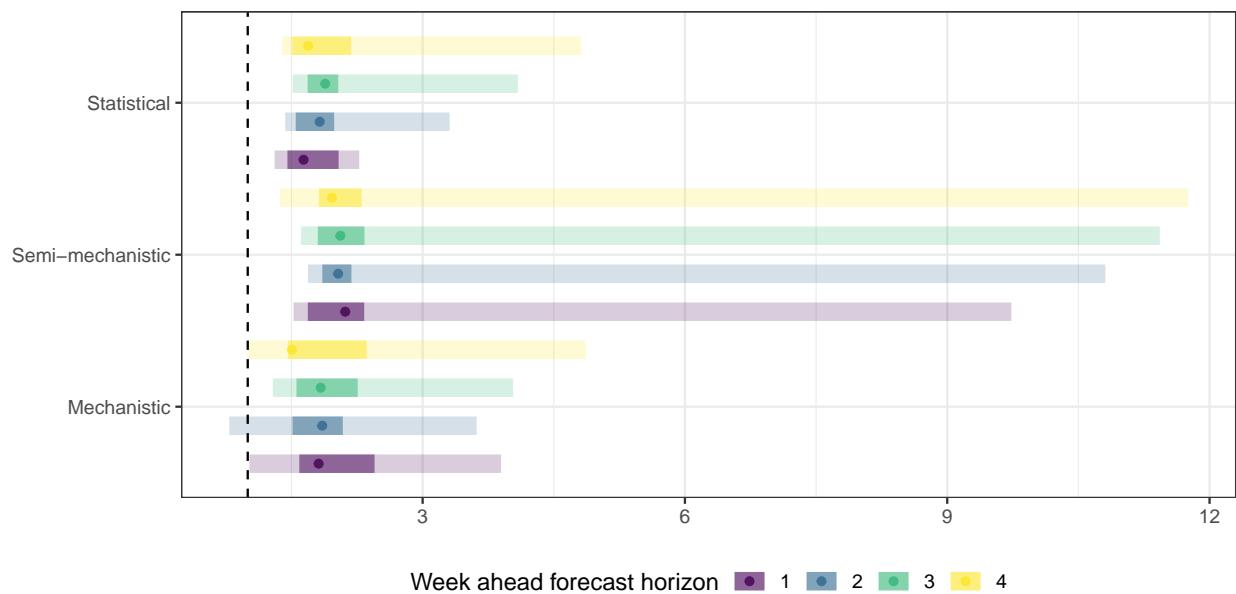


Supplement

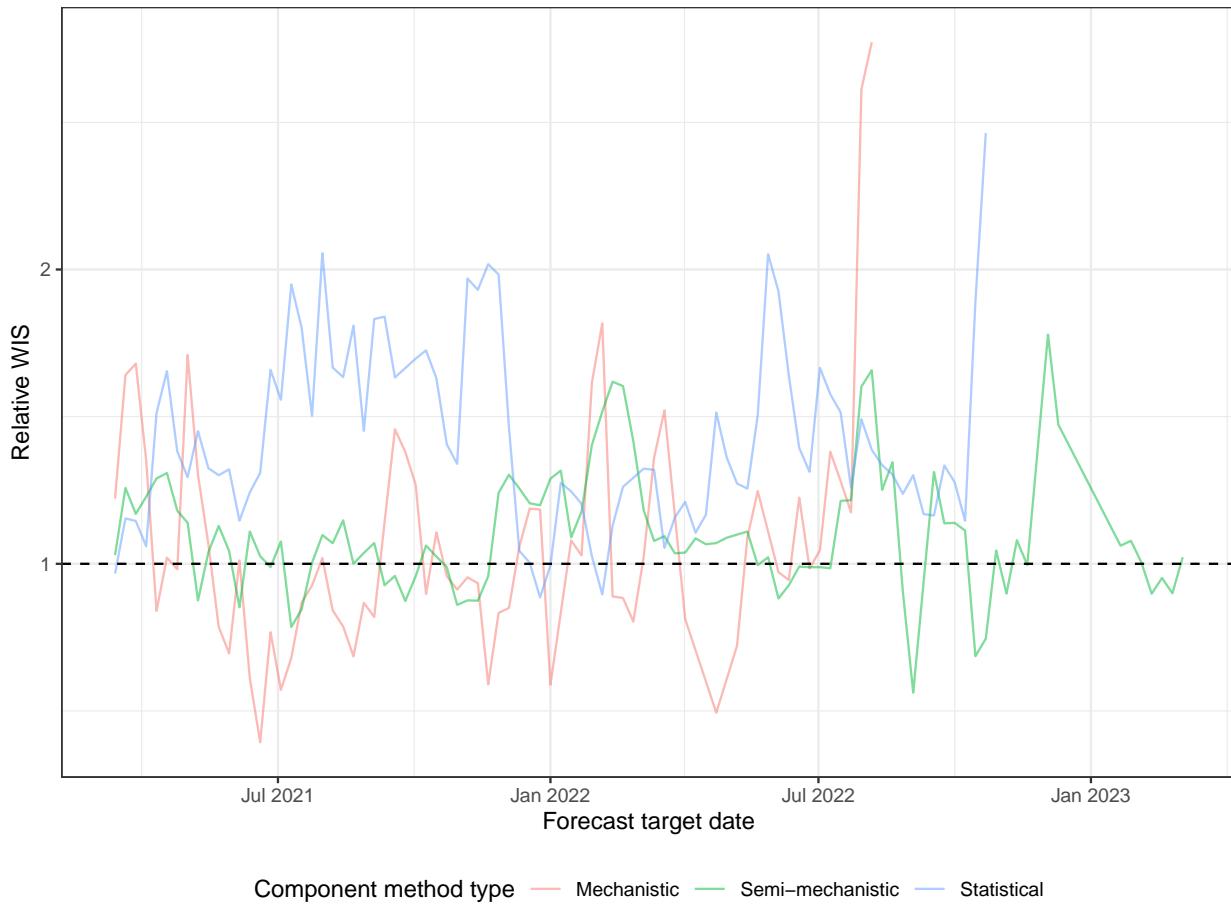
Contents

Scores on the natural scale	1
Ensemble variations	2
Team metadata	2
Model fitting	3

Scores on the natural scale



Ensemble variations



Team metadata

model	method_type	target_type	metadata
IEM_Health-CovidProject	Mechanistic	Multi-country	Link
JBUD-HMXK	Mechanistic	Multi-country	Link
Karlen-pypm	Mechanistic	Multi-country	Link
UMass-MechBayes	Mechanistic	Multi-country	Link
USC-SIkJalpha	Mechanistic	Multi-country	Link
bisop-seirfilterlite	Mechanistic	Multi-country	Link
FIAS_FZJ-Epi1Ger	Mechanistic	Single-country	Link
HZI-AgeExtendedSEIR	Mechanistic	Single-country	Link
ICM-agentModel	Mechanistic	Single-country	Link
LeipzigIMISE-SECIR	Mechanistic	Single-country	Link
MIMUW-StochSEIR	Mechanistic	Single-country	Link
MIT_CovidAnalytics-DELPHI	Mechanistic	Single-country	Link
MOCOS-agent1	Mechanistic	Single-country	Link
MUNI_DMS-SEIAR	Mechanistic	Single-country	Link
ULZF-SEIRC19SI	Mechanistic	Single-country	Link
bisop-seirfilter	Mechanistic	Single-country	Link
epiMOX-SUIHTER	Mechanistic	Single-country	Link

