Analysis of 5 Loxin® Treatment for Patients with Osteoarthritis in Clinical Trial using Power Filter

Sourish Das, Dipak Dey

Technical Report #2008-09
October 10, 2008

Statistical and Applied Mathematical Sciences Institute
PO Box 14006
Research Triangle Park, NC 27709-4006
www.samsi.info

Analysis of 5 Loxin® Treatment for Patients with Osteoarthritis in Clinical Trial using Power Filter

Sourish Das*   and   Dipak Dey

Abstract

5 Loxin® is a new and novel treatment for the Osteoarthritis (OA) of knee. In this paper, we present the efficacy study of 5 Loxin® treatment for OA and explain the underlying mechanism of a treatment to be successful for treating OA such as 5 Loxin®. In a double-blind, randomized, placebo-controlled, clinical trial, the modulation of pro-inflammatory bio-molecules like tumor necrosis factor α (TNFα) and interleukin-1β (IL-1β) were evaluated in the OA patients. We present the analysis of association between TNFα and IL-1β with the effect of 5 Loxin® treatment for OA. Based on this analysis, we describe the fundamental system of a successful treatment for treating OA, such as 5 Loxin®; otherwise any new therapy for OA in future would fail, if they fail to follow the mechanism. We accomplish this by introducing the power filter (PF) for dynamic generalized linear models, which extends the usual Kalman filter (KF) for dynamic linear models. We establish a relationship between the KF and the PF as well. An information processing optimality property of the PF is presented; which shows PF is optimum filter like KF. This optimum property gives PF an edge over the other suboptimal filters, such as extended Kalman filter.

KEY WORDS: Boswellia serrata; Dynamic model; Information processing; Kalman filter; TNF alpha .

1   Introduction

Osteoarthritis (OA) is the most common form of inflammatory joint disease characterized by articular cartilage degradation with an accompanying peri-articular bone response. OA affects nearly 21 million people in United States, accounting for 25% visits to primary care physicians (Felson, D. T, 2004). Clinical manifestations of OA of knee are pain in and around the joint, stiffness of the joint after rest, crepitation on motion and limited joint motion (Hochberg, M. C *et al* 1995). Recent recommendations for managing OA focus on relieving pain and stiffness and improving physical function as important objective of treatment (Pendleton *et al* 2000). Currently available medication of OA for most

cases include nonsteroidal anti-inflammatory drugs (NASIDs), including cyclo-oxygenase II inhabitors. This can reduce pain and inflammation effectively. However, long-term use of NASIDs has been found to be associated with adverse side-effect, including hypertension, congestive heart failure and renal insufficiency. Because of high incidence of adverse events associated with currently available treatment, an effective and safer alternative treatment for OA are needed.

In recent years, the gum resin extracted from an ancient herb Boswellia (*Boswellia serrata*, often known as "Indian Frankincense") has gained good attention as a potent anti-inflammatory, anti-arthritic and analgesic agent (Wright, L.M. 2002). 5-Loxin® is a novel treatment enriched with *Boswellia serrata* extract (US Patent 2004/0073060A1) which confers a significant improvement in treating OA . Pertinent references are Roy, *et al.* (2005) and Sengupta *et al.* (2008). Cell based *in vitro* studies and *in vivo* experiments conducted in Sprague Dawley rats demonstrate that 5-Loxin® potentially inhibits the pro-inflammatory cytokines such as tumor necrosis factor-α (TNFα). A recent study (Moxley *et al.* 2007) showed an association with OA in hand and the human interleukin-1 (IL-1) region, more specifically with IL-1β. Question is: can 5-Loxin® inhibits IL-1β also?

In the present research, a double-blind and placebo controlled clinical trial, was conducted to evaluate the mechanism of the effectiveness of the 5-Loxin® for treating of OA of knee. It is important to understand the mechanism of a successful treatment, such as 5 Loxin®, to prevent treatment failure. We assess the effectiveness of 5-Loxin® on the symptoms of pain, joint stiffness, mobility in OA patients. We also look at the association study of potent pro-inflammatory agent TNFα and IL-1β, which add the second line of evidence in support of likely molecular mechanism of 5-Loxin® in reducing pain through lowering the pro-inflammatory agent.

In this *longitudinal study*, each individual patients are followed over time and responses of interest, that is pain score in *visual analogue scale* (VAS) are recorded together with covariates information of biomarker measurement for pro-inflammatory agent TNFα and IL-1β. Here we have observed a group of 70 patients of OA, who are randomly assigned into two different levels of doses (either 100 mg/day or 250 mg/day) of 5 Loxin® and placebo. The observations are taken at $q(=5)$ points in time, t_1=Day 0, t_2 = Day 7, t_3 = Day 30, t_4 = Day 60 and t_5 = Day 90. Each time we measure patients' pain score in VAS together with their biomarker measurement for TNFα and IL-1β, as covariates. The major advantage of such longitudinal design is its capacity to separate the cohort effect from time effect.

Although longitudinal data are not comparable to typical long chain time series data; many ideas of analyzing time series data can be incorporated in analyzing such longitudinal study. One of the common approaches to model time series of counts, proportions, compositions and positive observations is exponential family state space dynamic model. Several researchers, including West, Harrison, Migon (1985), Harvey and Fernades (1989) and others, used exponential-conjugate family as the basis of Bayesian time series analysis. Many researchers used *power steady model* (PSM) of Smith (1979, 1981) as a basis for state space modeling of non-Gaussian time series. Harvey and Fernandes (1989) extended the idea of PSM to include explanatory variable in GLM frame work using the maximum likelihood method of fitting. Grunwald, Guttorp and Raftery (1993) have given some general results for estimation and forecasting. If the goal is filtering, (estimating an unobservable signal $p(\theta_t \mid D_t)$ for instance) Grunwald, Hamza and Hyndman (1997) recommended to use PSM type Bayesian models.

In this paper, we present the method for modeling longitudinal studies as state space dynamic model. In order to achieve this we extend the horizon of PSM to the dynamic generalized linear models (DGLM); hence naming it as "power filter" (PF) which extends the usual "Kalman filter" (KF) for dynamic linear models (DLM). The format of the paper is as follows. In section 2, we introduce the idea of PF for DGLM. In section 3, we present the PF for binomial response longitudinal data with covariates information. In section 4, we present the detail analysis of 5 Loxin® treatment using PF for dynamic logistic regression model. Section 5 concludes the paper with brief discussion and in appendix, we showed optimal property of PF and discuss the connection between KF and PF.

2   Power Filter for Longitudinal Data

First we introduce the *observe part* of the longitudinal system. Suppose the distribution of $y_i^{(t)}$ belong to the natural exponential family with parameters $\theta_i^{(t)}$ and $\phi^{(t)}$, *i.e.*,

$$f(y_i^{(t)} \mid \theta_i^{(t)}) \propto exp\Big\{ \frac{1}{\phi^{(t)}}\Big(y_i^{(t)}\theta_i^{(t)} - \psi(\theta_i^{(t)})\Big)\Big\}, \quad i = 1, ..., n;$$

where $y_i^{(t)}$ is the longitudinal response at stage t, $\theta_i^{(t)}$ is the canonical parameter at stage t, also known as *state of nature* and $\phi^{(t)}$ is known as dispersion parameter. Now the posterior distribution of θ_i at stage $(t-1)$ is

$$\pi\Big(\theta_i^{(t-1)} \mid y_i^{(t-1)}\Big) \propto exp\Big\{ \frac{1}{\tau_i^{(t-1)}}\Big(\mu_{0i}^{(t-1)}\theta_i^{(t-1)} - \psi(\theta_i^{(t-1)})\Big)\Big\},$$

where $E\Big(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)}\Big) = \mu_{0i}^{(t-1)}$ and $\tau_i^{(t-1)}$ is the corresponding convolution parameter at stage $(t-1)$; see Diaconis and Ylvisker (1979) for clarification.

