Working Paper No. 8208

THE VALUE OF INFORMATION ON THE RESPONSE FUNCTION OF CROPS TO SOIL SALINITY

by

E. Feinerman and D. Yaron

1983
Rehovoth, Israel, P.O.B. 12
The working papers in this series are preliminary and circulated for the purpose of discussion. The views expressed in the papers do not reflect those of the Center for Agricultural Economic Research.
THE VALUE OF INFORMATION ON THE RESPONSE FUNCTION OF CROPS TO SOIL SALINITY

by

E. Feinerman and D. Yaron
THE VALUE OF INFORMATION ON THE RESPONSE
FUNCTION OF CROPS TO SOIL SALINITY

E. Feinerman and D. Yaron

1 The paper is based on parts of an unpublished Ph.D. thesis submitted by E. Feinerman to the Hebrew University. The authors express their thanks to Y. Yahav for his advice on statistical problems and to the Rockefeller Foundation and Israel Foundation Trustees for financial support.

2 The Hebrew University of Jerusalem, Department of Agricultural Economics and Management; presently Post Doctoral Fellow at the University of California, Berkeley and Riverside.

3 The Hebrew University of Jerusalem, Department of Agricultural Economics and Management.

Running Head:
VALUE OF INFORMATION ON RESP. FUN.

Mailing Address:
Eli Feinerman
University of California, Riverside
Department of Soil and Environmental Sciences
Riverside, CA 92521
ABSTRACT

The expected profitability to farmers from acquiring additional information on the biological response function of crop yield to soil-salinity is investigated. First, a switching regression approach to estimate piecewise linear response function with critical threshold level is presented. Then, an optimization irrigation model was developed, aimed at determining the optimal use of irrigation water for soil leaching. Finally, a loss function was defined, the expected value of sample information was calculated and the optimal number of additional needed observations was determined. At each stage, an empirical analysis, using data from potato field experiment in Israel, is presented.
INTRODUCTION

Agricultural production system involves a large number of random variables, numerous of them are physical and biological functions connected with the production process. The knowledge of the biological response function of crop yield to soil salinity is essential in decision making regarding irrigation with saline water. The paper investigates (analytically and empirically) the expected profitability to farmers (the decision makers) of acquiring additional information on this biological function. The true values of the response-function's parameters are usually unknown to the decision maker, and therefore he uses the parameters' estimates and may become a victim of a sub-optimal solution. The deviation from the optimum may be measured by a loss function and the calculation of its expectation. The parameters' estimates (which are arguments in the loss-function) are based on a priori information available to the decision maker. He can acquire additional information which will reduce the variances of these estimates and, hence, will improve his ability to choose a suitable strategy with resulting decrease of the expected loss (or, equivalently, increase of the expected profits). Value of sample information (EVSI) is defined as the difference between the reduction of the expected value of the loss-function due to the additional information and the cost of its acquisition. The optimal number of observations to be acquired is the one that maximizes EVSI.

An accepted hypothesis among soil researchers states that the yield of a given crop is a function of the average soil-salinity in the root
zone during the growing season. Increase of the average soil-salinity level slows down the rate of growth and reduces crop yield (e.g. [1, 11, 20]). The relationships between the soil-salinity level and the reduction of crop yields has been dealt with previously. Some works have shown these relationships in the form of tables [1, 3, 6]; in other studies, the response function was hand-fitted to the available observations [15, 23]. Only few publications report estimates of continuous response functions based on the "best linear unbiased estimates" (BLUE) criterion [14, 20, 22].

A detailed discussion of the response function is found in Maas and Hoffman's article (see [11]). They compiled data on relative yield losses due to salinity with respect to a wide range of crops, i.e., fruit crops, field crops and vegetables. They hypothesized a threshold soil salinity level, beyond which a linear decrease in relative yield is obtained. The critical threshold hypothesis is also presented by [1].

A broad theoretical presentation of decision theory, value of information and the Bayesian approach can be found in [7] and [16]. A number of studies deal with the value of information in farm management [13, 18] as well as in the management of water resources [5, 10]. It should be pointed out that most of these articles did not deal explicitly with the choice of the optimal estimate or with the optimal size of the sample. Furthermore, the articles that dealt with the management of irrigation systems did not refer to water quality.

To calculate the expected profitability to farmers of the additional acquired knowledge about the biological response function to soil salinity, the following stages were taken:
1. Using a switching regression approach (e.g. [17]) to estimate the response function's parameters, according to the response function as formulated by Maas and Hoffman.

2. An optimization irrigation model for a monoculture farm was developed, aimed at determining the optimal quantity of irrigation water from a given source for soil leaching (to reduce salinity).

3. A loss-function was defined, the expected value of sample information was calculated and the optimal number of additional needed observations was determined (taking into consideration their desired spread).

At each stage, an empirical analysis using data from potato field experiments carried out by Sadan and Berglas (see [19]) in the Negev area of Israel, is presented.