model	method_type	target_type	metadata
itwm-dSEIR	Mechanistic	Single-country	Link
epiforecasts-EpiExpert	Qualitative	Multi-country	Link
epiforecasts-EpiExpert_Rt	Qualitative	Multi-country	Link
epiforecasts-EpiExpert_direct	Qualitative	Multi-country	Link
ILM-EKF	Semi-mechanistic	Multi-country	Link
Imperial-DeCa	Semi-mechanistic	Multi-country	Link
Imperial-RtI0	Semi-mechanistic	Multi-country	Link
Imperial-sbkp	Semi-mechanistic	Multi-country	Link
LANL-GrowthRate	Semi-mechanistic	Multi-country	Link
UMass-SemiMech	Semi-mechanistic	Multi-country	Link
epiforecasts-EpiNow2	Semi-mechanistic	Multi-country	Link
prolix-euclidean	Semi-mechanistic	Multi-country	Link
ITWW-county_repro	Semi-mechanistic	Single-country	Link
UpgUmibUsi-MultiBayes	Semi-mechanistic	Single-country	Link
EuroCOVIDhub-baseline	Statistical	Multi-country	Link
MUNI-ARIMA	Statistical	Multi-country	Link
MUNI-LaggedRegARIMA	Statistical	Multi-country	Link
MUNI-VAR	Statistical	Multi-country	Link
RobertWalraven-ESG	Statistical	Multi-country	Link
SDSC_ISG-TrendModel	Statistical	Multi-country	Link
UB-BSLCoV	Statistical	Single-country	Link
UNED-PreCoV2	Statistical	Single-country	Link

