CELESTIAL SATELLITE AND EARTHLY CROP YIELD: INFORMATIONAL CONTENT OF SATELLITE-BASED CROP YIELD FORECASTS

by

Chen Zhen

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

July 2001

APPROVAL

of a thesis submitted by

Chen Zhen

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Co-Chair
Joseph Atwood ________________________________ ____________
(Signature) Date

Co-Chair
David Buschena ________________________________ ____________
(Signature) Date

Approved for the Department of Agricultural Economics and Economics

Myles J. Watts ________________________________ ____________
(Signature) Date

Approved for the College of Graduate Studies

Bruce McLeod ________________________________ ____________
(Signature) Date

STATEMENT OF PERMISSION TO USE

ACKNOWLEDGMENTS

First, I would like to thank Dr. Joe Atwood who brought me this topic (economics and satellite?). His insightful suggestions and heart-felt encouragement, for many times, went a long way toward relieving the excruciating pain of puzzlements. My sincere appreciation also goes to Dr. David Buschena who read every manuscript of this thesis and provided many detailed advices. He is always very patient with listening my (sometimes naïve) questions. Dr. Myles Watts also offered many useful suggestions, especially the inclusion of Montana winter wheat acreage information into the satellite-based model. I would also thank Donay Hansen for teaching me how to use the GIS software and for interpreting very succinctly principles of remote sensing from the satellite, and Mari Henry for making the satellite images and crop filter available and for explaining clearly the procedures in her thesis, of which my thesis is an economic and statistical extension. Donna Kelly was very helpful in preparing this manuscript. I would also thank Saleem Shaik and Roberto Valdivia for listening and offering many good advices throughout the completion of my thesis.

In addition, I wish to thank my college professor and friend, Dr. Charles Steele, who taught me, for the first time, the economics of property rights and recommended me the economics program at Montana State University, which has later proved to be a fantastic place both in terms of academics and recreation. At that time he was working with the University of Colorado at Denver (UCD) and teaching at my college in Beijing which held a joint undergraduate economics program with UCD. I would also thank Dr. Randal Rucker and other fantastic people at the department of economics who admitted me and provided me an incredible academic opportunity. In particular, I thank Dr. Ronald Johnson for reminding me, as an international student, the necessity for linguistic exhibition improvement.

A warm thank is due to Yi Jin who taught me what is more important.

Finally, I appreciate the altruism of my parents who have emotionally and financially directed and assisted me since the minute I was born on August 16th, 1977.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Since the late 70s, burgeoning efforts have been allocated to study the potential of monitoring crop conditions and forecasting crop yields via remote sensing from the satellite. An overwhelming majority of these studies shows that remote sensing from the satellite express high predictive power in crop forecasting.

In this thesis, using satellite images to forecast wheat yield from 1989 to 2000 in six Montana Crop Reporting Districts (CRD), several statistical improvements were achieved over extant crop forecasting models. First, different weights were allowed for satellite images obtained at different points of time, accounting for the likely heterogeneous contributions of various crop phenological stages to the final crop yield. Second, crop acreage information was directly modeled. This, to some extent, alleviates the low-resolution problem of existing satellite imagery. Third, jackknife out-of-sample forecasts were generated to formally measure the well-known instability problem of using satellite imagery in crop forecasting across seasons. In addition, the satellite-based crop yield forecasts were compared with those of the U.S. Department of Agriculture (USDA), whose forecasts were based on traditional methods. It is shown that although meaningful crop forecasts can be generated from the satellite imagery late season, the additional yield information that can be extracted from the satellite tends to be limited. Because in the major wheat producing CRDs, the USDA forecasts are already very accurate and little independent information is observed in the satellite-based forecasts.

Results suggest the needs to pinpoint crop phenological stages and to calibrate region-specific crop forecasting model.

CHAPTER 1

INTRODUCTION

The National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture (USDA) resumes the responsibility of regularly reporting pre-harvest production estimates for major U.S. crops. In recognition of the social welfare benefits arising from the informed decisions based on these forecasts, substantial public resources have been devoted to this endeavor.[1] Before the announcement, production forecasts for crops that are also traded in futures markets are surrounded by tight security to prevent information leakage,[2] the content of which might unfairly entitle the recipients to capture the speculative profits by adjusting trading positions in the futures market. The speculative nature of these production forecasts can be seen through the following NASS statement:

> NASS employees prepare the official [crop production] estimates in rooms that are kept locked and guarded by officers stationed outside in the hallways. Opaque vinyl shades with steel reinforcers are drawn over windows and sealed to prevent unauthorized observation. All telephones are disconnected, and computer systems are secured against tempering. The lockup area is monitored to detect the presence of electronic surveillance equipment. Once they have entered the area, employees preparing the report are prohibited from leaving the area or contacting anyone outside until the report has been released (NASS Agency Information, 2001, pp. 1, http://www.usda.gov/nassinfo/reports.htm).

---

[1] For example, the 2002 budget for NASS is $114 million (USDA, 2001).
[2] These crops are corn, soybeans, wheat, cotton, and oranges.

Since the mid-1970's private concerns have expended substantial resources to provide their subscribers crop forecasts in advance of USDA crop forecast announcements. These private forecasts are considered relatively accurate and believed to be one primary reason for the futures markets' minimal response to USDA forecasts (Garcia et al. 1997). Some authors (e.g., Just 1983) have argued that private concerns can replace the governmental crop information program. The success on the part of private enterprises in gathering and disseminating private crop information has been attributed to the technological advances during the past two decades. Among these technological improvements, remote sensing from the satellite is deemed to be promising in crop condition monitoring and production forecasting.

However, given the technical limitations found in our study, commercial promotions from private agencies claiming remote-sensing-based crop forecasts should be met with caution.[3]

This thesis examines the following questions:

1. Can remote sensing technology provide reasonably reliable "*ex ante*" forecasts, given its limitations?

2. Given that the remote-sensing-based forecasts can be made, do they have potential economic value to the market? In another word, do they contain valuable information that is *not* a subset of the current information set of the market?

This thesis is organized as follows. Chapter 2 briefly reviews the economic theory of value of information for crop forecasting. Chapter 3 introduces the technical

---

[3] The major obstacles are the low spatial and temporal resolution of the satellite imagery, lack of crop filter that can discriminate between crops, and sensors susceptible to weather interferences. These technicalities

underpinnings of forecasting crop yield using remote sensing images from the satellite. A crop-forecasting model for Montana wheat is developed in Chapter 4. Assessment of the USDA Montana wheat forecasts is presented in Chapter 5. Chapter 6 compares the informational content of the USDA and Satellite-based forecasts. Chapter 7 concludes and qualifies results found in this thesis.

---

will be dealt with in chapter 3 and 4. Because of these difficulties NASS has limited its experimental remote sensing program to one or two states.

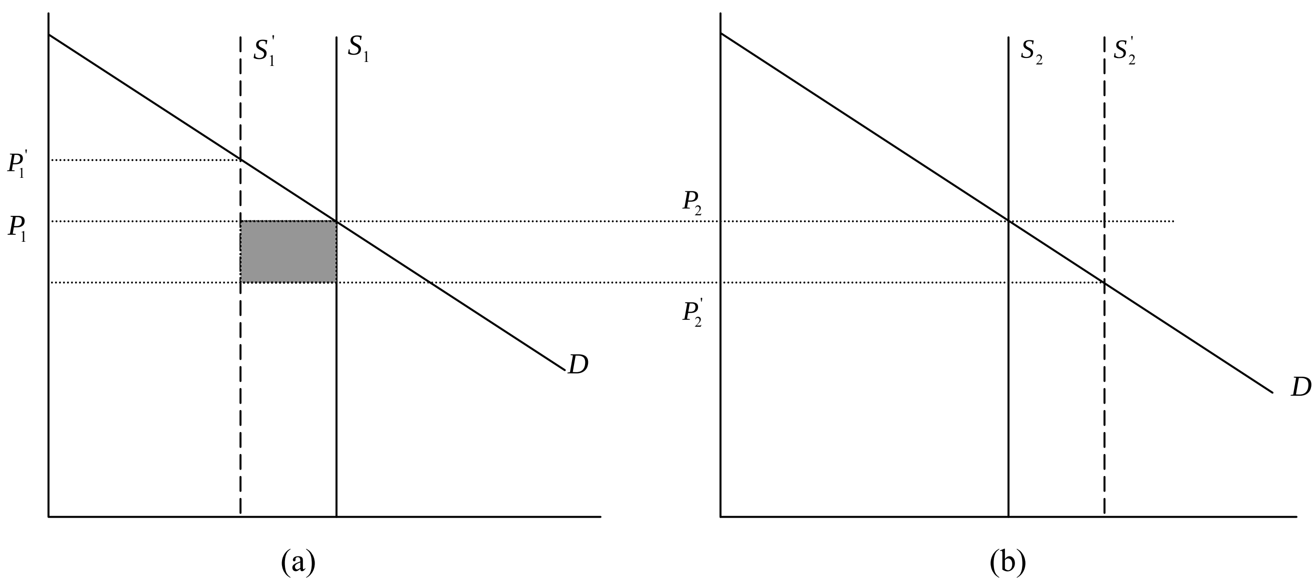
CHAPTER 2

THE VALUE OF INFORMATION FOR CROP FORECASTING

Leaving aside the gains from better resource allocation by public agencies, demand for crop forecasting services has stemmed from two distinct yet intertwined markets—the commercial and noncommercial markets.

Within the domain of the commercial market, gains from crop forecasting services is realized through the profit and utility maximizing behavior of marketing firms and consumers as they respond to crop production projections. The analysis by Hayami and Peterson (1972) serves as a useful illustration. They formulated a two-period inventory adjustment model germane to situations where production cannot be significantly adjusted at the time of forecasting. The harvest is realized in period 2, while in period 1 consumers live on grain inventory held by marketing firms. The resulting social welfare from this consumption pattern can be measured by summing the areas under the demand functions in the two periods. Abstracting from storage costs and interest rate discounts and assuming the same demand functions in the two periods, if the future harvest, H, could be determined with certainty, the social welfare would be maximized by equalizing the grain consumption between the two periods. This situation is illustrated in figure 1, when in the context of perfect foresight, grain supply in period 1, S_1, and period 2, S_2, are both equal to $(I+H)/2$, where I is the inventory available at the start of period 1. Now suppose that the forecasted harvest is H_f, which is ε_f lower than the actual but unknown quantity *H*. Further assuming that the marketing firms are endowed with knowledge of the grain market price determination mechanism, they will contract the grain supply in period 1 by an amount of $(S_1 - S_1^{'}) = \varepsilon_f / 2$ to equate period 1 price with the (erroneously) anticipated price in period 2. This inventory adjustment leads to the grain price, $P_1^{'}$, in period 1 and, once the actual harvest is realized, $P_2^{'}$ in period 2, in which $S_2^{'}$ is the actual supply and $(S_2^{'} - S_2) = \varepsilon_f / 2$ is the adjustment. The shaded area in figure 1 represents the resulting social welfare loss due to deviation from perfect foresight. Analogous results hold for situations of overly optimistic crops forecasts.[4]

Figure 1. The Hayami-Peterson Inventory Adjustment Model



In the context of a noncommercial market, the burgeoning futures trading activities during the last three decades have provided an unparalleled opportunity for agents with prior knowledge about the crop conditions to reap speculative profits by

4 For theoretical extensions of the Hayami-Peterson model, see Bradford and Kelejian (1977).

making speculative commitments before their private information is publicized. Some discussion of the informational role of crop forecasting in commodity futures markets is useful.

A traditional role of commodity futures markets, as suggested by Keynes (1930) and Hicks (1946), is that they provide a mechanism by which risk-averse speculators insure other hedgers—holders of the underlying commodities—against price fluctuation. Grossman (1977) provided an alternative view of the role of futures markets as "a place where information is exchanged, and where people who collect and analyze information about future states of the world can earn a return on their investment in information gathering" (p. 431). Consider a market in which traders are grouped according to their knowledge about the future circumstances of the world. Traders with information about the future are called "informed", while those who are ignorant are termed "uninformed". Before the initiation of a futures market, the informed traders make spot market operations (purchases or sales) in anticipation of a capital gain based on their information about the future states. Hence, in equilibrium the spot market price will be a function of the informed traders' information. By observing the present spot market price, the uninformed traders can infer some, but *not* all, of the informed traders' information. If the present spot price conveyed information to the uninformed perfectly, there would be no (private) incentive for the informed to involve in the costly process of procuring information in the first place. Luckily perhaps, the informed traders' information is not the sole determinant of present spot price; other factors ("noise") also contribute to the spot price formation. This implies that in equilibrium traders' beliefs will not be homogeneous. Grossman suggested that it is the information differentials that explain the development of futures markets in addition to the traditional hedging purposes. As long as the noisiness of the spot price formation mechanism precludes the full revelation of the inside information of the informed, there will be an incentive for the futures market to develop for information exchange and for the informed to arbitrage the information differentials. By adding a new statistic—the futures price—the uninformed now form their (rational) future spot price expectations conditional on both the present spot price and the futures price; and a more informative price system presumably leads to improved intertemporal resource allocation. Insofar as the futures price is not fully revealing, the informed will continue expending resources to obtain information. In this way the commercial and noncommercial markets are interwoven.

Grossman's concept of prices, as insufficient statistics for full information revelation, falls under the general heading of "noisy rational expectation equilibria"(NREE), upon which research efforts have extended from, to list a few, Lucas (1972), Grossman (1977), and Danthine (1978) to Grossman and Stiglitz (1980), Brown and Jennings (1989), and Wang (1994). These papers, despite their various focal points, were motivated by the recognition of the information exchange role of capital markets, i.e., asset prices aggregate traders' diverse information in a noisy manner—prices are imperfectly informative—hence traders with (expected) superior information can remunerate their information expenditure by arbitraging the information differentials among traders; in so doing, the traders' incentive for collecting information is not eliminated. Grossman and Stiglitz (1980) criticized Fama's[5] 1970 paper for conjecturing that costless information is a sufficient condition for capital market efficiency whose strong-form is defined as "… security prices at *any* time "fully reflect" *all* available information" (Fama 1970, p. 383). They demonstrated that costless information is not only a sufficient, but also a *necessary* condition for Fama's strong-form MEH.

Although being perhaps an imperfect indicator of currently available information, recent NREE models (see, for example, Kandel et al., 1993) show that uninformed traders can learn more about the informed traders' private information from a sequence of asset prices as the *rounds of trading increase*, even if no new (private or public) information enters the market since the first round. In the case of futures market, this implies that progressively higher proportion of the private information held by the informed becomes imbedded in the futures price as the informed continuously arbitrage the information differential.

An extensive literature has investigated the effects of various USDA commodity reports on futures markets (see, for example, Sumner and Mueller 1989; Baur and Orazem 1994). Most found that markets respond to the USDA reports in the sense that prices adjust to the unanticipated, but not anticipated, parts of the reports. Nevertheless, several studies (e.g. Fortenbery and Sumner 1993; Garcia et al. 1997) indicate that market reaction to USDA crop announcements plummeted since the mid-1980s.[6]

It has been suggested that the information collecting activity of private agencies since the mid-1970s has discounted the economic value of USDA crop announcements. These

---

[5] See Fama (1970, 1991) for reviews of Efficient Market Hypothesis.

[6] The minimal price reaction to the release of USDA crop announcements during the post-1985 period does not in itself imply that these crop reports do *not* have economic value. As McNew and Espinosa (1994) argued, USDA crop reports may still be informationally valuable in that they diminish market uncertainty level, even though the market expected prices may not change in the presence of these crop reports. Using implied volatility of futures option as a measure of market uncertainty level, they found that

private forecasts are usually distributed to their clients a few days before the USDA crop announcements. Hirshleifer (1971) suggested that there might be private overinvestment in procuring prior information about future state of the world when the private gains from such activity exceed the gains in social welfare. Despite the aforementioned revealing property of the futures price, its value in directing a better resource allocation in the public and commercial market domains is limited given the calendar closeness of the private and the USDA forecasts, and, henceforth, the benefits from acquisition of private forecasts seems to be largely redistributive (Just 1983).

---

the implied volatility of corn and soybean futures options declines substantially after the release of USDA corn and soybean forecasts.

CHAPTER 3

CROP YIELD AND SATELLITE IMAGERY

Physical and Physiological Basis

While an exhaustive review of the voluminous biophysical and geological literature on this subject is beyond the scope of this thesis, this chapter provides a brief discussion of the techniques involved.

It has long been demonstrated that plant canopies display heterogeneous reflectance patterns in various wavelengths of incident solar radiation (Knipling 1970). The high absorption in the visible red spectral region (0.6-0.7 μm) is primarily due, among other things, to the chlorophylls; while, in a large part, mesophylls are responsible for the high reflectance (40-60%) in the near infrared (0.7-1.3 μm). Furthermore, plant phenology and physiological stresses can be manifested through changes in canopy reflectance properties. For example, plant stresses (by diseases, insects, nutrient deficiency, etc.) may affect canopy reflectance from two reasons. First, these stresses could be felt through deteriorated visible red absorption efficiency of chlorophyll resulting from stress-generated metabolic disturbances. Second, these stresses could also cause direct loss of foliage density and plant growth suppression. What results is the increased exposure of plant branches and bare soil to solar radiation. Consequently, a smaller percentage of spectral reflectance is due to green vegetation, and more is due to bare soil and other non-foliage matters. Lillesand and Kiefer (1994) noted that bare soil has similar reflectance in both visible red and near infrared. Therefore, ceteris paribus, the reflectance pattern of the stressed plant-soil complex would be distinct from those complexes inhabited by healthy vegetation.

An overwhelming majority of studies on crop performance involving spectral reflectance assessment concentrates on multi-wavelength properties of crop canopies. In these cases various vegetation indices (VIs) are considered. VI is typically a sum, difference, ratio, or other linear combination of different spectral responses in two or more wavelengths (Wiegand et al. 1991). The benefit of analyzing VIs instead of individual wavelengths is at least threefold (Malingreau 1989). First, by appropriately combining different spectral bands, the VI should have a better-defined relationship with the physiological properties of the investigated crop than individual spectral measurements. Second, the standardization of multi-wavelength responses via VI facilitates cross-region/year comparisons. Third, VI reduces the n-dimensional (if n spectral bands measurements contribute to the calculation of the VI) data set to mono-dimensional. In general, VI does a better job in estimating the yield for fodder crops whose entire above ground green matters represent all production. For grain crops, the relation between VIs and yields are only indirect; their yields hinge on the efficiency of their photosynthetic apparatus to absorb CO_2 into their storage organs (Malingreau 1989; Benedetti and Rossini 1993). In fact, all empirical remote sensing based (grain) crop yield models depend on the assumption that the VIs are significantly linked with some crop parameters whose biological variations are decisive to the crop yield formation.

Among the various VIs, the normalized difference vegetation index (NDVI) is the most frequently examined in the literature of crop yield estimation. The NDVI is defined as the ratio of the difference between near infrared and visible red to their sum or:

$$\text{NDVI} = (\text{NIR} - \text{R}) / (\text{NIR} + \text{R}) \tag{3.1}$$

where NIR and R denote near infrared and red respectively. The unitless NDVI typically ranges from -1 to 1; negative values indicate non-vegetated scenes—i.e., snow, clouds, or bare ground, which have higher visible red reflectance than the near infrared; larger NDVI value would usually be an indication of greener vegetation.

