The 2016 Equibase data set of American Quarter Horse starts in North America was analyzed, with the purpose of ranking the sires of the racehorses. A speed z-score derived from the race times and distances was used as a racing performance measure. Mixed effects models were used on various subsets of the data based on race distance and sire offspring number. The sire categorical variable was considered as a random effect. Various statistical criteria were used to optimize the model. The constructed models were then varied in terms of the random and fixed effects included, and the conditional modes of the sire effects were extracted from these models. The benefit of the sire ranking that comes from this analysis is that it is controlled for track, jockey, trainer, weather, and several other variables that can impact speed. Sires are typically valued for high rankings for offspring earnings and winners. Yet a sire with a low stud fee may still produce offspring with a high ranking using our z-score model. The offspring of this bargain sire have the potential to produce fast offspring that could pay a dividend on a relatively low cost investment. The model sire ranking approach described in this paper is clearly bringing a new approach to the field of sire rankings.

Key words: horse, mixed models, racing performance, sire

We are interested in quantifying the breeding value of American Quarter Horse sires, or obtaining some proxy for this breeding value. The performance of a race horse is thought to depend partly on its genetics, thus the race performance of a horse gives indirect information about the value of its sire in breeding. To assess the sire’s contribution to racing performance, it is necessary to take the other important factors into consideration, and we are limited to the information available in accounting for these other factors. To account for factors other than the sire, we fit a variety of mixed models, in which we assume that it is adequate to treat most factors as contributing linearly to race performance. Mixed models have been employed in studies of racing performance in Quarter Horses [5, 6, 11]; though these studies are typically focused on estimating quantities such as the heritability of particular traits as opposed to comparing individual horses.

Materials and Methods

The raw data

The raw data set is a proprietary data set (Equibase Co., LLC, Lexington, KY, U.S.A.) and contained information about the 56,999 American Quarter Horse race starts in 2016 in the United States and Canada [7]. The information used in this analysis were: the horse name (13,367 unique horses in races), the sire (1,245 unique sires), the dam (8,804 unique dams), the sex of the horse (male, female, gelding), the birth year of the horse, the country of the race, the track ID (72 unique tracks), the racing date, the wind speed, the wind direction, the weather, the temperature, the track condition, the race distance, the race time, the jockey (665 unique jockeys), the trainer (1,593 unique trainers), and the weight carried by the horse. An earnings variable was also in the data set and it was used to calculate sire total offspring earnings and sire average offspring earnings as two methods of sire rankings (Table 1). The data set was also used to calculate the number of winners per sire as shown in Table 1. In Table 1 the rankings for total earnings, average earnings, average z-score, and number of winners
Table 1. Ranking of sires with minimum of 10 offspring derived from 350 yard races, 300 yard races, and all races compared to rankings by total offspring earnings, average earnings, average z-score of sire, and number of winners

