2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

Parameterizing Hierarchical Particle Filters with Concept Drift for Time-varying Parameter Estimation

Joshua Murphy*, Conor Rosato†, Andrew Millard* and Simon Maskell*
* Dept. of Electrical Engineering and Electronics, University of Liverpool, United Kingdom
E-mail: {joshua.murphy, andrew.millard, smaskell}@liverpool.ac.uk
† Dept. of Pharmacology and Therapeutics, University of Liverpool, United Kingdom
E-mail: cmrosa@liverpool.ac.uk

***Abstract*—Models used in real-world applications can experience concept drift, where the deployment environment differs from the training environment due to previously unseen input data distributions or changes in the data generation process. Sequential Monte Carlo squared (SMC^2) is a Bayesian algorithm which infers the distribution of both the states and parameters of non-linear, non-Gaussian state space models. A Sequential Monte Carlo (SMC) sampler operates on the static parameter space and samples the parameters with which a number of particle filters (PFs), tackle the dynamic state inference problem. PFs often do not consider the influence of concept drift on the parameters of interest. Models that consider drift will do so in a purely online context and can suffer from slow convergence. By formulating PFs as hierarchical models, with explicit dynamics to describe parameter evolution, the use of batch parameter estimation techniques like SMC^2 is enabled. We demonstrate our approach by applying five concept drift models to an epidemiological disease model with time-varying transmission rates.**

I. Introduction

Bayesian inference is well-suited for quantifying uncertainty because it outputs distributions rather than single-point estimates. This approach allows for the integration of prior knowledge, the connection of observations to quantities through likelihood distributions, and the generation of posterior distributions. It can be utilized to infer unknown quantities in both static and dynamic contexts [1].

Particle filters (PFs) [2] are a widely used Bayesian method for tracking the dynamic evolution of state spaces in non-linear and non-Gaussian state space models (SSMs). They have achieved widespread usage in a variety of fields such as epidemiology [3], robotics [4] and target tracking [5]. Typically, in PFs, the parameters governing the evolution of the dynamics are fixed. However, in some instances, PFs can be expanded to infer these parameters. One approach is to augment the state space with the parameters and consider their co-evolution [6], [7].

Bayesian parameter estimation using PFs can occur online [8]–[11] and offline. One such online recursive method is the nested particle filter (NPF) which uses one PF to estimate the parameters and an inner particle filter to track the states [12], [13]. Offline methods include particle marginal methods, such as particle Markov Chain Monte Carlo (p-MCMC) [14] and Sequential Monte Carlo squared (SMC^2) [15], which are batch methods that require access to the full data-set at each evaluation. Similarly to the NPF, an inner particle filter tracks the states but an outer static Monte Carlo method deals with parameter estimation. SMC^2 is similar to the NPF in that they both contain two Sequential Monte Carlo methods but SMC^2 possesses favourable convergence properties while the NPF has favourable time complexity [12].

Concept drift refers to changes in the statistical environment of the problem at hand, a term commonly used in classification contexts [16]–[18]. In the context of PFs, concept drift has been modelled in online streaming contexts [19]–[21] and rarely associated with batch inference [22]. A pertinent example of a changing statistical environment is the modelling of communicable diseases. It has been noted in [23] that the use of concept drift in compartmental disease modelling is under-utilized. Two related examples can be found in [22], [24] where the abrupt concept drift (see Section III) is used to define the sudden change in the transmissibility of Ebola and HIV. Particle-MCMC was employed to infer the parameters of the disease model as well as the hyper-parameters governing the sudden change in transmissibility. As interventions from policymakers are introduced, the disease dynamics changes over time. It is therefore common practice to define β as time-varying [22], [25]–[28].

The contribution of this paper is to explicitly define an epidemiological model with frequently encountered types of concept drift and incorporate them as part of a hierarchical particle filter to perform batch Bayesian inference with SMC^2.

The paper continues by describing the SMC^2 algorithm in Section II and defining the dynamic models used to represent concept drift in Section III. The compartmental epidemiological model is defined in in Section IV and numerical results are provided in Section VI. Conclusions and future work are given in Section VII.

II. SMC^2

SSMs have been used to model unobserved latent states of dynamic systems in a wide range of research fields [29]. An SSM consists of a state equation,

$$\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1} \sim p(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1}, \boldsymbol{\theta}), \tag{1}$$

which is parameterized here by $\boldsymbol{\theta}$ and models the Markovian movement of the dynamic system from the previous state, $\boldsymbol{x}_{t-1}$, to the current state, $\boldsymbol{x}_t$. SSMs also include an observation equation

$$\boldsymbol{y}_t \mid \boldsymbol{x}_t \sim p(\boldsymbol{y}_t \mid \boldsymbol{x}_t, \boldsymbol{\theta}), \tag{2}$$

which describes how the observation, $\boldsymbol{y}_t$, is linked to the state $\boldsymbol{x}_t$.

