HUCAER

WP-8208

המרכז למחקר בכלכלה חקלאית

Hebrew university

THE CENTER FOR AGRICULTURAL ECONOMIC RESEARCH

Working Paper No. 8208

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

by

E. Feinerman and D. Yaron

Rehovoth, Israel, P.O.B. 12

רחובות, ת.ד. 12

Working Paper No. 8208

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[1]

E. Feinerman[2] and D. Yaron[3]

[1]The paper is based on parts of an unpublished Ph.D. thesis submitted by E. Feinerman to the Hebrew University.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:

$$(1) \qquad Y = \begin{cases} b_0 + U_1 & \text{if} \quad S \leq S_0 \\ b_1 + aS + U_2 & \text{if} \quad S > S_0 \end{cases}$$

subject to:

$$(*)\ b_0 = aS_0 + b_1$$

See illustration in Figure 1.

[Figure 1 about here]

where:

S - average soil salinity level in the root zone [meq Cl/l] during the growing season;

S_0 - threshold-salinity of the soil [meq Cl/l];

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:

$$S_1 \leq S_2 \quad \cdots \quad \leq S_t \leq \quad \cdots \quad \leq S_T$$

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:

$$(2) \qquad Y_i = \begin{cases} b_1 + aS_0 + U_{1i} & \text{if} \quad S_i \leq S_0; \quad i \leq t \\ b_1 + aS_i + U_{2i} & \text{if} \quad S_i > S_0; \quad i > t \end{cases}$$

This model assumes independent normally distributed random deviations with mean zero and TxT diagonal variance-covariance matrix-Ω. Its first t diagonal elements are $V(U_{1i}) = \sigma_1^2$ and the other (T-t) elements are $V(U_{2i}) = \sigma_2^2$.

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

$$(3)\quad \ln L(Y|S_0, t) = -T \ln\sqrt{2\pi} - t \ln \sigma_1 - (T-t)\ln \sigma_2$$

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

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

$$(4)\qquad \hat{\underline{\beta}}\ (S_0, t) = (Z'\hat{\Omega}^{-1} Z)^{-1} Z'\hat{\Omega}^{-1} Y$$

where:

$$Z = \begin{bmatrix} 1 & S_0 \\ \vdots & \vdots \\ 1 & S_0 \\ 1 & S_{t+1} \\ \vdots & \vdots \\ 1 & S_T \end{bmatrix} ; \qquad Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_t \\ Y_{t+1} \\ \vdots \\ Y_T \end{bmatrix}$$

(For more detailed discussion see [8]).

The MLE of σ_1^2 and σ_2^2 can be obtained by differentiating (3) and equating to zero:

(5) $$\hat{\sigma}_1^2(S_0, t) = \sum_{i=1}^{t} (Y_i - \hat{a}S_0 - \hat{b}_1)^2/t$$

(6) $$\hat{\sigma}_2^2(S_0, t) = \sum_{j=t+1}^{T} (Y_j - \hat{a}S_j - \hat{b}_1)^2/(T-t)$$

(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)[4], (7) is maximized numerically over S_0 as follows:

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

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

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

$$\ln \hat{L}(Y|\hat{S}_0, \hat{t}) = L_{\hat{t}} (S_{\hat{t}-1}, S_{\hat{t}}) = \max_{2 < i < T-2} L_i(S_{i-1}, S_i) .$$

Let $\underline{\hat{\theta}} = [\hat{S}_0, \hat{b}_1, \hat{a}]$.

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

$\Sigma_{\underline{\hat{\theta}}} = \left[E \left[\frac{- \partial^2 \ln L}{\partial\theta_i \quad \partial\theta_j} \right]^{-1} \right]$. The variances and covariances can be calculated as functions of the observations S_i and of a, S_0, t, T, σ_1^2, σ_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):

	Asymptotic standard deviation
$\hat{S}_0 = 6.054$ [meq Cl/l]	(1.92)
$\hat{b}_1 = 52.55$ [tonnes/ha]	(3.2)
$\hat{a} = -1.09 \left[\frac{\text{tonnes/ha}}{\text{meq Cl/l}}\right]$	(0.2)

TABLE I

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

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Y_i [a]	45.4	46.2	46.2	48.7	45.8	48.1	33.2	45.4	38.9	31.8	34.3	33.6	34.9	25.3	28.5	29.1	32.4
S_i [b]	4.015	4.234	4.964	6.058	7.664	8.394	12.263	12.401	12.774	12.847	16.642	18.102	18.832	19.781	20.292	22.263	22.701

[a] Yield [tonnes/ha].

[b] Average soil salinity level [meq Cl/l] in the rootzone (0-60 cm), during the growing period.

Accordingly, $\hat{b}_0$ = 45.95 [tonnes/ha] and the estimated response function is:

$$i = 1, \ldots, 17 \qquad \hat{Y}_i = \begin{cases} 45{,}95 & \text{if } S_i \leq 6.054 \\ 52.55 - 1.09\, S_i & \text{if } S_i > 6.054 \end{cases}$$

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.

