On the Impact of Different Kernels and Training Data on a Gaussian Process Approach for Target Tracking

Waqas Aftab, Lyudmila Mihaylova

Department of Automatic Control and Systems Engineering, University of Sheffield, S1 3JD, UK
{waftab1,l.s.mihaylova}@sheffield.ac.uk

***Abstract*—The application of multiple target tracking algorithms has exponentially increased during the last two decades. Recently, model-free approaches, such as Gaussian process regression and convolutional neural networks, have been developed for target tracking. This paper presents a simulation-based study on the practical aspects of a very promising and recently proposed Gaussian process method, namely the Gaussian process motion tracker [1]. The paper also provides design guidelines on the various aspects of the above-mentioned tracking method.**

***Index Terms*—Target Tracking, Gaussian Process, Gaussian Process Motion Tracking, Nonlinear Estimation, Data Driven Methods**

I. Introduction

Target tracking deals with state estimation of the targets of interest using sensor data. These methods have been applied in various automation systems belonging to various fields such as air traffic control [2], sea surveillance [3], oceanography [4], autonomous vehicles [5] and many more. Historically, model-based approaches have been applied for solving target tracking problems. Recently, machine-learning based model-free methods have been proposed either as a complete solution [6] or in a hybrid setup [7], [8]. Hybrid methods combine model-based and model-free methods for target tracking.

Machine learning methods rely on the available data to learn an unknown input to output mapping. Various machine learning methods have been proposed in literature for different artificially intelligent systems. In the field of target tracking, deep learning and Gaussian process based methods have become popular recently. However, providing a level of trust in the developed solutions, in the presence of uncertainties, is still a problem that has not been widely studied. The tracking results are often an input to a human or computer based decision making system which cannot perform satisfactorily without uncertainty measures. Gaussian process methods, have been recently proposed as an efficient solution to different target tracking problems [6], [7], [9].

Target tracking methods can be classified as point, extended and group target tracking [10], [11] methods. Point target tracking requires target kinematics estimation, i.e. the estimation of positions, velocities and accelerations of the objects of interest. In tracking of extended and groups of objects, in addition to the kinematics, we are interested in the estimation of the target shape, orientation and size. Gaussian process based methods have been proposed for point, extended and group target tracking. In this paper, various aspects of the Gaussian process motion tracker (GPMT) [1] are studied. The GPMT provides point target state estimation in presence of unknown target dynamics and measurement noise. The measurement to target assignment (measurement origin) is assumed known.

A Gaussian process is a flexible stochastic process. It is a learning based method which relies on a given data for learning the unknown model. Some example tasks performed using a Gaussian process are classification, regression and pattern recognition. A Gaussian process has also been applied for the target shape estimation [9], [12], [13]. The GPMT employs the Gaussian process in a regression setting for the target kinematics estimation. The paper [1] does not discuss some important aspects of the approach. These include the choice of the covariance kernel, robustness of the approach to the measurement noise model and effect of the training data on the proposed method. This paper focuses on the above-mentioned aspects of the GPMT in an attempt to highlight the strength of the proposed method.

The rest of the paper is structured as follows. The background knowledge for a Gaussian process and the GPMT are given in Sections II and III, respectively. The kernel choice, the impact of the training data and the robustness of the GPMT to the measurement noise is studied in Sections V, VI and VII, respectively. The studies are followed by conclusions.

II. Gaussian Process

A Gaussian process (GP) is a distribution over functions and is defined by a mean and covariance kernel [14]. It is a powerful non-parametric method which has been applied to solve various problems in the domain of artificial intelligence such as heart rate analysis [15], classification [16] and pattern recognition [17]. The Gaussian process regression is briefly described next since the Gaussian process motion tracker is based upon it.

Consider a one-dimensional input u which relates non-linearly to an output $f(u)$ and is modelled using a Gaussian process. The measurement model is assumed with an additive Gaussian noise and given below:

$$z = f(u) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma^2), \tag{1}$$

$$f(u) \sim GP(m(u), k(u, u')), \tag{2}$$

where z is the measured output, ϵ denotes the zero-mean independent identically distributed (i.i.d.) Gaussian noise with variance σ^2 and $GP(m(u), k(u, u'))$ specifies the GP model with mean $m(u)$ and covariance kernel $k(u, u')$. A GP is a learning based method and requires data to learn the unknown function. Suppose $\boldsymbol{u} = [u_1, \cdots, u_n]^T$ and $\boldsymbol{z} = [z_1, \cdots, z_n]^T$ denote, respectively, the input and output vectors, also called *training data*. A GP regression on the testing input vector $\boldsymbol{u}^\star$ can be applied by using the property of GP models. This means that the realisations of the GP have a joint Gaussian distributed. This is mathematically expressed below:

$$\begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{z}^\star \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{m}(\boldsymbol{u}) \\ \boldsymbol{m}(\boldsymbol{u}^\star) \end{bmatrix}, \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{uu}} + \sigma^2 \boldsymbol{I}_n & \boldsymbol{K}_{\boldsymbol{uu}^\star} \\ \boldsymbol{K}_{\boldsymbol{u}^\star\boldsymbol{u}} & \boldsymbol{K}_{\boldsymbol{u}^\star\boldsymbol{u}^\star} \end{bmatrix} \right), \tag{3}$$

$$\boldsymbol{K}_{\boldsymbol{uu}} = \begin{bmatrix} k(u_1, u_1) & \cdots & k(u_1, u_n) \\ \vdots & \ddots & \vdots \\ k(u_n, u_1) & \cdots & k(u_n, u_n) \end{bmatrix}, \tag{4}$$

where $\boldsymbol{m}(\cdot)$ denotes the mean vector, $\boldsymbol{K}_{\boldsymbol{uu}^\star}$ is the covariance matrix between the input training and test vectors and $\boldsymbol{I}_n$ is an n-dimensional identity matrix. The GP prediction at the test vector is given below:

$$\mathbb{E}[\boldsymbol{z}^\star] = \boldsymbol{m}(\boldsymbol{u}^\star) + \boldsymbol{K}_{\boldsymbol{u}^\star\boldsymbol{u}} \left(\boldsymbol{K}_{\boldsymbol{uu}} + \sigma^2 \boldsymbol{I}_n \right)^{-1} (\boldsymbol{z} - \boldsymbol{m}(\boldsymbol{u})), \tag{5}$$

$$\mathbb{C}[\boldsymbol{z}^\star] = \boldsymbol{K}_{\boldsymbol{u}^\star\boldsymbol{u}^\star} - \boldsymbol{K}_{\boldsymbol{u}^\star\boldsymbol{u}} \left(\boldsymbol{K}_{\boldsymbol{uu}} + \sigma^2 \boldsymbol{I}_n \right)^{-1} \boldsymbol{K}_{\boldsymbol{uu}^\star}, \tag{6}$$

where $\mathbb{E}[\boldsymbol{z}^\star]$ and $\mathbb{C}[\boldsymbol{z}^\star]$ represent, respectively, the mean and the covariance of the output test vector and $(.)^{-1}$ is the matrix inverse. The above GP regression be easily extended to the multiple-input and multiple-output case.

The flexibility of the GP regression is linked with the mean and the covariance kernel, which encapsulate the prior. The average behaviour of most stochastic processes is unknown. Similar is the case of the GPMT [1], where the target trajectory is assumed unknown. In such cases, the mean of the GP is set to zero. It is important to understand that this does not restrict the GP regression, except for cases when GP predictions away from the training data would converge to different results. The covariance kernel, captures the correlations among the input space and it is an important design parameter of GP models.

Various covariance kernels have been proposed for the GP regression, some of which are described briefly below. The two common parameters of the covariance kernels, also called *hyperparameters*, are the magnitude variance σ_m^2 and the lengthscale l hyperparameters. The magnitude of the variance controls the average distance between the mean function and the mean of the GP regression. The lengthscale controls the correlation width of the input domain.

1) **Squared Exponential Kernel:**

$$k_{se}(u, u') = \sigma_m^2 \exp\left(- \frac{(u - u')^2}{2l^2} \right) \tag{7}$$

is the most commonly used kernel [14]. It is a very smooth kernel and is infinitely differentiable.

2) **Rational Quadratic Kernel:** is in the form

$$k_{rq}(u, u') = \sigma_m^2 \left(1 + \frac{(u - u')^2}{2\alpha l^2} \right)^{-\alpha}, \tag{8}$$

where α is a scaling factor. The rational quadratic kernel behaves as sum of squared exponential kernels with different lengthscales. The lengthscales are varied using the α hyperparameter. The rational quadratic kernel meets the squared exponential kernel as $\alpha \to \infty$.

