

# An Inverse Problem: Trappers Drove Hares to Eat Lynx

Bo Deng<sup>1</sup>

**Abstract:** The Canadian lynx and snowshoe hare pelt data by the Hudson’s Bay Company do not fit the classical predator-prey theory. Rather than following the peak density of the hares, the peak density of the lynx leads it, creating the hares-eat-lynx (HEL) paradox. Although trappers were suspected to be a cause, no mathematical models in the literature have demonstrated the HEL effect. In this paper we consider various multitrophic models with interactions in vegetation-hare-lynx, hare-lynx-coyote, hare-lynx-trapper, and hare-lynx-coyote-trapper, and use Newton’s gradient search method to best fit each model to the data and then select the one with the least error as the benchmark model for the data. We will conclude from the benchmark model and its sensitivity analysis the following: (a) the trappers as a participant rather than an observer of the system changed the observed; (b) the lynx and hare populations in the wild follow the lynx-eats-hares (LEH) cycle but the harvested animals follow the HEL cycle; (c) the benchmark fit is more sensitive to changes in all lynx-trapper interactions than the respective hare-trapper interactions; (d) trappers did not interfere each other’s trapping activities a century ago; (e) the Hudson’s Bay Company’s hare pelt number was severely fewer than it should be. These results together dispel a long-held hypothesis that the pelt data is a proxy of the hare and lynx populations in the wild. It also shows that theoretical ecology must move beyond the classical Lotka-Volterra paradigm into a contemporary framework in which Holling’s predation theory is central to population modeling. It also demonstrates to the ecologists that we do not have to collect data from all dimensions in order to gain a good understanding of complex systems, but instead we can systematically fit models to empirical data of any dimensions with identifiable uncertainties and a complete understanding of parameter sensitivities to the best fit.

*Key Words:* Canadian lynx and snowshoe hare, population modeling, ecological and economical modeling, Holling’s disc function, dimensional analysis, least square fit, Newton’s gradient search method, the line search method, global minimizer, parameter sensitivity to least error, benchmark model

*“I find the whole oscillation most mysterious.”* —The closing words of a 60-page mathematical analysis of the Hudson’s Bay Company’s hare and lynx fur data by Egbert R. Leigh, Jr. in 1968.

**1. Introduction.** The Hudson’s Bay Company’s fur trade data ([8, 25]) for the Canadian lynx and snowshoe hares is the oldest, the longest, and the most well-known data set in ecology. It has been extremely controversial and remains enigmatic to this day. Every book on introductory ecology must pay tribute to it, most (e.g. [26]) would cite it as an example to support the classical predator-prey theory by Lotka ([23]) and Volterra ([39]), but only a few (e.g. [2]) would point it out correctly that it does not. Unlike the classical theory which predicts the peak density of the predator follows that of the

---

<sup>1</sup>Department of Mathematics, University of Nebraska-Lincoln, Lincoln, NE 68588. Email: bdeng@math.unl.edu

prey, the peak volume of the hare's pelt on average follows that of the lynx's. This shocking paradox was discovered by Leigh ([22]) and made widely known by Gilpin with an article provocatively titled *Do hares eat lynx?* ([11]). The net phase difference is more than two years with the lynx pelt data phase-advanced by one year on average from the hare pelt ([22, 11, 3]) but the lynx's field population phase-lagged by one year from the hare's population ([17, 28, 35, 15]). Data aggregation cannot be the problem because pelts from different boreal regions of Canada were shown to be spatially synchronized in time ([1]). Data mismatch in volume cannot be the problem either as we will show that the amplitudes of the pelt series do not effect the phase of their oscillation. So could it be the trading practice by the native trappers, or the bookkeeping practices by the Hudson's Bay Company, or both that shifted the lynx phase by two years? This is highly unlikely for two reasons: statistical averaging over large data set usually erase peculiarities and the absence of evidence for sloppy or idiosyncratic practices by the company is probably the absence of such systematic practices. The phase divergence in opposite directions is so 'mysterious' ([22]) that it prompted a recent statistical study ([42]) to suggest that the pelt data is the result of 'intrinsic self-regulation' of each hare and lynx.

Although trappers were suspected to play a role for the hares-eat-lynx (HEL) paradox ([9, 40, 41]), all statistical and mathematical studies in the literature (e.g. [3, 33, 34, 35, 1, 10, 18, 37, 42, 38]) have made a tactical assumption that the pelt numbers is a direct, i.e. proportional, proxy of the populations in the wild, effectively rendering trappers role nonessential. Fig.2 of [35] was cited as an empirical basis for this assumption but the opposite can be equally inferred. Even at the conceptual level this proxy assumption is incredibly simplistic for obvious reasons. Just to name a few: trappers were not field ecologists but resource exploiters who were economically vested in if not entirely depended on the animals for survival — taking out the animals in large quantity irreplaceably for food and trade; and like a natural predator they adjusted their efforts in pursuit of the preys ([9, 40, 41]). These facts alone suggest that the trappers were too deeply embedded in the system to be excluded from any mathematical model aimed at explaining their catch data or the hare-lynx interaction in the wild that the data implies. This is exactly what we will do in this paper. We will demonstrate among other things that the trapped animals are expected to exhibit the paradoxical prey-eats-predator oscillation while the populations in the wild always follow the classical predator-eats-prey chase. That is, these two contradictory properties are for the first time unified by our benchmark model. The Hudson's Bay Company's data is qualitatively good, and had it not shown the HEL effect it would have been questionable, a conclusion contrary to the conventional approach to explain it away ([19, 38]).

Because of the HEL paradox the failure of the classical Lotka-Volterra model to the hare-lynx system was spectacular. Leigh's work ([22]) was not only the first attempt to fit a Lotka-Volterra model to the fur data but also the most comprehensive mathematical analysis of any model for the hare-lynx system in the past forty plus years. The HEL legacy he left behind was long lasting. For one thing mathematicians and ecologists since then have stopped fitting their conceptual models to the data. As a result we have failed to establish any followup benchmark model. As a more serious consequence we have failed to connect mathematics to reality, which biologists were right to question the role of mathematics in ecological research ([16, 12]). (Statistical analysis is not considered strictly as mathematical

modeling here but rather an extension of observation or a tool of observation by experimentalists.) The HEL problem has raised some basic questions as to can a piece of mathematics be called a mathematical model without ever being best-fitted to an empirical data? How to objectively select them from such numerous and often arbitrary choices? And what is knowable and what is not when a model of high dimensions is fitted to a low dimensional data set?

In this paper we will consider several models for the Canadian hare-lynx problem. We will make the following distinctions among models. A model without being best-fitted to a data is referred to as a *conceptual* model. A conceptual model that is best-fitted to a data is referred to as a *provisional* model. A provisional model that is the best amongst all provisional models is referred to as the *benchmark* model. To best fit a model to a data is to have the least error between the predicted by the model and the observed from the data. To determine system parameter values for the least error is to solve the so-called inverse problem in mathematics, and the most effective method to solve inverse problems is by Newton's gradient search method for which the most effective implementation is the line search method ([30, 32]). The model selection protocol outlined above is referred to as benchmarking. All provisional models in this paper are determined mainly by the line search method. We will demonstrate that all models without the trappers do not exhibit the HEL effect but all models with the trappers do, and of all the models with or without the trappers, the hare-lynx-coyote-trapper (HLCT) model has the least error, hence is qualified to be the benchmark model. We will comment on some seemingly 'bad' parameter fit but apparently for some perfectly good reasons which in turn lead us to conclude that the hare pelt series is severely under recorded for a few orders of magnitude. We will also introduce for the first time to our best knowledge a sensitivity analysis of the best fit to show expectedly that the native trappers seemed to value their lynx furs more than their hare furs, but also to show unexpectedly that they did not interfere each others trapping activities. We will also introduce for the first time an uncertainty analysis of fitting high dimensional models to low dimensional data to show that despite the dimensional deficiency some system parameters can be uniquely determined by the best fit.

**2. Method.** In this paper we will first introduce various conceptual models in differential equations with as much mechanistic justifications as possible. We will then use Newton's gradient search method and its most effective implementation — the line search method — to best fit each conceptual model to the lynx-hare data from [22, 11]. It is by the benchmark model that our observations and conclusions about the Canadian hare-lynx system will be derived.