The marginal distribution

$$p(y_i^{(s)} \mid y_i^{(t-1)}) = \int_\Theta f(y_i^{(s)} \mid \theta_i^{(s)}, y_i^{(t-1)})\pi(\theta_i^{(s)} \mid y_i^{(t-1)})d\theta_i^{(s)}, \qquad (2.1)$$

(where $s = t-1$ or t) is important since it does not condition on the unobservable state $\theta_i^{(s)}$. It is predictive distribution when $s = t$. The marginal mean $E[y_i^{(s)}]$ is defined using equation (2.1) (with $s = t-1$) and is predictive mean when $s = t$. Following the Diaconis and Ylvisker (1979) and Grunwald *et al* (1993) it follows that the marginal mean satisfies

$$E\big[y_i^{(t-1)}\big] = E\Big(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)}\Big) = \mu_{0i}^{(t-1)}.$$

We introduce the *system equation* as,

$$\psi'\big(\theta_i^{(t)}\big) = \rho_t \psi'\big(\theta_i^{(t-1)}\big), \qquad (2.2)$$

where ρ_t is known quantity. One step predicted mean and variances are respectively

$$E\Big(\psi'(\theta_i^{(t)}) \mid y_i^{(t-1)}\Big) = \rho_t E\Big(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)}\Big) = \rho_t \mu_{0i}^{(t-1)} = \tilde{\mu}_{0i}^{(t)}$$

and

$$V\left(\psi'(\theta_i^{(t)}) \mid y_i^{(t-1)}\right) = \rho_t^2 V\left(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)}\right) = \rho_t^2 V_{(t-1)i}.$$

We define $\tilde{\tau}_i^{(t)} = (\rho_t^2 \tau_i^{(t-1)} + a_t)$; where $a_t \geq 0$ is a known constant. The *prior distribution* of θ_i at stage t is defined as

$$\tilde{\pi}(\theta_i^{(t)}) = [\pi(\theta_i^{(t)})]^\delta, \qquad \text{such that} \quad 0 \leq \delta \leq 1,$$

where

$$\pi(\theta_i^t) \propto \exp\left\{ \frac{1}{\tilde{\tau}_i^{(t)}} \left(\tilde{\mu}_{0i}^{(t)} \theta_i^{(t)} - \psi(\theta_i^{(t)}) \right) \right\}.$$

Thus $\tilde{\pi}(\theta_i^{(t)})$ can be viewed as power prior, see Chen, Ibrahim and Shao (2000), Ibrahim, Chen and Sinha (2003), where the data up to stage $(t-1)$ can be considered as past (or historical) information, such that $\tilde{\mu}_0^{(t)}$ is the prior mean or predictive mean at stage t with prior convolution parameter $\frac{\tilde{\tau}_i^{(t)}}{\delta}$. Therefore, *posterior distribution* of θ_i at stage t is

$$\pi(\theta_i^{(t)} \mid y_i^{(t)}) \propto f(y_i^{(t)} \mid \theta_i^{(t)})\tilde{\pi}(\theta_i^{(t)}) \tag{2.3}$$

$$\propto exp\left\{ \frac{1}{\tau_i^{(t)}} \left(\mu_{0i}^{(t)} \theta_i^{(t)} - \psi(\theta_i^{(t)}) \right) \right\}.$$

The *posterior mean* at stage t is then given as

$$\mu_{0i}^{(t)} = E\left(\psi'(\theta_i^{(t)} \mid y_i^{(t)})\right) = \tilde{\mu}_{0i}^{(t)} + \frac{\tilde{\tau}_i^{(t)}}{\delta\phi^{(t)} + \tilde{\tau}_i^{(t)}}(y_i^{(t)} - \tilde{\mu}_{0i}^{(t)}). \tag{2.4}$$

Since we use the idea of the *power steady model* and the *power prior* to update from one step to another, we call this as *power filter* (PF). Note that filter mean equals the prediction mean plus correction depending on how much the new observation differs from its prediction. Also the *posterior variance* at stage t is

$$\tau_i^{(t)} = \left(\frac{1}{\phi^{(t)}} + \frac{\delta}{\tilde{\tau}_i^{(t)}} \right)^{-1} = \frac{1}{\delta}\left[\tilde{\tau}_i^{(t)} - \frac{\tilde{\tau}_i^{(t)2}}{\tilde{\tau}_i^{(t)} + \delta\phi^{(t)}} \right].$$

The systematic component of DGLM at stage t is

$$\theta_i^{(t)} = x_i^{(t)'} \beta^{(t)},$$

which in matrix notation is

$$\theta^{(t)} = X^{(t)} \beta^{(t)}.$$

Under the full column rank assumption of $X^{(t)}$, we can write $\theta^{(t)}$ as

$$\theta^{(t)} = P^{(t)} \theta^{(t)} + (I - P^{(t)})\theta^{(t)},$$

where $P^{(t)} = X^{(t)}(X^{(t)'} X^{(t)})^{-1} X^{(t)'}$. Note that $P^{(t)}$ is linear transformation matrix representing the orthogonal projection from $n-$dimensional space $\mathbf{R}^n$ onto canonical space $C(\theta^{(t)})$, while $(I - P^{(t)})$ represents the orthogonal projection of $\mathbf{R}^n$ onto the $C^\perp(\theta^{(t)})$. Hence we can write $\beta^{(t)}$ as

$$\beta^{(t)} = (X^{(t)'} X^{(t)})^{-1} X^{(t)'} \theta^{(t)}. \tag{2.5}$$

Thus we can generate samples from the posterior distribution of $\beta^{(t)}$ using Monte Carlo algorithm described below. Advantage of this algorithm is, as long as we know the posterior distribution of $\theta^{(t)}$, we do not need to know the posterior distribution of $\beta^{(t)}$.

Monte Carlo Algorithm:

1. Suppose we are at the rth iteration of stage t. Generate sample $\theta_i^{(t)r}$ from $\pi(\theta_i^{(t)} \mid y_i^{(t)})$ in (2.3) for $i = 1, ..., n$.

2. Calculate $\beta^{(t)r} = (X^{(t)'}X^{(t)})^{-1}X^{(t)'}\theta^{(t)r}$ for $r = 1, 2, ..., N$; where N is the simulation size.

Once we have the posterior samples of $\beta^{(t)}$, $\{\beta^{(t)r} \mid r = 1, ..., N\}$; we can do all the necessary inferences on $\beta^{(t)}$. Note that one can also implement the regular MCMC algorithm to draw samples from $\pi(\beta^{(t)} \mid y^{(t)}, X^{(t)})$.

2.1   Information Processing and Choice of δ

The power parameter δ is playing an interesting role of "discount factor" by incorporating the extra uncertainty with the change. Such extra uncertainty provides the required adaptibility and robustness, see West and Harrision (1997). Some researchers show concern for the *power steady model* and *power prior* as well, because of the ambiguity embedded in the choice of δ. Some approaches are available to go around this problem. One approach, along the line of Ibrahim *et al* (2003), is as follows: assume δ as unknown parameter and specify prior on δ. It can be shown that the PF minimizes $E(K_\pi)$, where expectation is taken with respect to prior of δ. In this paper, we suggest the second approach which is more appropriate for longitudinal study. Define

$$\delta = f(\mid t_i - t_j \mid), \qquad i \neq j = 1, 2, ..., n \quad \text{such that}$$
$$(i) \ \ 0 \leq \delta \leq 1,$$
$$(ii) \mid t_i - t_j \mid \to 0 \Rightarrow \delta \to 1,$$
$$(iii) \mid t_i - t_j \mid \to \infty \Rightarrow \delta \to 0.$$

Such specification of δ has a natural interpretation, because δ determines that how much information should thrive from one stage to another. Clearly, $\delta = 0$ means, no information is borrowed from the previous stage; while $\delta = 1$ implies that all the information is being used from the previous time point. Hence we should define δ as a function of difference between two time points. If time gap, between two observation, is small - then we should borrow as much information as possible from the previous time point. However, if the time gap is large then we should borrow less information from the previous time point. Based on the above mentioned definition of δ, some functions $\{f(.) \mid f(\mid t_i - t_j \mid), \mid t_i - t_j \mid \geq 0\}$ could be considered as

1. $f(\mid t_i - t_j \mid) = \frac{1}{2} - \frac{1}{\pi}\arctan(log(\mid t_i - t_j \mid))$, where π is universal constant 3.14159.

2. $f(\mid t_i - t_j \mid) = \frac{1}{1+\mid t_i - t_j \mid}$,

3. $f(\mid t_i - t_j \mid) = exp\{- \mid t_i - t_j \mid\}$,

However, while designing the longitudinal study, one should decide the function in advance or one can fit the model using different $\delta = f(\mid t_i - t_j \mid)$ and then apply Bayesian model monitoring technique as described in West (1986) and West and Harrision (1997).