3) **Matérn Kernel**.

$$k_\nu(u, u') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}(u - u')}{l} \right)^\nu K_\nu \left(\frac{\sqrt{2\nu}(u - u')}{l} \right),$$

where $\nu > 0$ and K_ν is a modified Bessel function. Unlike other kernels, this function gives a class of kernels. Various kernels belonging to the Matérn class can be built for different values of ν. As $\nu \to \infty$, the kernel approaches a squared exponential kernel. As $\nu \to 0$, the kernel approaches an exponential kernel. A well known kernel from this class is obtained by setting $\nu = \frac{3}{2}$ [14] and is given below:

$$k_{\frac{3}{2}}(u, u') = \left(1 + \frac{\sqrt{3}(u - u')}{l} \right) \exp\left(- \frac{\sqrt{3}(u - u')}{l} \right), \tag{9}$$

Although, the GP models are quite flexible for constant hyperparameters. The model adaptation can be improved by determining the hyperparameters based on the training data. This process is also called *learning*. The learning is performed through optimisation of the marginal likelihood with respect to the hyperparameters. The logarithm of the marginal likelihood function is given below:

$$\log p(\boldsymbol{z}|\boldsymbol{u}, \boldsymbol{\eta}) = -\frac{1}{2} \boldsymbol{z}^T (\boldsymbol{K}_{\boldsymbol{uu}} + \sigma^2 \boldsymbol{I}_n)^{-1} \boldsymbol{z} - \frac{1}{2} \log |\boldsymbol{K}_{\boldsymbol{uu}} + \sigma^2 \boldsymbol{I}_n| - \frac{n}{2} \log 2n, \tag{10}$$

where $p(\cdot)$ denotes the marginal likelihood, $\boldsymbol{\eta}$ denotes the hyperparameters vector and $|\cdot|$ is the matrix determinant.

III. Gaussian Process Motion Tracker

The Gaussian process motion tracker (GPMT), proposed in [1], estimates the two-dimensional kinematics of the point targets using noisy measurements. The tracker is based upon the following assumptions:

1) The kinematics in x and y are mutually uncorrelated.
2) The coordinates are temporally correlated.
3) The temporal correlation with points in the distant past is weak and these points can be ignored while training of the GP model.
4) The measurement noise is an i.i.d. process.

One of the most commonly observed target manoeuvre model is coordinated turn. The x and y coordinates are correlated during a coordinated turn. In GPMT, the coordinates are assumed mutually uncorrelated. The coordinate coupling can be introduced in GPMT using coupled GPs [18]. The GPMT in x-coordinate is given in this section. A similar tracker can be built for the y-coordinate. It can be extended to any number of dimensions. Suppose, f^x represents the nonlinear target dynamics function in x coordinate.

The GPMT system model is given below:

$$x = f^x(t), \quad f^x(t) \sim GP^x(0, k^x(t,t')), \tag{11}$$

$$z^x = x + \epsilon^x, \quad \epsilon^x \sim \mathcal{N}(0, \sigma_x^2), \tag{12}$$

where GP^x denotes the GP model of the x-coordinate with covariance kernel k^x, z^x is the measurement and ϵ^x represents the i.i.d. zero-mean measurement noise with variance σ_x^2. A typical radar and sonar reports measurements in polar coordinates. In such scenarios, the process and measurement models, (11) and (12), are modelled in the polar coordinates. An alternate approach can be to calculate the measurement pdf in Cartesian coordinates [19], [20]. The performance may be degraded in the latter case as the cross-correlation among the x and y coordinates is ignored according to 11.

The GP regression is a batch processing method. The GPMT is an online method where it requires a subset of measurements for the purposes of prediction and estimation. It requires the d most recent measurements, the position prediction and estimation, as given below:

$$\tilde{\mu}_x = \boldsymbol{K}_{t\boldsymbol{t}}[\boldsymbol{K}_{\boldsymbol{tt}} + \sigma_x^2 \boldsymbol{I}_d]^{-1} \boldsymbol{z}_{\boldsymbol{t}}^x, \tag{13}$$

$$\tilde{\phi}_x^2 = \boldsymbol{K}_{tt} - \boldsymbol{K}_{t\boldsymbol{t}}[\boldsymbol{K}_{\boldsymbol{tt}} + \sigma_x^2 \boldsymbol{I}_d]^{-1} \boldsymbol{K}_{t\boldsymbol{t}}^T, \tag{14}$$

$$\hat{\mu}_x = \boldsymbol{K}_{t\boldsymbol{t}'}[\boldsymbol{K}_{\boldsymbol{t}'\boldsymbol{t}'} + \sigma_x^2 \boldsymbol{I}_d]^{-1} \boldsymbol{z}_{\boldsymbol{t}'}^x, \tag{15}$$

$$\hat{\phi}_x^2 = \boldsymbol{K}_{tt} - \boldsymbol{K}_{t\boldsymbol{t}'}[\boldsymbol{K}_{\boldsymbol{t}'\boldsymbol{t}'} + \sigma_x^2 \boldsymbol{I}_d]^{-1} \boldsymbol{K}_{t\boldsymbol{t}'}^T, \tag{16}$$

where $t = k$, $\boldsymbol{t} = [k-d, k-d+1, \cdots, k-1]^T$, $\boldsymbol{t}' = [k-d+1, k-d+2, \cdots, k]^T$, μ_x and ϕ_x^2 denote, respectively, the positional mean and variance, $\tilde{}$ and $\hat{}$ represent, respectively, the predicted and the estimated values, $\boldsymbol{K}$ is the covariance matrix and is determined using (4) and $\boldsymbol{z}_{\boldsymbol{a}}^x$ represents the measurement vector consisting of samples corresponding to time vector $\boldsymbol{a}$. The GPMT proposes to determine d using an offline trial and error method. The method can be improved through online determination of d as proposed in [21].

