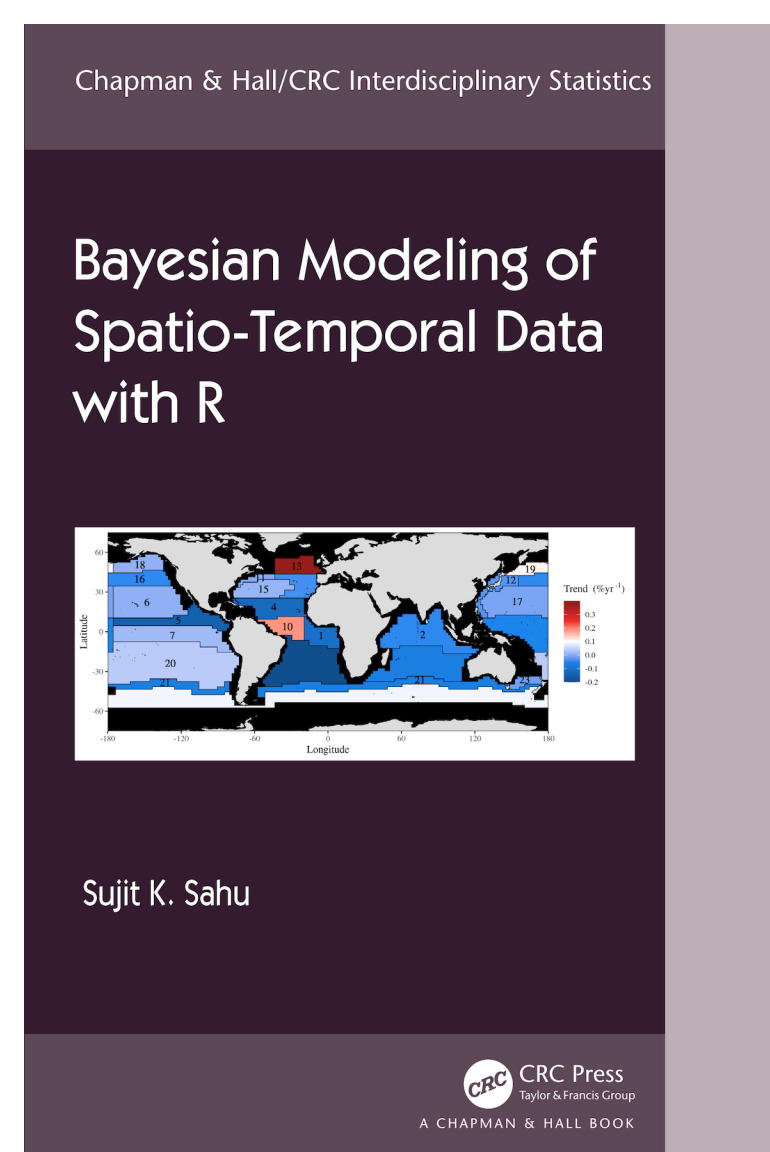


bmstdr: Bayesian Modeling of Spatio-Temporal Data with R



Sujit Sahu

University of Southampton, UK



The R package *bmstdr* facilitates Bayesian modeling of point referenced and areal unit spatial and spatio-temporal data. Using multiple packages, it allows model choice and validation as the user wishes to. The user does not need to write any package specific code.

bmstdr model fitting functions. Abbreviations: GP (Gaussian Process), GPP (Predictive GP), AR (AutoRegressive), BYM (Besag, York and Mollié).

Data type	Function	Available package options	Models(Book §)	Examples
1. point referenced (spatial)	<i>Bspatial</i>	spBayes, Stan, INLA	lm, GP (§6)	Air pollution
2. point referenced (spatio-temporal)	<i>Bsptime</i>	spTimer, spTDyn, spBayes, Stan, INLA	lm, GP, GPP, AR (§7, 8)	Air pollution, Precipitation volume
3. point referenced (moving sensors)	<i>Bmoving_sptime</i>	Stan	GPP (§8)	Temperature in deep ocean
4. Areal unit (both static & dynamic)	<i>Bcartime</i>	CARBayes, CARBayesST, INLA	GLM, BYM, Leroux, AR (§10, 11)	Cancer rates, hospitalization rates, child poverty, vaccination coverage

