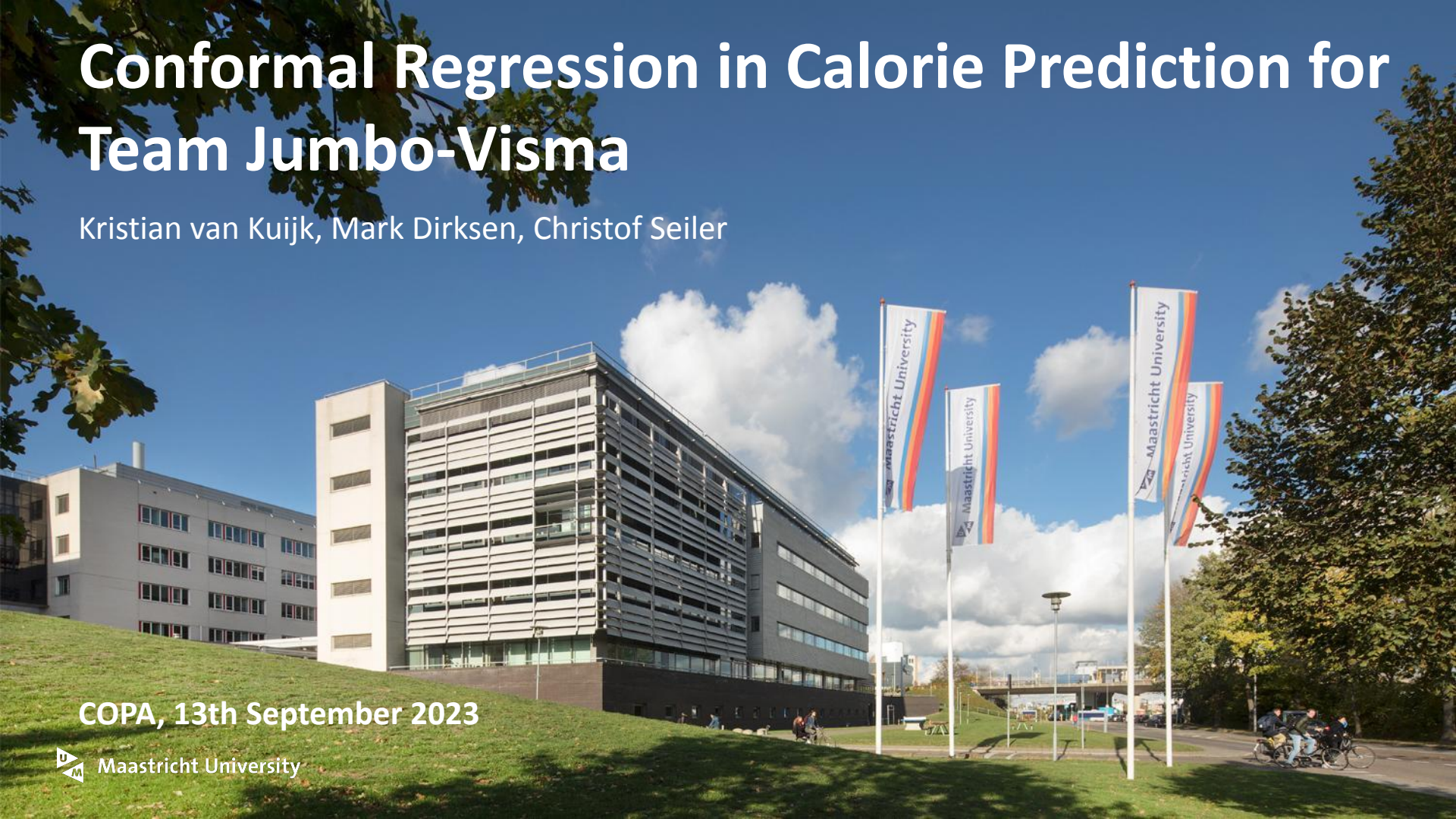
Conformal Regression in Calorie Prediction for Team Jumbo-Visma

