



State Estimation and Forecasting of the Ski-Slope Model Using an Improved Shadowing Filter

Auni Aslah Mat Daud*

*The University of Western Australia,
35 Stirling Highway Crawley, Perth,
Western Australia 6009, Australia*

*School of Informatics and Applied Mathematics,
Universiti Malaysia Terengganu,
21030 Kuala Terengganu, Terengganu, Malaysia
auni_aslah@yahoo.com*

Received February 18, 2015; Revised November 5, 2015

In this paper, we present the application of the gradient descent of indeterminism (GDI) shadowing filter to a chaotic system, that is the ski-slope model. The paper focuses on the quality of the estimated states and their usability for forecasting. One main problem is that the existing GDI shadowing filter fails to provide stability to the convergence of the root mean square error and the last point error of the ski-slope model. Furthermore, there are unexpected cases in which the better state estimates give worse forecasts than the worse state estimates. We investigate these unexpected cases in particular and show how the presence of the humps contributes to them. However, the results show that the GDI shadowing filter can successfully be applied to the ski-slope model with only slight modification, that is, by introducing the adaptive step-size to ensure the convergence of indeterminism. We investigate its advantages over fixed step-size and how it can improve the performance of our shadowing filter.

Keywords: Forecasting; gradient descent; shadowing filter; ski-slope; state estimation.

1. Introduction

In recent years, a number of techniques have been established for finding shadowing trajectories [Davies, 1994; Hammel, 1990; Kostelich & Schreiber, 1993; Walker & Mees, 1997] and state estimation [Judd & Smith, 2004] of nonlinear dynamical systems. One method that has proven very powerful recently is the gradient descent of indeterminism (GDI) shadowing filter [Judd, 2003]. Gradient descent is an optimization technique that minimizes a quantity by moving continuously in the direction of steepest descent [Judd *et al.*, 2004a].

The gradient descent methods were originally demonstrated for simple chaotic systems.

Shadowing trajectories method is an important technique to assess the quality and the reliability of forecasting models and numerically compute trajectories of chaotic systems [Judd *et al.*, 2004b]. Shadowing plays an important role in the theory of indistinguishable states [Judd & Smith, 2004, 2001], which is a new approach to state estimation, ensemble and probabilistic forecasting. In this study, the gradient descent algorithm will be used because results have shown that, for a perfect

*Permanent Address: School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 4.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.

deterministic model, it will converge to shadowing trajectories if the observational noise is sufficiently small [Ridout & Judd, 2002].

There are two objectives of this study. One is to see if the standard gradient descent of indeterminism (GDI) shadowing filter can successfully be applied to a real world dynamical system, particularly one that exhibits chaotic properties. An example of such a system is a ski-slope model. The model has been introduced by Lorenz [1993]. However, it appears that it has not been investigated by any researcher until a few years ago when the generation of the hump pattern on a ski-slope was studied by Egger [2002]. The results of this study, which will be discussed in detail later, show that the original GDI shadowing filter fails to provide stability to the convergence of the errors. Therefore, it requires further research on the cause of the problem and possible approach to address the problem. This leads to our second objective of the study, that is, to propose a potential solution to the problem, namely the introduction of the adaptive step-size in the algorithm.

In this study, the ski-slope model is selected for two main reasons. Firstly, as stated above, the objective of this study is to see if the gradient descent algorithm can be successfully applied to real world, chaotic system such as the ski-slope model. Secondly, the application of the original gradient descent shadowing filter to the ski-slope model has a convergence problem. Therefore, further research is required on the cause of the problem and the possible approach to address the problem.

In this study, we introduce in addition an adaptive step-size algorithm for the GDI shadowing filter. We will also discuss some analysis of the performance of modified GDI shadowing filter using adaptive step-size and a comparison with the performance of GDI shadowing filter using fixed step-size.

Note that this paper is only concerned with the perfect model scenario (PMS), that is, the system under study has known dynamics, identical to the model. We restrict our attention to a deterministic model, that is, the dynamics of the model do not involve any random elements, and only observations are influenced by measurement noise.

The paper is organized as follows. Section 2 provides a description of the formulation and governing equations of four-dimensional ski-slope model.

In Sec. 3, we discuss the standard shadowing filter using gradient descent of indeterminism (GDI). Section 4 introduces some required modifications of the shadowing filter for the investigation of the ski-slope model. The results of the implementation of the standard and modified algorithm on the ski-slope model for state estimation and their usability for forecasting, including some unavoidable problems in forecasting the model, are presented and analyzed in Sec. 5. The discussions on the limitations of the adaptive step-size and the ski-slope model follow in Sec. 6.

2. Ski-Slope Model

To the best of our knowledge, there have been only two studies investigating the motion of a skier on a ski-slope. Lorenz may have been the first to model the motion of skiers on ski-slopes. He introduced the ski-slope model [Lorenz, 1993], where he considered the motion of sled and board on a ski-slope. He formulated the governing equations of motion, and discussed the assumptions employed in the modeling process. He presented some discussions on several chaotic properties of the ski-slope, including several common features of a low-dimensional nonlinear dynamical system such as the bifurcations and the attractor of the system.

Egger [2002] presented a microscopic model where the humps are assumed to be generated by the action of individual skiers and the tracks of the skiers are modified by the humps in a nonlinear process. The model reproduces many of the observed features of mogul fields, in particular, regular patterns which do not occur in a linear version of the model. He demonstrated that the generation of regular hump patterns at ski-slopes can be achieved with a model of relatively low complexity. In his model, a skier is able to keep his course against gravity and to make turns. Mini humps and holes are generated at every turn which then evolve into regular hump patterns.

Lorenz was more interested in chaotic motion on slopes with prescribed regular hump patterns but not in the formation of these moguls while in Egger's model, humps are generated by the action of many individual skiers which in turn affect the tracks of the skiers. In this study, we will consider a four-dimensional ski-slope model that has been introduced by Lorenz [1993].

