ELSEVIER

Agricultural Economics 26 (2001) 227–236

AGRICULTURAL ECONOMICS

www.elsevier.com/locate/agecon

# Sense and sustainability revisited: the limits of total factor productivity measures of sustainable agricultural systems

Derek Byerlee [a,*], Rinku Murgai [b]

[a] *Rural Development Department, World Bank, 1818 H Street NW, Washington, DC 20433, USA*

[b] *Development Research Group World Bank, 1818 H Street NW, Washington, DC 20433, USA*

Received 5 November 1999; received in revised form 15 September 2000; accepted 15 October 2000

**Abstract**

Many economists have advocated and applied total social factor productivity (TSFP) (i.e., total factor productivity estimated with both market and non-market inputs and externalities, and with all factors valued at social prices) as a single all-embracing measure of agricultural sustainability. This paper reviews the conceptual and practical issues in measuring TSFP and shows that no one measure alone will be theoretically or empirically robust as an indicator of sustainability. TSFP is a conceptually flawed measure since inclusion of non-market inputs and outputs and social price-based valuation, in most cases, violates the theoretical basis underlying those estimates. Trends in TSFP also have limited value in diagnosing the nature of sustainability problems, unless changes in productivity are related to underlying changes in technology, human and physical infrastructure, and indicators of resource quality.

More attention needs to be given to defining key indicators of agro-ecosystem health and relating these measures to trends in productivity. This analysis must be sufficiently disaggregated and for a long enough time period to allow for spatial and temporal variability inherent in agricultural production. Secondary data at the district level on both conventional inputs and outputs and resource quality have recently allowed more quantitative estimates of sustainability and its causes. With limited data, yield growth decomposition analysis can often be used to provide valuable insights into sustainability problems. Meanwhile, there is a need to invest in long-term experimental and panel surveys of farmers and their fields for key production systems in order to provide long-term data that will allow full productivity accounting, using more formal statistical procedures. Regardless of the approach selected, the findings of this paper strongly suggest a need for economists, agronomists and soil scientists to collaborate in integrating approaches in order to provide more robust and informative measures of sustainability.


*Keywords:* Agricultural sustainability; Total factor productivity; Resource degradation; Resource quality indicators

## 1. Introduction

Since sustainable agriculture became the watch word for encapturing society's desire to better preserve the natural resource base for future generations, there have been debates about how to define and measure sustainable agricultural systems. It is now widely agreed that there are different dimensions of sustainability ranging from the biophysical dimensions to economic and social dimensions (Herdt and Lynam, 1992). The biophysical dimensions of sustainability relate to the long-term maintenance or enhancement of the productive capacity of the resource base. Economic and social dimensions relate to the

* Corresponding author. Tel.: +1-202-458-7287;
fax: +1-202-614-0065.
*E-mail address:* dbyerlee@worldbank.org (D. Byerlee).

0169-5150/01/$ – see front matter 
PII: S0169-5150(00)00131-6

long-term economic viability of farming and rural communities.

Naturally, biological and physical scientists have tended to focus on biophysical measures of sustainability, such as crop yields on the output side, and indicators of soil and water quality on the input side. However, there are serious limitations to these measures of biophysical sustainability since yields must be interpreted in relation to input use, and indicators of resource quality must be related to productivity. Despite these difficulties, a considerable body of work has been developed by crop and soil scientists to define and measure indicators of resource quality as a basis for tracking the sustainability of agricultural systems (e.g., Barnett et al., 1995; Pieri et al., 1995).

In order to develop a single unambiguous measure of biophysical sustainability, economists have proposed the use of an index of total factor productivity (TFP), since it explicitly accounts for changes in agricultural production in relation to changes in inputs. Lynam and Herdt (1989, p. 385) in an influential article, "Sense and sustainability" in this Journal, first proposed that:

> ... the appropriate measure of output by which to determine sustainability ... is total factor productivity ...; a sustainable system has a non-negative trend in total factor productivity over the period of concern.

A non-negative TFP growth implies that output is increasing at least as fast as inputs. A negative trend in TFP strongly implies that the quality of the resource base is being degraded although it says nothing about the causes of that degradation.

