

Estimation of Annual Trends and Environmental Effects on the Racing Records of Jeju Horses

Jongan Lee^{1†*}, Soo Hyun Lee^{2†}, Jae-Gu Lee², Nam-Young Kim³, Jae-Young Choi¹, Sang-Min Shin¹, Jung-Woo Choi⁴, In-Cheol Cho¹ and Byoung-Chul Yang¹

¹Subtropical Livestock Research Institute, National Institute of Animal Science, RDA, Jeju 63242, Korea

²Division of Animal Breeding and Genetics, National Institute of Animal Science, RDA, Cheonan 331-801, Korea

³National Institute of Animal Science, RDA, Wanju 565-851, Korea

⁴College of Animal Life Sciences, Kangwon National University, Chuncheon 24341, Korea

Received August 18, 2021 / Revised September 17, 2021 / Accepted September 23, 2021

This study was conducted to estimate annual trends and the environmental effects in the racing records of Jeju horses. The Korean Racing Authority (KRA) collected 48,645 observations for 2,167 Jeju horses from 2002 to 2019. Racing records were preprocessed to eliminate errors that occur during the data collection. Racing times were adjusted for comparison between race distances. A stepwise Akaike information criterion (AIC) variable selection method was applied to select appropriate environment variables affecting racing records. The annual improvement of the race time was -0.242 seconds. The model with the lowest AIC value was established when variables were selected in the following order: year, budam classification, jockey ranking, trainer ranking, track condition, weather, age, and gender. The most suitable model was constructed when the jockey ranking and age variables were considered as random effects. Our findings have potential for application as basic data when building models for evaluating genetic abilities of Jeju horses.

Key words : Annual trend, environmental effect, Jeju horse, racing record, stepwise variable selection

서 론

국제경마연맹 보고에 따르면 2018년 기준 세계 경마 시장의 매출 규모는 147조에 이르며 경마와 관련된 산업을 고려하면 360조원으로 추산되고 있다. 국내의 경우 2020년 기준 경마 산업 규모는 3조원대로 높은 부가가치를 창출하는 산업으로 인식되고 있으며 국민의 여가와 정부 세수 확보에 큰 기여를 하고 있다. 국내 경마는 서울, 부산, 제주를 포함한 세 지역에서 개최되고 있으며 제주경마공원에서는 제주마와 한라마 두 품종이 경주에 참여하고 있다. 제주마는 제주도에 토착한 독립적인 품종으로 1980년대 들어 개체수가 1천두 이하로 감소하자 개체수 확대를 위한 정책차원에서 경마용으로 보급하기 시작했다. 경마 초기에는 경주용 제주마의 수가 절대적으로 부족한 상황이었기 때문에 더러브렛과 교잡하여 경마용 한라마를 대량으로 생산했다. 제주마를 경주마로 보급하는 시점에서 한라마가 도입되었으며 국민 소득 증가로 경마에 관심이

높아지면서 제주마 개체수도 같이 증가하기 시작했다. 한라마는 제주마보다 체형이 크고 경주능력이 우월한 특성을 가지고 있어 경주마로 활용 가치는 제주마보다 뛰어나다. 하지만 한라마 경마 허용이 제주마 육성이라는 목표와 부합하지 않고 최근 제주마 개체수 증가로 인해 제주마 단독으로 경마가 가능한 상황이다. 제주마는 한라마에 비해 효용가치는 떨어지지만 개체수가 증가함에 따라 경주마로서 개량 가능한 수준이며 활용이 증가할 것으로 예측된다.

제주경마공원에서는 제주마 경주 자원 확산과 혈통 보존을 위해 2023년부터 한라마를 제외한 제주마 경마가 전면 시행될 예정이다. 2021년 8월 현재 제주마 사육농가는 734곳으로 총 6,886두가 통합관리시스템에 등록이 되어있으며 미경마 107두를 포함한 총 662두가 한국 마사회 제주경주마로 관리되고 있다. 이처럼 제주마 사육 규모가 작고 경주마로 활용되는 두수가 약 10% 내외인 점을 고려하면 2023년 제주마 경주 전면 시행을 앞두고 경주마로서 활용도를 높이고 개량방향 설정을 위한 유전능력 평가모형 구축이 필요한 상황이다. 국내 더러브렛 경주마의 속도형질에 대한 추세 분석과 체중 변화가 주파기록에 미치는 효과 규명 연구가 수행되었으며 과천과 부산 경남 경마공원 자료를 이용해 주파기록에 대한 유전력과 반복력 추정 및 소득상금간의 상관성 분석 등을 통해 더러브렛에 대한 유전능력 평가가 시도되었다[6, 7, 20]. 최근, 한국마사회는 유전체 기반 선발프로그램인 케이닉스(K-Nicks)를 활용해 우수 유전능력을 보유한 개체를 선발하고 계획교배 프로그램

† Authors contributed equally.

*Corresponding author

Tel : +82-64-754-5712, Fax : +82-64-754-5713

E-mail : amasss@korea.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

을 활용해 우량 경주마를 생산하고 있다. 국외의 경우 더러브렛의 경주기록에 대한 고정효과를 고려한 모형이 수립되었으며 단거리 경주마로 활용되는 퀴터호스는 임의효과와 영구적 환경효과를 고려한 변수 설정을 통해 유전력 및 반복력 추정 연구가 수행되었다[2, 26]. 하지만 제주마의 경우 경주마로서 선발 기준 및 개량방향 설정이 미흡한 상황이다. 이를 해결하기 위해 제주마 유전능력평가 모형 설정이 필요하며 모형 구축을 위한 선행연구로 경주능력에 대한 환경효과 및 경주형질 연간 개량량 분석을 통해 개량을 위한 초석을 마련해야 한다.

