Telemonitoring the Parkinson's Disease Progression by tracking speech test records

Ye Tian * **Haishan Huang** *

Abstract

Monitoring symptoms of Parkinson's disease traditionally requires frequent and costly clinic visits, making it challenging to track disease progression accurately and consistently. This study aims to develop predictive models based on speech recordings to estimate symptom severity scores, providing an accessible alternative to in-person assessments. We construct two models—Bayesian regression and a hierarchical Bayesian model—that account for individual patient variability in symptom progression. Additionally, we introduce a shrinkage prior to explore penalized regression from a Bayesian perspective. Finally, a time trend analysis is incorporated for model comparison, using visualizations to illustrate typical patient symptom patterns.

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder, with around 20% of cases undiagnosed. Key symptoms include tremor, rigidity, and vocal impairment. The Unified Parkinson's Disease Rating Scale (UPDRS) assesses the severity of PD but requires clinic visits and does not reveal the underlying causes. Noninvasive telemonitoring offers a cost-effective alternative, reducing clinic visits and easing healthcare burdens.

This study uses data from Tsanas et al. (2010) with 42 subjects diagnosed with idiopathic Parkinson's disease (PD). We model the relationship between voice symptoms and UPDRS scores using home-recorded data. We account for individual differences by applying both linear regression and a hierarchical model, based on previous research.

2. Data

For the measurement of progression, diagnosis is required for at least two symptoms: tremor, bradykinesia, or rigidity. Subjects completed at least 20 home-based sessions, with data transmitted online for UPDRS score prediction. The tasks measured tremor, motor function, speech, and voice.

In total, we have 5875 observations with 21 variables:

- Response: *motor* UPDRS (y).
- Independent variables: 16 voice symptoms features. $x' = (x_1, ..., x_J)', J = 16$.
- Demographic index: subject, sex, age.

3. Model

3.1. Model 1: Bayesian regression model

Suppose the variance of responses is the same and follows the normal distribution.

$$y \mid \alpha, \beta, \sigma, X \sim \mathcal{N}(\mu, \sigma^2)$$

where $\mu = \alpha + X\beta$.

Thus the parameter space is $\theta = (\alpha, \beta, \sigma)'$.

Consider a noninformative prior distribution below,

$$\alpha \sim \text{Uniform}(-\infty, \infty)$$

$$\beta \sim \text{Normal}\ (0, 1)$$

$$\sigma \sim \text{Student}'\text{s}\ \ t\ (3, 0, 2)$$

3.2. Model 2: Hierarchical Bayesian model

Since each patient has multiple voice measurements, we can use a hierarchical Bayesian model where the response (*motor* UPDRS) is modeled, accounting for patient-grouped(g th,$g = 1, ..., 42$) variation.

The hierarchy model is composed of two levels:

- Observation level:

  $$y_{ig} \mid \sigma, \beta_g, X \sim N(\mu_g, \sigma^2)$$

  where $\mu_g = \alpha + X\beta_g$.

- Population level:

  $$\beta_g \mid \mu_\beta, \tau_\beta \sim N(\mu_\beta, \tau_\beta^2)$$

Thus the parameter space is $\theta = (\alpha, \sigma, \mu_\beta, \tau_\beta)'$.

Consider a noninformative prior distribution below,

$$\mu_\beta \sim \text{Normal}\ (0, 2)$$

$$\mu_\beta \sim \text{Normal}\ (0, 2)$$

$$\sigma \sim \text{Student}'\text{s}\ \ t\ (3, 0, 2)$$

3.3. Model assessment and selection

To assess its predictive accuracy, we compare the predicted values of the model, obtained from a Stan simulation, with the actual observed values, which could visually verify the prediction performance of the model. Additionally, we also conduct a complementary comparison of predictive performance using Leave-One-Out Cross-Validation (LOO-CV).