1. *Bspatial*: Point referenced spatial data modeling

- Model: $\mathbf{Y} = X\beta + \mathbf{w} + \epsilon$, details in §6 of BMSTDR.
- Typical model fitting call requires the arguments: package, data, and formula and behaves like `lm`.

```
M1 <- Bspatial(package="spBayes", data=nyspatial, formula=yo3 ~ xmaxtemp+wdsp+
  xrh, coordtype="utm", coords=4:5, prior.phi=c(0.005, 2), mchoice=TRUE, N=5000,
  burn.in=1000, scale.transform="NONE", validrows=c(8,11,12,14,18,21,24,28))
print(M1); summary(M1); residuals(M1); plot(M1); coef(M1)
```

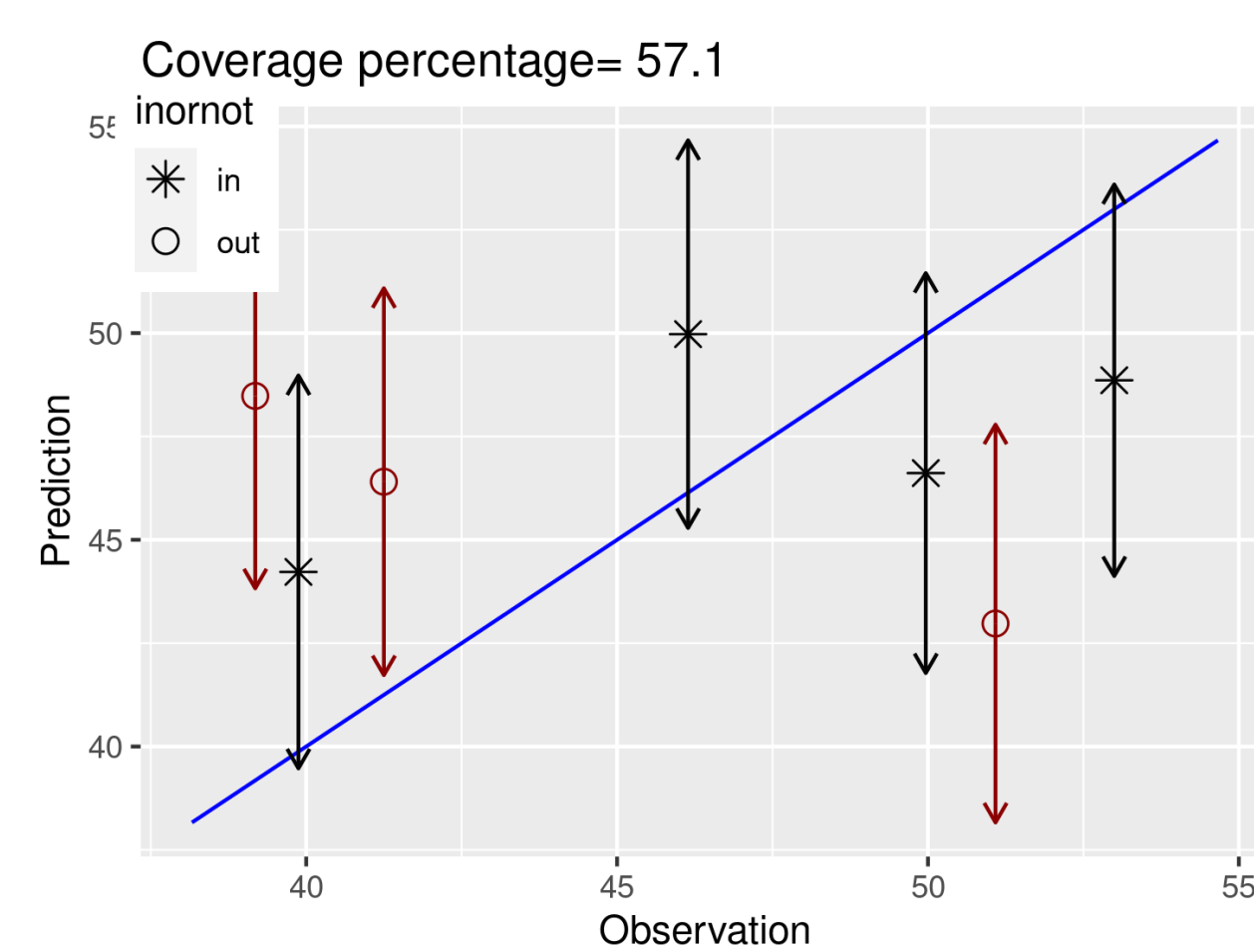
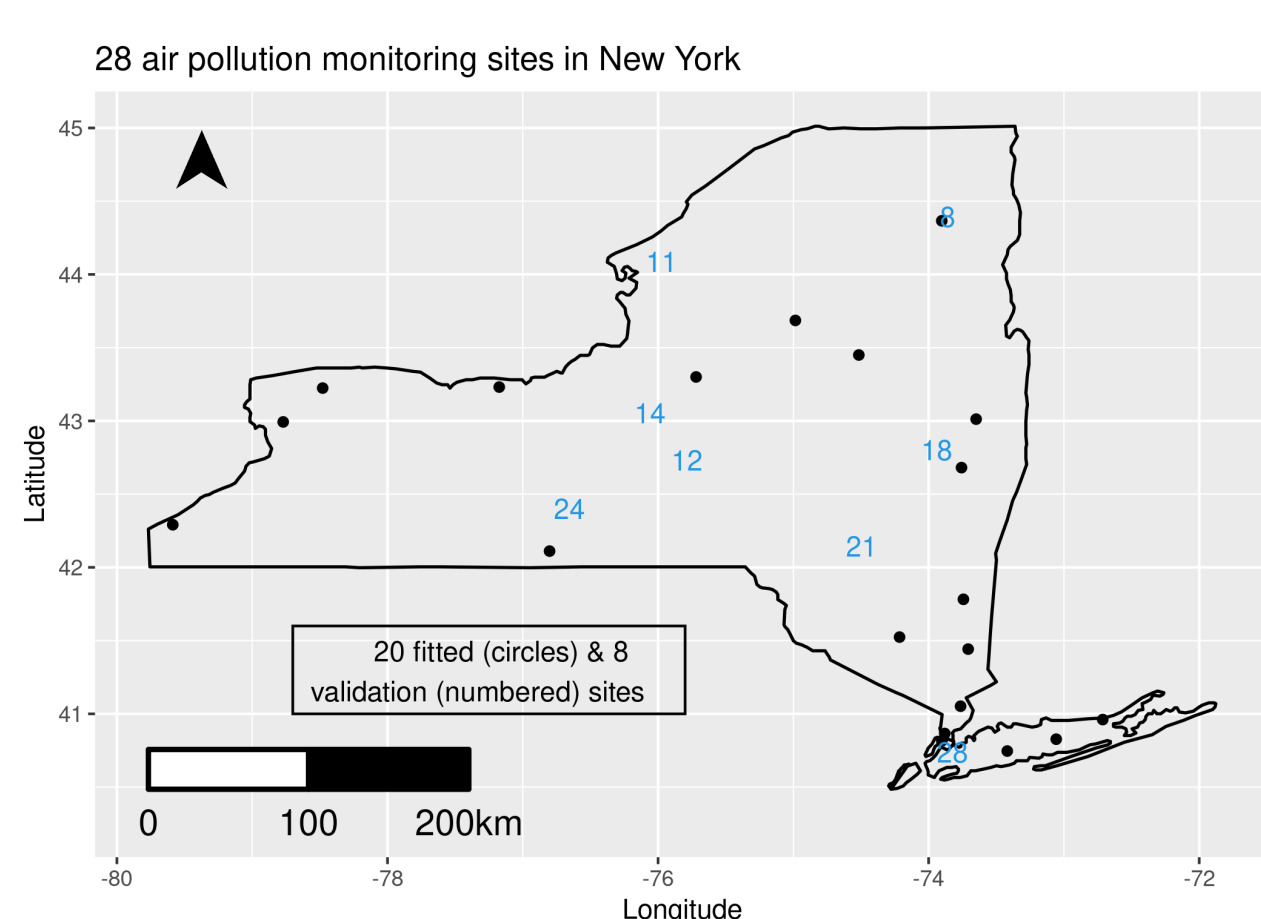


Figure 1. Pollution monitoring sites in New York and validation plot generated by *Bspatial*.