2.1. The formulation of the model

To illustrate the application and performance of the shadowing filter, we will use the motion of a board on a ski-slope. We reproduce the Lorenz's model, in which the focus is on the motion of a board sliding down a ski-slope. To simplify our mathematical model, we will employ some assumptions. As have been described by Lorenz [1993], the motion of a board on a ski-slope is considered as a scattering by smoothly rounded moguls. It is assumed that the ski-slope has plenty of round moguls. The moguls are assumed to be physically similar and are uniformly spaced. Unlike in Egger's model, we ignore that the humps are generated and altered by the action of the skiers. We instead will prescribe a fixed topography for the mogul.

It is assumed that the motion of the board will be governed by the action of three forces. The first force, F_1 is the gravity force which is directed vertically downward. The second force, F_2 is the friction force, which is directed against the velocity. The third force, F_3 is the force that the slope exerts against the board, which is directed normal to the slope's surface (that is, right angles to the slope), and opposing the effect of gravity to just the extent needed to keep the board sliding instead of sticking to the slope or taking off into the air [Lorenz, 1993]. The forces diagram is shown in Fig. 1.

2.2. The governing equations of motion

The following derivation follows Lorenz [1993]. The equations describing the motion of a board on a ski-slope obeys the Newton's Second Law of Motion: the board's acceleration is equal to the sum of the forces acting on the board, per unit mass. Suppose X, Y, Z are the southward, eastward, and normal (to the slope) distances, respectively, U, V, W are velocity components in the downslope, cross-slope and normal (to the slope) directions, respectively, r is the coefficient of friction, and $H(X, Y)$ is the height of the slope above some horizontal reference plane. The equations of motion of the board are then

$$\frac{dX}{dt} = U, \quad (1a)$$

$$\frac{dY}{dt} = V, \quad (1b)$$

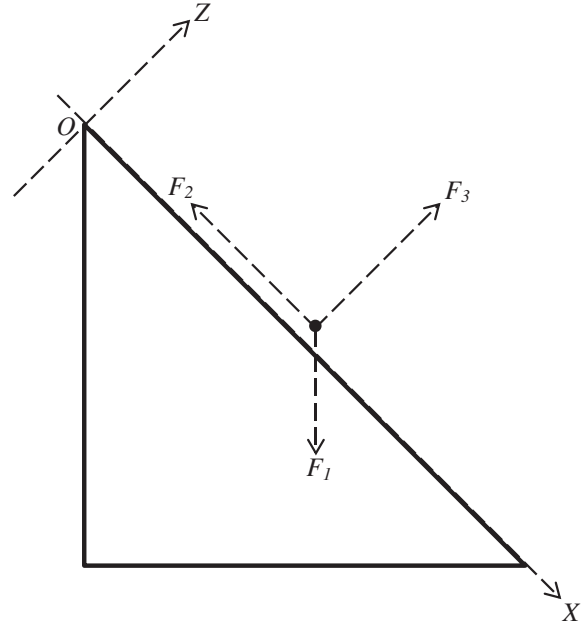


Fig. 1. A forces diagram. Note that the direction of arrows represent the direction of the forces, but the length of the arrows does not represent the magnitudes of the forces. The triangle represents the cross-section of the slope.

$$\frac{dZ}{dt} = W, \quad (1c)$$

$$\frac{dU}{dt} = -F \frac{\partial H}{\partial X} - rU, \quad (1d)$$

$$\frac{dV}{dt} = -F \frac{\partial H}{\partial Y} - rV, \quad (1e)$$

$$\frac{dW}{dt} = -g + F - rW, \quad (1f)$$

where F is the vertical component of the force of the slope against the board, g is the acceleration of gravity, $g = 9.81 \text{ m/s}^2$. Since

$$Z = H(X, Y), \quad (2)$$

on the slope, it follows that

$$W = U \frac{\partial H}{\partial X} + V \frac{\partial H}{\partial Y} \quad (3)$$

and

$$\begin{aligned} \frac{dW}{dt} = & -\frac{\partial H}{\partial X} \left(F \frac{\partial H}{\partial X} + rU \right) - \frac{\partial H}{\partial Y} \left(F \frac{\partial H}{\partial Y} + rV \right) \\ & + \left(\frac{\partial^2 H}{\partial X^2} U^2 + 2 \frac{\partial^2 H}{\partial X \partial Y} UV + \frac{\partial^2 H}{\partial Y^2} V^2 \right), \quad (4) \end{aligned}$$

eliminating W and dW/dt , gives

$$F = \frac{g + \frac{\partial^2 H}{\partial X^2} U^2 + 2 \frac{\partial^2 H}{\partial X \partial Y} UV + \frac{\partial^2 H}{\partial Y^2} V^2}{1 + \left(\frac{\partial H}{\partial X}\right)^2 + \left(\frac{\partial H}{\partial Y}\right)^2}. \tag{5}$$

We now define a four-dimensional ski-slope model as follows. Suppose r is chosen to be $r^{-1} = 2s$, and $H(X, Y)$ is given by

$$H = -aX - b \cos(pX) \cos(qY).$$

To be specific we choose in our numerical computation that $2\pi/p = 10.0$ m, $2\pi/q = 4.0$ m, $a = 0.25$ m and $b = 0.5$ m. Note that h , the height of a mogul above a neighboring pit is $2b$. From Eqs. (1)–(5) it follows that Eq. (1) can be reduced to the following set of governing equations of motion:

$$\frac{dX}{dt} = U, \tag{6a}$$

$$\frac{dY}{dt} = V, \tag{6b}$$

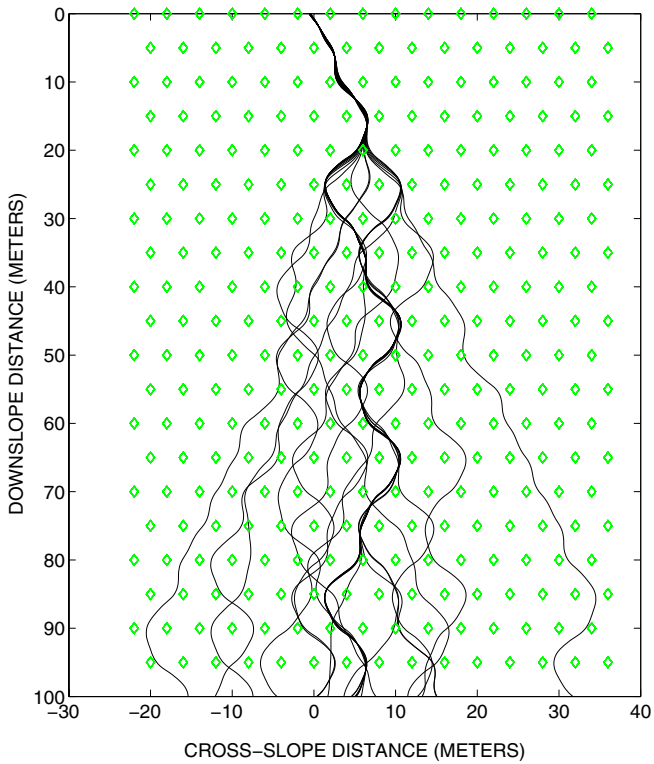


Fig. 2. Twenty typical simulated paths, computed numerically using Eqs. (6a)–(6d) above, where the diamonds represent the center of the moguls.

$$\frac{dU}{dt} = -F \frac{\partial H}{\partial X} - rU, \tag{6c}$$

$$\frac{dV}{dt} = -F \frac{\partial H}{\partial Y} - rV. \tag{6d}$$

2.3. The chaotic properties of the ski-slope model

An essential property of chaotic system is that nearby states will eventually diverge regardless of how small the initial differences are. As an example, we observe the simulated paths of 20 boards on a ski-slope model that start with identical velocities, but from slightly different points, where the starting points are chosen uniformly in an interval of 0.2 m. Let X , U , and V be 0.0 ms^{-1} , 4.0 ms^{-1} , and 2.0 ms^{-1} respectively, and Y varies from -0.2 to 0 . The paths of these boards are shown in Fig. 2. It can clearly be seen in the figure that, even by the first 25 m from the starting line, the small initial differences has been increased. Observe also that the tendency of the boards to deflect away from the moguls is clear, and the motion is chaotic.

3. Shadowing Filter by Gradient Descent of Indeterminism

Consider a discrete-time d -dynamical model given by $z_{i+1} = f(z_i)$. Given a sequence of observations $S = (s_1, s_2, \dots, s_n)$, a shadowing trajectory of S is defined as a sequence of states $Z = (z_1, z_2, \dots, z_n)$ such that Z is a trajectory of our model f and Z shadows S : for Z to be a trajectory requires $z_{i+1} = f(z_i)$, for $1, 2, \dots, n - 1$; and for Z to shadow S requires that distances $\|s_i - z_i\|$, for $i = 1, 2, \dots, n$, are small relative to a level of noise strength.

We assume the measurement noise to be additive and independent Gaussian random variates with mean 0 and variance σ^2 on each component. Therefore,

$$s_i = c_i + \xi_i$$

where the true trajectory is $C = (c_1, c_2, \dots, c_n)$, and the ξ_i are noise realizations.

Our aim now is to find a trajectory that is consistent with the observations, or shadowing trajectory. While there are a number of methods to do this, a straightforward method is by implementing gradient descent of indeterminism (GDI) shadowing filter [Judd & Smith, 2001; Stemler & Judd, 2009].

For any sequence of states Z the mean squared indeterminism function $I : \mathbb{R}^{nd} \rightarrow \mathbb{R}$ is defined by [Stemler & Judd, 2009]

$$I(Z) = \frac{1}{n-1} \sum_{i=1}^n \|z_{i+1} - f(z_i)\|^2. \quad (7)$$

Note that this scalar function is a measure of the average mismatch between states and forecasts. Therefore, it measures how far a sequence of states is from being a trajectory. A sequence of states will be a trajectory if and only if the indeterminism is zero.

If one wants to find a shadowing trajectory Z from given noisy observations S , the indeterminism can be used. Starting from $I(S) \neq 0$ (for the observation s the indeterminism $I(S)$ is almost surely nonzero), gradient descent can be used to minimize I . The gradient descent method then follows the steepest descent of the gradient of $I(z)$ down to a minimum where $I(z) = 0$. An explicit implementation of GDI is to solve the following differential equation

$$\frac{dz}{d\tau} = -\nabla I(z(\tau)), \quad (8)$$

where $z(0) = S$, τ is the time interval (or sampling frequency) and in the limit of $\tau \rightarrow \infty$ we obtain a shadowing trajectory. It can be shown that the GDI method always converges to a trajectory of the model [Ridout & Judd, 2002], that is, as $I(Z)$ converges monotonically to zero, then Z converges to a trajectory of f . A more detailed discussion of the properties of GDI shadowing filter can be found in [Judd & Smith, 2004; Judd *et al.*, 2004a; Judd & Smith, 2001; Stemler & Judd, 2009; Ridout & Judd, 2002].

While Eq. (8) can be solved using various methods, we used a fixed step Euler integration method as it is more straightforward. Letting $Z_m = (z_{1,m}, z_{2,m}, \dots, z_{n,m})$, where $z_0 = S$, the equation is reduced to

$$\begin{aligned} & z_{i,m+1} \\ &= z_{i,m} - \frac{2\Delta}{n-1} \\ & \times \begin{cases} -A(z_{i,m})(z_{i+1,m} - f(z_{i,m})), & i = 1 \\ z_{i,m} - f(z_{i-1,m}) \\ -A(z_{i,m})(z_{i+1,m} - f(z_{i,m})), & 1 < i < n \\ z_{i,m} - f(z_{i-1,m}), & i = n \end{cases} \end{aligned} \quad (9)$$

where the subscript m is the iteration number, Δ is a suitably chosen *step-size* and $A(z)$ is the adjoint of f (transpose of the Jacobian matrix) evaluated at z . $J(z)$ is the Jacobian of f at z , defined as

$$J_{ij}(z) = \frac{df_i}{dz_j}.$$

The algorithm may work for arbitrary positive choices of $2\Delta/(n-1)$ smaller than 1. The principal test to determine Δ is by observing the convergence of the indeterminism $I_m = I(Z_m)$, which should be strictly decreasing, as the reason GDI shadowing filter is used is to minimize I_m . Larger values of Δ tend to give faster convergence, but tend to increase the possibility of failure, that is, the indeterminism will not be strictly decreasing [Stemler & Judd, 2009].