2.2 Model Assessment

A simple method of monitoring the performance of a class of Bayesian models was introduced by West (1986) using Bayes factor. However recent development of deviance information criterion (DIC) (Spiegelhalter *et al.* 2002) provides us to measure the model complexity and effective number of parameters in a model at time t. DIC at time t is

$$DIC^{(t)} = 2\hat{D}_{avg}^{(t)} - D_{\hat{\theta}}^{(t)},$$

where $\hat{D}_{avg}^{(t)} = \frac{1}{L}\sum_{l=1}^{L} D(y^{(t)}, \theta_l^{(t)})$, $\theta_l^{(t)}$'s are the samples from the posterior distribution of θ at time t and $D(y^{(t)}, \theta^{(t)}) = -2.log\ f(y^{(t)} \mid \theta^{(t)})$ is the corresponding deviance function of time t. Measures of model complexity at time t is:

$$p_{D^{(t)}} = \hat{D}_{avg}^{(t)} - D_{\hat{\theta}}^{(t)}.$$

Note that intuitively we can say that as trial proceeds and observe more data on the subjects, based on the Bayesian learning as we learn more about the system, the predictive prior should be efficient enough to mimic the system accurately. As a result measure of model complexity $p_D^{(t)}$ should decease over time. But if the system experiences any change then measure of model complexity will increase. Efficacy study of 5 Loxin® treatment required dynamic logistic regression model to implement the analysis. In the next section we present the PF method for dynamic logistic regression model.

3 Power Filter for Dynamic Logistic Regression Model

Here we present the PF for binomial response model. *Observed part* at stage t: Suppose $y_i^{(t)} \sim Bin(n_i^{(t)}, p_i^{(t)})$, where $\theta_i^{(t)} = log\left(\frac{p_i^{(t)}}{1-p_i^{(t)}}\right)$, such that $\phi^{(t)} = 1$ and $\psi(\theta_i^{(t)}) = n_i^{(t)}log(1 + e^{\theta_i^{(t)}})$; $n_i^{(t)}$ is known and we have,

$$f(y_i^{(t)} \mid \theta_i^{(t)}) \propto \exp\{y_i^{(t)}\theta_i^{(t)} - n_i^{(t)}log(1 + e^{\theta_i^{(t)}})\}.$$

Since we restrict ourselves within the natural exponential conjugate family model, the posterior distribution at stage $(t-1)$ is:

$$p_i^{(t-1)} \mid y_i^{(t-1)} \sim Beta(a_i^{(t-1)}, b_i^{(t-1)}),$$

which in turn,

$$\pi(\theta^{(t-1)} \mid y_i^{(t-1)}) \propto \exp\{\frac{1}{\tau_i^{(t-1)}}(\mu_{0i}^{(t-1)}\theta_i^{(t-1)} - n_i^{(t-1)}log(1 + e^{\theta_i^{(t-1)}}))\},$$

where $\tau_i^{(t-1)} = \frac{n_i^{(t-1)}}{a_i^{(t-1)}+b_i^{(t-1)}}$, $\mu_{oi}^{(t-1)} = \frac{n_i^{(t-1)}a_i^{(t-1)}}{a_i^{(t-1)}+b_i^{(t-1)}} = n_i^{(t-1)}E(p_i^{(t-1)} \mid y_i^{(t-1)})$; *i.e.*, $E(p_i^{(t-1)} \mid y_i^{(t-1)}) = \frac{a_i^{(t-1)}}{a_i^{(t-1)}+b_i^{(t-1)}}$. The marginal predictive distribution is the beta-binomial distribution and the *system equation* of the model from (2.2) is $n_i^{(t)}p_i^{(t)} = \rho_t n_i^{(t-1)}p_i^{(t-1)}$, where ρ_t is a known constant. Hence, $E(p_i^{(t)} \mid y_i^{(t-1)}) = \rho_t\frac{n_i^{(t-1)}}{n_i^{(t)}}E(p_i^{(t-1)} \mid y_i^{(t-1)}) = \frac{\tilde{\mu}_{0i}^{(t)}}{n_i^{(t)}}$ and

$$Var(n_i^{(t)}p_i^{(t)} \mid y_i^{(t-1)}) = \rho_t^2 Var(n_i^{(t-1)}p_i^{(t-1)} \mid y_i^{(t-1)}) = \rho_t^2 V_{(t-1)i}.$$

This implies $Var(p_i^{(t)} \mid y_i^{(t-1)}) = \rho_t^2 \frac{n_i^{(t-1)2}}{n_i^{(t)2}} Var(p_i^{(t-1)} \mid y_i^{(t-1)})$. The updated convolution parameter at stage t is: $\tilde{\tau}_i^{(t)} = (\rho_t^2 \tau_i^{(t-1)} + a_t)$, $a_t \geq 0$ is a known constant and the updated prior distribution of θ_i at stage t is:

$$\tilde{\pi}(\theta_i^{(t)}) \propto \exp\left\{\frac{\delta}{\tilde{\tau}_i^{(t)}}((\tilde{\mu}_{0i}^{(t)}\theta_i^{(t)} - n_i^{(t)}log(1 + e^{\theta_i^{(t)}}))\right\},$$

where δ is the precision parameter such that $\delta = f(\mid t_i - t_j \mid)$, $0 \leq \delta \leq 1$. The posterior distribution at stage t is:

$$\pi(\theta_i^{(t)} \mid y_i^{(t)}) \propto \exp\left\{\left(y_i^{(t)} + \frac{\delta\tilde{\mu}_{0i}^{(t)}}{\tilde{\tau}_i^{(t)}}\right)\theta_i^{(t)} - \left(1 + \frac{\delta}{\tilde{\tau}_i^{(t)}}\right)log(1 + e^{\theta_i^{(t)}})\right\},$$

where the posterior mean at stage t from (2.4) is:

$$E(p_i^{(t)} \mid y_i^{(t)}) = \frac{\tilde{\mu}_{0i}^{(t)}}{n_i^{(t)}} + \frac{\tilde{\tau}_i^{(t)}}{\delta + \tilde{\tau}_i^{(t)}}\left(\frac{y_i^{(t)}}{n_i^{(t)}} - \frac{\tilde{\mu}_{0i}^{(t)}}{n_i^{(t)}}\right).$$

Thus for longitudinal studies, we have the dynamic logistic regression model as: $log\left(\frac{p_i^{(t)}}{1-p_i^{(t)}}\right) = \theta_i^{(t)} = x_i^{(t)'}\beta^{(t)}$. We can obtain samples from the posterior distribution of $\theta_i^{(t)}$ and can use the Monte Carlo algorithm, as describe in the section 2, to generate samples from the posterior distribution of $\beta^{(t)}$. Along the same line we can easily develop the PF method for dynamic Poisson regression model for count response longitudinal data. In appendix B we also established a relationship between power filter and Kalman filter.

In the next section we presented the analysis of the efficacy study of 5 Loxin® treatment with the association of TNFα and IL-1β for patients with OA of knee using the PF for the dynamic logistic regression model.

4   Analysis of the Efficacy Study of 5 Loxin®

The objective of this study was to reveal if 5 Loxin® is effective in reducing the pain and reveal the mechanism of its success. Pain score VAS lies between 0 and 100. So we categorize the pain score as strong and mild pain using the following scheme:

$$y_i^{(t)} = \begin{cases} 1 & \text{if VAS score for } i^{th} \text{ patient is greater than 50 at time point } t, \\ 0 & \text{otherwise.} \end{cases}$$

The dynamic logistic regression model for our case is,

$$\begin{aligned} logit(p_i^{(t)}) = \beta_1^{(t)} + \beta_2^{(t)}I(Low\ dose) + \beta_3^{(t)}I(High\ dose) \\ + \beta_4^{(t)}log(TNF\alpha_i^{(t)}) + \beta_5^{(t)}log(IL1\beta_i^{(t)}) \\ + \beta_6^{(t)}log(TNF\alpha_i^{(t)}) \times log(IL1\beta_i^{(t)}) \end{aligned}$$

where $I(A) = 1$ if the patient i belongs to group A or 0 otherwise. We implement the analysis of this study using the PF method for the dynamic logistic regression model. We fit the model using all three different choices of δ mentioned in section 2.1 and we also

fit the model by fixing the δ to be equal to 1. By fixing δ to be equal to 1, we are not allowing additional variability in the prior. We consider month as an unit of time for this analysis.

