STATISTICAL MODELS FOR PREDICTION OF FANTASY FOOTBALL POINTS

ADAM K. JAUREGUI AND STEVE S. CHUNG

Abstract. Fantasy football modeling has seen a sudden spike in interest during the past few years from academic researchers. The objective of this paper is to predict the National Football League (NFL) quarterbacks' fantasy points. To this end, we consider three different models and evaluate their predictive potentials. We use the maximum likelihood method to estimate the parameters and the Gibbs sampler to draw predictive values. The dataset consists of the whole-career fantasy points for twenty chosen NFL quarterbacks. The 2014 NFL season was used as a testing period to validate our models. The results show that one model consistently outperformed the other two. Additional tests were performed with the *inflation-adjusted* fantasy point data to test the hypothesis that we need to compensate for the fact that NFL quarterbacks average more fantasy points today than in years past.

1. Introduction

Fantasy football is a skill-based game that has exploded onto the scene in recent years. The Fantasy Trade Sports Association estimates that 51 million people play fantasy sports, with 33 million of them playing fantasy football alone [12]. This has driven the fantasy sports industry to become a $70 billion per year industry [4]. Clearly, predicting players' fantasy points is now a big business and various models will need to be tested to provide the most accurate fantasy football projections.

After years of letting fantasy football research remain within the confines of football blog websites, academia is finally coming around to the topic, as there has been a recent surge in research on the prediction of the NFL player's fantasy points, specifically for the quarterback position. Boyd [2] took the average of the quarterback's statistics and adjusted the expected production by using a scaled z-score on the quarterback and on the opposing defense. His predictions, however, were

Key words and phrases. NFL fantasy football; Gibbs sampler; posterior predictive distribution.

not as accurate as the ESPN ranking system. Lutz [6] used the support vector regression and the neural networks in his fantasy football research on quarterbacks, measuring his success by taking the mean absolute error (MAE) and mean relative error (MRE) of his predictions. He concluded that the errors were too high for practical use.

In this work, we considered three models under classical frequentist and Bayesian settings. Under the frequentist approach, we used a fixed effects model where the mean and the variance vary weekly. Our underlying assumption comes from the fact that weekly performance may change due to factors such as minor injuries, match location, quality of opposition, and crowd noise and hence, affecting the mean and the variance. Taylor et al. [13] and Nevill et al. [7] have examined the effects of these factors on football players' performances. Under the Bayesian setting, we examined two models: one with the prior distribution on mean and another with the prior distributions on both mean and variance. Bayesian models have been used in many different aspects of sports forecasting, such as in Baio and Blangiardo [1], who used a Bayesian hierarchical model for the prediction of a team's soccer goals scored and allowed over the course of a season. Shirley [11] also used a Bayesian hierarchical model to compute the probability that the Denver Broncos would defeat the New York Jets in week 11 of the 2011 NFL season. They created a posterior distribution on the Denver Broncos' margin of victory by using a Markov chain Monte Carlo (MCMC) method to simulate 3,000 games. In our work, we also incorporate an MCMC approach to simulate NFL quarterbacks' fantasy points.

This paper is organized as follows. In section 2, we describe the datasets and three models that we used to forecast the fantasy points for each of our twenty quarterbacks. Also, a brief explanation of a MCMC method, the Gibbs sampler, is presented. In section 3, we present the results. In section 4, we conclude and suggest possible future work.

2. Datasets and Models

We only selected the quarterbacks who had started a minimum of 30 NFL games prior to the 2014 NFL season. Applying this filter eliminated the possibility of collecting research on rookies and backup quarterbacks. In addition, since we considered the 2014 NFL season as our test data, this means that we only needed to collect data on quarterbacks who started in at least one NFL game for the 2014 season. Thus, this led to the testing of twenty different NFL quarterbacks. All fantasy points and NFL quarterback starts data were collected from `fantasydata.com` and `pro-football-reference.com` using standard NFL scoring.

There are sixteen games played in an NFL season over a seventeen week span. When formulating our models, we decided to work with the career mean of a quarterback's fantasy points in the i-th week of the season (i = 1, 2,..., 17). The reason we calculate a quarterback's career mean fantasy points and variance by week instead of over his whole career is twofold: (1) we need to account for the weekly variation in fantasy points that quarterbacks produce, and (2) it allows for easy comparison with our test data.