Kristian van Kuijk, Mark Dirksen, Christof Seiler



COPA, 13th September 2023

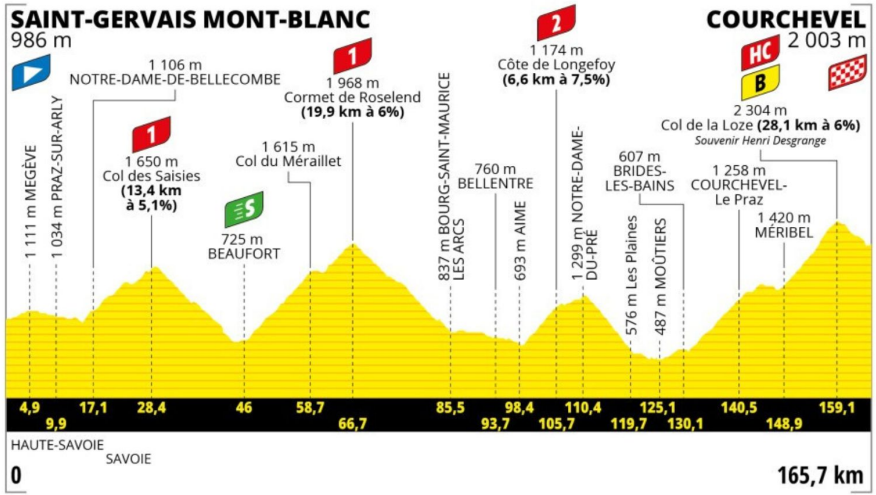
Maastricht University

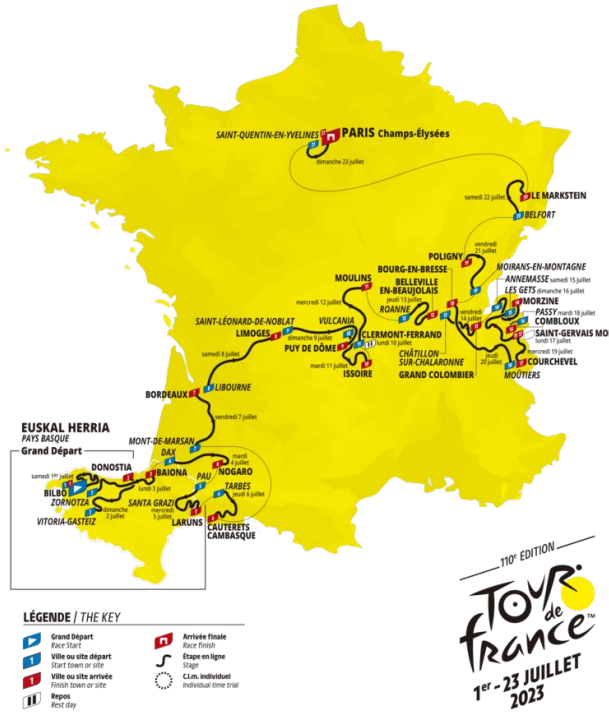
Imagine taking part in the Tour de France

165.5 km featuring 5100 metres of elevation

finish the stage in less than 5 hours to have a chance of winning

Tour de France is 23 days.





A game of Watts

Cycling power is the rate at which cyclists expend energy, which is obtained from the food they consume

Tour de France winner	Average recreational cyclist
325 watts for 80 hours	300 watts for 20 minutes

Nutrition is a number’s game

Tour de France cyclists expend around **120,000 calories** (roughly 210 big macs) overall (6,000 calories per stage)

The coaches have long been responsible for predicting the energy needs

Predicting energy needs has relied on judgement and experience of coaches

Team Jumbo-Visma

Winner of the Tour de France 2023 with Jonas Vingegaard

Winner completed the 3401 km in 82h 05' 42''. The runner up was just + 00h 07' 29'' behind, a difference of 0.2%

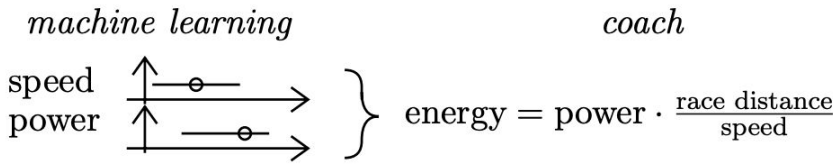
Difference between winning and losing comes down to the smallest details and margins

Why Conformal Prediction?

Coaches still tune the output predictions

More beneficial to predict a range of possibilities as coaches tend to tweak the models' output based on knowledge and previous experience for specific races

To achieve this, we employ methods from the conformal prediction framework

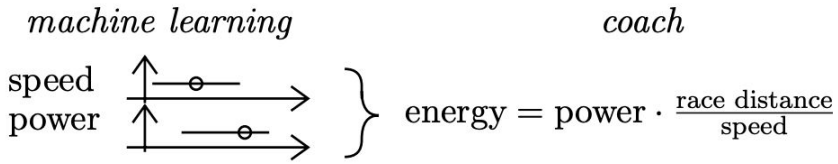


A concrete example of the 2022 season

Long-term power forecast bounds were [213, 265]

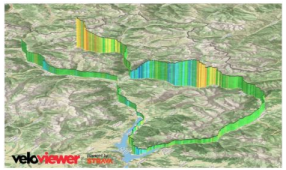
Predicted power of 245.17

Planned tactic + previous experience with this race, round the power to 250 watts. Combined with the predicted race time of 384 minutes, calorie forecast of 5760 kilocalories.



