Supplementary Material

A. Supplemental Data and Figures

1) Subject Demographics: Summary subject demographics are provided in Table -A.1. Data reported in this table relevant to subject age and the amount of time for which the subject had been blind reflects information from the earliest recording obtained from each subject. All subject demographics were provided by, or estimated from, previous studies [3], [11], [12].

TABLE -A.1: Subject Demographics

Feature	Mean ± Std Dev	Range	Frequency
Age (years)	58.85 ± 8.24	$[45, 72]$	–
Time Blind (years)	37.54 ± 10.71	$[20, 53]$	–
Gender: Male	–	–	11
Gender: Female	–	–	2
Implant Eye: Right	–	–	12
Implant Eye: Left	–	–	1

2) Subject Threshold Distributions: High variance in perceptual thresholds across subjects and among the electrodes of a single implant make threshold prediction a challenging task. Kernel density estimates of perceptual thresholds for each of the twelve subjects studied in this work are provided in Fig. -A.1. Perceptual threshold variance was lowest for subjects 12-005, 51-009, and 52-001. Qualitatively, higher R^2_{adj} values were typically observed in leave-one-subject-out (LOSO) regression evaluation for these subjects.

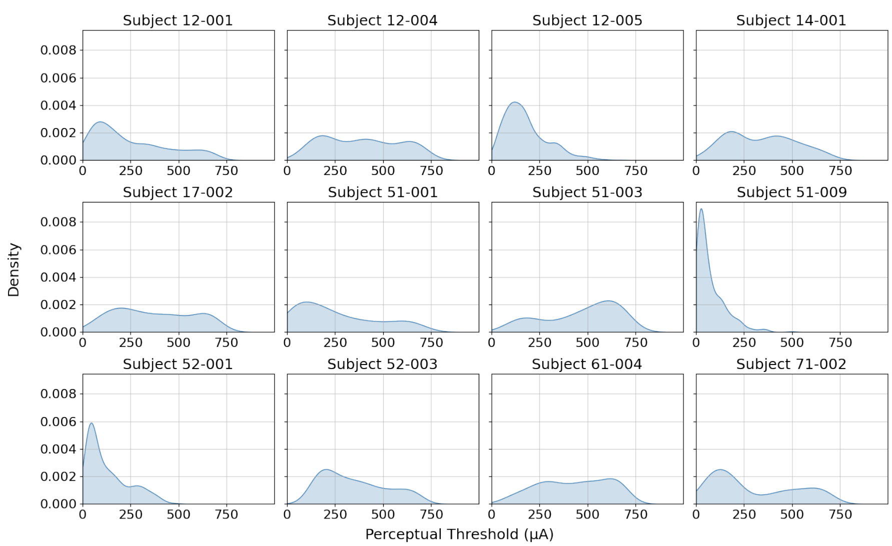

Fig. -A.1: Kernel density estimates of perceptual thresholds for each subject.

B. Per-subject Regression Model Fit

A closer inspection of regression model outputs at the subject level reveals additional insights into Elastic Net (EN) and XGBRegressor (XGB-R) model behavior when fitted to each subset of feature data. When fitted to 'Routine' features, the linear EN model predicted relatively consistent perceptual thresholds for nine of the twelve subjects (Fig. -B.2, top), likely suggesting poor linear relationships between 'Routine' predictors and perceptual thresholds. Similar behavior can be observed from the EN models fitted to 'System Fitting' data (Fig. -B.2, middle) and is additionally reflected in the low R^2_{adj} values for these models. Introducing 'Follow-Up' features significantly improved explained perceptual threshold variance for all subjects (Fig. -B.2, bottom), and enabled R^2_{adj} values greater than 0.7 for five of the twelve subjects. The higher variance XGB-R model also failed to capture generalizable relationships between 'Routine' predictors and perceptual thresholds (Fig. -B.3, top), despite our finding that some of these predictors (including, but not limited to, time since blindness onset, electrode impedance, and electrode-fovea distance) are important to threshold prediction. When 'System Fitting' features were included, improvements in explained variance can be observed for a few of the subjects but there were still a number of subjects whose perceptual thresholds were not well modeled (Fig. -B.3, middle). R^2_{adj} values for all subjects, with the exception of 51-001, improved when the XGB-R model was fitted with 'Follow-Up' features (Fig. -B.3, bottom). Similar to the results observed with the EN model, highest R^2_{adj} values were observed for subjects 12-001, 12-005, 52-001, and 71-002.

