Comparison of Parametric and Non-Parametric Population Modelling of Sport Performances

Stéphane Bermon
Health and Science Department,
Int. Ass. of Athletic Federations,
Monaco
stephane.bermon@iaaf.org

Asya Metelkina
Laboratoire I3S
CNRS,
Sophia Antipolis, France
metelkin@i3s.unice.fr

Maria João Rendas
Laboratoire I3S
CNRS,
Sophia Antipolis, France
rendas@i3s.unice.fr

***Abstract*—This work compares the performance of parametric mixed-effects models to a completely non-parametric (NP) approach for modelling life-long evolution of competition performance for athletes. The difficulty of the problem lies in the strongly unbalanced characteristics of the functional dataset. The prediction performance of the identified models is compared, revealing the advantages and limitations of the two approaches. As far as we know this is the first time NP modelling of athletic performance is attempted, our study confirming its appropriateness whenever sufficiently rich datasets are available.**

***Index Terms*—Functional data, longitudinal population models, mixed-effects models, Gaussian Processes, Hierarchical Bayesian Gauss-Wishart models, athletic performance.**

I. Introduction

Models of the temporal evolution of a given population are interesting in many different situations. Population models have traditionally resorted to parametric mixed-effects models [7] which have a long record of successful application in several domains, in particular in pharmacokinetic studies. These models parametrise the possible evolutions in time of a quantity of interest, distinct individuals of the population being described by different values of the parameters. The choice of the parametric family is usually dictated by existing knowledge about the expected patterns of variation, and population modelling amounts to estimation of the statistical distribution of the model parameters. A more recent trend is to rely on non-parametric models, which although computationally demanding do not require specification of the possible patterns of evolution, implicitly learning them from data. Our goal is to compare the flexibility and predictive power of these two approaches for modelling life-long athletic performance trajectories, highlighting their advantages and limitations. Significant efforts have recently been devoted to this problem, with both scientific (e.g. [6]) and forensic (e.g. [1]) goals, mostly using parametric models, and as far as we know this is the first time that non-parametric modelling is attempted in this domain.

II. Problem formulation

Let $\mathcal{P}$ be the population under study, and $\mathbf{Z}_i = \{Z_i(t), t \in T\}$, $i \in \mathcal{P}$, denote the (non-observable) evolution in T of the characteristic of interest for individual i. Let $\mathcal{Q} \subseteq \mathcal{P}$ be a representative sample of $\mathcal{P}$. Denote by $\mathbf{Y_i} = \{Y_i(t_{ik}), t_{ik} \in \mathbf{T}_i\}$ the vector of available observations for individual $i \in \mathcal{Q}$, where $\mathbf{T_i} \subset T$ are the times at which (s)he has been observed.

We assume that observations $Y_i(t)$ are a random function of the quantity of interest $Z_i(t)$, characterised by a parametric family of conditional distributions, such that

$$\mathbf{Y}_i|_{\mathbf{Z}_i} \sim p(\mathbf{Y}_i|\mathbf{Z}_i, \sigma_i), \qquad \sigma_i \in \mathbb{R}^+ \ .$$

Denote by $\mathcal{L} = \cup_{i\in\mathcal{Q}}\mathbf{Y}_i$ the set of available observations for model learning. Our goal to infer a probabilistic model for the individuals in $\mathcal{P}$: $p_{\mathcal{L}} \equiv p(\mathbf{Z}_i, \sigma_i|\mathcal{L}), i \in \mathcal{P}$.

In our application $Z_i(t)$ represents the fitness of athlete i at age t, and $Y_i(t_{ik})$ is the measured performance of athlete i in a competition done at age t_{ik}.

Figure 1 allows the appreciation of the major difficulties of the problem, on the example of men 400 meters running (similar patterns are observed for other distances). The top plot is the cloud of points of all available performances, the middle plot shows the individuals age spans, and in the bottom we plot the performances of three different athletes. We can remark that: *(i)* the temporal support of each series (middle plot) differs significantly across the population, with length ranging from 2 to 20 years (median: 8 years); *(ii)* each individual is non-uniformly sampled (bottom), with measurement concentration dictated by the competitions calendar; *(iii)* the variability of each athlete performances is strongly asymmetric (bottom): while extremely bad performances (large running times) can be due to injury or other accidental causes, exceptionally good performances have smaller departure from median performance.