TFP as conventionally measured by economists does not take account of non-market outputs and inputs, especially long-term resource degradation such as soil loss or nutrient mining, and environmental externalities (such as water pollution). Therefore, the definition of sustainability was soon modified to a more inclusive measure of total social factor productivity (TSFP), in which changes in non-market inputs and outputs, and externalities are explicitly incorporated (Herdt and Lynam, 1992; Crosson and Anderson, 1993). Expanding the criteria, Herdt and Steiner (1995) later argued that such a non-negative trend in TSFP must be achieved within acceptable limits of indicators of agro-ecosystem health for an economic system to be sustainable. While TSFP is intuitively appealing, the problems of measuring and valuing non-market inputs and outputs, such as resource degradation and environmental pollution, are formidable and as we show in this paper, the TSFP measure itself is conceptually flawed. [1]

Despite these difficulties, a non-negative trend in TFP or TSFP has been fairly widely accepted among economists, and many agronomists, as the measure of a sustainable agricultural system. Many recent studies have used TFP to analyze the sustainability of specific agricultural systems (Ali and Byerlee, 2000), or crops (Sidhu and Byerlee, 1992; Cassman and Pingali, 1995). A few studies have attempted to measure TSFP, usually using data from experimental plots (e.g., Ehui and Spencer, 1993; Whitaker, 1993), including one compendium based on long-term experiments from a number of sites (Barnett et al., 1995).

The purpose of this paper is to provide a conceptual and practical critique of the TFP and TSFP measures of sustainability and propose modifications of these approaches that might be applied to diagnose the presence or absence of a sustainability problem, as well as to understand the causes of the problem, if it exists. [2] We conclude that although TSFP is a conceptually flawed measure of sustainability, the trend in TFP, as conventionally defined, can be a useful starting point for measuring sustainability but it has little value unless it is interpreted in relation to trends in resource quality. That is, we advocate some combination of two of the main schools of thought in the debate on measures of sustainability of an agricultural system — those who focus on indicators of resource quality, and those who emphasize productivity measures. In particular, we argue that the main challenge now is to develop useful and cost-effective indicators to monitor long-term changes in resource quality and ways to relate these to changes in productivity.

[1] One of our referees has also pointed out that since agriculture is only one component of an ecosystem, agricultural sustainability cannot be used to judge environmental (resource) sustainability, and vice versa. While we accept this as an issue in multiple use systems, this paper largely relates to intensive agricultural systems in the developing world where agriculture is by far the dominant component of the ecosystem.

[2] A third level of analysis to evaluate interventions to improve sustainability is not treated in this paper.

## 2. Limitations of TSFP as a measure of sustainability

It has been argued that the use of TSFP as a measure of sustainability requires two types of changes in conventionally measured TFP. First, it is proposed that TFP measures should be broadened to include non-marketed inputs and outputs. Joint outputs or externalities of agricultural production such as soil erosion, depletion of soil fertility, and groundwater aquifer mining are usually not taken into account in productivity measurement yet are clearly central to tracking agricultural sustainability. To address these problems, several researchers have argued that changes in the quality of the natural resource base (including externalities) should be accounted for in measures of agricultural productivity indices for assessing sustainability (Antle and McGuckin, 1993; Harrington et al., 1994; Alston et al., 1995a; Herdt and Steiner, 1995; Repetto et al., 1997). The recent debate also suggests that costs of environmental regulation and abatement, as well as benefits of improved environmental quality, be included in TSFP (Repetto et al., 1997; Gollop and Swinand, 1998). Second, TSFP requires that both market and non-marketed inputs and outputs be valued at long-term *economic* prices (Herdt and Lynam, 1992). For example, Crosson and Anderson (1993) argue that non-marketed and therefore unpriced effects should be valued at their marginal *social* values that include impacts of production on the agricultural natural resource base and environmental consequences that reflect changes in the quantity or quality of resources not directly used in agriculture (e.g., value of wildlife habitat). In the remainder of this section, we note critical conceptual and practical difficulties with TSFP, as defined, as a measure of sustainability.

### 2.1. Conceptual issues

TFP growth measures the residual growth in output after controlling for the weighted growth in input use, where the weights are provided by the elasticity of output with respect to inputs. If we assume that producers maximize profits, output markets are competitive, and the production technology is characterized by constant returns to scale, the elasticity of output with respect to each input is equal to its share in total cost. Then, TFP growth can be estimated using observed input and output quantities and prices.

As noted earlier, TSFP modifies TFP by expanding the input indices to include non-marketed inputs and outputs, and weights the observed input quantities with their respective cost shares using social prices. Whether or not this is correct depends on the underlying behavioral assumptions and the nature of the production technology. If producers are profit-maximizers (or cost-minimizers), non-market factors such as resource stocks which are beyond the control of the producer but which affect the production environment, should be treated as a technological constraint rather than as conventional inputs in TFP estimates (McFadden, 1978; Squires, 1992). The resulting growth residual should be explained by both technical change and changes in resource stocks.

