Probabilistic climate emulation with physics-constrained Gaussian processes

Shahine Bouabid[1] Dino Sejdinovic[2] Duncan Watson-Parris[3]

[1] Department of Statistics, University of Oxford
[2] School of Mathematical Sciences, University of Adelaide
[3] Scripps Institution of Oceanography, University of California San Diego

This project receives funding from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 860100

Motivation

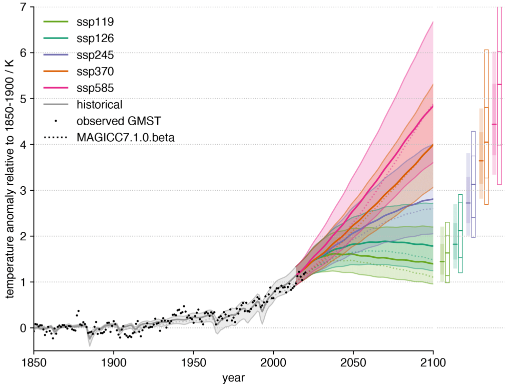


Figure 1: FaIRv2.0.0: a generalized impulse response model for climate uncertainty and future scenario exploration, Leach et al. (2021)

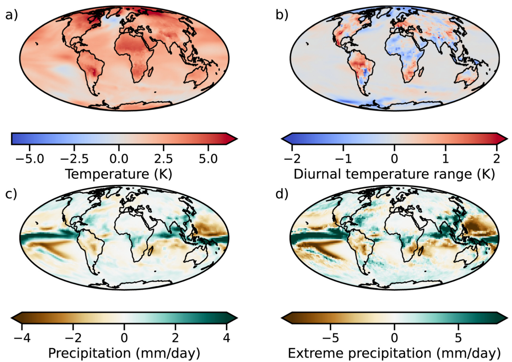
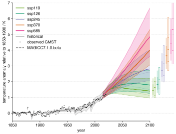


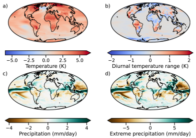
Figure 2: ClimateBench v1.0: A Benchmark for Data-Driven Climate Projections, Watson-Parris et al. (2021)

Motivation



Physics-driven emulation

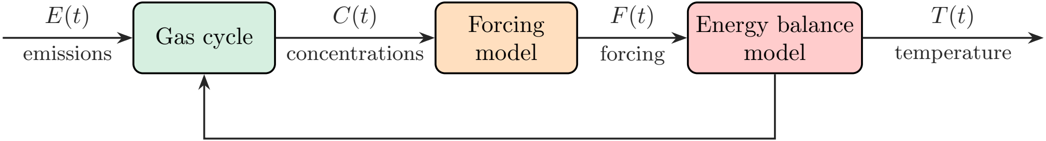
- $\oplus$ Robust intepretable physical modelling
- $\ominus$ Poor fit to some ESMs
- $\ominus$ Operates at global level



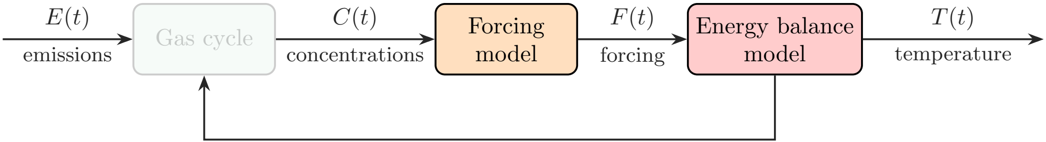
Data-driven emulation

- $\ominus$ Lack interpretability and robust physical grounding
- $\oplus$ Capture complex non-linear relationships from observations
- $\oplus$ Skilful spatial emulation

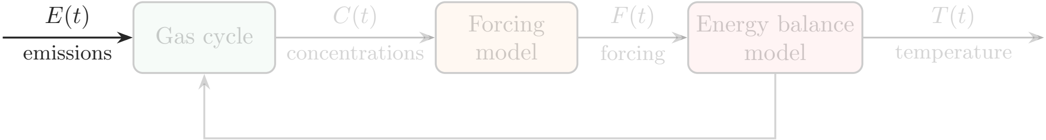
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)

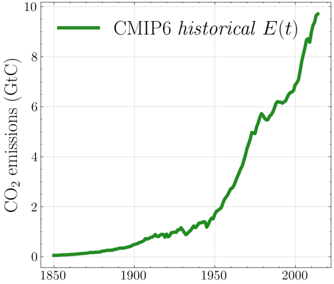


A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)

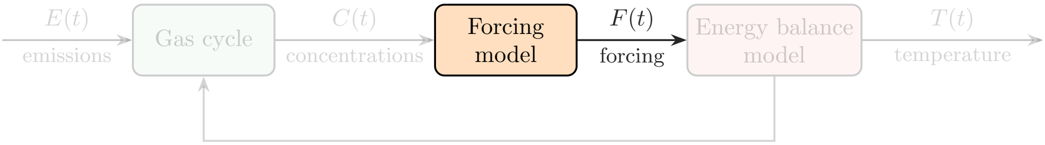


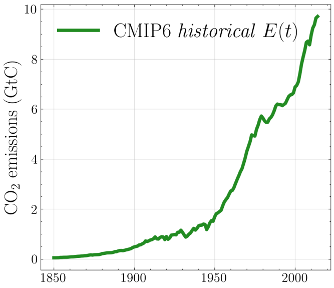
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)

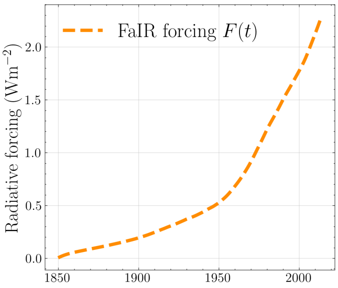




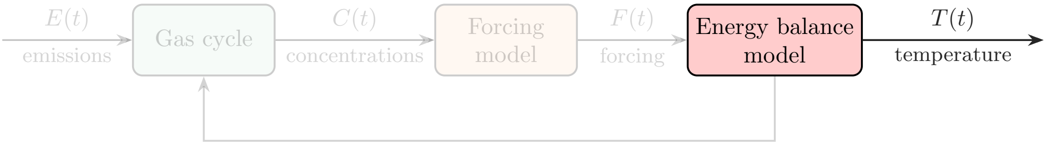
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)

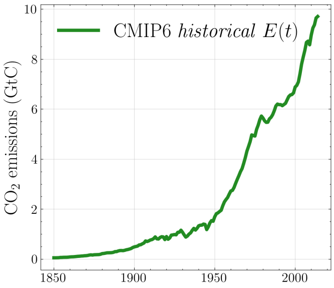


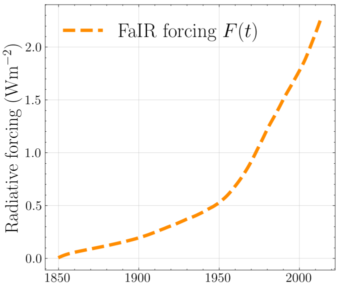


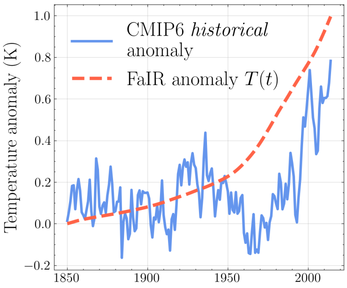


A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)