III. Population models

A. Parametric model

Mixed-effects models [7] are hierarchical models for multiple measurements on the same individuals. Inter- and intra-individual variability are described by a family of parametric distributions. One of their main advantages is that they do not require a balanced dataset: the sets $\mathbf{T}_i$, including their cardinality n_i, can vary from one individual to another as it is the case in our application. For each individual, observations

ISBN 978-90-827970-1-5

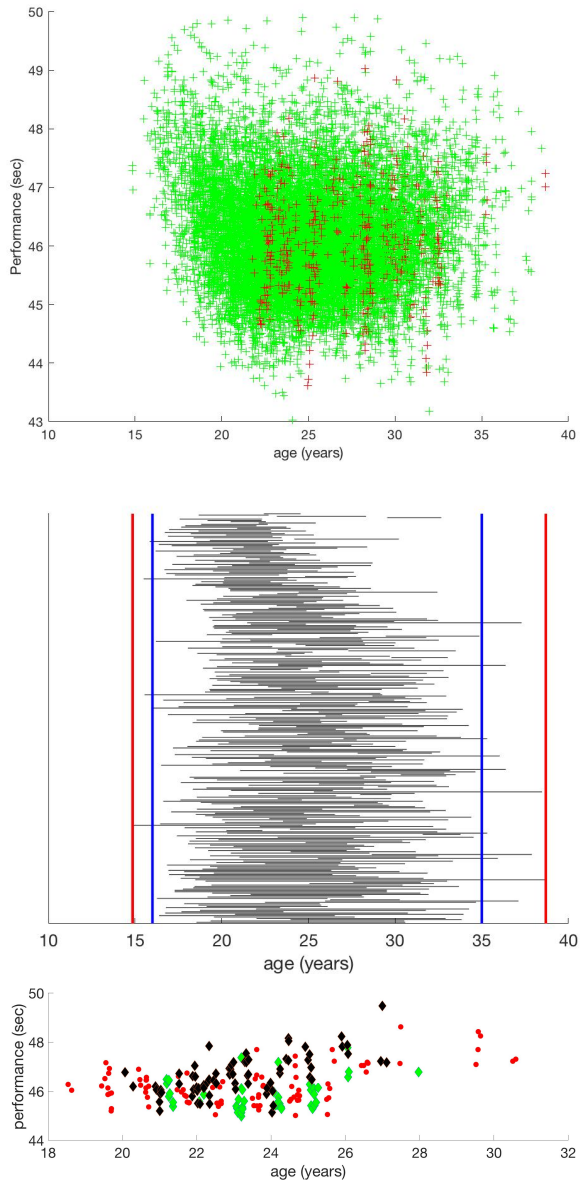

Fig. 1. 400 meters. Top: Learning (green) and testing (red) datasets for prediction test (see Section V-B). Horizontal axis: age (decimal years); vertical axis: performance time (in secs). Middle: Age span of athlete's time series. Red lines indicate the minimum and maximum ages of the dataset, blue lines indicate the support of the fitted GP model, see section IV. Bottom: measurements for three distinct athletes.

$\{Y_i(t_{ik})\}_{i=1}^{n_i}$ are modelled as noisy measurements of $Z_i(t_{ik})$ contaminated by additive noise

$$Y_i(t_{ik}) = Z_i(t_{ik}) + \varepsilon_{ik}, \quad Z_i(t) = Z(t, \alpha, \beta_i) \ .$$

The right-hand equation shows that $Z(t)$ is a parametric function of t with parameters $\alpha \in \mathbb{R}^p$ common to all the population (the fixed effects), and $\beta_i \in \mathbb{R}^q$, specific to each individual (the random effects). Choice of this parametric representation is commonly guided by domain knowledge.