Analysing a set of observations $\boldsymbol{y}_{1:T} \in \mathbb{R}^{D_y}$, SMC^2 uses an inner PF to estimate the unobserved states of a system, $\boldsymbol{x}_{1:T} \in \mathbb{R}^{D_x}$, (Section II-A) and an outer SMC sampler to estimate the parameters, $\boldsymbol{\theta} \in \mathbb{R}^D$ (Section II-B).

A. SMC Sampler

At each iteration k, SMC^2 targets

$$\pi(\boldsymbol{\theta}) \propto p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta})p(\boldsymbol{\theta}), \tag{3}$$

via sequential importance sampling and resampling steps. An unbiased estimate of the marginal likelihood $p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta})$ is provided by the PF. In this section, the posterior distribution over the $\pi(\boldsymbol{\theta}|\boldsymbol{y}_{1:T})$ is defined as $\pi(\boldsymbol{\theta})$.

The joint distribution of all states until $k = K$ is defined to be

$$\pi(\boldsymbol{\theta}_{1:K}) = \pi(\boldsymbol{\theta}_K) \prod_{k=2}^{K} L(\boldsymbol{\theta}_{k-1}|\boldsymbol{\theta}_k), \tag{4}$$

where $L(\boldsymbol{\theta}_{k-1}|\boldsymbol{\theta}_k)$ is the L-kernel, which is a user-defined probability distribution. The choice of this distribution can affect the efficiency of the sampler [30].

At $k = 1$, N samples $\forall i = 1, \ldots, N$ are drawn from a prior distribution $q_1(\cdot)$ as follows:

$$\boldsymbol{\theta}_1^i \sim q_1(\cdot), \quad \forall i, \tag{5}$$

and weighted according to

$$\boldsymbol{w}_1^i = \frac{\pi(\boldsymbol{\theta}_1^i)}{q_1(\boldsymbol{\theta}_1^i)}, \quad \forall i. \tag{6}$$

At $k > 1$, subsequent samples are proposed based on samples from the previous iteration via a proposal distribution, $q(\boldsymbol{\theta}_k^i|\boldsymbol{\theta}_{k-1}^i)$, as follows:

$$\boldsymbol{\theta}_k^i \sim q(\cdot|\boldsymbol{\theta}_{k-1}^i). \tag{7}$$

The proposal is commonly chosen to be Gaussian with a mean of $\boldsymbol{\theta}_{k-1}^i$ and a covariance of $\boldsymbol{\Sigma} \in \mathbb{R}^{DxD}$, such that

$$q(\boldsymbol{\theta}_k^i|\boldsymbol{\theta}_{k-1}^i) = \mathcal{N}(\boldsymbol{\theta}_k^i; \boldsymbol{\theta}_{k-1}^i, \Sigma), \quad \forall i. \tag{8}$$

These samples are weighted according to

$$\boldsymbol{w}_k^i = \boldsymbol{w}_{k-1}^i \frac{\pi(\boldsymbol{\theta}_k^i)}{\pi(\boldsymbol{\theta}_{k-1}^i)} \frac{L(\boldsymbol{\theta}_{k-1}^i|\boldsymbol{\theta}_k^i)}{q(\boldsymbol{\theta}_k^i|\boldsymbol{\theta}_{k-1}^i)}, \quad \forall i. \tag{9}$$

A common, suboptimal approach to selecting the L-kernel in (9) is to choose the same distribution as the forwards proposal distribution:

$$L(\boldsymbol{\theta}_{k-1}^i|\boldsymbol{\theta}_k^i) = q(\boldsymbol{\theta}_{k-1}^i|\boldsymbol{\theta}_k^i), \quad \forall i, \tag{10}$$

such that, when $q(\cdot|\cdot)$ is symmetric, the weight update in (9) simplifies to

$$\boldsymbol{w}_k^i = \boldsymbol{w}_{k-1}^i \frac{\pi(\boldsymbol{\theta}_k^i)}{\pi(\boldsymbol{\theta}_{k-1}^i)}, \quad \forall i. \tag{11}$$

Estimates of the expectations of functions, such as moments, on the distribution are realised by

$$\tilde{\mathbf{f}}_k = \sum\nolimits_{i=1}^{N} \tilde{\boldsymbol{w}}_k^i \boldsymbol{\theta}_{1:k}^i, \tag{12}$$

where the normalised weights $\tilde{\boldsymbol{w}}_k^i$ are calculated by

$$\tilde{\boldsymbol{w}}_k^i = \frac{\boldsymbol{w}_k^i}{\sum_{i=1}^N \boldsymbol{w}_k^i}, \quad \forall i. \tag{13}$$

SMC^2 computes the Effective Sample Size (ESS) as a measure of the efficacy of the sampler

$$N^{\text{eff}} = \frac{1}{\sum_{i=1}^N \left(\tilde{\boldsymbol{w}}_k^i\right)^2}. \tag{14}$$

As iterations continue, one weight tends to unity while all other to zero. This is known as particle degeneracy and can be mitigated by resampling.

