ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

Predicting Grain Protein Content of Winter Wheat

Mansouri Majdi, Dumont Benjamin, Destain Marie-France

Département des Sciences et Technologies de l'Environnement, GxABT,
Université de Liège, 2 Passage des Déportés, 5030 Gembloux, Belgium

Abstract. The objective of this paper is to propose to use a new Improved Particle Filtering (IPF) based on minimizing Kullback-Leibler divergence for crop models' predictions. The performances of the method are compared with those of the conventional Particle Filtering (PF) at a complex crop model (AZODYN) to predict an important winter-wheat quality criterion, namely the grain protein content. Furthermore, the effect of measurement noise (e.g., different signal-to-noise ratios) on the performances of PF and IPF is investigated. The results of the comparative studies show that the IPF provides a significant improvement over the PF because, unlike the PF which depends on the choice of sampling distribution used to estimate the posterior distribution, the IPF yields an optimum choice of the sampling distribution, which also accounts for the observed data. The efficiency of IPF is expressed in terms of estimation accuracy (root mean square error).

1. Introduction

Dynamic crop models such as EPIC [1], SALUS [2], and STICS [3] are non-linear models that describe the growth and development of a crop interacting with environmental factors (soil and climate) and agricultural practices (crop species, tillage type, fertilizer amount…). They are developed to predict crop yield and quality or to optimize the farming practices in order to satisfy agricultural objectives, as the reduction of nitrogen lixiviation. More recently, crop models are used to simulate the effects of climate changes on the agricultural production. Nevertheless, the prediction errors of these models may be important due to uncertainties in the estimates of initial values of the states, in input data, in the parameters, and in the equations. The measurements needed to run the model are sometimes not numerous, whereas the field spatial variability and the climatic temporal fluctuations over the field may be high. The degree of accuracy is therefore difficult to estimate, apart from numerous repetitions of measurements. For these reasons, the problem of state/parameter estimation represents a key issue in such nonlinear and non-Gaussian crop models including a large number of parameters, while measurement noise exists in the data.

For example, it is useful to predict the evolution of variables, such as the biomass and the grain protein content during the crop lifecycle. State estimation techniques can be of a great value to solve that problem since they have the potential to estimate simultaneously the variables and several parameters. As an example, involved parameters are the radiation use efficiency, the maximal value of the ratio of intercepted to incident radiation, the coefficient of extinction of radiation, the maximal value of LAI. Several estimation techniques, such as Particle filtering [4] method has been developed and utilized in many applications. PF methods approximate the posterior probability distribution by a set of weighted samples, called

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7. Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

particles. Since real world problems usually involve high dimensional random variables with complex uncertainty, the nonparametric and sample-based estimation of uncertainty has thus become quite popular to capture and represent the complex distribution in nonlinear and non-Gaussian models [5]. PF methods offer a number of significant advantages over other conventional methods. However, since they use the prior distribution as the importance distribution [6], the latest data observation is not considered and not taken into account when evaluating the weights of the particles. While the importance sampling distribution has computational advantages, it can cause filtering divergence. In cases where the likelihood distribution is too narrow compared to the prior distribution, few particles will have significant weights. Hence, a better proposal distribution that takes the latest observation data into account is needed. In other words, new adaptive methods that incorporate better feedback and smoothing in the selection or deletion of particles and their weights need to be investigated. The objectives of this paper are twofold. The first objective is to develop an improved Particle filtering (IPF) for improving nonlinear and non-Gaussian crop model predictions. In case of standard PF, the latest observation is not considered for the evaluation of the weights of the particles as the importance function is taken to be equal to the prior density function. This choice of importance sampling function simplifies the computation but can cause filtering divergence. In cases where the likelihood function is too narrow compared to the prior distribution, very few particles will have significant weights. Hence, a better proposal distribution that takes the latest observation into account is needed. The main novelty of this task is to develop new Bayesian algorithm for nonlinear and non-Gaussian state and parameter estimation with better proposal distribution based on minimizing Kullback-Leibler divergence. The second objective is to apply the state estimation techniques PF and IPF for predicting grain protein content. We present an application of the IPF for updating predictions of complex nonlinear crop models in order to predict protein grain content. The rest of the paper is organized as follows. In Section 2, a description of state estimation technique for nonlinear crop model prediction is presented. Then, in Section 3, the performances of the state estimation techniques are evaluated and compared through the application case. Finally, some concluding remarks are presented in Section 4.