본 연구에서는 한국마사회에서 수집한 제주마 경주기록을 이용해 데이터 전처리 및 보정 과정을 거치고 주과기록에 대한 연도별 추세와 환경변수 효과를 분석했다. 효과 분석을 통해 주과기록에 대한 통계적 유의성이 검증된 환경변수들에 대해 단계적 변수 선발법을 적용해 주과기록에 영향하는 요인을 가장 잘 설명할 수 있는 모형을 제시하였으며 임의효과를 특정 변수들에 적용하여 모형 적합도를 추정하고 설명력을 높였다. 제주마 경주기록에 대한 환경효과 분석 결과는 제주마의 유전능력 평가 시 개체모형(animal model) 설정 등에 기초자료로 활용이 가능할 것으로 사료된다.

재료 및 방법

제주마 경주 환경 및 선발 기준

제주경마공원 경주로는 직선 거리 493.7 m와 곡선 거리 306.3 m로 타원형 구조이며 주로 폭은 20 m로 제주마 10두가 경주에 참여한다. 경주로는 400 m를 제외한 나머지 경주거리에 곡선 구간이 포함되어 있으며 경주로 종단 경사도는 약 -2 m~+8 m에 이른다. 경마는 매년 1월부터 12월까지 계절이나 날씨의 상관없이 진행되며 경주로 단면은 상단부터 모래, 화강풍화토, 보조기층으로 구성되어 있다. 제주마 등록 관리 규정에 따르면 선대의 혈통 정보 유무를 기준으로 기초등록과 혈통등록으로 구분되며 혈통등록된 부모에서 생산된 36개월령 이상의 우수한 자마는 고등등록마로 등록된다. 제주마로 선발되기 위해서는 혈통등록 이상의 등급이 요구되며 137 cm 이하 체고 규정을 사용하는 한라마와 다르게 체형에 대한 제한은 없다.

공시재료 및 환경변수 범주

한국마사회 경마정보(<https://race.kra.co.kr/jejuMain.do>)에서 2002년 7월부터 2019년 10월까지 총 17년 4개월간의 제주마 경주기록 데이터가 수집되었다. 경마일수는 1,379일, 경마횟수는 4,673회로 총 48,645개의 관측치가 분석에 사용되었다. 연도별 관측치와 말 무게 정보는 Table 1과 같으며 분석에 사용된 총 2,167두의 성별은 수말 435두, 암말 996두, 거세 736두로 분류된다. 경주기록에 대한 환경효과 분석에 사용된 변수는 연도, 날씨, 부담구분, 주로상태, 마령, 수득상금 효과가

Table 1. Characteristics of Jeju horse records by race year

Year	No. of records	No. of horses	BW* (kg)	Race distance (m)					
				400	800	900	1,000	1,110	1,200
2002	48	12	256.1	0	48	0	0	0	0
2003	100	19	261.0	0	91	9	0	0	0
2004	463	46	266.2	0	368	95	0	0	0
2005	779	66	268.0	0	359	410	10	0	0
2006	1,186	103	264.7	101	422	612	51	0	0
2007	1,757	157	265.1	241	640	755	121	0	0
2008	2,188	212	264.5	288	775	859	266	0	0
2009	2,626	283	266.1	366	949	983	328	0	0
2010	2,503	304	271.5	269	949	726	559	0	0
2011	2,645	327	275.5	327	940	716	516	146	0
2012	2,896	352	271.3	285	1,234	603	436	251	87
2013	3,094	386	272.8	293	1,277	681	379	330	134
2014	3,429	436	273.7	478	1,442	522	594	313	80
2015	3,942	485	274.9	853	1,615	611	540	262	61
2016	4,802	541	272.7	767	2,326	929	504	220	56
2017	5,436	606	274.7	737	2,648	1,213	591	188	59
2018	6,025	656	276.1	519	2,001	1,536	981	704	284
2019	4,726	671	278.3	0	1,772	842	1,172	670	270
Total	48,645	2,167 [†]	-	5,524	19,856	12,102	7,048	3,084	1,031
(%)	(100%)			(11.3%)	(40.8%)	(24.8%)	(14.4%)	(6.3%)	(2.1%)

* , BW indicates the average of body weight of horses.

[†] , Total number of horses participated in race from 2002 to 2019 years, not the column sum.

반영된 기수와 조교사 순위, 성별이 있다. 연도는 경주가 시행된 2002년부터 2019년까지 해당 연도가 변수 값으로 사용되었다. 날씨 정보는 강풍, 눈, 맑음, 비, 안개, 흐림으로 구분되며 경주 시행일 당시 날씨가 변수 값으로 활용되었다. 부담구분은 3가지 부담중량체계에 따라 마령, 핸디캡, 별정경주로 분류되며 별정경주는 수득 규모, 승급 점수 등을 반영해 6단계로 세분화 된다. 부담구분 변수는 총 8단계(마령, 핸디캡, 별정A, 별정B, 별정C, 별정 I, 별정 II, 별정 III)로 구성되며 각 단계가 변수 값으로 사용되었다. 주로상태는 주로에 포함된 수분함량에 따라 건조(1~5%), 양호(6~9%), 다습(10~14%), 포화(15~19%), 불량(20%~)으로 나뉘며 함수율에 따른 주로상태가 변수 값으로 이용되었다. 주로상태에 대한 연도별 관측치 정보는 Table 2에 제시했다. 마령은 2세부터 27세까지 분포하고 있으며 개체별 마령 정보가 연속형 변수 값으로 적용되었다. 성별은 암, 수, 거세로 구분되며 성별 정보가 이산형 변수로 적용되었다.

