Week 14

Wentao Gao

Research Proposal

Traditional Statistical Methods

Physical Models

AI methods(ML and DL)

Dynamic Models

Emphasizing a bottom-up approach, physical models shifted the focus from statistical correlations to simulating the very hydrological processes that underpin droughts.

VIC (Variable Infiltration Capacity) The VIC model primarily simulates the water and energy balance at the surface and near-surface levels. It's especially suited for studying terrestrial hydrological processes.

Noah Land Surface Model: Noah incorporates detailed treatments of soil, vegetation, and snow. It's frequently coupled with climate models and weather forecasting models, providing them with surface boundary conditions

CLM (Community Land Model): CLM is crafted to simulate a spectrum of land processes, including hydrology, ecology, and biogeochemical cycles.

Reference:

1. Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research, 99(D7), 14415-14428.
2. Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review, 129(4), 569-585.
3. Oleson, K. W., et al. (2010). Technical description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-478+ STR.

Traditional statistical methods

Drought predictions heavily relied on time series and regression analyses. These methods utilize historical data, drawing on past patterns and relationships to predict future drought events

Time Series Models:

ARIMA: one of the pioneering efforts, excelled in handling non-stationary data, becoming a mainstay in early drought prediction endeavors (An operational method to forecast reservoir inflow using arima models.)

ETS models(Error, Trends, Seasonality): ETS models further refined this approach, emphasizing the modeling of time depen- dencies. However, their capabilities were often stymied by erratic climatic fluctuations (Forecasting with exponential smoothing: the state space approach.)

The Wavelet Transform: The Wavelet Transform brought a multi-scale perspective, breaking down time series data to unearth patterns at different resolutions. Such an approach proved invaluable for capturing the multifarious nature of hydrological processes (wavelet analysis of river discharge)

Machine learning based method

Feedforward Neural Networks emerged as a simplistic yet adaptable tool, casting a fresh perspective on hydrological patterns, especially in encapsulating concealed, non-linear intricacies prevalent in expansive datasets Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications.

Venturing further into the temporal domain, **Recurrent Neural Networks (RNN)** seamlessly bridged the chasm between static and evolving prediction paradigms, credited to their adeptness at integrating past input sequences(Artificial neural networks as rainfall-runoff models.)

Convolutional Neural Networks (CNNs) was not limited to just image data. They were adapted for sequential precipitation prediction, capitalizing on their ability to detect localized temporal patterns and achieve hierarchical feature extraction in time series data.(An empirical evaluation of generic convolutional and recurrent networks for sequence modeling)

Support Vector Machines (SVMs), with their inherent mathematical elegance, epitomized resilience in data-scarce environments, showcasing an enviable defense against overfitting (Multi-time scale stream flow predictions: The support vector machines approach.)

Deep learning based method

Long Short-Term Memory (LSTM) units stood out as an advanced iteration of RNNs, meticulously catering to long-term sequential data dependencies, while concurrently circumventing the vexing vanishing gradient conundrums that beleaguered their RNN counterparts (Long short-term memory. Neural computation)

Furthermore, the **Transformer** architectures, originally sculpted with linguistic tasks in mind, astonishingly manifested their versatility, adeptly discerning the significance of varied temporal sequences in hydrological data streams Attention is all you need.

Additionally, the Transformer architectures and their derivatives, such as the **Informer, ETSformer, FEDformer** have showcased considerable promise in time series forecasting, optimized to handle hydrological data, thus delivering commendable performances.

Research questions and expected outcomes

- Research Question**:**
- How can causal discovery techniques be leveraged to enhance feature selection, thereby augmenting the robustness and transferability of drought prediction models?

- Expected Outcome**:**
- Inspired by foundational works on causal discovery, such as the PC algorithm, **our methodology aims to extract causally relevant features from multidimensional climatic datasets related to drought conditions.** By emphasizing causally significant features, we anticipate enhanced feature selection leading to models that exhibit better robustness and transferability across various geographical and temporal domains. The outcome will not only improve model performance but also shed light on critical climatic drivers of drought, enabling better-informed intervention strategies.

- Research Question 2: How can the application of Independent Component Analysis (ICA) assist in identifying the intrinsic variables that exert genuine influence on drought phenomena, leading to enhanced predictive accuracy in drought forecasting models? Feature representation

- Expected Outcome:
- Building upon the foundational principles of ICA. Our anticipated approach seeks to decompose multi-dimensional climate datasets into statistically independent components. We expect to unveil latent structures and identify the intrinsic variables pivotal in influencing drought conditions. By isolating these variables, our goal is to achieve a more refined and accurate forecasting model. This streamlined model should not only exhibit superior predictive performance but also provide clearer insights into the interplay of diverse climatic factors that precipitate drought events. Furthermore, understanding these intrinsic variables can pave the way for more targeted and efficient interventions to mitigate drought impacts.

- Research Question 3: How can causal transfer learning be integrated to optimize domain adaptation and bolster the robustness of a model, ensuring that pre-trained architectures maintain high predictive accuracy on unseen datasets?

- The conceptual foundation for this query is rooted in the recent advances in domain adaptation and causal inference. **Our proposed approach will leverage the causal transfer learning framework to enhance domain adaptability.** The ambition is to devise an algorithm grounded in causal principles that can effectively adjust and adapt pre-trained models to new domains without direct access to the domain-specific training data, ultimately ensuring consistent performance across diverse datasets. This endeavor aims to not only bridge the performance gap in domain adaptation but also offer insights into the underlying causal dynamics that drive model generalization.