| Sire                | No. of race outcomes | No. of offspring | 350 yard median rank | 350 yard mean rank | 300 yard median rank | 300 yard mean rank | All distances median rank | All distances mean rank | Total earnings rank | Average earnings rank | Average z-score rank | No. of winners rank |
|---------------------|----------------------|------------------|----------------------|-------------------|----------------------|---------------------|-------------------------|------------------------|-------------------|----------------------|----------------------|---------------------|
| Apollitical Jess    | 233                  | 117              | 1                    | 1.0               | 7                    | 7.7                 | 2                       | 1.8                    | 2                 | 2                    | 9                    | 4                   |
| Fantastic Corona Jr | 81                   | 34               | 2                    | 2.0               | 6                    | 5.9                 | 1                       | 1.2                    | 42                | 20                   | 2                    | 47                  |
| One Sweet Jess      | 236                  | 109              | 3                    | 3.0               | 12                   | 12.2                | 13                      | 14.4                   | 11                | 22                   | 67                   | 5                   |
| Mr Jess Perry       | 175                  | 110              | 4                    | 4.7               | 33                   | 32.8                | 19                      | 18.1                   | 6                 | 13                   | 26                   | 13                  |
| Corona Cartel       | 204                  | 115              | 5                    | 5.3               | 3                    | 2.6                 | 6                       | 5.6                    | 5                 | 9                    | 11                   | 11                  |
| Carters Cartel      | 236                  | 118              | 6                    | 6.1               | 10                   | 9.7                 | 8                       | 8.7                    | 15                | 43                   | 13                   | 7                   |
| Desirio             | 76                   | 36               | 7                    | 7.6               | 14                   | 13.8                | 22                      | 22.5                   | 45                | 57                   | 84                   | 42                  |
| Foose               | 128                  | 62               | 8                    | 7.1               | 13                   | 14.5                | 24                      | 23.5                   | 20                | 34                   | 18                   | 20                  |
| Kiddy Up            | 44                   | 19               | 9                    | 10.9              | 3                    | 3.1                 | 3                       | 3.3                    | 73                | 69                   | 3                    | 61                  |
| Pyc Paint Your Wagon| 404                  | 212              | 10                   | 10.5              | 18                   | 16.7                | 11                      | 11.5                   | 4                 | 30                   | 12                   | 1                   |
| Walk Thru Fire      | 231                  | 127              | 10                   | 9.8               | 2                    | 2.6                 | 8                       | 7.9                    | 3                 | 14                   | 7                    | 3                   |
| Jess Zoomin         | 67                   | 28               | 11                   | 11.7              | 36                   | 36.0                | 32                      | 31.7                   | 76                | 97                   | 32                   | 52                  |
| Winners Version     | 133                  | 56               | 13                   | 12.6              | 66                   | 62.8                | 49                      | 43.6                   | 44                | 63                   | 25                   | 48                  |
| Good Reason Sa      | 63                   | 34               | 14                   | 14.5              | 16                   | 15.7                | 15                      | 15.3                   | 26                | 4                    | 5                    | 49                  |
| Favorite Cartel     | 156                  | 80               | 15                   | 15.3              | 30                   | 31.1                | 17                      | 16.8                   | 7                 | 7                    | 53                   | 16                  |
| Ivory James         | 303                  | 157              | 16                   | 15.5              | 1                    | 1.9                 | 16                      | 16.3                   | 9                 | 49                   | 14                   | 2                   |
| Dominyin            | 77                   | 42               | 17                   | 16.5              | 5                    | 4.9                 | 4                       | 4.5                    | 51                | 75                   | 42                   | 33                  |
| Rock Solid Jess     | 77                   | 34               | 18                   | 18.0              | 35                   | 34.7                | 20                      | 20.5                   | 52                | 45                   | 46                   | 59                  |
| Tres Seis           | 155                  | 87               | 20                   | 20.1              | 55                   | 52.5                | 23                      | 22.5                   | 19                | 53                   | 55                   | 17                  |
| One Famous Eagle    | 280                  | 151              | 21                   | 21.8              | 15                   | 12.8                | 12                      | 11.9                   | 1                 | 5                    | 36                   | 6                   |
| Chicks Regard       | 78                   | 41               | 22                   | 24.8              | 87                   | 84.4                | 85                      | 86.0                   | 58                | 96                   | 54                   | 55                  |
| Oak Tree Special    | 76                   | 46               | 22                   | 23.6              | 89                   | 85.8                | 52                      | 51.5                   | 31                | 42                   | 69                   | 29                  |
| Pool                | 39                   | 18               | 22                   | 21.9              | 25                   | 28.1                | 33                      | 37.9                   | 97                | 94                   | 112                  | 79                  |
| Capo de Capi        | 42                   | 21               | 25                   | 24.6              | 10                   | 9.9                 | 5                       | 5.5                    | 35                | 6                    | 16                   | 54                  |
| American Runaway    | 52                   | 25               | 26                   | 25.3              | 55                   | 52.5                | 40                      | 40.5                   | 27                | 1                    | 134                  | 71                  |
| Mighty B Valiant    | 25                   | 16               | 27                   | 26.1              | 42                   | 42.0                | 7                       | 7.7                    | 105               | 78                   | 20                   | 98                  |
| Tempting Dash       | 152                  | 87               | 27                   | 24.0              | 26                   | 24.4                | 18                      | 15.9                   | 30                | 76                   | 19                   | 15                  |
| Tr Dasher           | 146                  | 72               | 27                   | 28.4              | 50                   | 49.9                | 43                      | 43.5                   | 38                | 120                  | 61                   | 21                  |
| Heza Fast Dash      | 165                  | 90               | 29                   | 30.1              | 49                   | 48.1                | 44                      | 43.3                   | 12                | 29                   | 72                   | 10                  |
| Big Daddy Cartel    | 62                   | 33               | 31                   | 30.7              | 74                   | 74.4                | 35                      | 35.9                   | 49                | 24                   | 95                   | 77                  |

The sires of Table 1 were chosen because they were the top ranked thirty sires by median from the 350 yard data set model.

are based on races from all distances. The sires of Table 1 were chosen because they were the top ranked thirty sires by median from the 350 yard data set model.

Construction of multiple data sets

The data was processed using the R software environment [10]. In addition to the lme4 package, R tools and the R packages installr [9], openxlsx [12], and ggplot2 [13] were used during the analysis of the data.

