

Quantifying the Contributions of Teleconnections on Indian Summer Monsoon using Shapley Values

Deepayan Chakraborty , Adway Mitra , Bhupendra Nath Goswami
and P. V. Rajesh



Shapley Values

—

- Consider a game involving a set of N players, where the function $v(S) : 2^N \mapsto \mathbb{R}$ defines the total payoff generated by any subset (coalition) $S \subset N$ of these players.

- Shapley values $\phi_i(v) = \sum_{S \subset N \setminus i} ((|S|!(n-|S|-1)!)/n!)(v(S \cup i) - v(S))$

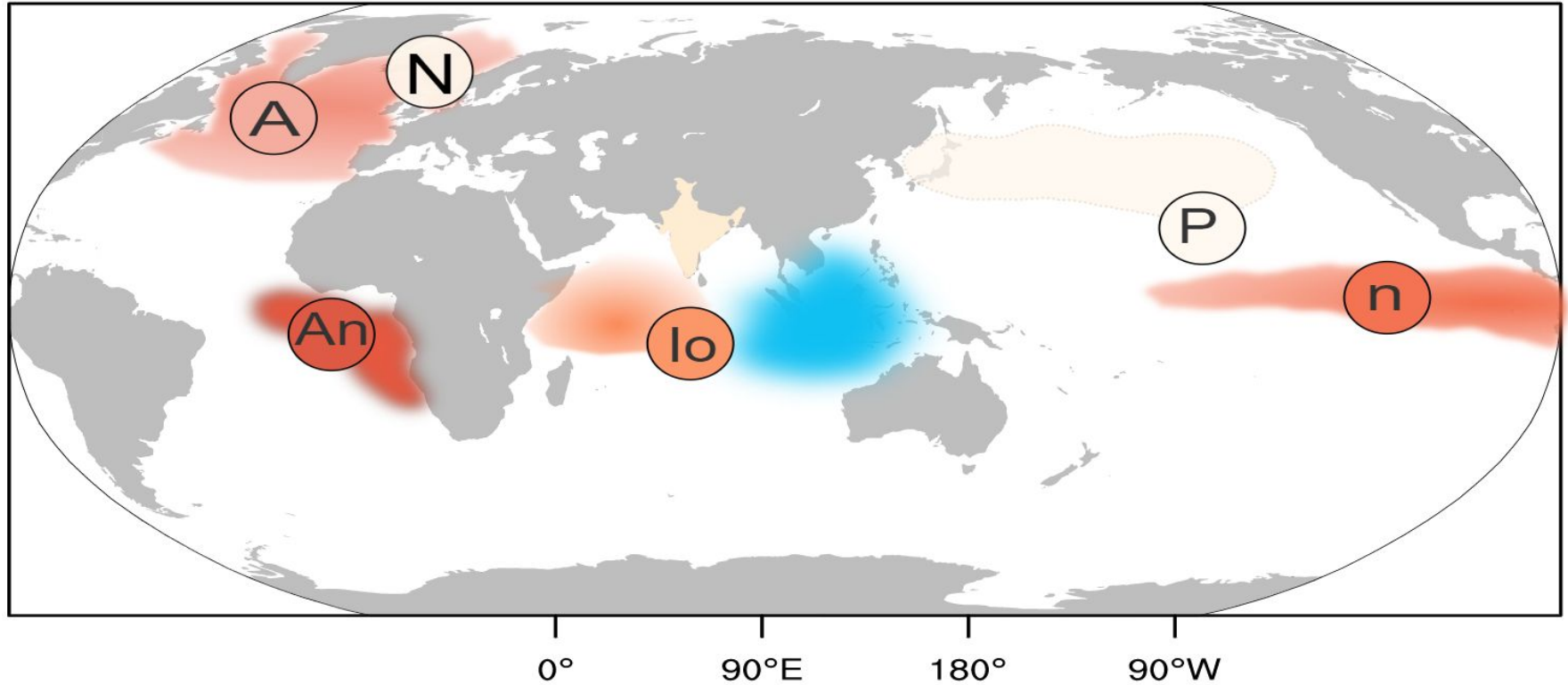
$$\sum_{i \in N} \phi_i(v) = v(N).$$

- The aim of Explainable Machine Learning is to assign importance scores to the features, through a simple linear function $f(x) = g(z) = b_0 + \sum_{i=1}^N b_i z_i$
- b_i can be the Shapley value $\phi_i(f)$ and $b_0 = E_X(f(x))$. (**Lundberg SM, & Lee SI (2017)**)

