



Full length article

A review of Bayes filters with machine learning techniques and their applications

Sukkeun Kim^a, Ivan Petrunin^a, Hyo-Sang Shin^{a,b,*}^a School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, MK43 0AL, UK^b Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon, 34141, Republic of Korea

ARTICLE INFO

MSC:

60J05

60J20

Keywords:

Bayes filter

Machine learning

Survey

Review

ABSTRACT

A Bayes filter is a widely used estimation algorithm, but it has inherent limitations. Performance can degrade when the dynamics are highly nonlinear or when the probability distribution of the state is unknown. To mitigate these issues, machine learning (ML) techniques have been incorporated into many Bayes filters, due to their advantage of being able to map between the input and the output without explicit instructions. In this review, we reviewed 90 papers that proposed the use of ML techniques with Bayes filters to improve estimation performance. This review provides an overview of Bayes filters with ML techniques, categorised according to the role of ML, remaining challenges and research gaps. In the concluding section of this review, we point out directions for future research.

1. Introduction

A Bayes filter, also known as recursive Bayesian estimation, is a powerful probabilistic recursive state estimator that utilises Bayes' theorem and Markov chain models. The Bayes' theorem enables Bayes filters to estimate the future state of the system based on prior probabilistic knowledge and given noisy measurements. In addition, the Markov property allows them to depend only on the present state rather than entire historical data. Despite the effectiveness of Bayes filters, several issues exist which are affecting their performance. For instance, the Kalman filter (KF) [1] requires accurate information on system dynamics and whiteness of error covariance, for its optimality. Performance degradation in Bayes filters may be caused by imperfect or nonlinear models and non-Gaussianity in the assumed error. To address these issues, many variations of Bayes filters such as extended Kalman filter (EKF) [2], unscented Kalman filter (UKF) [3], cubature Kalman filter (CKF) [4], adaptive KF (AKF) [5–8], sequential Monte Carlo (SMC) method or particle filter (PF) [9–13] etc. are proposed. However, some issues due to the highly nonlinear system dynamics or unknown probability distribution of the state still exist and degrade these filters.

Various machine learning (ML) techniques, such as artificial neural network (ANN) or simply neural network (NN) [14,15], including subtypes such as radial basis function neural network (RBFNN) [16] and recurrent neural network (RNN) including long short-term memory (LSTM) [17], are proposed. Beyond NNs, there are distinct methods such as reinforcement learning (RL) [18], Gaussian process (GP) [19],

kernel methods [20,21] and more. These are utilised in many research fields because of their capability for learning and inference based on data, without being explicitly instructed. Thanks to this advantage, researchers have proposed applying ML techniques to enhance Bayes filters and address their disadvantages as well. The heterogeneity of these approaches, however, makes it difficult to systematically assess current progress and formulate the research challenges related to the combination of Bayes filters and ML-based approaches.

There are few literature reviews of these hybrid approaches that provide an overview. For example, our previous work [22] included a review of KF with artificial intelligence (AI) techniques, but this is limited to KF and focused on the externally combined structure of KF and AI techniques. In this extension of our previous work, we explore PF and other types of Bayes filters, including those in kernel space. We focus not only on structures that combine them externally but also on internal combinations where ML operates within the filter loop. Feng et al. [23] conducted a review of the hybrid structures of KF and ANN, but this review is also limited only to KF and ANN. Additionally, this review [23] not only focused on utilising ANN to improve the performance of the KF but also on utilising KF for the ANN training. Jwo et al. [24] reviewed the application of ANN in navigation systems, specifically targeting the global positioning system (GPS) and inertial navigation system (INS). While this review offers valuable insights into the integration of ANN with navigation systems, its scope is limited to ANN and navigation applications. In addition, other reviews [25–27]

* Corresponding author at: School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, MK43 0AL, UK.

E-mail address: h.shin@cranfield.ac.uk (H.-S. Shin).

have introduced AI or learning-based solutions for PF degeneracy or impoverishment problems as a part of their scope, the primary focus of these reviews has been on PF, not the integration of Bayes filters and ML. Identifying these gaps prompted us to undertake a review, offering readers a comprehensive and systematic insight into approaches for incorporating ML into Bayes filters. This review, contrasting with previous studies, aims to investigate the approaches that leverage ML techniques, not only limited to ANN but including RL, GP, or kernel methods to enhance the estimation performance of Bayes filters.

This review focuses on approaches that: (1) aim to improve the state estimation with ML, not for the training of the ML and (2) employ Bayes filters, including KF families such as EKF and UKF, as well as SMC methods. The ML techniques under consideration in this review include supervised learning, unsupervised learning, reinforcement learning and probabilistic modelling. The general optimisation techniques like genetic algorithms (GA), particle swarm optimisation (PSO) and ant colony optimisation (ACO) are excluded since there are many arguments if those techniques can be included in ML and including them would expand the scope too broad. Within the defined scope, we carried out a literature survey and the bulk of the approaches were categorised into two primary groups: (1) internal and (2) external based on the position of the ML techniques, whether inside or outside of the estimation loop. Under these two main divisions, three respective subgroups exist, each classified according to the role of ML techniques. For these groups that have been used in different applications, we have summarised the characteristics of each approach. As a conclusion of the paper, we provided the classification of the approaches, suggestions for these approaches and ML techniques, remaining challenges, research gaps and future research direction in this study area. The main contributions of this paper are listed as follows:

1. We systematically reviewed 90 recent papers and introduced approaches that improve the performance of Bayes filters with ML techniques.
2. We categorised these approaches and provided a systematic overview of them, along with an analysis of the characteristics of each group.
3. The remaining challenges and gaps are discussed.

The rest of this paper is composed as follows: Section 2 briefly introduces the Bayes filters and ML techniques. Section 3 reviews the research of Bayes filters with ML techniques and categorises them. Section 4 discusses suggestions for approaches and ML techniques, remaining challenges, research gap and future research direction. Finally, we conclude this review in Section 5.

2. Bayes filter and machine learning techniques

This review is focused on approaches that aim to integrate ML techniques with Bayes filters. For clarity and a comprehensive understanding of the readers, let us briefly introduce these two principal techniques.

2.1. Bayes filter

A Bayes filter is a recursive estimator based on Bayes' theorem, which enables the filter to estimate the future state of a system using prior knowledge and noisy measurements. Additionally, given the Markov property, the probability of the current state $\mathbf{x}_k$ is conditionally independent of previous states, and the measurement $\mathbf{z}_k$ depends only on the current state, as denoted as follows:

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_0) = p(\mathbf{x}_k | \mathbf{x}_{k-1}), \quad (1)$$

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = p(\mathbf{z}_k | \mathbf{x}_k). \quad (2)$$

The recursive Bayesian estimation can be broken down into two steps: prediction and update steps. The prediction step calculates the

probability distribution of the predicted state using the Chapman–Kolmogorov equation to account for the transition probabilities, denoted as follows:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}, \quad (3)$$

where $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ represents the probability of the transition from time step $k-1$ to time step k under the assumed dynamic model and $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ denotes the posterior probability distribution of the previous time step. Subsequently, the update step computes the probability distribution of the updated state by incorporating the measurement likelihood using Bayes' theorem, denoted as follows:

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{\int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k}, \quad (4)$$

where $p(\mathbf{z}_k | \mathbf{x}_k)$ represents the measurement likelihood given the state, and the denominator serves as a normalising constant to ensure that the total probabilities sum to 1.

In this review, we consider hybrid approaches that combine KF families, such as EKF and UKF, or PF with ML techniques.

2.2. Machine learning techniques

ML is a popular computational methodology that is typically based on data and can infer patterns without explicit instructions. Techniques such as ANN, RNN, RL or GP are well-known examples. Since the majority of approaches reviewed here utilised ANNs or GPs, we will focus on providing a brief introduction to ANNs and GPs.

An ANN is one of the most commonly used ML techniques due to its ability to map inputs to outputs without a predefined model. An ANN consists of three parts: input, hidden and output layers. The outputs of the ANN are obtained through weights, which connect the input layer to the hidden layer, and the hidden layer to the output layer. These weights can be trained using backpropagation, a method associated with a backpropagation neural network (BPNN). A GP is another widely used ML technique due to its probabilistic nature, which provides uncertainty estimates along with predictions. This feature enables GPs to capture complex patterns in regression [28, Chapter 2.2].

Note that, the ML techniques under consideration in this review are supervised learning, unsupervised learning, reinforcement learning and probabilistic modelling. We explicitly exclude more general optimisation techniques like GA, PSO and ACO as we discussed in the introduction.

3. Bayes filters with machine learning techniques

3.1. An overview

The papers were identified using keywords such as Bayes filters (e.g., KF, EKF, UKF, PF) and ML techniques (e.g., ANN, RL, kernel, GP). From the papers found, some were included or excluded based on the following criteria:

1. Include if the approach utilises ML to address the shortcomings of Bayes filters in the estimation problem.
2. Exclude if the approach aims to train ML using the Bayes filters.
3. Exclude if Bayes filters are only used to post-process the estimation results from ML.

Based on the aforementioned inclusion and exclusion criteria, we reviewed a total of 90 papers, which were then categorised into multiple groups according to the role of ML. These approaches fall into two primary groups, determined by the position of ML techniques: (1) internal and (2) external. A conceptual illustration of these two primary groups is shown in Fig. 1.

A total of 51 papers (56.7%) are categorised as internal, where ML is positioned within and influences the performance of the Bayes filter

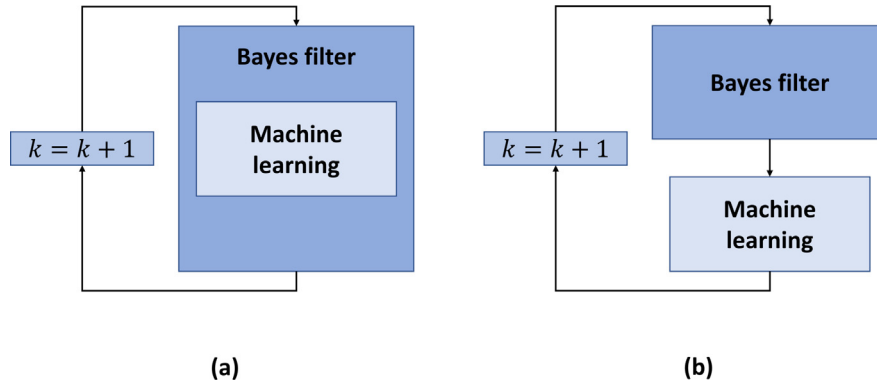


Fig. 1. A conceptual illustration of the two primary groups: (a) internal and (b) external.

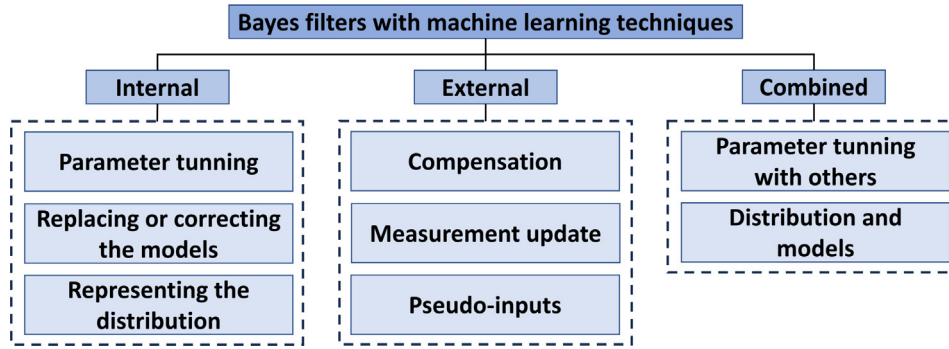


Fig. 2. A graphical classification of Bayes filters with machine learning techniques.

loop. Conversely, 27 papers (30.0%) are categorised as external where ML is positioned outside the filtering loop and affects the result. Each of these primary groups has three subgroups underneath. Lastly, the remaining 12 papers (13.3%) are categorised as combined, representing a combination of the two approaches from previous groups. A graphical classification of these groups of hybrid approaches is shown in Fig. 2.