In [1], a squared exponential covariance kernel has been proposed. The learning is done using the maximum likelihood approach. It has been shown in [1], that the proposed GPMT performs better than the model based approaches including fixed grid interacting multiple model in challenging scenarios. Tracking by using the position derivatives has also been proposed in [1]. However, here we restrict the study the position estimates only.

IV. Testing Scenarios and Performance Evaluation

The simulation-based studies are based upon the target scenarios and the evaluation methods described in this section. The root mean square error (RMSE) of the target predicted position is chosen as the main performance measure. A comprehensive database of the point target trajectories is not publicly available. Hence, the target trajectories are generated using the three most commonly used point target dynamics models. These are the nearly constant velocity (NCV) [22], the nearly coordinated turn (NCT) [22] and the Singer acceleration model [23]. The state transition and the process noise covariances of the three models are given below:

$$\boldsymbol{F}^{NCV} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \boldsymbol{Q}^{NCV} = q^{NCV} \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix}, \tag{17}$$

$$\boldsymbol{F}^{NCT} = \begin{bmatrix} 1 & \frac{\sin \omega T}{\omega} \\ 0 & \cos \omega T \end{bmatrix}, \boldsymbol{Q}^{NCT} = q^{NCT} \begin{bmatrix} \frac{T^4}{4} & \frac{T^3}{2} \\ \frac{T^3}{2} & T^2 \end{bmatrix}, \tag{18}$$

$$\boldsymbol{F}^{s} = \begin{bmatrix} 1 & T & \frac{\beta - 1 + \gamma}{\alpha^2} \\ 0 & 1 & \frac{1-\gamma}{\alpha} \\ 0 & 0 & \gamma \end{bmatrix}, \boldsymbol{Q}^{s} = \frac{2\sigma_m^2}{\alpha} \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}. \tag{19}$$

$$q_{11} = \frac{1 - \gamma^2 + 2\beta + \frac{2\beta^3}{3} - 2\beta^2 - 4\beta\gamma}{2\alpha^5}, \tag{20}$$

$$q_{12} = q_{21} = \frac{\gamma^2 + 1 - 2\gamma + 2\beta\gamma - 2\beta + \beta^2}{2\alpha^4}, \tag{21}$$

$$q_{13} = q_{31} = \frac{1 - \gamma^2 - 2\beta\gamma}{2\alpha^3}, \tag{22}$$

$$q_{22} = \frac{4\gamma - 3 - \gamma^2 + 2\beta}{2\alpha^3}, \tag{23}$$

$$q_{23} = q_{32} = \frac{\gamma^2 + 1 - 2\gamma}{2\alpha^2}, \tag{24}$$

$$q_{33} = \frac{1 - \gamma^2}{2\alpha}, \tag{25}$$

where $\boldsymbol{F}^{\cdot}$ and $\boldsymbol{Q}^{\cdot}$ represent, respectively, the state transition and the process noise covariance with variance $q^{\cdot}$, T denotes the sampling time, ω is the turn rate, σ_m^2 is the manoeuvre variance, $\alpha = \frac{1}{\tau_m}$ is the reciprocal of the manoeuvre sojourn time τ_m, $\beta = \alpha T$ and $\gamma = \exp(-\beta)$. The above three models can represent real target trajectories.

The sampling time is set to $T = 1s$, the total samples are $K = 100$, the measurement noise standard deviation is set to $\sigma = 25m$, the probability of detection is set to $p_d = 1$, the initial target velocity is randomly chosen in the limits $150m/s \leq v_0 \leq 250m/s$, the process noise variances of the NCV and the NCT models are set to $q^{NCV} = q^{NCT} = 1e{-}12$, the turn rate is set to $\omega = 15 \deg /s$, the manoeuvre variance is set to $\sigma_m^2 = 168.75 m^2/s^4$ and the manoeuvre sojourn time is set to $\tau_m = 8s$. The coordinated turn model based scenario switches between the NCV and the NCT models. The sojourn time of the NCT based manoeuvre is $8s$. The results are computed over 1000 Monte Carlo runs.