Average race speeds were calculated and abnormal times were removed, leaving 56,181 race outcomes. The data were subsetted in nine different ways by using two criteria: the choice of distances included and set of sires included. There were three different ways that the data was subsetted by distances: subsetting to include multiple distances, subsetting to include only race outcomes at 350 yards, and subsetting to include only race outcomes at 300 yards.

The subsetting to include multiple race distances was done by including only distances with at least one hundred race outcomes with finite speeds. The distances of 350 and 300 yards were chosen because these are the two race distances with the greatest numbers of race outcomes. After removing race outcomes with unrealistic, non-finite, or very slow speeds, the two distances had 14,127 and 13,143 race outcomes, respectively. (The process of removing very slow speeds is explained in the next paragraph.)

A z-scaling procedure was performed with the intent of making the speeds comparable between different race distances. In this z-scaling procedure, the mean speed was calculated for each racing outcome by dividing the race distance by the time. Then the mean speed for each distance was subtracted from each individual speed at this distance. This quantity was divided by the standard deviation of the speeds at that distance. If the value of this new variable was
less than negative four, (corresponding to a race outcome slower than four standard deviations below the mean) then the corresponding race outcome was removed. After this removal, the races at each distance were scaled a second time by subtracting the new mean and dividing by the new standard deviation at each distance. These speed z-scores were used as a measure of race performance in the fitted models.

After this z-scaling procedure was performed, the data was subsetted to include only sires that met cutoffs for number of race outcomes and/or number of offspring. There were three ways of subsetting by sire: (1) all sires, (2) only sires with at least ten race outcomes and at least three offspring, and (3) only sires with at least ten offspring. We are particularly interested in the sire effects for sires with at least a certain number of offspring or race outcomes, as we would expect our assessment of these sires to be more accurate; therefore, it is of interest to see whether this subsetting makes a difference in the model-based ranking of the sires. In the multiple-distance data sets, this subsetting by sire caused the number of race outcomes for some race distances to drop below one hundred. These distances were kept in the data. The three different ways of subsetting by distance and the three different ways of subsetting by sire result in nine data sets on which models were built.

Modifications of other variables

After the z-scaling and subsetting processes, additional new variables were constructed from the variables in the raw data.

A new variable (designated as age.est.day) describing the racing horse’s estimated age in days on the day of race was defined. This variable was calculated by taking the horse’s birth year, treating the horse as if it were born on July 2nd of that year (the middle-most day of years that are not leap years), and subtracting this date from the date of the race. Later scaling of the variables made this choice of day irrelevant. A month variable indicating the month in which the race occurred was also constructed from the race date.

The raw data contains a numeric variable for wind speed and a categorical variable for wind direction. A new cross-wind variable was constructed that has a value of zero when the wind direction is not “Cross” but has a value of the wind speed when the wind direction is “Cross”. Analogous variables were constructed for headwind and tailwind. Lastly, a parallel wind variable was constructed that takes the value of the tailwind speed when the wind direction is “Tail”, takes the negative of the value of the headwind speed when the wind direction is “Head”, and takes a value of zero otherwise. All wind speeds were converted to kilometers/hour and all temperatures were converted to degrees Celsius.

Scaling

Z-scaled versions of most numeric variables were created; each was centered by its mean and divided by its standard deviation. The exceptions to this were the new wind speed variables. These were scaled by dividing by their respective standard deviations, but they were not re-centered about their means. This was done to maintain the meaning of zero as no wind in the relevant direction; zero was the mode for these variables.

Method details

In some sense, it is problematic to estimate individual random effects, as they are not technically parameters of the model (i.e., the parameters of their distribution are the true model parameters associated with the random-effects terms) [8]. The values obtained from our analysis are not technically estimates; they are the conditional modes of the random effects [1, 3]. However, for brevity, throughout this paper, we refer to the conditional modes of the random effects as estimates, as the conditional modes of the random effects are the values we use to assess the individual levels (such as the individual sires) associated with the random-effects terms. More specifically, we use the conditional modes of the sire random effects to rank the sires. And we fit mixed models using the maximum likelihood (ML) method or the restricted maximum likelihood (REML) method.