[Fig. 2 about here]

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:

$$(8)\qquad EY = \begin{cases} aS_0 + b_1 & \text{if } S \leq S_0 \\ aS + b_1 & \text{if } S > S_0 \end{cases}$$

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

$$\underline{S} = \frac{Q\ Cl}{(N + \frac{\gamma}{2}Q)} + \bar{S}\,\frac{(N - \frac{\gamma}{2}Q)}{(N + \frac{\gamma}{2}Q)} \qquad \text{(see [4])}$$

(N and γ are known soil parameters).

Isolating Q gives:

$$(9)\qquad Q(\underline{S}) = \frac{N(\bar{S} - \underline{S})}{\frac{\gamma}{2}(\bar{S} + \underline{S}) - Cl}$$

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

$$(10)\qquad Q(\underline{S}) = K_1 + K_2^*(\bar{S} - \underline{S}) + K_3^*(\bar{S} - \underline{S})^2.$$

Let $\mathcal{S} = \frac{\overline{S} + \underline{S}}{2}$ be an approximation of the average salinity level (before and after leaching). Substituting $\underline{S} = 2\mathcal{S} - \overline{S}$ in (10) yields:

$$Q(\mathcal{S}) = K_1 + K_2(\overline{S} - \mathcal{S}) + K_3(\overline{S} - \mathcal{S})^2$$

where $K_2 = 2K_2^*$, $K_3 = 4K_3^*$.

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 $C(\overline{S}, \mathcal{S})$, is:

$$(11) \qquad C(\overline{S}, \mathcal{S}) = P.Q(\mathcal{S}) = PK_1 + PK_2(\overline{S} - \mathcal{S}) + PK_3(\overline{S} - \mathcal{S})^2.$$

A profit function is defined as:

$$(12) \qquad \pi(\mathcal{S}) = \begin{cases} R_1(aS_0 + b_1) - R_2 - C(\overline{S}, \mathcal{S}) & \text{if } \mathcal{S} \leq S_0 \\ R_1(a\mathcal{S} + b_1) - R_2 - C(\overline{S}, \mathcal{S}) & \text{if } \mathcal{S} > S_0 \end{cases}$$

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 $\frac{\partial \pi(\mathcal{S})}{\partial \mathcal{S}}$ (for $\mathcal{S} > S_0$) to zero, $\mathcal{S}^*$ which maximizes (12) is accepted:

$$\mathcal{S}^* = \overline{S} + \frac{R_1 a + PK_2}{2PK_3}$$

Since it is obvious that $S^* \geq S_0$, it can be written:

$$\text{(13)} \qquad S^* = \max \left(S_0, \bar{S} + \frac{R_1 a + PK_2}{2PK_3}\right) .$$

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

$$\text{(14)} \qquad \hat{S}^* = \max \left(\hat{S}_0, \bar{S} + \frac{R_1 \hat{a} + PK_2}{2PK_3}\right)$$

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/l] (A somewhat high initial soil salinity level was chosen in order to emphasize the need of soil-leaching), N = 3500 [m^3/ha], γ = 0.7 (average irrigated-soil parameters in the study area), and Cl = 5 [meq Cl/l], the following estimates were obtained:

$K_1 = 132$

$K_2 = 526$ $\qquad K_2^* = 263$

$K_3 = 146$ $\qquad K_3^* = 36.5$

(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/m^3][5]; R_1 = 161 [dollars/tonne], the following values for potatoes were obtained:

$$\hat{S}^* = 15.8 \text{ [meq Cl/l]}$$

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

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:

$$\text{LOSS } (S^*, \hat{S}^*) = \pi(S^*) - \pi(\hat{S}^*) \quad (15)$$

[6]

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:

$$P_r(S^* > S_0 \text{ and } \hat{S}^* \leq \hat{S}_0) \rightarrow 0 \quad (16)$$

$$P_r(S^* \leq S_0 \text{ and } \hat{S}^* > \hat{S}_0) \rightarrow 0 \quad (17)$$

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:

$$\text{LOSS}\ (S^*,\ \hat{S}^*) = \{R_1 a\ (\bar{S} + \frac{R_1 a + PK_2}{2PK_3}) + R_1 b_1 - R_2$$

$$- P[K_1 + K_2\ (\bar{S} - \frac{R_1 a + PK_2}{2PK_3} - \bar{S}) + K_3\ (\bar{S} - \frac{R_1 a + PK_2}{2PK_3} - \bar{S})^2]\}$$

$$- \{R_1 a\ (\bar{S} + \frac{R_1 \hat{a} + PK_2}{2PK_3}) + R_1 b_1 - R_2$$

$$- P[K_1 + K_2\ (\bar{S} - \frac{R_1 \hat{a} + PK_2}{2PK_3} - \bar{S}) + K_3\ (\bar{S} - \frac{R_1 \hat{a} + PK_2}{2PK_3} - \bar{S})^2]\}$$