V. Choice of Covariance Kernels

The GPMT is proposed using a squared exponential (SE) covariance kernel in [1]. This kernel is infinitely differentiable and it helps in tracking all the higher derivatives of the position coordinates. The kernel is, however, too smooth as compared to the real target dynamics. In this section, a simulation-based study is performed to compare the performance of the different covariance kernels. The two new kernels chosen in this study are the rational quadratic (RQ) and the Matérn (with $\nu = \frac{3}{2}$) kernels. The results are given in Fig. 1.

It can be observed that the SE and the RQ perform better than the Matérn for the NCV and Singer target dynamics models based trajectories. However, the Matérn kernel outperforms them for the NCT model based trajectory. The RQ based GPMT performs slightly worse as compared to the Matérn kernel. The performance of the SE based GPMT is significantly poor and could be a bad choice for this type of trajectory. Based on the above study, the following recommendations are made:

1) For the NCV and Singer based scenarios, the SE based GPMT should be chosen.
2) For applications involving target trajectories based on all three models, the RQ based GPMT is the preferred choice.

VI. Effect of the Training Data

The parameter d of the GPMT controls the size of the training data set. The evaluation of the GPMT in paper [1] is done by setting it as $d = 10$, that is, the 10 most recent measurements are considered for the training of the model. In this section, the performance of the GPMT for the different values of d is studied for the three kernels. The results are given in the Figs. 2, 3 and 4.

It can be observed, in Fig. 2, that the accuracy of the SE based GPMT increases with the increase in the training data for the NCV model based trajectories. For the remaining two scenarios, the accuracy decreases. It can be observed, in Fig. 3, that the accuracy of the RQ based GPMT increases with the increase in the training data for the NCV and the NCT model based trajectories. The performance degrades with the increase in the training data for the Singer model. In Fig, 4, it can be observed that the accuracy of the Matérn based GPMT increases with the increasing training data. Based on the above results, it is recommended to use a Matérn kernel based GPMT when the training data size is important for the application.

VII. Robustness to Measurement Noise Model

The measurement noise variance can be set as a hyperparameter of the GPMT [1] and learned recursively from the training data. In this way, the GPMT model is robust to the measurement noise variance. This section provides a simulation based study on the performance degradation of the GPMT with the increasing noise variance. The noise standard deviation is chosen as $\sigma = 25, 50, 75, 100$. The percentage increase in the standard deviation of the noise with respect to $\sigma = 25$ is $100\%, 200\%$ and 300%. The percentage degradation of the three kernels for the assumed scenarios is given in Table I. It can be observed that although the accuracy decreases with the increase in the noise, the filter does not diverge.

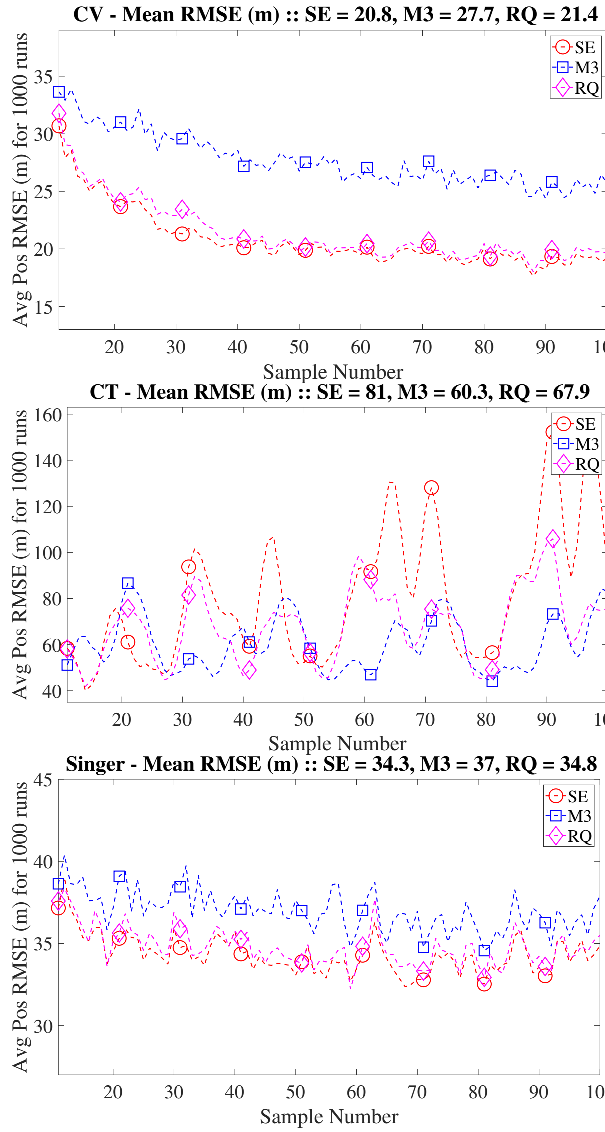

Fig. 1. **Comparison of covariance kernels**. This figure show results of the positional prediction using three different variants of the GPMT based on different covariance kernels. These are the squared exponential (SE), the Matérn with $\nu = \frac{3}{2}$ (M3) and the rational quadratic (RQ) kernels. The three plots correspond to three different target models which are the NCV (top), the NCT (middle) and the Singer (bottom).