1. ESTIMATION OF THE RESPONSE-FUNCTION PARAMETERS

The following model was formulated:

\[
Y = \begin{cases} 
  b_0 + U_1 & \text{if } S \leq S_0 \\
  b_1 + aS + U_2 & \text{if } S > S_0 
\end{cases}
\]

subject to:

\[(*) \quad b_0 = aS_0 + b_1\]

See illustration in Figure 1.
where:

- \( S \) - average soil salinity level in the root zone [meq Cl/1] during the growing season;
- \( S_0 \) - threshold salinity of the soil [meq Cl/1];
- \( Y \) - yield in tonnes per hectare (ha);
- \( U_1, U_2 \) - independent random variables, normally distributed with zero expectation;
- \( b_0, b_1, a, S_0 \) - the (unknown) parameters of the response function satisfying (*)

Assume that we have \( T \) observations \((S_i, Y_i)\) for estimating the above-mentioned parameters. Let us arrange the \( S_i \) in increasing order:

\[
S_1 \leq S_2 \cdots \leq S_t \leq \cdots \leq S_T
\]

where:
- \( S_i \leq S_0 \) for \( i \leq t \) (\( t \) unknown)
- \( S_i > S_0 \) for \( i > t \)

A regression model can be formulated:

\[
\begin{cases}
    b_1 + aS_i + U_{1i} & \text{if } S_i \leq S_0; \ i \leq t \\
    b_1 + aS_i + U_{2i} & \text{if } S_i > S_0; \ i > t
\end{cases}
\]

This model assumes independent normally distributed random deviations with mean zero and \( T \times T \) diagonal variance-covariance matrix-\( \Omega \). Its first \( t \)
diagonal elements are \( V(U_{11}) = \sigma_1^2 \) and the other (T-t) elements are \( V(U_{21}) = \sigma_2^2 \).

The logarithm of the likelihood function \( L(Y|S_0, t) \), given \( S_0 \) and \( t \), is

\[
\ln L(Y|S_0, t) = -T \ln Z_{-1} - t \ln \sigma_1 - (T-t) \ln \sigma_2
\]

\[
-\frac{1}{2\sigma_1^2} \sum_{i=1}^{t} (Y_i - aS_0 - b_1)^2 - \frac{1}{2\sigma_2^2} \sum_{j=t+1}^{T} (Y_j - aS_j - b_1)^2.
\]

Let \( \hat{\beta} = \begin{bmatrix} \hat{b}_1 \\ \hat{a} \end{bmatrix} \) be the vector of maximum likelihood estimates, i.e., \( \hat{\beta} \), which maximizes (3). As the square terms in (3) have negative signs, these estimates are identical to the least square estimates (given \( S_0, t \), so that one can write:

\[
\hat{\beta} (S_0, t) = (Z'\hat{\Omega}^{-1} Z)^{-1} Z'\hat{\Omega}^{-1} Y
\]

where:

\[
Z = \begin{bmatrix} 1 & S_0 \\ \vdots & \vdots \\ 1 & S_t \\ \vdots & \vdots \\ 1 & S_T \end{bmatrix}; \quad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_t \\ \vdots \\ Y_T \end{bmatrix}
\]

(For more detailed discussion see [8]).

The MLE of \( \sigma_1^2 \) and \( \sigma_2^2 \) can be obtained by differentiating (3) and equating to zero:
(5) \[ \hat{\sigma}_1^2(S_0, t) = \frac{1}{T} \sum_{i=1}^{T} (Y_i - \hat{a}S_0 - \hat{b}_i)^2/t \]

(6) \[ \hat{\sigma}_2^2(S_0, t) = \frac{1}{T} \sum_{j=t+1}^{T} (Y_j - \hat{a}S_j - \hat{b}_j)^2/(T-t) \]

(4), (5) and (6) form a set of 4 equations in 4 variables; these can be solved numerically with the aid of a computer [9].

Substituting these estimates into (3) yields:

(7) \[ \ln \hat{L}(Y|S_0, t) = -T \ln \sqrt{2\pi} - t \ln \hat{\sigma}_1 - (T-t)\ln \hat{\sigma}_2 - \frac{T}{2}. \]

Finally the estimates of \( S_0, t \) can be obtained as follows:

First step: Between every two consecutive observations \( S_{i-1}, S_i \) (starting at the third observation and stopping three observations before the end)\(^b\), (7) is maximized numerically over \( S_0 \) as follows:

\[ i = 3, \ldots, T-3 \quad L_i(S_{i-1}, S_i) = \max_{S_{i-1} < S_0 < S_i} \{ \ln \hat{L}(Y|S_0, i) \} \]

\[ L_{T-2}(S_{T-3}, S_{T-2}) = \max_{S_{T-3} < S_0 < S_{T-2}} \{ \ln \hat{L}(Y|S_0, T-2) \} \]

Second step: The optimal estimates of \( S_0, t \) are \( \hat{S}_0, \hat{t} \) which satisfy

\[ \ln \hat{L}(Y|\hat{S}_0, \hat{t}) = L_t(S_{t-1}, S_t) = \max_{2 < i < T-2} L_i(S_{i-1}, S_i) \]

Let \( \hat{\theta} = [\hat{S}_0, \hat{b}_1, \hat{a}] \).
From the properties of MLE (e.g. [12]), under fairly general conditions, 
\( \hat{\theta} \) is asymptotically normally distributed with mean \( \hat{\theta} = [S_0, b_1, a] \) and variance covariance matrix:

\[
E_{\theta} = \left[ E \left[ \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right] \right]^{-1}.
\]

The variances and covariances can be calculated as functions of the observations \( S_1 \) and of \( a, S_0, t, T, \sigma_1^2, \sigma_2^2 \).