Researchers have found that VIs are very useful in monitoring the photosynthetically active biomass of plant canopies (Tucker 1979) and, therefore, in conjecturing the final crop yield. In an effort to relate crop conditions at various stages to its final yield, numerous studies have constructed plant growth profiles by calculating sequential VIs over the growing seasons (see, for example, Tucker et al. 1981; Benedetti and Rossini 1993; Quarmby et al. 1993). These empirical examinations of sequential VIs can be classified in terms of the two distinguishable approaches each may assume: one approach tries to relate single VI observation at particular time to the final yield; the other (perhaps more plausible) studies the sequential VIs in a dynamic perspective—linking the integrated or summed VIs over a time interval (the entire growing season or a subset of it) with the final yield. This approach is illuminated by the work of Sellers (1986), and Tucker and Sellers (1986), who suggested that VIs provide more information about processes (Malingreau 1989).

Using integrated or summed VIs has the advantage of parsimony in terms of the degrees of freedom, especially when the sample size is small. This approach implicitly restricts the VI observations at different points of time (phenological stages of the crop) to have identical effects on the final yield. For example, it is well known that the most critical period for wheat yield is the period a few days after flowering and sustaining until the physiological ripeness, namely the grain filling stage. In this period the upper two leaves of the wheat plant, whose biological properties will be registered by the VI, generate the overwhelming majority of the substances accrued in the storage organs (Benedetti and Rossini 1993). Hence, in the regression of wheat yield on the sequential VIs, it is desirable to allow a larger weight for VIs in this period than VIs in other useful but less critical periods. Relying on a larger sample and pooled estimation we proceed to test this possibility in the next chapter.

AVHRR and Yield Estimation

Traditionally, hand-held or airborne spectral radiometer is used in recording spectral reflectance. Since the early 1970s applications of remotely sensed data taken from the satellite have mushroomed. Due to its regular synoptic coverage, remotely sensed data from satellite is a relatively efficient means of multi-temporal assessment of large-area terrestrial vegetation. The data set that is frequently used is extracted from the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) satellite/sensor-system. The AVHRR sensors were first carried on board the TIROS-N satellite launched on October 13, 1978. The sensor collects five channels of spectral wavebands ranging from visible red to thermal. The AVHRR NDVI is derived by:

$$\text{NDVI} = (\text{Channel } 2 - \text{Channel } 1) / (\text{Channel } 2 + \text{Channel } 1). \qquad (3.2)$$

Because continuous monitoring of crop conditions entails sequential satellite images with relatively short time intervals (high temporal resolution) to capture the temporary and permanent effects of various factors on crop vigor, the NDVI data of the NOAA AVHRR are almost the exclusive source for monitoring crop vigor from the satellite and forecasting the crop yields thereon.[7]

Among the vast array of studies within the remote sensing profession, the paper by Henry (1999) deserves particular attention. In her approach the predictive power of the NDVI data for wheat yields in six Montana Crop Reporting Districts (CRD's) and 39 counties are extensively examined. She found the strongest relation between integrated NDVI and wheat yields over the entire growing season (at the regional level $adj.R^2 = 0.753$, at the county level $adj.R^2 = 0.566$). In addition, she investigated the relation between integrated NDVI and the protein concentration at the farm level, enlightened by the fact that strong early growth of wheat, whose greener biomass is manifested by higher NDVI, will use up the soil nutrients necessary for the later grain filling stage and thus lower the protein level. As expected, significant negative relation is detected with the highest $adj.R^2$ equal to 0.789. Although her study is already very systematic, we feel there still are possible statistical improvements and, as agricultural economists, we wish to examine the informational content of the forecasts from these satellite-based models compared with those already existing in the market using more conventional crop assessment techniques. Since wheat forecasts based on more

[7] The following is quoted from NOAA website about the temporal resolution of the AVHRR sensor: "The AVHRR sensor provides global (pole to pole) on board collection of data from all spectral channels. Each pass of the satellite provides a 2399 km wide swath. The satellite orbits the Earth 14 times each day from 833 km above its surface" (http://www.ngdc.noaa.gov/seg/globsys/avhrr.shtml).

conventional methods which will serve as a benchmark for evaluating the satellite-based forecasts are only available for regional production estimates, we are only concerned with the informational content of regional wheat yield forecasts abstracting from the wheat yield forecasts at the county level and from the equally valuable protein concentration forecasts.[8]

[8] Henry argued that early warning of nutrient depletion is valuable in that farmers can apply supplemental N, accordingly, to maintain a desired level of protein. While the effectiveness of this early warning mechanism will be limited because the early season correlation between NDVI and protein concentration is low.

CHAPTER 4

SATELLITE-BASED YIELD FORECAST OF MONTANA WHEAT

The Data

Satellite images were taken from the U.S. Geological Survey's (USGS) Earth Resources Observations Systems (EROS) Data Center (EDC). EDC has developed a time series biweekly conterminous United States AVHRR maximum NDVI composite data set. Relatively cloud-free images were mosaiced by selecting the maximum NDVI of each pixel over the two-week-period. The EDC supplied NDVI values were rescaled to byte data range values of 0 to 200, in which 200 equals the original NDVI value of 1, 100 matches the original value of 0, and 0 corresponds to the original value of –1 (EDC, 1995). The image has a ground resolution of 1.1 km at nadir. Following Henry (1999), only Montana dryland acres were selected, since the majority of Montana wheat is grown under dryland[9] conditions (90+%). The 1973 polygon coverage of Montana agricultural dryland (MAPS, 1990) was imported into IMAGINEVisual GIS software and converted to an IMAGINE GRID format. The converted dryland coverage was used to create the area of interest (AOI) layer for each Montana crop-reporting district (CRD), which served as a filter to extract only the NDVI values of dryland vegetation. Wheat production and acreage data were located in various issues of *Montana Agricultural Statistics Service*. This data includes all wheat (winter wheat, spring wheat, durum, etc).

The nondiscrimination of wheat types is consistent with the agricultural dryland coverage that also includes all wheat. Although the forecast model can be more specific on the types of wheat, as a first approach we choose not to make this decomposition due to the coarseness of the satellite images and the imperfection of the crop filter to hand. Nevertheless, these images and the crop filter are probably the finest that are currently available. Six Montana CRDs were included in the study. These CRDs are Central, North Central, Northeast, South Central, Southeast and Southwest (figure 2).[10] The NDVI values under the AOI layer of each CRD were exported into an ASCII file. A simple average of these pixel NDVI values represents each CRD during a particular two-week-period. The NDVI values within the period of early April through mid-September, which roughly captures the Montana wheat-growing season, were used to build the wheat growth profile. This leaves 12 observations in each CRD per season. Typical date-of-coverage of the 12 observations in a season is presented in Table 1. Twelve years (1989-2000) of satellite imagery were used. To ameliorate the small sample problem, we stacked the observations for all CRD regions to form a panel data set with six cross-sectional units. For each variable, there are 72 observations.

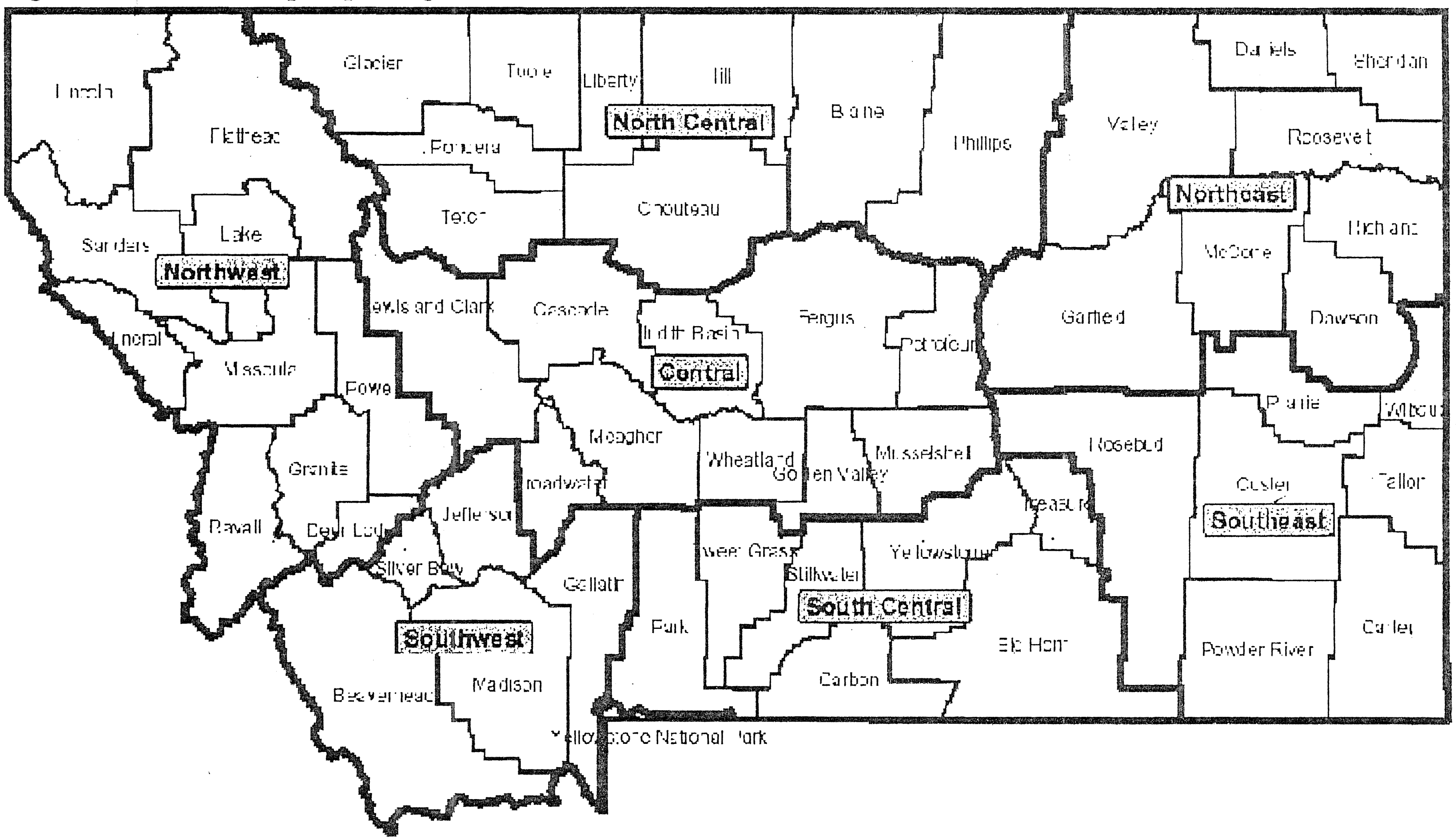
Table 1. Timing of NDVI Observations During A Representative Growing Season

	Period											
	1	2	3	4	5	6	7	8	9	10	11	12
Date of Coverage	4/01-4/14	4/15-4/28	4/29-5/12	5/13-5/26	5/27-6/09	6/10-6/23	6/24-7/07	7/08-7/21	7/22-8/04	8/05-8/18	8/19-9/01	9/02-9/15

Note—Date of Coverage in different years will slightly differ.

[9] Dryland here refers to nonirrigated cropland.
[10] The excluded Northwest CRDrepresents less than 3% of the state total wheat production.

Figure 2. Montana Crop Reporting Districts



The Model

As we have seen in chapter 2, a rubric of existing remote-sensing literature is to link the summed or integrated NDVI values over a growing season (or a subset) to the final crop yield. This assumption is, however, too restrictive. A more relaxed model should allow for the possibility of differential marginal effects of the NDVI values observed at different temporal points. Therefore, we regard each biweekly NDVI in a season as an individual right-hand regressor. One potential undesirable feature of panel data estimation is that it restricts *a priori* the marginal effect of each NDVI value at a time to be the same *across* regions (i.e., a common coefficient vector). Adding interaction terms in the forecast model could approximate this possibility. Still *ad hoc* restrictions are appended to keep the number of interaction terms low so as not to exhaust the degrees of freedom. Because the wheat yield also depends on some omitted time-invariant region-specific factors (e.g., topography, soil type, farming practice, etc.), six regional dummies are included. The ratio of all-wheat planted area to the area of dryland coverage (*AR*) in each CRD is also considered in the regression. We expect this variable to be significant—because due to the low resolution (1.1 km at nadir) of the AVHRR sensors the types of ground vegetation cannot be discriminated; we extracted, although reluctantly, the NDVI values from the entire dryland coverage in each CRD and used the simple average to build up the wheat growth profile. *AR* could measure the precision of this simple average as a representative of the actual wheat condition, which can never be exactly calibrated given the coarseness of the satellite imagery. In Montana winter wheat is usually harvested earlier than spring wheat and barley—a majority of which is also cultivated under dryland conditions. Removing winter wheat vegetation from the field should affect the ground vegetation reflectance pattern and can thus potentially impact the inference of the final crop yield drawn from the observed late season NDVI values. Accordingly, the ratio of the winter wheat planted-acres to the sum of all-wheat and barley planted acres (*WAR*) is incorporated in the regression. This variable is expected to be significant during the late season. The last variable recruited is the dummy variable for year 2000 (D_{00}) because the USGS has noted that there has been some orbital deterioration for NOAA-14 satellite in that year.[11] In addition to the possible impact from orbital deterioration, the year dummy may also capture the effects from omitted variables.

The unrestricted forecast model is formulated as follows:

$$y_{it} = \sum_{j=1}^{6} \alpha_j D_j + \sum_{k=1}^{p} \beta_k X_{it}^{k} + \sum_{j=1}^{5} (\beta^j D_j \sum_{k=1}^{p} X_{jt}^{k}) + \beta AR_{it} + \beta' WAR_{it} + \beta'' D_{00} + \varepsilon_{it} \qquad (4.1)$$

where y_{it} is the yield per planted acre of region *i* in year *t*.[12] The six regions are indexed as: 1=Central, 2=North Central, 3=Northeast, 4=South Central, 5=Southeast, 6=Southwest. The regional dummy, D_j, takes on the value of one if $j=i$ and zero otherwise. The NDVI value observed at period k is X^k. The variable *p* denotes the last

[11] It is not clear how this orbital deterioration may affect the NDVI values. NASS reported that in 2000 occasional inconsistency was observed between the NDVI values and the crop conditions reported by state extension agents and farmers in the mid-Atlantic and Midwest areas http://www.nass.usda.gov/research/avhrr/avhrrmnu.htm). Visual examination of our own data, however, does not reveal this inconsistency in Montana.

[12] Although the actual yield published by NASS is measured as yield per harvested acre, we prefer the use of yield per planted acre. This preference may be justified from two perspectives. First, it may better capture the overall crop condition. For instance, say, due to whatever reasons, half of the planted wheat perishes, while the rest grows with excellent condition until harvest. In this case, considering only yield per harvested acre would certainly distort the overall crop pattern. In formulating a crop forecast model based on the USDA weekly crop condition reports, Fackler and Norwood (2000) noted that if crop in the "poor" category is to be abandoned the yield per harvested acre might paradoxically rise when the proportion of crop rated poor increases. Second, we have deliberately kept the use of *ex post* information minimal. Here we assume that the acres planted estimates are available immediately after planting.

biweekly period whose NDVI observation is available by the time the forecast is made. A slope interaction term, $D_j \sum_{k=1}^{p} X_{jt}^{k}$, is included with the Southwest interaction effect excluded to avoid a singular data matrix. This interaction term to some extent customizes the forecast model for each region. The variables AR, WAR and D_{00} are the area-ratios and year dummy defined previously. The error term, ε, is drawn from a normal distribution. Descriptive statistics of the data set are reported in Table 2.

Table 2. Descriptive Statistics for Analysis of the Satellite-based Forecast Model

Variable		Mean	Std. Dev.	Maximum	Minimum
X^1 (April 1st – April 14th)		115.106	3.658	126.669	106.382
X^2 (April 15th – April 28th)		117.896	4.523	130.696	110.301
X^3 (April 29th – May 12th)		123.207	6.002	135.897	111.861
X^4 (May 13th – May 26th)		129.027	5.979	144.325	116.417
X^5 (May 27th – June 9th)		133.503	6.047	147.522	120.129
X^6 (June 10th – June 23rd)		138.638	5.466	148.289	120.506
X^7 (June 24th – July 7th)		139.726	6.200	152.008	118.759
X^8 (July 8th – July 21st)		138.091	6.073	150.493	126.097
X^9 (July 22nd – August 4th)		133.976	6.252	149.414	122.073
AR		0.278	0.071	0.396	0.128
WAR		0.340	0.181	0.694	0.024
Yield per planted acre by region	CN	32.013	5.569	42.246	23.396
	NC	31.818	6.785	41.278	19.090
	NE	24.816	5.518	32.052	12.817
	SC	31.909	4.448	39.640	25.829
	SE	25.456	3.355	30.083	20.990
	SW	49.790	8.706	62.419	33.243

Note—The typical dates of coverage in parentheses immediately following the biweekly NDVI's are not exactly the same across years. CN=Central, NC=North Central, CN=Central, NE=Northeast, SC=South Central, SE=Southeast, SW=Southwest.

This specification gives rise to a total of 12 forecast models, each of which is estimated the day after the last-day-of-coverage of the biweekly satellite image. For notational convenience, we denote each forecast model as Mp. For example, M6 is the model including NDVI observations up to period 6.

OLS estimates of the sequential forecast models indicate that their predictive power stabilizes by period 9 with the $\overline{R}^2$ equal to 0.889[13]. Similar results were reported by Quarmby et al. (1993), who found yield estimates for rice, cotton and wheat remain constant between 50 and 100 days prior to harvest in an area of Northern Greece. Since we are most interested in the information flow of these models, the analytic efforts discussed below are allocated exclusively to models prior to and including period 9.

Diagnosing The Data Matrix

Prior to conducting the detailed analysis, we need to resolve a concern that the columns of the data matrix may not be strictly or close to orthogonal; that is, the explanatory variables are likely plagued by multicollinearity. This concern stems from the intuitively plausible proposition that the sequential NDVI values within a season are serially correlated because of the biological nature of plant growth profile. The correlation matrix of the first nine biweekly NDVI values (Table 3) preliminarily confirms our suspicion. However, the seriousness and potential damage of this multicollinearity should be fully assessed with some more refined techniques.

Belsley et al. (1980) proposed the combined use of two diagnostic tools for multicollinearity: the singular-value decomposition (SVD) of the data matrix $\mathbf{X}$ and the regression coefficient variance decomposition.

[13] In Montana, by early August about 39%, 9% and 12% of winter wheat, spring wheat and barley are harvested respectively. These are twelve-year (1989-2000) average calculated from the weekly crop progress reports from NASS.