3.2. Internal approaches

The first of the two primary groups is internal, including the approaches that utilise ML within the filtering loop.

3.2.1. Parameter tuning

The first subgroup of the internal approaches is parameter tuning. The approaches in this subgroup use ML to tune the parameter of Bayes filters such as process noise covariance Q_k , measurement noise covariance R_k or both, as well as parameters for sigma points in the case of UKF, etc. Traditionally, these parameters are assumed to be constant, which can degrade the performance of the filter. However, in dynamic situations where the mathematical model may change or the noise characteristics of the measurement sensor are affected by external disturbance, this approach can enhance the performance. Most approaches in this subgroup aim to tune the Q_k or R_k in Bayes filters. In these approaches, ML techniques, particularly ANNs, are utilised to predict these parameters based on the measurement y_k or filter innovations. It is important to note that this approach is similar to adaptive methods for adjusting covariances, but it relies on ML trained with previously collected data. The primary distinction between ML-based parameter tuning and traditional adaptive methods is that the ML leverages training data from dynamic environments. This enables the system to update the covariance matrices based on nonlinear relationships based on training data. A conceptual architecture of the algorithms is described in Fig. 3.

The majority of approaches in this group focus on tuning Q_k , R_k or both, to prevent filter degradation, a method similar to the AKF. Early studies in this group can be seen as extensions of the AKF, using ML techniques, particularly ANNs, as a replacement for conventional adaptive noise covariance selection methods. Jwo and Huang [29] proposed an NN-aided adaptive EKF for differential GPS applications. In their approach, an ANN is used to predict Q_k and R_k based on the innovation from the update step of the EKF. The authors later extended this approach by incorporating a fuzzy NN for GPS navigation in [30] and further with PSO-RBFNN instead of fuzzy NN for GPS and INS integration in [31]. Other researchers have proposed similar methods. Wu and Min [32] introduced an approach that uses an ANN to predict the adaptive regulator factors S_k^Q and S_k^R . These regulator factors are multiplied by the original covariance matrices to adjust them as follows:

$$Q_k^{new} = S_k^Q Q_k^{old}, R_k^{new} = S_k^R R_k^{old}. \quad (5)$$

Lee and Jekeli [33] proposed a method similar to the previous one that utilises an ANN with adaptive UKF for the geolocation of unexploded ordnance using an inertial measurement unit (IMU) and GPS. This approach used an ANN based on measurements to predict Q_k of the adaptive UKF and combined it with the Rauch-Tung-Striebel (RTS) smoother for the post-estimation smoothing. Shi et al. [34] utilised a fuzzy NN to predict the weighting parameter α_w of the weighting KF, which tunes not only Q_k and R_k but also the covariance $P_{k|k}$ for a GPS/INS integrated navigation application. Xia et al. [35] proposed a method that uses an RNN to predict not only Q_k and R_k but also the signal autoregression (AR) parameter in the transition matrix of the KF for a speech enhancement application. Bekhtaoui et al. [36] proposed an approach that uses a Q-learning to switch Q_k between two predefined values for a target tracking problem using KF. Similarly, Xiong et al. [37] proposed the use of a Q-learning to select a set of Q_k and R_k based on measurements in EKF for a spacecraft autonomous navigation application. Brossard et al. [38] proposed the use of an ANN to predict

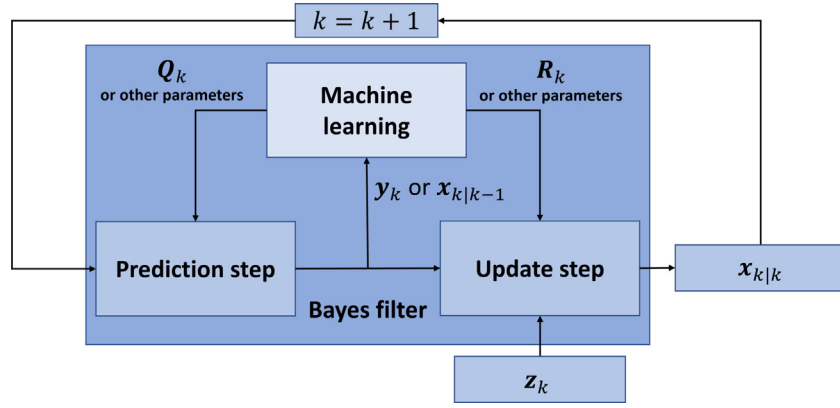


Fig. 3. A conceptual architecture of parameter tuning.

Q_k and R_k of the invariant EKF based on measurement for IMU dead-reckoning situations. Ullah et al. [39,40] proposed approaches that utilise an ANN to predict R_k of KF when the sensor is affected by dynamic conditions such as weather, with applications in temperature estimation and indoor environment prediction. Gao et al. [41] proposed the use of an RL to predict R_k for AKF ground vehicle navigation. Jiang and Nong [42] proposed the use of a group of ANNs to predict not only Q_k and R_k but also the control input of the filter and mean error value of the measurement in KF, based on an RNN structure for aviation surveillance. Jouaber et al. [43] proposed an approach that uses an ANN and an RNN to predict the parameter to tune the R_k of the AKF with innovation for a target tracking application. Wu et al. [44] utilised a multitask temporal convolution network (TCN) to predict Q_k and R_k of KF in a GPS/INS integrated navigation application.

Other approaches that utilise ML techniques to tune parameters other than noise covariance have also been proposed. Xinyuan et al. [45] employed an RBFNN to predict the forgetting parameter, which serves as a multiplier for the Kalman gain. This allows the filter to prioritise newer data by diminishing the influence of previous data in a central divided-central KF for satellite attitude determination. Turner and Rasmussen [46] used a GP-based optimiser to predict the key parameters of UKF, denoted as $\theta = \{\alpha, \beta, \kappa\}$, for sigma points selection. Motwani et al. [47] used an ANN to predict the weight factor to calculate the single point-valued estimate of the interval KF, instead of the weighted average in surface vehicle navigation. Finally, Revach et al. [48] utilised an RNN, specifically gated recurrent unit (GRU), to directly predict Kalman gain based on innovation and the posterior state from the previous time step for EKF and named it KalmanNet.

Most approaches within this group aim to enhance the robustness of the filter. They focus on tuning parameters in cases where the environment influences measurements or where nonlinearity might introduce errors. The common advantages of the approaches in this group include:

- **Robustness in dynamic environments:** The approaches in this group are particularly robust in dynamic environments by tuning the parameters when environments change.
- **Flexibility and applicability:** Straightforward frameworks of the approaches in this group enable them to be applied across various filter types and applications.
- **Ease of combination:** Due to the simple structure, the approaches in this group can be easily integrated with other methods to further improve performance, a topic to be considered in Section 3.4.

Despite these advantages, this group has limitations and disadvantages, including:

- **Similarity to conventional methods:** Most approaches are similar to traditional AKF. Consequently, they might underperform AKF, if the disparity between the environment or application from the training setup is huge.
- **Nonlinearity and noise limitations:** These approaches typically do not address the issues associated with nonlinearity and are limited to additive Gaussian noise.

To mitigate these challenges, other groups such as representing the distribution and substitution of the model, can be explored. These are discussed in subsequent sections.

3.2.2. Replacing or correcting the models

The second subgroup of the internal approaches is replacing or correcting the models. The approaches in this subgroup can be further divided into two categories: (1) using ML to replace the transition (or time propagation) model, the measurement model or both, and (2) using ML to correct the transition model. When the system, including the measurement model, is highly nonlinear or even unknown, the performance of the filter can degrade. To address this, many approaches have been proposed to replace dynamic or measurement models using ML, particularly GPs or ANNs. Additionally, some methods have been proposed to predict the error between the true transition model and known mathematical transition model using ML and correct them when the model is imperfect. A conceptual architecture of the algorithms is described in Fig. 4.

Most of the approaches in this group focus on replacing the transition model, measurement model or both using ML techniques. Vlassis et al. [49] proposed a nearest neighbour conditional density estimation-based nonparametric measurement model for an auxiliary particle filter (APF) [13] to address the robot kidnapping problem. This approach demonstrated enhancements both when the measurement model is high-dimensional and when it is unknown. Wang et al. [50] proposed the use of an RBFNN to construct the process model and propagate the samples of PF for target tracking with a sensor network. In this approach, the RBFNN is used as a transition function to propagate the previous samples to the current time step. Michel and Heiries [51] proposed an algorithm that uses a support vector regression (SVR) as a substitution for the transition model in a sigma point KF for battery state of health (SoH) and state of charge (SoC) estimation application, where these states are highly nonlinear. Wang [52] proposed a similar approach, using an ANN as a substitute for the transition model in the extreme learning KF for short-term wind speed prediction. Similarly, Zhang and Shin proposed an approach using an ANN as a substitute for the transition model in a Gaussian mixture filter [53].

GPs are frequently utilised in this group because of their inherent ability to perform function regression, where the regression is represented in functional form. Ko and Fox [54] proposed the use of a GP to substitute both the process and measurement models of Bayes

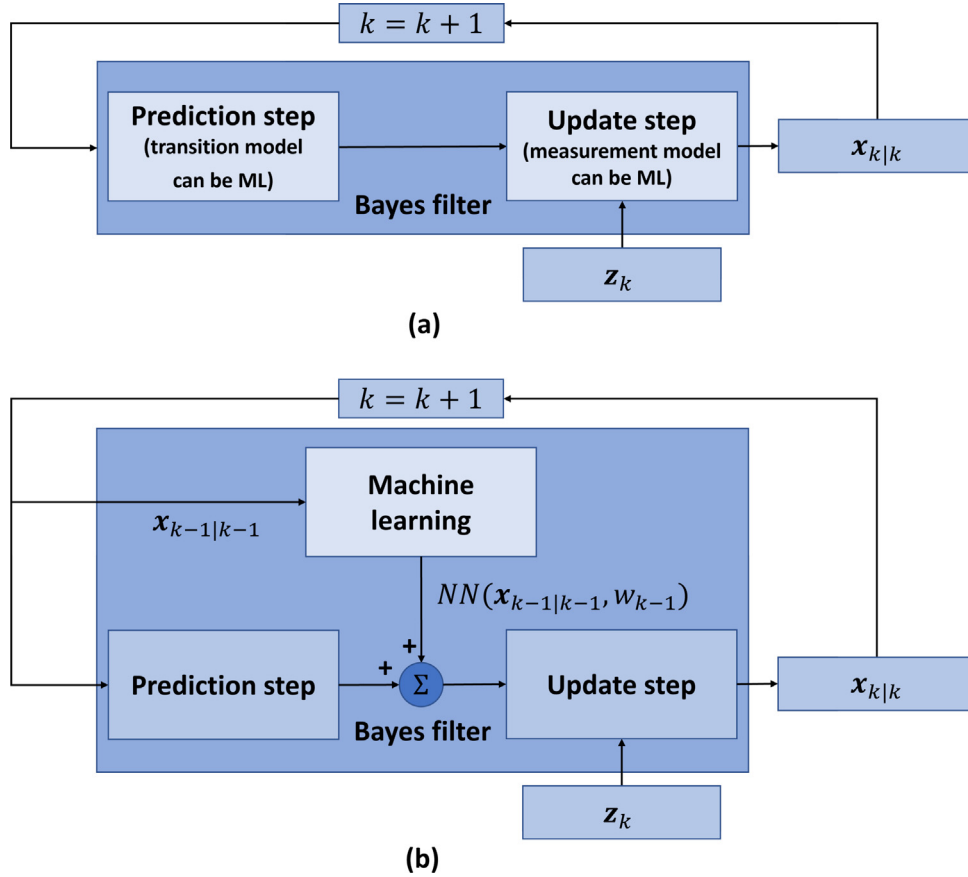


Fig. 4. A conceptual architecture of (a) replacing or (b) correcting the models.