Resampling is undertaken if N^{eff}. In this paper, systematic resampling scheme outlined in [31] is employed. Samples are assigned an unnormalized weight of $\frac{1}{N}$ after resampling.

B. Particle Filter

The PF uses a set of N_x particles to recursively represent any nonlinear, non-Gaussian SSM as $N_x \to \infty$. At every time step t, particles are drawn from a proposal distribution, $q\left(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t}, \boldsymbol{\theta}\right)$ and weighted according to

$$\boldsymbol{w}_t^j = \boldsymbol{w}_{t-1}^j \frac{p\left(\boldsymbol{y}_t|\boldsymbol{x}_t^j, \boldsymbol{\theta}\right) p\left(\boldsymbol{x}_t^j|\boldsymbol{x}_{t-1}^j, \boldsymbol{\theta}\right)}{q\left(\boldsymbol{x}_t^j|\boldsymbol{x}_{t-1}^j, \boldsymbol{y}_t\right)}, \tag{15}$$

In this example, we set the transmission model to equal the proposal which simplifies the weight update in (15) to

$$\mathbf{w}_t^j = \mathbf{w}_{t-1}^j p\left(\boldsymbol{y}_t|\mathbf{x}_t^j, \boldsymbol{\theta}\right). \tag{16}$$

State estimates are made by a weighted sum of functions on the particles

$$\int p\left(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:t}, \boldsymbol{\theta}\right) f\left(\boldsymbol{x}_{1:t}\right) d\boldsymbol{x}_{1:t} \approx \sum\nolimits_{j=1}^{N_x} \tilde{\boldsymbol{w}}_t^j f\left(\boldsymbol{x}_{1:t}^j\right), \tag{17}$$

where the normalized weights are calculated as in (13) (with $\tilde{\boldsymbol{w}}_t^j$ in place of $\tilde{\boldsymbol{w}}_t^i$).

An unbiased estimate of the marginal likelihood can be obtained from an average of the unnormalized weights

$$p\left(\boldsymbol{y}_{1:t}|\boldsymbol{\theta}\right) = \int p\left(\boldsymbol{y}_{1:t}, \boldsymbol{x}_{1:t}|\boldsymbol{\theta}\right) d\boldsymbol{x}_{1:t} \approx \frac{1}{N_x} \sum\nolimits_{j=1}^{N_x} \boldsymbol{w}_t^j. \tag{18}$$

As with the SMC sampler, the ESS in the PF can be calculated using (14) and resampling employed if N_x^{eff} is less than $N_x/2$. In this paper, multinomial resampling is chosen within the PF. To keep the total unnormalized weight constant (such that the approximation (18) is the same immediately before and after resampling), each newly-resampled sample is assigned an unnormalized weight

$$\frac{1}{N_x} \sum\nolimits_{j=1}^{N_x} \boldsymbol{w}_t^j, \tag{19}$$

such that the normalized weights after resampling are $1/N_x$.

III. Concept Drift

Concept drift is a broad term that, in a general machine learning context, describes changes over time in the relationship between input data and the target variable. Survey papers have sought to characterize the various types of drift and the timescales on which they occur [16], [17].

A change in the parameters governing the dynamic system in (1) (i.e. $\boldsymbol{\theta}_{t-1} \neq \boldsymbol{\theta}_t$) can be thought of as an example of real concept drift. We define $\boldsymbol{\theta} = \{\boldsymbol{\phi}_t, \boldsymbol{\psi}\}$, with $\boldsymbol{\phi}_t$ denoting time-varying parameters and $\boldsymbol{\psi}$ non-time-varying parameters. Equation (1) can be extended by (20) to include effective dynamics on the parameter space

$$\boldsymbol{\phi}_t \mid \boldsymbol{\phi}_{t-1} \sim p(\boldsymbol{\phi}_t \mid \boldsymbol{\phi}_{t-1}, \boldsymbol{\tau}). \tag{20}$$

We parameterise an SSM as hierarchical, with equations (1), (2) and (20), and perform inference on the now static set of parameters $\boldsymbol{\theta} = \{\boldsymbol{\tau}, \boldsymbol{\psi}\}$.