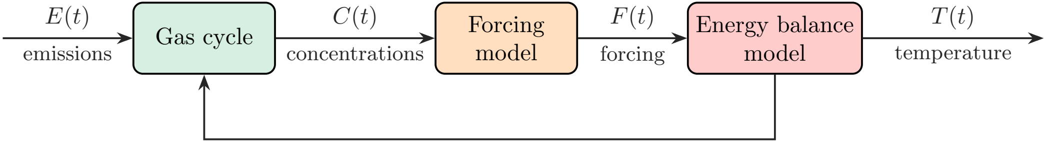


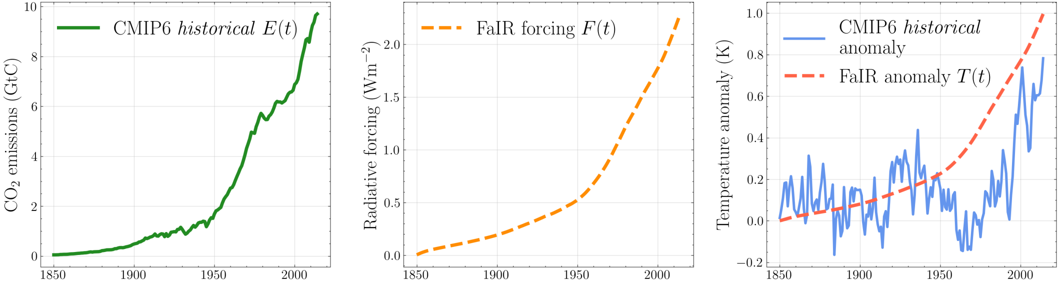




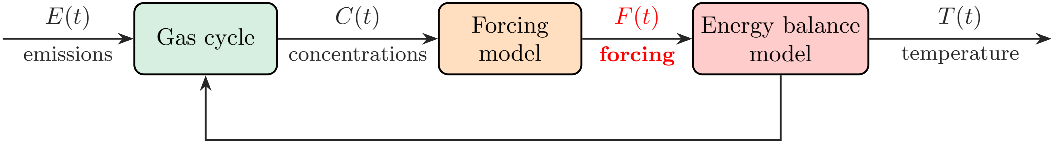


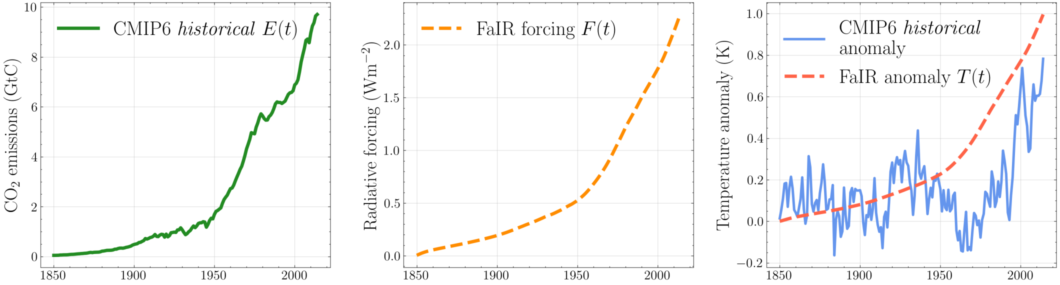
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)



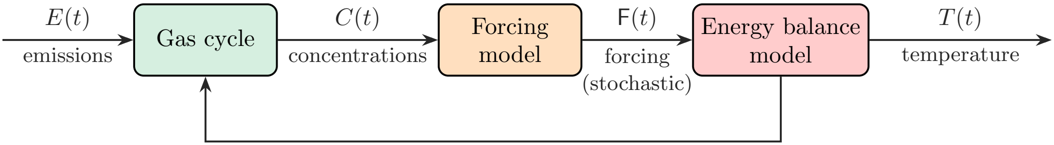


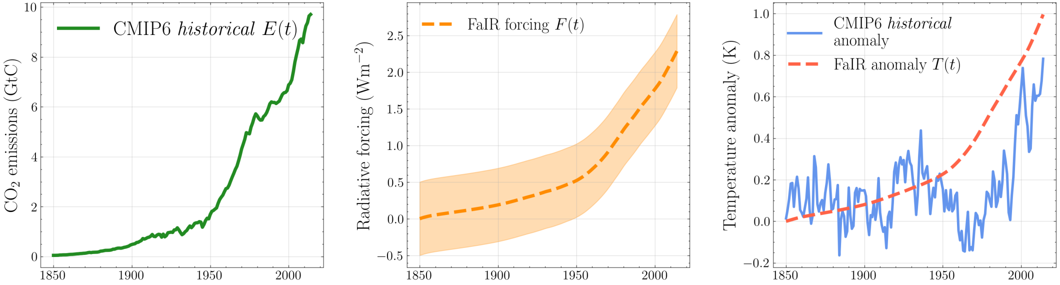
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)



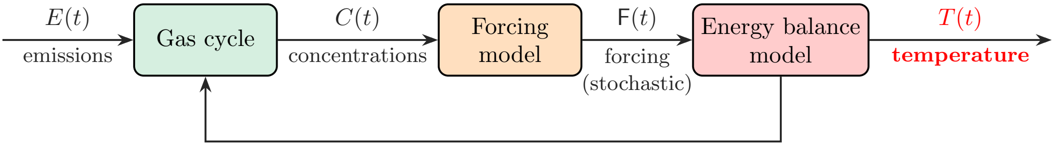


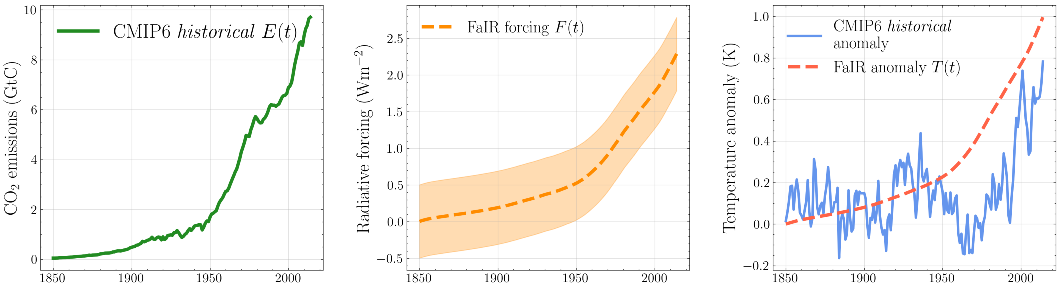
A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)