Table 3. Correlation Matrix for Biweekly AVHRR NDVI Observations

	X^1	X^2	X^3	X^4	X^5	X^6	X^7	X^8	X^9
X^1	1								
X^2	0.716	1							
X^3	0.508	0.579	1						
X^4	0.674	0.635	0.756	1					
X^5	0.558	0.525	0.588	0.783	1				
X^6	0.341	0.182	0.252	0.513	0.585	1			
X^7	0.040	-0.030	0.100	0.253	0.388	0.729	1		
X^8	0.267	0.209	-0.007	0.210	0.173	0.486	0.622	1	
X^9	0.010	0.017	-0.007	0.137	-0.018	0.366	0.425	0.773	1

The SVD can be defined as:

$$\mathbf{X} = \mathbf{UDV}^T \tag{4.2}$$

where $\mathbf{X}$ is the $n \times m$ data matrix; $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}_m$, $\mathbf{U}$ is $n \times m$, $\mathbf{V}$ is $m \times m$; $\mathbf{D}$ is diagonal with diagonal elements μ_r, $r = 1,...,m$, being nonnegative. The diagonal elements μ_r are called the singular values of $\mathbf{X}$, and it can be shown that these singular values are the square roots of the eigenvalues of the design matrix, $\mathbf{X}^T\mathbf{X}$.

The condition number of matrix $\mathbf{X}$ is:

$$\kappa(\mathbf{X}) = \frac{\mu_{\max}}{\mu_r} \qquad r = 1,...,m. \tag{4.3}$$

Belsley et al. demonstrated that for matrix $\mathbf{X}$ with unit column length[14] $\kappa(\mathbf{X})$ would be unity if its columns were orthogonal. A large value of $\kappa(\mathbf{X})$ (e.g., 20) would indicate severe multicollinearity. The *r*th condition index of $\mathbf{X}$ can be computed as:

$$\eta_r \equiv \frac{\mu_{\max}}{\mu_r} \qquad r = 1,...,m. \tag{4.4}$$

[14] A column vector $\mathbf{C}$ $(n \times 1)$ is said to have unit column length if $(c_1^2 + c_2^2 + ... + c_n^2)^{1/2} = 1$.

Therefore, there may be one or more condition indexes whose numerical values are greater than 20—the alarm level.

The variance-covariance matrix of the least squares estimator $\mathbf{b}$ can be expressed as:

$$\mathrm{var}(\mathbf{b}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1} = \sigma^2\mathbf{V}\mathbf{D}^{-2}\mathbf{V}^T \tag{4.5}$$

where σ^2 is the disturbance variance of the OLS regression. It can be shown:

$$\mathrm{var}(b_j) = \sigma^2\left(\frac{v_{j1}^2}{\mu_1^2} + \frac{v_{j2}^2}{\mu_2^2} + \ldots + \frac{v_{jm}^2}{\mu_m^2}\right) \qquad j = 1,\ldots,m \tag{4.6}$$

where b_j is the jth coefficient of the OLS regression; $v_{j1},\ldots,v_{jm}$ are the elements of the matrix $\mathbf{V}$. Define:

$$\phi_{jr} \equiv \frac{v_{jr}^2}{\mu_r^2} \quad \text{and} \quad \phi_j \equiv \sum_{r=1}^{m} \frac{v_{jr}^2}{\mu_r^2}, \tag{4.7}$$

then the proportion, $\pi_{rj} \equiv \phi_{jr}/\phi_j$, of b_j corresponding to μ_r is the element of the $m \times m$ variance-decomposition proportions matrix. If the proportions of two or more coefficients are larger than 0.5[15] and are associated with a singular value whose condition index is in excess of 20, then they are likely to be degraded by the multicollinearity in the $\mathbf{X}$ matrix.

Taking M9 as an example, we show how this procedure works.[16] The data matrix $\mathbf{X}$ of M9 is 72×23. It was scaled to unit column length before use to generate the variance-decomposition proportions matrix in Table 4. As can be readily seen, 15 out of 23 condition indexes are well in excess of the threshold number—20, suggesting 15 near dependencies among the columns of $\mathbf{X}$. However, determining which variables are

[15] This number, along with the threshold condition index value of 20, is based on a number of experiments by Belsley et al. (1980). They work very well from a pragmatic point of view.

involved in which near dependency is far more laborious. In the presence of at least two sets of competing dependencies (condition indexes of similar value)—45, 69, and 90 being one set, and the last 12 indexes the other—and some dependencies with condition indexes in the range of 250 to 529 possibly dominating those with condition indexes lower than 100, the proportions of each coefficient may be arbitrarily distributed among these indexes and no single proportion necessarily need to be over 0.5 to be identified the involvement in a specific dependency. In this case, the relevant detection procedure is to compute the *aggregate* proportions of each coefficient associated with these competing dependencies and still use 0.5 as the threshold value. The diagnostic results are somewhat frustrating: nearly 100% of the proportions of the coefficients on all but the winter wheat area ration (*WAR*) and the year dummy (D_{00}) are involved in one or both sets of competing dependencies, and about 64% of the aggregate proportions of the coefficient of *WAR* clusters in the weaker set of competing dependencies. We do not proceed to execute the auxiliary regressions necessary to identify which variables are involved in which dependencies. Because with so many variables, the number of auxiliary regressions we have to trial is huge.

The consequences of the multicollinearity is that the statistical inferences about individual coefficients based on student t-ratios may not be valid, and the confidence intervals constructed from the standard errors of these coefficients might be unnecessarily wide and therefore, not trustworthy. This observed multicollinearity should always be considered whenever the implications of individual coefficients are examined.

---

[16] Detailed discussion of this procedure is contained in Belsley et al. (1980).

Setting aside the problem of multicollinearity, efficient OLS estimates of the regression coefficients and the consistency of its estimated variance-covariance matrix hinge on the classical assumptions of serial independence, homoskedasticity, and cross-sectional independence. Violation of any of these assumptions will result in inefficient, although unbiased, coefficient estimates and invalid significance tests. The validity of each of these assumptions is addressed in turn below.

Testing For the Presence of Serially Correlated Disturbances

Serial independence is defined for time periods *t* and *s* in CRD *i* as:

$$E(\varepsilon_{it}\varepsilon_{is}) = 0 \qquad (t > s)\,. \tag{4.8}$$

When the equality does not hold, the disturbances are said to be autocorrelated. The most common alternative specification is:

$$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it} \tag{4.9}$$

where u_{it} are independently normally distributed with mean zero and variance σ_{ui}^2 and ε_{it} are normally distributed with variance $\sigma_{ui}^2/(1-\rho_i^2)$. In our case, an additional restriction is imposed that $\rho_i = \rho$ for all $i = 1,2,...,6$. This restriction is warranted by our short sample period (12 years) to preserve degrees of freedom for the estimation of the regression coefficient vector, $\mathbf{b}$.

A consistent estimate of ρ can be obtained by:

$$\hat{\rho} = \frac{\sum_{i=1}^{6}\sum_{t=2}^{12} e_{it}e_{i,t-1}}{\sum_{i=1}^{6}\sum_{t=2}^{12} e_{i,t-1}^2} \tag{4.10}$$

where e_{it} are the OLS estimates for ε_{it}. The $\hat{\rho}$'s for M1-9 are reported in the first row of Table 5. Under the null hypothesis of no intertemporal correlation, $\hat{\rho}$ has a standard error $(nT)^{-1/2}$, where n is the number of cross-sectional units and T is the number of time periods (see Lin 1992; and Judge et al. 1985 p. 319). In our case, n = 6, T = 12, the standard error of $\hat{\rho}$ under the null is 0.118. Therefore, the evidence suggests no discernable serial correlation in the sample.

Table 5. Test Statistics for Serial Correlation

	M1	M2	M3	M4	M5	M6	M7	M8	M9
$\hat{\rho}$	0.031	0.013	-0.111	-0.057	-0.045	-0.047	-0.064	-0.109	0.042
D-W	1.826	1.862	2.064	1.958	1.937	1.912	1.950	2.111	1.831
P-Value	0.085	0.112	0.415	0.250	0.257	0.250	0.324	0.556	0.166

Note—D-W=Durbin-Watson Statistic.

Alternatively, Bhargava et al. (1982) have generalized the Durbin-Watson test statistic to the fixed effect model. The relevant test statistic can be expressed as:

$$d = \frac{\sum_{i=1}^{6}\sum_{t=2}^{12}(e_{it} - e_{i,t-1})^2}{\sum_{i=1}^{6}\sum_{t=1}^{12} e_{it}^2}. \tag{4.11}$$

Calculation of its exact distribution is automated by SHAZAM 8.0. The d's for M1-9 are as shown in the second row of Table 5 and the corresponding p-values in the third row. As can be seen, none of the D-W statistic displays significant serial correlation at either the 1 percent or 5 percent level.

This lack of significant serial correlation is consistent with prior expectation, in the sense that there is no construction in our forecast models that would imply a moderate to strong pattern of interseasonal autocorrelation—any autoregressive effects resulting from extended drought, excessive moisture or fertilization lag upon the wheat yield would have presumably been registered through variations in the NDVI values.

Testing For Heteroskedasticity

The classical assumption of groupwise homoskedasticity is characterized as:

$$E(\varepsilon_{it}^2) = \sigma^2 \tag{4.12}$$

for all CRD's *i* and for all time periods *t*. Intuitively, we would expect this restriction to be violated. For instance, it is perhaps reasonable to assume cropping risk is different across regions[17], or the calibration of crop growth profile by the NDVI is subject to varying degrees of precision across locations; an example of such source of imprecision would be the area ratios (*AR* and *WAR*), which exhibit both regional and timewise fluctuations. If the heteroskedasticity arises from the regional differential preciseness of calibrating the crop growth profile via NDVI, we would expect it to persist throughout the season. If, on the other hand, the distinct levels of crop risk across regions are the culprits, the heteroskedasticity is expected to become less severe as the season approaches an end.[18] Because our sample spans a short period of time (12 years), we are only concerned with the presence of cross-regional (groupwise) heteroskedasticity, abstracting from the possibility of timewise heteroskedasticity.

Groupwise heteroskedasticity specifies:

$$E(\varepsilon_{it}^2) = \sigma_i^2 \,. \tag{4.13}$$

---

[17] For example, some regions might be more susceptible to natural disasters, such as hail or drought. From the early to mid-season when many of the unfavorable natural conditions have yet to occur, regions with higher cropping risk would have higher standard error of regression.

[18] This entails the further assumption that the crop damages can be ultimately captured by the satellite images.

Lagrange multiplier (LM) test has been proposed to test for groupwise heteroskedasticity (Greene 2000 p. 596). The test statistic can be computed as:

$$LM = \frac{T}{2}\sum_{i=1}^{n}\left[\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2}-1\right]^2 \quad (T = 12, n = 6) \tag{4.14}$$

where $\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T} e_{it}^2$; $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}_i^2$. Under the null hypothesis of homoskedasticity, the LM statistic has a chi-square distribution with $n-1$ degrees of freedom. Since the legitimacy of this LM test rests on the assumption of normally distributed disturbances, it is desirable to test it rather than to assume it *a priori*.

A method for testing the normality of least squares disturbances is to compute the measures of skewness and kurtosis (Kmenta, 1986 p. 266). Their sample estimates can be calculated as:

$$skewness_i = \frac{\hat{\mu}_{i3}}{\hat{\mu}_{i2}^{3/2}} \quad \text{and} \quad kurtosis_i = \frac{\hat{\mu}_{i4}}{\hat{\mu}_{i2}^{2}} \tag{4.15}$$

where $\hat{\mu}_{ir}$ is the estimated *r*th moment of the OLS disturbances of region *i*; $\hat{\mu}_{ir} = \frac{1}{T}\sum_{t=1}^{12} e_{it}^r$, *r*=2, 3, 4. The test of normality is defined for the joint null hypothesis: *skewnes* =0 and *kurtosis*=3. In practice these two measures can be combined to form the following test statistic (Kmenta 1986, p. 267):

$$normality = T\left[\frac{skewness^2}{6}+\frac{(kurtosis-3)^2}{24}\right]. \tag{4.16}$$

Under the null hypothesis, the normality statistic in (4.16) has a chi-square distribution with two degrees of freedom. Values of this normality statistic for each region and model are tabulated in the first six rows of Table 6. The critical value of chi-square distribution at the 5 and 10 percent level is 5.991 and 4.605 respectively. Hence, we would not reject normality for any of the regions and models.

Table 6. Test Statistics For Groupwise Heteroskedasticity

		M1	M2	M3	M4	M5	M6	M7	M8	M9
Normality Test	CN	1.112	1.114	0.607	1.312	1.493	1.066	0.724	2.120	0.785
	NC	1.562	2.291	1.241	0.230	0.179	0.372	1.052	1.155	2.687
	NE	1.112	0.792	0.237	0.321	0.328	0.252	0.268	1.224	0.907
	SC	0.581	0.152	0.680	0.698	0.756	0.328	0.439	0.881	1.015
	SE	0.532	0.744	0.476	1.033	1.138	0.631	0.513	0.986	0.774
	SW	1.029	0.225	0.743	0.744	0.820	0.560	1.050	0.554	0.583
	LM	15.090	13.464	12.639	8.890	8.577	10.565	11.796	4.777	16.349
	P-Value	0.010	0.019	0.027	0.114	0.127	0.061	0.038	0.444	0.006

Note—CN=Central, NC=Northcentral, NE=Northeast, SC=Southcentral, SE=Southeast, SW=Southwest, LM=Lagrange multiplier statistic.

The LM statistics for groupwise heteroskedasticity and their p-values for M1-9 are reported in row 7 and 8 of Table 6, respectively. The null hypothesis of homoskedasticity is rejected for M1, 2, 3, 7 and 9 at the 5 percent level and for M6 at the 10 percent level. These models span the periods from early April to mid-May, from a few days prior to mid-June to early July, and from late July to early August. Even in the case of M4 and 5 for which the null hypothesis cannot be rejected at the conventional levels the evidence for not rejecting the null is not very strong (p-value=0.114 for M4 and 0.127

for M5). Indeed, it is only in M8 (p-value=0.444), which roughly covers the period from July 8-21, that we have some confidence to assume groupwise homoskedasticity.

This rather erratic pattern of cross-sectional disturbances is not well predicted by either the “cropping risk” or the “calibration preciseness” argument. We are forced to consider alternatives.

Due to various agronomic and climatological reasons, the crop progress among Montana regions is not exactly synchronized, say, if the wheat vegetation has emerged in some locations (e.g., Southeast) in mid-April while others are still bare ground (e.g., North Central), the variance of the forecast errors in this period would be smaller for the early green-up regions than their bare-ground neighborhoods whose wheat yield potentials are totally unpredictable for the time being. Analogously, during the end of the season, the locations which have harvested their wheat would have their lowest forecast error variances possible, whilst the minimum forecast error variances for regions whose wheat is yet to be cut are still pending.

Examination of the reported OLS in-sample mean squared forecast errors (MSFE) of the sequential models by CRD regions (Table 7) suggests that the nonsynchronization argument seems to have some relevance. All except Southeast and Southwest region experience substantial improvement in predictive power as the forecasts move from period 8 to period 9. In contrast, the in-sample MSFE for Southwest region stabilizes by period 8 and (curious enough) for Southeast region by period 3! Furthermore, the in-sample MSFE for Southeast rises after period 3 till period 5, then decreases a little bit during period 6 and 7, rise again in period 8, and finally lowers in period 9 but never returns to its lowest level in period 3. In general, farms in the south have somewhat earlier harvest than the north. And the yield in Southeast has the lowest standard deviation in sample (Table 2). Note that even in M1 the in-sample MSFE for Southeast is quite small—smaller than that of Southwest in M9 and as small as that of Northeast in M9. This is because when substantial yield variation is lacking, a time-invariant region-specific intercept can gratifyingly predict the regional yield. Therefore, at the first sight, it seems to be the combined effects of minimal yield variance and earlier harvest that to a large degree result in this kind of forecast error behavior. But it remains baffling as to why the in-sample MSFE for Southeast has its lowest value as early as period 3, which is a few days prior to the middle of May, while, as one can see in section 4.2.7, its out-of-sample equivalent achieves the minimal as late as period 9, and before that period it has no predictive power!

Table 7. OLS Mean Squared Errors from the Satellite-based Forecast Model by CRD

Model	CN	NC	NE	SC	SE	SW
M1	21.133	31.834	30.638	14.866	7.157	64.803
M2	23.175	31.655	30.594	11.837	5.073	55.070
M3	18.142	28.292	28.351	12.092	4.104	46.418
M4	17.688	22.422	24.409	11.824	6.226	37.496
M5	16.561	22.802	24.544	12.245	6.590	37.002
M6	15.545	14.125	26.119	13.251	5.455	36.297
M7	15.119	14.320	24.683	12.201	5.446	38.031
M8	9.504	13.440	12.135	7.802	6.097	18.593
M9	3.941	6.734	7.129	3.726	5.702	18.930

Testing For Cross-Sectional Correlation

As a final diagnostic, we wish to relax the assumption that the disturbances between regions are contemporaneously uncorrelated, that is,

$$E(\varepsilon_{it}\varepsilon_{jt}) = 0 \qquad (i \neq j). \tag{4.17}$$

The unrestricted specification of this cross-sectional disturbance behavior can be expressed as:

$$E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij}^2 . \tag{4.18}$$

In view of the geographical contiguity of the six regions, it is anticipated that crop conditions in all regions can be affected by some common factors (e.g., a spell of unfavorable weather, pest conditions, etc.). This regional correlation should be most conspicuous in the early season when numerous natural mishaps detrimental to crop production are unforeseeable. Breusch and Pagan (1980) devised a Lagrange multiplier test for the null hypothesis of no cross-sectional correlation against the alternative. The Breusch-Pagan test statistic can be computed by:

$$\lambda_{LM} = T\sum_{i=2}^{6}\sum_{j=1}^{i-1} r_{ij}^2 , \tag{4.19}$$

where r_{ij} is the disturbance correlation coefficient between the *i*th and *j*th region. Under the null, the Breusch-Pagan test statistic has a chi-square distribution with degrees of freedom equal to the number of cross-sectional correlation coefficients. In our study, this number is 15.

SHAZAM 8.0 generates this test statistic using the OLS residuals to estimate r_{ij}. The LM statistics and their p-values for M1-9 are summarized in Table 8. We cannot reject the hypothesis of cross-sectional independence except for a few early-season models—M1 and 2. The sequence of forecast models with cross-sectionally correlated disturbances is truncated early in our sample.

Table 8. Test Statistics For Cross-sectional Correlation

	M1	M2	M3	M4	M5	M6	M7	M8	M9
LM	28.232	24.758	14.815	13.587	11.197	12.928	13.242	11.382	17.727
P-Value	0.020	0.053	0.465	0.557	0.739	0.608	0.584	0.725	0.277

Note — LM=Breusch-Pagan Lagrange multiplier statistic.

Alternative Estimations of the Forecast Models

In brief, statistical evidence suggests that the disturbance covariance structures of M1, 2, 3, 6, 7, and 9, and possibly 4 and 5 are groupwise heteroskedastic, with the added complication of cross-sectional correlation for M1 and 2.