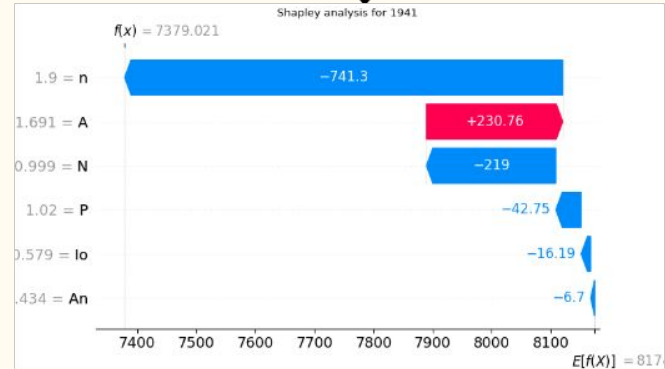
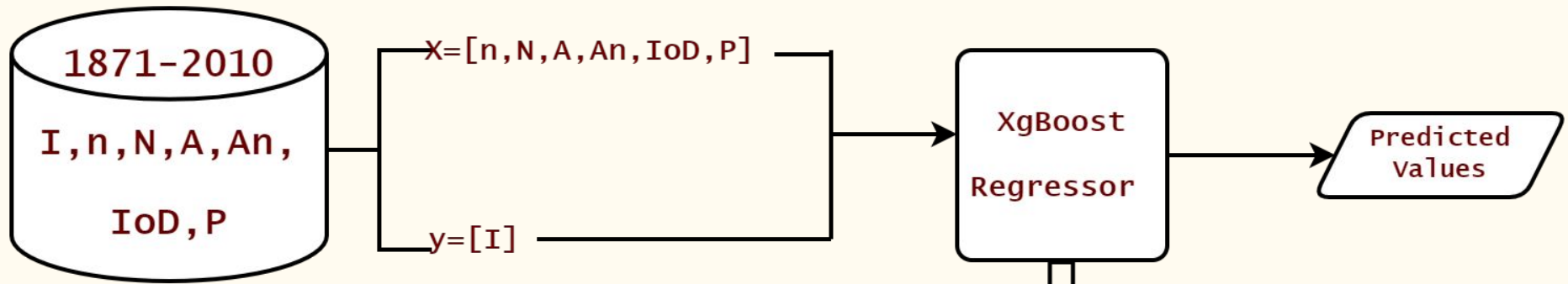
Probable Drivers