**Empirical Results**

The response function estimates are based on the experimental results of Sadan and Berglas (personal communication). Their one year experiment, conducted in the Northern Negev of Israel, provided a total of 17 observations \((Y_i, S_i)\) as presented in Table I.

[Table I about here]

The following estimates for the response function parameters (2) were obtained by substituting the data in equations (4), (5) and (6):

\[
\begin{align*}
\hat{S}_0 &= 6.054 \text{ [meq Cl/l]} \\
\hat{b}_1 &= 52.55 \text{ [tonnes/ha]} \\
\hat{a} &= -1.09 \text{ [tonnes/ha/meq Cl/l]}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Asymptotic standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{S}_0 = 6.054 \text{ [meq Cl/l]} )</td>
</tr>
<tr>
<td>( \hat{b}_1 = 52.55 \text{ [tonnes/ha]} )</td>
</tr>
<tr>
<td>( \hat{a} = -1.09 \text{ [tonnes/ha/meq Cl/l]} )</td>
</tr>
</tbody>
</table>
### TABLE I

Relationships Between Average Soil-Salinity ($S_i$) and Potato Yield ($Y_i$) - The Negev Area.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_i^a$</td>
<td>45.4</td>
<td>46.2</td>
<td>46.2</td>
<td>48.7</td>
<td>45.8</td>
<td>48.1</td>
<td>33.2</td>
<td>45.4</td>
<td>38.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_i$</td>
<td>31.8</td>
<td>34.3</td>
<td>33.6</td>
<td>34.9</td>
<td>25.3</td>
<td>28.5</td>
<td>29.1</td>
<td>32.4</td>
</tr>
</tbody>
</table>

---

*a* Yield [tonnes/ha].

*b* Average soil salinity level [meq Cl/l] in the rootzone (0-60 cm), during the growing period.
Accordingly, \( b_0 = 45.95 \) [tonnes/ha] and the estimated response function is:

\[
y_i = \begin{cases} 
45.95 & \text{if } S_i \leq 6.054 \\
52.55 - 1.09 S_i & \text{if } S_i > 6.054 
\end{cases}
\]

with other relevant statistics being:

\[ \hat{t} = 3, \quad \hat{\sigma}_1^2 = 0.14 \text{ [tonnes/ha]}^2; \quad \hat{\sigma}_2^2 = 15.78 \text{ [tonnes/ha]}^2 \]

and \( R^2 = 0.78 \).

The scatter-diagram of the observations and the fitted regression line are shown in Figure 2.

It should be noted that this response-function (2) was compared to estimation of alternative formulations (all of them monotonically decreasing, continuous and differentiable functions) and was judged "best" [8]. An "adjusted R square" \( R^2 \) adjusted for degrees of freedom) was used as a criterion in this comparison.
2. OPTIMAL USE OF IRRIGATION WATER FROM A GIVEN SOURCE
FOR A GIVEN CROP (POTATOES)

Consider the yield expectation:

\[
\text{EY} = \begin{cases} 
  aS_0 + b_1 & \text{if } S \leq S_0 \\
  aS + b_1 & \text{if } S > S_0 
\end{cases}
\]

Let $\overline{S}$ be the present soil salinity level, $\overline{S} > S_0$. Assume that there is a source of good quality water of salinity $C_1 \text{[meq C1/l]}$. Leaching the soil with $Q \text{ m}^3$ of this water per hectare will reduce the salinity to $S$, where:

\[
\overline{S} = \frac{Q C_1}{N + \frac{\gamma Q}{2}} + \overline{S} \left(1 - \frac{\gamma Q}{2(N + \frac{\gamma Q}{2})}\right) \quad (\text{see [4]})
\]

($N$ and $\gamma$ are known soil parameters).

Isolating $Q$ gives:

\[
Q(S) = \frac{N(\overline{S} - S)}{\frac{\gamma}{2} (\overline{S} + S) - C_1} \quad (9)
\]

A good empirical approximation is obtained by the following quadratic regression:

\[
Q(S) = k_1 + k_2(\overline{S} - S) + k_3(\overline{S} - S)^2 \quad (10)
\]
Let \( S = \frac{\bar{S} + \Sigma}{2} \) be an approximation of the average salinity level (before and after leaching). Substituting \( S = 2\bar{S} - \bar{S} \) in (10) yields:

\[
Q(S) = K_1 + K_2(\bar{S} - S) + K_3(\bar{S} - S)^2
\]

where \( K_2 = 2K^2_2, K_3 = 4K^3_2 \).

Let \( P \) be the cost of the above-mentioned water supply in dollars/m\(^3\). The cost of leaching the soil (with \( Q \) m\(^3\) per hectare) denoted by \( C(\bar{S}, S) \), is:

\[(11) \quad C(\bar{S}, S) = P \cdot Q(S) = PK_1 + PK_2(\bar{S} - S) + PK_3(\bar{S} - S)^2.\]

A profit function is defined as:

\[
\pi(S) = \begin{cases} 
R_1(aS_0 + b_1) - R_2 - C(\bar{S}, S) & \text{if } S \leq S_0 \\
R_1(aS + b_1) - R_2 - C(\bar{S}, S) & \text{if } S > S_0 
\end{cases}
\]

where:

- \( R_1 \) - net income in dollars per unit yield (tonnes) as a function of the yield (revenue, less variable cost dependent on yield, such as: harvesting, grading, packing and transportation).