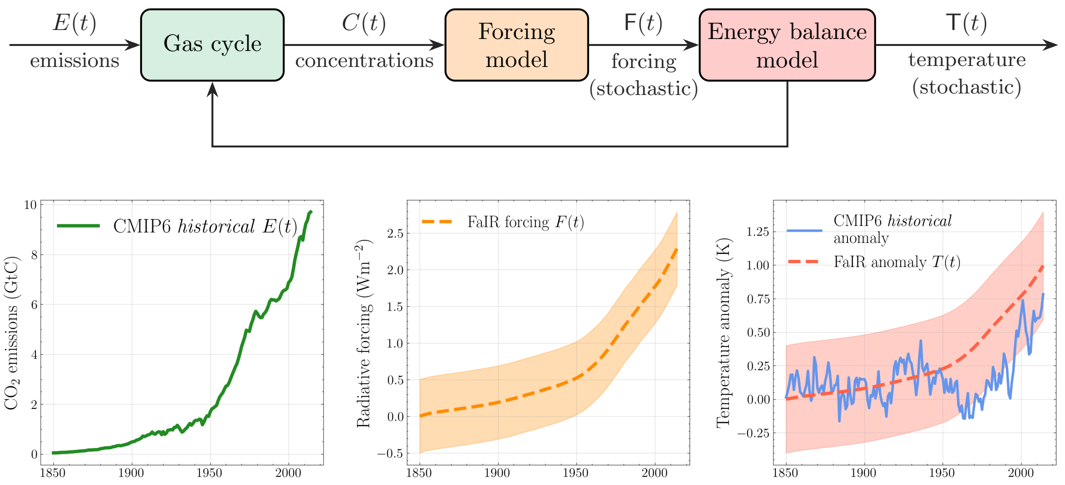


A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)





A physics-driven backbone: FaIRv2.0.0 (Leach et al., 2021)



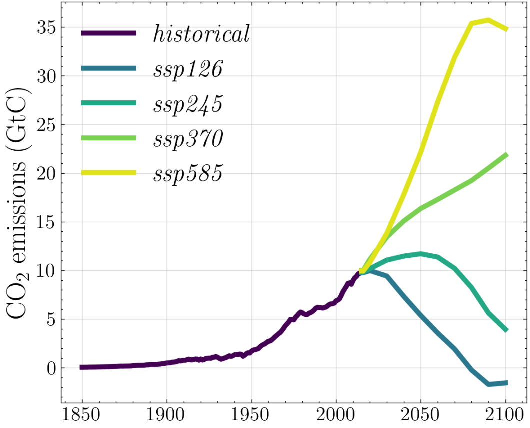
Solution to the energy balance model

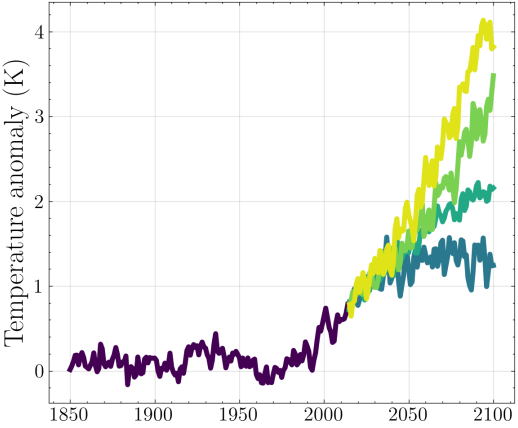
If $\mathsf{F}(t) \sim \mathrm{GP}(F, K)$ then $\mathsf{T}(t)$ is also a GP with

$$\mathsf{T}(t) \sim \mathrm{GP}\left(\sum_i m_i, \sum_{i,j} k_{ij}\right) \tag{1}$$

where

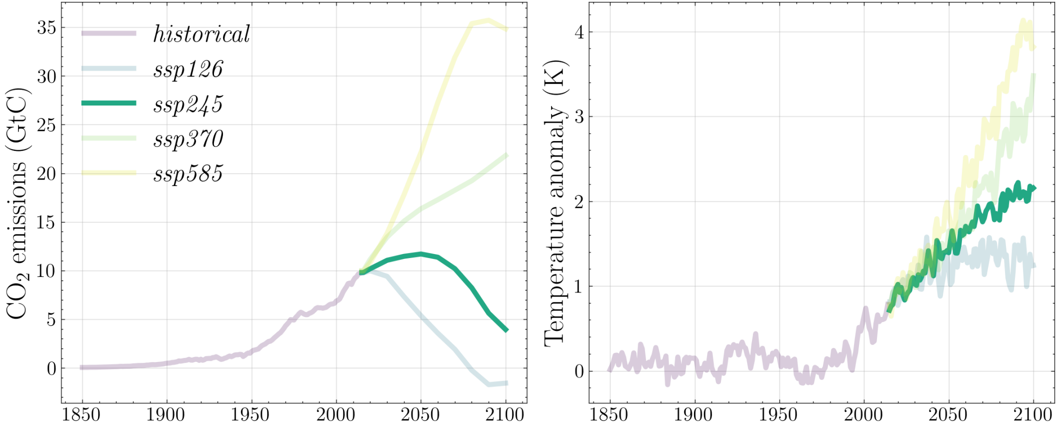
$$\begin{cases} m_i(t) = \dfrac{q_i}{d_i} \displaystyle\int_0^t F(s) e^{-(t-s)/d_i} \, \mathrm{d}s \\ k_{ij}(t, t') = \dfrac{q_i q_j}{d_i d_j} \displaystyle\int_0^t \int_0^{t'} K(s, s') e^{-(t-s)/d_i} e^{-(t'-s')/d_j} \, \mathrm{d}s \, \mathrm{d}s'. \end{cases} \tag{2}$$





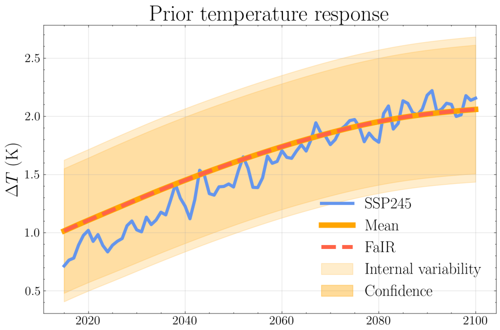
Watson-Parris et al. (2021)

Application: Shared socio-economic pathways emulation

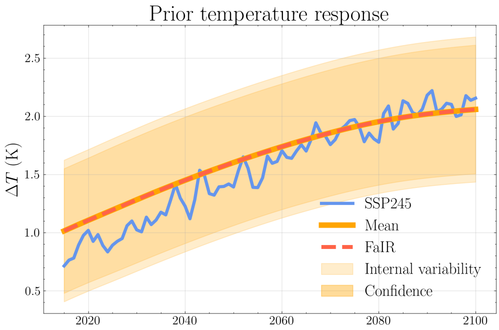


Watson-Parris et al. (2021)

