International Journal of Business and Development Studies **Vol. 12, No. 1,** (2020) pp 21-38

Comparison of Kullback-Leibler, Hellinger and LINEX with Quadratic Loss Function in Bayesian Dynamic Linear Models: Forecasting of Real Price of Oil

Esmaiel Abounoori[1] Mohsen Ali Heydari[2]

ARTICLE INFO

Article history
Received: 22/10/2019
Accepted: 10/04/2020

Keywords: Dynamic Linear Model, LINEX, Kullback-Leibler, Helliger, Quadratic Loss Function, Oil Price Forecasting

JEL classification: C11, C22, C53, C61, Q47

Abstract:
In this paper we intend to examine the application of Kullback-Leibler, Hellinger and LINEX loss function in Dynamic Linear Model using the real price of oil for 106 years of data from 1913 to 2018 concerning the asymmetric problem in filtering and forecasting. We use DLM form of the basic Hoteling Model under Quadratic loss function, Kullback-Leibler, Hellinger and LINEX trying to address the results if we treat the 'over-estimation' and 'under-estimation' differently. So, we drive one-step-ahead forecast for Dynamic Linear Model under quadratic, LINEX and Kullback-Leibler losses in Bayesian context. With Normal posterior distribution, our results suggest that, the LINEX loss function may provide better forecasts than conventional Quadratic loss function, Hellinger and Kullback-Leibler loss function, especially in case of having volatility and time-varying parameters.

1. Introduction

There are situation in which risk of "over-estimation" is much higher than the risk of "under-estimation". Therefore, the main aim in this paper is choosing the right loss function regarding the problem of asymmetry.

Bayesian point estimation of the normal distribution parameters is one of the non-trivial problems in mathematical statistics and yet, there is certainly no consensus about the most appropriate solution. Formally, point estimation may be seen as a decision problem where the action space is the set of possible values of the quantity on interest; foundations then dictate that the solution must depend on both the utility function and the prior distribution (Bernardo, 2007).

Due to the importance of forecasting oil price, many different econometric models of oil price forecasting have been used by researchers. In an overall view,

---

[1] Professor of Econometrics & Social Statistics, Department of Economics, Semnan University, Semnan, Iran, Corresponding author (Email: Esmaiel.abounoori@semnan.ac.ir).
[2] Ph. D Student of Econometrics, Semnan University, Semnan-Iran Aliheydari@Semnan.ac.ir.

despite the different types of model applied for oil price forecasting, applied models ignored the asymmetry problem and so overlooked the difference of "under-estimation" and "over-estimation" by using Quadratic Loss Function (QLS). So, we use Kullback-Leibler (KL), Hellinger (H) and LINEX loss function in Dynamic Linear Model (DLM) which is unprecedentedly new way to consider "asymmetry" issue in forecasting price trajectories.

We fit the basic Hoteling Model of a depletable resource production with DLM under QLS, LINEX, Hellinger and KL loss function (Zellner, 1986; Bernardo & Juarez, 2003; Robert,1996) trying to address the results concerning "over-estimation" and "under-estimation" differently.

2. Related literature review

2.1. Oil price forecasting

Due to importance of accurate prediction of oil price for producers and consumers, there are plenty of academic papers and related articles looking for the better fitness and forecasting models. Phillips and Loretan (1991) proposed a nonlinear model of long-run price of oil by estimating the lagged real prices. Green and Mark (1991) have used GMM for finding the related variable of price formation. Samii (1992) with linear approaches, Moosa and Al-Loghani(1994) with GARCH model for short-time and Schwarts (2003) with DLM for long-time investigate the real spot price fluctuation in time by taking future prices and long run changes through the time. Bahmani-Oskoee and Brown (2004) also use a Kalman Filter model to pursue time-dependency of parameters of demand for international oil demand.

Pindyck (1999) analyzes the stochastic dynamics of crude oil, coal and natural gas prices using a large data set covering 127 years, and tries to assess whether using Kalman filter in time series models are helpful in forecasting long horizons evolution. Morana (2001) proposed a methodology of semiparametric forecasting based on the bootstrap approach to short-term oil price. Ye et al. (2005) applied a dynamic autoregressive model of seasonal changes. Sadorsky (2006) uses several different univariate and multivariate statistical models (GARCH, TGARCH, VAR, bivariate GARCH) to estimate forecasts of daily volatility in petroleum futures price returns. Alquist and Kilian (2010) applied a structural model to predict spot price of oil. They suggest "oil futures prices tend to be less accurate in the mean-squared prediction error sense than no-change forecasts". And some of the most cited papers in oil price forecasting are Jammazi and aloui (2012), Arouri et al. (2012), Nordhaus(1987), Kilian and Hicks (2013), Dvir and Rogoff(2014), Jurado et al. (2015), Gao and Li(2017), Chen et al. (2018) Bakas and Triantafyllo (2019), Zhang and Wang(2019).

2.2. LINEX Loss Function Literature Review

LINEX loss function has been proposed by Varian (1975) and developed by Zellner(1986). In applied economics literature Cain and Janseen (1994) use

LINEX in real estate price prediction, Christoffersen & Diebold (1997) study the optimal prediction under asymmetric loss and indicate that GARCH are relevant for optimal point prediction under LINEX loss function, Batchelor & Peel (1998) applied an ARCH & ARCH-M test for Rational Expectation of agents under LINEX, Paton and Timmerman (2007) evaluate forecast optimality in economics and finance under asymmetric loss, Döpke et al. (2010) forecast German business cycle under an asymmetric loss function and Hyun-Jae, R. (2012) test the possibility of a monetary union in the ASEAN+3 countries: rationality and asymmetric loss functions.

There are, also, other articles about LINEX loss function concerning the fundamental statistics studies. Calabria & Pulicini (1996) test the performance of Bayes estimation of exponential distribution under LINEX compared to MLE, McCullough(2000) proposes a bootstrap method for prediction under asymmetric and nonlinear loss function, Hwang et al. (2001) derived optimal Forecasting Nonlinear Functions of returns, Li et al. (2007) investigate the accuracy of Bayes estimator using LINEX loss under progressively Type-II censored samples, Anatolyev (2009) propose an optimal predictor of the model of dynamic conditional expectation under LINEX, Franses et al.(2011) applied a Lin-Lin Model of experts Estimation of sales forecasts, Pandy et al. (2011) applied a statistical comparison between Bayesian and Maximum Likelihood Estimation of Scale Parameter in Weibull Distribution under LINEX, Jafari-Jozani et al.(2012) establish explicit connections between optimal actions derived under balanced and unbalanced losses under various robust Bayesian analysis criteria including posterior regret gamma-minimaxity, conditional gamma-minimaxed.