Place Table 1 here.

To illustrate our data collection process, Table 1 shows the fantasy point data for Tom Brady of the New England Patriots. The columns correspond to the i-th week of the NFL season, and the rows correspond to the n-th game that Tom Brady played during that i-th week. From the table, we can see that Tom Brady scored 23.76 fantasy points in his first game played during week 1, scored -3.08 fantasy points in his second game played during week 1, and so on. The bottom-most row shows his fantasy point averages for each week. Due to injuries and rest, some weeks have more data than others.

As previously mentioned, we used three modeling approaches: (i) a frequentist approach, (ii) a Bayesian approach with a prior distribution on the mean, and (iii) a Bayesian approach with prior distributions on the mean and standard deviation. The details are as follows.

2.1. **Frequentist Approach.** One of the most important assumptions in parametric modeling is distributional form of the data. We have examined both descriptive and graphical summaries of the quarterbacks' fantasy points by week for all quarterbacks considered. There were any noticeable outliers and each weekly set of observations showed a fairly symmetric and bell-shaped distribution. Hence, we assume normally distributed random errors. Although normality tests may be required to validate our assumption but we do not perform them in our work.

Let $\mathbf{X}_i = [X_{1i}, X_{2i}, ..., X_{ni}]'$ be a demeaned-vector of i-th week fantasy points where the weekly sample size n varies for each week. We put a structural form of the data using a fixed effects model of the form

$$X_{ki} = \theta_i + \epsilon_{ki}, \tag{2.1}$$

where $\epsilon_{ki} \sim N(0, \sigma_i^2)$ and these are identical and independently distributed. The parameters $\theta_i = E(X_{ki})$ and $\sigma_i^2 = Var(X_{ki})$ correspond to the quarterback's mean fantasy point and variance for the i-th week, respectively. We estimate these parameters using the the maximum likelihood (ML) estimation method. The likelihood function of θ_i and σ_i^2 is then

$$f(\mathbf{X}_i|\theta_i,\sigma_i^2) = \prod_{k=1}^{n} f(X_{ki}|\theta_i,\sigma_i^2) = \frac{1}{(2\pi\sigma_i^2)^{n/2}} exp\left(-\frac{\sum_{k=1}^{n}(X_{ki}-\theta_i)^2}{2\sigma_i^2}\right). \tag{2.2}$$

Therefore, the ML estimators from above likelihood are

$$\hat{\theta}_i = \bar{X}_i, \tag{2.3}$$

$$\hat{\sigma}_i^2 = \frac{n-1}{n} s_i^2, \tag{2.4}$$

where $\bar{X}_i$ is the mean of the i-th week, and s_i^2 is the i-th sample variance. Since we are concerned with the point-wise prediction of weekly fantasy points and hence, the estimator (2.3) is only used. However, we may incorporate the (2.4) (or a function of (2.4)) if we want to perform forecasting based on prediction intervals.

2.2. **Bayesian Approach.** In the Bayesian analysis, choosing a prior distribution is one of the most important modeling aspects since it heavily influences the posterior distribution and therefore, affects the inference and prediction. Robert [8] serves a good reference for Bayesian methodology. We take the same structural form as in (2.1) but in this case, we put a prior distribution on θ_i and this is specified as a normal with mean μ and variance τ^2. Then, $\bar{X}_i$ is normally distributed with mean θ_i and variance σ^2/n. That is,

$$\bar{X}_i \sim N(\theta_i, \sigma^2/n), \tag{2.5}$$

$$\theta_i \sim N(\mu, \tau^2). \tag{2.6}$$

Hence, the posterior distribution $p(\theta_i|\bar{X}_i)$ of $\theta_i|\bar{X}_i$ is

$$N\left(\frac{\tau^2}{\frac{\sigma^2}{n}+\tau^2}\bar{X}_i + \frac{\frac{\sigma^2}{n}}{\frac{\sigma^2}{n}+\tau^2}\mu, \frac{\sigma^2\tau^2}{\sigma^2+n\tau^2}\right). \tag{2.7}$$

