Inferring Release Characteristics From an Atmospheric Dispersion Model

Bruno Sansó

www.ams.ucsc.edu/~bruno
Department of Applied Mathematics and Statistics
University of California Santa Cruz

The Cast

The work presented in this talk was done in collaboration with

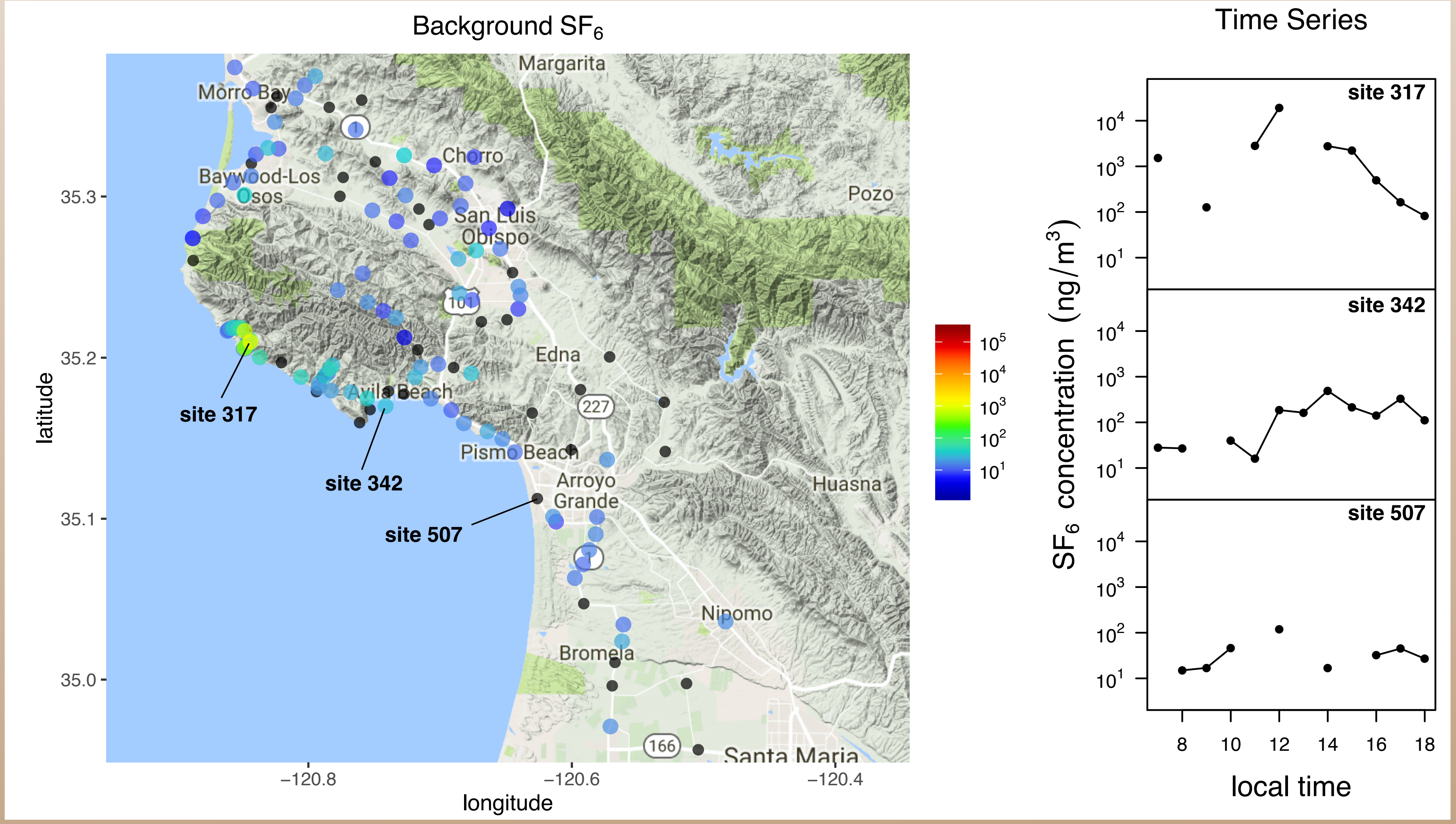
Devin Francom, LANL

Vera Bulaevskaya, TCC

Donald Lucas, LLNL

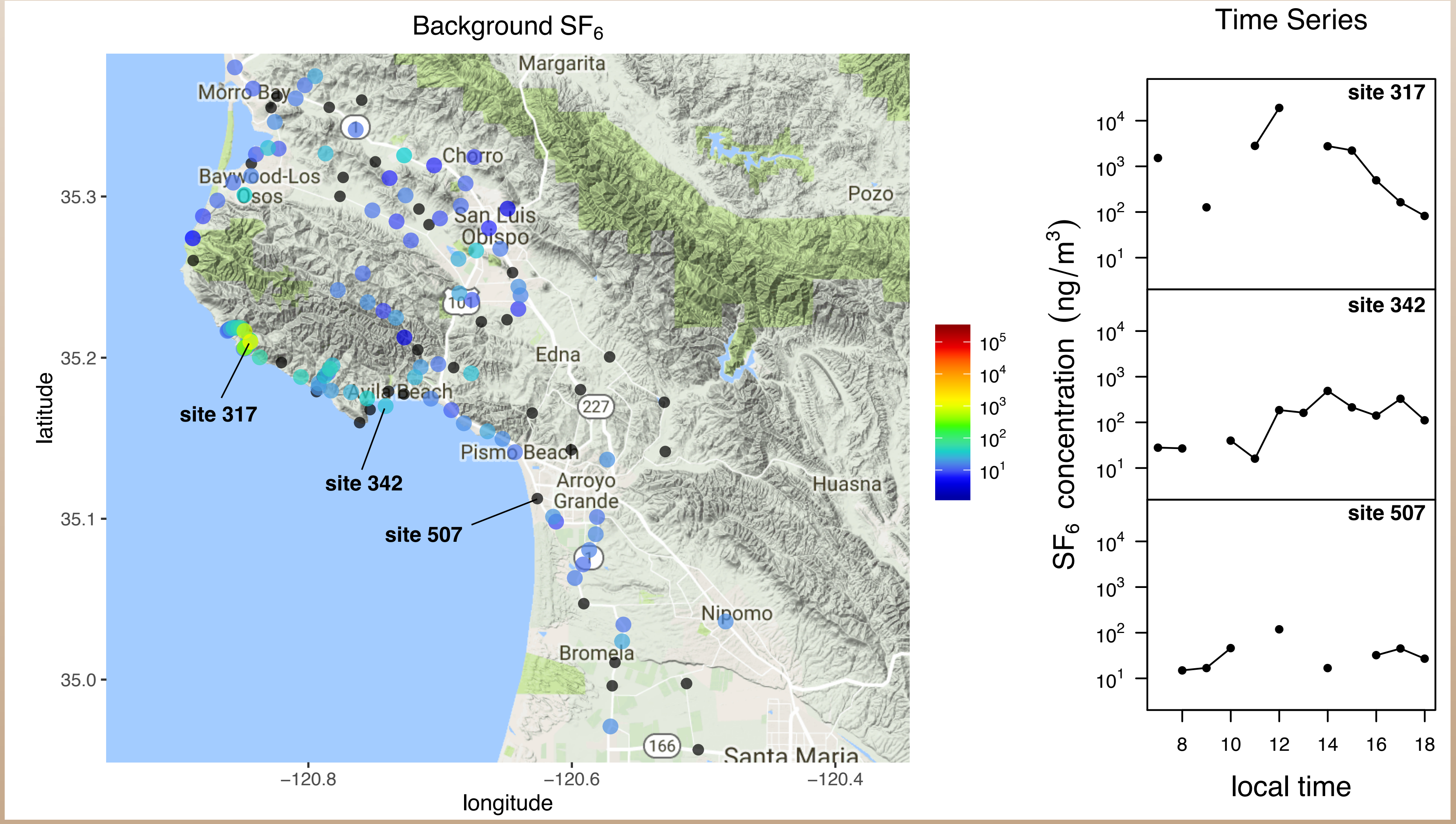
Matthew Simpson, LLNL

Diablo Canyon Nuclear Power Plant



Diablo Canyon Nuclear Power Plant

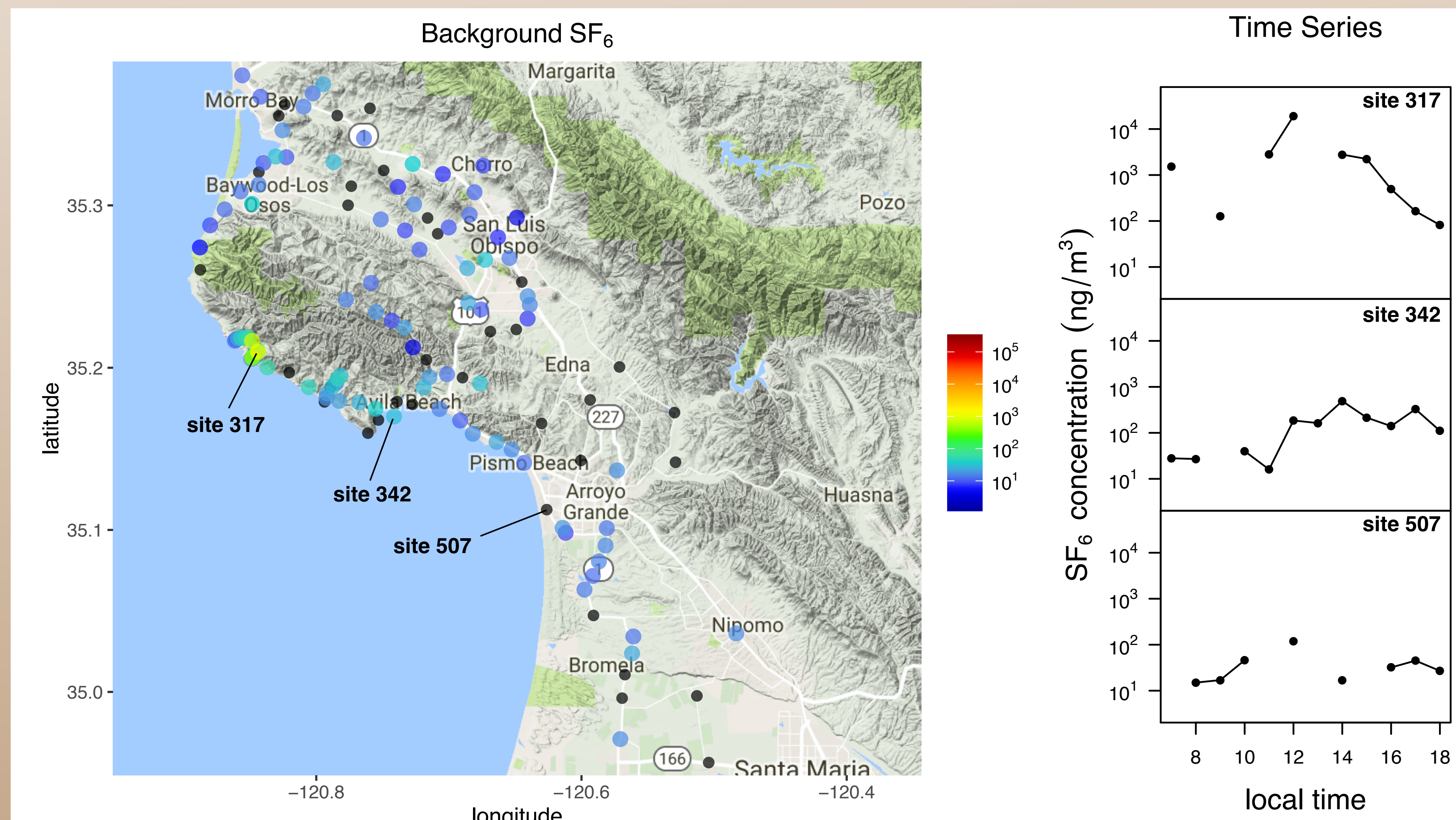
