Model simulations are key to understanding how climate impacts ecosystems over time, provided that we can trust them. As the demand for reliable projections is increasing, the systematic evaluation of model skill should be one of the main concerns of modellers. Such evaluation remains critical to build confidence in our models, and plays a crucial role in providing the most credible information on the impacts of climate change so that stakeholders can make informed decisions [@Dawson2011; @Mouquet2015].

Direct testing of the accuracy model future projections is infeasible as the unforeseen will always remain so. Hence, the most straightforward approach to evaluate models is to compare their output with what we know from the past. Plausibly reproducing the past (*hindcast performance/skill*) can be seen as a requisite condition to be considered a viable model for future projections (*forecast reliability*). While exact matches to 21st-century expected climatic conditions do not exist in historical records [@Burke2018], hindcasting exercises can still increase our confidence that the models capture, implicitly or explicitly, the essential processes for the simulation of future species range shifts.

In this regard, recent past observations (typically spanning a few decades) have been used for testing SDM predictions over time [e.g. @Araujo2005; @Kharouba2009; @Smith2013; @Illan2014]. However, as they were made in a limited climate range similar to calibration conditions, they do not enable fully independent model evaluation**.** By looking much further back in the past, paleoenvironments offer a unique framework to test species distribution model transferability in more challenging conditions, in the same way as climate models are evaluated using palaeodata [@Braconnot2012]. Simulations in distant past (spanning several millennia) allow for model evaluation under conditions significantly different from present-day [@Maguire2015], where climate variations were larger than those encountered during the last century. Taking advantage of the available paleo-archives is a great opportunity to understand long-term climate-biodiversity dynamics [@Fordham2020]. Numerous studies have tested the transferability of species distribution models using paleoclimate reconstructions and fossil records [e.g. @Veloz2012; @Pearman2008; @Williams2013, @Roberts2012]. They consistently reveal a decrease of the ability of SDMs to simulate past species distributions. This remind us the need for caution when interpreting their projections in novel climates that differ significantly from the present [@Maguire2016], especially as no-analogues climate conditions are forecasted to become common [@Williams2007], potentially compromising the accuracy of model predictions [@Fitzpatrick2018]. While these investigations have yielded valuable insights into the reliability of species distribution models, they have primarily focused on correlative models, despite the growing interest for process-based models in predictive ecology [@Connolly2017; @Urban2016; @Pilowksy2022].

This omission represents a notable gap in our understanding of the tenets of species distribution modelling, as only one side of the continuum between statistical and mechanistic approaches has been explored [@Dormann2012], neglecting the investigation of process-based model performance. Rather than inferring correlations between observations and potential environmental drivers, these models are built upon explicit causal relationships determined experimentally representing physiological, ecological and demographic processes. Their projections in response to climate change differ from correlative models, the latter being systematically more pessimistic [**références**]. However, very few studies have gone beyond these qualitative comparisons and really examined how they are performing, with virtual species [@Zurrel2016], exotic species in newly colonized areas [@Higgins2020], or in the recent past [@Fordham2018]. Therefore, despite process-based models have shown their usefulness for paleoecological studies [**références**], the extent to which the potential benefits of process-based models translate into improved predictions in really different climatic conditions remains unknown [@UribeRivera2022, @Briscoe2019]. Bridging this gap through a comprehensive evaluation of the different class of models is crucial [@Evans2016], given that process-based models hold the potential to enhance our ability to predict species responses to climate change and to provide more robust projections in novel conditions [@Evans2012; @Singer2016]. Long-term *retrospective* data offer a great opportunity to deepen our understanding of what conveys robustness to the different models and determine which model forecasts are appropriate to guide climate adaptation strategies [@Fordham2016].

Here, we propose a state-of-the-art comparison of the skills of correlative models and process-based models to simulate paleodistributions of emblematic tree species of Europe at a high-temporal resolution. In order to fully explore the different classes of models, we use different versions of the models that differ by their level of complexity and the methods of estimation of their parameters. In particular, as hybrid models [sensu @Dorman2012] have been raised as a potential avenue by borrowing strength from both statistical and process-based approaches [@Evans2016], we use inverse modelling to fit the process-based models in the same way as the correlative models [@VanderMeersch2023]. By encompassing the entire spectrum of models, from correlative models to process-based models and their hybrid data-driven counterparts, our comprehensive approach allows us to gain a holistic understanding of the key features necessary for building reliable models.