It is clear from figure 1, that for the model with $\delta = 1$, $DIC^{(t)}$ decreases quickly after first month. However, $p_D^{(t)}$, the measure of model complexity over time, diminishes most quickly for the model with $\delta = \exp\{- \mid t_i - t_j \mid\}$, which shows the robustness and adaptability of the model to the data. From this point all the analysis, we presented in this paper, are based on the model with $\delta = \exp\{- \mid t_i - t_j \mid\}$.

Based on the analysis as presented in the Table 1, we can say that by the 7th day of the treatment the high dose of 5 Loxin® treatment is significantly effective in reducing the pain. By the end of the first month of treatment both the low and high doses are significantly effective in reducing the pain. Again if we look at the figure 2, we can conclude that 5 Loxin® treatment is effective in reducing the pain from OA of knee. By the end of the study, we found that effect of TNFα over the pain is statistically significant with 95% credible interval $(1.085, 2.707)$ which does not include 0. In addition a positive estimate shows that there is a positive association between the pain from OA of knee and TNFα. We also found similar and significantly positive association between the pain from OA of knee and IL-1β. The most interesting feature we found from this analysis is that by the end of the study, the interaction effect between TNFα and IL-1β has a statistically significant effect over the pain from OA of knee. In order to understand the behavior of the statistically significant interaction effect between TNFα and IL-1β, we plot the interaction response surface for baseline and day 90 in figure 3. It is clear that at the end of the study period, patients with lower levels of both TNFα and IL-1β have lower probability of high pain. However if patient experiences higher level of any one of TNFα or IL-1β then those patients have higher probability of stronger pain. An additional regular repeated measure analysis (not presented here because of space constrained) and corresponding figure 4 from that analysis confirms that 5 Loxin® statistically significantly reduce both pro-inflammatory agent TNFα and IL-1β. Therefore we can successfully conclude that 5 Loxin® treatment reduces pain by reducing the pro-inflammatory agent TNFα and IL-1β simultaneously.

5   Conclusion

In this paper, we presented the method for modeling longitudinal studies as state space dynamic model. We implement this by introducing the PF for DGLM, which extends the usual KF for dynamic linear models. We establish a relationship between the KF and the PF. We showed (in appendix) that for a longitudinal design, PF yields 100% efficient Zellner's information processing rule. More importantly, we presented the analysis of the efficacy study of 5 Loxin®, which is enriched with an extract of ancient herb *Boswellia Serrata*. We implemented the analysis using the PF for dynamic logistic regression model. We presented the analysis of association between two important pro-inflammatory agent TNFα and IL-1β with the effect of 5 Loxin® treatment for OA of knee. We also explained the underlying mechanism of successful treatment, such as 5 Loxin®; otherwise any new therapy in future would fail, if they fail to follow the mechanism. We showed that any treatment would be effective if it can reduce the levels of both TNFα and IL-1β simultaneously. If therapy fails to reduce any one of them then that might leads to treat-

ment failure. An additional regular repeated measure ANOVA analysis confirms that 5 Loxin® significantly reduce the levels of both TNFα and IL-1β. This leads us to conclude that 5 Loxin® treatment reduces pain in OA of knee by reducing the pro-inflammatory agents TNFα and IL-1β simultaneously in patients. Therefore, in order to avoid treatment failure, in future any new therapy for OA should try to reduce the TNFα and IL-1β simultaneously.

Appendix A: Optimality of Power Filter and Information Processing

The main idea of PF is to propagate information from one stage to another. The posterior distribution at stage t is:

$$\pi(\theta^{(t)} \mid y^{(t)}) \propto f(y^{(t)} \mid \theta^{(t)})\tilde{\pi}(\theta^{(t)})$$
$$\propto f(y^{(t)} \mid \theta^{(t)})[\pi(\theta^{(t)})]^{\delta}, \text{ such that } 0 \leq \delta \leq 1.$$

We consider two extreme scenarios. If $\delta = 0$ then

$$\pi_0(\theta^{(t)} \mid y^{(t)}) \propto f(y^{(t)} \mid \theta^{(t)}),$$

and if $\delta = 1$ then

$$\pi_1(\theta^{(t)} \mid y^{(t)}) \propto f(y^{(t)} \mid \theta^{(t)})\pi(\theta^{(t)}).$$

Clearly, π_0 is not propagating any information from the previous stage and π_1 is propagating all the information from the previous stage. Now we assume the power parameter δ is fixed. Later we will relax this assumption. The PF can be justified as the minimizer of convex sum of Kullback Leibler (KL) divergence between the posterior densities π_0 and π_1.

Definition of KL divergence: If p and q are two densities with respect to Lebesgue measure, then KL-divergence between p and q is defined as

$$KL(p,q) = \int log\Big(\frac{p(\theta)}{q(\theta)}\Big)p(\theta)d\theta.$$

Theorem 5.1 *The density $\pi \equiv \pi(\theta^{(t)})$ that minimizes*

$$K_{\pi} = (1-\delta)KL(\pi,\pi_0) + \delta \ KL(\pi,\pi_1)$$

is

$$\pi_{opt} \propto f(y^{(t)} \mid \theta^{(t)})\pi(\theta^{(t)})^{\delta}.$$

Proof: We can write,

$$K_\pi = (1-a)K(\pi, \pi_0) + a\ K(\pi, \pi_1)$$

$$= (1-a)\int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi_0(\theta^{(t)})}\right)d\theta^{(t)} + a\ \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi_1(\theta^{(t)})}\right)d\theta^{(t)}$$

$$= \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi_0(\theta^{(t)})}\right)^{1-a}d\theta^{(t)} + \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi_1(\theta^{(t)})}\right)^{a}d\theta^{(t)}$$

$$= \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi_0(\theta^{(t)})^{1-a}\pi_1(\theta^{(1)})^a}\right)d\theta^{(t)}$$

$$= K\left(\pi(\theta^{(t)}), \frac{\pi_0(\theta^{(t)})^{1-a}\pi_1(\theta^{(t)})^a}{C}\right) - log(C),$$

where $C = \int \pi_0(\theta^{(t)})^{1-a}\pi_1(\theta^{(t)})^a$ is the normalizing constant. Now K_π is minimized when

$$\pi_{opt} = \frac{1}{C}f(y^{(t)} \mid \theta^{(t)})\pi(\theta^{(t)})^\delta.$$

Following Zellner(1988, 2002), the functional that analyzes the information processing for any longitudinal study is:

$$\Delta[\pi(\theta^{(t)} \mid y^{(t)})] = \text{Output Information} - \text{Input Information}$$

$$= \int \pi(\theta^{(t)})log\left(\pi(\theta^{(t)})\right)d\theta^{(t)} + \int \pi(\theta^{(t)})log\left(h(y^{(t)})\right)d\theta^{(t)} -$$

$$w_1 \int \pi(\theta^{(t)})log\left(f(y^{(t)} \mid \theta^{(t)})\right)d\theta^{(t)} - w_2 \int \pi(\theta^{(t)})log\left(\pi^*(\theta^{(t)})\right)d\theta^{(t)},$$

where $\pi^*(\theta^{(t)})$ is any prior density of θ at stage t, $\pi(\theta^{(t)})$ is proper probability density at stage t, $0 \le w_1 \le 1$ and $0 \le w_2 \le 1$. In our case we choose $w_1 = 1$ and $w_2 = \delta$ and $h(y^{(t)}) = \int f(y^{(t)} \mid \theta^{(t)})\pi^*(\theta^{(t)})d\theta^{(t)}$ is free of θ and $\int \pi(\theta^{(t)})d\theta^{(t)} = 1$

$$\Delta[\pi(\theta^{(t)} \mid y^{(t)})] = \int \pi(\theta^{(t)})log\left(\pi(\theta^{(t)})\right)d\theta^{(t)} + log\left(h(y^{(t)})\right) -$$

$$\int \pi(\theta^{(t)})log\left(f(y^{(t)} \mid \theta^{(t)})\right)d\theta^{(t)} - \delta \int \pi(\theta^{(t)})log\left(\pi^*(\theta^{(t)})\right)d\theta^{(t)}.$$