자료 전처리와 환경변수 보정

주파기록이 없는 394개 관측치와 거리별 전체 기록에서 99% 이상에 해당하는 주파기록은 오기나 비정상적인 경주로 판단하여 분석에서 제외했다. 연도별 주파기록 특성을 비교하기 위해 거리별 주파기록은 동일 경주거리에 대한 기록으로 변환하였으며 이를 위해 관측치가 가장 많은 800 m를 기준으로 주파기록을 보정했다[16]. 거리별 평균주파기록은 선형회귀모형을 통해 얻어진 아래 식에 따라 계산했다[19].

$$\text{평균주파기록(sec)} = 0.45464 + 0.08623 \times \text{경주거리(m)}$$

거리별 평균주파기록과 800 m 평균주파기록 차이를 해당 거리의 개별 관측치의 주파기록에 가감을 통해 반영했다. 기수와 조교사에 따른 수득상금 효과를 반영하기 위해 개별 수득상금은 해당연도의 수득상금 평균으로 나누어 연도별 수득상금 규모에 따른 차이를 보정해주었다[16]. 기수별, 조교사별 보정된 수득상금의 총합을 바탕으로 새로운 순위를 정하고 이를 변수값으로 활용했다.

주파기록에 대한 변수 선택과 모형 수립

보정된 주파기록에 대한 후보 환경변수의 범주 수준 간 차이를 확인하기 위해 일원분산분석(one-way ANOVA)을 수행했다. 유의성이 확인된 환경변수들을 대상으로 선형회귀모형을 이용한 Akaike information criterion (AIC) 값을 아래 식에 따라 계산하여 모형의 적합도를 평가했다[1].

$$AIC = -2 \ln(L) + 2K$$

L: maximum value of the likelihood function for the model

K: number of estimated parameters in the model

단계적 변수 선별법(stepwise variables selection)을 적용하여 다양한 변수 조합 중 AIC 값이 가장 낮은 모형을 선정했다. AIC 값 최소화 모형 선택은 우도(likelihood)를 가장 크게 하는 동시에 변수 개수가 가장 적은 최적의 모델을 의미한다[27].

Table 2. Track condition and moisture content by race year

Year	No. of records	Track condition (moisture content)				
		Heavy (1~5%)	Good (6~9%)	Soft (10~14%)	Yielding (15~19%)	Sloppy (20%~)
2002	48	0	16	0	17	15
2003	100	0	32	26	17	25
2004	463	0	244	45	55	119
2005	779	0	418	105	125	131
2006	1,186	0	606	206	125	249
2007	1,757	0	1,017	274	150	316
2008	2,188	0	1,252	320	137	479
2009	2,626	808	1,074	556	0	188
2010	2,503	995	842	518	148	0
2011	2,645	1,034	526	438	408	239
2012	2,896	901	767	730	315	183
2013	3,094	1,151	863	623	181	276
2014	3,429	935	1,164	580	411	339
2015	3,942	986	1,439	948	314	255
2016	4,802	1,341	1,595	1,019	442	405
2017	5,436	66	3,993	966	180	231
2018	6,025	0	3,139	1,847	442	597
2019	4,726	1,748	1,234	982	243	519
Total	48,645	9,965	20,221	10,183	3,710	4,566
(%)	(100%)	(20.4%)	(41.5%)	(20.9%)	(7.6%)	(9.3%)

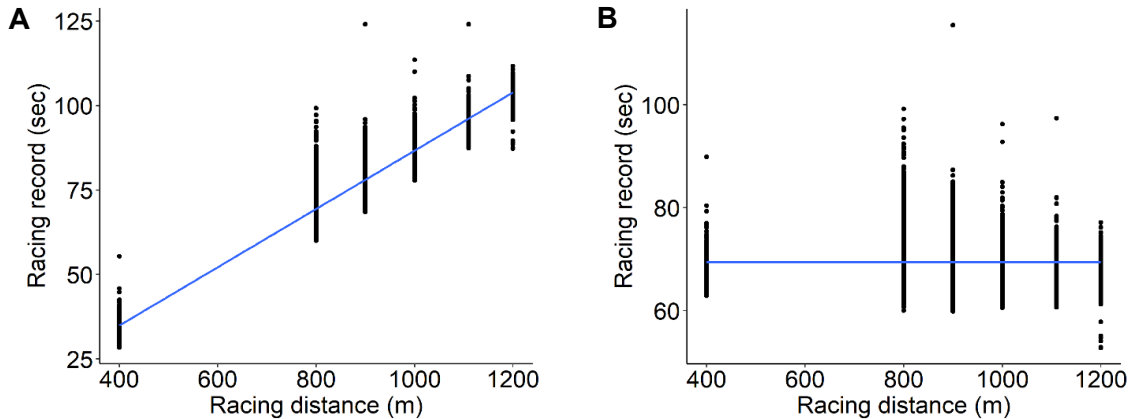


Fig. 1. Distribution of (A) original racing records and (B) adjusted racing records in different distances. Blue line indicates the linear regression line.