Intrinsic (KL and H) Loss Function

Bernardo and Juarez (2003) in Bayesian viewpoint introduces a reference posterior and defines the intrinsic estimator which minimizes the expected loss with respect to that reference posterior distribution. The resulting estimators are shown have attractive invariance properties. Bernardo (2007) propose an Intrinsic Point Estimator for the Normal Variance. Robert (1996) introduce the Bayes estimators related to entropy Hellinger losses for Normal, Gamma, Binominal and poison distribution, Jafari Jozani and Tabrizi (2013) estimate the Intrinsic posterior regret gamma-minimax for the exponential family of distributions, Hershey and Olsen(2007) by introducing two new methods, the variational approximation and the variation upper bound approximate the Kullback Leibler Divergence Between Gaussian Mixture Models (GMM), Erven and Harremos (2014) compare the Rényi entropy with KL, Smith et al. (2006) illustrate Markov-switching model selection using Kullback–Leibler divergence results and introduce a new information criterion based on Markov Switching. There are some other relevance papers on Hellinger loss function which are mentioned as "intrinsic" loss like Maasoumi, (1993), Afgani et al. (2008), Ryu(1993), Seghouane and Amari(2007 & 2012), Georgiou and A. Lindquist(2003).

3. Data and empirical methodology

3.1. Data

We examine the real price of oil over 106 year for period 1913-2018 which obtained from U.S. Crude Oil First Purchase Price (Dollars per Barrel) available in U.S. Energy Information Administration Historical Data. We then deflated this nominal series to 1982 dollars using the Producer Price Index for All Commodities obtained from Economic Research Division of U.S. Federal Reserve Bank.

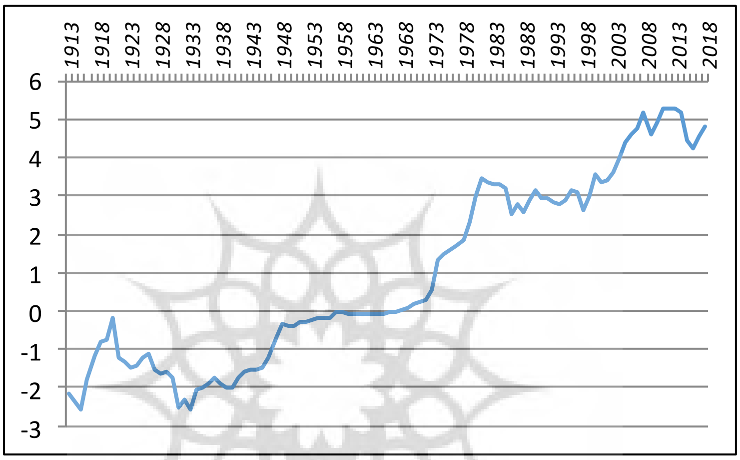


Figure 1: logarithm of Real Oil Price (WTI $Per Barrel)
Source: Findings of this article

Table 1: Descriptive statistic of log real oil price (deflated by 1982)

Mean	Median	Maximum	Minimum	Standard Deviation	Skewness	Kurtosis
0.91102	-0.0394	5.2742	-2.5742	2.39938	0.32461	1.726201

Source: Findings of this article

As shown in figure 1, the major changes (or structural breaks) can be distinguished in 1933, 1973 and 2002 which is important in choosing a proper Unit Root test. On the other hand we can observe a sharp and multiple slop changes of the curve from figure1 that is accompanied by high variance (SD) in data description table.

Table 2: Phillips-Perron Unit-Root Test Results

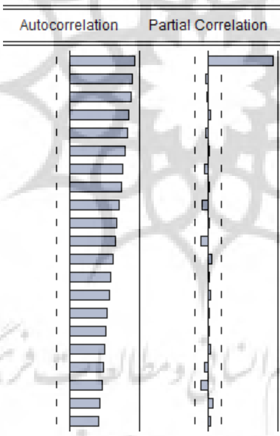
Phillips-Perron test statistic		NA	0.5000
Test critical values:	1% level	-4.047795	
	5% level	-3.453179	
	10% level	-3.152153	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.083818
HAC corrected variance (Bartlett kernel)	0.000000

Source: Eviews' output for Phillips- Perron Test of the data

In the next step, we run the Unit Root test for the logarithm of the real oil price. Since the results of the Dickey-Fuller test are feebly sensitive to structural breaks (Perron(1989), and undoubtedly, oil price data have several structural breaks (Sadorovsky, 2006; Pindyck, 1999), we used the Phillips-Perron test. The result, which has been shown in table2, indicates that the Unit Root of level cannot be rejected. Also, figure 2 which is the results of the correlogram of level for logarithm of real oil price, with 36 lags, confirms this. As we can see in the figure 2, due to the multiple shocks, the autocorrelation will slowly revert to the mean and the shocks' effects last for long time.



	AC	PAC	Q-Stat	Prob
1	0.971	0.971	102.90	0.000
2	0.941	-0.045	200.43	0.000
3	0.911	-0.013	292.70	0.000
4	0.884	0.031	380.34	0.000
5	0.855	-0.044	463.14	0.000
6	0.828	0.021	541.56	0.000
7	0.799	-0.051	615.30	0.000
8	0.771	0.013	684.77	0.000
9	0.740	-0.086	749.33	0.000
10	0.710	0.014	809.44	0.000
11	0.675	-0.112	864.32	0.000
12	0.644	0.052	914.83	0.000
13	0.615	0.017	961.37	0.000
14	0.589	0.030	1004.6	0.000
15	0.565	0.023	1044.8	0.000
16	0.544	0.020	1082.5	0.000
17	0.524	0.032	1117.8	0.000
18	0.503	-0.050	1150.8	0.000
19	0.476	-0.107	1180.6	0.000
20	0.454	0.073	1208.0	0.000
21	0.434	0.025	1233.5	0.000

Figure 2: Correlogram for real oil price deflated to 1982
Source: Eviews result for Unit Root Test of the data

So, the results indicate that the trajectory of real oil price, fluctuate continuously over time. So, in the next section, we use the theory of depletable resource production, completely capable of explaining the continuous and unpredictable fluctuation in level and slope of oil price trajectory through time.

