



University  
of Exeter

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



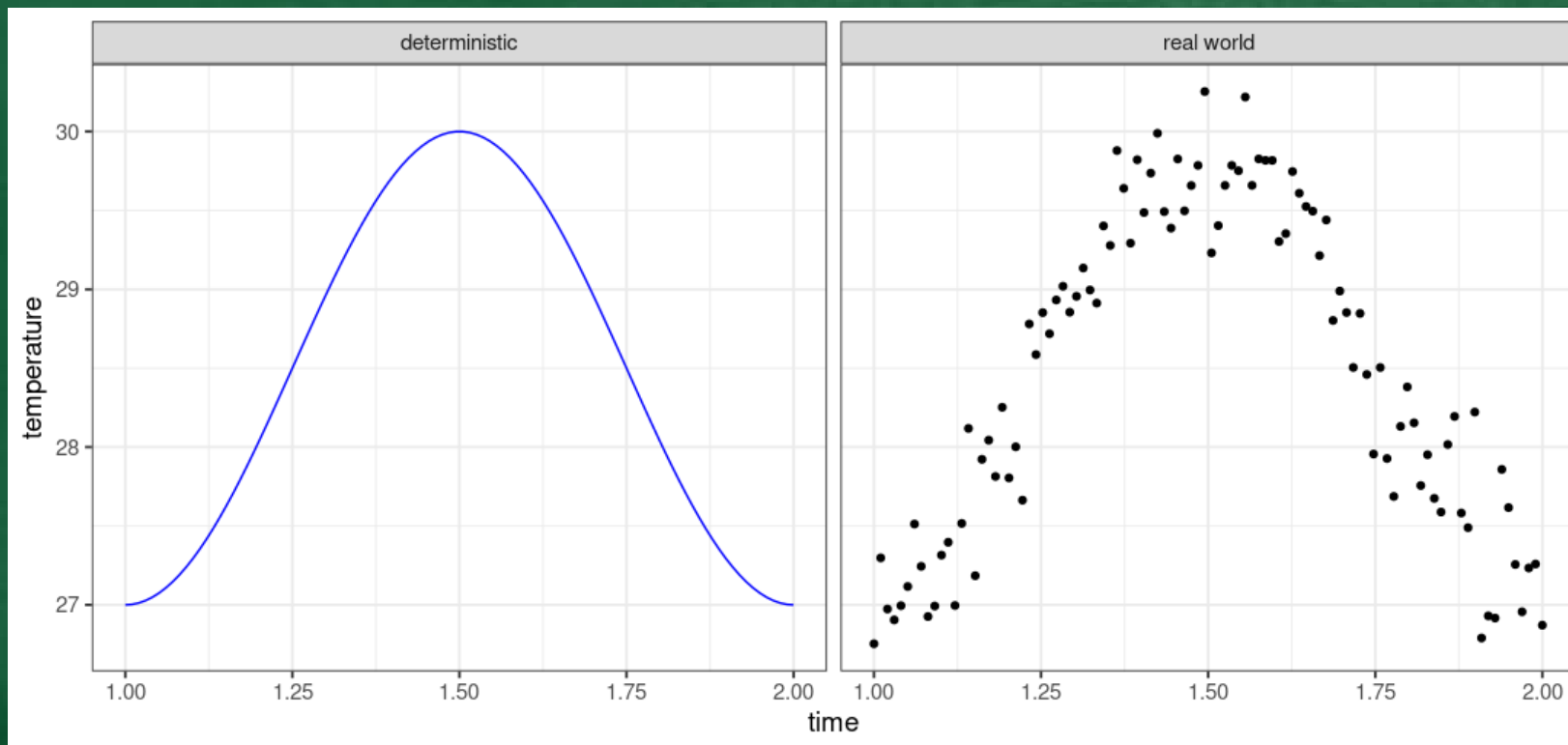
# Introduction to Statistical Data Analysis in R

Workshop CERMEL (21 July – 03 August, 2023)