With this method iterative time discrete models can be solved but our ski-slope models are flows of general form $\dot{z} = F(z)$. If we assume that the observations are with a constant sampling frequency τ , our task is to find a map f for the flow. Given a sequence of observations $S = (s_1, s_2, \dots, s_n)$, we will implement the GDI shadowing filter to find $z(t)$ of $\dot{z} = F(z)$, such that $Z = (z(0), z(\tau), \dots, z((n-1)\tau))$ shadow S . This can be done by integrating $\dot{z} = F(z)$, with $z(0) = x$, for a time interval τ to obtain a map $f(x) = z(\tau)$. Suppose $z(t)$, $0 \leq t \leq \tau$ is the solution, an adjoint product can be computed by solving the homogeneous differential equation

$$\frac{du}{dt} = J(z(\tau-t))^T u, \quad u(0) = v; \quad (10)$$

and the adjoint product is given by $u(\tau)$. Equations (6a)–(6d) define our flow $\dot{z} = F(z)$, and $z = (X, Y, U, V)$ for the ski-slope model.

The adjoint products are computed using analytic approximation.

3.1. Definitions of measured quantities

In this study, we assess the performance of the GDI shadowing filter using some measurements which have been defined in the previous study [Stemler & Judd, 2009]. The quality of the estimated states Z_m are investigated by measuring how close Z_m is to being a trajectory, that is, its indeterminism I_m [as defined in Eq. (7)], and the final state mismatch magnitude, $I_{n,m} = \|z_{n,m} - f(z_{n-1,m})\|$. We measure the distance between the estimated states Z

and the true trajectory $C = (c_1, c_2, \dots, c_n)$, by the root mean square error of states

$$E_m = \sqrt{\frac{1}{n} \sum_{i=1}^n \|z_i - c_i\|^2} \quad (11)$$

and the last point error $E_{n,m} = \|z_{n,m} - c_n\|$. We measure the usability of our state estimate for forecasting, using a quantity called separation time, or shadowing time, which is the largest lead time for which the forecast error remain less than some threshold.

$$T_m = \max\{T : \|c_{n+t} - f^t(z_{m,n})\| < 2\sigma, \forall 0 \leq t \leq T\}. \quad (12)$$

3.2. The windowing test

Before applying our shadowing filter, there are a number of adjustable parameters to be determined. The parameters are the step-size Δ , number of iterations to achieve convergence m , the integration interval τ , and the number of steps for adjoint computations w . We obtained the optimal values of these parameters using the windowing test. It is a basic method to find optimal or appropriate values of parameters for shadowing filters [Stemler & Judd, 2009]. For a trajectory of $n = 10$ points, we obtained the following optimal values of the parameters: $n = 10$, $m = 100$, $2\Delta/(n - 1) = 0.1$, $w = 20$, and $\tau = 1$ sec.

4. Improved GDI Shadowing Filter with Adaptive Step-Size

It is common feature of iterative optimization algorithms that there is a trade-off between the amount of computation and the accuracy of results. There is also often a trade-off between the stability of the algorithm and the rate of convergence. Recall the discussion on the algorithm of the GDI shadowing filter, a fixed step-size Δ is used in the iteration [see Eq. (9)]. The choice of Δ results in a trade-off between the stability and the rate of convergence. Generally, a small value of Δ will ensure stability, but provide a slow rate of convergence. The principle criteria of determining the value of the step-size is by observing the convergence of the indeterminism, which must always be decreasing.

Note that when a gradient descent problem is implemented as an ordinary differential equation (ODE) as in Eq. (9), then the rate of convergence

can be slow because this ODE is a stiff equation [Judd *et al.*, 2004a]. The gradient descent problem, Eq. (9), is inherently stiff, because the linearizations of the models have modes with small eigenvalues that may later become unstable. These potential instabilities in shadowing algorithms would result in poor quality of solutions, or worse, failure of convergence [Judd *et al.*, 2004a]. Furthermore, since $I(z)$ in Eq. (7) is defined by a quadratic form, the rate of convergence will slow or decrease as the solution is approached. Note that in Eq. (9) a fixed step-size is used in the iteration. We propose to employ adaptive step-sizes to overcome these difficulties. We define adaptive step-size Δ , as a number of different values of step-size in the iterative GDI shadowing filter (but must always be decreasing), which are changed to avoid the convergence failure of a shadowing filter. Adaptive step-size can decrease the rate convergence by reducing the value of the step-size.

GDI can fail if the step-size is too large, but is slow if step-size is too small. Failure here means that the indeterminism does not decrease when a step is taken. It is not necessarily a good idea to choose the largest step that results in a decrease of indeterminism, but certainly if a step-size does not cause a decrease, when a smaller step-size does, then the step-size should be reduced. This is the idea behind the approach implemented here.

4.1. The implementation

The flowchart in Fig. 3 summarizes the algorithm.

The essential idea is to adapt the step-size based on the convergence of the indeterminism to maintain a high rate of convergence without compromising the stability. That is, the need to reduce the step-size is based on the comparison between the indeterminism of the current and the previous iterations. If it appears that the current indeterminism is greater than or equal to the indeterminism in the previous iteration, then the step-size is reduced by a factor k in the next iteration and the state estimates of the previous iteration are filtered again but using a new reduced value of the step-size. On the other hand, if the indeterminism decreases, the step-size remains unchanged for the next iteration.