The fixed-effects α are unknown constants, and all other variables are independent and Gaussian:

$$\{\varepsilon_{ik}\} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2), \qquad \{\beta_i\} \overset{iid}{\sim} \mathcal{N}(0, \Sigma) \ .$$

We focus on Linear Mixed Effects (LME) models where mean and individual trends are expressed in terms of a finite number of known basis functions, $Z_i(t) = \sum_{\ell=1}^{p} G_\ell(t)\alpha_\ell + \sum_{\ell=1}^{q} F_\ell(t)\beta_{i,\ell}$ leading to the simple matrix form:

$$\mathbf{Y}_i = X_i\alpha + \tilde{X}_i\beta_i + \varepsilon_i \ ,$$

where X_i is the $n_i \times p$ fixed-effects design matrix with elements $[X_i]_{k,\ell} = G_\ell(t_{ik})$, $\tilde{X}$ is the $n_i \times q$ random-effects design matrix with elements $[\tilde{X}_i]_{k,\ell} = F_\ell(t_{ik})$ and $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})^t$ is the error vector. The functions $G_\ell(\cdot)$ can simultaneously model a mean population trend, as well as the effect of exogenous covariates, e.g. the olympic cycle, in the context of sport performances.

B. Non-parametric model

In this approach, the observations

$$Y_i(t_{ik}) = Z_i(t_{ik}) + \varepsilon_i(t_{ik}) \ ,$$

are noisy versions of the fitness trajectories $\{Z_i(t), t \in T\}$, which are modelled as realisations of a Gaussian Process (GP)

$$\{Z_i(t), t \in T\} \overset{iid}{\sim} GP(\mu(\cdot), \Sigma(\cdot,\cdot)) \ ,$$

specified by a mean function $\mu(t), t \in T$ and a covariance function $\Sigma(t,u), t,u \in T$. GPs are a complete characterisation of the distribution of an ensemble of functions $\{z_\omega(\cdot), t \in T\}_{\omega\in\Omega}$ defined over some index set T. For any finite $\mathbf{t} = \{t_i\}_{i=1}^n \subset T$ the joint distribution of the random variables $z_\omega(\mathbf{t}) = \{z_\omega(t_i), t_i \in \mathbf{t}\}$ is Gaussian: $z_\omega(\mathbf{t}) \sim \mathcal{N}(\mu(\mathbf{t}), \Sigma(\mathbf{t},\mathbf{t}))$, where $[\mu(\mathbf{t})]_i = \mu(t_i)$ and $[\Sigma(\mathbf{t},\mathbf{t}))]_{ij} = \Sigma(t_i, t_j)$. The observation noises $\{\varepsilon_i(\cdot)\}$ are Gaussian, white, mutually independent and zero-mean, with covariances $\sigma^2_{\varepsilon_i}$.

In this study we use a minor modification of the hierarchical Bayesian approach to the estimation of the GP moments proposed in [3], which relies on prior distributions over $\mu(\cdot)$ and $\Sigma(\cdot,\cdot)$ as Gaussian and Inverse Wishart (IW) processes, respectively. Traditional approaches to GP identification assume parametric models for both $\mu(\cdot)$ and $\Sigma(\cdot,\cdot)$, most often a polynomial model for $\mu(\cdot)$ and stationary parametric kernel for $\Sigma(\cdot,\cdot)$. Stationarity is mandatory when a single realisation of the process is observed but it is not required when a series of realisations is available. The use of an Inverse Wishart prior distribution [4], whose essential support is the entire cone of symmetric positive definite matrices, allows departure from the restrictive stationarity assumption and to fully capture the internal correlation structure of the dataset. More precisely, the following priors are assumed: $\mu(\cdot)|_{\mu_0,\Sigma} \sim GP(\mu_0(\cdot), \Sigma(\cdot,\cdot))$, where $\mu_0(\cdot)$ is an hyper-parameter to be specified; $\Sigma|_{\nu,\sigma_x^2,A(\cdot,\cdot)} \sim IW(\nu, \sigma_x^2 A(\cdot,\cdot))$, where $\nu \geq 4$ and $\sigma_x^2 A(\cdot,\cdot)$ are the scale and location parameters of the IW distribution. Scale parameter ν is set to 4, corresponding to the least informative prior, $A(\cdot,\cdot)$ is specified by the user, and σ_x^2 follows an inverse Gamma distribution.