단계적 변수 선발법은 전진선택법과 후진제거법을 결합한 형태로 고정효과 변수의 추가 및 삭제 과정을 반복적으로 수행하고 최종적으로 AIC 값이 최소가 되는 모형을 선택한다. 임의효과는 별도의 통제된 환경이나 실험을 수행하는 연구자의 의도에 따라 얻어지는 변수가 아니고 환경에 따라 변하는 특성을 가지고 있다. 특정 상황에서 발생하는 오차로부터 모형에 대한 설명력을 높이기 위해 임의효과를 고려한 모델 선택이 필요하다[21]. 마령, 기승자와 조교사 효과 같이 변동성이 있는 변수를 임의효과로 반영해 모형 설명력을 높였다. 이를 위해 R 프로그램의 lme4 패키지[3]를 사용해 고정효과와 임의효과를 동시에 반영할 수 있는 선형혼합모형을 수립했다.

결과 및 고찰

주파기록에 대한 보정 및 연도별 추세

제주마 경주기록은 900 m 이하 관측치가 전체 기록의 77.0%로 비교적 단거리 위주의 경마가 시행되고 있었으며 1,000 m 이상 중장거리는 23%로 경주거리가 증가할수록 경마 시행 횟수는 감소했다(Table 1). 이는 다양한 품종별 경주거리가 증가할수록 경기 횟수가 감소하는 추세와 일치했으며 승군제의 영향으로 장거리로 갈수록 경주편성수가 감소하기 때문이다[5, 12]. 제주마 주파기록은 단거리에서 중장거리로 갈수록 증가하므로 모든 관측치를 이용한 모형 수립을 위해 관측치가 40.8%로 가장 많은 비중을 차지하는 800 m를 기준 거리로 설정했다[16]. 경주거리별 평균기록과 기준 평균기록 사이의 차이를 해당 경주거리의 개별 관측치에 반영함으로써 주파기록 사이의 비교가 가능하도록 보정했다(Fig. 1). 연도별 평균 주파기록은 2000년대 초반에는 증가하였으나 2000년대 후반부터 2010년 중반까지 단축되는 경향을 보였다(Table 3). 이는 우수 씨수말을 활용한 교배설정으로 경주능력 개량이 이루어졌다고 생각된다. 하지만, 2010년 후반에 들어서면서 주파기록이 늘어나는 경향을 보였는데 이는 체계적인 개량시스템

구축이 미흡하기 때문에 나타난 결과라고 사료된다. 선형회귀 모형을 통해 얻어진 주파기록 변화 추이는 매년 0.242초 단축되었으며 결정계수는 0.6601로 중도의 집합도를 보였다(Fig. 2). 국내 더러브렛 경주마의 1990년도부터 2008년까지 경주기록 개량량은 -0.166초로 산출되었다[16]. 국내에서 시행된 두 품종에 대한 경주기록 추세를 비교해 볼 때 제주마를 활용한 경주형질 개량은 충분히 가능한 집단이라고 생각된다. 500두 내외의 터키 아랍종의 15년간 주파기록은 매년 0.284초 단축되었으며 소규모 집단인 제주마의 주파기록 추세와 비슷한 양상을 보였다[13]. 더러브렛의 경우 주파기록에 대한 기수와 조교사 영향은 경주거리에 상관없이 통계적 유의성이 확인되었다[25]. 제주마 경주기록에 기수와 조교사 효과를 반영하기 위해 평균수득상급에 따라 순위를 정하고 이를 기수와 조교사의 변수 값으로 변환했다(Fig. 3).

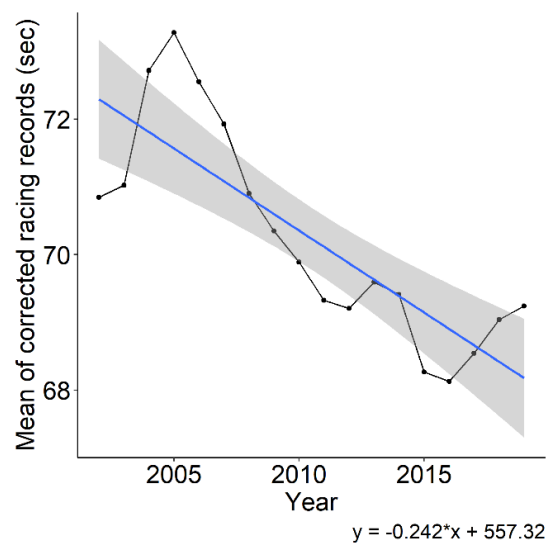


Fig. 2. Annual trends on adjusted racing records. Blue line indicates the linear regression line.