146 Kilos of SF_6 where released from the Diablo Canyon plant on Sept. 4, 1986 for 8 hours, starting at 8:00.



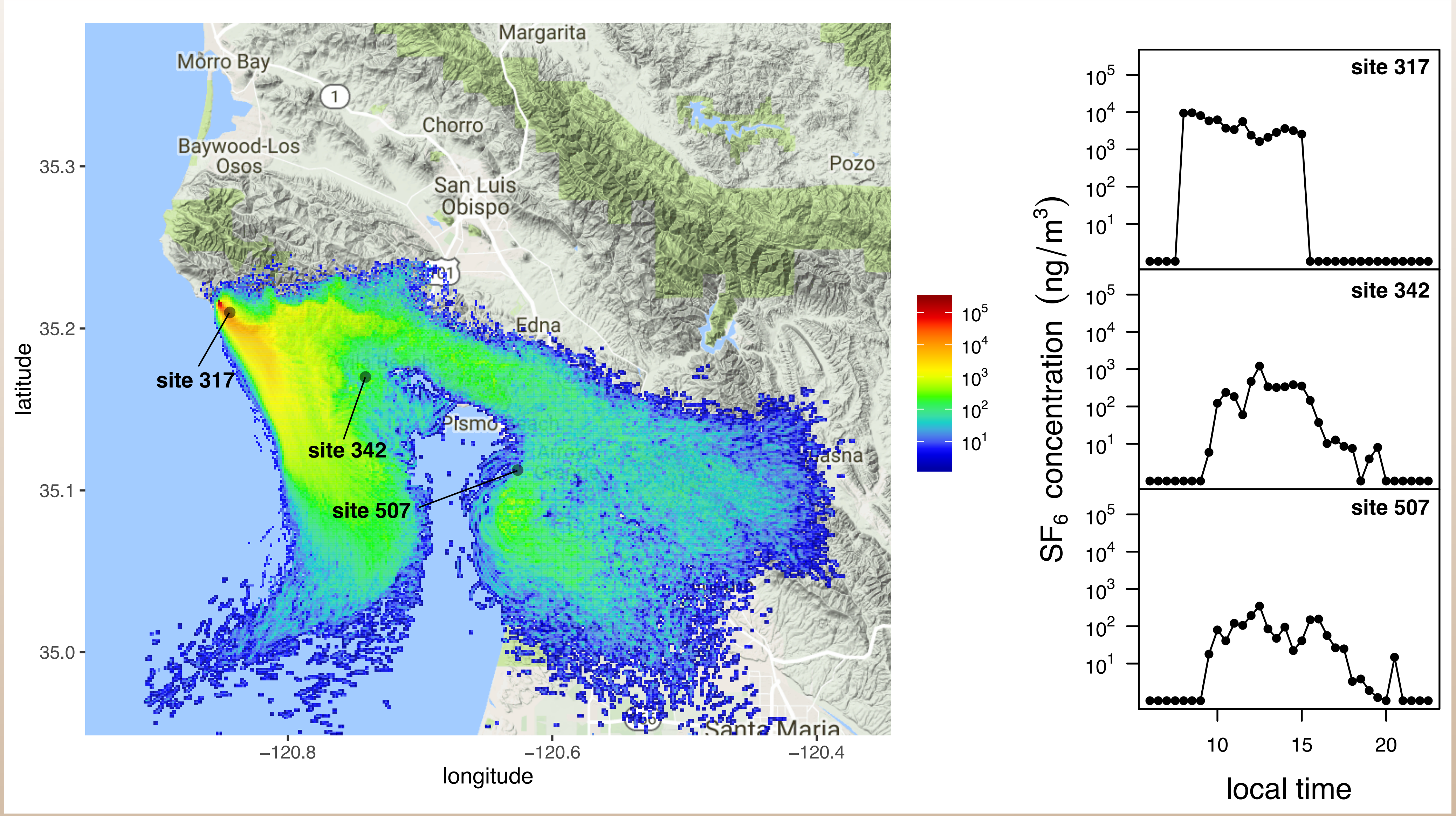
Diablo Canyon Nuclear Power Plant

146 Kilos of SF_6 where released from the Diablo Canyon plant on Sept. 4, 1986 for 8 hours, starting at 8:00.

Air samples were obtained from 7:00 to 18:00 at 150 sites. 24% are missing.

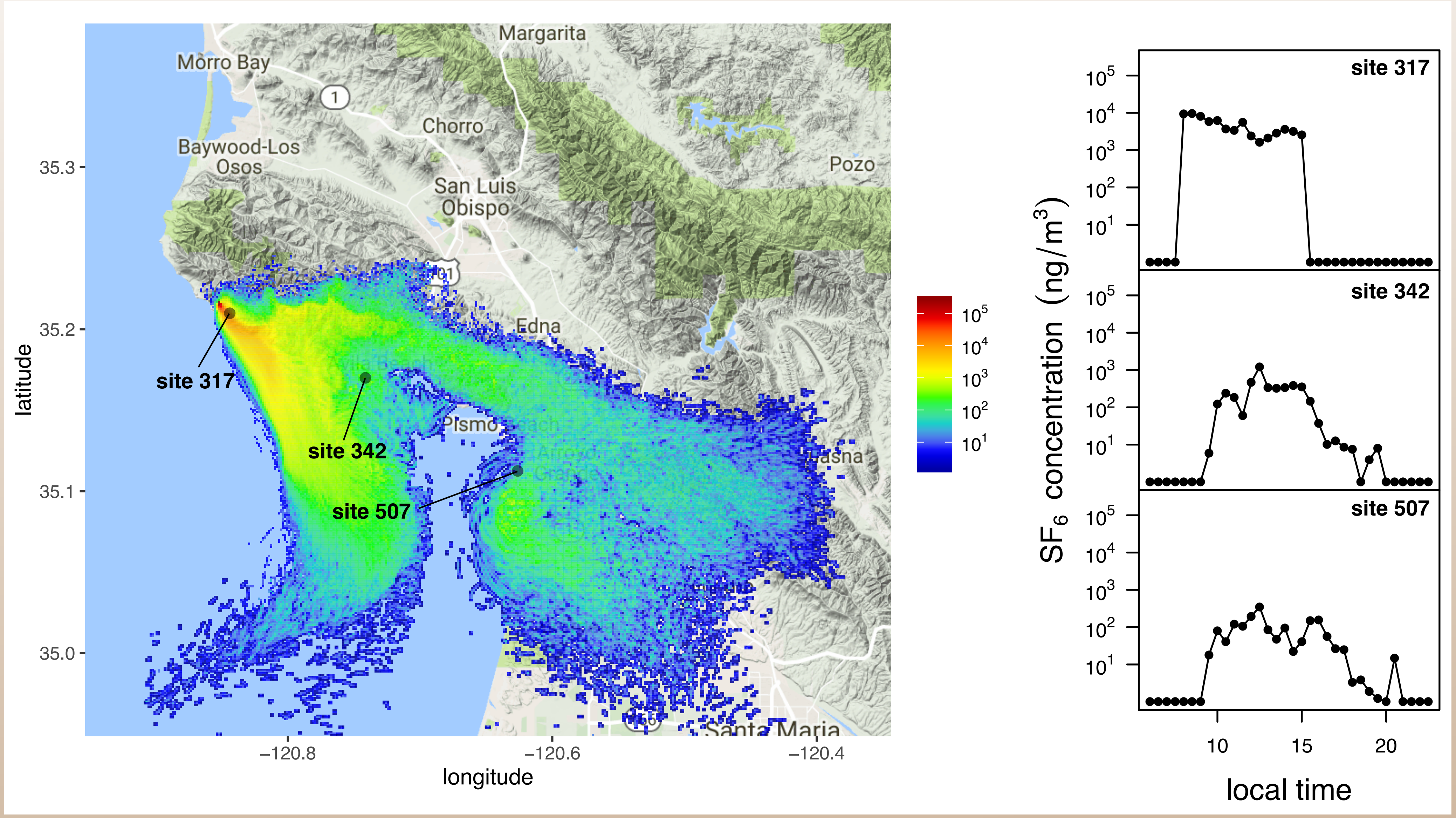


FLEXPART Simulations



The FLEXPART simulator is used to explore the dispersion of the release.