**Whether or not resource quality is included in TFP calculations also depends on the nature of the production technology. If water quality, e.g., is modeled as a joint output, it should be included.** If it is a pure externality that does not affect the system being evaluated (e.g., water quality for consumers), it should not enter the calculations although private regulatory costs may be included as an input into production. [3]

If instead, the behavioral assumption underlying TSFP is that of a central planner who faces social prices, for appropriate growth accounting, we need to estimate what levels of conventional and non-market inputs would have been chosen by the planner, and then use these estimated input quantities in the TSFP input index. *Observed* input quantities (i.e., those under profit maximization or cost minimization) only reflect optimal choices made by producers who ignore externalities and face private, not social, prices for conventional inputs.

For the same reason, observed input and output choices should be valued by private, rather than long-term economic or social prices. The role of valuation in TFP measures is to measure the gradients of

[3] In industrial contexts, regulatory costs have been incorporated into productivity measures. For example, Pittman (1983) used the costs of compliance with US pollution-control regulations as shadow prices with which to value pollutants (negatively, in the output index) in the paper industry. However, regulatory costs are less common in the agricultural sector, particularly in developing countries.

the technology (elasticities) at the observed producer output and input choices. However, in the absence of markets or in the presence of price distortions, the gradients will reflect only the producer's valuation, not society's, of the products (Weaver, 1998). Therefore, valuing production at the societal shadow prices of goods, or by consumer prices (e.g., Gollop and Swinand, 1998) rather than producer prices is theoretically incorrect. For example, in the case of overexploitation of a joint access resource such as groundwater, where the cost of using the resource is negligible to the individual — however large it might be for society or whatever the long-term implications for sustainability — the input should be valued at a negligible price in the input index.

In summary, in many instances, modifying TFP to a TSFP measure is likely to be conceptually flawed. Alternative approaches need to be developed for considering social costs and benefits when assessing sustainability and designing interventions. We turn to these in a later section in the paper.

### 2.2. Practical issues in using the TSFP approach

Aside from the above theoretical difficulties with using a TFP-based measure of sustainability, there are a number of practical issues in its application.

#### 2.2.1. Level of aggregation

Traditionally TFP has been measured at the national or state level since input and output data are most readily available for politically defined regions. This is appropriate since the main objective of these studies has been to assess sources of overall sector growth. However, TFP trends at this level are a blunt instrument for identifying particular production systems and regions with potential sustainability problems, especially given that such problems tend to be location specific.

By contrast, several recent studies have focused on estimates of TFP at the plot level using data from long-term experiments as a measure of sustainability. The problem with this approach is how to relate these trends to those in farmers' fields. Physical conditions often differ substantially between farmers' fields and plots on experimental stations, and more importantly, fixed experimental treatments do not allow for farmers' substitution among inputs and changing rotations in response to resource degradation.

As Lynam and Herdt (1989) note, there is no easy answer to this dilemma. They recommend that the appropriate level of analysis of sustainability is at the cropping or farming systems level defined in terms of a relatively homogeneous agro-ecological resource base that leads to similar choices of crop and livestock activities and inputs. Cropping or farming systems are a more powerful principle for classification than other scales or spatial dimensions since the differences across farming systems will tell us much about the critical variables in long-term sustainability (ICRISAT, 1998). Variation in cropping systems are likely to reflect differences in factors such as the types of new technologies that are available, investment in infrastructure, agro-ecosystem health, and particularly important for sustainability analysis, in the pressures exerted on the stock of natural resources through intensification of dominant enterprises in the system.

Where panel farm-level data are available, these can sometimes be aggregated at the system level. For example, Cassman and Pingali (1995) use panel farm-level data to estimate TFP for intensive irrigated rice systems in the Philippines. However, such data sets are rare, although increasing, in the developing world. A more likely alternative is to construct secondary data sets at the district level and then aggregate districts on the basis of the dominant farming systems. For example, Murgai (1999) and Ali and Byerlee (2000) aggregate district-level data by cropping systems to estimate trends in productivity growth during the green revolution and post-green revolution periods in the Indian and Pakistan Punjab, respectively. Significantly different patterns of TFP growth across cropping systems in both studies underscore the importance of disaggregating state-level or national estimates to identifiable production systems.

#### 2.2.2. Time period of analysis

A sufficient time period of analysis presents a problem in assessing sustainability. As Monteith (1990) has shown, there is the problem of defining the necessary number of years to estimate a trend with some degree of statistical confidence. In a variable rainfed environment with a low growth rate in TFP, the number of years required to estimate a statistically valid trend may be as high as 30 years. Even in irrigated areas, trends in TFP may be very sensitive to the presence of good or bad years at the beginning or end of the trend period, as shown by Tiongco and Dawe (2000).