$$y = -0.242 \cdot x + 557.32$$

Table 3. Statistics of adjusted racing records by race year

Year	Race time (seconds)					
	Minimum	1st quartile*	Median	Mean ± SD†	3rd quartile*	Maximum
2002	64.30	68.48	70.30	70.85±3.71	73.28	84.20
2003	64.20	69.00	70.00	71.03±3.37	72.75	81.20
2004	66.70	70.64	72.70	72.72±2.89	74.40	84.30
2005	62.68	71.20	73.08	73.28±3.08	74.98	87.28
2006	65.29	70.78	72.48	72.55±2.96	74.38	87.50
2007	64.59	69.85	71.90	71.93±3.01	73.68	115.48
2008	63.68	68.67	70.80	70.90±2.99	72.90	91.80
2009	61.88	67.99	70.38	70.35±3.24	72.45	92.75
2010	59.88	67.48	70.29	69.89±3.54	72.08	95.10
2011	60.78	67.19	69.28	69.32±3.12	71.30	86.60
2012	52.81	66.88	69.20	69.21±3.44	71.10	95.50
2013	61.20	67.37	69.40	69.60±3.33	71.40	99.20
2014	61.35	67.29	69.18	69.41±3.13	71.28	97.37
2015	60.67	66.49	67.89	68.27±2.87	69.90	97.20
2016	60.08	66.20	67.80	68.13±3.00	69.90	90.20
2017	60.00	66.69	68.30	68.54±2.87	70.38	90.50
2018	60.58	67.55	69.00	69.04±2.46	70.40	82.98
2019	60.55	67.95	69.35	69.24±2.38	70.67	87.28

* , 1st quartile indicates the median of the lower half of the data and 3rd quartile indicate the median of the upper half of the data.

† , Standard Deviation (SD)

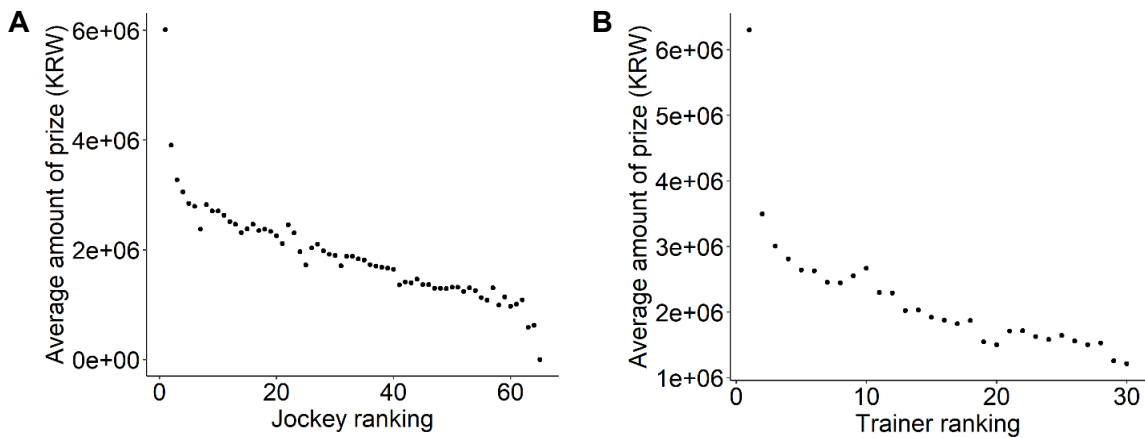


Fig. 3. Distribution of average earning prize by (A) jockey and (B) trainer ranking.

환경변수 효과 및 모형 수립

후보 환경변수의 특성에 따라 연도, 날씨, 부담의 구분, 주로 상태, 성별은 이산형 변수로 마령, 기수, 조교사 순위는 연속형 변수로 구분했다. 주파기록에 대한 환경변수들의 범주 효과 분석을 위해 일원분산분석을 사용하여 변수 범주간 통계적 유의성을 확인했다(Table 4). 분석 결과 8가지 환경변수 모두에서 보정된 주파기록에 대한 통계적 유의성($p < 0.001$)이 검증되었으며 이들 변수에 대해 단계적 변수 선별법을 활용하여 경주기록 환경효과 분석을 위한 모형을 수립했다. 변수 선별에 대한 모형 적합도를 측정하기 위해 AIC 값을 추정하였고 AIC 값이 낮을수록 모형의 적합도가 높은 것을 의미한다. 환

경변수를 고려하지 않은 상태에서의 추정된 AIC 값에서 8가지 변수 중 연도 변수가 추가될 때 가장 낮은 AIC 값이 산출되었다(Table 5). 연도별 제주마 개량 효과가 경주기록에 반영되었음을 의미하고 경주마의 평균 체중은 2002년부터 2019년까지 매년 꾸준히 증가했다(Table 1). 제주마의 체형 기능적 특성은 경주능력과 상관성이 있다는 결과를 미루어 볼 때 제주마 체형 변화는 경주기록에 영향을 준 것으로 생각된다[17]. 또한 아랍종의 경우 경마 출주연도가 경주 순위에는 영향이 없지만 경주기록과 수득상금에 영향을 미치는 것으로 보고되었다[11]. 2단계에서 나머지 7가지 변수 중 부담구분 변수가 추가될 때 AIC 값이 최소로 산출되었으며 부담중량체계에 따

Table 4. Effects of environment variables on adjusted racing records

Variables	Degrees of freedom	Sum of squares	Mean of squares	F-value	P-value
Year	17	66,004	3,883	510.57	< 2.00E-16
Weather	5	1,077	215	28.32	< 2.00E-16
Track condition	4	1,411	353	46.37	< 2.00E-16
Budam classification	7	41,342	5,906	776.66	< 2.00E-16
Gender	2	476	238	31.29	2.63E-14
Age	1	3,039	3,039	399.65	< 2.00E-16
Jockey ranking	1	8,028	8,028	1,055.65	< 2.00E-16
Trainer ranking	1	4,384	4,384	576.51	< 2.00E-16
Residuals	48,606	369,620	8		