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

Similarly:

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

$$\text{LOSS}\ (S^*,\ \hat{S}^*)^7 = \frac{R_1^2\ (\hat{a} - a)^2}{4PK_3}$$

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

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

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

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

(d) $\hat{S}_0 > S_0,\quad S^* = S_0,\quad \hat{S}^* = \hat{S}_0$

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

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

$$= - (R_1 a + PK_2)(\hat{S}_0 - S_0) - PK_3[(\bar{S} - S_0)^2 - (\bar{S} - \hat{S}_0)^2]$$

Using indicator functions we write the loss function concisely as:

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

$$- PK_2(\hat{S}_0 - S_0)\}\ I_{\{S^* = S_0\}} - R_1 a(\hat{S}_0 - S_0)\ I_{\{S^* = S_0\}} \cdot I_{\{S_0 < \hat{S}_0\}}$$

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 $\hat{a}$, $\hat{S}_0$ are random variables, the loss function is also random.
For given values of S_0, a, σ_1^2, σ_2^2, and for a given scatter $\underset{\sim}{S}_T$ of the observations $S_1, \ldots, S_T$, the conditional expectation of the loss function is:

$$(19)\quad \overline{\text{LOSS}}\ (\sigma_1^2,\ \sigma_2^2,\ a,\ S_0,\ \underset{\sim}{S}_T) = E[\text{LOSS}\ (S^*,\ \hat{S}^*)/\sigma_1^2,\ \sigma_2^2,\ a,\ S_0,\ \underset{\sim}{S}_T] =$$

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

$$+ \{ - 2\ K_3\bar{S} \int_{-\infty}^{\infty} (\hat{S}_0 - S_0) dN(S_0,\ V(\hat{S}_0)\ ) + PK_3 \int_{-\infty}^{\infty} (\hat{S}_0^2 - S_0^2) dN(S_0,\ V(\hat{S}_0))$$

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

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

$$+ \{PK_3 V(\hat{S}_0) - \frac{R_1 a}{\sqrt{2\pi}} \cdot \sqrt{V(\hat{S}_0)}\ \}.\ I_{\{S^* = S_0\}}$$

Let $G(\sigma_1^2,\ \sigma_2^2,\ a,\ S_0)$ be the joint prior distribution of σ_1^2, σ_2^2, a, S_0. Assume that these four variables are independent, and their marginal distributions known[8] :

(i) $P_r(\sigma_1^2 = d_1) = P_r(\sigma_2^2 = d_2) = 1$

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

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

where d_1, d_2, δ_1, δ_2, Δ_1, Δ_2 are known.

Under these assumptions, the expectation of (19) will be:

(20) $$H(T, \underset{\sim}{S}_T) = E\{\overline{LOSS}\,(\sigma_1^2, \sigma_2^2, a, S_0, \underset{\sim}{S}_T)\}$$

$$= \frac{R_1^2}{4K_3P} \int_{\{S^* > S_0\}} V(\hat{a})\,dG(\sigma_1^2, \sigma_2^2, a, S_0) + PK_3 \int_{\{S^* = S_0\}} V(\hat{S}_0)\,dG(\sigma_1^2, \sigma_2^2, a, S_0)$$

$$- \frac{R_1}{\sqrt{2\pi}} \int_{\{S^* = S_0\}} a\sqrt{V(\hat{S}_0)}\;dG(\sigma_1^2, \sigma_2^2, a, S_0)$$

Let us now calculate the profitability of acquiring additional observations (S_i, Y_i). For a sample of a given size T let $\underset{\sim}{S}^*_T$ 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:

$$f_{\frac{1}{3},\frac{1}{3}}(S) = \begin{cases} S^{\frac{1}{3}}(1-S)^{\frac{1}{3}}\,\dfrac{(5/3)!}{(\frac{1}{3})!\,(\frac{1}{3})!} & \text{if } 0 \le S \le 1 \\ 0 & \text{otherwise} \end{cases}$$

This solution was adapted to our case, where the number of the observations is discrete, and further:

(i) $\sigma_1^2 = d_1 \neq \sigma_2^2 = d_2$

(ii) $S_0 \sim U[\Delta_1, \Delta_2]$:

$\frac{1}{3}\frac{d_1}{d_2}$ observations were taken in the vicinity of Δ_1, $\frac{2}{3} - \frac{1}{3}\ \frac{d_1}{d_2}$ in the vicinity of Δ_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 $H(T, \underset{\sim}{S}^*_T)$ describing the situation a priori, the expected value of additional information - to the potato growers of that region - from n observations with spread $\underset{\sim}{S}^*_{T+n}$, is:

$$\text{(21)} \qquad EVSI(n) = D[H(T, \underset{\sim}{S}^*_T) - H(T+n, \underset{\sim}{S}^*_{T+n})] - C_0(n) \ .$$