Data

FUSION SPORT



DARK SKY

PCS
PRO CYCLING STATS

Data

Independent variables	Dependent variables
Race type **Stage profile** Ascent Descent Distance **Weather conditions** Temperature Humidity negative wind-effect rainfall **Attributes of the riders** BMI **Tactics** roles	Speed Power

Weather

Forecasts performed daily

Considering only one weather forecast for a race that takes place in more than 10 days is suboptimal

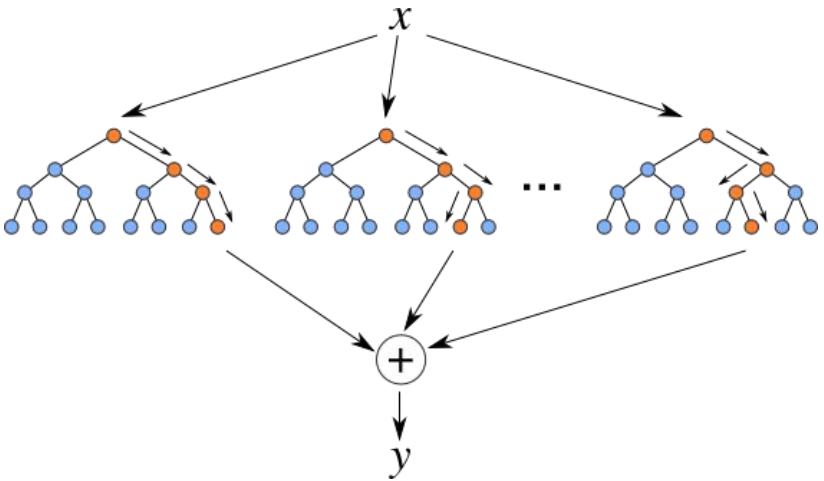
We assign weights for short-term forecasting based on how many days in advance the forecast is produced (NASA 2022)

Days to race	Weights weather model	Weights without weather model
10 days	0.5	0.5
5 days	0.9	0.1

Approach

Random forest model as the underlying regressor for both the power and speed

The energy is then obtained by multiplying both outputs. Allows to better understand predictions and tweaking them



Conformal Prediction

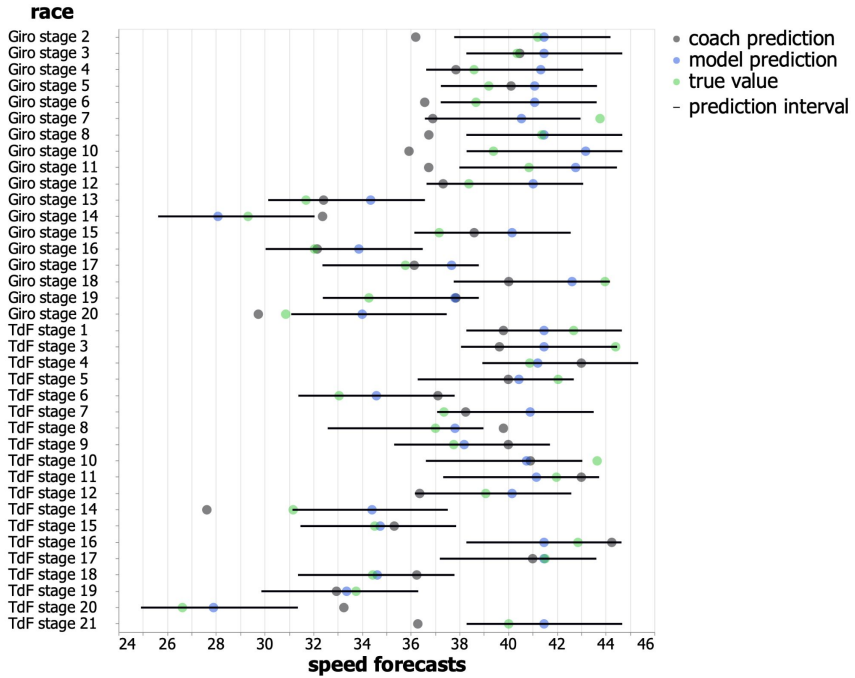
jackknife and its variations (jackknife+, jackknife-minmax, jackknife+-after-bootstrap, and jackknife -afterbootstrapminmax) (Barber et al. 2021)