Model selection using AIC and BIC

The Akaike Information Criterion (AIC) was used as a model selection criterion. For theoretical reasons, it was necessary to use the ML method (as opposed to the REML method) of fitting the models to obtain AIC values [8]. In the model selection procedure, we started with eighteen different top-level models. In R syntax, these starting models were of the following form. (Note the 1 at the beginning of the formula indicating an intercept, and note the formatting of the dam, sire, horse, trainer, jockey, and track ID variables, indicating treatment as random-effects terms.)

\[
\text{speed.z} \sim 1 + (1|\text{Dam}) + (1|\text{Sire}) + (1|\text{Horse}) + (1|\text{TRAINER}) + (1|\text{JOCKEY}) + (1|\text{Track.ID}) + A + B + \text{Weather} + C + \text{track.condition} + \text{weight.carried.scaled} + 1
\]

where \(A\) can take values of

- \(\text{age.est.days.scaled} + 1\) (\(\text{age.est.days.scaled}^2\)) + sex.category
- \(\text{age.est.days.scaled} + \text{sex.category} + \text{age.est.days.scaled};\text{sex.category},\)
- \(\text{age.est.days.scaled} + 1\) (\(\text{age.est.days.scaled}^2\)) + sex.category + sex.category;\text{age.est.days.scaled} + \text{sex.category};1

(\(\text{age.est.days.scaled}^2\),
model forms obtained from this procedure were used when path but ended on the same form of the model. The same distance data, the BIC selection procedure took a different the ten-offspring sire-subsetted version of the multiple race tion procedure (though the AIC- and BIC-selected forms of each time by the AIC selection procedure, and the same three data sets, the same form of the model was selected for these data sets that included multiple race distances. For these model selection procedure.

The procedure was then repeated using the Bayes Informa derived from was chosen, and the process came to an end. If the derived model with the lowest AIC was derived, then this model was selected to have models

A number of mixed effects models were fit on the data using the R package lme4 [2], with the speed z-score as the dependent variable. In these models, each dam, sire, racehorse, jockey, trainer, and track was treated as an associated random effect, and the other included variables were treated as associated fixed effects. From these eighteen models, the one with the lowest AIC was selected to have further models derived from it. Each of the derived models differed by having one of the many fixed-effects terms removed (when a particular term was removed, all terms in the model that were powers of that term, interactions with that term, or interactions with powers of that term were removed as well).

Comparison of AIC values between models with different random effects is not straightforward [8], so no random-effects terms were allowed to be removed during this part of the model selection procedure.

The AIC values for the derived models were compared to each other. If the derived model with the lowest AIC value had a lower AIC value than the model from which it was derived, then this model was selected to have models derived from it, and the process reiterated. If the derived model with the lowest AIC did not have a lower AIC value than the model it was derived from, then the model it was derived from was chosen, and the process came to an end. The procedure was then repeated using the Bayes Information Criterion (BIC) instead of the AIC.

This model selection procedure was performed only on the data sets that included multiple race distances. For these three data sets, the same form of the model was selected each time by the AIC selection procedure, and the same form of the model was selected each time by the BIC selection procedure (though the AIC- and BIC-selected forms of the model were different from each other). When applied to the ten-offspring sire-subsetted version of the multiple race distance data, the BIC selection procedure took a different path but ended on the same form of the model. The same model forms obtained from this procedure were used when fitting models on the 350 and 300 yard data, and this model selection procedure was not performed on the 350- and 300-yard-specific data sets.

Removal of random-effects terms

In the next section we see that we eventually arrive at 3 distinct model forms. Model versions without some of the random-effects terms were fit on each of the nine data sets. The five versions fitted were: all terms (dam, sire, horse, trainer, jockey and track), without dam (sire, horse, trainer, jockey and track); without dam and jockey (sire, horse, trainer and track); without dam and trainer (sire, horse, jockey and track); and without dam, trainer, and jockey (sire, horse, and track).

There is a specific reason for removing the dam term first. In a model with terms for horse and dam but no term for sire, the horses would be nested within dams. However, in this data, the number of dams in the data is large compared to the number of sires, as many of the dams have only one offspring racing in the year 2016 (In the processed multirace-distance data set with no subsetting based on sire, there are 13,294 racing horses and only 8,762 dams). In a model with terms for horse and sire but no term for dam, each horse has only one sire and each sire can have many offspring, so horse is nested within sire, and the interpretation of the horse effects is relatively straightforward, as the horse effects account for variation between horses that is not accounted for by their sire.

With these various subsets and approaches, we then name various models in a descriptive manner. For example, the model “mixedmod.350yards.no.sire.subsetting.014.reml.nodam”, is the dataset derived from 350 yard races, all sires independent of number of offspring, uses REML modeling, and excludes dam from the model.

