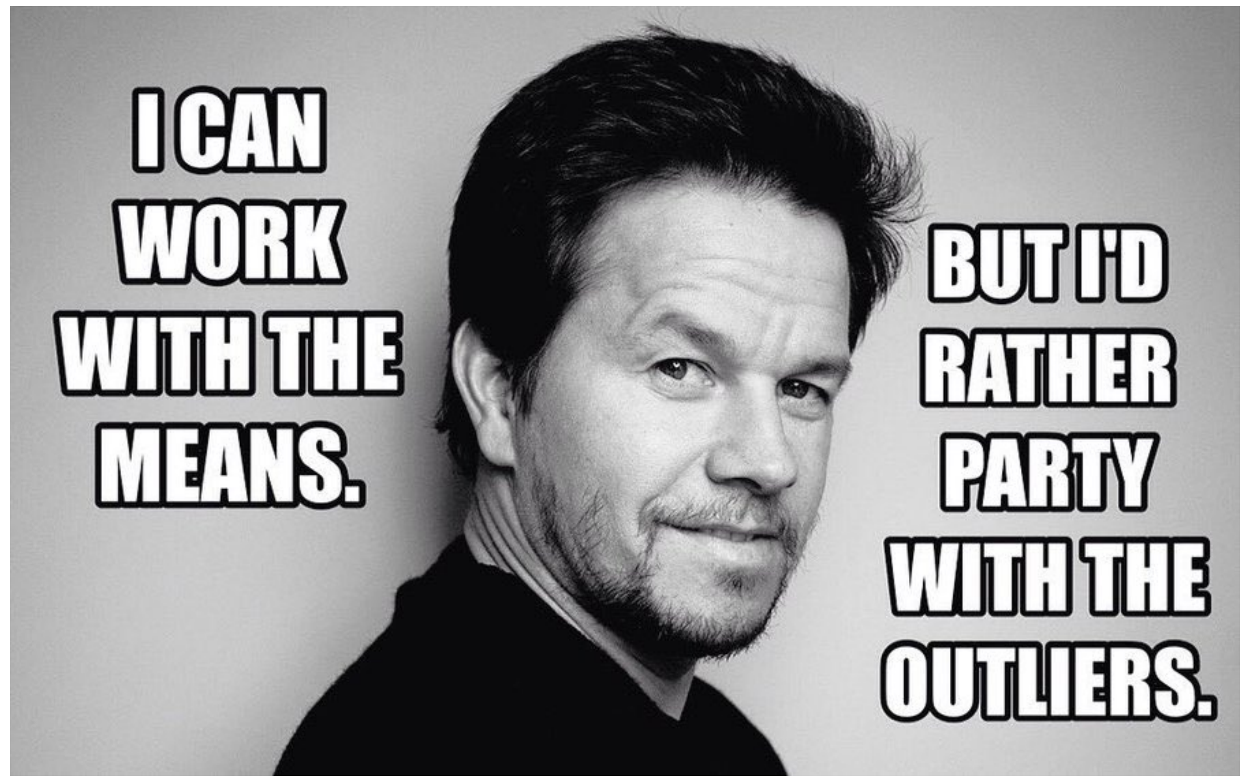
$$\Gamma(a) = \int_0^\infty t^{a-1}\exp(-t)$$

μ location parameter

σ scale parameter

ν shape parameter

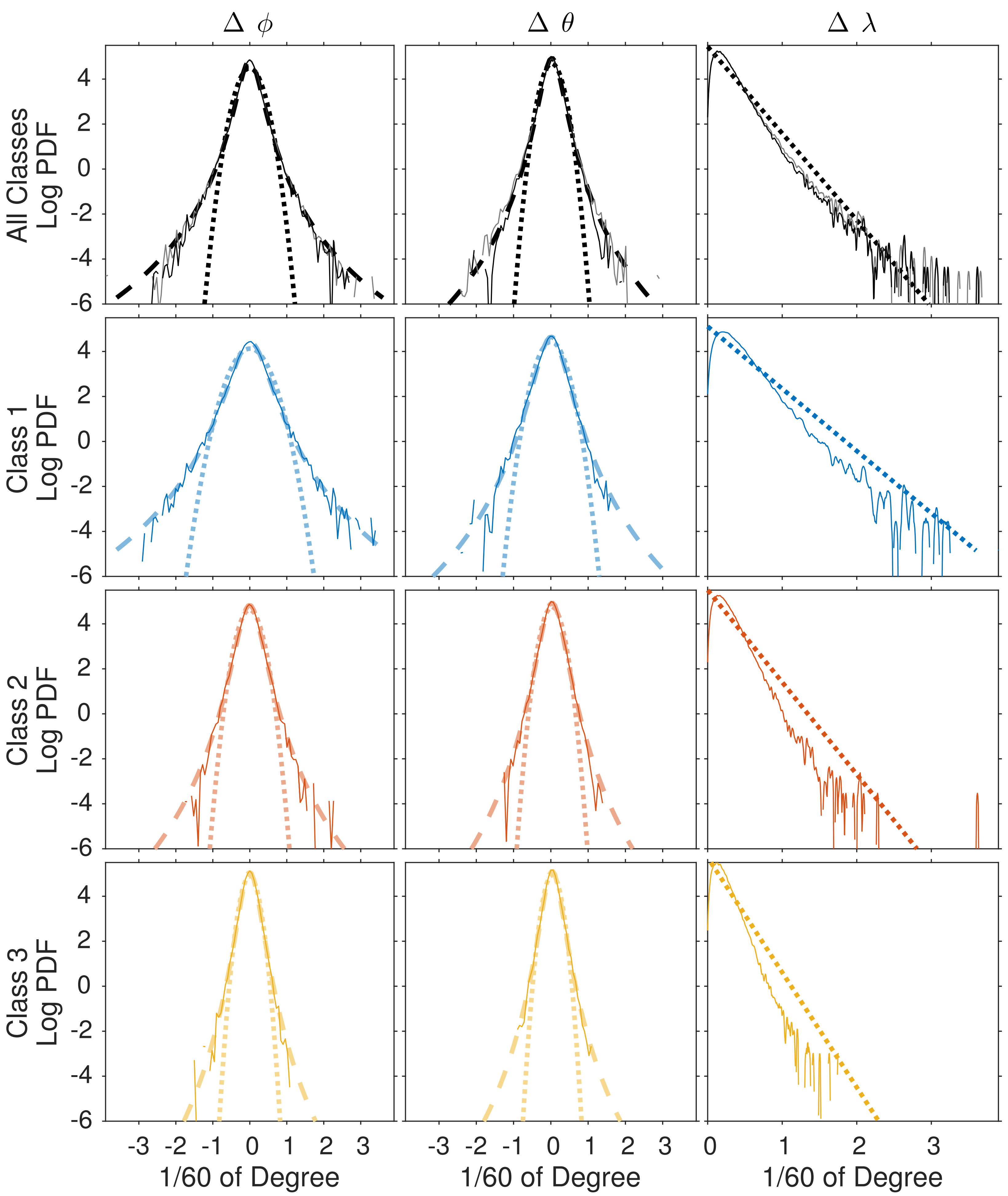




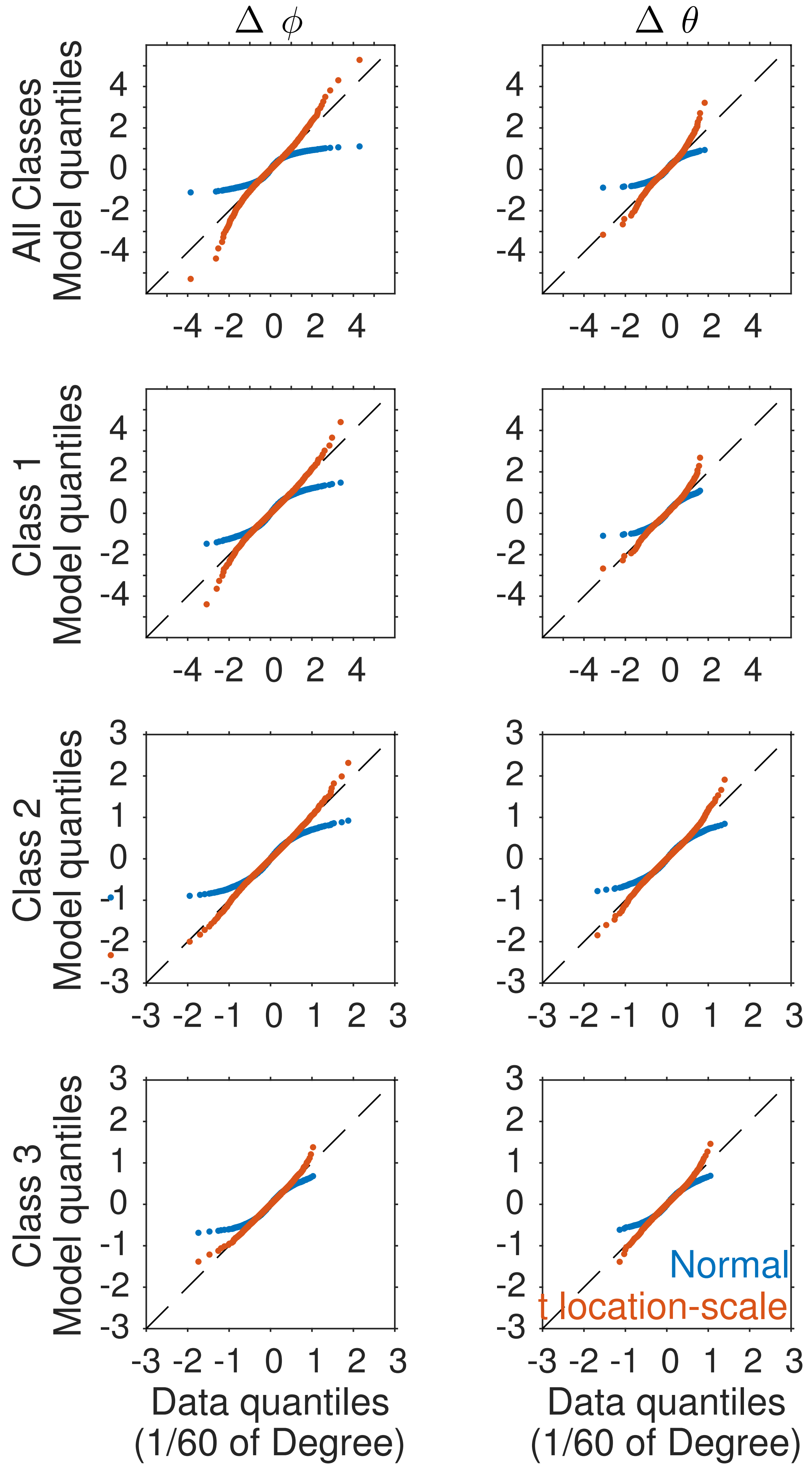
@ResearchMark

Characterizing drifter position errors

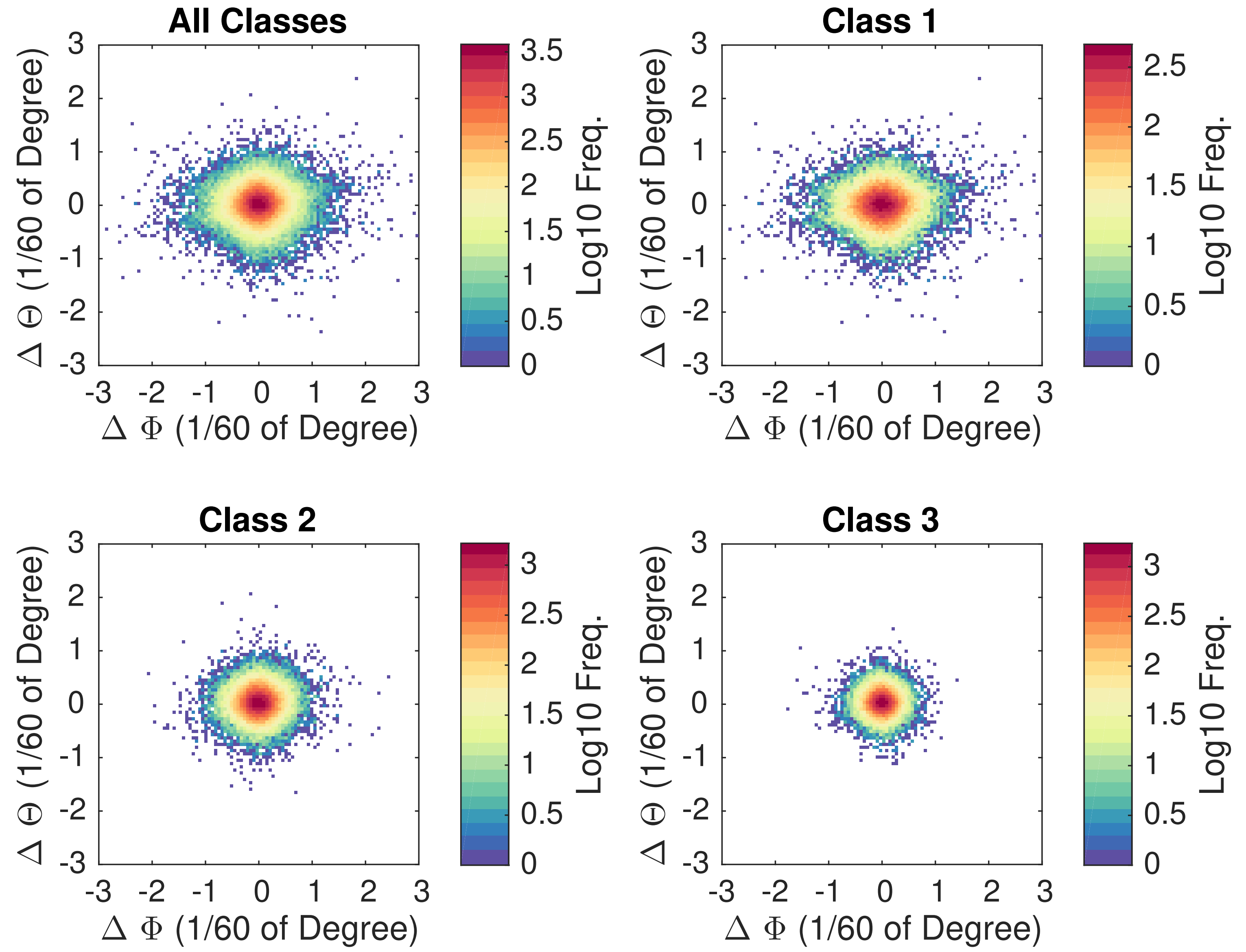
by fitting PDF models



Observed and fitted distribution



Quantile-quantile plots



2D histogram of errors

Elipot et al. 2016

Characterizing drifter position errors

by fitting PDF models

Weighted likelihood: $$L(\boldsymbol{\beta}) = \prod_{k=1}^{N} \left\{ p \left[\Phi_k, \Theta_k, \phi(t_k; \boldsymbol{\beta}^\phi), \theta(t_k; \boldsymbol{\beta}^\theta) \right] \right\}^{w_k}$$

log-likelihood: $$l(\boldsymbol{\beta}) = \sum_{k=1}^{N} w_k \ln p \left[\Phi_k, \Theta_k, \phi\left(t_k; \boldsymbol{\beta}^\phi\right), \theta\left(t_k; \boldsymbol{\beta}^\theta\right) \right]$$

lat, lon errors independent: $$p \left[\Phi_k, \Theta_k, \phi\left(t_k; \boldsymbol{\beta}^\phi\right), \theta\left(t_k; \boldsymbol{\beta}^\theta\right) \right] = p \left[\Phi_k, \phi\left(t_k; \boldsymbol{\beta}^\phi\right) \right] p \left[\Theta_k, \theta\left(t_k; \boldsymbol{\beta}^\theta\right) \right]$$

and consider that the errors are distributed like *t* PDFs ...

Characterizing drifter position errors

by fitting PDF models

$$l(\boldsymbol{\beta}) = \sum_{k=1}^{N}\left\{-w_k \ln\left[\frac{\sigma_{\Phi_k}\sqrt{\nu_{\Phi_k}\pi}\,\Gamma\left(\frac{\nu_{\Phi_k}}{2}\right)}{\Gamma\left(\frac{\nu_{\Phi_k}+1}{2}\right)}\right] - w_k \ln\left[\frac{\sigma_{\Theta_k}\sqrt{\nu_{\Theta_k}\pi}\,\Gamma\left(\frac{\nu_{\Theta_k}}{2}\right)}{\Gamma\left(\frac{\nu_{\Theta_k}+1}{2}\right)}\right]\right\}$$
$$+\sum_{k=1}^{N}\left\{-w_k\frac{\nu_{\Phi_k}+1}{2}\ln\left[1+\frac{1}{\nu_{\Phi_k}}\left(\frac{\Phi_k-\phi(t_k;\boldsymbol{\beta}^{\phi})-\mu_{\Phi_k}}{\sigma_{\Phi_k}}\right)^2\right]\right.$$
$$\left.-w_k\frac{\nu_{\Theta_k}+1}{2}\ln\left[1+\frac{1}{\nu_{\Theta_k}}\left(\frac{\Theta_k-\theta(t_k;\boldsymbol{\beta}^{\theta})-\mu_{\Theta_k}}{\sigma_{\Theta_k}}\right)^2\right]\right\}$$

$\mu_{\Phi_k}, \mu_{\Theta_k}, \sigma_{\Phi_k}, \sigma_{\Theta_k}, \nu_{\Phi_k}, \nu_{\Theta_k}$ location, scale, and shape parameters of the *t* PDF of longitude or latitude observations with index k.