2. *Bsptime*: Point referenced spatio-temporal data modeling

- Model: $\mathbf{Y}_t = X_t\beta + \mathbf{w}_t + \epsilon_t$, details in §7 and §8.
- Each column in Table 1 is output of a call such as:

```
M2 <- Bsptime(package="stan", data=nysptime, formula=y8hrmax ~ xmaxtemp+wdsp+
  xrh, coordtype="utm", coords=4:5, scale.transform="SQRT", validrows=sample(1736, 496))
```

	Independent	Separable	spTimerGP	stan	inla	spTimerAR	spTDyn	sptimerGPP
RMSE	9.35	6.49	6.40	6.42	9.73	6.46	6.59	6.36
MAE	7.54	5.00	4.94	4.85	7.64	4.99	5.11	4.85
CRPS	5.67	10.56	6.79	3.57	2.63	5.97	5.12	7.47
CVG	98.36	99.59	99.59	92.83	65.37	99.39	99.39	99.39
G	728.91	218.49	181.71	173.45	527.67	185.76	71.30	146.69
P	731.61	195.37	935.42	266.23	17.19	718.47	467.46	815.85
G+P	1460.52	413.86	1117.13	439.67	544.86	904.23	538.76	962.54

Table 1. Validation statistics for different models implemented using different packages.

Theory details are in the book.
Code and data are online.

3. *Bmoving_sptime*: Modeling data from moving sensors

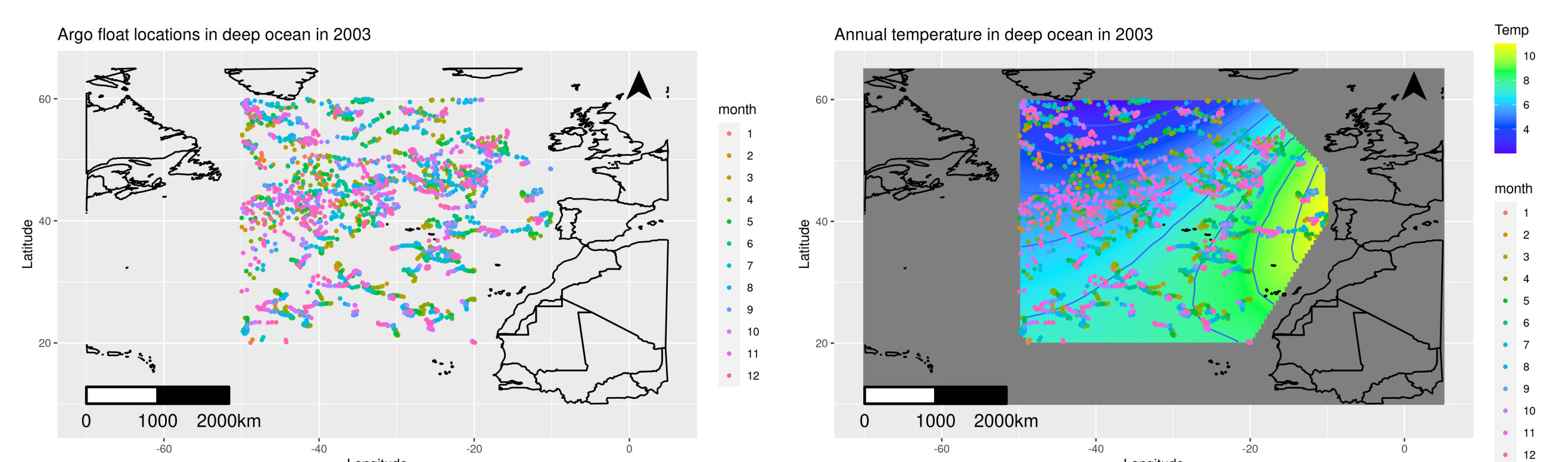


Figure 2. Locations of moving sensors and interpolated temperature map of the deep ocean

- Model data from Argo-floats in the North Atlantic.
- Fitted by *Bmoving_sptime* implemented in Stan.
- Details are in §8 of the BMSTDR book.