Results

The form of the model obtained by using AIC as the model selection criterion was:

\[ \text{speed.z} \sim 1 + (1|\text{Dam}) + (1|\text{Sire}) + (1|\text{Horse}) + (1|\text{TRAINER}) + (1|\text{JOCKEY}) + (1|\text{Track.ID}) + \text{age.est. days.scaled} + I(\text{age.est. days.scaled}^2) + \text{sex.category} + \text{sex. category:age.est. days.scaled} + \text{sex.category:age.est. days.scaled}^2) + \text{headwind.scaled} + \text{tailwind.scaled} + \text{crosswind.scaled} + \text{weather} + \text{temperature.celsius.scaled} + \text{track.condition} + \text{weight.carried.scaled} + 1 \] (weight.carried.scaled^2)

This AIC-selection model form was indicated in model names using the suffix “014”.

The form of the model obtained by using BIC as the model selection criterion was:

\[ \text{speed.z} \sim 1 + (1|\text{Dam}) + (1|\text{Sire}) + (1|\text{Horse}) + (1|\text{TRAINER}) + (1|\text{JOCKEY}) + (1|\text{Track.ID}) + \text{age.est. days.scaled} + I(\text{age.est. days.scaled}^2) + \text{headwind.scaled} + \text{tailwind.scaled} + \text{crosswind.scaled} + \text{weather} + \text{temperature.celsius.scaled} + \text{track.condition} + \text{weight.carried.scaled} + 1 \] (weight.carried.scaled^2)
days.scaled + 1 (age.est.days.scaled^2) + sex.category +
headwind.scaled + tailwind.scaled + crosswind.scaled +
temperature.celsius.scaled + track.condition + weight.
carried.scaled

This BIC selected model form is indicated by 002.07.10.

A third, relatively simple form of the model was also fit
for each of the nine data sets:
speed.z ~ 1 + (1|Dam) + (1|Sire) + (1|Horse) +
(1|TRAINER) + (1|JOCKEY) + (1|Track.ID) + age.est.
days.scaled + 1 (age.est.days.scaled^2) + sex.category +
headwind.scaled + tailwind.scaled + crosswind.scaled

This relatively simple model form was indicated in model
names using the suffix “simple”. The fixed-effects terms
in the “simple” model were chosen because they had the
highest correlation coefficients when speed z-scores were
plotted against them.

Removal of random-effects terms

A maximum likelihood analysis for the AIC and BIC
models produced two base models (designated ML). The
form of this ML AIC model was the same as the AIC model
described in the previous section. The form of this ML BIC
model was the same as the BIC model described in the
previous section. Using the three model forms, versions
without some of the random-effects terms were fit on each
of the nine data sets. The five versions fitted were: all terms
(dam, sire, horse, trainer, jockey and track), without dam
(sire, horse, trainer, jockey and track); without dam and
jockey (sire, horse, trainer and track); without dam and
trainer (sire, horse, jockey and track); and without dam,
trainer, and jockey (sire, horse and track). Using REML
optimization we then produce 5 versions of the modeling per
approach. As previously mentioned, the 3 model types are
“014”, “002.07.10”, and the “simple”. The 5 model versions
and the 3 base model types produce fifteen base ranking
approaches. If one then adds the previously mentioned two
maximum likelihood models (ML) applied to the “014”,
and “002.07.10” types, then seventeen different ranking
approaches are available. We then applied the fifteen base
rankings to the 300 yard and 350 yard data sets. We applied
all seventeen ranking approaches to the all distance data.
Averages and medians of the ranking are displayed in Tables
1 and 2.

The track variable was not omitted from the models, as
its variance component was large compared to the other
random-effects terms, including sire, at least in the models
checked. Table 3 gives the variance components associated
with the random effects in the model “mixedmod.350yards.
no.sire.subsetting.014.reml.nodam”, which has the highest
R^2-like value of the models fit after the initial model selec-
tion process (This R^2-like value is explained in a later
section). The relatively large value of the track variance
component can be observed.

Assessing the sire effects and ranking

(For brevity, we refer to the conditional modes of the
random effects as estimates of those random effects, even
though they are not technically estimates) Estimates of
the sire random effects were obtained from each model
using the ranef function. The ranef function in R extracts
the conditional modes of the random effects from a fitted
model. The sire random-effects estimates were sorted for
each model in order to obtain sire ranks. Within each set of
REML models fit on one of the nine data subsets, the median
and mean ranks were calculated for each sire. Generally, the
list of top ranking sires obtained from models fit on the same
data set were similar.

Diagnostics

Diagnostic plots were made for a broad sampling of
the models to check underlying assumptions. There are
multiple ways of defining residuals and fitted values for
mixed models, and the fitted values used here include the
random effects such that the residuals can be considered
estimates of the errors [8]. From residual versus fitted values
plots it was seen that the residuals increase slightly with
the fitted values from the normal quantile-quantile plot
of the residuals and the corresponding histogram of the
residuals, it was apparent that the residuals were slightly
left-skewed; therefore, there is a violation of the assumption
of the normality of the errors. These diagnostic plots were
not constructed for every model, but this left-skewness was
apparent in the few checked.