Table 5. Evaluation of the model fit on adjusted racing records

Step	Variables	Sum of squares	Residual sum of squares	AIC
0	none		495,381	112,896
1	year	42,848	452,533	108,498
2	year + bc ¹⁾	33,952	418,581	104,718
3	year + bc + jr ²⁾	8,999.2	409,582	103,662
4	year + bc + jr + tr ³⁾	4,831.8	404,750	103,087
5	year + bc + jr + tr + tc ⁴⁾	3,541.7	401,208	102,668
6	year + bc + jr + tr + tc + weather	2,211.14	398,997	102,409
7	year + bc + jr + tr + tc + weather + age	1,061.1	397,936	102,281
8	year + bc + jr + tr + tc + weather + age + gender	497.94	397,438	102,224

¹⁾bc, budam classification; ²⁾jr, jockey ranking; ³⁾tr, trainer ranking; ⁴⁾tc, track condition

른 패널티 부여가 주파기록에 영향 미치는 변수임을 의미한다. 일본 더러브렛 경주마에서 부담중량에 대한 경주기록 영향력을 평가하기 위해 분산 분석이 수행되었으며 1,200 m에서 2,000 m에 이르기까지 부담중량이 경주시간에 영향을 주는 것으로 확인되었다[18]. 3단계와 4단계에서 기수와 조교사 순위 변수가 선택되었고 경주마의 선천적 경주능력 이외에 기승자의 능력이나 조교사의 훈련 절차와 방법이 주파기록에 중요한 영향을 미칠 수 있다. 더러브렛의 경우 기수 1·2위 비율, 기수 1년 1위 등 과거 우승 경력이 주파기록에 영향을 미치는 것으로 나타났다[9]. 더러브렛 50두 이상 조련 경력이 있는 조교사에게 훈련 받은 경주마는 5두 이하 경력을 가진 조교사에게 훈련 받은 경주마보다 수득상금이 두배 많은 것으로 보고되었다[23]. 이는 조교사의 능력과 경험에 따라 경주마의 경주능력이 향상됨을 의미한다. 5단계에서 선택된 주로상태 변수는 주로 함수율 변화가 주파기록에 영향을 끼칠 수 있으며 더러브렛의 경우 주로상태에 따라 주파기록은 거리별 평균 0.37초에서 2.28초까지 단축됨을 보고했다[8]. 경주를 위한 이상적인 주로상태는 양호 또는 다습으로 경주하기에 완벽한 토양 표면을 제공하고 자연스러운 다리 동작에 적합하다고 알려져 있다[22]. 6단계에서 선정된 날씨의 경우 주로상태와 상관성이 높을 것으로 판단되는 변수이다. 모형에 유사한 특성을 가지는 변수들이 사용되면 향후 유전능력평가를 하는데

있어서 다중공선성(multicollinearity) 문제가 발생할 여지가 있고 이로 인해 부적합한 회귀결과를 도출시킬 수 있다. 날씨와 주로상태 범주간 상관성 분석은 Cramer's V [10]로 수행하였으며 상관계수 0.3251로 다중공선성 문제를 일으키지 않는 것으로 확인되었다. 추후 다른 변수를 추가하여 모형 설정 시 특성이 유사하다고 판단되는 변수들은 상관관계를 미리 확인하고 보완함으로써 다중공선성 문제를 해결할 수 있을 것으로 사료된다. 7단계에서 선택된 마령 변수의 경우 제주마 고령 출주 빈도가 타 품종에 비해 상대적으로 높기 때문에 연령 효과에 따른 모형 설명력이 떨어지는 결과를 얻었다. 6세 이상 제주마 출주빈도는 18.4%로 한라마와 더러브렛의 1.1%와 7.1%보다 높은 경향을 보였으며 더러브렛의 경주능력은 2세에서 4.5세까지 상승하고 4.5세 이후부터 하락하는 특성을 보였다[12, 15]. 이외에 성별 변수 추가는 이전 단계 AIC 값 변화보다 감소폭이 가장 적었으며 주파기록을 설명하는데 영향이 미비한 것으로 나타났다. 하지만 더러브렛과 스탠다드브레드에서는 마령이 5세, 성별이 수말인 경우 수득상금이 증가하는 것으로 보고되었다[4]. 이러한 상반된 결과는 품종별 차이를 고려한 변수 선택과 모형 수립이 필요하다는 것을 반증한다고 볼 수 있다.

단계적 변수 선별법을 통해 수립된 모형에 임의효과로 고려할 수 있는 마령, 기수와 조교사 순위 3가지 변수에 대해 임의

Table 6. Estimation of random effects on the established model

	Random effects	Fixed effects	AIC
Model A	jr ¹⁾	year, bc ²⁾ , tr ³⁾ , tc ⁴⁾ , weather, age, gender	236,964
Model B	jr, tr	year, bc, tc, weather, age, gender	236,875
Model C	tr	year, bc, tr, tc, weather, age, gender	236,845
Model D	age, tr	year, bc, jr, tc, weather, gender	236,253
Model E	age	year, bc, jr, tr, tc, weather, gender	236,239
Model F	age, jr, tr	year, bc, tc, weather, gender	236,164
Model G	age, jr	year, bc, tr, tc, weather, gender	236,140

¹⁾jr, jockey ranking; ²⁾bc, budam classification; ³⁾tr, trainer ranking; ⁴⁾tc, track condition