- \( R_2 \) - variable costs in dollars/ha, independent of yield.

By substituting (11) into (12) and then equating the derivative \( \frac{\partial \pi(S)}{\partial S} \) (for \( S > S_0 \)) to zero, \( S^* \) which maximizes (12) is accepted:

\[
S^* = \bar{S} + \frac{R_1a + PK_2}{2PK_3}
\]
Since it is obvious that \( S^* > S_0 \), it can be written:

\[
S^* = \max \left( S_0, \bar{S} + \frac{R_1 \hat{a} + PK_2}{2PK_3} \right).
\]

And, by substituting MLE \( \hat{a} \), \( \hat{S}_0 \) for the unknown parameters \( a \), \( S_0 \):

\[
\hat{S}^* = \max \left( \hat{S}_0, \bar{S} + \frac{R_1 \hat{a} + PK_2}{2PK_3} \right).
\]

**Empirical Results**

The empirical approximation (10) to the leaching function (9), was achieved by dividing the relevant range of soil-salinity into a large number of discrete points, calculating the value of \( Q(S) \) by (9) for each point and estimating the regression-line (10).

For \( \bar{S} = 20 \) [meq Cl/1] (A somewhat high initial soil salinity level was chosen in order to emphasize the need of soil-leaching), \( N = 3500 \) [m³/ha], \( \gamma = 0.7 \) (average irrigated-soil parameters in the study area), and \( Cl = 5 \) [meq Cl/1], the following estimates were obtained:

\[
\begin{align*}
K_1 &= 132 \\
K_2 &= 526 \quad K_2^* = 263 \\
K_3 &= 146 \quad K_3^* = 36.5
\end{align*}
\]

(9) and (10) are presented in Figure 3 on the same set of axes (function (9) is marked by the numeral 1, and (10) by the numeral 2).

[Fig. 3 about here]

With \( P = 0.1 \) [dollars/m³] ; \( R_1 = 161 \) [dollars/tonne], the following values for potatoes were obtained:
\[ \hat{S}^* = 15.8 \text{ [meq Cl/1]} \]

\[ Q(\hat{S}^*) = 4917 \text{ [m}^3/\text{ha}] \]

3. THE LOSS FUNCTION AND THE VALUE OF ADDITIONAL INFORMATION

In the following section, the loss function and its possible situations are defined, the expected value of sample information is calculated and the optimal sample size is determined (taking into consideration the desired spread of the additional observations).

Let us define a loss function:

\[ \text{LOSS}(S^*, \hat{S}^*) = \pi(S^*) - \pi(\hat{S}^*) \]

One may distinguish between 8 alternatives associated with its possible values, based on all possible combinations of the relationships between \( \hat{S}_0 \) and \( S_0 \), \( S^* \) and \( S_0 \), \( \hat{S}^* \) and \( \hat{S}_0 \). But, four of them can be disregarded, since \( \hat{a}, \hat{S}_0 \) are consistent estimates (being MLE) and therefore tend to \( a, S_0 \), respectively, so that:

\[ P_r(\hat{S}^* > S_0 \text{ and } \hat{S}^* \leq \hat{S}_0) \to 0 \]

(16)

\[ P_r(\hat{S}^* \leq S_0 \text{ and } \hat{S}^* > \hat{S}_0) \to 0 \]

(17)

The four remaining alternatives can be described as follows:

(a) \( \hat{S}_0 > S_0, \ S^* > S_0, \ \hat{S}^* > \hat{S}_0 \)

Substituting in (15) yields:
\text{LOSS}(S^*, \hat{S}^*) = \{R_1 a \left(\bar{s} + \frac{R_1 a + PK_2}{2PK_3}\right) + R_1 b_1 - R_2 \}

- \{R_1 a \left(\bar{s} + \frac{R_1 \hat{a} + PK_2}{2PK_3}\right) + R_1 b_1 - R_2 \}

- \{R_1 a \left(\bar{s} - \frac{R_1 \hat{a} + PK_2}{2PK_3}\right) - \bar{s} + \frac{R_1 a + PK_2}{2PK_3} - \bar{s}\}^2\}

= \frac{R_1^2 (a - a)^2}{4PK_3}

Similarly:

(b) \ \hat{s}_0 \leq S_0, \ S^* > S_0, \ \hat{S}^* > \hat{s}_0

\text{LOSS}(S^*, \hat{S}^*) = \frac{R_1^2 (a - a)^2}{4PK_3}

(c) \ \hat{s}_0 \leq S_0, \ S^* = S_0, \ \hat{S}^* = \hat{s}_0

\text{LOSS}(S^*, \hat{S}^*) = \{R_1 (aS_0 + b_1) - R_2 - P[K_1 + K_2 (\bar{s} - S_0) + K_3 (\bar{s} - S_0)^2]\}