The optimal number of observations, n*, can be determined by:

$$\text{(22)} \qquad EVSI(n^*) = \max_n EVSI(n)$$

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, δ_1, δ_2, Δ_1 and Δ_2:

$$d_1 = 0.1 \simeq \hat{\sigma}_1^2 \; ; \qquad d_2 = 16 \simeq \hat{\sigma}_2^2$$

$$\delta_1 = -0.88 = \hat{a} + 0.21 \; ; \qquad \delta_2 = -1.30 = \hat{a} - 0.21$$

$$\Delta_1 = 4.5 \simeq \hat{S}_0 - 1.5 \; ; \qquad \Delta_2 = 7.5 \simeq \hat{S}_0 + 1.5$$

Accordingly, the marginal a priori distributions are:

$$P_r(\sigma_1^2 = 0.1) = 1 \; ; \qquad P_r(\sigma_2^2 = 16) = 1$$

$$a \sim U(-0.88, -1.30)$$

$$S_0 \sim U(4.5, 7.5)$$

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 $S_0 = 20(\bar{S})$. For P=0.1[dollars/m^3], R_1 = 161[dollars/tonne], D = 2000[ha] (the Northern Negev, where this experiment was conducted, is the main producer area of potatoes in Israel), and $C_0(n) = 260\ N$[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[9]. Since a

TABLE II

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

Three Alternative Observations Spreads (Dollars)

Number of Observations - n	3	6	9	12	15	18	21	24	27	30	33
EVSI(n), "Spread I"	14364	17185	22290	23658	23992	23898	22993	21058	19902	19600	19216
EVSI(n), "Spread II"	4610	8843	11277	12782	13614	14133	14398	14482	14424	14272	14032
EVSI(n), "Spread III"	6562	9149	11253	13550	14631	15467	16330	16738	17022	16742	16298

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

1. L. Berstein, Effects of salinity and soil water regime on crop fields, in "Salinity in Irrigation and Water Resources" (D. Yaron, ed.), Marcel Dekker, New York (1980).

2. P. I. Bickel and J. A. Yahav, Some contributions to the asymptotic theory of Bayes solution, Z. Wahrscheinlich Keitstheorie Verw. Geb. II, 257-276 (1969).

3. J.F. Bierhuizen, Water quality and yield depression, Inst. Land and Water Manag. Res., Wageningen, The Netherlands (1969).

4. E. Bresler, A model for tracing salt distribution in the soil profile and estimating the efficient combination of water quality and quantity under varying field conditions, Soil Science, 104, 227-233 (1967).

5. D. R. Davis and W. M. Dvoranchik, Evaluation of the worth of additional data, Water Res. Bul. 7, 700-707 (1971).

6. J. M. De Forges, Research on the utilization of saline water for irrigation in Tunisia, Nature and Resources 6, 2-6 (1970).

7. M. H. DeGroot, "Optimal Statistical Decisions," McGraw-Hill, Carnegie Mellon University, New York (1970).

8. E. Feinerman, "Economic Analysis of Irrigation with Saline Water Under Conditions of Uncertainty," Unpublished Ph.D. Dissertation (Mimeo in Hebrew), The Hebrew University of Jerusalem, Rehovot (1980).

9. D. A. Harville, Maximum likelihood approaches to variance component estimation and related problems, Journal of the American Statistical Association 72, 320-340 (1977).

10. V. Klemes, Value of information in reservoir optimization, Water Res. Res. 13, 837-850 (1977).

11. E. V. Maas and G. J. Hoffman, Crop salt tolerance - Current assessment, ASCE, J. of the Irrigation and Drainage Division IR2, 103, 115-134 (1977).

12. G. S. Maddala, "Econometrica," 176-181, McGraw-Hill, New York (1977).

13. T. Maddock, Management model as a tool for studying the worth of data, Water Res. Res. 9, 270-280 (1973).

14. A.K.H. Nouri, J. V. Drew, D. Knudson and A. Olson, Influence of soil salinity on production of dry matter and uptake and distribution of nutrients in barley and corn. I. Barley, Agronomy 62, 43-48 (1970).

15. A. Polovin, "Economic Analysis of Irrigation with Saline Water on a Kibbutz Farm," M.Sc. Thesis (Mimeo in Hebrew), The Hebrew University of Jerusalem (1974).

16. J. W. Pratt, H. Raiffa and R. Schlaifer, "Introduction to Statistical Decision Theory," McGraw-Hill, New York (1965).

17. R. E. Quandt, Tests of the hypothesis that a linear regression system obeys two separate regimes, Journal of the American Statistical Association 55, 324-340 (1960).

18. J. G. Ryan and R. K. Perrin, Fertilizer response information and income gains: The case of potatoes in Peru, Amer. J. Agr. Econ. 56, 337-343 (1974).

19. D. Sadan and M. Berglas, Personal Communication, Ministry of Agricultural, Extension Service, Negev (1980).

20. J. Shalhevet and L. Bernstein, Effects of vertically heterogeneous soil salinity on plant growth and water uptake, Soil Science 106, 85-93 (1968).