4. *Bcartime*: CAR modeling of areal (static & dynamic) data

- We model Covid-19 death data from 313 English areas for 20 weeks and validate 10% data, see §10.

```
M4 <- Bcartime(package="inla", model=c("bym", "ar1"), data=engdeaths, formula=
  covid ~ jsa, scol="spaceid", tcol="Weeknumber", W=Weng, validrows=sample(6260, 626))
```

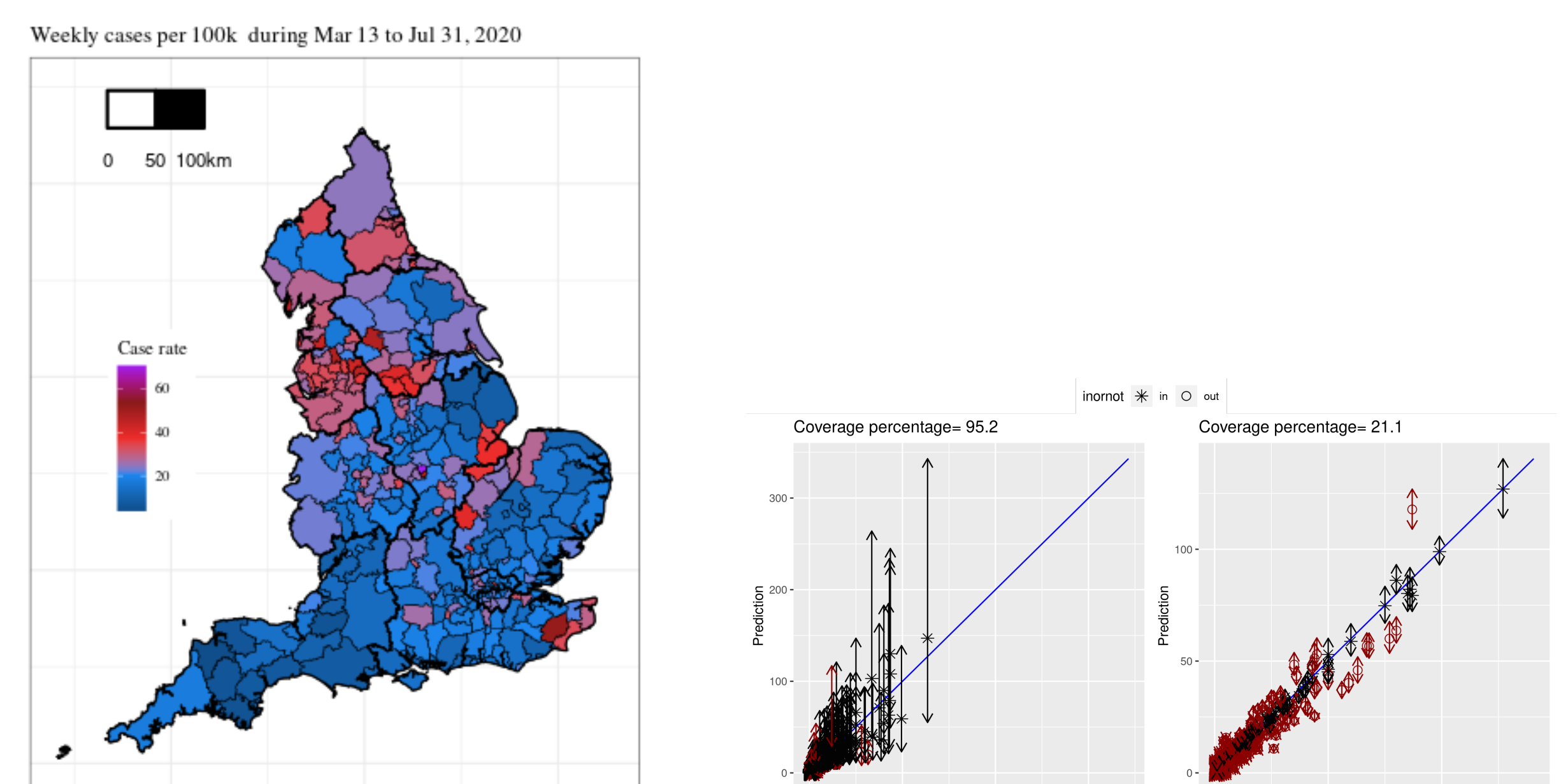


Figure 3. Validation plots for models based on CARBayesST and INLA packages.

- Bcartime* also includes most of the CARBayes & CARBayesST fitting functions for areal data (§11).
- Only required to change a few arguments in the model fitting call (see §10 and 11).