As a first Bayesian model, denoted BM1, we first estimate σ^2 using the maximum likelihood assuming constant variance for the whole career. It is easy to sample from the joint posterior distribution $p(\Theta|\bar{X}_i)$ using the Markov chain Monte Carlo (MCMC) methods. Many MCMC approaches are possible but we chose the Gibbs sampler. To simplify the notation, we let $\Theta = [\theta_i, \mu, \tau^2]$. The Gibbs sampling algorithm, at each iteration, cycles through the vector Θ and draws each subset conditional on the value of all the others. Geman and Geman [3] first considered this algorithm when making an analogy between images and statistical mechanics systems. One can show that the resulting set of samples converges in distribution to the true joint posterior distribution $p(\Theta|\bar{X}_i)$ for sufficiently large number of iterations; see [3], [10], and [9] for results and further discussions.

To apply the Gibbs sampler in our prediction, we compute the posterior predictive distribution and it is defined as

$$p(y|\bar{X}_i) = \int_{\Theta} p(y|\Theta, \bar{X}_i) p(\Theta|\bar{X}_i) d\Theta \tag{2.8}$$

$$= \int_{\Theta} p(y|\Theta) p(\Theta|\bar{X}_i) d\Theta, \tag{2.9}$$

where y denotes the predictive value. In general, this integral is not possible to evaluate analytically but by resorting to the Gibbs sampling, we can obtain the random samples from (2.9). As a starting point, denote $\Theta^{(t)} = [\theta_i^{(t)}, \mu^{(t)}, \tau^{2(t)}]$ as the state vector of the Markov chain at transition t. Then, we can generate $\Theta^{(t+1)}$ and y as follows.

1. $\theta_i^{(t+1)} \sim p(\theta_i|\bar{X}_i, \mu^{(t)}, \tau^{2(t)})$,
2. $\mu^{(t+1)} \sim p(\mu|\bar{X}_i, \theta_i^{(t+1)}, \tau^{2(t)})$,
3. $\tau^{2(t+1)} \sim p(\tau|\bar{X}_i, \theta_i^{(t+1)}, \mu^{(t+1)})$,
4. $y \sim p(y|\bar{X}_i, \theta_i^{(t+1)}, \mu^{(t+1)}, \tau^{2(t+1)})$.

Steps 1 through 3 give the samples from the joint posterior distribution $p(\Theta|\bar{X}_i)$ and Step 4 gives the samples from $p(y|\bar{X}_i, \Theta)$. Hence, averaging y's gives the mean of samples from the posterior predictive distribution $p(y|\bar{X}_i)$.

As another Bayesian model, denoted BM2, σ^2 is assigned a gamma prior distribution. That is,

$$\sigma^2 \sim Gam(1, 1), \tag{2.10}$$

with the shape and scale parameters 1 and 1, respectively. This is simply an exponential distribution with parameter 1. The Gibbs sampler is again applied under this setting and obtaining the predictive value y is same as above where the Markov chain state vector has another dimension (i.e., $\Theta^{(t)} = [\theta_i^{(t)}, \mu^{(t)}, \sigma^{2(t)}, \tau^{2(t)}]$). The Gibbs sampler was run through the `JAGS` (Just Another Gibbs Sampler) in the R package `rjags`.

3. Results

The objective of our modeling is to determine which of our three approaches is most accurate in predicting a quarterback's weekly fantasy points. To test the accuracy of each, we first sample $T = 1,000$ times from the frequentist distribution as well as the respective posterior predictive distributions from BM1 and BM2 for each week for each quarterback. After gathering predictive values, we test their accuracies by computing the mean square error (MSE) and mean absolute deviation (MAD), whose general formulas are

$$MSE = \frac{1}{17}\sum_{i=1}^{17}(y_i - \hat{y}_i)^2, \tag{3.1}$$

and

$$MAD = \frac{1}{17}\sum_{i=1}^{17}|y_i - \hat{y}_i|, \tag{3.2}$$

where y_i's are the the weekly 2014 NFL Quarterback Fantasy Football points, and $\hat{y}_i$'s are the predictive values.