효과로 가정하고 이들 변수 조합에 따른 모형 적합도를 추정했다(Table 6). 일반적으로 수집되는 성별, 날씨 등과 같은 고정효과 외에 범주가 다양하고 연도별 빈도가 불규칙적인 기승자나 조교사 효과를 임의효과에 반영함으로써 모형에서 설명하지 못하는 잔차와 같은 나머지 값 혹은 오차로부터 모형에 대한 설명력을 높일 수 있다. 더러브렛의 경우 경주속도에 대한 기승자 변수를 임의효과로 반영했을 때 통계적 유의성이 확인되었다[24]. 3가지 임의효과를 조합한 7가지 모형 중 마령과 조교사 순위가 임의효과로 반영된 모형 G가 주파기록에 대한 설명력이 가장 높았다. 향후 유전평가 시 표현형(주파기록)에 대한 잔차분산(error variance)을 최소화하기 위해서는 환경변수의 수집형태에 따라 적절한 임의효과 설정이 필요할 것으로 사료된다. 경주마 개량을 위한 개체선발에 앞서 유전모수와 육종가 추정을 통해 유전능력평가가 수행되어야 한다. 제주마의 경주형질 유전능력평가를 위한 통계 모형 구축 시 다양한 환경변수들에 대해 고정효과로 포함시킬 것인지 아니면 임의효과로 포함시킬 것인가에 대한 결정이 필요하다. 본 연구의 결과를 기초자료로 활용하면 이에 대한 판단 및 모형에 대한 설명력을 높이고 최적의 모형을 수립할 수 있을 것이다. 예를 들어 기승자의 효과를 고정효과로 판단하고 모형 수립 시 추정된 유전모수나 육종가는 부정확할 가능성이 있고 이를 활용한 개체선발은 개량방향 설정에 오류가 생길 가능성이 있다.

현재 수집된 제주마 혈통정보는 세대 간격이 짧고 혈연 정보량이 충분하지 않아 혈연관계의 설명력이 부족한 상황이다. 제주마 혈통 정보의 유실 또는 정보의 오류가 상가적 유전분산량과 유전력 추정치를 왜곡하는 것으로 나타났다[14]. 따라서 제주마 혈통 정보에 대한 충분한 데이터 확보가 필요하고 경주능력이 우수한 말을 고등마로 지정 시 수정되는 등록번호를 지정 전 경주기록 데이터에 반영하는 등 동일개체가 다른 혈통정보를 가지게 되는 오류를 줄이는 노력이 필요하다고 사료된다. 이번 연구에서 수립된 주파기록에 영향을 미치는 환경효과 분석결과는 혈통이나 유전체 정보를 활용한 제주마 유전능력 평가에 초석이 될 것으로 기대한다.

감사의 글

본 연구는 농촌진흥청 연구사업(주관과제명: 제주마 경주능력에 대한 능력검정 연구, 주관과제 번호: PJ014281, 세부과제명: 제주마 경마 형질 유전능력 추정 기법 개발 연구, 세부과제 번호: PJ01428102)의 지원에 의해 이루어진 것임.

The Conflict of Interest Statement

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Akaike, H., Parzen, E., Tanabe, K. and Kitagawa, G., Selected papers of hirotugu akaike. 1998: Springer Science & Business Media.
2. Bakhtiari, J. and Kashan, N. 2009. Estimation of genetic parameters of racing performance in Iranian Thoroughbred horses. *Livest. Sci.* **120**, 151-157.
3. Bates, D., Mächler, M., Bolker, B. and Walker, S. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
4. Cheetham, J., Riordan, A., Mohammed, H., McIlwraith, C. and Fortier, L. 2010. Relationships between race earnings and horse age, sex, gait, track surface and number of race starts for Thoroughbred and Standardbred racehorses in North America. *Equine Vet. J.* **42**, 346-350.
5. Cho, B. W., Ha, T. Y., Cho, K. H., Kim, S. D., Lee, H. K., Kong, H. S. and Park, K. D. 2009. Estimation of genetic parameters for finished and furlong times in Thoroughbred racehorses. *Asian-Australas J. Anim. Sci.* **22**, 1609-1613.
6. Cho, K. H., Son, S. K., Cho, B. W., Kim, J. G., Kong, H. S., Lee, H. K. and Park, K. D. 2010. Genetic evaluation of Thoroughbred Racehorses using the integrated racing records collected from different racetracks. *J. Anim. Sci. Technol.* **52**, 97-102.
7. Cho, K. H., Son, S. K., Cho, B. W., Lee, H. K., Kong, H. S., Jeon, G. J. and Park, K. D. 2008. Effects of change of body weight on racing time in Thoroughbred racehorses. *J. Anim. Sci. Technol.* **50**, 741-746.