FLEXPART Simulations



The FLEXPART simulator is used to explore the dispersion of the release.

18,000 different combinations of the 11 input parameters of FLEXPART are sampled from a latin hypercube. These result in 18,000 plumes varying in space and time.

Input Parameters

Continuous Input Parameters

Input	Lower Bound	Upper Bound	True Value
Latitude	35.1977	35.2250	35.2111
Longitude	-120.87	-120.83	-120.8543
Altitude	1	10	2
Start Time	7:00	9:00	8:00
Duration	6	10	8
Amount	10	1000	146.016

Categorical Inputs

Input	Number of values
Pre-release Initialization time	2
Boundary Layer Model	3
Nudging	3
Reanalysis	3
Land Model	3

There are five nested domains for WRF models. Each combination of the five categorical variables produces a different wind field at 300 meters resolution.

Emulator

We build an emulator for the computer output corresponding to location s, time t and input values x, by using the representation on empirical orthogonal functions

$$y^c(s,t,x) = \sum_{i=1}^{k} K_i(s,t) w_i(x) + u(s,t)$$

The EOFs are calculated from the model runs

Emulator

We build an emulator for the computer output corresponding to location s, time t and input values x, by using the representation on empirical orthogonal functions

$$y^c(s,t,x) = \sum_{i=1}^{k} K_i(s,t)w_i(x) + u(s,t)$$

The EOFs are calculated from the model runs

For the truncation we take a non-Gaussian error

$$u(s,t) \sim Unif[y^c(s,t,x_j) - \sum_{i=1}^{k} K_i(s,t)w_i(x_j),\; j = 1,\ldots,n_x]$$

This is important in order to propagate truncation uncertainty

Estimating the EOF coefficients

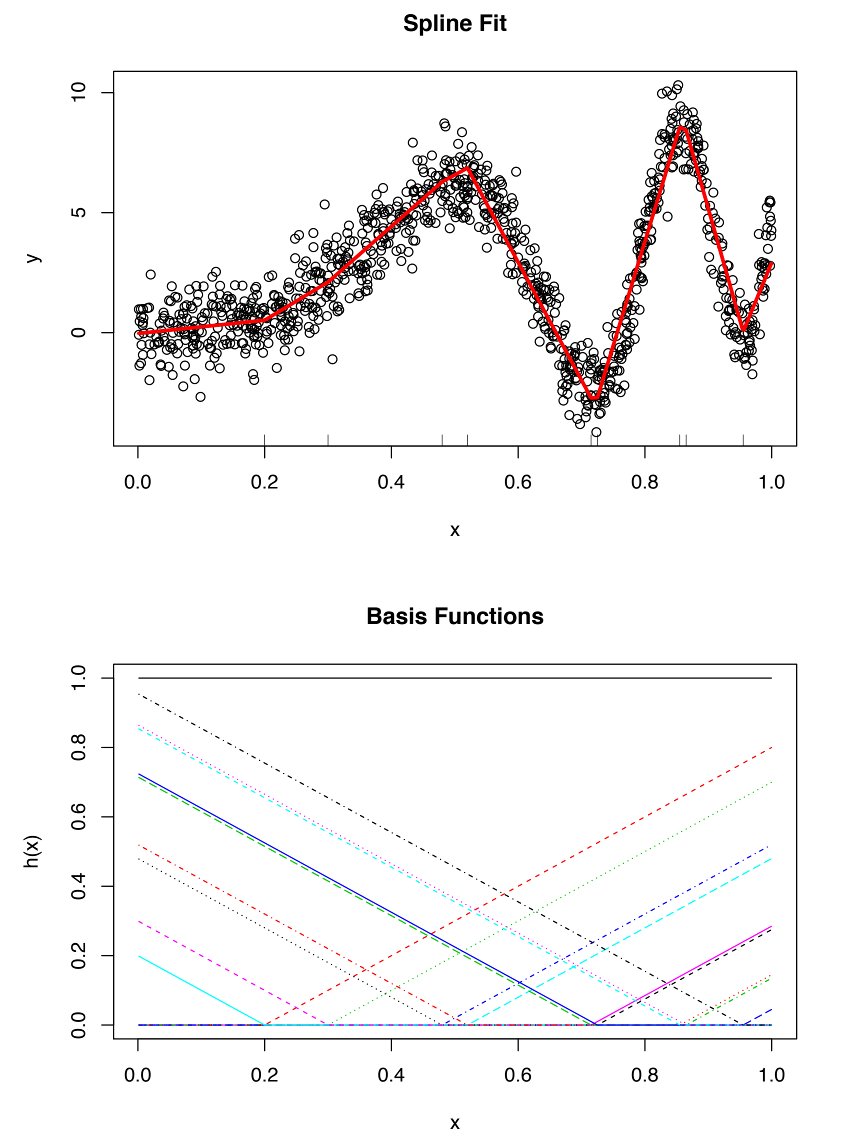
To estimate the coefficients in the EOF we set:

$$w_i(x) = \eta_i(x) + \epsilon_i$$

where

$$\eta(x) = a_0 + \sum_{m=1}^{M} a_m B_m(x)$$

is a representation on **adaptive spline** basis composed of M (unknown) terms, that uses products of hockey sticks with varying signs, number of interactions, and unknown knots.



Categorical an Continuous Inputs

Our application requires the emulator to handle continuous and categorical inputs. Assume that x_1 and x_2 are continuous, and x_3 and x_4 are categorical, then

$$B(x) = [s_1(x_1 - t_1)]_+^\alpha [s_2(x_2 - t_2)]_+^\alpha \mathbf{1}_{x_3 \in C_3} \mathbf{1}_{x_4 \in C_4}$$

where t_1 and t_2 are the knots and

$$s_i = \pm 1$$

and C_i corresponds to one or more categories of the i-th variable.

Categorical an Continuous Inputs

Our application requires the emulator to handle continuous and categorical inputs. Assume that x_1 and x_2 are continuous, and x_3 and x_4 are categorical, then

$$B(x) = [s_1(x_1 - t_1)]_+^\alpha [s_2(x_2 - t_2)]_+^\alpha \mathbf{1}_{x_3 \in C_3} \mathbf{1}_{x_4 \in C_4}$$

