Research Statement

Mauricio Tec

Goal and ongoing work My ultimate research goal is to develop trustworthy AI systems that make decisions for and with humans in real-world applications. My current work seeks to advance the applicability of reinforcement learning in real-world settings by considering a problem's structure and leveraging tools from Bayesian inference, causality, and deep learning.

Allowing me to contribute to this mission is my dexterity to flexibly integrate AI tools with statistical techniques, which I have acquired through my uniquely interdisciplinary experience in statistics and computer science – including my doctoral studies in Statistics at UT Austin, postdoctoral work at Harvard University, robotics competitions, internships in top tech companies, industry experience in data science, and collaborations over a wide a spectrum of application domains. My statistical foundations allow me to consider fundamental aspects of decision-making, such as uncertainty, robustness, efficiency, and causality. My experience in AI focuses particularly on reinforcement and active learning (RL), which are essential for designing adaptive decision-making systems.

In my current position at Harvard University, I focus on decision-making methods motivated by problems in climate and health. In Considine et al. (2023), I am leading an effort using RL to learn heat warning issuance policies that reduce hospitalization during heatwaves in the United States. Its application holds an enormous potential to mitigate the health impacts of climate change by developing adaptive policies determining when and how best to intervene. Using Bayesian statistical methods, our work synthesizes an interactive environment for training decision-making agents from observational data. Evaluation in this environment shows that out-of-the-box RL methods struggle to learn effective heat warning policies. We offer practical modifications to the training framework to improve their performance, showcasing the potential of AI to improve the current rule-based system of the National Weather Service.

Complementary to sequential decision-making, I am advancing deep-learning methods for spatial and higher-order network data. In Tec et al. (2023a), I introduced self-supervised methods for estimating causal effects under spatial lagged dependencies. In Tec et al. (2023b), I proposed the first benchmark toolkit with realistic semi-synthetic datasets affected by spatial confounding (Tec et al., 2023b), contributing to a much-needed growing ecosystem of machine-learning tools for causal inference. In more recent work, I am investigating novel equivariant topological deep-learning architectures to address limitations in graph neural networks to handle higher-order, multi-resolution data. Ultimately, this work will enable learning representations of complex spatiotemporal data to improve AI decision-making systems in public health applications.

Agenda In my research at Harvard, I have identified several challenges in RL methods for the heat warning setting. These challenges include constrained budgets over long-term horizons, reward sparsity, small sample sizes, low signal-to-noise rewards, and causal confounding. *These issues are not unique to the heat warning setting*. Addressing them can benefit business and technological AI applications within leading companies. For instance, constrained budgets are common in resource allocation problems, such as cloud computing, data centers, and notification and recommendation systems. Reward sparsity and low signal-to-noise ratios are common issues in many RL applications, including robotics, gaming, and learning from human feedback. Small sample sizes and causal confounding are prevalent in offline learning when using non-experimental data, particularly for consumer behavior and health applications.

I propose to tackle these issues from a generalizable perspective. Initially, I will concentrate on the widely applicable setting of episodic constrained RL problem with exogenous information (entailed by the heat alerts problem) (Efroni et al., 2021, 2022; Sinclair et al., 2022).

One issue I am interested in is off-policy evaluation (OPE) and offline RL, where an agent learns from observational data. OPE is a thriving research. However, current offline RL methods do not consider the availability of exogenous information, such as the climate. I am currently exploring augmenting the agent's decision space with belief states of incomplete exogenous information. A related approach was recently explored in the RL literature by Walsman et al. (2022) using imitation learning from "impossibly good experts". Another idea we adopt in our heat alert work is using a Bayesian model of the rewards to create a carefully crafted simulator that reuses real climate data and the known dynamics of the endogenous decision variables. The properties of this framework need further investigation. In particular, concerning related work at leading companies, one can explore the hindsight approach of Sinclair et al. (2022) to learn approximately exact solvers from planning methods from this simulator. We can also derive sample-efficiency results analogous to those of Efroni et al. (2022), adapting their setting to consider the error in the rewards model and a reward function depending on the exogenous information. Lastly, while we have used a Bayesian model of the rewards from aggregated data, future work could explore simulators of human response to climate and warning systems inspired by the generative models of human behavior (Pearce et al., 2023).