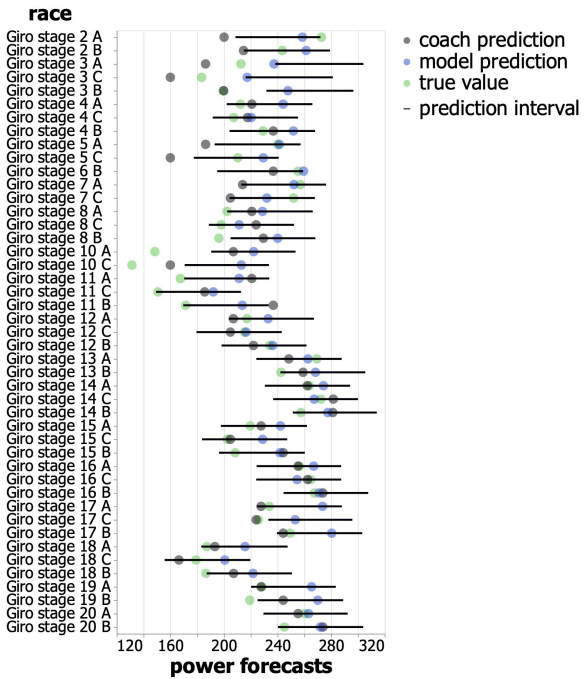
cross-validation (CV) and its variations (CV+ and CV-minmax) (Barber et al. 2021)

conformalized quantile regression (CQR) (Yaniv Romano et al. 2019)

inductive conformal prediction (ICP) (Vovk et al. 2012)

Coaches and Machine Learning comparison





Thank you!

References

Rina Barber, Emmanuel Candes, Aaditya Ramdas, and Ryan Tibshirani. Predictive inference with the jackknife+. Annals of Statistics, 49:486–507, 02 2021. doi: 10.1214/ 20-AOS1965.

Yaniv Romano, Evan Patterson, and Emmanuel Candes. Conformalized quantile regression. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alche-Buc, E. Fox, and R. Garnett, editors, Advances in Neural Information Processing Systems, volume 32. Curran Associates, Inc., 2019.

V. Vovk. Conditional validity of inductive conformal predictors. Proceedings of the Asian Conference on Machine Learning, PMLR 25:475-490, 2012.

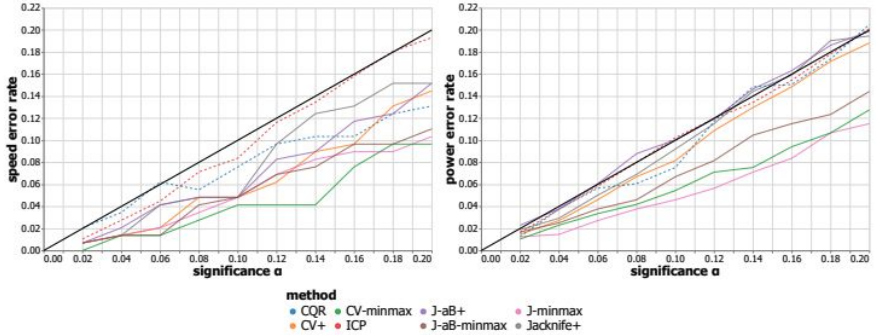
NASA. How reliable are weather forecasts?, Jun 2022. URL https://scijinks.gov/ forecast-reliability/.

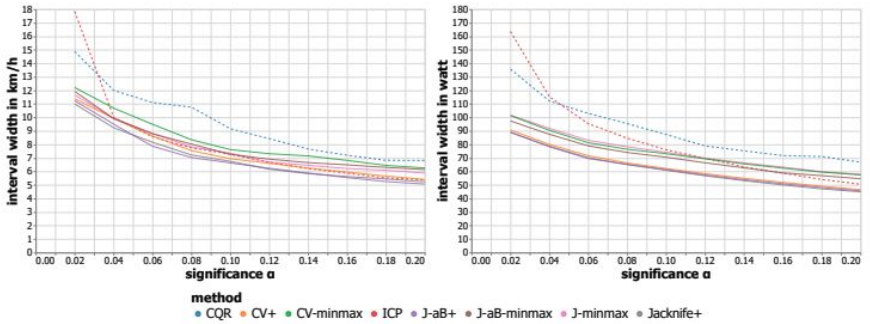
Appendix

For all the experiments, we repeat five-fold cross-validation five times and report the average. All experiments are performed on an Intel i7 with 8 CPU cores at 3GHz and 16GB of RAM.

As significance levels larger than 0.20 are very unusual, since the error rate becomes too large for the prediction intervals to be used in practice, all figures only include $\alpha \leq 0.20$. To differentiate constant and non-constant interval size prediction intervals methods, the two methods computing non-constant interval size prediction intervals (CQR and ICP methods) are depicted by dashed lines.

Appendix


speed error rate
power error rate
significance α
method
CQR
CV-minmax
J-aB+
J-minmax
CV+
ICP
J-aB-minmax
Jacknife+


interval width in km/h
interval width in watt
significance α
method
CQR
CV+
CV-minmax
ICP
J-aB+
J-aB-minmax
J-minmax
Jacknife+