

Supplement

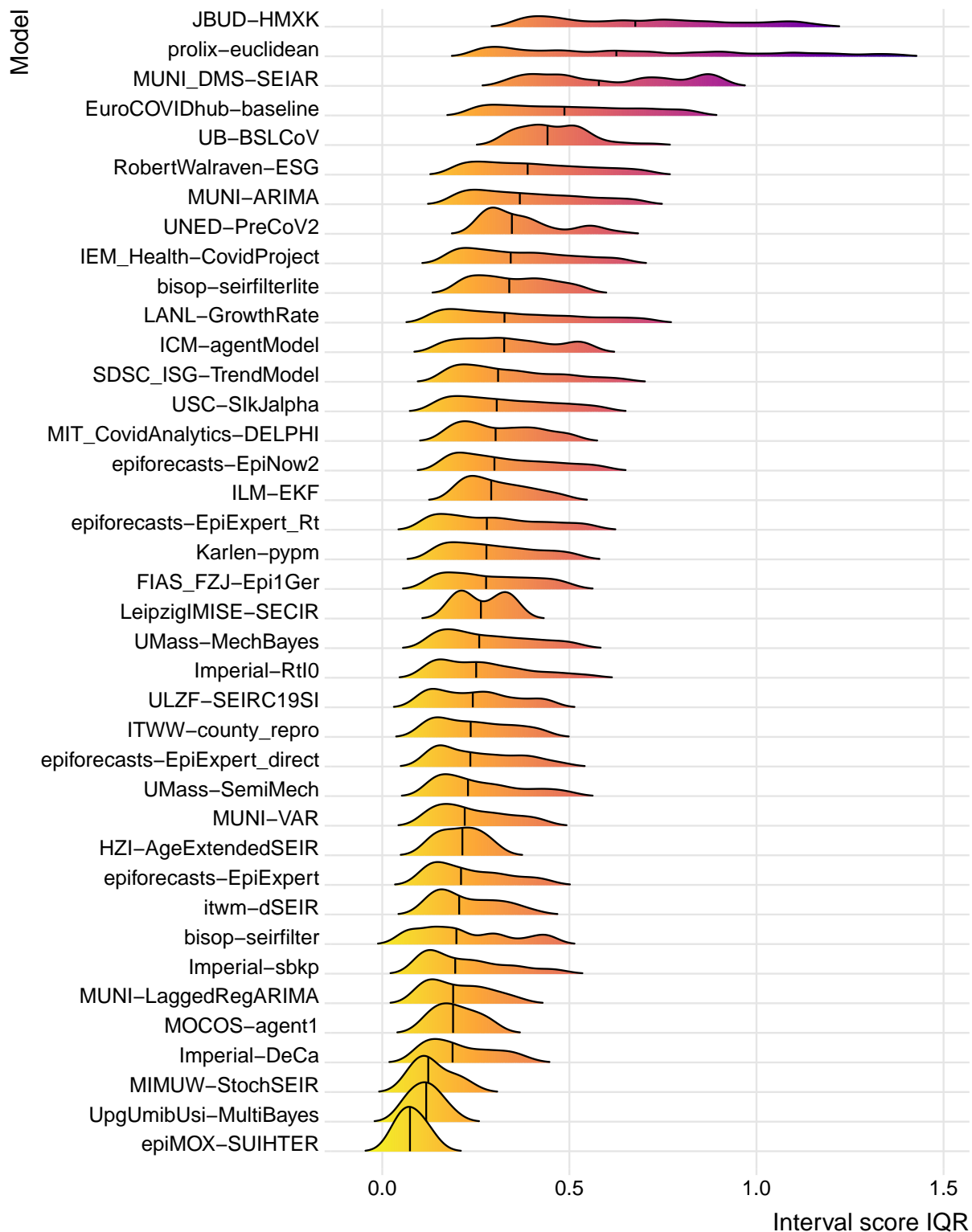


Figure 1: Distribution of forecast scores for 1 to 4 week ahead forecasts across 32 locations over 104 weeks (N=103249). Each distribution shows the interquartile range and median (vertical line) of interval scores across forecasts made by each model, with lower interval score indicating better predictive accuracy. Each model forecast for a different combination of targets, with some models contributing very few forecasts, meaning that forecast scores are not directly comparable.

Model fitting

Code is available at: <https://github.com/epiforecasts/model-structure-evaluation>.