Log-likelihood expression to maximize $\boldsymbol{\beta} = \left[\boldsymbol{\beta}^{\phi}, \boldsymbol{\beta}^{\theta}\right] = [\beta_0^{\phi}, \beta_1^{\phi}, \beta_0^{\theta}, \beta_1^{\theta}] \longrightarrow$

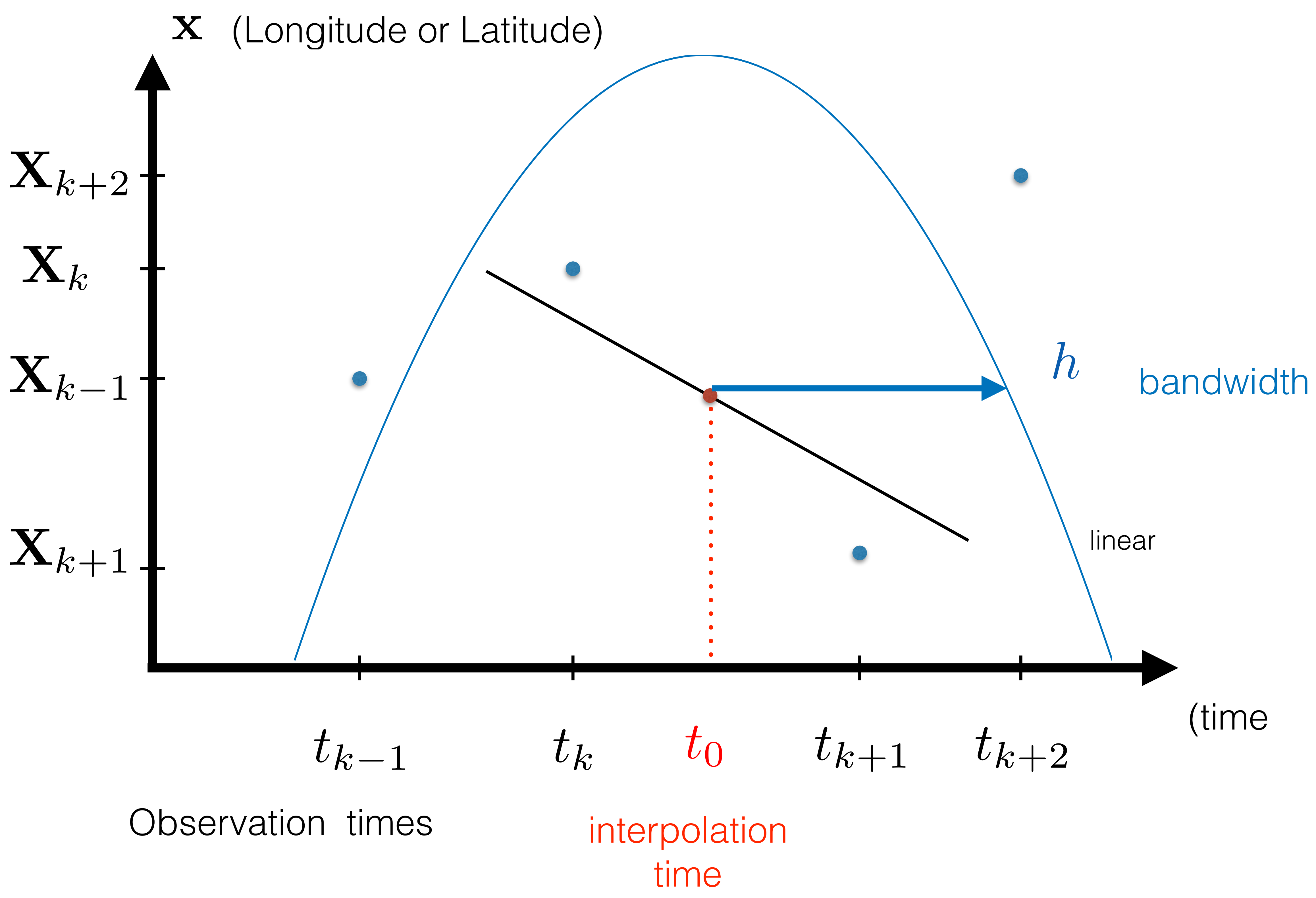
$$\begin{aligned}\phi(t;\boldsymbol{\beta}^{\phi}) &= \beta_0^{\phi} + \beta_1^{\phi}(t-t_0)\\ \theta(t;\boldsymbol{\beta}^{\theta}) &= \beta_0^{\theta} + \beta_1^{\theta}(t-t_0)\end{aligned}$$

Evaluate at $t = t_0$ to obtain fitted/estimated positions ...

Example 2:
Deriving uncertainty estimates

Deriving uncertainty estimates of the position estimates

Using a bootstrap method



We chose to use *N = 4* observation points to obtain each $\hat{\boldsymbol{\beta}} = [\hat{\beta}_0^{\phi}, \hat{\beta}_1^{\phi}, \hat{\beta}_0^{\theta}, \hat{\beta}_1^{\theta}]$ at hourly time scale

How to get $\mathrm{Var}[\hat{\beta}_j^x]$?

Asymptotic theory:

$$N \to +\infty, \quad \hat{\beta}_j - \beta_j \sim \mathcal{N}(0, \sigma)$$

Instead, use a bootstrap method, the jackknife!

Re-estimate $\hat{\boldsymbol{\beta}}$ 4 times using three observations out of four to obtain $\mathrm{Var}[\hat{\beta}_j^x]$

Deriving uncertainty estimates of the position estimates

Using a bootstrap method

Re-calculate estimates 4 times using three observations out of four, in order to obtain $s_j = \mathrm{Var}\left[\hat{\beta}_j\right]$

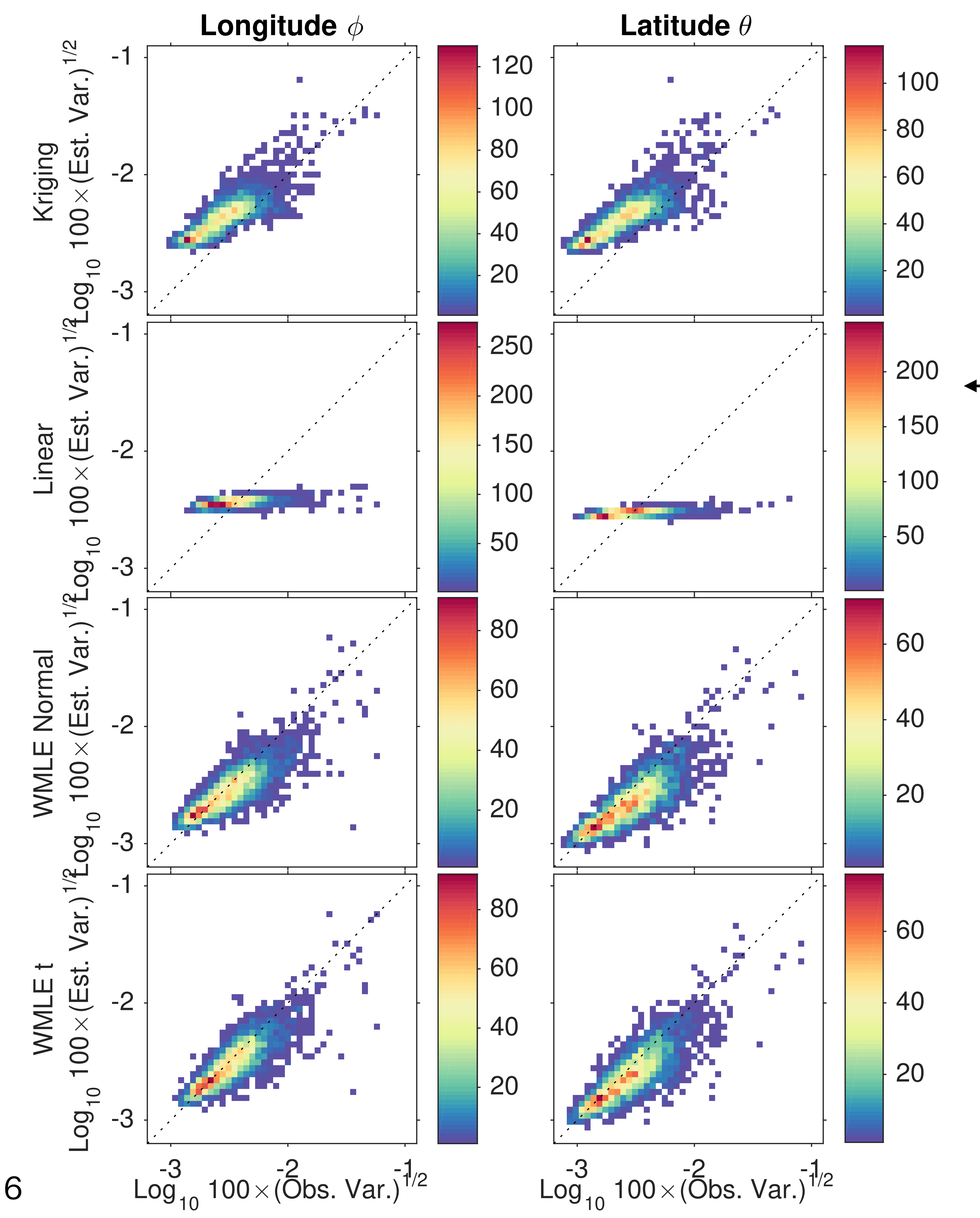
"t statistic": $t_j = \frac{\widehat{\beta_j} - \beta_j}{s_j/\sqrt{N}}$

95% confidence interval:

$$[\hat{\beta}_j - t_{\alpha/2,N-1}\frac{s_j}{\sqrt{N}}, \hat{\beta}_j + t_{1-\alpha/2,N-1}\frac{s_j}{\sqrt{N}}]$$

$$t_{\alpha/2,N-1} = t_{1-\alpha/2,N-1} = 3.1824 \quad \text{with } \alpha = 0.05$$

Validation of uncertainty estimates



Elipot et al. 2016

Two-dimensional histograms of the square root of **estimated** versus **observed error variance** for each different methods of estimation.

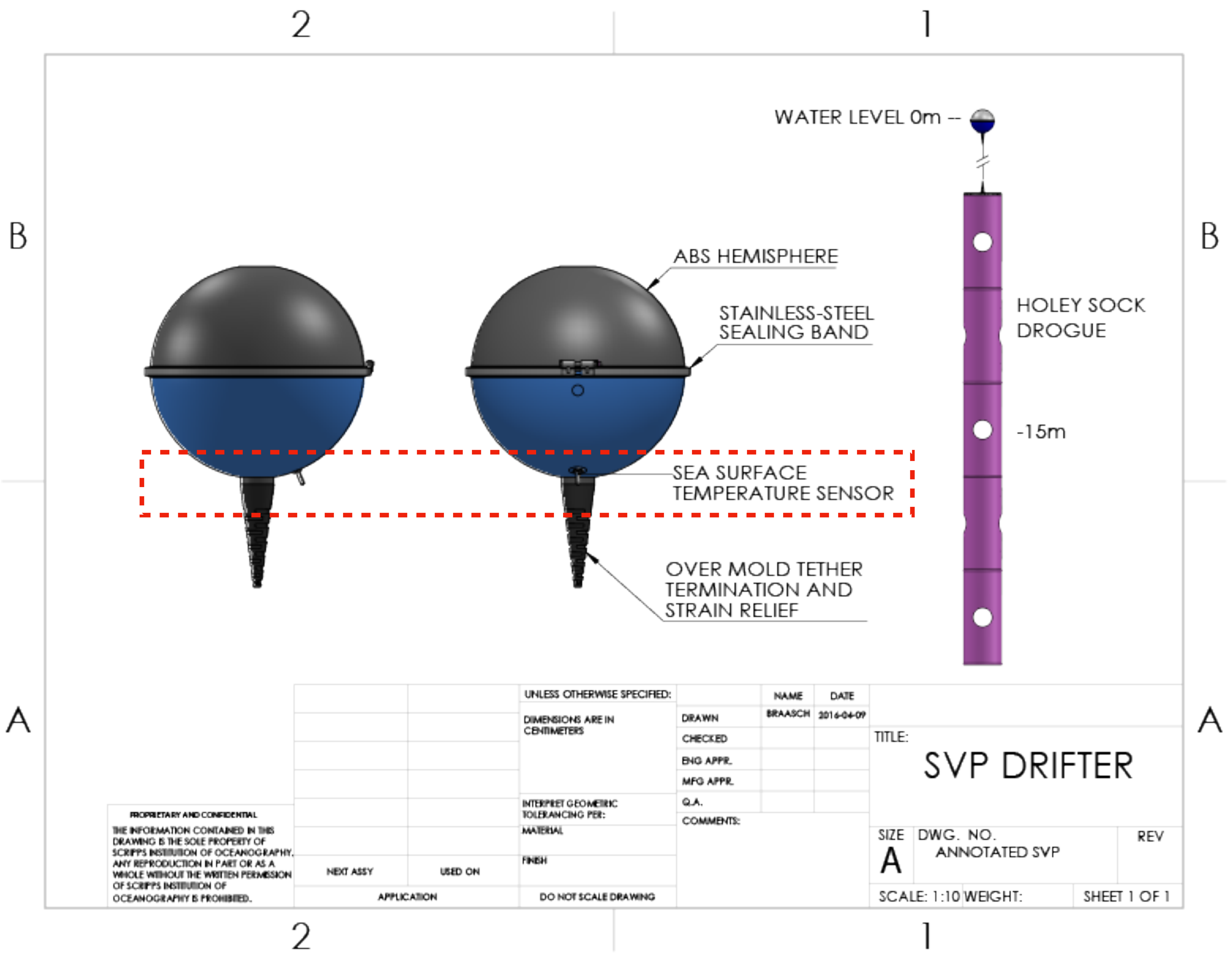
We considered 3366 10-day trajectory segments with 50% overlap in order to calculate the **variance of the observed longitude and latitude error time series** for each segment, which we compare to **the mean of the estimated variances for longitude and latitude**

Example 3:
Deriving uncertainty estimates

A dataset of hourly sea surface temperature from drifting buoys

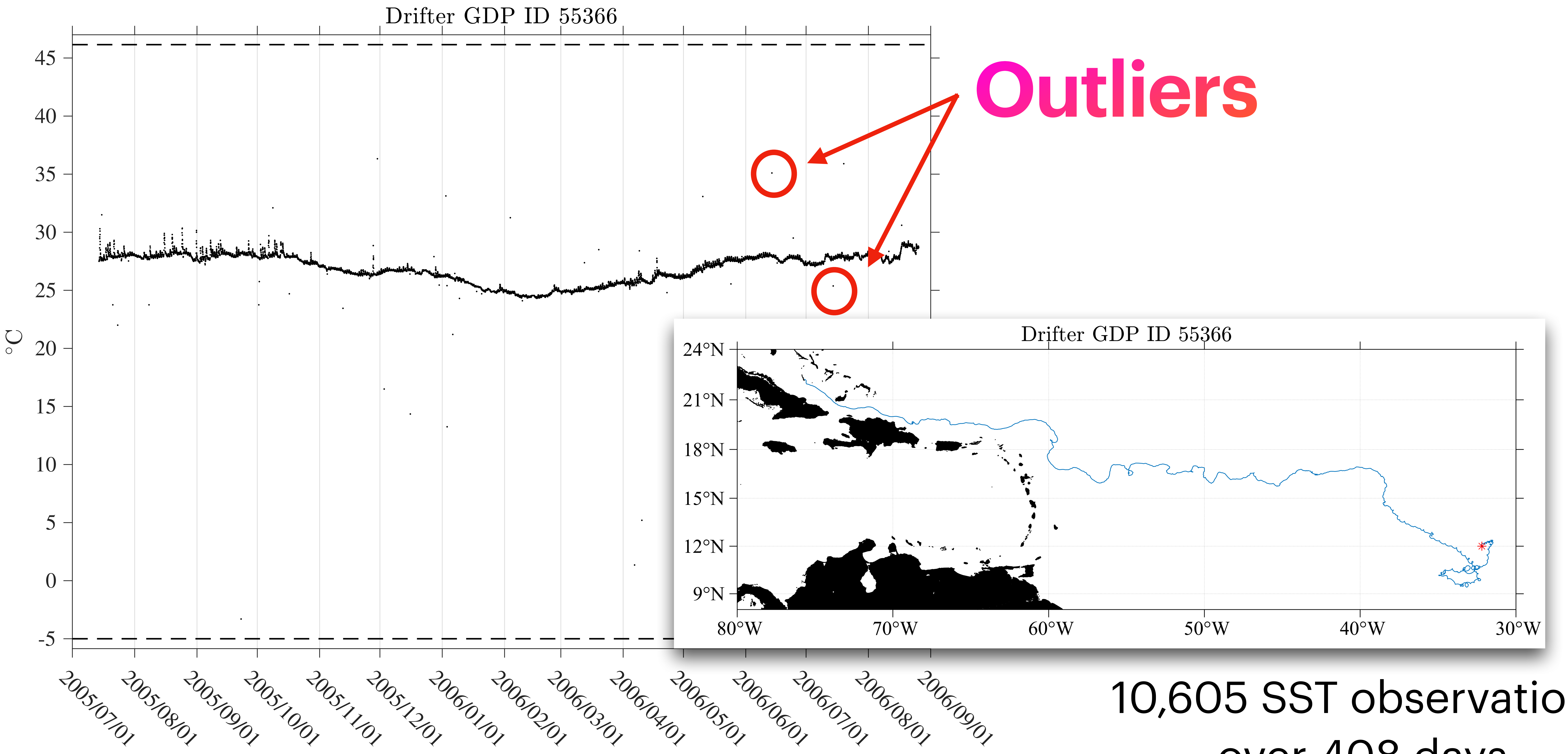
Elipot et al. 2022

- <u>Goal</u>: produce time series of drifter SST estimates, with uncertainty estimates, at regular, hourly intervals.
- <u>Obstacle</u>: Many!