**2.1. Models.** For variable notations we will denote by  $t$  the time in year with  $t = 0$  corresponding to 1875 the first data year of [22, 11]. We will use  $H(t), L(t)$  for the head counts of hare and lynx in the wild at time  $t$  that we often breviate to  $H, L$ , suppressing the time variable. According to [28], coyotes is a main predator of the hares just as lynx is among many other ancillary predators ([30, 34, 28]). Hence variable  $C = C(t)$  is used as a proxy for all predators, the coyotes in particular, other than the lynx and instead of a natural number it is simply a nonnegative natural number using as an index to measure the predatory effect on the hares. It can be interpreted as the total biomass if the average size of the coyotes is not specified. Similarly, variable  $T = T(t)$  is used as an index for the trappers, a proxy for the trapping effect on the hare and the lynx rather than the head count or family count or tribe count

Table 1: Model Variables and Parameters

Parameter	Definition	Unit
$H$	Hare's population	Natural Number
$L$	Lynx's population	Natural Number
$C$	Coyote's population index	Scalar
$T$	Trapper's population index	Scalar
$b$	Hare's per-capita growth rate	1/[year]
$m$	Hare's intraspecific competition rate	1/[ $H$ ]/[year]
$a_1$	Hare's encounter rate with Lynx	1/[year]
$h_1$	Lynx's handling time of Hare	[year]/[ $H$ ]
$a_2$	Hare's encounter rate with Coyote	1/[year]
$h_2$	Coyote's handling time of Hare	[year]/[ $H$ ]
$u_1$	Hare's encounter rate with Trapping	1/[year]
$v_1$	Trapper's handling time of Hare	[year]/[ $H$ ]
$u_2$	Lynx's encounter rate with Trapping	1/[year]
$v_2$	Trapper's handling time of Lynx	[year]/[ $L$ ]
$b_1$	Lynx's consumption-to-birth ratio	Scalar
$d_1$	Lynx's per-capita death rate	1/[year]
$m_1$	Lynx's intraspecific competition rate	1/[ $L$ ]/[year]
$b_2$	Coyote's consumption-to-birth ratio	Scalar
$d_2$	Coyote's per-capita death rate	1/[year]
$m_2$	Coyote's intraspecific competition rate	1/[ $C$ ]/[year]
$r_1$	Trapper's Hare-to-recruitment ratio	Scalar
$r_2$	Trapper's Lynx-to-recruitment ratio	Scalar
$d_3$	Trapper's per-capita quit rate	1/[year]
$m_3$	Trapper's intraspecific competition rate	1/[ $T$ ]/[year]

of trappers. The variables and parameters are listed in Table 1, and the theoretical model is given as follows:

$$\begin{cases} \frac{dH}{dt} = H \left( b - mH - \frac{a_1 L}{1 + h_1 a_1 H} - \frac{a_2 C}{1 + h_2 a_2 H} - \frac{u_1 T}{1 + v_1 u_1 H + v_2 u_2 L} \right) \\ \frac{dL}{dt} = L \left( \frac{b_1 a_1 H}{1 + h_1 a_1 H} - d_1 - m_1 L - \frac{u_2 T}{1 + v_1 u_1 H + v_2 u_2 L} \right) \\ \frac{dC}{dt} = C \left( \frac{b_2 a_2 H}{1 + h_2 a_2 H} - d_2 - m_2 C \right) \\ \frac{dT}{dt} = T \left( \frac{r_1 u_1 H + r_2 u_2 L}{1 + v_1 u_1 H + v_2 u_2 L} - d_3 - m_3 T \right). \end{cases} \quad (1)$$

Explanation for the model is as follows. Without the predators and trappers ( $L = C = T = 0$ ), the hare population is modeled as a logistic growth with the intrinsic growth rate  $b$  and the intraspecific competition coefficient  $m$  (which can be justified by the field study of [20]). We will use Holling's Type II functional form ([14]) for the predation rate of the lynx on the hare with the encounter rate  $a_1$  and the handling time  $h_1$ . Parameter  $b_1$  is the consumption-to-birth ratio (biomass conversion coefficient) of the lynx and parameters  $d_1, m_1$  are its natural death rate and intraspecific competition rate respectively. Similar parameters notations,  $a_2, h_2, b_2, d_2, m_2$ , apply to the coyotes equation.

Like the predators, the trapping rates (the pelt harvest rates) of the hares and the lynx per unit index of trappers are the joint Holling Type II functional forms ([27, 21]),  $\frac{u_1 H}{1 + v_1 u_1 H + v_2 u_2 L}$ ,  $\frac{u_2 L}{1 + v_1 u_1 H + v_2 u_2 L}$ , with the encounter rates  $u_1, u_2$  and handling times  $v_1, v_2$  of the hares and lynx, respectively. However, unlike the predators, parameters  $r_1, r_2$  for the trappers equation are the intrinsic pelt-to-recruit ratios,  $d_3$  is trappers quit rate, and  $m_3$  is trappers intraspecific competition rate. That is, the system of equations models not only the ecological interactions of the hares, lynx, and coyotes, but also the economical interactions of the trappers with the natural system. The trappers equation is the same as those for the predators in form but the interpretations for its parameters are economical rather than ecological.

The continuous model Eq.(1) is referred to as the hare-lynx-coyote-trapper (HLCT) model. We will also consider the following models which it contains: the hare-lynx-trapper (HLT) model with  $C \equiv 0$  set for the HLCT equation; the hare-lynx-coyote (HLC) model with  $T \equiv 0$ ; and the HLCT1 model which has the same equation (1) except that the trapping rates are of Holling's Type I forms with the zero handling times  $v_1 = v_2 = 0$ . For comparison purpose we will also consider the vegetation-hare-lynx (VHL) model used in [1, 18, 37].

**2.2. Gradient Search and Line Search for Least Error.** Empirical data for a physical process  $P$  is a collection of time and real numbers, denoted in general by

$$(t_{ij}, y_{ij}), \quad i = 1, 2, \dots, k_j, \quad j = 1, 2, \dots, \ell$$

Here the second subindex  $j$  is for different type of data, say  $j = 1$  for the population of a prey and  $j = 2$  for the population of a predator. We will refer to it as the  $j$ th data type for a total of  $\ell$  many types. Each data type is collected at the same or distinct data acquisition times but we will assume without loss of generality that  $t_{ij}$  is increasing in  $i$  and the earliest collecting time is set to 0, i.e.  $t_{(i+1)j} > t_{ij} \geq 0$ .

In this paper we will only consider differential equations as mathematical models for the process,

$$\begin{cases} \frac{dx}{dt} = F(x, p) \\ x(0) = x_0 \end{cases} \quad (2)$$

where  $t$  has the same time dimensional unit as  $t_{ij}$ ,  $p$  the model parameters,  $x(t) = (x_1, x_2, \dots, x_n)(t)$  is the state variables of the model at time  $t$ . For each  $j$ , we consider a fit functional,  $f_j(t_{ij}, p, x_0)$ , to the  $j$ th data type  $(y_{1j}, y_{2j}, \dots, y_{k_jj})$ , and consider the weighted Euclidean error between the predicted and the observed:

$$E_{(F,f)}(p, x_0) = \sqrt{\sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 |f_j(t_{ij}, p, x_0) - y_{ij}|^2}$$

where the weight parameter  $w_{ij}$  has the reciprocal unit of  $y_{ij}$  to scale each term dimensionless. For example, we can use  $w_{ij} \equiv 1/\max_{1 \leq i \leq k_j} \{|y_{ij}|\}$  assuming not all  $y_{ij} = 0$  in  $i$ , or analogous to the  $\chi$ -square test we can use  $w_{ij} = 1/|y_{ij}|$  assuming all  $y_{ij} \neq 0$ . The usage of dimensional weights is essential when the best fit is sought for multiple data types for which the error,  $E_{(F,f)}(p, x_0)$ , has to be dimensional free for consistency. We also note that the state variable  $x(t)$  may or may not coincide in part or whole with the data type  $y$ . That is,  $f_j$  for any  $1 \leq j \leq \ell$  may or may not have the same dimensional unit as  $x_k$  for any  $1 \leq k \leq n$ .

By definition, the *best fit* of the model  $F$  to the data has the least error

$$\varepsilon(F, f) = \min_{(p, x_0)} E_{(F,f)}(p, x_0) = E_{(F,f)}(p^*, x_0^*)$$

at some  $(p^*, x_0^*)$ , referred to as the *global minimizer*, among all choices of the initial conditions  $x_0$  and parameter values  $p$ . Therefore, by definition, a model  $F$  is a benchmark model if

$$\varepsilon(F, f) \leq \varepsilon(G, g)$$

holds for all provisional models  $G$  (with the same fit weights  $w_{ij}$ ). A benchmark model is only temporary as it can be replaced by new and better provisional models. We note that it is often the case that we cannot prove a minimizer we found by a particular method is indeed the global minimizer but instead the best local minimizer with respect to the search method. Thus, the provision and benchmark designation to a model in this paper is contingent upon the search method we used for all models.

Finding local minima of the error function  $E(p, x_0)$  is the same as finding local minima of the error function squared  $E^2(p, x_0)$ . The search is done in the parameter and initial state space  $(p, x_0)$ , often along a fastest descending path. The methods we will use are all based on Newton's gradient search method. That is, we seek to determine a path in the parameter and initial state space,  $(p, x_0)(s)$ , so that it follows the negative gradient of  $E^2(p(s), x_0(s))$  in search of a local minimum of the squared error:

$$\begin{cases} \frac{\partial(p, x_0)}{\partial s} = -\nabla E^2(p, x_0) = -2 \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 [f_j(t_{ij}, p, x_0) - y_{ij}] D_{(p, x_0)}(f_j(t_{ij}, p, x_0)) \\ (p(0), x_0(0)) = (p_0, x_{0,0}) \end{cases}$$

where  $D_z f(z)$  denotes the derivative of function  $f$  with respect to its variable  $z$ , and  $(p_0, x_{0,0})$  denotes the initial search point. A local minimizer is found if the path converges

$$(p^*, x_0^*) = \lim_{s \rightarrow \infty} (p(s), x_0(s))$$

and a local minimum is declared numerically after a sufficiently large number  $s$ . We note that at each search point,  $x(t, x_0(s), p(s))$  is a solution to the model differential equations Eq.(2). Thus, as a function of  $(t, s)$ ,  $x(t, x_0(s), p(s))$  in fact is the solution to a partial differential equation induced from the gradient search for which more details are given in the Appendix.