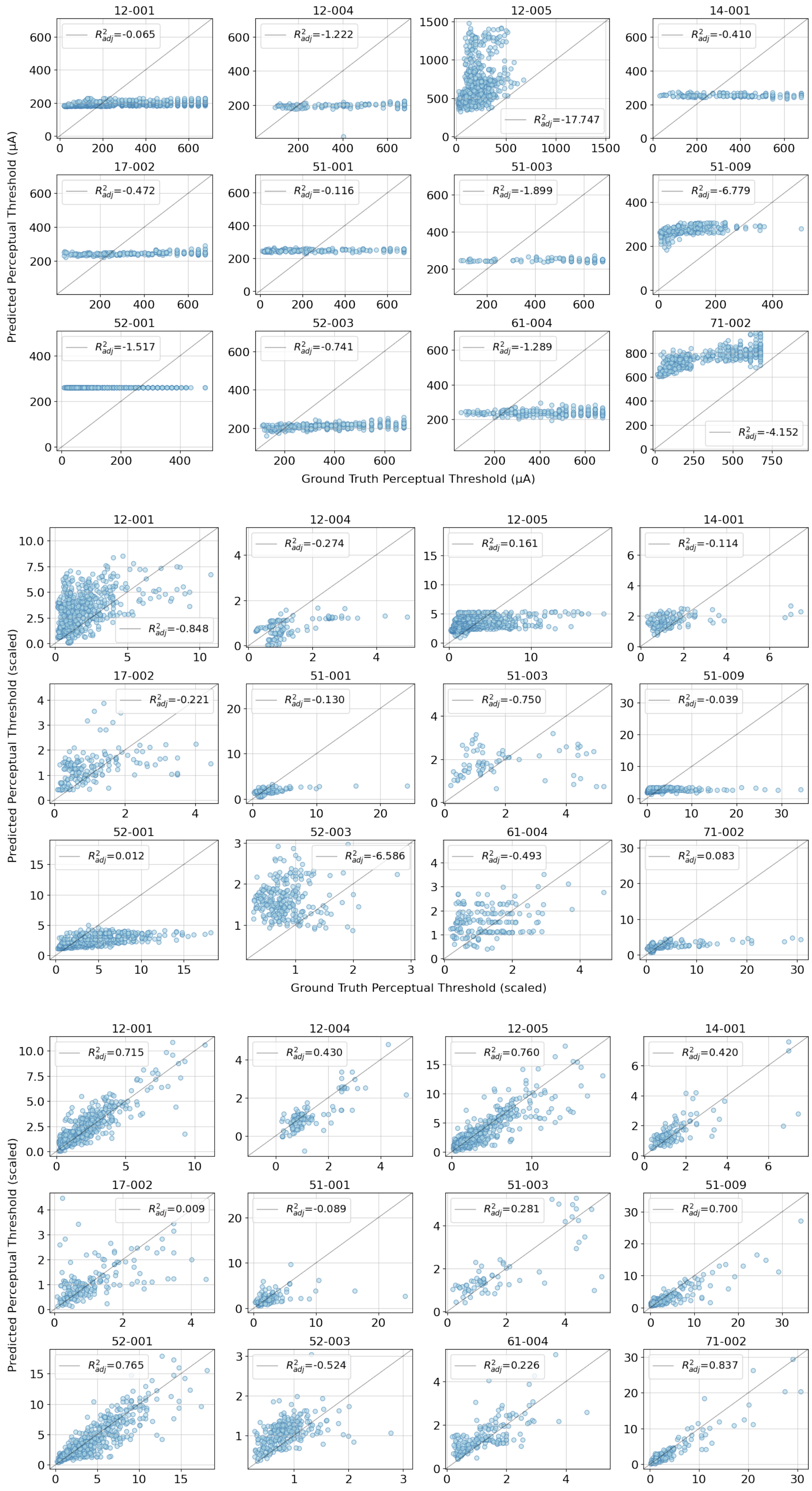

Fig. -B.2: Ground truth and EN predicted perceptual thresholds based on 'Routine' measures (top), 'System Fitting' measures (middle), and 'Follow-Up' measures (bottom).