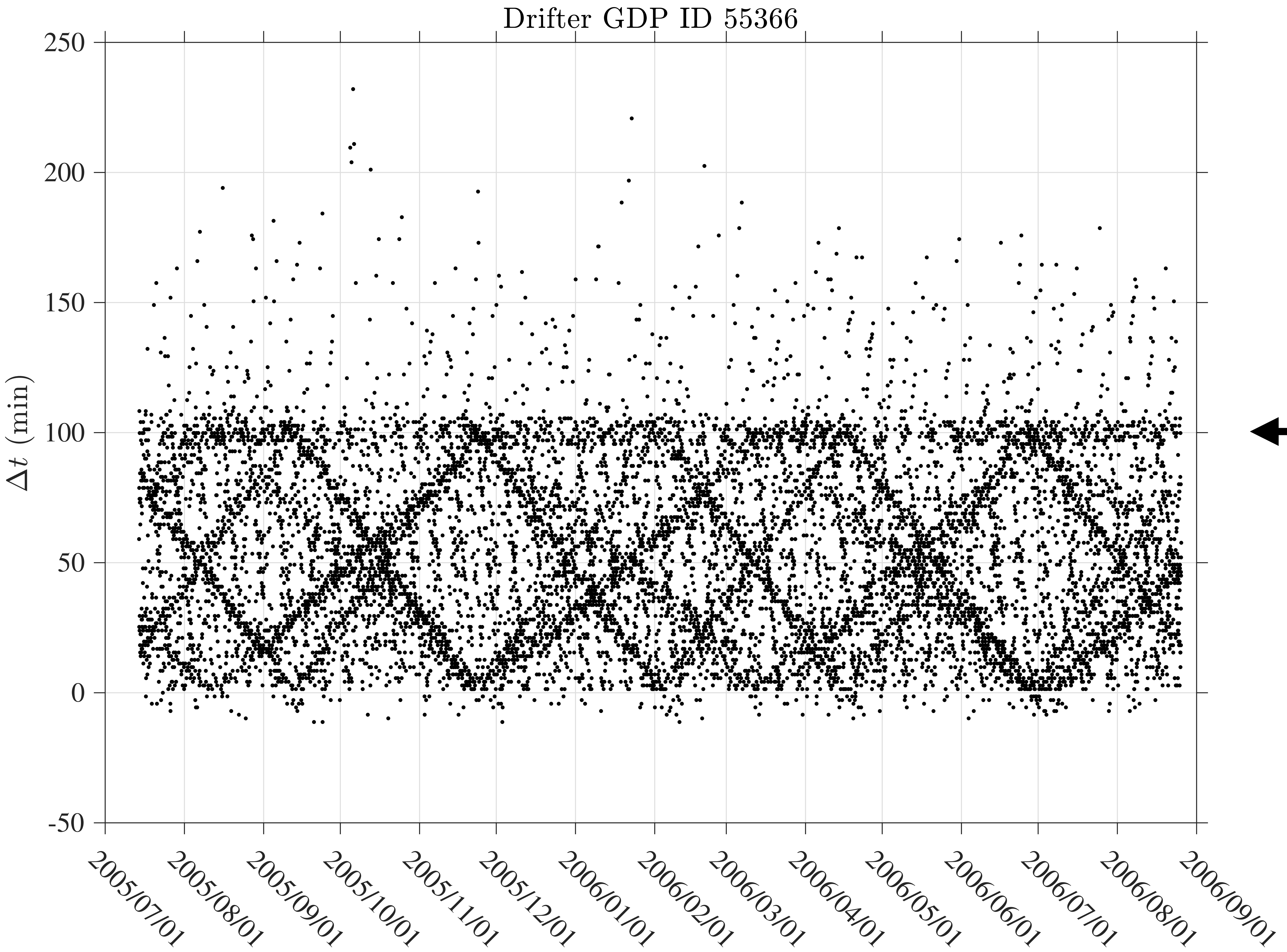
Time series of SST observations from drifter ID 55366

[Argos-tracked drifter, SVP type, built by Pacific Gyre]



10,605 SST observations
over 408 days

Time series of temporal intervals (Δt) between consecutive observations:



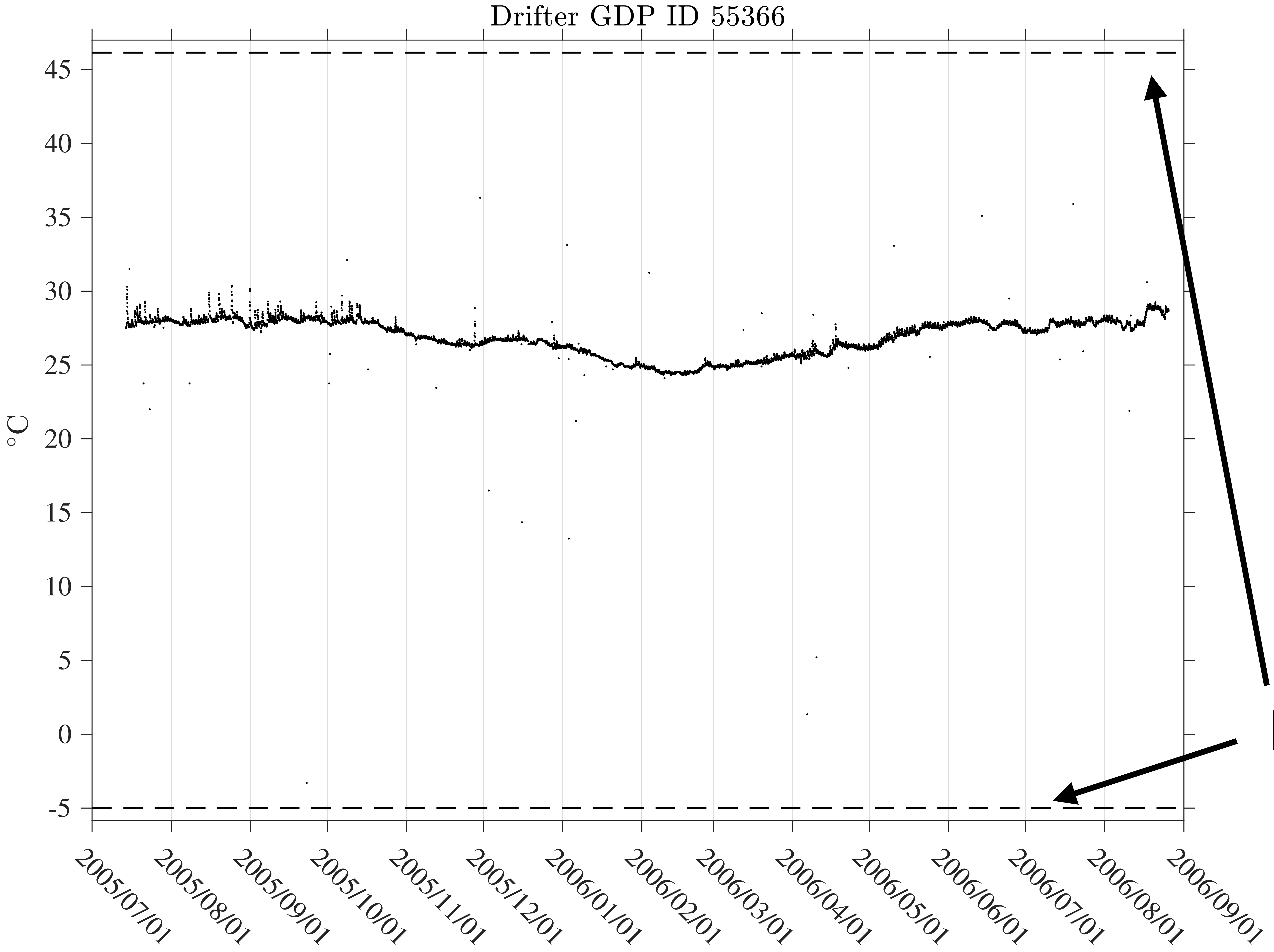
This drifter sampled every 60 s, averaged every 15 samples, and transmitted every 90 s.

Argos satellite orbital period (101.47 min)

Uneven sampling

Time series of SST observations from drifter ID 55366

For 85% of drifters, observations are derived from a SST equation:



$$\mathrm{SST} = an + b$$

a is the *resolution*

n is a binary bit count

Here $a = 0.05$

$b = -5$

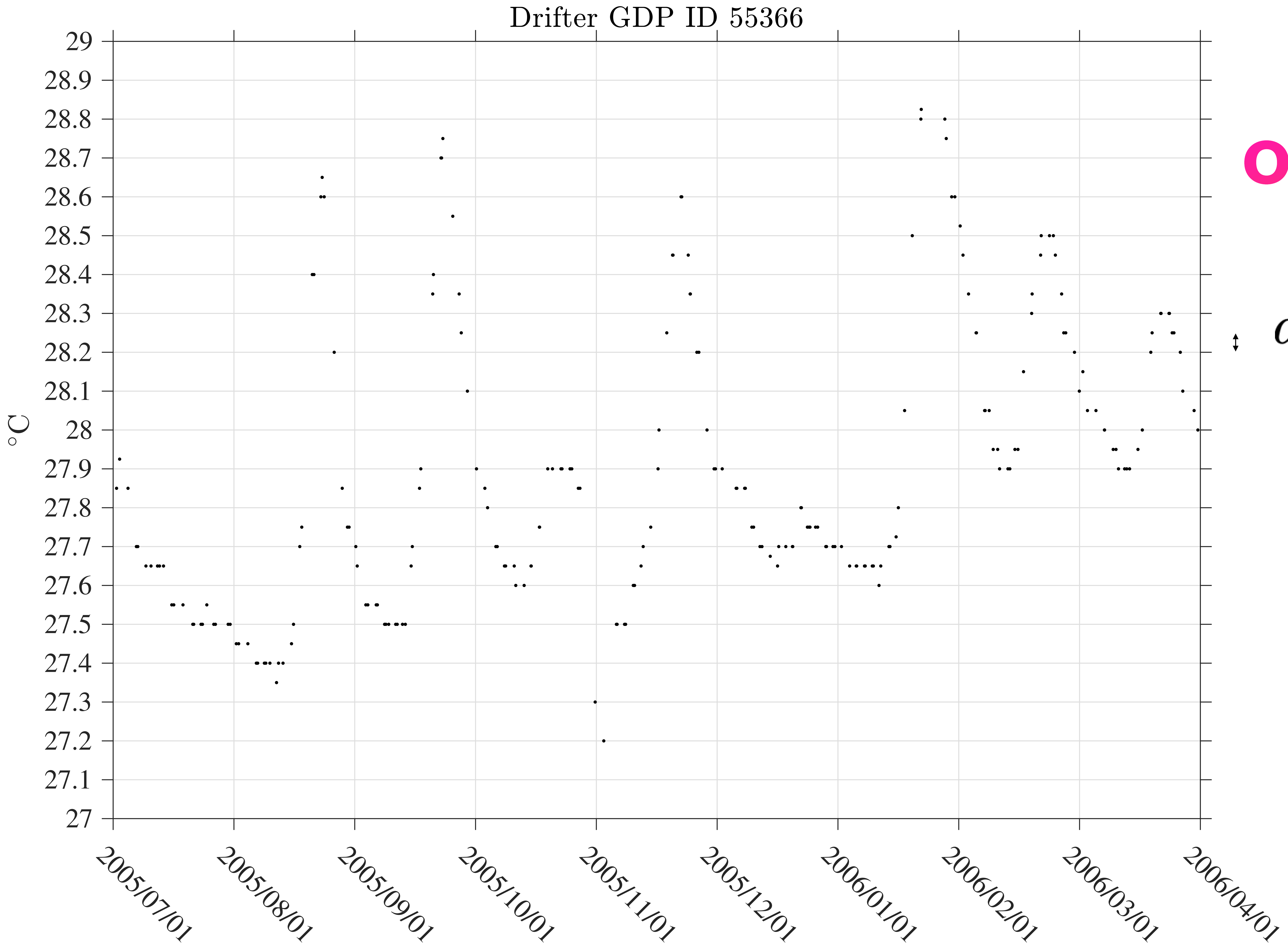
n is coded over 10 bits

Range of observable values:

$$0.05 \times [0,\ (2^{10} - 1)] - 5$$

$$= [-5,\ 46.15]$$

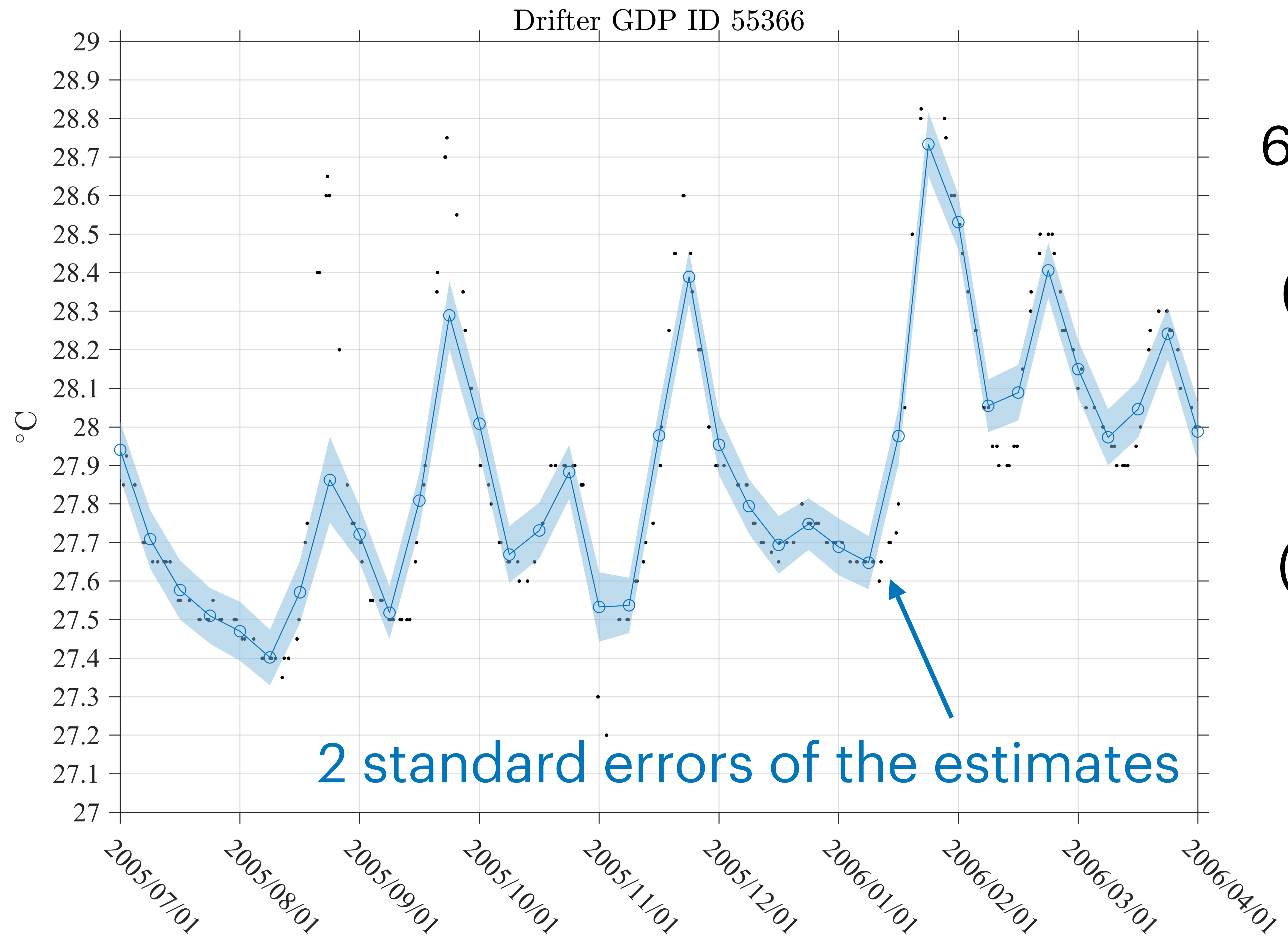
Quantized observations



$a = 0.05$ resolution

Causes an added error called **quantization error**

“Kriging” solution at 6-hour time steps



6-hourly “kriged” estimates @ 00:00, 06:00, 12:00. 18:00 (Hansen and Poulain , 1996):

≈ Optimal interpolation method

(using 5 observations before and after estimation times)

Methodology: process model

The goal is to estimate SST at time $t = t_k$
(which is observation time or hourly time steps)

process model $s_i = s_m(t_i; t_k) + \sigma_k \varepsilon_k$

observations at times t_i = SST model + noise

The noise component, ε_k, is expected to be zero-mean, to have unit variance

… locally scaled by σ_k where σ^2_k is the error variance of the observations around time t_k.