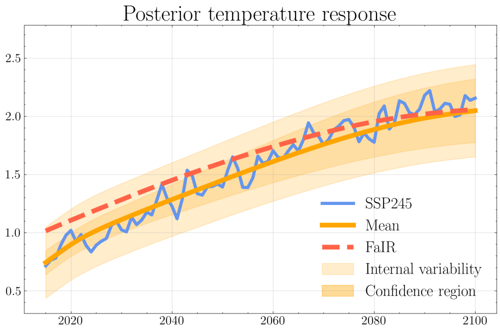
Application: Shared socio-economic pathways emulation



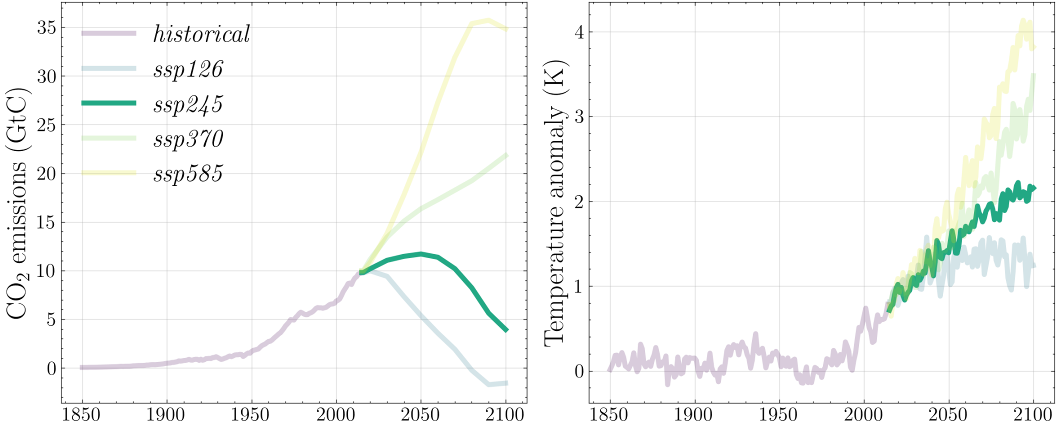
Application: Shared socio-economic pathways emulation



update with
historical, SSP126, SSP370, SSP585

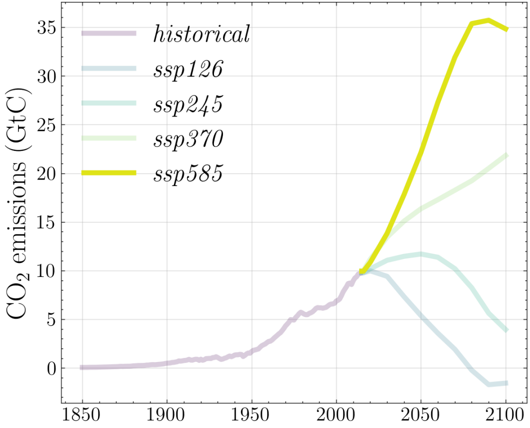


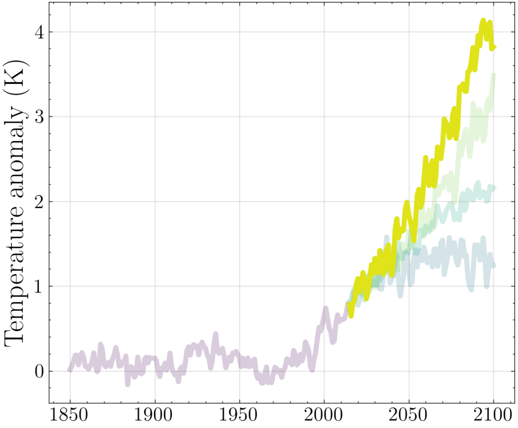
Application: Shared socio-economic pathways emulation



Watson-Parris et al. (2021)

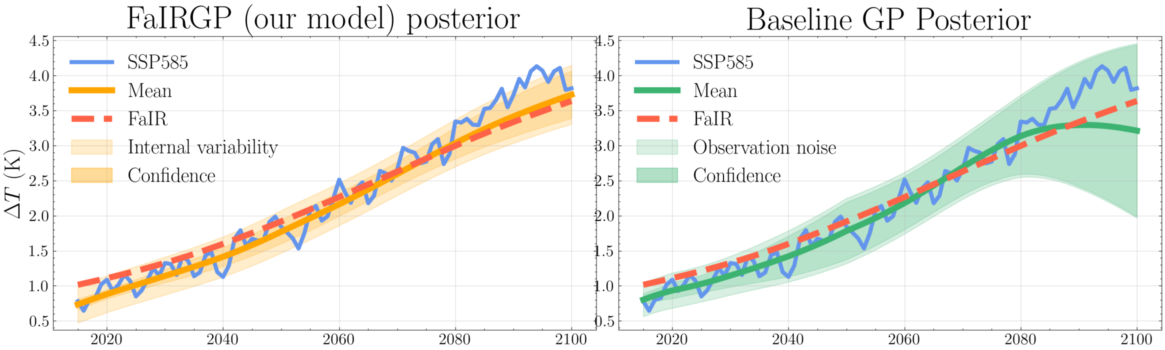
Application: Shared socio-economic pathways emulation

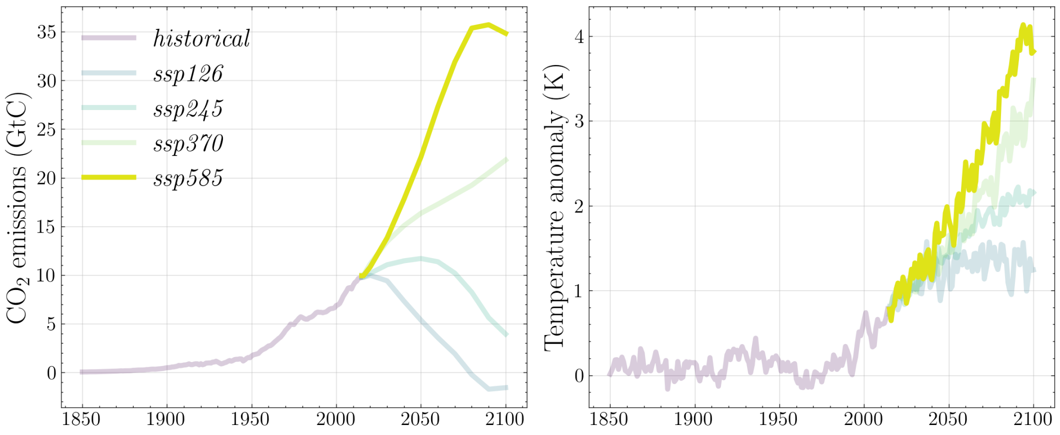




Watson-Parris et al. (2021)