The quantile-quantile plots of the random-effects
estimates from the model “mixedmod.350yards.no.sire.
subsetting.014.reml.nodam”, allowed assessment of the
normality of the random effects. The horse random effects
were somewhat left-skewed, but generally these random-
effects estimates had fairly close to normal distributions.
Diagnostic plots of the residuals on the individual fixed-
effects predictor variables revealed no obvious issues,
though some of the plots show some slight asymmetry in
the residual density in regions with fewer points.

To check whether the normality of the random-effects
estimates was merely due to the fact that the corresponding
terms were treated as random effects in the model, a few
alternative models were fit in which some terms were
treated as fixed-effects terms instead of random-effects
terms. All of these alternative models were based on the
“mixedmod.350yards.no.sire.subsetting.014.reml.nodam”
model; i.e., this was the only model that was checked in
this manner. Normal quantile-quantile plots of these fixed-
effects estimates showed a good approximation to linearity.
Goodness of fit

To characterize the goodness of fit of the mixed models, a quantity based on $R^2$ for simple linear models was utilized. For mixed models, there are a number of $R^2$ analogs, which serve a variety of different purposes [4]. In our case, we desired to quantify the closeness between the actual values of the response and the mixed model fitted values, with the random-effects estimates included in the calculation of those fitted values. We utilized a straightforward method contributed by Jarrett Byrnes in the r-sig-mixed-models mailing list; in this method, simple linear regression is performed between the fitted values and the actual values of the response, and the $R^2$ value obtained from this simple regression is used as a goodness of fit measure of the mixed model [4]. Among the models fitted after the initial variable selection process, these $R^2$ values range from 0.600 to 0.792.

We now present the sire ranking results. Recall that there are 9 datasets comprised of three variations on race distance: 1) 350 yard races, 2) 300 yard races, and 3) all distances combined; and three variations on sampling: 1) minimum of 10 offspring, 2) minimum of 3 offspring and 10 racing outcomes. We now present the sire ranking results. Recall that there are 9 datasets comprised of three variations on race distance: 1) 350 yard races, 2) 300 yard races, and 3) all distances combined; and three variations on sampling: 1) minimum of 10 offspring, 2) minimum of 3 offspring and 10 racing outcomes.

| Sire               | 350 yards, 10 offspring median rank | 350 yards, 10 offspring mean rank | 350 yards, all offspring median rank | 350 yards, all offspring mean rank | 350 yards, 3 offspring median rank | 350 yards, 3 offspring mean rank |
|--------------------|-------------------------------------|----------------------------------|-------------------------------------|-----------------------------------|-----------------------------------|----------------------------------|
| Apollitical Jess   | 1                                  | 1.0                              | 2                                   | 1.6                               | 1                                 | 1.0                              |
| Fantastic Corona Jr| 2                                  | 2.0                              | 1                                   | 1.4                               | 2                                 | 2.0                              |
| One Sweet Jess     | 3                                  | 3.0                              | 3                                   | 3.0                               | 3                                 | 3.0                              |
| Mr Jess Perry      | 4                                  | 4.7                              | 7                                   | 7.3                               | 5                                 | 5.3                              |
| Corona Cartel      | 5                                  | 5.1                              | 6                                   | 5.9                               | 6                                 | 5.3                              |
| Carters Cartel     | 6                                  | 6.1                              | 5                                   | 5.5                               | 5                                 | 5.4                              |
| Desirio            | 7                                  | 7.6                              | 7                                   | 7.3                               | 8                                 | 7.9                              |
| Foose              | 8                                  | 7.1                              | 8                                   | 7.3                               | 7                                 | 6.7                              |
| Kiddy Up           | 9                                  | 10.9                             | 5                                   | 7.7                               | 9                                 | 10.7                             |
| Pyc Paint Your Wagon| 10                                | 10.5                             | 12                                  | 11.3                              | 10                                | 10.5                             |
| Walk Thru Fire     | 10                                 | 9.8                              | 11                                  | 9.9                               | 10                                | 9.5                              |
| Jess Zoomin        | 11                                 | 11.7                             | 12                                  | 12.3                              | 12                                | 12.7                             |
| Winners Version    | 13                                 | 12.6                             | 15                                  | 14.5                              | 12                                | 12.4                             |
| Good Reason Sa     | 14                                 | 14.5                             | 12                                  | 11.5                              | 14                                | 13.7                             |
| Favorite Cartel    | 15                                 | 15.3                             | 14                                  | 14.7                              | 15                                | 14.9                             |
| Ivory James        | 16                                 | 15.5                             | 18                                  | 17.0                              | 17                                | 16.2                             |
| Dominyun          | 17                                 | 16.5                             | 17                                  | 17.1                              | 17                                | 16.9                             |
| Rock Solid Jess    | 18                                 | 18.0                             | 16                                  | 16.5                              | 17                                | 17.8                             |
| Tres Seis          | 20                                 | 20.1                             | 22                                  | 22.1                              | 19                                | 19.9                             |
| One Famous Eagle   | 21                                 | 21.8                             | 24                                  | 24.5                              | 21                                | 20.9                             |
| Chicks Regard      | 22                                 | 24.8                             | 22                                  | 26.0                              | 22                                | 25.4                             |
| Oak Tree Special   | 22                                 | 23.6                             | 29                                  | 28.9                              | 24                                | 24.4                             |
| Pool               | 22                                 | 21.9                             | 25                                  | 25.3                              | 22                                | 22.1                             |
| Capo de Capi       | 25                                 | 24.6                             | 23                                  | 22.9                              | 27                                | 26.6                             |
| American Runaway   | 26                                 | 25.3                             | 29                                  | 29.3                              | 24                                | 24.9                             |
| Mighty B Valiant   | 27                                 | 26.1                             | 25                                  | 25.1                              | 26                                | 26.3                             |
| Tempting Dash      | 27                                 | 24.0                             | 27                                  | 24.2                              | 26                                | 23.5                             |
| Tr Dasher          | 27                                 | 28.4                             | 27                                  | 27.1                              | 27                                | 27.6                             |
| Heza Fast Dash     | 29                                 | 30.1                             | 38                                  | 38.7                              | 30                                | 30.5                             |
| Big Daddy Cartel   | 31                                 | 30.7                             | 32                                  | 32.5                              | 30                                | 29.9                             |