filters. Furthermore, by employing enhanced GP models that combine GP with a parametric model, they achieved further improvements. The original idea in this study, combining a GP with a UKF to replace the transition model of UKF, denoted as GP-UKF, will be discussed further in Section 3.4.2. Deisenroth et al. [55] proposed a method similar to that of Ko and Fox, using a GP to replace both models in an assumed density filtering (ADF). This filter is consistent and moment-preserving in test cases and showed improvements compared to Ko's GP-UKF. Neerukatti et al. [56] also proposed a similar approach, using a GP-based prognosis model as a substitute for the transition model in a PF for real-time damage prediction application. Sun et al. [57] proposed a method that uses a GP with a PF for multiple-model (MM) tracking problems, instead of conventional interacting MM (IMM). A notable difference in Sun et al.'s approach compared to previous ones is that it considered a mixture of models, such as the constant velocity (CV) and the coordinated turn (CT) models. Ferris et al. [58] proposed a distinct approach, employing a GP to extract the likelihood of Wi-Fi signal strength for use in a PF-based localisation problem. This approach demonstrated that a GP can be utilised to predict the likelihood of localisation and replace the model.

The remaining approaches in this group employ ML to correct errors during time propagation. Stubberud et al. [59] proposed an approach using an ANN to predict the discrepancy between the true and known mathematical models, allowing for correction before the update step of the EKF. The model correcting step of this approach can be simply denoted as follows:

$$\mathbf{x}_{k|k-1} = \hat{\mathbf{f}}(\mathbf{x}_{k-1|k-1}) + NN(\mathbf{x}_{k-1|k-1}, \mathbf{w}_{k-1}), \quad (6)$$

where $\hat{\mathbf{f}}(\cdot)$ is the known mathematical transition model, $NN(\cdot, \cdot)$ is the prediction from the ANN, which is derived to correct the discrepancy between the true and known mathematical models and $\mathbf{w}$ is the weight

of the ANN. Other approaches aimed at correcting errors in the transition model follow a similar structure. Zhan and Wan [60] proposed an ANN-aided AUKF, where the error between the true and known mathematical models is predicted by an ANN and then corrected in the predicted state before the update step. Kramer and Stubberud [61] proposed a neural EKF (NEKF) where the ANN is utilised for error prediction to correct the transition model of the MTT problem. This approach is widely appreciated in simultaneous localisation and mapping (SLAM) problems. Choi et al. [62] applied this hybrid method, combining an ANN and an EKF for SLAM, while Kotov et al. [63] applied an RNN in place of a feedforward NN and Miljković et al. [64] adopted this approach for monocular SLAM in indoor environment. Miao et al. [65] demonstrated further improvements with a variational Bayesian adaptive cubature KF by utilising an ANN to predict the error between the true and known models in bearing-only tracking problems. Pesce et al. [66] used an RBFNN to predict the error between models in an AEKF for spacecraft relative navigation.

The approaches proposed in this group are designed to address gaps in understanding or inaccuracy in the transition or measurement model using machine learning. The common advantages of the approaches in this group include:

- **Unknown or inaccurate models:** The approaches in this group can be beneficial with data when the transition or measurement model is unknown or inaccurate.
- **High-dimensional scenarios:** The approaches in this group can be advantageous in scenarios involving high-dimensional states or measurements.
- **Robustness with GP:** The use of GP provides some robustness to the approaches in this group, mainly due to its probabilistic nature, reflecting variance in predictions.

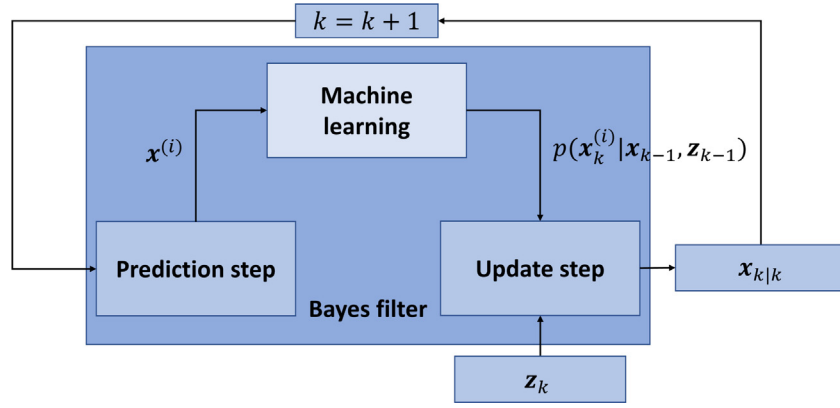


Fig. 5. A conceptual architecture of representing the distribution.

Despite these advantages, this group has limitations and disadvantages, including:

- **Nonlinearity limitation:** The approaches in this group might not be able to replace the highly nonlinear dynamics due to limited training data.
- **Rapid motion dynamics:** The approaches in this group struggle with rapid changes in motion dynamics, such as stunt manoeuvres of multirotor.

3.2.3. Representing the distribution

The last subgroup of the internal approaches is representing the state probability distribution with ML, such as GPs or GMMs. The approaches in this subgroup use ML to represent the posterior probability distribution of the state. When the system is highly nonlinear or when the posterior probability distribution of the state is even multimodal, it is hard to estimate the state using the EKF or UKF. In these cases, the SMC method, or PF, is widely applied to estimate the state. However, PF also has issues or limitations in either computation burden or discretisation problems. Furthermore, the PF can be degraded when sample degeneracy or impoverishment issues arise. Many approaches have been proposed to overcome these issues using various resampling methods, but PF still suffers from computational burden (see [25–27] for details). To alleviate these issues, and eliminate the resampling step as well in some approaches, many approaches utilise ML, particularly GPs or kernel methods, where these ML techniques represent the posterior probability distribution of the state in PF. In addition, some approaches have proposed the use of ML techniques to construct the Gaussian mixture model (GMM) to represent the posterior probability distribution of the state. A conceptual architecture of the algorithms is described in Fig. 5.

Many approaches which can be considered similar to the resampling in PF have been proposed. Qin and Peng [67] proposed an approach that utilises a general regression NN (GRNN) to adjust the position of the samples drawn from the prior density in PF and provide an approximation of the importance density. Similarly, Pang et al. [68,69] proposed using an ANN to adjust the low-weight particles in the tail area to increase their weights and split the particles with large weights into two particles in PF. Both algorithms have demonstrated improved performance of the PF in tracking problems. Xiaojing et al. [70] utilised a GRNN to adjust the particles, bringing them closer to the posterior probability density in PF. Imbiriba and Closas [71] introduced another approach which utilises a GP to calculate the minimum mean squared error (MMSE) and maximum a posteriori (MAP) for the filter and, additionally, to derive the approximated posterior probability distribution to draw new samples. Comandur et al. [72] proposed a method of using an invertible particle flow to approximate the proposal distribution for the importance sampling and combined it with Gaussian PF. This

approach also maps the particles to a new point to bring them closer to the distribution which is similar to [70], but with particle flow.

As previously mentioned, many approaches have utilised ML techniques to construct the GMM, assuming the posterior probability distribution as a mixture of multiple Gaussian distributions. Kotecha and Djurić [73] proposed a Gaussian sum particle filtering based on the Gaussian particle filter (GPF) [12] where the posterior probability density is assumed to be GMM constructed by the expectation-maximisation (EM) algorithm. Psiaki [74] proposed a similar filter with a GMM constructed by the EM algorithm but incorporated an additional resampling method to constrain the covariance of the mixture and economise the number of mixtures simultaneously. Raihan and Chakravorty [75,76] proposed particle Gaussian mixture (PGM) filters, which allow the algorithm to modify the number of the mixtures during the prediction step, update step or both. This modification uses a K-means cluster-based Gaussian mixture weight calculation and can adjust the number of mixtures, given a predetermined maximum number of the mixture. Furthermore, Yun and Zanetti [77] introduced two new clustering algorithms based on the fuzzy C-means (FCM) clustering algorithm for the GMM filters. The algorithms, based on FCM with regularisation by the Kullback–Leibler information (KLFCM), utilise a fuzzifier for soft clustering based on membership. This demonstrated an improvement over the limitations of the K-means cluster and the EM algorithm in the context of particle clustering and estimation.

Last but not least, some approaches utilise kernel methods to represent the probability distribution in the reproducing kernel Hilbert space (RKHS) [20]. Sun et al. [78,79] proposed the adaptive kernel Kalman filter (AKKF) to approximate the probability distribution using kernel mean embedding (KME) [21] for tracking and sensor fusion problems. KME is a powerful kernel-based method that utilises the kernel trick for representing distribution in RKHS and is applied to other estimation techniques. They will be discussed more in detail later in Section 3.4.2.

The approaches within this group are beneficial because of their ability to approximate closer to the real posterior probability with greater accuracy, leading to enhanced performance. The common advantages of the approaches in this group include:

- **Enhanced posterior approximation:** The approaches in this group can improve performance by using samples that are closer to the real distribution or more accurately approximate the posterior probability distribution.
- **Highly complex systems:** The approaches in this group demonstrate significant improvements in high-dimensional, nonlinear or even multimodal systems.

Despite these advantages, this group has limitations and disadvantages, including:

- **Computational complexity:** One key challenge with the approaches in this group is their high computational cost.

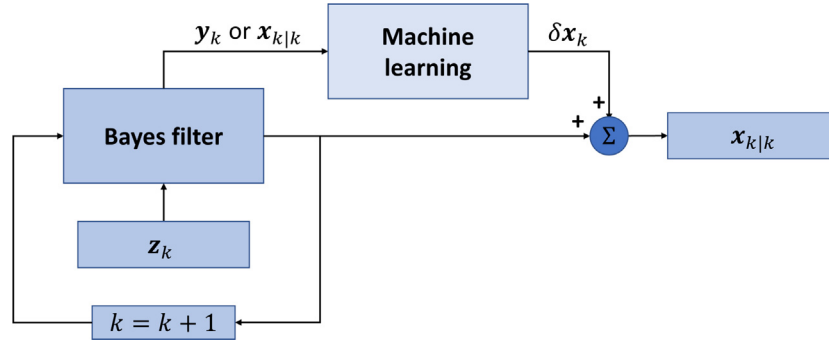


Fig. 6. A conceptual architecture of compensation.

- **Practical issues with GP or GMM:** While the approaches in this group are promising, practical challenges, such as model selection in GP or clustering in GMM, can lead to performance degradation.

It is worth considering the data acquisition and training of the ML approaches utilised in the internal approaches. The approaches, parameter tuning and replacing or correcting models, utilised ANNs or GPs to enhance the performance of the Bayes filters, which requires the training data. Most of the approaches in the parameter tuning group considered offline training with true data acquired from hardware, such as GPS signals. The ANNs were trained with this true data and hand-tuned parameters so that they could produce adequate parameters for filters with a given input, such as current state or innovation. In the case of replacing or correcting the models group, true data are also acquired from hardware or test experiments to train ANNs or GPs offline. The supervised learning algorithms were trained with this true data to produce the time-propagated state or error in the propagated state. The representing the distribution group utilised unsupervised learning techniques such as GMM or kernel methods which have advantages in that they usually do not require the training data.