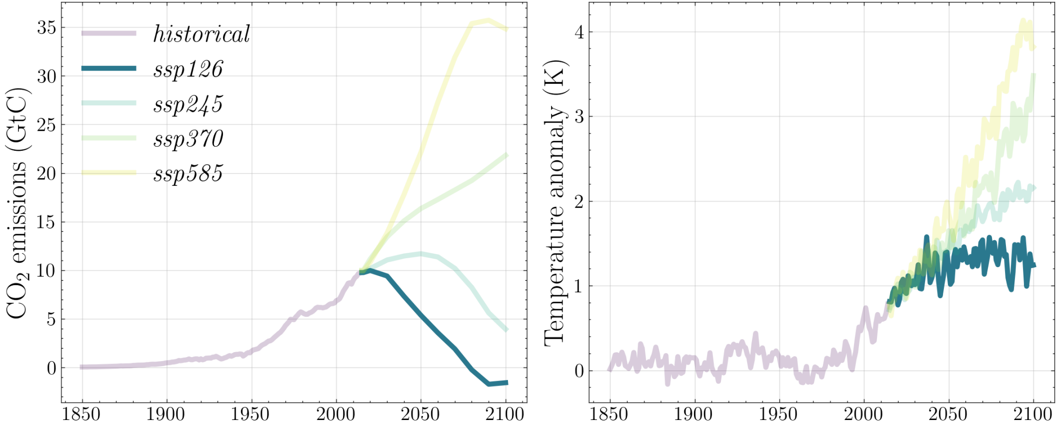
Application: Comparing with baseline GP





Watson-Parris et al. (2021)

Application: Shared socio-economic pathways emulation



Watson-Parris et al. (2021)

Which one is *SSP126*?

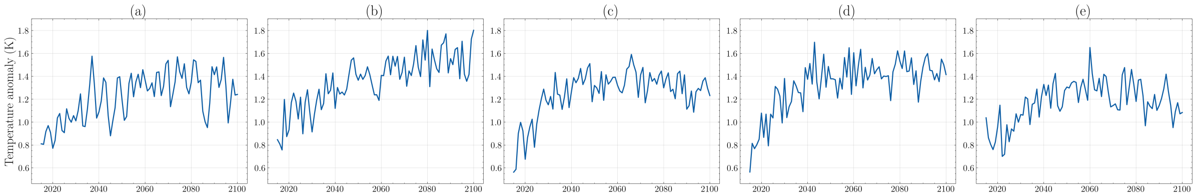
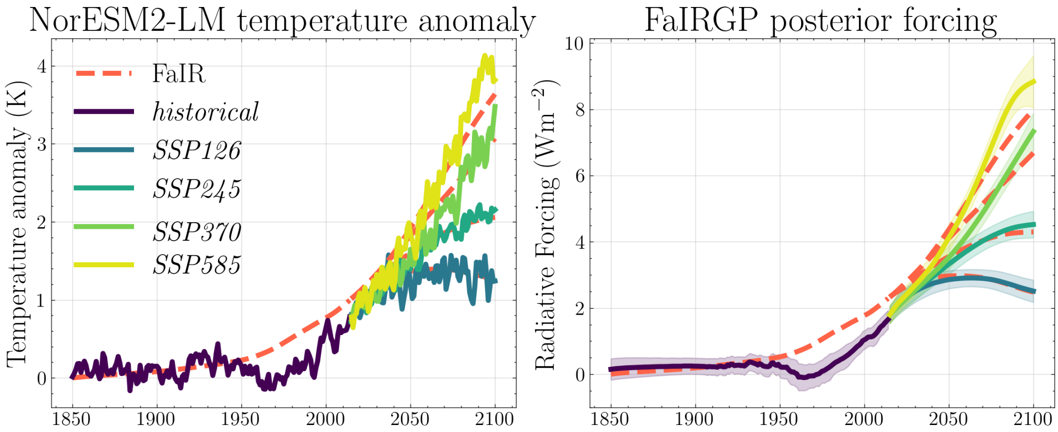


Figure 3: 4 draws from our model + the NorESM2-LM *SSP126*.

SSP emulation: it works!

Emulator	RMSE ↓	MAE ↓	Bias ↓	Log-likelihood ↑	Calib95
FaIR (physics-driven)	0.22±0.06	0.18±0.05	0.07±0.08	-	-
GP (data-driven)	0.20±0.09	0.15±0.06	-0.04±0.11	0.30±0.25	1.0±0.0
FaIRGP (hybrid)	**0.16±0.05**	**0.13±0.04**	**-0.01±0.07**	**0.41±0.24**	**0.94±0.06**

Application: Emulating radiative forcing from temperatures



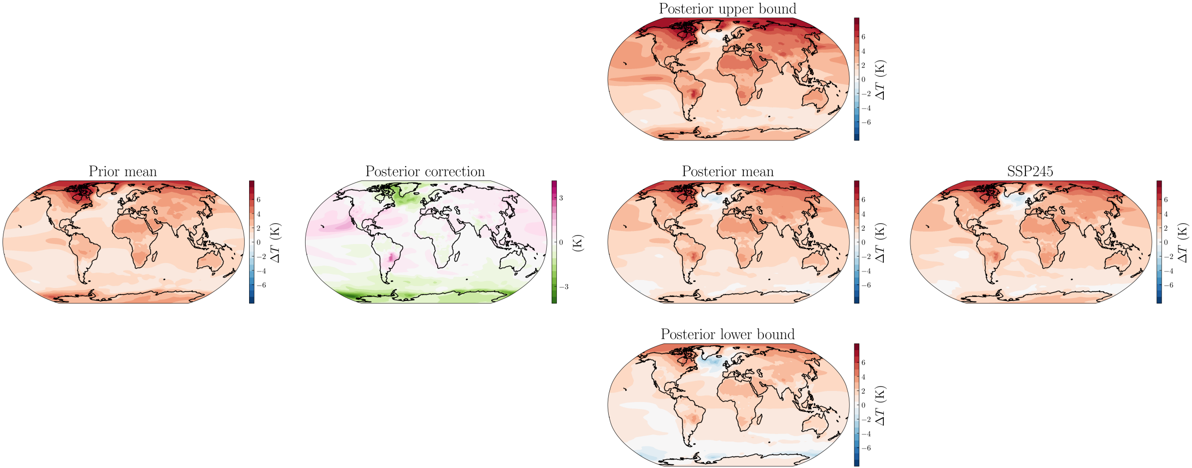


Figure 4: Spatial emulation of NorESM2-LM *SSP245* outputs. Maps are averaged over 2080-2100 period.