- \{R_1 (aS_0 + b_1) - R_2 - P[K_1 + K_2 (\bar{s} - \hat{s}_0) + K_3 (\bar{s} - \hat{s}_0)^2]\}

= - PK_2 (\hat{s}_0 - S_0) - PK_3 [(\bar{s} - S_0)^2 - (\bar{s} - \hat{s}_0)^2]

(d) \ \hat{s}_0 > S_0, \ S^* = S_0, \ \hat{S}^* = \hat{s}_0
\[
\text{LOSS}(\hat{S}^*, S^*) = \{R_1(aS_0 + b) - R_2 - P[K_1 + K_2(\tilde{S} - S_0) + K_3(\tilde{S} - S_0)^2]\}
\]
\[
- \{R_1(a\hat{S}_0 + b) - R_2 - P[K_1 + K_2(\tilde{S} - S_0) + K_3(\tilde{S} - \hat{S}_0)^2]\}
\]
\[
= - (R_1a + PK_2)(\hat{S}_0 - S_0) - PK_3[(\tilde{S} - S_0)^2 - (\tilde{S} - \hat{S}_0)^2]
\]

Using indicator functions we write the loss function concisely as:

\[
(18) \text{LOSS}(\hat{S}^*, S^*) = \frac{R_1^2(a-a)^2}{4K_3P} \cdot I_{\{S^* > S_0\}} + \{- PK_3[(\tilde{S} - S_0)^2 - (\tilde{S} - \hat{S}_0)^2]\}
\]
\[
- PK_2(\hat{S}_0 - S_0) I_{\{S^* = S_0\}} - R_1a(\hat{S}_0 - S_0) I_{\{S^* = S_0\}} I_{\{S_0 < \hat{S}_0\}}
\]

where \( I \) takes values of 1 or 0 as follows:

\[
I_{\{\text{expression}\}} = \begin{cases} 1 & \text{expression true} \\ 0 & \text{otherwise} \end{cases}
\]

As \( \hat{S}_0 \) are random variables, the loss function is also random.

For given values of \( S_0, a, \sigma_1^2, \sigma_2^2 \), and for a given scatter \( \bar{S}_T \) of the observations \( S_1, ..., S_T \), the conditional expectation of the loss function is:

\[ \]
\( \text{LOSS} \left( \sigma^2_1, \sigma^2_2, a, S_0, S_{\infty} \right) = E[\text{LOSS} \left( S^*, \hat{S}^* \right) / \sigma^2_1, \sigma^2_2, a, S_0, S_{\infty}] = \)

\[
= \frac{R_1^2}{4K_3P} \int_{-\infty}^{\infty} (a - \hat{a})^2 dN(a, V(\hat{a})) I_{\{S^* > S_0\}}
\]

\[+ \left\{ -2k_3S \int_{-\infty}^{\infty} (\hat{S}_0 - S_0) dN(S_0, V(\hat{S}_0)) + PK_3 \int_{-\infty}^{\infty} (\hat{S}_0^2 - S_0^2) dN(S_0, V(\hat{S}_0)) \right\}
\]

\[
- PK_2 \int_{-\infty}^{\infty} (\hat{S}_0 - S_0) dN(S_0, V(\hat{S}_0)) I_{\{S^* = S_0\}} - R_1 a \int_{S_0}^{\infty} (\hat{S}_0 - S_0) dN(S_0, V(\hat{S}_0)) I_{\{S^* = S_0\}}
\]

\[
= \frac{R_1^2}{4K_3P} V(\hat{a}) I_{\{S^* > S_0\}}
\]

\[+ \left\{ PK_3 V(\hat{S}_0) - \frac{R_1 a}{\sqrt{2\pi} V(\hat{S}_0)} \right\} I_{\{S^* = S_0\}}
\]

Let \( G(\sigma^2_1, \sigma^2_2, a, S_0) \) be the joint prior distribution of \( \sigma^2_1, \sigma^2_2, a, S_0 \). Assume that these four variables are independent, and their marginal distributions known:

(i) \( P_r(\sigma^2_1 = d_1) = P_r(\sigma^2_2 = d_2) = 1 \)

(ii) \( a \sim U[\delta_1, \delta_2], \delta_1 \leq \delta_2 \leq 0 \)

(iii) \( S_0 \sim U[\Delta_1, \Delta_2], 0 \leq \Delta_1 \leq \Delta_2 \)

where \( d_1, d_2, \delta_1, \delta_2, \Delta_1, \Delta_2 \) are known.

