Some body text here to show the figure placement at the top of the page. Lorem ipsum dolor sit amet, consectetur adipiscing elit.
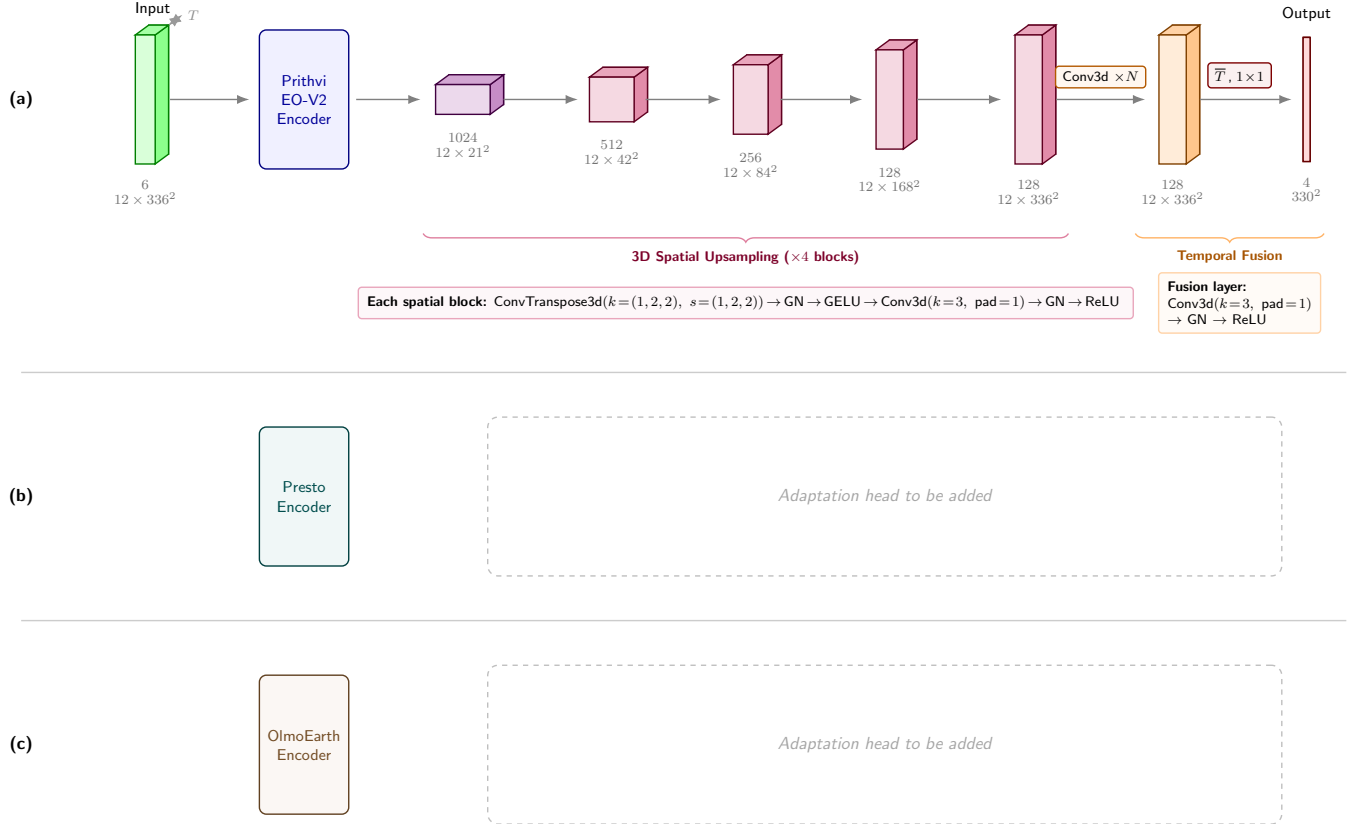
Figure 1: Downstream adaptation architectures for phenology prediction. (a) The Prithvi-EO-V2 encoder extracts spatio-temporal features, which are spatially upsampled through four 3D transposed convolution blocks while preserving the temporal dimension T. A 3D convolutional fusion head mixes temporal information, followed by mean pooling over T and a $1{\times}1$ projection to predict four phenological transition dates per pixel. (b) Presto adaptation (to be added). (c) OlmoEarth adaptation (to be added).