Outlook

- Bayesian version of an energy balance model
- Maintains robustness of the impulse response model
- Gains flexibility with possibility to inform with observations

Advantages of Gaussian process approach

- Principled uncertainty quantification (not sampling based)
- Allows to sample and evaluate likelihoods (analytical densities)
- Can naturally way to account for climate internal variability

Applications

- Spatially-resolved temperatures emulation
- Detection/attribution studies (analytical $\mathbb{P}(T > T^* \mid \text{scenario})$)
- Study the climate system (e.g. posterior on forcing)

N. J. Leach, S. Jenkins, Z. Nicholls, C. J. Smith, J. Lynch, M. Cain, T. Walsh, B. Wu, J. Tsutsui, and M. R. Allen. Fairv2.0.0: a generalized impulse response model for climate uncertainty and future scenario exploration. *Geoscientific Model Development*, 2021.

Duncan Watson-Parris, Yuhan Rao, Dirk Olivié, Øyvind Seland, Peer J Nowack, Gustau Camps-Valls, Philip Stier, Shahine Bouabid, Maura Dewey, Emilie Fons, and et al. Climatebench: A benchmark dataset for data-driven climate projections. *Earth and Space Science Open Archive*, 2021.

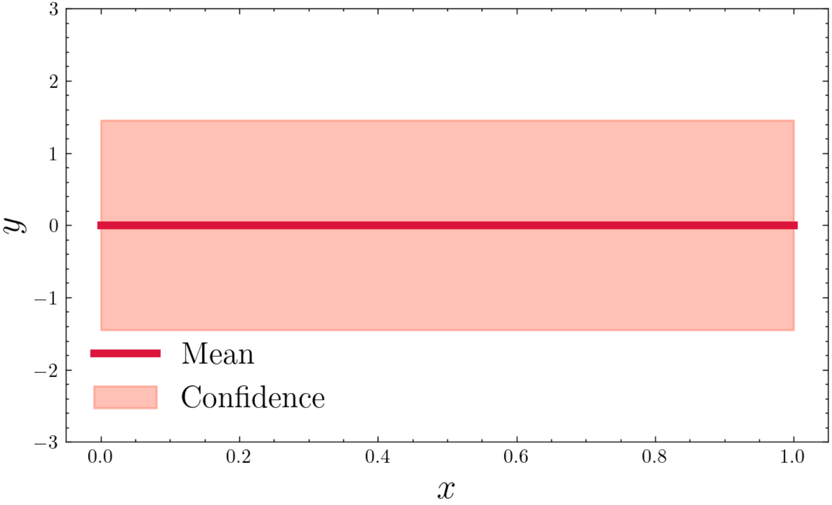
R. J. Millar, Z. R. Nicholls, P. Friedlingstein, and M. R. Allen. A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions. *Atmospheric Chemistry and Physics*, 2017.

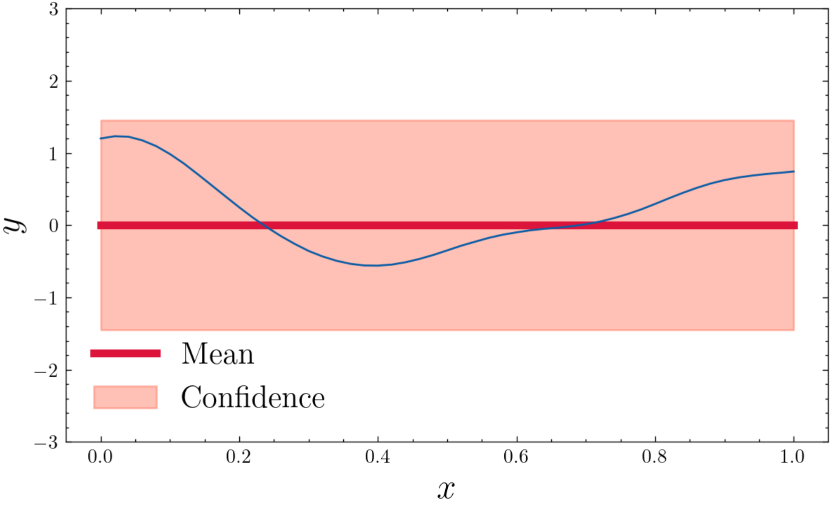
- GPs are Bayesian prior over classes of functions (generalisation of Gaussian random variables to functions)
- A GP $\mathsf{f}(x)$ is a stochastic function which is fully characterised by its *mean function* $m(x)$ and *covariance function* $K(x, x')$.

$$m(x) = \mathbb{E}[\mathsf{f}(x)] \qquad K(x, x') = \mathrm{Cov}(\mathsf{f}(x), \mathsf{f}(x')). \tag{3}$$

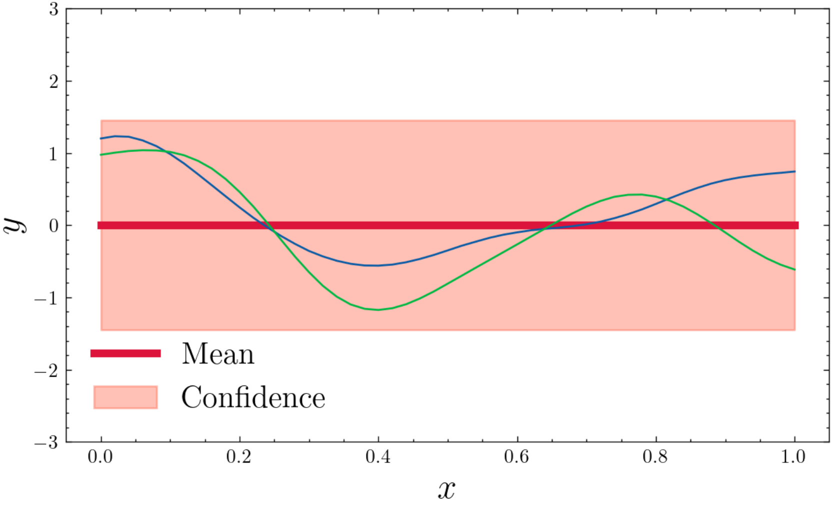
We denote,

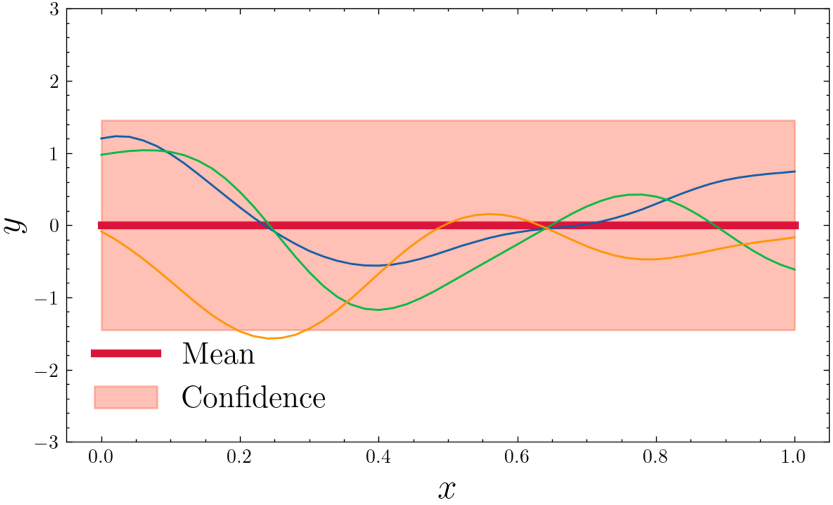
$$\mathsf{f}(x) \sim \mathrm{GP}(m, K) \tag{4}$$