TABLE I
Performance Degradation with Increased Noise Variance

	NCV			**NCT**			**Singer**		
	100	***200***	***300***	***100***	***200***	***300***	***100***	***200***	***300***
SE	92	181	270	34	64	88	64	127	187
RQ	93	183	273	41	80	112	65	127	187
M3	78	150	217	49	94	132	70	131	187

VIII. Conclusions

This paper presents a simulation based study of different aspects of the Gaussian process approach proposed in [1] for the point target tracking. The study demonstrates that the rational quadratic kernel is a better choice, as compared to

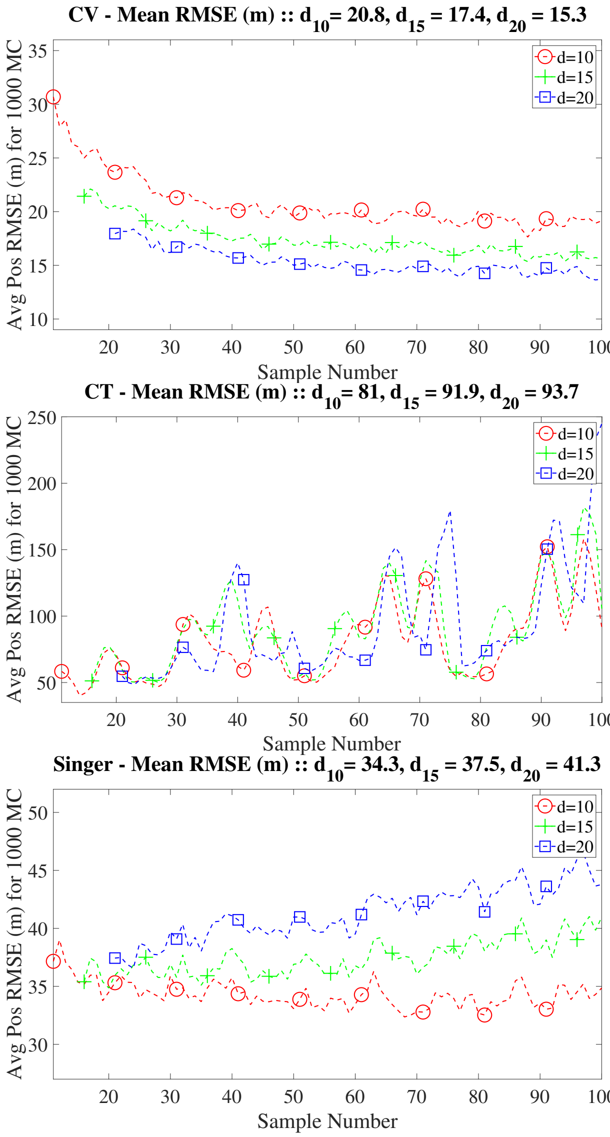

Fig. 2. **Effect of the training data on the SE kernel**. This figure shows the results for the three scenarios, as explained in Fig.1. The three different values of the parameter are $d = 10, 15, 20$.