ISBN 978-90-827970-1-5

As the variability of the observations varies significantly from individual to individual we modified the original BABF model to use distinct $\sigma^2_{\varepsilon_i}$ for each individual.

IV. Model identification

A. Mixed-effects model

The parameters (Σ, σ^2) of the LME model are estimated by the restricted maximum likelihood method (REML). The fixed-effects are obtained by the generalized least squares formula using $\hat{\Sigma}$ and $\hat{\sigma}^2$:

$$\hat{\alpha} = \left(\sum_{i=1}^{N} X_i^t \hat{W}_i X_i \right)^{-1} \sum_{i=1}^{N} X_i^t \hat{W}_i Y_i \ ,$$

where $N = |\mathcal{Q}|$ is the number of observed individuals, and $\hat{W}_i = (\hat{\sigma}^2 I_{n_i} + \tilde{X}_i \hat{\Sigma} \tilde{X}_i^t)^{-1}$ is the inverse of the estimated covariance matrix of Y_i. Individual parameters are predicted by the empirical Bayes formula:

$$\hat{\beta}_i = \hat{\Sigma} \tilde{X}_i^t \hat{W}_i (Y_i - X_i \hat{\alpha}) \ .$$

The variance $\mathrm{Var}(\hat{Z}_i(\tau))$ of the prediction error for fitness $Z_i(\tau)$ at time τ takes into account the joint posterior distribution of the random- and fixed-effects, and is computed using the formulae in [8]. When predicting performance over an unobserved population this would lead to a slightly optimistic error characterisation, ignoring the errors in the estimation of the covariance Σ (see [9] for discussion). However, in the framework of our study, where the estimated model is used to predict future performances, these errors are best characterised by the induced biases.

In the results presented in section VI $q = p = 3$, and $[X_i]_{k\ell} = [\tilde{X}_i]_{k\ell} = t_{ik}^{\ell-1}, \ell = 1, \ldots, 3$ are polynomial design matrices.

B. GP model

Conjugacy of the chosen priors allows inference of the GP moments $\mu(\cdot), \Sigma(\cdot,\cdot)$ and of the intra-individual variabilities $\{\sigma^2_{\varepsilon_i}\}_{i=1}^N$ by using a Gibbs sampler, treating the noiseless versions of the observations $(Z_i(t_{ik}))$ as latent variables.

The distribution of $Y_i(t_{ik})$ depends only on the values of $\mu(t_{ik})$ and $\Sigma(t_{ik}, t_{ik'})$ at which the time series has been observed, and thus the distribution of the entire dataset depends only on the values of the process moments in $\cup_{i=1}^N \mathbf{T}_i$. Contrary to the LME model, the GP model is valid only in the age interval spanned by the data. If we denote by $I_i = [\min(T_i), \max(T_i)]$, the GP model temporal support must be $T \subset \cup_i I_i$.

BABF overcomes the large numerical complexity induced by the variation of the sets $\mathbf{T}_i$ by assuming that $Z_i(\cdot)$ belong to the span of a number $q \ll |\bigcup_i \mathbf{T}_i|$ of fixed knots B-splines.

Although the impact of the prior hyper-parameters $\mu_0(t)$ and $A(t,u)$, $t, u \in \mathbf{S}$ is low if the available dataset $\mathcal{L}$ is large, faster convergence is obtained if they are carefully initialised using $\mathcal{L}$. A common choice for μ_0 is the empirical mean of the available series. Since the sets $\mathbf{T}_i$ are distinct, we set

$$\mu_0(t) = \frac{1}{N_t} \sum_{i \in \mathcal{Q}_t} \tilde{Y}_i(t), \qquad t \in \mathbf{S} \ , \tag{1}$$