Methodology: models

process model: $s_i = s_m(t_i; t_k) + \sigma_k \varepsilon_k$

model of SST temporal evolution:

$$
\begin{aligned}
s_m(t_i;t_k) &= s_P(t_i;t_k) + s_D(t_i;t_k) \\
&= \sum_{p=0}^{P} s_{p,k}(t_i - t_k)^p + \sum_{n=1}^{N} A_{n,k} \cos[n\omega(t_i - t_k) + \phi_{n,k}] \\
&= \sum_{p=0}^{P} s_{p,k}(t_i - t_k)^p + \sum_{n=1}^{N} [\alpha_{n,k} \cos n\omega(t_i - t_k) + \beta_{n,k} \sin n\omega(t_i - t_k)]
\end{aligned}
$$

P+1+2N parameters to estimate

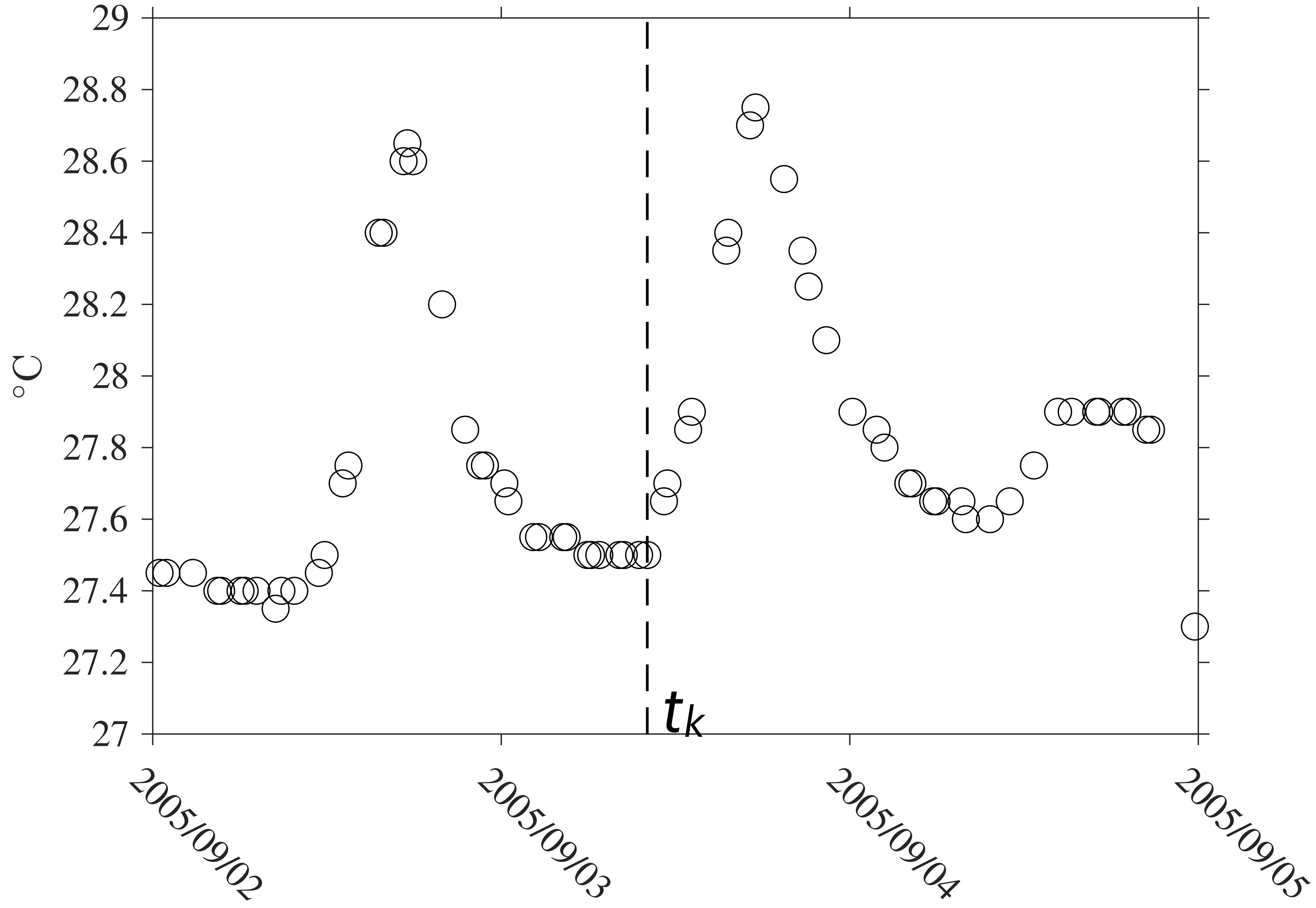
low-frequency component

diurnal evolution, ω = 1 cpd
e.g. *Gentemann (2003)*

Estimate at $t_i = t_k$: $\widehat{s}_{m,k} \equiv s_m(t_k; t_k) = s_{0,k} + \sum_{n=1}^{N} \alpha_{n,k}$

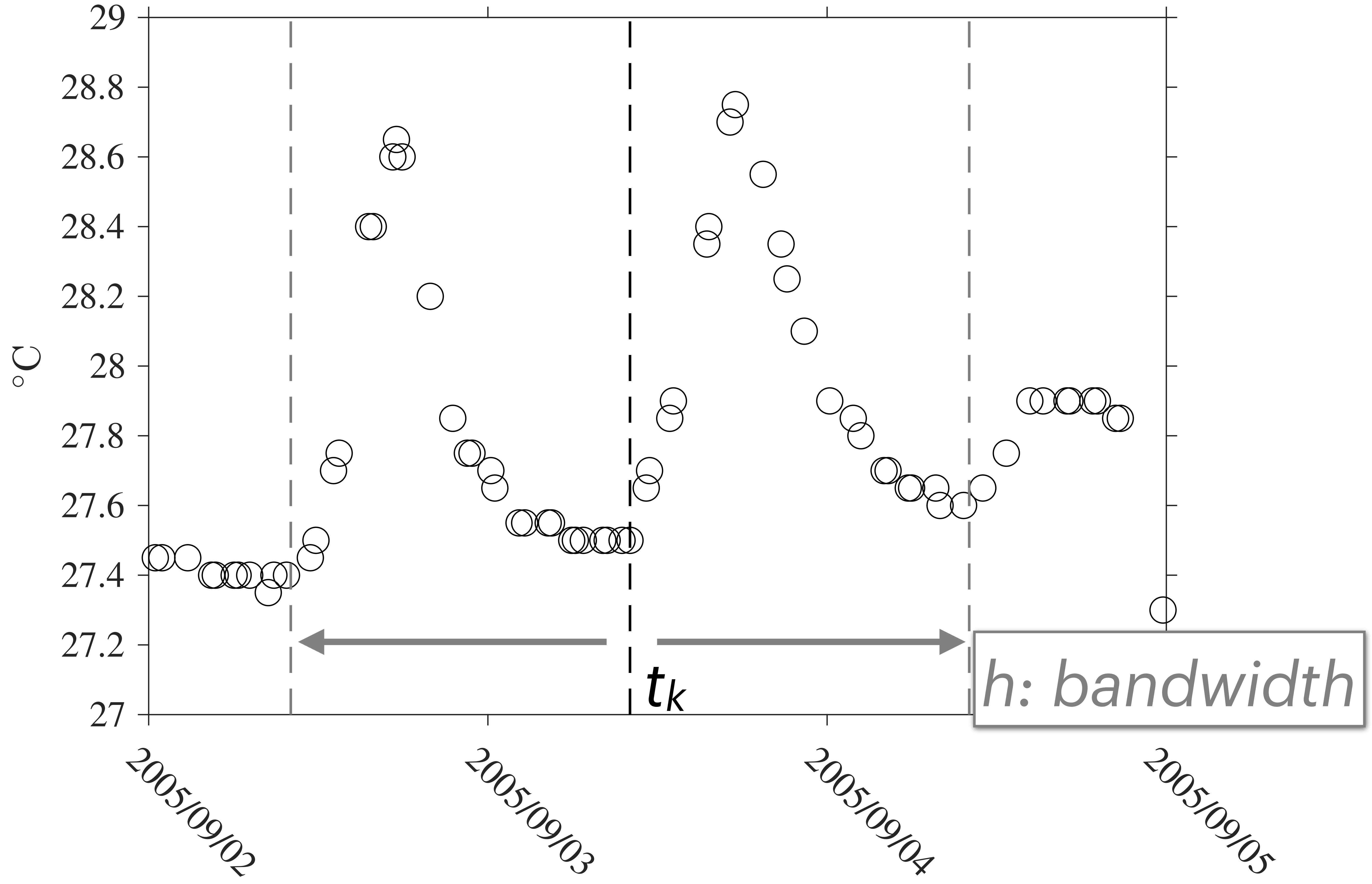
Methodology: models

$$s_i = s_m(t_i; t_k) + \sigma_k \varepsilon_k$$



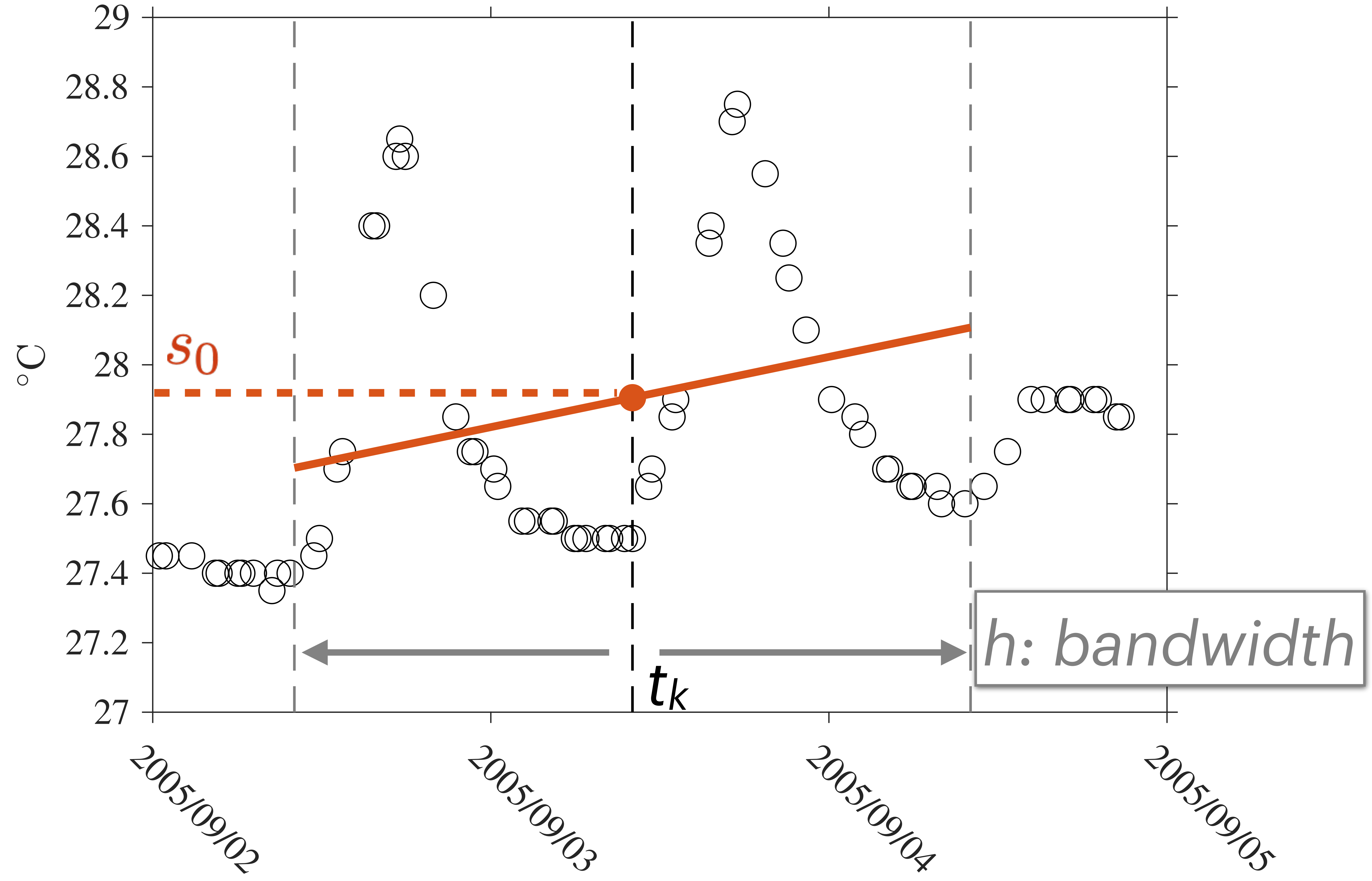
Methodology: models

$$s_i = s_m(t_i; t_k) + \sigma_k \varepsilon_k$$



Methodology: models

$$s_m(t_i; t_k) = s_0 + s_1(t_i - t_k)$$



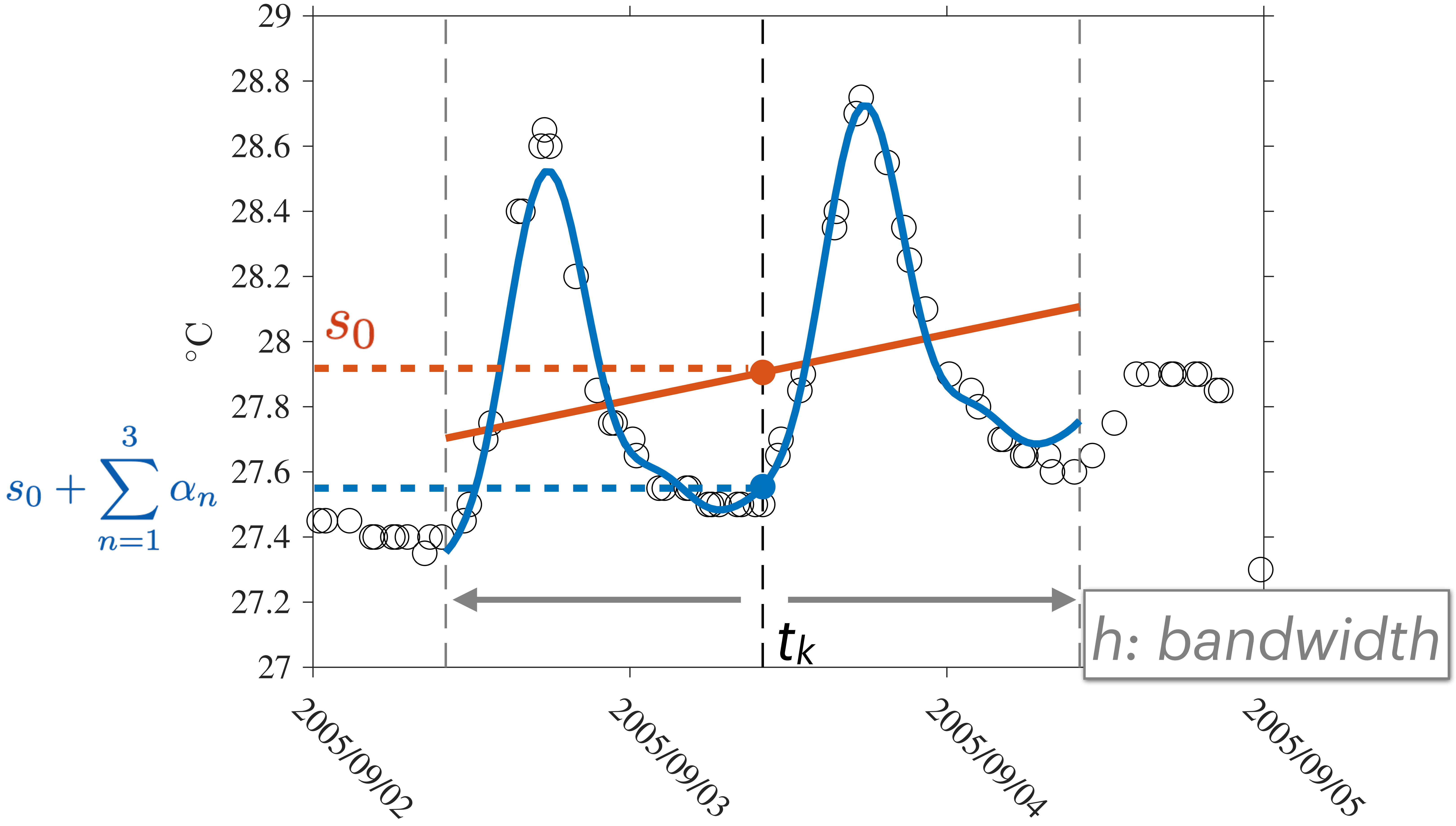
$$s_m(t_k; t_k) = s_0$$

$P = 1$

$N = 0$

Methodology: models

$$s_m(t_i; t_k) = s_{0,k} + s_{1,k}(t_i - t_k) + \sum_{n=1}^{3} [\alpha_{n,k} \cos n\omega(t_i - t_k) + \beta_{n,k} \sin n\omega(t_i - t_k)]$$