This problem is compounded by the fact that, in practice, some systems have undergone several stages of technical change in a short period. In Asia, for example, the green revolution propelled irrigated systems from very low external input use to high input use in a period of two decades. Since sustainability is likely to relate to the particular stage of technical change, it may be difficult to detect underlying trends before the system evolves into a new stage of change. Thus, over a relatively long period of time, a system experiencing rapid technical change may appear to be sustainable as measured by trends in TFP, even though underlying trends in resource quality suggest that such a system is not sustainable (see, e.g., Duff et al., 1995). By the time that these resource quality problems are reflected in TFP trends, resource degradation may be very serious and even irreversible (e.g., some types of soil salinity).

### 2.2.3. *Confounding of labor-saving and land-saving changes*

A further difficulty in using a TFP-based measure of sustainability is that trends in TFP are due to both land-saving and labor-saving technical change. In practice, the main interest in estimating TFP to assess sustainability is to explore the balance between land-saving technical change (positive effect on TFP) and land degradation (negative effect on TFP). A situation where land-saving technical change is able to compensate for resource degradation over a sufficiently long period of time provides confidence that a system is sustainable, even if there is evidence of resource degradation. However, increasingly in the developing world, trends in TFP are driven by labor-saving technical change, especially where mechanization is proceeding rapidly, as in much of Asia. Labor-saving technical change reflected in positive TFP growth may disguise serious problems of resource degradation that do not become apparent until after the period of rapid labor-saving technical change has been completed. For example, Traxler et al. (1995) found that harvest mechanization was the main driving force in TFP growth in cotton in the post-war period in US, while other evidence shows that this was a period of considerable resource degradation. In the Pakistan's Punjab, labor productivity has increased faster than land productivity in the post-green revolution period, allowing positive TFP growth in most systems despite evidence of serious resource degradation (Ali and Byerlee, 2000).

### 2.2.4. *Measurement and valuation issues in estimating TSFP*

There are many practical problems in measuring and valuing non-market inputs and outputs related to resource degradation, even where it is theoretically appropriate to include them. While it is conceptually straightforward, although often costly, to quantify non-market inputs and externalities, the valuation of these inputs and outputs is challenging. Even in valuing *private* shadow prices for on-site effects, the analyst must decide whether to use foregone output, replacement costs, user costs, or option and existence value to measure the costs of resource degradation (Harrington et al., 1994).

As we have discussed above, it is theoretically valid to include some on-site non-market effects in estimates of TFP, where these are reflected in farmers' decision making. The most relevant case for agriculture is the accounting for soil nutrient mining and the on-site cost of erosion. Changes in on-site soil nutrient stocks are a valid candidate for inclusion in a TFP index since these are probably internalized by farmers in decision making. Measurement and valuation of changes in nutrient stocks is relatively straightforward but entails intensive data collection. In some cases, the cost of soil nutrient 'mining' can be approximated by its replacement cost, i.e., the market price of chemical or organic fertilizers needed to fully replenish the nutrient. For example, using experimental data and plot soil tests, Ehui and Spencer (1993) compute TSFP for four production systems in Nigeria in which nutrients applied and extracted had been carefully recorded. The results illustrate that a system with apparent productivity growth but large depletion of resources is in fact unlikely to be sustainable when proper account is taken of the cost of replacing the nutrients extracted. Likewise attempts have been made to value soil erosion effects by estimating erosion losses and output foregone using simulation models of soil run-off and crop productivity (Steiner et al., 1995).

Several studies have tried to account for externalities arising from use of agricultural chemicals in TFP measures, but these efforts have often been very crude and have also made little difference to the overall conclusions (Repetto et al., 1997, and various chapters in Barnett et al., 1995). Steiner et al. (1995) classify externality costs of pesticide and fertilizer use into regulatory costs, health-related costs, and environmental costs. Among these, costs of compliance with government regulations and health costs to farmers of sustained contact with pesticides are largely borne by the farmer and could be included in a TFP index. However, other costs that do not influence the farmer's production choices (either because the perceived cost is zero or because the farmer is not aware of the cost) should not be included. As a result, unpriced environmental costs such as wildlife losses and water pollution — notwithstanding their important social costs — are clearly misplaced in a TFP index. In addition, valuation of these externalities in the existing literature has often been estimated as an arbitrary share of private costs of using the input (e.g., a mark up of 20–100% on market prices) (e.g., most of the case studies in Barnett et al., 1995).