8. Cho, K. H., Son, S. K., Cho, B. W., Kim, J. G., Lee, H. K. and Park, K. D. 2008. Effects of the condition of racetrack on racing time in Thoroughbred racehorses. *J. Anim. Sci. Technol.* **50**, 309-314.
9. Choe, H., Hwang, N., Hwang, C. and Song, J. 2015. Analysis of horse races: prediction of winning horses in horse races using statistical models. *KJAS.* **28**, 1133-1146.
10. Cramér, H., *Mathematical Methods of Statistics (PMS-9)*, Volume 9. 2016: Princeton university press.
11. Ekiz, B. and Kocak, O. 2005. Phenotypic and genetic parameter estimates for racing traits of Arabian horses in Turkey. *J. Anim. Breed. Genet.* **122**, 349-356.
12. Gramm, M. and Marksteiner, R. 2010. The effect of age on thoroughbred racing performance. *J. Equine Sci.* **21**, 73-78.
13. Kaygisiz, A. 2011. Genetic trends for racing performance in Turk Arabian horses raised at Anatolian state farm. *Indian J. Anim. Sci.* **81**, 853.
14. Kim, N. Y., Lee, S. S. and Yang, Y. H. 2014. Influence of amount of pedigree information and parental misidentification of progeny on estimates of genetic parameters in Jeju Race Horses. *J. Anim. Reprod. Biotechnol.* **29**, 289-296.
15. Kong, H. S., Lee, H. K., Park, K. D. and Cho, B. W. 2011. A breed comparison on the finishing times of racehorses. *J. Anim. Sci. Technol.* **53**, 23-27.
16. Oh, S. Y., Park, J. E., Lee, J. R., Lee, J. W., Oh, H. S. and Kim, H. B. 2009. Estimation of genetic parameters and annual trends for racing times of Thoroughbred racehorses. *J. Anim. Sci. Technol.* **51**, 129-134.
17. Oh, W., Kim, B., Cho, H., Shin, T., Cho, S. and Cho, B. 2014. Analysis of associated race performance and functional characterization of conformation in Jeju Horse. *J. Agric. Life Sci.* **48**, 99-106.
18. Oki, H., Sasaki, Y. and Willham, R. 1994. Genetics of racing performance in the Japanese Thoroughbred horse: II. Environmental variation of racing time on turf and dirt tracks and the influence of sex, age, and weight carried on racing time. *J. Anim. Breed. Genet.* **111**, 128-137.
19. Park, J. E., Lee, J. R., Oh, S., Lee, J. W., Oh, H. S. and Kim, H. 2011. Principal components analysis applied to genetic evaluation of racing performance of Thoroughbred race horses in Korea. *Livest. Sci.* **135**, 293-299.
20. Park, K. D., Lee, H. K., Jeon, G. J., Kong, H. S., Cho, B. W., Cho, K. H. and Son, S. K. 2008. Trends on racing speed traits in thoroughbred racehorses. *J. Anim. Sci. Technol.* **50**, 747-752.
21. Peng, H. and Lu, Y. 2012. Model selection in linear mixed effect models. *J. Multivar. Anal.* **109**, 109-129.
22. Sheridan, M. and Sweeney, J. 2001. Weather and horse racing: Towards a more objective prediction of the going. *Weather* **56**, 48-55.
23. Sobczynska, M. 2007. The effect of selected factors on length of racing career in Thoroughbred racehorses in Poland. *Anim. Sci. Pap. Rep.* **25**, 131-141.
24. Sobczyńska, M. 2011. Environmental factors affecting the speed of Thoroughbred horses competing in Poland. *Anim. Sci. Pap. Rep.* **29**, 303-312.
25. Velie, B., Hamilton, N. and Wade, C. 2015. Heritability of racing performance in the Australian Thoroughbred racing population. *Anim. Genet.* **46**, 23-29.
26. Villela, L., Mota, M. and Oliveira, H. 2002. Genetic parameters of racing performance traits of Quarter horses in Brazil. *J. Anim. Breed. Genet.* **119**, 229-234.
27. Yamashita, T., Yamashita, K. and Kamimura, R. 2007. A stepwise AIC method for variable selection in linear regression. *Commun. Stat. Theory Methods* **36**, 2395-2403.

초록 : 제주마 주파기록에 대한 연도별 추세 및 환경효과 분석

이종안^{1*} · 이수현^{2*} · 이재구² · 김남영³ · 최재영¹ · 신상민¹ · 최정우⁴ · 조인철¹ · 양병철¹

(¹농촌진흥청 국립축산과학원 난지축산연구소, ²농촌진흥청 국립축산과학원 가축개량평가과, ³농촌진흥청 국립축산과학원, ⁴강원대학교 동물생명과학대학 동물자원과학과)

본 연구는 제주마의 연도별 주파기록 특성을 파악하고 경주기록에 대한 환경효과 분석을 위해 수행되었다. 한국마사회가 2002년부터 2019년까지 수집한 제주마 2,167두의 48,645개 관측치 정보가 분석에 사용되었다. 환경효과 분석을 위하여 관측치 빈도가 가장 높은 800 m 기준으로 주파기록은 보정되었다. 거리별 경주기록은 동일거리 기준으로 변환되었으며 연도별 비교가 가능했다. 그 결과 표현형에서 매년 주파기록이 0.242초($R^2=0.66$) 단축됨을 알 수 있었다. 보정된 주파기록에 대한 환경효과의 유의성 검증 위해 일원분산분석을 수행하였고 모형에서 설정한 모든 변수에서 고도의 유의성을 나타내었다($p<0.001$). 이들 변수에 대해 단계적 변수 선발법을 적용, AIC 값 산출을 통해 모형 적합도를 평가하였으며 연도, 부담구분, 기수 순위, 조교사 순위, 주로상태, 날씨, 마령, 성별 순으로 변수 선택 시 가장 낮은 AIC 값을 갖는 모형이 수립되었다. 최종적으로 주파기록에 영향을 미치는 환경효과 분석을 위하여 기수, 조교사 순위 및 마령 3가지 변수를 임의효과로 가정하였다. 그 결과 기수와 마령을 임의효과로 설정하였을 때 주파기록에 영향을 미치는 환경효과 분석에 가장 적합한 모형인 것으로 나타났다. 본 연구에서 도출된 결과는 제주마 유전능력 평가를 위한 모형설정 시 기초자료로 활용 가능할 것으로 사료된다.