The adaptive step-size can be implemented by the following algorithm:

- (1) Set the initial state estimates, $Z_0 = S$ and the initial step-size Δ_0 . Compute the initial

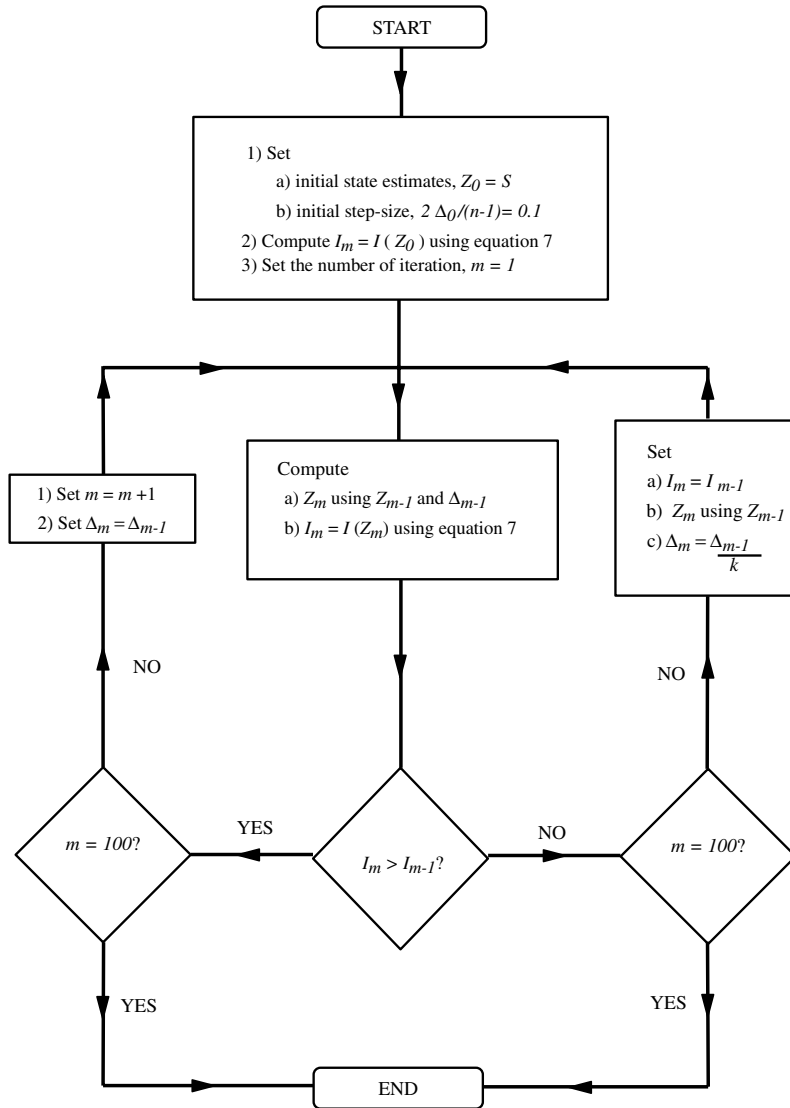


Fig. 3. The flow chart to show the steps involved in a GDI shadowing filter with adaptive step-size.

indeterminism $I_0 = I(Z_0)$. Then set the number of iterations, $m = 1$.

- (2) Apply GDI shadowing filter to obtain Z_m using Z_{m-1} and Δ_{m-1} . Compute the indeterminism $I_m = I(Z_m)$. If $I_m \geq I_{m-1}$, then divide Δ_m by a factor k and set $m = m - 1$, the state estimates $Z_m = Z_{m-1}$ and $I_m = I_{m-1}$. If $I_m < I_{m-1}$ then set $m = m + 1$ and $\Delta_m = \Delta_{m-1}$.
- (3) Repeat Step 2 for 100 number of iterations.

4.2. Choice of adaptive factor

If the step-size is divided by a factor k when the indeterminism increases, then what is the optimal or appropriate value for k ? Although we initially considered several values of k , we found by experiment that there was no significant difference in any

of them. It appears that the algorithm works for any positive choices of k greater than 1. Our results suggest that, in most cases, a value of 2 will provide good solutions. Therefore, we use $k = 2$ in the following experiments. If $k \gg 1$ then there is no progress to convergence, because the step becomes too small too quickly, hence k should not be too large.

4.3. The initial step-size

Since the step-size is varied throughout the iteration, what is the optimal value for the initial step-size Δ_0 ? The convergence is ensured, that is, the indeterminism will always be decreasing or at least in the worst case, nonincreasing. The only concern is the speed of the convergence. Although the GDI

shadowing filter using adaptive step-size may work for arbitrary positive choices of $2\Delta_0/(n-1)$ less than 1, we find that a value around $2\Delta_0/(n-1) = 0.1$ generally gives the fastest rate of convergence.

5. Results and Discussions

5.1. State estimation using fixed step-size

To demonstrate the quality of the GDI shadowing filter to the ski-slope model, we consider 100 different trajectories of the ski-slope model, which are generated using different initial conditions. The trajectory or the ski-slope data are generated by

solving Eqs. (6a)–(6d) using the function ode45 of MATLAB, which is based on an explicit Runge–Kutta formula. Every trajectory has ten points. We test the iterative GDI shadowing filter where $2\Delta/(n-1) = 0.1$ for observation data having three different noise strengths, expressed in relative standard deviation (in %) $RSD = 5, 10$ and 15 . Figure 4 shows the average values of $I_m, E_m, I_{n,m}, E_{n,m}$ and Δ_m as a function of the number of iterations m .

