I. ABSTRACT
- Airline pilots experience acute and chronic sleep deprivation, varying sleep inertia, and circadian desynchrony due to the need to schedule flight operations around the clock (Ceziëls et al., 2007).
- Biomathematical fatigue models (McCauley Model, Harvard’s Matlab Model, and the privately-sold SAFTE-FAST Model) have been useful in predicting pilot performance (Vann Dongen, 2011).
- We compared model outputs to a pilot airline dataset of alertness and performance to find their strengths and weaknesses.
- Our findings will aid operational decision-makers in determining the reliability of each model under real-world scheduling situations.

II. INTRODUCTION
- Sleep loss and circadian desynchrony lead to cognitive impairments, reduced vigilance and inconsistent performance (Ceziëls et al., 2007).
- Several biomathematical models, based principally on patterns observed in circadian rhythms and homeostatic drive, have been developed to predict a pilot’s levels of fatigue or alertness. These models inform the Federal Aviation Administration (FAA) and commercial airlines of pilot capabilities and flight schedules (Mallis et al., 2004).
- These models have not yet been thoroughly tested in operational environments where uncontrolled factors, such as environmental sleep disruptors, caffeine use, and napping, may impact actual pilot alertness and performance.

III. METHODS
- The dataset used was previously collected from 44 pilots during normal schedule operations in a short-haul commercial airline over a period of one month and contained sleep logs, movement and light recordings, psychomotor vigilance task (PVT), and urinary melatonin (a marker of circadian phase).
- Pilots flew to a fixed pattern design (FPD) roster schedule, so there were several days of early, mid and late start days (see right).
- We preprocessed the dataset several times to serve as inputs for the three biomathematical models. In the McCauley Model, we did this with and without naps. PVT is the gold-standard measure of sleep deprivation-related performance impairment. This task measures the speed with which subjects respond to a visual stimulus and the number of times the subjects fail to respond to the appearance of the stimulus within 500ms. The PVT consists of responding to a small stimulus on the computer screen by pressing a response button as soon as the stimulus appears, which stops the stimulus counter and displays the reaction time (RT) in milliseconds.
- Validated PVT response time is displayed accurately through devices by scoring a video experiment of PVT tests.

IV. ANALYSIS/DISCUSSION
- Statistically compared the values of performance that were either outputs from the three models or recorded from the PVT itself during the pilot data collection process for early, midday, and late shifts.
- Calculated sensitivity and specificity of each model prediction.
- Observed accurate shift when including nap input for McCauley Model.