Reference [8] proposes the concept drift models defined in (22)-(25) as ways of evolving a parameter to generate synthetic data. These models are analogous to the aforementioned types of drift timescales and are presented below:

1) Drift-less:
$$\boldsymbol{\phi}_t = l, \tag{21}$$

2) Incremental:
$$\boldsymbol{\phi}_t = l + u\boldsymbol{\tau}t, \tag{22}$$

3) Recurring:
$$\boldsymbol{\phi}_t = \frac{u-l}{2}\mathrm{s}(\frac{2\pi t}{\boldsymbol{\tau}}) + \frac{u+l}{2}, \tag{23}$$

4) Outlier:
$$\boldsymbol{\phi}_t = \begin{cases} l, & t = \boldsymbol{\tau} \\ u, & else \end{cases}, \tag{24}$$

5) Abrupt:
$$\boldsymbol{\phi}_t = \begin{cases} l, & t < \boldsymbol{\tau} \\ u, & else \end{cases}, \tag{25}$$

6) Gradual:
$$\boldsymbol{\phi}_t = \frac{u-l}{2}\mathrm{s}(\frac{2\pi t}{\boldsymbol{\tau}_0 + \boldsymbol{\tau}_1 t}) + \frac{u+l}{2}, \tag{26}$$

where $s(t)$ is the square wave function, u and l are the upper and lower values the parameter may take, respectively and $\boldsymbol{\tau}$ is the drift parameter of interest. A graphical representation of how these concept drift models change a parameter can be seen in Figure 1.

The red lines in Figure 1 show the different timescales on which concept drift occurs for $\boldsymbol{\beta}_t$ in the SEIS disease model outlined in Section IV. Figure 1(a) shows a non-time varying parameter (i.e. $\boldsymbol{\beta}_t$ is fixed for the duration of the simulation), Figure 1(b) depicts a sudden shift from one regime to another at a specific time point, while Figure 1(c) illustrates a gradual drift from one regime to another and back, with the parameter spending increasingly longer periods in the upper regime over time. In Figure 1(d), a slow linear variation in the parameter is displayed. Figure 1(e) depicts a sudden jump to another regime followed by an immediate return, while Figure 1(f) illustrates equal time spent in both regimes.

IV. SEIS model

The stochastic Susceptible, Exposed, Infected, Susceptible (SEIS) disease model utilized in this paper is a variant of the SIR model [32]. It simulates the movement of N_{pop} individuals between the Susceptible, Exposed and Infected latent state compartments which take values $\mathbf{S}_t, \mathbf{E}_t, \mathbf{I}_t \in [0..N_{pop}]$

The total count of individuals moving compartments is governed by the parameters $\boldsymbol{\theta} = \{\boldsymbol{\tau}, \boldsymbol{\gamma}, \boldsymbol{\sigma}\}$. We note that $\boldsymbol{\tau}$ is a static parameter which governs the rate of change of $\boldsymbol{\beta}_t$ i.e. $\phi_t = \beta_t$. The corresponding binomial distributions are

$$n(\mathbf{S} \to \mathbf{E}) \sim \text{Binomial}(\mathbf{S}_{t-1}, 1 - \exp\left(-\frac{\boldsymbol{\beta}_{t-1}\mathbf{I}_{t-1}\mathbf{S}_{t-1}}{N_{pop}}\right)), \tag{27}$$

$$n(\mathbf{E} \to \mathbf{I}) \sim \text{Binomial}(\mathbf{E}_{t-1}, 1 - \exp(-\boldsymbol{\gamma})), \tag{28}$$

$$n(\mathbf{I} \to \mathbf{S}) \sim \text{Binomial}(\mathbf{I}_{t-1}, 1 - \exp(\boldsymbol{\sigma})). \tag{29}$$

The complete discrete, stochastic SEIS model is

$$\mathbf{S}_t = \mathbf{S}_{t-1} - n(\mathbf{S} \to \mathbf{E}) + n(\mathbf{I} \to \mathbf{S}), \tag{30}$$

$$\mathbf{E}_t = \mathbf{E}_{t-1} + n(\mathbf{S} \to \mathbf{E}) - n(\mathbf{E} \to \mathbf{I}), \tag{31}$$

$$\mathbf{I}_t = \mathbf{I}_{t-1} + n(\mathbf{E} \to \mathbf{I}) - n(\mathbf{I} \to \mathbf{S}). \tag{32}$$

The likelihood is

$$\boldsymbol{y}_t | \boldsymbol{x}_t \sim \text{Poisson}\left(\boldsymbol{y}_t; \boldsymbol{I}_t\right). \tag{33}$$

V. Computational Setup

The experiments were conducted using the distributed memory SMC^2 framework outlined in [33]. The code is provided here[1]. The analysis presented in this paper was performed on a distributed memory cluster equipped with two Xeon Gold 6138 CPUs, providing 384GB of memory and 40 cores.

