

ASSIMILAÇÃO DE DADOS POR REDES NEURAIIS NO MODELO WRF-NCAR

Gerônimo Gallarreta Zubiaurre Lemos¹ (UFPEL, Bolsista PIBIC INPE/CNPq)
Haroldo Fraga de Campos Velho² (INPE, Orientador)
Juliana Aparecida Anochi³ (INPE, Coorientadora)

RESUMO

O presente trabalho propõe o uso de método de aprendizado de máquina para o problema de assimilação de dados (DA: *Data Assimilation*) aplicado ao modelo de meso-escala WRF. Para o desenvolvimento do projeto, foi estruturado um banco de dados contendo análises do sistema 3D-Var, previsões do modelo WRF (*Weather Research and Forecasting System*) e um conjunto de observações. Para gerar essas análises, foram obtidos dados de previsão do modelo GFS-NOAA (*Global Forecast System*) e dados de observações para execução do módulo de assimilação do modelo WRF (WRF-DA). Os dados para o experimento numérico foram coletados dos meses de Junho, Julho e Agosto dos anos de 2015, 2016, 2017 e 2018. O domínio estudado abrange a região sul do Brasil, Rio Grande do Sul, Santa Catarina e Paraná, além do Uruguai e parte da Argentina e sul do Paraguai. O método de aprendizado de máquina é chamado de *árvores de decisão* e usou-se a codificação da biblioteca XGBoost, que implementa uma versão otimizada do algoritmo *Gradient Boosting*, um algoritmo supervisionado. Para configurar a biblioteca XGBoost para a aplicação, o conjunto de dados foi subdividido em dois conjuntos distintos: conjunto de treinamento e conjunto de teste. O conjunto de treinamento é constituído pelos dados dos anos de 2015, 2016 e 2017, enquanto o conjunto de teste possui os dados do ano de 2018. A técnica de *K-fold Cross-Validation* foi utilizada para o treinamento do modelo, com $k = 3$, sendo os conjuntos de validação para cada iteração um dos anos do conjunto de treinamento. Ou seja, o modelo XGBoost-DA foi configurado por um procedimento em etapas: inicialmente, treinado com os anos de 2015 e 2016 e avaliado no ano de 2017; após, treinado com os anos de 2015-2017 e avaliado no ano de 2017 e, por fim, o último modelo foi treinado nos anos de 2016 e 2017 e avaliado no ano de 2015. Ao fim desse processo, foi escolhido o modelo com menor erro sobre o conjunto de validação. Para avaliação final do modelo, foi executado um exemplo de previsão de 24 horas no ano de 2018, com análises geradas pelo módulo de 3D-Var (nativo no WRF) e análise geradas pelo XGBoost-DA. Previsões geradas com as duas análises se mostraram muito similares, evidenciando a aplicação do XGBoost como uma ferramenta promissora para assimilação de dados.

Palavras-chave: Inteligência Artificial. Aprendizado de Máquina. Assimilação de dados. Previsão Numérica do Tempo.

¹Aluno do curso de bacharelado em Ciência da Computação - E-mail: ggzlemos@inf.ufpel.edu.br

²Pesquisador do INPE - E-mail: haroldo.camposvelho@inpe.br

³Pesquisadora do INPE - E-mail: juliana.anochi@inpe.br