4. Result

4.1. Predict performance

4.1.1. Model 1: Bayesian regression model

To demonstrate predictive performance, we compare the predicted values from the Stan simulation with the actual values, along with results from a linear regression model as a frequentist benchmark. The simulation generated a total of 23,500,000 predicted values from four chains, with 1,000 post-warmup iterations each.

For the true value of the response variable, *motor* UPDRS (blue line), Figure 1 shows that the Bayesian linear regression (grey line) fits the observed data (blue line) better compared to the linear regression model using OLS (orange line). However, the former does not capture the bimodal characteristic. This suggests that the true population consists of at least two groups, each following a similar normal distribution. This observation motivates the introduction of a hierarchical model for further analysis.

Also, we do inference on the parameters of interest, which is shown in Table 1. To determine the appropriate number of significant digits, the number of decimal places is adjusted based on the magnitude of the Monte Carlo SE.

4.1.2. Model 2: Hierarchical Bayesian model

We continue to use the prediction value to visualize the performance of the model. From Figure 2, compared to the Bayes linear regression, the Bayes hierarchy model has an advantage in capturing a wider range of variation, which can be useful in expressing the difference between groups. In our case, it successfully shows the bimodal distribution of the true value which benefits from the additional information from the subjects themselves.

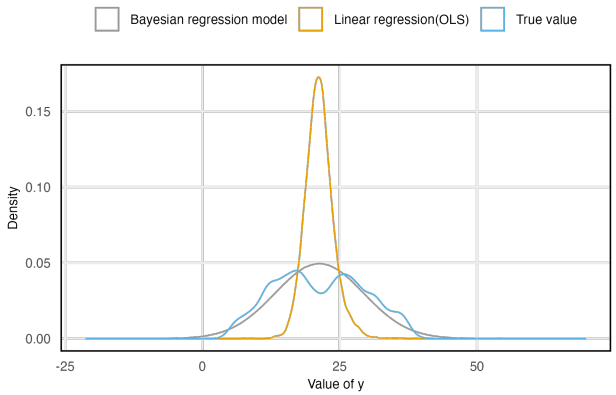

Figure 1. Predict comparison of *motor* UPDRS between Bayes linear regression model and linear regression(OLS).

Table 1. Parameter estimates of Bayes linear regression model.

Parameter	Estimate	MCSE	Rhat
alpha	32.12	0.032	1.002
beta[1]	-0.08	0.014	0.999
beta[2]	0.01	0.013	0.999
beta[3]	-0.05	0.013	1.000
beta[4]	-0.06	0.014	1.000
beta[5]	-0.13	0.013	0.999
beta[6]	-0.24	0.014	1.000
beta[7]	-1.10	0.011	1.000
beta[8]	-0.26	0.012	0.999
beta[9]	-0.29	0.013	0.999
beta[10]	0.36	0.013	0.999
beta[11]	-0.83	0.013	1.000
beta[12]	-1.43	0.013	1.000
beta[13]	-0.34	0.001	1.001
beta[14]	1.02	0.013	1.000
beta[15]	-6.52	0.013	1.000
beta[16]	2.78	0.013	1.000

4.1.3. Model assessment and selection

As a complementary predictive performance comparison in Section 4.1.2, we assess the predictive accuracy using computation results from the LOO-CV(Table 2), we could determine the Hierarchical model is optimal. It achieves a better predictive accuracy to the data, making it more suitable, especially considering that there are meaningful differences between subjects (as indicated by the higher complexity and p_{loo} value).

4.2. Model exploration

Based on the previous research about this data set, within the frequentist framework, there are two fields of interest to explore more. One is **feature selection**, by simplifying the model to solve the underlying collinearity among existing

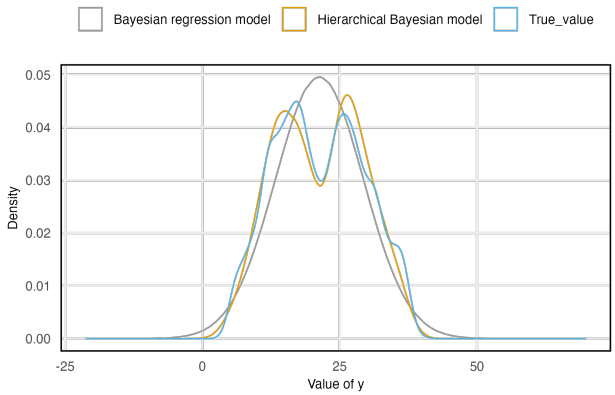

Figure 2. Predict comparison of *motor* UPDRS between Bayesian regression model and hierarchical Bayesian model.