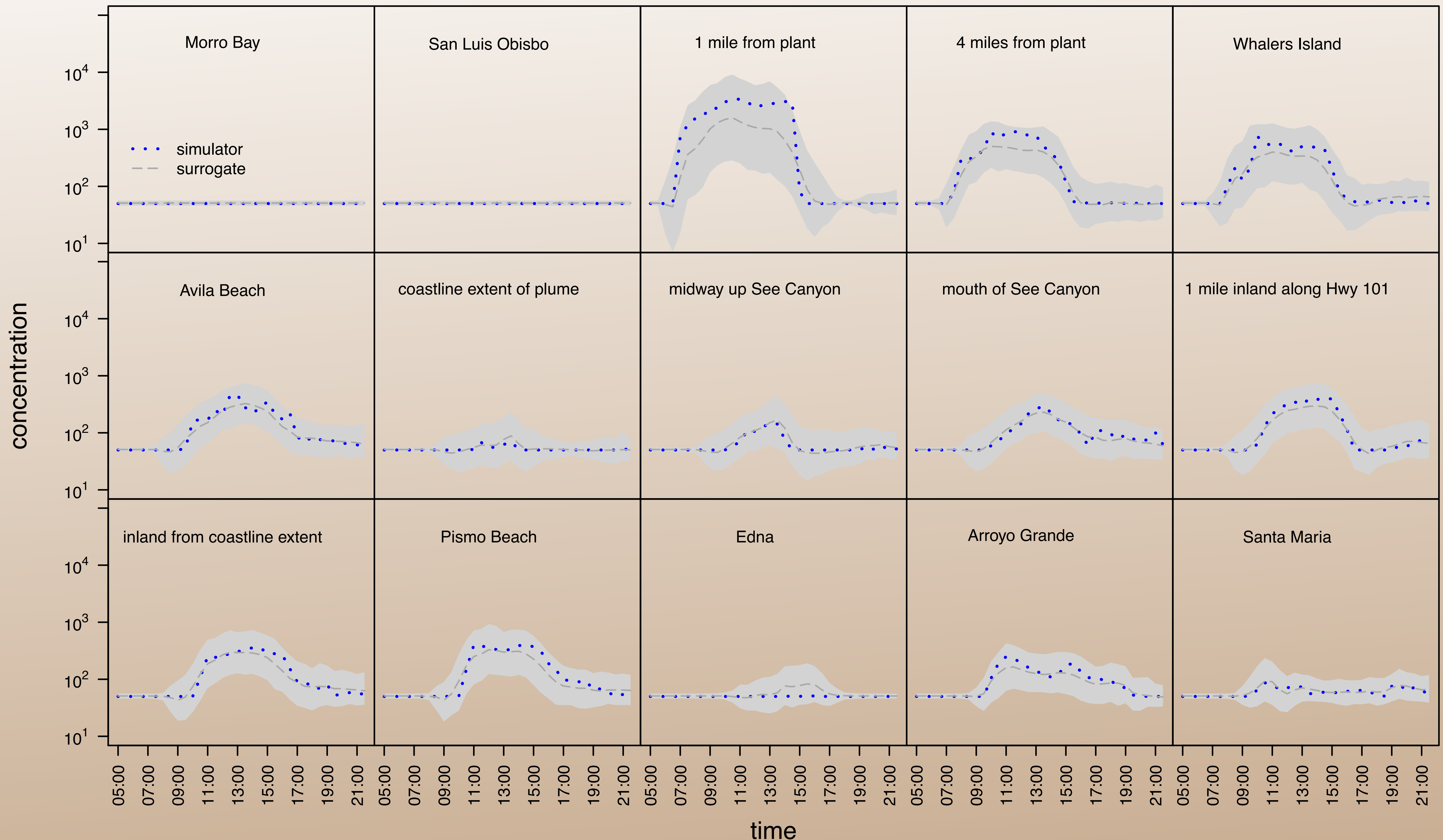
where t_1 and t_2 are the knots and

$$s_i = \pm 1$$

and C_i corresponds to one or more categories of the i-th variable.

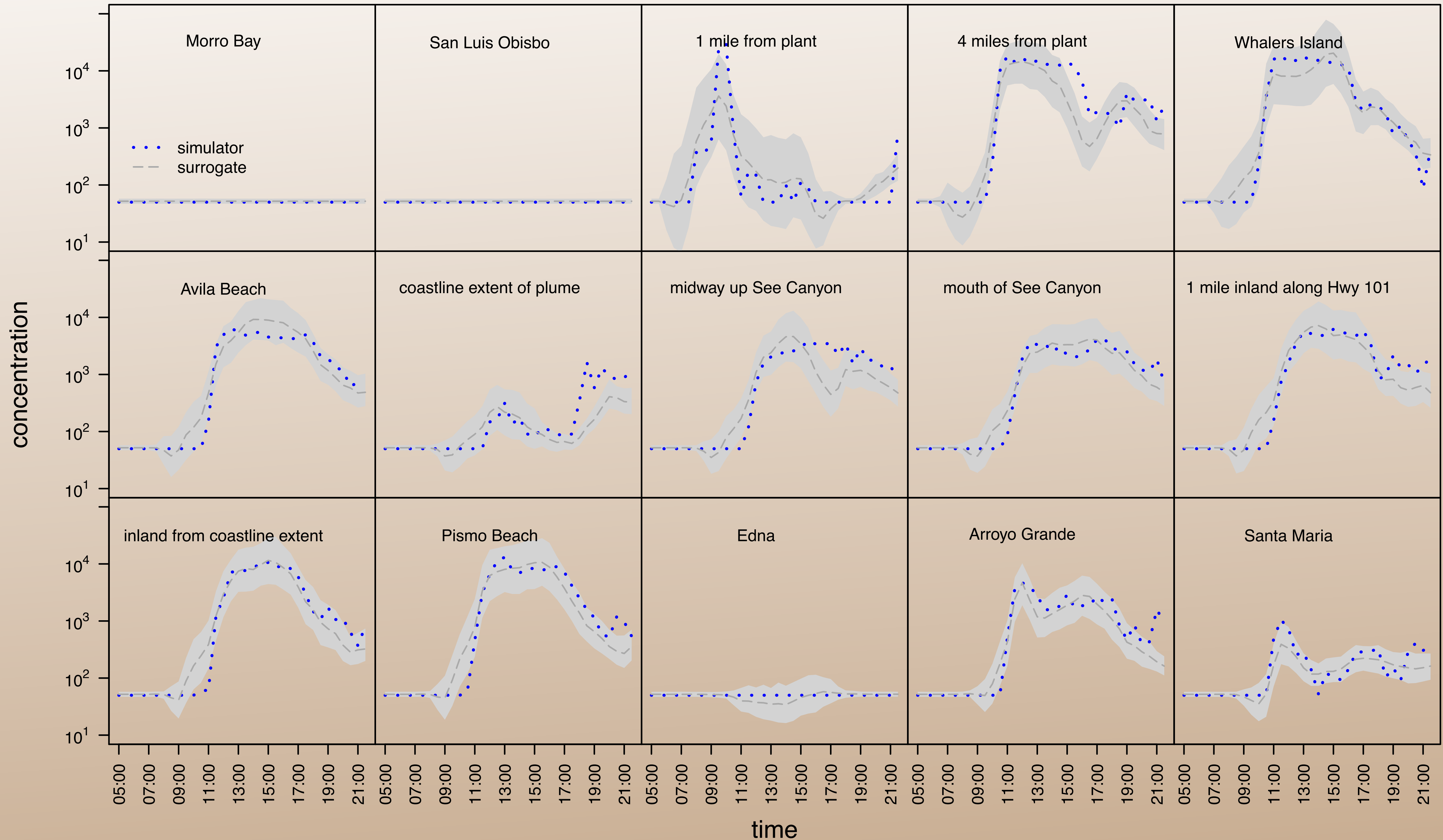
- Allows for basis functions specific to some categories
- Allows for basis functions common to all categories
- Learn from the data the categorical variables in each basis function, if any

Emulation Performance



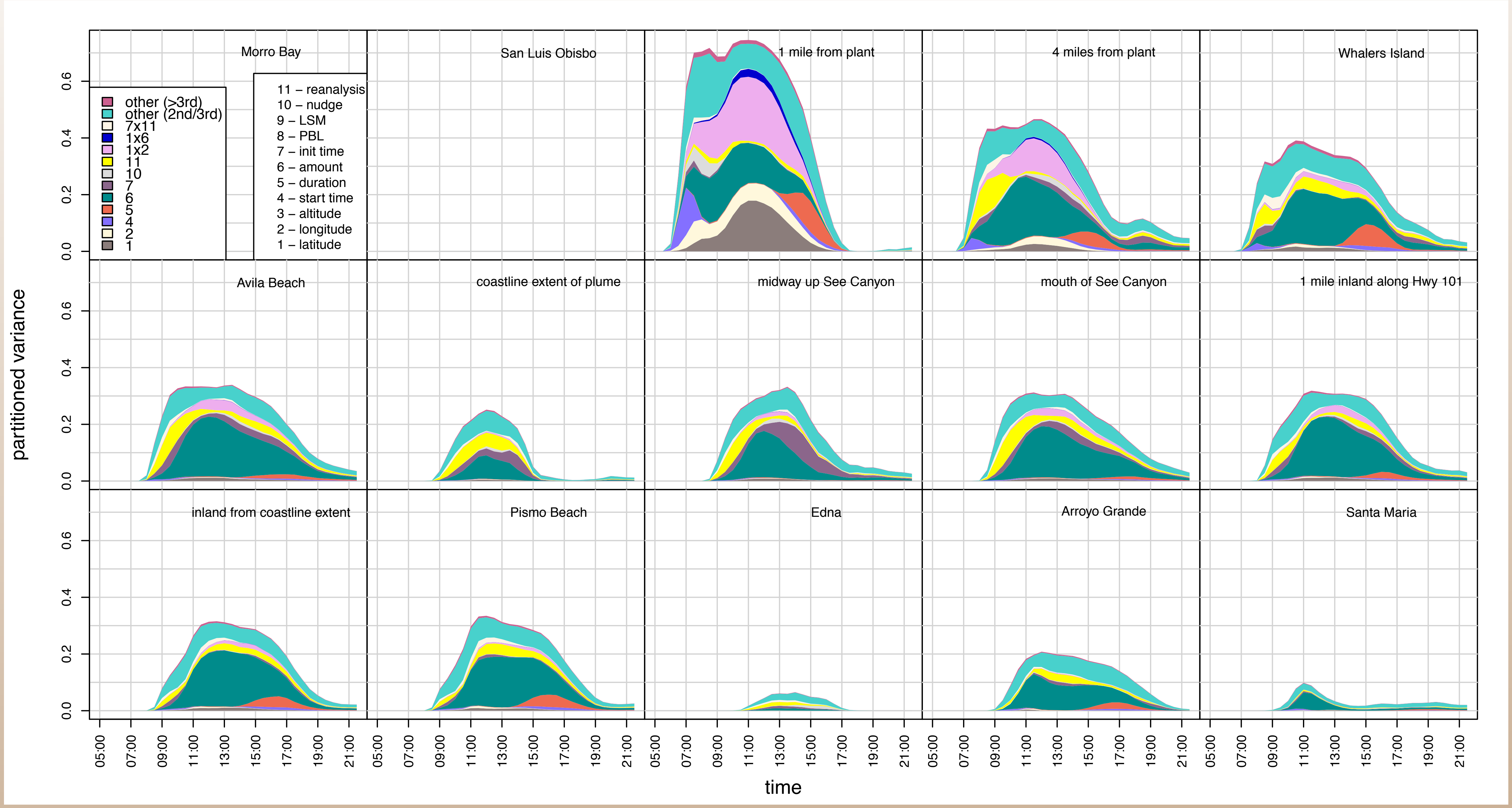
Predictions at 15 of the of the 137 locations considered for a held out configuration of the input parameters.

Emulation Performance



Predictions at 15 of the of the 137 locations considered for another held out configuration of the input parameters.