It is known that if the squared error has non-unique local extrema, a gradient search may not yield the global minimizer. In fact, finding the global minimizer is still an active research in the area of scientific computations. Another drawback for the gradient search method is that it can be time consuming in solving the resulting PDEs. A practical approach to both speeding up the search and to finding a better minimizer, which we will also adopt, is the line search method. Without loss of generality, we assume all the parameters and the initial states are non-negative. The line search method we will use in this paper works as follows.

For every initial guess  $(p_0, x_{0,0})$ , we consider a hypercube centered at the initial guess with  $0 < p < 2p_0, 0 < x_0 < 2x_{0,0}$ , componentwise. We will then partition each interval into a fixed even number, say  $2N$ , of subintervals of equal length, with  $N$  discrete partitioning points to each side of the center. We will then search for a smaller error  $E(p, x_0)$  along this and each coordinate line through the center  $(p_0, x_{0,0})$  at these discrete points. For example, for the first parameter  $p_1$  there are  $2N + 1$  discrete partitioning points  $q_i$  with  $q_0 = 0, q_N = p_{1,0}$  the initial guess, and  $q_{2N} = 2p_{1,0}$  the end of the line search segment for the parameter  $p_1$ . With all other parameters and initial states fixed at the initial guess value  $(p_0, x_{0,0})$ , we compute  $E(p, x_0)$  with  $p_i = p_{i,0}, i \neq 1, x_0 = x_{0,0}$  but  $p_1 = q_k$  for all  $k = 0, 1, 2, \dots, 2N$ . This generates  $2N + 1$  many values for  $E(p, x_0)$ . Do the same for all other parameters and initial states to generate a total of  $(2N + 1) \times [\text{number of parameters and initial states}]$ . Of which we select the smallest value of  $E(p, x_0)$  and thus the next new initial guess  $(p_0, x_{0,0})$ . We repeat this process only to stop the search if either the successive errors are within certain stoppage tolerance or if it runs out a predetermined number of iterations. The output of this line search is our provisional global minimizer  $(p^*, x_0^*)$ .

One can also run a gradient search after the line search just to increase the accuracy of such provisional minimizer, which we did use. Notice that if we *know* the error function  $E(p, x_0)$  has all its local minima inside a bounded region, then both the gradient search and the line search must converge to a local minimizer. In fact, all searches carried out for this paper converged, and it is in this sense each best fitted model is the provisional model for the Canadian hare-lynx system.

**2.3. Best-Fit Sensitivity.** Suppose a provisional global minimizer  $(p^*, x_0^*)$  has been found for the error function  $E(p, x_0)$ , a next question is how sensitive does the error depends on changes in the parameters and initial states? This question can be formulated by the Taylor expansion of the error function. To be more specific, we first assume without the loss of generality that the minimizer occurs in interior of the parameters and initial states space  $(p, x_0) > 0$  componentwise. The justification is as follows.

If the minimizer occurs on a boundary with one of the parameters  $p_i = 0$ , then that parameter can be effectively dropped from the model and we can restrict the model only to those system parameters whose minimizer components are strictly greater than zero. Similarly if the minimizer occurs on a boundary with one of the initial states  $x_j = 0$ , then the state of the model will stay invariant with the  $x_j$ -population zero for all future time, and hence it can be dropped from the model to consider only an effectively smaller system of equations. Hence, sensitivity of the best fit is referred in this paper to only those effective parameters and initial states for which the minimizer occurs in their interiors of definition. That is, the first partial derivatives of the error function  $E$  at the minimizer are zeros.

We now define the sensitivity of the best fit. As an example, consider the case of the first parameter  $p_1$  and expand  $E$  at the minimizer

$$E(p, x_0) = E(p^*, x_0^*) + \frac{1}{2} \frac{\partial^2 E}{\partial p_1^2}(p^*, x_0^*) (p_1 - p_1^*)^2 + \dots,$$

where the dots represents the expanding terms for other parameters and initial states. Because  $p_1^* > 0$  we can rewrite it as follows making the squared change dimensionless

$$E(p, x_0) = E(p^*, x_0^*) + \frac{(p_1^*)^2}{2} \frac{\partial^2 E}{\partial p_1^2}(p^*, x_0^*) \frac{(p_1 - p_1^*)^2}{(p_1^*)^2} + \dots,$$

By definition, the coefficient of the squared percentage change  $\frac{(p_1 - p_1^*)^2}{(p_1^*)^2}$  is the *sensitivity* of the error with respect to the  $p_1$  parameter:

$$S(p_1) := \frac{(p_1^*)^2}{2} \frac{\partial^2 E}{\partial p_1^2}(p^*, x_0^*) \quad (3)$$

Similar definition applies to other parameters and initial states, denoted by  $S(p_i)$  and  $S(x_{j,0})$  respectively. Note that all sensitivities are greater than or equal to zero because  $E$  is an interior local minimum at the point  $(p^*, x_0^*)$ .

It is important to note that the sensitivity can be used to compare deviations of the error from the best fit with changes of all parameters and initial states. For example, for the same squared relative changes in parameter  $p_1$  and  $p_2$  with  $\frac{(p_1 - p_1^*)^2}{(p_1^*)^2} = \frac{(p_2 - p_2^*)^2}{(p_2^*)^2}$ , the inequality  $S(p_1) > S(p_2)$  implies that the error function  $E(p, x_0)$  is farther greater than its minimum  $E(p^*, x_0^*)$  along the  $p_1$  axis than along the  $p_2$  axis. In this sense we can say the best fit of the model to the data is more sensitive to the parameter  $p_1$  than to the parameter  $p_2$ . Similar pair-wise comparison applies to all parameters and initial states. We also note that the  $S$ -sensitivity can be easily approximated from the line search method when at least three discrete points are used for each of the search range  $[0, 2(p^*, x_0^*)]$  componentwise, enough for a discrete approximation of that component's second order partial derivative of the error function.

**2.4. Dimensional Analysis: Best-Fit Uncertainty and Sensitivity Certainty.** It is often the case that due to practical limitations, empirical data are collected in fewer independent dimensions than the dimensions of the physical system. The Canadian lynx-hare system is such an example for which the pelt data in lynx and hares are available but in reality the foodweb to which these two species are embedded has far more independently state variables from vegetation to competing herbivores and carnivores and

to trappers. Intuitively, there ought to be some degree of freedom allowed for the best-fitted parameter values of any provisional model. The questions are which parameters can be uniquely determined and which parameters cannot be, and for the latter what is the degree of uncertainty, and will such uncertainty effect the best-fit sensitivity  $S(p, x_0)$ ? These questions can be answered by the following theorem of dimensional analysis whose proof is a straightforward application of the Buckingham's  $\pi$  Theorem (e.g. [24]).

**Theorem:** Consider an  $\ell$ -dimensional data set

$$(t_{ij}, y_{ij}) \text{ for } i = 1, 2, \dots, k_j, j = 1, 2, \dots, \ell,$$

an  $n$ -dimensional differential equation model  $x' = F(x, p)$  with  $x \in \mathbb{R}^n$ ,  $p \in \mathbb{R}^m$ ,  $n \geq \ell$ , and scalar fit functionals  $f_j(t_{ij}, p, x_0)$ . Assume the differential equations and the errors of fit,  $f_j(t_{ij}, p, x_0) - y_{ij}$ , are unit-free. Then there exist scalings  $\tau(p)$ ,  $K_i(p)$ ,  $q_j = g_j(p)$ , and  $s_j(p)$  so that the model and the weighted squared error

$$E^2(p, x_0) = \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 (f_j(t_{ij}, p, x_0) - y_{ij})^2$$

is transformed to a dimensionless model  $\bar{x}' = G(\bar{x}, q)$  and its corresponding squared error becomes

$$E^2(q, \bar{x}_0, s, \tau) = \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 (s_j \bar{f}_j(t_{ij}\tau, \bar{x}_0, q) - y_{ij})^2$$

with  $\bar{f}_j$  being dimensionless,  $s_j$  having the same dimensional unit as  $y_{ij}$ ,  $\bar{t} = t\tau(p)$ ,  $\bar{x}(\bar{t}) = x(\bar{t}/\tau(p))/K_i(p)$ ,  $q = g(p) \in \mathbb{R}^{[m-n-1]_+}$ , where  $[m-n-1]_+$  is zero if  $m-n-1 \leq 0$  and  $m-n-1$  otherwise.

We note that all physical systems are unit-free, namely equivalent under dimensional unit conversions, and hence the theorem should applies to mechanistic conceptual models. The degree of freedom for the best fit is explained as follows. Notice that when  $m-n-1 \geq 0$ , a best fit by the dimensionless model to the data in the scaled  $(m-n-1) + n + \ell + 1 = m + \ell$  many quantities  $(q, \bar{x}_0, s, \tau)$  corresponds to an  $(n-\ell)$ -dimensional manifold of the same error value in the  $m+n$  dimensional parameter and initial condition space in  $(p, x_0)$ . That is,  $n-\ell$ , which is the difference between the dimensional dimension  $m+n$  and the scaled, dimensionless dimension  $m+\ell$ , is the degree of freedom for the best fit of the model to the dimensional data. In other words, if  $n > \ell$ , we must expect infinitely many choices in the dimensional parameters to give the same best error fit. For particular model, the question is to determine which parameters can be uniquely determined for the best fit and which parameters cannot but embody the degree of freedom for the fit.