The filter gives around an order of magnitude reduction of $I_{n,m}$ and I_m , for all of the three noise levels. However, it can be clearly observed in the figure that the GDI shadowing filter is unstable and fails to provide the convergence, that is, E_m increases as m increases, while $E_{n,m}$ decreases for

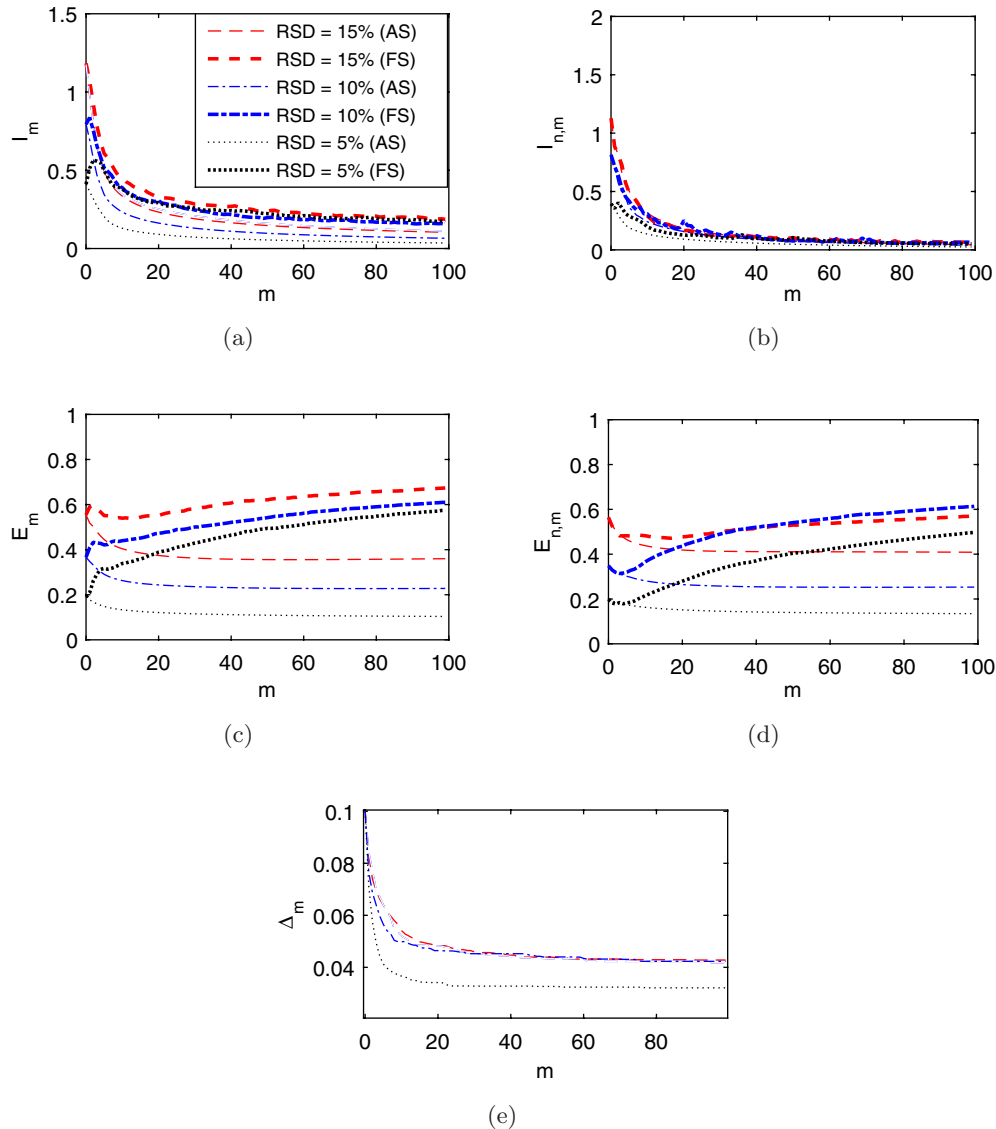


Fig. 4. The average value of $I_m, E_m, I_{n,m}, E_{n,m}$ and the term Δ_m as a function of m , for $RSD = 5, 10$ and $15, m = 100, k = 0.5$, and the $2\Delta_0/(n-1) = 0.1$. AS represents adaptive step-size and FS represents fixed step-size.

the first few iterations and then begin to increase as m increases. Again, this convergence issue requires some modification to the filter.

Observe also in Fig. 4 that, on average, $I_{n,m} < I_m$. This feature is a consequence of the limited information available to the final state of the sequence. The final state $z_{m,n}$ only has to make adjustments to mismatches on one side, whereas a state in the middle of the sequence has to make adjustments to mismatches on both sides, which is usually harder to achieve and requires more iterations to converge, and hence $I_{n,m} < I_m$.

5.2. State estimation using adaptive step-size

To investigate the performance of the GDI shadowing filter using adaptive step-size, we computed the average values for I_m , E_m , $I_{n,m}$ and $E_{n,m}$ using the same initial observation data as the previous experiment. Figure 4 shows the average value of I_m , E_m , $I_{n,m}$, $E_{n,m}$ and the term $2\Delta/(n-1)$ as a function of m , for RSD = 5, 10 and 15, where the same data from the experiment in the previous experiment are used.

It can be clearly observed from Fig. 4 that, using adaptive step-size, the average value of all four quantities decreases as the number of GDI iterative steps, m increases. Also observe in Fig. 4 that using adaptive step-size, after $m = 100$ iterations, I_m and $I_{n,m}$ are reduced slightly more than the fixed step-size. E_m decreased by a factor of about 1/3, and $E_{n,m}$ decreased by about 1/4. There is a slower rate of decrease after around ten iterations, with almost the same shape of decrease for all three noise levels. This feature can be understood as follows. Typically, in the first ten iterations of the shadowing filter algorithm using adaptive step-size, the big step-sizes removes large mismatches from the states that are caused by large observational errors and in the following iterations the algorithm is adjusting states using smaller step-sizes, to achieve convergence to a trajectory.

5.3. Forecasting using fixed step-size

We now consider the quality of forecasts. Recall that we defined the separation time T_m as the maximum time for which the forecast error remains less than a threshold of 2σ . It is useful to compare T_m , for

$m > 0$ with the value T_0 , that is, the forecast from the raw unfiltered observation. Using the same data as the previous subsection, we computed the average value of T_m for the ski-slope model, plotted as a function of the number of iterations in Fig. 5.