3.3. External approaches

The second of the two primary groups is external, including the approaches that utilise ML outside the filtering loop.

3.3.1. Compensation

The first subgroup of the external approaches is compensation. The approaches in this subgroup use ML to predict the error in the estimated state of Bayes filters and to compensate for the error in the estimation result. Due to various reasons, such as nonlinearity, non-Gaussian noise or imperfect models, the filters' performance can degrade, leading to the estimation results deviating from the actual state. To address this issue, many strategies have been proposed employing ML techniques, particularly ANNs, to predict errors in the estimation. These methods then adjust the results to bring the estimates closer to the true state. It is worth noting that many of these approaches can be viewed as post-processing, as in many proposed cases, the compensated state is not used recursively. A conceptual architecture of the algorithms is described in Fig. 6.

Chin [80] proposed a method using an ANN to predict the error in the estimated value and compensate for the KF's estimated state for an MTT problem. This approach outperformed the conventional joint probabilistic data association (JPDA) filter. Vaidehi et al. [81] built upon the previous ANN-aided KF MTT approach, adapting it for a three-dimensional, two-state problem involving highly manoeuvring multiple targets. Similarly, Yu et al. [82] used an ANN to predict and correct estimation errors from the UKF, applying this method to a distributed acoustic sensor network for MTT. Li et al. [83] introduced an approach using an ANN to estimate the error and combined it with a federated KF for enhanced robustness in terrain-integrated navigation systems.

Gao et al. [84] developed an approach using an RBFNN to predict and compensate for errors in the KF, which can be beneficial when the system exhibits coloured noise in laser welding applications. Evidence indicates that this method remains robust even with changes in the noise covariances Q_k and R_k . Zhang et al. [85] extended this approach for PFs, focusing on automatic gap tracking during welding. Nguyen et al. [86] designed a method using an ANN to forecast the unmodelled errors in the EKF, aiming to boost robot accuracy. The proposed algorithm was particularly effective at managing unmodelled errors arising from the robot's highly nonlinear dynamics. Sun et al. [87] proposed a method using an Elman NN to predict errors in the IMM KF for manoeuvring target tracking applications. The feedback-based architecture of the Elman NN demonstrated enhanced performance compared to the RBFNN-aided IMM KF. Zhao et al. [88] employed this architecture in conjunction with an ANN and an AKF for battery SoC estimation problems. In a related effort, Song et al. [89] applied a similar structure for velocity and attitude matching during the initialisation and calibration of INS, leveraging multiple KFs that were corrected using an ANN to establish a federated KF. Hosseinyalamdary [90] utilised an RNN to compensate for IMU errors during GNSS/IMU integration using a KF. Similarly, Li et al. [91] proposed an RNN-augmented UKF to correct random drift errors in IMUs. Chen et al. [92] presented a RBF and an extreme learning machine (ELM) NN-assisted AUKF to compensate for filter estimation errors for AUV navigation, while Tang et al. [93] incorporated an RL to correct attitude estimation discrepancies.

The biggest advantage of this group is the simple architecture that allows it to be applied to various applications. The common advantages of the approaches in this group include:

- **Flexibility and applicability:** The straightforward frameworks of the approaches in this group can be applied across various filter types and applications, allowing for broader applicability.

Despite these advantages, this group has an disadvantage:

- **Estimation divergence:** In situations where the estimation error becomes excessively large, some approaches in this group might encounter divergence because many methods do not utilise the compensated estimate as the prior state in the next time step.

3.3.2. Measurement update

The second subgroup of the external approaches is measurement update. The approaches in this subgroup use ML to predict the error in the current measurement and update it based on the estimation from the Bayes filters. Once the error in the measurement accumulates, the Bayes filters can diverge. For example, when the INS is used for the localisation or navigation standalone in cases such as a GPS outage, the accumulated error, as known as drift error, can cause the system to diverge, a well-known issue. To address this, many approaches have been proposed to update the measurement with the error predicted using ML techniques, particularly ANNs. Most of the approaches in this group

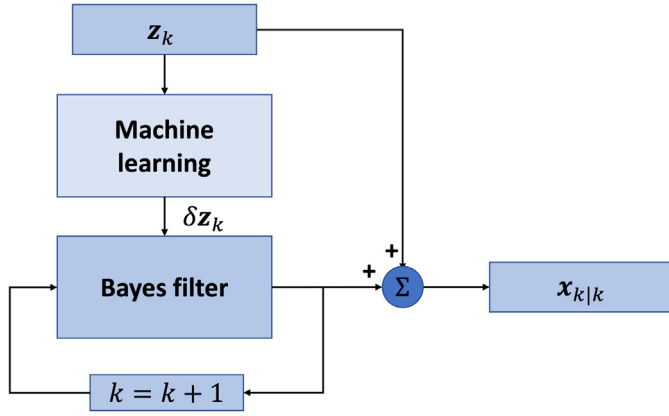


Fig. 7. A conceptual architecture of measurement update.

are applied in sensor fusion, especially in GPS/INS integration for GPS outage issues, with error-state KF (ESKF). A conceptual architecture of the algorithms is described in Fig. 7.

The approaches in this group are straightforward. For example, in the case of GPS/INS integration, position data from the GPS is assumed to be accurate, and the INS is the sensor used during GPS outages. The ANN predicts the error between the true data from GPS and the INS data in the event of GPS outages. This error is then used as a measurement in the ESKF. The ESKF estimation is then added directly to the INS data to update the measurement. The approaches in this group are similar to each other but are applied with different sensors or networks. Wang et al. [94] utilised an ANN to predict the error in the INS and update the measurement of the EKF for the GPS/INS integrated navigation problem when the GPS is unavailable. Zhang and Xu [95] combined an ANN for measurement updates and wavelet multi-resolution analysis (WMRA) for de-noising with the KF for GPS/strap-down INS (SINS)/magnetometer integrated navigation during GPS outages. Chen and Fang [96] used an RBFNN for GPS/INS integrated position and orientation estimation of aircraft during GPS outages. Jingsen et al. [97] proposed a similar approach but used two ELMs to reduce the prediction error, which can be affected by acceleration, for GPS/SINS integration during GPS outages. Zhang et al. [98] introduced an approach that uses an RNN to predict the error in measurements. This error can be directly added to the measurement and used to update the EKF for AUV navigation. Shaikat et al. [99] utilised an ANN to predict measurement errors and integrate multiple sensors, such as IMU, underwater acoustic positioning, Doppler velocity log, etc., for underwater vehicle localisation with an indirect KF during GPS outages. This study was further extended to multi-sensor fusion with a RBFN and an ESKF for underwater vehicle localisation during GPS outages [100]. Tang et al. [101] proposed the use of a gated recurrent unit (GRU) for measurement updates with an AKF for GPS/INS integration during GPS outages. Taghizadeh and Safabakhsh [102] proposed an attention-based convolutional LSTM approach for measurement updates with EKF in GPS/INS integration during GPS outages.

The common advantages of the approaches in this group include:

- **Reduction of filter divergence:** The approaches in this group can diminish the divergence of the filter by enhancing the measurements with less noise.

Despite these advantages, this group has an disadvantage:

- **Application specificity:** The approaches in this group are limited to specific applications.

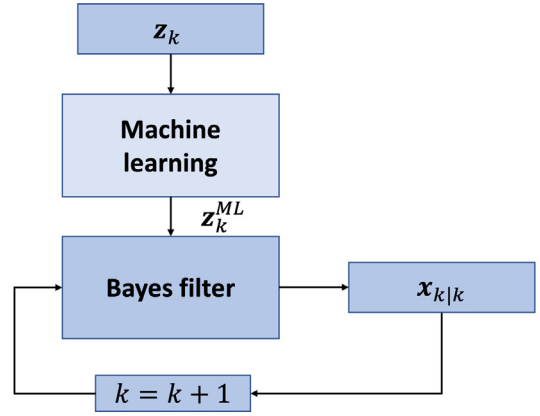


Fig. 8. A conceptual architecture of pseudo-input.

3.3.3. Pseudo-inputs

The last subgroup of external approaches is pseudo-inputs. The approaches in this subgroup use ML to produce pseudo-inputs before they are fed into the Bayes filters. When the dynamics model is highly nonlinear, or the measurement contains errors, the performance of the filter can degrade or, in the worst case, diverge. To alleviate this issue, some approaches propose using ML, particularly ANNs, to preprocess the estimated state from the previous time step or the measurement. A conceptual architecture of the algorithms is described in Fig. 8.

Vargas-Meléndez et al. [103] proposed the use of an ANN to produce a pseudo-roll angle of a ground vehicle, which is then used as a measurement in the KF for roll angle estimation. This approach enables the application of a linear KF to a highly nonlinear system, such as the roll angle of vehicle dynamics. Similarly, Novi et al. [104] employed an ANN to produce a pseudo-sideslip angle of a ground vehicle, which is subsequently used as a measurement in the UKF for sideslip angle estimation. Additional hybrid approaches for estimating the vehicle sideslip angle will be discussed in Section 3.4.1. Furthermore, Zhong et al. [105] integrated an ANN with a KF for water level forecasting, and Yu et al. [106] combined a DNN with a KF for a speech enhancement application.

The common advantages of the approaches in this group include:

- **Reduction of nonlinearity:** By pre-processing measurements using ML techniques, the approaches in this group can mitigate some nonlinearity, enhancing the overall estimation performance.
- **Cost-effective performance boost:** The approaches in this group allow the improved performance with more economical sensors.

Despite these advantages, this group has an disadvantage:

- **Data preparation:** The preparation of the training data for the approaches in this group can be challenging given that training data is customised for specific systems and sensors. The approaches in this group may also encounter limitations when used with different configurations.

As discussed in the previous section, it is important to consider how data is acquired and how machine learning models are trained in external approaches as well. In the external approaches, supervised learning techniques such as ANNs are commonly used. The compensation and the pseudo-inputs groups mostly utilised ANNs, acquiring real-world data through various sensors and training the network offline. The output of the original filter is used as a training input and the error from the true data is used as training output in the compensation group. Similarly, the measurement from the sensor is used as a training input and the desired de-noised measurement is used as a training output in the pseudo-input group. The measurement update approaches utilised

either ANNs or RNNs, acquiring real-world data primarily from GPS and INS, since most applications in these approaches involve GPS and INS integration, and training the network online. The INS input is used as a training input and the difference between GPS and INS is used as a training output.

3.4. Combined approaches

Some approaches have suggested using multiple ML techniques for various purposes to enhance the performance of estimation. We reviewed those methods that are the combinations of the previously introduced subgroups.

3.4.1. Parameter tuning with others

The parameter tuning approach is one of the most commonly combined methods with other approaches. The rationale is that parameter tuning, similar to AKF, enhances performance in dynamic situations and applies to many applications. Most algorithms in this group use ML techniques, particularly ANNs, to predict the noise covariances Q_k and R_k , along with other previously mentioned approaches.