The heat alert problem can be a valuable testbed for new offline methods. For instance, current OPE methods were unsuccessful in the heat alerts problem due to a lack of stochasticity and overlap in the observed policies. My connection with Dominici's lab at Harvard University can provide access to a rich, fine-grained health and climate data platform to continue investigating this setup and other social impact applications in public health.

Another issue that I aim to tackle is exploration in budgeted long-horizon settings. Standard exploration strategies struggle to reach states where the budget is saved for late-stage interventions. Exploration bonuses and intrinsic motivation could provide solutions. I found these methods useful in my previous work on goal-conditioned RL (Durugkar et al., 2021).

There are many other exciting directions to explore. One that is important is information sharing and policy transfer between multiple independent agents. Policies learned on shared data can increase learning efficiency but also introduce bias. In particular, causal confounding bias could be introduced offline due to spurious correlations between the observed policies and individual-specific rewards. To my knowledge, this problem has not been addressed in the RL literature. In the heat alerts, this problem would manifest as *spatial confounding*, which I have investigated in recent work (Tec et al., 2023a,b).

Collaborations and previous work In my previous experience, I have found that the most exciting and impactful research often comes from collaborations. I have substantial experience working with experts from various domains, including computer science, robotics, healthcare, transportation engineering, and epidemiology. Three collaborations were particularly influential in informing my current goals and interests.

The first influential experience occurred in the Learning Agents Research Group (LARG) at UT Austin, directed by Dr. Peter Stone. I joined this group to pursue more RL experience after finishing an internship at Facebook AI Research (FAIR) in the Summer of 2020, conducting RL experiments using world probabilistic models from image-based states. In one project at LARG, I provided critical insights on the statistical behavior of the Wasserstein distance between distributions of the states visited by an agent in goal-conditioned RL, leading to the publication of a new method at NeurIPS 2021 (Durugkar et al., 2021). Other collaborations with Dr. Stone leading to publications include (1) a statistical analysis of human-robot interaction through gaze (Holman et al., 2021) and (2) developing a lightweight real-time object detection system for soccer robots (Narayanaswami et al., 2022), allowing us to obtain 4th and 5th place in the 2021 and 2022 Robocups. My passion for AI and robotics was consolidated during my stay at LARG. Here, I also gained confidence in my ability to quickly adapt to new domains and provide critical insights based on my statistical background.

The second experience came from collaborating with Memorial Sloan Kettering Center researchers. In Tosh et al. (2023), we introduced a novel active learning to accelerate drug discovery for combination therapies in cancer treatments, where a sequential decision step chooses the next best batch of drug pairs for testing based on the information gain predicted by a drug synergy Bayesian model that I designed.

My implementation code was used to run weekly drug screening real experiments. This work fueled my interest in using Bayesian tools for sequential decision-making, which I have continued to pursue in other work (Tec et al., 2022; Müller et al., 2022; Considine et al., 2023). Batch active learning is an area of active research.

The third experience was contributing to the UT Austin's Covid-19 Modeling Consortium during my early doctoral studies. Here, I made significant contributions to models used by city officials to guide policy decisions during the pandemic. My main contribution was to develop a hybrid Bayesian and epidemiological model and scalable code using cell phone mobility data as a proxy for social distancing. This work was crucial in deciding stay-at-home orders policy in Austin, TX, and subsequently published at the *Proceedings of the National Academy of Sciences* (Fox et al., 2021; Cramer et al., 2021). This experience showed me the importance of efficient, robust, and uncertainty-aware methods in critical real-world settings like the pandemic. It also taught me that AI methods were, for the most part, not yet ready to be deployed in such settings, motivating me to pursue research to close this gap.