21. J. A. Yahav, Personal Communication, The Hebrew University of Jerusalem (1980).

22. D. Yaron, H. Bielorai, J. Shalhevet and Y. Gavish, Estimation procedures for response functions of crops to soil water content and salinity, Water Res. Res. 8, 291-300 (1972).

23. D. Yaron and A. Olian, Application of dynamic programming in Markov chains to the evaluation of water quality in irrigation, Amer. J. Agr. Econ. 55, 467-471 (1973).

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 ($\hat{S}^* \leq S_0$, $S^* > S_0$). 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.

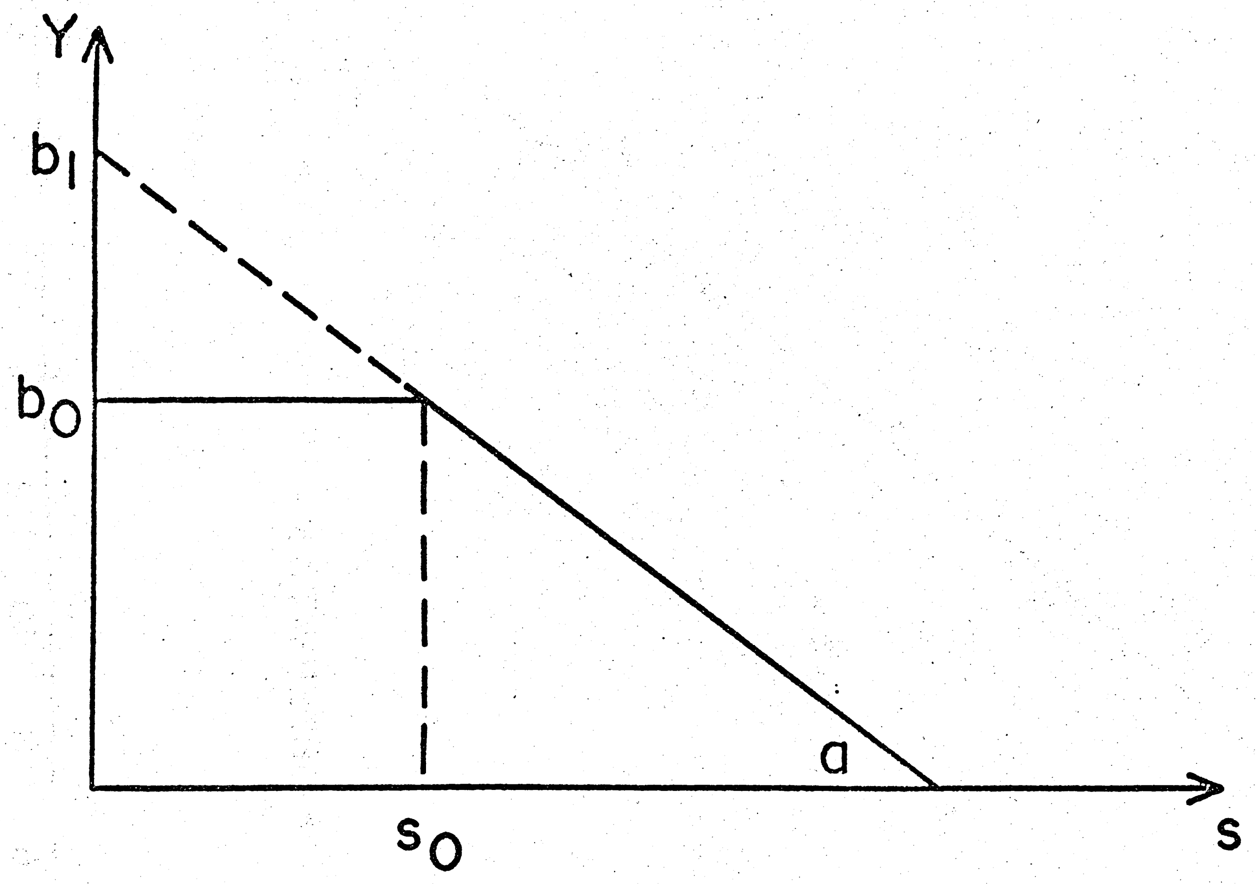
[9]The optimal EVSI (23992 dollars/region) found in Table 2 constitute less than one percent of the total revenue of the potato growers.