Global Sensitivity



Analytic expressions are available for the time and space-varying Sobol coefficients for the different inputs and interactions.

Calibration: $p(\theta|Y^F, Y^C)$

Observation Equation with Gaussian error

$$\overbrace{y^F(s,t)}^{Observations} = \underbrace{\zeta(s,t)}_{True\ System} + \overbrace{v(s,t)}^{Oberv.\ Error}$$

System Equation with additive discrepancy with U[0,2] multiplication factor

$$\zeta(s,t) = \underbrace{y^P(s,t,\theta)}_{Best\ Estimate} + \overbrace{\gamma\delta(s,t)}^{Discrepancy}$$

Calibration: $p(\theta|Y^F, Y^C)$

Observation Equation with Gaussian error

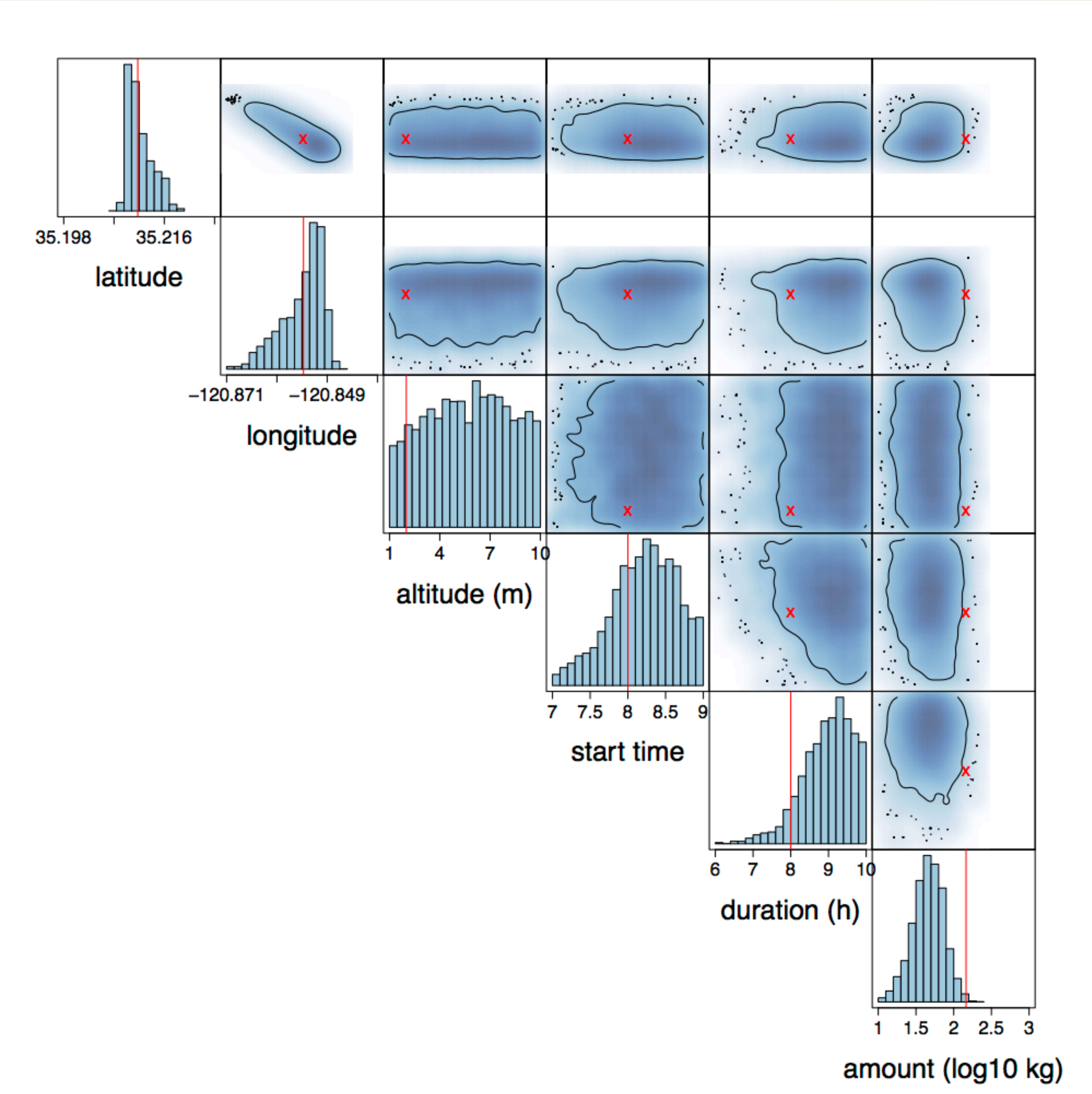
$$\overbrace{y^F(s,t)}^{Observations} = \underbrace{\zeta(s,t)}_{True\ System} + \overbrace{v(s,t)}^{Oberv.\ Error}$$

System Equation with additive discrepancy with U[0,2] multiplication factor

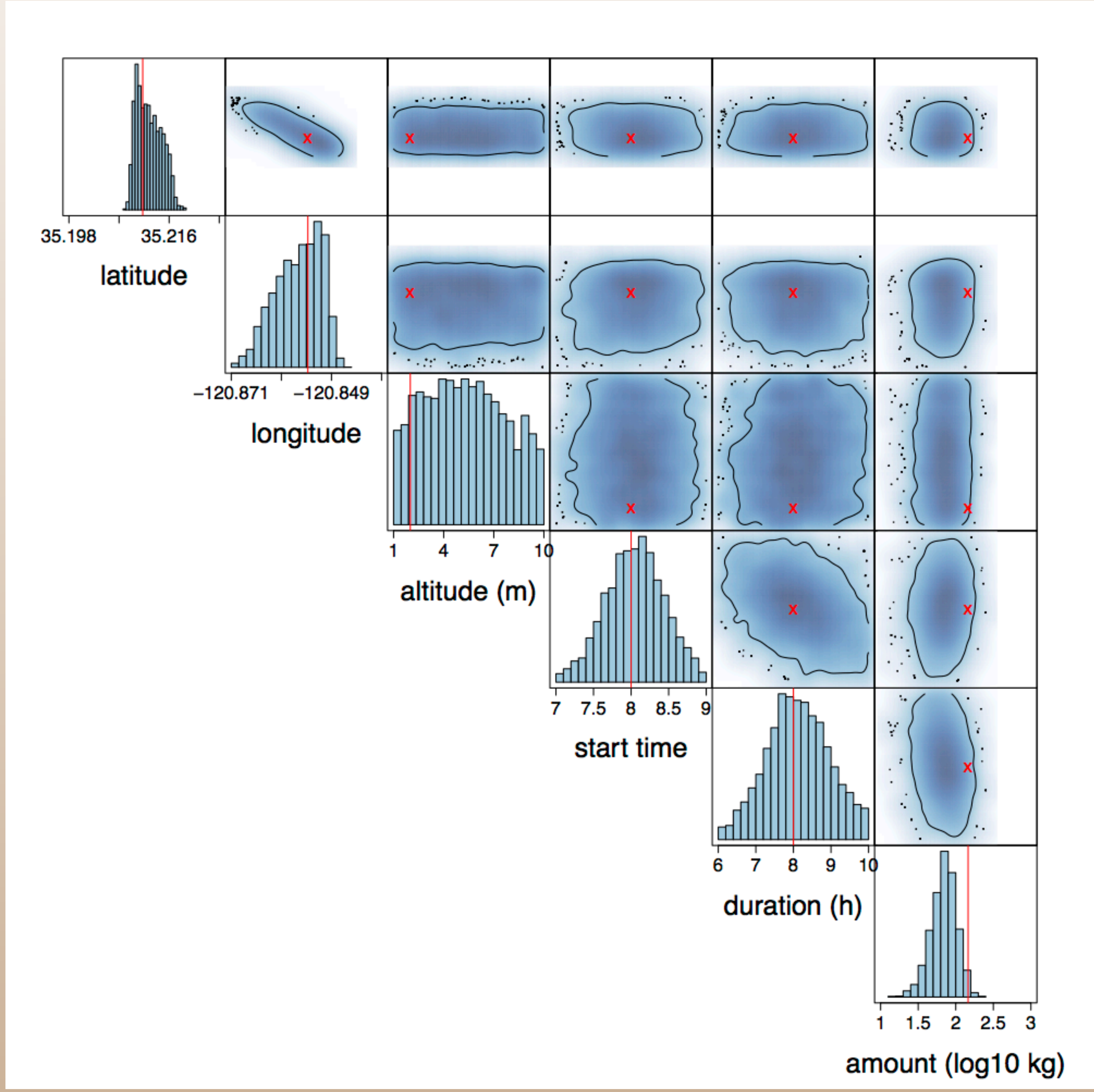
$$\zeta(s,t) = \underbrace{y^P(s,t,\theta)}_{Best\ Estimate} + \overbrace{\gamma\delta(s,t)}^{Discrepancy}$$

We estimate the discrepancy by fitting the emulator at the prior mean of the inputs. We then fit the discrepancy using adaptive splines (BASS). γ provides information about the relevance of the discrepancy.