3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence



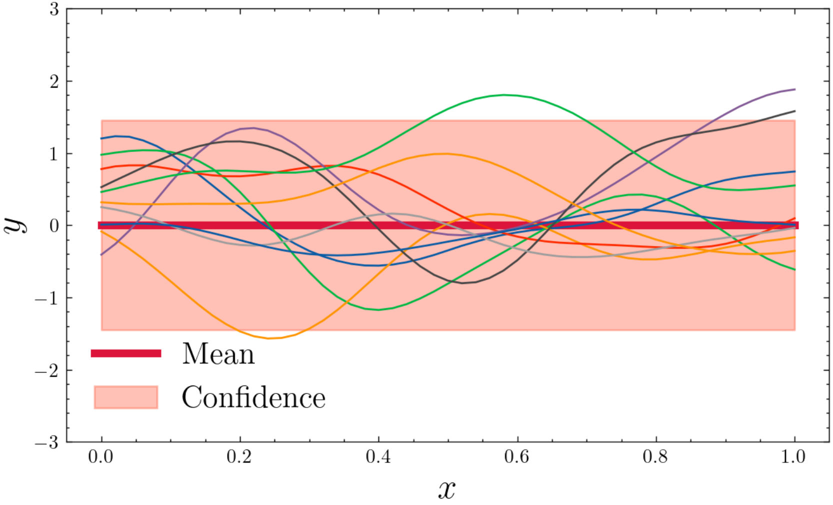
3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence

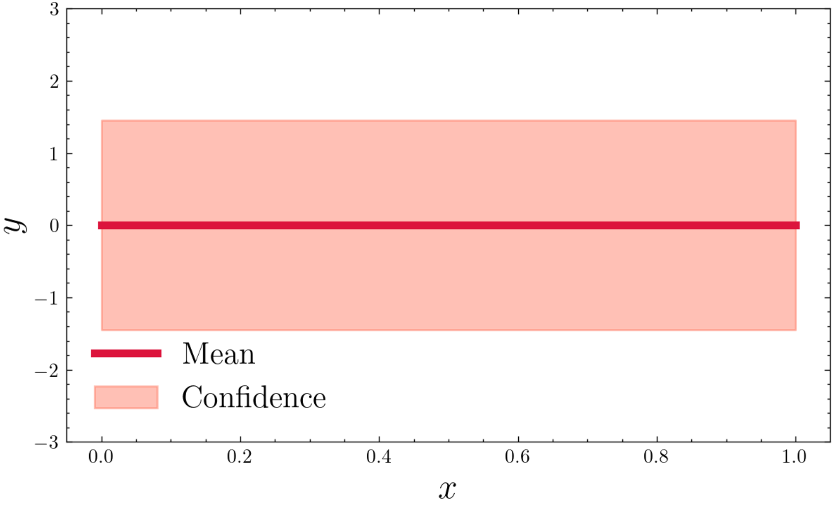

3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence

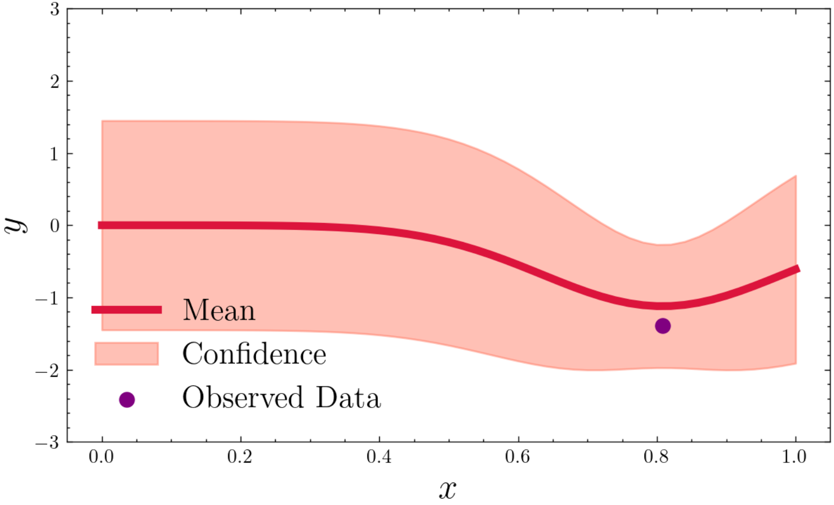


3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence

Gaussian processes (GPs)