The results are seen in Table 2, with Figures 1 and 2 providing an even more in-depth look into how the three models performed by week. The quarterbacks are arranged from most to least NFL starts to see if there is any decrease in the error based off of an increase in sample size, but there does not appear to be any relation. In Figure 1, as one can see, BM2 (the green line) clearly outperformed the other two models in terms of producing the least amount of error. The frequentist model (in black) performed better than BM1 (in red). Similar results can be concluded from Figure 2. From these results, it seems to reaffirm the importance of assigning a prior distribution on the variance whenever running Bayesian modeling, as well as suggesting that perhaps our choice for the prior distribution on the parameter μ was suspect.

Place Table 2 here.

In our research, some people may argue against the idea of using fantasy point data that goes back many years ago to predict a quarterback's fantasy points for the 2014 season. Peyton Manning's fantasy point data, for instance, goes back to 1998! Lutz [6] only included quarterback data from 2009 - 2013 as a precautionary measure, since he reasoned that going back even further might compromise his fantasy

Place Figure 1 here.

Place Figure 2 here.

football projections. This is due to the NFL recently enforcing rules that make it easier for quarterbacks to gain passing yards and touchdowns, consequently producing higher fantasy point scores than in the past. This is a seemingly valid point to make, but does it have merit? We decided to test this hypothesis with what we call "inflation-adjusted" fantasy points. The equation for it is a duplicate of how economists' compute the monetary inflation rate:

$$Y = \frac{100(B - A)}{A}, \tag{3.3}$$

where B in our case is the average quarterback fantasy points scored per game in the current season, A is the average quarterback fantasy score in the earlier season, and Y is our inflation rate.

As an example, in the 2013 NFL season, quarterbacks averaged 16.10 fantasy points per game. However, back in 1999, quarterbacks averaged 14.12 fantasy points per game. Thus, $Y = \frac{100(16.10-14.12)}{14.12} = 1.39$. This means that the inflation rate is 1.39. We would multiply all quarterbacks' fantasy points from the 1999 season by 1.39 to "inflation-adjust" the fantasy points.

We applied this formula to all quarterbacks tested to create a new model called Infl-BM2. It is an exact replica of BM2, except the fantasy point data for all quarterbacks are multiplied by that corresponding season's inflation rate. In Table 3, we compare the results of Infl-BM2 with the already existing results from BM2. 9 of the 20 quarterbacks benefited from having their fantasy points inflation-adjusted. It also appeared to benefit more so those quarterbacks who have had longer careers than those with shorter careers, as one would expect. Thus, there does appear to be some merit in only using recent quarterback data. Perhaps future work can insert a time-dependent parameter in the analysis that puts more weight into a quarterback's recent data and less weight into their old data.

Place Table 3 here.

4. Conclusion and Discussion

The objective of our research was to formulate a model that could predict an NFL quarterback's weekly fantasy points. To accomplish this objective, we gathered weekly NFL quarterback fantasy point data and built statistical models under the Bayesian and frequentist settings. In the Bayesian approach, we used two models: the first (BM1) had a prior distribution on the mean while the second (BM2) had prior distributions on both the mean and the variance. Using the Gibbs sampling technique, the results were encouraging as BM2 outperformed both BM1 and the frequentist approach. The inflation-adjusted BM2 model also scored encouraging results when working with quarterbacks who have had longer careers.

Though we expected all of our Bayesian models to outperform the frequentist approach, the fact that the frequentist model outperformed BM1 reaffirms the notion that even with our prior knowledge, having a prior distribution on the variance, which BM2 did, is important in obtaining optimum results. It also seems to suggest that the prior distribution placed on the mean parameter μ was, at best, suspect. And though we are satisfied with the success of BM2, much like Lutz [6] noted with his own fantasy football model, our errors, with BM2 averaging an error range between 5-8 fantasy points, are probably too high for practical use. Nonetheless, the fact that a Bayesian model with priors on the mean and variance improved upon the frequentist approach is a good stepping stone for researchers who intend to use Bayesian analysis with an MCMC approach in the modeling of fantasy football.