## 3. Toward a set of indicators on productivity and resource quality

The main conclusions from the above discussion are as follows. First, TSFP is a conceptually flawed measure of sustainability. Inclusion of non-market inputs and outputs and use of social prices for valuing inputs in TFP estimates in many and probably most cases, violates the assumptions underlying those estimates. Second, attempts to value non-market inputs and outputs have often been meaningless, given the arbitrary assumptions that have been employed. Finally, trends in TFP have limited value in diagnosing the nature of sustainability problems, without relating changes in TFP to underlying changes in technology, human and physical infrastructure, and especially indicators of resource quality.

Thus a positive trend in TFP (conventionally measured) is not a good indicator of sustainability, except over the very long run in a situation in which externalities are minimal. Nevertheless, a statistically robust negative trend in TFP over a period of say a decade or more, is very likely a good indicator of sustainability problems related to underlying resource degradation. As recognized by Lynam and Herdt (1989), however, the finding of negative TFP trends is of little value in decision making, unless it can be interpreted in light of data on underlying trends in resource quality that help identify the factors causing the negative TFP trend and provide guidance on possible measures to mitigate or reverse the problem.

In sum, the recent emphasis in the economics literature of searching for an all embracing single measure of sustainability in the form of TSFP has not been helpful. Rather attention in diagnosing sustainability should now turn to doing a better job of measuring productivity and trends in resource quality and relating the two.

### *3.1. Decomposing trends in productivity growth*

We suggest that diagnosis and understanding of sustainability should be conducted within the framework of a production function that integrates economic, agronomic and resource quality variables. Such a function would include conventional inputs (land, labor, etc.), non-conventional inputs (education, infrastructure, etc.), technology variables (e.g., use of improved varieties), resource degradation variables (e.g., soil erosion, nutrient status, etc.), and weather variables (to reduce weather-induced variability and allow trends to be picked up with a shorter period of data). [4] Inclusion of resource quality variables in the production function is also an advantage in that it allows valuation of resource degradation through the estimated marginal product. [5] Conceptually, and perhaps empirically, outputs could be modeled as a joint production process of agricultural outputs and environmental externalities (Jaenicke and Lengnick, 1999).

Ideally, this production function would be estimated for the system of interest using time series information. In practice, time series data are rarely long enough or complete enough to allow econometric estimation.

[4] Technology variables could be incorporated as a knowledge stock (e.g., investment in research) but because of long and uncertain lags (Alston et al., 1995b), direct inclusion of technology variables is usually preferable if the main interest is system sustainability.

[5] Either a production, profit or cost function can be employed to estimate effects of changes in resource quality on productivity.

A compromise is to use panel data to capture both cross-sectional and time series variability. Secondary panel data for many of the variables are increasingly available at the district level, especially production and input data. Several recent studies have also included district-wise data on resource quality — Huang and Rozelle (1995) and Lindert (1999) for China, and Ali and Byerlee (2000) for Pakistan. The latter two studies are unique in using direct estimates of soil quality from soil sampling (and water quality in the case of Pakistan). All three studies show significant effects of resource degradation on productivity.

In a more advanced approach, resource quality would be modeled endogenously as a function of past productivity and management practices (Byerlee et al., 1991). This approach was used by Lindert (1999) to analyze trends in soil quality and agricultural productivity in China. Soil quality was modeled as a function of crop yields and cropping patterns in previous years and several exogenous factors. As might be expected, soil quality was positively enhanced by shifts to enterprises such as legumes and livestock.

More in-depth analysis will require primary panel data at the farm or plot level. Panel data on rice farmers collected by IRRI over a period of 30 years have been analyzed to test for resource degradation as indicated by a continuous shift downward in the production function (Cassman and Pingali, 1995). However, this data set does not include indicators of resource quality that would identify the causes of that degradation. At the plot level, there are a number of data sets that have been generated by long-term experiments that are good candidates for production function analysis (Barnett et al., 1995). These data sets often include detailed information on resource quality, but are often constrained by information on key inputs, especially labor.

Meanwhile, a less data intensive but useful approach is for agronomists and economists to collaborate in a yield-accounting analysis. In this approach, available agronomic response information (e.g., fertilizer response) and information on farmers' practices and yields at a minimum of two points in time, is used to decompose actual yield growth into various sources — genetic, input use, changes in management practices, and changes in soil parameters. The approach is more qualitative but much cheaper, less data intensive, and in the few cases in which it has been applied, has provided valuable insights on system sustainability.

The best example of such an approach is provided by Cardwell (1982) for maize in Minnesota over a 50-year period. Cardwell was able to include factors leading to negative yield effects such as the decline in soil organic matter. Other recent examples include Bell et al. (1995) for wheat in northwest Mexico, and Byerlee and Siddiq (1994) for wheat in Punjab, Pakistan. In the latter case, the yield accounting left a significant negative residual that was attributed to resource degradation, but the data available did not allow an identification of the precise yield-reducing factors.