The specification of groupwise heteroskedasticity and cross-sectional correlation says that the variance-covariance matrix of the disturbances ε_{it} has the following form:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2\mathbf{I} & \sigma_{12}^2\mathbf{I} & . & . & . & \sigma_{1n}^2\mathbf{I} \\ \sigma_{21}^2\mathbf{I} & \sigma_2^2\mathbf{I} & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & \sigma_{n-1,n}^2\mathbf{I} \\ \sigma_{n1}^2\mathbf{I} & . & . & . & \sigma_{n,n-1}^2\mathbf{I} & \sigma_n^2\mathbf{I} \end{bmatrix}, \tag{4.20}$$

where I is $T \times T$. Absent cross-sectional correlation, V is block diagonal. Estimation strategies other than straight OLS were considered.

As a first approach, the Parks estimator (see, Parks, 1967; Kmenta, 1986 p. 616-25; and Greene, 2000 p. 592-608) is used. The Parks estimator is in essence an FGLS procedure. Consistent estimators of elements in V are given by:

$$\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T} e_{it}^2 \tag{4.21}$$

and

$$\hat{\sigma}_{ij}^2 = \frac{1}{T}\sum_{t=1}^{T} e_{it}e_{jt} \quad (i \neq j), \tag{4.22}$$

where e are OLS residuals. The FGLS estimator is then:

$$\mathbf{b} = \left[\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X}\right]^{-1} \left[\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{y}\right]. \tag{4.23}$$

It is known that asymptotically the FGLS estimator is fully efficient. But little is known about its small sample properties except that it yields unbiased estimators under very general conditions (Kakwani 1967).

Beck and Katz (1995) have shown that in very small samples, such as in our case, the Parks estimator could substantially underestimate the sampling variability of the parameters. For the case of cross-sectionally correlated errors, the potential source of inaccuracy of Parks estimator comes from the large number of elements (in our case, this number is 21) contained in V that have to be empirically estimated. However, if only groupwise heteroskedasticity is detected, the Parks estimator may not subject to the above extreme problem, because now the number of elements in V that have to be estimated drops to the number of cross-sections (6, in our case) in sample. Instead of FGLS estimation they proposed to use, despite its inefficiency, the OLS parameter estimates with a corrected formula for the sampling variability of the OLS parameter estimates, what they called the panel-corrected standard errors (PCSE), defined as:

$$\text{var}(\mathbf{b}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}} \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1}, \tag{4.24}$$

where each term is as defined previously. Elements in $\hat{\mathbf{V}}$ are again computed from the OLS residuals following the formulas in (4.21) and (4.22). They showed that although the OLS parameter estimator is somewhat less efficient that its Parks counterpart, the difference is not large when the contemporaneous cross-sectional correlation is not extremely high and the number of time periods is not very long.[19] Moreover, the PCSE provides very accurate estimates of the variability of the OLS parameter estimates.

Results for M1-9 using both OLS parameter estimator with PCSE and Parks estimator are reported in Table 9, 10, and 11 for presentational ease. Column (1) present the OLS parameter estimates with the corresponding PCSE for all models except M8, because previous tests do not suggest complicated error structure in this model. Among these columns, the PCSE for M1 and 2 accounted for cross-sectional correlation of the errors, while the PCSE for M3, 4, 5, 6, 7 and 9 only corrected groupwise heteroskedasticity. The corresponding column (2) report the FGLS coefficient estimates correcting for groupwise heteroskedasticity and cross-sectional correlation for M1 and 2, and for groupwise heteroskedasticity for M3, 4, 5, 6, 7, and 9. Both estimators provide similar parameter estimates and model fits.

Empirical Results

Although the collinearity problem precludes us from completely disentangling the information attached to individual coefficients, a number of inferences still can be drawn.

Table 9. Coefficient Estimates of the Satellite-based Forecast Model (M1-3)

	Regression Coefficients					
	M1		M2		M3	
Explanatory	(April 14)		(April 28)		(May 12)	
Variable	(1)	(2)	(1)	(2)	(1)	(2)
X^1	0.206	0.450	-0.253	-0.109	-0.492	-0.502***
	(0.707)	(0.582)	(0.443)	(0.322)	(0.318)	(0.276)

[19] They suggested that one should find OLS with PCSE gratifying unless the average contemporaneous cross-sectional correlation is over 0.5 and the number of time periods is at least twice as large as the number of cross sections. These threshold conditions are often not met in empirical research situations. For our case, the average cross sectional correlation is 0.321 for M1 and 0.287 for M2, and the number of time periods (12 years) exactly doubles the number of regions (6 CRD's).

Table 9. Coefficient Estimates of the Satellite-based Forecast Model (M1-3) (continued):

X^2			0.713***	0.794*	0.501***	0.601**
			(0.409)	(0.280)	(0.277)	(0.244)
X^3					0.717*	0.655*
					(0.242)	(0.222)
I_CN	-0.397	-0.300	-0.384	-0.243	-0.307	-0.299
	(0.629)	(0.529)	(0.277)	(0.237)	(0.225)	(0.225)
I_NC	-1.025	-1.253***	-0.599***	-0.669***	-0.452***	-0.445***
	(0.913)	(0.744)	(0.418)	(0.359)	(0.267)	(0.266)
I_NE	0.147	0.014	0.002	-0.156	-0.109	-0.083
	(0.711)	(0.626)	(0.321)	(0.286)	(0.257)	(0.256)
I_SC	-0.645	-0.734	-0.364	-0.423	-0.376***	-0.377***
	(0.672)	(0.577)	(0.292)	(0.260)	(0.214)	(0.213)
I_SE	-0.212	-0.308	-0.187	-0.252	-0.250	-0.244
	(0.672)	(0.551)	(0.290)	(0.251)	(0.204)	(0.203)
AR	-29.701	-52.076*	-38.609	-61.339*	-63.324**	-82.314*
	(32.74)	(19.199)	(33.44)	(19.091)	(27.18)	(21.79)
WAR	-2.099	6.350	-0.069	8.047	-11.651	-6.325
	(10.74)	(4.865)	(11.19)	(4.888)	(8.380)	(6.205)
D_{00}	-6.328***	-8.924*	-7.627**	-10.453*	-8.331*	-9.015*
	(3.405)	(1.854)	(3.561)	(1.931)	(2.346)	(1.762)
$\overline{R}^2$	0.649	0.632	0.670	0.651	0.707	0.702
	2045	2144	1889	2001	1649	1678

Note—Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10% level, respectively. Six CRD dummies are not reported. The calendar dates in parentheses underneath the model number are the typical last dates of coverage which are included in the corresponding forecast models. These dates will slightly differ across years. *I* represents the interaction term. For example, *I_CN* is the interaction term for Central CRD. $\overline{R}^2$ denotes the adjusted R^2. SSR = Sum squared residual. The $\overline{R}^2$ and SSR in column (2) of each model are the unweighted statistics.

Table 10. Coefficient Estimates of the Satellite-based Forecast Model (M4-6)

	Regression Coefficients					
	M4		M5		M6	
Explanatory	(May 26)		(June 9)		(June 23)	
Variable	(1)	(2)	(1)	(2)	(1)	(2)
X^1	-0.720*	-0.596**	-0.763*	-0.641*	-0.944*	-0.914*
	(0.278)	(0.248)	(0.266)	(0.237)	(0.270)	(0.240)
X^2	0.467**	0.504**	0.421***	0.440**	0.573*	0.519*
	(0.235)	(0.211)	(0.222)	(0.198)	(0.223)	(0.193)
X^3	0.365***	0.418**	0.316	0.365**	0.360***	0.463*
	(0.223)	(0.203)	(0.205)	(0.185)	(0.201)	(0.181)
X^4	0.890*	0.685*	0.888*	0.650*	0.757*	0.605*
	(0.253)	(0.225)	(0.269)	(0.238)	(0.256)	(0.223)

Table 10. Coefficient Estimates of the Satellite-based Forecast Model (M4-6) (continued):

X^5			0.166	0.206	0.078	0.005
			(0.182)	(0.171)	(0.178)	(0.165)
X^6					0.571*	0.580*
					(0.178)	(0.163)
I_CN	-0.254	-0.252	-0.249***	-0.237***	-0.206***	-0.203
	(0.164)	(0.164)	(0.136)	(0.135)	(0.129)	(0.128)
I_NC	-0.344***	-0.338***	-0.262***	-0.253***	-0.212***	-0.214***
	(0.183)	(0.182)	(0.147)	(0.147)	(0.124)	(0.123)
I_NE	-0.217	-0.178	-0.174	-0.142	-0.184	-0.162
	(0.177)	(0.176)	(0.139)	(0.138)	(0.128)	(0.128)
I_SC	-0.328**	-0.326**	-0.272**	-0.258**	-0.263**	-0.256**
	(0.154)	(0.153)	(0.127)	(0.126)	(0.122)	(0.122)
I_SE	-0.293**	-0.272***	-0.242**	-0.220***	-0.210***	-0.195***
	(0.148)	(0.147)	(0.119)	(0.119)	(0.112)	(0.111)
AR	-33.232	-55.511**	-30.117	-52.953**	-37.045	-66.340*
	(26.50)	(23.61)	(26.86)	(24.03)	(26.33)	(23.06)
WAR	-5.107	-3.288	-5.086	-3.222	-3.021	-1.693
	(8.242)	(7.101)	(8.270)	(7.222)	(7.539)	(6.642)
D_{00}	-6.683*	-7.426*	-6.699*	-7.385*	-6.025*	-6.358*
	(2.254)	(1.949)	(2.245)	(1.963)	(2.242)	(1.924)
$\overline{R}^2$	0.739	0.733	0.735	0.728	0.750	0.744
SSR	1441	1477	1437	1473	1330	1363

Note—Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10% level respectively. Six CRD dummies are not reported. The calendar dates in parentheses underneath the model number are the typical last dates of coverage which are included in the corresponding forecast models. These dates will slightly differ across years. *I* represents the interaction term. For example, *I_NC* is the interaction term for North Central CRD. $\overline{R}^2$ denotes the adjusted R^2. SSR = Sum squared residual. The $\overline{R}^2$ and SSR in column (2) of each model are the unweighted statistics.

Table 11. Coefficient Estimates of the Satellite-based Forecast Model (M7-9)

	Regression Coefficients				
	M7		M8	M9	
Explanatory	(July 7)		(July 21)	(August 4)	
Variable	(1)	(2)	(1)	(1)	(2)
X^1	-0.888*	-0.854*	-1.325*	-1.023*	-0.867*
	(0.274)	(0.240)	(0.213)	(0.195)	(0.170)
X^2	0.555*	0.527*	0.279***	0.314**	0.350*
	(0.222)	(0.191)	(0.165)	(0.145)	(0.124)
X^3	0.328	0.396**	0.387*	0.268**	0.239**
	(0.207)	(0.183)	(0.141)	(0.137)	(0.116)
X^4	0.702*	0.579*	0.451*	0.169	0.151
	(0.250)	(0.216)	(0.180)	(0.168)	(0.143)

Table 11. Coefficient Estimates of the Satellite-based Forecast Model (M7-9) (continued):

X^5	0.032	0.083	0.388*	0.653*	0.620*
	(0.178)	(0.164)	(0.141)	(0.134)	(0.117)
X^6	0.409***	0.410**	0.369**	0.159	0.075
	(0.227)	(0.200)	(0.163)	(0.153)	(0.133)
X^7	0.317***	0.310**	-0.270***	-0.138	-0.008
	(0.171)	(0.158)	(0.145)	(0.130)	(0.115)
X^8			1.009*	0.506*	0.344**
			(0.140)	(0.150)	(0.139)
X^9				0.701*	0.769*
				(0.122)	(0.112)
I_CN	-0.181	-0.181	-0.080	-0.006	-0.003
	(0.121)	(0.121)	(0.073)	(0.065)	(0.06)
I_NC	-0.153	-0.152	-0.046	0.016	0.017
	(0.115)	(0.115)	(0.069)	(0.067)	(0.066)
I_NE	-0.150	-0.135	-0.185*	-0.174*	-0.162*
	(0.120)	(0.120)	(0.061)	(0.061)	(0.060)
I_SC	-0.229**	-0.223***	-0.119***	-0.071	-0.074
	(0.115)	(0.115)	(0.071)	(0.063)	(0.062)
I_SE	-0.190***	-0.183***	-0.125**	-0.112***	-0.110***
	(0.106)	(0.105)	(0.062)	(0.061)	(0.061)
AR	-39.492	-65.914*	-46.318**	-54.509*	-58.979*
	(25.95)	(22.83)	(20.49)	(16.05)	(14.66)
WAR	-1.549	-0.382	15.354**	19.926*	20.789*
	(7.736)	(6.840)	(7.649)	(5.728)	(5.396)
D_{00}	-5.724*	-6.290*	-4.154**	-2.250	-2.150
	(2.251)	(1.934)	(1.753)	(1.520)	(1.337)
$\overline{R}^2$	0.747	0.742	0.841	0.889	0.886
SSR	1318	1344	811	554	569

Note—Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10% level respectively. Six CRD dummies are not reported. The calendar dates in parentheses underneath the model number are the typical last dates of coverage which are included in the corresponding forecast models. These dates will slightly differ across years. *I* represents the interaction term. For example, *I_NE* is the interaction term for Northeast CRD. $\overline{R}^2$ denotes the adjusted R^2. SSR = Sum squared residual. The $\overline{R}^2$ and SSR in column (2) of each model are the unweighted statistics.

Of particular interest is a test of the null hypothesis of homogenous intertemporal marginal NDVI effects. Allowing different marginal NDVI effects across time is what distinguishes our model from the common practice of existing remote sensing literature.

This test is performed by examining $H_0 : \beta_1 = \beta_2 = ... = \beta_k , k = 2,...,9$, and these results are reported in Table 12. These results are based on both OLS with PCSE and the Parks estimation. Because M8 was estimated under straight OLS, the corresponding test statistics for this model is based on straight OLS estimation. Homogeneity of intertemporal marginal NDVI effects is uniformly rejected at the 1 percent level for all but M2 under OLS with PCSE, the null hypothesis for which can be rejected at the 10 percent level. Therefore, there is strong evidence supporting our proposition that the crop yield model should be constructed to allow for different marginal NDVI effects at different time points.

Table 12. Test Statistics for Linear Restrictions

		M1	M2	M3	M4	M5	M6	M7	M8	M9
OLS with PCSE	χ^2	—	2.683	12.997	23.182	23.250	27.589	25.115	77.792*	112.170
	p-value	—	0.101	0.002	0.000	0.000	0.000	0.000	0.000*	0.000
Parks	χ^2	—	9.659	20.004	23.124	23.512	34.003	31.159	—	144.004
	p-value	—	0.002	0.000	0.000	0.000	0.000	0.000	—	0.000

Note—$H_0 : \beta_1 = \beta_2 = ... = \beta_k , k = 2,...,9$. * indicates that the statistic is computed under straight OLS.

Although the coefficient on the all-wheat area ratio (*AR*) is only statistically significant in M3 and 9 under OLS with PCSE and in M8 under straight OLS, its Parks counterpart shows up being statistically significant in all models. Because it is very possible that the Parks estimator, when correcting for cross-sectional correlation, could substantially underestimate the sampling variability and hence lead to overconfidence in the estimated coefficients, the reported statistical significance of *AR* in M1 and 2 under Parks estimation should be interpreted with caution. Despite the disagreement between

OLS with PCSE and the Parks estimation over the coefficient significance, in both estimations, this coefficient is consistently negative, conforming with our expectation that, if wheat is the greenest vegetation during the season, then a higher *AR* would mean a higher regional NDVI value (remember our regional NDVI is a simple *average* of values extracted from *all* pixels of the agricultural dryland in that region). If the regional NDVI values (the X^k's) are held constant, a lower *AR* would necessarily require higher NDVI from per unit area of wheat vegetation, and presumably higher yield per acre, to compensate the decreased proportion of wheat vegetation in the calculation of the regional NDVI values used in our estimation.

An analogous argument holds for the positive and statistically significant coefficient of the winter wheat area ratio (*WAR*) in M8 and 9, which span the period from early July to early August. If the regional NDVI values are held fixed, higher NDVI from per unit area of spring wheat is necessary to compensate for the decreased proportion of wheat vegetation in the calculation of the regional NDVI values after the harvest of the winter wheat; then, *ceteris paribus*, the higher the *WAR*, presumably the higher the per acre spring wheat yield and perforce the per acre all-wheat yield. Indeed, AR and WAR are the weighting factors that control the marginal effects of the NDVI value on the predicted yield.

The coefficient of the year 2000 dummy (D_{00}) is negative in all models and is statistically significant except for M9. Its negative coefficient, however, does not necessarily suggests that the suspected satellite orbital deterioration in 2000 causes higher than usual NDVI values, and hence the year dummy serves as a down-weighting factor. The 2000 crop year is characterized by low wheat yields in part due to the lack of precipitation.[20] Except in the case of a completely specified model, the effects from the omitted variables which account for the low yields will be captured by the year dummy with negative coefficient. But for the summer forest fire in 2000,[21] it would appear that this line of reasoning is almost exhaustive for the statistical insignificance of this coefficient on the year dummy in M9 because its specification is more complete compared with models in earlier periods. The 2000 Montana forest fire situation culminated during the period of late July and the whole month of August. It is well known that a smoke contaminated image has lower NDVI value. It is perhaps for this reason that the coefficient of the year dummy was estimated to be less negative to up-weight the predicted yield. Obviously, there is too much noise to draw a determinate inference about the effect of satellite orbital deterioration in year 2000 on the observed NDVI values.

Another noteworthy feature of the parameter estimates is that the coefficient on the first NDVI observation (in early April) during the season, X^1, is consistently negative in all models except for being positive but not statistically significant in M1. We attribute this on a large part to the collinear nature of the data set. It is totally not surprising to expect a strong near dependency between the region-specific intercepts and the first NDVI value, because the first NDVI value exhibits little variation and if normalized would take on values around unity. Johnston (1984, p. 240) showed that if two variates are positively correlated then a negative covariance for their estimated coefficients should be anticipated. When more later-in-the-season NDVI values with which the first NDVI is

[20] The all wheat yield per planted acre in all but Northeast CRD is below the twelve-year-average (1989-2000).

positively correlated (see Table 3) are included in the model, it is very likely for the coefficient on the first NDVI to be less than its true (and perhaps a positive) value if the coefficients of the later-in-the-season NDVI observations exceed their true values. This characteristic of collinearity seems to explain the increasing (in absolute value) pattern of the coefficient of the first NDVI observation reasonably well.[22]

Likewise, because the seventh NDVI observation (X^7) is positively correlated with the eighth (X^8) and ninth (X^9) NDVI observations, when the forecast moves from M7 to M8 the coefficient on the seventh NDVI changes from being positive and significant at the 10 percent level under OLS with PCSE and at 5 percent level under the Parks estimation to being negative and significant at the 10 percent level under straight OLS. But as the ninth NDVI observation is added to the forecast model, the coefficient on the seventh NDVI observation is no longer statistically significant, perhaps because the eighth and ninth NDVI observations "dominate" the seventh in terms of the informational content in predicting the actual yield.