Fantasy football modeling is still within its infancy (at least to the academic researchers). The motivation of our work have come from the fact that, to the best of our knowledge, the Bayesian modeling is novel in fantasy football research. Our models were simple enough to make comparisons and build other models based on our models. There are many aspects of our model that can be expounded upon, such as adding more parameters. Obvious ones that come to mind are the effects of playing at home, the effect of playing a good/bad defense, and week-to-week dependence. As noted earlier, future research can look into adding a time-dependent parameter to a quarterback's fantasy point data set that puts more weight into their recent performance than in their earlier career games. Even the actual fantasy point numbers themselves can be broken down and studied, as a quarterback's weekly fantasy points are actually the sum of six different parameters! And last but not least, the choice of a suitable prior needs further research as well. It seemed logical to hypothesize that having the prior distribution as a quarterback's entire career fantasy point average might be a good predictor of his upcoming weekly fantasy point scoring, but the results of our research suggest otherwise.

5. Appendix

Game/Week	Wk1	Wk2	Wk3	Wk4	Wk5	Wk6	...	Wk17
1	23.76	17.26	6.92	4.34	22.56	19.88	...	8.12
2	-3.08	22.90	30.60	18.42	11.60	7.92	...	13.64
3	23.30	11.06	11.14	13.76	12.96	4.48	...	24.46
4	20.14	10.80	12.88	20.12	8.94	9.94	...	13.34
5	10.82	8.70	17.50	11.36	24.10	17.16	...	3.88
6	24.28	21.36	26.64	15.62	13.30	34.02	...	12.90
7	3.04	6.84	15.08	19.04	22.60	39.20	...	21.84
8	22.02	11.92	22.68	19.42	14.50	11.78	...	5.44
9	22.32	29.22	23.98	10.72	15.14	17.26	...	15.96
10	34.98	14.54	18.10	16.94	18.72	19.80	...	23.52
11	17.54	11.20	15.50	32.00	5.88	14.36	...	19.36
12	15.12	–	–	20.44	–	–	...	8.58
Mean	17.85	15.07	18.27	16.84	15.48	17.80	...	14.23

Table 1. Fantasy point data for Tom Brady of the New England Patriots.

Quarterback (from most to least NFL starts)	Mean MSE Freq.	Mean MSE BM1	Mean MSE BM2	Mean MAD Freq.	Mean MAD BM1	Mean MAD BM2
Manning, Peyton	114.89	124.3	79.46	8.70	9.06	7.37
Brady, Tom	129.05	137.47	86.01	9.11	9.43	7.67
Brees, Drew	120.82	131.36	56.36	8.55	9.10	5.92
Manning, Eli	123.42	133.24	86.08	8.78	9.23	7.57
Roethlisberger, Ben	162.54	172.69	124.94	9.59	9.90	7.83
Palmer, Carson	66.49	83.48	42.84	6.61	7.33	5.66
Rivers, Philip	111.76	120.94	81.42	8.44	8.82	7.48
Romo, Tony	91.53	99.78	48.08	7.26	7.73	5.40
Cutler, Jay	82.89	103.74	56.06	7.24	8.15	6.24
Flacco, Joe	103.93	18.28	68.82	8.14	8.66	6.46
Ryan, Matt	80.96	93.16	58.04	7.11	7.71	5.80
Smith, Alex	105.21	120.43	42.66	8.38	8.85	5.31
Rodgers, Aaron	144.20	157.68	77.00	9.38	10.07	7.20
Orton, Kyle	92.75	107.11	57.95	07.46	8.23	6.33
Stafford, Matthew	134.96	157.76	70.24	9.37	10.14	7.03
Dalton, Andy	99.67	122.13	59.35	7.71	8.74	5.85
Newton, Cam	136.31	184.43	80.04	9.35	10.90	7.40
Luck, Andrew	102.33	137.51	86.25	7.37	9.33	7.36
Tannehill, Ryan	90.95	114.60	62.43	7.94	8.57	5.96
Wilson, Russell	208.27	236.44	101.49	11.65	12.52	07.73

Table 2. Mean MSE and MAD for each quarterback by model. Each value is the average of the weekly estimates.