The most common combination is between parameter tuning and replacing the models. Ko et al. [107] proposed GP-UKF, which utilises GP to replace the transition and measurement models in tracking an autonomous micro-blimp problem with UKF. The means of GPs replace the output of the transition and measurement models, while the covariances of GPs are used for the noise covariance Q_k and R_k . Ramirez et al. applied GP-UKF to the AUV position estimation problem, particularly when GPS is also occasionally unavailable [108]. Wu et al. [109] proposed an approach that uses an ANN to express the nonlinear measurement model in the master-slave (MS)-AUKF where the master UKF estimates the battery SoC and the slave UKF estimates Q_k and R_k . Jiang and Nong [110] utilised multiple ANNs to replace the transition model for time propagation in KF and multiple generalised regression NNs (GRNN) to estimate the mean error and error covariance of the transition and measurement models, respectively, for aerial surveillance. Song et al. [111] proposed a similar approach with three LSTMs to replace the transition model, predict the process noise covariance Q_k and predict the measurement noise covariance R_k for the MTT problem with KF.

A combination of parameter tuning and measurement updates is also found. Bitar and Gavrilov [112] proposed the use of a nonlinear autoregressive NN with external inputs (NARX) for GPS/INS integration problems with UKF. In this approach, NARX estimates the INS error and updates the UKF measurement during GPS outages. The combination of parameter tuning and pseudo-inputs is also found. Kim et al. [113] proposed the use of a deep ensemble network, comprising multiple LSTMs, to produce the pseudo-sideslip angle of a ground vehicle and process noise covariance Q_k for either EKF or UKF. This approach is employed for sideslip angle estimation using a simple kinematics model. Kim et al. [114] further extended this approach by incorporating an additional sensor filtering network that de-noises and de-biases the measurement using multiple LSTMs for UKF.

The common advantages of the approaches in this group include:

- **Synergy of approaches:** The approaches in this group could be beneficial by combining two existing approaches so that one can appreciate the advantages of both groups.
- **Adapting to dynamic environments:** By using ML to predict the noise covariance, the performance of the approaches in this group can be improved more significantly when the environment or sensor changes dynamically.
- **Versatility of application:** The approaches in this group can be combined with many other groups and are easily applicable to various applications.

Despite these advantages, this group has an disadvantage:

- **Data preparation:** Data preparation for the approaches in this group can be more complicated since it involves the combination of multiple approaches.

3.4.2. Distribution and models

The approaches that utilise ML techniques, particularly kernel methods, to represent the posterior probability distribution and to replace the models at the same time are found. The strategies in this group are based on the kernel method (or kernel trick) and Bayes' theorem. The state and distribution can be embedded into a higher-dimension feature space (RKHS [20]). In this RKHS space, some nonlinear operations can be linearly derived using KME [21]. Also, by applying Bayes' rule or KF in the KME, one can estimate the states without models or a known posterior probability distribution. Leveraging this characteristic, Fukumizu et al. [115] proposed a non-parametric, kernel-based realisation of Bayes' rule, termed the kernel Bayes' rule (KBR). In this method, the prior and posterior probabilities can be expressed with RKHS functions, even when the transition model is unknown. Based on KBR, Kanagawa et al. [116,117] proposed PF for scenarios where the measurement model (likelihood) is unknown. The measurement model is learned from training data, and the particles are drawn from the posterior distribution constructed with KME. Gebhardt [118] combined KBR with the concept of KF to develop a method called the kernel Kalman rule (KKR), which is computationally more efficient and numerically more stable than KBR.

The common advantages of the approaches in this group include:

- **Combination of the kernel and Bayes:** The approaches in this group leverage the kernel trick with Bayes' rule which can enhance the estimation performance in higher dimensions.
- **Model-free estimation:** The approaches in this group can estimate the state without relying on models or a known posterior probability distribution.

Despite these advantages, this group has an disadvantage:

- **Dependence on kernel function:** The effectiveness of the approaches in this group is closely tied to the kernel function, emphasising the importance of selecting an appropriate kernel and its parameter based on the dimension of the state and the system characteristics.

4. Discussion

We reviewed and categorised the approaches integrating ML techniques to improve the estimation performance of the Bayes filters. Despite the improvements shown by all approaches compared to using Bayes filters alone, there are some points that need further study. In this section, we first summarise the review and discuss the suggestions for approaches and ML techniques by application, remaining challenges, research gaps and proposals for future research direction.

4.1. Summary of the review

We reviewed and categorised approaches combining Bayes filters and ML techniques, and pointed out the common characteristics, advantages and disadvantages of these approaches according to each group in the previous section. A summary of the groups, along with their advantages and disadvantages is shown in Tables 1–3. In addition, Tables 4 and 5 provide a summary of the publications from the recent five years (2019–2023) and their research groups. A total of 36 papers out of 90 were published within this period, indicating a growing interest in hybrid approaches from multiple research groups. Also, 8 papers out of 12 in the combined group were published within this period, indicating many research groups are extending their research from simple hybrid approaches to combined approaches lately.

Table 1
A comprehensive overview of approaches categorised under the internal group.

Groups	Internal		
	Parameter tuning	Replacing or correcting models	Representing distributions
Reference	[29]-[48]	[49]-[66]	[67]-[79]
No. of papers	20 (22.2%)	18 (20.0%)	13 (14.4%)
Roles of ML	ML predicts and tunes noise covariances or other parameters	ML replaces the model or predicts and corrects the error in the model	ML represents the posterior probability distribution
Advantages	Can improve robustness in a dynamic environment and can be applied to various applications	Can estimate without known transition and/or measurement model (or imperfect model)	Can improve the sampling and reduce the resampling with a distribution close to the real posterior distribution
Disadvantages	Limited to additive Gaussian noise	Still not able to replace highly nonlinear dynamics	High computation and practical issues in GMM
Common applications	Localisation and navigation	Localisation, navigation and SLAM	Dynamic systems and tracking problem
Common ML	ANN	GP and ANN (RBFNN)	GMM and kernel method

Table 2
A comprehensive overview of approaches categorised under the external group.

Groups	External		
	Compensation	Measurement update	Pseudo-inputs
Reference	[80]-[93]	[94]-[102]	[103]-[106]
No. of papers	14 (15.6%)	9 (10.0%)	4 (4.4%)
Roles of ML	ML predicts and compensates the error in the estimation	ML predicts and updates the error in the measurement	ML refines the inputs before they are used in the filters
Advantages	Can improve performance with simple architecture and can be applied to various applications	Can reduce divergence and can be applied to error state filters easily	Can reduce the nonlinearity and better performance with cheaper sensor
Disadvantages	Still can diverge when estimation error is big	Limited to specific applications	Generalisation issue is more critical
Common applications	Localisation, navigation and tracking	GPS/INS integration and sensor fusion	Vehicle attitude estimation
Common ML	ANN	ANN and RNN	ANN and DNN

It is important to point out the summary of the research topic trend within the recent five years (2019–2023) period. One of the most notable trends we found from the review is the change in the type of ML techniques applied. The ML techniques applied for this topic have been changed with the development of ML techniques or authors applied additional networks in their framework to improve the performance. For example, Kim et al. utilised an ensemble of DNN in their research [113] and proposed an enhanced approach with an additional filter network based on LSTM [114]. The second trend is the extension of the research to the combined group. It is shown that authors extend their works to the combined group which initially used one type of hybridisation. For example, Jiang and Nong proposed the combination of a KF and a group of ANNs where it predicts the parameters in KF for aviation surveillance [42] (Parameter tuning) but extended the work not only to tune the parameter in the filter but also to replace the transition model for time propagation [110] (Parameter tuning with others).

4.2. Suggestions for approaches and ML techniques by application

Suggestions of categorised approaches and ML techniques for some common applications are discussed in this section. For applications

Table 3
A comprehensive overview of approaches categorised under the combined group.

Groups	Combined	
	Parameter tuning with others	Distribution and models
Reference	[107]-[114]	[115]-[118]
No. of papers	8 (8.9%)	4 (4.4%)
Roles of ML	ML predicts noise covariance and replaces the model or refine the input	ML represents the probability distribution and replace the model
Advantages	Advantages of each group	Can improve the numerical stability and enables estimation without models
Disadvantages	Data preparation and generalisation issue	Kernel selection
Common applications	Navigation and tracking	Dynamic systems
Common ML	GP and RNN	Kernel method

Table 4

A summary of publications within the recent five years (2019–2023) of internal group.

Groups	Reference	First author	Research group	Year
Internal	[37]	K. Xiong		2021
	[38]	M. Brossard		2019
	[39]	I. Ullah	D. Kim	2019
	[40]	I. Ullah	D. Kim	2020
	[41]	X. Gao	H. Luo	2020
	[42]	Y. Jiang	X. Nong	2020
	[43]	S. Jouaber		2021
	[44]	F. Wu	H. Luo	2021
	[48]	G. Revach		2022
	[52]	H. Wang		2023
	[53]	B. Zhang		2022
	[57]	M. Sun	J. R. Hopgood	2020
	[66]	V. Pesce		2020
	[70]	Y. Xiaojing		2019
	[71]	T. Imbiriba		2020
	[72]	K. Comandur		2022
	[77]	S. Yun		2022
	[78]	M. Sun	J. R. Hopgood	2021
	[79]	M. Sun	J. R. Hopgood	2023

Table 5

A summary of publications within the recent five years (2019–2023) of external and combined groups.

Groups	Reference	First author	Research group	Year
External	[91]	D. Li		2021
	[92]	X. Chen	B. He	2020
	[93]	Y. Tang		2021
	[98]	X. Zhang	B. He	2019
	[99]	N. Shaukat	P. Otero	2021
	[100]	N. Shaukat	P. Otero	2021
	[101]	Y. Tang		2022
	[102]	S. Taghizadeh		2023
	[106]	H. Yu		2019
	[108]	W. A. Ramirez		2019
Combined	[109]	Z. Wu		2019
	[110]	Y. Jiang	X. Nong	2020
	[111]	F. Song		2022
	[112]	N. A. Bitar		2020
	[113]	D. Kim	K. Huh	2020
	[114]	D. Kim	K. Huh	2021
	[118]	G. H. W. Gebhardt		2019

where the dynamics or measurement model is not accurate, simplified or cannot be acquired properly, the parameter tuning approach (Section 3.2.1) or replacing or correcting models approach (Section 3.2.2) are particularly beneficial. ANNs and GPs have been shown to be effective with Bayes filters in these approaches. These approaches can enhance the performance of localisation, navigation or SLAM applications. For applications with high nonlinearity, representing the distribution approach (Section 3.2.3) is particularly beneficial. GMMs and kernel methods have shown to be effective with Bayes filters in this approach. This approach can enhance the performance of tracking problems. For signal processing applications where noise reduction is important, the compensation approach (Section 3.3.1) is particularly beneficial. Simple ANNs have shown to be effective with Bayes filters in this approach. However, the simple structure of this approach means it is not limited to signal processing and can be applied to broader applications. For sensor fusion applications, especially GPS and INS integration, the measurement update approach (Section 3.3.2) is particularly beneficial. ANNs and RNNs have been shown to be effective with the Bayes filter in this approach. RNNs with memory capability may offer additional benefits due to the temporal nature of the data. Finally, for vehicle attitude estimation problems, the pseudo-inputs approach (Section 3.3.3) is particularly beneficial. ANNs, especially DNNs, have shown effectiveness with Bayes filters in this approach.

4.3. Remaining challenges and research gaps

The remaining challenges and research gaps in Bayes filters with ML techniques are identified. It is worth mentioning some general problems aside from technical aspects. First of all, many approaches did not address computational complexity, even though ML techniques can be computationally demanding in many scenarios. This is crucial, as estimation problems often require real-time applications. In addition, comparisons with previously proposed methods within the same group could provide valuable insights for researchers when choosing an algorithm for one's applications. The lack of comparison makes it difficult to find the research trend and set the future research direction. This is one of the reasons why we proposed the categorisation and provided a list of the research papers within each group. We believe this categorisation and providing a list of the research paper can serve as a roadmap for future research, enabling the identification of development trends and relationships between different hybrid approaches. Moreover, the absence of a benchmark problem makes it challenging to compare performance. Lastly, many papers have gaps in the rationale behind their approach, not fully explaining why and how the proposed method can improve performance.