Mario Recker

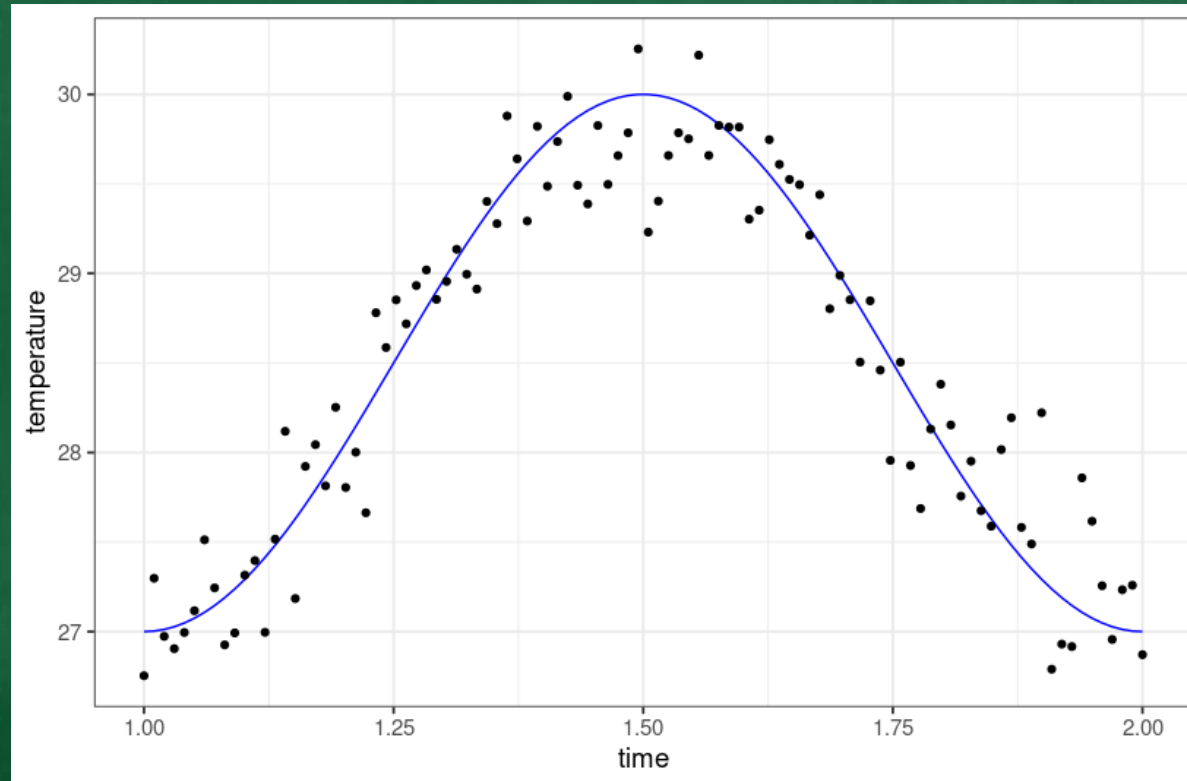
# Why statistical modelling?

- Mathematical models can help us to *describe*, *understand* and *predict* the world
- Although the process that we are trying to model might exhibit a systematic pattern that we can describe *deterministically*, the real world rarely matches such patterns exactly



# Why statistical modelling?

- Statistical models combine deterministic descriptions of systematic patterns with stochastic (probabilistic) descriptions of unexplained variation
- Quantifying the unexplained variation is important if we want to make useful assessments of uncertainty and risk



# Why R?

(taken from <https://www.r-bloggers.com/2014/03/why-use-r-five-reasons/> )

- R is free
- R is popular
- R is powerful
- R is flexible
- R is well supported

# Workshop structure

- Mix of lecture-like presentation / live-coding and hands-on practicals
- Plenty of opportunities to test your understanding
- Training datasets are provided and can be downloaded from <https://statistics.ecovir.net/download.php> but you are welcome to analyse your own data!
- **Please ask questions!**

# Workshop structure

## (1) General introduction to R/RStudio

- *variables and data frames*
- *functions*
- *basic plotting*

## (2) Data handling and visualisation

- *importing / exporting data*
- *visualisation with ggplot*
- *data wrangling*

## (3) Statistical analysis

- *descriptive statistics*
- *linear models*
- *mixed effect models*
- *generalised linear models*
- *survival analysis*

# Workshop structure

To access the workshop please go to

[https://mrecker.github.io/Intro\\_data\\_analysis\\_in\\_R/index.html](https://mrecker.github.io/Intro_data_analysis_in_R/index.html)

To download the (data) files needed for this workshop please visit

<https://statistics.ecovir.net/download.php>