3.2. Empirical methodology

3.2.1. Oil price model

We use the basic Hotelling model of the depletable resource production in competitive price market with constant marginal cost, $\boldsymbol{c}$.

$$\frac{d\,log(p_t)}{dt} = \frac{rc}{(e^{(\frac{rcR_0}{A})}-1)ce^{-rt}+c} \qquad \text{Eq .(1)}$$

Where $\frac{dp}{dt} = \boldsymbol{r}(\boldsymbol{p}-\boldsymbol{c})$, and $\boldsymbol{r}$ is constant, $\boldsymbol{c}$ is extraction cost and $\boldsymbol{p}$ is price level, the demand function designated by $\boldsymbol{Q}$ which is $\boldsymbol{Q} = \boldsymbol{A}(\boldsymbol{p}_0' \boldsymbol{e}^{rt} + \boldsymbol{c})^{-\eta}$, so, $\boldsymbol{\eta}$ is elasticity of demand, $\boldsymbol{A}$ is shift parameter of demand and $\boldsymbol{R_0}$ is initial level of reserve in the market.[1]

As it shown in eq.1, the increase of $\boldsymbol{A}$ (i.e. the transfer of the demand curve upwards) leads to an increase in the level and slop of the logarithm of the price. An increase in the cost of extraction, $\boldsymbol{c}$, shall increase the oil price, but the slope of the price trajectory, and accordingly, the logarithm of the price decreases. The sudden rise in $\boldsymbol{R_0}$, means discovery of new resources, will also lower the price.

To track the evolution of parameters, we use the process of the Ornstein-Uhlenbeck (OU) transformations (Pindyck, 1999, Radchecko, 2005).

$$\begin{aligned} p_t &= \rho p_{t-1} + b_1 + b_2 t + \varphi_{1,t} + \varphi_{2,t} t + \nu_t \\ \varphi_{1,t} &= c_1 \varphi_{1,t-1} + \omega_{1,t} \\ \varphi_{2,t} &= c_2 \varphi_{2,t-1} + \omega_{2,t} \end{aligned} \qquad \text{Eq .(2)}$$

Where $\varphi_{1,t}$ and $\varphi_{2,t}$ are unobservable variables.[2]

For dynamic linear model estimation, we eliminate the trend by replacing the $b_2 t$ and $\varphi_{2,t} t$ with a Local Level and Linear Growth equations accordingly (Petris et al., 2010).

$$\begin{aligned} p_t &= \rho p_{t-1} + b_1 + \varphi_{1,t} + \varphi_{2,t} + \nu_t \\ \varphi_{1,t} &= c_1 \varphi_{1,t-1} + \omega_{1,t} \\ \varphi_{2,t} &= c_2 \varphi_{2,t-1} + \beta_t + \omega_{2,t} \\ \beta_t &= \beta_{t-1} + \omega_{3,t} \end{aligned} \qquad \text{Eq . (3)}$$

3.2.2. LINEX Loss Function

There are several loss functions in statistics. For instance, some of the most popular loss functions are: the Absolute, All-or-Nothing, Quadratic Loss, LINEX and Intrinsic Loss Function[3].

[1] For a further proof and discussion, see Pindyck (1999).
[2] For a detailed proof and discussion, see Lo and Wang (1995).
[3] For further reading, see Koop et al. (2007).

In this paper, due to implying the different weight on “under-estimation” and “over-estimation” for real oil price estimation, we first use the Linear Exponential loss function (LINEX) which have been proposed by Varian (1975) and developed by Zellner (1986).

If let the $\Delta = \widehat{\theta} - \theta$ which denote the scalar estimation error in using $\widehat{\theta}$ to estimate θ, the $L(\Delta) = be^{a-\Delta} - c\Delta - b$ *where* $a, c \neq 0$, $b > 0$, is the Varian (1975) convex loss function, so, if let the $ab = c$, we shall have:

$$L(\Delta) = b[e^{a-\Delta} - a\Delta - 1] \qquad a \neq 0 \ , b > 0 \qquad \text{Eq .(4)}$$

For the posterior pdf, $p\left(\theta \middle| D\right)$ is $N(\bar{y}, \sigma^2)$, the value of $\widehat{\theta}$ that minimize Eq.(4) is $\widehat{\theta}_B = \bar{y} - \frac{a\sigma^2}{2n}$ (Zellner, 1986).

Table 3: Admisible Risk and Point Estimation under QLS and LINEX

Loss Function	Square Error	LINEX
Admissible point estimation of y	$\bar{y}$	$\hat{\theta}_B = \bar{y} - \frac{a\sigma^2}{2n}$
Admissible risk of estimation	$R_s(\bar{y}) = \frac{\sigma^2}{n}$	$R_L(\hat{\theta}_B) = b\frac{a^2\sigma^2}{2n}$

Note: these risk functions, denoted by **R**, where the subscript **L** denotes risk relative to the LINEX loss function and **S** denoting square error loss function.