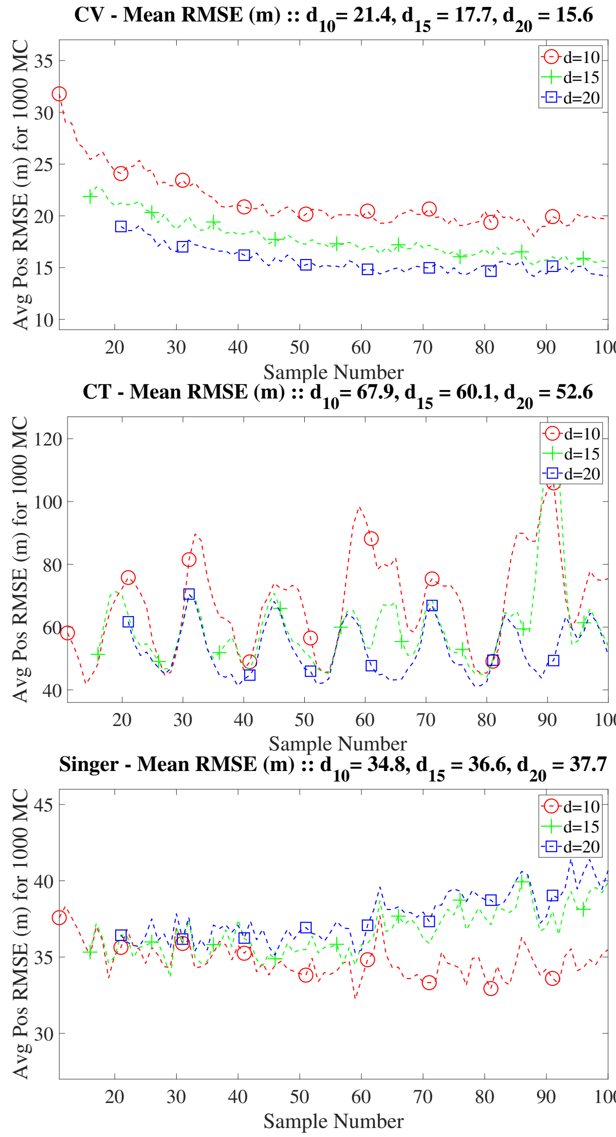

Fig. 3. **Effect of the training data on the RQ kernel**. This figure shows the results for the three scenarios, as explained in Fig.1. The three different values of the parameter are $d = 10, 15, 20$.

the originally proposed squared exponential kernel for the commonly observed point target tracking dynamics. Unlike the squared exponential and the rational quadratic kernels, the accuracy of the Matérn kernel improves consistently with the increase in the training data. Lastly, the robustness of the approach is demonstrated by assuming unknown noise variances. Current work is focused on theoretical studies of the impact of uncertainties on Gaussian process methods.

Acknowledgment

We acknowledge the support from the Govt. of Pakistan and the Dept. of ACSE (University of Sheffield, UK) for granting scholarships to the first author. We are grateful to the EPSRC for funding this work through EP/T013265/1 project NSF-EPSRC:ShiRAS. Towards Safe and Reliable Autonomy in Sensor Driven Systems. This work was also supported by the National Science Foundation under Grant NSF ECCS 1903466.

References

[1] W. Aftab and L. Mihaylova, "A Gaussian process regression approach for point target tracking," in *Proceedings of the 22nd International Conference on Information Fusion (FUSION)*. Ottawa, Canada: IEEE, 2019, pp. 1–8.

[2] I. Hwang, H. Balakrishnan, K. Roy, J. Shin, L. Guibas, and C. Tomlin, "Multiple-target tracking and identity management," in *SENSORS, 2003 IEEE*, vol. 1. Toronto, Canada: IEEE, 2003, pp. 36–41.

[3] K. Granström, A. Natale, P. Braca, G. Ludeno, and F. Serafino, "Gamma Gaussian inverse Wishart probability hypothesis density for extended target tracking using X-band marine radar data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 12, pp. 6617–6631, 2015.

[4] M.-C. Chuang, J.-N. Hwang, K. Williams, and R. Towler, "Tracking live fish from low-contrast and low-frame-rate stereo videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 1, pp. 167–179, 2014.