2. Improved Particle Filtering Description

The choice of optimal proposal function is one of the most critical design issues in importance sampling schemes. In [7], the optimal proposal distribution $\hat{p}(z_k | z_{0:k-1}, y_{0:k})$ is obtained by minimizing the variance of the importance weights given the states $z_{0:k-1}$ and the observations data $y_{0:k}$. This selection has also been studied by other researchers. However, this optimal choice suffers from one major drawback. The particles are sampled from the prior density $p(z_k | z_{0:k-1})$ and the integral over the new state need to be computed. In the general case, closed form analytic expression of the posterior distribution of the state is untractable [8]. Therefore, the distribution $p(z_k | z_{0:k-1})$ is the most popular choice of proposal distribution. One of its advantages is its simplicity in sampling from the prior functions $p\left(z_k \middle| z_{0:k-1}\right)$ and the evaluation of weights $l_k^{(i)}$ (as presented in the previous section). However, the latest observation is not considered for the computation of the weights of the particles as the importance density is taken to be equal to the prior density ([8]). The transition prior $p\left(z_k \middle| z_{0:k-1}\right)$does not take into account the current observation data y_k, and many particles can be wasted in low likelihood areas. This choice of importance sampling function simplifies the computational complexity but can cause filtering divergence [8]). In cases where the likelihood density is too narrow as compared to the prior function, very few particles will have considerable weights. Next, we present an overview of KLD-based improved particle filter.

2.1 Improved Particle Filter based on KLD minimization

As mentioned above, the distribution of interest for the state takes the form of a marginal posterior distribution $p\left(z_k \middle| y_{0:k}\right)$. The proposed extended Bayesian sampling algorithm (also named as improved particle filtering, IPF) is proposed for approximating intractable integrals arising in Bayesian statistics. By using a separable approximating distribution $\hat{q}(z_k) = \hat{p}\left(z_k \middle| z_{0:k-1}, y_{0:k}\right) = \Pi_i\, p(z_k^i)$ to lower bound the marginal likelihood, an analytical approximation to the posterior probability $p\left(z_k \middle| y_{0:k}\right)$ is provided by minimizing the Kullback-Leibler divergence (KLD):

$$D_{KL}\left(\hat{q} \| p\right) = \int \hat{q}\left(z_k\right) \log \frac{\hat{q}\left(z_k\right)}{p\left(z_k \middle| z_{0:k-1}, y_{0:k}\right)} dz_k \tag{1}$$

where, $$\hat{q}(z_k) = \Pi_i \hat{q}\left(z_k^i \middle| z_{0:k-1}, y_{0:k}\right) = \hat{q}(z_k)\hat{q}(\mu_k)\hat{q}(\lambda_k), \tag{2}$$

μ_k is the expectation of z_k and λ_k is the covariance matrix of z_k.

Minimizing the KLD subject to the constraint $\int q(z_k) dz_k = \Pi_i \int q(z_k^i) dz_k^i = 1$, the Lagrange multiplier scheme is used to yield the following approximate distribution,

$$\hat{q}(z_k^i) \propto \exp\left\lfloor E\left(\log\left(p(y_{0:k}, z_k)\right)\right)_{\Pi j \neq i\, \hat{q}(z_k^j)} \right\rfloor \tag{3}$$

where $E\left(\log\left(p(y_{0:k}, z_k)\right)\right)_{\Pi j \neq i\, \hat{q}(z_k^j)}$ denotes the expectation operator relative to the distribution $\hat{q}(z_k^j)$. Therefore, these dependent parameters can be jointly and iteratively updated. Taking into account the separable approximate distribution $\hat{q}(z_{k-1}^i)$ at time $k-1$, the posterior distribution $p(z_k | y_{0:k})$ is sequentially approximated according to the following scheme:

$$\hat{p}(z_k | y_{0:k}) \propto p(y_k | z_k)\, p(z_k, \lambda_k | \mu_k)\, q_p(\mu_k) \tag{4}$$

where, $$q_p(\mu_k) = \int p(\mu_k | \mu_{k-1}) \hat{q}(\mu_{k-1}) d\mu_{k-1}$$

Hence, the particles $\left\{z_{0:k}^{(i)}, l_k^{(i)}\right\}_{i=0}^{N}$ (where $l_k^{(i)}$ denotes the importance weight of the sample $z_{0:k}^{(i)}$ and N is the total number of the samples) are sampled according to the following optimal function:

$$\hat{q}(z_k^i) = \int \mathrm{N}(z_k^i \,|\, \mu_k, \lambda_k)\, p(\mu_k, \lambda_k \,|\, z_k^i) p(z_k \,|\, z_k^i) \hat{q}(\mu_{k-1}) d\mu_k d\lambda_k \tag{5}$$