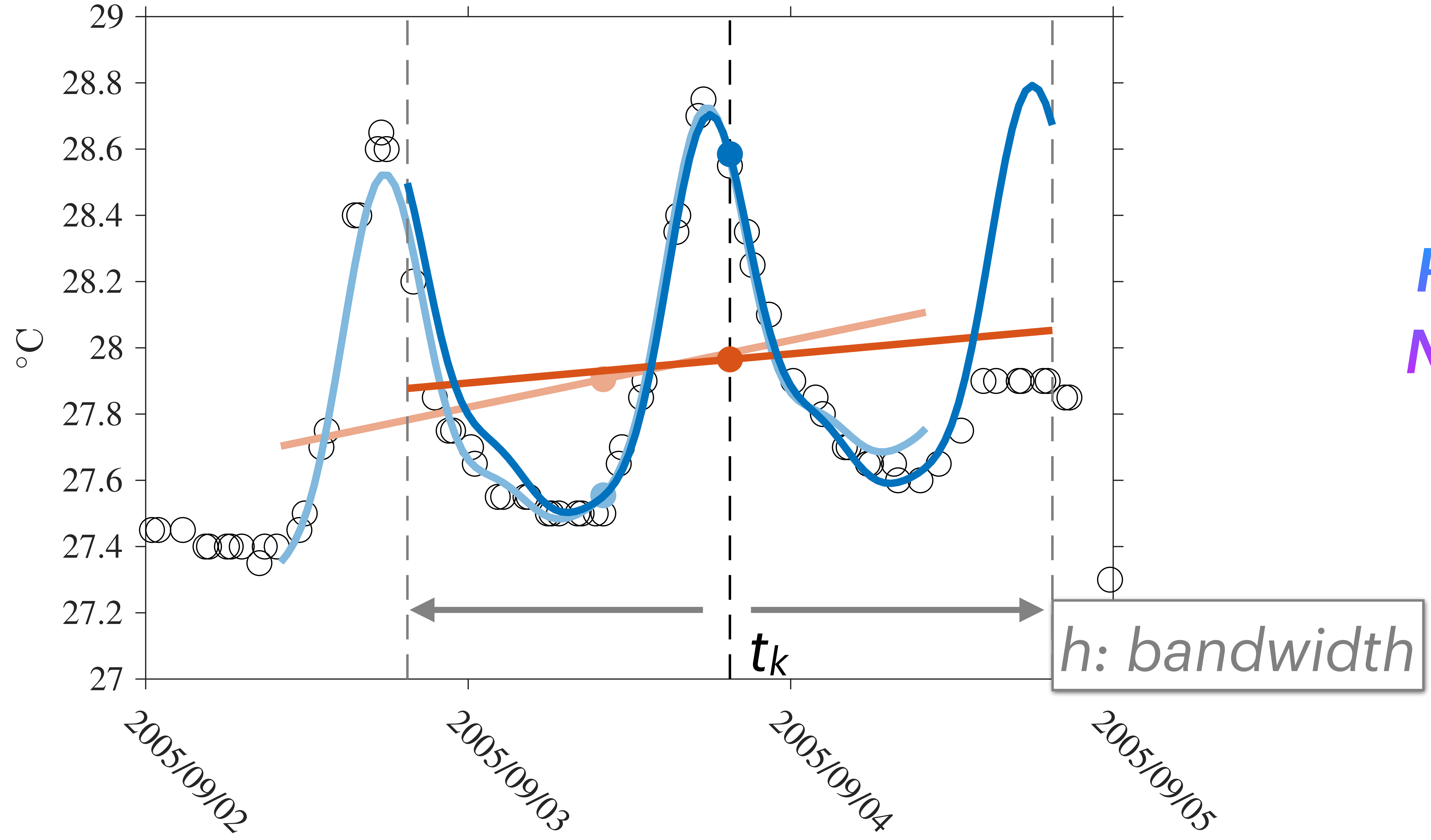
$$s_m(t_k; t_k) = s_0 + \sum_{n=1}^{3} \alpha_n$$

P = 1

N = 3

Methodology: models

$$s_m(t_i;t_k) = s_{0,k} + s_{1,k}(t_i - t_k) + \sum_{n=1}^{3} [\alpha_{n,k} \cos n\omega(t_i - t_k) + \beta_{n,k} \sin n\omega(t_i - t_k)]$$

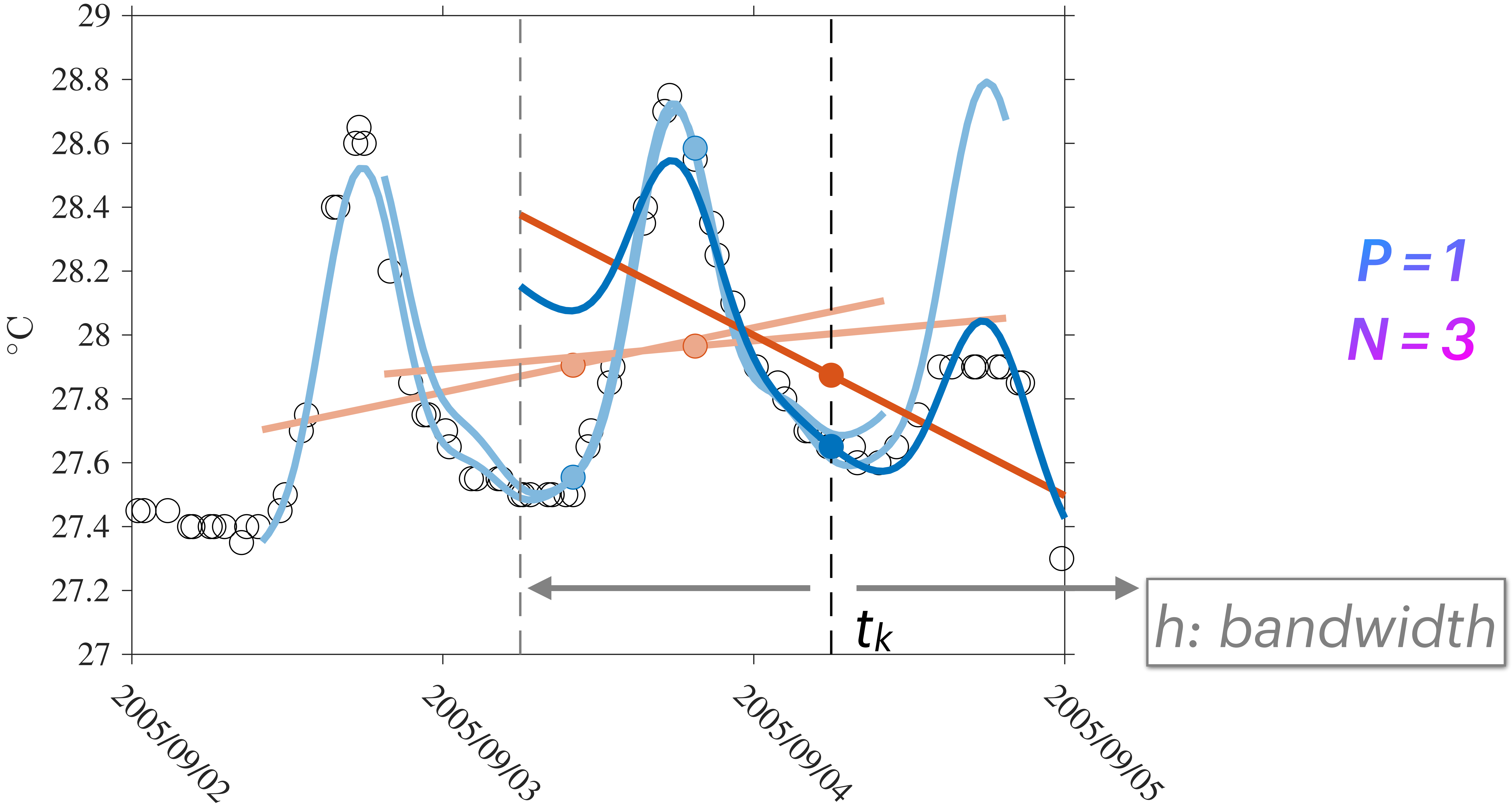


$P = 1$

$N = 3$

Methodology: models

$$s_m(t_i;t_k) = s_{0,k} + s_{1,k}(t_i - t_k) + \sum_{n=1}^{3}\left[\alpha_{n,k}\cos n\omega(t_i - t_k) + \beta_{n,k}\sin n\omega(t_i - t_k)\right]$$



P = 1

N = 3

Methodology: deriving uncertainties

$$s_m(t_i;t_k) = s_{0,k} + s_{1,k}(t_i - t_k) + \sum_{n=1}^{3}[\alpha_{n,k}\cos n\omega(t_i - t_k) + \beta_{n,k}\sin n\omega(t_i - t_k)]$$

Method to fit model to observations is adapted from the LOcally WEighted Scatterplot Smoothing estimator (LOWESS), *Cleveland (1979); essentially an iterative least squares method*

Theoretical formula for the variance of the estimates, i.e squared **standard error** of the estimates:

$$\mathbf{C}_{\boldsymbol{\beta}} \equiv \mathrm{Var}(\widehat{\boldsymbol{\beta}}) = (\mathbf{X}\mathbf{W}^*\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{W}^*\underset{\uparrow}{\boldsymbol{\Sigma}}\mathbf{W}^*\mathbf{X})(\mathbf{X}\mathbf{W}^*\mathbf{X})^{-1}$$

$$\boldsymbol{\beta} = \begin{bmatrix} s_{0,k} \\ \vdots \\ s_{P,k} \\ \alpha_{1,k} \\ \vdots \\ \alpha_{N,k} \\ \beta_{1,k} \\ \vdots \\ \beta_{N,k} \end{bmatrix}$$

Unknown covariance matrix (~22² terms!) of the observation errors from the process model

$$s_i = s_m(t_i;t_k) + \sigma_k\varepsilon_k$$

How to move forward? See Fan & Gijbels (2018) but ...

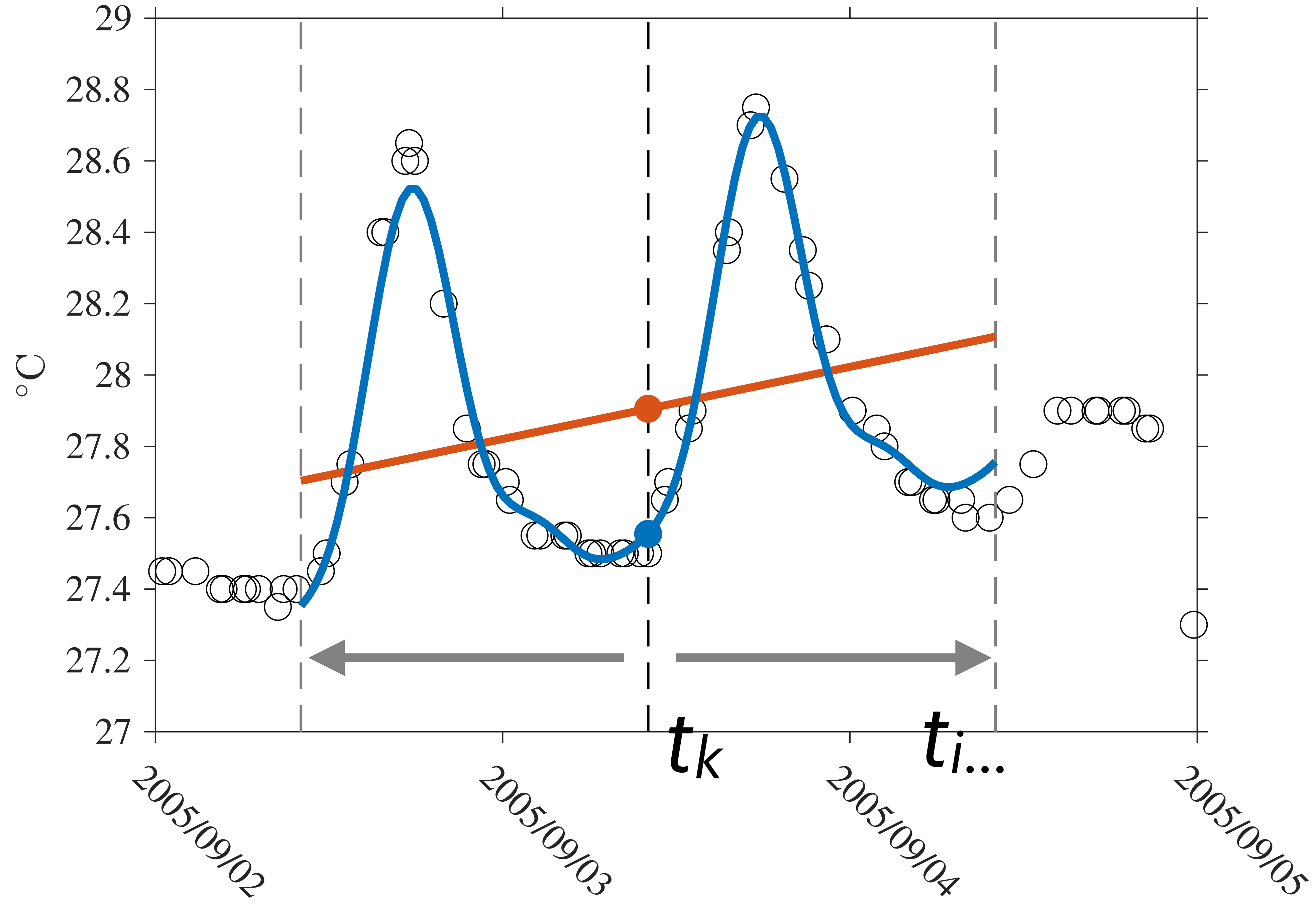
Methodology: deriving uncertainties

$$\mathbf{C}_{\boldsymbol{\beta}} \equiv \mathrm{Var}(\widehat{\boldsymbol{\beta}}) = (\mathbf{X}\mathbf{W}^*\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{W}^*\boldsymbol{\Sigma}\mathbf{W}^*\mathbf{X})(\mathbf{X}\mathbf{W}^*\mathbf{X})^{-1}$$

- Assume **homoscedasticity** and that the errors are independent

$$\longrightarrow \quad \boldsymbol{\Sigma} = \sigma^2(t_k)\mathbf{I} = (\sigma_1^2(t_k) + \sigma_q^2)\mathbf{I}$$

- Estimate **error variance** directly from residuals of the fit



$$\widehat{\sigma}_1^2(t_k) = \frac{(\mathbf{s} - \mathbf{X}\widehat{\boldsymbol{\beta}})^T\mathbf{W}^*(\mathbf{s} - \mathbf{X}\widehat{\boldsymbol{\beta}})}{\mathrm{tr}\left\{\mathbf{W}^* - \mathbf{W}^*\mathbf{X}(\mathbf{X}^T\mathbf{W}^*\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}^*\right\}}$$
$$= \frac{\Sigma_i\left[s_i - \widehat{s}_m(t_i;t_k)\right]^2 \delta_i K_{h_k}(t_i - t_k)}{\nu}$$

effective number of degrees of freedom for the residuals for weighted least squares