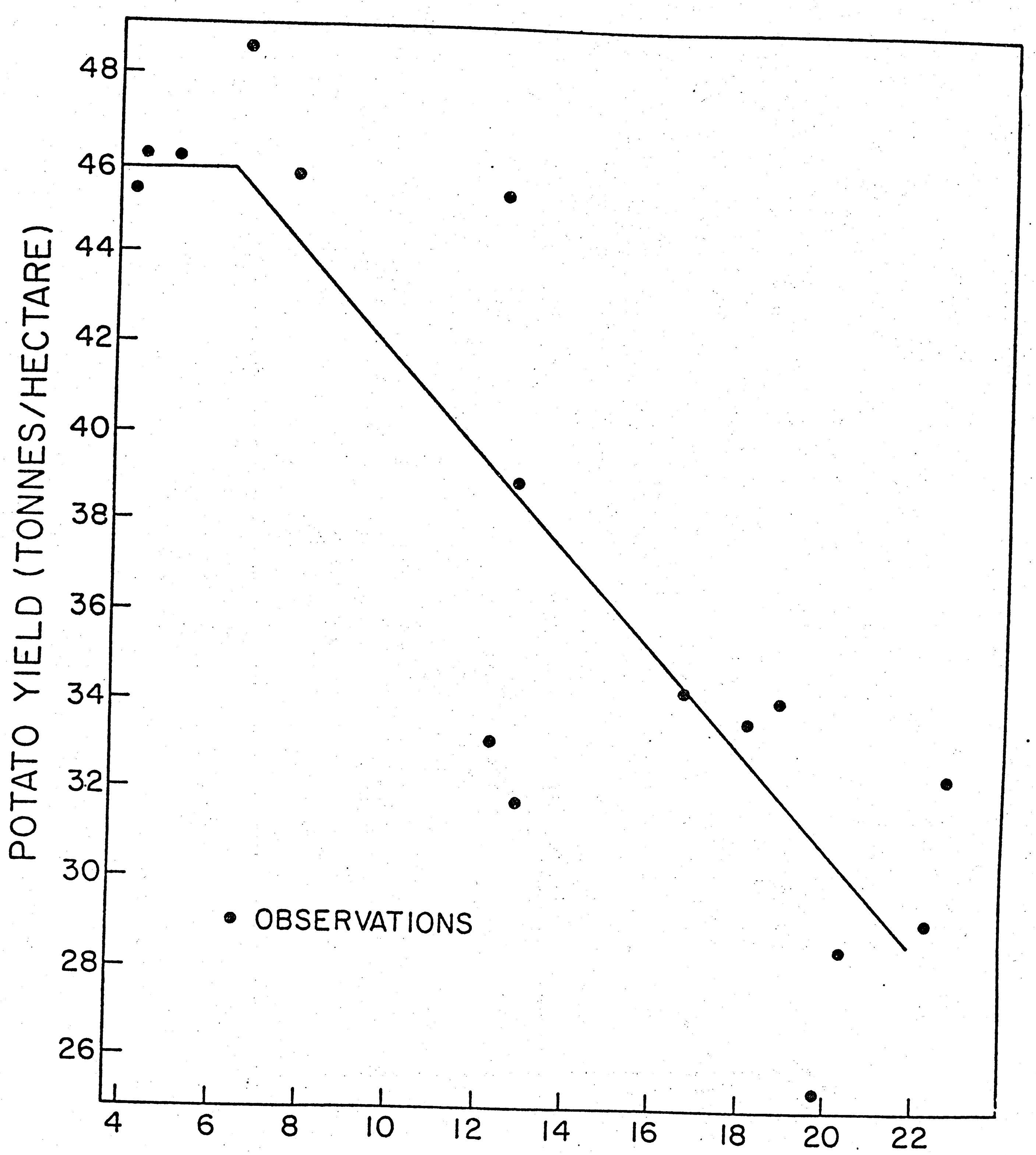
FIGURE LEGENDS

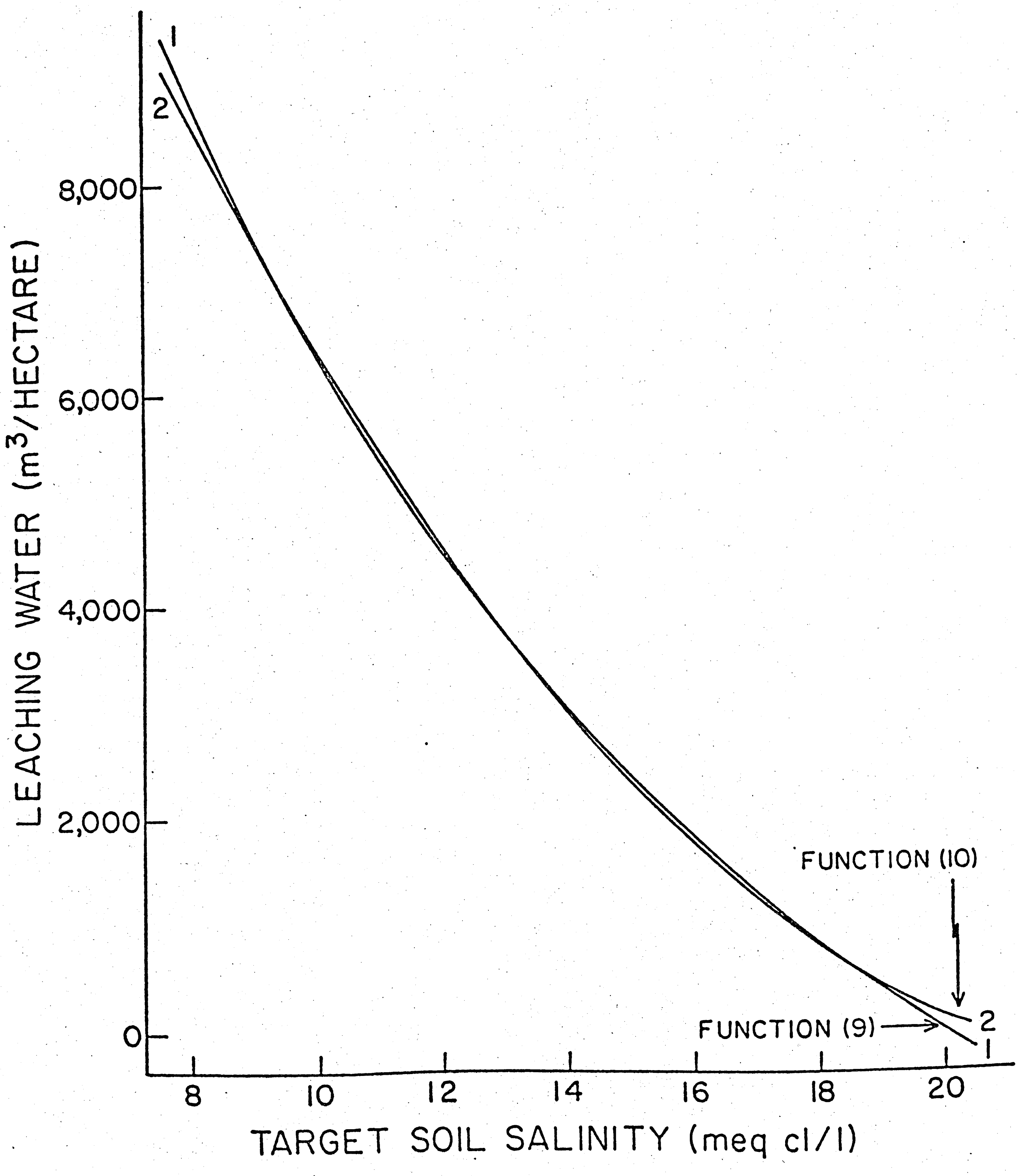
Figure 1: The Response Function.

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 $\bar{S}$ = 20 meq Cl/l).


Y
b1
b0
a
s0
s


POTATO YIELD (TONNES/HECTARE)
48
46
44
42
40
38
36
34
32
30
28
26
OBSERVATIONS
4
6
8
10
12
14
16
18
20
22
SOIL SALINITY (meq cl/l)


LEACHING WATER (m3/HECTARE)
8,000
6,000
4,000
2,000
0
1
2
FUNCTION (10)
FUNCTION (9)
2
1
8
10
12
14
16
18
20
TARGET SOIL SALINITY (meq cl/l)

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).

7002 Yakir Plessner - Computing Equilibrium Solutions for Various Market Structures.

7003 Yoav Kislev and Yeshayahu Nun - Economic Analysis of Flood Control Projects in the Hula Valley, Stage One - Final Report. (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 Meri G. Kohn - Monopolistic Behavior in Situations of Expectation Motivated Demand.

7103 Yakir Plessner and Meir G. Kohn - A Model of Optimal Marketing Policy.

7104 Yoav Kislev and Yakir Plessner - An Applicable Linear Programming Model of Inter-Temporal Equilibrium.

7105 Aharon Ben-Tal and Eitan Hochman - Bounds on the Expectation of a Convex Function of a Random Variable with Applications to Decision Making Under Uncertainty.

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.

7204 Yoav Kislev and Benjamin Nadel - Economic Analysis of Flood Control Project in the Hula Basin. (Hebrew).

7301 Yigal Danin and Yair Mundlak - The Effect of Capital Accumulation on a Well Behaved n-Sector Economy.

7302 Pinhas Zusman - Power Measurement in Economic Models.

7303 Aba Schwartz, Uri Regev and Shmuel Goldman - Estimation of Production Functions Free of Aggregation Bias with an Application to the Israeli Agriculture.

7401 Yakir Plessner - <u>A Theory of the Dynamic Competitive Firm under Uncertainty</u>.

7402 Robert E. Evenson and Yoav Kislev - <u>A Stochastic Model of Applied Research</u>.

7501 Meir G. Kohn - <u>Competitive Speculation</u>.

7601 Yoav Kislev and Uri Rabiner - <u>Animal Breeding -- A Case Study of Applied Research</u>.

7602 Jack Habib, Meir Kohn and Robert Lerman - <u>The Effect on Poverty Status in Israel of Considering Wealth and Variability of Income</u>.