Note that the optimal number of iterations $m = 100$ is obtained using the windowing test, as mentioned in Sec. 3.2. That is, sufficient convergence is achieved and the algorithm will be stopped after 100 iterations. As m increases, better state estimates are obtained as shown in Fig. 4. Therefore, it should be expected that the separation time for better state estimates will increase as m increases.

There are two main points that can be observed from the figure. First, the shadowing filter provides improved or longer separation time, sometimes more than two units. This is because the forecast using the state estimates (the noise-reduced trajectory) will remain close to the future state of the true trajectory longer than the observations (noisy trajectory). Second, note from the figure that the average separation time is not strictly increasing, for all three noise levels, for example, there are a number of spikes in the average separation time. The spiking phenomenon observed in the figure is generally attributed to forecasted trajectories being deflected to the wrong side of a hump on the ski-slope that is directly in the path of the skier. For all noise levels, there can be a situation where better estimated states (the ones with smaller error) give worse forecast than the worse estimated states. Investigations show us that this rare case only happens if the trajectory forecasted by the better estimated states happen to be moving closer to a hump and deflected to the wrong side of a hump. On average, such cases happen 26 times, out of 100 sequences tried. An example of such unavoidable phenomena is illustrated in Fig. 6. The black (thick) trajectory is the truth, the red (solid) trajectory is the trajectory forecasted using state estimates after fifth iteration and the blue (dashed) trajectory is the forecasted trajectory using seventh iteration. Note that the hump at $(-20, -55)$ is the one that causes the problem. Therefore, we conclude that the phenomenon does not indicate the failure of the shadowing filter to obtain better estimate, but rather merely a consequence of the tendency of a skier to be deflected on the wrong side of a hump. This behavior is unavoidable, because in our formulation of the model, it is assumed that a skier might tend to avoid moving over a hump.

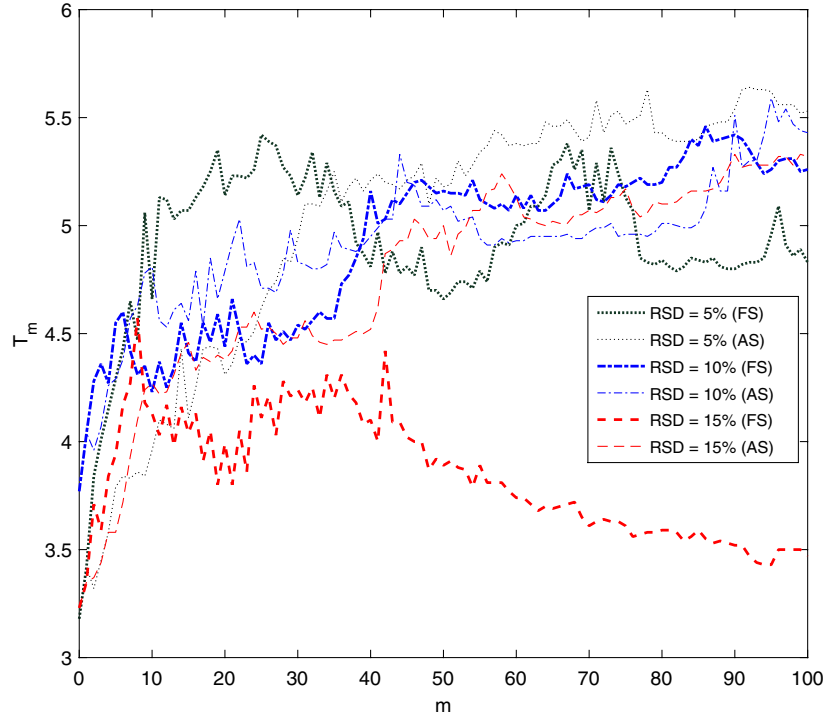


Fig. 5. The average separation time T_m as a function of number of iterations m . AS represents adaptive step-size and FS represents fixed step-size.

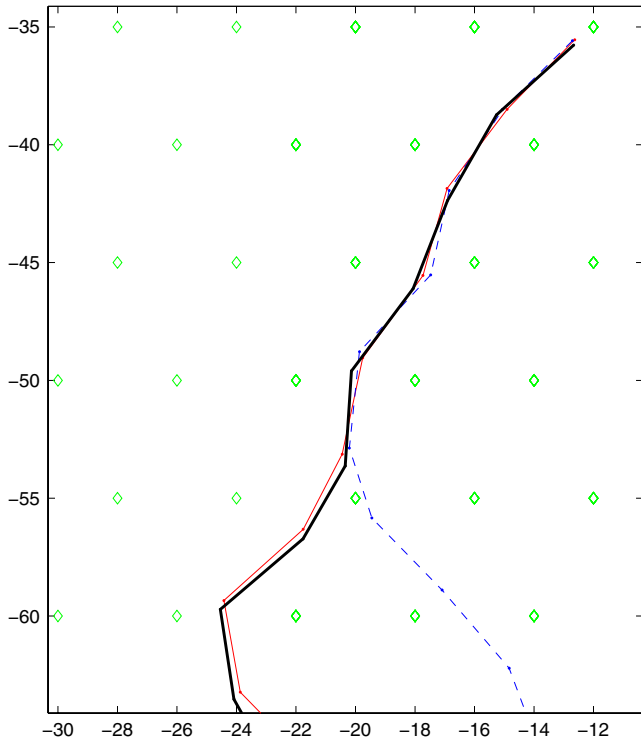


Fig. 6. An example of a ski-slope trajectory where better state estimates give worse forecast than worse state estimates.

5.4. Forecasting using adaptive step-size

Figure 5 shows the average separation time T_m as a function of the number of iterations m , for adaptive step-size, using the same data used in the experiment for the ski-slope model in the previous experiment.