Under these assumptions, the expectation of (19) will be:
Let us now calculate the profitability of acquiring additional observations \((S_1, Y_1)\). For a sample of a given size \(T\) let \(S^*\) be the scatter which minimizes (20). According to Yahav (see [21]) under the assumptions (i) - (iii):

(i) \(\sigma_1^2 = \sigma_2^2\)

(ii) \(S_0 \sim U[0, 1]\)

(iii) there is a continuum of observations,

the optimal spread will be according to the Beta \((\frac{1}{3}, \frac{1}{3})\) density:

\[
f_{\frac{1}{3}, \frac{1}{3}}(S) = \begin{cases} \frac{1}{S^3} \frac{1}{(1-S)^3} \frac{(5/3)!}{(\frac{1}{3})! (\frac{1}{3})!} & \text{if } 0 \leq S \leq 1 \\ 0 & \text{otherwise} \end{cases}
\]
(i) \( \sigma_1^2 = d_1 \neq \sigma_2^2 = d_2 \)

(ii) \( S_0 \sim U[\Lambda_1, \Lambda_2] : \)

\( \frac{1}{3} d_1, \frac{2}{3} - \frac{1}{3} \frac{d_1}{d_2} \) observations were taken in the vicinity of \( \Lambda_1 \), \( \frac{2}{3} - \frac{1}{3} \frac{d_1}{d_2} \) in the vicinity of \( \Lambda_2 \), and 1/3 uniform spread between the above two groups.

At this stage we have no proof that this spread will indeed minimize (20), and it is only an intuitive approximation to Yahav's result.

Finding the optimal spread is a complicated statistical problem, whose analytical solution would not be discussed here. However, the suggested spread (hereafter "spread I") was empirically compared with two other alternative spreads of additional observations.

Assume that \( D \) hectares of potatoes are grown in the region where the \((S_i, Y_i)\) observations were taken, with the same technology, soil and climate conditions. Let \( n \) be the number of additional observations to be taken and \( C_0(n) \) - the cost of their acquisition. With \( H(T, S^*_1) \) describing the situation a priori, the expected value of additional information - to the potato growers of that region - from \( n \) observations with spread \( S^*_{T+n} \) is:

\[
(21) \quad EVSI(n) = D[H(T, S^*_1) - H(T+n, S^*_{T+n})] - C_0(n) .
\]

The optimal number of observations, \( n^* \), can be determined by:

\[
(22) \quad EVSI(n^*) = \max_n EVSI(n) .
\]
Empirical Results

The empirical application of the above described analysis, to a potato growing region in the Northern Negev area of Israel, is presented below. The aggregate expected benefits to the region's potato growers, from improved response data are calculated and the optimal sample size is determined.

Based on the empirical estimates of the response-function parameters, the following approximate values were assigned to $d_1$, $d_2$, $\delta_1$, $\delta_2$, $\Lambda_1$ and $\Lambda_2$:

\[ d_1 = 0.1 = \hat{\sigma}_1^2, \quad d_2 = 16 = \hat{\sigma}_2^2 \]
\[ \delta_1 = -0.88 = \hat{a} + 0.21, \quad \delta_2 = -1.30 = \hat{a} - 0.21 \]
\[ \Lambda_1 = 4.5 = \hat{s}_0 - 1.5, \quad \Lambda_2 = 7.5 = \hat{s}_0 + 1.5 \]

Accordingly, the marginal a priori distributions are:

\[ Pr(\sigma_1^2 = 0.1) = 1, \quad Pr(\sigma_2^2 = 16) = 1 \]
\[ a \sim U(-0.88, -1.30) \]
\[ s_0 \sim U(4.5, 7.5) \]

The suggested "spread I" of additional observations was compared with two alternative spreads, "spread II" uniformly scattered in the range $s_0 = 4.5$ to $s_0 = 7.5$, and "spread III" uniformly scattered from $s_0 = 6.054 (\hat{s}_0)$ to $s_0 = 20(\hat{s})$. For $P=0.1[\text{dollars/m}^3]$, $R_1 = 161[\text{dollars/tonne}]$, $D = 2000[\text{ha}]$ (the Northern Negev, where this experiment was
conducted, is the main producer area of potatoes in Israel), and
\[ C_0(n) = 260 \text{ N[dollars]} \] (these expenses constitute 130 dollar/observation of direct costs + 130 dollars/observation due to the opportunity cost of the research personnel); the obtained values of EVSI(n) are presented in Table II.

[Table II about here]

Based on the results of Table II the following should be emphasized:

(i) "Spread I" is substantially superior over the two other spreads for all values of n. Based on "spread I", which is an intuitive approximation of Yahav's [21] findings, the estimated optimal sample size is \( n^* = 15 \) and the expected value of additional information is 23992 dollars.

(ii) Since some approximations were used and since the underlying statistical theory is mainly asymptotic and assumes the use of large sample, the results, which are based on medium size sample, must be regarded as approximate. Their main value is that they enable us to learn the order of magnitude of EVSI and to draw operative conclusions about the benefit of additional sampling.

(iii) The present level of knowledge of the potato growers in that region is relatively high and any improvement in production due to a better knowledge of the response-function is expected to be relatively small. Since a
TABLE II

The Expected Value of Sample Information (EVSI(n)) for the

Three Alternative Observations Spreads (Dollars)