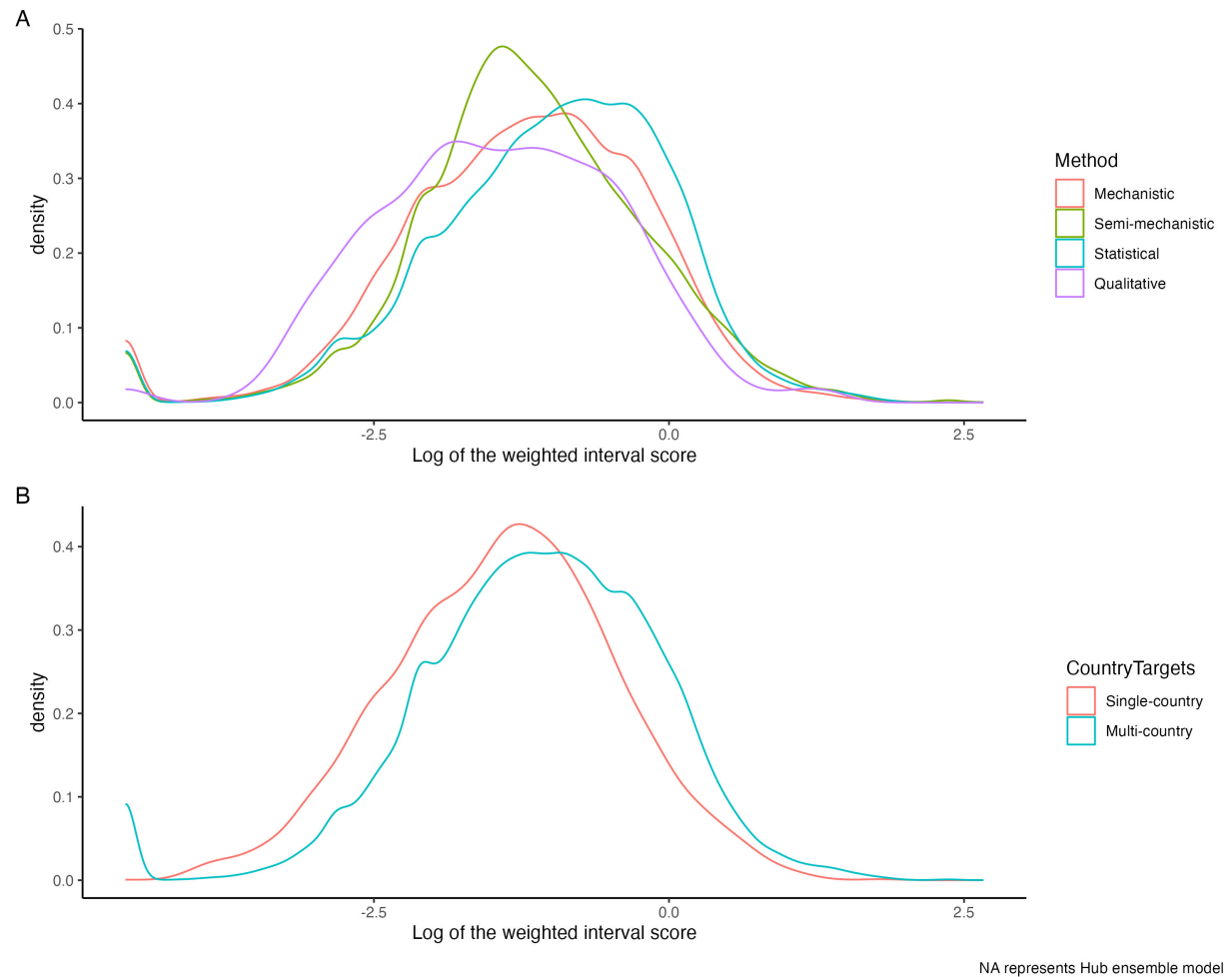


Figure 2: Density plot

1. Conditional distribution of outcomes

2. Model formula \sim , log_interval_score, Method + CountryTargets + s(Incidence) + Trend + Horizon + s(Model, bs = “re”)

3. Summary of estimates Reference levels for categorical variables are: - Mechanistic model method - Single-country target - Modelling team affiliation is located in the forecast target country - Trend (of observed incidence) is stable - Horizon is one-week-ahead