The experimental configuration comprises $T = 150$ observations, $N_x = 4096$ particles in the PF, $N = 256$ samples in the SMC sampler and $K = 50$ iterations. The step-size for the RW proposal was 0.1. The non-time varying parameters are $\boldsymbol{\gamma} = 0.1$ and $\boldsymbol{\sigma} = 0.1$ with priors $q_1(\boldsymbol{\theta}_1^i) = \Gamma(1, 0.1)$. The drift parameters have bounds of $l = 0.1$ and $u = 0.8$ and priors of $q_1(\boldsymbol{\theta}_1^i) = \Gamma(1, 0.1)$ and $q_1(\boldsymbol{\theta}_1^i) = \Gamma(8, 0.1)$. The drift parameters have true values of $\boldsymbol{\tau} = 0.3$, $[\frac{1}{3}, 0.3]$, $\frac{u-l}{T}$, 0.3, $\frac{1}{3}$ for abrupt, gradual, incremental, outlier and recurring scenarios, respectively. The prior distributions are $q_1(\boldsymbol{\theta}_1^i) = \Gamma(3, 0.1)$ for all scenarios except incremental drift which is given the prior $q_1(\boldsymbol{\theta}_1^i) = \Gamma(0.1, 0.1)$. The initial distribution of $N_{pop} = 1000$ individuals in each compartment at t_0 was drawn from

$$\text{Multinomial}(\boldsymbol{x}_0; N_{pop}, \boldsymbol{p}), \tag{34}$$

where $\boldsymbol{p} = [0.9, 0.05, 0.05]$.

[1]https://github.com/j-j-murphy/Explicit-Concept-Drift

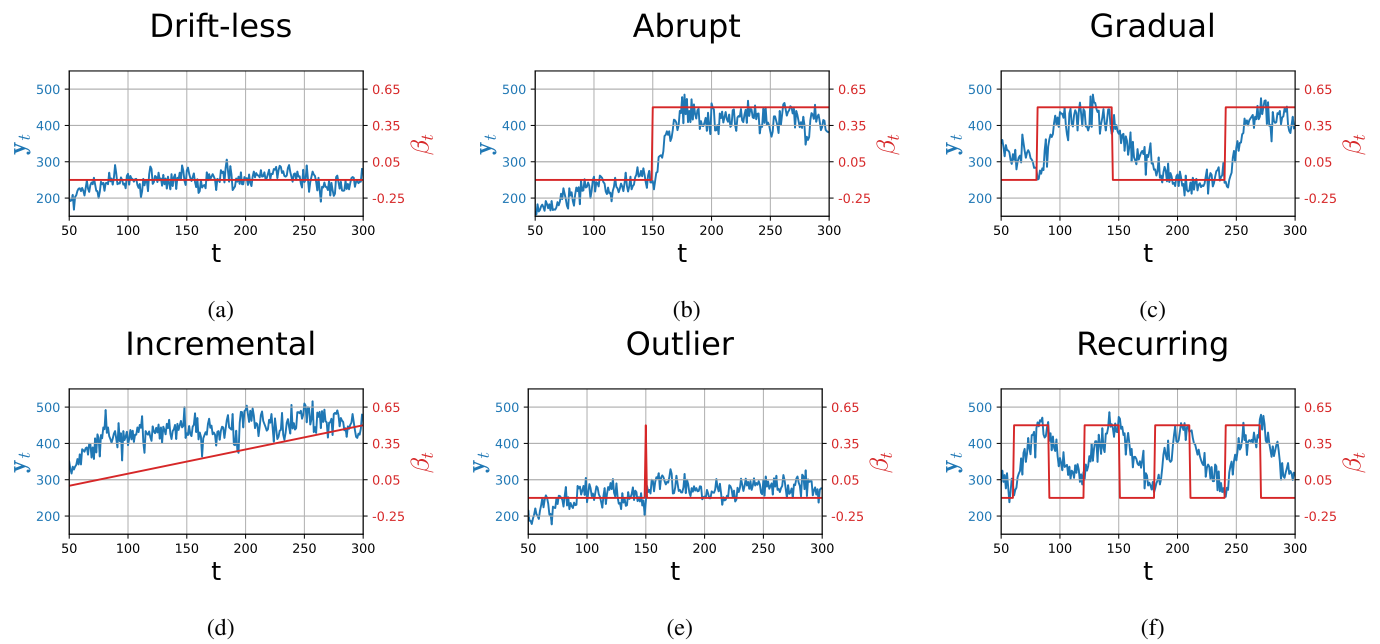

Fig. 1: Infection observations (blue) and corresponding β_t (red) generated for a trajectory of length $T = 300$ with $\gamma = 0.1$, $\sigma = 0.1$, $l = 0.2$ and $u = 0.8$. (a) In the drift-less scenario there is no τ (b) in abrupt $\tau = 0.5$, (c) in gradual $\tau = [0.2, 0.25]$, (d) in incremental $\tau = 0.005$, (e) in outlier $\tau = 0.5$ and (f) in the recurring drift scenario $\tau = 0.2$.

