How to Predict Nesting Sites?

Stephen Escarzaga  
*The University of Texas at El Paso*, smescarzaga@utep.edu

Craig Tweedie  
*The University of Texas at El Paso*, ctweedie@utep.edu

Vladik Kreinovich  
*The University of Texas at El Paso*, vladik@utep.edu

Follow this and additional works at: https://scholarworks.utep.edu/cs_techrep

Part of the Mathematics Commons

Comments:


**Recommended Citation**

Escarzaga, Stephen; Tweedie, Craig; and Kreinovich, Vladik, "How to Predict Nesting Sites?" (2015). *Departmental Technical Reports (CS)*. 984.  
https://scholarworks.utep.edu/cs_techrep/984

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.
How to Predict Nesting Sites?

Stephen Escarzaga$^{1,2}$, Craig Tweedie$^{1,2}$, and Vladik Kreinovich$^{1,3}$

$^1$Cyber-ShARE Center
$^2$Department of Biological Sciences
$^3$Department of Computer Science

University of Texas at El Paso
El Paso, TX 79968, USA

smescarzaga@utep.edu, ctweedie@utep.edu, vladik@utep.edu

Abstract

How to predict nesting sites? Usually, all we know is the past nesting sites, and the fact that the birds select a site which is optimal for them (in some reasonable sense), but we do not know the exact objective function describing this optimality. In this paper, we propose a way to make predictions in such a situation.

1 Formulation of the Biological Problem

We observe nesting sites for a certain bird species. Our goals (see, e.g., [4, 8]) are:

- to analyze which criteria are important for selecting nesting sites, and
- to come up with formulas that would enable us to predict nesting sites.

2 Reformulating the Problem in Precise Terms

General description. Let $v_1, \ldots, v_n$ be parameters that may influence the selection of the nesting site: e.g., elevation, hydrology, vegetation level, etc. For each geographic location $\vec{x}$, we record the values of all these variables $v_1(\vec{x}), \ldots, v_n(\vec{x})$.

Main assumption. We assume that the birds select a nesting site based on the values of (some of) these quantities. Namely, a bird tries to maximize the value of some objective function $F(v_1, \ldots, v_n)$ depending on these values.
Simplifying assumption. Let us start with the simplest case, when the objective function is linear, i.e., when

$$F(v_1, \ldots, v_n) = \sum_{i=1}^{n} w_i \cdot v_i$$  \hspace{1cm} (1)$$

for some weights $w_i$.

We assume that each year, each of the observed nesting sites $\vec{x}_j$ has the largest possible value of this objective function among all locations within the corresponding Voronoi cell $C_j$ (see, e.g., [2, 3, 5] and references therein) – i.e., among all locations $\vec{x}$ which are closer to $\vec{x}_j$ than to any other nesting location.

Under this assumption, we would like to find the weights $w_1, \ldots, w_n$ that best explain the observed nesting sites.

3 Analysis of the Problem

The fact that on the cell $C_j$, the linear function (1) attains its largest value at the site $\vec{x}_j$, means that

$$\sum_{i=1}^{n} w_i \cdot v_i(\vec{x}_j) \geq \sum_{i=1}^{n} w_i \cdot v_i(\vec{x}) \text{ for all } \vec{x} \in C_j.$$  

In other words, we should have

$$\vec{w} \cdot \vec{a}(\vec{x}) \overset{\text{def}}{=} \sum_{i=1}^{n} w_i \cdot a_i(\vec{x}) \geq 0,$$  \hspace{1cm} (2)$$

where we denoted $\vec{w} = (w_1, \ldots, w_n)$, $\vec{a}(\vec{x}) = (a_1(\vec{x}), \ldots, a_n(\vec{x}))$, and $a_i(\vec{x}) \overset{\text{def}}{=} v_i(\vec{x}_j) - v_i(\vec{x})$.

Similarly, we should have $w \cdot (-\vec{a}(\vec{x})) \leq 0$ for all $\vec{x}$.

4 How Can We Solve This Problem?

This can be reduced to a known problem. From the mathematical viewpoint, this problem is similar to a linear discriminant analysis (see, e.g., [1, 6, 7]), when:

• we have two sets $A$ and $B$ and
• we need to select a hyperplane that separates them, i.e., a vector $\vec{w}$ for which $\vec{w} \cdot \vec{a} \geq 0$ for all $\vec{a} \in A$ and $\vec{w} \cdot \vec{b} \leq 0$ for all $\vec{b} \in B$.

In our case:

• $A$ is the set of all the vectors $\vec{a}(\vec{x})$, and
• $B$ is the set of all the vectors $-\vec{a}(\vec{x})$.

**How to solve our problem.** The standard way of solving this problem is to compute the mean $\vec{\mu}$ of all the vectors $\vec{a} \in A$, the covariance matrix $\Sigma$, and then to take $\vec{w} = \Sigma^{-1} \vec{\mu}$. So, in our case, we should do the following:

• compute all the vectors $\vec{a}(x)$ with components $a_i(\vec{x}) = v_i(\vec{x}_j) - v_i(\vec{x})$, where $\vec{x} \in C_j$; let $N$ be the total number of such vectors;

• compute the average $\vec{\mu} = \frac{1}{N} \sum_{\vec{x}} \vec{a}(\vec{x})$ of these vectors;

• compute the corresponding covariance matrix $\Sigma$ with components

$$
\Sigma_{ik} = \frac{1}{N} \sum_{\vec{x}} (a_i(\vec{x}) - \mu_i) \cdot (a_k(\vec{x}) - \mu_k);
$$

(4)

• compute the desired weights as $\vec{w} = \Sigma^{-1} \vec{\mu}$, i.e., as a solution to a linear system $\Sigma \vec{w} = \vec{\mu}$.

5 Auxiliary Question: How Can We Gauge the Quality of the Resulting Prediction

To gauge the quality of the resulting prediction, for each cell $C_j$, we compute the location $\vec{c}_j$ at which the weighted combination $\vec{w} \cdot \vec{v}(\vec{x})$ attains its maximum. The mean square distance between these predicted nesting sites $\vec{c}_j$ and the actual nesting sites $\vec{x}_j$ can serve as a natural measure of prediction accuracy.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

References