We also note that for unit-free models the best fit sensitivity  $S(p, x_0)$  is independent of the best fit uncertainty. This can be easily proved by the same argument for the Buckingham's  $\pi$  Theorem. More specifically, the relationship between the dimensional and the dimensionless parameters and variables are algebraic, and the dependence of the uncertain parameters and initial states on the free parameters and initial states is also algebraic. As a result the free parameters and initial states are canceled out in

the sensitivity values with respect to all the uncertain parameters and initial states. That is, even though the global minimizers in the parameters and initial states are not unique, their sensitivities are.

**2.5. Chirality: HEL v.s. LEH Orientation.** When the time-dependent populations or pelts of hares and lynx are plotted in the  $HL$ -plane, the trajectory may proceed in a general counterclockwise direction, i.e. the lynx-eats-hare (LEH) orientation, or respectively in a general clockwise direction, i.e. the HEL orientation. Describing it differently the LEH chase is also right-handed, or *right chiral*, and the HEL chase is left-handed, or *left chiral*. Chirality is a quantity designed for the handedness of the orientation. In particular, a *positive* chirality is for a right-handed LEH chase and a *negative* chirality is for a left-handed HEL chase. Here is how the chirality of the hare-lynx trajectory is defined.

Let  $t_i, i = 0, 1, 2, \dots, n$  be an increasing sequence in time, and  $H_i, L_i$  be the population for the hares and lynx respectively at the time  $t_i$ . To define their chirality, let  $\vec{v}_i = (H_i - H_{i-1}, L_i - L_{i-1})$  be the direction or secant vector from point  $(H_{i-1}, L_{i-1})$  on the projected  $HL$ -plane to point  $(H_i, L_i)$ . Now with respect to the direction  $\vec{v}_i$  the next movement by the projected  $HL$ -trajectory takes place in the direction of  $\vec{v}_{i+1}$ , which can either right-handedly (counterclockwise) rotate up or left-handedly (clockwise) rotate down, or neither. The *chirality*,  $c_i$ , at the time  $t_i$ , is defined to be the coefficient of the curl vector from  $\vec{v}_i$  to  $\vec{v}_{i+1}$ . More specifically, if  $\vec{a} = (a_1, a_2)$  and  $\vec{b} = (b_1, b_2)$ , then the curl of  $\vec{a}$  to  $\vec{b}$  is  $\mathbf{curl}(\vec{a}, \vec{b}) := (a_1b_2 - a_2b_1)\mathbf{k}$  with  $\mathbf{k} = (0, 0, 1)$  the standard vector base for the  $z$ -axis, and the sign of the curl coefficient,  $a_1b_2 - a_2b_1 = \mathbf{curl}(\vec{a}, \vec{b}) \cdot \mathbf{k}$ , tells whether the orientation from  $\vec{a}$  to  $\vec{b}$  is right chiral ( $\mathbf{curl}(\vec{a}, \vec{b}) \cdot \mathbf{k} > 0$ ) or left chiral ( $\mathbf{curl}(\vec{a}, \vec{b}) \cdot \mathbf{k} < 0$ ). That is, we define the local or point chirality of the  $HL$ -trajectory at time  $t_i$  to be

$$c_i = \mathbf{curl}(\vec{v}_i, \vec{v}_{i+1}) \cdot \mathbf{k}, \text{ for } i = 1, 2, \dots, n.$$

The chirality for the trajectory is defined to be the time-averaged point chirality:

$$\bar{c}(H, L) = \frac{1}{t_n - t_0} \sum_{i=1}^n c_i \Delta t_i$$

with  $\Delta t_i = t_i - t_{i-1}$ . Note that this definition applies to sequences from numerical simulations as well as to the pelt data. It is in this sense that we say the  $HL$ -trajectory or data is right chiral if  $\bar{c}(H, L) > 0$  or left chiral if  $\bar{c}(H, L) < 0$  for the rest of the paper.

**3. Result.** We now apply the method outlined above to the HLCT conceptual model Eq.(1) and its various subsystems for comparison purposes. First we will use the lynx and hare pelt data used by [22, 11, 2, 42], and denote them by  $H_{T,i}, L_{T,i}$ , the trapped pelt for the hares and lynx respective in the  $i$ th year,  $t_i = i, i = 0, 1, 2, \dots, 31$ , since the year of 1875. Since they are tallied annually we can take them as the annual catch rates by the trappers. As a result we will use the trappers predation rate functionals for the fit functionals:

$$H_T(t_i) := f_H(t_i, x_0, p) = \frac{u_1 H(t_i, x_0, p) T(t_i, x_0, p)}{1 + v_1 u_1 H(t_i, x_0, p) + v_2 u_2 L(t_i, x_0, p)}$$

$$L_T(t_i) := f_L(t_i, x_0, p) = \frac{u_2 L(t_i, x_0, p) T(t_i, x_0, p)}{1 + v_1 u_1 H(t_i, x_0, p) + v_2 u_2 L(t_i, x_0, p)}$$

Table 2: Dimensional and Dimensionless Scalings

Dimensional			Best-Fit	Scaled		
Parameter	Best Fit	Scaling	Sensitivity	Parameter	Best Fit	Scaling
$b$	4.9399	$\tau$	238.59	$\tau$	4.9399	$b$
$m$	0.0467	$\tau/K_H$	173.97			
$a_1$	0.0965	$\tau/K_L$	174.34			
$h_1$	0.2942	$\eta_1 K_L/(\tau K_H)$	134.13	$\eta_1$	3	$h_1 a_1 K_H$
$a_2$	$4.9399/K_C$	$\tau/K_C$	66.3512			
$h_2$	$0.0057 K_C$	$\eta_2 K_C/(\tau K_H)$	127.23	$\eta_2$	3	$h_2 a_2 K_H$
$u_1$	$2.4684/K_T$	$\alpha_1 \tau/K_T$	5.14	$\alpha_1$	0.4997	$u_1 K_T/b$
$v_1$	$0.0019 K_T$	$\tau_1 K_T/s_1$	11.22	$\tau_1$	0.5	$v_1 u_1 K_H$
$u_2$	$13.6498/K_T$	$\beta_1 \tau/K_T$	31.29			
$v_2$	$0.0157 K_T$	$\tau_2 K_T/s_2$	171.84	$\tau_2$	11	$v_2 u_2 K_L$
$b_1$	1.3386	$\beta_1 K_L/K_H$	175.46	$\beta_1$	2.7632	$b_1 a_1 K_H/b$
$d_1$	1.9592	$\delta_1 \tau \beta_1$	128.50	$\delta_1$	0.1435	$d_1/(b_1 a_1 K_H)$
$m_1$	0.0037	$\mu_1 \tau \beta_1/K_L$	1.57	$\mu_1$	0.0139	$m_1 K_L/(b_1 a_1 K_H)$
$b_2$	$0.005 K_C$	$\beta_2 K_C/K_H$	104.72	$\beta_2$	0.5263	$b_2 a_2 K_H/b$
$d_2$	0.0269	$\delta_2 \tau \beta_2$	1.64	$\delta_2$	0.0103	$d_2/(b_2 a_2 K_H)$
$m_2$	$0.5207/K_C$	$\mu_2 \tau \beta_2/K_C$	82.46	$\mu_2$	0.2003	$m_2 K_C/(b_2 a_2 K_H)$
$r_1$	$0.0103 K_T$	$\gamma_1 \rho \tau K_T/s_1$	22.25	$\gamma_1$	0.0687	$r_1 u_1 K_H/(r_2 u_2 K_L)$
$r_2$	$0.0558 K_T$	$\rho \tau K_T/s_2$	170.33	$\rho$	7.8948	$r_2 u_2 K_L/b$
$d_3$	2.6982	$\delta_3 \rho \tau$	134.63	$\delta_3$	0.0692	$d_3/(r_2 u_2 K_L)$
$m_3$	$0.0511/K_T$	$\mu_3 \rho \tau/K_T$	1.02	$\mu_3$	0.0013	$m_3 K_T/(r_2 u_2 K_L)$
				$s_1$	260.8767	$u_1 K_H K_T$
				$s_2$	698.8768	$u_2 K_L K_T$
$K_H$	105.6875	$s_1/(\alpha_1 \tau)$		$K_H$	105.6875	$b/m$
$K_L$	51.2005	$s_2/(\tau \beta_1)$		$K_L$	51.2005	$b/a_1$
$K_C$	1	free		$K_C$	1	$b/a_2$
$K_T$	1	free		$K_T$	1	$b_1 a_1 K_H/u_2$
Dimensional Initials						
$H_0$	52.0003		7.41			
$L_0$	9.7764		30.56			
$C_0$	$0.9051 K_C$		57.69			
$T_0$	$0.2747 K_T$		94.95			

and the corresponding the squared error:

$$E^2(p, x_0) = \sum_{i=0}^{31} \left[ \left( \frac{f_H(t_i, x_0, p) - H_{T,i}}{H_T^*} \right)^2 + \left( \frac{f_L(t_i, x_0, p) - L_{T,i}}{L_T^*} \right)^2 \right]$$

for which  $t_{i1} = t_{i2} = t_i = i$ ,  $w_{i1} = 1/H_T^*$ ,  $w_{i2} = 1/L_T^*$ ,  $i = 0, 1, 2, \dots, 31$  and

$$x_0 = (H_0, L_0, C_0, T_0), \quad H_T^* = \max\{H_{T,i}\}, \quad L_T^* = \max\{L_{T,i}\}.$$