The recursive estimate of the importance weights can be derived as follows:

$$l_k^{(i)} = l_k^{(i-1)} \frac{p(y_{0:k} \,|\, z_{0:k}^{(i)}) p(z_k^{(i)} \,|\, z_{0:k-1}^{(i)})}{\hat{q}(z_{0:k}^{(i)} \,|\, y_{0:k})} \tag{6}$$

Equation (6) provides a mechanism to sequentially update the importance weights, given an appropriate choice of proposal distribution, $\hat{q}(z_{0:k}^{(i)} \,|\, y_{0:k})$. Then, the estimate of the augmented state $\hat{z}_k$ can be approximated by a Monte Carlo scheme as follows:

$$\hat{z}_k = \sum_{i=0}^{N} l_k^{(i)} z_k^{(i)} \tag{7}$$

3. Simulation Results Analysis

3.1 Application to a crop model predicting grain protein content

The AZODYN crop model ([9]) is a nonlinear dynamic model simulating winter-wheat crop in function of environmental variables (characteristics of the crop at the end of winter, soil characteristics, climate) and of nitrogen fertilization (dates and rates of fertilizer applications). We consider a particular site-year (2008-2009) This model can be used to predict grain yield, soil mineral nitrogen, and grain protein content at harvest. AZODYN is a useful tool for studying the effects of nitrogen management on crop yield, grain quality and risk of pollution by nitrate ([10]). Before flowering, five state variables are simulated each day by AZODYN: nitrogen uptake (NU), dry matter (DM), nitrogen-nutrition index (NNI), leaf-area index (LAI), soil mineral nitrogen supply (SNS). We consider chlorophyll-content measurements obtained with a chlorophyll meter. These measurements are correlated to one of the model state variables, namely nitrogen uptake, and can be easily performed by farmers, collecting-firm operators, or farmers' advisors. Here, we suppose that only one chlorophyll-content measurement is performed at flowering and that this measurement is linearly related to the model state variables as follows:

$$\mathrm{y}_{m_k} = \mu + H\mathrm{x}_{m_k} + \mathrm{v}_{m_k} \tag{8}$$

where y_{m_k} and x_{m_k} are, respectively, the chlorophyll-content measurement and the (5×1) vector of the true state-variable values at flowering, μ is an intercept parameter, and H is a one-row matrix defined by $H = (\alpha, 0, 0, 0)$ where α is the slope of the linear equation relating the measurement to nitrogen uptake. We assume that the error term v_{m_k} is normally distributed, $v_{m_k} \sim N(0, R)$. The IPF is used to

ESANN 2014 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 23-25 April 2014, i6doc.com publ., ISBN 978-287419095-7.
Available from http://www.i6doc.com/fr/livre/?GCOI=28001100432440.

update the five states variables nitrogen uptake (NU), dry matter (DM), nitrogen-nutrition index (NNI), leaf-area index (LAI), soil mineral nitrogen supply (SNS) given a single chlorophyll-content measurement y_{m_k} performed at flowering. Yield and grain protein content at harvest are then estimated from the updated state variables. Eventually, to perform comparison between the techniques, the estimation root mean square errors (RMSE) criteria will be used and calculated on the states (with respect to the noise free data).

$$RMSE = \sqrt{E((z - \hat{z})^2)} \tag{9}$$

Where z (resp. $\hat{z}$) is the true parameter/state (resp. the estimated parameter/state). Figures 1, 2 and Table 1 show the estimation of the two states variables Yield and grain protein content using PF and IPF. The results show the performance of IPF over PF, the efficiency of IPF is due to the fact it uses the KLD divergence to compute the optimum sampling distribution used to approximate the posterior density function, which also accounts for the observed data.