TABLE I
Size of Datasets

dataset	all		train		test	
event (m)	N obs.	N ind.	N obs.	N ind	N obs.	N ind
400	15474	308	14979	308	493	88
800	12201	226	11822	226	379	64
1500	11975	327	11627	327	348	72
5000	6741	345	6586	345	155	62
10000	1918	181	1872	181	46	33

where $N_t = |\mathcal{Q}_t|$ and $\mathcal{Q}_t \subset \mathcal{Q}$ is the set of trajectories for which $t \in [\min \mathbf{T_i}, \max \mathbf{T_i}]$, and $\tilde{Y}_i(t)$ is obtained by interpolating $\{Y_i(t_{ik})\}_{k=1}^{n_i}$ at t. Matrix $A(t,u), t, u \in \mathbf{S}$ is set to the empirical correlation, estimated in the same manner by further imposing positive definiteness. The hyper-parameters of the prior distribution for σ_x^2 and of the common prior for $\{\sigma^2_{\varepsilon_i}\}_{i=1}^N$ are inferred from the available data, see [3] for details.

GP modelling requires that N_t be sufficiently large over the entire modelled interval. In our application we set $S = [16, 35]$ (indicated by the blue vertical lines in Figure 1), $q = 10$ and $\mathbf{S}$ is a uniform 20-points grid spanning T.

V. Design of tests for performance analysis

A. Datasets

From the database AllAthletics.com of performances in finals of official competitions on 5 outdoor running events (400, 800, 1500, 5000 and 10000 meters) in the years 1997–2017, we extracted the records of men athletes *(i)* who are high-level, i.e., whose personal best is above entry standard for the World Championships of 2007; *(ii)* have measurements at least for three distinct years in the age interval $[20, 35]$. Table I gives the number of observations (N obs) and the number of athletes (N ind) in each dataset.

B. Prediction performance

We compare the in-sample prediction performance of our models. Consider distance $D \in \{400, 800, 1500, 5000, 10000\}$, and let $\mathcal{Q}_D$ be the corresponding dataset. Using the LME and the GP models learned on all available measurements for years up to 2016, we compute the one-step ahead prediction estimates for the athletes' fitness during year 2017. The top plot in Figure 1 shows the learning (green) and testing (red) datasets for $D = 400$. Note the diversity of the athlete's ages in the predicted year (2017).

Predictive performance has been tested in the following recursive manner. Denote by $\mathbf{Y}_i^T$ the set of measurements for athlete i in year 2017, occurring at ages $\mathbf{T}_i^T = \{t_1 \leq t_2 \leq \ldots, \leq t_{r_i}\}$. For each $k = 1, \ldots, r_i$ we used all available measurements $\mathbf{T}_{ik}^T = \mathbf{T_i}^L \cup \{t_1, \ldots, t_{k-1}\}$ for athlete i previous to t_k, to predict the fitness at age t_{ik}, $Z_i^M(t_k)$.

Figure 2 illustrates the results for one athlete for both LME (top) and GP (bottom), for $D = 400$. The red stars ⋆ indicate the measures used to learn the model and the red circles ∘ the test set, i.e. the performances being predicted. The blue stars

ISBN 978-90-827970-1-5 © EURASIP 2018

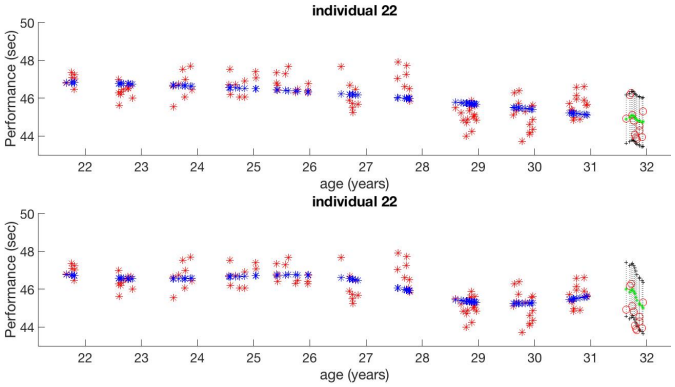

Fig. 2. Learning ($\star$) and testing ($\circ$) datasets; estimated ($\star$) and predicted ($\star$) fitness; confidence intervals for $Y_i(t_{ik})$ (black vertical lines), $D = 400$ meters. Top: LME; bottom: GP.