Consistent with the common sense that as the season progresses information concerning the potential crop yield become more and more abundant and thus crop

---

[21] In that fire season around 4,000 fires burned over 947,044 acres in Montana (Keep Montana Green Association, http://www.keepgreen.org/START.htm).

[22] It seems tempting to venture an alternative argument that would predict a negative *true* coefficient and hence explain the observed negative coefficient on the first NDVI value in M2-9. Sources within the remote-sensing profession inform us that the NDVI value from a snow-covered scene is likely to be lower than the NDVI value from a bare ground or vegetated ground. It seems plausible that a low NDVI in the early April (period 1) may be resulted from snow cover and portend a bumper year assuming normal conditions later in the season, because in Montana enough snow cover in the early spring usually heralds propitious moisture condition during the growing season. Although theoretically appealing we can offer little empirical support for this argument. In a revised M1 in which all constants are eliminated to buffer the first NDVI value from collinearity, the coefficient on the first NDVI value is positive and statistically significant. Moreover, as we introduce the second NDVI (X^2) the coefficient on the first NDVI reverses to a negative but not statistically significant number. Therefore, statistical evidence lends more credits to the collinearity argument in the text.

forecasts are increasingly accurate, the adjusted R^2 from the sequential forecast models increases monotonically. But the increment in the adjusted R^2 from period to period is not commensurate. The largest increase happens in period 8 (approximately from July 8 to 21), when the adjusted R^2 jumps from 0.747 in M7 to 0.841 in M8 under OLS estimation. The second largest increment occurs in period 9 (approximately from July 22 to August 4), during which the adjusted R^2 increases by 0.048 under OLS estimation. Minimal increments are observed in M5 (around May 27 to June 9) and 7 (around June 24 to July 7), the adjusted R^2's of which actually fall slightly, suggesting these two periods may be less critical in predicting the final wheat yield.

Finally, note the occasional statistical significance of the regional interactions in M1-9, Table 9, 10 and 11. These, along with the previous reasoning for the observed groupwise heteroskedasticity, substantiate the ideal of modeling crop yield model for each region separately. However given the limited sample size such individual calibration is not possible without exhausting the degrees of freedom.

<u>Out-of-sample Forecasting</u>

While these sequential models seem to be promising in terms of applying remote-sensing imagery in crop forecasting, an obvious drawback of in-sample forecasting is that the forecast errors can be made arbitrarily small by adding more explanatory variables. This problem could be even more serious with small samples where the low degrees-of-freedom is always a concern. In the extreme case in which the degrees of freedom is completely depleted, a spurious perfect model fit will be observed (Fair and Shiller 1989). Obviously out-of-sample verification is required to establish the efficacy of these models in generating *ex ante* forecasts and to repudiate the possible accusation of data mining. The conventional time series out-of-sample estimation is not feasible because of the short data period. In the absence of autoregressive effects from our estimating equations we can perform jackknife out-of-sample estimation by forecasting one-year wheat yields using parameters estimated from the sample with the predicted year removed. For computational ease, we apply simple OLS to all out-of-sample estimations which is acceptable in view of the unbiasedness, although possible inefficiency, of OLS. Forecasts from M7 and 9 are taken as representative for July and August forecasts respectively to be comparable to the corresponding monthly USDA forecasts that we shall deal with in the next chapter. The absolute forecast errors (AFE) for M7 are shown in Table 13, and for M9 in Table 14. The forecasted yields are plotted against actual yields in Figures 3-20. As expected the out-of-sample forecast errors are substantially higher than their in-sample counterparts.[23]

The considerable deviation from in-sample estimates is in a large part due to the regional and temporal shifts of wheat growth profiles in Montana, namely a more general nonsynchronization problem than the one discussed with groupwise heteroskedasticity which considers only regional shift of wheat growth profile. The soundness of our estimating equations rests on the assumption that there is a chain of biologically determined relations among the sequential biweekly NDVI observations and their contributions in predicting the final yield is *inherently* well defined, although econometrically these contributions can not be precisely allocated to individual NDVI observations because of their collinear nature. However, the fixed region-specific effects

[23] Note that it is no longer the case that the Southeast can stabilize its forecast errors by period 3.

and climatological conditions cause variations in crop progresses across regions and seasons—some years and/or regions may have early planting and others late, or some years and/or regions exhibit same-time planting but are different in terms of the length of the growing season. Modeling the NDVI observations mechanically in calendar sequence, although having the advantage of simplicity, is *not* impeccable.

