Model comparison with data augmentation and window aggregation