On a technical side, although many proposed approaches aimed to improve the performance of estimation by integrating with ML, each group has its remaining issues or disadvantages, as mentioned in each section. Apart from its issues, the biggest challenge of hybridising Bayes

filters with ML techniques across all groups is generalisation. Except for a few approaches, the generalisation to other applications or untrained data is rarely mentioned. In other words, the proposed approaches can improve performance under certain circumstances, sensors or dynamics covered by the training data but not well in general setups. This issue is particularly critical in external approaches since they are mostly tested in specific applications or conditions.

4.4. Proposal for the future research direction

Based on the remaining challenges and research gaps mentioned above, the following future research directions are recommended. From a general perspective, more attention should be given to evaluation aspects such as computational complexity and comparisons with previously proposed algorithms, as discussed in the previous section. In addition to that it is crucial to provide the rationale behind their proposed algorithms. On a technical side, addressing estimation problems with highly nonlinear dynamics and non-Gaussian noise requires further research. Researchers should also need to focus on the challenge of generalisation, a well-known issue in implementing ML techniques. Conducting an ablation study related to training data or sensor configuration can offer valuable insights into the robustness of a given approach. Exploring the use of GMMs or kernel methods to replace dynamics or represent the posterior distribution is also recommended, given their recent significant improvements. Finally, research on combined approaches is also encouraged, as these have demonstrated enhanced performance by leveraging the strengths of individual components.

5. Conclusion

In this review, we introduced and categorised approaches that integrate Bayes filters with ML techniques to enhance the performance of state estimation. The fusion approaches discussed in this review are categorised into three main groups, each containing two or three subgroups. Based on these groups, we analysed the advantages, disadvantages and limitations of each approach. The review also offered suggestions for approaches and ML techniques and highlighted remaining challenges and research gaps in Bayes filters with ML techniques. These gaps include often-overlooked computational costs, a lack of benchmark problems for performance comparison and missing rationales behind many approaches. It is recommended to evaluate computational complexity, compare with existing algorithms within the same group and clarify the reasoning behind new methods. Additionally, addressing challenges such as nonlinear dynamics, exploring the potential of the GMM or kernel method, investigating combined strategies and tackling the generalisation issue in ML with ablation studies can enhance algorithmic robustness and applicability. Finally, we believe this article serves as a guide for readers in designing their own hybrid filters or applications.

CRediT authorship contribution statement

Sukkeun Kim: Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Conceptualisation, Visualisation. **Ivan Petrunin:** Writing – review & editing, Supervision. **Hyo-Sang Shin:** Writing – review & editing, Supervision, Investigation, Conceptualisation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgement

This research was supported by the Inha University funded project Precision Guidance and Navigation for Deep Space Exploration (project number P20962)

References

- [1] R.E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.* 82 (1) (1960) 35–45, <http://dx.doi.org/10.1115/1.3662552>.
- [2] Y. Kim, H. Bang, et al., *Introduction to Kalman filter and its applications*, in: *Introduction and Implementations of the Kalman Filter*, Vol. 1, IntechOpen London, UK, 2018, pp. 1–16.
- [3] S. Julier, J. Uhlmann, Unscented filtering and nonlinear estimation, *Proc. IEEE* 92 (3) (2004) 401–422, <http://dx.doi.org/10.1109/JPROC.2003.823141>.
- [4] I. Arasaratnam, S. Haykin, Cubature Kalman filters, *IEEE Trans. Autom. Control* 54 (6) (2009) 1254–1269, <http://dx.doi.org/10.1109/tac.2009.2019800>.
- [5] R. Mehra, On the identification of variances and adaptive Kalman filtering, *IEEE Trans. Autom. Control* 15 (2) (1970) 175–184, <http://dx.doi.org/10.1109/tac.1970.1099422>.
- [6] S.D. Brown, S.C. Rutan, Adaptive Kalman filtering, *J. Res. Natl. Bureau Standards* (1977) 90 (6) (1985) 403–407, <http://dx.doi.org/10.6028/jres.090.032>.
- [7] Q. Xia, M. Rao, Y. Ying, X. Shen, Adaptive fading Kalman filter with an application, *Automatica* 30 (8) (1994) 1333–1338, [http://dx.doi.org/10.1016/0005-1098\(94\)90112-0](http://dx.doi.org/10.1016/0005-1098(94)90112-0).
- [8] R. Kottath, S. Poddar, A. Das, V. Kumar, Window based multiple model adaptive estimation for navigational framework, *Aerosp. Sci. Technol.* 50 (2016) 88–95, <http://dx.doi.org/10.1016/j.ast.2015.12.025>.
- [9] N.J. Gordon, D.J. Salmond, A.F.M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc.* 140 (2) (1993) 107, <http://dx.doi.org/10.1049/ip-f-2.1993.0015>.
- [10] J.S. Liu, R. Chen, Sequential Monte Carlo methods for dynamic systems, *J. Amer. Statist. Assoc.* 93 (443) (1998) 1032–1044, <http://dx.doi.org/10.1080/01621459.1998.10473765>.
- [11] J. Candy, Bootstrap particle filtering, *IEEE Signal Process. Mag.* 24 (4) (2007) 73–85, <http://dx.doi.org/10.1109/msp.2007.4286566>.
- [12] J.H. Kotecha, P.M. Djuric, Gaussian particle filtering, *IEEE Trans. Signal Process.* 51 (10) (2003) 2592–2601, <http://dx.doi.org/10.1109/tsp.2003.816758>.
- [13] M.K. Pitt, N. Shephard, Filtering via simulation: Auxiliary particle filters, *J. Am. Stat. Assoc.* 94 (446) (1999) 590–599.
- [14] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci.* 79 (8) (1982) 2554–2558, <http://dx.doi.org/10.1073/pnas.79.8.2554>.
- [15] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: A survey, *Heliyon* 4 (11) (2018) e00938, <http://dx.doi.org/10.1016/j.heliyon.2018.e00938>.
- [16] V.T. S. Elanayar, Y.C. Shin, Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems, *IEEE Trans. Neural Netw.* 5 (4) (1994) 594–603, <http://dx.doi.org/10.1109/72.298229>.
- [17] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780, <http://dx.doi.org/10.1162/neco.1997.9.8.1735>.
- [18] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: A survey, *J. Artificial Intelligence Res.* 4 (1996) 237–285, <http://dx.doi.org/10.1613/jair.301>.
- [19] C. Williams, C. Rasmussen, Gaussian processes for regression, in: D. Touretzky, M. Mozer, M. Hasselmo (Eds.), *Advances in Neural Information Processing Systems*, Vol. 8, MIT Press, 1995.
- [20] A. Danyluk, L. Bottou, M. Littman, L. Song, J. Huang, A. Smola, K. Fukumizu, Hilbert space embeddings of conditional distributions with applications to dynamical systems, in: *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 961–968, <http://dx.doi.org/10.1145/1553374.1553497>.
- [21] K. Muandet, K. Fukumizu, B. Sriperumbudur, B. Schölkopf, Kernel mean embedding of distributions: A review and beyond, *Found. Trends Mach. Learn.* 10 (1–2) (2017) 1–141, <http://dx.doi.org/10.1561/22000000060>, [arXiv:1605.09522](https://arxiv.org/abs/1605.09522).
- [22] S. Kim, I. Petrunin, H.-S. Shin, A review of Kalman filter with artificial intelligence techniques, in: *2022 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, 2022, pp. 1–12, <http://dx.doi.org/10.1109/icns54818.2022.9771520>.
- [23] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan, Z. Wang, A review: state estimation based on hybrid models of Kalman filter and neural network, *Syst. Sci. Control Eng.* 11 (1) (2023) 2173682, <http://dx.doi.org/10.1080/21642583.2023.2173682>.