Figure 3. July Forecasts for Central CRD

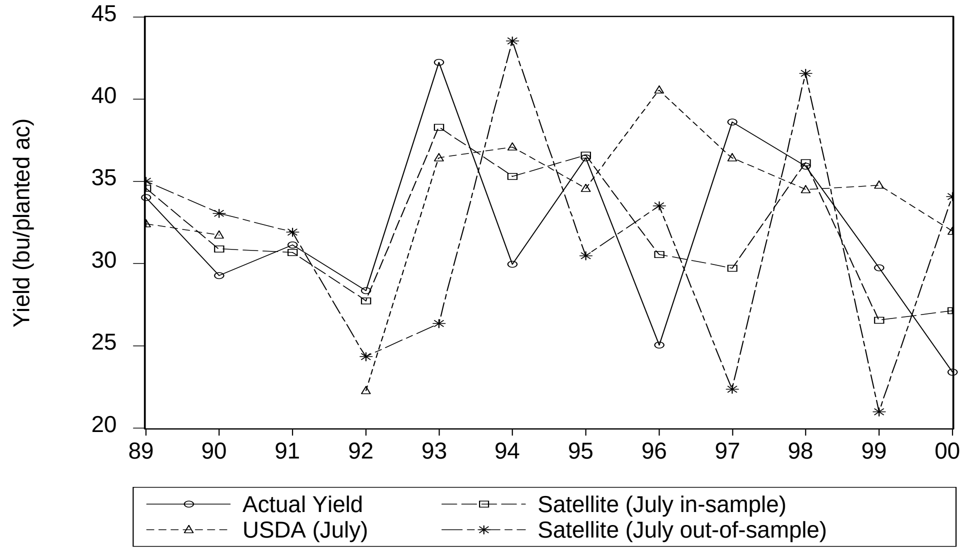


Figure 4. August Forecasts for Central CRD

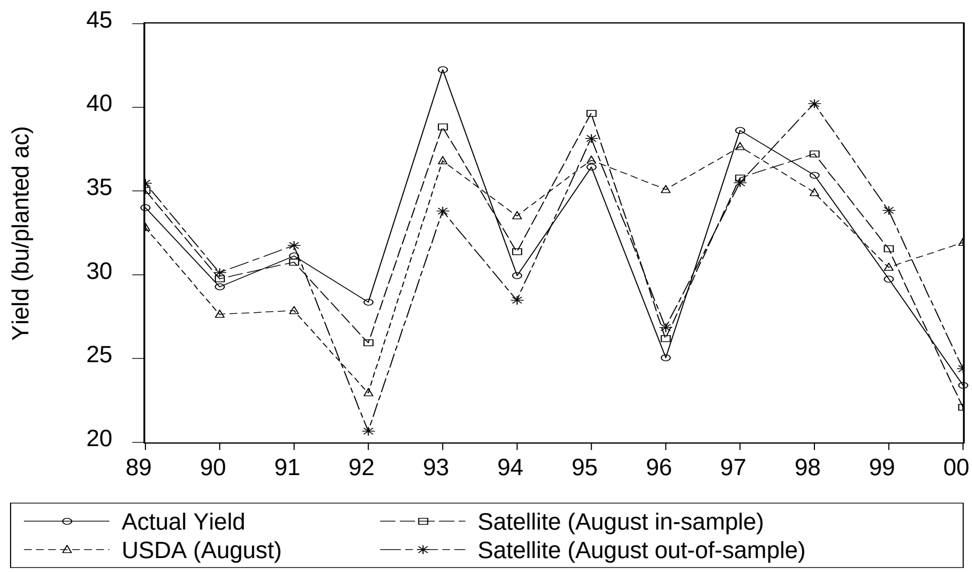


Figure 5. September Forecasts for Central CRD

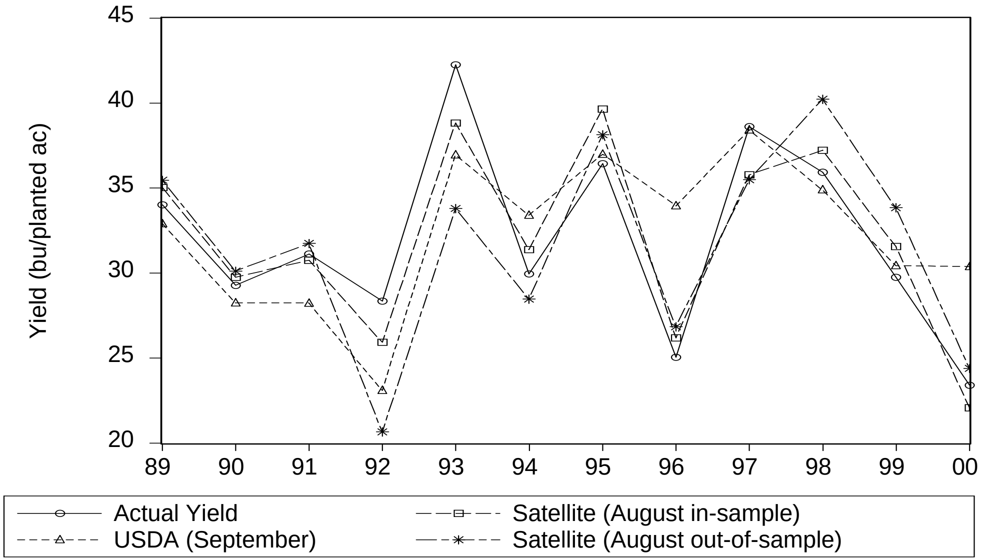


Figure 6. July Forecasts for North Central CRD

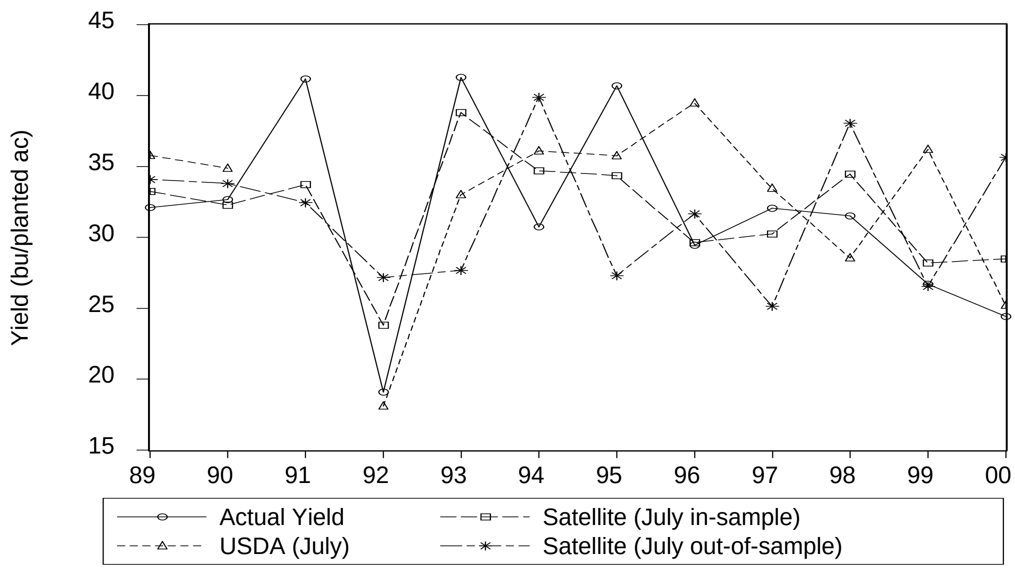


Figure 7. August Forecasts for North Central CRD

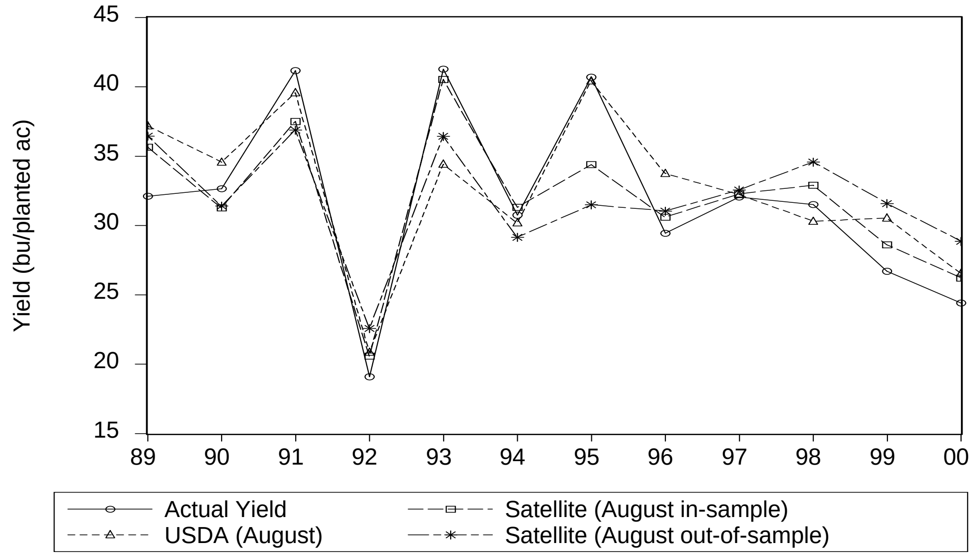


Figure 8. September Forecasts for North Central CRD

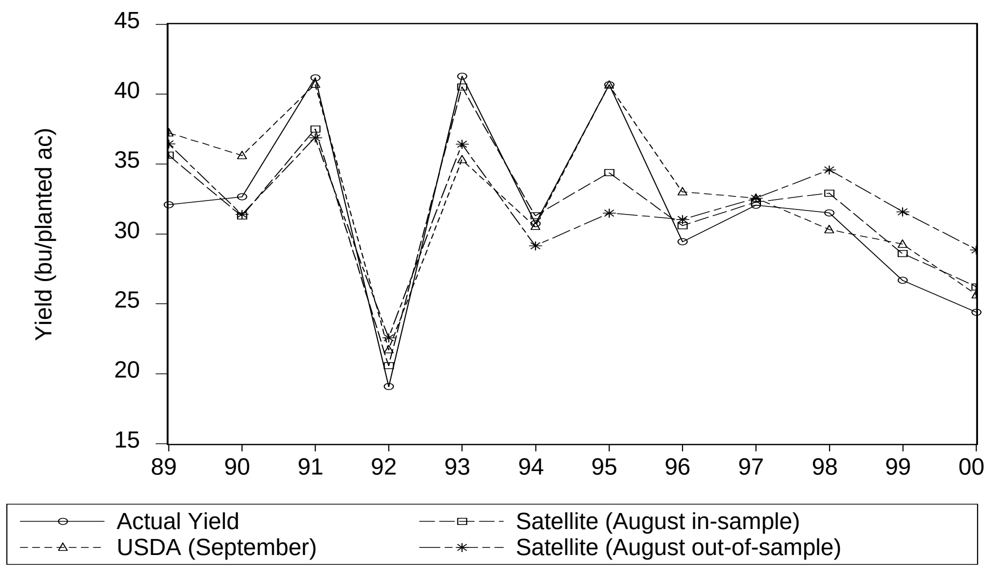


Figure 9. July Forecasts for Northeast CRD

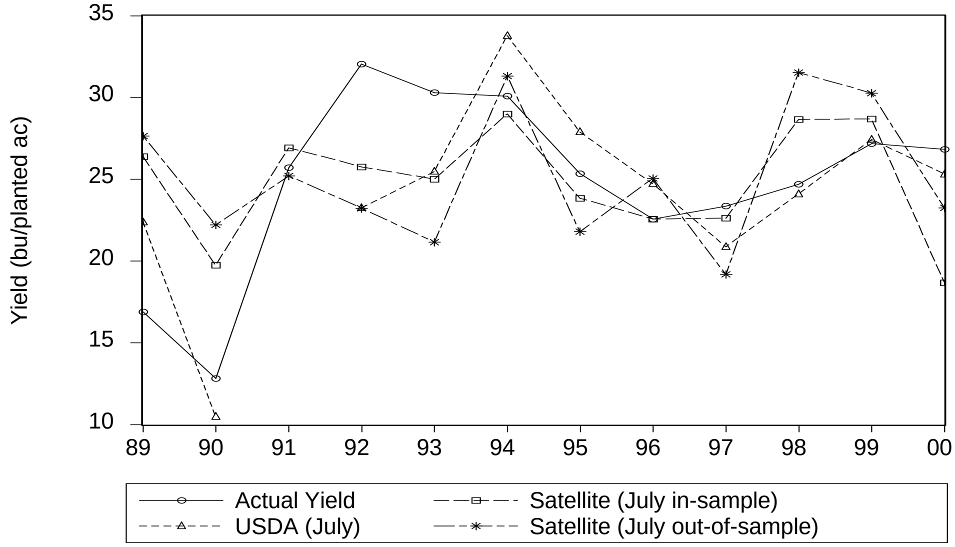


Figure 10. August Forecasts for Northeast CRD

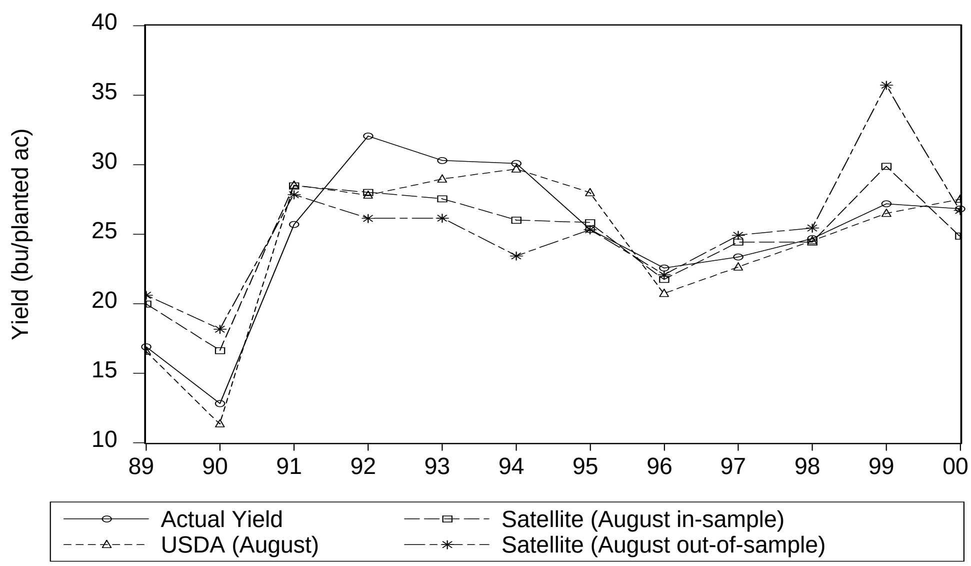


Figure 11. September Forecasts for Northeast CRD

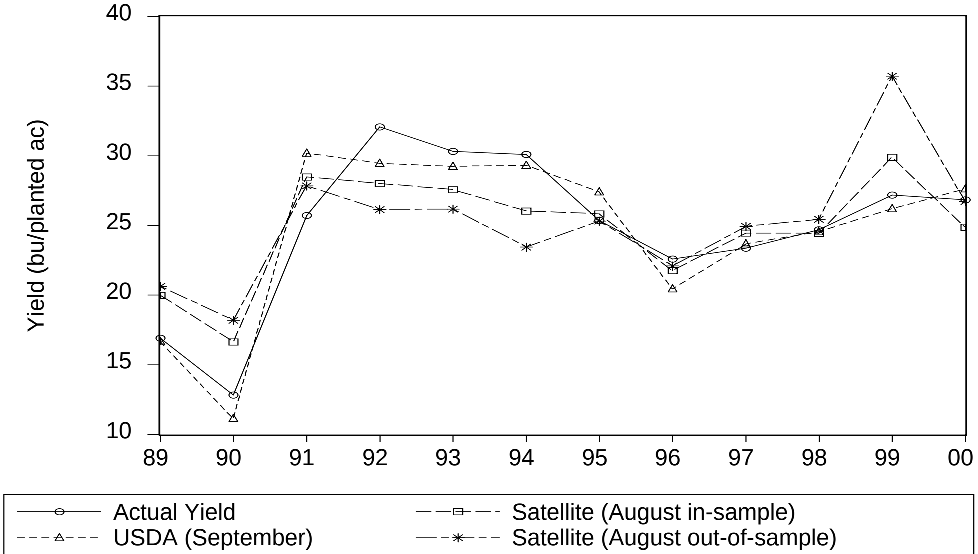


Figure 12. July Forecasts for South Central CRD

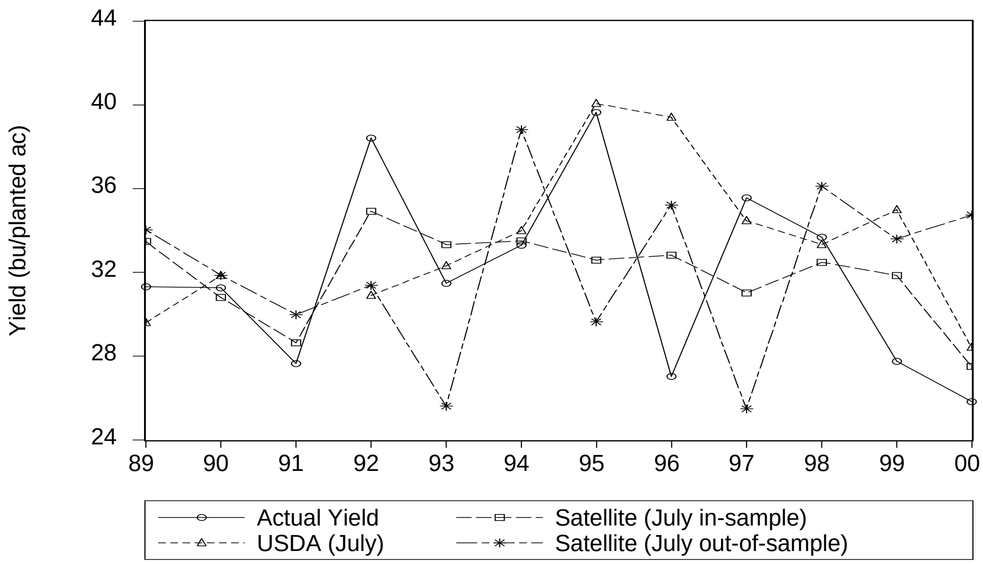


Figure 13. August Forecasts for South Central CRD

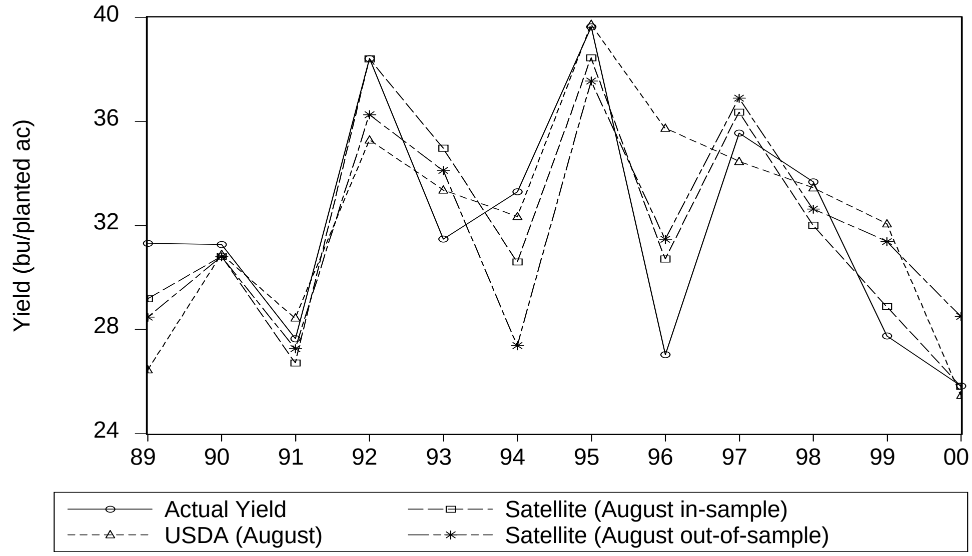


Figure 14. September Forecasts for South Central CRD

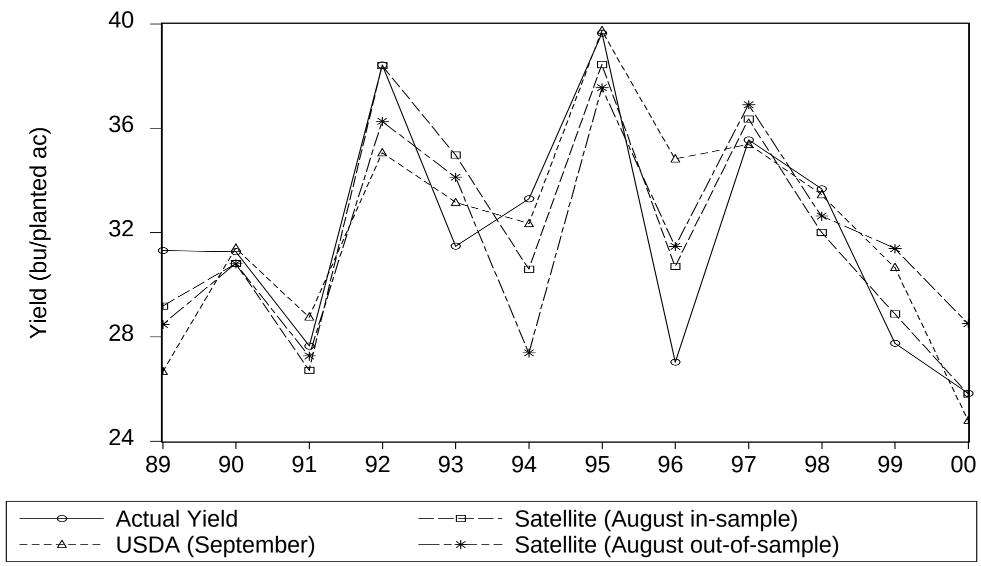


Figure 15. July Forecasts for Southeast CRD

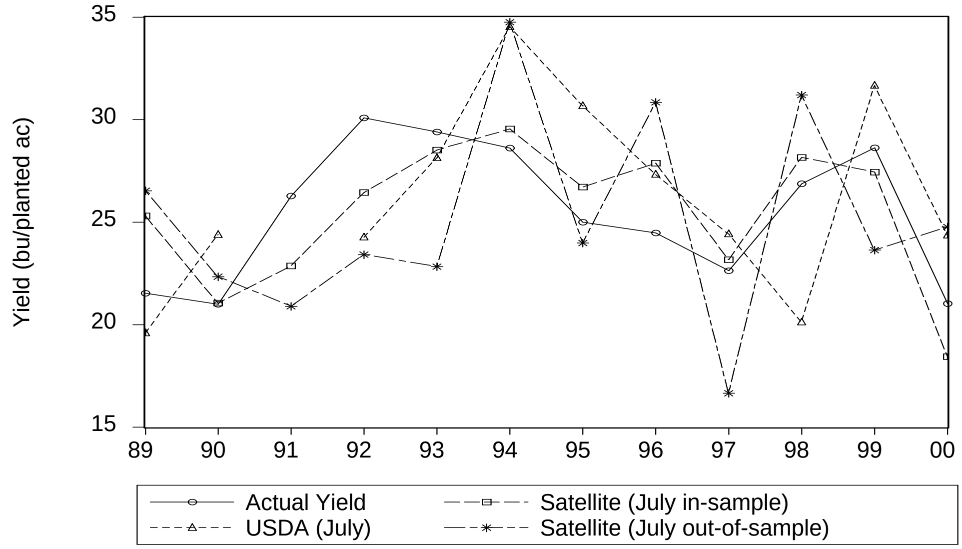


Figure 16. August Forecasts for Southeast CRD

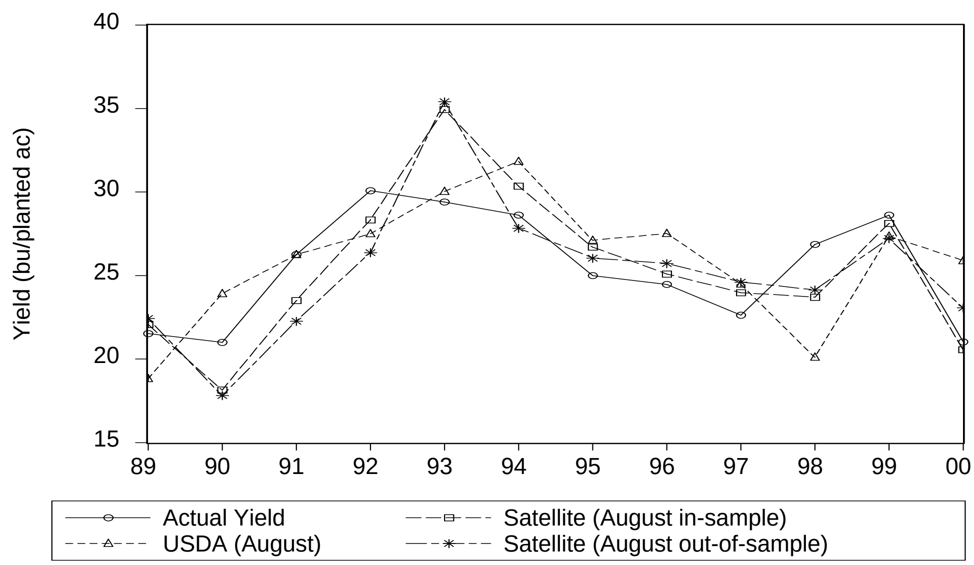


Figure 17. September Forecasts for Southeast CRD

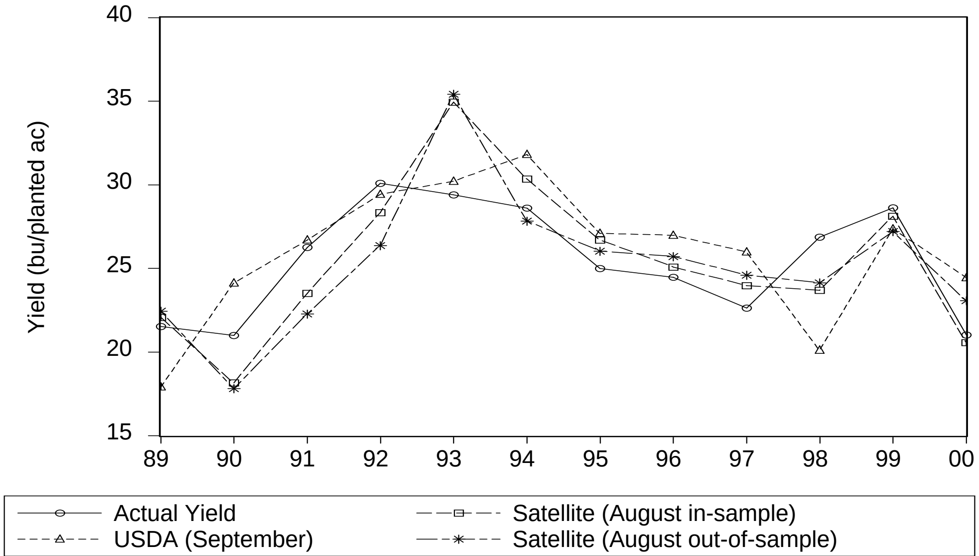


Figure 18. July Forecasts for Southwest CRD

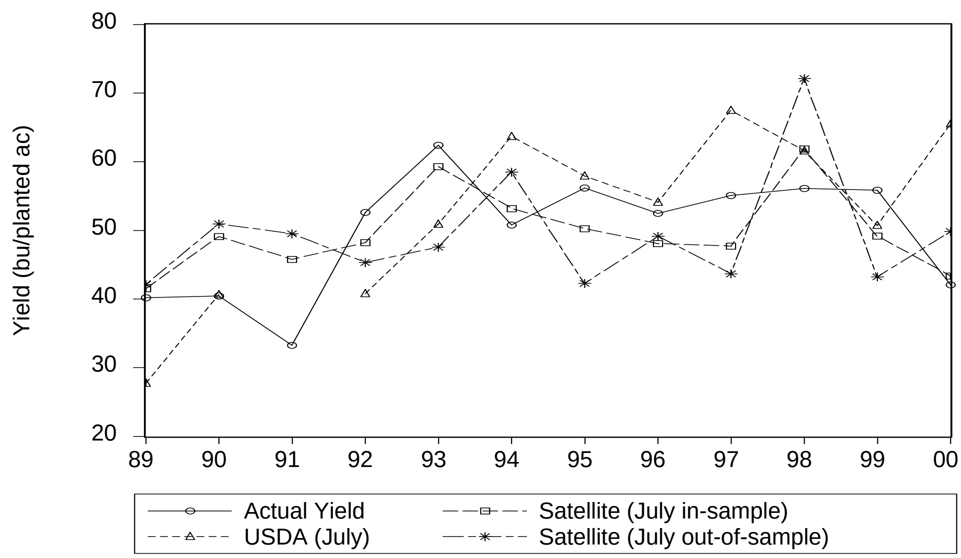


Figure 19. August Forecasts for Southwest CRD

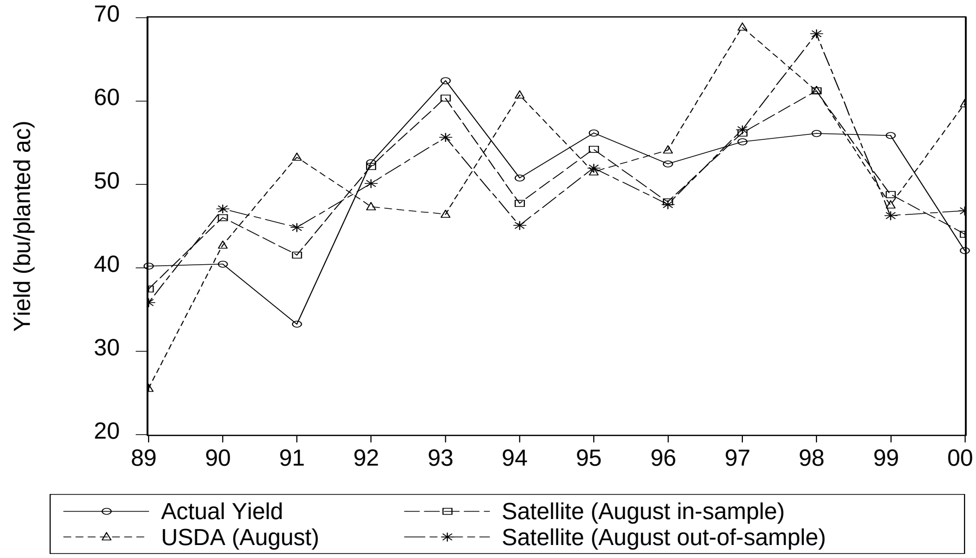
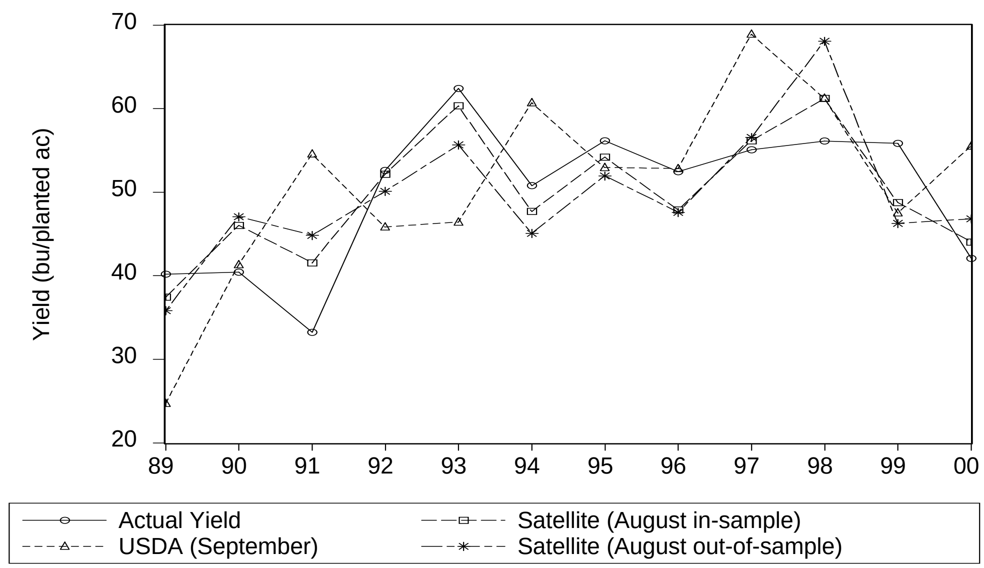


Figure 20. September Forecasts for Southwest CRD



The strictly appropriate estimating equations should be modeled in a way that incorporates information from *in situ* surveys to *match* the right NDVI values with the corresponding wheat growth stages[24]. Clearly more research efforts in the area of plant science should be called upon.

However, it still astonishes us that the July out-of-sample forecasts appear to have no predictive power.[25] If the nonsynchronization alone is culpable, it seems to us that the August out-of-sample forecasts should be as bad. But perusal of the August plots (Figure 4, 7, 10, 13, 16 and 19) indicates that they work reasonably well. Also it could be that the NDVI values included in M7 (up to early July) are not decisive to the actual yield because there is still a month or so to come before harvest, and the spurious good in-sample fit (OLS $\overline{R}^2 = 0.747$) is due to data mining. If this argument is largely legitimate then it should be the case that the predictive power of August out-of-sample forecasts derives almost exclusively from NDVI values in period 8 and 9. We test this by eliciting the jackknife out-of-sample forecasts from the following equation:

[24] Some authors (e.g., Groten, 1993, and Henry, 1999) have suggested using solely the NDVI values to identify the onset and the end of a growing season. Their approach can be seen as pinpointing the time when the photosynthetic activity first increases coupled with a "take-off" of the NDVI value as the start of a season, and the time when the photosynthetic activity declines considerably accompanied by a "dive" of the NDVI value as the end. However, existing criteria for determining the "take-off" and "dive" points of a NDVI series cannot be strictly implemented. Because each requires two consecutive NDVI changes in bounded values, whereas we cannot always fulfill these bounded values in our NDVI series. Even with some subjective discretions or some smoothing methods the start and end of a season can be determined, it is not clear how to assign these NDVI observations within a season as individual regressors if the growing season in some years and/or regions involves, say, ten biweekly NDVI observations while the season in some other years and/or regions is only eight-biweekly long.