The major missing element in the above approaches is the failure to account for off-site externalities. In those systems where specific externalities are hypothesized to be important, they need to be quantified. Externalities might enter the production function as a joint output, although this still begs the problem of valuation. Alternatively, the cost of the externality can be 'normalized' by comparing the cost of externalities produced with the value of productivity gains realized (but not by the computation of a single all-inclusive index). Knowledge of such externalities, including social costs, will be especially important in evaluating interventions to enhance system sustainability.

### 3.2. Minimum data sets for monitoring sustainability

The major challenge to implementing the above framework is the lack of comprehensive data sets to analyze sustainability of key production systems. Panel input–output data on production are increasingly becoming available at the plot level (i.e., long-term experiments), farm level (i.e., panel farm-household and plot survey data), and district level (i.e., secondary statistics). However, except for the district level, most of these data sets are limited by the relatively short-time period over which the data have been collected. Both international and national research institutes are now recognizing the value of long-term experiments and many have been initiated in the past few years. However, analysis of some of these results after only 3–5 years is probably premature (e.g., Whitaker, 1993; Ehui and Spencer, 1993). International agricultural research centers have been in the forefront in setting up farm-household and farm plot panels (such as the ICRISAT village surveys, the IRRI rice farmer surveys, and CIMMYT's wheat field panel in northwest Mexico, and wheat–rice field panel in the Terai of Nepal). However, even some of these panel data sets have not been sustained over time (e.g., the ICRISAT data sets have not been updated for several years). [6]

In the meantime, district-level secondary data offer the best opportunity to analyze long-term sustainability provided more emphasis is given to finding innovative means to track resource quality. For example, in Pakistan, soil and water test data have been collected continuously for over 30 years in response to demands from farmers and extensionists but had not been aggregated and made available. While not a random sample, these data did seem to explain important trends in productivity of Punjab agriculture (Ali and Byerlee, 2000). Similar data for parts of China have also provided new insights into long-term trends in soil quality (Lindert, 1999). Likewise, Stoorvogel and Smaling (1990) have demonstrated the value of computing nutrient balances at the system or even national level, based on secondary data on nutrients applied, crop yields and nutrient content, livestock numbers and crop residue management. Measures of genetic diversity in farmers' fields have recently been developed which may be proxies for potential pest and disease losses (Smale et al., 1998).

Simulation models offer much potential to infer trends in resource quality. For example, soil run-off models combined with data on topography, soil type and crop management information may provide reasonable estimates of soil erosion. Crop simulation models may help detect widening yield gaps between what are expected yields and actual yields with the difference being attributed to unmeasured trends in resource quality. Recent work has applied these modeling approaches to assessment of impacts of technological change (Dyke, 2000).

Nonetheless, there is little doubt that there is an urgent priority to invest in regular monitoring of resource quality at the level of farmers' fields. We believe that a relatively modest investment in monitoring soil quality could provide high payoffs in identifying sustainability problems and allowing more timely design of remedial actions by both farmers and policy makers. Collection of such data should become an integral part of natural resource management research programs at major research stations (Byerlee, 1991). The exact data to be collected will depend on hypotheses on the main factors likely to affect system sustainability and will necessarily be system specific. In all situations, data would be collected on carefully selected soil physical, chemical and biological parameters, with parameters such as soil organic matter likely to be included in most situations. In some cases, pest populations might be monitored through qualitative or quantitative scoring. Although some have argued for composite indicators, such as a single soil quality index (Jaenicke and Lengnick, 1999), we believe a disaggregated approach provides greater flexibility and will help to better understand the nature of the problem.

Such data collection systems necessarily require considerable up-front investment to define the key parameters, sample size, frequency of data collection, and benchmark sites. However, once established, the collection of data at regular intervals has been shown to be a relatively low cost exercise.

## 4. Conclusion

Long-term trend in productivity is an important, but only one measure, of biophysical sustainability. This review of the conceptual and practical issues in measuring productivity has shown that no one measure alone will be theoretically or empirically robust as a measure of sustainability, even in the initial diagnosis stage. Nor do single measures help in understanding the underlying factors influencing sustainability. Much more attention needs to be given to defining key indicators of agro-ecosystem health and relating these measures to trends in productivity. This analysis must be sufficiently disaggregated and for a long enough time period to allow for spatial and temporal variability inherent in agricultural production.