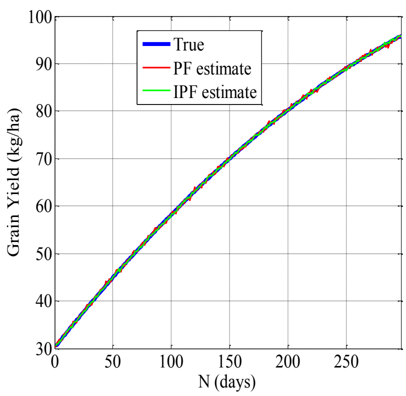

Fig. 1: Updated value of grain protein content (kg/ha) (Days) using PF and IPF techniques

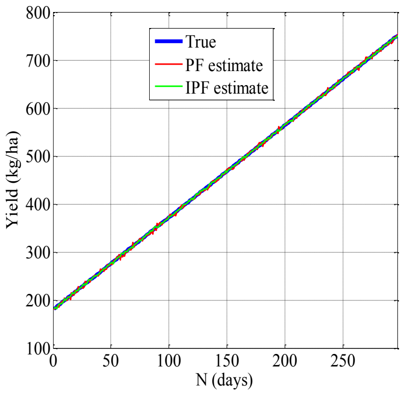

Fig. 2: Updated value of yield (kg/ha) versus N versus N (days) using PF and IPF techniques.

Technique	ERROR	
	Yield	grain protein content
PF	1.0761	0. 0622
IPF	0.4376	0.0192

Table 1: ERROR of estimated states.

4. Conclusions

In this paper, we developed a state estimation techniques for crop model prediction. In the case study, we have used Bayesian methods PF and IPF for updating predictions of complex nonlinear crop models. In this case, the proposed IPF is applied at a complex crop model (AZODYN) to predict an important winter-wheat quality criterion, namely the grain protein content. The results of the comparative study show that the IPF provides a significant improvement over the PF because, unlike the PF which depends on the choice of sampling distribution used to estimate the posterior distribution, the IPF yields an optimum choice of the sampling distribution, which also accounts for the observed data. The performance of PF and IPF is evaluated on a synthetic example in terms of estimation accuracy, and root mean square error.

References

[1] J. Williams, C. Jones, J. Kiniry, and D. Spanel, “The Epic crop growth model,” Trans. ASAE, vol. 32, no. 2, pp. 497–511, 1989.
[2] B. Basso and J. Ritchie, “Impact of compost, manure and inorganic fertilizer on nitrate leaching and yield for a 6-year maize-alfalfa rotation in michigan,” Agriculture, ecosystems & environment, vol. 108, no. 4, pp. 329–341, 2005.
[3] N. Brisson, B. Mary, D. Ripoche, M. Jeuffroy, F. Ruget, B. Nicoullaud, P. Gate, F. Devienne-Barret, R. Antonioletti, C. Durr et al., “Stics: a generic model for the simulation of crops and their water and nitrogen balances. i. theory, and parameterization applied to wheat and corn,” Agronomie, vol. 18, no. 5-6, pp. 311–346, 1998.
[4] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking,” Signal Processing, IEEE Transactions on, vol. 50, no. 2, pp. 174–188, 2002.
[5] P. Djuric, M. Vemula, and M. Bugallo, “Target tracking by particle filtering in binary sensor networks,” IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2229–2238, 2008.
[6] E. Wan and R. V. D. Merwe, “The unscented kalman filter for nonlinear estimation,” Adaptive Systems for Signal Processing, Communications, and Control Symposium, pp. 153–158, 2000.
[7] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking,” Signal Processing, IEEE Transactions on, vol. 50, no. 2, pp. 174–188, 2002.
[8] A. Doucet and V.B. Tadi´c. Parameter estimation in general state-space models using particle methods. *Annals of the institute of Statistical Mathematics*, 55(2):409–22, 2003.
[9] M.-H. Jeuffroy and S. Recous, “Azodyn: a simple model simulating the date of nitrogen deficiency for decision support in wheat fertilization,” European journal of Agronomy, vol. 10, no. 2, pp. 129–144, 1999.
[10] C., Varlet-Grancher, R., Bonhomme, M., Chartier, et al., Efficience de la conversion de l'énergie solarie par un couvert végétal. Acta Oecologica, Oecologia Plantarum, 3 (1), 3-26, 1982.