[25] In a seemingly unrelated regression framework, the actual yield per planted acre was regressed on July satellite out-of-sample forecasted yield allowing for region-specific coefficient vector. We found all six regional slope coefficients are statistically insignificant and five of them have negative sign. This negativity in slope coefficient can be noted in Figure 3, 6, 9, 12, 15 and 18, where the July forecasted yields often point to the opposite direction as the actual yields (the "diamond" pattern). No substantive changes were observed when we estimated each region separately under OLS, except that the number of regions with negative but insignificant slope coefficient declined to three.

$$y_{it} = \sum_{j=1}^{6} \alpha_j D_j + \sum_{k=8}^{9} \beta_k X_{it}^k + \sum_{j=1}^{5} (\beta^j D_j \sum_{k=8}^{9} X_{jt}^k) + \beta AR_{it} + \beta' WAR_{it} + \beta'' D_{00} + \varepsilon_{it} \qquad (4.25)$$

where definition of terms follows from equation (4.1). If the forecasts from equation (4.25) perform as well as their counterparts from (4.1), it would suggest that the NDVI values prior to period 8, on which the July forecast is based, have no informational value in providing *ex ante* yield forecasts. However, both absolute mean forecast error (AMFE) and the adjusted R^2 in the regression of actual yield on forecasted yield tend to confirm the informational value contained in NDVI values prior to period 8.[26]

It remains puzzling that the satellite images prior to period 8 have no predictive power in the July forecast (M7) while exhibiting considerable value of information in August (M9) which is only one month later.[27] This finding has profound implication on crop forecasting efforts using remote sensing technology; during the early season *ex ante* forecasts might be erroneous in that they might occasionally provide yield information contrary to the *ex post* actual production situation despite the appealing in-sample fits which, to our knowledge, is the dominant focus of the remote sensing literature.[28]

---

[26] The AMFE for all regions pooled is 4.45 for forecasts from (4.25), 3.51 for forecasts from M9, (4.1). In pooled regression of the actual yield on forecasted yield specifying region-specific intercepts but restricting common slope coefficient, the OLS adjusted R^2 is 0.708 for forecasts from (4.25), and 0.803 for forecasts from M9, (4.1).

[27] We also experimented out-of-sample forecasts from M4, M6 and M7 but excluding 2000 data. Unfortunately no forecast exhibits predictive power.

[28] It is well known that, in addition to the shifting of phenological stages across locations and years, atmospheric perturbations, soil differences, satellite sensor viewing geometry and cropping pattern, among many others, can all affect the relation between NDVI and biophysical variables, and these potential factors have been noted in most empirical research in the remote sensing literature. However we have not seen any quantitative approach of these deleterious effects on the out-of-sample forecasts.

CHAPTER 5

THE USDA/NASS MONTHLY FORECAST OF MONTANA WHEAT

The Data

NASS conducts its monthly (July, August, and September) wheat production forecasts in cooperation with its state offices in major wheat producing states. Wheat conditions up to the first day of the announcement month are collected by interviews, probability surveys and *in situ* observations. NASS sends its enumerators to thousands of fields nationally to objectively measure such factors as plant density, number of wheat heads and spikelets. Several visits to the fields are required during each growing season. With crop information from the carefully designed samples NASS officials are then able to forecast yield per acre for the population in each area. Acreage information, which is also collected by mail surveys and face-to-face interviews, is used to scale yield forecasts to total production forecasts.

We acquired the USDA/NASS July, August, and September forecasts for Montana wheat from the Montana Agricultural Statistical Service in Helena. The sample period is identical to its satellite counterpart, that is, from 1989 to 2000.[29] Each monthly forecast consists of production estimates for winter wheat and all spring wheat at the

[29] The July 1991 forecasts were missing.

CRD level. The production forecasts for all-wheat are obtained by adding the winter and all spring wheat forecasts.

Forecast Assessment

The information flow of the USDA forecasts at various temporal points can be examined by the following regression:

$$y_{it} = \alpha_i + \beta_{USDA,i} USDA_{m,i,t} + \eta_{m,i,t} \tag{5.1}$$

where y_{it} is actual yield per planted acre defined in equation (4.1), $USDA_{m,i,t}$ represents the USDA forecast for all-wheat announced in month *m*, *m* = July, August and September, each monthly forecast is assessed independently, η is the error term. Because the area planted as wheat in each region varies from year to year, to avoid timewise heteroskedasticity, the monthly forecast used here are normalized by the all-wheat planted acres in that year.

The forecast unbiasedness test holds that:

$$USDA_{m,i,t} = E(y_{it} | \Omega_{m,i,t}) \tag{5.2}$$

where *E* is the expectation operator; $\Omega_{m,i,t}$ is the information set in the *m*th month for CRD *i*. This formulation leads to a test of the hypothesis $\alpha_i = 0$ and $\beta_{USDA,i} = 1$ in equation 5.1. Departure from these values would indicate bias.

The satellite-based yield-estimating model developed in Chapter 4 and the USDA forecasting systems are operating under the influences of the same agronomic and climatological forces. The rationale that led us to expect cross-regionally correlated disturbances in the early to mid-season satellite-based forecast models applies to the

USDA forecast as well. We formed a system of regressions by stacking equation (5.1) for each region but allowing for a region-specific coefficient vector. Breusch-Pagan tests for contemporaneous correlation resulted in values of 36.772, 22.278 and 20.725 for July, August, and September regressions, respectively. Under the hypothesis of contemporaneous independence, this statistic has a chi-square distribution with $N(N-1)/2$ degrees of freedom where N is the number of cross-sections. There are six regions in sample, and with 15 degrees of freedom the tabulated values of χ^2 are 24.996 and 22.307 at the 5 and 10 percent level, respectively. There is some evidence that the disturbances of the system of equation (5.1) might be contemporaneously correlated in July and possibly in August as well. Accordingly, we estimated the July and August regressions in the seemingly unrelated regression (SUR) framework as well as through OLS. Results are given in Table 15. Because the SUR estimation has to estimate the 21 elements in V of equation (4.20) using OLS residuals, the same reasoning that envisages the potential risk of considerably underestimating parameter sampling variability by the Parks estimator in small samples also applies to the SUR estimator. Therefore, although there may be gain in efficiency, one should be wary in interpreting the coefficient estimates of SUR estimation. Column (1) of each regression are the OLS estimates and column (2) reports the SUR equivalents.

USDA forecast performance appears to be strikingly different across regions.[30] In terms of the adjusted R^2, the best forecasts occur in the North Central and Northeast CRD whose production accounts for nearly three-quarters of the state total wheat crop.

[30] Note for this case, in which each region is assigned a separate coefficient vector, computation of PCSE for OLS is superfluous, because the standard errors of straight OLS are accurate.

Table 15. Coefficient Estimates of the USDA Forecast Assessment Regressions

	Explanatory Variable	Regression Coefficients				
		July		August		September
		(1)	(2)	(1)	(2)	(1)
Central	C	23.388	20.149*	9.898	13.206**	5.853
		(14.04)	(7.591)	(10.81)	(6.043)	(9.876)
	USDA	0.257	0.353	0.683***	0.581*	0.809**
		(0.411)	(0.219)	(0.331)	(0.182)	(0.303)
	$\overline{R}^2$	-0.065	-0.071	0.228	0.221	0.358
	SSR	326	328	239	242	199
	Test stat for H_0	2.122	9.952	0.489	5.351	0.228
		(0.176)	(0.007)	(0.627)	(0.069)	(0.800)
North Central	C	10.871	9.746	-3.893	-0.135	-3.603
		(9.261)	(5.994)	(6.143)	(4.571)	(5.250)
	USDA	0.621**	0.656*	1.098*	0.983*	1.083*
		(0.281)	(0.179)	(0.186)	(0.138)	(0.158)
	$\overline{R}^2$	0.282	0.281	0.755	0.746	0.808
	SSR	266	266	113	117	88
	Test stat for H_0	1.276	4.593	0.398	0.638	0.657
		(0.325)	(0.101)	(0.682)	(0.727)	(0.540)
Northeast	C	6.379	11.111*	2.405	4.202**	3.001
		(5.589)	(4.008)	(2.586)	(2.037)	(2.481)
	USDA	0.760*	0.564*	0.919*	0.845*	0.885*
		(0.226)	(0.160)	(0.104)	(0.081)	(0.098)
	$\overline{R}^2$	0.508	0.467	0.876	0.870	0.880
	SSR	148	160	38	40	37
	Test stat for H_0	0.678	7.747	0.589	4.351	0.733
		(0.532)	(0.021)	(0.573)	(0.114)	(0.505)
South Central	C	22.951	12.798	8.179	7.675	6.179
		(13.36)	(9.207)	(8.374)	(6.077)	(7.535)
	USDA	0.278	0.581**	0.735**	0.750*	0.799*
		(0.396)	(0.272)	(0.257)	(0.186)	(0.232)
	$\overline{R}^2$	-0.053	-0.122	0.394	0.394	0.496
	SSR	188	200	120	120	100
	Test stat for H_0	2.092	3.430	0.609	1.990	0.418
		(0.179)	(0.180)	(0.563)	(0.370)	(0.670)
Southeast	C	16.007**	19.605*	12.100***	14.160*	11.538***
		(5.881)	(3.215)	(6.159)	(4.344)	(5.426)
	USDA	0.356	0.220***	0.516**	0.436*	0.535**
		(0.220)	(0.117)	(0.236)	(0.165)	(0.206)
	$\overline{R}^2$	0.139	0.102	0.256	0.248	0.342
	SSR	95	99	84	85	74
	Test stat for H_0	4.713	45.261	2.254	11.998	2.785
		(0.040)	(0.000)	(0.156)	(0.002)	(0.109)

Table 15. Coefficient Estimates of the USDA Forecast Assessment Regressions (continued):

Southwest	C	38.672*	38.374*	38.139*	39.323*	37.371*
		(9.760)	(5.261)	(12.44)	(9.124)	(12.06)
	USDA	0.239	0.245*	0.226	0.203	0.243
		(0.180)	(0.093)	(0.236)	(0.171)	(0.231)
	$\overline{R}^2$	0.070	0.070	-0.008	-0.009	0.010
	SSR	448	448	764	765	751
	Test stat for H_0	9.157	67.012	5.610	22.253	5.486
		(0.007)	(0.000)	(0.023)	(0.000)	(0.025)

Note—Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10% level, respectively. $\overline{R}^2$ denotes the adjusted R^2. SSR = Sum squared residual. The $\overline{R}^2$ and SSR in column (2) of each regression are unweighted statistics. Test stat = test statistic for H_0, it is the F statistic for OLS (column 1) and chi-square statistic for SUR (column 2). $H_0\!:\alpha_i = 0, \beta_{USDA,i} = 1.$ C=Constant, USDA=Normalized USDA production forecast. Significance levels are given under the relevant statistics.

In July, the USDA forecasts do not appear to be informative in Central and South Central.[31] But, substantial improvements are observed for these two regions as the USDA forecasts moves into the month of August. The less accurate regional forecast happens in the Southwest CRD where the forecast appears to have no predictive power. The coefficient on the USDA forecasts in this region is not statistically significant for all time periods, with the exception that it is significant at 1 percent level in the July regression under SUR. However, given our small sample size and the insignificance of the coefficient on the USDA forecasts in Southwest in later months (August and September), its significance under SUR in July regression is not credible. Although the Southwest CRD has the highest yield per acre among the six regions in sample, its total production is the lowest due to the low acreage committed to wheat. There is also some disagreement on the predictive power of the July USDA forecasts in Southeast between OLS and

[31] Although the coefficient on the USDA forecasts for South Central in this month is statistically significant at the 5 percent level under SUR, the unweighted adjusted R^2 for this regression is negative (-0.122). Hence, this coefficient significance is likely due to the overconfidence by SUR in very small sample size.

SUR—the coefficient of the USDA forecasts is statistically significant at the 10 percent level under SUR but not at any conventional level under OLS.

The substantial discrepancy in forecast performance among regions might be, in part, due to the greater NASS efforts allocated to regions whose production vagaries are more likely to have substantial impacts on the state wheat crop.

In July, although both OLS and SUR confirm the coefficient significance of the USDA forecasts for North Central and Northeast, these two estimation schemes disagree on the forecast unbiasedness test. Under SUR, unbiasedness is rejected at the 5 percent level for Northeast and is marginally rejected at the 10 percent level for North Central, while neither is rejected under OLS. Clearly, one can hardly draw definite conclusions from this small sample size. We are inclined to give more weight to the OLS estimates because the average cross-sectional correlation is 0.427 for July regression and 0.294 for August regression. Neither of these two numbers exceeds the suggested threshold value of 0.5 by Beck and Katz (1995) for consideration of alternative to OLS, and the number of time periods exactly doubles the number of cross-sections, in which case the sample periods would not be consider "very long" by them. In contrast to the estimation problem of July satellite-based out-of-sample forecasts and considering the relatively scant information available at that time, the USDA July forecasts appear to be performing reasonably well.

In August, for all regions except for Southwest, both OLS and SUR indicate that the coefficient on the USDA forecasts is statistically significant at least at the 10 percent level. In terms of the forecast unbiasedness test, the null hypothesis is not rejected by OLS for Central and Southeast CRD, while SUR rejects the null for Central at the 10 percent level and for Southeast at the 1 percent level. If one is willing to acknowledge the superiority of OLS over SUR in providing an appropriately balanced tradeoff between efficiency and accuracy of estimating sampling variability, then the only problematic regions in terms of forecast unbiasedness are Southwest and possibly Southeast both in August and September. Southeast is included because under OLS although forecast unbiasedness cannot be rejected at the conventional levels in August (p-value = 0.156), it is marginally rejected at the 10 percent level in September. Since the credibility of statistical finding at the margin of significance level is debatable, there is no definite conclusion on the unbiasedness of the USDA forecasts for the Southeast CRD. Wheat production in the Southeast and Southwest region altogether amount to less than 10 percent of the state total. Therefore, the mediocre USDA forecasts for Southeast and even worse ones for Southwest are likely to have small influences on the state aggregate production forecasts.

The statistical assessment of disaggregated USDA wheat forecasts at the regional level in Table 15 yields results consistent with previous studies which evaluate various USDA crop reports at the country level (see, e.g., Sumner and Mueller 1989; Garcia et al. 1997). These papers conclude that the USDA crop reports are reasonably accurate given the crop conditions at the time of forecasting and become increasingly precise as the season progresses.

CHAPTER 6

ENCOMPASSING THE USDA/NASS AND THE SATELLITE-BASED FORECASTS

The relative merits in forecasting of the satellite-based vs. the more conventional forecast procedure, such as that of the USDA, should be carefully evaluated before anything can be said about the feasibility of substituting the conventional forecasts with the satellite-based forecasts.

For purposes of comparisons, the satellite-based forecast is generated by including the most recent biweekly NDVI observation whose last day of coverage is at least two days before the USDA announcement date.[32] This is satisfied by M7 for the July forecasts, and M9 for the August forecasts. Because the out-of-sample satellite-based forecasts from M7 were previously shown to lack predictive power, only out-of-sample forecasts from M9 are compared with the USDA August and September forecasts.

At first glance it seems that the August satellite-based forecasts are given an unfair advantage over the USDA forecasts, in that the former incorporates the most updated (i.e., the early announcement month) information; whereas, the latter is the forecast conditional on the crop condition up to the first day of the announcement month. This unfairness, however, may not be so protrusive because timeliness is the distinct

[32] The USDA crop reports are released around the 11th of the announcement month, after the 2:45 PM closing of the futures markets.

feature of the satellite-based crop forecasts. Therefore, this technical superiority should be formally accounted for in comparisons.

In terms of the absolute forecast error (see Table 14), the out-of-sample satellite-based forecasts appear to be better than their USDA counterparts in Central and Southwest, as good in South Central and Southeast, but are inferior to the USDA forecasts in North Central and Northeast. This is somewhat disappointing because albeit the satellite-based model and the USDA draw on number of regions in which they are better forecasters, the regions where the satellite-based model excels are areas whose wheat productions represent only a minor proportion of the state total.

Despite this disappointment, opportunity still exists that the satellite-based forecasts contain information that is independent of those reported by the USDA, that is, the information set of the August satellite-based forecast model, $\Omega^{S}_{m,i,t}$, is not completely subsumed by that of the USDA, say, $\Omega_{m,i,t}$. Plausibility of this argument derives from the fact that the USDA is primarily using field surveys and questionnaires to form its estimates, and it is very likely that the information from the more traditional assessment procedures and the state-of-the-art satellite imagery are complementary. This comparison of forecasts from alternative models can be executed by the well-established "encompassing regression" test. Using macroeconomic forecasts from different models, Fair and Shiller (1989, 1990) provided detailed discussion of this technique.

For our study, the regression can be formulated as follows:

$$y_{it} = \alpha_i + \beta_{USDA,i} USDA_{m,i,t} + \beta_{SATE,i} SATE_{Aug,i,t} + \eta_{m,i,t} \tag{6.1}$$

where y_{it} and $USDA_{m,i,t}$ are defined in equation (4.1) and (5.1), respectively. $SATE_{Aug,i,t}$ is the satellite-based yield forecast in August. If the USDA and the satellite-based forecasts can be optimally combined to achieve forecast unbiasedness, the null hypothesis $H_0 : \alpha_i = 0, \beta_{USDA,i} + \beta_{SATE,i} = 1$ would not be rejected. If $\beta_{USDA,i} \neq 0, \beta_{SATE,i} = 0$, it suggests that there is no independent information in the satellite-based forecasts that is *not* contained in the USDA forecasts and in the constant term if it is statistically significant; and vice versa, if $\beta_{USDA,i} = 0, \beta_{SATE,i} \neq 0$. If $\beta_{USDA,i} \neq 0, \beta_{SATE,i} \neq 0$, it indicates that both the USDA and the satellite-based forecasts contain independent information that is not in each other's information set.

The August satellite-based forecasts are compared with the August and September USDA forecasts independently. An individual coefficient vector is estimated for each region. A Breusch-Pagan test for contemporaneous correlation does not reject the hypothesis that the disturbances of equation (6.1) are independent among regions. Accordingly, only OLS estimation is performed.