3.2.3. Intrinsic Loss Function

Kullback-Leibler: *If* $p(x \mid \theta_1)$ *and* $p(x \mid \theta_2)$ *are probability densities with the same support X, the directed logarithmic of* $p(x \mid \theta_2)$ *from* $p(x \mid \theta_1)$ *is defined as*

$$k_X(\theta_1 \mid \theta_2) = \int_X p(x \mid \theta_1) \, log \frac{p(x \mid \theta_1)}{p(x \mid \theta_2)} dx \qquad \text{Eq .(5)}$$

The directed logarithmic divergence (often referred to as Kullback-Leibler information) is non-negative, and it is invariant under bijections of both x and θ. It is additive in the sense that, if $x \in X$ and $y \in Y$ are conditionally independent given θ_1, then the divergence $k_{X,Y}(\theta_1, \theta_2)$ of $p(x, y \mid \theta_2)$ from $p(x, y \mid \theta_1)$ is simply $k_X(\theta_1 \mid \theta_2) + k_Y(\theta_1 \mid \theta_2)$; if data x are assumed to be a random sample $x = \{x_1, \ldots . n\}$ from $p(x \mid \theta)$, *then the divergence of* $p(x \mid \theta_2)$ from $p(x \mid \theta_1)$ is simply n times the divergence of $p(x \mid \theta_2)$ from $p(x \mid \theta_1)$. Under appropriate regularity conditions, there are many connections between the logarithmic divergence and Fisher’s information (see e.g. Stine, 1959; Bernardo and Juarez, 2003). Furthermore, $k_X(\theta_1, \theta_2)$ has an attractive interpretation in information-theoretical terms: it is expected amount of information (in natural units, *nits*) necessary to recover $p(x \mid \theta_1)$ from $p(x \mid \theta_2)$.

However, the directed logarithmic divergence is not symmetric and diverges if the support of $p(x \mid \theta_2)$ is an strict subset of the support of $p(x \mid \theta_1)$. To simultaneously address those two unwelcome feature we propose to use the symmetric intrinsic discrepancy $\delta_X(\theta_1 \mid \theta_2) = \min\{k_X(\theta_1 \mid \theta_2), k_X(\theta_2 \mid \theta_1)\}$. To simply the notation, the subindex X we will dropped from both $\delta_X(\theta_2 \mid \theta_1)$ and $k_X(\theta_2 \mid \theta_1)$ whenever there is no danger of confusion. (Robert, 1996).

Let x = $\{x_1, \dots, x_n\}$ be a random sample from a Normal $N(x \mid \mu, \sigma^2)$ distribution, and let $\bar{x}$ and s^2 respectively be the corresponding sample mean and variance, with $n\bar{x} = \sum_j x_j$, and $ns^2 = \sum_j (x_j - \bar{x})^2$. In terms of precisions, $\lambda_j = \sigma^{-2}$, the directed logarithmic divergence $k\{\mu_2, \lambda_2 \mid \mu_1, \lambda_1\}$ of $N(x \mid \mu_2, \lambda_2)$ from $N(x \mid \mu_1, \lambda_1)$ is

$$\int_{-\infty}^{+\infty} N(x \mid \mu_1, \lambda_1^{-1}) log \frac{N(x \mid \mu_1, \lambda_1^{-1})}{N(x \mid \mu_2, \lambda_2^{-1})} dx \quad \text{Eq .(6)}$$

$$= \frac{1}{2}[log \frac{\lambda_1}{\lambda_2} - 1 + \frac{\lambda_1}{\lambda_2} + \lambda_2(\mu_1 - \mu_2)^2]$$

And the intrinsic discrepancy between the estimated model $N(x \mid \mu^e, \lambda^e)$and the assumed model $N(x \mid \mu, \lambda)$ is

$$\delta\{\lambda^e, \mu^e, \lambda, \mu\} = \min[k\{\lambda^e, \mu^e \mid \lambda, \mu\}, k\{\lambda, \mu \mid \lambda^e, \mu^e\}] \quad \text{Eq .(7)}$$

The reference prior when both μ and λ are of interest is $\pi(\lambda, \mu) = \lambda^{-1}$, and the corresponding (joint) reference posterior is Normal-Gamma $\pi(\lambda, \mu \mid x) = N(\mu \mid \bar{x}, (n\lambda)^{-1}) Ga(\lambda \mid \frac{n-1}{2}, \frac{ns^2}{2})$

Thus, the reference posterior expected intrinsic loss may then be computed as

$$d(\lambda_e, \mu_e \mid x) = \int_0^{\infty} \int_{-\infty}^{+\infty} \delta\{\mu_e, \lambda_e, \mu, \lambda\} \pi(\mu, \lambda \mid x) d\lambda d\mu \quad \text{Eq .(8)}$$

a concave function. The intrinsic estimator $\{\mu^*, \lambda^*\}$ is its unique minimum $\{\bar{x}, \lambda^*(x)\}$, where the exact value of $\lambda^*(x)$ requires one-dimensional numeric integration, but which is very well approximated by

$$\lambda^*(x) \approx \frac{n-2}{ns^2} = \frac{1}{2}(E[\lambda \mid x] + Mo[\lambda \mid x]) \quad \text{Eq .(9)}$$

Bernardo & Juarez (2003).

Hellinger: Given two densities $p(x \mid \theta_1)$ and $p(x \mid \theta_2)$ with respect to a dominating measure ν, the Hellinger distance $H(p(x \mid \theta_1), p(x \mid \theta_2))$ is defined by

$$H^2(p(x \mid \theta_1), p(x \mid \theta_2)) = \frac{1}{2}\int_x (\sqrt{p(x \mid \theta_1)} - \sqrt{p(x \mid \theta_1)})^2 \, dv(x) \qquad \text{Eq .(10)}$$

If $f(x \mid \theta)$ is the normal $N(\theta, 1)$ distribution, the Hellinger loss is $L_H(\theta, d) = 1 - \exp\{-(\theta - d)^2/8\}$.

Consider an $N(\mu, \tau^2)$ prior on θ. The Bayes estimator of θ is then

$$Arg \;\; max_d \int_{-\infty}^{+\infty} \exp\{-\frac{(\theta - d)^2}{8} - \frac{(\theta - x)^2}{2} - \frac{(\theta - \mu)^2}{2\tau^2}\} d\theta$$

$$= Arg \;\; max_d \int_{-\infty}^{+\infty} \exp\{-\frac{\theta^2}{8}\left(5 + -\frac{4}{\tau^2}\right) + \theta(\frac{d}{4} + x + \frac{\mu}{\tau^2})\} d\theta e^{-d^2/8}$$

$$= Arg \;\; max_d \exp\left\{\left(\frac{d}{8} + \frac{x}{2} + \frac{\mu}{2\tau^2}\right)^2 \frac{8}{5 + 4/\tau^2} - \frac{d^2}{8}\right\}$$

$$= Arg \;\; max_d \left(d - \frac{\tau^2 x + \mu}{\tau^2 + 1}\right)^2 \qquad \text{Eq .(11)}$$

And we recover the usual posterior mean, $\frac{\tau^2 x + \mu}{\tau^2 + 1}$. Robert (1996)

4. Empirical results

We applied the DLM to estimate Equation (3), which is a dynamic linear form of space-state models. The basic model is AR (1) and the values of coefficients and parameters are unknown. We run the maximum likelihood to estimate the unknown parameters and to avoid the “local maximum” we used different initial values, and finally we selected the unknown values of equation (3) which had the highest likelihood values. We, then, implemented the conventional DLM (under least squares loss function) in the State- Space model with **R** programming using the "dlm" package[1]. This was done to pave the way for comparing the results of the DLM under the alternative loss functions, i.e. QLS and LINEX, KL and Hellinger.