Quarterback (from most to least NFL starts)	Mean MSE Infl-BM2	Mean MSE BM2	Mean MAD Infl-BM2	Mean MAD BM2
Manning, Peyton	81.41	79.46	7.29	7.37
Brady, Tom	96.37	86.01	8.01	7.67
Brees, Drew	56.87	56.36	6.09	5.92
Manning, Eli	75.93	86.08	7.15	7.57
Roethlisberger, Ben	114.42	124.94	7.86	7.83
Palmer, Carson	35.50	42.84	4.86	5.66
Rivers, Philip	81.38	81.42	7.45	7.48
Romo, Tony	51.88	48.08	5.54	5.40
Cutler, Jay	50.49	56.06	5.88	6.24
Flacco, Joe	64.82	68.82	6.34	6.46
Ryan, Matt	58.52	58.04	5.98	5.80
Smith, Alex	38.31	42.66	4.94	5.31
Rodgers, Aaron	79.30	77.00	7.02	7.20
Orton, Kyle	67.60	57.95	6.63	6.33
Stafford, Matthew	77.11	70.24	7.31	7.03
Dalton, Andy	60.16	59.35	5.87	5.85
Newton, Cam	96.92	80.04	8.22	7.40
Luck, Andrew	84.88	86.25	7.31	7.36
Tannehill, Ryan	69.16	62.43	6.25	5.96
Wilson, Russell	106.96	101.49	7.97	7.73

TABLE 3. Infl-BM2 and BM2 Mean MSE and MAD comparison for each quarterback. Each value is the average of the weekly estimates.

FIGURE 1. MSE graphs comparing the frequentist model (black line), BM1 (in red), and BM2 (in green).