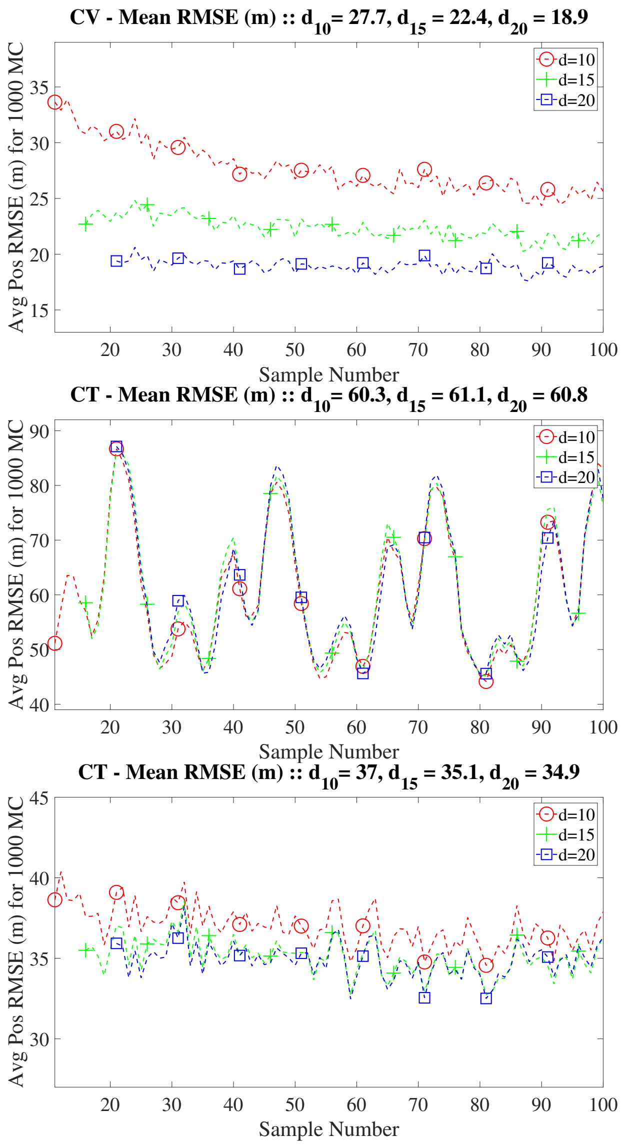

Fig. 4. **Effect of the training data on the Matérn with $\nu = \frac{3}{2}$ (M3) kernel**. This figure shows the results for the three scenarios, as explained in Fig.1. The three different values of the parameter are $d = 10, 15, 20$.

[5] J. Choi, S. Ulbrich, B. Lichte, and M. Maurer, "Multi-target tracking using a 3D-Lidar sensor for autonomous vehicles," in *Proceedings of the International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. The Hague, Netherlands: IEEE, 2013, pp. 881–886.

[6] M. Lázaro-Gredilla, S. Van Vaerenbergh, and N. D. Lawrence, "Overlapping mixtures of Gaussian processes for the data association problem," *Pattern Recognition*, vol. 45, no. 4, pp. 1386–1395, 2012.

[7] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, "Multiple extended object tracking using Gaussian processes," in *Proceedings of the 19th International Conference on Information Fusion (FUSION)*. Heidelberg, Germany: IEEE, 2016, pp. 868–875.

[8] W. Aftab, A. De Freitas, M. Arvaneh, and L. Mihaylova, "A Gaussian process convolution particle filter for multiple extended objects tracking with non-regular shapes," in *Proceedings of the 21st International Conference on Information Fusion (FUSION)*. Cambridge, UK: IEEE, 2018, pp. 1–8.

[9] W. Aftab, R. Hostettler, A. De Freitas, M. Arvaneh, and L. Mihaylova, "Spatio-temporal Gaussian process models for extended and group object tracking with irregular shapes," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2137–2151, 2019.

[10] L. Mihaylova, A. Y. Carmi, F. Septier, A. Gning, S. K. Pang, and S. Godsill, "Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking," *Digital Signal Processing*, vol. 25, pp. 1–16, 2014.

[11] K. Granstrom, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview and applications," *ISIF Journal of Advances in Information Fusion*, vol. 12, no. 2, pp. 139–174, 2017.

[12] N. Wahlström and E. Özkan, "Extended target tracking using Gaussian processes," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4165–4178, 2015.

[13] W. Aftab, A. De Freitas, M. Arvaneh, and L. Mihaylova, "A Gaussian process approach for extended object tracking with random shapes and for dealing with intractable likelihoods," in *Proceedings of the 22nd International Conference on Digital Signal Processing (DSP)*. IEEE, 2017, pp. 1–5.

[14] C. E. Rasmussen and C. K. Williams, *Gaussian processes for machine learning*. MIT press Cambridge, 2006, vol. 1.

[15] O. Stegle, S. V. Fallert, D. J. MacKay, and S. Brage, "Gaussian process robust regression for noisy heart rate data," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2143–2151, 2008.

[16] J. Hensman, A. Matthews, and Z. Ghahramani, "Scalable variational Gaussian process classification," 2015.

[17] A. Wilson and R. Adams, "Gaussian process kernels for pattern discovery and extrapolation," in *Proceedings of the International Conference on Machine Learning*, Atlanta, USA, 2013, pp. 1067–1075.

[18] O. Rudovic, M. Pantic, and I. Patras, "Coupled Gaussian processes for pose-invariant facial expression recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1357–1369, 2013.

[19] Z. Duan, C. Han, and X. R. Li, "Comments on " unbiased converted measurements for tracking"," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 4, p. 1374, 2004.

[20] S. V. Bordonaro, P. Willett, and Y. Bar-Shalom, "Unbiased tracking with converted measurements," in *Proceedings of the Radar Conf. (RADAR), 2012 IEEE*. IEEE, 2012, pp. 0741–0745.

[21] H.-M. Kim, B. K. Mallick, and C. Holmes, "Analyzing nonstationary spatial data using piecewise Gaussian processes," *Journal of the American Statistical Association*, vol. 100, no. 470, pp. 653–668, 2005.

[22] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

[23] R. A. Singer, "Estimating optimal tracking filter performance for manned maneuvering targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 6, no. 4, pp. 473–483, 1970.