The key constraint now to assessing sustainability of agricultural systems is the lack of disaggregated data on resource quality for a sufficient time period. Agricultural research systems are beginning to generate such data through long-term experimental and panel surveys of farmers and their fields, but the effort to date is very modest in relation to the need. In the

[6] In addition, to our knowledge, none of these data sets have as yet been geo-referenced using global positioning technology to allow future researchers and even future generations of researchers, to accurately return to the same fields and farms in order to update information.

meantime, analysts can often do much to identify sustainability problems through partial approaches, such as yield decomposition analysis using snapshots of data on production practices, yields and resource quality at key points in time. The few examples available of this type of analysis have provided more valuable insights than a fixation on a single measure of total factor productivity. Over time, as more data becomes available, full productivity accounting using more formal statistical procedures can become routine for key production systems. Regardless of the approach selected, the findings of this paper strongly suggest a need for economists, agronomists and soil scientists to collaborate in integrating approaches in order to provide more robust and informative measures of sustainability.

## Acknowledgements

The authors appreciate helpful comments from Robert Herdt, John Lynam, Paul Heisey, Jock Anderson, Loris Strappazzon and an anonymous reviewer on an earlier draft of this paper.

## References

Ali, M., Byerlee, D., 2000. Productivity growth and resource degradation in Pakistan's Punjab: a decomposition analysis. Policy Research Working Paper No. 2480. World Bank, Washington, DC.

Alston, J.M., Anderson, J.R., Pardey, P.G., 1995a. Perceived productivity foregone, future farm fruitfulness, and rural research rationalization, 1994. In: Peters, G.H., Hedley, D. (Eds.), Agricultural Competitiveness: Market Forces and Policy Choice. Proceedings of the 22nd International Conference of Agricultural Economists, Harare, Zimbabwe, August 22–29, 1994, Aldershot, UK.

Alston, J.M., Norton, G.W., Pardey, P.G., 1995b. Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting. Cornell University Press, Ithaca, NY.

Antle, J.M., McGuckin, T., 1993. Technological innovation, agricultural productivity, and environmental quality. In: Carlson, G.A., Zilberman, D., Miranowski, J.A. (Eds.), Agricultural and Environmental Resource Economics. Oxford University Press, New York.

Barnett, V., Payne, R., Steiner, R. (Eds.), 1995. Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Bell, M.A., Fischer, R.A., Byerlee, D., Sayre, K., 1995. Genetic and agronomic contributions to yield gains: a case study for wheat. Field Crops Res. 44, 55–65.

Byerlee, D., 1991. Monitoring farmers' practices and fields as a continuing activity in onfarm research for evolving cropping systems. J. Asian Farming Syst. Res. 1, 179–190.

Byerlee, D., Siddiq, A.K., 1994. Has the green revolution been sustained: the quantitative impacts of the seed–fertilizer technology in Pakistan revisited. World Dev. 22 (9), 1345–1361.

Byerlee, D., Triomphe, B., Sebillotte, M., 1991. Integrating agronomic and economic perspectives into the diagnostic stage of on-farm research. Exp. Agric. 26, 1–20.

Cardwell, B.V., 1982. Fifty years of Minnesota corn production: sources of yield increase. Agron. J. 74, 984–990.

Cassman, K.G., Pingali, P.L., 1995. Extrapolating trends from long-term experiments to farmers' fields: the case of irrigated rice systems in Asia. In: Barnett, V., Payne, R., Steiner, R. (Eds.), Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Crosson, P., Anderson, J.R., 1993. Concerns for sustainability: integration of natural resource and environmental issues in the research agendas of NARS. ISNAR Research Report No. 4. International Service for National Agricultural Research, The Hague.

Duff, B., Rasmussen, P.F., Smiley, R.W., 1995. Wheat-fallow systems in semi-arid regions of the Pacific northwest America. In: Barnett, V., Payne, R., Steiner, R. (Eds.), Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Dyke, P., 2000. Development of IMPACT in west Africa: EPIC and SWAT. Paper Presented at a Workshop on New Developments in Estimating Impacts of Agricultural Research, Washington, DC, July 19–21.

Ehui, S.K., Spencer, D.S.C., 1993. Measuring the sustainability and economic viability of tropical farming systems: a model from sub-Saharan Africa. Agric. Econ. 9, 279–296.

Gollop, F.M., Swinand, G.P., 1998. From total factor productivity to total resource productivity: an application to agriculture. Am. J. Agric. Econ. 80 (3), 577–583.

Harrington, L., Jones, P., Winograd, M., 1994. Operationalizing sustainability: a total productivity approach. Paper Presented at the Land Quality Indicators Conference. CIAT, Cali, Colombia, June 1994.