- [24] D.-J. Jwo, A. Biswal, I.A. Mir, Artificial neural networks for navigation systems: A review of recent research, *Appl. Sci.* 13 (7) (2023) <http://dx.doi.org/10.3390/app13074475>.
- [25] T. Li, S. Sun, T.P. Sattar, J.M. Corchado, Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches, *Expert Syst. Appl.* 41 (8) (2014) 3944–3954, <http://dx.doi.org/10.1016/j.eswa.2013.12.031>, [arXiv:1308.2443](https://arxiv.org/abs/1308.2443).
- [26] T. Li, M. Bolic, P.M. Djuric, Resampling Methods for Particle Filtering: Classification, implementation, and strategies, *IEEE Signal Process. Mag.* 32 (3) (2015) 70–86, <http://dx.doi.org/10.1109/msp.2014.2330626>.
- [27] C. Kuptamete, N. Aunsri, A review of resampling techniques in particle filtering framework, *Measurement* 193 (2022) 110836, <http://dx.doi.org/10.1016/j.measurement.2022.110836>.
- [28] C.K. Williams, C.E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2, (3) MIT press Cambridge, MA, 2006.
- [29] D.-J. Jwo, H.-C. Huang, Neural network aided adaptive extended Kalman filtering approach for DGPS positioning, *J. Navig.* 57 (3) (2004) 449–463, <http://dx.doi.org/10.1017/s0373463304002814>.
- [30] D.-J. Jwo, H.-C. Huang, GPS navigation using fuzzy neural network aided adaptive extended Kalman filter, in: *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005, pp. 7840–7845, <http://dx.doi.org/10.1109/cdc.2005.1583429>.
- [31] D.-J. Jwo, J.-J. Chen, Neural network aided adaptive Kalman filter for GPS/INS navigation system design, in: *9th IFAC Workshop*, 2011, pp. 1–7.
- [32] W. Wu, W. Min, The mobile robot GPS position based on neural network adaptive Kalman filter, in: *2009 International Conference on Computational Intelligence and Natural Computing*, Vol. 1, 2009, pp. 26–29, <http://dx.doi.org/10.1109/cinc.2009.257>.
- [33] J.K. Lee, C. Jekeli, Neural network aided adaptive filtering and smoothing for an integrated INS/GPS unexploded ordnance geolocation system, *J. Navig.* 63 (2) (2010) 251–267, <http://dx.doi.org/10.1017/s037346330990397>.
- [34] Z. Shi, P. Yue, X. Wang, Research on adaptive Kalman filter algorithm based on fuzzy neural network, in: *The 2010 IEEE International Conference on Information and Automation*, 2010, pp. 1636–1640, <http://dx.doi.org/10.1109/icinf.2010.5512237>.
- [35] Y. Xia, J. Wang, Low-dimensional recurrent neural network-based Kalman filter for speech enhancement, *Neural Netw.* 67 (2015) 131–139, <http://dx.doi.org/10.1016/j.neunet.2015.03.008>.
- [36] Z. Bekhtaoui, A. Meche, M. Dahmani, K.A. Meraim, Maneuvering target tracking using Q-learning based Kalman filter, in: *2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B)*, 2017, pp. 1–5, <http://dx.doi.org/10.1109/icee-b.2017.8192005>.
- [37] K. Xiong, C. Wei, H. Zhang, Q-learning for noise covariance adaptation in extended KALMAN filter, *Asian J. Control* 23 (4) (2021) 1803–1816, <http://dx.doi.org/10.1002/asjc.2336>.
- [38] M. Brossard, A. Barrau, S. Bonnabel, AI-IMU dead-reckoning, *IEEE Trans. Intell. Veh.* 5 (4) (2019) 585–595, <http://dx.doi.org/10.1109/tiv.2020.2980758>.
- [39] I. Ullah, M. Fayaz, D. Kim, Improving accuracy of the Kalman filter algorithm in dynamic conditions using ANN-based learning module, *Symmetry* 11 (1) (2019) 94, <http://dx.doi.org/10.3390/sym11010094>.
- [40] I. Ullah, M. Fayaz, N. Naved, D. Kim, ANN based learning to Kalman filter algorithm for indoor environment prediction in smart greenhouse, *IEEE Access* 8 (2020) 159371–159388, <http://dx.doi.org/10.1109/access.2020.3016277>.
- [41] X. Gao, H. Luo, B. Ning, F. Zhao, L. Bao, Y. Gong, Y. Xiao, J. Jiang, RL-AKF: An adaptive Kalman filter navigation algorithm based on reinforcement learning for ground vehicles, *Remote Sens.* 12 (11) (2020) 1704, <http://dx.doi.org/10.3390/rs12111704>.
- [42] Y. Jiang, X. Nong, Trainable Kalman filter based on recurrent neural network and its application in aviation surveillance, *J. Phys. Conf. Ser.* 1642 (1) (2020) 012010, <http://dx.doi.org/10.1088/1742-6596/1642/1/012010>.
- [43] S. Jouaber, S. Bonnabel, S. Velasco-Forero, M. Pilte, NNAKF: A neural network adapted Kalman filter for target tracking, in: *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 4075–4079, <http://dx.doi.org/10.1109/icassp39728.2021.9414681>.
- [44] F. Wu, H. Luo, H. Jia, F. Zhao, Y. Xiao, X. Gao, Predicting the noise covariance with a multitask learning model for Kalman filter-based GNSS/INS integrated navigation, *IEEE Trans. Instrum. Meas.* 70 (2021) 1–13, <http://dx.doi.org/10.1109/tim.2020.3024357>.
- [45] D. Xinyuan, W. Jinjie, W. Sufeng, C. Ting, An improved CDKF algorithm based on RBF neural network for satellite attitude determination, in: *2012 International Conference on Image Analysis and Signal Processing*, 2012, pp. 1–7, <http://dx.doi.org/10.1109/iasp.2012.6425013>.
- [46] R. Turner, C.E. Rasmussen, Model based learning of sigma points in unscented Kalman filtering, *Neurocomputing* 80 (2012) 47–53, <http://dx.doi.org/10.1016/j.neucom.2011.07.029>.
- [47] A. Motwani, S.K. Sharma, R. Sutton, P. Culverhouse, Application of artificial neural networks to weighted interval Kalman filtering, *Proc. Inst. Mech. Eng. I* 228 (5) (2013) 267–277, <http://dx.doi.org/10.1177/0959651813520148>.
- [48] G. Revach, N. Shlezinger, X. Ni, A.L. Escoriza, R.J.G.v. Sloun, Y.C. Eldar, KalmanNet: Neural network aided Kalman filtering for partially known dynamics, *IEEE Trans. Signal Process.* 70 (2022) 1532–1547, <http://dx.doi.org/10.1109/tsp.2022.3158588>, [arXiv:2107.10043](https://arxiv.org/abs/2107.10043).
- [49] N. Vlassis, B. Terwijn, B. Kröse, Auxiliary particle filter robot localization from high-dimensional sensor observations, in: *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292)*, Vol. 1, 2002, pp. 7–12, <http://dx.doi.org/10.1109/robot.2002.1013331>.
- [50] X. Wang, S. Wang, J.-J. Ma, An improved particle filter for target tracking in sensor systems, *Sensors* 7 (1) (2007) 144–156, <http://dx.doi.org/10.3390/s7010144>.
- [51] P.-H. Michel, V. Heiries, An adaptive sigma point Kalman filter hybridized by support vector machine algorithm for battery SoC and SoH estimation, in: *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–7, <http://dx.doi.org/10.1109/vtcspring.2015.7145678>.
- [52] H. Wang, Extreme learning Kalman filter for short-term wind speed prediction, *Front. Energy Res.* 10 (2023) <http://dx.doi.org/10.3389/fenrg.2022.1047381>.
- [53] B. Zhang, Y.C. Shin, A Gaussian mixture filter with adaptive refinement for nonlinear state estimation, *Signal Process.* 201 (2022) 108677, <http://dx.doi.org/10.1016/j.sigpro.2022.108677>.
- [54] J. Ko, D. Fox, GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models, *Auton. Robots* 27 (1) (2009) 75–90, <http://dx.doi.org/10.1007/s10514-009-9119-x>.
- [55] M.P. Deisenroth, M.F. Huber, U.D. Hanebeck, Analytic moment-based Gaussian process filtering, in: *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, 2009, pp. 1–8, <http://dx.doi.org/10.1145/1553374.1553403>.
- [56] R.K. Neerukatti, M.Y. Fard, A. Chattopadhyay, Gaussian process-based particle-filtering approach for real-time damage prediction with application, *J. Aerosp. Eng.* 30 (1) (2017) 04016080, [http://dx.doi.org/10.1061/\(asce\)as.1943-5525.0000680](http://dx.doi.org/10.1061/(asce)as.1943-5525.0000680).
- [57] M. Sun, M.E. Davies, I. Proudler, J.R. Hopgood, A Gaussian process based method for multiple model tracking, in: *2020 Sensor Signal Processing for Defence Conference (SSPD)*, 2020, pp. 1–5, <http://dx.doi.org/10.1109/sspd47486.2020.9272174>.
- [58] B. Ferris, D. Haehnel, D. Fox, Gaussian processes for signal strength-based location estimation, *Robot.: Sci. Syst. II* (2006) <http://dx.doi.org/10.15607/rss.2006.ii.039>.
- [59] S. Stuberud, R. Lobb, M. Owen, An adaptive extended Kalman filter using artificial neural networks, in: *Proceedings of 1995 34th IEEE Conference on Decision and Control*, Vol. 2 (1995) 1852–1856, <http://dx.doi.org/10.1109/cdc.1995.480611>.
- [60] R. Zhan, J. Wan, Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation, *IEEE Signal Process. Lett.* 13 (7) (2006) 445–448, <http://dx.doi.org/10.1109/lsp.2006.871854>.
- [61] K. Kramer, S. Stuberud, Tracking of multiple target types with a single neural extended Kalman filter, in: *2006 3rd International IEEE Conference Intelligent Systems*, 2006, pp. 463–468, <http://dx.doi.org/10.1109/is.2006.348463>.
- [62] M. Choi, R. Sakthivel, W.K. Chung, Neural network-aided extended Kalman filter for SLAM problem, in: *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 2007, pp. 1686–1690, <http://dx.doi.org/10.1109/robot.2007.363565>.
- [63] K.Y. Kotov, A.S. Maltsev, M.A. Sobolev, Recurrent neural network and extended Kalman filter in SLAM problem, *IFAC Proc. Vol.* 46 (20) (2013) 23–26, <http://dx.doi.org/10.3182/20130902-3-cn-3020.00057>.
- [64] Z. Miljković, N. Vuković, M. Mitić, Neural extended Kalman filter for monocular SLAM in indoor environment, *Proc. Inst. Mech. Eng. C* 230 (5) (2015) 856–866, <http://dx.doi.org/10.1177/0954406215586589>.
- [65] Z. Miao, H. Shi, Y. Zhang, F. Xu, Neural network-aided variational Bayesian adaptive cubature Kalman filtering for nonlinear state estimation, *Meas. Sci. Technol.* 28 (10) (2017) 105003, <http://dx.doi.org/10.1088/1361-6501/aa7d70>.
- [66] V. Pesce, S. Silvestrini, M. Lavagna, Radial basis function neural network aided adaptive extended Kalman filter for spacecraft relative navigation, *Aerosp. Sci. Technol.* 96 (2020) 105527, <http://dx.doi.org/10.1016/j.ast.2019.105527>.
- [67] W. Qin, Q. Peng, An improved particle filter algorithm based on neural network for visual tracking, in: *2007 International Conference on Communications, Circuits and Systems*, 2007, pp. 765–768, <http://dx.doi.org/10.1109/icccas.2007.4348162>.
- [68] Z. Pang, D. Liu, N. Jin, Z. Wang, Neural network strategy for sampling of particle filters on the tracking problem, in: *2007 International Joint Conference on Neural Networks*, 2007, pp. 254–259, <http://dx.doi.org/10.1109/ijcnn.2007.4370964>.
- [69] Z. Pang, D. Liu, N. Jin, Z. Wang, A Monte Carlo particle model associated with neural networks for tracking problem, *IEEE Trans. Circuits Syst. I. Regul. Pap.* 55 (11) (2008) 3421–3429, <http://dx.doi.org/10.1109/tcsi.2008.924894>.
- [70] Y. Xiaojing, J. Qingju, L. Xinke, Research on GRNN-based particle filter algorithm, in: *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2019, pp. 2053–2057, <http://dx.doi.org/10.1109/itnec.2019.8729218>.
- [71] T. Imbiriba, P. Closas, Enhancing particle filtering using Gaussian processes, in: *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, 2020, pp. 1–7, <http://dx.doi.org/10.23919/fusion45008.2020.9190631>.
- [72] K. Comandur, Y. Li, S. Nannuru, Particle flow Gaussian particle filter, in: *2022 25th International Conference on Information Fusion (FUSION)*, 2022, pp. 1–6, <http://dx.doi.org/10.23919/fusion49751.2022.9841354>.