By the dimensional analysis theorem above we know for the HLCT model it has a degree-2 uncertainty for the best fit. To determine those uncertain parameters and initial states we transform the dimensional model Eq.(1) into a dimensionless model with a change of parameters and states. More specifically, the transformation and the inverse transformation are given in Table 2. For example, the entries from Scaled Parameters column are defined by the last Scaling column which defines transformation from the dimensional parameters to the dimensionless ones, such as  $\eta_1 = h_1 a_1 K_H$  with  $K_H = b/m$ . Similarly, the third column defines the inverse transformation from the dimensionless parameters to the dimensional ones in the first column, such as  $h_1 = \eta_1 K_L / (\tau K_H)$  with  $K_H = s_1 / (\alpha_1 \tau)$ ,  $K_L = s_2 / (\tau \beta_1)$ . The dimensionless variables and dimensionless time are  $\bar{H} = H/K_H$ ,  $\bar{L} = L/K_L$ ,  $\bar{C} = C/K_C$ ,  $\bar{T} = T/K_T$ ,  $\bar{t} = t\tau$ . To simplify notations we drop the bars for the new variables and time and obtain the following dimensionless model for Eq.(1):

$$\begin{cases} \frac{dH}{dt} = H \left( 1 - H - \frac{L}{1 + \eta_1 H} - \frac{C}{1 + \eta_2 H} - \frac{\alpha_1 T}{1 + \tau_1 H + \tau_2 L} \right) \\ \frac{dL}{dt} = \beta_1 L \left( \frac{H}{1 + \eta_1 H} - \delta_1 - \mu_1 L - \frac{T}{1 + \tau_1 H + \tau_2 L} \right) \\ \frac{dC}{dt} = \beta_2 C \left( \frac{H}{1 + \eta_2 H} - \delta_2 - \mu_2 C \right) \\ \frac{dT}{dt} = \rho T \left( \frac{\gamma_1 H + L}{1 + \tau_1 H + \tau_2 L} - \delta_3 - \mu_3 T \right) \end{cases} \quad (4)$$

together with the fit functionals and the squared error

$$\begin{aligned} \bar{f}_H(\bar{t}_i, \bar{x}_0, q) &= \frac{H(\bar{t}_i, \bar{x}_0, q)T(\bar{t}_i, \bar{x}_0, q)}{1 + \tau_1 H(\bar{t}_i, \bar{x}_0, q) + \tau_2 L(\bar{t}_i, \bar{x}_0, q)} \\ \bar{f}_L(\bar{t}_i, \bar{x}_0, q) &= \frac{L(\bar{t}_i, \bar{x}_0, q)T(\bar{t}_i, \bar{x}_0, q)}{1 + \tau_1 H(\bar{t}_i, \bar{x}_0, q) + \tau_2 L(\bar{t}_i, \bar{x}_0, q)} \\ E^2(q, \bar{x}_0, s, \tau) &= \sum_{i=0}^{31} \left[ \left( \frac{s_1 \bar{f}_H(t_i \tau, \bar{x}_0, q) - H_{T,i}}{H_T^*} \right)^2 + \left( \frac{s_2 \bar{f}_L(t_i \tau, \bar{x}_0, q) - L_{T,i}}{L_T^*} \right)^2 \right] \\ \bar{x}_0 &= (H_0/K_H, L_0/K_L, C_0/K_C, T_0/K_T), \quad \bar{t}_i = t_i \tau. \end{aligned}$$

Notice here that  $t_i = i$  and  $\bar{t}_i = i\tau$  retain their dimensional and dimensionless identities for being a fixed sequence each rather than a variable.

A combined PDE search and line search for the dimensionless model yielded a provisional global minimizer for the dimensionless parameters listed in the second last column of Table 2 from  $\tau$  down to  $s_2$ . Translating it to the dimensional variables and state scalings we obtain the parameterized values in the second column of Table 2. As predicted by the theorem, four parameters ( $m, a_1, a_2, u_2$ ) are scaled away but two more,  $s_1, s_2$ , are created by the transformation as can be seen in the Scaled Parameter column, which in turn creates two free, parameterizing, auxiliary parameters which we take them to be  $K_C, K_T$ , the ‘carrying capacities’ for the coyotes and the trappers respectively. Notice that their units remain to be free — they can be head-count, biomass, or for the case of  $K_T$ , a pure index for the trapping business from the perspective of Hudson’s Bay Company. For the provisional global minimizer in the space of the dimensional parameters and initials, we see clearly from the second and the third columns that some parameters are uniquely determined by the global minimization of the squared error but some are not, namely the uncertain parameters and initial states. That is, different choices for the parameterizing pair,  $K_C, K_T$ , will give rise to different values for those uncertain parameters but to the same minimum error value  $E(p^*, x_0^*) = 1.1677$ .

The minimizer in the parameters and initial states was found for the dimensionless model first, and then translated for dimensional model. The dimensional minimizer was then checked and re-searched independently by both methods for the dimensional model, only after which were the sensitivities calculated and listed in the Best-Fit Sensitivity column.

Figs.1,2,3 highlight part of the numerical result. Fig.1 clearly shows the hare and lynx populations left in the wild is right chiral and the respective catches by the trappers is left chiral. Fig.1(d) also shows a typical gradient or line search in action. Fig.2 shows part of the result of the line search. Twenty subinterval partitions are used for each of the line segment,  $(0, 2p_0)$  and  $(0, 2x_0)$ , and as a result the provisional minimizer sits at the center of each segment with ten searching points to each side. Because the same proportionality is used for each interval length and the same window size is used for all plots of the error function against the search intervals regardless the values of the components of the minimizer, these plots give a graphical depiction to the dimensionless sensitivity and a graphical comparison of the sensitivities among all parameters and initials. For example, between the recruitment parameters  $r_1, r_2$  of the hares and lynx, respectively, the best fit is less sensitive to hare pelt than to lynx pelt because  $S(r_1) < S(r_2)$ . This is represented by the top two graphs of Fig.2 for which the concavity is more pronounced for the  $r_2$  parameter than to the  $r_1$  parameter. Similar comparisons can be done for all parameters and initial states, and are captured by the sensitivity scores from Table 2. Notice also for each of the plots the trajectory chiralities are plotted.

One general conclusion is that the best fit is more sensitive to all trapping parameters and initials in the lynx than the hare. This may not be that surprising because it is consistent with the fact that the lynx fur was economically more valuable than the hare fur to the trappers and the Hudson’s Bay Company. However, what is surprising is the finding from the last plot of Fig.2 that the intra-competition parameter  $m_3$  for the trappers can be set to zero with little effect on the best fit. That is, there was little interference among the trappers. Although the sensitivities in parameters  $m_1, d_2$  are comparable to that in  $m_3$ , none of the parameterized error function plots remains as flat as that with  $m_3$  all way to the parameter zero.

That is, for the HLCT model,  $m_3$  is the only parameter that can be effectively dropped.

Fig.3 summarizes the result of comparing various provisional models. Among which are the VHL model from [1], the HLC model without the trappers, the HLT model without the coyotes, and the HLCT1 model with the Holling Type I predation form for the trappers. The last four models are simply subsystems of the HLCT model Eq.(1) setting  $T_0 = 0, C_0 = 0$ , and  $v_1 = v_2 = 0$ , respectively. For the first two models without the trappers, the pelt quantities are set to be proportional to the populations as  $H_T = s_1 H, L_T = s_2 L$ , respectively, with  $s_i$  being the proportion parameters. As shown in the bar plot of Fig.3, the two models without the trappers do not share the same chirality as the pelt data but all models with the trappers do. Also notice that having more parameters or bigger systems does not necessarily yield a smaller least error. A comparison of the least errors among the HLT and HLCT1 models against the HLC model demonstrates this point. Notice also from the figure that the best-fit exercise by the gradient and line searches places the  $H_T L_T$ -trajectory in the middle of the pelt data for the VHL, HLT, and HLCT1 models, but cannot at the same time make the trajectory oscillate left chiral. For the VHL model without the trappers, it cannot because the error will only get greater when its catch trajectory oscillation in the opposite direction of the pelt trajectory. (The VHL model can generate chaotic oscillations as shown in [1, 37], but they are far away from the pelt oscillation and in the wrong direction.) For the HLT and HLCT1 models, although each can generate the co-chiral oscillations as the pelt but not at the right location in the  $HL$ -plane. The important conclusion is the HLCT model becomes the benchmark model for the pelt data set.