—

Lat spacing = 30, Lon spacing = 90



Methodology



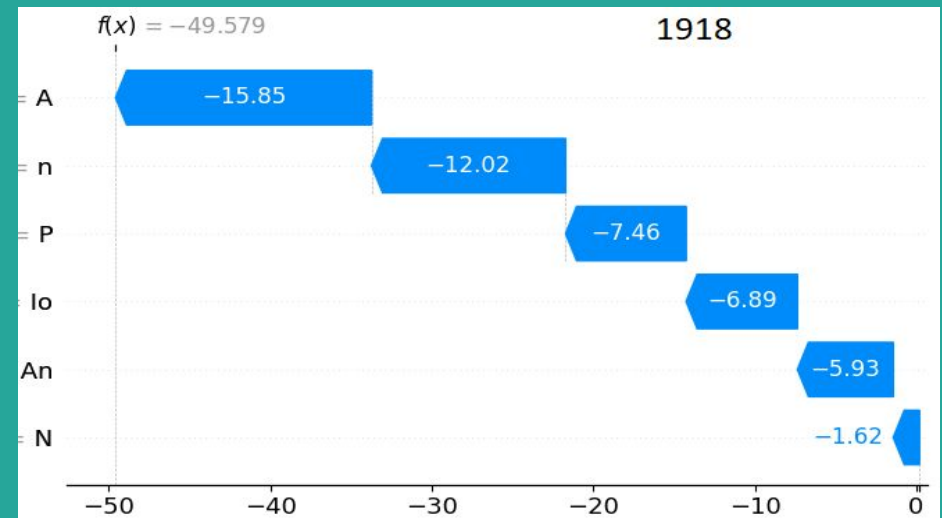
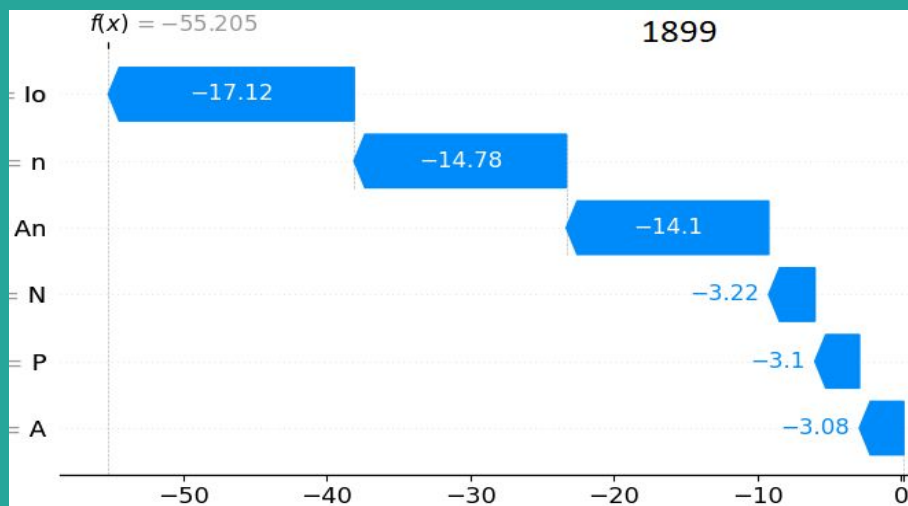
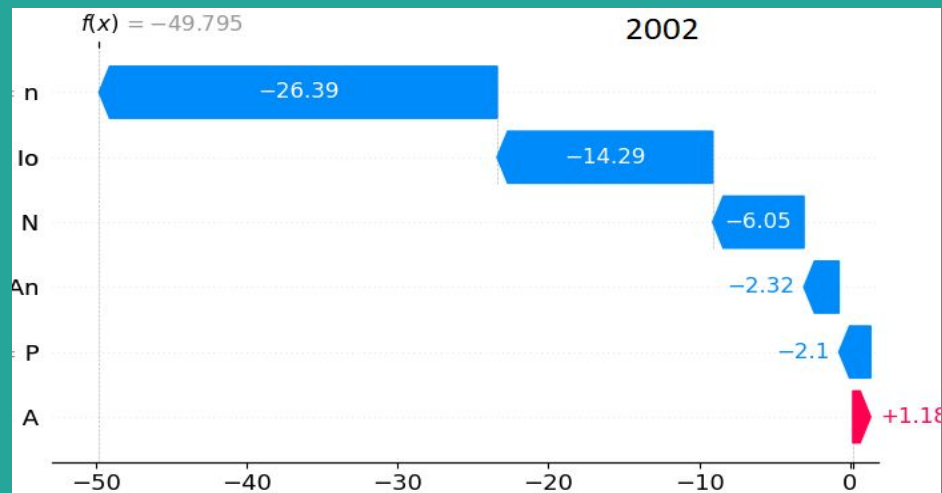
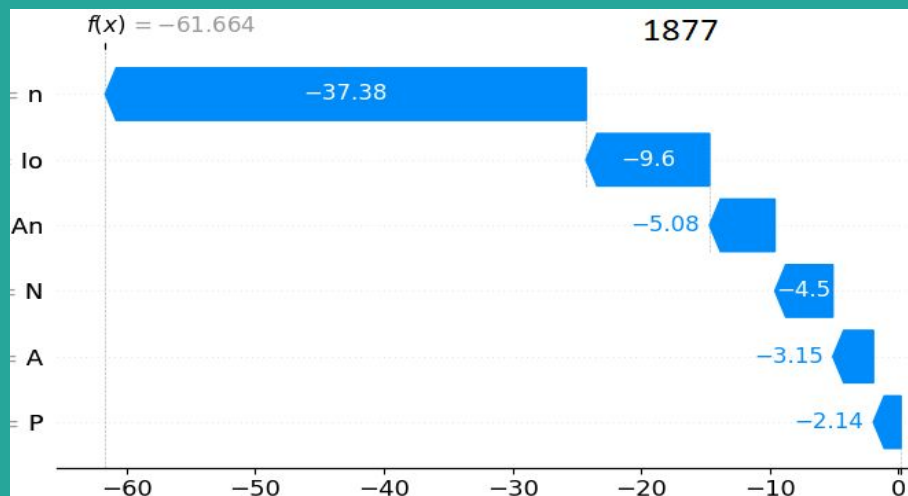
shapley values to analyze contribution of each predictor