<table>
<thead>
<tr>
<th>Number of Observations - n</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>15</th>
<th>18</th>
<th>21</th>
<th>24</th>
<th>27</th>
<th>30</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVSI(n), &quot;Spread I&quot;</td>
<td>14364</td>
<td>17185</td>
<td>22290</td>
<td>23658</td>
<td>23992</td>
<td>23898</td>
<td>22993</td>
<td>21058</td>
<td>19902</td>
<td>19600</td>
<td>19216</td>
</tr>
<tr>
<td>EVSI(n), &quot;Spread II&quot;</td>
<td>4610</td>
<td>8843</td>
<td>11277</td>
<td>12782</td>
<td>13614</td>
<td>14133</td>
<td>14398</td>
<td>14482</td>
<td>14424</td>
<td>14272</td>
<td>14032</td>
</tr>
<tr>
<td>EVSI(n), &quot;Spread III&quot;</td>
<td>6562</td>
<td>9149</td>
<td>11253</td>
<td>13550</td>
<td>14631</td>
<td>15467</td>
<td>16330</td>
<td>16738</td>
<td>17022</td>
<td>16742</td>
<td>16298</td>
</tr>
</tbody>
</table>
short run optimization model was assumed, the results might be regarded as downward biased: actually, improved estimates of the response function's parameters may contribute to the region farmers' benefits for more than one growing season. It should be expected, however, that multi-period optimization model (with soil salinity as dynamic state variable), may yield a higher values of ESVI.

SUMMARY

The estimation of the "response function" of a given crop to soil salinity and the calculation of the expected value of additional information on the parameters of this function are important steps in the process of decision making, regarding irrigation with saline water under conditions of uncertainty.

A method for the estimation of a "response function" following Maas and Hoffman's (1977) specification was developed and the stochastic properties of the estimated parameters were discussed.

An optimization model for the determination of the optimal quantity of water from a given source needed to leach the soil was formulated. A loss function was constructed, its possible states were defined and its expectation derived.

Finally, the expected value of sample information (EVSI) on the response function parameters was calculated. The optimal sample size was determined, with regard to the preferable spread of the additional observations.
The empirical results are depending on the physical data and the assumptions made and are therefore relevant to the region under consideration. Although the procedures introduced are quite complex, once a computer program is written, they can be applied to other crops and regions at a relatively low cost.

There are at least two directions for possible extension of the analysis: (i) Computing expected value of sample information (EVSI), in a long-run analysis, referring to the water-soil-crop farm system over a sequence of several irrigation seasons, taking into consideration the long-run soil-leaching process. (ii) Computing EVSI for a single crop within a multi-culture farm framework, with several crop alternatives, several water sources differing in quality, quantity and prices, and several fields differing in area and initial salinity (Feinerman, 1980). The analysis presented in this paper can serve as a building block in such extended analyses. Its main advantage seems to be in providing conceptual and methodological framework to investigate the problem as well as an efficient tool for empirical analysis.
REFERENCES


FOOTNOTES

4 Deleting the first two and the last three observation intervals eliminates the possibility \( \hat{\sigma}_1 = 0 \) or \( \hat{\sigma}_2 = 0 \) which would make (7) equal infinity for any \( \hat{S}_0 \) in these intervals.

5 All the monetary values are at January 1978 price level.

6 The observations \( Y_i \) are normally distributed, the prior density function is everywhere positive, and the loss function is proportional to the squared error. It is therefore asymptotically true that Bayes estimates (the parameter estimates which minimize the expected loss) are identical to the MLE, (e.g. [2]).

7 Based on (16) and (17), one may disregard the case \( S^* < S_0, S^* > S_0 \). Hence, the loss function for (b) should be the same as the loss function for (a).

8 Based on conversations with soil researchers and our a priori knowledge, we believe that it is possible to assign closed intervals to the unknown true parameters. Since we are not able to assign different probabilities to subset lengths of the above mentioned interval, we assume prior uniform distributions.

9 The optimal EVSI (23992 dollars/region) found in Table 2 constitute less than one percent of the total revenue of the potato growers.
FIGURE LEGENDS

Figure 1: The Response Function.

Figure 2: The Estimated Response Function of Potatoes.

Figure 3: Quantity of Leaching Water (Q) as a Function of the Target Soil Salinity (S) (Initial Soil Salinity $S = 20$ meq Cl/L).
FUNCTION (10)

FUNCTION (9)

LEACHING WATER (m³/HECTARE)

TARGET SOIL SALINITY (meq cl/1)
PREVIOUS WORKING PAPERS

6901 Yoav Kislev and Hanna Lifson - An Economic Analysis of Drainage Projects.
6902 Yair Mundlak and Ran Mosenson - Two-Sector Model with Generalized Demand.
6903 Yoav Kislev - The Economics of the Agricultural Extension Service. (Also in Hebrew).
7001 Dan Yaron and Gideon Fishelson - A Survey of Water Mobility on Moshav Villages. (Hebrew).
7004 Yoav Kislev and Hanna Lifson - Capital Adjustment with U-Shaped Average Cost of Investment.
7005 Yair Mundlak - Empirical Production Functions with a Variable Firm Effect.
7006 Yair Mundlak - On Some Implications of Maximization with Several Objective Functions.
7101 Yair Mundlak and Assaf Razin - On Multistage Multiproduct Production Function.
7102 Yakir Plessner and Meir G. Kohn - Monopolistic Behavior in Situations of Expectation Motivated Demand.
7104 Yoav Kislev and Yakir Plessner - An Applicable Linear Programming Model of Inter-Temporal Equilibrium.
7106 Yair Mundlak and Zvi Volcani - The Correspondence of Efficiency Frontier as a Generalization of the Cost Function.
7107 Uri Regev and Aba Schwartz - Optimal Path of Interregional Investment and Allocation of Water.
7108 Eitan Hochman and Hanna Lifson - Optimal Control Theory Applied to a Problem of an Agricultural Marketing Board Acting as a Monopolist.
7201 Mordechai Weisbrod, Gad Stretiner, Dan Yaron, Dan Shimshi, Eshel Bresler - A Simulation Model of Soil Variation Moisture. (Hebrew).
7202 Yoav Kislev, Yakir Plessner, Aharon Perahia - Multi-Period Linear Programming with a Consumption Application. (Hebrew).
7203 Ran Mosenson - Fundamental Dual Price-Rent Relations in Input-Output Analysis - Theory and Application.
7301 Yigal Danin and Yair Mundlak - The Effect of Capital Accumulation on a Well Behaved n-Sector Economy.
7303 Aba Schwartz, Uri Regev and Shmuel Goldman - Estimation of Production Functions Free of Aggregation Bias with an Application to the Israeli Agriculture.