TABLE I: Mean Square Error (MSE) of different drift scenarios.

Scenario	Including Drift	Drift-less
Abrupt	1.655×10^3	5.502×10^4
Gradual	7.051×10^3	7.754×10^4
Incremental	5.791×10^4	3.426×10^4
Outlier	1.150×10^3	1.263×10^3
Recurring	1.922×10^4	6.745×10^4

VI. Results

Over 10 Monte carlo runs, the parameters when including drift ($\theta = \{\tau, \gamma, \sigma, u, l\}$) and when drift-less ($\theta = \{\gamma, \sigma, l\}$) were estimated with SMC^2 for each concept drift scenario. Using the parameter estimates, the PF predicts the states of the disease model, which can be seen in Figure 2. It is evident that when including drift, the PF can more accurately estimate the true states as outlined in Table I when comparing the Mean Square Error (MSE). However, this is not the case with incremental drift which actually performs worse than the drift-less PF. We hypothesis that the is due to the estimate of the susceptible compartment changing too quickly. This is likely due to the random walk proposal which is suitable for the other parameters but not the far finer-scaled τ parameter, giving an estimate an order of magnitude lower of 5.155×10^{-2} against a true value of 4.667×10^{-3}.

VII. Conclusion

In this paper, we explicitly define a hierarchical PF with time-varying parameters governed by concept drift models. We apply these methods to a compartmental disease model, demonstrating that the PF can track unobserved states and estimate parameters using SMC^2.

The scenarios in this paper are simplified, by assuming that the drift type is known apriori and considering only one type of drift at a time. Natural extensions include a model selection method like the likelihood ratio outlined in [34], modeling multiple types of drift simultaneously and exploring drift beyond the binary high-low framework. The SEIS example model is discrete so prevents the the use of improved, gradient-based proposals in the PF and SMC sampler as in [35], [36]. One method to overcome this would be to use HINTS [37]. Further verification of the approach could be found in applications to real world scenarios.

References

[1] A. Doucet, N. De Freitas, and N. Gordon, "An introduction to sequential monte carlo methods," *Sequential Monte Carlo methods in practice*, pp. 3–14, 2001.

[2] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 174–188, 2002.

[3] D. M. Sheinson, J. Niemi, and W. Meiring, "Comparison of the performance of particle filter algorithms applied to tracking of a disease epidemic," *Mathematical biosciences*, vol. 255, pp. 21–32, 2014.

[4] Q.-b. Zhang, P. Wang, and Z.-h. Chen, "An improved particle filter for mobile robot localization based on particle swarm optimization," *Expert Systems with Applications*, vol. 135, pp. 181–193, 2019.