Table 4: MLE values of the unknown parameters

σ_{ω_p}	$\sigma_{\omega_{\varphi 1}}$	$\sigma_{\omega_{\varphi 1}}$	σ_{ω_β}	σ_{v}	ρ	ML	Convergence
5.86	8.92	2.45	18.7	0.00013	0.9997	-62.748	0

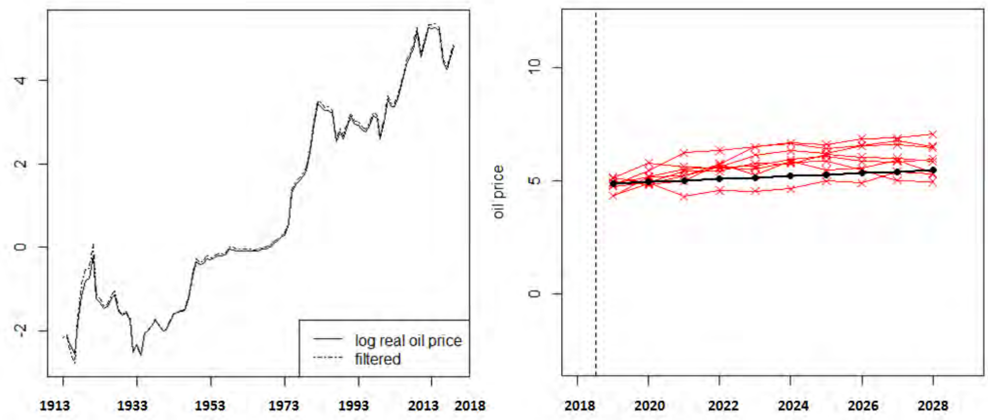
Source: Findings of this article

Now consider a general form of State-Space Model where $\boldsymbol{\theta}$ represent the state variable which is latent, and $\boldsymbol{y}$ denoting the observed variable.

$$\begin{cases} \theta_t = G_k \theta_{t-1} + \omega_t \\ y_k = F_t \theta_t + \nu_t \end{cases} \qquad \text{Eq .(12)}$$

[1] Further details are available in Petris et al. (2010).

DLM Oil price filter (left) and forecast (right) under QLS



Source: Findings of the article based on DLM under QLS

4.1. DLM under LINEX loss

To enter the LINEX loss function into the DLM, assuming the normal conjugate of the posterior distribution, we fulfill the process of filtering of the following calculations by replacing the admissible LINEX point estimation amount into the each following values

I. mean and conditional variance $\pi\ (\theta_t \mid y_{1:t-1})$
II. mean and conditional variance $\pi\ (Y_t \mid y_{1:t-1})$
III. mean and conditional variance $\pi\ (\theta_t \mid y_{1:t})$

For initiation of filtering, and afterward the forecasting, we put the initial values of mean by **(price of 1913, average of first 5 years, 0 ,0)** and the variance by **(2.5, 2.5, 0,0)** into the one-step-ahead filtering 1913- 2013 and afterward forecasting the data of real price of oil for 2014 -2020, *for LINEX parameters*. The results are as follows:

Table 5: One-step-ahead forecast values under QLS and LINEX

YEAR	OBSERVATION	QLS FORECASTING	LINEX FORECASTING A=1,B=1	LINEX FORECASTING A=2,B=2	LINEX FORECASTING A=3,B=3
2014	5.18976417	5.263746	5.263713	5.263214	5.261058
2015	4.43718996	5.258482	5.258449	5.257951	5.255796
2016	4.26239938	5.253224	5.25319	5.252693	5.250541
2017	4.53252171	5.24797	5.247937	5.247441	5.24529
2018	4.82087644	5.242722	5.242689	5.242193	5.240045
2019	-	5.23748	5.237447	5.236951	5.234805
2020	-	5.23224	5.232209	5.231714	5.22957

Source: Findings of the article.

Based on the results of table 5, we calculate the Theil's inequality coefficient (table. 6) to compare the results of two alternatives loss functions in which lower values of Theil indicate the better results of forecasting under LINEX loss function.

Table 6: Theil coefficient for the Loss Function forecasting

THEIL INEQUALITY INDEX	UNDER QLS	UNDER LINEX A=1,B=1	UNDER LINEX A=2,B=2	UNDER LINEX A=3,B=3
Theil	0.069179808	0.069177075	0.069136347	0.068960005

Source: Findings of the article.

4.2. DLM under Hellinger and Kullback-Leibler loss

To capture the dynamic linear model of eq.3 under the Intrinsic losses like Hellinger and Kullback-Leibler, in accordance with eq.9, we applied the Mean for point estimation on KL loss function and simply extract the related amount of the loss-fitted estimation taking the sample size of 1000 and 10,000 and replace the posterior mean of the date which minimized the related loss function into the calculation of three steps of DLM estimation, and follow the procedure to finally find the best point-estimation based on the KL loss function.

To find point estimation of DLM under Hellinger loss, we need to find the appropriate amount of the parameters of prior distribution for eq.11. To do so, we need two assumptions which can help us to determine the parameters of prior distribution. By assuming $-0.9x \leq \frac{\tau^2 x+\mu}{\tau^2+1} \leq +0.9x$ and based on experienced Bayesian method, we assumed $\tau^2 = 6.09$ and $\mu = 5.8$ which will give well-approximation amount of prior for the first observation. To replace the Hellinger-fitted amount into DLM, we use the prior parameters in calculation of $\pi\left(\theta_t \mid y_{1:t-1}\right)$.