The sires of Table 2 were chosen because they were the top ranked thirty sires by median from the 350 yard data set model.

| Groups               | Variance | Standard deviation |
|----------------------|----------|--------------------|
| Horse                | 0.305    | 0.552              |
| Trainer              | 0.064    | 0.252              |
| Sire                 | 0.083    | 0.288              |
| Jockey               | 0.022    | 0.148              |
| Track ID (Track)     | 0.157    | 0.396              |
| Residual             | 0.362    | 0.602              |

The column labeled “Standard deviation” lists the square roots of the variances (i.e., it does not represent an uncertainty estimate of the variance column) [10].

Goodness of fit

To characterize the goodness of fit of the mixed models, a quantity based on $R^2$ for simple linear models was utilized. For mixed models, there are a number of $R^2$ analogs, which serve a variety of different purposes [4]. In our case, we desired to quantify the closeness between the actual values of the response and the mixed model fitted values, with the random-effects estimates included in the calculation of those fitted values. We utilized a straightforward method contributed by Jarrett Byrnes in the r-sig-mixed-models mailing list; in this method, simple linear regression is performed between the fitted values and the actual values of the response, and the $R^2$ value obtained from this simple regression is used as a goodness of fit measure of the mixed model [4]. Among the models fitted after the initial variable selection process, these $R^2$ values range from 0.600 to 0.792.

We now present the sire ranking results. Recall that there are 9 datasets comprised of three variations on race distance: 1) 350 yard races, 2) 300 yard races, and 3) all distances combined; and three variations on sampling: 1) minimum of 10 offspring, 2) minimum of 3 offspring and 10 racing outcomes.
10 racing outcomes, and 3) all offspring. As mentioned earlier, seventeen different versions of the models were run on each dataset. As we present the data, the medians and means in the tables represent the averages over these seventeen versions.

First we examine the impact of the three methods of sampling on the 350 yard dataset. The results are shown in Table 2. The sires of Table 2 were chosen because they were the top ranked thirty sires by median from the 350 yard data set model. We note the similarity of the results for the three methods of sampling. Similar results were seen for the 300 yards and all distances datasets. The reason for this result is that a relatively small number of sires account for the vast majority of racing outcomes. Therefore as we proceed, we will only discuss the minimum of 10 offspring results.

Next we present the sire rankings for the minimum of 10 offspring datasets at 300 yards, 350 yards, and all distances. The results are presented in Table 1 and compared to rankings by total offspring earnings, average earnings, average z-score, and number of winners. We note the significant difference in sire rankings between our modeling results and rankings by total offspring earnings, average earnings, average z-score, and number of winners. From Table 1 we note that the very small difference between the median and mean values for the model results indicate a very close correspondence among the model outcomes.