Table 2. Comparison of Bayesian Regression Models on Prediction Performance Indicators, Model 1 refers to Bayesian regression model and Model 2 refers to Hierarchy Bayesian model

INDICATORS	MODEL 1	MODEL 2
KHAT < 0.7 %	100%	99.96%
ELPD_LOO	-20513.6	-12959.1
P_EFF	3.8	178.1
LOO_IC	41027.2	25918.1
ELPD_DIFF	0.0	-7554.6
SE_DIFF	0.0	80.7

variables. Another is to describe the **time trend**; since these data come from repeated measurement and the subjects are all in the earlier stage of PD, a clinical benefit could be the feasibility of tracking UPDRS changes in time.

Hence, we update our model to explore the feature selection but from a Bayesian point of view by using shrinkage prior. Then we chose three subjects with small, mediate, and big fluctuations throughout the six-month trial to verify whether our model could describe the time trend.

4.2.1. SHRINKAGE PRIOR

As an analogy to penalized regression techniques in the frequentist framework, including lasso and ridge regression, shrinkage priors in Bayesian penalization aim to shrink small effects to zero while maintaining true large effects. Meanwhile, shrinkage priors fit the Bayesian framework in a quite natural way since a prior is needed anyway, and shrinkage towards zero can be directly achieved by choosing a specific parameter from the prior(Van Erp et al. (2019)).

Here, we introduce the Horseshoes prior, which is a popular choice in Bayesian literature. For simplicity, we update Model 1, i.e. Bayesian regression model by adding the shrinkage prior to β as below,

$$\beta_j \mid \tau_j^2 \sim N(0, \tau_j^2)$$

$$\tau_j \mid \lambda \sim \text{Half} - \text{Cauchy}(0, \lambda), \text{ for } j = 1, ...p$$

$$\lambda \mid \sigma \sim \text{Half} - \text{Cauchy}(0, \sigma)$$

We then obtain the updated coefficients, which are presented in Table 3. To assess prediction performance, we compare the distribution of predicted values shown in Figure 3. This comparison indicates that incorporating the shrinkage prior makes it overly informative, resulting in poorer predictive performance compared to a noninformative prior in the Bayesian linear regression. Additionally, despite the updated coefficients, only two parameters are significantly reduced with a coefficient near zero, providing minimal improvement to the model.

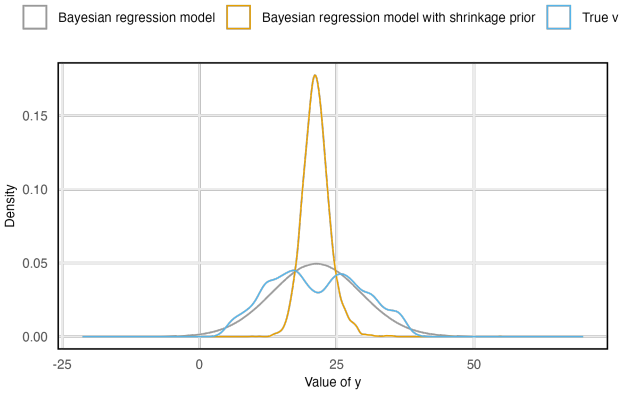

Figure 3. Predict comparison of *motor* UPDRS between Bayes linear regression model and Bayes linear regression model with shrinkage prior.