Conclusion My multiple collaborations and a strong background in both statistics and computer science have prepared me to pursue my proposed research agenda. By pursuing this agenda, we will work towards enabling the real-world deployment of heat warning systems and other mitigation policies to minimize the health impacts of climate change in the United States, with the potential of saving thousands of lives every year. At the same time, these developments go beyond public health and carry benefits for the science of reinforcement learning and applications throughout systems essential to technology products.

References

* indicates equal first-author contribution; ** indicates senior authorship.

E. Considine, R. Nethery, G. Wellenius, F. Dominici, and M. Tec**. Optimizing heat alert issuance for public health in the united states with reinforcement learning, 2023. In preparation for submission to JASA. An advanced draft is attached for your review.

E. Cramer, ..., M. Tec, ..., and N. Reich. Evaluation of individual and ensemble probabilistic forecasts of covid-19 mortality in the us. *Proceedings of the National Academy of Sciences (PNAS)*, 2021.

I. Durugkar, M. Tec, S. Niekum, and P. Stone. Adversarial intrinsic motivation for reinforcement learning. *Neural Information Processing Systems (NeurIPS)*, 2021.

Y. Efroni, D. Misra, A. Krishnamurthy, A. Agarwal, and J. Langford. Provable rl with exogenous distractors via multistep inverse dynamics. In *ICLR 2022*, October 2021.

Y. Efroni, D. Foster, D. Misra, A. Krishnamurthy, and J. Langford. Sample-efficient reinforcement learning in the presence of exogenous information. In *2022 Conference on Learning Theory*, May 2022.

S. Fox, M. Lachmann, M. Tec, ..., and L. Meyers. Real-time pandemic surveillance using hospital admissions and mobility data. *Proceedings of the National Academy of Sciences (PNAS)*, 2021.

B. Holman, A. Anwar, S. Akash, M. Tec, J. Hart, and P. Stone. Watch where you're going! gaze and head orientation as predictors for social robot navigation. *IEEE International Conference on Robotics and Automation (ICRA)*, 2021.

P. Müller, Y. Duan, and M. Tec. Simulation-based sequential design. *Pharmaceutical Statistics*, 21(4): 729–739, 2022.

S. Narayanaswami, M. Tec, ..., and P. Stone. Towards a real-time, low-resource, end-to-end object detection pipeline for robot soccer. In *Robot World Cup XXV Proceedings*, 2022.

T. Pearce, T. Rashid, A. Kanervisto, D. Bignell, M. Sun, R. Georgescu, S. V. Macua, S. Z. Tan, I. Momennejad, K. Hofmann, et al. Imitating human behaviour with diffusion models. *ICLR*, 2023.

S. R. Sinclair, F. Frujeri, C.-A. Cheng, and A. Swaminathan. Hindsight Learning for MDPs with Exogenous Inputs. *arXiv:2207.06272*, 2022.

M. Tec, Y. Duan, and P. Müller. Bayesian sequential design and reinforcement learning: A comparative tutorial. *The American Statistician*, 2022.

M. Tec, J. Scott, and C. Zigler. Weather2vec: Representation learning for causal inference with nonlocal confounding in air pollution and climate studies. *Association for the Advancement of Artificial Intelligence (AAAI)*, 2023a.

M. Tec, A. Trisovic, M. Audirac, S. Woodward, J. K. Hu, N. Khoshnevis, and F. Dominici. Space: The spatial confounding environment. *arXiv:2312.00710. Under review at ICLR and published as a workshop paper at the Causal Learning and Reasoning (CLeaR) conference*, 2023b.

C. Tosh, M. Tec, J. White, J. F. Quinn, G. I. Sanchez, P. Calder, A. L. Kung, F. S. D. Cruz, and W. Tansey. A bayesian active learning platform for scalable combination drug screens. *Under review at Nature Biotechnology*, 2023.

A. Walsman, M. Zhang, S. Choudhury, D. Fox, and A. Farhadi. Impossibly good experts and how to follow them. In *The Eleventh International Conference on Learning Representations*, 2022.