Table 7: One-step-ahead forecast values under QLS and LINEX

YEAR	OBSERVATION	QLS FORECASTING	KL FORECASTING N=1000	KL FORECASTING N=10000	HELLINGER
2014	5.18976417	5.263746	5.2638	5.263751	5.271815
2015	4.43718996	5.258482	5.258536	5.258487	5.266543
2016	4.26239938	5.253224	5.253277	5.253229	5.261277
2017	4.53252171	5.24797	5.248024	5.247976	5.256016
2018	4.82087644	5.242722	5.242776	5.242728	5.25076
2019	-	5.23748	5.237533	5.237485	5.245509
2020	-	5.23224	5.232296	5.232248	5.240263

Source: Findings of the article

Based on the results of table 7, we calculate the Theil's inequality coefficient (table. 6) to compare the results of two alternatives loss functions.

Table 8: Theil coefficient for the Loss Function forecasting

THEIL INEQUALITY INDEX	UNDER QLS	UNDER KL N=1000	UNDER KL N=10000	UNDER HELLINGER
Theil	0.069179808	0.069184207	0.069180249	0.06984

Source: Findings of the article

5. Conclusion and Discussion

1. The findings of table 6 and 8, Theil index, indicate that better results of forecasting are captured through LINEX, than QLS or H and KL. As an example, for logarithm of oil price in 2014, the difference between conventional loss function (QLS) and LINEX is 0.027, if we convert the logarithm value into the real value of oil price, the difference will be $1.028, which is a considerable improvement in accuracy of forecasting.

Also, the difference between Theil indexes is not remarkably high, but considers the fact that the estimation takes place under logarithm transformation and it will be remarkably meaningful for oil price forecasting.

2. According to the results the DLM under the conventional Quadratic Loss function and LINEX loss functions, admissible variance of the estimation under LINEX are less than the QLS. Note that the simulation results are susceptible to the variance of the prediction error; therefore, improvement of simulation results can be expected.

3. LINEX loss function for $a = 1$, is quite asymmetric with overestimation being more costly than underestimation (Zellner, 1986). It would be significantly important, especially when one intended to assess the price of oil in a tough situation, like international boycott of an oil exporter country. So, as the dependency of governments' budget to the oil revenues deepens, the over-estimation of oil price gets more sensitive and even more vulnerable. The LINEX loss function allows the researcher to choose between the risks, lower the prediction error variance occasionally, and adjust the precisions beside the preferences of the research.

4. Since the asymmetry is more important in socio-economic studies, choosing the appropriate loss function will also be more significant. For example, when we fit a model on the number of political assassinations or casualties on anti-government clashes over an economic variable such as uncertainty, investment and growth, then researcher can no longer, and should not, be indifferent between a unit of "positive error" and "negative error" in the estimation. Therefore, in the socio-economic studies, choosing the appropriate loss function will be much more significant.

5. Our findings confirm that in the DLM forecasting, the LINEX loss function has more accurate forecasting results than other loss function.

References:

1. Aboonouri, E., Ali-Heydari, M. (under publication). “Kalman Filter under the LINEX Loss Function: Forecasting Real Price of Oil” (in Persian).
2. Alquist, R., Kilian, L. & Vigfusson, L. (2011). “Forecasting the Price of Oil”. *Staff Working Papers* No.11-15, Bank of Canada.
3. Anatolyev, S.(2009). “Dynamic modeling under linear-exponential loss”. *Economic Modeling, Elsevier*, 26(1): 82-89.
4. Bahmani-Oskooee, M. & Brown, F. (2004). “Kalman filter approach to estimate the demand for international reserves”. *Applied Economics*, 36:1655-1668.
5. Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K. (1998). “Don’t Look Back”. *Risk* 11:100–104.
6. Barro, R.J. & Martin, X.S. (1990). “World Real Interest Rate,”. *RCER Working Papers 227*, University of Rochester Center for Economic Research (RCER).
7. Batchelor,R. & Peel,D.A.(1998). “Rationality testing under asymmetric loss”. *Economics Letter.Elsevier*,61:49-54.
8. Bernardo, J.M., Juarez, M. (2003). “Intrinsic Estimation”. *Bayesian Statistics*, 7:465–476.
9. Bernardo,J.M., (2007). “Objective Bayesian point and region estimation in location-scale models”. *SORT*,31(1):3-44.
10. Bernardo, J.M., (2007). Modern Bayesian Inference: Foundations and Objective Methods. In *Philosophy of Statistics*, eds. Bandyopadhyay, P. and Forster, M., Amsterdam: Elsevier.
11. Birg´e, L. (2006). “Model selection via testing: an alternative to (penalized) maximum likelihood estimators”. *Ann. Inst. H. Poincar´e Probab. Statist.*, 42(3):273–325.
12. Cain, M. & Janssen, C. (1995). Ann Inst Stat Math, 47:390- 401. https://doi.org/10.1007/BF00773391
13. Calabria,R. & Pulicini,G. (1996). “Point estimation under asymmetric loss functions for left-truncated exponential samples”, *Communications in Statistics: Theory and Methods*, 25(3): 585-600
14. Chen, Y. & Zhang, C., He, K. & Zheng, A. (2018). “Multi-step-ahead crude oil price forecasting using a hybrid grey wave model”. *Physica A: Statistical Mechanics and its Applications, Elsevier*, 501(C):98 -110.
15. Christoffersen, P.F. & Diebold, F.X. (1997). “Optimal prediction under asymmetric loss”. *Econometric Theory*, 13: 808-817.
16. Dees, S., Karadeloglou, P., Kaufmann, R.K. & Sanchez, M. (2007). “Modeling the World Oil Market: Assessment of a Quarterly Econometric Model”. *Energy Policy*, 35:178–191.