**4. Discussion.** Model parameters for a physical system can be estimated by many ways. One of which is by the best-fit process presented here. By this approach the parameter values are forced to give the smallest error between the predicated and the observed data. For an ideal situation in which both the theoretical model and the data are good representations of the process, the best-fitted parameter values achieve some intrinsic status in the sense that they can be used elsewhere independent of the observed data from which they are determined. For example, if our HLCT model and the data are both good for the hare-lynx system, then we can use, say, lynx's birth-to-consumption ratio  $b_1$  for the hare-lynx interaction in all theoretical models with or without the trappers. However, upon a closer examination of the value,  $b_1 = 1.3386$ , we know it is a bad estimate. It cannot be right whichever ways you look at it. If both populations are measured in the same unit of an element, say carbon, then  $b_1$  must be  $\leq 1$ , because of the conservation of mass transferred across trophic levels. If both are measured in their head count which is the case, it just does not make sense that one hare kill can give birth to 1.3386 lynx. So is the data or the model bad?

The answer is probably both. Our dimensional analysis implies the data is bad and points to a possible solution. In fact, in terms of the dimensionless parameters, we have  $b_1 = \beta_1 K_L / K_H$  from Table 2. The value of  $b_1$  becomes smaller if the lynx-to-hare capacity ratio  $K_L / K_H$  becomes smaller. Yet the ratio  $K_L / K_H = 51.2005 / 105.6875 \approx 1:2$  is too high. It is a rule of thumb that as the biomass across trophic boundary from resources to consumers, the biomass difference is of several orders in magnitude (e.g. [6]). In particular, the lynx to hare population in the wild should at least differ by a magnitude of three orders (see Fig.5 of [28]). By the definitions of the capacity parameters we see

that the ratio is directly proportional to the scaling parameter ratio  $s_2/s_1$ , which is directly related to the magnitudes of the pelt data for the lynx and hares. Thus, the ‘bad’ fit can be the result of an underestimate or an unreported or an unaccounted return in hare’s pelt. In fact, according to [40] hares were trapped for their meat as well as for their fur and therefore it was highly probable that the hare pelt brought in to the Hudson’s Bay Company was significantly fewer than the hares actually trapped. Also, as pointed out in [34, 35] the hare data came from one of the fourteen regions from which the lynx pelt were combined. Although the aggregation does not cause appreciable phase shift since those regional data were synchronized in time ([1]), the aggregated data do flip the amplitude at least one order of magnitude in lynx’s favor. If this indeed is the case, the ‘badly’ fitted parameter  $b_1$  can be fixed to fall inside the plausible range  $b_1 < 1$  by simply increasing the  $s_1$  parameter (i.e. scaling up the hare data) alone. There are other unreasonable fits (e.g.  $h_1$ ) which share the same problem and therefore can be fixed by the same solution as well. In fact, by multiplying a scaling factor to the hare pelt  $\{H_{T,i}\}$  we will have exactly the same best fit by multiplying the same factor to the parameter  $s_1$  without changing the sensitivities for the best fit nor any parameter values not related to  $s_1$ . For example, Fig.1 will look exactly the same except for the scale for the hare variable which is scaled accordingly by the same factor. In other words we can drive the  $K_L:K_H$  ratio to any value by simply scaling the hare pelt data without changing the essentials of the best fit.

The other equal culprit for the high  $b_1$  value lies in the model. It is known that the lynx is a generalist preying on other animals beside the hares ([28]), but instead our model modeled it as a specialist. The benchmarking process forced a best fit to the data by unduely crediting the hares for other prey’s contribution to lynx’s good fortune. That is, the value  $b_1$  is expected to drop if alternative food sources for the lynx are incorporated to the HLCT model. In conclusion it will take a combination of matching up the hare and lynx pelt region-to-region as well as including alternative preys for the lynx to the model to improve the estimate for parameter  $b_1$  and to improve the benchmark model as a result.

There were considerable interests in the question of whether or not the Canadian hare-lynx system is chaotic. A three-dimensional time-delayed embedding ([33]) suggested the pelt time-series sit on a strange attractor but a quantitative analysis on the Lyapunov exponents of the pelt series concluded otherwise ([7]). Nonetheless, chaos was actively pursued for the conceptual VHL model in [1, 10, 37]. Our benchmark model supports the non-chaos conclusion of [7], but also suggests the possibility of chaos at parameter values away from the best-fit minimizer as shown in Fig.4. The question of how close is the pelt series to the onset of chaos can be the topic of another future exploration.

We deliberately left out discrete models (e.g. [34]) for benchmarking but instead considered exclusively differential equation models. This is due to the fact almost all discrete models in ecology violate the Time Invariance Principle ([5, 6]) without which a model cannot be independently validated by experiments. Our result also supports the One-Life Rule postulation ([5, 6]) that every organism has a finite life span which is guaranteed by models with carrying capacities for all their species. This in turn is guaranteed by non-vanishing intraspecific competition parameters  $m, m_1, m_2 > 0$ . Curiously this rule does not necessarily apply to the economical-ecological interaction between the trappers and their habitat because our trapper equation does not model the birth and death of the trappers as a species but

rather the rise and fall of their trapping business. Our result shows that interferences among trappers can be absent ( $m_3 = 0$ ) and be insensitive to variations. Our result also supports another fundamental theory in ecology — Holling’s theory of predation. Our best-fit shows that all the predation handling times ( $h_i, v_j > 0$ ) must be non-zero and be sensitive to changes. These results suggest that ecological modeling must move beyond the Rosenzweig-MacArthur producer-consumer model ([31]) as well as from the Lotka-Volterra model for competitive species.

The benchmarking method presented here did exactly what was intended — to force the best fit to a data to expose new problems for further investigations. It produces ‘bad’ fits for good reasons, a property of a good and useful method. All mathematical models are constructed to fail against the test of time, but with this method the cycle of falsification can run faster. This method literally rolled out answers by mathematical inference alone before we could even imagine to ask the right questions, such as were the lynx furs more valuable? or was the hare pelt under represented? or was the lynx a generalist? or did the trappers interfere each others business? We expect the method to be equally effective for other problems (e.g. [36, 13, 4]) in ecology and in biology in general.

We have given a numerical demonstration for the phenomenon that the kill rates by a top predator on a predator-prey chain must rotate in opposite direction against the populations of the predator and prey. We believe this anti-chirality property can be proved mathematically for such three-trophic food chains, which we will leave it open nonetheless. In conclusion, what Leigh found so ‘mysterious’ about the Canadian hare-lynx oscillation some forty plus years ago can be understood systematically by benchmarking modern ecological-economical models we presented here.

## References

- [1] Blasius, B., A. Huppert, & L. Stone, *Complex dynamics and phase synchronization in spatially extended ecological systems*, Nature, **399**(1999), pp.354–359.
- [2] Britton, N.F., *Essential Mathematical Biology*, Springer-Verlag, London, 2003.
- [3] Bulmer, M.G., *A statistical analysis of the 10-year cycle in Canada*, J. Anim. Ecol., **43**(1974), pp.701–718.
- [4] Cattadori, I.M., D.T. Haydon, S.J. Thirgood, & P.J. Hudson, *Are indirect measures of abundance a useful index of population density? The case of red grouse harvesting records*. Oikos, **100**(2002), pp.439–446.
- [5] Deng, B., *The Time Invariance Principle, the absence of ecological chaos, and a fundamental pitfall of discrete modeling*, Ecological Modeling, **215**(2008), pp.287–292.
- [6] Deng, B., *From energy gradient and natural selection to biodiversity and stability of ecosystems*, The Open Ecology Journal, **3**(2010), pp.95–110, DOI: 10.2174/1874213001003010095
- [7] Ellner, S. & P. Turchin, *Chaos in a noisy world: new methods and evidence from time-series analysis*, Am. Nat., **145**(1995), pp.343–375.

- [8] Elton, C. & M. Nicholson, *The ten-year cycle in numbers of the lynx in Canada*, J. Anim. Ecol. **11**(1942), pp.215-244.
- [9] Finerty, J.P., *Cycles in Canadian lynx*, Am. Nat., **114**(1979), pp.453-455.
- [10] Gamarra, J.G.P. & R.V. Solé, *Bifurcations and chaos in ecology: lynx returns revisited*, Ecology Letter, **3**(2000), pp.114–121.
- [11] Gilpin, M.E., *Do hares eat lynx?* Amer. Nat., **107**(1973), pp.727–730.
- [12] Hall, C.A.S, *An assessment of several of the historically most influential theoretical-models used in ecology and of the data provided in their support*, Ecological Modelling, **43**(1988), pp.5–31.
- [13] Hanski, I., H. Henttonen, E. Korpimäki, L. Oksanen, & P. Turchin, *Small-rodent dynamics and predation*, Ecology, **82**(2001), pp.1505-1520.
- [14] Holling, C.S., *Some characteristics of simple types of predation and parasitism*, Canadian Entomologist, **91**(1959), pp.385–398.
- [15] Huffaker, C.B., *Experimental studies on predation: dispersion factors and predator-prey oscillations*, Hilgardia, **27**(1958), pp.795–834.
- [16] Hutchinson, G.E., *Variations on a theme by Robert MacArthur*, pp.492–521 in M.L. Cody and J.M. Diamond, eds., *Ecology and the evolution of communities*. Belknap, Cambridge, 1975.
- [17] Keith, L.B., *Wildlife's ten-year cycle*. Univ. Wisconsin Press, Madison 1963.
- [18] King, A.A. & W.M. Schaffer, *The geometry of a population cycle: a mechanistic model of snowshoe hare demography*, Ecology, **82**(2001), pp.814–830.
- [19] Krebs, C.J., R. Boonstra, S. Boutin, and A.R.E. Sinclair, *What drives the 10-year cycle of snowshoe hares?* BioScience, **51**(2001), pp.25–35.
- [20] Krebs, C.J., S. Boutin, R. Boonstra, A.R.E. Sinclair, N.N.M. Smith, M.R.T. Dale, K. Martin, & R. Turkington, *Impact of food and predation on the snowshoe hare cycle*, Science, **269**(1995), pp.112–115.
- [21] Lawton, J.H., J.R. Beddington, & R. Bonser, *Switching in invertebrate predators*, pp.141–158 in *Ecological Stability* M.B. Usher and M.H. Williamson (eds.), Chapman and Hall, London, 1974.
- [22] Leigh, E.R., *The ecological role of Volterra's equations*, pp.1–61 in M. Gerstenhaber, ed., *Some mathematical problems in biology*. American Mathematical Society, Providence, R.I. 1968.
- [23] Lotka, A.J., *Elements of Physical Biology*, Williams and Wilkins, Baltimore, Md., 1925.
- [24] Logan, D., *Applied Mathematics*, 2nd Ed., Wiley-Interscience, 1996.
- [25] MacLulich, D.A., *Fluctuation in the number of the varying hare*, Univ. of Toronto Press, Toronto, Ontario, Canada, 1937.
- [26] May, R.M., *Stability and Complexity in Model Ecosystem*, Princeton University Press, Princeton, 1973.