We can write $log\left(h(y^{(t)})\right)$ as

$$log\left(h(y^{(t)})\right) = (1-\delta)log(h^0(y^{(t)})) + \delta log(h^1(y^{(t)})),$$

where $h^0(y^{(t)}) = \int f(y^{(t)} \mid \theta^{(t)})d\theta^{(t)}$ and $h^1(y^{(t)}) = \int f(y^{(t)} \mid \theta^{(t)})\pi^*(\theta^{(t)})d\theta^{(t)}$. Hence we can write $\Delta[\pi(\theta^{(t)})]$ as

$$\Delta[\pi(\theta^{(t)})] = (1-\delta)\int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})h^0(y^{(t)})}{f(y^{(t)|\theta^{(t)}})}\right)d\theta^{(t)} + \delta \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})h^1(y^{(t)})}{f(y^{(t)|\theta^{(t)}})\pi^*(\theta^{(t)})}\right)d\theta^{(t)}$$

$$= (1-\delta)\int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi^0(\theta^{(t)})}\right)d\theta^{(t)} + \delta \int \pi(\theta^{(t)})log\left(\frac{\pi(\theta^{(t)})}{\pi^1(\theta^{(t)})}\right)d\theta^{(t)} = K_\pi.$$

This shows that $\Delta[\pi(\theta^{(t)})]$ and K_π have the same minimizer and summarizing these in the following theorem as:

Theorem 5.2 *For any longitudinal design, the* power filter *yields 100% efficient Zellner's Information Processing Rule (ZIPR).*

Appendix B: Relation Between Power Filter and Kalman Filter

Here we establish the relationship between PF for DGLM and KF for DLM. First we present the observed part at stage t. Suppose $y_i^{(t)} \sim Normal(\theta_i^{(t)}, \phi^{(t)})$, *i.e.*,

$$f(y_i^{(t)} \mid \theta_i^{(t)}) \propto \exp \left\{ \frac{1}{\phi^{(t)}} \left(y_i^{(t)} \theta_i^{(t)} - \frac{\theta_i^{(t)2}}{2} \right) \right\}.$$

In matrix notation, $Y^{(t)} \sim Normal(\theta^{(t)}, \Phi^{(t)})$. The posterior distribution at stage $(t-1)$ is

$$\pi(\theta^{(t-1)} \mid y_i^{(t-1)}) \propto \exp \left\{ \frac{1}{\tau_i^{(t-1)}} \left(\mu_{0i}^{(t-1)} \theta_i^{(t-1)} - \frac{\theta_i^{(t-1)2}}{2} \right) \right\},$$

where

$$E\left(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)} \right) = \mu_{0i}^{(t-1)} \Rightarrow E\left(\theta_i^{(t-1)2} \mid y_i^{(t-1)} \right) = \mu_{0i}^{(t-1)}$$

and $\tau_i^{(t-1)}$ is the dispersion parameter at stage $(t-1)$. We obtain the *system equation* of the PF same as that of the KF from Meinhold and Singpurwalla (1983), *i.e.*, system equations for KF is $\theta_i^{(t)} = \rho_t \theta_i^{(t-1)}$, where ρ_t is known constant and the updated mean and variances are $E(\theta_i^{(t)} \mid y_i^{(t-1)}) = \rho_t E(\theta_i^{(t-1)} \mid y_i^{(t-1)}) = \rho_t \mu_{0i}^{(t-1)} = \tilde{\mu}_{0i}^{(t)}$ and $Var(\theta_i^{(t)} \mid y_i^{(t-1)}) = \rho_t^2 Var(\theta_i^{(t-1)} \mid y_i^{(t-1)}) = \rho_t^2 \tau_i^{(t-1)}$. We can present the same in matrix notation as, $Var(\theta^{(t)} \mid y^{(t-1)}) = \rho_t \Sigma_\tau^{(t-1)} \rho_t'$. Now we have $\tilde{\tau}_i^{(t)} = (\rho_t^2 \tau_i^{(t-1)} + a_t)$, $a_t \geq 0$ is known constant. The prior distribution of θ_i at stage t is:

$$\tilde{\pi}(\theta_i^{(t)}) \propto \exp \left\{ \frac{\delta}{\tilde{\tau}_i^{(t)}} \left(\tilde{\mu}_{0i}^{(t)} \theta_i^{(t)} - \frac{\theta_i^{(t)2}}{2} \right) \right\},$$

where δ is the precision parameter, $0 \leq \delta \leq 1$. In matrix notation, the prior distribution at stage t is: $\theta^{(t)} \sim N(\tilde{\mu}_0^{(t)}, R_t)$, where $R_t = \delta^{-1}[\rho_t \Sigma_\tau^{(t-1)} \rho_t' + A_t]$. Note that $\delta = 0$ will lead to non-informative improper prior distribution for stage t. The posterior distribution at stage t is:

$$\pi(\theta_i^{(t)} \mid y_i^{(t)}) \propto \exp \left\{ \left(\frac{y_i^{(t)}}{\phi^{(t)}} + \frac{\delta \tilde{\mu}_{0i}^{(t)}}{\tilde{\tau}_i^{(t)}} \right) \theta_i^{(t)} - \left(\frac{1}{\phi^{(t)}} + \frac{\delta}{\tilde{\tau}_i^{(t)}} \right) \frac{\theta_i^{(t)2}}{2} \right\},$$

where $\theta_i^{(t)} \sim N(\mu_{0i}^{(t)}, \tau_i^{(t)})$ such that $\mu_{0i}^{(t)} = \tilde{\mu}_{0i}^{(t)} + \frac{\tilde{\tau}_i^{(t)}}{\tilde{\tau}_i^{(t)} + \delta \phi^{(t)}} (y_i^{(t)} - \tilde{\mu}_{0i}^{(t)})$ and $\tau_i^{(t)} = \frac{1}{\delta} \left[\tilde{\tau}_i^{(t)} - \frac{\tilde{\tau}_i^{(t)2}}{\tilde{\tau}_i^{(t)} + \delta \phi^{(t)}} \right]$. Therefore, in matrix notation, the posterior at stage t is: $\theta^{(t)} \sim N(\mu_0^{(t)}, \Sigma^{(t)})$, where $\mu_0^{(t)} = \tilde{\mu}_0^{(t)} + R_t (\delta \Phi^{(t)} + R_t)^{-1} (Y^{(t)} - \tilde{\mu}_0^{(t)})$, and $\Sigma^{(t)} = \delta^{-1} [R_t - R_t'(R_t + \delta \Phi^{(t)})^{-1} R_t]$. Now if $y_i^{(t)} \sim N(\theta_i^{(t)}, \tau^{(t)})$, such that $\theta_i^{(t)} = x_i^{(t)'} \beta^{(t)}$, *i.e.*, $\theta^{(t)} = X^{(t)} \beta^{(t)}$ and $\delta = 1$. Then following Menihold and Singpurwalla (1983), we have the Kalman filter as $\beta_0^{(t)} = \tilde{\beta}_0^{(t)} + R_t X_t'(\Phi^t + X_t R_t X_t')^{-1} (Y^{(t)} - X^{(t)} \tilde{\beta}_0^{(t)})$ and $\Sigma_\beta^{(t)} = R_t - R_t' X_t'(V_t + X_t R_t X_t')^{-1} X_t R_t'$.

Appendix C: Power Filter for Dynamic Poisson Regression Model

Here we present the power filter for Poisson response model. *Observed part* at stage t: Suppose $y_i^{(t)} \sim Poisson(\lambda_i^{(t)})$, such that $\theta_i^{(t)} = log(\lambda_i^{(t)})$ and $\phi^{(t)} = 1$, $\psi(\theta_i^{(t)}) = e^{\theta_i^{(t)}}$. Hence we have,

$$f(y_i^{(t)} \mid \theta_i^{(t)}) \propto \exp\left\{ y_i^{(t)} \theta_i^{(t)} - e^{\theta_i^{(t)}} \right\}.$$

The posterior distribution at stage $(t-1)$ is: $\pi(\theta^{(t-1)} \mid y_i^{(t-1)}) \propto \exp\left\{ \frac{1}{\tau_i^{(t-1)}} (\mu_{0i}^{(t-1)} \theta_i^{(t-1)} - e^{\theta_i^{(t)}}) \right\}$,

where

$$E\left(\psi'(\theta_i^{(t-1)}) \mid y_i^{(t-1)} \right) = \mu_{0i}^{(t-1)} \Rightarrow E\left(e^{\theta_i^{(t-1)}} \mid y_i^{(t-1)} \right) = \mu_0^{(t-1)} \Rightarrow E(\lambda_i^{(t-1)} \mid y_i^{(t-1)}) = \mu_{0i}^{(t-1)}$$

and $\tau_i^{(t-1)}$ is the convolution parameter at stage $(t-1)$. The predictive distribution is negative binomial and the *system equation* of the model from (2.2) is:

$$e^{\theta_i^{(t)}} = \rho_t e^{\theta_i^{(t-1)}} \quad \Rightarrow \quad \lambda_i^{(t)} = \rho_t \lambda_i^{(t-1)}, \quad \text{where } \rho_t \text{ is known constant.}$$