Model fitting

We fit a generalised additive mixed effects model using the `mgcv` package in R. Code is available at: <https://github.com/epiforecasts/model-structure-evaluation>.

We used the following model formula:

$\sim, \text{log_interval_score}, \text{Method} + \text{Targets} + s(\text{Incidence}) + \text{Trend} + \text{Horizon} + s(\text{Model}, \text{bs} = "re")$

Summary of estimated coefficients

Component	Term	Estimate	Std Error	t-value	p-value
A. parametric coefficients	(Intercept)	-1.722	0.102	-16.888	0.0000 ***
	Method2	0.050	0.160	0.309	0.7570
	Method3	0.213	0.174	1.224	0.2209
	Targets2	0.077	0.141	0.549	0.5830
	Trend2	0.610	0.010	61.686	0.0000 ***
	Trend3	0.524	0.010	55.083	0.0000 ***
	Horizon.L	0.521	0.007	73.913	0.0000 ***
	Horizon.Q	-0.021	0.007	-3.009	0.0026 **
	Horizon.C	0.005	0.007	0.707	0.4793
Component	Term	edf	Ref. df	F-value	p-value
	s(Incidence)	8.209	8.823	229.213	0.0000 ***

Component terms	Term	Estimate	Std Error	t-value	p-value
	s(Model)	29.898	31.000	87.123	0.0000 ***
Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05					

Adjusted R-squared: 0.159, Deviance explained 0.160

-REML : 108273.655, Scale est: 0.928, N: 78277

Key to categorical variables (* indicates reference level):

- Method: 1* = Mechanistic, 2 = Semi-mechanistic, 3 = Statistical
- Target: 1* = Single country, 2 = Multi-country
- Trend: 1* = Stable, 2 = Increasing, 3 = Decreasing
- Horizon: L indicates a linear fit, Q a quadratic and C a cubic fit

Table 3: ANOVA for parametric terms

	df	F	p-value
Method	2	0.770	0.463
Targets	1	0.301	0.583
Trend	2	2069.151	0.000
Horizon	3	1825.622	0.000

