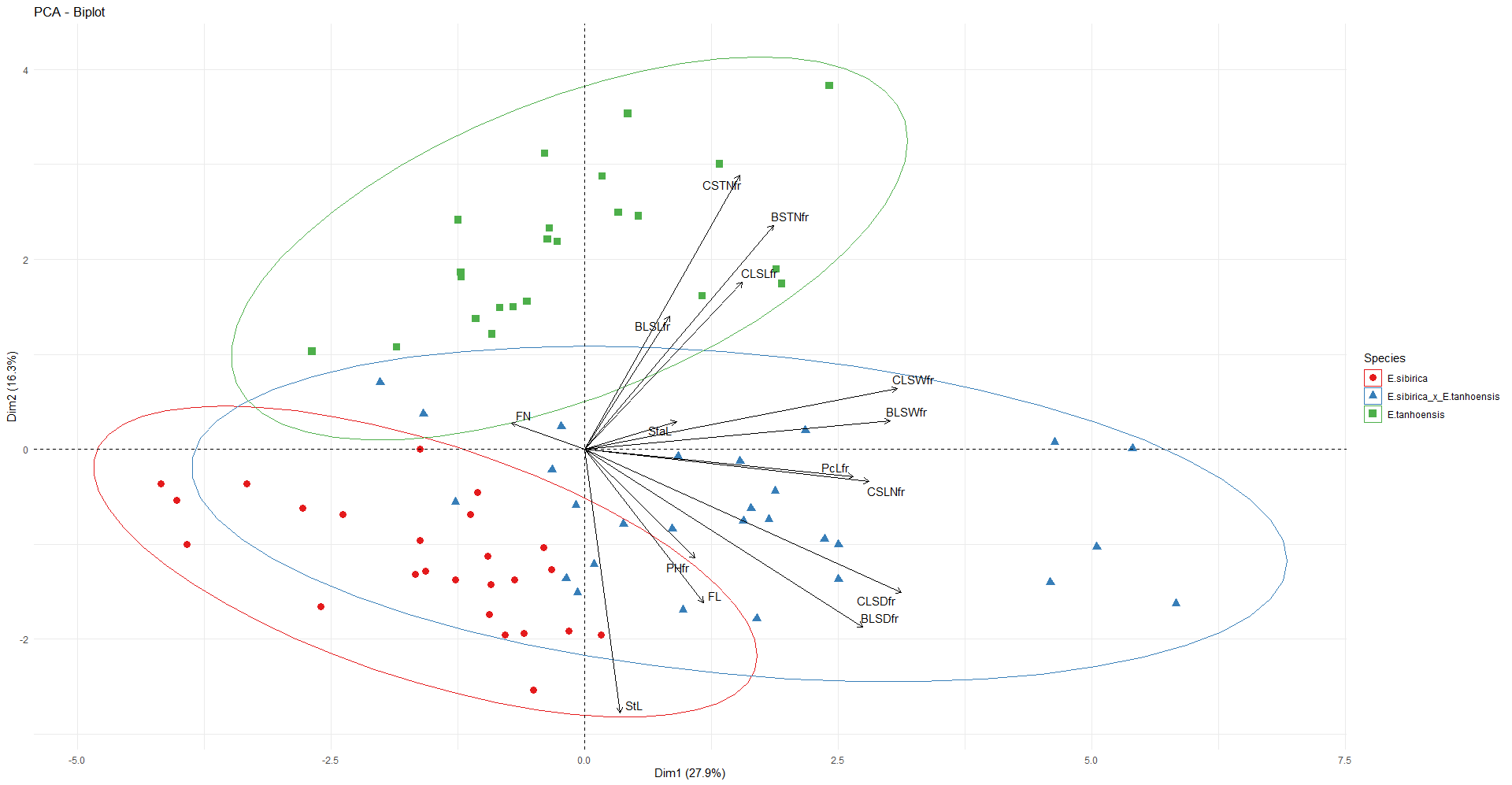
To prepare the data for analysis, we employed multiple linear regression modeling to facilitate the imputation of missing values for subsequent analysis. Following this, we conducted principal component analysis (PCA). A pivotal advantage of PCA lies in its capacity to standardize variables of differing scales, thus facilitating meaningful comparisons across them. Through PCA, we can discern the distinctive features associated with specific species and explore their intercorrelations. To provide a more descriptive visualization, we utilized box plots.

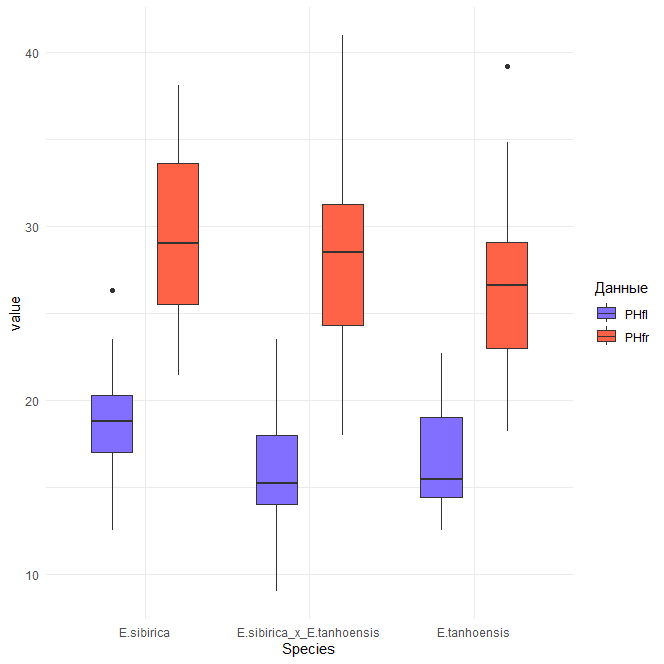
For PCA we divided data to fruting and flowering, but on box plots we can compare it.

Examples



Fruiting of E. sibirica, E. tanhoensis and E. hybrid

So if variable less than 90 degrees its positively related, 90 degrees – unrelated and greater than 90 degrees – negatively related. Also, we can see how hybrid located between parent species.



Comparison of two variables Plant height (flowering) (PHfl), Plant height (fruiting) (PHfr)

As can be seen in this box plot, the hybrid has clear differences in both populations flowering and fruiting.

So we make this analyze on all white-flowered erathis.