Posteriors for Continuous Inputs

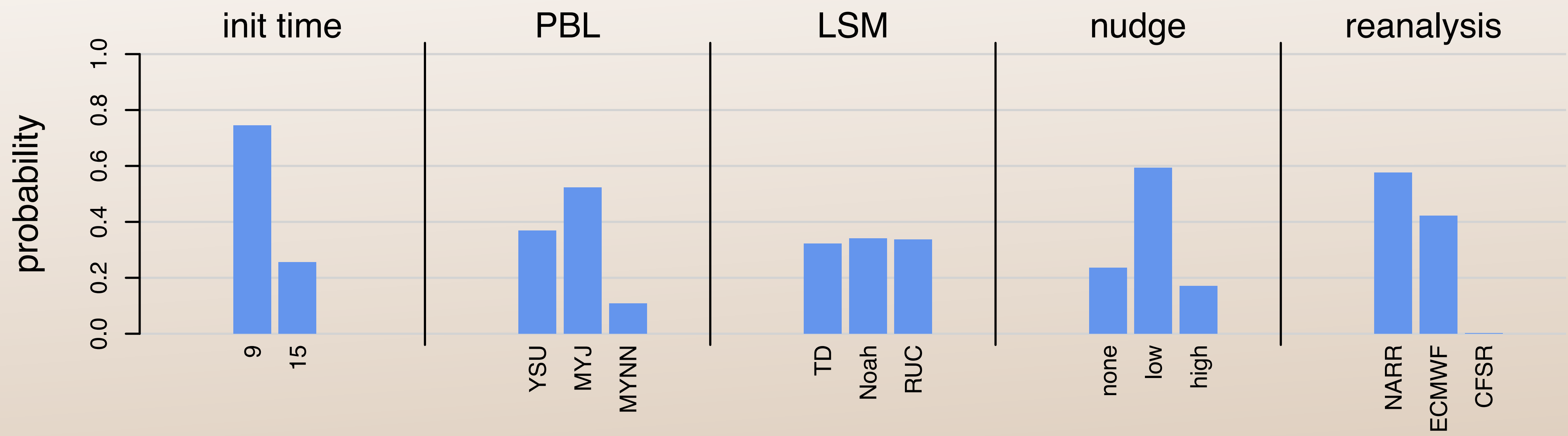


No discrepancy

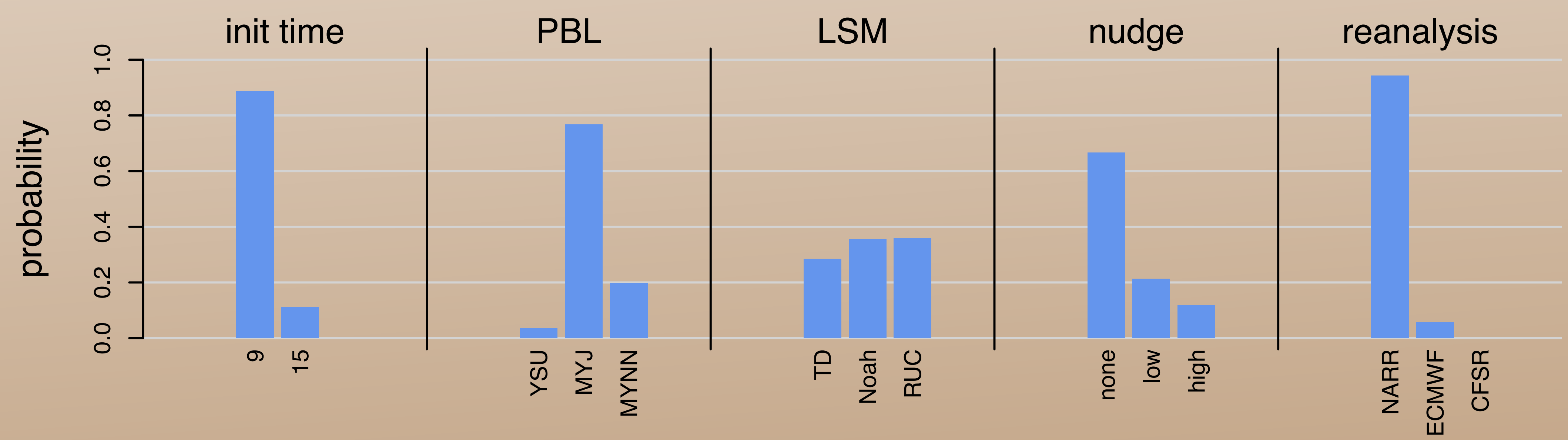


Discrepancy

Posteriors for Categorical Inputs

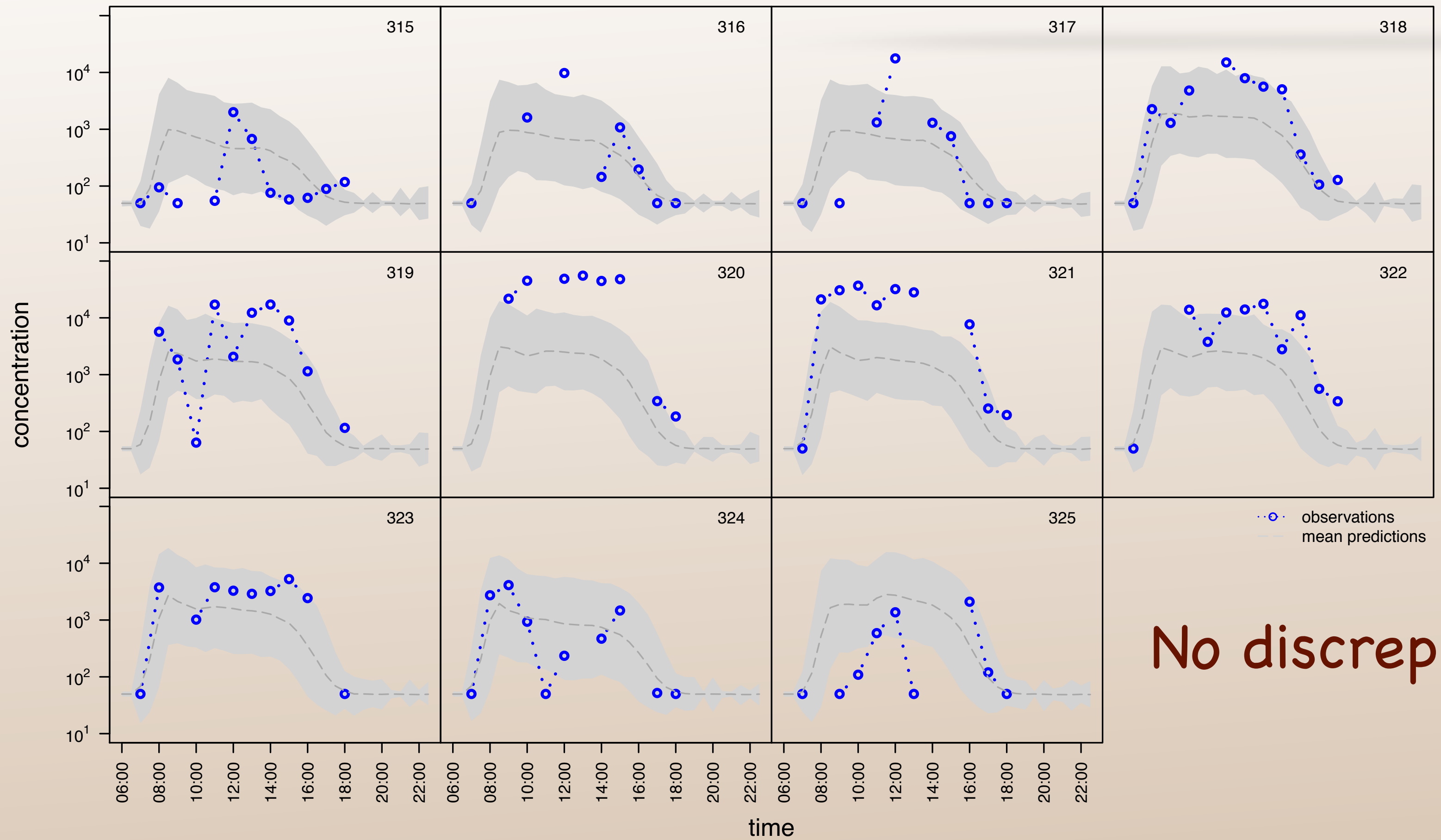


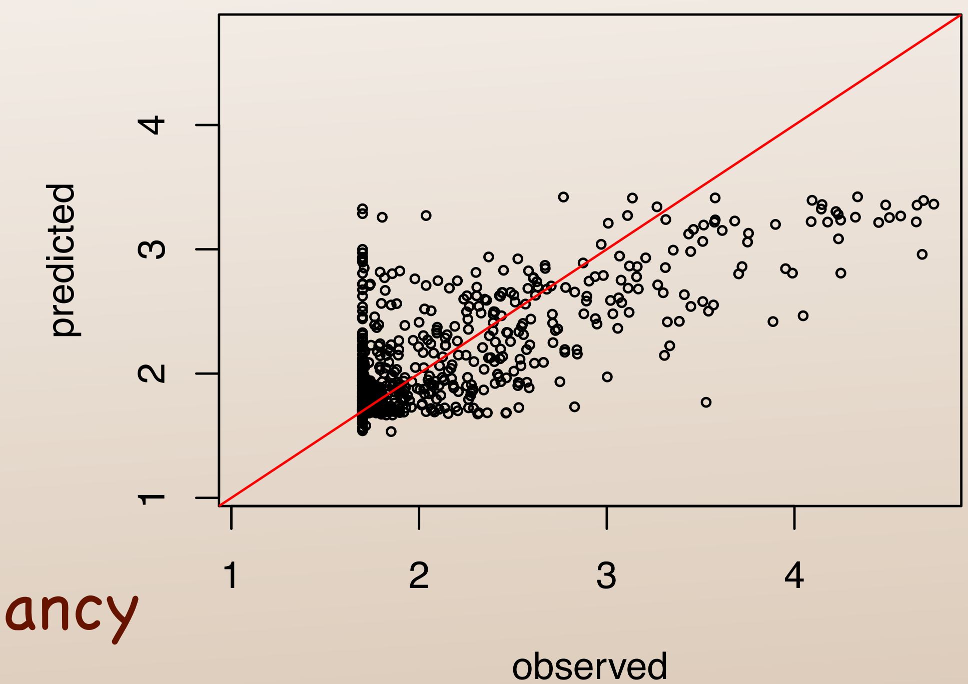
No discrepancy



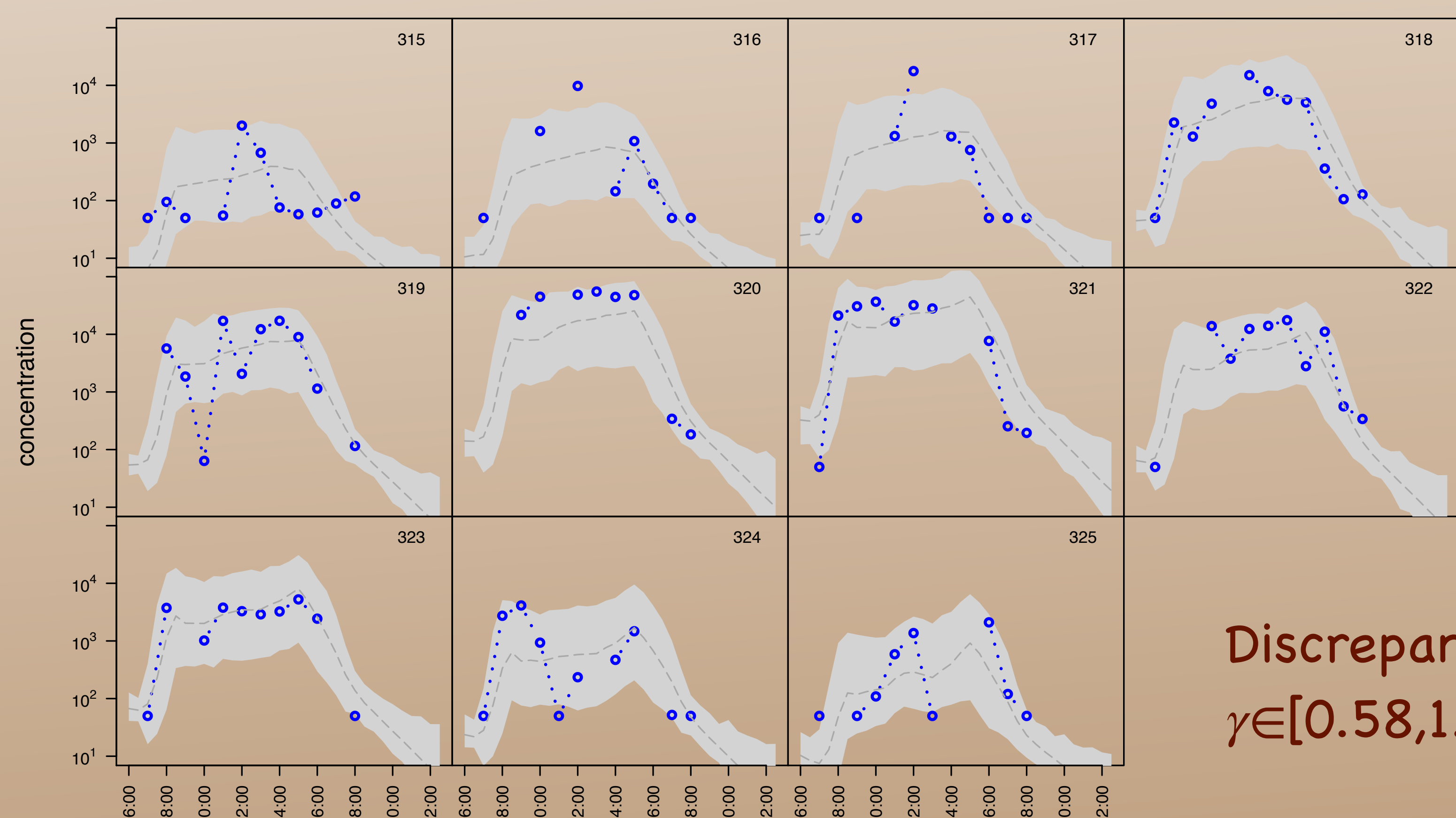
Discrepancy

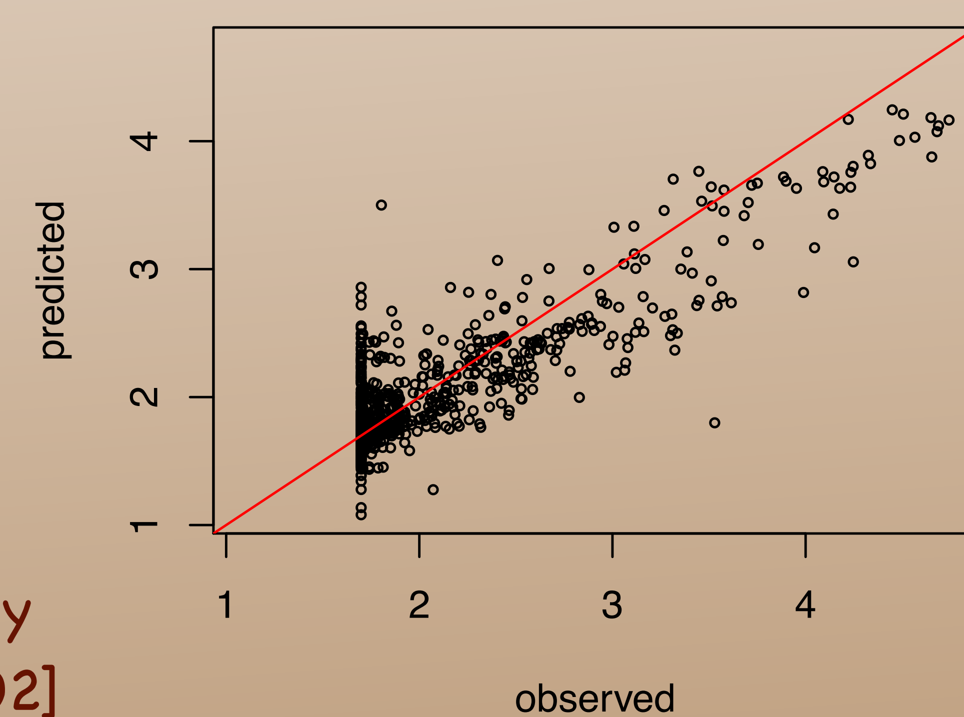
Calibrated Predictions





No discrepancy

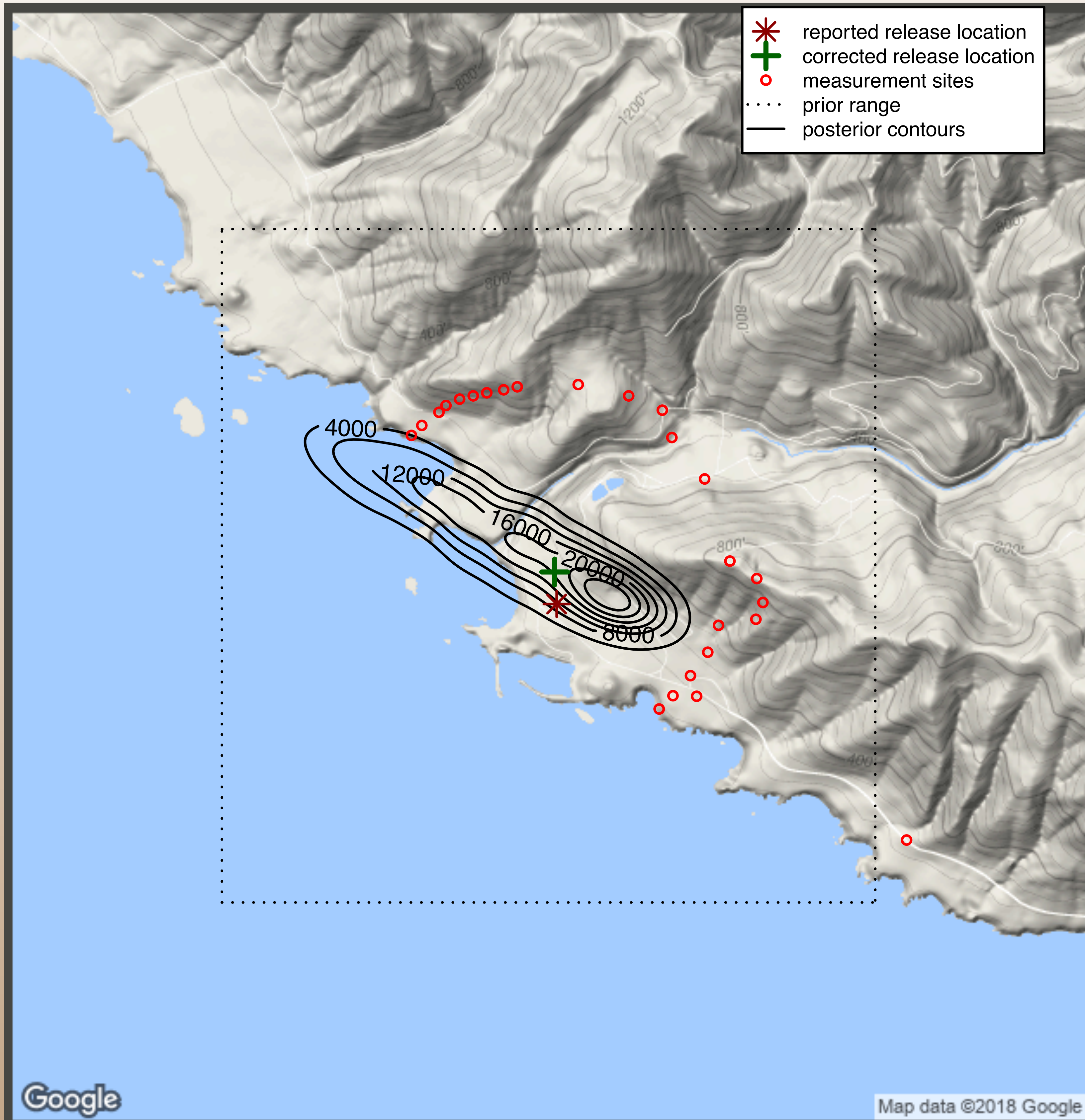




Discrepancy

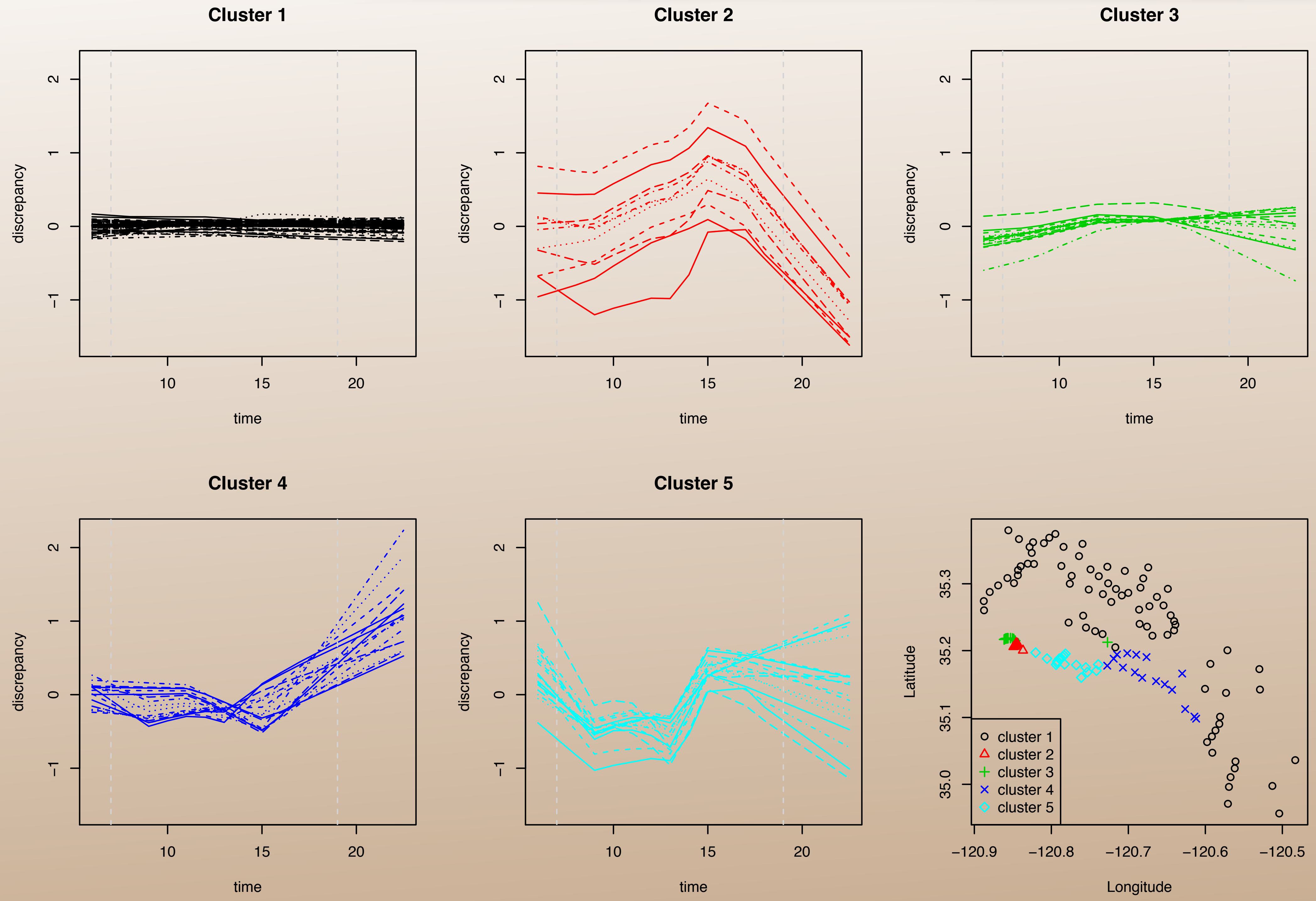
$\gamma \in [0.58, 1.02]$

Calibrated Release Location



The release location was originally mis-reported. Our posterior distribution reveals that a second source of information corresponds to a much more probable location

Analysis of Discrepancies



Clusters of discrepancy curves identifying clear location patterns.

Conclusions

Conclusions

- Adaptive regression splines have been effectively used to tackle an emulation and calibration problem for a massive computer experiment with large spatio-temporal output.

Conclusions

- Adaptive regression splines have been effectively used to tackle an emulation and calibration problem for a massive computer experiment with large spatio-temporal output.
- Our method scales well to large amounts of data, providing accurate emulation and prediction.

Conclusions

- Adaptive regression splines have been effectively used to tackle an emulation and calibration problem for a massive computer experiment with large spatio-temporal output.
- Our method scales well to large amounts of data, providing accurate emulation and prediction.
- Our method can handle continuous and categorical inputs.

Conclusions

- Adaptive regression splines have been effectively used to tackle an emulation and calibration problem for a massive computer experiment with large spatio-temporal output.
- Our method scales well to large amounts of data, providing accurate emulation and prediction.
- Our method can handle continuous and categorical inputs.