- [27] Murdoch, W.W., *The functional response of predators*, J. Appl. Eco., **10**(1973), pp.335–342.
- [28] O’Donoghue, M., S. Boutin, C.J. Krebs, G. Zuleta, D.L. Murray, & E.J. Hofer, *Functional responses of coyotes and lynx to the snowshoe hare cycle*, Ecology, **79**(1998), pp.1193–1208.
- [29] Peressini, A., F. Sullivan, & J. Uhl, *The Mathematics of Nonlinear Programming*, Springer, 1993.
- [30] Rohner, C., *The numerical response of great horned owls to the snowshoe hare cycle: consequences of non-territorial ‘floaters’ on demography*, J. Ani. Ecol., **65**(1996), pp.359–370.
- [31] Rosenzweig, M.L. & R.H. MacArthur, *Graphical representation and stability conditions of predator-prey interactions*, Am. Nat., **97**(1963), pp.209–223.
- [32] Ruszczyński, A.P., *Nonlinear Optimization*, Princeton University Press, Princeton, N.J., 2006.
- [33] Schaffer, W.M., *Stretching and folding in lynx fur returns - evidence for a strange attractor*, Amer. Nat., **124**(1984), pp.798–820.
- [34] Stenseth, N.C., W. Falck, O.N. Bjørnstad, & C.J. Krebs, *Population regulation in snowshoe hare and Canadian lynx: Asymmetric food web configurations between hare and lynx*, Proc. Natl. Acad. Sci., **94**(1997), pp.5147–5152.
- [35] Stenseth, N.C., W. Falck, K.-S. Chan, O.N. Bjørnstad, M. O’Donoghue, H. Tong, R. Boonstra, S. Boutin, C.J. Krebs, & N.G. Yoccoz, *From patterns to processes: Phase and density dependencies in the Canadian lynx cycle*, Proc. Natl. Acad. Sci., **95**(1998), pp.15430–15435.
- [36] Stenseth, N.C., K.-S. Chan, E. Framstad, & H. Tong, *Phase- and density-dependent population dynamics in norwegian lemmings: interaction between deterministic and stochastic processes*, Proc. R. Soc. Lond. B., **265**(1998), pp.1957–1968.
- [37] Stone, L. & D. He, *Chaotic oscillations and cycles in multi-trophic ecological systems*, J. Theo. Bio., **248**(2007), pp.382–390.
- [38] Vik, J.O., C.N. Brinch, S. Boutin, & N. Stenseth, *Interlinking hare and lynx dynamics using a century’s worth of annual data*, Popul. Ecol., **50**(2008), pp.267–274.
- [39] Volterra, V., *Fluctuations in the abundance of species, considered mathematically*, Nature, **118**(1926), pp.558–560.
- [40] Weinstein, M.S., *Hares, lynx, and trappers*, Amer. Nat., **111**(1977), pp.806–808.
- [41] Winterhalder, B.P., *Canadian fur bearer cycles and Cree-Ojibwa hunting and trapping practices*, Amer. Nat., **115**(1980), pp.870–879.
- [42] Zhang, Z.B., Y. Tao, & Z.Q. Li, *Factors affecting harelynx dynamics in the classic time series of the Hudson Bay Company, Canada*, Clim. Res. **34**(2007), pp.83-89.

**Appendix.** Gradient Search for local minimum of the squared error  $E^2(p, x_0)$  in the parameters and initial states space  $(p, x_0)$

$$\begin{cases} \frac{\partial(p, x_0)}{\partial s} = -\nabla E^2(p, x_0) = -2 \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 [f_j(t_{ij}, p, x_0) - y_{ij}] D_{(p, x_0)}(f_j(t_{ij}, p, x_0)) \\ (p(0), x_0(0)) = (p_0, x_{0,0}) \end{cases}$$

Formally, we let

$$\begin{cases} x(t, s) = x(t, x_0(s), p(s)) \\ x(0, s) = x_0(s) \\ X(t, s) = D_{(p, x_0)}x(t, x_0(s), p(s)) = (X_1(t, s), X_2(t, s)) \\ X(0, s) = (0, I) \\ p(s) = p(s, x_{0,0}, p_0) \\ x_0(s) = x_0(s, x_{0,0}, p_0) \end{cases}$$

Then they satisfy the following system of PDEs

$$\begin{cases} \frac{\partial x(t, s)}{\partial t} = F(x(t, s), p(s)) \\ \frac{\partial X(t, s)}{\partial t} = [D_x F(x(t, s), p(s))X_1(t, s) + D_p F(x(t, s), p(s)), D_x F(x(t, s), p(s))X_2(t, s)] \\ x(0, s) = x_0(s), \quad X(0, s) = (0, I) \\ \frac{\partial p}{\partial s} = -\nabla_p E^2(p, x_0) = -2 \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 [f_j(t_{ij}, p, x_0) - y_{ij}] D_p(f_j(t_{ij}, p, x_0)) \\ \frac{\partial x_0}{\partial s} = -\nabla_{x_0} E^2(p, x_0) = -2 \sum_{j=1}^{\ell} \sum_{i=1}^{k_j} w_{ij}^2 [f_j(t_{ij}, p, x_0) - y_{ij}] D_{x_0}(f_j(t_{ij}, p, x_0)) \\ (p(0), x_0(0)) = (p_0, x_{0,0}) \end{cases}$$

The reason we need the variational equations in  $X$  for  $x$  is because a further expansion of  $D_{(p, x_0)}f_j$  usually involves the variations of  $x$  in both the parameters and the initial states. One can show or verify directly that for the standard linear regression,  $x(t_i) = a t_i + b \sim y_i$  with  $p = a$ ,  $x_0 = b$ , we have the corresponding gradient search PDEs as follows:

$$\begin{cases} \frac{\partial x(t, s)}{\partial t} = a(s), \quad \frac{\partial X(t, s)}{\partial t} = (1, 0) \\ x(0, s) = b(s), \quad X(0, s) = (0, 1) \\ \frac{\partial(a, b)(s)}{\partial s} = -2 \sum_{i=1}^k [a(s)t_i + b(s) - y_i][t_i, 1] \\ (a(0), b(0)) = (a_0, b_0) \end{cases}$$

It is a linear system of differential equations and it can be easily shown the global minimum exists, which can be solved explicitly as

$$a^* = \frac{k \sum t_i y_i - \sum t_i \sum y_i}{k \sum t_i^2 - (\sum t_i)^2}, \quad b^* = \frac{\sum t_i^2 \sum y_i - \sum t_i \sum t_i y_i}{k \sum t_i^2 - (\sum t_i)^2}$$