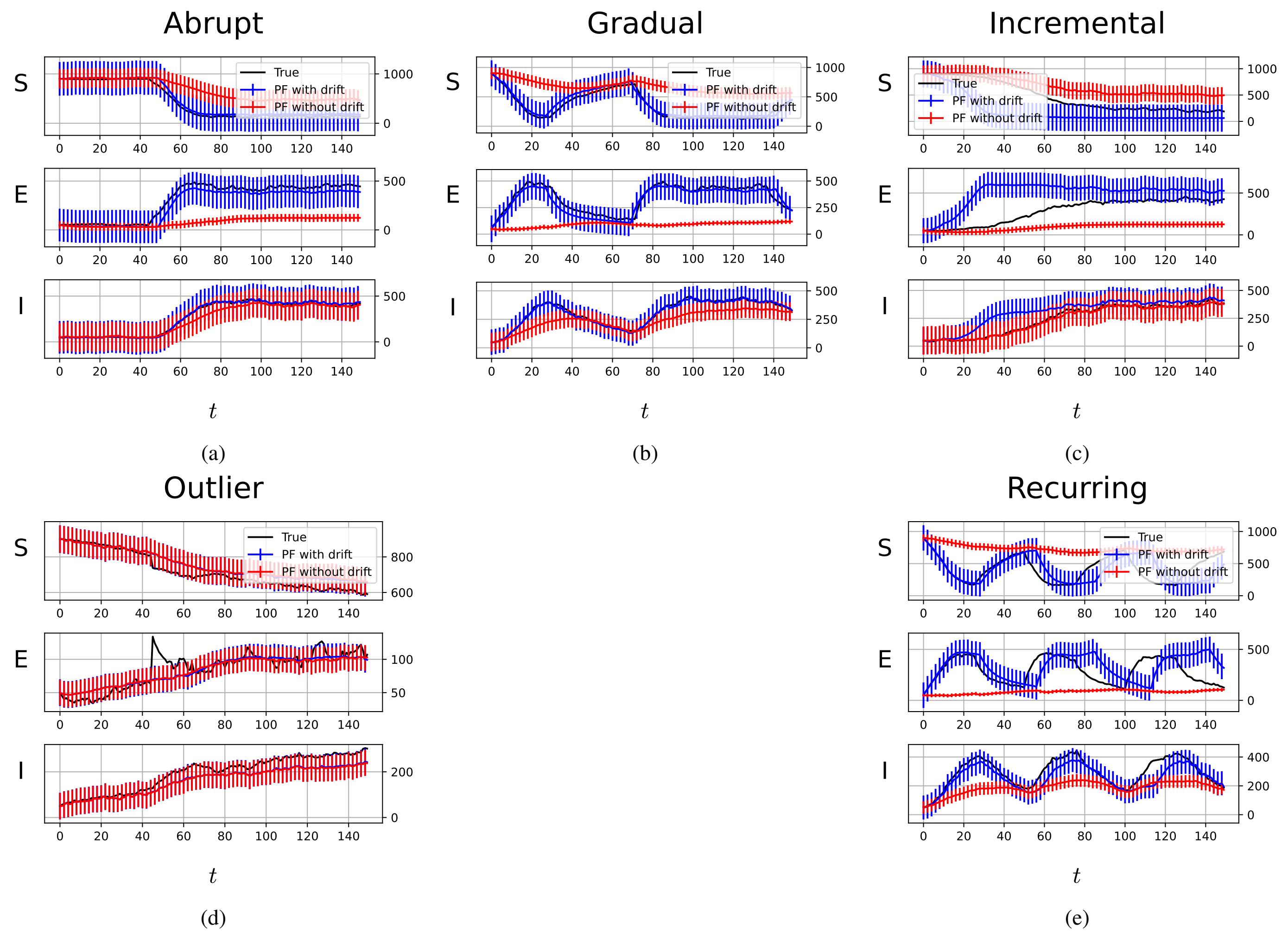

Fig. 2: State estimates of PFs over 5 Monte Carlo runs in (a) abrupt, (b) gradual, (c) incremental, (d) outlier and (e) recurring drift scenarios. The mean estimate is the blue line and the error bars give the standard deviation.

[5] P. M. Djuric, M. Vemula, and M. F. Bugallo, "Target tracking by particle filtering in binary sensor networks," *IEEE Transactions on signal processing*, vol. 56, no. 6, pp. 2229–2238, 2008.

[6] N. Kantas, A. Doucet, S. S. Singh, J. Maciejowski, and N. Chopin, "On Particle Methods for Parameter Estimation in State-Space Models," *Statistical Science*, vol. 30, no. 3, pp. 328–351, 2015. DOI: 10.1214/14-STS511. [Online]. Available: https://doi.org/10.1214/14-STS511.

[7] G. Kitagawa, "A self-organizing state-space model," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1203–1215, 1998, ISSN: 01621459. [Online]. Available: http://www.jstor.org/stable/2669862 (visited on 05/24/2024).

[8] A. Arnold, "When artificial parameter evolution gets real: Particle filtering for time-varying parameter estimation in deterministic dynamical systems," *Inverse Problems*, vol. 39, no. 1, p. 014 002, 2022.

[9] C. Nemeth, P. Fearnhead, and L. Mihaylova, "Sequential monte carlo methods for state and parameter estimation in abruptly changing environments," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1245–1255, Mar. 2014, ISSN: 1941-0476. DOI: 10.1109/tsp.2013.2296278. [Online]. Available: http://dx.doi.org/10.1109/TSP.2013.2296278.

[10] M. Satoh, P. J. Van Leeuwen, and S. Nakano, "Online state and time-varying parameter estimation using the implicit equal-weights particle filter," *Quarterly Journal of the Royal Meteorological Society*, 2024.

[11] T. Matsumoto and K. Yosui, "Adaptation and change detection with a sequential monte carlo scheme," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 3, pp. 592–606, 2007.

[12] D. Crisan and J. Miguez, *Nested particle filters for online parameter estimation in discrete-time state-space markov models*, 2017. arXiv: 1308.1883 [stat.CO].

[13] S. Pérez-Vieites, I. P. Mariño, and J. Míguez, "Probabilistic scheme for joint parameter estimation and state

prediction in complex dynamical systems," *Phys. Rev. E*, vol. 98, p. 063 305, 6 Dec. 2018. DOI: 10.1103/PhysRevE.98.063305. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.98.063305.

[14] C. Andrieu, A. Doucet, and R. Holenstein, "Particle markov chain monte carlo methods," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 72, no. 3, pp. 269–342, 2010.

[15] N. Chopin, P. E. Jacob, and O. Papaspiliopoulos, "Smc2: An efficient algorithm for sequential analysis of state space models," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 75, no. 3, pp. 397–426, 2013.

[16] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM computing surveys (CSUR)*, vol. 46, no. 4, pp. 1–37, 2014.