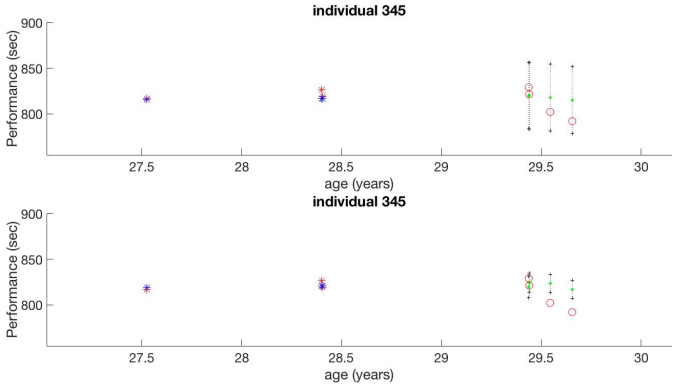

Fig. 3. Learning ($\star$) and testing ($\circ$) datasets; estimated ($\star$) and predicted ($\star$) fitness; corresponding confidence intervals for $Y_i(t_{ik})$ (black vertical lines), $D = 5000$ meters. Top: LME; bottom: GP.

$\star$ are the fitness estimates using the entire learning set. The green stars $\star$ are the predictions for 2017, using the identified population model and the measures in $\mathbf{T}_{ik}^T$. The black vertical lines are the 95% estimated confidence intervals according to the estimated models.

Denote by $\delta_{ik} = Y_i(t_k) - \hat{Z}_i(t_k)$ the prediction residuals, $E[\delta]^M$ their empirical mean, and Δ_{ik}^M their normalised version: $\Delta_{ik}^M = {}^{\delta_{ik}}/_{\sigma_{ik}^M}$, $M \in \{\text{LME}, \text{GP}\}$, where σ_{ik}^M is the estimated standard deviation of δ_{ik} :

$$\sigma_{ik}^M = \sqrt{(\sigma_{\varepsilon_i}^M)^2 + \sigma_{Z_i}^M(t_{ik})^2} \ . \tag{2}$$

Performance is assessed by the normalised error metric

$$C_{msqe}^M = \frac{1}{K} \sum_{i \in \mathcal{Q}_D} \sum_{t_{ik} \in \mathbf{T}_{ik}^T} (\Delta_{ik}^M)^2 \ , \tag{3}$$

where $K = \sum_{i \in \mathcal{Q}_D} |T_{ik}^T|$, which should ideally be close to one if the uncertainty intervals are correctly estimated.

Interesting indicators of the asymmetry of large residuals are

$$P_+^M = \frac{1}{K} \sum_{i = \in \mathcal{Q}_D, k \in \{1,\dots,r_i\}} \mathbb{I}\left[\Delta_{ik}^M > 1.96\right] \ , \tag{4}$$

where $\mathbb{I}(\cdot)$ is the indicator function, and P_-^M, which is defined analogously for $\Delta_{ik}^M < -1.96$.

VI. Model comparison

The two fitted models differ in the following aspects: *(i)* LME strongly constrains the possible evolutions of fitness; *(ii)* a common variability of the performance measures is assumed for LME (a common σ_ε^2), while distinct $\sigma_{\varepsilon_i}^2$ are used in GP.

The top plot in Fig. 5 shows, for 400 m, the empirical mean population performance trajectory $\mu_0(t)$ (dotted black) and the density of the ages in the complete learning set (in yellow, values adjusted for visualisation). The GP mean (red curve) captures closely the observed pattern, while the LME mean (blue curve) shows a weak agreement with the empirical curve in the less frequent age ranges. The LME constraint is less penalising when the population evolution is closer to a quadratic curve, as for $D = 5000$ meters.

Table II summarises the performance criteria, for two normalisations of the residuals, and Fig. 4 shows the histograms of the prediction residuals δ_{ik} for LME (left) and GP (right) for 400 meters. Both methods lead to a similar residual distribution, with small biases that are larger for GP than for LME, see the third column in Table II. A similar behaviour is observed for other distances.