An alternative approach to the encompassing test is the relative forecast accuracy test, initially proposed by Baur and Orazem (1994). They noted that the social value of the crop reports should be a monotonically increasing function of the reduction in market supply forecast variance due to the new information brought by these reports. This test can be implemented by the following two successive regressions:

$$y_{it} = \alpha_i + \beta_{USDA,i} USDA_{m,i,t} + \eta_{m,i,t} \tag{6.2}$$

$$y_{it} = \alpha_i + \beta_{USDA,i} USDA_{m,i,t} + \beta_{SATE,i} SATE_{Aug,i,t} + \eta_{m,i,t} \, . \tag{6.3}$$

Let $\sigma^{''}_{m,i,t}$ denote the sample variance of wheat supply forecast error in CRD *i* with only the USDA forecasts on hand, $\sigma^{'}_{m,i,t}$ represents the sample variance of wheat supply forecast error in CRD *i* with both the USDA and the satellite-based forecasts available, and $\sigma_{m,i,t}$ is the actual wheat supply sample variance. Then, the informational newsworthiness of the satellite-based forecasts can be measured by the increment in the OLS adjusted R^2 from equation 5.1 to 6.1, which is

$$\overline{R}^2_{6.1} - \overline{R}^2_{5.1} = \frac{\sigma^{''}_{m,i,t} - \sigma^{'}_{m,i,t}}{\sigma_{m,i,t}}. \qquad (6.4)$$

Results of both the forecast encompassing test and the forecast accuracy test are presented in Table 16. In addition to the August regression, the August satellite-based forecasts are also compared with the September USDA forecasts in the September regression.

For the Central region, the satellite-based forecasts seem to dominate the USDA forecasts not only in August but also in September—an important result given that the September USDA forecast is one month later than the satellite-based forecasts.[33] The informational value of the satellite-based forecast in this region is further substantiated by the notable change in adjusted R^2, which is 0.228 in August and 0.126 in September. In the Southeast region where the USDA forecasts are possibly biased, the satellite-based forecasts tend to add substantial information evident both from the substantial gain in adjusted R^2 and from the statistical significance of the coefficient on the satellite-based

[33] Conversations with NASS officials indicate that because at the time the September forecasts are set there is a better survey that will be mailed about two weeks later and be the basis for the final production estimates. They are inclined to await for this better survey before any substantial changes from the August

forecasts. Moreover, in the August regression as well as the September regression the unbiasedness test cannot reject the hypothesis of unbiased forecasts for this region with high p-values (0.333 in August and 0.334 in September), suggesting that it may be

Table 16. Coefficient Estimates of the Forecast Encompassing Regressions

	Explanatory Variable	Regression Coefficients August (1)	Regression Coefficients September (1)
Central	C	6.356	4.403
		(9.211)	(8.889)
	USDA	0.182	0.328
		(0.355)	(0.376)
	Satellite	0.626**	0.538***
		(0.275)	(0.290)
	$\overline{R}^2$	0.456	0.484
	Change	0.228	0.126
	SSR	152	144
	F test for H_0	0.239	0.123
		(0.792)	(0.886)
North Central	C	-9.209	-8.652
		(8.143)	(7.210)
	USDA	0.872**	0.897*
		(0.294)	(0.242)
	Satellite	0.397	0.349
		(0.399)	(0.343)
	$\overline{R}^2$	0.756	0.810
	Change	0.001	0.002
	SSR	101	79
	F test for H_0	0.766	1.023
		(0.493)	(0.398)
Northeast	C	2.444	2.940
		(3.674)	(3.616)
	USDA	0.920*	0.883*
		(0.144)	(0.136)
		-0.003	0.005
	Satellite	(0.189)	(0.185)
	$\overline{R}^2$	0.862	0.866
	Change	-0.014	-0.013

forecasts can be made. This may partially explain the informational value of August satellite-based forecasts in September.

Table 16. Coefficient Estimates of the Forecast Encompassing Regressions (continued):

	SSR	38	37
	F test for H_0	0.347	0.333
		(0.716)	(0.725)
South Central	C	2.437	2.031
		(8.776)	(8.330)
	USDA	0.259	0.448
		(0.400)	(0.391)
	Satellite	0.662	0.485
		(0.442)	(0.436)
	$\overline{R}^2$	0.461	0.508
	Change	0.067	0.012
	SSR	96	88
	F test for H_0	0.048	0.424
		(0.954)	(0.959)
Southeast	C	8.731	7.911
		(5.629)	(5.185)
	USDA	0.193	0.273
		(0.260)	(0.229)
	Satellite	0.465***	0.414***
		(0.229)	(0.217)
	$\overline{R}^2$	0.433	0.480
	Change	0.177	0.138
	SSR	57	53
	F test for H_0	1.244	1.243
		(0.333)	(0.334)
Southwest	C	16.980	16.620
		(14.47)	(14.42)
	USDA	-0.100	-0.084
		(0.252)	(0.254)
	Satellite	0.765***	0.754***
		(0.354)	(0.362)
	$\overline{R}^2$	0.262	0.258
	Change	0.270	0.249
	SSR	503	506
	F test for H_0	0.691	0.668
		(0.526)	(0.537)

Note—Standard errors in parentheses. *, **, *** indicate significance at 1%, 5% and 10% level, respectively. $\overline{R}^2$ denotes the adjusted R^2. SSR = Sum squared residual. Significance levels are given under F statistics. $H_0 : \alpha_i = 0, \beta_{USDA,i} + \beta_{SATE,i} = 1.$ C = Constant.

USDA = Normalized USDA production forecast. Satellite = Satellite-based forecast. Change = Change in OLS $\overline{R}^2$ from equation (5.1) to (6.1).

beneficial to optimally combine these two kinds of forecasts.[34] The greatest improvement in forecast performance is achieved through remote sensing the Southwest region. In this area the satellite-based forecasts appear to be the only meaningful information available throughout the season. Comparisons of the coefficient estimates on the satellite-based and the USDA forecasts for the South Central region, in both August and September regressions, show that these two types of forecasts are too collinear to suggest any strong claims. However, the satellite-based forecasts do carry a weight of about seven tenths in the August regression and of one half in the September regression. The statistical insignificance of the coefficient on the satellite-based forecasts in August and September regression for North Central and Northeast indicates that the USDA forecasts for these two regions are already so accurate that the satellite-based forecasts could offer little independent information.

It is shown that in several regions the satellite-based forecasts contain substantial independent information from the USDA forecasts, and this information tends to persist throughout the rest of the season. Yet, in view of the fact that the regions for which the satellite-based forecasts are most valuable are also where variations in wheat production are less likely to have substantial impact on the state total wheat crop, the potential value of information encapsulated in the satellite-based forecasts is limited.

Apart from comparing the relative information content in each forecasting system, it might be plausible to argue that the satellite-based forecasting system can be operated under much lower budget than the USDA/NASS forecasts. This is particularly important

---

[34] Although the coefficient on the USDA forecasts in this region is not statistically significant, these forecasts may still be valuable. Because in an unbiasedness test for the August satellite-based forecasts

in the presence of the ensued budget squeeze on the public data gathering agencies, which is one of the stimuli for a number of papers reappraising the value of information contained in various USDA reports over the last two decades. It is not clear what is the NASS budget allocated specifically to the crop forecasting system. Yet, it should represent a substantial portion of the agency's total expenditure. Currently, the compact disc storing two months of conterminous US biweekly AVHRR NDVI images are sold for about 32 dollars and the software needed to process these discs can be purchased for several hundred dollars. Of course, we do not intend to declare that this is the only cost associated with the satellite-based forecasts as one must recognize the astronomical budget for the space program.[35] Nevertheless, in a formal cost/benefit analysis it may be argued that the marginal costs of additional information from the satellite-based model can be lower than that from the USDA crop assessment system. Macauley and Toman have argued that, " [the "value-added" processing and analyzing of the satellite imagery] approximate a constant returns technology given the modularity of workstations and software" (p 39, 1991).

---

alone, the hypothesis of forecast unbiasedness is rejected only for the Southeast region at the 5 percent level, while for other regions the test statistics are insignificant with high p-values.

[35] The budget estimated by GAO for the Earth Observing System (EOS) of NASA is 33 billion dollars through 2020. The EOS program is launched to cope with the burgeoning need for global environmental monitoring and has been object of grumbles for federal budget cuts (http://www.sdsc.edu/SDSCwire/v2.6/41_policy.html).

CHAPTER 7

CONCLUSIONS AND QUALIFICATIONS

Traditional procedures of forecasting crop production are somewhat subjective and plagued by the nonresponse bias. In contrast to the traditional field surveys and questionnaires, remote sensing from the satellite has been claimed to present the crop forecaster a more timely and panoramic image of the crop condition.

Interest in these claims prompted a systematic study of the economics of remote sensing technology as applied to crop forecasting. Specifically, this thesis explored the gain in information from satellite images at various stages of the crop production. To the best of my knowledge, the potential informational value contained in the satellite image has not riveted the attention of agricultural economists. The most extensive discussion of remote sensing technology is, of course, in the department of earth and natural resource sciences. As the remote sensing specialists strive to refine their satellite images, statistical inferences from the existing images have not been seriously explored. Using the Normalized Difference Vegetation Index (NDVI) from the Advanced Very High Resolution Radiometer (AVHRR) this thesis was able to develop a series of crop growth profiles; whereby, the information from the satellite about Montana wheat conditions and production were extracted.

Results are mixed and vaguely encouraging. On the one hand, comparisons of these satellite-based wheat forecasts with those of the USDA establish the informational usefulness of the satellite-based forecasts. Statistical tests suggest these two kinds of forecasts could be complementary at certain stages of wheat production. On the other, however, contrary to the widely held belief that remote sensing from the satellite can provide updated crop forecasts in high temporal frequency throughout the season, it is not until July that meaningful wheat forecasts can be generated in Montana for the sample period included here. Further research should be directed into investigating why the satellite images before July exhibit no predictive power in the forecast model using satellite images up to the end of June, while, puzzlingly, contribute considerable value of information as the model includes two more biweekly images. More data is required to better calibrate the crop model developed here and to eliminate inappropriate conclusions reached at this point from the small sample. Since the raw AVHRR NDVI images are available as early as 1979, resources should be diverted to process the images of 1979-88 into the ready-to-use format for researchers outside the remote-sensing discipline.

Refining the extant satellite-based forecasting models is still possible but entails greater interdisciplinary research efforts. Specifically, a better crop yield model is possible if the forecaster can accurately identify the timing of various crop phenological stages and relating the corresponding NDVI observations with the actual yield. However procuring this kind of information entails *in situ* observation, which is both costly and time-consuming. Therefore, the timeliness and the cost competitiveness of a more carefully constructed satellite-based model will be compromised.

The results of this study need to be qualified with an important caveat. The USDA production forecast errors could be decomposed into forecasting error in forecasting the yield per harvested acre and measurement error in estimating the number of acres to be harvested. However, in our satellite-based model, the pre-harvest measurement error in total planted acres is deliberately spelled out by incorporating the published USDA acreage data months after the harvest. It is on this score that the satellite-based forecasts are granted unfair advantage over its USDA counterparts.[36] Yet, the revealed value of the acreage information in the satellite-based forecast model suggests that greater efforts should be devoted to collecting finer acreage information prior to harvest.

While it is the case that numerous satellites configured for accomplishing a vast array of missions have been launched, the only system capable of providing data in sufficiently high temporal frequency for precise crop forecasting is the NOAA AVHRR satellite/sensor-system. Our findings indicate useful information in the satellite-based crop forecasts. However, it does not mean that the more traditional crop forecasting system can be replaced. The virtue of the conventional forecasting system lies *inter alia* in ground-truthing for image interpretation and estimating crop acreage that is beyond the capability of the current state of remote sensing technology.[37] If the results discussed here can be generalized to other crops and regions, it would indeed cast doubt on the commercial promotions claiming satellite-based crop forecasts.

[36] We contacted NASS officials about the availability of their acreage estimates at the time of forecasting. Unfortunately this kind of information cannot be retrieved more than a few years back.

[37] Research within the remote sensing industry has been focusing on the possibility of estimating crop acreage from the satellite with finer resolution but much lower temporal frequency. However, it is doubtful that, given the status quo, this kind of information can be extracted from the satellite in a timely manner.

REFERENCES

Baur, R. F., Orazem, P. F. (1994): "The Rationality and Price effects of U.S. Department of Agriculture Forecasts of Oranges," *Journal of Finance*, 49: 681-695.

Beck, N., Katz, J. N. (1995): "What to do (and not to do) with Time-series Cross-Section Data," *The American Political Science Review*, 89: 634-647.

Belsley, D. A., Kuh, E., Welsch, R. E. (1980): *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Benedetti, R., Rossini, P., (1993): "On the Use of NDVI Profiles as a Tool for Agricultural Statistics: The Case Study of Wheat Yield Estimate and Forecast in Emilia Romagna," *Remote Sensing of Environment*, 45: 311-326.

Brown, D. P., Jennings, R. H. (1989): "On Technical Analysis," *Review of Financial Studies*, 2: 527-551.

Danthine, J-P. (1978): "Information, Futures Prices, and Stabilizing Speculation," *Journal of Economic Theory*, 17: 79-98.

Daraiswamy, P. C., Cook, P. W. (1995): "Spring Wheat Yield Assessment Using NOAA AVHRR Data," *Canadian Journal of Remote Sensing*, 21: 43-51.

Bradford, D. F., Kelejian, H. H. (1977): "The Value of Information for Crop Forecasting in a Market System: Some Theoretical Issues," *Review of Economic Studies*, 44: 519-531.

Fackler, P. L., Norwood, Bailey. (2000): "Forecasting Crop Yields with Condition Indices," Working Paper, North Carolina State University.

EDC. Readme. 1st, *The Conterminous U. S. AVHRR Biweekly Composites*. CD-ROM. National Mapping Division U. S. Geological Survey EROS Data Center, Sioux Falls, South Dakota

Fair, R. C., Shiller, R. J. (1989): "The Informational Content of Ex Ante Forecasts," *Review of Economics and Statistics*, 71: 325-331.

____________________ (1990): "Comparing Information in Forecasts from Econometric Models," *American Economic Review*, 80: 375-389.

Fama, E. F. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25: 383-417.

_________ (1991): "Efficient Capital Markets: II," *Journal of Finance*, 46: 1575-1616.

Fortenbery, T. R., Sumner, D. A. (1993): "The Effect of USDA Reports in Futures and Options Markets," *The Journal of Futures Markets*, 13: 157-174.

Garcia, P., Irwin, S. H., Leuthold, R. M., Yang, L. (1997): "The Value of Public Information in Commodity Futures Markets," *Journal of Economic Behavior and Organization*, 32: 559-570.

Greene, W. H. (2000): *Econometric Analysis*, 4th ed. Prentice Hall.

Grossman, S. J. (1977): "The Existence of Futures Markets, Noisy Expectations and Informational Externalities," *Review of Economic Studies*, 44: 431-449.

Grossman, S. J., Stiglitz, J. E. (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70: 393-408.

Groten, S. M. E. (1993): "NDVI-crop monitoring and early yield assessment of Burkina Faso," *International Journal of Remote Sensing*, 14: 1495-1515.

Hayami, Y., Peterson, W. (1972): "Social Returns to Public Information Services: Statistical Reporting of U.S. Farm Commodities," *American Economic Review*, 62: 119-130.

Henry, M. P. (1999): "Wheat Yield Estimates Using Multi-temporal AVHRR-NDVI Satellite Imagery," A Master Thesis in the Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT.

Hicks, J. R. (1946): *Value and Capital*, 2nd ed,: esp. 136-139, Oxford: Clarendon Press.

Hirshleifer, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61: 561-574.

Just, R. E. (1983): "The Impact of Less Data on the Agricultural Economy and Society," *American Journal of Agricultural Economics*, 65, 872-881.

Kakwani, N. (1967): "The Unbiasedness of Zellner's Seemingly Unrelated Regression Equation Estimators." *Journal of the American Statistical Association*, 62: 141-142.

Kandel, S., Ofer, A. R., Sarig, O. (1993): "Learning from Trading," *Review of Financial Studies*, 6: 507-526.

Keynes, J. M. (1930): *Treatise on Money*, vol 11: *The Applied Theory of Money*, Harcourt, New York.

Kmenta, J. (1986): *Elements of Econometrics*. New York: Macmillan.

Knipling, E. B. (1970): “Physical and Physiological Basis for the Reflectance of Visible and Near-Infrared Radiation from Vegetation,” *Remote Sensing of Environment*, 1: 155-159.

Lillesand, T. M., Kiefer, R. W. (1994): *Remote Sensing and Image Interpretation*, 3rd ed. John Wiley and Sons, Inc., New York.

Lucas, R. E. (1972): “Expectations and the Neutrality of Money,” *Journal of Economic Theory*, 4: 103-124.

Macauley, M. K., Toman, M. A. (1991): “Providing Earth Observation Data from Space: Economics and Institutions,” *American Economic Review*, 81: 38-41.

Malingreau, J. P. (1989): “The Vegetation Index and the Study of Vegetation Dynamics,” In F. Toselli (ed.), *Application of Remote Sensing to Agrometeorology*, 285-303. ECSC, EEC, EAEC, Brussels and Luxembourg.

MAPS (Montana Agricultural Potential System). (1990). MAPS Mailbox, A Land and Climate Information System. Montana State University, Bozeman, MT.

McNew, K. P., Espinosa, J. A. (1994): “The Informational Content of USDA Crop Reports: Impact on Uncertainty and Expectations in Grain Futures Markets,” *The Journal of Futures Markets*, 14: 475-492.

Montana Agricultural Statistics (1989-2000), Montana Agricultural Statistics Services, Helena, MT.

Parks, R. W. (1967): “Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated,” *Journal of the American Statistical Association*, 62: 500-509.

Quarmby, N. A., Milnes, M., Hindle, T. L., Silleos, N. (1993): “The Use of Multi-temporal NDVI Measurements from AVHRR Data for Crop Yield Estimation and Prediction,” *International Journal of Remote Sensing*, 14: 199-210.

Sellers, P. J. (1986): “Canopy Reflectance, Photosynthesis and Transpiration,” *International Journal of Remote Sensing*, 6: 1335.

Sumner, D. A., Mueller, R. A. E. (1989): " Are Harvest Forecasts news? USDA Announcement and Futures market reactions," *American Journal of Agricultural Economics*, 71: 1-8.

Tucker, C. J. (1979): "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation," *Remote Sensing of Environment*, 8: 127-150.

Tucker, C. J., Holben, B. N., Elgin, J. H., McMurtrey, J. E., (1981): "Remote Sensing of Total Dry Matter Accumulation in Winter Wheat," *Remote Sensing of Environment*, 11: 171-189.

Tucker, C. J., Sellers, P. J. (1986): "Satellite Remote Sensing of Primary Production," *International Journal of Remote Sensing*, 7: 11.

U.S. Department of Agriculture, Office of Budget and Program Analysis (2001): *FY 2002 Budget Summary*. Washington DC.

Wang, J. (1994): "A Model of Intertemporal Asset Prices Under Asymmetric Information," *Review of Economic Studies*, 60: 249-282.

Wiegand, C. L., Richardson, A. J., Escobar, D. E., Gerbermann, A. H. (1991): "Vegetation Indices in Crop Assessments," *Remote Sensing of Environment*, 35: 105-119.