- [73] J.H. Kotecha, P.M. Djurić, Gaussian sum particle filtering, *IEEE Trans. Signal Process.* 51 (10) (2003) 2602, <http://dx.doi.org/10.1109/tsp.2003.816754>.
- [74] M.L. Psiaki, Gaussian mixture nonlinear filtering with resampling for mixand narrowing, *IEEE Trans. Signal Process.* 64 (21) (2015) 5499–5512, <http://dx.doi.org/10.1109/tsp.2016.2595503>.
- [75] D. Raihan, S. Chakravorty, Particle Gaussian mixture filters-I, *Automatica* 98 (2018) 331–340, <http://dx.doi.org/10.1016/j.automatica.2018.07.023>.
- [76] D. Raihan, S. Chakravorty, Particle Gaussian mixture filters-II, *Automatica* 98 (2018) 341–349, <http://dx.doi.org/10.1016/j.automatica.2018.07.024>.
- [77] S. Yun, R. Zanetti, Clustering methods for particle filters with Gaussian mixture models, *IEEE Trans. Aerosp. Electron. Syst.* 58 (2) (2022) 1109–1118, <http://dx.doi.org/10.1109/taes.2021.3117655>.
- [78] M. Sun, M.E. Davies, J.R. Hopgood, I. Proudler, Adaptive kernel Kalman filter multi-sensor fusion, in: 2021 IEEE 24th International Conference on Information Fusion (FUSION), 2021, pp. 1–8, <http://dx.doi.org/10.23919/fusion49465.2021.9627030>.
- [79] M. Sun, M.E. Davies, I.K. Proudler, J.R. Hopgood, Adaptive kernel Kalman filter, *IEEE Trans. Signal Process.* 71 (2023) 713–726, <http://dx.doi.org/10.1109/tsp.2023.3250829>, arXiv:2203.08300.
- [80] L. Chin, Application of neural networks in target tracking data fusion, *IEEE Trans. Aerosp. Electron. Syst.* 30 (1) (1994) 281–287, <http://dx.doi.org/10.1109/7.250437>.
- [81] V. Vaidehi, N. Chitra, M. Chokkalingam, C. Krishnan, Neural network aided Kalman filtering for multitarget tracking applications, *Comput. Electr. Eng.* 27 (2) (2001) 217–228, [http://dx.doi.org/10.1016/s0045-7906\(00\)00013-6](http://dx.doi.org/10.1016/s0045-7906(00)00013-6).
- [82] Z.-J. Yu, S.-L. Dong, J.-M. Wei, T. Xing, H.-T. Liu, Neural network aided unscented Kalman filter for maneuvering target tracking in distributed acoustic sensor networks, in: 2007 International Conference on Computing: Theory and Applications (ICCTA'07), 2007, pp. 245–249, <http://dx.doi.org/10.1109/iccta.2007.88>.
- [83] P. Li, X. Zhang, X. Xu, Novel terrain integrated navigation system using neural network aided Kalman filter, in: 2010 Sixth International Conference on Natural Computation, Vol. 1, 2010, pp. 445–448, <http://dx.doi.org/10.1109/icnc.2010.5583345>.
- [84] X. Gao, X. Zhong, D. You, S. Katayama, Kalman filtering compensated by radial basis function neural network for seam tracking of laser welding, *IEEE Trans. Control Syst. Technol.* 21 (5) (2013) 1916–1923, <http://dx.doi.org/10.1109/tcst.2012.2219861>.
- [85] Y.-x. Zhang, D.-y. You, X.-d. Gao, S.-J. Na, Automatic gap tracking during high power laser welding based on particle filtering method and BP neural network, *Int. J. Adv. Manuf. Technol.* 96 (1–4) (2018) 685–696, <http://dx.doi.org/10.1007/s00170-018-1636-3>.
- [86] H.-N. Nguyen, J. Zhou, H.-J. Kang, A calibration method for enhancing robot accuracy through integration of an extended Kalman filter algorithm and an artificial neural network, *Neurocomputing* 151 (2015) 996–1005, <http://dx.doi.org/10.1016/j.neucom.2014.03.085>.
- [87] M. Sun, Z. Ma, Y. Li, Maneuvering target tracking using IMM Kalman filter aided by elman neural network, in: 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, Vol. 1, 2015, pp. 144–148, <http://dx.doi.org/10.1109/ihmsc.2015.241>.
- [88] W. Zhao, X. Kong, C. Wang, Combined estimation of the state of charge of a lithium battery based on a back-propagation–adaptive Kalman filter algorithm, *Proc. Inst. Mech. Eng. D* 232 (3) (2017) 357–366, <http://dx.doi.org/10.1177/0954407017701533>.
- [89] L. Song, Z. Duan, B. He, Z. Li, Application of federal Kalman filter with neural networks in the velocity and attitude matching of transfer alignment, *Complexity* 2018 (2018) 1–7, <http://dx.doi.org/10.1155/2018/3039061>.
- [90] S. Hosseinalamdary, Deep Kalman filter: Simultaneous multi-sensor integration and modelling: a GNSS/IMU case study, *Sensors* (Basel, Switzerland) 18 (5) (2018) 1316, <http://dx.doi.org/10.3390/s18051316>.
- [91] D. Li, J. Zhou, Y. Liu, Recurrent-neural-network-based unscented Kalman filter for estimating and compensating the random drift of MEMS gyroscopes in real time, *Mech. Syst. Signal Process.* 147 (2021) 107057, <http://dx.doi.org/10.1016/j.ymssp.2020.107057>.
- [92] X. Chen, S. Wu, X. Zhang, X. Mu, T. Yan, B. He, Neural network assisted adaptive unscented Kalman filter for AUV, in: *Global Oceans 2020: Singapore–U.S. Gulf Coast*, 2020, pp. 1–4, <http://dx.doi.org/10.1109/ieeconf38699.2020.9389271>.
- [93] Y. Tang, L. Hu, Q. Zhang, W. Pan, Reinforcement learning compensated extended Kalman filter for attitude estimation, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021, pp. 6854–6859, <http://dx.doi.org/10.1109/iros51168.2021.9635963>.
- [94] J.J. Wang, J. Wang, D. Sinclair, L. Watts, Neural network aided Kalman filtering for integrated GPS/INS geo-referencing platform, in: *5th Int. Symp. Mobile Mapping Technology*, 2007, pp. 1–6.
- [95] T. Zhang, X. Xu, A new method of seamless land navigation for GPS/INS integrated system, *Measurement* 45 (4) (2012) 691–701, <http://dx.doi.org/10.1016/j.measurement.2011.12.021>.
- [96] L. Chen, J. Fang, A hybrid prediction method for bridging GPS outages in high-precision POS application, *IEEE Trans. Instrum. Meas.* 63 (6) (2014) 1656–1665, <http://dx.doi.org/10.1109/tim.2013.2292277>.
- [97] Z. Jingsen, Z. Wenjie, H. Bo, W. Yali, Integrating extreme learning machine with Kalman filter to bridge GPS outages, in: 2016 3rd International Conference on Information Science and Control Engineering (ICISCE), 2016, pp. 420–424, <http://dx.doi.org/10.1109/icisce.2016.98>.
- [98] X. Zhang, X. Mu, H. Liu, B. He, T. Yan, Application of modified EKF based on intelligent data fusion in AUV navigation, in: 2019 IEEE Underwater Technology (UT), 2019, pp. 1–4, <http://dx.doi.org/10.1109/ut.2019.8734414>.
- [99] N. Shaukat, A. Ali, M. Moinuddin, P. Otero, Underwater vehicle localization by hybridization of indirect Kalman filter and neural network, in: 2021 7th International Conference on Mechatronics and Robotics Engineering (ICMRE), 2021, pp. 111–115, <http://dx.doi.org/10.1109/icmre51691.2021.9384844>.
- [100] N. Shaukat, A. Ali, M.J. Iqbal, M. Moinuddin, P. Otero, Multi-sensor fusion for underwater vehicle localization by augmentation of RBF neural network and error-state Kalman filter, *Sensors* 21 (4) (2021) 1149, <http://dx.doi.org/10.3390/s21041149>.
- [101] Y. Tang, J. Jiang, J. Liu, P. Yan, Y. Tao, J. Liu, A GRU and AKF-based hybrid algorithm for improving INS/GNSS navigation accuracy during GNSS outage, *Remote Sens.* 14 (3) (2022) 752, <http://dx.doi.org/10.3390/rs14030752>.
- [102] S. Taghizadeh, R. Safabakhsh, An integrated INS/GNSS system with an attention-based deep network for drones in GNSS denied environments, *IEEE Aerosp. Electron. Syst. Magaz.* 38 (8) (2023) 14–25, <http://dx.doi.org/10.1109/MAES.2023.3266180>.
- [103] L. Vargas-Meléndez, B.L. Boada, M.J.L. Boada, A. Gauchía, V. Díaz, A sensor fusion method based on an integrated neural network and Kalman filter for vehicle roll angle estimation, *Sensors* (Basel, Switzerland) 16 (9) (2016) 1400, <http://dx.doi.org/10.3390/s16091400>.
- [104] T. Novi, R. Capitani, C. Annicchiarico, An integrated artificial neural network–unscented Kalman filter vehicle sideslip angle estimation based on inertial measurement unit measurements, *Proc. Inst. Mech. Eng. D* 233 (7) (2018) 1864–1878, <http://dx.doi.org/10.1177/0954407018790646>.
- [105] C. Zhong, Z. Jiang, X. Chu, T. Guo, Q. Wen, Water level forecasting using a hybrid algorithm of artificial neural networks and local Kalman filtering, *Proc. Inst. Mech. Eng. M* 233 (1) (2017) 174–185, <http://dx.doi.org/10.1177/1475090217727135>.
- [106] H. Yu, Z. Ouyang, W.-P. Zhu, B. Champagne, Y. Ji, A deep neural network based Kalman filter for time domain speech enhancement, in: 2019 IEEE International Symposium on Circuits and Systems (ISCAS), 2019, pp. 1–5, <http://dx.doi.org/10.1109/iscas.2019.8702161>.
- [107] J. Ko, D.J. Klein, D. Fox, D. Haehnel, GP-UKF: Unscented kalman filters with Gaussian process prediction and observation models, in: 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, pp. 1901–1907, <http://dx.doi.org/10.1109/iro.2007.4399284>.
- [108] W.A. Ramirez, Z.Q. Leong, H. Nguyen, S.G. Jayasinghe, Position estimation for underwater vehicles using unscented Kalman filter with Gaussian process prediction, *Underw. Technol.* 36 (2) (2019) 28–34, <http://dx.doi.org/10.3723/ut.36.029>.
- [109] Z. Wu, M. Shang, D. Shen, S. Qi, SOC estimation for batteries using MS-AUKF and neural network, *J. Renew. Sustain. Energy* 11 (2) (2019) 024103, <http://dx.doi.org/10.1063/1.5064479>.
- [110] Y. Jiang, X. Nong, A radar filtering model for aerial surveillance base on Kalman filter and neural network, in: 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS), 2020, pp. 57–60, <http://dx.doi.org/10.1109/icseess49938.2020.9237716>.
- [111] F. Song, Y. Li, W. Cheng, L. Dong, M. Li, J. Li, An improved Kalman filter based on long short-memory recurrent neural network for nonlinear radar target tracking, *Wirel. Commun. Mob. Comput.* 2022 (2022) 1–10, <http://dx.doi.org/10.1155/2022/8280428>.
- [112] N.A. Bitar, A. Gavrilov, Neural networks aided unscented Kalman filter for integrated INS/GNSS systems, in: 2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), 2020, pp. 1–4, <http://dx.doi.org/10.23919/icins43215.2020.9133878>.
- [113] D. Kim, K. Min, H. Kim, K. Huh, Vehicle sideslip angle estimation using deep ensemble-based adaptive Kalman filter, *Mech. Syst. Signal Process.* 144 (2020) 106862, <http://dx.doi.org/10.1016/j.ymssp.2020.106862>.
- [114] D. Kim, G. Kim, S. Choi, K. Huh, An integrated deep ensemble-unscented Kalman filter for sideslip angle estimation with sensor filtering network, *IEEE Access* PP (99) (2021) 1, <http://dx.doi.org/10.1109/access.2021.3125351>.
- [115] K. Fukumizu, L. Song, A. Gretton, Kernel Bayes' rule: Bayesian inference with positive definite kernels, *J. Mach. Learn. Res.* 14 (1) (2013) 3753–3783.
- [116] M. Kanagawa, Y. Nishiyama, A. Gretton, K. Fukumizu, Monte Carlo filtering using kernel embedding of distributions, in: *Proceedings of the AAAI Conference on Artificial Intelligence*, 28, (1) 2014, <http://dx.doi.org/10.1609/aaai.v28i1.8984>.
- [117] M. Kanagawa, Y. Nishiyama, A. Gretton, K. Fukumizu, Filtering with state-observation examples via kernel Monte Carlo filter, *Neural Comput.* 28 (2) (2016) 382–444, http://dx.doi.org/10.1162/neco_a.00806.
- [118] G.H.W. Gebhardt, A. Kupcsik, G. Neumann, The kernel Kalman rule, *Mach. Learn.* 108 (12) (2019) 2113–2157, <http://dx.doi.org/10.1007/s10994-019-05816-z>.



DSPACE

<https://dspace.org/>

A review of Bayes filters with machine learning techniques and their applications

Kim, Sukkeun; Petrunin, Ivan; Shin, Hyo-Sang

2025-02-01

Elsevier

<https://doi.org/10.1016/j.inffus.2024.102707>

Kim S, Petrunin I, Shin H-S. (2025) A review of Bayes filters with machine learning techniques and their applications. Information Fusion, Volume 114, February 2025, Article number 102707