TABLE II
Next year (2017) performance

D	M	$E[\delta]^M$	C_{msqe}^M	P_+^M, %	P_-^M %
$\sigma_{ik}^M = \sqrt{(\sigma_{\varepsilon_i}^M)^2 + \sigma_{Z_i}^M(t_{ik})^2}$					
400	LME	-0.07	0.89	2.84	0.41
	GP	-0.06	1.13	4.52	2.05
800	LME	-0.02	0.77	2.90	0.0
	GP	-0.03	1.40	4.49	0.79
1500	LME	-0.03	0.98	4.60	0.0
	GP	-0.10	1.15	5.48	0.29
5000	LME	-0.05	0.69	3.23	0.0
	GP	-2.15	1.32	4.7	2.68
10000	LME	0.03	0.93	0.0	0.0
	GP	2.78	2.30	9.52	7.14
$\sigma_{ik}^M = \sigma_{\varepsilon_i}^M$					
400	LME	-0.07	1.04	3.45	1.42
	GP	-0.06	1.39	6.57	3.29
800	LME	-0.02	0.87	3.17	0.0x
	GP	-0.03	1.13	5.54	0.79
1500	LME	-0.03	1.07	5.17	0.0
	GP	-0.10	1.30	6.34	0.86
5000	LME	-0.05	0.77	4.52	0.0
	GP	-2.15	2.07	5.37	3.36
10000	LME	0.03	1.08	2.17	4.35
	GP	-2.78	5.43	14.29	11.90

As the last three columns of Table II show the estimated variance of the LME residuals tends to be larger than for GP, i.e., $\sigma_{ik}^{LME} > \sigma_{ik}^{GP}$. Consequently, P_-^{GP} and P_+^{GP} are always larger than for the LME model and $C_{msqe}^{LME} < C_{msqe}^{GP}$ in all cases. The asymmetry of athlete's performance variability is well reflected in the fact that $P_+ > P_-$ almost always, confirming that exceptionally bad performances occur more often than good ones.

ISBN 978-90-827970-1-5 © EURASIP 2018

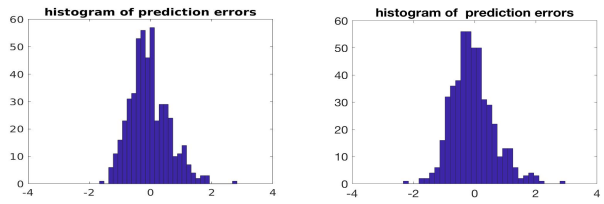

Fig. 4. Histograms of residuals ($D = 400$ meters). Left: LME, right: GP.

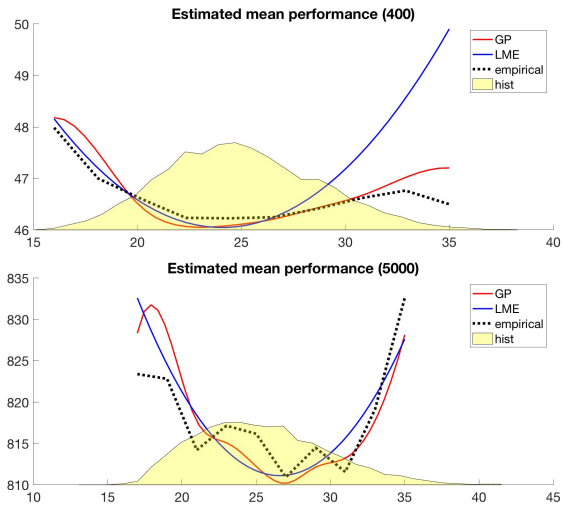

Fig. 5. Estimated population mean performance trajectory (top: 400 m, bottom 5000 m). Red: GP $\mu(t)$; blue: LME $Z(t, \alpha, \beta_0)$; black dashed: empirical estimate $\mu_0(t)$; yellow: histogram of ages in the dataset.