- We able to perform a time and space-varying sensitivity analysis of the inputs based on accurate analytic expressions for the global sensitivity coefficients.

Conclusions

- Adaptive regression splines have been effectively used to tackle an emulation and calibration problem for a massive computer experiment with large spatio-temporal output.
- Our method scales well to large amounts of data, providing accurate emulation and prediction.
- Our method can handle continuous and categorical inputs.
- We able to perform a time and space-varying sensitivity analysis of the inputs based on accurate analytic expressions for the global sensitivity coefficients.
- The method uses a fully probabilistic approach that allows to account for all sources of variability and provides a coherent quantification of the uncertainty.

References

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411-428.

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.
- Francom, D., Sansó B., Kupresanin, A., and Johannesson, G. (2018), "Sensitivity Analysis and Emulation for Functional Data using Bayesian Adaptive Splines," Statistica Sinica, 28, 791-816.

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.
- Francom, D., Sansó B., Kupresanin, A., and Johannesson, G. (2018), "Sensitivity Analysis and Emulation for Functional Data using Bayesian Adaptive Splines," Statistica Sinica, 28, 791-816.
- Francom, D., Sansó B., Bulaevskaya, V., Lucas, D., Simpson, M. (2018), "Inferring Atmospheric Release Characteristics in a Large Computer Experiment using Bayesian Adaptive Splines", submitted

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.
- Francom, D., Sansó B., Kupresanin, A., and Johannesson, G. (2018), "Sensitivity Analysis and Emulation for Functional Data using Bayesian Adaptive Splines," Statistica Sinica, 28, 791-816.
- Francom, D., Sansó B., Bulaevskaya, V., Lucas, D., Simpson, M. (2018), "Inferring Atmospheric Release Characteristics in a Large Computer Experiment using Bayesian Adaptive Splines", submitted
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), "Computer model calibration using high-dimensional output," Journal of the American Statistical Association, 103.