Similar to logistic regression model, it can be shown that $E(\lambda_i^{(t)} \mid y_i^{(t-1)}) = \tilde{\mu}_{0i}^{(t)}$ and

$$Var(\lambda_i^{(t)} \mid y_i^{(t-1)}) = \rho_t^2 Var(\lambda_i^{(t-1)} \mid y_i^{(t-1)}).$$

Now $\tilde{\tau}^{(t)} = (\rho_t^2 \tau_i^{(t-1)} + a_t)$, $a_t \geq 0$ is known constant. Hence prior distribution of θ_i at stage t is:

$$\tilde{\pi}(\theta_i^{(t)}) \propto \exp\left\{ \frac{\delta}{\tilde{\tau}_i^{(t)}} (\tilde{\mu}_{0i}^{(t)} \theta_i^{(t)} - e^{\theta_i^{(t)}}) \right\},$$

where δ is the precision parameter, with $0 \leq \delta \leq 1$. Now posterior distribution at stage t is

$$\pi(\theta_i^{(t)} \mid y_i^{(t)}) \propto \exp\left\{ \left(y_i^{(t)} + \frac{\delta \tilde{\mu}_{0i}^{(t)}}{\tilde{\tau}_i^{(t)}} \right) \theta_i^{(t)} - \left(1 + \frac{\delta}{\tilde{\tau}_i^{(t)}} \right) e^{\theta_i^{(t)}} \right\},$$

where the posterior mean at stage t from (2.4) is:

$$E(\lambda_i^{(t)} \mid y_i^{(t)}) = \tilde{\mu}_{0i}^{(t)} + \frac{\tilde{\tau}_i^{(t)}}{\delta + \tilde{\tau}_i^{(t)}} \left(y_i^{(t)} - \tilde{\mu}_{0i}^{(t)} \right).$$

Hence, for the dynamic Poisson regression model $log(\lambda_i^{(t)}) = \theta_i^{(t)} = x_i^{(t)'} \beta^{(t)}$, we can implement the Monte Carlo algorithm to generate samples from the posterior distribution of $\beta^{(t)}$.

Hurricane activity in Atlantic ocean

Suppose $y^{(t)}$ denotes number of storms that reached hurricane intensity on a particular year t in Atlantic ocean. *Observed part* at stage t: Suppose $y^{(t)} \sim Poisson(\lambda^{(t)})$. The

posterior distribution of λ at stage $(t-1)$ is $Gamma(a^{(t-1)}, b^{(t-1)})$. Now the *system equation* of the dynamic Poisson model from (2.2) is:

$$e^{\theta^{(t)}} = \rho_t e^{\theta^{(t-1)}} \Rightarrow \quad \lambda^{(t)} = \rho_t \lambda^{(t-1)}, \quad \text{where } \rho_t \text{ is known constant.}$$

Here choose $\rho_t = 1$, *i.e.*, we assume the rate of occurrence of hurricane at year t is same as that of previous year. Hence, $E(\lambda^{(t)} \mid y^{(t-1)}) = \rho_t \frac{a^{(t-1)}}{b^{(t-1)}} = \tilde{\mu}_0^{(t)}$. In order to determine the value of power parameter δ, we again use the Cauchy type function $\delta = \frac{1}{2} - \frac{1}{\pi} \arctan(log(\mid t_i - t_j \mid))$, π is universal constant 3.14159. Then the posterior distribution at stage t is

$$\pi(\lambda^{(t)} \mid y^{(t)}) \propto exp\Big\{ (y^{(t)} + \delta a^{(t-1)}) e^{\lambda^{(t)}} - (1 + \delta b^{(t-1)}) \lambda^{(t-1)}) \Big\}$$
$$\propto exp\Big\{ a^{(t)} exp\{\lambda^{(t)}\} - b^{(t)} \lambda^{(t)} \Big\},$$

where $a^{(t)} = y^{(t)} + \delta a^{(t-1)}$ and $b^{(t)} = 1 + \delta b^{(t-1)}$ and posterior mean at stage t from (2.4) is:

$$E(\lambda^{(t)} \mid y^{(t)}) = \tilde{\mu}_0^{(t)} + \frac{\tilde{\tau}^{(t)}}{\delta + \tilde{\tau}^{(t)}} \Big(y^{(t)} - \tilde{\mu}_0^{(t)} \Big).$$

We consider the satellite era of 1980 to 2007 for this analysis. Red line plot in figure 5 and in figure 6 presents the observed total number of hurricanes for a particular year. Blue line plot represents corresponding predictive and posterior mean and dashed line represents the corresponding 95% predictive and credible interval. Figure 5 presents the predictive and posterior inference for total number of storms that reaches the hurricane intensity in the Atlantic ocean; when we fixed the power parameter as $\delta = 1$. It is clear that if we don't allow δ to be flexible; regular filter technique (when $\delta = 1$) will take care of trend part and ignore the variability of the data. However, in figure 6, when we allow δ to be flexible and PF kicks in - it is clear that most of the time observed number of hurricanes falls well within the 95% predictive interval. This indicates that power filter method can take care of the high variability in the data well.

References

[1] Chen, M.H., Ibrahim, J.G., and Shao, Q. M. (2000) "Power prior distributions for generalized linear models", *J. Statist. Planning. Inf.*, **84**, 121–137.

[2] Diaconis, P. and Ylvisaker. D. (1979) "Conjugate priors for exponential families" *Ann. Statist.*, **17**, 269–281.

[3] Felson DT., (2004) "An update on the pathogenesis and epidemiology of osteoarthritis", *Radiologic Clinics of North America* **42(1)**, 1-9.

[4] Grunwald, G.K., Hamaza, K., and Hyndman, R. J. (1997) "Some properties and generalizations of non-negative time series models" *J. R. Statist. Soc.* B, **97**, 615 – 626.

[5] Grunwald, G.K., Guttorp, P. and Rafftery, A.E, (1993) "Prediction rules for exponential family state space models" *J. R. Statist. Soc.* B, **55**, 937–943.

[6] Harvey, A. C. and Fernandes, C. (1989) "Time series models for count or qualitative observations" *J.Bus. Econ. Statist.*, **7** ,407–422.

[7] Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, Moskowitz RW, Schnitzer TJ. "Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology" *Arthritis. Rheum.*; **38(11)**, 1541-1546.

[8] Ibrahim, J.G., Chen, M.H., and Sinha, D. (2003) "On optimality properties of the power prior" *J. Am. Statist. Assoc.*, **98**, 204–213.

[9] Meinhold, R. J. and Singpurwalla, N. D. (1983) "Understanding the Kalman Filter" *The Am. Statist.*, **37**, 123–127.

[10] Moxley, G., Han, J., Stern, A.G. and Riley, B.P., (2007) "Potential influence of IL1B hapalotype and IL1A-IL1B-IL-1RN extended hapalotype on hand osteoarthritis risk" *Osteoarthritis and Cartilage*, **15**, 1106–1112.

[11] Pendleton A, Ardon N, Dougados M, Doherty M, Bannwarth B, Bijlsma JW *et al.* (2000) "EULAR recommendations for the management of osteoarthritis: report of task force standing committee for International Clinical Studies including Therapeutic Trials (ESCISIT)" *Annals of Rheumatic Diseases*; **59** 936-944.

[12] Roy, S., Khanna, S., Shah, H., Rink, C., Phillips, C., Pressus, H., Subbaraju, G. V., Trimuttulu, G., Krishnaraju, A. V., Bagchi, M., Bagchi, D., and Sen, C.K. (2005) "Human Genome Screen to Identify the Genetic Basis of the Anti-Inflammatory effects of *Boswellia* in microvascular endothelial cells" *DNA Cell Biol*, 24, 244–255.