Component	Term	Estimate	Std Error	t-value
A. parametric coefficients	(Intercept)	-1.737	0.100	-17.413
	MethodSemi-mechanistic	0.046	0.158	0.289
	MethodStatistical	0.217	0.166	1.310
	CountryTargetsMulti-country	0.106	0.138	0.771
	TrendIncreasing	0.624	0.008	73.470
	TrendDecreasing	0.539	0.008	65.807
	Horizon.L	0.544	0.006	89.804
	Horizon.Q	-0.036	0.006	-5.994
	Horizon.C	0.006	0.006	1.017
Component	Term	edf	Ref. df	F-value
B. smooth terms	s(Incidence)	8.114	8.779	303.545
	s(Model)	31.028	32.000	107.352

Signif. codes: 0 ‘***’ < 0.001 ‘**’ < 0.01 ‘*’ < 0.05 ‘.’ < 0.1 ‘ ’ > 0.1

Adjusted R-squared: 0.175, Deviance explained 0.175

-REML : 137956.780, Scale est: 0.893, N: 101181

Table 2: ANOVA for parametric terms

	df	F	p-value
Method	2	0.903	0.405
CountryTargets	1	0.594	0.441
Trend	2	2944.876	0.000
Horizon	3	2703.946	0.000

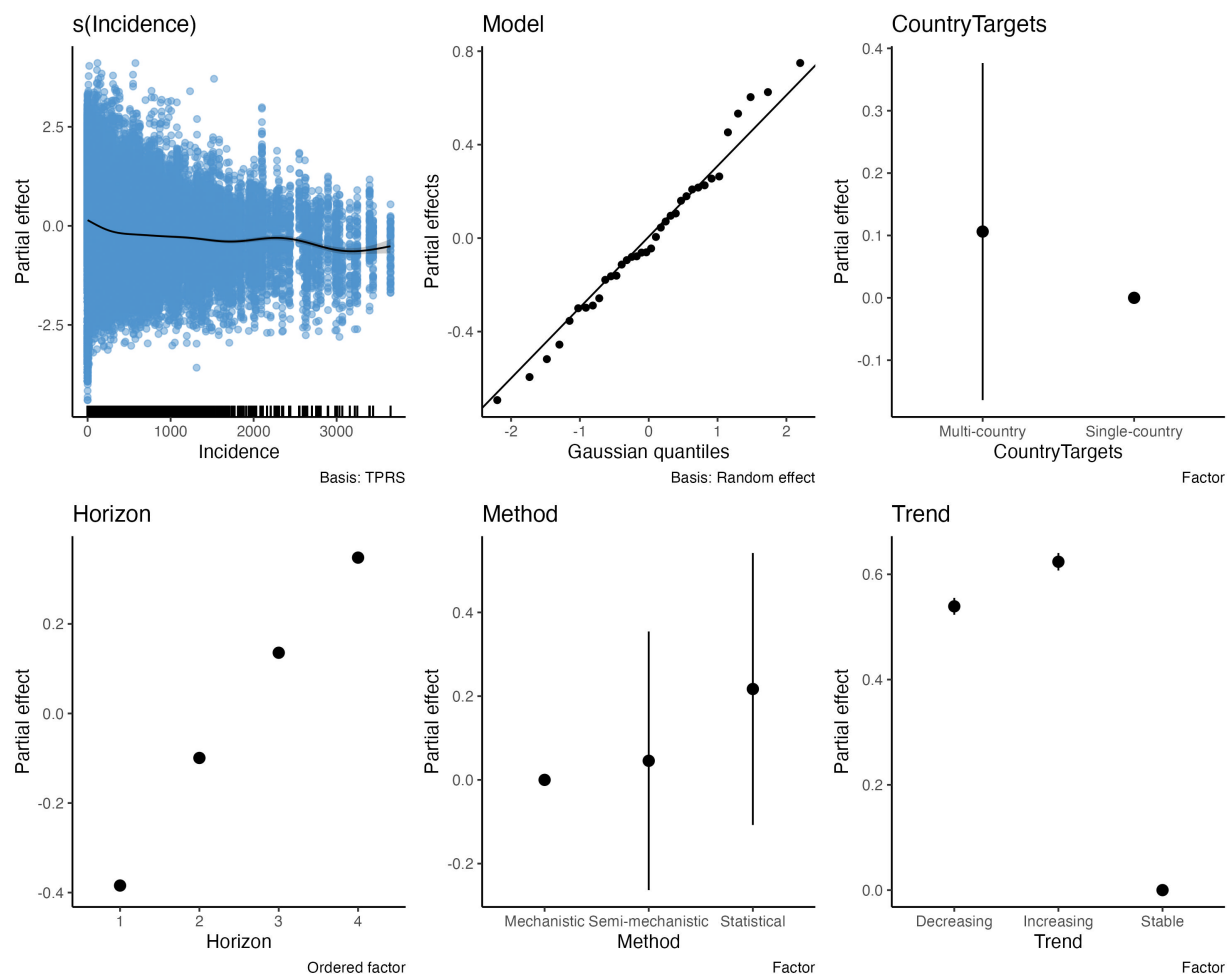


Figure 3: Model estimates

4. Partial effects plots

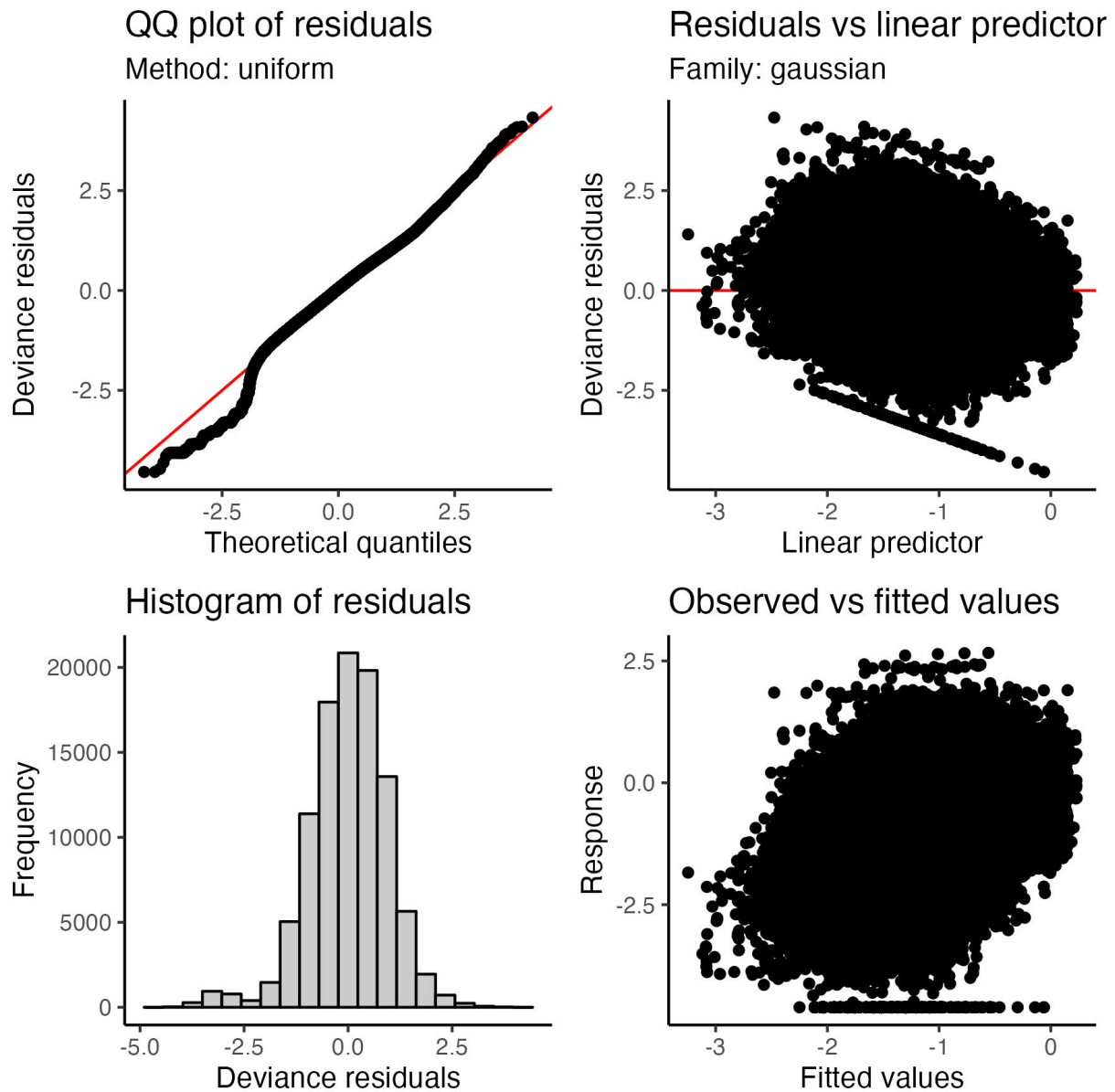


Figure 4: Diagnostics

5. Model diagnostics