Results and Analysis

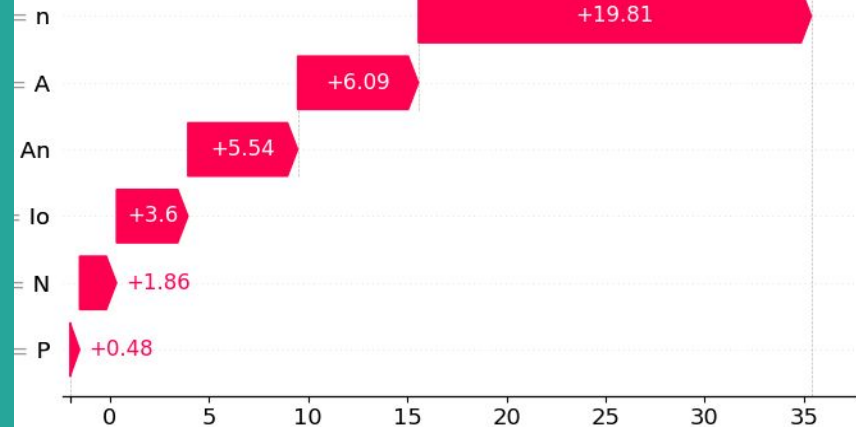
—

Results and Analysis

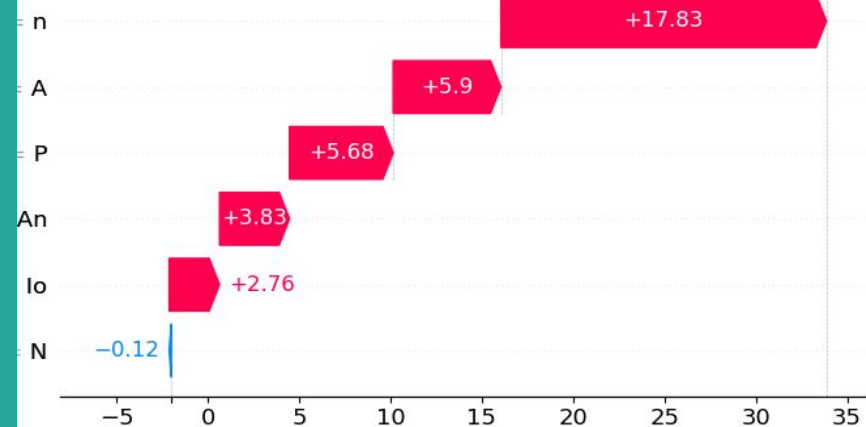
- Waterfall plots of Shapley values for strongly Deficient Rainfall years.
 - Waterfall plots of Shapley values for strongly excess ISMR years.
 - Waterfall plots of Shapley values for normal ISMR years.
-



