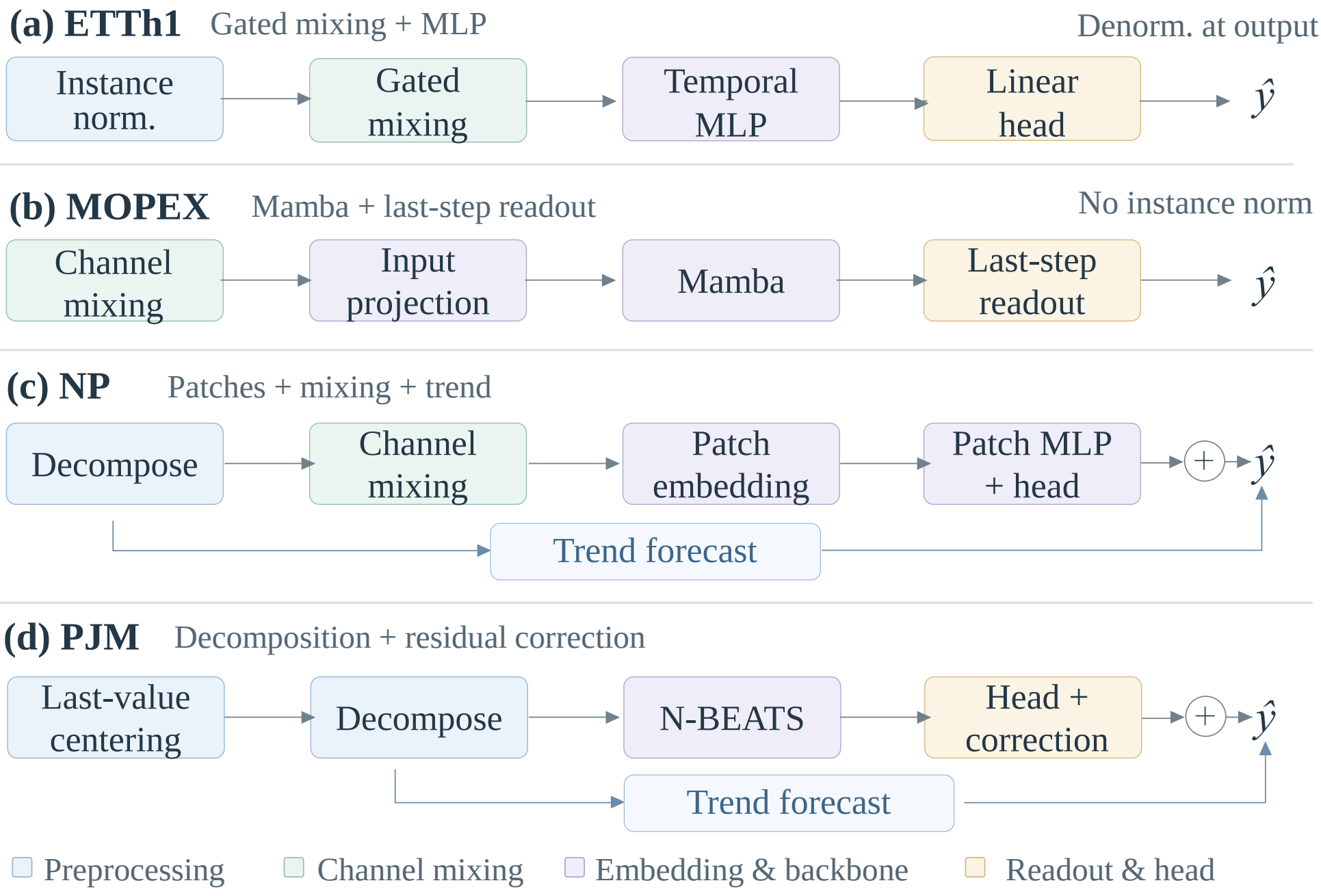

(a) ETTh1
Gated mixing + MLP
Denorm. at output
Instance norm.
Gated mixing
Temporal MLP
Linear head
$\hat{y}$
(b) MOPEX
Mamba + last-step readout
No instance norm
Channel mixing
Input projection
Mamba
Last-step readout
$\hat{y}$
(c) NP
Patches + mixing + trend
Decompose
Channel mixing
Patch embedding
Patch MLP + head
$\hat{y}$
Trend forecast
(d) PJM
Decomposition + residual correction
Last-value centering
Decompose
N-BEATS
Head + correction
$\hat{y}$
Trend forecast
Preprocessing
Channel mixing
Embedding & backbone
Readout & head