4.2.2. TIME TREND

To specify the model performance in time tracking, we choose subjects $16, 4$ and 17 to represent the subjects with small, mediate, and large fluctuations of *motor* UPDRS correspondingly throughout the six-month trial.

From Figure 4, we can determine that the Bayesian hierarchy model consistently shows a better ability to capture time trends across all subjects compared to the other two models. It adapts to both increases and decreases in the data.

5. Conclusion

In this study, we explored the use of sustained vowel measurements as predictors for the progression of average Parkinson's disease (PD) symptoms, linking dysphonia measures to UPDRS scores. We introduced two Bayesian

Table 3. Parameter estimates of Bayes linear regression model with shrinkage prior.

PARAMETER	ESTIMATE	SE
JITTER...	204.300	138.309
JITTER.ABS.	-46421.579	7587.423
JITTER.RAP	46.481	284.257
JITTER.PPQ5	-46.756	107.895
JITTER.DDP	16.742	89.742
SHIMMER	42.506	41.061
SHIMMER.DB.	-0.171	2.264
SHIMMER.APQ3	-33.186	86.039
SHIMMER.APQ5	-159.004	42.895
SHIMMER.APQ11	112.716	20.115
SHIMMER.DDA	-11.132	29.957
NHR	-21.345	4.826
HNR	-0.407	0.055
RPDE	1.055	1.291
DFA	-28.625	1.760
PPE	20.477	2.315
(INTERCEPT)	45.076	2.313

models—a Bayesian regression model and a hierarchical Bayesian model—and demonstrated that both outperform ordinary least squares (OLS) linear regression in predictive accuracy. Notably, the hierarchical Bayesian model further enhances predictive accuracy by capturing group-level variability within the data.

Additionally, we conducted feature selection in a Bayesian framework, building on the approach taken in prior research but adapting it to utilize Bayesian shrinkage priors rather than traditional frequentist methods.

Our data primarily included subjects with moderate symptoms (maximum *motor* UPDRS of 39.5, on a scale of 0 to 108), suggesting that our models currently apply best to early-stage PD. Whether these models can generalize to patients with more severe symptoms remains uncertain. Future studies involving patients with advanced PD could enhance the complexity and robustness of our predictive models.

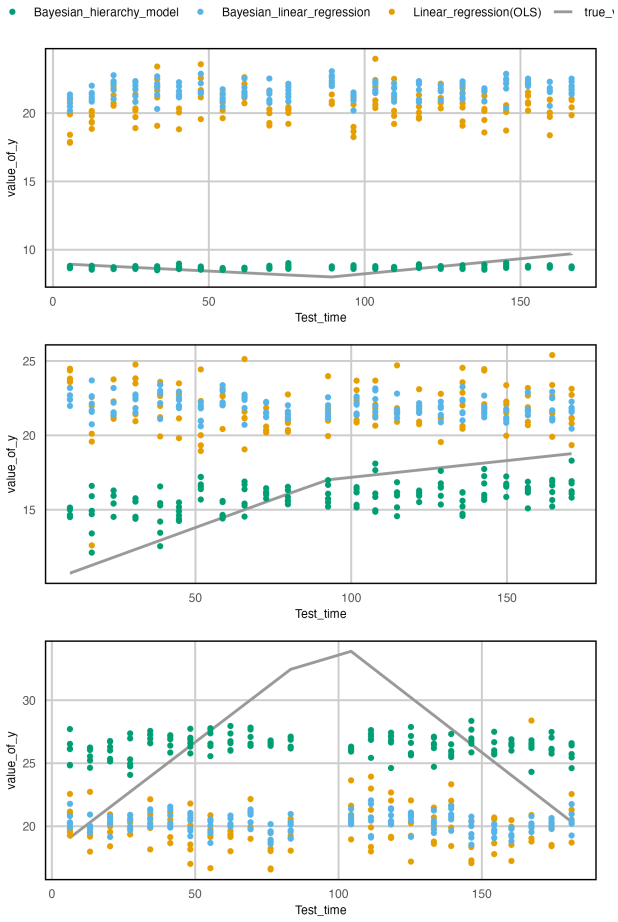

Figure 4. Predict comparison of *motor* UPDRS for subjects 16(top, with small fluctuation), 4(middle, with mediate fluctuation), and 17(bottom, with big fluctuation).