[17] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning under concept drift: A review," *IEEE transactions on knowledge and data engineering*, vol. 31, no. 12, pp. 2346–2363, 2018.

[18] S. Liu, S. Xue, J. Wu, *et al.*, "Online active learning for drifting data streams," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 1, pp. 186–200, 2021.

[19] S. Mohamad, M. S. Mouchaweh, and A. Bouchachia, "Active learning for data streams under concept drift and concept evolution," in *STREAMEVOLV at ECML-PKDD*, 2016. [Online]. Available: https://api.semanticscholar.org/CorpusID:32173629.

[20] L. Chen, Y. Zhou, and J. Yang, "Object tracking within the framework of concept drift," in *Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5-9, 2012, Revised Selected Papers, Part III 11*, Springer, 2013, pp. 152–162.

[21] R. Fok, A. An, and X. Wang, "Mining evolving data streams with particle filters," *Computational Intelligence*, vol. 33, no. 2, pp. 147–180, 2017.

[22] D. J. Morris, "Inference on historical ebola outbreaks using hierarchical models: A particle filtering approach," Ph.D. dissertation, 2021.

[23] T. Lazebnik, "Computational applications of extended sir models: A review focused on airborne pandemics," *Ecological Modelling*, vol. 483, p. 110 422, 2023.

[24] A. Gill, J. Koskela, X. Didelot, and R. G. Everitt, "Bayesian inference of reproduction number from epidemiological and genetic data using particle mcmc," *arXiv preprint arXiv:2311.09838*, 2023.

[25] G. Storvik, A. Diz-Lois Palomares, S. Engebretsen, *et al.*, "A sequential monte carlo approach to estimate a time-varying reproduction number in infectious disease models: The covid-19 case," *Journal of the Royal Statistical Society Series A: Statistics in Society*, vol. 186, no. 4, pp. 616–632, 2023.

[26] X. Liu and P. Stechlinski, "Infectious disease models with time-varying parameters and general nonlinear incidence rate," *Applied Mathematical Modelling*, vol. 36, no. 5, pp. 1974–1994, 2012.

[27] P. Birrell, J. Blake, E. Van Leeuwen, N. Gent, and D. De Angelis, "Real-time nowcasting and forecasting of covid-19 dynamics in england: The first wave," *Philosophical Transactions of the Royal Society B*, vol. 376, no. 1829, p. 20 200 279, 2021.

[28] J. Dureau, K. Kalogeropoulos, and M. Baguelin, "Capturing the time-varying drivers of an epidemic using stochastic dynamical systems," *Biostatistics*, vol. 14, no. 3, pp. 541–555, 2013.

[29] A. Doucet, A. M. Johansen, *et al.*, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of nonlinear filtering*, vol. 12, no. 656-704, p. 3, 2009.

[30] P. L. Green, L. Devlin, R. E. Moore, R. J. Jackson, J. Li, and S. Maskell, "Increasing the efficiency of sequential monte carlo samplers through the use of approximately optimal l-kernels," *Mechanical Systems and Signal Processing*, vol. 162, p. 108 028, 2022.

[31] A. Varsi, S. Maskell, and P. G. Spirakis, "An o (log2 n) fully-balanced resampling algorithm for particle filters on distributed memory architectures," *Algorithms*, vol. 14, no. 12, p. 342, 2021.

[32] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.

[33] C. Rosato, A. Varsi, J. Murphy, and S. Maskell, "An o(log2 n) smc2 algorithm on distributed memory with an approx. optimal l-kernel," in *2023 IEEE Symposium Sensor Data Fusion and International Conference on Multisensor Fusion and Integration (SDF-MFI)*, 2023, pp. 1–8. DOI: 10.1109/SDF-MFI59545.2023.10361452.

[34] N. J. Gordon, S. Maskell, and T. Kirubarajan, "Efficient particle filters for joint tracking and classification," in *Signal and Data Processing of Small Targets 2002*, O. E. Drummond, Ed., International Society for Optics and Photonics, vol. 4728, SPIE, 2002, pp. 439–449. DOI: 10.1117/12.478524. [Online]. Available: https://doi.org/10.1117/12.478524.

[35] C. Rosato, J. Harris, J. Panovska-Griffiths, and S. Maskell, "Inference of stochastic disease transmission models using particle-mcmc and a gradient based proposal," in *2022 25th International Conference on Information Fusion (FUSION)*, IEEE, 2022, pp. 1–8.

[36] A. Varsi, L. Devlin, P. Horridge, and S. Maskell, "A general-purpose fixed-lag no-u-turn sampler for nonlinear non-gaussian state space models," *IEEE Transactions on Aerospace and Electronic Systems*, 2024.

[37] M. Strens, "Efficient hierarchical mcmc for policy search," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 97.