Ideally, $C^M_{msqe} \simeq 1$. Table II shows that C^{LME}_{msqe} tends to be below 1, indicating an overestimation of uncertainty, and that C^{LME}_{msqe} is, with a single exception, larger than 1, indicating an underestimation of uncertainty. Neglecting the uncertainty of the fitness predictions leads to values closer to 1 for LME, indicating that the noise variance is sufficient to model expected variability, and fitness uncertainty is often over-estimated for this model. On the contrary the values of C^{GP}_{msqe} are, with one exception, always greater than one, indicating an uncertainty under-estimation, worst in the bottom sub-table.

The degradation with distance of the GP models is explained by the smaller sizes of the learning datasets for larger values of D (see Table I). LME is much less sensitive to this problem.

The different behaviour of the two methods can be appreciated in Figs. 2 and 3, for 400 and 5000 meters, respectively, representative of small and large distances. In the plot for 400 meters the different "plasticity" of the two models is obvious, the GP model being best able to track the athlete trajectory. This higher plasticity induces also a more reactive behaviour of the GP model to the observed residuals, adapting faster to the athlete's performance variations than LME. Fig. 3, for 5000 meters, is a good example of the situation that is frequent for larger distances, where a very small number of observations per individual may be available. The LME model has a wide uncertainty interval, while the GP model has been able to closely fit the few observed performances, and predicts a too small variability.

VII. Conclusions

The results presented confirm that the constrained LME model offers a more robust approach to population modelling, while when the available dataset conveniently samples the trajectories of the observed individuals, the GP model is best able to express the dataset detailed characteristics.

A major conclusion of the performance analysis of both methods concerns the importance of the statistical modelisation of the individual variability of performance, the $\{\varepsilon_{ik}\}$. Both our models are based on a simple Gaussian assumption, which, as the plots in the paper show, is a poor model for the actual observed variability. This is increasingly important for long running distances, where very large positive outliers (worst performance, i.e. larger time) are frequent. We consider that the key point for improving prediction performance in this application relies on abandoning the assumption of normality of the ε_{ik}, using for instance mixture models, that can best capture the asymmetric intra-individual variability of athlete's performances.

Finally, the study confirms the high plasticity of the GP model, which if enough data is available actually captures the trends in the set of processed individual trajectories. A present line of research is to combine the two modelling approaches, using the GP model to identify a convenient basis for a mixed-effects model. This would further facilitate the consideration of exogenous factors, which cannot be naturally incorporated in the GP model.

References

[1] S. Iljukov, Y. O. Schumacher, "Performance Profiling-Perspectives for Anti-doping and beyond". Front Physiol. vol. 8:1102. doi: 10.3389/fphys.2017.01102, December 22, 2017.

[2] P. E. Sottas, N. Robinson, M. Saugy, and O. Niggli, "A forensic approach to the interpretation of blood doping markers", Law, Probability and Risk, vol. 7(3), pp. 191–210, 2008.

[3] J. Yang, D. D. Cox, J. S. Lee, P. Ren, T. Choi, "Efficient Bayesian hierarchical functional data analysis with basis function approximations using GaussianWishart processes", Biometrics, vol. 73(4), pp. 1082–1091, December 2017.

[4] J. Yang, H. Zhu, T. Choi and D. D. Cox, "Smoothing and MeanCovariance Estimation of Functional Data with a Bayesian Hierarchical Model. Bayesian Analysis", Bayesian Analysis, vol. 11(3), pp. 649–670, November 2016.

[5] C. E. McCulloch, J. M. Neuhaus, "Generalized linear mixed models". John Wiley & Sons, Ltd, 2001.

[6] R. M. Malcata, W. G. Hopkins, S. N. Pearson, "Tracking career performance of successful triathletes", Med. Sci. Sports Exerc. , vol. 46(6), pp. 1227–34, June 2014.

[7] E. Demidenko, E., "Mixed models: theory and applications with R", John Wiley & Sons, Ltd, 2014.

[8] D. Harville, "Extension of the Gauss-markov theorem to include the estimation of random effects", The Annals of Statistics, vol. 4(2), pp 384-395, 1976

[9] K. Das, J. Jiang and J. N. K. Rao, "Mean squared error of empirical predictor", Annals of Statistics, vol. 32(2), pp. 818–840, 2004.

ISBN 978-90-827970-1-5 © EURASIP 2018