[13] Sengupta, K., Alluri, K. V., Satish, A.R., Mishra, S., Golakoti, T., Sarma, K.V., Dey, D., and Raychaudhuri, S. B., (2008) "A double blind, randomized, placebo controlled study of the efficacy and safety of 5 Loxin$^®$ for treatment of osteoarthritis of knee" *Arthritis Research and Therapy*, Available online: `http://arthritis-research.com/content/10/4/R85`

[14] Smith, J. Q. (1979) "A generalization of the Bayesian steady forecasting model" *J.R. Statist. Soc. B*, **41**, 375–387.

[15] Smith, J. Q. (1981) "The multiparameter steady model" *J.R. Statist. Soc.* B, **43**,256–260.

[16] West, M. (1986) "Bayesian Model Monitoring". *J.R. Statist. Soc.* B, **48**, 70–78.

[17] West, M., Harrison, P. J. and Migon, H. S. (1985) "Dynamic generalized linear models and Bayesian forecasting" (with discussion). *J. Am. Statist. Assoc.*,**80**, 73–96.

[18] West, M., Harrison, P. J. (1997) "Bayesian Forecasting and Dynamic Models", Second Ed, Springer.

[19] Wright, J.M. (2002) "Double-edged sword of COX-2 selective NSAIDs" *Canadian Medical Association Journal* **167**, 1131-1137.

[20] Zellner, A. (1988) "Optimal information processing and Bayes's theorem" *The Am. Statist.*,**42**, 278–280.

[21] Zellner, A. (2002) "Information processing and Bayesian analysis" *J. Econometrics*, **107**, 41–50.

Table 1: Estimates and 95% credible interval of the regression parameters from the dynamic logistic regression model for pain in 5 Loxin® study

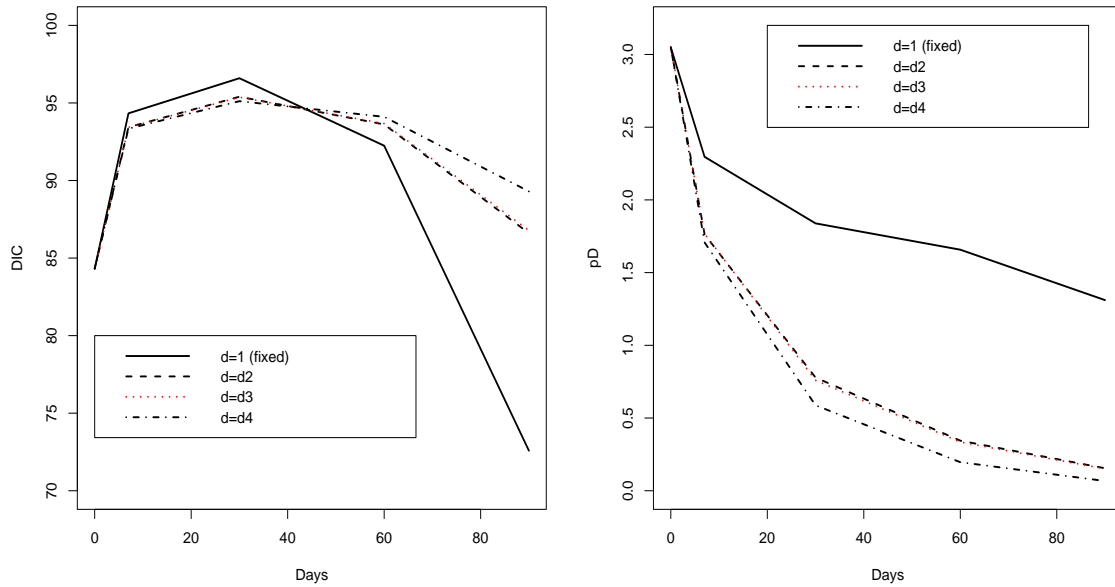|	Baseline	Day 7	Day 30	Day 60	Day 90
Intercept	-2.5399	-26.6535	1.5186	-1.0358	-8.3357
	(-41.821,36.549)	(-67.161,11.232)	(-8.963,11.757)	(-6.166, 4.231)	(-12.076,-4.637)
Low dose	-0.03318	-0.4898	-0.5484	-0.4941	-0.6018
	(-0.906,0.875)	(-1.131, 0.152)	(-0.917,-0.149)	(-0.707, -0.254)	(-0.765,-0.439)
High dose	-0.1648	-0.9294	-0.980	-0.8307	-0.8434
	(-1.047,0.713)	(-1.592,-0.276)	(-1.367,-0.593)	(-1.068,-0.587)	(-1.057,-0.636)
TNFα	0.5049	5.5150	-0.2066	0.2022	1.8864
	(-7.1824, 8.2667)	(-2.006, 13.517)	(-2.449, 2.049)	(-0.932, 1.332)	(1.085, 2.707)
IL-1β	1.4983	8.9364	0.0594	0.5656	3.7256
	(-11.383,14.349)	(-3.731, 22.496)	(-3.556, 3.699)	(-1.423, 2.544)	(2.320, 5.173)
TNF$\alpha\times$	-0.2576	-1.7931	-0.0077	-0.0769	-0.7953
IL-1β	(-2.806, 2.315)	(-4.463, 0.716)	(-0.793, 0.782)	(-0.510, 0.354)	(-1.110, -0.490)

Figure 1: **Left window**: DIC over time for models with different choices of δ. $d_1 = 1$, $d_2 = 0.5 - (\arctan\{log(\mid t_i - t_j \mid)\}/\pi)$, $d_3 = 1/(1 + (\mid t_i - t_j \mid))$, $d_4 = \exp\{- \mid t_i - t_j \mid\}$. **Right widow**: $p_{D^{(t)}}$: Measure of Bayesian complexity over time for the same four different models.
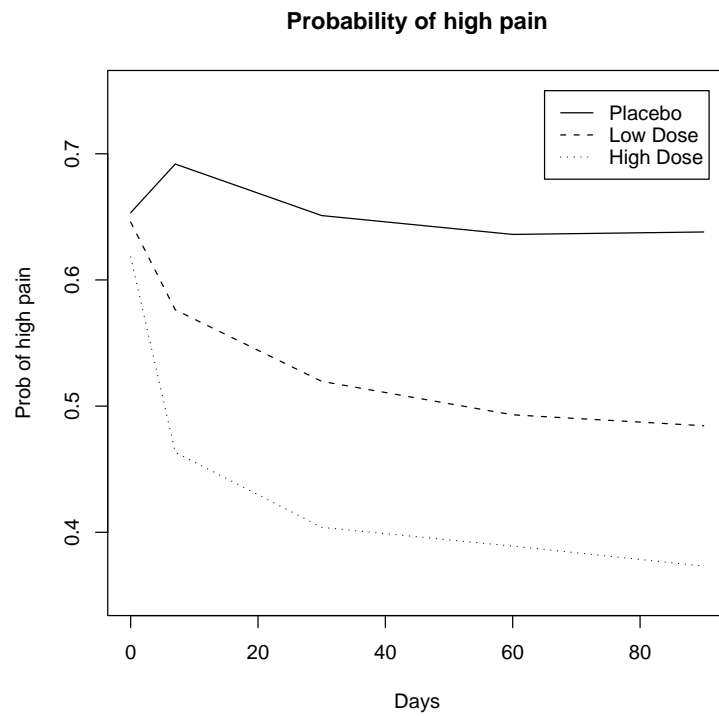
Figure 2: Comparison of the effect of treatment (5 Loxin® with low or high dose) and control (placebo) over the probability of high pain during the study period.
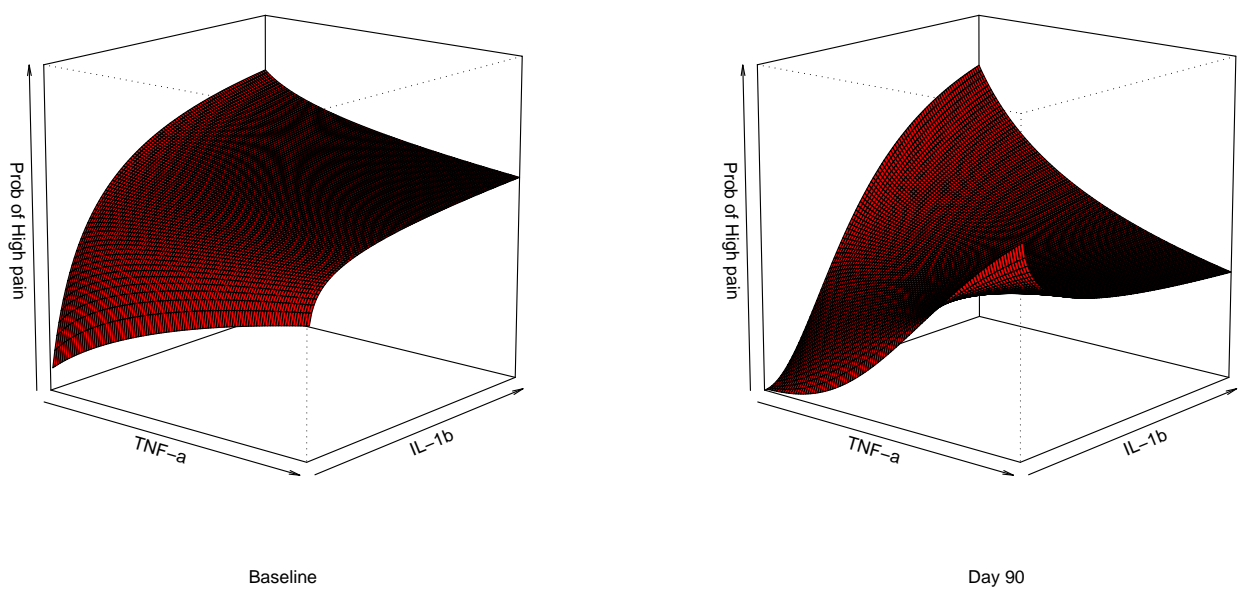