6. Appendix

6.1. Model 1: Stan code

```
// Stan model code
data {
  int<lower=0> N;
  int<lower=0> J;
  vector[J] x[N];
  vector[N] y;
  response vector
}

parameters {
  vector[J] beta;
  real<lower=0> sigma;
  real alpha;
}

model {
  // Priors
  beta ~ normal(0, 1);
  sigma ~ student_t(3, 0, 2);

  vector[N] mu;

  for (n in 1:N) {
    mu[n] = alpha +dot_product(beta, x[n]);
  }

  y ~ normal(mu, sigma);
}

generated quantities {
  vector[N] y_pred;
  vector[N] log_lik;
```

```
  for (n in 1:N) {
    y_pred[n] = normal_rng(alpha +
        dot_product(beta, x[n]), sigma);
    log_lik[n] = normal_lpdf(y[n] | alpha +
        dot_product(beta, x[n]), sigma);
  }
}
```

6.2. Model 2: Stan code

```
// Stan model code
data {
  int<lower=0> N;
  int<lower=0> J;
  int<lower=1> G;
  int<lower=1, upper=G> group[N];
  vector[J] x[N];
  vector[N] y;
}

parameters {
  // Group-level beta coefficients
  matrix[G, J] beta_raw;
  vector[G] alpha_raw;

  real<lower=0> sigma;
  real mu_alpha;
  real<lower=0> tau_alpha;
  vector[J] mu_beta;
  vector<lower=0>[J] tau_beta;
}

transformed parameters {
  matrix[G, J] beta;
  vector[G] alpha;

  // Transform raw parameters to obtain
     actual group-level coefficients
  for (g in 1:G) {
    alpha[g] = mu_alpha + tau_alpha *
        alpha_raw[g];
    for (j in 1:J) {
      beta[g, j] = mu_beta[j] + tau_beta[j]
          * beta_raw[g, j];
    }
  }
}

model {
  // Priors for hyperparameters
  mu_alpha ~ normal(0, 2);
  tau_alpha ~ normal(0, 2);
  mu_beta ~ normal(0, 2);
  tau_beta ~ normal(0, 2);

  // Priors for group-level raw parameters
     (non-centered parameterization)
  alpha_raw ~ normal(0, 1);
  to_vector(beta_raw) ~ normal(0, 1);

  // Prior for residual standard deviation
  sigma ~ student_t(3, 0, 2);

  vector[N] mu;

  for (n in 1:N) {
    int g = group[n];
    mu[n] = alpha[g] + dot_product(beta[g],
        x[n]); // group-specific linear
       predictor
  }

  // Likelihood
  y ~ normal(mu, sigma);
}

generated quantities {
  vector[N] y_pred;
  vector[N] log_lik;

  for (n in 1:N) {
    int g = group[n];
    y_pred[n] = normal_rng(alpha[g] +
        dot_product(beta[g], x[n]), sigma);
    log_lik[n] = normal_lpdf(y[n] | alpha[g
        ] + dot_product(beta[g], x[n]),
        sigma); // log-likelihood
  }
}
```

References

A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig. Accurate telemonitoring of parkinson's disease progression by noninvasive speech tests. *IEEE Transactions on Biomedical Engineering*, 57(4):884–893, 2010. doi: 10.1109/TBME.2009.2036000. URL https://doi.org/10.1109/TBME.2009.2036000.

S. Van Erp, D. L. Oberski, and J. Mulder. Shrinkage priors for bayesian penalized regression. *Journal of Mathematical Psychology*, 89:31–50, 2019. doi: 10.1016/j.jmp.2018.12.004. URL https://doi.org/10.1016/j.jmp.2018.12.004.