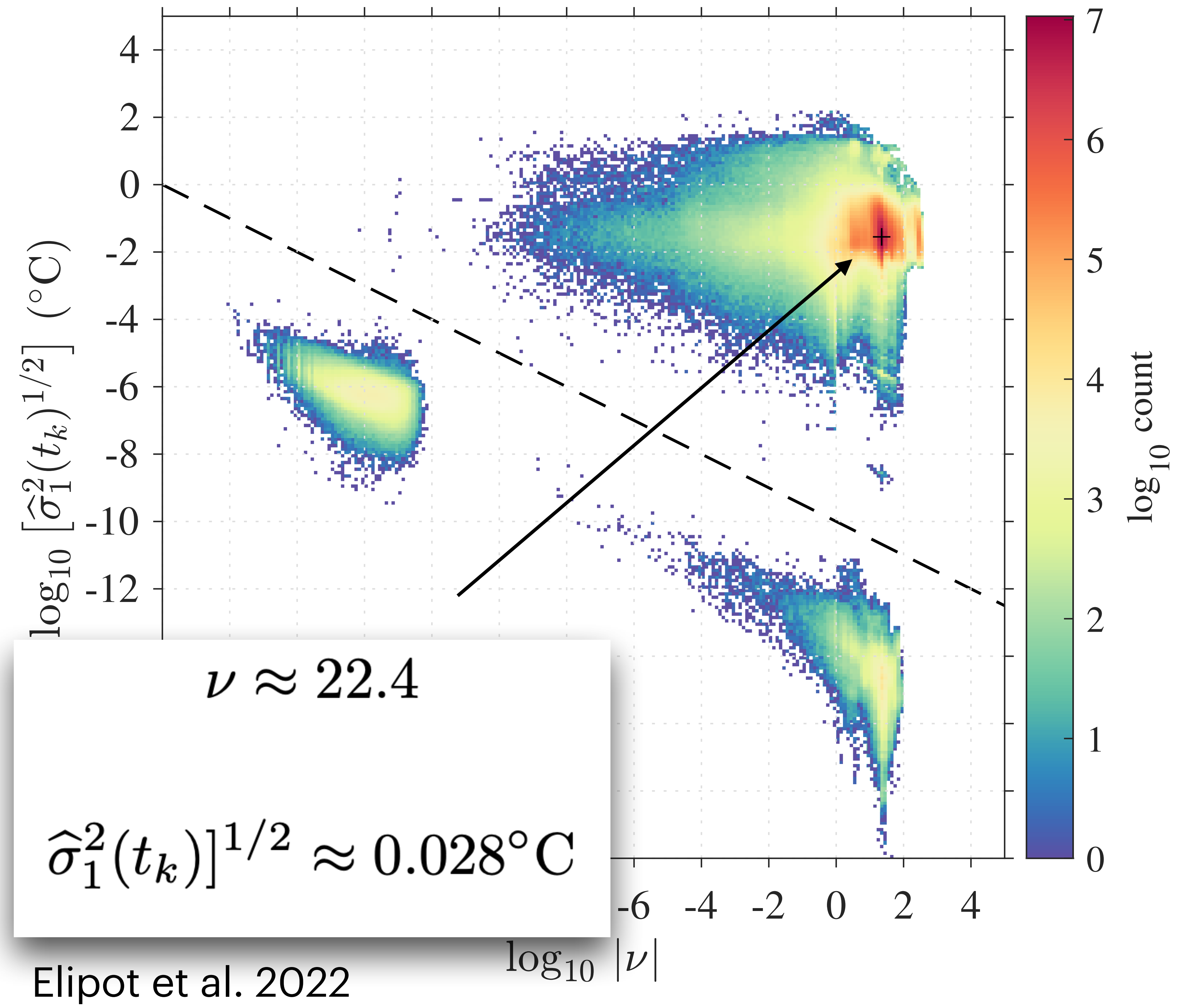
Methodology: deriving uncertainties

$SST = an + b$ SST sensor equation

error variance estimate:

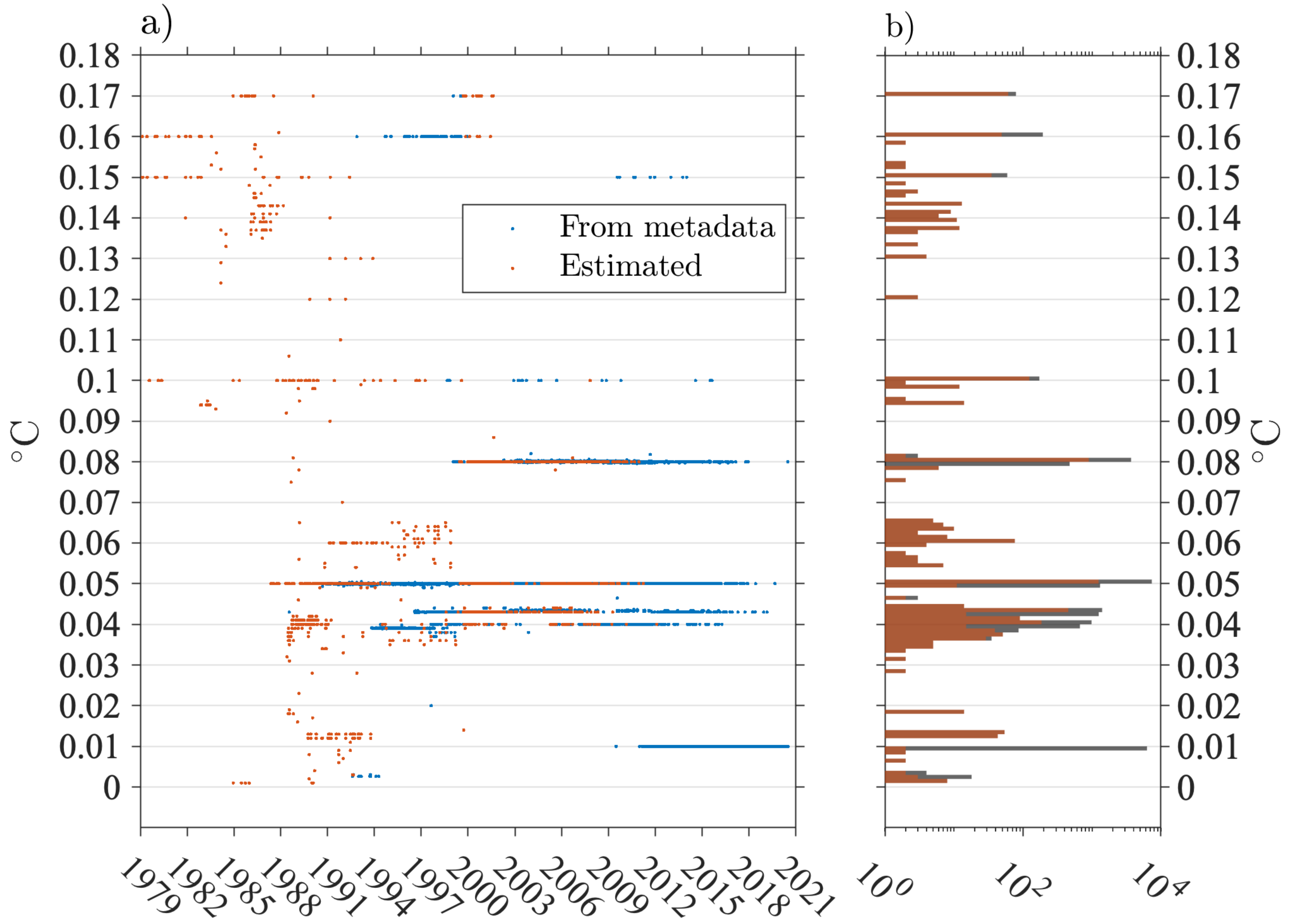
$$\widehat{\sigma}^2(t_k) = \widehat{\sigma}_1^2(t_k) + \sigma_q^2 = \widehat{\sigma}_1^2(t_k) + \frac{a^2}{12}$$

See Chiorboli (2003)



Elipot et al. 2022

a : resolution distributions



Elipot et al. 2022

Methodology: deriving uncertainties

SST estimate:

$$\widehat{s}_{m,k} \equiv s_m(t_k;t_k) = s_{0,k} + \sum_{n=1}^{3} \alpha_{n,k}$$

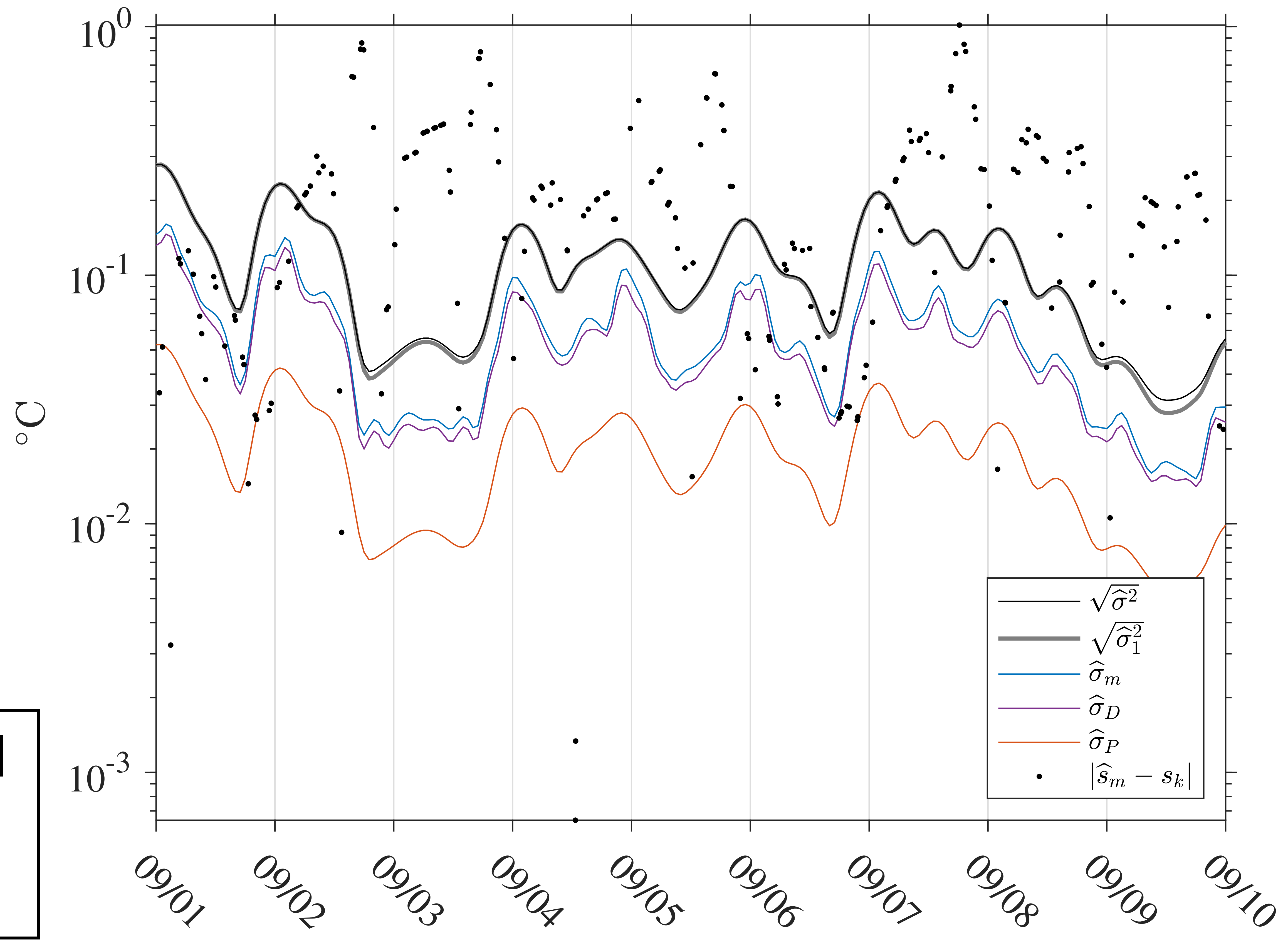
Variance of estimate:

$$\begin{aligned}\sigma_m^2 \equiv \mathrm{Var}\left[s_m(t_k;t_k)\right] &= \mathrm{Var}\left[s_{0,k} + \Sigma_{n=1}^{3}\alpha_{n,k}\right] \\ &= \sigma_P^2 + 2\mathrm{Cov}\left[\left(s_{0,k}\right)\left(\Sigma_{n=1}^{3}\alpha_{n,k}\right)\right] + \sigma_D^2\end{aligned}$$

Variance of low-frequency component (1 term)

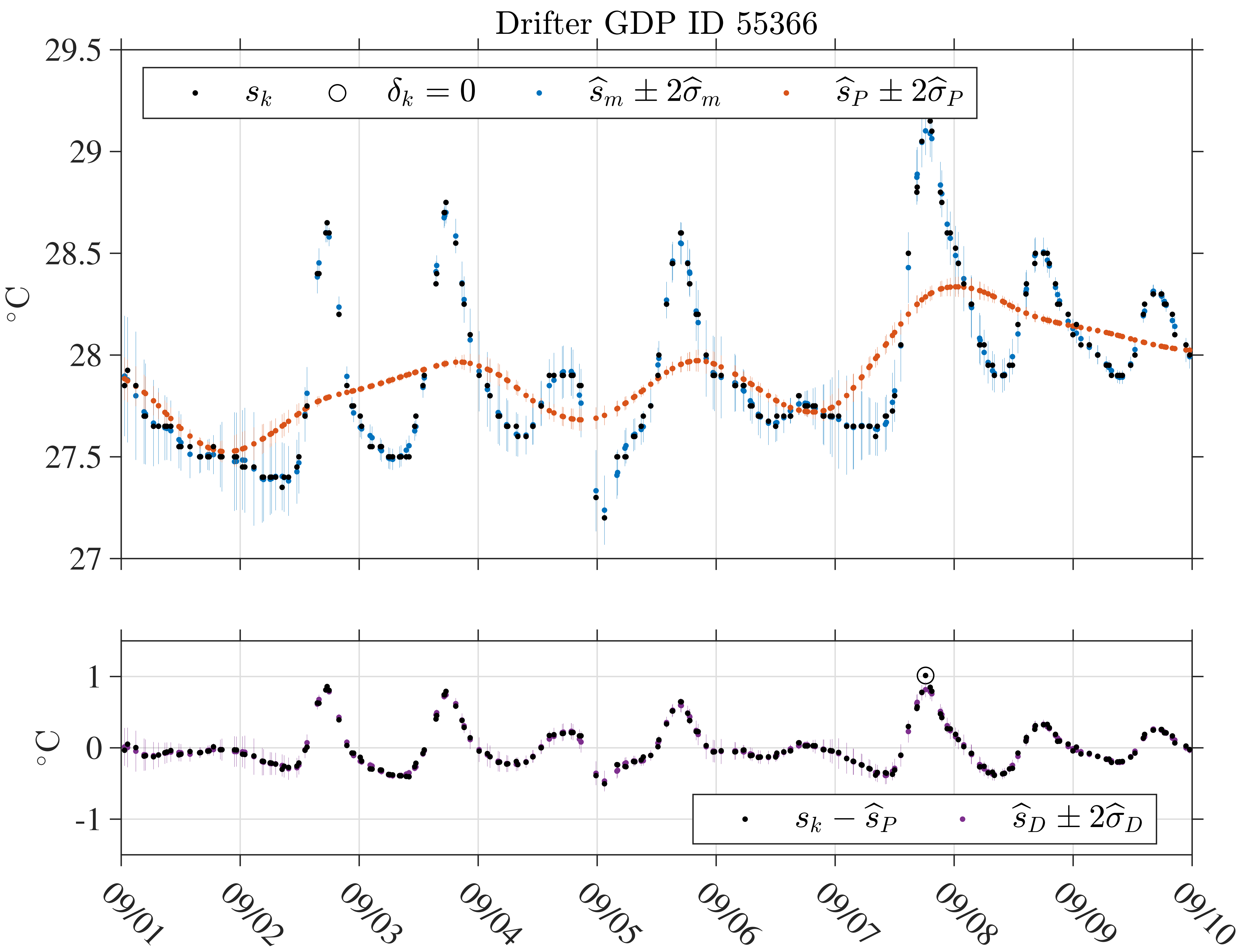
2x3 terms

Variance of diurnal component (3^2 terms)