17.Döpke, J., Fritsche, U. & Siliverstovs, B. (2010). “Evaluating German business cycle forecasts under an asymmetric loss function”. *Macroeconomics and Finance Series 200905:1-18,* University of Hamburg, Department of Socioeconomics. https://doi.org/10.1787/jbcma-2010-5kmlj35rx10s.
18.Dvir, E. & Rogoff, K.(2014). “Demand effects and speculation in oil markets: Theory and evidence”. *Journal of International Money and Finance*, 42:113–128.
19.El Hédi Arouri, M., Lahiani, A., Lévy, A. & Khuong-Nguyen, D. (2012). “Forecasting the conditional volatility of oil spot and futures prices with structural breaks and long memory models”. *Energy Economics*, 34:283–293.
20.Fama, E.F. & French, K. (1987). “Commodity Future Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage”. *The Journal of Business,* 60(1):55-73.
21.Fama, E.F. (1965). “The Behaviour of Stock Market Prices”. *Journal of Business*, 38(1):34-105.
*22.*Franses, P.H., Legerstee, R.& Paap, R.(2011). “Estimating Loss Functions of Experts”. *Tinbergen Institute Discussion Paper No. 2011-177/4. Econometric Institute, Erasmus School of Economics, Erasmus University of Rotterdam & Tinbergen Institute.*
23.Gao, S. & Lei Y. (2017). “A new approach for crude oil price prediction based on stream learning”. *Geoscience Frontiers* ,8:183-187.
24.Greenand, S. L. & Mork, K. A. (1991). “Towards Efficiency in the Crude Oil Market”. *Journal of Applied Econometrics*, 6:45-66.
25.Gurvich, E. T. & Prilepskiy, I. V. (2018) . “Analysis of expert and official oil price forecasts”. *voprosy economiki Redaktsiya zhurnala*, 4:83-56.
26.Hamilton, J.D. (2014). “Oil Prices as an Indicator of Global Economic Conditions”. Econ Browser Blog Entry, 14 December.
27.Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge, UK: Cambridge University Press.
28.Hwang, S., Knight, J. & Satchell, S. (2001). “Forecasting nonlinear functions of returns using LINEX loss function,” Annals of Economics and Finance, 2, 187-213.
29.Hyun-Jae, R. (2012). “Testing for the possibility of a monetary union in the ASEAN+3 countries: rationality and asymmetric loss functions”. *Applied Economics Letters, Taylor & Francis Journals,* 19(3):261-268.
30.Jafari-Jozani,M., Marchand, É. & Parsian, A.(2012). “Bayesian and Robust Bayesian analysis under a general class of balanced loss functions”. *Stat Papers, Springer,* 53:51–60.DOI 10.1007/s00362-010-0307-8
31.Jammazi ,R. & Aloui, C. (2012). “Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling”. *Energy Economics*, 34:828–841.

32.Kalman, R. E. (1960). “A new approach to linear Filtering and prediction problems: Transactions of the ASME”. *Journal of Basic Engineering*, 82 (Series D):35-45.

33.Kaufmann, R. K. (1995). “ A Model of the World Oil Market for Project LINK: Integrating Economics, Geology, and Politics”. *Economic Modeling*, 12:165–178.

34.Kaufmann, R. K. (2004). “Does OPEC Matter?”. An Econometric Analysis of Oil Prices. *Energy Journal*, 25:67–91.

35.Kilian, L. & Hicks, B. (2013). “Did Unexpectedly Strong Economic Growth Cause the Oil Price Shock of 2003–2008? ”. *Journal of Forecasting*, 32(5):385-394.

36.Kilian, L. & Thomas K.(2014). “Quantifying the speculative component in the real price of oil: The role of global oil inventories”. *Journal of International Money and Finance*, 42:71–87.

37.Koop, G., Poirier, D. J., & Tobias, J.L. (2007). *Bayesian Econometric Methods*. Cambridge ,UK: Cambridge University Press.

38.Koopman, S. J. , & Durbin j.(1998). “Fast filtering and smoothing for multivariate state space models”. *Journal of time series analysis*, 21:45-89.

39.Lalonde, R., Zhu, Z. & Demers, F. (2003). “Forecasting and Analyzing, World Commodity Prices”.*Working Paper 2003–24*, *Bank of Canada.*

40.Leung, G.C.K. (2010). “China's Oil Use in 1990–2008”. *Energy Policy*, 38:932–944.

41.Li, X., Shang, W. & Wanga, Sh. (2019). “Text-based crude oil price forecasting: A deep learning approach”. *International Journal of Forecasting*, in Press.

42.Li, X., Shi, Y., Wei,J. & Chai, J. (2007). “Empirical Bayes estimators of reliability performances using LINEX loss under progressively Type-II censored samples”. *Mathematics and Computers in Simulation*, 73: 320–326.

43.Lo,A.W., & Wang, J. (1995). “Implementing Option Pricing Models when Asset Returns are Predictable”. *Journal of Finance*, March 50:87-129.

44.Mamipour, S. & Vaezi-Jezeie, F. (2015). “Non-Linear Relationships among the Oil Price, the Gold Price and Stock Market Returns in Iran: a Multivariate Regime-Switching Approach”. *Iranian Journal of Economic Studies*, 4(1):101-128.

45.Mc Cullough(2000). Optimal Prediction with a General loss function, Journal of Combinatorics, Information & System Sciences, 25(1-4):207-221.

46.Miller, J. I., & Ni, Sh., (2011). “Long-Term Oil Price Forecasts: A New Perspective on Oil and the Macro economy”. *Macroeconomic Dynamics*, 15 (Supplement 3):396–415.

47.Moosa, I. A., & Al-Loughani, N. E. (1994). “Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market”. *Energy Economics*, 16:99–105.

