Predicting late arrivals: Model performance