Observe in the figure that adaptive step-size performs better than fixed step-size for all noise levels. This can be explained by the better performance of adaptive step-size in reducing the average value of I_m , $I_{n,m}$, E_m and $E_{n,m}$, compared to fixed step-size. However, the spiking phenomenon still occurs even when we employ adaptive step-size, because as previously explained this is an unavoidable property of the system.

6. Limitations

The limitation of this research is that so far we have only focused on the perfect model scenario, which is a mere hypothetical. In reality, the model and the observations of the systems are never perfect. Therefore, it is hoped that in the future, some of

the results presented in this note can be generalized to the imperfect model scenario (IPMS). Some limitations of this method can be improved in the future. There are a variety of possible methods to adapt the step-size, but the approach proposed in this paper is an example of a simple and straightforward method. One of the simplest methods that has been proposed is scaling the step-size by the gradient vector, that is, divide Δ by $\|\partial I/\partial z\|$. The method provides a more uniform, almost linear rate of convergence. However, it has been shown that this simple adaptive step-size scheme excited instabilities in the jet-stream of the atmosphere when applied to a weather model [Judd *et al.*, 2004a], and other schemes should be considered.

There are at least two limitations of the model. Firstly, the ski-slope model would be more realistic if the friction force were made to be proportional to the force of the slope against the board. For example, when the board is nearly taking off, presumably because it is shooting over a mogul, the frictional effect will be greatly reduced. The assumption that the force of friction is proportional to velocity is convenient, but is quite controversial on account of its oversimplicity. It has been shown that in some experiments, that it is a plausible assumption, at least as an approximation [Brauer, 2001]. Another noticeable limitation of the model is that it does not deal with sliding objects of various sizes and shapes. Therefore, if a skier slides down a slope using a slightly longer board, this will not make any difference in terms of the trajectory. We treat the board as if it is a single particle, and disregard the flexibility of real boards and the variability of their orientations.

7. Conclusion

This paper presents two important findings: the study of the application of the GDI shadowing filter to the ski-slope model and the refinement of the standard shadowing algorithm by introducing a new approach in selecting the step-size. The application of gradient descent shadowing algorithms to the ski-slope model appears to be successful and it has shown that the method can be applied practically. However a minor modification, that is by using adaptive step-size, instead of fixed step-size, in the gradient descent algorithm can further improve the performance of the algorithm, particularly in

estimating states of true trajectories. It is designed to reduce the indeterminism at each iteration, eliminate the uncertainty in the selection of the optimal value for step-size parameter, and to increase the speed of convergence without compromising the stability.

It has been demonstrated that the adaptive step-size performs significantly better than the fixed step-size. One main problem is that the original GDI shadowing filter fails to provide stability to the convergence of the root mean square error and the last point error of the ski-slope model. A simple solution to this problem is proposed, that is, by employing a new adaptive step-size approach designed to ensure the convergence of indeterminism during each iteration. The adaptive step-size will guarantee the average value of I_m and E_m to be strictly decreasing, or at least nonincreasing. Although a bigger step-size will generally give faster convergence, the possibility of failure is higher. That is, the convergence of the indeterminism may not be strictly decreasing. Although fixed step-size has a faster initial convergence in some cases, it is eventually surpassed by adaptive step-size. This step-size is important to ensure successful application of GDI shadowing filter and hopefully will further enhance the performance of the original shadowing filter.

There are also unexpected results with the quality of forecasts for the ski-slope model. There are cases where the improved or better state estimates do not improve the quality of forecasts. After further investigations, we found out that this phenomenon is due to the existence of the moguls which change the direction of the skier and hence reduce the separation time.

Acknowledgments

I am very grateful to Prof. Kevin Judd and Dr. Thomas Stemler for supervising this research. This study was conducted at The University of Western Australia (UWA) as part of my doctoral research [Mat Daud, 2012]. The study was financially sponsored by Ministry of Higher Education, Government of Malaysia.

References

Brauer, F. [2001] "What goes up must go down, eventually," *Amer. Math. Month.* **108**, 437–440.

- Davies, M. [1994] “Noise reduction schemes for chaotic time series,” *Physica D* **79**, 174–192.
- Egger, J. [2002] “Hump pattern generation on ski-slopes,” *Physica D* **165**, 127–136.
- Hammel, S. M. [1990] “A noise-reduction method for chaotic systems,” *Phys. Lett. A* **148**, 421–428.
- Judd, K. & Smith, L. A. [2001] “Indistinguishable states I: Perfect model scenario,” *Physica D* **151**, 125–141.
- Judd, K. [2003] “Nonlinear state estimation, indistinguishable states and the extended Kalman filter,” *Physica D* **183**, 273–281.
- Judd, K. & Smith, L. A. [2004] “Indistinguishable states II: The imperfect model scenario,” *Physica D* **196**, 224–242.
- Judd, K., Reynolds, C. & Rosmond, T. [2004a] “Towards shadowing in operational weather prediction,” Tech. Rep. Technical Report NRL/MR/7530-04-18, Naval Research Laboratory, Monterey, CA, USA.
- Judd, K., Smith, L. A. & Weisheimer, A. [2004b] “Gradient free descent: Shadowing and state estimation using limited derivative information,” *Physica D* **190**, 153–166.
- Kostelich, E. J. & Schreiber, T. [1993] “Noise reduction in chaotic time-series data: A survey of common methods,” *Phys. Rev. E* **48**, 1752–1763.
- Lorenz, E. N. [1993] *The Essence of Chaos* (University of Washington Press).
- Mat Daud, A. A. [2012] “State estimation and forecasting using a shadowing filter applied to quincunx and ski-slope models,” PhD thesis.
- Ridout, D. & Judd, K. [2002] “Convergence properties of a gradient descent noise reduction,” *Physica D* **165**, 26–47.
- Stemler, T. & Judd, K. [2009] “A guide to using shadowing filters for forecasting and state estimation,” *Physica D* **238**, 1260–1273.
- Walker, D. M. & Mees, A. I. [1997] “Noise reduction of chaotic systems by Kalman filtering and by shadowing,” *Int. J. Bifurcation and Chaos* **7**, 769–779.