7701 Yoav Kislev, Michal Meisels, Shmuel Amir - <u>The Dairy Industry of Israel</u>.

7702 Yair Mundlak - <u>Agricultural Growth in the Context of Economic Growth</u>.

7703 Meir Kohn - <u>Beyond Regression: A Guide to Conditional Probability Models in Econometrics</u>.

7801 Yair Mundlak - <u>Models with Variable Coefficients - Integration and Extension</u>.

7802 Yigal Danin and Meir G. Kohn - <u>An Analysis of the Israeli Grain Market and Purchasing Policy</u>.

7803 Yoav Kislev - <u>The Monetary Approach to the Israeli Balance of Payments</u>.

7804 Meir Kohn - <u>A Theory of Innovative Investment</u>.

7805 Yair Mundlak and Joseph Yahav - <u>ANOVA, Convolution and Separation, A Fresh View at Old Problems</u>.

7806 Meir Kohn - <u>Why the Dynamic Competitive Producer Should Not Carry Stocks of his Product</u>.

7901 Yair Mundlak - <u>Agricultural Growth - Formulation, Evaluation and Policy Consequences</u>.

7902 Dan Yaron, A. Dinar and S. Shamlah - <u>First Estimates of Prospective Income Losses Due to Increasing Salinity of Irrigation Water in the South and the Negev Regions of Israel</u>. (Hebrew).

7903 Yair Mundlak - <u>On the Concept of Non-Significant Functions and its Implications for Regression Analysis</u>.

7904 Pinhas Zusman and Michael Etgar - <u>The Marketing Channel as an Equilibrium Set of Contracts</u>.

7905 Yakir Plessner and Shlomo Yitzhaki - <u>The Firm's Employment Policy as a Function of Labor Cost Structure</u>.

7906 Yoav Kislev - <u>Management, Risk and Competitive Equilibrium</u>.

7907 Yigal Danin and Yair Mundlak - <u>The Introduction of New Techniques and Capital Accumulation</u>.

7908 Yair Mundlak - <u>Elements of a Pure Theory of Forecasting and the "After Keynesian Macroeconometrics</u>".

8001 Yoav Kislev and Willis Peterson - <u>Prices, Technology and Farm Size</u>.

8002 David Bigman and Haim Shalit - <u>Applied Welfare Analysis for a Consumer Whose Income is in Commodities</u>.

8003 David Bigman - <u>Semi-Rational Expectations and Exchange Rate Dynamics</u>.

8004 Joel M. Guttman - <u>Can Political Entrepreneurs Solve the Free-Rider Problem?</u>

8005 Yakir Plessner and Haim Shalit - Investment and the Rate of Interest Under Inflation: Analysis of the Loanable Funds Market.

8006 Haim Shalit - Who Should Pay for Price Stabilization?

8007 David Bigman - Stabilization and Welfare with Trade, Variable Levies and Internal Price Policies.

8008 Haim Shalit, Andrew Schmitz and David Zilberman - Uncertainty, Instability and the Competitive Firm.

8009 David Bigman - Buffer Stocks and Domestic Price Policies.

8101 David Bigman - National Food Policies in Developing Countries: The Experience and the Lesson.

8102 David Bigman - The Theory of Commodity Price Stabilization and Buffer Stocks Operation: A Survey Article.

8103 Yoav Kislev and Willis Peterson - Induced Innovations and Farm Mechanization.

8104 Yoav Kislev and Yakir Plessner - Recent Inflationary Experience in Israel.

8105 Yair Mundlak - Cross Country Comparison of Agricultural Productivity.

8106 Michael Etgar & Ilan Peretz - The Preference of the German Market for Quality Tomatoes (Hebrew).

8107 Tzvi Sinai - The Profitability of Land Development for Agriculture in Israel (Hebrew).

8108 Ilan Beeri - Economic Aspects of the Settlement Project in Yamit (Hebrew).

8119 David Bigman - Stabilization and International Trade.

8110 Nava Haruvi and Yoav Kislev - Cooperation in the Moshav.

8111 Michal Meisels-Reis - Specialization and Efficient in the Poultry Industry in Israel (Hebrew).

8112 Joel M. Guttman - Matching Behavior and Collective Action: Theory and Experiments.

8113 Yair Mundlak - Various Aspects of the Profitability of Milk Production. (Hebrew)

8114 Yair Mundlak & Joseph Yahav - Inference with Stochastic Regressors.

8201 Pinhas Zusman & Clive Bell - The Equilibrium Set of Dyadic Contracts.

8202 Yoav Kislev & Shlomit Farbstein - Capital Intensity and Product Composition in the Kibbutz and the Moshav in Israel.

8203 David Bigman - Food Aid and Food Distribution.

8204 Haim Shalit and Shlomo Yitzhaki - Mean-Gini, Portfolio Theory and the Pricing of Risky Assets.

8205 Rafi Melnick & Haim Shalit - The Market for Tomatoes: An Empirical Analysis. (hebrew)

8206 Dan Yaron & Hillary Voet - Optimal Irrigation With Dual Quality (Salinity) Water Supply and the Value of Information.

8207 David Bigman & Itzhak Weksler - Strategies for Emergency Stock Planning.