Herdt, R.W., Lynam, J.K., 1992. Sustainable development and the changing needs of international agricultural research. In: Lee, D.R., Kearl, S., Uphoff, N. (Eds.), Assessing the Importance of International Agricultural Research for Sustainable Development. Cornell University Press, Ithaca, NY.

Herdt, R.W., Steiner, R.A., 1995. Agricultural sustainability: concepts and conundrums. In: Barnett, V., Payne, R., Steiner, R. (Eds.), Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Huang, J., Rozelle, S., 1995. Environmental stress and grain yields in China. Am. J. Agric. Econ. 77, 853–864.

ICRISAT, 1998. Database development, typology construction and economic policy analysis for sustainable rainfed agriculture. Sustainable Rainfed Agriculture Research and Development Project (Module I). Indian Council of Agricultural Research, New Delhi.

Jaenicke, E.C., Lengnick, L.L., 1999. A soil-quality index and its relationship to efficiency and productivity growth measures: two decompositions. Am. J. Agric. Econ. 81, 881–893.

Lindert, P., 1999. Bad earth: China's soils and agricultural development since the 1930s. Econ. Dev. Cult. Change 47 (4), 701–736.

Lynam, J.K., Herdt, R.W., 1989. Sense and sustainability: sustainability as an objective in international agricultural research. Agric. Econ. 3, 381–398.

McFadden, D., 1978. Cost, revenue, and profit functions. In: Fuss, M.A., McFadden, D. (Eds.), Production Economics: A Dual Approach to Theory and Applications. North-Holland, Amsterdam.

Monteith, J.L., 1990. Can sustainability be quantified? Indian J. Dryland Res. Dev. 5 (1–2), 1–15.

Murgai, R., 1999. The green revolution and the productivity paradox: evidence from the Indian Punjab. Policy Research Working Paper No. 2234. World Bank, Washington, DC.

Pieri, C., Dumanski, J., Hamblin, A., Young, A., 1995. Land quality indicators. World Bank Discussion Paper No. 315. World Bank, Washington, DC.

Pittman, R., 1983. Multilateral productivity comparisons with undesirable outputs. Econ. J. 93, 883–891.

Repetto, R., Rothman, D., Faeth, P., Austin, D., 1997. Productivity measures miss the value of environmental protection. Choices Fourth Quarter, pp. 16–19.

Sidhu, D.S., Byerlee, D., 1992. Technical change and wheat productivity in the Indian Punjab in the post-green revolution period. CIMMYT Economics Working Paper No. 92-02. Mexico, DF.

Smale, M., Hartell, J., Heisey, P.W., Senauer, B., 1998. The contribution of genetic resources and diversity to wheat production in the Punjab of Pakistan. Am. J. Agric. Econ. 80 (3), 482–493.

Squires, D., 1992. Productivity measurement in common property resource industries: an application to the Pacific Coast trawl fishery. RAND J. Econ. 23 (2), 221–236.

Steiner, R., McLaughlin, P.F., Janke, R.R., 1995. Incorporating externality costs into productivity measures: a case study using US agriculture. In: Barnett, V., Payne, R., Steiner, R. (Eds.), Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Stoorvogel, J.J., Smaling, E.M.A., 1990. Assessment of soil nutrient depletion in sub-Saharan Africa, 1983–2000. Volume 1: Main Report (2nd Edition). Report 28. Winand Staring Centre for Integrated Land, Soil and Water Research, Wageningen, the Netherlands.

Tiongco, M., Dawe, D., 2000. Is the green revolution sustainable? Long-term productivity trends in a sample of Philippine rice farms. Poster Paper Presented at the 24th International Conference of Agricultural Economists, Berlin, August 13–18.

Traxler, G., Novak, J., Mitchell, C.C., Runge, M., 1995. Long-term cotton productivity under organic, chemical, and no nitrogen fertilizer treatments, 1896–1992. In: Barnett, V., Payne, R., Steiner, R. (Eds.), Agricultural Sustainability: Economic, Environmental, and Statistical Considerations. Wiley, London.

Weaver, R.D., 1998. Measuring productivity of environmentally interactive technologies: the case of agriculture and the environment. Am. J. Agric. Econ. 80 (3), 595–599.

Whitaker, M., 1993. Using long-term experiment data in evaluating productivity and sustainability of alternative crop technologies. In: Wood, R.C., Dumanski, J. (Eds.), Sustainable Land Management for the 21st Century. Proceedings of the International Workshop on Sustainable Land Management for the 21st Century. University of Lethbridge, Lethbridge, Canada, June 20–26.