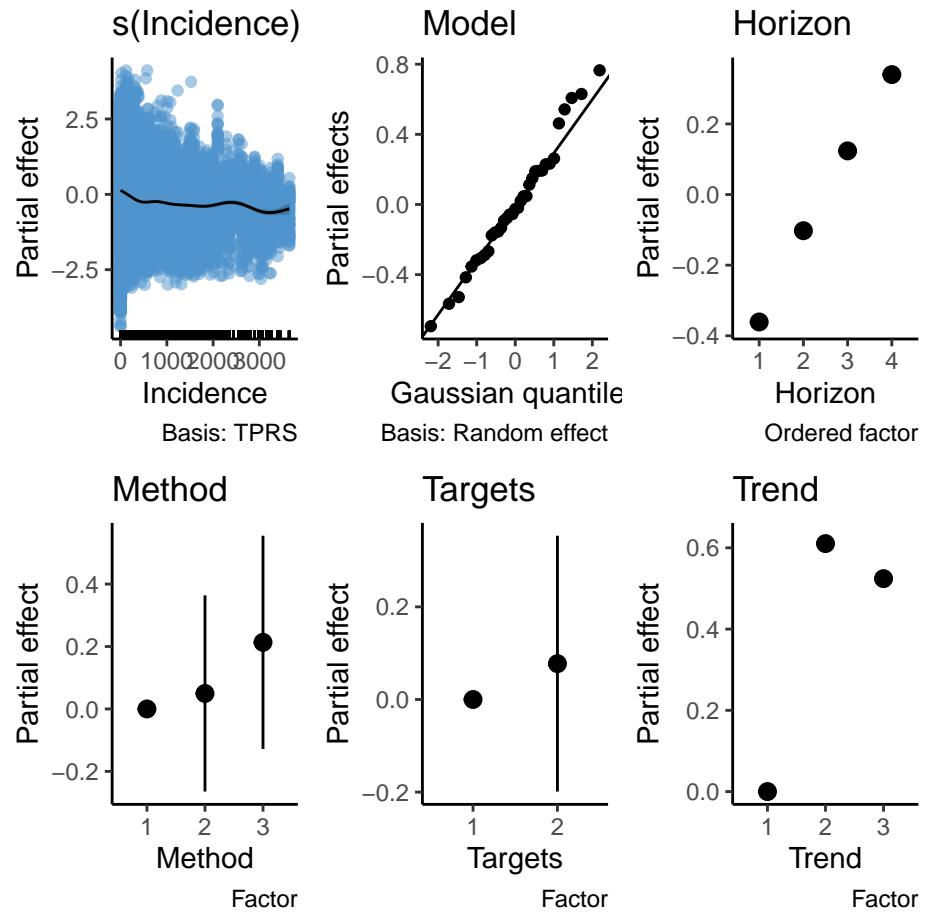


Figure 1: Partial residual plots for all terms

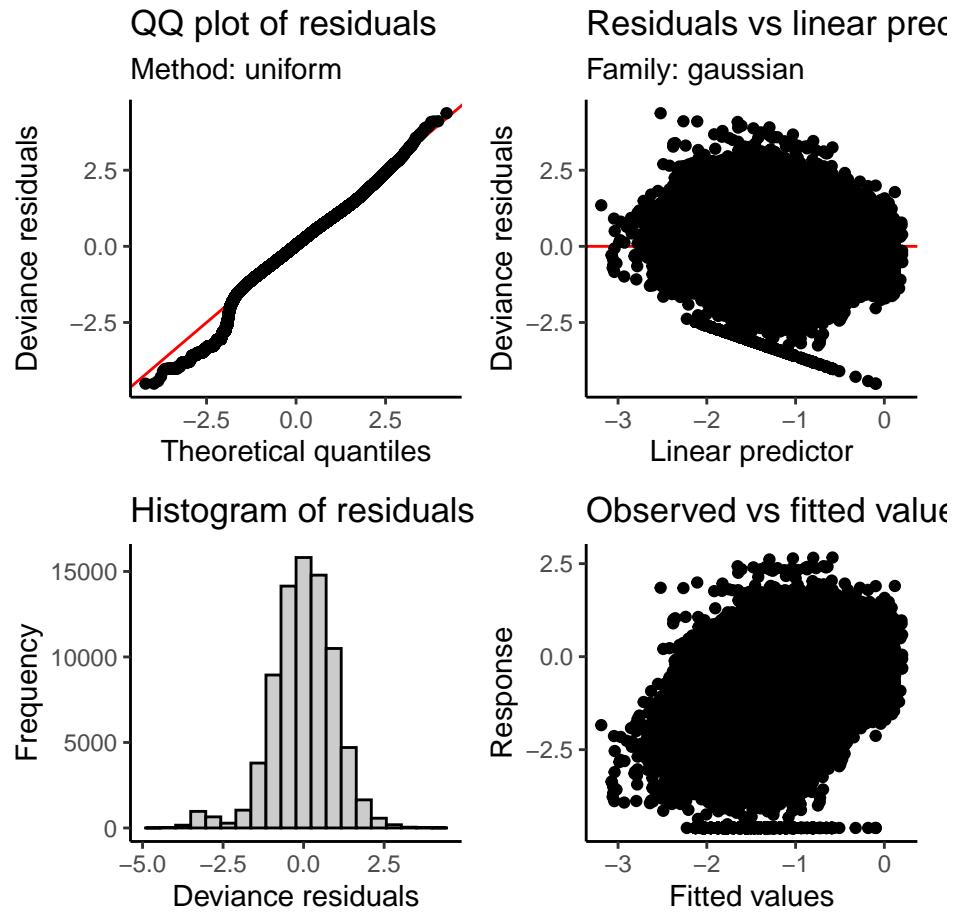


Figure 2: Model diagnostics