48.Morana, C., (2001). “A Semi-parametric Approach to Short-term Oil Price Forecasting”. *Energy Economics*, 23:325–338.
49.Morana,C.,(2014). “Understanding the Decline in the Price of Oil since June: A semiparametric approach to short-term oil price forecasting”. *Energy Economics*, 23(Issue 3):325-338.
*50.*Moshiri, S., & Foroutan, F., (2006). “Forecasting Nonlinear Crude Oil Prices”. *The energy journal*, 27(4):81-95.
51.Narayan, P.K., & Narayan, S. (2007). “Modeling Oil Price Volatility”. *Energy Policy*, 35:6549–6553.
52.Pandey, B.N., Dwivedi, N. & Bandyopadhyay, P. (2011). “Comparison between Bayesian and Maximum Likelihood Estimation of Scale Parameter in Weibull Distribution with known shape under Linex loss function”. *Journal of Scientific Research,* 55: 163-172.
53.Patton, A. & Timmerman, A.(2007). “Properties of optimal forecasts under asymmetric loss and nonlinearity”. *Journal of Econometrics*, 140(2): Pages 884-918.
54.Perron,P.,(1989). “The great crash, the oil price shock and the Unit Root Hypothesis”. *Journal of Political Economy*,57:1361-1401.
55.Petris, G., Petrone, S. & Campagnoli, P., (2010).Dynamic Linear Models with R.. London: Springer Publication.
56.Phillips, P.C.B. & Loretan, M. (1991). “Estimating Long-run Economic Equilibria”. *Review of Economic Studies*, 58: 407–436.
57.Pindyck, R. S. (1999). “The Long-run Evolution of Energy Prices”. *The Energy Journal*, 20:1–27.
58.Radchenko, S. (2005). “ The Long-run Forecasting of Energy Prices: Using the Model of Shifting Trend”. Working Paper No.1240, *University of North Carolina at Charlotte*.
59.Robert C. (1996). “Intrinsic losses”. *Theory and Decision* 40:191–214.
60.Robert C. (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. 2nd ed. New-York: Springer-Verlag, Springer Texts in Statistics.
61.Roio ,J.(1987). “ On the admissibility of c[Xbar] + d with respect to the linex loss function”. *Communications in Statistics: Theory and Methods,* 16(12), 3745-3748, DOI: 10.1080/03610928708829603.
62.Sadorsky P. (2006). “Modeling and forecasting petroleum futures volatility”. *Energy Economics*, 28:467–488.
63.Safari, A. & Davallou, M. (2018). “Oil price forecasting using a hybrid model". *Energy Elsevier*, 148(C):49-58.
64.Samii, M. V. (1992). “Oil Futures and Spot Markets”. *OPEC Review,* No. 4:409–417.
65.Schwartz, E. & Smith, J. E. (2000). “Short-term Variations and Long-term Dynamics in Commodity Prices”. *Management Science,* 46:893–911.

66.Stéphane, L., (2010). "Intrinsic Bayesian inference on a Poisson rate and on the ratio of two Poisson rates". *Journal of Statistical Planning and Inference, Elsevier,* doi:10.1016/j.jspi.2012.02.040. 10.1016/j.jspi.2012.02.040. hal-00505234v3.

67.Wan, A. T. K. & Hiroko, K. (1999). "An iterative feasible minimum mean squared error estimator of the disturbance variance in linear regression under asymmetric loss". *Statistics & Probability Letters, Elsevier*, 45(3):253-259.

68.Yannick Baraud.(2011). "Estimator selection with respect to Hellinger-type risks". *Probability Theory and Related Fields, Springer Verlag, 151* (1-2):353-401.

69.Zellner, A. (1986). "Bayesian Estimation and Prediction Using Asymmetric Loss Functions". *Journal of the American Statistical Association,* 81:446-451.

70.Zamani, M. (2004). An Econometrics Forecasting Model of Short Term Oil Spot Price. *Paper presented at the 6th IAEE European Conference,* Zurich, 2–3 September.

71.Zeng, T. & Swanson, N.R.(1998). "Predictive Evaluation of Econometric Forecasting Models in Commodity Future Markets". *Studies in Nonlinear Dynamics and Econometrics*, 2:159–177.

72.Zhang, Y.J. & Wang, J. (2015). "Exploring the WTI Crude Oil Price Bubble Process Using the Markov Regime Switching Model". *Physica A: Stat. Mech. its Appl*, 421:377–

مقایسه توابع زیان کولبک-لایبلر، هلینگر و لاینکس با تابع زیان حداقل مربعات در مدل خطی پویای بیزین: پیش بینی قیمت حقیقی نفت

چکیده:

ما در این مقاله، قصد داریم تا کاربرد توابع زیان متفاوتی همچون تابع زیان کولبک-لایبلر، هلینگر و لاینکس را در مدلهای خطی پویا مورد آزمون قرار دهیم و بدان وسیله قیمت حقیقی نفت را طی ۱۰۶سال از ۱۹۱۳ تا ۲۰۱۸ با توجه به مسئله "عدم تقارن"، فیلترینگ و سپس پیش بینی نمائیم. ما از مدل خطی پویا (DLM) برای برآورد "مدل منابع تجدیدناپذیر هتلینگ"، تحت توابع زیان مختلف استفاده کردیم. این برآوردها، بر اساس "عدم‌تقارن" جملات خطا صورت گرفته‌اند، و هدف اساسی در این مقاله پیش‌بینی قیمت نفت برپایه مدلهای خطی پویا تحت تابع زیان حداقل مربعات، کولبک-لایبلر، هلینگر و لاینکس است که در آن تابع زیان در برابر "بیش برآورد" یا "کم برآورد" حساس می‌باشند و ریسک بیش‌برآورد را در برآورد با شدت بیشتری جریمه می‌کند. برای این منظور، مدل خطی پویا بیزین را تحت توابع زیان فوق، برای پیش بینی یک گام جلوتر تصریح و برآورد نمودیم، و در نهایت نتایج پیش بینی به‌دست آمده را با یکدیگر مقایسه نمودیم. یافته‌های مقاله تاکید دارند که پیش بینی قیمت تحت تابع زیان غیرمتقارن لاینکس، از تابع زیان متعارف حداقل مربعات، هلینگر و کولبک-لایبلر، نتایج به مراتب دقیق‌تری از خود نشان می‌دهد. بویژه در حالتی که شاهد نوسانات و پارامترهای متغیر طی زمان باشیم. از آنجا که مسئله"عدم تقارن"در مطالعات اجتماعی-اقتصادی دارای اهمیت بیشتری است، انتخاب تابع زیان مناسب در این مطالعات بسیار معنادار می‌باشد.

کلمات کلیدی: مدل خطی پویا، توابع زیان لاینکس، کولبک-لایبلر، هلینگر، تابع زیان حداقل مربعات، پیش بینی قیمت نفت.

طبقه بندی JEL : C11, C22, C53, C61, Q47