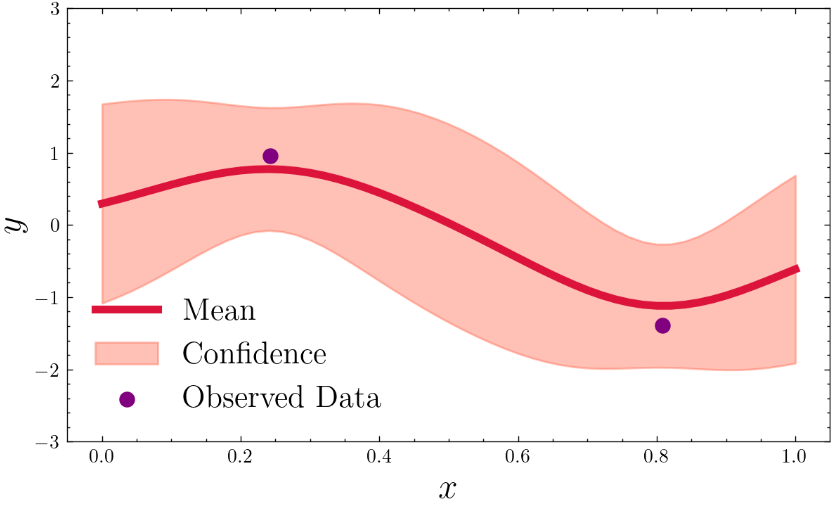



3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence

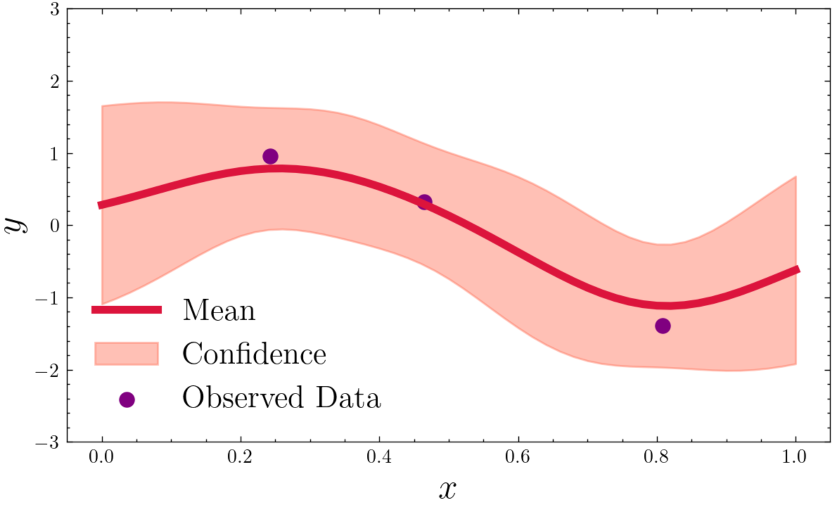


3
2
1
0
−1
−2
−3
y
0.0
0.2
0.4
0.6
0.8
1.0
x
Mean
Confidence
Observed Data

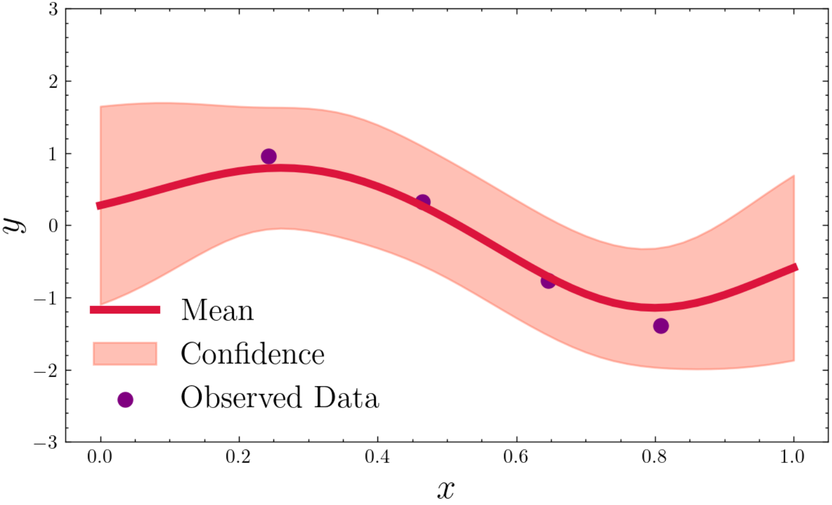
Gaussian processes (GPs)



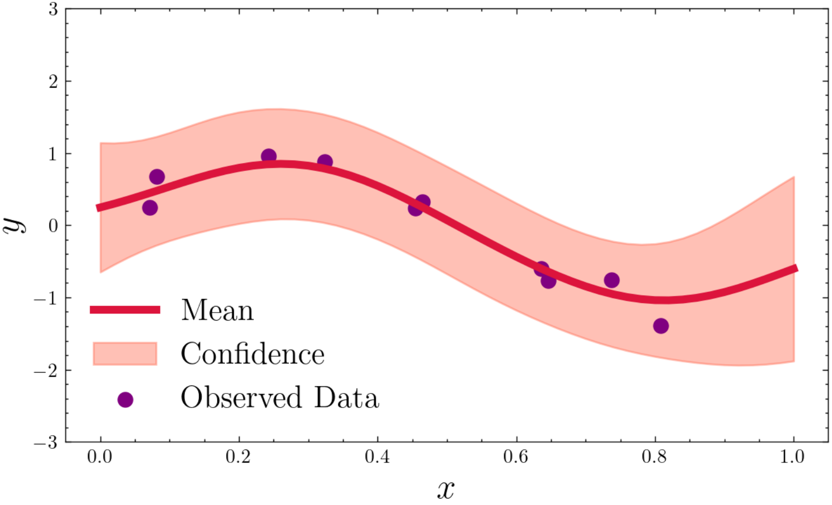
Gaussian processes (GPs)



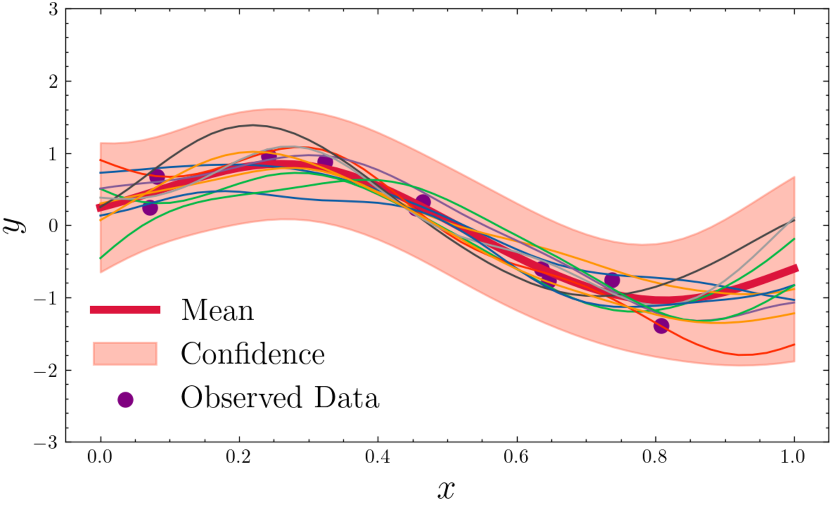
Gaussian processes (GPs)



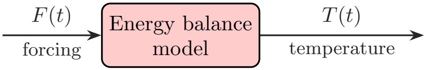
Gaussian processes (GPs)



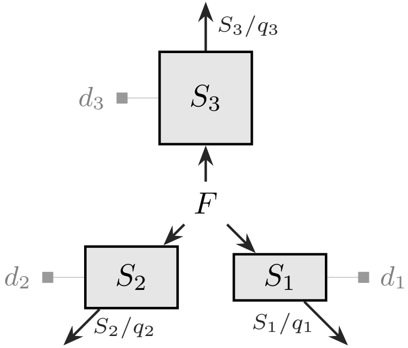
Gaussian processes (GPs)



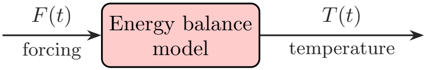
Thermal response model



$F(t)$	Radiative forcing (Wm^{-2})
d_i	Response timescale (years)
q_i	Equilibrium response (KW^{-1}m^2)
$S_i(t)$	Temperature of i^{th} thermal box (K)



Thermal response model



$F(t)$	Radiative forcing (Wm^{-2})
d_i	Response timescale (years)
q_i	Equilibrium response (KW^{-1}m^2)
$S_i(t)$	Temperature of i^{th} thermal box (K)

Thermal impulse response model

$$\frac{\mathrm{d}S_i(t)}{\mathrm{d}t} = \frac{1}{d_i}(q_i F(t) - S_i(t)) \qquad (5)$$

with $T(t) = \sum_i S_i(t)$.