Model performance comparison — MAE / RMSE with R²