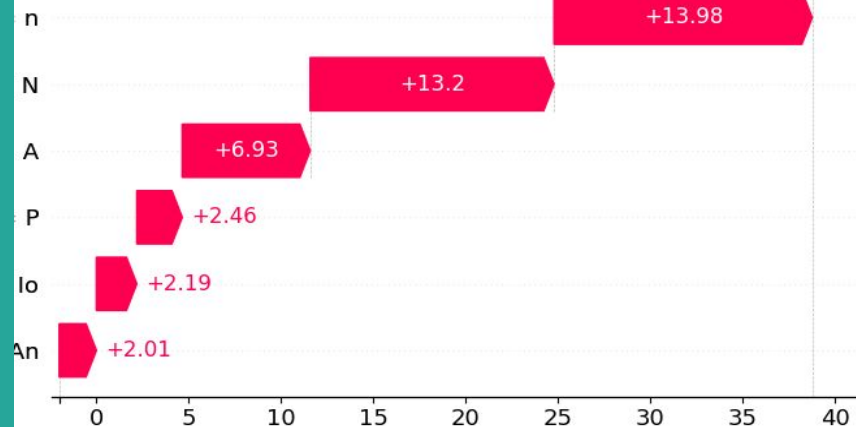
1892

 $f(x) = 35.37$ 

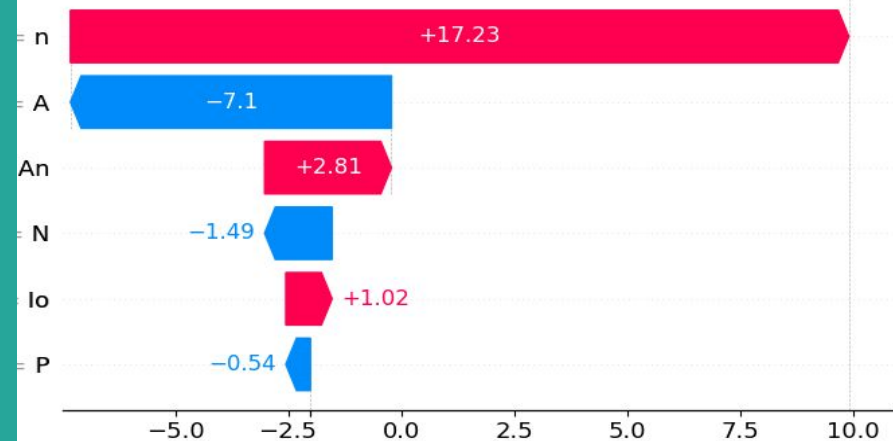
1956

 $f(x) = 33.87$ 

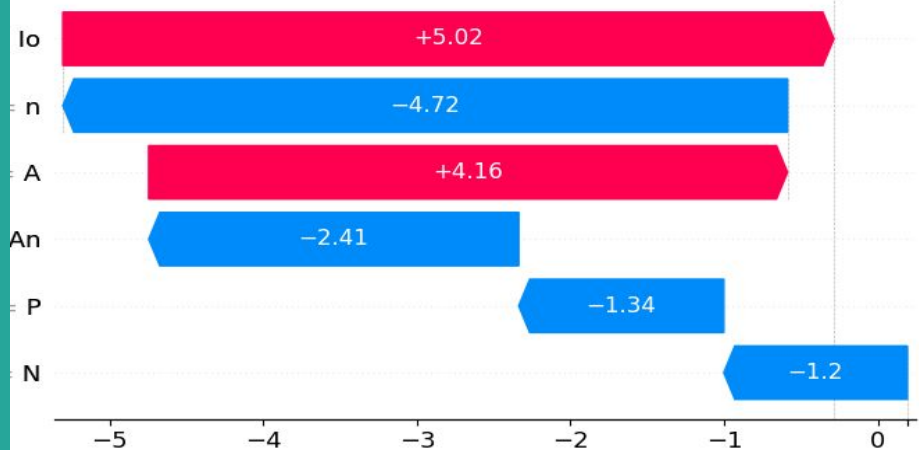
1917

 $f(x) = 38.75$ 

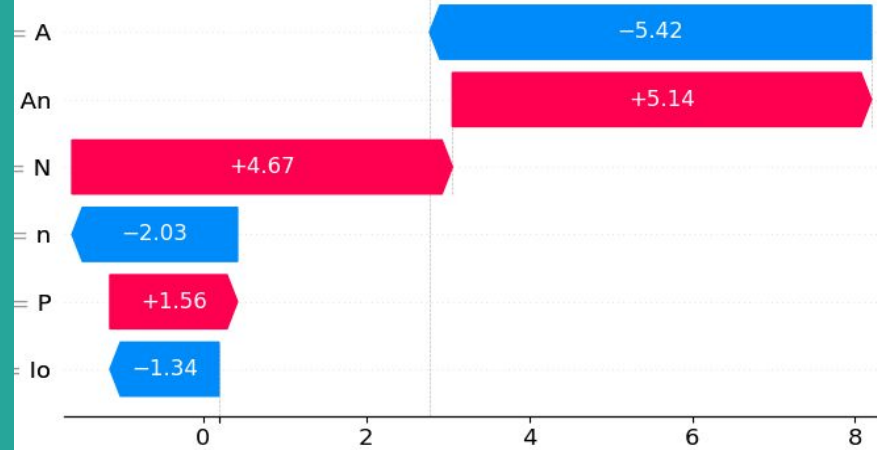
1971

 $f(x) = 9.9$ 

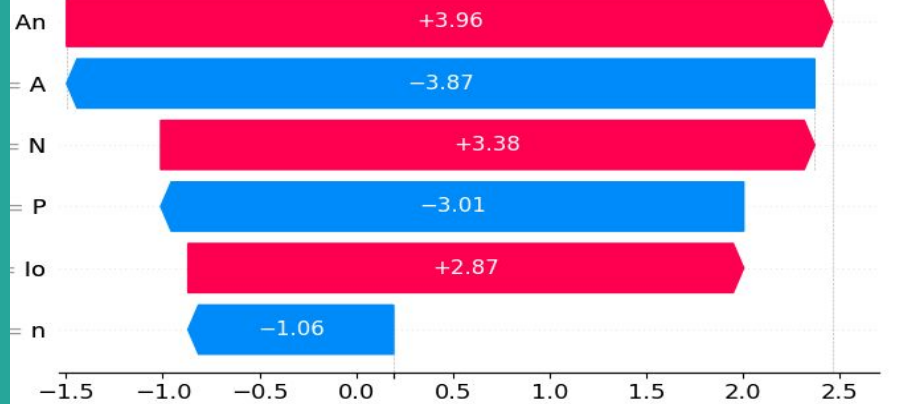
1883

 $f(x) = -0.2$ 

1903

 $f(x) = 2.779$ 

1976

 $f(x) = 2.46$ 

1997

 $f(x) = 2.083$ 

Thank you

—