### Discussion

The data subsetting by sire (various conditions on offspring numbers) produced a small effect on the variance values of the models’ random effects. This was not examined in depth, though it has implications for the generalizability of those variances values when attempting to draw conclusions about their degree of contribution to race performance. This subsetting may affect the validity of the model assumptions as well.

It is possible that structural issues in the data negatively impacted the sire rankings and conclusions. It is possible that the variables were not sufficiently crossed; an example problematic scenario would be a situation in which the data can be partitioned by track into two parts for which there is a set of horses, jockeys, and trainers that are associated with the tracks that make up the first part, a different set of horses, jockeys, and trainers that are associated with the tracks that make up the second part, and no horses, jockeys, or trainers that are associated with both sets of tracks. The final models all included horse and track. Some models included jockey and some did not. No significant ranking difference was seen between models with and without jockey.

We also note that different ranking methods have different goals. If the goal is to maximize total earnings, one should base breeding decisions on earnings rankings. If the breeding goal is to produce winners, one should base breeding decisions on rankings derived from the number of winners. Our goal was to derive a sire ranking system based on speed z-scores adjusted for various factors: wind, track, jockey, etc. We have accomplished that goal. From Table 1 we note the considerable difference between our model rankings for all distances and the more traditional rankings. For One Sweet Jess our all distances ranking is 13th, but it is 67th for average z-score. For Kiddy Up our all distances ranking is 3rd, but it is 61st for total winners. For Dominyun our all distances ranking is 4th, but it is 75th for average earnings. For Mighty B Valiant our all distances ranking is 7th, but it is 105th for total earnings. The significant difference in these rankings tells us that the factors we have adjusted for have a significant impact on the sire rankings. These differences support the validity of using a model that compensates for various factors: wind, track, jockey, etc. and supports the view that our modeling is independent of the standard ranking approaches. The benefit of the sire ranking that comes from this analysis is that it is controlled for track, jockey, trainer, weather, and several other variables that can impact speed. Sires are typically valued for high rankings for offspring earnings and winners. Yet a sire with a low stud fee may still produce offspring with a high ranking using our z-score model. The offspring of this bargain sire have the potential to produce fast offspring that could pay a dividend on a relatively low cost investment. The model sire ranking approach described in this paper is clearly bringing a new approach to the field of sire rankings.

### Acknowledgment

We would like to thank Dr. Matthew Jones for lending his resources.

### References

1. Bates, D.M. 2010. Lme4: Mixed-effects Modeling with R. http://lme4.r-forge.r-project.org/lMMwR/lrgprt.pdf [accessed on January 1, 2018].
2. Bates, D., Maechler, M., Bolker, B. and Walker, S. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67: 1–48 [CrossRef].
3. Bates, D., Maechler, M., Bolker, B. and Walker, S. 2017. Package lme4. https://cran.r-project.org/web/packages/lme4/lme4.pdf [accessed on January 1, 2018].
4. Bolker, B. 2017. GLMM FAQ. http://bbolker.github.io/mixedmodels-misc/glmmFAQ.html#how-do-i-compute-a-coefficient-of-determination-r2-or-an-analogue-for-glmms [accessed on August 7, 2017].
5. Buttram, S.T. 1987. Genetics of Racing Performance in...
the American Quarter Horse. J. Anim. Sci. 66: 2800–2807. [CrossRef]

6. Corrêa, M.I.M. and da Mota, M.D.S. 2007. Genetic evaluation of performance traits in Brazilian Quarter Horse. J. Appl. Genet. 48: 145–151. [Medline] [CrossRef]

7. Equibase. 2017. Equibase: North American records. http://www.equibase.com/about/northamericanrecords.cfm [accessed on July 8, 2017].

8. Faraway, J.J. 2016. Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. 2nd ed. CRC Press, Boca Raton.

9. Galili, T. 2017. Installr: using R to install stuff (such as: R, ‘Rtools’, ‘RStudio’, ‘Git’, and more). https://CRAN.R-project.org/package=installr [accessed on July 8, 2017].

10. R Core Team 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/ [accessed on July 8, 2017].

11. Villela, L.C.V., Mota, M.D.S. and Oliveira, H.N. 2002. Genetic parameters of racing performance traits of Quarter horses in Brazil. J. Anim. Breed. Genet. 119: 229–234. [CrossRef]

12. Walker, A. 2017. Openxlsx: Read, Write and Edit Xlsx Files. https://CRAN.R-project.org/package=openxlsx [accessed on July 8, 2017].

13. Wickham, H. 2009. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York. http://ggplot2.org [accessed on July 8, 2017].