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.
- Francom, D., Sansó B., Kupresanin, A., and Johannesson, G. (2018), "Sensitivity Analysis and Emulation for Functional Data using Bayesian Adaptive Splines," Statistica Sinica, 28, 791-816.
- Francom, D., Sansó B., Bulaevskaya, V., Lucas, D., Simpson, M. (2018), "Inferring Atmospheric Release Characteristics in a Large Computer Experiment using Bayesian Adaptive Splines", submitted
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), "Computer model calibration using high-dimensional output," Journal of the American Statistical Association, 103.
- Kennedy, M. C., and O'Hagan, A. (2001), "Bayesian calibration of computer models," Journal of the Royal Statistical Society. Series B, Statistical Methodology, 425–464.

References

- Chakraborty, A., Mallick, B. K., Mcclarren, R. G., Kuranz, C. C., Bingham, D., Grosskopf, M. J., Rutter, E. M., Stripling, H. F., and Drake, R. P. (2013), "Spline-based emulators for radiative shock experiments with measurement error," Journal of the American Statistical Association, 108, 411–428.
- Francom, D., Sansó, B. (2018), "BASS: An R Package for Fitting and Performing Sensitivity Analysis of Bayesian Adaptive Spline Surfaces", Journal of Statistical Software, to appear.
- Francom, D., Sansó B., Kupresanin, A., and Johannesson, G. (2018), "Sensitivity Analysis and Emulation for Functional Data using Bayesian Adaptive Splines," Statistica Sinica, 28, 791-816.
- Francom, D., Sansó B., Bulaevskaya, V., Lucas, D., Simpson, M. (2018), "Inferring Atmospheric Release Characteristics in a Large Computer Experiment using Bayesian Adaptive Splines", submitted
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), "Computer model calibration using high-dimensional output," Journal of the American Statistical Association, 103.
- Kennedy, M. C., and O'Hagan, A. (2001), "Bayesian calibration of computer models," Journal of the Royal Statistical Society. Series B, Statistical Methodology, 425–464.