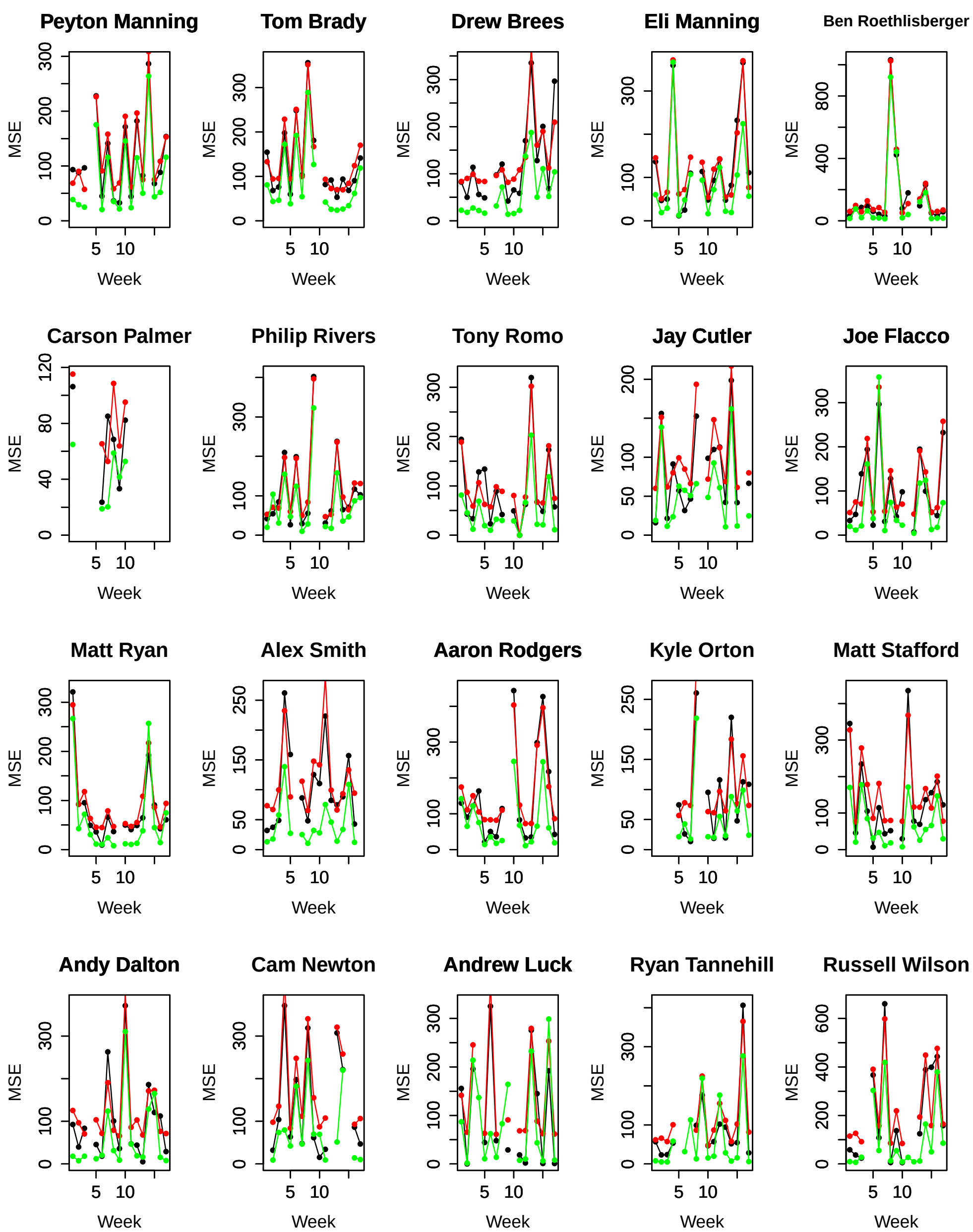
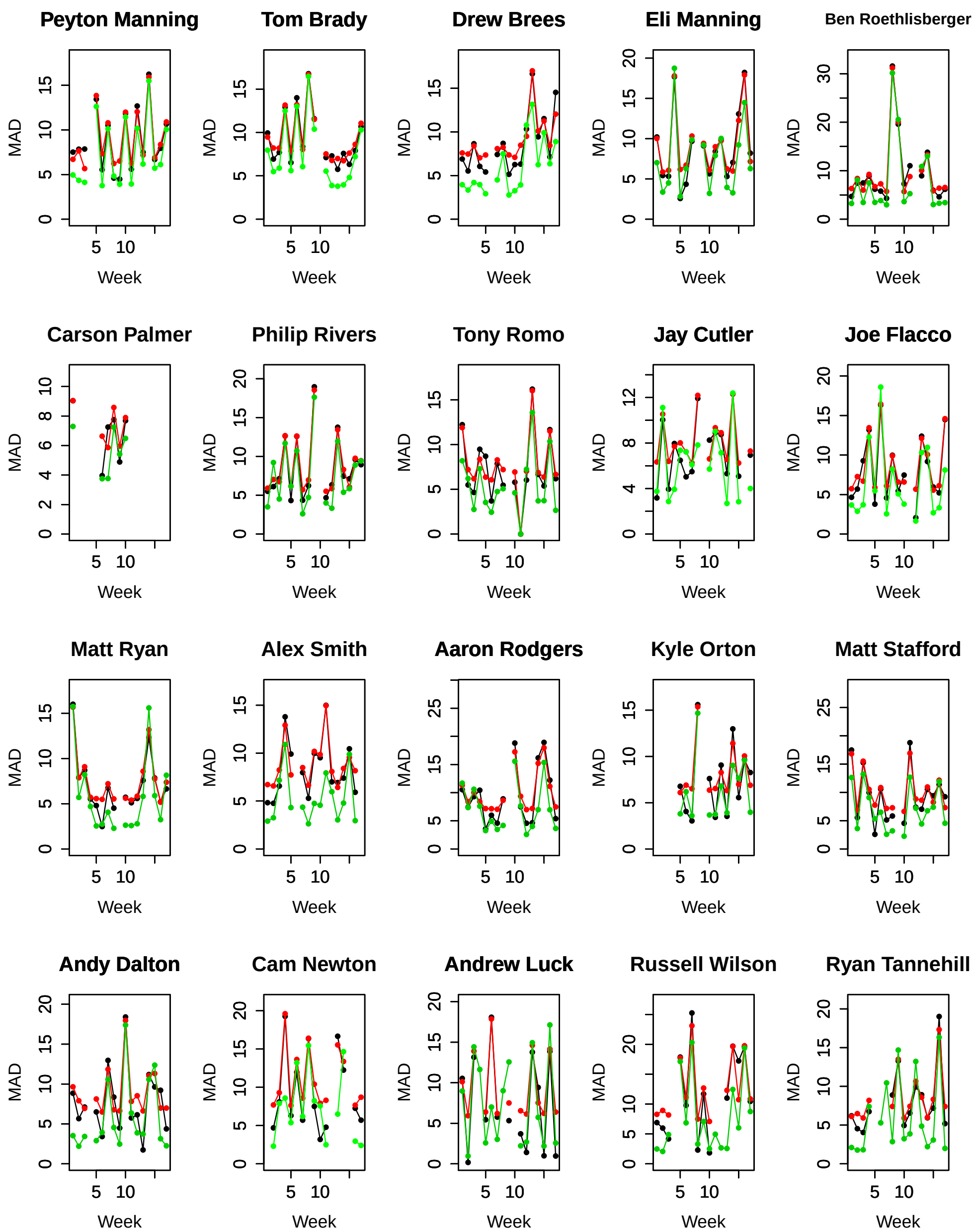