Meir G. Kohn - Competitive Speculation.


Jack Habib, Meir Kohn and Robert Lerman - The Effect on Poverty Status in Israel of Considering Wealth and Variability of Income.

Yoav Kislev, Michal Meisels, Shmuel Amir - The Dairy Industry of Israel.

Yair Mundlak - Agricultural Growth in the Context of Economic Growth.


Yair Mundlak - Models with Variable Coefficients - Integration and Extension.


Yoav Kislev - The Monetary Approach to the Israeli Balance of Payments.


Yair Mundlak and Joseph Yahav - ANOVA, Convolution and Separation, A Fresh View at Old Problems.

Meir Kohn - Why the Dynamic Competitive Producer Should Not Carry Stocks of his Product.

Yair Mundlak - Agricultural Growth - Formulation, Evaluation and Policy Consequences.

Dan Yaron, A. Dinar and S. Shamlah - First Estimates of Prospective Income Losses Due to Increasing Salinity of Irrigation Water in the South and the Negev Regions of Israel. (Hebrew).

Yair Mundlak - On the Concept of Non-Significant Functions and its Implications for Regression Analysis.

Pinhas Zusman and Michael Etgar - The Marketing Channel as an Equilibrium Set of Contracts.

Yakir Plessner and Shlomo Yitzhaki - The Firm's Employment Policy as a Function of Labor Cost Structure.


Yigal Danin and Yair Mundlak - The Introduction of New Techniques and Capital Accumulation.

Yair Mundlak - Elements of a Pure Theory of Forecasting and the "After Keynesian Macroeconometrics".

Yoav Kislev and Willis Peterson - Prices, Technology and Farm Size.

David Bigman and Haim Shalit - Applied Welfare Analysis for a Consumer Whose Income is in Commodities.

David Bigman - Semi-Rational Expectations and Exchange Rate Dynamics.

Joel M. Guttmann - Can Political Entrepreneurs Solve the Free-Rider Problem?
Yakir Plessner and Haim Shalit - Investment and the Rate of Interest Under Inflation: Analysis of the Loanable Funds Market.

Haim Shalit - Who Should Pay for Price Stabilization?

David Bigman - Stabilization and Welfare with Trade, Variable Levies and Internal Price Policies.

Haim Shalit, Andrew Schmitz and David Zilberman - Uncertainty, Instability and the Competitive Firm.

David Bigman - Buffer Stocks and Domestic Price Policies.

David Bigman - National Food Policies in Developing Countries: The Experience and the Lesson.


Yoav Kislev and Willis Peterson - Induced Innovations and Farm Mechanization.

Yoav Kislev and Yakir Plessner - Recent Inflationary Experience in Israel.

Yair Mundlak - Cross Country Comparison of Agricultural Productivity.

Michael Etgar & Ilan Peretz - The Preference of the German Market for Quality Tomatoes (Hebrew).

Tzvi Sinai - The Profitability of Land Development for Agriculture in Israel (Hebrew).

Ilan Beer - Economic Aspects of the Settlement Project in Yamit (Hebrew).

David Bigman - Stabilization and International Trade.

Navá Haruvi and Yoav Kislev - Cooperation in the Moshav.

Michal Meisels-Reis - Specialization and Efficient in the Poultry Industry in Israel (Hebrew).

Ilan Been - Economic Aspects of the Settlement Project in Yamit (Hebrew).


Yair Mundlak - Various Aspects of the Profitability of Milk Production. (Hebrew)

Yair Mundlak & Joseph Yahav - Inference with Stochastic Regressors.

Pinhas Zusman & Clive Bell - The Equilibrium Set of Dyadic Contracts.

Yoav Kislev & Shlomit Farbstein - Capital Intensity and Product Composition in the Kibbutz and the Moshav in Israel.

David Bigman - Food Aid and Food Distribution.

Haim Shalit and Shlomo Yitzhaki - Mean-Gini, Portfolio Theory and the Pricing of Risky Assets.

Rafi Melnick & Haim Shalit - The Market for Tomatoes: An Empirical Analysis. (Hebrew)

Dan Yaron & Hillary Voet - Optimal Irrigation With Dual Quality (Salinity) Water Supply and the Value of Information.

David Bigman & Itzhak Weksler - Strategies for Emergency Stock Planning.