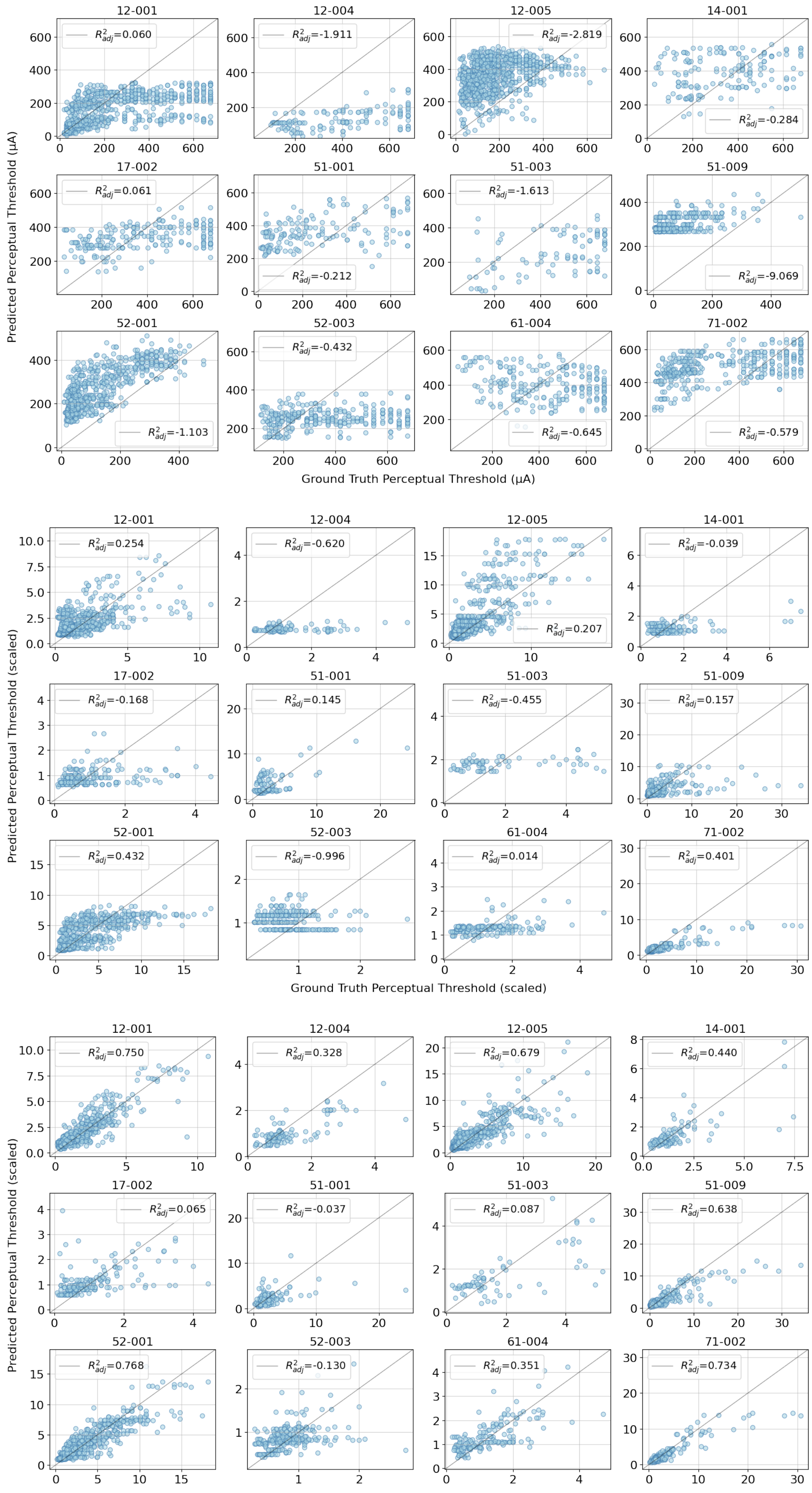

Fig. -B.3: Ground truth and XGB-R predicted perceptual thresholds based on 'Routine' measures (top), 'System Fitting' measures (middle), and 'Follow-Up' measures (bottom).