FIGURE 2. MAD graphs comparing the frequentist model (black line), BM1 (in red), and BM2 (in green).



References

[1] Baio, G, & Blangiardo, M. (2010). *Bayesian hierarchical model for the prediction of football results.* Journal of Applied Statistics, 37.2, 253-264.

[2] Boyd, E. (2014). *A New Method for Ranking Quarterback Fantasy Performance with Assessment Using Distances Between Rankings.* Theses, Dissertations and Capstone. Paper 827. http://mds.marshall.edu/etd827

[3] Geman, S. & Geman, D. (1984). *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images.* IEEE Transaction on Pattern Analysis and Machine Intelligence, 6, 721-741.

[4] Goff, B. (2013, August). *The $70 Billion Fantasy Football Market.* Retrieved from http://www.forbes.com/sites/briangoff/2013/08/20/the-70-billion-fantasy-football-market/

[5] Lee, K. (2015, October). *Watch Chris Christie Sack Jeb Bush In Fantasy Football Debate.* Retrieved from http://www.latimes.com/nation/politics/la-na-bush-fantasy-football-20151028-htmlstory.html

[6] Lutz, R. (2015). *Fantasy Football Prediction.* arXiv preprint. arXiv:1505.06918.

[7] Nevill, A. M., Balmer, N. J., & Williams, A. M. (2002). *The influence of crowd noise and experience upon refereeing decisions in football.* Psychology of Sport and Exercise, 3, 261-272.

[8] Robert, C. P. (2001). *The Bayesian Choice (2nd ed.).* New York, NY: Springer.

[9] Roberts, G. O. & Smith, A. F. M. (1993). *Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hasting algorithms.* Stochastic Processes and their Applications, 49, 207-216.

[10] Schervish, M. J. & Carlin, B. P. (1992). *On the convergence of successive substitution sampling.* Journal of Computational and Graphical Statistics, 1, 111-127.

[11] Shirley, K. (2011, November). *Hierarchical Bayesian Modeling with MCMC.* Retrieved from http://www2.research.att.com/$\sim$volinsky/DataMining/Columbia2011/Slides/Topic11-Bayesian.pdf

[12] Steinberg, L. (2014, August). *The Fantasy Football Explosion.* Retrieved from http://www.forbes.com/sites/leighsteinberg/2014/08/29/the-fantasy-football-explosion/

[13] Taylor, J. B., Mellalleu, S. D., James, N., & Shearer, D. A. (2008). *The influence of match location, quality of opposition, and match status on technical performance in professional association football.* Journal of Sports Sciences, 26, 885-895.

Department of Mathematics, California State University, Fresno, Peters Business Bldg., Rm 341, 5245 N. Backer Ave. M/S PB108, Fresno, CA 93740

E-mail address: freebird@mail.fresnostate.edu

Department of Mathematics, California State University, Fresno, Peters Business Bldg., Rm 341, 5245 N. Backer Ave. M/S PB108, Fresno, CA 93740

E-mail address: schung@csufresno.edu