Post-doctoral proposal: Conformal inference for drone trajectories forecast

Fabrice Gamboa* Adrien Mazoyer †

May 12, 2023

Scientific context and place

This position is a part of the ANR Project ASTRID-ROMEO dealing with the integrity of drone sheep. It is dedicated to the development of robust methods for the detection of outliers in drones trajectories, combinig physics-constrained neural network and uncertainty quantification using conformal inference methods. Thales Research and Technology (TRT) brings its expertise for the first part, whereas Toulouse Institute of Mathematics (IMT) carries out the second one.

The post-doctoral trainee will mainly take place at IMT, with several meetings in Paris Saclay. It will be supervised by Fabrice Gamboa (IMT) and Adrien Mazoyer (IMT), and by TRT in particular for the programming aspects.

Description and objectives

Conformal inference

Consider n **exchangeable** (typically i.i.d.) pairs of random variables $(X_i, Y_i)_{i=1\dots n} \in \mathcal{X} \times \mathcal{Y}$ (hereinafter referred to as the *calibration set*) and a new input X_{n+1}. Our objective is to construct a prediction set for the output Y_{n+1} with an aimed level of confidence, given the calibration set and a predictor function $\hat{f} : \mathcal{X} \to \mathcal{Y}$. That is we aim to construct a function C defined on $(\mathcal{X} \times \mathcal{Y})^n \times \mathcal{X}$ that will return a conservatively valid subset of $\mathcal{Y}$ for Y_{n+1}:

$$\mathbb{P}\left[Y_{n+1} \in C((X_i, Y_i)_{i=1\dots n}, X_{n+1})\right] \geq 1 - \alpha$$

The conformal inference process consists in the following. Let us give a score function $S : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ and compute $R_i = S(\hat{f}(X_i, Y_i))$. For example, if $S(Y, Y') = \|Y - Y'\|^2$, then the R_i's correspond to the squared errors. We define then $Q(1-\alpha) = R_{(\lceil (n+1)(1-\alpha) \rceil)}$, with $R_{(1)} \leq \dots \leq R_{(n)}$ and finally construct

$$C(X_{n+1}) = \left\{ y \in \mathcal{Y} \quad \text{s.t.} \quad S(\hat{f}(X_{n+1}, y)) \leq Q(1-\alpha) \right\}$$

Therefore we can show that the above set is conservatively valid for Y_{n+1}, as long as $1/(n+1) \leq \alpha < 1$.

This approach requires only that the calibration data are exchangeable. This is why since the sequential former approach proposed in [16], it has been widely studied and extended, in particular in both machine learning and statistics [1, 7].

Conformal inference for time series

Applying conformal inference to time series has been very active these last years, in particular under an online context, (data arrives in a sequential way), see e.g. [8, 9, 17, 2]. Despite the non-exchangeability of data, these works, that we can gather as *adaptive conformal inference* methods, achieve to provide conservative sets, but in an asymptotical and very specific way: the coverage is guaranteed over the entire time horizon, or over sub-intervals of time. Moreover, these methods may return as prediction the entire set or a trivial set. These drawbacks make such methods unsatisfying for some applications, in particular when it involves safety of person or material, such as self-driving cars or drones.

Instead of considering the previous locations of the current trajectory for predicting the future positions, an other context relies on a exchangeable calibration set of time series [14, 11, 13, 3]. In other

*Institut de Mathématiques de Toulouse. fabrice.gamboa@math.univ-toulouse.fr
†Institut de Mathématiques de Toulouse. adrien.mazoyer@math.univ-toulouse.fr

words, given $\left(Z^{(1)},\ldots,Z^{(n)}\right)$, where each $Z^{(i)}$ consists in $Z^{(i)}=\left(Z_0^{(i)},\ldots,Z_T^{(i)},\ldots,Z_{T+h}^{(i)}\right)$, and a trajectory $Z^{(*)}=\left(Z_0^{(*)},\ldots,Z_T^{(*)}\right)$, the objective is be to constuct a conformal prediction set for the next positions $\left(Z_{T+1}^{(*)},\ldots,Z_{T+h}^{(*)}\right)$. The main difficulty then is to choose a relevant score function, in order to get informative prediction sets.

In the specific context of drones trajectories prediction, the state of "outlier" might depend on specific conditions, such as weather conditions for example. It is possible then to apply a rough stratification with different scenarios, but the authors of [15] have proposed the usage of conformal inference when the distribution of the new input is different from the calibration set, in the case of a covariate shift, which could be extended in our time series context.