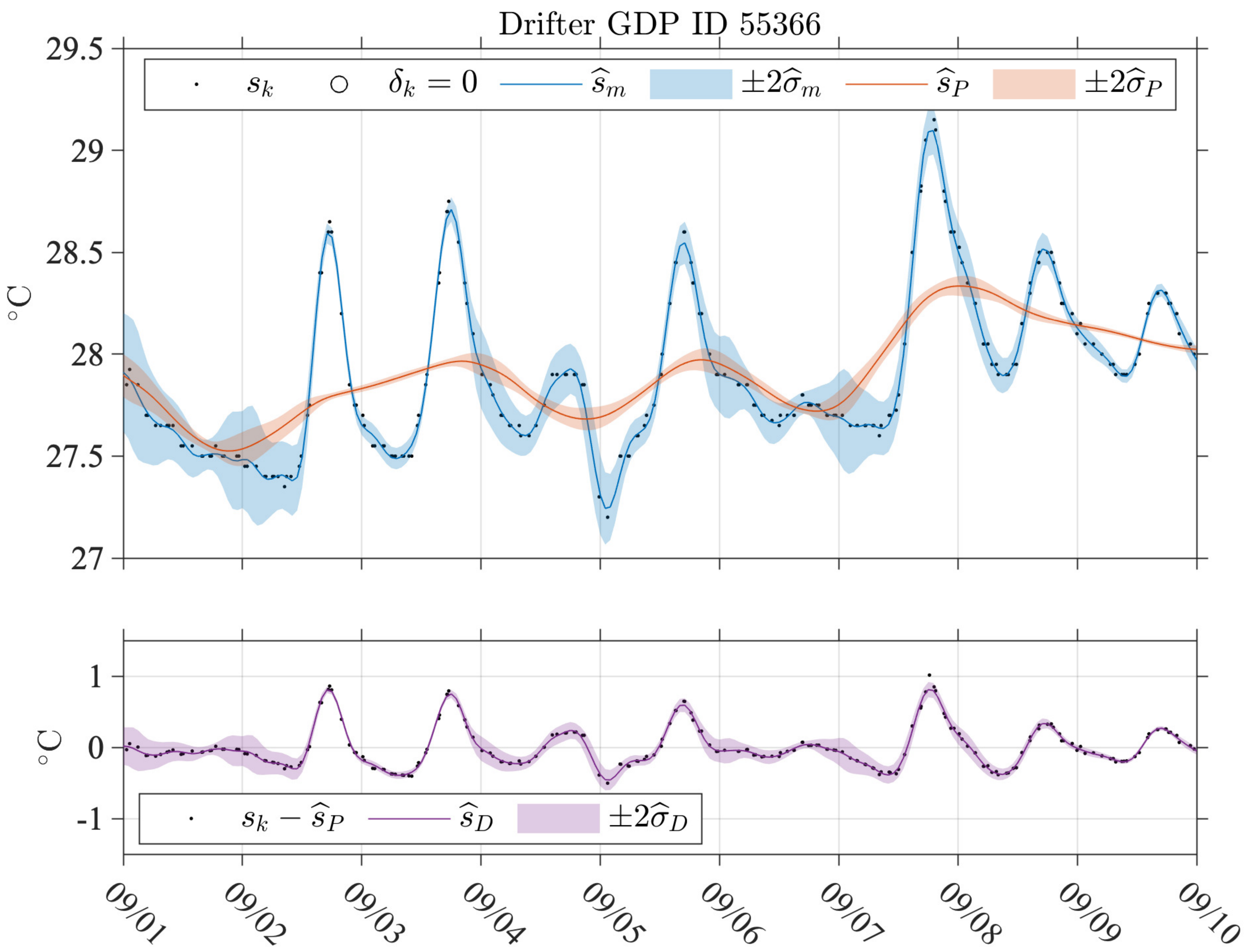


$(N+1)^2$ terms extracted from the error covariance matrix with $(2N+P)^2$ terms

Methodology: deriving uncertainties



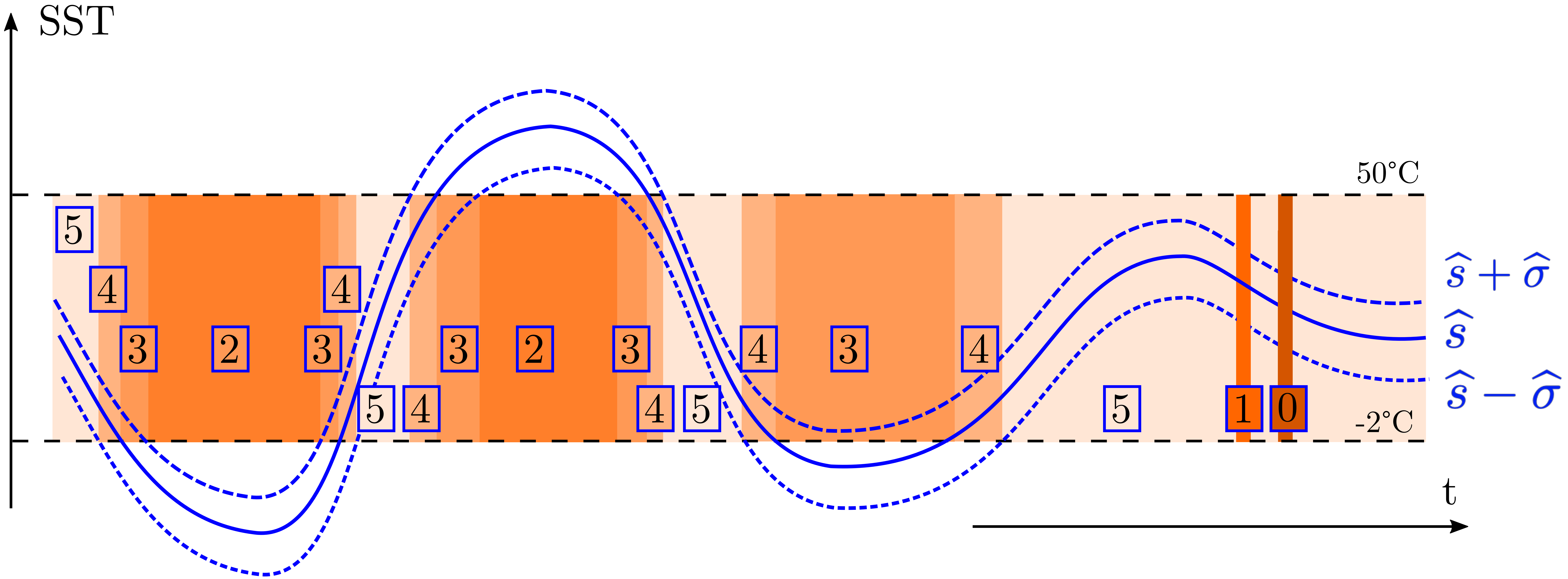
Estimates at observation times



Estimates at hourly times

Methodology: deriving uncertainties

ad hoc derivation of a quality flag based on estimates and uncertainty estimates



Elipot et al. 2022

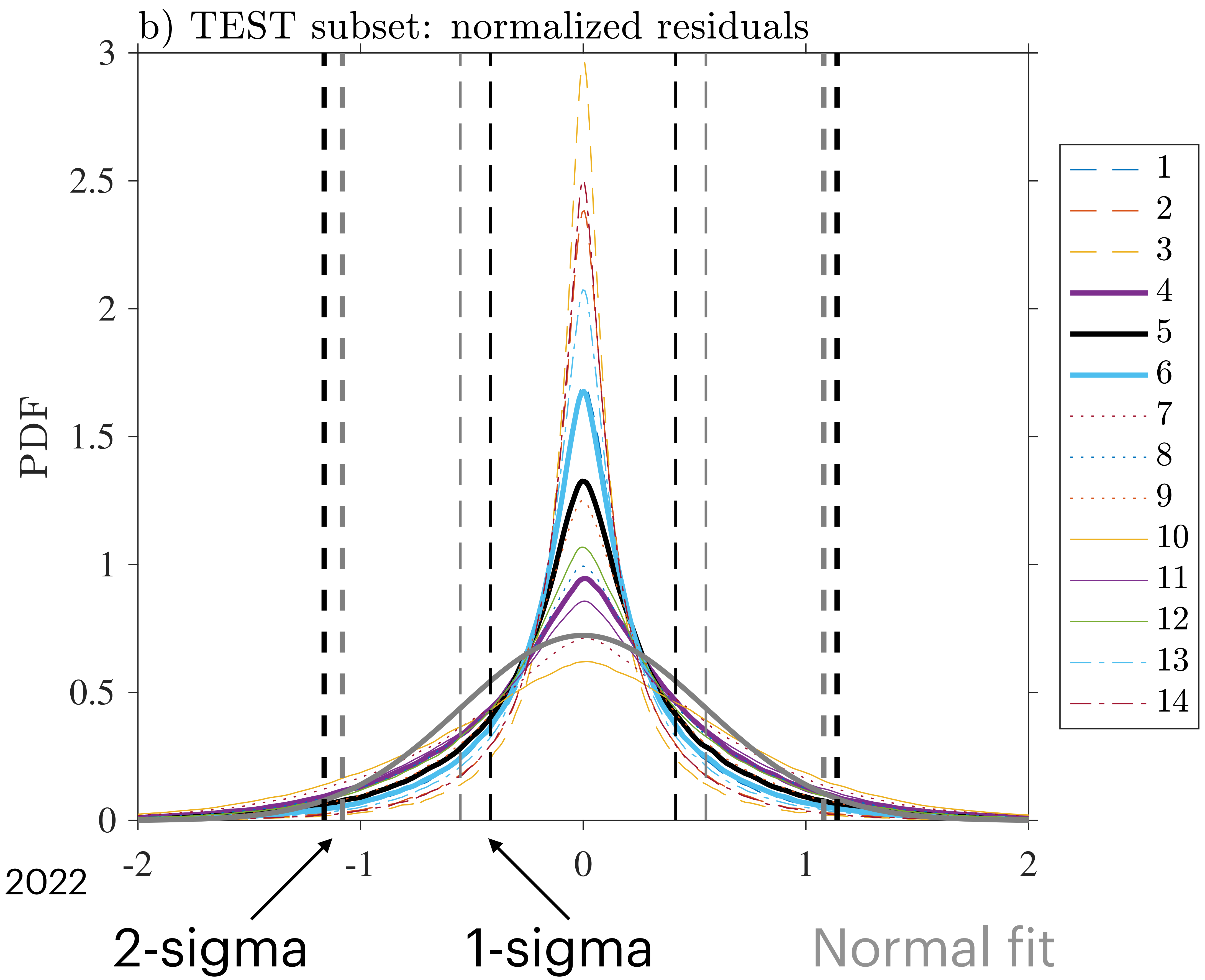
Interpreting uncertainties

process model: $s_i = s_m(t_i; t_k) + \sigma_k \varepsilon_k$ $\longrightarrow$ $\widehat{\varepsilon}_i = \dfrac{s_k - \widehat{s}_{m,k}}{\widehat{\sigma}(t_k)}$

Examine residuals normalized by their estimates of error standard deviations: ε are not normally distributed:

- A least squares method would put too much weight on outliers …
- A standard error therefore represents an interval encompassing more probable values of the true unknown values of a quantity, thus a more conservative confidence interval (78% confidence interval rather than a 68% confidence interval for Normal).
- But 1.96 standard deviation encompasses approximately 94% of the distribution of the residuals, almost like a normal distribution

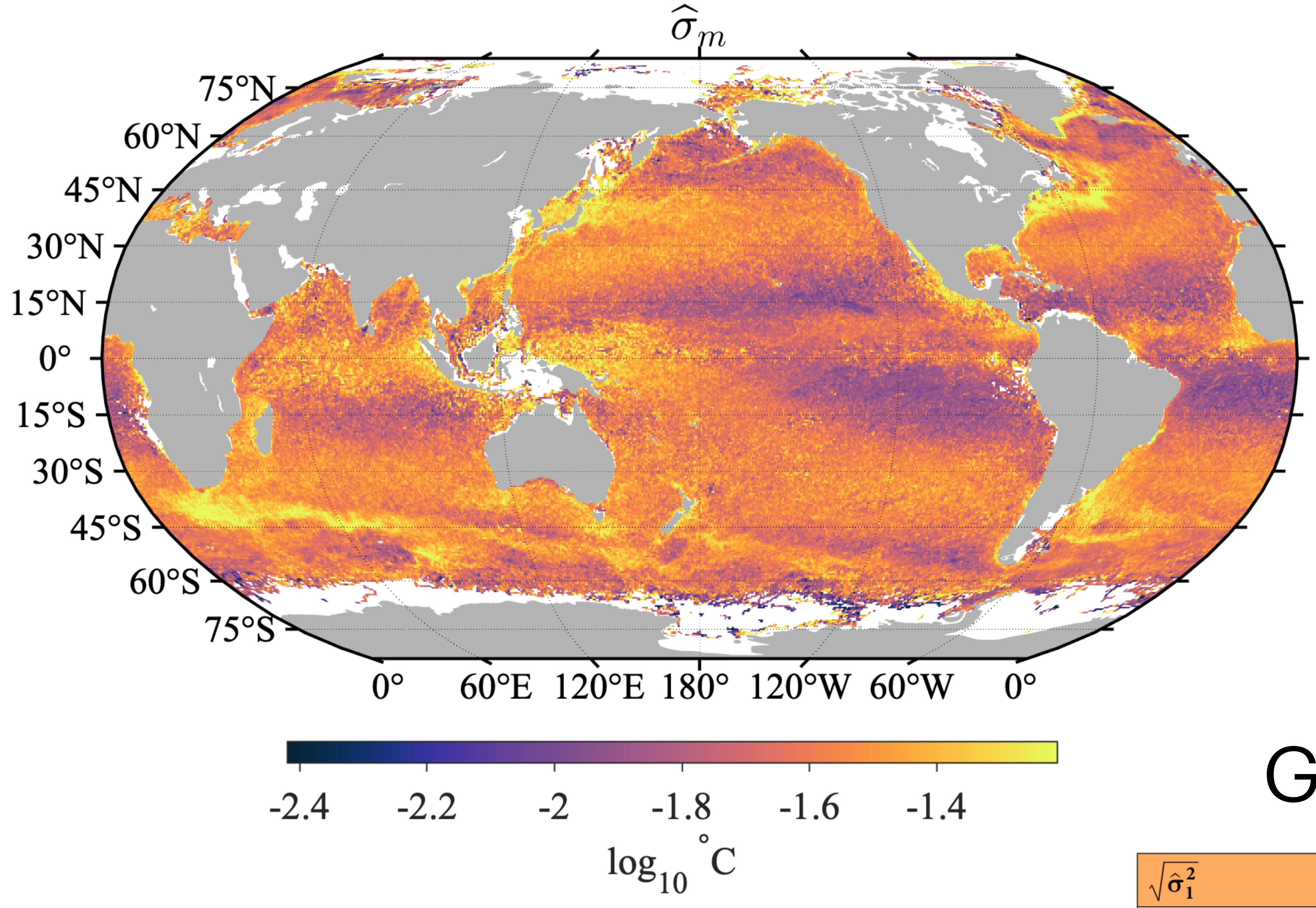
For 12 models tested on a subset:



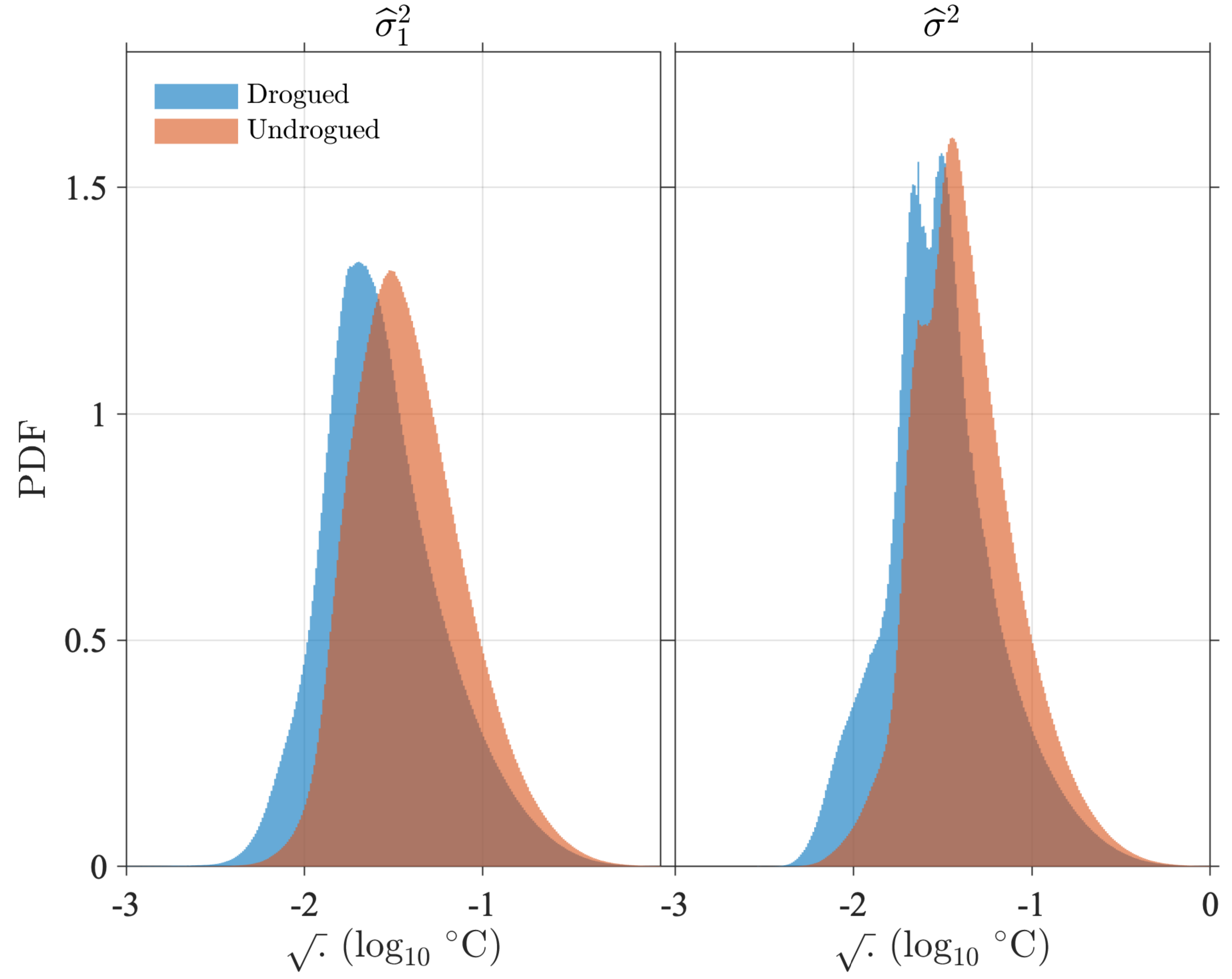
Elipot et al. 2022

Interpreting uncertainties

Spatial distribution



Drogued vs undrogued



Global averages

$\sqrt{\widehat{\sigma}_1^2}$		**All**	**Drogued**	**Undrogued**
Mode	Level-2	0.026	0.020	0.030
	Level-3	**0.026**	**0.020**	**0.030**
50-th percentile	Level-2	0.031	0.025	0.036
	Level-3	**0.031**	**0.025**	**0.036**
$\sqrt{\widehat{\sigma}^2}$		**All**	**Drogued**	**Undrogued**
Mode	Level-2	0.033	0.031	0.036
	Level-2	**0.033**	**0.031**	**0.036**
50-th percentile	Level-2	0.035	0.030	0.040
	Level-3	**0.036**	**0.030**	**0.040**

Elipot et al. 2022

Recap and summary

- Example 1: Characterizing and modeling errors for the purpose of conducting parameter estimation
- Example 2: Deriving uncertainty estimates using a bootstrap method
- Example 3: Deriving and interpreting uncertainty estimates

Thank you! Shane Elipot, selipot@miami.edu

Introduction to afternoon activity

Go to Google slides

Tuesday activity!