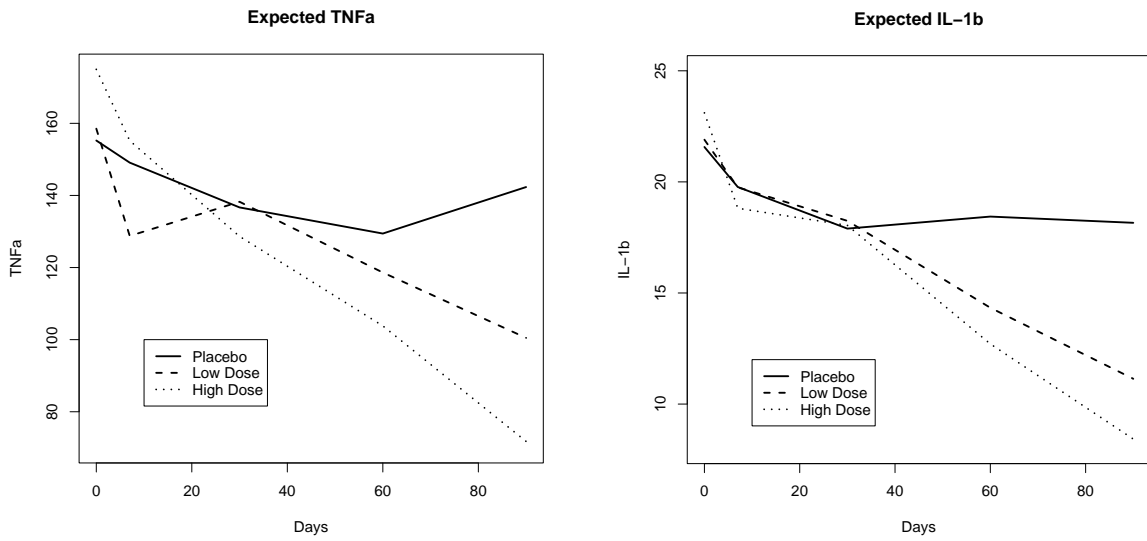Figure 3: Interaction response surface at baseline and end of the study period.

Figure 4: Expected levels of TNFα and IL-1β under placebo and treatment(= low/high dose of 5 Loxin®) group during the study period.
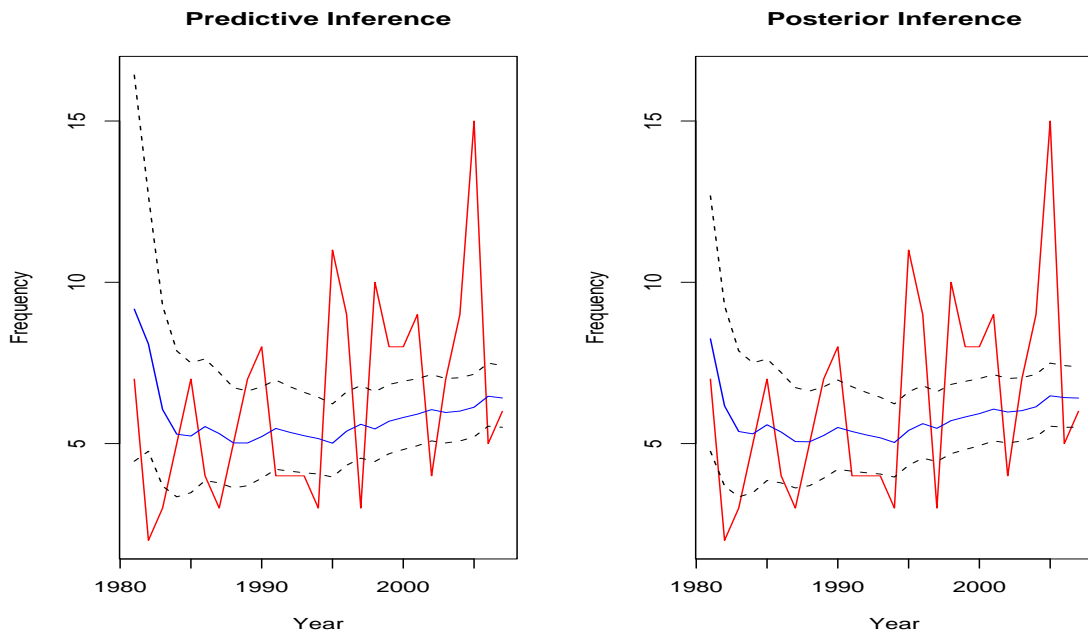

Figure 5: Predictive and Posterior Inference of the Atlantic Ocean when we force $\delta = 1$.
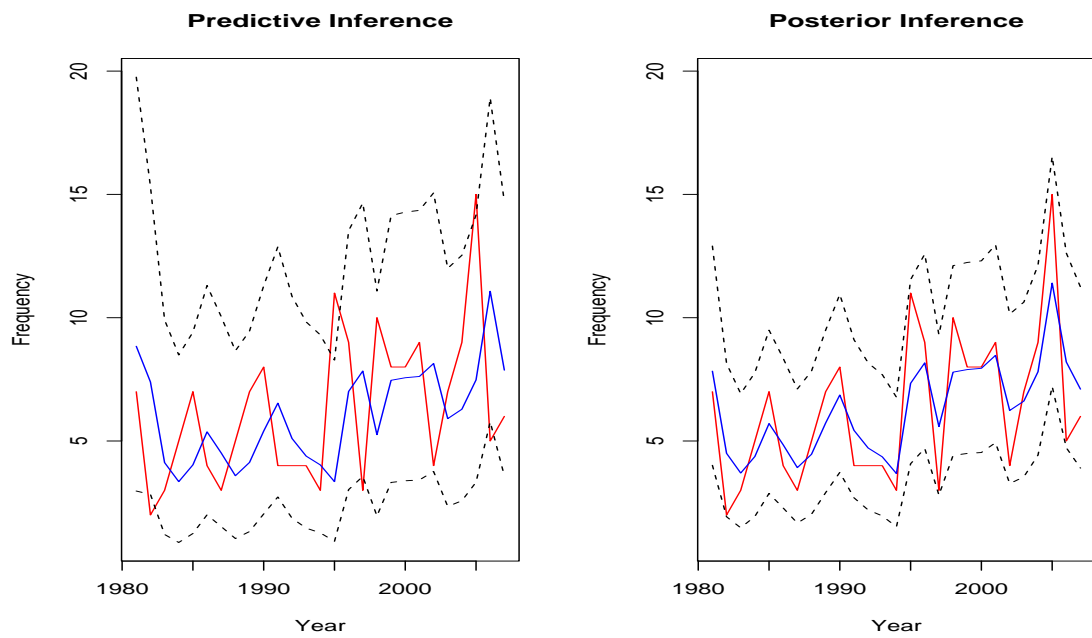
Figure 6: Predictive and Posterior Inference of the Atlantic Ocean.