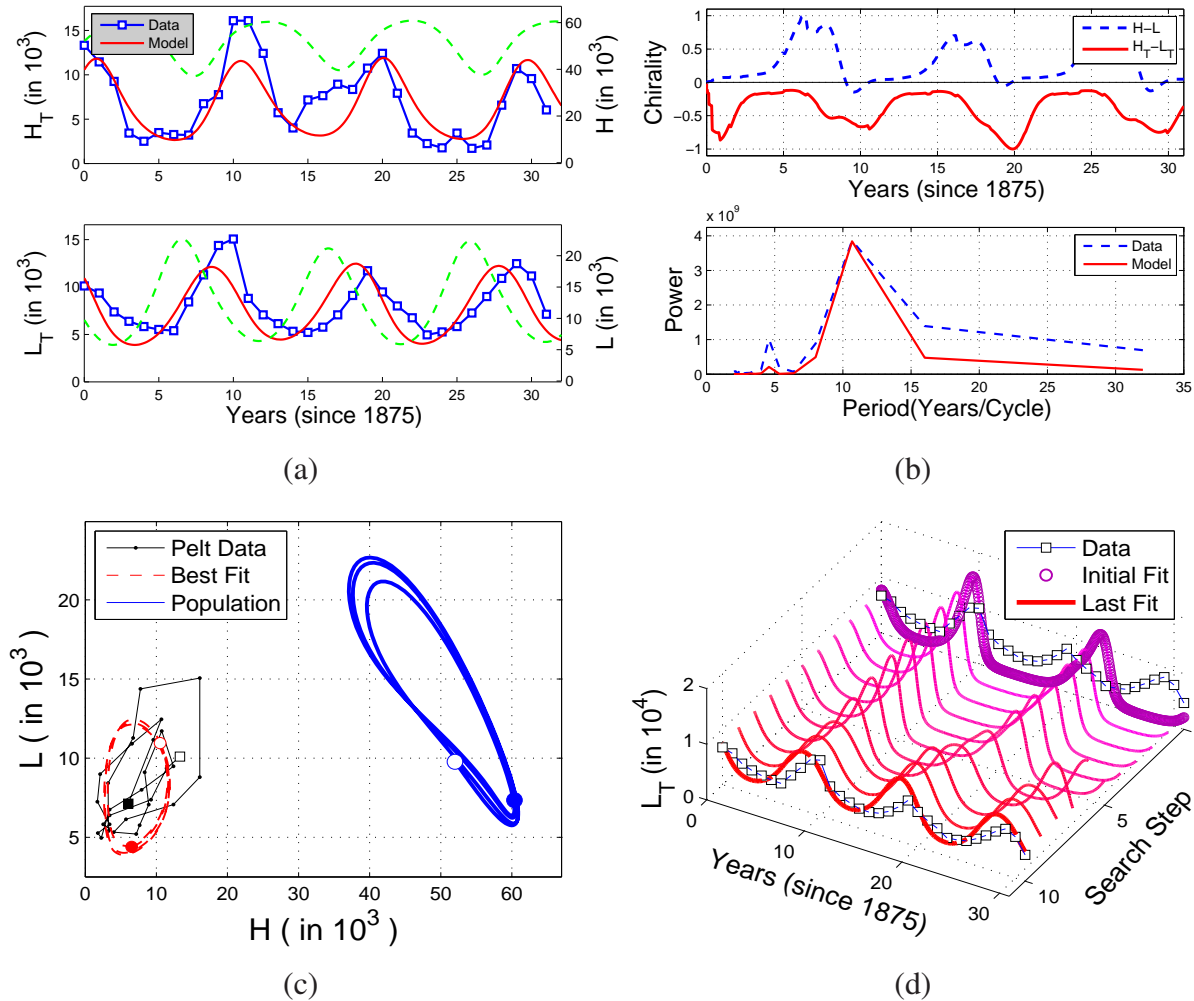


Figure 1: (a) The best fit of the HLCT model in  $H_T, L_T$  to the pelt data, together with the hare and lynx populations,  $H, L$ , left in the wild (dashed curves). (b) Top panel: it shows the point-wise right chirality of the hare-lynx population in the wild and the left chirality of the hare-lynx pelt. Bottom panel: the period-power plots of the pelt data and the best fit approximation. It shows both match exactly at the principle period mode around a 11-year cycle. (c) A hare-lynx phase plot, which shows a 5:1 peak ratio between the wild population and the pelt for hares and a weak 2:1 peak ratio for the lynx, both are unreasonably low as discussed in the text. (d) As part of the outcome of the PDE gradient search method, it shows the trapped lynx  $L_T$  as a function of the search variable  $s$  showing in 10 discrete steps and the time  $t$ . It suggests that a local minimum of  $E(p, x_0)$  was reached as the 7th, 8th, and 9th search step yield essentially the same time profile.

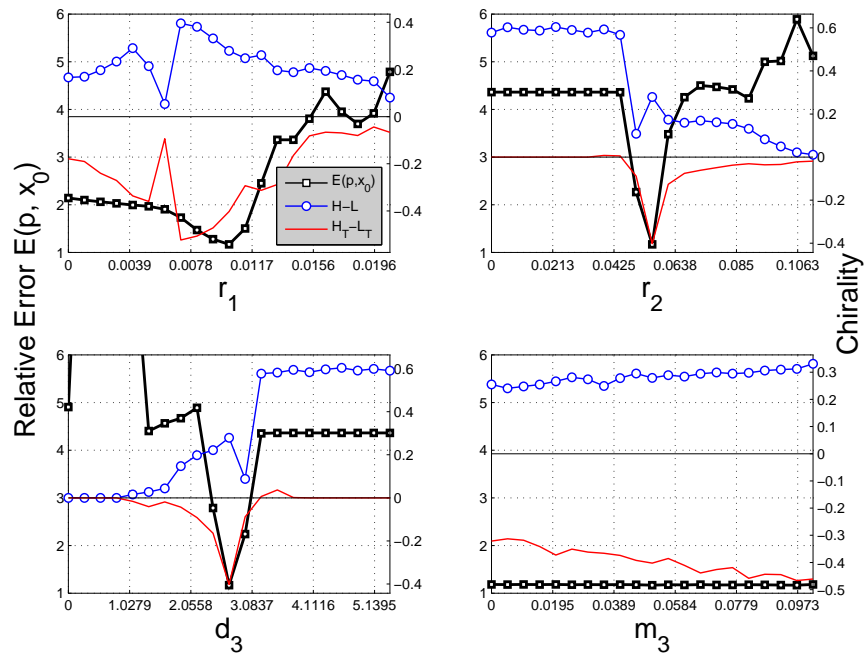


Figure 2: Part of the line search result for the parameters shown. As explained in the text when a solution is found by the algorithm, the relative error function  $E(p, x_0)$  against every parameter and initial condition must have the best fit minimum at the center of the searching intervals.

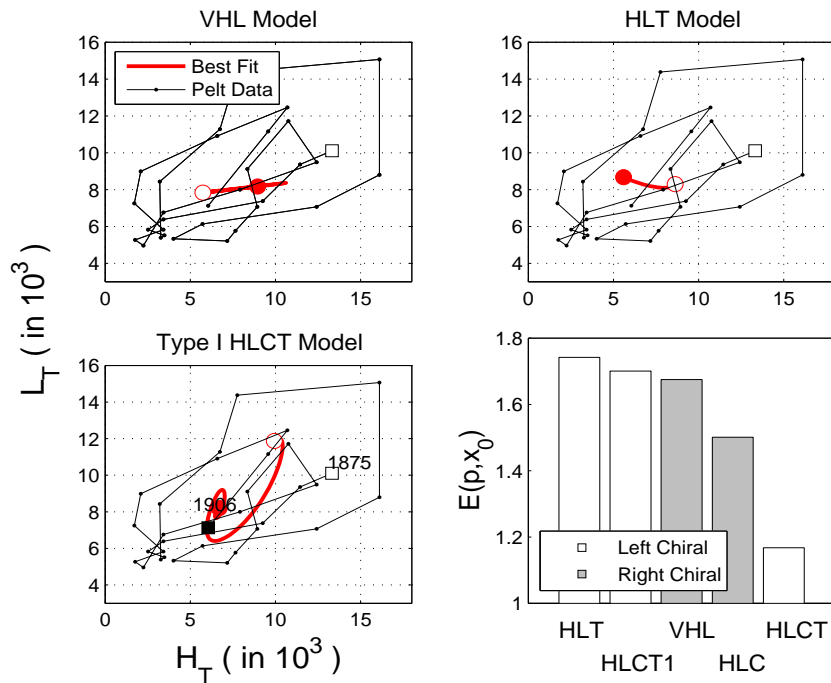


Figure 3: The best-fit results by the method of line search for three different models and a summary bar plot showing the HLCT model has not only the smallest error against other provisional models but also the correct left chirality.

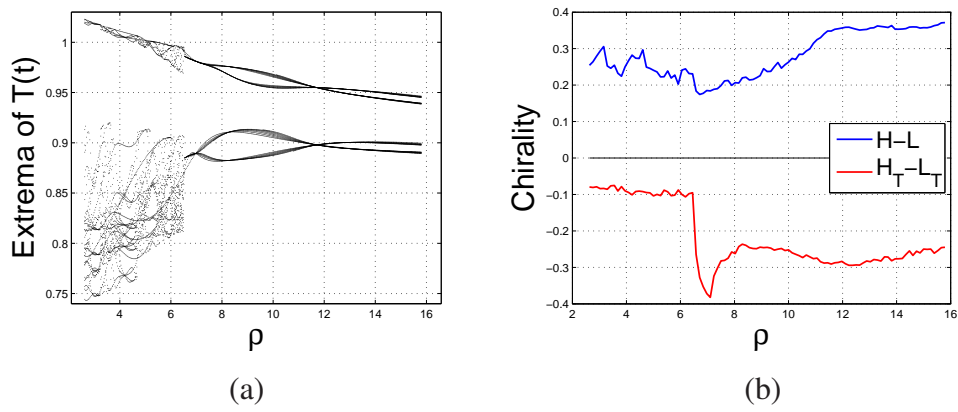


Figure 4: Bifurcation diagrams in (a) the local extrema of the trapper time-series  $T(t)$  and in (b) the trajectory chiralities, respectively, v.s.  $\rho$  for the dimensionless system Eq.(4).