Objectives

The first objective of the post-doctoral trainee is to implement a module performing conditional conformal prediction in a trajectory predictor. Depending on the profile of the selected candidate, they will also explore alternative possibilities in the latter context. Here are some examples of potential axes (non-exhaustive list):

- The usage of (ensemble or particular) Kalman filters for conformal inference appears at the moment only in the context of online data for the definition of the scores in applied fields ([5, 12]), but not yet in exchangeable context;
- Some works about the usage of conformal methods for functional data can be found (see e.g. [10, 4]), but the usage of signature (see [6] for usage of signature in functional context) for conformal inference remains quite open.
- Since physical constraints neural network are considered in the ROMEO project for the prediction of drones trajectories, it might be interesting to study ways to embed physical constraints in the conformity score.

Wished profile, salary, duration

The ideal candidate will have confirmed skills both in statistics and programming. We are looking for a doctorate owner in statistics with skills in programming, or a doctorate owner in computer sciences with apetancy for statistics. The development will be done in Python.

This is a contract of **one year**, that can start any time after October 2023.

The annual salary is around 30k€ after taxes, including prime.

If you feel interested please get in touch by e-mail with F. Gamboa and A. Mazoyer.

References

[1] V. N. Balasubramanian, S.-S. Ho, and V. Vovk (Eds.) *Conformal Prediction for Reliable Machine Learning: theory, adaptations and applications*. Morgan Kaufmann, 2014.

[2] A. Bhatnagar, H. Wang, C. Xiong, and Y. Bai. "Improved Online Conformal Prediction via Strongly Adaptive Online Learning". arXiv preprint arXiv:2302.07869. 2023.

[3] M. Cleaveland, I. Lee, G. J. Pappas, and L. Lindemann. "Conformal Prediction Regions for Time Series using Linear Complementarity Programming". arXiv preprint arXiv:2304.01075. 2023.

[4] J. Diquigiovanni, M. Fontana, and S. Vantini. "Conformal prediction bands for multivariate functional data". In: *Journal of Multivariate Analysis* 189 (2022).

[5] A. Dixit, L. Lindemann, S. Wei, M. Cleaveland, G. J. Pappas, and J. W. Burdick. "Adaptive Conformal Prediction for Motion Planning among Dynamic Agents". (arxov preprint arXiv:2212.00278). 2022.

[6] A. Fermanian. "Functional linear regression with truncated signatures". In: *Journal of Multivariate Analysis* 192 (2022).

[7] M. Fontana, G. Zeni, and S. Vantini. "Conformal Prediction: a Unified Review of Theory and New Challenges". In: *Bernoulli* 29.1 (2023), pp. 1–23.

[8] I. Gibbs and E. Candès. "Adaptive Conformal Inference Under Distribution Shift". In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 1660–1672.

[9] I. Gibbs and E. Candès. "Conformal Inference for Online Prediction with Arbitrary Distribution Shifts". arXiv preprint arXiv:2208.08401. 2022.

[10] J. lei, A. Rinaldo, and L. Wasserman. "A conformal prediction approach to explore functional data". In: *Annals of Mathematics and Artificial Intelligence* 74 (2018), pp. 29–43.

[11] Z. Lin, S. Trivedi, and J. Sun. "Conformal Prediction Intervals with Temporal Dependence". In: *Transactions on Machine Learning Research* (2022).

[12] S. Park, O. Bastani, and T. Kim. "ACon2: Adaptive Conformal Consensus for Provable Blockchain Oracles". (arxiv preprint arXiv:2211.09330). 2023.

[13] F. Schlembach, E. Smirnov, and I. Koprinska. "Conformal Multistep-Ahead Multivariate Time-Series Forecasting". In: *Proceedings of the Eleventh Symposium on Conformal and Probabilistic Prediction with Applications*. Vol. 179. 2022, pp. 316–318.

[14] K. Stankevičiūtė, A. M. Alaa, and M. van der Schaar. "Conformal Time-Series Forecasting". In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 6216–6228.

[15] R. J. Tibshirani, R. Foygel Barber, E. J. Candès, and A. Ramdas. "Conformal Prediction Under Covariate Shift". In: *Advances in Neural Information Processing Systems*. Vol. 32. 2019.

[16] V. Vovk, A. Gammerman, and G. Shafer. *Algorithmic learning in a random world*. Springer, 2005.

[17] M. Zaffran, O. Feron, Y. Goude, J. Josse, and A. Dieuleveut. "Adaptive Conformal Predictions for Time Series". In: *Proceedings of the 39th International Conference on Machine Learning*. Vol. 162. 2022, pp. 25834–25866.