- Goal of the activity is to compare two SST products: **The Global Drifter Program hourly SST product** (*Elipot et al. 2022*) and the **The Multi-scale Ultra-high Resolution (MUR) SST analysis** (Chin et al. 2017), a mostly satellite-based product.
- Both products provide **uncertainty estimates** and we would like you to use these estimates to inform the comparisons.
- A preliminary step of the comparison is “matching” the two datasets in space (and time). This has been prepared for you so that you can focus on the statistical and oceanographic aspects of the comparisons

Tuesday activity!

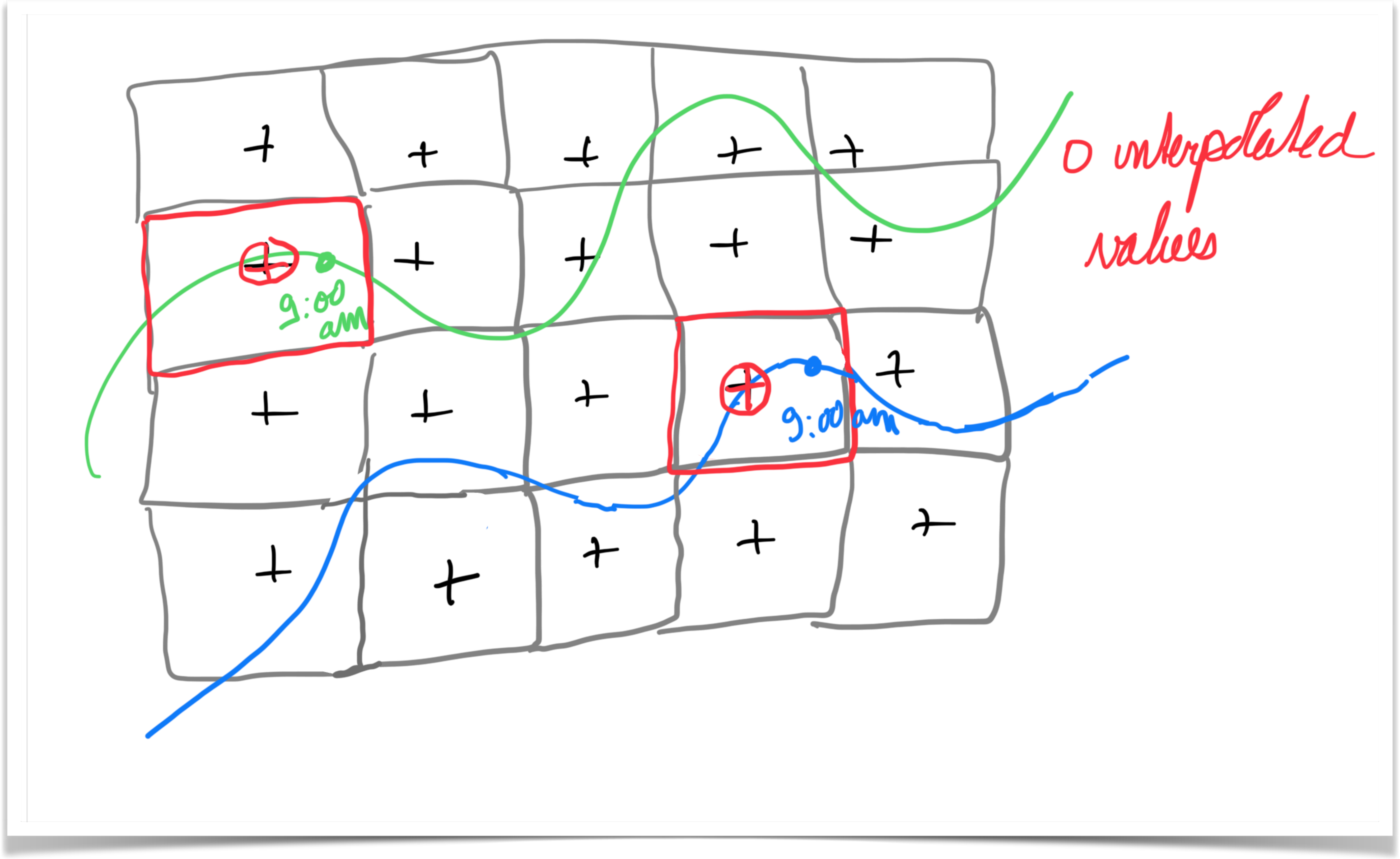
- MUR is available publicly on an AWS S3 bucket but only from 2002 to 2020 (https://registry.opendata.aws/mur/) ….
- Alternatively, it is is available from NASA PO.DACC from 2002 to present (https://podaac.jpl.nasa.gov/dataset/MUR-JPL-L4-GLOB-v4.1). Datasets for you to play with today were prepared from data downloaded from PODAAC …
- Global Drifter Data hourly product is available from AWS S3 bucket https://registry.opendata.aws/noaa-oar-hourly-gdp/

Tuesday activity!

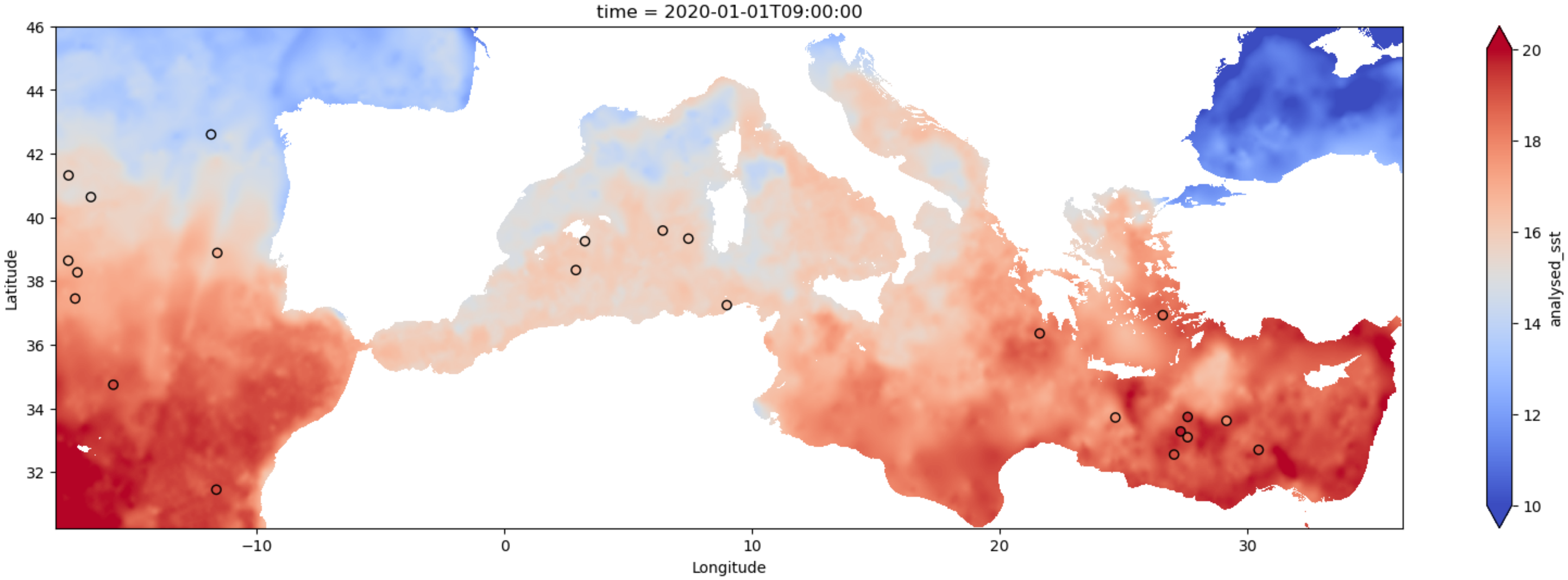
- MUR is the result of the application of the Multi-Resolution Variational Analysis (MRVA) method. This method fits a basis of wavelet functions to multiple satellite and in situ datasets within multiple spatial and temporal windows.
- The method includes a least square estimation of the coefficients multiplying basis functions, the sum of which leads to an estimate of the SST field *T(x,y)* which can be evaluated anywhere and continuously on the globe. The MUR product consists of an estimation of *T(x,y)* on a 0.01 by 0.01 geographical grid, so approximately 1~km scale resolution (at best).
- As stated in Chin et al. (2017), **the standard deviation of the formal estimation error is provided at each grid point as an estimate of analysis uncertainty.** The analysis time is **09:00 UTC.** One global estimate per day is available.

Tuesday activity!

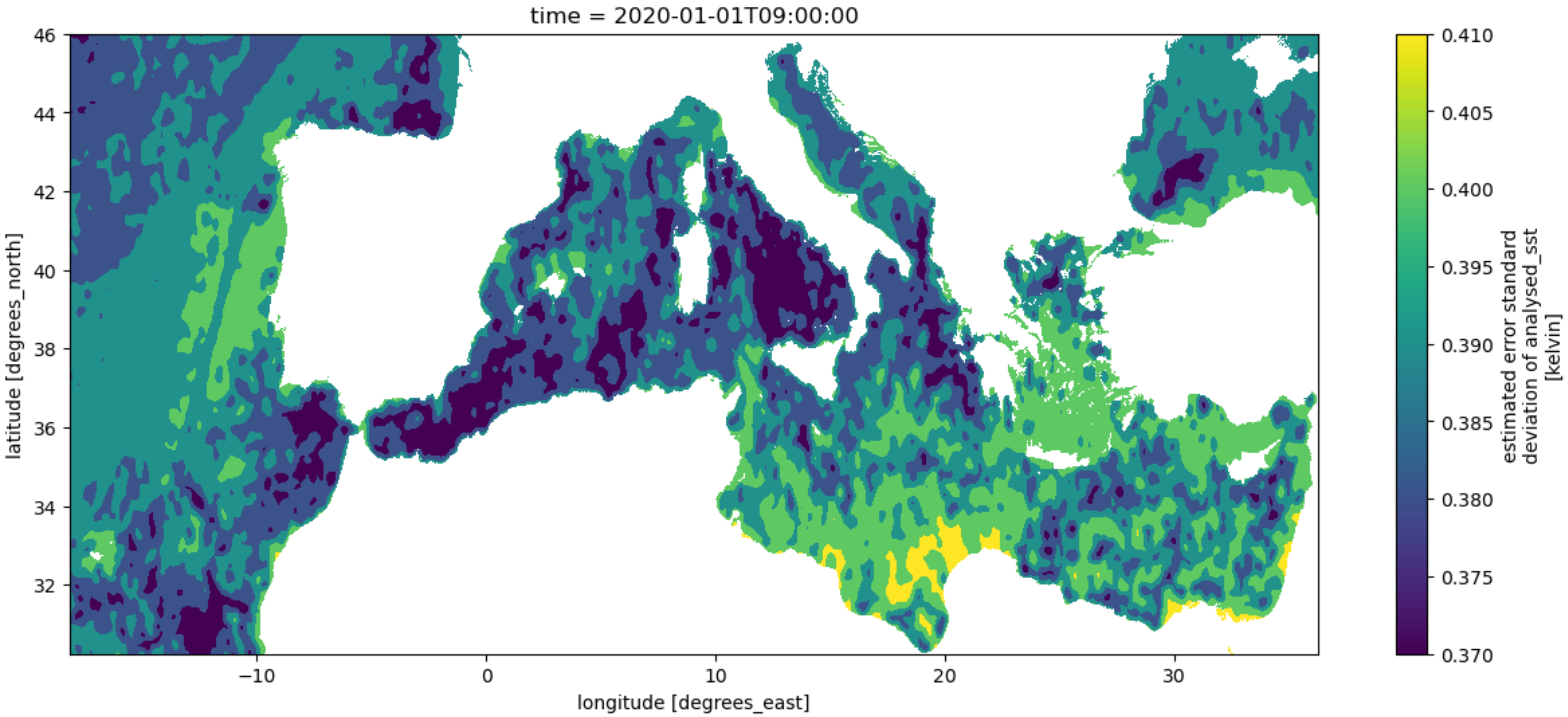
- To prepare a match-up dataset, we matched drifter and MUR data by doing a nearest neighbor interpolation of MUR data (SST estimates and uncertainty estimates) at the 9:00 am drifter locations.
- Limited ourselves to small region in North Atlantic (60-50W, 30-40N)
- **Matched 74,921 9:00am drifter SST estimates from 797 trajectories**
- For bonus activity, also matched 1,799,525 hourly drifter SST estimates from 804 trajectories by expanding to nearest neighbor interpolation in time ...



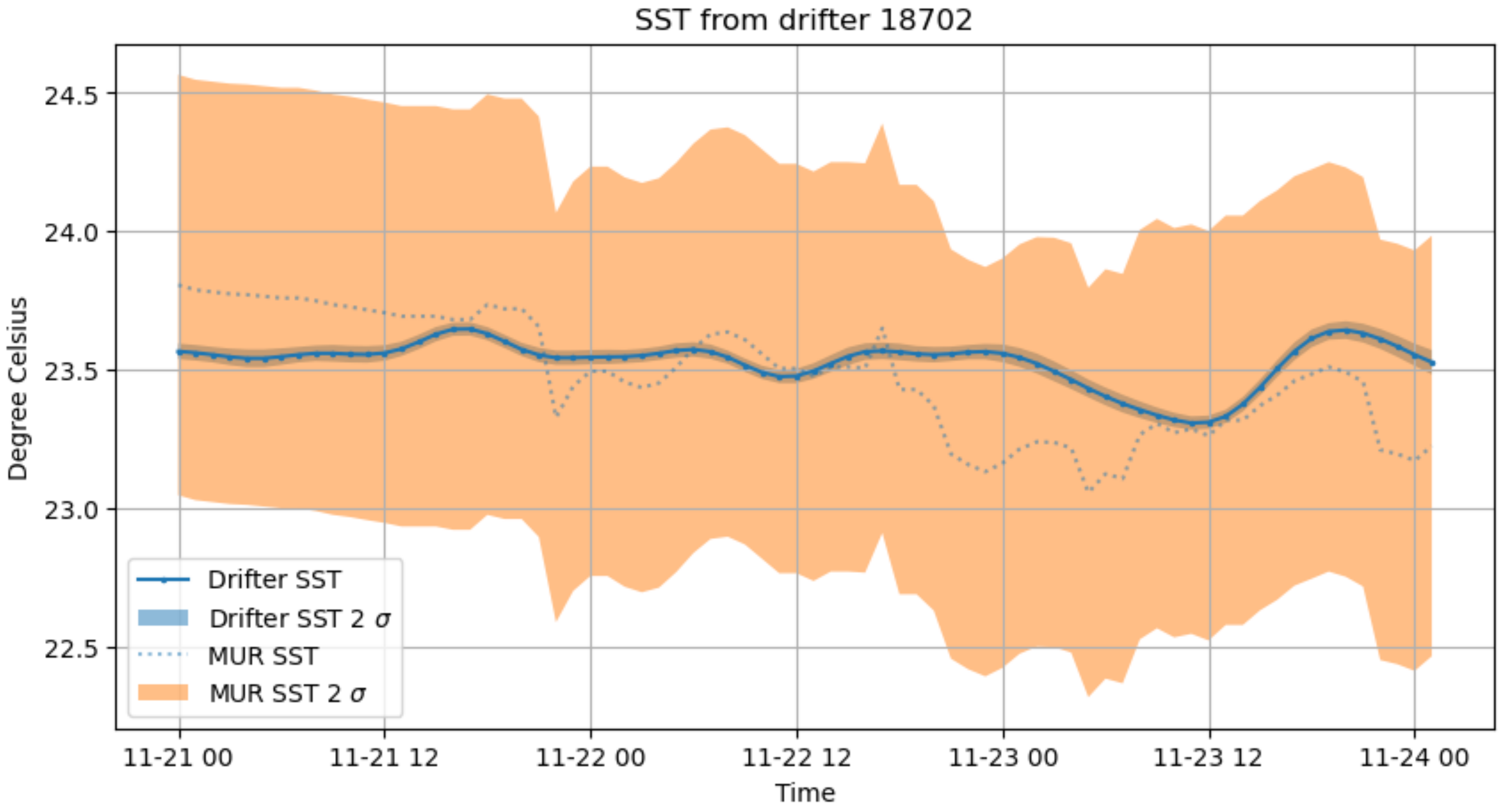
Tuesday activity!



Tuesday activity!



Tuesday activity!



Questions?

Adam S. and Shane E. will be around the classrooms from 1:30pm to 3:30pm

4:00pm we all re-convene in SLAB103 for Science Talk by Mohamed Iskandarani